SETAC PRESS

Environmental Toxicology and Chemistry, Vol. 21, No. 8, pp. 1578–1590, 2002

Printed in the USA
0730-7268/02 $9.00 + .00

# VALIDATION OF PESTICIDE ROOT ZONE MODEL 3.12: EMPLOYING UNCERTAINTY ANALYSIS

JOHN P. CARBONE,*† PATRICK L. HAVENS,‡ and WILLIAM WARREN-HICKS§
†Toxicology Department, Rohm and Haas Company, Spring House, Pennsylvania 19477-0904, USA
‡Global Environmental Chemistry Laboratory, Dow AgroSciences, 9330 Zionsville Road, Indianapolis, Indiana 46268, USA
§The Cadmus Group, 1920 Highway 54, Durham, North Carolina 27713, USA

(*Received* 19 *March* 2001; *Accepted* 21 *January* 2002)

**Abstract**—Computer models are being increasingly used to provide an efficient cost-effective means of evaluating the fate and behavior of chemicals in the environment. These models can be used in lieu of or in conjunction with field studies. Because of the increasing reliance on models for critical regulatory decision making, the need arose to assess the validity of regulatory models via an analysis of the correlation of model response estimates with measured data. In conjunction with the evaluation of the correlation of model response estimates and measured field data, a rigorous statistically based validation was also warranted. Because of the unique nature of the correlative exercise using modeled and measured data, standard statistical analyses, while informative, failed to encompass factors associated with the uncertainty of measured environmental fate data and potential model inputs. In an effort to evaluate this uncertainty, an initial sensitivity analysis was performed where key model input parameters for runoff and leaching simulations were identified. Once the sensitive input parameters were identified, a Monte Carlo–based preprocessor was developed whereby the sampling distributions of these parameters were used to propagate uncertainty in the input parameters into error in model predictions. Importantly, assumptions about parameter distributions for input into the Monte Carlo tool were made only after a formal detailed site-specific analysis of measured field data. Employing the functionality of the Crystal Ball® Pro development environment, the pesticide root zone model (PRZM) 3.12 was run iteratively for 500 trials, and model output was collated and analyzed. The model predictions were considered reasonably accurate for most regulatory requirements, and the model prediction error was considered acceptable.

**Keywords**—Uncertainty analysis Exposure modeling Model validation

## INTRODUCTION

In 1992, the U.S. Environmental Protection Agency (U.S. EPA) established a new paradigm for the evaluation of ecological risk that identified computer modeling as a fast and cost-effective exposure assessment tool compared to field studies. To address concerns regarding confidence in the use of modeling as a regulatory tool, the need arose to assess the correlation of model response estimates and measured data. In conjunction with the evaluation of the correlation of model response estimates and measured field data, a rigorous statistically based validation was also warranted. Inherent within any model validation process are uncertainties associated with the model code, accepted modeling paradigms, model input parameters, and the measured responses. For the purposes of this validation exercise, the uncertainty associated with the model code, accepted modeling paradigms, and measured responses were not addressed. Those uncertainties associated with model input parameters were, however, closely examined.

An approach for the evaluation of the PRZM 1.0 performance based on graphical comparisons of estimated and measured pesticide movement versus depth has been provided [1]. Observed and predicted response profiles demonstrated equivalent pesticide mass over time. In a similar qualitative approach, the use of PRZM 1.0 was advocated for estimating the magnitude of the depth and timing of the leading edge of pesticide movement but not the ability of the model to predict absolute concentrations [2]. The work of Leonard et al. [3] with the groundwater loading effects of agriculture management system (GLEAMS) have shown via graphical means that estimated values for mass of the parent and metabolites of concern were comparable to measured data within the variability of the measured data. Simulated and observed concentrations at depth in the soil at selected dates also closely corresponded. The modified chemical runoff and erosion from agriculture management system (CREAMS)/GLEAMS model output response has also been shown via graphical means to be comparable to measured data from field studies conducted in Finland [4].

An extensive discussion regarding the implementation and procedures for the validation of PRZM 1.0 and the risk of unsaturated/saturated transport and transformation of chemical concentrations (RUSTIC) model [5] have been presented [6]. The authors note that technical issues that require consideration prior to a validation exercise include, among others, a well-defined performance and acceptance criteria. The definition of performance criteria is a recurrent theme found throughout the model validation literature. Defining the performance criteria allows for the relative inaccuracies of model responses versus measured data. In conjunction with an unpublished U.S. EPA report regarding the validation status of the PRZM model, C.N. Smith et al. (1990) define performance criteria based on the purpose of the modeling analysis. For screening level analyses, a level of accuracy should be expected to be within an order of magnitude. For site-specific or higher-tiered modeling exercises, the authors suggest that a factor of two to four can be sufficient in certain instances but that a factor of less than two may be appropriate in others.

* To whom correspondence should be addressed (jcarbone@rohmhaas.com).

Statistical measures useful for the comparison of model estimates versus measured data should include paired comparisons of predicted and observed values in space and time, integrated comparisons relating to spatially or temporally composited data such as monthly or annual means or totals versus corresponding model estimates, and comparisons of cumulative frequency distributions of observed data and model predictions in stochastic situations [6]. Statistical measures for paired data and spatially and temporally integrated performance tests include descriptive statistics, error and regression analyses, and correlation coefficients. Plots of observed versus predicted values were also advocated as visual indications of agreement. The authors note that the variability associated with model input can be addressed by employing Monte Carlo analyses. Uncertainty analysis was discussed within the context of the approaches outlined by Carsel et al. [6].

The performance of PRZM 1.0, GLEAMS 1.8.54, and the pesticide component of the leaching estimation and chemistry model (LEACHM 1.0) have been compared on the basis of water mass balance, the transport of bromide, and the transport and degradation of the reactive solute aldicarb [7]. Objective criteria used to validate and compare the models included root-mean-square error, normalized objective function, and reduced error estimates. A description of model evaluation procedures based on graphical displays and statistical criteria have similarly been provided [8]. The statistical analyses discussed encompass the analysis of residual errors and the differences between observed and predicted values. These include maximum error, root-mean-square error, coefficient of determination, modeling efficiency, and coefficient of residual mass.

The factor-of-*f* technique and measures of goodness of fit have been employed to evaluate the predictive capabilities of PRZM and the aggregate model for field transport and transformation (AGGR) versus measured data [9,10]. The factor-of-*f* technique defined the validity criteria based on the regulatory function of the models (U.S. EPA, 1982, unpublished data) and reflects the 10× or 2× to 4× factor approach discussed previously. In addition, measures of goodness of fit were employed based on the approach described by Loague and Green [8]. Validation criteria have been discussed (B. Allen et al., 1990, unpublished data) in conjunction with the pesticide and industrial chemical risk analysis and hazard assessment modeling suite (PIRANHA) [11]. The validity criteria was defined such that the model was deemed valid if predictions were within a factor of two of the measured data at least 95% of the time. Both subjective (graphical) and objective (parametric statistics) criteria have been employed to compare simulated versus measured data [12,13]. Zacharias et al. [13] also presented robust quantitative techniques where the validity of the statistical procedures was not dependent on the assumption of a specific probability model of the population. For model validation, this distribution-free analysis was based on nonparametric techniques. Based on the study, GLEAMS and PRZM performed well in predicting pesticide mass but were less reliable in predicting pesticide concentration distributions in soil.

A stepwise process for the performance of model validation under the auspices of the European Union Environmental Research Programme has been described [14]. The models evaluated included PRZM 2.0, the pesticide component of LEACHM [15], and the variable leaching model (VARLEACH 2.0). The authors suggest the use of preliminary subjective graphical analysis in conjunction with objective statistical

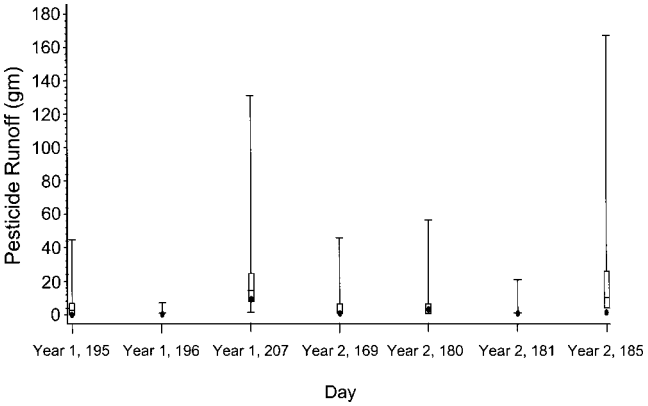


Fig. 1. Iowa, USA, corn site (IA2R) pesticide runoff mass. Monte Carlo uncertainty analysis with all variables sampled simultaneously. Box-and-whisker plots are model predictions; solid circles are measured values.

analyses encompassing tests for evaluation of overall model fit, degradational fit, and distributional fit. An approach for evaluating and comparing the outcome from multiple models has been described [16]. The paradigm for the validation procedure using field and lysimeter data sets entails a calibration or model error minimization step and a subsequent model test versus independent data. Models were evaluated using subjective graphical and objective statistical measures. A multitiered approach to model validation using both graphical and statistical testing has been advocated [17]. The components of the multitiered approach include a parameterization of the model using only independently measured parameters, validation of water movement and water content of the soil, validation of conserved solute movement, validation of pesticides fate in the soil using parameters reflective of independently measured fate information, and finally validation of pesticide leaching in terms of comparisons of model predictions with respect to patterns and orders of magnitude of occurrences.

It has been widely recognized that model input parameter variability in both the spatial and the temporal sense have significant impact on model response output. Initial efforts to capture the effects of variability in the temporal sense focused on the consequences of the long-term climatologic variation [1]. Model parameterization reflecting, as examples, adsorption/desorption behavior and degradation were held constant, and variability was introduced via the climate data. Thus, the model could be used to infer the behavior of pesticide movements under varied climatologic conditions. The drawback of this approach is that it does not account for the inherent spatial variability of, for example, degradation and adsorption/desorption that is associated with field studies and the behavior of chemicals under environmentally relevant conditions. Subsequent work [6] employed Monte Carlo techniques to account for the spatial variability noted with respect to soils in field studies. Statistical transformation paradigms, particularly the Johnson transformation technique and log transformations, were used to convert original random variable, nonnormal soil characterization data derived from a national database into normal distributions. Monte Carlo techniques were used to vary field capacity, wilting point and organic matter, soil hydrologic group, weather year, and degradation rate (triangular distribution). While no validation exercise was performed, results pertinent to the current discussion demonstrated cumulative probability distribution functions for annual pesticide

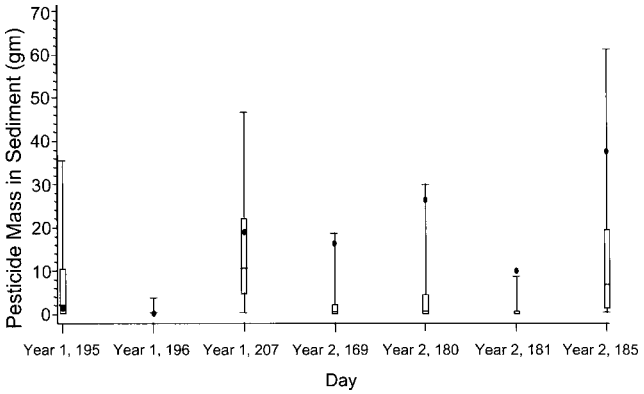


Fig. 2. Iowa, USA, corn site (IA2R) pesticide mass in sediment. Monte Carlo uncertainty analysis with all variables sampled simultaneously. Box-and-whisker plots are model predictions; solid circles are measured values.

movement and the influence of model input variability on model response estimates.

A stochastic approach for the pesticide leaching model (PELMO) has been developed and termed MCPELMO [18]. The MCPELMO is based on the deterministic PELMO model but in addition has an incorporated shell allowing for the stochastic simulations for a number of geographically diverse regions. The authors concluded that the temporal variation afforded by the weather data had a chemical specific impact on model outcomes and also noted that spatial variability in soil may be more influential than the temporal variability on model estimates. A validation methodology for the pesticide leaching and accumulation in soil model (PESTLA) has been developed that includes both statistical and graphic analyses [19]. The factor-of-*f* approach [10] was again used. The authors also discuss an approach based not on the derivation of the predicted mean from a deterministic simulation but on the estimate of the predicted mean and predicted standard deviation of model output based on a stochastic sampling approach for model input variables. In addition to typical subjective and objective measures, the authors advocate the use of comparisons of cumulative probability distribution functions of observed and predicted data. As a follow-up to the previous study, van den Bosch and Boesten [20] provide an approach for validation of the PESTLA model where both graphical and objective statistics were utilized. Estimated peak concentration values were log transformed, and the average and standard deviation of the average were calculated. The confidence intervals around the average of the measured peak concentration of one or two times the standard deviation were used to provide a measure of the uncertainty associated with measured data. The factor-of-*f* approach was applied where $f = 2$ and $f = 5$.

Eckhardt and Wagenet [21] evaluated the consequences of the inherent variability in soil hydrology and chemical applications and the uncertainty of measurements of soil and chemical properties on the leaching potential of atrazine. The model employed was the pesticide component of LEACHM [15]. Following a calibration step, the impact on model output re-

Table 1. Daily comparison of model predictions and measured values: Iowa, USA, corn (IA2R)[a]

| Date | Runoff variables | Measured values | Predictions exceeding the measured value (%) |
|---|---|---|---|
| 1992: Day 195 | Runoff volume ($m^3$) | 319.200 | 69.2 |
| | Sediment yield (kg) | 1,543.000 | 59.6 |
| | Pesticide runoff mass (g) | 0.330 | 82.8 |
| | Pesticide mass in sediment (g) | 1.380 | 59.6 |
| 1992: Day 196 | Runoff volume ($m^3$) | 21.900 | 38.6 |
| | Sediment yield (kg) | 73.000 | 30.0 |
| | Pesticide runoff mass (g) | 0.029 | 43.8 |
| | Pesticide mass in sediment (g) | 0.062 | 34.8 |
| 1992: Day 207 | Runoff volume ($m^3$) | 3,170.800 | 53.0 |
| | Sediment yield (kg) | 10,022.000 | 68.6 |
| | Pesticide runoff mass (g) | 9.000 | 67.4 |
| | Pesticide mass in sediment (g) | 18.900 | 30.6 |
| 1993: Day 169 | Runoff volume ($m^3$) | 688.700 | 45.0 |
| | Sediment yield (kg) | 16,980.000 | 0.0 |
| | Pesticide runoff mass (g) | 0.792 | 60.6 |
| | Pesticide mass in sediment (g) | 16.100 | 0.2 |
| 1993: Day 180 | Runoff volume ($m^3$) | 185.700 | 36.4 |
| | Sediment yield (kg) | 2,619.000 | 8.2 |
| | Pesticide runoff mass (g) | 3.270 | 27.2 |
| | Pesticide mass in sediment (g) | 26.200 | 1.0 |
| 1993: Day 181 | Runoff volume ($m^3$) | 11.700 | 36.4 |
| | Sediment yield (kg) | 4,208.000 | 0.0 |
| | Pesticide runoff mass (g) | 0.046 | 38.4 |
| | Pesticide mass in sediment (g) | 9.550 | 0.0 |
| 1993: Day 185 | Runoff volume ($m^3$) | 1,604.500 | 40.4 |
| | Sediment yield (kg) | 14,272.000 | 2.4 |
| | Pesticide runoff mass (g) | 0.477 | 94.6 |
| | Pesticide mass in sediment (g) | 37.000 | 10.4 |

[a] Studies are identified by the state name, number of study, and the letter R, indicating a runoff study.

Table 2. Daily comparison of model predictions and measured values: Georgia, USA, cotton (GA1R)[a]

| Date | Runoff variables | Measured values | Predictions exceeding the measured value (%) |
|---|---|---|---|
| 1989: Day 220 | Pesticide runoff mass (g) | 3.620 | 44.8 |
| 1989: Day 238 | Pesticide runoff mass (g) | 8.330 | 60.6 |
| 1989: Day 243 | Pesticide runoff mass (g) | 1.320 | 46.0 |
| 1989: Day 274 | Pesticide runoff mass (g) | 0.012 | 57.0 |

[a] See footnote in Table 1.

sponse due to the uncertainty associated with two critical transport parameters, unsaturated hydraulic conductivity and degradation rates, were evaluated. Additionally, the effect on model output response due to the uncertainty associated with the spatially heterogeneity of pesticide applications was evaluated. The effects of spatial variability in the hydraulic conductivity of the soil and the uncertainty of degradation rates below the root zone were represented through discrete sampling from probability density functions. The probability density functions from which the discrete samples for hydraulic conductivity and pesticide degradation rate were selected were defined on the basis of empirical data. The results clearly indicate that the uncertainty associated with hydraulic conductivity, pesticide degradation rate, and application rates substantially influenced model results.

Haan et al. [22] provide a discussion regarding the evaluation of model performance in a situation where no data are observed on the quantities being modeled. Measured data was therefore unavailable to assist in input parameter estimation or model calibration. The procedures employed include the conduct of a sensitivity analysis on model input parameters, generation of probability distributions functions of input parameters, generation of probability distributions of model output based on the input probability density functions, and finally the use of the output probability distributions to assess the model. The authors advocate that the use of the measured data be held in reserve so that the measured data do not enter the parameter estimation process. The authors note that the measured responses are also subject to uncertainty. The uncertainty can also be quantified using probability density functions. The overlap of the model output and measured response probability density functions can, if available, be utilized to evaluate model performance.

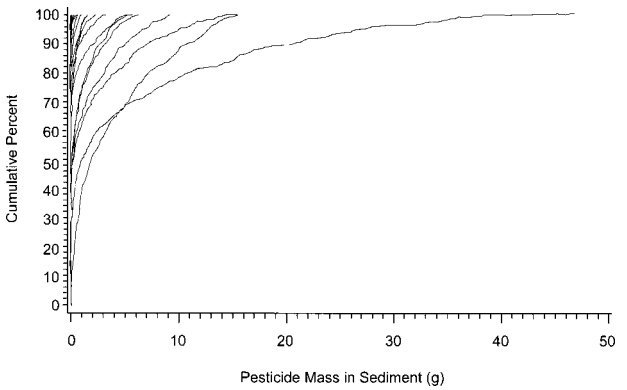


Fig. 3. Iowa, USA, corn site (IA2R) pesticide mass in sediment. Monte Carlo uncertainty analysis with all variables sampled simultaneously. Model predictions for 1992 with no measured runoff. Each plot on the graph is an individual day.

## APPROACH

As an initial step to evaluate the impact of uncertainty on model performance, a sensitivity analysis was performed using an approach based on that of Plackett and Burman [23] where key model input parameters for runoff and leaching simulations were identified [24]. The Plackett–Burman analysis results were used to develop a final list of most influential model inputs for evaluation using Monte Carlo techniques. For each of the final parameters, an attempt was made to define the nature of the sampling distribution for use in the Monte Carlo uncertainty analysis.

The results of the Plackett–Burman efforts to define the most sensitive PRZM 3.12 model input parameters are presented here. The following were defined as the most sensitive inputs, in descending order of sensitivity, affecting leaching endpoints: bulk density (representing available water-holding capacity), time-zero application rate variations, decay rate (soil layer 1), decay rate (soil layers 2 and 3), maximum rooting depth, adsorption coefficient ($K_d$ illustrating in-field variability), curve numbers 1 and 2, and pan evaporation factor.

For runoff analyses, the following were defined as the most sensitive inputs, again in descending order of sensitivity: curve numbers 1 and 2, adsorption coefficient ($K_d$ illustrating in-field variability), decay rate (soil layer 1), bulk density (representing available water-holding capacity), foliar decay rate (the assumption is that the foliar decay rate exhibits a similar coefficient of variation and distribution as that for soil degradation in soil layer 1), plant uptake factor (did not pursue uncertainty analysis for this input), management factor 2, and time-zero application rate variations.

For each of the preceding sensitive PRZM 3.12 input parameters, the nature of the sampling distribution for use in the Monte Carlo uncertainty analysis was defined [25]. Specific criteria were developed for establishing these sampling distributions. These criteria were used to ensure consistency in the procedures for evaluating model prediction error across sites. The criteria also ensured that the sampling distributions represented, to the degree possible, the actual site-specific uncertainty and variation in the parameters. Therefore, the criteria effectively increased the confidence that the Monte Carlo uncertainty analysis results reflect the true model prediction error associated with a specific site and parameter set.

A set of interface tools were built to implement the Monte Carlo sampling and analysis techniques with the PRZM 3.12 model [26]. The software chosen was the Crystal Ball Pro package, manufactured by Decisioneering (Denver, CO), along with some additional FORTRAN programs.

Table 3. Distribution of model predictions for days with no measured values: Iowa, USA, corn (IA2R)[a]

| | Percentage of all data with zeros | Value of the model prediction at selected percentiles | | | | |
|---|---|---|---|---|---|---|
| | | 25% | 50% | 75% | 90% | 95% |
| Runoff volume ($m^3$) | | | | | | |
| 1992 | 69.00 | 0.0 | 0.0 | 22.1 | 411.8 | 979.2 |
| 1993 | 70.00 | 0.0 | 0.0 | 4.9 | 238.0 | 674.3 |
| Sediment yield (kg) | | | | | | |
| 1992 | 69.00 | 0.0 | 0.0 | 8.6 | 646.7 | 3,797.5 |
| 1993 | 70.00 | 0.0 | 0.0 | 0.8 | 202.7 | 1,523.0 |
| Pesticide runoff mass (g) | | | | | | |
| 1992 | 70.00 | 0.0 | 0.0 | 0.0 | 1.0 | 3.4 |
| 1993 | 75.00 | 0.0 | 0.0 | 0.0 | 0.2 | 1.0 |
| Pesticide mass in sediment (g) | | | | | | |
| 1992 | 74.00 | 0.0 | 0.0 | 0.0 | 0.2 | 1.3 |
| 1993 | 83.00 | 0.0 | 0.0 | 0.0 | 0.0 | 0.3 |

[a] See footnote in Table 1.

## RESULTS

For the uncertainty analysis, model predictions were compared to actual groundwater or surface water measurements. Model predictions of interest for runoff analyses include the following: runoff volume ($m^3$/d), sediment yield (kg/d), pesticide runoff mass (g/d), and pesticide mass in sediment (g/d). For leaching simulations, the model predictions of interest include pesticide mass in soil (g/kg), pesticide in pore water (μg/L), and bromide in pore water (mg/L).

Graphical and tabular information have been generated for comparing the measured field information and model predictions. The ability of the model to predict runoff and leaching factors on a daily basis was evaluated. For the current exercise, no attempt was made to scale up the analyses to monthly or yearly comparisons. In addition, no attempt was made to compare results across sites. By using the smallest time-scale available (days), the effect of uncontrolled temporal and spatial influences on the comparison results was reduced. However, the error in model parameterization was incorporated into the comparison through the use of Monte Carlo analysis and generation of a prediction distribution, which represents the uncertainty in model predictions, conditional on the understanding and measurement of the model input parameters.

### *Variability about the measured data*

Because the measured data used in the studies consisted of proprietary information generated in support of product registrations, it was required that the current analyses be conducted in a manner that ensured data confidentiality. Additionally, in order to limit modeler bias, model simulations and all subsequent analyses were conducted without access to the original study reports. Because of these constraints, measured data for the runoff studies were provided as a mean of the analytical replicates for each specific sample date. For the leaching analyses, minimum, maximum, and mean values were provided for each sample date. For the leaching studies typical minimum and maximum pesticide soil core concentrations ranged from 25.4 to 28.4 μg/kg (mean 27.4 μg/kg) and 25.4 to 95.8 μg/kg (mean 67 μg/kg). Typical minimum and maximum pesticide soil pore-water concentrations ranged from 0.05 to 0.928 μg/L (mean 0.36 μg/L) and 0.121 to 3.1 μg/L (mean 1.16 μg/L). Typical minimum and maximum bromide soil pore-water concentrations ranged from 46 to 90 mg/L (mean 68 mg/L) and 0.25 to 100 mg/L (mean 43.4 mg/L). Importantly, significant percentages of the measured data exhibited no variability with the minimum, maximum, and mean values reported to be at the level of quantification (LOQ). Because of these limitations (data accessibility), the requirement for confidentiality, and the associated limited variability, uncertainties associated with daily field measurements were not included in the analysis.

Sources of uncertainty associated with the measured data include the spatial variation within the field site and sample variation, that is, the variability associated with multiple soil core, soil pore-water, or runoff measurements and the variability associated with analytical methods. One would anticipate that the variability associated with the former two sources could be considerable, while that for the latter would be minimal. In view of the uncertainty associated with the measured data, the current analysis is therefore conservative. Allowing

Table 4. Distribution of model predictions for days with no measured values: Georgia, USA, cotton (GA1R)[a]

| | Percentage of all data with zeros | Value of the model prediction at selected percentiles | | | | |
|---|---|---|---|---|---|---|
| | | 25% | 50% | 75% | 90% | 95% |
| Pesticide runoff mass (g) | | | | | | |
| 1989 | 78.00 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 |

[a] See footnote in Table 1.

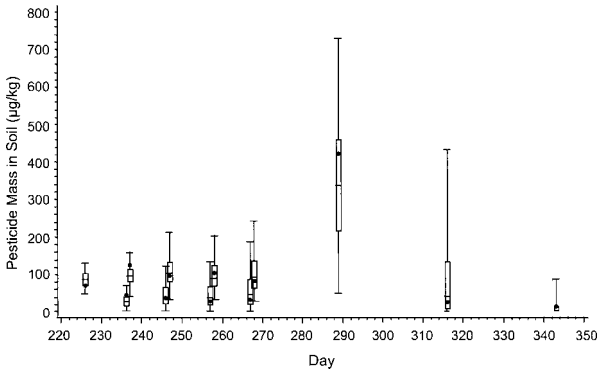


Fig. 4. Georgia, USA, sweet corn site (GA1L) pesticide mass in soil 0 to 15 cm (year 1). Monte Carlo uncertainty analysis with all variables sampled simultaneously. Box-and-whisker plots are model predictions; solid circles are measured values.

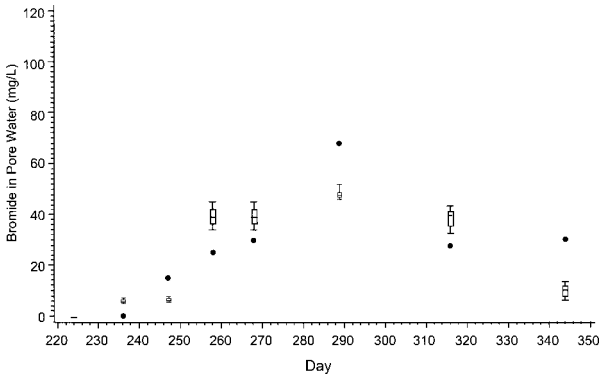


Fig. 6. Georgia, USA, sweet corn site (GA1L) bromide in pore water 180 cm (year 1). Monte Carlo uncertainty analysis with all variables sampled simultaneously. Box-and-whisker plots are model predictions; solid circles are measured values.

for the uncertainty associated with the measured values would increase the overlap of model outcome distributions with distributions representative of the measured data.

Box-and-whisker plots for representative analyses are presented to graphically compare measured field information and model predictions. For each day that a field measurement was available, a box-and-whisker plot of the model predictions was overlain on a marker for the value of the field measurement. The box-and-whisker plot displays the lowest, 25th-percentile, median, 75th-percentile, and maximum value for the model predictions (based on 500 iterations). Examination of the plot shows the relative number of model predictions below and above the measured value as well as the relationship of the measured value to specific statistics of the prediction distribution (median, 25th percentile, and so on). Model accuracy was evaluated by examining the percentage of model predictions below and above the measured value. When the measured field value was shown to be in the general center of the prediction distribution, the model can be considered reasonably predictive. When the measured value was seen in the lower or upper portions of the prediction distribution, the model can be considered less accurate (within the bounds of uncertainty) but acceptable given the variability in the model parameters. If the entire prediction distribution is above or below the measured value, the model may be considered inaccurate for that day. However, in some circumstances this latter scenario does not hold. In particular, for very small measured values (near the LOQ), the model is frequently shown to predict into the range below the detection level or only slightly above the detection value. By convention the measured value was always reported as one-half the LOQ. Subsequently, model outcomes that predict below one-half the LOQ may show a disagreement between the measured and the predicted values. Practically, however, the model and measured values show good agreement in this case, with the comparison truncated at the limit of quantification. In addition, the model predictions and measured values were frequently found to significantly disagree when the values were above but very near the LOQ. While the measured and predicted values differ, in many cases the difference was not practically relevant given the small magnitude of the recorded numbers. Measured values near the limit of quantification provide a basis for judging the ability of the model to predict pesticide concentrations at low magnitude.

Tables are provided that show the percentage of model predictions (out of 500 iterations) that exceed the measured field value. A 100% exceedence indicates that the model predicted high, and a 0% exceedence indicates that the model predicted low. Again, the magnitude of the measured values should be used to evaluate the significance of these extreme scenarios.

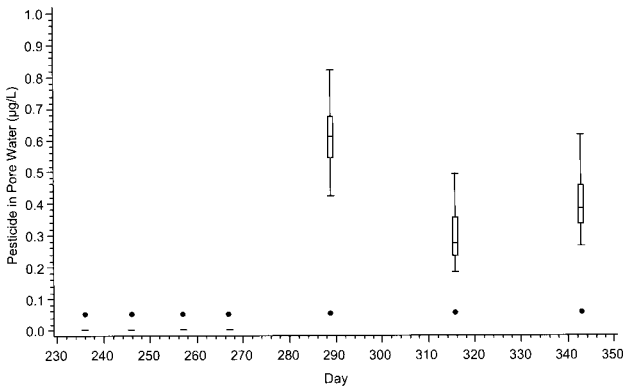


Fig. 5. Georgia, USA, sweet corn site (GA1L) pesticide in pore water 90 cm (year 1). Monte Carlo uncertainty analysis with all variables sampled simultaneously. Box-and-whisker plots are model predictions; solid circles are measured values.

*Runoff: Iowa, USA, corn (IA2R)*

The Iowa site (IA2R) was a field-scale runoff study with natural rainfall conducted on corn in a large (7 ha) watershed in Oskaloosa, Iowa (Mahaska County). Figures 1 and 2 present the Monte Carlo results for pesticide mass in runoff and sediment for site IA2R. Seven days, spanning two years, had measured runoff values. Table 1 presents information on the number of predictions exceeding the measured value. For runoff volume, all measured values fell within the interquartile range (between the 25th and the 75th percentile of the prediction distribution) of the model predictions, indicating that the model was very reliable. For sediment yield, measured values fell within the interquartile range for 3 d, within the bounds of the distribution for 2 d, and outside the bounds of the distribution for 2 d. For dissolved pesticide runoff mass, five of seven measured values fell within the interquartile range, and the remaining measurements fell within the bounds of the distribution. For pesticide mass in sediment, three measured values fell within the interquartile range of the predictions, three fell within the bounds of the predictions, and one fell outside the bounds of the model predictions.

Table 5. Daily comparison of model predictions and measured values: Georgia, USA, sweet corn (GA1L)[a]

| Julian day | Soil horizon depth (cm) | Measured pesticide (μg/kg) | Percentage exceedence | Lysimeter depth | Measured pesticide (μg/L) | Percentage exceedence | Measured bromide (mg/L) | Percentage exceedence |
|---|---|---|---|---|---|---|---|---|
| 224 | | | | 90 | | | 0.25 | 0 |
| 224 | | | | 180 | | | NA[b] | NA |
| 224 | | | | 270 | | | 0.38 | 0 |
| 226 | 0–15 | 67 | 80.2 | 90 | | | | |
| 226 | 15–30 | 0.5 | 0 | 180 | | | | |
| 226 | 30–45 | 0.5 | 0 | 270 | | | | |
| 226 | 45–60 | NA | NA | | | | | |
| 236 | 0–15 | 40.33 | 17.4 | | | | | |
| 236 | 15–30 | 0.5 | 98.2 | 90 | 0.05 | 0 | 3.58 | 100 |
| 236 | 30–45 | 0.5 | 68.2 | 180 | 0.05 | 0 | 0.25 | 100 |
| 236 | 45–60 | NA | NA | 270 | 0.05 | 0 | 0.38 | 0 |
| 237 | 0–15 | 122.8 | 15.4 | 360 | 0.05 | 0 | | |
| 237 | 15–30 | 0.5 | 97 | | | | | |
| 237 | 30–45 | 0.5 | 65.6 | | | | | |
| 237 | 45–60 | NA | NA | | | | | |
| 246 | 0–15 | 35.77 | 52.8 | 90 | 0.05 | 0 | | |
| 246 | 15–30 | 3.53 | 61.1 | 180 | 0.05 | 0 | | |
| 246 | 30–45 | 0.5 | 47.8 | 270 | 0.05 | 0 | | |
| 246 | 45–60 | 0.5 | 2.6 | 360 | 0.05 | 0 | | |
| 247 | 0–15 | 94.33 | 57.7 | 90 | | | 43.42 | 100 |
| 247 | 15–30 | 1.3 | 73.7 | 180 | | | 15.13 | 0 |
| 247 | 30–45 | 0.5 | 45.9 | 270 | | | 0.25 | 0 |
| 247 | 45–60 | 0.5 | 2.6 | | | | | |
| 257 | 0–15 | 27.4 | 60.9 | 90 | 0.05 | 0 | | |
| 257 | 15–30 | 0.71 | 91.1 | 180 | 0.05 | 0 | | |
| 257 | 30–45 | 0.5 | 69.2 | 270 | 0.05 | 0 | | |
| 257 | 45–60 | 0.5 | 43.7 | 360 | 0.05 | 0 | | |
| 258 | 0–15 | 102.4 | 39.5 | 90 | | | 93.33 | 88.6 |
| 258 | 15–30 | 12.33 | 42.5 | 180 | | | 25.13 | 100 |
| 258 | 30–45 | 0.5 | 68.2 | 270 | | | 0.25 | 100 |
| 258 | 45–60 | 0.5 | 42.7 | | | | | |
| 267 | 0–15 | 30.5 | 62.8 | 90 | 0.05 | 0 | | |
| 267 | 15–30 | 0.5 | 78.7 | 180 | 0.05 | 0 | | |
| 267 | 30–45 | 0.5 | 56.1 | 270 | 0.05 | 0 | | |
| 267 | 45–60 | NA | NA | 360 | 0.05 | 0 | | |
| 268 | 0–15 | 80.93 | 58.7 | 90 | | | 113.3 | 11.4 |
| 268 | 15–30 | 0.5 | 76.9 | 180 | | | 30.13 | 100 |
| 268 | 30–45 | 0.5 | 54.7 | 270 | | | 0.25 | 100 |
| 268 | 45–60 | NA | NA | | | | | |
| 289 | 0–15 | 422 | 32.4 | 90 | 0.05 | 100 | 51.2 | 0 |
| 289 | 15–30 | 13.07 | 81.6 | 180 | 0.32 | 0 | 68 | 0 |
| 289 | 30–45 | 0.5 | 95.7 | 270 | 0.08 | 0 | 33 | 100 |
| 289 | 45–60 | 0.5 | 97.6 | 360 | 0.05 | 0 | | |
| 316 | 0–15 | 22.47 | 59.5 | 90 | 0.05 | 100 | 59 | 0 |
| 316 | 15–30 | 2.96 | 68.8 | 180 | 0.05 | 0 | 27.5 | 100 |
| 316 | 30–45 | 0.5 | 70.8 | 270 | 0.05 | 0 | 12 | 100 |
| 316 | 45–60 | 0.5 | 78.1 | 360 | 0.05 | 0 | | |
| 343 | 0–15 | 10.11 | 26.7 | 90 | 0.05 | 100 | | |
| 343 | 15–30 | 2.18 | 53.6 | 180 | 0.05 | 0 | | |
| 343 | 30–45 | 0.85 | 55.9 | 270 | 0.48 | 0 | | |
| 343 | 45–60 | 0.5 | 61.1 | 360 | 0.05 | 0 | | |
| 344 | | | | 90 | | | 8.37 | 0 |
| 344 | | | | 180 | | | 30 | 0 |
| 344 | | | | 270 | | | 20 | 100 |
| 371 | 0–15 | 5.28 | 11.6 | 90 | 0.05 | 100 | 20 | 0 |
| 371 | 15–30 | 2.96 | 32.2 | 180 | 0.05 | 0 | 14 | 0 |
| 371 | 30–45 | 1.27 | 36.8 | 270 | 0.33 | 0 | 9 | 100 |
| 371 | 45–60 | 0.5 | 50 | 360 | 0.05 | 0 | 0.05 | 0 |
| 399 | 0–15 | 3.91 | 0 | 90 | 0.12 | 97.8 | 10 | 0 |
| 399 | 15–30 | 2.45 | 9.6 | 180 | 0.05 | 0 | 12 | 0 |
| 399 | 30–45 | 0.88 | 24.8 | 270 | 0.15 | 0 | 8 | 0 |
| 399 | 45–60 | 0.75 | 32 | 360 | 0.05 | 0 | 0.05 | 0 |
| 433 | 0–15 | 3.27 | 0 | | | | | |
| 433 | 15–30 | 1.1 | 1.6 | | | | | |
| 433 | 30–45 | 0.5 | 14.4 | | | | | |
| 433 | 45–60 | 0.5 | 21.2 | | | | | |
| 434 | | | | 90 | 0.05 | 0 | 2.8 | 0 |
| 434 | | | | 180 | 0.05 | 0 | 10.33 | 0 |
| 434 | | | | 270 | 0.17 | 0 | 5.5 | 0 |
| 434 | | | | 360 | 0.05 | 0 | 0.05 | 0 |

Table 5. Continued

| Julian day | Soil horizon depth (cm) | Measured pesticide (μg/kg) | Percentage exceedence | Lysimeter depth | Measured pesticide (μg/L) | Percentage exceedence | Measured bromide (mg/L) | Percentage exceedence |
|---|---|---|---|---|---|---|---|---|
| 468 | 0–15 | 3.37 | 0 | | | | | |
| 468 | 15–30 | 0.94 | 0 | | | | | |
| 468 | 30–45 | 0.5 | 2 | | | | | |
| 468 | 45–60 | 0.5 | 8.8 | | | | | |
| 469 | | | | 90 | 0.13 | 0 | NA | NA |
| 469 | | | | 180 | 0.05 | 0 | 6.7 | 0 |
| 469 | | | | 270 | 0.23 | 0 | 3.33 | 0 |
| 469 | | | | 360 | 0.1 | 0 | 0.1 | 0 |
| 496 | 0–15 | 4.96 | 0 | | | | | |
| 496 | 15–30 | 1.85 | 0 | | | | | |
| 496 | 30–45 | 0.5 | 1.4 | | | | | |
| 496 | 45–60 | 0.5 | 7.6 | | | | | |
| 497 | | | | 90 | 0.3 | 0 | 1.1 | 0 |
| 497 | | | | 180 | 0.05 | 0 | 0.7 | 0 |
| 497 | | | | 270 | 0.31 | 0 | 2.83 | 0 |
| 497 | | | | 360 | 0.16 | 0 | 0.16 | 0 |
| 525 | 0–15 | 2.31 | 0 | | | | | |
| 525 | 15–30 | 0.5 | 0 | | | | | |
| 525 | 30–45 | 0.5 | 0 | | | | | |
| 525 | 45–60 | NA | NA | | | | | |
| 526 | | | | 90 | 0.21 | 0 | 0.25 | 0 |
| 526 | | | | 180 | 1.16 | 0 | 2.5 | 0 |
| 526 | | | | 270 | 0.25 | 0 | 2.7 | 0 |
| 526 | | | | 360 | 0.05 | 0 | 0.05 | 0 |
| 554 | 0–15 | 0.5 | 0 | 90 | 0.05 | 0 | | |
| 554 | 15–30 | 0.5 | 0 | 180 | 0.38 | 0 | | |
| 554 | 30–45 | 0.5 | 0 | 270 | 0.36 | 0 | | |
| 554 | 45–60 | NA | NA | 360 | 0.18 | 0 | 0.18 | 0 |
| 587 | 0–15 | 0.5 | 0 | | | | | |
| 587 | 15–30 | 0.5 | 0 | | | | | |
| 587 | 30–45 | 0.5 | 0 | | | | | |
| 587 | 45–60 | NA | NA | | | | | |
| 625 | 0–15 | 0.5 | 0 | | | | | |
| 625 | 15–30 | 0.5 | 0 | | | | | |
| 625 | 30–45 | 0.5 | 0 | | | | | |
| 625 | 45–60 | NA | NA | | | | | |
| 652 | 0–15 | 0.5 | 0 | | | | | |
| 652 | 15–30 | 0.5 | 0 | | | | | |
| 652 | 30–45 | 0.5 | 0 | | | | | |
| 652 | 45–60 | 0.5 | 0 | | | | | |

[a] See footnote in Table 1.
[b] NA = data not available.

*Runoff: Georgia, USA, cotton (GA1R)*

Study site GA1R was a field-scale runoff study with natural rainfall conducted on cotton in a small watershed in southern Georgia (Colquitt County). Pesticide runoff mass was the only value measured at site GA1R. Table 2 shows the percentage exceedence values for the 4 d on which measurements were available. For all days, the measured value fell within the interquartile range of the model predictions.

*Runoff: Days with zero runoff measurements*

For the runoff field studies, runoff volume, sediment yield, pesticide runoff mass, and pesticide mass in sediment were monitored continuously for each day of the sampling periods at both the IA2R and the GA1R site. Those days with positive measurements were evaluated in the preceding tables and figures. For all other days, the measured values were assumed to be zero. At issue is whether PRZM 3.12 was able to predict zero or low values on those days where no runoff, sediment loss, runoff flux, or pesticide mass in sediment were recorded in the field.

Cumulative distributions of the model predictions (over the 500 Monte Carlo iterations) on those days where no runoff volume, sediment yield, dissolved pesticide runoff mass, or pesticide mass adsorbed to sediment were recorded at site IA2R in 1992. Figure 3 illustrates the results of the Monte Carlo uncertainty analysis for those days where no pesticide mass in sediment were recorded in the field. Similar data were generated for runoff volume, sediment yield, and dissolved pesticide mass for both 1992 and 1993 but have not been presented. Table 3 presents summary statistics of the information. Specifically, Table 3 presents the percentage of all model predictions that were zero or greater. Table 3 illustrates the frequency with which the model confirmed the measured values of zero for the four runoff variables. The information demonstrates that for some days the model clearly predicts positive values when the field monitors showed no measurable results. For most of these days, however, the model predicted values of small magnitude, near the limit of detection. For a small number of days, the model had large positive predictions. Intuitively, these data are not unanticipated given the nature of the Monte Carlo sampling procedure. For each input parameter defined via sensitivity analysis as exerting a significant

influence on model outcome, Monte Carlo sampling from a defined distribution would invariably produce a combination of values with a small likelihood of true occurrence. The combination of these input parameters would likely generate a relatively small number of predictions that would occur with low probability. Those low probability input combinations may produce those traces with high positive values (although the cause of these high positive values was not rigorously evaluated). However, overall at least 69% of all model predictions for a specific runoff variable were zero. The highest concordance was seen in the pesticide mass in sediment variable in 1983, where 83% of the model predictions equaled zero on days where no pesticide mass in sediment was measured.

Table 4 illustrates the distribution of Monte Carlo predictions for runoff mass for days with no measured values for the GA1R study. The data indicate that 78% of the Monte Carlo predictions were zero.

*Leaching: Georgia sweet corn (GA1L)*

Field study GA1L was a conventional small-scale prospective groundwater study conducted in a highly vulnerable agronomic setting on sweet corn in south-central Georgia. Box-and-whisker plots were generated for the leaching variables. Incremental depth intervals equaled 0 to 15, 15 to 30, 30 to 45, and 45 to 60 cm for the leaching variable pesticide mass in soil and depths of 90, 180, 270, and 360 cm for the leaching variables pesticide in pore water and bromide in pore water. Representative box-and-whisker plots are presented for the leaching variables pesticide mass in soil (0–15 cm, year 1) (Fig. 4), pesticide in pore water (90 cm) (Fig. 5), and bromide in pore water (180 cm) (Fig. 6). Percentage exceedence calculations for GA1L are presented in Table 5.

Of the 43 pesticide mass in soil values in year 1, 27 fell within the interquartile range of the model predictions; that is, the percentage exceedence values ranged between 25 and 75% (Table 5). Eleven of the remaining days at which the measured data fell outside the interquartile range had measured values at or below the LOQ and were therefore set to 0.5 μg/kg for purposes of this analysis. Only 5 d with measurements greater than the LOQ fell outside the interquartile range. All measurements were within the bounds of the prediction interval. In year 2, four of 36 measured values fell within the interquartile range. But 22 of the days exhibited values less than or equal to the LOQ. The remaining measured values were less than 5 μg/kg.

Only three measured values for pesticide in pore water were greater than the LOQ (set to one-half the LOQ, or 0.05 μg/L for purposes of the analysis) in year 1, and the largest of these three values was 0.48 μg/L. In year 2, 12 of the 28 measured values were less than or equal to the LOQ. The model underpredicted the measured values in 26 of the 28 possible cases; that is, the percentage exceedence was less than 25% (Table 5). Of the 16 measured values greater than one-half the LOQ, the largest value was 1.16 μg/L, and the remainder of the measured values were below 0.38 μg/L. Importantly, these measured concentrations are very small and are likely to be environmentally irrelevant.

For bromide in pore water in both years 1 and 2, the measured values did not fall within the interquartile range of model outcome distributions. Eight measured values were equivalent to one-half the LOQ (0.05 mg/L), and the remaining 49 measured values ranged from 0.10 to 113.30 mg/L. Model distribution outcomes underpredicted the measured values in 34 of the 47 cases and overpredicted the remaining 13 cases (Table 5). Measurements and model distribution outcomes, while typically not overlapping, were increasingly more correlative with depth and time. Importantly, the estimated spatial and temporal profile or pattern of pore-water bromide movement through the soil core was highly correlated to the measured data. Several conclusions can be drawn from the data: In this instance, the model can be considered inaccurate with regard to estimating the magnitude of the bromide pore-water concentration on a daily basis, and the model can be considered accurate in estimating the spatial and temporal movement of the tracer. The discrepancy in the magnitude of the estimated and measured pore-water bromide concentrations is likely due to the inability to precisely simulate bromide uptake by plant material, the variability associated with sampling of soil pore water via suction lysimeters, and its associated uncertainty and discrepancies related to estimating evapotranspiration.

*Leaching: North Carolina, USA, soybeans (NC4L)*

Field study NC4L was a conventional small-scale prospective groundwater study conducted in a highly vulnerable agronomic setting on soybeans in North Carolina. Percentage exceedence calculations are presented in Table 6.

In year 1, the predictive pattern of the model for large measured values of pesticide mass in soil (≥150 μg/kg) was inconsistent with model outcome distributions. Measured values fell outside the model prediction interquartile range with estimates either greater than or less than measured data. For smaller measured values (≤50 μg/kg), the model outcome distributions tended to underpredict the measured data. Of the 53 nonzero measured values equaling ≤50 μg/kg, two fell within the interquartile range, and 43 values were underpredicted where the percentage exceedence was less than 25%. As the depth increased and the measured concentrations decreased, the model did predict small concentrations. In year 2, a similar pattern held with the model underpredicting small values. Of the 26 measured values ranging from 21.5 to 0.5 μg/kg, all were underpredicted. However, as the depth of the soil profile increased and the measured concentrations became increasingly small, the model prediction error decreased. Beginning at the 45- to 60-cm soil core segment, any discrepancies between model estimates and measured soil core pesticide concentrations were not substantial.

Of the year 1 pesticide in pore water values greater than 10 μg/L, one fell within the interquartile range of the model predictions, and two more fell within the model prediction bounds. Of those concentrations greater than zero and less than 10 μg/L, seven of the eight measured values were overpredicted by the model. No measured values were greater than 10 μg/L in year 2. The model underpredicted 16 of the 17 available data points that were, however, small in magnitude. During the course of the two-year study, several pesticide soil pore-water measurements resulted in nondetectable residues. The model overpredicted three of seven nondetects. Three of the remaining estimates were underpredictive, and for one measurement the model percentage exceedence was within the interquartile range.

Bromide soil pore-water prediction distributions and measured data generally show a similar pattern with depth and time. Of the 18 available measured data points, one fell within the interquartile range of the model distribution outcomes. The model underpredicted eight values. However, the magnitude

Table 6. Daily comparison of model predictions and measured values: North Carolina, USA, soybeans (NC4L)[a]

| Julian day | Soil horizon depth (cm) | Measured pesticide (μg/kg) | Percentage exceedence | Lysimeter depth | Measured pesticide (μg/L) | Percentage exceedence | Measured bromide (mg/L) | Percentage exceedence |
|---|---|---|---|---|---|---|---|---|
| 136 | 0–15 | 330.3 | 0 | | | | | |
| 136 | 15–30 | 3.53 | 0 | | | | | |
| 136 | 30–45 | 0.67 | 0 | | | | | |
| 136 | 45–60 | 0 | 0 | | | | | |
| 136 | 60–75 | 0 | 0 | | | | | |
| 136 | 75–90 | NA[b] | NA | | | | | |
| 136 | 90–105 | NA | NA | | | | | |
| 136 | 105–120 | NA | NA | | | | | |
| 137 | 0–15 | 157.4 | 100 | | | | | |
| 137 | 15–30 | 2.23 | 0 | | | | | |
| 137 | 30–45 | 0 | 0 | | | | | |
| 137 | 45–60 | 0 | 0 | | | | | |
| 137 | 60–75 | 0 | 0 | | | | | |
| 137 | 75–90 | NA | NA | | | | | |
| 137 | 90–105 | NA | NA | | | | | |
| 137 | 105–120 | NA | NA | | | | | |
| 143 | 0–15 | 176.6 | 98.2 | | | | | |
| 143 | 15–30 | 2.53 | 0 | | | | | |
| 143 | 30–45 | 0 | 0 | | | | | |
| 143 | 45–60 | 0 | 0 | | | | | |
| 143 | 60–75 | 0 | 0 | | | | | |
| 143 | 75–90 | NA | NA | | | | | |
| 143 | 90–105 | NA | NA | | | | | |
| 143 | 105–120 | NA | NA | | | | | |
| 149 | | | | 90 | 0 | 0 | 0.09 | 0 |
| 149 | | | | 150 | 0 | 0 | 0.11 | 0 |
| 149 | | | | 210 | 0 | 0 | 0.12 | 0 |
| 150 | 0–15 | 203.5 | 92 | | | | | |
| 150 | 15–30 | 1.83 | 0 | | | | | |
| 150 | 30–45 | 0 | 0 | | | | | |
| 150 | 45–60 | 0 | 0 | | | | | |
| 150 | 60–75 | 0 | 0 | | | | | |
| 150 | 75–90 | NA | NA | | | | | |
| 150 | 90–105 | NA | NA | | | | | |
| 150 | 105–120 | NA | NA | | | | | |
| 168 | 0–15 | 145.8 | 0 | 90 | 0 | 100 | 1.18 | 100 |
| 168 | 15–30 | 16.17 | 99.4 | 150 | 0 | 99.8 | 0.07 | 100 |
| 168 | 30–45 | 1.7 | 99.6 | 210 | 0 | 71.4 | 0.09 | 100 |
| 168 | 45–60 | 0 | 100 | | | | | |
| 168 | 60–75 | 0 | 100 | | | | | |
| 168 | 75–90 | 4.4 | 13 | | | | | |
| 168 | 90–105 | 0 | 99.8 | | | | | |
| 168 | 105–120 | 0 | 99.8 | | | | | |
| 197 | 0–15 | 56.67 | 0 | 90 | 14.44 | 99.2 | 30.28 | 100 |
| 197 | 15–30 | 36.67 | 0 | 150 | 0.5 | 99.8 | 19.82 | 100 |
| 197 | 30–45 | 10.93 | 16.4 | 210 | 0.25 | 99.8 | 11.86 | 100 |
| 197 | 45–60 | 6.63 | 62.2 | | | | | |
| 197 | 60–75 | 5.23 | 92.6 | | | | | |
| 197 | 75–90 | 10.83 | 76 | | | | | |
| 197 | 90–105 | 2.5 | 98.2 | | | | | |
| 197 | 105–120 | 0.7 | 99.6 | | | | | |
| 239 | 0–15 | 21 | 0 | 90 | 15.98 | 0 | 2.09 | 0 |
| 239 | 15–30 | 16.77 | 0 | 150 | 11.69 | 57.8 | 3.39 | 0 |
| 239 | 30–45 | 5.6 | 0 | 210 | 3.65 | 98.8 | 1.72 | 100 |
| 239 | 45–60 | 3.2 | 0 | | | | | |
| 239 | 60–75 | 2.87 | 0 | | | | | |
| 239 | 75–90 | 2.73 | 0 | | | | | |
| 239 | 90–105 | 2.13 | 0 | | | | | |
| 239 | 105–120 | 1.37 | 0 | | | | | |
| 260 | 0–15 | 21.97 | 0 | 90 | 19.7 | 0 | NA | NA |
| 260 | 15–30 | 15.6 | 0 | 150 | 12.57 | 0 | 1.38 | 0 |
| 260 | 30–45 | 5.23 | 0 | 210 | 8 | 97.8 | 1.28 | 100 |
| 260 | 45–60 | 3.2 | 0 | | | | | |
| 260 | 60–75 | 2.9 | 0 | | | | | |
| 260 | 75–90 | 2.87 | 0 | | | | | |
| 260 | 90–105 | 1.23 | 0 | | | | | |
| 260 | 105–120 | 0 | 98.4 | | | | | |
| 295 | | | | 90 | 11.21 | 0 | 0.54 | 0 |
| 295 | | | | 150 | 7.26 | 99.6 | 1.34 | 0 |
| 295 | | | | 210 | 5.4 | 96.8 | 1.16 | 62.6 |

Table 6. Continued

| Julian day | Soil horizon depth (cm) | Measured pesticide (μg/kg) | Percentage exceedence | Lysimeter depth | Measured pesticide (μg/L) | Percentage exceedence | Measured bromide (mg/L) | Percentage exceedence |
|---|---|---|---|---|---|---|---|---|
| 296 | 0–15 | 22.03 | 0 | | | | | |
| 296 | 15–30 | 15.8 | 0 | | | | | |
| 296 | 30–45 | 4 | 0 | | | | | |
| 296 | 45–60 | 3.03 | 0 | | | | | |
| 296 | 60–75 | 2.03 | 0 | | | | | |
| 296 | 75–90 | 2.37 | 30.2 | | | | | |
| 296 | 90–105 | 1.4 | 0 | | | | | |
| 296 | 105–120 | 0.8 | 0 | | | | | |
| 322 | 0–15 | 17.87 | 0 | 90 | 12.34 | 0 | | |
| 322 | 15–30 | 10.83 | 0 | 150 | 15.05 | 99.2 | | |
| 322 | 30–45 | 2.83 | 0 | 210 | 9.22 | 48.4 | | |
| 322 | 45–60 | 1.43 | 0 | | | | | |
| 322 | 60–75 | 0.73 | 0 | | | | | |
| 322 | 75–90 | 1.27 | 99.6 | | | | | |
| 322 | 90–105 | 0.43 | 0 | | | | | |
| 322 | 105–120 | 0.33 | 0 | | | | | |
| 350 | 0–15 | 15.1 | 0 | 90 | 11.58 | 0 | | |
| 350 | 15–30 | 11.57 | 0 | 150 | 6.44 | 75.2 | | |
| 350 | 30–45 | 3.77 | 0 | 210 | 12.25 | 0 | | |
| 350 | 45–60 | 2 | 0 | | | | | |
| 350 | 60–75 | 0.9 | 0 | | | | | |
| 350 | 75–90 | 1.47 | 94 | | | | | |
| 350 | 90–105 | 0.37 | 0 | | | | | |
| 350 | 105–120 | 0 | 100 | | | | | |
| 384 | 0–15 | 17.9 | 0 | 90 | 2.77 | 0 | | |
| 384 | 15–30 | 6.53 | 0 | 150 | 9.16 | 0 | | |
| 384 | 30–45 | 2.3 | 0 | 210 | 2.81 | 0 | | |
| 384 | 45–60 | 1.47 | 0 | | | | | |
| 384 | 60–75 | 0.57 | 0 | | | | | |
| 384 | 75–90 | 0.37 | 0 | | | | | |
| 384 | 90–105 | 0 | 99.8 | | | | | |
| 384 | 105–120 | 0 | 99.8 | | | | | |
| 413 | 0–15 | 16.87 | 0 | 90 | 2.74 | 0 | | |
| 413 | 15–30 | 13.43 | 0 | 150 | 6.59 | 0 | | |
| 413 | 30–45 | 4.87 | 0 | 210 | 1.32 | 0 | | |
| 413 | 45–60 | 1.6 | 0 | | | | | |
| 413 | 60–75 | 2 | 0 | | | | | |
| 413 | 75–90 | 0.37 | 0 | | | | | |
| 413 | 90–105 | 0.57 | 0 | | | | | |
| 413 | 105–120 | 0.4 | 0 | | | | | |
| 440 | 0–15 | 17.57 | 0 | 90 | 2.24 | 0 | | |
| 440 | 15–30 | 9.6 | 0 | 150 | 4.19 | 0 | | |
| 440 | 30–45 | 2.2 | 0 | 210 | 2.81 | 0 | | |
| 440 | 45–60 | 1.47 | 0 | | | | | |
| 440 | 60–75 | 0 | 66.4 | | | | | |
| 440 | 75–90 | 0 | 82.8 | | | | | |
| 440 | 90–105 | 0 | 98.6 | | | | | |
| 440 | 105–120 | 0 | 99.8 | | | | | |
| 474 | | | | 90 | NA | NA | | |
| 474 | | | | 150 | 0.56 | 0 | | |
| 474 | | | | 210 | 0 | 100 | | |
| 475 | 0–15 | 13.43 | 0 | | | | | |
| 475 | 15–30 | 8.8 | 0 | | | | | |
| 475 | 30–45 | 0.97 | 0 | | | | | |
| 475 | 45–60 | 0.5 | 0 | | | | | |
| 475 | 60–75 | 0.37 | 0 | | | | | |
| 475 | 75–90 | 0 | 43 | | | | | |
| 475 | 90–105 | 0 | 61.4 | | | | | |
| 475 | 105–120 | 0 | 77.6 | | | | | |
| 502 | | | | 90 | 3.52 | 0 | | |
| 502 | | | | 150 | 4.71 | 0 | | |
| 502 | | | | 210 | 1.82 | 0 | | |
| 503 | 0–15 | 21.5 | 0 | | | | | |
| 503 | 15–30 | 14.1 | 0 | | | | | |
| 503 | 30–45 | 7.25 | 0 | | | | | |
| 503 | 45–60 | 3.35 | 0 | | | | | |
| 503 | 60–75 | 4.63 | 0 | | | | | |
| 503 | 75–90 | 2.43 | 0 | | | | | |
| 503 | 90–105 | 0.9 | 0 | | | | | |
| 503 | 105–120 | 0.37 | 0 | | | | | |

Table 6. Continued

| Julian day | Soil horizon depth (cm) | Measured pesticide (μg/kg) | Percentage exceedence | Lysimeter depth | Measured pesticide (μg/L) | Percentage exceedence | Measured bromide (mg/L) | Percentage exceedence |
|---|---|---|---|---|---|---|---|---|
| 530 | | | | 90 | 3.2 | 0 | | |
| 530 | | | | 150 | 5.85 | 0 | | |
| 530 | | | | 210 | 4.83 | 0 | | |

[a] See footnote in Table 1.
[b] NA = data not available.

of the differences between the measured data and estimate distributions was generally minimal.

## DISCUSSION

These results serve to demonstrate the feasibility and utility of evaluating the effects of model input uncertainty on PRZM 3.12 outcomes. In general, when model input uncertainty was accounted for, the correlation of model outcome distributions and measured data was reasonably to exceptionally well correlated. This conclusion can be drawn despite the fact that the uncertainty bounding the measured values was not factored into the analysis. Pennell et al. [7] conclude that the ability to validate model predictions of concentration distributions may ultimately be limited by the inability to account for the uncertainty in measured data from within the field. Given the expected uncertainty in the measured data, the degree of prediction error and measurement error would make it increasingly difficult to detect differences.

The current state of the science with regard to exposure analysis is such that evaluation of model predictive accuracy is often assessed via the factor-of-*f* approach [9,10,20]. Comparisons of model estimates versus measured values are often considered successful within two-, five- and 10-fold differences. The Monte Carlo–driven output distribution approach extends the factor-of-*f* approach discussed within the literature because it adds an empirical aspect to the analysis. Rather than set an arbitrary level for accuracy, as an example, a factor of five, this approach allows the nature of the measured data serving as input to set the bounds that define the precision of the model. Measured values falling within the interquartile range of an outcome distribution lead to the conclusion that the model is reasonably predictive. Given the state of the science of exposure analysis, even when measured values fall within the outcome distribution bounds, the model should be considered predictive. It is important to note, however, that the scale of the measurement influences the degree of required accuracy. Based on the current analysis, it has been shown that for small concentrations (e.g., less than 5 μg/L of pesticide or pesticide concentrations approximating the LOQ), the criteria for accuracy need not be as rigorous. Differences in model outcome distributions and measured data in instances where the magnitude of the scale of the measured data is small or approaches the LOQ become less critical. Typically, the magnitude of those differences is beyond the desired level of model accuracy and environmental relevance.

An important aspect of the current approach that should be emphasized is that the nature of the input distributions defines the output distributions. Subjective and incorrect assumptions about the nature of the input distributions, while allowing for the generation of seemingly accurate output distributions, can provide spurious results. In the process outlined in this discussion, the nature of the input parameter distributions was carefully explored as deeply as the data would allow. In those instances where considerable uncertainty existed about the input parameter distribution, the conservative assumption was taken. Typically, a uniform distribution was assigned to those uncertain input parameter distributions where any one value within the bounds of the distribution had an equal probability of selection. One flaw in the current analysis is the depth of information about each of the available input parameters. Future work should focus on enriching the database from which these assumptions about distributions can be made.

Loague and Green [8] and others note that statistical analyses using pairwise correlation or hypothesis testing can suffer from potential serious flaws because of sample size deficiencies. Preliminary efforts for this study centered on pairwise correlation and hypothesis-testing statistical approaches for estimating model accuracy. Ultimately, the efforts refocused on the Monte Carlo approach because the classical statistical approach was hampered by small sample sizes and differences in phase timing that led to conclusions of reduced model accuracy.

Importantly, the Monte Carlo approach lends itself to the current trend in environmental risk analyses where stochastic predictions are favored over single-point deterministic results. Clearly, under environmental conditions, the magnitude of associated uncertainties makes the utility of a single deterministic model prediction debatable.

*Acknowledgement*—We are especially grateful to Blanche Dean for graphics assistance.

## REFERENCES

1. Carsel RF, Nixon WB, Ballantine LG. 1986. Comparison of pesticide root zone model predictions with observed concentrations for the tobacco pesticide metalaxyl in unsaturated zone soils. *Environ Toxicol Chem* 5:345–353.
2. Jones RL, Black GW, Estes TL. 1986. Comparison of computer model predictions with unsaturated zone field data for aldicarb and aldoxycarb. *Environ Toxicol Chem* 5:1027–1037.
3. Leonard RA, Knisel WG, Davis FM, Johnson AW. 1990. Validating GLEAMS with field data for fenamiphos and its metabolites. *Journal of Irrigation and Drainage Engineering* 116:24–35.
4. Salo S, Posch M, Rekolainen S. 1994. Testing the modified CREAMS/GLEAMS model for pesticide concentration in soil. *Agric Sci Finl* 3:59–67.
5. Dean JD, Huyakorn PS, Donigian AS Jr, Voos KA, Schanz RW, Meeks YJ, Carsel RF. 1989. Risk of unsaturated/saturated transport and transformation of chemical concentrations (RUSTIC). I. EPA/600/3-89/048a and II. EPA/600/3-89/048b. U.S. Environmental Protection Agency, Athens, GA.
6. Carsel RF, Parrish RS, Jones RL, Hansen JL, Lamb RL. 1988. Characterizing the uncertainty of pesticide leaching in agricultural soils. *J Contam Hydrol* 2:111–124.
7. Pennell KD, Hornsby AG, Jesup RE, Rao PSC. 1990. Evaluation of five simulation models for predicting aldicarb and bromide behavior under field conditions. *Water Resour Res* 6:2679–2693.
8. Loague K, Green RE. 1991. Statistical and graphical methods for

evaluating solute transport models: Overview and application. *J Contam Hydrol* 7:51–73.
9. Parrish RS, Smith CN, Fong FK. 1992. Tests of the pesticide root zone model for transport and transformation of aldicarb, metolachlor, and bromide. *J Environ Qual* 21:685–697.
10. Parrish RS, Smith CN. 1990. A method for testing whether model predictions fall within a prescribed factor of true values with an application to pesticide leaching. *Ecol Model* 51:59–72.
11. Burns LA. 1991. PIRANHA: Pesticide and industrial chemical risk analysis and hazard assessment, Ver 2.0. U.S. Environmental Protection Agency, Athens, GA.
12. Zacharias S, Heatwole CD. 1994. Evaluation of GLEAMS and PRZM for predicting pesticide leaching under field conditions. *Trans ASAE* 37:439–451.
13. Zacharias S, Heatwole CD, Coakley CW. 1994. Comparison of quantitative techniques for pesticide model validation. ASAE Paper 93-2506. American Society of Agricultural Engineers, St. Joseph, MI.
14. Melacini A, Walker A, Gunther P. 1995. Preliminary methods for assessing and ranking model performance in the prediction of pesticide soil residues. In Walker A, Calvet R, Del Re AAM, Pestemer W, Hollis JM, eds, *Evaluation and Improvement of Mathematical Models of Pesticide Mobility in Soils and Assessment of Their Potential to Predict Contamination of Water Systems*. Biologische Bundesanstalt für Land und Forstwirtschaft, Berlin, Germany, pp 23–31.
15. Hutson JL, Wagenet RJ. 1992. LEACHM: Leaching estimation and chemistry model. Department of Soil, Crop and Atmospheric Sciences, Research Series 92-3. New York State College of Agriculture and Life Sciences, Cornell University, Ithaca, NY, USA.
16. Vanclooster M, Boesten JJTI, Trevisan M, Brown CD, Capri E, Eklo OM, Gottesburen B, Gouy V, van der Linden AMA. 2000. A European test of pesticide-leaching models: Methodology and major recommendations. *Agric Water Manag* 44:1–19.
17. Armstrong AC, Portwood AM, Leeds-Harrison PB, Harris LG, Catt JA. 1996. The validation of pesticide leaching models. *Pestic Sci* 48:47–55.
18. Klein M. 1999. Monte Carlo analysis using pesticide fate models. *J Pestic Sci* 24:55–59.
19. Boekhold AE, van den Bosch H, Boesten JJTI, Leistra M, Swartjes FA, van der Linden AMA. 1993. Validation of the PESTLA model: Definitions, objectives and procedure. Report 715802001. National Institute of Public Health and Environmental Protection, Bilthoven, The Netherlands.
20. Van den Bosch H, Boesten JJTI. 1994. Validation of the PESTLA model: Field test for leaching of two pesticides in a humic sandy soil in Vredepeel (The Netherlands). Report 82. The Winand Staring Centre for Integrated Land, Soil and Water Research, Wageningen, The Netherlands.
21. Eckhardt DAV, Wagenet RJ. 1996. Estimation of the potential for atrazine transport in a silt loam soil. In Meyer MT, Thurman EM, eds, *Herbicide Metabolites in Surface Water and Groundwater*. ACS Symposium Series 630. American Chemical Society, Washington, DC.
22. Haan CT, Allred B, Storm DE, Sabbagh GJ, Prabhu S. 1995. Statistical procedure for evaluating hydrologic/water quality models. *Trans ASAE* 38:725–733.
23. Plackett RL, Burman JP. 1946. The design of optimum multifactorial experiments. *Biometrika* 33:305–325.
24. Wolt J, Singh P, Cryer S, Lin J. 2001. Sensitivity analysis for validating expert opinion as to ideal data set criteria for transport modeling. *Environ Toxicol Chem* 21:1558–1565.
25. Warren-Hicks W, Havens PL, Carbone JP. 2001. Using Monte Carlo techniques to judge model prediction accuracy: Validation of pesticide root zone model 3.12. *Environ Toxicol Chem* 21: 1570–1577.
26. Havens PL, Carbone JP, Warren-Hicks W, Fouch MA. 2001. Development of a Monte Carlo sampling shell for the pesticide root zone model and its application by the Federal Insecticide, Fungicide, and Rodenticide Act Environmental Modeling Validation Task Force. *Environ Toxicol Chem* 21:1566–1569.