Integrating Economic and Environmental Process Models: An Application of CEEPES to Atrazine

Richard Cabe, Peter J. Kuch, and Jason F. Shogren

Staff Report 91-SR 54
May 1991

Integrating Economic and Environmental Process Models: An Application of CEEPES to Atrazine

by Richard Cabe,
Peter J. Kuch, and Jason F. Shogren

Staff Report 91-SR 54
May 1991

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011

Richard Cabe is assistant professor of economics, New Mexico State University; Peter J. Kuch is a senior agricultural economist with the Office of Policy Analysis, EPA; and Jason Shogren is assistant professor of economics and head of the Resource and Environmental Policy Division, CARD.

This paper reports progress in a research project designed to provide information for EPA decision making regarding atrazine and related herbicides. The present state of development of the analytical system is the result of efforts of many individuals who comprise the CEEPES research team. Special acknowledgment is due to Alicia Carriquiry, Bob Carsel, Phil Gassman, Chuck Hintermeister, Stan Johnson, Andy Manale, and Mike McCorkle. Richard Cabe would also like to acknowledge financial support of the Resources and Technology Division, Economic Research Service, USDA. Any opinions expressed are those of the authors and should not be attributed to the individuals or organizations acknowledged here. This paper may not be quoted without permission of the authors.

CONTENTS

FIGURE

ABSTRACT

Atrazine is the most widely used herbicide for corn and sorghum and the most commonly encountered in surface water and groundwater. In addition to water quality problems, atrazine poses hazards through atmospheric transport, food residues, and exposure of applicators and wildlife. If atrazine use is restricted, substitute herbicides will come into wider use, increasing the likelihood of occurrence of their own sets of undesirable side effects and imposing cost or efficacy penalties. This paper describes a configuration of the Comprehensive Environmental Economic Policy Evaluation System (CEEPES) to provide decision support for regulation of atrazine and related corn and sorghum herbicides.

Comprehensive evaluation of herbicide policies requires assessment of producers' substitution responses and their market consequences; simulation of environmental fate and transport processes that lead to herbicide concentrations in environmental media; evaluation of implied human health risks and other damages caused by these concentrations; and comparison of quantitative measures of the policies' market and nonmarket consequences. The present approach uses a mathematical programming model of agricultural decision making to simulate producer response to alternative policies and uses simulation models of the physical processes of herbicide fate and transport in multiple media. The system uses a novel approach to integration of models and aggregation of outcomes based on experimental execution of simulation models and statistical estimation of response functions that summarize required process model outcomes over a space of admissible policies. Aggregation is then accomplished by integrating estimated parametric forms over the empirical distribution of land attributes relevant to agricultural productivity, environmental fate and transport, and consequent damages.

INTEGRATING ECONOMIC AND ENVIRONMENTAL PROCESS MODELS: AN APPLICATION OF CEEPES TO ATRAZINE

There are two senses of the word *integration* crucial to a simulation approach to developing information for environmental policy decisions. One necessary integration is the "bringing together into a unified whole" simulation models that differ greatly in structure and purpose but need to be used in conjunction to solve a specific problem. The other requisite "integration" is the adding up of effects that differ over a distribution of heterogeneous polluters or sufferers of environmental damage.

The purpose of this paper is to discuss an approach to these integration problems that applies response function methods, and to describe current research being conducted to configure CEEPES to support EPA decisions regarding atrazine and related herbicides. The application of the method of response functions involves estimating relatively simple parametric forms to approximate a restricted set of outcomes of complex process models, thereby abstracting from detail not needed for intended policy evaluations, and using interpolation to estimate outcomes for experimental conditions that have not been directly simulated with the process model. Simple response functions calibrated against process model experiments approximate required outcomes from the complex process models, and the simplification afforded by the response functions allows integrated evaluation of the consequences of altering policy parameters, without resorting to the original process models.

The following sections contain a statement of the problem, a discussion of the configuration of models needed to address the problem, and the role of response functions in accomplishing the two necessary types of integration. The final section offers a summary and conclusions.

Problem Statement

Atrazine often appears at the tops of lists of pesticides observed through groundwater or surface water monitoring. In a survey of 150 sampling sites in 10 midwestern states, Goolsby and Thurman (1990) report detection of atrazine in 90 percent of samples taken before the crop year's application date, and in 98 percent of samples collected during periods of runoff immediately following application. In more than one-half of the post-application samples, atrazine concentrations exceeded the EPA's lifetime health advisory level. In February 1990, the Iowa Department of Natural Resources, in conjunction with the University of Iowa Center for Health Effects of Environmental Contamination, released results of a survey of 686 private wells in rural Iowa. Atrazine was detected in 8 percent of the sampled wells, with levels exceeding the lifetime health advisory limit in 1.2 percent (Iowa Department of Natural Resources 1990). At the same time, atrazine's widespread use (it is the most widely used herbicide on corn and sorghum) suggests its economic importance. Osteen and Kuchler (1986) find that "banning atrazine would have smaller effects than would banning all triazines (of which atrazine is a member), but a greater effect than would banning all thiocarbamates, dinitroanilines, foliar insecticides, nematicides, or any other herbicide examined."

Thus, there appears to be a need for careful comparison of the risks and benefits of atrazine use. Since eliminating atrazine use would lead corn and sorghum producers to adopt alternative herbicides and alter practices in other ways, it is important to realize that a new policy toward atrazine does not simply remove both risks and benefits. If atrazine use is restricted, substitute herbicides will be used more frequently, increasing the likelihood of new sets of undesirable side effects and imposing cost or efficacy penalties. Analysis of the atrazine problem should therefore focus on policy alternatives rather than on atrazine. What is needed is an analysis of the risks and benefits of alternative policies in comparison to the risks and benefits of present atrazine use.

In addition to water quality concerns, in which atrazine is prominent, atrazine and its potential substitutes pose hazards through atmospheric transport, food residues, and exposure of applicators and wildlife. Comprehensive evaluation of herbicide policies therefore requires assessing producers' substitution responses and their market consequences; simulating environmental fate and transport processes that lead to herbicide concentrations in several environmental media; evaluating the implied human health risks and other damage caused by those concentrations; and comparing quantitative measures of the policies' market and nonmarket consequences.

Configuration of CEEPES for Atrazine Decision Support

CEEPES is a suite of models that can be configured to address a wide variety of policy evaluation questions concerning environmental consequences of agricultural activity. This section outlines a configuration of CEEPES intended to evaluate alternative policies concerning atrazine and functionally related chemicals used in corn and sorghum production. The configuration allows damages to occur through multiple media, with the possibility of interregional transport through some media, and it allows farmers to substitute among inputs, crops, and agricultural practices in response to a specified policy. Emphasis is on the evaluation of policy alternatives to the status quo and takes into account the site-specific nature of fate and transport, agricultural production, and damage mechanisms.

Range of Policies to Be Evaluated

The set of policies susceptible to evaluation is referred to as the system's policy space. This set of admissible policies is considered in configuring the system, and becomes limited by the final system implementation. Input variables or experimental conditions for the simulation models are ultimately determined by the range of policies to be evaluated by the system. Thus, if the configuration is designed to evaluate policies that take the form of selective banning of atrazine for specified crop/geographic area combinations, then the input variables of the necessary response functions, and the experimental design of simulation runs, would be determined in accordance with that range of policies. If a subsequent policy proposal is based on restricted timing of application, it would probably not be susceptible to evaluation under the same configuration. Many details of system design depend critically on the domain of policies specified, and there is a clear trade-off between the dimensions of policy space to be accommodated and the extent and difficulty of process modeling required.

The guiding concept in the definition of policy space for this configuration of CEEPES is "targeting." The objective of targeted policies is to interfere as little as possible with individual decisions. Policies can be targeted in a number of dimensions, the most important probably being geographic targeting. Targeting restrictions to specific geographic areas is promising when the potential harm or control cost differs greatly among the soils, climatic regimes, population densities, or settlement patterns associated with different geographic locations. The system is being configured to allow geographic targeting at the county level on the basis of monitoring evidence or prediction of great hazard. Current plans allow restrictions to be targeted to specific technologies (tillage practice or application method) and to a limited extent, according to timing of application within the season. Other forms of temporal targeting may be promising, but current plans only consider within-season timing implications of substitution among preplant, preemergence, and postemergence application of herbicides.

Model Specification

The system consists of a chain of simulation models, linked together as tightly as possible to form an integrated evaluation system to estimate indicators of both market and nonmarket consequences of alternative policies toward atrazine and related herbicides. Figure 1 is a schematic representation of the evaluation system. Initially, the blocks in this diagram were interdependent research projects. As the research progresses they become blocks of code in a computer program designed to evaluate alternative policies for atrazine and functionally related chemicals.

Policy Specifications. This component specifies the policy to be evaluated in a form suitable for subsequent economic modeling to determine choices of agricultural practices. As discussed above, the class of alternative policies to be analyzed must be specified as an element of the design of simulation experiments and linkages among models.

ALMANAC. This model (Williams et al. 1990), a successor of EPIC, is used to estimate the physical productivity attributes of production regimes that could be chosen in response to policy alternatives of interest. The model simulates the simultaneous growth of a crop and a weed, including competition for water, nutrients, and solar radiation.

RAMS. This component models the choice of agricultural practices under alternative policies. Outcomes to be determined include acreage planted, rotation, tillage practice, chemical regime, yield, and cost of production for each geographic unit in the study area. The model relies on data developed to support ARIMS, the CARD linear programming model of U.S. agriculture, but uses profit maximization rather than cost minimization as the objective function, allowing more realistic modeling of government agricultural programs.

Fate and Transport. These models use information on agricultural activity in each geographic unit and produce damage-relevant concentration measures, for each damage category, for the

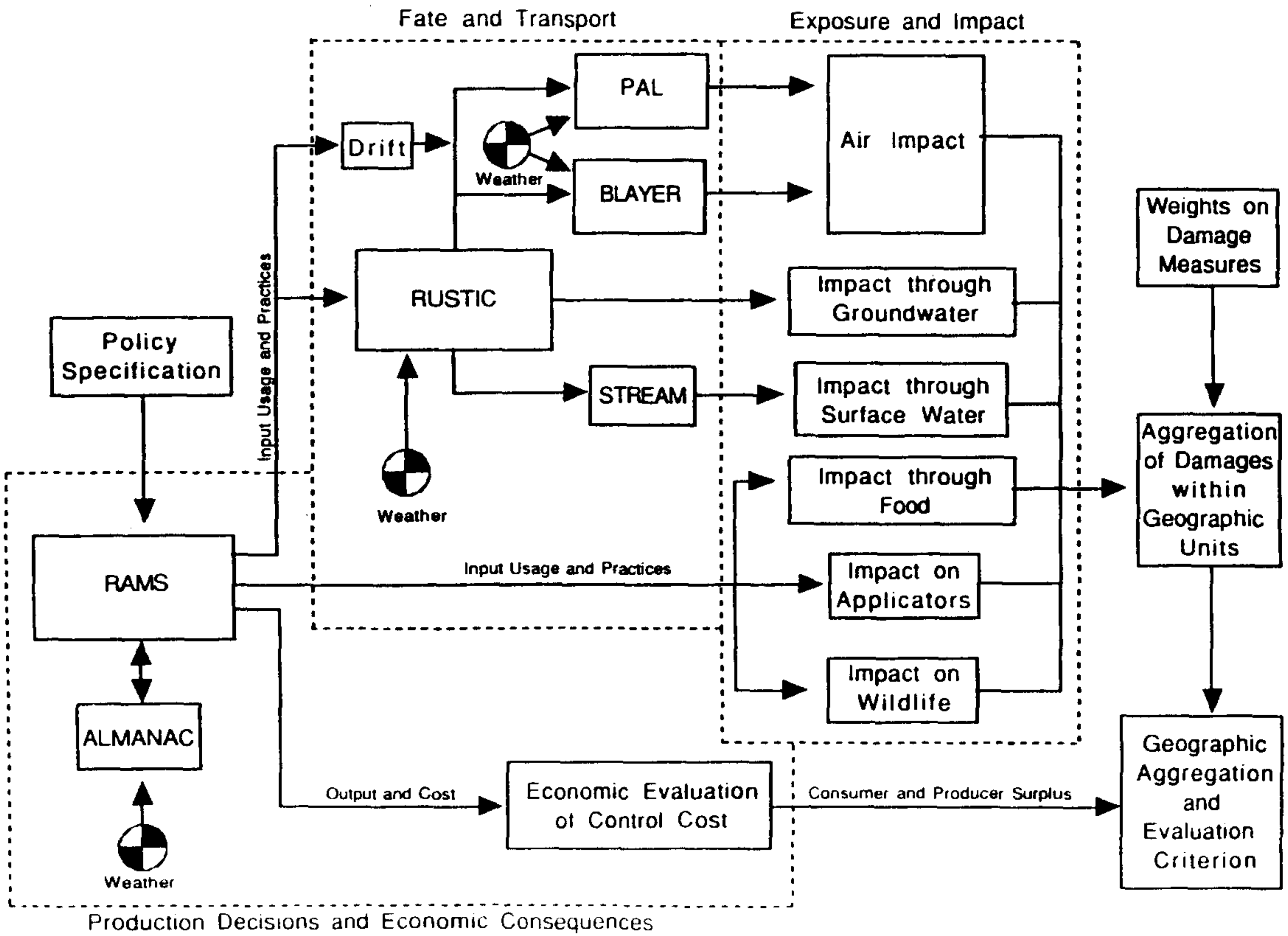

Figure 1. Configuration of CEEPES to evaluate atrazine policies

geographic unit where the chemical was applied, as well as other geographic units that may be affected by pollutant transport. For example, given a suitable description of agricultural activity in a rural central Corn Belt geographic unit, the fate and transport components estimate shallow groundwater concentrations relevant to domestic wells in that geographic unit, surface water concentrations in the area, and contributions to air concentrations in Chicago. The fate and transport component transforms a vector describing agricultural activity in all geographic units into a vector of ambient concentration measures for each medium in all geographic units. The concentration measures reported by the fate and transport component can include expected values, or they may convey information on the probabilities of various concentrations, or the distributions of concentrations over time. Outcomes of greatest interest will be 24-hour peak concentrations for acute toxicity and annual average concentrations for long-term exposure.

Exposure and Impact Component. These algorithms estimate physical measures of impact resulting from concentrations reported by the fate and transport component. These models incorporate such considerations as the distribution of population over the geographic units, toxicity of the chemical whose concentration is being evaluated, and the distribution of behavior that translates ambient concentration into the pertinent measure of damage, such as drinking water from a shallow rural domestic well, breathing polluted air, or eating foods containing herbicide residuals. Damage measures take the form of risk indicators, such as expected excess of various types of morbidity or mortality for each geographic unit.

Aggregation of Damages Within Geographic Units. This component calculates a weighted sum of physical damage measures for each geographic area using weights determined in another block. The output of this block is a damage measure suitable for aggregation over geographic areas. For a cost/benefit evaluation criterion the damage measure must be denominated in dollars to allow comparison with the control cost imposed by the policy under evaluation. For a cost-effectiveness evaluation criterion, the damage measure is regarded as a quantitative measure of the extent of improvement or risk reduction accomplished by the policy being evaluated.

Weights on Damage Measures. This component determines weights to be applied in the aggregation of physical measures of impacts. For a cost/benefit criterion these weights can be interpreted as the amount of control cost the decision maker is willing to impose on society in order to achieve a one-unit decrease in the associated impact measure. For a cost-effectiveness criterion, the weights indicate relative importance of the various physical impact measures without necessarily determining the importance of impact measures relative to control costs. In either case, the ratio of weights for two impact measures indicates the amount of increase of one measure the decision maker would tolerate in order to achieve a one-unit decrease in the other measure.

Geographic Aggregation and Reporting Outcomes. This component calculates simple sums as needed. The primary outcomes to report are the aggregate net benefit of the policy under evaluation and the geographic distribution of damage reductions and control costs associated with the policy.

Defining Geographic Units of Analysis

Each of the models used in this configuration of CEEPES relies on inputs that, in general, vary from one location to another. The natural geographic units of analysis are areas that are considered homogeneous for each model being evaluated; and the scope of this homogeneity is determined both by the model's structure and by the nature of available data. For example, a model that must rely on data only available at the scale of soil series maps cannot distinguish between two fields in different locations within the same soil series, even though the model structure is based on field-sized geographic units. Aggregation from the field level used in the model to the geographic scale available for input data may be a nontrivial task, but without additional data such a model cannot provide information of finer resolution than the level of units on the soil map used for input data, and these geographic units can be regarded as natural geographic units for this application. If additional data are available to make the model's outcome depend on location within a soil map unit--for example, two weather series might be used for different parts of a soil map unit--then the geographic resolution of the model is finer, and its natural units are the intersections of the soil map units and the areas associated with the two weather stations.

Natural geographic units implied by the structures of the various models, and by the various sources of data, will not generally fall into direct correspondence with one another. In going from a model of finer geographic resolution to one of coarser resolution, it will be necessary to aggregate over geographic space to produce an outcome that is meaningful to the model of coarser resolution. On the other hand, when a model of coarse geographic resolution generates input for a model of finer resolution, a procedure must be adopted to estimate the spatial distribution of the coarser model's output. There may be no basis on which to improve the initial approximation that the spatial distribution is uniform over the geographic units of the model of coarser resolution. In this case, disaggregation of the large area into suitable smaller areas is accomplished in accordance with this uniform distribution.

Linkages may also need to be established between models with natural geographic units that do not coincide, and neither aggregation nor disaggregation of geographic units suffices to accomplish the desired linkage between models. Such instances of overlap of natural geographic units of different models reduce to a combination of aggregation and disaggregation problems. Given the constraint of available data, basic geographic units for the integrated system are defined so that no component uses geographic units smaller than these basic units. Such units are defined by starting with the entire study area and then allowing each model to subdivide the area into units corresponding to that model's structure and available data. Basic units are then defined as the mutually exclusive and exhaustive areas resulting from this process of subdividing the study area. When this process of subdividing is complete, each model uses units that are an aggregation of basic units, and the problem of noncorresponding geographic units of interacting models is solved by disaggregating outcomes from one model to the basic geographic units and then aggregating over those basic units to create suitable input for the next component.

The Role of Response Functions

The problem of ease of computation must be overcome in order to accomplish both the integration of diverse process models and the integration of outcomes over a distribution of diverse input vectors. With the ability to easily call forth an execution of a simulation model every time its predicted outcome is required for the next model in a chain, complex simulation models can be linked as closely as their structures allow. To integrate outcomes over a distribution of input conditions, the system must be able to easily produce the simulated outcome for any specified set of input conditions so that every observation in an empirical distribution of input vectors can be transformed to the desired outcome measure, weighted appropriately, and added into the sum. In practice, such hammer-and-tongs methods are rarely feasible, either because of computation time limitations, or because of some manual operations required to transform one model's output into a form suitable for the next model's input. The method of response functions addresses both of these difficulties (also see Box and Draper 1987).

Estimating simple response functions to approximate outcomes of complex process models is a powerful means of simplifying computation necessary for both types of integration required by the policy evaluation system described above. Further, from the point of view of organizing interdisciplinary research, there is an intrinsic advantage in the response function method's requirement of careful specification of functional relationships that one work group will derive, and others will rely on for further analysis. These advantages come from being very explicit about what outcome measures are needed from each model, what input conditions will be allowed to vary in response to policy changes, and by using statistical methods to approximate the model's performance in this narrowly prescribed domain, neglecting all the model's other unneeded capabilities.

An example of this method is ALMANAC (Williams et al. 1990), a plant growth model used in this configuration of CEEPES to approximate the effect of varied herbicide regimes on agricultural productivity. This information is needed to approximate farmers' herbicide application decisions and to evaluate the consequences of those decisions. The model simulates the process of plant growth on a daily basis and allows the calculation of several end-of-session magnitudes, including harvest. By repeating the simulation process with the uncertain weather of many years, it is possible to estimate the expected value of yield. Further, the model can use information on the effect of weed control practice on weeds present or appearing, so the model can be used to estimate expected yield for specified herbicide types and application rates. Given unlimited resources for computation, it would be possible to use the process model itself to make these estimates whenever they are called for by the evaluation system. In practice, certain evaluations would be extremely cumbersome, expensive, or impractical without some means of reducing the burden of computation implied by this direct method.

The method of response functions is designed to accomplish the same purpose in a more computationally efficient way. If expectation of yield is the desired outcome to be measured, and herbicide type and application rate are the varying conditions whose effects are to be estimated, the multiyear simulation can be conducted several times, under varying regimes of herbicide type and application rate. For each multiyear simulation, the herbicide type and application rate and the calculated expected value of yield are recorded. The data from this experimenting with the process model can be used to estimate a convenient form of response function designed to summarize the relationship among variables predicted by the process model. This response function may then be used as a surrogate for the model itself, for the very limited purpose required by the policy evaluation system.

In this example, a complex process model was used to estimate the relationship between a few input variables and a single outcome measure. Each time the process model is used to directly estimate the desired outcome measure for specified values of the input variables, something on the order of a few million calculations is required. If many of these estimates are required, it quickly becomes impractical to use the process model directly, and a statistically estimated response function becomes attractive. Estimation of the response function requires careful specification of the relationship to be approximated in terms of the required outcome measure, input variables and their relevant ranges, and the nature of the functional relationship between input variables and the outcome measure. From this specification, an experimental design can be derived that calls for computation-intensive simulations to be undertaken for only the most informative specifications of input variables. Once the experiment has been carried out and the response function estimated, the required outcome measure can be approximated by a number of calculations many orders of magnitude smaller than required by direct application of the process model. It should be emphasized that the process model has been replaced by the response function only for the very limited purposes for which the experimental executions of the model and estimated response function were specified.

An example of the use of response function estimation to avoid manual operations sometimes required in linkages between models is afforded by the use of STREAM (Donigian et al. 1984) to transform runoff estimates from RUSTIC (Dean et al. 1989) into estimates of surface water concentrations suitable for impact analysis. STREAM is not a computer model at all, but a graphical representation of sensitivity runs of HSPF. Use of STREAM to transform runoff measures into surface water concentrations is described in Donigian and Carsel (1987), but the manual approach taken there does not permit the integration of models desired in this configuration of CEEPES. Instead, the model was executed by working through the graphical look-up and interpolation procedures for experimentally chosen values of inputs to get the model's predicted value of the desired outcome, surface water concentration. The estimated response function included all parameter estimates significant at levels of .001 or better, and R^2 for the regression of .90. It should be noted that this simple parametric form does not replace STREAM for all purposes, but only for those purposes required for this problem.

Summary and Conclusions

This paper has discussed an approach to certain integration problems important in simulating the consequences of alternative environmental policies, and it has described current related research applying CEEPES to support EPA decisions regarding atrazine and related herbicides. This application follows the market consequences of policies through simulation of producers' economic choice of inputs and technology, and follows nonmarket consequences of policies through simulation of fate and transport of applied herbicides. The system follows environmental consequences of herbicide use through short- and long-range air transport, and through various processes leading to herbicide concentrations in groundwater and surface water. In general, peak concentrations are estimated to assess acute toxicity impacts, and mean concentrations are estimated to assess chronic effects.

While development of this system is ongoing, enough has been accomplished to allow some confidence in the general approach. Although the method of response functions is by no means a panacea applicable to all natural resource modeling problems, it does hold promise for integrating models that would otherwise be too unwieldy or inappropriately specified to operate together. It also allows numerical integration over a distribution of heterogeneous input vectors in cases that would be computationally infeasible if every evaluation of the function required execution of a complex numerical simulation.

REFERENCES

Box, George E. P. and Norman R. Draper. 1987. *Empirical Model-Building and Response Surfaces*. New York: John Wiley and Sons

Dean, J. D., P. S. Huyakorn, A. S. Donigian, Jr., K. A. Voos, R. W. Schanz, Y. J. Meeks, and R. F. Carsel. 1989. *Risk of Unsaturated/Saturated Transport and Transformation of Chemical Concentrations (RUSTIC)*. Athens, Georgia: U.S. Environmental Protection Agency.

Donigian, Anthony S., Jr., and Robert F. Carsel. 1987. "Modeling the Impact of Conservation Tillage Practices on Pesticide Concentrations in Ground and Surface Waters." *Environmental Toxicology and Chemistry* 6:241-50.

Donigian, A. S., Jr., D. W. Meier, and P. P. Jowise. 1984. *Stream Transport and Agricultural Runoff of Pesticides for Exposure Assessment: A Methodology*. EPA/600/3-86/011a. Athens, Georgia: U.S. Environmental Protection Agency.

Goolsby, Donald A. and E. Michael Thurman. 1990. "Herbicides in Rivers and Streams of the Upper Midwestern United States." Forthcoming in *Proceedings of the 46th Annual Meeting of the Upper Mississippi River Conservation Committee*. March 13-15, 1990. Bettendorf, Iowa.

Iowa Department of Natural Resources. 1990. "The Iowa State-wide Rural Well Water Survey Design Report: A Systematic Sample of Domestic Drinking Water Quality." *Technical Information Series* 17:February.

McCorcle, M. D. 1988. "Simulation of Surface Moisture Effects on the Great Plains Low Level Jet." *Monthly Weather Review* 116:1705-20.

Osteen, Craig and Fred Kuchler. 1986. "Potential Bans of Corn and Soybean Pesticides: Economic Implications for Farmers and Consumers." *Agricultural Economic Report* 546:1-24.

Williams, J. R., J. R. Kiniry, P. W. Gassman, and P. deBaeke. 1990. "ALMANAC/WEED: A Model for Weed-Crop Interaction." American Society of Agronomy Meeting, San Antonio, Texas. October.