2014 NATA – Monitoring Data Information

The monitoring data used in the 2014 NATA model-to-monitor comparisons is available as a layer in the NATA Map app. This layer provides annual concentrations measured at monitoring sites for air toxics covered in NATA. These sites were not used directly to generate the NATA risk results because the number of sites and air toxics monitored are each too few for their direct use in estimating risk.

For sites and pollutants with sufficient monitored values, we summarize the measurements to help you better understand the distribution of measured concentrations. We provide concentrations for the NATA year, 2014, plus other years with enough monitoring data for us to compute annual concentration averages. Where available, we also provide measurements taken after 2014 to give you more recent information on pollutant concentrations at the monitor. In addition, we provide trends information so you can see how concentrations have changed over time.

In this layer, users can explore for each site and monitored pollutant:

1) Descriptive statistics that show the distribution of the data for 2014 (where available) and the most recent year of monitoring data available. We only display sites with 2008 data or later. You can obtain the detailed daily (or hourly) measurement for any monitor from the air toxics archive.
2) Monitored and modeled values used in the NATA model evaluation. Note that for modeled pollutant groups such as xylenes, the modeled value is not shown, as it reflects total xylenes and not the monitored individual components within the xylenes group (e.g., m/p-xylene and o-xylene). Also note that we did not use some sites for the model evaluation due to the impact of non-detected values. In these cases, both the modeled and monitored values will be missing. More information on model evaluation is provided in Section 3.7 of the NATA Technical Support Document.
3) Trends plots that display the trends in annual concentrations over time. These are shown as box plots. Boxes display the 25th, 50th (median) and 75th percentile of the daily averages at the monitoring site. The whiskers show the 10th and 90th percentiles of the measurements. The dot is the annual mean. In these plots, the descriptive statistics are computed using a value of zero where the pollutant was not detected. Plots are not provided for sites with only 1 year of data or for which the 90th percentile concentration for all years is a non-detect.

Below is a description of fields in the attribute table for this map app layer:

STATE_ABBR	Two-letter state abbreviation
AMA_SITE_CODE	Site Code in the Air Toxics Archive
AQS_PARAMETER_CODE	EPA's Air Quality System (AQS) code describing measured parameter
AQS_PARAMETER_NAME	Name of the pollutant associated with the AQS PARAMETER CODE
DURATION_DESC	Duration of the sampling (24-hour, 1-hour, etc.)

NEI_pollutant	Pollutant name in the National Emissions Inventory (NEI)
mod_conc2014_ugm3	Modeled concentration, ug/m3
median2014_ugm3	Median annual measured ambient concentration value for year 2014, computed via the Regression on Order Statistics (ROS) Approach via the NADA package in R
mean2014_ugm3	Mean annual measured ambient concentration value for year 2014, computed via the Regression on Order Statistics (ROS) Approach via the NADA package in R
max2014_ugm3	Maximum daily concentration for year 2014
Variance2014	Variance of the concentrations for 2014, computed from daily values (days with a concentration of ND are treated as a zero value)
meanMDL2014_ugm3	The mean of the daily method detection limits (MDL) for year 2014
PERCENTlessthanMDL2014	Percent of days in 2014 for which the pollutant was less than MDL
PERCENT_NonDETECT2014	Percent of days in 2014 for which the pollutant was not detected
dailycount2014	Number of days averaged to get the annual value for year 2014
mostrecentYear	The most recent year in which a complete year of data were available to compute annual statistics based on the data in the air toxics archive
RECENTmedian_ugm3	median annual measured ambient concentration value for the most recent year, computed via the Regression on Order Statistics (ROS) Approach via the NADA package in R
RECENTmean_ugm3	mean annual measured ambient concentration value for the most recent year, computed via the Regression on Order Statistics (ROS) Approach via the NADA package in R
RECENTmax_ugm3	Maximum daily concentration for the most recent year
RECENTvariance	Variance of the concentrations for the most recent year, computed from daily values (days with a concentration of ND are treated as a zero value)
RECENTmeanMDL_ugm3	Mean of the daily MDLs for the most recent year
RECENTlessthanMDL_percent	Percent of days for the most recent year for which the pollutant was less than MDL
RECENTnonDETECT_percent	Percent of days for the most recent year for which the pollutant was not detected
RECENTdailycount	Number of days averaged to get the annual value for the most recent year

MONITOR_LATITUDE	Latitude of site location, decimal degrees
MONITOR_LONGITUDE	Longitude of site location, decimal degrees
Location	Location is filled out only for monitors at National Air Toxics Trends Sites (NATTS)
Setting	Urban or rural – filled out only at NATTS
LOCATION_TYPE	From the Air Monitoring Archive (also called Air Toxics Archive) Data Dictionary which provides metadata associated with each site
CITY	ditto
LAND_USE	ditto
cancerbenchmark1permillion_ugm3	Value of the concentration that represents 1-in-1 million cancer risk
noncancerbenchmark_HQis1_ugm3	Value of the concentration that represents the noncancer risk threshold. Above this threshold indicates potential for noncancer health effects
NEI_pollutant_code	Pollutant code in the NEI that matches the AQS_PARAMETER
CMAQhap	Yes or No, depending on whether this pollutant is modeled in CMAQ for NATA
URE	Unit risk estimate used in NATA
RFC	Reference concentration used in NATA
resp	Yes, if it has respiratory noncancer health impacts
neuro	Yes, if it has neurological noncancer health impacts
PAHgroup	Yes, if it is a polycyclic aromatic hydrocarbon
NATA_file_NAME	Name of the file containing modeled results for the pollutant
NATA_POLL	Name of the modelled NATA pollutant
SMOKENAME	Name of the modelled pollutant
Spearmancorrelation__mean	The correlation of annual mean values from years 2008 and 2016, using the Spearman rank order approach
pvalue	Significance (statistical) of the Spearman correlation from years 2008 and 2016
trend_of_mean	Direction of the trend from years 2008 and 2016 produced by the Spearman rank correlation.
N_years	Number of sufficiently complete years of data