

8.2 Population Development Models

S.A. WARD, R. RABBINGE and A.F.G. DIXON

INTRODUCTION

This section introduces the types of mathematical model available for use in studies of aphid population biology, shows when and how they may be used, and discusses their merits and limitations. First, however, the function of modelling should be considered.

Most research can be broadly classified as “scientific”, which seeks explanations for observable phenomena, or “technological”, which seeks methods for manipulating real “systems” (Jeger, 1983). [In this section, a system is defined as a limited part of reality with well defined limits which are dictated by the objectives of the model.] Both types require prediction: the scientific because explanations (and predictions, which are logically equivalent) are developed, ideally, by a cycle of prediction and falsification; the technological because manipulative decisions must be based on their predicted consequences. Scientific studies of complex biological systems may also require a framework within which new observations can be interpreted. Neither frameworks nor predictions can be easily provided without the use of models: “substitutes”, capable of replicating the important features of systems’ behaviour, but much more rapidly.

The role of modelling is clearly more far-reaching than that of a distinct component in a research programme. In fact, although much of this section is inevitably concerned with the techniques involved, “modelling” can be considered more as the expression of an attitude of mind than as a specific tool (although the resulting model may have a very precise function). It is difficult, therefore, to give general rules on why and how models are developed. Three points should now be made.

First, substitute systems can be formulated either verbally or mathematically. Where possible, the latter is preferable, since it permits much more precise (and therefore falsifiable) predictions to be made.

Secondly, a model in biological sciences is a representation of only the most “important” features of the system – a perfect replica would behave as slowly as the system itself. The considerable simplifications in its construction mean that no model is completely true; it has only more or less accuracy, generality, and realism (an expression of its structural similarity with reality) (Conway and Murdie, 1972).

Finally, modelling is not, or should not be, an end in itself. A model can be judged only in terms of its success as a tool for the job in hand. In “science”, this job may be to suggest general laws or syntheses, to stimulate empirical research in important areas, to test our knowledge and insights for consistency and completeness, or to rule out suggested explanations. For “technological”

use, a model needs only to be accurate; realism or explanatory power are valuable only if they improve this (although in fact they often do).

We will now consider the various types of mathematical models.

DESCRIPTIVE MODELS

Attempts to describe the behaviour of a system as a whole must often rely on purely descriptive models of the relations between distantly connected variables. Two examples will illustrate the scientific and technological application of these "statistical" models.

First, in a study of the lime aphid, *Eucallipterus tiliæ* (Linnaeus), Barlow and Dixon (1980) showed that the peak density of oviparae in autumn could be related to the maximum population of fundatrices in the preceding spring, by the regression equation

$$\log(\text{peak oviparae})_t = 2.43 - 0.60 \log(\text{peak fundatrices})_t \quad (1)$$

(where t refers to the year); and that

$$\log(\text{peak fundatrices})_{t+1} = 0.66 + 0.78 \log(\text{peak oviparae})_t \quad (2)$$

In this scientific context, the value of such a simple description of the annual population changes is two-fold. First, it reveals a pattern at the level of the whole system (in this case the lime tree and the aphid population), and demonstrates, for example, that the aphid population is regulated in a density-dependent way. (This does not mean that there is a causal relation between the two densities: this remains to be analysed with reference to the basic biological processes.) Secondly, it narrows down the field of search for the mechanisms of population regulation: only in the summer generations is the rate of increase reduced at high densities.

A technological example of the use of descriptive models is the work of Walters (1983). Seeking a method for making long-term predictions of outbreaks of the cereal aphid, *Sitobion avenae* (Fabricius), he showed that the timing and size of the population peak could be predicted from the mean temperature in January.

In science, therefore, purely descriptive models reveal a pattern and suggest questions; in technology they allow a degree of prediction without requiring an understanding of the detailed working of the system.

The limitations of such models stem precisely from this lack of understanding. Since the causal mechanisms relating the variables are not included in the model, it cannot be used to explain failures in prediction. After suggesting the initial questions, it cannot be used later in a scientific study, except perhaps to check that hypothesized mechanisms *are* able to produce the observed whole-system behaviour (e.g. Barlow and Dixon, 1980). For technological use, such unrealistic models may fail in that they cannot be extrapolated to new circumstances, and that their predictions may therefore be unreliable: e.g., unusually high densities of natural enemies may greatly influence both size and timing of peak aphid populations (Chambers et al., 1982). In addition, predictions such as those of Walters cannot easily be used in the management of individual fields, as they concern the *average* conditions in a large region.

In terms of the properties mentioned in the introduction, such models are unrealistic, lack generality, and may be inaccurate. Ideally, their use should be restricted to the opening stages of an investigation, when too little is known about the system to permit the inclusion of all important relationships. During these opening stages, however, purely descriptive models do often

prove invaluable; and, given the limited information available on many aphid species, are likely to do so for some time to come.

ANALYTICAL EXPLANATORY MODELS

This section discusses a wide range of models with only one common feature: they incorporate some of the causal relations in the system, but remain simple enough to be analytically tractable. Their generality is their major strength, but also the main limitation to their ability to explain real systems. The main function of such models is the generation of hypotheses about general ecological principles. Since their only common characteristic is their tractability, they can be discussed only by means of examples illustrating their use in suggesting principles.

The simplest models with any claim to realism are those describing the changes in single-species populations. Of these, perhaps the best-known is the logistic equation:

$$dN/dt = rN(1 - N/K) \quad (3)$$

This states that the rate of increase of the population depends on the current size of the population (N) and describes the per capita rate of increase as a declining linear function of density. The terms r and K now have a biological meaning: r is the per capita rate of increase when N is low, i.e. in the absence of competition, and K is the “carrying capacity” of the environment, the density at which the population is maintained at a constant level. While r and K are both summaries of the effects of a complex array of causal factors, their inclusion in the model represents a considerable advance over holistic descriptive models (e.g. equation (1)). Both r and K can be measured or estimated separately – they need not be estimated by curve-fitting – and from these independent measurements it is possible to predict not only the relation between “initial” and “final” density, but also the temporal changes in population. In addition, examination of the related model

$$N_{t+1} = N_t \exp(r - rN_t/K) \quad (4)$$

(May, 1974), suggests properties of general ecological systems. This simple model can be used to study the behaviour of a population perturbed from its equilibrium (K): convergence to K ; convergent oscillations; limit cycles; divergent oscillations; divergence or chaos.

Though useful for suggesting general laws, these analytical whole-system models make too many simplifying assumptions to describe any real system. Perhaps, for this reason, “many practical ecologists both past and present have spent a lot of time looking for evidence of population effects whose existence is mere conjecture” (Hughes, 1972). Whether you regard this remark as pejorative or not depends on whether your priorities are scientific or technological.

The second group of whole models in this category are, at least from the population biologist’s viewpoint, models of sub-systems. If a model’s realism is to be measured by the level at which it becomes purely descriptive, these are far more realistic than any generalized logistic equations. The example here is the set of equations describing the effects of predation on a prey population (for detailed discussions, see Hassell et al., 1976; Hassell, 1978).

The initial Nicholson–Bailey equations (Nicholson and Bailey, 1935) assume that the rate at which prey is captured is proportional to the density of prey and that of predators. The capture rate equation was then extended (Holling, 1959; Rogers, 1972) to incorporate the fact that at high prey densities predation is

limited by the "handling time" required to consume a prey item. Failures in prediction stimulated further experimental work, which was incorporated into the model: e.g. "mutual interference" between predators or parasites (Hassell, 1971) and the changes in search rate and handling time as the predator becomes satiated (Mills, 1982).

From a strictly scientific point of view, these models have proved fairly successful as catalysts for, and frameworks for an understanding of, increasingly detailed empirical work. However, although this body of theory has succeeded in expanding "downwards" into ethology and physiology, it has failed to percolate "upwards" into whole-system ecology. Its use in technology has therefore been restricted to suggestions of criteria for biological control agents. The main reason for this is that though we can now understand reasonably well what happens when a predator enters a "patch" of prey, we cannot predict how often this will occur without relying on difficult field observations (the fact that we now realize the importance of this information gap illustrates again the scientific value of these models).

Returning to the subject of analytical models in general, a very serious difficulty now arises. Increasing realism requires the introduction of an increasing number of expressions and parameters. This parameter proliferation means that, even if analytical solutions can be found, e.g. for conditions, they may be so complex as to be indecipherable, or, at least, useless. The attempt for generality must be abandoned. Between adjacent levels of organization, general models are extremely useful guides for research; but attempts to combine generality with a reasonable degree of realism are doomed to failure, serving only to illustrate the "physics-envy" that is "the curse of biology" (Cohen, cited in Oster and Wilson, 1978).

If we are to understand or manipulate an ecological system, we require accuracy and realism. (As mentioned above, although technology *needs* only accuracy, this can rarely be achieved without realism.) It is at this point that analytical techniques become useless: any natural system is so complex that a realistic analytical model is intractable; while any tractable model must ignore so many influences on the population that it is quite unrealistic.

NUMERICAL SIMULATION MODELS

The development of powerful computers removed the need for mathematical tractability, and therefore also for the drastic oversimplification of whole-system models. It became possible to construct models incorporating a vastly greater range of causal mechanisms, and to use them for either scientific or technological work. These models could be based on a "bottom upwards" approach: the synthesis of detailed quantitative information about a wide range of physiological processes, for example.

Since, in contrast with analytical models, numerical simulation models share many common principles, this section first considers some general rules before discussing the models' strengths and limitations. The basic procedure for construction of such a model involves the following sequence of steps (see Rabbinge, 1976; Jeger, 1983; Rabbinge and Carter, 1983).

(1) The system must be defined (e.g. the aphid population, the host plant, and the predators or parasites of the aphids), with its main components and relationships. It is often convenient to express this "qualitative model" using relational diagrams. Fig. 8.2.1 illustrates an example of such a diagram (Rabbinge et al., 1979).

The system is described in terms of three types of variables, and their interrelations. "State variables" are those which are, in theory, measurable at any

the analytically derived disc equation (Holling, 1959) in the predation sub-model.)

(3) The parameters of the various relationships must now be quantified with reference to experimental data. At this stage, the modelling effort can already highlight areas of ignorance, i.e. parameters yet to be measured empirically; however, the assigning of *priorities* is not possible until a preliminary computer model is available for sensitivity analysis (see below).

(4) The model is now written in the form of a computer program. The program proceeds iteratively by reading the values of the driving variables calculating the rate variables, and updating the values of the state variables. A very important point must be made here. Since continuous processes are being approximated by changes during discrete time-steps, the length of the time-step can have a considerable effect on the model's outcomes. If, for example, the daily mortality is 50%, the use of a three-day time-step would give 150% mortality. Clearly, long time-steps can result in gross inaccuracy. Use of a very short time-step, however, greatly increases the computation time required for the simulation. Two approaches have been used here. In one important group of models, rate variables are integrated not over real time, but over "physiological time", expressed in units of day-degrees above the temperature threshold for development, or of fractions of the duration of an instar (Hughes and Gilbert, 1968; Gilbert and Hughes, 1971; Gilbert et al., 1976). This method has proved invaluable in modelling systems whose rate variables increase linearly with temperature, but is more difficult to apply if non-linear relations (e.g. predation rates) are important, or where different processes have different temperature thresholds. In this case, it is preferable to integrate over real time (Rabbinge et al., 1979; Barlow and Dixon, 1980; Carter et al., 1982). Here the appropriate step length depends on the smallest "time coefficient" in the model: if the rate of change of a state variable, x (e.g. aphid density), is

$$dx/dt = kx$$

then the "rate coefficient" is k , and the time coefficient is $1/k$; the longest acceptable time-step is $1/4$ of the time coefficient, or, in this case, $1/4k^{-1}$. (For more detailed discussion of this subject, see De Wit and Goudriaan, 1978; Rabbinge, 1976.) In relatively simple models, the time-step can often be calculated directly, but in large complex models it is sometimes necessary to proceed by trial and error (Hughes and Gilbert, 1968; Rabbinge and Carter, 1983).

(5) The model produced must now be "verified", before being tested against real data (Jeger, 1983). This involves checking that it operates as intended: that the computer program is internally consistent and corresponds to the model formulation; and that the model itself does not contain inconsistencies.

(6) "Validation" is the comparison of the program's output with real data. It is important that not only the whole model but also its component sub-models be validated. If errors in sub-models counterbalance each other, the predictions of the whole model may be accurate in the conditions tested but fail under different circumstances.

(7) Preliminary testing usually reveals that the model's predictions are inaccurate. If the aims were scientific this is a mark of success, and should mean that the model is functioning in its capacity as a guide for further study. An example of this is the work of Watson and Carter (1983), who showed that the usually successful model of Carter et al. (1982) greatly overestimated populations of *S. avenae* if there is heavy rainfall throughout the spring colonization period. This directed attention towards a previously unresearched topic – mortality due to rain – and illustrates clearly one of the scientific uses of numerical simulation models. For the technologist, however, inaccuracy is failure; the model must be revised before it can become useful. This revision

can be accomplished in two ways: by re-examining the model's structure and parameters with reference to detailed empirical research, or by juggling parameters until the predictions are in reasonable agreement with whole-system data. In the latter case, unless the newly estimated parameters are checked experimentally the process has become a cumbersome curve-fitting exercise: while possibly valuable to the technologist, the model has lost its claim to being explanatory.

A technique of value in suggesting which of the many parameters should be re-examined is sensitivity analysis. Here, model parameters are varied, or relationships altered, to determine their effect on the system's behaviour. Thus, only the parameters affecting the inaccurate aspects of the predictions need to be investigated. In addition, this technique reveals the degree of accuracy required in measurements of field variables. For example, if peak aphid density is highly sensitive to parasite density, but barely influenced by the level of generalist predators, parasite density must be measured with greater accuracy.

The scientific value of these numerical simulation models can now be summarized. First, they provide a synthesis of an otherwise bewildering and unmanageable mass of observational data, and a framework for understanding particular experimental results. Secondly, by means of sensitivity analyses, they can be used to rule out suggested explanations for population regulation or cyclical fluctuations (Van den Bos and Rabbinge, 1976); or they can be used to provide preliminary estimates, e.g. of the efficiency of predation or parasitism in the field (see section 9.2). They can also be used to narrow down the search for explanations, to suggest priorities for applied research (e.g. which "resistance" factors should be selected by plant breeders (Carter and Dixon, 1981)) and to pinpoint important gaps in current knowledge: they act as guides for empirical research. For the same reason (their structural similarity with the real system), they can easily be extended to include new experimental results within the existing sub-models, or to include additional sub-systems (e.g. plant growth simulators). Finally, since they need not be tractable or elegant (analytically), they can be used to examine systems with stochastic components, or with complex spatial variation (Rabbinge et al., 1984). The main drawback to this approach is that it is time-consuming: in terms of both the real time required to create the synthesis, and the expensive computer time needed, particularly where stochasticity is to be incorporated.

It is this high cost that limits the technological value of models designed for complete explanation. They can at this stage be used to suggest priorities for purely applied research, but before they can be used as a basis for practical decisions, they must be simplified.

SUMMARY MODELS AND THEIR USE IN TECHNOLOGY

The first step in the production of a summary model is sensitivity analysis. As described above, this reveals the effects of various parameters and sub-models on features of the system's behaviour. It also, therefore, shows which parts of the model are unimportant in determining the results of greatest interest (often the yield of a crop). After these components have been excised, descriptive functions are sought, to replace complex causal networks (for a detailed example, see De Wit et al. 1978). It should be noted that this is not simply a return to the purely descriptive modelling discussed above, which often involves description of the most readily available data; here we are describing the relationships we know to be important.

When this stage has been reached, the model can be used in management decision-making: it is now a relatively inexpensive substitute system, which

does in minutes what the real system does in months. In passing, two types of management decision should be distinguished: strategic and tactical. Strategic decisions concern long-term policy, including, e.g., crop rotation schemes. Inevitably, however, management of highly migratory colonizing species, such as aphids, must rely mainly on shorter-term "tactical" decisions: the recommendations of the EIPRE system (Rabbinge and Rijdsdijk, 1983) are based on predictions of cereal pest levels approximately a week in advance. In this respect, aphids are similar to fungal epidemics; only the rate of increase of the epidemic and its maximum are of practical importance. In cereal aphids it has been demonstrated that the upsurge and the timing of the peak depend heavily on crop development and crop condition (Rabbinge et al., 1979; Carter et al., 1982), so in the simplified summary model, crop development drives the epidemic. This yields predictions of cereal aphid population changes that are now widely used in a supervised control system for pests and diseases of wheat (Rabbinge and Carter, 1983).

For strategic or tactical planning, summary models are used in basically the same way: to predict the outcomes of alternative decisions. Here, the model's function as a substitute system is apparent. Once a goal has been identified – maximization of expected yield or minimization of the risk of financial loss (Rossing, 1983) – the optimal decision can be determined by direct comparison of model outputs, or, if a range or sequence of decisions is available, by deterministic or stochastic dynamic programming (Wagner, 1969; Van Beek and Hendriks, 1983; Onstad and Rabbinge, 1985).

DISCUSSION AND CONCLUSIONS

This section has considered a range of techniques for producing substitutes for reality. The techniques vary according to the degree of realism of the resulting models. The simplest statistical models are only able to describe the main features of the behaviour of a system and are often inaccurate if conditions change. These models are, however, a useful first step in attempting to understand the system, and can often be used to determine the need for various measures to control aphid populations.

Analytical explanatory models incorporate more causal relations in the system and are invaluable as a means of suggesting general ecological rules. In order to remain analytically tractable, however, they require considerable simplification. Such models cannot, therefore, be used to understand the detailed functioning of any real system, so their use in technological research is confined to the derivation of analytical "short cuts" for numerical models. To understand or predict the detailed behaviour of complex systems, the desire for generality, tractability or elegance must be abandoned. Instead, numerical simulation models can be used to synthesize information and explain diverse aspects of the behaviour of particular systems.

The final stage in technological modelling (i.e. that directed towards manipulation of the system) is the construction of a summary model. This approximates the output of the explanatory simulation model, but requires less computer time, so it can be used repeatedly for optimization of management practices.

The main virtues of all these types of substitute are that they work much faster than the real system, so they can be more easily used for experiments (scientific or technological), and that failures in correspondence between the model and the system expose ignorance. Uncritical use of such a substitute, however, carries severe penalties. No model is a perfect replica, so, although a model is a substitute for reality, modelling can never be a

substitute for empiricism. Instead, the two pursuits should complement one another, each stimulating and redirecting the other.

REFERENCES

- Barlow, N.D. and Dixon, A.F.G., 1980. Simulation of Lime Aphid Population Dynamics. Simulation Monographs. Pudoc, Wageningen, 171 pp.
- Carter, N. and Dixon, A. F. G., 1981. The use of insect population simulation models in breeding for resistance. Bulletin IOBC/WPRS, 1981/IV/1: 21–24.
- Carter, N., Dixon, A. F. G. and Rabbinge, R., 1982. Cereal Aphid Populations: Biology, Simulation and Prediction. Simulation Monographs. Pudoc, Wageningen, 91 pp.
- Chambers, R.J., Sunderland, K.D., Stacey, D.L. and Wyatt, I. J., 1982. A survey of cereal aphids and their natural enemies in winter wheat in 1980. Annals of Applied Biology, 101: 175–177.
- Conway, G.R. and Murdie, G., 1972. Population models as a basis for pest control. In: J.N.R. Jeffers (Editor), Mathematical Models in Ecology. Blackwell, Oxford, pp. 195–213.
- De Wit, C.T. and Goudriaan, J., 1978. Simulation of Ecological Processes. Simulation Monographs. Pudoc, Wageningen, 2nd edn., 175 pp.
- De Wit, C.T. et al., 1978. Simulation of Assimilation, Respiration and Transpiration of Crops. Simulation Monographs. Pudoc, Wageningen, 141 pp.
- Gilbert, N. and Hughes, R.D., 1971. A model of an aphid population – three adventures. Journal of Animal Ecology, 40: 525–534.
- Gilbert, N., Gutierrez, A.P., Frazer, B.D. and Jones, R.E., 1976. Ecological Relationships. Freeman, Reading, 156 pp.
- Hassell, M.P., 1971. Mutual interference between searching insect parasites. Journal of Animal Ecology, 40: 473–486.
- Hassell, M.P., 1978. The dynamics of arthropod predator–prey systems. Princeton University Press, Princeton, NJ, 237 pp.
- Hassell, M.P., Lawton, J.H. and Beddington, J.R., 1976. The components of arthropod predation. I. The prey death rate. Journal of Animal Ecology, 45: 135–164.
- Holling, C.S., 1959. Some characteristics of simple types of predation and parasitism. Canadian Entomologist, 91: 385–398.
- Hughes, R.D., 1972. Computer simulations of aphid populations. In: A.D. Lowe (Editor), Perspectives in Aphid Biology. New Zealand Entomological Society, Bulletin No. 2, pp. 85–91.
- Hughes, R.D. and Gilbert, N., 1968. A model of an aphid population – a general statement. Journal of Animal Ecology, 37: 553–563.
- Jeger, M.J., 1983. Mathematical models: their nature and development to practical ends in crop protection. Bulletin IOBC/WPRS, 1983/VI/2: 3–17.
- May, R.M., 1974. Stability and Complexity in Model Ecosystems. Princeton University Press, Princeton, NJ, 2nd edn., 265 pp.
- Mills, N.J., 1982. Satiation and the functional response: a test of a new model. Ecological Entomology, 7: 305–315.
- Nicholson, A.J. and Bailey, V.A., 1935. The balance of animal populations. Part I. Proceedings of the Zoological Society of London, 1935: 551–598.
- Onstad, D. and Rabbinge, R., 1985. Dynamic programming and the computation of economic injury levels for crop disease control. Agricultural Systems, 18: 207–226.
- Oster, G.F. and Wilson, E.O., 1978. Caste and ecology in the social insects. Princeton University Press, Princeton, NJ.
- Rabbinge, R., 1976. Biological Control of Fruit-tree Red Spider Mite. Simulation Monographs. Pudoc, Wageningen. 228 pp.
- Rabbinge, R. and Carter, N., 1983. Application of simulation models in the epidemiology of pests and diseases; an introductory review. Bulletin IOBC/WRPS, 1983/VI/2: 18–30.
- Rabbinge, R. and Rijdsdijk, F.H., 1983. EPIPRE: a disease and pest management for winter wheat taking account of micrometeorological factors. EPPO Bulletin, 13: 297–305.
- Rabbinge, R., Ankersmit, G.W. and Pak, G.A., 1979. Epidemiology and simulation of population development of *Sitobion avenae* in winter wheat. Netherlands Journal of Plant Pathology, 85: 197–220.
- Rabbinge, R., Kroon, A.G. and Driessen, M., 1984. Consequences of clustering in parasite-host relations of the cereal aphid *Sitobion avenae*, a simulation study. Synopsis. Netherlands Journal of Agricultural Sciences, 32: 237–239.
- Rogers, D.J., 1972. Random search and insect population models. Journal of Animal Ecology, 41: 369–383.
- Rossing, W.A.H., 1983. A stochastic dynamic programming model for optimizing control of aphids in winter wheat, with a proposal for the quantification of risk. Internal Report, Vakgroep Theoretische Teeltkunde, Landbouwhogeschool, Wageningen, 36 pp (in Dutch).

-
- Van Beek, P. and Hendriks, Th.H.B., 1983. Optimization Techniques: Principles and Applications. Bohn, Scheltema and Holkema, Utrecht, 277 pp. (in Dutch).
- Van den Bos, J. and Rabbinge, R., 1976. Simulation of the Fluctuations of the Grey Larch Bud Moth. Simulation Monographs. Pudoc, Wageningen, 83 pp.
- Wagner, H.M., 1969. Principles of Operations Research: with Applications to Managerial Decisions. Prentice-Hall, Englewood Cliffs, NJ, 937 pp.
- Walters, K.F.A., 1983. Ph.D. thesis, School of Biological Sciences, University of East Anglia, Norwich.
- Watson, S.J. and Carter, N., 1983. Weather and modelling cereal aphid populations in Norfolk (UK). EPPO Bulletin, 13: 223-227.