Department of Pesticide Regulation

Paul Helliker
Director

Gray Davis
Governor

Winston H. Hickox
Secretary, California Environmental Protection Agency

M E M O R A N D U M

TO: Pam Wofford, Associate Environmental Research Scientist **HSM-02026**
Environmental Monitoring Branch

FROM: Sally Powell, Senior Environmental Research Scientist
Worker Health and Safety Branch

DATE: August 21, 2002

SUBJECT: STATISTICAL CONFIDENCE THAT THE MAXIMUM 24-HOUR MITC CONCENTRATION WAS CAPTURED IN THE 2000 LOMPOC FUMIGANT MONITORING

This is in response to the question from a reviewer of the draft report on air monitoring for fumigants in Lompoc, which you forwarded to me. Ray Chavira asked that we show statistically whether the maximum concentrations were in fact captured at the monitoring sites.

Method

Even with large numbers of samples, it is never likely that the highest measured value is the highest that is possible. There are statistical methods for estimating the percentiles of a parent population from sample data; some assume a specific statistical distribution for the parent population, others are distribution-free methods.

It is the practice of the Worker Health and Safety Branch (WHS) to assume lognormality for environmental contaminants in most cases. DPR's experience with many large environmental datasets has shown that they are usually well described by the lognormal distribution. In addition, WHS prefers to avoid the inconsistency of using different exposure statistics based on sample characteristics.

In order to estimate the "upper bound" of daily exposure, WHS generally uses the estimated population 95th percentile of daily exposure. Sample maxima and upper-end percentiles are both statistically unstable and known to underestimate the population values. The population estimate, on the other hand, is more stable because it is based on all the observations rather than a single value; moreover, it is adjusted for the number of samples, correcting some of the underestimation bias due to small numbers. A high percentile is estimated, rather than the maximum itself, because in theory, the maximum value of a lognormal population is infinitely large. In practice, exposures must be bounded because a finite amount of pesticide is applied. The use of a high percentile acknowledges that the assumed lognormal distribution is probably not a perfect description of the population, especially at the upper extremes. The population 95th, rather than a higher percentile, is used because the higher the percentile the less reliably it can be estimated.

The 95% tolerance limit is the concentration that, with given probability, will be exceeded in no more than 5% of future samples (Hahn and Meeker, 1991). It is equivalent to a 90% upper confidence limit on the population 95th percentile. For the lognormal distribution, it is calculated as:

$$95\%\ \text{tolerance limit} = \exp\{\text{arithmetic mean of } \log_e \text{ concentrations} + g_{(0.90;0.95;\, n)} * (\text{sd of } \log_e \text{ concentrations})\}.$$

The multiplier g for 90% probability is tabled in Hahn and Meeker (1991).

This analysis was done on MITC concentrations, since no other fumigant was found above trace levels.

Twenty-four-hour average MITC concentrations were used because the acute screening level was based on a 24-hour NOEL. The twenty-four average was calculated as the time-weighted average of total (front + back tubes) ng m^{-3} in the 8- and the 16-hr sample from each calendar day. If there was only one sample for a day, it was used as the 24-hr concentration. Before the 8- and 16-hr samples were combined, one-half the LOD was substituted for measurements below the LOD (0.003 was substituted for an ND 8-hr sample, 0.006 for a 16-hr sample), and the midpoint between the LOD and LOQ was substituted for measurements below the LOQ (0.051 was substituted for a trace-level 8-hr sample, 0.026 for a 16-hr sample).

The frequency distribution of 24-hr concentrations is shown in Fig. 1.

Results

The maximum observed 24-hr concentration was 592 ng m^{-3}. The estimated 95th percentile 24-hr air concentration was 2,537 ng m^{-3}. The 90 percent tolerance limit for the 95th percentile was 7,615 ng m^{-3}. That is, if these concentrations do come from a lognormal population, the probability is 0.90 that at least 95 percent of 24-hr MITC concentrations are below 7,615 ng m^{-3}.

Limitations

The validity of the calculated tolerance limit depends, first, on the assumption that the population of air concentrations in the Lompoc area during fumigations is lognormal and, second, on the samples taken being representative of the population both temporally and spatially. If the lognormality assumption is unfounded, the most likely alternative is a less highly skewed distribution. In this case, the calculated tolerance limit would be health-protective. With respect to the second requirement, since the siting and timing of samples could not be done either randomly or systematically, the representativeness of the data is unknown.

Reference

Hahn, G.J., and Meeker, W.Q. 1991. *Statistical Intervals: A Guide for Practitioners*. New York, John Wiley & Sons, Inc.

cc: Randy Segawa
Joe Frank

Fig. 1. Distribution of 24-hr MITC concentrations, Lompoc, 2000.