The Oilspill Risk Analysis Model of the U.S. Geological Survey

GEOLOGICAL SURVEY **PROFESSIONAL PAPER 1227**

The Oilspill Risk Analysis Model of the U.S. Geological Survey

By RICHARD A. SMITH, JAMES R. SLACK, TIMOTHY WYANT, *and*
KENNETH J. LANFEAR

GEOLOGICAL SURVEY PROFESSIONAL PAPER 1227

UNITED STATES GOVERNMENT PRINTING OFFICE, WASHINGTON: 1982

UNITED STATES DEPARTMENT OF THE INTERIOR

JAMES G. WATT, *Secretary*

GEOLOGICAL SURVEY

Dallas L. Peck, *Director*

Library of Congress Cataloging in Publication Data
Main entry under title:
The Oilspill risk analysis model of the U.S. Geological Survey
(Geological Survey Professional Paper 1227)
Bibliography: p. 35
Supt. of Docs. no.: I 19.16:1227
1. Oil spills—Mathematical models. 2. Oil spills—Data processing.
I. Smith, Richard A. II. Series.
GC1085.O44 363.7'394 80-606812 AACR2

For sale by the Branch of Distribution, U.S. Geological Survey.
604 South Pickett Street. Alexandria, VA 22304

CONTENTS

ILLUSTRATIONS

TABLES

METRIC CONVERSION FACTORS

S1 (International System of Units) is a modernized metric system of measurement. An asterisk after the last digit of the factor indicates that the conversion factor is exact and that all subsequent digits are zero, all other conversion factors have been rounded to four significant digits.

To convert from	To	Multiply by
inch (in)	millimeter (mm)	25.4*
foot (ft)	meter (m)	0.3048
mile (mi)	kilometer (km)	1.609
nautical mile	kilometer (km}	1.852*
acre	meter^2 (m^2)	4,047
mile^2 (m^2)	kilometer^2 (km^2)	2.590
gallon (gal)	meter^3 (m^3)	0.003785
barrel (bbl) (petroleum, 1 bbl = 42 gal)	meter^3 (m^3)	0.1590
million gallons per day (Mgal/d)	meter’ per second (m^3/s)	0.04381

THE OILSPILL RISK ANALYSIS MODEL OF THE U.S. GEOLOGICAL SURVEY

By RICHARD A. SMITH, JAMES R. SLACK, TIMOTHY WYANT, and KENNETH J. LANFEAR

ABSTRACT

The U.S. Geological Survey has developed an oilspill risk analysis model to aid in estimating the environmental hazards of developing oil resources in Outer Continental Shelf (OCS) lease areas. The large. computerized model analyzes the probability of spill occurrence, as well as the likely paths or trajectories of spills in relation to the locations of recreational and biological resources which may be vulnerable. The analytical methodology can easily incorporate estimates of weathering rates, stick dispersion, and possible mitigating effects of cleanup.

The probability of spill occurrence is estimated from information on the anticipated level of oil production and method and route of transport. Spill movement is modeled in Monte Carlo fashion with a sample of 500 spills per season, each transported by monthly surface-current vectors and wind velocities sampled from 3-hour wind-transition matrices. Transition matrices are based on historic wind records grouped in 41 wind velocity classes, and are constructed seasonally for up to six wind stations. Locations and monthly vulnerabilities of up to 31 categories of environmental resources are digitized within an 800,000 km^2 study area. Model output includes tables of conditional impact probabilities (that is, the probability of hitting a resource, given that a spill has occurred), as well as probability distributions for oilspills occurring and contacting environmental resources within preselected vulnerability time horizons.

The model provides the U.S. Department of the Interior with a method for realistically assessing oilspill risks associated with OCS development. To date, it has been used in oilspill risk assessments for eight OCS lease sales with the results reported in Federal environmental impact statements. .4 summary of results is presented herein. A ‘<real time” version was also used to forecast the movement of oil from the 1976-77 Argo Merchant oilspill. Additional model runs are planned for future OCS lease sales in frontier areas, Other possible applications include analysis of OCS development alternatives and site selection for oilspill cleanup equipment.

INTRODUCTION

The past decade has been a period of rapid growth in the offshore petroleum industry. The Department of the Interior currently conducts sales of mineral leases for specific areas of the Outer Continental Shelf at the rate of more than two per year, and it is anticipated that lease sales will continue, perhaps even at an increased rate, well into the 1980’s.

Oilspills are one of the major concerns associated with offshore oil development in all OCS lease sale areas. Concern is clearly strongest among those who live in coastal areas and who depend, directly or indirectly, on coastal zone resources other than oil for a livelihood. Controversy over the risks and benefits of off-shore oil development inevitably gives rise to a need for quantitative estimates of the oilspill risk involved in a particular development proposal. Within the Federal Government, oilspill risk estimates are required prior to holding an OCS lease sale, at the time the Secretary of the Interior makes decisions on tracts to be withheld from leasing because of unacceptable oilspill risk to specific environmental resources in the proposed sale area. At issue in the decisionmaking for a typical OCS lease sale are anywhere from 100 to 500 nine-square-mile tracts which have been identified as possible production areas by interested oil companies. Also at issue areas many as *20* or *30* specific resources which have been identified by the Bureau of Land Management or the U.S. Geological Survey as vulnerable to oilspills on the basis of research and communication with local authorities.

An important fact that stands out when one attempts to predict oilspill damages for a proposed OCS lease area is that the problem is fundamentally probabilistic. A great deal of uncertainty exists not only with regard to the location, number, and size of spills that will occur during the course of development, but also with regard to the wind and current conditions that will exist and give direction to the oil at the particular times spills occur. While some of the uncertainty reflects incomplete or imperfect data, considerable uncertainty is simply inherent in the problem.

The Geological Survey has developed a model for assessing the oilspill risks associated with petroleum development in Federal OCS lease areas. The model is constructed to deal with three fundamental and essentially independent factors which comprise the total oilspill risk to coastal zone resources: (1) the probability of spill occurrence as a function of the quantity of oil which is to be produced and handled at individual production sites, pipelines, and tanker routes; (2) the probabilities of occurrence of various spill trajectories from production sites and transportation routes as a function of historical wind and current patterns for the area: and (3) the location in space and time of vulnerable resources defined according to the same coordinate system used in spill trajectory simulation. Results of the individual parts of the analysis are combined to estimate the total oilspill risk associated with production and transportation at locations within a proposed lease area. This information is then used in making final tract selections prior to leasing. To date, risk analyses have been conducted for seven Federal lease areas, including sites offshore the North-, Mid-, and South-Atlantic Coasts, the Eastern Gulf of Mexico, Southern California, and the Western and Northern Gulf of Alaska.

The purpose of this report is to describe how the Oilspill Risk Analysis Model of the U.S. Geological Survey works, both in theory and in actual operation. It discusses the assumptions used in developing the model and defines the role of each computer program. While not a detailed operating instruction manual, it provides the broad understanding of the model which is necessary for operating the model and properly interpreting the results.

The report begins with a discussion of how the data base is developed, proceeds to describe how oilspills are simulated, and then reviews the results to date. The section, “Representations of Physical Data,”’ describes how winds, currents, and the locations of environmental resources, or targets, are represented as data and put in the proper form for analysis. Simulation of oilspill movement is the topic of the section, “Oilspill Trajectory Simulation,” and the probabilistic calculations of oilspill risk is covered in “Risk Calculation.” The section, “Model Verification and Limitations,” places the accuracy of risk calculations in perspective with discussions of sensitivity and verification studies. A summary of past results and ideas for future uses of the model are presented in the section, “Model Output and Case Examples." Discussion of “Practical Aspects of Operating and Managing the Model” concludes the paper.

REPRESENTATIONS OF PHYSICAL DATA

The model of the U.S. Geological Survey is designed to use a large amount of information about the physical environment, including sizable files of wind and current data and the locations of numerous environmental resources which may be adversely affected by oilspills. Model programs process all of this data and store it in computer files before any trajectories are computed. All of the files are designed to allow rapid access to the data by subsequent computer programs. An extensive system of internal checks, along with graphic displays and printouts, help ensure that physical data are represented correctly. The following section describes how physical data are collected, processed, checked, and stored.

BASE MAP

A system for representing spatial locations is the foundation of the trajectory simulation model. The model employs a Cartesian coordinate system superimposed over a base map of the study area. All stored data are referenced to this system, and it is used for all internal calculations.

The initial step in establishing a coordinate system is the delineation of the area to be modeled. This area must be large enough so that all oilspill targets likely to be affected, such as land or biological resources, are included; at the same time, the map scale must not be so large that essential details are obscured. Previous OCS lease sale analyses have typically examined areas of about 800 km by 800 km, and included 1,000 km of coastline. The base map boundaries are usually chosen so that the major origins of potential spills, such as the lease area and transportation routes, are centered; if winds or currents are expected to drive spills predominantly in a certain direction, the map is shifted accordingly. Land need only be included to the extent necessary to define the shoreline, and to aid in visual recognition of the map.

Choice of a projection for the base map is particularly important, since representing the surface of the earth by a planar surface necessarily introduces some distortion in scale, or direction, or both. The Universal Transverse Mercator (UTM) projection system has relatively little scale distortion but has a directional distortion of about 10 degrees. Because the equations for correcting this distortion are lengthy and too expensive to perform for each trajectory movement, earlier OCS lease sale analyses used UTM or Lambert projections and neglected distortion. However, neglecting distortion caused serious difficulties in combining data obtained from different maps, and necessitated use of a more general mapping system.

A useful property of the Mercator projection is that there is no distortion in direction; that is, a constant compass direction is a straight line. This makes it extremely easy to align a Mercator projection with a Cartesian coordinate system. The penalty for this, however, is extreme distortion in scale, particularly at high latitudes. Fortunately, the correction factor is a relatively simple function of latitude, which the computer can calculate quickly and easily. Because of these properties, the Mercator projection is ideal for oilspill modeling purposes, and is now used by the model whenever possible.

Once the base map has been selected, a Cartesian coordinate system is superimposed with its origin at the lower left-hand (southwest) corner of the map. The longest side of the map is usually assigned a length of 480 units. The whole study area is then divided into a matrix of square cells of one unit each; the maximum size of this matrix is 480X480 cells. For a typical analysis, each cell represents an area of approximately 2 to 4 km’, which is thus the basic unit of resolution for spatial data.

Spatial data is stored in a set of 480x480 matrices. Elements of the matrices define, for every cell:

- Presence or absence of land, and land segments.
- Presence or absence of up to 31 targets,
- Identification of a wind station, for determining the appropriate wind vector for oilspill movement (see subsection-Wind Data).
- Identification of a current polygon, for determining the appropriate current vector for oilspill movement (see subsection—Current Data Checking).

Processing data to construct large arrays is a complicated task requiring a great deal of automation. Likewise, the practical limitations of computers require an efficient, though sometimes complex, storage and paging scheme for handling these matrices. Other sections describe the matrices in “more detail.

LAND AND TARGETS

One primary function of the model is to relate oilspill trajectory movements to the locations of wildlife populations, fishing areas, and other potential “targets” in coastal and continental shelf areas. Environmental impact statements for Federal OCS leasing require collecting an enormous quantity of data about these resources, and a substantial part of this data base becomes input for the model.

STORAGE OF TARGETS

The model stores indicators of the presence or absence of land and up to 31 other targets in each of a quarter million grid cells. This is done in such a way that each of perhaps 150,000 simulated spills are quickly checked at each step in the trajectory for possible impact on each target.

Two features of the model allow a high level of performance in checking cells. When trajectories are being simulated by program SPILL (see section on "Oilspill Movement,”) a paging system burdens computer memory with only a small, easily accessible fraction of the total grid at any time. Additionally, an effective exploitation of IBM storage attributes provides a compact and efficient mechanism for handling data which resides either in main memory or on permanent storage devices.

More technically, each grid cell is assigned one 4-byte integer to indicate the presence of up to 31 categories of targets, and land. Each of the 32 bits (numbered O-31) corresponds to a different target, or land. Bit O, the sign bit, corresponds to land, and is “on” when land is present in the cell. Bit i represents the target number i and the interger value 2**(3l-i); “on” signals that the target is present in the cell. Thus an integer value of, say, 9 (binary 0000000000000000 00000000 00001001) would indicate that targets 28 and 31 are present. Simple subroutines can decode these integers to suit various purposes.

TYPES OF TARGETS

Examples of spill-vulnerable targets which have been included in past analyses appear in tables 1 and 2. Sample targets are shown in figures 1 and 2. A simulated spill registers either “hit” or’ ‘no hit” on a target. A hit is scored as soon as the simulated spill crosses a cell occupied by the target. Multiple crossings by the same spill count as a single hit.

The selection of targets is clearly of critical importance if the model is to produce useful results, The section, “Model Output and Case Examples” further discusses the targets considered in past risk analyses.

FURTHER REFINEMENTS OF TARGETS-SEASONAL VULNERABILITY

Passage of spilled oil through a target location does not necessarily imply an adverse impact on the target, since vulnerability of a single target may vary according to time of year. Many wildlife populations undergo migrations during the year, and seasonal reproductive activities are often more susceptible to damage from spilled oil than other parts of the life cycle, The economic impact of spilled oil on such targets as beaches may also differ seasonally.

TABLE 1.–*Targets for a risk analysis in the Western Gulf of Alaska (from Slack, Smith, and Wyant, 1977)*

Salmon purse seining and set net areas
Pink and chum salmon intertidal spawning areas
Dungeness crab spawning, rearing, and catch areas
Tanner crab fishing areas
Tanner crab mating and hatching areas
Tanner crab vital rearing areas
Tanner crab important rearing areas
King crab mating and hatching areas
King crab vital rearing areas
King crab important rearing areas
Shrimp fishing areas
Shrimp production rearing areas
Seabird colonies
Summer bird distribution (June, July, August)
Fall bird distribution (September, October, November)
Winter bird distribution (December, January, February)
Spring bird distribution (March, April, May)
Marine mammal foraging areas
Sea lion rookeries and hauling grounds
Harbor seal rookeries and hauling grounds
Sea otter concentration areas
Kelp beds
Foreign fishing areas
Archeological sites

The model accounts for seasonal vulnerability by associating with each target a vector specifying "home" or "away" for each month. When a simulated trajectory crosses a cell which the target matrix indicates may be occupied by a target, program SPILL checks to see if the target is home before registering a hit. Figure 2 shows a blue crab migration route in the Gulf of Mexico. A spill crossing this path might be assumed to not affect the crabs at times other than the migratory period. In assessing risk to migrating blue crabs from proposed offshore oil production in this area, hits on migrating crabs were recorded only when simulated spills contacted this path from September through February.

Modeling seasonal vulnerability inevitably requires some degree of professional judgment since assumptions must be made about the longevity of oilspill impacts. For example, an oilspill hitting a beach in May could still affect recreation in June.

LAND SEGMENTS

The model uses a special accounting system for simulated spills which hit land. The land areas near proposed oil production sites can be arbitrarily divided into two independent sets of land segments, with each set containing up to *99* segments. When a simulated oilspill hits a cell containing land, program SPILL checks to see which land segment contains this cell. The number of simulated spills hitting the shore (broken down into time-to-shore categories) are counted and stored by land segment.

TABLE 2.–Targets *for a risk analysis in the Eastern Gulf of Mexico (from Bryant and Slack, 1978)*

Coral areas
Manatee concentrations
Brown pelican rookeries
Wading or pelagic bird rookeries
Dusky seaside sparrow habitat
Marine turtle nesting areas
American alligator habitat
Mangroves or tidal marsh
Estuarine nursery areas
West Florida adult female blue crab migration route
West Florida blue crab larval transport route
Tortugas pink shrimp nursery grounds
Calico scallops
Oysters and bay scallops
Seagrass beds
Spiny lobster
Sandy beaches
Florida Straits
High density use shoreline
National register sites
Designated wildlife, natural, and conservation areas
Designated national wildlife areas
National marine and estuarine sanctuaries
Florida aquatic preserves
Designated shoreline, national. and State parks
Ports
Foreign islands

Figure 3 shows a typical division of the shoreline of an analysis area into 52 land segments. The example comes from a risk analysis for a proposed Eastern Gulf of Mexico offshore oil production area (Wyant and Slack, 1978).

Compact storage of land segment numbers corresponding to each grid cell is achieved by breaking down IBM computer words in the 480X480 array, The word-breakdown method for overall targets was described earlier; the method for land segments differs, but is similar in principle. The computer time required to access land segment information during a trajectory run is much less than that required for targets, as the land segment array need be consulted only when land is hit. Program SEGMATRX inserts the land segment information into the model in the appropriate format.

A few examples will clarify how an analyst might use the land segment feature of the model. If the estimated overall spatial distribution of spills hitting shore is desired, one set of land segments can simply divide the shore into equal-length units; counts of simulated spills hitting each equal-length segment provide the necessary information. If risk analyses are needed for each individual political jurisdiction in the overall analysis area, the second set of land segments could divide the shore into counties or other political units.

A further advantage of land segments is that they allow consideration of risks to targets which may not have been included in the model runs. For example,

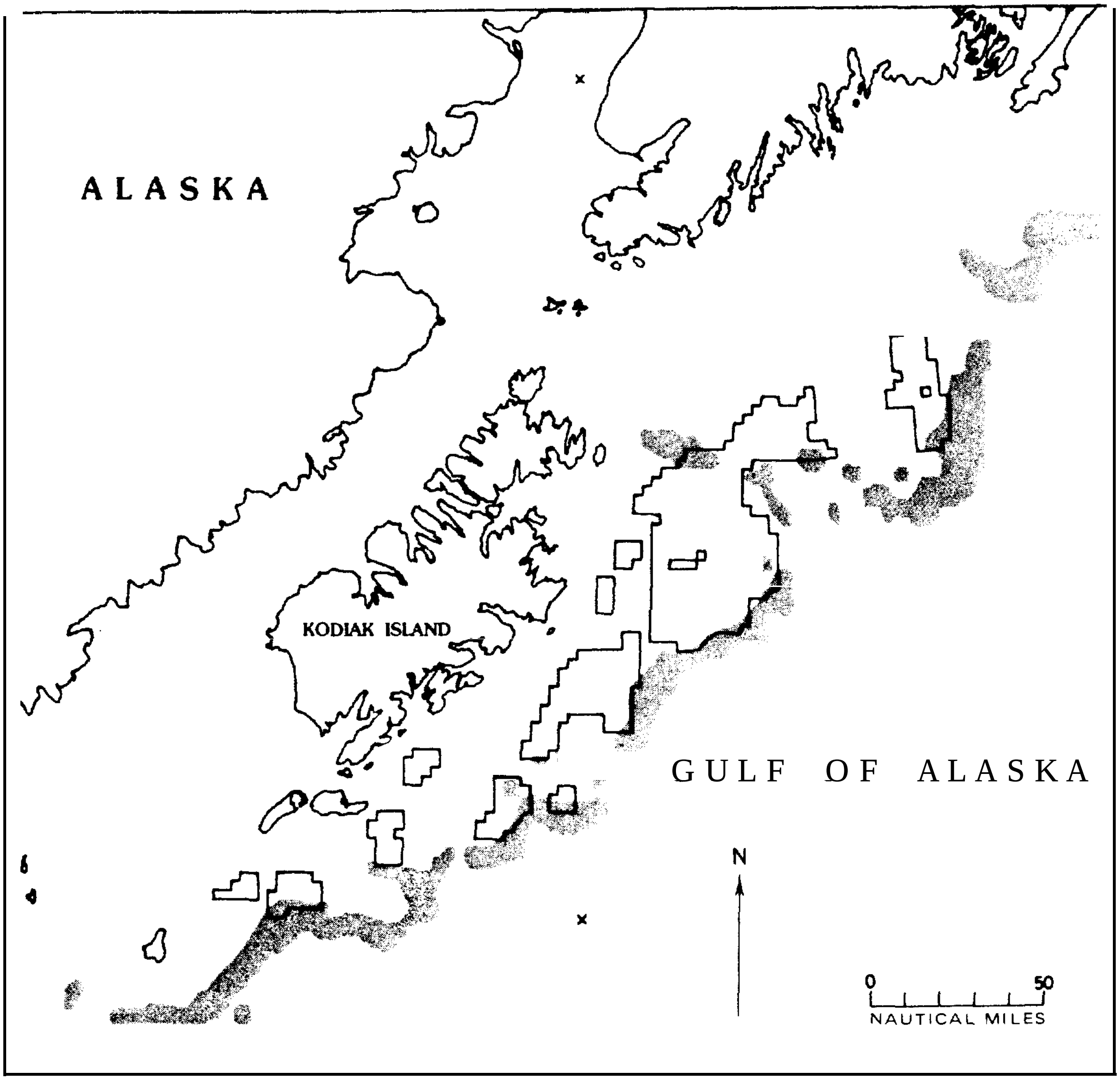

FIGURE 1.–.Map showing a sample target in the Western Gulf of Alaska. Hatched areas indicate foreign fishing areas. Rectangles are proposed lease *tracts (Slack,* et al, 19'77)

suppose that after the model has been run, a shoreline species is added to endangered species lists. Risk to the species can be estimated by examining the land segments in which the species resides.

Finally, the model is not applicable in many bays and estuaries. In a risk analysis for the Mid-Atlantic coast (Slack and Wyant, 1978), simulated spills were not permitted to enter the Chesapeake or Delaware estuaries where the trajectory assumptions of the model are not applicable. To count simulated spills which would have entered the bays, the bay entrances were treated as parts of the shoreline, and a land segment was associated with each bay entrance. Counts of simulated spills hitting these land segments allowed analysis of risk to the bays as a whole without addressing the further problems of spill movements within the bays.

CHECKING TARGETS AND LAND SEGMENT D\TA

The model is designed to allow treatment of extensive and intricate spatial information. In addition to creating computer storage and run-time problems, the size and complexity of the model's basic data structure creates validation problems. Inattention to errors in data input can often lead to disastrously misleading output. Given the time and tedium required for data

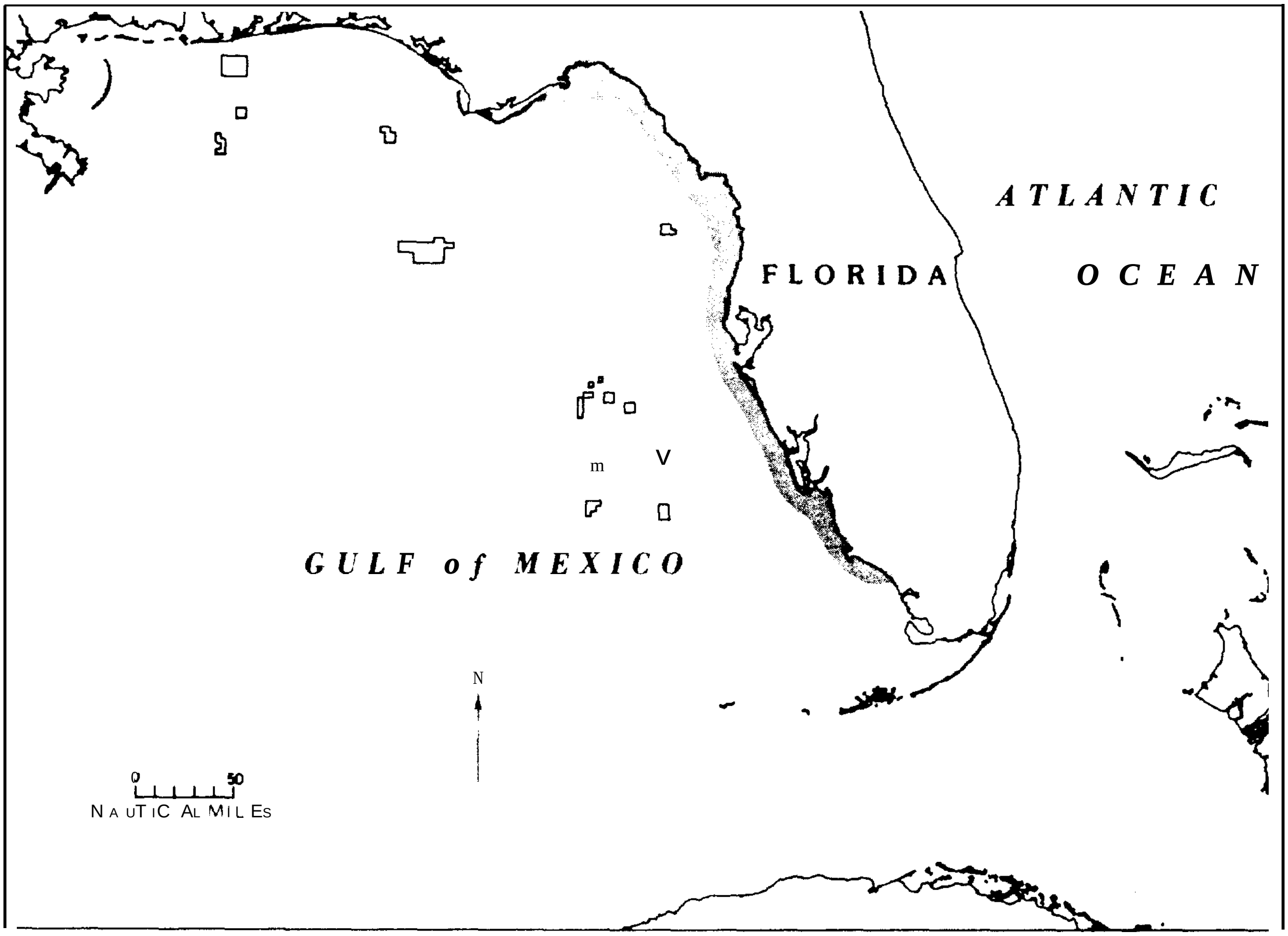

FIGURE 2.–Map showing a sample target in the Eastern Gulf of Mexico. Hatched area indicates blue crab migration route. Rectangles are proposed lease tracts (Wyant and Slack, 1978),

checking and the *greater* intellectual satisfaction of tinkering with the analytical specifications of the model, it is always tempting to pay too little attention to this possibility. The computer programs have been designed to make data checking as complete and convenient as possible, and to prompt modelers to thoroughly carry out this phase of an analysis. OBJECTS and other spatial data entry programs routinely provide diagnostic information such as the number of points in the overall grid system used to represent each target. A coded version of the array used to store the target locations is routinely printed in each run of OBJECTS. The most important checking routines, however, are graphical.

Computer graphics provide a powerful tool for quickly and fully examining complex spatial data. This tool is exploited throughout the data entry phases of a model run. Program DIGIPLOT plots each target as it resides on computer tape immediately after entry from a digitizer. (The target's location at this stage is stored as a string of x-y coordinates representing locations along the boundary of the target area on a map laid on the plane of the digitizer table. See the next subsection, "Insertion of Spatial Data into the Model, " for more detail.) Timely examination of freshly entered spatial data using DIGIPLOT speeds the data entry phase of a model run and prevents costly cascading of errors through subsequent programs.

When program OBJECTS has inserted target locations into the final grid system, program OBJPLOT produces plots such as figures 1 and 2. These plots allow quick appraisal of how faithfully and completely target location in the final coordinate system agrees with the target location on the original map. These plots also provide an immediate check on the correctness of the various map scalings, rotations, and projections required to combine spatial information from different maps and different sessions on the digitizer.

In addition to providing the key to thorough and economical data checking, these computer-graphics programs are an invaluable tool for communicating the content and output of the Survey model. Model results

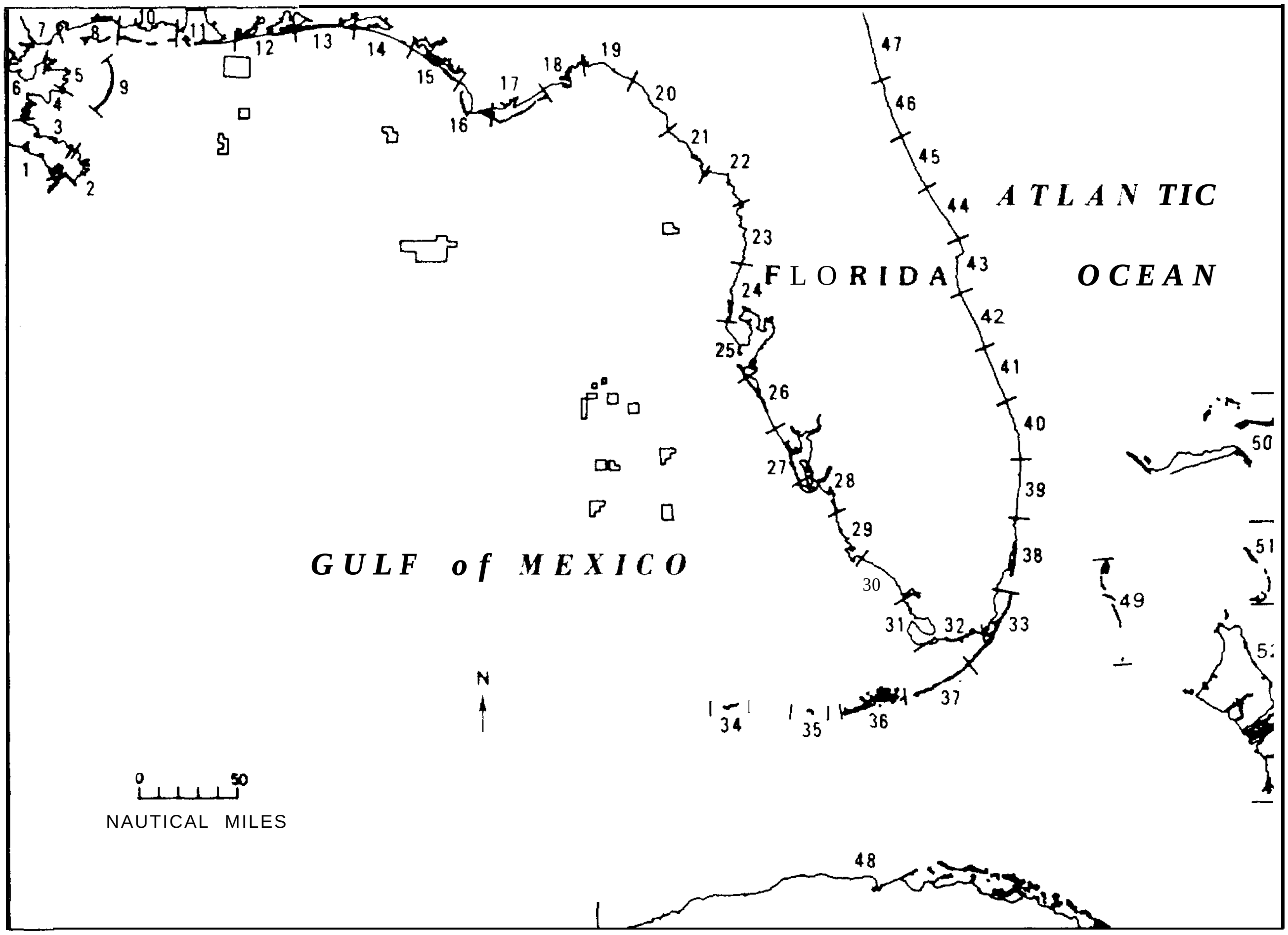

FIGURE 3.–Map showing a typical division of the shoreline into land segments (Wyant and Slack, 1978).

must often be presented to users from a variety of technical backgrounds. Pictures such as figures 1 and 2 are easily understood.

INSERTION OF SPATIAL DATA INTO THE MODEL

This subsection describes the mechanical details of inserting spatial data into the model.

Target location is originally provided on a map of part of the overall analysis area. Each map must have a pair of reference points corresponding to a pair of reference points on the overall map of the area.

The map is laid on a digitizer table, and the outline of a target is traced with the digitizers electrical cross-hairs. This converts the image of the outline to a sequence of points expressed in digitizer table coordinates. The digitizer stores this sequence of coordinates on computer tape.

Program DIGIPRE screens the digitized locations of reference points, targets, and shorelines and stores them on a direct access disk pack in a form accessible to program OBJECTS. Program DIG I PLOT creates diagnostic plots of target locations from these disk files to check against the original maps.

Several options are available for entering spatial data. Correct use of the options speeds the entry process and simplifies data organization and storage. Programs DIG I PLOT and OBJECTS automatically check for large gaps in the point sequences representing target outlines. Thus, the outline of a target with many discrete subareas, such as an island chain, can be traced on the digitizer table and the model will automatically recognize the individual islands. Targets representable as polygons can be entered simply by digitizing the polygon vertices; they need not be traced in their entirety. Some targets can also be entered as isolated points, but this presents some theoretical difficulties, since oilspills are also represented as points.

Land segments are entered much the same as polygonal targets. The order in which the polygon vertices are digitized is important—a specific order is needed to let subsequent programs produce plots like figure 3. Land segments are stored somewhat differently from overall targets and are processed by programs DIGICOPY and SEGMATRX rather than OBJECTS and DIGIPRES.

OBJECTS performs several functions. The sequence of points representing a target outline on the digitizer table is scaled, rotated, and projected into the final coordinate system. The grid cells occupied by these points are noted, and the grid outline of the target is completely connected using subroutines GETLIM, NBR, and TRACK. The grid cells inside the outline are then found using subroutine FILL. Grid locations of the targets, or segments, are compactly stored in arrays using the compaction methods described previously. The arrays are then stored on a direct access disk in such a way that they are accessible to the trajectory program SPILL via subroutine NEWBLK's paging system. They are also accessed by program OBJPLOT, to produce drawings of the target locations, as in figures 1 and *2*. SEGMATRX performs an identical function for land segments.

PAGING SYSTEM FOR LARGE ARRAYS

The 480 x *480* arrays identifying targets and land segments occupy almost 1.4 megabytes of storage. Since only a small portion of each array is needed at any time, a paging system has proven economical in reducing computer core requirements.

Each large array is divided into 30X 30 blocks, which are stored as direct-access records on a disk. The paging system will retain the most recently used blocks in core, and access the others as needed. A further refinement for the array of targets is to construct a 16 x 16 array (one element for each block of the target array) that indicates presence or absence of target categories in each of the blocks. Thus, by checking the smaller array, the computer can determine whether or not it is necessary to read a block of the larger array.

WINDS

The subsections which follow explain how wind data is put into a form that can be used for oilspill simulation. Movement of oil under the influence of wind is covered in "Oilspill Trajectory Simulation."

STOCHASTIC MODEL OF WIND DATA

The variation in the wind is represented as a first-order Markov process. That is, the wind in one time-step is a random function of the wind in the previous time-step. This reflects one's experience that if the wind is presently out of the north at 5 knots, the wind 3 hours or so from now will quite likely be the same, though there is a smaller chance of a large wind shift. A probability transition matrix, constructed from the historic wind record is used to model this Markov process. An example of a wind transition matrix is shown in figure 4, and provides for 41 wind velocity states (8 directions time 5 speed classes, plus the calm state). The elements of this matrix are the probabilities that a particular wind velocity will be succeeded by another wind velocity in the next time step. For example, if the wind is now from the north at 10 knots, row 2 of the matrix shows there is a 22 percent chance that, 3 hours from now, the wind will still be from the north at 10 knots, and that there is a 9 percent chance it will be From the northwest at 5 knots. If the present state of the wind is i, then the next wind state, j, can be randomly chosen by procedures described in the subsection, "Constructing Wind Transition Matrices: Programs RAWWIND, LISTWIND, and WINDTRAN. "

Program W INDTRAN constructs the wind transition matrix from the historic record at a wind station. The resulting matrix is a description of the frequencies of wind velocity transitions that have occurred during the period of record. Probabilities of transitions not occurring in the record are assigned the value of zero.

There is an important difference between sampling winds from a Markov transition matrix constructed in this manner, and simply reading the historic wind record with randomly selected starting days. Although neither technique will model an individual transition which has not occurred in the past, the Markov process model can yield sequences of transitions which have not been observed in the historic record. Since a 30day oilspill trajectory, with winds sampled every 3 hours, will involve a sequence of 240 wind transitions, a far greater number of sequences can be sampled from the transition matrix without repetition than is available from reading the historic record. In effect the difference is that reading solely from the historic record assumes that only wind transition sequences that have occurred in the past can happen in the future, whereas sampling from a transition matrix assumes that the sequence of wind transitions observed in the historic record is only a sampling of some underlying distribution. Considering that usually only 5 to 10 years of historic record is available, and that the oilspill simulation is to represent an exploration and production period of 20 to 25 years, this assumption seems appropriate.

The ideal wind data would be obtained from long-term weather stations located in the area of interest measuring wind velocity at the surface of the ocean. Unfortunately, there are few permanent stations at

NEXT WIND IS OUT OF ..., AT SPEED ... IN PERCENT.

LAST FROM DIR	SPD	CALM	N	NE	E	SE	S	SW	W	NW
			510152025	510152025	510152025	510152025	510152025	510152025	510152025	510152025
CALM	0	27	2 1 0 0 0	3 0 0 0 0	34 3 0 0 0	4 0 0 0 0	1 0 0 0 0	1 0 0 0 0	6 4 1 0 0	7 3 1 0 0
N	5	28	1013 1 0 0	6 1 0 0 0	11 0 0 0 0	2 0 0 0 0	0 0 0 0 0	1 0 0 0 0	0 2 0 0 0	1211 4 0 0
N	10	10	1522 6 0 0	9 1 0 0 0	6 0 0 0 0	0 0 0 0 0	0 0 0 0 0	1 1 0 0 0	0 0 3 0 0	412 7 1 0
N	15	6	011 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	6 0 0 0 0	0 0 6 0 0	111144 6 0
N	20	50	050 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
N	25	0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
NE	5	26	6 1 0 0 0	9 1 0 0 0	16 4 0 0 0	8 1 0 0 0	0 0 0 0 0	1 0 0 0 0	8 4 1 0 0	3 5 6 0 0
NE	10	22	022 0 0 0	0 0 0 0 0	011 0 0 0	011 0 0 0	011 0 0 0	0 0 0 0 0	11 0 0 0 0	11 0 0 0 0
NE	15	0	0** 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
NE	20	0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
NE	25	0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
E	5	11	1 0 0 0 0	1 0 0 0 0	5011 0 0 0	4 1 0 0 0	0 1 0 0 0	1 0 0 0 0	8 5 0 0 0	2 2 0 0 0
E	10	7	0 0 0 0 0	2 0 0 0 0	4027 4 0 0	5 2 1 0 0	1 0 0 0 0	0 0 0 0 0	4 2 0 0 0	2 1 0 0 0
E	15	0	0 0 0 0 0	0 4 0 0 0	42824 0 0	0 41612 4	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	4 0 0 0 0
E	20	0	0 0 0 0 0	0 0 0 0 0	01325 013	0 0132513	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
E	25	0	0 0 0 0 0	0 0 0 0 0	0 0333333	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
SE	5	13	3 0 0 0 0	2 0 0 0 0	46 6 1 0 0	10 4 1 0 0	1 2 0 0 0	1 0 0 0 0	4 1 0 0 0	4 2 1 0 0
SE	10	2	0 0 0 0 0	0 0 0 0 0	9 7 2 0 0	112018 7 2	4 2 4 0 0	2 2 0 2 0	2 0 2 0 0	0 0 0 0 0
SE	15	0	0 0 0 0 0	3 0 0 0 0	10 3 8 8 0	0102618 3	0 3 8 0 0	0 3 0 0 0	0 0 0 0 0	0 0 0 0 0
SE	20	0	0 0 0 0 0	0 0 0 0 0	0 0 8 8 4	0 81923 8	0 0 412 0	0 0 0 4 0	0 0 0 4 0	0 0 0 0 0
SE	25	0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	014292914	014 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
S	5	26	3 0 0 0 0	3 0 0 0 0	24 0 0 0 0	3 0 3 0 0	15 6 0 3 0	9 6 0 0 0	0 0 0 0 0	0 0 0 0 0
S	10	9	0 0 0 0 0	0 0 0 0 0	3 0 0 0 0	9 3 3 0 0	121212 3 0	9 9 3 0 0	6 9 0 0 0	0 0 0 0 0
S	15	0	0 0 0 0 0	0 0 0 0 0	6 0 0 0 0	6 6 0 0 0	61919 0 0	01313 6 0	0 0 6 0 0	0 0 0 0 0
S	20	0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	014 029 0	0 0 029 0	0 0 0 0 0	01414 0 0	0 0 0 0 0
S	25	0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
SW	5	27	0 2 0 0 0	5 0 0 0 0	15 0 0 0 0	6 0 0 0 0	2 5 3 0 0	8 5 0 0 0	8 6 0 0 0	5 3 0 0 0
SW	10	20	3 0 0 0 0	0 0 0 0 0	10 0 0 0 0	0 0 0 0 0	7 7 0 0 0	23 7 0 0 0	313 0 0 0	3 0 3 0 0
SW	15	0	0 0 0 0 0	0 0 0 0 0	13 0 0 0 0	13 0 0 0 0	13 0 0 0 0	02513 0 0	1313 0 0 0	0 0 0 0 0
SW	20	0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	033 0 0 0	0 033 0 0	033 0 0 0	0 0 0 0 0
SW	25	0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
W	5	33	1 0 0 0 0	0 0 0 0 0	9 0 0 0 0	3 0 0 0 0	1 0 0 0 0	2 1 0 0 0	23 7 1 0 0	10 7 2 1 0
W	10	15	0 1 1 0 0	1 0 0 0 0	3 0 0 0 0	1 0 0 0 0	1 0 0 0 0	3 1 1 0 0	2517 2 2 0	1513 3 1 0
W	15	0	0 0 0 0 0	3 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 3 0 0 0	013 8 3 0	11292111 0
W	20	7	0 0 7 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 714 7 7	7 72114 0
W	25	0	050 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 050 0	0 0 0 0 0
NW	5	27	6 1 1 0 0	2 0 0 0 0	10 0 0 0 0	1 0 0 0 0	0 0 0 0 0	1 0 0 0 0	8 4 1 0 0	1914 4 1 0
NW	10	7	4 2 1 0 0	1 0 0 0 0	2 0 0 0 0	0 0 0 0 0	0 0 0 0 0	2 0 0 0 0	7 4 1 0 0	203015 2 0
NW	15	2	5 2 1 0 0	1 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	2 0 2 0 0	11293110 0
NW	20	0	7 7 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 2 3 0	2133826 3
NW	25	0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 025 0	0 0 02550
OVERALL		16	2 2 0 0 0	2 0 0 0 0	24 5 1 0 0	3 1 1 1 0	1 1 0 0 0	1 1 0 0 0	8 5 1 0 0	8 8 5 1 0

* Indicates greater than 99 percent probability

FIGURE 4.—An example of a wind transition matrix. A 3-hour wind transition matrix for winter at Point Arguello, California (See also fig. 5).

sea—a handful of lightships and "Texas towers." Thus, it is necessary to find other sources of data that can reasonably portray winds at sea.

Wind data recorded by ships have the advantage of originating from the area of interest and of (usually) being measured near the sea surface. For determining average wind conditions, these data are quite useful, although they may reflect the tendency of ships to avoid adverse weather. However, because the data are not collected on a consistent, regular basis, they are not suitable for calculating wind transition probabilities.

Permanent weather stations provide long-term, continuous records of winds. It is not difficult to find stations that sample wind velocity every three hours and have been in operation for over five years. Such data are needed for constructing wind-transition matrices. However, permanent weather stations are usually located onshore (often at airports), away from the areas of oilspill interest, and may also be influenced by topographic effects, such as mountains.

The model combines the advantages of both types of data by comparing averaged ships' data, such as wind roses, for different parts of the study area, with the same data for permanent weather stations. In this manner, each part of the study area can be associated with the permanent weather station that most closely matches the ships' data in that region. Although this does not necessarily mean that the wind transition data are exactly the same, it appears to be the best that can be done with the available data.

CONSTRUCTING WIND TRANSITION MATRICES: PROGRAM RAWWIND, LISTWIND, AND WINDTRAN

Data collected by permanent weather stations are available on magnetic tapes in standardized formats. Program RAWWIND reads these tapes, excludes extraneous data, and stores the wind record for each day on a disk file in a compact, unformatted form. Once processed onto the disk file there is no further need for the rather cumbersome weather station tapes.

Wind sampling procedures may differ among weather stations. For example, some collect data at hourly intervals, others at six-hour intervals; some only collect data in the daytime. To ensure that the weather station record is suitable for sampling wind transitions, program LISTWIND provides a compact printout of the wind data. By examining this printout, the analyst can decide upon the appropriate course for further wind data analysis.

Once the wind data for a station are stored on a disk file in suitable form, a wind transition probability matrix is constructed for each season by program WINDTRAN. This program reads the wind record at a specified sampling interval, and classifies the wind into one of 41 wind velocity classes (eight directions times five speeds, plus the calm). It then looks ahead to the next sample to determine how the wind has changed. When data input is completed, WINDTRAN computes a wind transition probability matrix, with elements, $a_{i,j}$ as follows:

$$a_{i,j} = \begin{cases} \dfrac{n_{i,j} \times 10{,}000}{\sum_{k=1}^{41} n_{i,k}}, & \text{if } \sum_{k=1}^{41} n_{i,k} > 0 \\ 0, & \text{if } \sum_{k=1}^{41} n_{i,k} = 0, \end{cases} \tag{1}$$

where

$a_{i,j}$ = probability (times 10,000) that, if the wind is in state i, the next sample will be in state j,

$n_{i,j}$ = observed number of transitions observed from state i to state j.

Thus, if R is a random number between 1 and 10,000, and, the starting state is i, the next state, k, can be found by summing the elements of row i such that:

$$\sum_{j=1}^{k+1} a_{i,j} \geqslant R > \sum_{j=1}^{k} a_{i,j} \tag{2}$$

WINDTRAN must perform several other operations, in addition to constructing the transition matrices. First, it calculates an average speed and direction for all of the observations within each velocity class. This is done because selection of the classes is somewhat arbitrary, and given a finite number of samples, could introduce a bias in the simulated wind record. By using actual averages for each class (rather than only the nominal speed and direction), a simulated wind record should, in the long run, reproduce the averages of the observed winds. Thus the nominal designation of a class as "from the north at 5 knots" may actually mean "from the direction 2 degrees at 5.3 knots." Of course, as the number of observations increases, the two will become more and more alike.

The assumed wind drift angle (usually 20 degrees clockwise, in the northern hemisphere) is added directly to the average direction determined for each category. Then, the average wind vector is divided into x and y components of the coordinate system. Thus, the velocity class is found by a random sampling of the wind transition probability matrix, and the wind vectors for computing oilspill movement are found in a table for the appropriate class.

The final operation of WINDTRAN is performed to enhance the computational speed of later programs. Equation 2, which uses the ordinary transition matrix, is unnecessarily cumbersome for fast calculations: one must try an average of 20 values of k for each solution. Greater speed is attained by sorting and summing across the rows so that, in effect, the most likely transitions are sampled first. An additional matrix is needed as an index to the sorting, but since most wind transition probability matrices are strongly diagonal, the net result allows a much faster search.

The sorted wind transition probability matrices, along with the corresponding indices, and the x and y oilspill movement vectors, are all stored on a disk file.

DEFINING WIND ZONES: PROGRAM WINDZONE

As explained earlier, winds in different parts of the study area may be simulated using the records of different permanent weather stations. Program WIND-

ZONE assigns to each 10X lO block of grid cells a selected wind station number. By reading the wind station number, program SPILL can find the correct wind station to use for any location. Up to six sets of wind transition probability matrices, constructed from the records of six permanent weather stations, are permitted. Figure 5 shows an example of wind zones used for OCS Lease Sale 48; this particular analysis used four weather stations.

CURRENTS

Ocean currents are represented in the model as varying from month to month in a deterministic fashion. This is in contrast to the winds, which vary randomly over a relatively short time period. Spatial variation of currents is incorporated by dividing the study area into as many as 600 subareas, and assigning monthly current vectors to each of these subareas.

The model does not actually model ocean currents but utilizes a current field determined by other means. Input data for currents, whether derived from mathematical models or from direct measurements, must conform to the assumption (made in the preceding sections) that winds and currents are uncorrelated within a given month. Therefore, the current field used for the model is the baroclinic current, and all wind-induced currents must be represented with the wind data.

Tidal currents are also not included in the model. Generally, the waters in which tides are an important transport mechanism are not within the model's intended scope of analysis.

CURRENT DATA ENTRY AND STORAGE: PROGRAMS CURPOLY AND CURMATRX

Current data are made available to the model from maps of the study area showing the current field for each month. The study area is then divided into subareas or polygons, with 600 the maximum number. Each polygon is assigned a current vector for every month. The polygons are, therefore, a finite-element representation of the current field.

The polygon configuration must be able to adequately characterize the overall monthly current fields with the fewest possible polygons, At present, the judgments of the analysts and modelers are the sole determinants of the polygon configuration for each analysis. No mechanical polygon construction routine exists. Figure 6 shows a monthly current field of 518 current vectors used by the model in a run for *a* southern California risk analysis (see Slack, Wyant, and Lanfear. 1978): figure 7 shows che 518 polygons.

The vertices of the current polygons are first digitized in the same manner *as* land segment polygons (see "Land and Targets"). Program DIG I COPY transfers the digitized information into direct access disk datasets. Program CURPOLY combines the information in these datasets into a single current polygon file. Program CURMATRX then carries out the final processing steps to make current data accessible to program SPILL: (1) The polygon vertices are scaled, rotated, and projected from digitizer table coordinates to the final model coordinate system; (2) a 480 *X* 480 2-byte integer array is filled with the current polygon identification numbers corresponding to each grid cell: (3) this array *is* stored on a disk file accessible to program SPILL's paging system; (4) the monthly current velocities associated with each polygon are read from cards and stored in another disk file. CURMATRX also generates diagnostic plots for data checking and a printout of the grid system showing the current polygon associated with each cell.

CURRENT DATA CHECKING

The special requirements of the model and the *occasional* need to obtain current data from several sources necessitate the translation of large amounts of current data to the appropriate format at the beginning of every analysis. Several programs of the model enable quick and thorough graphical checking of the final-format current data to detect translation errors. Graphics are especially effective in detecting major errors. Programs CURPLOT and CURMATRX provide plots of polygon locations, plots of spatial current fields, by month, and diagnostic printouts of the final 480 *X 480* current grid.

OILSPILL TRAJECTORY SIMULATION

MONTE CARLO SIMULATIONS OF OILSPILL TRAJECTORIES

For each selected launch point, a large number of hypothetical oilspills are released at randomly chosen days within the year and are moved about by randomly sampled winds and currents. With sufficient trials, the statistical behavior of the trial spills will approximate the statistical behavior of spills integrated over all possible combinations of release times, winds, and currents.

The model analyzes oilspill trajectories from a set of potential launch points which are chosen to represent different proposed oil production sites in the OCS lease area and proposed transportation routes. A total of 100 launch points may be selected. From each launch point, 500 hypothetical oilspills are simulated for each season of the year, resulting in as many as 200,000 simulated oilspills for a model run, The next section shows how these simulation results are further analyzed to determine risks from various parts of the lease area.

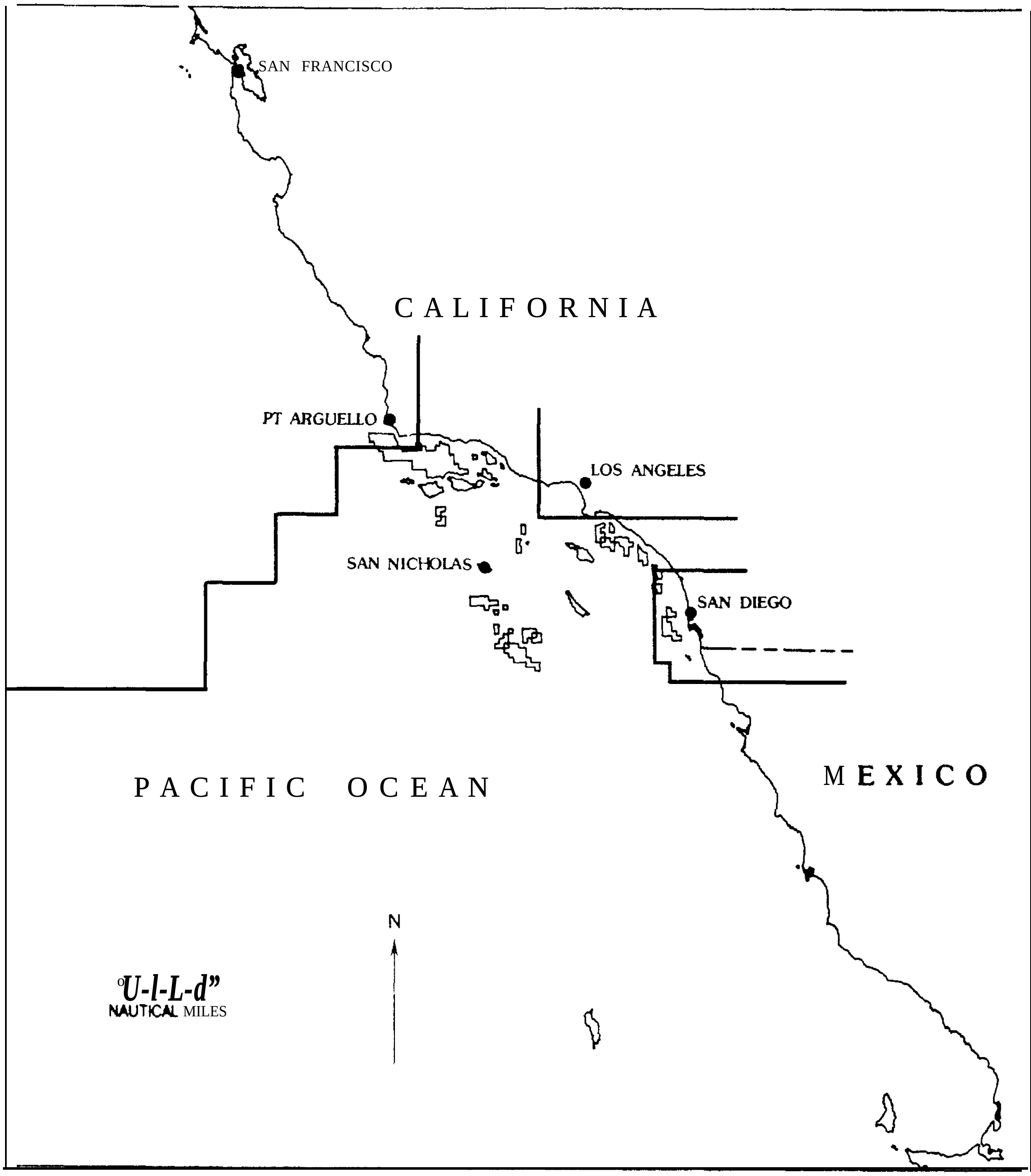

FIGURE 5.—Wind zones used for OCS Lease Sale 48–southern California. Four weather stations were used instead of six. Rectangles are proposed lease tracts (Slack, Wyant, and Lanfear, 1978).

A single launch point may adequately portray a group of proposed lease tracts, but additional points are often needed to represent pipelines and tanker routes. An option of the model allows a launch point to be specified as a straight line, rather than a single point: the 500 spills per season are started from 100 uniformly spaced locations along the line, 5 spills at each location.

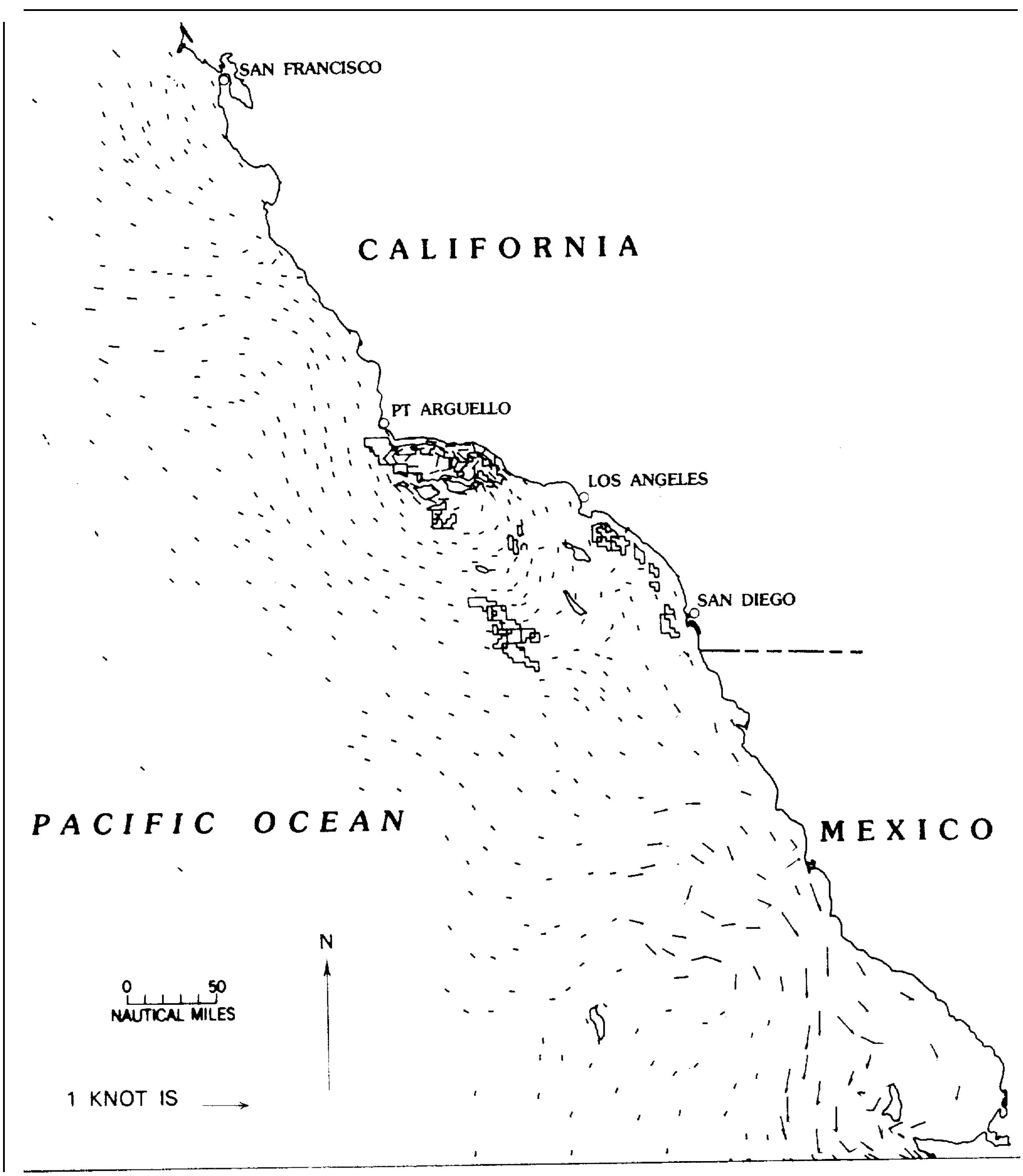

FIGURE 6.– Monthly current field for southern California (March). Rectangles are proposed lease tracts. Lines represent current vectors (Slack, Wyant, and Lanfear, 1978).

EFFECTS OF WIND AND CURRENT ON OILSPILL MOVEMENT

The effects of wind on a parcel of oil flowing on the sea surface have been studied by a number of investigators (Murray and others, 1970; Murty and Khandekar, 1973; Allen and Thanarajah, 1977: Phillips and Groseva, 1977; Stolzenbach and others, 19’77; Zilitinkevich, 1978). There appears to be only partial agreement on the general theory of wind-induced oil-spill movement, probably because of the complexity Of the subject. Winds may transport oilspills through wind-generated currents, wind-induced waves, and by

FIGURE 7.—Current polygons for southern California developed from the monthly current field. Rectangles not containing a vector are proposed lease tracts (Slack, Wyant, and Lanfear, 1978).

direct wind shear; these effects can combine in different ways, depending on characteristics of the oil, sea conditions, and ocean-bottom topography.

Despite the theoretical difficulties, an empirical approach to predicting oilspill movement has proved quite successful in practical trajectory modeling. First described by Smith (1968) in a study of the Torrey Canyon spill off the coast of Great Britain, the method requires the following simple assumptions:

- The effects of wind and current on the oilspill act independently, and can thus be described as a simple vector sum of velocities.
- The wind vector is a constant small fraction of the wind speed, but the direction of oilspill mo-

tion induced by the wind is at a nonzero angle to the direction of the wind due to Coriolis forces.

- The current vector is equal to the current velocity.

Regarding the second assumption, the wind vector has been estimated empirically to equal 3.5 percent of the wind speed with a drift angle of 20 degrees to the right (clockwise) of the wind direction for the northern hemisphere (Smith. 1968; Stolzenbach and others, 1977),

The independence of the effects of wind and current allows the forces to be calculated separately and the resultant motion of the oil to be taken as the vector sum. This requires, of course, that the current field be free of wind effects. Data on currents for past trajectory analyses using the model have come from many sources. Results of drift studies and the output of computer models have been used. Precise assessments of the validity of either drift study results or mathematical model outputs are hard to come by, but something can usually be said about the sensitivity of a set of oilspill risk analysis model results to assumptions regarding currents. More exactly, it is important throughout an analysis to remain aware of whether oilspill movement would be current-dominated or wind-dominated. Often, dominance differs both seasonally and spatially.

Figure 8 contrasts spill movements dominated by each mechanism. The figure shows 10 simulated trajectories launched from a point in a proposed oil production area off the Mid-Atlantic coast (see Slack and Wyant, 1978), For this area as a whole, average wind speed is 12.3 knots (based on lightship data). Assuming winds move oilspills at about 3.5 percent of the winds' own velocities, the winds in this area would, on the average, induce a 0.43 knot speed in spill movements. The currents in the immediate vicinity of the proposed production area are weak—0.1 to 0.3 knot s—and the meanderings in simulated trajectories induced by shifts in the winds can be readily seen in the figure, The Gulf Stream runs to the east of the proposed production area, with currents at 1.0-2.0 knots. As simulated spills leave the lease area to the east, currents dominate over winds in influencing movement. Thus, the simulated spills in the eastern portion of the area move rapidly eastward into the Atlantic Ocean, with little meandering,

OILSPILL MOVEMENT

Program SPILL simulates oilspill movement as a series of displacements over finite time-intervals. For each time step in the duration of a hypothetical spill, two vectors—one representing the effect of the wind and the other that of the current—are summed to obtain the displacement of the spill's center of mass. The spill is then moved in a straight line between its old and new grid coordinates as illustrated in figure 9, and any cells through which the spill passes are checked for the presence of targets. The tracking of a hypothetical spill continues in this manner until a time limit (usually 30 or 60 days) is exceeded, or until the spill contacts land or leaves the area being modeled.

The choice of the time step is based on the sizes of the current polygons, the persistence of the wind data, and practical limits for computer run time. Since a current vector is selected only at the beginning of a time step, a time step short enough to consider the smaller current polygons must be chosen, or they will be skipped over and ignored. Assuming a spill movement speed of 0.5 to 1.0 m/s, a 3-hour time step usually fulfills this condition; where current polygons are larger, a 6-hour time step may be satisfactory. A 3- to 6-hour time step also appears to adequately characterize the wind data in that it makes the model sensitive to the variability in synoptic weather patterns. Finally, a 3-hour time step is a realistic limitation considering the computational speed of program SPILL using existing Geological Survey computer facilities.

Although program SPILL's function–moving a simulated oilspill through cells, checking each cell for targets, and counting hits *on* the targets-is simple in concept, it is a tedious and time-consuming task. A detailed explanation of how program SPILL accomplishes this, using a variety of programming techniques to increase its running speed, is not included in this paper. To understand the probability calculations, however, it is important to know the rules used for recording contacts (hits) of simulated oilspills with the targets. These rules apply to each simulated oil spill:

- The spill may only be designated as hitting or missing each of the 31 target classes: multiple hits on the same target class count as only one hit.
- SPI L'L automat ically determines which months a target is vulnerable to oilspill damage, and counts hits only during these months.
- Upon first contact of the spill with each target class, the simulated age of the spill is recorded.
- If a spill contacts a cell containing land, its simulation is terminated, and the land segment code of that cell is noted: thus the spill may hit no more than one land segment in each set of land segments.
- If the spill moves off of the grid. its simulation is terminated and the direction in which it left the grid is recorded,

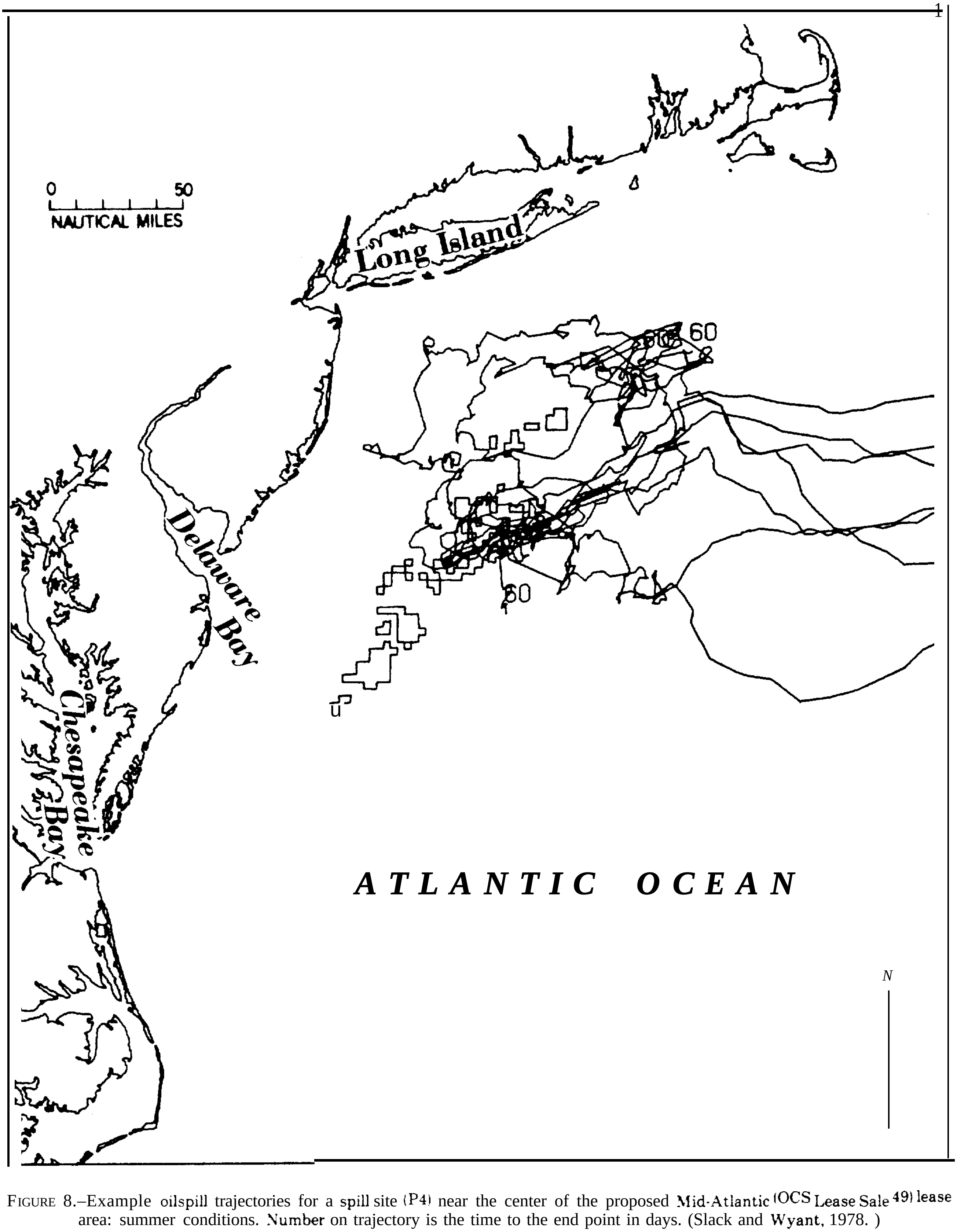

FIGURE 8.–Example oilspill trajectories for a spill site (P4) near the center of the proposed Mid-Atlantic (OCS Lease Sale 49) lease area: summer conditions. Number on trajectory is the time to the end point in days. (Slack and Wyant, 1978.)

- If a spill continues beyond a fixed Lime limit (usually 30 or 60 days), it is assumed to have decayed, and its simulation is terminated.
- The final grid location of the spill is recorded.

Program SPILL produces a record, on a disk file, of the behavior of each hypothetical spill. A summary of SPILL's output is created by program SUMMARY, which shows the results in groups of 100 spills, so that the variability of the Monte Carlo simulation can be checked.

A variation of SPILL (identical to the main program, but containing plotting subroutines) is used to produce graphical displays of trajectory runs. Graphical displays help the analyst ensure that simulated spills behave in a logical manner, and effectively detect errors such as improper scaling factors and reversed wind or current fields. These displays, such as those shown in figures 8 and 10 are also useful as examples of the performance of the model. Conclusions about oilspill behavior from such displays should be cautioned against, since a figure showing 10 spills represents only 0.5 percent of the 2000 spills launched. While average probabilities of hits by oilspills is a meaningful concept, there is no such thing as an "average" trajectory.

A paging system for storing and retrieving the large matrices containing current and land segment data holds down the size of SPILL. Even so, its 500-kilobyte size makes it the second largest of the model's programs. For a large OCS lease sale analysis, SPILL may require more computer operating time than all of the model's other programs combined. Because of its long running time, SPILL is usually run in 5 to 20 independent jobs, so that no one job uses more than one-half hour of central processing unit (CPU) time. The output files of all the jobs are concatenated to form the complete output file. On an IBM 370/155 computer, SPILL can operate at a speed of 1 millisecond per time step, or about ¼ second for a single 30-day oilspill simulation.

RISK CALCULATION

CONDITIONAL PROBABILITIES

Program SPILL records, on a disk file, data about the trajectories of 2,000 hypothetical oilspills from each launch point and the contacts made by these trajectories on targets and land segments. SPILL does not perform any analysis or interpretation of this data; summations and statistical analyses are performed by a subsequent set of programs. These programs determine the likelihood, or conditional probability. that certain events, such as contact with targets or land segments, will occur if an oilspill occurs at a given launch point.

Separating the conditional probability analysis from the Monte Carlo simulation permits the large and time-consuming program SPILL to remain relatively straightforward, while its output can be tailored to user requirements with small, easily modified programs. Two programs, HITPROB and LANDSEG, are used to calculate the conditional probabilities of spills contacting targets and land segments, respectively. A third program, FIRSTPAS, analyzes the travel times oilspills need to reach targets. All three programs operate in a similar manner, scanning the disk output of SPILL to review the results of each trajectory run, and selecting and tabulating the necessary information.

CONDITIONAL PROBABILITIES OF CONTACTING TARGETS: PROGRAM HITPROB

Program H ITPROB calculates the probability that, if an oilspill occurs at a given launch point, it will, within a specified period of time, contact a target. Conditional probabilities are calculated for each launch point for oilspills with maximum ages of 3, 10, 30, and 60 days. Typical output from HITPROB is shown in table 3; this same information is stored on a disk for use by program NU, which calculates overall risks. (See section on "Probability y that an Oilspill will Occur and Contact a Target, " for elaboration of program NU). Since SPILL records a target as "hit" only if contact occurs during a month in which it is vulnerable to oilspill damage, the condition probabilities automatically reflect any seasonal vulnerability.

It is important to recognize that the conditional probabilities calculated by program HITPROB refer to each target as a whole and imply nothing about the distribution of risk among any subdivision of that target. For example, the target "sandy beaches" may extend for several hundred miles of coastline, and risks to particular beaches may differ. Program HITPROB would only calculate the conditional probability that an oilspill originating at a given point would land on a sandy beach somewhere in the study area; it would tell nothing about the likelihood of contacting a specific beach (except that it is less than or equal to that of contatting "sandy beaches"). If such differentation is desired, then each item should be defined as a single target or the land segment feature of the model should be used.

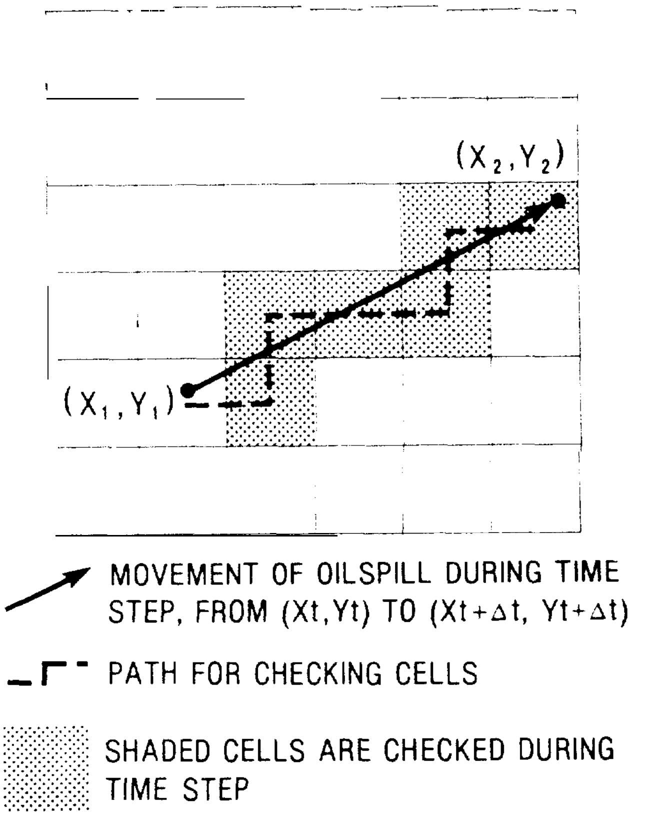

FIGURE 9.–Movement of a hypothetical oilspill through the model's grid system during one time step.

CONDITIONAL PROBABILITIES OF CONTACTING LAND SEGMENTS: PROGRAM LANDSEG

Program LANDSEG calculates the probability that, if an oilspill occurs at a given launch point, it will, within a specified period of time, contact a particular segment of coastline. As in HITPROB, conditional probabilities are calculated for each launch point for oilspills with maximum ages of 3, 10, 30, and 60 days. Each of the two sets of land segments is processed independently, using two slightly different versions of LANDSEG, called LANDSEG 1 and LANDSEG 2. Typical output from LANDSEG is shown in table 4: the same type of information is stored on a disk for use by program NU, which calculates overall risks.

Identification of land segments does not explicity account for spreading of the oil. Although a large oilspill, in reality, could affect several land segments, a "hit" is scored on only one; the user must examine neighboring segments, as well as oilspill travel times, and separately calculate the possible extent of spreading.

TRAVEL TIMES FOR OIL SPILLS CONTACTING TARGETS: PROGRAM FIRSTPAS

Program FIRSTPAS calculates the average, minimum and maximum times-of-travel for oilspills occurring at a given launch point to make first contact with a target. This tabulation, an example of which is shown in table 5, is presented by season as well as for the entire year. Spills which do not contact a target are not included in the statistics for that target.

Program FIRSTPAS was, in earlier versions of the model, the only means of accounting for oilspill age. When HITPROB was revised to present its results for spills of different travel times, FIRSTPAS became partially obsolete. However, it has still proven to be a useful program for checking the behavior of the model and for helping to understand oilspill transport.

SPILL OCCURRENCE

This section describes how spill occurrence probabilities are estimated.

To construct the estimated probability distribution on spill occurrence for a fixed class of spills, certain simplifying assumptions must be used which may be unsatisfactory in particular instances. The forecasting method used in the model is sufficiently flexible for incorporation of new and specific assumptions, however,

The following were considered some desirable features of a spill occurrence forecasting method:

1. The method should include an estimate of the uncertainty in the forecast by providing a probability distribution rather than a predicted number of spills.
2. The method should be consistent with past observations and intuitively reasonable.
3. The dependence on past occurrence rates should be clear and explicit.
4. The method should be flexible; that is, changes in the assumptions concerning use of past occurrence rates should be easily accommodated, and the method should be easy to update as new data are accumulated.

SOME BASIC FEATURES OF SPILL OCCURRENCE FORECASTING

Forecasts of oilspill occurrence are made via a predicted probability distribution on the number of spills which might occur during the production life of a lease area. The predicted distributions are constructed using Bayesian methods to incorporate the uncertainty due to limited historic spill-incidence data. The appendix describes this method in detail.

Simple summary statistics to describe the frequency of spills expected to occur during the production life of a lease area must. be chosen to reflect, as best as possible, the shape of the probability distribution. Considerable uncertainty in forecasting for a new offshore lease area is reflected in a predicted probability distri-

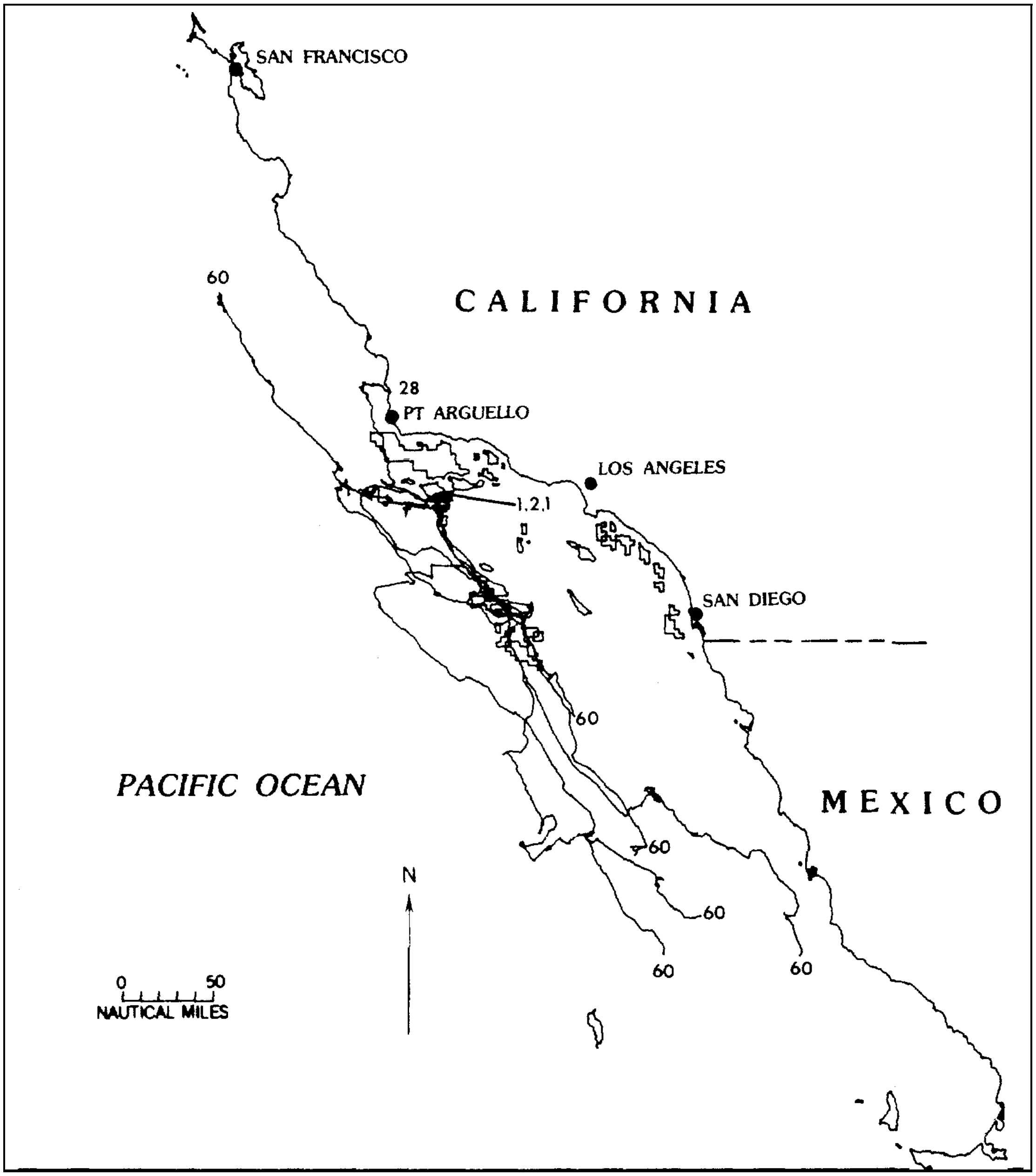

FIGURE 10.–Example oilspill trajectories for a spill site near southern California {OCS Lease Sale 48). Rectangles are proposed lease tracts. Number on trajectory is the time to the end point in days (Slack. Wyant, and Lanfear. 1978).

bution with high variance, implying that one *cannot* forecast a single number of future spills with much confidence. Presenting the "'expected number" of spill can be misleading, as a wide range of possible spill totals may be as likely to occur over the life of the lease area as the "expected number, " which is the hypothetical average over many lease area lifetimes. Thus, model forecasts are presented in terms of the most likely number of spills based on the predicted probability distribution *(in* statistical *terms,* the mode rather than the mean) as well as the predicted probability that one *or* more spills of a given size will occur in the lifetime of a lease area.

Spill occurrence forecasts are made separately for different spill-size categories. Oilspills of different magnitudes have different damage potentials, and may be expected to exhibit different statistical properties in their occurrence. The largest spills occur relatively rarely, but account for a large proportion of the total volume spilled. For example, the Argo Merchant

TABLE 3.–*Example of typical output from program HITPROB, showing probabilities (expressed in percent chance) that an oilspill starting at a particular location will reach a certain target in 30 days*

[n, less than 0.5 percent chance; ●, greater than 995 percent chance]

Target	Hypothetical spill location (launch point) P1	P2 -E2	P3 -E3	P4 -E4	P5	P6	T1	T2	T3	T4	T5	T6	T7
Land	3	2	n	1	1	1	4	11	39	5	17	2	11
1	1	1	n	n	n	n	2	3	6	3	6	1	8
2	n	1	n	1	1	n	2	5	15	5	19	2	6
3	2	1	n	1	1	n	1	4	11	4	12	1	5
4	n	n	n	n	n	n	n	n	n	2	3	n	1
5	n	1	n	n	n	n	2	4	4	n	n	n	2
6	n	1	n	1	n	n	1	2	2	2	6	1	3
7	n	n	n	n	n	n	n	n	n	2	4	n	1
8	n	n	n	n	n	n	n	n	n	n	n	n	n
9	1	1	n	n	n	n	1	1	5	2	5	1	5
10	n	n	n	n	n	n	n	1	1	n	n	n	n
11	n	n	n	n	n	n	1	3	4	n	n	n	n
12	n	n	n	n	n	n	n	n	n	1	4	n	1
13	n	n	n	n	n	n	n	n	n	n	n	n	n
14	n	n	n	n	n	n	n	n	n	n	n	n	n
15	3	2	n	1	1	1	4	12	44	6	19	3	11
16	n	1	n	n	n	n	1	3	9	1	3	1	4
17	n	1	n	n	n	n	n	1	5	2	5	1	1
18	n	n	n	n	n	n	1	3	11	1	5	n	1
19	1	1	n	n	n	n	2	4	4	2	4	1	3
20	n	n	n	n	n	n	n	n	n	n	n	n	2
21	3	n	n	n	n	n	1	n	n	n	n	n	n
22	n	n	n	n	n	n	n	n	n	1	3	n	1
23	n	n	n	n	n	n	n	n	n	n	n	n	n
24	n	n	n	n	n	n	n	n	n	n	n	n	n
25	n	n	n	n	n	n	n	n	n	n	1	n	n
26	n	n	n	n	n	n	n	n	n	n	n	n	n
27	n	n	n	n	n	n	n	n	1	n	n	n	n
28	n	n	n	n	n	n	n	3	*	n	n	n	n

n - less than 0.5 percent chance
* - greater than 99.5 percent chance

spilled 7.7 million gallons when it broke up off Nantucket in December 1976 (Grose and Mattson, eds., 1977, p. 1); by comparison, the total volume spilled in 1975 by U.S. tankers was 7.8 million gallons in 587 incidents (Stewart, 1976, p. 60). The largest spills have the most damage potential, and generally occur under different circumstances from smaller spills. Major blowouts of wells or complete ship breakups, for instance, are somewhat distinct from minor collisions or equipment malfunctions.

Spill occurrence forecasts are made separately for different types of spill sources–tankers, pipeLines, and platforms. It is reasonable to expect that spill occurrence rates will differ for the various modes of production and transport, and past data support this contention (see table 6). A principal use of the risk analysis model has been to help compare transportation mechanisms for given lease areas.

Continued accumulation of data may enable greater refinement of spill-source categorization in the future. For example, tankers might be considered separately by age class (Stewart and Kennedy, 1978, p. 25), or deep-water production rigs with single-buoy moorings might be considered separately from rigid, near-shore structures. The exact approach taken for a given risk analysis should depend on available data, the precise concerns of the analysis, and how the model results are to be interpreted; the model can be straightforwardly applied, using the same methodology, programs, and reporting structures as for the present pipeline-tanker breakdown.

Spill occurrence forecasts, as well as any assessment of risk from a given development program, depend fundamentally on the estimated amount of oil to be produced in a lease area. First described by Devanney and Stewart (1974), the Bayesian methodology used to construct the probability distributions on spill occurrence utilizes past production and oilspill occurrence data and future production estimates in a straight-forward way. The following sections provide further details.

TABLE 4.—*Example of typical output from program LANDSEG, showing probabilities (expressed in percent chance) that an oilspill starting at a particular location will reach a certain land segment in 30 days*

[n, less than 0.5 percent chance]

Land Segment	Hypothetical spill location												
	P1	P2 -E2	P3 -E3	P4 -F4	P5	P6	T1	T2	T3	T4	T5	T6	T7
A1	n	n	n	n	n	n	n	n	n	n	n	n	n
A2	n	n	n	n	n	n	n	n	n	n	n	n	n
A3	n	n	n	n	n	n	n	n	n	1	2	n	1
A4	n	n	n	n	n	n	n	n	n	n	1	n	n
A5	n	n	n	n	n	n	n	n	n	n	n	n	n
A6	n	n	n	n	n	n	n	n	n	n	n	n	n
A7	n	n	n	n	n	n	n	n	n	1	1	n	n
A8	n	n	n	n	n	n	n	n	n	1	2	n	1
A9	n	n	n	n	n	n	n	n	n	1	5	n	1
A10	n	n	n	n	n	n	n	n	n	n	5	n	1
A11	n	n	n	n	n	n	n	n	n	n	1	n	n
A12	n	n	n	n	n	n	n	n	n	n	n	n	3
A13	n	n	n	n	n	n	1	1	1	n	n	1	3
A14	n	n	n	n	n	n	2	6	8	n	n	n	1
A15	n	n	n	n	n	n	n	2	18	n	n	n	n
A16	n	n	n	n	n	n	n	n	1	n	n	n	n
A17	n	n	n	n	n	n	n	n	3	n	n	n	n
A18	n	n	n	n	n	n	n	2	8	n	n	n	n
A19	n	n	n	n	n	n	n	n	n	n	n	n	n
A20	n	n	n	n	n	n	n	n	n	n	n	n	n
A21	n	n	n	n	n	n	n	n	n	n	n	n	n
A22	n	n	n	n	n	n	n	n	n	n	n	n	n
A23	n	n	n	n	n	n	n	n	n	n	n	n	n
A24	n	n	n	n	n	n	n	n	n	n	n	n	n
A25	n	n	n	n	n	n	n	n	n	n	n	n	n
A26	n	n	n	n	n	n	n	n	n	n	n	n	n
A27	n	n	n	n	n	n	n	n	n	n	n	n	n
A28	1	n	n	n	n	n	n	n	n	n	n	n	n
A29	1	n	n	n	n	n	n	n	n	n	n	n	n
A30	n	n	n	n	n	n	n	n	n	n	n	n	n

n - less than 0.5 percent chance

PREDICTED PROBABILITY DISTRIBUTION FOR A FIXED CLASS OF SPILLS

This subsection describes the predicted probability distribution for spill occurrence within a fixed class of spills. A fixed class of spills consists of spills in a single size range (say, spills larger than a thousand barrels) originating from a single spill source (say, tankers).

A basic assumption of the method is that spills occur as a Poisson process, with volume of oil produced or handled as the exposure variable. (Other exposure variables can also be considered, as discussed in the next subsection.) That is, the probability P, of observing n spills in the course of handling t barrels of oil is

$$P(n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}, \qquad (3)$$

where λ is the spill occurrence rate per unit exposure, (spills per million barrels of oil).

The Poisson assumption requires that spills occur independently of each other. One could clearly question this assumption — if, for example, safety and inspection standards were improved as a result of a particular spill, several potentially subsequent spills might be averted. Nonetheless, there is evidence that a Poisson model for spill occurrence provides a reasonable approximation (see Stewart and Kennedy, 1978, p. 36).

The spill occurrence rate, λ, is unknown. A Bayesian methodology, described in detail in appendix A, provides one way to w-eight the different possible values of λ, given the past frequency of spill occurrence for a fixed class of spills by taking a weighted average of the distributions (equation 3) over different values of λ. If ν is the number of past spills in the fixed spill class in the

TABLE 5.—*Example of output from program FIRSTPAS. Average minimum, maximum, and mean time-of-travel for oilspills occurring at site P1 to contact targets. Winter season and all seasons.*

LAUNCH POINT: P1
500 SPILLS PER SEASON

TARGET		WINTER NO.	WINTER MIN	WINTER MAX	WINTER MEAN	ALL NO.	ALL MIN	ALL MAX	ALL MEAN
1	TANNER AND CORTEZ BANKS	64	11	59	33.3	350	6	59	25.5
2	RANGER BANK	3	53	56	54.6	13	35	56	46.9
3	MAJOR MARKETFISH	296	0	60	14.4	1626	0	60	7.0
4	MAJOR COM PELAGIC FISH	97	10	60	32.9	441	5	60	26.5
5	SALMON	0	0	0	0.0	11	7	20	12.7
6	ALBACORE	0	0	0	0.0	372	3	60	27.5
7	BONITO	0	0	0	0.0	11	5	56	35.2
8	TUNA	0	0	0	0.0	78	15	60	41.3
9	SWORDFISH	0	0	0	0.0	18	34	58	45.2
10	COMMERCIAL SHELLFISH	192	1	60	21.8	1189	1	60	11.2
11	SEABIRDS (APRIL-JUNE)	0	0	0	0.0	400	1	19	2.6
12	SEABIRDS (JULY-SEP)	0	0	0	0.0	42	0	58	16.2
13	SEABIRDS (OCT-DEC)	0	0	0	0.0	1	19	19	18.6
14	SEABIRDS (JAN-MAR)	129	5	60	27.1	141	4	60	26.0

course of handling τ barrels of oil, the estimated predicted probability that there will be n spills in the next t barrels handled is

$$P(n)=\frac{(n+\nu-1)!\,t^n\tau^\nu}{n!(\nu-1)!(t+\tau)^{n+\nu}} \tag{4}$$

This is the negative binomial distribution with expectation

$$E=\nu t/\tau \tag{5}$$

and variance

$$V=\frac{\nu t}{\tau}\left(1+\frac{t}{\tau}\right) \tag{6}$$

The probability of one or more spills is

$$P(n\geq 1)=1-\left(\frac{1}{1+t/\tau}\right)^{\nu} \tag{7}$$

Thus, the predicted probability distribution equation on spill occurrence for a fixed class of spills (a single spill-size range, a single spill-source category) incorporates the predicted volume to be handled, t, the past occurrence rate, (ν/τ), and the uncertainty which stems from the fact that $(\nu t/\tau)$ is not likely to equal the true occurrence rate, λ, exactly.

CHOOSING AN EXPOSURE VARIABLE

Fundamental to the spill occurrence forecasting method is the notion of an exposure variable, An exposure variable is some quantity related to oil production or transportation which has a precise statistical relationship to spill occurrence. In the past, the exposure variable used in the model has been volume of oil handled. Predicted probability distributions have been constructed by utilizing past rates of spills per volume of oil handled and the projected volume of oil to be handled.

Other exposure variables could be used. In the case of tankers, for example, number of port calls and numbers of tanker years have been contemplated (Stewart, 1976, p. 53, and Stewart and Kennedy, 1978, p. 23). The model described here permits the use of any exposure variable without major alteration of programs or other parts of the analysis.

An exposure variable should measure some aspect of oil production or oil transport such that for an amount of exposure t the probability of n spills occurring is given by the Poisson distribution:

$$P(n)=\frac{(\lambda t)^n e^{-\lambda t}}{n!}, \tag{3}$$

where λ is the average rate of spill occurrence per unit exposure and t is the exposure. This implies the following technical assumptions:

- The mean and variance of spills for a given amount of exposure should be λt.
- Spills must occur independent y.

In practice, this relationship holds only approximately for any specific exposure variable, and it may be impossible to reject any of several alternative exposure variables simply on the basis of analysis of past data. Further criteria for choosing exposure variables are:

TABLE 6.—*Historic spill occurrence rates used in the Oilspill Risk Analysis Model*

Spill source	Number of spills	Volume handled (millions of bbls)	Data source	Area covered	Time period covered	Model runs used in
			Spills over 1,000 barrels			
Platforms	9	3,927	Devanney and Stewart (1974, p. 89).	U.S.	1964-72	Before May 1977.
Do	10	5,338	Stewart (1975, p. 32) and personal commun. May 1977.	U.S.	1964-75	After May 1977.
Pipelines	8	3,169	Devanney and Stewart (1974, p. 94).	U.S.	1964-72	Before May 1977.
Do	11	4,780	Stewart (1975, p. 33) and personal commun. May 1977.	U.S.	1964-75	After May 1977.
Tankers	99	29,326	Devanney and Stewart (1974, p. 49).	World	1969-72	Before May 1977.
Do	178	45,941	Stewart (1976, p. 66)	World	1969-73	After May 1977.
			Spills under 1,000 barrels			
Platforms and pipelines.	1,230	4,396	Devanney and Stewart (1974, p. 102).	U.S.	1971-72	All.
Tankers	624	1,412	Devanney and Stewart	U.S.	1971-72	All.

- The exposure should be simple.
- It should not intuitively violate the preceding technical assumptions to any significant extent.
- It should be a quantity which is predictable in the future.

The last criterion is particularly important in forecasting applications. If the analyst has an estimate of future production from a large area, but no specific information on how the area is to be developed, in terms of number of platforms, etc., then volume produced might be preferable as an exposure variable over platform-years, even if platform-years appear to be a better exposure variable based on past data.

How can a contemplated exposure variable be checked using past data? One way is by testing the assumption that the mean and the variance are both equal to λt. The linear relationship between the expected number of spills and the exposure variable suggests the use of least-squares regression techniques; weighted least squares should be used because the variance of the number of spills is not constant, and is also linearly related to exposure (see Draper and Smith, 1966, p. 77). Thus, if $(\tau_1, \tau_2, \ldots, \tau_k)$ are the exposures in regions $(r_1, r_2, \ldots, r_n)$ during some year, and $(\nu_1, \nu_2, \ldots, \nu_k)$ are the respective numbers of spills observed, then a regression of $\nu_i/\sqrt{\tau_i}$ vs. $\sqrt{\tau_i}$ checks the first technical assumption. This gives $\Sigma\nu_i/\Sigma\tau_i$ as the true rate of spill occurrence per unit of exposure. The usual tests of a regression fit can be used to evaluate the appropriateness of this assumption. Mean and variance equal to λt is a necessary but not sufficient condition for the Poisson model. (Stewart and Kennedy, 1978, p. 60, present this point quite forcefully). Devanney and Stewart (1974, p. 45) give some examples of regression investigations where volume of oil handled is used as an exposure variable.

Occasionally it will be possible to test directly the Poisson assumption in its entirety. If there are numerous observations, each with the same exposure, then the associated numbers of spills represent independent observations from a single Poisson distribution, and the standard statistical tests for goodness-of-fit can be employed. A possible base is tanker spills, where a contemplated exposure variable is tanker-years. Every tanker which has been in service for the same time period will have the same exposure. Stewart and Kennedy (1978, p. 24) performed goodness-of-fit tests in this situation and felt the Poisson model to be acceptable,

In practice, however, these statistical testing procedures rarely demonstrate unequivocally that a given exposure variable is "correct." They provide one way to rank contemplated exposure variables based on past experience. The ultimate choice of an exposure variable will rest largely on the judgment of the analyst.

The regression work of Devanney and Stewart (1974, p. 26) indicated that volume of oil handled is at least a reasonable exposure variable. The variable is simple, bears a good intuitive connection to the number of spills, and is relatively easy to predict in advance within known error limits. Recently, though, Stewart and Kennedy (1978) have investigated the use of other exposure variables.

SPILL OCCURRENCE RATES AND EXPOSURE VARIABLES

The sources for the spill occurrence rates used in the model are Devanney and Stewart (1974), and Stewart (1975 and 1976). Those authors obtained data primarily from three sources: the Conservation Division of the U.S. Geological Survey, the U.S. Coast Guard, and a survey of world-wide major tanker spills in 1969-1972 (Devanney and Stewart, 1974, p. 1). In the past, there have been many problems in screening and reconciling the information in these data sources; Stewart and Devanney have done much in this area and describe it in the above-cited reports. Table 6 gives the spill occurrence information used to date in runs of the model. The occurrence rates were used to construct predicted probability distributions on spill occurrence as described in the earlier subsection, "Predicted Probability Distribution for a Fixed Class of Spills." For small spills, pipelines and platform occurrences are lumped together due to the data base ambiguity concerning the precise division point between a platform and a pipeline spill.

TRANSPORTATION SCENARIOS

The previous section presented a method for constructing a probability distribution on spill occurrence. The next logical step is to show how site-specific details are applied to calculations of spill occurrence.

COXSTRUCTION' OF TRANSPORTATION SCENARIOS: PROGRAM SCEN.$R1O

The risks of oilspills resulting from OCS development do not arise solely from platform operations. Transporting oil to shore entails additional risks which can exceed the risks of extracting the oil. Therefore, each group of leasing tracts must be considered as part of an integrated production and transportation system; program SCENARIO provides a means of describing this system so that spill occurrence probability distributions can be calculated.

For each production site, a transportation route must be defined by linking together any of the launch points analyzed by program SPILL with destination points (see figure 11), The method of transport (that is, pipeline or tanker) must be specified for each transportation route segment. It is not necessary for the route to be strictly continuous, since this description is only an approximation of an actual route. The modeler must use judgment in striking a balance between a precise route description and a reasonable computational effort, and should be watchful against specifying a transportation route more detailed than justified by the resolution of the model. Figure 11 shows how oil produced from lease tract group P14 would be brought to land in tankers following a route starting at P14 and continuing through T21 and T20. At least one transportation route must be defined for every lease tract group contained in an analysis, and the complete set of transportation routes is called the transportation scenario. The coding of program SCENARIO allows the inclusion of sources of oilspill risk other than OCS leasing in a transportation scenario.

In the preceding subsection, "Spill Occurrence Rates and Exposure Variables,'' it was explained that the exposure variable for transporting oil is the volume of oil handled. That is, a given volume of oil, t, moved from A to B can be expected to result in 1 t spills, regardless of the distance between A and B. The route from A to B can be described as a series of launch points with the oilspill risks distributed among the route segments according to their length. (In figure 11, for example, typical weights may be 20 percent for P14, 40 percent for T19, and 40 percent for T20, demonstrating a rough proportioning of risk to length.) Use of other exposure variables would require a similar weighting of transportation route segments. To accomplish this, program SCENARIO is designed with a highly flexible weighting process that allows the user to assign an arbitrary weight to each segment of a transportation route. This flexibility allows the user to specify a complicated transportation route that involves multiple movements of oil (e.g. "pipeline to port A, then tanker from port A to port B' '), or to divide the oil from a lease tract among several different transportation routes (that is, "half to port A, half to port B"). If deemed justifiable, "high risk" transportation segments can even be assigned higher weights.

ESTIMATED VOLUMES OF OIL RESERVES

For calculating actual oilspill risks, it is necessary to include the volume of oil that is expected to be produced from each lease tract group as input to program SCENARIO. This information is compiled by the Conservation Division of the U.S. Geological Survey and is considered proprietary information.

PROBABILITY THAT AN OILSPILL WILL OCCUR AND CONTACT A TARGET

The model produces as an end result an estimated probability distribution for the number of spills con-

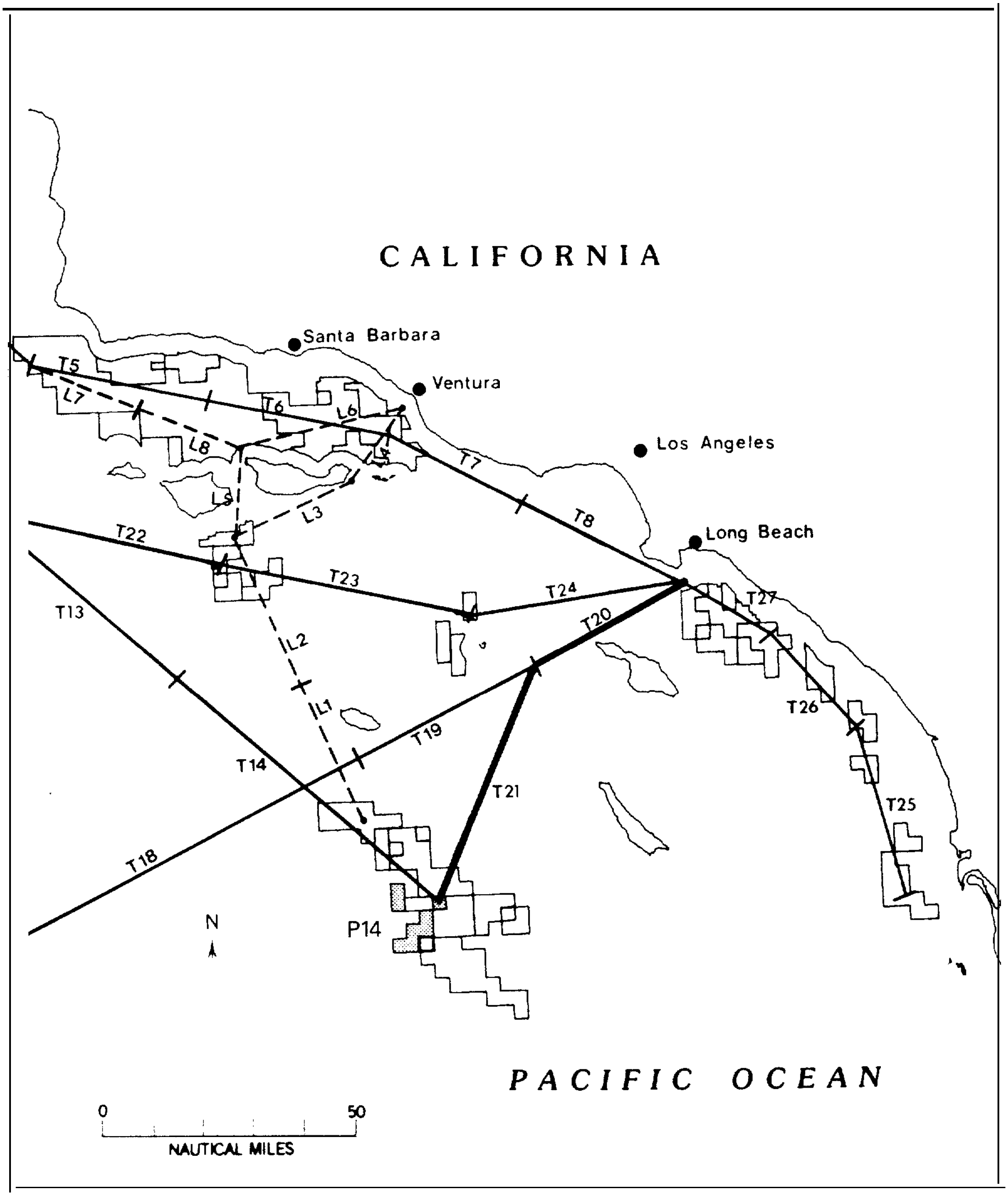

FIGURE 11 –Potential transportation route segments for southern California, showing how oil from tract group P14 would be brought to Long Beach via segments T20 and T 21 [Slack, Wyant, and Lanfear, 1978].

tacting each target or land segment over the production life of a lease area. This final calculation entails three steps which are performed by program NU:

- For each production site or transportation route, the "conditional probability–the probability that a spill, having occurred, will contact the given target or land segment-must be extracted from the output of program HITPROB or program LANDSEG. (The operation of these programs was described in an earlier section.)
- For each production site or transportation route, the conditional probability must be combined with the probability distribution of spill occur-

rence (estimated using the methods described in the section "Spill Occurrence") to yield a single-source probability distribution for the number of hits on the target or land segment. This distribution may be arrived at by one of two methods, according to whether the single source is a "point" source, such as a production platform, or a "distributed" source, such as a tanker route where the oil could be released anywhere along the route.

- All single-source estimated probability distributions in a scenario (see the previous section) must be combined to yield the overall estimated probability distribution for the number of hits on each target and(or) land segment.

The following subsections describe, in more detail, the methods employed,

PROBABILITY OF HITS ON A TARGET FROM A SINGLE SOURCE

Programs HITPROB, LANDSEG, SCENARIO, and NU communicate through files stored on permanently mounted disk packs. After obtaining the conditional probabilities for all targets and land segments for each launch point, program NU begins to process the transportation scenario, segment by segment.

Suppose that for production platforms, the estimated probability distribution of N', the number of spills occurring, is negative binomial. Then, following from the section "Spill Occurrence,"

$$P(N'=n)=\frac{(n+\nu-1)!\,t^n\tau^\nu}{n!\,(\nu-1)!\,(t+\tau)^{n+\nu}} \tag{4}$$

Suppose, further, that the conditional probability (obtained from HITPROB or LANDSEG) that a spill from a point source will hit a given target is *p*. Then, the estimated probability distribution of N', the number of spills which both occur and hit the target over exposure t, is negative binomial

$$P(N'=n)=\frac{(n+\nu-1)!\,(pt)^n\tau^\nu}{n!\,(\nu-1)!\,(pt+\tau)^{n+\nu}} \tag{8}$$

with mean

$$\lambda=\nu pt/\tau \tag{9}$$

and variance

$$\sigma^2=\frac{\nu pt}{\tau}\left(1+\frac{pt}{\tau}\right). \tag{10}$$

Appendix A contains a rigorous derivation of this result.

For a distributed source defined by several transportation route segments, suppose that the estimated probability distribution of spills is the negative binomial with parameters as before: ν, the number of past observed tanker spills, τ, the amount of past exposure observed, and t, the predicted future exposure. These spills could occur at any point along the route. As the previous sections pointed out, it is often desirable in a risk analysis to be able to weight points along a route in terms of that likelihood of spill occurrence. The constraint on the weights is that the distribution of the total number of spills along the route must be the above-mentioned negative binomial distribution, with mean

$$\lambda=\nu t/\tau \tag{11}$$

and variance

$$\sigma^2=\frac{\nu t}{\tau}\left(1+\frac{t}{\tau}\right). \tag{12}$$

This constraint is satisfied by assuming that the distribution of spills at each transportation route segment is negative binomial with parameters ν_i, τ_i, and t_i, where the sum of the ν_i must be ν. Appendix A demonstrates that this structure satisfies the constraint.

To determine the estimated probability distribution for hits on a target from spills along the whole route, the model first constructs the hit distribution for each separate point source along the route (using the previously described methods for single point sources). The model then combines these distributions as described in the next section.

PROBABILITY OF HITS ON A TARGET FROM MULTIPLE SPILL SOURCES

The overall estimated probability distribution for the number of hits on a target is constructed as the convolution of the appropriate single-source distributions derived in previous steps. The meaning of this statement is best conveyed through an example: Let $P_{1,n}$ be the probability of *n* spills hitting the target from the first source, and $P_{2,n}$ be the probability of *n* spills hitting the target from a second source,

$$P_{i,n}=\frac{(n+\nu_i-1)!\,(p_it_i)^n\tau_i^{\nu_i}}{n!\,(\nu_i-1)!\,(p_it_i+\tau_i)^{n+\nu_i}} \tag{13}$$

Let P_n represent the probability that *n* spills hit the target from both sources combined. Then,

$$P_0=P_{1,0}P_{2,0}, \tag{14}$$

$$P_1=P_{1,1}P_{2,0}+P_{1,0}P_{2,1}, \tag{15}$$

$$P_2 = P_{1,2}P_{2,0} + P_{1,1}P_{2,1} + P_{1,0}P_{2,2}, \tag{16}$$

and,

$$P_n = \sum_{i=1}^{n} P_{1,i}P_{2(n-i)} . \tag{17}$$

The extension to more than two sources is similar.

Program NU carries out these calculations in the model. Its design is such that the effects on a risk analysis of different assumptions concerning incidence rates, production and transportation scenarios, or resource estimates, can be determined simply and straightforwardly by rerunning the program with different parameter values.

MODEL VERIFICATION AND LIMITATIONS

FORMAL ERROR ANALYSIS

Results of complex systems models are seldom amenable to formal error analysis, that is, to the expression of error in the end result as a function of errors in input quantities. Often, the error as an input quantity will be unobtainable or unquantifiable, or the error's effect on the overall analysis will be too ambiguous. Furthermore, it is especially difficult to attach a single number representing "standard error" to the results of a model run, when the results consist of a set of predicted probability distributions. In fact, the Bayesian methods used in constructing the distributions described here explicitly incorporate some elements of uncertainty (notably those in estimating spill incidence rate) and were developed, in part, for situations where classical error analysis seemed unsatisfactory.

This does not imply that a useful assessment of model reliability cannot be made, or that the results are categorically unreliable. Three modes of testing are available to the model user: (1) an informal assessment of individual model components is often satisfactory: (2) the sensitivity of the model results to particular assumptions can be tested by repeated runs with differing inputs; and (3) parts of the model can be directly tested by comparison with actual spill trajectories. The following sections contain discussions of how these modes of evaluation were applied to the model described here.

INFORMAL ERROR ANALYSIS

Several factors constrain the effective breadth of a model's applicability, The model's structure (how it works, what it includes or excludes), the refinement of the driving data, and the analytical treatment of the component oilspill occurrence and movement processes all play a role. In a general purpose model, these limitations will differ for each application of the model. This requires that the assumptions necessary for specific areas be readily testable; different factors may limit the precision of the results from case to case. As an example, unreliable current data may not seriously affect model output in situations where the movement of oil is largely wind-dominated (as in some proposed North Atlantic oil production areas–see Smith and others, 1976b), but may critically affect model output where currents are the primary mover (as in many proposed Gulf of Mexico oil production areas–see Wyant and Slack, 1978), The computer programs of the model have been built to facilitate case-specific testing. This has been done by modularizing computer programs, by concentrating on simple parameterizations of processes, and by restricting analytical representations of physical processes to those which are relatively simple, general, and widely accepted.

SPATIAL RESOLUTION

The model cannot represent the locations of oilspills or targets with any finer resolution than the cell size (about 1 nautical mile square) of the grid system. This is an artificial restriction, of course, in the sense that the model could be simply modified to diminish the cell size. Increasing the spatial resolution of the model by this means would, however, lead to a spurious and misleading impression of accuracy in the output, given the present accuracy with which the location of many targets and the spreading of large spills can be depicted.

RISK FOR NEAR-SHORE AND CONFINED-AREA SPILL SOURCES

The spill transport equation used in the model has several virtues. It is simple, is widely accepted as a reasonable representation of oilspill movement in open water (Stolzenbach and others, 1977, p. 5-47), and is void of any special assumptions which would disqualify the model for risk forecasting in most proposed offshore production areas. However, due to the fact that the basic oil transport equation is designed to represent the "average" movement of large spills in fairly open waters, the model cannot adequately represent the detailed movement of spills close to shore or in confined estuaries or bays, where tides and highly localized currents may dominate the movement of spilled oil.

SPREADING

The model does not explicitly incorporate spreading. This deficiency is mitigated by several factors. First, because of the large regions over which the model is designed to operate and the resulting scale of model resolution, spreading is less important than overall advection in determining risks. Second, the original digitization of targets and their insertion into the grid system tends to expand the areas occupied by targets (any cell partially occupied is treated as fully occupied), and causes "near misses" of oilspills to be counted as hits. Third, recording the time-of-contact for each hit enables the analyst to estimate spreading effects independently, given information on oil type, sea state, and so on.

DECAY

The modeling of spill decay presents the same difficulties as the modeling of spreading, in that knowledge of factors such as oil type are integral to analytical descriptions of the physical process. As with spreading, the model does not explicitly calculate decay, but is constructed to provide information on spill travel time, thus enabling assessment of the extent to which decay might mitigate predicted impacts, Contacts of spills with targets are compiled in several elapsed-time categories–up to 3 days, 3 to 10 days, 10 to 30 days, and 30 to 60 days–to assist the analyst in this assessment.

SENSITIVITY ANALYSIS

Sensitivity analysis of model assumptions is a useful technique for assessing model strengths and weaknesses. As suggested above, such analyses should be tailored to particular situations: different features of the model are critical in different situations, and the dictates of economy require the appropriate selection from the many possible sensitivity analyses. Design features of the model make such analyses easy to carry out.

Two sensitivity analyses performed during a risk analysis for proposed North Atlantic OCS production areas (Smith and other, 1976 b), exemplify the kinds of analyses which can be readily performed. The basic transport equation includes the wind drift angle, which is the number of degrees wind-induced oil movements are deflected from the direction of the wind by Coriolis acceleration. Some controversy surrounds the optimal value of this parameter for spill modeling, but most suggested values fall between O and 20 degrees clockwise in the Northern Hemisphere (Stolzenbach and others, 1977, p. 81). For the North Atlantic risk analysis, two separate model runs were made using drift angles of O and 20 degrees clockwise. The resulting estimates of probabilities of spills from the proposed oil production areas hitting shore were 21 and 8 percent, respectively.

In another sensitivity analysis for the same area, actual historic wind sequences were substituted for stochastically generated ones (see the section on "Winds' '). Table 7 shows the estimated probability of hitting land for spills from one proposed North Atlantic oil production site by the two modes of model operation.

These two studies convey the kinds of evaluations which can be conducted in the course of a risk analysis, and the sensitivity of results to certain key assumptions. Different sensitivities to these particular assumptions can be obtained in different OCS areas.

DIRECT MODEL VERIFICATION

Clearly, predictions of expected numbers or probabilities of spill impacts for a given place and time cannot be "proved" or "disproved" by a single spill. Nonetheless, a limited verification of the trajectory model was achieved for the area covered by the North Atlantic OCS oilspill risk analysis (Smith and others, 1976b). In December 1976 the Argo Merchant spilled 7.7 million gallons of oil and the spill traveled in the direction that the model indicated was most likely (see fig. 12). Extensive overflights and monitoring of this spill provialed data for a more thorough evaluation of components of the model. In particular, by comparing actual and simulated spill locations, the validity of current assumptions, transport equations, and wind data source choices could be examined. This work, presented in detail in Grose and Mattson (1977), Pollack and Stolzenbach (1978), and Wyant and Smith (1978), supported the general adequacy of the transport segment of the model.

The spill occurred not as an idealized instantaneous point spill but rather was released over an extended period of time. To facilitate comparisons of simulated trajectories with the actual spill, the spill was modeled as a set of sequentially released points, with each 3-hourly wind applied to the entire, gradually enlarging set. This enabled 2-dimensional construction of spill representations such as that in figure 12. Runs were made using a variety of different parameter values. Graphical output such as that in figure 12 seems to be a particularly appropriate way to communicate the validity of risk forecasts to potential model users in that it quickly and concisely gives a feeling for the model's level of approximation.

TABLE 7.—*Sensitivity of predicted oilspills risks for the North Atlantic study area to the assumption that winds can be modeled as a first-order Markov process.*

Wind sequence	Number of simulated trajectories per season	Percent of simulated trajectories hitting shore				
		Winter	Spring	Summer	Fall	Total
Generated from first order Markov process.	.500	1	8	19	3	8
Taken directly from historic record.	300	0	12	8	0	5

MODEL OUTPUT AND CASE EXAMPLES

REPORTS FOR OCS LEASING

For each application of the model to a Federal OCS lease sale, a final report is produced which includes the following items:

- A discussion of the data sources which were used.
- Maps showing the location of the study area and the locations of the targets and land segments.
- Tables of conditional probabilities giving, for each launch point, the probabilities that an oilspill Occurnng at a given production site will contact targets or land segments within 3, 10, 30, and 60 days.
- Tables and graphs showing the probabilities of oilspills occurring.
- Tables showing the overall probabilities of oilspills occurring and contacting targets or land segments within 3, 10, 30, and 60 days.

A list of reports prepared for seven previous analyses is presented in table 8.

SUMMARY OF RESULTS TO DATE

The model has been used to conduct oilspill risk analyses for eight OCS lease sales in six Federal lease areas, which together represent only a small fraction of the total number of offshore tracts that may be developed eventually. Nevertheless, the six areas studied thus far are distributed among all four of the major

TABLE 8.—*Reports prepared for OCS lease sale* analyses *using the Oilspill Risk Analysis Model of the U.S. Geological Survey*

An Oilspill Risk Analysis for the Southern California (Proposed Sale 48) Outer Continental Shelf Lease Area: James R. Slack. Timothy Wyant, Kennth J. Lanfear; U.S. Geological Survey Water-Resources Investigations 78-80; 1978: 101 p. (Available from NTIS.)

An Oilspill Risk Analysis for the Mid-Atlantic (Proposed Sale 49) Outer Continental Shelf Lease Area; James R, Slack and Timothy Wyant; U.S. Geological Survey Water-Resources Investigation 78-56: 1978; 79 p. (Available from NTIS,)

An Oilspill Risk Analysis for the Eastern Gulf of Mexico (Proposed Sale 65) Outer Continental Shelf Lease Area; Timothy Wyant and James R. Slack. U.S. Geological Survey Open-File Report 78-132; 1978; 72 p.

An Oilspill Risk Analysis for the Western Gulf of Alaska (Kodiak Island) Outer Continental Shelf Lease Area; James R. Slack, Richard A. Smith. and Timothy Wyant; U.S. Geological Survey Open-File Report 77-212; 1977; 57 p.

An Oilspill Risk Analysis for the South Atlantic Outer Continental Shelf Lease Area; James R. Slack and Richard A. Smith; U.S. Geological Survey Open-File Report 76-653; 1976:54 p.

An Oilspill Risk Analysis for the Mid- Atlantic Outer Continental Shelf Lease Area; Richard A. Smith, James R. Slack, and Robert K. Davis: U.S. Geological Survey Open-File Report 76-451: 1976a: 24 p.

An Oilspill Risk Analysis for the North Atlantic Outer Continental Shelf Lease Area; Richard A. Smith, James R. Slack, and Robert K. Davis; U.S. Geological Survey Open-File Report 76-620: 1976b; 25 p.

OCS regions which will experience oil and gas development (the Atlantic and Pacific coasts, the Gulf of Mexice, and the Alaskan Peninsula) and will serve as focal points for further development in those regions.

The primary objective of oilspill risk analyses conducted by the Geological Survey is to determine the risks of petroleum development for the tracts within a given lease area. Such information is useful to the Federal Government in selecting tracts to offer for sale from a list of tracts proposed for development by the oil industry. It is also of interest, however, to make comparisons in oilspill risk between lease areas, since the sites represent the four major OCS regions and the possibility of large differences in risk exists. An interregional comparison will be the emphasis of the summary presented here. (For more detailed descriptions of studies of individual lease areas, the reader is directed to the bibliography of oilspill risk reports in table 8.)

An important question concerning oilspill risk in Federal areas is whether there are significant geographic differences in spill risk per unit of expected oil production. (Risk per unit production can be measured as expected number of spill impacts on a given resource or shoreline segment per billion barrels pro-

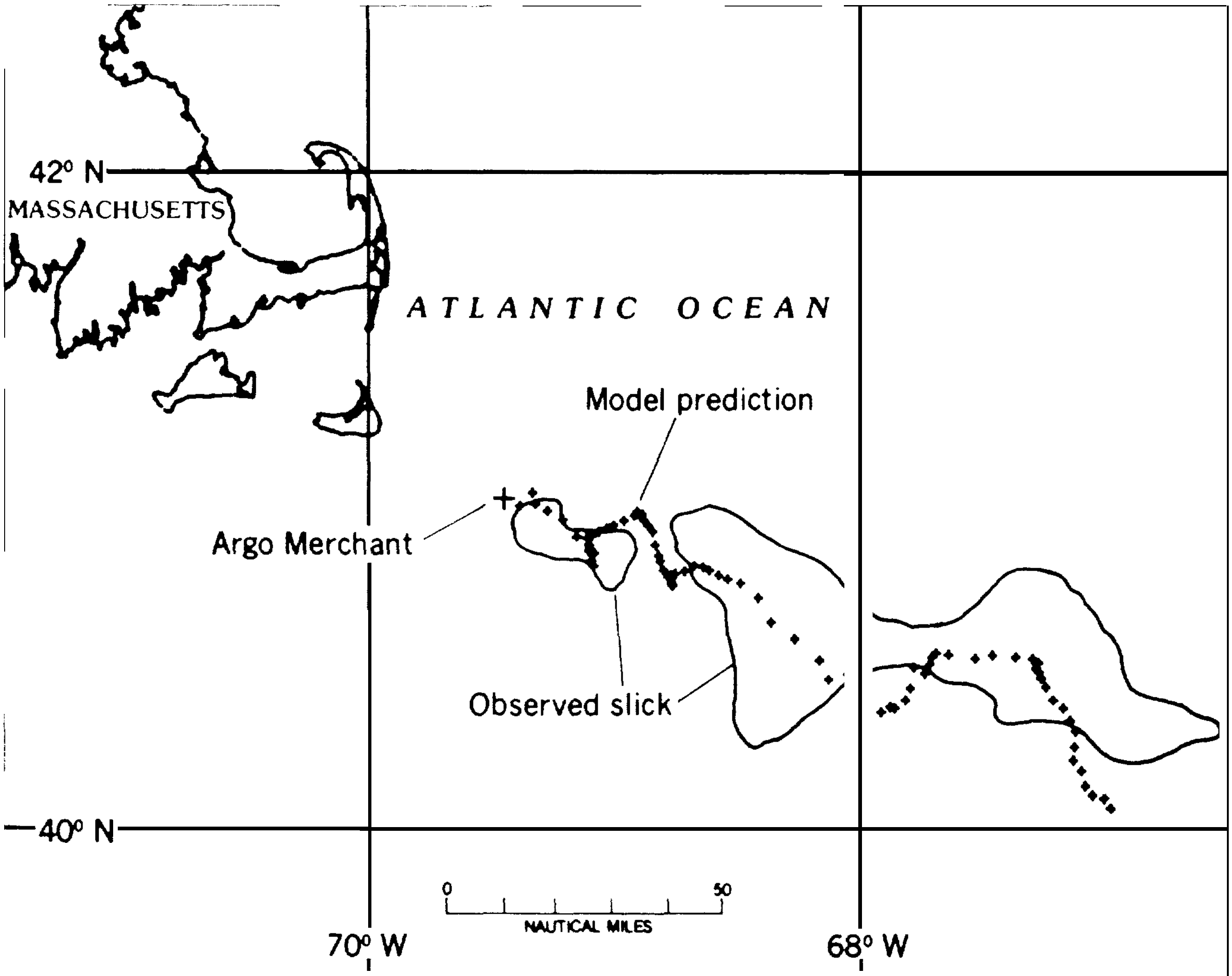

FIGURE 12.–Comparison of the observed slick from the Argo Merchant with a prediction of the Oilspill Risk Analysis Model of the U.S. Geological Survey (Wyant and Smith, 1978).

duced and transported to shore.) Differences in risk per unit production among sites could influence the scheduling of future lease sales. One logical policy, for example, would be to develop the sites bearing the least risk per unit production first, in anticipation of continual improvement in spill prevention and cleanup technology,

Table 9 gives the expected number of oilspills larger than 1,000 barrels occurring and reaching shore during the production life of the six lease areas studied. Where applicable, data for existing and proposed tracts are presented separately. Column 1 summarizes the results of trajectory model runs, and gives the range in conditional probability of spills reaching shore from individual production sites and transportation routes within each lease area, assuming a spill occurs. Column 2 gives the expected number of spill occurrences associated with both production and transportation for the six lease areas. Column 3 gives expected spills reaching shore during the production life of the six lease areas, and represents the sum of the products of conditional probabilities and expected numbers of occurrences of oilspills for individual tract groupings and transportation routes within each lease area. Column 4 gives total estimated oil production for each of the six lease areas. Column 5 gives risk per unit production expressed as expected number of spills reaching shore per billion barrels of oil produced, and is calculated as the quotient of column 3 by column 4.

A value for the average conditional probability of contacting land from spill sites within a given lease area can be obtained by dividing the expected number

TABLE 9.—*Estimated conditional probabilities and expected numbers of spills larger than 1,000 barrels reaching shore as a result of petroleum development in each of the six Federal lease areas*

Lease area	Range of conditional probability of reaching shore from anticipated production areas and transportation routes (percent)	Expected number of spill occurrences	Expected number of spills reaching shore	Estimated oil production (million bbls)	Expected number of spills ashore per billion (bbls. produced)
North Atlantic	2-79	2.4	1.1	500	2.2
Mid-Atlantic					
Existing leases	2-42	3.3	.17	800	0.21
Proposed leases	1-42	0.86	.02	150	0.13
South Atlantic	69-97	3.2	2.5	660	3.8
Eastern Gulf of Mexico					
Existing leases	17-99	7.4	5.1	1800	2.8
Proposed leases	10-94	0.4	.17	70	2.5
Southern California					
Existing leases	1-97	9.3	7.0	1330	5.3
Proposed leases	1-97	5.0	3.8	720	5.3
Gulf of Alaska	28-70	7.5	3.5	1550	2.3

of spills reaching shore (column 3) by the expected number of spills occurring (column 2).[1] In the North Atlantic, for example, the average conditional probability of a spill reaching shore, given that one has occurred on a randomly selected tract, is 46 percent (1.1 - 2.4). It can be seen from table 9, column 1, that even within the same lease area, the probability of oilspills reaching shore from different tracts and transportation routes is quite variable. Ranging from less than 20 percent to nearly 80 percent in a majority of lease areas, the spread in conditional probability reflects variability in wind and current patterns within each area as well as geographic differences, such as the distance of potential spill sites from shore. The variation in risk among different potential drilling sites and transportation routes is, in itself, evidence of the need for an effective methodology for estimating risk prior to tract selection.

More to the point of the present summary, however, are the large differences in oilspill risk between the lease areas, as seen in table 9. By far, the lowest risk of spills reaching shore exists in the Mid-Atlantic area, where the total expected number of spills reaching shore over the production life of both existing and proposed leases is only 0.19, In all other areas, the expectation of spills reaching shore is at least 6 times higher than in the Mid-Atlantic, and for southern California, the expectation is more than 50 times higher. A major reason for low risk values in the Mid-Atlantic area is clear in the results of trajectory model runs for that area (Smith and others, 1976; Slack and Wyant, 1978): the predominance of westerly winds and the great distance of tracts from shore (50 to 100 miles) combine to make the conditional probability of reaching shore comparatively low (1 to 42 percent).

The most significant comparison of oilspill risk among Federal OCS areas is given in the figures for expected contacts with shoreline per unit production (table 9, column 5). It is worth noting that values for expected impacts per unit production are nearly independent of estimated oil production since production estimates appear both in the denominator and numerator of the calculation. Thus any *errors in* predicting oil production are not carried over into this measure of oilspill risk.

In terms of risk per unit production, the greatest contrast is, again, between the Mid-Atlantic lease area, where the expectation of spills occurring and reaching shore is less than one per billion barrels produced, and the southern California lease area, where the expectation is more than five landings per billion barrels produced. Overall, three lease areas stand out as having comparatively high risks of onshore impacts per unit production. These are the southern California, eastern Gulf of Mexico, and South Atlantic areas-each with risk values greater than 2,5 landings per billion barrels produced. The Gulf of Alaska and the North Atlantic together compose a sort of medium risk category, with landing expectations of 2.3 and 2,2 per billion barrels, respectively.

All of the above statistics refer to the risk of oilspills reaching the shoreline within the boundaries of the digital map used to track spill trajectories. Since so many resources vulnerable to spilled oil are located on or

[1] It should be pointed out that averages calculated in this way amount to a weighting of the conditional probabilities of landing from each possible spill site by the relative likelihood of spills occurring at that location.

near the shoreline, probability of contacting land is perhaps the best single descriptor of the risk of oilspill damages in OCS lease areas. However, the probability of contacting land is not always an indicator of the probability of impact on all the resources, and it is advisable to avoid condensing the description of oilspill risk into a single number. For this reason oilspill risk analyses have considered risk to an extensive list of specific resources (typically 20-30) for each lease area.

Table 10 compares OCS Lease areas on the basis of oilspill risk to six general categories of coastal and marine resources. The second part shows the expected number of contacts with each resource category per billion barrels produced. For the most part, the oilspill risk values in table 10 follow the same pattern established in table 9; that is, the lowest impact probabilities appear for the Mid-Atlantic lease area, and the highest appear for the southern California lease area. There are three important instances in table 10, how ever, where oilspill risk is not highest for the southern California area. These are high-density resort and recreation areas (highest for the South Atlantic), critical waterfowl and seabird habitat (highest for the Gulf of Alaska), and marine mammal concentration areas (also highest for the Gulf of Alaska).

OTHER POSSIBLE USES OF THE MODEL

Although the primary purpose of the model is to assess oilspill risks from OCS lease sales, it has several other potential applications. Wyant and Smith (1978) described how the model was used in a "real time" mode to predict movement of oil spilled from the tanker Argo Merchant. A lease sale analysis had ordy recently been completed that included the area of the grounding, and the necessary data files were already in existence. Because subsequent model runs have expanded the model's data base to include major portions of the U.S. Outer Continental Shelf, operation in the real time mode would be possible in many other situations. Conversion to real time operation is relatively simple: data files must be retrieved from tape archives, and program SPILL must be modified so that each Monte Carlo trajectory run begins with a "present" wind velocity. However, it must be emphasized that such use is an extension beyond the original model design, and may not be as efficient nor as technically sound as using models designed specifically for oilspill cleanup.

The model's risk assessment capabilities are not limited to risks of OCS lease sales; other potential sources of oilspills, such as tanker import routes, can be analyzed as well. Since data files must be established for OCS lease sales in any case, the marginal costs of including other oilspill risks are small.

PRACTICAL ASPECTS OF OPERATING AND MANAGING THE MODEL

The model has been used to analyze oilspill risks in eight OCS lease sales, and its continued use is anticipated for future sales. To give potential users a realistic appraisal of the effort involved in model operation, this section discusses the practical aspects of operating the model. The management system which has evolved over three years of modeling operations is described, and the necessary software and hardware support for the model is identified.

MANAGEMENT SYSTEM

The model is constructed as a network of modules, or tasks. Each module is designed to accomplish a single specific objective using, as input, output produced by earlier modules. The major elements of a complete model run are illustrated in figure 13.

Modular construction is not unusual for large models, as it greatly simplifies the modification process. The internal workings of any module may be freely changed, as long as its input and output remain compatible with associated modules.

The network shown in figure 13 can produce an OCS lease sale analysis–from data input to final report–in four months. Initial priorities are to establish a refined set of input data files for program SPILL, and to identify alternative leasing and transportation scenarios. As program SPILL requires a substantial amount of computer time, every effort is made to find and correct errors in the data submitted to SPILL before the latter is executed. Trajectory test runs help to spot data errors and to identify a satis factory set of launch points for potential spills. Program SPILL produces nothing more than a disk file containing the outcome of each Monte Carlo trajectory run, which subsequent programs use to generate conditional probabilities. The latter are combined with leasing and transportation scenarios to determine overall probabilities.

The different stages of model development for a typical sale may produce as many as 50 files. All of these are saved on disk, so that the analysis can be restarted at any intermediate point. Printouts associated with creating these files serve a valuable function in quality control, and help to document the progress of a model run.

SOFTWARE

There are 21 computer programs used in the present version of the model, all written in IBM FORTRAN IV, Level H. An extensive library of subroutines and functions (written in either assembly language or FORTRAN), in addition to the system libraries, is also em-

TABLE 10.—*Expected number of oilspill contacts with coastal and marine resources in six Federal lease areas*

[*, less than 0.01 –, not evaluated]

Lease area	High density resort & recreation areas	Commercial fish. shellfish areas	Wildlife refuges, sanctuaries	Critical waterfowl, seabird habitat	Marine mammal concentrating areas	Critical habitat of rare or endangered species.
	Contact, during production life					
North Atlantic	1.0	1.38	0,43	0.75	0.1	0.17
Mid-Atlantic						
Existing	.02	.08	.12	.15	.01	.10
Proposed	*	.01	.02	.02	*	.01
South Atlantic	1.4	1.45	.08	.08	.03	.03
Eastern Gulf of Mexico						
Existing	.92	4.75	2,64	.56	*	1.14
Proposed	.06	.24	0.10	.02	*	.08
Southern California						
Existing	1.9	7.0	5.7	3.00	2.66	3.91
Proposed	1.1	3.6	3.0	1.45	1.43	1.97
Gulf of Alaska	—	5.5	0.45	6.62	7.5	—
	Contacts per billion barrels of oil produced					
North Atlantic	1.9	2.76	0.86	1.5	0.2	0.35
Mid-Atlantic						
Existing	.03	.10	.15	.19	.06	.13
Proposed	*	.07	.13	.13	*	.07
South Atlantic	2.1	2.18	.12	.12	.05	.05
Eastern Gulf of Mexico						
Existing	.51	2.63	1.46	.31	*	.63
Proposed	.91	3.53	1.47	.29	*	1.23
Southern California						
Existing	1.4	5.3	4.3	2.26	2.01	2.95
Proposed	1.5	5.0	4.2	2.03	2.00	2.75
Gulf of Alaska	—	3.5	.29	4.27	4.8	—

ployed. Proprietary, commercially available subroutine packages are used to control the plotting equipment.

Many of the21 programs involve relatively straightforward processing of digitized raw data. The output from the present digitizing equipment used by the model requires considerable programmer intervention to correct both human and machine errors. In addition, the raw data do not always arrive in a standard format and often need manipulations such as map projection transformations, Therefore, the “front end” programs of the model are usually recompiled, with the necessary modifications, for each individual run. Complete, formal documentation is obviously difficult to achieve under these circumstances and is not expected to be completed until planned improvements in digitizing equipment are accomplished

The remaining programs of the model, including such major programs as WINDTRAN, SPILL, and NU, are stored on a disk in a partitioned data set. Their operation is controlled by cataloged procedures, and extensive checks and interlocks help to ensure correct usage. Although design improvements still result in program changes, careful documentation is maintained for this group of programs.

HARDWARE

The computer hardware requirements for a model this size are substantial. The model is designed to be run on an IBM System/370, Model 155 computer. (The U.S. Geological Survey National Center in Reston, Virginia, has three such computer systems.) Eight hundred kilobytes (800K) of core storage is required for the largest program, although most of the programs need less than 200 kilobytes. All of the model’s files are designed to be stored on a dedicated, on-line 333o disk· storage unit; about 2,000 tracks are required for each OCS lease sale analysis. Tape drives, a plotter, and a digitizer are also required for full operation of the model. When using the model to its full capability, an analysis for an OCS lease sale can require a total of up to 12 hours of CPU time, although no single program is designed to run for more than 30 minutes of CPU time.

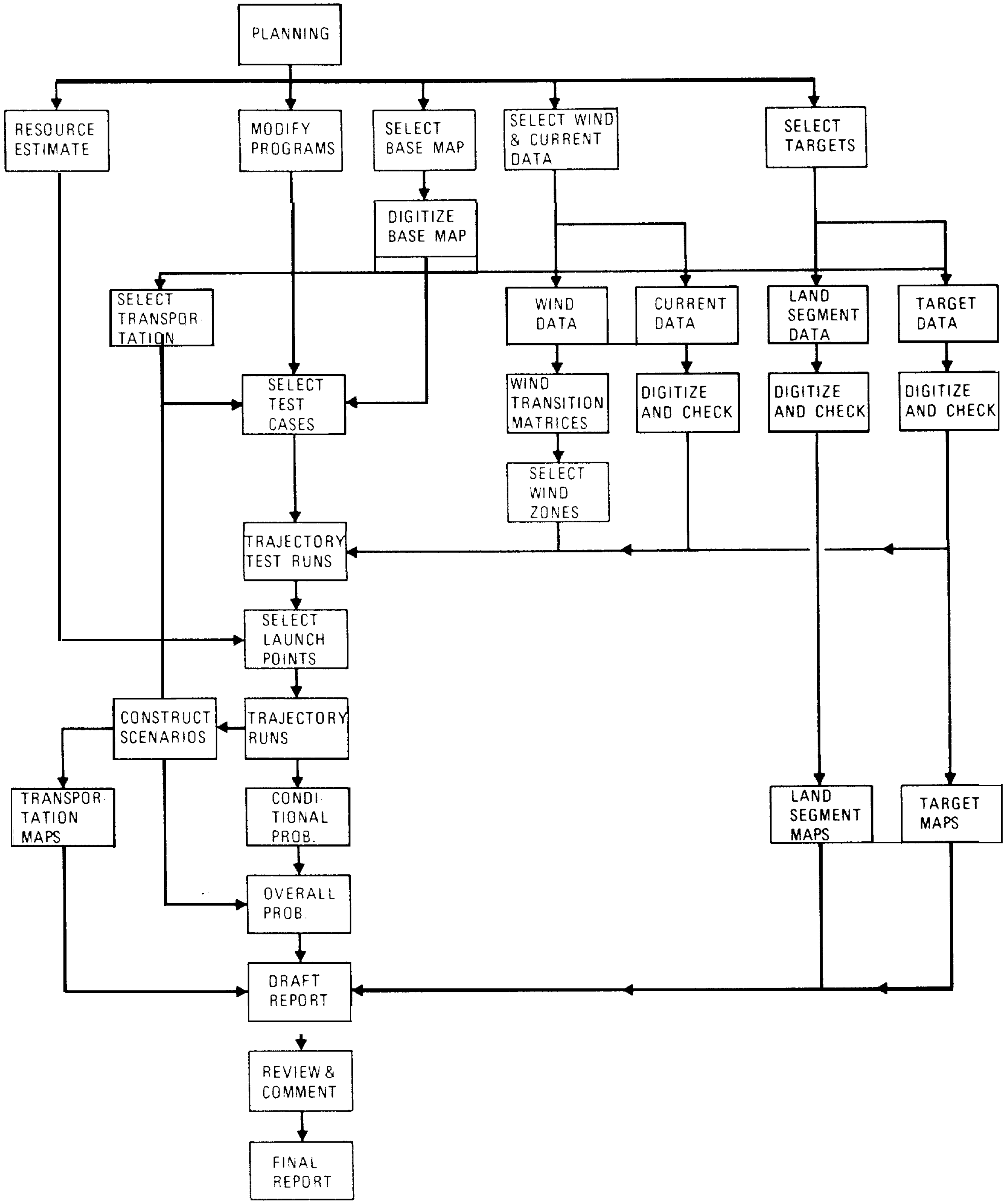

FIGURE 13.– Flow chart illustrating the major elements of a complete model run.

Programs requiring a longer time are broken into several jobs and the output files are concatenated. Some programs, particularly program SPILL, are capable of placing a noticeable strain on computer center operations, and are usually scheduled for execution at off-peak hours.

SELECTED REFERENCES

Allen, J., and Thanarajah, J. C. M., 1977, Laboratory studies of the velocity of windon the movement of oil slicks: *in* Journal of Hydraulic Research, v. 15, no. 4.

Blummer, M., Sonza, S., and Sass, J., 1970, Hydro- carbon pollution of edible shellfish by an oil spill: Marine Biology International Journal on Life in Oceans and Coastal Waters, v. 5, no. 3, p. 195-202.

Box, G. E. P., and Tiao, G. C., 1973, Bayesian inference in statistical analysis: Reading, Massachusetts, Addison-Wesley Publishing Company, 588 p.

Conrad, J. M., 1977, Oil spills: The policy of prevention and the strategy of recovery: Water Resources Research Center, University of Massachusetts, Amherst, Publication No. 93.

Csanady, G. T., 1973, Turbulent diffusion in the environment: Boston, D. Reidel Publishing Co., Geophysics and Astrophysics Monographs, v. 3.

Danenberger, E. P.. 1976, Oil spills, 1971-75, Gulf of Mexico Outer Continental Shelf: U.S. Geological Survey Circular 741, 47 p.

Devanney, J. W., III, and Stewart, R. J., 1974, Analysis of oilspill statistics: Report to Council on Environmental Quality, Washington D. C., 126 p.

Draper, N. R., and Smith, H., 1966, Applied regression analysis: New York, John Wiley, 407 p.

Feller, William, 1966, An introduction to probability theory and its applications: New York, John Wiley, 461 p,

Grose, P. L., and Mattson, J. S. eds., 1977, The Argo Merchant Oilspill; A Preliminary Scientific Report: NOAA

Jeffery, P. G.. 1973, Large-scale experiment on the spreading of oil at sea and its disappearance by natural forces, *in* Proceedings of Conference on Prevention and Control of Oil Spills (1973): p. 469-474.

Kirwan, A. D., Jr., McNally, G., Pazan, S., and Wert, R., 1979, Analysis of surface current response to wind: *in* Journal of Physical Oceanography, American Meteorological Society, v. 9, no. 2.

Murray, S. P., Smith, W. G., and Shaw, C. J., 1970, Oceanographic observations and theoretical analysis of oil slicks during the Chevron spill, March 1970: Coastal Studies Institute, Louisiana State University, Baton Rouge, Louisiana, Technical Report no. 87.

Murty, T. S., and Khandekar, M. L., 1973, Simulation of movement of oil slicks in the Strait of Georgia using simple atmospheric and ocean dynamics: *in* Proceedings of the 1973 Conference on the Prevention and Control of Oil Spills.

Nelson, W. L., 1958, Petroleum refinery engineering, New York, McGraw-Hill.

Offshore Oil Task Group. 1973, The Georges Bank petroleum study, v. 11: Massachusetts Institute of Technology Sea Grant Report, 311 p.

PhiUips, C. R., and Groseva, V. M., 1977, The spreading of crude oil spills across a lake: in Water, Air, and Soil Pollution, v. 8, no. 3, p. 353-360.

Pollack, A. M., and Stolzenbach, K. D., 1978, Crisis science: investigations in response to the Argo Merchant oilspill: Massachusetts Institute of Technology, Cambridge, Massachusetts.

Royer, T. R., 1979, Personal communication to D. Amstutz, U.S. Bureau of Land Management, March 23, 1979.

Slack, J. R., and Smith, R. A., 1976, An oilspill risk analysis for the South Atlantic Outer Continental Shelf lease area: U.S. Geological Survey Open-File Report 76-653, 54 p.

Slack, J. R., Smith, R. A., and Wyant, Timothy, 1977, An oilspill risk analysis for the Western Gulf of Alaska (Kodiak Island} Outer Continental Shelf lease area: U.S. Geological Survey Open-File Report 77-212.

Slack, J. R., and Wyant, Timothy, 1978, An oilspill risk analysis for the Mid-Atlantic (Proposed Sale 49) Outer Continental Shelf lease area: U.S. Geological Survey Open-File Report 78-56.

Slack, J. R. Wyant, Timothy, and Lanfear, K. J., 1978, An oilspill risk analysis for the Southern California (Proposed Sale 48) Outer Continental Shelf lease area: U.S. Geological Survey Open-File Report 78-80,

Smith, J. E., cd., 1968, Torrey Canyon pollution and marine life: Cambridge Univ ersity Press, Cambridge, United Kingdom.

Smith, R. A., Slack, J. R., and Davis, R, K., 1976a, An oilspill risk analysis for the Mid-Atlantic Outer Continental Shelf lease area: U.S. Geological Survey Open-File Report 76-451, 24 p.

Smith, R. .4., Slack, J. R., and Davis, R. K., 1976b, .4n oilspill risk analysis for the North Atlantic Outer Continental Shelf lease area: U.S. Geological Survey Open-File Report 76-620, 50 p.

Stewart, R. J., 1975, Oil spillage associated with the development of offshore petroleum resources, in Report to Organization for Economic Co-operation and Development, 49 p.

Stewart, R. J., 1976, A survey and critical review of U.S. oil spill data resources with application to the tanker pipeline controversy: Report to U.S. Department of the Interior, Washing ton D. C., 69 p.

Stewart, R. J., and Kennedy, M. B., 1978, An analysis of U.S. tanker and offshore petroleum production oil spillage through 1975: Report to Office of Policy Analysis, U.S. Department of the Interior, Contract Number 14-01-0001-2193.

Stolzenbach, K. D., Madsen, O. S., Adams, E. E., Pollack, A. M., and Cooper, C. K., 1977, A review and evaluation of basic techniques for predicting the behavior of surface oil slicks: Ralph M. Parsons Laboratory, Report no. 222,

U.S. Department of the *Interior*, Bureau of Land Management, 1977, Proposed 1977 Outer Continental Shelf oil and gas lease sale in the Western Gulf of Alaska: (OCS Draft Environmental Impact Statement), 4 volumes,

Wardley-Smith, J., cd., 1976, The control of oil pollution on the sea and inland waters, Graham and Trotman, Ltd., United King dom, 251 p.

Wyant, Timothy, and Slack, J. R.. 1978, An oilspill risk analysis for the Eastern Gulf of Mexico (Proposed Sale 65) Outer Continental Shelf lease area: U.S. Geological Survey OpenFife Report 78-132.

Wyant, Timothy, and Smith, R. A., 1978, Risk forecasting for the Argo Merchant spill: *in* Proceedings of a Symposium held January 11-13, 1978, at the Center for Ocean Management Studies. University of Rhode Island. p. 28.

Zilitinkevich, S. S., 1978, An evaluation of the oceanic surface drift current speed and direction: in Boundary Layer Methodology, v. l-i. no. 1.

APPENDIX

DISTRIBUTION THEORY OF SPILL INCIDENCE

1. The derivation of the predicted probability distribution

This appendix describes rigorously the derivation of the predicted probability distribution on spill occurrence given as equation 4, in subsection "Predicted Probability y Distributions for a Fixed Class of Spills. " The development is a Bayesian one; a good general description of these Bayesian inference techniques may be found in Box and Tiao, (1973, p. 1-73). The application of these methods to oilspill occurrence forecasting was proposed and described in Devanney and Stewart, (1974).

We will use the following terminology:

n = number of future spills,
t = future exposure,
λ = true rate of spill occurrence per unit exposure,
u = number of spills observed in past,
τ = past exposure,
$f(n)$ = a marginal probability density on n, and
$f(n|y)$ = the conditional probability density of n given that the random variable y = y.

Assume that spills occur at random with some intensity, $p(n)$:

$$P[n \text{ spills over exposure } t] = \frac{(\lambda t)^n e^{-\lambda t}}{\quad} \tag{A-1}$$

Suppose that, in the absence of information about λ, we choose to represent our uncertainty about this parameter in the form of an "improper" prior density on λ:

$$f(\lambda \mid \text{no data}) = 1/\lambda \tag{A-2}$$

This says, in effect, that with no spills ever having been observed, we place a good deal of faith on λ being equal to O, although we allow *a priori* the possibility that it may be any positive number. This may seem artificial (as is often the case with Bayesian ignorance priors), but note that in any case all it takes is one observation of a spill to refute the notion that λ = O. Our previous feelings in the absence of any data will be overwhelmed by minimal experimental evidence.

Suppose we then observe v spills in τ exposure and wish to update our estimate of λ. The Bayesian approach is to represent our new estimate by a posterior density on λ derived from our ignorance prior density on λ combined with experimental evidence. This is accomplished through use of Bayes theorem:

$$f(\lambda|\nu,\tau) = \frac{f(\nu|\lambda,\tau)f(\lambda|\text{no data})}{f(\nu,\tau)}$$

$$= \frac{f(\nu|\lambda,\tau)f(\lambda|\text{no data})}{\int_0^\infty f(\nu|\lambda,\tau)f(\lambda|\text{no data})d\lambda}$$

$$= \frac{\left[\frac{(\lambda\tau)^\nu e^{-\lambda\tau}}{\nu!}\right]\frac{1}{\lambda}}{\int_0^\infty f(\nu|\lambda,\tau)f(\lambda|\text{no data})d\lambda}$$

$$= \frac{\left[\frac{(\lambda\tau)^\nu e^{-\lambda\tau}}{\nu!}\right]\frac{1}{\lambda}}{\int_0^\infty \frac{(\lambda\tau)^r e^{-\lambda\tau}}{\nu!}\frac{1}{\lambda}d\lambda}$$

$$= \frac{(\lambda\tau)^\nu e^{-\lambda\tau}\frac{1}{\lambda}}{\int_0^\infty x^{\nu-1}e^{-x}dx}$$

$$= \frac{(\lambda\tau)^{\nu-1}e^{-\lambda\tau}\tau}{(\nu-1)!}$$

This is the density on λ in Devanney and Stewart (1974, p. 28) "through which our past spill experience enters the analysis. " It is, in Bayesian terms, the posterior density on λ.

If we were to gather more evidence, this posterior would now become the prior, and the same reasoning would apply:

$$f(\lambda|\nu_2,\tau_2) = \frac{f(\nu_2|\lambda,\tau_2)\times f(\lambda|\nu_1,\tau_1)}{\int_0^\infty f(\nu_2|\lambda,\tau_2)f(\lambda|\nu_1,\tau_1)d\lambda} \tag{A-4}$$

$$= \frac{e^{-\lambda(\tau_1+\tau_2)}(\lambda(\tau_1+\tau_2))^{\nu_1+\nu_2-1}(\tau_1+\tau_2)}{(\nu_1+\nu_2-1)!}$$

Note that this is exactly the same density on A we would have obtained by adding the two exposures, τ_1, *and* τ_2, and the two numbers of spills, ν_1 and ν_2, and treating it all as one piece of data.

Having done all this, if we desire the density of the phenomenon (oilspill occurrence) given our current uncertainty about A, we take the average of the Poisson densities weighted according to the posterior on A:

$$f(n|t,\nu,\tau) = \int_0^\infty f(n|\lambda,t)f(\lambda|\nu,\tau)d\lambda$$

$$= \int_0^\infty \frac{(\lambda t)^n e^{-\lambda t}}{n!}\frac{[\lambda(\tau_1+\tau_2)]^{\nu_1+\nu_2-1}(\tau_1+\tau_2)}{(\nu_1+\nu_2-1)!}$$

$$= \frac{(n+\nu-1)!\,t^n\tau^\nu}{n!(\nu-1)!(t+\tau)^{n+\nu}} \tag{A-5}$$

This is the negative binomial distribution given as equation 4, in the subsection "Predicted Probability Distributions for a Fixed Class of Spills."

2. Moment-generating functions

Results in the remainder of this appendix depend on the use of generating functions. Some standard results from probability theory will be reviewed,

If X is a discrete random variable with $P(X = n) = p_n$, the generating function of X (Feller, 1957, p. 249) is

$$\Phi_X(s) = \sum_{n=0}^{\infty} p_n s^n \tag{A-6}$$

Moment generating functions for some common distributions used in this analysis are as follows:

Bernoulli random variable with probability p of "success":

$$\Phi_X(s) = 1 - p + ps \tag{A-7}$$

Poisson random variable with mean λt:

$$\Phi_Y(s) = \exp(\lambda t(s-1))$$

Negative binomial random variable with mean $\nu t/\tau$ and variance $\nu t/\tau(1+t/\tau)$

$$\Phi_N(s) = \left(\frac{\tau}{t+\tau-ts}\right)^{\nu} \tag{A-8}$$

If X_k is a sequence of random variables with $P(X_k = n) = p_{kn}$, and X is a random variable such that $P(X = n) = p_n$, in order that $p_{kn} \to p_n$ for any fixed n, it is necessary and sufficient that

$$\Phi_{X_k}(s) \to \Phi_X(s) \tag{A-9}$$

for all s in [0, 1] (Feller, 1957, p. 262).

If Z = X + Y, and X and Y are independent, then

$$\Phi_Z(s) = \Phi_X(s)\Phi_Y(s) \tag{A-10}$$

(Feller, 1957, p. 250). If Xi, $i = 1, 2, 3, \ldots$ are independent and identically distributed,

$Z = \sum_{i=1}^{N} X_i$, and N is independent of the X_i, then

$$\Phi_Z(s) = \Phi_N(\Phi_X(s)) \tag{A-11}$$

(Feller, 1957, p. 268).

3. Convergence of the negative binomial to the Poisson

Let N be the number of spills in an exposure t, and assume (following the first-part of this appendix) that N is a Poisson random variable with generating function

$$\Phi_N(s) = \exp(\lambda t(s-1)), \tag{A-12}$$

and that the predicted number of spills N' is a negative binomial random variable with generating function

$$\Phi_{N'}(s) = \left(\frac{\tau}{t+\tau-ts}\right)^{\nu}, \tag{A-13}$$

where ν is the number of spills observed in the past in the course of exposure τ, If the Poisson model holds, then the Law of Large Numbers guarantees that as $\tau \to \infty$ then $\nu/\tau \to \lambda$. Suppose we had adopted the negative binomial model. Then

$$\Phi_{N'}(s) = \left(\frac{\tau}{t+\tau-ts}\right)^{\nu}$$
$$= \left(\frac{\nu/\lambda}{t+\nu/\lambda-ts}\right)^{\nu} \tag{A-14}$$

and, as τ grows larger,

$$\Phi_{N'}(s) = \left(\frac{1}{\lambda t/\nu + 1 - \frac{\lambda t}{\nu}s}\right)^{\nu}$$
$$= \left(1 + \frac{\lambda t}{\nu}(1-s)\right)^{-\nu} \tag{A-15}$$

which approaches

$$\Phi_{N'}(s) = \exp\left(\lambda t(s-1)\right) \tag{A-16}$$

as τ (and hence ν) grows larger.

Thus, if the Poisson model is correct, the analyst will be led to the Poisson model as enough data accumulates even while formally adopting the negative binomial model. Spill incidence could be modeled quite simply and directly using the Poisson distribution with λ set equal to v/T, This convergence to the true model is an example of "Bayesian consistency." The advantage of the negative binomial model, as derived through the Bayesian methodology of this appendix, is that it incorporates the uncertainty about λ for a finite exposure τ, since ν/t will never equal λ exactly. The uncertainty is reflected in a broader distribution on spill incidence due to the larger variance of the negative binomial distribution—a wider range of spill incident totals has non-neglible probability. The variance of N' is $\frac{\nu t}{\tau}\left(1+\frac{t}{\tau}\right)$, the variance of N is $\frac{\nu t}{\tau}$, and the difference is $\nu\left(\frac{t}{\tau}\right)^2$. Thus the increase in uncertainty (as measured by the difference in variances) is proportional to the squared ratio of estimated future exposure, t, to observed past exposure, τ.

This is only one measure of the closeness of the two models, of course. Of more interest is a direct comparison of the summary features presented in the Oilspill Risk Analysis Model of the U.S. Geological Survey, particularly in calculating the probability that no spills will occur. The expectations of N and N' are the same under the two models, $\nu t/\tau$. Let

$$P_p = P(0 \text{ spills [Poisson model]}) = e^{-\nu t/\tau} \tag{A-17}$$

$$P_{nb} = P(0 \text{ spills/Negative binomial model}) =$$

$$\left(\frac{1}{1+t/\tau}\right)^{\nu}. \tag{A-18}$$

Consequently, dividing the two equations and applying a Taylor's expansion yields

$$\ln\frac{P_{nb}}{P_p} = -\left(-\frac{\nu t}{\tau}\right) + \left(-\frac{\nu t}{\tau} + \tfrac{1}{2}\nu\Theta^2\right)$$

$$= +\tfrac{1}{2}\nu\Theta^2, \tag{A-19}$$

where

$$0 \leqslant \Theta \leqslant (t/\tau)^2 \tag{A-20}$$

or

$$1 \leqslant \frac{P_{nb}}{P_p} \leqslant e^{\frac{1}{2}\nu (t/\tau)^2} \tag{A-21}$$

Thus the difference in the probability of no spills occurring under two models (and hence the difference in the probabilities of one or more spills) is again directly related to the size of $(t/\tau)^2$.

4. Distribution of the total number of spills from multiple sources

Let N_1 and N_2 be negative binomial distributed random variables with generating functions

$$\Phi_{N_1}(s) = \left(\frac{\tau_1}{t_1+\tau_1-t_1 s}\right)^{\nu_1}, \tag{A-22}$$

$$\Phi_{N_2}(s) = \left(\frac{\tau_2}{t_2+\tau_2-t_2 s}\right)^{\nu_2}. \tag{A-23}$$

Then, if $N = N_1 + N_2$,

$$\Phi_N(s) = \Phi_{N_1}(s)\Phi_{N_2}(s), \tag{A-24}$$

following equation A-10. In general, this will not be a simple distribution. However, if $t_1 = t_2 = t$ and $\tau_1 = \tau_2 = \tau$, then

$$\Phi_N(s) = \left(\frac{\tau}{t+\tau-ts}\right)^{\nu_1+\nu_2}, \tag{A-25}$$

so N is distributed as a negative binomial random variable with mean

$$\lambda = (\nu_1+\nu_2)\frac{t}{\tau} \tag{A-26}$$

and variance

$$\sigma^2 = (\nu_1+\nu_2)\frac{t}{\tau}\left(1+\frac{t}{\tau}\right). \tag{A-27}$$

5. Distribution of the number **of hits**

Let N, the total number of spills, be distributed as above, that is, negative binomial with parameters ν, t, τ. For each spill that occurs, associate a random variable X which takes the value 1 if a specified event occurs (such as the spill hitting land and O otherwise. Let X be a Bernoulli random variable.

$$P(X=1) = P \tag{A-28}$$

Let T be the total number of events that occur when spills originate from a single source,

$$T = \sum_{i=1}^{N} X_i \tag{A-29}$$

From section 2 of this appendix, the generating function of T is

$$\Phi_T(s) = \Phi_N(\Phi_X(s)) \tag{A-30}$$

From equations A-6 and A-8

$$\Phi_T(s) = \left(\frac{\tau}{t+\tau-t(1-p+ps)}\right)^{\nu}$$

$$= \left(\frac{\tau}{pt+\tau-pts}\right)^{\nu} \tag{A-31}$$

Thus, the distribution of T, the number of events, is in turn negative binomial, but with parameters ν, pt, and τ.

U. S. GOVERNMENT PRINTING OFFICE 1982—576-034/116