March 25, 2015

**EPA RUTGERS RARITAN RIVER PROJECT – Data Gaps**

Project-Defined Missing Data:

* As outlined in the EPA-approved QAPP, the following data criteria were established for inclusion in the EPA Rutgers Raritan River Project database and subsequent analyses:
  + Data have been generated under an approved QAPP or other sampling document.
  + Data have been generated by a reliable source {i.e., the data generator is generally trusted and respected (federal, state, and local agencies or research institutions)}.
  + Data have been published in peer-reviewed articles or publications.
  + Data have been widely used and/or is trusted by scientists and professionals.
  + Data have been collected for similar purposes (i.e., to assess environmental health) to describe pollution sources and sites.
  + Data have been collected along the Lower Raritan main stem and along the main tributaries in Watershed Management Area 9 (WMA09).
  + Analytical methods were sufficiently sensitive to support data reporting to regulatory criteria.
  + Reported data include laboratory qualifiers and qualifier definitions.

If data did not meet this set of criteria, they were not included within the final database and analyses.

* It was established by the Project Team (the Pi and Co-PIs) that the data collection efforts would focus on the surface waters of the Lower Raritan River and the Raritan Bay, where available. This excluded groundwater and drinking water from the data gathering work.

Data Gaps:

* Data collection was carried out until June 2014, and all data are accurate up to that point in time. Data created or updated after June 2014 are not included in the project database.
* There is a lack as to the current ‘status’ of sites associated with the point sources of pollution (known contaminated sites, Brownfields, Superfund, etc.) therefore it is difficult to determine if a site is closed from remediation, part of an open/active investigation, or the remediation efforts are pending.
* Data on the levels and/or concentrations of pollutants and contaminants was lacking for point source pollution sites (i.e., Brownfields, known contaminated sites, etc.). The data indicating what caused the location to be listed as such a site is missing from the database.
* There is a paucity of data related to any of the biological or ecological communities in the project area (excluding bacteria and aquatic macroinvertebrates). Populations of commercially and recreationally important species (fish, shell fish, mammals, and birds) are available in few publications with only summary data reported.
* There is a lack of complete spatial coverage for some areas for some sets of data, most likely due to the time and funding needed provide comprehensive coverage. More sites are located in the lower portion (Middlesex County) of the Raritan River Project Area than in the upper portion (Somerset County). In addition, there are few sites that are consistently monitored in the Raritan Bay.
* While the Lower Raritan River and the Raritan Bay have undergone dredging in the past, data on dredging activities are missing from general searches conducted for the project database. Information regarding timing, cost, reason, and size of dredging projects is not included in the project database.
* While there is a large amount of data available through the EPA’s Storage and Retrieval Data Warehouse (STORET), the data files and location/site files are separate downloads and present a challenge in linking data to particular sites. There is a lack of consistency in naming sites or agencies responsible for including data in STORET which made associating the data to the sites difficult and time consuming.