SETAC PRESS

Environmental Toxicology and Chemistry, Vol. 21, No. 8, pp. 1566–1569, 2002

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

# 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

PATRICK L. HAVENS,*† JOHN P. CARBONE,‡ WILLIAM WARREN-HICKS,§ and MICHAEL A. FOUCH‖
†Regulatory Laboratories, Dow AgroSciences, 9330 Zionsville Road, Indianapolis, Indiana 46268, USA
‡Toxicology Department, Rohm and Haas, Spring House, Pennsylvania 19477-0904, USA
§The Cadmus Group, 1920 Highway 54, Durham, North Carolina 27713, USA
‖Sphere Software, 9655 North 750 East, Brownsburg, Indiana 46112, USA

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

**Abstract**—A user interface to the U.S. Environmental Protection Agency pesticide root zone model (PRZM) was constructed to allow Monte Carlo sampling of input parameter distributions. The interface was constructed employing the Visual Basic for Applications development environment, along with the functionality of the Crystal Ball Professional forecasting and risk analysis package. The tool has been utilized by the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) Environmental Model Validation Task Force to perform detailed statistical analyses of model input parameter uncertainty and the propagation of this uncertainty on the model outputs as well as comparisons of modeled and field-measured data.

**Keywords**—Federal Insecticide, Fungicide, and Rodenticide Act  Model validation  Monte Carlo

## INTRODUCTION

The FIFRA Environmental Model Validation Task Force (FEMVTF) was established in late 1995 to improve regulatory confidence in the use of simulation models [1]. The task force, a joint effort of the U.S. Environmental Protection Agency and scientists from crop protection companies, has compared measured data from a large number of field leaching and runoff studies with predictions from the PRZM model [2,3]. The task force has also invested in the development of tools for estimating the sensitivity of the model outputs to input variation [4] and for characterizing model input and output uncertainty [5,6].

Many real-world simulation problems involving uncertainty, such as environmental fate as described by PRZM, are too complex to be solved by conventional analytical methods. Even with well-defined uncertainties, simulating all the combinations of inputs to a model can be computationally prohibitive. A powerful statistical approach to this problem is the application of Monte Carlo (MC) methods [7]. The MC analysis allows simple characterization of the effects of input uncertainty on deterministic model outputs. In the context of PRZM model validation, these methods allowed a direct comparison of the range of model outputs with field-measured data [5].

As has been described previously [1], the PRZM 3.12 model [8,9] was selected as the primary model for the FEMVTF effort. The PRZM model does contain an MC module, although its functionality is limited in the input parameters that can be varied as well as in the choice of input distributions [9]. Although additional parameters may be added by modifying the PRZM code, code modification was determined to be an unacceptable option since code integrity is a critical concern for regulatory models. Therefore, the development of a more complete MC shell to meet the needs of a robust evaluation of model uncertainty was accomplished with standard tools, building on the previous development of a tool for PRZM sensitivity analysis [4].

The environment chosen for the MC shell development was the Crystal Ball Professional® (CB) package produced by Decisioneering (Denver, CO, USA). The CB is a set of Visual Basic programs that function as an add-on to the Microsoft Excel (Redlands, WA, USA) spreadsheet program. The CB contains functionality for defining input distributions, sampling from them to run a deterministic model, and analyzing and storing the results. The CB allows selection of input distributions from a library of 17 distribution function types, including user-defined and noncontinuous distributions. The CB was integrated with the PRZM model via the Microsoft Visual Basic for Applications (VBA) functionality that is integrated into Excel. The Developer's Kit for CB allows control of most CB functions via VBA objects; external system calls allow the execution of batch and FORTRAN programs external to Excel. The shell allows sampling from a large number of PRZM input parameters, as shown in Table 1.

As a set of simulations is run within the CB/VBA environment, CB maintains the output distributions and accumulates statistics about these outputs (such as mean, standard deviation, and coefficient of variation) that can be compared against a user-defined criterion for convergence. For computational efficiency, output is calculated not after each simulation but for sets of simulations, implemented as the CB *multisimulation* functionality. The input distributions used in this analysis were based on field-measured data and are described in detail by Warren-Hicks et al. [6]. By sampling from the real-world measurements, measurements of the variability in outputs could be produced, and measures of the statistical validity of the model predictions were assembled.

* To whom correspondence should be addressed (phavens@dow.com).

Table 1. Pesticide root zone model input parameters that can be varied in the Monte Carlo shell

| Parameter[a] | Descriptions | Subscript 1 | Subscript 2 |
|---|---|---|---|
| AD(NHORIZ) | Soil drainage parameter | Soil horizon number (1–10) | — |
| ANETD | Minimum evaporation extraction depth | — | — |
| AMXDR(NCD) | Maximum rooting depth | Crop number (1–5) | — |
| BD(NHORIZ) | Bulk density | Soil horizon number (1–10) | — |
| CINTCP(NDC) | Maximum crop interception storage | Crop number (1–5) | — |
| CLAY(NHORIZ) | Clay % | Soil horizon number (1–10) | — |
| CN(NDC, CN_NUMBER) | Runoff curve number, antecedent moisture condition II | Crop number (1–5) | Cropping stage (1–3) |
| COVMAX(NDC) | Maximum areal coverage of the canopy | Crop number (1–5) | — |
| DAIR(NCHEM) | Diffusion coefficient in air | Chemical number (1–3) | — |
| DEPI(NCHEM, NAPS) | Depth of incorporation | Chemical number (1–3) | Application number (1–50) |
| DGRATE(NCHEM, NHORIZ) | Vapor phase chemical decay rate | Chemical number (1–3) | Soil horizon number (1–10) |
| DISP(NCHEM, NHORIZ) | Chemical hydrodynamic solute dispersion coefficient | Chemical number (1–3) | Soil horizon number (1–10) |
| DPN(NHORIZ) | Compartment thickness | Soil horizon number (1–10) | — |
| DSRATE(NCHEM, NHORIZ) | Adsorbed phase chemical decay rate | Chemical number (1–3) | Soil horizon number (1–10) |
| DWRATE(NCHEM, NHORIZ) | Dissolve phase chemical decay rate | Chemical number (1–3) | Soil horizon number (1–10) |
| EMPY(NCHEM) | Enthalpy of vaporization | Chemical number (1–3) | — |
| FEXTRC(NCHEM) | Foliar extraction coefficient for chemical wash-off | Chemical number (1–3) | — |
| FILTRA | Filtration parameter | — | — |
| HENRYK(NCHEM) | Henry's law constant | Chemical number (1–3) | — |
| HL | Hydraulic length | — | — |
| HTMAX(NDC) | Maximum canopy height at maturation | Crop number (1–5) | — |
| KD(NCHEM, NHORIZ) | Chemical partition coefficient | Chemical number (1–3) | Soil horizon number (1–10) |
| MNGN(NDC, NUSLEC) | Manning's N | Crop number (1–5) | Universal Soil Loss Equation ''C'' factor (USLEC) number |
| OC(NHORIZ) | Organic carbon | Soil horizon number (1–10) | — |
| PFAC | Pan factor | — | — |
| PLDKRT(NCHEM) | Pesticide decay rate on plant foliage | Chemical number (1–3) | — |
| PLVKRT(NCHEM) | Chemical volatilization decay rate on plant foliage | Chemical number (1–3) | — |
| SAND(NHORIZ) | Sand % | Soil horizon number (1–10) | — |
| SLP | Land slope | — | — |
| SOL(NCHEM) | Solubility in water | Chemical number (1–3) | — |
| TAPP(NCHEM, NAPS) | Chemical application rate | Chemical number (1–3) | Application number (1–50) |
| THETO(NHORIZ) | Initial soil water content | Soil horizon number (1–10) | — |
| UPTKF(NCHEM) | Plant uptake factor | Chemical number (1–3) | — |
| USLEC(NDC, NUSLEC) | Universal soil loss cover management | Crop number (1–5) | USLEC factor number |

[a] Parameter names correspond to PRZM 3.12 inputs [9].

## SYSTEM DESIGN

The general design of the system is shown in Figure 1, with overall program control being set up in an Excel dialog (or form). The parameters entered into the dialog are the minimum and maximum number of simulation sets and the number of simulations per set as well as the convergence statistic and the criterion for convergence. This dialog's start button is linked to the VBA script *start*, which first prompts for the output cell for statistical calculations and begins execution of a CB multisimulation set within the multisimulation set, *CB*. The *Simulation* command first samples from the input distributions (*assumptions* in CB's terminology) previously defined by the user. By setting the *run preferences* in CB, the VBA program *write_file* carries out a series of commands to load the sampled inputs into a PRZM input file, execute PRZM, and load the outputs back into CB *forecasts*, which are vectors of the outputs. Simulations then continue, up to the number of simulations per set defined in the setup dialog. At this point, statistics are calculated for the output cell defined in the dialog, and the value is compared to the convergence criterion. If this value is less than the criterion, program control returns to the user. If the criterion is not met, the next multisimulation set is initiated, up to the maximum number of simulation sets defined in the setup dialog.

### *Setup of input parameter distributions and outputs*

The input parameter distributions are defined on the *inputs* worksheet as CB assumptions. For each assumption, a corresponding forecast is defined; this allows tracking of outputs versus the inputs used to calculate them by employing CB's

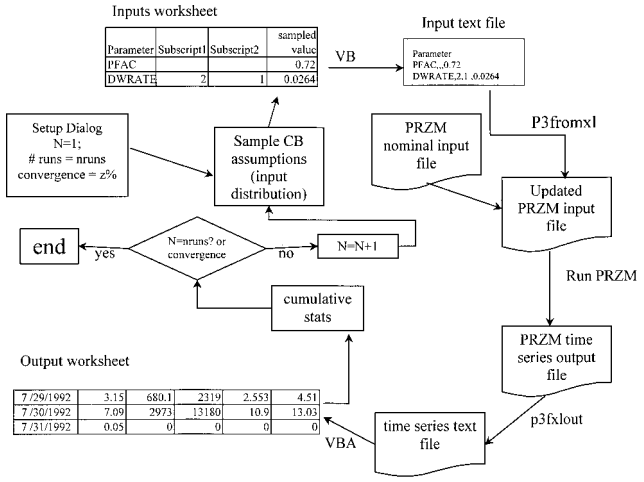


Fig. 1. General system design. CB = Crystal Ball Professional®; PRZM = pesticide root zone model; VBA = Visual Basic for Applications. Refer to Table 1 for explanation of abbreviations.

*keep sorted trials* option. The inputs worksheet contains the columns *parameter, subscript 1, subscript 2, formatted sampled value, sampled value,* and *input distribution.* The values in the *sampled value* column are defined as CB forecasts, while the *formatted sampled value* entries are needed for entry into the fixed-format PRZM input file. The *input distribution* cells are CB assumptions, defined as described by Warren-Hicks et al. [6].

### *Populating PRZM input files, executing PRZM, and processing outputs*

Once a set of input values has been defined by MC sampling, the contents of the *inputs* worksheet is saved by a VBA statement into an external, comma-delimited text file. An external batch file then executes a FORTRAN program that reads the text file; parses the parameter names, subscripts, and values; and writes the inputs into the appropriate fields in a PRZM input file, termed the *nominal* file. The nominal file is a functioning PRZM input file, taken from the FEMVTF site-specific comparison exercises [2,3]. The appropriate weather and run control files must also be available to the program.

Once the input file has been populated, the batch program executes PRZM. The PRZM runs under the DOS operating system, so a DOS window is spawned. A series of VBA routines tracks the progress of the program, and when the DOS window menu bar reads *Finished,* the DOS window is closed and another FORTRAN program is run that parses the PRZM time-series output file and writes the output into another comma-delimited text file. A VBA statement then imports this file into the *outputs* worksheet. The cell on the *outputs* worksheet corresponding to outputs of interest are predefined by the user as CB forecasts. As the CB runs continue, the cells are constantly overwritten, but CB retains the values in memory, forming a vector for each of the outputs. The forecasts of inputs and outputs can be extracted via a CB report for further processing and statistical analyses.

## RESULTS AND DISCUSSION

### *Number of simulations and convergence*

To examine convergence in runoff simulations [2], statistics were accumulated on the sum of the edge-of-field runoff flux. A series of simulations was carried out with the IA2R data set, varying up to 19 input parameters simultaneously. This scenario represents multiple applications of a pesticide to corn in Iowa, USA. Figure 2 shows the percentage change in the output statistic from set to set (in this case, the sets were 10 simulations each). It can be seen that the change of both statistics, standard deviations, and coefficient of variation decreases rapidly and that little change is seen beyond 40 or so simulation sets (400 total simulations).

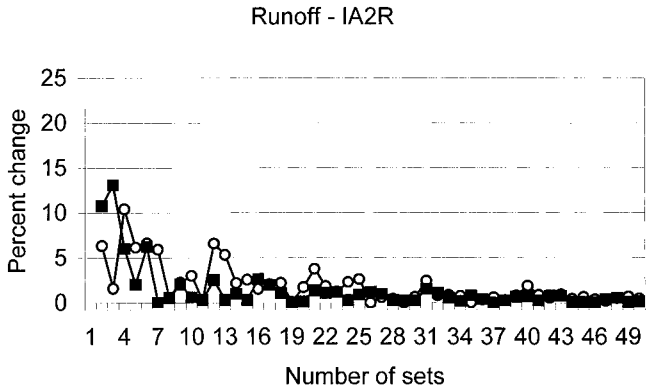


Fig. 2. Effect of number of simulations on convergence, IA2R (IA, USA). ○ = standard deviation, ■ = coefficient of variation.

For leaching simulations [3], the output tracked was the average surface soil layer concentration of pesticide or bromide tracer over the first year of simulation. Convergence was tested for scenario GA1L, representing 25 chemical applications to sweet corn in Georgia, USA; bromide tracer was applied once. An example of this convergence of the pesticide concentration is shown in Figure 3 for this scenario; bromide tracer concentration convergence is shown in Figure 4. Primarily because of the smaller number of inputs varied, even faster convergence can be seen for these leaching scenarios than was exhibited in Figure 2; based on this, 500 simulations (50 sets of 10) were deemed sufficient to characterize the uncertainty for all the scenarios.

### *Output reporting*

For leaching simulations, total soil and soil water concentrations for pesticide were captured (PRZM time-series output variables TCON and DLYS) at 30-cm increments down the simulated soil profile; in addition, pore-water concentration were simulated for bromide tracer. In the runoff simulations, edge-of-field water and sediment flux (RUNF and ESLS) as

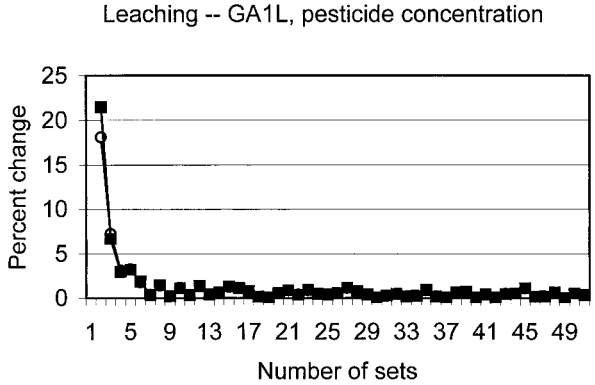


Fig. 3. Effect of number of simulations on convergence, GA1L (GA, USA), total soil pesticide concentrations. ○ = standard deviation, ■ = coefficient of variation.

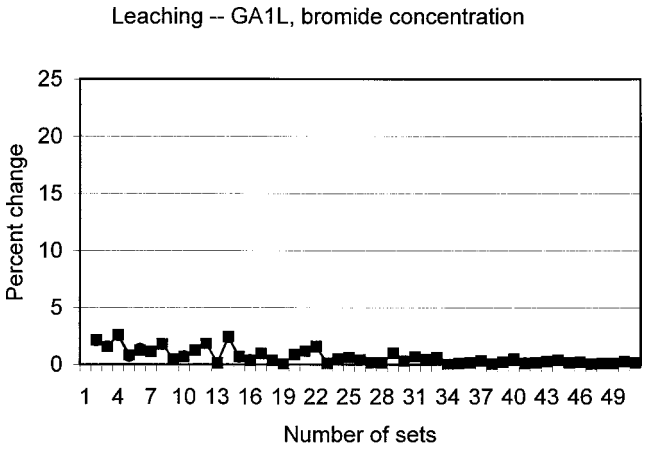


Fig. 4. Effect of number of simulations on convergence, GA1L (GA, USA), total soil bromide concentrations. ○ = standard deviation, ■ = coefficient of variation.

well as pesticide flux in water and sediment (RFLX and EFLX) were collected. All the outputs were collected on a daily time scale over a two-year span for comparison against the experimental values. The CB *generate report* function was used to write the simulation outputs and corresponding inputs into external files for further processing [6].

## CONCLUSIONS

A flexible Monte Carlo interface has been built for the PRZM 3.12 pesticide transport model that allowed the FEMVTF to perform detailed analyses of input uncertainty and the corresponding effects on model outputs. The system uses standard tools, supports a wide range of input parameters and input distributions, and allows easy export of results. Further developments of the interface have included incorporation of the EXAMS II aquatic fate model [10] as well as the ability to vary any model input.

## REFERENCES

1. Jones R, Mangels G. 2002. Review of the validation of models used in Federal Insecticide, Fungicide, and, Rodenticide Act environmental exposure assessments. *Environ Toxicol Chem* 21: 1535–1544.
2. Singh P, Jones R. 2002. Comparison of pesticide root zone model 3.12: Runoff predictions with field data. *Environ Toxicol Chem* 21:1545–1551.
3. Russell MH, Jones R. 2002. Comparison of pesticide root zone model 3.12: Leaching predictions with field data. *Environ Toxicol Chem* 21:1552–1557.
4. Wolt JD, Singh P, Cryer SA, Lin J. 2002. Sensitivity analysis for validating expert opinion as to ideal data set criteria for transport modeling. *Environ Toxicol Chem* 21:1558–1565.
5. Carbone JP, Havens PL, Warren-Hicks W. 2002. Validation of pesticide root zone model 3.12: Employing uncertainty analysis. *Environ Toxicol Chem* 21:1578–1590.
6. Warren-Hicks W, Havens PL, Carbone JP. 2002. Using Monte Carlo techniques to judge model prediction accuracy: Validation of pesticide root zone model 3.12. *Environ Toxicol Chem* 21: 1570–1577.
7. U.S. Environmental Protection Agency. 1997. Guiding principles for Monte-Carlo analysis. EPA/630/R-97/001. Risk Assessment Forum, Washington, DC.
8. Carsel RF, Mulkey LA, Lorber MN, Baskin LB. 1985. The pesticide root zone model (PRZM): A procedure for evaluating pesticide leaching threats to groundwater. *Ecol Model* 30:56–69.
9. Carsel RF, Imhoff JC, Hummel PR, Cheplick JM, Donigian AS Jr. 1998. *PRZM-3, A Model for Predicting Pesticide and Nitrogen Fate in the Crop Root and Unsaturated Soil Zones: User's Manual for Release 3.12.* U.S. Environmental Protection Agency, National Exposure Research Laboratory, Athens, GA.
10. Burns LA, Cline DM, Lassiter RR. 1982. *Exposure Analysis Modeling System (EXAMS): User Manual and System Documentation.* EPA-600/3-82-023. U.S. Environmental Protection Agency, Washington, DC.