FAF2 Pilot Project – Utilization of FAF2 Data by State and Local Governmental Agencies

Final Report for:
Alabama Department of Transportation
Research Project 930-682

Project Advisory Committee
Mr. Bob Jilla – Chair
Mr. Don Vaughn
Dr. Emmanuel Oranika
Dr. Tianjia Tang – FHWA
Ms. Ivy Harris – RAC Liaison
Mr. Clint Andrews – FHWA Liaison

February 19, 2009

Principal Investigator:
Gregory A. Harris, Ph.D., P.E.

Co-Investigator:
Michael D. Anderson, Ph.D., P.E.

FAF2 Pilot Project – Utilization of FAF2 Data by State and Local Governmental Agencies

TABLE OF CONTENTS

Executive Summary………………………………………………………….. 4
1. Introduction…………………………………………………………….... 5
2. Disaggregation of the Freight Analysis Framework Version 2 Database... 6
2.1 Experiment Design…………………………………………………….. 9
2.2 Modeling Tools………………………………………………………… 11
2.3 Analysis……………………………………………………………….... 11
2.3.1 *Setting up the Experiment.…………………………………………..* 12
2.3.2 *10 tons per Vehicle Analysis.……………………………………….* 15
2.4 Results………………………………………………………………… 16
3. Application of Disaggregated FAF2 Data in Alabama……………………. 17
3.1 Utilizing FAF2 Data to Analyze the Addition of a New Facility............... 17
3.2 Forecasting Freight Traffic Using FAF2 Data…………………………… 20
3.2.1 *Model Validation.…………………………………………………….* 20
4. Conclusions……………………………………………………………… 31

TABLE OF TABLES

1. Listing of Commodities from FAF2 Database……………………………………… 7
2. Listing of Transportation Modes from FAF2 Database…………………………….. 8
3. Set of Runs Containing Various Factor Levels……………………………………… 13
4. Run Results for the Tons/Vehicle Data…………………………………………... 16
5. Total Annual Tons of Freight Passing Through Alabama…………………………… 19
6. Difference in Annual Tons for each Alabama Interstate……………………………. 19

TABLE OF FIGURES

1. Geographic Locations for FAF2 data……………………………………………… 5
2. Experimental Procedure………………………………………………………… 9
3. Modeling Network……………………………………………………………… 11
4. Nash Sutcliffe Values for Various Tonnages……………………………………… 15
5. Location of Interstate 22………………………………………………………... 18
6. Nationwide Travel Demand Model Network……………………………………… 18
7. Freight Flow without Interstate 22………………………………………………… 19
8. Freight Flow with Interstate 22…………………………………………………… 20
9. Current Congestion Locations…………………………………………………….. 22
10. Congested Locations Using Trend Line Analysis………………………………… 23
11. Congested Locations Using the FAF2 2015 Projection…………………………… 24
12. Locations Where Travel Time Exceeds 25% Using the FAF2 2015 Projection….. 24
13. Congested Locations Doubling the FAF2 2002 Volumes………………………… 25
14. Locations Where Travel Time Exceeds 25% When Truck Traffic is Doubled from FAF2 2002 Volumes………………………………………………………………. 26
15. Locations Where Travel Time Exceeds 100% When FAF2 2002 Volumes are Doubled…………………………………………………………………………. 27
16. Congested Locations When Truck Traffic is Doubled in Selected Counties……… 28
17. Locations Where Travel Time Exceeds 25% When Truck Traffic is Doubled for Selected Counties……………………………………………………………….. 28
18. Congested Locations Using the 2035 FAF2 Projection…………………………… 29
19. Locations Where Travel Time Exceeds 25% Using the 2035 FAF2 Projection…... 30
20. Locations Where Travel Time Exceeds 100% Using the 2035 FAF2 Projection…. 30
21. UAHuntsville Freight Planning Framework………………………………………. 32

FAF2 Pilot Project – Utilization of FAF2 Data by State and Local Governmental Agencies

Executive Summary

It is difficult to incorporate freight information into transportation models and plans because freight data is proprietary and the release of that data is considered to be detrimental to the company's competitive position. Due to the difficulty in acquiring freight data, the inclusion of freight in most transportation plans and models has either been limited in scope or based upon limited sample sizes without knowledge of contents. In the United States, many national freight databases aggregate information to the individual states, or major communities in the states. For example, the Freight Analysis Framework, Version 2 Database (FAF2) developed and distributed by the Federal Highway Administration (FHWA) contains freight flows for 114 zones at the national level.

The use of national freight data at the local level is challenging due to the high level of aggregation. In most instances the disaggregation of freight data from national levels for use in local areas has been based on the factor "employment" by prorating the employment in the local area to the total employment in the study region. The use of employment as a planning factor has come under scrutiny as this factor does not accurately estimate the effect of productivity improvements implemented to increase production without increasing the number of employees.

The purpose of this research, in support of the Federal Highway Administration FAF2 Pilot Project, was to develop methods and procedures to enable State and local governmental agencies to utilize the Freight Analysis Framework 2 (FAF2) commodity origin-destination data and the FAF2 freight network during the transportation planning process. The University of Alabama in Huntsville (UAHuntsville) has fairly quickly developed a reputation for innovative, forward-looking ideas, theories and methods in the area of freight forecasting and systems modeling. Because of the work previously performed and ongoing research, the opportunity to participate in this national level pilot study emerged. This opportunity has allowed Alabama and UAH to establish a lead position in setting the research direction of the freight forecasting and planning efforts of the nation.

During the course of this research a methodology for disaggregating national level data to the state and local levels was developed, tested and applied under several scenarios. The results of the development of this disaggregation method has led to expanded use of the FAF2 database for research into freight related activities at ports, in metropolitan planning organizations, and "what if" scenarios from the addition of facilities to the interstate system to the forecasting of industry growth patterns and the affect on the congestion of a facility.

This research has been successful in accomplishing the goals originally established and has progressed farther into the areas of application, and now, refinement of the methods developed.

1. Introduction

Freight transportation is vital to the growth and economic development of a region or state. Local transportation professionals often have problems incorporating freight into transportation plans and models because freight data is proprietary at local levels and requires extensive aggregation to national levels before being released to the public. Understanding freight activity and the factors affecting freight activity are important for planning infrastructure supply to transport demand and for assessing potential investment and operational strategies. In the United States, many national freight databases aggregate information to the individual states, or major communities in the states. For example, the Freight Analysis Framework, Version 2 Database (FAF2) developed and distributed by the Federal Highway Administration (FHWA) contains freight flows for 114 zones at the national level, as shown in Figure 1. A benefit of using the FAF2 database for analysis of freight is that the database includes freight flow data for base-year, 2002, as well as forecasts for 2010 through 2035 in 5-year increments (FAF2 2007)[1].

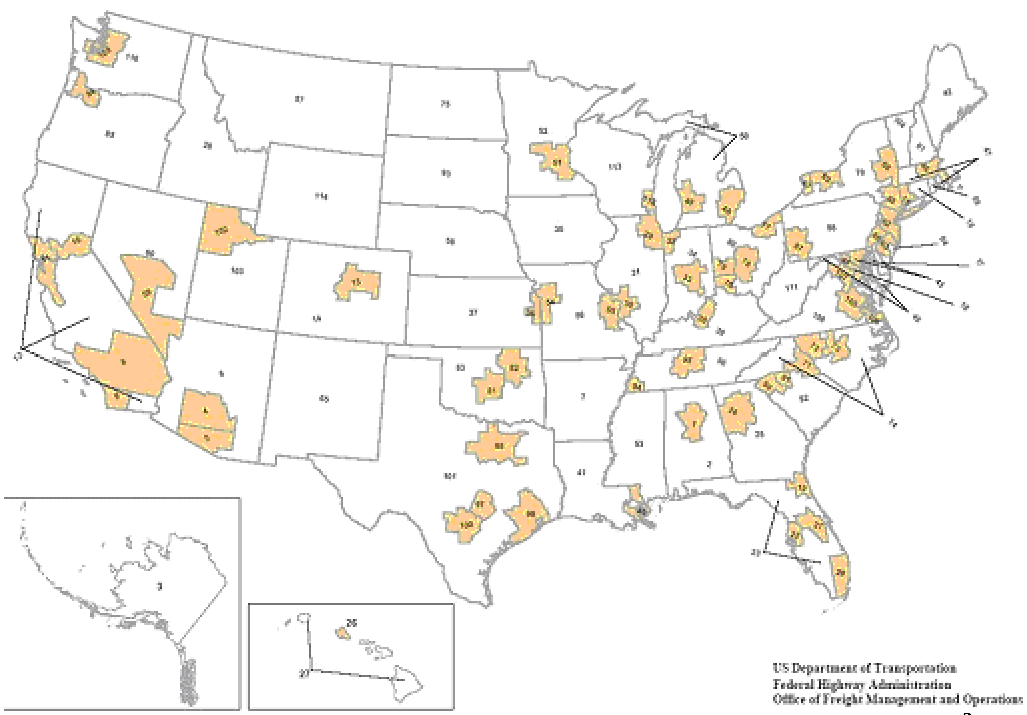

Figure 1. Geographic Locations for FAF2 data (FAF2 Areas 2007)[2]

But, the application of freight data to the local level is challenging due to the aggregation of the data. The disaggregation freight from national levels for use in local areas has been based on the relative employment in the local area to the total employment in the zone. This disaggregation technique has limitations in that productivity improvements which allow manufacturers to produce more products that require more freight shipments using fewer employees are often undetected (UAH 2005)[3]. This report presents research into using a national freight origin/destination database and various socio-economic factors to perform disaggregation to the local level. The factors considered in this research include population, employment, personal income, and value of shipments.

A case study is used to address the disaggregation for Alabama from 2 zones at the national level into 67 individual counties at the state level. The case study uses a CUBE/TRANPLAN to model disaggregated freight data on a statewide network with a variety of weighting factors placed on the four socio-economic data elements.

2. Disaggregation of the Freight Analysis Framework Version 2 Database

The FAF2 Commodity Origin-Destination Database contains estimates of tonnage and value of goods shipped by type of commodity (see Table 1) and mode of transportation (see Table 2) for 114 FAF2 zones (shown in Figure 1), 7 international trading regions, and 17 additional international gateways. The 2002 estimate is based on the Commodity Flow Survey (CFS) with some of the data voids in the CFS filled in by analysis of the Economic Census and additional data sources. Forecasts are included for 2010 to 2035 in 5-year increments. The FAF2 database contains origin and destination values for tonnage and value of shipment identified for six unique transport modes and 42 individual commodities identified using the Standard Classification for Transported Goods (SCTG).

The FAF2 database is a continuation of the original Freight Analysis Framework developed by the U.S. Department of Transportation, FHWA. Whereas the original FAF provided the public with generalized freight movement and highway congestion maps without disclosing the underlying data, FAF2 provides commodity flow origin-destination (O-D) data and freight movement data on all highways within the FAF2 highway network. To maintain confidentiality of the data providers, the FAF2 data is extensively aggregated within each state. For example, freight data for Alabama is categorized as AL BHM, containing an eight county region around Birmingham, AL, and AL REM, containing the remaining 59 counties in Alabama. At such a high level of aggregation it is often difficult to incorporate FAF2 data into transportation models and plans at the state and local levels.

As stated previously, there are two identified zones for Alabama in the FAF2 database. The disaggregation of this data is not merely a distillation of data; UAH researchers developed a repeatable process of defining the data into nine unique trip purposes.

- Internal-Internal for Zone 1 and Zone 2. The internal trips for the individual zones are defined as the total trips that are both produced and attracted in the zone of interest. These trips are disaggregated into production and attraction values for the individual zones using the socio-economic factors.
    - Internal to Zone 1
    - Internal to Zone 2

Table 1 - Listing of Commodities from FAF2 Database

BTS/Census Full Commodity Name	**FAF Abbreviation**
Live animals and live fish	Live animals/fish
Cereal grains	Cereal grains
Other agricultural products	Other ag prods.
Animal feed and products of animal origin, n.e.c.1	Animal feed
Meat, fish, seafood, and their preparations	Meat/seafood
Milled grain products and preparations, bakery products	Milled grain prods.
Other prepared foodstuffs and fats and oils	Other foodstuffs
Alcoholic beverages	Alcoholic beverages
Tobacco products	Tobacco prods.
Monumental or building stone	Building stone
Natural sands	Natural sands
Gravel and crushed stone	Gravel
Nonmetallic minerals n.e.c.1	Nonmetallic minerals
Metallic ores and concentrates	Metallic ores
Coal	Coal
Crude Petroleum	Crude petroleum
Gasoline and aviation turbine fuel	Gasoline
Fuel oils	Fuel oils
Coal & petroleum products, n.e.c.1 (Note: primarily natural gas, sel. coal products, & products of petroleum refining, excluding gasoline, aviation fuel, and fuel oil.)	Coal-n.e.c.1
Basic chemicals	Basic chemicals
Pharmaceutical products	Pharmaceuticals
Fertilizers	Fertilizers
Chemical products and preparations, n.e.c.1	Chemical prods.
Plastics and rubber	Plastics/rubber
Logs and other wood in the rough	Logs
Wood products	Wood prods.
Pulp, newsprint, paper, and paperboard	Newsprint/paper
Paper or paperboard articles	Paper articles
Printed products	Printed prods.
Textiles, leather, and articles of textiles or leather	Textiles/leather
Nonmetallic mineral products	Nonmetal min. prods.
Base metal in primary or semi-finished forms and in finished basic shapes	Base metals
Articles of base metal	Articles-base metal
Machinery	Machinery
Electronic and other electrical equipment and components and office equipment	Electronics
Motorized and other vehicles (including parts)	Motorized vehicles
Transportation equipment, n.e.c.1	Transport equip.
Precision instruments and apparatus	Precision instruments
Furniture, mattresses and mattress supports, lamps, lighting fittings	Furniture
Miscellaneous manufactured products	Misc. mfg. prods.
Waste and scrap	Waste/scrap
Mixed freight	Mixed freight
Commodity unknown	Unknown

Table 2 - Listing of Transportation Modes from FAF2 Database

Truck. Includes private and for-hire truck. Private trucks are operated by a temporary or permanent employee of an establishment or the buyer/receiver of the shipment. For-hire trucks carry freight for a fee collected from the shipper, recipient of the shipment, or an arranger of the transportation.
Rail. Any common carrier or private railroad.
Water. Includes shallow draft, deep draft and Great Lakes shipments. FAF2 uses definitions by the U.S. Army Corps of Engineers. Shallow draft includes barges, ships, or ferries operating primarily on rivers and canals; in harbors; the Saint Lawrence Seaway; the Intra-coastal Waterway; the Inside Passage to Alaska; major bays and inlets; or in the ocean close to the shoreline. Deep draft includes barges, ships, or ferries operating primarily in the open ocean.
Air (includes truck-air). Includes shipments by air or a combination of truck and air. Commercial or private aircraft and all air service for shipments that typically weigh more than 100 pounds. Includes air freight and air express.
Truck-Rail Intermodal. Includes shipments by a combination of truck and rail.
Other Multiple Modes. Includes shipments typically weighing less than 100 pounds by Parcel, U.S. Postal Service, or Courier, as well as shipments of all sizes by truck-water, water-rail, and other intermodal combinations.
Pipeline and Unknown. Pipeline is included with unknown because region-to-region flows by pipeline are subject to large uncertainty.

- Values exchanged between Zone 1 and Zone 2. The freight values produced in one Alabama zone and attracted to the other Alabama zone are handled by applying the disaggregation factors to both the counties as a function of the total trips produced or attracted.
    - From Zone 1 to Zone 2
    - From Zone 2 to Zone 1
- Values exchanged between Alabama and the Remainder of the Country. The freight values are disaggregated through the use of the socio-economic factors for Alabama counties.
    - From Zone 1 to locations outside Alabama
    - From Zone 2 to locations outside Alabama
    - From outside Alabama to Zone 1
    - From outside Alabama to Zone 2
- Alabama pass through. Encompasses those freight values that neither originate nor terminate in Alabama, but travel on Alabama roadways because the preferred route is through Alabama. These trips are defined using the following relationship:

$$FAF2(ee) = [\ FAF2 - FAF2\ (origin\ AL) - FAF2\ (to\ AL) - FAF2\ (not\ AL)] \quad (1)$$

Where: FAF2(ee) = pass through on Alabama Roadways
FAF2 = entire database
FAF2 (origin AL) = values originating in Alabama
FAF2 (to AL) = values terminating in Alabama
FAF (not AL) = values that do not travel through Alabama.

- Pass through values

2.1 Experiment Design

The purpose for disaggregation is to develop data to use in transportation modeling that better represents existing conditions and provides a more accurate forecast. To accomplish this, a testing framework was developed utilizing TRANPLAN to develop truck counts. For a better understanding, the procedure is displayed in Figure 2.

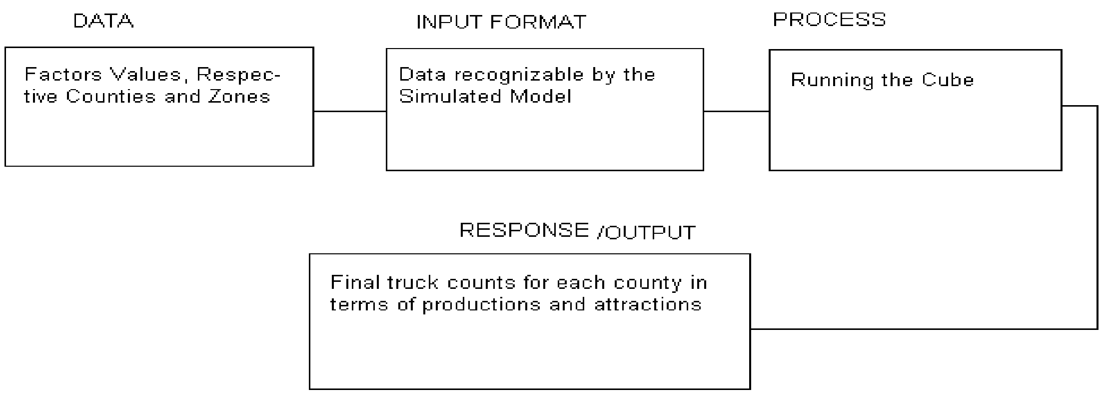

Figure 2. Experimental Procedure

As listed in Figure 2, the methodology for this research consists of three major tasks, generating the input (INPUT), running the modeling software (PROCESS) and analyzing the output (OUTPUT). Generating the input was thought to be the most crucial step in this process, and is discussed below in detail followed by the succeeding two tasks.

The number of freight carrying trucks visiting each county, $PA_{i,}$ is the input to the model. Productions are referred as the number of trucks going out and attractions are the number of trucks coming to each county. These productions and attractions are a function of the factors used to disaggregate freight traffic. Zonal truck counts in AL and the freight factor amounts for each county by population, employment, personal income and value of shipment are available. Based on the data available, Equation (2) was used to disaggregate the zonal truck counts to county level.

$$PA_i = (NFD_{ab}) * \frac{(WF) * Factor_i}{\sum Factor_{ij}} \quad (2)$$

Where,

PA_i = Truck passing County i
NFD_{ab} = Truck Counts from Zone-a to Zone-b taken from the National Freight Flow
WF = Weight of the factor (or) importance of the factor (or) proportion of the factor considered for disaggregating
$Factor_i$ = Factor level for county i
$\sum Factor_{ij}$ = Total Factor level for the corresponding Zone of county i
i= county number (1, 2, 3, 4……67)
j= Zone number (1, 2)

When the factors population, personal income, employment, and value of shipment were substituted in place of 'Factor$_i$' in Equation (2), it becomes:

$$PA_i = (NFD) * \left[\frac{W_1 * P_i}{\sum P_j} + \frac{W_2 * PI_i}{\sum I_j} + \frac{W_3 * E_i}{\sum E_j} + \frac{W_4 * VOS_i}{\sum VOS_j} \right] \quad (3)$$

P = population, PI = Personal Income, E = Employment, VOS = Value of shipment

W_1, W_2, W_3, W_4 are the weights or contribution levels of population, personal income, employment and value of shipment respectively in calculating the county level truck counts (input). The amount of each factor used for disaggregating the National Freight Flow data is given by the above weights. For example, if the contribution of each factor is considered to be the same in calculation of truck counts, then $W_1=W_2= W_3=W_4=0.25$. This equation is used to disaggregate or distribute the zonal truck counts from the National Freight data to the county level. Therefore, the total number of trucks before and after disaggregating must be equal, shown in Equation (4).

$$\sum PA_i = \sum NFD_{ab} \quad (4)$$

For satisfying criterion, there are two constraints in the equation involving W_1, W_2, W_3 and W_4

1. $$\sum_{i=1}^{4} W_i = 1 \quad (5)$$

2. $$W_i = Range(0,1) \quad (6)$$

These levels sum to 1 because if $\sum W_i>1$, the total number of modeled trucks would exceed the total actual trucks. By assigning a number within the range (0, 1) to these weights, we are actually choosing the contribution level or the importance of each factor in generating the input, which is then entered into the modeling software.

The next task after generating the input (disaggregated zonal truck counts) is to enter the data into the modeling software and extract the output. The output is the freight truck traffic generated on Alabama roadways. This is displayed in the form of an excel file containing various roadways numbered from 1 to 383, the total number of roadways modeled. The assignment of the forecasted truck counts for each roadway is contingent to the input PA_i entered.

One way to measure the impact on output of this model is to measure the deviation or difference of each data point with respect to the actual counts. The yard stick in this case is the actual truck traffic in the Alabama network provided by the Alabama Department of Transportation (ALDOT). The closer the actual counts are to the modeled values, the better is our forecast and thus the factor contributions. Minimizing the difference between actual counts and modeled values can be achieved by varying the factor contribution in disaggregating the zonal truck counts. By this analysis we can deduce a combination of factor contributions that aid us in forecasting the truck counts.

2.2 Modeling Tools

A travel demand model network was developed in CUBE/TRANPLAN and used to assign the trips obtained from the FAF2 database. The model contains all Interstates, U.S. Highways and many Alabama Highways totaling nearly 5,000 miles of roadway in the state. The roadways are attributed with posted speed limits and capacities, using approved ALDOT capacities for travel modeling purposes (Figure 3). The model contains 67 internal zones, representing each county in Alabama and has 15 external roadways connecting Alabama with the remainder of the nation. A gravity distribution model has been incorporated to distribute the trips between the counties using the nine trip purposed previously described. The assignment is performed using an all-or-nothing assignment as the assumption is made that freight will not deviate from the shortest path because there is not necessarily knowledge regarding shortest path alternative when assigning trips for potential out-of-town shippers.

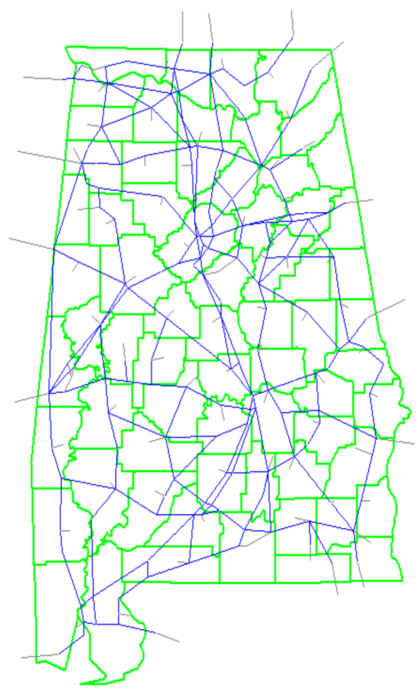

Figure 3. Modeling Network.

2.3 Analysis

Three metrics were identified that could give a measure of accuracy of the forecast; Root Mean Square Error, the Nash-Sutcliff Coefficient, and Percent Error.

Root Mean Square Error (RMSE) is a common measure of the variability in the error (difference between model and actual counts) of a model. The greater the RSME, less accurate is the model.

$$RMSE = \frac{\sqrt{\sum (Model_i - Ground_i)^2 / (NumofCounts - 1)} * (100)}{\sum Ground_i / NumofCounts}, \text{ (Monsere 2001)} \quad (7)$$

Where:
RMSE = root mean square error
$Model_i$ = Modeled Value for the roadway i
$Ground_i$ = Actual Counts for the roadway i.

The second measure used in this analysis was the Nash Sutcliffe (NS) coefficient which can range from -∞ to 1. An efficiency of 1 (E=1) corresponds to a perfect match of forecasted counts to the actual counts. An efficiency of 0 (E=0) indicates that the forecasted values are as accurate as the mean of the actual counts, whereas an efficiency less than zero ($-\infty < E < 0$) occurs when the forecasted mean is less than the actual values. This coefficient gives a measure of scatter variation from the 1:1 slope line of modeled truck counts vs the actual data. The greater the NS-value the better the forecast. It can be calculated using equation (8):

$$\text{NS-Coefficient} = 1 - \frac{\sum_1^n (ModeledCounts - GroundCounts)^2}{\sum_1^n (GroundCounts - MeanGoundCounts)^2} \text{ (Monsere 2001)[4]} \quad (8)$$

The Nash Sutcliffe statistic is considered the best measure of deviation between two data sets and used in many similar instances.

The third measure used was the percent error between the forecasted and the actual counts. It has given the percentage of difference between both the data sets (equation 9).

$$Percenterror = \left(\frac{Model(i) - Ground(i)}{Ground(i)}\right)(100)/N \text{ (Monsere 2001)[4]} \quad (9)$$

Where,
$Model_i$ = Modeled Value for the roadway i
$Ground_i$ = Actual Counts for the roadway i
N = Total number of modeled values.

2.3.1 Setting up the Experiment

The purpose of this research is to determine relevant factors for disaggregating the zonal truck values, testing various combinations of the contribution of each factor or the importance of each factor. Table 3 displays the results for which all the three metrics have been calculated.

An experiment in which the response is assumed to depend only on the relative proportions of the factors present is a mixture experiment. In a mixture experiment, the total amount of mixture is held constant and the value of the response changes when changes are made in the relative proportions of those ingredients making up the mixture (Cornell 1990)[5]. The total number of trucks used for disaggregating the zonal values is always constant and the ingredients to make up this constant value are the factor contribution levels in this case.

Below is a set of model runs containing a combination of factor proportions (W_i) for each run in the experimental design which was generated from Minitab® 14.0 under the Mixture Experiments option. The column under each factor represents the contribution level of each factor in disaggregating the zonal truck counts. Since all the weights must sum to one, the run totals are equal to one and no single factor exceeds this value. Note: The values under each column of P, PI, E, VOS are the corresponding Weights (W_i): (P=population, PI=personal income, E = employment and VOS = value of shipment)

Table 3. Set of Runs Containing Various Factor Levels

RUN	P	PI	E	VOS	NS-Value	RMSE	%Error
1	1	0	0	0	0.195821	105.92	86.44
2	0	1	0	0	0.197551	105.8	87.97
3	0	0	1	0	0.195821	105.92	86.36
4	0	0	0	1	0.193561	105.92	86.44
5	0.5	0.5	0	0	0.195821	105.85	86.39
6	0.5	0	0.5	0	0.196825	105.84	86.31
7	0.5	0	0	0.5	0.196985	105.86	86.26
8	0	0.5	0.5	0	0.196642	105.8	87.97
9	0	0.5	0	0.5	0.197551	105.82	87.12
10	0	0	0.5	0.5	0.197239	105.87	86.61
11	0.33333	0.33333	0.33333	0	0.1965	105.85	86.04
12	0.33333	0.33333	0	0.33333	0.196835	105.92	86.36
13	0.33333	0	0.33333	0.33333	0.195821	105.91	86.42
14	0	0.33333	0.33333	0.33333	0.195952	105.9	86.11
15	0.25	0.25	0.25	0.25	0.19606	105.89	86.18
16	0.625	0.125	0.125	0.125	0.196219	106.07	85.57
17	0.125	0.625	0.125	0.125	0.193561	106	85.67
18	0.125	0.125	0.625	0.125	0.194573	105.83	85.96
19	0.125	0.125	0.125	0.625	0.197182	105.82	86.53

The first four runs gave held each of the four factors at highest importance. The next six observations took combinations of two out of four in each run, giving them a weight of zero and 0.5 to the remaining variables. Four trials were also allotted to a scheme, giving a weight of zero to each variable in turn and equal weights to the remaining four. A run was also apportioned equal weight to all the four input variables.

The three metrics were calculated for each run after it was executed from the model. Using the Nash-Sutcliffe coefficient, the immediate step would be to trace a combination of factor contributions that have a positive effect over the response.

An F-test was employed to test the hypothesis and Minitab® yielded a p-value of 0.279 which is greater than the 5% significance level. The null hypothesis could not be rejected in this case according to 5% significance. As a result, no matter whatever the importance level we assign to each of the factors, population, personal income, employment and value of shipment, it is not varying modeled truck traffic inside the Alabama network.

When the scatter plots were graphed between the modeled trucks counts versus the actual counts from the ALDOT, for all the runs, there was not much of a difference in the scatter pattern for all the runs. Even this shows that there is not much impact by varying the factor levels for generating the input. In addition, the RSME (root mean square error) and the percent error showed constant results, giving less scope for variation.

This could have been the conclusion for this investigation whether the population, personal income, employment and value of shipment of each county effected the disaggregation of zonal truck counts, and thus the modeled truck traffic. However, when observed carefully, there was not a single combination of factor contributions that yielded a NS-value close to 1, which means that regardless of which combination of factors were considered, the modeled truck traffic is not matching with the actual counts (actual freight carrying truck counts provided by ALDOT). If the modeling software was built appropriately, there must have been some point where the modeled truck counts were close to the actual counts. As a result, speculation was aroused as to whether some of the factors did really have an impact over the freight flow.

Before it could be concluded that none of the initially assumed factors influenced the freight flow in Alabama, there is one plausible argument which negates the previous inference that none of the factors help in disaggregating the zonal truck counts and the modeled truck traffic. All the above runs were executed for a freight transfer of 30 tons. The truck capacity is an attribute in the modeling software that can be manipulated. Until this point it was to disaggregate the trucks to the Alabama network which had an assumed freight carrying capacity of 30 tons. This is an attribute setting in the forecasting software where we can vary the tonnage from 0 to 30 tons with a fixed interval of 5 tons.

Since the varying of coefficients had a limited impact on the final truck counts for the 30 tons/vehicle, one combination of the coefficients was predefined and a set of runs were carried by varying the tonnage of the trucks. Below is the graph of how the Nash Sutcliffe coefficient varied with the change in tonnage. As an initial step for the second phase of analysis, the Nash-Sutcliffe coefficient was calculated for all the truck capacities for the same combination of factors levels. Figure 4 shows the variation of this coefficient with the tons/vehicle.

From the above graph, it is evident that this macro level change in the software impacted the network and a highest value was recorded for the 10 tons. When trucks with a capacity of 10 tons/vehicle were used for the modeling network, it yielded the truck counts closest to the actual truck counts (actual counts) provided by the Alabama Department of Transportation (ALDOT). A micro level analysis performed for the 30 tons/vehicle was again performed for the 10 tons/vehicle model and analyzed if disaggregating based on county level factors impacted the response.

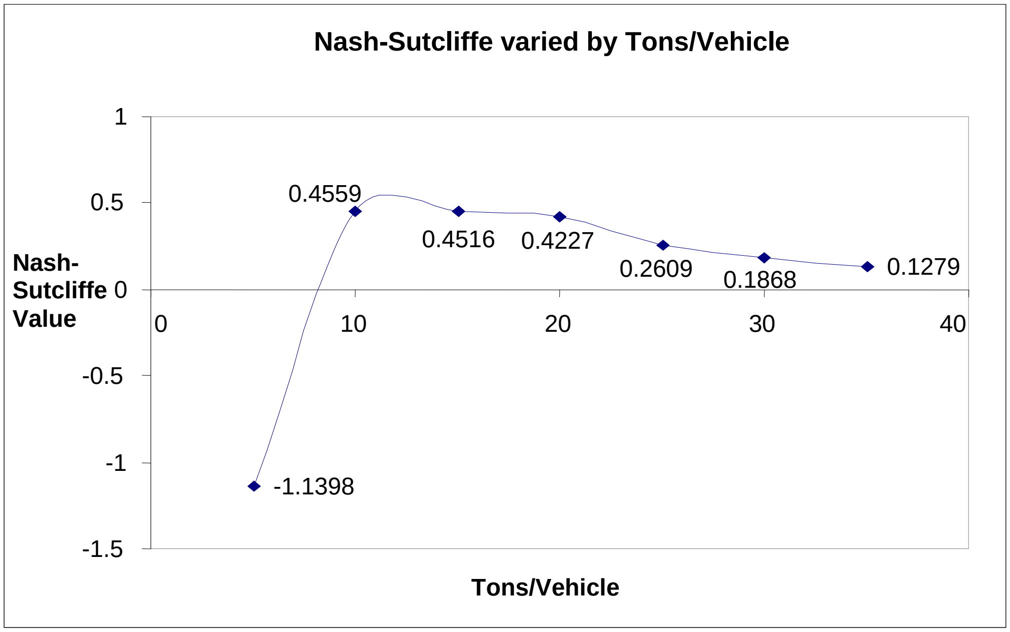

Figure 4. Nash Sutcliffe Values for Various Tonnages

2.3.2 10 tons per Vehicle Analysis

A design similar to the initial analysis was performed. A simplex centroid design with 19 runs was setup in the Minitab® 14.0 and a regression equation for this design was setup, see Table 4. From this regression equation, we can make some initial conclusions as to what factors really impact the modeled traffic flow.

The Minitab® output indicates that all the coefficients of W_i produce NS-Values that are approximately equal to 0.45, an improvement over the initial experiment, but also exemplifying that the effect of all the variables is essentially equal and choosing any one for calculating the county level truck counts would work. When considering the interactions in this experiment, magnitudes of coefficients in the regression equation have little variation, indicating that none of the interactions are significant. A hypothesis test (F-test) was conducted for this equation with the null and alternative hypotheses as:

H_0 = Response does not depend on the mixture components
H_A = Response does depend upon the mixture components

Minitab® yields a P-value 0.385 which was greater than the 5% significance indicating that the null hypothesis could not be rejected with that significance level. This indicates that changing the factor proportions did not impact the modeled truck traffic with a capacity of 10 tons/vehicle. Scatter plots were graphed between the modeled trucks counts versus the actual counts from the ALDOT for all model runs with little variation. This indicates that factor levels are not impacted by variation.

Table 4. Run Results for the 10 Tons/Vehicle Data

RUN	P	PI	E	VOS	NS-Value	RMSE
1	1	0	0	0	0.4559	77.4425
2	0	1	0	0	0.460142	77.1463
3	0	0	1	0	0.462018	75.5991
4	0	0	0	1	0.47171	74.5071
5	0.5	0.5	0	0	0.459501	77.1895
6	0.5	0	0.5	0	0.472021	76.3102
7	0.5	0	0	0.5	0.469717	75.7614
8	0	0.5	0.5	0	0.463931	76.1724
9	0	0.5	0	0.5	0.481489	75.6334
10	0	0	0.5	0.5	0.469971	74.9775
11	0.33333	0.33333	0.33333	0	0.468871	76.5275
12	0.33333	0.33333	0	0.33333	0.464233	76.1502
13	0.33333	0	0.33333	0.33333	0.468771	75.6167
14	0	0.33333	0.33333	0.33333	0.469662	75.5556
15	0.25	0.25	0.25	0.25	0.467459	75.9246
16	0.625	0.125	0.125	0.125	0.467388	76.6304
17	0.125	0.625	0.125	0.125	0.469834	76.4629
18	0.125	0.125	0.625	0.125	0.469302	75.7198
19	0.125	0.125	0.125	0.625	0.46922	75.1712

For the 30 tons/vehicle runs, the plotted data points had an increasing trend indicating that the error was higher for larger actual values. The error in the modeled values was larger for busier roadways. When the 10 tons/vehicle residual plots were plotted, there was a different scatter when compared to original runs, but the same conclusion that changing factor proportions did not matter. As a result, varying the factors in the calculation of county level truck counts did not have an impact over the modeled truck traffic in the state.

2.4 Results

This research determined what factors were the best considerations for disaggregating the national freight flow data, which can be used as input for Alabama freight flow modeling software. After the initial analysis, it was determined that individual factors considered for disaggregating the national freight flow data did not impact the modeled freight flow inside AL. An attribute regarding tonnage was changed and the Nash-Sutcliffe efficiency was calculated for different tonnages. The tonnage yielding the best Nash-Sutcliffe efficiency was performed using the same methodology applied for the runs with 30 tons/vehicle. The best Nash Sutcliffe value was recorded for the attribute 10 tons/vehicle and the experiment was performed for this case. No change was observed even in this case and the same conclusion that population, personal income, employment and value of shipment did not affect the desegregation of freight flow to the county level was concluded.

In both the cases with different tonnage capacities, the Nash Sutcliffe statistic was different but a higher value was derived for the 10 tons/vehicle runs. This indicated that the modeled values were much closer to the actual counts meaning a better batch of runs containing the 10 tons/vehicle trucks. By varying the factors and the attribute within software, the approximate maximum Nash-Sutcliffe Value was around 0.47. The highest achieved values were when Value of Shipment and Personal Income were used in the analysis, although the number of vehicles developed by the model is still short of the actual counts. These findings indicate that the research has not yet found the best freight factors to use but it would appear that the researchers are actually on the trail of some promising results.

3. Application of Disaggregated FAF2 Data in Alabama

To accomplish the application of FAF2 data, the research team built upon the existing transportation analysis and planning tools developed at UAHuntsville, the Alabama Transportation Infrastructure Model (ATIM) and a statewide highway, rail and waterway network developed in TRANPLAN that was enhanced to support statewide freight analysis. The research team developed a seamless interface between the two models to allow for easy sharing of Origin/Destination, route and volume data. The integration of these models produced a tool capable of quickly analyzing scenarios and events on the transportation infrastructure and can be used to evaluate alternative solutions.

Two applications of the disaggregated FAF2 data will be presented here. The first is the use of FAF2 data to determine the level of pass through freight traffic that can be expected due to the completion of a new Interstate highway (I-22) between Birmingham, Alabama and Memphis, Tennessee. Finally, the FAF2 data will be used to forecast a level of congestion on Alabama Interstates and major highways.

3.1 Utilizing FAF2 Data to Analyze the Addition of a New Facility

Interstate 22 will be constructed to connect the urban areas of Birmingham, AL and Memphis, TN and is expected to be open for traffic between 2010 and 2015. The interstate is intended to replace U.S. Highway 78 with a high-speed, controlled access facility. The location of the interstate can be seen in Figure 5.

To model the freight flow patterns that will be impacted by the construction of Interstate 22, a national-level travel demand model was created using the FAF2 zone descriptions and existing interstate facilities. The 114 zones were selected as the traffic analysis zones for the model. The existing interstate infrastructure was used to represent the transportation network in the travel model. The roadway segments in the model were attributed with distance (using a scale factor because of the large distances) and the speeds for vehicles on the interstates. The nationwide network used in the process can be seen in Figure 6. An alternative network was developed for the purpose of this study that included a link for Interstate 22 from Birmingham to Memphis.

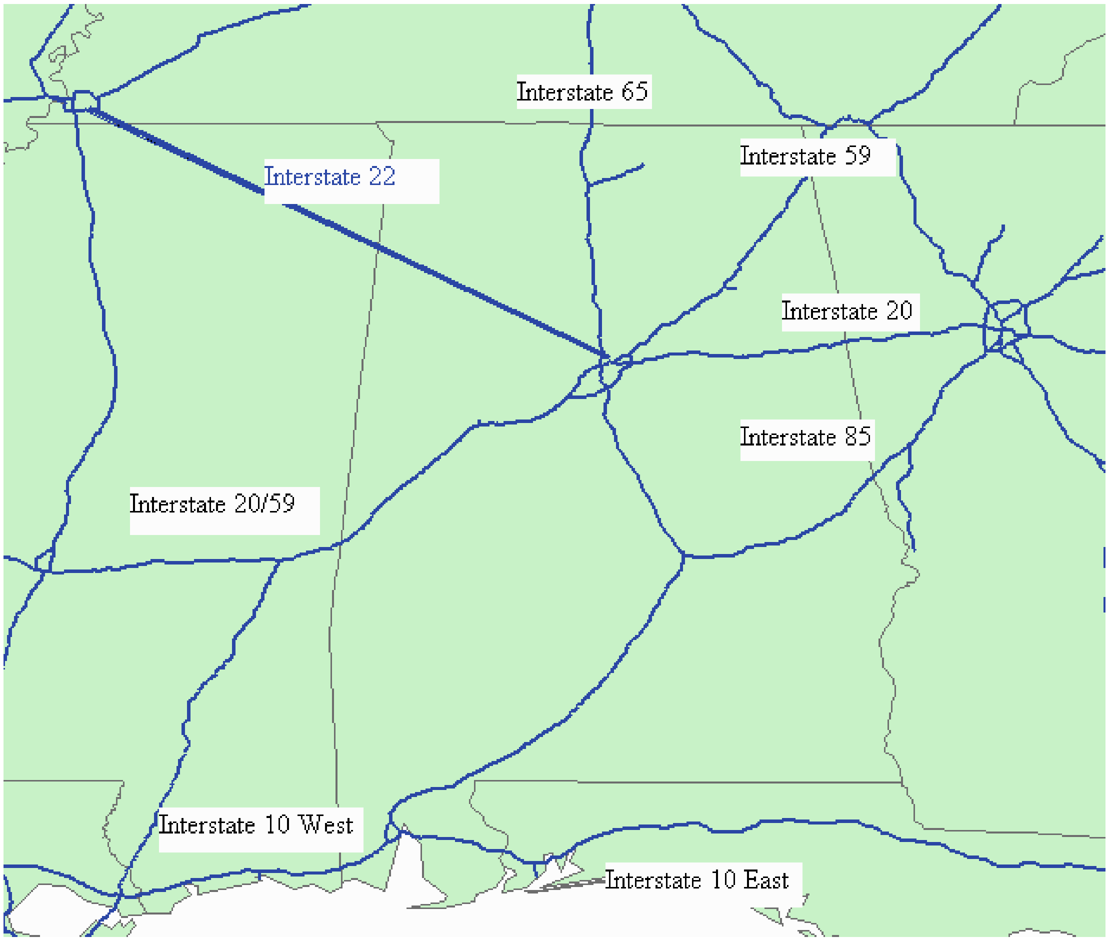

Figure 5. Location of Interstate 22

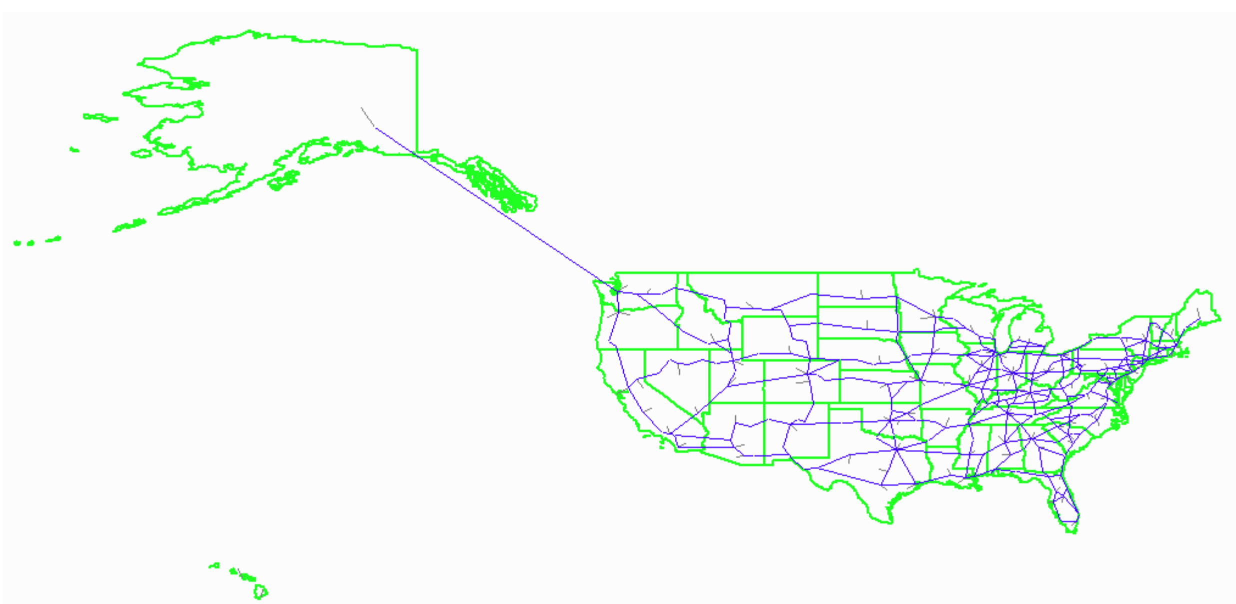

Figure 6. Nationwide Travel Demand Model Network

Assigning the data from the FAF2 database and focusing on the amount of freight passing through Alabama, it can be seen in Table 5 that there is an increase in annual tons of freight expected to pass through Alabama as a result of the completion of the interstate. Table 5 shows the increases expected to cross the state in five year increments, as collected and maintained in the FAF2 database.

Table 5. Total Annual Tons of Freight Passing Through Alabama

Model	2015	2020	2025	2030	2035
Without I-22	432,874,110	484,783,310	552,931,760	639,775,080	744,630,330
With I-22	537,344,830	600,803,970	684,808,050	789,047,250	911,172,470
% Increase	24	24	24	23	22

Examining the road-by-road impact of Interstate 22, Table 6 presents a variety of increases and decreases. Table 6 depicts the differences in annual tons expected for each roadway after the completion of Interstate 22 (a positive number means additional freight; a negative numbers means a decrease in freight). Figures 7 and 8 show a visual representation of the flow of freight for the travel demand models without Interstate 22 and with Interstate 22, respectively.

Table 6. Difference in Annual Tons for each Alabama Interstate

Interstate	2015	2020	2025	2030	2035
I-22	17655520	19464200	21624380	23953370	26446160
I-65	-6176640	-6949890	-7744380	-8589850	-9473720
I-59	3723030	4178960	4889210	5779650	6676530
I-20	10318240	11264150	12481860	13835430	15333990
I59 & I20	2562390	2928800	3491070	4251560	5038080

Figure 7. Freight Flow without Interstate 22

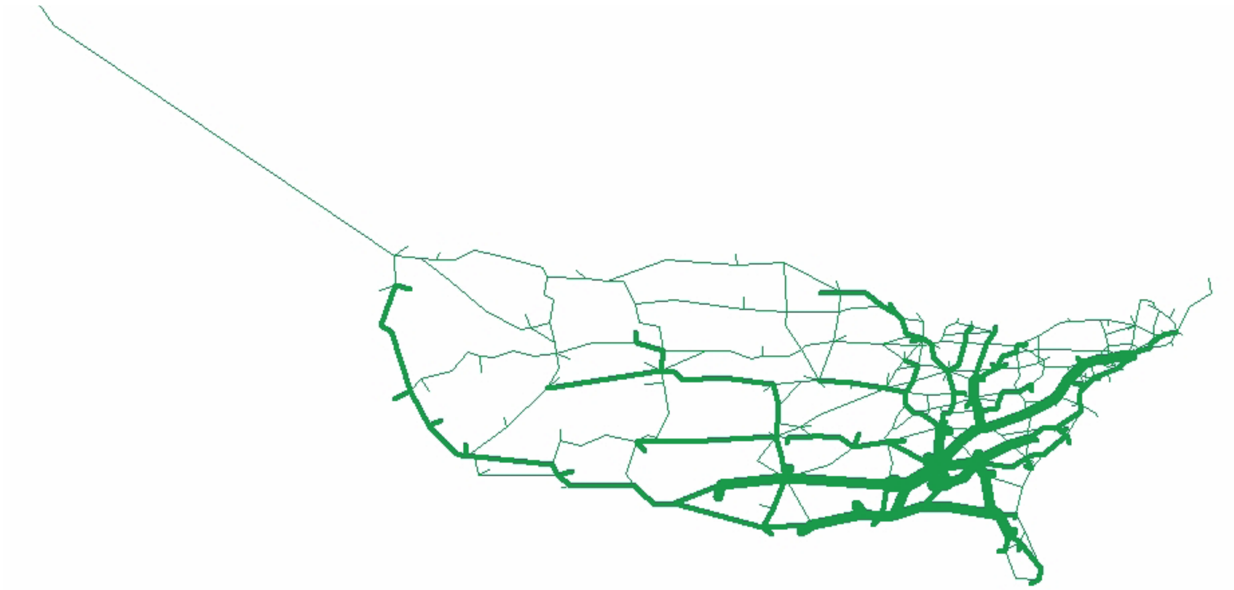

Figure 8. Freight Flow with Interstate 22

The completion of Interstate 22 will reduce freight transportation only slightly on Interstate 10 and somewhat significantly on Interstate 65. This reduction in freight will not only affect Alabama, but will also impact roadways in Tennessee and Kentucky due to the effect on Interstate 65. These future reductions in freight needs will have beneficial impacts regarding congestion and maintenance in Alabama as well as the neighboring states.

The other interstates show a significant increase in freight transportation. The Birmingham area will experience significant growth in freight transportation traffic since many of these facilities intersect within the city. These increases will also affect how Alabama allocates infrastructure money, but will also impact Georgia, Mississippi and Tennessee as these interstates extend beyond the state line.

3.2 Forecasting Freight Traffic Using FAF2 Data

National projections are that freight shipments will double in the next ten years. The increase in freight will have a significant impact on the level of congestion along the national transportation infrastructure and will require innovative congestion mitigation solutions. A detailed understanding of the impact of the projected increase in truck traffic on the existing highway system is needed to examine in the potential outcomes and develop a focused plan to accommodate the anticipated increase.

3.2.1 Model Validation

The use of freight data in transportation modeling requires confidence in the data being used for decision making. The method in transportation planning activities to gain confidence in the data is through a validation process, essentially, determining how well the freight data developed from the model matches the freight data observed through actual traffic counts. To perform this validation, a complete model for freight data was needed to generate freight at the county level, distribute freight between counties and assign freight to expected roadways in Alabama to determine if the assignment met closely with the actual volumes. Only after this preliminary validation of the freight volume, could the model be trusted to provide accurate future volumes when the various scenarios were developed

3.2.2 Development of Freight Scenarios

The scenarios developed to explore the impact of increasing the number of trucks on the state's highway infrastructure were intended to provide a snap-shot of "what-ifs" to the current capacity available. It is important to note that the Alabama Department of Transportation is constantly adding capacity to the roadway infrastructure and the analysis performed assumes that the state's infrastructure is held constant, as a mechanism to identify potential choke-points and assist in focusing the scarce resources of the state.

Using the original ALDOT volumes and capacities and the freight data provided from the FAF 2, there were five growth scenarios examined in this work. The first scenario was a trend line projected growth using only the existing traffic count and growth percentage obtained from historical traffic counts to the year 2015. This scenario was the anticipation of what would happen if traffic volumes grew as traffic volumes had been growing in the past.

The next four scenarios all took different routes to forecast the amount of freight expected on the state's roadway infrastructure. However, the constant in the scenarios was that the number of passenger cars expected on the roadways was all based on the trend line projection of current passenger car level forecasted to 2015 using the historical growth factors.
The second scenario involved forecasting the freight using the projection for 2015 that was developed within the FAF2 database. The procedure for disaggregating the freight data to the county level identified earlier was followed. Additionally, the passenger car volumes were added afterwards.

The third scenario involved a doubling of the 2002 FAF2 freight flow data. This doubling was performed such that each county's contribution of freight to the entire state was doubled. Additionally, the passenger car volumes were added afterwards. The fourth scenario involved double selected counties in Alabama, not the state as a whole. The top sixteen counties, representing 25 percent of the state, were modeled with a doubling of freight volumes. Additionally, the passenger car volumes were added afterwards.

The fifth scenario involved using the FAF2 database forecast of 2035 freight volumes and making the assertion that the state would reach this level of freight activity by 2015, due to the tremendous growth in industry and freight movements. As with scenario 2, the procedure for disaggregating freight data to the county level identified in chapter 3 was followed. Additionally, the passenger car volumes were added afterwards.

3.2.3 Run Scenarios and Identify Congestion Chokepoints

The data developed were forecasted in TRANPLAN and ATIM to identify congestion chokepoints. The advantage of operating the two models together is the ability to identify two different measures of congestion. TRANPLAN, which runs a static daily assignment, is used to determine the locations where the forecasted daily volumes are approaching the daily capacities. ATIM, which runs a discrete event simulation, is used to determine the travel times for vehicles on specific roadways segments during the peak period of the day. For the analysis, locations are defined as congested if the volume to capacity ratio exceeds 0.9 on a daily basis in TRANPLAN, and the travel time during the peak congested hour of the day exceeds 25 percent of the travel time that would be achieved if a vehicle could travel at the posted speed limit. Again, it is important to note that these identified chokepoints are based in existing capacity levels, and changes in capacity would alleviate the congestion and improve travel time at these locations. Additionally, varying the definition of congestion would identify different locations and amounts of congestion.

Initially, it is important to know what level of congestion is currently being experienced in Alabama. The TRANPLAN model indicates that there are 329 miles total lane miles of congestion, shown in Figure 9. Using the increased travel time method of calculating congestion, there are 159 center line miles of roadway where actual travel time is 25% greater than travel time at free flow speeds. Based upon this starting point, several scenario based analyses were performed utilizing both the TRANPLAN model and ATIM.

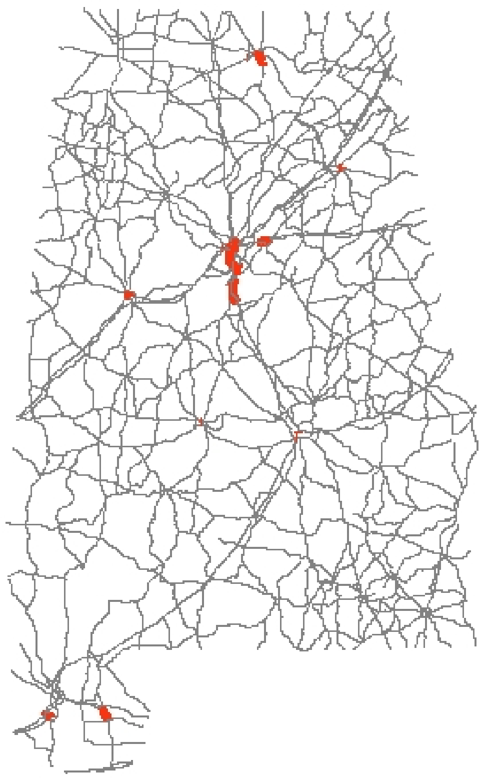

Figure 9. Current Congestion Locations.

3.2.4 First Scenario - ALDOT 2015 Forecast

The first scenario used only the ALDOT 2015 forecast based upon trend line analysis. The traffic growth rate method by which the future network demand would be determined is based upon historical traffic. From the data, the total lane miles of congestion was determined to be 1,421 miles and the congested locations are shown in Figure 10. Based upon the current state of 329 congested lane miles, congestion is projected to grow by 332%.

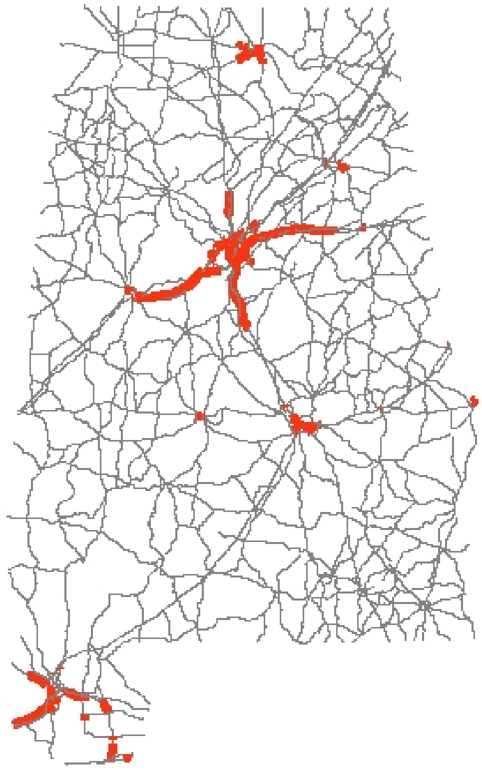

Figure 10. Congested Locations Using Trend Line Analysis.

3.2.5 Second Scenario - FAF2 2015 Forecast

The second scenario used the 2015 projection from the FAF2 database. The origin/destination table for the trucks for this, and all other, scenarios was obtained using the data disaggregating procedure identified previously in this report. The truck data was added to the passenger car projection for 2015 provided by ALDOT. In this scenario the TRANPLAN model resulted in 1,813 total lane miles of congestion. The congested locations are identified in Figure 11. ATIM was used to determine the maximum departure from free-flow travel speed to the travel speed the vehicle actual encounters during the worst period of the day. Figure 12 shows the locations where the actual travel time is 25 percent greater than the free-flow travel time.

These two figures clearly show the value in using the gravity distribution model (TRANPLAN) and a discrete event simulation (ATIM) concurrently to communicate transportation system issues. The TRANPLAN model communicates that the projected increase in congestion, based upon volume to capacity ratio is 451%. The ATIM model communicates that it will take at least 25% longer to travel than expected on 292 centerline miles of roadways, an 84% increase. For freight, the expected travel time is most probably the more important metric.

It would appear that there is discrepancy between the congestion calculations. In reality, the models are indicating that there are locations of congestion as defined by the volume to capacity ratio, where vehicles travel at times greater than free flow speed but less than the 25% threshold.

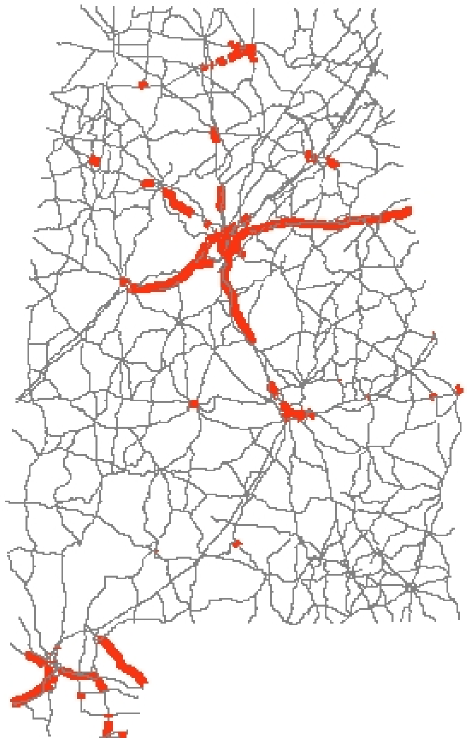

Figure 11. Congested Locations Using the FAF2 2015 Projection.

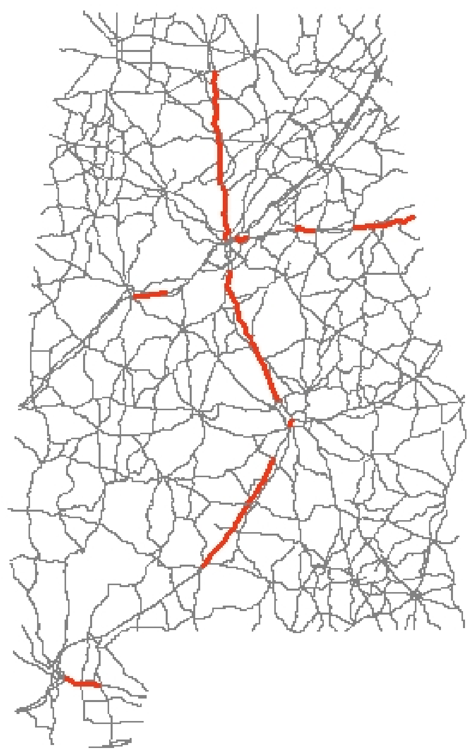

Figure 12. Locations Where Travel Time Exceeds 25% Using the FAF2 2015 Projection.

3.2.6 Third Scenario – Forecast Doubling the FAF2 2002 Truck Traffic Values

The third scenario used the 2002 truck data from the FAF2 database, but the values for each county in Alabama were doubled, which simulates the projection that freight traffic will double in the US by 2020. The truck data was added to the passenger car projection for 2015 provided by ALDOT. This scenario resulted in 2,191 total lane miles of congestion and the congested locations are identified in Figure 13. The resulting growth in congestion, based upon the volume to capacity ration and the TRANPLAN model is 566%. It is important to state that this congestion metric assumes that capacity remains constant.

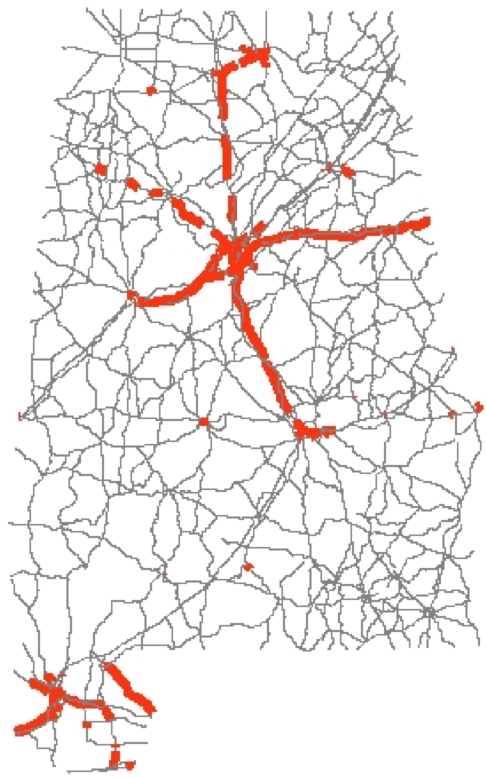

Figure 13. Congested Locations Doubling the FAF2 2002 Volumes.

ATIM was used to determine the maximum departure from free-flow travel speed to the travel speed the vehicle actual encounters during the worst period of the day. Figure 14 shows the locations where the actual travel time is 25 percent greater than the free-flow travel time. The resulting growth in congestion based upon travel time is 692 miles, or 335%.

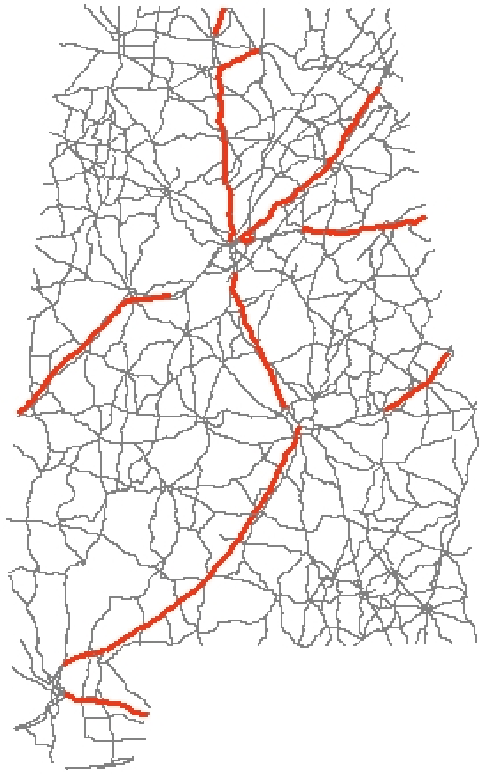

Figure 14. Locations Where Travel Time Exceeds 25% When Truck Traffic is Doubled From FAF2 2002 Volumes.

Figure 15 shows the locations where the actual travel time is greater than 100 percent of the free-flow travel time. This scenario describes a situation where it will take twice as long as to travel I-65 as one would expect at free flow speeds. The resulting impact of freight flow would be significant.

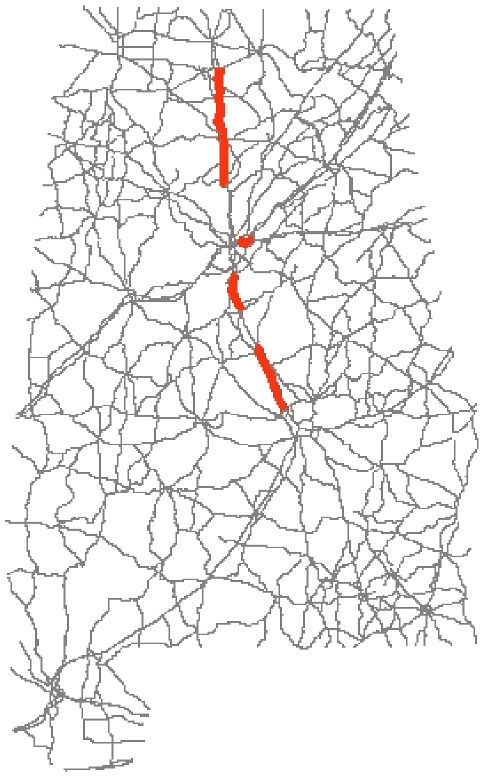

Figure 15. Locations Where Travel Time Exceeds 100% When FAF2 2002 Volumes are Doubled.

3.2.7 Fourth Scenario - FAF2 2002 Forecast Doubling Truck Traffic in Selected Counties

The fourth scenario used the 2002 truck data from the FAF2 database, but only the values for 16 counties with the most economic activity as measured by employment and value of shipments in Alabama were doubled. The truck data was added to the passenger car projection for 2015 provided by ALDOT. This scenario resulted in 2,101 total lane miles of congestion. The congested locations are identified in Figure 16. The resulting growth in congestion, measured by volume to capacity ratio, is 538%. Note that this is not significantly different than scenario three where truck traffic in all counties was doubled, indicating that there are at least 51 counties where the capacity is available to absorb significant levels of growth before infrastructure improvements are required.

ATIM was used to determine the maximum departure from free-flow travel speed to the travel speed the vehicle actual encounter during the worst period of the day. Figure 17 shows the locations where the actual travel time is greater than 25 percent of the free-flow travel time. The resulting growth in congestion, measured by travel time exceeding the free flow time by more than 25 percent is 614 centerline miles or 286%.

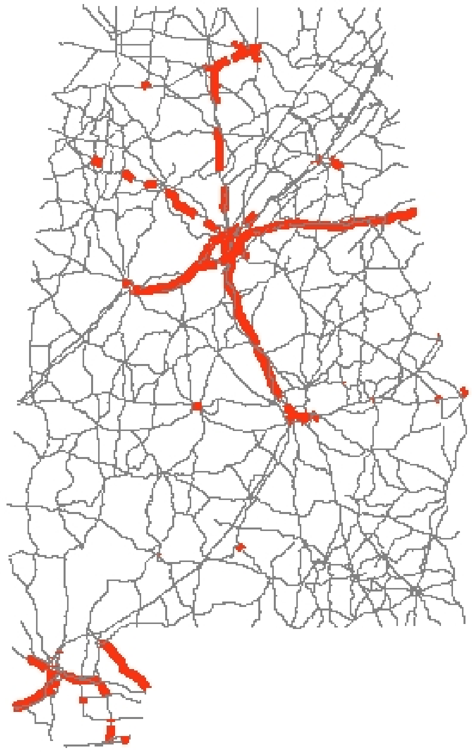

Figure 16. Congested Locations When Truck Traffic is Doubled in Selected Counties.

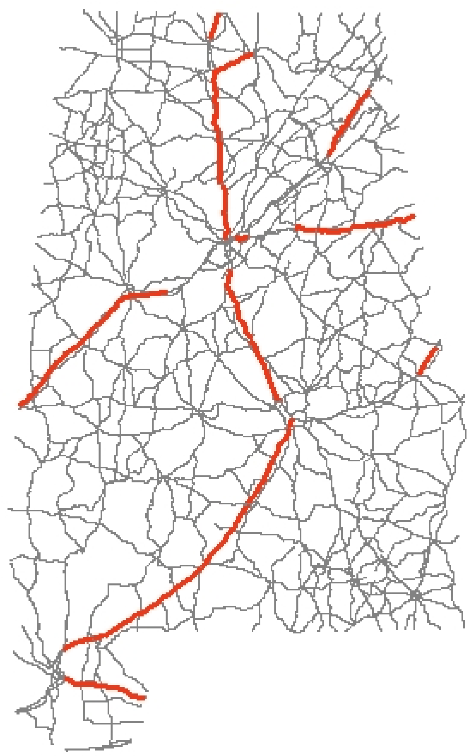

Figure 17. Locations Where Travel Time Exceeds 25% When Truck Traffic is Doubled for Selected Counties.

3.2.8 Fifth Scenario - FAF2 2035 Forecast

The fifth scenario tested used the 2035 truck data from the FAF2 database. The truck data was added to the passenger car projection for 2015 provided by ALDOT. This scenario resulted in 2,105 total lane miles of congestion and the congested locations are identified in Figure 18. The resulting growth in congestion, measured by volume to capacity ratio, is 539%.

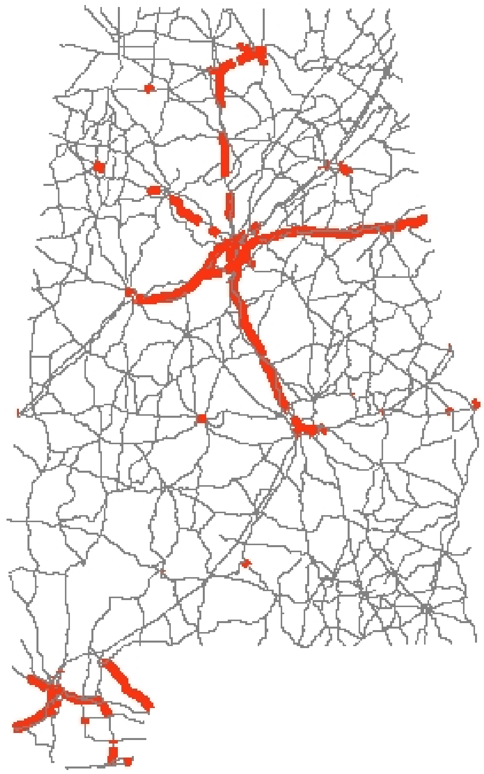

Figure 18. Congested Locations Using the 2035 FAF2 Projection.

ATIM was used to determine the maximum departure from free flow travel speed to the travel speed the vehicle actual encounters during the worst period of the day. Figure 19 shows the locations where the actual travel time is 25 percent greater than the free flow travel time. The resulting growth in congestion, measured by travel time exceeding the free flow time by more than 25 percent, is 694 centerline miles or 336%. Figure 20 shows the locations where the actual travel time is greater than 100 percent of the free-flow travel time.

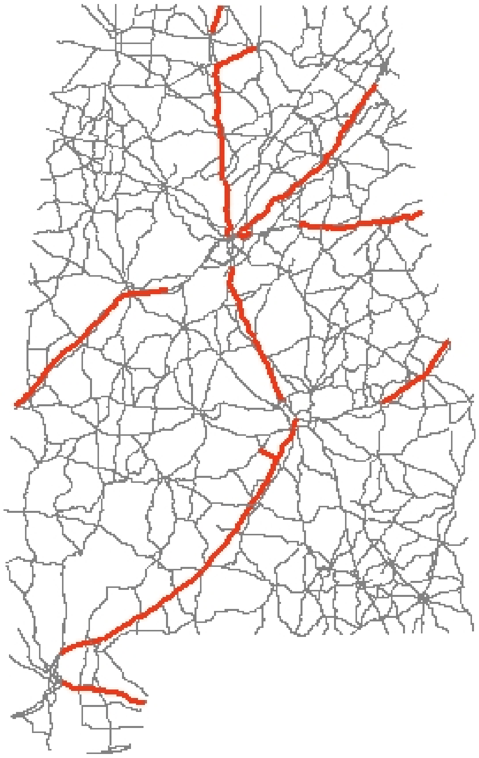

Figure 19. Locations Where Travel Time Exceeds 25% Using the 2035 FAF2 Projection.

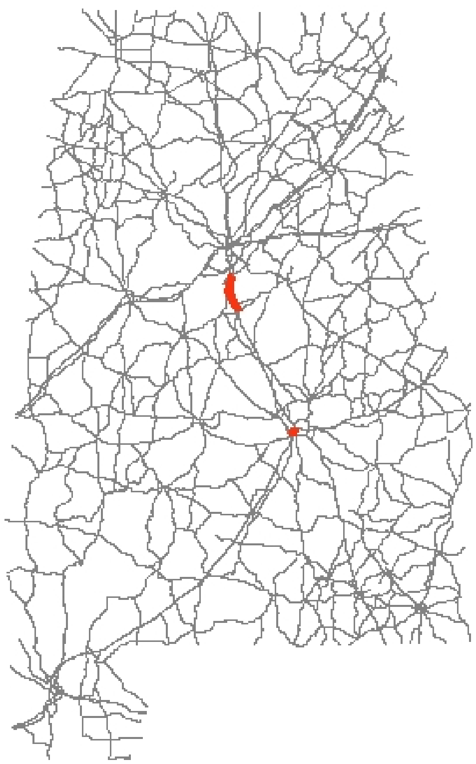

Figure 20. Locations Where Travel Time exceeds 100% Using the 2035 FAF2 Projection.

This ability to quickly develop and test scenarios in the two modeling packages represents a significant improvement, without reducing the accuracy of the models. The ability to identify congestion chokepoints/bottlenecks is important when considering that scarce resources are available for highway improvements and the cost for improvements continues to rise. The scenarios developed for this effort represent only possible growth options developed to test the development of the software packages and interface between them. The scenarios need not be limited after the development of the tools are established and multiple options could have been attempted. The scenarios are important to illustrate what conditions might occur due to funding limitations and other unforeseen situations.

4. Conclusions

The disaggregation methodology developed as a result of this sponsored research has been shown to be effective in creating information for use at the state and local level that was not previously available to transportation planners and modelers. As such, this research has provided a tool that can help the local or state transportation planner integrate freight into the study area transportation plans. This was the ultimate goal of this research and it has been successful. The UAHuntsville research team has developed a core capability in the investigation of interrelationships of freight factors and development of those factors for use in freight planning and modeling.

The disaggregation method developed is part of the Freight Planning Framework (Figure 21) developed by researchers at UAHuntsville. This methodology has been employed in the investigation of freight demand at the state level in Alabama and at the local metropolitan planning organization (MPO) level. The research performed by UAHuntsville on freight demand planning has been incorporated into the Alabama Statewide Transportation Plan and is being used to integrate freight into the South Alabama Regional Planning Commission's (Mobile, AL) 5-year transportation plan.

The foundation of the FPF methodology is the use of industry sectors to focus the understanding and analysis of the economic factors in an area that permits knowledgeable and informed decisions on transportation infrastructure issues. The concept is that if the underlying principles of freight demand generation can be discovered for a particular industry, the ability to predict accurately the infrastructure requirements due to the need of that industry sector to access the freight transportation system is enhanced. Once the freight generation principles of an industry are determined, it is theoretically possible to apply those principles anywhere the industry exists to estimate the demand for freight system requirements. In the FPF methodology, Value of Shipments (VoS), Personal Income, Population, and Employment are utilized as planning factors. One factor alone cannot adequately define the demand for freight system requirements. The planning factors used are capable of describing the freight generation characteristics of a region and the freight attraction characteristics of that region.
This discussion of the FPF is simply to provide evidence that the investigation of freight demand factors for prediction and allocation of freight is an area of significant research

opportunity. This area of research opportunity will grow as we get better at developing the factors that best replicate freight levels on the Alabama roadways, thus providing a higher quality of information that helps to better direct the resources of the state to those projects that support the goals and objectives of the transportation network. The methodology used in the FPF can be used to analyze other freight factors, whether they are global (predictive) in nature or more aligned for allocation of freight in a modeling environment.

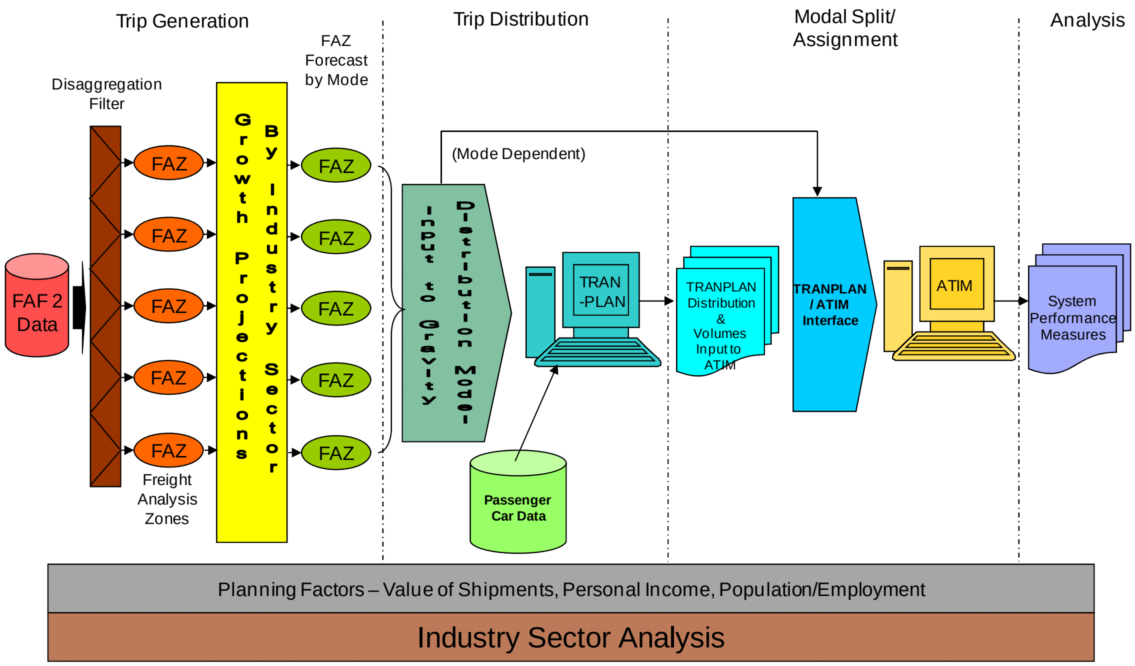

Figure 21. The UAHuntsville Freight Planning Framework

The initial use of counties as the disaggregation level for the freight data appeared promising and has easy initial understanding until the number of counties creates a data matrix that becomes excessively large and unwieldy. The UAHuntsville research team believes that the ability to organize counties into Freight Analysis Zones provides a more efficient and effective way to organize the data into a user-friendly form. Some preliminary research into this concept developed an initial methodology for developing Freight Analysis Zones at a State level. The results found indicate that the development and use of Freight Analysis Zones for including freight in the overall transportation plan provides value and can improve the planning process without increasing the cost and time required to disaggregate all data down to the county level. Future research into the concepts of Freight Analysis Zones needs to continue through the examination of ways to improve the freight data disaggregation methods developed under this research and the associated travel model results.

References

[1] FAF2 Internet Page. http://ops.fhwa.dot.gov/freight/freight_analysis/faf/index.htm (August 1, 2007).

[2] FAF2 Areas. http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/cfs_faf_areas.htm (August 1, 2007).

[3] UAH Office for Economic Development (2005), *Transportation Infrastructure In Alabama -Meeting the Needs for Economic Growth*, Final Report on the Requirements for Infrastructure and Transportation to Support the Transformation of the Alabama Economy. Prepared for the Office of the Secretary, U.S. Department of Transportation, Grant No. DTTS59-03-G-00008.

[4] Monsere, C.M.(2001). "A GIS based multi-commodity freight model: Typology, model refinement and field validation" Dissertation for Degree in Civil Engineering, Iowa State University.

[5] Cornell, J.A. (1990). "Experiments with Mixtures" Second Edition, John Wiley and Sons, Inc. 1990.