ABSTRACT

Title of Document: INTEGER PROGRAM OPTIMIZATION OF COACH ASSIGNMENT WITHIN AN OVERLAPPED NETWORK

Steven A. Kolarz, MSCE 08

Directed By: Dr. Ali Haghani, Chair, Dept. of Civil Engineering

In these days of increasing traffic congestion, increasing energy prices, and decreasing transportation funding it is imperative that efficient, alternate transportation be maintained. It is therefore the goal of this thesis to propose an Integer Program model for optimizing train consists (the number of cars assigned to a particular passenger train) to lower the operational costs while still meeting demand. Further benefits are the increased utilization of the existing car fleet of the service optimized and the reduction of the overall car fleet required. All of these goals are met by the model contained here-in, and validated through an optimization of Amtrak's Northeast Operations. The model shows distinct improvements in lowering operational costs, reducing the overall fleet required, and increasing car utilization for all cases optimized. These include cases to determine sensitivity analysis, where a minimum train length is imposed and where a maximum terminal capacity is imposed.

INTEGER PROGRAM OPTIMIZATION OF COACH ASSIGNMENT WITHIN A COMPLEX NETWORK

By

Steven Adam Kolarz

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Masters of Science, Civil
Engineering
2008

Advisory Committee:
Professor Ali Haghani Chair
Professor Paul Schonfeld
Professor Lei Zhang

Acknowledgements

First, I wish to thank Dr. Ali Haghani. He has endured my foolhardy accelerated schedule and distractions to see this thesis produced. He has provided constant advice and guidance through the many intricacies of Graduate School so that I may not simply graduate, but that I may expand my knowledge.

A second thank you goes to Masoud Hamedi for his constant help and assistance through the technical aspects of integer programming and its application. Without his assistance, I am sure that this thesis would have taken atleast twice as long to produce.

Finally, I wish to thank Dr. Zhang and Dr. Schonfeld for their review and participation in my thesis examination committee.

Table of Contents

List of Figures

List of Tables

Chapter One: Introduction

With increasing environmental conscience and recent volatility in energy prices, the world has been reminded both morally and financially to pursue efficient transportation. According to an Oak Ridge National Laboratory report inter-city rail is only surpassed by motorcycles and vanpools in energy efficiency for passenger transport (Davis et. al. p.2-14). This would leave inter-city rail as the logical choice for transporting large numbers of people long distances whilst minimizing energy usage.

However, passenger rail is “subsidized throughout the world” (Karush). It is therefore in the best interest of the government agency or operating company to minimize costs in order to reduce subsidies and remain competitive within a government’s budget. Since the capital costs of a railroad (tracks, signals, etc.) require more time to change than current operations, it is of interest to study the operational costs up front. The operational costs of a railroad are largely driven by the number of trains and the consists (number of cars) of each train operated.

It is with this in mind that this thesis casts a more detailed look unto consisting and ridership, so that efficiency might be maximized. This is proposed to be accomplished by better fitting the consists of individual trains to their demand, whether it be reducing the consist on a low demand train or increasing the consist on a high demand train. By doing this there will be less unused capacity on each train operated. A train with fewer cars will then operate with lower operational costs than the existing train. By making these adjustments across an entire service area, significant cost savings can then be realized.

The actual approach utilized is a Mixed Integer Program to match the consists to the demand while still balancing the flow of cars across the system. The balancing of car flows is a necessary feature of any approach utilized to this problem since it is inappropriate to assume otherwise. This program utilizes several assumptions to accomplish its goals. These assumptions allow a lean model to be utilized, making its implementation feasible for intercity passenger rail agencies and operators without intensive resources.

The proposed model is then applied to source data from Amtrak for the month of October, 2005. Though October, 2005 is not recent data, it is still useful for comparative purposes. Specifically, it is used to compare the model's proposed car assigments to existing car assignments on an intercity passenger rail system. These are compared in terms of operating costs, fleet requirements, and car utilization. These are all common measures to determine the costs and efficiency of a passenger rail service.

Finally, a set of conclusions and further recommendations is offered. The conclusions summarize the presented model and its usefulness to an intercity passenger rail operator. The recommendations are offered in order to facilitate further research and expansion of the model presented here, as well as to guide the implementation of this model within industry.

Chapter Two: Literature Review

2.1 Overview

Though passenger rail has existed for over a century, research relating to passenger rail optimization is a much younger field. Cordeau et al. (1998) cites "early research" as occurring in 1957 for locomotive assignment, though he also explains that, "Very little work has been accomplished concerning the assignment of locomotives and cars in the context of passenger transportation" (Cordeau 1998 p.380). This may be attributable to the fact that passenger trains have existed far longer than the study of integer programming and linear programming. Whatever the case may be, there are few passenger-rail specific research papers that may be directly referenced here. Even when the topic is expanded to include transferable freight-rail research, the research field remains narrow.

2.2 Existing Passenger Rail Assignment Research

Cordeau et al. (2001) and Cordeau et al. (2000) both discuss simultaneous locomotive and car assignment heuristics. However, both of these problems are for VIA Rail (the Canadian equivalent of Amtrak) which has a smaller service density with far less route overlap. This makes the model inappropriate for use on the network presented here that does feature high service densities and significant route overlaps. Furthermore, VIA Rail had a greater heterogeneity to their locomotive and car fleet at the time of both these papers' publications. This prevented certain assumptions from being made that are made here, such as universal inter-operability.

Amtrak's modern fleet is more homogeneous, allowing the more generalized approach proposed here that does feature universal interoperability. Finally, these papers also consider locomotive assignment. Locomotive Assignment has been omitted from this paper to allow a better focus upon coach-car operations. We believe that this focus will allow an intercity railroad to more easily apply the model and produce cost savings with utilization improvements.

Ramani et al. 1992 proposes a Decision Support System (DSS) for Indian Railways that could be of use. However, their approach analyzes links instead of trains. In order to deal with links the system requires accounting for maintenance intervals and other periodic occurrences not dealt with here. Further, the focus on links for sizing consists leads to far longer runs for each train set. Since a longer run will most certainly feature greater variations in ridership than a simple train would, this approach would allow more operations with lower load factors. This translates to lower utilization, which is an integral part of what this thesis is attempting to improve.

Hong et al. presents an interesting set partition approach to solve train-set assignment in Korea. Train sets are essentially a fixed set of cars, so this is a similar problem to the one contained here. However, Hong et al. approach the problem as a weekly-repeating problem with specific equipment requirements. It therefore becomes their goal to minimize the total in-service fleet on a given day, while the model approached here is concerned with minimizing operating costs first.

Similarly, Cacchiani et al. proposes a model for Train-Units. However, Cacchiana et al. limit any train to a maximum of 2 TUs and incorporate a maintenance constraint into their model. Similar to Hong et al., this model is also built to minimize the total number of TUs required across the modeled area.

Bussieck et al. describes a similar problem to the one contained here, except on a single line (rather than a network). In Bussieck et al.'s model the demand for separate classes of service is accounted for, but the simplicity of a single, cyclic line hampers the applicability of this model.

2.3 Existing Engine Assignment Research

Kuo et al. present an interesting mixed integer linear program to model freight engine allocation. Though a freight model for locomotives, the approach and actual model used is fairly similar to the one proposed here-in for passenger coaches. However, Kuo et al. uses a fairly simplistic service area (three nodes) for testing and validation, making its implementation less rigorous than the model contained here (nine nodes).

Florian et al., as well as Ziarati et al. 1997 propose a similar model for freight locomotive assignment for application on Canadian National. However, both models utilize multiple engine classes, which must be appropriately matched to the route and tonnage of a particular train. This reflects a heterogeneous equipment fleet, which differs from the services modeled here. Furthermore, Ziarati et al. 1997 splits the model into several smaller overlapping problems to make a solution feasible, a move that is inappropriate to the problem size proposed here.

Ziarati et al. 2005 revisits the work of Ziarati 1997 (freight locomotive assignment on Canadian National) with a Genetic Algorithm approach. However, this approach assumes cyclic trains (all trains are matched by a similar train in the reverse direction). This approach would is inappropriate for the model proposed here since it would mute the effectiveness of the model's matching of consists to ridership.

Likewise, Wright and Forbes et al. each propose a model to assign locomotives to a daily-repeating schedule. However, this assumption (of a daily-repeating schedule) is not utilized here. But both of these models do assume a single locomotive class for assignment, similar to the homogeneous single car class utilized here.

Ahuja et al. define each train individually (as opposed to recurring daily or weekly with identical assignments), a useful approach which is repeated here. However, their model for freight locomotive assignment is still inappropriate when compared to this problem. This is because their model allows deadheading. Though common in freight locomotive assignment, deadheading is highly undesirable for passenger car assignment, as it represents wasted capacity. Furthermore, since the operating cost of a passenger train is based upon the consist length this approach would still incur greater operational costs for deadheaded equipment despite the potential for crew cost savings. This is because deadheaded cars would not require additional crew members, but their weight would still be a part of the train and therefore still add to the fuel costs to operate the train.

Charnes et al. propose to minimize the operational costs of a terminal railway operation. This is accomplished through a model to assign work to various crews and engines, with the timing of shipments accounted for. The model does consider numerous constraints beyond those incorporated here, such as crew ability and engine type. Further, it differs by allowing deadheading, similar to Ahuja.

2.4 Other Existing Research

Booler presents and solves a simplistic model to schedule railway locomotives (though not explicitly stated, this appears to be a passenger railway oriented model). The approach used is similar to the one presented here despite being a locomotive assignment problem. However, Booler's model does incorporate multiple locomotive classes with restrictions on the work each class can perform. This restriction is not present in the model proposed here.

Ramani 1981 proposes an alternate approach to quantify passenger coach utilization. They propose utilizing the ratio of time a car spends in service rather than utilizing a distance or a passenger load based system. Though appropriate for an extremely high-density situation (such as on Indian Railways, where Ramani is modeling), this approach is wholly inappropriate for application to an intercity system with varied ridership levels. This is because the time approach would encourage each car to be operated as much as possible, a useful approach in response to overwhelming demand. However, on an intercity line this would assign too many cars to most trains, driving up operational costs.

Chapter Three: Formulation

3.1 Background

3.1.1 Assumptions

Several assumptions were needed in order to formulate the Mixed Integer Program. A basic, initial assumption is that all trains will continue their existing schedules and motive power assignments. This allows the model to simply focus upon coach assignments without crew scheduling and locomotive assignment considerations. This assumption does force the Integer Program to assume that all trains operate, no matter the ridership. However, an alternate case (discussed below) was studied that identify low ridership trains for possible consolidation/elimination. This was done to determine if benefits could be obtained through violating this constraint.

Since the existing schedules are assumed to remain, this also allows for existing terminal operations to be assumed. With existing terminal operations assumed, the model does not need to consider turn-around times for returning cars to service or other constraints on terminal operations.

3.1.2 Parameters Used

The following is a list of parameters used in development and application of the model. It was the intent of the listed parameters that the minimum amount of detail be used to describe each train so that the model can function flexibly. In order to allow this, features such as route mileage and route time have been indexed to origin-destination stations so that the information can be reused for multiple trains that use that segment. This leaves ticket sales, fuel costs, and crew costs indexed to specific trains and dates.

d = day

t = time

S = total station list

i = origin of segment, from S

j = destination of segment, from S

k = train #

p = car capacity

F_k = Fuel costs of train *k* (\$/car-mile)

U_{ij} = Mileage of *i* to *j* (miles)

C_{kd} = crew costs of train *k* on day *d* (\$/mile)

C_{kd}' = additional crew costs of train *k* on day *d* (\$/mile)

T_{ij}^{dt} = Demand (Ticket Sales) at time *t*, day *d* for *i* to *j*

N_{ij} = Time to operate *i* to *j*

M = A very large number

3.1.3 Variables Used

The following is a list of variables used in the development and implementation of the model. The key decision variable is the consist length (Y^{kd}). It is this variable that gives the length of each train, matched to its specific ridership, and it is upon this value that the individual crew costs and fuel costs of each train are determined. I_i^{dt} is another variable that is determined by the consist length. This variable tracks inventories (the number of cars available for immediate service) at terminal stations, so that no train is assigned a consist for departure that is more than the available cars.

Y^{kd} = Consist length of train *k* on day *d*

δ^{kd} = 1 if Y^{kd} > cutoff length

0 otherwise

I_i^{dt} = Inventory of cars at *i*, day *d*, time *t*

3.2 Development

The model was approached as a traditional Minimization Integer Program problem. This entailed generating a cost function with various parameters to limit the reduction of values. This took shape in a cost function based upon operating costs. Operating costs was broken down into Crew Costs (Engineer, Conductor, and Assistant Conductors) and Fuel Costs (Diesel Fuel or Electricity). The value of the fuel varies directly with the consist length (more cars requires more fuel), while the crew costs are a step-wise function related to consist length.

The constraints began with the assumption that train capacity must meet or exceed demand for all trip segments. When applied to sample data, this would require a minority of trains to receive an increase in cars to accommodate existing ridership due to overcrowding, but the vast majority of trains' ridership will allow the consist to be reduced. A further constraint was then added that all trip segments must have the same consist (i.e. – that the train cannot change consist enroute). This restricted any switching activities to terminal stations once a train has terminated its revenue run. This constraint is in line with the assumption that existing schedules and terminal operations will be maintained.

It was at this point that a need for tracking cars at terminal stations was noticed. This tracking has been dubbed "inventories" within this thesis. It is necessary to track car inventories so that enough cars are on hand to allow the prescribed consist to operate for a train. Constraints were then added to the model that allow tracking of inventories across each day. Since the inventories vary significantly across a service day at each station, it was also at this point that time elements were added to the model. Utilizing the time element of the inventory tracking, it was then possible to add constraints to track the arrival and departure of trains from terminals stations. This was accomplished by subtracting the consist of a train from the appropriate terminal's inventory upon its departure, and likewise adding the consist to another terminal's inventory upon arrival after the appropriate time interval for a train to traverse its route. Between the arrival and departure of trains the inventory is simply carried over to the next time slot. Though the incorporation of a time element complicates the model, it is necessary in order to ensure an inventory is available to originate each train.

As discussed above, crew costs are represented as a step-wise function. New constraints were added to the model as a final step of development that allow for the crew costs to be calculated in a step-wise fashion to better reflect reality. This was accomplished by breaking crew costs into a base crew cost (C_{kd}) and an additional crew cost (C_{kd}'). The additional crew cost is only added when a consist length exceeds the cutoff value. It is once a consist length exceeds this cutoff value that operating rules and union agreements require an additional crew member to be added.

3.3 Completed Form

3.3.1 Objective Function

The model's objective function is as follows:

$$\text{(1) Min. } \Sigma_i\Sigma_j\Sigma_d\Sigma_k\, Y_{ij}^{kd}\cdot F_k\cdot U_{ij} + \Sigma_i\Sigma_j\Sigma_d\Sigma_k\, C_{kd}\cdot U_{ij} + \Sigma_i\Sigma_j\Sigma_d\Sigma_k\, C_{kd}'\cdot\delta^{kd}\cdot U_{ij}$$

The first part of Term (1) sums all car-miles accrued within the model then multiplies it by the fuel costs to give a total fuel cost. The fuel costs are uniquely set for each train to account for diesel or electric operations. The mileage is determined by the unique route of each train. The consist length (Y_{ij}^{kd}) is the decision variable of the model, largely determined by demand (ridership). The second part of Term (1) sums all train-miles accrued within the model, then multiplies it by the base crew costs. This is then added to the third part (1) which gives the additional crew costs incurred by each train with a consist longer than the cut off length. Together, both the second and third parts provide the total crew costs of the modeled services.

Together, (1) represents the direct operational costs of the trains modeled. It is considered appropriate practice within industry to solely model the operational costs of a train, leaving the accounting of infrastructural costs for elsewhere. These costs include trackage, overhead power systems, vehicle maintenance facilities, and stations. It is appropriate to leave their accounting for elsewhere since it is difficult to assign costs and "necessity" of these individual pieces to the various trains that service them. Further, the majority of the physical plant within the area modeled is also extensively used by commuter and freight rail railroads and agencies. Since these additional operations were omitted from this study, their impacts and use of facilities would be difficult to quantify here-in.

3.3.2 Constraints

The following constraints are included in the model:

(2) $p \cdot Y^{kd} \geq \Sigma_j T_{ij}^{dt}$ for all i, d, t, j <{destinations of k beyond i}

(3) $\Sigma_j Y_{ij}^{td} \leq I_{ij}^{dt}$, for all d, t, i

(4) $I_i^{d,t+1} = I_i^{d,t} - \Sigma_j Y_{ij}^{td}$

(5) $I_j^{d,t+N} = I_j^{d,t+N-1} + \Sigma_i Y_{ij}^{td}$

(6) $I_i^{d,\max t} = I_i^{d+1,\min t}$

(7) $Y^{kd} - M \cdot \delta^{kd} <= \text{cutoff}$

(8) $Y^{kd} - M \cdot \delta^{kd} >= (\text{cutoff} - M) + 1$

Constraint (2) establishes that any consist must be greater than the demand for that particular train. This constraint is written such that the consist must accommodate the peak demand segment of the train's route. Though certain trains have demand greater than their existing capacity, this constraint was set as a requirement by the case data provider (Amtrak). This means that a minority of trains will have their consist length increased in order to accommodate ridership demand that exceeds existing capacity. But the overwhelming majority of trains will have their consist length reduced because the demand is less than existing capacity, in some cases significantly so.

Constraint (3) establishes that the consist of a departing train must be less than the available inventory at that terminal. The following two constraints then allow for the tracking of departures and arrivals. Constraint (4) accomplishes this by subtracting the consist of a departing train from the inventory, then setting this as the value of the following time slot's inventory. Likewise, Constraint (5) adds a consist into the inventory of the terminal station after a train has completed its run of N time-slots length. Constraint (6) then forces the initial inventory of a day to equal the final inventory of the previous day.

Finally, Constraints (7) and (8) allows for crew costs based upon consist length. This works by using a binary variable (δ^{kd}) to track whether the consist length (Y^{kd}) is greater than, or less than/equal to the cutoff value. The δ^{kd} term is then incorporated into the objective function where it triggers the inclusion of the additional cost incurred by the additional crew member necessary for the additional car(s).

Chapter Four: Case Study Data

In the United States of America inter-city passenger rail service is provided by the National Railroad Passenger Corporation, doing business as Amtrak. Amtrak was created by Congress in 1970 with operations commencing in 1971 to relieve freight railroads of required passenger service, as it was seen as a burden to their profitability. Though established as a for-profit corporation, Amtrak has never turned a profit and has instead been reliant upon subsidies from various levels of government. It is therefore useful to minimize the operating cost and to maximize efficiencies of the services offered. The minimized operating costs shall allow Amtrak to operate with less government subsidies, while maximized efficiencies shall better make the case for those subsidies to continue. (Amtrak p.6)

Though Amtrak's routes and equipment at inception were a hodge-podge of various heritages and conditions, the system has now largely stabilized and standardized. Within the study area (detailed below), the trackage is now built to consistent standards that allow reliable, high-speed operation. The equipment is now primarily equipment that was built for Amtrak (as opposed to inherited) with few barriers to interoperability. Figure 4.1 shows a typical train within the study area. It is these facts that now allow a study of this nature to even be possible.

Figure 4.1 – Typical Amtrak train

4.1 Northeast Network

The Northeast Corridor (NEC) is the backbone of Amtrak's operations. Utilizing this corridor, Amtrak's Regional, Inland, Keystone, and Tidewater Services operate approximately 97 trains on weekdays, 57 on Saturdays, and 62 on Sundays. These services carry approximately 10 million passengers annually. The timetable is currently organized into "Weekdays," "Saturdays," and "Sundays."

The existing corridor consists of several important rail lines (distances shown in Table 4.1). The Northeast Corridor mainline (NEC) runs from Boston, MA to Washington, DC, a distance of 457 miles. The Tidewater Route extends beyond Washington, DC to Newport News, VA for 187 miles. From Philadelphia, PA the Keystone Corridor extends 104 miles to Harrisburg, PA, while the Inland Route extends 58 miles from New Haven, CT to Springfield, MA. This network can be seen in diagram form in Fig. 4.2. All trackage is electrified except for the Tidewater Route and the Inland Route.

Station		Boston, MA	Harrisburg, PA	New Haven, CT	New York, NY	Newport News, VA	Philadelphia, PA	Springfield, MA
City	*Code*	BOS	HAR	NHV	NYP	NPN	PHL	SPG
Harrisburg, PA	HAR	426						
New Haven, CT	NHV	156	270					
New York, NY	NYP	231	195	75				
Newport News, VA	NPN	644	426	488	413			
Philadelphia, PA	PHL	322	104	166	91	322		
Springfield, MA	SPG	214	328	58	133	546	224	
Washington, DC	WAS	457	239	301	226	187	135	359

Table 4.1 - Amtrak Northeast Network Distances (miles)

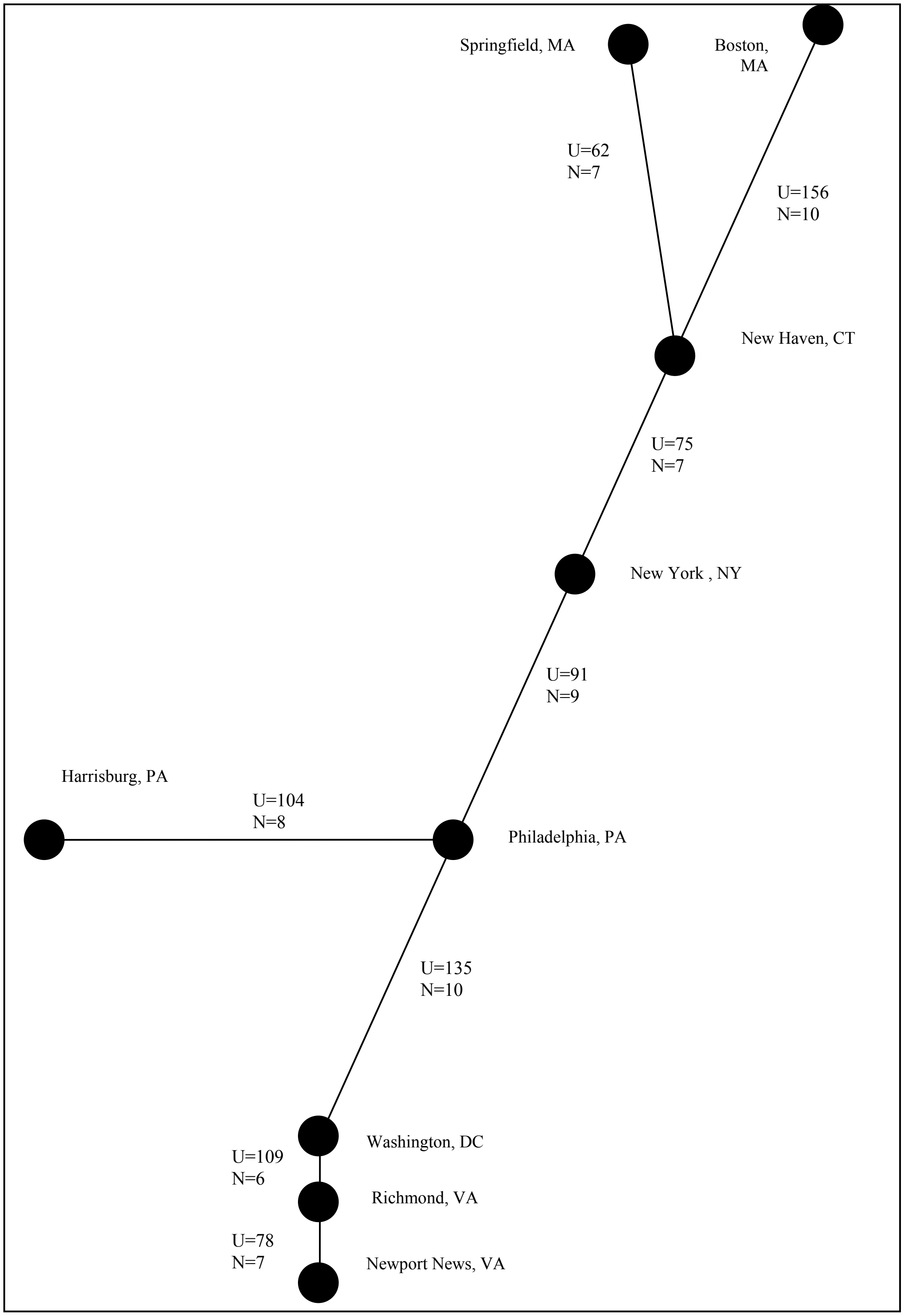

Figure 4.2 – Diagram of Amtrak's Northeast Corridor

Amtrak has assigned a three-letter station code to each station that their trains serve, to allow for speedy and accurate station identification. The following is a list of the station codes and locations for the terminal stations utilized here:

BOS – South Station, Boston, MA

HAR – Transportation Center, Harrisburg, PA

NHV – Union Station, New Haven, CT

NPN – Newport News, VA

NYP – Penn Station, New York City, NY

PHL – 30^{th} Street Station, Philadelphia, PA

RVR – Staples Mill Road Station, Richmond, VA

SPG – Union Station, Springfield, MA

WAS – Union Station, Washington, DC

4.2 Ridership/Demand Data

4.2.1 Background

Ridership data (to be specific, ticket-sales) has been used for all analysis contained in this report. This is because no true demand data is known to exist, and it is beyond the scope of this study to generate such. So it has been assumed that existing ridership is the demand. For most trains the ticket sales is indeed the ridership. However, this does present a problem in relation to sold-out trains, as ridership has been artificially limited. However, we believe that overall the impact of this limitation is minimal.

The specific data provided by Amtrak gives basic information for each existing train for the month of October, 2005. For each train it gives the ridership and capacity (seats) on each segment of the route. A separately provided timetable then gives data on each train's origination and termination times and stations, and the mileage of each route. These sources can then be combined to give an overall picture of ridership amounts, times, and locations across the system modeled.

4.2.2 Analysis

Ridership varies wildly across the study schedule. This revelation was anticipated, since Amtrak's practice of running an identical schedule each weekday would produce variations when applied to ridership demands that vary across the week. Figure 4.3 exemplifies this phenomenon through Train #198 (daily, 8:30pm New York Penn departure, 11:53pm Washington Union Station arrival). The chart displays tickets sold (Riders) and total capacity. It is easy to notice that the ridership varies from a high of 498 on Friday, October 7th to a low of 93 on Monday, October 31st. When analyzed individually, the data still shows distinct variations. If only Mondays on Train #198 were to be studied (Oct. 3, 10, 17, 24, 31), ridership still varies from 184 to 93 riders.

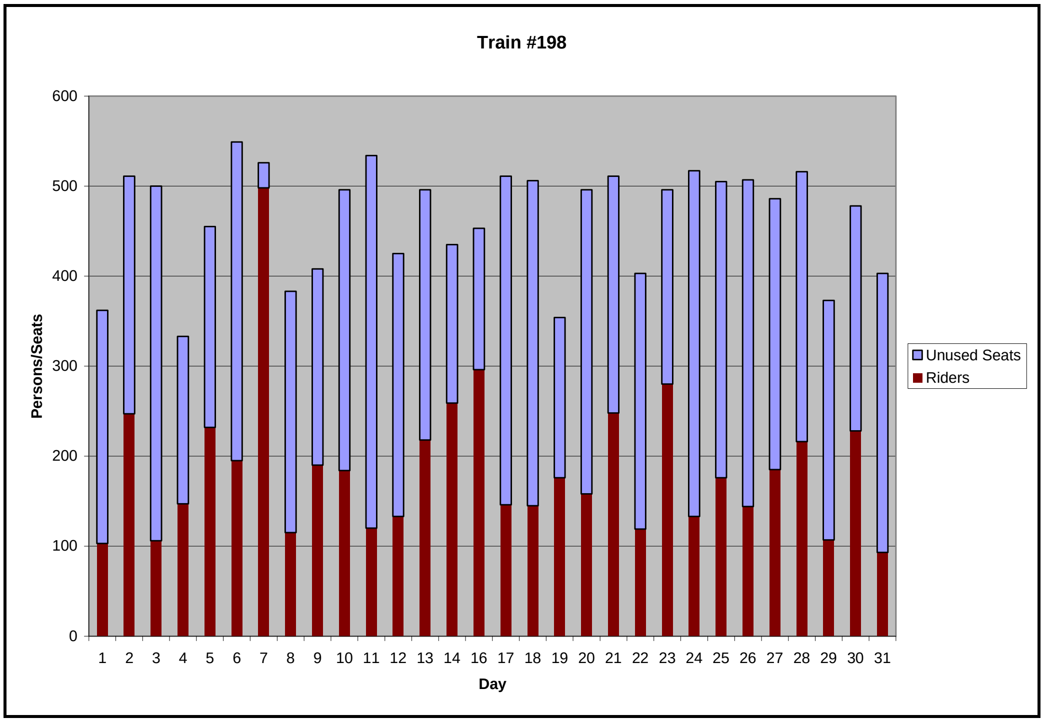

Fig 4.3 - Train #198 Ridership vs. Capacity

Though the data does appear to vary in a cyclic fashion both weekly (a relative low on Monday, increasing to a high on Friday) and daily (highs during rush-hours, lower mid-day and late-night), the data does still vary considerably. This would lead to an analysis approach that treats each instance of a train (by date and by number) individually, rather than treating them solely by train number. An analysis that treats ridership solely by train number would have to assume that ridership demands do not vary significantly enough to warrant modifying a consist on different days. This idea obviously does not reconcile easily with the existing ridership data. This approach appears appropriate, since the source data shows that the existing approach practiced by Amtrak is based on a similar assumption.

4.3 Existing Service

4.3.1 Rolling Stock Utilized

Existing service is provided by a fairly homogeneous fleet of coaches, café cars, and locomotives. For the purposes of this study, locomotive and café car assignments are assumed to remain unchanged. Since café cars are not counted in train capacities under Amtrak's existing practices, their omission will not affect the model's results. This allows the model to strictly focus upon coach assignment.

The portion of the network studied is served almost exclusively by Amfleet coaches, a fairly homogenous fleet of cars. Amfleet is a class of cars built in the 1970s and 80s to upgrade and modernize Amtrak's fleet. Though these cars were constructed in two sets and come in both coach and lounge varieties, they are wholly compatible with one-another. The Amfleet coach car capacities vary from 55 to 84 seats per car. At the directive of Amtrak, an assumed value of 72 seats per car is used for all modeling.

Further, the existing engine fleet is also fairly homogeneous within the study area. From the existing car assignment data provided, it appears that significant inter-mixing of various car-types and engine-types in use within the study area is already practiced. This means that any engine is allowed to couple to any car in use, and that the cars are capable of operation in any order.

4.3.2 Service Patterns

Existing service in the study area provides dense service coverage. Figure 4.4 shows a schematic of the services modeled here. It is important to note their overlap along the NEC mainline, particularly between NYP and PHL. This is due to that stretch having the highest ridership of the entire area modeled.

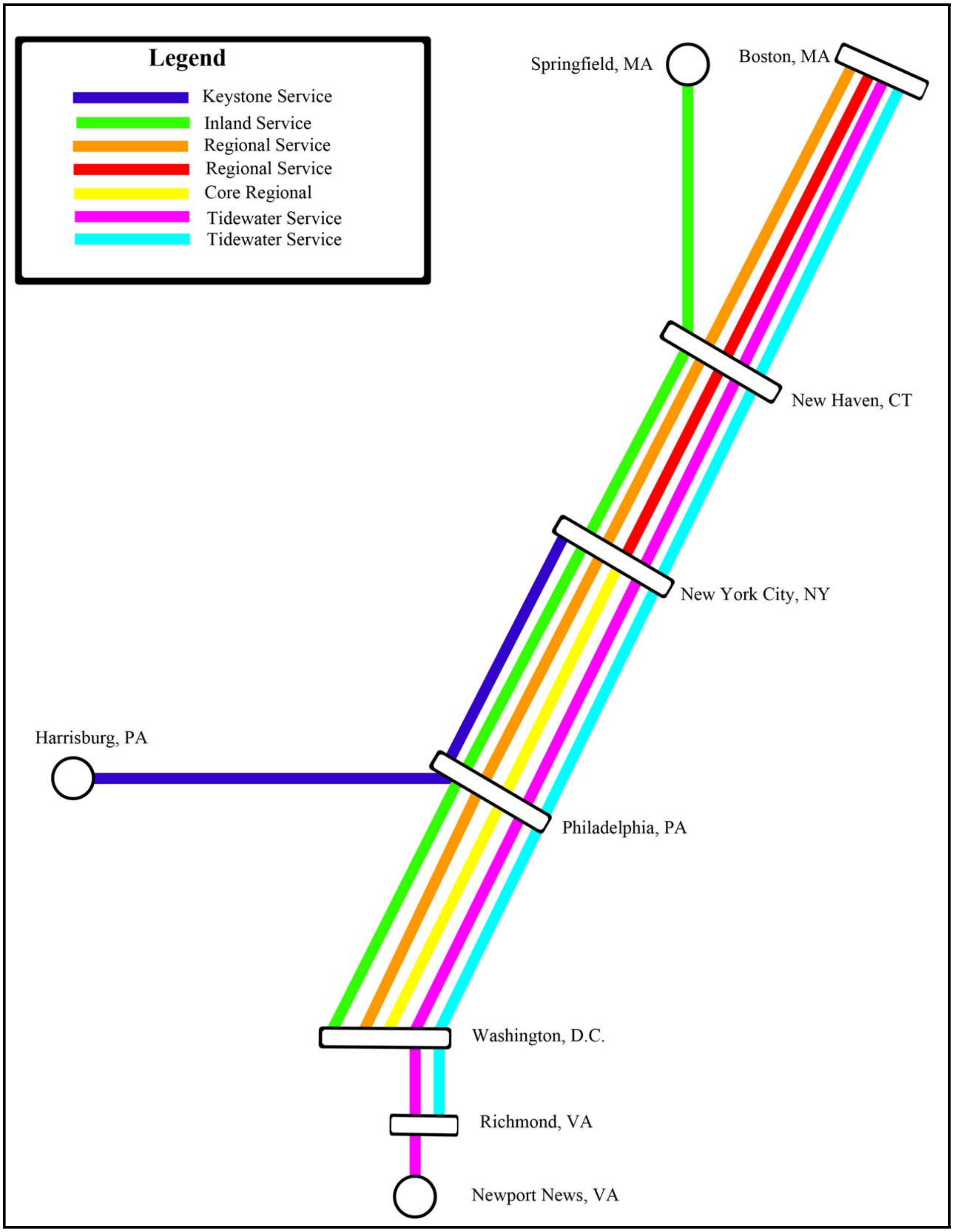

Figure 4.4 – Graphic of Amtrak's Northeast Services

The existing Regional and Core Regional service operates approximately hourly along the Northeast Corridor mainline. Figure 4.5 shows the existing consist length distribution for these services. As can be seen, 7-car consists largely dominate this service, but 5 and 6-car consists are seen. A maximum length of 11-cars occurs eleven times in the month-long period.

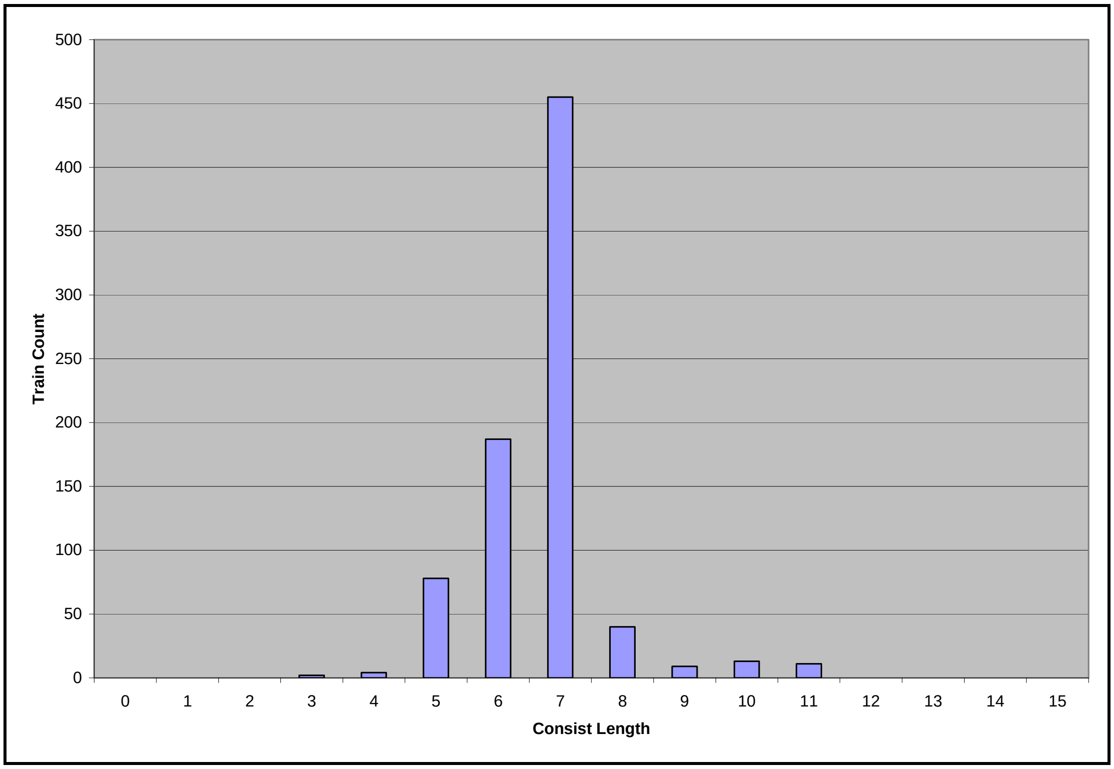

Fig 4.5 – Existing Regional Service Consist Distribution

Tidewater Service operates approximately hourly during rush-hour. Figure 4.6 shows the existing consist length distribution for these services. As can be seen, 7-car consists largely dominate this service, but 5-car consists are also prevalent. A maximum length of 9-cars occurs once in the month-long period.

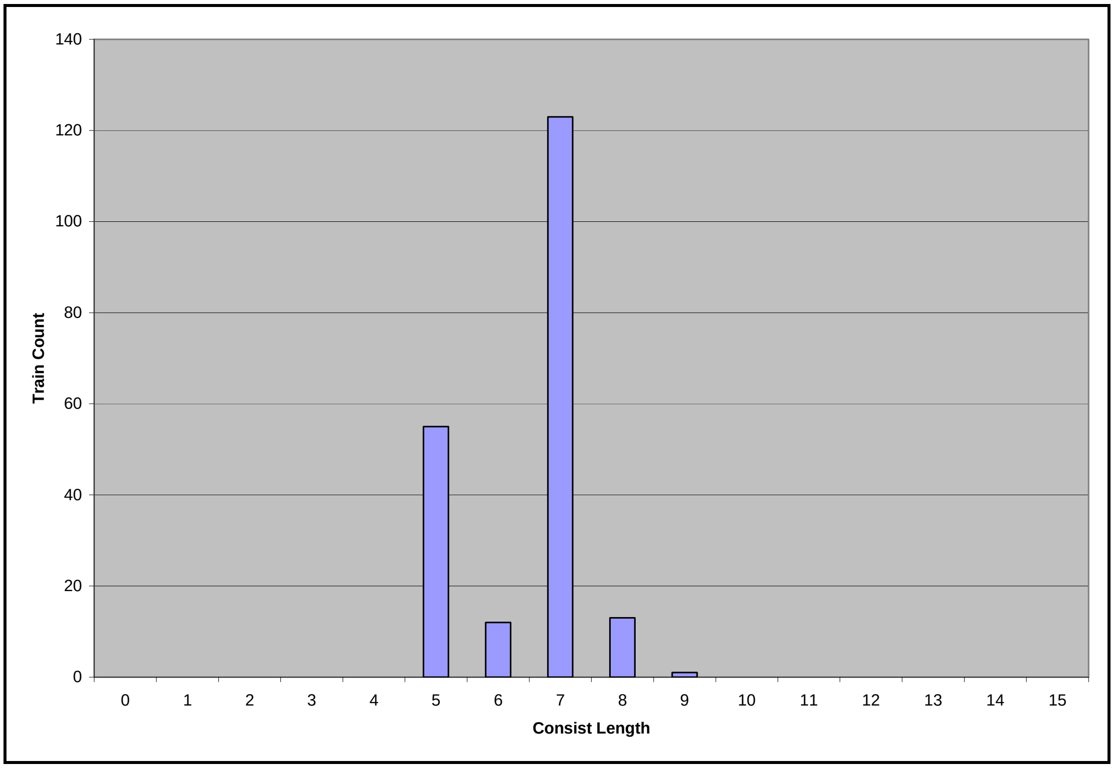

Fig 4.6 – Existing Tidewater Service Consist Distribution

Much as the Tidewater Service, the Inland Route Service also operates hourly during rush-hour. Figure 4.7 shows the existing consist length distribution for these services. As can be seen, 7-car consists largely dominate this service, with other consist lengths occurring rarely. A maximum length of 10-cars occurs once in the month-long period.

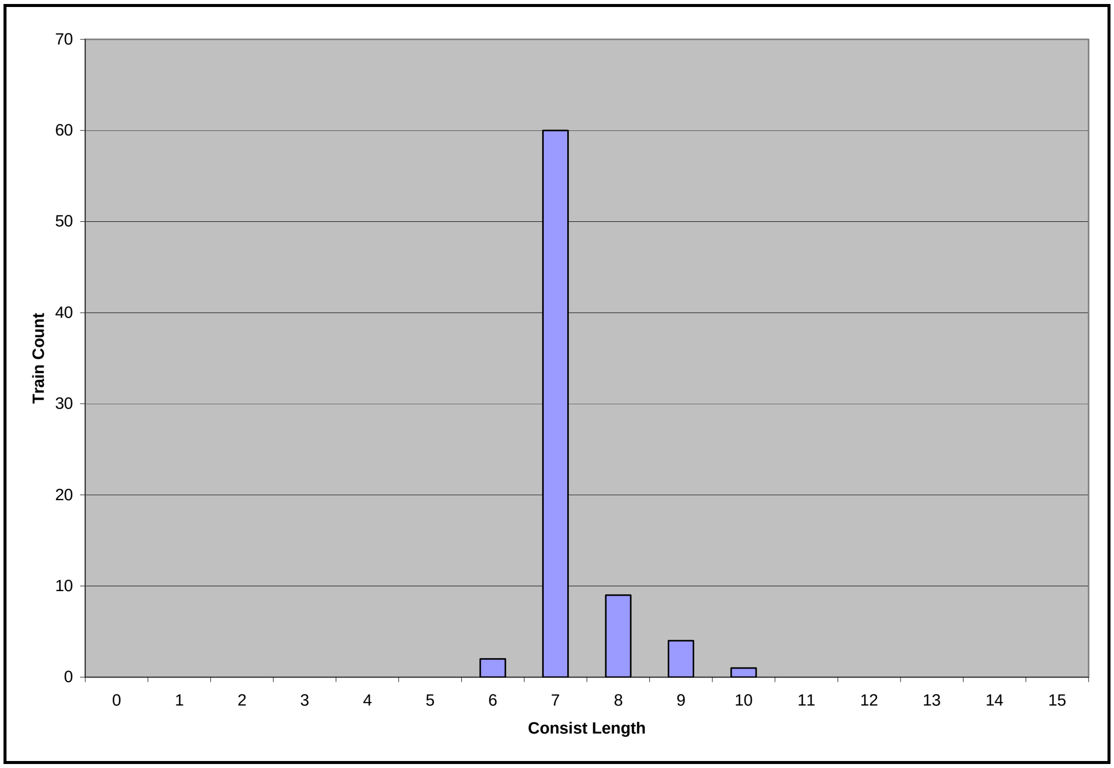

Fig 4.7 - Existing Inland Route Service Consist Distribution

Keystone Service operates approximately hourly with half-hourly rush-hour service. Figure 4.8 shows the existing consist length distribution for these services. As can be seen, 3- and 4-car consists largely dominate this service, with 2-car consists also prevalent. A maximum length of 9-cars occurs five times in the month-long period.

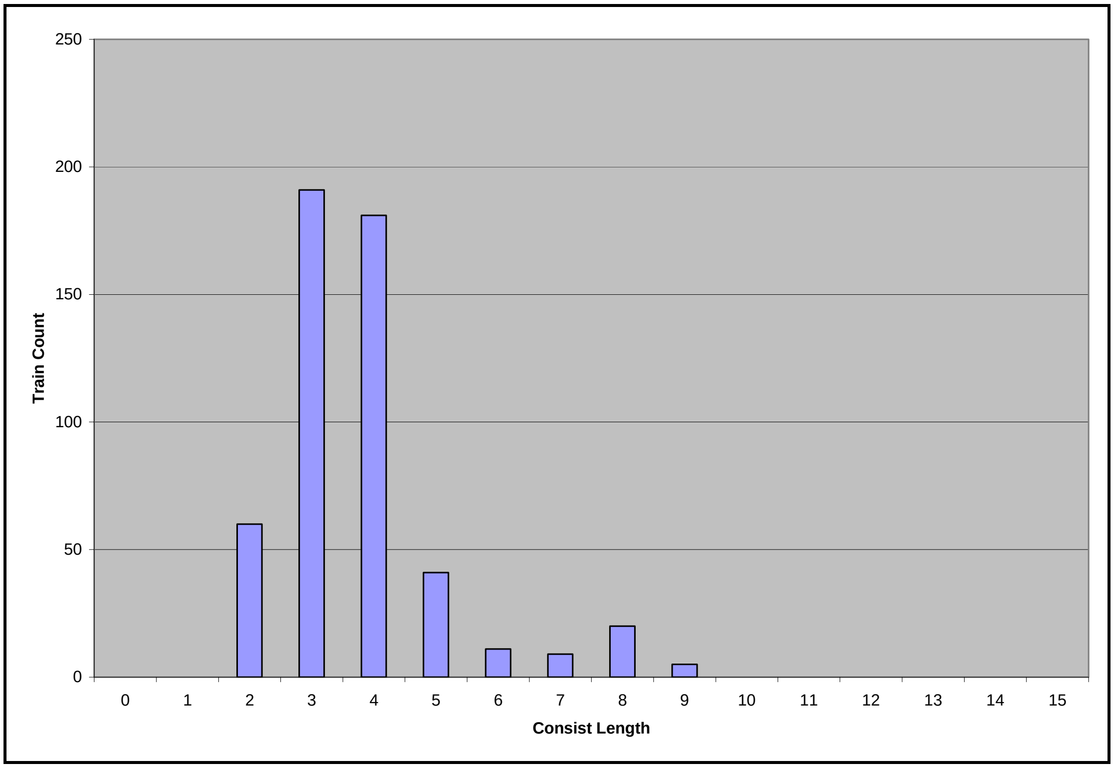

Fig 4.8 - Existing Keystone Service Consist Distribution

These multiple services mesh to provide frequencies up to a train every 15-minutes along certain stretches of the corridor, notably between Philadelphia and New York City. Figure 4.9 shows the existing consist length distribution for the entire study area. As can be seen, 7-car consists largely dominate all services. A maximum length of 11-cars occurs eleven times in the month-long period, all on the Regional Service.

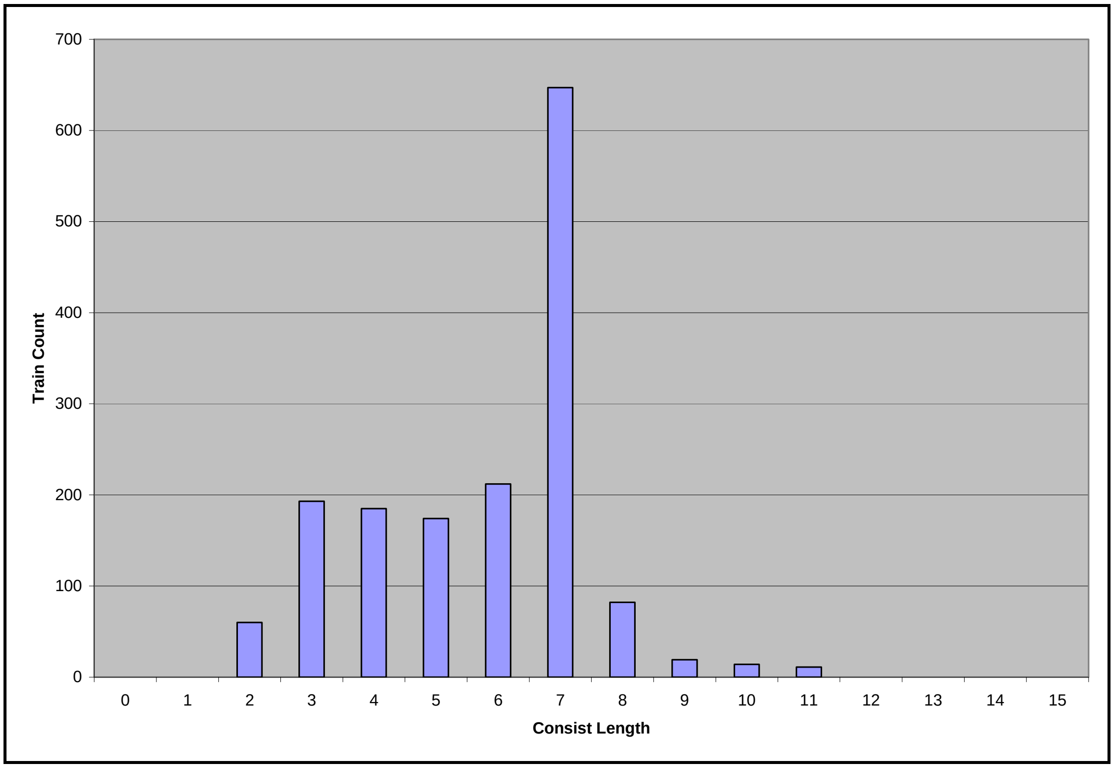

Fig 4.9 - Existing Overall Consist Distribution

4.3.3 Load Factors

Though the data is provided and detailed elsewhere, it is important to discuss LDF or Load Factors here. This is a measure of utilization commonly used in the transit industry. There are two approaches to calculating LDF (Vuchic 13), both of which are used throughout this thesis. Peak LDF is an approach where the maximum ridership along a route is used, to give the peak utilization of the available capacity (in this case, seats). Though useful to determine the maximum loading, this approach does not explain utilization along an entire route. That is where LDF-Miles comes in handy. In this approach the seat-miles of each train is calculated to give available capacity, then compared to rider-miles of the riders. The results give a better picture of utilization across the entire route, but omit any consideration of peaks in the loading.

Hence, both approaches must be presented to give a complete analysis. Figure 4.10 illustrates these differences for Train #95 on October 4th. Note that the existing Peak LDF is 42.94%, but the existing LDF-Miles is 28.49%. This discrepancy is easily visible given the variation in the ridership (blue line) across the route, versus the constant capacity (green line proposed, red line existing).

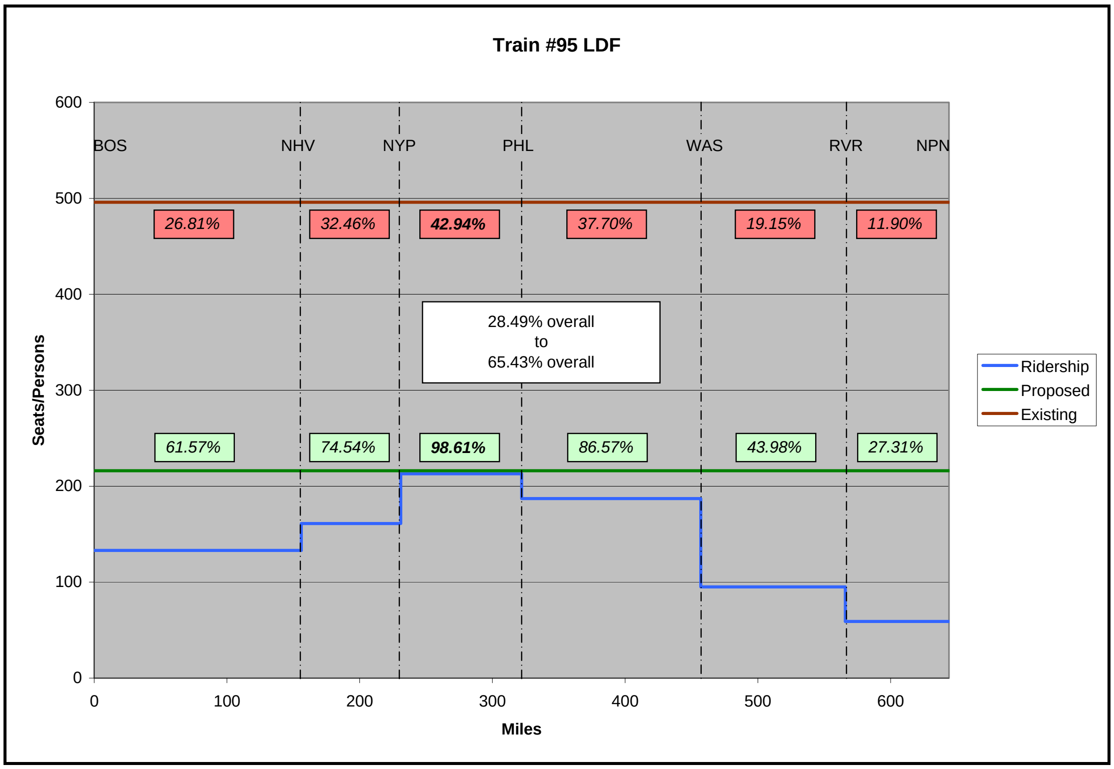

Fig 4.10 – Train #95 LDF Comparison

4.4 Existing Crewing Patterns

Existing trains are crewed based upon consist lengths. A train of 7 revenue cars or less has a crew consisting of a conductor and an assistant conductor in the passenger cars and an engineer in the locomotive. Beyond 7 revenue cars another assistant conductor is added to the crew. An extra crew member is required for the longer consists by union operating agreements to assist in handling the extra cars and ridership. Assuming that all trains have a diner or cafe car assigned (as discussed above), this leaves the cutoff at 6 coaches for use within this thesis. Therefore, all trains up to and including 6 coaches have a smaller crew than a train with more than 6 coaches in the coding implementation. This important point is used to implement the stepwise crew-costing element of the Integer Program proposed.

4.5 Assumptions Validity

As assumed in the formulation, all existing schedules are to remain. This allows the model to work strictly on coach assignments with existing locomotive and crew assignments to remain.

As assumed in the formulation, all existing terminal operations are to remain unchanged. Since all terminals studied (see p. 19) allow for some form of car storage, this is deemed appropriate. The major terminals (BOS, NYP, PHL, and WAS) also have switchers to provide consist make-up and break-down services, but these services could be provided by the mainline engine and train crew at other terminals as well. Though terminal capacity is assumed unconstrained for the Full Model Run, an alternate case was performed to determine the sensitivity of the results to terminal capacity limits.

4.6 Summary

Amtrak's Northeast Operations represent the epitome of rail passenger transport in the United States, but there is still room for improvement. The presented data represents the existing operations of this service area for October 2005 – a dense network of trains with varied consist lengths. The existing network and coach assignments reflect a heritage approach to coach assignment that does leave excess capacity. It utilizes an assignment method that is based upon historic practices rather than a programmed approach.

However, the ridership in this area presents a varied picture. By nature the ridership varies with the time of day and day of the week, but it also varies across the month. Though the existing coach assignments do somewhat mimic the rise and fall of ridership, they do not closely match the actual demand. It is because of this that the existing trains have excess capacity and room for improvement.

The existing crewing patterns are based upon a stepwise function determined by the consist length. This leads to crew levels linked, but not linearly determined by the consist length. Though the crews are determined by revenue cars, a basic assumption can adapt the model to correctly cost the crew levels.

Chapter Five: Case Study Implementation and Results

In order to test the efficacy of the proposed model, it was applied to the real world Case Study data as presented in Chapter 4. Though not projections of future ridership, this still allows the model to be compared against existing assignment practices to determine the magnitude of potential savings possible. The actual mechanics of this implementation and the results it produces are detailed here.

The application is specifically applied to several unique cases. The first is a Base Case, simply intended to determine the existing car flows and inventories to establish a baseline for comparison. The Full Model is then optimized unconstrained to determine the largest possible improvements accomplished by this model. To determine sensitivity the model is then rerun with constraints added. In the first case a Minimum consist length of 3 is applied (MIN 3), with ridership reassigned from dropped trains to determine savings possible with consolidation. The second case returns the consists to unconstrained and applies a maximum inventory to each station (Term Cap) to determine the effects of a real-world constraint. Finally, other ideas that were deemed infeasible (and therefore un-implemented) are also discussed.

For each of the cases optimized results are presented in both discussions and graphs. Since all cases discussed provide improvements over the Base Case, comparisons between the various cases are also offered. This allows for determination of the model's sensitivity to other factors and full consideration of the assumptions utilized prior to implementation.

5.1 Parameter Values

Since the model is being applied to Amtrak's Northeast Operations, the values of the parameters were directly dictated by Amtrak. It should be noted that the last train arrival within the modeled area occurs at 2:00am (there are no overnight trains modeled). To allow the model to properly account for arrivals at precisely 2:00am, the model's time frame is then extended to 2:15am. The time slots from 5am to 2:15am are treated as a single service day, despite straddling midnight. Below is a list of values for each parameter as used in all of the cases discussed:

d = 1 to 31 (Day of the month of October)
t = 0 to 93 (5am to 2:15am in 15-minute increments)
S = 9 total stations (See p.18 for Stations)
p = 72 seats
F_k = \$0.9141/car-mile electric traction, \$1.1246/car-mile diesel traction
U_{ij} = See chart on p.2
C_{kd} = \$1.57/train-mile
C_{kd}' = \$0.50/train-mile

5.2 Initial Validation

As an initial validation of concept, the model was initially tested in LINDO. Though the actual model formulation was in flux, this validation was undertaken in April 2008. Because of its high frequency and variation of demand, the Keystone Service between Harrisburg, PA and Philadelphia, PA was chosen for validation. This choice necessitated the truncation of any train operations beyond Philadelphia, a choice that likely stilted the results. However, it was considered an appropriate validation at the time that the model did indeed show improvements both in cost function and utilization (through Peak LDF). LINDO required 1132 iterations with 53 branches to produce a solution. This solution lowered the objective function from \$74,214 to \$41,792 and improved the Peak LDF from 30.7% to 54.8%. With this "Proof of Concept," development continued and programmed proceeded with CPLEX.

5.3 Coding Implementation

The model was programmed in a multi-step process. First, the source data (provided in Excel and Access format) was converted into a text file. Within this file each line of input is considered a train, with various parameters in a predetermined order. This data was then read-in by a C++ program created to synthesize and properly interpret the source data into an actual IP file for CPLEX input. The C++ program then wrote this model to a second output text file. This file was then read into CPLEX and optimized. From CPLEX, a log file was produced detailing the optimization process and listing all non-zero variables. All post-processing was accomplished by importing this log file into Excel.

As can be seen in Table 5.1, the Full Model Optimization input file encompassed 4,310 lines of data. This file length was repeated for all input files except the MIN 3 cases since the number of trains modeled only changed for those. This was processed by a C++ file of approximately 270 lines length (321 lines in the case of the Term Cap case) to produce a C++ Code. The length of the CPLEX Code varied from 39,021 for the MIN 3 case to 60,725 for the Base Case and Full Model Optimization to 86,953 for the Term Cap case. As can be seen in Table 5.2, the C++ programs had a typical runtime of 7 minutes. The MIN 3 case had a shorter run time due to the consolidation eliminating several trains from processing. CPLEX had a typical runtime of a half second for all cases except the Term Cap Case. Even then, all run times remained under 1 second.

	Input File	C++ Code	CPLEX Code
Base Case	4310	269	60725
Full Model Opt	4310	269	60725
MIN 3 Opt	1597	274	39021
Term Cap	4310	321	86953

Table 5.1 – Lines of Code

	C++ Runtime	CPLEX Runtime
Base Case	406 s	0.47 s
Full Model Opt	440 s	0.58 s
MIN 3 Opt	155 s	0.38 s
Term Cap	425 s	0.84 s

Table 5.2 – Run Times

5.4 Base Case Results

5.4.1 Purpose

The purpose of the Base Case is to establish the existing service attributes for comparison of all proposed cases. Though the existing consists were provided by Amtrak, this analysis was conducted to determine inventories at terminal stations and operational costs (the objective function value).

5.4.2 Implementation

The Base Case was implemented through the standard model. However, the restraint to ensure capacity is greater than ridership was changed to force the consist to be greater than or equal to that given by Amtrak. The use of "greater than or equal to" allowed deadheaded cars to appear in order to balance inventories. The capacities given by Amtrak were still used to calculate load factors, but the inventories given by this model were used to calculate fleet requirements.

5.4.3 Results

The model returned an objective function value of $3,173,269.06 to operate the trains for a month. This cost is a base value to compare all further models for cost improvement. Based upon the inventories returned by the model, the existing service requires 799 coaches to operate. The inventories required across the entire month can be seen in Figure 5.1. This number will also be used for all future comparisons to determine fleet reductions achieved.

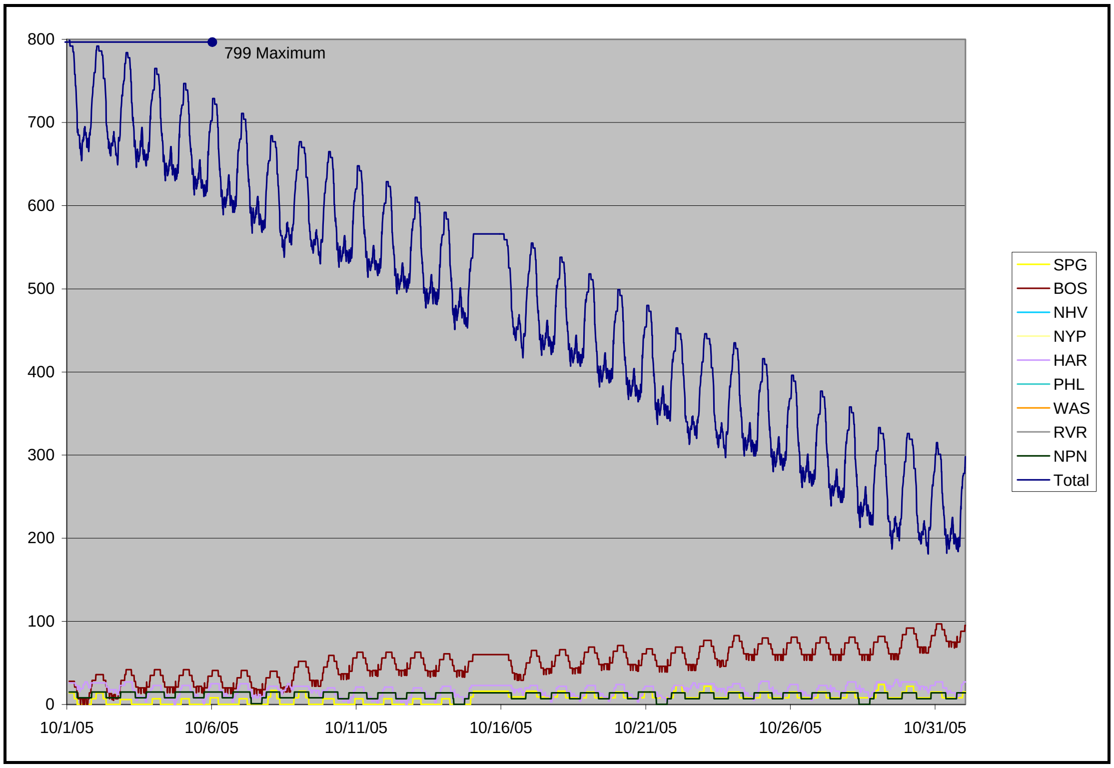

Fig 5.1 – Base Case Inventory Chart

5.5 Full Model Optimization

5.5.1 Purpose

The purpose of the Full Model Optimization was to evenly apply the proposed model to all of the existing trains within the study limits. This run then produces a full set of optimized trains throughout the study limits for further analysis. Existing schedules were assumed to remain.

5.5.2 Implementation

The Full Model Optimization was implemented through the standard model with no changes. The constraints on consist length were not modified to allow the model to freely propose whatever consist length was appropriate to the given input data. Constraints on Inventories at terminal stations were also left unconstrained to allow the model to freely modify inventories and consists to match the ridership.

5.5.3 Results

The Full Model Optimization returned a more varied distribution of consist lengths than the Base Case. This can be observed in Figure 5.2. Whereas the Base Case was dominated by consists of 7 coaches, the model has proposed distribution featuring far more consist of 2 through 5 coaches. It should also be noted that based upon source data of zero ridership for three trains, the model has proposed these trains receive consist lengths of 0. This is assumed to mean that the train does not operate, and that the train is dropped from further analysis, such as load factors, for this case.

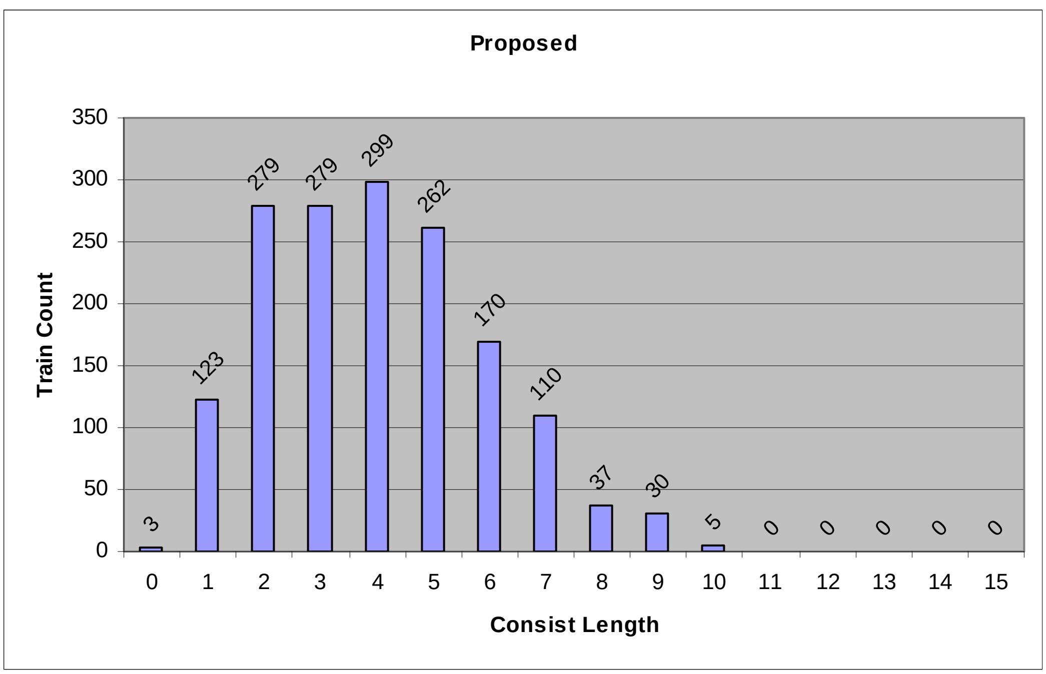

Fig 5.2 – Full Model Optimization Consist Distribution

For a true understanding of what the model has proposed, it is also useful to study Proposed Deltas. Proposed Deltas are the difference between the existing consist length and the proposed consist length. Therefore, a plot of the distribution of deltas for all trains can show how many trains are having a certain number of cars removed or added from their existing consists. This is shown in Figure 5.3. The distribution produces an apparent bell curve about -2. This means that the most common occurrence is for an existing train to loose 2 coaches from its consist. It is also interesting that +5 and +6 both have values of 1, meaning that there are 2 existing trains that require a significant increase in consist length.

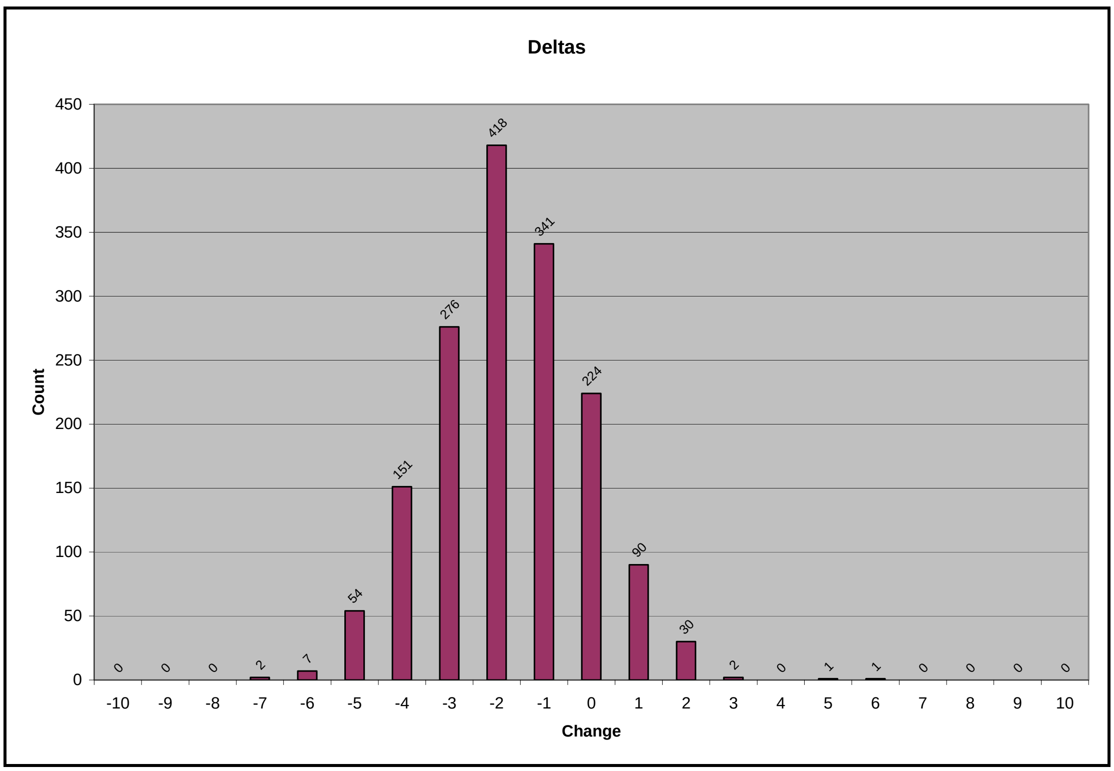

Fig 5.3 –Full Model Optimization Deltas Distribution

The Full Model Optimization returned an objective function value of $2,484,920.98 to operate all trains within the study area for a month. This represents a $688,348.07 or 21.7% reduction in cost from the Base Case. If projected to an annual savings, this totals approximately $8.2 million.

Further, after analysis the total fleet required to operate all trains is 564 coaches. This represents a 29.4% (235 coaches) reduction compared to existing requirements. The inventory chart across the entire month modeled can be observed in Figure 5.4.

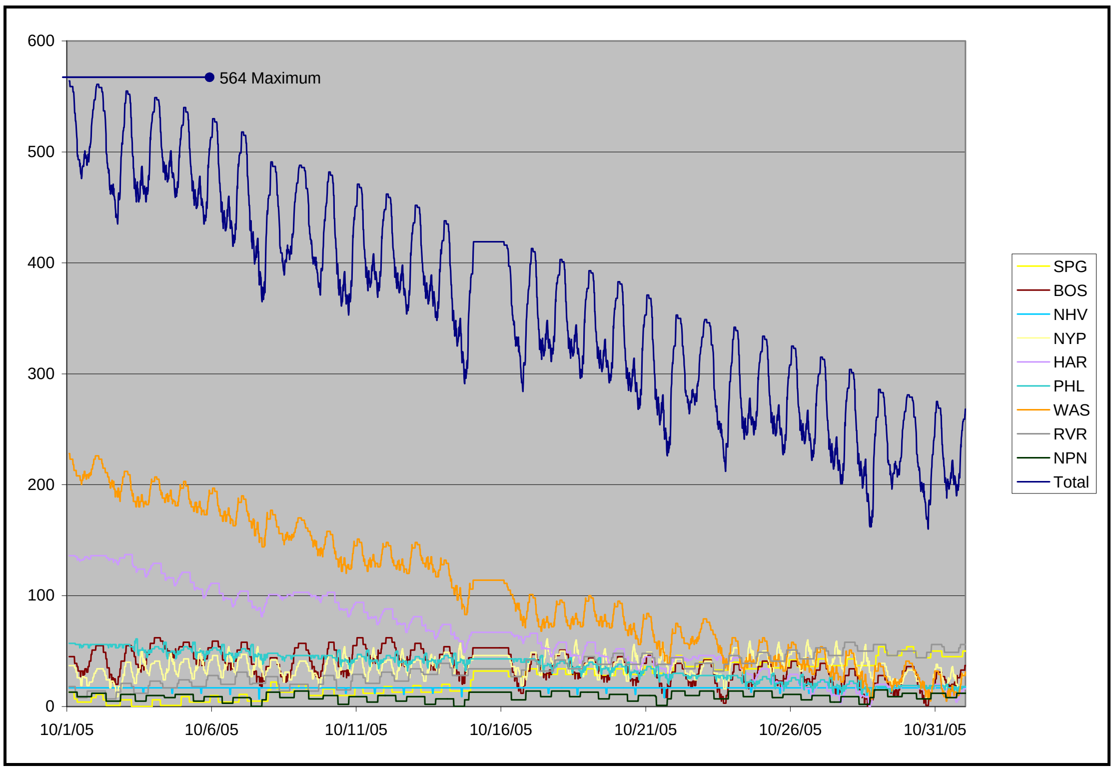

Fig 5.4 –Full Model Optimization Inventory Chart

Load Factor, or LDF was also improved by the Full Model Optimization. The Peak Load Factor improved from an overall average value of 60.48% to 84.35%. As can be seen in Figure 5.5, the distribution also improved from a series of low peaks between 40% and 70% to two higher peaks at 90% and 100%. This represents a marked improvement in utilization of the peak capacity.

Likewise, LDF-Miles showed marked improvement from 48.53% to 66.90%. A distribution of LDF-Miles can be seen in Figure 5.6. This figure shows an improvement from an even distribution around a peak near 40% to several peaks near 70%.

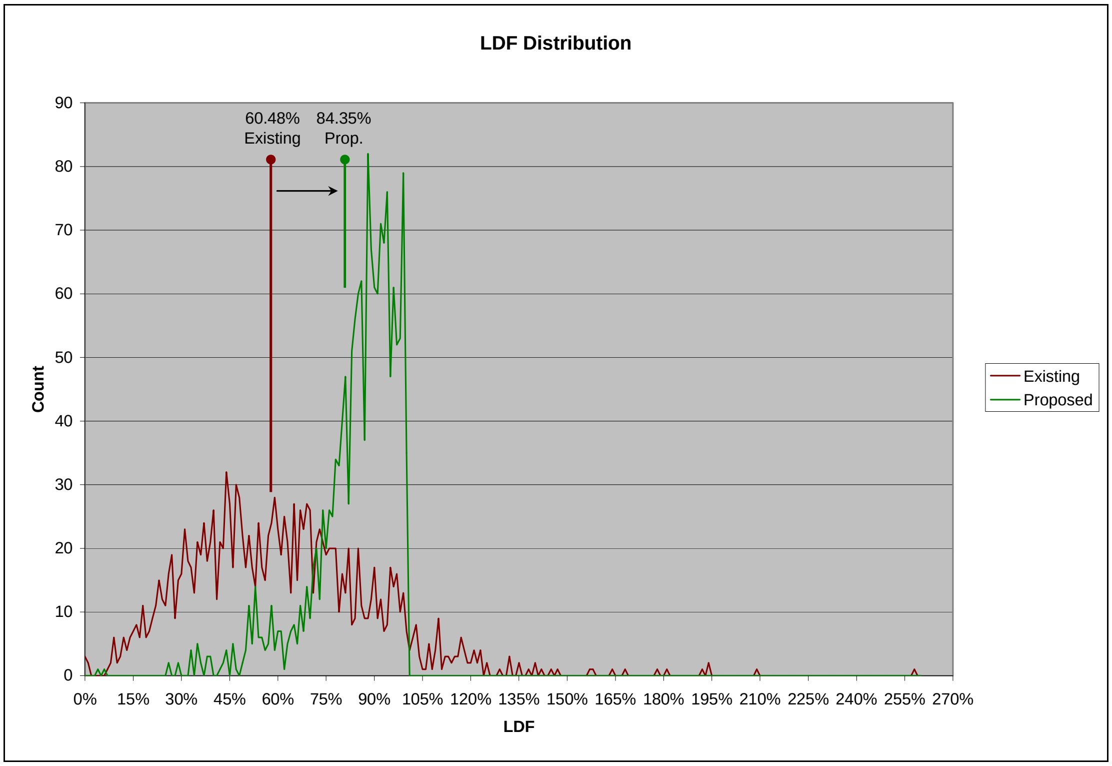

Fig 5.5 –Full Model Optimization Peak LDF Distribution

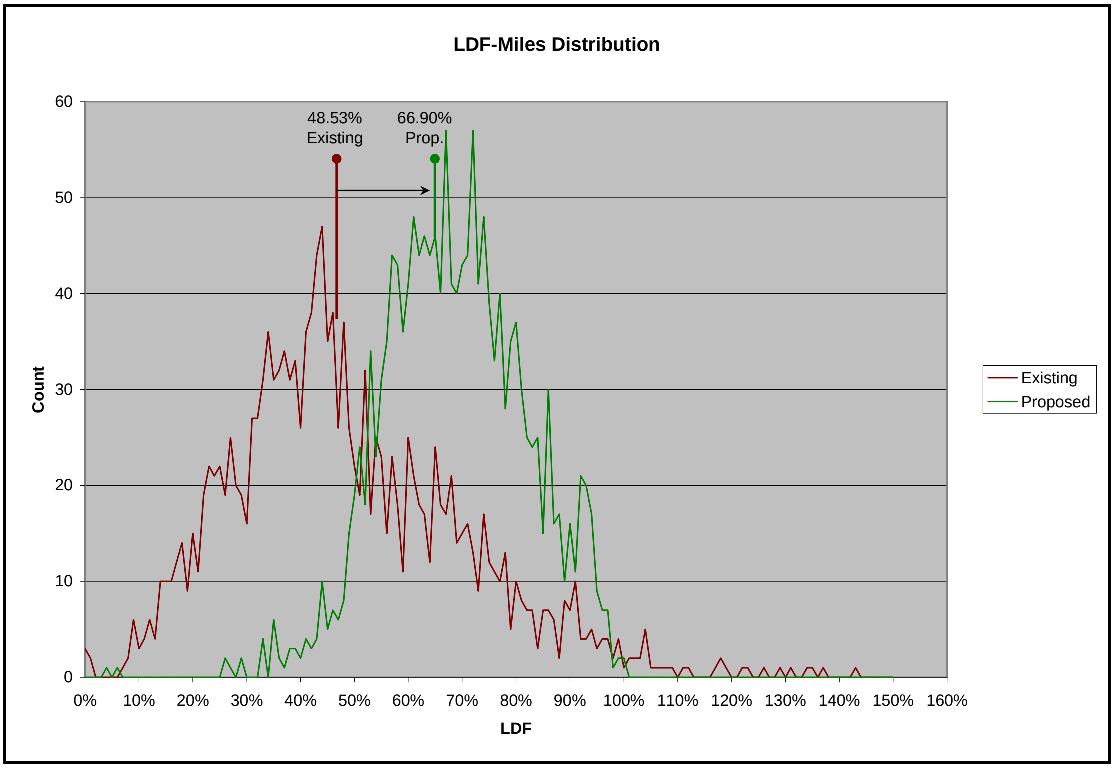

Fig 5.6 – Full Model Optimization LDF-Miles Distribution

5.6 MIN 3 Case Optimization

5.6.1 Purpose

The initial Full Model Optimization proposes 405 trains with a consist of 2 or less cars (including 3 with a consist of 0). Since a train of less than 3 cars is considered uneconomical, this model was run to determine benefits of a "Minimum 3 Car" case. This was applied by assuming that any train with 144 (two cars) ridership or less at any point would be dropped, while any train with greater than 144 ridership at any point would operate. An algorithm was applied to the ridership demand to transfer ridership from the dropped trains to the next available train on that route at 75% and 100% retention rates. If no later train was available, then the previous train was utilized. If no other train existed on the route for that date, then the ridership was considered lost (this did not occur in the application though). The approach of the algorithm is:

```
k = train #
k_prev / k_later = previous / later train to transfer ridership to
i = origin station of train k
j = destination station of train k
T = ticket sales of train k
P = percentage transferred (75% or 100%)

If Train k has low ridership (T <= 144)
        {
        Find next train (t > t_k) on route or longer (k_later includes i,j of k)
                {
        Transfer p*T to k_later
                }
        Else, find previous (t < t_k) on route or longer (k_prev includes i,j of k)
                {
                        Transfer p*T to k_prev
                }
        Else, if no other train that includes i, j on d
                {
                        Consider Ridership lost
                }
```

5.6.2 Implementation

Once the ridership data had been modified by the algorithm, the MIN 3 Case was implemented through the standard model. To allow counting of dropped trains, they were left in the model with a ridership of zero. The model did not propose that any of the trains with ridership zero receive any cars.

5.6.3 Results

The model returned consist results notably different from the Full Model Optimization. Due to the consolidation algorithm, the 75% retention rate resulted in 305 trains assigned a consist length of zero, meaning they were dropped from operation. Since the Base Case had 3 trains with consist zero (meaning zero ridership), it can be derived that 302 trains were consolidated out of the schedule due to low ridership. The consist counts for 75% retention are shown in Fig. 5.7. It can also be observed that all consist counts except 1 and 2 coaches have higher values than the Full Model Optimization to account for the shifted ridership.

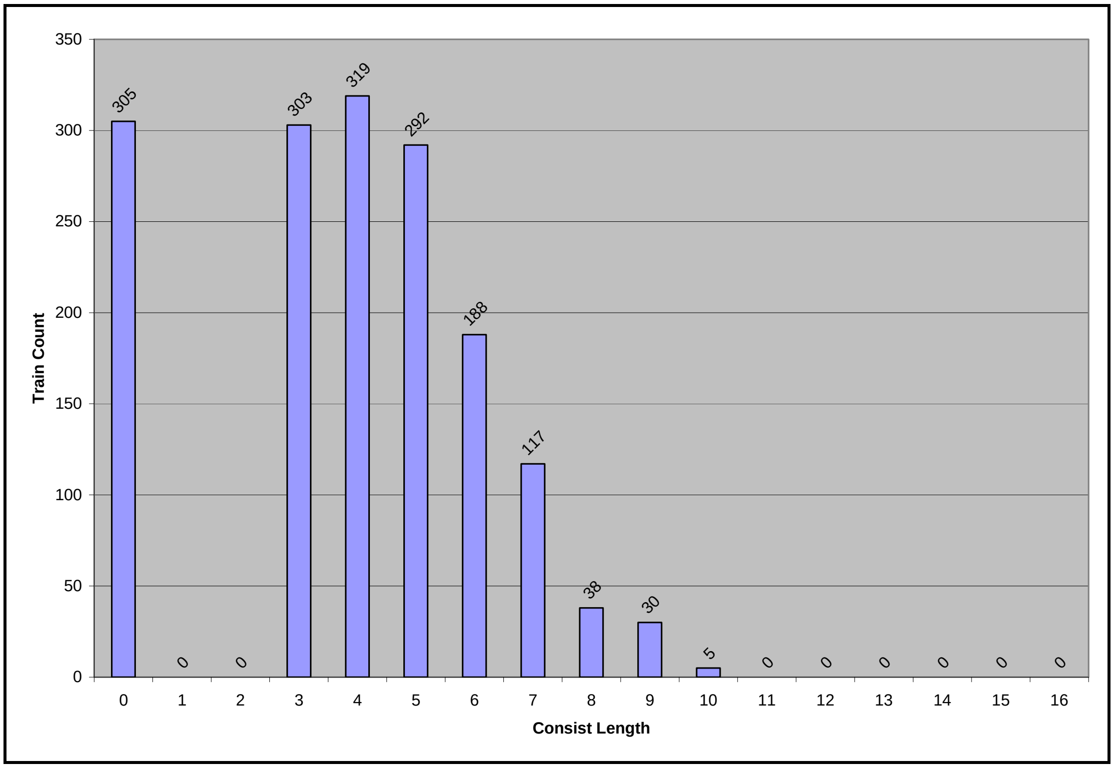

Fig 5.7 – MIN 3 – 75% Consist Distribution

The 100% retention rate produced similar results to the 75% retention rate. The largest change realized is the decrease of 3-car consists by 13. The rest of the changes in consist counts are less than 10 trains. These results can be seen in Figure 5.8.

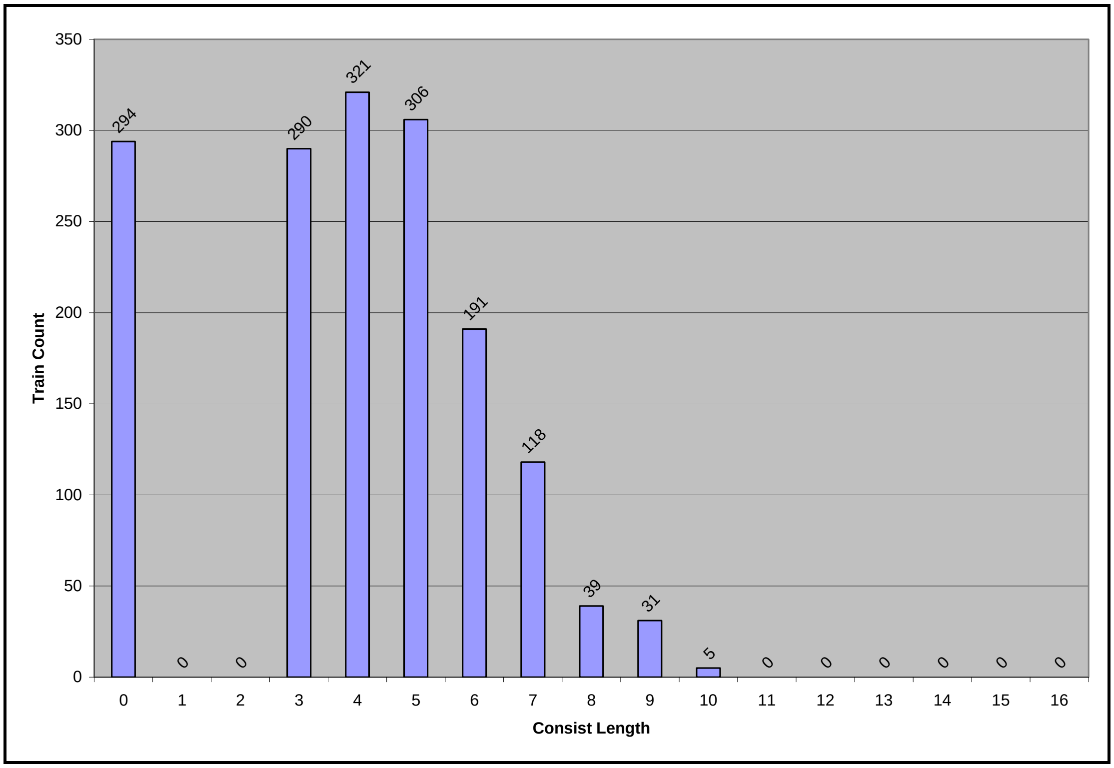

Fig 5.8 – MIN 3 – 100% Consist Distribution

It is also useful to study the Deltas Distribution. The Deltas for 75% retention can be observed in Figure 5.9. Though still an approximate bell curve, there is now an increase in the larger reduction values (notably -6 and -7), and a decrease in the lower reduction values (notable -2 and -1). Figure 5.10 shows the Delta Distribution for a 100% retention rate. Following the similarities of the Consist Counts, this figure is also similar to the 75% retention rate's Delta Distribution. There are only minor decreases in the negative deltas and minor increases in the positive deltas present. The shifts in either retention rate compared to the Base Case are logical since dropping shorter consists would favor increases to the larger reduction values from lower ridership trains, while shorter deltas would lower due to more cars needed on the remaining trains.

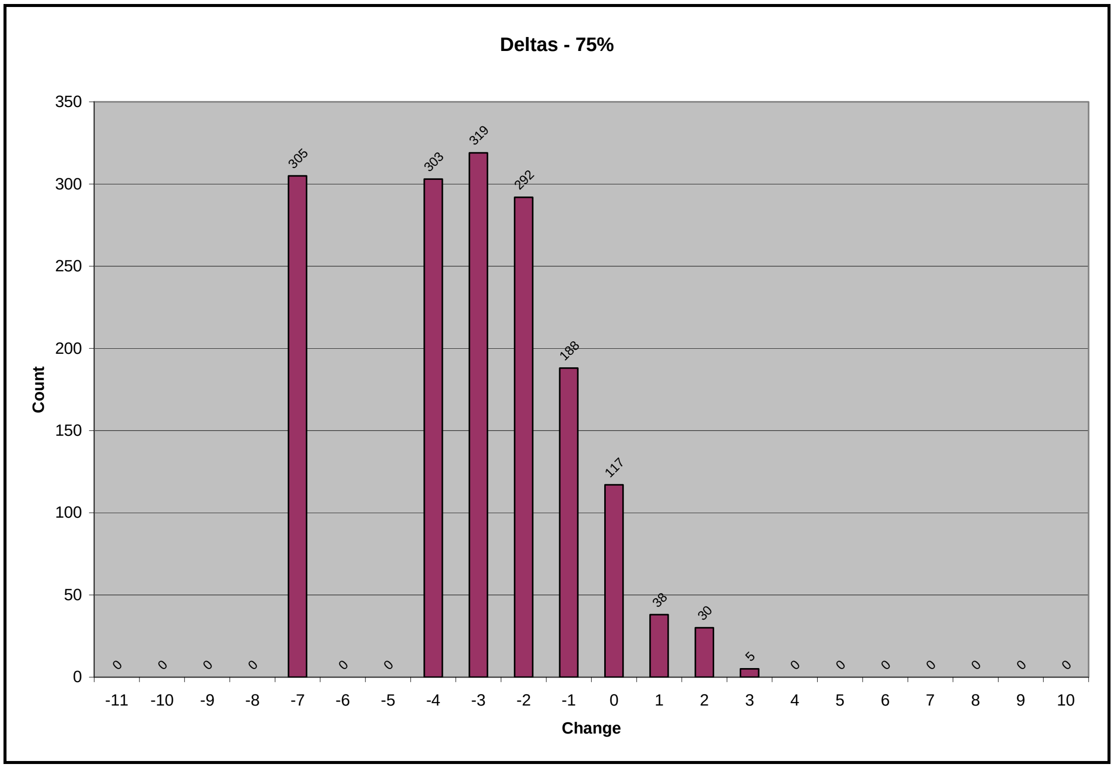

Fig 5.9 – MIN 3 – 75% Deltas Distribution

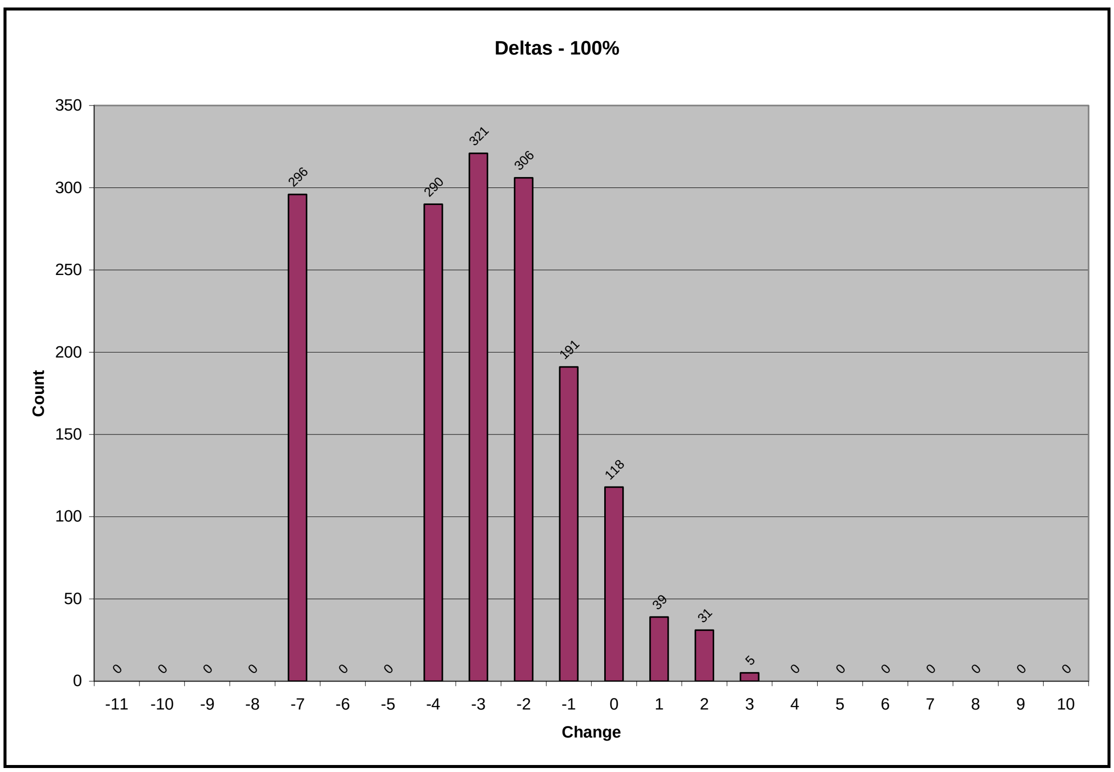

Fig 5.10 – MIN 3 – 100% Deltas Distribution

The model returned an objective function of $2,687,483.99 for 75% retention and $2,705,761.82 for 100% retention to operate the remaining trains within the month long period. This represents a $485,785.07 or 15.31% and $467,507.23 or 14.73% (respectively) savings over the existing the Base Case operational costs. If projected out to an annual cost savings, the savings could reach approximately $5.5 to $6 Million for either retention rate.

After analysis, the overall inventory requirement for the 75% retention case was determined to be 407 coaches. This represents a 49.1% (392 coaches) reduction over the Base Case. This chart can be seen in Figure 5.11.

The overall inventory requirement for the 100% retention case was found to be similar at 395 coaches. This represents a 50.6% (404 coaches) reduction over the Base Case. It is interesting to note that this case produced the largest reduction in inventory of any of the cases studied. This is likely due to slightly more trains operating than the other retention rates leaving more opportunities to balance cars between terminals, while still consolidating low demand trains. This chart can be seen in Figure 5.12.

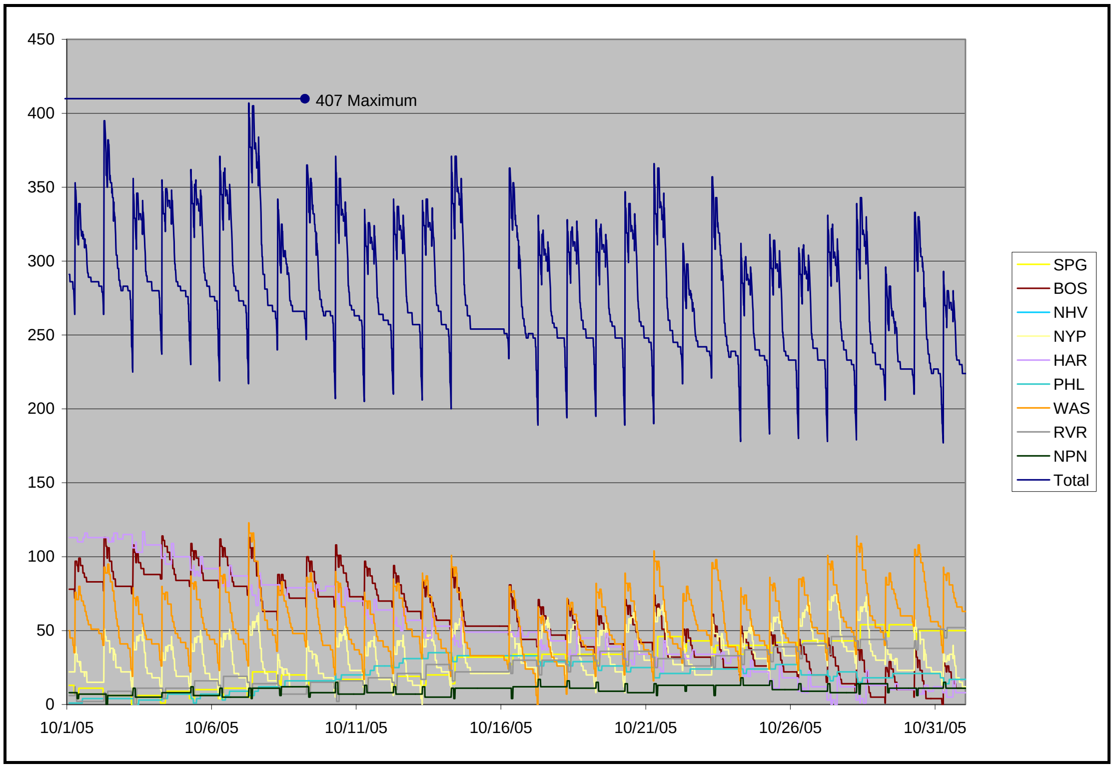

Fig 5.11 –MIN 3 – 75% Inventory Chart

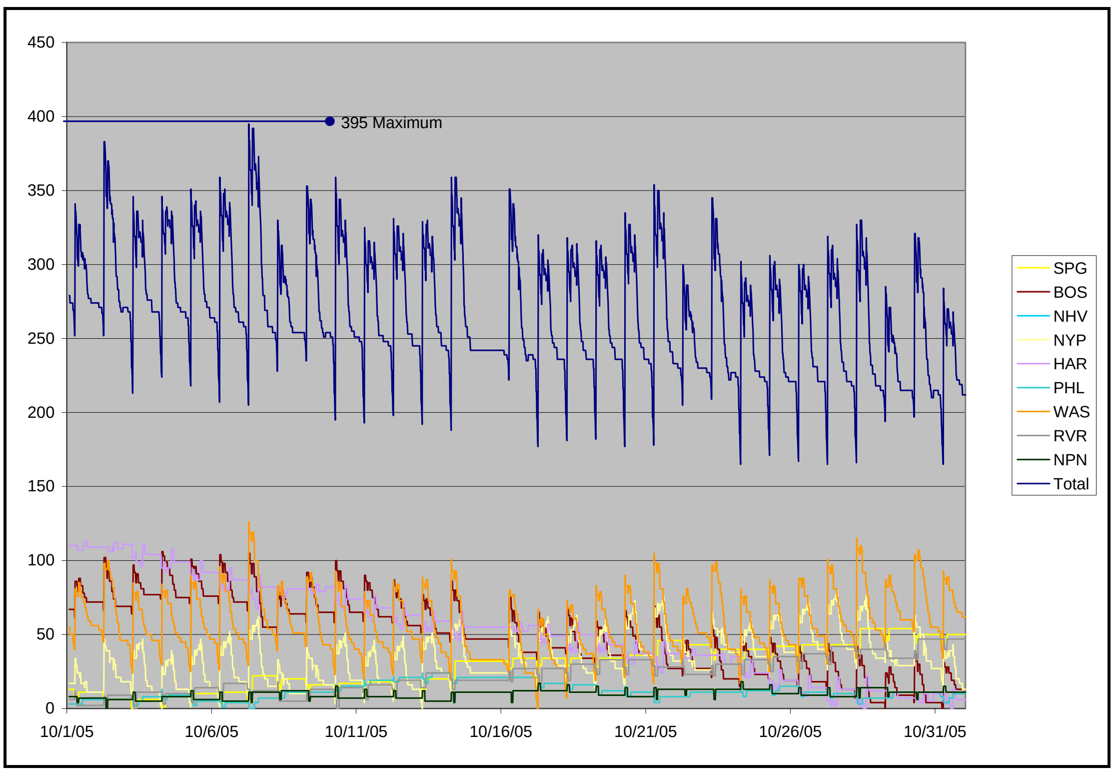

Fig. 5.12 – MIN 3 – 100% Inventory Chart

Additionally, any of the retention rates studied provides the greatest improvement in utilization of any case analyzed. As can be seen in Figure 5.13, the Peak LDF improves from 60.48% (Base Case) to 88.54% for 75% retention rate. In Figure 5.14 it can be seen that the 100% retention rate further improves the Peak LDF to 88.80%. Both of these are a larger improvement than the 84.35% achieved by the Full Model Optimization. Similarly to the Full Model Optimization, the proposed Peak LDFs peak between 85% and 100%.

An LDF-Miles Distribution was not prepared for the MIN 3 Case. This is due to the complexities of the algorithm's reassignment of ridership. In order for LDF-Miles to be calculated the algorithm would need to reassign ridership for each individual leg of a train's route rather than reassign the total ridership from one train to another (with a retention rate applied).

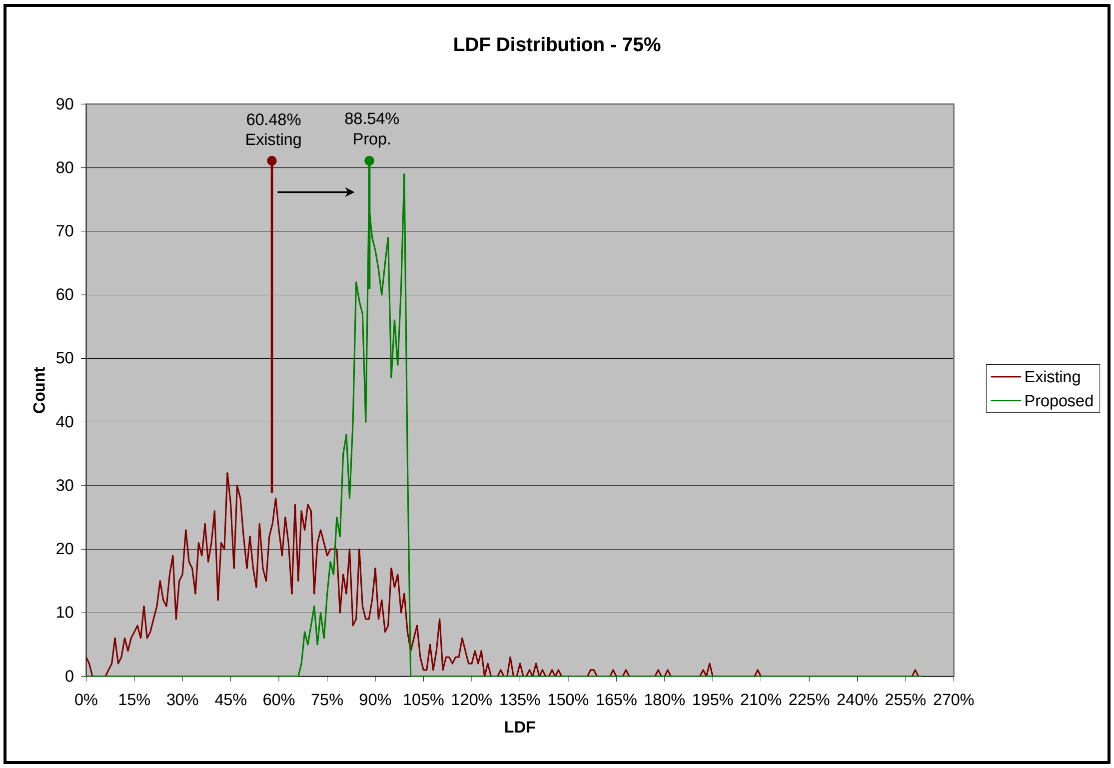

Fig 5.13 – MIN 3 Peak LDF – 75% Distribution

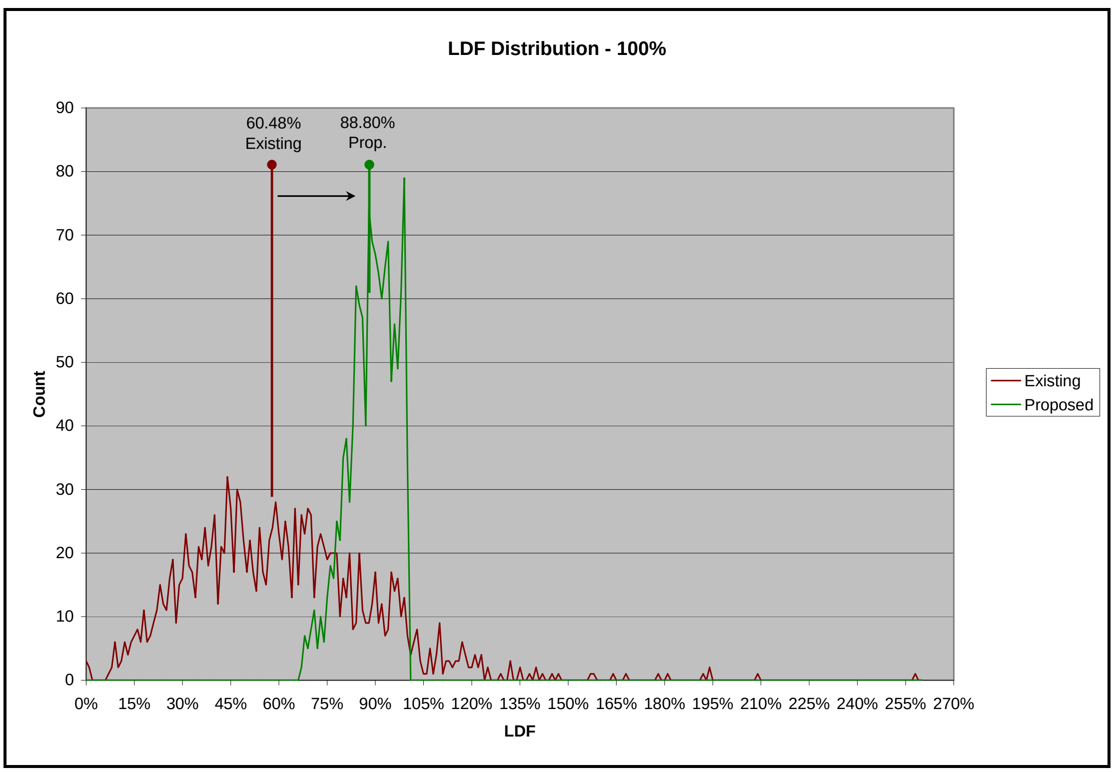

Fig 5.14 – MIN 3 Peak LDF – 100% Distribution

Since the 100% retention variation assumes all riders will transfer to another train, there is no ridership lost. However, under the 75% retention variation 6,116 or 1.50% of existing ridership is assumed lost. This assumed loss may occur due to inflexibility of schedule, inconvenience of fewer trains running (and therefore fewer options for travel), or other reasons.

Since any loss of ridership may affect the economic viability of a service, it useful to further examine the revenue versus operating costs of the remaining trains. The retention rates of 0%, 25%, 50%, 75%, and 100% were analyzed for ridership lost and revenues in order to generate enough data for proper analysis. As can be seen in Figure 5.15 as the Retention Rate is increased the ticket revenues increase far faster than operational costs increase. This means that the additional ridership generates more revenue than the additional costs to accommodate it. The data was generated by utilizing the Objective Function value as the operational costs. The revenue was generated by multiplying an average ticket revenue of $58.915 ($1.52 billion revenue divided by 25.8 million ridership, Amtrak p.10) by the ridership served.

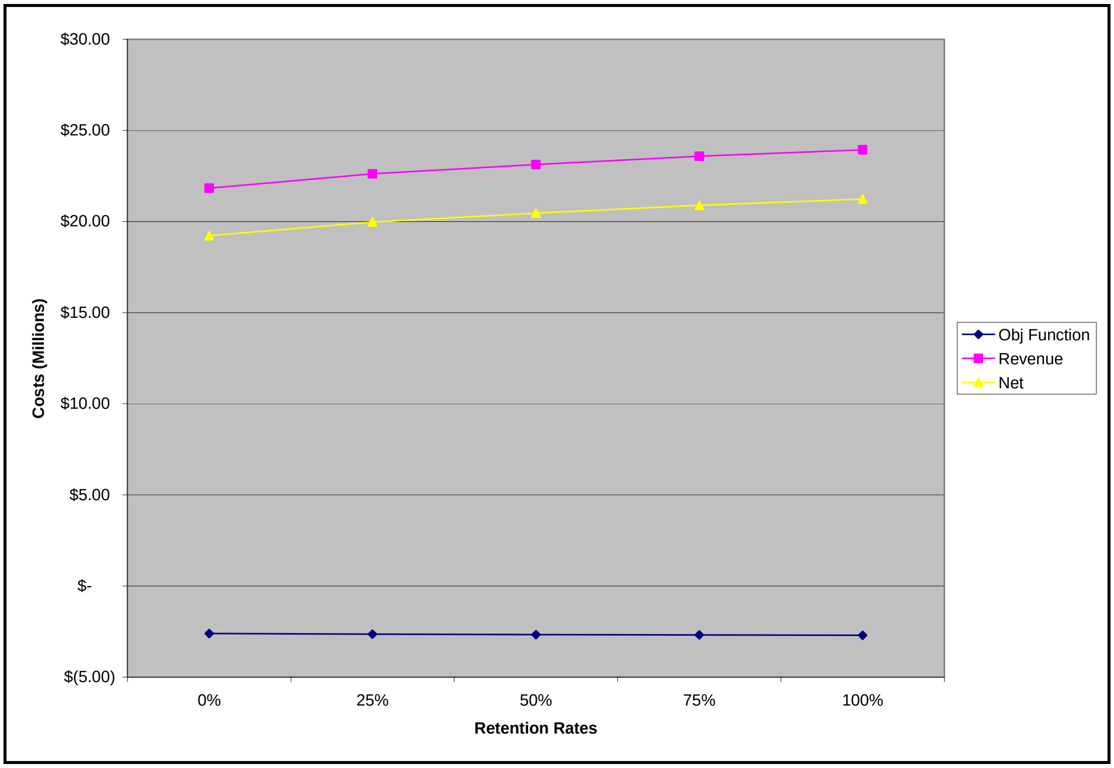

Fig 5.15 – MIN 3 Retention vs. Costs

The count of dropped trains is unevenly distributed amongst the various services. The overwhelming majority of the trains dropped from the schedule are from the Keystone Service, with 229 trains with 75% retention, 220 trains with 100% retention. This represents 75.08% and 74.32% respectively of all trains proposed for dropping, and in fact represents 44.21% and 42.47% of all the Keystone Service trains. The algorithm primarily recommends the elimination of westbound trains in the morning, the elimination of eastbound trains in the evening, and the sporadic elimination of trains in either direction during the mid-day. These recommendations are logical since the eastbound-AM, westbound-PM flows represents the logical commuter flows along the route (into and out from Philadelphia respectively). However, enough trains in the reverse direction and terminal inventories remain to allow the inventory balance constraint to be satisfied.

A smaller portion of the trains proposed for dropping are Regional Service (66 or 8.06% for both retention values). These trains represent 21.64% of all dropped trains for 75% retention and 22.30% of all dropped trains for 100% retention. Primarily the trains recommended for elimination are sporadic occurrences of early morning or late evening services in either direction. The only train that the algorithm outright eliminated is Train #151, the first train of the morning from NYP to WAS. This appears logical since the train has a 4:40am departure, 8:10am arrival – a very early time that gives it limited appeal to riders.

A small number of the Tidewater Service (10 or 5.43% for either retention value) trains are proposed for dropping. These dropped trains represent 3.28% of all trains proposed for dropping under 75% retention or 3.38% of trains proposed for dropping under 100% retention. The only train proposed for outright elimination by the algorithm is a Friday night train between Newport News, VA and Richmond, VA. It is assumed that this train is currently operated in order to balance cars, a maneuver that is determined unnecessary under this model. This theory is given credence by the existing ridership data – 3 riders on Oct. 7th, 0 riders on Oct. 14th, 4 riders on Oct. 21st, and 0 riders on Oct. 28th. It is interesting to note that all trains selected for elimination are northbounds. The model did not select any southbound trains for elimination as they all have a higher ridership level.

There are no trains on the Inland Service proposed for removal from the schedule. In fact, only 10% of Inland Service trains are proposed to operate with a consist of 3, the minimum under this case. The rest operate with longer consists due to higher ridership.

5.7 Term Cap Case

5.7.1 Purpose

Since the Full Model Optimization uses unconstrained terminal inventory capacity, it is logical to test the sensitivity of the results to imposing terminal capacities. This was accomplished by measuring existing trackage capacities for each terminal, then dividing the length by 85' (length of a typical passenger car) to arrive at the capacity. These capacities are shown in Table 5.3.

#	Code	Full Description	Cap Length	Cap Cars
1	SPG	Union Station, Springfield, MA	2,407'	28
2	BOS	South Station, Boston, MA	7,862'	92
3	NHV	Union Station, New Haven, CT	5,600'	65
4	NYP	Penn Station, New York City, NY	25,200'	296
5	HAR	Transportation Center, Harrisburg, PA	5,200'	61
6	PHL	30th Street Station, Philadelphia, PA	10,400'	122
7	WAS	Union Station, Washington, DC	7,400'	87
8	RVR	Staples Mill Road Station, Richmond, VA	2,000'	23
9	NPN	Newport News, VA	2,000'	23
		Overall Total	68,069'	797

Table 5.3 – Terminal Capacities

5.7.2 Implementation

The Terminal Capacity Case was implemented through the standard model with constraints added to the inventory terms. These constraints limited the inventory to being less than or equal to the capacity as measured in Table 5.3. Otherwise, no changes were made.

5.7.3 Results

The model returned consist results similar to the Full Model Optimization, but with lower counts for consists less than or equal to 6 cars and greater counts at longer consist lengths. The Consist Length Distribution can be observed in Figure 5.16. This is due to the model needing to send more cars through the system to avoid violating terminal capacities.

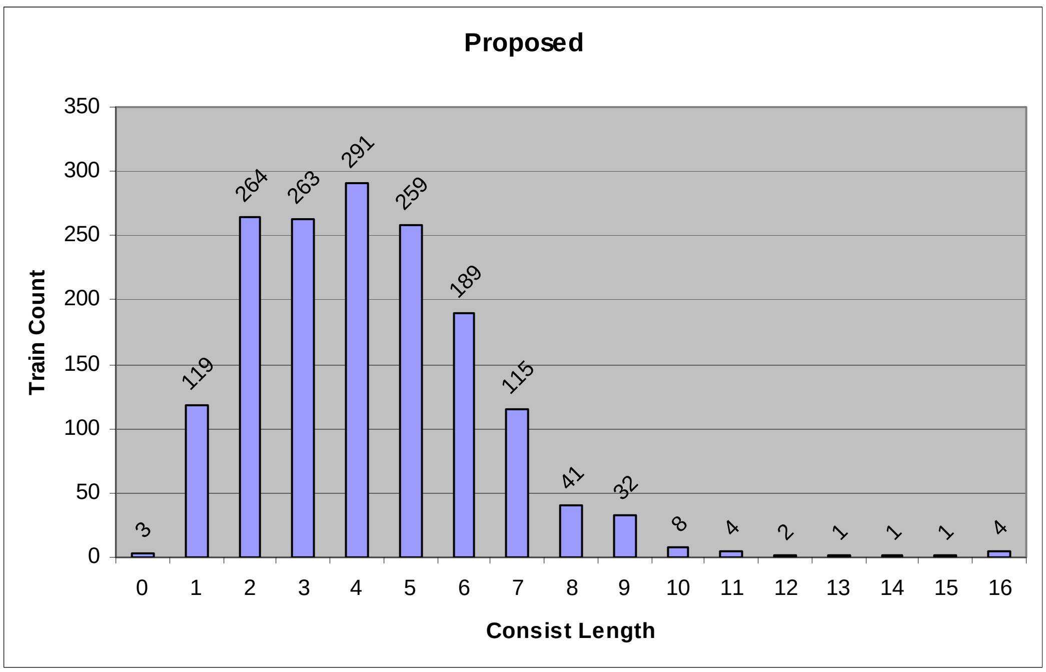

Fig 5.16 – Term Cap Consist Distribution

It is also useful for this case to study the Deltas Distribution. The Deltas can be observed in Figure 5.17. The figure bears a striking resemblance to the Full Model Delta Distribution, though with a slightly higher peak, and lower values at the outlying deltas. These shifts are logical, since a case with Terminal Capacities enforced would tend to stay closer to the existing car assignments than an unconstrained case would.

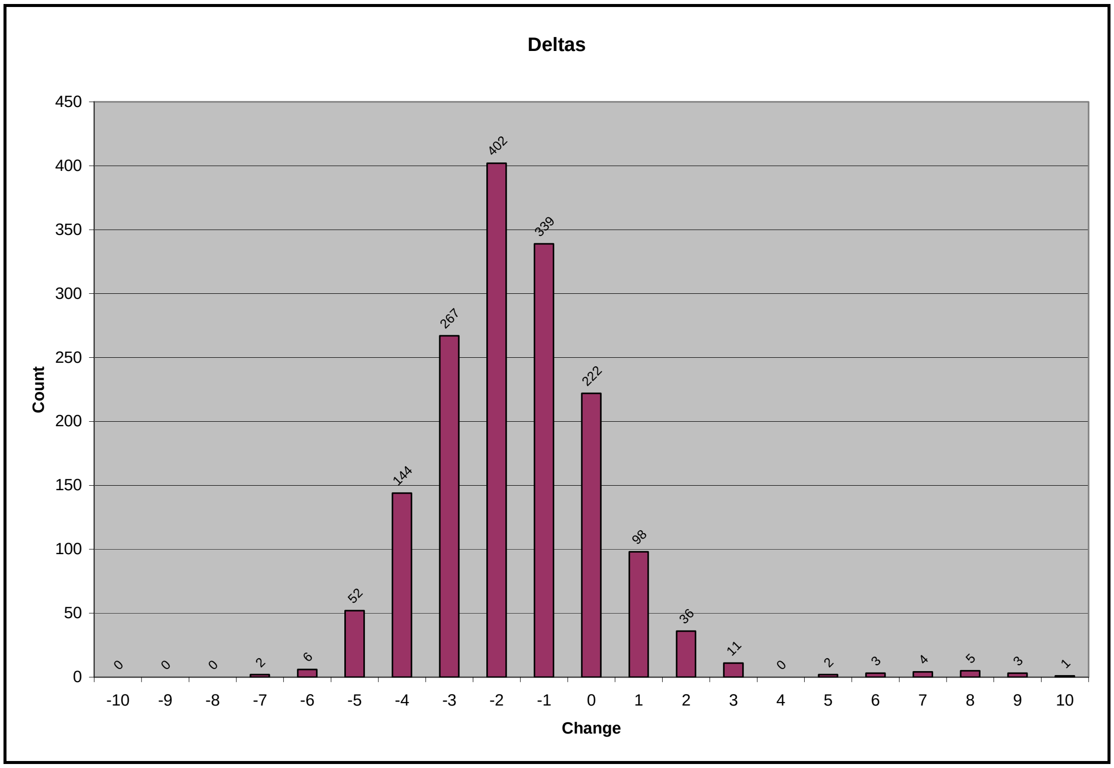

Fig 5.17 – Term Cap Deltas Distribution

The model returned an objective function of $2,540,471.48 to operate the trains within the month long period. This represents a $632,797.58 or 19.9% savings over the existing the Base Case operational costs. If projected out to an annual cost savings, the savings could reach approximately $7.6 Million. This is a smaller cost savings than the Full Model Optimization.

After analysis, the overall inventory requirement was determined to be 503 coaches. This represents a 37.0% (296 coaches) reduction over the Base Case. This reduction is larger than the reduction produced by the Full Model Optimization. This chart can be seen in Figure 5.18.

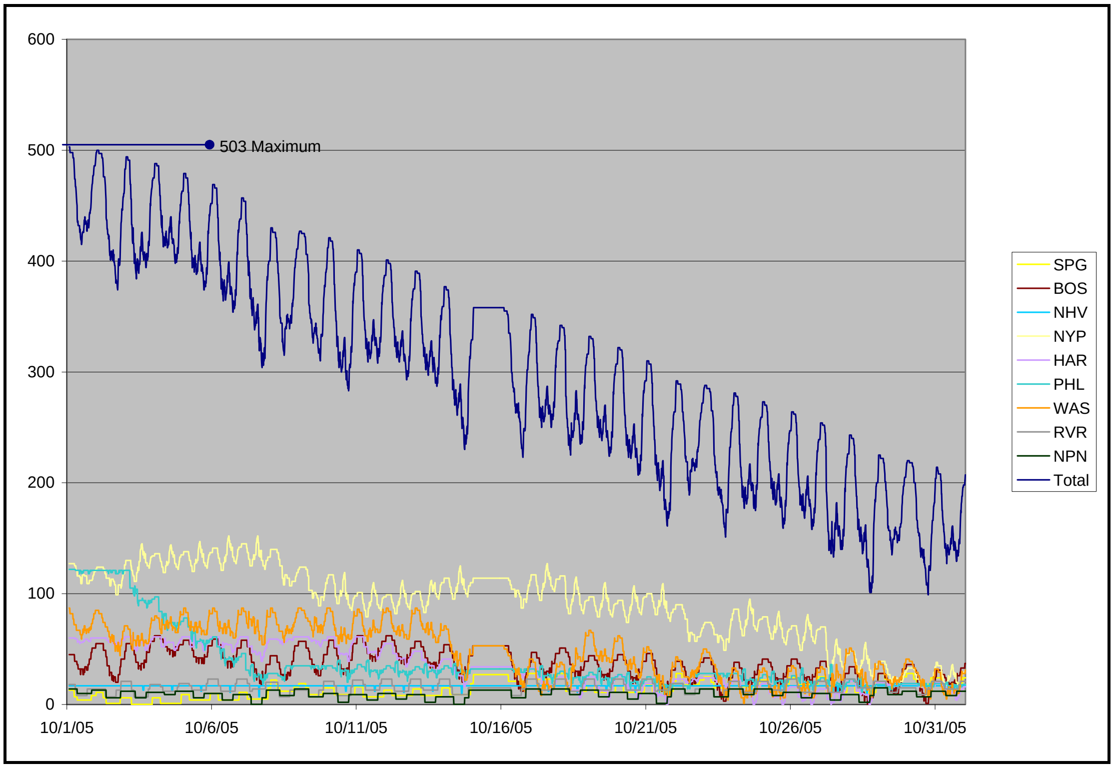

Fig 5.18 –Term Cap Inventory Chart

Also, this case does provide an improvement in utilization smaller than the Base Case. As can be seen in Figure 5.19, the Peak LDF improves from 60.48% to 82.77%. This is a smaller improvement than the 84.35% achieved by the Full Model Optimization. Similarly to the Full Model Optimization, the proposed Peak LDFs peak between 80% and 100%.

Likewise, the LDF-Miles Distribution (shown in Figure 5.20) shows an improvement. This factor improves from 48.53% overall to 65.18% overall, with the peak of the distribution now occurring between 65% and 75%. This is a slightly lower result than the Full Model Optimization's 66.90% overall LDF-Miles.

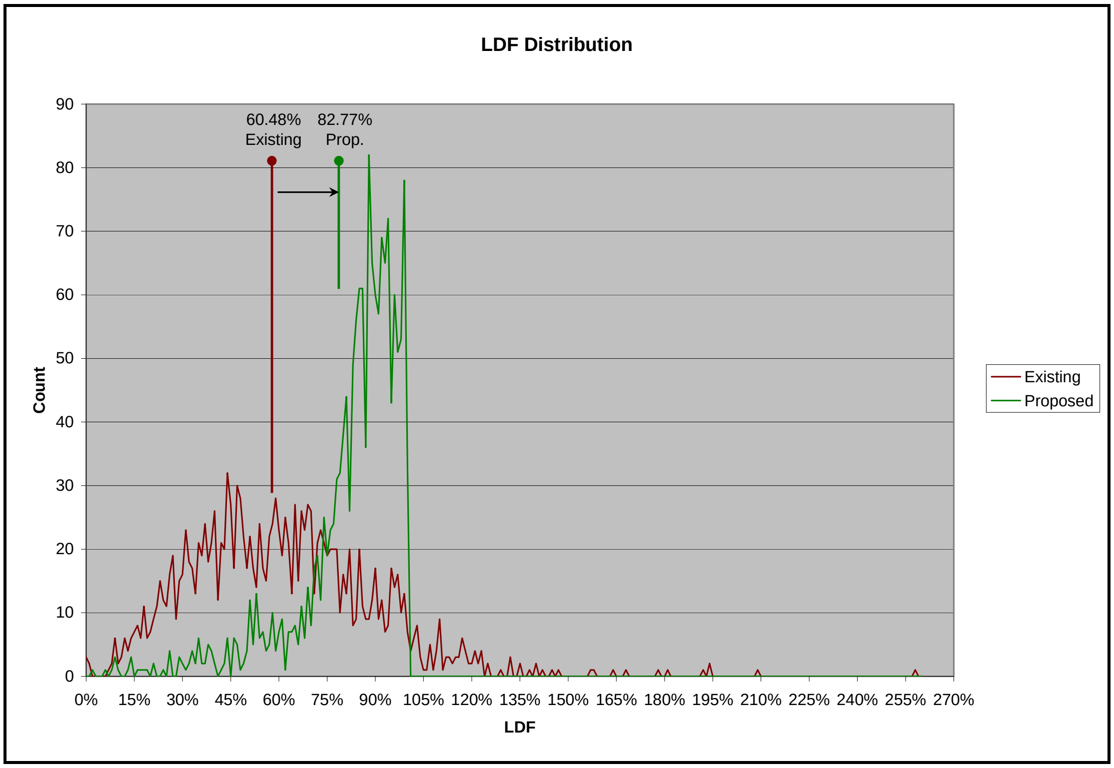

Fig 5.19 –Term Cap Peak LDF Distribution

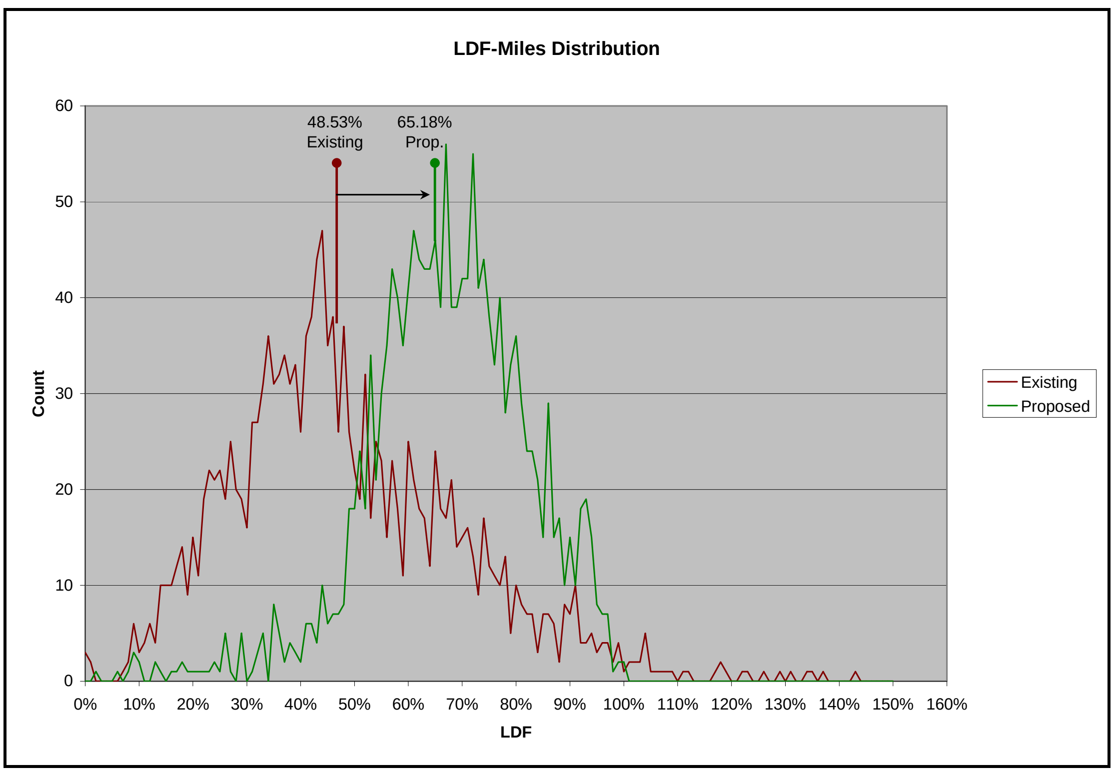

Fig 5.20 – Term Cap LDF-Miles Distribution

5.8 Other Attempted Cases

5.8.1 MIN 3 MAX 6

Since the minimum consist length considered economical is 3 cars, and the maximum number of coaches before another crew member is added is 6 cars, it is logical to attempt a case where consist lengths only within these limits are allowed. However, this case does disadvantage trains with high ridership. These longer, higher-demand trains require longer consists to meet demand, and under this case that demand would either be lost, or multiple trains would need to be operated. Since multiple trains is far less economical than adding an additional crew member to an existing train, the logical solution appears to be simply to add cars to existing trains. The added ridership gained by this additional capacity should offset the added ridership garnered. A more detailed analysis of revenue versus ridership could provide an answer as to exactly how many seats must be filled before a car breaks even.

5.8.2 Terminal 0

Though the Full Model Optimization assumed unconstrained capacity at Terminal stations for storing coaches, the purpose of this run was to determine the effects of limiting overnight capacity at Terminal Stations located outside the NEC Mainline. Therefore, the terminal stations at HAR, RVR, NPN, and SPG were modeled with an overnight capacity of zero. CPLEX immediately returned that the model is infeasible. Upon more detailed inspection, it was observed that at each station a morning outbound train originates before a morning train terminates at the given station. This does not allow any chance for a train to arrive and provide the needed inventory to commence outbound service in the morning. Without adding additional inbound morning trains, it is therefore impossible to hold any of these stations to a zero overnight inventory. The addition of inbound morning trains to each of these terminals would distinctly modify the existing service patterns and require ridership generation modeling. In consideration of these issues, it was determined that this case should therefore not be pursued further.

5.9 Summary

As shown, each Optimization Case attempted does offer improvements over the Base Case results. These improvements vary based upon which specific assumptions and constraints are utilized.

Based upon the objective function (operational costs) alone, the MIN 3 Case offers the most improvement. However since this is at the inconvenience of travelers (through the modification of the schedule), it should be noted that the Full Model Optimization offers the second highest savings. These results are shown in Table 5.4.

		Obj Function	% Diff	Proj Monthly	Proj Annual
Base Case		$3,173,269.06			
Full Model		$2,484,920.98	-21.69%	-$688,348.07	-$8,260,176.89
MIN 3	00% Retention	$2,609,577.34	-17.76%	-$563,691.71	-$6,764,300.56
MIN 3	25% Retention	$2,644,958.64	-16.65%	-$528,310.42	-$6,339,725.06
MIN 3	50% Retention	$2,669,560.23	-15.87%	-$503,708.82	-$6,044,505.90
MIN 3	75% Retention	$2,687,483.99	-15.31%	-$485,785.07	-$5,829,420.79
MIN 3	100% Retention	$2,705,761.82	-14.73%	-$467,507.23	-$5,610,086.80
Term Cap		$2,540,471.48	-19.94%	-$632,797.58	-$7,593,570.95

Table 5.4 – Objective Function Comparisons

Based upon the overall car fleet required to operate the service, the MIN 3 Case once again generates the largest savings. Without modifying the existing schedule, the next largest savings is accomplished under the Term Cap Case. These results are seen in Table 5.5.

		Max Inv	Cnt Diff	% Diff
Base Case		799		
Full Model		564	-235	-29.4118%
MIN 3	00% Retention	464	-335	-41.9274%
	25% Retention	403	-396	-49.5620%
	50% Retention	396	-403	-50.4380%
	75% Retention	407	-392	-49.0613%
	100% Retention	395	-404	-50.5632%
Term Cap		503	-296	-37.0463%

Table 5.5 – Car Fleet Comparisons

Based upon the Load Factor of the entire system, the Full Model Optimization Case offers the most improvement without modifying the existing schedule. Once again, if the schedule is allowed to be modified, then the MIN 3 Case produces large increases.

		Peak LDF	LDF-Miles
Base Case		60.48%	48.53%
Full Model		84.35%	66.90%
MIN 3	00% Retention	88.63%	-
	25% Retention	88.56%	-
	50% Retention	88.77%	-
	75% Retention	88.54%	-
	100% Retention	88.80%	-
Term Cap		82.77%	65.18%

Table 5.6 – Load Factor Comparisons

Since there is a different optimum case if the existing schedule is maintained, it is difficult to select a specific case for sole recommendation. If the schedule is allowed to be modified, then it becomes obvious that the MIN 3 Case produces superior results. However, since all cases show improvement over the Base Case then any of the cases analyzed are appropriate for implementation. It is likely that in a real-world situation a hybrid of all the cases analyzed may need to be developed for actual implementation. But the overall results are clear, that the model proposed is indeed effective at generating improvements over the existing car assignment approach.

Chapter Six: Conclusions and Recommendations

6.1 Conclusions

As demonstrated, the model can generate distinct improvements over existing car assignments. These improvements are present even when additional constraints (such as a minimum consist length and maximum terminal capacities) are added to the model beyond its basic form. These improvements take the form of lowered operating costs (varying from 19.9% to 71.0% reduction), reduced overall fleet requirements (varying from 29.4% to 52.7% reduction), and improved car utilization.

The model itself represents a basic approach to car assignment for simplified application. This is seen in the assumptions that allow the model to focus upon strictly coach assignment, which led to the development of a sparse model for coach assignments. This can allow quick modeling for optimization, but is inappropriate for longer-term planning because of the omitted considerations. One important omitted consideration is motive power assignment. A more thorough optimization model would need to consider the assignment of motive power as well as cars in order to properly consider the utilization of all equipment, not just the passenger coaches. Also, the model assumes that there is no cost for car storage, which is simply not the case. In a real world situation, there are inspection and maintenance cycles that generate costs, as well as the infrastructure costs of the storage facilities. The implications for the results presented here are felt to be minor, but a model with a larger modeled area or longer planning horizon would need to account for the costs of equipment storage as well as equipment operation. These assumptions do narrow the applicability of the model, but it is still a useful tool for minor changes and readjustments due to shifting ridership patterns.

It is the hope of this thesis that the simplicity of the model presented assists in either the quick adoption of this model, or its continued development for application. With the simplified form of the model an intercity passenger rail operator could easily, quickly, and cheaply implement the model on their services. It is conceivable that the specific implementation could be adapted to a more widely available program than CPLEX, perhaps even into EXCEL through the use of macros, to reduce barriers to its use at operators or agencies with limited resources.

6.2 Amtrak Recommendations

6.2.1 Service Changes

It is apparent from the conducted work that through the application of a ridership/demand based model efficiencies can be realized. These efficiencies would amount to lower operational costs, fewer equipment needed overall, and better utilization of the existing fleet.

A specific approach to realize these improvements is to reduce consist lengths as recommended by the model. If doubt remains as to its efficacy, then it would be beneficial to target those trains with particularly large negative deltas. It is these trains that have significant excess capacity and therefore would produce noticeable improvements to the system as a whole if modified. Once the efficiencies are proven, the modification of the consists of a wider number of trains can then proceed.

6.2.2 Data

A specific recommendation is that Amtrak develops a comprehensive demand survey for their service area. This would be able to generate far more reliable source data for models such as this one. Consisting and car optimizations could then be based upon predicted ridership and accurate predictions of true demand rather than historic ridership data. Reliable demand data would also be useful in determining sensitivity to timetable changes and pricing changes, particularly those pursued under the MIN 3 Case described in the Implementation and Results Chapter.

It is particularly important to note that all of the work presented here is conducted on historic ridership data. Though the model shows improvement over the existing approaches, it is of little use to improve previously occurring services. Therefore, an accurate ridership prediction model would provide the needed future ridership data to input into the proposed model, allowing it to correctly propose future consists.

6.3 Further Optimization

6.3.1 Amtrak

It is recommended that Amtrak continue to re-optimize its car assignments in order to ensure their matching to current ridership. These re-optimizations would need to occur periodically, possibly as a second step after the results of an updated demand survey are available. That way the consists are updated to match the demands of the travelling public.

6.3.2 Other Applications

Obviously the model developed here can be easily applied to any application with regularly scheduled service and uniform capacity vehicles. This means that within rail applications, the model could be applied to any other passenger system easily (both transit and commuter), as well as to freight rail applications. For example, the model could be utilized to model coal cars assigned to service between a network of coal mines and a network of power plants. The model could also be applied to high-demand bus operations where multiple vehicles per timetable slot are required. However, there is a limit to the size network that the model can be implemented on, due to the assumptions about car maintenance, motive power assignment, and crew scheduling. A larger network would necessitate the addition of maintenance planning to ensure cars and motive power are at the maintenance facility when inspections are due. Crew movements would also need to be incorporated to ensure adherence to service limits and required rest periods.

However, a less obvious application would be to motive power operations. As cited in the Literature Review, Kuo utilizes a very similar model to approach freight locomotive assignment. This could also be utilized for application to freight locomotive assignment, but care must be exercised when assessing the homogeneity of the locomotive fleet in use.

6.4 Further Research

6.4.1 Amtrak

Further research into the time-sensitivity of Amtrak's ridership would be beneficial. This would allow decisions to be made as to how much of a time change for departure would be acceptable to the ridership. This knowledge could then drive the allocation of cars and trains to increase efficiency. Specifically, this knowledge would allow a better analysis of the MIN 3 Case presented above and its specific effects upon ridership. If the ridership is willing to shift departure/arrival times, the effects of train elimination upon the ridership could be minimized. This would then further increase the efficiencies of the trains that would benefit from ridership increases from eliminated trains.

An interesting wrinkle to the optimization problem requiring more research is the State Sponsored Corridor. Amtrak operates several services through funding from States (including the Keystone Service discussed above). This funding dictates that a certain number of trains must be operated on a route, no matter their inefficiency. It therefore would be of great use to conduct further research into appropriate methods to schedule and assign cars to these trains while still meeting the requirements of the Sponsor State.

6.4.2 Academia

Obviously Academia would be of great assistance in the further Amtrak Research recommended above. This includes extending and expanding the model proposed here to model a national rail network. This would entail adding constraints for crew scheduling, motive power assignment, and accommodations for multiple, unique car-types. Academia would also be instrumental in collecting and analyzing data to determine attributes necessary to generate an expanded model. This is particularly true for a ridership demand model, but would also be necessary to determine the effects of train consolidation upon ridership.

In addition to this work, there is further research that academia could conduct independently. This research would need to be of a network nature, targeted at the overall transportation system rather than a specific mode. Therefore, a beneficial study would be the total passenger demand between various metropolitan areas. This demand could then be used to determine inter-city passenger rail's current mode split. With the knowledge of existing mode splits, consist extensions could then be targeted to underserved markets, while consist reduction could be targeted to over-served markets.

Glossary

α (Load Factor) or LDF – A measure of utilization, calculated by dividing the ridership by the available capacity.

Amtrak – Business name of the National Railroad Passenger Corporation, a for-profit state related passenger railroad company.

Consist – the set of equipment that forms a train. For this study, the term merely means the set of coaches assigned to each train.

Couple – the connecting of 2 pieces of rail equipment.

Deadheading – The practice of operating equipment in nonrevenue service. This could be as a separate train, or as part of a train that is in revenue service (i.e. – a closed car on the end of a passenger train)

(Train) Links – A linked series of train #s operated by the same consist.

NEC – Northeast Corridor, Amtrak's highest ridership and service corridor, primarily between Boston, MA and Washington, DC.

Terminal Station – A major station where infrastructure exists to modify consists and store unused coaches.

Bibliography

Amtrak. 2007 Annual Report. Amtrak Website, http://www.amtrak.com/pdf/AmtrakAnnualReport_2007.pdf.

Ahuja, Ravindra K., Liu, Jian, Orlin, James B., Sharma, Dushyant, and Shughart, Larry A., 2005. Solving Real-Life Locomotive-Scheduling Problems. Transportation Science Vol. 39, No. 4, 503-517.

Booler, J. M. P., 1980. The Solution of a Railway Locomotive Scheduling Problem. Journal of Operational Research Society Vol. 31, pp.943-948.

Bussieck, M. R., Winter, T., and Zimmermann, U. T., 1997. Discrete Optimization in Public Rail Transport. Mathematically Programming 79, p.415-444.

Cacciani, Valentina, Caprara, Alberto, and Toth, Paolo. Solving a Real-World Train Unit Assignment Problem. ATMOS 2007 - 7th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems."

Charnes, A. and Miller, M. H. A Model for the Optimal Programming of Railway Freight Train Movements. Management Science 1957 No. 3, p.74-92

Cordeau, Jean-Francois, Desaulniers, Guy, Lingaya, Norbert, Soumis, Francois, and Desrosiers, Jacques, 2001. Simultaneous Locomotive and Car Assignment at Via Rail Canada. Transportation Research Part B 35, 767-787.

Cordeau, Jean-Francois, Soumis, Francois, and Desrosiers, Jacques, 2000. A Benders Decomposition Approach for the Locomotive and Car Assignment Problem. Transportation Science Vol. 34, No. 2, 133-149.

Cordeau, Jean-Francois, Toth, Paolo, and Vigo, Daniele, 1998. A Survey of Optimization Models for Train Routing an Scheduling. Transportation Science Vol. 32, No. 4, 380-404.

Davis, Stacy C., Diegel, Susan W., and Boundy, Robert G., 2008. Transportation Energy Book: Edition 27. Oak Ridge National Laboratory Website, http://cta.ornl.gov/data/Index.shtml.

Florian, M., Bushell, G., Ferland, J., Guerin, G., and Nastansky, L., 1976. The Engine Scheduling Problem in a Railway Network. INFORMS vol. 14, no. 2, p.121-138.

Forbes, M. A., Holt, J. N., and Watts, A. M., 1991. Exact Solution of Locomotive Scheduling Problems. The Journal of the Operational Research Society, Vol. 42, No. 10, p.825-831.

Hong, Sung-Pil, Kim, Kyung Min, Lee, Kyungsik, and Park, Bum Hwan, 2008. A Pragmatic Algorithm for the Train-Set Routing: The Case of Korea High-Speed Railway. Omega: The International Journal of Management Science 37, p.637-645.

Karush, Sarah, 2008. Amtrak: Business, politics in train tug-of-war: Associated Press, http://ap.google.com/article/ALeqM5hCED5U9ot1BNfd4qyiyvbH5f1LdwD94ASGSG0.

Kuo, Ching-Chung and Nicholls, Gillian M., 2005. A mathematical modeling approach to improving locomotive utilization at a freight railroad. Omega: The International Journal of Management Science 35, 472-485.

Ramani, K.V., 1981. An Information System for Allocating Coach Stock on Indian Railways. Interfaces Vol. 11, No. 3, p.44-51.

Ramani, K. V. and Mandal, B. K., 1992. Operational Planning of Passenger Trains in Indian Railways. INTERFACES 22, 39-51.

Wright, M. B., 1989. Applying Stochastic Algorithms to a Locomotive Scheduling Problem. The Journal of the Operational Research Society Vol. 40, No. 2, p.187-192.

Vuchic, Vukan, 2005. Urban Transit: Operations, Planning, and Economics. John Wiley & Sons, Inc., Hoboken, NJ.

Ziarati, K., Chizari, H., and Nezhad, A. Mohammadi, 2005. Locomotive Optimization Using Artificial Intelligence Approach. Iranian Journal of Science & Technology, Transaction B, Engineering, Vol. 29. No. B1, p.93-105.

Ziarati, Koorush, Soumis, Francois, Desrosiers, Jacques, Gelinas, Sylvie, and Saintonge, Andre, 1997. Locomotive Assignment with Heterogeneous Consists at CN North America. European Journal of Operational Research 97 p.281-292.