THE USE OF SHORT-INTERVAL GPS DATA IN CONSTRUCTION OPERATIONS ANALYSIS

John C. Hildreth

Dissertation submitted to the faculty of
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Civil Engineering

Michael Vorster, Chair
Jesus de la Garza
Julio Martinez
Anthony Songer
Mehdi Ahmadian
Randy Dymond

January 29, 2003

Blacksburg, Virginia

Keywords: Data Collection, Construction Operations, Data Reduction, Field Data, Correlation, Field Verification

THE USE OF SHORT-INTERVAL GPS DATA IN CONSTRUCTION OPERATIONS ANALYSIS

John C. Hildreth

Abstract

The global positioning system (GPS) makes use of extremely accurate measures of the time to determine position. The times required for electronic signals to travel at the speed of light from at least four orbiting satellites to a receiver on earth is measured precisely and used to calculate the distances from the satellites to the receiver. The calculated distances are used to determine the position of the receiver through triangulation.

This research takes an approach opposite the original GPS research, focusing on the use of position to determine the time at which events occur. Specifically, this work addresses the question: *Can the information pertaining to position and speed contained in a GPS record be used to autonomously identify the times at which critical events occur within a production cycle?*

The research question was answered by determining the hardware needs for collecting the desired data in a useable format and developing a unique data collection tool to meet those needs. The tool was field evaluated and the data collected was used to determine the software needs for automated reduction of the data to the times at which key events occurred. The software tools were developed in the form of Time Identification Modules (TIMs). The TIMs were used to reduce data collected from a load and haul earthmoving operation to duration measures for the load, haul, dump, and return activities.

The value of the developed system was demonstrated by investigating correlations between performance times in construction operations and by using field data to verify the results obtained from productivity estimating tools. Use of the system was shown to improve knowledge and provide additional insight into operations analysis studies.

Acknowledgment

I would like to express my gratitude to the many individuals and organizations that have contributed to this work. I sincerely appreciate Dr. Mike Vorster for his vast amounts of guidance and encouragement throughout this journey. He contributed considerable time and effort to keeping this work focused and on going. A great man on and off campus, he is truly a champion of both academia and construction.

A sincere thanks goes to Dr. Julio Martinez for guidance through the process, and particularly for his assistance with chapter four of this dissertation. His insights have proved invaluable throughout the duration of this journey. Also, sincere thanks to Drs. de la Garza, Songer, Dymond, and Ahmadian for their contributions to shaping this research.

I would like to express my appreciation for the companies that assisted with this work. A special thanks goes to Branch Highways and Acco Stone Quarry (Salem Stone Company) for providing the laboratory for this research. Without access to their sites and equipment, this work would not have been possible. And a special thanks goes to Terradon Communications for their assistance with the programming of the software.

Perhaps the greatest resources a graduate student can have are friends and colleagues. I would like to mention a special few by name, but I truly appreciate all of the folks who have passed through the 321 Patton complex, and the support and assistance they have provided. Those special few include, but are not limited to, John Draper, Greg Wagner, Kirsten Davis, Ponch Frank, Josh Lawrence, Roxene Thompson, Tim Kastens, Walid Matta, and Saral Shodhan.

Finally, my deepest gratitude and most sincere appreciation is for my family. To my parents for their unconditional love and support over the many years, without them none of this would have been possible. And to my wife, Kari, for her amazing love and support through the most difficult portions of this journey. I owe everything to these three very special people.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. THE USE OF SHORT-INTERVAL GPS DATA IN CONSTRUCTION OPERATIONS ANALYSIS

The global positioning system (GPS) makes use of extremely accurate measures of the time to determine position. The times required for electronic signals to travel at the speed of light from at least four orbiting satellites to a receiver on earth is measured precisely and used to calculate the distances from the satellites to the receiver. The calculated distances are used to determine the position of the receiver through triangulation.

One can envision that the question addressed by the original GPS researchers was: *Can extremely accurate measurements of time be used to determine position?* The global positioning system that resulted has revolutionized surveying and mapping techniques. GPS makes date, time, position, and speed information available worldwide.

This research takes an approach opposite the original GPS research, focusing on the use of position to determine the time at which events occur. Specifically, this work addresses the question: *Can the information pertaining to position and speed contained in a GPS record be used to autonomously identify the times at which critical events occur within a production cycle?*

The construction industry has traditionally relied on an observer in the field using a stopwatch to record these times. The observer uses vast amounts of visual information regarding position and speed to determine when a critical event has occurred. For example, the observer would draw an imaginary line between haul road segments and record the time at which, in their mind, the truck crosses the line. Similarly, speed may be used to indicate an event has occurred, when the observer records the time at which the equipment either comes to a stop, speed equals zero, or resumes travel, speed greater than zero.

Video recording of production operations removed the stopwatch, but replaced it with cameras, tapes, and sophisticated computer programs to record time. Rather than observe the operation in the field, the analyst observes a recording of the operation in an office and uses keystrokes to record the time at which critical events occur. However, the observer continues to rely on visual information regarding position and speed and follows the same decision making process used when recording with a stopwatch.

Whether an observer is in the field or in the office, manual time studies are subject to significant limitations. The observer is required to make instantaneous decisions to identify the critical times. Decisions made by the same observer may vary over time, or different observers may not base decisions on the same criterion, producing results that are not repeatable. Also, the information available is strictly limited to the field of view of the observer, or the camera. A single observer cannot simultaneously study portions of the subject operation performed out of sight. Therefore, operations performed over a large area require multiple analysts, or multiple days of analysis.

GPS is a technology that exhibits potential for use in an autonomous data collection system for analysis of construction operations, which relies on knowledge of when a particular resource is at a location of interest. GPS data includes knowledge of when (as UTM time or Greenwich Mean Time) and knowledge of where (as latitude and longitude). The accuracy of non-differential (static) horizontal positions determined by GPS receivers has increased from approximately 100 meters to approximately 10 meters with the removal of an intentional signal degradation known as selective availability (SA) [Angelo 2000]. In addition to providing current time and horizontal position, modern tracking grade GPS receivers are capable of reporting current date, speed, and direction of travel.

The research performed for this dissertation results in a new process for identifying the critical times in a production cycle by using GPS data recorded at a fixed, short time interval. The use of GPS data enables the analyst to autonomously and continuously collect position, time, and speed data. Software and appropriate algorithms can reduce the data to determine critical times of interest. Decisions made by computer programs based on mathematical algorithms can be exactly repeated. The GPS “field of view” is also much greater than that of an observer or camera; GPS data can be gathered at essentially any point on the earth with an open view of the sky.

This chapter introduces the dissertation by providing background information pertaining to field data collection. The methods of collecting field data and the evolution of methods from manual to automated means will be discussed. GPS and its application to field data collection is also described. The goal of the research is stated and the objectives completed to obtain the goal are enumerated. The scope of the research is described and an outline of the dissertation provided.

1.1. Background

Studies of construction operations are undertaken with the intent of improving performance and performed in four steps: *record, analyze, devise*, and *implement* [Oglesby et al. 1989]. The four steps are described as:

1. *Record* - data is collected for analysis, possibly from many different sources using many different means
2. *Analyze* – based on data collected, the operation is modeled and examined in detail
3. *Devise* - strategies and recommendations are formulated based on results of the analyses
4. *Implement* - devised strategies and recommendations are implemented through the sharing of information and a common interest in the results of the process

These steps are performed to develop a better understanding of the operation, such that improvements can be realized. The extent to which a better understanding can be realized is directly influenced by the ability to collect data and analyze the operation.

Recent advances in sensing and computing technologies have enhanced the methods used in the data collection and operations analysis phases. Modern automated data collection techniques rely on sensors and instruments placed on the equipment to monitor and record data, effectively rending the traditional manual methods obsolete. This has made it possible for techniques used in operations analysis to shift from paper-based deterministic methods to computer-based modeling and simulation methods.

The relationship between construction operations, data collection, and simulation is shown in Figure 1.1, which also shows how the four steps of the productivity improvement process (record, analyze, devise, and implement) occur at the interfaces of two or more of the three areas. For example, the recording phase is the overlap of the data collection and construction operations.

Data, including measures of production, cost, and the time required to complete critical tasks, are collected from construction operations to provide a better understanding of the operation and to facilitate analysis of the operation. Records of production and cost are often maintained, but the time required to complete individual tasks is typically obtained through the use of observational techniques. To adequately capture the variation inherent in construction operations, a large statistically valid data set is required.

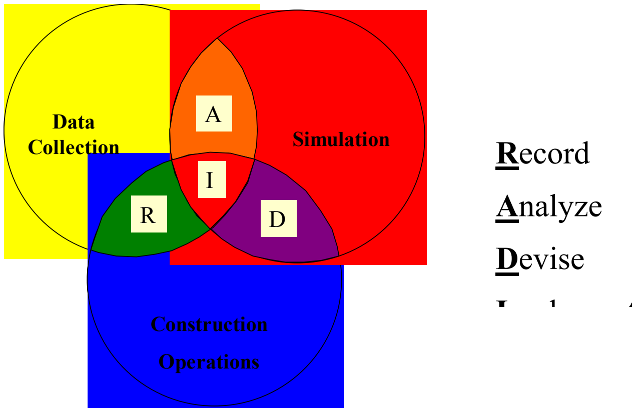


Figure 1.1: Research Realm

1.2. Field Data Collection

Field data is collected from construction operations as part of time studies performed to record the time required to complete various tasks [Oglesby et al. 1989]. The systems available to analysts have historically been manual in nature and limited to the field of view of an observer. Analysts have recently turned to technology for automated means for collecting field data. Instrumentation placed on-board construction equipment is replacing traditional manual methods, such as stopwatch studies, photographic, and video recordings.

1.2.1. Manual Methods

The stopwatch is the simplest tool available for time studies. While the stopwatch is simple and requires minimal capital investment, it requires a maximum of human involvement in the study. This tradeoff between simplicity and labor intensity makes elaborate stopwatch studies infeasible, but studies of limited scope can be quite valuable [Oglesby et al. 1989].

Oglesby et al. [1989] note that there are significant limitations to stopwatch studies. The observer is tasked with instantaneously deciding when an activity stops and the next starts. A single observer has difficulty watching several components of an operation simultaneously. The data gathered is restricted to that available within the field of view of the observer, and the information extracted is limited to the detail with which the data was recorded.

These limitations are not unique to stopwatch studies; in fact all of the described manual methods of field data collection are subject to the same limitations. Regardless of whether on observer is in the field with a stopwatch, or is reviewing a recording of the operation in an office, the data recorded is limited to the field of view and the information available is limited to the detail of the data recorded.

Time-lapse photography eventually replaced the stopwatch as the tool of choice for field data collection. Parker and Oglesby [1972] enumerate the advantages over stopwatches. The system is less labor intensive, thereby offsetting somewhat the greater capital investment. Multiple components of an operation can be observed simultaneously, with the added benefit of recording the interrelationships among the components.

Sprinkle [1972] noted that a single camera could be used in place of several observers with stopwatches, and that the record created is far superior to tabulated data sheets. The photographs form a comprehensive record that can be repeatedly reviewed as necessary, and without question as to the accuracy of the visual data.

Research into the use of time-lapse photography to collect data for simulation was performed at Stanford University [Paulson 1978; Paulson et al. 1983; Paulson et al. 1987]. Initially investigated as an educational tool, the research evolved into a search for an economical means of data collection, statistical analysis, and graphical simulation of operations. It was concluded that time-lapse photography with video data acquisition is an economically feasible means of data collection for simulation.

Videotape replaced film as the media of choice by analysts in that it reduced cost and provided the advantage of instant replay [Oglesby et al. 1989]. However, the tradeoff brought with the media was that the time compression required to reduce viewing time could no longer be easily achieved.

Bjornsson and Sagert [1994] made significant advances with PAVIC+ in the means for extracting data from videotape. PAVIC+ places a timestamp on the audio track of the tape and uses the stamp to reference points of interest on the tape. The analyst views the tape through a computer and marks the points of interest through keystrokes. The referenced times are summarized and can be exported for further analysis. Frank [2001] presented DVAT, a similar system for digital video that was developed based on PAVIC+.

Everett [1998] used time-lapse video recording methods to document entire construction projects. An excellent discussion of available equipment and its use are presented with several case studies of using time-lapse video to provide an as-built record of sorts. However, the extraction of information for analysis was not addressed, as the recording was used only as historical documentation of the project.

1.2.2. Automated Methods

Manual methods can provide the data necessary for simple deterministic job studies, but modeling and simulation techniques require large volumes of statistically valid data that are prohibitively expensive to gather with traditional means. Therefore, analysts have turned to new automated field data collection tools able to gather the large volumes of data necessary, and overcome the limitations of manual methods. Instrumentation of equipment with electronic sensors and the circuitry to record their output has been widely used to monitor the mechanical health of machines, and occasionally for field data collection [Kannan and Vorster 2000].

While the more widely known systems have been for monitoring machine health, other systems have been developed to improve machine productivity. Sotoodeh-Khoo and Paulson [1989] developed and used an on-board instrumentation (OBI) system to optimize the load time for a scraper. While deterministic methods were used to calculate activity times, the calculated load time was compared in real-time with the actual load time. Chronis [1987] describes the use of instrumentation for several earthmoving applications, including controlling scraper wheel slippage and transmission shifting, maximizing dozer production, preventing loader rollovers, providing increased traction to haul trucks, monitoring dragline production, and monitoring drilling operations. The mining industry has been a champion of OBI use, and has used systems to prevent the overloading of trucks [Chronis 1984] and to assign trucks to loading, dumping, fueling, and maintenance areas [Chronis 1985].

The Vital Information Management System (VIMS) produced by Caterpillar, Inc. is an established machine monitoring system that monitors both machine systems and productivity [Caterpillar 1999]. VIMS collects and records sensed, internal, communicated, and calculated data to provide faster access to real-time data and aid in service diagnostics. Schexnayder et al. [1999] used productivity data from VIMS to study the relation between truck payload and productivity. In addition to the conclusions drawn regarding the effect of payload on fleet productivity and truck economics, the authors conclude that “data in the computer is of no value; it must be extracted and presented in a clear format so [those] who understand the physical processes can discern the effects of their decisions”.

Kannan and Vorster [2000] attempted to do just that when they used VIMS data to investigate the relationship between payload and load time. A payload-load time (PLT) map was used to graphically depict the joint frequency distribution of load time and payload. The PLT map uses contours to indicate combinations of payload and load time with equal probability of occurring.

OBI systems as field data collection tools overcome the field of view limitation of manual methods by replacing the observer with sensors and circuitry. This also serves to make OBI systems autonomous data collection tools. The data collected by autonomous systems is not inherently biased, as is data manually recorded. Such systems are capable of recording the large volumes of data required for modeling and simulation, but bring with them issues associated with managing the data.

1.2.3. Global Positioning System

GPS is the universal name for the navigational system owned and operated by the U.S. Department of Defense that allows users to determine position, speed, and time using GPS receivers. Originally intended solely for military purposes, use of an intentionally degraded signal was opened for general public use in the 1980's. This intentional degradation was known as selective availability (SA) and was removed in 2000, increasing the accuracy of receivers from approximately 100 meters to approximately 10 meters [Angelo 2000]. This increased accuracy and the ability to incorporate the technology into OBI systems makes GPS a potential technology for field data collection.

GPS has been used as a field data collection tool in the agricultural, transportation, and construction industries. Borgelt et al. [1996] describe the use of GPS in an OBI system to map variations in crop yield. Data was logged at one-second intervals, including position from GPS, speed, and grain yield from other sensors. Uses of GPS by the transportation industry to gather field data are well documented [Goodchild and Fairhead 1993; Zito and Taylor 1994; D'Este et al. 1999]. The principal use of the technology is to provide position and speed data pertaining to study vehicles.

The construction industry has used GPS to provide data to navigational systems and to operations monitoring systems. Lothon and Akel [1996] describe a system to audibly alert operators of excavating equipment when in danger of striking a pipeline network. The current machine position was determined using GPS and compared to the known location of pipelines. The authors note the importance of a common coordinate system for comparison purposes. GPS was used to periodically correct for drift in an inertial navigation system used to guide an autonomously operated excavator [Crane et al. 1995]. IMPACT, an instantaneous motion planning and controlling tool, relies on GPS positional data to efficiently route construction equipment [Tserng and Russell 1997]. Compaction operations have been monitored by GPS based systems; Oloufa et al. [1997] and Pampagnin et al. [1998] describe similar systems for monitoring asphalt compaction.

GPS applications in earthmoving operations has mostly centered on providing positional data to systems designed to automate a portion of the work, and limited use for monitoring operations. Grade control systems, such as SiteVision by Trimble Navigation Limited and the Computer Aided Earthmoving System (CAES) by Caterpillar, Inc., rely on GPS technology for current and accurate position data.

The Morenci copper mine has reported improved performance and cost savings as results of instrumenting shovels and trucks with GPS [Phelps 1998; Shields et al. 2000]. Shovels use CAES to guide excavations, while trucks rely on GPS to track routes, destinations, and travel times. Shields et al. [2000] note the importance of system mobility, as the mine operates with fewer GPS receivers than trucks. Ackroyd [1998] describes using GPS to monitor construction earthwork operations. Data was collected from trucks and scrapers, but was not analyzed to identify potential operational improvements. Regardless, this work represents the first documented use of GPS to collect productivity data from construction earthmoving operations.

1.3. Research Goal and Objectives

The principal goal of this research is to determine whether the information pertaining to position and speed contained in GPS data can be used to autonomously identify the critical times within a production cycle. This goal will be achieved through the completion of the following objectives:

1. To design, build, and test a data collection tool to record a stream of GPS data at a short, fixed time interval
2. To show how the GPS data stream can be used to autonomously identify the times at which critical events occur in the production cycle by developing algorithms to autonomously identify the critical records
3. To demonstrate the value of the research by showing how it improves knowledge and provides additional insight into operations analysis studies.

1.4. Scope and Limitations

The scope of this work to answer the research question within the context of a load-haul-dump-return earthmoving operation. Such operations are relatively simple, making them easily understood and modeled. The large equipment used in earthmoving operations can be easily instrumented with a GPS data collection system. A simple GPS record containing time, position, and speed data will be collected. Non-differential GPS techniques will be used to record horizontal position with an accuracy of approximately 10 meters. The data will be recorded with a short, fixed time interval between two successive recordings. The short, fixed time interval used in this work will be two-seconds.

1.5. Outline

This dissertation is structured as an introductory chapter, four independent chapters, and a summary chapter. The introduction lays the framework for the research that is detailed in the following chapters. Chapters two through five are designed as independent manuscripts, each providing a literature review, methodology, results, and concluding remarks for their respective portions of the work.

- Chapter One – Describes methods for collecting field data from construction operations, and introduces GPS and its applications. The goal of the research is articulated and the objectives through which the goal will be attained are enumerated.

- Chapter Two – *Development of a Short-Interval GPS Based Data Collection System* – The development of the data collection tool to record a stream of short-interval GPS data is described.

- Chapter Three - *Reduction of Short-Interval GPS Data for Construction Operations Analysis* – The algorithms developed to autonomously identify the GPS records associated with critical events in the production cycle are presented and described. A case study is provided to demonstrate application of the system for construction operations analysis.

- Chapter Four – *A Methodology for Assessing Correlation in Construction Operations* – Demonstrates the value of the new system by using the data to provide insight into data correlation.
- Chapter Five – *The Use of Short-Interval GPS Data as a Model Validation Tool* – Data from the new system is used to provide insight into whether cycle time estimating tools provide reasonable estimates of travel time for roads of varying grades and lengths.
- Chapter Six – Concluding Summary and Recommendations – Provides a comprehensive review of whether the objectives have been achieved and clear statements regarding future research.

CHAPTER 2. DEVELOPMENT OF A SHORT-INTERVAL GPS BASED DATA COLLECTION SYSTEM

2.1. Introduction

Studies of construction operations undertaken with the intent of improving performance are performed in four steps: record, analyze, devise, and implement [Oglesby et al. 1989]. The four steps are described as:

1. *Record* - data is collected for analysis, possibly from many different sources using many different means
2. *Analyze* – based on data collected, the operation is modeled and examined in detail
3. *Devise* - strategies and recommendations are formulated based on results of the analyses
4. *Implement* - devised strategies and recommendations are implemented through the sharing of information and a common interest in the results of the process

These steps are performed to develop a better understanding of the operation, such that improvements can be realized. The extent to which a better understanding can be realized is directly influenced by the ability to collect data and analyze the operation. Recent advances in sensing and computing technologies have enhanced the methods used in the data collection and operations analysis phases. Modern data collection techniques rely on sensors and instruments placed on the equipment to monitor and record data, effectively rendering the traditional manual methods obsolete.

The global positioning system (GPS) is a technology that exhibits potential for use in an automated data collection system for analysis of construction operations. The accuracy of non-differential horizontal positions determined by GPS receivers has increased from approximately 100 meters to approximately 10 meters with the removal of an intentional signal degradation known as selective availability (SA) [Angelo 2000]. In addition to providing current horizontal position, modern tracking grade GPS receivers are capable of reporting current date, time, speed, direction of travel, and many GPS related parameters.

2.2. Literature Review

A review of available literature has been performed to establish the current body of knowledge pertaining to data collection for construction operations analysis and automated collection of GPS data.

2.2.1. Manual Methods

Data has historically been collected from construction operations through the use of manual techniques. Such techniques require a high degree of human involvement; the labor requirements making them costly. Common techniques used and reviewed as part of this work are stopwatch studies, time-lapse photography, and video recording.

2.2.1.1. Stopwatches

Time studies using stopwatches, or other such interval timers, are the simplest studies. While elaborate stopwatch studies are not feasible, studies of limited scope can be very valuable [Oglesby et al. 1989]. Such studies require the least capital investment, but require a high degree of human involvement. This human requirement is but one limitation from which stopwatch studies suffer.

Oglesby noted that such studies are severely limited because the exact point in time that one cycle or activity stops and another begins must be instantly chosen by the observer. Activities involving several components are difficult to accurately examine by a single observer and information available from the study is strictly limited to the information recorded by the observer and extracted from the study notes. The information recorded is also subject the physical limitations and biases of the observer [Oglesby et al. 1989].

2.2.1.2. Time-Lapse Photography

Time-lapse photography used in time studies provides all of the information that makes such studies beneficial. Parker and Oglesby [1972] enumerate the advantages of this method over stopwatch studies. They note that the process is relatively inexpensive, able to simultaneously record the activities of many components, able to record the interrelationships between the observed components, and able to provide an east to understand permanent record. Further benefits are that the record can be compressed to minimize reviewing time, contains information more detailed and dependable than stopwatch studies, and can be reviewed and studied by analysts, management, and others away from the work site [Oglesby et al. 1989].

Sprinkle [1972] investigated the use of time-lapse photography for collecting data from construction operations and found it to have definite advantages over conventional stopwatch techniques. He notes that one camera can replace several observers, the film as a record is superior to tabulated data sheets, and the film is a complete record that can be reviewed as often as is desired without doubt in regard to the accuracy of the visual data.

Time-lapse photography is not with out limitation. Similar to a stopwatch study, the available information is limited to that recorded [Oglesby et al. 1989]. In the case of photography, the field of view and any objects serving as obstructions limit the recorded information.

Much early research of time-lapse photography and its usefulness for simulation was performed at Stanford University and is described in [Paulson 1978], [Paulson et al. 1983], and [Paulson et al. 1987]. Initially investigated as an educational tool, the research evolved into seeking an economical means of field data collection, statistical data reduction and analysis, and graphical simulation of operations. The authors conclude that the use of time-lapse photographic and video data acquisition methods linked to a computer for data reduction and statistical analysis is an economically feasible manner in which to collect real data for simulation.

2.2.1.3. Video Recording

Oglesby et al. [1989] notes that video techniques are rapidly replacing photographic methods due to the advantages of instant replay along with less expensive and more reliable equipment. The authors also point out that video can be used to continuously record movements of workers and machines. The principle disadvantage of such real-time recording is that the period of time required to review the tape is equal to the recording time. It is noted that such an approach is rarely practical or cost-effective.

Everett [1998] focused on the use of time-lapse video for observing and documenting the construction of entire projects. The author provides an excellent discussion of available equipment and its use, along with several case studies of using time-lapse photography to provide an as-built record of actual construction operations. While the author details the methods of recording time-lapse video and many of its uses, he does not address extraction of information from the video.

The development of PAVIC+, the successor to Productivity Analysis with Video and Computer (PAVIC), at Chalmers University of Technology in Sweden made a significant advancement towards data extraction from video [Bjornsson and Sagert 1994]. PAVIC was developed in 1987 for use with traditional time-lapse photography and PAVIC+ updated the technology to integrate VHS video with the PC. The PAVIC+ system allows the user to view the video on the computer and to reference points on the tape through a standard time stamp placed on the audio track of the videotape. The referenced points in time can be exported to spreadsheet programs for further analysis.

Frank [2001] discusses the development and use of DVAT (Digital Video Analysis Tool) for analyzing digital video in a manner very similar to PAVIC+. DVAT allows the user to view the digital video and reference points on the video by using the time code created during recording. As with PAVIC+, DVAT can export the referenced times to spreadsheet programs for further analysis.

Senior and Swanberg-Mee [1997] discuss the methods of creating time-lapse video from real-time video. The hardware and software necessary to convert analog video to digital format at a specified time interval, and then back to analog, are described. The issues associated with time stamping the video are discussed. The authors evaluated the technology through field trial and presented the results. It is important to note that the authors acknowledge the practical hindrances associated with time-lapse techniques.

2.2.2. Automated Methods

While the manual methods of data collection presented are sufficient for deterministic job studies, they are not capable of supplying the large statistically valid data sets necessary for modeling and simulation techniques of analysis. Automated methods for data collection have relied on the use of vehicles instrumented with sensors.

Instrumented equipment have been used in earthmoving operations in the past. Systems have been used in attempts to improve performance and to monitor machine status. The largest and most prevalent current use of instrumented vehicles is for monitoring the mechanical health of equipment. The mining industry has been the leader in instrumentation use for monitoring production and improving performance.

2.2.2.1. Machine Monitoring

The Vital Information Management System (VIMS) is an example of a system designed for monitoring machine health. VIMS is an on-board system capable of monitoring machine systems, measuring and recording payload productivity, identifying abnormal machine conditions, and providing prognostic information [Caterpillar 1999]. VIMS data consists of four types of data: sensed, internal, communicated, and calculated. Sensed data is read from sensors and switches. Data generated by the VIMS module, such as date and time, is internal data. Data transferred to VIMS from other machine systems is communicated data. Data that VIMS arrives at mathematically is calculated data. Benefits of the system are noted to be faster service diagnostics and improved access to information in real time.

Caterpillar also offers other vehicle instrumentation systems such as Truck Payload Management System (TPMS), Computer Aided Earthmoving System (CAES), and MineStar [www.cat.com]. Examples of similar systems offered by other manufacturers are CONTRONICS by Volvo [www.volvoce.com] and INTELLIMINE and DISPATCHER by Modular Mining Systems [www.mmsi.com].

2.2.2.2. Performance Improvement

Sotoodeh-Khoo and Paulson [1989] were perhaps the first to use on-board instrumentation for collection of construction data. The primary goal of the work was to interface selected sensor-based data acquisition technologies to microcomputer-based software. This was done to allow for real time automatic collection, storage, retrieval, preprocessing, statistical analysis, and control decision-making. This was accomplished through a series of fully integrated knowledge based expert systems for data handling and decision-making logic. The system was used to optimize the load time of a scraper. Deterministic methods were used to calculate haul, dump, and return times, and a theoretical load time is calculated to balance the cycle. Actual production is compared to theoretical production to enable the system to modify the load time for the subsequent load activity.

The mining industry has been the leader of instrumentation use for performance improvement. Instrumentation has been used to prevent the overloading of haulage trucks at surface mines [Chironis 1984]. Modular Mining Systems, Inc. developed a system for the managing the assignment of haul trucks to shovels, overburden dump sites, refueling sites, and maintenance centers [Chironis 1985]. Chironis [1987] detailed the use of instrumentation for controlling scraper wheel slippage and transmission shifting, maximizing dozer production, preventing loader rollovers, providing increased traction to haul trucks, monitoring dragline production, and monitoring drilling operations.

Data from machine health monitoring systems has been used as an economical means for collecting large statistically valid sets of production data. Schexnayder et al. [1999] used data from VIMS to study the relation between truck payload and productivity. A significant volume of data, representing 54,300 truck cycles, was analyzed to determine the optimum truck payload. Despite the large volume of data, the variance in determined parameters, such as payload and productivity, was not emphasized. The relationship between payload and average productivity was found to not be linear and the marginal production was much greater when payload does not exceed the rated truck capacity. The relationship between payload and load time was investigated, and the results indicated that load time decreased as the payload range increased. A recommendation is made that management attempt to match the number of bucket loads to fill a truck to an integer number that satisfies both volumetric and gravimetric constraints. They also note that "data in the computer is of no value" and to be of use must be extracted and presented to discern the effects of decisions.

Kannan and Vorster [2000] were able to quantify the relationship between payload and load time in statistical terms using VIMS data collected from a fleet of large haulers. Production data was used to build an experience database to organize performance data on the basis of operating conditions and provides realistic data for planning and estimating. The work promotes and facilitates a shift in the source of data for productivity studies from generic and non-electronic standards and manuals to company- or project-specific, dynamic, and electronic archives of actual field experience. A convincing argument is made for the use of a continuous form of data collection to provide the probabilistic information necessary for simulation.

The authors used the data collected to investigate the joint frequency distribution of load time and payload. A payload-load time (PLT) map, provided as Figure 2.1, was used to graphically represent the joint frequency. The PLT map is a contour plot that uses color to indicate areas of equal probability. For example, the payload-load time pairs marked with red in Figure 2.1 occur with a probability of 1-2 percent. The PLT map is an excellent format for displaying the joint frequency distribution because the modal region and any relation between payload and load time are easily and quickly identified.

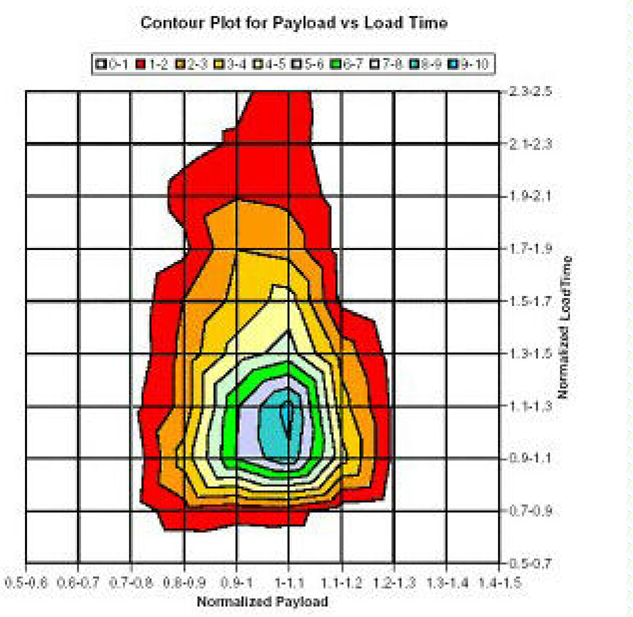


Figure 2.1: Payload – Load Time (PLT) Frequency Contour Map [Kannan 1999]

2.2.3. Global Positioning System

The global positioning system (GPS) is a sensor technology that is seeing rapid widespread use to provide positional data. Officially called Navigation Satellite Timing and Ranging (NAVSTAR), GPS is the universal name for the navigational system owned and operated by the U.S. Department of Defense and that allows users to determine position, speed, and time using GPS receivers. The system was originally intended for military use, but was opened to civilian use in the 1980's.

Garmin [2000a], a leading GPS manufacturer, describes the system in terms of its three parts: a space segment, a control segment, and a user segment. The space segment is comprised of 24 satellites, 21 active and three spares that orbit the earth at an altitude of approximately 12,000 miles. Each satellite transmits a low power radio signal to which GPS receivers "listen". There are five control stations making up the control segment. These stations track the orbit of the satellites and provide them with corrected orbital and clock information. The user segment consists of the user and GPS receiver.

The satellites transmit an ephemeris, which is its exact position in space and time, and a pseudo-range code, which the receiver uses to calculate the distance to the satellite. The receiver generates a signal identical to the pseudo-range code, and the delay between the signals provides a measurement of the distance to the satellite. The signal travels in a straight line between the satellite and the receiver and is obstructed by solid objects. Therefore, it is necessary to have a clear line of sight to four satellites in order to resolve a three-dimensional position, since time must also be resolved.

2.2.3.1. Agriculture and Transportation Industry GPS Uses

The agriculture and transportation industries have made widespread use of GPS in data collection systems. GPS has been used to provide positional data for mapping, as well as, travel speed data for transportation studies.

Borgelt et al. [1996] used an instrumented combine to map variations in crop yield in a soybean farm. Instrumentation consisted of a system to monitor grain yield and other parameters, a calibrated speed sensor to record the distance traveled, and a real-time differential GPS system to provide positioning data. The data acquisition system was a portable computer that logged data at one-second intervals. The authors note the importance of GPS accuracy relative to the scale of the operation, though they found the GPS accuracy to be adequate. They also found inaccuracies in the speed and distance traveled data provided by the GPS.

Goodchild and Fairhead [1993] describe the use of GPS in controlling and monitoring a public bus system. The bus navigation system used data from the odometer, GPS receiver, and a switch indicating that the bus doors had been opened. The accuracy of the GPS system was approximately 100 meters. Therefore, the odometer provided a more accurate measure of location, provided the bus was not diverted from the established route. The GPS positional data was used to determine if the bus had been diverted when the odometer data indicated a stop should have been reached but the door was not opened.

Zito and Taylor [1994] describe the use of GPS to collect data for travel-time surveys. Data from a GPS system providing date, time, and positional data was logged by a laptop computer, along with the speed and distance traveled data from a system monitoring the vehicle speedometer. The GPS data was logged at a three-second interval, while the speedometer data was logged at a one-second interval. The authors note that while the velocities reported by the GPS tended to be overestimated, the GPS system did accurately report the portion of time the vehicle was stopped. They acknowledge the limitations of GPS, but conclude that the amount of detailed data collected exceeds that of any previous data collection methods.

D’Este et al. [1999] discuss the use of a similar system to collect real-time data regarding traffic flow and performance. The authors note that while there have been many applications of GPS in transportation, but most do not make use of GPS data in both the spatial and temporal domains. The authors recognize the problems associated with communicating large volumes of data, and not that it is preferable to minimize the volume of data transmitted. The authors made use of modern GPS receiver technology, and note that in absolute positioning mode very accurate measurements of vehicle speed were provided.

2.2.3.2. Construction Industry GPS Uses

The use of GPS for data collection has not been limited to the agricultural and transportation industries, the construction industry has also seen the potential of GPS. The technology has been used to provide data for navigation systems and to provide positional data to operations monitoring systems.

Lothon and Akel [1996] describe a system to prevent accidental damage to buried utility pipelines. The potential system would compare the position of excavating equipment determined by GPS to a digitized map of the pipeline network. The operator would be alerted via an audible alarm in the event the equipment poses a risk to the pipeline. The authors note that the advantage is that it is only necessary to instrument the equipment with a GPS receiver to obtain positioning data. They also acknowledge the need to convert positional data to a common coordinate system for comparison.

GPS was used to periodically correct for drift in an inertial navigation system used to guide an autonomously operated excavator [Crane et al. 1995]. The instrumented excavator was navigated with an extremely high degree of accuracy, and deviations from the planned path were noted only at the corners of the path.

Tserng and Russell [1997] presented an instantaneous motion planning and controlling tool (IMPACT) for efficient routing construction equipment. Real-time kinematic GPS methods were used to provide equipment position data. A quadtree network was developed for each piece of equipment and potential paths were calculated for traveling machines. The paths were evaluated to identify potential collisions and determine the shortest collision-free path.

GPS has also been used on road compaction equipment, both to monitor equipment movements and provide navigation. Oloufa et al. [1997] describe an automated system to monitor and record the areas compacted by an asphalt roller. The authors used real-time differential GPS techniques to achieve a high degree of positional accuracy. A similar system described by Pampagnin et al. [1998] used GPS to provide positional data of asphalt rollers. Dead reckoning based on other instrumentation was also used for times at which GPS could not be used due to obstructions. The system was able to archive compaction progress and provide the operator with real-time information.

2.2.3.3. GPS Applications in Earthwork

GPS applications for earthwork operations have mostly centered on providing positioning information to systems designed to automate a portion of the work. It has seen limited use for monitoring earthwork operations, but no detailed analysis has been performed using data collected.

SiteVision™, produced by Trimble Navigation Limited [www.trimble.com], and the Computer Aided Earthmoving System (CAES), produced by Caterpillar, Inc. [www.cat.com], are examples of grade control systems for earthmoving equipment that rely on GPS for position data. In both systems, information pertaining to the desired finish grade is uploaded into the on-board system and compared to the existing equipment position to provide information to the operator. The Trimble system can also be integrated into the hydraulic system of the equipment to automatically control the blade.

The Morenci copper mine has reported performance improvement and cost savings as a result of instrumenting its equipment with GPS [Phelps 1998; Shields et al. 2000]. Electric shovels use the CAES to guide excavation operations, while haul trucks are equipped with GPS to track haul routes, destinations, and travel times. Truck position is updated at 30-second intervals and information regarding speed and congestion. By using the GPS equipped trucks to record the points of ore excavation and stockpiling, the mine has been able to increase the portion of ore recovered and decrease the time to recovery. Shields et al. [2000] note the importance of system mobility, as the mine operates with fewer GPS units than vehicles. The ability to move the system from one vehicle to another in a short time period eliminates the need for a substantial GPS hardware investment. They also point out difficulties associated with GPS use in an open pit mine, including multipath interference and a constantly decreasing satellite visibility due to pit deepening.

Ackroyd [1998] describes the use of GPS for monitoring earthwork operations. Haul trucks and scrapers were equipped with a GPS system and a cell phone to report and transmit position every two minutes. The GPS system was capable of accepting input from contact switches and load sensors. The author reports an ability to determine payload with a precision of five percent, although no data is presented to support this claim. Despite the long time interval between, it was concluded that timely information of earthworks operations could add significant value to the task of earthworks monitoring and scheduling.

2.3. Opportunity

Despite the wide range of applications both in and out of construction, GPS technology has not been used to collect statistically valid activity duration data for construction operations analysis. Such datasets are necessary when using modeling and simulation techniques to analyze construction operations. This need for data, combined with the available GPS technology, led to this research into the use of GPS for construction operations analysis. This research builds on the work by Kannan and Vorster [2000] in the area of autonomous data collection. GPS was the technology chosen for this work on the basis of it current applications in earthwork [Ackroyd 1998; Phelps 1998; Shields et al. 2000], as well as transportation and agricultural applications [Borgelt et al. 1996; D’Este et al. 1999; Goodchild and Fairhead 1993; Zito and Taylor 1994].

2.3.1. CrossCheck AMPS System Description

The CrossCheck AMPS vehicle tracking system manufactured by Trimble Navigation Limited was identified as a potentially suitable technology. Thus, it was chosen for the preliminary field evaluation; the objectives were to test the concept and develop specifications for the hardware, as well as system performance. The system is shown in Figures 2.2 and 2.3. The unit integrates GPS, cellular, and computing technologies onto a single board, which is enclosed in a low-profile housing. Trimble [2000] describes the system as ideal for asset management and security, as it can notify an operations center of both normal and unauthorized activity.

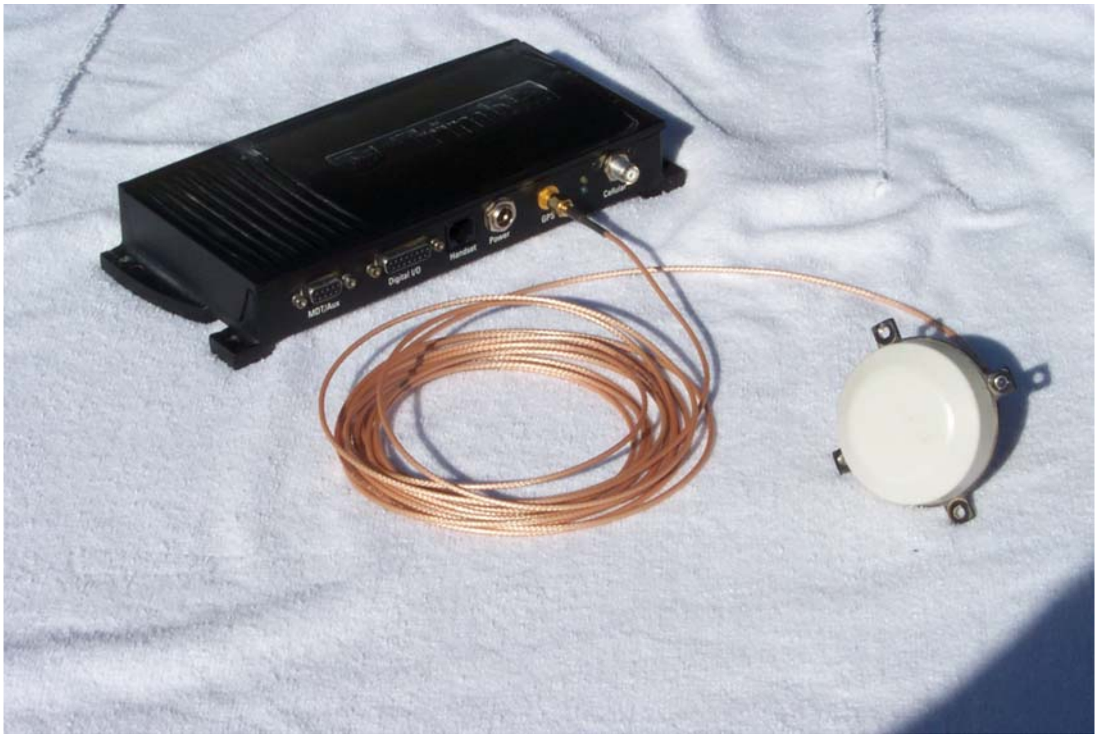


Figure 2.2: Trimble CrossCheck AMPS System

Figure 2.3: CrossCheck AMPS Connections

Intelligence is built into the system with the IQ Event Engine™ firmware, which is based on the concept of event reporting. Event reporting allows the unit to transmit reports when the vehicle status meets user-defined criteria. The IQ Event Engine™ provides the on-board intelligence to allow the unit to evaluate the vehicle status and report when criteria are met. Signals are binary indicators used to trigger events. A large number of signals are built into the system, including signals based on voltage supply, GPS fix, location, time, and distance traveled.

Event definitions set the criteria to be met and the resulting action taken when the event is triggered. The unit is typically set to provide a report as a result of an event. The user can choose the desired report from many pre-defined formats. The reports are in the form of an ASCII text data string, and are either sent to an operations center via cellular transmission or logged to the on-board memory. The on-board memory is 256 kilobytes of RAM and can store approximately 4,100 reports. Data logged to memory can be downloaded to a PC through a serial connection.

2.4. Preliminary Field Evaluation

Although the CrossCheck AMPS system was not designed to collect short-interval GPS data, it was chosen for the preliminary field evaluation. The ruggedness of the system and the ability to record short-interval GPS data were sufficient to test the concepts and evaluate the hardware requirements.

To evaluate the CrossCheck AMPS system, it was installed on five pieces of earthmoving equipment engaged in a load-and-haul earthmoving operation. Four Volvo A25C six-wheeled articulated haulers and a Caterpillar 350L excavator, a portion of a larger earthmoving equipment fleet, were instrumented and data was collected.

The fleet was used to move overburden soil and shot rock approximately 1,500 feet to either of two dump locations. Rock was stockpiled in one location to be crushed, while soil was stockpiled in an adjacent area. At both stockpiles, a dozer maintained the areas in a level condition. Hauling activities occurred over a relatively short distance, on well-maintained roadways, and without major obstacles.

As the project progressed, the fleet switched to a longer haul of approximately 9,000 feet in each direction to a roadway fill being constructed. The material being moved continued to be overburden soil and shot rock, but was no longer dumped in separate areas. Additional articulated haulers were gradually incorporated into the fleet to compensate for the long haul distance. Loading operations continued with the two hydraulic excavators, while a dozer maintained the dump area. The long haul route crossed three public roadways and other minor access ways. Each of the three roadway crossings was limited to one-way traffic. While the haul road surface was well maintained, the long haul distance and one-way segments presented hauling obstacles.

2.4.1. Data Collection

Data collected from the system included date, time, latitude, longitude, speed, and direction of travel. The time-distance signal was used to trigger an event when two seconds had elapsed, without regard to distance traveled, and resulted in the reporting of the short format of the event message, known as the EV message. This message was logged to the on-board memory for download to a laptop computer. The format of the message is:

AABBBBCDDDDDEEEFFFFFGGGGHHHHHIIIJJJKL

Where: AA is the event number
BBBB is the GPS week
C is an integer representing the day of the week
DDDDD is the current time Greenwich Mean Time (GMT) expressed in seconds since midnight
EEE.FFFFF represents degrees of latitude with north positive
GGGG.HHHHH represents degrees of longitude with east positive
III is vehicle speed in miles per hour
JJJ is the vehicle heading in degrees with zero at true north and increasing eastwardly
K is an integer representing the quality of the data recorded
L is an integer representing the age of the data

A sample of collected data is shown below.

```
>REV001065269090+3719328-0804034100000412;ID=0012;*5E<
>REV001065269092+3719328-0804034100000412;ID=0012;*5C<
>REV001065269094+3719328-0804034100000412;ID=0012;*5A<
>REV001065269096+3719328-0804034100000412;ID=0012;*58<
>REV001065269098+3719328-0804034100000412;ID=0012;*56<
```

The on-board memory can store approximately 4,100 records, which translates to an observation period of slightly less than 2 ¼ hours when data is logged at two-second intervals. Attempts were made to extend the observation period by altering the event criteria. During the shorter hauling operation, a speed threshold was added to prevent the logging of data when the truck was traveling at speeds greater than the threshold. This approach did not significantly increase the observation period, as the portion of time the trucks spent traveling at speeds above the threshold value was not significant.

A separate attempt was made to eliminate the logging of repetitive data during the longer hauling operation. Criteria were set to trigger an event when the truck stopped and every two-seconds when the truck was traveling. This eliminated the logging of data every two-seconds during the times at which the truck was stopped. This attempt also did not result in a significant increase in observation time, as the longer haul distance increased the portion of time the truck spent traveling relative to the time spent stopped.

Data could not be downloaded when the on-board memory was filled, approximately every 2 ¼ hours, as this would disrupt the operations. To avoid disrupting the operation, data was downloaded twice daily, while the trucks were parked for lunch and at the end of the workday. It was determined early in the evaluation that cellular transmission of the data was not feasible due to the slow cellular transmission rate of 1200 bauds. Therefore, data was downloaded to a laptop PC via a physical serial connection at a rate of 9600 bits per second. This reduced the time required to download the data to approximately five minutes per unit.

2.4.2. Preliminary Evaluation Results

The principal result of the preliminary field evaluation was that GPS data collected at short intervals does yield valuable information regarding the observed earthmoving operation. In addition to validating the research concept, the CrossCheck AMPS system was evaluated in terms of hardware capabilities and found to be somewhat limited. The noted system limitations were the basis for the developed hardware specifications.

The cyclic nature of operations performed by earthmoving haulers is evident when the collected data is presented graphically, as in Figure 2.4. The abscissa is represented by GPS time, which is commonly measured in seconds since midnight, and the ordinate is the linear distance to the load area. It is necessary to represent the load area as a single point to facilitate the distance calculations; the point selected was the approximate center of the load area. Data is plotted for those records where the hauler was stationary, which was indicated by a zero speed. The hauler cycle time and haul distance can be approximated from this plot. Interestingly, the stair stepping of the data when the hauler is near the load area can identify the queuing of haulers at the load area.

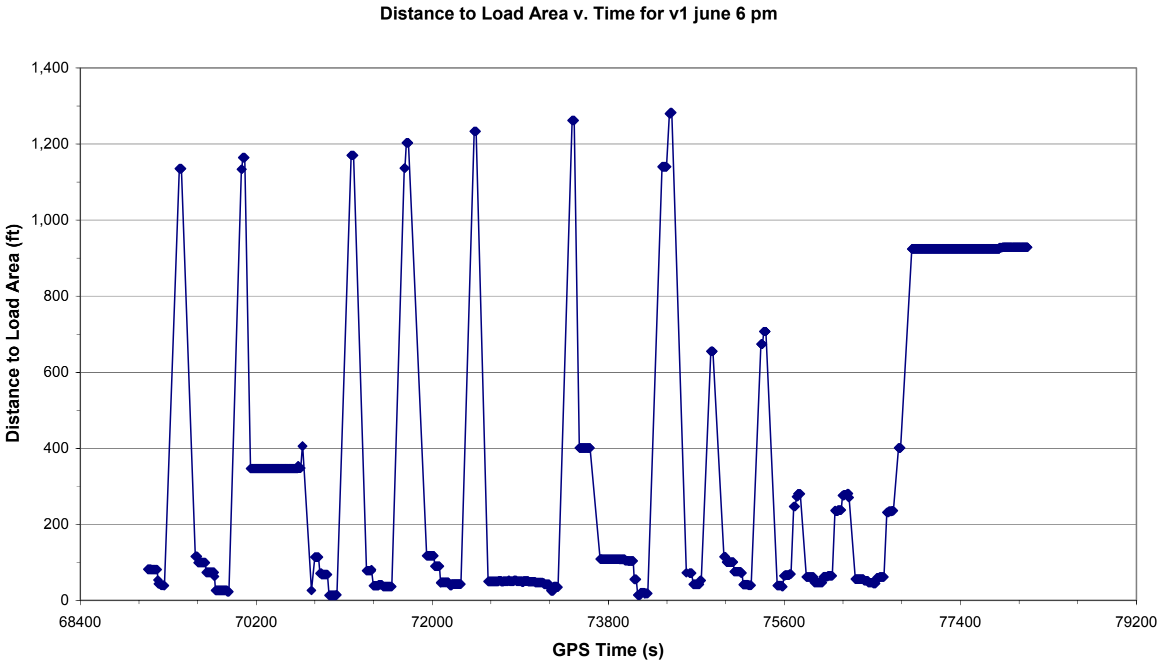


Figure 2.4: Hauler Distance to Load Area versus Time from Preliminary Evaluation

The field evaluation of the CrossCheck AMPS system indicated that the system was capable of collecting short-interval GPS data. However, deficiencies in memory, data transmission, and flexibility were noted in system.

The on-board memory was found to have insufficient capacity for logging GPS data at short-intervals. The lack of data storage capacity and infeasible cellular transmission made it necessary to download the data twice daily, which made the data collection process labor intensive. Even when downloaded twice daily, only approximately five hours of data was recorded. This corresponds to approximately half of the workday.

It was also found that the system was fixed in terms of the format of data and ability to accept other sensors as input. While a large number of output strings were available for use, the user is not given the ability to construct a unique output string, nor to alter any of the provided data strings. Also, the system was only capable of accepting inputs from digital devices, which precludes the use of more versatile analog devices.

While it is important to note that the CrossCheck AMPS system was capable of collecting the desired GPS data at short-intervals, it was subject to the above-described limitations. Therefore, hardware specifications were developed from the noted limitations. It was decided that a unique system would be built from these specifications with the purpose of collecting the desired data.

2.4.3. System Recommendations

The limitations noted during the preliminary field evaluation resulted in the development of recommendations for an autonomous short-interval GPS data collection system. The limitations were discussed in the previous section, and are summarized in Table 2.1 along with the resulting recommendation.

Table 2.1: CrossCheck AMPS Limitations and System Recommendations

System Attribute	CrossCheck AMPS	System Recommendation
Power Requirement	12 volts DC	12 - 24 volts DC
Data Format	ASCII text	ASCII text, comma delimited
Data String	Chosen from many pre-defined strings	User defined
Data Logging Interval	1 sec +, user defined, requires system re-programming	1 sec +, user defined at system power up
Data Storage Capacity	Approx. 4,100 records, stored in on-board RAM	Approx. 90,000 records, stored on removable Compact Flash media
Data Transmission	Cellular or serial connection	Data radio capable, removable storage media
Additional Sensors	Up to 4 digital	Up to 8 digital and 5 analog

2.5. System Description

The fixed configuration of the automated data collection system was designed and built by Mark Wright dba Wright Technical Systems, and was based on the above recommendations. The mobile configuration was further developed as part of this research. Both consist of a data acquisition and storage box, a sensor pack, and a wiring junction box. The system was developed to operate on a power range of 12 to 28 volts so that it may be powered by battery or draw power from construction equipment, which commonly operates on a 24 volt system.

2.6. Data Box

The data acquisition and storage box, shown in Figure 2.5, is a weather-tight enclosure that houses a micro-controller, a PC/104 embedded PC, and a CompactFlash media writer. The micro-controller gathers the sensor outputs, assembles the data string, and passes the string to the PC. The PC interfaces with the CompactFlash media writer to store the data on CompactFlash cards. Sensor outputs and power are supplied to the data box through a 25-pin connection on the side of the box. The box also contains power switch for the system.

Figure 2.5: Data Box

The micro-controller is a PIC16F876 device with 256 bytes of memory, 368 bytes of RAM, and a five channel analog-to-digital converter. The user can set the time interval at which data is logged by interfacing with the micro-controller through a text emulation program such as HyperTerminal or ProComm for Windows. A nine-pin serial port is provided on the outside of the data box for this interface. The user can set the logging interval to any integer number of seconds from one to 999. The flexibility in logging interval makes the system applicable to a wide range of applications.

The micro-controller also gathers the sensor outputs at the specified time interval and assembles the data string. The data string is then passed to the PC and echoed to the serial port for capture using text emulation programs. The firmware on the micro-controller can also be flash upgraded through the serial port.

The embedded PC is a 386SX 33MHz PC with two megabytes of DRAM and 900 kilobytes of Flash hard disk memory. MS-DOS version 6.22 is the operating system for the PC, which allows it to execute a small QBASIC program to accept the data string from the micro-controller and pass it to the media writer.

CompactFlash media is a solid-state media available in a wide range of sizes, from eight megabytes to 400 megabytes. The size of card chosen for use was 64 megabytes, as it will hold data from approximately 20 10-hour working days. The small size, approximately one-inch square, and large data volume make CompactFlash media the ideal storage medium for the system. The cards can be quickly and easily replaced, and readers are readily available to transmit the stored data to a PC. Such readers commonly use the USB interface for desktop computers, and the PCMCIA interface for laptop computers.

2.6.1. Sensor Pack

The sensor pack currently consists of a Garmin GPS-35 LVC receiver and a Vector 2X magnetic compass manufactured by Precision Navigation, Inc. The sensor pack is shown in Figure 2.6. While the GPS receiver is capable of reporting direction of travel, a separate magnetic compass was selected to eliminate the variations in the reported direction when the equipment is not in motion. The GPS receiver calculates direction of travel based on the determined position. Reported position can vary while the equipment is not moving due to GPS errors.

The system has been developed in a manner that allows additional analog and digital sensors to be incorporated. As many as five analog and eight digital sensors can be included. The data string has been configured to include these additional sensors, thus the sensors only need to be plugged into the junction box.

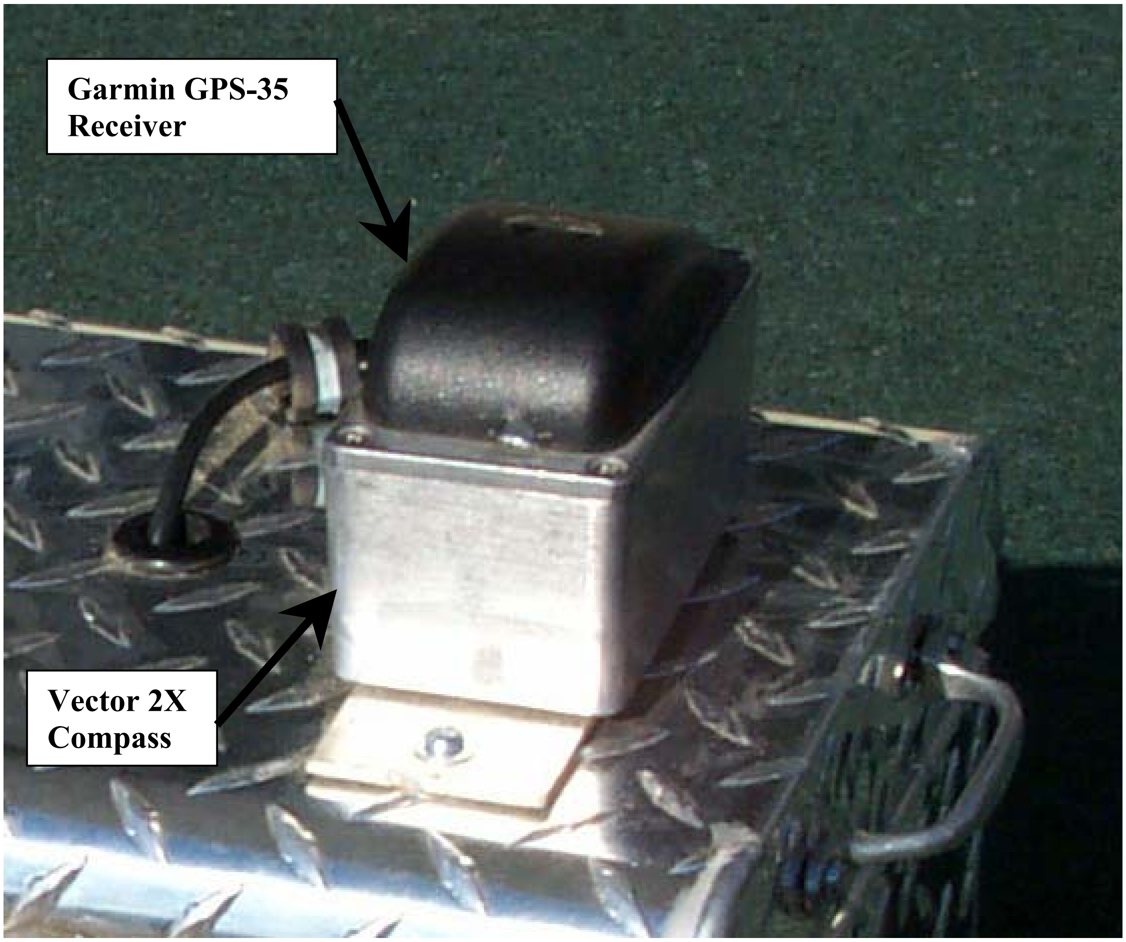


Figure 2.6: Sensor Pack on the Mobile Configuration

2.6.2. Junction Box

The wiring junction box provides a central location in which all electrical connections can be made. The box is shown in Figure 2.7. Power is brought into the box and distributed to the data box and sensor pack. Additionally, a five-volt power point is provided for analog sensors. Two modular connections are provided to quickly and easily connect the GPS receiver and a two-dimensional accelerometer, which will be added at a later time.

Figure 2.7: Wiring Junction Box

2.6.3. Mobile Configuration

The mobile system includes the previously described data box, junction box, and sensor pack, and also incorporates a 12-volt battery for power. All components are housed in an aluminum box for easy installation and removal. An aluminum box was chosen to minimize magnetic interference with the compass. Attachment points are provided on both sides of the box to allow installation using common cargo straps. The mobile system may be attached to the equipment at any location. However, the GPS receiver attached to the top of the box requires a clear view of the sky. The mobile configuration is shown in Figures 2.8 and 2.9.

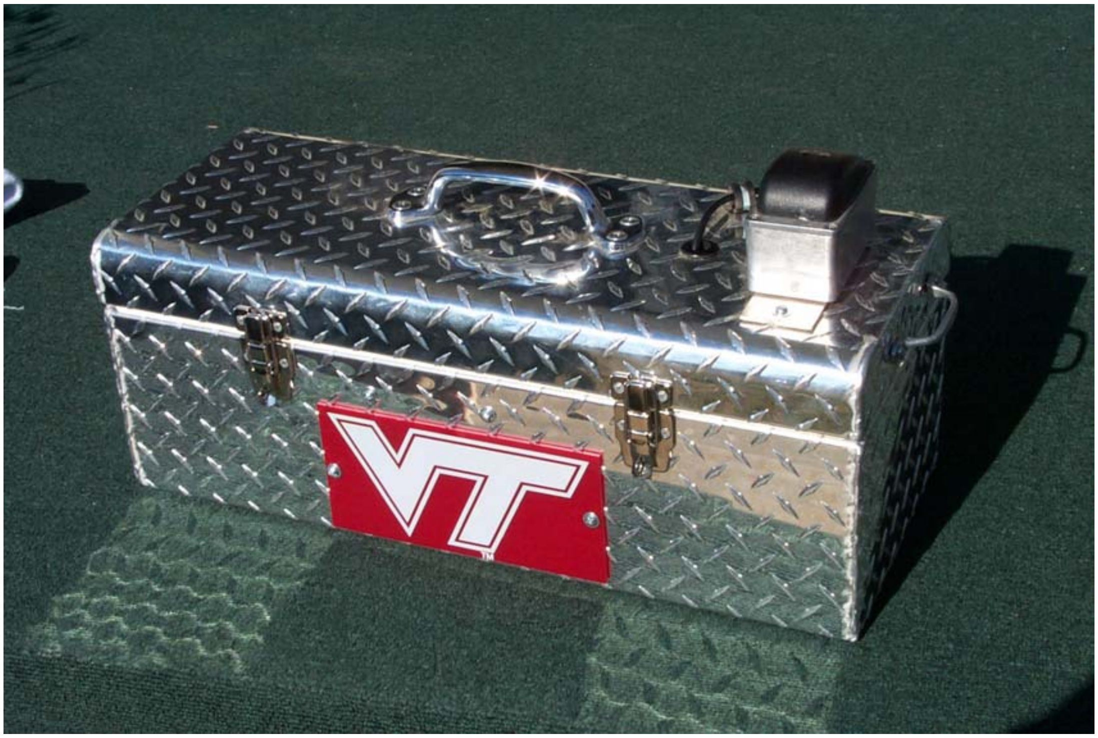


Figure 2.8: Mobile Configuration

Figure 2.9: Volvo A40D Articulated Hauler with Mobile System Attached

2.6.4. Fixed Configuration

The fixed configuration, like the mobile configuration, consists of the data box, junction box, and sensor pack, but without the system housed in an aluminum box. Rather, the data and junction boxes are mechanically attached in the cab of the equipment and power is drawn from the construction equipment. The GPS receiver requires a clear view of the sky, and therefore is mounted on the exterior of the equipment. The sensor pack has been attached to a piece of aluminum channel approximately two feet in length. Then the channel was mechanically attached to the construction equipment. The aluminum was chosen to minimize compass error resulting from magnetic interference. The data and junction boxes installed in the equipment cab are shown in Figure 2.10, and the attached sensor pack is shown in Figure 2.11.

Figure 2.10: Data and Junction Boxes Inside the Equipment Cab

Figure 2.11: Sensor Pack Located on Catwalk Handrail

2.6.5. Data String

The data string is recorded in comma delimited ASCII text format to minimize the memory usage and to facilitate use with a wide range of software programs. Data is written to a single file that is named on the basis of the unit identification number. The parameters that are recorded are box identification number, date, UTM time, latitude, longitude, speed, type of GPS fix, output from the five analog and eight digital sensors, and direction of travel from the compass. A sample data string is

```
06,190201,234250,A,3709.3967,N,08033.2691,W,012.4,0000,0001,0000,004,022,00000000,110
```

Where 06 is the unit identification number
190201 is the date – Feb. 19, 2001
234250 is the UTM time – 11:42:50 pm
A represents the GPS fix as three-dimensional
3709.3967,N is the latitude – 37° 9.3967" North
08033.2691,W is the longitude - 80° 33.2691" West
012.4 is the speed in kilometers per hour
0000,0001,0000,004,022 are the outputs from the five analog sensors
00000000 are the outputs from the eight digital sensors
110 is the direction of travel in degrees from north

While the use of ASCII text format minimizes the volume of memory needed, the volume of data collected over a 10-hour workday is very large. When collected at a two-second interval, a total of 18,000 lines of data are recorded. The resulting file size is approximately 3.3 megabytes in size. However, a file containing data from multiple days can quickly gain size and becomes increasingly difficult to manage.

Additional data is available from the GPS-35 receiver, but is not recorded in the interest of maintaining reasonable file sizes. A large number of GPS related parameters, such as dilution of precision values and total number of satellites in view, are available from the receiver. However, these parameters yield no information regarding the instrumented equipment.

2.7. Accuracy Test

The spatial position determined by any GPS receiver is subject to error from many sources. Therefore, the resolved position is not absolutely accurate. Error sources can be generally categorized as variations in the GPS, environmental conditions, and receiver noise. Error from GPS variations and environmental conditions is comprised of many components, several of which are time dependent. Therefore, the resultant error varies with time. An investigation was undertaken to estimate the non-differential horizontal static accuracy of the GPS units proposed for research use, and will be described in detail next. The published non-differential horizontal accuracy is 15 meters RMS [Garmin 2000b].

2.7.1. Methodology

Two GPS35-LVC receivers were placed on two different known points to simultaneously record observations. GPS control points established on the Virginia Tech (VT) Blacksburg campus were used. Latitude and longitude for each point, as well as state plane coordinates, were obtained from the firm responsible for establishing the control points in 1999.

Each observation period was one hour in duration. Horizontal position was recorded as latitude and longitude at one-second time intervals, producing 3,600 records for each one-hour observation period.

The effect of "observation time" on the accuracy was investigated. Observation time is the time period during which the unit receives position data from the satellites before determining its position. The reported position for longer observation times is typically an average of the positions determined during the observation time. The GPS35-LVC unit is not capable of averaging positions. Therefore, pseudo-observation periods were constructed. While the true observation period for all recorded positions was one second, pseudo-observation periods of 2, 5, 10, 30, 60, and 120 seconds were constructed by averaging the appropriate number of previous readings. For example, a 30-second observation period position is the average of the previous 30 one-second positions determined by the unit.

The horizontal position determined by GPS in latitude and longitude was projected to the Virginia South State Plane Coordinate System. The calculated coordinates were in units of meters. This allowed for quantification of error in meters as well. Error was defined as the arithmetic difference between the observed position and the actual position and was calculated for each determined position in terms of easting error and northing error.

Descriptive statistics, including mean, median, and standard deviation, were generated for all data sets so the effect of observation period could be seen. Additionally, statistical t-tests were performed to evaluate the difference in error between the two receivers.

2.7.2. Results

The effect of observation time on the mean and standard deviation of the error is shown in Table 2. 2. While the error populations were not normally distributed, they were nearly normally distributed. Therefore, the mean and standard deviation of the populations are meaningful measures.

Table 2.2: Effect of Observation Time on Mean and Standard Deviation of Error

Observation Period (sec)	Unit 3 Easting		Unit 4 Easting		Unit 3 Northing		Unit 4 Northing	
	Mean (m)	Std. Dev.	Mean (m)	Std. Dev.	Mean (m)	Std. Dev.	Mean (m)	Std. Dev.
1	8.1044	0.50971	8.8846	1.16825	2.0297	1.23580	2.6697	1.0206
2	8.1044	0.50891	8.8839	1.16636	2.0294	1.23513	2.6695	1.0195
5	8.1047	0.50697	8.8814	1.16035	2.0284	1.23327	2.6688	1.0169
10	8.1051	0.50281	8.8771	1.14888	2.0266	1.22832	2.6675	1.0123
30	8.1054	0.47583	8.8613	1.09483	2.0185	1.19600	2.6632	0.9883
60	8.1039	0.42432	8.8455	1.02086	2.0064	1.15125	2.6580	0.9454
120	8.0999	0.32046	8.8345	0.88750	1.9854	1.09222	2.6498	0.8693

The results indicate a significant difference between the errors in northing and easting. The average easting error for both units was greater than 8 meters, while the northing error for both units was less than 3 meters. Both northing and easting errors were positive, meaning that the determined position was north and east of the actual position.

There is also a difference indicated by the results between the accuracy of the two receivers used. The average easting error for unit 3 was approximately 8.1 meters, while the same error for unit 4 was approximately 8.8 meters. Similarly, the average northing errors for units 3 and 4 were approximately 2.0 and 2.6 meters, respectively.

Increasing the observation period from 1-second to 120-seconds does not significantly increase the accuracy of the determined position. It will, however, significantly decrease the standard deviation of the positional error. The reduction in error mean and standard deviation is shown in Table 2. 3 as a percentage of the 1-second error mean and standard deviation.

Table 2.3: Percent Reduction in Error Mean and Standard Deviation by Increasing the Observation Period from 1 to 120 seconds

Error	Reduction in	
	Mean	Std. Dev.
Unit 3 Easting	0.056%	37.129%
Unit 4 Easting	0.564%	24.032%
Unit 3 Northing	2.185%	11.618%
Unit 4 Northing	0.747%	14.822%

Paired t-tests at the 95 percent confidence level were performed to determine if there was a significant difference in error between the receivers. Easting and northing errors were tested separately. The paired t-test was chosen because it does not require knowledge of the distributions underlying the variables, but rather that the difference between the pairs be normally distributed [Analyse-It 2000]. The difference in both northing and easting coordinate pairs were tested for normality. The p-values of Shapiro-Wilk tests were less than 0.0001 for both northing differences and easting differences. These results indicate non-normal underlying distributions. However, the normal plots, shown in Figure 2.12, for both difference samples indicate that a majority of data is normally distributed, with deviations from normalcy at the tails, or extremes of the samples.

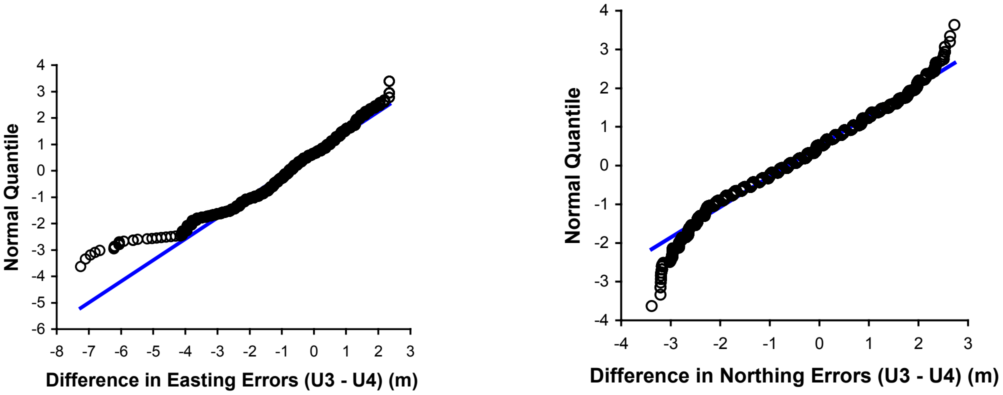


Figure 2.12: Normal Plots of Differences in Errors

The t-statistic resulting from the easting and northing tests was 37.59 and 30.25, respectively. Both of these values are significantly greater than the $t_{0.025}$ value of 1.96. Therefore, the null hypothesis that there is no difference in the populations is rejected. The results of the statistical tests lead to the conclusion that there are statistically significant differences in the accuracy of the two tested GPS receivers. However, it is also important to evaluate the practical significance of the difference between the receivers.

The receivers were found to be accurate in the easting direction to approximately eight meters, and to approximately 2 meters in the northing direction. The difference in receivers, in both directions, was found to be approximately 0.5 meters. This difference is of little practical significance when compared to the determined errors, and 0.5 meters is of no real practical significance when determining the position of large earthmoving equipment.

2.8. System Field Evaluation

The system was tested at a local quarry producing crushed aggregate. A fleet of three Caterpillar 771 quarry trucks and a 988 rubber-tired loader was equipped with the fixed configuration of the data collection system. The equipment was used to load and haul shot rock to feed crushers and overburden material to a waste area.

Management personnel and equipment operators were informed of the research in advance of the installation to gain their support and cooperation. It was explained that a history of position and speed was to be recorded and used to measure performance times. It was stressed that the purpose of the study was operations analysis and not operator evaluation. All personnel appeared supportive of the research and performed their daily duties with the same diligence that was observed prior to the study.

2.8.1. System Installation

The data and junction boxes were installed inside the cab of each piece of equipment. The data box was mechanically attached and the junction box was attached using self-adhesive Velcro strips in an area behind the operators seat, as seen in Figure 2.5. This was an unobtrusive location that provided access to the boxes while the operator’s seat was occupied. The system was wired to the equipment ignition switch to provide the necessary power only during the operating times of the equipment.

The sensor pack was mounted near the end of a 24-inch section of aluminum channel, which was attached to the equipment. The channel allowed the pack to be located away from the equipment, thereby minimizing magnetic interference with the compass and increasing the view of the sky from the GPS receiver. The channel was mechanically attached to the handrails along the equipment catwalks, as seen in Figure 2.6. All wiring to the junction box was protected with wire loom.

2.8.2. *Data Collection*

Data was collected for a one-week period before the data recorded on each system was transferred to a PC. The data string was recorded at a two-second interval in comma delimited text format, as previously described. The equipment operated approximately 11 hours each day, producing approximately 19,800 data records. Over the entire week, nearly 100,000 data records were logged by each system.

The system logged all of the data as a single text file, which was comma delimited intentionally to facilitate import into Microsoft Excel. The Excel format was chosen because it provides both data analysis and presentation capabilities. However, the approximately 100,000 line text files exceeded the maximum line capacity of Excel and had to be divided into individual days for use. The data can be separated manually using text editors such as Notepad or Wordpad, or shareware programs are also available. The data from a few days can easily be separated manually. A small sample of the collected data is shown below.

```
06,010301,113548,A,3710.7763,N,08023.5561,W,012.4,0001,0000,0000,008,022,00000000,139
06,010301,113550,A,3710.7697,N,08023.5585,W,012.4,0000,0000,0000,008,022,00000000,142
06,010301,113552,A,3710.7633,N,08023.5596,W,011.6,0001,0001,0000,012,022,00000000,149
06,010301,113554,A,3710.7573,N,08023.5584,W,010.7,0000,0000,0001,008,018,00000000,158
06,010301,113556,A,3710.7521,N,08023.5556,W,010.7,0001,0001,0000,008,022,00000000,165
06,010301,113558,A,3710.7472,N,08023.5512,W,010.9,0000,0000,0000,008,026,00000000,172
06,010301,113600,A,3710.7424,N,08023.5467,W,010.7,0000,0000,0001,008,030,00000000,170
```

The data files were imported into Excel for presentation. The data format was designed to allow easy import into Excel, with each data string occupying a single row in the spreadsheet. The pertinent data recorded is summarized in Table 2. 4.

Table 2.4: Summary of Data Collected

Unit	Date	UTM Time	Fix	Latitude		Longitude		Speed (km/h)	Heading
6	10301	11:35:48	A	37d 10.7763'	N	80d 23.5561'	W	12.4	139
6	10301	11:35:50	A	37d 10.7697'	N	80d 23.5585'	W	12.4	142
6	10301	11:35:52	A	37d 10.7633'	N	80d 23.5596'	W	11.6	149
6	10301	11:35:54	A	37d 10.7573'	N	80d 23.5584'	W	10.7	158
6	10301	11:35:56	A	37d 10.7521'	N	80d 23.5556'	W	10.7	165
6	10301	11:35:58	A	37d 10.7472'	N	80d 23.5512'	W	10.9	172
6	10301	11:36:00	A	37d 10.7424'	N	80d 23.5467'	W	10.7	170

Position is recorded as geodetic coordinates, along with speed and heading, and each record is stamped with a unit identification number, date, and time. The logging of both spatial and temporal data will allow for analysis of equipment position throughout the workday.

2.8.3. Field Evaluation Results

The field evaluation results indicate that the system performs as designed and is capable of collecting the desired data and surviving in the harsh earthmoving environment. A plot of hauler distance to the load area versus time, as previously described, can graphically represent the collected data. Such a plot is provided as Figure 2.13. In addition to the cyclic nature of earthmoving operations, it can be seen that the system recorded data for the entirety of the nearly 12-hour workday. Two distinct cycles are evident, one with an approximate 400 meter haul distance and another with an approximate 750 meter haul distance.

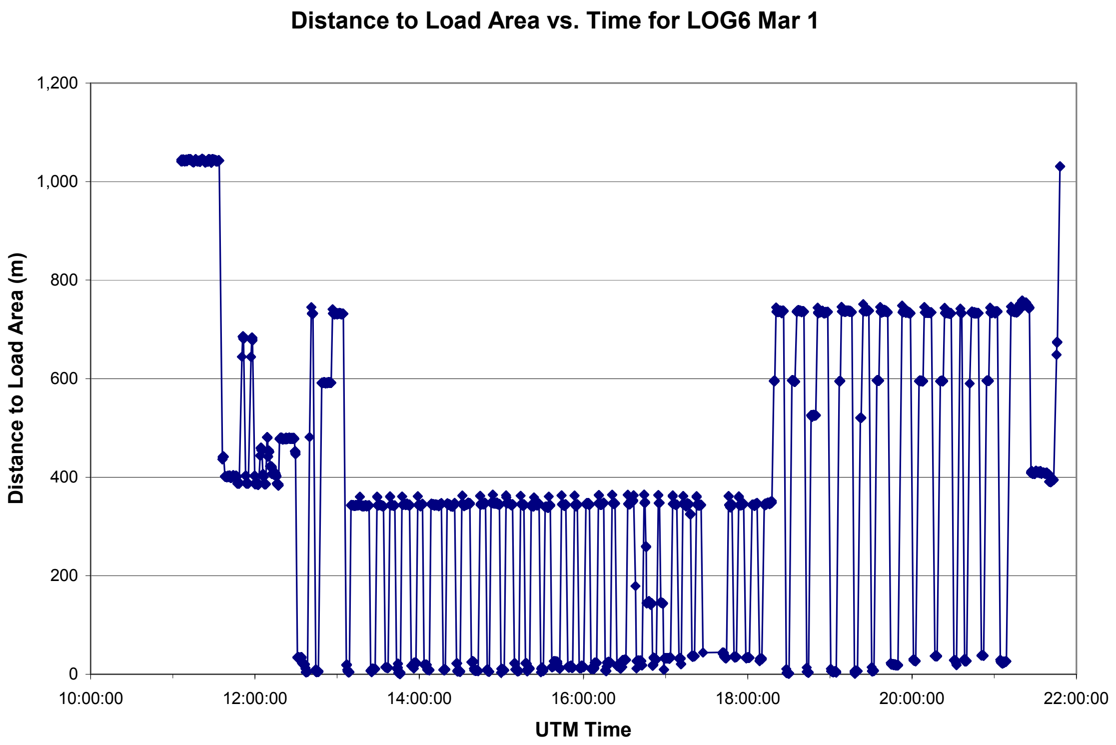


Figure 2.13: Hauler Distance to Load Area versus Time from Field Evaluation

A plot of the two-dimensional coordinates recorded was prepared to depict where the hauler was operating. This plan view of hauler operations is provided as Figure 2.14. The positions were recorded as latitude and longitude, and have been converted to state plane coordinates (Virginia south, NAD83) to provide the use of common linear units such as meters. Color is used to differentiate those positions recorded when speed was zero from those when speed was greater than zero. This allows the positions at which the hauler was stopped to be easily identified, along with the travel routes.

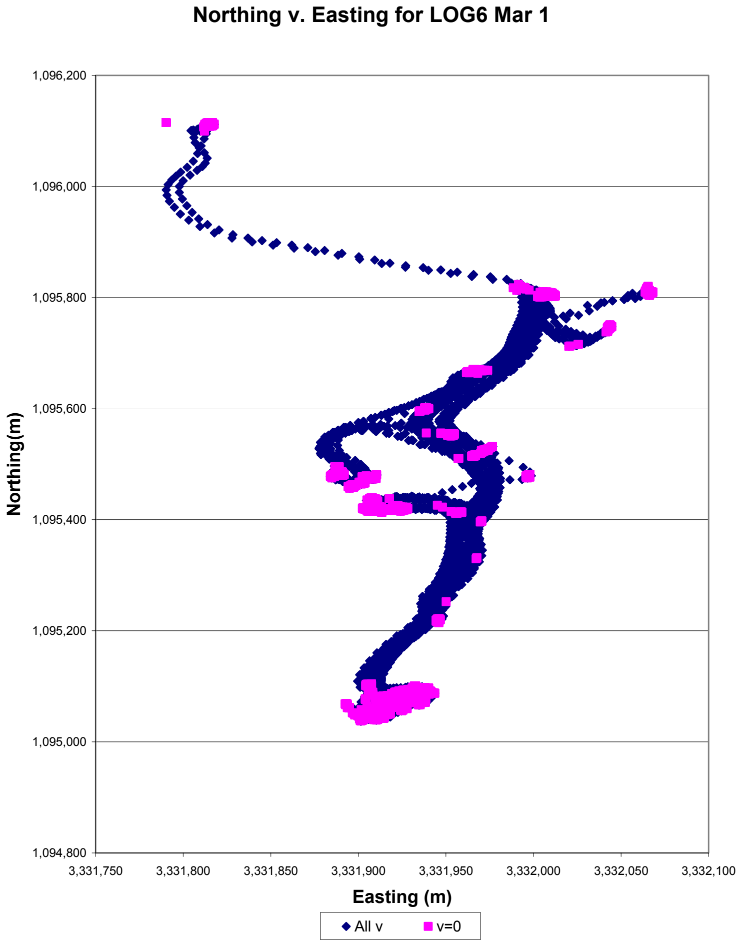


Figure 2.14: Plan View of Field Evaluation Data

Recommendations for improving system operation also resulted from the field evaluation. Areas of noted concern were the required installation time, very large data files, and the format and flexibility of the data string.

Considerable time and effort were required to mount the systems on the equipment. While the time and effort invested is appropriate for collecting data over long periods of time, it is not appropriate for data collection efforts lasting only a few days. Therefore, the mobile configuration fills a very large need in terms of ease of installation.

Storing data from multiple days as a single text file results in extremely large files. These large text files can be difficult to manage and separate by day. Recording the data from each day separately was desirable. The system firmware was updated to allow the user to log data based on the current date, thus a new data file is created for each day.

The format of the recorded time was hhmmss, without any separating characters. While the user can easily identify the time, Excel views this data as an integer. It is necessary to insert colons in the appropriate locations for Excel to view the data as a time. A simple macro was developed to loop through data in the time column and separate the string and insert the colons. However, this is much better accomplished when the data is initially recorded. Therefore, the data string was altered to include colons as data was recorded.

Any alteration to the data string requires an update to the system firmware. Initially the systems did not have flash-upgrade capabilities. To alter the data string, it was necessary for the manufacturer's representative to reprogram each box. This required either a visit by the representative to the quarry when the equipment was not operating, or removal of the systems and transport to the company. Therefore, the system firmware was updated to provide flash-upgrade capabilities.

The user can flash-upgrade the firmware through the serial connection in the field using a notebook computer. The small software program melabs Loader produced by microEngineering Labs, Inc. is used upgrade and verify the firmware. New firmware versions are produced by the manufacturer and provided to the user via email.

The recommendations developed and the resulting improvements to the system represent significant improvements in the system ease of use. The mobile configuration is easily installed on equipment for short-term data collection. Logging daily data files separately eliminates a slow and tedious manual step in the process. The ability to flash-upgrade the firmware provides the flexibility needed for a widely used system.

2.9. Conclusions

Data collection is a fundamental step in the construction operations analysis process. GPS is a technology exhibiting much potential for use in automated data collection system for construction operations analysis. The positional data provided potentially yields useful information regarding spatially oriented operations, such as earthmoving operations. Equipment engaged in such operations can be instrumented with GPS and data recorded at short intervals to potentially provide the information necessary for operations analysis. The uncorrected horizontal positional accuracy of the GPS-35 receiver was less than 10 meters, which should be sufficient to provide data for operations analysis. The system developed and field-tested is capable of recording such data at short intervals. The system has been observed to perform as designed and improved in terms of ease of use. Provided that the large volumes of data collected can be reduced to the information necessary, the system represents a new method of collecting data for construction operations analysis.

2.9.1. Future Research

Future research stemming from this work is the developing automated data reduction methods and evaluating the applicability of geographical information systems (GIS). The large volumes of data resulting from automated collection methods dictate the use of automated data reduction methods. It is envisioned that research efforts will be focused using Microsoft Excel as the data reduction engine. Excel can be programmed with relative ease using the Visual Basic for Applications language. Additionally, Excel is a common spreadsheet program with which many people are familiar and comfortable. Data reduction will involve the identification of data records logged at the start and stop of activities such that activity durations can be calculated. The ability of the system to record data throughout the entire workday and over several consecutive days will result in large samples of activity durations. From these samples, information regarding the operations can be extracted. Additionally, the samples can be used to develop stochastic duration definitions.

GIS is a tool for managing information based on spatial and attribute data, and thus may be applicable. GIS may be suitable for data management, presentation, and reduction. While not as common as spreadsheets, GIS software is used for a wide range of engineering applications. The ability to combine many different types of data makes it attractive as an information system.

2.10. References

Ackroyd, N. (1998), "Earthworks Scheduling and Monitoring Using GPS," in the *Proceedings of the 1998 Trimble User Conference*, pp. 1-6.

Analyse-It (2000), *Analyse-It Electronic Help*, Analyse-It for Microsoft Excel, v. 1.62, Analyse-It Software, Ltd.

Angelo, W. (2000), "Order to Unscramble GPS Signals to Boost Equipment Productivity" *Engineering News Record*, 244(19), p 32.

Bjornsson, H. C., and P. Sagert, (1994), "Multimedia based system for analyzing production organizations," *Computing in Civil Engineering*, 2, pp1405-1411.

Borgelt, S. C., J. Harrison, K. Sudduth, and S. Birrell (1996), "Evaluation of GPS for Applications in Precision Agriculture," *Applied Engineering in Agriculture*, v 12 n 6, pp 633-638.

Caterpillar (1999), *Systems Operation Testing and Analysis – Vital Information Management System (VIMS),* Caterpillar, Inc., Peoria, IL.

Chironis, N. P., (1984), "Detectors Estimate Truck Overloads," *Coal Age*, 89(8), p125.

Chironis, N. P., (1985), "Computer monitors and controls all truck-shovel operations," *Coal Age*, 90(3), pp50-55.

Chironis, N. P., (1987), "Microprocessor-aided equipment proves productive and reliable," *Coal Age*, 92(7), pp48-56.

Crane, C.D., D. Armstrong, and A. Rankin (1995), "Autonomous Navigation of Heavy Construction Equipment," *Microcomputing in Civil Engineering,* 10(5), pp. 357-70.

D'Este, G., R. Zito, and M. Taylor (1999), "Using GPS to Measure Traffic System Performance," *Computer-Aided Civil and Infrastructure Engineering,* v 14, pp255-265.

Everett, J. G., H. Halkali, T. G. Schlaff, (1998), "Time-lapse video applications for construction project management," *Journal of Construction Engineering and Management*, ASCE, 124(3), pp204-209.

Frank, S. (2001), "Field Operations Analysis Laboratory – A Tutorial on the Review and Analysis Aspects of Operations Analysis," Project and Report, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Garmin (2000a), "*GPS Guide for Beginners*" Garmin International, Inc., Olathe, Kansas.

Garmin (2000b), "*GPS 35LP Series Technical Specification*," Garmin International, Inc. Olathe, Kansas.

Goodchild, B and S. Fairhead (1993) "Using GPS to enhance the control and effectiveness of a public transport system," *Electronic Engineering*, v 65 n 796, pp 49-53.

Kannan, G. (1999), "A Methodology for the Development of a Production Experience Database for Earthmoving Operations using Automated Data Collection," PhD Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Kannan, G. and M.C. Vorster (2000), "Development of an Experience Database for Truck Loading Operations," *Journal of Construction Engineering and Management*, ASCE, 165(3), pp201-209.

Lothon, A. and S. Akel (1996), "Techniques for Preventing Accidental Damage to Pipelines," in the *Proceedings of the 1st International Pipeline Conference,* v2, American Society of Mechanical Engineers, New York, NY, pp. 643-650.

Oglesby, C. H., H. W. Parker, and G. A. Howell, (1989), *Productivity Improvement in Construction*, McGraw Hill Companies, New York, NY.

Oloufa, A.A., Do, W. and Thomas, W.R. (1997) "Automated Monitoring of Compaction using GPS," in the *Proceedings of the 1997 Construction Congress*, ASCE, Reston, VA., pp. 1004-1011.

Pampagnin, L., Peyret and Garcia G. (1998), "Architecture of a GPS-Based Guiding System for Road Compaction," in the *Proceedings of the 1998 IEEE International Conference on Robotics and Automation*, Leuven, Belgium, pp. 2242-2427.

Parker, H. W. and C. H. Oglesby, (1972), Methods Improvement for Construction Managers, McGraw-Hill Company, New York, NY.

Phelps, R.W. (1998), "GPS – Now for Excavation Morenci Mine Improves Ore Control," *Engineering & Mining Journal*, 199(7), pp. 34-6.

Paulson, Jr., B. C., (1978), "Interactive graphics for simulating construction operations," *Journal of Construction Division*, ASCE, 104(**CO**1), pp69-76.

Paulson, Jr., B. C., S. A. Douglas, A. Kalk, A. Touran, G. A. Victor, (1983), "Simulation and analysis of construction operations, *Journal of Technical Topics in Civil Engineering*, ASCE, 109(2), pp89-104.

Paulson, Jr., B. C., W. T. Chan and C. C. Khoo, (1987), "Construction operations simulation by microcomputer," *Journal of Construction Engineering and Management*, ASCE, 113(2), pp302-314.

Schexnayder, C., Weber, S. L., and Brooks, B. T., (1999), "Effect of truck payload weight on production," *Journal of Construction Engineering and Management*, ASCE, 125(1), pp1-7.

Senior, B. A. and A. Swanberg-Mee, (1997), "Activity Analysis using computer-processed time lapse video," in the *Proceedings of the 1997 Construction Congress*, ASCE, Reston, VA.

Shields, S., J. Flinn, and A. Obregon (2000), “GPS in the Pits: Differential GPS Applications at the Morenci Copper Mine,” *GPS World,* v 11, n 10, Oct 2000, pp 34-38.

Sotoodeh, H. and B. C. Paulson, (1989), “Sensors and expert systems in production optimization of earthmoving scrapers,” *in Proceedings of the Sixth Conference on Computing in Civil Engineering*, Ed. Thomas O. Branwell, Jr., ASCE, Reston, VA.

Sprinkle, H. (1972), “Analysis of Time-Lapse Construction Films,” *Journal of the Construction Division*, ASCE, 98(**CO**2), pp183-199.

Tserng, H. and J. Russell (1997), “ An Instantaneous Motion Planning and Controlling Tool (IMPACT) for Construction Equipment,” in the *Proceedings of the 1997 Construction Congress*, ASCE, Reston, VA

Zito, R. and M. A. Taylor, (1994), “The use of GPS in travel-time surveys,” *Traffic Engineering & Control,* 35(12), pp685-690.

CHAPTER 3. REDUCTION OF SHORT-INTERVAL GPS DATA FOR CONSTRUCTION OPERATIONS ANALYSIS

3.1. Introduction

Time studies have historically been performed to record the time required to complete various construction tasks [Oglesby et. al 1989]. The systems used to collect time study data have primarily been manual in nature and limited to the field of view of the observer. The original time study system was the stopwatch, which has since been replaced by time-lapse video recordings. Recently, analysts have turned to technology for tools not limited to field of view. Regardless of the tool implemented, the process is performed in four phases: data capture, preliminary review, data reduction, and data analysis. The tools used in each phase for various systems can be seen in Table 3. 1.

Table 3.1: Time Study Tools

Phase/System	Stopwatch Study	Time-Lapse Video	Short-Interval GPS
Data Capture	Stopwatch	Camera	GPS Box
Preliminary Review	N/A	Watch video	Graphical presentations
Data Reduction	By the user during data capture	PAVIC/DVAT	TIMs
Data Analysis	Statistical Analysis Software		

Whether an observer performs the study in the field with a stopwatch or the operation is filmed for preliminary review and data reduction in the office, the information is limited to the field of view of the observer or camera. Analysts have looked to technology in the form of on-board instrumentation as a data collection tool not limited to field of view.

GPS can be integrated into an on-board instrumentation system to provide position and speed data. While GPS provides a potential solution to the field of view issue, it has other issues associated with it. Recording data frequently enough to provide good results produces a very large volume of data. With several thousand data records produced daily on each piece of instrumented equipment, the issue becomes how to manage the data and identify the relatively few key records that mark the beginning and end of the activity being studied.

These key points are identified in real-time when personnel observe an operation, whether in the field or reviewing videotape in the office. The observer bases the instantaneous decisions of when one activity stops and the next begins on enormous volumes of visual information. This work develops a methodology for making equivalent decisions based not on visual information, but rather on GPS data. Specifically, this work answers the question: How can the large volume of short-interval GPS data collected be automatically reduced to the small volume of key discrete points that mark the start and stop of activities?

This chapter briefly describes the hardware used to record the necessary GPS data and presents procedures developed to identify the key records necessary to calculate activity durations. A case study is used to show the application of the system to an earthmoving operation. This paper also postulates how the information from the data can be used in discrete event simulations. Descriptions of discrete event simulation applied to construction operations can be found in [Martinez and Ioannou 1995, Ioannou and Martinez 1996a, Ioannou and Martinez 1996b, Martinez 1998, Sangarayakul 1998, Cor and Martinez 1999, Ioannou 1999].

3.2. Previous Work

A review of literature pertaining to the use of GPS to collect productivity data from earthmoving operations and the reduction of the data for construction operations analysis was performed to establish the state of the art. This specific literature review yielded a single reference that describes a pilot study designed to collect productivity data regarding earthwork operations and transmit the data in real Ackroyd [1998]. Position information was transmitted via radio and recorded by a remote PC. Position was recorded at a two-minute interval and at the closure of proximity switches on levers of articulated haulers and motor scrapers. In the initial phase, analysis of the data focused on measuring cycle times, but no real analysis was performed of positions between events. In a second phase, strain gauges were used to estimate payload, and thereby provide productivity data. This work depicts the potential of GPS for recording data from earthmoving operations, but focuses on the hardware issues rather than the data reduction issues. The author concludes that the information may add value to monitoring and scheduling tasks, but does not address use of the data for analyzing earthmoving operations.

3.2.1. Field Data Reduction

Oglesby et al. [1989] point out that the purpose of a time study is to record the times of various activities that make up an operation. The detailed record produced shows the duration of each task.

The simplest time studies, stopwatch studies, rely on the observer to decide the point in time at which activities start and stop. Oglesby et al. [1989] note that placing this responsibility on the observer limits the usefulness of the data. Observers can have differences in opinion, data recorded by a single observer may vary over time, available information is strictly limited to that recorded, and is subject to the physical limitations of the observer. Additionally, the information recorded is strictly limited to that field of view of the observer.

Bjornsson and Sagert [1994] presented PAVIC+ as computer-based system to extract data from video recordings of construction operations, including activity durations. Each activity instance is called a segment and defined by a start and stop time. Three tools are available for registering segments: start-stop time, cycle time, and consecutive time. The start-stop time tool operates like a stopwatch, the user indicates through keystrokes both the start and stop time. PAVIC+ reads the time of the keystroke from a time stamp placed on the audio track and records it. The cycle time tool is used when the same work sequence is repeated. The keystroke used to indicate the stop of an activity also indicates the start of the successive activity. This eliminates the need to press two keys simultaneously. The consecutive time tool is appropriate for a group of correlated activities, such as a crane serving multiple crews. Each activity is assigned a unique key that is struck when the activity ends. This keystroke indicates the stop of the associated activity and the start of the successive activity.

Frank [2001] describes a very similar system used to extract information from digital video. The author presents and discusses the Digital Video Analysis Tool (DVAT), and notes that PAVIC+ was the basis from which DVAT was developed. DVAT is a software package specifically developed for stripping information from digital video. The extracted information can be used to generate crew balance charts or to develop input for computer simulation models.

The traditional data collection methods have relied upon manual tools limited to the field of view of the observer. Researchers have turned to technology for tools that are both automated and not restricted to field of view. Kannan [1999] describes the use of on-board instrumentation systems for recording data from earthmoving operations. The author used mechanical parameters recorded by the Vital Information Management System (VIMS) and the Total Payload Management System (TPMS) produced by Caterpillar to obtain operational data. Such systems place a virtual observer in the equipment, and thus the equipment is always in the virtual field of view.

Kannan [1999] states that when using on-board instrumentation, the sensors must be able to detect a change in the status of the truck. It is this change in status that indicates the start or stop of an activity. Kannan and Vorster [2000] state that mechanical parameters can be translated to production measures through the use of surrogate measures and protocol rules. Kannan [1998] describes the use of suspension strut pressure, gear control lever position, bed raise switch position, and speed to derive the duration of truck cycle components. Truck cycle components, or activities, are defined in terms of the parameters.

It is evident from the literature reviewed that traditional data collection methods are limited by the ability of an observer to instantaneously decide activity start and stop times within a narrow and static field of view. Modern techniques remove this limitation by using sensor-based data collection techniques to identify key times through changes in sensor output.

3.3. Data Capture

An automated data capture system based on GPS technology was used to record the raw data. The system was installed on equipment engaged in a load and haul earthmoving operation. Data recorded included date, time, speed, and horizontal position. The system was designed to log GPS and other sensor data at a specified time interval.

3.3.1. Hardware Description

Commercially available systems were field evaluated and found to be capable of collecting the desired data, but deficient in critical areas. The systems are designed to operate on a 12-volt system, not the 24-volt systems of construction equipment. Data collection at two-second intervals fills the memory buffer in only a few hours; data from an entire workday cannot be stored. The cellular means of data transmission built into the systems is not sufficient for transmitting the large volumes of data intended to be recorded, nor to transmit the small volumes of data that constituted a full memory bank.

The evaluations showed that the technology existed to record GPS data at a short-interval, but a system satisfying the demands of the construction industry was not readily available. Therefore, the results of the evaluations were used to form the basis of recommendations for a unique GPS data capturing system for earthmoving construction equipment. The system was designed to log a user-defined string of GPS and other sensor data at a user-defined time interval. The data string is recorded in comma delimited ASCII text format. The system was designed to operate on the 24-volt electrical system of construction equipment.

The system hardware developed as a result of lessons learned is shown in Figure 3.1 and consists of a data box, junction box, and sensor pack. Chapter 2 provides a detailed description of the system hardware. The data acquisition and storage box is a weather-tight enclosure that houses a micro-controller, PC-104 embedded PC, and a CompactFlash media writer. The micro-controller gathers the sensor outputs, assembles the string, and passes the string to the PC. The PC interfaces with the media writer to store the data on a removable CompactFlash card.

The sensor pack consists of a Garmin GPS-35 LVC receiver and a Vector 2X magnetic compass manufactured by Precision Navigation, Inc. In addition to the GPS and compass, the system has been developed such that additional analog and digital sensors can be incorporated into the system for future research.

The junction box serves as a central location for all electrical connections to be made. Power is wired into the box and distributed to the data box and sensor pack. Power is also available at five volts for any additional analog sensors incorporated into the system.

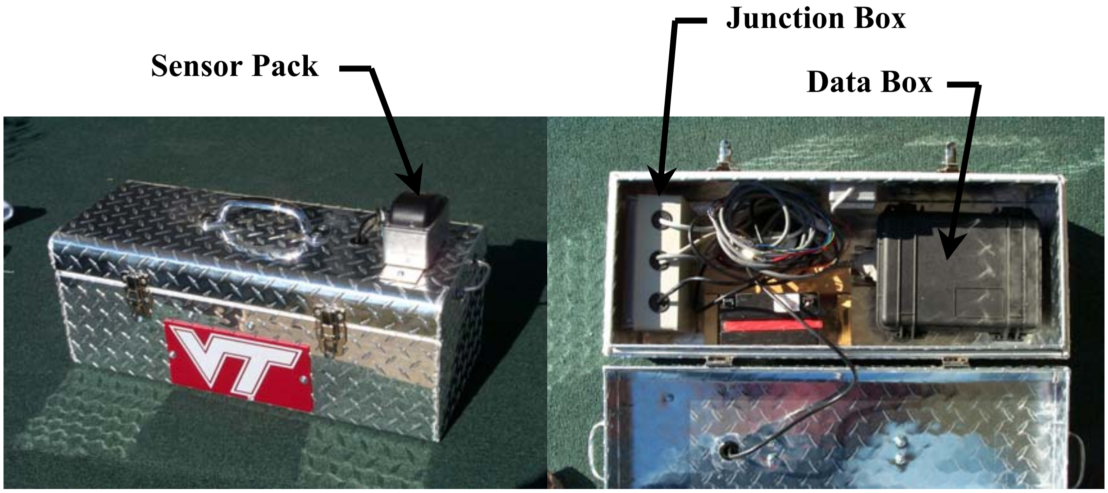


Figure 3.1: Mobile Hardware Configuration

3.3.2. Data Description

Data is recorded in comma delimited ASCII text format, and temporarily stored by the system on the CompactFlash media card. A separate data file is created on the card for each day of collection. Dividing the data by date of recording maintains manageable files sizes and eliminates the need to divide the data during processing.

The parameters recorded by the system include unit identification number, date, UTM time, latitude, longitude, speed, type of GPS fix, output from the five analog and eight digital sensors, and direction of travel from the compass. A sample data string is

```
06,160801,10:15:22,A,3710.8110,N,08023.5380,W,008.6,0000,0000,0000,023,037,0,0,0,0,0,0,0,0,139
```

Where **06** is the unit identification number
160801 is the date – Aug. 16, 2001
10:15:22 is UTM time
A represents the GPS-fix as three-dimensional
3710.8110, N is the latitude - 37° 10.8110’ North
08023.5380, W is the longitude - 80° 23.5380’ West
008.6 is speed in kilometers per hour
0000,0000,0000,023,037 are the outputs from the five analog sensors (future)
0,0,0,0,0,0,0,0 are the outputs from the eight digital sensors (future)
139 is the direction of travel in degrees from north

3.4. Preliminary Review

A trade-off exists between the volume of data recorded and the level of detail provided by the data. Data recorded at a short interval provides sufficient detail for operations analysis, but results in a large volume of data. The several thousand records produced daily on each machine must be reduced to the relatively few key records necessary to calculate activity times. Additionally, the volume of data dictates that the data reduction process be automated.

The format of the data, a text string repeatedly recorded at a set time interval, is conducive to analysis in a spreadsheet environment supported by capable graphics. Microsoft Excel is a common and robust spreadsheet application that can be automated through the Visual Basic for Applications (VBA) language and provides the means to graphically present data.

Excel was chosen as the data processing engine and is a widely used and recognized spreadsheet program already familiar to many potential users, thereby minimizing or eliminating the learning curve.

Processing the collected data for evaluation and reduction is performed in three steps:

1. Imported the data into Excel,
2. Convert recorded horizontal position from geodetic to planar coordinates,
3. Prepare graphical representations of the data.

These three steps have been automated using VBA code, and this routine is available from the system user interface, provided as Figure 3.2. The process of importing, converting, and graphing is executed by selecting the Import Text File option, which allows the user to point to the desired data file.

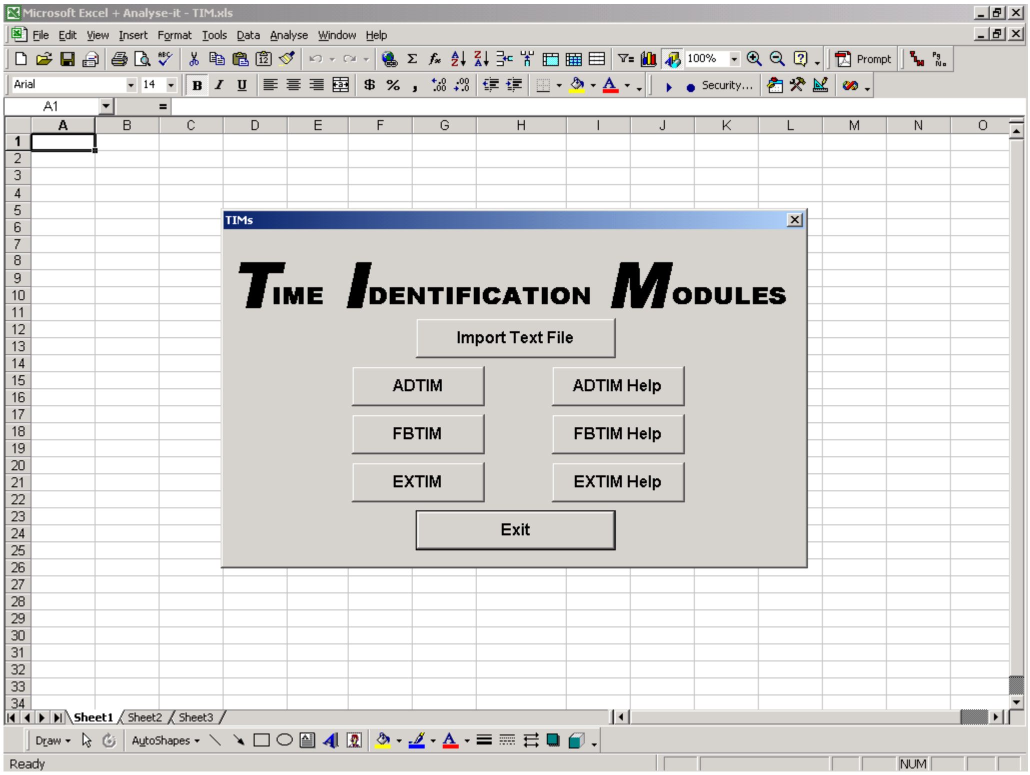
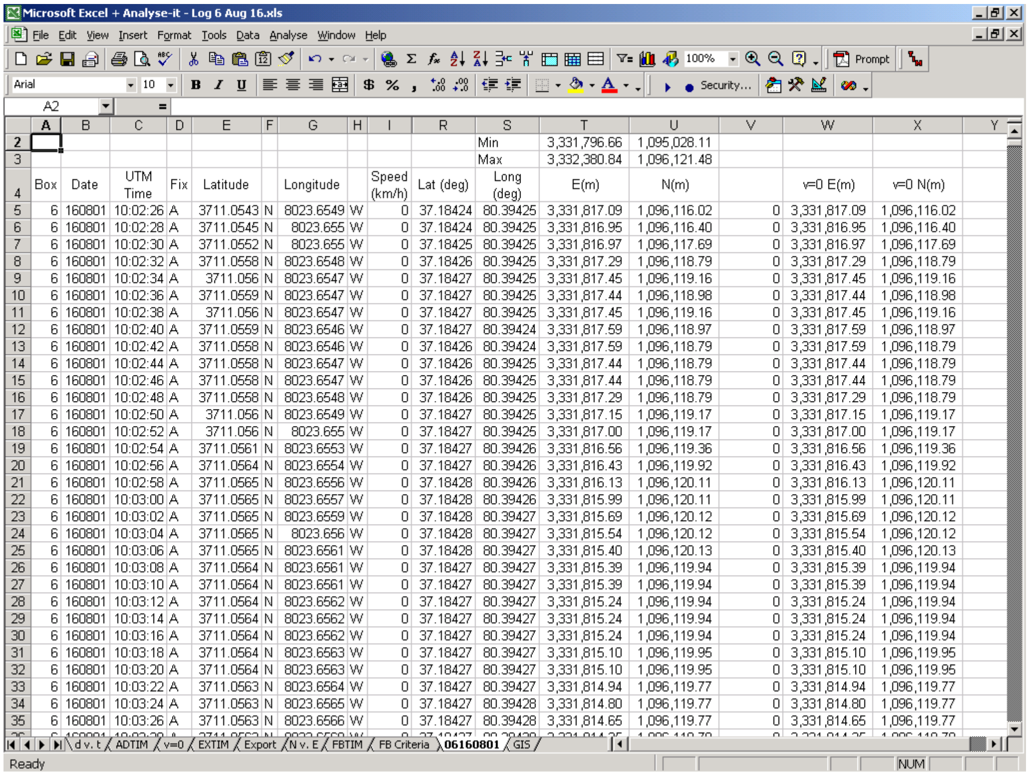


Figure 3.2: User Interface

Once the file is identified, the data is imported into Excel with each record in a separate row and each parameter in a separate column, as seen in Figure 3.3. The geodetic positions are recorded as latitude and longitude in units of degrees and minutes. The scale of even a very large construction earthmoving operation is sufficiently small to make geodetic coordinates less than meaningful. Positions expressed in units of meters or feet are much more appropriate. Therefore, the recorded geodetic coordinates are converted to a planar coordinate system. For the data recorded near Blacksburg, Virginia, the Virginia South State Plane Coordinate System was chosen. The methods to convert from geodetic to State Plane Coordinates are outlined in several standard surveying texts [Stem 1990]. The raw positional data recorded can be seen in columns E through H in Figure 3.3, while the converted planar coordinates are in columns T and U.



	A	B	C	D	E	F	G	H	I	R	S	T	U	V	W	X	Y
2											Min	3,331,796.66	1,095,028.11				
3											Max	3,332,380.84	1,096,121.48				
4	Box	Date	UTM Time	Fix	Latitude		Longitude		Speed (km/h)	Lat (deg)	Long (deg)	E(m)	N(m)		v=0 E(m)	v=0 N(m)	
5	6	160801	10:02:26	A	3711.0543	N	8023.6549	W	0	37.18424	80.39425	3,331,817.09	1,096,116.02	0	3,331,817.09	1,096,116.02	
6	6	160801	10:02:28	A	3711.0545	N	8023.655	W	0	37.18424	80.39425	3,331,816.95	1,096,116.40	0	3,331,816.95	1,096,116.40	
7	6	160801	10:02:30	A	3711.0552	N	8023.655	W	0	37.18425	80.39425	3,331,816.97	1,096,117.69	0	3,331,816.97	1,096,117.69	
8	6	160801	10:02:32	A	3711.0558	N	8023.6548	W	0	37.18426	80.39425	3,331,817.29	1,096,118.79	0	3,331,817.29	1,096,118.79	
9	6	160801	10:02:34	A	3711.056	N	8023.6547	W	0	37.18427	80.39425	3,331,817.45	1,096,119.16	0	3,331,817.45	1,096,119.16	
10	6	160801	10:02:36	A	3711.0559	N	8023.6547	W	0	37.18427	80.39425	3,331,817.44	1,096,118.98	0	3,331,817.44	1,096,118.98	
11	6	160801	10:02:38	A	3711.056	N	8023.6547	W	0	37.18427	80.39425	3,331,817.45	1,096,119.16	0	3,331,817.45	1,096,119.16	
12	6	160801	10:02:40	A	3711.0559	N	8023.6546	W	0	37.18427	80.39424	3,331,817.59	1,096,118.97	0	3,331,817.59	1,096,118.97	
13	6	160801	10:02:42	A	3711.0558	N	8023.6546	W	0	37.18426	80.39424	3,331,817.59	1,096,118.79	0	3,331,817.59	1,096,118.79	
14	6	160801	10:02:44	A	3711.0558	N	8023.6547	W	0	37.18426	80.39425	3,331,817.44	1,096,118.79	0	3,331,817.44	1,096,118.79	
15	6	160801	10:02:46	A	3711.0558	N	8023.6547	W	0	37.18426	80.39425	3,331,817.44	1,096,118.79	0	3,331,817.44	1,096,118.79	
16	6	160801	10:02:48	A	3711.0558	N	8023.6548	W	0	37.18426	80.39425	3,331,817.29	1,096,118.79	0	3,331,817.29	1,096,118.79	
17	6	160801	10:02:50	A	3711.056	N	8023.6549	W	0	37.18427	80.39425	3,331,817.15	1,096,119.17	0	3,331,817.15	1,096,119.17	
18	6	160801	10:02:52	A	3711.056	N	8023.655	W	0	37.18427	80.39425	3,331,817.00	1,096,119.17	0	3,331,817.00	1,096,119.17	
19	6	160801	10:02:54	A	3711.0561	N	8023.6553	W	0	37.18427	80.39426	3,331,816.56	1,096,119.36	0	3,331,816.56	1,096,119.36	
20	6	160801	10:02:56	A	3711.0564	N	8023.6554	W	0	37.18427	80.39426	3,331,816.43	1,096,119.92	0	3,331,816.43	1,096,119.92	
21	6	160801	10:02:58	A	3711.0565	N	8023.6556	W	0	37.18428	80.39426	3,331,816.13	1,096,120.11	0	3,331,816.13	1,096,120.11	
22	6	160801	10:03:00	A	3711.0565	N	8023.6557	W	0	37.18428	80.39426	3,331,815.99	1,096,120.11	0	3,331,815.99	1,096,120.11	
23	6	160801	10:03:02	A	3711.0565	N	8023.6559	W	0	37.18428	80.39427	3,331,815.69	1,096,120.12	0	3,331,815.69	1,096,120.12	
24	6	160801	10:03:04	A	3711.0565	N	8023.656	W	0	37.18428	80.39427	3,331,815.54	1,096,120.12	0	3,331,815.54	1,096,120.12	
25	6	160801	10:03:06	A	3711.0565	N	8023.6561	W	0	37.18428	80.39427	3,331,815.40	1,096,120.13	0	3,331,815.40	1,096,120.13	
26	6	160801	10:03:08	A	3711.0564	N	8023.6561	W	0	37.18427	80.39427	3,331,815.39	1,096,119.94	0	3,331,815.39	1,096,119.94	
27	6	160801	10:03:10	A	3711.0564	N	8023.6561	W	0	37.18427	80.39427	3,331,815.39	1,096,119.94	0	3,331,815.39	1,096,119.94	
28	6	160801	10:03:12	A	3711.0564	N	8023.6562	W	0	37.18427	80.39427	3,331,815.24	1,096,119.94	0	3,331,815.24	1,096,119.94	
29	6	160801	10:03:14	A	3711.0564	N	8023.6562	W	0	37.18427	80.39427	3,331,815.24	1,096,119.94	0	3,331,815.24	1,096,119.94	
30	6	160801	10:03:16	A	3711.0564	N	8023.6562	W	0	37.18427	80.39427	3,331,815.24	1,096,119.94	0	3,331,815.24	1,096,119.94	
31	6	160801	10:03:18	A	3711.0564	N	8023.6563	W	0	37.18427	80.39427	3,331,815.10	1,096,119.95	0	3,331,815.10	1,096,119.95	
32	6	160801	10:03:20	A	3711.0564	N	8023.6563	W	0	37.18427	80.39427	3,331,815.10	1,096,119.95	0	3,331,815.10	1,096,119.95	
33	6	160801	10:03:22	A	3711.0563	N	8023.6564	W	0	37.18427	80.39427	3,331,814.94	1,096,119.77	0	3,331,814.94	1,096,119.77	
34	6	160801	10:03:24	A	3711.0563	N	8023.6565	W	0	37.18427	80.39428	3,331,814.80	1,096,119.77	0	3,331,814.80	1,096,119.77	
35	6	160801	10:03:26	A	3711.0563	N	8023.6566	W	0	37.18427	80.39428	3,331,814.65	1,096,119.77	0	3,331,814.65	1,096,119.77	

Figure 3.3: Imported Data and Converted Coordinates

In addition to importing the data and converting the coordinates, two separate graphs of the data are prepared during the Import Text routine to assist the user in reviewing and understanding the data. Mapping the recorded coordinates creates a plan view of vehicle position throughout the workday. This graph is enhanced by using color to distinguish between positions recorded when speed is below a user defined value, most frequently zero, and those recorded when speed was greater than the value. A plan view of data collected during one day of the field application is provided as Figure 3.4; those points in grey are when speed was zero, those in black represent velocities greater than zero.

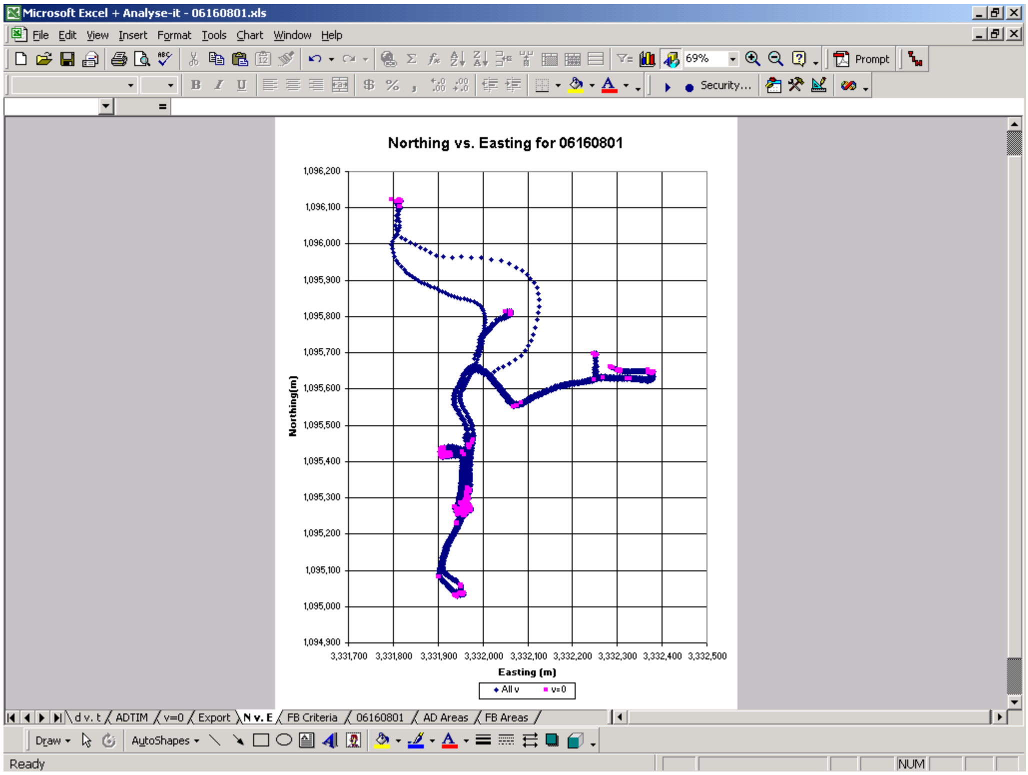


Figure 3.4: Plan View of Recorded Positions

Time is of paramount interest from the data, and yet the most straightforward graphic, the plan view, does not contain time information. Therefore, a graph of distance from a user specified point versus time is also prepared to aid in understanding the data. To provide a two-dimensional plot of position versus time, it is necessary to have a single value represent position. This is achieved by using the distance to a fixed-point to represent position. The point can be arbitrarily chosen, but choosing the approximate center of the loading area will provide additional meaning to the graph. As the truck returns to the load area, the distance approaches zero and allows the user to determine a load count from the graph. The distance is calculated and plotted against time for those records when speed was equal to zero. Such a graph is provided as Figure 3.5.

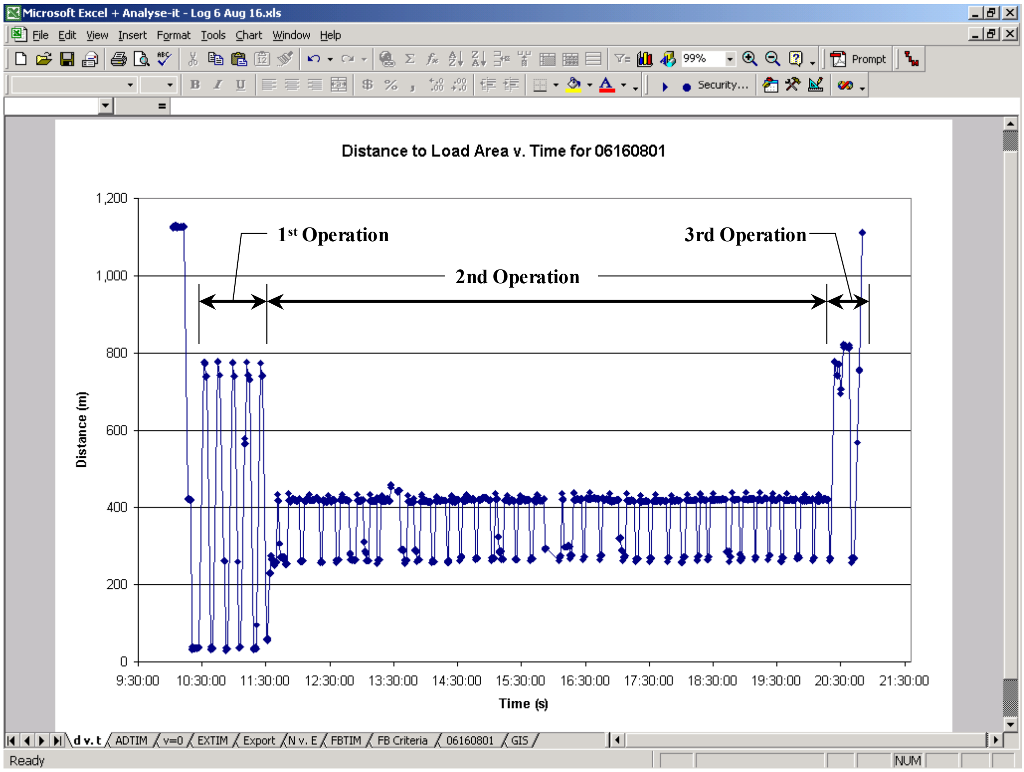


Figure 3.5: Position versus Time Graph

The value of this graphic is the fact that a steady pattern depicts a steady cyclic operation, and that changes in pattern indicate changes in the cyclic operation. Three separate operations were identified in Figure 3.5, and the number of peaks associated with the operation indicates the number of cycles performed.

Preparations are also made during the Import Text process for data reduction. Three additional sheets are placed in the workbook and populated with data to be used by the data reduction modules developed as a part of this work and described in the following section.

3.5. Data Reduction

Many approaches were explored in attempts to reduce the data file to a manageable size and identify the records that mark critical aspects of truck cycles. Initially, attempts were made to review the collected data in its raw form. The enormous data volume made review in a timely manner infeasible and led the user into a state of information overload. It was recognized that user review was necessary, but that data reduction was required first. Early investigations into the data were focused on analyzing the speed data. Plan view plots, not unlike Figure 3.4, were generated with point color used to distinguish between speed magnitudes. Such plots allowed the user to identify locations of interest, but provided no information as to when interesting things happened.

This led to a shift in the focus of research efforts to analysis of the position data collected. Attempts to identify the key records were not entirely successful based solely on position data. It was found that both position and speed criteria were necessary to successfully identify the key records.

The process led to the understanding that the use of mobile vehicles can be better understood if data is obtained for the following three conditions:

1. Time spent with speed equal to zero within a specified fixed location – e.g. an aggregate haul truck loading at a bin.
2. Time spent with speed equal to zero within a fixed distance of a moving location – e.g. a haul truck loading at a loader.
3. Time required to travel through or between fixed areas – e.g. a haul truck traveling between load and dump areas.

Time Identification Modules (TIMs) were developed to identify the GPS data records corresponding to the beginning and end of activities associated with each of the three specific conditions. These modules have been developed in the Visual Basic for Applications (VBA) language for execution in Excel.

The results produced by TIMs were compared to those produced by PAVIC+ analysis of the videotaped operation in an effort to validate the TIMs results. The standard methods for video recording of construction operations and PAVIC+ data reduction were applied. GPS data was recorded at the same time and reduced using the TIMs. The activity durations resulting from both analyses were compared, and are presented as Table 3. 2.

Table 3.2: Activity Duration Comparison

	Duration				Duration		
Activity	PAVIC+	TIMs	Difference	Activity	PAVIC+	TIMs	Difference
Load		0:07:08		Dump		0:02:04	
Load	0:03:37	0:03:42	0:00:05	Dump	0:05:55	0:05:56	0:00:01
Load	0:07:33	0:07:34	0:00:01	Dump	0:01:54	0:01:58	0:00:04
Load	0:05:37	0:05:40	0:00:03	Dump	0:04:33	0:04:32	0:00:01
Load	0:02:36	0:02:38	0:00:02	Dump	0:01:41	0:01:42	0:00:01
Load	0:04:45	0:04:54	0:00:09	Dump	0:02:24	0:02:26	0:00:02
Load	0:03:57	0:04:00	0:00:03	Dump	0:03:36	0:03:38	0:00:02
Load		0:05:32					
Haul		0:00:58		Return		0:00:54	
Haul	0:01:03	0:00:58	0:00:05	Return		0:00:50	
Haul	0:01:01	0:01:00	0:00:01	Return	0:00:51	0:00:50	0:00:01
Haul	0:01:04	0:01:04	0:00:00	Return	0:00:56	0:00:52	0:00:04
Haul	0:01:02	0:01:00	0:00:02	Return	0:00:50	0:00:50	0:00:00
Haul	0:01:10	0:01:02	0:00:08	Return	0:01:16	0:01:14	0:00:02
Haul	0:01:03	0:01:00	0:00:03	Return	0:00:54	0:00:54	0:00:00
Haul		0:01:02		Return	0:00:51	0:00:48	0:00:03

The results of the analyses show that the activity durations, as well as the activity order, determined from the GPS data agree with those determined from PAVIC+ analysis of videotape. These results lead to the conclusion that the short-interval GPS data can be reduced to the key records using the TIMs.

The Arrive and Depart TIM (ADTIM) was developed to identify the first (arrive) and last (depart) times at which the truck was stopped within a given distance of a user-defined point. ADTIM was used to determine dump times from the data collected.

The External Data TIM (EXTIM) was developed to identify the first (arrive) and last (depart) times at which the truck stops within a given distance of a moving point. Load times were determined from the collected data using the EXTIM.

The Fixed Boundary TIM (FBTIM) was developed to identify entry and exit times for a user-defined rectangular area. Haul and return times were determined from the collected data using the FBTIM.

3.5.1. ADTIM

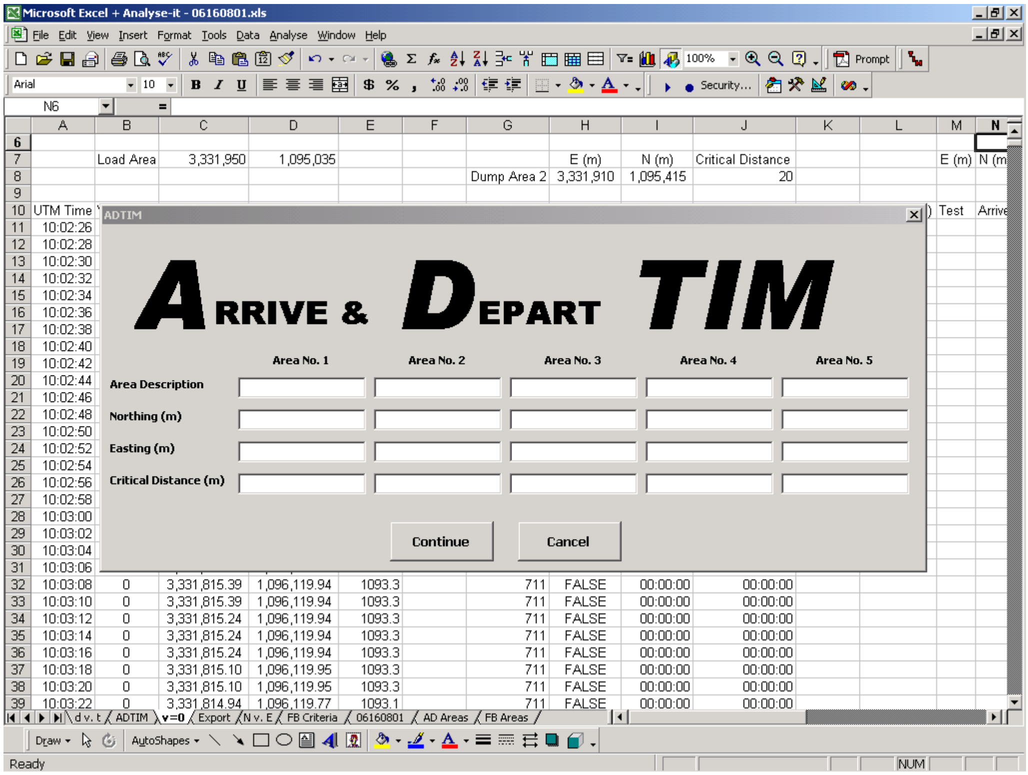
The duration of construction activities that repeatedly occur in a fixed location can be characterized as the time elapsed while the vehicle was stopped within a given distance from a fixed point. A truck loaded from a hot mix asphalt bin or a truck dumping into the bin of a rock crusher are examples of such activities. The Arrive and Depart Time Identification Module (ADTIM) was developed for this condition and to identify the first (arrive) and last (depart) times at which the truck was stopped within a given distance of a user-defined point. The time difference between arriving and departing represents the activity duration.

Three criteria must be satisfied for a record to be considered as an Arrival or Departure time:

1. The recorded speed must be zero (i.e. the truck is stopped)
2. The position of the truck must be within a user-defined area
3. The previous record must be outside of the area for Arrival times, or the next record must be outside of the area for Departure times.

The ADTIM evaluates each of the three criteria in the order in which they were presented. To ensure that the recorded speed was zero, all calculations are performed on a "v=0" sheet which contains only those records where speed equaled zero.

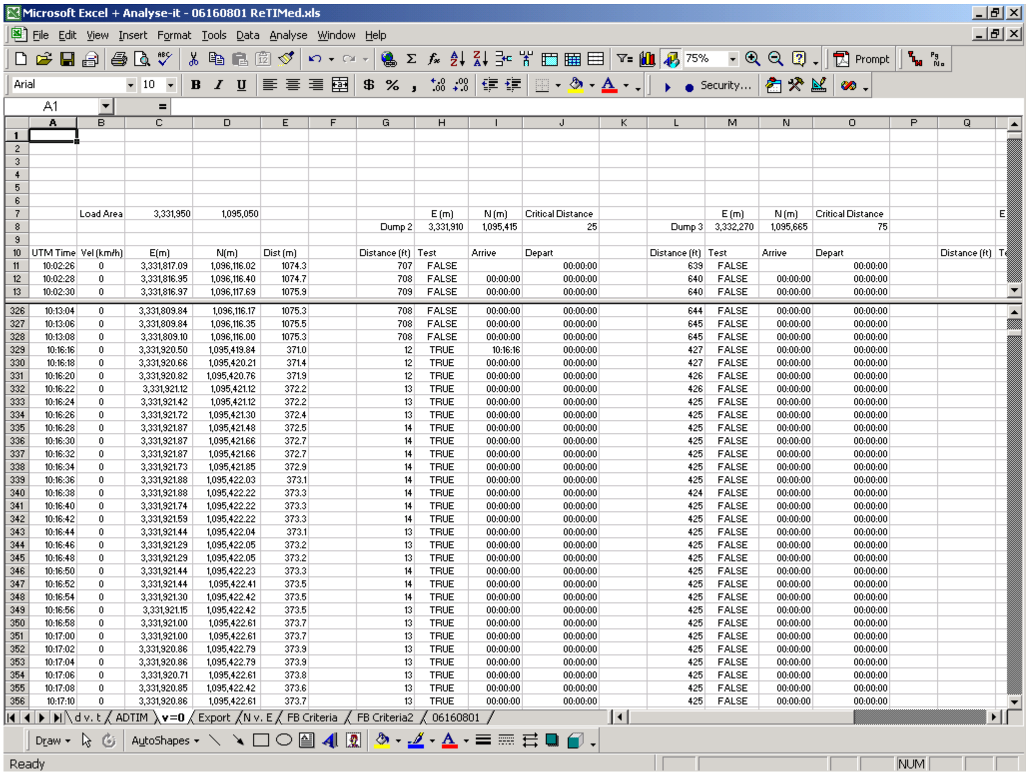
The user defines the second criterion, the area within which the truck was stopped, the area by providing a set of central coordinates and a radial distance on the ADTIM Input form, shown as Figure 3.6. The user may define as many as five ADTIM areas.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
6														
7		Load Area	3,331,950	1,095,035				E (m)	N (m)	Critical Distance			E (m)	N (m
8							Dump Area 2	3,331,910	1,095,415	20				
9														
10	UTM Time											)	Test	Arrive
11	10:02:26													
12	10:02:28													
13	10:02:30													
14	10:02:32													
15	10:02:34													
16	10:02:36													
17	10:02:38													
18	10:02:40													
19	10:02:42													
20	10:02:44													
21	10:02:46													
22	10:02:48													
23	10:02:50													
24	10:02:52													
25	10:02:54													
26	10:02:56													
27	10:02:58													
28	10:03:00													
29	10:03:02													
30	10:03:04													
31	10:03:06													
32	10:03:08	0	3,331,815.39	1,096,119.94	1093.3		711	FALSE	00:00:00	00:00:00				
33	10:03:10	0	3,331,815.39	1,096,119.94	1093.3		711	FALSE	00:00:00	00:00:00				
34	10:03:12	0	3,331,815.24	1,096,119.94	1093.3		711	FALSE	00:00:00	00:00:00				
35	10:03:14	0	3,331,815.24	1,096,119.94	1093.3		711	FALSE	00:00:00	00:00:00				
36	10:03:16	0	3,331,815.24	1,096,119.94	1093.3		711	FALSE	00:00:00	00:00:00				
37	10:03:18	0	3,331,815.10	1,096,119.95	1093.3		711	FALSE	00:00:00	00:00:00				
38	10:03:20	0	3,331,815.10	1,096,119.95	1093.3		711	FALSE	00:00:00	00:00:00				
39	10:03:22	0	3,331,814.94	1,096,119.77	1093.1		711	FALSE	00:00:00	00:00:00				

Figure 3.6: ADTIM Input Form

The third criterion is evaluated within the software. For each defined area, each position recorded with zero speed is used to calculate a distance between the truck and the defined point. This calculated distance is compared to the user supplied critical distance to determine if the truck is within the defined area, and each record is marked appropriately. The records are then checked again to identify changes in truck status and those records corresponding to arrive and depart times are marked with the time at which the position was recorded. This can be seen in Figure 3.7.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	
1																		
2																		
3																		
4																		
5																		
6																		
7		Load Area	3,331,950	1,095,050				E (m)	N (m)	Critical Distance			E (m)	N (m)	Critical Distance			E
8							Dump 2	3,331,910	1,095,415	25		Dump 3	3,332,270	1,095,665	75			
9																		
10	UTM Time	Vel (km/h)	E(m)	N(m)	Dist (m)		Distance (ft)	Test	Arrive	Depart		Distance (ft)	Test	Arrive	Depart		Distance (ft)	Te
11	10:02:26	0	3,331,817.09	1,096,116.02	1074.3		707	FALSE		00:00:00		639	FALSE		00:00:00			
12	10:02:28	0	3,331,816.95	1,096,116.40	1074.7		708	FALSE	00:00:00	00:00:00		640	FALSE	00:00:00	00:00:00			
13	10:02:30	0	3,331,816.97	1,096,117.69	1075.9		709	FALSE	00:00:00	00:00:00		640	FALSE	00:00:00	00:00:00			
326	10:13:04	0	3,331,809.84	1,096,116.17	1075.3		708	FALSE	00:00:00	00:00:00		644	FALSE	00:00:00	00:00:00			
327	10:13:06	0	3,331,809.84	1,096,116.35	1075.5		708	FALSE	00:00:00	00:00:00		645	FALSE	00:00:00	00:00:00			
328	10:13:08	0	3,331,809.10	1,096,116.00	1075.3		708	FALSE	00:00:00	00:00:00		645	FALSE	00:00:00	00:00:00			
329	10:16:16	0	3,331,920.50	1,095,419.84	371.0		12	TRUE	10:16:16	00:00:00		427	FALSE	00:00:00	00:00:00			
330	10:16:18	0	3,331,920.66	1,095,420.21	371.4		12	TRUE	00:00:00	00:00:00		427	FALSE	00:00:00	00:00:00			
331	10:16:20	0	3,331,920.82	1,095,420.76	371.9		12	TRUE	00:00:00	00:00:00		426	FALSE	00:00:00	00:00:00			
332	10:16:22	0	3,331,921.12	1,095,421.12	372.2		13	TRUE	00:00:00	00:00:00		426	FALSE	00:00:00	00:00:00			
333	10:16:24	0	3,331,921.42	1,095,421.12	372.2		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
334	10:16:26	0	3,331,921.72	1,095,421.30	372.4		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
335	10:16:28	0	3,331,921.87	1,095,421.48	372.5		14	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
336	10:16:30	0	3,331,921.87	1,095,421.66	372.7		14	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
337	10:16:32	0	3,331,921.87	1,095,421.66	372.7		14	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
338	10:16:34	0	3,331,921.73	1,095,421.85	372.9		14	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
339	10:16:36	0	3,331,921.88	1,095,422.03	373.1		14	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
340	10:16:38	0	3,331,921.88	1,095,422.22	373.3		14	TRUE	00:00:00	00:00:00		424	FALSE	00:00:00	00:00:00			
341	10:16:40	0	3,331,921.74	1,095,422.22	373.3		14	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
342	10:16:42	0	3,331,921.59	1,095,422.22	373.3		14	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
343	10:16:44	0	3,331,921.44	1,095,422.04	373.1		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
344	10:16:46	0	3,331,921.29	1,095,422.05	373.2		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
345	10:16:48	0	3,331,921.29	1,095,422.05	373.2		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
346	10:16:50	0	3,331,921.44	1,095,422.23	373.3		14	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
347	10:16:52	0	3,331,921.44	1,095,422.41	373.5		14	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
348	10:16:54	0	3,331,921.30	1,095,422.42	373.5		14	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
349	10:16:56	0	3,331,921.15	1,095,422.42	373.5		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
350	10:16:58	0	3,331,921.00	1,095,422.61	373.7		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
351	10:17:00	0	3,331,921.00	1,095,422.61	373.7		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
352	10:17:02	0	3,331,920.86	1,095,422.79	373.9		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
353	10:17:04	0	3,331,920.86	1,095,422.79	373.9		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
354	10:17:06	0	3,331,920.71	1,095,422.61	373.8		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
355	10:17:08	0	3,331,920.85	1,095,422.42	373.6		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			
356	10:17:10	0	3,331,920.86	1,095,422.61	373.7		13	TRUE	00:00:00	00:00:00		425	FALSE	00:00:00	00:00:00			

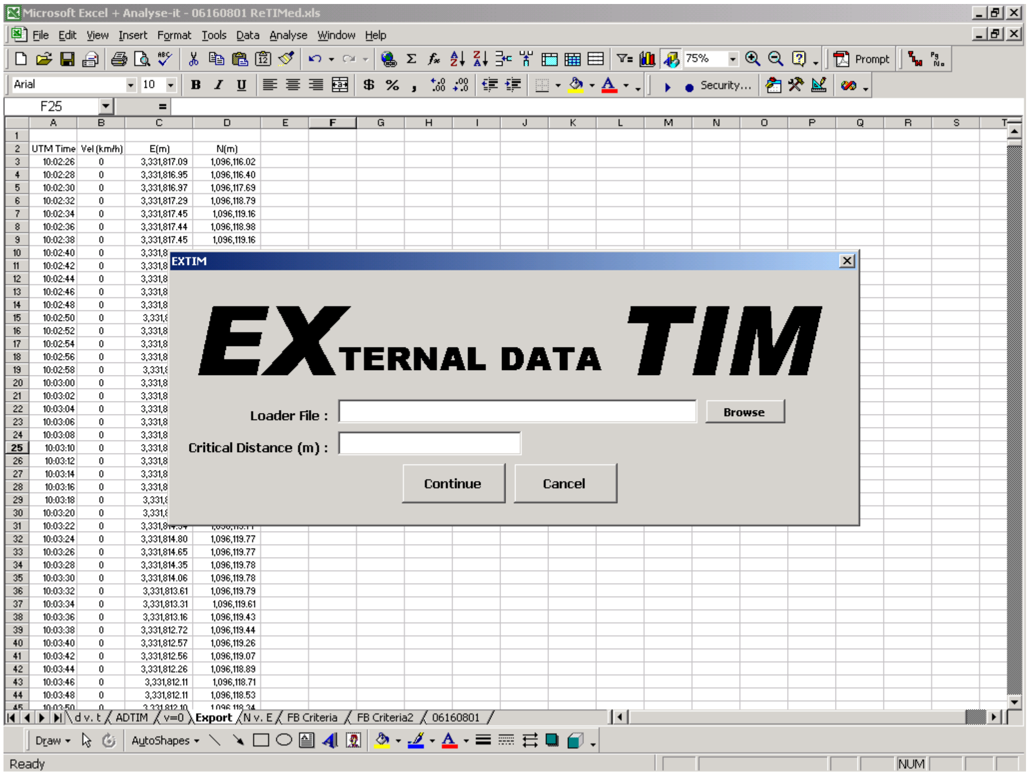
Figure 3.7: ADTIM Data on the “v=0” Sheet

3.5.2. EXTIM

Some earthmoving activities do not always occur in the same location, but rather occur relative to a second location that moves over time. An example of such an activity is a truck loaded by a rubber-tired loader. While the location of the loading activity may not move significantly from one instance to the next, it may very well change significantly over the course of a day or several days. The area in which the load activity occurs thus moves depending on the location of the loader. The External Data TIM (EXTIM) was developed to return the first (arrive) and last (depart) times at which the truck stops within a given distance of a moving point defined by the loader position, or any other instrumented equipment.

The EXTIM operates very similarly to the ADTIM; with the exception that the user defines only a single area and that the central coordinates of the area are defined by the positions collected for the moving point. The coordinates are read from those recorded by the system installed on another piece of equipment, such as a rubber-tired loader used to load trucks. Therefore, prior to executing EXTIM, it is necessary to have executed the Import Text routine on the data file collected from the loader system.

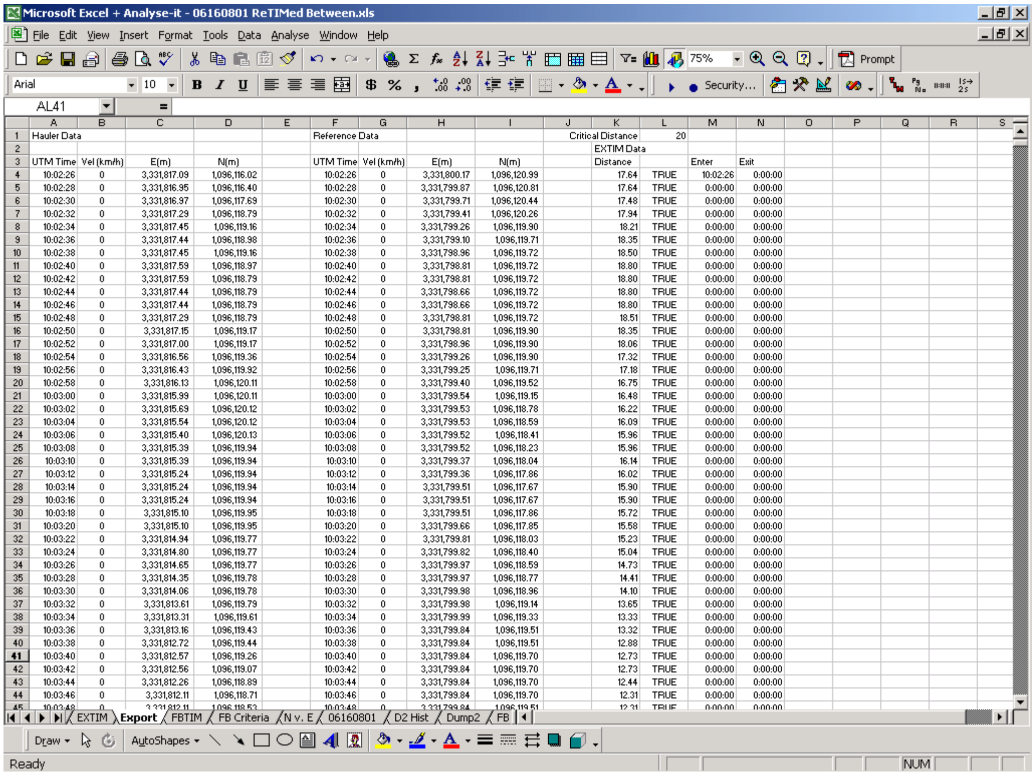
When the EXTIM button is selected from the user interface, the EXTIM Input dialog box appears and prompts the user for the reference data file to be read and the associated critical distance. The EXTIM Input dialog is shown in Figure 3.8.



	A	B	C	D
1				
2	UTM Time	Vel (km/h)	E(m)	N(m)
3	10:02:26	0	3,331,817.09	1,096,116.02
4	10:02:28	0	3,331,816.95	1,096,116.40
5	10:02:30	0	3,331,816.97	1,096,117.69
6	10:02:32	0	3,331,817.29	1,096,118.79
7	10:02:34	0	3,331,817.45	1,096,119.16
8	10:02:36	0	3,331,817.44	1,096,118.98
9	10:02:38	0	3,331,817.45	1,096,119.16
10	10:02:40	0	3,331,8	
11	10:02:42	0	3,331,8	
12	10:02:44	0	3,331,8	
13	10:02:46	0	3,331,8	
14	10:02:48	0	3,331,8	
15	10:02:50	0	3,331,	
16	10:02:52	0	3,331,8	
17	10:02:54	0	3,331,8	
18	10:02:56	0	3,331,8	
19	10:02:58	0	3,331,	
20	10:03:00	0	3,331,8	
21	10:03:02	0	3,331,8	
22	10:03:04	0	3,331,8	
23	10:03:06	0	3,331,8	
24	10:03:08	0	3,331,8	
25	10:03:10	0	3,331,8	
26	10:03:12	0	3,331,8	
27	10:03:14	0	3,331,8	
28	10:03:16	0	3,331,8	
29	10:03:18	0	3,331,	
30	10:03:20	0	3,331,	
31	10:03:22	0	3,331,8	
32	10:03:24	0	3,331,814.80	1,096,119.77
33	10:03:26	0	3,331,814.65	1,096,119.77
34	10:03:28	0	3,331,814.35	1,096,119.78
35	10:03:30	0	3,331,814.06	1,096,119.78
36	10:03:32	0	3,331,813.61	1,096,119.79
37	10:03:34	0	3,331,813.31	1,096,119.61
38	10:03:36	0	3,331,813.16	1,096,119.43
39	10:03:38	0	3,331,812.72	1,096,119.44
40	10:03:40	0	3,331,812.57	1,096,119.26
41	10:03:42	0	3,331,812.56	1,096,119.07
42	10:03:44	0	3,331,812.26	1,096,118.89
43	10:03:46	0	3,331,812.11	1,096,118.71
44	10:03:48	0	3,331,812.11	1,096,118.53

Figure 3.8: EXTIM Input Dialog Box

The EXTIM calculates the distance between the truck and loader for each position recorded, and compares this calculated distance to the critical distance. Because there is a comparison made between the position recorded on the truck and that recorded on the loader, it is necessary to time match the records. The data contained on the Export sheet of the loader file is copied and pasted onto the Export sheet of the truck file so that the time match can be performed. To time match the data, EXTIM loops through the recorded truck times and compares it to the loader time in the same row. Provided that the times are not equal, the data associated with the earlier of the two times is deleted and the same row rechecked to ensure a time match. A complete time match is ensured by looping through all of the truck data, not just providing an initial time match. The results of the time matching can be seen in Figure 3.9.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Hauler Data					Reference Data					Critical Distance	20		
2											EXTIM Data			
3	UTM Time	Vel (km/h)	E(m)	N(m)		UTM Time	Vel (km/h)	E(m)	N(m)		Distance		Enter	Exit
4	10:02:26	0	3,331,817.09	1,096,116.02		10:02:26	0	3,331,800.17	1,096,120.99		17.64	TRUE	10:02:26	0:00:00
5	10:02:28	0	3,331,816.95	1,096,116.40		10:02:28	0	3,331,799.87	1,096,120.81		17.64	TRUE	0:00:00	0:00:00
6	10:02:30	0	3,331,816.97	1,096,117.69		10:02:30	0	3,331,799.71	1,096,120.44		17.48	TRUE	0:00:00	0:00:00
7	10:02:32	0	3,331,817.29	1,096,118.79		10:02:32	0	3,331,799.41	1,096,120.26		17.94	TRUE	0:00:00	0:00:00
8	10:02:34	0	3,331,817.45	1,096,119.16		10:02:34	0	3,331,799.26	1,096,119.90		18.21	TRUE	0:00:00	0:00:00
9	10:02:36	0	3,331,817.44	1,096,118.98		10:02:36	0	3,331,799.10	1,096,119.71		18.35	TRUE	0:00:00	0:00:00
10	10:02:38	0	3,331,817.45	1,096,119.16		10:02:38	0	3,331,798.96	1,096,119.72		18.50	TRUE	0:00:00	0:00:00
11	10:02:40	0	3,331,817.59	1,096,118.97		10:02:40	0	3,331,798.81	1,096,119.72		18.80	TRUE	0:00:00	0:00:00
12	10:02:42	0	3,331,817.59	1,096,118.79		10:02:42	0	3,331,798.81	1,096,119.72		18.80	TRUE	0:00:00	0:00:00
13	10:02:44	0	3,331,817.44	1,096,118.79		10:02:44	0	3,331,798.66	1,096,119.72		18.80	TRUE	0:00:00	0:00:00
14	10:02:46	0	3,331,817.44	1,096,118.79		10:02:46	0	3,331,798.66	1,096,119.72		18.80	TRUE	0:00:00	0:00:00
15	10:02:48	0	3,331,817.29	1,096,118.79		10:02:48	0	3,331,798.81	1,096,119.72		18.51	TRUE	0:00:00	0:00:00
16	10:02:50	0	3,331,817.15	1,096,119.17		10:02:50	0	3,331,798.81	1,096,119.90		18.35	TRUE	0:00:00	0:00:00
17	10:02:52	0	3,331,817.00	1,096,119.17		10:02:52	0	3,331,798.96	1,096,119.90		18.06	TRUE	0:00:00	0:00:00
18	10:02:54	0	3,331,816.56	1,096,119.36		10:02:54	0	3,331,799.26	1,096,119.90		17.32	TRUE	0:00:00	0:00:00
19	10:02:56	0	3,331,816.43	1,096,119.92		10:02:56	0	3,331,799.25	1,096,119.71		17.18	TRUE	0:00:00	0:00:00
20	10:02:58	0	3,331,816.13	1,096,120.11		10:02:58	0	3,331,799.40	1,096,119.52		16.75	TRUE	0:00:00	0:00:00
21	10:03:00	0	3,331,815.99	1,096,120.11		10:03:00	0	3,331,799.54	1,096,119.15		16.48	TRUE	0:00:00	0:00:00
22	10:03:02	0	3,331,815.69	1,096,120.12		10:03:02	0	3,331,799.53	1,096,118.78		16.22	TRUE	0:00:00	0:00:00
23	10:03:04	0	3,331,815.54	1,096,120.12		10:03:04	0	3,331,799.53	1,096,118.59		16.09	TRUE	0:00:00	0:00:00
24	10:03:06	0	3,331,815.40	1,096,120.13		10:03:06	0	3,331,799.52	1,096,118.41		15.96	TRUE	0:00:00	0:00:00
25	10:03:08	0	3,331,815.39	1,096,119.94		10:03:08	0	3,331,799.52	1,096,118.23		15.96	TRUE	0:00:00	0:00:00
26	10:03:10	0	3,331,815.39	1,096,119.94		10:03:10	0	3,331,799.37	1,096,118.04		16.14	TRUE	0:00:00	0:00:00
27	10:03:12	0	3,331,815.24	1,096,119.94		10:03:12	0	3,331,799.36	1,096,117.86		16.02	TRUE	0:00:00	0:00:00
28	10:03:14	0	3,331,815.24	1,096,119.94		10:03:14	0	3,331,799.51	1,096,117.67		15.90	TRUE	0:00:00	0:00:00
29	10:03:16	0	3,331,815.24	1,096,119.94		10:03:16	0	3,331,799.51	1,096,117.67		15.90	TRUE	0:00:00	0:00:00
30	10:03:18	0	3,331,815.10	1,096,119.95		10:03:18	0	3,331,799.51	1,096,117.86		15.72	TRUE	0:00:00	0:00:00
31	10:03:20	0	3,331,815.10	1,096,119.95		10:03:20	0	3,331,799.66	1,096,117.85		15.58	TRUE	0:00:00	0:00:00
32	10:03:22	0	3,331,814.94	1,096,119.77		10:03:22	0	3,331,799.81	1,096,118.03		15.23	TRUE	0:00:00	0:00:00
33	10:03:24	0	3,331,814.80	1,096,119.77		10:03:24	0	3,331,799.82	1,096,118.40		15.04	TRUE	0:00:00	0:00:00
34	10:03:26	0	3,331,814.65	1,096,119.77		10:03:26	0	3,331,799.97	1,096,118.59		14.73	TRUE	0:00:00	0:00:00
35	10:03:28	0	3,331,814.35	1,096,119.78		10:03:28	0	3,331,799.97	1,096,118.77		14.41	TRUE	0:00:00	0:00:00
36	10:03:30	0	3,331,814.06	1,096,119.78		10:03:30	0	3,331,799.98	1,096,118.96		14.10	TRUE	0:00:00	0:00:00
37	10:03:32	0	3,331,813.61	1,096,119.79		10:03:32	0	3,331,799.98	1,096,119.14		13.65	TRUE	0:00:00	0:00:00
38	10:03:34	0	3,331,813.31	1,096,119.61		10:03:34	0	3,331,799.99	1,096,119.33		13.33	TRUE	0:00:00	0:00:00
39	10:03:36	0	3,331,813.16	1,096,119.43		10:03:36	0	3,331,799.84	1,096,119.51		13.32	TRUE	0:00:00	0:00:00
40	10:03:38	0	3,331,812.72	1,096,119.44		10:03:38	0	3,331,799.84	1,096,119.51		12.88	TRUE	0:00:00	0:00:00
41	10:03:40	0	3,331,812.57	1,096,119.26		10:03:40	0	3,331,799.84	1,096,119.70		12.73	TRUE	0:00:00	0:00:00
42	10:03:42	0	3,331,812.56	1,096,119.07		10:03:42	0	3,331,799.84	1,096,119.70		12.73	TRUE	0:00:00	0:00:00
43	10:03:44	0	3,331,812.26	1,096,118.89		10:03:44	0	3,331,799.84	1,096,119.70		12.44	TRUE	0:00:00	0:00:00
44	10:03:46	0	3,331,812.11	1,096,118.71		10:03:46	0	3,331,799.84	1,096,119.70		12.31	TRUE	0:00:00	0:00:00
45	10:03:48	0	3,331,812.11	1,096,118.53		10:03:48	0	3,331,799.84	1,096,119.51		12.31	TRUE	0:00:00	0:00:00

Figure 3.9: Time Matched EXTIM Data on Export Sheet

The criteria associated with the EXTIM are the same as that associated with the ADTIM. However, the speed and distance criteria are evaluated simultaneously. The status of the truck is then checked to identify the key records.

3.5.3. FBTIM

Not all earthmoving activities are associated with the truck stopped in a certain area, but rather can be associated with the truck traveling in or through a certain location. The haul and return activities, or travel times for roadway segments, are examples of truck activities for which the truck is moving. The Fixed Boundary TIM (FBTIM) was developed to identify the times at which a truck crosses user-defined boundaries, thus entering or exiting the defined area. Establishing proper boundaries will allow the user to calculate transit or travel times.

Two criteria need to be satisfied for a record to be considered an Enter or Exit time by the FBTIM:

1. Position of the truck is within the defined boundaries
2. The previous record was outside of the area for Enter times, or the next record is outside of the area for Exit times.

The FBTIM first evaluates position to determine whether the recorded position is within defined boundaries. Then it determines whether the record represents an Enter or Exit time by examining the status of the previous or next record.

The user defines the boundaries on the FBTIM Input dialog box, shown as Figure 3.10, that appears after the FBTIM option is selected from the user interface. The user may define as many as four areas by providing a minimum and maximum value for both northing and easting coordinates for each area. A brief textual description to be used to label the results can also be provided.

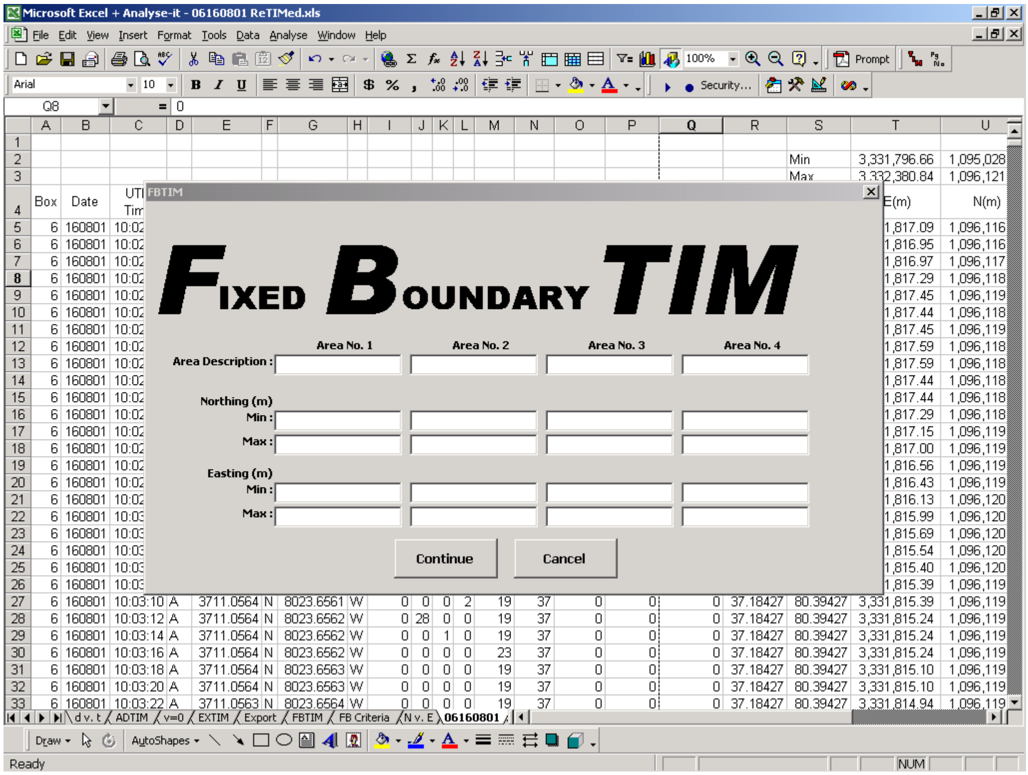
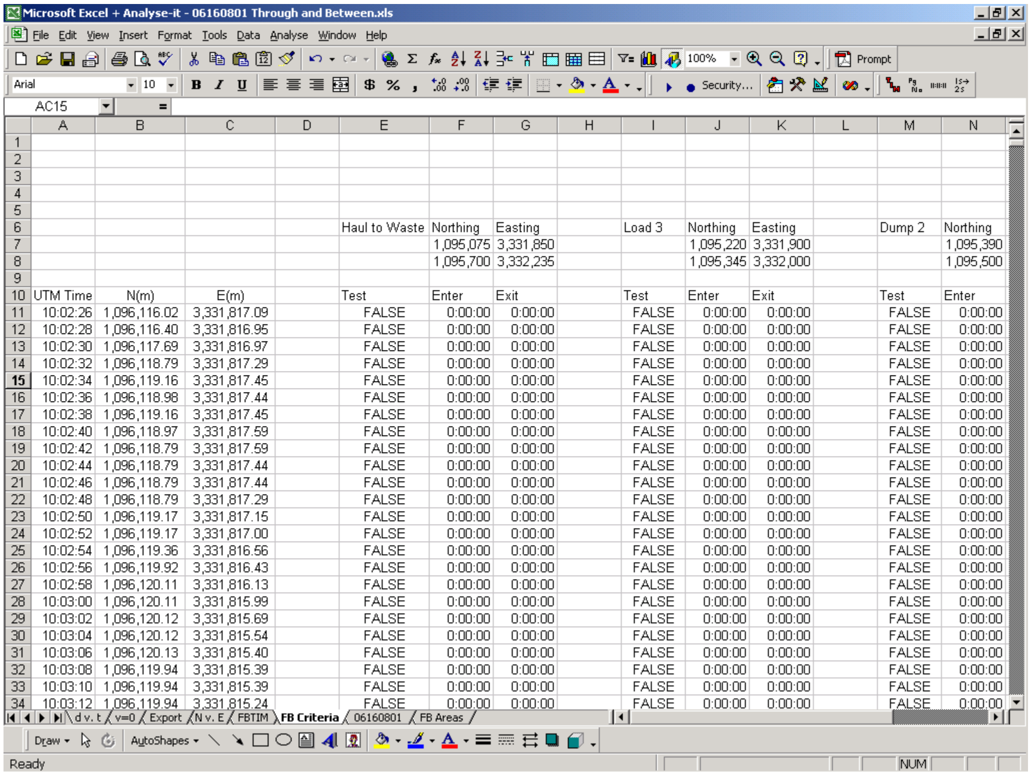


Figure 3.10: FBTIM Input Dialog Box

The criteria are evaluated on a FB Criteria sheet. For each area defined by the user, each recorded position is compared to the provided minimum and maximum coordinate values, and each record marked appropriately. The truck status is then reviewed again to identify those records corresponding to entry and exit times, and those key records marked with the time at which the position was recorded. This can be seen in Figure 3.11.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2														
3														
4														
5														
6					Haul to Waste	Northing	Easting		Load 3	Northing	Easting		Dump 2	Northing
7						1,095,075	3,331,850			1,095,220	3,331,900			1,095,390
8						1,095,700	3,332,235			1,095,345	3,332,000			1,095,500
9														
10	UTM Time	N(m)	E(m)		Test	Enter	Exit		Test	Enter	Exit		Test	Enter
11	10:02:26	1,096,116.02	3,331,817.09		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
12	10:02:28	1,096,116.40	3,331,816.95		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
13	10:02:30	1,096,117.69	3,331,816.97		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
14	10:02:32	1,096,118.79	3,331,817.29		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
15	10:02:34	1,096,119.16	3,331,817.45		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
16	10:02:36	1,096,118.98	3,331,817.44		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
17	10:02:38	1,096,119.16	3,331,817.45		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
18	10:02:40	1,096,118.97	3,331,817.59		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
19	10:02:42	1,096,118.79	3,331,817.59		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
20	10:02:44	1,096,118.79	3,331,817.44		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
21	10:02:46	1,096,118.79	3,331,817.44		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
22	10:02:48	1,096,118.79	3,331,817.29		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
23	10:02:50	1,096,119.17	3,331,817.15		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
24	10:02:52	1,096,119.17	3,331,817.00		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
25	10:02:54	1,096,119.36	3,331,816.56		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
26	10:02:56	1,096,119.92	3,331,816.43		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
27	10:02:58	1,096,120.11	3,331,816.13		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
28	10:03:00	1,096,120.11	3,331,815.99		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
29	10:03:02	1,096,120.12	3,331,815.69		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
30	10:03:04	1,096,120.12	3,331,815.54		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
31	10:03:06	1,096,120.13	3,331,815.40		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
32	10:03:08	1,096,119.94	3,331,815.39		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
33	10:03:10	1,096,119.94	3,331,815.39		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00
34	10:03:12	1,096,119.94	3,331,815.24		FALSE	0:00:00	0:00:00		FALSE	0:00:00	0:00:00		FALSE	0:00:00

Figure 3.11: FBTIM Data on FB Criteria Sheet

3.6. Case Study

The system was field tested by collecting data from three Caterpillar 771 quarry trucks and a 988 rubber-tired loader used to move material at a local aggregate quarry. Shot rock was hauled to crushers, while overburden material was hauled to an on-site waste area. Three load areas and three dump areas were used. The load areas were active rock faces, while the dump areas consisted of two rock crusher bins and the waste area.

The system was installed on the four pieces of equipment and wired to draw power from the equipment battery through the ignition switch. Wiring through the ignition switch allowed the system to operate only when the equipment was operating, thus not collecting extraneous and useless data. The equipment battery provided a reliable power source that allowed the system to record over the full workday.

Data was collected over several days and was recorded at a two-second time interval. The equipment operated approximately 11 hours each day, thereby producing approximately 19,800 records daily. This large volume of recorded data resulted in a very clear need for automated data reduction methods.

The system simply and efficiently recorded position and speed data at two-second intervals. The challenge then became how to present the data in a graphical format for a preliminary review and then to identify and select the relatively small number of key records needed to calculate activity times. The remainder of this paper addresses this challenge and describes the methods developed.

This case study provides an understanding of the TIMs and how they work. A preliminary review is necessary for the user to evaluate the data to determine in which activities the truck was engaged and the location of each activity. Graphical data representations and knowledge of the project are the most appropriate sources from which to gather this information.

With knowledge of the project and a plan view of the data, the areas in which earthmoving activities were performed can be determined. These areas and the associated activities are identified for the sample data in Figure 3.12. The truck from which the sample data was obtained moved material from two separate loading areas to two separate dumping areas. The truck hauled material from loading areas 1 and 3 and dumped the material at dump areas 2 and 3.

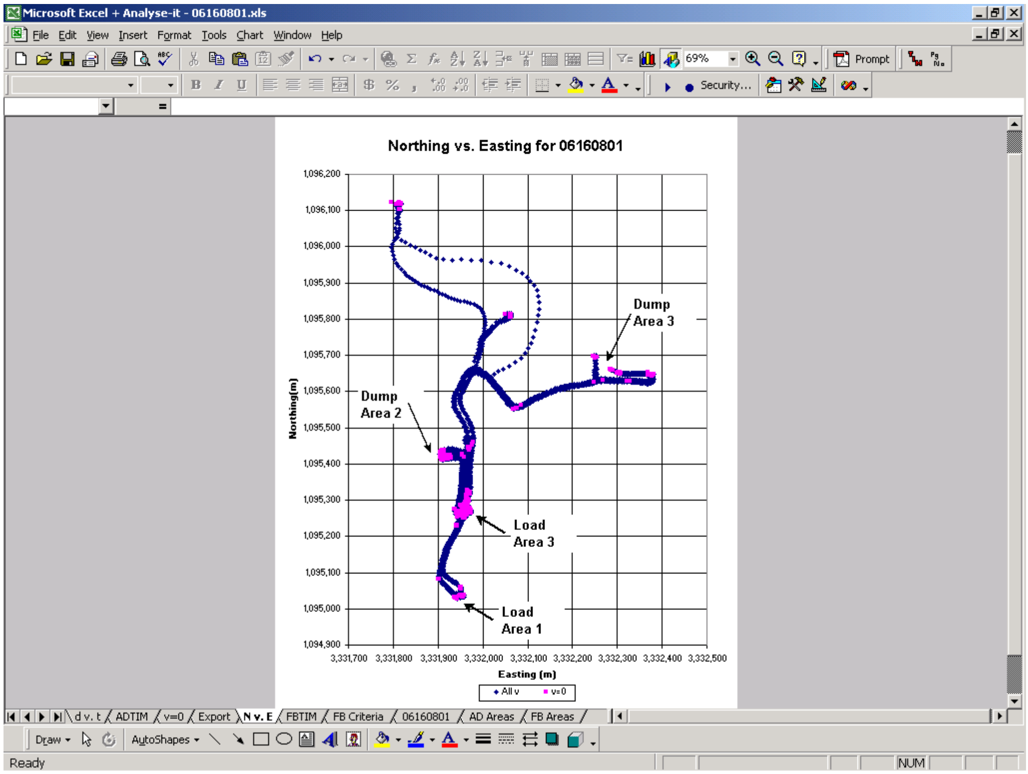


Figure 3.12: Plan View with Noted Activity Areas

Further information can be gathered from the plot of position versus time, shown in Figure 3.5. It is also important to note that the load area coordinates used to generate this plot were the approximate center of Load Area 1, as shown in Figure 3.12. It can be seen from Figure 3.5 that the truck was engaged in three separate operations on the day in which the sample data was recorded, the first two of which are of importance. A single cycle of the third operation was performed, and therefore is of little importance. Approximately 5 loads were hauled from Load Area 1 a distance of approximately 800 meters, which must be Dump Area 3 as Dump Area 2 is much less than 800 meters from the loading area, between approximately 10:20 and 11:30. Following this first operation was a second hauling operation, during which approximately 36 loads were hauled a distance of approximately 100 meters. Based on the information from the plan view plot, this operation must have loaded at Loading Area 3 and hauled to Dump Area 2. From Figure 3.5 it is determined that this second operation was performed from approximately 10:30 to 20:20.

Therefore, two load and haul operations consisting of a four activities each have been identified from the preliminary review. Each operation consists of four activities: load, haul, dump, and return. The activity naming convention adopted for this work is the activity name followed by the associated areas. For example, loading in area 3 is labeled Load 3, while hauling from load area 1 to dump area 2 is labeled Haul 12. The truck activities identified from the sample data are Load 1, Haul 13, Dump 3, and Return 31 for the first operation, and Load 3, Haul 32, Dump 2, and Return 23 for the second operation. The TIMs can be applied to the data to identify the beginning and end times for each of the identified activities.

The dump activities occur in fixed locations and with zero speed. Therefore, the ADTIM will be used to identify the records key to those activities. The ADTIM areas to be applied to the data are shown in Figure 3.13.

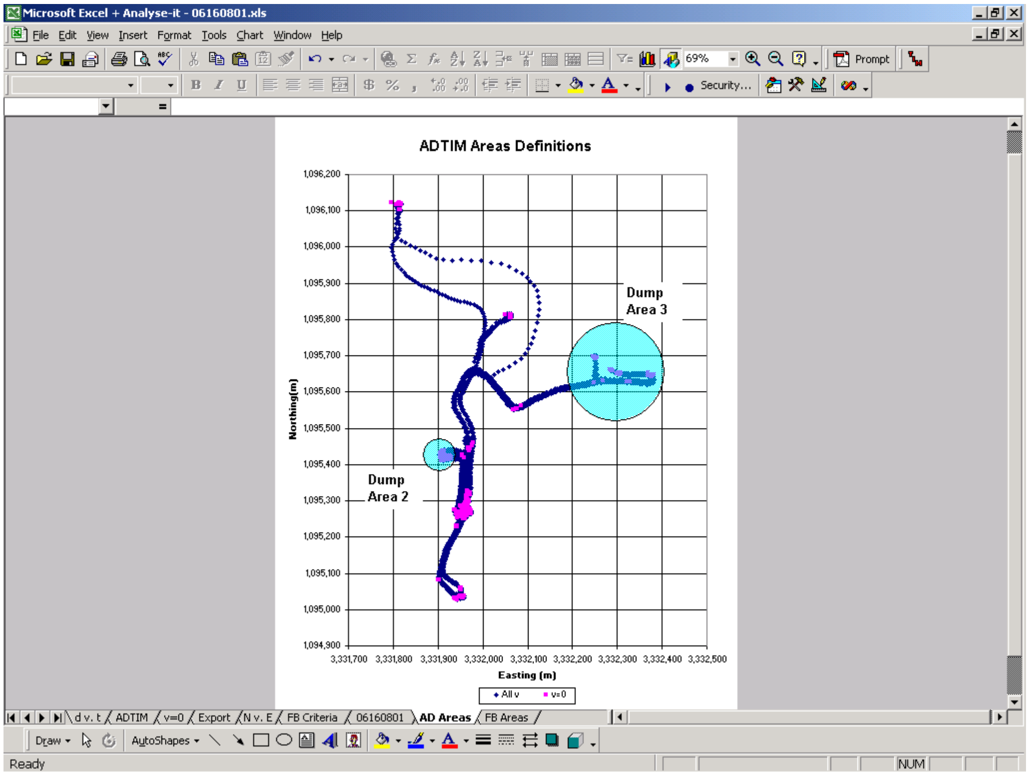
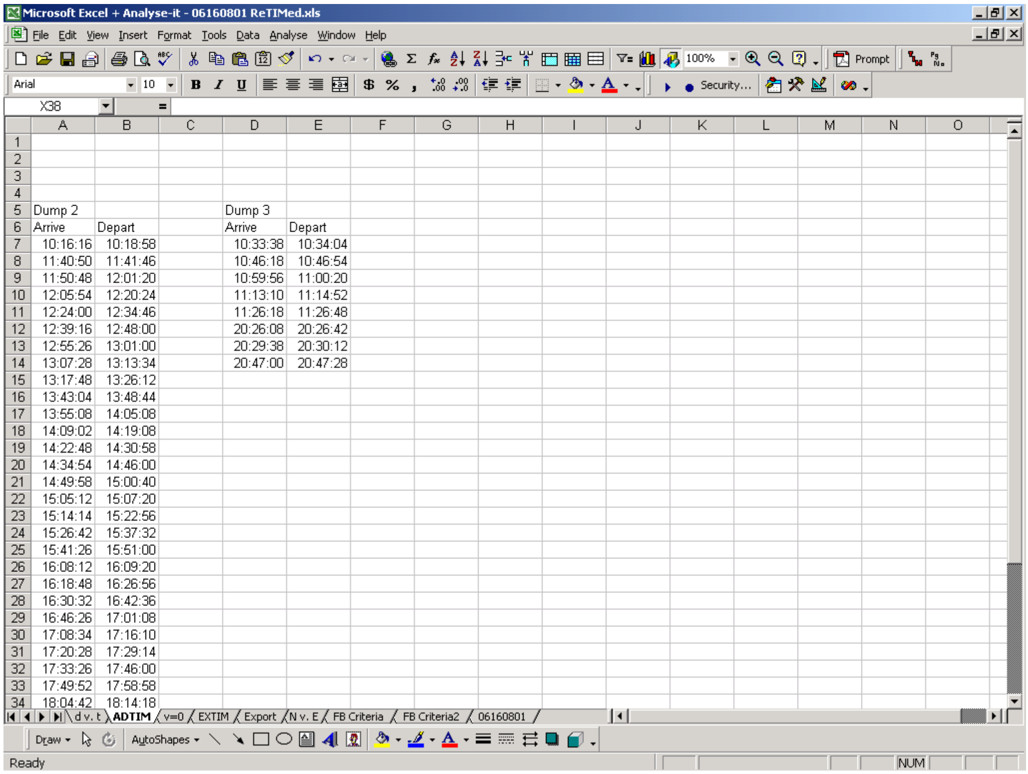


Figure 3.13: ADTIM Areas

The user can determine the central coordinates and radial distance of each ADTIM area from the plan view of the recorded data. Once the input data is provided, the ADTIM is executed and the results are placed onto an ADTIM sheet. The ADTIM results from the case study are provided as Figure 3.14.



	A	B	C	D	E
1					
2					
3					
4					
5	Dump 2			Dump 3	
6	Arrive	Depart		Arrive	Depart
7	10:16:16	10:18:58		10:33:38	10:34:04
8	11:40:50	11:41:46		10:46:18	10:46:54
9	11:50:48	12:01:20		10:59:56	11:00:20
10	12:05:54	12:20:24		11:13:10	11:14:52
11	12:24:00	12:34:46		11:26:18	11:26:48
12	12:39:16	12:48:00		20:26:08	20:26:42
13	12:55:26	13:01:00		20:29:38	20:30:12
14	13:07:28	13:13:34		20:47:00	20:47:28
15	13:17:48	13:26:12			
16	13:43:04	13:48:44			
17	13:55:08	14:05:08			
18	14:09:02	14:19:08			
19	14:22:48	14:30:58			
20	14:34:54	14:46:00			
21	14:49:58	15:00:40			
22	15:05:12	15:07:20			
23	15:14:14	15:22:56			
24	15:26:42	15:37:32			
25	15:41:26	15:51:00			
26	16:08:12	16:09:20			
27	16:18:48	16:26:56			
28	16:30:32	16:42:36			
29	16:46:26	17:01:08			
30	17:08:34	17:16:10			
31	17:20:28	17:29:14			
32	17:33:26	17:46:00			
33	17:49:52	17:58:58			
34	18:04:42	18:14:18			

Figure 3.14: ADTIM Results

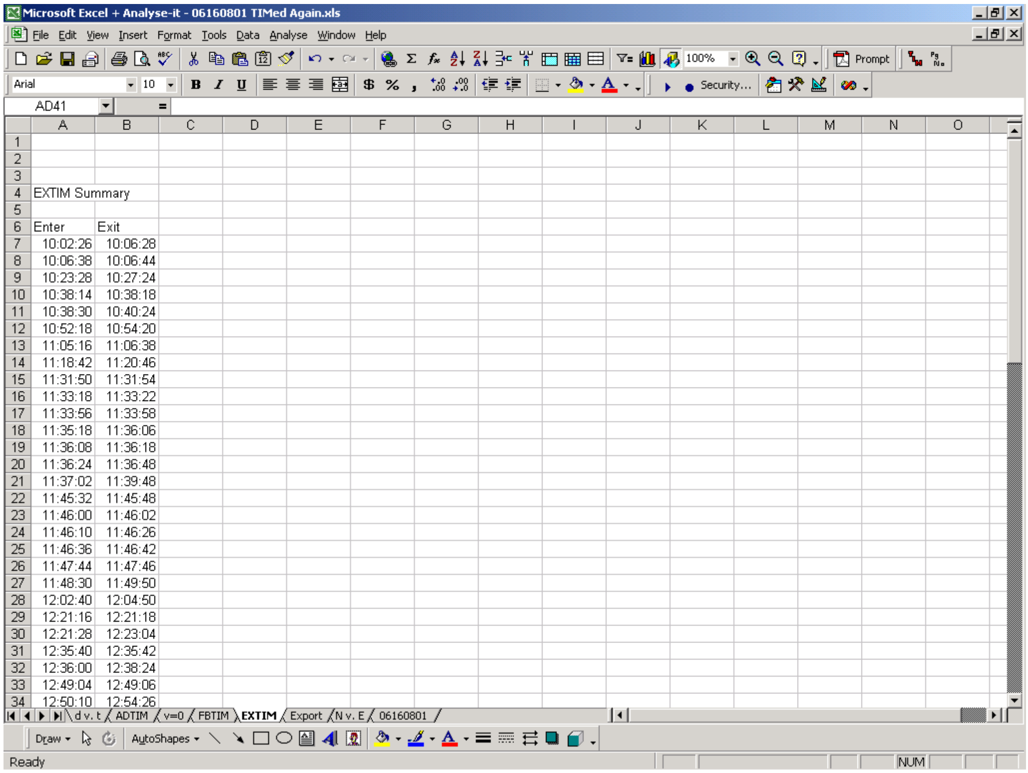
A secondary review is necessary to compare the results with those of the preliminary review. This will verify the ADTIM process and ensure that the total instances and timeframe of the associated activities concur with the results of the preliminary review.

It was determined during the preliminary review that the truck initially hauled approximately five loads from Load Area 1 to Dump Area 3. However, the ADTIM results indicate that eight loads were hauled throughout the day. On further inspection, it can be seen that five loads were hauled prior to 11:30, and the remaining three loads nearly nine hours later. This timeframe corresponds to that of the preliminary review. Thus the five results are considered as the start and end times of the Dump 3 activity associated with the first operation, and the additional arrive and depart times neglected.

The results also indicate that the truck was in the Dump 2 area at a time well before the time identified during the preliminary review. Neglecting this result yields 36 instances of the Dump 2 activity that occur during the identified timeframe. Therefore, the subject record was neglected.

A total of approximately 18,850 data records were collected and analyzed. The set of records where speed equaled zero numbered approximately 13,900, a reduction of approximately 25 percent. Further reduction by ADTIM identified 90 key records that mark the start and end of activities. These 90 records represent less than one-half of one percent of the total records collected; the collected data was reduced by more than 99.5 percent.

The EXTIM will be used to identify the records key to the load activities, as the area in which the truck was loaded changed over time. The data file recorded on the loader was used, and a critical distance of 20 meters was used. This distance was selected based on previous experience with data from the same or similar operations. The results of the EXTIM are provided as Figure 3.15.



EXTIM Summary

Enter	Exit
10:02:26	10:06:28
10:06:38	10:06:44
10:23:28	10:27:24
10:38:14	10:38:18
10:38:30	10:40:24
10:52:18	10:54:20
11:05:16	11:06:38
11:18:42	11:20:46
11:31:50	11:31:54
11:33:18	11:33:22
11:33:56	11:33:58
11:35:18	11:36:06
11:36:08	11:36:18
11:36:24	11:36:48
11:37:02	11:39:48
11:45:32	11:45:48
11:46:00	11:46:02
11:46:10	11:46:26
11:46:36	11:46:42
11:47:44	11:47:46
11:48:30	11:49:50
12:02:40	12:04:50
12:21:16	12:21:18
12:21:28	12:23:04
12:35:40	12:35:42
12:36:00	12:38:24
12:49:04	12:49:06
12:50:10	12:54:26

Figure 3.15: EXTIM Results

It is also necessary to review the EXTIM results to identify and eliminate any extraneous results. Such results can be produced by actions occurring outside of timeframes identified during data evaluation and by trucks queuing at the loader at a distance less than the specified critical distance. Results outside of an operational timeframe are easily identified. Those produced by queuing can be identified as those preceding another result at approximate same time. Examples of both can be seen in the EXTIM results from the sample data, presented as Figure 3.15.

The first of the operations identified began shortly before 10:30. The first two EXTIM results are from approximately 10:05, while the third result is at approximately 10:25. Therefore, the first two results are not part of the identified operation and should be neglected.

The situation produced by queuing can be seen in the fourth and fifth EXTIM results. The fourth result indicates a four second activity duration and is followed by the fifth result only 12 seconds later. Therefore, the fourth result should be neglected, and the fifth result considered valid. Evaluating the results in this manner produces a total of 44 load activity instances. This matches very well with the approximately 42 instances identified during data evaluation.

EXTIM identified 150 records that mark the start and stop of the load activities. Further review and reduction by the user of the identified records resulted in 84 key records. This level of reduction from 18,850 to 84 records is very similar to that of the ADTIM and represents greater than 99.5 percent reduction.

FBTIM will be used to identify the records key to the haul and return activities. The duration of the haul and return activities can be characterized as the time required for the truck to travel through or between FBTIM areas. For illustration purposes, the FBTIM areas will be applied in both manners. The FBTIM areas are shown in Figure 3.16.

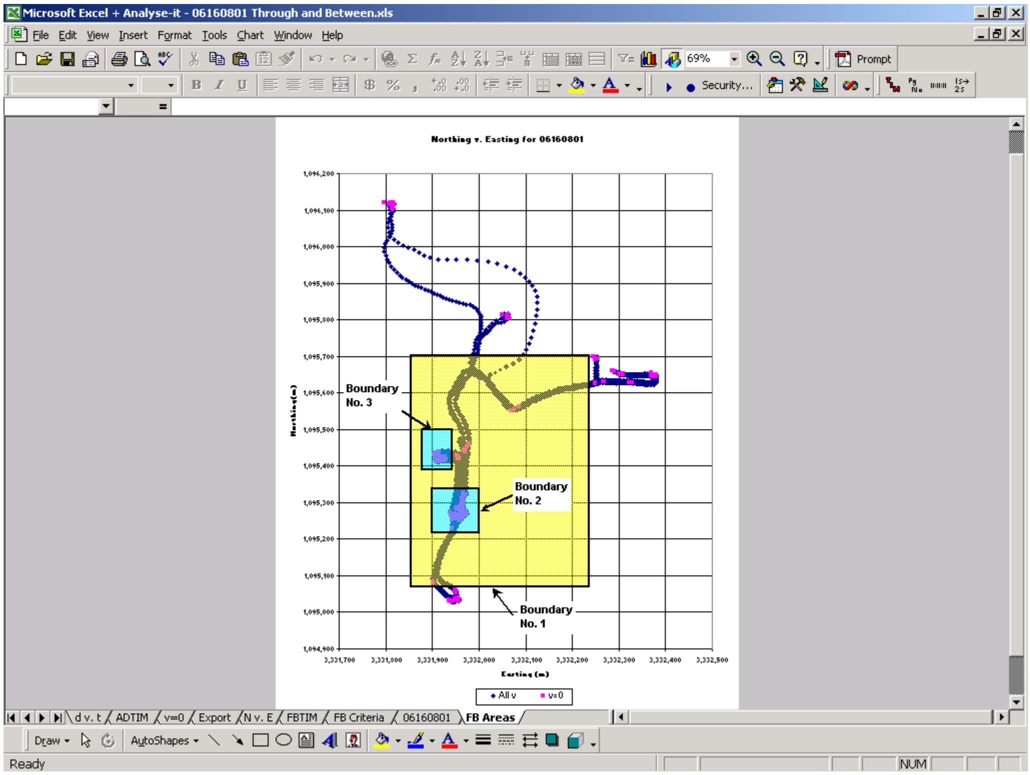
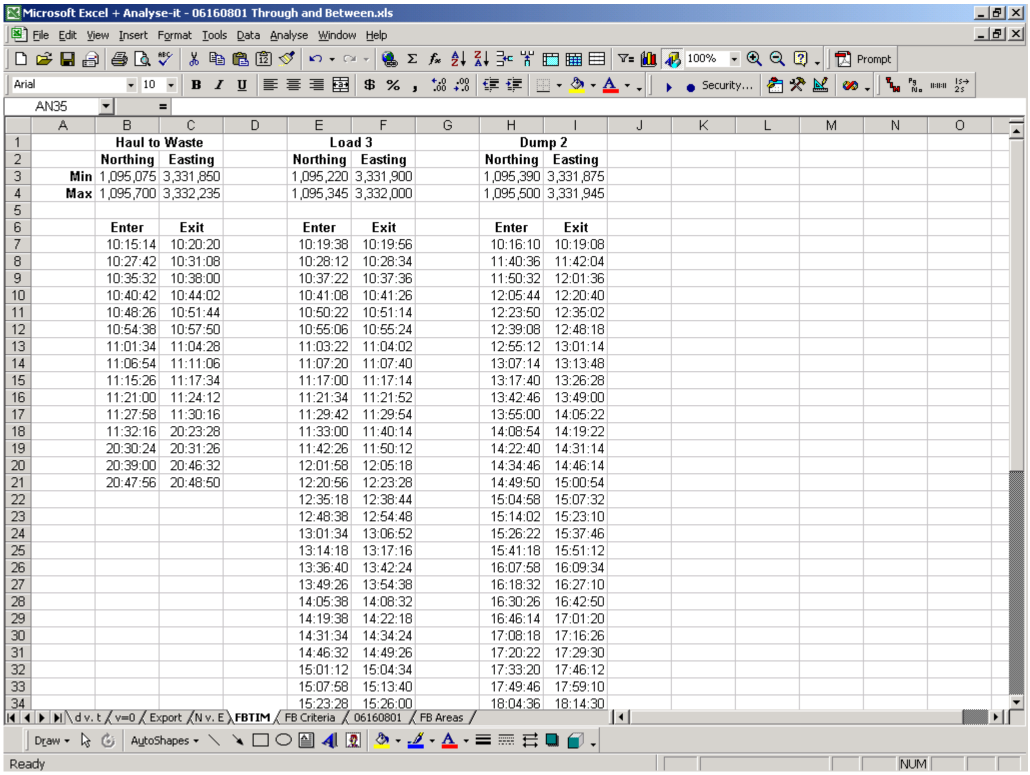


Figure 3.16: FBTIM Boundaries/Areas

Further manual reduction of the data resulting from the TIMs may be necessary to produce a set of critical records. Overlapping areas can be defined for the FBTIM and produce results that are both accurate and useful. It is important to understand that the results produced may also overlap and the appropriate overlapping results neglected.

The truck traveled through the area defined by Boundary No. 1 during the first operation, see Figure 3.16. As it passed through the area, it also traveled through the area defined by Boundary No. 2. Because the areas overlap, the FBTIM results will overlap as well. This can be seen in the FBTIM results presented as Figure 3.17.



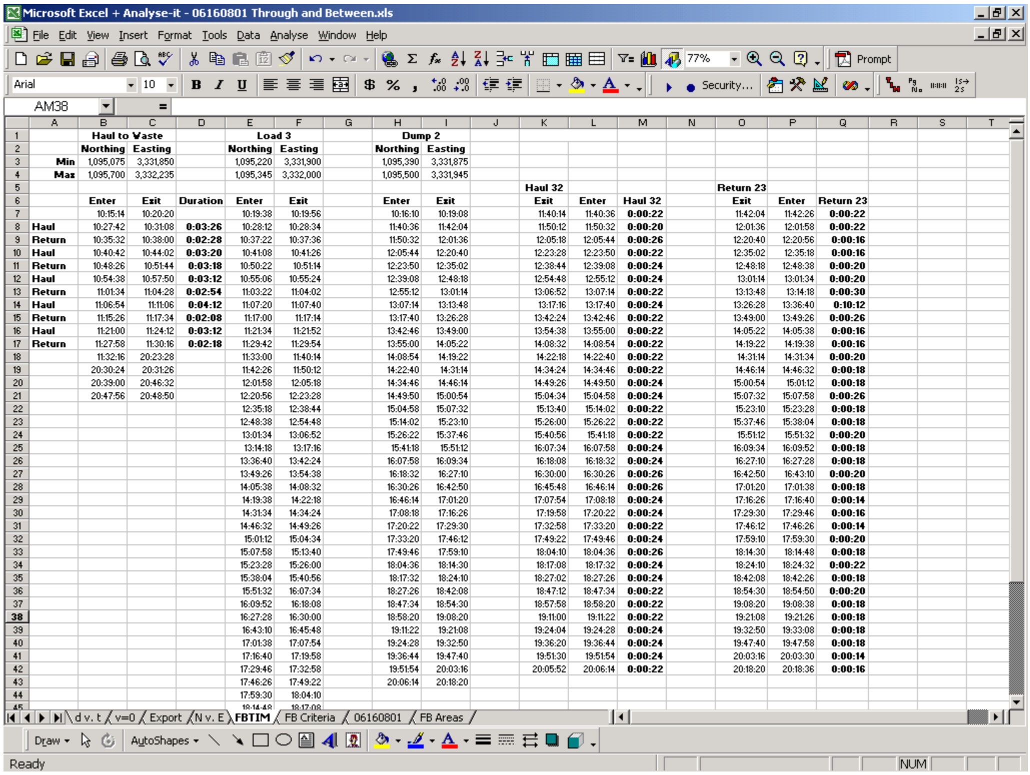
	A	B	C	D	E	F	G	H	I
1		**Haul to Waste**			**Load 3**			**Dump 2**	
2		**Northing**	**Easting**		**Northing**	**Easting**		**Northing**	**Easting**
3	**Min**	1,095,075	3,331,850		1,095,220	3,331,900		1,095,390	3,331,875
4	**Max**	1,095,700	3,332,235		1,095,345	3,332,000		1,095,500	3,331,945
5									
6		**Enter**	**Exit**		**Enter**	**Exit**		**Enter**	**Exit**
7		10:15:14	10:20:20		10:19:38	10:19:56		10:16:10	10:19:08
8		10:27:42	10:31:08		10:28:12	10:28:34		11:40:36	11:42:04
9		10:35:32	10:38:00		10:37:22	10:37:36		11:50:32	12:01:36
10		10:40:42	10:44:02		10:41:08	10:41:26		12:05:44	12:20:40
11		10:48:26	10:51:44		10:50:22	10:51:14		12:23:50	12:35:02
12		10:54:38	10:57:50		10:55:06	10:55:24		12:39:08	12:48:18
13		11:01:34	11:04:28		11:03:22	11:04:02		12:55:12	13:01:14
14		11:06:54	11:11:06		11:07:20	11:07:40		13:07:14	13:13:48
15		11:15:26	11:17:34		11:17:00	11:17:14		13:17:40	13:26:28
16		11:21:00	11:24:12		11:21:34	11:21:52		13:42:46	13:49:00
17		11:27:58	11:30:16		11:29:42	11:29:54		13:55:00	14:05:22
18		11:32:16	20:23:28		11:33:00	11:40:14		14:08:54	14:19:22
19		20:30:24	20:31:26		11:42:26	11:50:12		14:22:40	14:31:14
20		20:39:00	20:46:32		12:01:58	12:05:18		14:34:46	14:46:14
21		20:47:56	20:48:50		12:20:56	12:23:28		14:49:50	15:00:54
22					12:35:18	12:38:44		15:04:58	15:07:32
23					12:48:38	12:54:48		15:14:02	15:23:10
24					13:01:34	13:06:52		15:26:22	15:37:46
25					13:14:18	13:17:16		15:41:18	15:51:12
26					13:36:40	13:42:24		16:07:58	16:09:34
27					13:49:26	13:54:38		16:18:32	16:27:10
28					14:05:38	14:08:32		16:30:26	16:42:50
29					14:19:38	14:22:18		16:46:14	17:01:20
30					14:31:34	14:34:24		17:08:18	17:16:26
31					14:46:32	14:49:26		17:20:22	17:29:30
32					15:01:12	15:04:34		17:33:20	17:46:12
33					15:07:58	15:13:40		17:49:46	17:59:10
34					15:23:28	15:26:00		18:04:36	18:14:30

Figure 3.17: FBTIM Results

The FBTIM results indicate that the truck entered and exited Load Area 3 during the timeframe associated with the first operation. These results are accurate, but are of little significance as Load Area 3 is not associated with the first operation. The truck did in fact travel through Load Area 3 on its way from Load Area 1 to Dump Area 3. Therefore, the results associated with Load Area 3 before the start of the second operation at approximately 11:30 should be neglected, and only those results within the timeframe identified by the preliminary review will be used.

When boundaries are defined such that transit time is the time required for the truck to travel through the area, the FBTIM results must be evaluated to distinguish between haul and return times. Boundary No. 1 defined an area through with the truck traveled, and the results can be seen in columns B and C in Figure 3.18. Activity duration can be found by subtracting the entry time from the exit time, but this gives no indication of the activity. It can be seen from the graph of position versus time that an instance of the Haul 13 activity began at approximately 10:27. This corresponds to the second entry time identified by the FBTIM. Therefore, these results can be associated with the haul activity and the next results associated with the return activity. The results will alternate between haul and return as the truck performs the operation. This can be seen in Figure 3.18.

To determine the duration of the Haul 32 activity, the time at which the truck exited Load Area 3 is subtracted from the time at which the truck entered Dump Area 2. The Dump Area 2 exit time is subtracted from the Load Area 3 entry time to determine the duration of the Return 23 duration. This can also be seen in Figure 3.18.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		**Haul to Vaste**			**Load 3**			**Dump 2**									
2		**Northing**	**Easting**		**Northing**	**Easting**		**Northing**	**Easting**								
3	**Min**	1,095,075	3,331,850		1,095,220	3,331,900		1,095,390	3,331,875								
4	**Max**	1,095,700	3,332,235		1,095,345	3,332,000		1,095,500	3,331,945								
5											**Haul 32**				**Return 23**		
6		**Enter**	**Exit**	**Duration**	**Enter**	**Exit**		**Enter**	**Exit**		**Exit**	**Enter**	**Haul 32**		**Exit**	**Enter**	**Return 23**
7		10:15:14	10:20:20		10:19:38	10:19:56		10:16:10	10:19:08		11:40:14	11:40:36	**0:00:22**		11:42:04	11:42:26	**0:00:22**
8	**Haul**	10:27:42	10:31:08	**0:03:26**	10:28:12	10:28:34		11:40:36	11:42:04		11:50:12	11:50:32	**0:00:20**		12:01:36	12:01:58	**0:00:22**
9	**Return**	10:35:32	10:38:00	**0:02:28**	10:37:22	10:37:36		11:50:32	12:01:36		12:05:18	12:05:44	**0:00:26**		12:20:40	12:20:56	**0:00:16**
10	**Haul**	10:40:42	10:44:02	**0:03:20**	10:41:08	10:41:26		12:05:44	12:20:40		12:23:28	12:23:50	**0:00:22**		12:35:02	12:35:18	**0:00:16**
11	**Return**	10:48:26	10:51:44	**0:03:18**	10:50:22	10:51:14		12:23:50	12:35:02		12:38:44	12:39:08	**0:00:24**		12:48:18	12:48:38	**0:00:20**
12	**Haul**	10:54:38	10:57:50	**0:03:12**	10:55:06	10:55:24		12:39:08	12:48:18		12:54:48	12:55:12	**0:00:24**		13:01:14	13:01:34	**0:00:20**
13	**Return**	11:01:34	11:04:28	**0:02:54**	11:03:22	11:04:02		12:55:12	13:01:14		13:06:52	13:07:14	**0:00:22**		13:13:48	13:14:18	**0:00:30**
14	**Haul**	11:06:54	11:11:06	**0:04:12**	11:07:20	11:07:40		13:07:14	13:13:48		13:17:16	13:17:40	**0:00:24**		13:26:28	13:36:40	**0:10:12**
15	**Return**	11:15:26	11:17:34	**0:02:08**	11:17:00	11:17:14		13:17:40	13:26:28		13:42:24	13:42:46	**0:00:22**		13:49:00	13:49:26	**0:00:26**
16	**Haul**	11:21:00	11:24:12	**0:03:12**	11:21:34	11:21:52		13:42:46	13:49:00		13:54:38	13:55:00	**0:00:22**		14:05:22	14:05:38	**0:00:16**
17	**Return**	11:27:58	11:30:16	**0:02:18**	11:29:42	11:29:54		13:55:00	14:05:22		14:08:32	14:08:54	**0:00:22**		14:19:22	14:19:38	**0:00:16**
18		11:32:16	20:23:28		11:33:00	11:40:14		14:08:54	14:19:22		14:22:18	14:22:40	**0:00:22**		14:31:14	14:31:34	**0:00:20**
19		20:30:24	20:31:26		11:42:26	11:50:12		14:22:40	14:31:14		14:34:24	14:34:46	**0:00:22**		14:46:14	14:46:32	**0:00:18**
20		20:39:00	20:46:32		12:01:58	12:05:18		14:34:46	14:46:14		14:49:26	14:49:50	**0:00:24**		15:00:54	15:01:12	**0:00:18**
21		20:47:56	20:48:50		12:20:56	12:23:28		14:49:50	15:00:54		15:04:34	15:04:58	**0:00:24**		15:07:32	15:07:58	**0:00:26**
22					12:35:18	12:38:44		15:04:58	15:07:32		15:13:40	15:14:02	**0:00:22**		15:23:10	15:23:28	**0:00:18**
23					12:48:38	12:54:48		15:14:02	15:23:10		15:26:00	15:26:22	**0:00:22**		15:37:46	15:38:04	**0:00:18**
24					13:01:34	13:06:52		15:26:22	15:37:46		15:40:56	15:41:18	**0:00:22**		15:51:12	15:51:32	**0:00:20**
25					13:14:18	13:17:16		15:41:18	15:51:12		16:07:34	16:07:58	**0:00:24**		16:09:34	16:09:52	**0:00:18**
26					13:36:40	13:42:24		16:07:58	16:09:34		16:18:08	16:18:32	**0:00:24**		16:27:10	16:27:28	**0:00:18**
27					13:49:26	13:54:38		16:18:32	16:27:10		16:30:00	16:30:26	**0:00:26**		16:42:50	16:43:10	**0:00:20**
28					14:05:38	14:08:32		16:30:26	16:42:50		16:45:48	16:46:14	**0:00:26**		17:01:20	17:01:38	**0:00:18**
29					14:19:38	14:22:18		16:46:14	17:01:20		17:07:54	17:08:18	**0:00:24**		17:16:26	17:16:40	**0:00:14**
30					14:31:34	14:34:24		17:08:18	17:16:26		17:19:58	17:20:22	**0:00:24**		17:29:30	17:29:46	**0:00:16**
31					14:46:32	14:49:26		17:20:22	17:29:30		17:32:58	17:33:20	**0:00:22**		17:46:12	17:46:26	**0:00:14**
32					15:01:12	15:04:34		17:33:20	17:46:12		17:49:22	17:49:46	**0:00:24**		17:59:10	17:59:30	**0:00:20**
33					15:07:58	15:13:40		17:49:46	17:59:10		18:04:10	18:04:36	**0:00:26**		18:14:30	18:14:48	**0:00:18**
34					15:23:28	15:26:00		18:04:36	18:14:30		18:17:08	18:17:32	**0:00:24**		18:24:10	18:24:32	**0:00:22**
35					15:38:04	15:40:56		18:17:32	18:24:10		18:27:02	18:27:26	**0:00:24**		18:42:08	18:42:26	**0:00:18**
36					15:51:32	16:07:34		18:27:26	18:42:08		18:47:12	18:47:34	**0:00:22**		18:54:30	18:54:50	**0:00:20**
37					16:09:52	16:18:08		18:47:34	18:54:30		18:57:58	18:58:20	**0:00:22**		19:08:20	19:08:38	**0:00:18**
38					16:27:28	16:30:00		18:58:20	19:08:20		19:11:00	19:11:22	**0:00:22**		19:21:08	19:21:26	**0:00:18**
39					16:43:10	16:45:48		19:11:22	19:21:08		19:24:04	19:24:28	**0:00:24**		19:32:50	19:33:08	**0:00:18**
40					17:01:38	17:07:54		19:24:28	19:32:50		19:36:20	19:36:44	**0:00:24**		19:47:40	19:47:58	**0:00:18**
41					17:16:40	17:19:58		19:36:44	19:47:40		19:51:30	19:51:54	**0:00:24**		20:03:16	20:03:30	**0:00:14**
42					17:29:46	17:32:58		19:51:54	20:03:16		20:05:52	20:06:14	**0:00:22**		20:18:20	20:18:36	**0:00:16**
43					17:46:26	17:49:22		20:06:14	20:18:20								
44					17:59:30	18:04:10											
45					18:14:48	18:17:08											

Figure 3.18: Activity Durations Determined from FBTIM Results

Evaluating the FBTIM results in this manner produced data for five haul and return activities associated with the first identified operation, and data for 36 haul and 35 return instances for the second operation.

Data reduction by FBTIM reduced the number of records from 18,850 to 202. Further review and reduction by the user resulted in the identification of 164 key records that mark the start and stop of activities. These 164 records represent less than one percent of the total number of data records collected; the data was reduced by greater than 99 percent.

3.7. Data Analysis

The results produced can be used to calculate the collection of durations for each activity over the workday. These collections of durations can be used to provide insight into both the activity and the parent operation as a whole. They can also be used to develop best-fit statistical functions for use in computer simulation and modeling techniques.

For the sake of brevity, this will be demonstrated only for the Dump 2 activity results produced by the ADTIM. The durations calculated based on the ADTIM results are shown in Table 3. 3.

Table 3.3: ADTIM Dump 2 Activity Durations

Dump 2							
Instance	**Arrive**	**Depart**	**Duration**	**Instance**	**Arrive**	**Depart**	**Duration**
1	11:40:50	11:41:46	00:00:56	19	16:08:12	16:09:20	00:01:08
2	11:50:48	12:01:20	00:10:32	20	16:18:48	16:26:56	00:08:08
3	12:05:54	12:20:24	00:14:30	21	16:30:32	16:42:36	00:12:04
4	12:24:00	12:34:46	00:10:46	22	16:46:26	17:01:08	00:14:42
5	12:39:16	12:48:00	00:08:44	23	17:08:34	17:16:10	00:07:36
6	12:55:26	13:01:00	00:05:34	24	17:20:28	17:29:14	00:08:46
7	13:07:28	13:13:34	00:06:06	25	17:33:26	17:46:00	00:12:34
8	13:17:48	13:26:12	00:08:24	26	17:49:52	17:58:58	00:09:06
9	13:43:04	13:48:44	00:05:40	27	18:04:42	18:14:18	00:09:36
10	13:55:08	14:05:08	00:10:00	28	18:17:48	18:23:56	00:06:08
11	14:09:02	14:19:08	00:10:06	29	18:27:34	18:41:52	00:14:18
12	14:22:48	14:30:58	00:08:10	30	18:47:48	18:54:16	00:06:28
13	14:34:54	14:46:00	00:11:06	31	18:58:28	19:08:06	00:09:38
14	14:49:58	15:00:40	00:10:42	32	19:11:30	19:20:54	00:09:24
15	15:05:12	15:07:20	00:02:08	33	19:24:36	19:32:36	00:08:00
16	15:14:14	15:22:56	00:08:42	34	19:36:52	19:47:28	00:10:36
17	15:26:42	15:37:32	00:10:50	35	19:52:00	20:03:02	00:11:02
18	15:41:26	15:51:00	00:09:34	36	20:06:22	20:18:04	00:11:42

As a collection of numbers in a table, the durations are not very meaningful. Again, graphical representations of the data portray a clearer image. A histogram of the duration, as shown in Figure 3.19, indicates the shape of the statistical distribution underlying the duration data.

It can be seen from the histogram that the majority of durations were between six and 12 minutes. It is also evident that the histogram is skewed to the left. It does not require a Caterpillar 771 quarry truck six minutes to raise the dump body and dump the load. In fact, from Table 3.3 it can be seen that the dump activity can be completed in less than one minute. The long average duration and the left skewed histogram indicate that truck is not the constraining equipment in the Dump 2 activity. Also indicated is the fact that the constraining equipment, a rock crusher into which the truck dumps, increases the duration many fold over the capability of the truck.

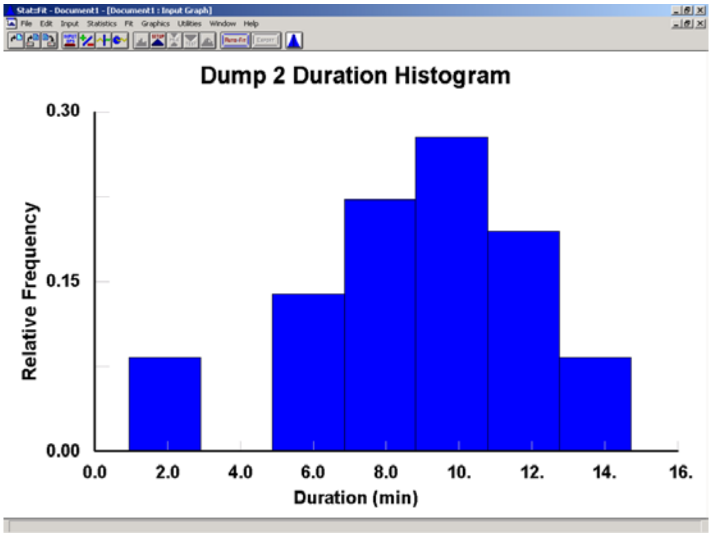
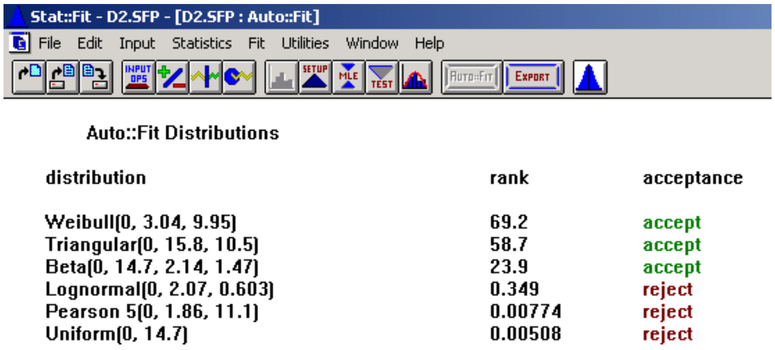


Figure 3.19: Dump 2 Duration Histogram

The collection of activity durations can be further used to determine statistical functions that fit the sample data within statistical limits. The Stat::Fit program accepts the sample data as input, estimates the parameters of several potentially fitting statistical functions, evaluates the appropriateness of the function through goodness-of-fit tests, and provides a recommendation to the user regarding each function. Such functions can be used to model and simulate the operation.

Modeling and simulation techniques can be used to analyze construction operations, and use as input stochastic definitions of activity durations. To determine the duration of each activity instance, the definitions are randomly sampled during simulation. Therefore, the simulation results will most accurately reflect the actual operation when the duration definitions are based on actual duration measures.

The 36 durations calculated for the Dump 2 activity were input into Stat::Fit and used to determine suitable statistical functions. The parameters for six potential functions were estimated using the maximum likelihood method, and the functions evaluated for goodness-of-fit with chi-squared, Kolmogorov-Smirnov (KS), and Anderson-Darling (AD) tests. A recommendation as to whether the user should accept or reject each function is based on the goodness-of-fit tests results and is provided. The potential functions and recommendations are provided as Figure 3.20.



Auto::Fit Distributions

distribution	rank	acceptance
Weibull(0, 3.04, 9.95)	69.2	accept
Triangular(0, 15.8, 10.5)	58.7	accept
Beta(0, 14.7, 2.14, 1.47)	23.9	accept
Lognormal(0, 2.07, 0.603)	0.349	reject
Pearson 5(0, 1.86, 11.1)	0.00774	reject
Uniform(0, 14.7)	0.00508	reject

Figure 3.20: Potential Statistical Functions for Dump 2 Activity

A graphical comparison of each distribution and the sample data is provided by Stat::Fit. Figure 3.21 is such a comparison for the best fitting distribution, Weibull(0, 3.04, 9.95). It can be seen that the function reasonably fits the data; this can be further reinforced with the results of the goodness-of-fit tests.

A level of significance is set by the user, and was set to 0.05 in this example. A p-value less than 0.05 resulting from a goodness-of-fit test would indicate that the function does not adequately fit the data and should be rejected. The p-values for the test of the Weibull distribution were 0.203, 0.335, and 0.289 for the chi-squared, KS, and AD tests respectively. All values were well above the 0.05 limit, indicating the distribution fits the data well and should not be rejected.

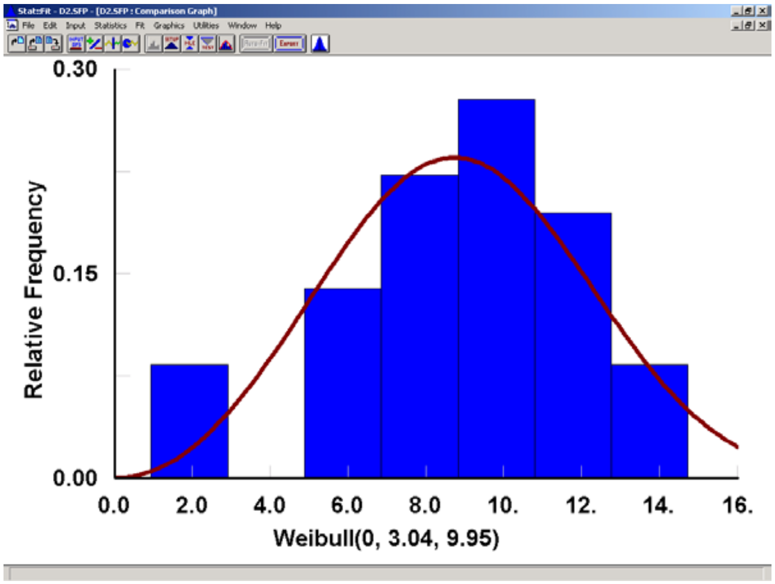


Figure 3.21: Comparison of Data and Fit Distribution

3.8. Conclusion

The tools available for time studies of construction operations have historically been manual in nature and limited to the field of view of the observer. The short-interval GPS system presented is an automated on-board instrumentation system not limited by field of view. The system places a virtual observer in the cab of the equipment, thereby keeping the equipment within the virtual field of view.

The system was presented in terms of the traditional data capture, preliminary review, data reduction, and data analysis steps of the time study process. The data captured by the system is position and speed data, and is autonomously recorded. The data is represented graphically to assist the user in reviewing and understanding the data. TIMs were presented and described to identify the key records associated with activities performed under the following conditions:

1. Time spent with speed equal to zero within a specified fixed location
2. Time spent with speed equal to zero within a fixed distance of a moving location
3. Time required to travel through or between fixed areas

It is necessary to use automated methods to identify the key records that mark the start and stop of activities due to the large volumes of data collected. A case study was presented to show the application of the system to an earthmoving operation. Each of the TIMS reduced the number of data records in the case study by more than 99 percent; collectively the TIMs reduced the number of records by more than 98 percent. It was postulated as to how the data can be used to provide insight into the activity and parent operation, as well for discrete event simulation.

3.9. Acknowledgement

The author wishes to acknowledge Terradon Communications for their assistance in developing the TIMs.

3.10. References

Ackroyd, N. (1998), "Earthworks Scheduling and Monitoring Using GPS," in the *Proceedings of the 1998 Trimble User Conference*, pp. 1-6.

Bjornsson, H. C., and P. Sagert, (1994), "Multimedia based system for analyzing production organizations," *Computing in Civil Engineering*, 2, pp1405-1411.

Cor, H. and J.C. Martinez (1999), "A Case Study in the Quantification of a Change in the Conditions of a Highway Construction Operation," in *Proceedings of the 1999 Winter Simulation Conference*, P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans (Ed.), Phoenix, AZ, pp1007-1009.

Frank, S. (2001), "Field Operations Analysis Laboratory – A Tutorial on the Review and Analysis Aspects of Operations Analysis," Project and Report, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Hildreth, J. and M. Vorster, (in review), *Development of a Short-Interval GPS Based Data Collection System*, Computer-Aided Civil and Infrastructure Engineering, Blackwell Publishers, Boston.

Ioannou, P. G. and J. C. Martinez, (1996a) "Comparison of Construction Alternatives Using Matched Simulation Experiments," *Journal of Construction Engineering and Management*, ASCE, 112(3), ASCE, Reston, VA.

Ioannou, P.G. and J.C. Martinez (1996b), “Simulation of Complex Construction Processes,” in *Proceedings of the 1996 Winter Simulation Conference*, J.M. Charnes, D.J. Morrice, D.T. Brunner, and J.J. Swain (Ed.), San Diego, CA, pp1329-1336.

Ioannou, P.G. (1999), “Construction of a Dam Embankment with Nonstationary Queues,” in *Proceedings of the 1999 Winter Simulation Conference*, P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans (Ed.), Phoenix, AZ, pp921-928.

Kannan, G. (1998), “Automated Data Collection for Earthwork Estimating,” in *Proceeding of Computing in Civil Engineering*, K. Wang (Ed.), 457-465.

Kannan, G. (1999), “A Methodology for the Development of a Production Experience Database for Earthmoving Operations using Automated Data Collection,” PhD Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Kannan, G. and M.C. Vorster (2000), “Development of an Experience Database for Truck Loading Operations,” *Journal of Construction Engineering and Management*, ASCE, 165(3), pp201-209.

Martinez, J. C. and P. G. Ioannou, (1995), “Advantages of the Activity Scanning Approach in the Modeling of Complex Construction Processes,” in *Proceedings of the 1995 Winter Simulation Conference*, Society for Computer Simulation, San Diego, CA.

Martinez, J. C., (1998), “EarthMover - Simulation Tool for Earthwork Planning,” in *Proceedings of the 1998 CIB W78 Conference*, Department of Construction Management and Economics, Royal Institute of Technology, Stockholm, Sweden.

Oglesby, C. H., H. W. Parker, and G. A. Howell, (1989), *Productivity Improvement in Construction*, McGraw Hill Companies, New York, NY.

Sangarayakul, B. (1998), “Use of Simulation to Analyze Block Manufacturing Methods.” MS Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Stem, J., (1990), *State Plane Coordinate System of 1983*, NOAA Manual NOS NGS 5, US Department of Commerce, Rockville, MD.

CHAPTER 4. AN INVESTIGATION INTO CORRELATIONS BETWEEN PERFORMANCE TIMES IN CONSTRUCTION OPERATIONS

4.1. Introduction

Modeling and simulating techniques have been applied to construction operations for many years, and have become an accepted means for analyzing complex operations. The process requires that a valid model of the operation be developed and used to test alternate scenarios. Data regarding operational variables is gathered and used to drive the model, and this is, to a large degree, determines the analyst's ability to produce results that accurately represent the operation. Descriptions of discrete event simulation applied to construction operations can be found in [Martinez and Ioannou 1995, Ioannou and Martinez 1996a, Ioannou and Martinez 1996b, Martinez 1998, Sangarayakul 1998, Cor and Martinez 1999, Ioannou 1999].

Traditionally, the statistical performance time techniques used to select data from a distribution of possible values assume no relationship between performance time and factors such as the immediately preceding performance time or time of day. Typically, the duration of each activity is selected independently from a distribution of likely performance times according to the probability of occurrence. Correlation is a statistical measure of the linear relationship between an independent and dependent variable that can be used to quantify the strength of such relationships [Benjamin and Cornell 1970]. It is known that the potential for relationships exist among construction data. Law and Kelton [2000] note that correlation between input random variables, such as performance times, should be considered during simulation where evident and present three methods for considering correlation: joint probability functions, marginal probability functions, and specification of the stochastic process.

Previous systems for collecting data from construction operations were not able to autonomously collect accurate performance time data over long periods of time. This made it difficult to analyze long-term phenomena like correlation.

This chapter will investigate the validity of the assumption that performance times can be independently and randomly selected from distributions according to the probability of occurrence by using short-interval GPS data to autonomously collect the large volumes of data needed to analyze the correlation phenomena. GPS data will be reduced to performance times for the haul and return activities of an earthmoving operation. The haul and return times will be used to:

1. investigate the correlation between performance times (i.e. haul and return times)
2. investigate the auto-correlation in performance times (e.g. haul time and previous haul time)
3. investigate the correlation between performance time and the time of day (e.g. haul time and the time of day at which the activity began)

Conclusions will be drawn regarding whether significant correlation exists in construction data and the validity of the independent performance time assumption.

4.2. Background

Research into construction activity durations with respect to simulation has focused on determining the appropriate probabilistic distribution [AbouRizk and Halpin 1990; AbouRizk and Halpin 1992a; AbouRizk and Halpin 1992b; Fente et al 1999; Fente et al. 2000; Maio et al. 2000]. The issue of correlation among or between performance times has not been addressed in the literature.

Correlated variables tend to vary together, either directly or inversely. The degree to which variables are linearly dependent can be measured by computing the correlation coefficient from an observed sample. Benjamin and Cornell [1970] note that it is important to know what the coefficient does and does not say about the joint behavior of the variables to properly interpret the correlation coefficient. Correlation coefficients range in value from 1 to –1, with a value of zero indicating no correlation. Values greater than zero indicate positive correlation, or large values of one variable tend to occur with large values of another, with a value of 1 indicating perfect positive correlation. Similarly, values less than zero indicates negative correlation, or large values of one variable tend to occur with small values of another, and a value of –1 indicates perfect negative correlation. A coefficient of small magnitude does not imply independence, but rather a weak relationship.

Scatter plots of the data are useful in identifying the nature of the relationship, be it linear or non-linear. Sheskin [1997] points out that correlation does not imply causation, the information used to quantify correlation does not provide for such a conclusion to be drawn.

Both parametric and non-parametric correlation tests are available, with parametric tests based on the assumption of normally distributed data. Therefore, it is necessary to determine the whether the normal distribution underlies the data. The Shapiro-Wilk test is currently the best available normality test [Miller 1986], and data to support this claim is provided by Shapiro et al. [1968]. Shapiro et al. compared eight normality tests and found the Shapiro-Wilk W statistic is sensitive to non-normality over a wide range of distributions. The power of the Shapiro-Wilk test was found to generally be as good as or better than that of other normality tests. The W statistic ranges in value from zero to one, where low values indicate non-normality [D'Agostino and Stephens 1986].

Correlation is typically measured by the Pearson product-moment correlation coefficient "r" when the data is normally distributed. The coefficient "r" is an estimate of the population correlation ρ, and is based on five assumptions:

1. the sample from which r is computed was randomly selected from the population
2. each variable is measured on a scale where equal differences in measurements corresponds to equal differences in the attribute measured
3. the two variables are bivariate normally distributed, which implies that for any value of X the Y variable will be normally distributed, and vice versa
4. the set of sample data exhibits homoscedasticity, which implies that the relation between variables is the same over the entire range of both variables
5. the data is non-autoregressive, or auto-correlation is not present [Sheskin 1997].

The “r” statistic follows a t-distribution with n-2 degrees of freedom, which can be used to evaluate the statistical significance of the computed r statistic. The null hypothesis that the population correlation is zero can be tested against any of three alternative hypotheses: that it is not zero, is greater than zero, or is less than zero.

When data is not normally distributed, either Spearman’s coefficient r_s or Kendall’s coefficient τ is typically calculated. Both methods are rank-order methods and estimate the population rank-order correlation ρ_s. Conover [1980] notes that the tests produce nearly identical results and that there is no strong reason to prefer one test to the other.

Spearman’s coefficient assesses the degree to which two variables are monotonically related [Sheskin 1997]. Tables of critical values of the r_s statistic are available for testing the null hypothesis of zero correlation, and evaluate the statistical significance of the computed value.

Kendall’s τ is similar to Spearman’s r_s, but based on slightly different logic [Sheskin 1997]. As a result, the two measures have differing underlying scales and thus will have differing when calculated from the same sample data. As a rule of thumb, τ will be less than r_s in magnitude. However, both are equally likely to detect statistically significant correlation. Tables of critical τ values are available for evaluating the statistical significance of computed values.

Sheskin [1997] notes that Spearman’s coefficient is more commonly used as a bivariate measure of correlation for ranked data, and provides two reasons for this. First, the computation of τ is more complex than that of r_s. Additionally, when the samples are drawn from normally distributed data, Spearman’s coefficient is a reasonably good estimation of Pearson’s r statistic, while Kendall’s τ is not.

4.3. Field Data Collection

Measures of haul and return times were obtained through a field study of a conventional load and haul earthmoving operation. The studied operation involved loading overburden soil and shot rock onto Volvo A25C articulated haulers, transporting the material approximately 450 meters (1,500 feet), and dumping it at a stockpile. The trucks traveled a well-established and maintained haul road.

GPS data was collected at a short time interval by instrumentation installed one of the trucks engaged in the studied operation. The CrossCheck AMPS system manufactured by Trimble Navigation, Limited, was used to record the GPS data at a two-second interval. The data collected in ASCII text format, and included time, horizontal position, and speed. The system was programmed to record the following data string:

AABBBBCDDDDDEEEFFFFFGGGGHHHHHIIIJJJKL

Where: AA is the event number
BBBB is the GPS week
C is an integer representing the day of the week
DDDDD is the current time Greenwich Mean Time (GMT) expressed in seconds since midnight
EEE.FFFFF represents degrees of latitude with north positive
GGGG.HHHHH represents degrees of longitude with east positive
III is vehicle speed in miles per hour
JJJ is the vehicle heading in degrees with zero at true north and increasing eastwardly
K is an integer representing the quality of the data recorded
L is an integer representing the age of the data

A sample of collected data is shown below.

```
>REV001065269090+3719328-0804034100000412;ID=0012;*5E<
>REV001065269092+3719328-0804034100000412;ID=0012;*5C<
>REV001065269094+3719328-0804034100000412;ID=0012;*5A<
```

The Fixed Boundary Time Identification Module (FBTIM) described in Chapter 3 was used to calculate the haul and return activities. The FBTIM was used to define a boundary around the haul road on which the haulers traveled and to identify the times at which the hauler entered and exited the haul road. The identified entry and exit times were used to calculate performance times for the haul and return activities. Performance times were reviewed and analyzed to determine the strength and statistical significance of correlation present within the data.

The data collected was reviewed and it was determined that the truck traveled the studied haul road on 11 of the 20 working days studied. The number of cycles performed by the hauler and for which data was gathered varied from eight on June 15 to 42 on June 21. Haul times were calculated for 214 cycles, while return times were calculated for 204 cycles. Summary statistics consisting of the mean, 95 percent confidence interval about the mean, and standard deviation were calculated on a daily basis for both haul and return times. The daily datasets were checked for normality using the Shapiro-Wilk test with a significance level of 0.20. The normality check was necessary to determine whether parametric or non-parametric correlation measures should be applied. The performance times were predominantly non-normally distributed, as indicated by Shapiro-Wilk p-values less than the 0.20 significance level. Summary statistics and Shapiro-Wilk test results are presented as Tables 4.1 and 4.2.

Table 4.1: Daily Haul Duration Summary Statistics

Date (1)	**Haul Duration (s)**						
	n (2)	**Mean (3)**	**Std Dev. (4)**	**95% CI of Mean (5)**	**(6)**	**Shapiro-Wilk W (7)**	**p-value (8)**
6-Jun	23	45.913	5.3421	43.603	48.223	0.8860	0.0132
7-Jun	12	44.167	4.2176	41.487	46.846	0.9080	**0.2013**
8-Jun	12	40.333	4.8866	37.229	43.438	0.8540	0.0412
9-Jun	24	45.917	5.2247	43.710	48.123	0.7613	0.0000
12-Jun	15	43.200	2.4842	41.824	44.576	0.9461	**0.4648**
13-Jun	26	44.538	4.9818	42.526	46.551	0.7749	0.0000
15-Jun	8	43.250	3.0119	40.732	45.768	0.7150	0.0033
16-Jun	28	43.286	6.0666	40.933	45.638	0.6873	0.0000
19-Jun	12	50.333	19.1042	38.195	62.472	0.8037	0.0103
20-Jun	12	40.667	6.9978	36.220	45.113	0.8746	0.0749
21-Jun	42	44.095	7.8765	41.641	46.550	0.8878	0.0006
Total	**214**						

Table 4.2: Daily Return Duration Summary Statistics

Date (1)	Return Duration (s)						
	n (2)	Mean (3)	Std Dev. (4)	95% CI of Mean		Shapiro-Wilk	
				(5)	(6)	W (7)	p-value (8)
6-Jun	22	43.091	12.1377	37.709	48.472	0.8007	0.0005
7-Jun	12	43.167	3.1286	41.179	45.154	0.9298	**0.3779**
8-Jun	11	38.909	12.2756	30.662	47.156	0.6317	0.0000
9-Jun	22	40.000	5.1640	37.710	42.290	0.6464	0.0000
12-Jun	15	38.400	10.6690	32.492	44.308	0.6784	0.0001
13-Jun	24	39.417	4.3129	37.595	41.238	0.8978	0.0194
15-Jun	7	32.857	4.5981	28.605	37.110	0.7231	0.0066
16-Jun	26	33.615	6.6937	30.912	36.319	0.6598	0.0000
19-Jun	12	39.667	7.7146	34.765	44.568	0.9251	**0.3312**
20-Jun	12	46.333	12.4706	38.410	54.257	0.8863	0.1056
21-Jun	41	39.902	11.0086	36.428	43.377	0.5661	0.0000
Total	**204**						

The daily datasets were aggregated to produce monthly datasets of haul and return performance times. This was done by testing the homogeneity of the daily datasets to identify daily combinations that were similarly distributed, and discarding the rest. The Kruskal-Wallis (KW) test of homogeneity was used because the performance times were predominantly non-normally distributed. A resulting p-value less than the 0.05 significance level would indicate that at least one of the aggregated datasets was dissimilarly distributed.

It was desirable to produce monthly datasets with the largest number of performance times. Therefore, initially all 11 daily datasets were aggregated and tested for homogeneity. Test results indicated that at least one of the daily datasets differed, and the results were reviewed to identify the dataset with an anomalously high or low mean rank. The identified dataset was removed and the remaining 10 datasets were aggregated and tested again for homogeneity. This process was repeated until the test results showed that all of the remaining datasets were similarly distributed. It was necessary to disregard two daily haul datasets (June 8 and 20), leaving 9 daily datasets to be combined into a monthly haul data set. Five daily return datasets (June 7, 12, 15, 16, and 20) were disregarded, leaving 6 daily datasets to be combined into a monthly return data set. The results of the KW analysis are presented as Table 4. 3.

Table 4.3: Kruskal-Wallis Homogeneity Analysis Results

Activity (1)	No. of Days Combined (2)	No. of Performance Measures (3)	Kruskal-Wallis	
			kw (4)	p-value (5)
Haul	11	214	20.27	0.0268
	10	202	16.87	0.0508
	9	***190***	***11.93***	***0.1545***
Return	11	204	55.79	<0.0001
	10	197	49.77	<0.0001
	9	171	25.59	0.0012
	8	159	16.80	0.0187
	7	147	13.19	0.0401
	6	***132***	***9.87***	***0.0790***

4.4. Data Analysis

The assumption of independently distributed performance times was assessed by investigating the data for five specific correlations:

1. Paired haul and return performance times were used to investigate possible correlations between haul and return activities. The data used were198 haul times that were immediately followed by return times as indicated by the time at which the activities occurred. All 214 haul times collected could not be used because in some instances the truck was assigned to a different operation and therefore did not return to the same load area, while in other instances the haul activity was the last activity of the day.

2. Haul time and the previous haul time were used to investigate possible auto-correlation in the haul activity. The data used to investigate haul auto-correlation was the monthly haul dataset. Of the 190 haul times in the monthly dataset, 179 could be associated with a previous haul time. All 190 haul times could not be used because the first haul time of the day, or the first time after changing operations, could not be associated with a previous haul time.

3. Return time and the previous return time were used to investigate possible auto-correlation in the return activity. The data used to investigate return auto-correlation was the monthly return dataset. Of the 132 return times in the monthly dataset, 125 could be associated with a previous return time.
4. Haul time and the time at which the haul activity started were used to investigate possible correlation between time of day and the haul activity. The data used was the monthly haul dataset. All 190 haul times and the time at which the activity began were used.
5. Return time and the time at which the return activity started were used to investigate possible correlation between time of day and the return activity. The data use was the monthly return dataset. All 132 return times and the time at which the activity began were used.

The data for each of the above five specific correlations were presented as scatter plots to view the data prior to using statistical correlation measures to quantify each. Altman [1991] notes that scatter plots are useful tools when dealing with correlation. Thus, scatter plots were created from each of the five datasets and are presented as Figures 4.1 through 4.5.

4.4.1. Analysis of Correlation between Haul and Return Times

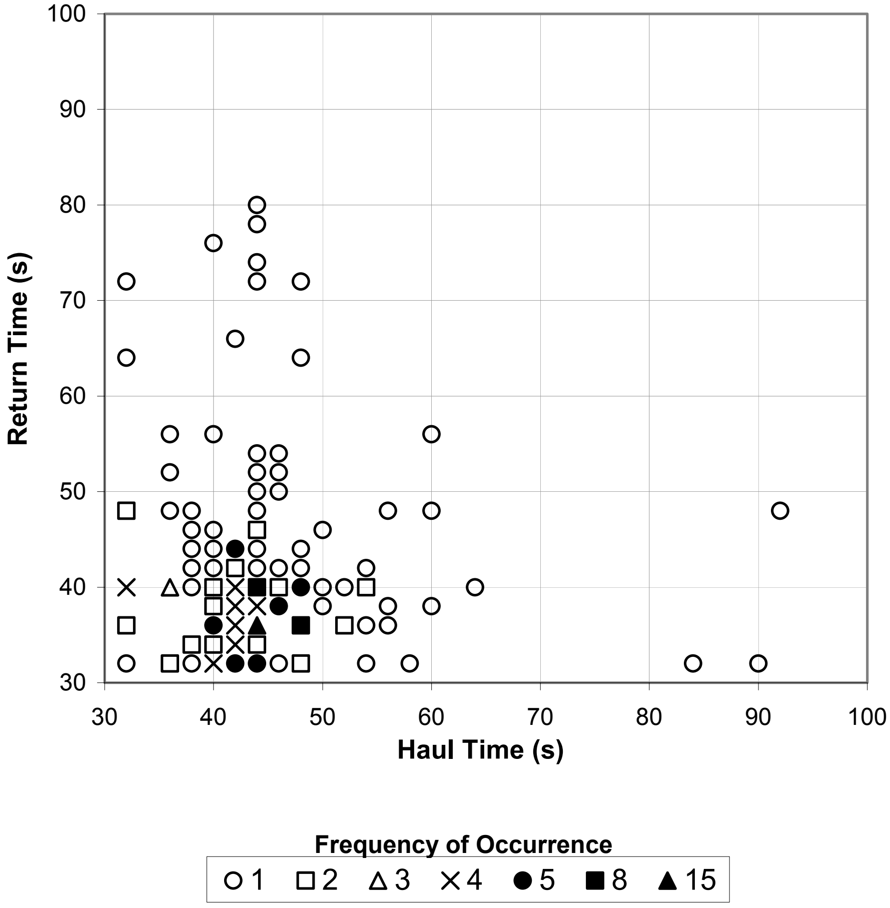


Figure 4.1: Scatter Plot of Activity Correlation Dataset

It can be seen from Figure 4.1 that the data points were clustered about the most frequently occurring data point, which was a haul time of 44 seconds followed by a return time of 36 seconds. Figure 4.2 also shows the joint distribution of the activity correlation dataset. Long haul times were defined as those more than one standard deviation above the mean (45 + 7 = 52 seconds). Sixteen of the 198 haul times (8.1 percent) fell into this category. Long return times were also defined as those more than one standard deviation above the mean (40 + 9 = 49 seconds), and 21 of 198 return times (10.1 percent) fell into this category. Only one data point (0.5 percent) could be classified as having both a long haul and a long return time.

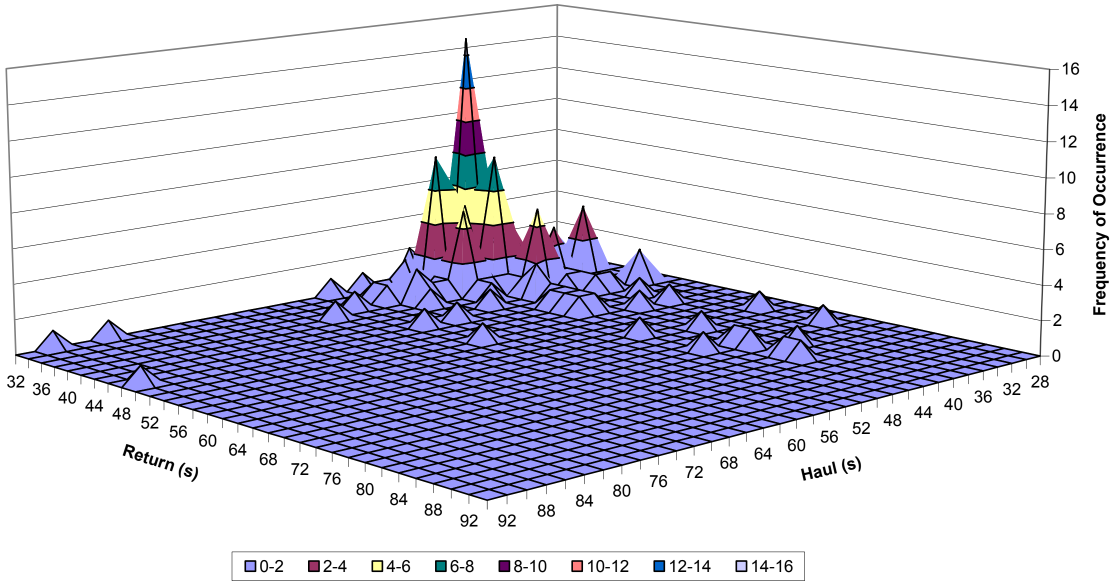


Figure 4.2: 3D Surface Plot of Activity Correlation Dataset

The appearance of Figure 4.1 indicates that haul and return times may be negatively correlated, although the magnitude of the correlation does not appear great. The Shapiro-Wilk test was used to test whether the return times (the dependent variable) were normally distributed with a significance level of 0.05. The resulting W statistic was 0.7442 with a significance level of less than 0.0001, indicating that return times were not normally distributed. Therefore, Spearman's Rank correlation coefficient was used to quantify the correlation between haul and return times, and test its significance at the 0.05 level. The calculated correlation coefficient was –0.01 and a significance level of 0.8460, which is much greater than the 0.05 level. The results indicate that there is a weak negative correlation between haul and return times, but that the correlation is not statistically significant.

4.4.2. Analysis of Auto-Correlation in Haul Times

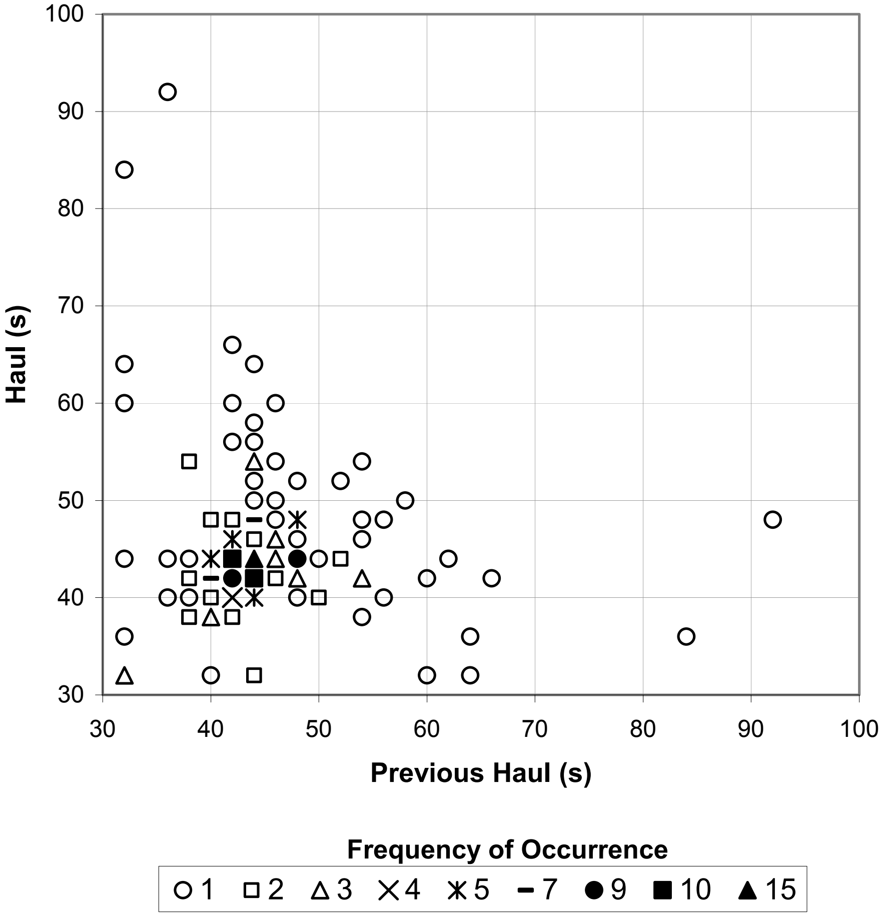


Figure 4.3: Scatter Plot of Haul Auto-Correlation Dataset

The data points in Figure 4.3 were clustered about the most frequently occurring data point, which was a haul time of 44 seconds preceded by a haul time also equal to 44 seconds. The joint distribution of the haul auto-correlation dataset is shown in Figure 4.4. Long haul times were again defined as those more than one standard deviation above the mean, or greater than 52 seconds. Eighteen of the 179 data points (10 percent), and 18 previous haul times (10 percent) fell into this category. Only one data point of the 179 (0.6 percent) could be classified as having both a long haul time and a long previous haul time.

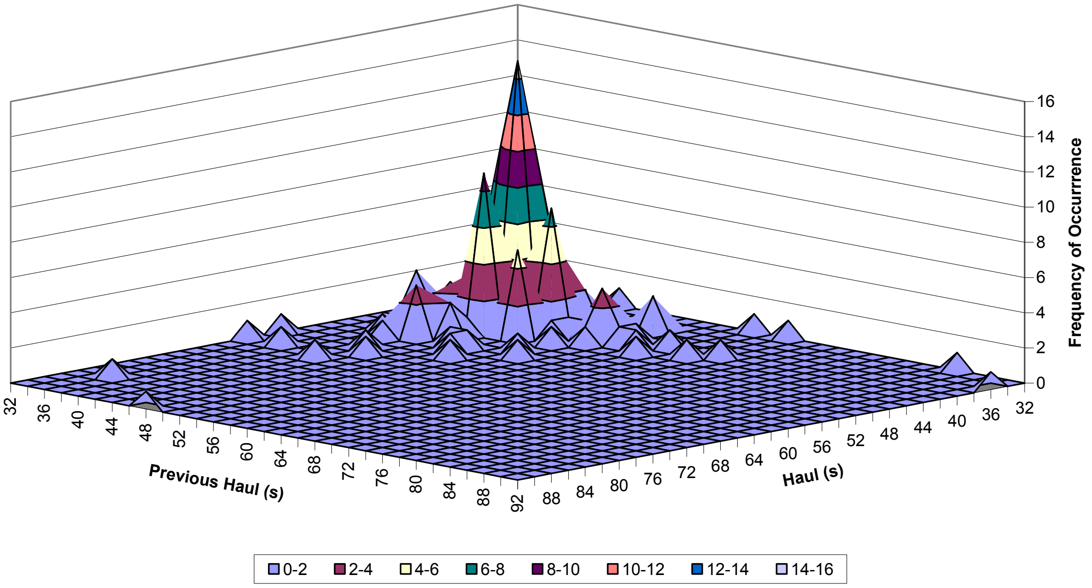


Figure 4.4: 3D Surface Plot of Haul Auto-Correlation Dataset

The appearance of Figure 4.3 indicates that there may be some positive auto-correlation in the haul performance times. The strength of the correlation does not appear great, but does appear greater than that observed in Figure 4.1. The Shapiro-Wilk test was used to test whether the haul times (the dependent variable) were normally distributed with a significance level of 0.05. The resulting W statistic was 0.7754 with a significance level of less than 0.0001, indicating that haul times were not normally distributed. Therefore, Spearman's Rank correlation coefficient was used to quantify the auto-correlation in haul times, and test its significance at the 0.05 level. The calculated correlation coefficient was 0.12 and a significance level of 0.1179, which is greater than the 0.05 level. The results indicate that haul times are positively auto-correlated, but that the correlation is not statistically significant.

4.4.3. Analysis of Auto-Correlation in Return Times

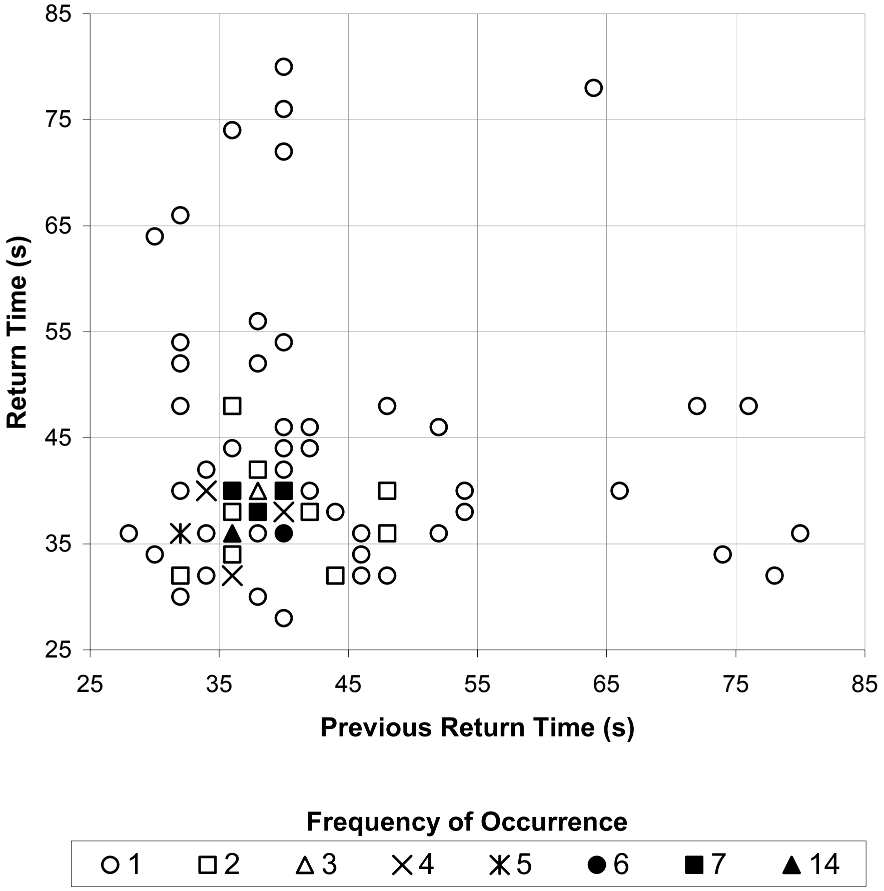


Figure 4.5: Scatter Plot of Return Auto-Correlation Dataset

The data points in Figure 4.5 were clustered about the most frequently occurring data point, which was a return time of 36 seconds preceded by a return time also equal to 36 seconds. Figure 4.6 shows the joint distribution of the return auto-correlation datset. Long return times were again defined as those more than one standard deviation above the mean, or greater than 49 seconds. Twelve of the 125 data points (10 percent), and 11 previous return times (9 percent) fell into this category. Only one data point of the 125 (0.8 percent) could be classified as having both a long return time and a long previous return time.

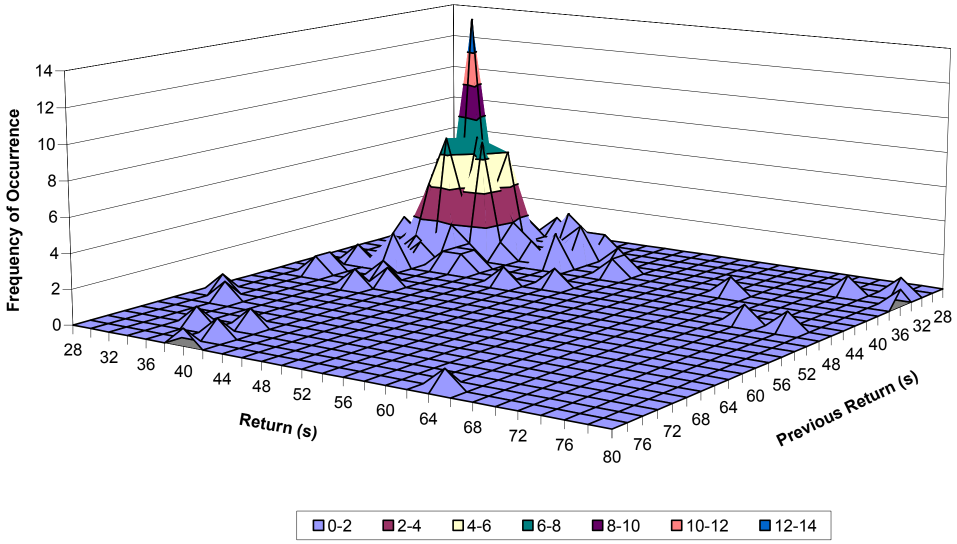


Figure 4.6 3D Surface Plot of Return Auto-Correlation Dataset

The appearance of Figure 4.5 indicates that there may be some positive auto-correlation in the return performance times. The Shapiro-Wilk test was used to test whether the return times (the dependent variable) were normally distributed with a significance level of 0.05. The resulting W statistic was 0.7042 with a significance level of less than 0.0001, indicating that return times were not normally distributed. Therefore, Spearman's Rank correlation coefficient was used to quantify the auto-correlation in return times, and test its significance at the 0.05 level. The calculated correlation coefficient was 0.09 and a significance level of 0.3111, which is greater than the 0.05 level. The results indicate that return times are positively auto-correlated, but that the correlation is not statistically significant.

4.4.4. Analysis of Correlation between Haul Time and Time of Day

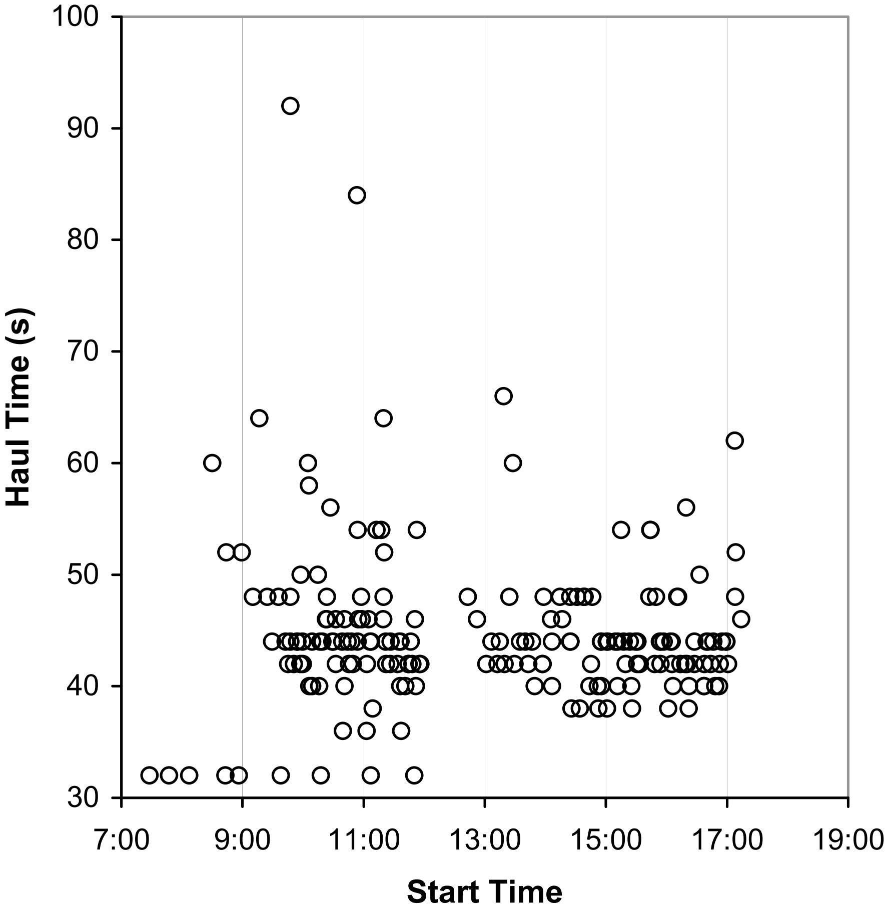


Figure 4.7: Scatter Plot of Time of Day and Haul Time

A majority of the haul times were measured to be between 42 and 46 seconds, and these times occurred throughout the morning and afternoon. This indicates no correlation between haul time and time of day, and the appearance of Figure 4.7 gives the same indication. The Shapiro-Wilk test was used to test whether the haul times (the dependent variable) were normally distributed with a significance level of 0.05. The resulting W statistic was 0.7838 with a significance level of less than 0.0001, indicating that haul times were not normally distributed. Therefore, Spearman's Rank correlation coefficient was used to quantify the correlation between haul time and time of day, and test its significance at the 0.05 level. The calculated correlation coefficient was –0.07 and a significance level of 0.3301, which is greater than the 0.05 level. The results indicate that haul times are negatively correlated with time of day, but that the correlation is not statistically significant.

4.4.5. Analysis of Correlation between Return Time and Time of Day

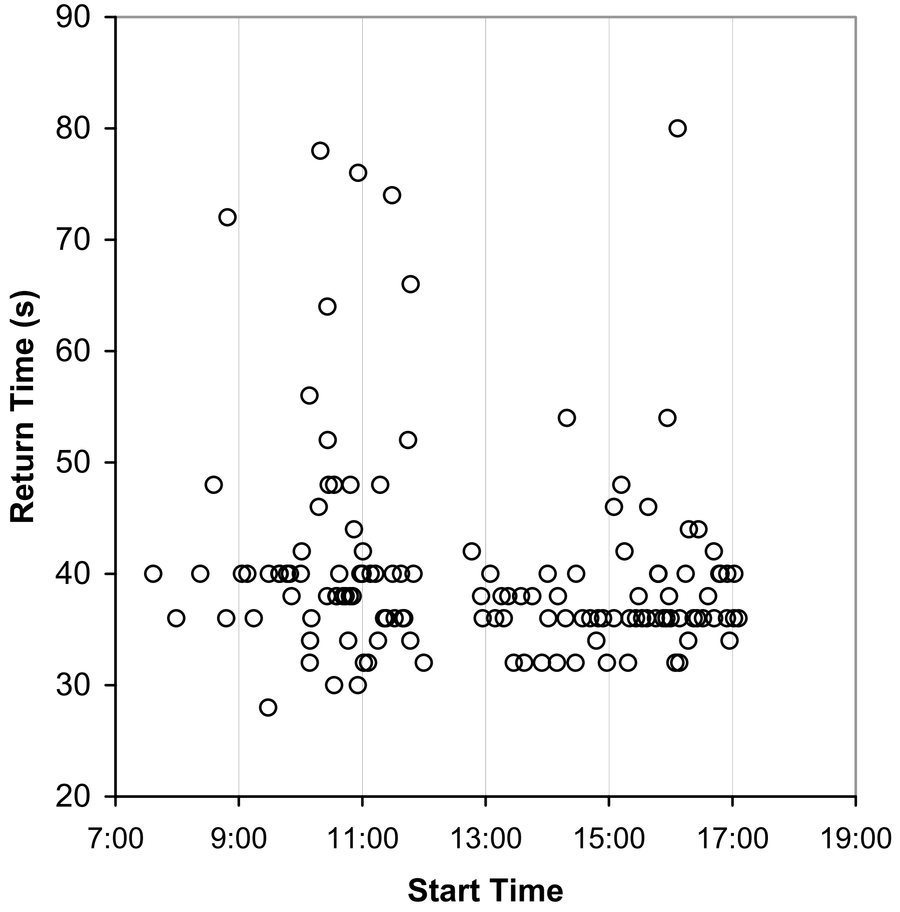


Figure 4.8: Scatter Plot of Time of Day and Return Time

Return times occurring before noon clustered at 38 and 40 seconds, while in the afternoon times clustered at 36 seconds. This indicates a negative correlation between return time and time of day, and the appearance of Figure 4.8 gives the same indication. The Shapiro-Wilk test was used to test whether the return times (the dependent variable) were normally distributed with a significance level of 0.05. The resulting W statistic was 0.7068 with a significance level of less than 0.0001, indicating that return times were not normally distributed. Therefore, Spearman's Rank correlation coefficient was used to quantify the correlation between return time and time of day, and test its significance at the 0.05 level. The calculated correlation coefficient was –0.19 and a significance level of 0.0266, which is less than the 0.05 level. The results indicate that return times are negatively correlated with time of day, and that the correlation is statistically significant. This negative correlation means that longer return times tend to occur early in the day and longer return times occur later in the day.

The results of the correlation investigation are summarized in Table 4. 4.

Table 4.4: Normality and Correlation Test Results

Correlation	**Data Set**		**No. of Samples**	**Shapiro-Wilk**		**Spearman's Rank Correlation**	
	Dependent Variable	Independent Variable		W	p-value	r	p-value
Activity	Return	Haul	198	0.7442	<0.0001	-0.01	0.8460
Auto-Correlation	Haul	Previous	179	0.7754	<0.0001	0.12	0.1179
	Return	Previous	125	0.7042	<0.0001	0.09	0.3111
Time	Haul	Time	190	0.7838	<0.0001	-0.07	0.3301
	Return	Time	132	0.7068	<0.0001	**-0.19**	**0.0266**

4.5. Conclusion

Modeling and simulation techniques typically assume that performance times are independently distributed. This work investigated auto-correlation in both haul and return times, correlation between haul and return times, and correlation between both haul and return times and time of day. Short-interval GPS data was used to autonomously collect the large volumes of data needed to investigate phenomena like correlation.

The correlation coefficients calculated were greater than zero in magnitude, indicating that interdependence exists in the haul and return performance time data collected. However, the measured correlations were not sufficiently strong to conclude that studies based on the concept of independent data are flawed. The presence of statistically significant correlation supports the recommendation that performance data be analyzed for the presence of correlation before an assumption of independent data is made.

The presence of interdependence in the data places a requirement on data collection methods. It is no longer sufficient to collect only performance time, rather it is necessary to also collect the time of day at which the performance was realized so that the data points can be paired for analysis. The collection of short-interval GPS data provides the necessary time of day information, as well as the large volumes of data necessary.

4.6. References

AbouRizk, S. M. and D. W. Halpin (1990), "Probabilistic Simulation Studies for Repetitive Construction Processes," *Journal of Construction Engineering and Management*, ASCE, 116(4), pp575-594.

AbouRizk, S. M. and D. W. Halpin, (1992a), "Modeling input data for construction simulation," in *Proceedings of Computing in Civil Engineering and Geographic Information Systems Symposium*, pp1147-1154.

AbouRizk, S. M. and D. W. Halpin, (1992b), "Statistical properties of construction duration data," *Journal of Construction Engineering and Management*, ASCE, 118(3), pp525-544.

Altman, D. G. (1991), *Practical Statistics for Medical Research*, Chapman and Hall, London.

Benjamin, J. and C. Cornell (1970), *Probability, Statistics, and Decision for Civil Engineers,* McGraw-Hill, Inc., New York, NY.

Conover, W. J. (1980), *Practical Nonparametric Statistics*, 2nd Ed., John Wiley & Sons, New York.

Cor, H. and J.C. Martinez (1999), "A Case Study in the Quantification of a Change in the Conditions of a Highway Construction Operation," in *Proceedings of the 1999 Winter Simulation Conference*, P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans (Ed.), Phoenix, AZ, pp1007-1009.

D'Agostino, R. and M. Stephens (1986), *Goodness-of-Fit Techniques*, Marcel Dekker, Inc., New York.

Fente, J., K. Knudson, and C. Schexnayder (1999), "Defining a Beta Distribution for Construction Simulation," in *Proceedings of the 1999 Winter Simulation Conference*, P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans (Ed.), Phoenix, AZ, pp1010-1015.

Fente, J., C. Schexnayder, and K. Knudson (2000), "Defining a Probability Distribution Function for Construction Simulation," *Journal of Construction Engineering and Management*, ASCE, 126(3), pp234-241.

Ioannou, P.G. (1999), "Construction of a Dam Embankment with Nonstationary Queues," in *Proceedings of the 1999 Winter Simulation Conference*, P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans (Ed.), Phoenix, AZ, pp921-928.

Ioannou, P. G. and J. C. Martinez, (1996a) "Comparison of Construction Alternatives Using Matched Simulation Experiments," *Journal of Construction Engineering and Management*, ASCE, 112(3), ASCE, Reston, VA.

Ioannou, P.G. and J.C. Martinez (1996b), "Simulation of Complex Construction Processes," in *Proceedings of the 1996 Winter Simulation Conference*, J.M. Charnes, D.J. Morrice, D.T. Brunner, and J.J. Swain (Ed.), San Diego, CA, pp1329-1336.

Law, A. M. and W. D. Kelton, (2000), *Simulation Modeling and Analysis*, 3rd Ed., McGraw-Hill Higher Education, Boston.

Maio, C., C. Schexnayder, K. Knutson, and S. Weber (2000), "Probability Distribution Functions for Construction Simulation" *Journal of Construction Engineering and Management*, ASCE, 126(4), pp285-292.

Martinez, J. C., (1998), "EarthMover - Simulation Tool for Earthwork Planning," in *Proceedings of the 1998 CIB W78 Conference*, Department of Construction Management and Economics, Royal Institute of Technology, Stockholm, Sweden.

Martinez, J. C. and P. G. Ioannou, (1995), "Advantages of the Activity Scanning Approach in the Modeling of Complex Construction Processes," in *Proceedings of the 1995 Winter Simulation Conference*, Society for Computer Simulation, San Diego, CA.

Miller, R. G. (1986), *Beyond ANOVA, Basics of Applied Statistics*, John Wiley & Sons, New York.

Sangarayakul, B. (1998), "Use of Simulation to Analyze Block Manufacturing Methods." MS Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Shapiro, S., Wilk, M., and Chen, H. (1968), "A Comparative Study of Various Tests for Normality," *Journal of the American Statistical Association*, v 63 n 324, pp 1343-1372.

Sheskin, D. J. (1997), *Handbook of Parametric and Nonparametric Statistical Procedures*, CRC Press, Boca Raton, FL.

CHAPTER 5. FIELD VERIFICATION OF CYCLE TIME ESTIMATING TOOLS

5.1. Introduction

Modeling and simulation techniques can be used to plan operations or to evaluate alternatives. Regardless of their use, models should be verified if at all possible and the results validated before being used as the basis for decisions. Verification is the process of ensuring that the model functions as designed, while validation is the process of determining whether the results obtained from the model accurately represent the actual operation [Law and Kelton 2000]. Additionally, Law and Kelton [2000] note that the most definitive test for validity is whether the simulation results closely resemble the actual output of the system that was modeled.

The performance of construction equipment can be predicted using a number of analytical tools, including software packages to estimate productivity. These estimating tools use data input by the user and mathematical abstractions, or models, of both the production system and equipment to predict productivity. Models of construction equipment can be designed to work from first principles or from derived measures. Models developed from first principles use characteristics of the engine and power train as inputs to predict machine performance parameters such as acceleration, braking, and maximum speed. Models developed from derived measures make use of machine characteristics such as rimpull and retarder performance curves to predict steady state speed under given conditions.

This chapter will investigate the validity of haul times estimated from derived measures by comparing them with measures of actual haul times. Haul times will be estimated using VolvoSim and compared to measures of actual haul time reduced from short-interval GPS data. The results of each are compared and used to determine the validity of the estimates.

5.2. Background

Models are used to simulate and predict the outcome of a real world operations performed under certain conditions. These models are a mathematical abstraction that approximates the operation as closely as possible. The expected conditions under which the operation is to be performed are input as data to produce results designed to estimate the output of the operation under the given set of conditions. The resulting estimates are used as the basis for decisions regarding productivity, performance, and efficiency of the system.

The most definitive test of model validity is whether the simulation results closely resemble the expected output of the actual system [Law and Kelton 2000]. One manner in which this can be determined is through comparison of simulation results under given conditions with the values obtained from the actual system under similar conditions. While this is a highly desirable validation method, it should be noted that it can often not be used because the simulated operation has not yet, or never will be, performed in the field.

Figure 4.1 is a graphical depiction of the challenges involved in simulating an actual earthmoving operation. The top portion depicts the steps needed to simulate the operation detailed in the bottom portion of the figure. The simulation input data are values for factors that influence the performance of the field operation. These factors include equipment performance parameters, such as engine, transmission, braking system, as well as operator characteristics. Operational factors considered include roadway curvature, grade, and surface condition, as well as material, and management characteristics.

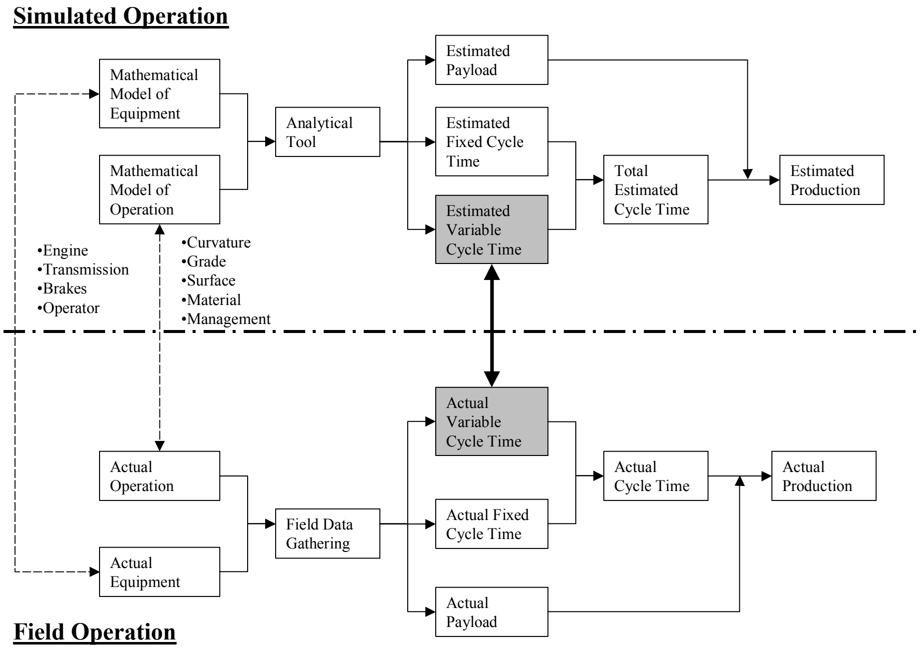


Figure 5.1: Simulated and Field Operation Schematic

The factor values are input into the analytical tool to produce estimates of payload, fixed cycle time, and variable cycle time. The fixed and variable cycle time estimates together form the total estimated cycle time. Portions of the cycle, such as maneuver and dump times are relatively fixed, while the time required to travel between the load and dump areas varies with payload, distance, and road parameters such as gradient, curvature, and surface condition. The estimated production rate is a function of the estimated payload and the total estimated cycle time.

This shows that cycle time estimates are an extremely important component of production estimates. Therefore, manufacturers and analysts have developed analytical tools able to estimate the variable travel times based on the equipment used and the characteristics of the operation. This paper seeks to compare the estimates of variable cycle times calculated using VolvoSim with actual cycle times measured in the field using short-interval GPS data logging devices.

5.2.1. Variable Cycle Time Estimates

Variable cycle time estimates for both haul and return are an integral part of production estimates for earthmoving operations, and they often comprise a majority of the cycle time. Travel times are initially estimated as part of the production and cost estimates needed for bid preparation. Estimates are also made as construction progresses as a means for evaluating equipment types and routing alternatives. Originally, estimating methods were manual in nature involving the use of rimpull charts and other standards. Computer based methods have also been developed to perform essentially the same calculations at a more detailed level. Equipment specifications and project conditions are input data for most methods, and computer programs are often pre-packaged with a database of equipment specifications to aid the estimator.

5.2.1.1. Manual Method

The estimator is required to gather the equipment and project information needed to manually calculate travel times. Manufacturers publish equipment handbooks containing information such as rimpull curves and operating weight. The estimator uses this information along with roadway information to predict truck speed, which is then used to calculate the time required to travel a known distance [Peurifoy et. al 1996]. Such methods are time consuming and tedious for large projects.

The accuracy of these methods is also suspect for a number of reasons. Gransberg [1998] points out the resulting estimates are based on theoretical maximum instantaneous values, which differ substantially from the results obtained from sustained operations. Additionally, most manual methods are based on steady state conditions and neglect acceleration.

5.2.1.2. Computer Programs

Analytical tools to model the performance of construction equipment have been developed in electronic format for use with computers. Originally developed by equipment manufacturers as sales tools, productivity planning software is now used by estimators to predict travel times. Two types of equipment performance models available: those working from knowledge of the mechanical components and those working from measures that represent the performance of the machine as a whole.

Perdomo [2001] developed a model of hauling equipment in the MATLAB environment based on knowledge of engine characteristics and transmission ratios. The primary result of the model is travel time for a specified haul unit and route. The author noted difficulties in modeling torque converter performance as well as difficulties in determining transmission losses.

Torque converter performance, transmission losses, and gear reductions are parameters fundamental to the performance of a machine. While these values are measured by and known to the equipment manufacturers, they are considered proprietary information and not published. However, the values are used to derive rimpull curves for the equipment, and these curves are published. Perdomo [2001] describes a method of calculating the necessary data from the rimpull curves, rated RPM of the engine, and the maximum power generated at that RPM and was able to approximate values of torque converter efficiencies and reductions using an empirical approach.

The alternative approach to working from first principles is to develop a performance model of the machine as a whole, essentially producing an electronic version of the charts and graphs used to manually estimate travel time. Equipment manufactures are able to measure machine performance, such as the relationship between rimpull and speed in a given gear, with precision in a laboratory environment and many have developed software for estimating production based on this approach.

The calculations associated with this approach are similar to the hand calculations, but are performed in an incremental fashion. Speed is estimated and assumed constant over the distance traveled before a change in status warrants recalculation of the speed. This recalculation is often performed when the speed changes by a pre-defined magnitude.

VolvoSim is an example of a software program that uses this approach in estimating the productivity and economics of operations [VolvoSim 2001]. The user specifies the haul route, equipment types, material type, work schedule, and any operating limitations such as road surface conditions, loading strategy, or speed limits. VolvoSim uses the input data along with data contained in the equipment and material databases to perform production, loading, and cost analyses.

VolvoSim is capable of providing both a quick estimate and a full simulation of operational productivity. The quick estimate is a deterministic calculation, thereby neglecting any variation in the data. The advantages of such a calculation are that they can be made relatively quickly and produce repeatable results. Although quick and repeatable, the results produced may not be realistic. A full simulation makes use of stochastic data definitions, thereby incorporating data variations and producing varying results. More realistic results are produced, but the calculations take longer to perform, and are not necessarily repeatable.

5.2.2. Travel Time Observations

The tools available for measuring activity performance times from production operations include stopwatches, video cameras, and GPS receivers. Stopwatch studies are by far the simplest studies, requiring the least capital investment and the most human involvement [Oglesby et. al 1989]. In such a study, an observer in the field records the time required to complete tasks of interest. The introduction of video cameras to time studies made it possible to record operations and perform repeated reviews in an office environment. Computer programs such as PAVIC+ [Bjornsson and Sagert 1994] and DVAT [Frank 2001] were developed to facilitate videotape analysis using computers. Both methods are limited to line of sight of the observer, whether the observer be a person or a camera, and rely on the analyst to instantaneously decide when one task stops and the next starts.

Recently analysts have turned to technology for tools not subject to these limitations and have used on-board instrumentation (OBI) systems in the form of machine health monitoring systems, as well as GPS data collection systems. The Vital Information Management System (VIMS) produced by Caterpillar, Inc. was used to gather payload and load time information for analysis [Kannan and Vorster 2000]. The use of short-interval GPS data logging devices to measure activity duration was described in Chapter 3.

5.3. Field Verification

A field study of a conventional load and haul earthmoving operation was performed to compare estimated haul times produced by VolvoSim with travel times recorded using short-interval GPS data. The operation entailed loading overburden and shot rock into Volvo A25C six-wheeled articulated haulers using a Caterpillar 350 excavator, hauling the material approximately 2,500 meters, and dumping it to construct a roadway embankment. The haul road was well maintained and established, but crossed three public roads. It carried traffic in both directions, except for two of the road crossings and near the load area. Traffic control personnel monitored all road crossing and gave priority to the trucks.

5.3.1. Determining Haul Road Parameters

The haul road was surveyed to establish the centerline profile and alignment. The survey was performed with an electronic total station and tied to established project control points with known Virginia South state plane coordinates. The survey data was used to divide the roadway into segments of approximate constant grade. This resulted in 19 road segments of varying lengths. Fifteen of the 19 segments formed the principal haul path between the load and dump areas, and therefore were studied in this work. The remaining four segments formed a loop around the load area and carried only intermittent traffic. In addition to the 15 segments identified from the survey, a segment was used to model each of the three public road crossings.

Parameters regarding horizontal curvature and surface condition were also determined, in addition to the length and grade of each segment. Only one of the 15 studied segments contained horizontal curvature. This curve was minor and did not influence the manner in which the trucks were operated, and was therefore not considered. The road surface condition was described by the ground structure, traction coefficient, and rolling resistance. Ground structure refers to the size of obstacles in the roadway, and a structure class of 0.4 corresponding to obstacles 3 to 4 cm in size was used to represent all but the road crossing segments. A structure class of 0.2 corresponding to 2 to 3 cm was used to represent the road crossing segments.

The haul road had a well-compacted dirt surface, thus a traction coefficient of 0.4 and a rolling resistance of three percent were used to represent the actual conditions. The traction coefficient describes the portion of the generated rimpull force that is available to do work, while rolling resistance describes the friction between the tires and the road surface. Parameters of the studied segments are provided as Table 5. 1.

Table 5.1: Haul Road Parameters

Segment	Length (m)	Grade (%)	Curve Angle	Ground Structure	Traction Coefficient	Rolling Resist. (%)
Haul 5	58.9	12.9	0	Class 0.4	0.4	3
Haul 6	231.3	-3.6	0	Class 0.4	0.4	3
Haul 7	78.6	5.8	0	Class 0.4	0.4	3
Cross South Main St	10	0	0	Class 0.2	0.4	3
Haul 8	282.2	-3.6	0	Class 0.4	0.4	3
Haul 9	269.5	-6.7	0	Class 0.4	0.4	3
Cross Yellow Sulphur Rd	5	0	0	Class 0.2	0.4	3
Haul 10	67.4	5.6	0	Class 0.4	0.4	3
Haul 11	188	-4	0	Class 0.4	0.4	3
Haul 12	88.5	-12.5	0	Class 0.4	0.4	3
Cross Jennell Rd	5	0	0	Class 0.2	0.4	3
Haul 13	164.6	2.9	0	Class 0.4	0.4	3
Haul 14	294.5	5.8	0	Class 0.4	0.4	3
Haul 15	64.3	-6.1	0	Class 0.4	0.4	3
Haul 16	212.2	0.6	0	Class 0.4	0.4	3
Haul 17	60.7	11.6	0	Class 0.4	0.4	3
Haul 18	152.1	-4	0	Class 0.4	0.4	3
Haul 19	140.9	0	0	Class 0.4	0.4	3

5.3.2. Estimating Haul Times

VolvoSim was used to perform a haul analysis and estimate the mean time required for the trucks to travel each roadway segment. In addition to the haul road parameters described, VolvoSim requires information regarding the maximum and final speed, load status, and throttle factor for each segment. Maximum speed is the greatest attainable speed at which the trucks can travel. The trucks were mechanically fixed to operate only in low range, where the maximum attainable speed is 34 kph [Volvo 2002]. Therefore, 34 kph was used for all segments except the road crossings, where the maximum speed was 5 kph. This 5-kph speed limit represented the actual operation of the trucks, as they were observed to cross roads slowly. Final speed is the maximum speed at which the truck may be traveling when the end of the segment is reached. The trucks were observed slowing at the first road crossing and stopping at the others. Therefore, a final speed of 5 kph was assigned to segment number 7, which preceded the first road crossing. The final speed for segment numbers 9 and 12, which preceded the other crossings, was set at 0 kph. For all other segments, the final speed was set at the maximum speed of 34 kph.

The load status for all segments was "Full", as it was haul times that were estimated. The throttle factor for all segments was 100 percent to indicate that the trucks were operated aggressively to achieve maximum production. With the exception of the road crossings, nothing precluded the trucks from operating at or near full throttle. The VolvoSim input data is provided as Table 5. 2.

Table 5.2: Haul Road Parameters and VolvoSim Input Data

Segment	Length (m)	Grade (%)	Curve Angle	Ground Structure	Traction Coefficient	Rolling Resist. (%)	Velocity (kph)		Load Status	Throttle Factor (%)
							Max	Final		
Haul 5	58.9	12.9	0	Class 0.4	0.4	3	34	34	Full	100
Haul 6	231.3	-3.6	0	Class 0.4	0.4	3	34	34	Full	100
Haul 7	78.6	5.8	0	Class 0.4	0.4	3	34	5	Full	100
Cross South Main St	10	0	0	Class 0.2	0.4	3	5	34	Full	100
Haul 8	282.2	-3.6	0	Class 0.4	0.4	3	34	34	Full	100
Haul 9	269.5	-6.7	0	Class 0.4	0.4	3	34	0	Full	100
Cross Yellow Sulphur Rd	5	0	0	Class 0.2	0.4	3	5	34	Full	100
Haul 10	67.4	5.6	0	Class 0.4	0.4	3	34	34	Full	100
Haul 11	188	-4	0	Class 0.4	0.4	3	34	34	Full	100
Haul 12	88.5	-12.5	0	Class 0.4	0.4	3	34	0	Full	100
Cross Jennell Rd	5	0	0	Class 0.2	0.4	3	5	34	Full	100
Haul 13	164.6	2.9	0	Class 0.4	0.4	3	34	34	Full	100
Haul 14	294.5	5.8	0	Class 0.4	0.4	3	34	34	Full	100
Haul 15	64.3	-6.1	0	Class 0.4	0.4	3	34	34	Full	100
Haul 16	212.2	0.6	0	Class 0.4	0.4	3	34	34	Full	100
Haul 17	60.7	11.6	0	Class 0.4	0.4	3	34	34	Full	100
Haul 18	152.1	-4	0	Class 0.4	0.4	3	34	34	Full	100
Haul 19	140.9	0	0	Class 0.4	0.4	3	34	0	Full	100

When VolvoSim is used to simulate the operation, the haul analysis results in a distribution of travel times for each roadway segment, however only the mean value is reported. VolvoSim varies the payload as each cycle is simulated to more accurately simulate the operation. Payload is an input data for calculating travel time, and it is this input variation that propagates to the results producing a distribution of travel times. The estimated mean travel time for each segment is provided in Table 5. 3, along with the haul road parameters and VolvoSim input data.

Table 5.3: Road Parameters, Input Data, and Estimated Mean Haul Times

Segment	Length (m)	Grade (%)	Curve Angle	Ground Structure	Traction Coefficient	Rolling Resist. (%)	Velocity (kph)		Load Status	Throttle Factor (%)	Estimated Time (s)
							Max	Final			
Haul 5	58.9	12.9	0	Class 0.4	0.4	3	34	34	Full	100	29
Haul 6	231.3	-3.6	0	Class 0.4	0.4	3	34	34	Full	100	36
Haul 7	78.6	5.8	0	Class 0.4	0.4	3	34	5	Full	100	16
Cross South Main St	10	0	0	Class 0.2	0.4	3	5	34	Full	100	-
Haul 8	282.2	-3.6	0	Class 0.4	0.4	3	34	34	Full	100	38
Haul 9	269.5	-6.7	0	Class 0.4	0.4	3	34	0	Full	100	40
Cross Yellow Sulphur Rd	5	0	0	Class 0.2	0.4	3	5	34	Full	100	-
Haul 10	67.4	5.6	0	Class 0.4	0.4	3	34	34	Full	100	19
Haul 11	188	-4	0	Class 0.4	0.4	3	34	34	Full	100	27
Haul 12	88.5	-12.5	0	Class 0.4	0.4	3	34	0	Full	100	26
Cross Jennell Rd	5	0	0	Class 0.2	0.4	3	5	34	Full	100	-
Haul 13	164.6	2.9	0	Class 0.4	0.4	3	34	34	Full	100	32
Haul 14	294.5	5.8	0	Class 0.4	0.4	3	34	34	Full	100	68
Haul 15	64.3	-6.1	0	Class 0.4	0.4	3	34	34	Full	100	10
Haul 16	212.2	0.6	0	Class 0.4	0.4	3	34	34	Full	100	23
Haul 17	60.7	11.6	0	Class 0.4	0.4	3	34	34	Full	100	8
Haul 18	152.1	-4	0	Class 0.4	0.4	3	34	34	Full	100	17
Haul 19	140.9	0	0	Class 0.4	0.4	3	34	0	Full	100	26

5.3.3. *Measuring Haul Times*

GPS data was collected at a short time interval by instrumentation installed on-board four of the trucks engaged in the studied operation. The CrossCheck AMPS system manufactured by Trimble Navigation, Limited, was used to record time, horizontal position, speed, and other data in ASCII text format at a two-second time interval. The system was programmed to record the following data string:

AABBBBCDDDDDEEEFFFFFGGGGHHHHHIIIJJJKL

Where:
AA is the event number
BBBB is the GPS week
C is an integer representing the day of the week
DDDDD is the current time Greenwich Mean Time (GMT) expressed in seconds since midnight
EEE.FFFFF represents degrees of latitude with north positive
GGGG.HHHHH represents degrees of longitude with east positive
III is vehicle speed in miles per hour
JJJ is the vehicle heading in degrees with zero at true north and increasing eastwardly
K is an integer representing the quality of the data recorded
L is an integer representing the age of the data

A sample of collected data is shown below.

```
>REV001065269090+3719328-0804034100000412;ID=0012;*5E<
>REV001065269092+3719328-0804034100000412;ID=0012;*5C<
>REV001065269094+3719328-0804034100000412;ID=0012;*5A<
```

Data was collected as part of a field evaluation of the system that is described and the results presented in Chapter 2. The buffer capacity of the system was approximately 2.5 hours, therefore data was downloaded from the system on each truck twice daily: at lunch break and end of shift. Data was gathered twice daily over an approximate six-week period from four trucks resulting in 192 data files. The collected data was imported into the Microsoft Excel spreadsheet program and reduced to measures of actual travel time for each roadway segment identified from the survey. These are the same segments on which the VolvoSim analysis was based.

The data reduction methods employed were an extension of the Fixed Boundary Time Identification Module (FBTIM) described in Chapter 3. FBTIM allows the user to define rectangular areas and autonomously identify the times at which the truck entered and exited each defined area.

The FBTIM was extended beyond the limit of four rectangular areas to allow for non-orthogonal boundaries defining the area around each of the 15 roadway segments. The surveyed roadway alignment was used to delineate the segments and define fixed linear boundaries around each segment. Each boundary was defined as a mathematical inequality. The inequalities were compared to the recorded truck positions to determine which if any of the areas the truck was located. The status of the truck changed as it crossed the boundary from one area to another, and the time at which the boundary was crossed was identified and summarized for further analysis. The identified times were then used to calculate the time required for the truck to travel between boundaries, or to travel the segment between boundaries.

GPS data was downloaded twice daily from the each of the four trucks for a period of 31 days (24 working days), resulting in 192 data files. The number of cycles captured in each file varied from zero when the truck was assigned to a different operation to eight when the truck worked full time on the studied operation. This resulted in 9,044 haul times for each of the 15 segments studied. A sample of the measured actual haul times is presented as Table 5. 4, which shows approximately 180 of the 9,044 total times collected.

Table 5.4: Sample of Measured Actual Haul Times from GPS Data

Segment	Estimated Haul Time (s)	Measured Haul Times (s)															
Haul 5	29	20		12						22	18			10			
Haul 6	36	52	34	42						40	48						
Haul 7	16	14	12	16						14	16	12	12	14	12	16	12
Haul 8	38	40	34	50						40	36	36	38	34	38	34	46
Haul 9	40	46	40	38	40	40	40	38	36	100	42	42	40	48	38	40	68
Haul 10	19	22	26	24	18	18	20	20	20	18	18	20	18	20	20	20	18
Haul 11	27	24	24	24	30	26	24	26	28	30	30	28	30	30	32	30	24
Haul 12	26	24	26	22	22	20	20	20	32	24	30	22	28	54	26	28	28
Haul 13	32	24	28	28	22	24	22	26	26	22	26					78	20
Haul 14	68	80	96	88	70	84	76	82	88	74	74						84
Haul 15	10	6	8	8	8	8	8	8	8	8	8			8			8
Haul 16	23	26	26	28	26	30	28	28	28	24	24			24			24
Haul 17	8	18	20	16	16	18	16	18	16	14	14			12			14
Haul 18	17	20	22	22	20	22	24	26	22	20	20			18			22
Haul 19	26	20	18	18	22	20	22	24	20	20	20			22			20

5.4. Data Analysis

Data analysis was performed in three steps: normalize the data to produce accuracy ratios, identify and eliminate outlying data points, and analyze the accuracy ratios by segment. An accuracy ratio was defined as:

$$\text{Accuracy Ratio} = \frac{Actual \quad Haul \quad Time}{Estimated \quad Haul \quad Time}$$

The accuracy ratio was calculated for each of the 9,044 actual haul times measured. Table 5. 5 provides the calculated ratios for the measured haul times presented in Table 5.4.

Table 5.5: Sample of Accuracy Ratios

Segment	Estimated Haul Time (s)	Accuracy Ratios															
Haul 5	29	0.69		0.41						0.76	0.62			0.34			
Haul 6	36	1.44	0.94	1.17						1.11	1.33						
Haul 7	16	0.86	0.74	0.99						0.86	0.99	0.74	0.74	0.86	0.74	0.99	0.74
Haul 8	38	1.04	0.89	1.30						1.04	0.94	0.94	0.99	0.89	0.99	0.89	1.20
Haul 9	40	1.16	1.01	0.96	1.01	1.01	1.01	0.96	0.91	2.53	1.06	1.06	1.01	1.21	0.96	1.01	1.72
Haul 10	19	1.18	1.40	1.29	0.97	0.97	1.08	1.08	1.08	0.97	0.97	1.08	0.97	1.08	1.08	1.08	0.97
Haul 11	27	0.89	0.89	0.89	1.11	0.96	0.89	0.96	1.04	1.11	1.11	1.04	1.11	1.11	1.19	1.11	0.89
Haul 12	26	0.93	1.01	0.85	0.85	0.78	0.78	0.78	1.24	0.93	1.16	0.85	1.09	2.09	1.01	1.09	1.09
Haul 13	32	0.74	0.86	0.86	0.68	0.74	0.68	0.80	0.80	0.68	0.80					2.41	0.62
Haul 14	68	1.17	1.40	1.29	1.02	1.23	1.11	1.20	1.29	1.08	1.08						1.23
Haul 15	10	0.59	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78			0.78			0.78
Haul 16	23	1.11	1.11	1.20	1.11	1.28	1.20	1.20	1.20	1.03	1.03			1.03			1.03
Haul 17	8	2.31	2.56	2.05	2.05	2.31	2.05	2.31	2.05	1.79	1.79			1.54			1.79
Haul 18	17	1.15	1.26	1.26	1.15	1.26	1.38	1.49	1.26	1.15	1.15			1.03			1.26
Haul 19	26	0.76	0.68	0.68	0.83	0.76	0.83	0.91	0.76	0.76	0.76			0.83			0.76

5.4.1. Identifying Outliers

It was assumed that ratios less than 0.50 and greater than 3.00 were influenced by irregular events and were neglected. A total of 246 haul times comprising approximately three percent of the 9,044 total haul times were neglected as outlying data points.

The accuracy ratios ranged in value from 0.28 to 24.14. Extreme values such as these are clearly erroneous, as the estimates are based on the laws of physics. These outlying data points are a result of the autonomous data reduction methods employed. Any of a number of irregular events could have produced long haul times and thus high ratios, such as the operator stopping to talk with other personnel or stopping due to mechanical problems. Short haul times and thus small ratios could result from the truck traveling unloaded. Load status was not collected as part of the GPS data, but rather was inferred based on the direction of travel. These irregular events were not considered when the data was reduced to measures of haul time. Therefore, it was necessary to identify and eliminate outlying data points.

5.4.2. Analyzing Accuracy Ratios by Segment

The remaining 8,798 accuracy ratios were used to analyze the haul times by segment. Theoretical distributions were fit to the measured accuracy ratios for each segment using Stat::Fit, a computer program designed for distribution fitting. AbouRizk and Halpin addressed how to fit distributions to construction data [1990] and what distributions are appropriate to use [1992a and 1992b]. AbouRizk and Halpin [1990] describe the three steps in the fitting process:

1. Identify an appropriate distribution
2. Estimate parameters of the selected distribution
3. Test the goodness-of-fit.

Flexible distributions, such as the beta distribution, are recommended by AbouRizk and Halpin [1992a and 1992b] for modeling cycle time data, as the underlying probability density functions can take many differing shapes. Fente et al. [1999 and 2000] researched fitting theoretical distributions to construction data and also concluded that the beta distribution is suitable for modeling construction activity durations.

Maio et al. [2000] fit theoretical distributions to large sets of construction data gathered using automated data collection techniques. Data was gathered from 54,000 truck cycles and the BestFit computer program was used to estimate parameter distributions and test the goodness-of-fit. Maio et al. [2000] point out that a large number of observations may lead to problems with goodness-of-fit tests. The tests provide a clear indication of how well the observations match the theoretical distributions when the number of observations is small. However, the results can be misleading when a large number of observations is used as all distributions fit to large data sets were rejected based on goodness-of-fit test results. The Kolmogorov-Smirnov (KS) goodness-of-fit test was found to be appropriate for use with large data sets.

Stat::Fit was used to fit several continuous distributions to the accuracy ratios for each segment. A minimum value of 0.50 was assumed for each distribution, corresponding to the minimum accuracy ratio deemed valid. Parameters for each distribution were estimated by Stat::Fit using the maximum likelihood method, and the chi-squared, Kolmogorov-Smirnov (KS), and Anderson-Darling (AD) goodness-of-fit tests were used to quantify the fit between the theoretical distribution and the observed data. A significance level of 0.05 was used for each goodness-of-fit test. The three best fitting distributions for each segment are provided in Table 5.6. Those distributions that passed the goodness-of-fit tests are italicized in the table.

Segment number five was the only segment for which the theoretical distributions passed the goodness-of-fit tests; this is also the segment with the least number of observations. This is in agreement with the results obtained by Maio et al. [2000] that the number of observations influences the results of goodness-of-fit tests. The Log-Logistic distribution was the distribution best fitting the data; it was fit to data from 10 of 15 segments, and was the distribution best fitting 9 segments. The Pearson 5 and Pearson 6 distributions were also noted to fit 9 and 10 of the 15 segments, respectively. However, no single distribution was found in the top three fitting distributions for each segment. This indicates that no single distribution should be used, but rather a number of distributions should be checked to find the best fitting distribution.

The histograms and best fitting theoretical distribution are presented as Figures 5.2.1 to 5.2.15, and the cumulative frequency curves are presented as Figures 5.3.1 to 5.3.15.

Table 5.6: Distributions Fit to Measured Haul Times

Segment	Number of Measured Samples	Fitted Distribution	KW-Statistic	p-value
Haul 5	169	***Weibull (0.5, 1.6, 0.601)***	0.0690	0.371
		Gamma (0.5, 2.15, 0.251)	0.0853	0.162
		Pearson 6 (0.5, 1.82e+04, 2.14, 7.21e+4)	0.0859	0.158
Haul 6	345	Log-Logistic (0.5, 4.68, 0.522)	0.0972	0.00269
		Pearson 5 (0.5, 7.08, 3.54)	0.109	0.0132
		Pearson 6 (0.5, 0.0526, 73.8, 7.67)	0.113	0.000249
Haul 7	453	Lognormal (0.5, -0.879, 0.735)	0.0989	0.000258
		Pearson 6 (0.5, 0.477, 3.9, 4.41)	0.104	9.74E-05
		Log-Logistic (0.5, 2.34, 0.418)	0.113	1.78E-05
Haul 8	481	Log-Logistic (0.5, 8.03, 0.466)	0.131	9.96E-08
		Pearson 5 (0.5, 17.9, 8.19)	0.144	3.72E-09
		Pearson 6 (0.5, 0.245, 50.8, 26.6)	0.154	1.72E-10
Haul 9	743	Log-Logistic (0.5, 4.35, 0.534)	0.0964	1.75E-06
		Pearson 5 (0.5, 6.52, 3.34)	0.106	9.21E-08
		Pearson 6 (0.5, 0.121, 31.5, 7.27)	0.116	3.18E-09
Haul 10	764	Beta (0.5, 2.9, 6.16, 19.4)	0.165	1.38E-18
		Log-Logistic (0.5, 5.44, 0.561)	0.178	0
		Pearson 6 (0.5, 16.2, 8.48, 238)	0.180	0
Haul 11	759	Log-Logistic (0.5, 7.05, 0.457)	0.203	0
		Pearson 5 (0.5, 13.9, 6.32)	0.234	0
		Pearson 6 (0.5, 0.0677, 105, 15.5)	0.238	0
Haul 12	750	Erlang (0.5, 3, 0.116)	0.131	8.71E-12
		Weibull (0.5, 1.68, 0.391)	0.135	2.47E-12
		Beta (0.5, 2.95, 2.27, 13.7)	0.138	6.09E-13
Haul 13	636	Log-Logistic (0.5, 3.81, 0.229)	0.164	1.71E-15
		Pearson 5 (0.5, 4.92, 1.07)	0.164	2.14E-15
		Pearson 6 (0.5, 0.00531, 204, 4.98)	0.165	1.05E-15
Haul 14	622	Log-Logistic (0.5, 5.86, 0.641)	0.0873	0.000138
		Pearson 5 (0.5, 11.7, 7.41)	0.0992	8.55E-06
		Pearson 6 (0.5, 0.0104, 720, 11.9)	0.0997	7.55E-06
Haul 15	635	Normal (0.752, 0.114)	0.359	0
		Weibull (0.5, 2.29, 0.284)	0.381	0
		Beta (0.5, 1.57, 3.44, 11.2)	0.384	0
Haul 16	632	Log-Logistic (0.5, 10.9, 0.567)	0.245	0
		Pearson 5 (0.5, 28.6, 16.3)	0.245	0
		Lognormal (0.5, -0.547, 0.197)	0.244	0
Haul 17	596	Pearson 5 (0.5, 14.4, 20.2)	0.138	2.45E-10
		Inverse Gaussian (0.5, 21.5, 1.5)	0.152	1.79E-12
		Lognormal (0.5, 0.37, 0.257)	0.157	3.07E-13
Haul 18	635	Pearson 6 (0.5, 0.306, 98.4, 42)	0.225	0
		Lognormal (0.5, -0.325, 0.181)	0.231	0
		Pearson 5 (0.5, 33.5, 23.8)	0.233	0
Haul 19	578	Pearson 6 (0.5, 0.108, 15.1, 6.85)	0.167	1.52E-14
		Log-Logistic (0.5, 3.98, 0.244)	0.168	1.07E-14
		Lognormal (0.5, -1.39, 0.484)	0.169	7.49E-15

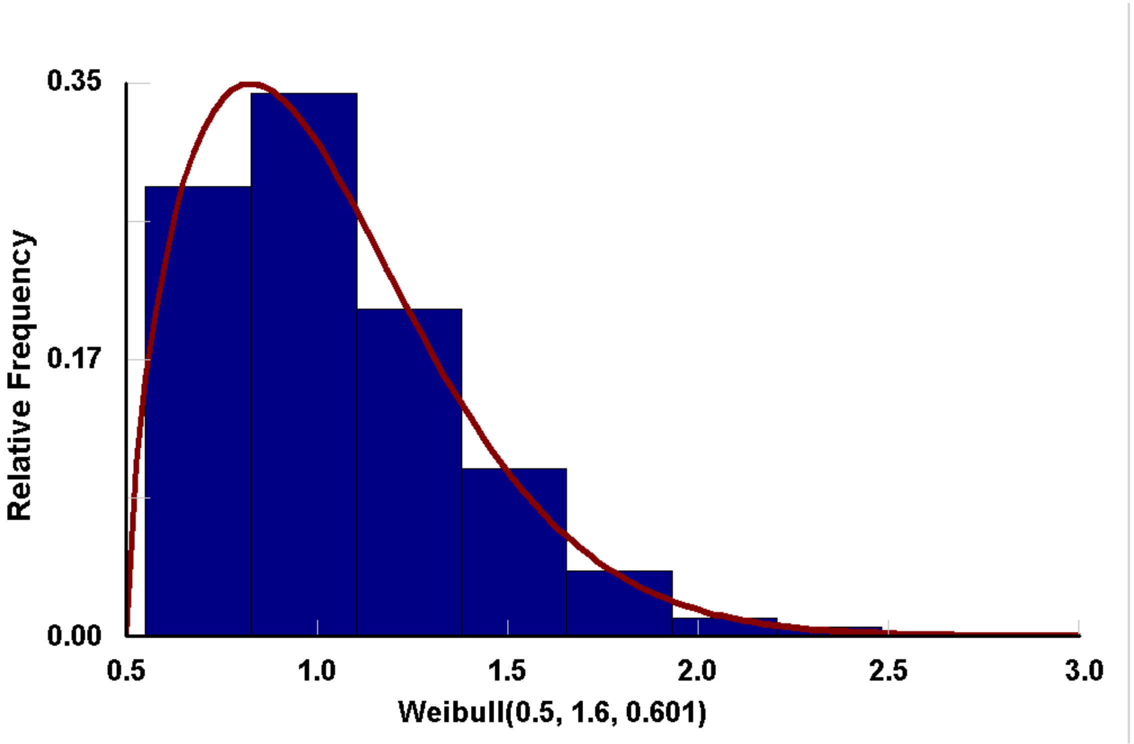


Figure 5.2.1: Histogram of Accuracy Ratios for Segment Number 5

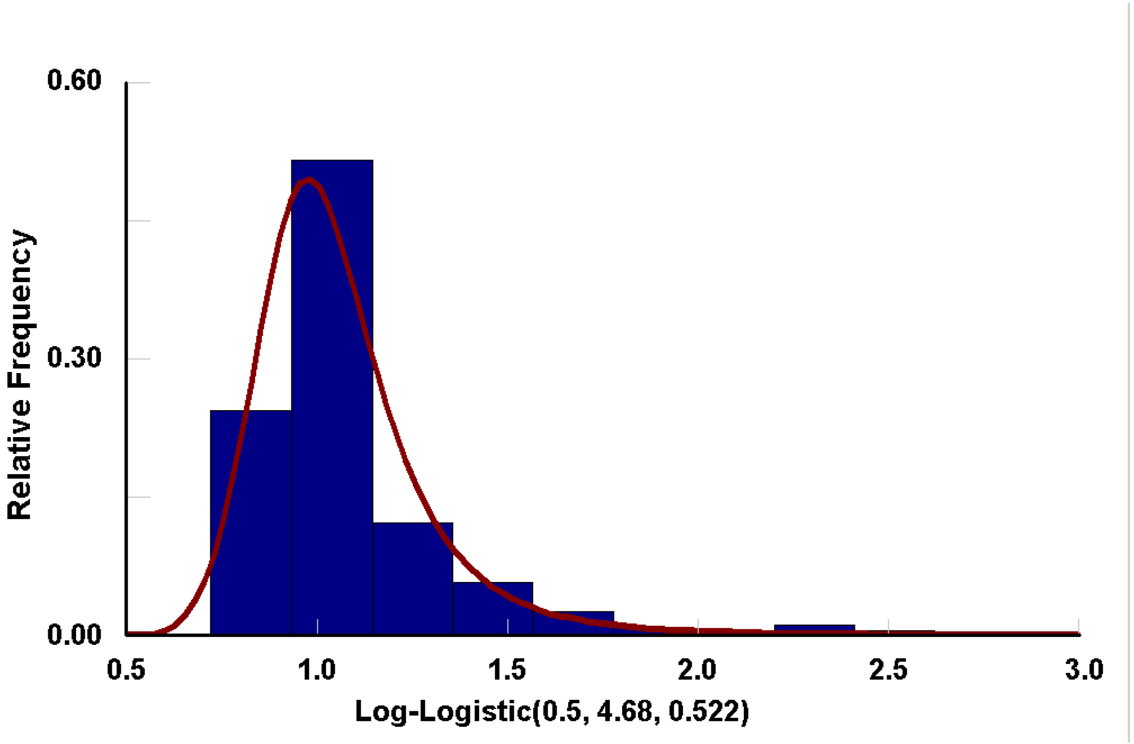


Figure 5.2.2: Histogram of Accuracy Ratios for Segment Number 6

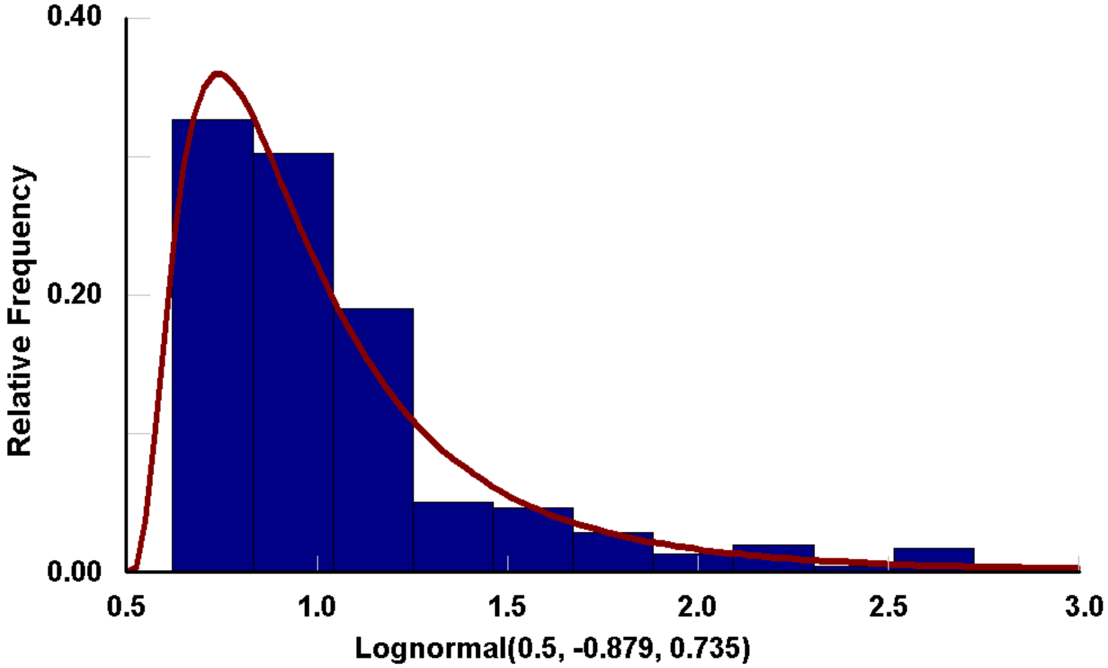


Figure 5.2.3: Histogram of Accuracy Ratios for Segment Number 7

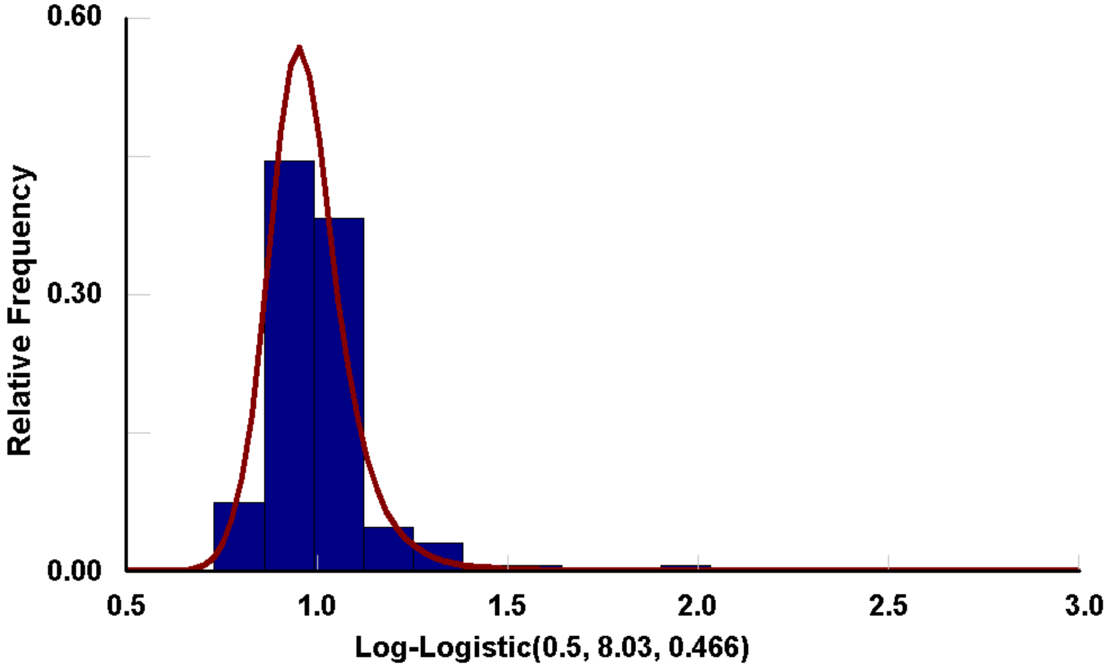


Figure 5.2.4: Histogram of Accuracy Ratios for Segment Number 8

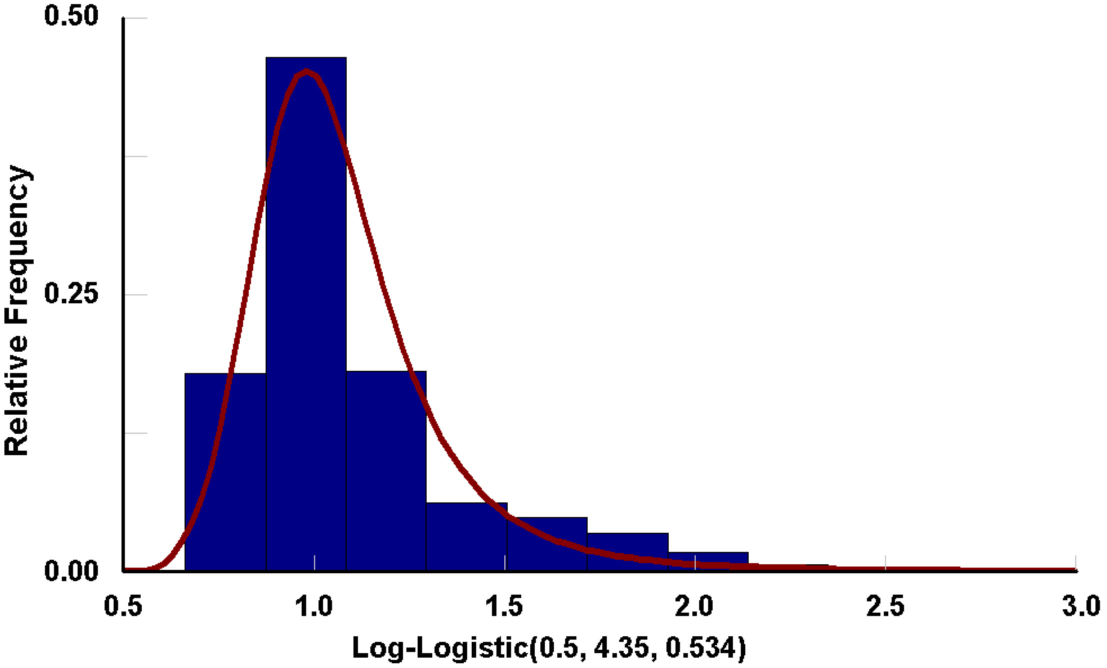


Figure 5.2.5: Histogram of Accuracy Ratios for Segment Number 9

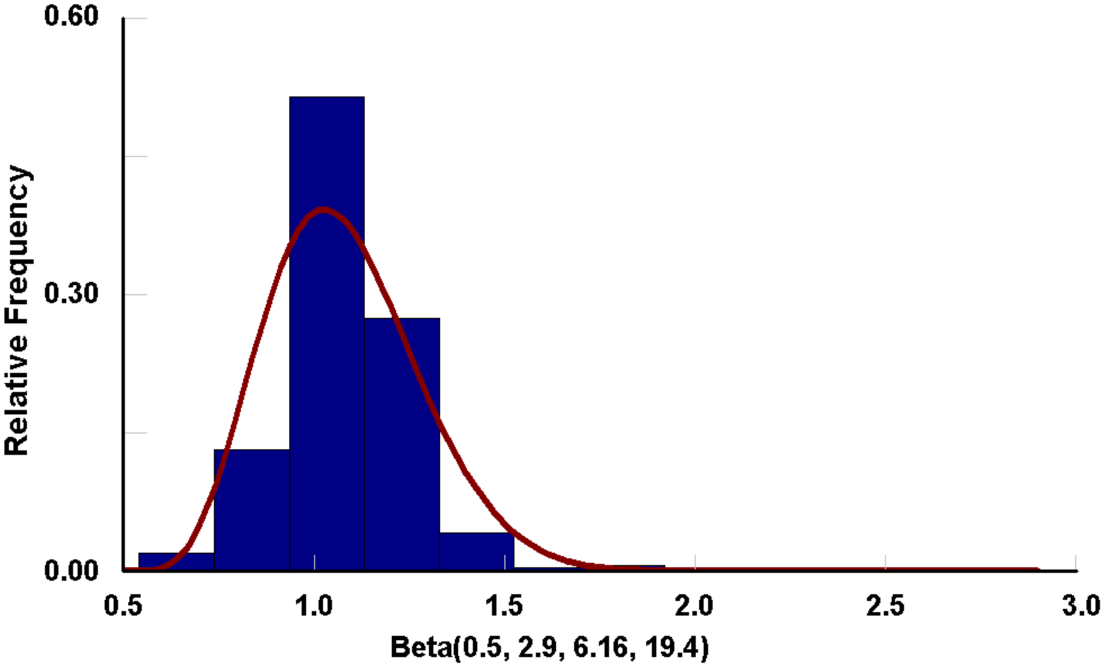


Figure 5.2.6: Histogram of Accuracy Ratios for Segment Number 10

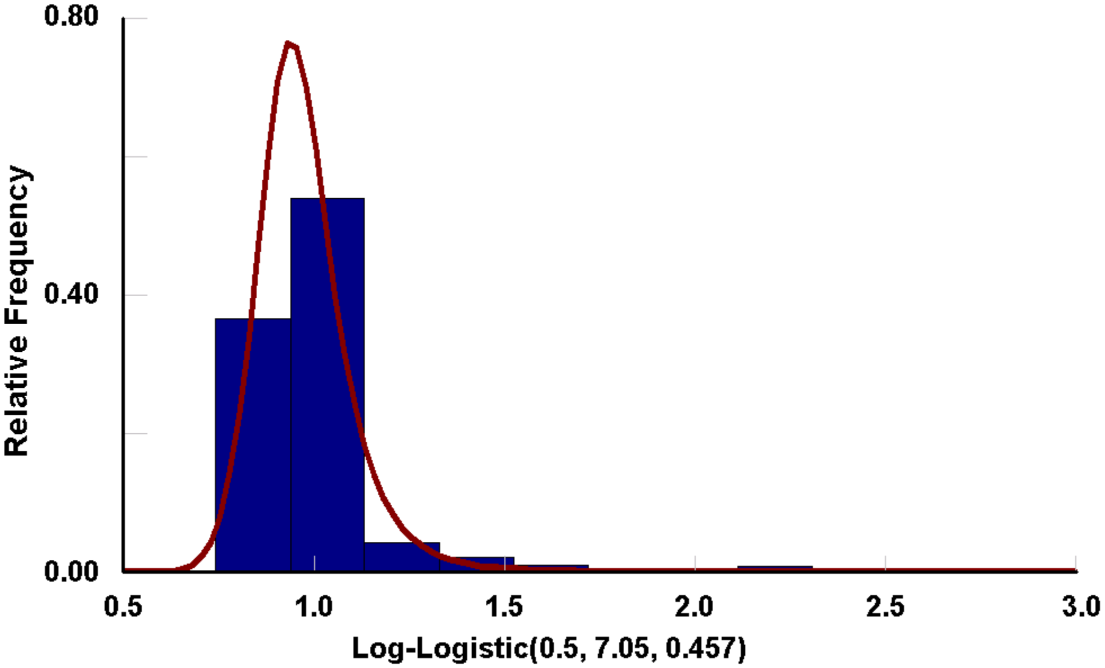


Figure 5.2.7: Histogram of Accuracy Ratios for Segment Number 11

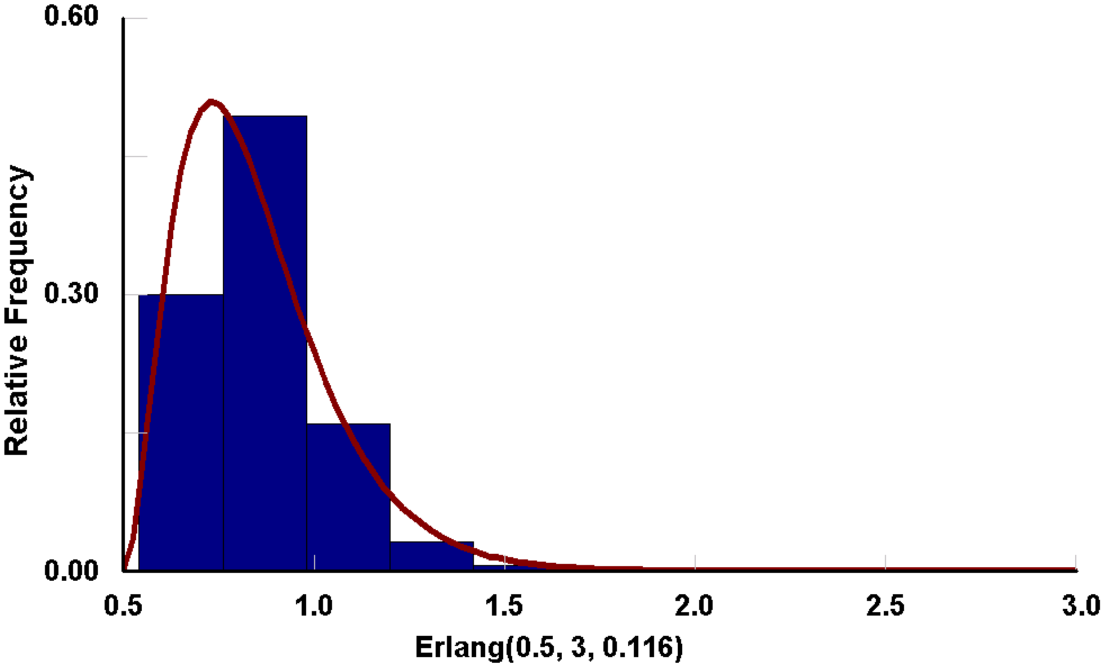


Figure 5.2.8: Histogram of Accuracy Ratios for Segment Number 12

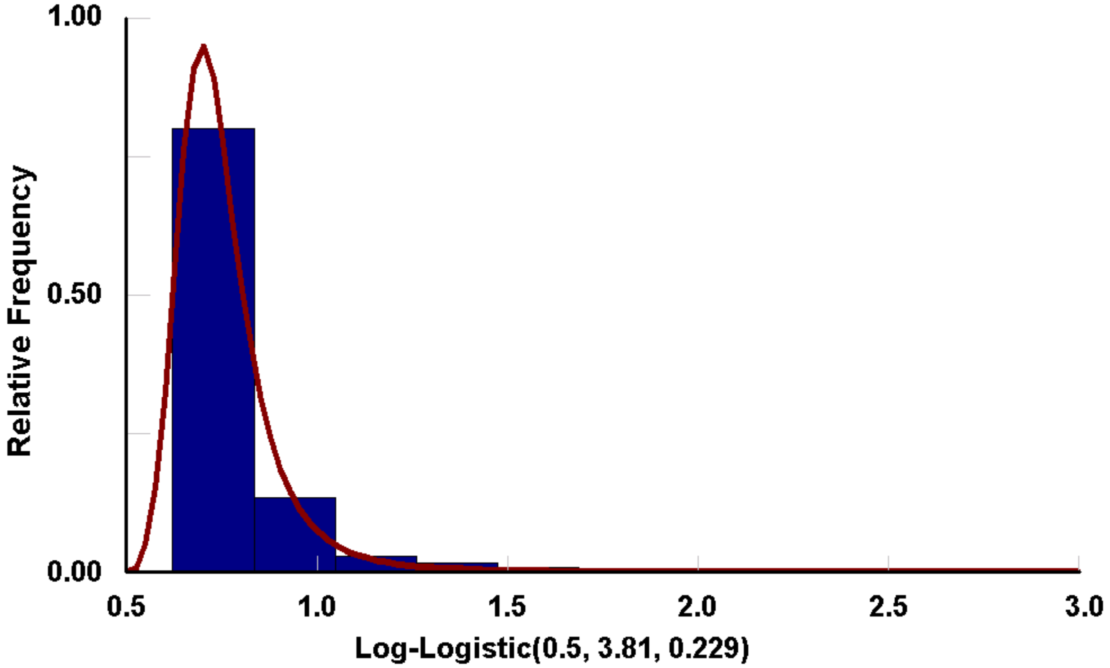


Figure 5.2.9: Histogram of Accuracy Ratios for Segment Number 13

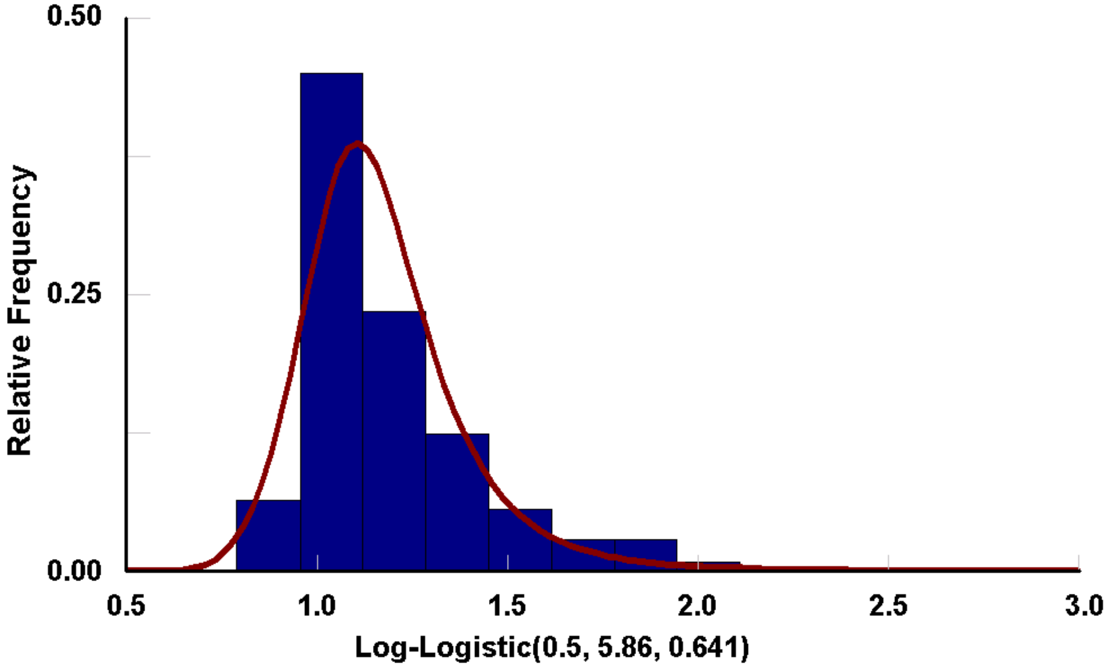


Figure 5.2.10: Histogram of Accuracy Ratios for Segment Number 14

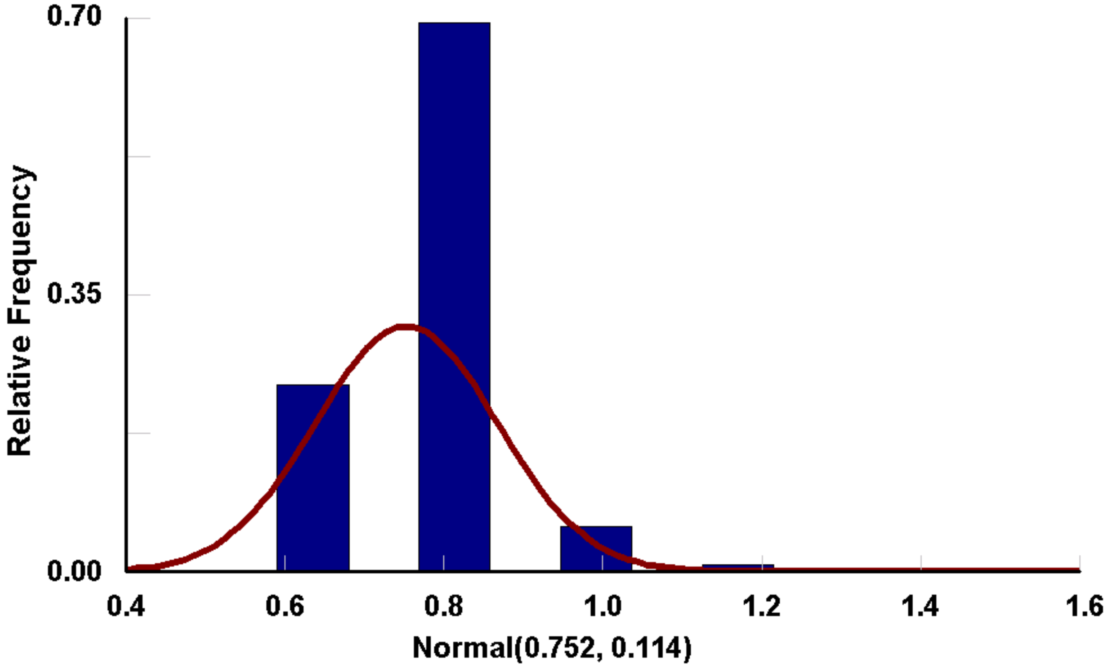


Figure 5.2.11 Histogram of Accuracy Ratios for Segment Number 15

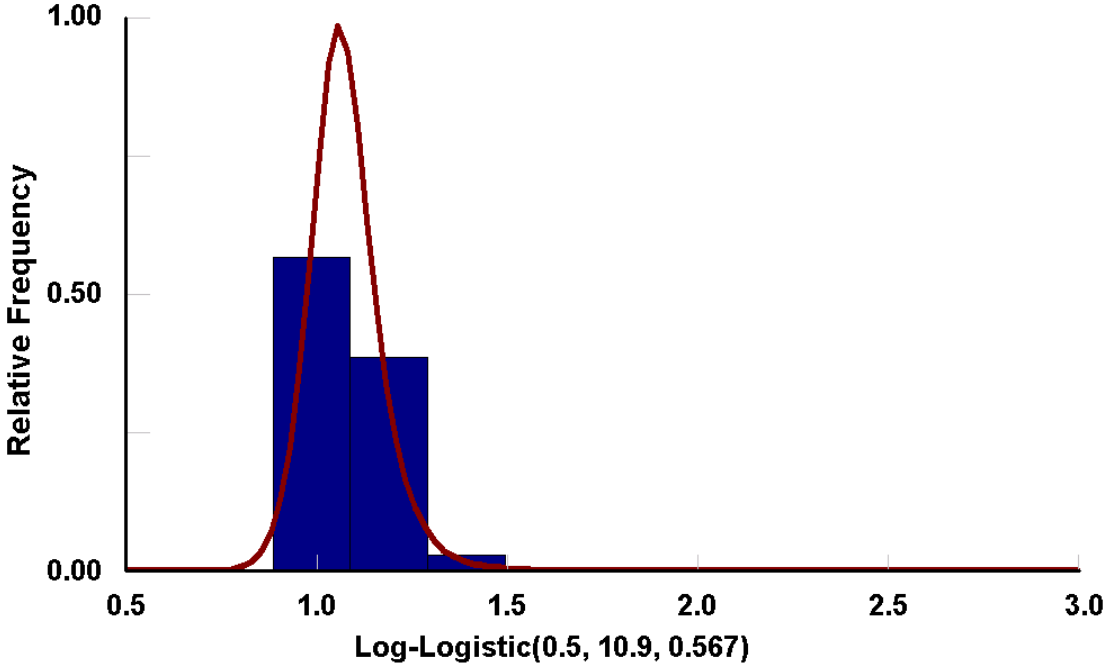


Figure 5.2.12 Histogram of Accuracy Ratios for Segment Number 16

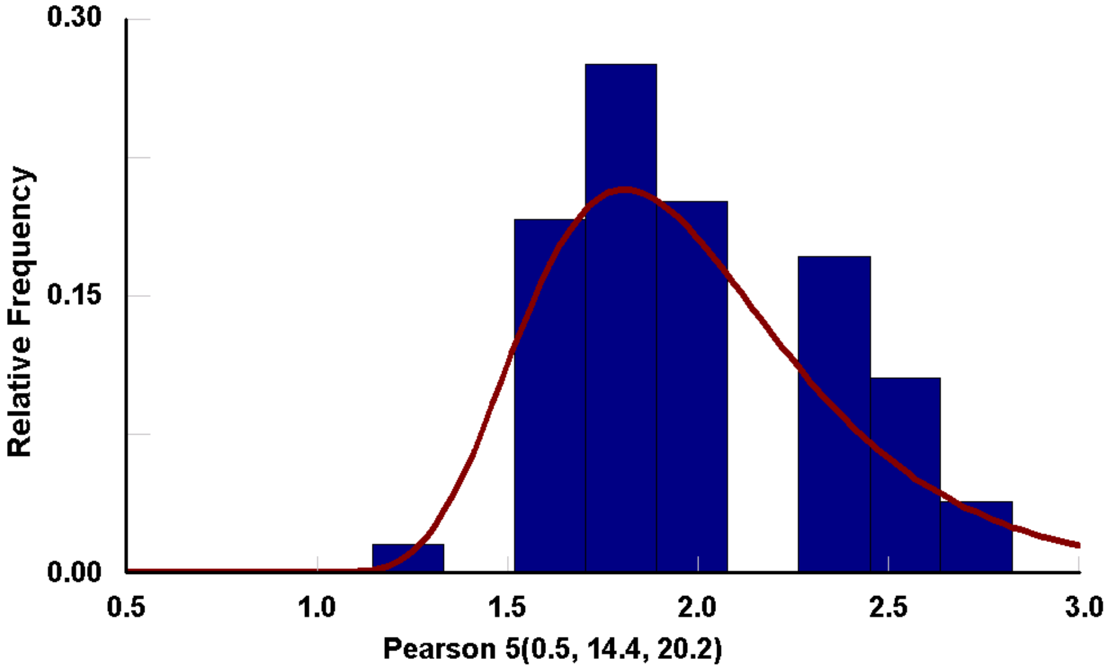


Figure 5.2.13 Histogram of Accuracy Ratios for Segment Number 17

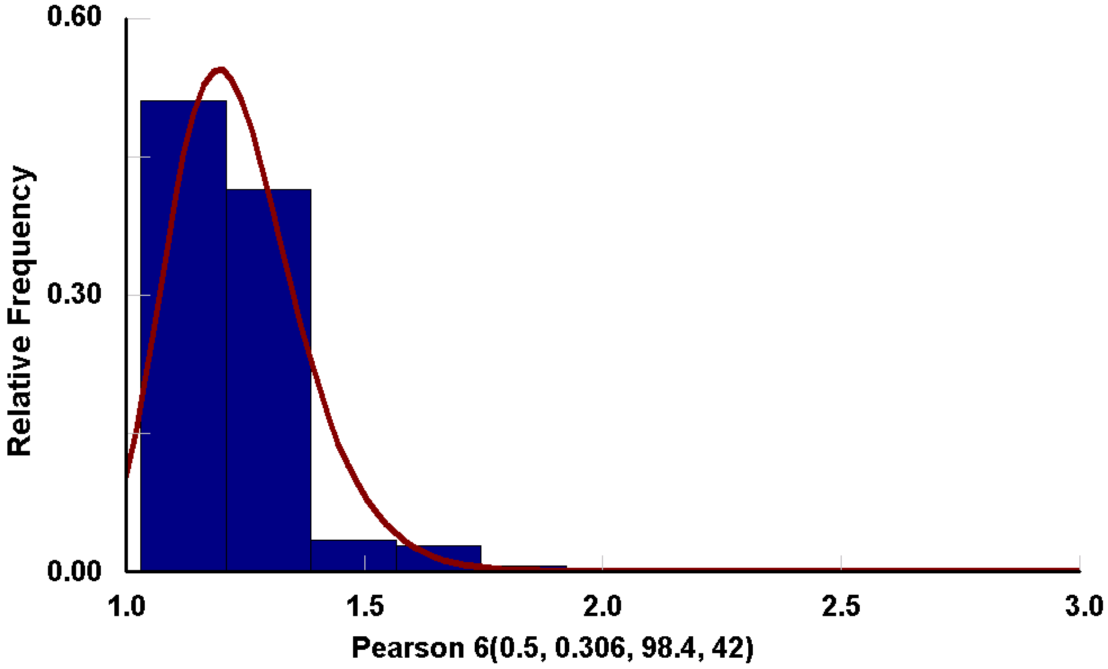


Figure 5.2.14 Histogram of Accuracy Ratios for Segment Number 18

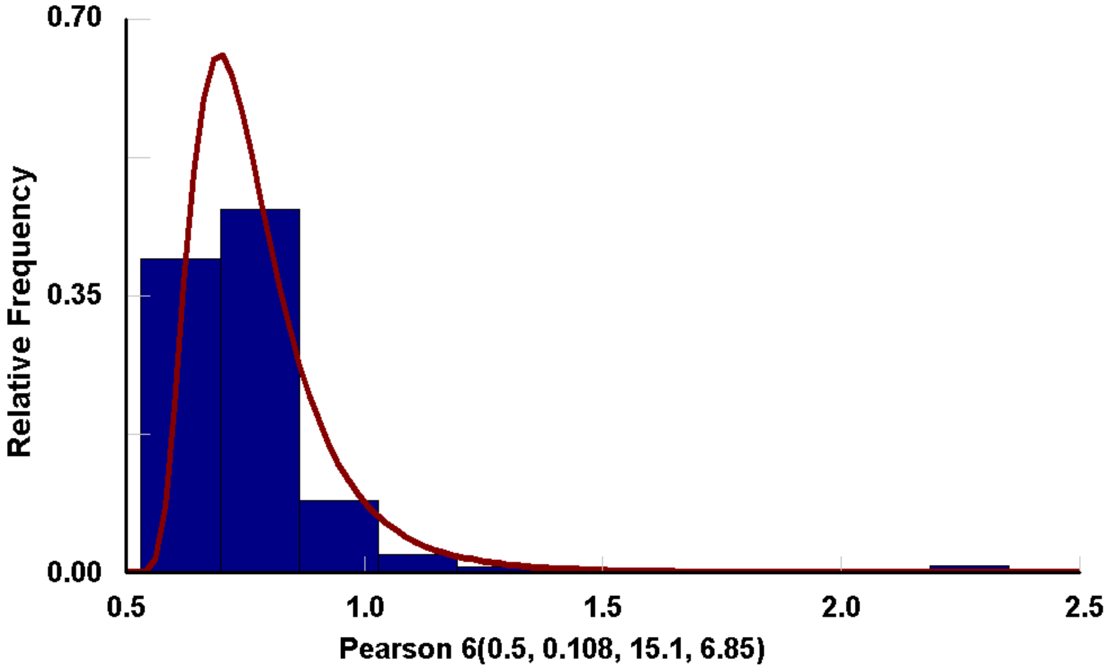


Figure 5.2.15 Histogram of Accuracy Ratios for Segment Number 19

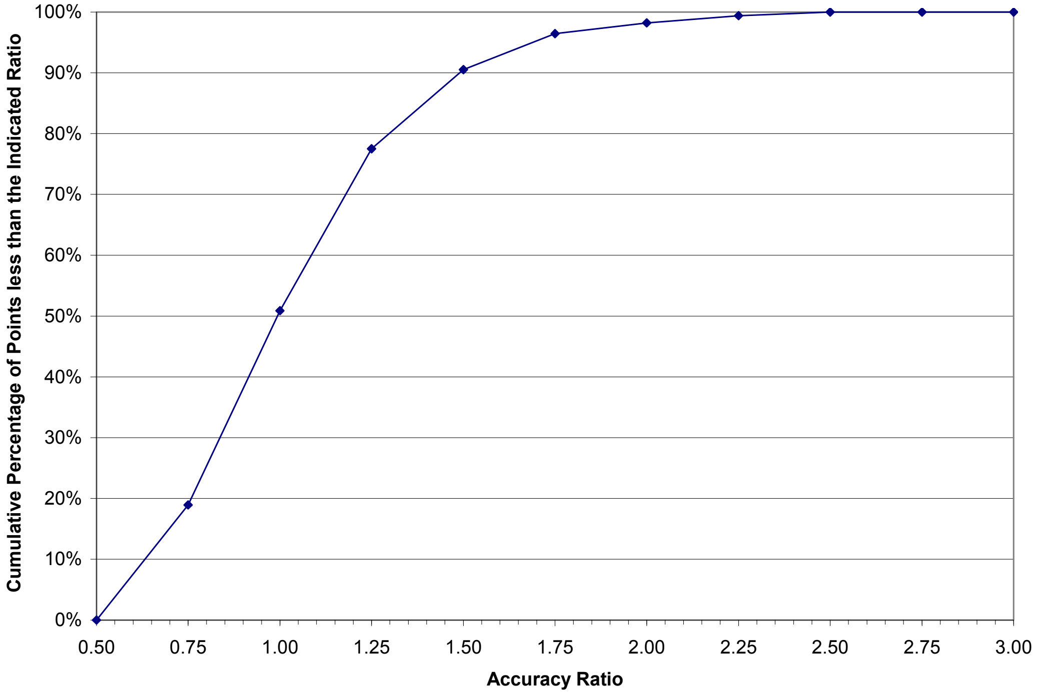


Figure 5.3.1: Cumulative Distribution of Accuracy Ratios for Segment Number 5

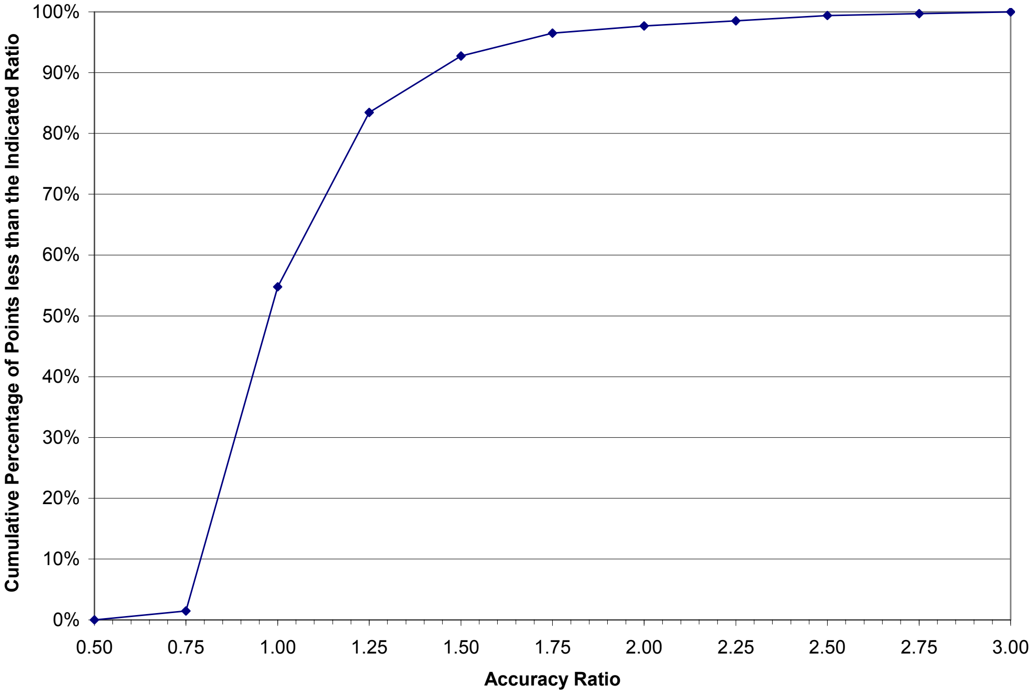


Figure 5.3.2: Cumulative Distribution of Accuracy Ratios for Segment Number 6

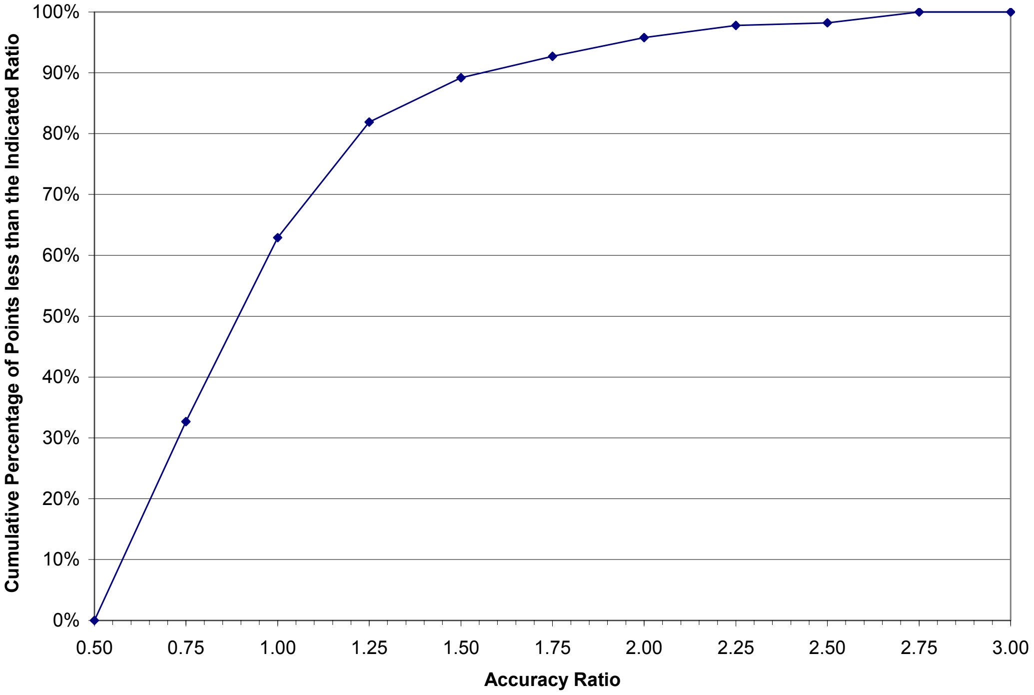


Figure 5.3.3: Cumulative Distribution of Accuracy Ratios for Segment Number 7

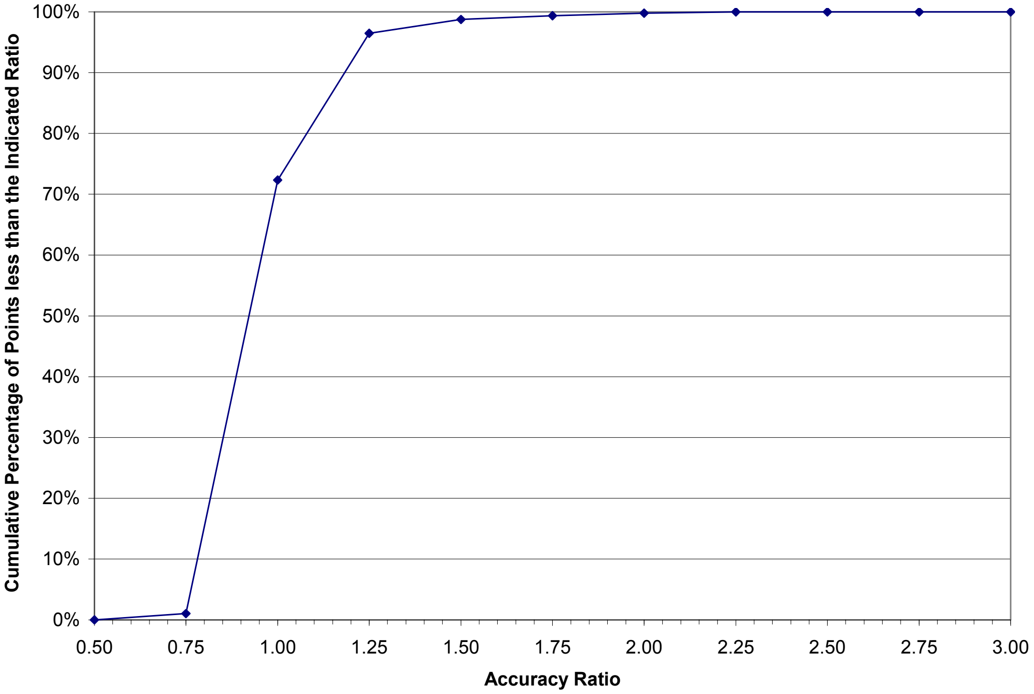


Figure 5.3.4: Cumulative Distribution of Accuracy Ratios for Segment Number 8

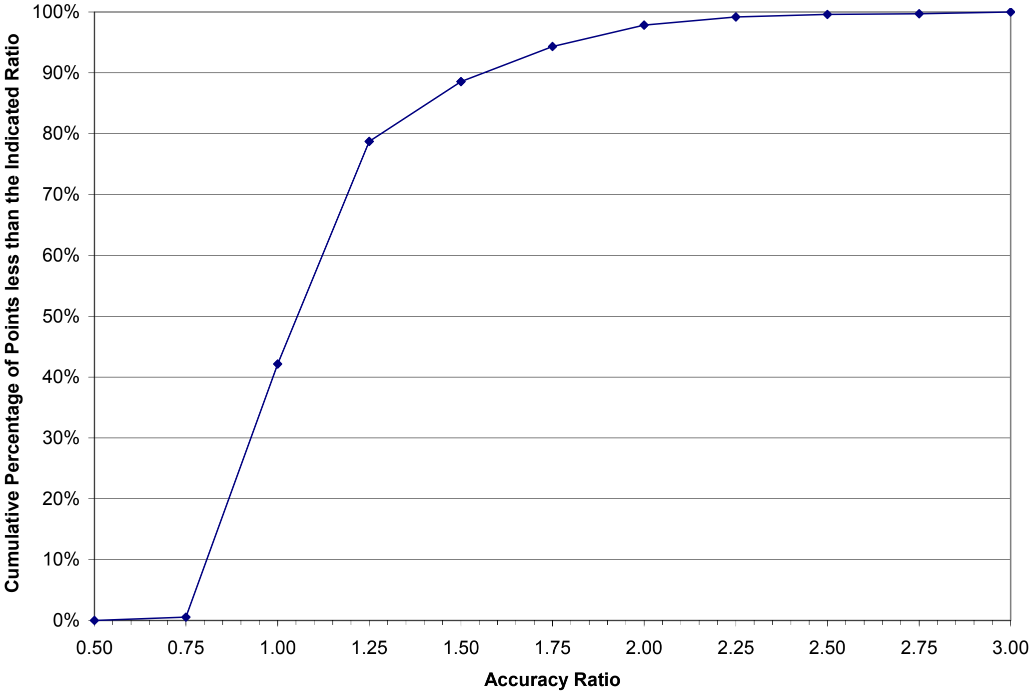


Figure 5.3.5: Cumulative Distribution of Accuracy Ratios for Segment Number 9

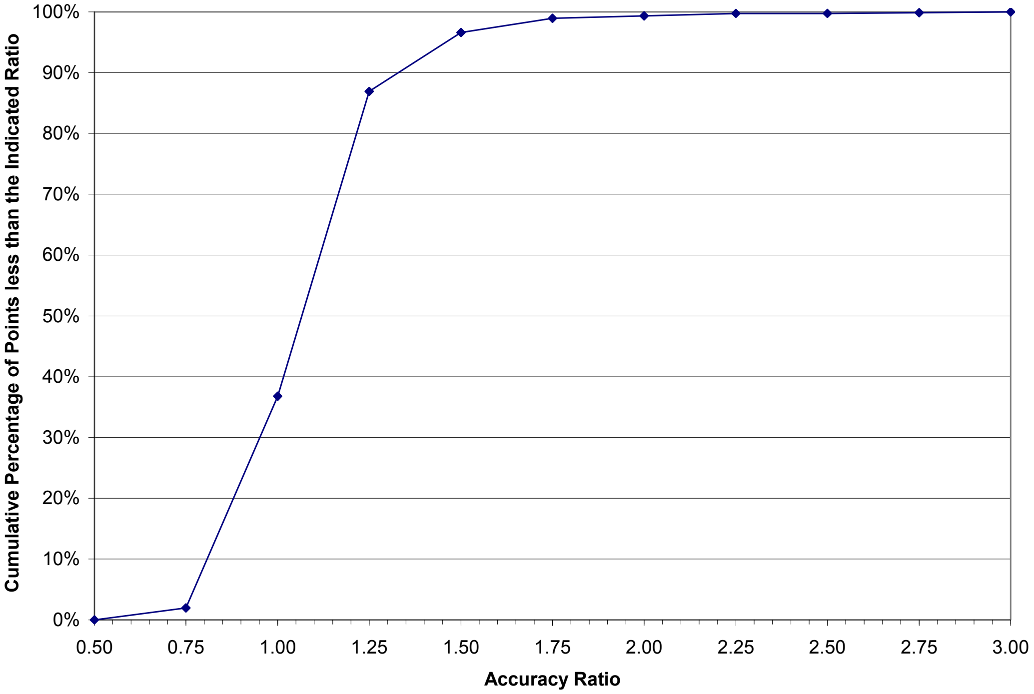


Figure 5.3.6: Cumulative Distribution of Accuracy Ratios for Segment Number 10

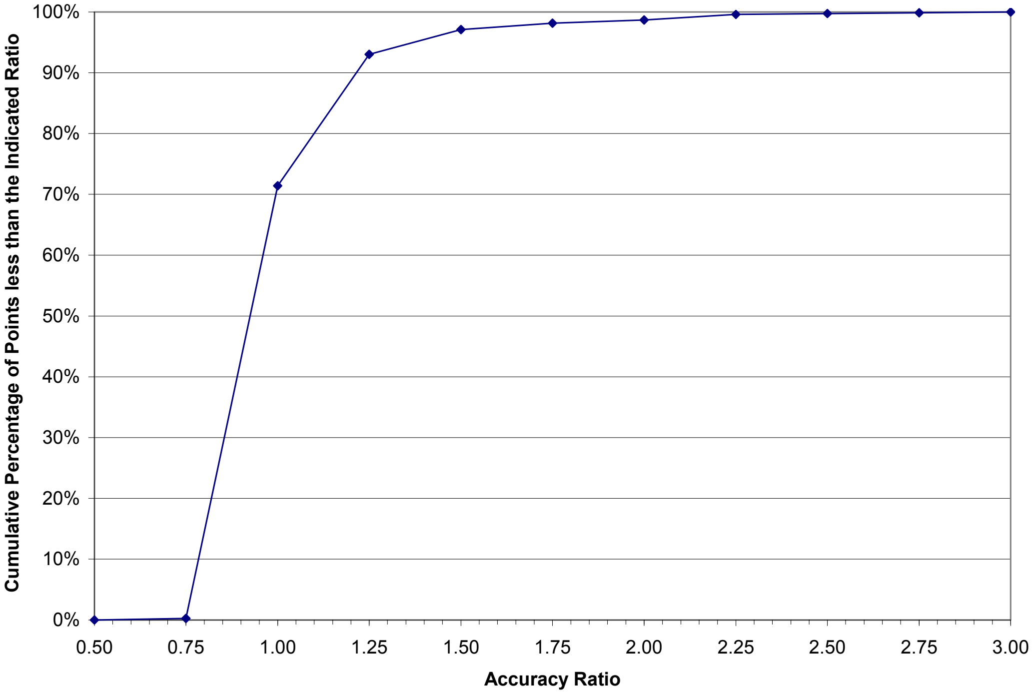


Figure 5.3.7: Cumulative Distribution of Accuracy Ratios for Segment Number 11

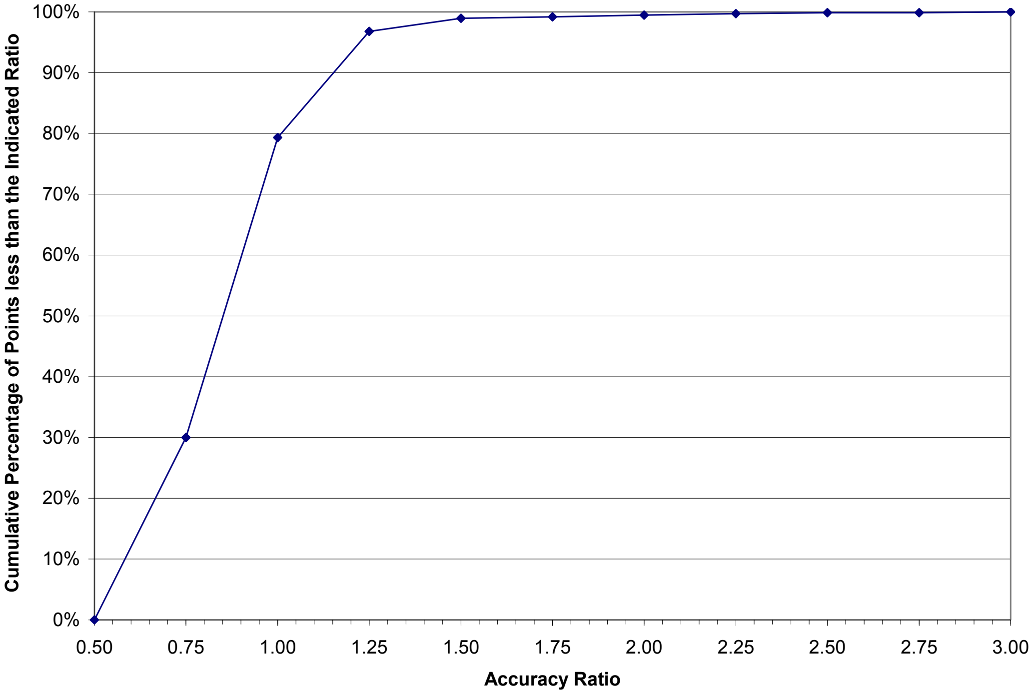


Figure 5.3.8: Cumulative Distribution of Accuracy Ratios for Segment Number 12

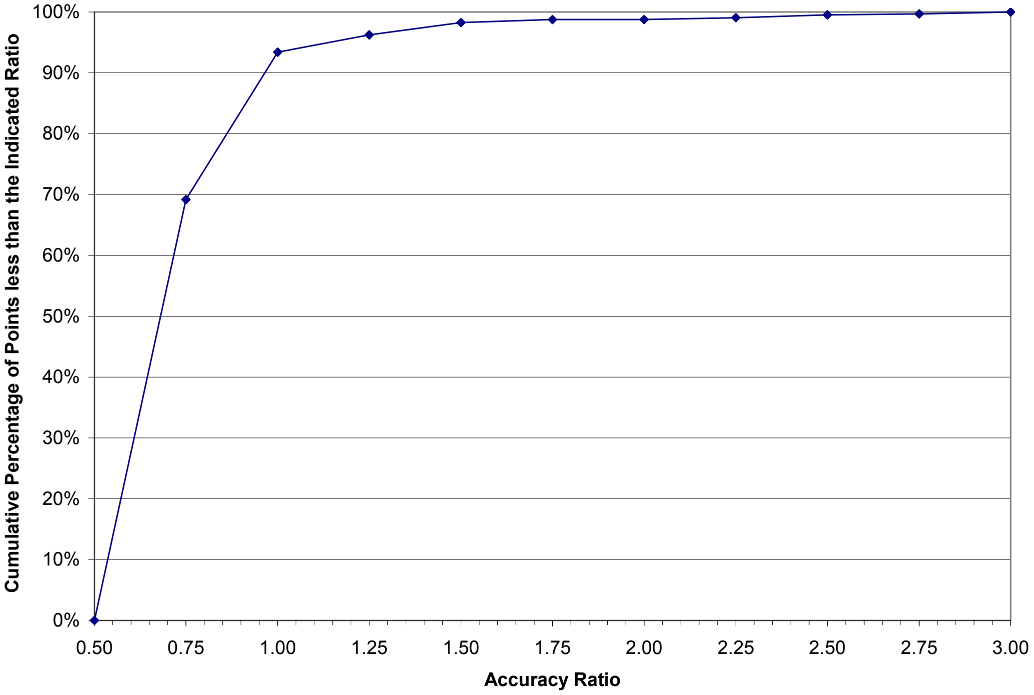


Figure 5.3.9: Cumulative Distribution of Accuracy Ratios for Segment Number 13

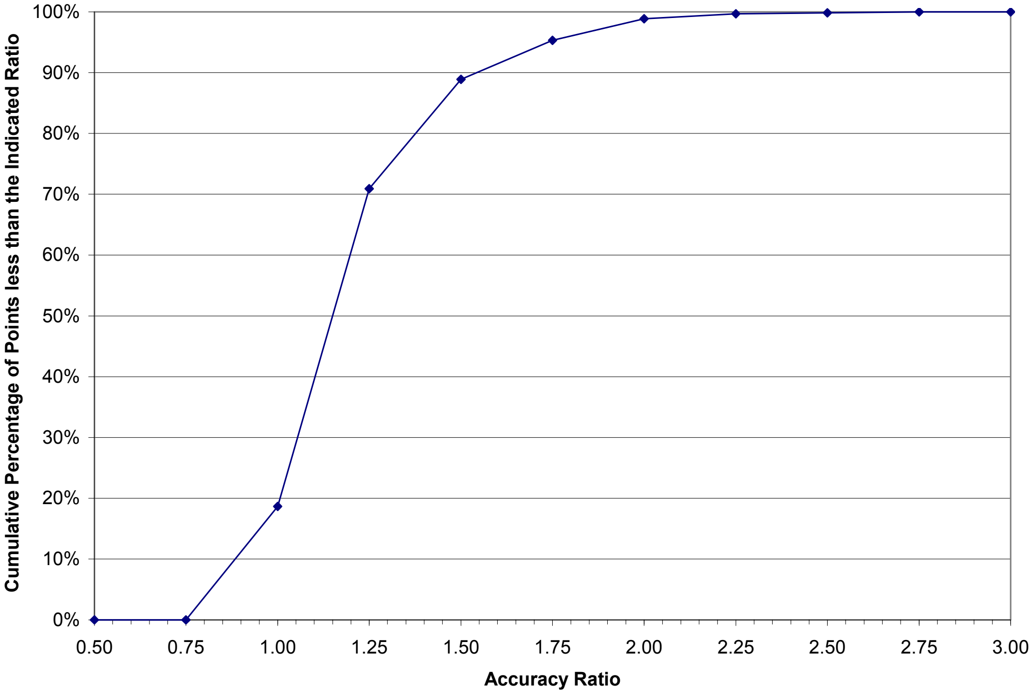


Figure 5.3.10: Cumulative Distribution of Accuracy Ratios for Segment Number 14

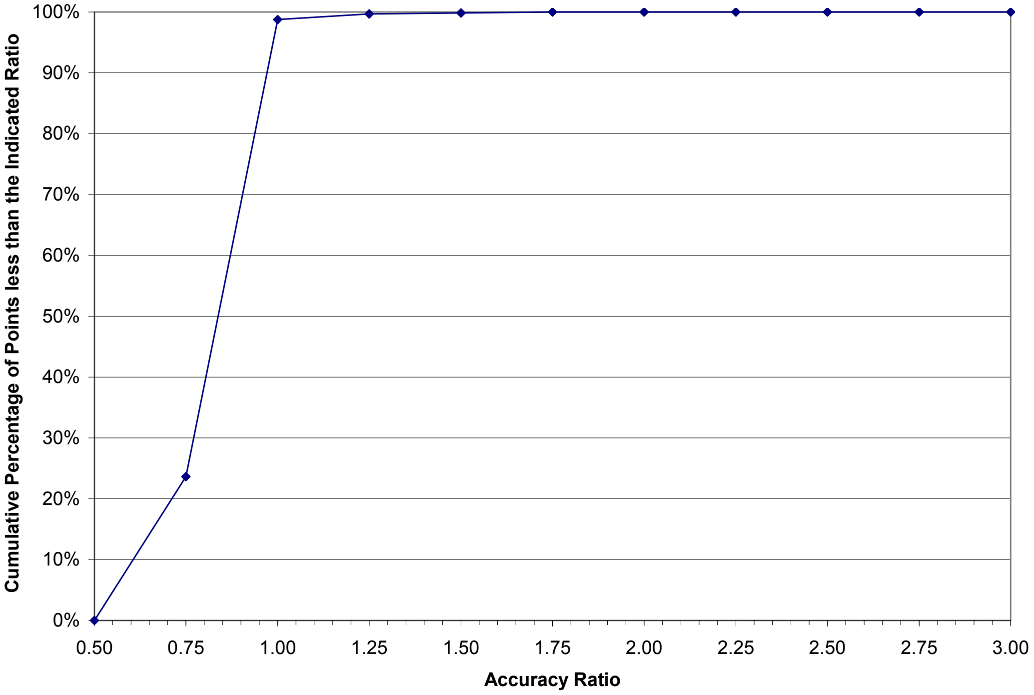


Figure 5.3.11: Cumulative Distribution of Accuracy Ratios for Segment Number 15

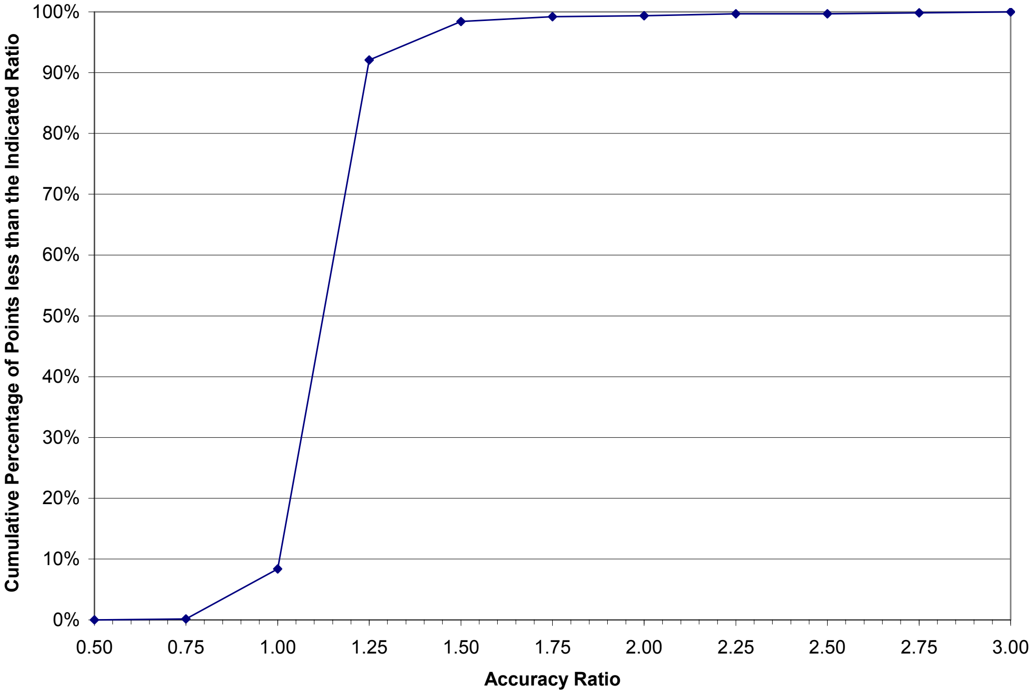


Figure 5.3.12: Cumulative Distribution of Accuracy Ratios for Segment Number 16

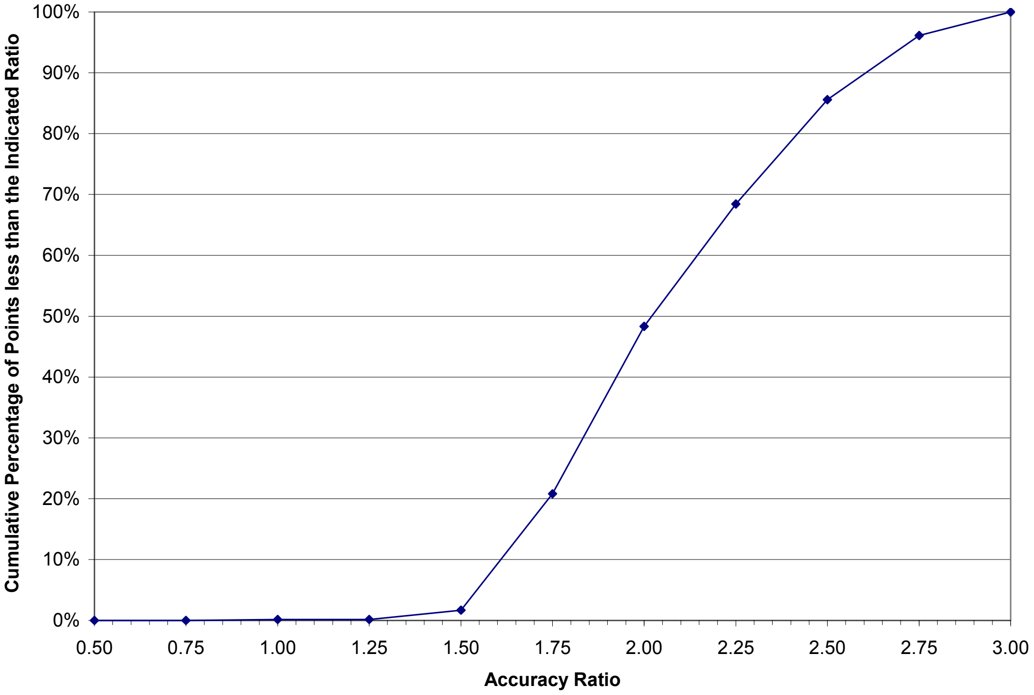


Figure 5.3.13: Cumulative Distribution of Accuracy Ratios for Segment Number 17

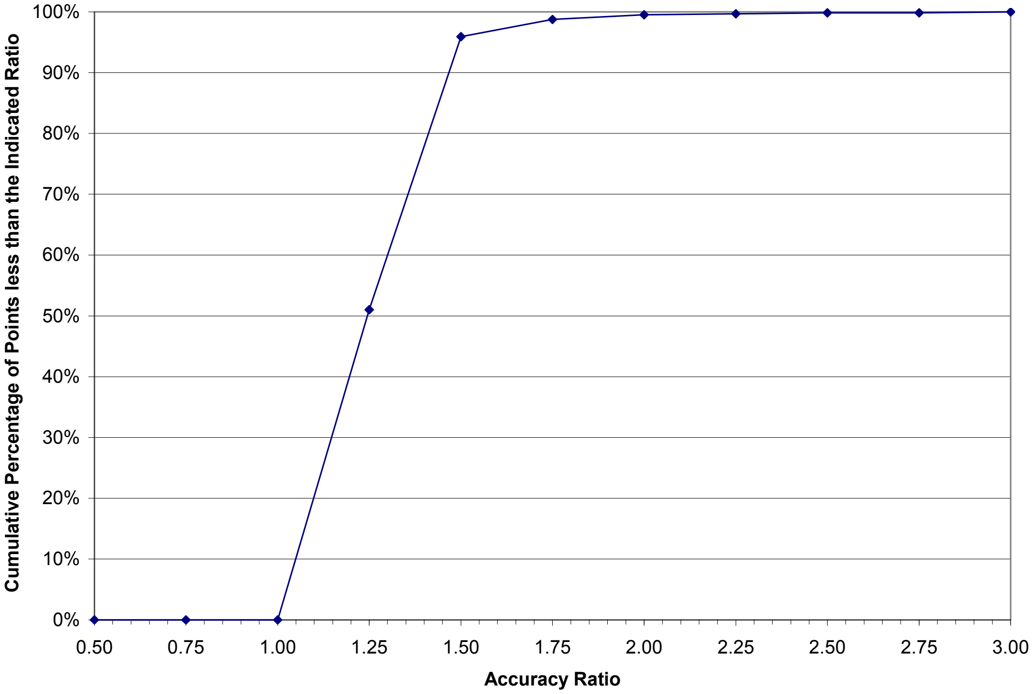


Figure 5.3.14: Cumulative Distribution of Accuracy Ratios for Segment Number 18

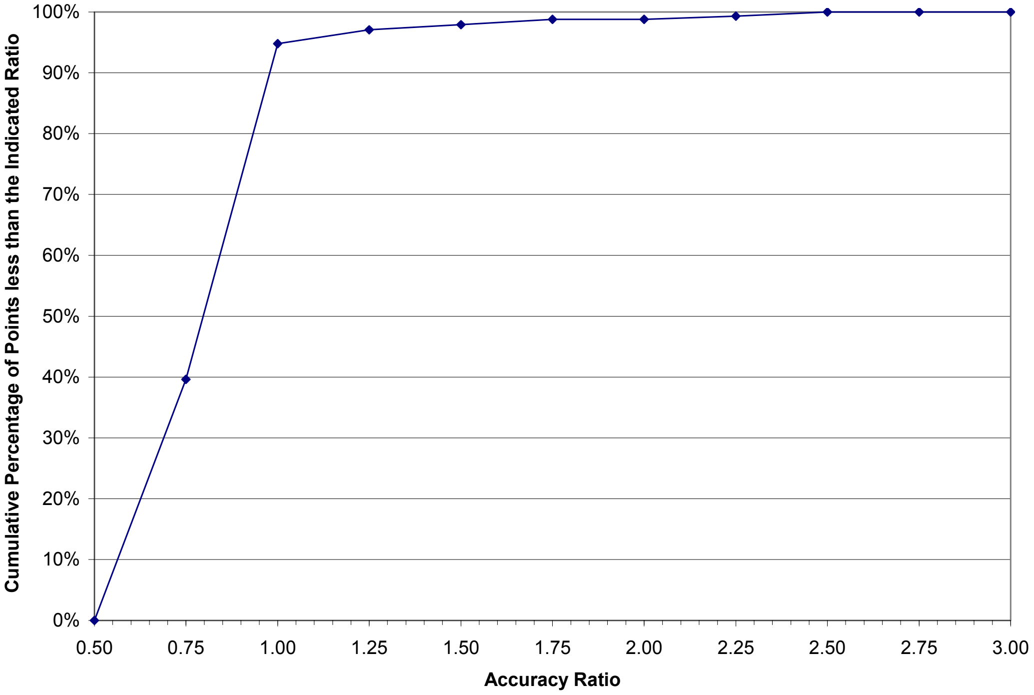


Figure 5.3.15: Cumulative Distribution of Accuracy Ratios for Segment Number 19

The cumulative distribution curves were used to determine the portion of points within 5 and 10 percent of the estimated haul time. Those points within 5 percent have accuracy ratios between 0.95 and 1.05, while those within 10 percent have ratios between 0.90 and 1.10. The mean accuracy ratio was calculated as a measure of the central tendency associated with each segment. These results are presented as Table 5.7.

Table 5.7: Summary of Results

Segment	Estimated Haul Time (s)	Mean Accuracy Ratio
Haul 5	29	1.04
Haul 6	36	1.09
Haul 7	16	1.04
Haul 8	38	0.99
Haul 9	40	1.11
Haul 10	19	1.08
Haul 11	27	1.00
Haul 12	26	0.85
Haul 13	32	0.78
Haul 14	68	1.19
Haul 15	10	0.75
Haul 16	23	1.09
Haul 17	8	2.00
Haul 18	17	1.24
Haul 19	26	0.79

5.5. Investigation into Influencing Factors

Statistical analysis techniques were used to investigate the influence of segment parameters, speed restrictions, and the operator on the results. Linear regression techniques were used to investigate the influence of segment parameters on the mean accuracy ratio. The Kruskal-Wallis (KW) test for homogeneity was used to determine whether speed restrictions resulted in significantly different ratios, and to assess the difference between ratios from different trucks.

5.5.1. Influence of Segment Parameters

Linear regression analyses were used to determine whether the mean accuracy ratio varied with segment length, grade, or estimated haul time. Each factor was analyzed separately with a significance level of 0.05, and the resulting scatter plots are presented as Figures 5.4 to 5.6. The statistical analyses results are summarized in Table 5.8.

Table 5.8: Linear Regression Analyses Results

Independent Variable	F-Statistic	p-value
Segment Length	0.27	0.6101
Segment Grade	*5.94*	*0.0299*
Estimated Haul Time	0.47	0.5028

The results indicate a statistically significant positive linear relationship between the mean calculated ratio and segment grade. This positive relationship indicates that VolvoSim tended to overestimate haul times for downhill segments, while underestimating haul times for uphill segments. While the relationship was significant, it was not strong. The p-value was 0.0299, less than the test significance level of 0.05, but the adjusted R^2 was 0.26, indicating that segment grade accounts for little of the total variation. No significant relationship was found between the mean calculated ratio and either estimated travel time or segment length.

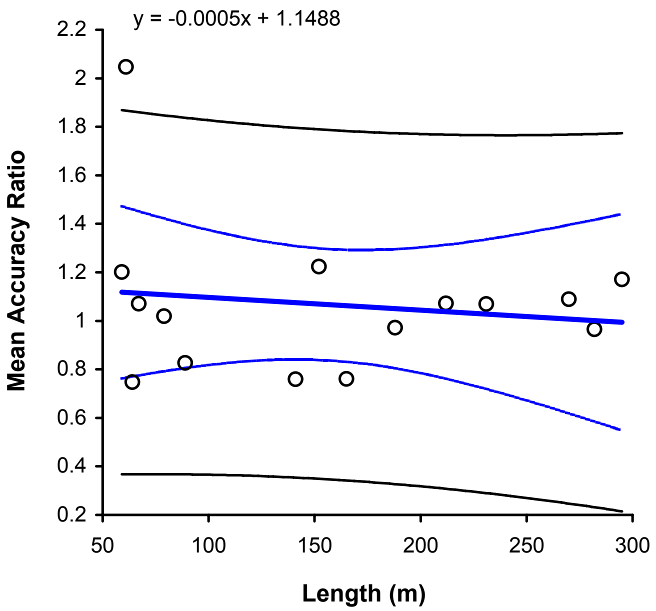


Figure 5.4 Scatter Plot of Mean Accuracy Ratio versus Segment Length

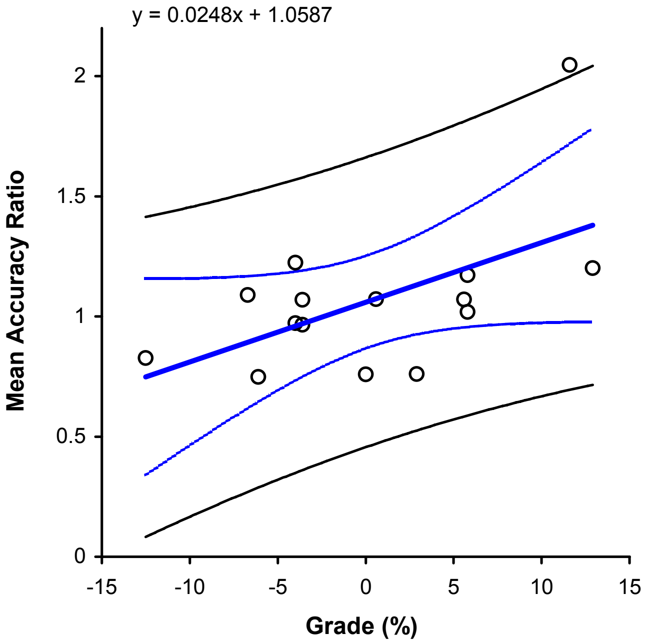


Figure 5.5 Scatter Plot of Mean Accuracy Ratio versus Segment Grade

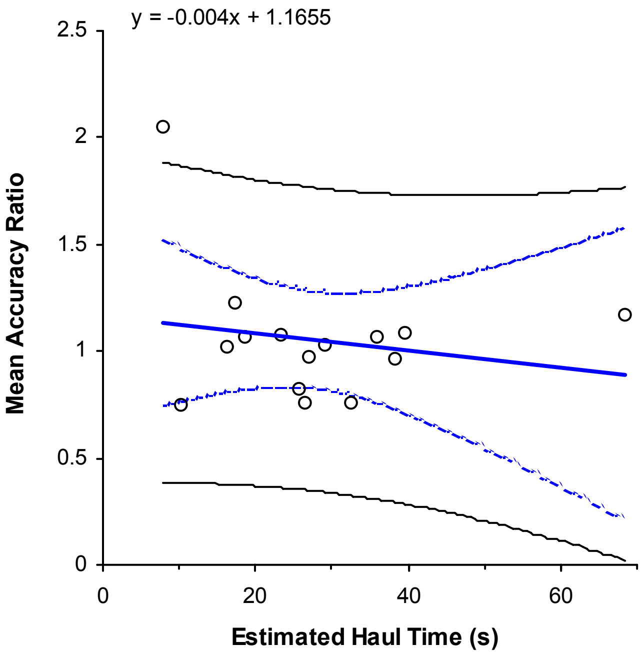


Figure 5.6 Scatter Plot of Mean Accuracy Ratio versus Estimated Haul Time

5.5.2. Influence of Speed Restrictions

The non-parametric Kruskal-Wallis (KW) test for homogeneity was used at a 0.05 significance level to determine whether there was a significant difference between mean accuracy ratios from segments with initial, final, and no speed restrictions. The calculated KW statistic was 1.56, which corresponds to a p-value of 0.4574. This indicates that the mean accuracy ratio does not significantly differ as a result of speed restrictions.

5.5.3. Influence of the Operator

The influence, if any that the operator has on the value of the mean accuracy ratio was also investigated. The same person operated each truck throughout the study period; only if one operator was away from work ill, did another person operate the truck. The KW test was used to determine whether the accuracy ratios were distributed differently for each truck. Table 5.9 summarizes the results; trucks identically distributed for a segment are marked with an "X". In some instances such as segment 13, the accuracy ratios for all trucks were identically distributed. This indicates that the operator had no significant influence on the accuracy ratio for this segment. In other instances such as segment 16, the accuracy ratios for all trucks were distributed differently. This indicates that perhaps the operator did have some influence on the accuracy ratio for this segment. However, when taken as a whole, the results indicate that there was no real systematic influence on the accuracy ratios by the operator.

Table 5.9: Summary of Identically Distributed Trucks by Segment

Segment	Length (m)	Grade (%)	Estimated Haul Time (s)	Mean Accuracy Ratio	Truck Number			
					1	2	3	4
Haul 5	58.9	12.9	29	1.04	x	x	x	
Haul 6	231.3	-3.6	36	1.09	x	x		x
Haul 7	78.6	5.8	16	1.04	x	x	x	
Haul 8	282.2	-3.6	38	0.99	x	x		
Haul 9	269.5	-6.7	40	1.11				
Haul 10	67.4	5.6	19	1.08	x			x
Haul 11	188	-4	27	1.00	x		x	
Haul 12	88.5	-12.5	26	0.85	x			x
Haul 13	164.6	2.9	32	0.78	x	x	x	x
Haul 14	294.5	5.8	68	1.19		x	x	
Haul 15	64.3	-6.1	10	0.75		x		x
Haul 16	212.2	0.6	23	1.09				
Haul 17	60.7	11.6	8	2.00	x	x	x	x
Haul 18	152.1	-4	17	1.24	x		x	x
Haul 19	140.9	0	26	0.79	x			x

5.6. Conclusion

Modeling and simulation techniques can be used to plan earthmoving operations, but should be verified and validated before being used as the basis for decisions. Productivity estimating software packages can be used to predict the performance of construction equipment. Such packages rely on mathematical models of the equipment and the expected conditions input by the analyst. Equipment models can be based on first principles and represent components of the equipment, or can be based on derived measures and represents the performance of the equipment as a whole.

The studied earthmoving operation was modeled and simulated using VolvoSim to produce haul time estimates. Field data was also collected in the form of short-interval GPS data and reduced to measures of actual haul time. An accuracy ratio was used to compare the actual and estimated haul times for each haul road segment studied. The estimated produced by VolvoSim were found to be reasonably accurate, as the estimated haul time was within 10 percent of the mean actual haul time for 8 of the 15 segments. Estimates for 14 of the 15 segments were within 25 percent of the mean actual haul time.

Factors potentially influencing the mean accuracy ratio, such as haul road parameters, speed restrictions, and equipment operator, were also investigated. It was found that a statistically significant linear relationship existed between the mean accuracy ratio and road segment grade. No significant relationship was found between the mean accuracy ratio and road segment grade, estimated haul time, presence of speed restrictions, or equipment operator.

Theoretical distributions were fit to the accuracy ratios observed for each segment. No single distribution best fit all the segments, but rather that the Log-Logistic, Pearson 5, and Pearson 6 distributions fit the data from most segments. This indicates that several distributions should be considered when fitting theoretical distributions to construction data. It was also found that distributions fit to a large number of observations were rejected based on the goodness-of-fit test results, but that those distributions fit to a relatively small number of distributions were not rejected. This agrees with the conclusion by Maio et al. [2000] that the results of the fitting procedure appear to be influenced by the dimension of the data set.

This work provided insight into the value of productivity estimating tools, such as VolvoSim. Identifying the significant relationship between mean accuracy ratio and segment grade provides and area of potential improvement that the program developers may wish to investigate further.

5.7. References

Bjornsson, H. C., and P. Sagert, (1994), “Multimedia based system for analyzing production organizations,” *Computing in Civil Engineering*, 2, pp1405-1411.

Frank, S. (2001), “Field Operations Analysis Laboratory – A Tutorial on the Review and Analysis Aspects of Operations Analysis,” Project and Report, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Gransberg, D. (1998), "A Review of Methods to Estimate Haul Fleet Production," *Civil Engineering,* v 40 n 3, pp. 31-35.

Hildreth, J. and M. Vorster, (in review (a)), "Development of a Short-Interval GPS Based Data Collection System", *Computer-Aided Civil and Infrastructure Engineering*, Blackwell Publishers, Boston.

Hildreth, J. and M. Vorster, (in review (b)), "A Methodology for the Reduction of Short-Interval GPS Data", *Journal of Construction Engineering and Management*, ASCE, Reston, VA.

Kannan, G. and M.C. Vorster (2000), "Development of an Experience Database for Truck Loading Operations," *Journal of Construction Engineering and Management*, ASCE, 165(3), pp201-209.

Law, A. M. and W. D. Kelton, (2000), *Simulation Modeling and Analysis*, 3rd Ed., McGraw-Hill Higher Education, Boston.

Oglesby, C. H., H. W. Parker, and G. A. Howell, (1989), *Productivity Improvement in Construction*, McGraw Hill Companies, New York, NY.

Perdomo, J. (2001), "Detailed Haul Unit Performance Model," MS Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Peurifoy, R. L., W.B. Ledbetter, and C. J. Schexnayder, (1996), *Construction planning, equipment, and methods,* 5th Ed., McGraw-Hill, New York.

Volvo (2002), A25C Operator's Manual, Volvo Construction Equipment Group, Asheville, NC.

VolvoSim (2001) – On-line Help, Volvo Construction Equipment Group, (http://www.runge.com.au/volvosim/).

CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

The principal goal of this research was to determine whether the information pertaining to position and speed contained in GPS data could be used to autonomously identify the critical times within a production cycle. The construction industry has traditionally relied on an observer, either in the field or office, to identify the critical times based on vast amounts of visual information. The observer uses positional information to determine when, in their mind, the studied equipment crosses an imaginary line. Likewise, the observer may use speed information to determine when the equipment either comes to a stop (speed equals zero), or resumes travel (speed greater than zero). There are two significant limitations to such manual time studies: the observer is required to make instantaneous decisions and the information available for making such decisions is strictly limited to that within the observer's field of view.

The research goal was achieved through the completion of the following objectives:

1. To design, build, and test a data collection tool to record a stream of GPS data at a short, fixed time interval
2. To show how the GPS data stream can be used to autonomously identify the times at which critical events occur in the production cycle by developing algorithms to identify the critical records
3. To demonstrate the value of the research by showing how it improves knowledge and provides additional insight into operations analysis studies.

The first objective to design, build, and test a data collection tool to record a stream of GPS data at a short, fixed time interval was completed by field evaluating commercially available tools and using the knowledge gained to design, build, and test a unique GPS data collection tool. The tool, shown as Figure 6.1, records data from a Garmin GPS35-LVC GPS receiver on CompactFlash media in comma delimited ASCII text format. Data is recorded at a user-specified, fixed time interval and includes date, time, latitude, longitude, and speed.

Figure 6.1: Fixed Configuration of the Data Collection Tool

The collected data was presented graphically in two forms. A scatter plot of horizontal position provided a plan view of the operation, while a plot of position versus time showed the cyclic nature of the operation and provided a means for distinguishing between operations by the distance between load and dump points.

A field evaluation of the tool resulted in the following improvements:

1. A mobile configuration of the tool was developed. This configuration, shown in Figure 6.2, is battery powered and is easily mounted on equipment using cargo straps. Considerable time and effort was required to mount the fixed configuration during the field evaluation, and the mobile configuration fills a very large need in terms of ease of installation.
2. A data file is created for each day data is recorded. Files containing data from multiple days of data collection are very large and require the user to separate the data prior to processing. Logging daily data files produces files of manageable size and eliminates a tedious manual task.
3. System firmware can be flash upgraded. This eliminates the need to return the system to the builder to be reprogrammed.

The second objective to show how the GPS data stream can be used to autonomously identify the times at which critical events occur in the production cycle by developing algorithms to identify the critical records was completed by developing the time identification modules (TIMs) to reduce the GPS data. The data is represented graphically to assist the user in reviewing and understanding the data. The TIMs were developed to identify the key records associated with activities performed under the following conditions:

1. Time spent with speed equal to zero within a specified fixed location
2. Time spent with speed equal to zero within a fixed distance of a moving location
3. Time required to travel through or between fixed areas

The Arrive and Depart TIM (ADTIM) was developed to identify the first (arrive) and last (depart) times at which the truck was stopped within a given distance of a user-defined point. ADTIM was used to determine dump times from the data collected.

The External Data TIM (EXTIM) was developed to identify the first (arrive) and last (depart) times at which the truck stops within a given distance of a moving point. Load times were determined from the collected data using the EXTIM.

The Fixed Boundary TIM (FBTIM) was developed to identify entry and exit times for a user-defined rectangular area. Haul and return times were determined from the collected data using the FBTIM.

Collectively, the TIMs were used to reduce data collected from a load and haul operation to measures of load, haul, dump, and return duration. It was postulated as to how the data can be used to provide insight into the activity and parent operation, as well for discrete event simulation by fitting theoretical statistical distributions to the activity durations.

The third objective to demonstrate the value of the research by showing how it improves knowledge and provides additional insight into operations analysis studies was completed by investigating correlations between performance times in construction operations and by using field data to verify the results obtained from productivity estimating tools.

The validity of the assumption that performance times can be independently and randomly selected from distributions according to the probability of occurrence was tested by investigating correlations between haul and return times determined from GPS data. The haul and return times were used to:

1. investigate the correlation between performance times (i.e. haul and return times)
2. investigate the auto-correlation in performance times (e.g. haul time and previous haul time)
3. investigate the correlation between performance time and the time of day (e.g. haul time and the time of day at which the activity began)

The results of the investigation are provided in Table 6.1.

Table 6.1: Normality and Correlation Test Results

Correlation	**Data Set**		**No. of Samples**	**Shapiro-Wilk**		**Spearman's Rank Correlation**	
	Dependent Variable	Independent Variable		W	p-value	r	p-value
Activity	Return	Haul	198	0.7442	<0.0001	-0.01	0.8460
Auto-Correlation	Haul	Previous	179	0.7754	<0.0001	0.12	0.1179
	Return	Previous	125	0.7042	<0.0001	0.09	0.3111
Time	Haul	Time	190	0.7838	<0.0001	-0.07	0.3301
	Return	Time	132	0.7068	<0.0001	**-0.19**	**0.0266**

The results indicate a statistically significant correlation between return time and time of day and non-significant interdependence in other data sets. The measured correlations were not sufficiently strong to conclude that studies based on the concept of independent data are flawed. However, the presence of statistically significant correlation supports the recommendation that performance data be analyzed for the presence of correlation before an assumption of independent data is made. Analyzing performance data for correlation requires large volumes of data that includes time of day so that data points can be paired for analysis.

The value of the field data collected was demonstrated by using the data to investigate the validity of haul times estimated from derived measures. Actual haul times determined from collected data were compared to estimates of haul time produced by VolvoSim. Observed and measured conditions were input into VolvoSim to accurately represent the conditions under which the observed load and haul earthmoving operation was performed. GPS data collected was reduced to measures of actual haul time for each haul road segment. These measures were compared to the haul time estimates to verify the accuracy of the estimates. Theoretical statistical distributions were fit to the distribution of measured haul times for each road segment. The data was also used to investigate the influence of factors such as segment length and grade, segment speed restrictions, and the operator. A statistically significant linear relationship was found between the mean accuracy ratio and the grade of each segment. The relationship indicated that VolvoSim tended to overestimate downhill haul times and underestimate uphill haul times, however, overall the estimates were found to be reasonably accurate.

The demonstrated values of the tool were its ability to collect the large volumes of time-stamped data required to investigate correlations among performance times and its ability to collect performance time data required to verify the results of productivity estimating tools.

This research performed to complete the objectives, when taken together, creates an integrated body of knowledge pertaining to the use of automated data collection for construction operations analysis. The concept of GPS data recorded at a short time interval to determine performance times from construction operations led to a review of available data collection tools and a field evaluation of the CrossCheck AMPS system. The knowledge gained from the test was used to design and build the hardware for a unique GPS data collection tool. The new tool was also field tested and software needed to reduce the large volumes of data was developed. Use of the hardware and software together as a system was demonstrated by using the developed TIMs to reduce the collected GPS data to measures of activity durations. The value of the system was demonstrated by using the data collected to investigate correlations among performance times and to verify the results produced by productivity estimating tools.

6.1. Contributions

This research has contributed to the body of knowledge pertaining to the collection and analysis of data from construction operations. The principal contribution is a system for recording a stream of GPS data and using that data to autonomously identify the times at which critical events occur within a production cycle. Due to the autonomous nature of the system, it is not subject to the same human limitations as traditional data collection systems, such as stopwatches and time-lapse video.

The field of view of the GPS system is vastly greater than that of an observer or camera in the field. Traditionally, data from operations performed in multiple locations are analyzed by observing each location separately or using multiple observers. The developed system allows analysis of an entire operation by a single analyst. The algorithms developed to identify the critical times produce results that are repeatable, thereby eliminating any variation resulting from multiple observers.

The system is capable of collecting large volumes of data that allow analysis in terms of distributions rather than point measurements. Data can be autonomously collected from many pieces of equipment for an entire day or multiple days and analyzed by a single person to produce large volumes of information in a timely manner. Data is collected by stopwatch studies in real time and analyzing a video recording of an entire day can take many hours. GPS data from an entire day can be analyzed in a few minutes.

Use of the system is technologically and economically feasible. The components and configuration of the data collection system are relatively simple. The components can be purchased for a few hundred dollars and the price of additional systems has been quoted to Virginia Tech by the original system developer at several hundred dollars each. The algorithms have been developed for Microsoft Excel, a widely used spreadsheet program.

6.2. Future Research

Future research stemming from this work will likely focus on improving the function and capabilities of the system and exploring the information contained in the gathered data. Potential research projects include both those with goals that may be immediately realized, and those with goals which might be pursued in the long term.

6.2.1. Potential Immediate Projects

Research that may be immediately pursued includes developing VBA algorithms to allow the analyst to graphically define the areas used to reduce the data. This can be accomplished by using object and chart properties to calculate the necessary coordinates. The coordinates would be used with the TIMs algorithms to reduce the data. This work will require formatting the plan view data plot and revising the TIMs algorithms to redirect them to the newly available coordinates. It is necessary to format the plot such that circle objects placed by the user define circular areas with respect to the data coordinate system.

Data is also available to immediately explore the information contained in the speed data. This research did not focus on use of the speed data. However, it did produce a large volume of speed data. The data can be used to generate speed profiles and calculate actual acceleration and deceleration rates. Speed profiles can be compared with those produced with programs such as VolvoSim. Actual acceleration and deceleration rates can be compared with values used to calculate estimates of speed and travel time.

6.2.2. Potential Long Term Projects

Research projects that may be pursued long term include improvements to both the system hardware and software, as well as employing the system to study various aspects of production operations. New technologies may be used to develop a smaller, more energy efficient and more accurate data collection system. Computing capabilities may be enhanced to allow on-board data reduction, which will result in smaller volumes of data to be transmitted. One can envision an analyst walking by a fleet of parked equipment and wirelessly downloading cycle times from data collected during the previous day, week or month and using that data to identify modifications to the operation to increase production or decrease cost.

Potential software enhancements include developing the algorithms to reduce the data within a GIS environment. The GIS environment will allow the data to be incorporated with other data formats such as digital terrain models and aerial photographs. GIS softwares are well suited for managing large volumes of GPS position data and can be automated to reduce the data through filters and queries. Analysts can graphically define areas of interest. They are capable of graphically representing the reduced data in the form of histograms.

The position and time data collected may be used to drive a time-compressed animation of an analyzed operation. One can envision an animation of equipment icons traveling on a digital aerial photo draped over a digital terrain model. Data from an entire day or week could be visually reviewed in the span of a few minutes. While this would not provide the complete view of field operations that video recording does, it would allow the analyst to visualize the data.

The combination of time, position, and speed data allows for the study of operations performed under varying conditions to determine the influence of the conditions on measures such as production and cost. Making use of other available data such as weather records and data that can be readily gathered such as equipment type, roadway super elevation, the list of conditions that can be studied include roadway geometry, super elevation, road surface conditions, weather conditions, operator fatigue, operator experience, and load/dump configurations. For example, data collected from equipment operated by well-trained personnel could be used to produce acceleration measures and speed profiles. These might be compared with data from other equipment and used to identify patterns that can be used to train other personnel. Or rubber-tired and track loaders might be studied performing various loading configurations in varying material types under various weather conditions to identify the rules for equipment selection. The combinations of conditions under which operations can be studied is at this point practically limitless, and application of the system most likely lead to more possibilities.

CHAPTER 7. BIBLIOGRAPHY

Abdelhamid, T.S., J. Everett (1999). "Time Series Analysis for Construction Productivity Experiments," *Journal of Construction Engineering and Management*, ASCE, 125(2), ASCE, Reston, VA.

AbouRizk, S. M., A. Hijazi, and D. W. Halpin (1989), "Effect of input modeling on simulation output," in *Proceeding of Computing in Civil Engineering*, T. O. Barnwell (Ed.), pp605-612

AbouRizk, S. M. and D. W. Halpin (1990), "Probabilistic Simulation Studies for Repetitive Construction Processes," *Journal of Construction Engineering and Management*, ASCE, 116(4), pp575-594.

AbouRizk, S. M. and D. W. Halpin, (1992a), "Modeling input data for construction simulation," in *Proceedings of Computing in Civil Engineering and Geographic Information Systems Symposium*, pp1147-1154.

AbouRizk, S. M. and D. W. Halpin, (1992b), "Statistical properties of construction duration data," *Journal of Construction Engineering and Management*, ASCE, 118(3), pp525-544.

Ackroyd, N. (1998), "Earthworks Scheduling and Monitoring Using GPS," in the *Proceedings of the 1998 Trimble User Conference*, pp. 1-6.

Ahmad, I. U. (1999), "Information as a Resource," *Journal of Management in Engineering*, ASCE, 15(4), ASCE, Reston, VA.

Ahuja, H.N. and V. Nandakumar, (1985) "Simulation Model to Forecast Project Completion Time," *Journal of Construction Engineering and Management*, ASCE, 111(4), ASCE, Reston, VA.

Alexander, J. F., R. J. Coble, and B. R. Elliott (1997), "Hand-Held Communication for Construction Supervision," in the *Proceedings of the 1997 Construction Congress*, ASCE, Reston, VA., pp. 972-979.

Altman, D. G. (1991), *Practical Statistics for Medical Research*, Chapman and Hall, London.

Angelo, W. (2000), “Order to Unscramble GPS Signals to Boost Equipment Productivity” *Engineering News Record*, 244(19), p 32.

Anonymous (1986), “Intelligent equipment: machines under electronic controls,” *World Construction*, v 39 n 3, Mar 1986, p 112-114, 116.

Back, W.E. and L.C. Bell, (1995) “Monte Carlo Simulation as Tool for Process Reengineering,” *Journal of Management in Engineering*, ASCE, 11(5), ASCE, Reston, VA.

Benjamin, J. and C. Cornell (1970), *Probability, Statistics, and Decision for Civil Engineers,* McGraw-Hill, Inc., New York, NY.

Borgelt, S. C., J. Harrison, K. Sudduth, and S. Birrell (1996), “Evaluation of GPS for Applications in Precision Agriculture,” *Applied Engineering in Agriculture*, v 12 n 6, pp 633-638.

Bjornsson, H. C., and P. Sagert, (1994), “Multimedia based system for analyzing production organizations,” *Computing in Civil Engineering*, 2, pp1405-1411.

Burton, F. M., (1991), “Method for measuring construction productivity,” 1991 Transactions of American Association of Cost Engineers, Morgantown, WV, paper no. L.3.

Carr, R. (1979), “Simulation of Construction Project Duration,” *Journal of Construction Division*, 105(CO2), pp117-128.

Carrier, J. (1998), “Satellites Guiding Industries on the Move,” New York Times, September 28.

Caterpillar (1999), *Systems Operation Testing and Analysis – Vital Information Management System (VIMS),* Caterpillar, Inc., Peoria, IL.

Cheng, M. and S. Yang, (2001), “GIS-Based Cost Estimates Integrating with Material Layout Planning,” *Journal of Construction Engineering and Management*, ASCE, 127(4), pp291-299.

Chironis, N. P., (1984), “Detectors Estimate Truck Overloads,” *Coal Age*, 89(8), p125.

Chironis, N. P., (1985), "Computer monitors and controls all truck-shovel operations," *Coal Age*, 90(3), pp50-55.

Chironis, N. P., (1987), "Microprocessor-aided equipment proves productive and reliable," *Coal Age*, 92(7), pp48-56.

Christian, J. and D. Hachey (1996), "A computer-aided system to improve production rates in construction," *Advances in Engineering Software*, v 25, pp207-213.

Conover, W. J. (1980), *Practical Nonparametric Statistics*, 2nd Ed., John Wiley & Sons, New York.

Cor, H. and J.C. Martinez (1999), "A Case Study in the Quantification of a Change in the Conditions of a Highway Construction Operation," in *Proceedings of the 1999 Winter Simulation Conference*, P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans (Ed.), Phoenix, AZ, pp1007-1009.

Crane, C.D., D. Armstrong, and A. Rankin (1995), "Autonomous Navigation of Heavy Construction Equipment," *Microcomputing in Civil Engineering,* 10(5), pp. 357-70.

D'Agostino, R. and M. Stephens (1986), *Goodness-of-Fit Techniques*, Marcel Dekker, Inc., New York.

D'Este, G., R. Zito, and M. Taylor (1999), "Using GPS to Measure Traffic System Performance," *Computer-Aided Civil and Infrastructure Engineering,* v 14, pp255-265.

DeMers, M. N., (1997), *Fundamentals of geographic information systems,* Wiley & Sons, New York.

Dinovitser, A. and M. Taylor (1997), "Simulating Truck Performance in Mining and Earthmoving," in the *Proceedings of the International Simulation Modelling Congress MODSIM9*, Hobart, Tasmania, vol 4, pp.1391-1396.

ESRI, (1998), *Getting to know ArcView GIS,* Environmental Systems Research Institute, Inc., Redlands, CA.

Everett, J. G., H. Halkali, T. G. Schlaff, (1998), "Time-lapse video applications for construction project management," *Journal of Construction Engineering and Management*, ASCE, 124(3), pp204-209.

Fente, J., K. Knudson, and C. Schexnayder (1999), "Defining a Beta Distribution for Construction Simulation," in *Proceedings of the 1999 Winter Simulation Conference*, P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans (Ed.), Phoenix, AZ, pp1010-1015.

Fente, J., C. Schexnayder, and K. Knudson (2000), "Defining a Probability Distribution Function for Construction Simulation," *Journal of Construction Engineering and Management*, ASCE, 126(3), pp234-241.

Frank, S. (2001), "Field Operations Analysis Laboratory – A Tutorial on the Review and Analysis Aspects of Operations Analysis," Project and Report, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Garmin (2000), "*GPS Guide for Beginners*" Garmin International, Inc., Olathe, Kansas.

Geaney, C.M. (1997), "A Strategy for the Development of a Cost Estimating Database," *Cost Engineering*, 39(4), pp30-35.

Gonzalez-Quevedo, A. (1995), "Sensitivity Analysis of Construction Simulation," *Computing in Civil Engineering*, v 2, pp1427-1434.

Goodchild, B and S. Fairhead (1993) "Using GPS to enhance the control and effectiveness of a public transport system," *Electronic Engineering*, v 65 n 796, pp 49-53.

Guo, B. and A. Poling (1995), "Geographic Information Systems/Global Positioning Systems Design for Network Travel Time Study" *Transportation Research Record*, n 1497, Jul 1995, pp135-139.

Gransberg, D. (1998), "A Review of Methods to Estimate Haul Fleet Production," *Civil Engineering,* v 40 n 3, pp. 31-35.

Hajjar, D. and AbouRizk, S. M., (1997), "AP2-Earth: A simulation based system for the estimating and planning of earthmoving operations," in *Proceedings of the 1997 Winter Simulation Conference*, S. Andradottir, K. J. Healy, D. H. Whithers and B. L. Nelson (Ed.), San Diego, CA, pp1103-1110.

Halpin, D. W., (1977), "CYCLONE- Methods for modeling job site processes," *Journal of Construction Division*, 103(3), pp489-499.

Halpin, D. W. and L. S. Riggs, (1992), *Planning and analysis of construction operations*, John Wiley & Sons, Inc, New York, NY.

Halton, D. and R. Johnson (1984), "The Road to Increased Mine Productivity," in the *1984 Fall Industrial Engineering Conference Proceedings,* IIE, Norcross, GA, pp499-506.

Hildreth, J. and M. Vorster, (in review), "Development of a Short-Interval GPS Based Data Collection System", *Computer-Aided Civil and Infrastructure Engineering*, Blackwell Publishers, Boston.

Hildreth, J. and M. Vorster, (in review), "A Methodology for the Reduction of Short-Interval GPS Data", *Journal of Construction Engineering and Management*, ASCE, Reston, VA.

Howard, C. J., (1991), "Project-Specific Knowledge Bases in AEC Industry," *Journal of Computing in Civil Engineerin*g, ASCE, 5(1), pp25-41.

Ioannou, P. G., (1990), *UM-CYCLONE- Discrete Event Simulation System—Reference Manual*, UMCEE Report No. 89-11, Dept. of Civil Engineering, University of Michigan, Ann Arbor, MI.

Ioannou, P. G. and J. C. Martinez, (1996a) "Comparison of Construction Alternatives Using Matched Simulation Experiments," *Journal of Construction Engineering and Management*, ASCE, 112(3), ASCE, Reston, VA.

Ioannou, P.G. and J.C. Martinez (1996b), "Simulation of Complex Construction Processes," in *Proceedings of the 1996 Winter Simulation Conference*, J.M. Charnes, D.J. Morrice, D.T. Brunner, and J.J. Swain (Ed.), San Diego, CA, pp1329-1336.

Ioannou, P.G. (1999), “Construction of a Dam Embankment with Nonstationary Queues,” in *Proceedings of the 1999 Winter Simulation Conference*, P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans (Ed.), Phoenix, AZ, pp921-928.

Jeljeli, M. N., J. S. Russell, G. Meyer, and A. P. Vonderohe, (1993), “Potential applications of geographic information systems to construction industry,” *Journal of Construction Engineering and Management*, ASCE, 119(1), pp72-86.

Kannan, G. (1998), “Automated Data Collection for Earthwork Estimating,” in *Proceeding of Computing in Civil Engineering*, K. Wang (Ed.), 457-465.

Kannan, G. (1999), “A Methodology for the Development of a Production Experience Database for Earthmoving Operations using Automated Data Collection,” PhD Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Kannan, G. and M.C. Vorster (2000a), “Development of an Experience Database for Truck Loading Operations,” *Journal of Construction Engineering and Management*, ASCE, 165(3), pp201-209.

Kannan, G. and M.C. Voster (2000b), “Image Database for Characterizing Field Cycle Time Data” in the *Proceedings of the 2000 Construction Congress*, ASCE, Reston, VA., pp. 59-67.

Kannan, G., M. Vorster, J. Martinez, and T. Rakes (1999a), “Improved Mine Management Through Data Mining,” presented at SME Mining Conference, Denver, CO.

Kannan, G., M. Vorster, and J. Martinez (1999b), “Developing the Statistical Parameters for Simultaneous Variation in Final Payload and Total Load Time,” in *Proceedings of the 1999 Winter Simulation Conference*, P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans (Ed.), Phoenix, AZ, pp1016-1022.

Kimmance, J. P., M. P. Bradshaw, and H. H. Seetoh (1999), “Geographical Information System (GIS) Application to Construction and Geotechnical Data Management on MRT Construction Projects in Singapore,” *Tunnelling and Underground Space Technology,* v 14, n 4, pp469-479.

Korte, G.B., (1994), *The GIS Book,* 3rd Ed., On-World Press, Sante Fe, NM.

Law, A. M. and W. D. Kelton, (2000), *Simulation Modeling and Analysis*, 3rd Ed., McGraw-Hill Higher Education, Boston.

Leick, A. (1992), "Beyond Push-Button GPS," *Civil Engineering*, v 62 n 6, Jun 1992, pp75-76.

Lewis, Chris R. and Cliff J. Schexnayder (1986), "Production Analysis of the Cat 245 Hydraulic Hoe" *in the Proceedings of the Earthmoving and Heavy Equipment Conference,* New York, NY, p 88-94.

Li, C., A. Oloufa, and R. Thomas (1996), "A GIS-based system for tracking pavement compaction," *Automation in Construction,* v 5, pp51-59.

Liu, L. (1995), "Construction Crane Operation Simulation," *Computing in Civil Engineering,* v 2, pp1420-1426.

Lluch, J. and D. W. Halpin, (1982), "Construction operations and microcomputers," *Journal of Construction Division*, ASCE, 108(**CO**1), pp129-145.

Lothon, A. and S. Akel (1996), "Techniques for Preventing Accidental Damage to Pipelines," in the *Proceedings of the 1st International Pipeline Conference,* v2, American Society of Mechanical Engineers, New York, NY, pp. 643-650.

Maio, C., C. Schexnayder, K. Knutson, and S. Weber (2000), "Probability Distribution Functions for Construction Simulation" *Journal of Construction Engineering and Management*, ASCE, 126(4), pp285-292.

Martinez, J. C. and P. G. Ioannou, (1995), "Advantages of the Activity Scanning Approach in the Modeling of Complex Construction Processes," in *Proceedings of the 1995 Winter Simulation Conference*, Society for Computer Simulation, San Diego, CA.

Martinez, J. C. and P. G. Ioannou, (1999), "General-purpose systems for effective construction simulation," *Journal of Construction Engineering and Management*, ASCE, 125(4), pp265-276.

Martinez, J. C., (1996), Stroboscope: State and resource based simulation of construction processes, Ph.D. Thesis submitted to the University of Michigan.

Martinez, J. C., (1998), "EarthMover - Simulation Tool for Earthwork Planning," in *Proceedings of the 1998 CIB W78 Conference*, Department of Construction Management and Economics, Royal Institute of Technology, Stockholm, Sweden.

McCahill, D. F. and L. E. Bernhold, (1993), "Resource-oriented modeling and simulation in construction," *Journal of Construction Engineering and Management*, ASCE, 119(3), September, pp590-606.

McCahill, D. F. and L. E. Bernhold (1993), "Simulation for Construction Planning and Control," in *Proceedings of Construction Congress I: Constructed Project Excellence,* ASCE, Reston, VA, pp349-359.

Means, (2000), *Means Heavy Construction Cost Data*, R. S. Means Co. Inc, Kingston, MA.

Miller, R. G. (1986), *Beyond ANOVA, Basics of Applied Statistics*, John Wiley & Sons, New York.

Moffitt, F. H., and H. Bouchard, (1992), *Surveying*, *9th Edition*, Harper Collins Publishers, New York, NY.

Mueller, G.R. and J. T. Knepler (1984), "A Digital Instrumentation System for Agricultural and Construction Equipment," in *ASAE Publication 84-8 Proceedings of the National Conference on Agricultural Electronics Applications*, Dec 11-13, 1983, Hyatt Regency Illinois.

Murakami, E. and D. Wagner (1999), "Can using global positioning system (GPS) improve trip reporting?" *Transportation Research Record*, n 1660, pp 149-165.

Nasland, D.K. and D.P. Johnson, (1996), "Real-Time Construction Staking," *Civil Engineering*, ASCE, 66(6), pp 46-9.

Oglesby, C. H., H. W. Parker, and G. A. Howell, (1989), *Productivity Improvement in Construction*, McGraw Hill Companies, New York, NY.

Oloufa, A. (1993), “Modeling and simulation of construction operations,” *Automation in Construction,* v 1, pp351-359.

Oloufa, A.A., Do, W. and Thomas, W.R. (1997) “Automated Monitoring of Compaction using GPS,” in the *Proceedings of the 1997 Construction Congress*, ASCE, Reston, VA., pp. 1004-1011.

Ott, R. and M. Longnecker (2001), *An Introduction to Statistical Methods and Data Analysis,* 5th Ed., Duxbury, Pacific Grove, CA.

Paksoylu, M, D. Halpin, and D. Abraham (1995), “Productivity Analysis Using Simulation,” *Computing in Civil Engineering,* v 1, pp743-750.

Pampagnin, L., Peyret and Garcia G. (1998), “Architecture of a GPS-Based Guiding System for Road Compaction,” in the *Proceedings of the 1998 IEEE International Conference on Robotics and Automation*, Leuven, Belgium, pp. 2242-2427.

Paulson, Jr., B. C., (1978), “Interactive graphics for simulating construction operations,” *Journal of Construction Division*, ASCE, 104(**CO**1), pp69-76.

Paulson, Jr., B. C., S. A. Douglas, A. Kalk, A. Touran, G. A. Victor, (1983), “Simulation and analysis of construction operations, *Journal of Technical Topics in Civil Engineering*, ASCE, 109(2), pp89-104.

Paulson, Jr., B. C., W. T. Chan and C. C. Khoo, (1987), “Construction operations simulation by microcomputer,” *Journal of Construction Engineering and Management*, ASCE, 113(2), pp302-314.

Parker, H. W. and C. H. Oglesby, (1972), Methods Improvement for Construction Managers, McGraw-Hill Company, New York, NY.

Perdomo, J. (2001), “Detailed Haul Unit Performance Model.” MS Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Peurifoy, R. L., W.B. Ledbetter, and C. J. Schexnayder, (1996), *Construction planning, equipment, and methods,* 5th Ed., McGraw-Hill, New York.

Peyret, F., D. Betaille, and G. Hintzy (2000), “High-precision application of GPS in the field of real-time equipment positioning,” *Automation in Construction*, v 9, pp299-314.

Phelps, R.W. (1998), “GPS – Now for Excavation Morenci Mine Improves Ore Control,” *Engineering & Mining Journal*, 199(7), pp. 34-6.

Quiroga, C. and D. Bullock (1998), “Travel time studies with global positioning and geographic information systems: an integrated technology,” *Transportation Research Part C-Emerging Technologies,* v 7 n 2-3, Apr-Jun 1999, p 101-127.

Sangarayakul, B. (1998), “Use of Simulation to Analyze Block Manufacturing Methods.” MS Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Schexnayder, C., Weber, S. L., and Brooks, B. T., (1999), “Effect of truck payload weight on production,” *Journal of Construction Engineering and Management*, ASCE, 125(1), pp1-7.

Senior, B. A. and A. Swanberg-Mee, (1997), “Activity Analysis using computer-processed time lapse video,” in the *Proceedings of the 1997 Construction Congress*, ASCE, Reston, VA.

Shapiro, S., Wilk, M., and Chen, H. (1968), “A Comparative Study of Various Tests for Normality,” *Journal of the American Statistical Association*, v 63 n 324, pp 1343-1372.

Sheskin, D. J. (1997), *Handbook of Parametric and Nonparametric Statistical Procedures*, CRC Press, Boca Raton, FL.

Shi, J. and S. AbouRizk (1998), “An Automated Modeling System for Simulating Earthmoving Operations,” *Computer-Aided Civil and Infrastructure Engineering,* v. 13, pp 121-130.

Shields, S., J. Flinn, and A. Obregon (2000), “GPS in the Pits: Differential GPS Applications at the Morenci Copper Mine,” *GPS World,* v 11, n 10, Oct 2000, pp 34-38.

Smith, R. S., (1997), “Factored estimating guidelines— Refined/ petrochemical projects,” *Cost Engineering*, 39(4), pp36-41.

Sotoodeh, H. and B. C. Paulson, (1989), “Sensors and expert systems in production optimization of earthmoving scrapers,” *in Proceedings of the Sixth Conference on Computing in Civil Engineering*, Ed. Thomas O. Branwell, Jr., ASCE, Reston, VA.

Sprinkle, H. (1972), “Analysis of Time-Lapse Construction Films,” *Journal of the Construction Division*, ASCE, 98(**CO**2), pp183-199.

Statfit, (1997), *Stat::Fit Users Guide*, Geer Mountain Software Corporation, Inc.

Stem, J. E. (1990), *State Plane Coordinate System of 1983*, NOAA Manual NOS NGS 5, US Department of Commerce, Rockville, MD.

Stewart, L. (1998), “How to Get Value From On-Board Computers,” *Construction Equipment,* Jan, pp42-50.

Stewart, L. (1998), “Monitor Machine Condition Via Satellite,” *Construction Equipment,* Mar, pp204-205.

Tavakoli, A. (1985), “Productivity Analysis of Construction Operations,” *Journal of Construction Engineering and Management*, ASCE, 111(1), pp31-39.

Thomas, Jr. H. R., M. P. Holland and C. T. Gustenhoven, (1982), “Games people play with work sampling,” *Journal of the Construction Division*, ASCE, 108(**CO**1), pp13-22.

Touran, A. (1993), “Probabilistic Cost Estimates with Subjective Correlation,” *Journal of Construction Engineering and Management*, ASCE, 119(1), pp58-71.

Trimble (2000), “CrossCheck AMPS,” Datasheet and Specifications, Trimble Navigation Limited, Sunnyvale, CA.

Tserng, H. and J. Russell (1997), “ An Instantaneous Motion Planning and Controlling Tool (IMPACT) for Construction Equipment,” in the *Proceedings of the 1997 Construction Congress*, ASCE, Reston, VA

Volvo (2002), A25C Operator's Manual, Volvo Construction Equipment Group, Asheville, NC.

VolvoSim (2001), On-line Help Manual, Volvo Construction Equipment Group, (http://www.runge.com.au/volvosim/).

Wang, W. and L. Demsetz (2000a), "Application Example for Evaluating Networks Considering Correlation," *Journal of Construction Engineering and Management*, ASCE, 126(6), pp467-474.

Wang, W. and L. Demsetz (2000b), "Model for Evaluating Networks under Correlated Uncertainty – NETCOR," *Journal of Construction Engineering and Management*, ASCE, 126(6), pp458-466.

Wolf, J., S. Hallmark, M. Oliveira, R. Guensler, and W. Sarasua (1999), "Accuracy issues with route choice data collection by using global positioning system" *Transportation Research Record*, n 1660, pp 66-74.

Zito, R. and M. A. Taylor, (1994), "The use of GPS in travel-time surveys," *Traffic Engineering & Control,* 35(12), pp685-690.

VITA

John C. Hildreth was born April 30, 1975 in Beckley, West Virgina; the son of James and Shirley Hildreth. He graduated from Spencer High School in 1993. He studied civil engineering at West Virginia University, graduating Cum Laude in the spring of 1997 with a Bachelor of Science in Civil Engineering. He continued his studies and earned his Master of Science degree in Civil Engineering in 1999. He was selected as a Charles E. Via Doctoral Fellow and began his doctoral studies at Virginia Polytechnic Institute and State University in the fall of 1999 under the direction of Dr. Michael C. Vorster. He returned to Spencer in the fall of 2002 with his wife Kari, and entered private engineering practice with Boyles and Hildreth, Consulting Engineers, where he is currently employed