# Use of ARB's Ambient Toxics Data

The toxics monitoring schedule calls for samples to be collected every 12 days, currently at 17 sites throughout California. Consequently, there are a maximum of 31 values a year for a given toxics substance at a given site, or 2 or 3 values per month. Since there are occasionally problems with the collection or the analysis, it is not unusual for there to be an occasional month or so missing from the data record for a given substance at a given site.

The toxics sampling network was designed to produce a statewide annual average. Because an occasional missing month of data at a site is relatively unimportant when averaged over the entire network, the production of a statewide annual average is an appropriate use of the data. The network was not designed to produce statistics representing smaller geographical areas or shorter time periods.

We include two values for the annual mean on the DVD/CD. The first is an arithmetic average ("Average") of all measurements of a substance during the year. The standard deviation is based on this mean. The second annual average is the mean of the monthly means ("Average of Monthly Average). The mean of monthly means is listed only when there is at least one measurement in each month of the year, whereas the arithmetic average is the simple average of all measurements in a year, whether or not all months are present.

Most of the toxics substances show some seasonal variation, and some substances differ by as much as two orders of magnitude between the high and low periods of the year. In addition, the data for some substances can vary widely from sample to sample, creating non-systematic variations from month to month. If a month's data are missing, the calculated average could be significantly different from the real average, i.e., the average that would have been calculated had the missing month's data been available. The absence of a mean of monthly means indicates that the annual average for that year may not be representative of this "real" average. You should carefully consider the data for the individual months and use the annual average in the context of what months are or are not present.

Since the sampling program was designed to produce statewide annual averages, it is important that the data from a single site not be used too broadly. We have observed that data from sites that are as little as 20 miles from each other can differ drastically. As a result, data from one sampling site should not be used to characterize the annual average of a location beyond the general area of the sampling site.

Similarly, when sites are moved, added, or discontinued, this could also affect the statewide annual average. If a site with lower annual concentrations is replaced by a site with higher annual concentrations, then the statewide annual average can be expected to increase, and vice versa. The effect could inappropriately be interpreted as a change in emissions.

While the monitoring schedule calls for samples to be collected every 12 days, the schedules for sites in northern vs. southern California are offset by 6 days. As a result, monitoring never occurs on the same day at all sites in the state. At best, monitoring occurs at half the sites in the monitoring network on any given day.

In addition, on many scheduled sampling days, data are not available for all of the sites that were scheduled to sample on that day. Technical problems often result in a sample being invalid, which means that a make-up sample is collected during one of the days following the scheduled sampling day. When that occurs, the make-up sample may be the only sample collected in the state on that day.

When viewing the daily toxics measurements, you should keep in mind that the California-Daily highest value represents the highest measurement of the sites in California that were collected samples on that day. On days when only one site collected a sample, the daily highest value represents only that one site. Because low measurements are more commonly seen than are high measurements, it's likely that this "highest" measurement is lower that it would have been had measurements been available from more sites.

When we calculated the statistics for the toxics substances, we substituted one-half the limit of detection (LOD) for those measurements that were below the limit of detection. This practice, while widely practiced, can lead to a misleading situation when the LOD changes. When a large percentage of the measurements are below the LOD and the LOD changes, the means of monthly means and the annual averages will change as the result of the change in the LOD. This effect should not be mistaken as a change in emissions. The current LODs and the LOD history for each compound are listed in LOD.PDF on the DVD/CD.

The risks shown in ARB's annual toxics summary pages are estimated chronic cancer risk (acute risks and non-cancer risks are not shown). These risks are expressed in terms of expected cancer cases per million population based on exposure to the annual mean concentration over 70 years. They are calculated using unit risk factors provided to the Air Resources Board by the California Office of Environmental Health Hazard Assessment.

The monitoring sites in Mexico and the Calexico site in California are included in the annual toxics file, but the California statewide value only includes California sites. Also, both the Mexico and Calexico sites are included in the daily toxics file.

We encourage you to use the data appropriately. If you have questions about whether or not your use of the data is appropriate, please feel free to call Mike Redgrave of the Air Quality Data Review Section of ARB's Planning and Technical Support Division at (916) 322-7071. Additional information from the ARB's Monitoring and Laboratory Division regarding the quality of the data that can be found in the document titled "Data Quality" is located in D:\DOC\QA.DOC or D:\DOC\QA.PDF.

NOTE: All pre-1999 benzene and 1,3-butadiene data in the primary toxics data file, DLYTOXIC.DBF, have been adjusted, as described in D:\Doc\BzButAdj.PDF or D:\DOC\BzButAdj.Doc.

NOTE: 1995 and earlier acetaldehyde and formaldehyde data are the result of a sample collection method known to underestimate ambient concentrations. This method allowed an unknown proportion of ambient acetaldehyde and formaldehyde to be lost. As of January 1, 1996, a revised collection method was adopted that does not have this problem. We recommend that pre-1996 and 1996 and later formaldehyde and acetaldehyde data be treated as separate data sets and not be compared. For more information about the nature of the method's problem, refer to http://www.arb.ca.gov/aaqm/toxics.htm#carbonylstudy96.

NOTE: The toxics data from the Bay Area Air Quality Management District (BAAQMD) are included as a convenience to our users. Their inclusion does not imply that the Air Resources Board makes any endorsement of the quality of those data. Any questions about the quality of the BAAQMD toxics data should be directed to Nancy Balberan, Air Quality Instrument Specialist, at (415) 749-4630 at the BAAQMD.

NOTE: When Toxics data are compared with NMOC data on this DVD/CD, the Toxics gases units are in ppbV while the NMOC units are in ppbC, therefore a conversion from ppbV to ppbC needs to be applied based upon the formula:
(ppbV) X (number of carbons in the compound) = ppbC.

February 2008