

US EPA ARCHIVE DOCUMENT

Pesticide Usage Estimation and Use in Estimating Dietary Risk

**Economic Analysis Branch
Biological and Economic Analysis Division
Office of Pesticide Programs
August 2002**

Executive Summary

Pesticides are widely used to control undesirable organisms in agriculture and in other parts of society. As well as controlling pests these chemicals may have harmful effects on non-target organisms including humans. Because of these possible harmful effects, residues from pesticides applied to food crops are permitted to be present in the harvested crop and in processed products made from that crop only at or below a certain fraction of the weight of the harvested crop. Under the Federal Food, Drug, and Cosmetic Act (FFDCA), the Environmental Protection Agency may authorize a tolerance or exemption from the requirement of a tolerance, to allow a pesticide residue in food, only if the Agency determines that such residues would be “safe” (FFDCA sec. 408(b)(2)(A)(i)).¹ The term “safe” is defined as a “reasonable certainty that no harm will result from aggregate exposure to the pesticide chemical residue, including dietary exposures and all other exposures for which there is reliable information” (FFDCA sec. 408(b)(2)(A)(ii)). In order to make safety findings, EPA conducts dietary risk assessments for pesticides. These risk assessment may depend, in part, on estimates of the percent of a crop that is treated with a particular pesticide. Analysts within OPP provide estimates of percent crop treated (PCT) to support these assessments. The purpose of this paper is to describe the present methods for estimating PCT, as background information to support OPP’s proposed refinements in calculating PCT.

This paper will first briefly discuss the estimation of dietary pesticide exposure, one of several possible routes of pesticide exposure. The various types of data used in this estimation will be covered. This paper then discusses various aspects of the determination of total individual pesticide exposure to show how percent crop treated data contributes to the determination of the “reasonable certainty” criteria listed above. Analysts’ estimates of percent crop treated data are used by dietary risk assessors in OPP. OPP uses a number of data sources and computer algorithms to develop these estimates. The present methods of estimating percent crop treated will be discussed, and this discussion is intended to help in forming improvements in these methods. One expected improvement will be a new algorithm that will aid in predicting average future percent crop treated with a pesticide, as well as possible maximum percent crop treated.

¹ Tolerance levels are set by EPA and FDA and represent a maximum allowable level of residue of a pesticide in a food.

Pesticide Usage Estimation and Use in Estimating Dietary Risk

- I. Pesticide Toxicity, Exposure and Risk** 1
- II. Dietary Pesticide Exposure** 2
 - B. Pesticide Residue Data** 3
 - A. Combining Residue Data with Pesticide Use Data** 5
- IV. Short History of Pesticide Usage Estimation in OPP** 6
 - A. Quality, Quantity and Availability of Pesticide Usage Data** 6
 - 1. Survey and census data** 6
 - B. Quantitative Usage Analysis (QUA)** 7
 - 1. Existing QUA Program** 7
 - 2. Present Methodology for Estimates of Percent Crop Treated** 7
 - 3. Estimating the Likely Average** 8
 - 4. Estimating the Likely Maximum** 9
 - 5. Limitations** 10
 - 6. Factors Influencing the Use of a Pesticide** 11
- Conclusions** 11

I. Pesticide Toxicity, Exposure and Risk

This section describes the concept of dietary risk and discusses several of the data types needed to estimate dietary risk. Pesticide residues are explained in more depth because percent crop treated is most commonly used in dietary risk assessment to calculate the probability of encountering residues. The size of the potential risks depends on the toxicity of the pesticide and the amount of that pesticide to which the individual is exposed. In general,

$$\text{Risk} = f(\text{toxicity, exposure})$$

For any individual there are several possible routes of exposure to a pesticide. An individual may be exposed to residues from pesticides applied at home or work; the individual may also consume food and water containing pesticide residues.

EPA is required by FQPA to determine that the use of a pesticide will only occur in such a manner as to cause a “reasonable certainty of no harm.” An estimate of the quantity of a pesticide to which an individual could be exposed, without exceeding this measure, may be determined by assessing animal toxicity data to determine a “no observable adverse effect level” (NOAEL) or, when the data does not provide a NOAEL, the “lowest observable adverse effect level” (LOAEL²). NOAELs and LOAELs are determined for both acute and chronic effects for each pesticide registered by EPA. NOAELs and LOAELs are normally determined by dosing groups of experimental animals with varying amounts of the pesticide. The animals are observed and examined for numerous possible pesticide effects. Both acute and chronic NOAELs are determined for each pesticide: an “acute NOAEL” is calculated following a one-time dosing while a “chronic NOAEL” is calculated following longer-term dosing. The more toxic the pesticide, the lower the NOAEL.

The NOAEL or LOAEL doses are adjusted downward by uncertainty factors to allow for inter- and intra-species variability in sensitivity to the pesticide. An uncertainty factor of 100 is generally applied to account for inter- and intra-species differences.³ An additional FQPA safety factor is also required under certain circumstances to ensure that infants and children are adequately protected. This safety factor may be reduced or removed if adequate data exists to show that infants and children are not more sensitive than adults. For example, if the NOAEL for a rat was determined to be a daily dose of 1 mg/kg of body weight and the FQPA factor was removed, the allowable daily dose for a human would be set at 0.01 mg/kg of body weight. The allowable daily dose (which accounts for the inter- and intra-species factors and the FQPA uncertainty factor) is termed the “population adjusted dose”

² Alternatively, when information is available, EPA can use a Benchmark Dose procedure.

³ A factor of 10 to allow for the possibility that humans as a species are more sensitive to the animal species that is considered to serve as an appropriate human surrogate for the harmful effect under consideration and a factor of 10 to allow for a range in human sensitivity.

(PAD). There are normally both an “acute Population Adjusted Dose” (aPAD) and a “chronic Population Adjusted Dose” (cPAD) [based on the respective acute and chronic NOAELs] used as toxicological “benchmarks” to compare to estimated short-term and longer-term exposures, respectively. Estimated exposure (at any given percentile of the distribution) is compared to these toxicological benchmarks (or reference points) and expressed as a percentage (e.g., an estimated exposure of 0.005 mg/kg of body weight would “occupy” 50% of the 0.01 mg/kg body weight PAD and EPA would say that 50% of the PAD is occupied).⁴

II. Dietary Pesticide Exposure

Exposure to a pesticide in the food supply depends, in turn, on two factors: the amount of the pesticide present in food and how much food a person eats. It is impossible to know precisely how much food every individual consumes, either over a lifetime or even on a single day. Similarly, it is impossible to know how much residue each specific item of food contains. Thus, the Agency must use available data that are reliable and representative to develop estimates of such exposure so that the %aPAD and %cPAD from dietary exposure can be estimated. Information on how much food an individual consumes is available from a food consumption survey performed by the U.S. Department of Agriculture (USDA’s 1994-96/1998 Continuing Survey of Food Intake by Individuals, or CSFII). Information on the pesticide residue concentrations present on food is available from a variety of sources including USDA’s Pesticide Data Program (PDP), FDA monitoring data, pesticide registrant-sponsored market basket surveys, and crop field trials performed in support of pesticide registration by pesticide registrants/manufacturers. Each of these sources of data is discussed briefly below.

A. Food Consumption Data

The above procedure depends on reasonable food consumption data and reasonable pesticide residue data. The food consumption data set contains information on food consumption from two or three day food consumption surveys conducted by USDA. This Continuing Survey of Food Intake by Individuals (CSFII) is a nationwide survey performed by USDA to measure what Americans eat and drink. The 1989-1991 database contains information for a three day period for nearly 12,000 individuals. The 1994-1996/1998 database contains information for a two day period for over 20,000 individuals. Each individual records the identity and quantities of food consumed (e.g, 100 g of pizza, 120 grams of apples pie, 90 g of raw apple). This information is “translated” by USDA by means of standard recipes to the raw agricultural commodities that make up or comprise the food item (in the case of pizza, for example, this would be tomatoes, milk, wheat, etc.). It is the translation of these food items to agricultural commodities which

⁴ For more details see “Additional Information on Assessing Exposure from Pesticides in Food” <http://www.epa.gov/fedrgstr/EPA-PEST/2000/July/Day-12/6061.pdf>

permits the random “matching” of food items consumed (expressed in terms of raw agricultural commodities) to measured pesticide residues (see below).

B. Pesticide Residue Data

Pesticide residue monitoring data are available from a number of sources. The obvious, and best, way to determine the average level and distribution of pesticide residues in food is to analyze an adequate number of food samples using food that is near the end of the distribution chain and ready to be consumed. A considerable amount of pesticide residue information is collected in this manner by the USDA Pesticide Data Program (PDP) and by other residue monitoring programs. The PDP monitoring data are the broadest and most comprehensive set of publically available monitoring data. Each year, since 1991, in consultation with EPA and other interested parties, the PDP selects a number of commodities for residue analysis. Several hundred samples of each commodity are selected randomly from distribution warehouses and analyzed for the presence of, and the residue levels of, a large number of pesticides and pesticide degradates. Individual commodities are normally sampled for several consecutive years.⁵ The presence of each pesticide can normally be detected at levels of parts per million or much lower. The Level of Quantification (LOQ) is the level at which a determination of the amount of the pesticide that is present can be made. The Level of Detection (LOD) is the level above which the pesticide can be said to be present, but the actual residue level can not be accurately determined except that it can be determined to be below the LOQ.

The Food and Drug Administration (FDA) also samples food commodities for residues. Under regulatory monitoring, FDA samples individual lots of domestically produced and imported foods and analyzes them for pesticide residues⁶. A complimentary approach to regulatory monitoring, known as incidence/level monitoring, has been used to increase FDA's knowledge about particular pesticide/commodity combinations by analyzing certain foods to determine the presence and levels of selected pesticides. (For example, from 1995 to 1997, a survey of triazines was done.)

Various other market basket monitoring studies have also been conducted by pesticide and food industry companies and associations. They are usually not as comprehensive as the PDP studies since they may have sampled one or more foods and looked for the presence of one or more pesticides.

⁵ PDP has now sampled over 40 commodities for the presence of more than 160 pesticides.

⁶ Domestic shipments are collected as closely as possible to the point of production in the distribution system; import samples are collected at the point of entry into U.S. commerce. This protocol differs from the PDP which collects samples near the point of consumption. FDA samples on the raw agricultural products, which are analyzed as the unwashed, whole (unpeeled), raw commodity. Processed foods are also included in this FDA monitoring. Domestic and import food samples collected for analysis are classified as either "surveillance" or "compliance." Most samples collected by FDA are the surveillance type; that is, there is no prior knowledge or evidence that a specific food shipment contains illegal pesticide residues. Compliance samples are collected as follow-up to the finding of an illegal residue or when there is other evidence of a pesticide problem.

Field trials are conducted to determine possible pesticide residues for each crop on which the pesticide will be used. Pesticides in these studies are normally applied at the maximum use rate and as close to crop harvest as is allowed by the pesticide label. These are conditions that can be expected to lead to highest residues on the harvested crop. When PDP, FDA, or other market basket residue data are not available, these field trial data may be used.

III. Use of the Dietary Exposure Evaluation Model (DEEM™)

In order to understand the importance of percent crop treated in dietary risk assessment, it is worth reviewing briefly the dietary risk assessment model which OPP uses to perform its estimation of dietary exposure to pesticides. OPP uses the Dietary Exposure Evaluation Model (DEEM) to estimate dietary exposure and risk assessments for pesticides in food. Much of the data described in the preceding section regarding food consumption surveys and pesticide residues plays a role in the dietary risk assessment. The DEEM™ software employs Monte Carlo techniques to generate estimates of the *distribution* of pesticide dietary exposures.

Data required to run DEEM include detailed information on food consumption and pesticide residue distributions for each food in the food consumption database. As stated above, the food consumption data used are from the USDA's CSFII, while the pesticide residue data can be from the USDA's PDP program, from FDA monitoring data, from market basket surveys, or from field trials. DEEM operates as follows to obtain a %aPAD:

- 1) Actual food consumption for an individual for a single day is drawn from the USDA Food Consumption Survey database.
- 2) For each food consumed that day, a residue number is drawn from the distribution of residues found in the residue database using PDP or other monitoring data or appropriate field trial data. This residue database is constructed so as to appropriately apportion the residues between treated commodities (which bear residues) and untreated commodities (which do not). For example, if 50% of oranges are treated crop with specific pesticide of interest, then there should be only a 50% probability of drawing a residue consistent with a treated commodity (and a 50% probability of drawing a zero value). Percent crop treated is key factor which is used to build these residue files such that they appropriately reflect the probability of consuming treated vs. untreated commodities.
- 3) The residues from each food (which may be zero in the case of untreated commodities) and the reported consumption amounts are multiplied together (for each individual and each reported consumption amount) to obtain an estimated exposure from that commodity and these exposures are summed (within an individual) to obtain a total one-day acute exposure estimate.

4) Steps (2) and (3) are repeated many times for each individual's reported diet, with each "iteration" randomly drawing and matching a pesticide residue with that individual's reported consumption amount. Each of these iterations results in an estimated total exposure.

5) Steps (1) through (4) are repeated for all individuals in the CSFII survey.

The result of this process is a distribution containing hundreds of thousands (or millions) of predicted (potential) exposure estimates (i.e., a distribution of predicted exposures). EPA evaluates the distribution of estimated exposures in making regulatory decisions.

A. Combining Residue Data with Pesticide Use Data

The last section briefly outlined EPA's dietary risk model, which depends critically on pesticide residue values (and the appropriate proportionate representation between treated and non-treated commodities). This section describes in additional detail the relationship between percent crop treated and residues in the dietary assessment model. The residue data from PDP and other monitoring studies can be classified in three ways:

- (a) observations with detectable residues,
- (b) observations with no detectable residues

For example: assume that 10% of the samples of a given commodity have detectable residues of a particular pesticide, and the remaining 90% may have no detectable residues. What assumption should be made about possible residue levels on the 90% of samples without detectable residues?

When available, percent crop treated data can provide information about what fraction of the 90% of the samples with no detectable residues may actually contain some low level of residues. From our example, residue levels on 10% of the samples are known. Suppose that we know that 35% of the crop was treated with a pesticide. It would be reasonable to assume that 65% of the samples would therefore contain no pesticide residues (by virtue of the fact that they were not treated with pesticide). This leaves 25% of the samples that were treated but have non-detectable residues (35% treated less 10% detected residues). For these 25% of samples which have non-detectable residues, but were treated, a decision has to be made about what level of residues should be assumed.

OPP has examined this problem and has developed methods for assigning values to these possible residues. OPP uses, as an initial step in the exposure assessment process, a residue value of $\frac{1}{2}$ LOD (or $\frac{1}{2}$ LOQ if an LOD has not been determined), for samples with no detectable residues if these samples have been treated with a pesticide. In this example, therefore, 25% of the samples in the residue data base would be assigned a residue of $\frac{1}{2}$ LOD, and 65% would be assigned a value of zero.

In summary, PCT data are primarily used in dietary risk assessment to determine how often to sample from distributions of residue data, and to determine the frequency of assigning positive residue values (e.g., ½ LOD) to non-detects in monitoring data.

IV. Short History of Pesticide Usage Estimation in OPP

EPA estimates percent crop treated based on a number of public and proprietary data sources. Pesticide usage data are available for both agricultural and non-agricultural use, although agricultural use data are usually available in much more detail than non-agricultural data. Only agricultural pesticide use data will be discussed below.

Beginning in 1987 pesticide usage data began to be available in electronic form. In addition pesticide usage data became available for more crops and in much more detail. Also, it was determined that the users of pesticide usage data usually either wished to know the average level of usage or the highest likely level of usage. This increased availability of data combined with a better understanding of the needs of the users of this data within OPP led to a revision of the process of providing pesticide usage estimates.

Procedures were developed that took advantage of increased data quantity and quality, and the ability to manipulate these data electronically. These procedures have been used, with periodic improvements, from approximately 1993 until the present. These analytical approaches and the output obtained are described in more detail below. Estimates include parameters for the “likely average” and the “likely maximum” for a crop/pesticide combination in a typical recent year.

A. Quality, Quantity and Availability of Pesticide Usage Data

This section describes the source and reliability of pesticide use data that EPA uses to estimate percent crop treated. Most pesticide usage data used by EAB are collected by either survey or census. EAB does not normally carry out surveys itself, but instead obtains data from various public and private sources to estimate percent crop treated.

1. Survey and census data

Surveys are the primary sources of detailed, quantitative information reflecting actual product use-patterns of pesticide users. In many cases, surveys collect information from a statistically representative sample of pesticide users, and then use information about the whole community of pesticide users to infer total use patterns. Most commonly, surveys are done at the state level, which can then be aggregated to the national level. In rare cases (California is the only current example), census data are available which represent a comprehensive accounting of all agricultural pesticide users in a given area or state. Some surveys are performed regularly (annually or bi-annually) and are available for a number

of years on numerous sites in numerous locations.⁷ Other surveys are conducted for specific industries or user groups and may be available only for a single, fixed period of time and location.

Survey and census data have several major advantages and some disadvantages. Surveys and censuses contain detailed, quantitative information – representative of the whole population of pesticide users – which can be used in quantitative analytical tools. By collecting data directly from users, these data collection methods may also generate the most accurate data reflecting actual pesticide use. Moreover, in many cases, one can also obtain the range within which estimates are statistically valid (e.g., confidence intervals).

The main disadvantage of survey data is that surveys and censuses can be quite costly. Not only are they costly to produce, but they also require time and effort from people who participate as survey respondents. There are few reliable ways to assure the reliability of a user's response, short of detailed interviews with respondents, so the most reliable and extensive surveys are generally the most costly. Unfortunately, relative to the wide range of sites in which pesticides are used, the number of surveys is limited. Proprietary data is principally purchased by pesticide producers and marketers, and these companies are usually interested in crops which could provide a large market for pesticides. Such crops tend to be either large acreage crops or smaller acreage crops that have high rates of pesticide use. Public sector surveys survey large acreage and high value crops and also survey crops that are grown by well organized user groups that understand the value of this data.

B. Quantitative Usage Analysis (QUA)

1. Existing QUA Program

For a given agricultural chemical the OPP develops a Quantitative Usage Analysis (QUA) that provides estimates – for each site where the chemical is used – of the following information regarding actual use: acres grown annually, percent of crop or units treated (both average and maximum), total annual use (average and maximum), application rates, number of applications, and states where the pesticide is most used. For non-agricultural sites, data are more limited and the QUA usually lists only the site and the national annual use of the pesticide on that site.

2. Present Methodology for Estimates of Percent Crop Treated

Generating a QUA is a procedure involving several steps and this section provides an overview of those steps.

⁷ As an example for the 2000 crop year NASS surveyed 35 vegetable crops. For seven of these crops data was collected for both fresh market and processing production. State coverage varied from one to 12 states depending on the crop. For fresh sweet corn 12 states were surveyed while for spinach-for-processing only one state was surveyed.

An analyst generates a list of alternative names for a given chemical and the sites (crops and other situations) on which it may legally be used. This information is contained in EPA's Reference Files System (REFS)⁸, a data base of pesticide names, product brands, use sites, target pests, and other similar data.

Using the list of names and sites as input, an EPA analyst then uses an automated program to generate a preliminary QUA. The *Automated Quantitative Usage Analysis (AQUA)* is a computer program that extracts – for a given chemical – quantitative use-related data from EPA's pesticide use data bases.⁹

3. Estimating the Likely Average

The AQUA program first standardizes data to generate consistent crop, product, and pesticide identifiers that often vary across different sources. For example, some sources name pesticides by their brand name and others name them by their active ingredient. Specific crop names often differ across sources. As an example, one source of usage data provides data for "cherries", another source provides data for "sweet cherries" and "tart cherries". The list of states covered by the two sources also differs and sometimes also varies between years for a single source. The analyst has to decide how differing estimates are to be modified to be consistent. Then, for agricultural pesticides, the program extracts *state-level estimates* of percent of crop treated, application rates, and number of applications from each source for each of the past 10 years if such data are available. For some crops, more than 10 years of data are available and could be used, if such use would improve the quality of the estimates. Estimates from different sources are then averaged *within each state*, with weights used that reflect the size of the survey sample from a given source. Surveys with larger sample populations in a particular state are given greater weight since the confidence limits around the estimates from those surveys are narrower. Yearly state estimates are then aggregated across states to generate national estimates (for each of several years) of quantitative use-related data. Although ultimately the program reports usage at a national level this procedure is used because of inconsistent state coverage among sources and within sources from year to year.

8

REFS will soon be replaced by OPPIN which will link a number of EPA pesticide related databases.

9

Not all the EPA sources described above provide their data in electronic form, and manually entering all data into electronic format is not feasible for a number of reasons (e.g., volume and heterogeneity of data, differences in definitions across sources, narrative nature of some sources). Some but not all of the non-electronic data has been entered into the system. For crops where many of the use data are not in electronic form, the QUA generated by AQUA is preliminary and must be supplemented manually with data from printed sources. At this time, most data obtained by the Agency is available in electronic form, although significant data transformation may be required to make it compatible with other data.

The estimates for the different years are then averaged using larger weights for more recent years, based on the observation that recent years are more representative of current normal use patterns than earlier year estimates.¹⁰ The weighting function is:

$$C_{est} = \left(\sum_{i=0}^9 (c_{t-i} \times 0.8^i) \right) \div \left(\sum_{i=0}^9 0.8^i \right)$$

where: C_{est} is the weighted estimate

0.8 is the yearly weighting factor

c_t is the observed usage level for the t^{th} year.

This weighting scheme was chosen using the following reasoning. Use of a pesticide on a crop varies from year to year and may also have a general trend up or down. It is likely that some of the factors influencing use of an individual pesticide at the present time, such as the price of competitive products and farmer's knowledge of the characteristics of the pesticide, are more likely to be similar to the factors for recent years than for less recent years. Other factors such as weather, while possibly very important influences on the usage of a pesticide, may be considered to be random across years. This suggested that a weighting scheme that gives more weight to recent years might be appropriate. Only if there are factors affecting use of the pesticide that are inversely correlated between two consecutive years will this weighting scheme yield worse estimates than weighting all years equally.¹¹ The size of the weights was chosen by examining the "reasonableness" of predictions obtained using different weights.

The final result is a national estimate of percent crop treated on each crop that includes the following parameters:

- (a) number of U.S. acres grown
- (b) average number of acres treated
- (c) average percent of crop treated

Estimates of acres grown and acres treated are calculated using the weighting procedure. The percent crop treated is calculated from the two estimates.¹²

¹⁰ A period of 10 years was chosen because the 10th year has a weight of approximately 3% in the value of C_{est} and earlier years would provide little additional contribution to the estimate.

¹¹ The statement applies only to the estimate. Nothing is being said here about the confidence limits around that estimate.

¹² Some sources report only acres grown and percent crop treated. Acres treated is calculated from these numbers so the

(continued...)

4. Estimating the Likely Maximum

In addition to the estimates of weighted averages, the AQUA program also estimates the likely maximum number of acres treated, and hence, maximum percent of crop treated.

For these three additional parameters, one option would be to report the maximum *observed* level in the sample period. This option was not chosen because a particular maximum observed during the survey period may not represent conditions of possible high pest infestation that would lead to the maximum likely pesticide use.

The AQUA estimates a *probable (or likely)* maximum percent crop treated. This likely maximum is calculated using statistical procedures that set a usage level below which actual usage might be expected to fall 95% of the time. This likely maximum is obtained by the following procedure:

- (1) Estimate a linear regression (OLS) of usage versus time to account for trends over time. Usage of a pesticide usually fluctuates from year to year. In addition, for many pesticides, there is an underlying trend, up or down. A linear regression can capture this trend.
- (2) Using the regression results, the procedure calculates an approximate 95% confidence interval around the linear trend at the last (most recent) observation. The difference between the expected value for the most recent year and the upper confidence limit is added to the weighted average generated earlier.¹³

All the output information discussed above is provided to the EPA analyst in the form of a spreadsheet. The analyst then adds non-electronic data and, in cases where crops and subsets of crops appear (such the example for cherries given earlier), reconciles the estimates for the different crop definitions. The estimate for tart cherries added to estimate for sweet cherries may not be equal to the estimate for cherries. The analyst has to decide how differing estimates are modified to be consistent.

5. Limitations

Several issues arise with respect to making statistical inferences from the pesticide data available to EPA. Ideally, pesticide use surveys are based upon a randomly drawn sample. Random sampling from the population of pesticide users assures that the data represent the actual distribution of pesticide use. Although some of EPA's sources are not strictly random (as explained below) the Agency believes them to

¹²(...continued)

information can be merged with data from other sources.

¹³ The regression estimate is not used to obtain average current use because individual point estimates based on linear regression for a short time series can be strongly influenced by outliers near either end of the regression series.

be functionally equivalent to random. For these sources, the survey designers maintain and update lists of farmers and other pesticide users. The lists are categorized according to farm size and location, and then used to identify cooperating pesticide users and gather data. Although sampling from a fixed list – or subset – of pesticide users may not be random in the strictest sense, given the very large size of these lists (thousands of respondents), the stratification process, and the verification procedures used by these data providers, EPA analysts consider the data to meet randomness criteria for statistical inference.

A second issue centers on the distributional properties of data generated by the estimation procedure. This is important because statistical inference from least squares regressions are often based on the assumption of normally distributed residuals, and this assumption influences the estimate of the 95% confidence interval. With a limited number of observations, one may not be able to verify the normality assumption. Therefore, EPA analysts address this issue by using a deterministic, rather than probabilistic, likely maximum where there are few observations. Mindful of these issues, EPA analysts try to use statistical methods that are appropriate to the available data in calculating likely maxima.

After the automated program has been run, EPA analysts refine the automated, preliminary QUA. Analysts may check estimates against other sources – including expert judgement inside and outside EPA, such as the information provided in the USDA Crop Profiles – to assure that estimates are consistent with observed patterns of pesticide use. In some cases, there may be evidence that the dynamics of pest control are changing rapidly due to such factors as newly introduced pesticides or changing pest populations. When evidence for such changes occurs, but is not reflected in historical use patterns, EPA analysts take this evidence into account in making (or annotating) estimates of percent crop treated.

6. Factors Influencing the Use of a Pesticide

There are a number of factors that influence the use of a pesticide. These include the acreage and value of crops to which the pesticide is applied, the occurrence of pests and the damage to the crop caused by these pests, the cost of the pesticide and cost of alternative pesticides and non-chemical control measures and the value of crop damage avoided by the use of the pesticide. Past pesticide use levels may be determined by these as well as other relevant factors not listed above. If it can be assumed that these factors, including yearly or other fluctuations, will continue in the future as in the past, estimates of future pesticide use can be made using only information on past pesticide use. These estimates can either be used by themselves or can be modified using knowledge about likely future values of any of the factors that influence pesticide use. There are sufficient time-series of pesticide use data to forecast percent crop treated, and methods for doing so are described in companion papers to this document.

Conclusions

The need to determine possible residues in food when a “maximum” percent of crop treated in future years is an important reason for developing refined methods of forecasting future pesticide use. The methodology presented in companion papers is expected to meet that requirement. Estimates obtained using past

information must always be critically examined and modified when other information suggests that such modification would be appropriate; such estimates, however, can serve as an excellent starting point.