



## Memo

38 Chauncy Street  
Suite 200  
Boston, MA 02111

T 857-453-5450  
F 857-453-5451

[www.transystems.com](http://www.transystems.com)

**To:** Steve Dudley, SCRCOG  
Jim Rode, SCRCOG

**From:** Jim Wensley

**Date:** June 27, 2012

**Subject:** Regional Transit Data Acquisition

TranSystems has completed the task to assemble the transit data needed for future enhancements to the transit component of the regional travel demand forecasting model. Transit ridership and service data has been assembled for all public transit providers operating in the SCRCOG region. Base year (2010) level of service data has been assembled so that it can be readily used as an input to the model. Transit ridership data has also been assembled and summarized so that it can be used to help calibrate and validate the mode share and transit assignment portions of the model.

The project included several tasks:

**Task 1: Determine Data Needs** - TranSystems identified the data needed regarding bus services provided in the region by CTTransit New Haven, as well as CTTransit Wallingford, CTTransit Meriden, Milford Transit and the multiple-operator Coastal Link service. Data on routes entering the region and operated by Middletown Transit, Estuary Transit and the Hartford express carriers was also desired, as was data on rail services. The needs identified included level of service data (peak and off-peak headways) and current fare information plus ridership data, including route-level daily ridership for bus services and station level daily ridership for rail services. Bus ridership at the TAZ level was also needed, as well as route-to-route bus transfer data. The data to be assembled, sources of the data, and the data output format was reviewed with SCRCOG staff before the data was assembled.

**Task 2: Data Assembly** - The data was assembled by contacting the individual bus and rail operators, as well as Connecticut DOT. Schedule and fare data was generally obtained from agency web pages since service levels have not changed on most services since 2010. Route-level ridership and stop-level ridership was collected from CTTransit New Haven and CTTransit Wallingford. **Only Route-level ridership has been obtained from the Coastal Link but stop-level data is still expected.** Only Route-level ridership was available from CTTransit Meriden and Milford Transit. Farebox data from which bus route transfer data could be estimated was obtained from CTTransit New Haven for a full week in April 2010. Transfers counts were not available from other bus operators. GIS files of bus stop locations were obtained from CTTransit and a GIS file of TAZs was obtained from SCRCOG. Rail data was gathered from ConnDOT for Shore Line East, Metro-North New Haven Line, and Amtrak. Wherever possible, data was assembled for year 2010.

**Task 3: TAZ Estimates** – Using the GIS files of bus stops and TAZs TranSystems was able to identify in which TAZ each bus stop is located. For the majority of stops, the stops had to be assigned to two or more TAZs, since many stops lay on the boundary between TAZs, and the ridership had to be split among them. The process for doing this is detailed in the attached documentation. The ridership obtained from those operators who could provide stop-level data was adjusted to represent a full weekday of ridership. The stop-level ridership was then summed by TAZ to obtain daily transit ridership by TAZ.

**Task 4: Transfer Data** - A route-to-route matrix of transfers for CTTtransit's New Haven Division was created through manipulation of detailed farebox data. One week of cash boardings data and one week of pass boardings data was used to develop a daily transfer matrix that can be used for future model validation. The process followed is detailed in the attached documentation.

**Task 5: Documentation** - The data assembled and prepared for use in the regional model is provided in three spreadsheet files, consistent with the format reviewed with SCRCOG staff in Task 1. The data assembled, the procedures used to prepare the data in its final format, and a description of the final output format is included in the documentation that follows.

## Route-Level Data

### Data Sources:

- Online schedules for CTTransit Hartford, New Haven, Wallingford and Meriden (headways and fares).
- Online schedule for Milford Transit, Estuary Transit, Middletown Area Transit, and Greater Bridgeport Transit (headways and fares)
- Online schedules for Shore Line East, Metro-North New Haven Line and Amtrak (headways and fares).
- CTTransit “Route Summary Productivity 2010”. Includes estimates of 2010 weekday passengers by route for CTTransit New Haven factored from a sample of trips.
- Tabulations of farebox data from October 2010 provided by CTTransit Meriden (CTTransit Meriden ridership).
- Nelson Nygaard “Wallingford Bus Ridership and Routing Study” Wallingford ride check data summary spreadsheet. (CTTransit Wallingford ridership from September 2011)
- AECOM “Coastal Corridor Bus Study, Recommended Service Plan” January 2012 (2011 Coastal Link ridership).
- E-mail from Milford Transit with 2009/2010 average weekday ridership by route for Milford Transit.
- Shore Line East, Metro-North and Amtrak station boardings through e-mail from CTDOT
  - Shore Line East information from 2010 Daily Rider monthly reports
  - Metro-North information from 2010 TOTAL WEEKDAY BOARDINGS BY STATION
  - Amtrak information from Amtrak Fact Sheet, Fiscal Year 2010 State of Connecticut
- Ridership data for the following routes was not available:
  - Meriden-Hartford Express (Route 19) operated by Peter Pan
  - New Haven-Hartford Express (Route 20) operated by DATTCO
  - Middletown Area Transit M-Link
  - Estuary Transit Shoreline Shuttle

### Processing:

- Current (2012) headways were used for all services. CTTransit indicated that the only bus route that has changed service levels since 2010 is the S Route, which was operated at a lower frequency of service in 2010 before being taken over by CTTransit in 2011.
- Ridership is from 2010 for CTTransit New Haven and Meriden, Milford Transit and the rail services. Ridership from 2011 was used for CTTransit Wallingford and the Coastal Link.

### Output Table:

- Data is contained in “SCRCOG Model Data.xlsx”.
- Average weekday ridership, peak and off-peak (midday) headways, and base fare for each bus route operating in the SCRCOG region are listed in the table. Data is for 2010 where possible; however some data are more recent.

- For rail services, average weekday station boardings are given along with line headways and fares to New Haven (for Shore Line East and Amtrak from Meriden and Wallingford) and fares to New York on Metro-North from New Haven.

## Transit Ridership by TAZ

### Data Sources:

- CTTTransit “NH Weekday Route Profile 2010” text file. Consists of a text file version of CTTTransit’s Route Profile Report. Report includes trips checked from 1/1/10 through 12/31/10. Most scheduled trips were checked once during the year. The report shows total ons and offs recorded by stop by time period on each route and direction for a typical day. Stops were identified only by a text description; there was no numeric ID or alphanumeric stop code.
- Nelson Nygaard “Wallingford Bus Ridership and Routing Study” Wallingford ride check data summary spreadsheet. Contained typical weekday ridership by stop for CTTTransit Wallingford collected in September 2011.
- **AECOM Coastal Link data files**
- Stop level data was not available for CTTTransit Meriden or Milford Transit
- CTTTransit “New Haven Bus Stops March 18, 2012” GIS shape file. Contained geocoded locations of all bus stops in the CTTTransit New Haven Division (including CTTTransit Wallingford and Meriden stops) as of March 18, 2012. CTTTransit does not maintain historical GIS records of bus stops so there was no way to get a 2010 version. The file contained text descriptions and a numeric ID for each stop.
- SCRCOG TAZ file. Geographic file of all TAZs in the SCRCOG region in TransCAD format.

### Processing:

- The route profile text file report was imported into a spreadsheet and parsed to extract the number of scheduled trips and the number of observed trips by route, direction and time period. In most cases 100% of scheduled trips were sampled. With the exception of the M Route, all sample sizes were at least 67% (M Route samples ranged from 17% to 88%). The number of complete samples and smallest sample sizes by period are shown below:

Period	Time	Route/Dirs. Operated	Route/Dirs with 100% Sampled	Smallest Sample (excl. M)
AM Peak	5:30 – 8:30 AM	47	35	67%
Mid Peak	8:30 AM – 2:30 PM	43	35	83%
PM Peak	2:30 – 5:30 PM	46	32	78%
Off Peak	5:30 PM – 5:30 AM	43	36	83%

- The text file was imported into another spreadsheet and parsed to extract the number of recorded ons and offs by stop for each route, direction, and time period.
- In cases (i.e. route/direction/period) where fewer than 100% of trips were sampled, the observed ons and offs were factored up to represent a full typical day using the ratio of trips scheduled to trips observed.

- The text descriptions of stops in the profile data and in the GIS bus stop file were different length fields so the descriptions in the GIS file had to be shortened and have some spaces removed in order to match the route profile descriptions. The edited GIS file stop descriptions were then used to lookup and attach the numeric stop ID to each stop in the route profile data.
- Because the GIS file was more recent than the route profile, there were 59 stops in the route profiles that were not in the GIS file; however only 18 of these had non-zero ridership data. These were assigned an unused ID number and were addressed separately later.
- CTTTransit Wallingford stop level ridership and Coastal Link stop level ridership were appended to the CTTTransit New Haven data.
- The GIS bus stop file was converted to TransCAD and overlaid with the TAZs. (Several versions of the bus stop file had to be obtained from CTTTransit before receiving once that could be successfully overlaid.)
- Because it was noted that many bus stops lie along the streets that form the boundaries between TAZs, TransCAD was used to determine the number of TAZs that lie at least partially within 200 feet of each bus stop. This distance was chosen after examining the number of stops serving one, two, or three or more TAZs for several longer and shorter distances. This distance was found to identify stops lying along the dividing street between two TAZs even if the stop wasn't perfectly coded along the line. It also appears to identify stops serving three or more TAZs only if the stop is really at the junction of three or more zones. The table below shows the number of stops identified to serve each number of TAZs.

Number of TAZs	Number of Stops
0	300
1	1,457
2	1,975
3	492
4	124

- Stops serving zero TAZs were confirmed to be stops lying outside the SCRCOG region on the J and F routes.
- TransCAD was used to automatically tag each stop with the ID number(s) of up to two TAZs (that were within 200 feet). The stops serving three or more TAZs had to be checked manually to add the third and fourth TAZ ID.
- The 19 stops not in the GIS file were manually assigned to one or more TAZs based on the stop description and their order in the stoplist for the route.
- Stops on the Milford portion of the Coastal Link were manually assigned to one or more TAZs based on the stop description.
- The number and list of TAZs served by each bus stop was joined to the bus stop ridership from the route profiles using the bus stop ID number.
- The ridership (ons and offs by time period) for each stop was then split evenly between the TAZs served by that stop. While more sophisticated methods of allocation using population data and size

of the catchment area within each zone could be envisioned, these would require substantial geographic analysis involving over 2,500 stops. Splitting the ridership evenly seems to be a reasonable approximation, especially given that most cases with multiple TAZs were simply stops on a straight line dividing two zones.

- Matrices were then produced of ridership by TAZ and route, and ridership by TAZ and time period.

#### **Output Tables:**

- Several matrices of ridership by TAZ and route were produced and are included in “[SCRCOG Transit Ridership by TAZ.xlsx](#)”. Rows represent TAZs listed by their ID numbers. Columns represent the CTTransit New Haven and Wallingford routes. Data represents a typical weekday in 2010. Each matrix is on a separate tab:
  - **Daily Route Riders** – Daily ridership (total ons; offs can be assumed to be similar)
  - **AM Route Ons** – AM peak (5:30-8:30) boardings
  - **AM Route Offs** – AM peak (5:30-8:30) alightings
- A fourth tab (**Period Riders**) shows ridership (both ons and offs) by TAZ and time period. TAZ data in this table is summed over all routes.

## Transfer Data

### Data Sources:

- CTTTransit “New Haven Transaction Details – Transfer” data from fareboxes for April 9-13, 2010. This text file contains a record for each cash transfer issued and each transfer received. Records for transfers received contain the route codes for both the issuing and receiving route.
- CTTTransit “New Haven Transaction Details – Period Pass” data from fareboxes for April 9-13, 2010. Spreadsheet contains a record for each boarding using a time-based pass (e.g. a monthly pass).
- Similar data was not available for CTransit Meriden or Wallingford, or for Milford Transit.

### Processing:

- Only the “Transfer Received” records were used for cash-based transfers. The two routes codes (issuing and receiving) on these records were checked for validity. Incorrect codes were re-coded to “999”. Valid codes included CTTTransit New Haven (issuing and receiving) and CTTTransit Meriden (issuing only) routes. Transfers from CTTTransit Wallingford did not appear in the data. All records were tabulated into a route-to-route matrix for cash transfers for one week. The data for transfers issued in Meriden were duplicated in the opposite direction to account for transfers issued on CTTTransit New Haven and likely received on CTTTransit Meriden.
- Period pass data included records for every boarding using a pass. This data was sorted by pass ID, date and time. Two consecutive boardings were considered a “pass transfer” if they met the following four criteria:
  1. They were made using the same pass ID
  2. They were made on the same date
  3. They were made within 90 minutes of each other (the valid window for “cash transfers”)
  4. They were on different routes
- Adjustments were made in cases where the direction for the originating trip was “outbound” or the receiving trip was “inbound”. These are believed to often be cases where the driver switched the route entered in the farebox to the opposite half of the same route (e.g. switching from B Whalley to B Congress Avenue) either before the initial boarding or after the transfer boarding. In these cases the route was swapped to the other “half-route”. (The methodology was copied from the methodology used to create a similar transfer matrix for the 2008 *Implementation of the Regional Transit Study* done by TranSystems.) All records were then tabulated into a route-to-route matrix for “pass transfers” for one week.
- The Cash and pass transfer matrices were added together.
- Cash transfers to and from the Meriden routes were factored up using the systemwide ratio of total transfers to cash transfers (1.39) since the pass transfer data did not include Meriden routes.
- The figures in the combined weekly matrix were divided by five to yield an average daily transfer estimate for the week analyzed.



**Output Table:**

- A single weekday transfer matrix is provided in "[SCRCOG Region Transfers.xlsx](#)" representing an average weekday in a typical week in April 2010. The rows represent the originating routes and the columns represent the receiving routes.