United States Environmental Protection Agency
Research and Development

Risk Reduction Engineering Laboratory
Cincinnati, OH 45268

EPA/600/SR-92/145 September 1992

EPA

Project Summary

AutoMOUSE An Improvement to the MOUSE Computerized Uncertainty Analysis System Operational Manual

Albert J. Klee

Environmental engineering calculations involving uncertainties in data are well beyond the capabilities of conventional analysis for any but the simplest models. There exists a number of general-purpose computer simulation languages that are capable of such analysis, but these languages are difficult to learn and implement quickly.

The original MOUSE (Modular Oriented Uncertainty SystEm) system was designed to deal with the problem of uncertainties in static mathematical models, such as a set of engineering cost or risk analysis equations. It was especially intended for use by individuals with little or no knowledge of computer languages, programming, or simulation. The MOUSE system runs on MS-DOS-based personal computers. It is easy and fast to learn and has all of the features needed for substantive uncertainty analysis, such as built-in probability distributions, plotting and graphing capabilities, sensitivity analysis, and interest functions for cost analyses. A series of unique companion utility programs help (1) analyze sample data to determine the probability distributions that best fit those data and (2) check each program for errors in syntax.

AutoMOUSE is a significant improvement to the original MOUSE system. It actually writes the computer program necessary to carry out the uncertainty analysis. The input to AutoMOUSE consists of the equations of the model and requires no knowledge of computer programming. It is designed primarily for beginners but is also of value for those who have some programming experience and wish to construct MOUSE programs more quickly and with fewer errors.

Some typical examples of the use of MOUSE within the U.S. Environmental Protection Agency include studying the migration of pollution plumes in streams, establishing regulations for hazardous wastes in landfills, and estimating pollution control costs.

This Project Summary was developed by EPA's Risk Reduction Engineering Laboratory, Cincinnati, OH, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Introduction

Models consisting of one or more mathematical equations are extremely important for they are used in a number of professional disciplines including economics, engineering, and the health sciences. In environmental engineering, for example, they are common in cost and risk analysis calculations. Such models can be either deterministic (input variables are single numbers) or stochastic (input variables are in the form of probability distributions that reflect the uncertainty about their values). Although the solution of deterministic models is well understood, stochastic models are more difficult. The following simple example will illustrate this point. It is taken from an actual EPA problem, where the objective was to develop a method to establish the regulatory

Printed on Recycled Paper

level of wastes disposed of in a landfill which pose a hazard due to toxic organic constituents. The environment modeled was that of a non-secure sanitary landfill that receives a small amount of toxic industrial waste.

In the model, the regulatory level (REGLEV) of a toxic constituent was obtained by multiplying the maximum permitted daily exposure level (DELMAX) for the toxic constituent in question by a suitable factor, i.e.,

$$\text{REGLEV(ppm)} = \text{DELMAX(mg/day)} \times \text{FACTOR(days/kg)} \quad [1]$$

The factor involved four variables as follows:

FACTOR =
[(attenuation factor x (liters leached/year) x years]/[(liters consumed/day) x (industrial waste quantity, kg)]. [2]

Defining PRECIP as the annual precipitation, RLEACH as the rate of leaching, DEPTHL as the landfill depth, and WDENS as waste density, and assuming that: (1) the (dimensionless) attenuation factor is 100, (2) there are 2 liters of water consumed per day, (3) the exposure period is 70 yr, and (4) 5% of the waste is industrial, the factor becomes:

$$\text{FACTOR} = 70{,}000^*[\text{PRECIP}^*\text{RLEACH}]/[\text{DEPTHL}^*\text{WDENS}] \quad [3]$$

To determine a regulatory level using the model of Equation 1, a value of FACTOR must first be determined. In Equation 3, using average estimates of annual precipitation of 1,143 liters/meter per year, rate of leaching of 0.75, landfill depth of 8.35 meters, and waste density of 415.3 kg/meter^3, FACTOR equals 17,303. Based upon this value, Table 1 shows the regulatory levels for compounds with varying DELMAX's.

Table 1, however, is a *deterministic* solution; it does not take into account any uncertainty in the four stochastic input variables. Table 2 shows the probability distributions for these variables which reflect the best judgment of the investigators. The problem was to calculate the value of FACTOR, and hence the value of REGLEV, given these input probability distributions. A common approach to the solution of mathematical models which involve uncertainty is a form of Monte Carlo simulation known as Model Sampling:

(1) A value for each of the input variables is drawn at random from their respective probability distributions, and the model's output is computed using this particular set of values.
(2) The above process is repeated many times. Since the results vary with each iteration, the output is gathered in the form of a probability distribution.

Figure 1 shows a Monte Carlo solution (using MOUSE) to the distribution of the variable FACTOR. The statistics reported include the mean, standard deviation, coefficient of variation, and minimum and maximum of the output variable, FACTOR. Graphical output is provided in the form of (a) the frequency distribution (asterisks) and (b) the cumulative frequency distribution (circles) of FACTOR. Histogram statistics are also provided, such as number of entries, percent entries, cumulative percent entries, and complement of the cumulative percent entries for each interval of the histogram.

AutoMOUSE

The MOUSE program used to obtain Figure 1 is shown in Figure 2. It can be seen that even this simple uncertainty problem poses a complex task for those unfamiliar with programming. It is no wonder then that uncertainty analyses are either left up to highly skilled professionals, frequently at high costs and long lead times, or else omitted altogether. For this reason, AutoMOUSE was developed to provide a simple, clean, non-programming interface to the MOUSE model. With AutoMOUSE, anyone can develop a complete uncertainty analysis model without any knowledge of, or contact with, the programming language itself.

There are two major input stages to AutoMOUSE. The first consists of the model equations. This is followed by the specification of the probability distributions for the uncertain variables in the model and the arguments for these distributions. AutoMOUSE operates in either of two modes: (a) a Non-Expert mode and (b) an Expert mode. The former, intended for beginners to the MOUSE system and which places certain restrictions on the model, will be used in the following exposition.

The Non-Expert mode equation entry screen is shown in Figure 3, with the simple FACTOR=70000*(PRECIP*RLEACH)/(DEPTHL*WDENS) equation example entered. The screen consists of three parts: (1) Top: equation/error reporting area, (2) Middle: equation entry area, and (3) Bottom: equation instruction area. Each equation is allotted 240 characters, denoted by the shaded lines. The instructions for moving around the equation are shown below the shaded lines. Entering Alt-F brings up a help screen that reminds the user of the symbols used for the various arithmetic operations and functions permitted in this mode. The INS key toggles the insert mode on and off (when

Table 1. *Regulatory Levels as a Function of DELMAX*

Compound	*DELMAX (mg/day)*	*Regulatory Level(ppm)*
Compound A	*.000001*	*.017*
Compound B	*.00001*	*.17*
Compound C	*.0001*	*1.7*
Compound D	*.001*	*17*
Compound E	*.01*	*170*
Compound F	*.1*	*1700*
Compound G	*1.0*	*17000*

Table 2. *Distributions of the Input Stochastic Variables*

Variable	*Type*	*Outcome*	*Cumulative Probability*
RLEACH	*Continuous*	*0.5*	*0.00*
	Uniform	*1.0*	*1.00*

Variable	*Type*	*Parameters*	
PRECIP	*Trapezoidal*	*762*	*(= lowest value)*
		1016	*(= most likely value, lower boundary)*
		1270	*(= most likely value, upper boundary)*
		1524	*(= highest value)*
DEPTHL	*Trapezoidal*	*3.0*	*(= lowest value)*
		6.1	*(= most likely value, lower boundary)*
		9.1	*(= most likely value, upper boundary)*
		15.2	*(= highest value)*
WDENS	*Triangular*	*297*	*(= lowest value)*
		415	*(= most likely value)*
		534	*(= highest value)*

insert mode is on, characters are inserted into the lines; when insert mode is off, new characters overwrite the old). After an equation has been entered and edited, the user presses the ESC key to continue to the next equation. After entering all of the model's equations, the ESC key is pressed twice to end model entry. An equation can be entered anywhere within the shaded three lines. AutoMOUSE will remove unnecessary spaces when it scans the equation for errors.

During equation input and editing, if the user attempts to enter an illegitimate character or a character that cannot legitimately follow the preceding character, AutoMOUSE will beep and reject the character. An explanation of the error is also displayed on the screen, just below the equation entry area. AutoMOUSE makes two checks for errors: (1) the preliminary check, just described, as the user enters the equation, and (2) a more detailed scan after the equation is entered and the user has pressed the ESC key. The more detailed check involves checking the equation for legitimate numbers or variable names; correct use of parentheses, commas and equal signs; allowable functions and function arguments; etc. If

DISTRIBUTION FOR QUANTITY FACTOR

NUMBER OF ITERATIONS = 5000

MEAN =	20361.21000	STANDARD DEVIATION =	12154.16000
MINIMUM =	4219.64600	COEFFICIENT OF VARIATION, % =	59.69275
MAXIMUM =	82200.32000		

LOWER LIMIT	NUMBER OF ENTRIES	PERCENT ENTRIES	CUMULATIVE % ENTRIES	CUMULATIVE COMPLEMENT	DISTRIBUTIONS * = FREQUENCY DISTRIBUTION O = CUMULATIVE DISTRIBUTION
4200.	60.	1.20	1.20	98.80	*O******
6100.	245.	4.90	6.10	93.90	****O*************************
8000.	443.	8.86	14.96	85.04	**********O**
9900.	585.	11.70	26.66	73.34	******************O***
11800.	486.	9.72	36.38	63.62	*************************O********************************
13700.	467.	9.34	45.72	54.28	*******************************O************************
15600.	381.	7.62	53.34	46.66	*************************************O********
17500.	374.	7.48	60.82	39.18	***O**
19400.	248.	4.96	65.78	34.22	***************************** O
21300.	239.	4.78	70.56	29.44	**************************** O
23200.	194.	3.88	74.44	25.56	*********************** O
25100.	158.	3.16	77.60	22.40	****************** O
27000.	127.	2.54	80.14	19.86	*************** O
28900.	121.	2.42	82.56	17.44	*************** O
30800.	107.	2.14	84.70	15.30	************* O
32700.	94.	1.88	86.58	13.42	************ O
34600.	95.	1.90	88.48	11.52	************ O
36500.	69.	1.38	89.86	10.14	******** O
38400.	75.	1.50	91.36	8.64	********* O
40300.	70.	1.40	92.76	7.24	******** O
42200.	56.	1.12	93.88	6.12	******* O
44100.	56.	1.12	95.00	5.00	******* O
46000.	55.	1.10	96.10	3.90	******* O
47900.	35.	.70	96.80	3.20	***** O
49800.	26.	.52	97.32	2.68	**** O
51700.	24.	.48	97.80	2.20	**** O
53600.	28.	.56	98.36	1.64	**** O
55500.	17.	.34	98.70	1.30	*** O
57400.	11.	.22	98.92	1.08	** O
59300.	15.	.30	99.22	.78	*** O
61200.	5.	.10	99.32	.68	**
63100.	6.	.12	99.44	.56	**
65000.	9.	.18	99.62	.38	**
OVERFLOW	19.	.38	100.00	.00	***

CUMULATIVE % ENTRIES	CUMULATIVE COMPLEMENT	VALUE OF FACTOR
5.0	95.0	5673.4690
10.0	90.0	6936.3430
25.0	75.0	9630.4280
50.0	50.0	14767.1900
75.0	25.0	23536.7100
90.0	10.0	36677.3300
95.0	5.0	44100.0000
99.0	1.0	57906.6700

Figure 1. *Monte Carlo analysis for FACTOR.*

an error is detected, AutoMOUSE will beep, redisplay the equation, explain the error in the equation/error reporting area of the screen, and request correction of the equation.

If AutoMOUSE finds no problem with an equation, it will ask whether you wish to continue to the next equation or create a statistics summary and histogram table (of the kind shown in Figure 1) for the dependent variable of the equation. In our example, the following menu would appear (the arrow indicates a highlighted bar across the line):

```
1. Continue to next equation
2. Create table for variable factor <--
```

Pressing the Enter key selects the highlighted line). No additional input is required of the user to create a statistics summary and histogram table for this variable.

Stochastic Variable Specification

During the processing of equations, AutoMOUSE keeps track of the independent variables used in the equations. Since no numerical values have been specified for these variables, they are presumed to be stochastic or uncertain variables. When equation entry is completed, the user is prompted to specify for each stochastic variable the nature of its distribution and its parameters using the menu shown in Figure 4. The figure shows the entry for variable PRECIP where the trapezoidal probability distribution has been selected, with parameters 762, 1,016, 1,270, and 1,524. The entries for the other variables are similar.

When all of the probability distribution arguments have been supplied, a summary table appears on the screen indicating the number of equations and stochastic variables in the model and the number of table output lines requested. The MOUSE program shown in Figure 2 is written to a file, and AutoMOUSE is finished.

Error Checking and Avoidance in AutoMOUSE

There are two types of error checking in AutoMOUSE. The first scans for proper FORTRAN characters and syntax (since MOUSE is based upon that computer language). One cannot have, for example, the character "@" in a MOUSE program or have plus (+) follow minus (-) in an equation. AutoMOUSE goes even further than the usual FORTRAN compiler checker, in that it will not allow taking the square root of a negative number. In the Expert mode, the second type of error checker scans the arguments to MOUSE probability and interest functions. The probability of a "success" in a binomial distribution, for example, must not be equal to or greater than one or less than or equal to zero and there must be two arguments (i.e., the probability of a success and the number of trials). No other error checking is necessary since it is AutoMOUSE and not the user that writes all of the remaining code necessary to complete the program. Errors made by the user in writing an equation or in specifying a FORTRAN function or a MOUSE probability or interest function are caught immediately, and not after the user has entered the complete model.

Summary

A comparison with general purpose simulation languages has shown that the MOUSE/AutoMOUSE system is the quickest and easiest way to implement uncertainty in models involving equations. This is accomplished without sacrificing any of the power commonly associated with such languages. Most importantly, the MOUSE/AutoMOUSE system permits the user to focus on the model rather than on the details of coding a program. Without a system such as MOUSE, uncertainty analysis on any but the simplest waste management models would probably be skipped and the decisionmaking process would be noticeably less effective as a result.

```
      DIMENSION TAB(46)
      COMMON/Q1/C80/Q2/JUMP/Q3/LOOP/Q4/SENS/Q5/IDATA/Q6/IOUT/Q7/IPASS/Q8
     &/IRAN/Q9/IREAD/Q10/ISCR1/Q11/ISCR2/Q12/ISNAP/Q13/ITER/Q14/ITOG/Q15
     &/IWRIT/Q16/IXXX/Q17/OK1/Q18/OK2/Q19/OK3/Q20/OK4/Q21/OK5/Q22/OK6
      CALL NAME('(REGULATORY EXAMPLE)')
      ITER=5000
5     DO 10 IXXX=1,ITER
      PRECIP=TRA(762.0,1016.0,1270.0,1524.0)
      RLEACH=CUNI(.5,1.0)
      DEPTHL=TRA(3.0,6.1,9.1,15.2)
      WDENS=TRI(297.0,415.0,534.0)
      FACTOR=70000*(PRECIP*RLEACH)/(DEPTHL*WDENS)
      CALL TABLE(FACTOR,'(FACTOR)',DUM0,DUM1,40,TAB)
10    CONTINUE
      GOTO(5,15), JUMP
15    CONTINUE
      END
```

Figure 2. MOUSE program for the regulatory example.

ENTER EQUATION #1 OF THE MODEL ON THE FOLLOWING LINES.
(Leaving lines blank indicates that there are no more equations to enter.)

FACTOR=70000*(PRECIP*RLEACH)/(DEPTHL*WDENS)

To move around use ARROW keys plus HOME/END for start/end of last line, CONTROL LEFT ARROW/CONTROL RIGHT ARROW for end/beginning of line, ENTER for first column of next line, and TAB for right 8 spaces. DEL/BACKSPACE deletes the character to the right/left, ESCAPE ends the data entry, and \ toggles between erase/restore all lines. Enter Alt-F to view a summary of permissible functions and arithmetic operations in this Non-Expert mode.

INSERT MODE STATUS IS ON

Figure 3: *Equation entry screen, Non-Expert mode.*

Stochastic Variable Distribution Choices

1. Autocorrelation, Exponential	8. Discrete Uniform	17. Lognormal, Alternate Form
2. Autocorrelation, Linear	9. Empirical Beta	18. Normal
3. Beta, Standard Form	10. Empirical Cts	19. Pascal/Geometric
4. Binomial	11. Empirical Discrete	20. Poisson
5. Continuous Uniform	12. Erlang	21. Step Rectangular
6. Discrete Trapezoidal	13. Exponential	22. Trapezoidal
7. Discrete Triangular	14. Gamma	23. Triangular
	15. Hypergeometric	24. Set = to LOOP
	16. Lognormal	

Enter number of probability function desired for variable PRECIP >> 22

Trapezoidal Distribution - Example: PRECIP = TRA(2.,5.,7.,9.) where 2 and 9 are the lower and upper bounds (Parameters A and D), respectively, of the distribution, and 5 and 7 are the lower and upper bounds (Parameters B and C), respectively, of the middle of the distribution. Enter your own PARAMETERS A, B, C, and D, separated by commas or spaces, e.g., 2,5,7,9 (or Return to select a different probability distribution):

>> 762,1016,1270,1524

Figure 4. *Probability distribution specification menu.*

*U.S. Government Printing Office: 1992 — 648-080/60084

The EPA author, ***Albert J. Klee,*** *(also the EPA Project Officer, see below) is with the Risk Reduction Engineering Laboratory, Cincinnati, OH 45268*

The complete report consists of paper copy and diskette, entitled "AutoMOUSE, An Improvement to the MOUSE Computerized Uncertainty Analysis System Operational Manual".

Paper Copy (Order No. PB93-500007AS; Cost: $35.00, subject to change)

Diskette (Order No. PB93-100113AS; Cost: Contact NTIS for pricing information)

The above items will be available only from:

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: 703-487-4650

The EPA Project Officer can be contacted at:

Risk Reduction Engineering Laboratory
U.S. Environmental Protection Agency
Cincinnati, OH 45268

United States
Environmental Protection Agency
Center for Environmental Research Information
Cincinnati, OH 45268

BULK RATE
POSTAGE & FEES PAID
EPA
PERMIT No. G-35

Official Business
Penalty for Private Use
$300

EPA/600/SR-92/145