SEMINAR/WORKSHOP ON STATISTICAL ECOLOGY

SCHOOL OF ENVIRONMENTAL SCIENCES

UNIVERSITY OF POONA

16- 22 DECEMBER ,1985

- SPONSORED BY

UNIVERSITY GRANTS COMMISSIONS

AND

DEPARTMENT OF ENVIRONMENT

MODELLING DIFFUSION OF CROSS-BRED GOATS IN A RURAL COMMUNITY

A.P.Gore and U.A. Lavaraj

Department of Statistics
University of Poona
PUNE 411 007.

MODELLING DIFFUSION OF CROSS-BRED GOATS IN A RURAL COMMUNITY

By

A.P. Gore and U.A.Lavaraj
Department of Statistics
University of Poona.

1. Summary :

Diffusion of desirable innovations in a community is a phenomenon that interests all students of development. In the present article we discuss an exploratory attempt to understand the process of adoption of cross-bred goats during the last decade in the village Narayangaon, District Pune, Maharashtra State. The standard logistic growth model is fitted as also a model for marketing of new products, developed by Bass. Some possible implications are discussed.

2. Introduction :

Cross breeding of cows has become an established technology in India. However the capital investment involved in terms of either the price of a cross-bred/exotic cow or cost of rearing an heifer is high and participation in this program is difficult for weaker sections of the society. Goat keeping is relatively more common among these sections. Yet cross breeding of goats has not been taken up on any major scale. Rural Agricultural Institute, Narayangaon, Dist. Pune, is a voluntary

organisation that has introduced this innovation in the Ambegaon and Junner Talukas of Pune District since 1974. Male goats of Saanen breed imported from overseas have been made available for servicing female goats of local variety for a small fee and thus for about two thousand farmers have taken advantage of this facility. This number is divided about half and half between residents of Narayangaon village and other villages nearby.

The aim of our study is to understand the process of adoption of cross-breeding by the population of the area. This study is divided into three parts. First we report the findings of a mail questionnaire survey about the socioeconomic status, attitude and experience of the participants. Second and third part concern the modelling of the first adoption timing of participants. Two models have been fitted. The well known logistic growth model is used first and then a model due to Bass (1969) proposed for first purchase times for new products such as T.V. The last section concerns the implications of the study and further work.

3. Mail-Questionnaire Survey :

The Rural Agricultural Institute mailed blank questionnaires to several hundred goat owners who had used the available cross breeding service. Over two hundred and fifty of these were returned. Findings of this survey are reported below.

(a) Profile of respondents :

The following three tables describe the household size, land holding and livestock of the beneficieries.

Table 1

Distribution of households

Persons in a family	1	2-5	6-10	11 or more
Number of households	4	84	110	40

Table 2

Land holding (in acres)

Dryland	Irrigated		
	0	1 to 4	5 or more
0 to 5	70	127	22
6 to 25	5	10	9
26 or more	1	0	0

Table - 3

House size	Animals (excluding goats) 0	1	2 to 4	5 or 6	7 or more
1	3	0	0	0	0
2 to 5	30	4	26	9	15
6 to 22	22	8	27	13	36
11 or more	6	1	6	7	20

Table 1 suggests that beneficiaries are fairly distributed in terms of size of household and number of individuals in the family may not have much to do with adoption.

Table 2 seems to indicate that very few beneficiaries are from the 'big' landowning families. Nearly 30% have no irrigated land and only 12% own 5 or more acres of irrigated land. Perhaps this is as expected. Goat is essentially a poorman's interest.

From table 3 we find that about a quarter of the beneficiaries own no animal other than goat. About equal number own 7 more animals. This would be taken with a pinch of salt since the number includes chicken. On the whole the livestock holding (other than goat) seems to be modest too.

A tentative conclusion emerges that the programme seems to be catering to relatively weaker sections. Rich farmers seems not be interested in it.

(b) Number of goats owned :

Table -4

Number of cross bred goats	Number of Indigenous Goats			
	0	1 to 3	4 to 7	8 or more
1	64	37	17	5
2 - 4	16	2	5	1
5 or more	10	7	2	1

Over 50% of owners of cross-bred goats possess no local goats. Thus they seem to have switched over completely to cross-bred goats. Others who own one or more local goats, represent the growth potential for the cross breeding program. About 75% of the owners possess only 1 cross bred goat which again suggests meagre means.

Table - 5

	Number of services of Saanen male used.			
	1	2	3	4 or more
frequency	151	24	6	5

This table shows that the number of farmers who have come back for repeated service is only about 15%. This is a very important point. Further investigation is necessary to check whether there is a tendency to turn away from the innovation and if so why ?

(c) Milk yields reported :

Table - 6

Milk yield (in litres) of indigeneous goat and period of location

Location period in months	Yield (in litres) per day			
	0-0.25	0.25- 0.50	0.50 - 0.75	0.75 or more
1 - 3	7	11	3	6
4 - 6	10	64	12	38
7 or more	1	4	0	12

Table - 7

Milk yield (in litres) or cross-bred goats and period of location

Location period in months	Yield per day		
	0-1	1-2	more than 2
1 - 3	1	0	2
4 - 6	9	11	3
7 or more	11	32	15

Table 6 and 7 bring out the main advantage of cross bred goats. According to respondents, cross bred goats seems to have both longer period of location and higher milk yield. Only about 5% cross bred goats reportedly gave milk for less than 3 months while the corresponding figure for local goats is nearly 3 times as much. About 25% cross bred goats give 2 litres of milk or more per day and nearly 75% give more than a litre. In indigenous goats those giving 750 ml or more of milk per day are only about 30%. This difference is what should attract small farmers to the program.

(d) Feeding and health :

Farmers depend upon free grazing as well as stall feeding of goats. Various trees and residues of crops are mentioned as sources of fodder. 4 most commonly mentioned trees are banyon, mango, acacia and peepal. Nearly two thirds of the respondents did not answer questions about feeding practices. But among the 90 that did, as many as 40 reported total stall feeding and no free grazing. This is very remarkable. In the long run, from an ecological point of view, stall feeding must replace free grazing if goat breeding program is to be sustained. While questions were asked about fodder and concentrates purchased from market, 60% respondents gave no information at all and in the remaining

cases it is mostly scanty. This question therefore needs further investigation.

About 75% of those who answered the question felt that cross bred goats are not specially susceptible to diseases. The question of immunisation was not well responded to. The following table suggests that immunisation may be quite helpful but its awareness is perhaps still not adequate.

Table - 8

Immunisation and incidence of disease among cross bred goats

		Immunisation	
		Yes	No
Disease proneness	Yes	3	10
	No	22	20

(e) Selling price of goats :

Table - 9

Sale price in rupees of cross bred goat (male) by age in months (rounded off)

Price	Age		
	1 - 3	4-6	7 or more
upto 300	25	21	2
300 to 600	1	3	8
over 600	2	2	6

Table - 10

Sale price in rupees of cross bred goat (Female)

Price	age in completed years 1	2	3 or more
upto 300	7	1	1
300 to 600	7	2	2
600 to 800	1	2	1
above 800	1	12	12

The pattern in case of male goats seems to be to sell it in the first six months for a couple of hundred rupies while in case of females it may be best to sell it in the third or fourth year when it can bring as much as a thousand rupees while if sold in the second, it may still earn three or four hundred rupees.

As all text books of sampling warn, mail questionnaire surveys are often subject to major nonresponse bias. It is possible that experiences of nonrespondents may have been different. In depth investigation of this aspect has not been possible. But a small number of detailed interviews were conducted to supplement the survey.

(f) Supplementary interviews :

50% utilisation of the cross breeding service is in Narayangaon itself. So about half of our interviews were also in the same place while about equal number of interviews were taken in 3 other villages. Selection of interviewes was partially purposive in so far as we tried to interview several of the award winners of the first goat ralley held by the institute in the previous month. Two individuals connected with the institute were interviewed and also some other individuals available

In a typical interview, it was first ascertained that the family did own atleast one cross bred goat. Then the first question asked was how the family came to know about the cross breeding facility. The answers basically fall into 3 types. First, direct contact with the institute. Thus a barber who gave hair-cut to people in the institute heard about exotic goats from them. The second type is those who were impressed by the experience of their relatives. The third and perhaps the most common type is those who observed animals kept by others in their village. Whenever such random contact process is operative the number of new adopters is directly proportion to the existing number of adopters and also the size of the villages. Thus we expect that the logistic growth model should give good fit as well be seen in the next section.

It is interesting to note that speeches, pamphlets or other written publicity materials, propaganda programs or such other sources were not reported in our interviews. Nor did anyone mention influence of village or community leaders. This point needs further probing and we hope to have detailed discussions about it with some early adopters, in the next stage of the study.

We inquired about the benefit that the adopters hoped to get from the cross-bred goats. Invariably the answer was "more milk". Indigenous goats give typically about half a litre of milk while the cross-bred goat may give twice or thrice as much. By and large, this expectation seems to have been fulfilled and incidentally, since the cross-bred goat is stall fed, keepers collect some manure.

Regards feeding, it appears that the common practice is offering green fodder. Some owners reported use of grass and other fodder grown in their own fields but many seem to purchase it from market. As much as 3 or 4 rupees a day may be spent on this. Cross bred goats are not sent out for free grazing. The common explanation of this is that they are rather sensitive to sun and hence must be stall fed. This is quite interesting in so far as free grazing of local goats is considered by scientist,

to be environmentally undesirable. It also suggests that fodder is likely to be limiting factor in the spread of this innovation.

Next we discussed the difficulties, if any, in keeping cross-bred goats and reasons for not adopting. Firstly, a dam has to be taken on foot to the institute for service. The distance can be an impediment. Perhaps more people will join the program if the service is available in their own village. The institute envisages an artificial insemination program which will alleviate this difficulty. Secondly goats need attention and care. In many households women take up this responsibility. But if no adult can spare the time, adoption becomes difficult. In some well off households the view was that goat keeping is not sufficiently remunerative (and perhaps some what 'below dignity'). Available management skills are considered better deployed in keeping cattle or intensive irrigated farming. We repeatedly raised the question of taste of goatmilk. But this is not perceived to be a difficulty.

Lastly, we discussed the usefulness of the first goat rally held by the insitute. Goat keepers who won prizes on this occasion were very proud of their animals and discussed them at length with beaming faces. These individuals may play a useful role in further propagation of this innovation.

4. The Logistic Model :

Let N_t denote the total number of individuals who adopted during time interval (0,t). If we assume that the upward change in N_t is proportional to N_t, we get deterministic equation

$$\frac{d N_t}{dt} = a N_t$$

which simplifies to

$$N_t = b e^{ct} .$$

This is the well known equation for exponential growth. The main difficulty with this equation is that N_t explodes as t increases, which is unrealistic. So a damping factor $f(N_t)$ which is decreasing in N_t is introduced. Thus

$$\frac{d N_t}{dt} = N_t f(N_t) .$$

The simplest function f is the linear one with negative slope. It gives

$$\frac{d N_t}{dt} = N_t (a - bN_t)$$

which has the solution

$$N_t = \frac{c_1 e^{at}}{1+c_2 e^{at}} .$$

This is the classical logistic growth model. If we express growth of adoption in terms of proportion P_t of adopters in the population then in a changed notation.

$$P_t = \frac{K}{1+e^{-(a+bt)}} \quad ,$$

where K is the asymptotic upper limit on P_t (as t tends to infinity) and a,b are constants. With some elementary algebra, this can be written as

$$\log \frac{P_t}{K-P_t} = a + bt$$

This is the model that we try to fit to the data on cross-bred goat adoption. For application of this model to adoption of hybrid corn in U.S.A.See Griliches (1957) who uses the above form of the equation.

Our data are in the form of service records kept at the institute. From the inception of the program, all through the last 11 years record of all services is kept including date, name of the farmer, his village, number of female goats serviced and details of the exotic male used. In the first instance we have used data of adopters from Narayangaon, the headquarters of the institute.

Of the 2236 records of services between 17-9-73 and 25-1-85 1121 are from Narayangaon. After eliminating repetitions, the list contains 628 distinct names. These are the adopters. We shall assume that they come from different families. The population of Narayangaon (1971 census) is 10183 which we assume contains about 2037 families and divide the number N_t by it to obtain P_t. Given below are the annual increments in N_t.

Table 11

Yearly and commulative number of adopters

Year	73	74	75	76	77	78	90	80	81	82	83	84
number of new adopters in that year	8	65	24	~~77~~ 27	18	13	65	76	87	74	92	79
Comulative total	8	73	97	124	142	155	220	296	383	457	549	628

First notice that the value for 1973 is too low, since it is the value for 3½ months only. The value for 74 is distinctly high and is never attained again for 4 years. Perhaps in the initial enthusiasm to see their new program get established, the workers of the institute went all out to persuade the people and the enthusiasm later tapered off. We therefore decided to fit two curves, one including values for 73 and 74 and the other

excluding them. In each case value of K, the ceiling, was estimated by the following trial and error method. On a plot of p_t versus t a free hand curve was drawn through the points. This yielded an initial solution. Then, a and b were computed by the least squares method. Now the correlation coefficient between observed and estimated values of P_t was computed. The whole process was repeated by changing values of K by steps of .1 (and later, steps of 0.05) and chose that value of K for which the correlation coefficient was maximum. The resulting models are given below :

Table- 12

Fitting Logistic Model to Adoption Data.

	k	a	b	r
all data	0.5	-1.4523	0.1884	.9962
excluding 73, 74	0.35	-1.0483	0.2491	.9989

Clearly there is not much to choose in terms of r . However, the model using all data is more optimistic in that it predicts eventual adoption by half the population while the truncated model predicts that about one third of the population will adopt. It appears prudent to discount the unusually good results in 1974 and accept the second model. Here the attained proportion of

adoption is 0.27 while another 8% will be added to this in the next 5 years or so. The excellent fit of the model can be seen in the graph.

Another approach to these computations is the iterative method proposed by Nelder (1961) which avoids the shortcut of linearisation. Our computations reveal that this method also leads to virtually identical estimates in this case. Raktowsky (1983) also recommends this procedure.

5. Bass Model :

The second model we have fitted to the data is due to Bass (1969). This model attempts to fit a curve to sales of new products and to develop a long range forecaste for the demand in future years. Bass applied his model successfully to markets for consumer durables such as air conditioners, refrigerators and T.V. sets when they were first introduced in the U.S..

Consider a discrete time coordinate t, t= 0,1,2,..., which refers to successive years. Let f(t) be the probability that an individual adopts the innovation (for the first time) in year t. Let F(t) correspond to the corresponding cumulative probability. Let Y(t) denote the observed total number of adopters during the years 0,1,...,t, and S(t) is the observed number of those who

adopted during the year t. Thus $f(t) = F(t)-F(t-1)$ and $S(t) = Y(t)-Y(t-1)$ using this notation, Bass postulates that conditional probability of adoption in the year t (given failure to adopt upto that time) is given by

$$\frac{f(t)}{1-F(t-1)} = p + q\,F(t-1).$$

The idea behind this equation is that some individuals (socalled innovators) have a probability of adoption independent of decisions of others (hence remaining the same for all t) while for others the probability is proportional to the extent of adoption by others upto that point. These latter are therefore termed imitators. The model is intended to reflect this consideration. p and q are unknown constants of the model.

If m is the total number of adopters (those who will adopt sooner or later), an unknown quantity, then expected value of $S(t)$ is $mf(t)$ and expected value of $Y(t)$ is $m\,F(t)$. Then the sample analogue of the model is

$$\frac{S(t)/m}{1-Y(t-1)/m} = p+q\,Y(t-1)/m\ .$$

This simplifies to

$$S(t) = mp + (q-p)\ Y(t-1) - \frac{q}{m} Y^2(t-1)$$

$$= a + b\ Y(t-1) + c\ Y^2(t-1) \quad \text{say.}$$

This equation was fitted to the Narayangaon data by least squares. Notice that

$$cm^2 + bm + a = 0 .$$

This quadratic equation will yield an estimate of m given a,b,c. The results of the computation are as follows :

Table - 13

	Estimates for parameters of Bass model					
	a	b	c	p	q	m
for whole data	20.22	0.2025	-1.4×10^{-4}	0.0132	0.2156	1537.82
for data excluding 1973,74	13.80	0.4323	-6.2×10^{-4}	0.0191	0.4514	724.46

Again using the estimated parameter values, expected values of numbers of adopters in different years were calculated and their correlation with observed values was obtained. The values are

r(for the whole data) = 0.73

r(for data excluding 1973, 74) = 0.88 .

The high values of the correlation coefficient are reflected in the relevant graph. Here too, it seems prudent to use the second model since the correlation is higher and since the first two observations are suspected outliers. This model also turns out to be conservative in its projection of the total number of adopters, which is rather close to the value obtained using the logistic growth model.

6. Remarks :

It is clear that the innovation of goat cross breeding has taken root in Narayangaon. Its spread is not spectacular and only about one third of the population is expected to adopt it. The question of why others seem likely to stay out needs further probing. Also the question of repeated visits or lack of the same and reasons behind any drop out rate need to be looked into. Further it is necessary to extend this study to include other villages in which about half the adopters are located. This will necessitate analysis of the spatial aspect of diffusion say along the lines of the classical work by Hagerstrand (1952). The Rural Agricultural Institute, Narayangaon is also involved in promotion of biogas plants and smokeless chulhas. We hope to work on modelling of these diffusion phenomena in near future.

Fig. I FITTING LOGISTIC GROWTH MODEL FOR GOAT DIFFUSION IN NARAYANGAON

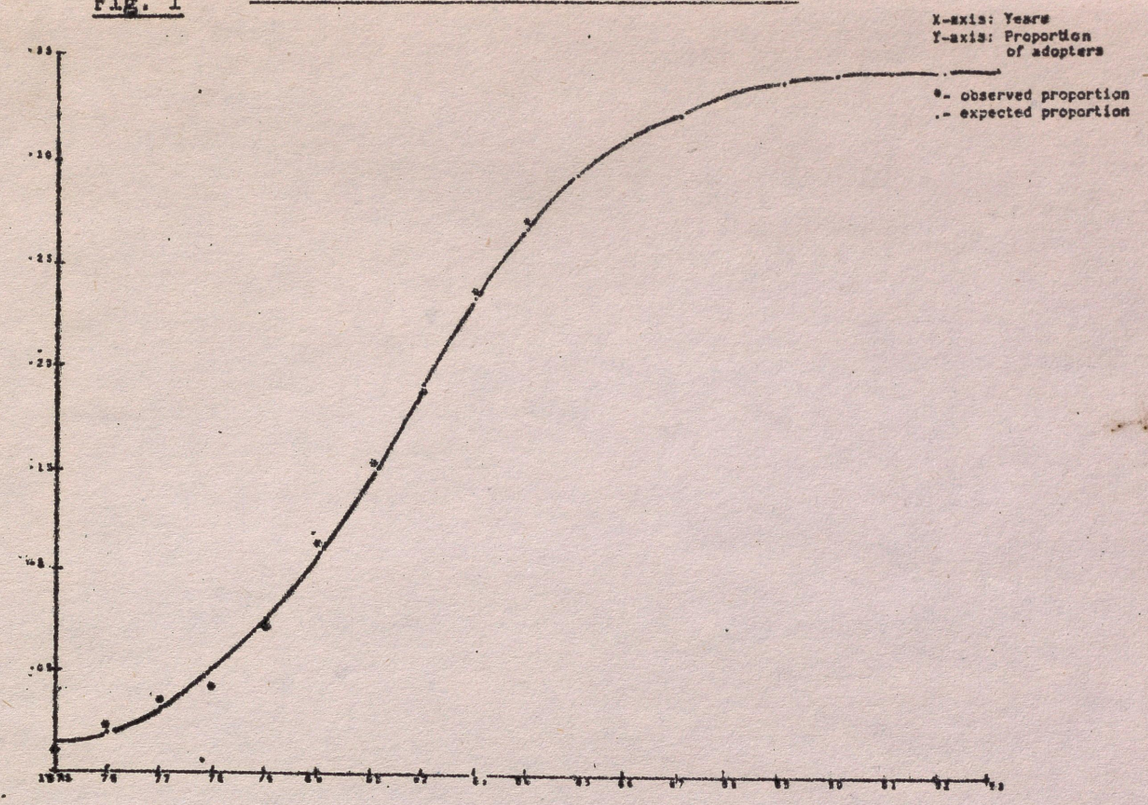
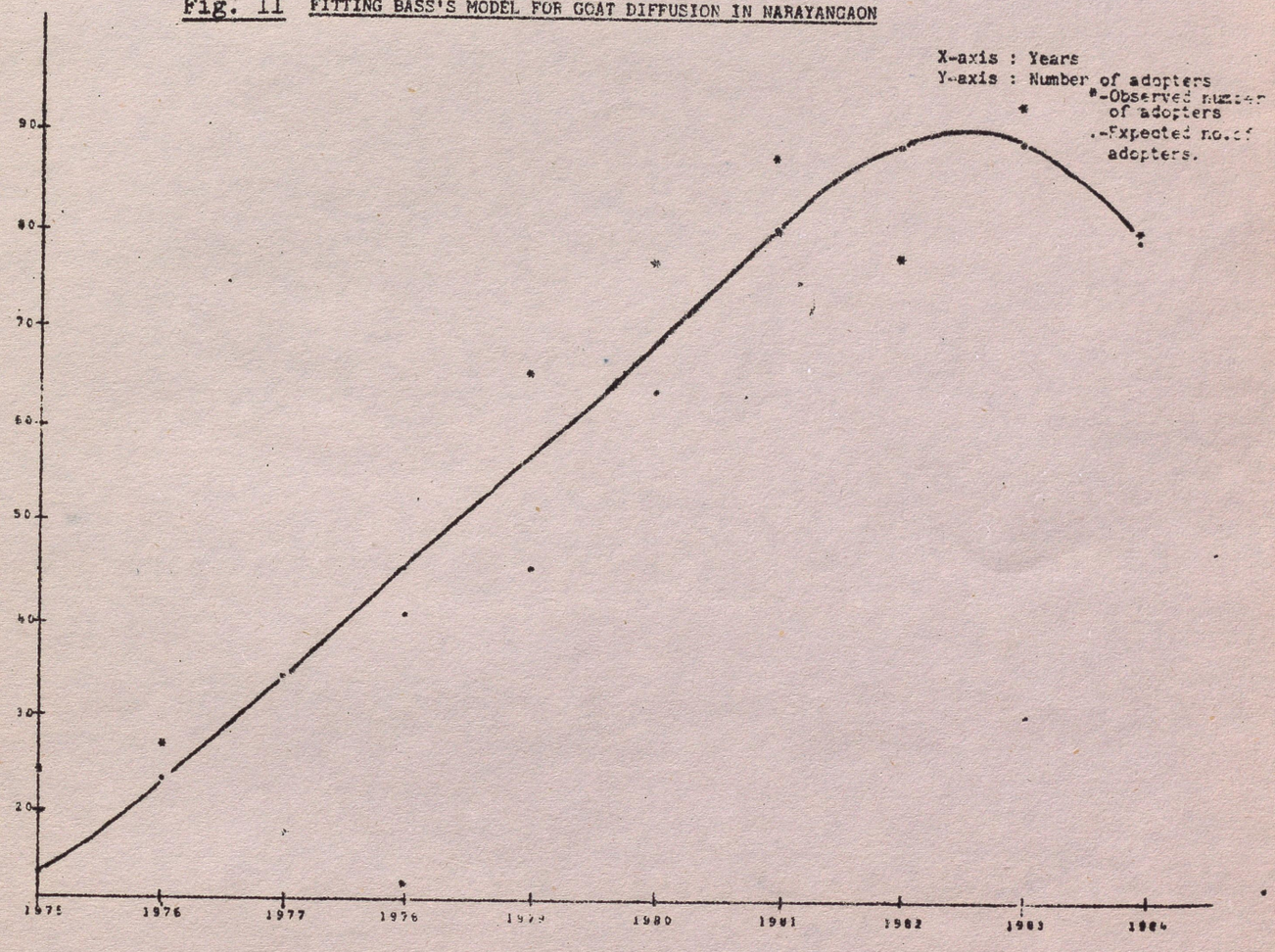


Fig. II FITTING BASS'S MODEL FOR GOAT DIFFUSION IN NARAYANGAON



7. Acknowledgement :

We are grateful to Mr. Shrikant Sabnis, Managing Trustee, Rural Agricultural Institute, Narayangaon(RAIN), District- Pune for extending all cooperation and for many useful discussions.

References :

Bass, F.M. (1969). A new product growth model for consumer durables Management Science,15 215-227.

Griliches, Z. (1957). Hybrid corn : an exploration in the economics of technological change. Econometrica,25, 501-522.

Hagerstrand, T. (1952). The propagation of innovation waves. Lund studies in Geography 4, Lund, Sweeden.

Nelder, J.A. (1961). The fitting of a generalisation of the logistic curve. Biometrics 17, 89-110.

Raktowsky, D.A. (1983). Nonlinear Regression Modelling Marcell Dekker, N.Y.

... ...

ses...

SEMINAR/WORKSHOP ON STATISTICAL ECOLOGY

SCHOOL OF ENVIRONMENTAL SCIENCES

UNIVERSITY OF POONA

16- 22 DECEMBER, 1985

SPONSORED BY

UNIVERSITY GRANTS COMMISSIONS

AND

DEPARTMENT OF ENVIRONMENT

MARINE ENVIRONMENTAL STRESSES AND IMPACT ASSESSMENT

ARUN PARULEKAR

NATIONAL INSTITUTE OF OCEANOGRAPHY

GOA

Marine Environmental Stresses and Impact Assessment

Arun Parulekar
National Institute of Oceanography
Goa.

Type - Environmental - Anthropogenic - Background levels - Stability - Magnitude and Tolerance - Quantification Methodology - Cause effect Approach - Sampling design - Analysis - Constraints - Reliability.

Case Study 1 : Environs of a highly industrialized metropolis- Demographic pressure - Type of stress - Sewage, Domestic and industrial wastes - Fate - Effects - Organic enrichment - Deoxygenation - Impoverishment - Time linked defaunation and extinction.

Case Study 2 : Inorganic enrichment - Mining and dredging effects on an estuarine and coral reef ecosystem - Generation and settlement of high volume of suspended load- Obliteration of light penetration - Reduction in autotrophic production- Near extinction of sensitive species - Prevalence of opportunistic organisms - Breakdown in higher link trophic production - Net loss - Quantification and conservation plan.

... ...

ses...

SEMINAR/WORKSHOP ON STATISTICAL ECOLOGY

SCHOOL OF ENVIRONMENTAL SCIENCES

UNIVERSITY OF POONA

16- 22 DECEMBER, 1985

SPONSORED BY

UNIVERSITY GRANTS COMMISSIONS

DEPARTMENT OF ENVIRONMENT

MARINE ECOSYSTEMS

ARUN PARULEKAR

NATIONAL INSTITUTE OF OCEANOGRAPHY

GOA

Marine Ecosystems

Arun Parulekar
National Institute of Oceanography
Goa.

Tropical - Uniformly high water temperatures - High evaporation rate - Variability in salinity - Long food chain - Fast growth Less longevity - More species less number - Special types coral reefs, Estuaries, Mangroves - Importance. Utilization and Conservation.

Temperate - High variability in temperature - Less number of species - High population size - Slow growth rate - High Longevity - Relatively high productivity - Less diverse more self contained biota - Special types - Fjords, Lochs, Salt Marsh and open ocean.

Polar - Extremely low temperature - Restricted photoperiod- Nutrient enrichment - Freeze thaw cycle - Wind chill effect - Survival mechanisms - Short food chain - Blooms and swarms - Surplus organics - Detrital pathways - Low species - High population - Most diverse - Special types - Exorheic lakes, ICY sea shelf, Polynya, Fast Ice, pack ice and Subpolar.

... ...

ses...

SEMINAR/WORKSHOP ON STATISTICAL ECOLOGY

SCHOOL OF ENVIRONMENTAL SCIENCES

UNIVERSITY OF POONA

16- 22 DECEMBER, 1985

SPONSORED BY

UNIVERSITY GRANTS COMMISSIONS

AND

DEPARTMENT OF ENVIRONMENT

QUANTITATIVE ANALYSIS IN MARINE ECOLOGY

Arun Parulekar

NATIONAL INSTITUTE OF OCEANOGRAPHY

GOA

Quantitative Analysis In Marine Ecology

Arun Paruleker
National Institute of Oceanography
GOA

Sea as a Biological Environment - Potentials and Economic importance of self generating living resources - Multi purpose utility and conflict of interest - significance of quantitative studies for the estimation of potential resources towards the development of relevant industries - long term projections of sustainable yield - Conservation of resources and monitoring of environment.

Practical difficulties in operating a statistically valid sampling design - wide expanse of the study ara - three dimensional environment - Intriguing nature of air - Land - Sea - Inter - Action - Wide spatial and temporal variability - Necessity of a multidisciplinary approach - High cost of sampling time - Realistic approach.

Ecosystem Approach - Food chain - Trophism - Energy pathways - Ecological efficiency of conversion - Gross and net production - P/B Ratio - Herbivore carnivore - Detritus linkage - Consumer concept - Plankton - Nekton - Benthos link - Degradation and Replinishment.

Abiotic Components - Meteorological, Physical, Chemical and Geological regimes - Geographical significance - Variability - Magnitude of interactions - Predictability.

Biotic Characterisation - Horizontal and Vertical gradients of distribution - Abundance of representative types (Picco, Nanno, Micro, Meio and Macro) - Community - Hierarchy - Affinity - Similarity - Diversity - Succession - Migration - Endemism - Survival, Mortality and Generation time - Turnover rate and Community metabolism.

SEMINAR/WORKSHOP ON STATISTICAL ECOLOGY

SCHOOL OF ENVIRONMENTAL SCIENCES

UNIVERSITY OF POONA

16- 22 DECEMBER , 1985

SPONSORED BY

UNIVERSITY GRANTS COMMISSIONS

AND

DEPARTMENT OF ENVIRONMENT

THE TRANSMISSION AND SENSING OF CHEMICAL SIGNALS

Vidyanand Nanjundiah

TATA INSTITUTE OF FUNDAMENTAL RESEARCH

BOMBAY 400 005.

THE TRANSIMISSION AND SENSING OF CHEMICAL SIGNALS

Vidyanand Nanjundiah
Tata Institute of Fundamental Research
Bombay 400 005.

Communication by means of chemical signals (e.g.pheromones) is widespread in living organisms, both plants and animals; amongst microbes it is the only means of long-range signalling. Human beings recognise chemicals via the sensory modalities of smell and taste (by the way, smell and taste do not differ only in respect of whether the stimulus is present in solution or not, as is seen by the fact that fish have both capabilities). These talks will be concerned with two aspects of chemical signalling: the mode of transmission of the signal, which is usually free diffusion with or without mass flow; and the primary event in sensing the signal, which in many (not all) cases is known to be a reversible chemical reaction involving the binding of one or more molecules of the signal to one- and perhaps, though this isn't known, more-molecules of a specific receptor. The aim of the talks will be to highlight certain features and implications of the models rather than to provide text-book descriptions. We will end by discussing a specific communicating system, the social amoeba Dictyostelium discoideum.

Talk 1:Diffusion

INTRODUCTION

Modern studies of diffusion start with Einstein's Ph.D. thesis, or more correctly, with the papers he wrote based on his thesis. As a result of this- and subsequent work by many others - three different situations were shown to have a common mathematical description : Brownian movement, the problem of random walks, and diffusion in dilute solutions. There are two underlying ideas. One, with its links to the kinetic theory of gases and thermodynamics, is that systems consisting of sufficiently many units can exhibit behavioural regularities even when the behaviour of a single unit is unpredictable (varies in a "random" fashion). The second is that the parameter which characterises these regularities- the coefficient of diffusion - is related to the frictional properties of the medium, the viscosity.

RESULTS

1. In an ideal gas at temperature T (absolute), a molecule of mass m has a mean-square velocity $\overline{v^2}$ given by $\frac{1}{2} m \overline{v^2} = \frac{3}{2} kT$; $k = 1.38 \times 10^{-16}$ erg/ ^{o}K . So at "room temperature" ($300^{o}K$) and for a molecule of nitrogen a typical speed is $\sqrt{\overline{v^2}} \simeq 5.10^{4}$ cm/sec. What will the corresponding speed be for

hydrogen ? Work out the thermal energy (in calories) of a nitrogen molecule (4.2×10^7 ergs = 1 cal) and of $N_{AV} = 6.10^{23}$ (=1 mole) nitrogen molecules. How does this compare with a chemical bond energy of a few kilo cal/mole ?

2. What is the speed of sound in air ? How does it compare with the speeds found earlier ?

3. A travelling molecule of radius 'a' will collide, each second, with all the molecules which lie within a cylinder of diameter 2a and length v. So the # of collisions per second is $f_{coll} = \pi a^2 v.n$ where n is the mean density of molecules. Then the average distance between collisions, or the mean-free path, is $\bar{l} = \frac{v}{(\pi a^2 vn)} = \frac{1}{(\pi a^2 n)}$ .Take $2a = 2.5 \times 10^{-8}$ cm. and $n = 3 \times 10^{19}$/cc show that $f_{coll} = 7.3 \times 10^8$/sec and $\bar{l} = 6.8 \times 10^{-5}$cm. In summary, a gas molecule performs a "walk" with an average step size 10^{-5}cm and abrupt and random changes in direction every 10^{-9}sec or so. Freely swimming E. coli do the same thing, but there are differences: their step length is 10μm (10^{-3} cm), and they change directions about once a second; more importantly, their movement is entirely autonomous.

4. Particles of dilute solution of one species in another will also perform a random walk. The mean square displacement $\overline{r^2}$ that they undergo in a time t is proportional to t:

$$\overline{r^2} = 2Dt \quad \text{(1-dimension)}$$
$$= 4Dt \quad \text{(2-dimensions)}$$
$$= 6Dt \quad \text{(3-dimensions)}$$

D is a constant known as the coefficient of diffusion.

D is releated to the viscosity of the medium η:

$D = \frac{kT}{6\pi\eta a}$ where a is once again the radius of the diffusing species. The c.g.s. unit on η is :'poise'; $\eta = 10^{-2}$ poise for water and 1.8×10^{-4} poise for air at $300^{o}K$. Note that the relation between $\overline{r^2}$ and t immediately implies that it is meaningless to talk of the speed of diffusion. Nonetheless, one sometimes says (loosely) that diffusion is "very fast over short distances".

5. The distribution which gives rise to $\overline{r^2} = 2Dt$ etc. can be expressed as follows : If M molecules of a certain species are concentrated at the origin at t = 0, their concentration at some other time t will be given, as a function of the distance r from the origin, by

$$C(r,t) = \frac{M}{(4\pi Dt)^{n/2}} \exp(-r^2/4Dt)$$

where n is the dimension.

6. Work out D for nitrogen in air ($\eta = 1.8 \times 10^{-4}$ poise) and in water ($\eta = 10^{-2}$ poise) at about 300°K. Would it make much of a difference if it were a pheromone (M.W = 300) rather than nitrogen ?

7. Some values of diffusion times for a small molecule (we take $D = 10^{-5}$ cm^2/sec in water and $D = 10^{-1}$ cm^2/sec in air) :

across a nerve-muscle synapse ($r \sim 100$ A°) 5.10^{-8} sec

one end of a bacterium to another ($r \sim 1\mu m$) 5.10^{-4} sec

one end of an emoeba to another ($r \sim 10\mu m$) 5.10^{-2} sec

nose to lungs ($r \sim 10$ cm) 500 sec

one end of room to another ($r \sim 500$ cm) 400 hr.

How long does it actually take to smell something across the distance of a room ? What do you infer ?

8. Steady and periodic release :

The result of 5 can be generalised for the case of a release rate Q(t) (molecules/sec) at the origin, giving

$$C(r,t) = \int_0^t dt^1 Q(t^1) \frac{1}{[4\pi D(t-t^1)]^{3/2}} \exp [-r^2/4D(t-t^1)]$$

$$= \int_0^t dx\, Q(t-x) \frac{1}{(4\pi Dx)^{3/2}} \exp (-r^2/4Dx)$$

Case (i) Steady release $Q = Q_0$, a constant

Then $C(r,t) = \frac{Q_o}{4\pi Dr}\ \mathrm{erfc}(\frac{r}{\sqrt{4Dt}})$ where erfc is known as the complementary error function; its value drops very rapidly from 1 when the argument is 0 to 0 when the argument $\to \infty$. Finally, for sufficiently large times ($t >> \frac{r^2}{4D}$), $C(r,t)$ equals $\frac{Q_o}{4\pi Dr}$ nearly. Note that we have assumed understricted diffusion in 3 dimensions; the matter is tricky in 1 and 2 dimensions (this is related to the problem of return to the origin in random walks of different dimensions). If the source is on the ground, the ground itself becomes a reflecting boundary, doubling the value of C to $\frac{Q_o}{2\pi Dr}$.

Case (ii) Periodic release Suppose $Q = Q_o\ (1+\cos wt)$. Once again, for times $t >> \frac{r^2}{4D}$ from the start, and assuming the ground to be a reflecting barrier,

$$C(r,t) = \frac{Q_o}{2\pi Dr}\ [1 + e^{-qr}\cos(wt-qr)] \text{ where } q = \sqrt{\frac{w}{2D}}$$

C continues to be periodic in time but assumes the form of a damped oscillation in space.

Numbers : In the case of the moth Utetheisa (Conner et al) one can make the following estimates. The active ingredient is apparently $CH_3\ CH_2\ CH = CHCH_2\ CH = CHCH_2\ CH = CH(CH_2)_{10}CH_3$ this gives a molecular weight of 290 and so a diffusion coefficient D in air of .056 cm^2/sec. Females pulse every evening

for 28.6 $\pm$ 1.6 min at a frequency of 1.5 $\pm$ 0.2 pulses/sec. (so w = 9.4/sec). 32 ng of active extract was recovered per female per scenting period. This gives us a rate of release Q_o of 5.10^{10} molecules per pulse or 3.3×10^{10} molecules/sec. Putting all this together, q = 9.16 cm^{-1}, and (at r=1 cm) $\frac{Q_o}{2\pi Dr}$ = 9.4×10^{10} molecules/cm^3 = 1.6×10^{-10} M nearly. The conclusion is that spatial variations in the signal get damped beyond about 1/q $\approx$0.1 cm. and that the steady concentration reached thereafter is very low indeed (Wilson's estimates of the threshold concentrations of alarm pheromones in the ant Acanthomyops range from 10^{12} molecules/cm^3 to 10^{16} molecules/cm^3).

Remark : Our expression for C shows the q increases with w. What does this imply ?

9. The Effect of Wind (Bossent, 1968).

For a wind speed v in the x direction one has to modify the earlier expression for C(r,t). Now the downstream concentration is given by

$$C(r,t) = \int_o^t dt^1 \, Q(t^1) \, \frac{1}{[4\pi D(t-t^1)]^{3/2}} \exp - \left[\frac{(x-v(t-t^1))^2+y^2+z^2}{4D(t-t^1)}\right]$$

where $r^2 = x^2 + y^2 + z^2$.

If Q is a constant,

$$C(r,t) = \frac{Q_o}{2\pi Dr} \exp\ [-(r-x)V/2D].$$

If $Q = Q_o\ (1 + \cos wt)$,

$$C(r,t) = \frac{Q_o}{4\pi Dr}[\exp\ (xv/2D)\ \exp(-^{rv}/2D)+\exp(-\alpha\cos\beta)\cos(wt-\alpha\sin\beta)]$$

$$\text{where } \alpha = \frac{vr}{2D}\ [1 + \frac{16w^2D^2}{v4}]^{\frac{1}{4}}.$$

$$\text{and } \beta = \frac{1}{2}\ \tan^{-1}\ (\frac{4wD}{v^2}).$$

These expressions look complicated, but we will show that they can by analysed rather easily to give estimates of ranges of attraction and active spaces (Wilson). Basically, the conclusions are that

(i) reasonable values for range are quite small, probably more than a few cm. (in air);

(ii)for the same average rate of release (= energy expenditure), patterning increases the range; and

(iii) for small wind speeds the range increases with wind speed but at larger speeds the range falls off on account of reasons (turbulent diffusion) beyond the scope of our model.

Talk 2

Binding and diffusion

We will assume that after travelling from source to receiver the chemical signal is sensed by specific receptors. Receptors are molecules capable of binding the signal rather specifically. Either the number of bound receptors or some function of the binding process constitutes the first step in signal detection.

1. Binding rates

$$A + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} AB$$

k_1 A.B is the rate at which the bound complex AB is formed;

k_{-1} AB is the rate at which it is borken up.

k_1(units: 1/conc. x time) is determined by the product of two factors, the frequency of A-B collisions and the fraction of such collisions which "succeed" in the sense of an AB complex being formed.

Therefore an upper limit for k_1 is given by the diffusional flux of 'A' molecules onto 'B' molecules (which, we shall assume, are immobile):

$$(k_1)_{max} = 4\pi D_A \sqrt{B}$$

Take $D_A = 10^{-5}$ cm^2/sec (water)

or 10^{-1} cm^2/sec (air)

and $r_B = 50$ Å

Then $(k_1)_{max} = 4 \times 10^{-5} \times 5 \times 10^{-7}$ cm^3/sec in water

$6 \times 10^{-11} \times 6 \times 10^{20}$ M^{-1} sec^{-1}

$= 3.6 \times 10^{10}$ M^{-1} sec^{-1}.

Measured values of k_1 are 1-2 orders of magnitude smaller. The value of k_{-1} needs a knowledge of binding energies for its calculation; experimental measurements are sometimes easier, A typical value of k_{-1} might be 1 sec^{-1} for rapid dissociation or 10^{-2} sec^{-1} for slow dissociation. $\frac{k_{-1}}{k_1} = K_D$ is called the equilibrium constant of the reaction and is a measure of how 'tight' the binding is. To see why, we compute the

2. Equilibrium binding

At equilibrium, $k_1 . A . B = k_1 [AB]$

Suppose we have a great excess of receptors (Bs) compared to signal molecules (As).

But $A = A_{total} - [AB] = A_o - [AB]$

$B = B_{total} - [AB] = B_o - [AB] \approx B_o$, say

Then $(A_o - [AB]) B_o = K_D [AB]$, which on simplification given

$$AB = A_o \left\{ \frac{B_o}{B_o + K_D} \right\} .$$

Show that if things are the other way round, i.e. if the signal is in excess,

$$[AB] \approx B_o \{ \frac{A_o}{A_o + K_D} \}.$$

In either case, 50% of the less common partner gets bound when the more common one is at a concentration of K_D; the binding keeps increasing with the concentration of the more common substance.

Exercise Suppose that (in appropriate units)

$A_o = 1$, $B_o = 10$ and (i) $K_D = 10^{-2}$, (ii) $K_D = 10^2$

Compute [AB].

3. How many receptors does a cell need ?

Consider a cell of radius a immersed in a medium containing a concentration C of some molecular species. We will assume that all molecules which hit the surface stick to it (or otherwise disappear), $c(r) = 0$ at $r=a$. Sufficiently far from the cell, say $C(r) = C_o$. Then it is easy to see that $C(r) = C_o(1 - \frac{a}{r})$. It follows that the flux on to the cell surface, which under our conditions gives an upper limit to the number of molecules the cell can detect (per unit area and time), is $D\left|\frac{dc}{dr}\right|_{r=a} = \frac{D.C_o}{a}$.

Therefore the total binding rate is limited by $\frac{DC_o}{a} \cdot 4\pi a^2 = 4\pi DC_o a$. Note that this expression is linear in a. Now let us consider a different situation; one in which not the entire cell, but a single receptor - a disc of radius s - can similarly bind the ligand. For $s << a$, one can show that the maximum rate of binding is now $4DsC_o$. If there are N such receptors, the rate becomes $N.4DsC_o$. Clearly, this cannot continue to hold good for indefinitely large N; not only will the receptors interfere with one another, their combined area will threaten to increase beyond the cell's surface area.

The appropriate expression, which for
$N.\pi s^2 << 4\pi a^2$ is nearly $N.4DsC_o$, and which for
$N.\pi s^2 >> 4\pi a^2$ is nearly $4\pi DC_o a$, turns out to be

$$4\pi DC_o a \frac{N.4DsC_o}{N.4DsC_o + 4\pi DC_o a} = 4\pi DC_o a \frac{NS}{Ns + \pi a}$$

From the expression $\frac{NS}{Ns+\pi a}$, one sees that 50% efficiency in binding is already achieved when $Ns = \pi a$, that is, $N = \frac{\pi a}{s}$. Consider a cell of radius 5μm; suppose s=50 A^o, then $N \approx 3.10^3$ receptors are sufficient. This is really a small number, as is seen by the fact that the fraction of the cell's surface covered by receptors is $\frac{N.\pi s^2}{4\pi a^2} = \frac{\pi^2}{4N}$ at $Ns = \pi a$. With our figures,

this fraction is just about 1/300 = 1/3%. The very important implication is that a cell can distribute a large number of different sensing devices on its surface and still leave a substantial part of its surface free for other purposes (e.g. transport).

Exercise Treat the same problem for clustered receptors (in the limit, all N receptors would be combined into one giant patch).

4. Energetics

Thermodynamics tells us that in a reaction of the type

$aX + bY \rightleftharpoons cZ$ and at equilibrium the ratio $\frac{n_x^a \cdot n_y^b}{n_z^c} = K$ is a constant depending (under idealised conditions) only on the pressure and the temperature. Here n_x, n_y and n_z are the respective mole-fractions of the species X,Y and Z; that is, if there are X molecules of X,------ in the system,

$$n_x = \frac{X}{X + Y + Z} \text{ , etc.}$$

Kis called the equilibrium constant of the system and is related to the free energy change occurring in the reaction:

$$\ln K = -\frac{1}{RT}(a g_x + b g_y - c g_z)$$

where the g s are free enthalpies (Cibbs free energies) for single moles

"standard" conditions - of X,Y and Z respectively taken individually. So the actual free energy change,

$$ag_x + bg_y - cg_z \text{ , equals } - RT \ln K.$$

This relationship is very popular as an exam question and, as a consequence, is widely misused. It is important to remember the assumptions involved in its derivation (see a good text on chemical thermodynamics).

Example : Consider the binding of cyclic AMP, an aggregation pheromone in Dictyostelium discoideum, to cell surface receptors. Represent this by $C + R \rightleftharpoons (CR)$ where C refers to cAMP and R to the receptor. Saturation experiments show that each cell has about 10^5 receptors. At 2.10^8 cells/ml and a cAMP concentration of 10^{-6}M, 50% of the receptors are occupied. The total receptor concentration is $\frac{2.10^8 \times 10^5}{6.10^{20}}$ M = 3.3×10^{-8}M, so that [CR] and R both equal 1.65×10^{-8}M, C equals 10^{-6}M - 1.65×10^{-8}M = 0.9835×10^{-6}M. Then the mole fractions are :

$$n_C = \frac{.98}{.98 + .033} = .97$$

$$n_R = \frac{.0165}{.98 + .033} = .015 = n_{CR}$$

$$\text{and} \quad -RT \ln K = -RT \ln \frac{.97 \times .015}{.015}$$

$$= -2 \frac{\text{Cal}}{^{o}K} \times 300^{o}K \times (-.03)$$

18 cal.

(useful figure : at "room temperature" RT 0.6K cal/mole)

What does this imply - particularly with reference to the value of RT ?

The standard free energy change, on the other hand, is $-RT\ln 10^{-6} = 8.3$ K cal nearly. This should be compared with chemical bonding energies, which are in the range of 20 to 200 K cal (per mole). What do you infer ?

5. Signal and Noise

Consider once again a ligand, present at a free concentration, C, and a receptor, present at a free concentration R; the concentration of the legand-receptor complex is denoted CR . As we have seen, under certain conditions $CR = R_o \frac{C}{C + K}$ where R is the total receptor concentration and K the equilibrium constant. This relationship can be interpreted to mean that the elementary binding event has a probability $p = \frac{C}{C+K}$ of occuring and a probability (1-p) of not occurring. νR_o can be interpreted as the equivalent of the 'number of trials' and CR as the mean binding.

Then the variance in binding will be given by $\nu R_o \cdot p \cdot (1-p)$ and the standard deviation by its square root. Therefore a measure of signal to noice is given by

$$\frac{S}{N} = \left(\frac{\nu R_o C}{C+K}\right) \Big/ \sqrt{\frac{\nu R_o KC}{(C+K)}} = \sqrt{\frac{\nu R_o C}{KK}}$$

with ν being the reaction volume.

We will examine certain implications of this result with particular reference to experimental measurements on the slime mold Dictyostelium. Also, we will look into the problem of measuring a spatial gradient of the signal (chemotaxis). A brief mention will be made of probabilistic approaches to binding and how they help in modelling the phenomenon of cooperativity between receptors.

... ...

ses...

SEMINAR/WORKSHOP ON STATISTICAL ECOLOGY

SCHOOL OF ENVIRONMENTAL SCIENCES

UNIVERSITY OF POONA

16- 22 DECEMBER, 1985

SPONSORED BY

UNIVERSITY GRANTS COMMISSIONS

AND

DEPARTMENT OF ENVIRONMENT

SOME MODELS IN MICROBIOLOGY

M.G. WATVE

DEPARTMENT OF MICROBIOLOGY

GARWARE COLLEGE, PUNE 4.

Some Models in Microbiology

M.G. Watve
Department of Microbiology
Garware College, Pune 411 004

(Summary)

Microbiologists seem to have developed different and distinct lines of thinking as compared to other biologists. Some of the microbiological approaches will be useful in biology and vice-versa; however there appears to be a communication gap. A microbiologist primarily deals with populations, while most other biologists deal primarily with individuals.

Counting rare species by MPN method :

The multiple tube fermentation technique or the most probable number (MPN) method is used to count coliform bacteria in water. A set of media tubes are innoculated with definite aliquotes of water sample. The resultant tubes are categorized as positive and negative and an estimate of the count is obtained. A standard table is available giving various combinations of positive and negative tubes and the respective MPN values. eg.

Most probable number of coliforms present in 100 ml of water sample.

No.of positive lactose broth tubes out of 5 tupes innoculated each:			No.of coliforms/ 100 ml.
10 ml	1 ml	0.1 ml	
0	0	0	0
0	1	0	2
0	1	1	4
1	0	0	2.2
1	0	1	4
1	1	0	4.4.
2	0	0	5
2	0	1	7
2	1	0	7.6
3	0	0	8.8
3	0	1	11
5	5	5	1800+

The test is based on the following assumptions :

(1) The coliform bacteria are distributed randomly

(2) If a tube receives one or more than one coliforms it gives positive results. If and only if a tube fails to receive any, it is negative.

Random distribution allows the use of Poisson series : The theoretical probability of getting zero organisms in the aliquote is

$$P_o = e^{-\mu}$$

where μ mean no. of organisms in the aliquote used.

P_c = the probability of getting negative tube.

$1-P_o$ = probability of getting positive result.

Once we have probabilities of two mutually exclusive events, probabilities of different combinations of positive and negative tubes can be plotted. On the other hand the probabilities of getting a single combination can be plotted against different values of μ. This plot will have a single peak. This peak is the MPN value.

In practice three sets of different aliquotes are used and the combined probability is plotted, against the mean density of organisms to increase the accuracy.

The principle of the method is applicable with due modifications to many situations where the estimate of a rare species is required, provided, the distribution is fairly random. For clumped distributions it gives an under estimate.

The Chemostat Model : The chemostat model is an 'open system' laboratory model, which tries to simulate growth condition in nature. To a fixed quantity of culture vessel, fresh nutrient medium is added at a fixed rate. Automatically the same quantity flows over. This dilutes the culture at an exponential rate.

If only growth occurs, population after time t will be given by the exponential equation.

$$X_t = X_o e^{\mu t}$$

where X_o = initial population.

and μ = specific growth rate.

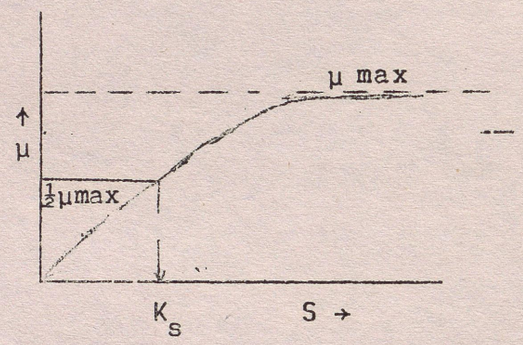
If only dilution is allowed, population after time t will be

$$X_t = X_o e^{-Dt} .$$

It follows that, when μ and D are numerically equal, X_t will be constant, this achieves a self adjusting steady state. If D increases, the availability of nutrients increases resulting into increase in μ till it reaches D.

However the growth rate has a genetically determined upper limit μ max. The relation of to the substrate concentration, similar to the relation of an enzyne towards its substrate is given by

$$\mu = \mu \text{ max } \frac{S}{K_s + s}$$



By working out the steady state state Kinetics in a chemostate we get the population as

$$\bar{X} = Y \quad S_R - [K_S \left(\frac{D}{\mu max - D}\right)] .$$

where Y = the yield value giving the population growth per unit substrate consumed.

and S_R = substrate concentration in the fresh medium.

A chemostat model is of interest to both commercial microbiologists and ecologists. It simulates an ecosystem being

an open system; replenishment of resources simulated by the medium inflow and death simulated by washout of cells.

Chemostat also supplies a working model for commercial microbiological processes and the continuous type of effluent treatment plants.

... ...

ses...

SEMINAR/WORKSHOP ON STATISTICAL ECOLOGY

SCHOLL OF ENVIRONMENTAL SCIENCES

UNIVERSITY OF POONA

16- 22 DECEMBER ,1985

SPONSORED BY

UNIVERSITY GRANTS COMMISSIONS

DEPARTMENT OF ENVIRONMENT

CLUSTER ANALYSIS

P.V. Nair

Kerala Forest Research Institute

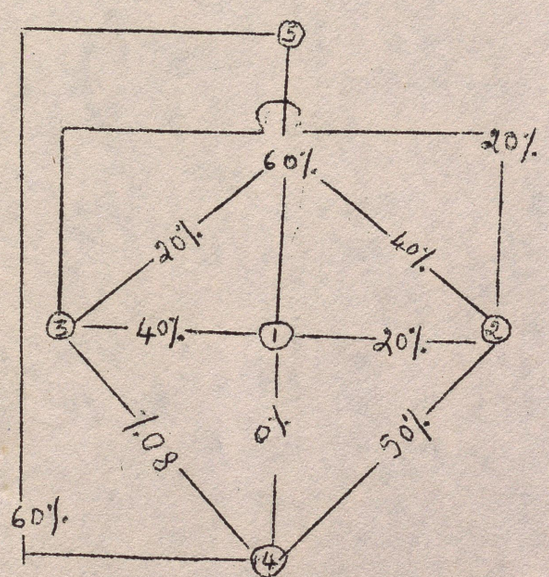
Peechi Karala

CLUSTER ANALYSIS

P.V. Nair
Kerala Forest Research Institute
Peechi Kerala.

Clustering classifies individuals or ojbects into groups. This method is usually employed when the relationship between the individuals are not apparent. In other words this technique tries to establish patterns from a given set of data. The technique generate a large number groups, some groups having only few individuals while others having many members at various levels of relation. We choose the level we require and test whether the group is significantly different from what is possible by chance alone.

Clustering techniques are used in many different fields such as library sience, behaviour analysis, ecology etc. Suppose there are five individuals with relation as shown



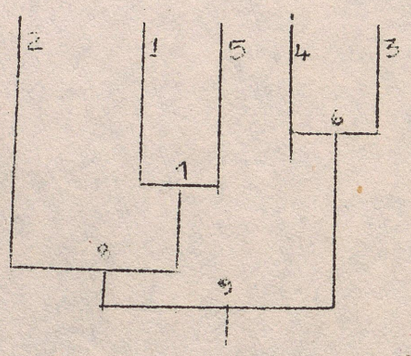
Individuals	1	2	3	4	5
1		.2	.4	0.	.6
2	.2		.2	.5	.4
3	.4	.2		.8	.2
4		.5	.8		.6
5	.6	.4	.2	.6	

Individuals

Similarity matrix

	1	2	3	4	5	6	7	8
1	\	.2	.4	0	.6	.2	.3	.25
2	.2	\	.2	.5	.4	2.5		
3	.4	.2	\	.8	.2			
4	0	.5	.8	\	.6			
5	.6	.4	.2	.6	\	.4		
6	.2	2.5			.4	\	.3	
7	.3					.3	\	
8	.25							\

(6) = 3 and 4 = 0.8
(7) = 1 and 5 = 0.6
(8) = 2 and 7 = 0.3
(9) = 6 and 8 - 0.25



The similarity matrix

If the similarity or dissimilarity between individuals are known, the hierarchic relation between the individuals can be computed. The similarity values can be derived in a variety of ways.

(1) Co-relation coefficient

(2) % similarity

(3) % distance table

(4) Co-occurrence

(5) Co-citation etc.

The similarity value can also be derived from multiple factors.

Habit	Species				
	1	2	3	4	5
Ground Dwell	20	80	30	40	60
Canopy Dwell	80	20	70	60	40
Feed on Grass	10	25	80	5	0
Feed on insects	90	75	20	95	100

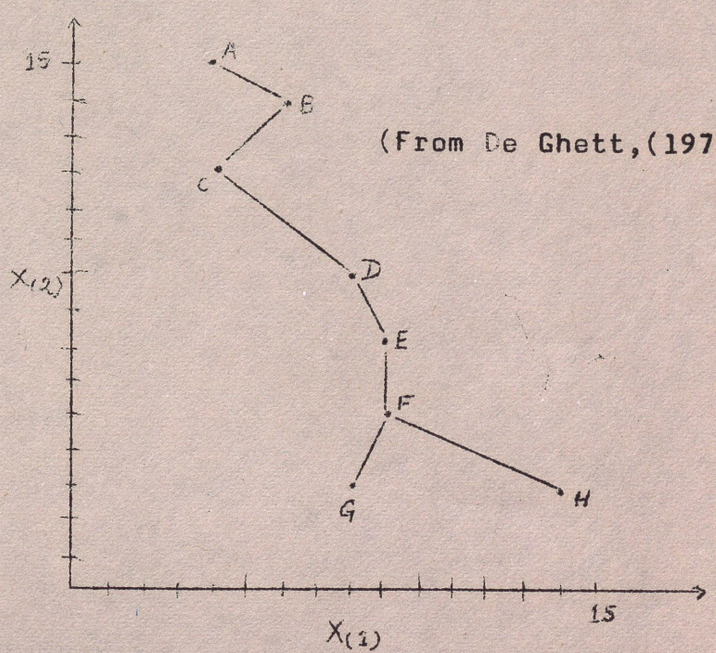
Species : 1 and 2 = 1/4 + 1/9 + 1/4 + 1/25 + 7.5/9

2 and 3 = 2/3 + 2/7 + 2.5/8 + 2/7.5

Distance

Consider eight individuals with two measures for each individuals

	$X_{(1)}$	$X_{(2)}$
A	4	15
B	6	14
C	4	12
D	8	9
E	9	7
F	9	5
G	8	3
H	14	3



(From De Ghett,(1972).

The similarity between the individuals can be defined as the distance between them

$$D_{AB} = \sqrt{(4-6)^2 + (15-14)^2} = \sqrt{5} = 2.2.$$

$$D_{AC} = \sqrt{(4-4)^2 + (15-12)^2} = \sqrt{9} = 3$$

A	B	C	D	E	F	G	H
A	2.2	3.0	7.2	9.4	11.2	12.7	15.6
B		2.8	5.4	7.6	9.5	11.2	11.6
C			5.0	7.1	8.6	9.9	13.5
D				2.2	4.1	6.0	8.5
E					2.0	4.1	6.4
F						2.2	5.4
G							6.0
H							

Similarity matrix from association:

Depending on the range of values needed and situation one of the following may be chosen.

	Assume
a = 1, 1 matches	4
b = 1, 0 "	4
c = 0, 1 "	3
d = 0, 0 "	1
m = a + d "	5
u = b + c "	7
n = m + u "	12

(1) Jaccards' association coef. $= \frac{a}{a+u} = 0.36$

(2) Dice's association coef. $= \frac{2a}{2a+n} = 0.53$

(3) Sokal and Michener's association coef. $= \frac{m}{n} = 0.42$

(4) Rogers' and Tanimoto's association coef. $= \frac{m}{m+2n} = 0.26$

(5) Yule's association coefficient $= \frac{ad - bc}{ad + bc} = -0.50$.

Dendrogram

The branching, tree like representation of the relation is called a dendrogram. Significance also can be shown on the dendrogram.

Limitations

The clusters are just groups. Their significance has to be explained from experience. Meaningless clusters also can result.

Significance test

In order to test whether the clusters obtained in statistically significant, i.e. whether they are more important than what is possible by combinatorial chance a Monte Carlo method is used. The computer is asked to generate all possible combinations of clusters. The distinctiveness (Sokal and Rhlf 1969) of all these clusters is compared with the same of the cluster we want to test.

5. Garfield, E. 1980

ABC of cluster mapping. Current contents 11 (40) : 5 - 12

11 (41) : 5 - 12

6. Dawkins, R. 1976.
Hierarchial organisation : a candidate principle for ethology. In growing points in ethology. Ed. Batesan, ppG and Hinde, R. A. Cambridge University Press.

7. Morgan, B.J.T. et al 1976. Visualising interactions and sequential data in animal behaviour. Theory and application of cluster analysis methods behaviour 56 : 1-43.

References

1. Cody, M.L. 1974.
Competition and the structure of bird communities.
Princeton, University Press. New Jersey. p 91-76; 278-788

2. Sokal, R.R. and Rohlf, F.J. 1969.
Biometry. W.H. Freeman and Co. Sanfrancisco.

3. Sokal, R.R. and Snealth, P.H.A. 1963.
Principles of numerical taxonomy W.H. Freeman and Co.
Sanfransisco.

4. De Gett, V. 1978. in COLGAN P.W.
Hierarchical cluster analysis In Quantitative Ethology
John Wiley and Sons, New York.

... ...

ses...

SEMINAR/WORKSHOP ON STATISTICAL ECOLOGY

SCHOOL OF ENVIRONMENTAL SCIENCES

UNIVERSITY OF POONA

16- 22 DECEMBER , 1985

SPONSORED BY

UNIVERSITY GRANTS COMMISSIONS

AND

DEPARTMENT OF ENVIRONMENT

REGRESSION ANALYSIS

Anand Karandikar

Indian Institute of Cost and
Management Studies and Research

1195/ 4 B Fergusson College Road,
PUNE 411 004.

REGRESSION ANALYSIS

INTRODUCTION :

ANAND KARANDIKAR

In a study of the relationship between two variables X and Y, there are basically two problems involved. The first is to find a measure of the degree of association or correlation between the values of X and Y. The second is to find the most suitable form of an equation for use in predicting the average, Y for a given X or in predicting the average X for a given Y, and to estimate the error in such Predictions. This latter problem is the problem of regression. In regression, we try to quantify the dependence of the variable on the other. Once this is done, we can know the changes in one variable induced by changes in the other variable.

The regression of Y on a single independent variable is often inadequate. Two or more X's may be available to give additional information about Y by means of a multiple regression on the X's. Among the principal uses of multiple regression are :

1) Constructing an equation in the X's that gives the best prediction of the values of Y.
2) When there are many X's, finding the subset that gives the best linear prediction equation. In predicting future weather conditions at an airport, there may be as many as 50 available X-variables, which measure different aspects of the present weather pattern at neighboring weather stations. A prediction equation with 50 variables is unwieldy, and is unwise if many of the X-variables contribute nothing to improved accuracy in the prediction. An equation based on the best three or four variables might be wise choice.
3) In some studies the objectives is not prediction, but instead to discover which variables are related to Y, and, if possible, to rate the variables in order to their importance.

Multiple regression is a complex subject. The calculations become lenghty when there are numerous X- variables, and it is hard to avoid mistakes in computation. Standard electronic computer programs, now becoming more readily available, are a major help. Equally important is an understanding of what a multiple regression equation means and what it does not mean.

Lecture Notes for :
Seminar/workshop on Statical Ecology, December 1985 at School of Environmental Sciences, University of Pune. Based on :'Cost Benefit Analysis of Cow Cross Breeding Programme in Maharashtra', INDSEARCH, PUNE, 1985.

SOME BASIC STASTICAL CONCEPTS :

The most common method employed to estimate the values of parameters in an equation is called the method of least squares. It estimates the values of parameters for which the sum total of squared deviations between observed and estimated values of the independent variable are minimised.

Either time-series or cross-sectional data may be used to estimate an equation. The larger the difference between the number of observations and number of variables used in the model, the more is the reliability of the model. Generally, extreme and exceptional observations are eliminated before the data is used for estimating an equation.

The explanatory power of the model is judged by the ratio of total explained variation to total observed variation in the independent variable. This ratio is generally represented by the symbol R^2. The value of R^2 ranges between 0 to 1, and the closer it is to 1, the better is the explanatory power of the model.

The possible variation in the estimated values of parameters with different samples judged by calculating the standard error in values of the parameters. The value of the standard error is generally written in brackets below the value of the parameter when presenting the estimated form of the equation. The value of the standard error should be small compared to the value of the parameter f r the parameter to be significantly different from zero. As a thumb rule, the standard error should be less than half the parameter value.

While deciding whether the estimated equation should be used for analytical purposes, the following aspects should be carefully considered :

1. The estimated form should be consistent with the theoretical and intuitive understanding of the relationship between the variables.
2. The magnitude and signs of the parameters should be consistent with theoretical and intuitive understanding.
3. The R^2, that is the explanatory power of the equation, should be sufficiently high. With larger samples, a smaller R^2 may still be significant.

4. The standard error associated with parameters should be small in comparison with the parameter value.
5. The independent variables should not be highly correlated amongst themselves. This is known as the problem of multi-collinearity.
6. The unexplained portion of variation in each of the observations, that is the error terms, should not be correlated amongst themselve.. This often occurs in time-series data and is known as auto-correlation.

If any one of the above conditions is not observed, it is often necessary to revise the model. Amongst alternatively acceptable models, simpler models should be preferred to complex models.

While using the estimated equation for analytical purposes, it is important to bear in mind that the regression methodology by itself does not establish a causal relationship amongst the variables.

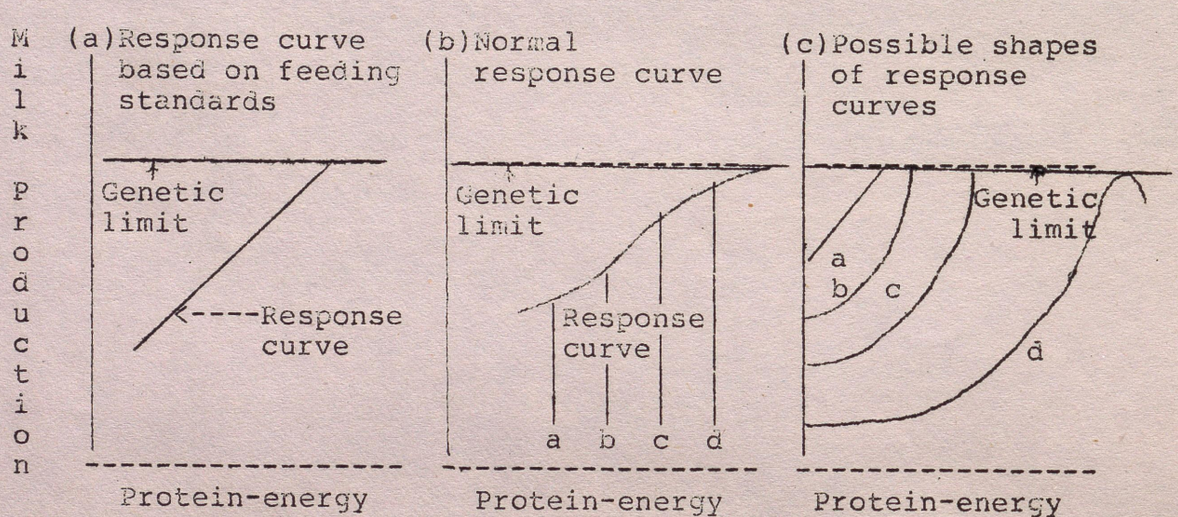
A) CASE STUDY :

The brief introductory remarks in the earlier paragraphs may be better understood with the help of a study which was recently carried out to find the relationship between feed intake and milk yield for cows in rural Maharashtra. The salient features of this study are presented in the following pages.

5.10 Response of Milk Yield to Feed : Among the various factors which affect the milk yield of a cow, the quantity and quality of feed are of extreme importance. Ceteris paribus, as the cow is fed more protein and energy, its milk yield will increase until it reaches the biologically possible limit. 'Scientific' feeding standards for cows are generally calculated by analysing the feed requirements for two major functions, viz., body maintenance and milk production, seperately. There are several standards used in different countries. The nutrient requirements in these standards are expressed in terms of digestible nutrients, metabolizable energy, starch equivalent, etc.. In India, Sen & Ray (16) have prescribed feeding standards for Zebu cattle and buffaloes. Studies on the nutrient requirements of different Indian bovine catagories have been initiated at IVRI, ~~NDRI~~ and other regional research centres. However, as rightly pointed out by the NDDB (17), these 'scientific' standards prove to be of limited use due to the following considerations.

"Diagram 1(a) below shows the milk yield response curve implicit in the feeding standards. It shows a constant increase of milk yield in response to unit increases in protein-energy until it abruptly reaches the genetic limit beyond which milk production cannot be increased by increasing the intake of proteins and energy. The response curve shown in 1(b) appears to be more realistic. It shows that when the animal is underfed (a-b), the efficiency of converting feed into milk is low as the bulk of the feed is used to increase the body weight to the normal level. Similarly, the conversion is less efficient as the milk yield reaches the genetic limit (d). In the intermediate range (b-c), there is greater efficiency of conversion of feed into milk.

Diagram 1



We are concerned with the shape of this curve because it

plays a very important role in determining optimal feed schedules. Diagram 1(c) on the previous page for instance, shows different response curves".

5.11 Our Model : Our primary interest is to ascertain the shape of the milk yield response curve for local and crossbred cows under field conditions.

As explained earlier, our data on feed inputs has some limitations. More specifically, we do not have accurate information about the nutritive content of each type of feed given to a cow. Our data is in terms of broad categories of feed such as dry fodder, green fodder and concentrates. Consequently, the response curves estimated from our data are only preliminary and indicative approximations. The results provide plausible hypotheses which need to be further investigated by controlled experiments.

We use the multi-variable regression technique to estimate the relationship between milk yield and feed for a particular breed.

Alternative Models : We initially postualated three alternative relationships between milk yield and feed.

(i) Linear model:
$Y = a + bD + cG + dC + eZ + u$

(ii) Constant elasticity model (log - log):
$\log Y = a + b\log D + c\log G + d\log C + e\log Z + y$

(iii) Semi-log model:
$\log Y = a + bD + cG + dC + eZ + W$

where the variables are :

Y = Milk yield of the cow (litres/day). No adjustments have been made for fat content. Milk consumed by calves is also not accounted for.

D = Quantity of dry fodder offered to cow (kgs/day). All types of dry fodder are assumed to be homogenous and of uniform quality. Wastage is not accounted for.

G = Quantity of green fodder offered to the cow (kgs/day).
The assumptions are similar to those made for dry fodder.

C = Quantity of concentrates given to the cow (kgs/day).

Z = Number of hours for which the cow is grazed in a day.

Extreme observations : To obtain results which represent

the central regressive tendency of the data, we identified the extreme observations. This was done by the following method. Initially, the equations were estimated using all the available observations. The error terms associated with each observation were calculated. Those observations for which the error terms were greater than their two sigma-limits were identified as extreme observations. The same equation was then estimated after eliminating these extreme observations. The procedure was repeated until no extreme observations could be identified. Upto four successive estimations were required to homogenize the data. Between 3% to 7% of the observations were eliminated by this procedure.

The eliminated observations were seperately investigated to seek causal explanations for their extreme values. This often revealed an error made by the field investigator while reporting quantities of feed and milk yield.

Multicollinearity : The combination of various types of feed forms a synergetic unit, wherein one type of feed influences the ability to eat and digest another. Consequently, there is a tendency to substitute one type of feed by another. For instance, if good grazing pastures are available, the quantity of green fodder fed in the stall would decline. Therefore some multicollinearity is expected amongst the independent variables. A substantial multicollinearity will reduce the reliability of the estimated coefficients which will be indicated by large standard errors associated with them.

Table 5.11 shows the simple correlation coefficients between the independent variables. As expected, a negative correlation is observed between the quantity of green fodder fed and the number of grazing hours. Similarly, there is negative correlation between the quantities of dry fodder and green fodder.

The extent of the correlation between the independent variables, though often statistically significant, is generally less than 0.5. Moreover, attempts to eliminate the multicollinearity by transforming, combining or eliminating independent variables did not improve the regression results.

Estimated Equations : We found that the results of the linear model were consistently better than either the semi-log or the log-log models. The linear model gave a higher R-square, together with coefficient signs that were consistent with a priori expectations. The standard errors associated with the coefficients were also comparatively lower in the linear model. We

Table 5.11

Simple Correlation Coefficients
(All zones, income groups, four representative months)

Breed :	Crossbred				Local			
Season :	Lean		Flush		Lean		Flush	
Month :	May	June	Sept	Oct	May	June	Sept	Oct
Correlation between :	(Figures in 1/100)							
Dry fodder & milk yield	0.2	6	-11	1	18	29	4	1
Green fodder & milk yield	37	50	37	23	32	24	22	17
Concentrates & milk yield	74	62	61	57	57	59	54	55
Grazing & milk yield	-13	-20	-10	-19	- 5	-27	-18	-30
Dry fodder & green fodder	-38	-36	-38	-26	-20	-32	-38	-27
Dry fodder & concentrates	-30	10	21	39	10	32	2	- 8
Dry fodder & grazing	42	56	- 9	-12	3	20	-29	-28
Green fodder& concentrates	35	57	13	4	33	8	24	27
Green fodder & grazing	-39	-44	-32	-28	-35	-31	-29	-37
Concentrates & grazing	-15	-16	- 4	-10	-11	-17	-13	-19
	...	...	...	...	...	...	...	...
1% significance value of simple correlation coefficients :	28	24	21	21	23	27	21	21
	...	...	...	...	...	...	...	...
Sample size (nos) :	95	117	159	155	124	97	136	139

therefore present the results of only the linear relationship in this section.

The model was estimated seperately for crossbred and local cows for each of the twelve months. The results of the regression were broadly similar within a particular breed and season. We therefore present and discuss the results for two typical months from each season, viz., May and June in the lean season and September and October during the flush season.

Results for Crossbred Cows : The equations estimated for the linear model are given below.

(i) May 1983 Sample size : 95

$$Y = -0.82 + 0.11\ D + 0.04\ G + 1.32\ C - 0.18\ Z$$
$$(0.02)\quad (0.03)\quad (0.09)\quad (0.08)$$

R square = 0.76
F statistic = 67.79

(ii) June 1983 Sample size = 117

$$Y = 1.35 + 0.07\ D + 0.17\ G + 0.51\ C - 0.10\ Z$$
$$(0.02)\quad (0.03)\quad (0.12)\quad (0.07)$$

R square = 0.61
F statistic = 40.07

(iii) September 1983 Sample size : 159

$$Y = 1.97 - 0.03\ D + 0.03\ G + 1.00\ C - 0.001\ Z$$
$$(0.01)\quad (0.01)\quad (0.08)\quad (0.04)$$

R square = 0.62
F statistic = 55.80

(iv) October 1983 Sample size : 155

$$Y = 3.22 - 0.06\ D - 0.02\ G + 1.20\ C - 0.16\ Z$$
$$(0.01)\quad (0.01)\quad (0.08)\quad (0.05)$$

R square = 0.64
F statistic = 52.20

Some of the important conclusions emerging from the above regression analysis are summarized below :

The values of the F statistic indicate that all the equations have significant R squares. The values of the R square indicate that around 60% to 75% of the variation in the milk yield can be explained by feed.

The comparatively low values of the standard errors indicate that most of the coefficients are significant. Moreover, the high values of the R square and low values of the standard errors indicate that multicollinearity may not be a major problem.

The values of the coefficients associated with grazing

are consistently low and negative. This suggests that grazing does not have a major impact on increasing the milk yield. The negative correlation may be because of the poor quality of grazing lands, so that the energy spent in grazing is not effectively compensated by the feed intake through grazing. However, the values of the standard errors associated with the grazing coefficients are relatively high. Further studies may be necessary to confirm the negative impact of grazing on milk yield.

Concentrates have the maximum positive and significant impact on milk yield. The values of the coefficient associated with concentrates range between 0.51 and 1.32. These coefficients indicate the incremental increase in the milk yield per kilogram of concentrate fed.

Both dry and green fodder have a signficant impact on the milk yield. However, the relationship between dry and green fodder and milk yield has one surprising feature. The equations estimated for the flush season reveal that during September, the contribution of dry fodder to milk yield is negative while that of green fodder is marginal. This phenomenon is observed to be more acute during October, when both dry and green fodder appear to reduce the milk yield.

Dry and green fodder are expected to increase milk yield during the flush season, particularly because of their greater availability. It is therefore reasonable to presume that the simple linear model involves a specification bias, at least for the flush season, and therefore needs to be altered.

In this context, it may be mentioned that a similar linear model was estimated in the study of the Indo-Swiss Project in Kerala by NDRI (18). This study also mentions occasional negative coefficients for dry and green fodder, but the matter was not further investigated by the authors.

5.12 Revised Model :

Some A Priori Considerations : One major reason for obtaining negative coefficients in the linear model may be that the response function of milk yield to dry and green fodder is curvilinear. The real response curve may have shape 'C' with a maximum at x, instead of the straight line 'A', as depicted in Diagram 2 on the following page.

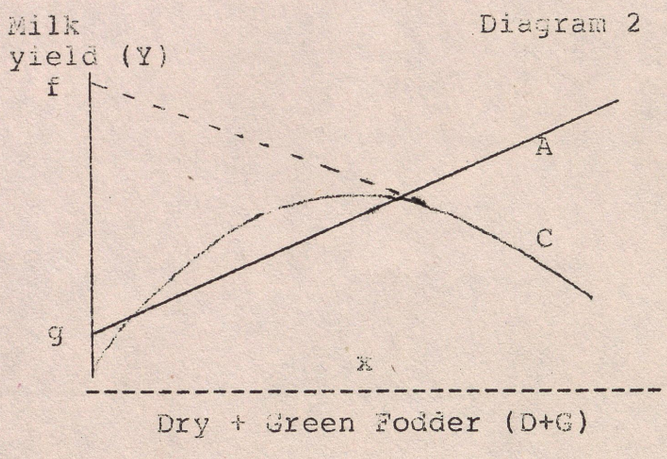


Diagram 2

If the relationship is non-linear, and if dry and green fodder are fed in quantities greater than x, a simple linear model will essentially estimate the downward sloping part of the curve C. This will result in negative values of the coefficients for dry and green fodder. Another consequence will be an overestimation of the value of the intercept (f instead of g).

A second factor contributing to the negative value of the coefficients could be the existence of multicollinearity between dry fodder, green fodder and grazing.

Taking these considerations into account, the model to estimate the relationship between milk yield and feed was revised in the following manner:

(i) To eliminate the effect of multicollinearity, grazing hours were eliminated from the equations because they had a marginal impact on milk yield.

(ii) To eliminate the multicollinearity between dry and green fodder, a new variable Q was defined as :

$$Q = D + w.G$$

where w is the relative weight of green fodder with respect to dry fodder.

The curvilinear response function was hypothesised as :

$$Y = e + m[Q] + n[Q.Q] + p[C] + q[C.C] + e$$

where Y and C are the quantities of milk yield and concentrates per day respectively and Q is as defined above.

If the relationship is curvilinear, we would expect a positive value for m and a negative value for n.

The equation was estimated for values of w ranging from 0.3 to 1.0. This sequential procedure revealed that the value of w which gives the best fit was within the range of 0.5 to 0.7. This implies that one kg of green fodder is equivalent to around 0.6 kgs of dry fodder (during the flush season). This value of w appears to be reasonable for the following reasons:

(i) The dry matter content of dry and green fodder is 30% and 90% respectively (19). The value of w calculated on the basis of dry matter equivalence alone would therefore be 0.33. However, the nutritive content of dry matter obtained from green fodder is greater than that obtained from dry fodder. Although the exact nutritive content varies between different types of dry and green fodder, the dry matter from green fodder contains roughly twice as much nutrient value (20). The value of w should therefore be approximately 0.66 if both nutrient value and digestibility are to be taken into account.

(ii) The market price of green fodder is roughly half the price of dry fodder. Taking into consideration their equivalence from a commercial point of view, the value of w should be approximately 0.5.

Results of the Revised Model : The following equations were obtained for alternative values of w during the flush season.

(v) For w = 0.75

$$Y = 2.19 + \underset{(0.027)}{0.024}[Q] - \underset{(0.0004)}{0.0007}[Q.Q] + \underset{(0.248)}{0.824}[C] + \underset{(0.038)}{0.049}[C.C]$$

R square = 0.63
F statistic = 55.80

(vi) For w = 0.50

$$Y = 2.37 + \underset{(0.030)}{0.025}[Q] - \underset{(0.004)}{0.002}[Q.Q] + \underset{(0.257)}{0.716}[C] + \underset{(0.041)}{0.075}[C.C]$$

R square = 0.63
F Statistic = 55.80

The value of the R square of the quadratic model is 0.63, which is not very different from the value obtained for the linear model. The explanatory power of the model is therefore practically unchanged even when grazing hours are totally eliminated and dry and green fodder are combined into a single variable.

The values of the coefficients associated with the linear variables (Q and C) are not very sensitive to values of w ranging from 0.5 and 0.75. However, the values of the coefficients associated with the quadratic terms ([Q.Q] and [C.C]) are sensitive to the value of w.

The values of m are greater than zero, while the values of n are less than zero. This supports our hypothesis that the response function of milk yield to dry and green fodder is curvilinear. Additionally, the values of the intercept of the quadratic model are lower than that of the linear model.

We find that the values of the coefficients associated with both linear and quadratic terms for concentrates are positive, indicating that within the range observed by us, the impact of concentrates on milk yield is a monotonically increasing function.

One possible source of the unexplained variation in the revised model estimated for crossbred cows is the type of the crossbred and the level of exotic inheritance. To understand their impact, we estimated the revised model seperately for HF and Jersey cows with 50% level of exotic inheritance.

50% Jersey Cows : Our results, based on a sample of 81 Jersey cows with 50% level of exotic inheritance observed during October 1983 are given below :

(vii) For w = 0.6

$$Y = 1.83 + \underset{(0.042)}{0.066}[Q] - \underset{(0.0007)}{0.0023}[Q.Q] + \underset{(0.332)}{0.803}[C] + \underset{(0.050)}{0.071}[C.C]$$

R square = 0.71
F statistic = 42.69

(viii) For w = 0.5

$$Y = 1.94 + \underset{(0.043)}{0.052}[Q] - \underset{(0.0007)}{0.002}[Q.Q] + \underset{(0.335)}{0.822}[C] + \underset{(0.051)}{0.070}[C.C]$$

R square = 0.71
F statistic = 42.91

We find that the explained variation increases by around 8% when the breed and the level of exotic inheritance are taken into account. The coefficients are relatively stable when w ranges between 0.5 and 0.6

By partially differentiating equation (vii) with respect to Q and equating the differential to zero, we obtain the maximum of the function at Q equal to 14.3 kgs. Our field data shows that on an average, 50% Jersey cows are fed 8 kgs of dry fodder and 18 kgs of green fodder during the flush season. Consequently, the value of (D + [0.6]G) is 18.8 kgs. This suggests that farmers tend to overfeed their cows during the flush season, perhaps because roughages are available in plenty at this time of the year. Farmers would perhaps be better off if they fed lower quantities of dry and green fodder and utilized the money saved to feed a greater amount of concentrates. We believe this is a critical aspect if the economics of crossbred cows is to be improved, and should be further investigated.

50% HF Cows : A similar curvilinear regression was carried out for a sample of 27 HF cows with 50% exotic inheritance observed in October 1983. The results obtained are given below :

(ix) For w = 0.6

$$Y = 5.25 - 0.100[Q] + 0.001[Q.Q] + 0.520[C] - 0.033[C.C]$$
$$\quad (0.04) \quad (0.0005) \quad (0.500) \quad (0.077)$$

R square = 0.29
F statistic = 1.87

The above results indicate that feed is not an adequate explanatory variable and the curvilinear model explains only 29% of the variation in milk yield of 50% HF cows. The value of the R square is low and the signs associated with coefficients are unacceptable on the basis of a priori considerations. One reason for the low explanatory power of the model could be the small sample size. However, it could also be due to the larger intrinsic biological variabilty in HF cows. This variability could perhaps be a consequence of the variation in the semen of the pure-bred HF sires. This conjecture may be plausible because 50% Jersey and HF cows are bred from the same mother stock and reared under similar conditions, yet HF cows exhibit a greater variation.

Our conjectures are based on a small sample, and need to be investigated further. Such an enquiry is important since the variation in milk yield is closely related to the risk an individual farmer faces in dairying, and the risk can apparently be reduced by the proper choice of breed and control on the quality of semen.

Results for Local Cows : The results of the linear equations estimated for local cows during representative months in the lean and flush seasons are given below :

(x) May 1983 Sample size : 137

$$Y = 0.64 + 0.02\,D + 0.02\,G + 0.35\,C + 0.04\,Z$$
$$\quad (0.01) \quad (0.01) \quad (0.07) \quad (0.02)$$

R square = 0.36
F statistic = 16.86

(xi) June 1983 Sample size : 123

$$Y = 0.93 + 0.02\,D + 0.002\,G + 0.42\,C - 0.06\,Z$$
$$\quad (0.01) \quad (0.015) \quad (0.06) \quad (0.02)$$

R square = 0.48
F statistic = 23.64

(xii) September 1983 Sample size : 166

$$Y = 1.45 - 0.01\,D + 0.02\,G + 0.65\,C - 0.02\,Z$$
$$\quad (0.01) \quad (0.01) \quad (0.06) \quad (0.02)$$

R square = 0.51
F statistic = 38.37

(viii) October 1983 Sample size : 160

$$Y = 1.69 + 0.002\,D - 0.01\,G + 0.53\,C - 0.09\,Z$$
$$\quad\quad (0.000) \quad (0.01) \quad (0.06) \quad (0.03)$$

R square = 0.33

F statistic = 59.33

The major conclusions emerging from the regression are summarised below :

We find that the values of the R square range between 0.30 and 0.63 and are statistically significant. The comparatively low values of the standard errors indicate that a majority of the coefficients are significant.

Concentrates have a positive and signficant impact on the milk yield of local cows, but the impact declines during the lean season.

Grazing appears to have a negative impact on milk yield, once again indicating the poor quality of pastures.

Dry and green fodder appear to have a marginal impact on milk yield during both seasons. We notice that the coefficient of dry fodder is negative in September, while that of green fodder is negative in October. This is similar to the results obtained for crossbrd cows and indicates overfeeding of roughages in the flush season.

A Comparison Between Crossbred and Local Cows : Table 5.12 below presents a comparative picture of the values of the coefficients associated with various types of feed during representative months in the lean season. Values for the flush season are not compared because the analysis is complicated by the problem of overfeeding.

Table 5.12

Response of Milk Yield to Feed : A Comparison Between Crossbred and Local Cows

Month :	Values of coefficients			
	May		June	
Breed :	Crossbred	Local	Crossbred	Local
Type of feed				
Dry fodder (D)	+ 0.11	+ 0.02	+ 0.07	+ 0.02
Green fodder (G)	+ 0.04	+ 0.02	+ 0.17	+ 0.002
Concentrates (C)	+ 1.32	+ 0.35	+ 0.51	+ 0.42
Grazing (Z)	- 0.10	+ 0.04	- 0.10	- 0.06

The comparison reveals that the coefficients associated with dry fodder, green fodder and concentratres are consistently and substantially lower for local cows. During the lean season, since there is little chance of overfeeding, the farmers essentially operate in the

linear region of the response curve. Our results indicate that crossbred cows are better converters of feed. It is important to note that crossbred cows have better marginal conversion efficiency and not only a better average conversion efficiency.

.....