

## **OCCASIONAL PAPER SERIES NO 6 – JULY 2010**

### **MODELS AND LIMITS TO PREDICTABILITY<sup>1</sup>**

**BY  
OSCAR GARCIA**

---

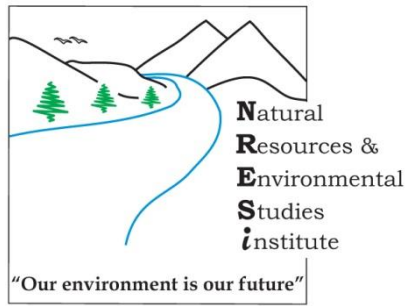
<sup>1</sup>Based on “Dimensionalidad en los modelos de crecimiento”, *Cuadernos de la Sociedad Española de Ciencias Forestales* 23, 19-25, 2007. Materials reproduced with permission

Dr. García is a Professor and FRBC / West Fraser Endowed Chair in Forest Growth and Yield within the Ecosystem Science and Management Program, University of Northern British Columbia, 3333 University Way, Prince George, B.C., V2N 4Z9 Canada.

The correct citation for this paper is:

**García, O. 2010. Models and limits to predictability. Natural Resources and Environmental Studies Institute Occasional Paper No. 6, University of Northern British Columbia, Prince George, B.C., Canada.**

This paper can be downloaded without charge from <http://www.unbc.ca/nres/occasional.html>



*The Natural Resources and Environmental Studies Institute (NRES Institute) is a formal association of UNBC faculty and affiliates that promotes integrative research to address natural resource systems and human uses of the environment, including issues pertinent to northern regions.*

*Founded on and governed by the strengths of its members, the NRES Institute creates collaborative opportunities for researchers to work on complex problems and disseminate results. The NRES Institute serves to extend associations among researchers, resource managers, representatives of governments and industry, communities, and First Nations. These alliances are necessary to integrate research into management, and to keep research relevant and applicable to problems that require innovative solutions.*

For more information about NRESI contact:

Natural Resources and Environmental Studies Institute  
University of Northern British Columbia  
3333 University Way  
Prince George, BC Canada  
V2N 4Z9

Phone: 250-960-5288  
Email: [nresi@unbc.ca](mailto:nresi@unbc.ca)  
URL: [www.unbc.ca/nres](http://www.unbc.ca/nres)

**CONTENTS**

Abstract .....2

Introduction.....3

Models.....3

Dynamic Forest Growth Models .....4

Aggregation .....5

Limits to predictability.....7

Conclusions.....10

References.....11

## **Abstract**

Models of a same system may differ greatly in scale and level of detail. The implications of this are examined, in general and more specifically in relation to forest growth models. The nature of modelling is discussed, distinguishing descriptive and predictive models, and briefly describing the concepts of dynamical system and state space. Through examples, I demonstrate limits to

predictability that can make reliable predictions impossible at the individual level. A complete understanding of the functioning of a system, or its computer simulation, do not imply being able to predict its behaviour. Although detailed models are useful for research purposes, low-dimensional aggregated models are generally more appropriate for decision-making.

## Introduction

Growth models are useful as research tools, and for predicting outcomes in forest management. Their level of detail, scale, or state space dimensionality, varies within wide limits. We shall examine some characteristics and consequences of the scale used. Much of the discussion is applicable to models in general, not just to growth models.

## Models

A model (or theory) is a partial representation of some aspect of “reality”. For instance, a scale model of a building or of an aircraft (material models), the manual for a DVD player (verbal model), or our subjective idea of the results of acting in a certain way in a given situation (mental model). Mathematical models are like verbal ones, but using mathematical language. They generally have the advantage of being less ambiguous. Perhaps more importantly, mathematical models allow the reuse of known recipes, rules, or previously established theorems, instead of having to reason starting from scratch each time. For making any decision it is necessary to employ some kind of model; it is the link that connects actions with consequences.

*“An engineer thinks that his equations are an approximation to reality. A physicist thinks that reality is an approximation to his equations. A mathematician doesn't care.”*

*Anonymous*

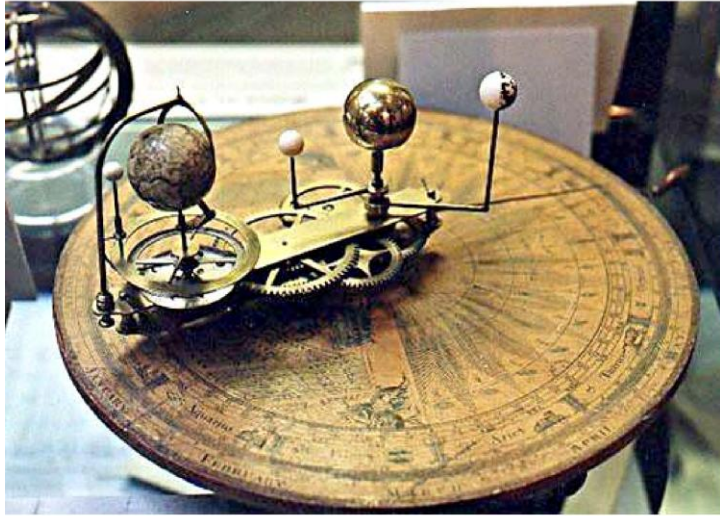
Although obviously a caricature, this quote reflects certain attitudes regarding the perception of models and theories. The last part is not particularly relevant here, it relates to the role of the mathematician, as such, in the study of formal relationships regardless of what they might represent (i.e. the development of the ‘recipes’ previously mentioned). There appear to be, however, differences in the way of thinking about the nature of models that may be associated with different experiences and traditions. One often talks of scientists

discovering natural laws, with the implication that such models have an independent existence. On the other hand, one can argue that humans can not comprehend all the details of a reality in which everything is connected to everything else; they can only reason through representations based on artificial classifications and definitions, ignoring the less important interactions. Models would have more to do with the structure of the human brain than with the reality “out there” (assuming that such reality exists!). At any rate, at least for the type of models in which we are interested, it seems more appropriate to think in this way, in which it makes no sense to speak of a true model. Quoting G. E. P. Box slightly out of context, “All models are wrong, but some are useful.”<sup>2</sup>

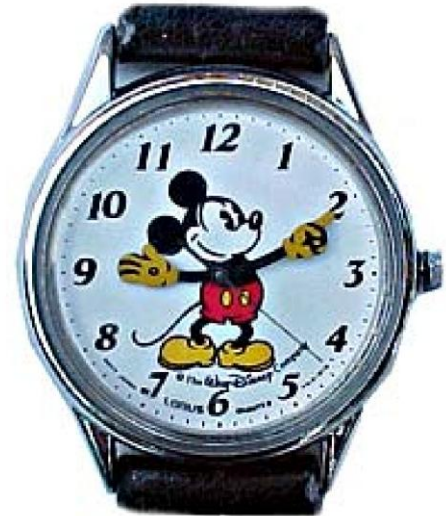
Useful for what? There is controversy about the superiority of various types of models (see for instance the special issue edited by Mohren and Burkhart 1994). Part of the problem is that the criteria for what makes a model superior are not clearly specified. One of the dichotomies is the use of models either as research tools, or as management tools. In research, the use of models is primarily descriptive; we are interested in understanding the functioning of a system, in synthesizing previously isolated facts, and in generating questions to guide future studies. The model is a working hypothesis, and major advances are achieved when the hypothesis fails. In management, the purpose is primarily predictive, to predict future behaviour, possibly in response to alternative treatments or events. Precision and accuracy in forecasting takes precedence over qualitative explanation.

---

<sup>2</sup>Economists have yet another view of models: they speak of “market failures”, it is reality which is wrong.



(a)



(b)

Figure 1: Terrestrial motion models: (a) Mechanistic, causal. (b) Empirical.

For example, Figure 1(a) is a mechanistic model that helps to understand the motion of the planets. It could be used to predict when it will become dark today. However, the empirical model of Figure 1(b), which has little or nothing to do with the functioning of the solar system, may be more accurate and convenient. Note in passing that the size and distance proportions in Figure 1(a) are far from real, and precisely that makes the model more understandable; realism in a model is not necessarily a virtue.

As explained later, another aspect related to model use is the most appropriate level of detail or complexity, or the dimensionality of the state space in dynamic models.

## Dynamic Forest Growth Models

The most traditional forest stand models are yield tables. These describe how volumes per hectare, mean diameter, dominant height, number of trees, or other variables, change as functions of time. Sufficient in many situations, their use is more problematic when there are silvicultural stand thinnings, or in other cases where a stand has deviated from the nominal trajectory provided in the table. Isaac Newton introduced an approach more flexible than

modelling functions of time directly: a dynamic model describes the rate of change of certain variables that define the state of the system. The state trajectories are synthesized by integration or accumulation of the rates of change, represented by differential or difference equations (Figure 2). Other variables of interest can be estimated from the state at any given time (Luenberger 1979, Garcia 1994, Garcia 2005b).

A system can be represented by dynamic models with various levels of detail or resolution, which differ in the dimension of the state vector (number of state variables). In particular, the standard classification of growth models (Goulding 1972, Munro 1974, Vanclay 1994) reflects differences in dimensionality (Garcia 1988). Whole-stand models use a few aggregate variables such as basal area, dominant height, number of trees. Individual-tree models, on the contrary, use tens to thousands of state variables, including sizes for each of the trees in a stand or sample plot. Individual-tree models may be distance-dependent, which use tree coordinates to calculate competition indices (Staeble 1951, Newnham 1964, Mitchell 1975), or distance-

independent, which ignore spatial structure (Goulding 1972, Stage 1973).

## Aggregation

With greater availability of computing power, there has been an increasing emphasis on the development of individual-tree models. There is a tendency to regard aggregation at the stand level as unnecessary and obsolete. On the other hand, self-thinning theories, which can be

interpreted as two-dimensional dynamic models, rather imprecise for managed stands (Garcia 2005a), remain popular. But models of intermediate dimensionality are now rare in the research literature, although still widely used in forest management practice. Something similar has happened more recently in population ecology modelling (e.g. Grimm 1999), and in other areas such as physics, economics, and sociology (Garcia 2001).

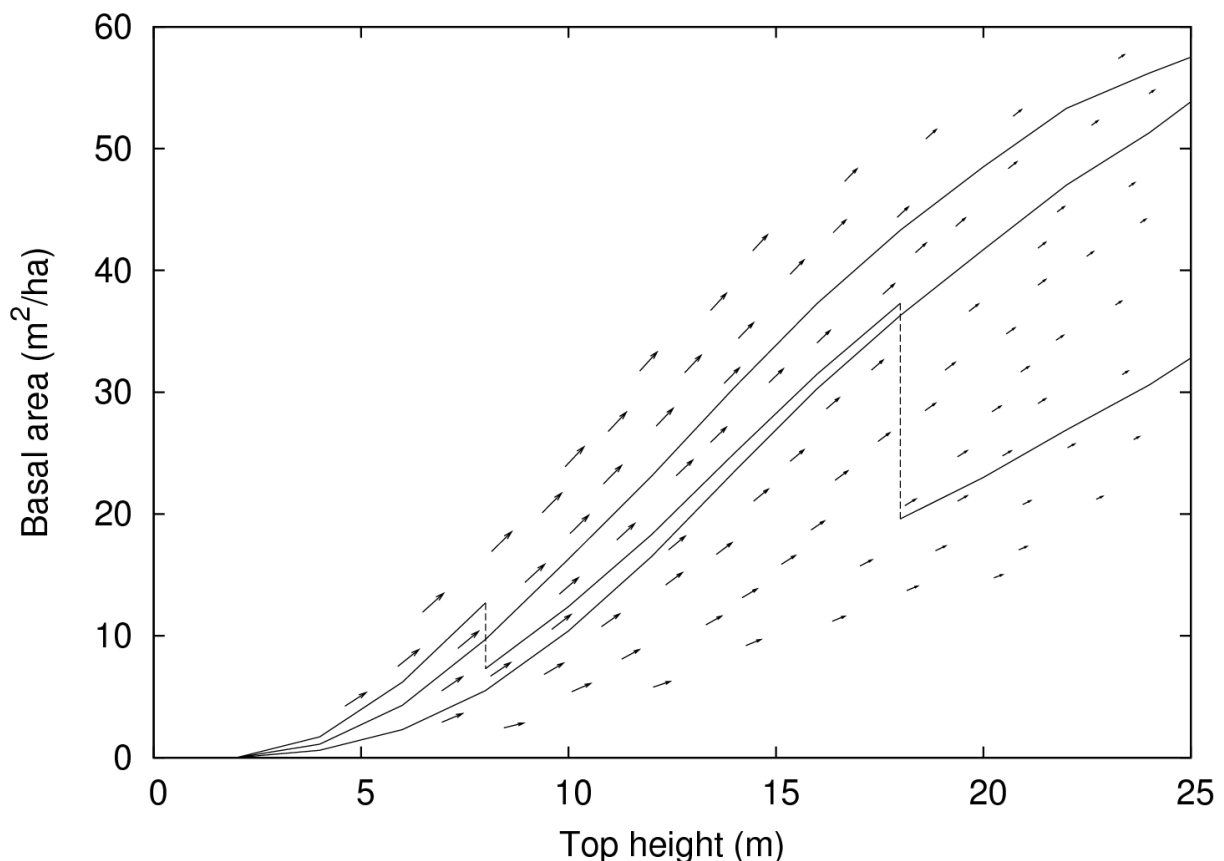


Figure 2: Example of dynamic growth model with two state variables. The arrows representing the rates of change for the two variables can be followed to generate a trajectory starting from any



Because of their flexibility, conceptual simplicity, and the representation of interaction hypotheses in the most natural and intuitive way, individual-tree models, individual-based models, or multi-agent models, are particularly attractive as research tools. Even for these purposes, however, high dimensionality has drawbacks that may make it advisable to complement the individual-based models with aggregate models. One of the disadvantages might be called the “so what?” effect: observing the results of complex simulations is not always very illuminating. A reductionist approach may also lose sight of so-called emergent properties. There is a growing and confusing literature on

these properties, that tend to be described as “the whole being greater than the sum of the parts”; in reality, they generally relate to the fact that correctly modelling individuals is usually easier than modelling the interactions between them. As stated by Levin and Pacala (1997): “individual-based models have the advantage that they are closer in detail to real systems; that advantage is also a disadvantage in that they retain all the details that may hide what is really important at broader scales”. For these and other reasons, there is growing interest on deriving aggregate models from individual-based models (Garcia 2001).

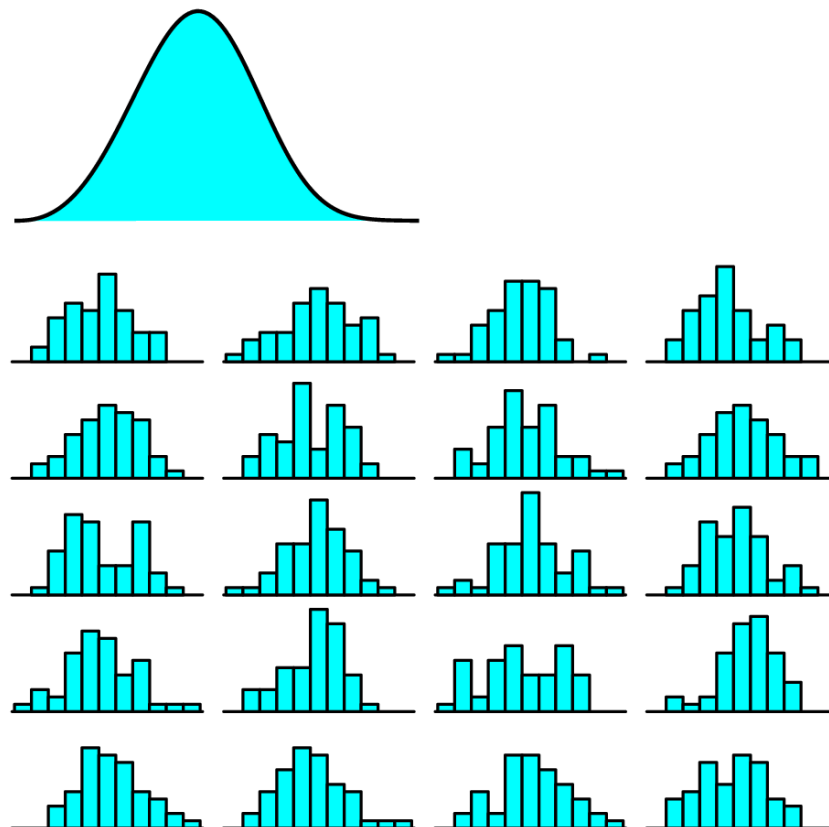


Figure 3: Histograms of 20 random samples of size 50 obtained from the distribution shown.

In the case of prediction for decision-making, the limitations of individual-based models are more serious. One relates to spatial correlations in the size and growth of neighbouring trees caused by competition, by similarities of micro-site, or by other factors. Tree sizes not being independently distributed on the ground, the concept of size distribution used by distance-independent models presents problems, and parameter estimation can have significant biases. It has also been found that the effects of micro-site often produce positive spatial correlations, masking any negative correlations due to competition, contrary to the assumptions in current distance-dependent models (Garcia 2006). A second problem with the application of highly detailed models is that often the initial state is not known with sufficient precision, which obviously does not allow a reliable prediction of future states. For instance, even assuming independence, it is known that obtaining reasonable estimates of higher moments, or of the shape of a probability distribution, requires very large samples (Kendall and Stuart 1976). Often this high variability (Figure 3) is not appreciated in the applications, and one can be misled as to the credibility of the projections. A third problem

with the use of complex models for making predictions is discussed in more detail next.

## Limits to predictability

This limitation of individual-based models goes beyond growth modelling, and has to do with what can and can not be predicted. To illustrate, think of the circles of Figure 4 as particles, balls, or pucks moving in the plane. An individual is thrust in a certain direction. It is easy to calculate the trajectory (Figure 4a). Now, what happens if we change slightly the launching angle? (Figure 4b). Even with an uncertainty of one millionth of a degree in the initial angle, the result becomes completely unpredictable after a few bounces: the “butterfly effect”, Chaos Theory, sensitive dependence on initial conditions.

The situation in forest growth modelling might not be as bad as this; or perhaps it might. Experiments with some models suggest that when altering the starting diameter of one of the trees by a few millimetres, the difference increases and spreads quickly to the rest. It may not be possible to project individual diameters as well as is generally believed.

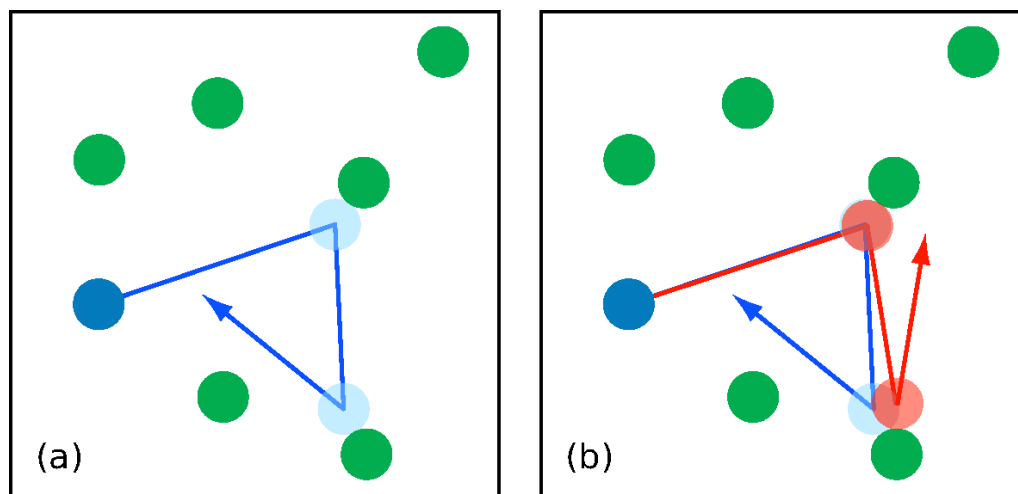


Figure 4: Trajectories in the plane. Effect of uncertainty in the initial angle.

What can be done? If Figure 4 represented a gas, the behaviour of the whole could be approximated by the equation  $PV = kT$ : pressure times volume is proportional to temperature. Note that these variables are properties of the aggregate, they do not exist at the molecular level. In fact, pressure and temperature are related to the mean and variance of the velocities. For an ideal gas, Statistical Mechanics is able to derive the aggregate equation from the dynamics of the individual molecules. In solids, the relationships between properties of the ensemble and molecular properties are still topics of research.

When designing a bridge or car component, in

principle one could model the trajectory of all the individual molecules. In practice, one would probably use an average position (centre of gravity), and apply an aggregate model proposed by Isaac Newton in the seventeenth century:  $d^2 \mathbf{x} / d t^2 = \mathbf{F} / m$ . This is an empirical model, based on observations, without any theoretical basis. And it is an approximation, it fails when going too fast or too small, although within a certain range it is pretty good.

Figure 5 depicts a pinball machine. Explanation for the younger audience: a ball is shot through the channel on the right-hand side, and drops down the slope colliding with various objects.



Figure 5: Pinball machine. Knowing its functioning does not guarantee predictability.



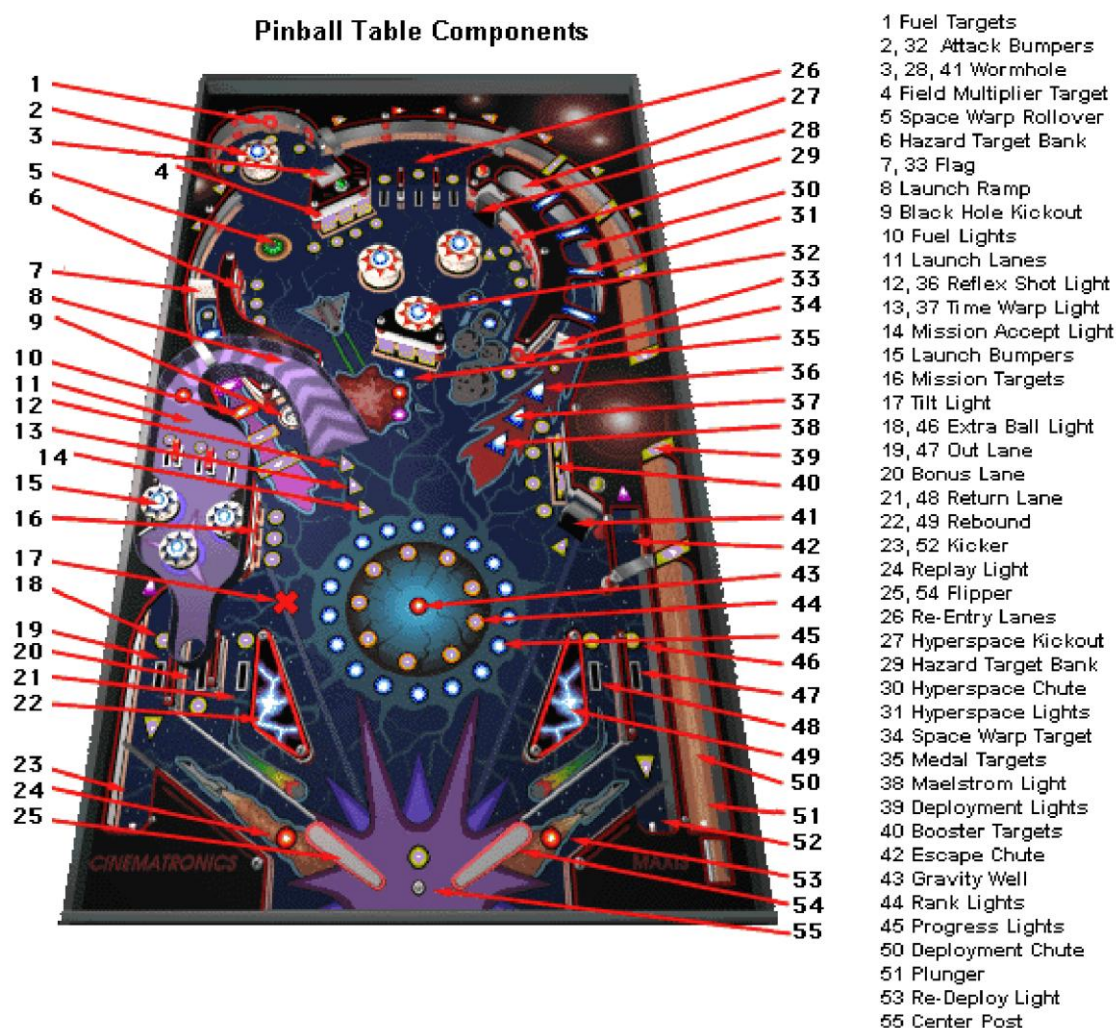


Figure 6: *Microsoft Pinball*, a pinball computer simulation. Simulating is not predicting.

The theory is well known, there is no mystery about its operation, but can we predict the trajectory of the ball? Understanding or explaining does not imply being able to predict. Figure 6 is an even better example: *Microsoft Pinball*. It is a fairly realistic computer simulation; apparently it contains no stochastic elements, only calculations based on physical laws. Given the time that the keyboard space bar is held down, the movement of the ball is perfectly predetermined. Can we predict it? It is sometimes argued that some complex process model is not yet very precise because of lack of knowledge about the functioning of some components, or of the values of certain

parameters; with further research it would become useful for management. In reality, a model can be very useful for understanding things better, but there are inherent limitations to predictability beyond a certain level.

A final example is shown in Figure 7. It is a device sometimes used in teaching probability. Steel bearings fall through a grid of pegs into the bottom compartments. It is hopeless trying to predict the fate of any of the balls individually. It is possible, however, to predict reasonably well the average final position, and to some extent, its variance. With an adequate sample it may be possible to get some idea of the distribution.



Figure 7: Device for demonstrating the binomial or normal distribution. Individual-level predictions are virtually impossible, but some statistical summaries can be predicted.

## Conclusions

Descriptive models, used primarily for research, should generally be mechanistic and detailed, with high dimensionality in their state space,

although some aggregated models can also be useful. Detailed process models contribute indirectly to the improvement of management-oriented models, but in these the priorities are different. For decision-making it is preferable to link decisions to consequences as directly as possible. Contrary to what is sometimes thought, the management of complex systems requires simple models.

## References

- García, O., 1988. Growth modelling - A (re)view. *New Zealand Forestry* 33 (3), 14-17.
- García, O., 1994. The state-space approach in growth modelling. *Canadian Journal of Forest Research* 24, 1894-1903.
- García, O., 2001. On bridging the gap between tree-level and stand-level models. In: Rennolls, K. (Ed.), *Proceedings of IUFRO 4.11 Conference "Forest Biometry, Modelling and Information Science"*, University of Greenwich, June 25-29, 2001. (<http://cms1.gre.ac.uk/conferences/iufro/proceedings>).
- García, O., 2005a. TADAM: A dynamic whole-stand approximation for the TASS growth model. *The Forestry Chronicle* 81 (4), 575-581, (Errata: 81(6), 815, 2005).
- García, O., 2005b. Thinking about time. In: Naito, K. (Ed.), *The Role of Forests for Coming Generations - Philosophy and Technology for Forest Resource Management*. Japan Society of Forest Planning Press, Utsunomiya, Japan, pp. 47-54.
- García, O., 2006. Scale and spatial structure effects on tree size distributions: Implications for growth and yield modelling. *Canadian Journal of Forest Research* 36 (11), 2983-2993.
- Goulding, C. J., 1972. Simulation techniques for a stochastic model of the growth of Douglas-fir. Ph.D. thesis, University of British Columbia.
- Grimm, V., 1999. Ten years of individual-based modelling in Ecology: What have we learned and what could we learn in the future? *Ecological Modelling* 115, 129-148.
- Kendall, M., Stuart, A., 1976. *The Advanced Theory of Statistics*, 4th Edition. Vol. 1: *Distribution Theory*. Griffin.
- Levin, S. A., Pacala, S. W., 1997. Theories of simplification and scaling of spatially distributed processes. In: Tilman, D., Kareiva, P. (Eds.), *Spatial Ecology. The Role of Space in Population Dynamics and Interspecific Interactions*. Princeton University Press, Princeton, New Jersey, Ch. 12, pp. 271-295.
- Luenberger, D., 1979. *Introduction to Dynamic Systems; Theory, Models and Applications*. Wiley, New York.
- Mitchell, K. J., 1975. Dynamics and simulated yield of Douglas-fir. *Forest Science Monograph* 17, Society of American Foresters.
- Mohren, G., Burkhart, H., 1994. Contrasts between biologically-based process models and management-oriented growth and yield models. *Forest Ecology and Management* 69, 1-5.
- Munro, D. D., 1974. Forest growth models: A prognosis. In: Fries, J. (Ed.), *Growth Models for Tree and Stand Simulation*. Royal College of Forestry, Research Note 30, Stockholm, Sweden, pp. 7-21.
- Newnham, R. M., 1964. The development of a stand model for Douglas fir. Ph.D. thesis, The University of British Columbia.
- Staebler, G. R., May 1951. Growth and spacing in an even-aged stand of Douglas-fir. Master's thesis, School of Natural Resources, University of Michigan.

Stage, A. R., 1973. Prognosis model for stand development. Research Paper INT-137, USDA Forest Service, Int. Northwest For. and Range Exp. Sta., Ogden, Utah.

Vanclay, J. K., 1994. Modelling Forest Growth and Yield: Applications to Mixed Tropical Forests. CABI International, Wallingford, UK, 312 p.