based upon the data submitted by one facility with three outfalls and EPA agrees that one facility should not be considered necessarily representative of an entire industry sector for the purposes of determining the need to monitor. If three facilities which discharge a pollutant, however, the pollutant is not unique to a particular facility and is indicative of the industrial activities conducted in the industry sector or subsector. EPA conducted the monitoring evaluation assuming both a normal distribution and a lognormal distribution of the data set. The results were not significantly different.

Quality of the Part II Database

The Part 2 group application database includes Part 2 monitoring data from participants which participated in the group application process. Statistical analyses (e.g., mean, median, 95th percentile, and 99th percentile values) of this data was conducted for each parameter within every industrial sector. These analyses were conducted assuming both a normal distribution to the data and a lognormal distribution. The results of the analyses were used in the methodology to determine the proposed monitoring requirements.

Several commenters stated that the database, which only included monitoring data received prior to January 1, 1993, was incomplete and/or contained errors. The commenters stated that the database should be expanded to include all the group application data, as well as further reviewed to eliminate duplications and inaccuracies. Other commenters requested that the methods used to develop the statistical evaluation of the data be revamped (e.g., use a lognormal distribution of the data). In addition, a few commenters stated that the analysis did not properly consider facilities which did not submit data for a pollutant listed in Part C of the Form 2F since these facilities had no reason to believe the pollutant was present in their discharge. Therefore, the commenters argued, EPA's analysis should assume that the discharge concentration of these pollutants is zero.

EPA has again reviewed and doublechecked the monitoring data analyzed for the development of the permit. EPA concludes that the monitoring data analyzed is representative of the industries evaluated. EPA analyzed data which was submitted months after the application deadline for the purpose of identifying pollutants of concern and developing monitoring requirements. In addition, on a sector-by-sector basis, EPA reviewed data that was submitted late to determine if the additional data was consistent with what had already been evaluated. Given this extra level of effort to analyze and consider all submitted data, even though some data was not loaded into the database that was publicly distributed, EPA believes that the analyses performed on the group application sampling data, and the results that were derived, are valid and reasonable.

EPA also believes that the concerns raised by commenters about the number of duplications and errors contained in the database which was distributed, is no longer warranted in that as errors were noted, EPA further screened and corrected the database. In response to the recommendation from commenters that a zero concentration value should be entered into the database every time a facility did not sample for a given pollutant because they did not believe it was present on their site, EPA does not agree. Obviously, assuming zero concentrations for these facilities would significantly reduce the mean and median concentrations. This would be imposing a major, unsupported assumption into the database. It cannot be assumed that facilities which did not submit data for a part B or C pollutant have a discharge concentration of zero for that pollutant. Facilities which did not sample for a pollutant because they did not believe it was present, may not have adequately considered all potential sources of these pollutants. In addition, facilities that did sample were supposed to be representative of the entire group in which they were located. This was a process determined by the group applicants themselves, with approval from EPA. Therefore, where facilities did sample and report for a given pollutant, and other facilities in the group did not, it could be assumed that the pollutant really was present at all other facilities. To be more accurate and unbiased in the analyses of the data, EPA chose not to assume either a zero value or an extrapolated value for pollutants that were not analyzed for by some facilities within a sector. EPA analyzed only actual data points that were submitted. Where a pollutant was tested for, and the result was below detection levels, EPA assumed these data points to be zero values for the pollutant.

Establishing Priority Monitoring Sectors

The multi-sector permit requires analytical monitoring only for 'priority' sectors. A sector was considered a 'priority' if, based on the Part II data for the sector, five or more pollutants sampled for had median concentrations above benchmark values. If the sector had median values greater than benchmark values for four or less parameters, only visual examinations would need to be conducted.

Several commenters stated that the methodology employed for establishing priority sectors was arbitrary and/or flawed (i.e there is no basis for choosing five as the number of parameters needed to be above benchmark levels to trigger sampling). Others indicated that the approach did not consider the relative impacts (e.g., toxicity) of the pollutants on receiving waters. Commenters also indicated that it was inappropriate to group together a wide range of industrial activity discharge data into one industry sector, and to use that data as a basis for comparison.

In response to these comments, EPA has revised the methodology for selecting which industries must conduct analytical monitoring. EPA reviewed the grouping of industries into sectors for statistical analysis. It was determined that in some cases a sector contained a grouping of industrial activities which may have different storm water discharges. In these cases EPA modified its analysis to statistically summarize the industry by subsectors. Division into industry sub-sectors was prepared in most cases based upon the three digit SIC codes provided by the group participants in their group application information. The results of the subsector analysis of the data were then used for comparison to the revised benchmarks (discussed above).

Today's permit also eliminates the five pollutant threshold for determining if a sector merited monitoring. For each subsector (or sector where it was not possible to further divide the sector into subsectors) EPA compared, on a pollutant by pollutant basis, the median concentration to the benchmark. Where the median concentration for a pollutant is higher than the benchmark, where there are likely sources of the pollutant associated with the industrial activity, and where the concentrations are high enough so as not to be due to "background" or natural sources, the subsector (or sector) is required to conduct analytical monitoring for the listed pollutant. This methodology is pollutant-specific and addresses the concerns that some commenters had that some industries within a sector may be inherently clean compared to other industries in the same sector. In addition, this approach is more environmentally protective in that the number of different pollutants in a discharge does not necessarily increase the risk posed by that discharge. It is possible that a receiving water may be significantly impacted by a discharge