October 17, 2013

To: Ian MacMillan
South Coast Air Quality Management District (SCAQMD)

From: Robert L. Russell & Kanok Boriboonsomsin
College of Engineering-Center for Environmental Research and Technology (CE-CERT)

Re: Analysis of Business Survey Data

The Business Survey Data included 63 sites with warehouse spaces from 6,000 to 2,100,000 square feet. Thirteen of the sites did not report the warehouse size thus only 50 sites were potentially available for analysis of trip generation rates. The Air Quality Management District (AQMD) Stakeholder Working Group has chosen the following key criteria for high cube warehouses:

- Greater than 200,000 square feet of building size.
- Minimum ceiling height of 24 feet
- Minimum dock door ratio of 1 per 10,000 square feet
- Used primarily for distribution of goods to stores or other warehouses
- High level of automation/mechanization

Only 34 of the 50 warehouses had a square footage greater than 200,000 sf. The only data available about vehicle entry or exit to the sites was “How many employee vehicles enter through your gate(s) on the busiest day of a typical week?” and “How many delivery trucks enter through your gate(s) on the busiest day of a typical week?” Only 23 of the 34 warehouses reported the number of employee vehicles and only 31 of the 34 warehouses reported the number of delivery trucks. We assumed that all the vehicles which entered through the gates on the busiest day of a typical week also exited through the gates on the same day. Therefore the average trip generation rates and the regression analyses were performed after doubling the number of employee vehicles and the number of delivery trucks. The results of the analysis are in Table 1.

Figure 1 shows the scattered plot of number of employee vehicle trips versus 1000 sf of gross floor area (GFA). Figure 2 shows the scattered plot of number of delivery truck trips versus 1000 sf of GFA. The regression program flagged one case for the employee vehicle trips as an outlier based on the Studentized residual and one case as having a large leverage effect. The regression program flagged two cases for the delivery truck trips as outliers and one case as having a large leverage effect. After the outliers and large leverage cases were removed and the regressions rerun: (1) the R^2 of the linear equation for the employee vehicle trips increased to 0.459 but one new case was flagged as an outlier and another one as having large leverage, (2) the R^2 of the logarithmic equation for the employee vehicle trips increased slightly to 0.390, (3) the R^2 of the linear equation for the delivery truck trips decreased to 0.001, and (4) the R^2 of the logarithmic equation for the delivery truck trips decreased to 0.005.

Table 1: Average and Regression Equation Results

		Equations			
Scenario	**Average Rates per 1000 sf of GFA**	**Linear** **T = No. of Trips** **X = 1000 sf GFA**	**R^2**	**Logarithmic** **T = No. of Trips** **X = 1000 sf GFA**	**R^2**
Employee vehicles on busiest day	0.663	T = 0.4526(X) + 117.97	0.276	Ln(T) = 1.6845(Ln(X))-5.1626	0.380
Delivery trucks on busiest day	0.534	T = 0.1568(X) + 189.35	0.069	Ln(T) = 0.3758(Ln(X)) + 2.8888	0.060

Number of Warehouses: 23

Average 1000 Sq. Feet GFA: 594

Trip Generation per 1000 Sq. Feet Gross Floor Area

Average Rate	Range of Rates	Standard Deviation
0.663	0.016 - 2.369	0.562

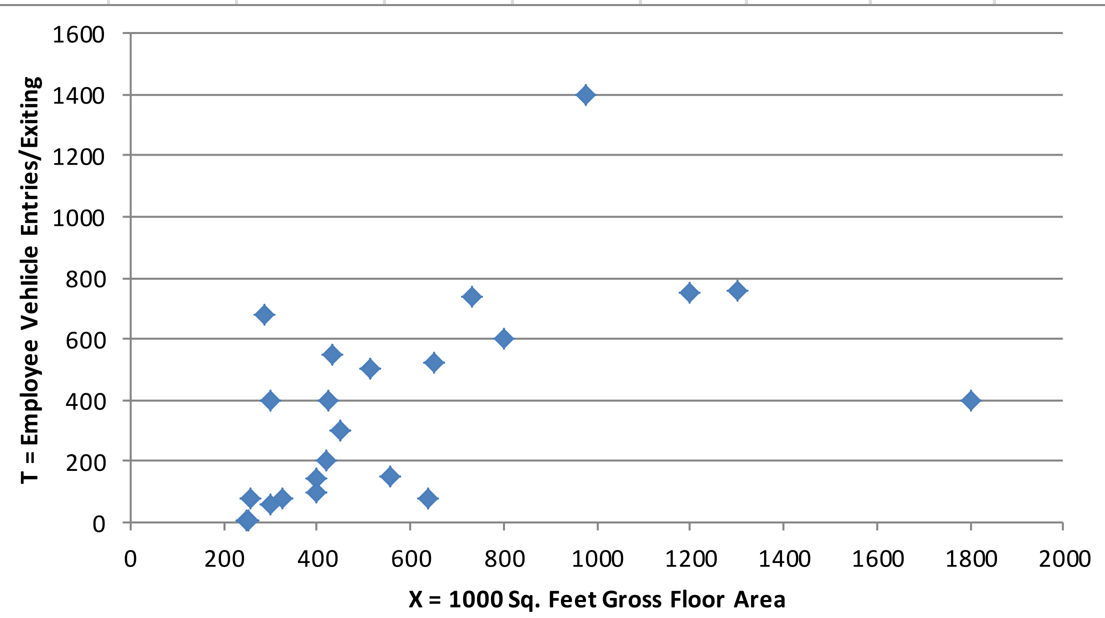

Figure 1: Number of Employee Vehicle Trips on Busiest Day vs 1000 sf of Gross Floor Area

Number of Warehouses: 31
Average 1000 Sq. Feet GFA: 713

Trip Generation per 1000 Sq. Feet Gross Floor Area

Average Rate	Range of Rates	Standard Deviation
0.534	0.048 - 1.755	1

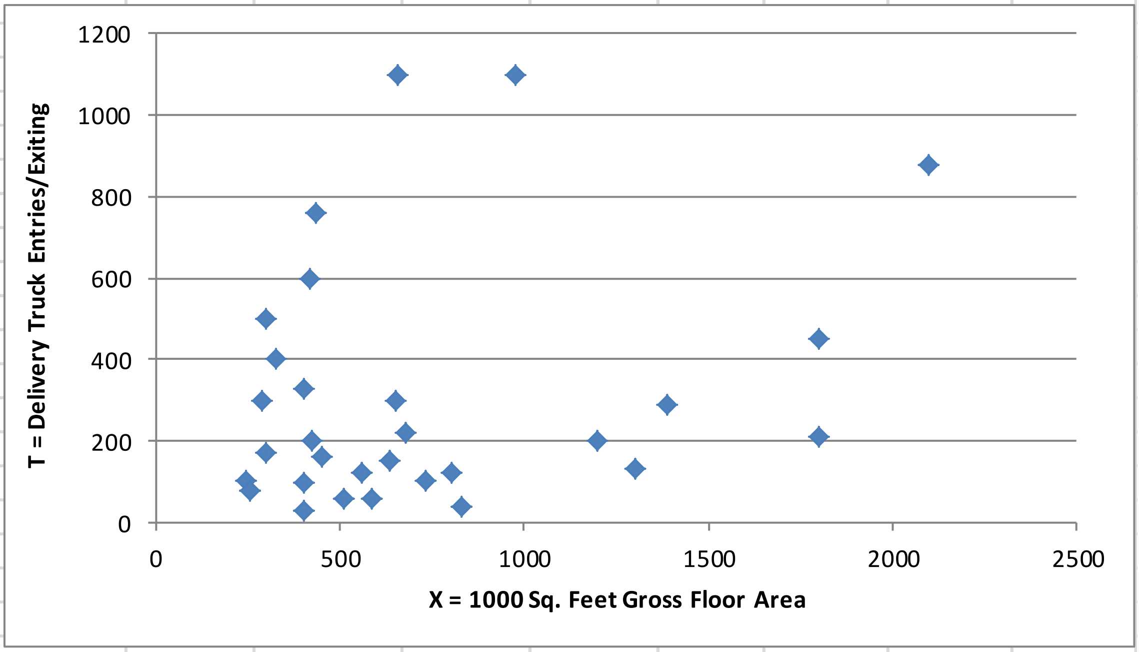

Figure 2: Number of Delivery Truck Trips on Busiest Day vs 1000 sf of Gross Floor Area

To determine whether adding additional independent variables would improve the correlation between number of delivery truck trips and 1000 sf of GFA, regressions were performed with the following equation specifications:

T = X + Y + constant
LN(T) = LN(X) + LN(Y) + constant

And for a couple of cases:

T = X + Y1 + Y2 + constant
LN(T) = LN(X) + LN(Y1) + LN(Y2) + constant

Where:

- T = Number of delivery truck trips
- X = 1000 sf of GFA
- Y = various variables per below
  - CHFT = Ceiling Height of warehouse in feet
  - WHDD = Number of dock doors for warehouse
  - WHPU = Percentage utilization of warehouse

ITOM = Number of months to turn over inventory
EWDR = Estimated Weighted Distance Received goods traveled
EWDS = Estimated Weighted Distance Shipped goods traveled
MFP = Miles from Ports (Los Angeles & Long Beach)
MOC = Miles of Conveyors

The approximate ceiling height of the warehouses and the number of dock doors for the warehouses were available in the survey data. The percentage utilization of the warehouses was reported as a range of values or as less than a value. We used the center of the range and 5% less than the reported less than value. The number of months to turn over the inventory was reported as less than a given number of months. We used the less than number as representing the approximate number of months to turn over inventory. The miles of conveyors were reported as a range or as “we don’t use conveyors”. If conveyors weren’t used, we set the value for miles of conveyors to 0.

The survey data reported the approximate percentage of deliveries that came from and were shipped to the following locations: Port of Los Angeles, Port of Long Beach, Port of San Diego, Port Hueneme, Other Warehouses, Rail Yard, Airport, Local Store (Southern California), Non-local Store (Other California), and Out of State. Some of the percentages totaled to more than or less than 100% and some warehouses reported total numbers instead of percentages. All percentages which totaled more than or less than 100% were normalized to total 100% and the numbers were converted to percentages. The addresses of the warehouses were available in the survey data so for the ports, Mapquest was used to determine the driving distance from the ports to the warehouse. For the other locations, since there were no specific addresses we used the following estimated distances: Other Warehouses (30 miles), Rail Yard (40 miles), Airport (30 miles), Local Store (50 miles), Non-local Store (250 miles), and Out of State (1000 miles). The EWDR and EWDS were then calculated for each warehouse by summing the percentage times the miles for each location.

MFP assumes that all goods are received from and delivered to either the Port of Los Angeles or the Port of Long Beach.

The R^2 of the regression equations after adding each of these independent variables are shown in Table 2. For all the variables except for MOC, we required that the data samples have a value greater than 0 to be included in the regression analysis. For MOC, we conducted the regression analysis both with and without the data samples having a value of 0 mile.

According to Table 2, adding ceiling height of warehouse, number of dock doors for warehouse, or miles of conveyors variable more than doubled the R^2 of both linear and logarithmic equations. Adding number of months to turn over inventory variable increased the R^2 of the logarithmic equation significantly, but only increased the R^2 of the linear equation slightly. These variables help improve the correlation between number of delivery truck trips and 1000 sf of GFA.

Table 2: R^2 values after adding independent variable(s)

	Linear		Logarithmic	
Y	N	R^2	N	R^2
CHFT	28	0.342	28	0.121
WHDD	31	0.248	31	0.152
WHPU	27	0.117	27	0.074
ITOM	24	0.082	24	0.304
EWDR	22	0.130	22	0.055
EWDS	22	0.031	22	0.048
MFP	31	0.083	31	0.066
MOC*	27 15	0.156 0.300	15	0.462
ITOM + MFP	24	0.098	24	0.304
WHPU +MFP	27	0.143	27	0.088

*N = 27 with data samples of 0 mi, N = 15 without data samples of 0 mi