

Development of a simulation model of Mallard Duck populations

**by Carl J. Walters,
Ray Hilborn, Emily Oguss,
Randall M. Peterman,
and Jeffrey M. Stander**



Environment
Canada

Environnement
Canada

0016388A S

OCCASIONAL PAPER (CANADIAN WILDLIFE
SERVICE)

**Occasional Paper
Number 20**

**Canadian
Wildlife Service**

SK
471
C33
No. 20

7003466H

0016388AS
**Development of
a simulation model
of Mallard Duck
populations**

by Carl J. Walters, Ray Hilborn, Emily Oguss,
Randall M. Peterman, and Jeffrey M. Stander*



Environment Canada
Wildlife Service

Environnement Canada
Service de la Faune

Canadian Wildlife Service
Occasional Paper Number 20



SK
471
C33
No. 20

*Institute of Animal Resource Ecology,
The University of British Columbia

SCF

Issued under the authority of the
Minister of the Environment

Canadian Wildlife Service

©Information Canada
Catalogue No. CW69-1/20
Ottawa, 1974

Contents

4	Acknowledgments
4	Abstract
4	Résumé
5	Introduction
6	Basic principles of simulation modelling
7	1. Notation and symbolism
8	2. Key components of models
8	3. Steps in model building
12	4. Models and data
14	5. Judging the performance of models
15	The Mallard simulation model
19	Weather and pond generation
19	1. Weather
19	2. Pond numbers
20	3. Data availability
20	4. Comments on weather and pond generation
21	Production and summer survival
21	1. Factors affecting production
23	2. Calculation of production
23	3. Summer survival and fall population structure
24	Hunting and southward migration
24	1. Migration movements
25	2. Calculation of the kill
27	3. Hunting effort and time distribution
28	Northward migration
30	Results and discussion
34	Literature cited

List of figures

10	Figure 1. Interaction tables provide a format for systematically identifying factors and relationships
----	--

11	Figure 2. Functional relationships may be used in models to represent the action of various biological mechanisms
11	Figure 3. Another way of representing the relationship shown in Fig. 2. With appropriate units, either relationship could be used in a model to give the same prediction
12	Figure 4. Sequence of calculations that might be performed in the simple duck population model given in eqs. 7-12
13	Figure 5. Most dynamic simulation models have the same basic format: rules for change that can be applied repeatedly
13	Figure 6. Field data are usually not adequate to estimate functional relationships in models. Field experiments involving deliberate manipulation of populations are necessary to fill in the gaps
16	Figure 7. Factors and interrelationships considered in the Mallard simulation model
17	Figure 8. Management or statistical areas represented in the Mallard simulation model. Each area was treated as homogeneous with respect to habitat quality and movement of birds during migrations
18	Figure 9. Sequence of calculations in the Mallard model. Each named box (i.e. PEEP, KABLUI) refers to a computer subroutine or sequence of equations
21	Figure 10. Relative reproductive success related to index of land use intensity in the Mallard model
22	Figure 11. Relationship between relative reproductive success and rate of pond loss in the Mallard model. This relationship results in decreased effects of pond loss in areas where absolute pond density is high

- 22 Figure 12. Reproductive success is assumed to decline when the density of females per pond becomes too high
- 23 Figure 13. Excess males are assumed to affect reproductive success
- 25 Figure 14. Migration movements each simulated fall are calculated from a series of monthly tables, where each table gives the proportions of birds moving from every area to every other area in a month. Such tables are called Markov matrices
- 25 Figure 15. Kill rate per hunter is represented in the model as a function of bird density according to a series of equations with the graphical form shown
- 28 Figure 16. Relative vulnerability of Mallards to hunting is assumed in the model to change in relation to how many birds have already been shot
- 29 Figure 17. Birds that cannot find breeding space in areas to which they home are assumed in the model to move on, with varying probabilities of going in different directions as shown in this example
- 30 Figure 18. Results from an early run of the Mallard model. Pond numbers were seriously overestimated at first, resulting in unrealistically high population size
- 31 Figure 19. The Mallard simulation predicted sudden and dramatic population declines under increasing hunting pressure. Notice that harvest and production curves give no early warning of the decline
- 31 Figure 20. The Mallard model predicts that no population decline will occur if habitat is improved and hunting effort is kept at present levels

List of appendices

- 34 Appendix I. A FORTRAN program for the simple duck model described in the text.

Acknowledgments

We wish to thank the Canadian Wildlife Service personnel who contributed their time, ideas, and data for our modelling work; each deserves co-authorship on this paper. Special appreciation is due Dr. J. B. Gollop and Dr. W. J. D. Stephen for organizing our interaction with the CWS and for providing criticism of this manuscript.

Abstract

This paper describes an attempt to develop a comprehensive computer model of Mallard Duck populations in North America. The model considers major population processes and spatial distribution patterns in an attempt to predict long term population trends. Though predictions of the model have been questionable, it has provided a valuable means to help identify critical data needs for future work.

Résumé

La présente étude décrit une tentative faite en vue de la création d'un modèle informatisé général des populations de Canards mallards en Amérique du Nord. Le modèle tient compte des principaux phénomènes qui se rattachent aux populations et des modes de distribution spatiale, dans un effort pour prévoir les tendances démographiques à long terme. Bien que les prévisions obtenues par le modèle soient contestables, ce dernier a permis de déterminer des besoins importants en données pour de futurs travaux.

Introduction

The past few years have seen rising interest among resource managers in the tools of systems analysis and computer simulation. These tools seem to offer a powerful means of tackling large-scale problems of information synthesis and resource forecasting. However, a major problem has been that biologists are seldom trained in quantitative methodology, so mathematical techniques and computers have been viewed with fear and distrust. "Garbage in, garbage out" is an axiom of computer programming, expressing the fact that good biological input is essential for good resource simulation. Some attempts have been made to bridge the communication gap between biologists and formal systems analysts, but these attempts have not been particularly successful. It appears that resource people must learn to do their own model building. Recently attempts have been made to develop teaching programs that avoid the jargon and confusion of standard mathematics and computer training.

Simulation modelling can provide special benefits beyond information handling and forecasting. These benefits are often the best justification for modelling activity and arise from two characteristics of resource problems. First, these problems are large and require teamwork in data collection and interpretation; generally specialists from several disciplines must work together. Second, the problems are difficult to define and the goals of management activity are thus hard to identify. Model building enters the picture by providing a common language, a focus for mutual attention, and a concrete goal (the simulation model) for the team.

Experience at the University of British Columbia (Holling and Goldberg, 1971) has suggested that problem-oriented workshops can be used to teach basic ideas in modelling, and at the same time produce useful simulation schemes. This report describes the teaching methods and

results of such a workshop, held for the Canadian Wildlife Service in May and June of 1971. The explicit goal of this workshop was to develop a simulation model of the population dynamics of the Mallard (*Anas platyrhynchos*) in North America, to make long term predictions of population response to changes in habitat and harvest.

The workshop was conducted in two one-week sessions, three weeks apart. In the first session we presented lectures on basic concepts of simulation, gave demonstrations of some existing computer models, and established the scope and conceptual framework for a Mallard simulation model. Between the sessions, a group of four U.B.C. graduate students (the junior authors) programmed the basic framework of the simulation model, while CWS participants collected data and decided on the biological relationships to be represented. In the second session we assembled and ran the final simulation model.

We have written this report in three sections. In the first, we review some basic principles of simulation as presented in the first workshop session. In the second, we describe the Mallard simulation model in its final form. In the third, we evaluate and discuss the usefulness of the workshop and the simulation model.

Basic principles of simulation modelling

Ecological simulation is so new a tool that no generally accepted synthesis of its principles and limitations has yet been published. We present only one viewpoint in this paper; other introductory discussions can be found in Watt (1968), Patten (1971), Jeffers (1972), Forrester (1971), Holling (1972), and Walters (1971). Our discussion is for readers who are unfamiliar with mathematics and computers. We are mainly interested in dispelling commonly held misconceptions about model building.

Let us examine a typical set of calculations made by resource managers who are trying to predict sustainable yield from a duck population. The simplest estimate of sustainable yield would be

$$\text{sustainable yield} = \% \text{ gain}/100 \times \text{population size}$$

where percent gain is estimated from an unhunted, growing population. This crude estimate requires little understanding of the dynamics of the duck population. A more accurate estimate breaks population gain into production and loss components:

$$\text{sustainable yield} = (\text{prodn. per duck} - \text{loss per duck}) \times \text{population size}$$

This second calculation would require considerably more data, although it gives little increase in precision of the sustained yield estimate. Next, the time sequence of events in the population can be considered, and a series of calculations generated to describe the annual sequence of changes in population:

$$\begin{aligned} \left(\begin{array}{c} \text{spring popn.} \\ \text{next year} \end{array} \right) &= \left(\begin{array}{c} \text{adult popn.} \\ \text{this spring} \end{array} + \text{prodn.} \right) \\ &\times \left(\begin{array}{c} \text{surv. rate} \\ \text{to fall} \end{array} \right) \times \left(\begin{array}{c} \text{survival rate} \\ \text{through hunting} \end{array} \right) \times \left(\begin{array}{c} \text{survival rate} \\ \text{through winter} \end{array} \right) \end{aligned}$$

$$\text{harvest} = \text{fall popn.} \times \left(1 - \begin{array}{c} \text{survival rate} \\ \text{through hunting} \end{array} \right)$$

These relationships can be used to generate survival-through-hunting data and to evaluate sustainable harvest by trying out different values for harvest rates and noting their relationship to population next spring. This calculation would allow consideration of replacement mortality (those birds that are killed which would otherwise have died naturally). Given still more information, one might differentiate the age classes in the population by assigning to each its own survival and birth rates and vulnerability.

$$\begin{aligned} [1] \text{ Prodn.} &= \left(\begin{array}{c} \text{new adult} \\ \text{in spring} \end{array} \times \begin{array}{c} \text{prodn. rate for} \\ \text{first breeding} \end{array} \right) \\ &+ \left(\begin{array}{c} \text{old adult} \\ \text{in spring} \end{array} \times \begin{array}{c} \text{adult prodn.} \\ \text{rate} \end{array} \right) \end{aligned}$$

$$[2] \text{ Fall popn.} = \left(\begin{array}{c} \text{new + old} \\ \text{spring adult} \end{array} \right) \times \left(\begin{array}{c} \text{adult summer} \\ \text{surv. rate} \end{array} \right)$$

$$[3] \text{ Fall juv. popn.} = \text{prodn.} \times \left(\begin{array}{c} \text{juv. summer} \\ \text{survival rate} \end{array} \right)$$

$$\begin{aligned} [4] \text{ Harvest} &= (\text{fall juv.} \times \text{juv. kill rate}) \\ &+ (\text{fall adult} \times \text{adult kill rate}) \end{aligned}$$

$$[5] \left(\begin{array}{c} \text{New adult} \\ \text{next spring} \end{array} \right) = \text{juv. surv.} \times \left(\begin{array}{c} \text{winter surv. rate} \\ \text{for juveniles} \end{array} \right)$$

$$[6] \left(\begin{array}{c} \text{Old adult} \\ \text{next spring} \end{array} \right) = \text{ad. surv.} \times \left(\begin{array}{c} \text{winter surv. rate} \\ \text{for adults} \end{array} \right)$$

This sequence of equations is about as complicated as would ever be attempted without resorting to a computer. An obvious extension would be to repeat eqs. 3-6 over several years, while varying production and harvest rates over the time

periods of the calculations. The rates appearing on the right sides of eqs. 3-6 are called *driving variables*. The other quantities are called *system state variables*. The extended sequence of calculations we call a simulation model. Given considerable time, we can do the calculation sequence for several species, in several areas, and even account for migration movements between areas.

Each set of calculations like the one above is what we call a model. The basic principles which emerge from such a set are:

(1) We can write down the calculations to be made without reference to any specific numbers, and some shorthand notation for the various variables would be useful.

(2) Calculations about biological systems can be organized into hierarchies of complexity, but there is no objective way to decide when to stop increasing the complexity of the calculations.

(3) As models become more detailed we need more information, we have to make more assumptions, and the possibility of errors in our predictions becomes more likely. In addition, we are likely to leave out some critical factor which may have a disproportionate effect.

(4) As we add more detail, it becomes harder to see intuitively the consequences of the model.

(5) More detailed models require that we define variables more accurately and pay more attention to logic and consistency.

The key point to be made is that there is no best way to describe a particular system; the value of each model or calculation sequence depends on the particular situation to which it is being applied.

A further important principle is that the boundaries of the system to be modelled are arbitrary; these boundaries must be carefully defined by the model builder. In eqs. 1-6 we have implicitly set one system boundary by saying that harvest shall be described in terms of constant kill rates only. Instead, we could have extended the

boundaries of the model to include calculations of potential and actual numbers of hunters and their kill by considering aspects of the human population and its growth. One way to define a driving variable is to say that it is some factor whose variation is determined by forces outside the arbitrary boundaries of the system under study, e.g. light conditions. When we change a model to include calculations or predictions about a factor that we have previously called a driving variable, then that factor is no longer called a driving variable but is instead part of the arbitrary system (a system or state variable).

1. Notation and symbolism

Two steps are necessary to rewrite eqs. 1-6 in order to condense them and make them easier to deal with:

(1) We must assign symbols to variables and constants.

(2) We must rewrite our basic sentences using the symbols.

Let us make a list of symbols to use in eqs. 1-6.

P	= total production
NAS	= old adults in spring
NAF	= fall adult population
NJS	= new adults in spring
NJF	= fall juvenile population
H	= harvest
KA	= adult kill rate
KJ	= juvenile kill rate
PA	= adult production rate
PJ	= production rate for first breeding
SAS	= adult summer survival rate
SJS	= juvenile summer survival rate
SWJ	= winter survival rate for juveniles
SWA	= winter survival rate for adults

With these symbols, we can rewrite eqs. 1-6 as eqs. 7-12 respectively:

- [7] $P = NJS \cdot PJ + NAS \cdot PA$
- [8] $NAF = (NJS + NAS) \cdot SAS$
- [9] $NJF = P \cdot SJS$
- [10] $H = NJF \cdot KJ + NAF \cdot KA$
- [11] $NJS = NJF \cdot (1 - KJ) \cdot SWJ$
- [12] $NAS = NAF \cdot (1 - KA) \cdot SWA$

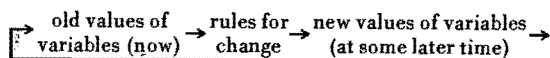
This is simply a condensed way of writing the series of rules for calculations and has the same biological meaning as the original sentences. Difficulty in understanding papers that contain equations is usually a problem of understanding the definitions that authors choose for their symbols. Notice also that the equations and symbols shown are meaningful only if presented in the proper order; this is often true of models.

2. Key components of models

We classify the elements in any calculation sequence or model as:

- (1) system state variables—the entities which the model tries to predict; indices of the state of the biological system;
- (2) parameters—constants, such as survival rates, which are necessary in the predictions;
- (3) equations—those shorthand sentences which say how system state variables and parameters are related and state the basic rules for the calculation;
- (4) driving variables—the factors, such as kill rates, that we want to manipulate or vary over time but that are not to be predicted within the basic calculation sequence.

The system state vector is the list of all the system state variables. Dynamic models are calculation sequences that try to predict change over time. The basic structure of any dynamic model can be shown as:



We usually try to make the rules for change (the model) fairly general so that we can have the new values of the variables become old values in a repeated sequence. The application of this repeated sequence is called a simulation.

Rules for change can be specified in a variety of ways that fall into three classes: continuous, event-oriented, and state-oriented.

We specify rules for continuous change in terms of differential equations that indicate how fast each variable is changing over every moment of time. We usually try to avoid continuous system models because they are often hard to formulate and solve.

In setting up event-oriented rules, we first specify how much variable change is to occur (e.g., loss of one animal); the rules are then stated in terms of the amount of time required before the change should occur. Event-oriented models are especially useful in describing processes like predation, where we want to calculate the amount of time between successive attacks by a predator.

State-oriented rules are usually the easiest to specify and form the basis for most biological simulation schemes. Here we start with the list of variables describing the state of a system at some time, and specify our rules so as to give the system state at a fixed later time directly in terms of the starting state. The population-harvest models given above are state-oriented models.

3. Steps in model building

3.1. Decide purpose and scope of model

The first step in model building is to decide exactly why the model is being built. We cannot go ahead until the following questions are answered. What predictions are wanted? How precise should these predictions be? Over what range of situations and for how long should the predictions be applicable? What information is available for inclusion in the calculation sequence? It is

obvious that we could continue to build models of increasing detail and complexity, without knowing when to stop, what to include or what to leave out. There are no formal rules or guidelines to help the model builder at this stage.

3.2. Choose variables to be included

Simulation models are always based on a set of numerical indices of system condition (just as we always measure indices of system condition in field or laboratory studies). Commonly used indices are numbers of animals in a population, numbers of hunters in an area, and numbers of ponds available for breeding birds. Indices or variables used in a simulation model need not necessarily be the best measures of the condition of the system to be simulated. To decide whether or not it is useful to include a particular variable, we have to know the specific purpose for which the model is being built.

How detailed should the model be? Again, this depends on the circumstances to which the model is applied, and the kinds of questions being asked. One of the factors which will influence our choice of variables is the fact that predictions are always conditional. Of necessity every prediction we make assumes certain regularities about the circumstances surrounding the study. For example, in developing duck population models we must assume regularities about recreational demand and about genetic composition of the duck population. Thus, our predictions are always in the form, "if the following circumstances occur, then we expect the following factors to change in such and such a way."

There are some criteria for choosing the variables. First, for highly correlated biological factors, only one factor need be represented in a model. For example, if survival rates for two age classes of animals are approximately equal, then a single survival rate parameter will suffice. As

another example, if pond drying rates are correlated with initial numbers of spring ponds, we need only include the latter (because it is easier to measure) for predicting potential production. Second, we can watch for factors which, when taken together, may qualify one another. For example, hunters vary in individual success, but statistically they may act as a unit whose success rate remains constant. Also, increasing the number of hunters may simultaneously lead both to interference in hunting activities and to their facilitation (due to increased numbers of birds in the air at any time). When the effects of two variables are expected to cancel one another, we can treat the total effect as being constant. In this example, we may be able to treat hunter success as constant and independent of the numbers of hunters. Third, we can ask what factors will have constant effects over all ranges of possible model application. We can treat these factors as parameters or determinants of parameters, and estimate their effects empirically. For example, we usually assume that populations will have constant genetic composition over periods of a few years; we can consider the effects of genetic factors on production in terms of empirical production rates that can be estimated from field data (for short predictions only).

In choosing variables we must be careful to distinguish between system state variables and driving variables. For example, in building a waterfowl harvest model we have the choice of trying to simulate recreational demand as a system state variable or treating demand as a driving variable. If demand is treated as a driving variable, then different demand patterns can be tested for their effects, giving a series of conditional predictions about population change. In general, more and more driving variables must be treated as system state variables as one increases the time span over which the model is to apply.

Figure 1
Interaction tables provide a format for systematically identifying factors and relationships

Figure 1

Effect "of"	Effect "on"			
	Total production	No. of adults	Effort level	Harvest
Total production			X	X
No. of adults	X	X	X	X
Effort level		X	X	X
Harvest		X	X	X

(X indicates direct effect of row variable on column variable)

3.3. State basic relationships among variables

Once a basic set of variables or factors has been chosen for simulation, one must decide what factors interact with one another, and in what time sequence and pattern relationships occur. While stating basic relationships, we may discover other variables which should be included in the model.

A useful device for helping to identify basic relationships is the interaction table, a cross-listing of the factors to be included in the model (Fig. 1). By checking each row against each column in the table, we can look at all possible interactions between system variables, and decide which interactions to include in the model. Such tables are particularly useful in designing models that describe flows of materials or individuals between different parts of a system or between spatial areas.

Once the basic lists of variables and their interactions have been established, we can concentrate on specific parts of the model, confident that a coherent picture is being maintained. This is one of the primary values of model building in resource management: with simulation schemes we can look very carefully at each part of a system while building a description of how the parts fit together.

In order to describe basic relationships among variables we must concentrate on one variable at a time, and be precise in its definition. When we are certain of the biological factor or event that is represented by the variable, we can state how the variable will change, considering the degree of resolution desired of the model. For example, if it is clear that production is to mean the number of newborn Mallards in Manitoba that survive to their first autumn, then we may describe production as the product of

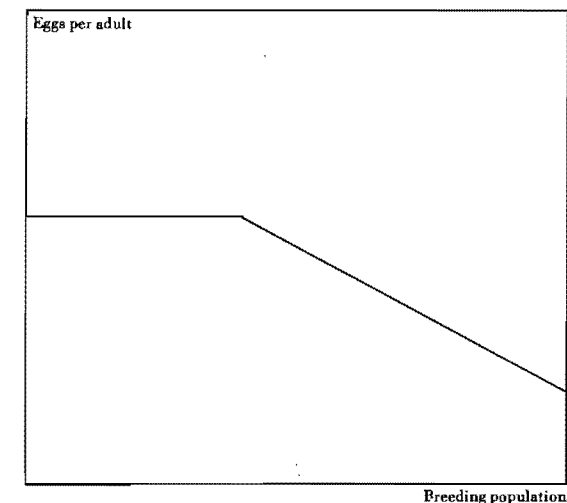
$$\left(\begin{array}{c} \text{spring breeding} \\ \text{population} \end{array} \right) \times \left(\begin{array}{c} \text{eggs prod.} \\ \text{per adult} \end{array} \right) \times \left(\begin{array}{c} \text{surv. rate of} \\ \text{eggs to hatching} \end{array} \right) \\ \times \left(\begin{array}{c} \text{survival rate of} \\ \text{chicks to fledging} \end{array} \right) \times \left(\begin{array}{c} \text{surv. rate through} \\ \text{early flight period} \end{array} \right)$$

Each of these factors can then be broken down into sub-factors, treated as constant, or related to other variables or factors in the model. As an example of this last method, we can describe by means of a graph the egg production per adult in terms of size of breeding population (Fig. 2). Then the population size beyond which production drops may in turn be described in terms of the availability of ponds or other factors. In this example, we are using breeding population size as an index to the conditions which birds will encounter in terms of factors such as competition for nesting sites or food supply.

If simple linear equations will not adequately describe a relationship, it is often best to express

Figure 2
Functional relationships may be used in models to represent the action of various biological mechanisms

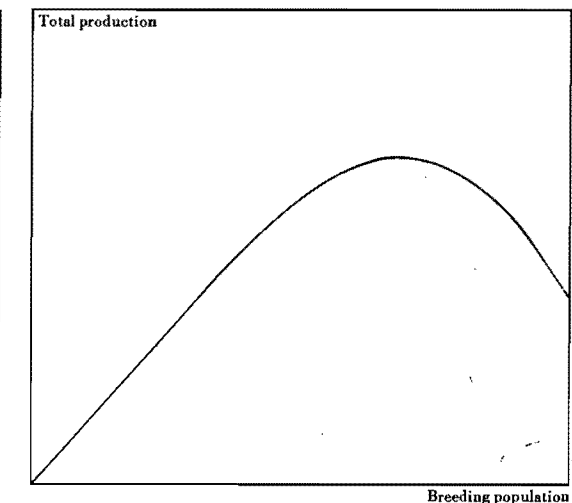
Figure 2



the relationship in terms of a graph. There are computer techniques for entering graphical relationships directly into models. Suppose we are studying waterfowl production, and we have information to support the following assumptions: that production is proportional to numbers of breeding adults for low population densities; that there is a maximum production, set by availability of ponds and by territorial characteristics of breeding birds; and that production will drop off at very high population densities, due to competition between breeding birds and failure of food supplies for young. We can represent all these assumptions at once by means of a simple graph of production versus breeding population size (Fig. 3). Specific data can then be used to scale the graph axes to give proper maximum production rates and breeding population sizes. In graphical representation no biological content is lost by stating a relationship in that form. Several

Figure 3
Another way of representing the relationship shown in Figure 2. With appropriate units, either relationship could be used in a model to give the same prediction

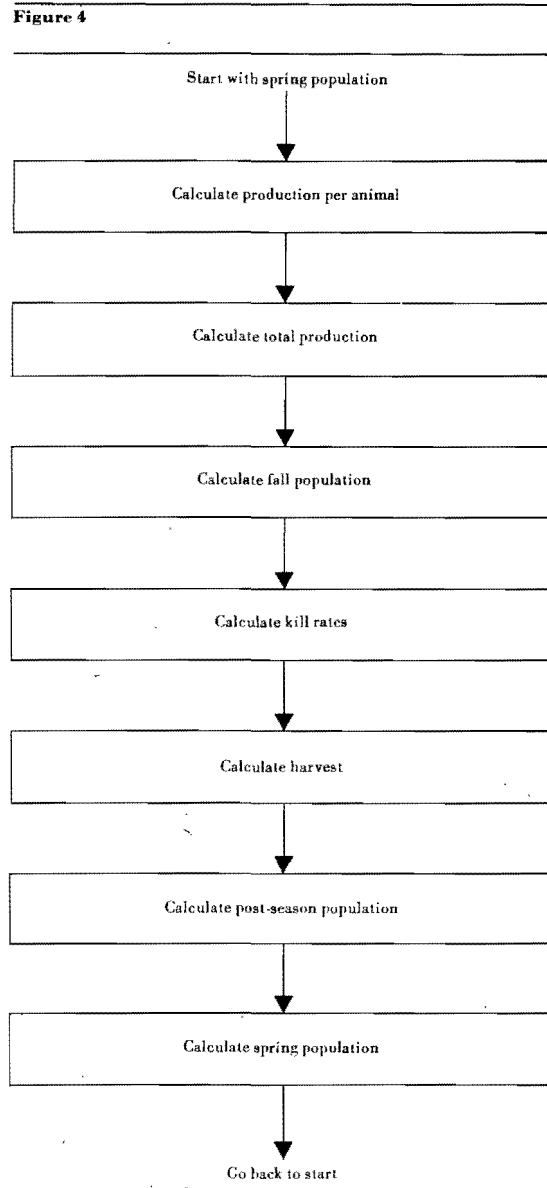
Figure 3



different biological relationships may result in the same graph. When this happens, the behaviour of the model is invariant to certain assumptions.

Frequently a particular relationship is not well understood or supported by data. For example, suppose we are trying to describe kill rates for a population in terms of the numbers of hunters. The problem is that increased numbers of hunters might result in decreased individual success, increased success, or no effect on individual success. Although more data might help to resolve this problem, decisions and predictions must be made in the meantime. Rather than ignore the problem, or use a simpler model, it seems best to develop the model, and test it with several alternative assumptions. Sensitivity analysis is the term used for the process of testing the effects of different assumptions and parameter values on model predictions.

Figure 4
Sequence of calculations that might be performed in the simple duck population model given in eqs. 7-12



3.4. Illustrate the basic relationships

A useful tool for illustrating relationships among variables is a flow chart (Fig. 4) showing the calculation sequence with boxes and arrows. This sequence will usually follow the real sequence of events that is to be simulated. Each box, in itself, may represent a whole series of calculations (e.g. "total production") that could be shown in another, more detailed, flow chart.

3.5. Program the model for the computer

Writing the computer program is relatively simple once basic relationships have been clearly stated and flow charts have been designed. A number of computer languages have been designed expressly for simulation (DYNAMO, GASP, SIMULA, etc.), but experience has shown that the standard FORTRAN or ALGOL usually give the best results.

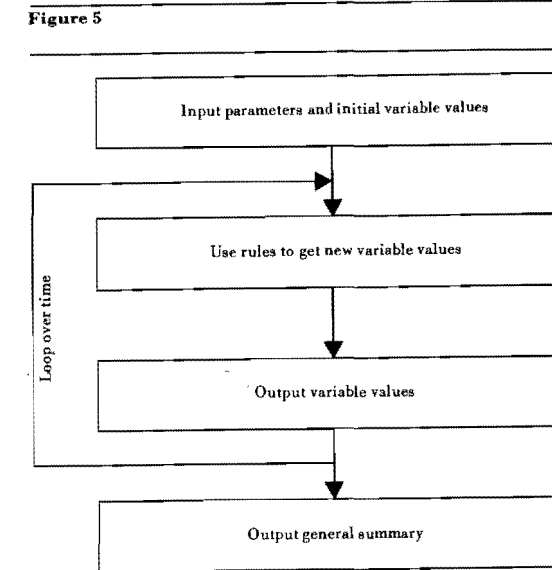
Most simulation programs have three basic parts: parameter and initial variable input; the simulation sequence; and variable output (Fig. 5).

A simple FORTRAN program, implementing the duck model in eqs. 7-12 is given in Appendix I. No special programming tricks are needed to develop simulation models. Repeated use of the same variable names in time sequences of calculations and simple looping and branching operations are the only essential programming conventions. For more complicated simulation models such output devices as plotters and cathode ray displays are useful.

4. Models and data

A major problem in model building is estimation of parameters, initial values of state variables, and driving variables. In some situations we avoid the problem by making only broad conditional predictions of the form: if parameter A is in the range X_1 to X_2 , then pattern Q will occur if parameter B is in the range Y_1 to Y_2 . For example,

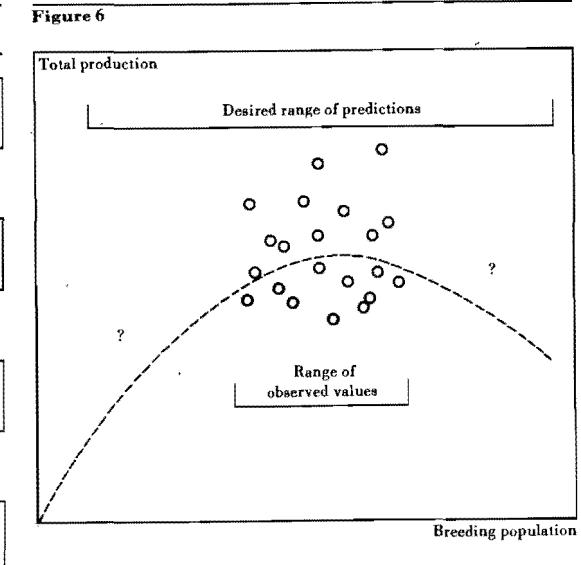
Figure 5
Most dynamic simulation models have the same basic format: rules for change that can be applied repeatedly



we may say that if production rate is in the range 0.9 to 1.3, then if each hunter kills between 0 and 3 birds, when there are 30 birds available, then the duck-hunter predation system will remain self-regulating. However, unless most of the parameters are well-established, conditional predictions are almost meaningless in complicated models.

Most field data are of limited value in parameter estimation. This is because when left alone natural systems usually do not vary over the full range of conditions that we might like to examine with a model. For example, in the problem of predicting production of ducks in terms of breeding population size, we may want to predict production for a wide range of breeding population sizes, although past data do not cover such a range (Fig. 6). However, resource systems that have experienced great changes in exploitation rates and management policies do give a wide range of

Figure 6
Field data are usually not adequate to estimate functional relationships in models. Field experiments involving deliberate manipulation of populations are necessary to fill in the gaps



past data. Studies of population response to progressive changes in exploitation have formed the basis for the few successful models that now exist, for example, in commercial fisheries management.

We can take three courses of action when dealing with a narrow range of field data on a particular relationship:

- (1) restrict our predictions to those situations for which data are available;
- (2) use our biological intuition to extrapolate beyond observed data;
- (3) try to resolve the overall relationship into simpler experimental components (Holling, 1972) for which better data may be available.

The first course of action is safest, but may defeat the purpose of the model. The second alternative is risky, but may prove best in many situations. Some relationships can be extrapolated with fair confidence, given some basic biological understanding about the system of study. For

example, we know that total production in the graph in Figure 6 must eventually fall off as breeding population decreases; if our predictions need not be too precise, we may assume that this drop will begin to occur at breeding populations just below those observed. A danger would be that in reality production might fall off very rapidly for low breeding populations, due to failures in mating or lack of social facilitation. Alternatively, we can use conditional predictions and base management policies on "least optimistic" assumptions.

The third course of action, experimental components analysis, is not necessarily best. It can greatly increase the number of assumptions in the model, without ensuring that model behaviour will not depend critically on just a few of these assumptions. In more complex models, the odds are greater that any one assumption will be incorrect; at the same time, there is no assurance that model predictions will not depend strongly on such erroneous assumptions. For example, in the problem of calculating total production for a duck population, our first step in an experimental components analysis would be to identify a series of time stages:

Mating → selection of nest area → egg laying → hatching → . . .

Each of these stages will provide a gain or loss factor. These factors, when multiplied together, give a final production rate of premating adults. If any one stage is inaccurately estimated, and if compensatory mechanisms do not operate in successive stages, then the resulting production calculation will be equally inaccurate. Luckily, nature seems to provide for compensation between life history stages. For example, low survival in one period may be followed by higher survival in later stages, so that overall survival is nearly constant. A good experimental components

analysis will reveal these compensatory mechanisms when they exist.

5. Judging the performance of models

We can never say that a model has been validated; its rules are always simplifications. Likewise, models should not be judged solely on their ability to fit past data and predict new observations. Models are intended to apply to situations that are in some respects novel (otherwise we would need no model, and could rely for decision-making on past data), and model predictions may fail in some but not all of these novel situations.

A model is not necessarily a bad one because it lacks numerical precision in fitting past data. For example, a waterfowl model should not be considered useless if it predicts a kill of 20,000 when the actual kill is 100,000. We make this assertion for two reasons. First, failure of the model may give us clues to errors in the formulation of the rules for change. If these rules embody our biological understanding, then the model is helping us to find errors in that understanding. Second, the model may predict the correct basic pattern of responses even if particular numerical results are in error. We can always re-scale or change the units of the model.

The model can be particularly useful if the patterns it predicts are counter to our intuitions. For example, consider a model of flyway harvest patterns in waterfowl management. Intuitively we may predict that some harvest pattern in one flyway will have a particular effect on subsequent yields in other flyways. The model may predict exactly the opposite effect if it is considering some interaction between flyways (e.g. through breeding populations) that we have omitted from intuitive consideration. A classic example of counter-intuitive model behaviour comes from aquatic biology. Limnologists have fertilized many lakes on the intuitive assumption that the effects of fer-

tilization should include increases in phytoplankton standing crops. Often these increases are not seen, so fertilization is discounted as a management tool for many situations. Recently, aquatic models have predicted that phytoplankton crops should rarely increase under fertilization and instead that only zooplankton standing crops should change (McAllister *et al.*, 1972). The reason is that potential increases in plant standing crop are quickly transmitted to zooplankton populations, and mean plant standing crop is determined by feeding and energetic characteristics of individual zooplankters rather than by phytoplankton productivity.

With these thoughts in mind, we should ask where models can go seriously wrong. Major errors seem to come when we badly misstate key rules of change or omit important factors from consideration. Minor errors (10–30%) in most parameter values usually have little effect on the patterns predicted by a model, although they may change the numerical results. Usually there are only a few critical parameters. Basing the model on the wrong factors is not necessarily bad, if these factors are strongly correlated with whatever variables are really important in the system. The biggest danger is that of omission. Suppose we are trying to predict recreational demand for a game population. We assume this demand is determined by the potential number of users and by past hunting success. We then get good correlation between these factors considering past data. But suppose that demand can be strongly influenced by communication and publicity, and when developing the model we assume these factors will remain constant. An unexpected series of newspaper articles or game management bulletins could make our predictions much too low, but there is no absolute standard for judging the merit of a particular model or decision-making method; there are only relative standards.

The model focuses on population biology and habitat, with those human factors (social, economic) that affect Mallard populations being treated as driving variables (Fig. 7).

The model explicitly takes into account spatial variation and movement of birds across North America. Thus, it allows detailed specification of harvest regulations by area and flyway, description of breeding habitats in terms of regional habitat types and land use patterns, and detailed representations of migration patterns in relation to these variables. Population and habitat factors are treated as being relatively homogeneous within each of 33 areas (Fig. 8) into which North America is divided. Movements of birds are represented as discrete jumps from one area to another. We did not program the model to be specific to any particular set of areas, so that future users would be free to split or combine areas for more detailed or simpler simulations.

The model also takes account of the sex-age class structure of the Mallard population, insofar as this relates to breeding, migration, and vulnerability to hunting. All calculations include five sex-age classes:

- 1—adult males
- 2—unsuccessful and non-breeding adult females
- 3—breeding adult females
- 4—juvenile females
- 5—juvenile males

In the program, the Mallard population is represented as a triple-subscripted array, POP (I,J,K). Each element in this array is the number of ducks of sex-age class K that last bred (or were born) in spatial area I and are now in spatial area J. The array is updated and changed throughout each simulated year as ducks are born, migrate across areas and die.

Figure 7
Factors and interrelationships considered in the Mallard simulation model

Figure 7

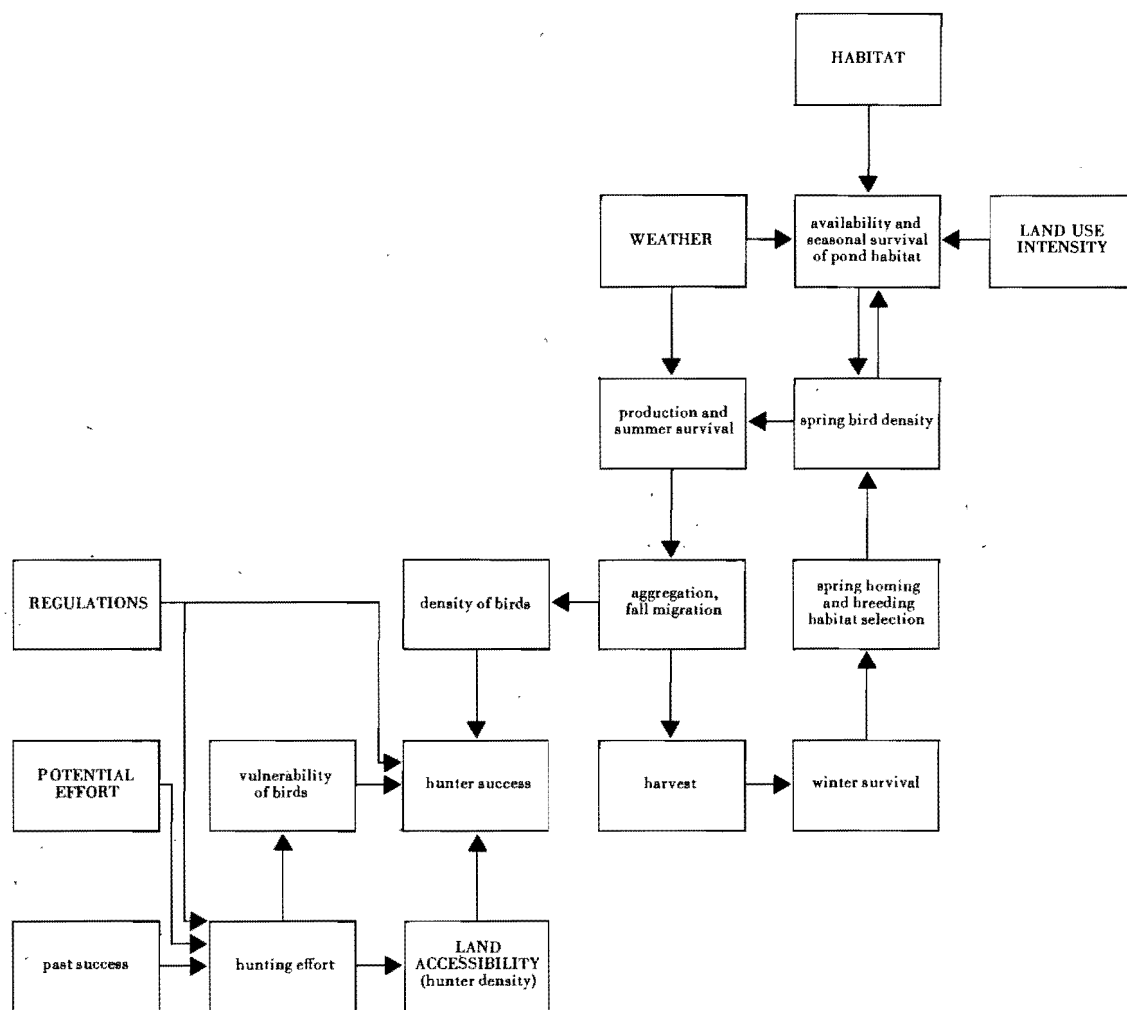


Figure 8
Management or statistical areas represented in the Mallard simulation model. Each area was treated as homogeneous with respect to habitat quality and movement of birds during migrations

Figure 8

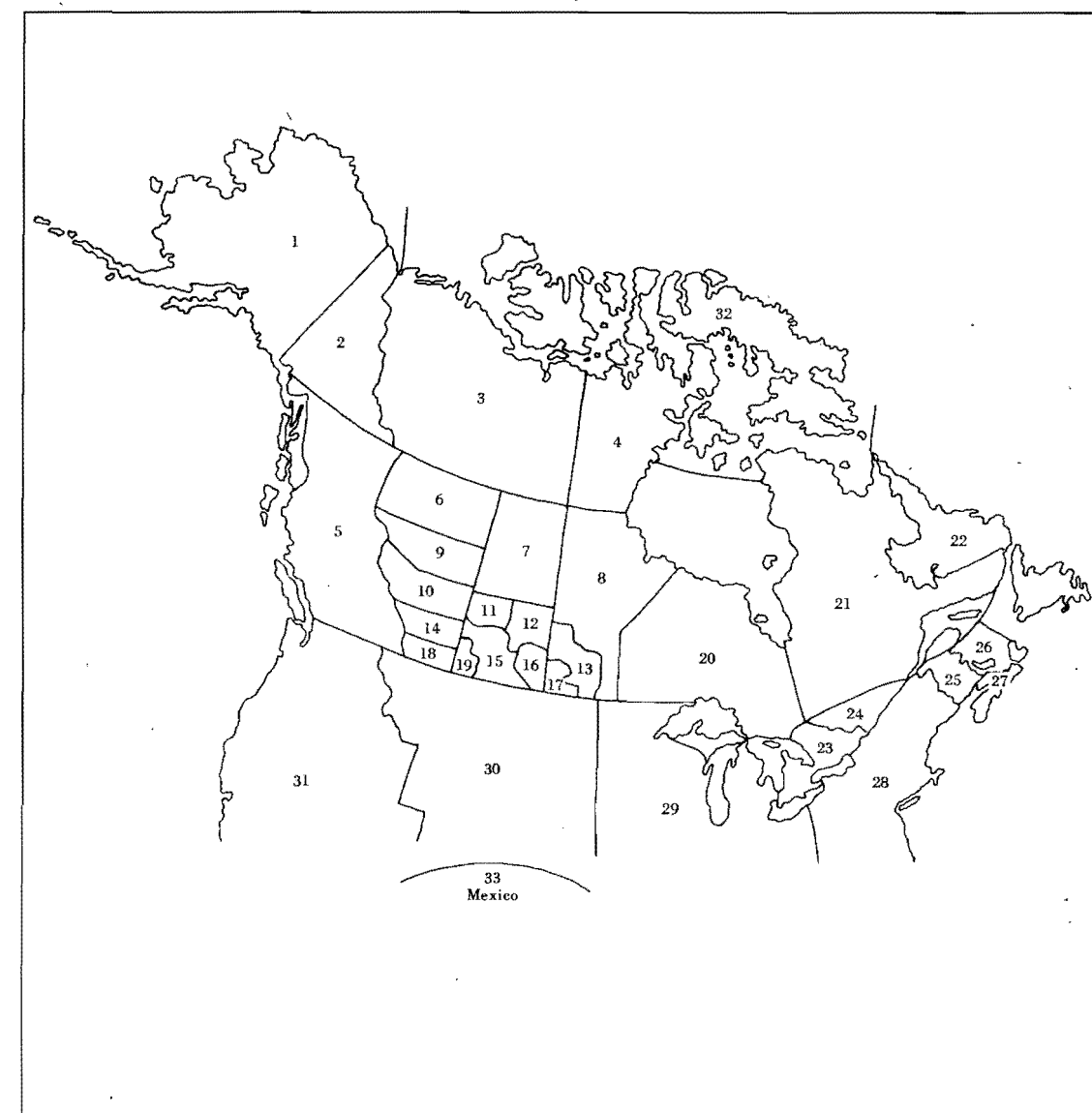
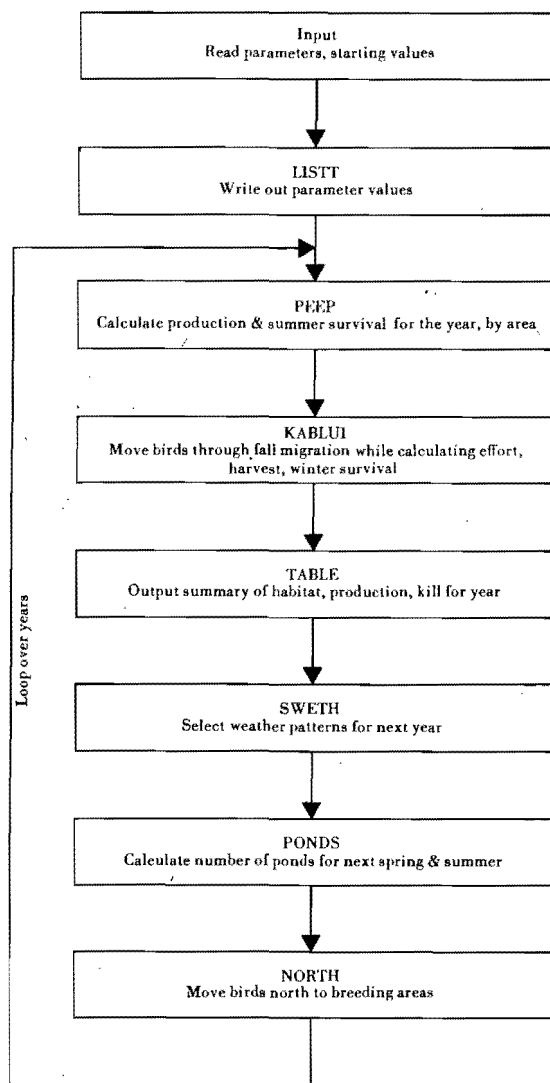


Figure 9
Sequence of calculations in the Mallard model. Each named box (i.e. PEEP, KABLUI) refers to a computer subroutine or sequence of equations.

Figure 9



The computer model consists of a series of subroutines, or submodels. Each submodel simulates a related set of habitat or population processes. The submodels are arranged so that the results from each provide starting values for the next submodel. The sequence of calculations in the complete simulation is shown in Figure 9.

The capitalized names like PEEP and NORTH are individual subroutines. Each simulated year starts in the late spring with production calculations (PEEP). This subroutine generates fall population size and sex-age structure for each spatial area. Then bird movement across the areas is simulated in five fall migration steps, each step representing one simulated month (KABLUI). Hunting effort and kill are simulated for each step and winter mortality is removed after the last migration step (KABLUI). The computer then prints out a winter summary of events for the year (TABLE). Next, spring weather and pond conditions for the next year are simulated (SWETH, PONDS). Finally, the northward movement of birds and their distribution across breeding areas are simulated for the next year (NORTH).

The following sections give more detailed descriptions of the submodels shown in the flow chart (Fig. 9). These descriptions are intended primarily for readers interested in the actual computer program, the assumptions underlying it, and the problems that we encountered in the estimation of model parameters from available data. For the casual reader, we begin each section with a general summary.

Weather and pond generation

1. Weather

Weather patterns in each simulated year are generated in subroutine SWETH. SWETH simulates three types of weather used in the other submodels: the weather that determines pond change from July to the following May; the weather that is concerned with spring timing (early, average, or late); and the weather associated with pond loss from May to July. We will call these run-off, spring, and evaporative weather. We assumed that weather patterns are the same throughout North America, so an early spring year in one management area will also be an early spring year in all others. Run-off and evaporative weather are each given a single index number in each simulated year, that index being the number of standard deviations above or below the average run-off or evaporation.

All weather can be treated in two ways: it can be specified as a parameter with the computer input data, or it can be randomly simulated in a way that permits many patterns to be created. The former method requires that run-off and evaporation are first determined as intensity indices, from 1 to 10 for run-off, and from 1 to 5 for evaporation. When these indices are input they are stored for up to 40 years, so we may utilize up to 40 years of weather.

If weather patterns are not provided to the computer as driving variable data, they are simulated in the following ways. For each type of weather, a matrix is read giving the probabilities of going from last year's run-off type I to this year's run-off, evaporative, or spring type J. This means we must design a 10 by 10 table for run-off types, a 10 by 5 table for evaporative changes, and a 10 by 3 table for spring weather. Finally, the various run-off and evaporative weather indices are converted into more meaningful units in terms of standard deviations above or below average weather conditions, by entering

as computer input the number of standard deviations corresponding to each index value.

2. Pond numbers

Simulation of annual and seasonal changes in pond numbers is done in the subroutine PONDS. This subroutine has three major sections. First, permanent losses or gains due to drainage and other factors are calculated for each management area, and added to or removed from the supply of potential ponds. Second, the number of ponds in May and July is calculated for each area. Third, the square miles of staging areas in August are calculated for each area.

We used six variables to characterize pond conditions in each management area: the maximum number of ponds, the minimum number of ponds, the average number of ponds lost from May to July, the standard deviation of the number of ponds lost from May to July, the average number of ponds gained from July to the following May, and the standard deviation of the number of ponds gained from the month of July to the following May.

The assumptions used to calculate drainage and other permanent basin changes are as follows. First, a certain number of ponds are added to the supply of ponds each year due to certain types of activities (gravel pits, farm ponds, etc.). This number is treated as a constant, and is read in as data. Second, a certain number of ponds are lost each year due to drainage. These losses are permanent basin changes, so both the maximum and minimum pond numbers are decreased accordingly. Finally, the proportion of ponds that is drained each year is a linear function of the proportion of the drainage cost that the farmer must pay.

The equation used is

$$\text{proportion of ponds drained} = a + bx$$

where

- a = the proportion of ponds drained with no subsidy
- b = the slope of the line, a measure of the additional number of farmers willing to drain ponds per unit increase in the subsidy rate
- x = the subsidy rate ($1 -$ proportion the farmer pays)

If we assume that $b = 0.05$ and $a = 0.0$, a 100% subsidy would cause 5% of the ponds in an area to be drained. Values for b were selected empirically from data on pond drainage rates in the Prairie Provinces.

The second stage of PONDJ calculates the actual number of ponds in May and July. The number of ponds in May is taken to be the sum of the number of ponds in the previous July and the number of ponds gained since then. The pond submodel receives from the weather submodel the type of run-off and evaporative weather in standard deviations above or below the average, and uses this information to calculate the actual number of ponds in May and July within the constraints imposed by the appropriate maximum and minimum. The number of ponds in July is taken to be the number in May minus the number lost between May and July. The number gained from July to May is the average number gained plus the standard deviation times the run-off weather type. The number of ponds lost from May to July is equal to the average number lost plus the standard deviation times the evaporative weather pattern. If the run-off weather (in standard deviations) is negative, the number of ponds gained will be less than average; if the run-off weather is positive the number of ponds gained will be more than average.

Even if we assume extreme weather conditions, the method of calculation used will continue to show a number of ponds in May and July somewhere between the maximum and minimum values initially established.

3. Data availability

In designing the pond submodel we used data that had been collected in the past. Starting values of the pond maxima and minima come from the U.S. Fish and Wildlife counts of ponds in May and July. Standard deviations were calculated empirically from those counts. The data needed for changes in permanent basins are the natural rate of gain, the slope of the drainage vs. subsidy line (see equation above), the rate of loss with no subsidy, and the present subsidy rate. Estimates of the rate of gain and the expected rate of loss with no subsidy were made by workshop participants based upon their work.

Where weather patterns were to be fixed as data rather than simulated, we estimated them by examining historical trends and by roughly calculating what each year represented in terms of standard deviations about the average.

4. Comments on weather and pond generation

We were disappointed with the weather and pond generation submodels. Estimates based on historical weather patterns consistently indicated more ponds than expected. We believe this is due to several false assumptions about weather and ponds: first, that weather is the same throughout North America (it is not necessarily true that a wet year in one place will be a wet year everywhere else); second, that the number of ponds fluctuates between a fixed maximum and minimum in any area. By imposing a minimum and maximum number the model artificially keeps the number at a higher end of the recorded range. A much more detailed model of pond response to weather is required to give reliable pond numbers applicable to 33 different areas.

Production and summer survival

The submodel for calculating production and summer survival, subroutine PEEP, begins with the ducks having already finished the annual northward migration, and their distribution among the management areas. In this submodel no migration between different areas is assumed to take place, so we can simplify discussion by referring only to a single area.

The model assumes that there are five factors which affect the overall production of young ducks each year.

1. Factors affecting production

(1) The first production factor is *brood* success. More broods are hatched in an early spring than a late spring; more broods are hatched in parkland than in other habitat types; an adult will have greater breeding success than a juvenile. Brood success is a tabled function and is selected from 18 possible values depending on the three habitat types, the three spring weather types, and two age classes. These data were provided by Canadian Wildlife Service workshop participants.

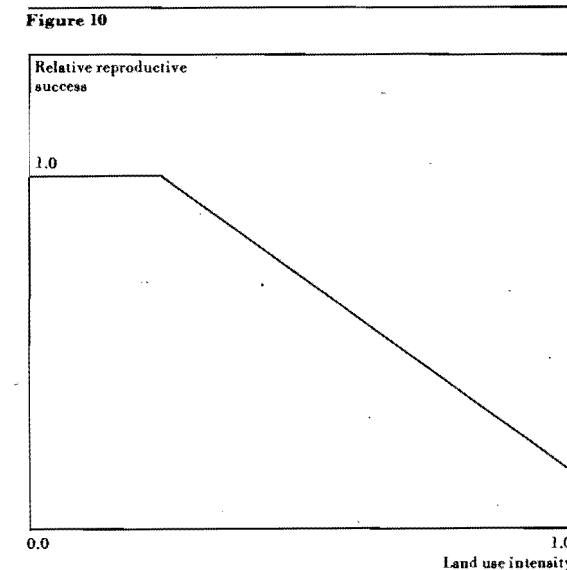
(2) *Nest-to-water* success is a function of land use intensity and habitat. Figure 10 shows the basic form of the function which varies with the three different habitat types. For each habitat type a maximum and a minimum success rate is specified. When land use is 1.0, i.e. 100% of land is used for agriculture, the minimum reproductive success rate is reached.

(3) Next, *water-to-water* success is evaluated. It is dependent on mean pond density and pond loss rates from May to July. Two parameters are used in evaluating this success function, PLC and PLD, and there is a different parameter pair for each habitat type.

The rate of pond loss is

$$RPL = 1 - (PONDJ/PONDM)$$

Figure 10
Relative reproductive success related to index of land use intensity in the Mallard model



where PONDJ and PONDM are the number of July and May ponds in our particular management area. RPL is confined in the model to positive values and is set at 0.0 for cases where July ponds exceed May ponds. The mean pond density, PDM, is the average number of ponds divided by the surface area in square miles of the breeding grounds.

The function for *water-to-water* success has the mathematical form

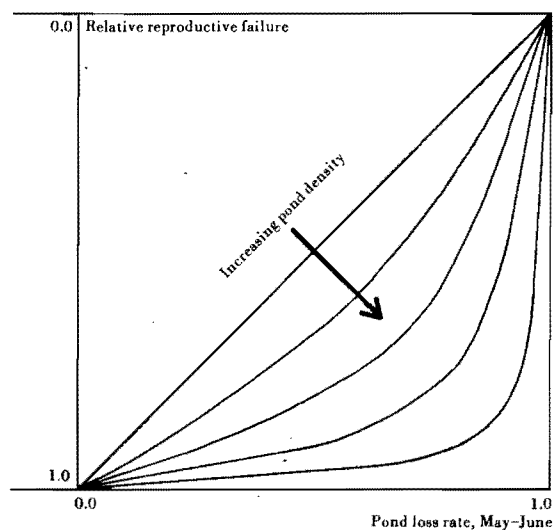
$$y = 1 - x^a$$

where x is between 0.0 and 1.0. If the exponent, a , is positive it is seen that y must also be between 0.0 and 1.0. For a given x , and 1.0, an increasing exponent implies a decrease in the value of y . In the model the function is expressed by the FORT-RAN statement

$$BS = 1 - RPL ** ((PDM - PLD + 1) * PLC)$$

Figure 11
Relationship between relative reproductive success and rate of pond loss in the Mallard model. This relationship results in decreased effects of pond loss in areas where absolute pond density is high

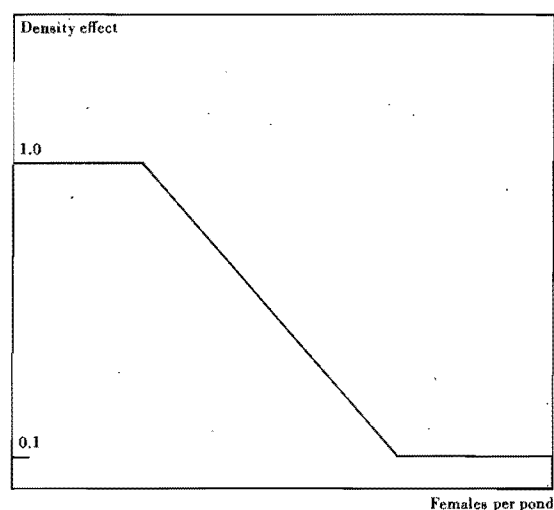
Figure 11



where the parameter PLD is a pond density threshold and PLC is a scaling factor for the exponent. If the pond density PDM is less than PLD then the success factor is set equal to $1 - \text{RPL}$. PLD was initially set equal to 3.0 ponds per square mile. Thus, if the pond density was less than this, success was linearly related to the pond loss rate, as indicated in Figure 11. For greater pond densities the exponent $(\text{PDM} - \text{PLD} + 1) * \text{PLC}$ causes drying to have less effect until a larger proportion of the ponds have dried up. The larger the exponent becomes, the more ponds must dry before BS drops significantly. Thus, a family of curves is defined by this function (Fig. 11) with increasing curvature as pond density rises. The idea is that when pond density is high, a high loss rate will only slightly affect duckling success up to the fledgling size. Conversely, a slight loss rate with low pond density will lower success very much. But as pond densities near the threshold

Figure 12
Reproductive success is assumed to decline when the density of females per pond becomes too high

Figure 12



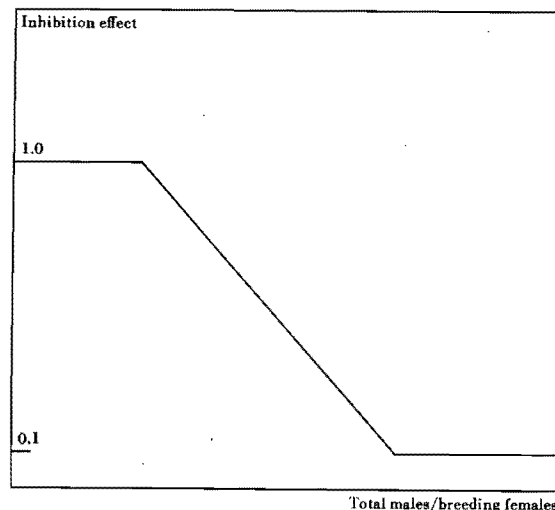
level, loss rates of young become close to the rate of pond loss.

(4) The next factor is the *density effect* of overcrowded ponds, which is graphically shown in Fig. 12. It is a possibility that a large population may result in dispersal of non-breeding females into an area (see the section on northward migration). In this case, total density of females per pond may exceed the pond's holding capacity. The density effect coefficient is equal to 1.0 minus the proportion of broods lost due to overcrowding.

(5) The presence of *excess males* (males without a mate) may cause a loss of broods due to harassment and other disturbances of the breeding females. A function similar to the density effect of overcrowded ponds was therefore used (Fig. 13). The success coefficient is equal to 1.0 minus the proportion of broods lost due to the excess male effect and it depends on the ratio of males to breeding females.

Figure 13
Excess males are assumed to affect reproductive success

Figure 13



2. Calculation of production

Production is calculated in each management area by assuming a hatching clutch size. This is dependent on habitat type (parkland, grassland, or boreal forest) and the age class of the hen. Within the calculations for each management area each of the five production functions outlined above is applied first to adult breeding females, then to juvenile breeding females. The number of fledglings surviving through the summer is estimated by computing the number of unsuccessful females from fractional brood losses and summer mortality. The fledglings are then distributed into their proper sex-age groups.

The fraction of early birds (those hatched before June 15) in the breeding population is also calculated. The ratio of male to female hatchlings is fixed at 1.0 but this can be altered by the model user.

The calculation sequence is given schematically below:

Hatching clutch size

times hen success coefficient
times excess male effect coefficient
times density effect coefficient
equals

Brood size

times nest-to-water success coefficient
times water-to-water success coefficient
equals

Final brood size

times number of breeding females
equals

Number of young

What we have done is to calculate an average surviving clutch size for a single breeding female, and to multiply this by the total of initial breeding females. This is a reasonable method, since we would obtain the same results if we had calculated a brood size for successful nests and subsequently multiplied it by the number of successful females.

3. Summer survival and fall population structure

Prior to summer, breeding females are lost only to the unsuccessful category or to predation and such losses are assumed to occur up through the time of hatching. Hen losses from the time the young are hatched until September 1 are subsumed under summer survival.

The proportion of unsuccessful adult and juvenile hens is multiplied by their respective population sizes to obtain the number of unsuccessful adults, which is then reduced by the proportion of unsuccessful females lost to predation, a factor that depends solely on habitat type.

25

the kill of each sex-age class of ducks as a function of (simulated) hunting effort, density of Mallards of all sex-age classes, and decreased vulnerability of all classes of birds due to previous hunting. Then the kills are subtracted from each class and the vulnerabilities of surviving ducks are reassessed.

We think the total kill (including cripples) per hunter-day is a function of the density of the ducks. We proposed a simple model for this which is shown graphically in Figure 15. Both the slope of the curve at low bird density and the maximum value can change during the season, and are dependent on the type of hunting (jump, pass or decoy) and skill of hunters, the vulnerability of the birds, and the distribution of the birds in relation to hunting areas.

Since kill rates have not been measured for a wide range of bird densities, only a few points along the curve could be obtained from existing hunter survey data, so it was necessary to identify the basic components or factors in the relationship. As a first approximation, we examined the time allocation of the hunter:

$$\left(\begin{array}{c} \text{Propn. of kill in} \\ \text{each class} \end{array} \right) = \left(\frac{\text{Total time spent hunting}}{\text{Time reqd. per bird of} \\ \text{that class killed}} \right)$$

Total time spent hunting was estimated from CWS data. The time required per bird killed was further broken down into:

$$\left(\begin{array}{c} \text{Time per} \\ \text{bird killed} \end{array} \right) = \left(\begin{array}{c} \text{Av. time btwn.} \\ \text{successful shots} \\ \text{at that class} \end{array} \right) + \left(\begin{array}{c} \text{Time to handle} \\ \text{and retrieve} \\ \text{each bird} \end{array} \right)$$

Handling time was estimated by CWS workshop participants from their own hunting experience. The time between successful shots (when birds are abundant) is a function of the wariness of

each class of birds and its relative density. We assumed that birds and hunters are randomly distributed in space within an area, and that the time between successful shots at one class could be described as a portion of a kill rate constant:

$$\left(\begin{array}{c} \text{Time btwn. successful} \\ \text{shots at one class} \end{array} \right) = \left(\begin{array}{c} \text{Kill rate} \\ \text{constant} \end{array} \right) \\ \times \left(\frac{\text{Density} \times \text{vulnerability of that class}}{\text{Density} \times \text{vulnerability of all ducks}} \right)$$

The kill rate constant is a function of the type of hunting, the average skill of the hunters and the size of the flocks; essentially it is the time between successful shots when there is one vulnerable bird per unit area. We estimated its value (different for each area) from typical data on opening day kill rates and from duck density estimates by the CWS. The method of shooting is an important factor because each method is believed to have a different efficiency. Because of lack of data, there is no separation of shooting types in the present model, and this factor is accounted for implicitly by adjusting the constant for each area.

It is assumed that decreased vulnerability due to avoidance learning can be expressed through decrease in the slope of the kill versus bird density curve (Fig. 15). This could occur either because each bird becomes more wary or because an increasing proportion of birds learns to keep out of the range of the hunters. In the model, the relative vulnerability of each sex-age class is expressed as a number which ranges from a maximum (0.4 for adults and 0.99 for juveniles) which applies during the first exposure of the class to hunting, to a minimum (0.2) which represents the vulnerability of the cleverest and most experienced ducks. A different kill versus density graph is generated for each sex-age class, where the maximum kill for each class is a function of the densities and vulnerabilities of all classes.

Within each area, the hunting submodel distributes ducks evenly over the total acreage of ponds and staging areas. We assumed that hunters have access to all land, i.e. there are no posted or protected areas, and that all sex-age classes are evenly distributed. Since it is known that ducks are not really distributed evenly (adult males go to staging areas earlier, etc.), the model could be improved if there were more information about the different methods of shooting. To calculate kill separately for the three basic types of hunting, the following information would be needed:

- (1) Acres of ponds and acres of staging water in each area.
- (2) Proportion of these acres accessible to hunters.
- (3) Proportion of ducks on ponds and staging areas (by age and sex) during monthly intervals.
- (4) Vulnerability and learning rates of juveniles and adults for each type of shooting.
- (5) Crippling rates for each type of shooting.
- (6) Hunter preferences for each type of shooting.
- (7) Effects of hunters on each other, based on the numbers of hunters per acre, for each type of shooting.

If the total kill per hunter-day (sum of kills of all sex-age classes) exceeds the bag limit for the area plus crippling loss, we reduce the figure to comply with the limit.

3. Hunting effort and time distribution

Data for the number of hunters came from the sales of waterfowl permits in Canada and of duck stamps in the U.S., and from the results of a survey conducted by the CWS. A number of factors influences the number of hunters each year, but since there has been little research in this area we elected to calculate each year's number of hunters simply as a constant proportion of the human population in each area. For

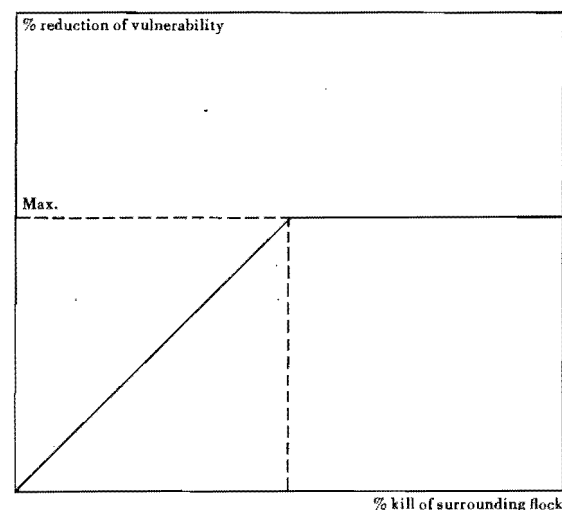
simulation of future years it was assumed that the population trends of the past 10 years will continue. We obtained the percentage of hunters in each area from a CWS survey of permit purchasers. The survey also provided the average number of days each hunter hunts. All those reporting one or more days of hunting were counted as active hunters. Multiplication of the number of active hunters by the average number of days spent hunting gives the total number of hunter-days for the year in each area.

Hunting effort is not distributed evenly throughout the season, so we assumed some fluctuation of effort during the semi-monthly hunting periods in each area. The distribution of hunting effort is important because the composition of the flock (age and sex ratios), and its distance from breeding grounds, is changing throughout the hunting season; higher mortality during one period can change these characteristics. In order to make the submodel as realistic as possible, we assumed that the opening weekend of the season would show a surge of hunting effort, but that changes in effort due to holidays and weather would cancel each other out. We simulated the opening weekend's surge of effort by arbitrarily assigning it a proportion of the year's total hunting effort. We divided the remaining effort evenly among the remaining periods. In the model, the assigned proportion for the opening weekend's effort is 0.40.

At the end of each hunting period, the bag is calculated as a proportion of the total kill, and kills are subtracted from the numbers in each age, sex and origin class. Then the vulnerability of each class is decreased as a function of the proportion killed in the flock, since the proportion killed is an indication of the number of learning experiences that the birds have had. The hypothesized relationship is shown in Figure 16 on the following page.

Figure 16
Relative vulnerability of Mallards to hunting is assumed in the model to change in relation to how many birds have already been shot

Figure 16



Virtually no data were available for the rate of change of vulnerability, so the parameters (Fig. 16) were adjusted to correspond with the results of wing survey data collected by the CWS. At the start of the hunting season there might be a juvenile to adult ratio of 1:1 in a particular area. In the first period's bag the ratio of kills will be 6:1, in the last period's bag the ratio will have become 2:1. We assumed that the maximum semi-monthly vulnerability change is 1.0 for adults and 0.5 for juveniles. This is achieved only if the kill in the surrounding flock equals or exceeds the threshold in Figure 16.

At the end of the last hunting period, winter mortality is subtracted from the remaining population. Winter survival rates take account of natural mortality for the whole year (except for the summer breeding season). Survival rates were assumed to be constant for each wintering area and to be independent of weather patterns.

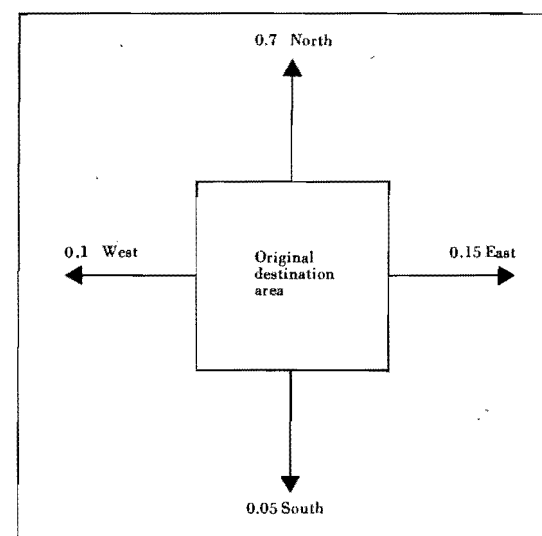
Northward migration

In this submodel, movement of ducks from wintering areas to breeding areas is simulated, allowing for homing and pioneering to occur in differing amounts for adults and juveniles. First, the breeding capacity of each management area is simulated as the maximum number of breeding females per pond for the area times the number of May ponds (provided by subroutine PONDS). Next, homing adults are moved to the areas occupied during the previous breeding season (rather than to the areas where they were born). There is a lack of data on actual homing habits. Pioneering adults are then distributed. Pioneering birds are assumed to come from two sources: non-homing birds, if any, and homing birds that are not able to establish breeding territories in their original areas of destination, i.e. where they bred the previous year. Then homing juveniles are moved to their proper areas, and finally pioneering juveniles are distributed. The result of the series of calculations is a simulated distribution of breeding and non-breeding birds at the start of the breeding season. This distribution is then used by the production submodel.

The first calculation, which finds the maximum number of breeding females that can fit into a management area, illustrates a serious deficiency in present data. Good estimates are needed of the maximum number of breeding (not breeding plus non-breeding) females that are willing to occupy a pond of average size and quality. Only rough estimates are now available from bird counts in years of various pond densities, and no reliable estimate can be made of how many of the birds counted were actually breeding. Data for 1957-1969 range from 0.16-1.3 females per May pond in Alberta to 0.09-0.49 in Manitoba (Crissey, pers. comm., 1969). These figures were found by dividing the number of birds counted in breeding areas by the number of ponds. The lack of data is critical since it concerns one of the most impor-

Figure 17
Birds that cannot find breeding space in areas to which they home are assumed in the model to move on, with varying probabilities of going in different directions as shown in this example

Figure 17



tant stages in the northward migration. Holding capacities of areas determine not only how many birds can breed in a given area, but also how many birds must move elsewhere to possibly less productive habitats. Small errors in estimates of capacity could be compounded over several simulated years, resulting in poor estimates of productivity.

In all calculations we assumed that males follow the females. Females are assumed to bring with them the number of males in accordance with the sex ratio prevailing in the areas from which the birds are moving.

In the model homing adults are moved into areas first, and the unused capacity of each area is updated after each series of movements. We assumed that all adults and 50% of the juveniles are homing, although these values can be varied. The model moves birds to a management area only if breeding space is still available.

If all females cannot be fitted into a homing area, a proportion of them is assumed to be non-breeding. We took this proportion to be 25%, although it can be varied. If the remaining 75% still cannot be accommodated, we assumed a further fraction fly elsewhere to seek breeding sites. We simulated the distribution of this fraction, plus the non-homing adults (if any), by means of a redistribution matrix. The matrix shows the probability of a bird moving from area I to area J if it cannot find a breeding site in area I. We assumed a bird is most likely to continue flying north, and least likely to return south. The way these probabilities are assigned is illustrated in Fig. 17. An implicit assumption in the calculations is that all birds redistributing to an area have an equal chance of obtaining whatever territories are still available. In other words, we assumed that ducks moving into an area from nearby have no greater chance of finding a breeding site than those coming from farther away. However, the redistribution table contains probabilities of movement that reflect distances between areas.

Results and discussion

Typical output from the Mallard model is presented in Figures 18–20. In complex simulation schemes such as that described above, it becomes a major problem to decide what variables best represent model behaviour. Figures 18–20 present only a superficial summary of the calculations. In early runs (Fig. 18) pond numbers were seriously overestimated, resulting in errors in all other predictions. Southward migration was simulated as a series of steps, and small errors in the proportions of birds moving in different directions at each step resulted in large errors in the number of birds wintering in each U.S. flyway.

In several test runs of the model, we predicted a sudden and dramatic decline in duck populations after about fifteen years (Fig. 19). Further, the model predicted that the decline should occur even if pond habitat were greatly improved at the last moment. We assumed that hunting pressure will grow geometrically at the same rate as human population, and that this increasing pressure eventually uses up all the resilience of the Mallard population to harvesting although it is unlikely that such a decline would actually occur.

More precisely, the decline indicated by the model occurred after increasing hunting pressure resulted in severe deterioration of the age structure of the duck population, so that most breeding birds were yearlings, with higher natural mortality and lower productivity per head. The sharpness of the decline was due to the combined effects of these characteristics. Similar responses have often been observed in fishery stock models (e.g. Ricker, 1963). It is unlikely that such a decline would actually occur, since hunting pressure would drop as birds became increasingly hard to find in some management areas. However, it may be of interest to managers that a sudden decline could at least start without prior warning in terms of declining productivity or kills.

Figure 18
Results from an early run of the Mallard model. Pond numbers were seriously overestimated at first, resulting in unrealistically high population size

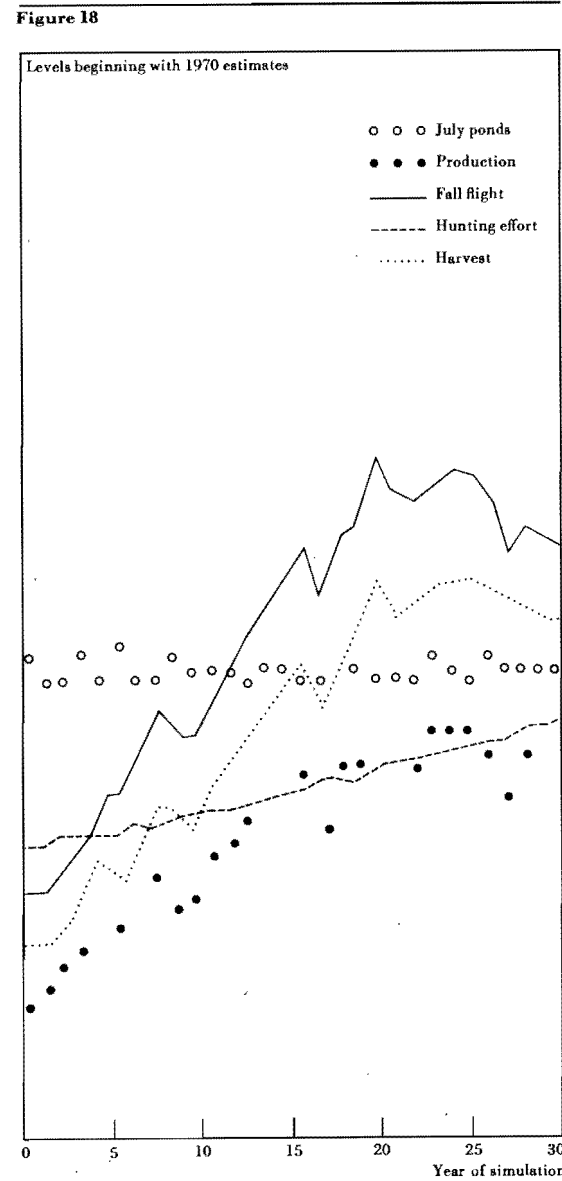


Figure 19
The Mallard simulation predicted sudden and dramatic population declines under increasing hunting pressure. Notice that harvest and production curves give no early warning of the decline

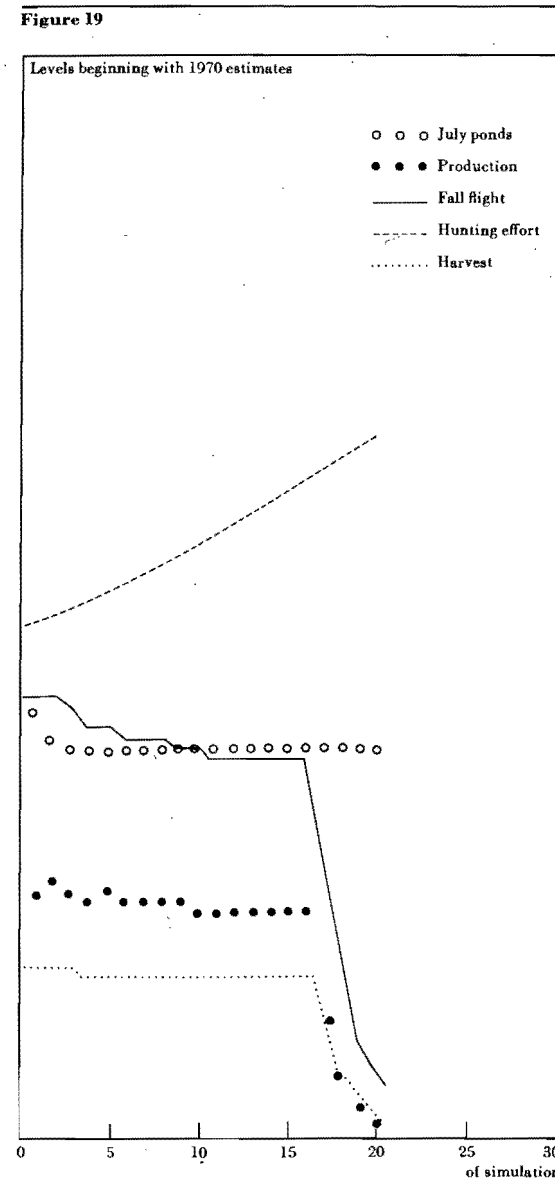
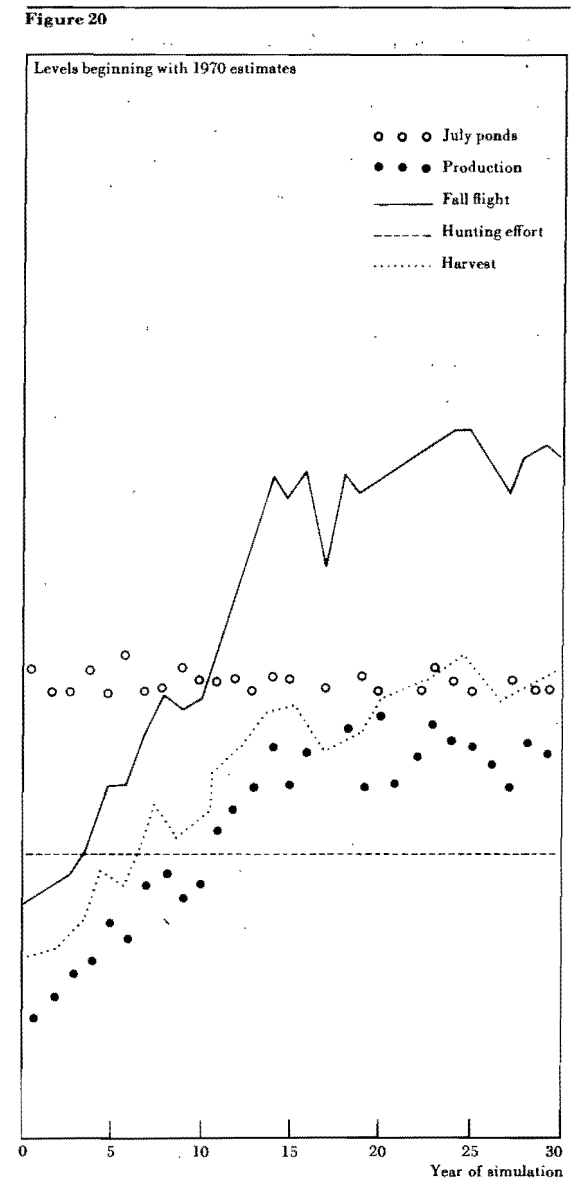


Figure 20
The Mallard model predicts that no population decline will occur if habitat is improved and hunting effort is kept at present levels



Results obtained in the workshop sessions illustrate an important principle about simulation models: increased detail in biological representation does not necessarily result in more accurate prediction. By adding detail in spatial representations, the number of parameters (such as movement rates) to be estimated becomes very large. But at the same time, the model does not necessarily become less dependent on the accuracy of any single parameter estimate. Similarly, by adding detail we make parameters more specific and harder to estimate, yet the model does not become any less dependent on some particular set of these parameters. For example, the model now gives rather poor estimates (wrong by up to 50%) of the relative number of birds wintering in each U.S. flyway. At present, each wintering population is calculated by simulating the movement of birds through several migration steps across many intermediate areas; the errors at each step have been small, but have a cumulative effect.

The model should not be expected to give accurate numerical predictions, even in the short run, because environmental conditions such as pond habitat are not predictable. Instead, the model should be used as an aid to answering broad questions of management policy and population dynamics. For example, there are a number of questions which the CWS might well wish to have answered. First, to what extent does spatial patterning and bird movement mitigate the effects of management regulations in specific areas or flyways? Second, considering normal sampling error, can poor management in one area even be detected? Third, how much resilience does the Mallard population have to bad management? Fourth, what effect would particular environmental patterns, e.g. a series of dry years, have on the population's resilience? In other words, must managers be particularly careful during certain weather sequences? Finally, what would the

effects be if hunters could travel more easily from one area to another and had better information about hunting conditions? The model cannot provide definitive answers to these questions, but it can certainly do better than simple intuition. The exercise of setting up the model to answer such questions is also a good way to gain some feel for the computer program and its operation, and a good way to sharpen and improve normal intuition.

The key value of the Mallard model is not its predictive power but its ability to act as a focus and concrete goal for research activity. The Mallard model is not merely an arbitrary mathematical or theoretical structure; it is a set of biological statements in shorthand notation. Any judgment on the validity of the model is a judgment on the validity of the biological concepts and parameters it embodies. The simulation scheme is a numerical device to fit several biological factors together simultaneously, so that each factor can be seen in relation to the others. Computers cannot operate with ambiguous instructions, so the biological factors and their relationships must be precisely defined in a simulation scheme. Models can provide a powerful tool for research organization, planning, and communication. Model-oriented research teams are more likely to succeed than teams working from a traditional biological viewpoint, simply because there is explicit numerical evaluation of each biological concept or parameter in relation to others.

As devices for research organization, the workshop and Mallard simulation model seemed fairly successful. Several major areas of data deficiency were clearly identified, and the predictive value of much data now being collected was made clear. Perhaps the most important deficiency is the lack of data on the relationship between duck abundance and hunter success, the critical link between people and birds. It seems

likely that the workshop will lead to research being concentrated on this link. Kill statistics now being collected shed little light on the abundance-success relationship, so their value for prediction is questionable. Furthermore, banding and migration data now being collected give little information on the dynamics of bird movement during the hunting season.

In the workshop, no particular premium was placed on the utilization only of past data, and little attempt was made to force the Mallard model to fit historical trends. Such an approach may seem alien to the biologist concerned with careful ecological description, and may be criticized as too theoretical. If we had been very careful to base the model only on established principles and data, and if we had made every effort to make it fit all historical population trends, two facts would soon have become apparent. First, the model would have been merely an alternative way to represent the data, and would thus have been as incomplete and as narrow in its range of application as the existing data itself. Second, since errors in data and principles can only be detected by contrasting them with other information, the model could not have pointed out errors, and would not have said anything new about the biological system.

In other words, the model would lose much of its value if it were strictly an empirical representation; failure to fit past data indicates our incomplete knowledge of the biological system, and for this reason such failure is instructive. The model is a complex set of hypotheses, which we can try to refute and improve by comparison with data; system approaches are not mutually exclusive alternatives for the "scientific method." There is an obvious extension to this reasoning: large data collections are not a prerequisite for model building, just as it is not necessary to have all the data for any experiment before the experi-

ment can be designed. Thus, we suggest that model building is likely to be most valuable in future research programs if it is included as an integral part of the work right from the start.

Literature cited

Forrester, J. 1971. Principles of systems. Wright-Allen Press, Cambridge, Mass. 178 p.

Holling, C. S. 1972. Ecological models; a status report. In Proc. International Symposium on Modelling Techniques in Water Resources Systems, Vol. 1. A. K. Biswas (ed.) Environment Canada, Ottawa.

Holling, C. S., and M. A. Goldberg. 1971. Resource science workshops. Resource Science Centre, Univ. of British Columbia. 33 p.

Jeffers, J. N. R. (ed.) 1972. Mathematical models in ecology. Blackwell Scientific Publications, Oxford. 398 p.

McAllister, C. D., R. J. LeBrasseur, and T. R. Parsons. 1972. Stability of enriched aquatic ecosystems. Science, 175:562-565.

Patten, B. C. (ed.) 1971. Systems analysis and simulation in ecology. Academic Press, Inc., New York. Vol. I. 607 p.

Ricker, W. E. 1963. Big effects from small causes: two examples from fish population dynamics. J. Fish. Res. Bd. Canada, 20:257-264.

Walters, C. J. 1971. Systems ecology: the systems approach and mathematical models in ecology. In Odum E.P., Fundamentals of Ecology, Chap. 10, W. B. Saunders Co., N.Y.

Watt, K. E. F. 1968. Ecology and resource management: a quantitative approach. McGraw-Hill Book Co., New York. 450 p.

Appendix I

FORTTRAN names and equations are the same as were used in the "Notation and Symbolism" section:

```
C Duck simulation model
C Read parameters and initial values
  READ (2, 10) NAS, NJS, KJ, KA, PA, PJ,
  SAS, SJS, SWJ, SWA
10 FORMAT (10F8.0)
C Set up loop over ten years
  DO 1 I = 1, 10
C Calculate production
  P = NJS*PJ + NAS*PA
C Calculate fall populations
  NAF = (NJS + NAS) * SAS
  NJF = P * SJS
C Calculate harvest
  H = NJF * KJ + NAF * KA
C Calculate new spring numbers from survivors
  of hunting
  NJS = NJF * (1 - KJ) * SWJ
  NAS = NAF * (1 - KA) * SWA
C Write out harvest and new spring
  populations for year
  WRITE (3, 11) I, NJS, NAS, H
11 FORMAT ('YEAR = ', I2, ' JUVENILES = ',
  F10.2,
  'ADULTS = ', F10.2, 'HARVEST = ',
  F10.2),
C Return for start of next year.
1 CONTINUE
STOP
END
```

The program shown above is not quite complete: not shown are system control cards and the data card to read in parameter values and initial conditions.

Other publications in the Occasional Papers Series

No. 1
Birds protected in Canada under the Migratory Birds Convention Act. 4th ed. Also available in French. Cat. No. R69-1/1

No. 2
Canadian bird names, French, English and scientific. Bilingual publication. Cat. No. R69-1/2

No. 3
Use of aerial surveys by the Canadian Wildlife Service by D. A. Benson. Out of print. Cat. No. R69-1/3

No. 4
Queen Elizabeth Islands game survey, 1961 by J. S. Tener. Cat. No. R69-1/4

No. 5
Age determination in the polar bear by T. H. Manning. Cat. No. R69-1/5

No. 6
A wildlife biologist looks at sampling, data processing and computers by D. A. Benson. Out of print. Cat. No. R69-1/6

No. 7
Preliminary report on the effects of phosphamidon on bird populations in New Brunswick by C. D. Fowle. Out of print. Cat. No. R60-1/7

No. 8
Birds of the Nova Scotia-New Brunswick border region by G. F. Boyer. Cat. No. R69-1/8

No. 9
Effects of dietary methylmercury on Ring-necked Pheasants, with special reference to reproduction by N. Fimreite. Cat. No. R69-1/9

No. 10
Trends in populations of barren-ground caribou over the last two decades: a re-evaluation of the evidence by G. R. Parker. Cat. No. R69-1/10

No. 11
The Canada migratory game bird hunting permit and related surveys by D. A. Benson. Cat. No. R69-1/11

No. 12
Observations on duck hunting in eastern Canada in 1968 and 1969 by H. J. Boyd. Cat. No. R69-1/12

No. 13
Evaluation of ecological effects of recent low water levels in the Peace-Athabasca Delta by H. J. Dirschl. Cat. No. CW69-1/13

No. 14
The Great Cormorants of eastern Canada by A. J. Erskine. Cat. No. CW69-1/14

No. 15
Distribution of barren-ground caribou harvest in northcentral Canada by G. R. Parker. Cat. No. CW69-1/15

No. 16
Bird migration forecasts for military air operations by H. Blokpoel. Cat. No. CW69-1/16

No. 17
Waterfowl populations on the Peace-Athabasca Delta, 1969 and 1970 by D. J. Nieman and H. J. Dirschl. Cat. No. CW69-1/17

No. 18
Gammarus predation and the possible effects of Gammarus and Chaoborus feeding on the zooplankton composition in some small lakes and ponds in western Canada by R. S. Anderson and L. G. Raasveldt. Cat. No. CW69-1/18

No. 19
A summary of DDE and PCB determinations in Canadian birds, 1969 to 1972 by M. Gilbertson and L. Reynolds. Cat. No. CW69-1/19

