Pre-Meeting Proceedings*

McVey, Pautsch, Fruin & Baumel
Agricultural Transportation Infrastructure Issues in the 21st Century

Transportation Research Forum

34th Annual Meeting

Editor:
Russell B. Capelle, Jr.

October 21-23, 1992
St. Louis, Missouri

*"The St. Louie Book"

Title:

Car Cycle Analysis and Network Modeling at CSX Transportation: Using the "QBASE" File

Author:

Edwin R. Kraft, Director, Operations Research, CSXT

Affiliations and Addresses:

CSX Transportation
Speed Code J230
500 Water Street
Jacksonville, Florida 32202
Tel: (904) 359-1171 Fax: (904) 366-5433

Equipment Request:

Transparency Projector

Summary:

This paper describes the implementation of the Service Planning Model family of tools at CSX Transportation, a major U.S. rail carrier. In order to successfully implement this simulation model, it was necessary to build a programmatic interface to collect required input data. An enhanced car cycle file called "QBase" was developed as a part of this effort. The paper describes how QBase is used in the context of the overall Service Planning process at CSXT; and describes a specialized data-driven QBase application which identifies bypass blocking opportunities.

Car Cycle Analysis and Network Modeling at CSX Transportation: Using the "QBASE" File

Edwin R. Kraft, Director, Operations Research, CSXT

Introduction

CSX Transportation and predecessor roads have a long history of building and utilizing Operations Research tools. As a member of rail industry user group committees, CSXT co-sponsored the development of the Automated Blocking Model (ABM)[1] , Train Scheduling System (TSS)[2] and Service Planning Model (SPM)[3]. These software packages are maintained by ALK Associates, Princeton, NJ. CSX has also licensed ALK's proprietary graphics system, the Princeton Transportation Network Model (PTNM)[4] .

In September 1989, the CSX Transportation Operations Research (OR) group was formed. Our group's primary focus has been the *implemention* of these existing programs, so we have not undertaken a great deal of new model development. We have, however, undertaken substantial *data feeder system* development to provide inputs for the models, linking them with internal CSX systems.

Along the way, we acquired a considerable capability in the analysis of historical car movement data. This paper will focus on those *data-driven* applications which were primarily designed and implemented by the author. At the same time, the foundation will be laid for future publications, describing other applications developed within the CSX Operations Research group.

Initial Efforts at Network Modeling at CSXT

Prior to the formation of the OR group, the implementation of a version of the ABM and TSS had been carried on by a single individual and applied successfully on a number of projects. Using these inherited ABM and TSS databases, we attempted to manually build an SPM base case for CSXT's 19,000 route mile rail network[5]. But, with new traffic data, many flows were stranded at intermediate points, needing additional blocks or trains to move all the way from origin to destination. An inordinate amount of manual effort would have been required to obtain and enter all the missing data, and that would only have supported a one-time use of the model.

We concluded that the sheer volume and complexity of model input required an automated data feed process be built, linking the Service Planning Model with CSX's production data systems. This was the genesis of the "QBase" file[6] at CSXT.

What is the "QBase" File ?

In the "QBase" file, individual car movement reports are linked together to form trips or "cycles". QBase contains movement detail for all rail equipment moving on CSXT. It is created by merging several data sources together to collect all the information we need. The context diagram for this process is shown in Figure 1.

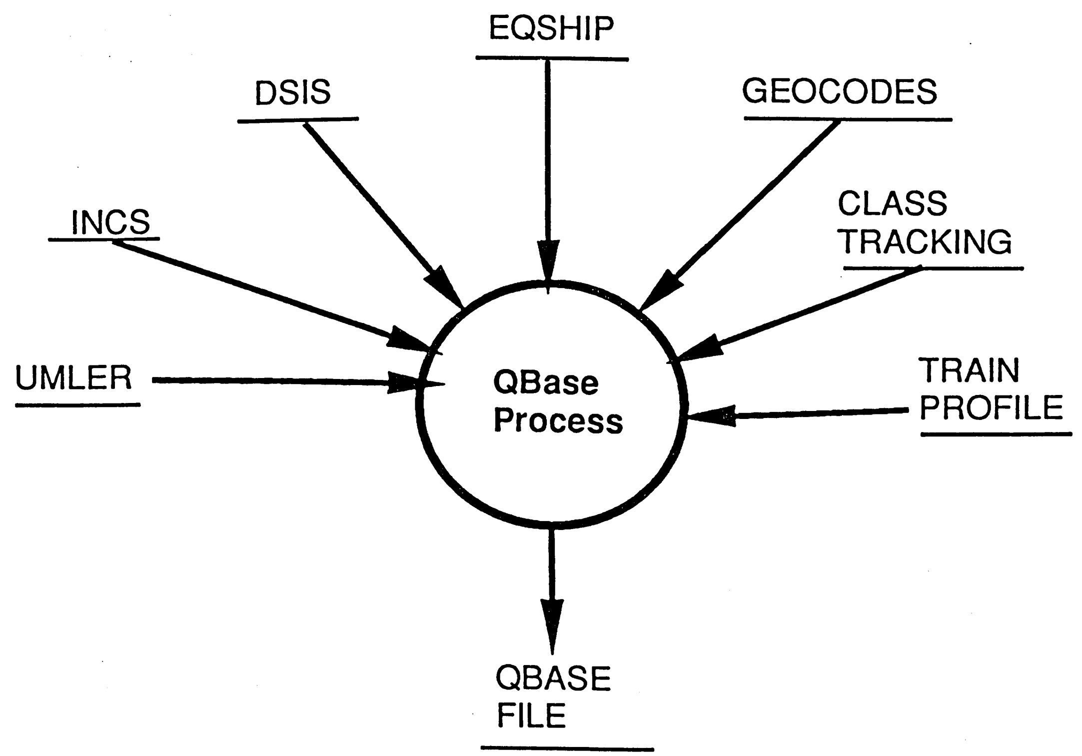
- The QBase process starts by extracting both loaded and empty car movement detail from the "EQSHIP" car cycle file.

- Then, revenue, variable cost, contribution and line of business codes are added onto each loaded record from "DSIS".

- Special car-related data such as length, tare weight and axle count are added by matching to the "UMLER" file.

- We "geocode" the file to convert city/state and milepost notation into PTNM and ABM network node numbers.

Figure 1- QBASE File Context Diagram



- Load/empty, empty/load and empty/empty association keys are added from "INCS". Using these keys, empty and loaded car cycle records are linked so that round-trip cycle time analysis can be performed. Associated load information, such as line of business code, is carried forward (or backwards) onto each empty record.

- Intermediate Yard Switching Codes (IYSC) and Train Block Codes are added from the Class Tracking and Train Profile data bases. In the future, this information will be captured directly from car scheduling trip plans.

- A "Handling Code" is calculated, using a combination of train ID and block coding information to *summarize the type of handling each car received at each location.* Two different kinds of handling codes are defined: Alphabetic codes, when special, customer or interchange activity occurs; and numerical codes, when only standard yard handling occurs. These codes are shown in Figure 2.

- As shown in Figure 3, the last step is final reformatting of the QBase file. While the original source data is in hierarchical header/trailer format, the QBase record is "flattened out" so that all information for the trip is contained on a single data record. Up to 15 handlings can be stored on each QBase record. In this format, the information is generally much easier to use.

Figure 2.- QBase Handling Code Values:

Alphabetical Values for Special "Zone" Activity:

Code Value	Definition
"P"	Patron Activity
"F"	Interchange
"R"	Repair Track
"C"	Cleaning Track
"S"	Storage Track
"A"	Abnormal Event

Numerical Values for Normal Activity:

Code Value	Definition
"0"	Intermediate Reporting Point Only-Car was Not Handled
"1"	Block Swapped - Train ID Changed- Blocking Remained Constant
"2"	Expedited Reclassify - Blocking changed, but Train ID Remained Constant
"3"	Standard Reclassify- Both the train ID and blocking changed

Figure 3.- Original Hierarchical vs Flattened QBase Formats

Source Data File Format: (Multiple records per cycle)

Header Record:	CSXT 123456 From A to D
Movement Record:	Customer Release at A
Movement Record:	Run from A to B
Movement Record:	Run from B to C
Movement Record:	Run from C to D
Movement Record:	Customer Placement at D

Reformatted QBase Record: (Single record per cycle, up to 15 Yards)

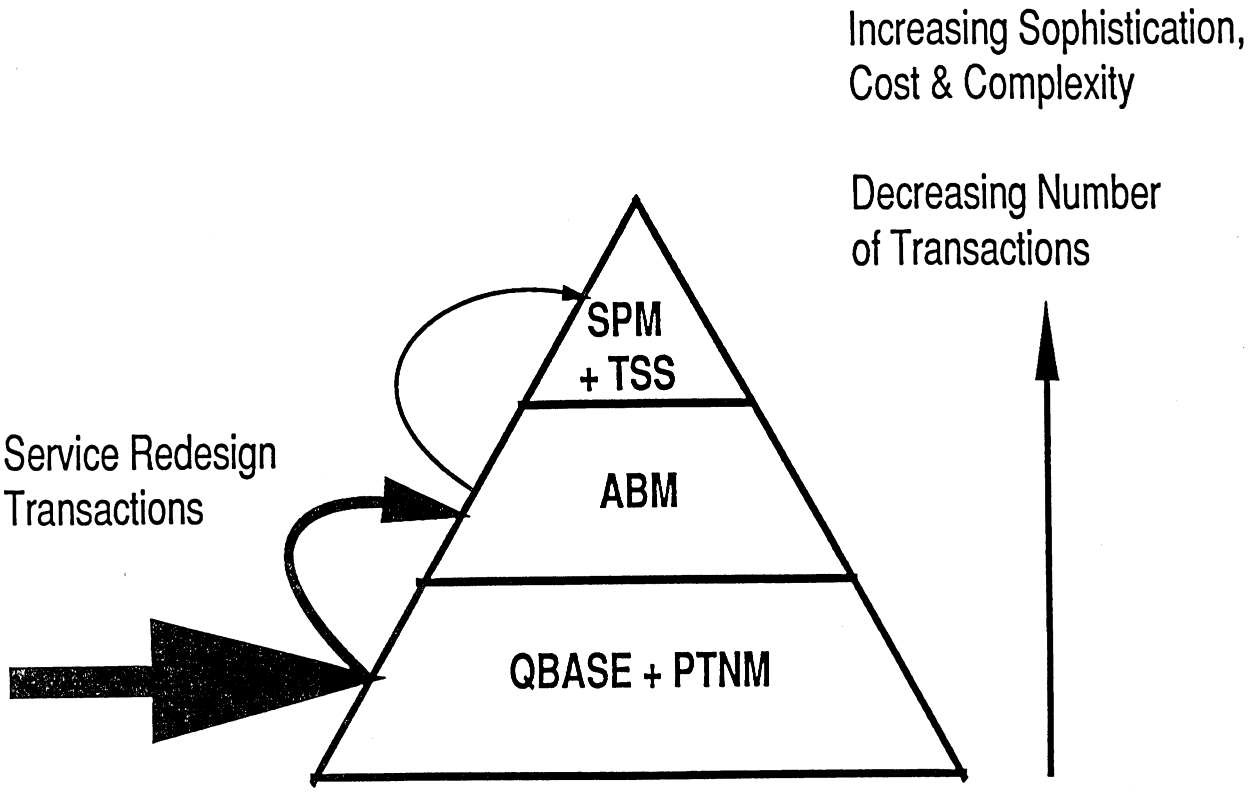
CSXT 123456	Yard A	Yard B	Yard C	Yard D
	Rel Dep	Arr Dep	Arr Dep	Arr Plc

How QBase is Used

We have found the vast majority of Service Design's information needs to be straightforward data requests. Users ask for traffic volume and car handling counts, trip and yard time distributions, flow density maps, and car mile calculations. These needs, which do not require sophisticated modeling or "what if" capability, are satisfied through a retrieval and reporting system using the QBase file.

Figure 4 shows the relationship between QBase and the ABM, TSS and SPM models. QBase serves as the foundation for the entire modeling pyramid, providing all input data to drive the higher elements of the modeling process. Nearly *all* service redesign projects begin by using QBase inquiries to gain basic knowledge of traffic flows and current handling patterns. The vast majority of service redesign projects are handled at this lowest level of the pyramid, never requiring analysis at a higher level.

Figure 4- Relationship of QBase to Modeling Tools



As shown in Figure 4, the cost per QBase transaction is low, and many alternatives can be quickly explored utilizing the PTNM flow mapping capability. QBase allows a user to "bootstrap" from a level of zero knowledge into a fairly complete understanding of traffic flows and routing options.

A moderate number of alternatives might survive into the partial specification stage, where they are evaluated by ABM at a moderate cost per transaction.

Only a few alternatives are likely to survive as fully specified operating plans, where they can be evaluated by SPM at a relatively high cost per transaction. As a service redesign project moves forward, increasing detail becomes available, but operating alternatives are also weeded out, so that *only one or two prospective operating plans are likely to survive to the fully specified level.*

Figure 5 -Degree of Operating Plan Specification Required by Network Models

Model	Level of Specification Required
QBase + PTNM	None
ABM	Partial - Blocking Plan Only
TSS + SPM	Complete - Full Operating Plan, including train schedules and block to train assignments must be furnished

As shown in Figure 5, *the extensive data requirements of the higher-level models effectively prevent their use during the initial, formative stages of a project—for the simple reason that the operating plan is simply not yet known, or not yet specified to the level of detail required by higher level models.*

Typical service design projects move through three distinct analytical stages:

- **Decomposition**

In the early stages, the analyst tries to understand what traffic volumes are to be handled, service levels currently provided, and what trains and blocks are currently moved. This stage is a discovery process: removing one layer of aggregation at a time, until the transportation requirements are fully understood at their lowest component level. QBase is the tool most appropriate for this phase.

- **Reaggregation**

The next task is a goal-driven reaggregation of flows into new blocks and trains. The objective might be expressed as cost minimization subject to meeting preestablished service constraints. The ABM and TSS are the tools most appropriate for this phase, and QBase PTNM flow density mapping[7] can also be used.

- **Evaluation**

The last step is a systematic evaluation of the final proposed operating plan, to make sure that all service and cost objectives have actually been met. This optional step may not always have been performed adequately in the past. Our objective is that evaluation (and, after implementation, post-audit) should be performed as standard procedure in the future. The SPM is the tool most appropriate for this phase.

When used together, these tools become a fully integrated decision support system, addressing differing user needs during *all* stages of a Service Design project.

SPM Model Interface

The SPM Model Interface builds all the input files required by the ABM, TSS and SPM from a QBase input file. In addition, since all railroad network models share common data requirements, the interface can easily be adapted to any other railway network model[8]. Our modeling techniques are now much more "reality based". The features of this interface are detailed in Figure 6 on the next page.

Direct QBase Applications

A variety of standardized, prepackaged QBase reports can be executed at any time. Users can easily customize these programs to their own applications, or write entirely new reports. Prepackaged PTNM graphics applications can be executed from a simple menu system. Output goes directly to a graphics printer. No knowledge of programming is required to utilize this system. It is just as easy to produce a graph as it is to make a printed report.

Figure 6- Inputs used in ABM, TSS, SPM Models Before and After Implementation of the SPM Interface

	Before	After
O-D Traffic Flows	Loaded Flows: ICC 1% Waybill Sample Empty Flows: Generated via Reverse route	Loaded & Empty: 100 % Sample from QBASE representing actual, rather than waybilled movement patterns
Classification & Line Haul Costs	Manually calibrated to make shortest-path flows match actual flows.	Productivity-based economic line haul and terminal costs[9]
Blocks Built at Each Yard	Manually maintained from Operating Plan	Programmatically built from QBase
Car to Block Assignment	Used shortest-path model-based assignments.	Use "tag table" based on actual movement patterns
Train Schedules	Manually maintained from Operating Plan: *regularly scheduled* trains only	Statistical analysis of actual movement data. *All* trains, with actual frequency, arrival and departure times
Block to Train Assignments	Manually maintained from Operating Plan	Programmatically built from QBase
Connection Cutoff Times	Based on Operating Plan	Based on Statistical "PMake" analysis of historical connection performance

A standardized data selection program, the "Generic Retriever", is typically used as a front-end to all applications. The *sole purpose* of this program is to retrieve and manipulate QBASE records. The output of the generic retriever is always another, subsetted, QBASE file.

The generic retriever extracts records based on a wide variety of car and trip-related header fields, but the real power of the retriever is in matching detailed car movement information. Geographical selection identifies records having up to three nodes in common. The nodes must occur in order, but do not necessarily have to be the origin, destination or adjacent to one another. Records can be extracted based on *intermediate* yards, corridors, train ID's and arrival/departure times.

Using the retriever, users can select records from:

- Any set of origin nodes,

- across a single intermediate "pivot" node,

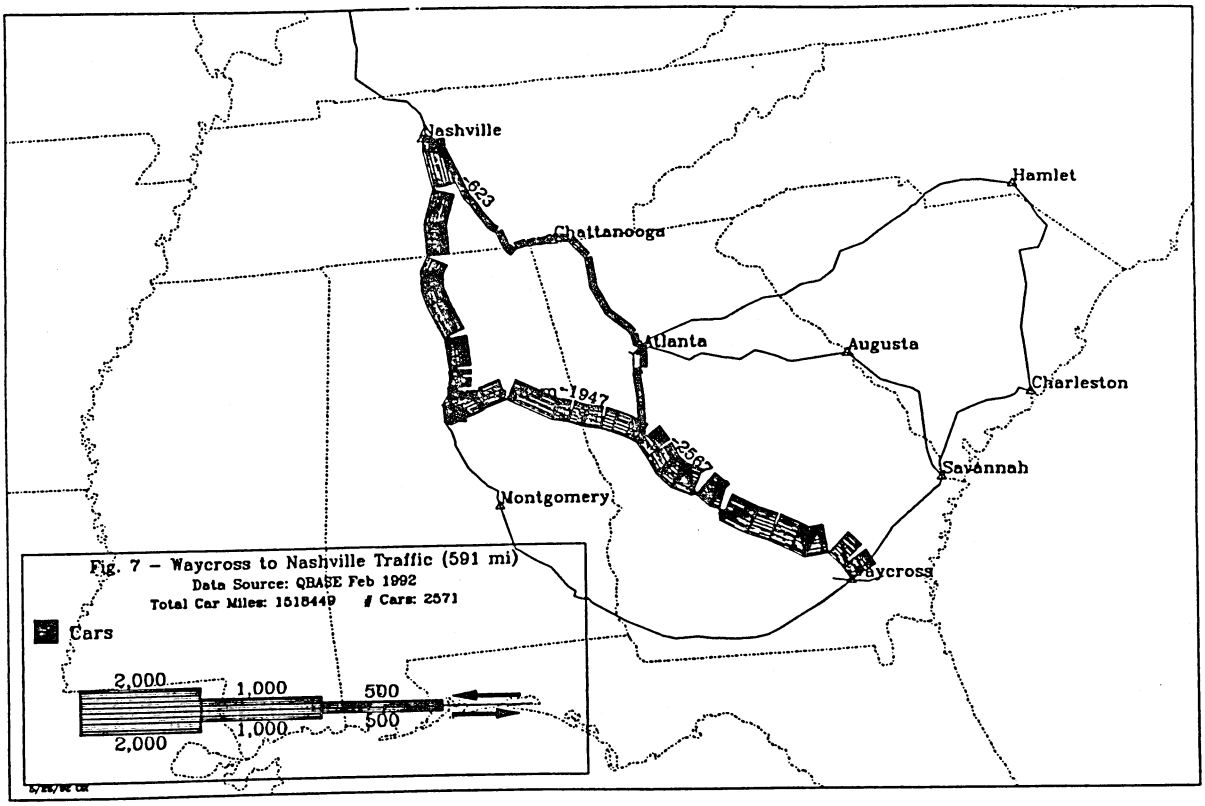
- to any set of destination nodes.

- or, any combination of the above.

Surgical functions allow users to modify or remove unwanted information from within a QBASE record. The most common surgical function is elimination of movement cells based on handling code. This is accomplished by sliding the rest of the QBASE record one position to the left, overwriting the previous contents of the field. The primary use of this feature is to eliminate undesired *intermediate* points, *such as locations where a car was not handled,* from the QBASE record.

The second most common surgical application is the clipping function. This function removes the *first or last part* of the record. Portions of trips that are prior to the desired "origin yard" or subsequent to the "destination yard" are clipped off. This has the effect of changing the origin and destination of the trip. *The primary use of this feature is to perform "corridor" or yard-to-yard flow analysis.* To demonstrate the clipping feature, a typical QBase-generated PTNM flow density map is shown in Figure 7.

Figure 7- Clipped Waycross to Nashville Traffic



The Blocking Opportunities Process

Most QBase reports are inquiry based, waiting for the user to initiate a transaction. As useful as this can be, we also wanted the QBase system to bring *significant* opportunities to a user's attention without having to be inquired against.

The Blocking Opportunities process is an exhaustive search through the entire QBase file, looking for opportunities to build new blocks, bypassing car-hours and handlings in intermediate yards, subject to maintaining a minimum block volume threshold.

- **Build Blocking Combinations**

First, the program identifies all locations where cars are currently handled, and builds *all possible combinations* of handling locations for every car. For example, a car which actually moved as follows:

Cincinnati, OH -> Louisville, KY -> Nashville, TN -> Waycross, GA

would generate the following six potential blocking combinations:

Cincinnati, OH -> Louisville, KY
Cincinnati, OH -> Nashville, TN
Cincinnati, OH -> Waycross, GA
Louisville, KY -> Nashville, TN
Louisville, KY -> Waycross, GA
Nashville, TN -> Waycross, GA

These combinations are built for every car in the file, and then the results are summed up across the whole file by from-to pair. *Combinations are only built for locations where the car was classified or block-swapped, according to the handling code, and where the car also spent a minimum of 10 hours.*

- **Identify Initial Opportunities**

An initial listing of opportunities is extracted from the combinations file based on the following criteria:

- The proposed block must attain a minimum volume of 15 cars/day

- At least 24 car hours per day intermediate yard time must be saved by making the proposed block. If a block is being made *some*, but not *all* the time, consolidating the remaining cars onto this existing block will still be identified as an opportunity.

- The block must "fit" reasonably with the current train operating plan. Any blocks which require more than two trains to move to destination are discarded.

- **Trial Masking Process**

In the Trial Masking phase, a "What If" analysis is performed. This step eliminates double-counted handlings and block volumes. This is accomplished by employing two conflict resolution rules when more than one blocking opportunity might apply:

1. *The first yard to receive the car classifies it.*

2. *Make the longest block possible.*

This process, also called "Block Simulation," projects how cars would actually move if all the proposed new blocks were built at the same time. Two examples of how the process works are shown in Figure 8 on the next page.

On each QBase record, the generic retriever labels or "masks" any handlings to be bypassed by overlaying a pound sign "#" into the handling code field.

Figure 8- Operation of Conflict Resolution Rules

Sample Car Movement Record:

Cincinnati, OH -> Louisville, KY -> Nashville, TN -> Waycross, GA

Example 1:

Blocks to be Made:

Cincinnati, OH -> Nashville, TN
- and -
Louisville, KY -> Waycross, GA

This car would be counted as a Cincinnati -> Nashville car, since Cincinnati received the car first. Louisville would be bypassed.

Example 2:

Blocks to be Made:

Cincinnati, OH -> Nashville, TN
- and -
Cincinnati, OH -> Waycross, GA

This car would be counted as a Cincinnati -> Waycross car, since that is a longer distance block than Cincinnati -> Nashville.

- **Identify Final Opportunities**

Based on the results of the Trial Mask, a "final cut" is made on candidate opportunities. After eliminating double-counts, this step verifies that all simulated block volumes *still meet* the required minimum thresholds. A two-step threshold is used:

- Blocks having 25 cars or more are permitted one blockswap to reach destination.

- Blocks having 15-25 cars are retained *only* if they can move from origin to destination on an existing train, without requiring a blockswap.

- Blocking opportunities having less than 15 cars per day are discarded.

- **Final Masking Process**

In this stage, the surviving opportunities are simulated once again to estimate block volumes and time savings for the final reports. The results of this step can be input into the SPM interface for precise operational, cost savings and service reliability impact evaluations.

A sample handling pattern report, automatically generated by the Blocking Opportunities system for the Chicago to Detroit block is shown in Figure 9.

- While most cars moved without intermediate handling, 70 cars from R326 were delayed in Grand Rapids, MI, averaging in excess of 26 hours. They then moved to Detroit on R326 the following day.

- An additional 11 cars arrived from Chicago on R336 connecting to R326, averaging 16 hours delay.

The handling pattern report is designed to give sufficient detail to both diagnose and correct service reliability problems. QBase inquiry-based reports can provide additional detail, right down to specific car initials, numbers and movement dates, if necessary.

```
1 PAGE      3

               FIGURE 9 - CAR ROUTINGS WITH FREQUENCY USED
                  FOR FEB92 BLOCKING OPPORTUNITIES REPORT
               OD PAIR: 100 CHICAGO  IL      240 DETROIT  MI

MOVEMENT
TYPE
----
     TOTAL  # HANDLINGS  AVG HRS
     CARS   BYPASSED     PER HNDLG  ROUTE
     ----   -----------  ---------  --------

MASKED
        70           70     26.38  CHICAGO   IL R326 GRARAPIDSMI R326 DETROIT   MI
        12           12      2.06  CHICAGO   IL R326 PLYMOUTH MI D999 DETROIT  MI
        11           11     15.92  CHICAGO   IL R336 GRARAPIDSMI R326 DETROIT   MI
         5           10      4.66  CHICAGO   IL R326 GRARAPIDSMI  FIL PLYMOUTH MI R326 DETROIT  MI
         4            4     13.80  CHICAGO   IL L336 GRARAPIDSMI R326 DETROIT  MI
         3            6    299.03  CHICAGO   IL R326 GRARAPIDSMI R336 WIXOM    MI D778 OAK      MI
         3            3       .64  CHICAGO   IL R326 PLYMOUTH MI R326 DETROIT  MI
         2            4     98.82  CHICAGO   IL R326 GRARAPIDSMI R326 PLYMOUTH MI D771 LIVONIA  MI
         2            2     52.12  CHICAGO   IL R326 PLYMOUTH MI D750 DETROIT  MI
         2            2     47.53  CHICAGO   IL R326 PLYMOUTH MI D771 LIVONIA  MI
         2            2     48.47  EASCHICAGIN R326 GRARAPIDSMI R326 DETROIT  MI
         1            2     30.96  CHICAGO   IL R326 GRARAPIDSMI R326 PLYMOUTH MI D763 LIVONIA  MI
         1            1    153.17  CHICAGO   IL R326 PLYMOUTH MI D772 LIVONIA  MI
         1            2     19.48  CHICAGO   IL R336 GRARAPIDSMI R326 PLYMOUTH MI D771 LIVONIA  MI
         1            1    171.58  CHICAGO   IL R382 WILLARD  OH R217 DETROIT  MI
         1            2     34.50  CHICAGO   IL R501 CINCINNATOH R514 PLYMOUTH MI D772 LIVONIA  MI

*TOTAL FIXIT MASKED
       121          134     38.29

UNMASKED
       430            0       .00  CHICAGO   IL R326 DETROIT  MI
        16            0       .00  CHIFORSTRIL R326 DETROIT  MI
         5            0       .00  CHICAGO   IL R336 GRARAPIDSMI R326 DETROIT  MI
         1            0       .00  CHICAGO   IL Q509 WALBRIDGEOH R390 DETROIT  MI
         1            0       .00  CHICAGO   IL R326 PLYMOUTH MI D772 LIVONIA  MI
         1            0       .00  CHICAGO   IL R336 GRARAPIDSMI R326 PLYMOUTH MI D772 LIVONIA  MI
         1            0       .00  CHICLEARIIL R326 DETROIT  MI
         1            0       .00  EASCHICAGIN R326 DETROIT  MI

*TOTAL FIXIT UNMASKED
       456            0       .00
*TOTAL ODPAIR 100 CHICAGO  IL      240 DETROIT  MI
       577          134     38.29
```

Implementation Experience

As of May, 1992, the QBase file is transitioning from the proof of concept stage into a widely used corporate database. With direct support from Operations Research, QBase has thus far been utilized on two major projects: a merchandise reroute, leading to the closing of a merchandise hump yard; and in ongoing plant rationalization studies including capacity sharing opportunities with other rail carriers. In addition, it has been utilized on a wide variety of smaller, adhoc projects.

Currently, only a few "power users" outside the Operations Research group are able to utilize the QBase file. During July 1992, we hope to offer training classes for a broad base of users in Service Design, Marketing, Cost & Economic Analysis and Cost & Budgets.

The Service Planning Model interface runs on a production basis each month to create a new model base case. Recent SPM projects have included an evaluation of the impact of proposed network train consolidations on operating costs, transit time and reliability. As of this writing, this SPM study is still ongoing.

An initial run of the Blocking Opportunities process has identified a *potential* car-hour savings equivalent to more than 3,000 freight cars. Even if only a small fraction of this potential can be realized, the savings to CSXT will still be substantial.

Most opportunities fall into the "better discipline" category (to more effectively build blocks which are already called for in the operating plan, or to adjust connections and train schedules to reduce intermediate yard delay). The blocking opportunities report serves primarily as a monthly "Plan Integrity" report, with an added bonus that it will also identify new opportunities which are not in the current plan.

The implementation status of QBase is clearly work-in-progress. We hope to be able to bring conference attendees up-to-date on our implementation experience at the TRF Saint Louis meeting in October, 1992.

Endnotes and References

1. Van Dyke, C., "The Automated Blocking Model: A Practical Approach to Freight Railroad Blocking Plan Development." *Proceedings, Transportation Research Forum,* Vol XXVII, No. 1, pp 116-121, 1986.

2. Van Dyke, C. and L. Davis, "Software Tools for Railway Operations/ Service Planning: the Service Planning Model Family of Software," T.K.S. Murthy, et al. (eds.), *Computer Applications on Railway Planning and Management* (1990).

3. J. Reilly McCarren and Carl D. Martland, "The MIT Service Planning Model," *Studies in Railroad Operations and Economics,* Vol 31, Cambridge, Ma.: MIT, Center for Transportation Studies, Report No. CTS 79- ,1979.

4. Griffin, Mary B, "The Princeton Transportation Network Model and Graphic Information System," *ALK Associates, Princeton, NJ,* 1992.

5. Our is the largest known application of the SPM, the network definition consisting of 2,000 trains, 3,000 blocks, 40 traffic classes and 12,000 OD pairs. Previously, SPM had been successfully built and maintained using manual methods at CSX Intermodal, a network 1/10 the size. However, the simple truth was that we did not have in place, at that time, a mechanism for collecting, formatting, and keeping up-to-date the sheer volume of data required to make the SPM effective for the entire CSX rail network.

6. Actually, this file is named after a similar data file with which the author worked as a student intern at ConRail in 1979.

7. The QBase flow mapping process includes a flow rerouting capability and a network editor function. Foreign line links can be added, CSX links deleted and network impedances modified. This provides a limited "what if" capability, to estimate car-miles and calculate flow densities using the PTNM network.

8. We have adapted SPM input data to the needs of a grain network planning model under development at University of Virginia.

9. Lawrence, Martha B and Shughart, Larry A, "A Railroad Terminal Cost Model to Support Network Management Strategy," *TRF Proceedings,* October 1992.