DEVELOPMENT OF THE GRAPHICAL USER INTERFACE AND IMPROVEMENT AND STREAMLINING OF NYMTC'S BEST PRACTICE MODEL

Jim Lam and Howard Slavin
Caliper Corporation

Introduction

The New York Metropolitan Transportation Council (NYMTC's) "Best Practice" (BPM) model is a large, deployed activity-based travel demand model. The model, however, is complex, difficult to run, and incurs long model run times.

This paper details the development of a new Graphical User Interface (GUI) for the NYMTC BPM model, and describes the streamlining steps developed to drastically improve model performance. The flow chart based GUI makes the model more user friendly, makes the model easier to use and comprehend, and simplifies the process of creating scenarios and alternatives. The paper also describes extra tools that were developed to create automated model reports, and to create a web-based interface that enables users to view and query model output data. Model streamlining steps included code optimization, multi-threading, and distributed processing. We demonstrate that activity models can run fast and be made user friendly thus clearing the way to more widespread use of better travel demand forecasting methods.

Current NYBPM Model

The New York Metropolitan Transportation Council's Best Practice Model (BPM) is a state of the art disaggregate travel demand model developed by PB Americas. The model was developed using an old version of TransCAD for network skimming and assignment and controlled through a GUI interfaced called CENTRAL. At the time of development, the recommended hardware requirement was a PC with 2 CPUs and 3-4 GB of RAM. With that hardware, the model requires long running times. The BPM is an activity-based model that models individual trips through the travel choice process. [1]

Since the initial development, there have been numerous advancements in both software and hardware technologies. On the hardware side, computer processors are now faster, but more importantly, more processors, or cores, are now available on a single computer. In addition, computers typically come equipped with more RAM memory, computer hard drives store more data and are faster, and computer networks can transmit information faster. On the software side, new developments in the travel demand software have taken advantage of these hardware improvements. New procedures allow certain model procedures to make use of the multiple cores in a computer simultaneously and to take advantage of the increased memory to improve model performance. In addition, the software advances allow multiple computers to be used in a distributed format to run models faster. [2] On the model interface side, new interfaces have been developed to improve the way users interact with model files and parameters to set up and run models. Our work and research in this paper mainly involves the identification and implementation of many of these performance and GUI improvements to the BPM model.

There were two main objectives identified in improving the NYBPM model:
- Optimizing model performance by updating the model to the latest version of the travel model software, utilizing multi-core computers and distributed computer environments, and streamlining and optimizing the model code
- Improving the usability and applicability of the model by:

- Designing an updated graphical user interface
- Creating automated and interactive reports
- Creating a dynamic web application

Each of these improvements is described in the sections below.

Optimizing Model Performance

One main goal was to reduce the running times of the BPM model. This was achieved by utilizing the following techniques:

- Updating the BPM model from an older version of TransCAD to Version 5.0, taking advantage of the internal speedups and multithreaded operations inherent to the new version.
- Taking advantage of hardware enhancements that house multiple processors and allow more memory on one computer.
- Distributing the model run over several computers using distributed processing.
- Optimizing and streamlining the model code

Effects of New Version of Travel Model Software

The BPM model was originally developed in TransCAD Version 4.5 and recently updated to Version 4.8 by PB Americas. In updating to Version 5.0, model performance is enhanced by the following features:

- Certain model procedures such as highway and transit assignments and skimming are multi-threaded. These procedures automatically detect the number of computing cores available on a computer and will distribute the procedure calculations simultaneously across the cores. In one extreme example, a highway assignment on a computer with 8 physical cores and hyperthreading ran twelve times faster than on a single core. Moving the model to the newest version thus enabled dramatically reduced running times for traffic assignments.
- Certain model procedures, such as transit skimming and highway assignment were optimized in the new version, making them faster to run, or faster to converge than the old version.

Effects of Multi-threading and Increased Computer RAM Memory

The focus on model performance improvements did not rely exclusively on the model software procedures. Key programs in the BPM such as the Mode-Destination Choice (MDC) model were also optimized. Specific portions of this program were multi-threaded so that they can run on multiple computer cores simultaneously. In addition, multiple instances of this program were called simultaneously, as MDC is called once per trip purpose. On a four-core machine, the MDC is invoked in three instances. Tests determined that any more than three simultaneous instances of this program leads to increased run times due to overhead expenses.

Effects of Distributed Computing

There were two main components of the BPM model that were distributed: the Mode Destination Choice (MDC) models and the Highway Assignment and Skimming models. Collectively, these models account for over 90% of the overall processing time of the BPM model when run on a single computer. The manner in which each of these models is distributed, and comparative model run times, are described in this section. All models were run using the following computer configuration:

Processor: Intel Xeon 4-Core Processor, 3.0GHZ,
RAM Memory: 3 GB
Hard Drive: ~300GB 10,000 RPM in Raid 0 configuration
OS: Windows XP Professional, 32-bit

Setting up Distributed Computers

The setup of the distributed computers is performed within a dialog box in the travel model software. In this dialog box, the remote computers are identified and added to the list of available computers:
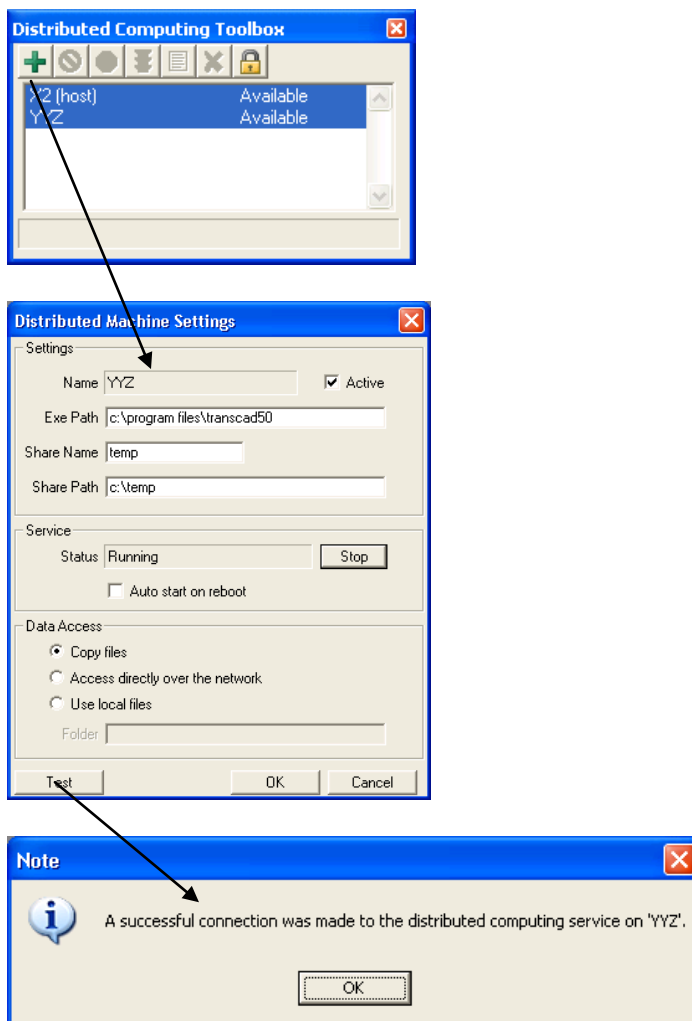
Distributing the MDC Model

The MDC model roughly accounts for about 67% of the overall processing time of the highway portion of the BPM model. There were several distributing and multi-core techniques that were implemented to reduce the processing time for MDC. There are 8 MDC trip purposes that are invoked in the model. The MDC trip purposes, and their individual running times when run sequentially, are listed in the table below:

Trip Purpose/Process	*Running Time*
Low income work	8 minutes
Medium income work	29 minutes
High income work	24 minutes
School	31 minutes
University	9 minutes
Maintenance	59 minutes
Discretionary	35 minutes
At work	9 minutes
Total	**204 minutes**

The model detects the number of computing processor cores available on the main computer running the model. The model also detects the number of distributed computers available and the number of cores available for each distributed computer. Based on this information, the trip purposes are divided up so that they can run simultaneously. Since some trip purposes take longer than others to run, the purposes are "load balanced" such that each computer and/or processor achieves a balanced processing time.

It was observed that as more trip purposes were simultaneously assigned to the same computer, the run times for each purpose would increase, due to memory sharing and file swapping issues. Adding more computers optimizes run times by reducing the simultaneous runs for each computer.

The effects of distributed computing are limited for the MDC model since one trip purpose (Maintenance) by itself takes the bulk of the run time. Due to this constraint, the model limits the number of distributed computers that can be used to 3, since extra computers would not be useful.

The model was run in a 2-computer and 3-computer setup. The resulting running times, compared with single computer configurations, are presented below:

Computer/Model Setup	*MDC Running Time*
1 Computer, All models run sequentially	204 minutes
1 Computer, MDC runs on two cores (4.5, 4.8)	154 minutes
1 Computer, MDC runs on three cores (5.0*)	95 minutes
2 Computers, MDC runs on up to three cores	77 minutes
3 Computers, MDC runs on up to three cores	75 minutes

* In addition to the time savings achieved from the extra processor, extra savings were achieved by optimizing the MDC program in TransCAD 5.0.

As can be seen, there is a significant time savings for the MDC model when distributing between two computers. Adding a third computer adds only a marginal time savings.

Distributing the Highway Assignment and Skimming Models

The highway assignment and skimming models account for about 26% of the overall processing time of the highway portion BPM model. For both highway assignment and skimming, TransCAD 5.0 already exploits the available multi-processors on a computer. In the distributed environment, the different assignment and skim models are run in parallel, and are graphically depicted in the flow chart interface:
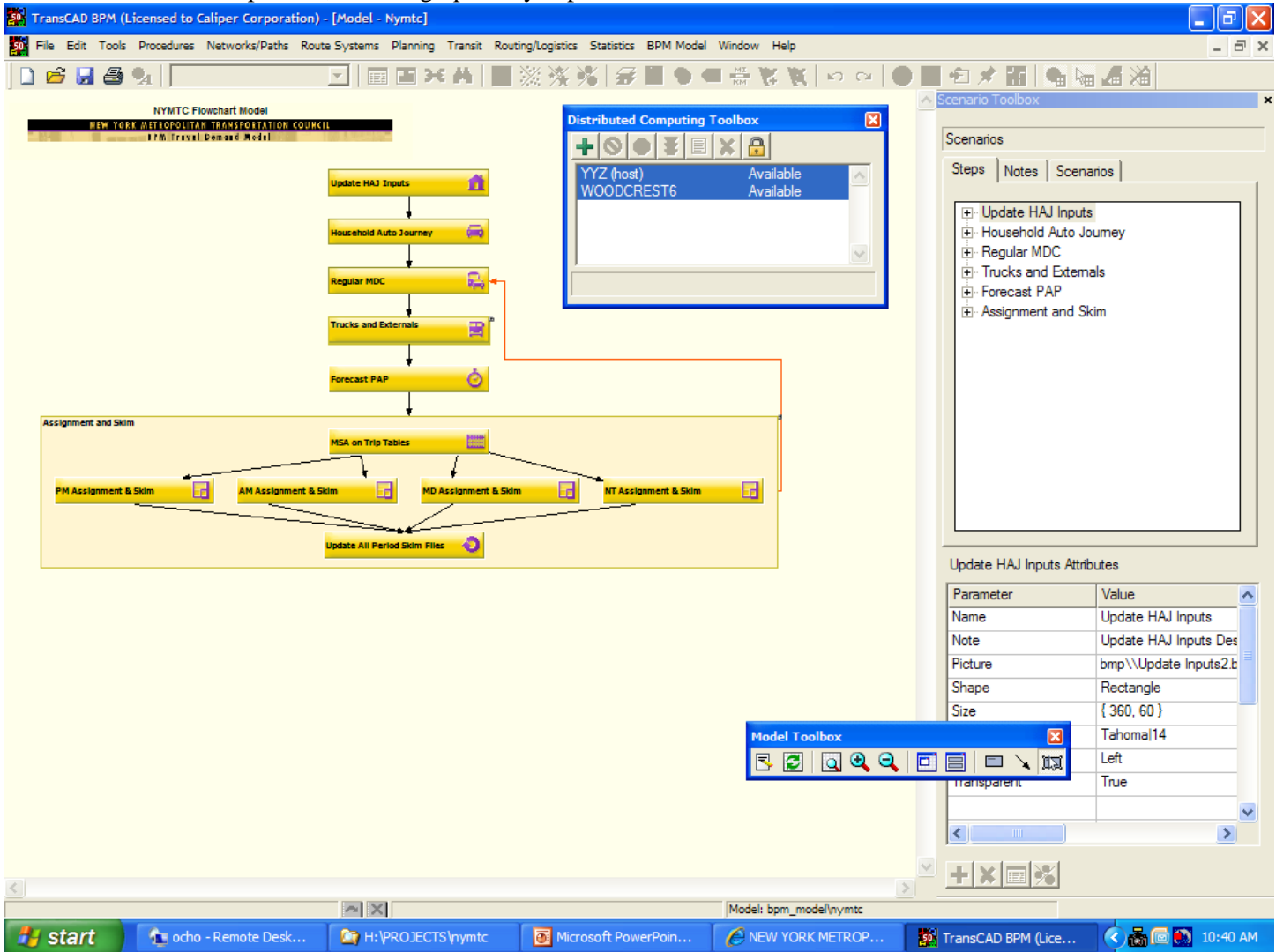
In the flow chart, AM, MD, PM and NT assignments are all set up and run in parallel to each other. When flow chart nodes are depicted in this fashion, the models will automatically be invoked in distributed mode.

When run sequentially, the individual time period assignments and skim run times are depicted below. These run times were computed for the final loop of a 3 feedback loop model:

Time Period	Iterations Run	Run Time
AM Assignment and Skim	35	34 minutes
MD Assignment and Skim	12	13 minutes
PM Assignment and Skim	35	33 minutes
NT Assignment and Skim	5	7 minutes
Total	**n/a**	**87 minutes**

The total run times for both a single computer and a 2 and 3 computer setup are shown for each feedback loop:

	Assignments and Skims Run Times		
Feedback Loop	**Single Computer**	**Two Computers**	**Three Computers**
Loop 1	40 minutes	34 minutes	34 minutes
Loop 2	41 minutes	33 minutes	33 minutes
Loop 3 (Final)	87 minutes	61 minutes	51 minutes
Total	**156 minutes**	**128 minutes**	**118 minutes**

To optimize the run times, the longer assignment processes, AM and PM are paired together to run simultaneously as much as possible. In the first two feedback loops, only AM and MD assignments are run, so there are no additional performance benefits of the third computer until the final loop when all assignments are run.

Effects of Other Code Optimization

Additional code optimization was performed to other model components. These included:
- Certain external programs being converted to GISDK code and being optimized. In many cases, these external programs would export matrices, read and process the matrices, and import them back into the model software. Eliminating the matrix imports and exports and reading the matrices natively dramatically reduced their running times.
- Certain inefficient functions and looping structures from within the model code were identified and optimized.

The following table lists the overall effects of the above model optimizations without distributed computing contributions. The run times trace only the highway-related models.

|	Model Run Times in TransCAD 5.0	Model Run Times in TransCAD 4.8	Time Savings	Percent Savings
Step	Time (minutes)	Time (minutes)	Time (minutes)	
Loop 1				
HAJ Inputs	3.93	3.93	0.00	0%
HAJ	3.18	2.82	-0.36	-13%
MDC	96.06	153.00	56.94	37%
Forecast External Trips	0.02	0.02	0.00	0%
Forecast Impedance	0.45	n/a	n/a	n/a
Forecast Truck Trips	2.16	18.62	16.46	88%
Forecast Van Trips	1.08	4.95	3.87	78%
PAP	5.23	12.87	7.64	59%
External Replace	0.11	n/a	n/a	n/a
Assignment and Skims	24.97	27.50	2.53	9%
Loop 2			0.00	
MDC	96.66	154.00	57.34	37%
Forecast External Trips	0.02	0.02	0.00	0%
Forecast Impedance	0.67	n/a	n/a	n/a
Forecast Truck Trips	2.23	18.60	16.37	88%
Forecast Van Trips	1.07	4.93	3.86	78%
PAP	5.24	12.91	7.67	59%
External Replace	0.10	n/a	n/a	n/a
Assignment and Skims	28.78	34.00	5.22	15%
Loop 3		n/a	n/a	n/a
MDC	97.71	153.00	55.29	36%
Forecast External Trips	0.01	0.02	0.01	50%
Forecast Impedance	0.68	n/a	n/a	n/a
Forecast Truck Trips	2.34	18.63	16.29	87%
Forecast Van Trips	1.15	4.96	3.81	77%
PAP	5.22	12.86	7.64	59%
External Replace	0.10	n/a	n/a	n/a
Assignment and Skims	56.53	56.00		0%
Total Run Time	**435.7**	**693.64**	**257.94**	**37%**
	(7 hours 15 min)	**(11 hours 34 min)**	**(4 hours, 18 mins)**	

With a distributed solution of two computers, model run times are reduced by about 21%.

Improving the Usability and Applicability of the BPM Model

The main objective in improving the usability and applicability of the BPM model was to improve the overall user experience with the model. The intent was to make it easier for the user to interact with the model: to define and view input and output files and parameters, to analyze and report model outputs, and to distribute model outputs to the wider modeling community.

Designing an Updated Graphical User Interface

The current NYBPM model is controlled by a controller program called Central, which launches both GISDK macros in the travel model software and external programs. Snapshots of the Central interface from the NYBPM User's Guide are shown below:
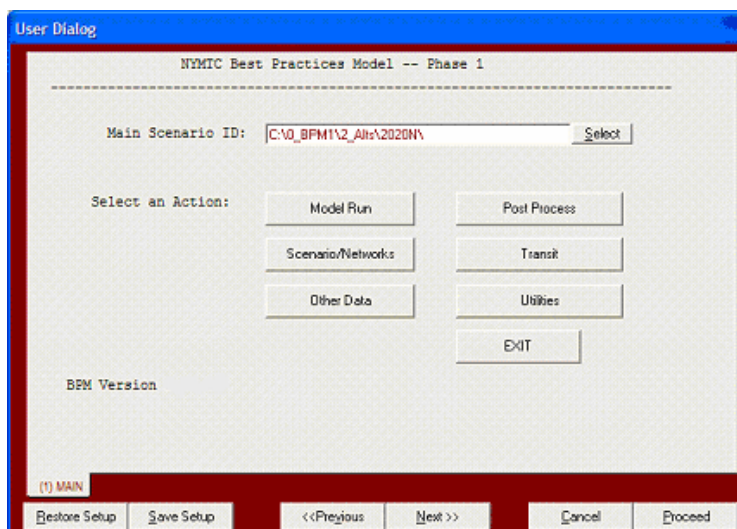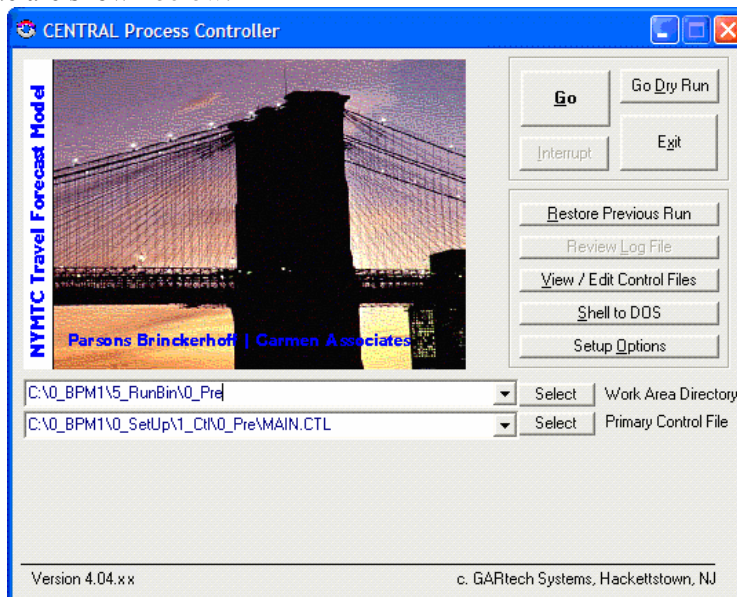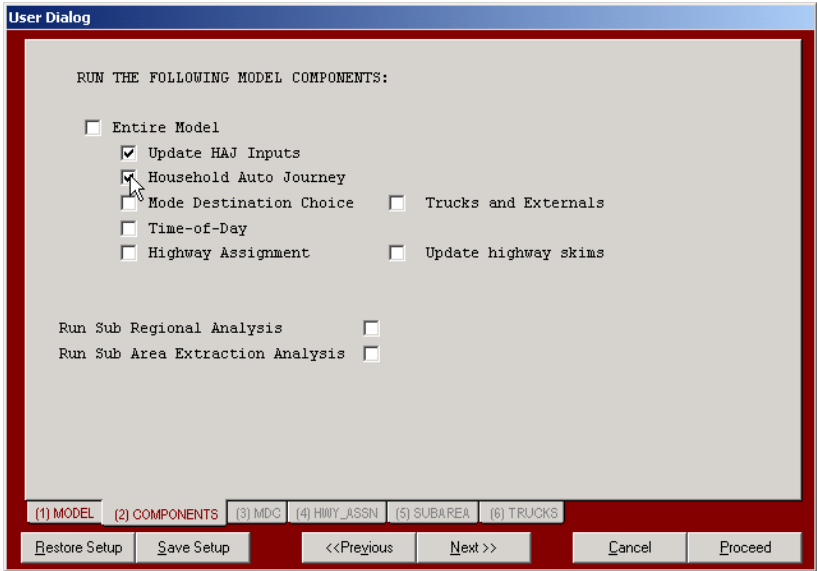
This interface is replaced with an easier to use and understand flow chart interface, which controls both the flow of the model and its input and output files and parameters. Each model step is represented by a flow chart node. Each flow chart node can drill down into sub-steps, and each flow chart node can contain specified input and output files and parameters. Clicking on the input or output file will open the file. Clicking on the node will allow the user to run the model code associated with the step or substep. An example of the interface is shown in the next graphic:
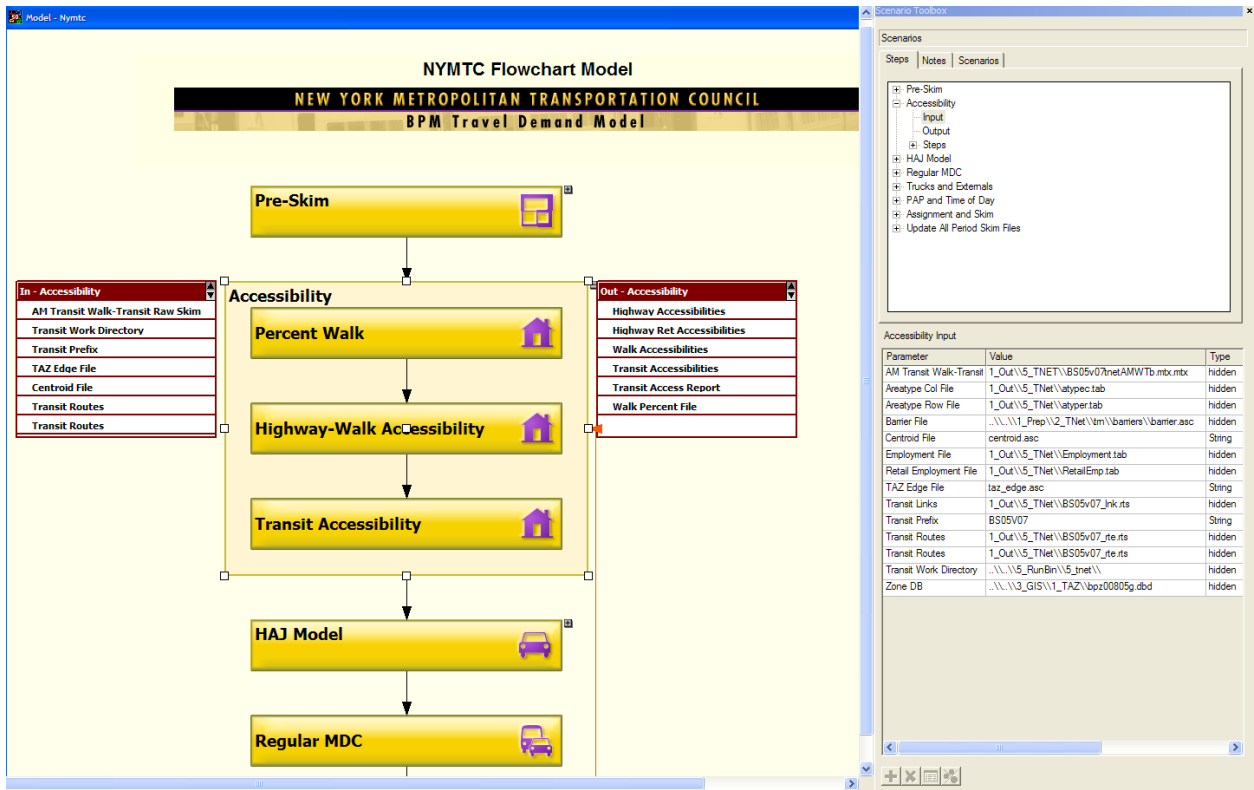
Creating Reports, and Designing a Web Application

To further enhance the usability of the model, both automated and interactive reports were added to the model stream and interface. Automated post-processing reports displaying demographic, origin-destination, and assignment output were designed using Crystal Reports. Reports are exported into PDF and Excel format. Report examples are shown below:

NEW YORK METROPOLITAN TRANSPORTATION COUNCIL

Average Trip Lengths and Times for 2005T

Trip Length Distribution by County, Mode and Time Period

|				AM	MD	NT	PM	Total
ALL	ALL	SOV	Time	20.25	16.64	17.09	18.38	18.09
			Distance	11.48	9.20	9.74	10.07	10.12
		HOV2	Time	17.68	16.32	15.97	17.22	16.80
			Distance	9.72	9.18	8.89	9.19	9.25
		HOV3	Time	16.04	15.65	16.12	17.25	16.26
			Distance	8.78	8.92	9.35	9.55	9.15
		Ext	Time	73.67	71.53	78.03	74.26	74.37
			Distance	52.61	51.42	59.41	53.09	54.13
		Comm	Time	20.03	19.13	18.87	20.10	19.53
			Distance	12.04	11.90	11.89	12.04	11.97
		Truck	Time	22.70	24.31	36.40	26.57	27.49
			Distance	15.40	17.10	27.41	18.66	19.64
		Total	**Time**	**28.40**	**27.26**	**30.41**	**28.96**	**28.76**
			Distance	**18.34**	**17.95**	**21.11**	**18.77**	**19.04**
	Total		**Time**	**28.40**	**27.26**	**30.41**	**28.96**	**28.76**
			Distance	**18.34**	**17.95**	**21.11**	**18.77**	**19.04**
CT	Fairfield	SOV	Time	16.58	12.99	13.37	14.52	14.37
			Distance	8.88	6.85	7.22	7.49	7.61
		HOV2	Time	14.68	13.58	12.46	13.37	13.52
			Distance	7.76	7.54	6.74	6.90	7.23
		HOV3	Time	12.95	12.59	12.15	12.90	12.64
			Distance	6.46	6.73	6.47	6.55	6.55
		Ext	Time	75.98	72.71	74.41	76.09	74.80
			Distance	51.27	48.89	53.36	49.25	50.69
		Comm	Time	42.01	39.81	39.78	40.79	40.60
			Distance	28.82	29.02	29.28	28.51	28.91

5/29/2008

Page 20 of 54

NEW YORK METROPOLITAN TRANSPORTATION COUNCIL

Demographic Data: Population for 2005T

| Name	HHPOP	GQPOP	GQPOPINS	GQPOPSTR	GQPOPOTH	HHNUM	HHSIZE
CT							
Fairfield	897,368	18,747	5,321	896	12,534	338,676	2.65
New Haven	820,697	29,677	12,474	190	17,021	331,247	2.47
State Total	**1,718,065**	**48,424**	**17,795**	**1,086**	**29,555**	**669,923**	**2.57**
NJ							
Bergen	895,606	9,768	5,738	18	4,010	338,928	2.72
Essex	790,472	18,971	9,480	1,035	8,464	289,469	2.76
Hudson	613,664	8,462	4,794	385	3,286	235,847	2.63
Hunterdon	125,994	4,012	3,711	0	302	46,749	2.61
Mercer	347,518	19,028	6,630	9	12,388	131,966	2.61
Middlesex	761,398	24,229	9,705	17	14,502	277,190	2.71
Monmouth	639,171	12,786	7,012	13	5,760	237,963	2.56
Morris	484,919	9,836	5,204	11	4,624	178,671	2.73
Ocean	545,383	7,554	5,680	5	1,867	217,379	2.41
Passaic	480,935	17,818	10,025	228	7,565	165,890	2.78
Somerset	309,159	5,740	5,008	63	672	115,485	2.62
Sussex	150,067	1,935	1,762	0	173	53,734	2.71
Union	524,065	6,731	4,027	68	2,638	188,965	2.68
Warren	109,815	1,632	1,106	1	523	42,176	2.52
State Total	**6,778,166**	**148,502**	**79,882**	**1,853**	**66,774**	**2,520,412**	**2.67**
NY							
Bronx	1,321,018	46,455	26,497	3,321	16,631	472,967	2.75
Dutchess	280,096	19,956	8,895	0	11,060	107,756	2.60
Kings	2,477,462	37,838	16,219	3,335	18,284	896,658	2.74
Nassau	1,334,247	18,765	8,838	202	9,725	452,500	2.98
New York	1,521,804	61,499	14,024	6,599	40,878	759,293	1.99
Orange	359,535	16,085	7,619	28	8,436	127,887	2.80
Putnam	100,344	1,570	579	0	990	35,127	2.84
Queens	2,243,398	29,279	18,897	482	9,898	795,896	2.85
Richmond	461,414	9,476	6,099	49	3,328	165,402	2.70
Rockland	289,726	9,181	4,938	0	4,242	97,440	2.93

5/29/2008   Page 1 of 54

Interactive reports let the user display maps of traffic flows, congestion, select link flows, and demographic information. The reports also enable the user to query shortest paths by various modes and display travel time maps. Interactive report examples are shown below:
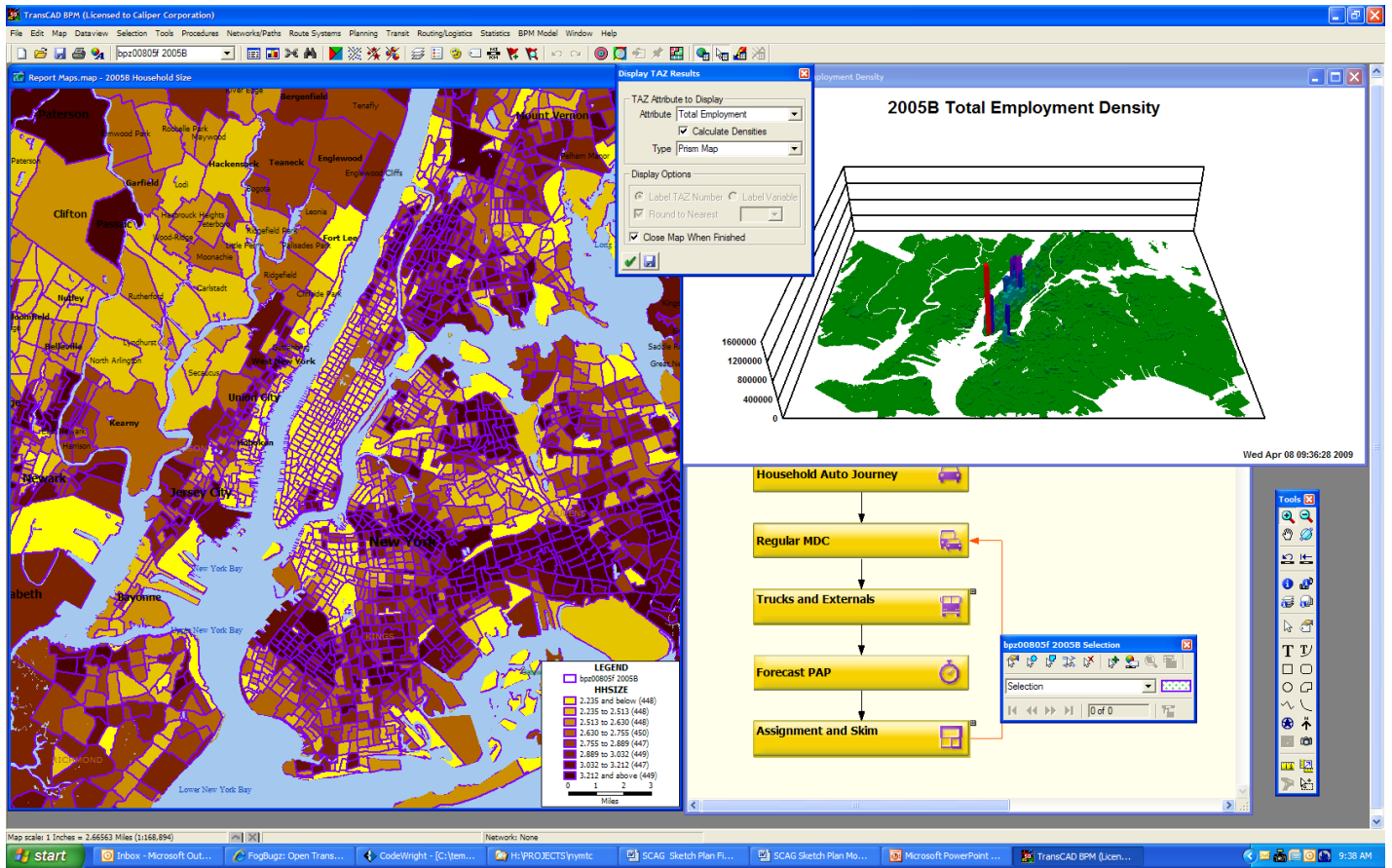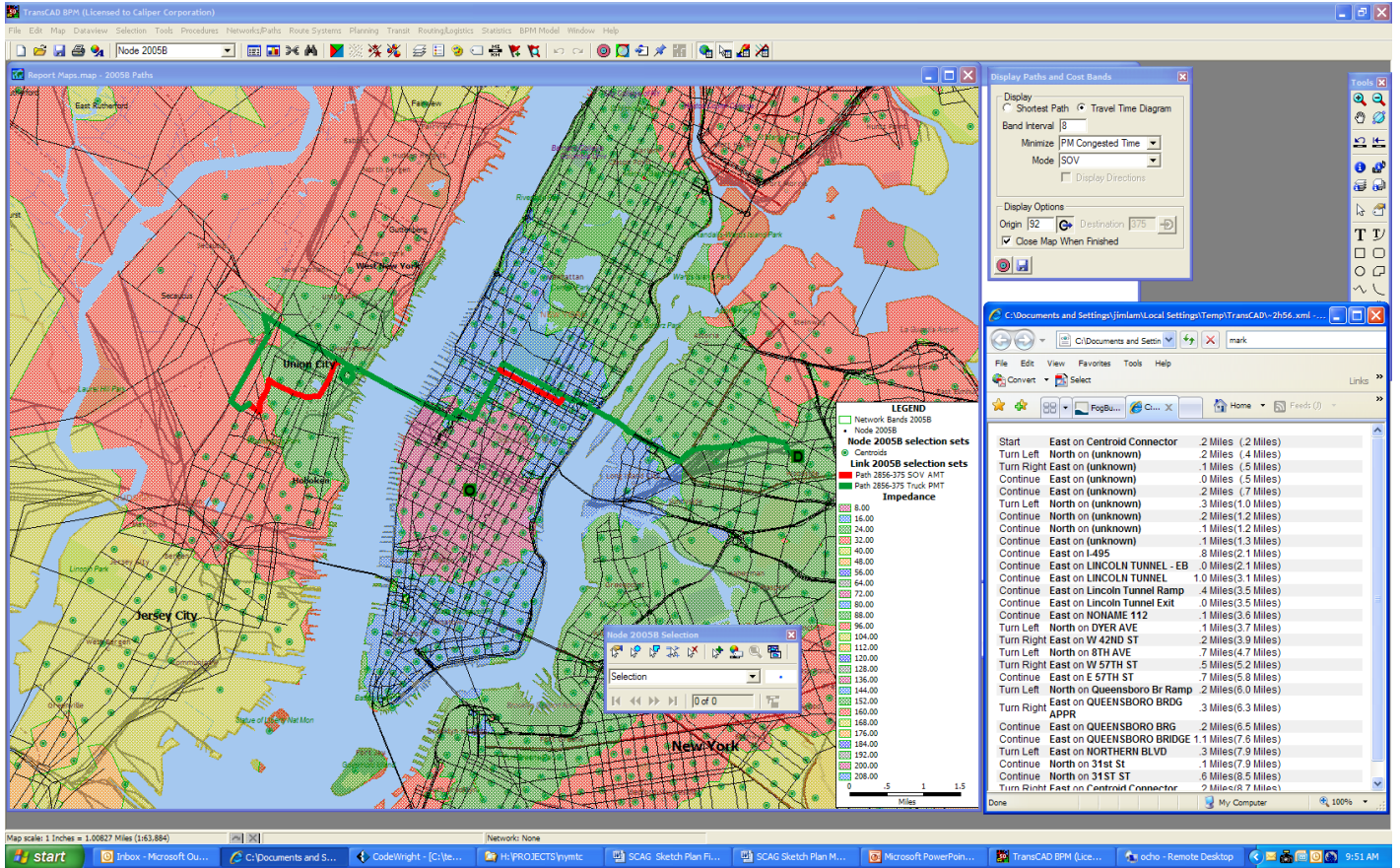
A dynamic web application was created that lets the public interactive maps and reports over a web browser. The web application gives the user limited GIS functionalities such as zooming, panning, finding locations, displaying and hiding map layers, and querying and displaying map feature attribute information. The reports and maps generated from the BPM model can be easily moved to the web application and organized through a series of web folders accessible to the browser user. Snapshots of the web application are shown below:
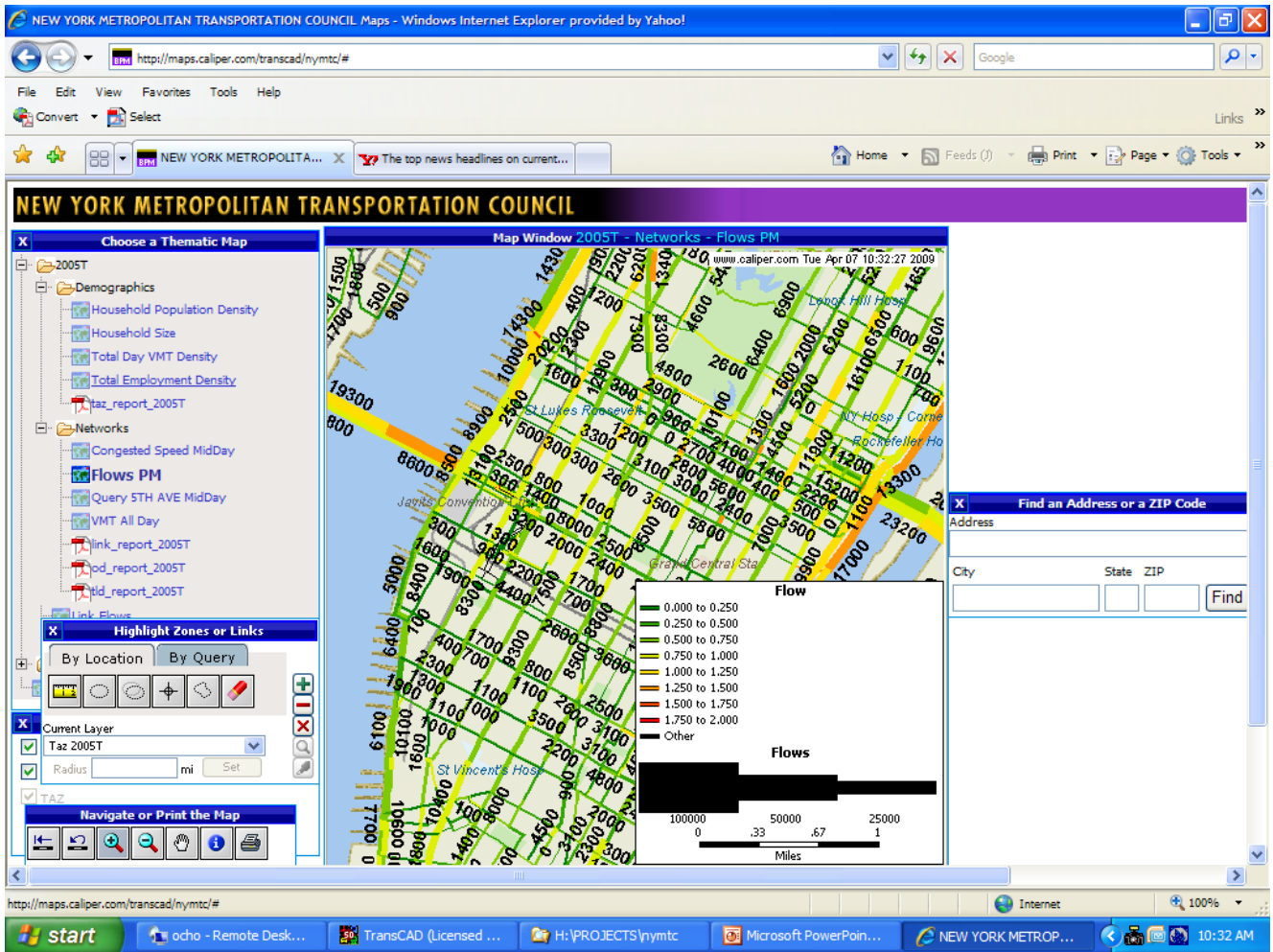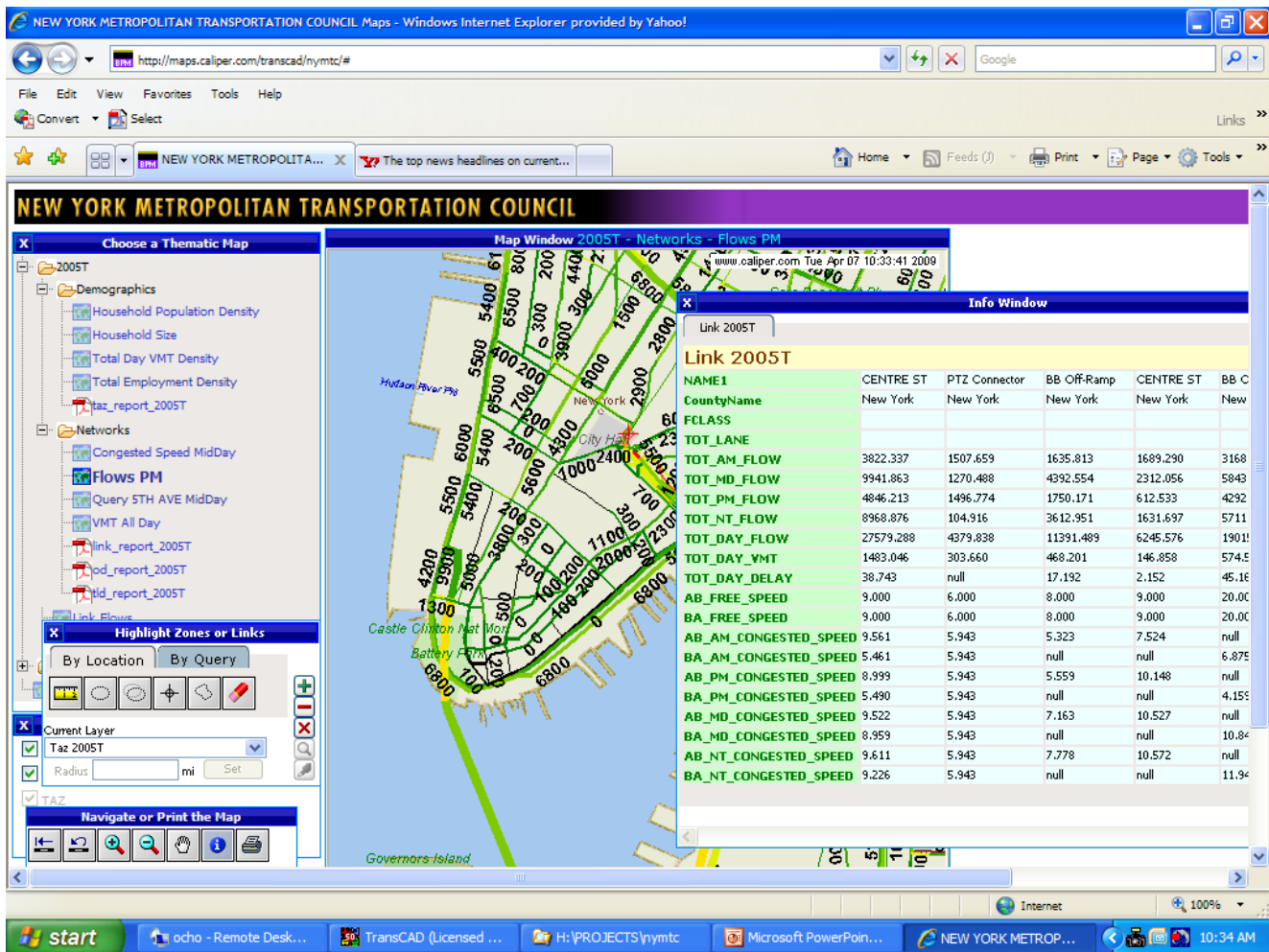
References

1. New York Best Practice Model (BPM) For Regional Travel Demand Forecasting: NYBPM User Documentation: Parsons Brinckerhoff/PB Consult, May, 2005.

2. Travel Demand Modeling with TransCAD 5.0 Users Guide: Caliper Corporation, Sept. 2007