MODELING: A BRIDGE BETWEEN NATURAL SCIENTISTS AND AGRICULTURAL ECONOMISTS

by

Caroline Hoisington

A PLAN B RESEARCH PAPER

Submitted to

Michigan State University
In Partial Fulfillment of the Requirements
For The Degree Of

MASTER OF SCIENCE

Department of Agricultural Economics

1982

ACKNOWLEDGMENT

I wish to express my sincere gratitude to the members of the faculty of the Department of Agricultural Economics for their encouragement and intellectual guidance. I wish especially to thank Drs. Bob Stevens and Jim Shaffer for taking particular interest in my program, and for their good humor, and help.

I particularly wish to thank Roy Black for conversations which were fascinating, once I began to understand them, and for continued support in this project.

And I am indebted to Deb Andrews for completing a typing marathon, the final step in this project.

Lastly, thank you to my fellow graduate students for being such a fine, cohesive group.

MODELING: A BRIDGE BETWEEN NATURAL SCIENTISTS AND AGRICULTURAL ECONOMISTS

Introduction

This paper has a broad topic: the use of modeling as an area of common interest and as a bridge for research between agricultural biological scientists and agricultural economists. The thesis is that modeling is a tool which offers some unique properties, both in itself for addressing research problems and as a mechanism for communication. To demonstrate this position, examples of current and projected uses of models in the various plant sciences and in agricultural economics will be discussed. Suggestions will be made as to how both groups of researchers can benefit from working together on models. Lastly some suggestions will be made as to how scientists and social scientists from various disciplines can work together as teams.

As a preface, it is important to ask what is meant by a "model". For the purposes of this paper, the delineation is not difficult. I will be discussing mainly explicit mathematical models of systems: biological and/or economic which relate to agriculture. At times I will also be referring to the "systems approach" to a problem or a discipline. Here using a "systems approach" or "systems analysis" means simply to envision not only a specific research topic but also its context: the larger environment of which it is a part and its influence on other areas. The "systems approach" is taken as synonymous with an "holistic" approach where the attempt is made to reintegrate things and pieces of things (plants, economic factors, etc.) into a broader viewpoint of system functioning. In the first chapters of the paper, the focus is on modeling with very little specific reference to the systems approach more broadly. In Chapters V and VI the concept of a systems approach is more explicit. Overall, the major focus of the paper is on modeling, as the major part of the systems approach and explicit systems analysis. Other aspects of building a systems approach are not examined.

It is important first to point out that the field of systems theory and mathematical modeling is relatively new and growing. The professionals who use models differ somewhat in their views of what models and systems analysis are and, particularly, what they can do.

> "Systems analysis has rarely been defined when introduced into ecological studies. Watt (1968) suggests that it is the determination of those variables which are important in a system, and further adds that systems simulation, systems optimization, and systems measurement are part of the approach. Others, such as Priban (1968), view model building as the essence of the systems approach. Morton (1964) has suggested that systems analysis is no more nor less than scientific method itself, and that the distinguishing feature of the systems approach is the conscious application of scientific method to complex organizations in order that no important factor be overlooked, a view expressed by Pascal as 'error comes from exclusion'. These viewpoints are not necessarily mutually exclusive. Systems analysis is the application of scientific method to complex programs, and this application is further distinguished by the use of advanced mathematical and statistical techniques and by the use of computers."[1]

> "A Model is a simplified representation of a real system (either physical or biological). There are many types of models, e.g., word models, picture models, physical models, etc. all of which play an important role in science. This review is limited to mathematical models, that class of models which can be described symbolically and discussed deductively' (Kac). Hence a mathematical model is simply an equation (or set of equations) containing variables with relationships given by the mathematical expressions, these variables are meant to be analogous to physical entities in the real world system of interest."[2]

Spedding points out that models "include those features that are essential for the purpose and they leave out those that are inessential. Without a clear purpose, there are no criteria for deciding what is and

---

[1]Dale, M.B. "Systems Analysis and Ecology". Ecology. Vol. 51, No. 1, Winter 1970, p. 2.

[2]Reynolds, James F. "Some Misconceptions of Mathematical Modeling". What's New In Plant Physiology, Vol. 10, No. 11, p. 41.

what is not essential."[3] This simple statement has very important implications, namely that system definition is inseparable from the purposes of the analyst.

Forrester points out that all models contain assumptions and that they must be explicitly stated and incorporated in a mathematical model.[4]

Baschlelet lists several requirements for biological models. They should be (1) reasonably simple, (2) logically consistent, (3) mathematically correct, (4) consistent with the physical sciences, (5) consistent with natural phenomena and (6) comprehensive. He also points out that some models do not satisfy all of these requirements, which is an indication that better specification and/or further research is necessary.[5]

Reynolds points out a number of misconceptions about models:

1. A model can substitute for lack of understanding of a system.
2. There is a unique model for a given system.
3. Mathematical models provide new facts.
4. Greater complexity leads to a "better" model. He argues that it is difficult to determine broadly applicable measures of how complex a model should be, but there are definitely times when simpler is better.

---

[3]Spedding, C.R.W. An Introduction to Agricultural Systems. Applied Science Publishers, Ltd., London, 1979, p. 19.

[4]Forrester, Vay W. "Counterintuitive Behavior of Social Systems". Technology Review, January 1971, p. 54.

[5]Baschelet, Edward. "The Application of Mathematics to Biological Problems". Bioscience. January 1966, p. 22-23.

5. The major goal of model-building is prediction. He lists synthesis of information, communication and problem solving, and design as other purposes. This question will be discussed at various points throughout this paper.

6. A mathematical model can be validated. He says that validation of empirical models is appropriate only for the level of prediction and then it is usually not absolute. For theoretical models validation can be attempted at the level of the predictions and at the assumptions. However, one must not assume that if predictions made on the basis of such a model are good, then the assumptions are correct. These models are of the nature of hypothesis, "tentative explanations to explain observed phenomena ... It is impossible to prove any hypothesis correct since a false hypothesis may lead to correct predictions and contradictory assumptions can sometimes lead to the same mathematical model".[6]

Reynold's argument summarized above, also brings up an important conceptual distinction found in virtually all the literature, but under different names. There is a fundamental distinction to be made between two general types of models, which has little to do with the nature of the mathematical techniques used, but rather with the degree to which the model attempts to mimic the inner workings of the system. If a mechanism is not understood, it can be treated as a "black-box" the insides of which are unknown. But to the extent that what goes into it and comes out of it can be measured, or correlations measured, the inner mechanisms may not need to be explained. One can still construct

---

[6] Reynolds, Op. Cit., p. 43.

mathematical relationships from the known values. This type of model (or part of a model) is referred to variously as "black-box", "associative", "empirical", "empirical-statistical", "descriptive", or "ad-hoc".

The other type of model is one which attempts to determine the mechanisms of what goes on in a given process and to mathematically mimic the steps. This kind of model is variously called "theoretical", "structural", "mechanistic", or "fundamental". Some writers (e.g., Baschelet, Kac)[7] dwell mostly on this type of modeling and therefore describe models as hypotheses which over time may become so well established as to be considered "principles", "theories", or even "laws".

The distinction between these two types of models is not absolute. Mechanistic models generally must contain some simplifications or aggregations either for the sake of workability or because some processes are not known. "'Nature' remarked Fourier 'is indifferent toward the difficulties it causes a mathematician,' and because of this, mathematical models must of necessity be greatly simplified."[8] But, as will be further discussed in this paper later on, empirically modeled functions at one level can become the basis for structurally modeled functions at the next.

There is one more point to be brought up in this section: writers and advocates of modeling vary in their description of the significance of modeling and systems. Thornley, for instance, says: "It is unfortunate that the current fashionable use of the term "modeling" suggests

---

[7] Baschelet, Ibid.
Kac, Mark. "Some Mathematical Models in Science". Science, Vol. 166, No. 3906, Nov. 7, 1969, pp. 695-699.

[8] Kac, Ibid., p. 695.

that a new tool is at hand. This is not the case, for modeling, which is nothing more than quantitative hypothesis testing, has been used as a matter of course (and with great success) in the more physical sciences for at least a century."[9]

Dillon, in contrast, takes a broader view and puts a great deal of emphasis on the importance of a purpose or goal in defining the limits of a model, and on the implications of systems approach. "I believe the systems approach constitutes a new technology for viewing the world. In recognizing the purposeful nature of much of the world it has substantive implications for science. Expansionism, teleology and synthesis must be admitted as valid elements of scientific methodology-- and science takes its place as an instrument of higher systems."[10]

These ideas will be taken up again repeatedly in this paper and are therefore merely mentioned now as an introduction to the topic.

In the process of writing this paper, it became apparent to me that I was trying at times to explain some basic concerns of biological scientists to social scientists: research interests, ways of addressing problems, etc., and also, at other times, to express simple economic concerns to the plant scientists.

The chapter on the agricultural economist--plant scientist relationship is addressed more toward the economist(s) than toward the biologist(s) in that the suggestions made therein involve more conciliatory moves on the part of the economist. This bias comes partly from my

---

[9]Thornley, J.H.M. Mathematical Models in Plant Physiology, A Quantitative Approach to Problems on Plant and Crop Physiology. Academic Press, London. 1976, p. 2.

[10]Dillon, John L. "The Economics of Systems Research". Agricultural Systems. (1) (1976), Vol. 1, No. 1, p. 20.

opinion that agricultural economists, who deal frequently with questions of social policies and administration of various types, are likely to be both better equipped to explore and more interested in the mechanics of setting up cross-disciplinary work. I think it is clear however that all members of multidisciplinary teams can benefit both from the process and the results, the biological scientists as fully as the social scientists, and in some spots in the paper there are simple economic explanations aimed at non-economists in an attempt to illustrate a bit of what tools economists bring to such work (e.g., II.B.1).

This paper, then, comes from someone with some experience as a graduate student in both agricultural sciences and agricultural economics. It is written in the hopes of offering information and suggestions of use to members of disciplines within the biological, agricultural sciences and in agricultural economics in their efforts to work with each other.

II. Modeling and Agronomy and Whole Farm Models

A. Agronomy (Applied Research)

Most agronomic testing is done without the aid or complement of a model run either previously or simultaneously, paralleling or complementing the field work. Traditionally such things as testing for response to levels of fertilizer application, timing of planting and harvest, irrigation rates and timing, etc. have been considered fairly straightforward research questions and not areas of inquiry for which models were needed or even useful. As agronomic practices become gradually more sophisticated, however, this is less the case. Changes in production at one time would have involved incremental additions of various inputs and discovering optimal, least-cost-combinations in a production function sense. These things are still done but there are now factors which make modeling of whole processes a more attractive option under some circumstances. These factors include:

1. the rising costs (in labor, materials, etc.) of doing field experiments;
2. the rising costs of certain inputs to production and the concern that some prices may continue to rise or even to rise at an increasing rate as certain materials (e.g., good quality water, petroleum products) become scarcer relative to the demand for them. This issue relates not only to the economic assessment of varying amounts of inputs but also to the fact that much more sophisticated technologies have grown up in response to these needs, and the level of management which they require is qualitatively different from the previous systems.

3. As the above-mentioned concerns grow, the need for researchers in agronomy to focus more on whole production systems and tradeoffs within whole systems rather than maximizing parts of systems increases. The variables and interactions among system parts can easily become so complex and difficult to handle simultaneously as to make a modeling effort look quite appealing.

Some examples and clarification of the above ideas follow:

1. Rising Costs of Doing Field Experiments. Not very much needs to be said about this conceptually simple idea. To the extent that models which provide fairly accurate predictions of a biological system's behavior exist, a certain amount of field work may be pre-tested. This implies that the models used are probably (but not necessarily) structural rather than empirical in nature, and that the researchers have confidence in the significance of model results. A simulation or other modeling attempt may show the likely economic limits of some production possibilities and obviate the need for some of the treatment levels in field testing (e.g., Black, J.R., personal communication regarding feedlot management).[11]

An example of a need for a systems approach analysis of this type was given to me by Dr. M. Wayne Adams recently. Dr. Adams, a plant-breeder, is in the process of releasing two new varieties of field beans for commercial production. They are both of a type which represents a new plant shape or "architype" (his word: architecture and archetype).

[11] Black, J. Roy, et al. "The Development of Software for Computer Assisted Agricultural Decision-Making". Michigan State University, Agricultural Economics Staff Paper No. 81-37. Revised June 9, 1981; and personal communication.

Dry bean plants are typically low and branching. If they are harvested in one pass by a combine, the number of pods missed by the combine, because they are below the blades, is significant. The generally used method of harvesting dry beans is in more than one pass: they are pulled, winrowed and then thrashed by different machines. The presumed advantages of the new varieties are based on the fact that they can be combine harvested. To use them will, however require an entirely new production package. They must be planted earlier and in narrower rows (because they don't branch, high yields require higher planting densities), weed control is different, etc. Yield trials are being run this summer with two different plant spacings and ten different herbicide treatments on each of four varieties (two of the older type and two of the newer type plant). This test will be a beginning of generating enough data to do an economic analysis of such questions as the trade-offs between pulling, winrowing and thrashing vs. combining (in the former case, each pass takes less energy than the combine, but the total requirement may be higher) and the effects of the various management changes.

If—as is not the case—a sufficiently developed model of the bean plant and a bean farm existed, one might be able to try a number of runs of this nature and test fewer situations in the field, thereby saving time and money. One could also make earlier estimates of the economic value of this particular research.

2. Managing Specific Production Packages. The second point relates to models which have grown up as part of the application of increasingly sophisticated technologies. An example which comes to mind is in soil science where the current state of the art of irrigation is

almost mindboggling. Changes in water delivery systems of recent years have included the giant pivot sprinklers, lower pressure units which move laterally, trickle irrigation and bubblers, and automated gravity flow open-ditch systems. Precision land grading with a laser has made a system called "dead level" possible which "divides a field into sections, their size depending on soils and topography, which are then leveled as smooth as a table top and surrounded by low levees. Water is applied in large streams to provide quick and uniform temporary flooding... increasing efficiency ... in one study 60-80 percent".[12]

Combined with this sort of technical expertise has been the use of models:

> "Aided by better experimental techniques, and especially, increased computer capability, we are now able to predict or describe the movement of water, salts, pesticides, and heavy metals with a reasonable degree of certainty. That movement is of paramount importance when dealing with irrigation and drainage, salinity management, and pollution control. We also have 'discovered' a new problem: use of such mathematical models requires quantitative knowledge of the pertinent soil properties. Whereas these often can be measured adequately on soil samples, soils are notoriously unhomogeneous (as well as variable in time), and serious questions arise as to the extrapolation to field-scale situations. Significant progress has recently been made in defining the nature of the spacial variability by drawing on concepts from statistics and geology and in interpreting the consequences of this variability ... Not until recently, however, has there been much success in effectively applying such theory to practical field design. A set of computer programs has been developed that are both simple enough and sufficiently realistic to enable relatively routine design that takes account of soil properties, weather variations, and anticipated crops to be grown."[13]

3. Integrating Agronomic Research. The third point relates to the fact that scientific research is generally a reductionistic process.

---

[12] Van Schilfgaarde, Jan. "Earth and Water". The Antioch Review., Vol. 38, No. 4, Fall 1980, p. 428.

[13] Ibid., p. 423, 429-430.

This point will be referred to repeatedly in this paper. The objects and processes are conceptually broken down into finer and finer parts, the measurements more delicate, the conceptual framework for understanding biological systems goes from the level of organisms to tissues to cells, subcellular particles, molecules. The tendency is to specialize both in the basic sciences and the applied, but for research results to be usable in applied work, whether in research on management of systems such as farm enterprises or in extension of the research, the pieces must be reintegrated into a whole: a whole plant, a whole field, a whole farm. In this context, the overriding advantage of modeling lies in the development of a framework through which to integrate the research pieces. In agronomic problems, this integration will often involve research and researchers from more than one discipline and may even take the form of "bio-economic models." Charlton and Street are fairly critical of specialization and research carried out in isolation which "has often been found to be irrelevant or meaningless when put into a practical context ... [and of] advisory and extension work [where] advice has often been given referring only to an isolated part of the system without regard for its significance within the whole farm situation."[14]

They further claim that "It is principally only by non-practitioners, that is, by academics and advisers, that the need to consider complete systems, rather than just their component parts, has

[14]Charlton, P.J. and Street, P.R. "The Practical Application of Bio-Economic Models", in Dalton, G.E., ed., Study of Agricultural Systems, Applied Science Publishers, Ltd., London, 1975, p. 236.

often been neglected," and they consider that "Traditionally, agricultural science was developed by researchers with a good understanding of all the components of agriculture and its related disciplines.

> "It was the recognition of the need to try once again to take an overview of a situation, simultaneously considering a wide range of factors, which produced the so-called 'systems approach'. This has been widely heralded by its protagonists as a totally original concept capable of giving insight into the behavior of highly complex situations. In fact it is important to recognize that it is really only a recognition of the inadequacy, in a practical situation, of much of the highly specialized academic research which ignores important reactions between system components. It is simply a reversion by academics to the more general or 'holistic' approach that farmers and other practitioners have always adopted. This is a very pragmatic view of the concept of the 'systems approach'. However, we believe it to be a very necessary one if the study land analysis of whole systems is to be restored to a realistic perspective."[15]

It is my view that the "systems approach" and also specifically building mathematical models can have other functions extending somewhat beyond that described by Charlton and Street above, including the use of models as a tool for hypothesis testing by basic science researchers, a means of trying to evaluate the relative importance of parts of some processes (and needed research in these areas) through sensitivity analysis, and in other ways which shall be discussed below, models may have a variety of useful roles in research, but the importance of Charlton and Street's point remains.

As the sciences progress, it becomes increasingly difficult or impossible to keep up broadly in one's own field, let alone other's fields. The traditional "good understanding of all the components of agriculture and its related disciplines" which loss they lament, may indeed be rare now partially due to an over-emphasis on specialization,

---

[15] Ibid., p. 236-237.

but it is no doubt also due to the wealth of research and depth of inquiry in the numerous agricultural fields, and the fact that there are tradeoffs between specific expertise and general broadly inclusive understanding. This is not to argue that working intensely in a professional niche surrounded by smooth walls of perfect ignorance is justifiable or unavoidable, but rather to argue that in order for experts in different fields to communicate adequately and to contribute to each others' understanding, to do complementary research work, such as work on different aspects of one project, it is likely that some structure will be needed. A model may be a way to provide the structure to enhance communication and to allow for handling large amounts of data or analyses, and to coordinate research on different aspects of a problem. In fact a number of authors mention that one of the major advantages of working with computerized models is that modeling can provide some understanding of interactions on various levels, one of them being the interdisciplinary. A corollary to this is that in the process of synthesis of information to generate a model of a functioning whole, information gaps come to light some of which may not have been previously recognized.

Dent and Blackie, referring specifically to the use of simulation models suggest the following advantages which, they say, have made simulation a standard procedure in many disciplinary areas within agriculture:

1. "It enables the study of systems where real-life experimentation would be either impossible, inordinately costly or disruptive.

2. By synthesizing systems in model-form, it permits the exploration of systems that do not exist.

3. It permits the study of long-term effects since the time horizon over which a model is run is within the control of the model-builder.

4. It forces those concerned with building the simulation model to examine the system objectively and consequently undertake a thorough and critical review of knowledge concerning the system. The enlightenment that this process provides is often surprising."[16]

We turn next to the application of models to farm situations, looking first at the more economic side of the "bio-economic" combination.

[16]Dent, J.B. and Blackie, M.J. Systems Simulation In Agriculture, Applied Science Publishers, London, 1979., p. 11.

II. B. "Bio-economic" and Whole Farm Models

1. "Bio-economic" Models vs. Production Functions

Agricultural economists have long been involved with farm management and production economics. There is, emerging from the need to analyze complex production systems, such as farms, a new aspect of model building as it relates to agronomy and agricultural processes generally: the "bio-economic model". In doing an economic analysis of inputs, outputs, assessing efficiency, reorganization of production, etc., economists have traditionally used "production functions". These may be visualized as coming in a variety of shapes and having properties depending on the extent to which the inputs are complements or substitutes, whether an excess of a given input can have a negative (toxic) effect on total product, etc.[17]

It occurs to me that the production function is a very simple mathematical model. It is a summary of a lot of information, which purports to demonstrate the relationships between inputs and outputs, somewhat selectively and with some awareness of the few types of interactions listed above (complementarity, etc.) and that it is basically a "black box" or empirical analysis as opposed to a structural one. To the extent that the attempt is made to simplify to a single equation the results of several processes, it will be rather too simplistic for predicting or understanding the inner workings, the rates of processes, interactions, etc. within the function. In fact, these pieces of information, to the extent which they are used, are all exogenous variables

[17] Johnson, Glenn. Michigan State University, Department of Agricultural Economics, Classroom lecture, AEC 805. 1980.

in production functions, supplied by the analyst through his/her or someone else's experience and knowledge about the system.

More complex systems of equations, whether in the form of mathematical programming, optimizing models, or of simulation models involving difference equations and stochastic elements, are more closely tied to the actual processes being modeled--at least to the extent that the models are structural. The empirical or "ad hoc" models would, in this analysis, be thought of as more sophisticated versions of the production functions. Their explanatory powers would be higher than the production functions in that they account for input and output of complex processes and in that both a number of steps and interactions between parts of the system can be modeled (whether or not the modeling reflects biological events accurately). To the extent that a model is structural it attempts to explain or mimic the inner workings of a system as well as the inputs and the outputs. The model will not substitute for knowledge. It cannot generate data where there was none. It can however be used as a means of hypothesis testing, exploring relationships, interactions and attempting to find mechanisms (expressed as mathematical relationships) which are consistent with the data and with current understanding of biological relationships.

To use a simple-minded example, to describe the differences between production functions and other more complex models, from "black-box" to "structural" we imagine the situation of researchers analyzing of the process of producing a dinner with a certain balance of different courses, in a restaurant and making an economic analysis.

The researcher working with production functions might gather data in the form of grocery lists and information about the technologies (ways of cooking, mixing, chopping, etc.) and also information about which items were complements (e.g., oil and vinegar or baked potatoes and sour cream) and which might be substitutes (fresh beans and frozen beans, etc.). This understanding of the process on the part of the analyst would have to be adequate for him/her to recognize which items in which combinations are crucial (if baking chocolate is essential for making chocolate cake and is not available, what substitutes are required, chocolate cake mix, ingredients for angle food cake, or potatoes?). But the final production function(s) could involve ingredients (groceries) as inputs in each of the processes (recipes), and one can stretch the imagination and envision optimal points if the output dinner courses could be appropriately priced. Clearly if one lumped all of the inputs and processes into one big equation with the output "dinner" the equation would have very little explanatory power, but it could conceivably carry information just relating inputs and outputs, a big "black-box". Making adjustments in production of such a complex system (multi-course dinner) on the basis of one or of many production functions simply doesn't seem feasible. The concepts of complementary and substitute inputs are not sufficient to deal with a several-stage process with several products to be produced in sequence. We can take the example further and suggest that maximizing production and least-cost combination of inputs are not of sufficient goals. Certain quality constraints must be built in. The whole farm is also a multi-stage, multi-product, multi-objective, multi-constraint, multi-interaction sort of system.

The researcher generating a "black box" <u>model</u> would presumably observe which and how much of the ingredients disappeared at various times in the process and which products (and how much) appeared and work on the basis of correlations, regressions, etc. A black box model might not represent the actual processes taking place, but it would have more explanatory power about the inner workings of the system--as more intermediate steps and separate processes would be explicitly included--than the production functions. It could be deterministic or have stochastic elements.

An LP could be generated with the objective of minimizing costs subject to the stipulation that a certain level of nutrients was obtained or maximizing nutrients given a certain level of costs. However the resulting ration formulation if a very large (and unwieldy) LP tableau were set up for an entire dinner would, likely generate some very odd combinations of foods. Separate LP's for various parts of a meal could be done, as could multiple objective programming.

A structural simulation model would likely take into account separate recipes, timing of steps, perhaps stochastic elements (e.g., a price generator), decision trees (e.g., if price of roast beef is greater than x, use pork), etc. The form and the complexity of the model would depend on the objectives of the modeler.

This example is rather far removed from those biological agricultural models which incorporate a biological crop model, partly because the nature of the dinner-making process is more of an assemblage of parts which are large and easy to quantify whereas a biological growth model may also be an assemblage of parts, but the parts are minute (molecules of sugars, etc.) and the pathways they take and the regulation of those pathways so complex as to be virtually impossible to trace and many are currently quite unknown. Also a growth model is a continuous process of accretion whereas the dinner example is not.

The simple point however remains that when examining productive processes a variety of analytic tools exists. The production function may be adequate to describe certain kinds of processes given known possible tradeoffs and complementarities, but its explanatory powers are limited in comparison to modeling procedures, which can handle multiple processes and their interrelationships explicitly. If processes are not adequately known or if a number of complex interactions exist and changes in a production process are envisioned, for instance in a multi-cropping situation, stronger analytical tools may be of greater use, including optimizing models or complex structural models. The dinner example, although not realistic, suggests the complexity of dealing with a number of in-place technologies (recipes for food or production packages for crops) which interlock with other parts of the whole (the rest of the dinner or the rest of the farm) and are subject to exogenous prices for inputs and outputs, and which may involve optimizing or at least considering criteria other than money (e.g., nutrition, taste, crop quality, soil erosion, etc.).

The need to work with more sophisticated understanding of complex systems raises the demand for more structural, less black-box, bio-economic models. The following excerpt supports this argument:

> "The typical agrosystem is controlled primarily by the driving force of the environment and the agronomic practices of the producer. The components of the system are either or both the crop and animal population as well as the pest(s), the beneficial organisms and the producer. The important weather factors are sunlight, humidity, air temperature, wind speed, rainfall and photoperiod. Of these six factors, only the photoperiod can be predicted consistently with reliable accuracy. In addition to the variation of inherited characteristics of individuals of the crop and animal

populations, variations occur in the individual microclimate, nutrient supply, pest densities and agricultural practices within the field. It is not difficult to identify areas where the effects of the environment on growth, development, reproduction, disease resistance and yield are not understood. Most of these effects have been measured only for certain specified conditions; outside these specified conditions the responses are essentially unknown.

"It is generally accepted random events do not occur, other than, perhaps at the level of individual nucleons and electrons. The apparent randomness observed in agrosystems stems from a lack of detailed knowledge about the mechanics of the system. If we know nothing about a system, the expected response must be treated as completely random, with every outcome equally likely. As we being to acquire data and develop theories, the outcomes, although still random, are not equally likely. A non-uniform distribution function then describes the likelihood of an outcome. A system whose outcome is predictable is one where all mechanisms and inputs to the system are known. The distribution function becomes a delta function where the probability of all but one of the outcomes is zero ... Thus, until much more is known, the agrosystem must be viewed as a stochastic or random system and the response to any specified agronomic or pest management practice should be treated as a random variable. Models, particularly biophysical models, sharpen the distribution thereby improving the accuracy of forecasts."[18]

A final and important point about the choice of analytical tools should be mentioned here. As the degree of model complexity increases (thereby reducing the amount of abstraction from the real world information) so does the cost of the model. One can envision a hierarchy of abstraction, with the "real world" (zero abstraction) at the top, and various kinds of models--operational exercises, operational games, computer simulations with human decision makers, computer simulations with decision rules incorporated, various mathematical and statistical analytic tools— following, with an increasing degree of abstraction or simplification from reality. The simpler models will most likely cost

---

[18] DeMichele, Don W. "An Evaluation of Modeling, Systems Analysis and Operations Research in Defining Agricultural Research Needs and Priorities in Pest Management". Iowa State Journal of Research., Vol. 49, No. 4, Pt. 2, May 1975, p. 603.

less to construct, but contain less information and more room for error. Clearly this is one very important aspect of how the purpose of the model determines how it is constructed. For some purposes, relatively simple relationships will be adequate and for others not. The decision then relates to the level of complexity of the problem or research question and to the cost of "being wrong" either from omitting important relationships or from over-abstracting and misspecifying relationships.

The selection of the type of analytical tool then becomes a kind of economic question, subject itself to analysis of what the costs of being wrong are relative to the costs of developing more complex analytic tools, more complex models. Such decisions are made all the time, probably more subjectively than analytically, but an analytical framework could be developed. In many areas, however, the trend toward building even more complex analytic tools seems clear, such as those discussed in the following two sections.

II. B. 2. Whole Farm Models and Farm Management

In whole farm models, the combination of biology (agronomy, botany, etc.) and economics brings in a number of issues which agronomists classically have not considered to be within their domain, particularly economic tradeoffs which involve settling for less than potentially optimal yields. These may include such things as labor bottlenecks forcing a choice between haying at the optimal time and planting a late crop or, more simply, accepting a less than optimal machinery component for producing (harvesting, etc.) a particular crop because the change in equipment will cost more than it can generate in improved operations, or because a farmer's preference is for slightly higher risk (e.g., due to a longer planting time with a slower/smaller machiner) over a larger debt-load, etc. Agronomists, while being fully mindful that on-farm conditions are likely to be less than optimal at some times and certainly aware that experiment station results are often better than on-farm results (although the opposite is occasionally true!) still generally gear their research towards optimizing and in extension situations will advocate reaching the highest possible production.

To the extent that principles of economic levels and tradeoffs are inherent in whole farm models, these models may be a useful tool for production agronomists who are now approaching such issues as multiple cropping or looking at cropping rotations in a new light (e.g., in relation to relatively new developments such as no-till methods of growing corn, etc.) or in light of the concern growing in some quarters, over erosion control.

M.B. Dale describes aspects of using a model of a system to manage that system:

> "This will involve experimenting with the real and model systems, identifying the parameters of the system which will enable it to be controlled, choosing a value function by which the performance of the system is to be measured, selecting the route to some desired state, and maintaining the system at or near this desired point. The advantages of using the model system lie in the ease and rapidity with which experiments may be carried out, and the possibility of including experiments which might be totally destructive in the real system. The disadvantages lie in the restricted range of confirmed validity of the model and in its fidelity even within this range to the real system which it is desired to control."[19]

The more economic side of the "bio-economic" models of farms already exists in the form of the farm management decision-assisting models which are currently being used and/or under development in a number of states.[20]

This large and important area, however, is considered to be beyond the scope of this paper; these economic management models are generally well known among agriculturalists, especially among the agricultural economists who generate and run them, and this paper is focused more on the questions which include biological modeling. The question of whole farm modeling in relation to developing "farming systems models" for use in developing countries will, however, be looked at briefly as some unique issues come up in this context, such as modeling of multi-cropping and mixed farming systems. In the attempt to deal with the complexity and multiplicity of traditional (largely subsistence) agricultural production systems, and because of difficulties encountered in the past by agronomists and other development workers where seemingly superior crop varieties or management practices were not adopted by

---

[19] Dale, M.B. Op. Cit., p. 11.

[20] Debertin, et al. "Impact on Farmers of a Computerized Management Decision-Making Model". American Journal of Agricultural Economics. Vol. 63, No. 2, May 1981, p. 270.

farmers in various areas around the world because they did not allow the whole farm system to function smoothly if used (e.g., much improved cotton varieties which require a longer growing season and therefore crowd out a later food grain crop) an entire sub-field has grown up in development literature which is treated in the following sections--the "farming-systems approach" and "farming systems research".

II. B. 3. FSR: Farming Systems Research

> "The international brand of FSR was developed largely to address the problem of lack of adoption of improved agricultural technology. Low adoption rates were a sign that important factors had been excluded in the technology design process. FSR was intended to account for these missing factors, such as: (1) interactions involving crop and animal enterprises and farm and non-farm activities; (2) the performance of the technology under actual on-farm conditions; and (3) economic and socio-cultural factors affecting acceptability. Whether, FSR—so defined—will in fact successfully overcome the 'adoption problem' remains to be seen.".[21]

FSR probably has a number of definitions and has been applied to programs which vary significantly in several dimensions. Both research which includes the "systems approach" of thinking in terms of a whole system with many interacting sub-parts and that which involves actually building mathematical models of various types can be included as FSR.

Crawford's discussion of the use of quantitative models in FSR is both instructive of the state of the art and suggestive of future research needs, and is therefore reproduced at length here:

> "Econometric models are attractive in part because comparatively well-accepted procedures are available for estimating and evaluating their structural parameters. Econometric models can potentially incorporate features such as behavior over time and stochastic variability. For example, random coefficients production models have been discussed by Swamy (1974), Mount (1974), and Harville (1977), although apparently they have not yet been applied to an integrated farm household model.
>
> "Linear programming (LP) has been a powerful, widely used tool. It is flexible enough to incorporate features such as multiple inputs and outputs, behavior over time and the effect of uncertainty. One drawback of LP models is their inherent tendency to give unrealistically one-sided optimal solutions, e.g., overspecialized cropping patterns. Also very careful scrutiny is

---

[21] Crawford, Eric W. "Farming Systems Research and Agricultural Economics". Michigan State University Working Paper No. 1, June 1981, p. 5.

needed to establish whether an LP model is sound, or whether apparently plausible results were "forced" by artificial, a-theoretical manipulation of constraints. (This drawback is shared to some extent by econometric and systems simulation models).

"Systems simulation models offer even greater flexibility of form. Complex features can be readily accommodated and solutions still obtained (Johnson and Rausser, 1977; Crawford, 1980a). Model specification can be eclectic and behavioral, facilitating use in problem-solving research (Johnson, 1977). The principal drawback is that simulation models often cannot be proven theoretically consistent. A related problem is the inadequacy of standard statistical procedures for evaluating how well a simulation model performs; considerable subjective judgement is also required.

"Other difficulties arise in modeling the link between events at the individual farm level and aggregate effects at the macro level. FSR focuses primarily on the farm level, but the literature generally recognizes the importance of national policies and the regional agricultural economy as factors influencing the appropriate direction of new technology development. However, both theoretical and quantitative models are limited in their ability to predict the macro effects of introduced new technology or policy interventions. At the formal level, there are problems of bias in aggregating the results of individual farm models to obtain a picture of regional impact (Day, 1963); in general, it is not legitimate to assume that the whole is equal to the sum of the parts. In addition, evaluating the impact of new technology or policy interventions depends on the ability to analyze regional product and factor markets. This moves the domain of the analysis from a partial to a more general equilibrium framework, unless the situation can safely be simplified. Also, it is clear that the economic impact of a given development intervention is mainly influenced by institutional and socio-cultural variables. However, predicting socio-economic effects over time and on a macro level is not yet feasible given available theory and analytical methods in the social sciences."[22]

Despite limitations in methodology and the conceptual complexities of trying to balance the necessity of incorporating behavioral, institutional and political variables to make models run realistically with the practical necessity of limiting scope, FSR is seen by many as the best presently available approach to analyzing and surmounting the enormous range of difficulties and sometimes high rate of failure of development projects. In the literature of development economics developed since

---

[22] Ibid., pp. 14-15.

World War II, there have been a variety of theories and approaches. In a very general way it is correct to say that there have been swings back and forth between the approach which emphasized anthropology and sociology—the years of "community development" or "CD"--and the technical emphasis—the "Green Revolution." The difficulty is that neither is totally successful. CD emphasized local organization and doing more with what was available; development advisors were "catalysts for change." The outlook was implicitly condescending and paternalistic in the unspoken assumption that the barriers to change were social/political/organizational and that an outsider could do a better job than the local leadership of reorganizing what was available. The technical emphasis of the Green Revolution was—in hindsight—simplistic in its (also implicit) assumption that supplying improved technologies and access to necessary inputs would solve the problem of impediments to development. Even in the most successful showpieces of the Green Revolution, unexpected problems were encountered. Some were social problems such as those in the Punjab in India caused by massive unemployment among the rural landless who had been tenant farmers before the advent of the Green Revolution and the new technologies, including tractors, which made farming the land themselves economically feasible for the first time for the large landowners. This situation became so acute that the Indian government at one point actually forbade the importation of more tractors, a seemingly unusual act for a developing country![23]. Technical problems of non-adoption of new inputs, referred to above, are ubiquitous in the literature. They are attributed to a number of

[23]Eicher, Carl, Michigan State University, Department of Agricultural Economics, Classroom lecture, AEC 862. 1979.

"constraints" sometimes operating singly (e.g., inadequate access to reliable water supplies) but more often in groups (e.g., inadequate access to reliable water supply and to adequate credit to buy inputs, and inadequate input markets, and inability to assume extra risk, etc.). Thus the technical approach has encountered its own frustrations.

FSR, has thus grown out of the more sophisticated realization that fostering major changes in a society through direct involvement is an extremely complex task. It is multidisciplinary in conception and projects will often require multidisciplinary teams in order to facilitate changes in major socio-economic endeavors such as agricultural production, it is necessary (usually) to adequately understand the existing system(s). Otherwise, proposed and advocated changes may be infeasible or in conflict with what people (farmers, consumers, governmental officials, etc.) want or are willing to accept.

The question, of course, remains closely tied to a moral dilemma of the extent to which it is not just possible but desirable for an outside individual or culture to attempt to try to change another. For now I will skirt that issue by simply asserting that the pragmatically clear need for the world's growing populations to feed themselves demands that some kind of economic development take place. What kind of development and by whose definition is not settled. Still, the needs of Third World people and the demands for aid from their governments are concrete. For the moment we turn to some existing models.

An interesting and unique LP model of a Chinese commune has been produced by Tom Weins.[24] The model included a variety of cropping sequences (thirteen of them) and variations of pig raising, vegetable production and included several different composting, green manure and chemical fertilizer options. It is an optimization model, and generates an optimal mix of enterprises, and also shadow prices, and by varying prices and constraints (which are exogenous) one can obtain ideas about not only the values of certain resources and their allocation but also about choice of technologies (e.g., various composting techniques) and tradeoffs inherent in complex cropping schemes. When I saw the model it was in a research stage--not developed to the point of being a management tool. This is only natural as the project is a pioneering effort based on a unique research opportunity. This approach could be of considerable value in other parts of the world where multiple cropping (both sequential and intercropping) is an essential part of farming practice and where agronomists have encountered difficulties in introducing either improved seed varieties or improved production techniques because of the complexities of timing, planting and harvesting various crops, due to both the biological demands of the crops and to the--sometimes strict--labor bottlenecks involved.

This methodology could also presumably be of use to planners and decision-makers who wish to examine such questions as the impact upon a production system of changing a governmentally controlled pricing structure or a particular policy. In the Weins' study, several less than economically-optimal practices were discovered.

---

[24]Weins, Tom. Personal Communication, 1981, study soon to be published.

Among the efforts to simulate whole farms or whole farming regions, a study by Crawford and Milligan on Nigeria is of particular interest.[25] The study emphasizes the use and value of experimental design and particularly examines the usefulness of simulation models for farm level analysis.

This study is of more of an economic than a "bio-economic" simulation in that crop yields were fed in, not generated by the model. Soil, climate and agronomic data were not included--which Crawford laments: "[a]accordingly, it proved very difficult to fully analyze the sources of observed variability--in terms of labor inputs and crop output per hectare--among the different fields growing a given crop mixture type."[26] The focus is on the economic/social criteria of "Accumulated Net Capital" and "Household Consumption Per Consumer," which allowed an examination of some issues critical to the question of adoption of new technologies. Particularly, the role of resource endowment was examined by projecting the financial viability of families with different resources (basically land/person ratio) over a period of years under different scenarios of circumstances. The four experimental factors

---

[25] Crawford, Eric W. and Robert A. Milligan. "A Stochastic Farm System Simulation Model with Emphasis on Experimentation". Michigan State University, Agricultural Economics Staff Paper No. 80-81, Nov. 1980, revised July 1981. Crawford, Eric W. "Understanding, Quantification, and Modeling in Farming Systems Research: Results of a Simulation Study in Northern Nigeria". Agricultural Economics Staff Paper No. 80-19, March 1980.

[26] Crawford, Op. Cit.., AEC Staff Paper No. 80-19., p. 10.

were (1) household type (three levels of family size and endowment combinations); (2) crop enterprises used to approximate different management levels (two levels); (3) marginal propensity to consume (two levels); and (4) stochastic sequence (includes crop yields and investment returns, and special consumption expenditure) with four levels: base run, moderate loss, severe loss and moderate gain. The model was run first deterministically and then stochastically. The reader is referred to the cited documents for an analysis of the results. A brief mention only of the implications of the results is given here.

> "The chief advantage of this study is the multi-year framework it provides for examining the critical endogenous and exogenous variables of the farming system ... the relative importance of variables which are hypothesized to determine long-run economic success can be systematically and directly studied."[27]
>
> "It was anticipated that households with little land cash and food resources per family member would experience slower growing and more erratic incomes than better endowed households. It has been observed that purchased-input intensive "Green Revolution" technology is not scale-neutral, contrary to expectations. The point made here is that traditional production and investment opportunities are not scale neutral either."[28]

And finally, another suggested use of such models is suggested:

> "There are circumstances in which simulation modeling may be a worthwhile adjunct to experiment station or field-based FSR, despite the recognized data, expertise, and computer requirements of the modeling approach. For example, testing prototype technology via on-farm trials is an essential process, but a slow, sometimes expensive, and expertise-demanding method itself (Collinson, p. 11).
>
> "Where the manpower and computational resources are available (e.g., at an international research center, or middle-income or developed country institution), simulation studies might usefully proceed side by side with on-farm trials, facilitating a timely assessment of the sensitivity of new technology under different

---

[27] Crawford, E.W. and Milligan, R.A. <u>Op</u>. <u>Cit</u>., pp. 20-21.

[28] <u>Ibid</u>., p. 2.

assumptions regarding household type and environmental variability. The analysis can easily be extended over a five to ten year period, guarding against the possibility that the modified farm system might be subject to unacceptable cumulative, seasonal, or bad-year losses. Such losses are unlikely to be picked up in a single-year analysis which relies on average returns to evaluate profitability. The improved evaluation of the dependability of returns which the simulation approach can afford would be particularly helpful where there are many marginal farmers, and/or where weather conditions are highly variable."[29]

Farming Systems Research has developed in the context of development economics.

"(T)he capacity of FSR to examine complex multi-enterprise farming systems and to 'give a voice' to the farmer is probably more beneficial in Third World agriculture than in North America and Europe, where farms tend to be more specialized and where farmers have the resources and education needed for them to represent their own interests to the research and extension establishment. For small or part-time farmers in developed countries, FSR may have the same benefits it does in Third World countries... However, the FSR methodology is potentially suited to small and large farms alike, and to North American and European farms as well as those in the Third World ... A domestic U.S. application of the systems approach to problem-solving by a continuation of biological and chemical controls, changes in crop mix and cultivation, and more careful monitoring of pest populations. This approach relies on information from several disciplines, including entomology, soil and plant science, agricultural economics, and agricultural engineering, as well as from the farmer."[30]

Thus Crawford makes the point that IPM, while not generally considered a branch of FSR—the latter term being most commonly used in international work—is similar in its focus on the farmer, on farm-level application and in its interdisciplinary nature. Indeed, there is nothing intrinsically international about an FSR type approach.

This leads into a question of the farming-systems-research which might be developed domestically in the U.S. Why has an "integrated pest

---

[29]<u>Ibid</u>., pp. 10-11.

[30]Crawford, <u>Op</u>. <u>Cit</u>., Working Paper No. 1, pp. 3-4.

management" scheme been developed for dealing with insect pests and not, as yet, an "integrated weed management" system? Perhaps because the research on allelopathy and the use of rotations to control weeds has not progressed as far as research on biological control of insects. Perhaps concern over herbicides in the environment has not been as strong as concern over insecticides. Research in weed control through rotations and allelopathy is, however, underway at IRRI in the Philippines.[31]

One more comment about FSR follows. It seems clear to me that one of the main goals of FSR offers something of a paradox, or at least an unresolved dilemma, relating to the mechanisms of (and rate of) change. Lack of adoption of new technologies, as cited above, is seen as a major impetus behind the felt need to understand more completely the entire farming system in traditional agricultural contexts. Because the traditional agricultural systems have been developed under conditions where farmers had relatively little control over their environments (without chemical fertilizers, pesticides, and physical ability to make extensive land and water management schemes work, with some exceptions) traditional farmers have had to make more accommodations to the environments they couldn't change. (They also generally had fewer opportunities to use the disease and pest resistant plant varieties which have been provided by modern plant breeders.) Traditional farming systems are often found to be more intimately bound up with a number of physical and sociological cycles including not only cropping activities which may be

---

[31]Harwood, Richard R. "Natural Weed Controls Are Looking Good". New Farm., Vol. 1, No. 6, September-October, 1979, pp. 56-58, and "Seasonality in Organic Weed Maintenance", New Farm, Vol. 2, No. 4, May-June 1980, pp. 61-64.

very complex and rigidly demanding in their time sequences, but also livestock operations, off-farm enterprises, domestic tasks, etc. The whole of the farming system has developed as a delicate network of balanced activities, recycling of nutrients, (e.g., bush fallow, slash and burn, composting, rotations) division of labor throughout the year, etc. Systems may be moderately productive or marginally so, but they have evolved and survived over long periods of time, (e.g., King, Farms of Forty Centuries).[32] Modern agriculture, by contrast, has afforded the farmers much more control and freed them to some extent from the need to adapt to small regional variations in soil, climate, etc. and thereby has allowed a certain standardization of procedures. This is not to say that modern agriculture is a well-oiled assembly line where inputs of one's choosing may be used and the output controlled; far from it, although the trend is in that direction. (Think, for instance of the extreme of hydroponic greenhouse production of tomatoes.)

The paradox, then, is that development agencies and experts have found that in order to facilitate a change to a system which allows greater standardization of procedures and freedom from constraints, it is necessary to better understand the existing systems, the delicately balanced networks. People do want "development" and yet changing the old systems involves breaking parts of them. The question of how to manage difficult transitions arises. The difficulties lie not only in the complexities of the systems involved but also because the choices involve very tough political, social, institutional changes and the decisions are related to questions of goals.

---

[32]King, F.H. Farmers of Forty Centuries. Madison, Wisconsin 1911, recently reprinted by Rodale Press.

The paradox is then that FSR has to some extent the goal of putting itself out of business (changing fundamentally the systems it studies). The unresolved dilemma is that of the relationship of FSR and development generally to change, how to promote it, how quickly and along whose guidelines. It is quite simply not a value-free series of decisions.

One might postulate, for example, that development should be gentle and non-disruptive of existing social and cultural environments, a process of gradual and selective innovation, structured in such a fashion that the benefits of new technologies are distributed throughout a society (or accrue mainly to the "poorest of the poor") rather than going mostly or exclusively to the larger and richer farmers, landowners, etc. This is a tall order at best, and the way in which aid or development projects are structured will have political consequences regardless of whether such issues are explicitly considered (e.g., the Punjab experience cited above).

Alternatively, one might suggest that development should proceed at the fastest possible pace to bring the greatest economic good to the greatest number of people subject to the limitation of some upper limit in the amount of political despotism or cultural and social upheaval, or environmental degradation, etc., which would be considered the acceptable limit.

The point of this argument is to suggest that even though it is possible to phrase the problem so that it sounds rather like a maximization exercise, in fact every definition and level of goal or constraint is subjectively defined. The guidelines are all subject to the needs and interests of different groups or decision-makers. The techniques, including FSR models, will not define, in themselves, the direction(s) in which to proceed.

It should be clear, however, that given set goals and some decision making capabilities, the type of research which FSR workers are promoting should afford better tools to allow understanding of some of these systems and bring focus to some of the most urgent questions. To the extent that models which include simulations of farms or regional farmings systems, interactions between sectors, allocation of resources within households which are simultaneously producing and conserving units, etc., are developed, they should be of great use to decision-makers and planners in understanding the ramifications of their acts or their proposals.

Here again, it is perhaps appropriate to say that a model cannot provide data which do not exist, and in that sense, cannot generate new knowledge. A model cannot be expected to give a single best solution to a complex problem, because the questions of "best for whom" and "optimal in what sense" remain. But a good model may considerably extend what is known by allowing for hypothesis testing, by clarifying relationships, by highlighting what is not known and, mostly by offering a framework within which to deal with an extremely complex reality.

The next section of this paper returns to the more strictly biological uses of modeling.

III. Modeling and Crop Growth and Physiology

A. Crop Growth-Whole Plant Models

The descriptive progression in this paper is from the macro to the micro; moving from whole farm and agronomic models, involving various management techniques toward the more particular and detailed models of plant growth. Plant growth can be examined in terms of whole fields of a crop, including the crop-environment interactions overall and accumulated total dry matter (yield). These models may rely on models of single plant growth, which will be quite unalike for different plant species and various modeling approaches. On a still more micro level, models which attempt to explain particular processes within plants are built. These models can have finer and finer turnings, theoretically down through the various metabolic pathways and even to the molecular level. Models then, can and have been built to examine plant growth on any of a very large number of different levels. Again, the purpose of the research effort determines how the model will be built. There is no strict delineating between the various levels of organization of the models and the treatment here will be somewhat generally divided into models of crop growth and models of physiological processes. (Single plant models being relevant to both.) As a beginning, we will refer to Gates:

> "Mathematics abbreviates the lengthy thought processes involved in logic and extends these thought processes to extrapolation and prediction. It is for this reason that mathematics is applied as an analytical tool in the solution of biological problems.
>
> "Not only do all organisms live in a physical world: in every respect, they utilize the basic physical mechanisms for their viability and reproducibility. As remarkable as the biological world seems to us, I do not believe that its workings are more than an

> incredible number of physical mechanisms interacting in a large number of subtle combinations. The complexities involved are enormous and our ability to understand these is limited. Nevertheless, certain mechanisms, forces and processes may dominate the performance and behavior of organisms. Our task in the study of biology is to understand these and to recognize those of primary importance first, then those of secondary or tertiary importance. This viewpoint does not deny that every possible kind of cell-to-cell, organ-to-organ, or organism-to-organism interaction may exist. A community of organisms has many remarkable properties some of which may not be characteristic of any kind of assemblage in the universe.
>
> "A reproductionistic approach to biology, or specifically to ecology, by no means excludes a holistic viewpoint. I am convinced that a great deal of biological understanding will be achieved through analysis based on mechanisms; at the same time other approaches are worthwhile and necessary."[33]

Here again is brought up the fact that basic scientific research is by nature reductionistic: to understand whole entities, the scientist subdivides and examines the parts and subdivides again. This, as mentioned in earlier sections of this paper, brings with it the problem of reintegration of the parts to understand the whole. My contention is that one of the most fundamental offerings which systems science can bring to basic biological—including plant—sciences is that of offering a framework and a methodology for reintegrating separate pieces of research and bringing a counterpoint to the reductionistic approach.

Gates continues:

> "Relatively few people have come to grips with the most difficult and challenging ecological problems. For the most part, they have been satisfied with the quantitative aspects in terms of numbers and rates. The new extremely worth-while work concerning systems ecology provides insights to the interrelations among many components of ecosystems. Yet within such interrelations of trophic levels, the flow of energy, the flow and cycling of minerals, and the gains and losses of biomass are the fundamental mechanisms that control, regulate, and influence them. These

---

[33]Gates, David M. and Rodolf B. Schmerl, eds. <u>Perspectives of Biophysical Ecology</u>. Springer-Vertag, New York, 1975, pp. 1-2.

fundamental mechanisms are physiological and physical; they involve organisms and their environment. Once the coupling mechanisms between an organism and its environment are thoroughly understood, an extremely critical domain of physiological ecology will still remain to be worked out. This domain is in the biochemistry of metabolism and growth, resistance to heat and cold, fertility, germination and a whole complex of important biological events. Many of these events are mediated through enzymes, and the incredible number of complex, closely related, biochemical reactions staggers the imagination ... No scientific problem could be more difficult or more challenging, yet modern science is fully capable of addressing it. The analysis of this problem involves taxonomy, systematics, physiology, biochemistry, biophysics, physics, meteorology, climatology, mathematics, engineering and other disciplines. Clearly, a single investigator cannot learn all these things well ... "[34]

That is enough of an introduction for now of the difficulties and complexities of the holistic approach in biology. A number of models have been built which simulate crop plants or communities (fields) of crop plants and it is to these we now turn.

"One approach is through whole plant research but there are only a few examples where information on organ or tissue level physiology has been translated into explanations and predictions of field behavior (e.g., Ludwig et al's (1965) study of light adaptation in cotton communities). Such research with whole plants is extremely difficult and we lack the skill and genetic materials for rapid progress. Perhaps also, an important element of the system is missing—a structural hypothesis of the physiology of the community system.

"It is our contention that explanatory models of community behavior can be structured from physiological information and should go hand-in-hand, as integrative tool, with cellular, tissue and whole-plant physiology. To fulfill this research role, the models should predict field or phytotron performance of whole plants and communities. That is, they should be multi-level, with clear structural and quantitative correspondence to the real biological system; they should deal with the dynamics of development over time; and they should explain how changes occur in the system. The single-level, associative systems models of technology have a role in management but they fail as scientific research tools for

[34] Ibid., pp. 2-3.

> integrative studies of the complex interactions of morphology, physiology, and environment."[35]

There are currently several models of crop plant growth extant, including several generations of models of corn, sugar beet and alfalfa models and probably more. The models usually in effect 'grow' the plants by determining various state variables and rates which will cause incremental growth to occur as functions of light, water, temperature, nutrients, etc. The explanatory power of the models is related to how many physiological processes are included (e.g., photosynthetic rates, source-sink allocations, translocation, respiration, cell division and expansion, leaf initiation, etc.). Since not all processes can be modeled—nor are all necessary for the purposes of most models-- "black box" type associations are generally incorporated (e.g., leaf initiation as a function of temperature).

> "The need for an associative, integrative nature of many of the inputs would be true regardless of the starting level of the model. Fortunately, it is also theoretically sound when used in multilevel models since associative relations at one level become explanatory at higher levels."[36]

The way in which these processes are approximated in "associative" relationships depends partly, also, on the crop plant being modeled and the interest of the researchers. For instance, Fick's model of alfalfa growth is concerned with whole plant, dry matter accumulation as the whole plant is harvested as hay (minus root growth).[37] There is concern

---

[35]Fick, Gary G., R.S. Loomis, and W.A. Williams. "Sugar Beet", Chapter 9 in Crop Physiology, L.T. Evans, ed. Cambridge University Press, Cambridge, 1975, p. 259.

[36]*Ibid*., p. 267.

with flowering and seed formation, only as these affect total plant growth. In corn, by contrast, the ear formation is critical. In sugar beets, the root growth is of primary interest although again, of course, the whole plant is modeled.

In this light, Dr. Adams was interviewed about his interest in a model of bean plant growth as his research concerns the economic crop of dry beans. He was interested in the idea, and, in fact stated that he had had a graduate student (Carlos Antonio Burga) who wanted to construct a model of bean plant growth. The graduate student had to content himself finally with gathering information which would be relevant to constructing such a model as the current state of information was simply inadequate, and his Ph.D. thesis (1978) consisted of measuring and analyzing a number of processes in a way which would provide information necessary for building such a model. This brings up a point of interest: despite enormous amounts of information gathered on every major crop plant, and, of course, much current research, it will often be seen that constructing a model will require information about particular relationships which have not been quantified by plant researchers. This may be related to the nature of the aggregation—or disaggregation—needed by a model-builder, for instance relating the rate of a particular process to temperature, whereas a botanist or plant scientist has been concerned more with the enzymes which control pathways that cause the process. However, the modeling process may also, in

[37] Fick, Gary W. "The Mechanism of Alfalfa Regrowth: A Computer Simulation Approach". SEARCH Agricultural Agronomy 7, Lot 7, No. 4, 1977, pp. 1-27.

fact, uncover gaps in information or uncertainties which were not previously pointed out. This is easily understood in light of Gates' comments, quoted above, about the intricacies of relationships and interrelationships within biological systems. It should not be surprising that some relationships are not fully known. This will be mentioned again later in the context of how plant model-building can add to the knowledge of plant researchers.

The bean plant will be touched upon briefly as explained by Dr. Adams because he was very interested in the possibility of developing a model, and it is suggestive of the differences in requirements that different species models would have. In general it seems to be true that the first model or models of a given crop have largely to do with growing the whole plant and aggregating a great deal--expressing dry matter accumulation in the top of the plant as a function of total photosynthates and partitioning to roots, shoots and leaves for instances (e.g., Fick's alfalfa model). This is in itself a complex task. As the later generations of models are developed, more relationships are defined, more parameters discovered, more specific inputs required. (As an example see the corn model, CORNF by Stapper and Arkin which requires more inputs than earlier corn models including such things as soil data on "potential plant extractable moisture per layer," "initial plant extractable moisture per layer," "upper limit of Stage 1 cumulative evaporation" and also "leaf area of the first leaf" as well as the more usual data on climate, latitude, planting date and population, etc.).[38]

---

[38]Stapper, Maarten and Gerald F. Arkin. "CORNF: A Dynamic Growth and Development Model for Maize (Zea Mays L.)". Texas Agricultural Experiment Station, Blackland Research Center, Temple, Texas, December 1980.

The bean plant's physiology is sufficiently different from corn or alfalfa, mentioned above, to suggest the reasons why plant models are not interchangeable. A critical question is that of the photosynthetic capacity of the crop. Light interception is a function of leaf area, number of leaves, size and orientation of leaves, the vertical distance between leaves, branching, and planting density. In bean plant breeding, an objective has been to maximize light absorption and transmission within the canopy (direct and indirect or scattered light). This is largely because of the nature of the bean plant and the way in which beans are formed, which has to do with what are called "phytometric source-sink units". In non-technical language, the phytometric unit consists of a short side stem (raceme) which has a trifoliate leaf and the bean pods which develop at its base. The critical point is that the pods on any given raceme are formed and filled by photosynthates generated in the leaves on that same raceme (with some exceptions). Thus, generally, it is not sufficient for the plant to receive a certain amount of light intercepted, say, by the top leaves for good productivity. Each phytometric units develops somewhat separately--at least as far as the seeds (beans) are concerned. This means that "plant architecture" is critical, not only in the sense of maximizing total photosynthetic capacity but also in terms of how the light is distributed within the plant canopy. Dr. Adams felt that if an adequate model of bean plant development existed, it might be of considerable use in deciding the plant architecture or "architype" to breed for. He has, in fact, been breeding along these lines, selecting for vertical, nonbranching varieties which show promise for maximizing light interception to phytometric units (and also ease of harvesting as mentioned above, section II.A.). But, he still felt that modeling work, developed along these lines, would be of great use to him in that it might help clarify ways to maximize light interception to photometric units.

And yet such a project is not as straightforward as one might suppose or hope, as the following different example explains:

> "Suppose a cotton leaf is to be modeled, in which process the biophysical approach is used. The environmental input data would be sunlight, wind, humidity, etc. The fundamental constants would include leaf size, enzyme levels, chloroplast density, and reaction rate constants for various biological processes occurring within the leaf. If the concepts are correct and the various biological constants have been measured properly, the model should be able to predict the range of possible responses of a cotton leaf at a given stage of development for all reasonable environmental conditions. With a high humidity and low wind speed, equations of the model will yield one solution; whereas, for a low humidity and wind speed another solution will be obtained. At another stage of development the cotton-leaf model will have another set of fundamental constants because of physiological differences between leaves at different ages. The result will be a different solution, even for the same environmental conditions. The difficulties existing in this approach are obvious. Hypotheses, theories or established laws are often nonexistent. Sometimes the fundamental constants will not have been measured and may not be measurable except indirectly through the model itself. As the system becomes more complex, these models become unwieldy."[39]

And yet, such difficulties notwithstanding, various researchers profess their interest in modeling:

> "It seems to us that with the potato, as with all crops, we would obtain a much clearer understanding of the above issues as well as a satisfactory procedure for predicting yields, from the construction of a quantitative simulation model. Current interest in this activity and the degree of success achieved in defining the light relationships, photosynthesis, and water relations (Setlek, 1970) augur well for accelerating success although many issues have yet to be faced in handling the supply of mineral nutrients and of

---

[39] DeMichele, Op. Cit., p. 599.

generating the plant (Milthorpe and Moorby, 1974). The large degree of plasticity shown by the potato makes this a more formidable challenge than for many other crop species."[40]

The above discussion begins to lead towards a treatment of the use of models in basic research. First, though, a reemphasis: the development and form of the model will depend upon the use to which it is to be put. For instance, in modeling the rice crop as part of his Ph.D. thesis, Tirst B. Paris used a much simpler "black box" approach. The approach used in the prediction of yield is the reduction rates approach, with potential yield as the initial point and then a series of reduction factors applied to the potential yield reflecting environmental influences which depart from optimal levels.[41]

Whole plant/crop models, then, may be constructed for any of a number of reasons and their complexity will vary markedly accordingly. In some instances, as discussed in the next section, the researchers' focus is on processes which take place within the plant and on structural models which can be used to elucidate them.

---

[40]Moorby, V. and F.L. Milton. "Potato" in Evans, L.J. Crop Physiology. Cambridge University Press, Cambridge, Mass. 1975, p. 250.

[41]Paris, Tirso B., Jr. "Systems Analysis and Simulation of Upland Rice-Based Cropping Systems in the Philippines". Ph.D. Thesis, Michigan State University, 1978.

III. B. Modeling and Plant Physiology

We turn, now, to a discussion of the use of models in the even more micro context of specific physiological processes, and how the models are used. Thornley, for example, says:

> " ... modeling, which is nothing more or less than quantitative hypothesis testing, has been used as a matter of course (and with great success) in the more physical sciences for at least a century. What is novel is the more deliberate and intensive application of this method to plant physiology alongside the traditional approach."[42]

Here we are turning to a discussion not of modeling as an integrative tool, coordinating research or facilitating communication between workers delving into different aspects of a given crop or separate parts of a complex chain of processes, but those doing the most basic and reductionistic analyses. Of what use is modeling to them?

It appears that in this context, of basic research, models can have their best use as means of hypothesis testing. For example, from a recent article in "What's New in Plant Physiology",

> "Preliminary computer simulation of the reductive pentose phosphate cycle suggests that this reaction is one of the most important points of control of the reductive pentose phosphate pathway."[43]

In other words this laboratory research is to some extent being accompanied by a parallel or preliminary simulation study. This use of models is of interest, and somewhat different than that reported above. A model of a farm which projects a management change ex-ante is also an

---

[42]Thornley, Op. Cit., p. 2.

[43]Anderson, Louise E. "Metabolic Regulation of the Reductive Pentose Phosphate Cycle". What's New in Plant Physiology., Vol. 10, No. 10, October 1979, p. 39.

hypothesis testing. The difference here is simply one of application to a different type of research. The concept is the same or similar but the application is developed differently. The use of modeling in the basic plant sciences as a tool for research which complements laboratory research is of fairly recent origin. It is applied largely in areas where a great deal of information about biological pathways is known (such as nitrogen metabolism and photosynthesis). In the last few years as computer capabilities have developed and researchers have developed the skills to apply them, this use was begun.

The following example elucidates the process: a basic scientist with good information about the biological steps in a given pathway, say nitrogen metabolism, can model the pathway, the nitrogen is taken up by the plant (in most instances) as nitrate. Once within the plant, the nitrate is transformed into nitrite, then to ammonia and then into the various amino acids--building blocks for proteins, enzymes, structural and metabolic agents which do a wide variety of biological work. The steps in each process are complex and mediated by enzymes. One area of interest along these lines (one of many) is in drought tolerance in small grains. Plants which are subjected to drought stress have been found to have very high (100x normal) concentrations of one amino acid: proline (and other imbalances). By building a mathematical model of the entire process as it is currently understood, it has been possible for a researchers[44] to examine the various steps and rates and begin to determine which of the enzymes are most sensitive to drought stress (the

[44]Rhodes, David reference: Peter S. Carlson, Michigan State University, Department of Crop and Soil Sciences, personal communication, 1981.

ones which cause the build up of proline). The model in this way can aid in pointing out the next step to pursue in research. The next step, after finding out which part of the process is most sensitive to drought is to attempt to modify that particular step via genetic manipulation.

Generally, models are of use in basic science in that they can test the current state of knowledge about a given process. A researcher (or team) who feels that the various parts are understood can put them together, including estimates and measurements of various rates and test the model's working against the laboratory results. If the model predicts incorrectly, evidently something is not understood or is misspecified. Also, a kind of sensitivity-testing can be done: parameters can be varied to discover which ones are most important in changing the outcome, and the importance of research into uncertain areas assessed.

A great deal of modeling research in plant physiology is underway (see for example: Thornley[45]) particularly in the areas mentioned. As the example cited above relating to nitrogen metabolism and stress tolerance of plants is meant to show, even the most basic and technical of applications may soon have relevance for the more applied areas of agronomy.

---

[45] Thornley, Op. Cit..

III. C. The Relevance of Plant Models for Agricultural Economists.

This brings us to the question of how much of this is relevant to an agricultural economist?

The economist or agricultural economist should probably be interested both in the simplest "black-box", input-output types of models and in the more complex ones for reasons which will be examined now.

The purposes of modeling whole farms or farming regions whether under highly developed modern agricultural conditions in the U.S. or in the intricate traditional agricultural context described above under "Farming Systems Research", is generally to look at economic or agri-economic or ecological questions of some complexity. There have been instances where the research focus was on a use of a resource measured in physical terms, not assigned monetary values, such as studies of energy budgeting in agriculture. Such studies, I would argue, are still economic in nature.

More generally, however, the focus will be on criteria such as testing new or alternate existing management possibilities, looking at the long run economic viability of a type of farm or faming system, or at the value a projected new input might have, and other clearly economic issues.

The question then becomes: to what extent are the biological models capable of generating information needed by an economic analyst? What does the economic analyst need? In a study involving a crop plant, an economist would likely need at least something which can predict average yields with some level of variance or probability associated with the prediction. The simplest way to do this is most likely a "black-box" approach, perhaps, in econometric terms, a time-series model rather than a structural one. Time-series data may give reasonably reliable estimates for some work, particularly when the analyst is working with fairly large aggregations and if good information is available. Major stochastic variables such as weather may not be predictable to a degree of accuracy necessary to make more structural models better predictors than time-series models.

Technological changes can be included in econometric models in various ways. In ex poste work, a major technological change, such as the introduction of an improved variety which is widely and quickly adapted, may be treated as a dummy variable for the year it was introduced, for example. Obviously this is more difficult to predict accurately, although time-series projections might add a parameter to indicate a gradual trend variable increasing production over a period of time based on a knowledgeable guess, with reasonable likelihood of getting good estimates.

If, however, the analyst's area of interest is more specifically micro, attempts to model growth of different crops, their interactions, yield prospects changing plans throughout a season, or general farmer decision-making, more structural models will be of greater use. For example, in attempting to analyze cropping choices in a given region and farmers' decisions about planting, a large amount of biological information about crop responses may be of use. This will presumably be true particularly in areas where farmers have decisions to make about which crop to plant relative to last year's yields and what the farmer thinks this year's weather or market will be like. Other factors, such as soil types (e.g., the soil's ability to hold water in a dry year) and their suitability for crops which have different requirements, will be very important to a farmer deciding between alfalfa and corn for silage in a given year or between wheat and barley. Obviously the presence of an irrigation system would change this scenario: it would give a farmer more freedom to choose to grow the crops he felt best for other reasons besides water, and it would likely also pressure him/her to grow one which s/he thinks will return enough income to cover some of his/her fixed costs.

My contention is that, at least at the current time, the degree of model specificity derived, particularly the degree to which agricultural economists are interested in biological models or in developing bio-economical ones will probably be closely related to how "micro" the focus of his/her research is.

In farming systems research, as cited above, the focus is very micro-economic. The trend over recent years has been to attempt to define very specific interactions down to such issues as household consumption: distribution of food among household members and among various priorities (food, feed, seed, storage reserves, cash sales, etc.) as a necessary condition for understanding decision-making by small farmers who are both producers and consumers.

To the extent that researchers in this area are trying to both understand existing relationships and to change them, the need for good crop models is growing. Particularly in areas where several or many crops are grown, it is necessary to include some mechanism which can be used to look at decisions which are made more than once a year and which may include other than strict optimizing criteria for individual crops. In other words, the interactions between parts of the system such as different labor requirements for and behavior of various crops under different conditions become crucial. (See for example, Gypmantasiri et al.[46]) It may be helpful--or essential--to know, for instance, in analyzing some African cropping systems, that cowpeas require that a certain amount of water be available in the top inches of soil at planting time for the seeds to emerge and that a farmer who feels from past experiences that the water available is inadequate will not plant cowpeas but rather switch to another crop. Alternatively it may be important to know that under some conditions a farmer may plant cowpeas late and rely upon them to grow as an intercrop providing green leaves for use as a vegetable crop (while also shading the ground and discouraging weed growth) without much regard to whether or not they produce peas.

An analyst attempting to suggest an innovation, whether introducing a totally new crop or tillage equipment, or trying to rearrange an existing farming sequence particularly needs good information on crop biology, without which s/he will be unable to suggest feasible plans. The question then becomes, how much can be supplied directly through the presence of plant and soil scientists and experienced agronomists and/or farmers in the region, and what role could a model play?

My argument is that model-building may eventually be able to fulfill a function of cataloguing and making accessible separate pieces of information, which may overwise be unavailable. As research expands, and as areas where experimentation is proceeding increase in number, so

---

[46]Gypmantasin, Purek et al., An Interdisciplinary Perspective of Cropping Systems in the Chiang Mai Valley: Key Questions for Research, Faculty of Agriculture, University of Chiang Mai, Thailand, June 1980.

does the need for qualified personnel. A model cannot take the place of direct information or on the spot personnel, but can potentially, as stated before, expand the use of existing knowledge and may be an excellent vehicle of communication. For example, ideally a good biological model of a given crop would be able to give a good estimate of yield under a wide variety of climatic, soil and management conditions. It would also be able to given an estimate of how some factors would contribute to variability of yields. This is true despite the fact that there is no single unique or even best way to build a model of a given system. A model would need to be built for a specific purpose or interpreted according to circumstances. A pre-existing crop model would most likely not obviate the need for an agronomist on a given project for instance, but on some projects it might be a help both in informing the agronomist of economic issues (e.g., assessing the potential value of research such as developing a high yielding variety of one crop and the characteristics needed for the new variety to fit into a given cropping system). A model of crop growth could include information on the biological demands of that crop at given times, which would need to be supplied by purchased inputs: fertilizers, etc.—and labor: transplanting, weeding, and of likely returns on the crop under different circumstances—and thereby inform the economist of the agronomical requirements. A model could provide an information system for both sides to work with, finally facilitating the translation of field research into management practices.

Much of this sort of interaction between fields of research fields can and does take place without the aid of models but rather through direct communication between researchers and between researchers and people with field experience. I think it would be foolish to suggest that a model can provide better information than that supplied by an intelligent, educated or semi-educated person with a great deal of experience about a given area. A model, however, may be a record of information, much like a reference work or library on a particular subject which is potentially more accessible, accessible to greater numbers of people over greater time and space, just as a textbook is more accessible to people generally than its author. Then also, beyond the cataloguing and supplying of existing information a model will allow testing of new ideas, not as a total substitute for field tests, but as a preliminary test and a way to focus field research.

Some of this type of work has been done in the U.S. in the animal sciences (e.g., ruminant nutrition, herd replacement models) in forestry and in such areas as pest management. It has been done to a lesser degree in agronomy and plant sciences, but is developing now. Given the tractable nature of most field crops and the relative availability of inputs to agriculture, it has been possible to have agronomic and basic plant science and farm management research coexist somewhat separately, sharing information where appropriate. With the increasing demands placed on agriculture, changing availability of natural resources, land and labor costs and continually expanding research, biological and social, the prospect of bioeconomic modeling gains in appeal to agronomists and agricultural economists alike.

IV. Modeling and Plant Breeding.

IV. A. Mathematical Models and Crop Breeding Goals.

We turn now to a different area of plant science, that of crop breeding. Crop improvement through the application of genetic understanding is an interesting synthesis of basic science (genetic theory and plant physiology) and applied (knowledge of agronomy is absolutely essential, and a crop breeders work is judged by practical results: new varieties of crop plants released for production).

The use of modeling here is described in a somewhat different way than in the above sections of this paper. Mathematical and statistical tools, which take the form of a conceptual model of a system can be used to both describe and as a tool to help alter a system which is fundamentally not visible (the genetic makeup of the crop plant). Despite the fact that a plant breeder or a geneticist can "map" chromosomes through various indirect means (e.g., measuring recombination as an indicator of how far apart various genes must be on a given chromosome), s/he cannot simply take the cell or nucleus or DNA apart and see where the control for a given process exists. Genetics is, by its fundamental nature a process of mental modeling, and crop breeding is also.

This is true of course, somewhat, in a number of areas of basic research: a botonist explores a metabolic pathway indirectly through chemical analysis at various stages, for instance. The difference suggested here is that the crop breeder is in a somewhat unique position of trying to do a blend of basic and applied research: using the most basic research in genetics (largely but not necessarily done by other researchers) which s/he has learned and continues to learn, and to which s/he may contribute, to maintain a mental model of how the genetics of the particular crop is known to function. S/he combines that with a strategy—which is also a mental model, but a different one--of how s/he will combine the sources of genetic variability available in nature or artificially induced, to produce a better--and immediately usable product. The process however is not immediate. Depending on the biology of the crop it may take several years--or with trees, human generations--to complete a number of the matings necessary for the recombinations. (This is one of the reasons that cellular-level manipulations, recombinent DNA, etc. are currently of such interest to researchers.)

The crop breeder will generally work, with these mental models, both learned and constructed, to progress toward yet another mental construct or "model": that of an ideal plant or "archetype" or "ideotype" of a given crop species.

The archetype or ideal plant must be realistic. It is not the Woody Allen movie ("Sleeper") vision of a giant banana or bunch of celery lying in a field with a growing tube of some sort connected. It is based, rather, on the crop breeders extensive knowledge of the biology of the crop, the agronomic practices used to produce it (including what changes are feasible), the requirements of various users (e.g., consumers' tastes, millers' specifications, etc.) and his/her judgement of what delicately balanced changes are realistically possible.

Most plant breeding work takes place along fairly classical lines as outlined in a number of very good texts on the subject. In this paper, two somewhat unusual examples will be given of the use of mathematical models as part of plant breeding programs at Michigan State University.

IV. B. Mathematical Modeling and Crop Breeding: Grafius' Model

In addition to the models currently in use or being developed in crop growth, forage management, etc., there are areas of research in which model building has had little development and which offer interesting possibilities as areas for future cooperations between plant scientists and agricultural economists. One such area is in crop plant breeding. Traditionally this has been the domain of the biologists (the plant breeders, aided perhaps by plant pathologists, etc) alone, and, more recently of the biologist aided by a programmer who can write a program to catalogue and sort in various ways the hundreds or thousands of breeding lines, hybridizations, various generations and trials (for yield, resistance to pests, diseases, etc.) which the crop breeder manipulates. The plant breeder proceeds according to the system for genetic recombination and selection which s/he has worked out and his/her subjective assessment of the plant materials. Plant breeding is, thus, a mixture of a rigorous scientific understanding and personal experience, an "eye" for one's plants, imagination sufficient to envision an ideal and to develop new paths to try to attain that ideal. In other words, plant breeding has elements of both scientific rigour and of being an art. And plant breeders will at times lay claim to being both scientist and artist. Generalizations are dangerous however as plant breeders are a varied bunch.

The variety of approaches which have been developed by different breeders is somewhat surprising to someone with just a rudimentary understanding of the field--and yet there is no reason to believe that all the possible avenues of genetic selection systems have already been invented/discovered and tried, although the profession as a whole seems to lean toward the classical techniques and the well-recognized breeding systems. In general, novel breeding approaches such as the tetraploid scheme for breeding potatoes,[47] or the mathematical ones proposed below are not quickly adopted by the profession as a whole.

Some of the more novel systems which have been developed are the result of the biologists advanced understanding of the genetics of the crop and his/her own ingenuity. A few represent a line of thought which incorporates a mathematical/statistical/model-building expertise in full coordination with the biological understanding to build a new synthesis. This is qualitatively different than the situation of a biologist hiring a computer programmer or even a biometrician to work out the technical details of keeping track of a breeding program.

The example which is best known to this author is controversial among plant breeders and yet seems to have generated good results unusually quickly. The plant breeder in this instance did not develop his system in coordination with mathematicians or statisticians, rather he became an expert in these fields himself.

The question which intrigues this author is: could something similar to this have been thought out by an agricultural economist with a model-building and systems analysis background working in concert with a plant breeder or would such a high degree or coordination be impossible, or perhaps of no interest to an economist? (I am assuming the economist is not interested in acting only as a model-builder, programmer for the

---

[47]Peliquin, Stanley, University of Wisconsin.

biologist.) I am suggesting that a project which involved building a mathematical model of a plant-breeding system might incorporate information of interest to the economist as well. Such a joint research project might include an analysis of the potential value of incorporating certain characteristics into new varieties and on the likelihood of achieving such results in various time frames. It might be of interest to an economist with focus on a particular crop to deal with questions of potential changes in production packages, and information relevant to projecting what macro, long run, quantities produced and price might be.

The late Dr. John Grafius developed this model and applied it to his oat and barley breeding program at MSU. It has been continued by this graduate student Jim Nelson, who explained the system to me. First, a general background treatment will be given.

Plant improvement (crop breeding) programs evolved from a prehistoric process (for which we now believe stone age women were largely responsible): gradual development of all the major crop species by observing and selecting seed (or cuttings in the tropics) from superior plants. Many of today's breeding programs are more elaborate versions of the same process, with the additional and important steps of making deliberate sexual crosses between plants by transferring pollen and creating new combinations of characteristics.

There are fundamentally two reasons to do this. One is to attempt to combine superior characteristics of various plants into one highly superior plant. That is largely an additive process and a major part of plant breeding. It implies that there must be a pre-existing store of variability in the gene pool—or that new sources will somewhat be brought in. Creating new variability is one very important direction of modern genetic research.

The other reason for this crossing process is to discover situations in which "the whole is more than the sum of the parts" or the offspring of a cross is better (e.g., more vigorous, higher yielding) than either of it's parents, or "transgressive" rather than intermediate in a particular characteristic. This phenomenon is called heterosis or hybrid vigor. It is used in two ways: One way is as an initial step in a breeding program. Crosses are made, progeny grown out for several generations and selections made according to the breeder's goals and mental model of how best to achieve them. (Many selection systems exist.) In the second way, various crosses and selections are made to develop two (or more) pure breeding parental lines, which are known to produce offspring showing heterosis (good combining ability) when crossed. The final cross of the pure parental lines produces hybrid seed. This is the basis for hybrid corn seed and for other hybrid seed production. The seed must be produced each year because the next generation would not breed true to type, and the commercially produced seed is bought each year by the farmer. Crops which are self-fertilizing rather than out-crossing (as is corn) also show heterotic effects, but the production of hybrid seed presents large problems which are currently of great interest. Because it is difficult, almost impossible, to produce hybrid seed commercially in self-fertilizing crops, most breeding programs for small grains (self-fertilizing) for instance, involve making crosses and then growing several generations of the offspring until an approximately homozygous (true breeding) condition is reached, making selections along the way as the breeding scheme dictates. The advantages gained are partly of the additive type and partly transgressive (showing heterosis). Although the phenomenon of heterosis is well known and theories as to how it works have been suggested, the actual mechanism of heterosis is not fully understood. Nor is good combining ability generally considered to be predictable.

Most breeding programs for small grains, for these reasons—and more not mentioned—involve making very large numbers of crosses and screening the large numbers of resultant progeny for desired characteristics, at one or more points in a sequence of generations. Dr. Grafius' breeding program is unique in a number of ways, one of which being that in any given year, his program only involved making 5-10 crosses rather than the 2-3000 routinely made in some programs. In one year eleven parents were used, in a total of eight crosses, seven of which survived and yet five out of the seven outperformed not only the parents but also all other known varieties. The system is complex and involves a combination of fairly sophisticated statistical/mathematical techniques in the parental selection and crossing stages with some somewhat unusual ways of growing and selecting at later generations. The former are of more interest here and will be discussed briefly. A more complete description of the program is included as Technical Appendix 1.

The first novel technique is a method of allowing a computer to make the "crosses" and evaluate them <u>ex ante.</u> The procedure is to decide upon ideal types of plants which would be desired which combine various specific traits and which are considered to be a realistic goal. The various potential parents are evaluated from past field data as regards the various possible characteristics. Each plant's genotype (set of genetic material) is considered a vector of n dimensions where n = the number of traits being selected for which combine to an overall worth score.

The genotypes are graded for all the various traits and the grades are converted to standard deviation units. These are then weighted according to 1) the economic value of a given trait and 2) the heritability of that trait (degree to which it is transmitted genetically).

The computer "crosses" the vectors which represent the genotypes in all combinations and then "backcrosses" all hypothetical Fl's (offspring, first generation) to both parents. The combinations here represent an assessment of additivity thus far not any heterotic effects.

The computer then will generate correlation coefficients between each of the hypothetical crosses and backcrosses and each of the ideals. Those displaying the highest positive correlations will, it is hoped, give the plant breeder the best chance of progressing toward the ideal plant:

The advantages of the vector method are given as:

"

1. Can breed simultaneously for as many traits as desired.
2. Number of hybridizations is reduced.
3. Considerable time and money are saved.
4. Can evaluate fewer segregating generations.
5. Traits can be coded or modified by some measure of their economic value."[48]

---

[48]Nelson, James L. "Michigan Oat and Barley Breeding". Michigan State University, Department of Crop and Soil Sciences, CSS-408, March 11, 1981., p. 3, and Technical Appendix No. 1 of this paper.

The disadvantages, it seems to this reviewer, include:

1. There is no allowance for the unexpected, because all gene action is assumed to be additive, and predictions are made on that basis. The unexplained will, of course, not appear in the model. The savings in time and energy may be a tradeoff, somewhat, for missing a great unpredicted plant. It may be that those are so rare as to justify the tradeoff, however.
2. The model is an empirical or "black box" model in the sense that some of the genetic mechanisms are not fully incorporated. For instance, two traits (or components of traits) of the "vector" might be controlled by genes on the same chromosome and close enough in distance to be positively linked or correlated. To the degree that the traits discussed are polygenic this may not be a problem. Gene interaction (heterosis, epistasis) would also not be modeled. Gene additivity is assumed with the exception of the yield component part, discussed below.

Plant breeders consider yield an "artifact" in that it is a by-product of a number of processes. It is possible (sometimes) to select directly for disease resistance by infecting a plot and finding which plants remain alive and healthy and similarly for other resistances to stress. Yield, however, is affected by photosynthetic pathways and rates, translocation of photosynthate within the plant, storage capacity of various organs, etc. and these are in turn complex pathways under the control of a large number of genes. Much research has, as mentioned above, gone into the understanding of these and other physiological processes, and some attempts have been made to consider developing plants with particularly efficiency in one pathway (e.g., recent work on

C-3, C-4 and CAM photosynthetic pathways, efficient nitrate uptake, etc.).

Dr. Grafius' strategy was to look at yield in a different and in one sense more simplistic way—breaking the "artifact" or "black box" into sub-artifacts or smaller black boxes which can interact: the "principle components of yield", where

$$W = X \quad Y \quad Z$$

W = yield (by weight)

X = number of tillers (shoots bearing "spikes" or seed heads) per unit area

Y = number of seeds/spikes

Z = weight/seed

(Obviously this formulation applies to small grains,not alfalfa, sugar cane, fruit, etc.).

X, Y and Z are multiplicative, can be viewed as sequential (not strictly true biologically) and highly negatively correlated. If this last, referred to as "component compensation" were not true, increasing yields would be much simpler, in fact with positive correlations, yield would go to infinity, which is clearly not the case.

In this model, correlations are determined between the components, using Sewell-Wright pathway coefficient analysis. Environments are broken down into R_1, R_2 and R_3, and genetic-environmental interactions assayed.

Ultimately simple linear regressions are run of yield components on one another, say X against Y, to look for "outliers" or observations off the X, Y line and with a high Z value. It is considered that these "outliers" are the parents to be used to increase yields. They will given the transgressive jumps in yield.

These X, Y, Z weights are entered into the vector parental selection program, a few crosses are actually made, seeds are increased and sent through sometimes novel (e.g., "single seed descent" and "head hills") forms of selection which retain a lot of genetic variability while culling some less-desirable traits and succeed in creating a plant population where the entire distribution of desired traits has been significantly improved, as the results have shown.

This novel breeding technique has not yet been widely applied, and remains to be demonstrated in other crops. The results in the oat and barley program have, however, been very good.

IV. C. Mathematical Model of Genetic Distance: Adams' Model.

A second example will be given below of innovative and fairly sophisticated use of statistical tools to extend both the theory and the practice of crop genetics and crop breeding.

Dr. M. Wayne Adams has been working with the concept of genetic distance between the varieties within a given species, particularly in *Phaseolus vulgare* (beans). A little bit of background explanation follows. Individuals within a species represent members of a gene pool in a large sense, in that they are historically related and are capable of crossing sexually and producing viable offspring or seeds, whereas individuals from different species generally do not mate successfully. Still, there exists a very great deal of variation within a given species both "phenotypically"--in terms of visible or measurable plant characteristics such as seed coat color red, white, black beans--and "genotypically"--the actual genetic makeup, differences in the alleles of the genes. (The genotypic differences cause the phenotypic ones. Because of the various ways in which genes interact [dominance, epistasis, etc.] genotype is not always immediately obvious from phenotype. This distinction will be brought up again below.)

Variations between different varieties or races or lines of a crop species grown today are the result of many years of evolution, natural or directed by agriculturalists, in separate ecological niches from some original population. Wheat for example developed from a genetic fusing of three different wild ancestors--most likely a series of natural but rather uncommon events and early planters who observed and saved seed from plants with larger or different seed heads. Since that ancient time wheat has been grown for thousands of years and carried around the world. Under different climatic and soil conditions and different cultivation practices, different qualities within the original, probably quite heterogeneous (mixed), gene pool were favored. That, plus occasional mutations, plus selection by humans for favored characteristics all account for the large differences among varieties of wheats, and similarly for other crops, beans, etc.

Dr. Adams' research on genetic distance involves an attempt to analyze the extent to which some major types of beans are genetically alike or "close" or unalike and "distant". Before explaining how this research was approached, we should explain why it is of interest. There are several possible reasons. The main one cited by Dr. Adams[49] had to do with the concern, often voiced in agricultural circles, particularly after the southern corn leaf blight epidemic that the genetic base of the major crop plants is too narrow. In other words, that the varieties of corn or wheat or beans, etc., grown are so alike and carry so many of the same genes that it might easily happen that a very large proportion of one crop would be destroyed by one disease or insect pest. This was true in 1970 in the U.S. when a large portion of the corn crop was destroyed, the new pathogen being able to travel with great speed through a population all of which carried susceptibility in one common gene (very closely linked with the gene for T-type male sterile cytoplasm which had been incorporated in most of the major varieties of hybrid corn). It

[49]Adams, M.W. and J.V. Wiersma, "An Adaptation of Principle Compnents Analysis to an Assessment of Genetic Distance". Michigan State University, Agricltural Experiment Station, Research Report No. 347, March 1978, and Technical Appendix No. 2 of this paper.

happened in Ireland in the potato blight of 1848 with more severe consequences for that population, as potatoes were the universal staple and no comparable (inexpensive, nutritious) staples were available.

The question of just how closely our crop varieties are related is not always easily answered. "Pedigrees of modern varieties are often complex, or do not exist due to the particular breeding method followed."[50] Dr. Adams' analysis of genetic distance offers the possibility of tackling this problem with quantitative rather than subjective analysis.

Much might be learned about the process of genetic diversification if it were possible to find a situation in which a particular variety or type of a crop species had been introduced at one time into an area where none of that crop was grown before and if it were then grown locally for many generations within a physically defined area. Dr. Adams has in fact found such a situation in Malawi, East Africa, where a great number of different types of beans are grown within a relatively small geographic area. The different varieties are prized for various qualities and an astonishing number are grown, and yet is is known that beans were unknown until introduced into this area by colonizers, approximately 300 years ago: a perfect laboratory for studying genetic distance (at least until the farmers all switch to one or two new high yielding varieties).

One further avenue of research will be pointed out. A very important, but not thoroughly understood element in crop breeding is the phenomenon known as heterosis. As mentioned in IV A and B above,

---

[50] <u>Ibid</u>., p. 2

heterosis is the name given to describe what happened when the offspring of a mating is better (from the crop breeder's point of view--but often such things as more vigorous and faster growing, more or larger seeds, etc.) than either parent, rather than being somewhere intermediate between the two as might be expected. Various possible explanations for the workings of heterosis exist but none has been proven. As a rule of thumb it is true that the more widely divergent in origin (and sometimes in physiology) the parents, the more likely heterosis is to be present. This is generally, but far from universally, true. A more precise measure of genetic distance than now exists might shed some light on this mechanism.

Dr. Adams developed a very interesting statistical analysis of genetic distance. It is included here as Technical Appendix No. 2.

The next section of this paper will approach a very different topic. Having considered the use of biological and bio-economic models in various scientific and production and management contexts, I wish to turn briefly to models used as policy tools.

V. Modeling as a Policy Tool.

A. The Link Between Policy Framework and Systems Approach.

The prospect of using modeling as a tool for policy analysis *ex-ante* or *ex-poste* is particularly interesting. It could be said that since very few actions are devoid of political and social implications--and research is, as is contended at various points throughout this paper, laden with social effects and sometimes political ramifications--modeling as a research tool also has policy implications by definition. Also, to the extent that any theory can be considered a model, an organized series of postulates about how the world--or some part of it--functions, policy, which involves a lot of theorizing, and systematic thought and building a framework or "paradigm" with which to analyze a system, is not far separate from modeling. What this section will attempt to address mostly, however, is the explicit use of system analysis and mathematical models to examine issues which are generally considered under the rubrick of "policy".

An immediate--if simple--example of a quasi-mathematical model comes to mind. In a simple economic analysis often used for teaching purposes, the "Edgeworth box" in this instance, used in a classroom example to demonstrate the basic value of trade in a "two-consumer, two-good world" the instructor said, quite simply: "Once you give me initial endowment, it is all determined: you give me initial endowment and I will give you the outcome."[51] This very simple example is cited because it is indicative to my mind of how the most basic and overly simplified of deterministic models (here portrayed either graphically or

---

[51] Koo, Anthony, Michigan State University Department of Economics, Classroom lecture, EC 800, 1980.

through simple arithmetic) can be, at once a matter of simple logistics and also a means to begin to examine the underlying structure of a system—in other words, to approach questions of choice and of policy. From this model come questions of the importance both of initial endowment and the framework (here, the Edgeworth box, the individuals' preference functions, and the rules of trade and ownership). The next step is simply to ask: if this is what they are, how did they get this way, do we want them to continue like this, or if not, how do we change them? Policy questions are born.

In the two person world example, given initial endowment of two goods and preferences (the "utility functions") and rates of trade along the line of "two apples for three oranges", implicit prices can be determined and one (or more if goods are not divisible) solution(s) determined. It is generally assumed that the individual with the better endowment is in a stronger bargaining position, of course. However this is only true if a number of other assumptions hold. For instance, if one individual is better at bargaining, or if initial endowments are not known the solution may change. If for instance, both individuals knew the total number of apples and oranges but not their initial endowments (or if their initial endowments were both zero and they were splitting up a found pile of apples and oranges) the outcome would be based on preferences only and might be more equitable in terms of equal satisfaction. Bargaining ability, and any other rules of trade or ownership they might make up would be relevant also. As another alternative, if initial endowment and preferences were known but if, after trading the individuals had to draw straws to see who got which pile (ownership of the new distribution) the outcome of the trading would probably be different again: it might be to equalize apples and oranges and then plan to bargain again, or it might not, perhaps depending on predilections for gambling (or "risk").

One point of this very simplistic argument is that if the different situations are modeled before the trading takes place, and the probable outcomes examined, more informed choices about policy--the guidelines for setting up the system--can be made. The choices are economic and political. And the difference here between an economic theory and a mathematical model is one of degree. In such a simplistic example, the theory can be easily transformed into a model. In the real world, the complexities make that task considerable.

The second point is that, even in a simple model of this nature, one must be aware of the underlying assumptions and the resulting implications when dealing with social rather than biological systems. It is easier to keep up one's scientific rigor and question all assumptions when dealing with a biological system and not one's own societal framework.

In the sense of being able to test out alternative scenarios and to experimentally alter existing assumptions and frameworks, mathematical models are already becoming more and more a part of strictly policy analysis. Several approaches are taken below.

V. B. Information Systems for Policy Analysis.

> "As science develops new information that applies to the growth and development of man, animals, plants and pests, the number of alternatives increase for solving some of the most complex problems facing society, in general, and agriculture, in particular."[52]

On the simplest level of looking at agricultural models, here particularly farm management type models, they can be considered information systems. As mentioned before, a model could be considered as a library or file of information. Where systems such as Michigan's TELFARM or Canada's CANFARM exist, large data files on farm size and type are built up—and constantly updated by the farmers themselves. Dent and Blackie make a strong argument for the potential strengthy of modeling as part of—and as a means of generating—information systems for agriculture.

The idea is that while an individual farmer's records would be confidential, there is no similar need to keep confidentiality about aggregation of data from a large number of farms.

> "Data concerning yields of crops, livestock performance, area of land under different crops, number of different types of livestock, quantities of different types of inputs are involved as well, in some cases financial information about input costs and selling prices: management practices are also detailed. The computer files therefore hold current information in the form of both raw and analyzed farm records ... Data retrieval from an industry as widely dispersed as agriculture is a major problem for policy makers in government, commerce, and industry serving agriculture. The demand for information to permit rational policy and marketing decision making by these agencies is growing in all countries of the world and the farm-based management-information system potentially provides a unique and detailed data bank ... The attraction of this concept is that it reduces the need for statistical survey

---

[52] Mahlstede, J.P. "Concepts, Organization and Funding of Interdisciplinary Research." Iowa State Journal of Research. May 1975, Vol. 49, No. 4, Pt. 2., pp. 629-633.

> (although obviously it does not eliminate such work). Because the farm manager has a direct return from his recording efforts (in terms of valuable on-farm information for business control) his records are likely to be accurate. The data then available for policy decisions are current and reflect the existing environment of the industry."[53]

Such a scenario presents some problems, particularly the possibility that the sample of farmers using such a system, might for a number of reasons, not be representative of the total farming population, but such problems are not insurmountable. Here the advantages of a largely user-generated information system, which are considerable, would have the effect of also benefiting—providing information to—other users.

---

[53]Dent, J.B. and Blackie, M.J. *Op. Cit.*, p. 168.

V. C. Models to Test Policy Prescription.

Models, then, can be of benefit as sources of information to policy makers. To the extent that information from farm management models can be incorporated into larger, macro models of the functioning of whole farming regions or commodities or the agricultural sector as a whole, the policy effects are more direct. Just as pieces of biological research can be combined into whole plant models and various management schemes tested, it should be possible--if difficult--to form aggregated models and test policy options in areas such as performance of marketing boards structured in various ways, or analysis of industry structure, or direction of research.

Clearly some very complex problems exist. For one, there is a question of intricacy of an economy and what level or levels of organization to work from. Conceptually a plant model could eventually be devised which would be specified right down to the molecular level and the model then run under "every" conceivable condition. This would be a truly massive undertaking and the costs would presumably outweigh the benefits by several magnitudes, but short of that level of specification, some relationships will be "black-box" or associative ones and sources of stochastic variance. Just as plant scientists generally can operate at a level of aggregation which relies on some "black-box" associations at a lower level, the social scientist can legitimately work at various levels of aggregation. Furthermore, with a model of an economy, efforts to model every individual within the economy are clearly unreasonable. Another problem is raised here. It may* be reasonable to proclaim as did DeMichele and Gates (quoted above, p.21) that "random

*Although many would contest this philosophically and scientifically.

events do not occur" and that the workings of the biological world are not "more than an incredible number of physical mechanisms interacting in a large number of subtle combinations," but it is more difficult to be this deterministic with people. Economic theory does, however, include a number of behavioral theories—neoclassical, institutional and others—which somewhat explain, or predict, the behavior of groups: consumers, producers, managers, labor, etc. These theories involve and evolve from a systematic manner of thinking, a structural framework, parts of which are continually being researched and subjected to hypothesis testing. Econometric studies are frequently along these lines. In specific contexts, and given the existence of an economic model which behaves fairly well, it is possible to model, particularly using simulation techniques and stochastic variables, various policy prescriptions in an effort to project likely outcomes of alternate choices under different states of the world (which are also judged to have various probabilities of occurrence) each of which will likely have different costs and benefits.

An example of this is given by Christensen and Mitchell.[54] The authors used the MSU Agricultural Model,[55] which is an annual model of U.S. agriculture and of world grain trade to predict what would be the likely outcome of a proposed institutional innovation: a farmer-held soybean reserve. The authors caution that this simulation included the

---

[54] Christensen, Thomas and Donald O. Mitchell. "An Analysis of a Farmer-Held Soybean Reserve," mimeo, Michigan State University Department of Agricultural Economics, August 29, 1980.

[55] Mitchell, Donald O. and John W. Ross. "A Forecast of Grain and Soybean Exports in the Year 2000". Michigan State University, Department of Agricultural Economics, July 20, 1981.

assumption of continuing trends in U.S. and world yields, but a stochastic version of the simulation allowing yields to vary is being developed. Recognizing the assumptions of any model (or any theory) is crucial to using it properly. This simulation suggests that (given circumstances similar to those assumptions) the proposed board would be expected to improve soybean price stability--and corn price stability--and to reduce total government payments associated with wheat, feedgrains and oilseeds. This type of analysis is clearly of great use for policy makers. Another proposed use of systems-approach and modeling which attempts to move the use of the model beyond being a tool for analysis is cited next.

Dillon talks at some length about using a systems approach as a way to organize and to direct research. His emphasis is largely on the idea that the hierarchical organization of the systems approach allows examination of what he feels <u>should</u> be the orientation of scientific research: goal oriented problem-solving where research directions in the physical (which he calls "passive and reactive") sciences should be directed by social (he includes the biological) sciences (which he calls "purposive systems") which are in turn directed by politicians ("employed by the overriding purposive system").[56] I think his analysis of this hierarchy is built on a false distinction. He confuses the lower to higher order systems classification with the practitioners of the levels. (Politics may involve more goal setting and decision-making and purpose in that sense than geology or physics, but this certainly

---

[56]Dillon, <u>Op</u>. <u>Cit</u>., p. 8.

does not mean that the physicist is "passive" and the politician "purposeful", nor that the politician is able to make reasonable judgements about research in physics.) Dillon feels that all science, basic and applied, should be integrated into a great system directed from the top by goal-setting there, as the levels below determine how to execute the goals. Science becomes a part of normative problem solving. Despite my beliefs that science is not neutral politically or economically, I consider Dillon's proposal to be behaviorally very shortsighted and an excellent way to reduce rather than to enhance the productivity of basic sciences. The image of centrally directed research is not unlike that of centrally directed economies in that when the autonomy is removed from an entrepreneur or a scientist's world, and decisions are made from afar by those who cannot have adequate information about or understand the intricate workings of a particular firm or market—or scientific experimental system—the directive are likely to be both heavy-handed and the cause of a decrease in productivity, due both to ignorant bureaucratic regulations which reinforce the wrong behavior to get the ostensibly desired results, and the subsequent decrease in motivation of the individuals.* Without good progress in the basic sciences, Dillon's view of problem-solving research will come grinding to a halt.

Nonetheless he addresses some interesting questions with important implications for applied scientific research. They are policy questions which bring up any number of difficult dilemmas. As an issue for further discussion, part of his argument is reproduced here:

*Since the initial writing of this paper in 1981, I have spent ten months in China, where scientific research is centrally directed. That experience supported this argument to a much greater degree than I had expected.

"As soon as we look at a piece of agricultural research either in the context of the higher systems in which it is embedded or in terms of the lower systems which it influences, we can see implications which go unrecognized in the traditional disciplinary approach to research. The reason is that optimization or improvement to merely a part of a system or a sub-system cannot be presumed to lead to an enhanced performance of the overall system--indeed it will generally result in overall suboptimization ... Traditionally agricultural research has been production oriented ... The goal and the result have been either the same output from fewer inputs or more output from the same inputs. To no significant degree has our research been oriented to encompass aims of social justice for farmers or domestic consumers. Benefits in the main have gone to well-to-do producers. The poor and disadvantaged, whether inside or outside agriculture, have been neglected if not positively disadvantaged. In the USA, for example, the development of cotton varieties specially adapted to the Southwest, along with the development of irrigation there, led to a much more efficient production of cotton. As a result, cotton production shifted westward and whole communities in the Southeast were left with unemployed farmers and workers, not to mention shopkeepers and parsons. Many of the displaced migrated to city ghettos in search of non-existent jobs for which they lack training anyway. The net result was that a piece of agricultural research which was excellent from a disciplinary point of view generated great socio-economic inequities and added further to already overcrowded city slums. As some have argued, it is indeed a tenable view that the terrible riots of a few years ago in Watts, California, were basically caused by the 'success' of production-oriented research which led to the displacement of agricultural labour in the USA—for example, the mechanical cotton picker, hard tomatoes, etc. The results of such research gave too much to the wrong people and ignored significant (human) costs.

"The first part of my argument is that the systems approach can provide an indication of where research is needed--not just in the trivial sense of being able to say to a plant physiologist that we need an estimate of such and such a coefficient, but in the far more significant sense of recognizing disadvantaged regions or groups and problems of farmer and rural equity or security that lie outside what has been the traditional focus of research interest. We need to recognize—as a systems approach emphasizing purposive higher—level systems makes explicit—that there is more than a material product to agriculture. It may be that in the future the question to be answered will be 'How much would it cost, how much are we willing to pay and where should the money be spent in order deliberately to foster a less economically efficient system whose social values add greatly to our entire national quality of life?' That question is already a live one in the USA. Its evaluation will certainly need a systems approach.

> "The second part of my argument is that the systems approach provides a logical and workable procedure for obtaining a guide to the likely ramifications of research."[57]

What he calls here the second part of his argument is, I think, the stronger argument. As discussed above, to the extent that systems analysis can be developed to act as a tool to increase policy analysts' or decision-makers understanding of the likely outcomes of their recommendations, their decisions, (in this case which research is to be funded) it can be enormously useful (assuming here, as in much of this section, that simulations with good predictive powers can be built for such complex systems).

The first part of Dillon's argument, here cited at length because it brings up interesting and important questions, contains some implicit assumptions which are troublesome. The argument I wish to pursue here is essentially the same as that followed under the FSR section, namely that one should not be gulled into thinking that the model or the process of modeling will solve what is fundamentally a question of subjective judgements. Dillon cites "National quality of life" but I maintain that it is simply not a neatly quantifiable or definable entity: it is a function of the point of view of the people involved. Some would prefer lower priced but hard tomatoes and some not. Who's view counts on the quality of life scale? There is another very important aspect to this question which was not brought up by Dillon, but it is totally analagous to the problem as presented under the FSR section: it is the question of change. Dillon objects that the social costs of the change in cotton production technology were too high. But if traditional agricultural

---

[57] Ibid., pp. 11-12.

research has been too narrow in focusing on only one thing, production, Dillon is focusing on only two: production and social costs, and not including social benefits (he may be assuming there are none to hard tomatoes and mechanical cotton pickers); nor does he touch on the much more difficult question about how to make choices about change, how to evaluate social costs and benefits equitably when there are no adequate numeraires for some of the issues. If the choice had been at some point not to develop a mechanical cotton picker, producting cotton would then on necessarily have involved hand harvesting. Cotton-picking has been amply celebrated in song as physically exhausting work. A given worker may prefer a job picking cotton to no job at all, but in the long run it may be more satisfactory to people seeking jobs to have choices which include other possibilities but where cotton picking has not become, by (governmental?) decision a job which will never be outmoded. Dillon's argument includes the fact that the displaced workers have no job training for other jobs--but is institutionalizing a hard labor job the best solution to that? Is it not rather condescending in the assumption that this job will continue to be needed? There are reasons other than profit maximization for allowing mechanical cotton pickers.

It seems to me that the real questions raised here are about rate and form of change and who benefits and who is hurt, just as they are in international development work. These questions are far more complex in conception than what Dillon described. The interactions which should be in the system to be modeled would include questions about job training, education or lack of it in underdeveloped and ghetto areas, job availability, discrimination, desirable and not desirable work, (by whose preferences), etc. My argument again is that the systems approach does not solve these questions. All that the systems approach and mathematical modeling specifically may be able to do is to handle greater numbers of variables and more interactions than most human minds can manage, and to allow experimentation with a model to try to gather insights into how systems work and how they can be changed. This is a very important point. Unless this is understood there is a hidden danger here: to build models of such complex social interactions would be chancy in the sense that such models would of necessity include large numbers of judgemental elements and the results might be anything but determinant, and yet be given weight of undue authority as a "mathematical (i.e., proven) solution" in peoples' minds.

V. D. Numeraires Other Than Money.

In the discussion so far it has been implied, but not stated outright, that an advantage of modeling is that it offers a possibility of looking at flow measurements and analyzing a system using measurements other than the simple "numeraire" of money. Where there are incidents of "market failure" in the generally accepted sense that for some reason (e.g., externalities, public goods, poor information, dissonance between long run and short run reinforcers, etc.) the market prices do not adequately reflect preferences to get individually or socially desired outcomes, this quality might be of use. The simplest example I know of this currently is in the use of energy accounting to look at various processes in agriculture and in other areas. The argument of some proponents of this measurement of production and consumption in terms of energy units rather than in terms of dollars is that current market prices are too low, they do not adequately describe the dependence of our agricultural processes on petrochemicals and therefore give biased results.

Evans suggests a similar idea:

> "In Biblical times yield was assessed in terms of the number of grains harvested for each grain sown, and thus foregone as food, even in years of famine. As agriculture became more settled and good arable land more scarce, yield per unit ground area became of prime importance and, with increasing pressure of population, remains so. We are reaching the stage, however, when other bases of yield estimation must also be considered, such as yield per unit of water or phosphorus or energy consumed or pesticide discharged into the environment, which may lead to a reversal of some past trends."[58]

---

[58] Evans, L.T., ed. Crop Physiology, Some Case Histories, Cambridge University Press, Cambridge, 1975.

There are, of course, economic ways to do accounting of such things: as resources become scarcer, their prices go up, when pollution is perceived to be a problem, institutional arrangements can be made which cause polluting to become a monetary cost for someone (e.g., pollution control standards, environmental clean-up campaigns). The point here is simply that modeling has inherently the flexibility to follow any numbers of different kinds of flows through various processes and to account for them in different measurements, which can be used rather like the traditional numeraire, money, in various analyses.

V. E. Modeling to Explore and Generate New Systems.

Modeling also offers, to an extent, a way to explore adjustments to systems or sophisticated changes in technologies and to work for their implementation, sometimes before the market has had a chance to act in its famous role of inducing technological change. I am thinking here of the example of IPM - integrated pest management mentioned above. The need was felt, in a number of quarters, to reduce the use of pesticides on crops and pesticide residues in the environment, and to maximize the effectiveness (efficiency) of the pesticides used. IPM was developed not by the private sector, where it might be anticipated since the greatest expertise in managing pests was in the chemical companies manufacturing pesticides. The motivation to reduce pesticide use was only as strong as public opinion ("voice") made it and the research took place in the semi-public domain of the land grant colleges. My suggestion is that this is a good thing. American agriculture and agribusiness are both developed to the point where extensive research and development capabilities exist in the private sector and also good contact between agribusinesses and farmers. The major agribusiness and seed companies have taken on some of the functions (fertilizer and pesticide recommendations, new seed production) that once were the sole domain of the state experiment station. This can be seen as part of the function of the experiment stations: to pass on some duties to the private sector, including the research and development and extension that these companies can afford to do as part of their production development and promotion. The experiment stations and colleges can then focus not on doing cheap research and development for the companies but in exploring new areas of research such as IWM or "Integrated Weed

Management" as suggested above, and further research in modeling agricultural systems.

V. F. General Policy Questions.

There is one last point to be brought up here relating to policy and modeling: it is what we began with in this section, that it is possible that modeling can be of use in examining some general policy questions--which would become less general as knowledge builds up. As an example, there are experts working in areas of development (e.g., Stillman Bradfield)[59] who have suggested that a goal of agronomists and economists working on projects designed to help small farmers should be to concentrate on diversification in enterprises for these farmers as a means of improving their options and ability to adapt to unstable and unpredictable markets, governments. Diversification in this context would mean a number of crops, livestock, storage facilities and also perhaps locally-based institutions such as buying co-ops. To the extent that farmers can sell, or feed, or store, or consume at home their various crops they are better protected from adversities than if they grow one or two crops which must be sold to generate any income. But this philosophy is not shared by all: there are also, after all, economies of scale and advantages to specializing and acquiring particular management skills. Also it may be easier to develop export or domestic markets, etc., as farmers adopt a given crop and produce it consistently for sale. Many development projects are very clearly aimed at increasing a farmer's ability to produce one or perhaps two crops and at simultaneously building a local marketing capacity (which may feed into a national or international one). Bradfield's arguments are almost

---

[59]Bradfield, Stillman. "Appropriate Methodology for an Appropriate Technology", paper delivered at Annual Meeting of American Society of Agronomy, 1978, or mimeo, Kalamazoo College, Kalamazoo, Michigan.

anarchistic by comparison. And yet there is a strength also in diversification. It is stressed particularly in biological systems: the need for a wide (diverse) genetic base, the strengths of multiple cropping. Monoculture requires a lot of technology and management biologically (better and more complete insect and weed control and a research system which can keep one jump ahead of new strains of pests) and more infrastructure (markets, etc.) economically. Diversification in some instances has its costs in productivity, but its returns in terms of resilience and adaptability. The difference between successful specialization and unsuccessful specialization (and therefore extinction) of a biological species may rest largely on the degree to which the environment remains friendly. If a species' environment changes significantly, it must either adapt, move or die. Similarly the economic forces of change and development make some modes of production obsolete, but as Dillon (and everyone else) points out, there are social costs and they are distributed unevenly. Such trade-off questions are part of the essence of policy. It is reasonable to hope that modeling of various scenarios will be able at some point not to provide answers, but to enhance the information and understanding of those analyzing and making policy.

VI. The Agricultural Economist—Plant Scientist Relationship

A. Multidisciplinary Teams and the Degree of Research Coordination

The subject of this paper has been the use of models in agricultural sciences. The focus extends from the plant sciences into agricultural economics and the argument is that the use of a systems approach and, in particular, constructing mathematical models offers a structure to enable researchers in these fields to communicate and to work on common areas of interest—which are sometimes called problems. The aim of the paper is to offer some insights into common ground and into the thought processes of the different areas.

At issue are some long-recognized difficulties, in particular the conflicting needs of those engaged in disciplinary research and the demands of real world problems which often do not respect disciplinary boundaries. There are many demands that research be applicable to the problem at hand—in both the basic and applied sciences—and that the researchers produce solutions to the problems. The results have been mixed, as the literature shows, but few concrete solutions are offered to the problems of facilitating interdisciplinary work specifically between plant scientists and agricultural economists. Several points and suggestions will be made here.

A good place to begin is with the question of how much technical background in plant sciences is needed by an agricultural economist who finds himself working with plant scientists. (The converse: how much economic understanding is needed by other agricultural, biological scientists, is not dealt with here. This is my bias. I think that both groups would gain from some broadening of understanding but I am here simply postulating that the agricultural economist is being the more intellectually flexible team member!) The agricultural economist was a hybrid from professional birth—a mixture of plant and animal sciences, farm background and economic theory—but displaying "heterosis" (as the field grew) in the ability to deal with the economics (policy, marketing, production) of agriculture in a very broad sense, and now is often far removed in training from the farm and from biological and technical sciences.

The extent to which an agricultural economist now needs a technical understanding depends on several things, most notably the kind and nature of work s/he wishes to do with agricultural scientists, which can be roughly divided into three categories:

1. Economists are often called upon to do an <u>ex-poste facto</u> analysis. After a project is done the group or some administrator may want "the economics of the thing" looked at. A hypothetical—but realistic—example is of an economic analysis, after several years, of the impact of high-yielding varieties of rice on a country or a region. An economist is on relatively safe home ground in doing costs and benefits in monetary terms, examining changes in production and perhaps social changes. Depending on the depth of the survey, s/he may have to learn a fair amount about rice production methods, but probably not a great deal about the rice plant. His/her dealing with rice plant physiologists, breeders, etc. might be extensive, but probably not.
2. The second scenario is of a predictive or <u>ex-ante</u> analysis. It is possible to analyze the expected results of research and to attempt

to direct the research accordingly. As an example (again hypothetical) one can image the initiation of research on a plant which is very important in the diet of a large number of people but which has as yet received little attention from crop breeders. Several tropical root crops fit this description. If the goal of a research establishment were to improve, as rapidly as possible, the planting stock available of such a crop, a two-pronged investigation could be begun: several plant scientists including a plant breeder would be asked to look at the stock currently being grown by farmers in a certain region and assess what might be done, through crop breeding methods, to improve it and to offer a judgement as to the relative difficulties—or likelihood of success, and how long breakthroughs might be expected to take—in dealing with the apparent problems.

The second prong of the investigation would be agricultural economist(s) and agronomist(s) examining the farming system, interviewing farmers, and those with local knowledge about the crop, as to the greatest problems in production: how often are the diseases an important problem, where do they get their planting stock, can they get fertilizers or pesticides, can they get credit to buy them, how do they use the crop, how important is it, do they have good markets or other uses which would allow them to use more of it or to use it in different ways (e.g., feed it to livestock). How much is more production worth if it is based on more expense or more labor?

The various team members then can consult with each other and combine the information. The economists could set up estimates of budgets and begin to determine how much various improvements would be worth to the farmer, the consumer, the government etc., and these estimates could be used by the plant breeders, as a guide in setting their research goals. It might be, for instance, that some varieties of a crop could more easily be improved than others but it would be important to know how acceptable they were—if one variety tastes as good, makes as desirable a product when cooked, stores as well or has as ready acceptance in the market—as another. Using such information as guidelines for a research program would be expected to yield more usable results faster than would a program without such a study.

This ex-ante research would also require a broader knowledge on the part of the economist who would have to communicate well with the plant scientists and the local farmers, market people etc. S/he would not have to know the genetics of the crop but would need to be able to understand the projections of the plant breeder and to be able to relate what the scientists and the farmers said. In sum, it would require more involvement in the technical side than an ex poste analysis. For a more involved scenario like this, see Cock, "Biologists and Economists in Bongoland"[60]

3. The third scenario is of a cooperation extending from beginning to the end of a project, where the agricultural economist and the plant scientist work together in formulating, designing and carrying out some research together. The area of model-building offers the prime example of this. If a plant scientist and an economist and a modeler set out to build a model together, the economist's need for

---

[60] Cock, James H. "Biologists and Economists in Bongoland", in Valdes, Scobie and Dillon, Economics and the Design of Small Farmer Technology, Iowa State University Press, Ames, Iowa. 1979.

technical understanding will be considerably greater than in the two preceding cases. These thoughts are expanded in the next section.

VI. B. Modeling and Communication, The Modeler as Translator.

In model building, probably the ideal situation, or at least the simplest, would be for the scientific, economic, and mathematical/statistical/modeling expertise to all be available in one highly trained individual, or a group of individuals all of whom understand these areas and were ready to translate their understanding of the crop biology and the economic environment into mathematical functions. At least this should minimize communications problems. The difficulties with this approach, of course, are that such individuals are not always available, and that training of single-field specialists has certain advantages in economies in division of labor over trying to train individuals to be multi-field experts.

It is fairly clear however that there has to be some common ground of understanding in order to communicate. In this argument, the economist is hypothesized to be an economist-modeler. Alternatively, the economist could be talking to a biologist-modeler or several economists and biologists with various specialties could all be talking with a modeler as interpreter. The arrangement is not crucial for the argument. The point is that the model-building process could create a bridge to carry information, create communication between disciplines and the practitioners of those disciplines where such communication is difficult because of either the specialization of technical languages and concepts or because of the numbers of people and specialties involved.

The scenario in this example is that an economist/modeler sits down with a plant scientist who explains the workings of the plant in terms of physical processes, rates of change, flows of materials, stages of growth, and the economist/modeler translates these into a series of mathematical/statistical functions which can represent the described processes. Various plant scientists might become involved in the effort of organizing and ordering data, hypotheses of how parts of the system work. The modeler in turn presents a mathematical description of what the plant scientists have reported, asking for clarification of some parts, for consensus on some process which has been described by different people in different ways, for an assessment of how sure some numbers are. In this sense model-building is rather like being a translator taking a large amount of information written—or reported—in one language and writing it out in another. But the languages are of biological plant sciences and mathematical model building, not English and German.

The analogy is that a model-builder is translating information from a biological system, which comes to him/her in biological language into a mathematical/statistical language so that it can be presented as a mathematical system—the model. The analogy is not arbitrary. It is quite evident that different disciplines do develop not only their own technical vocabulary but also unique frameworks of thought and familiarity with the language is the key to these. One of the sources of difficulty in communication is these frameworks: they are not innately learned or intuitively obvious. A mathematician does not automatically know how a geneticist sees the world even though there are areas where their conceptualizations may be very similar. The key to mutual understanding, I believe, is in becoming familiar with each other's language. It is not necessary to become an expert in the other's field in order to communicate—it is however necessary to go beyond the assumption that there is a common language in standard English which, in technical (scientific and economic) work is simply no longer totally true.

Every field develops its own terminology which is unique to it. From the names of subatomic particles in physics to the long and hopelessly complex names of organic compounds in biochemistry, it is fairly easy to understand the need for specific names and ways to talk about newly discovered things. Despite the fact that these names may be cumbersome and totally obscure to the uninitiated, they are not the main problem. The main problems in communicating arise, I would say, in more subtle ways--in the areas where common standard English is used by both parties but not to refer to quite the same thing, or where what is said sounds like standard English to the listener, but still doesn't make sense to him/her. This is particularly a problem for scientists (or laymen) listening to social scientists speak, where words which sound fairly standard turn out to have very specific meanings. (E.g., to the uninitiated the phrase "structure-conduct-performance paradigm" is totally gobbledygook despite the fact that the words are deceptively familiar. The result for the listener is frustration.) Computer jargon can have similar problems.

The point here is not that the problem is enormous or unsurmountable. It is neither. In fact it is fairly easy to learn enough of a language of a different discipline by working repeatedly with members of that discipline and listening well. It helps considerably and may be imperative to have some idea of the nature of research in the field but this will come with the contacts and interactions. It is not necessary to be a practitioner. In this again the translator/interpreter is a good analogy in that the interpreter has to be able to judge a speaker's meaning well enough to be able to choose between two possible translations for a word which have similar but not identical meanings, and this type of subtle judgement is often acquired in the field while working. The more a person in a given field works with researchers in another field and becomes familiar with the language and processes, the more that individual becomes a bridge between fields—a bridge for other scientists/researchers as well as him/her self. A model becomes, in this analogy, like a library or set of reference works in a commonly understood language. The modeler has interpreted the various languages so that the various users—including him/her self—can use it to gain access to the needed information. It can be set up to recombine information, test hypotheses, try alternate scenarios, etc.

While an individual may become something of a "bridge-scientist" through years of experience, there may also be ways of accelerating the process, which the following will address.

VI. C. Training for Multidisciplinary Research

Given that multidisciplinary teams do experience problems in getting to an operational phase, some preparation in understanding other disciplines would seem to be useful. Dillon suggests that education of professionals should be primarily multidisciplinary including perhaps two years in ones chosen field followed by multi-disciplinary problem solving, just as he suggests that research priorities should be set with reference to problem solving (see Section V.C). Again I disagree with him, not fully, but for the most part. I think that a thorough grounding in one field, in one specialty, is a good prerequisite for professional work. It is also true though, that encouraging specialists to broaden their scope into areas outside their specialty is crucial if "bridge scientists" capable of guiding or leading problem-oriented multidisciplinary research are to be found.

The first question is one of training for multidisciplinary team work, particularly on the graduate student level (since undergraduate education tends to be broader anyway). I would not advocate any special courses in how to work in teams or on task forces. Nor would I advocate changing the basic requirements for a Ph.D. radically. A requirement of a field outside one's own discipline is a good one and should provide a measure of understanding of at least one other field.

I am going to suggest one institutional change as described herein: I would suggest that throughout the departments which find themselves increasingly involved in multidisciplinary teams, that short "Interdisciplinary Lecture Courses" (ILC's) be offered, probably one per department for one term per year by each department.

The ILC's would be taught in five two-hour meetings, or ten one-hour, or eight lectures and four field trips, something along such lines as determined by the department offering the course. The subject would be "What is Agricultural Economics?" or "What Crop and Soil Scientists Do", etc. and would cover topics such as the areas of specialization within the department (for instance in agricultural economics, production, marketing, policy, etc.), the tools of the trade, (e.g., supply and demand analysis, modeling, industrial organization framework), who hires the experts in the field and to do what jobs, etc.

The classes might be taught by various individuals at department discretion but along some guidelines. The guidelines would be that the instructors should be chosen, one or more, from each of two groups:

1. The department chairman and emeritus professors or possibly visiting professors or faculty members who are engaged in 100 percent research or extension work. The idea here is that these people, particularly the first two groups, have good experience of the field and probably an overview of the whole field, but generally are not actively engaged in their own research but more as administrators and advisors. The latter two groups, visitors and active non-teaching faculty, are included because they would offer different perspectives.
2. The second group of potential teachers for the ILC's would be the most recently hired faculty members and Ph.D. graduate students who are somewhere in their dissertation work, but beyond classes and prelims. These people would offer the balance of newness, the experience both from other institutions and the home institution of

the state of current research and the most exciting or controversial topics, and perhaps fresh enthusiasm.

The advantages of this mixture are several:

1. The different experiences and viewpoints of the speakers should enrich the class for those attending and give a sense of the breadth of the discipline.
2. New faculty members get a chance to hear what their chairman or almost retired predecessors have to say and also to "try their wings" in the department on a relatively simple teaching assignment, but one which allows them to talk about their own experiences to a varied audience.
3. Practice for graduate students doing same.
4. By concentrating on these people, the teaching loads of those individuals who are probably feeling themselves to be the busiest, those in mid-career, are not increased. (Although individuals from this group need not be excluded from lecturing in such a class).

The ILC's would be offered with attendance by the following groups in mind:

1. Graduate students from other departments. The various departments might require that a graduate student attend one or two ILC's.
2. Incoming graduate students in the same department who do not have a strong background in that department.
3. Faculty members from other departments who for one reason or another have an interest.

The objectives of these courses would be, not to provide an in-depth analysis of the field in such a short time, nor to provide a boring list of disciplinary divisions or past achievements of the type sometimes given to visiting dignitaries.

The objective as I see it would be more of an advertisement of services available, a chance to affect the opinions of those in other fields, to explain how one—or one's profession—sees the world and tries to change it and what kinds of skills or means are generally used in that attempt. The tone would be better if it were interesting to someone _not_ in the field: discussion of ideas, anecdotal examples of what is done rather than a long list of research topics or history.

Something rather like this has been requested in the past by graduate students in agricultural economics from other technical agricultural science departments. The request was just that one course in technical agriculture be offered as an overview for graduate students in agricultural economics—or other fields—who feel the need to know something about technical agriculture. The response has been surprisingly vehement in its rejection from the technical sciences. The main objection seems to have been that asking for such a class, the agricultural economics graduate students were implying that one can summarize the high points, the main ideas of a technical discipline and then forget the rest, that the goal in asking for an overview or succinct description of what experts in those fields _do,_ is to learn the "answers" in order to be able to do without the fundamental research or the researchers.

It seems to me that this also is a communications problem and certainty points up the need for increased understanding among disciplines.

The goal of taking such a class, were it available, would be not to be able to do without the technical scientists, but to be better prepared to work with them. There is perhaps the point of view that "a little knowledge is a dangerous thing." One would simply have to assume that graduate students become sophisticated enough as their professional training proceeds not to assume that a series of brief introductory lectures told them everything they need to know about a field. This is not a difficult assumption, and yet the resistance to "summary" courses has remained. The feeling seems to be that if one wants to know about a field, one should go back and take all of the necessary courses to really know about it. There is merit to this approach under some circumstances—particularly if the goal is individuals thoroughly trained in two or more fields. Such individuals offer excellent prospects as "bridge" scientists as mentioned above. But such an all-or-nothing approach will do little to alleviate the problems of communication between disciplines, in times when the need for and interest in multidisciplinary research is growing.

Another way to look at the proposed ILC format is as a way to do some language teaching and learning—to allow some familiarity with technical language, research interests and therefore indirectly, at least a little of the thought processes of other disciplines. All of this could be regarded as "language" training against the time when the professional finds him/her self away from the sheltering home of his/her own department or discipline and, instead, living in the midst of a "polyglot" team speaking everything from biochemistry to fertilizers and s/he is trying to explain the concept of "opportunity cost" to those who only "speak" entomology.

VI. D. Making the Team Functional - Tools

Despite a certain mistrust in some quarters among plant scientists of both economics and mathematics as tools which may run rough shod over the subleties of biology, there is a certain and sometimes high degree of interest in modeling, in many areas, as this entire paper has tried to show.

> "The great complexity which life scientists encounter calls for more and more sophisticated mathematical methods. This tendency will cause many difficulties. A biologist, who has usually to spend even his spare time in the laboratory, might be reluctant to concentrate on mathematics which cannot be acquired quickly and which often involves time-consuming computations. However, we have good reason to assume that in the near future there will be more and more biologically oriented mathematicians willing to cooperate.
>
> ". . . when a scientist states a theory precisely and is interested in knowing just exactly what his theory involves, he is practicing mathematics.
>
> "The more precisely a scientist is working, the easier it is for him to make the proper decision at each step of his research. Therefore, mathematics may also be conceived as a tool for decision finding. Even in desperate situations of scientific research, when the investigator fails to find a good result, a careful mathematical study of his problem could be the stem for an inspiration of future work. If this were to happen, mathematics would have contributed indirectly to break mental barriers and, while in itself not creative, would have been a major tool in stimulating creativity."[61]

In general, I think it will be found that plant scientists are more interested in working jointly on models than in work pertaining to prices or to economic theory or almost anything else the agricultural economist has to offer.

However, the modeling aproach of the people inolved may diverge. For the most part, a biological scientist is generaly going to see any mathematical, statistical tool as useful it helps further his/her

---

[61]Batschelet, Op. Cit., p. 24.

reseach in understanding the process s/he is investigating. Black-box or associative models are of use in the more applied sciences and in management, but the more basic a scientist's work, the more likely s/he is to be interested primarily in structural models. For example, a black-box model which predicts winter-kill as a function of growing degree days, temperature, hard freezes, snow cover and soil moisture with fair accuracy would probably be of interest to a basic scientist in the sense that any statistical tool which shows correlations, unusual deviations, etc., is, but it would not be of very great interest as it would have little to do with discovering or explaining how the mechanisms worked. To a plant breeder it could be of some use if the statistical workings of the model could point out which of various factors were most significant in causing—or in avoiding—winterkill and relating this to plant characteristics to breed for, assuming that this information wasn't already known. This type of model would probably be of most use, however, to a very applied scientist, say to an agronomist working on improving a "production package" for winter wheat, or working to analyze the potential for growing winter wheat in a cold region which hadn't grown it previously, or working in extension helping local farmers to grow the crop. A farm manager would also be very interested in such an associative sort of model if one could be developed with good predictive capabilities.

The point is, then, that the more applied scientists in biology or in economics may have good use for an associative or black-box type models. Basic researchers are likely to be less interested. A correlation between two measurements may in fact be an artifact. The basic researcher is only really interested in a correlation if it can be used to help discover the inner workings of a process, and so, in most cases, if it is to aid in basic research, a model must be a structural one which models what is known or hypothesized about a system to date, and allows the theory to be tested against the behavior of the real biological system.

The economists' traditional approach of working from that fairly simple black-box, the production function, represents a very different point of view from the basic scientists'. The production function represents looking for the best way to combine inputs within the confines of the current technology and given prices. One can make some changes, adjust prices, change input mixes somewhat and then examine more closely the relevant parts. The starting point is the economic as much as the physical workings of the system. This is a major difference in approach between the economist and the scientist, and is indeed a source of mistrust on the side of the scientist who is apt to feel that the economist has no understanding of the actual workings of his system.

The difference in approach is understandable. The farm management production economist historically has been working with making production systems function in the present with some projections about the future, but based on extrapolations from current situations. The basic scientist on the other hand, is often really not dealing with such practical questions. It has been shown that most real break throughs in technology come as a result of basic science and only after an interval of some years, if at all.

The applied technical scientist is somewhere in-between the production economist and the basic scientist, making changes and adjustments in current production packages, sometimes helping to introduce new ones as they are developed. His/her approach is still likely more nearly that of the basic scientist: to find out how a system works, and then maximize production, rather than the economist's view of finding what is significant and researching that in order to maximize.

Here again, then, the argument is that the type of tool chosen for a research project is a function of the type of work to be done. My contention is that there are some areas where the economist will now find it of interest to be behave more like the biological scientist in approach and to take a structural rather than an associative (black-box) approach. I think that this is particularly true where an effort is being made to make large changes, to really change technologies, rather than to make adjustments in production packages. An excellent example of this is in development work. There is a perceived need to actually understand the innerworkings of very complex biological/economic/social systems with many complex interactions. Black-box associations do not carry enough information. The development economist is in the position somewhat like that of the basic scientist: s/he is finding it necessary to try to unravel complex workings of the parts of a system in order to try to find what changes will be effective in reorganizing the workings of the whole.

This is also true in areas other than development. Probably for more interdisciplinary work, whether it be on projects such as developing ruminant nutrition models and exploring new feeds, or in resource management, or in projecting world food needs and how to meet them, the need is more for structural than for black-box approaches. This is for two fundamental reasons:

1. In this type of analysis, it is not sufficient to optimize the parts of a system. Indeed, the danger of optimizing parts leading to suboptimizing the whole system is now widely understood. The simpler tools such as production functions are not sufficiently complex to handle a multi-stage system with many interactions. An LP is a simultaneous run of a number of production functions, more nearly maximizing the whole of the farm, for instance. Some types of analyses, especially those involving actually looking at biological systems directly (not just looking at what goes into the field and out, but "growing" the plants by computer under different conditions or management schemes) will require more structurally evolved analyses such as simulation studies.
2. To the extent that models are to act as holders and "translators" of information for users from various disciplines, they will have to deal with the actual processes.

VI. E. Making the Team Functional: Psychological Elements.

The literature provides a few basic, global, statements of guidelines for multidisciplinary work. The following summarizes some points which will be examined further.

> "The organizational pattern wherein the interdisciplinary research efforts can function in Iowa or other state agricultural experiment stations probably will be a mix of the kinds of research efforts that now exist. The question of structure might be posed in another way, such as: On what basis should an experiment station select a limited number of areas for research emphasis and encourage a complement of compatible, competent staff members to dedicate a certain amount of their research talent in working through the team approach to provide answers to a certain set of problems? In my opinion, this must be done in such a way as to permit individual team members access to their scientific discipline and to provide for interaction with their departmental colleagues, without sacrificing valuable time in an excessive number of meetings or commuting between research facilities. It should preserve the intensity of departments and the access to administrative talent that I believe exists in the department leadership in the Colleges of Agriculture. It must identify polydisciplinary, "bridge" scientists to provide the integration necessary to interpret the contributions of the individual scientists. To be successful, the system must not create animosities between scientists or departments or colleges. It must provide adequate resources commensurate with the challenge at hand."[62]

A number of points touched upon above will be expanded. The psychological questions, seeming obvious are extremely important. Members of various disciplines, without intending to do so, at times, usually because of lack of understanding or careless speech, actively discourage team work or even insult each others' fields. In the long section cited from Dillon, under Section V.C., Dillon makes two such blunders in my opinion. One is his ranking of research priorities as discussed under Section V.C. The simplest blunder is to say "in the trivial sense of being able to say to a plant physiologist that we need

---

[62] Mahlstede, Op. Cit., p. 632.

an estimate of such and such a coefficient, but in the far more <u>significant</u> sense of recognizing disadvantaged regions or groups . . .," [Emphasis mine]. In one sense any single number is trivial in the larger picture, but it is important not to categorize some kinds of research as trivial and others significant, particularly if one hopes to have further productive exchange with the practitioners in the "trivial" area. All scientifically conducted research involves examination of small details out of which to build larger understanding. A price for corn on a futures market of a given day is trivial to most of the world, but not to certain research. The point is that it is easy when dealing with disciplines other than ones own to make sweeping and even derisive judgements. It is simple but also important to avoid generalizations such as "Economists don't understand any values other than price" or "agronomists don't understand anything except maximizing yields" or "soil scientists just dig holes in the ground". These things are obvious and easy to avoid, but the kind of thoughts which they express are not as rare as one might imagine. There is probably some truth to all of them: all disciplines are limited, and a given expert probably finds his to be the most interesting or important. Thoughtlessly and usually ignorant attitudes about other disciplines must not survive long in an multidisciplinary framework.

There is also some problem between basic science and applied science when the former is characterized as aimlessly playing around or pursuing unusable avenues and the latter as "stomping out brush fires" implying little coherent thought about longer range or more important problems. Mahlstede says:

> "The importance of basic research to the development of new breakthroughs was well documented in a 1968 National Science Foundation report, "Technology in Retrospect and Critical Events in Sciences", (see Thompson, 1969). In the study, five technological innovations were traced back to the series of experiments that led to a particular development. Of the 340 or more key events documented, the study found that 70 percent were nonmission or basic research, 20 percent were mission oriented, and 10 percent were involved with development and/or application. Ninety percent of the basic research that led to a breakthrough was completed 10 years before the introduction of the technology that stemmed from it. One of the conclusions that can be drawn from this study is that undirected research with the primary goal of uncovering new information is the basis for understanding the applied research that leads to the solution of a relevant problem or the development of a new innovation."[63]

And in the same article Mahlstede also reports a growing trend for applied, multidisciplinary research: "The trend is partly a response to the nature of the problems we face and the recognition by the staff that multidisciplinary research is a bona fide approach to the solution of these kinds of problems."[64]

It seems clear that both basic and applied research are needed and that, in fact, dwelling on the distinction to attempt to say that one is more important than the other is not useful. Dillon's suggestions on ordering all research to be problem solving have been criticized to some length under Section V.C (with the proviso that organizing some applied research along the lines he suggests makes sense).

My suggestion would be that those interested in multidisciplinary work, which is clearly very important not waste time in contrasting the relative merits of basic and applied research, but try to make the most of both.

---

[63] Ibid., p. 631.

[64] Ibid., p. 632.

VI. F. Making the Team Functional—Professional Reinforcement

The area or school of institutional economic theory, which looks at many aspects of political/economic systems dwells, in part, on the functioning of individuals in various systems. One of the school's tenets is that to change a system, it is not sufficient to understand it or to understand peoples' motivations; it is necessary also to redesign the system so that people are motivated or "reinforced" in certain behavior.

This is clearly true in an academic context. Scientists and researchers are very like "entrepreneurs" in economic theory in that they have goals or "utility functions" which they attempt to reach or to "maximize". Generally, within an academic context the rewards are not strictly profit-maximization but a mixture of economic motivation (salary optimization or security) and the means to make research progress. The two are neither mutually exclusive nor a tradeoff—except occasionally. Generally a scientist or other professional within a university or research establishment can maximize salary and security and also his/her access to needed support for research, be it in terms of funds for equipment, salaries for graduate students and technicians or autonomous decision making, through achievement which is recognized by his/her peers (and department chairman, dean, research director, etc.). The recognition takes the form of published articles in journals, research grants, and other forms of professional recognition—honors, widespread unofficial recognition of research, success on a project, excellence at teaching, etc.

One major problem for multidisciplinary work is that most of the existing reward systems (although not all) are built upon the lines of disciplinary research. This is a particular problem for young professionals who are trying to attain tenure in a university or similar security in another institution. The reward system is often such that disciplinary expertise is more highly rewarded than good cooperative research in a multidisciplinary project. This situation is often cited, for instance in international development work. Development theory involving micro-studies in a given area is a topic of concern in developing economics. For a young agricultural economist to spend two years "in the field" working on a cotton project or a "rural small scale industry" project certainly does not guarantee professional acceptance. It may help to make him/her specialize more than s/he might choose as a "development economist", but it is professionally somewhat acceptable and perhaps very much so. In contrast an agronomist going overseas to a developing country for two years, even to work on an "underexploited" crop will likely find that his/her daily research activities will be the most mundane of soil testing and variety trials. His/her work may be crucial to a project, it may be very important to the region where he works, but when s/he returns to the United States, it most likely will not be seen as important to the discipline, because it will generally have fairly straightforward, maybe even elementary, research. As the focus of researchers in agronomy becomes more international in nature—which I see it doing—this may be less of a problem, but it will not disappear.

There is a similar problem in other disciplines. When an economist or an agronomist joins a multidisciplinary team it is often to use skills which are already well established within his/her field. It may be that his/her research will be such that it approaches new areas of disciplinary concern, but it is also likely that it will not. A professional person is then faced with the choice of doing disciplinary research, with its reward system well established, or multidisciplinary work with questionable reward system.

The point of this argument is simply that if multidisciplinary research is considered to be important, and worthy of the best research efforts that can be mustered, then we will have to carefully rethink the rewards system to encourage professionals—young and otherwise—to participate in multidisciplinary research.

A few specific suggestions follow:

1. Organize a section of the national annual meetings to discuss multidisciplinary work.
2. Publish the papers from the meetings and see what interest there is.
3. Encourage the professional journal(s) to include work of this nature (or start a new journal).

In a paper delivered at the 1981 American Agricultural Economics Association Annual Meetings, Tim Hammonds criticizes the profession generally land the Journal particularly for not being relevant to the needs of the business community but instead carrying articles advancing tools of analysis rather than their substantive application.

"To be sure, there is academic work which is timely and relevant for the business community. Several universities have a small cadre of agricultural economists who are respected and appreciated by those in food distribution. Some of these are found in research and teaching, some in extension. But the truth is that we have not found a constructive, highly visible outlet for their work.

"Published research remains the glamour child of our profession, so much that we reward it without regard to content relevance. Extension economists, along with those who work closely with business in research or policy, have not found an equal set at the academic table. This does not leave us well-positioned to deal with the future."[65]

Hammonds is cited here, not because I particularly agree with his argument that business economists need an "equal seat at the economic table". Business and academic economists should surely offer each other professional respect but separation of their interests somewhat may not be a bad thing at all. The business community is well equiped to take care of its own interests and to put to work both the research and the researchers coming out of universities. The university needs to have enough separation from the business community and applied research generally to allow basic and theoretical research to continue unfettered.

I am interested, though, in the importance he places on the journal and on finding a "constructive, highly visible outlet."

Perhaps there is a need for a "Journal of Diversified Agricultural Economics" to provide a means of publishing the work of academic professionals who are engaged in applied research relevant to the business community, doing extension work, or working at home

---

[65] Hammonds, Tim. "Agricultural Economists: Our Quantitative Historians", paper presented at Annual Meeting of American Agricultural Economics Association, July 28, 1981.

or overseas on multidisciplinary teams, and thereby provide access to the reward system of the university for them.

4. Encourage university personnel who are requesting sabbatical leave to branch out and broaden their academic competence by spending their sabbatical time working in another field.

5. Re-examine the university organization, remembering that disciplinary, sub-system optimalization may fail to optimize the functioning of the whole system. Once funding entities are established, they tend to perpetuate themselves and resist changes. Michigan State University College of Agriculture and Natural Resources has a relatively small number of departments in the natural sciences with a large faculty in each, under the direction of the Department Chairman. The Universities of Wisconsin, Madison and California, Davis, by contrast have a large number of small departments under the control of the Dean who can create new departments or phase out existing ones. One example of how the administrative structure can be stultifying is that at MSU, starting new groups and new interactions can be very difficult. MSU has not been able to form a department of genetics or of molecular biology, despite faculty attempts to do so. The "genetics group" ran largely on the initiative of the individuals and very small funding. A few graduate students were accepted into the genetics program but they were shifted into various departments where the genetics group lost all funding.

In summary, some institutional arrangements can be counterproductive. It was argued previously that to require all research to be "goal seeking" will certainly damage individual or department's ability to carry on disciplinary basic research. On the other hand, rigid departmental boundaries and a system which rewards mainly disciplinary basic research will inhibit forming multi- or cross-disciplinary research efforts or even of redefining the boundaries of disciplines as fields grow and change. A balance is necessary. The institutional arrangement of the university system must be such that both basic and applied, disciplinary and interdisciplinary research can be rewarded if all are considered necessary, as this paper contends.

VII. Conclusion.

In developing the material for this paper, and in the course of organizing a framework to deal with a multiplicity of ideas about modeling and the systems approach, about the relationship between the research interests of plant scientists working largely in the realm of biological understanding and those of agricultural economists who work more often with social systems, the theme of commonality of interest and thought processes has grown for me. Both the natural scientists and the social scientists are concerned with the interaction of the needs of society and the management of systems--biological and social--to improve human well being. Both groups of professionals are also very much concerned with the scientific approach, the need to expand quantifiably testable hypotheses and with developing tools, methodologies to do accurate, testable, repeatable and applicable research. Both groups are concerned with synthesizing from their reductionist work, larger hypotheses, theories which explain the larger workings of systems. Both groups are confronted with the problems of balancing specific basic research, subdivided by disciplines and the need to reintegrate and recombine new and existing knowledge. Both are concerned with the dynamics of systems and the changes which take place, both within the areas they study and in the world at large. Some of the most interesting and complex problems in both disciplines arise because the world is changing: extent of knowledge, technologies, social systems, people's needs. In fact if this were not the case, (although it is rather difficult to imagine that), it would be far easier to subdivide and categorize and characterize the elements of the whole.

One of the arguments then, of this paper, is that changes as listed above, are bringing the concerns of the biological and physical sciences closer together. Both increasing knowledge and growing social and political interactions are part of a process of expanding our collective awareness and increasing our interactions. This is taking place both physically, in increased communication, transportation, trade, etc. among the members of an—also expanding—world population and intellectually in what one philosopher, Teilhard de Chardin,[66] described as a "noosphere" an intangible network or web of informational connections which he envisioned as an (intangible) sort of sphere surrounding the earth. Whether one accepts his presentation of the concept or not, the tenets seem to be correct: that human knowledge, interconnections and understanding of our pressing needs are growing—as are our tools and concepts to work with them.

In this framework, modeling of systems becomes a multi-purpose tool. It is a technology developing to enable us to handle greater quantities of information simultaneously, and it is, as such, a research tool to work with and test the implications of interactions, and it is also, to some extent, a framework which can help in communicating among researchers, between disciplines, and to test what we know both on the very micro level about minutely small movements of matter and about the very large implications of social decisions.

Forrester[67] dwells on the "counterintuitive" behavior of social systems and makes the claim that the human mind, when holding a lot of

---

[66]Teilhard de Chardin, Pierre. Building the Earth. Dimension Books, Wilkes-Barre, PA. 1965.

[67]Forrester, Op. Cit.

information and assumptions simultaneously will often draw conclusions about complex systems which are not consistent with all the assumptions and data. He makes the claim that computer technology, properly used, can both handle all these variables and allow us to see the implications of our intuitive reactions in solving problems, which are often in conflict with the real results of social policies. He says that, given the nature of the way the human mind works, it is not surprising that our policy prescriptions get results which are the opposite of what was intended. Personally, I see fit to challenge some of his computer-generated model results of world projections as being too global, too simplified (of necessity) to be reliable for predictions (of quality of life, population growth, pollution, etc.) at this time, but I am willing to concede that such failings may be largely resulting from the limitations of the current state of the art of computer simulation in dealing with systems of the magnitude which he is modeling--and also the size of the assumptions he, of necessity again, makes.

How does this relate to the questions about modeling and scientific interactions posed in this paper? In practical terms, I see an historical hierarchy or progression of the use of modeling in agricultural sciences. The sequence is somewhat as follows:

1. In developing countries today, and historically in this and other economically developed countries, the most basic and useful application of computer modeling in agriculture can be as a means of handling information. Indeed this stage is not so much one of modeling as of data handling, and yet in developing nations, where technology changes happening in one person's lifetime can literally span what

took hundreds of years historically in Europe and the United States, the two are not far apart.

The most gainful initial use, then, of computer technology applied to agriculture in developing countries can be that of handling large amounts of data, on crops grown, on information about size of farms, names of farmers who wish to be cooperators with extension services: what they grow, how much, where they market their goods and how. This obviously does not happen overnight, but such information can be a major aid to an experiment station worker, a researcher, etc. Inventories of biological information: crop biology, insect pests, disease, can, when available, be both a reference library and, when combined with information about farmers, a powerful tool for agricultural development work. FAO is currently working on developing a common farm-survey skeletal framework with the aim of making micro-data on farming (and crop biology) gathered in one region of the world more accessible and possibly applicable in other areas.

2. Rapidly behind, or possibly concomitant with the data-handling systems envisioned above can come the farm budgeting and management models which have been developed in the U.S. and elsewhere. Farming systems researchers are now working to develop models of farms in a wide variety of ecological niches and social conditions. It is at this level that the bio-economic variety of modeling becomes relevant. As this paper has reported, this work is really just beginning. In parts of this country farmers are actively using economic farm management models as aids to business management and decision-making. In these models the biological variables are generally fed in highly aggregated terms ("management level" or yield-

per-acre rather than biological crop inputs). But there also exist biologically based management models such as Integrated Pest Management, and models which relate directly to physical science aspects of farm management such as those for irrigation management. The bio-economic models of whole farms are still in their infancy (see for example Parsch, L.)[68] and the problems to be surmounted considerable but much work is going on both from the site of models of agricultural crops and the more economic farm management models to integrate the two.

3. The area of using models for biological research is not necessarily dependent upon those areas listed as numbers (1) and (2) above., It can easily develop separately and simultaneously. However, in that the command of the tools of basic research and the amount of laboratory and field data necessary is generally only available under conditions of a highly developed research community, it is considered as a part of a more complex and highly developed agriculture and agricultural research complex.

Modeling of complex and minute biological processes—taken as a complement to field and laboratory work is of fairly recent development in the plant sciences. This paper has dwelled at some length on this use of modeling. The suggestion is that this type of research is increasingly being recognized as a most valuable tool in highly sophisticated areas of basic research.

---

[68]Parsch, Lucas D. "Economic Analysis of Dairy Farm Designs with Emphasis on Alternative Forage Systems". Type III Seminar, Department of Agricultural Economics, Michigan State University, April 8, 1981.

4. Finally, using agricultural models as tools of policy is, again not a strictly chronological sequence. It is quite likely that as soon as data gathering systems about farming in a given area exist, the information can be used by policy makers. That is, of course, one of the major goals of such an undertaking. However the explicit use of modeling to forecast the outcomes of various policy prescriptions under various "states of the world" or scenarios of future events is not only a complex undertaking, but involves, of necessity the development--or the pre-existence--of models like the MSU Agriculture Model of national and/or world, or regional, agricultural production, markets, imports and exports, etc.

This area of modeling use is also considered last in the sequence because it offers the possibility of becoming an information-generating tool for researchers and policy makers who are grappling with some very large--macro and long range--problems about the nature of whole agricultural production and distribution systems, issues such as the "World Food Problem" and choices of technology. In this context, policy issues which can be examined through the use of modeling range from the very specific (e.g., Christensen and Mitchell's analysis of a hypothetical farmer-held soybean reserve V.C above) to the broader issues (e.g., Dillon's suggestion on ordering research priorities, or Forrester's analysis of the world's support systems, including agriculture.) My contention is that at the present time, the specific projections of the results of particular policies, projected within the framework of well-tested and reasonably well-behaved models are probably more reliable and more useful than the more global models.

However, I expect that in the future, the more global will continue to be built because the need to analyze global problems exists. The reports of the World Food Council and of the Club of Rome offer different projections of the future and different attempts at policy prescriptions, and yet they are not totally divergent. Questions have been raised in many quarters about both the future of our society and the best course for the agricultural production system as part of it. It seems clear that modeling by agricultural specialists, biological and economic will have a large role in these projections.

The above sequence of the uses of modeling in agriculture, as I see it, is intended as a summary of the information in this paper. There are a large number of areas pointed out in which those concerned with questions of plant biology and of agricultural economics share a common ground. The hypothesis is that there are expanding areas where interests are mutual and multidisciplinary work can usefully occur. It is hoped that this paper will be of some use to those engaged in these projects. The hypothesis is still being tested.

BIBLIOGRAPHY

Adams, M.W. and J.V. Wiersma. "An Adaptation of Principal Components Analysis to an Assessment of Genetic Distance." Michigan State University, Agricultural Experiment Station, Research Report No. 347. March 1978.

Anderson, Louise E. "Metabolic Regulation of the Reductive Pentose Phosphate Cycle." What's New In Plant Physiology, Vol. 10, No. 10, October 1979, pp. 37-40.

Baschelet, Edward. "The Application of Mathematics to Biological Problems." BioScience, January 1966, pp. 22-24.

Bader, F.G. "A Structural Model of Photosynthesis in Anacystic Nidulans." (Communications to the Editor) Biotechnology and Bioengineering, Vol. XX (1978), pp. 119-125.

Black, J. Roy, John Waller and Roger Brook. "The Development of Software for Computer Assisted Agricultural Decision-Making." Michigan State University, Agricultural Economics Staff Paper No. 81-37. Revised June 9, 1981.

Bradfield, Stillman. "Appropriate Methodology for an Appropriate Technology", paper delivered at Annual Meeting of American Society of Agronomy, 1978, or mimeo, Kalamazoo College, Kalamazoo, MI.

Charlton, P.J. and Street, P.R. "The Practical Application of Bio-Economic Models", in Dalton, G.E., ed., Study of Agricultural Systems. Applied Science Publishers, Ltd., London, 1975.

Christensen, Thomas, and Donald O. Mitchell. "An Analysis of a Farmer-Held Soybean Reserve", mimeo, Department of Agricultural Economics, Michigan State University, August 29, 1980.

Cock, James H. "Biologists and Economists in Bongoland", in Valdes, Scobie and Dillon, Economics and the Design of Small Farmer Technology. Iowa State University Press, Ames, Iowa. 1979.

Crawford, Eric W. "Farming Systems Research and Agricultural Economics." Michigan State University Working Paper No. 1., June 1981.

Crawford, Eric W. and Robert A. Milligan. "Stochastic Farm System Simulation Model with Emphasis on Experimentation." Michigan State University, Agricultural Economics Staff Paper No. 80-81.

Crawford, Eric W. "Understanding, Quantification, and Modeling in Farming Systems Research: Results of a Simulation Study in Northern Nigeria." Michigan State University, Agricultural Economics Staff Paper No. 80-19.

Cunningham, Gary L. and James F. Reynolds. "A Simulation Model of Primary Production and Carbon Allocation in the Creosotebush." (Larreatridentata (DC) Cov.). Ecology, Vol. 59, No. 1, 1978., pp. 37-52.

Dale, Jacqueline G., Donald A. Holt and Robert M. Peart. "A Model of Alfalfa Harvest and Loss." ASAE Technical Paper No. 78-5030.

Dale, M.B. "Systems Analysis and Ecology." Ecology, Vol. 51, No. 1., Winter 1970, pp. 2-16.

Dalton, G.E., ed. Study of Agricultural Systems. Applied Science Publishers Ltd. London, 1975.

Debertin, et al. "Impacts on Farmers of a Computerized Managment Decision-Making Model." American Journal of Agricultural Economics, Vol. 63, No. 2, May 1981, pp. 270-274.

DeMichele, Don W. "An Evaluation of Modeling, Systems Analysis and Operations Research in Defining Agricultural Research Needs and Priorities in Pest Management". Iowa State Journal of Research, Vol. 49, No. 4, pt. 2., May 1975, pp. 597-621.

Dent, J.B. and J.R. Anderson. Systems Analysis in Agricultural Managment. John Wiley and Sons, Australasia Pyt. Ltd. Sydney, 1971.

Dent, J.B. and Blackie, M.J. Systems Simulation in Agriculture. Applied Science Pub., London, 1979.

deWit, C.T. and J. Boudriaan. Simulation of Ecological Processes. Centre for Agricultural Publishing and Documentation, Wageningen, Holland, 1974.

Dillon, John L. "The Economics of Systems Research". Agricultural Systems (1) (1976), Vol. 1, No. 1, pp. 5-22.

Erikson, Ralph O. "Modeling of Plant Growth". Annual Review of Plant Physiology, 1976. 27:407-34.

Evans, L.T., ed. Crop Physiology, Some Case Histories, Cambridge University Press, Cambridge, 1975.

Fick, Gary W. "The Mechanism of Alfalfa Regrowth: A Computer Simulation Approach". SEARCH Agriculture Agronomy 7, Lot 7, No. 4., 1977., pp. 1-27.

Fick, Gary G., R.S. Loomis and W.A. Williams. "Sugar Beet". Chapter 9 in Crop Physiology. L.T. Evans, ed. Cambridge University Press, 1975.

Forrester, Jay W. "Counterintuitive Behavior of Social Systems". Technology Review. January 1971, pp. 52-68.

Gates, David M. and Rudolf B. Schmerl, eds. Perspectives of Biophysical Ecology. Springer-Vertag, New York, 1975.

Gypmantasiri, Purek, et al. An Interdisciplinary Perspective of Cropping Systems in the Chiang Mai Valley: Key Questions for Research, Faculty of Agriculture, University of Chiang Mai, Thailand, June 1980.

Hammonds, Tim. "Agricultural Economists: Our Quantitative Historians", paper presented at annual meeting of American Agricultural Economics Association, July 28, 1981.

Harwood, Richard R. "Natural Weed Controls Are Looking Good," New Farm, Vol. 1, No. 6, September-October 1979,, pp. 56-58, and "Seasonality in Organic Weed Maintenance", New Farm, Vol. 2, No. 4, May-June, 1980, pp. 61-64.

Kac, Mark. "Some Mathematical Models in Science". Science, November 7, 1979, Vol. 166, No. 3906, pp. 695-699.

King, F.H., Farmers of Forty Centuries, Madison, Wisconsin, 1911., recently reprinted by Rodale Press.

Mahlstede, J.P. "Concepts, Organization, and Funding of Interdisciplinary Research". Iowa State Journal of Research. May 1975, Vol. 49, No. 4, pt. 2, pp. 629-633.

Mitchell, Donald O. and John W. Ross. "A Forecast of Grain and Soybean Exports in the Year 2000". Department of Agricultural Economics, Michigan State University, July 20, 1981.

Moorby, J. and Molton, J.L. "Potato" in Evans, L.J., Crop Physiology, Some Case Histories. Cambridge University Press, Cambridge. 1975, pp. 225-257.

Nelson, James L. "Michigan Oat and Barley Breeding". (CSS-408), March 11, 1981, Michigan State University, Department of Crop and Soil Sciences.

Paris, Tirso B., Jr., "Systems Analysis and Simulation of Upland Rice-Based Cropping Systems in the Phillippines." Ph.D. Thesis, Michigan State University, 1978.

Parsch, Lucas D. "Economic Analysis of Dairy Farm Designs with Emphasis on Alternative Forage Systems." Type III Seminar, Department of Agricultural Economics, Michigan State University, April 8, 1981.

Reynolds, James F. "Some Misconceptions of Mathematical Modeling." What's New in Plant Physiology, Vol. 10, No. 11, November 1979, pp. 41-44.

Spedding, C.R.W. An Introduction to Agricultural Systems. Applied Science Publishers, Ltd. London, 1979.

Stapper, Maarten and Gerald F. Arkin. "CORNF: A Dynamic Growth and Development Model for Maize. (Zea mays L.)" Texas Agricultural Experiment Station, Blackland Research Center, Temple, Texas, December 1980.

Teilhard de Chardin, Pierre, Building the Earth, Dimension Books, Wilkes-Barre, Pa., 1965.

Thesen, Arne. "Some Notes on Systems and Modeling". International Journal of Systems Science, 1974, Vol. 5, No. 2, pp. 145-152.

Thornley, J.H.M. Mathematical Models in Plant Physiology, A Quantitative Approach to Problems on Plant and Crop Physiology. Academic Press, London, 1976.

Van Schilfgaarde, Jan. "Earth and Water". The Antioch Review., Vol. 38, No. 4, Fall 1980, pp. 421-435.

TECHNICAL APPENDIX I

MICHIGAN OAT AND BARLEY BREEDING

CSS-408

James L. Nelson

3/11/81

I. INTRODUCTION

The role of the plant breeder is to take genetic variability and organize it into new genotypes which are superior in their phenotypic expression over time and over different environments. It matters not what the source of the variability is - wide hybridization, induced mutations, chloroplast and mitochondrial genomes, intraspecific variability or variability derived through molecular techniques - all of these sources of genetic variability must ultimately be organized into superior genotypes. Dr. Freed has presented the basic methods for such organization in this course. It is my purpose to present an holistic system which has evolved over the years and upon which the oat and barley projects in Michigan are based. It is by no means the only method for successful plant breeding, but it is a system that is highly successful and which rests upon the fundamental principles exposited in this course.

The system which the late John Grafius developed is a recurrent selection scheme coupled with novel parental selection. For a primary gene pool it utilizes a limited number of genotypes, all of which are adapted to the cool, mid-season environment of Michigan. Additions and deletions to this population are governed by performance over years and locations and the additions typically come from the cooperative nurseries of which Michigan is a participant. The intent of our programs is to consistently shift the mean overall performance of the populations to a higher level each year. From these populations are drawn progeny which are evaluated intensely and eventually released as new varieties transgressive for yield.

II. METHOD

A. Parental Selection

1. Vectors

   a. Each genotype is considered a vector of n dimensions, where $n = \#$ of traits being selected for. In general,

   $$A \cdot B = |A|\ |B| \cos\theta \text{ or } \cos\theta = \frac{A \cdot B}{|A|\ |B|}$$, where A and B are any two genotypes of n dimensions.

   b. Based upon additive gene action.

   c. Ideals are created.

      1. Optimal plants you wish to create.
      2. Vary for a number of traits.
      3. Must be realistic.

   d. Each line and ideal is assigned an overall worth score and they are preceived by the computer program in the following generalized format.

   U (overall worth) = (yield components) + (disease resistance) + (lodging resis.) + (various quality traits) + ...

   e. Traits are weighted according to their economic value and their heritability.

   $X_i = \frac{|\beta_i|}{\sigma_i} r_i \ \Delta X_i$, for a complex trait and $X_j = \frac{|\beta_i \ \beta_j|}{\sigma_j} \ \Delta X_j$, for a component of a complex trait.

   Where $X_i = i^{th}$ complex trait,

   $X_j = j^{th}$ component of the ith trait,

   $|\beta_i|$ = absolute value of the standard partial regression coefficient,

ΔX_i = observed value minus the mean,

r_i = correlation between seasons and/or locations for a trait (= heritability)

f. Computer "crosses" all entries and then backcrosses all hypothetical F_1's to both parents.

g. Computer output prints correlation coefficients which compare the degree of association of each hypothetical cross and backcross to each of the ideals. In other words, the output has now presented the breeder with a means to maximize his chance of recovering the ideal or created plant. Chosing the highest positive correlations among both the straight and backcross progeny will indicate which crosses should be made with which parents. (Refer to the attached sample output).

h. Advantages of vector method.

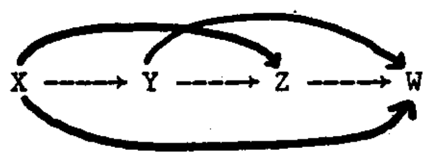
1. Can breed simultaneously for as many traits as desired.
2. Number of hybridizations is reduced.
3. Considerable time and money are saved.
4. Can evaluate fewer segregating generations.
5. Traits can be coded or modified by some measure of their economic value.

2. Outliers

a. Primary yield components:

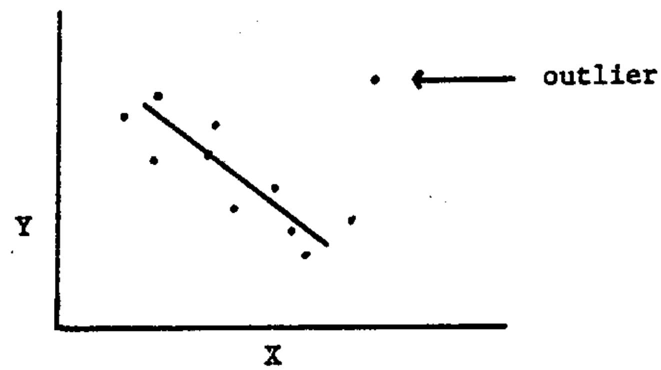
$X \cdot Y \cdot Z = W$, where X = # of spikes/unit area, Y = # of seeds/spike, Z = weight/seed and W = yield (by weight).

b. Developmental allometry



c. Negative correlations between the components are the primary impediments.

to yield improvement. As you breed for an increase in one component there will be a concomitant reduction in another component. What we seek, then, is a method to identify putative parents such that when crossed they will generate progeny which are superior to the parents in yield components and thus in yield. The method which we use is to regress the various yield components on one another.



d. Utilizing these outliers in the crossing program will increase yields as demonstrated by Menominee and Heritage outs and Bowers barley. Please observe the results of my experiment 8095 in which the unselected bulks outyielded all parents.

B. F_1 to F_2 Seed increase.

C. Single seed descent

1. Retains the range of genetic variability while advancing homozygosity.
2. Reduces variability within individuals but increases variability between eventual families.
3. Seed is bulked in F_4 generation.

D. F_4 bulk (approximately 10^3 seeds) planted in single plot for seed increase.

E. F_5 bulk planted in 4 replicated plots.

1. 25 spikes from each border row are randomly chosen in all 4 reps for a total of 200.

2. Preliminary yield trial of center two rows of each rep.

F. Head hills:

1. Heads (or spikes) from superior bulks are planted in a row with about 12" seperation.
2. Inferior hills (for quality factors - not yield) are culled as the season progresses.
3. At maturity only the fifty or so superior hills descended from any given cross are saved.
4. Based upon high degree of genetic differences between families, but not within.

G. Seed Increase.

1. 50 superior selections from selected crosses are increased in single plots.

H. Performance Trials.

1. Selections are evaluated for yield and agronomic performance in replicated plots over multiple locations.

I. Superior individual lines are submitted to uniform regional nurseries.

J. Eventually, yield selections which are stably superior in yield, lodging resistance, uniformity, disease resistance and quality are singled out for potential varietal release.

END OF LECTURE.

SELECTED REFERENCES: Grafius, J. E. 1952. A statistical model for estimating components of genetic variance in bulk yield tests of self-pollinating small grains. S. Dakot. Agric. Exp. Sta. Tech. Bull. #9, 13 pp.

---. 1956. Components of yield in oats: A geometrical interpretation. Agron. J. 48:419-423.

---. 1959. Heterosis in barley. Agron. J. 51:551-554.

---. 1960. Does overdominance exist for yield in Corn. Agron. J. 52:361.

---. 1963. Vecotr analysis applied to crop eugenics and genotype - environment interaction. In: Statistical Genetics and Plant Breeding, NAS/NRC Publication 982. pp. 197-217.

---. A geometry for plant breeding. Crop Sci. 4:241-246.

---. 1965. Short cuts in plant breeding. Crop Sci. 5:377.

--- and R. L. Thomas. 1971. Parental selection by computer. Res. Report 139, MSU Agric. Exp. Sta., 12 pp.

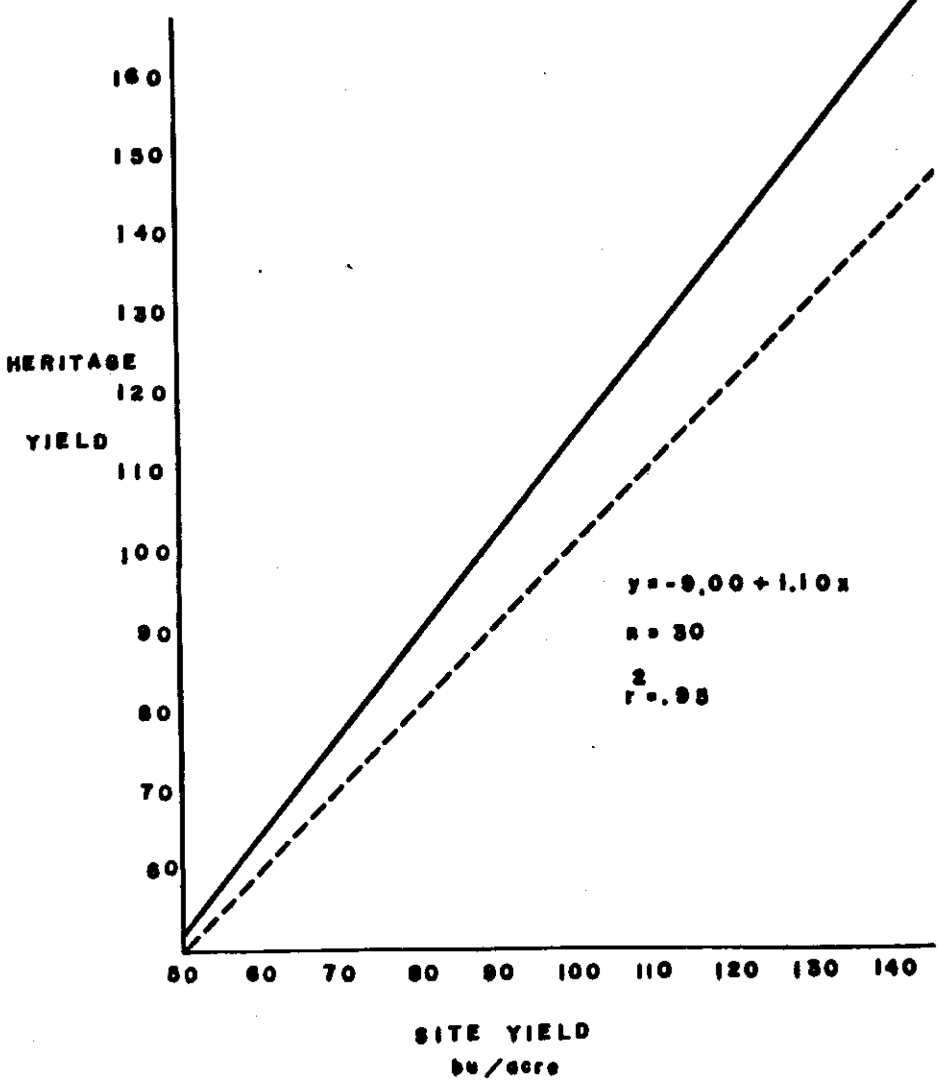
Thomas, R. L., J. E. Grafius and S. K. Hahn. 1971. Genetic analysis of correlated sequential characters. Heredity 26:177-188.

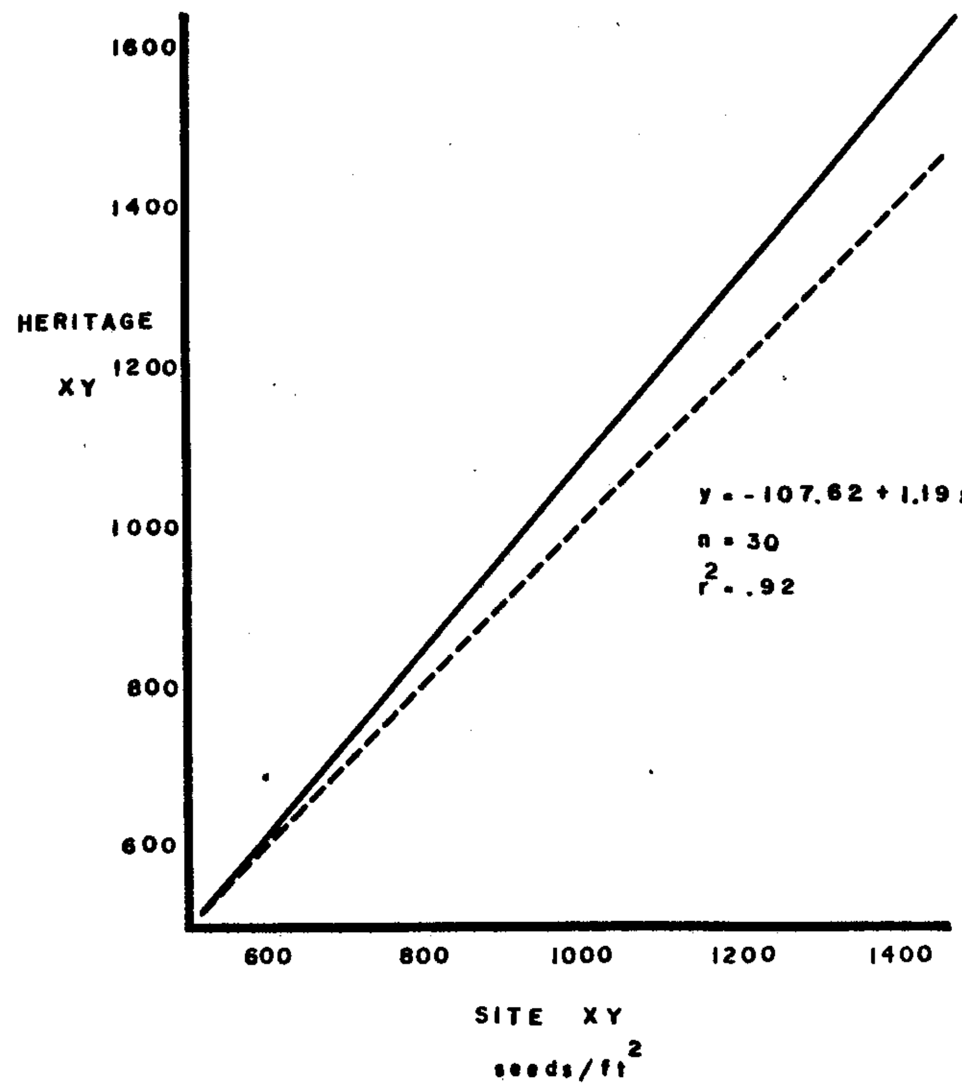
Pedersen, A., E. E. Everson and J. E. Grafius. 1978. The gene pool concept as a basis for cultivar selection and recommendation. Crop Sci. 18:883-886.

Grafius, J. E. Multiple characters and correlated response. Crop Sci. 18:931-934.

CORRELATIONS OF PARENTAL COMBINATION VECTORS WITH IDEAL VECTORS

PARENTS		CROSS					BACKCROSS TO B					BACKCROSS TO A				
A	B	IDEALS= 1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	2	0.5652	0.5879	0.5830			0.4297	0.4493	0.4439			0.6765	0.7013	0.6971		
1	3	0.3227	0.3540	0.3499			-0.0916	-0.0647	-0.0692			0.6233	0.6536	0.6493		
1	4	0.5649	0.5939	0.5861			0.2622	0.2809	0.2784			0.7000	0.7270	0.7226		
1	5	0.8862	0.8995	0.8976			0.9813	0.9836	0.9831			0.8252	0.8462	0.8433		
1	6	0.6841	0.7036	0.6956			0.5158	0.5264	0.5156			0.7436	0.7676	0.7622		
1	7	0.6787	0.6729	0.6724			0.3408	0.3168	0.3186			0.8129	0.8282	0.8257		
1	8	0.7693	0.8089	0.7973			0.6471	0.7000	0.6782			0.7687	0.7996	0.7933		
1	9	0.4908	0.5678	0.5011			0.1482	0.1405	0.1336			0.7034	0.7237	0.7187		
1	10	0.7049	0.7247	0.7287			0.3520	0.3539	0.3507			0.7536	0.7780	0.7744		
1	11	0.2797	0.3052	0.3039			-0.2520	-0.2303	-0.2371			0.6286	0.6563	0.6534		
1	12	0.6467	0.6724	0.6770			-0.5210	-0.5360	-0.5273			0.7690	0.7961	0.7939		
1	13	0.6714	0.6759	0.6774			0.5652	0.5579	0.5619			0.7371	0.7534	0.7522		
1	14	0.5311	0.5406	0.5477			-0.0440	-0.0597	-0.0458			0.7263	0.7484	0.7482		
1	15	0.6475	0.6531	0.6555			0.4943	0.4836	0.4902			0.7248	0.7426	0.7416		
1	16	0.5932	0.5741	0.5766			0.3171	0.3318	0.3376			0.6916	0.7162	0.7152		
1	17	0.6729	0.6894	0.6837			0.5464	0.5566	0.5506			0.7574	0.7809	0.7751		
1	18	0.9184	0.9264	0.9193			0.9359	0.9445	0.9352			0.8390	0.8613	0.8561		
1	19	0.8654	0.8766	0.8801			0.8159	0.8155	0.8231			0.8305	0.8587	0.8503		
1	20	0.8298	0.8241	0.8285			0.6660	0.6447	0.6524			0.8522	0.8653	0.8652		
1	21	0.6375	0.6708	0.5656			0.4077	0.5207	0.5151			0.7298	0.7603	0.7558		
1	22	0.6684	0.6905	0.6885			0.5251	0.5418	0.5410			0.7483	0.7654	0.7625		
1	23	0.1834	0.2452	0.2264			-0.6409	-0.6074	-0.6224			0.6795	0.7171	0.7094		
1	24	0.0193	0.0686	0.0187			-0.4557	-0.4809	-0.4703			0.6777	0.6966	0.6975		
1	25	0.0747	0.1121	0.1059			-0.3960	-0.3681	-0.3733			0.5596	0.5939	0.5880		
1	26	0.2584	0.3122	0.3081			-0.4672	-0.4238	-0.4254			0.6560	0.6935	0.6896		
1	27	0.4765	0.5021	0.4957			-0.0089	0.0810	-0.0037			0.7560	0.7864	0.7813		
1	28	0.4287	0.4598	0.4550			-0.1647	-0.1497	-0.1482			0.6841	0.7129	0.7891		
1	29	0.6689	0.7023	0.6969			0.5268	0.5641	0.5574			0.7300	0.7592	0.7550		
1	30	0.6888	0.7111	0.7065			0.3031	0.3127	0.3097			0.8224	0.8507	0.8463		
1	31	0.7827	0.8058	0.8093			0.6476	0.6600	0.6739			0.7767	0.8018	0.8010		
1	32	0.4063	0.4497	0.4441			-0.0463	-0.0026	-0.0050			0.6685	0.6945	0.6901		
1	33	0.7266	0.7504	0.7500			-0.2079	-0.2167	-0.2117			0.7791	0.8056	0.8027		
2	3	-0.0418	-0.0198	-0.0259			-0.2422	-0.2213	-0.2261			0.1502	0.1603	0.1633		
2	4	0.1496	0.1681	0.1595			0.0176	0.0366	0.0263			0.2338	0.2511	0.2442		
2	5	0.6812	0.6886	0.6845			0.8509	0.8489	0.8473			0.4530	0.4672	0.4619		
2	6	0.3319	0.3440	0.3338			0.3130	0.3193	0.3068			0.3117	0.3266	0.3189		
2	7	0.2832	0.2697	0.2678			0.1801	0.1537	0.1547			0.3164	0.3210	0.3164		
2	8	0.3333	0.3657	0.3587			0.3288	0.3796	0.3528			0.3049	0.3266	0.3177		
2	9	0.1045	0.1043	0.0973			-0.0229	-0.0330	-0.0398			0.2141	0.2234	0.2169		
2	10	0.1819	3.1893	0.1831			-0.5390	-0.0491	-0.0498			0.2570	0.2780	0.2640		
2	11	-0.1353	-0.1214	-0.1243			-0.4092	-0.4010	-0.4087			0.1247	0.1407	0.1359		
2	12	-0.4479	-0.4486	-0.4491			-0.7000	-0.8004	-0.7938			0.1439	0.1574	0.1525		
2	13	0.4196	0.4182	0.4186			0.4456	0.4353	0.4380			0.3610	0.3698	0.3669		
2	14	-0.0848	-0.0076	-0.0016			-0.3923	-0.3757	-0.3685			0.2186	0.2382	0.2276		
2	15	0.3025	0.3009	0.3018			0.2966	0.2825	0.2881			0.2967	0.3051	0.3023		
2	16	0.2047	0.2176	0.2186			0.1380	0.1305	0.1450			0.2559	0.2705	0.2477		
2	17	0.4773	0.4897	0.4827			0.4646	0.4729	0.4661			0.4145	0.4382	0.4230		
2	18	0.7457	0.7578	0.7471			0.9108	0.9166	0.9048			0.4891	0.5043	0.4963		
2	19	0.6746	0.6796	0.6822			0.7529	0.7403	0.7563			0.4670	0.4802	0.4777		
2	20	0.5741	0.5604	0.5638			0.5571	0.5309	0.5383			0.4515	0.4553	0.4532		
2	21	0.3912	0.4211	0.4140			0.3740	0.4059	0.3993			0.3531	0.3767	0.3699		
2	22	0.3725	0.3884	0.3850			0.3799	0.3890	0.3876			0.3355	0.3520	0.3471		
2	23	-0.5214	-0.4828	-0.5028			-0.0402	-0.0144	-0.0294			0.0769	0.1010	0.0907		
2	24	-0.4382	-0.4489	-0.4424			-0.5720	-0.5989	-0.5895			0.0362	0.0407	0.0390		
2	25	-0.2639	-0.2380	-0.2454			-0.4990	-0.4760	-0.4821			0.0519	0.0733	0.0667		
2	26	-0.2589	-0.2261	-0.2313			-0.6180	-0.5858	-0.5881			0.1029	0.1259	0.1200		
2	27	0.0175	0.0364	0.0262			-0.1928	-0.1895	-0.1912			0.2407	0.2607	0.2524		
2	28	-0.0792	-0.0570	-0.0628			-0.4073	-0.3950	-0.3984			0.1713	0.1890	0.1829		
2	29	0.2966	0.3214	0.3146			0.2893	0.3205	0.3131			0.2920	0.3124	0.3062		
2	30	0.3915	0.4114	0.4024			0.1571	0.1640	0.1597			0.3741	0.3944	0.3865		
2	31	0.3293	0.3418	0.3442			0.3160	0.3212	0.3336			0.3057	0.3208	0.3178		
2	32	-0.0182	0.0251	0.0176			-0.2637	-0.2242	-0.2306			0.1871	0.2115	0.2048		
2	33	-0.1446	-0.1375	-0.1427			-0.6478	-0.6667	-0.6631			0.2117	0.2269	0.2208		
3	4	-0.3129	-0.2941	-0.3007			-0.2461	-0.2277	-0.2365			-0.3587	-0.3400	-0.3449		
3	5	0.2024	0.2118	0.2095			0.6431	0.6489	0.6404			-0.1862	-0.1701	-0.1734		
3	6	-0.3180	-0.2878	-0.3030			0.0431	0.0536	0.0394			-0.3879	-0.3671	-0.3742		
3	7	-0.1898	-0.2039	-0.2039			-0.0101	-0.0369	-0.0350			-0.3492	-0.3430	-0.3461		
3	8	-0.3799	-0.3414	-0.3558			-0.2323	-0.1725	-0.2009			-0.3929	-0.3678	-0.3748		
3	9	-0.3049	-0.3044	-0.3099			-0.2188	-0.2280	-0.2340			-0.3649	-0.3545	-0.3591		
3	10	-0.4040	-0.3968	-0.3997			-0.3746	-0.3790	-0.3810			-0.4009	-0.3867	-0.3899		
3	11	-0.6979	-0.6819	-0.6828			-0.6884	-0.6715	-0.6703			-0.5516	-0.5328	-0.5354		
3	12	-0.7991	-0.7991	-0.7969			-0.7924	-0.8060	-0.8085			-0.5896	-0.5772	-0.5787		
3	13	0.1307	0.1284	0.1320			0.3769	0.3633	0.3689			-0.2446	-0.2313	-0.2325		
3	14	-0.5957	-0.5572	-0.5509			-0.5261	-0.5426	-0.5306			-0.4798	-0.4686	-0.4682		
3	15	-0.1648	-0.1670	-0.1629			0.1018	0.0825	0.0915			-0.3298	-0.3180	-0.3180		
3	16	-0.3149	-0.2947	-0.2905			-0.1175	-0.1047	-0.0961			-0.3988	-0.3703	-0.3710		
3	17	0.2824	0.3084	0.2968			0.4086	0.4192	0.4115			-0.2438	-0.2124	-0.2221		
3	18	0.4364	0.4577	0.4428			0.8898	0.8978	0.8833			-0.1593	-0.1381	-0.1456		
3	19	0.1989	0.1959	0.2007			0.5129	0.5086	0.5173			-0.1574	-0.1436	-0.1436		
3	20	0.3180	0.2884	0.2997			0.4807	0.4475	0.4578			-0.2597	-0.2527	-0.2519		
3	21	-0.0042	0.0239	0.0188			0.1819	0.2120	0.2065			-0.2083	-0.1842	-0.1887		
3	22	-0.0308	-0.0148	-0.0165			0.1690	0.1822	0.1816			-0.2256	-0.2080	-0.2107		
3	23	-0.8156	-0.7926	-0.8017			-0.8515	-0.8323	-0.8425			-0.6080	-0.5869	-0.5932		
3	24	-0.9241	-0.9411	-0.9329			-0.7229	-0.7486	-0.7389			-0.7657	-0.7404	-0.7509		
3	25	-0.5291	-0.5100	-0.5138			-0.5753	-0.5564	-0.5683			-0.4676	-0.4487	-0.4523		
3	26	-0.6063	-0.5827	-0.5847			-0.6833	-0.6587	-0.6596			-0.5085	-0.4792	-0.4821		
3	27	-0.3823	-0.3718	-0.3757			-0.3165	-0.3116	-0.3151			-0.4188	-0.3952	-0.3991		
3	28	-0.8217	-0.7995	-0.8035			-0.7858	-0.7723	-0.7746			-0.5687	-0.5476	-0.5515		
3	29	-0.1862	-0.1574	-0.1631			0.0111	0.0459	0.0389			-0.3129	-0.2900	-0.2944		
3	30	-0.2057	-0.1941	-0.1974			-0.0823	-0.0569	-0.0595			-0.3334	-0.3167	-0.3204		
3	31	-0.2924	-0.2782	-0.2730			-0.1054	-0.0998	-0.0851			-0.3628	-0.3456	-0.3460		
3	32	-0.6409	-0.6185	-0.6247			-0.5829	-0.5401	-0.5458			-0.5242	-0.4952	-0.5000		
3	33	-0.6535	-0.6508	-0.6498			-0.6521	-0.6623	-0.6591			-0.5171	-0.5036	-0.5058		
4	5	0.5487	0.5514	0.5452			0.7601	0.7546	0.7527			0.1762	0.1881	0.1782		
4	6	0.0989	0.1145	0.0983			0.2118	0.2156	0.1976			-0.0782	-0.0589	-0.0769		
4	7	0.0217	-0.0028	-0.0047			0.0727	0.0431	0.0441			-0.0841	-0.0884	-0.0973		
4	8	-0.0027	-0.0388	-0.0653			0.0250	0.0800	0.0476			-0.1336	-0.1062	-0.1232		
4	9	-0.1657	-0.1714	-0.1808			-0.1479	-0.1612	-0.1686			-0.1720	-0.1666	-0.1777		
4	10	-0.2360	-0.2347	-0.2422			-0.2655	-0.2752	-0.2793			-0.1933	-0.1827	-0.1929		
4	11	-0.6836	-0.6694	-0.6743			-0.6341	-0.6274	-0.6273			-0.4684	-0.4488	-0.4592		
4	12	-0.7566	-0.7656	-0.7662			-0.7217	-0.7394	-0.7348			-0.4983	-0.4909	-0.4990		
4	13	0.3933	0.3841	0.3845			0.4406	0.4254	3.4298			0.1481	0.1540	0.1466		
4	14	-0.4844	-0.4756	-0.4782			-0.4567	-0.4785	-0.4666			-0.3270	-0.3219	-0.3266		
4	15	0.1229	0.1118	0.1131			0.2278	0.2054	0.2128			-0.0350	-0.0296	-0.0359		
4	16	-0.0628	-0.0460	-0.0445			0.0849	0.0145	0.0218			-0.1312	-0.1119	-0.1184		
4	17	0.4665	0.4811	0.4684			0.4447	0.4526	0.4447			0.2929	0.3209	0.2996		
4	18	0.7712	0.7926	0.7643			0.9016	0.9053	0.8912			0.2807	0.2985	0.2815		
4	19	0.4614	0.4618	0.4644			0.6187	0.6117	0.6195			0.1753	0.1845	0.1800		
4	20	0.5354	0.5093	0.5109			0.5120	0.4801	0.4884			0.2991	0.2921	0.2461		
4	21	0.1784	0.2060	0.1988			0.2685	0.2984	0.2918			0.0396	0.0642	0.0544		


HERITAGE
YIELD
160
150
140
130
120
110
100
90
80
70
60
50
60
70
80
90
100
110
120
130
140
SITE YIELD
bu/acre
y = -9.00 + 1.10x
n = 30
r² = .95


HERITAGE
XY
1600
1400
1200
1000
800
600
600
800
1000
1200
1400
SITE XY
seeds/ft²
y = -107.62 + 1.19x
n = 30
r² = .92

EXPERIMENT 8095 1979 BULKS DATE 02/27/81 PAGE 11 COPY 1

TABLE OF ADJUSTED MEANS

(SORTED ON VARIABLE 12 IN DESCENDING ORDER)

ENTRY NAME		NO.	31 X TILLERS/PLOT	32 Y SEEDS/HEAD	14 Z GMS/SEED	12 W GMS
879-1 03	unselected bulks	3	16.3231	1254.4444	.0404	824.0000
879-1 07		7	17.3524	1125.8111	.0421	808.0000
879-1 04		4	17.7331	1172.1333	.0391	804.0000
879-1 01		1	15.4906	1289.0333	.0400	795.0000
879-1 02		2	16.9532	1223.9778	.0388	793.3333
68-10 5-15 *		9	15.8290	1199.1000	.0413	775.6667
68-10 5-17		17	16.5736	1136.7111	.0409	767.0000
68-10 5-15 *		10	17.1405	1219.3111	.0364	756.0000
M33 *		8	16.5523	1121.2556	.0409	756.6667
879-1 06		6	15.6755	1205.5222	.0395	746.0000
M37		14	16.0250	1167.3000	.0395	739.0000
68-10 3-1 *		11	15.5912	1185.2444	.0391	722.6667
879-1 05		5	15.8577	1087.5889	.0415	714.3333
BONAN ZA		19	16.3335	1188.4667	.0368	712.3333
68-10 5-1 *		13	16.0242	1024.6222	.0435	711.0000
68-10 4-7 *	* parents	12	16.5197	1041.3778	.0409	704.6667
M36		15	14.7641	1168.7667	.0402	691.3333
COHO		20	28.9779	466.5667	.0495	671.6667
LARKE R		18	15.3931	1076.4111	.0400	659.3333
M34		16	14.3475	1066.4556	.0423	646.6667
AVERAGE OF PROCEDING 20 MEANS			16.7729	1120.0050	.0406	739.9833
LSD (P=.05)			2.8649	161.2660	.0018	100.6874
(P=.01)			3.7244	209.6457	.0024	130.8936
COEFFICIENT OF VARIATION			10.4597	8.9174	2.7667	8.3324

8095

EXPERIMENT 8095 1979 BULKS DATE 02/27/81 PAGE 11 COPY

TECHNICAL APPENDIX II

MARCH 1978

RESEARCH REPORT

FROM THE MICHIGAN STATE UNIVERSITY
AGRICULTURAL EXPERIMENT STATION EAST LANSING

An Adaptation of Principal Components Analysis to an Assessment of Genetic Distance

TECHNICAL Information

An Adaptation of Principal Components Analysis to an Assessment of Genetic Distance

By
M. W. Adams and J. V. Wiersma
Department of Crop and Soil Sciences

INTRODUCTION

Plant breeders and plant pathologists, since the epidemic of southern corn leaf blight of 1970, have become acutely aware of the high risk of widespread disease in crops that are highly uniform genetically. The authors of "Genetic Vulnerability in Major Crops" (1) concluded that most crops in the U.S. are "impressively uniform genetically and impressively vulnerable". This opinion was reached on the basis of long experience and personal knowledge of the ancestry of most present-day varieties of field crops. Pedigrees of modern varieties are often complex, or do not exist due to the particular breeding method followed, and breeder selection has almost certainly biased the expected relationships. The judgments of experts, therefore, tend to be qualitative and highly subjective when it comes to estimating genetic homogeneity among varieties of a crop in a major production region.

A general method of quantitatively assessing genetic similarity among a set of varieties of a given crop has been proposed (2) and its application to a particular crop demonstrated. The method is an adaptation of the multi-variate technique of Principal Components Analysis (PCA).

The first intent of this paper is to describe the method in more detail, not necessarily in mathematical terms since these are found in textbooks (2), but in terms that will allow the plant geneticist, breeder, or pathologist to judge for himself the validity and relevance of the method. Second, the rationale of the method devised for weighting the calculated genetic distances between varieties by their relative production acreages in a region is presented. And thirdly, we present the computer program that we have used to calculate genetic distances, with all the essential steps of a worked example.

PART I

Description of the Method

Consider a set of n varieties, each of which has been scored for several, m, metrical traits. With the m variables as coordinate axes, the varieties may be positioned in the multi-dimensional space, their pattern of dispersion resembling the multi-dimensional analogue of an ordinary 2-way scattergram. When the raw data are untransformed and with correlation among the variables, the multi-dimensional scattergram will resemble an ellipsoidal figure. The first and subsequent axes of this ellipsoidal configuration comprise the principal components in a Principal Components Analysis.

PCA consists of a standardization and orthogonal angular rotation of the original axes (variables) to a new set of axes, the principal components. Each principal component is, in reality, a linear combination of the varietal scores on the original variables. A set of m homogeneous equations in m unknowns is generated, expressible ultimately in matrix form as

$$[R_e - \lambda I] = O,$$

where: λ = the diagonal matrix of the latent roots; in Fig. 1 this matrix is marked "M". (Latent roots are also known as eigenroots or eigenvalues. There are m roots, the first root is equivalent to the variance accounted for by the first principal component, the second root is the variance attributable to the second principal component and so on to the m^{th} root and component.)

b = matrix of latent vectors; in Fig. 1 this is denoted as the "E" matrix. (Latent vectors are also known as eigenvectors, and comprise the orthogonal transformation matrix by which the original standardized variates must be multiplied to produce the transformed variates.)

I = the identity matrix, and
R_e = the matrix of correlation coefficients between pairs of variables. (The variance-covariance matrix may replace R_e, depending on the objectives of the user.)

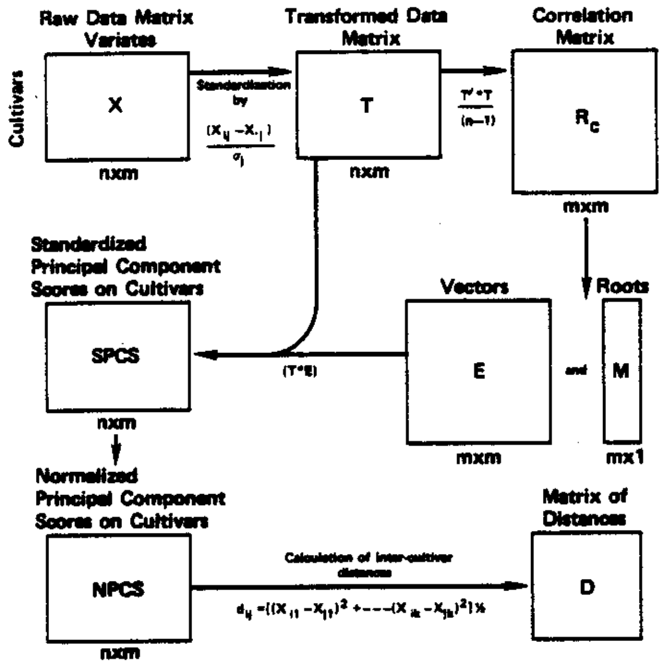


Fig. 1. Flow chart of major steps in a Principal Components Distance analysis.

With varieties simulating points in the multi-dimensional space as determined by their scores on k principal components ($k \leq m$), the geometric distance between any two varieties may be calculated. This distance is termed Euclidean and represents the similarity or non-similarity of the pair of varieties. If two varieties are highly related genetically they are expected to occupy a similar region in the k-space, the distance between them being small. With increasing distance the varieties are more remotely related.

As employed in this application, Principal Components Distance Analysis consists of seven steps (Fig. 1).

The Raw Data Matrix

The initial data may consist of measurements on any number of quantitatively-varying characteristics, such as agronomic data, quality components, and chemical composition of seed or plant parts. Measurements should be made on replicated plots or samples to provide a reliable mean value. Characteristics should be selected that are thought to be representative of the fundamental structure of the biological system being studied, and sufficiently diverse as to represent at least the most important dimensions of the system. A set of traits all of which are highly correlated *inter se* should not be selected, since this would amount to a redundancy of information about one principal axis and nothing for delineating other dimensions.

The genotypes to be measured may be a selected set of varieties, a random sample of varieties from a larger population, or all of the standard varieties of a particular crop grown in a region or country (2).

It is at this stage that one may conceptualize the varieties as an ellipsoidal cloud of points in the multi-dimensional space. Principal Components Analysis consists of finding the principal axes of this ellipsoidal cloud, or of the modified cloud following data transformation.

Transformation of Raw Data

Since the variables may be measured on different scales and be characterized by greatly differing variances, it is generally advisable to transform the data by standardizing to zero mean and unit variance. The matrix so produced is labelled T. This has the effect of moving the axes into the centroid of the data points. It can hardly be thought reasonable to retain the original units of measurements since these are generally chosen arbitrarily and for the convenience of the investigator. After standardization each

Table 1. Raw Data Matrix, X. (Rows and Columns represent cultivars and variables, respectively).

	Col. 1 Col. 10	Col. 2 Col. 11	Col. 3 Col. 12	Col. 4 Col. 13	Col. 5 Col. 14	Col. 6 Col. 15	Col. 7 Col. 16	Col. 8 Col. 17	Col. 9 Col. 18
Row 1	21.1	4.6	0.42	5.31	2.02	0.56	29.3	31.34	19.46
	45	41	8.13	4.1	195	187.3	58	64.98	81.63
Row 2	19.4	5.8	0.35	5.41	2.01	0.47	30	24.91	17.16
	39	45	8.66	4.4	208	162.5	60	55.97	59.57
Row 3	20.4	2.6	0.31	4.92	2.12	0.54	29.8	52.55	29.36
	43	49	10.25	3.7	332	97	50	50	50
Row 4	18.7	1.8	0.36	5.47	2.31	0.53	29.8	32.03	15.32
	55	52	7.4	4.4	137	281.3	44	75	75
Row 5	19.7	2.4	0.38	5.42	2.3	0.58	30.7	46.34	27.95
	54	35	8.06	5.7	145	200.8	54	86.64	75.36

Table 2. Transformed Data Matrix, T (figures arbitrarily limited here to 4 decimals)

	Col. 1 Col. 10	Col. 2 Col. 11	Col. 3 Col. 12	Col. 4 Col. 13	Col. 5 Col. 14	Col. 6 Col. 15	Col. 7 Col. 16	Col. 8 Col. 17	Col. 9 Col. 18
Row 1	1.3426	0.6873	1.3870	0.0179	−0.9029	0.5770	−1.2230	−0.5289	−0.3734
	−0.3136	−0.5079	−0.3439	−0.4797	−0.1074	0.0228	0.7478	−0.1046	1.0188
Row 2	−0.4980	1.3984	−0.3467	0.4654	−0.9713	−1.5868	0.1578	−1.0869	−0.7325
	−1.1690	0.0896	0.1487	−0.0799	0.0588	−0.3492	1.0594	−0.7179	−0.6687
Row 3	0.5846	−0.4977	−1.3375	−1.7274	−0.2189	0.0961	−0.2367	1.3119	1.1722
	−0.5987	0.6872	1.6268	−1.0128	1.6449	−1.3317	−0.4985	−1.1243	−1.4008
Row 4	−1.2559	−0.9717	−0.0990	0.7339	1.0808	−0.1442	−0.2367	−0.4690	−1.0198
	1.1120	1.1354	−1.0225	−0.0799	−0.8493	1.4328	−1.4333	0.5773	0.5116
Row 5	−0.1732	−0.6162	0.3963	0.5101	1.0124	1.0578	1.5386	0.7729	0.9536
	0.9694	−1.4043	−0.4090	1.6526	−0.7469	0.2253	0.1246	1.3696	0.5391

variable contributes equally to the total variance. The varieties may still be visualized as an ellipsoidal cloud, since standardization has not removed correlation among variables, nor have the variety vectors themselves been rescaled to unit dimensions. If there were no correlations among the variables, the transformation would have the effect of reshaping the raw data ellipsoidal figure to a spheroid.

The Correlation Matrix

Multiplication of the transformed data matrix, T, by its transpose, T′, and division by n-1, produces the correlation matrix, R_c. Had not standardization been carried out or had the first transformation been to logarithms, the product of the multiplication would have been a sum of squares-sum of products matrix. At this stage the matrix relationship has reached the form previously given: $[R_c - \lambda I]\ b = 0$. The values of λ and b that make this equation valid are sought next.

Table 3. The Correlation Matrix, R.
(Correlations among variables)

	Col. 1	Col. 2	Col. 3	—	—	—	Col. 18
Row 1	1	0.3157	0.3272	.	.	.	0.0365
Row 2	0.3157	1	0.2466	.	.	.	−0.0918
Row 3	0.3272	0.2466	1	.	.	.	0.9204
—	—	—	—	.	.	.	—
—	—	—	—	.	.	.	—
Row 18	0.0365	−0.0918	0.9204	.	.	.	1

Latent Roots and Vectors, and Normalization

The above equation represents a set of m homogeneous equations in m unknowns. Solution depends on the requirement that the determinant $|R_c - \lambda I|$ equal zero, since b is a matrix of vectors that cannot equal zero. An m^{th} degree polynomial in λ is generated and solved for λ in the computer to produce m latent roots (M). Re-insertion of values of λ into the original set of homogeneous equations produces the vector values b (Matrix E).

Table 4. Eigenroots, M

Row 1	7.9521
Row 2	4.6930
Row 3	3.3471
Row 4	2.0079
Row 5	—
Row 6	—
—	—
—	—
Row 18	−851967E-21

Table 5. Eigenvectors, E

	Col. 1	Col. 2	Col. 3	—	—	—	Col. 18
Row 1	0.1745	−0.0225	−0.3841	.	.	.	−0.1961
Row 2	0.1367	−0.3818	−0.1713	.	.	.	−0.0740
Row 3	−0.2010	−0.2211	−0.2965	.	.	.	0.0156
—	—	—	—	.	.	.	—
—	—	—	—	.	.	.	—
Row 18	−0.2876	−0.1320	−0.1627	.	.	.	−0.1523

Standardized Principal Component Scores on Cultivars

Multiplication of the standardized data matrix (T) by the matrix of vectors (E), symbolized in the matrix notation as T*E, has the effect of rotating the axes of the standardized ellipsoidal data swarm into new directions, orthogonal to each other, such that the first axis is now oriented as the major axis of the ellipsoid, the second at right angles to the first, and each successive axis now an orthogonal axis of the ellipsoid, and so on until all axes have been accounted for.

Table 6. Standardized Principal Component Scores, SPCS

	Col. 1	Col. 2	Col. 3	—	—	—	Col. 18
Row 1	0.0965	1.9032	−1.6452	.	.	.	
Row 2	1.3944	−2.4471	0.3675	.	.	.	
Row 3	3.9085	2.3291	0.4498	.	.	.	
Row 4	−2.6721	0.0903	2.6502	.	.	.	
Row 5	−2.7273	1.9309	−1.8223	.	.	.	

Table 7. Adjusted Eigenroots, N

Row 1	31.8083
Row 2	18.7719
Row 3	13.3882
Row 4	8.0317
Row 5	—
Row 6	—
—	—
—	—
Row 18	—340787E-20
Trace = 72	

Table 8. Square Roots of Retained Eigenroots, S

Row 1	5.6399
Row 2	4.3326
Row 3	3.6590

Each translated axis is equivalent to a principal component. Cultivars have scores on each principal component, just as they had values on the original axes. The first principal component, PC_1, since it is the major axis of the *m*-dimensional ellipsoid, is associated with the largest λ, thus accounting for the largest fraction of total variance in the ellipsoidal cloud. Each successive PC is associated with successively smaller values of λ and thus smaller amounts of variance, until all variance is accounted for.

There are two reasons why interest usually resides in the first *k* of the PC — those that account for some 90-95% of the variance. First, the variance, once dispersed among *m* correlated variables, is now attributable to a much reduced number, perhaps only 3-6, which collectively account for nearly all the original variance. Secondly, they are orthogonal to each other, therefore representing independent genetic contributions to variance.

Normalized Principal Component Scores

The PC's, as first calculated, are associated with different amounts of variance because several of the original variables are highly correlated and therefore contribute redundant information. In the smaller PC's, fewer original variables are contributing to the size of the associated λ's. There is, however, no reason to think that any more genetic information is represented by the larger PC's than by the smaller (significant) PC's, only that the same genetic information (or systematic environmental effects) is represented several times over by the more numerous correlated variables contributing to the larger PC's. To remove this effect so that all PC's are rendered equivalent in effect on the scores to be calculated for each variety, each PC must be divided by the square root of its associated λ, a process called "normalizing".

Geometrically, this has the effect of reshaping the multi-dimensional ellipsoid into a multi-dimensional spheroid, where each axis is of equal length in terms of the variance for which it accounts. In normalized column vectors the sum of squares of the elements in each vector equals unity.

If $k = 8$ PC's have been retained there will be 8 normalized PC scores for each variety. These scores are the values on each axis by which each variety is located in the new transformed space of *k* dimensions. It must be noted that the varieties themselves were never "standardized", only the variable values, the *m* columns of the raw data matrix. Consequently, the varieties are not of unit length in the transformed space. Their magnitudes are retained in the normalized scores.

Table 9. Normalized Principal Component Scores, NPCS

	Col. 1	Col. 2	Col. 3
Row 1	0.017113	−0.439279	−0.449626
Row 2	0.247237	−0.564801	0.100446
Row 3	0.69301	0.53757	0.122927
Row 4	−0.473793	0.0208433	0.72429
Row 5	−0.483567	0.445666	−0.498037

Calculation of Distances

Inter-varietal distances are calculated as the multiple axis analogue of a normal Pythagorean distance:

$$d_{ij} = [(X_{i1} - X_{j1})^2 + \ldots (X_{ik} - X_{jk})^2]^{1/2},$$

where d_{ij} equals the Euclidean distance between varieties i and j on principal axes 1 through *k*, and X_{i1} is the normalized score of variety i on axis 1.

Since the varieties were not constrained to be of equal length, the correlation between any two of them obtained from the sum of cross-products of corresponding PC scores, and equalling the cosine of the angle between the varieties as vectors, does not usually provide the same information as the Euclidean distance. For present purposes the distance value is more appropriate.

A widely used measure of distance between two populations is the D^2 statistic of Mahalanobis (4). Its calculations requires the inverse of the pooled covariance matrix. Kendall (3) has pointed out that it is unsuitable for finding the distance between two

Table 10. Matrix of Distances Between Cultivars, D

	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5
Row 1	0	0.609337	1.31867	1.35306	1.01792
Row 2	0.609337	0	1.1893	1.11895	1.38322
Row 3	1.31867	1.1893	0	1.4017	1.33356
Row 4	1.35306	1.11895	1.4107	0	1.29408
Row 5	1.01792	1.38322	1.33356	1.29408	0

points, such as two varieties, because for a pair of points the covariance matrix is degenerate, and technically a D^2 for those points does not exist.

PART II

Weighting of the Distances

In order to assess the level of genetic vulnerability within a region for a particular crop, a method of weighting of the inter-varietal distances must be found that takes account of perhaps widely unequal acreages among varieties in the region.

For example, other factors being unchanged, a region composed of two widely divergent varieties grown in a 2:1 proportion as respects their acreages would be more vulnerable than if the ratio were 1:1, in keeping with the principle that the more homogeneous the region is for a given crop, the more vulnerable it becomes.

Consider a simple case. Assume a region is composed of two varieties V_1 and V_2 in densities (proportions) v_1 and v_2 ($v_1 + v_2 = 1$) where the distance between them is d_{12}. The distance within varieties equals $d_{11} = d_{22} = O$. Assuming a random distribution of V_1 and V_2 in the region, either as plants or acres (fields would be an approximation), ask as follows: given that an infectious agent is located on a plant (or an acre) of V_1, what is the average expected "genetic distance" to the next plant (or acre)? With the densities treated as probabilities, the answer summed over both initial states is:

$$v_1[(v_1 d_{11}) + (v_2 d_{12})] + v_2[(v_2 d_{22}) + (v_1 d_{12})],$$

which, since $d_{11} = d_{22} = 0$, becomes $2v_1v_2d_{12}$. This procedure can be extended to more than two varieties, to any density, and any inter-varietal distance.

PART III

The Program for Calculating Principal Component Distances

This program is written in the language of SAS-76, the statistical analysis system created at North Carolina State University, and can be run on computer installations containing the SAS package of programs. Slight modifications in the use of certain operational signs or symbols may be necessary dependent upon the particular computer installation available to the user.

Within the text of the program listed below, *any words enclosed in brackets and printed in italics are words of explanation only and are not to be taken as part of the program.*

```
1      DATA PRINCOMP; INPUT
[2-5]  [list here the traits by abbreviation and
       card column numbers, following SAS rules
       for indication of decimals.]
6      CARDS; [data cards follow this statement.]
[7-28]
29     PROC PRINT; TITLE 'PRINCIPAL
       COMPONENT ANALYSIS [number]
       VARIETIES [number] VARIABLES';
30     PROC MATRIX;
31     FETCH X; NOTE RAW DATA MATRIX;
       PRINT X;
32     NOBS=NROW(X); NVAR=NCOL(X);
33     MEAN=J(1,NOBS)*X#/NOBS;
34     X=X-J(NOBS,1)*MEAN;
35     SS=J(1,NOBS)*(X#X);
36     STD=SQRT(SS'#/(NOBS-1));
37     X=X*DIAG(1#/STD);
38     T=X; NOTE TRANSFORMED (STAND-
       ARDIZED) DATA MATRIX; PRINT T;
39     TPT=T'*T;
40     R=TPT#/(NOBS-1); NOTE CORRE-
       LATION MATRIX; PRINT R;
41     EIGEN M E R; NOTE EIGENROOTS;
       PRINT M; NOTE E=EIGENVECTORS;
       PRINT E;
42     SPCS=T*E; NOTE STANDARDIZED
       PRINCIPAL COMPONENT SCORES;
       PRINT SPCS;
43     N=M#(NOBS-1); NOTE ADJUSTED
       EIGENROOTS; PRINT N;
44     TRACE=SUM(N); PRINT TRACE;
45     A=1;
46     LOOP1: H=N(A,1); I=I//H;
       K=SUM(I);
47     IF K>=.95*TRACE THEN GO TO
       PROCEED;
48     Q=SQRT(N(A,1)); S=S//Q;
49     A=A+1;
50     IF A<=NROW(N) THEN GO TO
       LOOP1;
51     PROCEED: NOTE SQUAREROOTS OF
       THE RETAINED EIGENROOTS;
       PRINT S;
52     A=1;
53     LOOP2: B=S(A,*);
54     P=SPCS(*,A);
55     C=P#/B;
56     NPCS=NPCS||C;
57     A=A+1;
58     IF A<=NROW(S) THEN GO TO
       LOOP2;
59     NOTE NORMALIZED PRINCIPAL
       COMPONENT SCORES; PRINT
       NPCS;
60     COL=O;
61     LPA: COL=COL+1;
```

```
62   VEC=VEC||O;
63   IF COL<NOBS THEN GO TO LPA;
64   ROW=O;
65   LPB: ROW=ROW+1;
66   D=D//VEC;
67   IF ROW<NOBS THEN GO TO LPB;
68   L=1;
69   LOOP3: Y=NPCS(L,*);
70   G=L+1;
71   LOOP4: W=NPCS(G,*);
72   F=Y—W;
73   U=SSQ(F);
74   V=SQRT(U); D(L,G)=V; D(G,L)=V;
75   G=G+1;
76   IF G<=NOBS THEN GO TO LOOP4;
77   L=L+1;
78   IF L<NOBS THEN GO TO LOOP3;
79   NOTE MATRIX OF DISTANCES;
     PRINT D;
80   /*
[END OF FILE]
```

The following description pertains to the numbered lines in the above program and to the major steps in the flow-chart of Fig. 1.

Program lines	Fig. 1 step #	Function
1–5		Describes data by name or abbreviation, location by card columns, and number of decimals in the data.
6–28		Data cards read in. In this example there are 22 varieties or observations and 18 variables (NOBS and NVAR, respectively).
29		Prints title.
30–31	1	Calls Matrix Procedure; calls and prints Raw Data Matrix X.
32		Declares form and dimensions of data matrix.
33–37	2	Standardizes raw data matrix.
38		Identifies and prints the standardized Matrix T.
39–40	3	Calculates and prints the Correlation Matrix R.
41	4	Extracts and prints Eigenroots and Eigenvectors M and E.
42	5	Calculates and prints Standardized Principal Component Scores on cultivars.
43		Adjusts eigenroots to equal roots of variance-covariance matrix.
44		Calculates and prints Trace, as sum of adjusted eigenroots.
45–50		Loop for eliminating all roots that remain (m-k) after nearly 95% of Trace has been accounted for by first *k* roots.
51		Calculates and prints square roots of retained eigenroots (S); these values needed in normalization of SPCS.
52–59		Calculates and prints Normalized Principal Component Scores, NPCS.
60–67		Forms a square matrix of zeros, to be replaced by individual distance values as these are calculated.
68–78		Inter-cultivar distances are calculated and placed in the previously created square matrix.
79–80		Labels and prints matrix of distances and terminates the program.

A Worked Example

In this publication, the use of the PCA distance method has been illustrated by a worked example. Such an example should convey a greater sense of practicality of the method and also make it possible for interested users to check their results with ours. The tables distributed throughout this report (Tables 1-10) constitute a worked example, providing a sample analysis at each significant step in the procedure. To keep the tables small, data from only 5 varieties and 18 variables were used in the example. The 18 variables are the chemical compositional and agronomic variables described in reference 2.

References

1. Anonymous (1972). Genetic Vulnerability of Major Crops. Nat. Acad. Sciences, Washington, D.C.
2. Adams, M. W. (1977). An estimation of homogeneity in crop plants, with special reference to genetic vulnerability in the dry bean, *Phaseolus vulgaris* L. Euphytica (in press).
3. Kendall, Sir Maurice. (1977). Multivariate Analysis. Charles Griffin and Co., Ltd.
4. Murty, B. R. and V. Arunachalam. (1967). Computer programs for some problems in biometrical genetics I. Use of Mahalanobis' D^2 in classificatory problems. Ind. J. Gen. and Pl. Br. 27: 60-69.

3-78-2M