USERS' MANUAL

NEURAL NETWORK APPLICATION FOR PREDICTING IMPACT OF TRIP REDUCTION STRATEGIES

Project Team
Philip L. Winters
Francis Cleland
Mark Burris
Dr. Rafael Perez
Michael Pietrzyk

Center for Urban Transportation Research
University of South Florida
4202 E. Fowler Ave., CUT 100
Tampa, FL 33620-5375
(813) 974-3120

February 1998

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the Florida Department of Transportation or the U.S. Department of Transportation. This report has been prepared in cooperation with the State of Florida Department of Transportation and the U.S. Department of Transportation, in partial fulfillment of HPR Study No. 0763, WPI No. 0510763, State Job No. 99700-3337-119, Contract No. B-9896, CUTR Account No. 21-17-189-LO, entitled "Neural Network Technology for the Evaluation of Trip Reduction Programs" (Philip L. Winters, CUTR, Dr. Rafael Perez, USF's Department of Computer Science and Engineering and Michael Pietrzyk, CUTR, are the co-Principal Investigators). Project Manager is Ike Ubaka, FDOT Public Transit Office.

Contents

Overview

This program requires the user to enter data on their company, current employee travel patterns, and the proposed transportation demand management (TDM) incentives that will be in place. The program then uses this information to calculate the company's current average vehicle ridership (AVR). Next, the program uses three neural network algorithms to calculate the company's predicted AVR after 1 year of having the TDM incentives in place. This allows users to make informed decisions regarding the effectiveness of TDM incentives at their worksite.

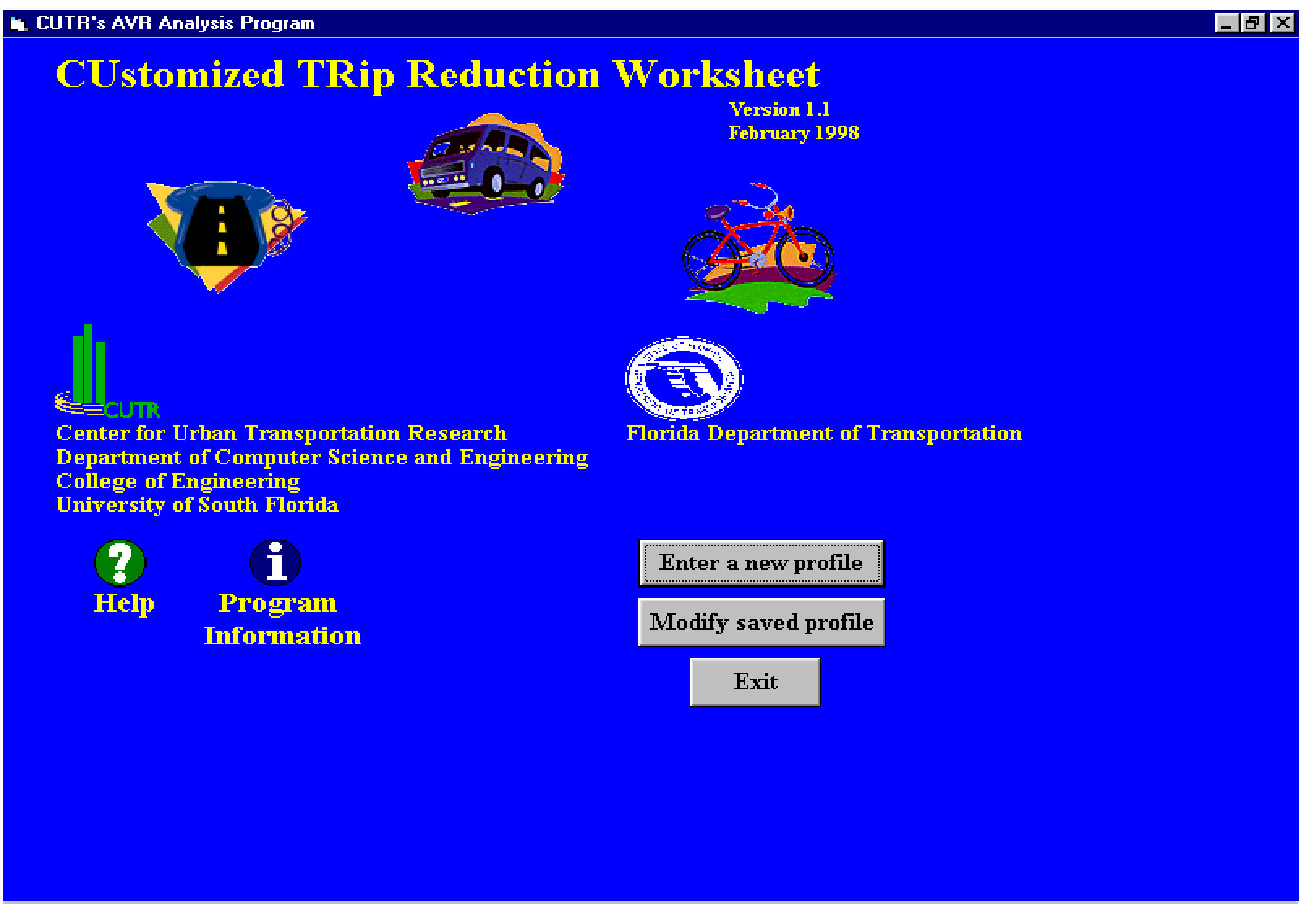
Information regarding this program has been placed in two categories. The first is information that is designed to help the user work with the program that information is contained here. The second category of information gives the user information on the program itself and details on its creation — that information is contained under the **?** information button.

The following pages will lead you step by step through each of the screens.

Title Screen

The user has 5 options on this screen:

? - HELP	provides access to this manual on-line
i - Information	provides information on the design and creation of this program
Enter a new profile	allows the user to start a new profile in a new file.
Modify saved profile	allows the user to modify already saved profiles.
Exit	ends the program.



Profiles contain all the information for a particular company's AVR reduction plan. Each profile contains the following information:

Company Name
Profile Name/Identifier
Address
Plan Start Date
Site ID
Employees
Local Area Population
Local Area Size (square miles)
Person who entered the data
Drive Alone
in 2 Person Carpools
in 3 Person Carpools
in 4+ Person Carpools
in Vanpools
in Buspools
using Transit
Biking
Walking
Telecommute/Compressed Work Week
Sick/Vacation
With Commutes > 40 Minutes
Higher Costs for Driving Alone
Guaranteed Ride Home
Ridematching Program
Alternative Mode Use Subsidies
Compressed Work Week
Current (calculated) AVR
Predicted AVR

A single company can have many different profiles. For example, a user could vary the incentive plans being used by the company and save each profile under a new name. The above inputs to the model are entered through the following screens.

Profiles are saved in profile (.pf1) files on your computer. Each file can contain up to 100 different profiles. Users can create as many different files as they like. The files are comma delimited text files containing the 28 variables listed above.

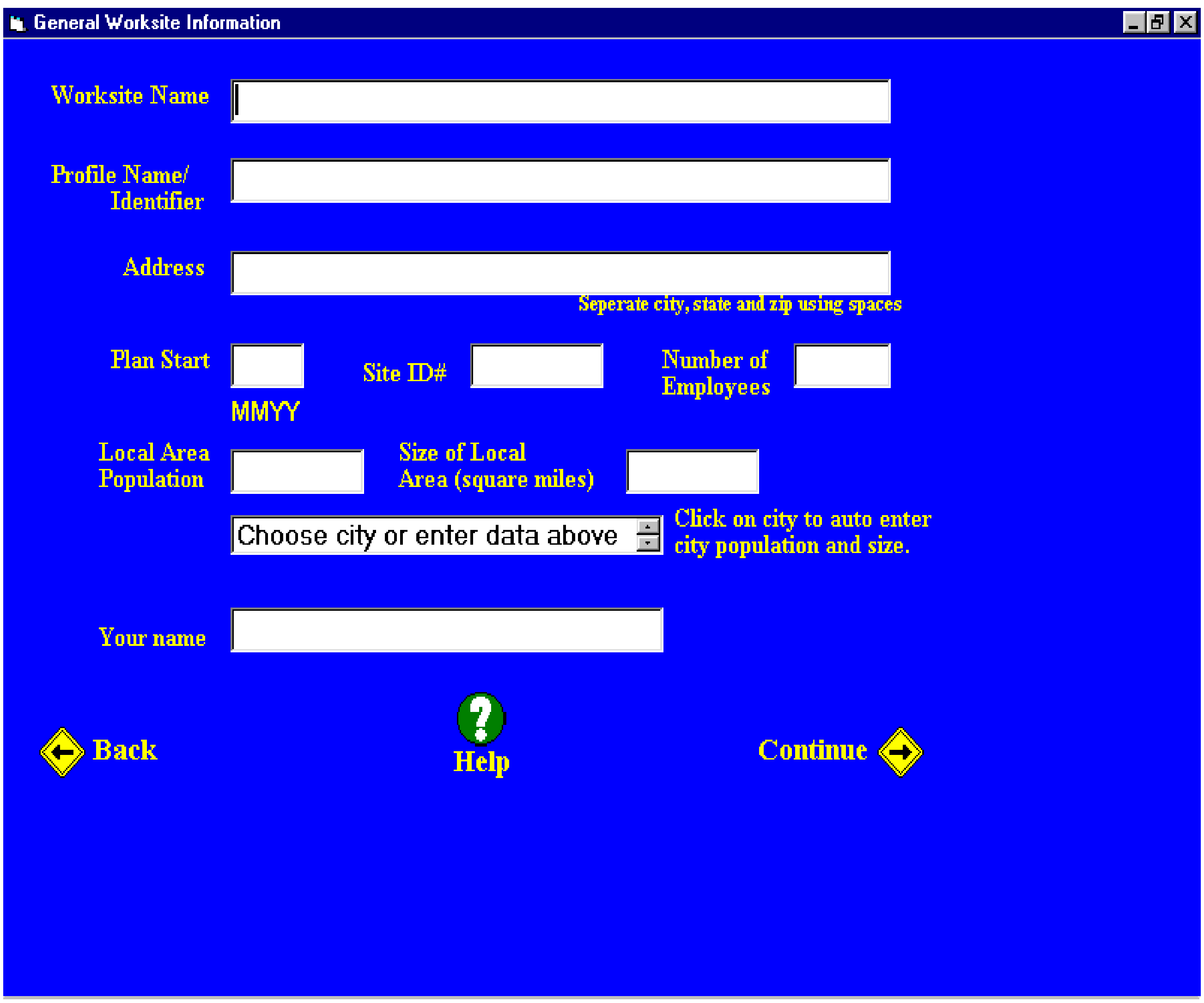
General Worksite Information Screen

On this screen the user is prompted to enter general data on the specific worksite in question. The worksite name, profile name, address, site ID #, and your name fields are for record keeping purposes. You can enter letters or numbers in these fields.

The Plan Start field requires 4 numbers. The first two being the month and the second two being the year. Do not separate the two numbers with a slash.

Hitting back on this screen will take you to the title screen - causing you to lose all data that you entered.

Hitting continue on this screen takes you to the Employee Commute Information Page 1.



Employee Commuter Information Page 1 Screen

In this screen, the user must enter travel characteristics of employees at the worksite. The first section of the screen asks for data on the number of employees that commute to work between 6:00 a.m. and 10:00 a.m. on an average day. In all cases, use data that are indicative of an average day at that worksite.

The next section of the screen asks for information regarding the typical number of employees that do not commute to work. For example, if your company had 100 employees that took an average of 15 vacation and 8 sick days per year, then: assume 260 work days / year

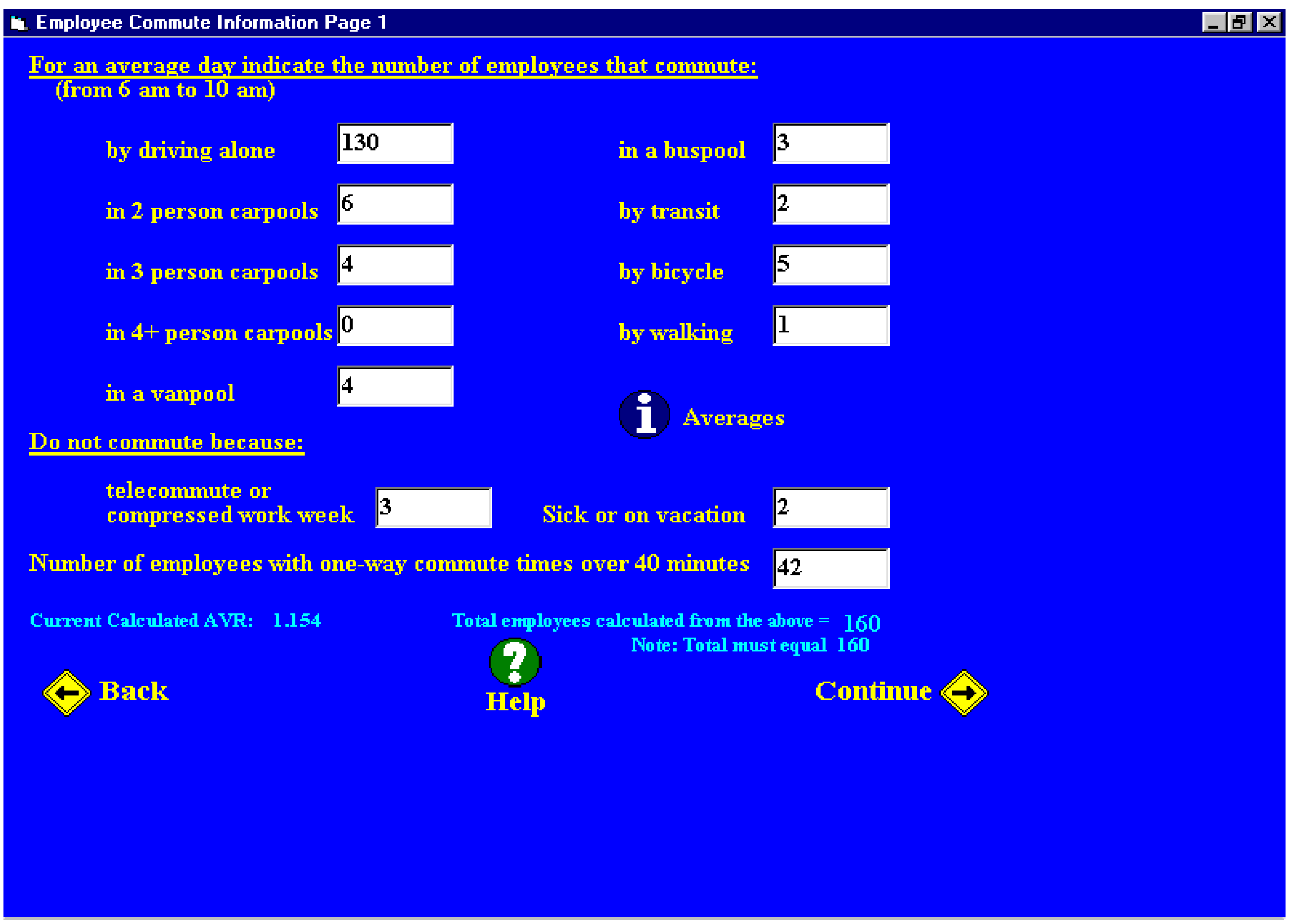
(15 + 8) / 260 = 8.8% of the time an employee is either out sick or on vacation

8.8% * 100 employees = approximately 9 employees are out sick or on vacation on an average day.

All of the data that has been entered on this page (except the number of employees with commute times greater than 40 minutes) must equal the total number of employees you have entered on the General Worksite Information Screen.

Hitting the back button will take the user to the General Worksite Information Screen, without losing any of the data that was entered.

Hitting the continue button takes the user to the Proposed Incentive Plans Screen.

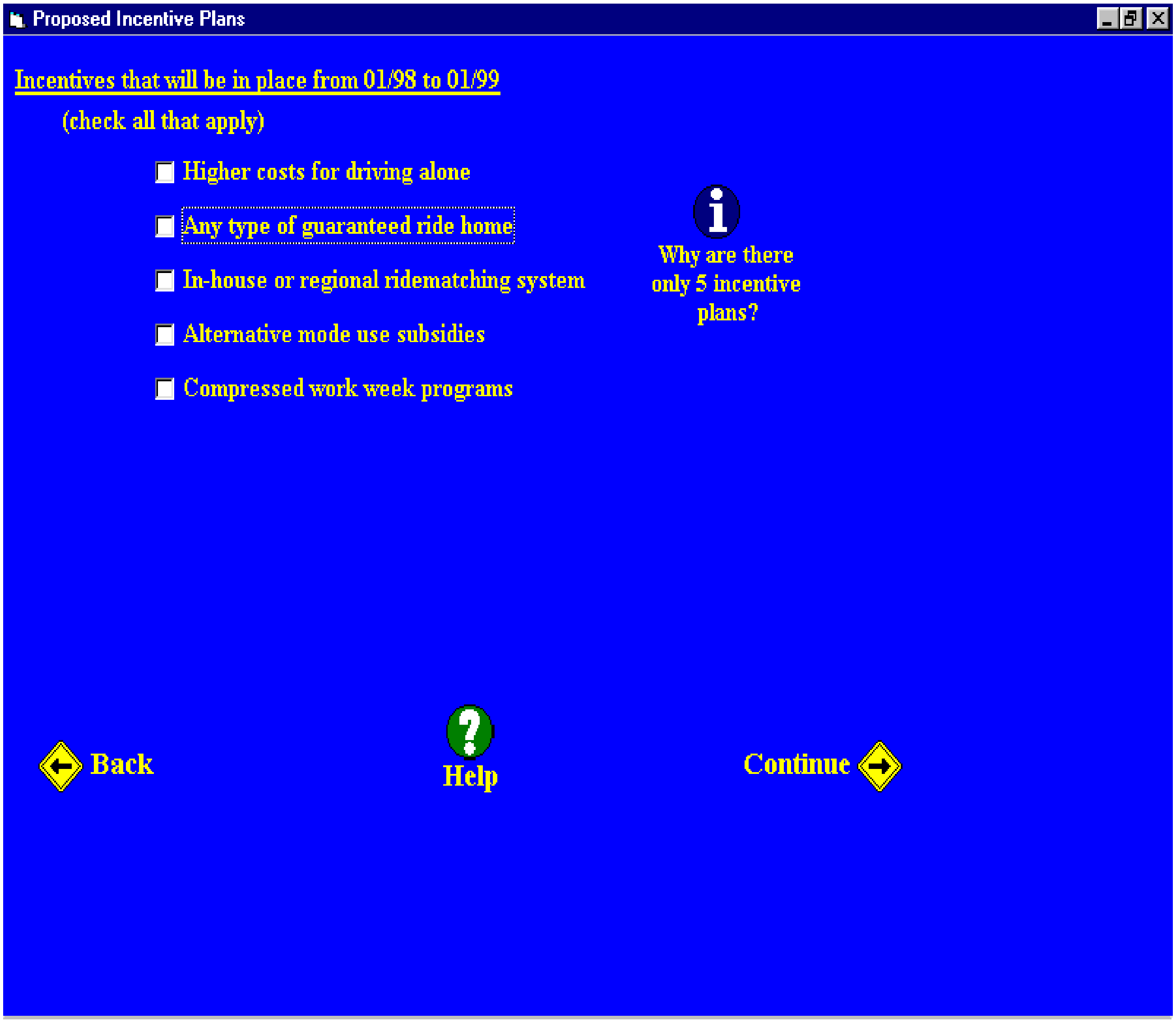


Proposed Incentive Plans Screen

On this screen the user simply clicks on the check box for each incentive plan that will be in place during the time under study (1-year periods).

Hitting the back button will take the user to the General Worksite Information Screen, without losing any of the data that was entered.

Hitting the continue button takes the user to the AVR Calculations Screen.



How Were the Incentives Chosen?

Picking the right input variables is critical to model development. A good subset of variables can substantially improve the performance of the neural network model. The challenge is finding ways to pick good subsets of variables to predict the change in average vehicle ridership (AVR), while keeping the number of input variables to a manageable level.

The neural network software uses a genetic algorithm that selects the variables. This algorithm is looking for sets of inputs (e.g., site characteristics and incentives) that act in a synergistic manner as good predictors of the output (i.e., change in AVR) rather than predicting the impact of every potential variable. The algorithm begins with a population of random variable sets of limited size. As the algorithm progresses, the size of these variable sets will tend to increase if the problem requires larger data sets.

The idea of discarding potentially substantial number of variables is sometimes hard to accept. However, there are plausible reasons for their exclusion by the algorithm.

Why are there only Five Incentives?

It might seem unrealistic that only five TDM incentives can impact employee choice of how to commute. Where are the marketing programs? What about having an Employee Transportation Coordinator in place?

For several reasons, some incentives that might seem effective, or even absolutely necessary, may not appear as options under this program.

1. Some incentives, particularly marketing materials and having Employee Transportation Coordinators (ETCs) in place, were common to many companies in our database. This situation made it impossible for any modeling procedure to determine where marketing worked and where it did not, and, therefore, seemed to have an unpredictable impact on AVR.

ETCs and focused marketing materials are key elements of any TDM program. This fact is one reason why ETCs and marketing materials were common to all of the employer plan submissions that were analyzed.

It is essential that marketing materials and ETCs be put in place to support ongoing TDM programs, to improve awareness and understanding of any of the other incentives (from the list of five that are included in this software) that might be provided in an employer's trip reduction program. Therefore, it is assumed that your company will have marketing materials and an ETC.

2. Some incentives (such as facility improvements) may have been offered by so few companies that it was impossible to accurately determine their impact. Rather than provide an extremely unreliable estimate of the impact of that incentive, we are searching for more data before providing an estimate.

How Much of a Subsidy Needs to be Provided?

The amount of financial subsidy provided is another area where the nature of the data we were using to build our model hampered our efforts to provide an estimate. The extent of financial incentives offered by companies was effectively constrained by the tax code (i.e., employers were less likely to offer more than the nontaxable amount allowed by the Internal Revenue Service. At the time of the plan submittal, transit subsidies were limited to $15 to $21 per month for all plans prior to 1993 and any vanpool subsidy was subject to tax. Hence, we only specify a generic "subsidy", and give no estimate of the impact of increasing the amount of the incentive. It is assumed that when variable indicating a financial subsidy is offered that it is at least $15 to $20 per month per employee using the incentive. Subsidies offered for multiple modes (e.g., transit, vanpool, etc.) could be expected to make a larger impact than the same subsidy for a single mode.

In 1998, the nontaxable transit and vanpool benefit limit that employers may provide employees is $65 per month. See the Internal Revenue Code Section 132(f)(2) for details.

What Incentives Were Considered?

Table 1 shows the number of plans with a given incentive from the data used to build and validate the model. Data from Los Angeles and Tucson were used to build the model. Data from Phoenix was used to validate the model. Those variables included in the final model are shown in the last column with the letter "I" and those that were available for selection but not included are shown with the letter "X".

Table 1 - Frequency of Incentives

Incentive	No. of Records with Incentive	No. of Records without Incentive	Final Model
Rideshare Matching	3644	3336	I
Guaranteed Ride Home	3486	3494	I
Alternative mode subsidies	3227	3744	I
Compressed work week	1769	5211	I
High parking costs for SOV	76	6904	I
Marketing & Promo Activities	4459	2521	X
Preferential Parking	2721	4259	X
Other Services	2655	4325	X
Bike racks and lockers	2620	4360	X
Flexible Work Arrangements	1914	5066	X
Showers & clothing lockers	1554	5426	X
Telecommuting	1058	5922	X
Cafeteria, ATMs, Post Office,	1019	5961	X
Other on-site Services	920	6060	X
Free Meals	771	6209	X
Other compressed work week	675	6305	X
Child care Service	597	6383	X
Walk to work subsidies	454	6526	X
Catalog Points	354	6626	X
Service (unspecified)	320	6660	X
Gift Certificates	304	6676	X
Auto Services	221	6759	X
Additional time off with pay	153	6827	X
Other non-financial Incentives	127	6853	X
Other Facility Improvements	117	6863	X
Other parking Strategies	116	6864	X
Company Vanpools	98	6882	X
Facility Improvements	33	6947	X
Prize Drawings	0	6980	X

AVR Calculations Screen

This screen displays the results of the program's calculations. The top section shows the critical information regarding the average vehicle ridership (AVR) and the number of vehicles per employees. The average vehicle ridership was calculated using the following formula:

AVR = EMP/VEH

VEH = # employees who drive alone + # employees in 2 person car pools/2 + # employees in 3 person car pools/3 + # employees in 4 person car pools/4

EMP = # employees - # employees out sick or on vacation - # employees telecommuting

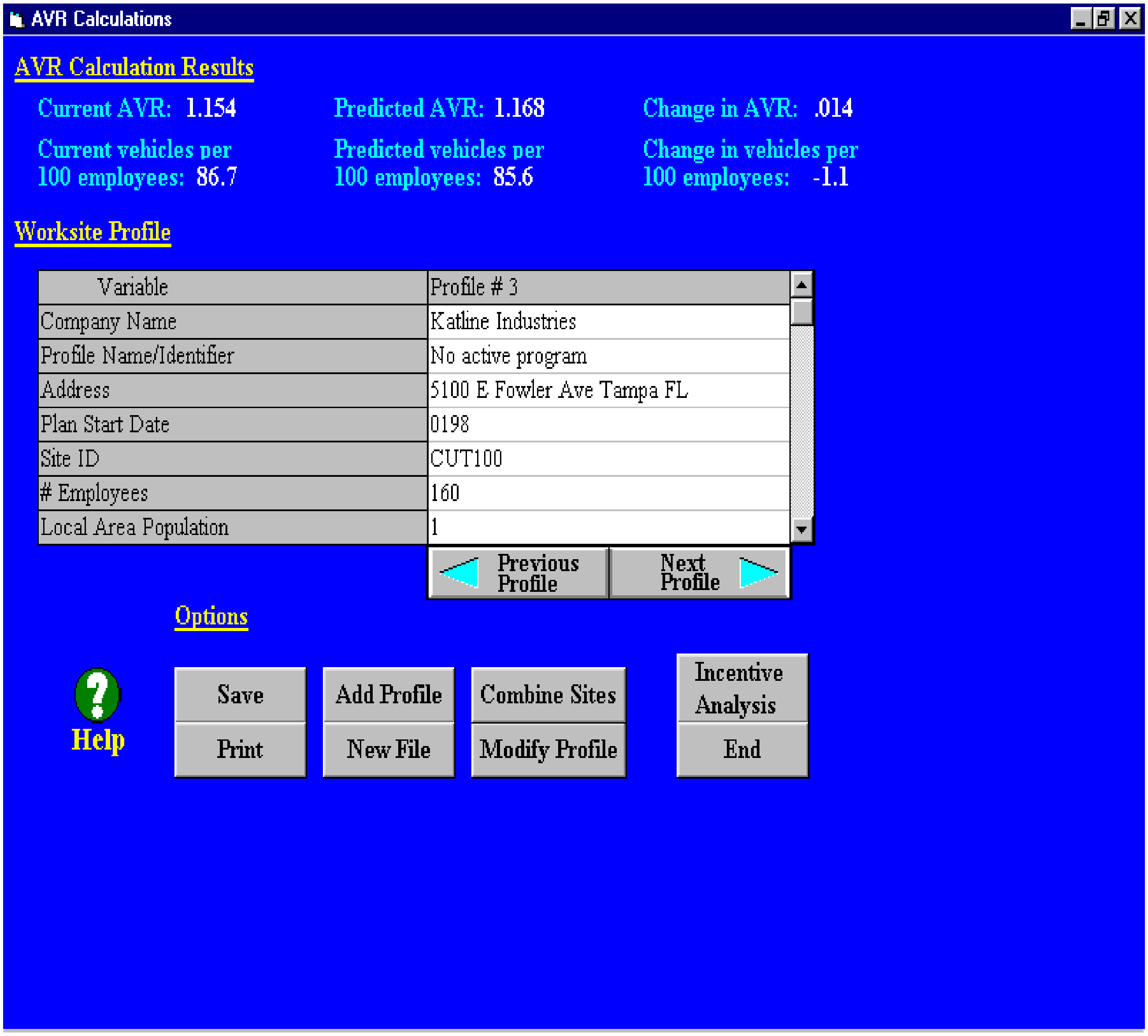
The number of vehicles per 100 employees (NV100E) was calculated using the following formula:
NV100E = 100 / AVR

Under worksite profile the user can see the characteristics of the worksite on which the above calculations are based.

If the user scrolls through the worksite using the previous and next profile arrows, the user will see the other profiles saved in that same (.pfl) file. Therefore, the calculations shown above will be updated to always show the correct calculations for the profile shown in the table. The user can also scroll through the data in each profile using the up and down scroll bar on the right side of the table.

The user also has seven options on this screen. They are:

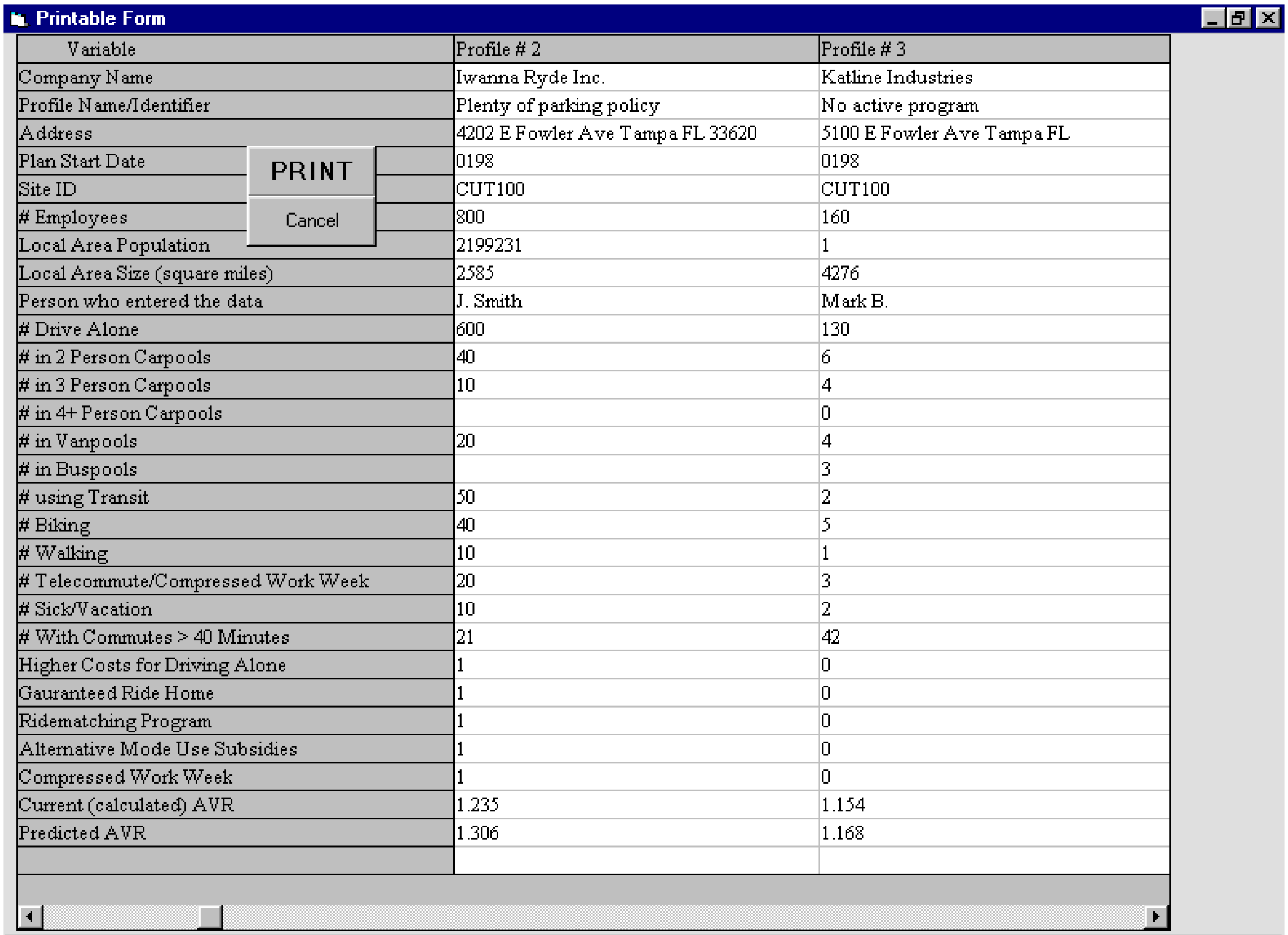
Save	save the entire file (.pfl) shown on the screen.
Print	go to the Printable Form screen.
Add Profile	this option allows the user to add an additional worksite profile to the current file. The user will have the option of having the new profile exactly like the one shown on the screen (and then the user can then alter this to create the new profile) or the user can start with a blank profile.
New Profile	this option allows the user to start a new profile in a new .pfl file.
Combine Sites	the user is taken to the Combine Profiles Screen.
Modify Profile	the user to can modify the profile that is currently shown on the screen.
End	ends the program.


AVR Calculations
AVR Calculation Results
Current AVR: 1.154
Predicted AVR: 1.168
Change in AVR: .014
Current vehicles per 100 employees: 86.7
Predicted vehicles per 100 employees: 85.6
Change in vehicles per 100 employees: -1.1
Worksite Profile
Variable
Profile # 3
Company Name
Katline Industries
Profile Name/Identifier
No active program
Address
5100 E Fowler Ave Tampa FL
Plan Start Date
0198
Site ID
CUT100
Employees
160
Local Area Population
1
Previous Profile
Next Profile
Options
Help
Save
Print
Add Profile
New File
Combine Sites
Modify Profile
Incentive Analysis
End

Printable Form Screen

If the user hits Print the screen will be printed to the default printer. The user has the ability to alter what is shown on the screen, and hence what is printed. The scroll bar(s) allow the user to navigate from profile to profile and, if necessary, scroll up and down through the data. The user can also adjust the width of the columns.

Hit cancel to go back to the AVR Calculations Screen without printing.

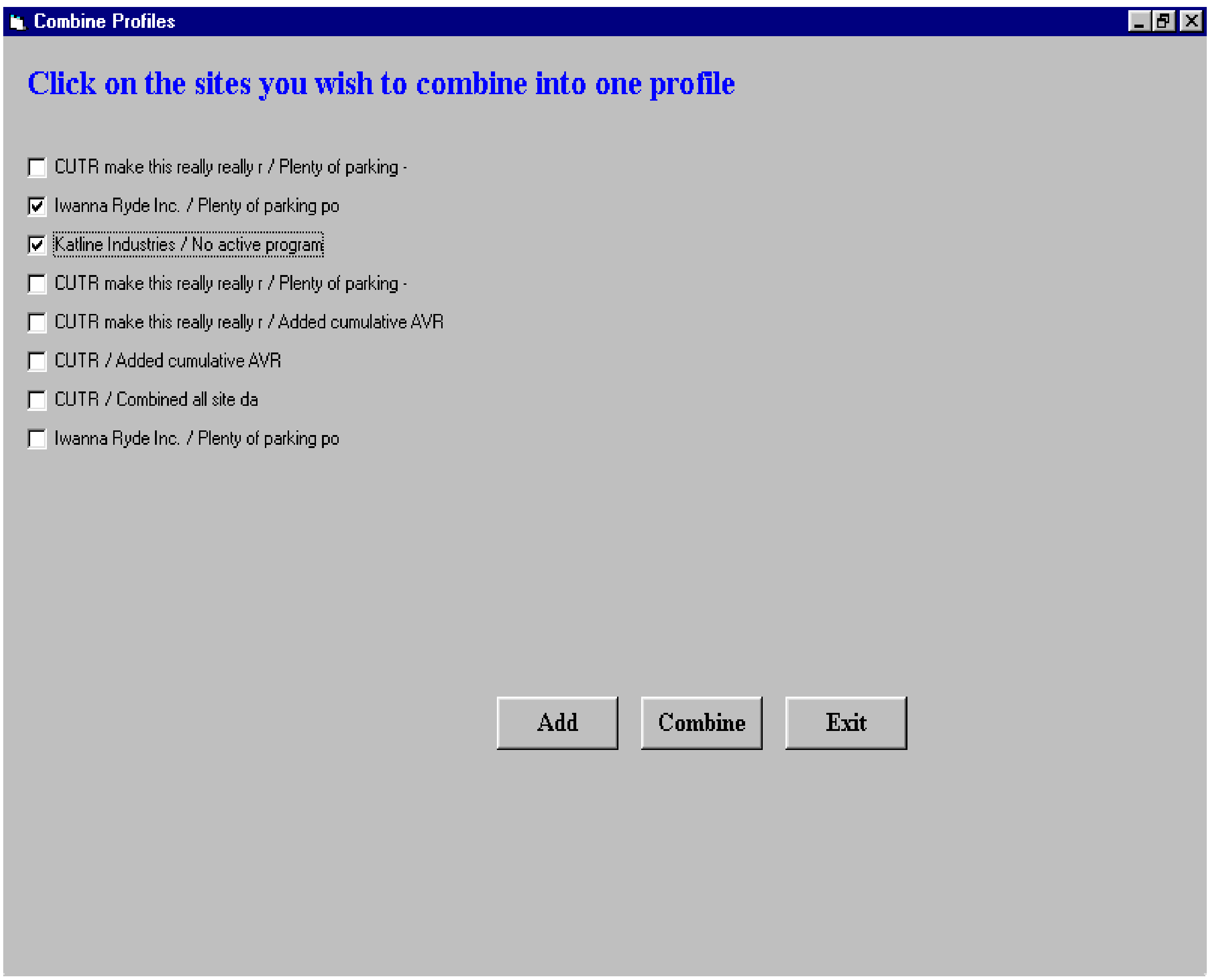


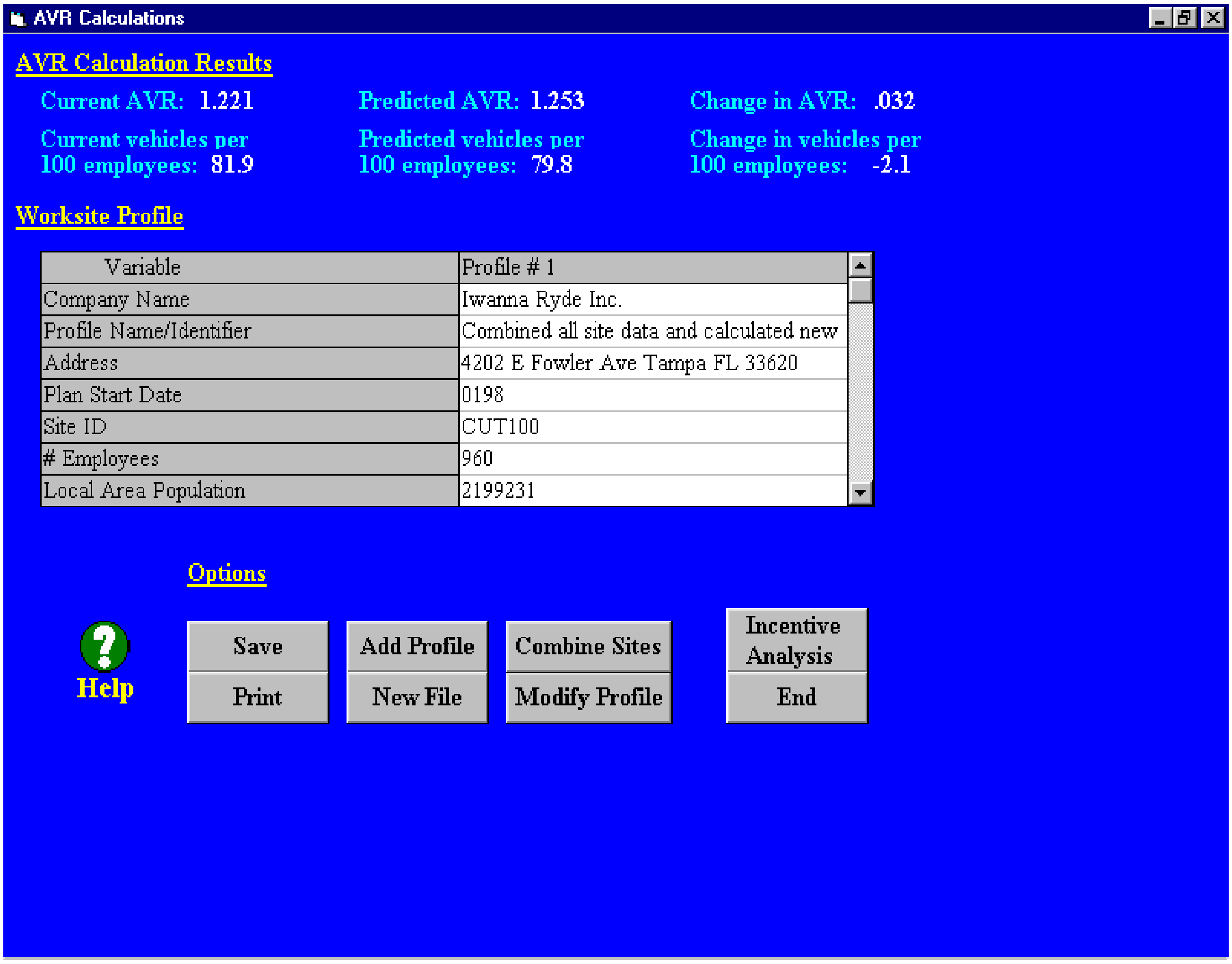
Variable	Profile # 2	Profile # 3
Company Name	Iwanna Ryde Inc.	Katline Industries
Profile Name/Identifier	Plenty of parking policy	No active program
Address	4202 E Fowler Ave Tampa FL 33620	5100 E Fowler Ave Tampa FL
Plan Start Date	0198	0198
Site ID	CUT100	CUT100
# Employees	800	160
Local Area Population	2199231	1
Local Area Size (square miles)	2585	4276
Person who entered the data	J. Smith	Mark B.
# Drive Alone	600	130
# in 2 Person Carpools	40	6
# in 3 Person Carpools	10	4
# in 4+ Person Carpools		0
# in Vanpools	20	4
# in Buspools		3
# using Transit	50	2
# Biking	40	5
# Walking	10	1
# Telecommute/Compressed Work Week	20	3
# Sick/Vacation	10	2
# With Commutes > 40 Minutes	21	42
Higher Costs for Driving Alone	1	0
Gauranteed Ride Home	1	0
Ridematching Program	1	0
Alternative Mode Use Subsidies	1	0
Compressed Work Week	1	0
Current (calculated) AVR	1.235	1.154
Predicted AVR	1.306	1.168

Combine Profiles Screen

This part of the program allows the user to combine worksites (that are contained in the same .pfl file) and create one large worksite. This can be used to combine all the worksites for an area, or for example, for an entire college campus. This can be accomplished in one of two ways.

The user begins by selecting 2 or more of the profiles listed by clicking on the check box. The program will then ask the user a series of questions to accurately create the new profile. This new profile will then be stored in the current .pfl file. If the user choose to add profiles together then the sum of the already calculated AVRs will be shown. If the user choose to combine profiles, then the companies will be merged as one and a new AVR will be calculated for that combined company. In both cases all incentives are turned off and it is up to the user to modify the profile and turn on the various incentives they wish to have at all combined companies.




AVR Calculations
AVR Calculation Results
Current AVR: 1.221
Predicted AVR: 1.253
Change in AVR: .032
Current vehicles per 100 employees: 81.9
Predicted vehicles per 100 employees: 79.8
Change in vehicles per 100 employees: -2.1
Worksite Profile
Variable
Profile # 1
Company Name
Iwanna Ryde Inc.
Profile Name/Identifier
Combined all site data and calculated new
Address
4202 E Fowler Ave Tampa FL 33620
Plan Start Date
0198
Site ID
CUT100
Employees
960
Local Area Population
2199231
Options
Help
Save
Print
Add Profile
New File
Combine Sites
Modify Profile
Incentive Analysis
End

APPENDIX A - PROJECT SUMMARY

Background

Rising traffic congestion and air quality problems had contributed to federal, state, and regional efforts to reduce vehicle emissions by requiring large employers to develop programs to reduce vehicle trips. In areas with the worst air pollution, the program's goal was to reduce driving-and pollution-by increasing the average number of employees in vehicles commuting to work (i.e., average vehicle ridership or AVR). Employers were targeted by these regulations as employer policies such as work location, work schedule, and parking policies strongly influence transportation mode choice decisions made by employees.

In several of the major urban areas of the country (e.g., Los Angeles, Phoenix, Seattle), large employers with 100 or more employees were required by federal, state or local regulation to submit detailed plans for influencing employee travel behavior in order to reduce air pollution and/or traffic congestion. Over the years, these metropolitan areas collected a large amount of data from these companies. Information was obtained that described different company site characteristics and the alternative modes of transportation available to the employees. The data also included information on the types of financial and non-financial incentives employers offered to employees. Employers provided information on work schedules and alternative work arrangements such as telecommuting and compressed work weeks. They also collected information from employees on the different modes of transportation selected by the employees and estimated the site's AVR.

Though areas such as Los Angeles had thousands of employer plans submitted under these regulations, the regulators have not been able to produce a predictive models using the aggregate data sets collected from employers or the disaggregate employee data information processed by the regulatory agency. Part of the reason rests with the complexity of the data. The Los Angeles area database, for example, includes 62 different incentives that employers can select to increase AVR in their work sites. Some incentives are offered by relatively few employers. Even when condensing the incentives into 28 categories, the plans represented about 1,500 different combinations of incentives. When the type of employer, current mode split, and employee composition is considered, the uniqueness of each plan becomes evident.

At the same time, the current models (e.g., FHWA TDM Model) are based on disaggregate data collected through relatively small samples of employers but augmented by employee surveys. Specifically, model predictions were not compared with actual results for any data that had not been used in the model building process.

Project Objective

Under this Florida Department of Transportation (FDOT) Research Idea project, the project team of the Center for Urban Transportation Research (CUTR) and the Department of Computer Science and Engineering at the University of South Florida applied neural network technology to use the large, data rich, aggregate data from employers to predict the impacts of various trip reduction strategies on changes in commute behavior. These changes in commute behavior were measured as changes in average vehicle ridership (AVR). The performance, and selection of the best model, was based on comparing neural network output to actual AVR observations. The neural network training (or learning) process allows the neural network model to predict the correct response to combinations of input data values not previously seen by the network. The benefits of developing such a model would be to streamline development of trip reduction plans for employers, increase effectiveness of those plans, and provide a basis for consistent review by the regulating agencies. It should also improve efficiency by reducing regulatory staff time in the review of employer/developer trip reduction plans.

A Primer on Neural Networks in Transportation: Concepts and Applications Artificial neural networks (ANN), synonymous with neural networks, represent a form of computer intelligence and operate similarly to the human brain, but on a very reduced scale. Artificial neural networks are being used today to predict results by learning from existing input and resulting output data in science, engineering, medicine, banking, management, marketing, manufacturing, and sports wagering.

To develop or "train" the model, the data set is usually divided into two groups one group for training the network and another group for testing how well the network has learned. A third independent data set is often reserved for validation. Each training set of data is presented to the network. If the output of the network differs from the correct output then the weights of individual network nodes are changed. Training a neural network requires many cycles until the cumulative errors of all training sets are below an acceptable level, as pre-defined by the neural network builder. The lower this number the better the network is able to duplicate the associations between inputs and outputs in the training data. It is expected that once the network is able to duplicate the associations between inputs and outputs in the training data, it will be able to produce correct outputs for input data not specifically included previously as part of the training data. The training set of data uses a list of paired input and desired output patterns to avoid overfitting the data. Overfitting the training data occurs when the neural network produces a nonlinear model that fits the training data perfectly, but fits the test data very poorly.

Training a network using back propagation (the method used in this project) consists of finding the weight values so that the associations between input and output in an existing data set can be duplicated by the network. Since each neuron implements a non-linear mapping between its inputs and output neural networks are capable of learning non-linear relationships that may exist in the data. This makes neural networks adaptable and especially useful in environments where the relationships between inputs and outputs change over time.

Comparison of Neural Networks to Other Modeling Techniques

Neural networks deal with a broad range of problems. Artificial neural networks are known to be good at classification, evaluation, optimization, decision-making, pattern recognition, behavior trend prediction, image analysis, filtering, and modeling control systems.

There are some significant differences between expert systems and neural networks. Expert systems require that the relationships between the input data and the conclusions to be derived from that data be established before the expert system is built. The neural network needs the data from which it can uncover the relationships while the expert system needs the expert who has already learned those relationships. Another important difference can be found in the encoding of the data. Expert systems encode their knowledge in terms of results, object descriptions, and procedures. After training, neural networks encode their knowledge of relationships in terms of weight values and in the inter-connection between the neurons.

Updating the knowledge in the system is another area where neural networks and expert systems differ. If the problem domain changes and new knowledge is required, this knowledge must be obtained from the human expert and carefully crafted into the already existing software knowledge structures of the expert system. A neural network would need input that reflects the changes in the problem domain with the corresponding conclusions that can be drawn from the data in order to retrain itself.

There are other machine learning techniques in addition to neural networks and expert systems. Neural networks form a category of learning techniques called "connectionist". This term emphasizes the dependency of neural networks on the connectivity of a large number of computational units. Other machine learning techniques rely on the manipulation of symbols used to create rules similar to the "IF-THEN" rules used widely in expert systems and are grouped under the category of "symbolic" learning techniques. One of the main differences between neural networks and these other symbolic learning techniques is in the form of the knowledge that they learn from the data presented to them.

Another important difference is in the range of problem domains that they can effectively deal with. Symbolic learning methods deal mostly with classification problems. The assignment of a class label to an object or situation based on the specific values of a set of parameters. The neural network models can learn not only to classify data into different categories but to predict the numerical value of outputs (e.g., level-of-service classification based on volume to capacity ratios or average travel speeds), learn to interpret a visual image, etc.

Probably the most important similarity between neural networks and symbolic learning methods is that they both require a set of representative data from the problem domain in order to learn the relationships that exist between inputs and outputs. There is a need for explicit knowledge of these relationships as long as training and testing data exist.

There are also differences between neural networks and linear regression modeling. Linear regression modeling uses a strictly linear combination of independent variables. Neural networks, on the other hand, provide weights that represent non-linear functions of the input variables. For example, the ANN models are trained to predict deterioration based on various samples of pavement condition data (inputs) that correspond to pavement roughness coefficients (outputs).

Alternative Modeling Procedures

To provide an indication of the relative ability of the neural networks to predict changes in AVR, and to show the reduction in data needed to conduct this analysis, three methods of alternative modeling were developed. The first was a standard linear regression analysis. This was used because the initial neural networks were essentially attempts to predict the value of the AVR change using a method similar to linear regression. The second method was a linear discriminant analysis, which was used to show the relative ability of the neural network to classify observations into ranges correctly. The SCAQMD data was converted into inputs to the FHWA TDM Model to compare the neural network with this commonly used analytical tool for predicting results of trip reduction strategies.

To get a more comprehensive evaluation of the network's effectiveness, it was determined that an examination of the network's ability to correctly classify each prediction into a range (or a category) of AVR change would be conducted. The ranges were developed by partitioning the data into equal sized groups based on the value of the dependent variable (see Table 1 below).

Table A-1: AVR Change Range Categories for Model Evaluation

AVR change category	Change in AVR category range	Acceptable range for model evaluation
Large decrease	-0.08 or less	Cl. 1: Any change less than -0.03
Moderate decrease	-0.03 to -0.079	Cl. 2: Any decrease
Small decrease	0 to -0.029	Cl. 3: Any change less than +0.03
Neutral	0 to 0.029	Cl. 4: >-0.03, <0.06
Small increase	0.03 to 0.059	Cl. 5: >0.00, <0.12
Moderate increase	0.06 to 0.119	Cl. 6: Any change more than +0.03
Large increase	0.12 or more	Cl. 7: Any change more than +0.06

In effect, the evaluation centers on the model's ability to predict whether a given combination of site characteristics and incentives will produce a large increase in AVR, a small increase, virtually no increase, a small decrease, or a large decrease in AVR. Models were evaluated both through comparison of R (linear correlation) values of predicted and actual change in AVR and by their ability to classify an observation into the correct group or into an adjacent group. This was termed "acceptable" (as opposed to "correct") classification.

Results

After the neural network was built, its performance was compared with the FHWA TDM Model and the alternative modeling procedures.

Table A-2: Acceptable Range Correct Classification by Final Models for TDM Model Validation Data Set (N=432)

MODEL	Inputs	Percent
Neural network	16	54.2
Discriminant	23	58.1
Regression	31	50.2
FHWA TDM Model	N/A	39.6

Table A-3: Linear correlation of prediction and actual output TDM Model Validation Data Set (N=432)

MODEL	R
Neural network	0.312
Regression	0.544
FHWA TDM Model	0.032

The models built were clearly superior to the alternative of using the FHWA TDM Model. As to correct classification, the neural network was superior to the regression procedure in classifying results into the proper ranges to regression, although the correlation of predicted to actual results was lower. The SCAQMD data contained many observations (more than 500) where employers had either a very large increase or very large decrease in AVR. Nevertheless, the vast majority (almost 90 percent) of the data falls near -0.10 to +0.20 change in AVR. Models built on prediction error minimization criteria may force their predictions to the middle of the range (i.e., predict little or no change in AVR). This approach causes the models to have much more accuracy in the middle ranges of AVR change than with the outliers (i.e., large changes in AVR). Preferably, a model should interpolate well over the entire range of the input values. Therefore, the correlation values are not necessarily the most appropriate way to evaluate the model's performance.

Given the neural network's ability to predict change at about the same level as a linear regression, and its ability to do so more efficiently (fewer variables as inputs compared with the linear regression model), it was determined that the best performing neural network should be used.

Approach to Field Testing

The approach for evaluating the transferability of the ANN model was to use the Los Angeles-based ANN model to predict change in AVR using data from another city (i.e., Tucson and Phoenix). Fortunately, many similarities were found between data collected from the Arizona programs and with the SCAQMD data.

Various combinations of the models were built with data from two of the cities and validated with the data from the remaining city. The data included 29 incentive fields. The data contained 5,001 employer plans from LA, 1,103 employer plans from Tucson used to build the new ANN model and 878 employer plans from Phoenix were used to validate the model.

The final ANN model was actually built as three sequential ANN models. All the variables were made available to build the first model to predict change in AVR. The second model was built to explain the residual value (i.e., actual AVR change less predicted AVR from the first model) using only the combined incentive groups (e.g., any guaranteed ride home, financial incentives, etc.). The final predicted value of AVR change is the sum of these three models.

The LA-Tucson based model performed the best in predicting change in AVR (see Table 4). The results of this task show that, based on the data from these three cities, the ANN model is transferable. More observations were acceptably classified in the validation sets than in the ANN modeling data set. Only at the large increase in AVR range did the validation data under perform the base model. This was partially due to the few employer records from Phoenix in that category.

Table A-4: Final Model Performance vs. Validation Data

Model	Inputs	Overall acceptable classification	Cl. 1	Cl. 2	Cl. 3	Cl. 4	Cl. 5	Cl. 6	Cl. 7
Train and Test Data Set (LA - Tucson)	13	49%	31%	15%	27%	67%	91%	78%	33%
Validation Data Set (Phoenix)	13	58%	43%	25%	63%	95%	87%	25%	2%

Though there were nearly 7,000 trip reduction plans used to build and validate the model, there are two points that should be made: (1) some incentives are offered by relatively few employers (see Help screen on Incentives page) and (2) many combinations of the plans illustrate the challenge in finding the "best" plan. Only "marketing incentives" was included by more than half of the filed plans. The plans also represent 1,163 different combinations of incentives when marketing activities were considered to be only present or absent. If the number of marketing activities is considered then there were about 1,500 combinations of incentives. This situation may have had a tendency to reduce the ability of the ANN model to detect any significant change based on the presence or absence of any given incentive.