

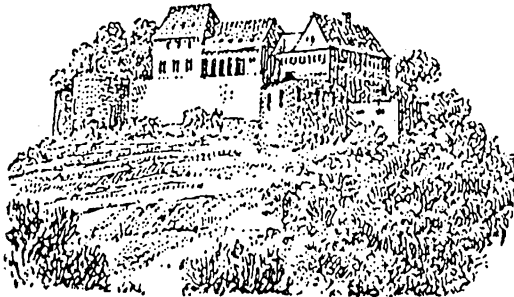
Dietmar P.F. Möller (Ed.)

2nd Ebernburger Working Conference

**Erwin-Riesch Workshop
System Analysis of Biological
Processes**

Bad Münster am Stein-Ebernburg

9.-11. April 1986



Ebernburg



Friedr. Vieweg & Sohn Braunschweig / Wiesbaden

Contents

Tutorial	1
Simulation of Biological Systems with Petri-Nets – Introduction to Modelling of Distributed Systems	3
<i>H. Fuss</i>	
Invited Papers	13
Synergetics and its Applications to Biological Systems	15
<i>H. Haken, A. Wunderlin</i>	
Das Osnabrücker Biosphären Modell als Simulationsmodell zur Beschreibung der globalen Änderung des Kohlenstoffkreislaufs	29
<i>H. Lieth</i>	
A Dynamic Simulation Model of Tree Development under Pollution Stress	46
<i>H. Bossel</i>	
Selected Papers	59
Methods	61
Fit, Fitter, The Fittest – Methods for Modeling and Identification of Dynamical Systems	63
<i>H. G. Bock, J. P. Schlöder</i>	
Strange Limit Cycles, Chaos, and Invariant Measure for a Simple Differential-Delay Equation	73
<i>U. an der Heiden</i>	
The use of Computer Simulation to Evaluate the Testability of a new Fitness Concept	82
<i>W. Gabriel</i>	
Texture Analysis using Random Field Models exemplified on Ultrasonic Images of the Liver	89
<i>U. Ranft</i>	
Smooth Descriptive Modelling of Multifactorial Systems Responses	95
<i>G. H. Klein</i>	
Zur Beschreibung offener thermodynamischer Prozesse durch Bindungsdiagramme	101
<i>A. Schöne</i>	

Biological and physiological Systems	113
The use of Artificial Intelligence for Simulation of Metabolic Processes	115
<i>J. R. Reichl</i>	
Population dynamics of <i>Daphnia magna</i> – Simulations using the individuals' approach	127
<i>V. Fitsch, H. Kaiser</i>	
Estimation of Individual Growth Curves from Aggregate Data	133
<i>W. Wosniok</i>	
Sleeping Stem Cells – A Model of Of Stem Cells under Continuous Stress	140
<i>G. Pabst</i>	
A Mathematical Method of Modelling and Simulating Biological Structure Control Systems	148
<i>T. Vogelsaenger</i>	
Simulation in electrophysiological pharmacology: Specific interactions of antiarrhythmic agents with ion channels of the cardia cell membrane	154
<i>D. Hafner, F. Berger, U. Borchard</i>	
Simulation of the Human Blood Circulatory System with the help of an Uncontrolled Pulsatile Model and its Validation	165
<i>T. Sikora, D. P. F. Möller, V. Pohl, E. Hennig</i>	
Model-Reduction for the Parameter Identification of an Uncontrolled Pulsatile Model of the Cardiovascular System	173
<i>V. Pohl, D. P. F. Möller, T. Sikora, E. Hennig</i>	
On the improved Estimation of the Compliance-Parameters of the Physiologically Closed Cardiovascular System	179
<i>A. Tanha, H. Maftoon, G. Thiele, D. P. F. Möller, D. Popovic</i>	
Some aspects of the application of Neurodynamical Models for the Simulation of Central Regulation and Dysregulation	188
<i>O. Hoffmann</i>	
Measuring Symptoms in Parkinson's Disease with a tracking device	194
<i>S. S. Hacisalihzade, C. Albani, M. A. Mueller</i>	
Round Table Discussion	201
<i>G. K. Wolff</i>	
Summary and Discussion	207
<i>O. Richter</i>	
List of Speakers	213
Index Register	215

The use of Computer Simulation to Evaluate the Testability of a new Fitness Concept

Wilfried Gabriel, Plön

Abstract

Using a model for the evolution of fitness, we show how computer simulation can be used to examine whether a theoretical model can be tested by experimental data. A maximum-likelihood-fit on simulated measurements was able to separate the parameters. Such measurements can be made in a realistic experimental set-up. Therefore, the concept and the predictions of the model can be experimentally tested.

Zusammenfassung

An dem Beispiel eines neuen Fitnesskonzepts, wo der Verlauf der Fitnessfunktion selbst der Evolution unterworfen ist, wird gezeigt, wie Computer-Simulation hilfreich sein kann, um die Testbarkeit von Modellvorstellungen zu untersuchen. In diesem Modellkonzept mit quantitativ genetischen Ansätzen war es ungewiß, ob die wesentlichen Modellparameter überhaupt aus experimentellen Daten bestimmbar sind. Mit einem Maximum-Likelihood-Verfahren gelingt es, diese Modellparameter zu separieren. Simulierte Messungen demonstrieren, daß der dazu notwendige Meßaufwand in realisierbarem Rahmen bleibt. Damit erweisen sich die grundlegenden Modellvorstellungen und Modellvorhersagen als experimentell testbar.

1. Introduction

The assumptions underlying many mathematical models in biology are such simplifications of reality that the models cannot be tested by empirical data. Nevertheless, such models do clarify ideas and enable precise definitions, as well as promoting a qualitative understanding necessary for the development of new concepts. Realistic models in biology, however, require a set of testable predictions. Crucial experiments or observations should be able to falsify a model or to suggest useful modifications. For that purpose a biological interpretation of all model parameters is required. If these parameters are not directly measurable, a description of a realistic experimental set-up and of a statistical procedure to extract the parameters from data are necessary to achieve creditibility with experimenters.

Here we present an example of a model concept for which it was uncertain, whether the parameters could be estimated from experimental data. In almost all genetic models (e.g. *Lande 1982, Lynch and Gabriel 1983*) the individual fitness function (or equivalently the individual niche width) is fixed and, therefore, independent of evolution. Using concepts from quantitative genetics, we redefine fitness as a tolerance curve which itself changes during the evolutionary process. Adaptation is driven in the direction of maximal fitness. Therefore, the relative fitness contribution of any quantitative trait is calculable by its difference from the possible optimal value and by the shape of the fitness function. However, in temporally and spatially variable environments natural selection changes the breadth of adaptation in relation to expected variations of the optimum. Thus, both the optimum value and the whole shape of a tolerance curve or fitness function are evolving. By incorporating this into a model concept under quite general conditions, the optimal breadth of adaptation can be predicted (most simple in clonal populations) as a function of the variability of the environment (*Lynch and Gabriel 1986a,b*). Thereby, according to our model, spatial heterogeneity, temporal variation within a generation, and temporal variation between generations act independently. The produced variances represent, in general terms, contributions to the fitness with multiplicative and additive effects.

2. Model Concept

We start by taking the simplest possible assumptions and the minimum number of parameters. Let us assume that the tolerance curve of an individual over an environmental gradient can be represented by two quantitative genetic characters: g_1 describes the optimum and g_2 the variance of the tolerance curve. Each of these characters may be the expression of a large number of genes and is measured on an environmental scale. However, this description is valid only for the average of a clonal population because of the inevitable variation due to development. Therefore, the fitness of an individual is determined rather by its phenotypic values z_1 and z_2 , than by the corresponding genotypic values g_1 and g_2 . We suppose that in most cases with a proper scale transformation the fitness function can be approximated by a normal distribution:

$$w(z_1, z_2, \phi) = (2\pi z_2)^{-1/2} \exp(-(z_1 - \phi)^2 / 2z_2) . \quad (1)$$

The actual environment (the model is most applicable to physical or chemical gradients like temperature or pH) is measured by ϕ . The variance of phenotypes with identical genotypes growing up under the same conditions is called developmental noise V_E (according to the traditional nomenclature). Quantitative genetics usually puts a normal distribution for the relationship between phenotypes and genotypes

$$p(z_1|g_1) = (2\pi V_{E1})^{-1/2} \exp(-(z_1 - g_1)^2 / 2V_{E1}) , \quad (2)$$

but as z_2 is a variance it always has to be greater than zero. For the distribution of z_2 around g_2 we use a beta distribution of the second kind (Kendall and Stuart 1977)

$$p(x) = x^{\alpha-1} (1+x)^{-(\alpha+\beta)} \Gamma(\alpha+\beta) / (\Gamma(\alpha)\Gamma(\beta)) .$$

For mathematical reasons it is helpful to set $x = z_2/V_{E1}$. By this substitution V_{E1} acts as a scaling factor. The z_2 -distribution, however, remains independent of V_{E1} . From the constraints $E(z_2) = g_2$ and $\text{Var}(z_2) = V_{E2}$, it follows that

$$p(z_2|g_2) = (z_2/V_{E1})^{\alpha-1} (1+(z_2/V_{E1}))^{-(\alpha+\beta)} \Gamma(\alpha+\beta) / (V_{E1}^{\alpha+\beta} \Gamma(\alpha)\Gamma(\beta)) \quad (3)$$

with

$$\alpha = (g_2(g_2 + V_{E1}) / V_{E2} + 1) g_2 / V_{E1}$$

$$\beta = g_2(g_2 + V_{E1}) / V_{E2} + 2 .$$

The four variables g_1 , g_2 , V_{E1} , and V_{E2} are essential for the understanding of the evolutionary process in question. Yet, as shown by the basic equations they have opposing influences on the effective fitness of a population. Therefore, as a first step it is necessary to demonstrate that each of these parameters can be measured under realistic conditions. In order to do this without appropriate experimental data, we simulated experiments with the expected range of variables. The parameter estimation is done by the following rather computer time consuming but effective maximum likelihood procedure.

3. Parameter Identification by the Maximum Likelihood Method

In the following we assume that measures of fitness w_i at environmental states ϕ_i are available for n individuals of a single genotype. The ϕ_i need

not all be unique, but - according to the experience of several simulations - a minimum of 3-4 environmental settings are recommended in a range where the fitness changes at least one order of magnitude. Each fitness value is a function of the corresponding ϕ_i as well as of the unknown parameters g_1 , g_2 , V_{E1} , and V_{E2} . We wish to estimate these four unknowns by maximizing the joint (a posteriori) probabilities of all observations, the product of all n p_i (or the sum of $n \log(p_i)$), where p_i , the conditional probability of observing (w_i, ϕ_i) , is a function of g_1 , g_2 , V_{E1} , and V_{E2} .

We will focus on the simple case in which individuals are exposed to constant ϕ throughout their lives as would be approximated in a controlled laboratory or greenhouse setting. A specific measure of individual i in environment ϕ_i can result from many different combinations of z_1 and z_2 , but equation (1) implies the constraint

$$z_1 = \phi_i \pm (-2z_2 \ln(w_i(2\pi z_2)^{1/2}))^{1/2} . \quad (4)$$

The a posteriori probability of the observation (w_i, ϕ_i) can now be expressed by weighted integration over all possible combinations of z_1 and z_2 . For each z_2 the probability of obtaining the z_1 that results in the observation (w_i, ϕ_i) can be calculated by substituting equation (4) into (2):

$$H(z_2, w_i, \phi_i) = (2\pi V_{E1})^{-1/2} (\exp(-(x-g_1)^2/2V_{E1}) + \exp(-(y-g_1)^2/2V_{E1})) \quad (5)$$

where $x = \phi_i + c$, $y = \phi_i - c$ and

$$c = (-2z_2 \ln(w_i(2\pi z_2)^{1/2}))^{1/2} .$$

The (a-posteriori) probability p_i of the measurement (w_i, ϕ_i) is then given with (3) and (5)

$$p_i = \int p(z_2 | g_2) H(z_2, w_i, \phi_i) dz_2 .$$

The boundaries of this integration are $z_2 = 0$ and $z_2 = 1/(2\pi w_i^2)$. Above this limit it is impossible to measure a fitness of the value w_i . Unfortunately, there is no analytical solution available for this integral. Another complication arises from the fact that the expectation value of the geometric mean fitness (f_i) varies slightly with ϕ . For combining measurements of different environmental settings, the contributions of distinct ϕ values to the joint probability have to be weighted by the expectation values f_i in order to avoid systematical errors. Given the above equations

the maximum likelihood estimates of the four parameters g_1 , g_2 , V_{E1} , and V_{E2} are obtainable using well-known Monte-Carlo-techniques for f_i , numerical integration for p_i and optimization procedures to maximize the joint probability. At the end of the maximization, estimates of the second derivatives at the maximum may also be computed to obtain (from the inverse of the Hessian matrix) estimates of the (sampling) variances and covariances of the four parameters.

This same analysis could be applied in situations in which ϕ is not constant within the lifetime of measured individuals but has for all ϕ settings the same variance V_ϕ . In this case, using equations derived in *Lynch and Gabriel (1986b)*, the equation (4) becomes

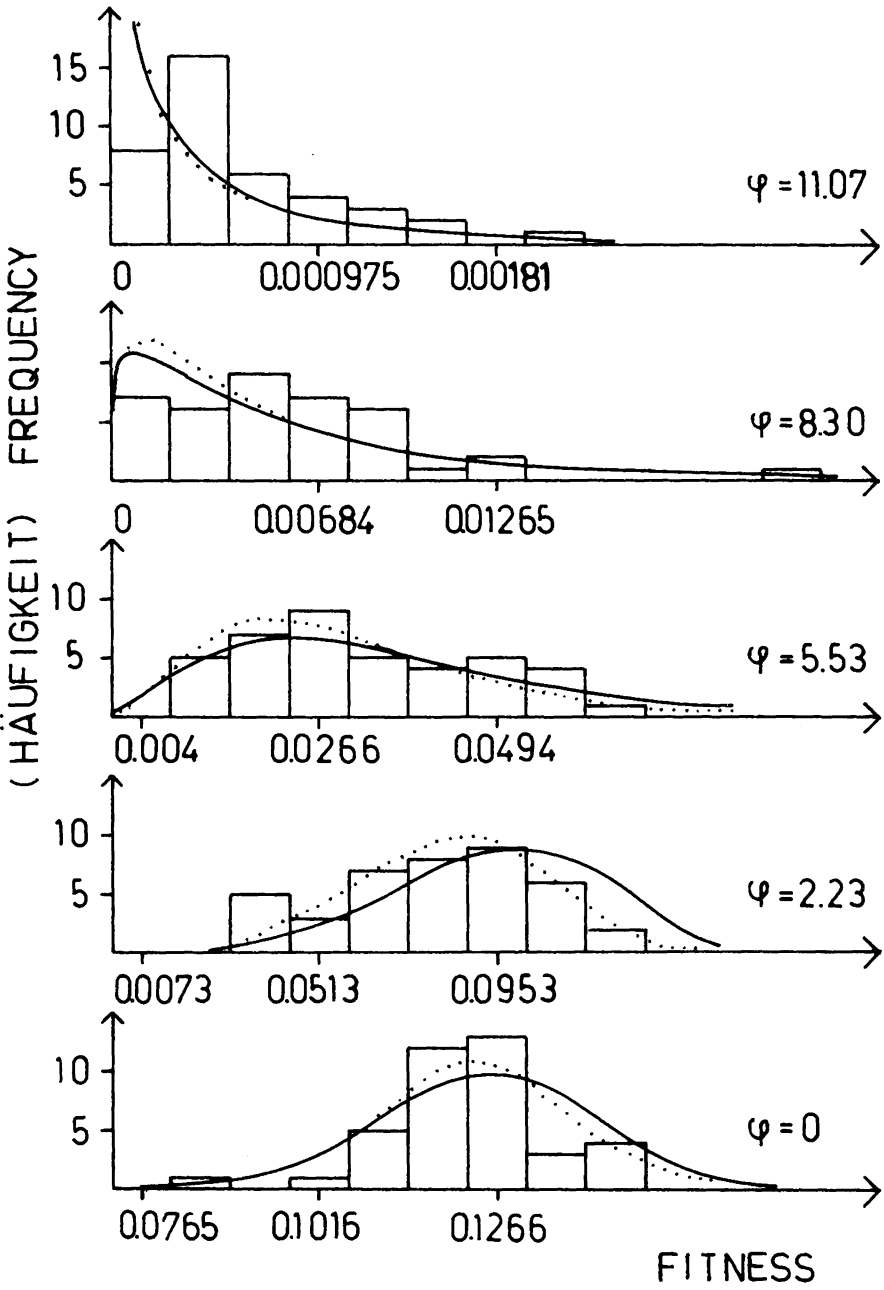
$$z_1 = \phi_i \pm (-2z_2 \ln(w_i(2\pi z_2)^{1/2} - V_\phi)^{1/2}) \quad (6)$$

With appropriate sample size, use of (6) in the definition of p_i would enable the investigator to derive an estimate of V_ϕ as well as g_1 , g_2 , V_{E1} , and V_{E2} .

4. Results

The dotted line in the figure represents the expected fitness distribution at five environmental values ($\phi = 0, 2.23, 5.53, 8.3, \text{ and } 11.07$) for a given parameter set ($g_1=0, g_2=10, V_{E1}=1, \text{ and } V_{E2}=3$). The scale of the environmental variable ϕ is chosen so that the fitness function of an average animal has its maximum at $\phi=0$. Besides the different numerical range the form of the distribution of fitness values changes with ϕ and shows distinct asymmetries. It can be shown that these distributions and especially the asymmetries depend characteristically on the parameters g_1 , g_2 , V_{E1} , and V_{E2} . Therefore, it is necessary to have the measurements of single individuals; the usually published mean values and standard deviations of fitness are not sufficient.

The figure also contains simulated measurements. These data are presented as histograms in order to facilitate graphical presentation and comparisons with expected values. For each of the five ϕ -values we have simulated 40 measurements. The procedure described above was then applied to the joint distribution of all 200 data points. The results on the presented set of simulated data are $g_1=0.51, g_2=8.85, V_{E1}=1.14, \text{ and } V_{E2}=3.30$. The



expected fitness distribution according to this maximum likelihood result are given by the unbroken line in the figure. Our shown results are atypical in that they appear to fit the data only poorly. However, this may arise from the use of too few statistics of simulated events, even with respect to the dotted line (the goodness of fit is still sufficient; a table with results of other simulations is given in *Lynch and Gabriel, 1986b*). Our aim is to demonstrate that with a manageable effort of 200 measurements over an environmental scale where the fitness changes by two orders of magnitude, the basic parameters for a new fitness concept can be estimated by experimental data. Thereby it is demonstrated that concepts and predictions of this model as given by *Lynch and Gabriel (1986a,b)* are testable.

Acknowledgements

We thank G. Carvalho and W.R. DeMott for improving the manuscript.

References

- Kendall, M. and Stuart, A. (1977):* The advanced theory of statistics, Vol. 1, 4th Ed., Charles Griffin, London.
- Lande, R. (1982):* A quantitative genetic theory of life history evolution. *Ecology* 63: 607-615.
- Lynch, M. and Gabriel, W. (1983):* Phenotypic evolution and parthenogenesis. *Am.Nat.* Vol. 122: 745-764.
- Lynch, M. and Gabriel, W. (1986a):* Evolution of the Breadth of Biochemical Adaptation. - In: P. Calow (Ed.): *Adaptational Aspects of Physiological Processes*. Cambridge University Press, Cambridge.
- Lynch, M. and Gabriel, W. (1986b):* Environmental Tolerance. (submitted to *Am.Nat.*)