APPENDIX G

Classification and Regression Tree Analysis

This page left intentionally blank.

Classification and Regression Tree Analysis

A classification and regression tree (CART) analysis was conducted by LADCO using 8-hour ozone monitoring data from three Cincinnati-area ozone sites: Colerain (ID number 39-061-00101), Sycamore (39-061-00061), and Taft (39-061-0040) measured from 2000-2014. The goal of the analysis was to determine the meteorological conditions associated with high ozone episodes in the Cincinnati air-shed and to construct trends for the days identified as sharing similar meteorological characteristics.

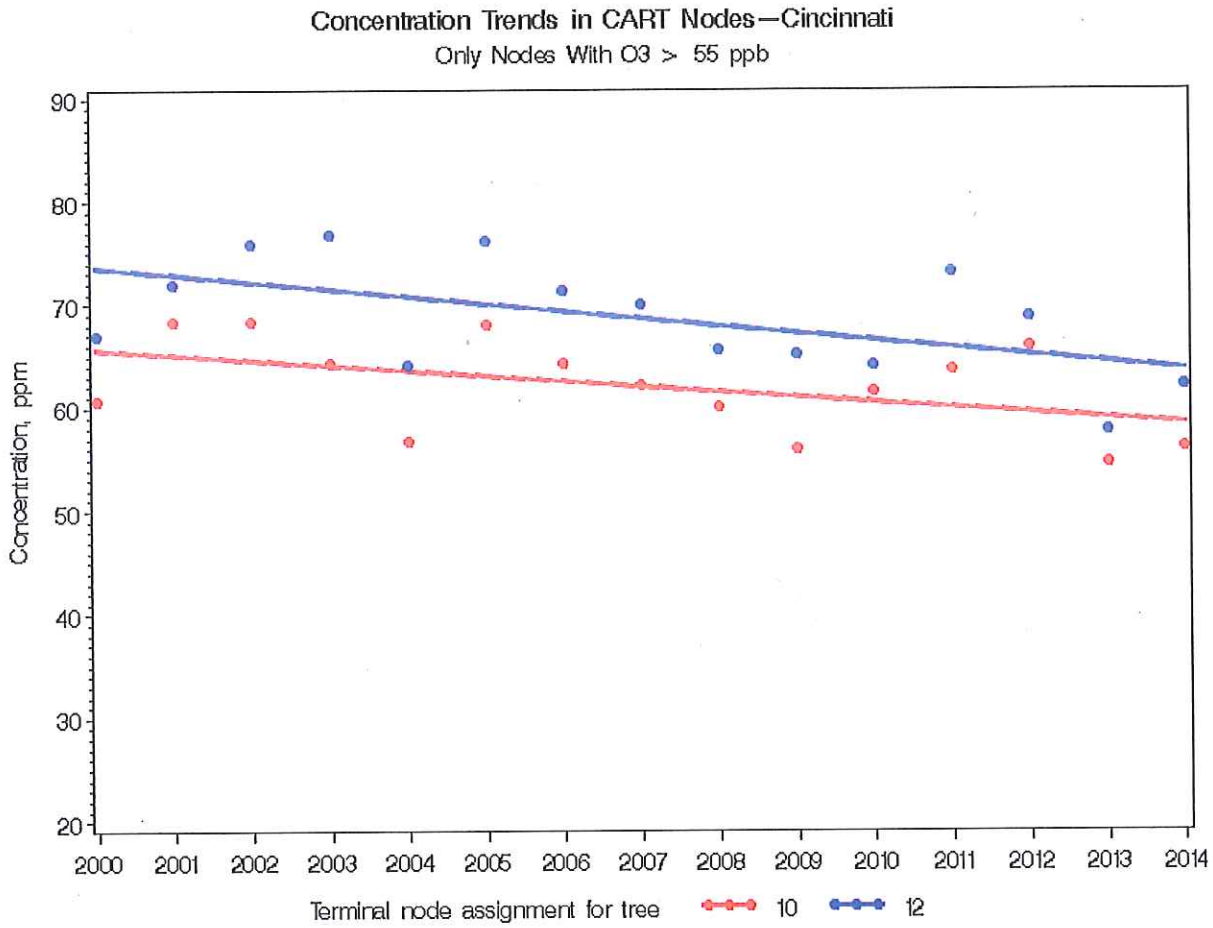
The CART analyses for the Cincinnati-area ozone study processed multiple meteorological variables for each day to determine which are the most effective at predicting ozone. Meteorological data collected for the Cincinnati CART analysis was taken from Cincinnati, Ohio - Municipal Airport's National Weather Service (NWS) station and processed by. Upper air observations, taken from the Dayton, Ohio - Wright Patterson Air Force Base Airport NWS site, were downloaded from the National Climatic Data's Center (NCDC) Integrated Global Radiosonde Archive. Meteorological variables included maximum and average daily temperatures, dew points, relative humidity and air pressure at the surface and different levels of the atmosphere, wind directions and wind speeds, change in temperatures and air pressure from the previous day, average wind speeds and temperatures over a 2 or 3-day period, day of the week, cloud cover, daily precipitation and many other parameters.

Regression trees, where each branch describes the meteorological conditions associated with different ozone concentrations, were developed to classify each summer day (May – September) by its ozone concentration and associated meteorological conditions. Although the exact selection of predictive variables changes from site to site, the universally common predictors are temperature, wind direction, and relative humidity. These are included in the dataset as daily averages and maximums as well as averages at specific times throughout the day (morning 7-10 am, afternoon 1-4 pm, etc.). Similar days were assigned to nodes, which are equivalent to branches of the regression tree. By grouping days with similar meteorology, the influence of meteorological variability on the underlying trend in ozone concentrations is partially removed; the remaining trend is presumed to be due to trends in precursor emissions or other non-meteorological influences. Ozone trends in these nodes were then plotted.

The CART analysis determined that two meteorological conditions had the strongest correlation with high ozone episodes: Node 10, representing south-southwest wind, and Node 12, representing a temperature difference between the surface and 925 mb greater than 2.26 K.

Chart 1, presented below, shows that, for both Node 10 and Node 12, monitored ozone values are trending lower for the most recent 15 years. This is to say, when wind directions are from the south-southwest and temperature changes at an altitude of 925 millibars are above 2.26° Kelvin (parameters associated with the highest ozone values in the past), ozone values are trending lower.

Chart 1



By using a CART analysis to analyze the Cincinnati-area 8-hour ozone data, the influence of variations in meteorology can be mitigated such that comparisons of high ozone days with similar meteorological conditions can be made to determine if ozone values have decreased over time due to anthropogenic emission reductions. In general, ozone trends in the Cincinnati-area have declined. Furthermore, under meteorological conditions when monitored 8-hour ozone has historically been at its highest, ozone concentrations are lower under similar meteorological conditions. This analysis demonstrates that lower ozone values are not caused by favorable meteorological conditions and that progress in reducing ozone precursor emissions is the primary reason for lower 8-hour ozone concentrations in the Cincinnati metropolitan area.