DATE: **Wednesday, May 16, 2007**
FROM: **Rhett Fussell, PB**

RE: **OD Matrix Estimation**

We wanted to be able to see if we could use the counts in Oahu to develop an OD matrix that would provide us with an estimate of the best zone to zone OD flows based on using the counts to develop that flow. Therefore using the Transcad Multi-class OD estimator seemed to be a logical choice. The estimator allowed us to use the OD matrices we were getting out of the model as our initial seed matrices. The OD pairs from the model are assigned to the network using UE (and the conical delay function), then re-factored based on the counts on the network links. This process was done for 30 iterations until it reached the .001 closure criteria. A separate memo addresses the detailed steps of performing these steps, therefore we will only discuss the current outcomes from the OD estimation.

We tried several different scenarios(different uses of the matrices) of running the OD estimation in order to try to determine the best method for replicating the counts/flows on the Oahu network. For all of these scenarios the 2+ links(HOV) were not allowed to be used in the assignment since we did not have counts for those links. The capacity values on the parallerl non-hov links were not adjusted either to reflect the additional capacity lost by not including those links. All of these scenarios are for the AM peak period. We used the 2 hr AM count and corresponding capacity for all the scenarios.

A brief description of the methods are shown:

A) Assigned each of the 6 matrices onto the network using the multi-class estimation. However, since the estimator needs a count for each type of class (or matrix) this resulted in overestimation of the OD pairs because each matrix was attempting to match the same count on the links. Basically each matrix tried to match the count on the link. Therefore, this method was not valid.

B) Combined all 6 matrices into one OD matrix. The truck matrix was multiplied by 1.9 to transform it into a passenger car equivalent before adding it to the other 5 matrices. Congested speeds from the full model run were used as the speeds on the network. Note: congested speed is not recommended to be used as the assigned speed when doing the OD estimator but we wanted to see if it did effect the estimation/assignment. By combining all the matrices we can use the 2 hr count since we only have one matrix to try to match the count.

C) Same as option B but used free flow speeds on the links instead of the congested speeds.

D) Preloaded the trucks onto the network assuming they would use the shortest path regardless of capacity/congestion. No PCE conversion was made in this run.

E) Same as option D except the preloaded volumes were passenger car equilvalents instead of trucks (the truck matrix was multiplied by 1.9).

After running all of these scenarios the total O/D pairs estimated by the runs were as follows:
Model = 288,000
Scenario B = 290,461
Scenario C = 284,200
Scenario D = 289,362
Scenario E = 292,603

The total OD pairs seemed to stay in line for the different scenarios, however, when comparing the zone to zone flows they were drastic differences. Over 10,000 OD pairs difference in some cases. We also compared district to district flows (25 districts).

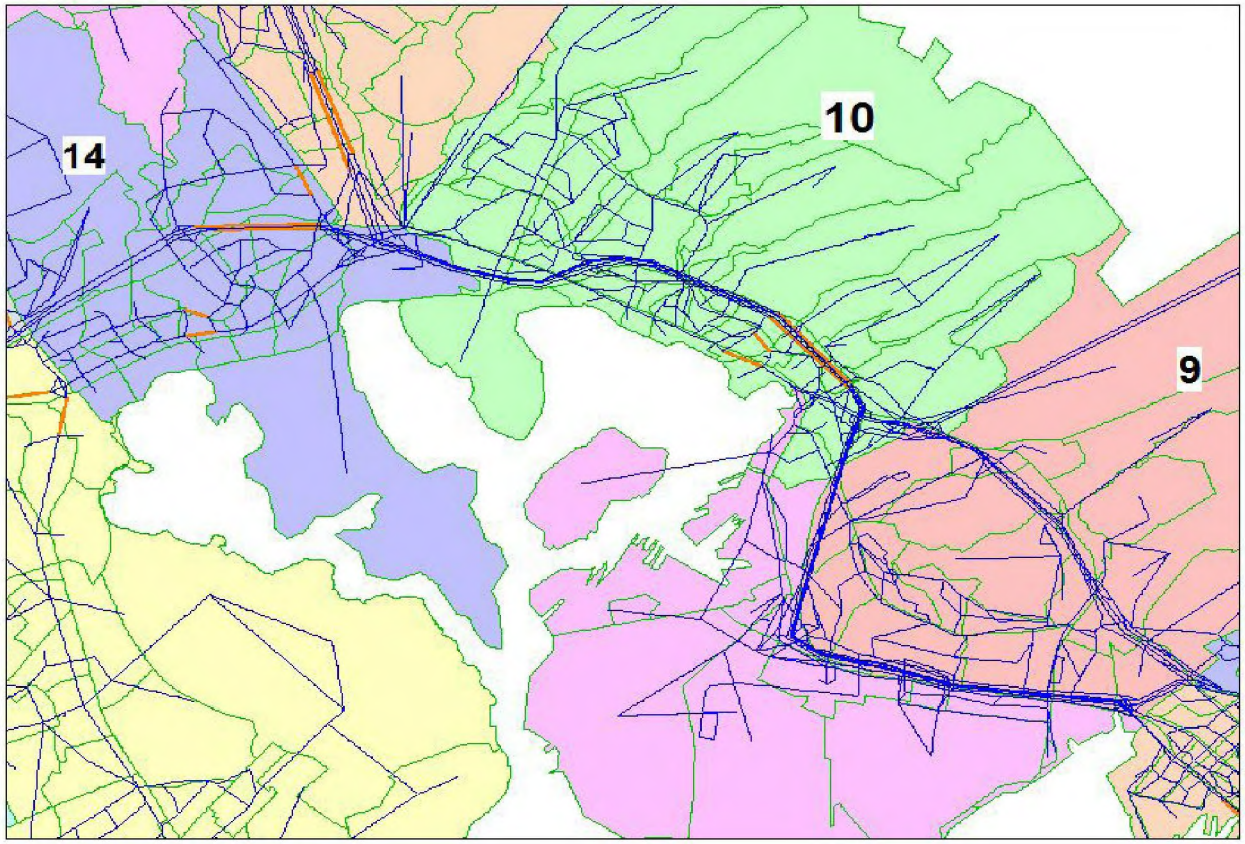
MCDIFF shows the difference between the model OD district flows and the scenario C district flows.

Results of some of these district flows are displayed below:

	idist	jdist	optC	model	MCdiff
51.	3	1	4384.53	1483	2901.53
103.	5	3	6386.92	2257	4129.92
151.	7	1	4329.76	1280	3049.76
156.	7	6	3654.16	1229	2425.16
157.	7	7	6229.58	2031	4198.58
207.	9	7	4269.17	1751	2518.17
214.	9	14	12587.43	355	12232.43
237.	10	12	3732.69	486	3246.69
334.	14	9	7226.17	874	6352.17
364.	15	14	5135.69	634	4501.69
384.	16	9	4610.5	893	3717.5

The MCdiff is quite high for some of the district to district flows. The map on the next page shows the counts for a small area. You can see that we really only have counts on the major facility between 9 & 14. There are no counts in district 9 yet there are some major facilities located in that district. Therefore the imbalance in OD flows from above is explainable because the counts are being matched on either side of district 9 since there is nothing inside that district to determine if the flow to that district is correct. The estimator is assuming the flow must be coming from district 9 since higher volumes are located on both sides of that district.

Throughout the network these larger counts dominate the OD estimation since we don't have many counts.


14
10
9

This chart shows the major overestimation of some of the district flows. There is also underestimation(negative values) in the scenarios but they are not displayed in this section.

	idist	jdist	model	MCdiff	MBdiff	MDdiff	MEdiff
51.	3	1	1483	2901.53	3455.35	3067.86	3856.04
103.	5	3	2257	4129.92	2926.11	4348.88	4425.8
151.	7	1	1280	3049.76	6086.52	4461.55	4312.41
156.	7	6	1229	2425.16	815.37	2508.39	1355.29
157.	7	7	2031	4198.58	4874.4	3977.57	2425.01
207.	9	7	1751	2518.17	3246.26	3600.58	3482.9
214.	9	14	355	12232.43	5367.86	12812.07	12735.19
237.	10	12	486	3246.69	-12.04	3350.53	3148.28
334.	14	9	874	6352.17	11142.13	9378.72	13111.31
364.	15	14	634	4501.69	8783.97	5649.51	6353.63
384.	16	9	893	3717.5	-310.86	2423.66	2263.5

The conclusion was that the current status of the counts on the network did not allow for the proper estimation of OD's. We have 72 counts on the network currently and 4414 links. That means that only 1.7% of the links have a count. And all of the links that do have counts are only on screenline locations which do not give us a good coverage of counts across facility types. Most of the counts we do have are on major facilities and so the zone/districts that fall within those limits receive the higher OD pair estimation because of the count located between them (insert the figure of district 3 to 9 or whatever it was) The traffic flows on the facilities with counts of course do a great job of matching the counts. However, they force the numbers to be correct on those facilities(due to the nature of concept) but because other facilities don't have count values we have unrealistic values for those areas.

It also hurts that we do not have truck counts specified separately in the database because we cannot use the multi-class estimator for trucks as their own matrix because we don't have a truck count to balance to.

The conclusion is that we need to have many more counts on the network in order to achieve a higher confidence in the flows and OD pairs that are predicted. We have about 250 more counts that are available to us and we will try to get them on the network and re-run the estimation. Some of those counts have truck information so we will test several scenarios once we have the count data on the links. We will assume fixed truck percentages from the counts based on the facility type. So looking at the truck counts we do have we might assume that all freeways have 10% trucks, local roads 3% trucks. This will allow us to assign the truck matrix onto the network and balance it with only the truck counts.

We will also use scenarios C & D for the OD estimation since they are the most reasonable assumptions and produced the most logical results.

For future updates to the model it is suggested that hourly classification counts are taken on a larger portion of the network so that estimation and overall calibration is easier to achieve. This will allow for a true multi-class OD estimation by functional class since we will have truck and vehicle counts for each type of facility.