

32729

APPROACHES TO LONG RANGE FORECASTING



DISTRIBUTION STATEMENT A

**Approved for public release;
Distribution Unlimited**

**A Symposium held at
Alamogordo, New Mexico**

29 - 30 APRIL 1969

19971208 069

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 29-30 April 1969	3. REPORT TYPE AND DATES COVERED Symposium, 29-30 Apr 69		
4. TITLE AND SUBTITLE Approaches to Long Range Forecasting - A Symposium held at Alamogordo, New Mexico, 29-30 April 1969			5. FUNDING NUMBERS	
6. AUTHOR(S) James P. Bennett; Noland D. Harris; Richard C. Taeuber; Robert W. Paterson; Ambrose B. Nutt; Marvin J. Cetron; Donald N. Dick; Rudolph J. Rummel; Frederick W. Ward, Jr.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Aerospace Studies (AFMC) 3550 Aberdeen Ave SE Kirtland AFB NM 87117-5776			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution unlimited.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) Table of Contents: Simulating Alternative Economic Futures; Forecasting Military Requirements: A Critical Viewpoint of the Industry's Approach; Election Night Forecasting; Forecasting Highway Demand; Technological Forecasting in an R&D Laboratory; Producing the First Navy Technological Forecast; Forecasting International Relations: A Proposed Investigation of Three-Mode Factor Analysis; Forecasting Solar Events.				
14. SUBJECT TERMS Operations Research Analysis			15. NUMBER OF PAGES	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL	

DTIC QUALITY INSPECTED 2

APPROACHES TO LONG RANGE FORECASTING



**A Symposium held at
Alamogordo, New Mexico**

29 - 30 APRIL 1969

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402—Price \$3.00

PREFACE

These Proceedings of the Third Symposium on Long Range Forecasting and Planning contain the papers which were presented and the discussion of those papers. The discussion was recorded during the Symposium and has been edited to clarify the discussants' meaning. However, during the editing process a conscious effort was made to retain the flavor of free discussion which marked the Symposium. We feel that the interchange of information and opinion during the discussion adds significantly to the formal presentations, and we hope that both the papers and the discussion will prove to be of benefit to all those concerned with methods of long range forecasting and planning. The reader should remember that the opinions presented by the speakers, particularly during the discussion, are their own and do not necessarily represent the official position of their organizations.

Special thanks are due to Mrs. F. E. Carter from the Office of the Staff Judge Advocate, Holloman AFB, for recording the discussion of the formal papers. We also wish to acknowledge the efforts of Mrs. Linda Martin and Mrs. Pat Montoya in preparing the manuscript for publication.

JOSEPH P. MARTINO, *Lt Colonel, USAF*
Assistant for Research Analysis

THOMAS E. OBERBECK, PhD
Advisor for Operations Research

This Page Intentionally
Left Blank

Contents

	<i>Page No.</i>
SIMULATING ALTERNATIVE ECONOMIC FUTURES by James P. Bennett.....	1
FORECASTING MILITARY REQUIREMENTS: A CRITICAL VIEW-POINT OF THE INDUSTRY'S APPROACH by Noland D. Harris.....	13
ELECTION NIGHT FORECASTING by Richard C. Taeuber.....	28
FORECASTING HIGHWAY DEMAND by Robert W. Paterson.....	45
TECHNOLOGICAL FORECASTING IN AN R&D LABORATORY by Ambrose B. Nutt.....	61
PRODUCING THE FIRST NAVY TECHNOLOGICAL FORECAST by Marvin J. Cetron and Donald N. Dick.....	78
FORECASTING INTERNATIONAL RELATIONS: A PROPOSED INVESTIGATION OF THREE-MODE FACTOR ANALYSIS by Rudolph J. Rummel.....	94
FORECASTING SOLAR EVENTS by Frederick W. Ward, Jr.....	118

LIST OF ILLUSTRATIONS

Figure III—

1	Illinois, Indiana and the U.S.....	31
2	Network Election System.....	32
3	ABC Network System.....	34
4	Form of ABC Projection Model.....	37
5	ABC Decision Groups.....	38

Figure V—

1	RDE Flow Diagram.....	63
2	TORQUE Simplified Operational Diagram.....	64
3	Speed Trends.....	65
4	Trends of Aircraft Noise.....	65
5	Variable Geometry Complexity.....	66
6	Structural Weight Fraction Trends.....	67
7	Gross Weight Trends.....	67

Figure VI—

1	The Navy Technological Forecast Program.....	79
2	The 1968 Navy Technological Forecast.....	80
3	Exploratory Development Planning Structure.....	82
4	Relative Quality of the 1968 Navy Technological Forecast.....	84
5	Technological Forecast: High Power Density Batteries.....	85
6	Technological Forecast in Breadth: Metallic Materials Technology.....	86

Figure VII—

1	Phases and Tasks of the DON Project.....	97
2	Three Modes of National Attributes or Behavior.....	98
3	Two-Dimensional Views of the Three Modes.....	100
4	Matrix Patterns for a Single Time Period.....	102
5	Change of Patterns Over Time.....	103
6	Supermatrix of Attributes, Years and Nations.....	105
7	Hypothetical Plots for Four Dyads.....	107
8	Hypothetical Attribute Patterns.....	108

Figure VIII—

1	The Sun.....	119
2	Distribution of Energy in the Solar Spectrum.....	120
3	Solar Corona—12 November 1966.....	121
4	Variation of the Sunspot Number During the Years 1700 to 1968.....	122
5	The Solar Disk as Observed in H-Alpha.....	123
6	Interaction of a Solar Flare with the Earth.....	125
7	Solar Flare Classification.....	126
8	Sudden Ionospheric Disturbances.....	127
9	Effects of Sudden Ionospheric Disturbances.....	127
10	Effects of Solar Particles.....	128
11	Polar Cap Absorption.....	129
12	HF Radio Propagation.....	130
13	The Auroral Zone.....	131
14	Frequency Ranges and Propagation Modes for Radio Communication.....	132
15	Military Radio Communication Systems.....	132
16	High Frequency Management.....	133
17	Solar Flare Effects on High Frequency Communication.....	134
18	Ionospheric Scatter Communication.....	134
19	Tropospheric Scatter Communication.....	135
20	Microwave Communication.....	135
21	Effects on Manned Space Program.....	136
22	New Solar Flare Reporting Code.....	140
23	Tabulation of Erroneous Flare Reports.....	142

Simulating Alternative Economic Futures

by JAMES P. BENNETT, U.S. Treasury Department

This paper treats a newly emerging technique of perceiving and analyzing likely economic futures. The technique is the econometric model which is a system of equations expressing quantitative links between alternative economic futures and their required government policies. Until very recently, any paper on methods of economic forecasting would have covered a wide variety of forecasting techniques including the equation. But for those more concerned with the future than with the past a statement and evaluation of this emerging technique seems appropriate since it is now becoming central to quantitative analysis of the economic future. Because of the time constraint, the discussion will deal only with the macroeconometric model, that is, of a system of equations describing the entire national economy.

To appreciate this new technique and its potential, it is important to understand four very significant changes that have occurred in 20th century economics. These are (1) the shift from economics as a mechanistic science to economics as a system science, (2) the change from a government policy of minimum interference in the economy to a policy of active intervention, (3) the incorporation of many different older techniques of forecasting into the equation method, and, most importantly of all, (4) the shift from forecasting one nearly inevitable

economic outcome to simulating a *set* of attainable economic futures.

The first part of this paper is a discussion of these four changes—mechanism to system, minimum interference to intervention, the development of the econometric model, and forecasting to simulating. The second part is a discussion of how equations are fitted and tested, how the full set of desired equations is tested as a complete econometric model, and then how the model is used to simulate the future. The last part treats various aspects of the reliability of economic simulation.

The first major developments in economics came largely in the 19th century. The ideas of Malthus, Ricardo, Marx, and Marshall were heavily influenced by the enormous success of classical physics. The results of laboratory control of experimental variables and the discovery of the laws of orbital mechanics, for instance, led 19th century social scientists to think that techniques of the physical sciences could easily uncover laws of society. This attempt at what we now call reductionism—the transfer of techniques of one science to another—led to concepts such as economic man, the stationary state, general equilibrium, and, in the case of Marx, iron laws of history. As a general rule, the classical methodology was to begin with assumptions or simple observations about individual behavior and derive, *a priori*, complex theories of economic behavior. Both as a result of data inadequacies and a lack of quantitative techniques, the behavioral hypotheses underlying the theories were never tested.

Theories, no matter how elegant, that fail to predict have to be modified or discarded. The classical economics and the Marxist model of social change failed to explain the development and behavior of the giant corporation, the chronic cycle of economic expansion and contraction, and continual economic growth and change.

With the arrival of the 20th century, dissatisfied

* Dr. Bennett received his Ph.D. degree in economics from the University of Michigan in 1964. He is a member of Phi Beta Kappa and an honors graduate of the Pennsylvania State University.

As a staff economist at the Board of Governors of the Federal Reserve System from 1964 to 1966, he had major responsibility for analysis and projection of the United States price level and development of the index of industrial production. Presently he is a forecaster, revenue estimator, and econometric model builder at the U.S. Treasury Department.

His major published articles deal chiefly with econometric forecasting and simulation. In addition, he is the economic forecasting editor of *The Futurist*, a journal of the World Future Society.

economists such as Wesley Mitchell, Thorstein Veblen, and Joseph Schumpeter began to produce empirical and historical studies of business cycles, large scale business organizations, and economic growth. Rather dramatically, existing economic theory was found totally unable to explain the tragic Great Depression of the 1930's, and with the publishing of Keynes *General Theory*, modern empirical macroeconomics emerged. In trying to explain why from 1929 to 1933 the output of the economy fell by over 30 percent and the unemployment rate rose to 25 percent of the labor force, the concepts and methodology of the new economics emerged as quite different from those of the old. The new macroeconomics began to explore empirical relationships of aggregate behavior which were part of a simultaneous system. The behavior of the system, furthermore, was best studied as a stochastic process, as a sequence of temporal events analyzed by the theory of probability rather than in some alternative deterministic framework. Guided by only a skeleton of theory, the search for stable empirical relationships in the form of regression equations began. One of the first discoveries was that the old puritan virtue of individual thrift easily explained by the older classical economics could, under certain conditions, be the cause of a depression or high level stagnation.

As the older, deductive economics gave way to the empirical testing of alternative behavioral hypotheses, developments in social psychology and sociology further reinforced the shift toward treating economics as a subsystem interacting with a very complex, evolving social system. As economics rejoined social science, new interdisciplinary studies emerged, dealing with such topics as the psychological determinants of economic behavior, the anthropological factors in the economic development of underdeveloped countries, the political behavior of economic elites, and the sociological determinants of labor force mobility.

Economics as a system science enables us to study—and use for forecasting—empirical regularities that have been evident in the past and which can be expected to persist into the future. The set of empirical equations, for example, relating the rate of unemployment to inflation has a tenuous link with older economic theory, yet the regularity describing observed behavior aids the Federal Government in making intelligent fiscal and monetary policy decisions.

At about the same time that economists began empirical studies of the economy, conventional ideas

of the desirability of laissez-faire were rapidly being replaced by ideas of economic intervention. Anti-trust laws, regulatory agencies, and the Federal Reserve System all came into being before the first world war. The process of increased economic control was considerably speeded up by the Great Depression. And the employment act of 1946 committed the Federal Government not only to preventing the recurrence of a serious depression but also to the more difficult goal of continuous full employment. Monetary and fiscal policies were to be used as the chief instruments of control; by varying tax rates, Federal expenditures, and the rate of increase of money and credit supplies, the target variables of output, employment, the growth rate, the price level, and the balance of payments would be maintained at or near their desired values.

In essence, American society began to take control over its economic future. No longer were random shocks, accidents of history, or other unknown forces given free rein. A new experiment began in which a society used sophisticated techniques of perceiving and influencing the future to choose the economic future that it desired.

The third major change in 20th century economics has been the absorption of a variety of forecasting techniques into one technique—that of the equation. For example, the technique that we know as extrapolation is essentially using intuitively some function, usually linear, to extend a regularity of observed behavior into the future. The use of a fitted trend function simply formalizes the extrapolation; a given variable in the past has been closely associated with the passage of time and use of the trend function for forecasting requires the assumption that the temporal relationship will persist into the future.

Another example of an older technique of economic forecasting is the leading indicator—a variable which in the past has preceded with acceptable regularity the variable to be forecast. A good example is new orders of machinery and equipment preceding corporate expenditures on new plant and equipment. Experience has shown, however, that if expenditures on plant and equipment are fitted to new orders the forecast errors of the equation are too large to be acceptable for forecasting. In other words, the use of mathematical statistics has shown that the older technique is inadequate for present needs.

Diffusion indexes, another older technique, measure the extent of change in the economy. To what degree an inflation is spreading, for example, is

measured by a diffusion index for perhaps 1,000 key industrial prices. An index of 50 means that one-half of the prices are rising and that one-half are falling; if the index rises to 90, then 90 percent of the prices are rising and the economy has moved from a state of price level stability to one of broadly based inflation. Again, the test of fitting a measure of inflation to a price diffusion index even with complex distributed lags will fail to uncover a stable empirical relationship.

In the same class as Delphi techniques are surveys of anticipations or expectations of informed individuals that are used as leading indicators or as inputs into equations. Two of the more important are the survey of consumer buying intentions and the survey of corporate plans for investment in new plant and equipment. Equation fits, as a general rule, lead to adequate forecasting results for up to no more than two years.

Great ratios, another technique, describe the outer limits within which variables may fall. From historical experience, for example, spending on new plant and equipment in the United States economy typically lies between 9.5 percent and 11.5 percent of gross national product. While ratio analysis has been useful in the past, it is clear that a 1 percentage point error in a projection of the investment ratio, given the \$1 trillion GNP of 1970, is equal to $2\frac{1}{2}$ times the cost of the space program or about one-half total Federal nondefense purchases. Thus, using seemingly stable ratios directly for forecasting or incorporating them into equations can easily give rise to errors of large dollar magnitude.

In general, the older techniques were useful in anticipating an economic future subject to little or no social control, but for present needs have become irrelevant, unreliable, or inputs into econometric equations. The new equation technique occasionally uses the information found in the older techniques, but more frequently, incorporates strong hypotheses of economic behavior that are consistent with economic theory.

The fourth significant change occurring in 20th century economics—one that is not yet completed—has been from the vague anticipation of the short run future of an uncontrolled economy to econometric simulation of highly plausible and attainable economic futures. Obviously, before this development could occur economics had to become a system science, social control over the vagaries of the business cycle had to be obtained, and the equation had to become central to forecasting.

In principle, any set of government policies

described by such factors as the degree of tightness of monetary policy and the size of the Federal budget surplus or deficit can be linked quantitatively to future outcomes such as full employment and a strong inflation, or a high rate of unemployment and price level stability. The object of the simulation is, of course, to identify those sets of policies required for maintaining the best possible economic outcome. Given the realities of the early 1970's, a very good outcome would be a growth rate of $4\frac{1}{2}$ percent, an unemployment rate of 3 percent, an inflation rate of no more than 2 percent, and equilibrium in the balance of international payments.

An econometric model is a set of simultaneous equations describing interrelated economic processes such as demand for the economy's output, income payments to economic participants, price and wage change, and financial flows. The equations are fitted by statistical techniques to historical data most of which are collected and published by the Federal Government. A typical equation might show that consumer demand for new cars depends on after tax income, bank deposits and other liquid assets, the price of new cars relative to the prices of other consumer goods and services, and the number of old cars on the road that need replacing. Given the estimated coefficients, the equation might state that, if nothing else changes, a \$1 billion rise in after tax income will result in new car outlays rising by about one-tenth of \$1 billion.

Each equation can contain three types of variables—endogenous variables, exogenous policy variables or instruments, and other exogenous variables. The endogenous variables are those to be predicted by the system of equations. They include major consumer outlays such as cars, services, and household durable goods; business spending on new plant and equipment; corporate profits; consumer income; Federal and state tax revenue; unemployment; and numerous other variables that make up or describe the circular flow of demand, production, and income.

The exogenous policy variables or "instruments" are those controlled or influenced by the government. They include government purchases of goods and services, old age pensions and aid to families with dependent children, all tax rates, the money supply, the availability of bank loans and long term money capital, and interest rates. The other exogenous variables are those neither predicted by the system nor under the control of the government. A good example is export demand which depends largely on economic activity abroad.

The initial step in developing an equation is to choose a set of variables expressing strong behavioral hypotheses. As an example, the change in the aggregate wage rate—a variable highly important in determining the price level and hence the rate of inflation—can be explained by three endogenous variables. These are the unemployment rate, the corporate profit rate, and the consumer price index. The unemployment rate measures the strength of demand for labor and consequently the rate at which wages rise. With a very low rate of unemployment, the number of available jobs will tend to exceed the number of people able to fill them; the competition of employers bidding for a limited supply will drive up wages and salaries. The corporate profit rate, on the other hand, serves both as a negotiation target for organized labor in wage contracts and also as a measure of the unwillingness of management to take a strike. With a fully employed economy permitting rapid advances in sales and profits, large wage increases can easily be passed on to consumers in the form of price increases. Finally, the consumer price index serves as a second target for organized labor; given an inflation, labor will attempt to negotiate either escalator clauses or wage packages that stay ahead of the increase in the cost of living. Thus, it is easy to recognize that the equation can express an inflationary dynamic in which each variable may take turns in driving the wage level upward.

Once a relevant set of variables is chosen, repeated experimental fits by computer are required to arrive at one specification of the equation that appears to be sufficiently stable for forecasting. The search for stability of the relationship—for indications that the regularity will persist into the future—is perhaps the single most important task in econometric model building.

In fitting and testing equations, quarterly data from 1947 to the present time are typically used. Data before 1947 are too strongly influenced by the high pressure economy of the World War II period, and during the Great Depression by the total collapse of the economy. Data before 1929 are either too unreliable or nonexistent.

The time period for fitting and testing equations is broken up into three parts—the sample or fitting period, the post-sample period, and the actual forecast period. The sample period to which the equation is fitted might be from 1947 to perhaps 1960. The critical assumption is required that this sample period contains all of the needed information; that

is, that the economic behavior expressed by the equation will persist into the future.

The next step is to test the stability of the equation in the post-sample period. Actual quarterly values for the independent variables are fed into the equation for the period 1961 through mid-1969, and predicted values for the dependent variable are compared to the actual values. If these residuals are not significantly larger than those for the sample period, a stable equation acceptable for forecasting has been found.

What happens in practice is typically discouraging. An equation that appears excellent during the sample period may perform fairly well up to perhaps 1965. Afterward, the equation may predict a world that never existed; in other words, the residuals become increasingly larger in one direction or another.

The final period is the actual forecast period—from mid-1969 into the future. There is, of course, no performance test except that deriving from the ordinary passage of time. However, one rough check on the reliability of an equation in the forecast period is to extrapolate by any method values of the independent variables and then compare the forecast values of the dependent variable with the independent estimates.

In a macroeconomic model of the national economy, equations are grouped by sectors. The aggregate demand for the output of the economy might be described by eight equations—business demand for structures, equipment, and inventories, and consumer demand for residential structures, autos, durables other than autos, nondurables, and services. To the sum of these spending categories one would add government purchases of goods and services and net exports to get the gross national product. Note that government purchases constitute an exogenous policy variable; by varying the rate of change of this variable, the Federal Government has considerable control over total demand and hence employment, capacity utilization, the economic growth rate and other target variables.

Another set of equations describing the components of aggregate supply such as labor force and capital stock interact with aggregate demand to form the basic model. Income equations are then added to explain wages and salaries, dividends, rent, and other types of income. Tax equations follow logically as claims on the income flows. And the unemployment rate, productivity, and other endogenous variables are used in still other equations to explain wage and price levels.

In the same way that each of the equations was tested for reliability in the post-sample period, so is the complete equation system. In particular, the period 1965 to 1969 is well suited for testing because of the severe shock to the economy resulting from the Vietnam buildup. At the same time that the economy was close to full employment, a very sharp increase in defense spending occurred. This abrupt increase in aggregate demand caused an immediate and sharp decline in the unemployment rate and severe stress in the material, product, and labor markets. The resulting inflationary dynamic bears considerable resemblance to those of other postwar inflationary periods.

Thus, if the complete equation system adequately predicts economic behavior during the Vietnam period in addition to the earlier part of the post-sample period, it is adequate for simulating the future. In other words, as the economic information embodied in the equations from the sample period continues to describe recent behavior of the post-sample period, the assumption can be made that these empirical regularities will persist.

The system of equations can be solved by computer with a relatively simple iteration program by placing values of the exogenous variables into the equations and solving for the endogenous variables. Note that alternative pasts can be analyzed as easily as alternative futures and typically produce quite useful results with regard to policy errors of the past.

The first short run simulation which is usually no more than two years into the future is called the control solution. Likely future values of exogenous variables are first extrapolated. Variables such as population, exports, and mandays idle due to strikes are projected with the advice of experts at organizations such as the Census Bureau, the Department of Commerce, and the Department of Labor. The exogenous policy variables—those under the control or influence of the Federal Government—are extrapolated after careful scrutiny of the Federal budget, pending economic legislation, and the stance of monetary policy. Typical projections might incorporate the assumptions that tax rates will remain unchanged, that government purchases and transfer payments will be at levels outlined in the budget, and that supplies of money and credit will continue to grow at recent rates.

The control solution might show that with the assumed government policies, the economy will move into a recession or a stage of high level stagna-

tion with the characteristic slow growth of output and unacceptably high rate of unemployment. One reason might be that a rapidly expanding U.S. economy raises Federal tax revenue by about \$10 to \$12 billion a year without any tax rate increase; this "fiscal drag" acts as a brake on the economy unless the increased revenues are returned to the economy through tax rate reductions, expenditure increases, or a combination of both.

The next sequence of simulations would involve experimenting with different sets of government policies in order to identify those particular sets required for achieving the desired economic outcome. From the previous example, a 10 percent reduction in personal and corporate tax rates might well achieve the targets of rapid growth and full employment, but might also create the adverse effect of a high rate of inflation. The next simulation would incorporate a change in the policy mix. Perhaps the cause of the inflation is that the increase in demand for the economy's output exceeds the increase in industrial capacity. In this simulation, a tax cut only for corporations along with a less restrictive monetary policy would create business incentives for a faster expansion of capacity, and hence a lower rate of inflation. In practice, the effects of numerous variations of government policies are studied at the same time. On a fast computer, twenty simulations might take ten minutes. It is important to understand, however, that the realism of the simulation depends on the care with which the input data are prepared, not on the speed of the computer.

A growing use of econometric models is exploring the path of the economy to a predetermined economic scenario. For example, a scenario describing a peacetime economy after Vietnam demobilization might be based on relatively stable empirical relationships which have been evident in the United States for generations. Relevant exogenous variables would be extrapolated and linked to the scenario, and the model would be solved for the path of the endogenous variables. A prime objective of the research is to determine the types and timing of economic policy changes such that demobilization would not interfere with continuous achievement of the policy targets.

Econometric models have been improved significantly in recent years, but are in what might be called the middle years of development. While the seeming precision of equation coefficients and the use of sophisticated computer programs tend to create an aura of scientific certainty around a model,

it is clear that simulation estimates will remain above desired error levels for a number of years.

This failure of economics with its advanced techniques to simulate the future with a high degree of reliability is frequently noted by noneconomists. Furthermore, with even very short run simulations correct little more than half the time, long run simulations can hardly do better.

The answer to this question of failure is twofold. First, variables from outside the economic subsystem and about which we know very little affect the equations in unknown ways. Many of these exogenous variables derive from technological change. One example is the extent to which emerging communications technology will affect the demand for all types of transportation. Communications is a partial substitute for transportation and how the widespread use of the large screen videophone, for instance, will affect the demand for new cars and commercial airline travel is unknown. At the response threshold, the decline in demand for airline travel may be quite dramatic, or alternatively, there may be no effect at all. Thus, the equations lack important variables entering from other parts of the social system.

The second part of the answer for the failure of economic simulation is considerably more fundamental than the failure to involve technological and other noneconomic variables in the equations. This is the question as to whether historical information will ever be adequate for discovering structural equations. In other words, it may be impossible to uncover stable empirical relationships expressing strong behavioral hypotheses, consistent with the best economic and social theory, and which will predict far into the future.

Nonlinearities, for example, that might not affect the equation during the sample period may prove to be of major importance in the forecast period. The logistic of "S" function is one example. It is hard to believe that the number of motor vehicles on the road will continue to increase at past rates; saturation may well be reached when each family has seven cars. The point is that an equation fitted

to 20 or 30 years of recent historical experience may totally fail to capture a long run nonlinearity and so be inadequate for long run forecasting.

Another example of the embodiment of incomplete information is an equation fitted to a period such as 1947 to 1961 that has been characterized by continual economic fluctuations. Using the equation for forecasting when a more perfectly controlled economy has achieved its growth and employment targets may lead the policy maker to make a serious error such as bringing about a recession. In other words, with more effective economic control, the relevance of the past for fitting equations decreases. Perhaps as a general proposition, the relevance of the past for policy making is inversely related to the speed of social change; history may become the study of dead cultures and sociology may become the study of recent history.

Thus, economists are forced to simulate the future with equations having not only incomplete information from within the economic subsystem but also notoriously weak empirical links to other subsystems. The equations currently used for economic simulation are in general only empirical regularities containing weak behavioral hypotheses.

The point in the future at which one can no longer have faith in the persistence of the regularity is called the forecast horizon. For the best econometric models in existence, the forecast horizon is about two or three years in the future. Using these models or some of the new long run growth models now in research stage for simulations beyond the forecast horizon requires the assumption that the economy of the far future will be very similar to that of today but much bigger. Given the massive and relentless change that the society is experiencing today, it is clear that while the economy will be bigger, it will also be very different.

Thus, further advances in simulating attainable economic futures require improved forecasting of technological change and its diffusion rate, the discovery of long run nonlinearities in the economic system, and considerably more knowledge of the effect of rapid social change on key economic variables.

DISCUSSION FOLLOWING PAPER BY DR. JAMES P. BENNETT, U.S. TREASURY DEPARTMENT, "SIMULATING ALTERNATIVE ECONOMIC FUTURES."

DR. FRED WARD: I have to confess that one of the biggest reasons for coming here is that I saw that there were going to be a number of people discussing

subjects about which I am almost totally ignorant. Yet this subject is one on which I have very strong feelings, and they will probably show in the question.

I will start off by asking: how many degrees of freedom do you have in most of your variables from 1947 to 1961?

DR. BENNETT: About fifty.

Q. Fifty? Five-o? It seems unbelievable. I would expect that you would have closer to five rather than fifty, just looking at the long term trends which are producing most of the correlations, and introducing a tremendous amount of scattering in the relationships.

A. Yes, but suppose we are dealing with the first difference in automobile consumption; the level of the automobile demand is at thirty billion dollars a year. The differences are averaging one and a half or two billion. If we can hold our forecasting down to a quarter of a billion, we will have it.

Q. You have a certain amount of information in your data. You can take first differences, second differences, fourth differences, and you are not increasing your total amount of information, period. There is just a certain amount of information available. Let me get on to my overall question. It seems to me from what you have said that you have violated almost every rule of statistics. You have a non-stationary time series; you have no measure of independence and no way of finding it; you have statistics which are much too fancy and much too complicated for the amount of data, i.e., the *independent* data that you have. And out of this you get results which, by your own admission, tend to be not tremendously good for projecting into the future—which is exactly what one would expect, given what you have done with the data. Now, if you say the reason you are doing this is because there are outside variables introducing other errors, I can agree there are a lot of outside variables. You also do not know, for example, whether the variables you are using—e.g., this consumer price index—have a feedback on the variables you are trying to find. And it is so complicated a variable in itself, I doubt if you could find out if it had a feedback, anyway.

A. That is what we use the two stage and the three stage least squares for.

Q. But you still have to satisfy the same basic conditions in order to develop and use a regression equation, no matter how many stages you use. For example, it must approach some degree of stationarity. You give all the reasons that your approach is effective even though you are not saying explicitly the reasons your time series are totally non-stationary. With all your time series the results you get on one sample should be about the same as if

you took another completely independent sample of the same series. Any sample should contain all of the interrelationships and in roughly the same form. In other words, you assume a relatively long series from minus infinity to positive infinity; and when you take a chunk out of this, you should get relatively the same results?

A. Yes.

Q. But by your own admission you are not getting the same results?

A. Yes.

Q. You have a tremendous non-stationarity in your series. Stationarity is an absolute requirement in order to extrapolate any results you get from the past into the future. Not only do you not have stationarity, but also it seems from everything you have said that you have very few degrees of freedom in your sample. It seems to me that your methods are covering up for the very basic difficulties in your statistics. Fancy statistics are only justified when you have tremendous amounts of data. You can get fancier and fancier, the better and more independent data that you have. Frankly, from what you said, I am surprised you have any forecast stability at all. I have worked with data roughly the length you have where the degrees of freedom must have been almost two orders of magnitude larger than yours, and even then there was some instability. I do not understand how you expect to get information out of your data when it is not there to start with.

A. I agree with much of what you say. All economists are aware of these problems; and in an economy like ours which has experienced some very bad depressions and recessions, this is one attempt to get at what will happen in the future. This is why I stressed the testing of the equation in the proposed sample period to see how it performs. The example I used was from 1960 quarterly to the present time. We find many or most of our equations do quite well—that is, the forecast errors tend to be small and be almost normally distributed.

Q. But you have a tremendous amount of persistence in this sequence. Persistence helps you when you are extrapolating a little bit into the future, but it guarantees that you do not have many degrees of freedom in your data. We all know persistence is the best forecast in economics or geophysics, but persistence can be used without any statistics. You do not need regression equations. For example, have you ever tested just pure persistence against whatever this fancy forecast is? Does your forecast do better than persistence? If so, what are the vari-

ables that are in this that are producing the gain over persistence?

A. We occasionally try a persistence forecast; but, as a general rule, the forecast breaks down after two quarters whereas an econometric forecast would do well for a much longer time.

Q. In other words, these are variables which you do not necessarily have to get out of any past sample, things where you can say "a certain cause has a certain effect," that you also put into your system?

A. Yes.

MR. FRESHMAN: Have you found certain variables that are so powerful that if we could predict the future with a control mechanism that these other variables have, we could get a decent hand on the future and we could control some of these variables? There are many variables in these equations. Are there a certain few of them that are very powerful that would really push the economic future? Can the government control some of these?

A. One of the more powerful is defense spending because it contains a very large fraction of hardware. The change in defense spending is probably the largest of all these multipliers.

MR. IRWIN: Dr. Ward's question seemed to focus on the idea that if the time series data were treated in some different way in the attempt to forecast, then the long range would be more successful but I think I would make a little different point which I thought you were making really, especially toward the end of the paper. You seemed to indicate in the early part of your remarks that this system of simultaneous equations and simulating the future—I should say concentrating on simulating the future—was going to get away from using past trends or an extrapolation of the past trends. But if a set of simultaneous equations which uses data from the historical period from 1947 to 1961 is used to project values into the future, you are assuming and I think you will agree with me—I mean, I got this from your last comments—you were assuming that certain relationships which occurred during the past will continue, so at the end then it doesn't work at the long range and the reason it doesn't work is that the past trends which you are using to extrapolate do not continue. I would phrase my question: do you see any road in the future which, using these simulated techniques, will get away from dependence on the extrapolation of past trends?

A. I would say that we cannot take this body of data from 1947 to 1961, handle it a different way, have different degrees of freedom, measure the auto-

correlation between the variables and in some way make the thing work. It will not work.

DR. WARD: What I said was: it does not make any difference what you do with it; you do not have the data in the sample. The number of degrees of freedom were fixed, and nonstationarity was in it no matter what you did.

A. This is the point I made when I raised the question: will historic information ever be adequate for fitting equations? I was very skeptical about it.

DR. WARD: So am I—even more so.

MR. IRWIN: I do not think getting more data for the period 1947 to 1961, which is more accurate and which is treated in more detail with better degrees of freedom, is going to make any difference.

DR. WARD: That is exactly my point, and I agree with you one hundred percent that there is nothing you can do with that data; but it should be tested to find out how many degrees of freedom it actually has before you proceed to do a lot of things which are not justified.

MR. DONAGHY: This discussion falls in line with my thought when you mentioned the word "control." I am taking "controlling the economy" in a very literal sense. Do you feel there is any bias in your modelling because you are trying to use this modelling technique to establish, as I understood you basically, methods of controlling economy rather than just seeing what it would do?

A. First of all, I should have qualified the word "control." I mean being able to control the macroeconomy, by maintaining full employment, or alternatively by preventing the recurrence of a depression or high level stagnation.

MR. DONAGHY: You mean controlling within a broad band? Would that be it perhaps?

A. Yes. For example, our population is growing such that our labor force and the increase in its productivity causes the potential output of our economy to grow at about four percent a year. It is an object of Federal policy to make sure that as potential GNP grows at four percent, we maintain the total demand for the output also growing at four percent. This is all I mean by control. How is this done? Well, we have a number of techniques such as fiscal policy and monetary policy. When and how to use these techniques derive from adequate forecasting. Perhaps six quarters from now, demand will start to fall while aggregate supply will keep growing at four percent a year. We know there will be unemployment; so by making a forecast, we still have time to ease monetary policy and perhaps cut tax rates until demand meets output.

MR. GOUMAS: You talked about general techniques (basic techniques) available for forecasting business which I think most of us are familiar with. Do you have some typical examples of those that you use in your line of business? For example, have you ever used the Delphi technique?

A. No.

Q. I have an intuitive feeling which has been brought out by others here in their line of questioning that ninety-nine percent of the data that you use is historically based in trending of extrapolations. Do you employ other techniques? In other words do you feel in the future you should use other techniques based on trend extrapolation type?

A. Yes, I would certainly like to get many more surveys of expectations, not only the two major ones that we have: the survey of corporate plans for investment in new plant and equipment and the survey of consumer buying intentions. You brought up Delphi. Those techniques certainly strike me as potentially very important to decision makers throughout the economy.

DR. PATERSON: I would like to try to clarify the point made about the level of sophistication of statistical techniques needed to handle what might be called rather crude historical time series. I am certain in my own mind the real difficulty here is that the state of economic theory is lagging way behind the state of what might be called data information systems. When the condition of economic theory is inadequate, you must handle some very sophisticated problems without an anchor.

DR. WARD: But you see he is using certain techniques which make certain assumptions which he has never shown to be valid; and which I doubt, on the basis of the information presented, are valid. He is using statistics, but they are not appropriate. I think he ought to try something else, but I do not know what.

DR. PATERSON: I do not think you can handle these problems until we know what we are doing.

DR. WARD: He does not have sufficient data—at least, he has not shown that he has sufficient data to use techniques which were developed and have been successfully applied in cases where the data was sufficient and the necessary information was available. Is that disagreeing with what you are saying?

DR. PATERSON: We are probably saying the same thing.

A. Our samples are too small, right?

DR. WARD: The historical data is too small to do the things you are trying to do.

A. Up to now the Federal Government has about four or five or six economic models which are used as checks on independent forecasts—that is, on the forecasts that are made by the older methods. It is only a matter of time before models will become central themselves. The major output of our Treasury equation system is forecasting tax revenue; and as a general rule, the results are pretty good. MAJOR MARTINO: This gives me an opportunity to ask a question. During the last eight years we had much talk of fine tuning the economy and so on. To what extent has the acceptability of your work been affected by the change in Administration? Do the economists brought in by the new Administration accept the theories implicit in your work?

A. I think they agree but I cannot speak for them and can only offer my opinion.

MR. FRESHMAN: According to McCracken's policy here in controlling the economy, both are using the same ideas, so there is the carryover.

MAJOR MARTINO: Is that your understanding?

A. I would say that there are not Democratic economists or Republican economists. It is as simple as that.

MAJOR MARTINO: That is a step forward, when economic science becomes free of politics.

DR. TAEUBER: With the Federal Government having a half dozen economic models; with various universities having others; and with, I suspect most major corporations who have a very major effect at least on a segment of the economy trying to control that segment and trying to use their own forecast to outguess Government planning, do these in fact tend to destroy any predictability in the series? Can we really forecast the future or should the effort from the Government point of view be placed on improving the measurement of what is happening and not have to rely on forecasts for what is happening today?

A. My answer lies, I think, in the discussion of lags. If we know that the economy is going to move into a recession six quarters from now and that lags in monetary policy are six quarters long, then we have to start to ease policy in such a way that by the time all these lags work themselves out we have the desired results.

DR. KNAUSENBERGER: I am sure you know what our eastern neighbors are doing. By nature they are geared to mechanistic models and control schemes. In the convention of the International

Federation of Automatic Controls in London, the leading paper was given by Trapeznikov. He had a prediction scheme which was less detailed and ambitious than yours, using essentially information theory starting out from entropy and introducing control processes. His mathematical relations were essentially exponential and the changing from one curve of the exponential to another was left to uncertainty. Surprisingly, as simple as this model is, at least in his notion, it could explain the behavior of the U.S. economy. I was wondering whether you know this scientist.

A. Yes. I think he came to the University of Pennsylvania several years ago and used the Wharton School Econometric Model. I've forgotten what kind of results he got but I do know that some of his ideas are pretty sophisticated.

DR. RUMMEL: I would like to return to this fascinating discussion of degrees of freedom. Perhaps some of the problem here is the lack of precision with which you discussed the forecasting that had been done using these models. The degrees of freedom are really relevant to stating the level of confidence about what ever projections you want to establish. So the question would be whether you have established confidence bounds about the projections you have made, and whether the actual event has fallen within these confidence bounds or not?

A. Although we have statistical tests, we usually do not bother using them in forecasting. They do not seem to be nearly as good as simply looking at the residuals.

Q. I think I have to agree with Dr. Ward. The degrees of freedom tell you how many independent pieces of information you have in your data. If you do not relate your prediction or your error to your independent pieces of information, then you have no way of gauging, it would seem to me, the effectiveness of the particular model that you are employing.

A. There are better ways of gauging effectiveness.

DR. WARD: I think the degrees of freedom are far more fundamental than just testing what the worth of the forecast is in the end. There is no point in going through a statistical regression scheme where your degrees of freedom are zero. I think this is what the situation is here. When you introduce new variables, the number of which are more than the sum total of the likely degrees of freedom in the data, then it is patently absurd to go through the analysis.

PROF. RUMMEL: Agreed, but I assume he has at least some degrees of freedom there. Even as little

as eight degrees of freedom will allow him to establish confidence bounds on his forecasts.

DR. WARD: He cannot use nine variables.

PROF. RUMMEL: Let us allow him eight degrees of freedom, or even five degrees of freedom and ask whether the bounds have been established.

A. Let me clear up some of the terminology. We are talking about the independence of successive observations and that there should be zero correlation between the successive observations. . .

DR. WARD: The correlation has to be less than one, and the lower it is the better off you are. If you compute, for example, one minus the correlation squared, this would give you the fractional amount of independent information in successive observations. For example, if the autocorrelation were 0.9, the square would be 0.8; 20 percent of each observation would be independent of the previous one. And part of this independence, as I stated in your case and other cases, is due merely to errors in measurement. So you have less than 20 percent real independent information. The point is that the absolute maximum of independent data with a 0.9 correlation is less than 20 percent. You take your total number of values and divide by five or more to find the total number of pieces of independent information that you have. I am sure you have a very persistent time series because when we have a depression, it lasts for a while, and when we have a nice inflation, that lasts too. If you have a one lag autocorrelation of 0.95, I would be absolutely amazed. That induces a tremendous one lag autocorrelation in your data. It is probably the most marked thing in your time series; and when you have that situation, you have to be very, very careful what you do with the analysis; extremely so.

A. What you have been saying appears in all the beginning textbooks on economic theory.

Q. And that is from Chapter 2 on.

A. Right. Nevertheless, what does one do?

Q. O.K. Now, that is getting back to the basic point. You do not use something which is not applicable. If you find that the procedures that you are using are not justified, then you do not use them. I would be much happier if you used something that "I dreamed of this last night," because then you would not be trying to cloak something which was the best you could do under the circumstances in a fancy disguise. I am not saying I could do any better, but at least you would not be cloaking it in something which has worked in other cases. I think that is the thrust of my objection. I am a little bit

unhappy that you are using something which applies in a lot of other cases and has been very successful, (e.g., the physical sciences), in a case where I am really uncertain that it works. And by your own admission, you have not tested it to see whether it works. Number one, you should test whether your data have enough degrees of freedom for the methods you want to apply; and if not, you have got to use something else.

MR. BISCI: This is all very fine theoretically, but how well have your predictions—you said you had a base period in which you collected data up to 1961 and then you applied it for 1962 up to 1969. Did you, in fact, predict this in a reasonably accurate manner?

A. Yes.

Q. Then regardless of the amount of correlation or anything else, the method is useful and a means of getting there. If it is useful to get the end product, then go on using it the way you are using it and refining it as you go along.

A. The output of our ordinary models are superior to any other forecasting system we have now.

COL. HAIDLER: You mentioned that your equations stand up well for six to eight quarters. Now this is about your forecast horizon, I believe?

A. Right.

Q. How far ahead have you attempted to go?

A. The econometric model used by the Department of Commerce is being run to the year 2000. Of course they are laughing about it, but they are doing it.

Q. If we look ahead to the year 2000 when we are talking about double the population that we have today in the United States, one might say, "There is the usual linear progression on some things." Is there a method or is there a means within your method of feeding in changes and social attitudes which may be non-linear?

A. No. This again is one of the weak links to other social systems, the other sub-systems. We do not know, for example, how attitudes toward leisure are going to affect the desire of people to work. It may well be that the number of people working will decline very sharply by the year 2000. We have no way of predicting peoples' attitudes toward work and leisure.

MR. IRWIN: You mentioned consumer anticipation surveys. Have you made any attempts to apply that kind of technique over the long range? I gather that you have talked about it and thought about it since it comes to the fore very quickly. Are there any plans for the Treasury to do something about it?

A. No. This is more the work of the Department of Commerce. For example, they might develop some modification of the Delphi technique for getting anticipations about the year 2000.

MR. IRWIN: Mr. McNeil in the consumer anticipation program, is right down the hall from me. He is not going to do anything like that, you know, in a dramatic kind of long range thing, unless somebody comes up and orders it. So I think the responsibility is not on census to come up with this concept. The census will run the survey and will do it right, I think. By the way, I have to call a point of order on population. It will not be doubled in 2000. It is two hundred million now and a reasonable forecast is three hundred million in 2000.

COL. HAIDLER: May I make a comment? I believe the UN is predicting twice the present population and this is also Hauser's estimates at the University of Chicago, Population Research Center, and these may be rather conservative.

MR. IRWIN: As a matter of fact, Hauser is predicting about three hundred million.

COL. HAIDLER: As I recall, he did not about two weeks ago in a prepared talk at the Air Force Academy.

MR. LINSTONE: We seem to be wandering between two year extrapolation and the year 2000. I would certainly have very strong reservations about anything more than a few years ahead where we have in fact—I don't even know for example if GNP, if you should consider the U.S. is approaching a post-industrial society—I doubt whether GNP is even the correct measure. But getting back to your statement about this being a superior forecast to others, could you explain this more? In what sense is it superior, for what year?

A. The studies that have been made of forecasts of the economy over the last several years—some of these surveys are put out by the Federal Reserve Bank of Richmond and the University of Pennsylvania all point to econometric results being superior to alternative methods of forecasting. One of the reasons is that with the older techniques feedback cannot be taken into account. I should have included some of these studies of forecasting results.

Q. To what degree are all of these good for? A two year period as opposed to a five year period?

A. For longer periods, it is obvious that you have to fall back on the older techniques of forecasting because the forecast horizon of the complete simultaneous equation is simply too short. It is like six to twelve quarters.

DR. SLAFKOSKY: I am going to try and wrap this

up for you, Mr. Bennett. You have been sort of itching to get away, and I do not blame you; but I would like to point out that Dr. Ward indicated that one might question your statistical techniques—i.e., their validity. I think Dr. Paterson, along with Mr. Irwin, stated that irrespective of this statistical validity there are other relevant factors missing. Dr. Paterson indicated that maybe it is economic theory which is lagging behind. Mr. Goumas suggested that in any kind of long range forecasting one might produce better results or do a more effective job if one ascertained what new factors or influences might, in fact, be affecting the future—factors or influences which one could not get out of the past. The most that we can ascribe to your forecasting (and this is not criticism—it is, I think, an assessment of what has been said here) is that it is short range forecasting. You are talking forecasting that uses the recent past, even the long past, and assumes the immediate future will resemble it. The “stuff” is in the mill, so to speak, and the mill is grinding away. The best it can do for you is to help you look ahead for maybe six months, eighteen months or two years, which for your purposes is quite fine when it works. For our purposes when we are concerned with long range forecasting, the two year period is nothing. It is, in effect, now. We are seeking techniques which will, in fact, enable us to project five, ten, fifteen, twenty years. What we are really looking for, in addition to the Delphi approach, is some other approaches which could, in fact, be put into a theory based on proper hypotheses. The nature of the hypotheses for a viable and acceptable theory might be obtained, in part, from reviewing the statistics and data of the recent past. In my judgment, at best, these things could only point to the nature or kinds of hypotheses that should provide the fundament for a theory. This forecasting technique should include an approach comparable to the Delphi process. Finally, I think, we must distinguish between long range forecasting of any kind as opposed to what Mr. Bennett has been talking about which, at best, is short range forecasting.

MR. BENNETT: I hope my point comes across that even with very sophisticated economic theory and with large simultaneous equation systems, economists find it very difficult to forecast adequately even a short into the future. As you return to long forecasting in your individual disciplines, consider the problems the economists are having with short term forecasting. You may be fooling yourself with long run forecasting.

Q. I think we have two things to remember here, that I tend to agree with. The first, as a matter of fact, a model is the way to forecast twenty years in advance. I'm not sure this is the case. I think the other thing is very well stated by Dr. Paterson when he pointed out—I don't know whether he is an economist or not but I trust maybe he is—that, you know, economic theory has in fact lagged behind and no matter how good a model you have, if you don't have any good theory behind it, good theory in effect with which to play around in order to understand it, you're not going to go for it.

MAJOR MARTINO: Let me take advantage by asking one more question. You mentioned in the introduction of several changes of the approach to economics in the 20th century, one of which being emphasis on such measures as growth. You may be aware of a recent book entitled “The Cost of Economic Growth” in which the author looks at the macroeconomic picture and charges that certain external diseconomies fail to be counted in, like the sewage disposal, waste disposal. Presumably we are now becoming aware of these. What impact would you see on your work which would come from starting to count these external diseconomies into the alleged growth of the economy?

A. That is a tough question. All that GNP measures is output. It does not measure the damage that we do to our environment at all. The only way this might show up in the future as we try to reduce the damage to our environment is perhaps relative increase in government spending on pollution control.

Q. You would not foresee a change in the models to account for some of these things?

A. No.

Forecasting Military Requirements: A Critical Viewpoint of the Industry's Approach

by N. D. HARRIS*

PREFACE

This paper was prepared while the author was a member of the Special Projects Group, DMS, Inc. Its successful completion is attributable to comments and observations of many DMS clients and associates as interpreted by the author, and does not necessarily represent the views of DMS, Inc. or North American Rockwell Corporation.

Opening Remarks

I think we all recognize that forecasting is not a brand-new discovery now a newly acquired gift of divine guidance. More often than not, it is just plain hard work. On the other hand, we seem quick to forget that the purpose of long-range planning and forecasting is to provide long-range planning guidance. There is no question that forecasters are more at ease discussing the academic side of the issue than becoming involved in the day-to-day task of improving forecasts as effective planning tools. Despite all the material which cites the significant advances we have made in methodology, there still remains a very real question as to the effectiveness of forecasts as devices to communicate ideas.

This problem is especially critical in the defense industry where the technical aspects of program

* Mr. Harris is currently involved in the analysis and communications of military requirements information at North American Rockwell, Columbus Division. His areas of interest include market research and information systems design. He was formally Director of Long Range Aircraft Forecast Programs at DMS, Inc. and headed DMS's "Free World Aircraft Forecast", a study of the non-communist world's civil and military aircraft inventory and production for the seventies. He has also served as a management consultant and has published extensively in the fields of forecasting and market research.

management and the sophistication of the products has far out-paced the executive's ability to keep abreast of change—let alone plan for the future. Despite the recent advances made in the application of economic theory to decision making, despite the involvement of more and more specialists in the planning phase of weapon systems development, and despite newly developed fields of management theory and long-range resource programming, the defense industry forecaster has yet to develop an effective way of communicating his concept of the future to the decision maker. For the most part, he is still using highly personalized techniques and his usefulness is almost totally dependent on his rapport with his management.

Generally, there is no methodology used in the sense that it is repeatable or follows the basic rules of scientific investigation.

Furthermore, to attempt to separate the planning problem from those associated with forecasting is completely unrealistic. Proper or not, the defense industry planner and forecaster are one and the same; and even where separate functions exist, they are forced to resolve their differences through committee action rather than well thought-out investigation.

To some extent this is acceptable. After all, long-range planning and forecasting are not yet purely mechanical. However, this does not excuse the fact that the defense industry is using techniques which are ten to twenty years behind their commercial counterparts.

Thus, rather than discuss specific techniques or how to plot trend lines, I will attempt to present what I choose to call a critical state-of-the-art report on the defense industry's approach to long-range planning and forecasting. I recognize that there is indeed a thin line between honest criticism and what may be called "negative thinking". However, I

hope that I can bring to the surface some of the fundamental problems associated with the defense industry's approach.

Introduction

In considering whether the development of a new weapons program should be pursued, interested companies naturally want to have at least some idea of the program's potential value. In some cases, the question may be: "will additional manpower and facilities be needed to place the company in a better competitive position?" or "should resources be diverted from less attractive programs to the development task?" In still other cases, the problem may be to determine if greater benefits can be achieved by pursuing programs further downstream. Value and benefits in this context involve not just profits but have bearing on the relative position of the company in the future market place. In other words, judgments must be made concerning the potential of current production programs, the desirability of pursuing one program as opposed to another, and the long-term effect such decisions will have on company goals.

To provide the base for these judgments, the defense industry executive calls upon his program managers to project their known business into the future, compares these projections to forecasts of expected business prepared by his sales and marketing departments, evaluates the resulting potentials—and these are usually many—against long-range company goals, and makes decisions about what course of action to follow. With only minor variations, this approach typifies the main thrust of today's defense industry planning—regardless of the company's size or its products.

In many ways this style of planning represents the personalized approach used by most small companies in the commercial market place: decisions are made by the company's chief executive on the basis of inputs solicited from line managers and from his own experience. Short-range planning serves as a way of backing into long-range planning, and forecasting is performed only when major market shifts or crises occur. When the product's life is measured in months and the company is small enough to adjust to change quickly, crisis forecasting may be acceptable. But where the product's life is measured in years and organizational changes affect tens of thousands of people and expenditures

run into the millions, other techniques seem appropriate.

While acknowledging that the forecasting methods used in the commercial industry sector may not work in the defense industry, it would seem reasonable to assume that the weapons producers would have developed their own arsenal of methods to deal with their unique problems. However, such is not the case. Despite the Department of Defense shift to highly developed long-range planning and resource allocation techniques, the defense industry still insists on using methods based more on chance than risk and profit. In short, the sophistication of long-range planning and forecasting in the defense industry is not even remotely related to the size of the company or the sophistication of the products it may be producing.

Forecasting is Implicit in Planning

Business leaders in every industry sector are quick to declare the need for a comprehensive look into the future. In the defense industry, it is accepted as a truism that no company—regardless of its size, products, or relationship in the contracting tier—can afford to plan improperly. Likewise, the popular school of business management tells us that successful business planning demands a complete interpretation, understanding, and dissemination of all the factors which affect the fields in which the company is now active or may enter within the next ten or twenty years. And, since forecasting the future is implicit in planning, the only choice seems to be whether the company's forecast will be made rationally and explicitly, based on best available talent and methods of analysis, or whether the forecast will be transparent and unimaginative. Clearly, the latter approach is unacceptable.

Nevertheless, while everyone is quick to point out that forecasting is essential to the company's future, there are fundamental questions about the quality of the effort needed and, more important, the role such efforts should play in shaping company policy. While we all recognize that these questions have persisted since the recognition of forecasting as a key ingredient in the successful business planning formula, they have increased in geometric proportion during the last decade. In the defense industry the issue is not so much just the reluctance to change, but involves the basic question of the adequacy or inadequacy of today's approach to long-range planning as a useful business tool.

The Defense Industry's Market Place

At the core of the issue is the unique character of the defense industry's market. As pointed out by Peck and Sherer in their analysis of the defense industry:

A market system exists only in a more or less atrophied form for the weapons acquisition process. Even though the weapons industry contains its share of entrepreneurial personalities, the buyer (that is, the military service) generally decides whether a new weapon is needed and in this sense takes the initiative on new products; furthermore, by making progress payments on development contracts and by furnishing government-owned plants and equipment, the buyer finances most development outlays. The seller does not offer a finished product which the buyer can either accept or reject. Rather, the Government pays development costs before it knows what the ultimate performance of the product or its desirability relative to other products will be. The Government can, and frequently does, change, reduce or cancel the project before its completion.

Furthermore . . . the price of a weapon system is not determined by market competition. Instead, the price is largely determined by reimbursement of costs actually experienced plus a fee bargained for in advance.

While contracting procedures have changed (a company can now *lose* money on a contract), the buyer's absolute power has not changed.

These same authors go on to point out that while there is an absence of competitive pricing in the traditional sense, this does not mean an absence of competition. But, this competition is essentially limited to a "closed club": experience is a necessary condition for admittance and experience cannot be gained without production contracts, and production contracts depend on previous contract experience. Major shifts in weapons technology can alter the line-up—the introduction of the missile, for example—but changes are short lived.

Daily Pressures in a Changing Environment

Today's business executive is faced with pressures which are causing him to question the relevance of his past experience and the applicability of traditional problem solving methods. Day after day, he is confronted with an almost endless flood of new information and an ever increasing number of variables which affect his company. In addition to this steady stream of facts and figures, he is being called upon to take a more active role in shaping our society. And he knows that he can no longer rely

as much on the judgments of his subordinates since they too are faced with similar pressures.

Once thought immune to such pressures, the defense industry executive now finds he too is faced with a host of—if not new—at least more complicated problems: methods of selecting new defense programs are becoming more complex; the cost of participation is higher and investment is required over longer periods of time; fewer major programs are being launched and those programs which are eventually selected are becoming technically, if not tactically, obsolete at shorter and shorter intervals; and in addition to all these pressures, defense programs are becoming more and more vulnerable to attacks from sources once thought to be outside of the defense community. The phrase "national defense" has lost its luster. Winning a program competition is not enough. It has to be justified over and over again, and there is just no such thing as a firm program.

More important, the phrase "defense industry" no longer holds the meaning it once did. There was a time—and not too long ago—when you could easily recognize a defense company. It worked almost exclusively for the Department of Defense and usually built aircraft and missiles. Sometimes, the less conspicuous term "aerospace industry" was used. Now, however, things have changed.

Changing Industry Patterns

Most aerospace companies grew up during World War II, nurtured with an almost endless stream of Government money and with an almost divine right of survival. This environment enhanced the position of the pioneers, made up of the industry's founders and a few key designers, who practiced a highly personalized style of management. It is only a slight exaggeration to say that many of these pioneers were more in love with technical challenge than making a profit or providing a base for long-range growth. Furthermore, this personalized style of management contributed little to the preparation of second generation executives and produced relatively weak financial positions. But, the financial community and stockholders, seeing long-term growth prospects and fascinated with the glamour of supersonic aircraft and space, were willing to accept low earnings and returns on assets managed.

In a similar fashion the military services were caught up with the urge to keep up with the latest technological promise. And, even when efforts

were made to improve the approach to making choices between technological alternatives, the specter of the Soviet threat was raised as justification for almost any expenditure.

The revolution brought to the Department of Defense by Secretary McNamara and his "Whiz Kids" changed all that—at least temporarily. His centralized style of management and insistence on formalizing the approach to long-range resource planning shook the defense establishment's foundation. Right or wrong, the overall effect was that there were fewer programs, and those companies which won these programs found that their return in dollars and cents depended upon their ability to meet their technological promises. Profits dipped lower and became, in some cases, substantial losses. Facility commitments were no longer made by the Government and defense companies were encouraged to buy those which they occupied. Even when the Vietnam insurgency became a war and the economy took off on an unparalleled climb, the Department of Defense stuck to its fixed price contract guns.

To some management at least, this situation was intolerable. They managed companies with billion dollar production capability, advanced technology, and facilities and realized that they were being left behind their commercial counterparts who were in the mainstream of a rapidly growing and diversifying economy. But, they had to harness their unique resources and change themselves from limited-outlook, Government-shielded companies to forward-looking, commercially oriented enterprises whose objectives were to be earning growth as well as technological progress.

This transition period has indeed been painful. For example, in 1959 there were 21 companies producing military aircraft. Today, there are only 12. Many have been acquired by or merged with conglomerates and super-corporations, while still others lost their capability to compete in the military market place. For stockholders, the effects have been generally a gain in their investment. For the military services, the results are measured in fewer qualified contractors and a shrinking production base. And for the the defense industry executive, it has meant losing his identity and acquiring the unpleasant prospect of being considered a liability rather than an asset. While these assertions may seem harsh, they are representative of the mood in much of today's defense industry. This is not so much a blanket condemnation of the management of these 12 companies or the more than 20 producers of missiles and spacecraft or the more

than 100 major suppliers and subcontractors, as it is an illustration of the backdrop for today's decision making. Clearly, it is a less-than-optimum atmosphere for well thought out long-range plans.

On the other hand, the defense industry prides itself in its ability to quickly adapt to new technology and shifts in national policy. While this can be debated, the fact remains that the defense industry has demonstrated its ability to accomplish almost anything. However, the individual company, no matter how vast its resources, cannot accomplish everything. The past approach to planning has been to overlook the simple fact that more "paper programs" can be conceived than justified, or even produced. The technical feasibility or the desires of the military service do not constitute a mandate for development: the program's military worth must be justified, resources must be allocated, and, as noted, all important commitments must be made over and over again over a period of years—sometimes in successive administrations.

The Problem of Compatibility in Futures

Here, too, we find a trait unique to the defense industry market. While the Department of Defense—the buyer—is concerned with acquiring a capability for some time in the future, no adequately reported concept of the future exists from which expected requirements can be derived. For instance, the individual military requirement generally infers that justification is made on explicit information, obtained from the best available sources, about the expected threat and about our own projected technological capabilities. But, in reality, military requirements are based on—at best—an implicit integration of future as conceived by their originators.

In still other cases, the originator may be attempting to deceive the enemy into believing we are going to develop some new capability in the hope of causing him to divert resources; or the goal may be to place a requirement into the hopper as a means of generating new research projects or to keep old ones alive a little longer. And, of course, there are instances where the originator may feel obliged to generate a requirement to satisfy some political pressure or to prevent encroachment on roles and missions by another service.

As may be expected, all this creates a great deal of incompatibility between futures—not to mention the problems encountered when trying to forecast

the requirements for a service or for the Department of Defense as a whole. While some of the discrepancies can be resolved by a detailed and honest appraisal of the reasons behind each program, there still remains the fact that once a program is launched, once a piece of paper is written, it tends to gain momentum, attracts its own supporters, and results in a chain of events which are almost irreversible. Even the most subtle of reasons for originating a requirement—deception or new technology—becomes lost.

It is important to remember that the planning style of most of the military and the industry is oriented towards the "engineering method". That is, planners and forecasters take it as axiomatic that explicit procedures, rather than implicit ones, are always desirable in the justification process. Add to this the popular usage of feasibility, effectiveness, and systems studies; and the most fallacious requirement takes on an air of respectability which is more than can be resisted by the most perceptive of observers.

Again, remember, we are talking about a process which often covers years. The original motives are forgotten, and what industry planner will not weaken when confronted with the official paper storm? The alarming truth is that once a military requirement is reported, it is interpreted as being firm, no matter how poor its original justification. No one need to look very far to find both military and industry supporter alike rushing to supply all the analytical documentation and four-color brochures necessary to keep some obviously questionable programs alive—if for no other reason than someone said that they wanted it; therefore, it must be necessary.

The sheer magnitude of the problem becomes apparent when considering that military requirements are probably this nation's most poorly kept secrets. The problem is no longer one of merely coordinating the exchange of informal viewpoints between offices or commands but involves the controlling of a host of eager contractors and military agencies, each with its own capability to generate additional supporting material. It is no wonder that the originator of one particularly shaky requirement remarked, "I never knew I had such a good idea."

Just as dangerous is the tendency to rely on information about present plans for future policy—particularly at the national and interservice levels. To do this is equivalent to saying, "I thought I would need \$200 a month for expenses, but I find that I will only have \$150 a month; therefore, I will

need \$150 a month." This is sound reasoning only to the extent that one can control requirements—in the example, control spending. But when requirements are controlled by the future environment, then the fact that you cannot meet them does not remove them as requirements. In other words, estimates of our future policy and probable assets, when compared to requirements, provide some guide as to our chances of being able to successfully meet needs. If the resulting judgment is that our capabilities are inadequate, there is reason to make allowances in our forecasts for changes in the current shopping list.

In short, the defense industry executive must not only be wary of reportedly firm requirements, he too must consider alternative programs which are not currently apparent or, for that matter, even popular. On the one hand, he has to sort out winners from losers; while on the other, he must make provisions for responding to future needs which are not now under study.

More uncertainty, more choices to be made; and there is nothing to suggest that all these trends— anomalies in the military market place and likelihood of further industry contraction—are going to abate. Clearly, then, if the problems are ones of uncertainty and choice, the solution, or at least the alternatives, should be more apparent when subjected to long-range planning and forecasting.

Common Approaches to Forecasting Requirements

For the school theorist, just the mention of military requirements, long-range planning, and forecasting conjures up visions of rooms full of specialists methodically examining reams of highly classified documents, carefully reducing every bit and piece of information to computer input forms so that the sensitivity of an almost endless number of variations can be fully displayed for analysis. Add to this a rustic setting, enlightened debate, and two typists for every researcher, and we have the classic picture of what everyone believes to be the long-range planner and forecaster's kingdom. Of course, we know this is not true. But in the back of our minds we still nurture the image in hopes that someday it will be realized.

Why? Because we are frustrated by the failure of our superiors to grasp the significance of our art. We spend more time educating than analyzing; even when we believe that everyone in the loop has been brought up-to-date, we find someone has

penetrated our arguments with his own style of persuasion or some unexpected event has altered the assumptions. To attempt to sell a faulty forecast to management, no matter how good the method, is equivalent to trying to shape the future the way we want it. To attempt to repudiate someone's judgment purely on the grounds that they have violated the method is just as foolhardy.

These conflicts between the theoretical and the everyday, the textbook approach and the actual approach, are alone sufficient reason for the executive to lose faith in the value of long-range planning and forecasting. On the other hand, since everyone is talking so much about forecasting, the executive is willing to support it, for if no other reason than that it has "sex appeal". In short, if everyone wants to see nice orderly charts displaying all the potential downstream programs, he will give it to them.

Of course, this leaves the planner and forecaster in the middle: should he prostitute his art and become a chartmaker or should he push for more support, more time, and a key slot in the decision-making loop. The end result of this dilemma is characterized by the current polarization of practitioners.

At one end of the spectrum we have the chief executive's private consultant who offers up reams of carefully worded documentation, all of which surprisingly seems to support his pet programs. To the company forecaster, the consultant is equivalent to a medium calling up ectoplasms and murmuring incantations and mumbo jumbo which cannot be argued with. Moving from the black art, we have a gray area populated by company experts and advisors. Here the style of forecasting is not much different; as before, the practitioner is in the fold of the revered few who have the ear of the chief executive. Maybe he is a retired admiral or general or a member of the past administrations, or he might even be an ex-whiz-kid. He, too, is beyond reproach.

Moving still further we find the day-to-day researcher. As opposed to his counterparts in the inner circle, he must be prepared to react to pressures from every corner. Frequently, his mission is more to perform the planning functions cast off by program managers and salesmen than to be concerned with the complete picture. More than likely he has come up through the ranks and has been passed over as good line-management material. A niche is created for him, or he creates a position for himself. Since we are still in our first generation of forecasters, he is generally self-educated and has created his own

style. Since he represents the largest share of today's defense industry practitioner, it is worthwhile to review some of his favorite approaches.

First, there is the "more of the same" approach. Threatened in the early 1960s, this approach regained popular support due to the Vietnam War. Basically, the current situation, or latest Vietnam War planning assumptions, are projected into the future. While sound as a short-range planning tool, many companies are using this as a base for mid-range and long-range planning. Ignored is the likelihood of post-Vietnam policy and force changes and the impact of obsolete inventories.

Second is the "one for one" method. The key assumption here is if the Soviets develop a weapon, we should develop one just like it. This approach may have some merit. However, the question of whether we should alter our current forces or tactics or whether we should develop some alternative type of weapons is overshadowed by the argument that we should meet the challenge—one for one. In other words, if the Soviets develop a short-range, high performance interceptor, so should we.

In the past this argument has been difficult to counter since most of its advocates are generally the more emotional proponents of new military hardware. This type of reasoning still prevails in much of the decision making apparatus and cannot be ignored as an influencing force. However, programs launched in this manner must be considered questionable and the probability of redirection is high.

A third approach is the "Big Bomber Club." This style draws heavily on the kind of logic previously cited in that its major premise rests on the theory that the methods used in the past (pre-McNamara) are good enough for tomorrow. Since it takes so long to affect a change in our military posture, the "Big Bomber Club" is offered ample opportunity to seek out soft spots or seize on crises to support their viewpoint.

A minor variation of this approach is the "big man" theme. Here the key is that there are certain leaders who determine future requirements. Identify them, find out what they want, and you will have the programs. Of course, there are always questions as to who exactly are the "big men", but just submitting to this type of approach throws the whole forecasting base open to debate and, more often than not, leads to arbitrary decisions.

All of these methods assume that past relationships will continue in the future—cause and effect will be

identical—and, although uncertainty may exist, there will be no quantum changes.

Next we have the “desirements” approach. This approach is very popular with companies who have large field service and sales groups who are in constant contact with the customer. Information from hot tips to black-market lists of programs filter back to various department heads at the home office. The outcome is predictable: a flood of information, most of which is unrelated or of marginal value, and the decision maker has to fight an uphill battle to sort out fact from fantasy.

A more refined variation of this approach is the “shopping list”. Requirements information is filtered through some type of selection process—generally a management committee—and a list of key programs is drawn up for action. Program development schedules and costs are estimated in a similar manner and the ‘shopping list’ becomes the plan.

This approach has one shortcoming which is prevalent throughout the industry—decision making by committee. To begin with, committees are not the best forums for making judgments about the validity of requirements. Generally, they have neither sufficient time nor information to conduct the background work. Furthermore, few line managers are willing (or allowed) to dispute the opinions of senior vice presidents.

While each of these approaches centers on a different theme, there are some obvious and not so obvious similarities. For instance, there is a common undercurrent which supports the old line concept of a few designers and engineers coming up with the ideas and, in turn, selling them to the admirals and generals. Then, too, there is a kind of fatalistic attitude or, “When it’s our turn, we’ll win.” And finally, not one of the approaches uses long-range forecasting as a starting point. If forecasts are made, they are the *end* result—backed into through the bottoms-up approach—not the *beginning*.

Moving still further, we find the “little DoD” approach. This camp is made up of the pro-McNamara club and lives and breathes the jargon of systems analysis. They are knee deep in FYDPs, PCPs, PCRAs, ADOs and a myriad of other very official looking documents, all of which spell out the future in dollars and cents. No one dares question if the bases are comparable for fear of being subjected to a three-hour treatise on how important each document is—and this short course in defense planning and programming is usually accompanied by a trip through some unbelievably complicated

flow charts which purport to pinpoint every minute step in the process from birth to death.

Just remember this: OSD employs easily 10,000 to 15,000 people who are actively involved in the weapons acquisition and planning process. Add to this a like number of service planners, a few thousand consultants, 10 to 20 quasi think-tanks, several hundred defense journalists and lobbyists, sprinkle with a few politicians, and you have the ingredients for a good planning hangover for the largest of defense contractors. In the total defense industry there are probably no more than 2,500 to 3,000 practitioners. In short, each industry planner and forecaster is outnumbered ten-to-one. Even in the best of Westerns, these are poor odds.

The fact is that most industry groups are small—no more than three or four practitioners per division in a large company—and they are usually trying to follow the total DoD process. The odds are now increased to 10,000 to 1 and there is still the problem of keeping abreast of the competition. Clearly, this presents an impossible task. Nevertheless, it is not uncommon to hear someone say this is the Air Force’s position, or DoD said this or that. Even the “little DoD” planners are beaten before they begin; they will never be able to get all the information they need; and even if they had it, they would not have time to organize it or process it.

Basic to all these styles—from that of the advisor at the top to the style of the “little DoD”—is one common weakness: translation into management action and company goals is difficult. To rebut that this is not the mission of the long-range planner and forecaster completely ignores the real world. Of the previously cited 2,500 to 3,000 practitioners, less than 5 percent are occupied full time with text-book type long-range looks. The remaining 95 percent are involved in the complete spectrum of planning tasks.

Side Effects

Regrettably, all of these factors contribute to the recent trend of “position forecasting.” Once considered members of a politically neutral discipline, long-range planners and forecasters are now projecting themselves into national issues and—as to be expected—there has developed a debate over whose concept of the future is right or real. Since there is no universally recognized method and since the methods which are commonly used are highly stylized, planners and forecasters are inclined to

take positions. Once a position is taken, all subsequent efforts are focused on justifying it. Lack of time and the desire to respond to the other camp lead to judgments based more on "gut" feelings than analysis. And, of course, this diminishes the usefulness of the forecaster's output.

At first glance this might be considered healthy expression. After all, who is more qualified to speculate about the future than the forecaster? Likewise, more than one viewpoint, more than one method, might be thought of as healthy. Sometimes this is called objectivity. However, when projecting this kind of expression and objectivity into the company's environment, three or four concepts of the future are more often than not equivalent to zero effectiveness.

This is a rather grim commentary on an industry with expected sales of more than \$550 billion during the next decade. Neither is it a complimentary view of DoD or the military services who are charged with defending this nation. And while criticism is certainly becoming more popular, the problem of better planning, as opposed to just reacting to plans, is still to be intelligently explored by policy makers. The idea of building more prototypes as a way of getting better weapons is not the solution. It just is a symptom of the confusion associated with long-range military planning. Two B-70s or two Skybolts will not solve the problem of changing strategy, nor will longer research and definition cycles. As one industry planner put it, "If we are not able to get a program defined, developed, and into the inventory within one administration, we are doomed to an endless cycle of research and justification."

Truly, we are faced with a crisis which calls for debate beyond the traditional "my model is better than your model" dialogue. Just trotting out pet ideas, no matter what their merit, or engaging in attacks on the executive for his failure to grasp the value of our art is no longer adequate. What is needed is a complete re-evaluation of the state-of-the-art, both academic and actual, and the development of methods which will help the decision maker project his company into the future. However, this does not mean more complexity.

Here lies a real danger. Since engineering knowledge has out-paced management techniques, there is always the temptation to substitute engineering methods for management methods. Good engineers are not necessarily good managers. Furthermore, mixing engineers and economists in an effort to obtain more cost-effective hardware has been shown

to be just as risky. Clearly, surrendering the design of planning methods to the engineer and economists has not made the process more explicit. If anything, response to management's immediate needs has become longer.

Steps Towards Useful Forecasts

Starting with the customer, I think one of the most important contributions which could be made is an "after"-Congressional-Hearings Posture Statement and Defense Budget. Right or wrong, most of the defense industry's short and mid-range planning is directly related to the current DoD budget plan. These industry plans in turn provide the basis for long-range planning. The February Posture Statement and the President's Budget issued in January frequently provide the only information on how force level and procurement issues raised at the service level have been resolved. Since the actions of the Armed Services Committees generally require a revision in the force posture and defense budget, it would seem appropriate for the Department of Defense to make known how they plan to implement the revised budget. This would eliminate the current problem of trying to deal with all the odds and ends which are published during Appropriations Hearings and provide a more realistic industry planning base.

In this same area, individual service force planning presentations would be of immense value. While the current long-range planning briefings for industry are indeed helpful, they are too heavily oriented towards long-range technical desires and almost completely ignore current force level and procurement planning. Again, such briefings should be timed to take advantage of the actual budget as opposed to the budget the services submit in September or October. It would also be desirable if these briefings were more informal or, at least, if more time were allowed for direct questioning. As an adjunct to the briefings, the services should provide some centrally located office which can serve as an industry focal point for force level and procurement planning information.

An additional input which is needed to augment both of these concepts is a more explicit definition of the long-range threat. Current industry practice centers on the generation of scenarios from open source material; and while many state that this is a more objective approach than that used in military circles, the fact remains that official documents

used to justify force levels should be examined for their long-range impact on future programs.

The next major area of improvement hinges on some method of upgrading our current approach to long-range planning and forecasting. In my judgment, this will have to be done on a massive scale if it is to count as a positive contribution. There are already too many small long-range planning short course sweat shops. Far too many enterprises have seized upon the confusion surrounding the problem to make a fast buck. What is needed now is an activity aimed at the practitioner and the executive which centers on the theme of joint learning and involvement in planning problems. Supported by a staff of well-trained and seasoned specialists and a cross section of other arts and sciences, this activity could tackle common industry problems and provide the catalysts for new long-range planning and forecasting techniques. Hopefully, this hypothetical activity would not acquire the attributes of an industry lobby. However, it should not exclude political and social issues from its curriculum.

Another area which needs to be revived is the joint military-industry forecast—such as the Air Force's Project Forecast. Until such an industry activity as just outlined could be established, intense long-range planning programs should be conducted as a means of generating some new concepts of the future. However, as opposed to earlier efforts, politicians and social leaders should be included to provide a better rounded forum. A taxpayer or two might also help to make the scenario more realistic.

A first order of business for one of these "Blue Ribbon" panels could be post-Vietnam planning. While there have been a couple of abortive attempts at attacking this problem, they have not attracted enough of the real policy makers to contribute to our vision. The weakness which appears most fre-

quently is a tendency simply to project current thinking into the future. Most of the policy and programs displayed are carry-overs from the thinking of the late 1950s and early 1960s. While there is always momentum, post-Vietnam planning should not be based just on today's decision making apparatus but on that which will exist five, ten, fifteen years from now. Like it or not, things are changing.

The most important point of all these proposals is that they should create an environment in which the planner and forecaster can think about the problem—openly and honestly. Hopefully, in this manner we can dispose of such false benchmarks as "if the military want it, it is justifiable" and "if it is secret, it is good information." Possibly, we can even prepare ourselves for such problems as conglomerate forecasting and genuine analysis of the role of civil programs in the defense industry.

I am firmly convinced that we are standing on the threshold of a future which, if not completely beyond our comprehension, will at least be extremely difficult to conceive. Our traditional approach to forecasting futures is rapidly becoming obsolete. The question seems to be: can we develop new methods which will allow us to forecast change or will we continue to forecast because of change. What we, and others involved in forecasting in government and industry, should be concerned with is how to make our art more responsive to the decision makers' needs. This may very well mean that we will have to retreat from our images of ivory towers and think-tanks and become intimately involved in the problems of translating plans into action.

In the defense industry, where we have entered a period when making the right decision about the future has become a matter of sheer survival, we have no choice.

DISCUSSION FOLLOWING PAPER GIVEN BY DR. NOLAND HARRIS, NORTH AMERICAN ROCKWELL, COLUMBUS, "FORECASTING MILITARY REQUIREMENTS, AN INDUSTRY VIEWPOINT."

DR. LINSTONE: I must say I was somewhat amazed to hear this presentation at a technical forecasting meeting, because there was no data base given for any of the statements made. I think many of them are either erroneous or superficial. For example, you discussed "the whole industry." I do not know how you define this industry. You suggest that it is be-

hind in the state-of-the-art in forecasting and planning. In the survey Jantsch made—being a European, and not part of the aerospace industry in America, I would say he is not biased—he discussed a number of techniques which were developed in our aerospace industry, such as Honeywell's PAT-TERN and our own MIRAGE work. In the area

of risk analysis, pioneering work is being done in this industry. We are doing some ourselves. We have set up a risk analysis model. Our discussions with people in the risk analysis field indicate that our techniques are not behind the state-of-the-art. In fact we have generated certain new concepts. You mention in your talk the need for a center for long range planning. The need has been expressed in more depth before. Nick Golovin, for example, published a paper on one way to get out of a major difficulty—lack of long range government policy analysis. He proposed setting up an evaluative branch of government which would have the responsibility for this function. We all recognize the need for it. Professor Yehezkel Dror has extensively written on the need in government for policy analysis as a separate discipline, somewhat beyond systems analysis but having some of its features.

You must be very careful in the data you obtain for corporate long range planning purposes. Consider the threat, for example. I think we can wring our hands all we want to; it is not going to do much good in getting the "right" data for long range analysts. The "threat" depends as much on what we do as on what the opponent does. The "threat" is what we as citizens want to consider as a threat. Threat data, unfortunately, has been used as a political tool. It is a very sad commentary on our culture that fear is the greatest stimulant for spending and public support generally. The missile gap discussions in the 1960 presidential campaign and more recently attest to this. If self-delusion in corporate planning is to be avoided, these considerations cannot be ignored.

DR. HARRIS: We are talking a lot about a lot of deception then. Maybe that is the wrong word to use.

DR. LINSTONE: I prefer the word self-delusion. I recall how, for planning purposes, we used to hunt in the old days for "a good threat estimate from the customer," and we found out the customer goes around to you in industry to learn what U.S. industry anticipates. So it is a kind of closed cycle or intellectual incest.

DR. HARRIS: Does it not seem appropriate we make these ambiguities known? Does it not seem worthwhile to explore some alternatives to that type of approach? It was pointed out by Dr. Rummel that soft sciences can contribute greatly to the problem of handling the vagueness surrounding the character of the future threat. Methods exist. But are they adequate, or more properly stated, are they being

used in the threat versus hardware analysis? In theory this type of investigation goes on during the requirements analysis process, but in practice this work is given a back seat during the requirements justification process. The emphasis then shifts to short range political considerations. Why can we not begin to make these judgements a little more honestly?

DR. LINSTONE: This is very laudable and I hope we get a more objective approach on all sides. There is in the enormous defense establishment a profusion of vested interests, each concerned with its own empire. Organizational or institutional obsolescence is the number one villain.

MR. FRESHMAN: I would like to address myself to the part of your talk where you said "alternatives," and I think implicit in more of the discussion here is the recognition that we have to have some better national goals and I will give you an example of some national goals that are accepted by all parties today. Today in economics it is a recognized national goal to have no unemployment greater than four percent. Thus the monetary policy is to reduce inflation and at the same time not have an unemployment rate greater than four percent. The educational goal used to be that the public would support government financed education through high school. Today we are saying, "Now we will support public education through junior colleges," and we are shortly approaching the time both in our state and national policies of education through a four year college. So we have stated national goals implied and accepted mostly by the public and by our politicians. It seems to me that before we are going to do forecasting and long range planning in individual companies and DOD we have got to have some national goals which would delineate the kind of nation we would like to be, keeping in mind the projected attainable world. Then below that is the national foreign policy which provided most of the requirements for the international security policy. These are not big thoughts but we have all been talking around it and we have all been saying that our politicians will not do it. Well, we are the people that could. I mean the public at large could indicate some clearcut definition of what we will support for our national security policy and what we will fund for other policies. If we would even go that far, we would have some delineation of where the forecasts are needed and what kind of plans we needed. Then our individual companies and DOD could fit better into the picture.

MAJOR MARTINO: I would like to add here the man

raising this point is in the planning office at a fairly high level in the Air Force organization dedicated to the advancement of technology. He is saying that what he needs is the same sort of thing you have been saying you need. I think this is significant.

MR. FRESHMAN: I will go a little further and talk about national intelligence policy. Our products divisions tried to secure acceptance of their long range systems (and they have not been proposing very many systems that are long range), let us say, based on a national threat policy. Every time they take it up, it is said, "Who told you that is the threat?" We have not got an officially accepted threat and maybe it is not possible to get it. Some people say it is. But somebody at the highest level in DOD or at the National Security Council level, if we are to do any forecasting, any planning, has to say what, in their estimation, is the DOD long range intelligence threat policy. If you do not say it, we are just going to play ball with each other, up and back. Like it or not, they are going to have to come out and say it. (The Defense Intelligence Agency, at present, is investigating the making and coordination of a long range intelligence forecast.)

MR. SCHNARE: I am from an Air Force Laboratory, the Rocket Propulsion Laboratory. My comments are on long range planning to establish objectives for research and development to be accomplished in the AFSC applied research laboratories and the importance of current and projected enemy capability.

During the past several years the AFSC laboratories have been seeking a method which be can used to establish long range objectives for technology. Because of limited resources, both manpower and funds, it is necessary to limit exploratory and advanced development to those technical efforts most needed to advance Air Force capability. Common goals must be established so that the technology needed will be available on a timely basis. Each AFSC laboratory has responsibility for a technical area such as guidance, propulsion, structures, etc. As the weapon system performance requirement increases, the technology from each laboratory must be advanced simultaneously to provide advanced subsystem capability at the time when a decision is dependent on enemy capability and will probably be responsive to a decision by a potential enemy.

New system capability objectives which are far enough into the future to allow time for applied research laboratories to complete technical programs must be established. Current weapon system

plans and objectives are too short term. They are based on technology which has been proven feasible. In other words, the laboratory has completed their programs on technology needed, and weapon development can proceed on a reasonably assured basis. Consequently, there is no additional technical effort for this system required of the laboratory. Because of the time required, 5 to 10 years, to complete research and development to prove feasibility of a new idea or to advance technology, system capability objectives which use this technology must be at least five, and hopefully, ten and fifteen years in the future, i.e., well beyond the next system being advocated for development.

The planning method which the AFSC laboratories advocate is based on an evolutionary process with allowance made for revolutionary ideas. The Defense Department and USAF missions are defined in the USAF Planning Concepts document. Our relationship to other countries is described and our operational forces are defined. This document is updated annually. Peaceful resolution of conflicting interests is desired but the communist doctrine does not necessarily support this objective; consequently, weapons of all types are required to deal with various levels of conflict.

Our Foreign Technology and intelligence organizations have a good understanding of enemy capability. The performance of their aircraft, missiles, anti-aircraft guns and weapons used by their military forces has been determined with reasonable certainty. Technology reports have been reviewed and improvements in weapons from the first model onward through subsequent models are noted. It is necessary to analyze and determine their motivation for improvement to be able to project their future trends. As our capability in this area improves, our projection of enemy capability becomes more reliable.

Knowing our potential enemies and their weapon capability provides us with the information necessary to plan our next system capability and the following ones; it allows us to establish improvement in performance characteristics necessary and the corresponding improvement in technology needed to achieve this system capability. This method is described as an evolutionary extrapolation of our current operational force which is responsive to current and projected enemy capability.

You made a statement that the aerospace industry can accomplish almost anything. I think this is erroneous. The naive attempts to apply the systems

approach to civil systems (e.g., poverty and ghetto problems) are unfortunate, albeit well meaning. We simply do not understand these "systems" well enough and you just cannot take aerospace engineers and apply the same approach as used in well-structured defense projects. One must get to some fundamental reasons and not be sidetracked by superficial points. In the testimony I gave to a Congressional committee in March, I mentioned four specific dilemmas. You referred to one of them: the long range objectives problem. The fact is that available long range objectives in many ways are unsuitable for the analysis of needs. In Kissinger's words, the United States does not have a foreign policy. Decisions on policy are made when immediate action is required. This creates grave problems for planners in industry; it creates problems for planning in DOD.

The second basic problem is the lack of systems understanding, which was already mentioned.

Another one is the implementation of new needs. We tend to buy and sell what we are comfortable with, rather than what is needed. The establishment—DOD and the aerospace industry—is growing too big and too old (or let us say mature). Whenever I talk to senior personnel, I always say "mature" because the other word makes them very nervous. Size and age lead to rigidity, particularly in the middle layer of the institutional hierarchy. It suffocates with obsolete functions and organizations. Innovation suffers.

These are some of the basic difficulties industry is wrestling with. I wish you had gotten more into these fundamentals. To return to my first point. I would like to know what your data base is for your statements.

MR. HARRIS: I do not think the problem can be reduced to the validity of a data base, or even the existence of one for that matter. As discussed, this is one of the fundamental problems of forecasting and long range planning; at some point in time the interpretation of results becomes a matter of personal judgement, personal feelings; how forecasters and planners go about supporting these judgments can always be questioned. This was amply demonstrated during the previous discussion concerning economic forecasting. The issue is not so much that the techniques do not exist, nor is it a lack of understanding or at least appreciation of their value. The point is that too much of the forecasting and long range planning being practiced in the industry today is not remotely approaching the degree of sophistication cited in your examples. I would sug-

gest that even if by some universal act we were to change over to some of these advanced techniques, we would still be confronted with the problem of translating the results into company goals.

The reasons for this are: (1) the nature of the defense industry's planning cycle—a cycle which involves from seven to fifteen years for a major weapon system, (2) as you point out, the age of the institution and (3) the simple fact that the forecaster and planner are not being rewarded when they employ them. Like it or not, we have to acknowledge that many of these practitioners have been pressed into service from other scientific and engineering skills and have evolved their own brand of forecasting and planning to satisfy their needs. This was the goal of my paper: to report to you my observations on the degree of personalization being practiced at the working level. I realized that this would be controversial at the outset, but it seems to me the most important contribution a group such as this can make is to develop techniques and information which can be used by the people who are actually engaged in forecasting and planning. I think it is imperative that we recognize that they are under almost constant pressure: they are dealing with a wide variety of problems ranging from following and predicting the outcome of DOD programs to interpreting social change. Their label may vary from market researcher to long range planner, but their problems are the same: communicating a concept of the future to management. In my judgement we need to begin to work on basic methods and agreements, naive as they may be, which will allow us to plan with more confidence. Put another way, the main theme of the approaches to forecasting military requirements I outlined centered on sorting out good programs from marginal ones. Of course your company is actively engaged in pursuing only good programs. That is a very naive approach. We all know it; but nevertheless, nine out of ten companies define future business in this manner. If they do not do it at the corporate level, they will at least do it at the division level. This is the problem that concern me.

DR. LINSTONE: The steps that you indicated will not give you an answer. You are looking for someone to give you this kind of information on a silver platter. First of all, I know that most companies are not as naive as you think they are. Second, the information that we are really concerned about is simply not there and I do not think just by setting up an agency you are going to get it. In many ways, it is as if you were trying to design an air defense

system for England in World War II when you do not have some of the basic fundamentals of energy propagation under your belt.

MR. HARRIS: I will grant you that all companies are not that naive, but we have to honestly admit that too much of the industry still approaches the problem in this manner. In any case, even when companies employ better forecasting and planning systems, there is still the problem of getting executives and decision makers, talking with forecasters and planners. I may be painting an overly bleak picture but it appears to me to be the only way we can get the participants thinking about the problem. Hopefully, in this manner we can begin to develop a better understanding of the future. Regarding future priorities, social needs as opposed to defense needs, we know that there are certain statements we can make and statements we cannot make. Regrettable as it may be, there is that ever present tendency to make statements which will be popular. This certainly is not a problem unique to the defense industry nor our DOD counterparts or the Congressman who is involved in the review and authorization of the defense expenditures. We all work for one master or another and we are all cautious about making unpopular statements or confusing the issue with more than one alternative. Unsatisfactory as it may be, we are a "yes" or "no" oriented society. We like simple answers and action—to suggest that a range of alternatives be explored almost always leads to no action. I have yet to see a defense program presented to Congress which was not billed as the ultimate solution: No one discusses the fact that a "hedge" system is also under development.

MR. SIMON: Most of the questions to yourself and the first speaker were so all-encompassing that to answer them would require almost a complete repeat of your discussion. So for a little relief, let me comment on a relatively minor point which you brought up. You stated that businesses in any generic area acted as a sort of "closed club." That is, if a company had not previously been in a particular business and had no experience, they could not get in it. This calls to mind the early days of the transistor development when every electronic manufacturer wanted to get into the transistor business. They did not have time to start from ground "zero," so they just located a relatively small "going" concern, bought them out and by acquisition were in the transistor business. This is true also in the defense industry at least as far as training devices are concerned. There have been instances where simulator manufacturer wanted to get in a par-

ticular line of simulation in which his company had no background experience. They could not bid in although they attempted to do so many times but were never considered. They were always told: "You are unresponsive; you have had no experience." So, they went the acquisition route. They bought a company that had existing contracts with the Naval Training Device Center and became accepted members of the hiterto "closed club." So, in essence, I feel that the point you made: "If you are not in, you can't get in" is not particularly valid. MR. HARRIS: This approach to penetrating the defense market is a common one and still is about the only way to overcome the "closed club" philosophy. However, probably more significant is recent trend of military aircraft and missile producers merging with non-defense companies. In short, the contraction we have witnessed in recent years is going to continue.

MR. IRWIN: I would like you to clarify what you think the role of the individual company in this defense industry business ought to be. In your prepared remarks, you seemed to indicate that the problem was to find out what DOD wanted and this was very difficult because they changed their plans. During your extemporaneous remarks, you talked about the difficulty of the forecaster and the working executive getting together, which sounds like an internal company problem. What do you think a company ought to do as far as taking a role in forecasting military requirements?

MR. HARRIS: I would suggest that the internal company problem you alluded to is not unique to any single company and is, in fact, a mirror image of the problems faced by every participant in the defense industry. Regarding an approach to forecasting military requirements, I would suggest the most important step that can be taken today involves the legitimizing of the company's forecasting and planning function and in turn organizing forecasting and planning resources—people, information and methods of analysis—into a regular, planned and coordinated attack on the problem. This is not any different from the approach used by the DOD or service long range planner. The basic theme of this activity ought to center on the weighing of reported plans against expectations. In other words, there are those in this room who are privileged to have access to detailed future DOD plans. There are others who, for one reason or another, do not have that opportunity. The first group has to avoid becoming captives of reported plans or "If it is secret, it is right." The second group, on the other

hand, need not throw up their hands in dismay: there are ways to put more order into the process. But the long range solution in my estimations—if you will permit me to quote from Kahn's book, *The Year 2000*—is to just think about the problem.

Properly organized, armed with basic methods of analysis and instructed to incorporate a wider perspective in future planning, these same forecasters and planners should be well qualified to evaluate a wider range of alternatives—including social issues and the value of acquisitions, mergers, diversification plans, organizational realignment, etc. This is not to diminish the importance of long range technological forecasting; but by expanding the scope of the planning function and by making it a recognized part of the decision making apparatus, the chances of anticipating change are increased and the fear of being surprised is diminished. The outcome should be more confidence in plans and planning with confidence.

DR. RUMMEL: There is a methodological bias in the defense industry—a bias that the techniques and methods that have been successfully applied to hardware problems can be equally applied to solve our social problems. At the same time, the industry and DOD should also be complimented. Let me give some perspective on this. I think that if you contrast the work of the defense industries on the area of soft science—that is, applying science to understanding international behavior and environments with that in the universities, you would find much pioneering work has been done by defense industries. I think the rigidity, the resistance to new techniques and problems in the defense industries, can also be found in the academic disciplines. Indeed, I think the pressures against change are greater in the universities. If one wants to look for large scale creative work attempting to meet the socioeconomic and politico-military problems of the defense analyst and policy maker, he has to look largely outside of the academic discipline to both government agencies and the defense industries.

DR. LINSTONE: I am a little bit surprised that you do not mention your own company's very interesting work in using input-output analysis as a forecasting tool. North American has made a real contribution here. I would, incidentally, suggest that you get hold of George Steiner's new book, *Top Management Planning*. It outlines the planning techniques used in many companies. I would like to point to the problem of self-delusion in corporate planning.

MR. SCHNARE: Industry can and does contribute to this planning. Each fall they are advised of potential future system capabilities and they respond with results of their studies and forecasts of advance technology. Many corporations have found the exercise profitable in bringing various corporation divisions together and consolidating various programs such as Independent Research and Development, product improvement, exploratory development, etc., toward common goals eliminating duplication within the corporation.

Finally, the AFRPL with the assistance of SAMSO and ASD has, during the past year, used this planning method to lay out long range System Capability Objectives (current, next generation and the following) which were used to define our technical efforts for Fiscal Year 1970 and 1971. These System Capability Objectives were prioritized and the priority list was used in determining which efforts would be funded. Those items on the bottom of the list were simply deleted from the program. The complete list had 23 System Capability Objectives of which technical effort for 13 was funded. The list was limited to those systems which were needed most. This information will be furnished to industry and their suggestions and comments requested.

MR. HARRIS: I could not agree with you more. But while the services have some excellent programs for reporting on long-range needs, they are still too broad in scope. As you pointed out, there still remains the problem of sorting out good programs from marginal ones and this requires heavy reliance on highly personalized lines of communications. As may be expected, this personalized approach leads to speculation and indecision. These conditions, of course, lead to duplication of effort. On the other hand this situation is not unique to industry or more specifically, the divisions which make up a corporation. There is duplication between services and duplication within services. As you indicated, your action is intended to overcome some of this. There still remains, however, a very serious question as to whose list of requirements industry will act on: SAMSO's or the Air Force. This brings me to another point you raised: The Department of Defense's role. I would suggest that the question I raised a few moments ago would have been easy to answer a few years ago. We would have accepted your list without question. But today, things are different. Not only has OSD emerged as a real force in determining which programs reach hardware, but the executive and legislative branches of govern-

ment are taking a more active role—they are saying not just how much money should be made available but which programs should be funded, how many should be built and how they should be deployed. No one is going to deny that this is the right or the duty of the President and the Congress, but I think many in industry and the services would like to ignore their impact and continue with the task of solving technical problems. The result is that we have to translate annual political statements into the make-up of the budget and in turn, into long-range

threat estimates. Again we are faced with evaluating the changes against technology which takes years to develop. As you can imagine, such decisions can mean life or death for a company which has spent millions of dollars developing technological capability. A few companies making the wrong decisions and you have lost the industry's ability to build the kind of hardware needed to keep your military forces creditable. There obviously ought to be a better way of making these decisions.

Election Night Forecasting

RICHARD C. TAEUBER*, Phoenix Systems International

Formerly with Leasco Systems & Research, and associated with the ABC election night efforts
in 1964, 1966 and 1968

Election forecasting can be broken down into three sub-categories: long range, which includes examination of historical perspectives and trends; medium to short range, which refers to the various opinion polls and their analyses which appear in various media over the days, weeks and months prior to an election; and finally the election night forecasting which is based on partial returns to permit various media (principally the three major TV-radio news networks) to inform their viewers/listeners/readers of the projected outcome prior to a semi-official complete count result which might be known 12 to 24 hours hence. This discussion will concern itself with the latter sub-category, and

* Dr. Richard C. Taeuber is Executive Vice-President of Phoenix Systems International, New York, New York. He received his PhD in experimental statistics from North Carolina State University and his B.A. in economics and mathematics from Swarthmore College.

Dr. Taeuber has done consulting with Phoenix Systems International, Leasco Systems & Research Corporation, EBS Management Consultants Incorporated, and C-E-I-R, Inc. Some of his more significant experiences include Elections '68 project for ABC TV, senior team member doing preliminary design of the National Highway Accident and Injury Analysis Center, consultant statistician to the Missouri Pacific Railroad for design and implementation of a freight traffic survey (also acted as expert witness concerning sampling methods), and consultant statistician to the American Research Bureau for the design and implementation of the NBC ARBITRON TV ratings sample. In addition, he has been a consultant to a variety of other private and governmental agencies and has taught graduate level statistical analysis.

The Elections '68 project forms the basis for Dr. Taeuber's presentation at the Third Symposium. On this project he was responsible for the projection model for early forecast of election night results, the design of the information output system to insure maximum flexibility in the ability to provide information where it is needed on the most current possible basis, and the development of parametric data files and test and rehearsal data.

explore the methodologies and the problems encountered.

Election night forecasting is a rather unique exercise. Elections, their forecasting and immediate analysis, happen all in one day. The speed of required reaction is also atypical, especially of business applications, for the projection models are updated and recycled based on new data every two to six minutes, rather than daily, weekly, or even monthly. (It might also be noted that in 1968 there were 106 statewide contests projected—51 presidential, 34 senatorial, and 21 gubernatorial contests.) But this response time does not imply any similarity to process control, for there is too much human involvement throughout the system for any real automation of the decision process. On the input side, the data flowing into the forecasting system is aggregated human responses, a rather poor, unstable data source. On the decision making side, much subjective information and intelligence interacts with the objectivity of the projection model to produce the projections: winners and estimated percentages, that are aired.

In a sense the election night forecasting on the TV networks is a competition, for it is somewhat assumed by the masses that the earlier the set of calls is made, the better the news department that is making them. Much is made over the remarkably small proportion, if any, of incorrect projections that are made during the course of the evening, and usually little attention is paid to the successes once the night (and portions of the next day) have faded into history. As a for instance, during the most recent election night (November 5-6, 1968) ABC was first with six of the eight major-state presidential calls, and thus also with the overall national call. But in reviews of the early portions of the coverage that were published by columnists the

next day, ABC was panned for being consistently slow in their calls which were aired during early evening on November 5. What is the explanation for this discrepancy between early slowness and late success? Are there differences in the forecasting models that are used by the three networks? Are there differences in the functioning of the various decision groups that make the calls? Are there differences in the guidelines that are imposed by the networks on the release of the projections over the air? The answer to all these questions is yes, and they will be explored later.

Looking back to the portion of the evening's activities for which they claim "victory", ABC News, in an internal "Backgrounder"; and Leasco Systems & Research the subcontractor who produced and programmed the projection model that ABC used, in an ad that appeared in five major newspapers on November 10-12; gave the following story of what happened that night.

"The Presidential election of 1968 will go down in history books as one of the closest of all time, rivaled in this century only by the Kennedy victory in 1960. But Election Night in 1968 was more dramatic than 1960 because the advances of electronic journalism in covering elections made the results known soon after the dawn following election night—if you were watching or listening to ABC News. . . The polls were closed. Millions of people sat avidly watching the election returns in one of the closest Presidential contests in history. But by morning there was still no winner. The 1968 election had turned into the most widely viewed cliffhanger in history.

Nevertheless, television was bringing the public the fastest, most accurate projections ever. To do this networks employed an elaborate and intricate computer system. Providing tabulated returns to all the networks and wire services was a central computer system (which contributed to the most publicized difficulties and delays). In addition, each network had its own computer system to provide information for analysis and projections.

The responsibility for creating and coordinating ABC's computer system had come to us. Leasco Systems & Research. . . When we set up the system we took great pains to cover every part of it in the event of unexpected problems on the big day. However there was one aspect over which we had no control—the central compilation of returns. Since this information (along with key precinct data) was crucial to the projections, we decided to be ready for any eventuality. We built a fail-safe into

the system. Now only did we catch mistakes, but we were able to react to the erroneous data, identify specifically what was wrong and assimilate other data to replace it without delays. We had set up alternative procedures in advance that allowed ABC to continue making projections.

It was an incredibly complex and demanding application. Yet, in spite of those difficulties on the input side, our system functioned continuously throughout the operation.

The result was that ABC was able to announce Nixon's victory with confidence at 8:19 a.m., Wednesday—two hours and sixteen minutes before one network, and two hours and twenty-six minutes before the other."

To show that there is still a human element in these projections, in those final minutes before the call that was released over the air at 8:19—a projection that Nixon would carry the State of Illinois, and thus would win sufficient electoral votes to capture the Presidency—one member of the group making the Presidential decisions arose, put on his suit coat which had been shed many hours earlier, announced to the others that he felt they should be suitably dressed to make a decision of this importance to the country, and said he then felt that the call should be made for Nixon.

Background

From the beginnings of elections politicians, political scientists, and in fact anyone with an interest and an access to partial information has wanted to use that partial information to get an early indication as to how the election might go. Various members of this interested group have their pet precincts, wards, cities, counties or states (in the Presidential contest—"As Maine goes, so goes the nation!?"") which they used and still use to extrapolate to the overall final outcome. Radio commentators used vote tabulations supplied by the wire services (AP and UPI) to bring subjective forecasts of the outcome. Beginning in 1952 televised election eve programs, complete with computers for tabulation and projection got into the act. From that beginning in 1952 computers have become at once both the core of the projections, at least the capability to do the projections, and the bogey man for the viewer who blames them for everything that may be going either wrong, or not as he individually wishes things to go. Much of the commentary goes along the vein that the computers

are actually making the decisions and projections, but as you all know, the computers must use techniques and procedures which are supplied from the minds of mortal man, to aid and assist the analytic staffs of the organizations doing the projecting.

Election evening forecasting involvement gives working (and inoperative) computers possibly their greatest single public exposure. Their use in assisting in election night forecasting and analysis is an application unlike no other in the computer business in that this application calls for a fully reliable operation, with no postponment of election day possible to permit a bit more time to find a program bug. No "T minus two days and holding" is available to save the projection teams/systems if last minute difficulties occur.

With this heavy involvement of computers in election coverage, and with the general public somewhat in awe of these electronic marvels, it should be noted that computers do not vote, they do not determine the outcome of any race, they do not merely total votes, nor do they guess at the projected results. Computers do assist the human analytical/decision staff in comparing present happenings with past occurrences, and projecting the final outcome of this particular election. Technically speaking, if the coverage was of but a single race, there would be very little need for computer involvement. However, with the network analytical staffs, in 1968, trying to keep track of what is happening in the Presidential contest in 50 states and the District of Columbia, in senatorial races in 34 states, gubernatorial races in 21 states, and in 435 congressional districts, the need for electronic assistance is evident. In 1970, 51 Presidential contests are deleted from the load, but there will be 35 senatorial contests, and also 35 gubernatorial contests—or 70 statewide contests instead of 106.

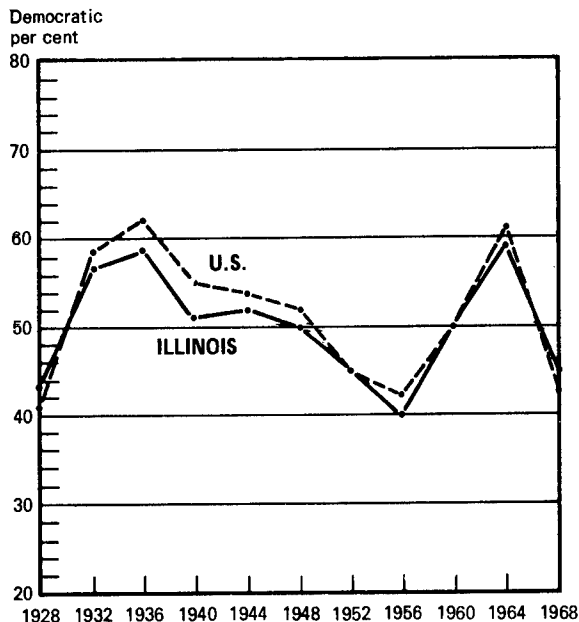
Earlier it was mentioned that there are three levels of election forecasting: long range or historical; medium range or opinion polling, and the short range or election night forecasting. The techniques that will be discussed are those used for the election night forecasting, but this forecasting uses both the others to some extent. Although the techniques used by the three networks differ in extent, they all make at least some use of prior or historic information, either directly in the model or in the selection of sample sizes. Historic patterns can be indicative of the potential closeness of the race, and there is a common tendency to concentrate sample points within states that are likely to be close, and thus more demanding of information to permit a decision

as to a likely winner. Dr. Louis H. Bean, in his book *How America Votes* (Scarecrow Press, Metuchen, New Jersey) has a series of historical comparisons: individual states versus the U.S., states versus neighboring states, and states versus a major city. And from his book one can readily see wherefrom comes the statement that as Maine goes, so goes the Nation—or as a good many states go, so go the Nation. The explanation of the statement is that it refers, not to absolutes, but to swings or deviations from a prior election: as Maine swings, so swings the Nation.

The accompanying figure (Figure III-1) shows the patterns for the Presidential race in Illinois and Indiana, both compared against the US. If the National contest were settled by direct popular vote, whereby the National vote percentage would have direct meaning, then one could come quite close to the National outcome merely by concentrating study/polling effort in the State of Illinois. Alternately, one could poll in Indiana, add about 3% to the Indiana estimates, and have an estimate for the U.S. which would be quite good, assuming that the historical consistency holds up when extrapolated into the future.

As regards the medium range forecasts of election results, they do enter directly into the forecasting model in providing a base from which the election night forecasts are made, but this will be discussed in more detail later. The one comment that will be mentioned here is in regard to the controversy over the effect that election forecasts have on the actual outcome. Based on two voters that were approached by Rockefeller supporters at 7:30 p.m. on the night of the 1964 Republican Presidential Primary, or seven minutes after one network had indicated that Goldwater was the indicated winner in the contest (a call made 37 minutes prior to the close of the machine precincts in the state, 23 minutes after paper precincts closed), voters who indicated that they then did not intend to vote because the issue had already been decided; a controversy has been raging over the effect of the early televised calls. Studies made in 1964 also in 1966 have indicated that any effect derives from the fact that the TV viewer of election programs is both more interested in politics and more likely to vote. There is no indicated effect on the actual turnout other than this correlation. But maybe it is the medium range forecasts which have the effect on voting patterns which should be of concern to our political system. Here again there are no definitive findings from studies or discussions—for it seems to be somewhat of a

ILLINOIS AND THE U.S.



INDIANA AND THE U.S.

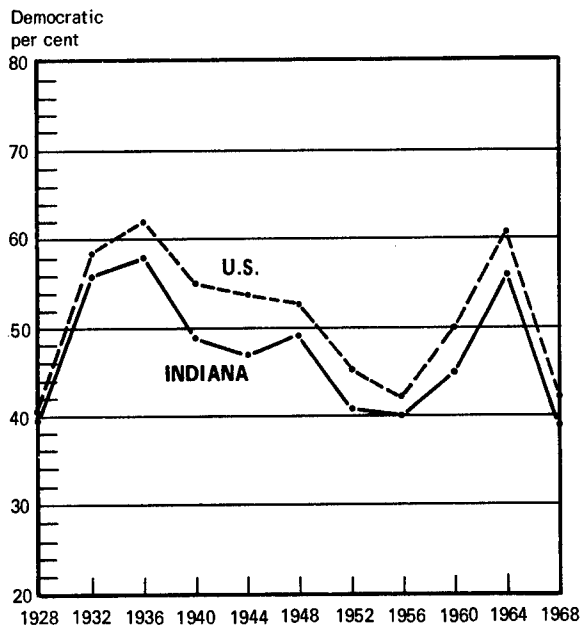


FIGURE III-1

cyclic argument as to whether the polls reflect or help mold opinion. Both areas are probably in need of continuing study, though at whose expense?

News Election Service

A basic element in election forecasting is the availability of new vote data as it reports and is tabulated after the polls close. Prior to 1964 for their election night programs, each of the networks set up its own vote tabulation system so as to be able to report the progress of the election. In the comparisons of the three networks performance, it was assumed, at least to some extent, that the network with the highest posted total vote figures was doing the best job, both in forecasting and in general election coverage. Add to this the fact that the two wire services were also running tabulation systems, and the whole mess would seem to have gotten out of hand. In the California primary in June of 1964, with 32,000 precincts in the state, each of the three networks (ABC, CBS, NBC), and both of the wire services (AP, UPI), had reporters in most, if not nearly all of the precincts. This meant that collectively they had better than 100,000 people out in the field reporting votes to five separate tabulation systems. Out of this fiasco came a proposal

that since the whole operation was getting out of hand, come 1966 they ought to do the tabulation on a pooled basis. Someone else suggested why wait? They should proceed immediately. Thus was born the Network Election Service whose purpose was to tabulate the raw vote on election night in November, 1964, with all five members agreeing that any vote totals relayed to viewers would be those released by NES. In 1965 the organization was made permanent, the name changed to News from Network, and Justice Department clearance secured so that the fledgling organization would not be in violation of anti-trust regulations. In 1966 NES conducted a pilot operation to tabulate 11 Western states by computer, and in 1968 the whole operation was switched to a computer system for tabulation and dissemination to the members (and 11 additional subscribers).

NES functions as indicated in Figure III-2, with the raw vote reports flowing in from precincts and counties, and the summarized vote data flowing out to the members and subscribers by either teletype or direct computer-to-computer transmission. In 1968 NES had reporters in most of the precincts in most states—missing were precincts in counties which have a centralized tabulation system, and a few isolated or rural precincts. The precinct totals were phoned in to the central system, presumably

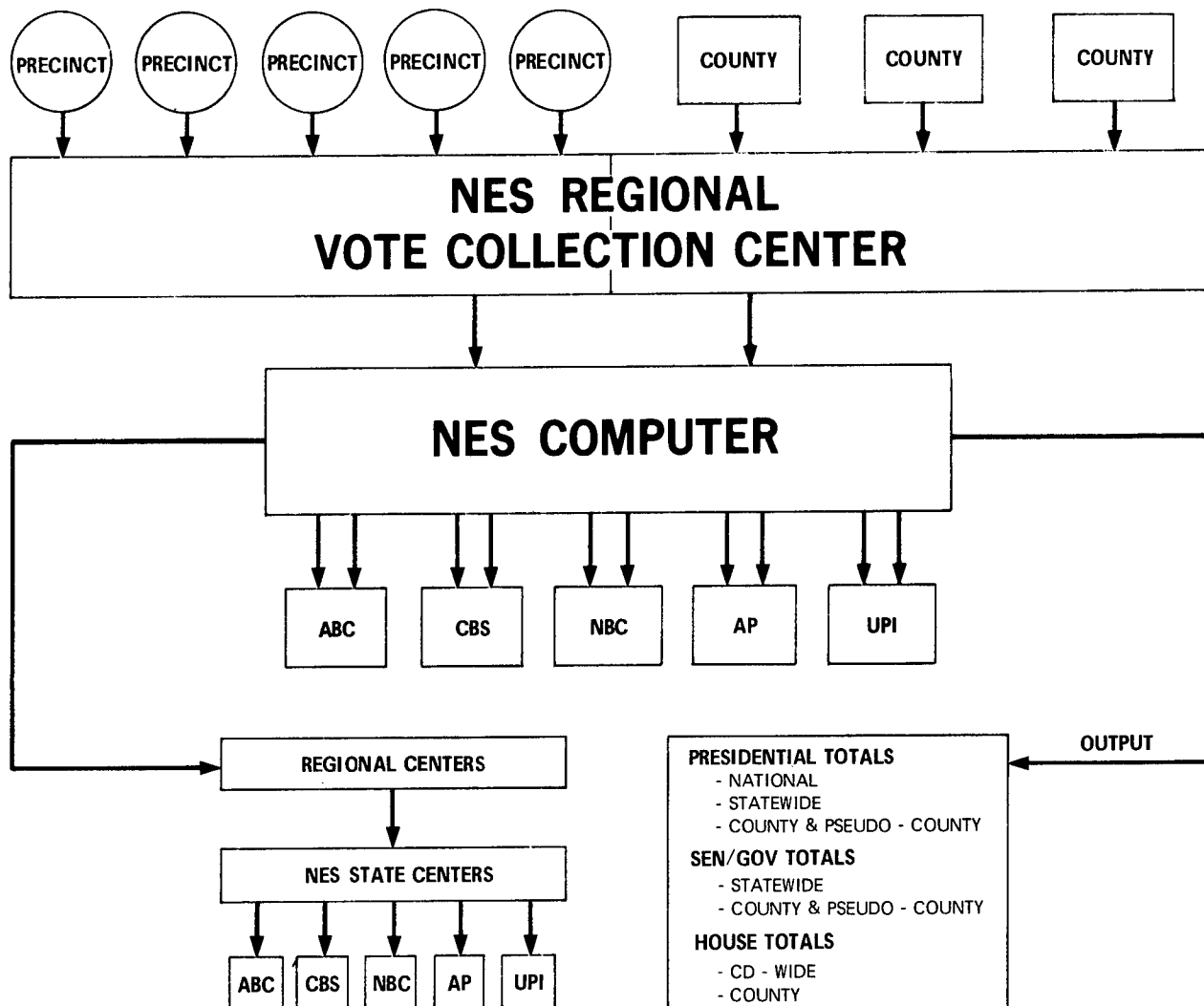


FIGURE III-2

faster than if they reported through the official tabulation system, and then were summarized on a county basis. County returns were also secured, to provide a check on the precinct accumulations, and also to provide a more complete figure late in the evening, when the number of precincts that had reported officially might well surpass the number that actually did report to NES directly. Within the NES system, the votes were accumulated on a county, congressional district, state, and national basis as appropriate, and the summarized data reported back to the members and subscribers on a periodic basis—with the periodicity dependent on the extent to which the data might have changed, and the importance of the particular data element. The output of this system, according to the ground rules under which NES

was established, are to be the only figures released for on the air viewing, thus there should be no difference in the total vote displayed on any of the television networks. The members are free to do with the data whatever they desire in the way of analysis and projection, and these areas are thus the areas of difference between the three networks.

On Figure III-2, the term pseudo-county is used to designate certain cities in various states which are tabulated separately from the county of which they are a part. This is a requirement imposed on the NES system by the members to provide information for analytical purposes. Such pseudo-counties, e.g., Chicago separately tabulated from the rest of Cook County, are treated just as any other county, and are given a numerical code at the end of the al-

phabetic listing of regular counties. As a note, Cook County would be tabulated as a single entity under its normal code using the county reporting system; Chicago and "the rest of Cook County" would be tabulated as two separate "pseudo-counties" from the precinct reporting system.

Figure 2 also indicates that the NES output is released in two locations—at the central computer, and also at the state centers. The release of data at the state is primarily intended to be a service to the wire service members for their relay to newspaper and broadcasting subscribers of detailed information. Relay of this detailed information back to the state of origin also permits the network forecasting systems to have a political scientist/analyst in the state receiving and analyzing this county level data.

Election Night Projections

With a threefold replication of this single application by the networks, one might raise the question of why the projections and analyses are not handled on a pooled basis as is the collection of the raw vote. Although, in some sense, the techniques used by the three networks have similar ingredients, the merging together of the basic informational ingredients is handled quite differently by the three networks, as we shall see in subsequent discussion.

What are these informational ingredients? As discussed above, all three networks have a common source of raw vote information, at the county level if they desire to use that level of detail in their individual projection models. Additionally there are two other sources of information for the models: prior or time-zero estimates and a sample of "key" precincts. All of the systems utilize some sort of prior estimate as a base of departure for the information that comes in during the evening—with that short of a time frame within which to operate, it is almost mandatory that there be an initial estimate of the final outcome, which is substantiated or modified by the actual flow of vote data.

The key precincts are the primary information source on which all three networks base both their projections and their analyses of the whys and wherefores of the election day happenings. The term "key precinct" refers to precincts which are selected by one of a variety of means, and then staffed by a network representative who phones the results directly to the network computer center. This pro-

vides information that can be received and analyzed as a sample of the total vote prior to its inclusion in the normal collection process. The key precincts are selected on a probabilistic basis, either completely randomly, or constrained in their selections to represent various strata or ethnic groups.

In utilizing these information/data sources for the projection of election results, the analytic staffs of the networks, served by computers which use procedures supplied from the minds of mortal man, face an interesting accuracy-tolerance problem in the calling of winners of the individual races. In a race where one candidate receives 65% of the vote, accuracy within 10% is more than sufficient to allow a correct call of the race. If that winning percentage slips to, say, 50.5% in a two-party situation, then accuracy within six-tenths of one percent may not be sufficient to prevent calling the wrong winner, although, from a statistical point of view, such accuracy (0.6%) would be rather remarkable for the amount of information generally available. The accuracy problem goes beyond merely the calling of a winner for projections are also made, in varying degrees by the different networks, of the vote-split, the plurality of margin or victory, and the total turnout. Further, there is an increasing amount of attempts to explain why a particular outcome happened in terms of the vote of various ethnic groups, or of various demographic or geographic strata.

Network Forecasting Systems: A B C

Because of direct involvement by the author with ABC over the past three national elections, this paper will give the greatest amount of space, and thus discussion, to the forecasting system utilized by ABC in the 1968 elections. A briefer discussion will be given of the systems/techniques utilized by CBS and by NBC. Actually these systems are evolving over time, and there will be changes in the systems utilized in 1970 and beyond—evolution both in projection techniques and the input/output aspects of the system. In this application, the actual projection models are but a small, but crucial, element of the total system, for the basic information must be fed into the system quickly and accurately—and the projection information, decisions, and massaged raw data must be routed to the analyst or commentator at the time when it is most useful.

The ABC system is described in Figure III-3. Briefly the system functions by first formulating a

ABC SYSTEM

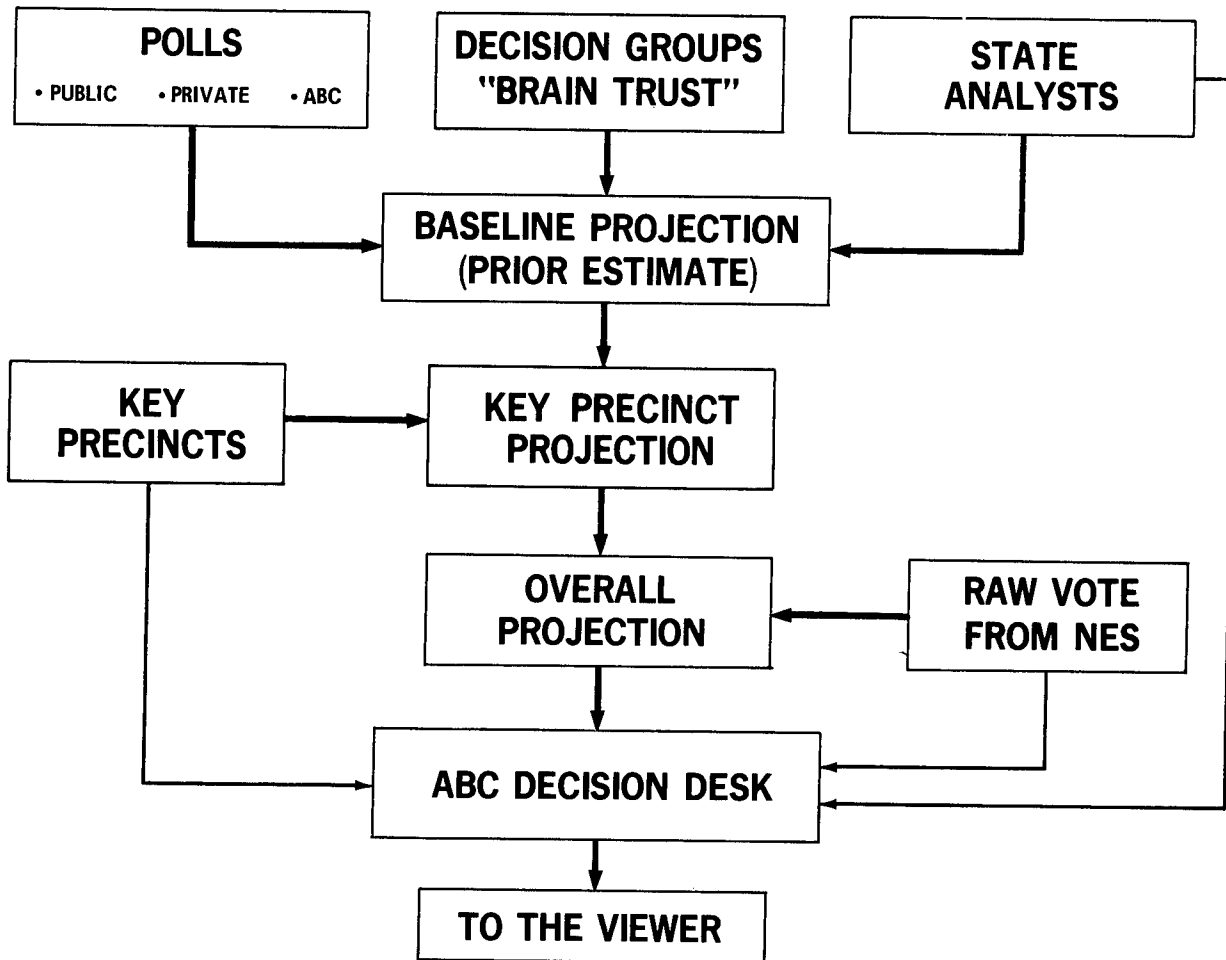


FIGURE III-3

baseline projection, generally over the weekend prior to the election, but subject to modification on election day. This prior estimate is put together utilizing polls, advice from analysts within each state, and the best judgment of the ABC decision team. The baseline estimate is modified by key precinct information and a combined estimate is produced. Finally, raw vote information is used to modify the baseline-key-precinct estimate to produce a final projection. This final projection by the model is then routed to a decision desk which evaluates the situation in the light of all the information, data, and intelligence that it has and makes or defers the call of a winner and any analytical estimates.

Let us look at the three major informational segments of this forecasting system individually.

Baseline Estimates

The baseline estimates are really a time zero projection based on any and all information available to the ABC staff prior to the actual tabulation of the votes that are cast on election day (or whenever absentee ballots are introduced into the tabulation system). They are, in a sense, the best subjective forecast by the projection staff based on polls, on educated opinions, on informed judgments, and any other sources of information available to the network staff. The estimates are numeric in nature and give the exact projected vote split, not merely an estimate as to who will be the winner. As indicated, they, and other parameters of the model which must have initial estimates, have traditionally been a task for Saturday evening (after

the usual Saturday afternoon system rehearsal) and all day Sunday. By Monday and Election Day there are many other things that need attention, and little additional prior information generally becomes available during those final two days. Such as does become known can be introduced into the system at the last minute on Tuesday.

Key Precincts and their Estimate

For the 1968 election coverage, ABC selected its precincts in a two-stage sampling process. After the decision as to the allocation of the total sample size which their available resources would permit them to select—an allocation based on the potential closeness of the race, in which case more information would be needed for decision purposes, and the importance of the state—the first stage of the sample within any individual state was to select communities or sub-elements of the state with probability proportional to the voting age population of that sub-element of the state. Within the selected communities, precincts were selected randomly, but were oversampled. The oversample was reduced to desired size based on considerations of historic performance, on the availability of historic data and current socio-econo-ethnic descriptive data, and on the potential availability of the results on election night. If the vote from an individual precinct would not be available within a reasonable time frame on election night (a somewhat open-ended time interval expiring sometime on Wednesday rather than strictly at midnight), if at all, that precinct was discarded from the list of key precincts. It does a forecasting model little good to receive the report of a key precinct after 80 to 90% or more of the total precincts in the state have been tabulated. (Of course, some precincts are selected for their potential assistance to analysis, rather than the quick forecast of the winner, but still it helps to be able to provide analysis at the time a decision is made.)

The precincts, after being selected, are researched to give both their historic performance in several immediately prior elections and also their ethnic composition. This latter information is used (by all three networks) in the portions of the model/system which project and help analyze the whys of the election and how the various ethnic blocs are voting or not voting (some comments on this approach later in the paper).

The need to obtain historic information also

provides certain problems with precincts whose boundaries or composition, in any one of a number of senses, have changed since the preceding election. For example, the precinct may have new boundaries, it may have a new high rise apartment, or it may have been partially cleared for redevelopment within the past two years. Because of the extent of mobility and dynamic changes so prevalent in urban portions of this country, one cannot simply rule out those precincts which have undergone such changes, for to do so would eliminate a very major portion of the electorate from objective consideration in the projections. Thus, in many cases, the historic information must be estimated, rather than taken from historic records. The ethnic data is generally estimated by some local person who is taken to be thoroughly familiar with the precinct area.

The statistical selection of the precincts, and their researching, are but part of the battle. These precincts must be staffed to report quickly and accurately directly to the studio. Here there are problems. ABC generally uses a staff of two on-site persons to handle each key precinct. They visit the precinct several weeks in advance to see if they can find it, and also to determine, among other things, the availability of a phone for the quick relay of the vote totals to the ABC studios. If there is no phone available for use in the immediate vicinity, then ABC has a private telephone installed for the use of the reporting team. In one instance, in 1966, even this presented a problem, for the precinct was in a rural farm house and the farmer would not permit a phone to be installed in his house. The resourceful team members looked around outside, but they could not find a suitable building in the vicinity. They did notice a very large oak tree, and, following a request to the phone company, a phone was installed in the oak tree.

Telephone installations and geographic locations, although providing some problems, were not the sole source of difficulties encountered by the field reporters. There also was problems in obtaining the vote, such as: a family of skunks under the floor of a precinct polling place; blown fuses; locked up telephones; precincts in which only Spanish was spoken; high wall phones for short women reporters; and several instances of teenagers monopolizing a public telephone booth that the reporter originally intended to use. The reporters also performed a variety of chores, including providing dinner or snacks to permit election officials to get on with the ballot counting without taking a dinner break.

Raw Vote Estimates

The incorporation of the raw vote into the projection model calls for an awareness of the patterns of reporting within whatever unit is being considered, especially if one is working at the state level. It is possible to assume that a county, or a major portion of a state, is reasonably homogeneous in its reporting pattern over time. However, in New York State it is well known that New York City reports before the rest of the state, and that it casts a higher proportion of Democratic votes than the upstate portions of the state. There is a similar well known pattern for Illinois, with Chicago or Cook County reporting earlier and being predominantly more Democratic than downstate Illinois. (Illinois can also provide another type of difficulty, such as in 1968 when the bosses of Chicago, the rest of Cook County, and some downstate counties all announced in advance that they would hold up the vote reporting from their areas, presumably to see what the rest of the state would be doing.) Patterns of reporting hold in varying degrees of existence and historic consistency in a good many states where the early vote may be more Republican or more Democratic than the final outcome and the discrepancy may disappear monotonically, or may, in fact, swing the other way during the middle of the tabulation process, and then later disappear. It might be noted parenthetically that the increasing mechanization of the local tabulation, and the faster higher level tabulations which are produced by a computerized NES add an additional element of uncertainty to the accuracy of any of these historic reporting patterns. One way to minimize the effect of this uncertainty in reporting patterns would be to use finer geographic breakdowns than the state level to permit the model to compensate for any such pattern. Such segmenting is being done within the ABC model, but even so the timeliness of the availability of aggregated data may force whole-state useage at various times throughout the evening, when the less frequently moved county level data may fall behind statewide data.

When statewide data is used to increase the total amount of information being used, even though a finer breakdown was desired, a correction factor may be entered to compensate for the historic pattern, or that vote-reporting pattern which is expected. Such a correction factor would normally be expected to go to zero at the end of the evening. Where absentee ballots are counted at the end of the evening, or even days later as with California

in the 1960 Kennedy-Nixon race, maybe there should even be a correction factor introduced with 100% of the precincts reporting in a state. (It should be noted that when the percent of the vote reported is given, that is for the number of precincts that have reported, not the total votes that will be cast, for where there is no manner in which the total vote to be cast can be known in advance to serve as the denominator for percentaging.)

Projection Model

Mathematically the general form of the projection model used by ABC is given in Figure III-4. The model functions by starting with a baseline estimate and using key precinct information to modify the baseline. This combined estimate is in turn modified by the raw vote information as it reports. The model is deliberately kept simple because of a strong feeling that it is important that the analytical personnel be able to understand not just the final numbers that come out of the model but also the individual components that go into that final projected outcome. This follows from the consideration that the projection model, as is the computer itself, is but a tool to lend assistance, but not to fully take over the decision function.

The actual model is based on what may be called deviation analysis, or swings, meaning that the model works with differences between historical data and present day happenings, rather than with the absolute levels of vote being recorded in those reporting units incorporated into the model. (The decision process itself might take into account the absolute levels of the reporting precincts.) The use of swings rather than absolute levels of vote may at times pose a problem for the analytical staff in years such as 1968, for there were no recent presidential races with three serious candidates and thus no useful three-party historic data available at the precinct level. Actually, this same lack of historic perspective exists in primary elections, where the number of candidates can vary widely, and the problem is solved merely by working with a swing from a zero historic vote, thus meaning that last November the minor party candidate (s) (all in a primary) worked with absolute levels. This modification for three or four party races thus places a greater burden on the humans who are interpreting the computer output, thus, again, the need for a comprehensible model. This approach made very little difference to the computational aspects of the

MODEL

$$P = \beta [\alpha B + (1 - \alpha) K] + (1 - \beta) R$$

Where: P - PROJECTION

B - BASELINE

K - KEY PRECINCT

R - RAW VOTE

α, β - WEIGHTS

FIGURE III-4

computer programs for 1968, but obviously had a rather major affect on the input of vote, and on the output of information, where all candidates had to be reported, rather than reporting merely the Democratic candidate with the Republican being derived by subtraction from one.

In the model described in Figure III-4, the weighting coefficients, α and β , are computed to reward consistency in that they are inversely proportional to the variances of the information element with which they are associated. They vary over time as the various forms of reports reach the ABC studios, but generally are monotonically decreasing. At the beginning of the evening, alpha equals one and the total weight is on the baseline, for neither key precinct information nor raw vote information is available. As the key precincts report, hopefully with at least some degree of consistency, alpha declines and $(1-\alpha)$ increases, thus

giving more weight to the key precinct information. By design the model does not completely eliminate the baseline estimate at this state. Also the beta weight starts out equal to one, until the raw vote starts to report more than a trickle, and then starts a decline, so that an increasing weight is given to the raw vote. Finally, in this model, at the end of the "evening", the coefficient for the raw vote information goes to one. Thus at the end of the election coverage, if the raw vote tabulation has also been completed, the model is in agreement with that final raw vote tabulation. It might be mentioned again that such a system presented a problem in 1960, for in that year in California the "final" tabulated vote showed Kennedy with a slight lead over Nixon in the Presidential contest. However, California, in that year but not thereafter, tabulated absentee ballots some two weeks after the election. The absentee ballots were so heavily for Nixon as

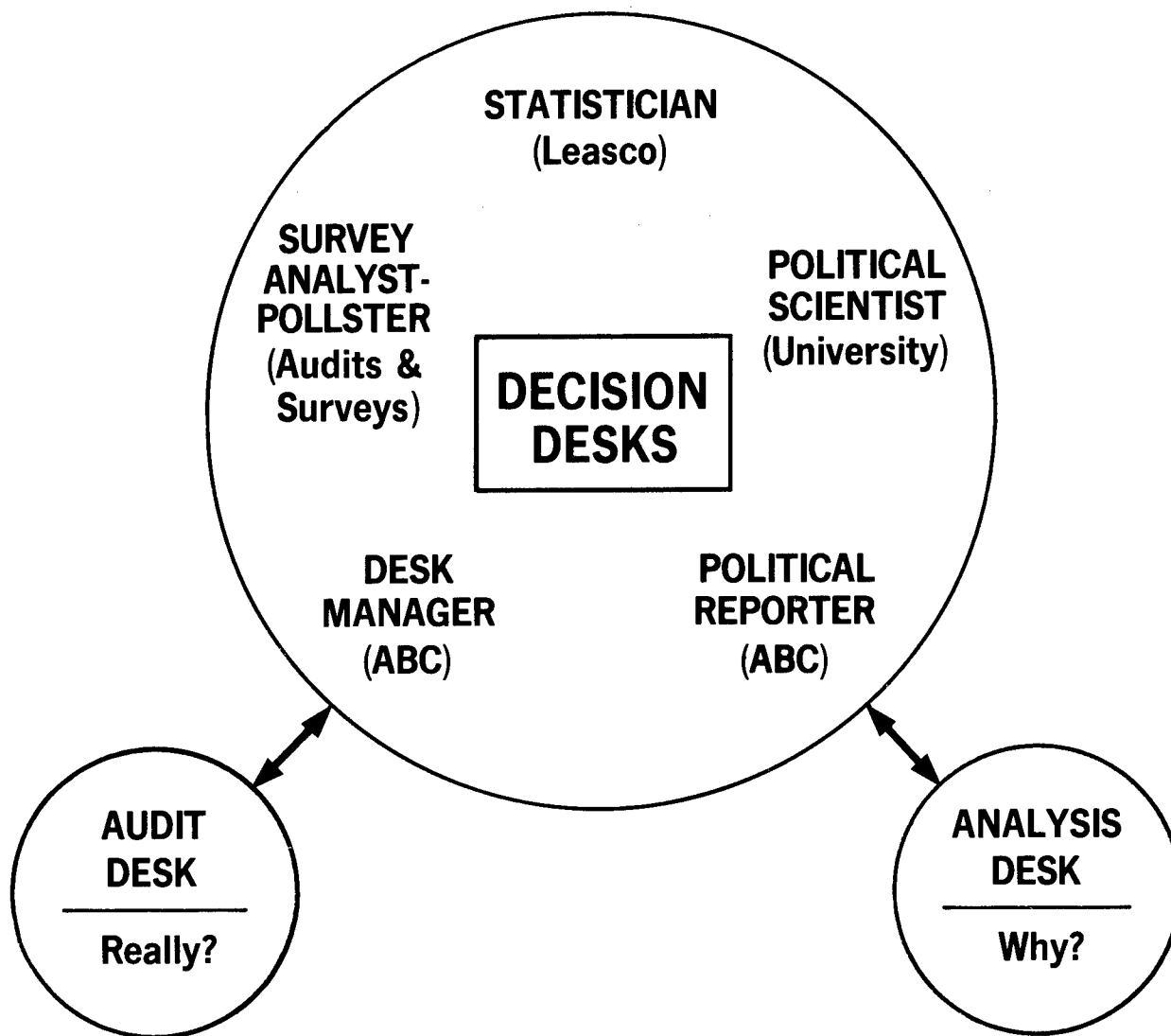


FIGURE III-5

to swing the victory in that state's Presidential contest from Kennedy to Nixon. A similar happening occurred in 1968 when the late tallied absentee ballots in Missouri switched an indicated victory for Humphrey to a victory for Nixon in that state.

ABC's Decision Process

As has been indicated elsewhere in this discussion, the final decisions as to who has won any individual statewide race, what his winning percentage is or will be, what his plurality is, and any other projections rests with the members of a "Decision Desk"—in 1968, one handling the various statewide Presidential contests and another handling the various Senatorial and Gubernatorial races. This

is not to say that the forecasting model could not be trusted, but that there frequently were other factors which might be considered in an evaluation of the model's results—factors which would not lend themselves to sufficiently rigorous handling procedures so as to permit their inclusion in a computer program. For instance, in the final states to be decided in the Presidential contest in 1968, one crucial factor was where the outstanding vote would come from, and what its split was likely to be, or how far it would have to go for the trailing candidate in order for him to catch up.

The composition of the decision groups that ABC utilized in 1968 is shown in Figure III-5. ABC News provided a desk manager, who had managerial responsibility for the functioning of the desk, and

had responsibility for insuring ABC News that there was sufficient evidence to substantiate the decisions being made. The four active decision makers included a political reporter from ABC News; a political scientist from academic world; a survey analyst/pollster from Audits & Surveys who did the key precinct sampling and research; and a statistician from Leasco Systems & Research Corporation, the 1968 employer of the team that did the modelling, systems designs, and programming for ABC from 1964 on.

Backing up the two decision desks were two support desks: an analysis desk that examined the performance of the ethnic and other blocs and generally delved into the why's of the various victories or defeats; and an audit desk which monitored the decisions that had been made, to keep an eye out for needed reexaminations or reversals of any decisions.

Projection Techniques: CBS

Under Lou Harris in 1964 and 1966, and an internal staff in 1968, CBS has concentrated their efforts on developing and utilizing essentially a key precinct system for forecasting. Their system works with stratified probabilistic samples of precincts and weighted estimates within the various states, treating each precinct as a point of estimate of the results for the state.

To quote from a paper given by Warren Mitofsky of CBS News at the American Statistical Association 1968 Annual Meeting, "The estimation of the outcome of an election is treated in many ways as a traditional sample design problem. A single state sample of precincts is selected with probability proportionate to size from a stratified frame which includes data from a past election for all precincts in a state. The stratification is based on past vote for precincts as well as county characteristics. Various estimators are available that will make maximum use of the usually high correlations of total vote or party vote with past elections, such that variance can be minimized. In fact, precise estimation of election results would not be conceptually difficult if it were not for the requirement imposed by the networks for calling the winner at the earliest possible time."

Projection Techniques: NBC

NBC uses basically two projection models: a precinct model and a county model. The precinct

model uses two types of samples, random samples and barometric samples. The barometric approach is used in states where there is sufficient past data to substantiate that in past elections these precincts mirrored the voting performance of the state as a whole. Where historic data is difficult to come by, they utilize a random sampling approach. In November of 1968, they utilized 2,681 precincts in 50 different state samples, mostly barometric. (There is one problem with dependence on barometricity in the key precincts in that many precincts have to be dropped from consideration due to a "reasonable suspicion the precinct no longer will vote in a way typical of the state"—which eliminates those areas that are undergoing growth or decay.)

Once the precincts report to NBC, then they are examined from three viewpoints within any given state:

1. the vote within the precincts that have reported are totaled by party, thus giving each vote an equal weight;
2. the percentages of vote reporting by party within the individual precincts are averaged across the reporting precincts, thus giving each precinct equal weight; and
3. the difference, or swing, between the election night results and the results in those same precincts in past election are averaged as in viewpoint 2.

The second of their two systems, the county system, "is a highly statistical, quite scientific, mathematical projection model. It can be worked only on a computer, because only a computer can perform the thousands of individual calculations this projection model demands within the time span—measured in seconds—for which this projection model was built." The county model works with all 4,500 counties in the United States, not on a sample basis. What it does is store up to 48 attributes on each county, and then uses consistencies in observed voting patterns for counties for which there are returns to adjust counties for which there are no returns in yet, and thus maintain and continuously update projections for all counties, and thus all states. NBC literature refers to the concept as that of "borrowing strength". This concept is based on the fact-borne out by past experience—that if you have a group of counties that are voting in the same way, you can assume the results in one of those counties, where most of the vote has been counted, will apply to the other counties, where only a fraction of the ballots have been counted."

The results from the two models were coordinated by two teams. The first decision team concentrated on selecting the indicated winner, and an initial estimate of his winning percentage. Then a second team took over, monitored the course of the election, and made any necessary adjustments in the percentages.

Needs

Other than the continuing evolutionary improvement in forecasting models by the various individual networks, what can be done to improve the entire process of election forecasting and examination over time spans not restricted to election night. Two areas stand out for development: a precinct level data bank and improvements in the analytical techniques used to estimate the voting patterns of ethnic groups.

With regards to precinct level data, for use election night, and also later by those interested in research in voting patterns, NES could be encouraged to move in the direction of setting up a national, real-time, precinct level data bank which could be used for:

- * tabulations on election night, at any desired level of detail,
- * key precinct reporting schemes for the various projecting organizations, wherein a stored index key for the precinct indicates which organizations desire to use it as a key precinct in their own projection model, so that each precinct would be staffed only once, and the reports to the networks would be on a computer driven real-time communications switching basis, and
- * to enhance post election research.

The analysis of the total vote picture and its various components has never been satisfactorily handled by the networks. The psychological analysis schemes used heretofore by the networks have a common major defect in that they all depend on a handful of precincts representing the totality of an ethnic bloc. The problems inherent to approaching

the "WHY?" problem in this manner are two-fold: (1) individual precincts may be assigned several labels such as Irish, Catholic and Urban for each precinct can represent several ethnic blocs or sub-blocs; and (2) if the precinct is sufficiently homogeneous, then it may be atypical of the bloc it allegedly represents, e.g., the poor northern Negro precinct in Harlem that is dominated by the Black Muslims.

With proper care in the selection process, and a little bit of luck, one can pick precincts whose voters predominantly represent that bloc; in fact only a few may if errors are made in the selection process such as a precinct that was picked to be Jewish when in fact it was predominantly Negro. In addition to the mixing of groups, however slight, another difficulty with this particular approach is that other effects may be confounded with the one under consideration and may have a greater effect on the voting of the precinct than that attribute in question, e.g., Negro and low income. One also, in a sample of five or so precincts may be lucky enough to include the one Negro precinct in the country that went for Goldwater (in 1964), which would, of course, play havoc with any small sample results of which it was a part.

Some research has indicated that regression or linear programming would yield a preferable approach to this analytical problem, linear programming slightly preferable in that the coefficients that represent the percentages voting in any given manner can be constrained to be between 0% and 100%, which are reasonable bounds to impose. More study needs to be devoted to this question.

The competitive pressures of election coverage have resulted in a willingness (albeit somewhat reluctant at budget setting times) of the national news networks to invest in the development of good projection methods, both total systems and models, proven under pressure. It is to be hoped that these methods will continue to improve and that an expanded use of these techniques can be made in other application areas.

DISCUSSION FOLLOWING PAPER GIVEN BY DR. RICHARD C. TAEUBER, PHOENIX SYSTEMS INTERNATIONAL, "ELECTION NIGHT FORECASTING"

MR. LINSTONE: A question arises from your very interesting discussion, by implication really. To what extent does the TV audience really base its

decision on which channel to watch on the accuracy, opposed to, let's say, the personality of the commentator?

DR. TAEUBER: I suspect strictly on personality. I do not know. I am at a disadvantage because I do not watch the coverage on election night. After 1966 I was told, for instance, the ABC coverage in 1966 had a large lead in the calls made first. I am also told that from a technical standpoint ABC had the most interesting show that night. The directors and the production people did a fairly good job of blending the actual coverage with interviews and that sort of thing. However, that did not seem to affect the audience size, so I suspect they go basically on personality. Also, from the reviews that came in the next morning, it appears they react to the speed of the early calls. Here is where you really get into the guidelines imposed by the networks. ABC had a strict policy of not releasing any projection prior to the final poll closing time in that state. CBS, the only other one I know about, had a policy where it was up to the individual station to decide whether an early call would be released in their area. In other words, they were not making it a network policy to hold back the call. They were placing the burden on the local stations. I have not seen NBC's policy in writing; but, presumably from all the pious pronouncements before the election, nobody was going to release any calls or projections prior to poll closing.

MR. LINSTONE: I thought there might be some profound reason such as which program they happened to be listening to before they thought about the election, and not change channels.

A. I suspect there is some lack of momentum in switching channels, especially where some parties watch all three networks.

MR. FRIEDMAN: The question I would like to ask does not apply as much to national elections—i.e., the presidential election—but more to local elections. It seems to me that if I can control the design of the ballot, the placement of names on the ballot, I would have a very significant effect on the outcome. For example, if I am asked to select eight out of sixteen people that are running in the primaries and if I can control the placement of those names on the ballot, I would have a significant effect on the outcome. Any comments on that?

DR. TAEUBER: I have seen no studies. There are variations by state as to whether you vote a straight ticket or whether you have to vote candidate by candidate.

DR. WARD: Is it sometimes determined by a lot drawn before the ballot is made up?

A. Yes. Also sometimes a specified party or the incumbent is always first; sometimes the candidates are listed alphabetically.

MR. LINSTONE: We found in California in one election, where a junior college board of trustees had 133 names on the ballot, the results were such that the highest 14 are in a run-off now, but they were not the first fourteen names on the ballot.

A. I suspect in some races, especially in local elections where several winners are to be selected and the candidates are not as well known, the position may have some effect as to which subset of the unknown candidates wins.

DR. SLAFKOSKY: Effect of position on the ballot depends on what kind of an election one is talking about. In an election of County Commissioners or membership on school board when one may have as many as five or six from each party actually to select from a number of about twenty, then position on the ballot may have a decisive effect. This holds for a person who may have decided that he is going to vote for A, B, and C and yet there are two or three slots to be voted for. Under these conditions the position of a name on a ballot could be advantageous.

MR. BOBROW: I am curious and this is more of a forecasting question and not so much short term: what kind of feeling do you have for the different degrees of accuracy you get out of your class of models on election night with less than, let's say arbitrarily, fifty percent of the precincts reporting. For instance, the accuracy you get from the last pre-election poll versus the accuracy you get from really pre-election procedures like in 1960 simulation, you know, at different levels in the electoral processing. There was how much gain in accuracy before you are told all the votes are in, when, of course, it loses its importance.

DR. WARD: How much improvement are you getting from your absolute valid point—you know, who is ahead and no matter by how much?

A. I would say, in general, our estimates in the middle of the cycle for in-state are probably the most accurate. Raw vote in most cases is fairly stable if you know that you can trust the raw vote that you are getting. In Illinois, in the Dirksen race, in which everyone knew months, weeks, days, hours in advance, what have you, Dirksen would win, there was almost no question and we were showing sixty percent of the vote for Dirksen. We still did not call it because we were getting word from our state representatives that there was a possibility that an upset was in the making and we also could not trust the figures they were listing because we were getting contrary evidence from the state that it was running below fifty for Dirksen than what

was coming over the wire at sixty percent. It turned out about fifty-four—a lot closer than it should have been. I suspect, although I haven't really studied it, I suspect there is something to be said, that the polls are potentially the least accurate—not the final polls, but I am thinking, like you are referring to the 1960 studies, that the longer range estimates of the percentages of vote by party do not change that much. There is some persistence in voting behavior, and I think, although it is not substantiated in any way, the polls tend to reflect changes of the moment, happenings of the recent past prior to the poll taker appearing and asking questions of the person being interviewed and that a good many of these disappear on election day when a person walks in to vote. I do not think, for instance, in the national percentage, which again does not mean anything, that Humphrey's percentage at any time would have been as low as the polls forecast in September. The same pattern occurred in 1964, in that Goldwater was not as bad off in 1964 as the polls showed him to be.

MR. RUMMEL: Was the base line projection on less than fifty percent of the returns very much off from your final projections?

A. I have not run a check on 1968, how accurate it was. The standard error ran below five percent in 1964 and 1966—somewhere about four to five. I have not really broken down the accuracy for those where the base line estimate was between forty-five and fifty-five percent Democratic, where accuracy is more critical for selecting a winner. In 1964, in setting base lines, we made an arbitrary decision which we lived to regret—that we would not set any base line higher than seventy percent for Johnson or for Goldwater in Mississippi. We just did not believe that Massachusetts, for instance, would go eighty percent Democratic, which they did. This enters into the variance result that comes out on post-analysis. I can think of others where the base line was right on. Removing the effect of the arbitrary limit, I would say probably within three percent—so we are about as accurate as the polls claim.

MR. RUMMEL: Was there any attempt before you built a particular model you work with to get regression equations based on the attributes—not the previous voting, just based on their attributes?

A. There was in 1962 with NBC. However, at ABC in 1964 we made the decision that we were not getting enough payoff by this attempted method of linking counties into what were called federations. I do not know what NBC calls them now, but

we called them federations then. They were linked by regression techniques or similar factors, responses and so forth. We decided there was not enough payoff in that approach to justify its use. We also had a feeling, especially in the closer range, that we needed a model which people making decisions could understand component by component and not just have to accept a final result at the end of a vast amount of statistical manipulation. The other argument, too, is on the precinct level: should we use weighted estimates or should we use simple unweighted averages?

MR. RUMMEL: Some of our post-graduate schools in political science concentrate on studying voting behavior in the United States. Would you care to give an assessment of how well political science has done in trying to project voting behavior from the characteristics of the voters—trying to project voting behavior from memory—that is, from persistence in voting as you have done?

A. It is not ducking the issue, but I do not think we have any way of measuring this now.

Q. Do you think good projections can be made from voter attributes?

A. I think it can be done. As far as I know, the post-election analyses at least to the extent that I have seen, are based on averages in appropriate ethnic precincts. A Negro precinct or any other ethnic group is defined as one having over sixty-five percent of the residents of that precinct having that attribute. Then we say how the ethnic bloc votes based on this average, which is not right, because we are getting confounded information. One would get a different effect in a precinct which is a hundred percent or ninety-five to a hundred percent of a block than you would with sixty percent or one that is forty percent—at least I think so. Intuition says that is so. I suspect most of the detailed studies or attempts to prove how members of any ethnic bloc vote are probing for better methods of analyzing ethnic vote. Maybe if we go to individual touch-tone voting with instantaneous compilation, we can identify the type of the individual and his attributes. We could really analyze it properly. As I say, this is one reason I would like to see a different analytic scheme come forward. It would help with political science, too.

MR. IRWIN: I assume that in getting characteristics of precincts or counties you would have to use the 1960 census because I do not believe this data comes from anywhere else. Did you notice any difference between 1964 and 1968 remembering the fact that

twenty percent of the people in this country change their house every year?

A. We basically do not use such information, even census information. Even 1960 census information does not go down to the precinct level. We need this level of information now, not only about the county but the precincts. Basically, the contact for needed information is a precinct official or someone who is supposed to know the attributes of the precinct. This is another reason I would like to get away from precinct level ethnic analysis and change to using the county level where you can pin it down better, at least to the last census. Presumably there is some consistency in the type of persons who live in a county. You can at least pull out counties that are reasonably stable in ethnic composition. That is why I would like to switch to that rather than using this precinct system, which again is based on estimates that are guesses.

MR. IRWIN: I am not sure that you are better off where you are. The point I was going to make is that this country is extremely mobile, and these characteristics are going to change. The only other way to get these characteristics in more detail is by a census, and you may be pretty well off the way you are.

A. But there may be some stability or you may be able to pull out counties which are more stable; given that there is a turnover, they are still likely to be replaced by a very similar type of people.

MR. IRWIN: Not necessarily, that is the point.

DR. SLAFKOSKY: I do not remember exactly, but if my memory serves me right, the prime bases for your election forecasts are the base line projection, the key precinct data, and the historical data. I would like to explore the very nature of your base line projections a little further. What are their prime ingredients? How do you arrive at a base line projection? What are its prime characteristics.

A. How do we come up with it? At ABC we have a worksheet for each race, including the presidential race, in each state. On that worksheet we keep track of all the polls we can get hold of: from state sources, some private polls that ABC commissions, and from newspapers. As an aside, going back to 1964, NBC's computers were down and they came on the air quite early and said their system projected that Johnson would get between sixty and seventy percent of the votes that night. I happened to have in my pocket a slip of paper which had the results of something like eight national polls that had been taken at the last moment. Some were private polls. (Rockefeller had a national poll commissioned at

the last minute. Why, I do not know, but he did it.) Those polls ranged from sixty-one to sixty-nine percent for LBJ. You can see where NBC's information came from. Anyway, we have a worksheet on which we have all the polls we can get listed. This is done on the weekend before the election. We start Saturday night—we have a big system rehearsal Saturday afternoon, go out for dinner and then work Saturday night and all day Sunday as needed. We have a report filed by our state political analyst, generally a political scientist from a university or, in some cases, political reporters to newspapers who file a report based on what they have observed within the state in which they give numeric estimates. Then we have a big round, really rectangular table with whoever is going to be involved in the decision process plus some research people sitting there coming up with a consensus. The detailed consideration that is given to each particular base line is greater earlier in the evening. We did make some attempt to take closer situations. First, we did not just go through alphabetically. Somehow, after six hours at a single sitting or after ten hours or so in two sittings, your willingness to devote a lot of attention to a particular item starts slipping—I hope I am not giving anything away by saying that. Generally the base line is based on any objective or subjective information we get our hands on.

Q. How do you mold this together in a rational way? You talk about eight different polls and talk about the Rockefeller poll and you talk about the political analyst. There has to be some way of ascertaining the base line projection itself. This is what I am asking for.

A. I would not say it was haphazard. There is not, in most cases, that much discrepancy around the table.

Q. You are all in accord very readily?

A. Generally the polls and the analysts are reasonable close to agreeing. For the base line we are not estimating just the winner, which makes it a little harder. We have to come up with a percentage. We usually do not try to go to the tenth or percent; we go to a whole percent which is sufficient accuracy for our needs for this is a starting point, not an estimate to use on the air.

MR. RUMMEL: What was your whole percent in the case of 1968?

A. You mean weighted over-all national?

Q. Yes.

A. The presidential base line is by state but we also estimate turnout in the states, so we use the turnout estimates to weight up to an overall na-

tional figure. I would say, if memory serves me right and I do not have anything with me I can check myself with, 46, 42, 12. We were a little over on Wallace. But I would say there is a large subjective element in the process-it is not a mechanized objective procedure.

DR. SLAFKOSKY: That is what I wanted to ascertain. I got the impression that it was primarily subjective.

DR. TAEUBER: I will say this: with any group where you get twelve people to reach a consensus, maybe the first few are really consensus of the group, but later on in the evening when the argument gets going a smaller subset tends to dominate. I could probably name by name the people sitting side by side who would say "How about this percent?" Someone would say 56 percent; another, 52 percent of something, and we would say, "O.K., how about 53?" and everybody would say, "O.K."

DR. SLAFKOSKY: There is no problem in an election which does not entail a close race; but in a closely contested election despite these agreements and despite the polls, this kind of consensus is just too subjective to warrant any significant confidence.

A. The base line, we feel, is very accurate, at least generally fairly good; it has been so historically. We try and knock it out of the projection model as soon as we can, but we do not take it all the way out until the raw vote dominates.

DR. SLAFKOSKY: You say you get rid of it as fast as you can. It seems to me the very factor you want to use as your prime measure is the one you want to prove most reliable. If, in fact, you try to eliminate or "knock it out" as quickly as possible, it does not seem to be your prime measure. Perhaps I am not understanding you.

A. We like to get actual data and we feel precincts give us a better handle than the base line. Again the problems of a close election: you really have to have a lot of raw vote before you make a call. Also, we retain the ability to look at the base line, the key precinct data, and the raw vote separately. Go back in the California primaries in 1964. Lew Harris, when he was doing the vote projections for CBS, did his political polling in the same precincts he used for his election night forecasting. He was polling in Los Angeles and in June of 1964 indicated a fifty-fifty split: Rockefeller and Goldwater in the Republican presidential primary. The ABC poll indicated a sixty-fourty split; Goldwater over Rockefeller in Los Angeles. The early returns came out at sixty-fourty. Harris looked at these results which were ten percent over what his research, what his polls in the precincts, showed he should be getting. Therefore, there was a Goldwater trend; and so he called the state for Goldwater. We were showing exactly the same thing he showed at the same time but it was right on our prior estimate. With the election going as we projected we did not call because the sixty-fourty in Los Angeles was indicating something like fifty-one and a half percent overall statewide. This application still is a human decision process. Some of the races can be automated as to the decision if they are very one-sided, but in general we still strongly prefer to let humans to the deciding and use computers to sort and process the vast flow of information and data so that the decision makers can have what they need where they want it at the time it can be most effectively used.

Forecasting Highway Demand*

ROBERT W. PATERSON**

INTRODUCTION

The Nation's highway network is an integral part of its transportation system. Consequently, long-range planning for highways must be done within the framework of overall transportation planning, which, in turn, must be geared to sound national, State, and community planning. As a result, in planning future highway facilities, transportation planners, including highway planners, must take cognizance of demographic, employment, income, and other socio-economic factors as well as trends in motor vehicle registrations, vehicles in use, and motor vehicle travel.

Good planning requires good forecasts. Estimates of future highway needs have consistently under-

rated the dynamic growth of the underlying factors that create the demand for highways. Examples of overestimation of these factors are so rare as to be almost unique. Fortunately, in recent years, with availability of more accurate and complete statistics, and better and more sophisticated techniques to be applied, the record of reasonably successful forecasting is improving.

Although most forecasts are a result of the blending of statistical processes and judgment there is an increasing interest among forecasters with mathematical processes. Much of the work being done in this area is still in a developmental state and there are very few studies available which explicitly employ operational mathematical models.

Forecasting and projecting are integral parts of the planning process. Planning involves, however, much more than the estimation of a set of probable future statistics. Nevertheless, because such statistics set the quantitative limits for the planning process, their accuracy is of concern to those who are involved in the broader problems of planning. And as the magnitude of a particular program expands, judgments about the future become increasingly significant emphasizing the search for better ways of making decisions. Simple linear extrapolations which may have been satisfactory for programming activities at one stage of development, often become inadequate at a later stage.

Time series analysis yields information about tendencies and trends as well as about the rate and direction of change. The economist or planner who utilizes the results of time series analysis may project a particular quantity into the future. In so doing, he assumes that if the events and influences of the historic period persist in the future, then his projected value will be accurate within a narrow range of variation. On the basis of, say a forecast of population, it would be possible to build a road system designed to provide service for a future time period. Superficially, it might appear that, in

* See Robert W. Paterson, *Forecasting Techniques for Determining the Potential for Highways* (Research Center, University of Missouri-Columbia), 1966, pp. 128.

** Robert W. Paterson is Professor of Economics and Director of the Business and Public Administration Research Center at the University of Missouri, Columbia. He has taught and conducted economic and public administration research projects at the University of Virginia and University of South Carolina. He received his PhD in economics at the University of Virginia.

Dr. Paterson was an officer in the Foreign Service of the U.S. Department of State from 1945-1949 and was stationed in Cuba, Brazil, and the British West Indies. During that time he was officer-in-charge of the U.S. Consulates at Victoria, Brazil, and at the Embassy at Barbados, B. W. I.

During his academic career Dr. Paterson has lectured, given papers, or been a discussant at conferences in South America, Europe, and Africa as well as in the United States. In addition, he serves as an economic and planning consultant to business firms and agencies of government.

Dr. Paterson is the author of *Forecasting Techniques* and *The Rise of Disorganization in Metropolitan Communities*. He is currently engaged in a study of *Congressional Response to U.S. Business Cycles*. He is a past president of the National Associated University Bureaus of Business and Economic Research and is a member of the National Academy of Science—Highway Research Board, Socio-Economic Aspects of Highways Committee.

deciding to act upon such a forecast, we were utilizing appropriate knowledge to provide a public asset of critical future value. But by basing the future location of roads upon nothing more than population trends of the past, one is in effect assuming that past patterns of expansion will not change in the years ahead. And, by its influence upon population movements, the very act of highway construction tends to lend credence to that assumption.

A. Economics and the Allocation of Resources

Economic analysis is concerned with the allocation of resources. Given a static general economic system in which the bounds of aggregate demand for goods and services are unchanging, the economist is primarily concerned with the way in which the coefficients of production (proportions of factors of production) may be altered to produce outputs at lower costs. A dynamic system, on the other hand, introduces a higher order of complexity into the problem of allocating resources in ways that yield lowest cost combinations for the millions of goods and services that are produced each day.

The problem of assessing consumer demand in a static economy where outputs, product choice, and technology are stable is sometimes difficult, but when one confronts the problem of measuring the efficiency of resource use in an economy of rapidly expanding product outputs, the problems mount rapidly. With the introduction into the system of fluctuating consumer tastes and preferences, as well as technological progress, it becomes of central importance to find means to measure the millions of transactions made each day in the market place in ways that will reveal the real costs of production of each item produced and that will also enable us to determine alternate cost combinations that would have resulted if a different mix of production factors had been employed.

In the private sector of the economy producers purchase the inputs at certain costs and provide products to consumers at selected prices designed to recover costs of raw materials, salaries and wages, depreciation on plant and equipment, return to investors, and profits. Individuals either buy or do not buy the outputs of firms. Their decisions are determined by the utility function for a particular good in relation to those of all other goods.

B. Government an Active Agent for Economic Growth

Turning to the question of the supply of public goods and services, but leaving aside the problems

posed by government subsidies to certain industries in the private sector, costs for final outputs are reflected, partially or wholly, through the tax system. User taxes for some government goods or services are similar to prices that are used in the private sector which must be paid only by those who wish to enjoy, for whatever reason, the fruits of a particular product or service. If there are no takers of the item it will leave the market and that is that. The governmental unit producing a product or service that is in excess of public demand may be guilty of having done a poor job of estimating (forecasting) the market, but society renders a judgment on its continuance through the ballot box. Reasonably accurate forecasts of public demands for government goods and services can minimize political problems of this sort.

It is not my present purpose to probe the philosophic questions surrounding the reasons why government expenditure programs have proliferated during the past generation. What is important is the fact that government has become an increasingly active agent for the expansion of the physical underpinnings of national growth. Some of these underpinnings consist of hospitals, educational and training facilities, scientific research centers, military establishments, and transportation depots, terminals, and road networks. These, plus all the other functions performed by governments compete for scarce revenues. Thus, studies of future public physical requirements need to be developed with greater precision if Society, through its public agencies, is to economize these resources.

C. Forecasting—An Art

Forecasting—the art of identifying potential magnitudes of change over a specified period of time in the future—is of importance in the construction of agency organization, planning and budgeting policies. It is often believed that forecasts, to be useful, must be accurate to within only a relatively small range of error. Yet, present techniques are not capable of incorporating all sources of possible change into a system for forecasting the demand for road networks with very great accuracy. But, barring catastrophes, such as wars and depressions, they can enable planners to view, within the constraints of certain assumptions, the general magnitude of road needs in the future. Planning officials charged with the responsibility for determining road needs cannot depend solely upon forecasts to establish the bounds of program demand

in the future. Forecasts must be interpreted by planners who have the technical skill, imagination, and boldness to translate them into facilities for human use.

A few comments on these qualities are justified in an exploration of forecasting techniques. They are justified on two grounds—historical and psychological.

In the past, most forecasts of population, income, and employment have understated actual changes. In searching for explanations why change is understated, one confronts a set of possibilities about some of which little is known. First, of course, is that of data availability. Second, many techniques used to forecast have not been constructed on a sound methodological base. This is not to say that they are necessarily logically inconsistent but rather they fail to incorporate analyses of alternative possibilities. Third, historical data yield clues to the future only when the future is similar to the past. To determine the future force of dynamic change in private investment, consumer expenditures, and income one must look behind these quantities and assess the quality and rapidity of technological change in tomorrow's economy. Fourth, planners like others are limited by the range of their experience and the particular patterns of their experience tend to become part of the psychological state which conditions their view of the future.

Many students of planning are articulate about the role entrepreneurial activity has had in the industrial development of the U.S. during the past century or so. Others are equally able to expound the technological basis for the rise of an entrepreneurial class. Still others find keys for the spread of American economic development in the rise of political processes through which the electorate institutes new government programs, in the timing of the industrial process itself, and in a wide assortment of other explanations. These, while they are of intellectual interest, do not provide quantitative indications of future changes.

Forecasts to be useful as policy aids must incorporate the bounds of policy assumptions within their own operational framework. This means that an interstate highway enactment requires operational forecasts of all travel, not just highways, if resources are to be economized. Planning must be comprehensive at the higher echelons of decision making and administration if a clear view of the production implications is to be gained. Those charged with the development of forecasts must be bold in challenging others who prefer to base programs on an

annual increment which varies little from year to year. This boldness should stem from a desire to incorporate history and intuition in the process of forming a solid assembly of assumptions for a forecast.

D. Forecasting Methods

Forecasting methods are not easily sorted into neatly arranged categories. There has developed in the last generation or so a wide variety of techniques for forecasting economic change. The methods used vary from highly intuitive processes that incorporate memory patterns (simple historical relationships) to rather sophisticated techniques for establishing patterns of relationships. The list of techniques does not exhaust the particular methods that ingenious analysts have devised for particular purposes. Nor does it often occur that any one technique will be sufficient for a particular purpose. Rather, some of the best work in forecasting incorporates salient features from several approaches. For example, in forecasting motor vehicle registrations in a state it is likely that most of the techniques listed would be utilized.

A competent forecaster would want to:

1. Review historical trends to identify and develop time series on population components, income, employment, productivity, residence, transportation modes, education, and recreation demand.
2. Employ an analysis of fluctuations in these indicators and develop indices of change.
3. Determine the probable impact of industrial and governmental activity by reviewing current plans for the future that will have an effect upon the indicators. Develop land-use data plans.
4. Determine the dominant forces that are likely to affect the indicators.
5. Determine the relationships between income and residence and motor vehicle ownership and travel.
6. Determine the relationship between city-size, income, and vehicle ownership.
7. Determine the relationship between transport modes and vehicle ownership.
8. Employ a model for testing the relationship of indicators.

Keeping in mind that forecasting is concerned with the identification of points in time when indicators "turn" or change direction will help the forecaster avoid the possibility of his becoming captivated by mechanical techniques for processing data. Of prime concern is the need to isolate the dynamic forces of change that are erupting and that are likely to shape the structure of the economic

system in the future and then to use this information in the forecast.

For purposes of highway planning forecasters have developed a method which incorporates regional analysis (measures of *volume* changes in economic activity rather than *qualitative* changes), land-use analysis (determination of economic use of land), and traffic flow analysis. Those favoring this methodology for predicting traffic flows depend upon aggregative regional economic base studies and land-use projections for setting the bounds of traffic generation. Some others favor utilizing an input-output analysis as the framework for the operational forecast on the ground that it is less general and therefore depicts industrial changes with greater precision than does the economic base concept. Unfortunately there are not many state and local input-output studies underway in the United States and only a few have been developed over a long enough period to provide adequate knowledge of structural change upon which to forecast.

If transport capacity and transport demand are to be balanced forecasters must assess the effect of an increase in overall economic growth on the composition of industrial activities to determine whether and to what extent proportionality shifts are taking place. The existence of present occupations may be threatened as time and change play upon the character of the productive process. But, since these factors do have a determining effect on future travel it is important to consider them, even though firm methodologies do not yet exist for their measurement.

It is apparent that methodologies for forecasting population, automobile registrations, travel per vehicle, and fuel consumption are in a state of development. There is no ready-made method that is presently sufficient for accommodating the increasing amount and variety of data into systematic organizational frameworks for forecasting that yields accurate estimates. It is debatable just how accurate estimates should be to be useful for planning purposes. Technicians (forecasters) concerned with the detailed day-to-day requirements of quantitative measurement necessarily are engaged in efforts to improve the quality of their products. At a higher level of administrative authority it is often possible (even necessary) to make decisions based on forecasts whose probable error of estimate is known to be relatively high. Improvements in accuracy would serve to buttress the arguments of those involved in the decision to build or not build a given road. What is of importance is that a parti-

cular forecast be accurate in determining, at least, the range (minimum-maximum) of probable future travel demand of a road. If, as mentioned earlier, a forecast cannot provide such an estimate then there is a high probability that investment decision makers may misallocate public resources.

E. Measurement of Forecast Results

In this presentation an attempt has been made to assess those techniques that have been developed to predict the future demand for roads. For present purposes "scientific" forecasting began in the 1920's. The development of a system of national accounts and a variety of statistical indicators, with regular reporting schedules, were the basic building blocks upon which present forecasting efforts have come to rest.

These tools made possible the undertaking of penetrating empirical studies and spurred many social scientists to believe that, once the data problems were solved, a new era would dawn in which public policy questions could be resolved by factual, objective, analytical processes. These hopes, it is now recognized, were largely unfounded. Yet, forecasting, while it is an art has nevertheless attracted able scholars who are contributing to its improvement and scientific development. Improvement, however, has been sporadic and this condition may be characteristic of forecasting for some time to come.

A review of some past forecasts leads to the conclusion that population estimates are very likely to be lower rather than higher than actual figures; motor vehicle registration estimates also strongly tend to be lower than actual; and about one-half the studies on travel per vehicles and total travel tend to be lower than actual.¹ These tendencies of long-term forecasts to underestimate the future suggest that reliance upon past trends is an inadequate base to forecast during a period of evolving economic change.

Another test of projection accuracy was made by classifying mathematical techniques and using each of them to project population at 10-year intervals beginning with 1910 and ending with 1960². Ten sets of projections were tested for the states of

¹ H.W. Hansen, "A Review of Travel Forecasts," *Highway Needs Studies 1967*, Highway Research Board Bulletin 158 (1957), pp. 104-108.

² Robert W. Paterson and Kenneth Lewis, Unpublished manuscript on *The Accuracy of Projection Techniques for Measuring Population Change* (1965), p. 79.

Arizona, Arkansas, California, Massachusetts, and Missouri. They showed that fewer than 50 percent of all projections were within 5 per cent of the actual population figures, while fewer than 80 per cent were within 10 per cent of actual data. Constant, arithmetic, geometric, and ratio techniques, with time variations, provided the set of techniques for testing. The states selected, having varying growth rates, provided a good range of characteristics for projection testing.

The conclusions from this study of technique accuracy were that projections err on the low side and that no one technique is "best" in all circumstances. The study does reveal which projection technique is "most accurate" under certain circumstances. The results are therefore useful in planning studies if one knows something about the probable changes in economic composition of the area to be analyzed.

Helen White's study of the accuracy of state forecasts for the period 1940 to 1950 revealed that the average percentage error between projections and actual data was 5 per cent for 10-year projections and almost 13 per cent for 20-year projections (1930-50).³ The greatest deficiency of the study is that its coverage of time is limited and shows no propensity of a particular technique to be "better" at one time than another.

Another set of forecasts made in the 1950's also reveals the slippery problems that seem always to confront estimators. The Nationwide Highway Finance (Section 13) Study forecast of automobile registrations was somewhat below (about 3 per cent) actual registration for the period 1953-60. In working out calculations of accuracy of the forecast for those years which had elapsed three writers indicated that 1959 preliminary registrations called for an increase of 2.1 million vehicles over 1958. The actual increase was 2.9 million.⁴ Even contemporary short-term estimates do not always adequately reflect the rate of change in the system.

The above citations of forecasting "errors" are not included here to imply a sense of pessimism or despair of ever obtaining usable estimates of the future. As mentioned earlier the art of forecasting is developing slowly and because of the possibility

of "forecast feedback" they may never be perfect instruments for piercing the clouds of doubt which hang over the future.

F. The Analysis of Forecasts

The case of forecasting as a policy assisting instrument of top management has been touched on. It is one that employs the broader questions of ends and means, and uses forecasting to help point toward the prospective quantitative bounds of those projects that fall within the purview of a particular policy or set of policies. The policy maker or executive is not necessarily concerned about the techniques used to derive those bounds nor does he necessarily concern himself about the technical development of the techniques themselves. What he is concerned about is the ability of a forecaster to provide him with a range of prospects and documentation that will enable him to act. To do this the forecaster is constantly engaged in efforts to improve techniques. However, because of the proliferation of interacting social, economic and political forces, problems of measurement become increasingly complex.

The development of the urban community is perhaps the most significant aspect of American life in the 20th century. Mass markets have emerged as a result of—and have, in turn, contributed to—financial, industrial, trade, educational, and cultural expansion. Forecasts have long been used as an aid in the making of private business decisions. In government, as long as it remained a passive force in the economy, there was little need for developing a highly sophisticated body of knowledge for decision making. Today all this has changed.

POPULATION ESTIMATION

A careful review of population studies was made to determine those that were based upon the most modern methodologies available at the time the studies themselves were made. A few of these are included here for illustrative purposes.

A. National Forecasts

1. Component Method

The component method is the most popular one for forecasting national population. It involves the components of population change—births,

³ Helen R. White, "Empirical Study of the Accuracy of Selected Methods of Projecting State Populations," *Journal of the American Statistical Association*, Vol. 49, No. 267 (September 1954), pp. 480-498.

⁴ E.L. Kanwit, C.A. Steele and T.R. Todd, "Need We Fail in Forecasting?," *Forecasting for Highways*, Highway Research Board Bulletin 257 (1960), p. 7.

deaths, and migration—in calculations of future population totals.

The traditional method of utilizing the component method is to obtain an initial enumeration of population. The total population figure is then reduced to age groups of 5 years, by sex, and, if necessary, by color. Each age group is then projected by using quinquennial survival rates that show the proportion of persons in a particular group who will survive for 5 years (cohort-survival).

In addition the number of births in a projected 5-year period is obtained by applying age-specific birth rates (births in a 1-year period per 1000 women of a specified 5-year age group) to the female population in the middle of the period (estimated as the arithmetic mean of the population at the beginning and end of the period) and multiplying the resulting annual births by 5. The product is then divided to obtain a breakdown by sex. This is done by using the existing sex ratio at birth and applying the ratio to the product of total births. Projections of birth products are obtained by using survival factors for each 5-year age group and reducing the births by this factor.

The data needed for the most accurate application of the component method are: annual mortality data by sex, color, age, and place of residence; recent life or survival tables; annual fertility rates by age, color, and place of residence of mother, by sex of child and by place of residence; annual net migration data by age, color, and sex.

Whelpton Study, 1936⁵

Six different series were presented in Whelpton's study. The highest projections assumed high fertility, low mortality, and 200,000 immigrants per year. The lowest projection assumed low fertility, high mortality, and no immigration. The table below shows the projected as per cents of actual U.S. population for these two series from 1940 to 1960.

Year	Low Series	High Series
1940	99.2%	100.3%
1945	95.8	99.4
1950	89.7	96.3
1955	82.9	92.4
1960	76.3	88.7
1965	70.1	85.6

⁵ P. K. Whelpton, "An Empirical Method of Calculating Future Population," *Journal of the American Statistical Association*, Volume 31, No. 195 (September 1936), pp. 457-473.

The projections are quite conservative. The three mortality series assumed that the expectation of life at birth in 1960 would be 65, 67, 68. The actual expectation of life at birth in 1960 was about 71. The three fertility projections assumed that in 1960 births per 1,000 women living through the child-bearing period would be 2,177, 2,000 and 1,750. The actual level for 1960 was over 2,700 and had been even higher in the 1950's. The series were inaccurate because of the forecaster's inability to foresee a rising fertility level when it had had a long history of decline.

Hornseth Study, 1953

In 1953 Richard Hornseth projected the population of the United States by 5-year periods to the year 1975.⁶ This study was a revision of an earlier work completed in August 1950.⁷ The revision is concerned essentially with fertility trends.

Accuracy of the Hornseth Study

Year/Series	A	B	C	D
1955	99.8	99.8	99.7	99.5
1960	98.5	98.5	97.8	96.5
1965	97.6	97.6	95.9	93.0

Most of the errors in the projections may be traced to inaccurate fertility forecasts. The death and immigration projections were quite good.

Greville Study, 1957

In 1957 T. N. E. Greville projected the population of the United States for the Social Security Administration.⁸ Because of evidence of an appre-

⁶ Richard A. Hornseth, "Illustrative Projections of the Population of the United States, By Age and Sex: 1955 to 1975," *Current Population Reports—Population Estimates*, Series P-25, No. 78, Bureau of the Census, U.S. Department of Commerce (Washington, D.C.: U.S. Government Printing Office, August 1953), 6 pp.

⁷ Jacob S. Siegel and Helen L. White, "Illustrative Projections of the Population of the United States 1950 to 1960," *Current Population Reports—Population Estimates*, Series P-25, No. 43, Bureau of the Census, U.S. Department of Commerce (Washington, D.C.: U.S. Government Printing Office, August 1950), 12 pp.

⁸ T. N. E. Greville, *Illustrative United States Population Projections*, Social Security Administration—Division of the Actuary, U.S. Department of Health, Education, and Welfare, Actuarial Study No. 46 (Washington: U.S. Government Printing Office, May 1957), 53 pp.

ciable amount of underenumeration in census enumerations it was decided to adjust upward the Census Bureau's estimates of 1955 population.

Accuracy of the Greville Study

Projection	Age Groups			
	Under 20	20-44	45-64	Over 65
I	98.4%	103.4%	103.9%	98.3%
II	98.3	103.4	103.6	97.8
III	104.0	103.4	103.9	98.3
IV	104.0	103.4	103.6	97.8
V	105.5	103.4	103.9	98.3
VI	105.5	103.4	103.6	97.8

Six series of projections were made. Projections I and II assumed low fertility; projections III and IV assumed high fertility; projections V and VI assumed very high fertility. The odd numbered projections assumed low mortality, the even numbered projections assumed high mortality.

B. State Forecasts

The general classes of techniques for forecasting state population have been constructed. These are: (1) simple methods, which include the fitting of various mathematical curves to total population figures; (2) ratio methods; and (3) component methods.

A. Simple Methods

William Spurr made his study of California's population and production in 1949.⁹ One of the techniques he used might be called a judgment-geometric method. Spurr observed that California's rate of growth had been almost constant since 1860, averaging 44.6 per cent per decade, or 3.8 per cent per year. Despite the consistency of past growth it seemed logical to assume that in the 1950's the rate of growth would begin to fall below the historical average because eventually population should taper off as an upper limit is approached.

A second method used by Spurr is an approximation of the Pearl-Reed logistic curve. The calculation of the logistic curve is complicated, so the

⁹ William A. Spurr, *Forecasts of California's Population and Production, 1950-1960*, Business Research Series No. 4, Graduate School of Business, Stanford University (Stanford, California: January 1949), 28 pp.

writer devised a shorter method of computing it.¹⁰ It involves the fitting of a curve to all decennial population data since 1860, rather than to only 3 points as in most older methods.

The results were that the 1950 projection was 98.3 per cent of the actual and the 1960 projection was 87.8 per cent of the actual population enumeration of California.

As a third technique Spurr plotted the population of California counties on a semi-logarithmic chart and fitted a logistic-type curve to the data. More than one-third (38 per cent) of the 1950 county projections were off by less than 5 per cent; 70.6 per cent were off by less than 10 per cent. Of the 1960 projections, 10 per cent were off by less than 5 per cent; 21 per cent were off by less than 10 per cent. The 1950 and 1960 records of accuracy of state projections were 98.6 per cent and 89.1 per cent, respectively.

2. Ratio Methods

To forecast a state's population by the ratio method it is necessary to project the state's percentage share of a larger region's population and multiply a population forecast of the region by that percentage. Iterative processes may be applied with the ratio method. For example, one might forecast New Hampshire's population by first forecasting the population of the United States, then forecasting New England's share of the U.S. population, and finally, convert New Hampshire's ratio of New England's population into a population number.

Census Bureau Ratio Method, 1952

A more complicated ratio method has been suggested by staff members of the U.S. Bureau of the Census.¹¹ In its study the Bureau divided the U.S. into nine regions. The ratio of the population of each region to the total population of the United States, and the ratio of the population of each state to the total population of its region, were computed for each decennial census year from 1920 to 1950.

¹⁰ See W. A. Spurr and D. R. Arnold, "A Short-Cut Method of Fitting a Logistic Curve," *Journal of the American Statistical Association*, Vol. 43, No. 241 (March 1948), pp. 127-134

¹¹ Helen L. White and Jacob S. Siegel, "Projections of the Population by States: 1955 to 1960," *Current Population Reports—Population Estimates*, Series P-25, No. 56, Bureau of the Census, U.S. Department of Commerce (Washington, D.C.: U.S. Government Printing Office, January 1952), 7 pp.

3. Component Methods

Zitter and Siegel Study, 1957

In 1957 Meyer Zitter and Jacob Siegel made a forecast of all states in the U.S.¹²

It may be concluded that while the projections possibly provide useful results, they do fail to set upper and lower limits which include more than one-half of the state projections.

Spurr Study, 1949

William A. Spurr's study of California's population and production used three series of population forecasts based upon the component method.¹³

Accuracy of Spurr's Component Method

Projection Type	1950	1960
Series I	99.1%	89.2%
II	98.4	88.8
III	98.9	89.3

Lewis Correlation, 1965

To evaluate a technique for forecasting migration it was decided to undertake a special study in connection with this project.¹⁴ Typically intercensal migration is estimated by determining change in an area's population between two censuses and then estimating the proportion of the change that is due to births and deaths. The difference or residual is then defined as the migration portion of the total change.

Six states are included in the analysis. Of these, Arkansas, Maine and Missouri are states which have experienced net out-migration while Arizona, California and Connecticut are states which have experienced net in-migration.

In every state the national income regression yielded a higher coefficient of determination than

either the arithmetic or geometric regression against time. The arithmetic regression for each state also yielded a higher coefficient than the geometric regression. It appears that migration comes closer to being an arithmetic than a geometric phenomenon. It also appears that it is desirable to take the level of income into account when projecting migration. The national income regression yielded very significant coefficients of determination for Arkansas, Arizona, California, and Connecticut. Thus, forecasts of national income would be helpful in projecting migration in these states, but the problem of finding adequate forecasts of income remains.

In 1962 the Connecticut Development Commission made a demographic analysis of Connecticut employing an interesting assumption concerning birth rates.¹⁵ The study included a suggestion by Richard Easterlin that birth rates are inversely related to the supply of young workers. Theoretically, a large supply of workers might be regarded as limiting future job opportunities, discourage family formation and, consequently, depress birth rates.

Lewis tested this possibility by correlating birth rates with the number of males aged 20 to 29 in the years 1920, 1930, 1940, 1950 and 1960 in the states of Connecticut, Maine, Massachusetts, and California. The coefficient of determination in the three New England States was .065, .051, and .005, respectively. The coefficient for California was .524; however, the correlation coefficient was positive, indicating that birth rates were directly rather than inversely related to the number of young males. Thus, the supply of young workers and its influence upon birth rates in California moves directly contrary to the theoretical pattern assumed in the Connecticut study. The very low coefficient in the New England States indicates that the supply of young workers had virtually no effect on birth rates in those states. Birth rates and the price adjusted level of net national income were also correlated to test the effect of income on fertility. Birth rates in one year were correlated with net national income the previous year. For Connecticut, Maine, Massachusetts, and California the analysis involved birth rates for each year from 1920 to 1960. For Arizona, Arkansas, and Missouri the test data were insufficient before 1927.

¹² Meyer Zitter and Jacob S. Siegel, *Current Population Reports—Population Estimates*, Series P-25, No. 160, Bureau of the Census, U.S. Department of Commerce (Washington, D.C.: U.S. Government Printing office, August 1957), 16 pp.

¹³ Spurr.

¹⁴ Kenneth Lewis, unpublished paper, "Correlation Analysis in Forecasts of Population Components," 1965.

¹⁵ *Population, A Demographic Analysis of Connecticut 1790-2000*, Connecticut Interregional Planning Program, Technical Report 131, Connecticut Development Commission (Hartford, Connecticut: November 1962), 105 pp.

C. Local Area Forecasts

Forecasting national and state populations, with all its complexities, is less difficult than predicting the growth or decline of local areas. There are many explanations for this difference, but much of the greater difficulty stems from two major problems. Firstly, there is a much more serious deficiency of data for local areas than for either national or state entities, and secondly, local planning at the community and neighborhood levels is either in its infancy or nonexistent.

In the preceding section on state forecasts the process of summing county population projections to obtain a state forecast was reviewed. The techniques of projecting county totals depended largely on linear relationships over time. In this section five forecasts made in connection with traffic studies have been reviewed. These involve more complex projection techniques than those discussed in connection with state forecasts and are of interest because, in one set of cases, future transportation needs were based on population projections and in the other set population growth and transportation facilities were considered to be interdependent.

1. Hartford Method¹⁶

Two projections of state population were made: (1) low estimates, based upon a straight line extrapolation of a semi-log plot for the period 1920-1960, but ignoring values at great variance from others; (2) high estimates, by fitting a nonlinear curve to the actual population (again ignoring large variance values) for the same 40-year period and extending this curve to the end of the projection period. A best estimate was selected within the high-low range.

2. Chicago Method

In the transportation study for the Chicago area, which includes most of Cook and DuPage counties, and small portions of Will and Lake Counties,¹⁷ it was determined that by 1880 all the forces making for Chicago's growth were operating at full strength.

¹⁶ *Hartford Area Traffic Study*, Vol. I. Prepared by Planning Research Unit, Division of Planning, Connecticut Highway Department in Cooperation with the Bureau of Public Roads (July 1961), 78 pages.

¹⁷ *Chicago Area Transportation Study, Three Volumes*. Study conducted under the sponsorship of the state of Illinois in cooperation with the Bureau of Public Roads, J. Douglas Carroll, Director.

These forces included: (1) regional and national population growth, (2) transportation advantages, (3) industrialization, (4) national policies, and (5) rapid and continuous technical innovations in manufacturing. Since 1880 these same forces have continued to work in favor of the expansion of the Chicago area. In the century between 1850 and 1950, U.S. population increased six and one-half fold and Illinois population increased ten fold. But the city of Chicago experienced a growth of 120 times its early population.

The population forecast was made with the following considerations in mind. In the past metropolitan areas have grown faster than other urban and rural areas. Because urban areas now contain such a large portion of the population, their growth rate in the future will probably be closer to that of the nation as a whole. Therefore, it was reasoned that the Chicago metropolitan area growth will about match the growth rate of the nation during the projection period.

Using the ratio method and employing estimates of U.S. population prepared by the Bureau of the Census, an estimate of metropolitan population was made. The study area's share of the metropolitan population was expected to decline from 85 percent in 1950 to 82 per cent in 1980. Historical data supports this contention, along with the fact that increasing proportions of additional population will settle in adjacent communities outside the study area.

Because the Hartford and Chicago studies have only recently been completed it is not possible to evaluate the methods employed. However, it would appear that judgment rather than technique played the crucial role.

3. Detroit Method¹⁸

Two studies of population were made in connection with the Detroit area. The first was a natural increase study; the second, one of future employment which was assumed to be a major determinant of net migration. These studies were compared with independent population forecasts of the Regional Planning Commission, and a final estimate was made based upon judgment.

¹⁸ *Report on the Detroit Metropolitan Area Traffic Study, Part II.*, "Future Traffic and a Long-Range Expressway Plan," study financed by Michigan State Highway Department, Wayne County Road Commission, City of Detroit, in cooperation with the Bureau of Public Roads, J. Douglas Carroll, Jr., Study Director (March 1956), 157 pp.

It was assumed that the population of a metropolitan area is largely based upon work opportunities. It was further assumed that the most active ingredient in the Detroit economy was manufacturing. It was reasoned that each production worker requires other workers in nonmanufacturing jobs. Therefore, if manufacturing and nonmanufacturing employment could be projected, this would produce the size of the total labor force, and the total labor force, being a known part of the total population, could be expanded to yield the projection to total population.

The best national indicator found was U.S. employment in durable goods industries. The correlation coefficient was .96, suggesting that as much as 92 per cent of the annual variation in industrial employment in the Detroit area was "explained" by factors that are also associated with fluctuations in national employment in durable goods industries. A 1975 forecast by the U.S. Bureau of the Census estimated national employment in durable goods industries to be 12.4 million persons. When entered in the predicting equation, this figure produced a Detroit area manufacturing employment estimate of 980,000 for 1980. Assuming manufacturing employment to be 46.5 per cent of the labor force (an average figure for 8 years), gave a total labor force estimate of 2,107,800 persons in 1975. Again, assuming 43.2 per cent of the population will be in the labor force (this is a current average percentage), the 1975 population estimate for the area was 4,880,000. By the conservative assumption of an annual population increment from 1975 to 1980 equal to the average annual increment from 1950 to 1975, the 1980 population would be expected to reach 5,253,000 persons. Comparisons with independent forecasts coupled with an explicit definition of the study area produced a final population forecast for the area amounting to 4,400,000 persons.

4. Pittsburgh Study

The Pittsburgh area transportation study was completed in 1963.¹⁹ The study proceeded on the assumption that a land-use forecast could be derived from projections of population and employ-

¹⁹ *Pittsburgh Area Transportation Study*. Study conducted under the sponsorship of the Commonwealth of Pennsylvania, County of Allegheny, City of Pittsburgh in cooperation with the Bureau of Public Roads, Volume I, *Study Findings*, November 1961, 120 pages; Volume II, *Forecasts and Plans*, February 1963, 177 pp.

ment; and, given the land-use forecast, projections of vehicle registrations and transportation demand could be made.

Historical data indicated that net outward migration had been the cause of Pittsburgh's relatively slow growth rate. The forecasters observed that in recent years energetic efforts had been made to attract a greater variety of industry and commerce to the area. Success in these ventures was expected to increase employment and reduce the chronically large number of jobless workers. At the same time, given a reasonable level of local prosperity, the employment base was expected to broaden automatically as a larger share of total employment comes to consist of service and consumer-oriented activities. Together these trends would make it less necessary for the younger generation to find employment outside the area, and thus reduce net outward migration. Accordingly, the Pittsburgh growth rate was expected to accelerate through the projection period.

5. Baltimore-Washington Study

The object of the 1960 Baltimore-Washington study was to project and analyze future patterns of population, employment, and automobile traffic growth in the Baltimore-Washington region to the year 2000.²⁰

Mathematical models incorporating various factors underlying the growth and location of population and employment between 1948 and 1957 were developed. By means of computer simulation techniques, data were processed on various growth patterns resulting from alternative public policies.

It is not possible to evaluate the simulation techniques scientifically at this early date. However, some intuitive comments may be in order. Because of the success of the computer in processing masses of data, one would expect the computer to be a great aid to the forecaster provided the programming is carefully designed.

FINAL COMMENT

In summary it may be of some usefulness to comment on two aspects of the forecasting problems

²⁰ *Baltimore-Washington Interregional Study, Land Use and Transportation*. Technical Report No. 7, Baltimore Regional Planning Council, National Capital Regional Planning Council (November 1960), 130 pages.

revealed in the foregoing analysis. First is that of the accuracy of past forecasts. Second is the influence of governmental organization or cooperation upon future forecasting activities.

A. The Problem of Accuracy

There has been increasing reliance upon relationships and improvements in data collection methods and processing, as well as in the use of these relationships for pinpointing changes in the direction of population shifts. For example, improved mortality statistics and fertility statistics and knowledge of the socio-economic factors that influence these variables may well be the very tools needed to provide more accurate national population forecasts. This is, of course, obvious. But strangely enough there has not been a coordinated concentration of effort to develop a strong socio-economic research program designed to uncover patterns of human fertility response. Until this is done, the assumptions used by forecasters for making population estimates may continue to be rather wide of the mark.

State population forecasts are plagued by the net migration problem and by faulty national forecasts (where the ratio method is applied). Both national and state income data are valuable indicators of state migration shifts and more and better research in income analysis would result in greatly improved state forecasts.

Local area population forecasts are hampered by two major deficiencies. On one hand, data for programming planning studies are insufficient and often of poor quality. In addition annexations of territory occur frequently and this requires that time series be adjusted—a laborious, difficult task. On the other hand, lack of effective community planning instruments means that no one can be very sure what area within a city, precinct, or ward will change in response to what stimulus, or at what time. Thus, the forecaster must depend almost solely on his particular knowledge of the area's history and market structure to formulate assumptions for his forecasting model.

B. The Problem of Governmental Cooperation

Since the days of colonization the people of the U.S. have been "future oriented." It was 1933 that the first comprehensive peacetime steps were taken to assess the strengths and weaknesses of the nation's resources. On a smaller scale, such assessments were continuously made by entrepreneurs in the

private sector from the earliest days of the nation's history.

In the days and years since 1933 an ever-expanding, albeit sporadic, organization of planning processes has been applied with the Federal administrative agencies, and there has slowly emerged a national acceptance of the Federal role in planning which has served to strengthen this development.

Politicians, academicians, and many knowledgeable laymen have despaired of ever achieving a meaningful restructuring of metropolitan community government. What this means, if they are right, is that it will be necessary for planners to work with the existing layers of metropolitan organization if they are to achieve progress in providing the services and facilities that are required in these population centers.

Still, national planning activities are relatively meaningless without the effective organization of and knowledge about the many local communities in the U.S. and how they combine to form unique sets of states and regions. To encourage local participation in national resource and facility planning Federal legislation has relied upon financial inducements in the form of whole or partial payment of research, planning, and/or operational costs. Some of these induced or forced planning efforts have resulted in massive cooperative undertakings, e.g. the Interstate Highway program, but there has not been marked improvement in planning and cooperation among metropolitan communities within any one SMSA. It is just here that many of the major problems of the nation germinate and grow. As has been shown in sections of this study labelled "Local Area Forecasts," it is almost impossible to forecast accurately population or land use quantities in sub-segments or metropolitan communities unless there is an effective community organization devoted to the overall planning of human development needs.

From the standpoint of the forecaster, the multiplicity of governments in even the smallest SMSA's makes the most general assumptions about future growth or decline extremely hazardous. On the other side of the coin is the fact that Federal agencies and programs designed to intervene in local community development are sometimes ill-designed and coordination among them almost non-existent. Thus, the forecaster cannot assume any continuity of policies that have deep and widespread impacts upon the land-uses in most urban communities. He must relate future prospects for land-use to the often quixotic relationships in the private real estate

market and hope for the best. Doing this, he admits into his scheme of adjustment incremental changes that are the stock in trade of the market system, but rules out the likelihood that drastic or revolutionary changes will upset the balance of present market processes.

Despite the pessimism that has been expressed about metropolitan governments' ability to gain a basis for unity of action there are some bright lights on the horizon. The technological revolu-

tion, in its most sophisticated form, has only just begun. In the fields of land transportation, technology has not yet had any appreciable impact. But, if one uses only a meager amount of imagination in projecting what the situation may be 20, 50, or 75 years hence, the possibilities for change mount rapidly. The point is that technology can and does infiltrate the market system, causing the creation of new political orientations and the regrouping of economic alliances.

DISCUSSION FOLLOWING PAPER GIVEN BY DR. ROBERT W. PATERSON, PROFESSOR OF ECONOMICS AND DIRECTOR, BUSINESS AND PUBLIC ADMINISTRATION RESEARCH CENTER, "FORECASTING HIGHWAY DEMAND."

MR IRWIN: I would like to comment on more recent publications and work which has been done by the Bureau of Census other than those you describe in your analysis. I am not going to claim that the more recent work is more accurate. We have publicly disclaimed the ability to say exactly how accurate our projections are, and we have said that we are unable to assign confidence limits. It would be very desirable if we could assign confidence limits or if we could name one projection as the most likely. With respect to the national projections, it is certainly very true that the projections which you analyzed showed very low results, really shockingly low. You could have mentioned another one in 1947 where the highest of four projections—uncomfortably below the actual 1960 population which emerged. Whelpton, a brilliant analyst in the fertility field, as a result of this failure, decided that error lay in not taking into account the behavior of women by so-called cohorts. We had not taken into account the fact that women had gone through a long period where they delayed having their babies due to the impact of the depression and of the war. So he went into a very detailed analysis and outlined an approach which we are now using.

However, in 1958 we still used the older method of projecting and came up with a set of projections more or less along the lines of the model which Mr. Paterson mentioned. In this case we were very high. We put out four series and the highest series turned out to be very very high. However, the low series projected lower than the present population. The present population falls between the lowest two series. For the highest series you could criticize us very justly and say that we never should have named

such a high series. We then changed to the cohort fertility method using tables developed by Whelpton to produce the present set of projections which I talked about at the last symposium. This is our publication No. 381. I would like to say the outcome was very different but it was not; we are still high and my most recent publication contains a comparison over the last two years of the projections and the current estimates. These projections were based on 1966. We would like to have the population fall somewhere in the middle range, but the lowest projection which we made does encompass the present experience. We would like to have it go down the middle. So we do not feel we have solved our problems. The reasons for this error go back to my comments this morning: we are relying on past trends. In the case of the earlier study, we were making a fairly detailed analysis. When that failed, we made an even more detailed analysis, taking in another dimension, a clearly important variable. We have made a significant attempt to get away from that by using Birth Expectation Data. There have been three major, national studies asking when and how many babies women expect to have. There were studies in 1955, 1960 and 1965. We have been quoting these to support our projections. We kept the projections high partly because the birth expectation answers are high. So again we do not feel we have the answer yet. The idea of trying to project ahead using birth expectations is a good one, I would say.

Finally, with respect to the local projections, we recently did a set of projections for all metropolitan areas. Here we are in beautiful agreement because our publication says that the rate of metropolitan

growth relative to non-metropolitan will taper off in the future, and that large metropolitan areas such as New York, Chicago, and Philadelphia cannot be expected to maintain the kind of growth they experienced in the past. What is indicated is a spreading out of population growth to smaller cities.

A. I will, if I may, make just one comment. During your statement I was speculating about the reasons for past failures and wondered will we not, twenty-five years from now, be speculating as to reasons for the failure of current long-range forecasts? Many of the aggregate measures of economic well-being were developed during the 1920's and thus are of very recent origin. Our first regularly published estimate of GNP dates from 1929. It also happens that shortly thereafter we ran into a depression which upset a number of relationships of the past in terms of the future. But forecasters, like most other people, tend to behave in terms of what their perception of the past was and it took a few years during the thirties to recognize the fact that things had been upset and to discover new relationships. Now we have been living through the fifties and the sixties, and there has been a prolonged period of prosperity. During this time some changes in behavior have taken place. Not many people in the general public realize, for example, that the birth rate has declined since 1957—sporadically, to be sure. MR. IRWIN: I think the fertility rate reached a high point in '57, but actually the number of births did not start dropping until 1962.

Now I would like to comment on the state projections which are our most recent work. We have tried a significant new development by treating migration separately as out- and in-migration. I do not think I can claim we are more accurate. What we are doing is more logical in that we first project out-migration from all states. We then develop a national pool of migrants which we allocate back to the states as in-migrants. By this method we have the same number of out-migrants as in-migrants, a logical relationship. That might not seem very important; but when you work with net migrants, the numbers come out very very differently because the big states that are growing demand more and more in-migrants and the small states that are losing are unable to produce the demanded out-migrants, creating a very serious difficulty.

As to Easterlin's thesis, which is a very important one, it is my understanding that Easterlin maintains that since 1953, the increase of income to young men is low relative to the increase in income of all people. So young workers are in a pro-

gressively deteriorating situation income-wise, with respect to the entire economy. This he claims explains the decrease in the birth rate. There is a good deal of interest among demographers in the theory.

I do not think there is one answer. The factors that influence the fertility rate of the country are very complex and go beyond just income. They come around to the things we were talking about this morning; that is, the world situation and how people feel about it, and how people in this country feel about the national situation. You have a big family, I feel, when you have a positive attitude, when you have a good feeling about how things are going. I cannot quantify that, and I am sure it can be attacked and criticized.

A. Nevertheless, I think our behavior has been tempered somewhat. But, it is likely that we are forecasting population change without proper regard for fertility-birth shifts since the latter 1950's. This will lead us to overestimate population in the absence of any adjusting change in the 1970's, 1980's, etc.

MR. IRWIN: That is what it has done. We have been unwilling or unable to say that the high birth rates in the immediate past are out. In 1958 when we made these projections, the range of the four series was very wide. It is not much to our credit that the present truth falls within those ranges, because they are so wide. Now that we look back, the high one was impossibly high. We predicted an impossibly large family size. I must add that they were not done by the family size methods you are doing now. The crucial input into the present projections is the estimated average family size which will be obtained by cohorts of women. And the low series, which are close to current data, assume that there would be 2.45 children per family on the average. In 1958, the high assumption was something like 4.3 children per family, well above what any cohort of women had ever achieved in the United States. We selected 2.45 as the low, because we could not believe that any cohort of women would go below the lowest we have ever had. This low of 2.2 occurred in the cohort of 1910, a group of women who reached twenty in 1930. At that time the depression, the biggest economic dislocation this country has ever had, started. Previously, there had been a long period of declining birth rates. The women delayed their child bearing as a result of the depression. World War II came and all the men were removed from their wife's loving arms for two to four years. When the war was over and the baby boom started, these women were too old to have

very many children. They averaged 2.2 children per woman. Then we made our crucial assumptions. (I was not there, but I would have said the same thing.) We said, "2.2 is very low; it just will not happen again. Therefore, 2.45 is going to be our low number." In 1969, the value is 2.45. At least the actual level of births is at this low assumption. There is a further postscript to add. The great argument among demographers now is: is family size really dropping to 2.4 children per woman or are the present groups of young women merely delaying the formation of their families? No one has a really qualified answer to this argument. It is now very clear that average family size is dropping. The cohorts that had their children through the fifties will reach about 3.2 children per woman. By examining current rates and the work we have done, I could almost certainly say that some cohorts will drop to at least 2.7 children per woman; but there is quite a range as to what could happen. It could be that these women are not delaying their children, but will never have them, in which case it could drop to 2.2 children per woman. Then the population would level off, and we would never reach three hundred million in 2000. That is why I objected this morning to the predicted doubling of the population. To double the population in 2000 would take approximately 5.0 children per woman, and that is impossible.

LT COL HAIDLER: The other end of this thing: the life span is increasing; your birth rate is going down and so is your death rate.

MR. IRWIN: But that does not really have that much impact. From the point of view of population projections, changes in the death rate for the last ten or fifteen years are not an influence. The death rate has been relatively constant, and the changes in the past twenty or thirty years do not really change the population levels very much. We are off on the births, not the deaths.

MR. SIMON: Getting back to the business of highway demand, you made a statement right at the start that people always underestimate long term needs; you hardly know of a case of overestimation. This brings to mind an article in Fortune Magazine, sometime back, having to do with toll roads. Certain state legislatures, seeing the great success of the Pennsylvania Turnpike and the New Jersey Turnpike said, "We want one!" and took action to build one in their state. The net result was a number of super highways in states where there was not sufficient traffic demand to pay off the bonded indebtedness. Due to the rising tide of

tourism, continued high prosperity and automobile production, most of these "non-pay-off" roads have finally become paying propositions. In the first analysis, however, they were built as a result of overestimation of need based upon the desire to make a fast buck.

A. I have not kept up with this as thoroughly as I possibly should. I know seven years ago some of the old toll road systems ran into trouble. I think it was revealed that they were all making money, and now a number of states are seeking to use this technique (toll roads) to develop facilities to handle increases in traffic. We have a controversy which is interesting because of the revenue expenditure problem that most of the states are running into right now. States simply do not want to increase, for instance, the state gas tax significantly to the point where they can undertake to build a new road needed in the state primary system. As a result they are thinking about toll roads. But experiences with toll roads indicate that some are making money for the states, paying the bonds and leaving a small residue in each case. This was not true in the early days: for instance, in the case of the West Virginia turnpike. The real problem relates to forecasting. How do we explain a rapid increase in mileage as portions of the Interstate system are completed? As portions of the Interstate system were completed, there was a spurt in the generation of traffic. It was not caused by new registrations but was simply caused by existing users. If it takes you, let us say, three and a half hours to go from where I live to St. Louis and you reduce the time element to go there to an hour and a half, you generate a new market. This market generation, in terms of time-saving, is more important than mileage. That is why some people drive two hundred and fifty miles via Interstate rather than a hundred and seventy-five miles via primary or secondary roads in order to get to a given place. This is an important economic fact.

DR. BENNET: I would like to get back to your original topic, that of highway demand, the highway demand function in which you imply that population is the major or chief explanatory variable. Is the relevant demand function you are looking for really the demand for the whole package of travel services?

A. There are a host of factors which really play on this. It is a question of whether you can get a real good simple denominator of demand which implies lots and lots of things about all the other variables associated with it. It is true you can move

from horses to automobiles and so forth as technology develops. But there are also effects. For example, how much truck usage of our highways today is related to the size of the community? In this sense highway demand is a function of the size of the community.

Q. Were you measuring demand for actual physical highways or the services that come from these highways?

A. Demand for the physical facilities.

MR. PYKE: It seems to me here is an area which is somewhat unique in forecasting in the sense that there appears to be physical limits. For example, it would be difficult to double the number of highways in downtown New York City. It would be even more difficult in Los Angeles County. What does one do with the cars at the terminus? So, my question is: has any study been made of the limiting parameters?

A. The Bureau of Public Roads during the period of its re-birth, which was about six years ago, I think, did some novel things. It held a conference with people who were interested in the total transportation problem—i.e., the allocation of resources to highways versus other kinds of transportation systems. One of the elements involved was the relationship between legal codes with respect to highway use and limitation on use which becomes immersed in what we Americans consider to be a right. One of these American "rights" is our right to drive a car anywhere we want to at any time, even though it may be stacked up in lines fifteen miles long. For some reason we do not behave rationally. We will line up driving from Washington and Baltimore with cars going bumper to bumper at fifty miles an hour and then take two hours off because two of them have an accident and we wait and wait and wait. We will stew a little bit, but nevertheless we will wait. The cost of such behavior is far and above the time that is warranted by our desire to use our private automobile under any circumstances. It was suggested at one of these meetings that one way of reducing the jam, particularly at peak hours in metropolitan communities, was simply to limit the right to drive to certain proportions of the population at certain times. I tell you, the lawyers threw up their hands. "You will never do this first." They said we might call a conference of city managers and city mayors and suggest this was one way out. A few mayors were consulted about it and they said no; that they would never do that. But you see, you could limit access to certain highways in the morning by asking the housewives not to go to the

grocery store until ten o'clock in the morning and you could resolve part of the peak hour demand situation which is really causing this tremendous bottleneck on facilities. It is the peak hour; nothing else.

MR. HARRIS: If you want to talk about Los Angeles, one of the problems I would say involved in that is we have a series of mountains scattered around the city and for some reason or another we cannot seem to get freeways built over the tops of those mountains, so we are constrained in little pockets. If you try to get from one pocket to another, that is where you get into trouble, and then everything backs up all over the place. Now, I do not understand personally why we have such difficulty in planning freeways in a city of that size.

A. I think it is very, very simple: people do not live where they work anymore. One of the problems associated with these big communities is that the older industrial districts used to be located where people lived. Now you have people, for one reason or another, who have vacated the city. That has caused people necessarily to travel five to ten miles to the place of work. In many cases, people are crisscrossing from home to work which causes a tremendous increase in the demand for facilities. We need—this is really what I would talk about in planning—to restructure mainly our metropolitan communities in such a way that people can live where they work. Then we would resolve some of this tremendously high cost of living, and it's not going to come about tomorrow. We just don't tell people what to do.

MR. IRWIN: I was wondering, Dr. Paterson, about enlarging on the remarks you made concerning the relationship between place of residence and place of work. It would seem to me, from the contacts I have had with the Bay Area Transportation Study and the Los Angeles Regional Transportation Study, that the Bureau of Public Roads has put a great deal of emphasis on population projections because of the question of the place of work as opposed to place of residence. Working from what was learned in those studies, what do you see as a pattern for the future with respect to this? Do you see workers living closer to their place of employment?

DR. PATERSON: These kinds of studies are socio-economic analyses of transportation behavior. I think in the last two and a half years the Bureau of Public Roads has taken a very active interest in them. They have even established a national analysis committee to advise the department on that. In this type of study, it is hoped that it would

be possible to reveal what would happen, in terms of location of workers, as a result of the application of a given public policy. Another thing has come out of all this effort. I suspect the key to any major change in the organization of our community has to come through the development of greatly increased effort in comprehensive planning means for states. You are all familiar with the fact that out of the requirements of each state to qualify for certain funds is that a state must engage in comprehensive planning. What this means at the moment is that each major division in a state government (education department, highway department, hospital department, mental welfare, mental health, welfare, etc.—all of the state departments which are engaged in capital expenditures and which are likely to change the economical environment of cities within the state) must work out a long range plan. I think the dates are 1965, 1980, and 2000. It does not matter if they are accurate or not. It is the process of planning that is really going to tell us more about the willingness of the general populace to support various kinds of efforts on the part of the states. This really does mean something and the first step is to crack the heads of the departments of the states; and the federal government is using the stick to force the states to set up agencies which will be responsible for organizing this planning operation. Right now it is in a state of chaos. I think Wisconsin probably does the best overall job of the states in this particular field, but it is a brand new activity for most states. We have only been engaged in it for about three years. It is not even well organized yet. But five or ten years from now, we are going to be in shape to at least assess and analyze what the most significant parts of the comprehensive plan are. At that time maybe we can, by building certain kinds of new facilities over here where the workers happen to be located today, encourage establishment of residence-employment into circles within populated communities where cross traffic patterns now prevailing will be reduced.

MR. IRWIN: I would make one short comment which I think may be of interest: it occurred to me while you were speaking, since 1960 everything seems to

be pointing toward more population growth in medium sized cities, and less population growth in large ones. Maybe the response of the population is, as you mention, that by moving to a smaller metropolitan area, you are going to cut down your commuting. The choice of residence in a big city is often dictated by one set of considerations—the wife wants certain schools or she wants certain characteristics of the neighborhood. The job location is fixed by others. I wonder if the decision to move will be influenced by the fact that in a smaller complex, job and house will not be far apart.

DR. PATERSON: I think you are getting into a different area which has to do with the economics of services by city size. You know we do not really know very much about what the optimum size city is. Some people say it is 250,000; others say 400,000; and some say a million. A study we are conducting now seeks to determine whether or not the cost of services per capita does rise as the size of the community increases. Preliminary data indicate it does. As you move from city size 300,000 to 500,000 to a million population, the cost of services rises something like twenty-nine percent. Now there is a lot of difficulty with this particular estimate. I hesitated to mention it primarily because nobody can really tell us what the revenue expenditures are from year to year. The definitions are changed. A problem is whether or not the figures include grants and aid from the state or federal government. A lot of federal aid is going directly to cities and not to the state and then to the city. So we really do not know precisely what revenues are. You say, "Well, the cost of maintaining a sewer system in the city of St. Louis is X." We do not know whether it is "X" or not because we do not really know whether or not certain other revenues have been applied to it, in which case the cost would rise. So, we have a tremendous problem here in just setting out what our base of comparison is and in getting the proper data which will enable us to make some analyses. I do not see any hope for improvement in this except going through the data carefully and painstakingly, hoping you can create some series that are reliable.

Technological Forecasting in an R&D Laboratory

by A. B. Nutt

Technological forecasting should be a prime component in carrying out the planning, programming, budgeting and program implementation functions of an R&D Laboratory. This paper will describe some actual uses of forecasting in the Laboratory with which the author is associated and will point out others, both explicit and implicit, that should be considered if the full utility of forecasting as a management tool is to be achieved. Actual techniques for technological forecasting will not be addressed in depth. Emphasis will be placed on the utilization of these techniques as an essential part of the R&D management process.

To put the presentation in context, the mission of the Laboratory will be briefly described as well

as the makeup of the resources made available to it for carrying out the mission. The relationship and essentiality of forecasting to each portion of the Laboratory's planning/programming/budgeting cycle will be indicated, although the description of detailed relationships will be confined to those areas where technological forecasting is performed by Laboratory personnel.

At this point, some definitions are in order. General agreement seems to have evolved in the literature around something close to the following definitions for types of technological forecasting:

Technological forecasting is the probabilistic assessment on a relatively high confidence level of future technology transfer. Exploratory technological forecasting starts from today's assured basis of knowledge and is oriented towards the future, while normative technological forecasting first assesses future goals, needs, desires, missions, etc., and works backward to the present. The subject of both types is a dynamic picture of a technology transfer process. Technological forecasting may be aided by anticipation and may "harden" to prediction.

The mission of the Air Force Flight Dynamics Laboratory encompasses the exploratory and advanced development for the entire aerospace flight vehicle, with the exception of the propulsion and airborne electronics (avionics) systems. It includes the planning and executing of the Air Force exploratory development programs in flight vehicle structures, aerodynamics, flight control dynamics, environmental control, crew escape and recovery and mechanical systems.

The Laboratory is the largest of the ten Laboratories of the United States Air Force which are under the Director of Laboratories in the Air Force Systems Command. The Laboratory has the equivalent of close to 1000 assigned personnel, with almost half of these being military and civilian scientists and engineers. It is located at Wright-Patterson Air Force Base, Ohio, and occupies 32 buildings of

* Ambrose B. Nutt was born in Milwaukee, Wisconsin, on March 16, 1920. He holds a Bachelor of Science degree in aeronautical engineering from the University of Michigan and a Master of Science degree from the Ohio State University. He is a graduate of the Air Force Command and Staff School and of the Industrial College of the Armed Forces.

His work experience includes 26 years of engineering and R&D management as project engineer, Section Chief, and Branch Chief in the Aircraft and Air Force Flight Dynamics Laboratories at Wright-Patterson Air Force Base, Dayton, Ohio, in the fields of structural analysis, flight vehicle mechanical and crew escape subsystems, and crew station design. As Chief of the Special Projects Branch in the Aircraft Laboratory, he had management and engineering responsibility for the Air Force Supersonic Military Air Research (SMART) Track at Hurricane, Utah, and for the Air Force Crew Escape Program. Since 1964 he has served as Senior Staff Plans Engineer in the Air Force Flight Dynamics Laboratory. His current research interests are in the applications of operations analysis techniques to research planning and management. He is the author of several recent papers in the field of R&D management and is co-author of the book, *Quantitative Methods in Technology Planning*, published in 1969 by the MIT Press.

Mr. Nutt is an Associate Fellow of the American Institute of Aeronautics and Astronautics; a member of the National Society of Professional Engineers, the American Ordnance Association, and the Research Society of America; and an affiliate of the Professional Group on Engineering Management of the IEEE.

various sizes devoted to a variety of purposes. These buildings include subsonic, supersonic, hypersonic and vertical wind tunnels, flight control simulation facilities, full scale landing gear test facilities, and full scale structural, environmental and acoustic test facilities and of course office space. The replacement cost of the physical plant is estimated at approximately one hundred twenty six million dollars (\$126,000,000). The current annual budget is in the neighborhood of thirty million dollars (\$30,000,000).

The planning/programming/budgeting (PPB) cycle of the Laboratory is divided into the following phases:

1. Long Range Plan Preparation
2. Preparation of Laboratory Technical Objective Document
3. Short Range Program Preparation
4. Budget Preparation
5. Program Implementation
6. Program Evaluation
7. Feedback to Long Range Plan

This is a conventional PPB cycle. However in actuality, the Long Range Plan part of the cycle looks more like the following (note the emphasis on future systems due to the vehicular-oriented nature of the Laboratory program):

- a. Long Range Plan Preparation
 - Define Goals
 - Operational Capability Objectives
 - Future System Characteristics, Options, and Alternate Approaches
- b. Determine Technology Needs
 - Implications of the Goals
- c. Assess Capabilities and Future Plans of the National Technical Community
- d. Assess Laboratory Internal Capabilities and Ongoing Program
- e. Prepare Plan

The sources of data for defining Laboratory goals include the Joint Chiefs at Staff Research and Development Document, which considers the R&D goals of the Army, Navy and Air Force to the 1987 time period. This document includes forecasts of the future military threat environment and assessment of the military and technological needs to counter this threat. Forecasting thus precedes the very initial effort to prepare the Laboratory long range plan.

Another source of guidance for the Laboratory in this regard is the Air Force Long Range Plan which provides an assessment of the world environment in which the Air Force is to operate through the 1981 time period. Again, the strategic appraisal

portion of the document provides a forecast to be used as a basis for conceptualizing future Air Force vehicle concepts. It also provides a description of the near term (five year) strategy and capability of the Air Force, based on existing policies, commitments and budgeting.

A major source of guidance is the Air Force Systems Command Planning Activity Report which gives the status of development planning activities in the Air Force, each of whose potential worth has been determined by the Air Force to justify R&D resources expenditures. These activities may include conceptualized flight vehicles, proposed sub-systems, proposed technical facilities and major technology effort areas.

A set of documents known as Technology Needs is furnished to the Laboratory (and is used by the Laboratory as another source of guidance for program planning) by the Systems Divisions, the organizations which develop new weapon systems for the Air Force. These documents are used to advise the Laboratories of specific technology deficiencies which, in the opinion of the systems analysts, may delay the planning, development or acquisition of qualitatively superior weapon systems. These are matched against the technical barriers addressed by the Laboratory to detect areas of technical deficiencies not previously known to Laboratory personnel.

Also used as a source of data for the Long Range Plan is the Laboratory's existing issue of its Technical Objectives Documents, one for each major technical domain in the Laboratory, each of which contains the following:

1. Quantitative statement of specific technological objectives in the area in question.
2. Assessments of the present state-of-the-art in the area.
3. A technical forecast for each objective, based on an assessment of the capability predicted for the national and international technical communities' current and planned efforts in the area (insofar as they can be determined) and on the past trends in technology in the area of the objective.

No specific technique of forecasting is required to be used, so that the full spectrum of techniques may be used in preparing the documents. All fall within the general major areas of technological forecasting currently available, namely:

1. Trend extrapolation—Developing past time-series parameter data and projecting with changes



SIMPLIFIED OPERATIONAL DIAGRAM

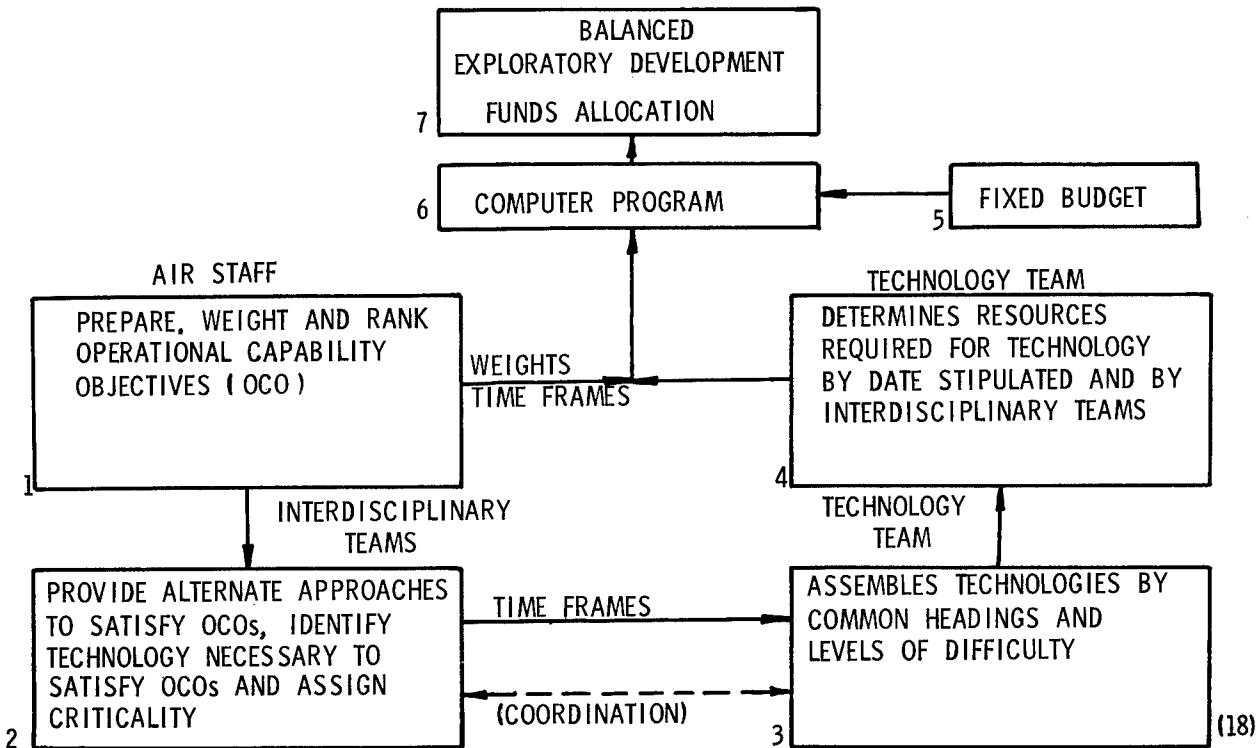


FIGURE V-1

in rates of advance or as an asymptote to a known limit.

2. Precursor events—Using an indicator that leads the parameter in question, then knowing the lead time and the time series data for the leading indicator, predicting the approximate future values of the parameter.

3. Envelope curves—A master curve tangent to a series of S curves for a set of related parameters which tapers off horizontally to known limits.

4. Analogies—Determining analogies to natural phenomena or past technological patterns of development.

5. Forecasts based on described future environments and from these extracting technical needs.

The Long Range Plan based on guidance documents and a technical forecast, could be used in combination with an assessment of current progress to update the forecast itself, since new objectives have been discerned and new technology needs

generated. However, thru use of operations research techniques, the Laboratory adds one more ingredient before updating the technical forecast, and that is the individual detailed plans for each task in the Laboratory. These are individual technical forecasts which have hardened into technological predictions. The manner in which this is done will be indicated later in the paper.

The next steps in the Laboratory PPB cycle are the preparation of the short range plan and the preparation of the budget. These are combined in the Laboratory by using as an aid to management, a program known as Research and Development Effectiveness (RDE). Here, thru use of a mathematical cost effectiveness simulation model (plus as an option, a linear programming technique), the proposed program is evaluated against the systems objectives as stipulated in the higher headquarters guidance documents and against the Technical Objectives of the Laboratory, to optimize

RDE FLOW DIAGRAM

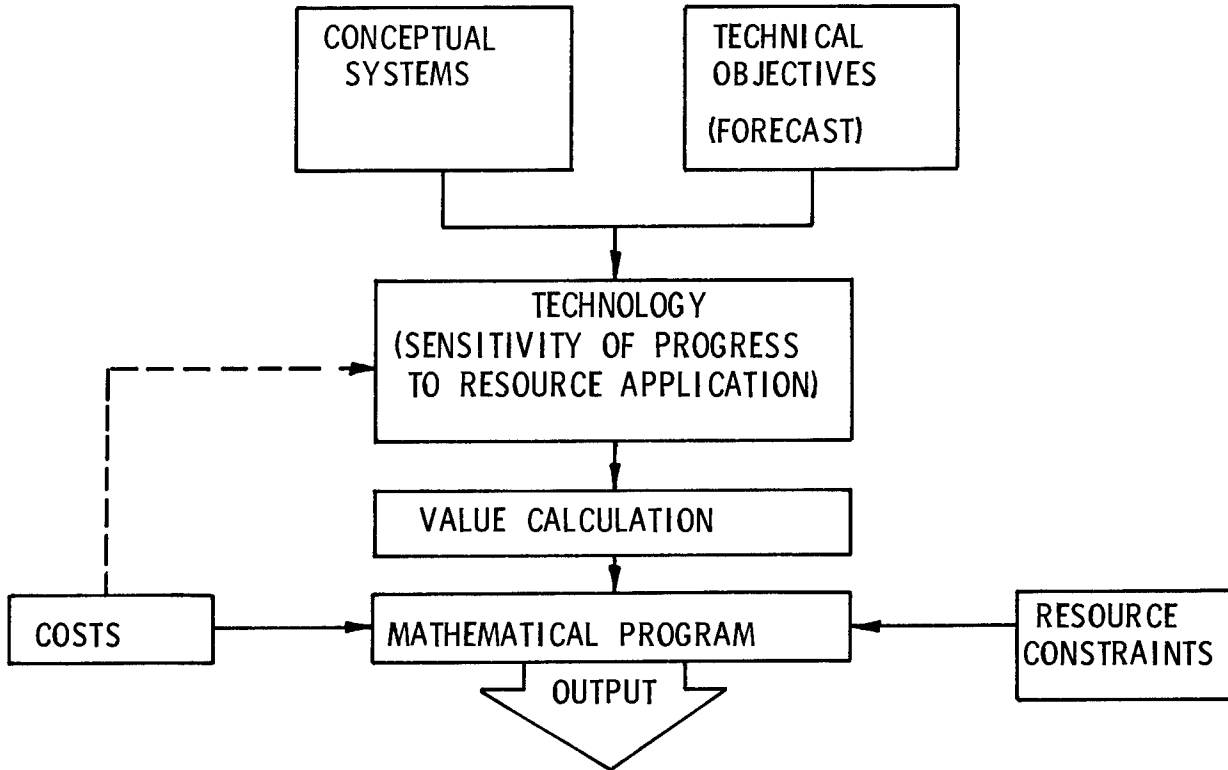


FIGURE V-2

the effectiveness of the Laboratory's exploratory development resources expenditures. The forecasts in the Technical Objectives Document now become individual quantitative goals for each task engineer in the Laboratory. The requirement in the RDE system for preparation of work packages for each task, along with the determination of relationship of resources available versus time required for each effort provides the "hard" prediction of task progress. This data becomes very valuable as a tool for updating the technological forecasts in the Laboratory Objectives documents as was mentioned previously herein. Tracking of past year's progress for each task as can be done in the RDE Program also aids in improving the accuracy of the technical forecasts.

An example of the implicit use of forecasting was the test of the DDR&E TORQUE resources allocation (Technology or Research Quantitative Utility Evaluation) system in 1968, directed by Head-

quarters, USAF, with the Air Force Flight Dynamics Laboratory as the test Laboratory. Referring to Figure V-1, in block 1 of the diagram, the definition of the Operational Capability Objectives (OCO's) required an initial time-phased forecast of the enemy capabilities. This was done in order to put the OCO's in individual time frames in terms of predicted needs. In blocks 2, 3, and 4, the process of technology delineation and costing required a match of need against forecasted technical possibilities in order to certify possible system development dates.

Figures V-3 thru V-7 are examples of technological forecasts prepared in the Laboratory which can have a significant impact on Laboratory planning and operations.

Figure V-3, showing the expansion trends in flight envelopes, would carry the following implications to laboratory management relative to future re-

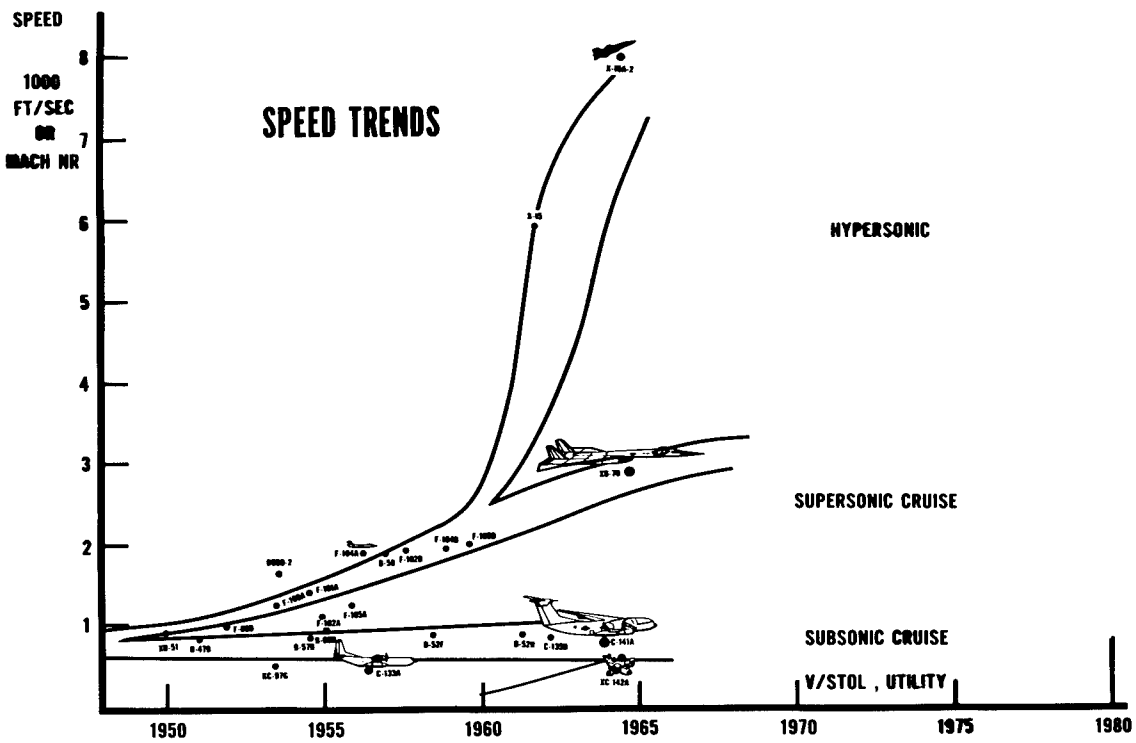


FIGURE V-3

TRENDS OF AIRCRAFT NOISE

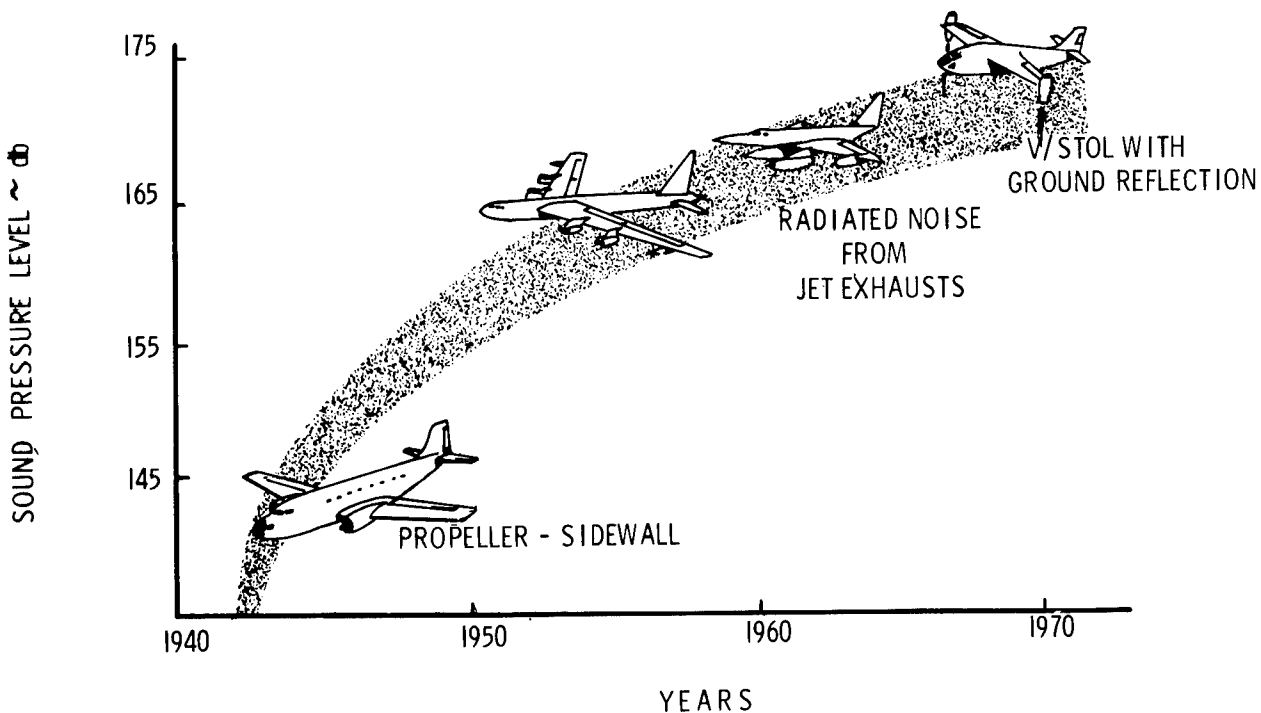


FIGURE V-4

VARIABLE GEOMETRY COMPLEXITY

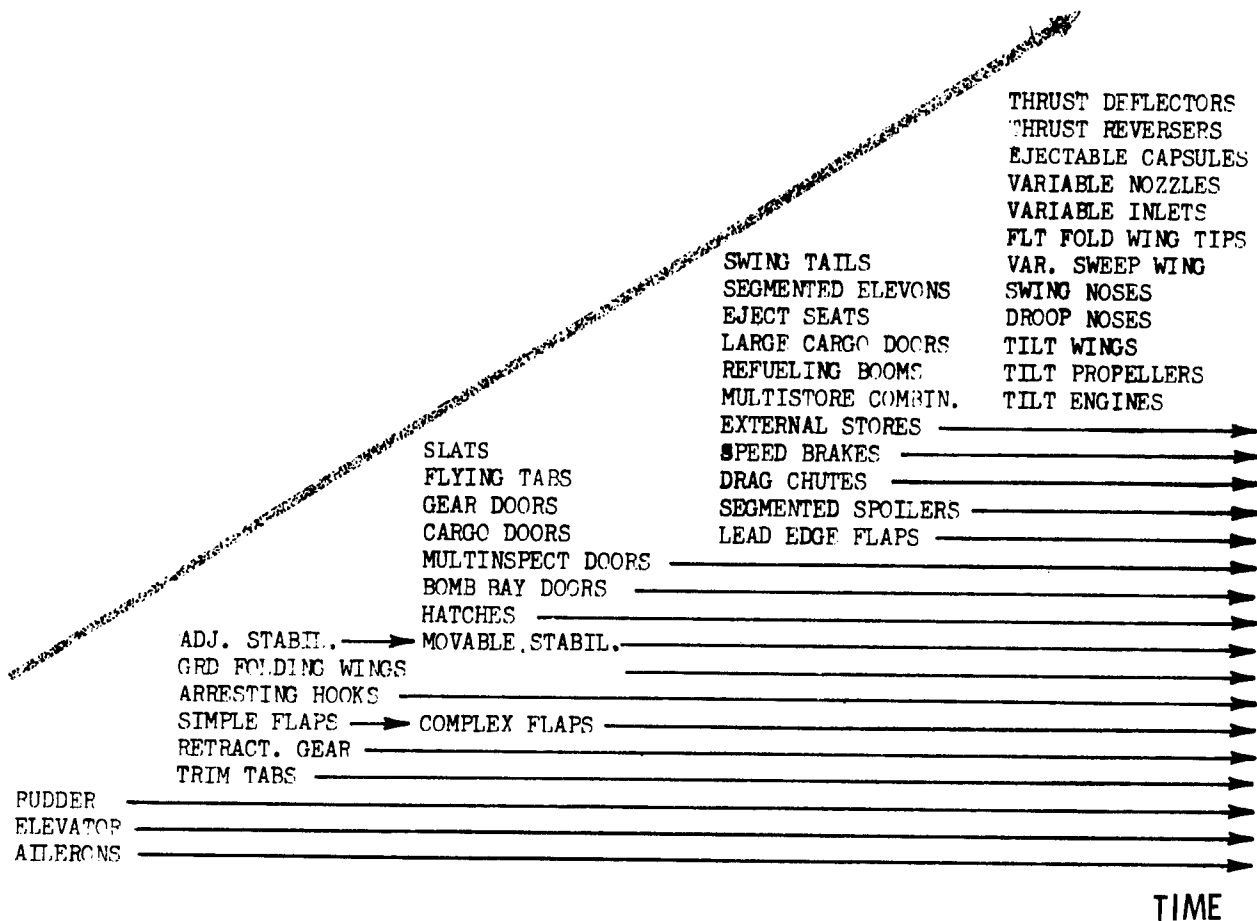


FIGURE V-5

requirements for program planning:

- Higher temperature structures
- New design criteria
- More potentially critical loading conditions
- Design iterations due to complexity of the operating environment
- New simulation and test methods and facilities

The possibility of an increase in sonic fatigue problems in flight vehicles is graphically demonstrated by the forecast in Figure V-4. The necessity of evaluating the capability of the sonic test facilities on hand as to their capability to provide the necessary sound levels would be underscored by this forecast.

Figure V-5, shows a trend in aircraft complexity with respect to "variable geometry" items or devices which are deployed or activated to give the vehicle a new or increased operational capability. Since each of these require additional structure in their vicinity to carry loads around the device, they show an undesirable trend of increased gross weight. This then would result in a requirement for increased

effort in the area of light weight structures for flight vehicles.

The forecast in Figure V-5, is borne out by the trends shown in Figure V-6, showing predicted increases in structural weight fraction (empty structural weight to T.O. gross weight) as vehicle complexity increases. The C-5, at the top of the curve in Figure V-6 with its large cargo doors fore and aft, shows partial confirmation of the Figure V-5 forecast.

Figure V-7, showing gross weight trends of flight vehicles, carries with it such implications to Laboratory management in their R&D planning as reduced stability and structural frequencies, increased capacity for test, assembly and tooling and a further reinforced requirement for efficient light-weight structures.

The ability to forecast technology advancement also lies at the core of special study/analysis efforts conducted in the Laboratory on new systems and subsystems throughout the program year. Given the requirement to satisfy a technical need at some

STRUCTURAL WEIGHT FRACTION (λ_s) TRENDS

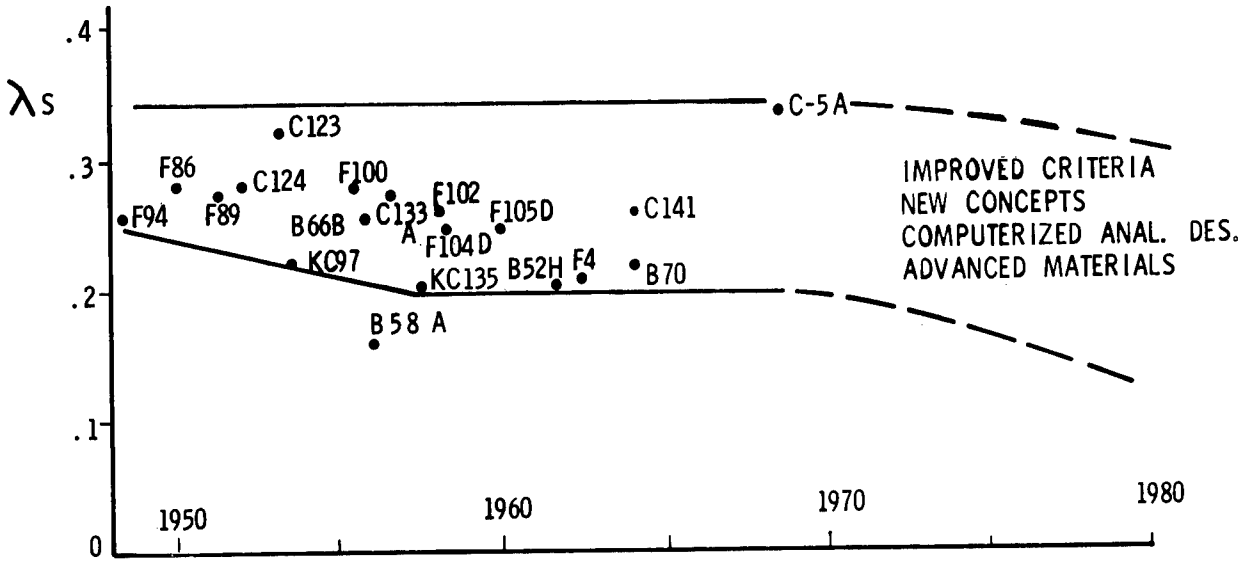


FIGURE V-6

GROSS WEIGHT TRENDS

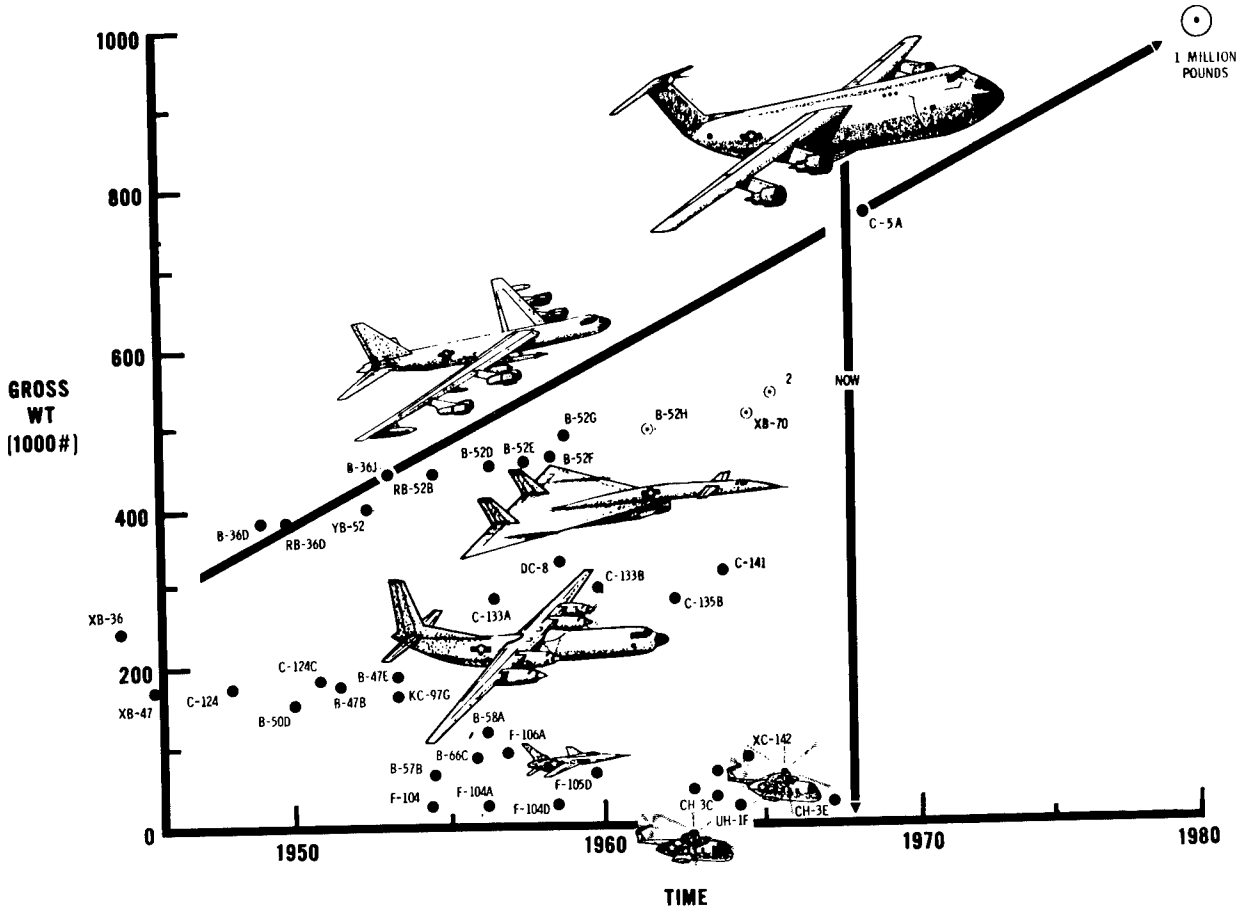


FIGURE V-7

specified date, the ability to forecast permits the making of a non-probabilistic prediction on a given confidence level to satisfy that need at the time required.

Typically the process begins with either a self-generated or assigned study requirement. This requires the assessment and definition of all parameters having an impact on the study area, as well as the delineation of interrelationships of these parameters. The timing of a study may vary from a determination of today's capability in the study area to the determination of level of achievement of a capability attainable at some future date. It is the latter case which requires the use of the technical forecast in the applicable parameter areas, in order to arrive at the set of technology predictions affecting the attainment of the desired capability. As is the case in the normal program cycle, these forecasts are often made not only implicitly, but often unconsciously, by personnel experienced in the area in question. More formalization of the forecasting function should lead to the preparation of more useful and accurate studies, in that the chances of inaccuracy or overlooking pertinent information would be correspondingly decreased. The requirements for having explicable logic and stated assumptions, forces better reasoning about the problem and permits better challenging of intuitive assumptions and their adequacy.

The foregoing has indicated some of the ways in which forecasting has become an integral and essential part of planning, programming and analysis in our Laboratory. It has helped guide us toward new technologies, helped to identify future problem areas, provided structured and quantified individual technical judgements and in summary, has made a major contribution to improvement in the quality and effectiveness of our R&D program. There is no claim made that it is always exact or infallible, but to quote one who has made substantial contributions to the science of forecasting, Prof. James Bright of the Harvard Business School, "technological forecasting forces us to challenge past practice and to look to the future in the hope of controverting the adage that history repeats itself, by introducing variations that will be to the benefit of all."

Acknowledgement:

Figures 3 thru 7, were extracted for the paper "Air Force Air Frame Technology" by W. E. Lamar, R. F. Hoener and W. J. Mykytow of the Air Force Flight Dynamics Laboratory. This paper was presented at the AIAA meeting on "Aircraft Design for 1980 Operations" in Washington, D.C., Feb. 1968.

RELATED REFERENCES

- Ackoff, R. L. et al, "*Scientific Method: Optimizing Applied Research Decisions*," John Wiley and Sons, New York, 1962, pp 87-91
- Air Force Systems Command Manual 375-3, Systems Program Management*, Department of the Air Force, Air Force Systems Command, Andrews AFB, Maryland
- Baker, N. R. and Pound, W. H., "R&D Project Selection: Where We Stand," *IEEE Transactions on Engineering Management*, December, 1964
- Buchheim, Robert W., *Problems of Planning and Decision in Military R&D* (Rand P-3021), Rand Corporation, Santa Monica, Calif., 1964
- Buck, G. T., Col., Evaluating R&D Effectiveness, *AIAA-Aeronautics and Astronautics*, April 1966
- Bright, James, R., Can We Forecast Technology, *Industrial Research*, March 1968
- Cetron, M. J., Martino, J. and Roepke, L., "The Selection of R&D Program Content—Survey of Quantitative Methods," *IEEE Transactions on Engineering Management*, December, 1964
- Cetron, M. J. and Monahan, Thomas I, An Evaluation and Appraisal of Various Approaches to Technological Forecasting, *First Annual Technology and Management Conference*, Harvard University, Cambridge, Massachusetts, May 1967
- Gordon, T. J. and Helmer, Olaf, *Report on a Long Range Forecasting Study* (Rand P-2982), Rand Corporation, Santa Monica, Calif., 1964
- Jantsch, Erich, *Technological Forecasting in Perspective*, Organization for Economic Cooperation and Development, Paris, France, 1967
- Joint AMC/NMC/AFSC Report on Technological Forecasting, 1967
- Long Range Forecasting and Planning, Proceedings of U.S. Air Force Office of Aerospace Research Symposium*, Clearinghouse U.S. Department of Commerce, Springfield, Va., 1966
- Nutt, A. B., An Approach to Research and Development Effectiveness, *I. E. E. E. Transactions on Engineering Management*, September, 1965
- Old, Bruce and Wheeler, William S., *Criteria for the Selection of Research Projects*, Arthur D. Little Co., Boston, Mass. 1966
- Pardee, Frederick S., *Technological Projection and Advanced Product Planning* (Rand P-3622), Rand Corporation, Santa Monica, Calif., 1967
- Richman, Barry, The Soviet Educational and Research Revolution, *California Management Review*, University of California at Los Angeles, Los Angeles, California, Summer 1967
- RTD Management Guide for Exploratory Development*, Department of the Air Force Systems Command, Andrews Air Force Base, Maryland
- Technology Transfer and Innovation, Proceedings of National Science Foundation Conference* (NSF 67-5), U.S. Government Printing Office, Washington, D.C., 1967
- The Planning-Programming-Budgeting System, Hearings before the Subcommittee on Economy in Government*, Joint Economic Committee, U.S. Congress, U.S. Government Printing Office, 1967
- The United States and the World in the 1985 Era*, Syracuse University Research Corporation, U.S. Department of Commerce, Washington, D.C., 1964

DISCUSSION FOLLOWING PAPER GIVEN BY MR. AMBROSE B. NUTT, AIR FORCE FLIGHT DYNAMICS LABORATORY, "TECHNOLOGICAL FORECASTING IN AN R&D LABORATORY"

MR. DARRACOTT: Did I understand you to say that the TORQUE experiment has not yet been completely evaluated?

A. That is my understanding, yes.

MR. FRESHMAN: I believe it has been completely evaluated. When I left last Friday, I think the final copy was being typed. There were just the last few pages to finish. It has been coordinated with the Air Staff. The Air Staff representative was impressed with the findings and conclusions that TORQUE might be a much improved method for resource allocation and for planning than our present methodology in the Air Force.

MR. IRWIN: You mentioned toward the end of your conclusions that you thought there should be stated assumptions rather than intuitive ones. This is quite interesting to me because I think I have been operating, without realizing, under unstated assumptions which I really ought to examine. I wonder if you would elaborate a little bit on some of the assumptions which in your work ought to have been stated that perhaps weren't.

A. Well, for instance, I think I was making that statement in the context of the fact that forecasting forces you to look at the total problem. In other words, suppose you are in the business of developing high temperature structures and you have gone down and talked to the people in the materials laboratory and you come up with a structure which looks real good or material which looks real good to you. Now this material, after you look at the test results, looks as though it might satisfy your problem. Your normal tendency now is to go home immediately and start working on it. The forecasting situation is such that we are required to look at all the available material and in forecasting determine whether or not there are other alternative ways of approaching this problem. So I think what I am saying is that it forces you to do a better job evaluating the technical possibilities in the particular area with which you are concerned.

DR. BENNETT: What specific steps do you take in forecasting the continued development of the material, say, like the boron deposit?

A. We do not really do the forecasting in materials. What we do is the forecasting for the structure which is a slightly different problem. In other words we are on step down the road and what we normally would do is go back and look, for instance,

at a set of capabilities to determine the particular structural problems. For instance, you might be interested in materials for the structure of a high altitude spacecraft operating in a certain time and temperature environment. Suppose one is examining a high lift over drag vehicle which comes gliding in like this (demonstrating)—it is exposed to the high temperature for a much longer time. You would look at the critical portions of that vehicle, namely the nose tip and possibly some portion of the leading edge or lower portion of the wing and determine what the temperature-time environment was. Then you would look at the material field and find out what material would be capable of withstanding this temperature and time. As a further step, you would look at the fabricating possibilities of these materials. Your forecast would show the various materials that you felt might do the job. Finally you would determine, in terms of what your research requirements were, which ones you would address in terms of looking at the structural problems, fabricating problems, maintenance problems, refurbishing problems, usability problems—all the things that go into a usable structure in a one-flight vehicle.

MR. LINSTONE: I think we all recognize that you have to have a two-way street in terms of needs in determining the capabilities and determining the needs. Most of your discussion was on the assumption that you start with the needs. What is the mechanism you have for introducing the new capabilities that are not coming out of these requirements but which can very well determine the systems which are not now on this list? How do you go up the ladder?

A. We call this our marketing responsibility. We have several ways of doing this. One is that in looking at a resources allocation system, like TORQUE, for instance, which is pretty much based on system's needs, our recommendation and conclusion in TORQUE was that certain laboratories' budgets not be subject to that at all but be free to leave space for breakthroughs of new ideas. So a certain part of the budget is completely free from any restraints and could come up with new capabilities. Another thing that is done in the Air Force is that we have a laboratory directors fund which is pretty much reported on after the fact in the terms of expenditure which is used specifically for the

purpose of looking at unsolicited proposals or new ideas within the laboratory, if you want to make a quick response, because the normal program is pretty much cast in concrete. Another way we make our new capabilities known in systems is in coupling meetings. We meet with the product divisions during the year and at that time we tell them what new capabilities we feel are coming out of our laboratory's program that are usable. We also use this technique with industry when we go out and review their internal R&D programs. One of the key parts of this latter procedure is the presentation by the laboratory of new possibilities and new techniques and capabilities that we feel have arisen or are about to arise based on our work or the work of someone else.

Another way we do this and this is sort of meaningful in terms of the immediate future is, for instance, whenever a new system is about to go into a definition phase or even a conceptual phase, we go over and talk to their chief engineer and say, "Here is a list of goodies from the laboratory which will come out in the next year or so that you should be able to use." For instance on the AMSA, there are about thirty-two different items that went into the request for proposal directly emanating from the laboratory—on the F-15 similar sets of new technical possibilities for the aircraft were presented. We also respond to advance system studies in the Aeronautical Systems Division and, as a matter of fact, participate in the studies and, of course, at that point we are now forwarding to them the latest results in terms of technology. TORQUE, I think, was very useful in this fashion in that they were not restrained in any way from conceiving new systems and, of course, it turned out that some of these new systems showed us new technical ideas. We make every effort to market our product at all times.

Q. One thing that is bothering us a great deal is the cost of fighter aircraft. We see a dismal record of constantly increasing complexity and costs. Solutions are likely to come from outside the industry. There may be new capabilities which could drastically change the manufacturing process.

A. As well as operational uses.

Q. Such capabilities might change out thinking from the bottom up. What is your view on incorporating this type of thinking?

A. We have a way of showing, and this is very difficult for a lot of people to understand, what exploratory development can do for us. We call it the "cascade" effect. On the one hand we put down

the list of operating environments such as air base capabilities, for instance, then we show certain subsystems and how they go to increase the capability of operating in such an environment and, based on these particular subsystems, we show the individual pieces of technology which go to make this up. It is a good example of addressing ourselves to the necessity for increasing simplicity in things such as generating oxygen in flight where you get away from all the ground LOX equipment and LOX plants that you see around bases, where you can actually use electrostatic means for concentrating the oxygen. This is one approach to the problem of simplicity. Of course, since I'm not in the avionics business, I could suggest we pull all the transistors out and save some of the cost.

MR. FRESHMAN: I would like to get to the point of where we get ideas for new systems capabilities under the TORQUE concept. There will be strict competition between the laboratories, which are the focal point for new technology under this TORQUE concept. Those specific parts of our organization which are best equipped by knowledge and capability to do certain jobs will do it. Now everybody knows that our laboratories are better equipped than our systems divisions for forecasting what technology will be or what it can do. It is fundamental under this TORQUE system that if they wish to get their technology funded under the TORQUE rationale, each and every laboratory should forecast in its areas of expertise where its technology will be in the time frame of the 1980's or when the systems division can use it and how they can use it to the best of their ability. One of the product divisions, Aeronautical Systems Division in this particular test case, can then take all these forecasts from all the competing laboratories and then, since they have the capability, they end up with integral conceptual design for a new system. The laboratories go back and compete again to get their pieces of technology included in the conceptualized advanced system, since there has to be a trade-off by the product division of various alternatives to achieve the aircraft. Many times we will go various competing technological ways. This is only a forecast. If there is a forecast of certain laboratories which will give us the same capabilities and both are effective and meet the time frame, then we might decide to go into competition until we found out which is really most advantageous. There is a constant reiteration. Now, the Air Staff, in all its wisdom, is not capable of thinking of future capabilities only by threat. We must also include the thoughts

of what else we can do regardless of threat. The initial technology forecast comes from the laboratories and it includes possible effects on components, subsystems and systems. It then goes to the product divisions; the product divisions offer possible systems to the Air Staff and back down it comes again with the operational capabilities. This goes on all the time. This is not a static thing. There is nothing new about it. It requires and insists that the laboratories do the forecast and structure dollar and time budgets to achieve specific technology objectives. It requires and insists that the Air Staff say what it wants and stipulate the environment for the operation of future systems capabilities so that each and every part of our organization must go out and do its work and make itself visible. In various ways we do that today but this requires and makes it completely visible so that our laboratory people are not required to gin up their own systems as many a laboratory does now. Also when a systems concept arrives upstairs, the higher headquarters says, "Who says that is the system capability I wanted? What are your assumptions here? We never told you that." So all of this stuff is eliminated and what did we find? We found that the laboratory spent less time in doing its work because it did the work it was most capable of doing and other parts of the Air Force did likewise. Nothing new, except complete visibility and cooperation on the part of our people working together.

MR. SCHNARE: You have not discussed the Technical Effort selection process. These people might be interested in hearing how these efforts are selected, since budgets are limited. We are all aware of this; it is a real practical hard fact that we have to deal with. How do we establish the priority of the various system capabilities and the importance of the technical effort to the achievement of the capability?

MR. NUTT: That is what Air Staff does as far as capability objectives are concerned and one thing in TORQUE is that we actually put a weight on the systems. The current TORQUE program does not have a relative weight in systems support of a single objective. It is our contention that this removes a very important factor from the program in that when an objective has five or six alternatives to support it, generally the technologies to support all five of these systems have five times as much utility. It is not true because the capability objective should be supported by the minimum amount of resources; therefore, you want to know which is the most efficient. So I think Mr. Schnare has a very important problem here. One has to make

judgements about the relative worth of these particular capability objectives on any system like this, and I should hasten to add at this point that none of these systems, RDE or TORQUE are really bibles. They are merely another management tool, another something the managers get to make better decisions. In the case of our laboratory, it has got to make a decision on the basis of possibly four hundred work units per year. It is absolutely impossible to simultaneously look at four hundred different increments and see which twenty or forty should be cut out. We have certain critical technical needs in the product divisions, and we will do these even though it is not cost effective because it would be impolitical for us not to. We rely very much on their backing to get our funds, and the pendulum is swinging more and more in that direction. Since the product division has control of the laboratory funds, they (the laboratories) jolly well better get on board with them and find out what they need and have a reputation for supporting them.

DR. SLAFKOSKY: I would like to know whether there is an agency or group in the Air Force which takes technological forecasts that the laboratory provides or takes forecasts or data that any other source provides and develops or generates a total concept for the operation of the Air Force for the mid-range or long range period. This can be accomplished in at least two or three ways. First, the new total concept can define or delineate new ways for accomplishing the current Air Force missions assigned it by the Joint Chiefs. This sort of concept may enable the Air Force to do its mission more effectively or more efficiently or both. Secondly, this new total concept may be one which would enable the Air Force to undertake totally new missions and accomplish them effectively. Or finally, it might be a combination of these two. Obviously these total concepts would have to be based on threat forecasts, politico-economical assessments, both national and international as well as the technological forecasts of our industrial and scientific bases. These concepts would comprehend strategic and tactical objectives and missions and would include the entire spectrum of aircraft, missiles, weapons and other systems necessary to implement them.

MR. FRESHMAN: Sure. The operational capability is made up, say, for this particular test case in the 1980's. It took it all into consideration even if we did not have a particular mission area right now, but it was not being addressed by the Army or the Navy, and the Air Force thought it was in its pur-

view to take. This method permits the Air Force to go into any area—there are no constraints whatsoever—to meet a problem that the Air Force thought had to be solved.

Q. What method are you talking about? Are you talking about one employed by the Air Staff? Before, you gave me the impression that the Air Staff is very busy and I am sure it is. All staffs on the highest level are busy. Therefore, I assume that you cannot be saying that the Air Staff at Headquarters is the agency which generates new or total concepts for the Air Force.

MR. FRESHMAN: Are you referring to this particular case? O.K. We have what is known as AFXDO which is the plans and concepts outfit. They make the Air Force Plan and they make the concepts for Operational Capabilities after the forecast, consulting with the divisions, consulting with DOD, considering the threat from all agencies, etc., and the national goals admissions, etc.—all things that you have iterated, and then some, and where the Air Force would like to get into certain areas. . .

Q. Is this accomplished annually?

A. No, it is not annual; it is continuous. The plan is annual; it comes out formally annually; but at any time they can throw into the hopper any new ideas they wish.

Q. What you are now telling me resembles the annual updating of current plans and incorporation of additional capabilities. This is not the activity I am inquiring about, for I am familiar with what you are describing as it occurs in all four Services. My question has to do with the existence of an agency or group to devise or generate new directions or goals for the Air Force and how to implement them. If as a result of their assessment of current goals and directions these are deemed adequate and proper, then this group might have as its goals better ways of accomplishing them. Such activity entails taking a new look at the Service and its structures, organizations, systems and interfaces with the other Services and ascertaining what technology and/or new tactics and techniques might be able to enable it to do its mission more effectively or more efficiently. Another way of stating this question is the following: Can the Service structure itself and develop systems to make it a much more effective instrument of national policy? Note, I am not speaking of the introduction of a new plane or weapon into the current structure to make it function.

MR. SCHNARE: What do you mean by “much more”

effective, “much more” cheaply; how much is “much more”?

DR. SLAFKOSKY: I think these words by their very definition would indicate fairly well what I mean by “much more effective.” For example, if you can, in fact, kill a target with one bomb or missile without having to drop or discharge tens or hundreds of them, that is being more effective especially if this is accomplished without additionally endangering the lives of other men. Or, as a matter of fact, if we could develop target acquisition systems which would enable us to detect and locate enemy targets quickly and to bring destructive firepower on them quickly, this would be much more effective. It is not being accomplished now.

MR. FRESHMAN: There are no constraints on where the technology can lead to. We have certain problems, whether it is protection of an air field or, as you say, to knock out convoys or to knock out cities with less tonnage and things like that, or to set up barriers, whatever you wish to say, or to set up an immediate airport. There are no constraints that the Air Staff puts down upon our people.

DR. SLAFKOSKY: Who are these people?

MR. FRESHMAN: The people who are in the laboratories and in the products divisions. These are the people most suited, not the Air Staff—the Air Staff does not prognosticate how to do these things; the Air Staff, even in all its wisdom, as I said before, cannot think of all the ways that technology can lead to, to do a different job or do a better job.

MAJOR MARTINO: As you may be aware, there is a continuing debate in the forecasting business as to which is better: to have a staff which produces an annual staff forecast or alternatively, every few years take people away from their desks and put them together some place on a non-routine basis and have them produce a forecast. Now I admit I have been in the first camp. I advocate setting up an organization which will produce annually a staff forecast and I think this has some advantages, but you have hit on the major disadvantage of the system. Once you institutionalize the annual production of a forecast, you have just about stifled the imagination that goes into it. I advocate doing this because we don't even have that much now but the practice the Air Force has followed in the past, assembling a massive forecast team every few years I think gets at the question you are asking.

DR. SLAFKOSKY: I obviously did not make myself clearly understood. What you are saying about institutionalization is fundamental to what I want

to talk about. I understand and understood the generation of the Air Staff's annual plan and the continuous process involved. What I am really asking is: is there any agency other than the Air Staff or other than the individual laboratories which develops concepts and then objectives based on these concepts,—i.e., unified concepts, not just one concept—that are inter-related in one over-all complete concept whereby the Air Force would hope to measurably improve its capabilities to perform its current missions or undertake new missions effectively in support of national policies? To no small extent such a concept would have to be based on the technology that technological forecasts predict will obtain.

MR. DARRACOTT: We are setting up an organization which we hope will achieve these results that the Colonel is talking about because we are trying to come up with the forecast and utilize the forecast of the Army, come up with concepts of systems and then the combat development commands, who are the user, will look at these and try to come up with a preferred manner of fighting on the battlefield. Now this is an experiment which is going ahead and probably within the next four years will either prove out to be magnificent or a failure, and that was in these papers I passed out yesterday.

DR. SLAFKOSKY: I would like to point out that in your laboratory, Mr. Nutt, within a limited area you sent to the industry and found out what they thought about technology would be providing for you. Then, within the limits of your responsibility, you developed some concepts and objectives and forwarded them up the line for approval. This makes sense. My question now is: Is there some group in the Air Force where the results of such efforts as Mr. Nutt describes are really unified into a cohesive, meaningful concept from which developmental objectives and requirements would flow?

MR. FRESHMAN: That was the improvement we have. Mr. Nutt's laboratory predicted its own systems. In each of our laboratories, they predicted their own to make sure their technology gets funded. Under the TORQUE concept, it is all being funneled into one area, objective people who are not concerned whether this one laboratory gets funded more than another, only which will give the Air Force a better capability and whether it is a new capability or an improvement over an old one. I do not see what you are driving at here. The Air Staff is considering all this through its products divisions, which has no limitations whatsoever, and taking advantage of what industry, universities, or

anyone else can suggest. We are not restricting ourselves.

MR. SCHNARE: Who has the great crystal ball on the Air Staff? Who is going to be all-seeing?

MR. FRESHMAN: They are not going to be all-seeing. It is coming up from the laboratories, up through our products divisions, and then finally from the Air Staff—but this is a continuous circle.

LT COL STEPHENSON: I think the key document is the USAF Planning Concepts which is prepared by AFXDOC, and the Chief of Staff's direction is that it should be used in this way. It consists of three major parts together with an introduction. Part 2 provides a best estimate of the future world situation during the planning period. Part 3 is a discussion of the Air Force approach to handling conflicts—i.e., how we do our job. Part 4 is where various concepts, both weapons systems and operational, suitable for the anticipated requirements and the appropriate time period are discussed. It is the longest part of the document. It goes into all aspects of Air Force type operations, both strategic and tactical, command and control, logistics, etc. In each of these, there are listed broad concepts which the Air Staff believes do show how well we can do the job, the way we can do it as related to the projected operational environment as we now see it. Now the next question concerns how the information is developed, because this is a key. One of my main concerns with TORQUE continues to be that it places a lot of responsibility on the Air Staff to identify the needed weapons systems. In the last two or three years, we have worked closely with the group preparing the planning concepts and it is apparent that the techniques for providing improved information are being developed and employed. But in answer to your specific questions, if we have anything in the Air Force which provides that information, I think it is the USAF Planning Concepts prepared by AFXDOC.

DR. LINSTONE: I think what Alex is getting at is a very fundamental problem. I think perhaps some of us still do not quite see it. A method like TORQUE is good to keep everybody on the track and coordinated. Like any bureaucratic tool, it is a useful, maybe even necessary, one. But who is handling the fundamental questions like: does the Army need divisions of the type they have now at all? The problem with methods such as TORQUE is that it uses as a basis the current view of the current weaknesses. I do not think any formal tool like TORQUE can introduce enough fresh thinking or innovation. You need the very opposite of a routine procedure.

DR. SLAFKOSKY: I would not agree that the Air Force cannot produce an organization which can generate new and even revolutionizing ideas. I do agree, however, that is is not the kind of activity that can be generated every year. Such an annual attempt would be self-defeating. What I had in mind was something that would be carried out every four or five, maybe six years. After the concept, plans, goals and objectives were completed, the next few years would be taken up with refinements. I do not know whether institutionalizing such activity would prove to be good policy. I know how rigor mortis sets in. I think this sort of group and activity can be set up. I am not convinced that this activity need be continuous. It might be activated on a fairly regular schedule. Another point that must be made about this long range concept development is that people concerned with it must realize and make others realize that long range concepts should at least provide general directions/goals in which and toward which movement should be made. Current and mid-range systems/capabilities will have to lead into the fruition of the long range systems, for transitions cannot be effected overnight and there are too many long-life systems which cannot be discarded efficiently.

MR. ELLISON: I would like to speak around the point a little from the work of creativity research-findings. There is a lot of evidence that supports your point of view. For example, a number of studies have been completed where quantity of education in a specialized field did not correlate highly with the criteria of success in science. The whole area of creativity training, whether you believe in it or not, is predicated on the idea that subject matter specialists often need a new point of view. You have to meet certain requirements of subject matter knowledge in the area and a specialist is usually required for that, but it is the input from outside people that sometimes makes the crucial difference in a creativity training program. There has also been some research done where actual quantity of knowledge has been related to problem solving tasks. Again, quantitative knowledge has not proved to be the essential ingredient. It only accounts for a small percentage. All of this research would seem to indicate that people outside of the normal frame work oftentimes are those who are essential. Although I am not a specialist in the history of science, I think we can argue there too, that some of the major breakthroughs have upset the whole apple cart of traditional science.

DR. SLAFKOSKY: Since I started this discussion,

let me try to bring it to an end. I get the impression that you think I am criticizing the laboratories and other divisions concerned with the future development of the Air Force. Really, I have no such intention. All I am trying to find out is if the Air Force, in fact, has this kind of concept development group. Obviously it does not. And it is understandable. When questioning you about it, I find that you misunderstand me or you are not responding to my questions. When I say you do not have such a group, that should not be construed as criticism of the Air Force.

MR. FRESHMAN: I do not agree with that at all. You draw your own hypothesis that we do not have this kind of group at all and we are saying to you that these people are thinking all the time. They are not constrained at all, and you are trying to set up this straw man and hitting it down yourself.

MR. SCHNARE: You have got to admit we are in the beginning phase of developing a usable planning method for Applied Research Laboratories. TORQUE is by no means the final answer and we all know that. Anybody that can make any comments or ask any questions which are going to help us determine how to do the planning job better should be heard.

MR. FRESHMAN: Let me point out one thing: under the TORQUE concept laboratories, for the first time, do not have to go begging up to the Air Staff. Their words were immediately up there and there were no constraints, no higher headquarters stopping it or anything else for the first time—easy, simpler, faster and with more authority behind it.

DR. SLAFKOSKY: Mr. Freshman, you are missing my point. I am not criticizing the laboratory.

MR. SIMON: I think between the two gentlemen there is little error, but perhaps what is needed is a fresh viewpoint. Dr. Slafkosky is talking about a super-high agency, not Air Force alone but looking at all the Services, that comes up and says, "Let's go in a different direction." It seems to me that creation of such authority would merely result in a duplication of the President's Scientific Advisors or something like that.

DR. SLAFKOSKY: I did not mean anything like a super agency or a Council of Scientific Advisors.

MR. SIMON: Let me continue on then. As was noted by the speaker, you can go along in the direction you are going and work very well in an institutionalized path; but the entrepreneurial approach is not really encouraged in governmental agencies. The place for that is in industry, but without it you will

never get the fresh approach that you are talking about.

MR. NUTT: Here is one thing: the laboratory director does have to maintain some sort of environment of creativity and receptivity to proposals, either internal or external and I think that is pretty much a matter of the individual director. I am very discouraged about the ability to program breakthroughs; but I think the internal DELPHI, to the extent they have gone, is where we are getting fresh insights and new directions. I wonder if you would like to say a word about that.

MR. DARRACOTT: I would like to make just a slight comment. I have watched Harold Linstone's work and I have watched TRW's work in two different areas, each coming up with forecasts and trying to determine areas which would be useful from an industrial standpoint from their companies. Now as I understand it, they have a group that sits off to one side that periodically reviews their work and then says, "All right, this is the way we want the company to go," or "These are the fields we feel are most promising for profit making for the company; therefore, this is what we will do." Now the Army is trying to do something similar in this respect in organizing a "three-ring circus" or "Troika," if you wish, by taking an agency from the Army Materiel Command (the builders of equipment), the Combat Development Command (the users of the equipment), and the Assistant Chief of Staff, Intelligence, which tries to determine the environment in which we are going to fight and combining them into an Advanced Concepts Organization.* We are trying to get together, sit off to one side and objectively look at the threat, man, and materiel information. As I said, it is a daring experiment and the next four years will tell the tale whether or not we are successful. So there are groups which are being formed like this. Now, I do not know about the Air Force or the Navy or the Marine Corps in this respect, but I have an idea that there is someone sitting back up here who can put his feet on the desk (he may be a retired Admiral or retired Air Force General) and say, "O.K. These are the things which I feel, from the technology which has been laid before me, will give us the best payoff in the years to come in the environment in which we may have to fight."

MR. NUTT: Not in the regular sense but certainly to some extent.

* This organization, formed from three independent elements of the Army, has no titular head.

MR. PYKE: Mr. Nutt has suggested that maybe Lockheed and TRW have a fresh new approach to the solution of the kind of problem raised here. Perhaps we do, but I think maybe John Gardner has put his finger on the problem when he raises the concept of "self-renewing institutions" and wonders if this is possible. It may not be—I cannot think of an example of a major industry replacing (renewing) itself. The semi-conductor replaced tubes, but the tube people did not invent it; it came from outside. I do not think internal DELPHI is the ultimate answer because here again we are talking to ourselves. TRW recognized this and we do plan to expand our DELPHI exercise (beyond corporate boundaries) at least to check some of our thinking with people from government, the academic world and from other industry; because, in the final analysis, it comes down to the decision maker who has to recognize, "What we are now making is going to be made obsolete by a new technology; thus, we must drop what we are now doing and embark on something new." That is a big decision and a hard one to make.

MR. KNAUSENBERGER: I just want to console you that such hope for getting new creativity in groups detached from the normal routine operation exists in NATO. These groups integrate knowledge represented by industry, universities and government people. They present to some extent each country's activity in cooperative science, and I think you would be interested in these examples of diverse interaction.

MR. SCHNARE: I would like to address a question you asked about: Who, what and where do new ideas come from and how do we provide for them? I have to go back a little bit in history in order to sort out the problem for you so that you can understand what the situation is like. Within the Air Force several years ago there was a feeling that our laboratories were too engrossed with the immediate problems of supporting the forces in the field and solving the unsatisfactory conditions and equipments already deployed. So, as a consequence of this, to make a long story short, the Applied Research Laboratories were set up to deal strictly with technology or advance development as the budget now lays these two programs out. As a consequence, we lost track, to a large degree, of the kind of things that were happening to the equipments that were being developed and moved to the field. We lost track of how well they were working or not working. Some old heads decided this was for the birds and they did continue to keep in touch with their old friends out in the field, finding out what was going

on and how weapons were performing. But generally, over this five to seven year time period, there has developed within the laboratories a feeling of being a group apart. This attitude is sort of reflected by the questions that Congress asks: why should we have two hundred and fifty million dollars in exploratory development budget? Why should we spend three hundred and fifty million dollars in an advanced development program? What good are these things and what do they contribute? Why should we have six thousand or ten thousand people assigned specifically to this research task? What purpose do they serve? What comes out of this kind of an organization? There are, also, the questions that laboratory people themselves are asking: how well are we responding? How well are we supporting the Air Force in its mission? A contributing factor was the McNamara era where new systems were not initiated at a very rapid rate, certainly not within the Air Force and certainly not with the objective of dealing with the limited war situation. So, as a consequence of this situation, revolutionary ideas have not received too much attention. The Planning Concepts Document by the Air Staff gives all kinds of opportunity for innovation. It gives some indication of what higher headquarters is thinking and what they would like to be able to do in terms of military capability. Historically, the military missions have not changed; the basic fundamentals of warfare have not changed very much; it is only the tools that we use to achieve the objectives that have changed. When you begin to look at the problem in this way, the revolutionary ideas such as highly concentrated light beam weapons and things of this sort begin to have applicability to the very fundamental objective of effectively destroying the other guy with the least amount of damage to yourself.

Our current environment is not too encouraging to allow for development of revolutionary weapons. People are getting tired of the Defense Department getting the bulk of the Federal budget. I think it is time the Defense Department took a good look at themselves to see just where they are going, why they are going there and how they intend to take care of the missions they have been assigned. As far as the future is concerned, I think we are going to have less money and less people in the Defense Department. I do not think there is any question about this. As a consequence, we have got to get more efficient and more effective in our use of people and in our ability to do the job. So we get to this particular point: how do we go about doing this? How do we bring new ideas into the picture when it takes

seven to ten years to develop a new idea to an actual operational system? It means the new idea must be compared to the old method of accomplishing the mission to determine if it really is a significant step forward or if is just an alternative way of doing the job. If it is good, we should communicate with the Systems Divisions of the Air Force who develop the systems and get their support for the new idea and perhaps eventually the Operational Commands to get their criticism. Then we must be sensitive to the fact that we do not want to start an arms race. We do not want to exceed the capability of the enemy by too great a degree for fear of challenging him at a particular point or forcing him into a corner. Now, if you understand that problem, you begin to understand why an evolutionary method of planning rather than a revolutionary approach is necessary. But in any case a new weapon must be significantly better in all respects as compared to the old one to be worthy of development. Otherwise, you do minor modification on the existing system or product improvement. But now the point of revolutionary ideas: revolutionary ideas are hard to come by, extremely hard to come by. We used to set aside ten percent of our budget to accommodate revolutionary ideas. Very rarely did our scientists or the industry come up with ideas that would use ten percent of our budget. However, the laboratories have set up within their organizations an accommodation for people with revolutionary ways of getting the mission accomplished. We do allow a certain amount of resources within the laboratory to exploit these ideas either inhouse or by contract. The new idea does spring, however; and this is the point Mr. Freshman is trying to get across, from or through the laboratory. The new idea can move forward in a direct way through this new communication channel that he talks about into the Air Staff for consideration of which mission and how this new technique can be applied as a weapon. TORQUE can provide system capability objectives and can provide a means of communication between various elements of the Air Force, but is is unreasonably complex. Simpler methods are needed.

DR. LINSTONE: I would like to comment on a point made by Don Pyke. In one sense our Corporate Development Planning group at Lockheed is an outsider. We worry about the divisions taking too parochial a look at their lines of business which are assumed to keep on going forever. We try to get the attention of top management where innovation is essential. Innovation very frequently does come from outside. Big changes in the steel industry

have come from individuals, not from large U.S. companies, or they have come from small European companies. The dilemma is not going to be solved by methodology. It can be defined in three words: age, size and success. All three are debilitating.

MR. Bisci: You talk about the five percent kitty that can be used in finance innovation.

MR. Nutt: That is all they would give us. We would like to have more.

MR. Bisci: Don't you have a committee of some sort which decides who will get the five percent?

MR. Nutt: Yes.

MR. Bisci: In order for a person to justify this five percent, they already have to have a prior amount of work already done?

MR. Nutt: No, not really. Actually this has been done by a committee of staff and chief scientists and they make a real effort to bend over backwards to any idea which shows promise, although there is no question in my mind that there is an unconscious filter they use, like "what kind of a guy is this," you know, "what kind of an idea does he have."

MR. Bisci: But a man had to come up with an idea, walk into your office and say, "I've got an idea this is going to happen." You are not financing him for innovation by saying, "We'll put you in a room and fund you for the next three weeks. Try and come up with some ideas."

MR. Nutt: We are very bureaucratic about it. We don't say it in that form.

Producing the First Navy Technological Forecast

by M. J. CETRON,* and D. N. DICK**

The views and opinions expressed herein are those of the authors and do not necessarily represent those of the U. S. Navy.

During 1968 the Navy prepared and published its first technological forecast. The Navy effort involved sixteen major Laboratory/Centers and eight System

* Mr. Marvin J. Cetron is the Head of the Technological Forecasting and Appraisal Group, Exploratory Development Division, Headquarters, Naval Material Command. He has been engaged in various technological planning roles in laboratories, Navy Headquarters, and in the Department of Defense for more than fifteen years. He served as Head of the Planning Branch at the United States Naval Marine Engineering Laboratory; Engineering Assistant to the Technical Director of the Navy Applied Science Laboratory; and Head of the Management Planning Review Branch of the Applied Science Laboratory. He is a member of the Military Operations Research Symposia, the American Institute of Industrial Engineers, the Interservice Technological Forecasting Methodology Group, the Interagency Group for Research on Research, and the American Association for the Advancement of Science.

Mr. Cetron has worked with various government agencies, both in this country and aboard, and has lectured at many academic institutions and to many professional groups. He has published numerous articles; coauthored several books; and contributed chapters to other books on long range planning, technological forecasting, and operations research in R&D.

He holds the B.S. degree from Pennsylvania State College, the M.S. degree from Columbia University, and is currently completing the requirements for the Ph.D. degree in R&D Management at American University.

** Mr. Donald N. Dick was born in York, Pennsylvania, on June 27, 1930. He received the Bachelor of Science in Electrical Engineering (1958) and the Master of Science in Electrical Engineering (1962) degrees from the Drexel Institute of Technology, Philadelphia, Pennsylvania.

From 1958 to 1965 he conducted research in the ship instrumentation area and performed system studies of ship control at the Marine Engineering Laboratory, Annapolis, Maryland. In 1965 he joined the technical staff of the Mitre Corporation, Bedford, Massachusetts, where he was engaged in communications studies for tactical and air defense command and control systems. In 1967 he joined the Advanced Planning and Analysis Staff at the Naval Ordnance Laboratory at White Oak, Maryland. His current tasks at NOL include technology assessment, weapon systems analysis,

Commands. The forecasts range from functional technologies to system options and the overall effort is comprised of approximately five hundred individual forecasts. It was implemented, prepared, and published in seven months and it is estimated the overall forecasting task cost approximately \$1.9 million. The implementation of such a task, requiring the efforts of a large number of activities in a relatively new field over a short period of time represented a challenge to all involved. We would like to share with you today the experiences gained in this task; producing the first Navy Technological Forecast (NTF).

We should start by acknowledging the assistance gained from those who formally prepared forecasts some time ago, notably the Air Force and Army. Although the Navy waited until 1968 to prepare a Navy-wide forecast, several years were spent studying other forecasts for the considerations of methodology, structures, and overall approaches. The end result of this study is contained in "A Proposal For a Navy Technological Forecast, Part II, Back Up Report"¹ which has, and still does, serve as the bible for the NTF as well as a good introduction to the subject of technological forecasting.

Organization of the NTF

A review of the above report will indicate the Navy chose an approach similar in part to both the Air Force and the Army. The result is a program with a

and advanced planning methodology for research and exploratory development.

Mr. Dick is a member of the Institute of Electrical and Electronic Engineers, the American Association for the Advancement of Science, the Society of General Systems Research, and Eta Kappa Nu.

¹ Footnotes refer to references.

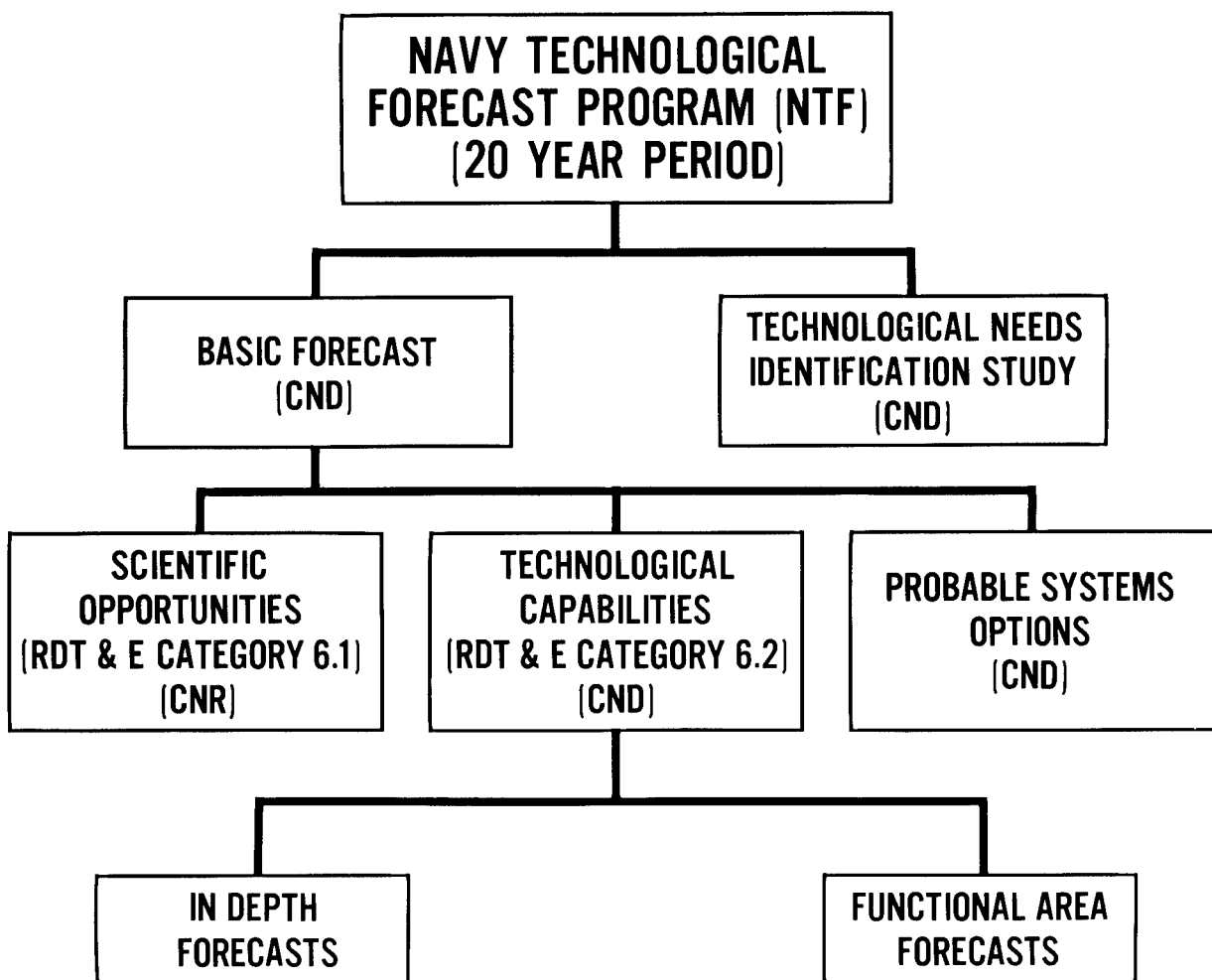


FIGURE VI-1.

very broad scope. The basic Navy Technological Forecast is composed of three major parts. These are: Scientific Opportunities, Technological Capabilities, and Probable System Options. As indicated in Figure VI-1, the three major areas constitute the basic forecast. In addition to these planned-for forecasts, provision has been made for specific technological projections for paramount needs. These are called Technological Needs Identification Studies (TENIS). The categories shown in Figure 1 indicate the complete program as it is envisioned at present.

During 1968 the Technological Capabilities and Probable Systems Options were prepared under the direction of the Chief of Naval Development (CND). The first forecast of Scientific Opportunities is being prepared by the Office of Naval Research (ONR) during this fiscal year (FY 69).

The Probable Systems Options section of the present Navy forecast suggests examples of systems

or subsystems which could be developed to satisfy broad mission requirements such as: strategic, amphibious, antisubmarine, etc. The forecast identifies known technological barriers and indicates fund-time requirements for system development. It differs from an unconstrained wish list in that concept feasibility has been investigated by analytic studies and/or experimentation to indicate gross operational capabilities. The Navy System Commands (Air, Ordnance, Ships) are responsible for preparation of the Systems Options forecasts. The 1968 NTF contains submissions from Air and Ordnance System Commands. The 1969 NTF will contain submissions from all Commands.

A summary of the volumes comprising the Navy Technological Forecast is shown in Figure VI-2. The technological capabilities portion, in terms of volume, is the largest contribution of the 1968 NTF. The Technological Capabilities forecast contains projec-

NAVY TECHNOLOGICAL FORECAST

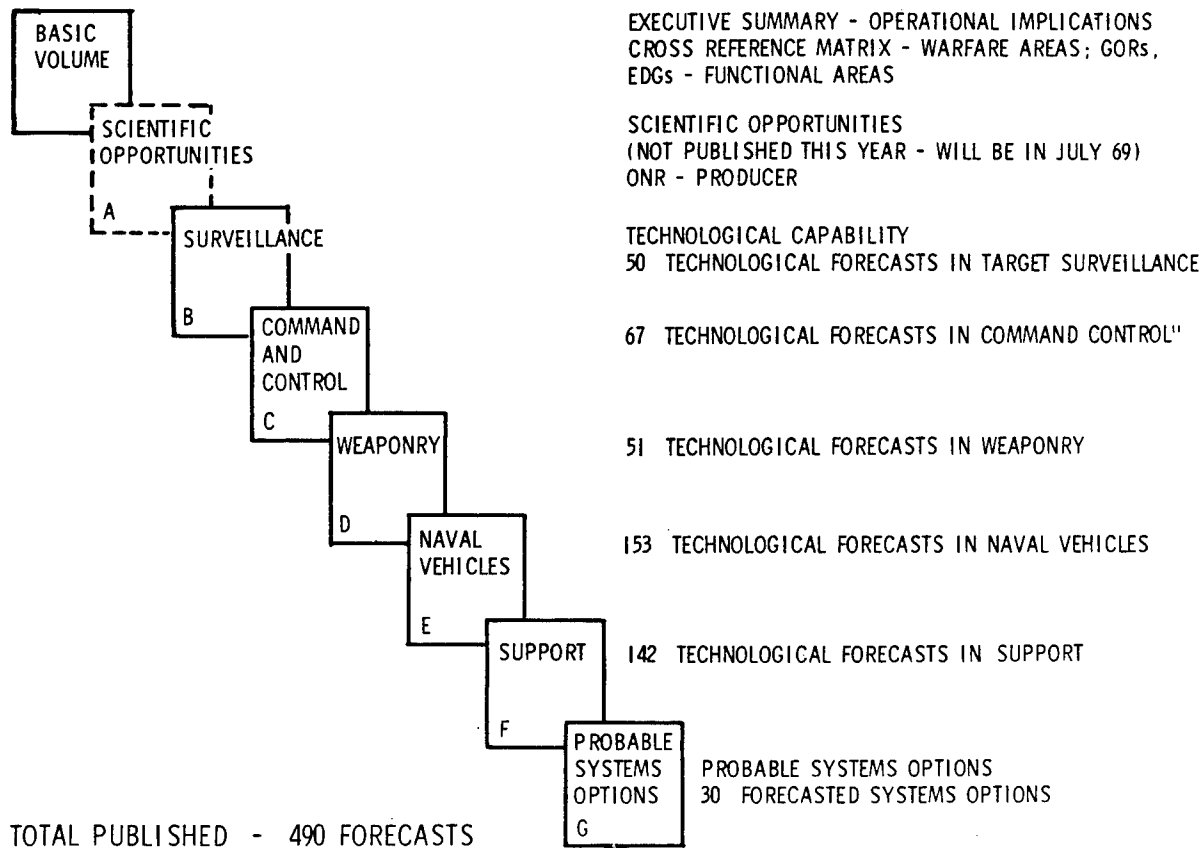


FIGURE VI-2.

tions of applied research and development activities of the Navy. The area covered is that normally associated with the RDT&E 6.2 planning series. In the system acquisition process it is the area of Exploratory Development. As such, it represents the technologies that convert basic science concepts into the fundamental building blocks of Navy systems.

The remainder of our discussions will be confined to the Navy experience in the Technological Capabilities forecast. The forecasts cover a broad spectrum of technologies, from nominal scientific approaches and specific Navy needs, to the broad functional technologies necessary to support general Navy missions. Because of the nature of this area of "technology" the problems of planning and providing a technological forecast are intensified. Any specific procedures or actions have broad implications.

Of the many factors considered in the initial stages of planning the Navy forecast, two were considered

fundamental. These are: (1) the forecast format, and (2) the technological structure. The determination of these two considerations will crystalize many aspects of forecasting such as: who will use the forecast, who will prepare the forecast, the overall utility, and the definition of technology.

Format

A major factor influencing the forecast format is the group of potential users. In general, these include groups from corporate executives through to the technical staff working at the bench level. The Navy group having the greatest need, and therefore the highest potential use, are the military planners in the Chief of Navy Operations Offices. This dictated that the Navy forecast be presented in a format meaningful to the operating Navy. At the same time, a

comprehensive technical projection was considered necessary to support any projections of Navy operational capabilities. Further, it was assumed that individual readers might not be familiar with Navy problem areas and be in doubt concerning the relatedness of technology areas, giving rise to the desirability of general background information.

Attempting to provide a forecast that satisfies all may provide a forecast that satisfies none. However, the format used by the Navy is designed to accommodate a diverse and wide audience. Table I indicates the major sections of a Technological Capabilities forecast. The amplifying statements under each section are the abbreviated directions given to each forecaster for the 1968 preparation.

As indicated in Table I, the Technological Capabilities forecast begins with a brief discussion of past applications and significance to the Navy, continues with a discussion of the present state-of-the-art, gives a technical prognostication, relates the

technical projection to operational considerations, and ends with credibility references and a listing of where, in the Navy, to look for such expertise.

One question immediately arising from the specification of this "total assessment of technology" type of forecast is—why such an extensive treatment, why not just a technical projection? Hopefully, it will suffice for now to say that in addition to any logical discussion of the utility of serving a wide audience, there is one economic fact-of-life that must be faced concerning technological forecasting. The price of admission for a technology projection (the Forecast section of the NTF) is a very large percentage of the cost for an NTF type production, approximately 80 percent. There are additional beneficial side effects gained by requiring the NTF broad-assessment-type of forecast which will become apparent only after acceptance and utilization by Navy planning and analysis groups.

NTF Structure

With a definite outline of the information desired about each technology, (as indicated in Table I) for each user, we can turn to the second fundamental issue: which technologies? Better yet, one can ask—what do we really mean by technology? To answer this question, it should be recognized that "technology" is a widely used and abused magic word of the day. It has recently seen expanded use in political circles as concern over the "technology base." We hear statements such as "technology is the solution to many of man's social ills," "technology made the moon shots possible," "technology is making possible the wide application of integrated circuits to everyone's packaging needs." The word is used at many levels of abstraction.

Through this haze there does seem to be a gross consensus of what technology "means." This meaning is summarized by Galbraith² as—"technology means the systematic application of scientific or other organized knowledge to practical tasks." In this sense, technology is a relative term and the full meaning requires association of scientific knowledge with a distinct problem or goal. This may seem somewhat picayune concerning definitions but it has been our experience that communication barriers become visible when trying to explain what one means by technology. To many, technology is used synonymously with areas of scientific activity, say superconductivity. The distinction we wish to make is; to say "superconductivity technology" is to *imply*

TABLE I

GUIDANCE FOR TECHNOLOGICAL FORECAST PREPARATION

1. *Background:* Identify, if possible, the Exploratory Development Goals or other objectives to which the Forecast will contribute, i.e. statements indicating why the forecast is being made. Describe briefly the significant factors which influenced past developments and those which will tend to emphasize or de-emphasize further developments.
2. *Present Status:* Describe briefly the present state-of-the-art. For hardware items cite, in vertically-parallel columns, advantages and disadvantages of present items. Include, where appropriate, a description of limitations (Technological barriers or gaps in technology) which are (or may become) troublesome.
3. *Forecast:* Utilize the functional parameters (i.e. specific weight, shaft horsepower/unit weight, shaft horsepower/unit volume, etc.) which are the most meaningful in your technological area. Describe anticipated changes in complexity, cost and where appropriate, physical characteristics and performance which have the potential for alleviating the limitations. (Charts and/or graphs should be employed where possible.)
4. *Operational Implications:* Utilize the appropriate operational parameters, (i.e. cruising range, CEP's, speed, operating depth or altitude, etc.). Describe the effect of the forecast changes on cost/effectiveness, manpower requirements, and any other factors affecting operational efficiency. (Charts and/or graphic techniques should be employed, where possible.)
5. *References:* Cite the publications from which authoritative direction has been elicited and list technical documents in this function subarea which tend to add credibility to the forecast.
6. *Associated Activities:* List the organizations who have contributed in this forecast area.

required the submission of forecasts for the areas. This constituted a major portion of the Technological Capabilities forecast.

Over the years, the Navy has evolved a structure of problem areas in the Exploratory Development phase of the system acquisition process. The structure, the Exploratory Development Planning Structure, has two very desirable features for long range technological forecasting. First, the problem areas have evolved to describe the unique needs of the Navy, giving a communicative base with the technical community. Second, the categorization of the structure is a functional description of Navy problem areas, independent of specific system concepts. The structure describes the generic, continuing problem areas of the Navy. It served as the base for the structure of the twenty-year Navy Technological Forecast.

Figure VI-3 shows the Navy Exploratory Development Planning Structure. It is a three level hierarchical structure, beginning with the military responsive functions of Command and Control, Target Surveillance, Weaponry, Vehicles, and Support. These accurately reflect problem areas of the Navy, however, they are too broad for a technology assessment of the type envisioned in the NTF. The second level is identified as the sub-areas of the Exploratory Development Planning Structure. It maintains the functional aspects and overall completeness required for long range forecasting. However, the sub-areas are still quite broad for overall NTF needs. At the third level of the structure, the problem areas begin to take on specific orientation for platforms, technical approach, or continue with a further breakdown of functions. Although the breakdown does reflect the uniqueness of each area, the third level does not represent a single specific orientation. There is no reason the present structure should; however, for technological forecasting it would be desirable to be comprehensive with all forecasting categories having the same orientation.

Specifying Forecasts

For the first Navy Technological Forecast, the Exploratory Development Planning Structure was modified slightly to serve as the definition of Navy functional technology. The second level was maintained for completeness, with the third level modified to give specific overall orientation. The orientation selected was platform/target problems; air, surface, undersea.

A technological forecast was requested for each third level category. For example, a forecast was requested in Weaponry Propulsion, Undersea. These forecasts are called Broad Area Forecasts of functional areas, indicating the intent to have a broad assessment of a general problem area.

The Broad Area forecasts represent the continuing problem areas in which the Navy required forecasts. Provision was made in implementing the NTF to encourage the submission of forecasts in areas the Laboratory/Centers felt were important. These forecasts are identified as In-Depth forecasts. One of the first acts in implementing the NTF was the solicitation of In-Depth forecasts. The Laboratory/Centers submitted approximately eighty proposals covering Probable System Options and Technological Capabilities. The Chief of Naval Development approved and funded eighteen In-Depth forecasts of Technological Capabilities.

The Broad Area forecasts were prepared by the Laboratory/Centers through overhead funding. A total of 75 Broad Area forecasts were requested. The assignment of activities to prepare the forecasts was based upon laboratory missions. Each laboratory was assigned a group of forecasts to prepare. Each laboratory was instructed to prepare the forecast with the entire Navy in mind, and to contact other laboratories for inputs, assistance, and review. The laboratories selected the individuals to prepare the forecasts. In general the forecasters were technical experts, as opposed to system analysts or operations researchers.

Response

In a period of approximately three months, the Laboratory/Centers prepared Technological Capabilities forecasts covering the assigned Broad Areas. The Technological Capabilities forecasts were submitted directly to CND from the laboratories. In response to the request for 75 Broad Area and 18 In-Depth forecasts, CND received approximately 600 forecasts. Of the 600 forecasts received, 300 were accepted and 300 were returned to the preparing activity for modification. Of the 300 returned, 190 were resubmitted and were used in the published NTF.

The submission of 600 forecasts in response to a request for 75 forecasts was not due to a zealous desire on the part of the laboratories to forecast. The large number of forecasts indicated an incompatibility or misinterpretation of the intent of Broad Area

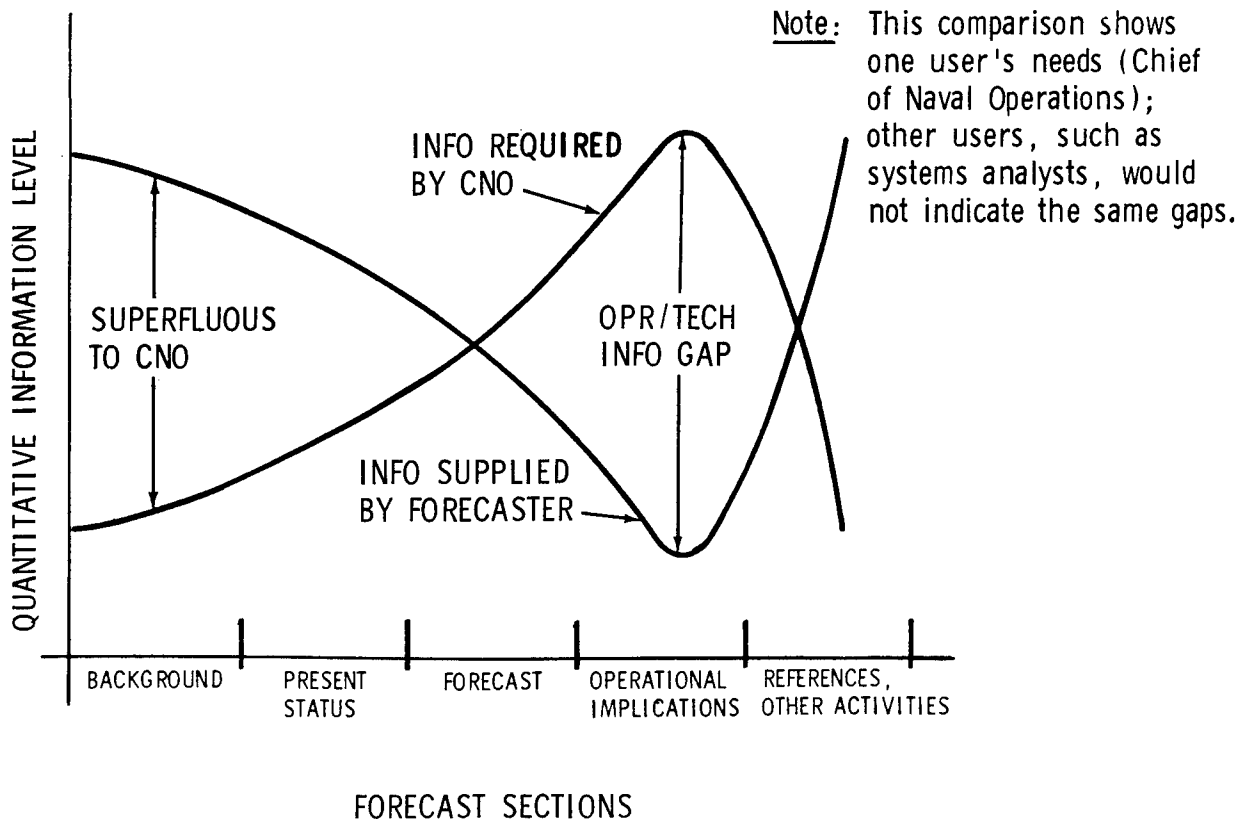


FIGURE VI-4.

forecasts. In those Broad Areas where a great number of forecasts were submitted the forecasts appeared to be prepared under the following interpretations:

- a. preparation of a forecast for each project in progress at the time of the forecast.
- b. preparation of a forecast for each known technical approach to the broad functional area.

Relative Quality

Each forecast received by CND was rated for forecast quality. *The rating made no attempt to determine technical validity.* Rather, the rating attempted to determine the extent to which the forecast met stated objectives. The objectives were:

1. the forecast is to be a function of time, resources, and confidence level,
2. compliance with overall format, and

3. quantitative in discussion in all forecast sections. A section-by-section indication of relative quality is shown in Figure VI-4. Also indicated on Figure VI-4 are the relative needs of one user, CNO. The Background and Present Status sections, in general, more than meet minimum levels of required information. The Forecast section drops off slightly and the Operational Implications section is by far the lowest in relative quality. This should not be interpreted as poor or invalid information in the Operational Implications section. It reflects an assessment of the *type* of information contained in the section which appears to reflect the type of individuals who prepared the forecast. As was indicated previously, the majority of forecasters were technical experts, people required to dig deeply into technological areas, having little experience in Navy operations. The best balance across the forecast sections was achieved by activities that used a technical expert for the Forecast section and a systems analyst (or operations researcher) for the Operational Implications section.

TECHNOLOGICAL FORECAST HIGH POWER DENSITY BATTERIES

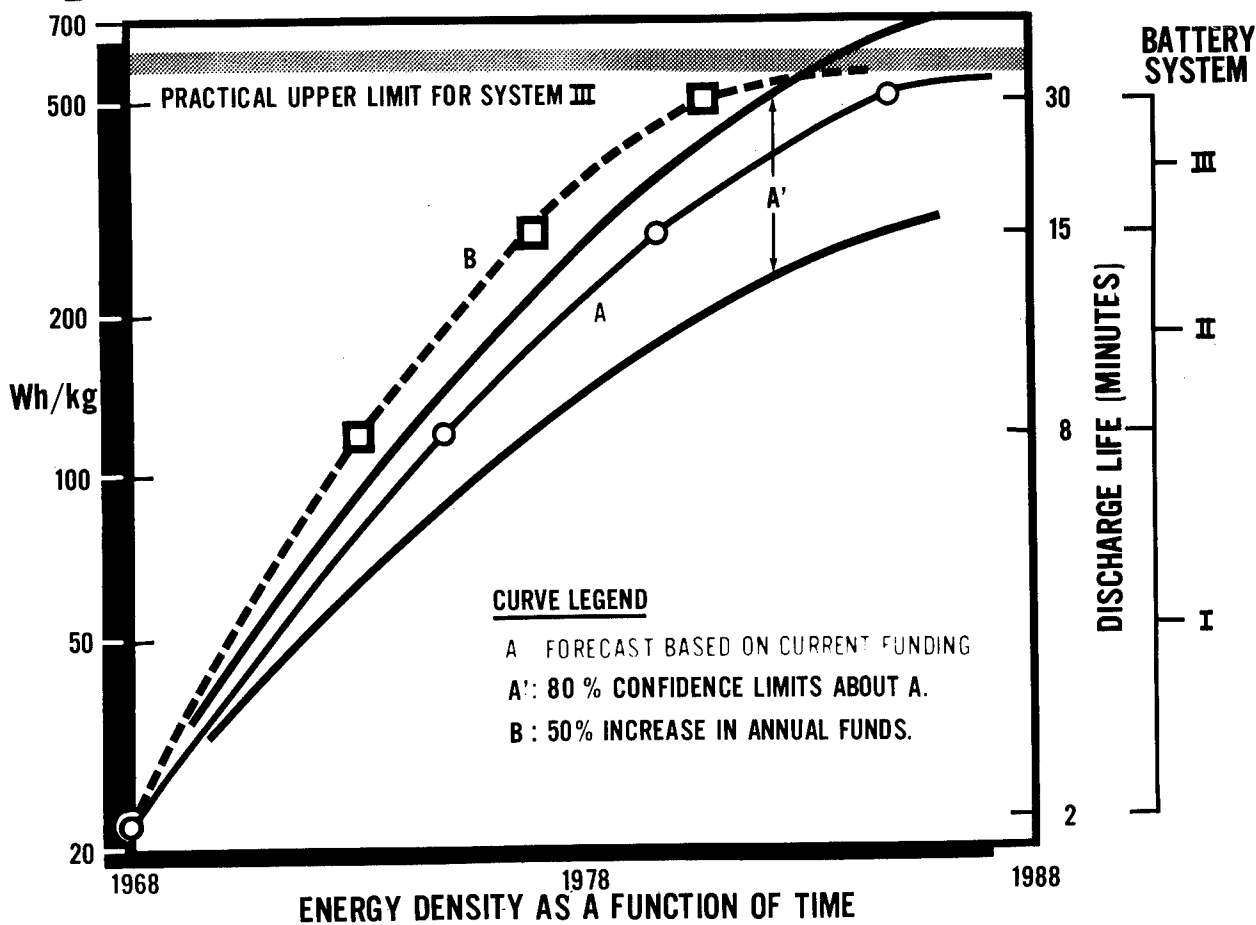


FIGURE VI-5.

Methodology

The format of the Technological Capabilities forecasts required each forecast to be both exploratory and normative. The forecast section was primarily exploratory while the Operational Implications section was need oriented. Several Forecast sections were oriented toward specific goals,³ however, most forecasts are exploratory.

The techniques used for forecasting were primarily trend extrapolation. The techniques used and the approximate percentage of forecasts are shown in Table II. There was one major Delphi effort.

Technique	Percentage
Trend Extrapolation	80%
Trend Correlation	10%
Growth Analogies	5%
Intuitive	5%

General Assessment

As an overall assessment of the first Navy Technological Forecast, as represented by Part II, Technological Capabilities, the verdict is good. The actual utility, in the end, depends upon the specific information required by a user, however, the NTF contains some excellent examples of forecasting. Figure VI-5 is an example (sanitized) of a projection that satisfies the requirements of the NTF. These are:

- a pacing parameter has been identified and is used as the basis of projection.
- the projection is a function of time (20 years).
- the projection indicates expected values and confidence limits.
- the variation to funding changes is indicated.
- the known practical limit is identified.

**TECHNOLOGICAL FORECAST IN BREADTH: METALLIC MATERIALS TECHNOLOGY
1968-1988**

METAL APPLICATION	FORECASTED PROPERTIES	OPERATIONAL IMPLICATIONS
Structural	50 - 100% increase in strength	50% increase in payload, range, speed or operating depth of military hardware
Machinery	50% increase in strength	15% reduction in weight and/or size of vehicle machinery
	25% increase in maximum useful temperature	7% increase in power plant efficiency
Corrosive Service (sea water)	100% immunity to sea water corrosion will be achieved with selected metals	Machinery and instrument failures due to sea water corrosion will be minimized or eliminated
	60% increase in sea water corrosion resistance of steel, magnesium and aluminum alloys	15% decrease in weight of ship hulls due to reduced corrosion allowances Reduced need for corrosion protection systems

FIGURE VI-6.

Another example is shown in Figure VI-6. This forecast demonstrates how the Forecast section and Operational Implications sections can be combined to give a concise, quantitative assessment.

Approximately two months after the first NTF was published and distributed, a critique of the overall effort was held. The critique was conducted as a workshop and attended by representatives of CND, the system commands, and the laboratory/centers. The meeting allowed a useful exchange of ideas and methods to be used for any future updatings of the NTF. Several of the items discussed were:

a. **Technology Areas**—The response of several hundred forecasts for Broad Functional Areas indicated the necessity to develop a technology structure compatible with expert working areas as well as general Navy needs.

b. **Pacing parameter**—In many forecasts a pacing parameter was not identified. This was expected, since the task of identifying pacing parameters is generally very difficult.

c. **Quantification**—The quantification of information is the key to forecast utility. Many forecasters are reluctant to be precise about a future where they doubt the overall accuracy of the projection. Both forecaster and user must recognize the need to think in terms of confidence levels of projections.

d. **Relative Quality**—It appears there could be a considerable increase in relative quality across forecast sections by utilizing different types of people to prepare different sections.

e. **Data Bank**—Many forecasts did not indicate specific data points in the projections. This could indicate the absence of any organized data bank for some of the technology areas.

Utility and Future Plans

Ultimately, the value of the NTF will be shown by the use of the document. Although we cannot indicate known specific uses here because of the unclassified nature of the paper, the NTF not yet one

year old, is being used extensively by some very influential (Navy-wise) groups. As could be expected, the first groups are those who must cut across technological lines or who must gain an indication of future operation implications.

As of this writing the laboratories are reviewing, and updating, where necessary, the forecasts submitted for the first Technological Capabilities forecast. We anticipate a relatively small percentage of significant changes, most will be updating parameter selection. The Office of Naval Research has completed the preliminary draft of Part I, Scientific Opportunities. The Navy Systems Commands are greatly expanding the systems contained in Part III, Probable Systems Options. In addition, the Systems Commands are referencing the options to the projections contained in the Part II Technological Capabilities forecast. In October 1969 the Navy will publish a complete assessment of technology—from scientific opportunities, through the functional technologies to a multitude of system options. Although there are no formal plans at present, this common experience across the entire Navy technical community will provide an opportunity during the upcoming year for a comprehensive

in-house review of all aspects of forecasting for the Navy. The Navy Technological Forecast, along with the Exploratory Development Goals, are fast becoming a formal part of the Navy planning system. The Probable Systems Options are presently being recognized by CNO as a predecessor to the preparation of Planned Technical Approach (PTAs). The Navy Technological Forecast is being used by people who live by the golden rule: "He who has the gold, makes the rules." As such, it is becoming a valuable aid to planning and funding technology.

REFERENCES

1. "A Proposal For a Navy Technological Forecast, Part II—Back Up Report," by M. J. Cetran, et al, Hdq NMC, May 1966, AD 659 200.
2. "The New Industrial State" by J. K. Balbraith, Boston, 1967 Houghton Mifflin.
3. "Navy Exploratory Development Goals," Chief of Naval Development, November 1967.
4. "Navy Technological Forecast," Basic Volume plus appendices B, C, D, E, F, and G, Chief of Naval Development, October 1968.

DISCUSSION FOLLOWING PAPER GIVEN BY DR. DONALD DICK, NAVAL ORDNANCE LABORATORY, WHITE OAK, "PRODUCING THE NAVY'S TECHNOLOGICAL FORECAST."

MR. DARRACOTT: In addition to what you mentioned, the other services do not have the Exploratory Development Goals that the Navy has, but they have similar documents. One is the Combat Development Objectives Guide (CDOG) which contains objectives in three different forms. The other is the R&D objectives documents of the Air Force, of which there are about sixty-five. They give the present state of the art, what the technical barriers apparently are and the forecasts and what their objectives should be. I was just adding information to your comment.

MR. DICK: I am not familiar with the Army and Air Force equivalents to the Navy's Exploratory Development Goals. The Navy Exploratory Development Goals quantitatively describe a specific military function necessary in the mid-1980's (such as counteraction against an undersea target by an airborne platform) without constraining any system

approach to the function. It is this "quantitative, yet open-to-all-solutions" feature that allows the Goals to be useful in guiding research and exploratory development.

MR. BENNETT: What is a pacing parameter?

MR. DICK: A few months ago I would not have hesitated to give you a precise definition of a pacing parameter. Today I say the definition will depend upon who you talk to. The reason I qualify the definition is this: At NOL we are engaged in a study that attempts to relate technological forecasts to the Exploratory Development Goals as a basis for resources allocation. One main consideration of this study is the identification and projection of a technology area's pacing parameter. Working with technology experts and systems analysts to identify the pacing parameter has indicated that a "pacing parameter" is a concept and is subject to personal interpretation. One interpretation is as follows: the

progress of a technology can be described by various technical parameters. Of these parameters, if a single parameter can be identified which is a barrier to progress, it would be the pacing parameter. In many technologies a single technical parameter does not convey information concerning progress. For this reason the growth or progress of a technology is sometimes typified by hybrid technical parameters or ratios. An example is materials technology where "strength/weight" is a widely used indicator of progress. It is recognized that any specific application will require consideration of additional factors. In most cases a single pacing parameter is difficult to identify. To me, a pacing parameter is a technically measurable, quantifiable indicator of progress. A pacing parameter's relationship to other parameters is analogous to the critical path's relationship to other paths in PERT.

MR. BENNETT: So in a sense a pacing parameter is a variable which is a constant over a certain range.

MR. DICK: No, it generally has a rate of change. The barrier can be a limit to the rate of change or a physical or theoretical limit which is approached asymptotically, as with the classic "S-type" saturation curve.

MR. BISCI: Don, you said pacing parameter; but in any given technology, there may actually be several parameters that could be so-called pacing parameters.

MR. DICK: The search for a single pacing parameter is very difficult. It is easier to describe a technology by many, seemingly independent parameters. A normative forecast should have only one pacing parameter. An exploratory forecast is more difficult but still should be typified by one projection. The identification of a pacing parameter is not a permanent situation. As breakthroughs are made in one pacing parameter, the application limit can shift to another. It is these fundamental issues of forecasting like the identification and projection of pacing parameters that forcefully drive home the significance of the selection of areas to be forecast and the specification of the types of forecasts.

MR. HARRIS: I was going to ask what kind of implications this forecast would have on your RDT&E budget? In other words, we are talking about a big RDT&E wave here, do we translate this into dollars and cents?

MR. DICK: The Navy Technological Forecast?

MR. HARRIS: Yes.

MR. DICK: At present, I do not know of any procedures that formally incorporate the Navy Technological Forecast into the RDT&E budget. Its

present place is as a reference for those who make budget decisions; however, it is a prime input to planning.

MR. RUMMEL: You have the category "operational implications?" Is this category of implications largely with reference to technological forecasts or technological hardware or does it take into account the political and military environment, or alternative environments?

MR. DICK: The operational implications section of the Navy Technological Forecast does not take into account any political or military scenarios per se. The extent to which these are incorporated depend upon the individual forecaster. The instructions for the '69 updating state that the threat environment is to be discussed.

MR. SIMON: You mentioned at the start that while you were asked for broad-area forecasts, in the main you received "in-depth" forecasts. This is not an apology but an explanation. Let's take the Naval Training Devices Center as an example. We procure or produce training devices across the board for Air, Surface and Sub-Surface Training. We also have Air Force and Army participation in our program. So, to give a broad-area forecast on training device technology alone would be so voluminous to prepare that, though we started with 57 different forecast areas, we ended up with just seven to cover all 50 areas. The idea is to update and expand on these as time went on. I do not see any possible way for us to give a broad across-the-board technological forecast in the training device area.

MR. DICK: Constraining and forecasting a broad area of technology requires the interpretation of data at a different level of abstraction than for an in-depth forecast. Many times the question of whether or not a broad area forecast is possible is a function of whether it is viewed at the suggested level of abstraction. In other cases where the output must await a specific system, perhaps the area should not be called a functional technology area in the sense the Navy uses the word technology. The Navy sense is that functional area technology is synonymous with exploratory development sub areas.

DR. LINSTONE: One of the problems we have been thinking about is that of technological failures, meaning operational failures of complex technological systems. For example, we are familiar with outstanding successes in strategic missiles and in space programs. Failures seem to be most apparent in areas such as air defense and tactical weapons. The enemy is not static but accommodates himself to the new U.S. weapons. There are lessons to be learned for

future planning. For example, one might consider a single-use weapon. With this concept there is no real time for the enemy to develop countermeasures. He has no time to learn first what is weak about the system. Another approach is the modular system which allows you to quickly change the modus operandi and keep the opponent from countering effectively. So this suggests new technological concepts. What I am wondering is: is this kind of technology, technology based on lessons learned, taken into account sufficiently? There are a lot of things implied by just looking at what has happened, but they involve directions that are somewhat different from just the extrapolation of the familiar technology.

MR. DICK: The Navy Technological Forecast does not specifically address total system theory and operations. There is a category of the Exploratory Development Planning Structure covering this area but the Forecast functional breakdown does not cover it at present. I agree with you that any functional description of a system must describe the operation of the system in the total environment. Total system concepts are a technology in themselves. My presentation was primarily concerned with the Navy's broad area forecasts. The Navy has an in-depth forecast which allows spontaneous contributions. Anyone can submit an in-depth forecast in the total system areas; however, there is no formal solicitation of them at present. There is an organization within the Navy under CNO that works in the area.

MR. LINSTONE: The people at the lab level very often may not see what has happened to their weapons in a previous generation of weapons, how they failed, why, and so on.

MR. DICK: Yes, this is only one part of the overall problem: what part should a laboratory play in the entire weapon system acquisition process? I think a laboratory should be involved but I do not know whether technological forecasting is the way to address it. This involves the concept of in-house capability. We do not want to go to the extreme and have a laboratory concerned entirely with "quick-fixes." A laboratory must be prepared to answer two questions: what could be? What is your need? The amount of work spent in each area depends upon the sponsor and the funding.

DR. LINSTONE: I would suggest that in-depth analysis of failures should be highlighted much more in developing new technology.

MR. DICK: Yes, I agree. The first step in developing any new techniques should be a thorough analysis of

present problems and procedures. The point I wanted to make concerning Navy laboratories is that the laboratories must be concerned with all aspects of "in-house capabilities," from research through the operational systems assistance. The labs must avoid the extremes on either end.

MR. SCHNARE: As you went through this technology forecast, did you find a lot of revolutionary ideas, a lot of extraordinary concepts of new weapons developing or falling out?

MR. DICK: The Navy Technological Forecast has one part concerned with Probable Systems Options. This section covers new concepts, at least from the system view. It is a matter of personal judgement whether or not they are revolutionary. I can give you an example of a "revolutionary" concept and how the Navy Technological Forecast can give visibility that might not exist otherwise. About two years ago DOD sponsored a study in an area of concern to the laboratories. The Navy Laboratories got together and proposed a solution. It was radical—radical in the sense that it was different from what exists today; however, problem areas were identified and the concept was based upon analysis. Because of time and other factors the concept was not evaluated properly. This concept is now in the Probable Systems Options of the NTF. The Probable Systems Options is a matter of official record. All concepts are available for review. As a matter of fact, this concept, because it was published in the NTF, has been reviewed along with the other thirty probable systems options. The review was made by the Deputy Chief of Naval Operations for Development (DCNO (D)). DCNO (D) recommended specific actions be taken for each concept such as prepare a PTA, update the threat, etc. In this sense the NTF does offer a visible platform for offbrand ideas. This is not perfect, but we now get visibility of ideas we might not have had otherwise. DCNO (D) is establishing formal procedures for annual review of the NTF. We in the laboratories know that every idea will be evaluated at least once each year, and supporting rationale for any decisions will be a matter of record. This is a great leap in objective decision-making concerning system concepts.

MR. IRWIN: I wonder if this revolutionary idea which you just cited could be called intuitive. I thought I detected in your discussion a deprecating reference to the idea of an intuitive forecast. You would almost have to be a little intuitive in order to transcend the boundaries of the individual areas and come up with a broad area idea. I wonder if you would comment in this specific case with respect to

these two aspects. Was it a broad area suggestion and was it intuitive?

MR. DICK: The Probable Systems Options is completely separate from the Broad Area Forecasts. For a concept to appear in the Probable Systems Options, it must be supported by analytical or experimental studies. It is not a wish list. There was some intuitive judgement that went into the analysis of the concept under discussion, as on any conceptual system. Concerning the remarks about intuitive forecasting, I hope I did not imply that there was anything inherently "bad" in this approach. Intuitive judgement enters every forecast, it is simply a matter of degree and whether it is visible.

MR. IRWIN: I noticed you were a little disturbed by this idea of an intuitive approach and you said so. Suppose you were given an intuitively good idea that was not backed up sufficiently by numbers. What would you do with it?

MR. DICK: I would say that is one expert's opinion. It would stand as a forecast until someone challenged it. I am a firm believer in the corrective aspect of the scientific method. As to your other comment, there are some intuitive forecasts by one person who does not use data or analytical methods when they are available. Furthermore, the lack of data or systematic procedure does not allow anyone else to arrive at the same conclusions. If data does not exist, it may be better to go to a systematic procedure with a group of experts. Of course even a high level group can and does go astray. For example, Dr. Von Karman, who was mentioned briefly, and a group of high level experts made a forecast concerning gas turbines in 1940. The committee, which included men like Hugh Dryden and Lee Dubridge, came up with the following prediction in June 1940; "In its present state, even considering the improvements possible when adopting the higher temperatures proposed for the immediate future: the gas turbine could hardly be considered a feasible application to airplanes, mainly because of the difficulty in complying with the stringent weight requirements." However, there were people in England and Germany who went ahead and developed a gas turbine for aircraft.

MR. BISCI: We got into intuition in our contribution to the Navy's forecast and I think you will agree that intuition is acceptable in technological forecasting, but in a probable system option, I agree with Dick here: if you propose a probable system and it is based strictly on intuition, "Go do a little bit of analysis and come back again and see if it is feasible."

MR. DICK: As far as the Probable Systems Options

of the NTF is concerned that is correct. At the NTF Critique the System Commands specified requirements that system options must meet to be included in the NTF. Supporting analysis was one of them.

MR. BISCI: Was it not more or less agreed that one just based on intuition was not acceptable? They want to know a few numbers to see if it is at least feasible

MR. DICK: The Probable Systems Options must be supported by feasibility and/or experimental analysis.

MR. STERLING: In talking about intuitive forecasts, I think the intent is not to criticize use of intuition alone but to use it with a rational process. If you have a collection of forecasts, say, 490, including many intuitive forecasts, it is hard to evaluate the forecasts. They do not spell out the assumptions or present the data they used or the logic process they went through. It is very difficult to really evaluate these things: on an intuitive basis, you will always have an intuitive gap to jump across. The point is: do you try to do a rational explicit way of developing a forecast and then explain it?

MR. IRWIN: The thing that strikes me though, it is kind of a key point: Don, you said the intuitive approach, the really original idea, will not come from the technical expert but from someone outside who, by definition, is not going to be able to back it up with figures as well as the technical expert. In some way you have to strike a balance here. It seems to me that if you are going to give the new original idea the acceptance it deserves, you have to downgrade the numbers a bit.

MR. DICK: Are you referring to this joint effort between two types of individuals preparing a forecast?

MR. IRWIN: Well, you mentioned that as one solution. I think you said the more imaginative ones covered broader areas than a collaboration between a technical expert and a research-oriented person. Maybe that is your solution.

MR. DICK: They satisfied what was asked for in the forecast, which may or may not have anything to do with creativity.

MR. IRWIN: That is a good point.

DR. SLAFKOSKY: I think we are confusing, at least in part, different sections of the technological forecast in question. The primary purpose of the forecast is to ascertain where the technologies in which the Navy Department is involved will be in a number of years, and what they can reasonably expect to produce. Of course we realize that at the laboratory

level many of the people are very familiar with current technology and are in an excellent position to state where it might go and at what rate. Yet these laboratory people would not necessarily be the ones who would generate the long range concept(s).

MR. DICK: I would like to add one further thought along this line; the sequence of the various Navy forecasts. The sequence was visualized as follows: ONR would prepare in-depth forecasts in scientific fields. With the scientific forecasts as a base and with defined functional problem areas, the laboratories would prepare Broad Area forecasts. Next, the System Commands with the Broad Area forecasts as input could gain insight into possible system alternatives. However, the Navy had to start somewhere, and it was the Broad Area forecasts. There were some forecasts of Probable System Options. For the FY 69 updating, the relationship between the three parts of the NTF will become apparent. The System Commands are referencing all system options to the supporting Broad Area forecasts.

MR. BROWN: Mr. Dick, do you have evidence that your testing facilities or your measurement ranges use your forecast in any way? In other words to get guidance on future instrumentation to satisfy some of the problems that you propounded here in the system efforts or otherwise?

MR. DICK: At present, the NTF is not formally tied to any future plans. However, at NOL we are using the Navy Technological Forecast in an attempt to develop an advanced plan. Hopefully, we will be able to gain from this study the type of information you suggest. Our method is still experimental. Outside the Laboratory, several groups have been using and are using the NTF for ship development and intelligence assessment.

MR. SCHNARE: You have someone in the Navy that does the mission analysis and weapons system analysis and defines the most worthwhile systems. I know you said it was headed by a committee but surely there must be some specific organization set up other than that.

MR. DICK: The Deputy Chief of Naval Material for Development is Admiral Goodfellow. I do not know of any formal methods he has incorporated. Members of his staff have published articles in this area and several techniques are being tested.

MR. SCHNARE: My second question in connection with this is: Since you have so many technology opportunities, as a result of this forecast situation and since we all live in a real world that requires either manpower, dollars or some other kind of resource in order to be able to carry out the necessary

work, how do you sort out the priority or that part which is to be funded over that part which is not to be funded?

MR. DICK: Of the Broad Area Technologies?

MR. SCHNARE: Yes.

MR. DICK: The funding priorities depend upon the balance between the near term and the long term.

MR. SCHNARE: I am not only thinking of the near future but some things which appear to be very desirable that require that you get started now, so that means far out as well.

MR. DICK: At the heart of the resource allocation method I have referred to is the time frame to be used. We are looking at 1980 and asking the question, "What capability should a laboratory have in the various technologies to support system development at that time?" We selected 1980 because it takes 7 to 8 years to develop a system, and also, we cannot do anything about resource allocation in the next few years. We are looking at the relationship of resource allocation-to-technology much like one views strategy-to-tactics: once someone decides to develop a system, we want to have the capability to support it, and help insure its possibility. It is long range planning, allocating resources today for future capabilities.

MR. DARRACOTT: Don, just as an answer to your comments about intuition, the National Research Council before World War II made a long range prediction of scientific activity and advances and they missed several things like atomic energy, jet propulsion, radar, and antibiotics, so people are not infallible.

MR. DICK: The American Association for Advancement of Science has shown over about fifty or sixty years that breakthroughs cannot be predicted, and I agree with this wholeheartedly. That is why I like the Navy's approach to total assessment; forecast the problem area and leave the door open for any technical approach, especially since the forecast will be updated every year. If we look through history, it is quite relative what we call breakthrough. Transistors are sometimes identified as one, lasers are identified as another. Our hindsight is always very good. These points in time are very clear in retrospect; but, it still takes a long time for the concepts to get into a system. Before anyone has a lock on the door, I think these major technical approaches tend to become obvious. What is required is for people to be prepared for such advances, be familiar with the problem area and apply new technical approaches as they become available.

MR. HARRIS: Could I ask a question? To what extent did industry formally and informally participate in this? I realize this was a six month effort overall. Let's say, in the future.

MR. DICK: It depended upon the activity that prepared the forecast and the individual forecaster. The forecasters were told to take the entire Navy and to the extent possible, the best people in the field into view, which meant the universities as well as private industry. One forecaster I know of used all sources. He prepared the forecast and sent it to all experts in the United States industry, universities and the Navy. He incorporated all comments into the final version of the forecast. There were other forecasts that were prepared totally in one room. Industry was not formally invited to participate by CNM, it was up to the forecaster. All forecasters for Part II were from Navy activities.

MR. SIMON: I would like to correct the impression that this has been a six-month effort in toto. The six-month effort was in preparing this first technological forecast. Actually, the groundwork for the forecast effort was started back in 1965 at the first Washington Meeting where the committee developed the guidelines. The Technological Forecast, as published, represents a six-month effort from when the word was "go". But, in my opinion, everybody had ample lead time. We at NTDC had so much advance notice that we were experimentally forecasting for almost one year. As I mentioned before, we started with 57 forecasts but finding we did not have the manpower to sustain such an effort, we boiled it down to a mere seven.

MR. DICK: Perhaps your laboratory was the exception to the 6 months. When I said 6 months, I am referring to the Broad Area forecast request which was made official in February, with forecasts due by July.

MR. BISCI: First of all, industry contact, industry participation in the forecast: indirectly industry participated by their various contacts with the laboratories in their fields of expertise and the laboratories are aware of the various industry proposals and the best foot that industry has put forth. So indirectly they are incorporated; their opinions were solicited in many cases. With respect to lead time; I know at NUWS, the only warning we received of forecasting was one day a memo came in the mail from the Technical Director designating personnel for forecasting, and that was well within the six month period we are talking about. At that time we had to indoctrinate ourselves with the literature avail-

able and do the forecast in a relatively short period of time.

MR. SIMON: It is surprising that you remember the story that was mentioned before: In 1965, Admiral Pinney, who headed the original forecast group, told this story to all the lab heads in attendance. Many of them were not particularly interested since they felt they had enough to do without forecasting. One laboratory head asked, "Do we have to make a forecast?", to which the Admiral replied, "No you don't have to make a forecast,, but let me tell you a story you can take back to your lab. There was a corporation president who was offered an extremely economical group insurance plan for his employees with only one stipulation. Every employee had to sign up all the way down the line. To make a long story short, everybody signed up except Joe down in supply. The president asked the foreman to talk to Joe, but with no success. He next asked the plant superintendent, but still no success. Five days and five layers of management supervision later, Joe was ushered into the President's office where he was told: "Here is your insurance form and here is your letter of resignation; sign one." He immediately signed his insurance slip whereupon the President said, "Holy, Moly! After six months of solid refusal you sign just like that, why?" Joe replied, "You are the first man to explain this insurance thing to me in terms I could understand." It is my contention that as far back as 1965 all Navy Labs should have known that the forecast effort was in earnest. As I mentioned before, we at the Naval Training Center had a year's drill in preparing forecasts before the word came out to go.

MR. SCHNARE: How did you establish communication with and what was the impact on the laboratory and the scientists? Are they convinced that forecasting is necessary for the future? Normally they take the attitude they know everything already and they do not need any external guidance.

MR. DICK: My personal experience indicates the response varies across the entire range. At NOL many of the forecasters were skeptical, and justifiably so, of this new demand on their time. Many of them still have this opinion. Some forecasters got the point right away, what was implied, and they went about in a scientific manner and were not emotionally involved. There was a group of people who were quite skeptical. I would like to use one as an example. This was a rather salty individual. He had seen planning ideas come and go, new ways of doing things, magic words of the year, and so on.

He knew his field and was assigned to prepare a broad area forecast. He was reluctant; he did not really like the idea, but he went along with it. At the end after he went through the process of preparing the forecast, he came up to me and said, "you know, I was really down on this forecasting when I

first started out; but now I can see some good in it. As a matter of fact, in trying to search for the information I needed for the forecast, I realized there were really no sets of comprehensive literature in the field." To the best of my knowledge, he is writing a book on the field.

Forecasting International Relations: A Proposed Investigation of Three-Mode Factor Analysis

by RUDOLPH J. RUMMEL*

INTRODUCTION

People in any society live out their daily lives within a field of expectations generally based on unconscious forecasts of each others behavior. We expect that our wives will have dinner for us when we get home, that our evening paper will arrive, and that our salary checks will be on time. As we calmly drive down any highway, we made implicit forecasts that others are going to stop at stop signs, drive on their side of the road, and pay attention to the traffic lights. Only when such predictions fail, sometimes with grave consequences, do we realize how much we take such forecasts for granted.

Nations, like people, also live out their daily affairs entwined in a field of expectations. And as with people some expectations are so without doubt that the forecasts involved are rarely surfaced. Canada will not attack the U.S., and so our common border is unfortified. The foreign ambassador wanting an audience with the President will not try to harm him, so the President is left unprotected. And foreign commercial airliners soaring above American cities are on legitimate business, and need not be guarded against. Unlike people living in some kind of community, however, the field of expectations for nations is more conscious. These fore-

casts are most often explicit. Two reasons account for this.

First, while the society of nations is the same as any community of people in that laws, roles, and norms, govern activity, international society is not as tightly laced with norms, derived from the culture, roles to enshrine the division of labor, and laws to explicitly demarcate and to regulate the permissible. For most communities the expectations that people have about each other's behavior—the forecasts—are taken for granted. People are acculturated to act in certain ways, to perform many roles, and sanctioned if they do not. Nations, however, are not firmly bound into such a normative system, although it is a mistake to deny that such a system exists. Therefore, national behavior is less predictable on the basis of shared expectations and forecasts taking into account the possibility of error have to be more explicit.

Second, the analogy between behavior of people and nations can easily be overplayed. Nations are deliberative bodies. Even long accepted policies come under frequent review (if not bombardment from interest groups) and decisions that are made are often filtered through a long decision-making administrative, and semi-legislative process. The unconscious expectations laid down by culture are closer to the surface for people than nations. Nations, because they exist in a looser society and because their actions result or are sanctioned by a deliberative process, must question what they are doing or what they are about to do. Their forecasts must often be made explicit.

Forecasts made by nations are of two types. One is the specific or point forecasts. If we give foreign aid to country X, she will vote with us in the U.N. on seating Communist China. If we raise duties on country Y's goods, she will do the same to ours. These point forecasts, however, are made against the backdrop of a second type of forecast—the

* R. J. Rummel is a professor of political science and the director of the Dimensionality of Nations Project, Department of Political Science, University of Hawaii. He has taught at Yale and Indiana Universities and received his PhD from Northwestern University in 1963. He has written *Applied Factor Analysis* and *Dimensions of Nations*, in addition to numerous articles on cross-national research, international relations, and conflict. Currently he is involved in developing a mathematical theory of international relations linking the behavior of nations, including their conflict behavior, to the geographic, social, economic, and political-military distances between nations.

trend forecast. The latter are predictions about what the international system will be like in 5, 10, or 25 years. It is along the lines of these trend forecasts that nations do their major foreign policy and defense planning. Should the U.S. play a larger diplomatic role in Africa, establish firmer commitments in Southeast Asia, reorient defense strategy towards deterring guerilla warfare, and base a weapons system on unconventional warfare requirements? These are questions involving forecasts about the direction of world politics and trends in international conflict.

Point forecasts are only so good as the trend forecasts against which they are made. As a busy man you may predict that 30 minutes is sufficient for you to get from your office to an appointment, but unless you take traffic patterns (trends) into account at that hour you may be embarrassingly late. As a nation we may predict that doing favors for the elite of a African nation will influence it to side with us on crucial issues in the United Nations, but if we disregard political trends in Africa we may find our friends on the execution block.

Trend forecasts in international relations are usually made by those who have long experience with the behavior of nations and the environment within which they interact. Through experience and training, statesman and politicians develop a conceptual framework within which they can describe world trends. This is an intuitive and judgmental process, and although the value and power of the sharp and experienced mind is not to be underestimated, there is a question as to whether projections about the gross trends of world politics and environments cannot be made more systematically.

The analogy with a farmer and weather trends is appropriate here. The aged farmer standing in his field, sniffing the air, studying the clouds, listening to the animals, and feeling the earth, was an excellent forecaster of the weather. Like the farmer, from a number of clues, the trained statesman can forecast the gathering of storm clouds in the Middle East or over Berlin. However, meteorologists can out-forecast the aged farmer, and hopefully, a science of international relations will eventually be able to similarly improve on the statesman.

Meteorology owes its success in the main to a growing empirical knowledge of weather patterns and trends, and a set of basic indicators (e.g., barometer, temperature) of weather conditions. Some research in international relations is now moving along similar lines. Attempting to work towards an ability to do "trend forecasts," work is being done on

defining the patterns in the behavior of nations, the trends in these patterns, and the basic indicators that can be employed to forecast these trends. With a reliable and systematic knowledge and understanding of world trends, it is hoped that a selection of alternative foreign and defense policies by nations will be made more rational, and that there will be a closer articulation between national goals and behavior.

The research described and proposed here is part of this movement to develop a "meteorology" of international relations. Towards this end, the major focus of the proposed research is to investigate a promising new method—three-mode factor analysis—for clarifying trends in national behavior and environments. To place the discussions of this method in context, the next section will describe the Dimensionality of Nations Project within which some of the research on patterns, trends, and indicators is being done. Subsequent sections will then describe two alternative methods for delineating time trends. Following these sections, the nature and possible value of three-mode factor analysis will be considered and the proposed research on it will be discussed.

THE CONTEXT: THE DON PROJECT

The interest in investigating three-mode factor analysis (TMFA) and the direction the proposed research will take are related to the Dimensionality of Nations (DON) Project. This project and its goals form the research context within which a concern with TMFA grew and with regard to which the approach will be studied. It is appropriate to discuss this context at this point, therefore, as a way of giving perspective to the subsequent discussion. First, an overview of DON will be given, with some historical background. Second, the tasks of the project for which it is currently funded will be discussed, and third, the problem of measuring change and forecasting for DON will be considered.

2.1 Overview. The DON Project began in 1962 with a grant from the National Science Foundation. The basic research goals at that time were to define the patterns (dimensions) of variation among nations on their diverse attributes and scores for nations on these patterns.

Research directed towards this goal involved collecting data on 236 attribute variables (such as defense expenditure, GNP per capita, literacy, men under arms, threats, population, and area) for 82

nations and a thorough methodological investigation of the application of factor analysis to such cross-national data as a method of delineating their patterns.

This phase of DON, which is called Phase I, ended in 1966 with very significant results (see Rummel, 1969b, for the most salient of these). Substantively, nations were found to distribute themselves on seven major attribute patterns (dimensions): economic development, size, political orientation, density, Catholic culture, foreign conflict, and domestic conflict. Moreover, economic development was found to be the most dominant pattern clustering national attributes, accounting for nearly twenty percent of the variation among 236 attributes analyzed. When size and political orientation (along a totalitarian, nontotalitarian continuum) are added to economic development, the three together describe almost fifty percent of the variation.

As a substudy, 94 international relations variables (such as treaties, trade, threats, wars, immigrants, tourists, and aid) were also analyzed to delimit the pattern dimensions of international behavior. The results pointed up a general participation dimension for nation. If nations have much activity of one kind, say trade for example, then they are likely to have many treaties, join many international organizations, and have many embassies and legations. Significantly, a second dimension was one of foreign conflict behavior, indicating that conflict behavior is quite independent (statistically) of the magnitude of a nation's participation in the international system.

Aside from methodological results, Phase I proved the usefulness of multivariate methods, especially factor analysis, on cross national data of both the hard (e.g., national income) and soft (e.g., freedom of group opposition to government) variety. Factor analysis was found flexible, efficient, and theoretically fruitful for building theory from the results. Out of the methodological experience with factor analysis during Phase I has grown two publications by the project (Rummel, 1967, 1969a).

As Phase I research was carried on, it became increasingly evident that a different theoretically important and policy-relevant line of research could be taken—a research direction that would deal more specifically with directed behavior *between* nations. Phase I involved analyzing the attributes and behavior of nations: the data matrix employed was a nation by attributes and behavior matrix. The new line of research would entail analyzing the behavior of specific nations to other specific nations, such as

the behavior of the U.S. to China, USSR to Poland, and France to U.K. This kind of analysis would be of dyads (nation pairs, $A \rightarrow B$, $B \rightarrow A$, $A \rightarrow C$, etc.) by behavior (such as exports $A \rightarrow B$, threats $A \rightarrow B$, and treaties $A \rightarrow B$) data matrix.

This new line of research was initiated in 1966 as Phase II of DON. Over a hundred variables were analyzed for around 400 dyads. The results subsequently found on 1955 data were that behavior between nations structures itself into several major dimensions, including exports, international organizations, transactions, deterrence behavior, and immigration.

During Phase II, a theoretical structure—called a social field theory—was formulated and has subsequently served to guide research. The theory is elaborated mathematically, drawing upon the theorems of n -dimensional spaces and linear algebra. It functions to link Phases I and II of DON together so that the research results and design have continuity that they would otherwise lack. Simply put, the theory states that nations are social units interacting in a social field (in some ways analogous to an electromagnetic field), and that this field can be analytically divided into two spaces, a space of behavior and a space of attributes.

Within the attribute space nations are located in terms of their characteristics; within the behavior space dyads such as $US \rightarrow USSR$, are located in terms of their behavior. Now, a distance vector between nations in attribute space can be measured and this vector can be treated as a force acting on the origin of behavior space. In other words, the location of dyads in behavior space can be treated as a function (resolution) of the distances (forces) between nations on their attributes. Mathematically,

$$\mathbf{W}_{A \rightarrow B} = \sum_{j=1}^p \alpha_j \mathbf{d}_j, A \rightarrow B,$$

where

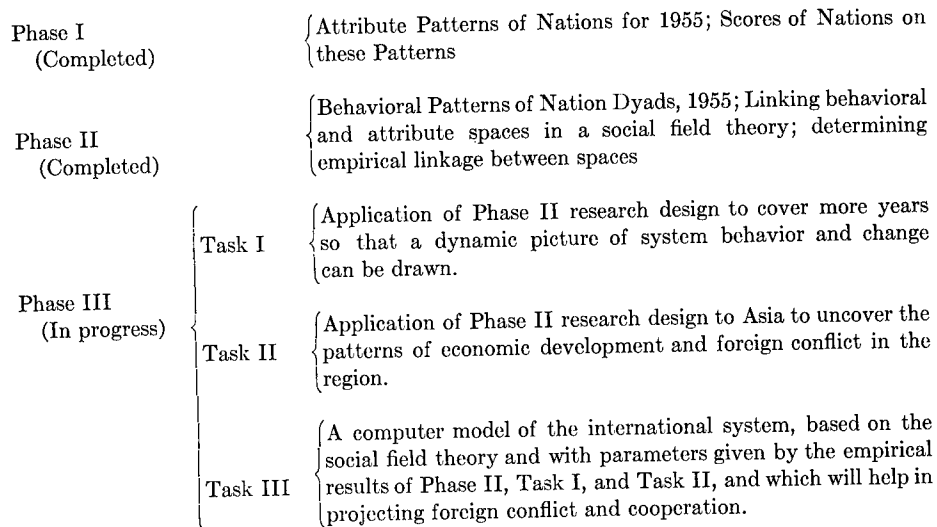
$\mathbf{W}_{A \rightarrow B}$ = location (as a vector) of dyad $A \rightarrow B$ in behavior space,

\mathbf{d}_j = distance vector between nation A and B on attribute dimension j of an attribute space of p dimensions.

α_j = a parameter for dimension j .

Leaving mathematical considerations aside, the field theory says simply that nations will behave towards each other in terms of their socio-economic, cultural, geographic, and political differences and similarities. Tests of the theory have so far been

FIGURE VII-1.



encouraging. In particular, one analysis of the linkage of behavior between nations to their differences showed different levels of economic development to be an especially important force influencing the direction and magnitude of behavior.

Experience with Phase I and II indicated the need for additional research funding and the possibility of relating more directly to foreign policy interests. Consequently, in 1967 a proposal was submitted to the Advanced Research Projects Agency and was subsequently approved to carry DON into Phase III by way of three Tasks. These will be discussed in the next section.

2.2 Current Research. Phase III efforts articulated within a five-year research plan (Rummel, 1968b) involving three tasks. Task I is an acceleration of the research carried out under Phase I and II of DON, and an extension of this research to four time periods. At the completion of this Task, the project will have defined the attribute and dyadic behavior patterns for 1950, 1955, 1960, 1963, 1965. Determination of patterns for each of these periods will allow us to 1) study the shift in interaction between nations, such as Cuba → US, US → France, or USSR → France, during 1950-1965, 2) relate these shifts to changing distances between nations on their attributes, such as economic development, power, or ideology, and 3) measure the existence of systematic change during the period and the interactions, such as conflict, and attributes, such as power, that are most related to the change.

Previous work has been concerned with the global system of nations. It may be that within regions, the behavior of nations distributed along different

patterns and the correlates of behavior are not the same. Yet, a knowledge of conflict and cooperation pattern within regions is particularly needed for rational policy making and theory construction. Accordingly, Task II focuses on the pattern dimensions of Asia. All nations in Asia will be analyzed to determine their cross-national patterns of interaction for 1950-1965. The results will be compared with the patterns of the global system of nations found in Phase I and II of DON, and Task I, above. The patterns found for Asian nations will enable the isolation of basic indicators of regional conflict and cooperation. These indicators will then be employed to forecast the direction of cooperation and conflict between pairs of Asian nations, such as China and Japan.

The accumulated findings from the previous tasks will enable a computer model to be built, with structure specified by the theory within which the research has been done, and parameters defined numerically by the empirical findings, 1950-1965. Therefore, Task III involves programming a computer model of the international system. The purpose of this model will be to evaluate to what degree the theory and findings of DON enable forecasting cooperation and conflict between nations. Tests of the model will be made against data on Asia from Task 2 and against data on nations collected by similar projects elsewhere. These tests will function both to test the model and to determine the degree of theoretical convergence in empirical findings about nations.

By way of summary, Figure VII-1 lists the Phases and Tasks of the DON Project.

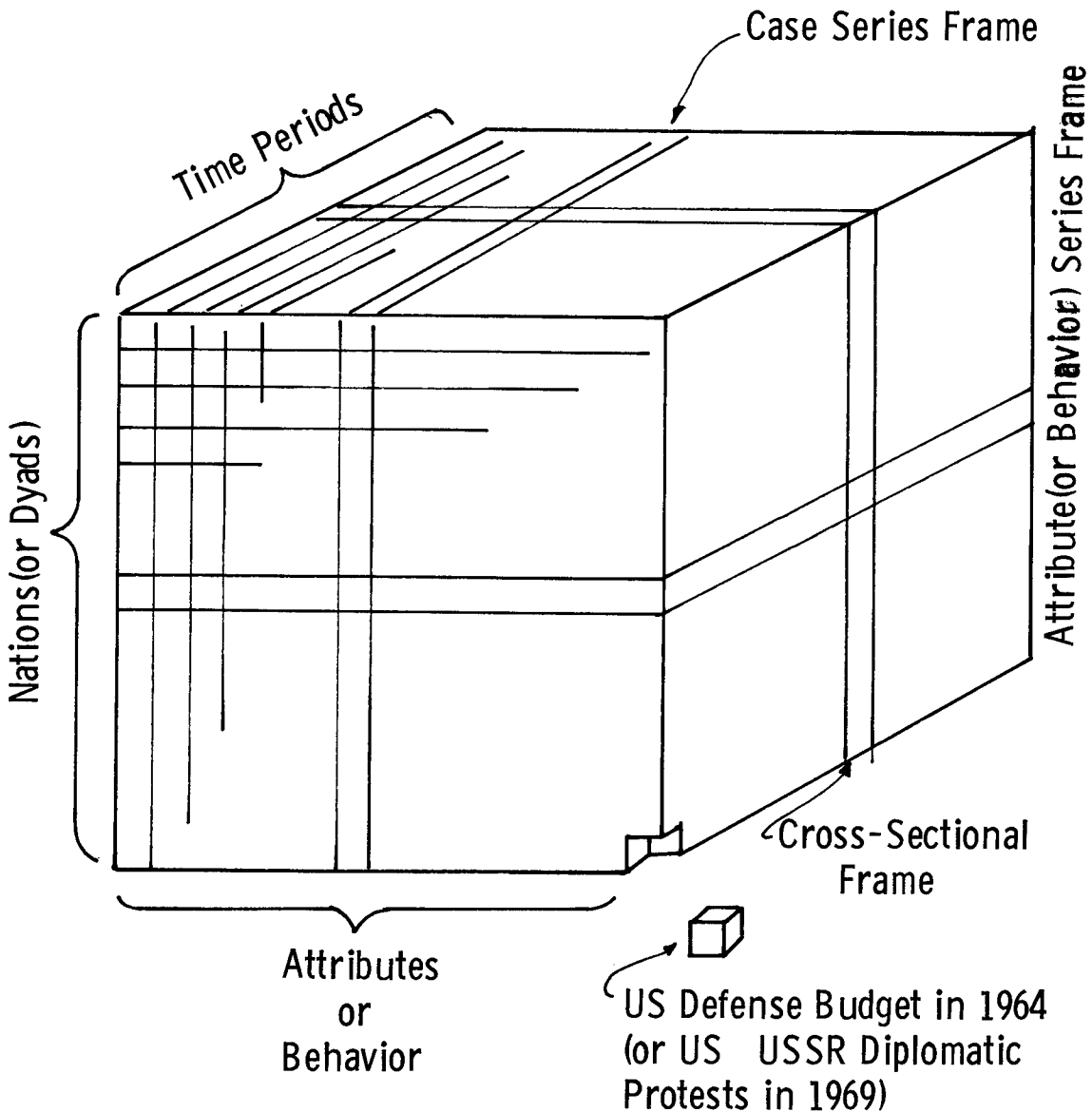


FIGURE VII-2.

2.3 *Problem of Measuring Change.* As can be noted from the previous sections, DON has a basic interest in forecasting to future international environments and behavior. In order to forecast, or to use a word more favored among scientists, to predict, requires a measurement of time patterns. Environments and behaviors interact within and between each other in complex ways. The seasoned foreign policy analyst has become trained to disentangle this complexity into a few basic trends to which foreign policies can be geared. One trend, for example, is assumed to be a growing political division between economically developed and underdeveloped na-

tions—a division that in effect is creating a two-party political system with regard to many world issues. Another trend the analyst sees is an increasing fragmentation of the East-West poles that dominated the world politics of the 1950's and a reshuffling of political orientations, as have been manifest by France, China, and Indonesia. A third trend is considered to be a continuation of instability in developing nations and the possibility of guerrilla and limited war in Africa and Southeast Asia.

The foreign policy analyst develops an intuitive feel for international trends through a lifetime of personal acquaintance with the behavior of nations.

This acquaintance may grow by participation in foreign affairs as a government official or representative, as with George Kennan, or through decades of specialized study of international relations, as with Professor Hans Morgenthau. Some analysts, like Dean Rusk, have both academic and scholarly credentials. After a lifetime of dealing with international relations, the policy analyst's judgments and forecasts are to him quite reasonable. They are conclusions that "sensible men" would draw. He doesn't worry about the number of variables involved, whether the forecasts are within the range of previous data, the functions giving the forecast, and so on. As we try to add precision and reliability to forecasting international relations, however, such questions become serious problems.

Since these problems are ones that DON will have to resolve if Phase III is to be at all successful and since the most promising solution points towards TMFA, these problems might be discussed more fully.

1. Because we wish to forecast the international environment and international behavior, we are faced with phenomena that are, statistically speaking, multifaceted. Consider Figure VII-2, in which national attributes or dyadic behavior lie along three modes:¹ nation (or dyads), attributes (or behavior), and time periods. A collection of data on these three modes would then comprise the cube shown in the Figure. In order to systematically forecast international environment (attributes) and behavior, variation along all three modes must be untangled and the time trends uncovered. For example, analysis along the three modes might well disclose that foreign conflict and domestic instability are distinct patterns of covariation apart from national socio-economic and political attributes. The analysis might also show that over time these two patterns are interlocked, such that a knowledge of domestic instability could enable a forecast of foreign conflict behavior.

2. Unfortunately, systematic techniques for analyzing variation in data, such as spectral analysis, regression (or time series) analysis, and P-factor analysis are geared to a two-mode *matrix*, or frame. Three typical two-mode frames are shown in Figure VII-3. These frames are sliced out of the three-mode

¹The term "dimension" would normally be preferable to "mode." Unfortunately dimension within the context of the overall discussion also refers to the factors delineated by a factor analysis. To avoid confusion between dimensions of a data cube and factor-dimensions, therefore, "mode" is used for the former concept.

data cube shown in Figure 2. Thus, each frame describes variation along two-modes while holding constant the third mode, such as the case-series frame describing variation of nations across time periods for a specific attribute. For example, the frame may describe the variation of wars for nations over the years 1945-1969. The specific attribute in this case would be number of wars.

Linear algebra and multivariate statistics have developed a number of ways by which variation in a two-mode frame can be analyzed and reduced to patterns. However, such analyses would be on the two-mode frame, and therefore can only be accomplished by holding constant the third mode. Variation in the cube along the three modes is thus partitioned and analyzed as though one were to determine human growth patterns by systematically analyzing only 18 years olds, analyzing the growth of only one baby, or analyzing only the growth of the big toe over the years for a large sample of children.

3. Realizing that the problem to be pointed out is general to the other frames, as well as to dyads and their behavior over time, consider only the attribute series frame, for the moment. Most methods for analyzing change or trends would take one or two attributes for a nation, say GNP per capita and number of wars, and plot their curves over time. One can then fit a time function for one attribute with regard to time, say $\text{wars} = a + b_1t + b_2t^2$. . . , where t = time period, or fit one of the attribute curves to the other over time, such as $\text{wars} = a + b \log(\text{GNP per capita})$.

For our purposes, two problems exist with this traditional approach. First, the number of nations in the international system, and thus the number of attribute series frames, ranges from around 70 to 110 over the period 1950 to 1965. Then, within each frame there are at least 100 socio-economic and political attributes to be analyzed. Thus, the number of functions relating attributes to time would be around 100×110 , or 11,000. This would present little difficulty for the computer of course, but *given that we want to forecast overall socio-economic and political changes*, and not the change on just one variable, these functions leave us with a formidable problem of evaluation.

This latter point brings us to the second problem with the traditional approach. It concerns itself with the changes in one or two specific variables. The variables, such as national income and savings in the U.S., are considered of intrinsic interest, the

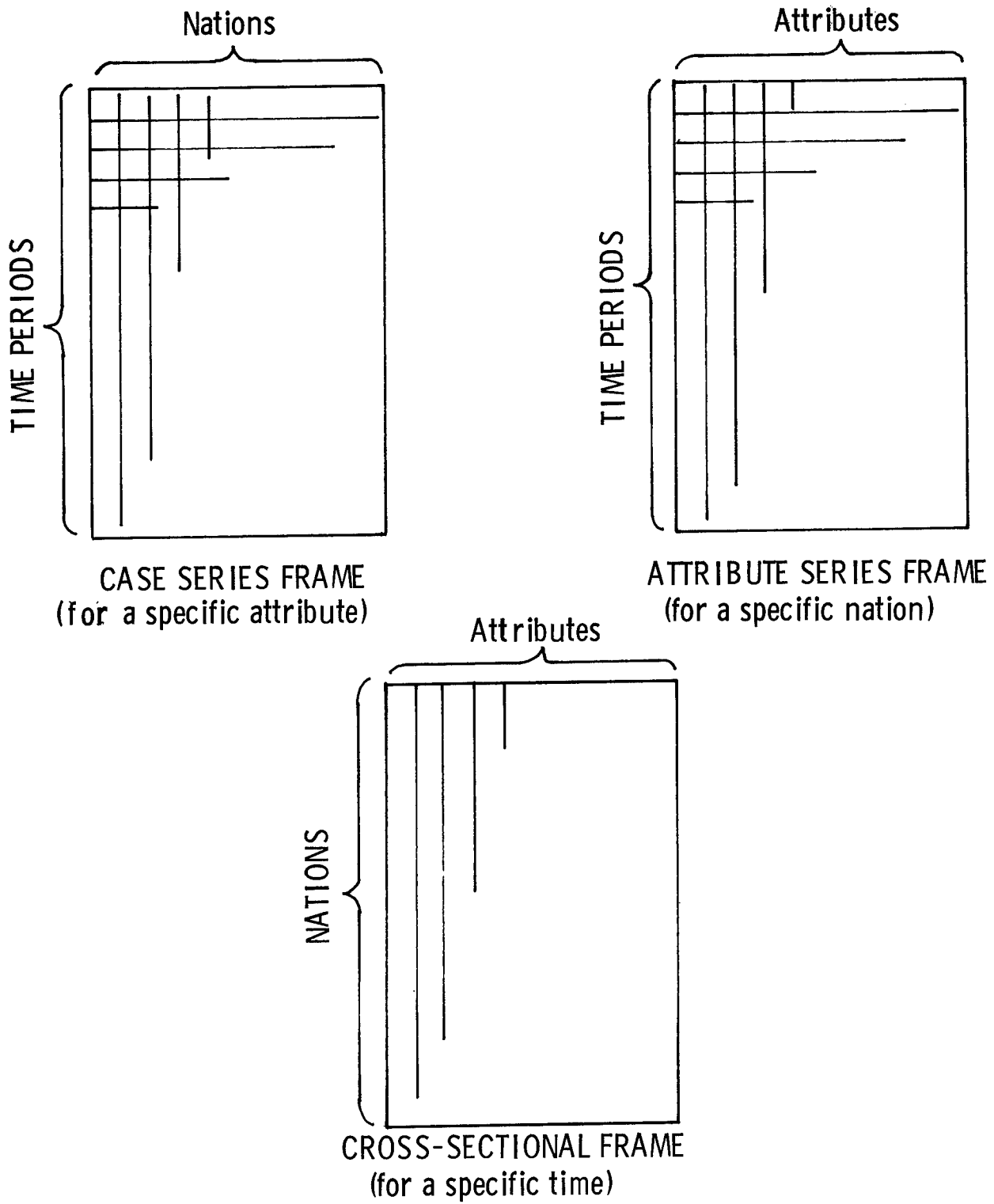


FIGURE VII-3.

measurements are metric (i.e., the numbers can be added and multiplied meaningfully), and are considered reliable. For the range of data we have available on nations, however, the variables (attributes in the attribute series frame) are not of intrinsic interest, but are important for the pattern they contribute to, are often measured on nominal (presence of absence) or ordinal (more or less than) scales, and are contaminated with random and systematic error. Let us consider these three aspects of our attributes in more detail.

In current work on international relations, of which the DON Project is a part, the primary interest is in mapping the major patterns or dimensions, of variation of nations. *A particular attribute, therefore, is significant only insofar as it contributes to defining a pattern of nations.* For example, conflict among nations is of great theoretical and policy significance. Threats, boycotts, severances of diplomatic relations, border clashes, and so forth, cannot themselves be taken as indicators defining conflict for they may represent only one aspect, one dimension, or one type of conflict behavior. Therefore, a range of conflict behaviors is analyzed along with other attributes of nations to determine whether a distinct pattern of conflict emerges and how specific behaviors relate to it.

A second example has to do with national political systems. Such systems are complex and involve many characteristics, no one of which will uniquely define a political dimension. As with conflict, a range of attributes, such as degree of censorship, number of political parties, freedom of group opposition, electoral procedures, independence of the courts, and so forth, has to be analyzed across nations to delineate a political system pattern.

The last example can also point up how our data are often measured. Censorship is usually measured on an ordinal scale: 0 = relatively no censorship, 1 = some censorship of the domestic press, 2 = considerable censorship of the domestic press and foreign dispatches, 3 = complete censorship. As cases in point, the U.S. would be coded 0, France 1, and China 3. Independence of the courts might be given a binary measurement: 0 = independent of political control; 1 = not independent. Number of political parties, however, would be an attribute measured metrically (on a ratio scale). A data cube for nations, as in Figure 2, is such a mixture of metric and non-metric data. Unfortunately, many standard techniques of time series analysis are not applicable to non-metric data.

As far as reliability of data are concerned, data

on nations are notorious for their unreliability and noncomparability. To assume that a figure of 16 riots for India, in 1956 collected from the *New York Times* is correct, or even close, is imprudent, to say the least. To assume that China really has a population of 646.5 million, as reported in the 1961 UN *Statistical Yearbook*, is probably to be incorrect by 10 to 20 percent. To assume that an ordinal measurement of censorship is unbiased, where values (as described above) are judgmentally assigned by the data collector, is to assume more knowledge and rationality on the part of investigators than any of us have. And to assume that GNP figures converted to common currency by exchange rates are directly comparable across nations is to be unaware of the ingredients going into GNP figures, the problems with shifting and multiple exchange rates, and the difficulties of assigning GNP to state controlled economies.

Thus, the signal to noise ratio for our international data is not good. Yet, most standard time series techniques would fit functions to the data as though they were all signal and no noise. In part, because of this high noise in international relations data, we have concentrated on methods that dampen the noise and bring out the signal—methods that deal with the patterns in the data rather than the precise data themselves.

Considering then that we deal with a range of attributes or behaviors, each one of interest only insofar as it contributes to defining an empirically distinct cluster, that our data are often at the lowest level of measurement not amenable to standard time series analysis, and that our data are contaminated with much error, we are forced into developing approaches that best delineate the reliable patterns—dimensions—of variation of nations. The next section will consider two approaches now being employed for this problem, and the subsequent section will consider TMFA as a particularly promising solution.

TRANSFORMATION AND P-FACTOR ANALYSIS

3.1 *Linear Transformation Analysis.* Given that we are concerned with patterns, rather than a set of attributes themselves, and want to track these patterns over time, we can use a modification of Ahmmavaara's (1954, 1957) transformation analysis. This approach has been programmed by DON and experimentally applied to several sets of data.

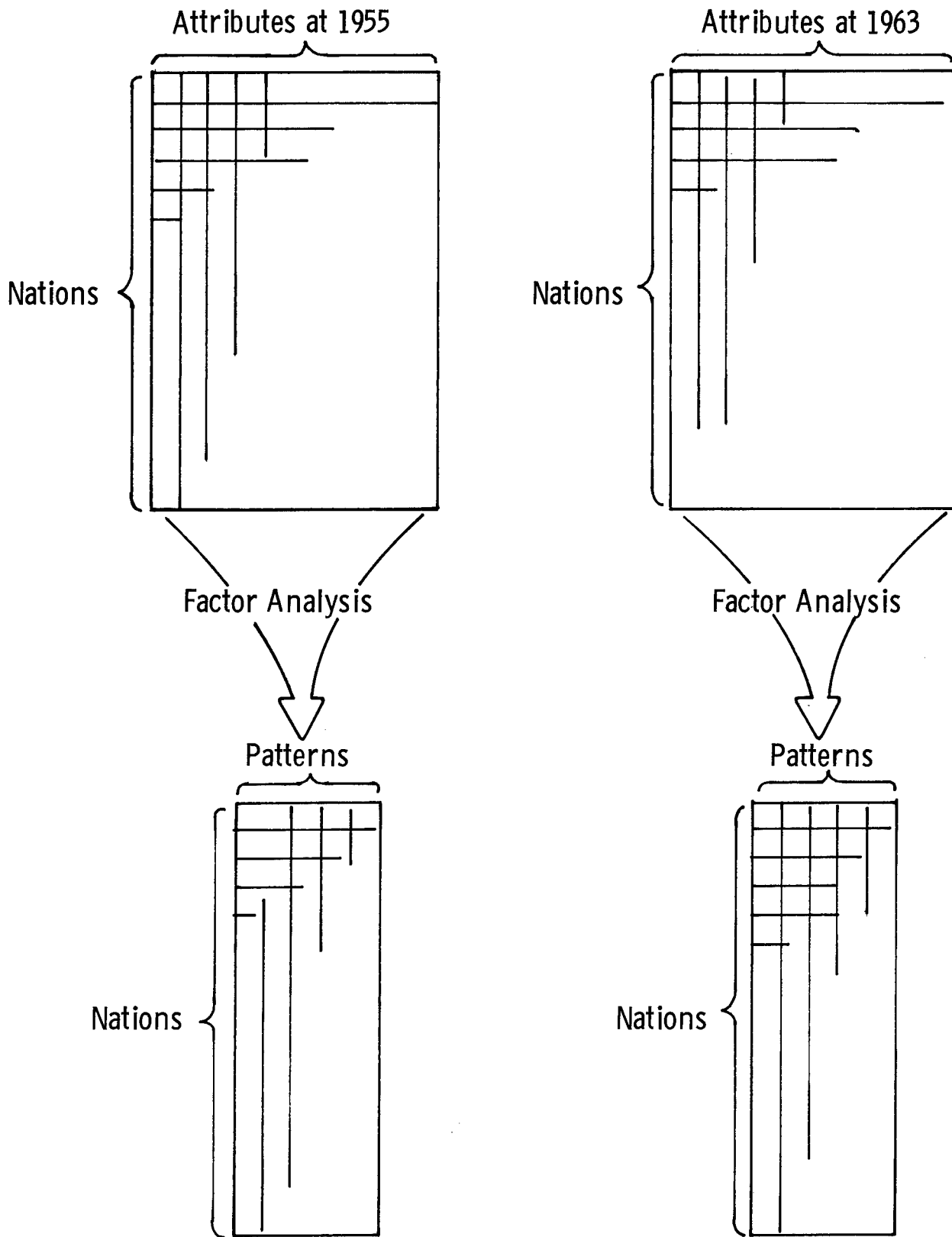


FIGURE VII-4.

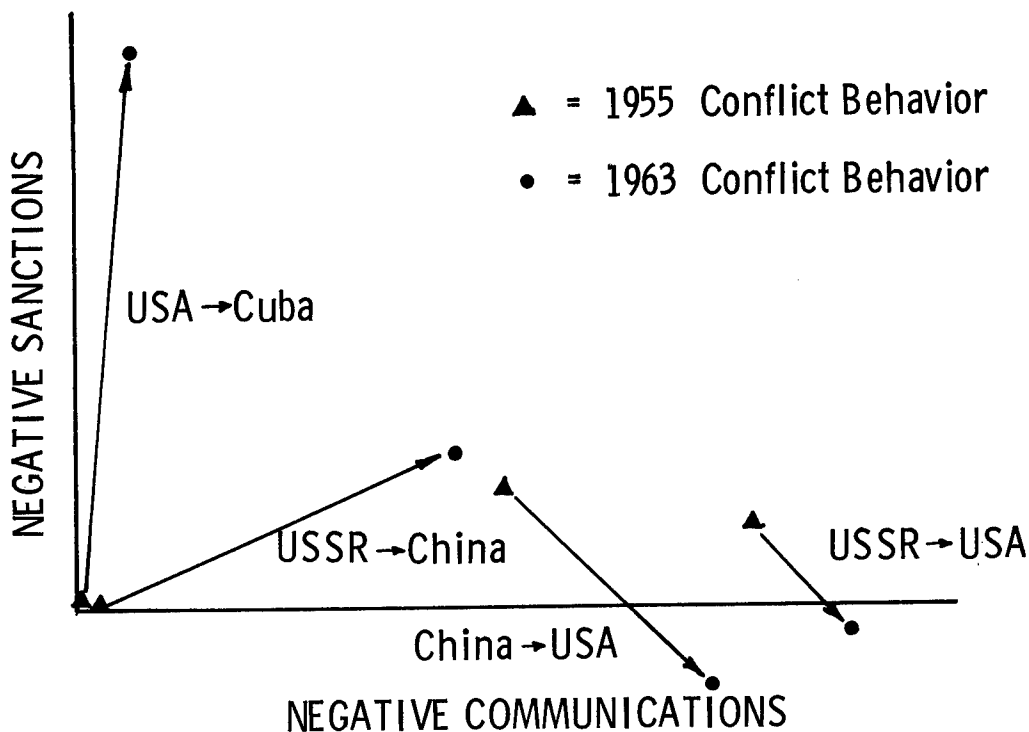


FIGURE VII-5.

To carry out a linear transformation analysis on the data cube of Figure 2, we partition the cube into successive cross-sectional frames of the type shown in Figure VII-3, each frame consisting of a discrete time period. If our time periods defining one mode of the cube are 1950, 1951, 1952, . . . , 1965, then we would have a cross-sectional frame for 1950, 1951, etc.

To each of these frames separately, we can apply factor analysis² to delineate the patterns in the data. Thus, for 1950 we might find economic development and foreign conflict patterns. When we analyze 1951, we might find the same patterns, as well as a third, say internal conflict. And for each successive year, we might similarly find the same as well as additional patterns. For each year, we can also compute a (factor) score of each nation on the patterns defined. These scores now give us a matrix of patterns for each time period, as exemplified in Figure VII-4 for just two years.

One difficulty with doing a separate factor analysis for each time period is that the average of the data will shift between the years, but it is the average that defines the origin of the space within

which the patterns for a year are delineated. A second difficulty is that the factor (rotated) solution is specific to a year. Each year may therefore have different patterns delineated that are only different within a linear transformation of each other. For example, if x is a pattern defined for one time period and y is a pattern defined for a subsequent time period and if $y = ax$, where a is a constant, then for our interests y does not represent a meaningful change in pattern over time from x .³ What we wish to define as change is $y - \hat{y}$ where \hat{y} is the best linear fit to y . Thus, the change in pattern $y - \hat{y}$ is change with regard to x after the linear relationship between x and y has been removed.

Because of the two difficulties—shifting averages across time and linear relationships between patterns for different years—we have to alter our patterns defined for each time period by (1) translating the pattern scores for each period to a common

² For a conceptual discussion of factor analysis, see Rummel (1967). For a technical elaboration, see Harman (1967).

³ The reason such change is considered trivial is because of the method employed to define patterns. For any one set of patterns defined for a time period by factor analysis, an infinite number of alternative patterns could be defined for the same data that are all within linear transformations of each other. Thus, to insure that different patterns for different years do indeed represent changes in underlying data, we first have to remove the linear relationships between the patterns.

origin, (2) doing a least squares fit of the (translated) patterns for each period to those (translated) for the first period. Change in translated patterns and scores is then measured as the deviation of each period from the least squares estimate for that period based on the first period.

Before giving the equations for this, it may be helpful to have an empirical example. In the manner of Figure VII-4, but analyzing dyads (e.g., US → USSR, China → Cuba, etc.) in place of nations, and conflict behavior (e.g., threats, warnings, boycotts, etc.) for 1955 and 1963 in place of attributes, factor analysis delineated five patterns of dyadic conflict behavior for each year. These patterns are intense military violence, incidence of violence, negative sanctions, negative communications, and unofficial violence (e.g., attacks on embassies). Given these patterns, a question is then: how did particular dyads change in their behavior on these patterns from 1955 to 1963? After translating the patterns for each year to a common origin and doing a least squares fit of the 1963 translated patterns to those translated for 1955, we have the result plotted in Figure VII-5 for four selected dyads on two of the five patterns.⁴ The coordinants are standardized and need not be shown here. As can be seen from the Figure, the linear transformation approach to studying change yields results that are intuitively satisfying. They are in accord with our knowledge of the contrast in behavior of these two dyads in 1955 and 1963. Were we now to similarly transform conflict data for 1958, 1960, 1965, and 1968, we would then have six points on the plot for each dyad and could thus track, across each of these years, their shifting conflict behavior.

A difficulty with the linear transformation approach, however, is that the patterns themselves may shift or disappear in time. While the plot of nations, or dyads, over time using this approach will be correct in displaying changes in their attributes or behavior, the posing of similar patterns for each year as coordinates may be misleading. While these patterns may exist for one or several time periods, for other periods they may have changed into other patterns or disappeared. Thus, a method is desirable that will establish patterns stable across time and against which can be cast such plots as in Figure VII-5.

⁴ The particular dynamic analysis from which this example was taken involved about 90 dyads. For simplicity, the plots for only four of these dyads on two patterns are shown. The full analysis is given in Hall and Rummel (1969).

Before going on to super P-factor analysis as a solution to this problem the mathematics of linear transformation analysis might be presented. Let F_1 be a matrix of pattern scores for period 1, and F_2 a pattern score matrix for period 2. Also, let us want to plot the change in pattern scores from period 1 to period 2. First, assume that F_1 and F_2 have been translated to a common origin.⁵

Now, we wish to determine the linear relationship between F_1 and F_2 . This is given by

$$(1) \quad F_2 = F_1 B + E$$

where B = a matrix of transformation coefficients
 E = errors of linear fit.

Then,

$$(2) \quad F_1' F_2 = F_1' F_1 B + F_1' E$$

Assume F_1 and the errors E are uncorrelated, in which case $F_1' E = 0$, where 0 is a null matrix. Assume also that the matrix $F_1' F_1$ is non-singular,⁶ in which case

$$(3) \quad B = (F_1' F_1)^{-1} F_1' F_2$$

The linear transformation of F_1 to F_2 is then,

$$(4) \quad \hat{F}_2 = F_1 B,$$

where \hat{F}_2 = least squares estimate of F_2 .
 The corresponding columns of F_2 and \hat{F}_2 is a pattern axis on which a nation or dyad (row) can be plotted in terms of its two values. If f_{ij} stands for the pattern score of nation or dyad i (the i^{th} row) on pattern j (the j^{th} column) of matrix F_2 , and \hat{f}_{ij} is the corresponding value in matrix \hat{F}_2 , then f_{ij} locates nation or dyad i on the j coordinate axis for period 2, and \hat{f}_{ij} locates nation or dyad i on the same axis for period 1.

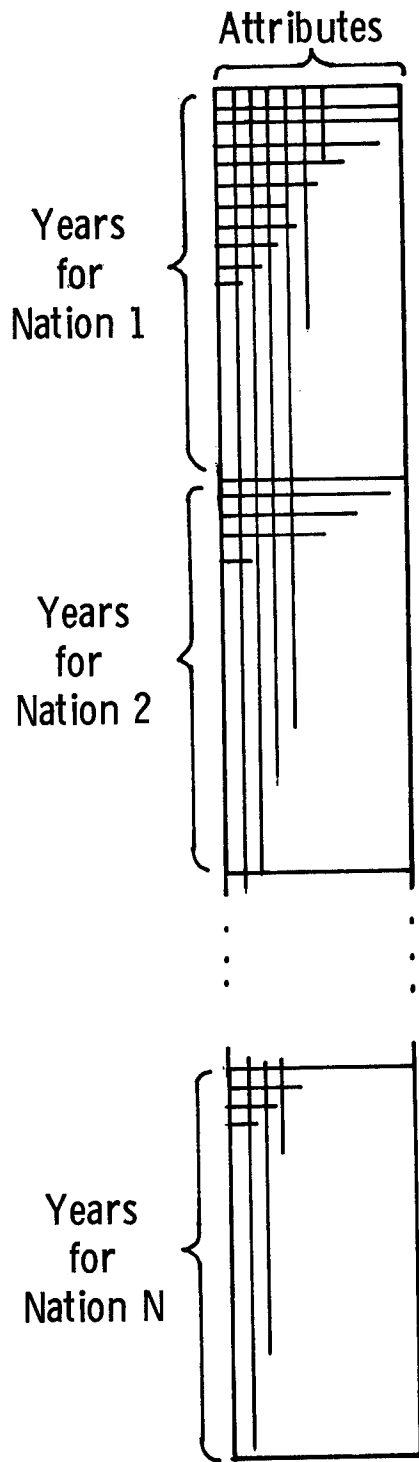
The shift, S , in pattern scores between the two periods is then

$$(5) \quad S = F_2 - \hat{F}_2.$$

S now measures behavior in period 2 that is linearly

⁵ In the dyadic conflict analysis, the results of which are shown in part in Figure 6, "peace," a dyad with all zero values on the conflict behavior variables, was included in the factor analyses for each year. The resulting pattern scores were then all translated to a common origin by using the peace dyad's scores. Thus, if the peace dyad had a pattern score of $-.43$ on pattern 1 for 1955, then $+.43$ was added to all pattern scores on pattern 1. Translating all patterns in this manner shifted the origin for the conflict space to non-conflict behavior for each year.

⁶ Since F_1 is a matrix of orthogonal pattern scores derived from a factor analysis, $F_1' F_1$ will always be nonsingular.



N = Number of Nations

FIGURE VII-6.

independent and uncorrelated with behavior in period 1.

3.2 *Super P- Factor Analysis.* An alternative to linear transformation analysis is super P-factor analysis. This approach gets us close to TMFA and indeed, the results of P-factor analysis are also a portion of the results of TMFA.

P-factor analysis involves determining the patterns of attributes or behavior over time. For simplicity, consider just a nation by attribute by year data cube. The data matrix analyzed is an attribute-series frame as shown in Figure VII-3—a horizontal frame cut from the data cube of Figure VII-2. P-factor analysis can be done on this frame above, if the concern is with time patterns of attributes for a specific nation.

Our interest, however, is time patterns for all nations. Therefore, we can cut a frame from the data cube for each nation, and then “freight-car” these frames one behind the other as in Figure VII-6. This produces a supermatrix, with columns equal in number to the attributes in the data cube and rows equal in number to the product of nations times the number of years.

The supermatrix can be factor analyzed (and since the P-factor analysis is applied to a supermatrix, it is called super P-factor analysis) to delineate the patterns of interrelationships between the attributes over both nations and years. A pattern (factor) score matrix can then be computed which will give a score for each nation for each year on these patterns. These scores can then be used to plot the change in behavior of nations over the years analyzed.

The DON Project is now experimenting with super P-factor analysis to determine how well it works in practice. A dissertation (Phillips, 1968) involving super P-factor analysis is being done on the conflict behavior of dyads by month for 1963 on some twenty variables (e.g., threats, warnings, clashes, expulsions, etc.). The dyadic data is being ordered by month and by dyad similar to Figure VII-6, where the columns for these data are conflict behaviors. The completion of the analysis of these data should give us some experience to gauge the value of super P-factor analysis as an alternative to linear transformation analysis.

The equations of factor analysis are well known⁷ and need not be presented here. An hypothetical example of what the results might look like, however, may be helpful in appreciating the possible value

of the approach. Assume that we have super P-factor analyzed the international behavior of dyads on such variables as trade, threats, treaties, immigrants, mail, and wars, for the years 1945 to 1965. Assume also that two of the behavior patterns found are transactions and conflict. The resulting pattern scores for each dyad on these two patterns should enable us to construct such an hypothetical plot as Figure VII-7 for the four dyads: USSR → China, USSR → US, France → U.S., China → U.S. The value of such time trends is that we could fit a function to them which would enable us to project these trends into the future—to forecast future behavior—and we could relate these behavioral trends to trends in attributes for these particular nations. In this way, we could begin forecasting behavior from basic indicators. A change in such an indicator could then forecast that particular behavioral changes will follow.

THREE-MODE FACTOR ANALYSIS

Three-mode factor analysis, or TMFA, has been largely the development of Ledyard Tucker (1963, 1964, 1965a) and examples of its substantive application have largely come from Tucker's projects ((1965b) Mills and Tucker, 1965). A discussion of the mathematical structure of TMFA appears in Tucker (1965a) and Levine (1963). TMFA has gone through stages of development representing clarification of notation and the development of alternative models that satisfied practical constraints. The model to be discussed here is Model III (Rummel, 1968).

TMFA begins with three super matrices cut from the data cube of Figure 2. One super matrix represents a stacking of attribute-series frames in the manner of Figure VII-6. A second supermatrix is of time-series (see Figure VII-3) frames stacked on top of each other; the third supermatrix is a similar stacking of cross-sectional frames. These matrices are then reduced to four pattern matrices. One gives the patterns in the interrelationships of attributes (again for simplification, only a nation by attribute by year data cube will be considered) across nations and years. These patterns are the same as those found by super P-factor analysis as discussed in the last section, and need not be further discussed here.

A second pattern matrix defines patterns (clusters) of similar nations across attributes and years. These patterns group nations such that the members of

⁷ See, for example, Harman (1967, Chapter 2).

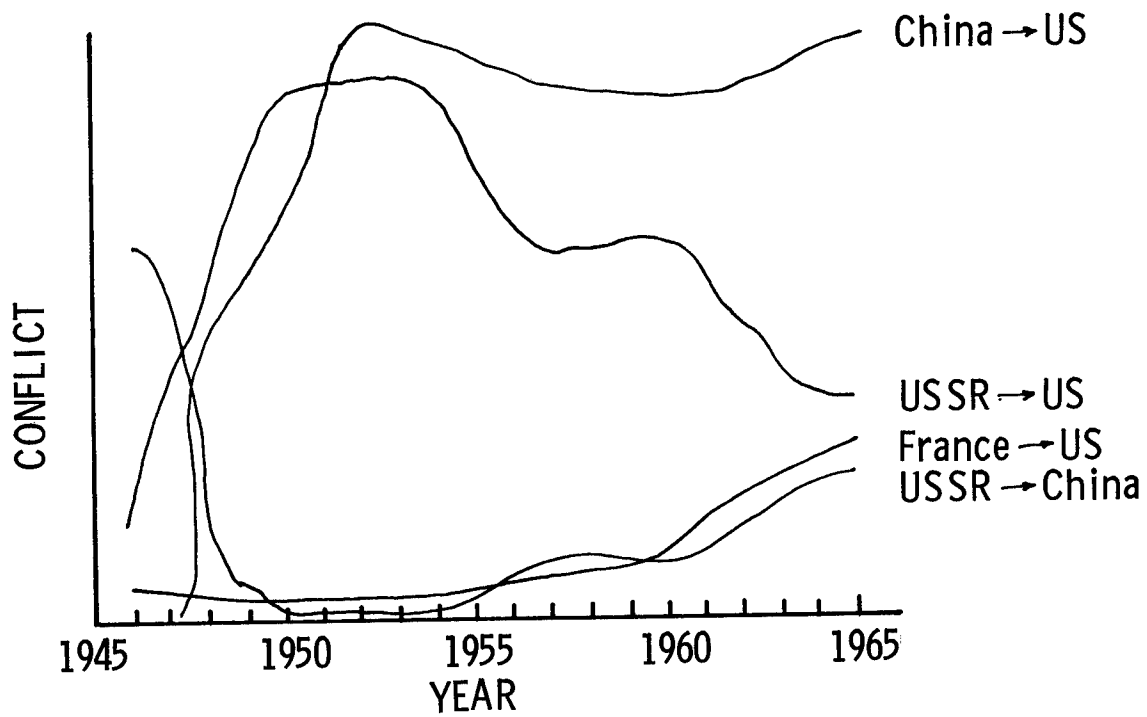
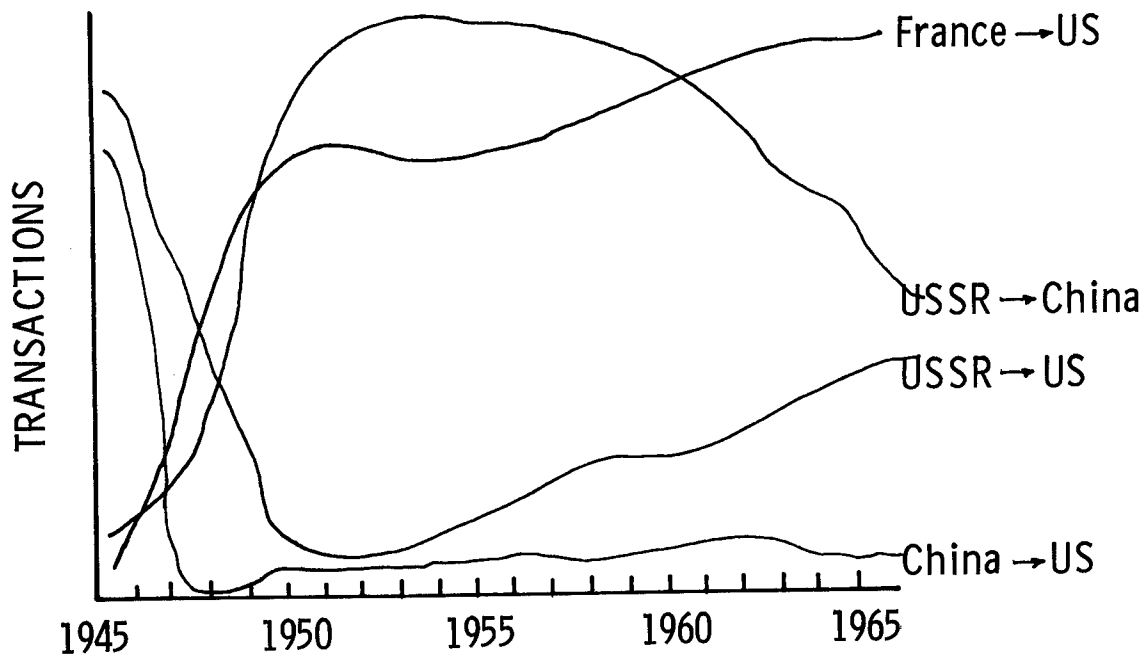


FIGURE VII-7.

<u>Nation Patterns</u>	<u>Attribute Patterns</u>	Time Patterns**		
		Growth	Foreign Crises	Foreign Conflict
Group I	Econ. Dev. Political Orient. * Domestic Conflict	H		
		H		
			H	
Group II	Econ. Dev. Political Orient. * Domestic Conflict	L		
		L		H
				H
Group III	Econ. Dev. Political Orient. * Domestic Conflict	H		
			H	
			M	H

* High values are in the non-totalitarian direction

** H = high relationship

L = some relationship

M = moderate relationship

blanks = no relationship

FIGURE VII-8.

a particular group have similar profiles on the attributes for each on the years. Thus, we might find a group involving Mexico, Ecuador, Honduras, and Peru, for example, who on their attributes are similarly peaceful, poor, Catholic nations and have manifested little change in these attributes over time. As a second example, we might also find the U.S. and U.S.S.R. forming a two-nation pattern, with similar conflict behavior, economic development, size, and participation in international relations over the years.

The third matrix defines the time patterns or trends among nations in their attributes. The fourth, and final matrix, also gives the time patterns of change in nations over attributes, but in addition the matrix shows how these time trends relate to

the nation and attribute patterns described by the other three matrices. The fourth pattern matrix is called the *core matrix*. At this point, an hypothetical example may help.

Let us say that we have applied TMFA to a range of political and socio-economic attributes of all nations for the years 1945 to 1965. The hypothetical attribute patterns to emerge are economic development, political orientation, and domestic conflict. Among the hypothetical nation patterns or groups, let us posit: Group I = Netherlands, Belgium, and West Germany; Group II = Albania, Yemen, Outer Mongolia, and Bulgaria; and Group III = USA, USSR, and U.K. Let us also assume that the time patterns in the core matrix are three: a growth trend, foreign crisis behavior, and a foreign conflict

trend. The core matrix can then relate these three sets of patterns as shown in Figure VII-8.

Now, the core matrix of Figure VII-8 tells us what nation groups had what attribute patterns correlated with what time trends. It neatly partitions our original data cube into subsamples that give us the maximum ability to forecast. For example, we would know, were Figure 9 the presentation of actual results⁸ that we could forecast that members of Group III will continue to see a growth over time in their economic development and political orientation (towards an increasing nontotalitarian orientation). Moreover, the crises behavior of Group III is related to their change in political orientation, while for Group I it is not. For members of Group II, however, their foreign conflict pattern over time is tied in with change in political orientation. We could go on in this manner, interpreting the meaning of the core matrix for our data and making forecasts on this basis.

Neither linear transformation analysis nor super P-factor analysis gives us the information about patterns over time that TMFA displays. In this sense, both are incomplete or partial solutions to an analysis of a data cube. An important, perhaps crucial, aspect of this information yielded by TMFA is the division of nations into groups, which are then interrelated with time and attribute trends. The delineation of homogeneous groups within which forecasts can take place is a basic aspect of science. In fact, the most important result of TMFA may be in this definition of nation groups or *types*.

The virtue of classifying nations by type is that it enables parsimonious description of nations and reliable forecasts about nations based on their type. For example, by being able to type a strange object as a fish, you can describe aspects of the object perhaps not immediately observable, such as being, cold-blooded, having gills, and a two-chambered heart. Moreover, you can make predictions about the object you have classified as a fish, such that if removed from water it will shortly die. If you can classify another object as a dog, you expect (predict) it to bark if surprised, growl if threatened, and wag its tail if happy. Moreover, you can anticipate that

it will chew on bones and also on your shoes if allowed to.

By establishing types, we simplify reality and increase our ability to explain phenomenon. Our types are boundaries defining different kinds of causal relationships. The variance discriminating between types (the characteristics discriminating between cats and dogs) is different than that discriminating subtypes. The variance relevant to distinguishing German shepherds from cocker spaniels is different than that distinguishing both as dogs from cats or monkeys.

Students of comparative and international relations have always dealt with nation types. One type that has played a dominant role in theoretical and applied international relations is that of the powerful nation. This type has become so widely recognized as implying set characteristics and international behavior that we readily employ the noun "powers" alone to refer to nations of this kind. Such nation "types" as modern, underdeveloped, constitutional, status quo nations, prismatic, aggressive, traditional, and nationalistic, have only to be mentioned to evidence the prevalence of typical distinctions.

The problem with the prevailing types is that the rationale underlying the categorization is not explicit (and that it is not clear whether the type really divides different kinds of variance). If we are to deal in types, a clear and empirical basis for the distinctions must be made. In taxonomic work in biology, where typical distinctions have been built in a non-systematic (non-quantitative) fashion, it is becoming increasingly clear that a more systematic basis is required for scientifically meaningful classification (Sokal and Sneath, 1963).

It is easy to assert that members of a type are more similar than non-members, but when one tries to make explicit the notion of similarity, a host of problems arise. Certainly, objects can be similar on some things and not similar on others; they can be similar on several dimensions, and dissimilar on others. Which are the relevant similarities to consider and dissimilarities to ignore? Indeed, how do we measure similarity at all?

TMFA gives us a precise way of defining types and an explicit answer to the question often asked about types: "So what?" The types are composed of those nations that have similar profiles on their attributes over time. The so what question is answered by the relationship of each type and its attribute patterns to the time trends. Then, so what if we find a group including the U.S. and USSR?

⁸ The actual results will give us weighting coefficients in the cells of the core matrix that show the degree of relationship between group-attribute patterns and time patterns. For clarity here, only the rough designations of high, medium, and low are shown.

⁹ This and the following paragraphs are from Rummel (1969b, Chapter 12).

According to our hypothetical Figure 9, the answer is that we now know that this type manifests a long run economic growth pattern, a foreign conflict pattern, and a crises pattern associated with its political orientation.

Although the interpretation of TMFA can be relatively simply put, the mathematics involved are complex and require considerable introductory notation and explanation prior to the derivations. The mathematical solution itself first involves computing the factors of the stacked attribute matrix of Figure VU-7. Then, the correlation matrices for the nation matrix and the time period matrix, each stacked like the attribute matrix, are computed. The eigenvalues and eigenvectors of these two correlation matrices are then computed. The core matrix is given as a product of the factor matrix for the attributes and the Kronecker product of the eigenvector matrices derived from the nation and time period matrices. Those interested in the derivations involved in these procedures should consult Tucker (1965a).

TMFA has only recently been developed in psychology and has seen less than a half-dozen applications. Its promise is great, however, and like factor analysis itself, may prove to be an important tool for describing regularities in the behavior of nations. Before applying TMFA, however, there should be considerable experimentation with the approach so that we understand precisely what is it that we are getting out of the computer and whether the model is really useful (rather than only apparently useful in the sense that TMFA generates matrices of numbers, which it would do even if applied to entirely random numbers) with regard to the difficulties we have in analyzing nations.

TMFA needs to be confronted with the particular difficulties we have in analyzing nation, such as poor data, which have been spelled out in Section 2, before we can accord much credibility to its results. Moreover, before we can have much confidence in our interpretation of these results we need to run three kinds of experiments.

First, we should analyze behavioral data on nations already analyzed using linear transformation analysis and super P-factor analysis. In this way, we can contrast TMFA results with those given on the same data by better known and understood methods. This we plan to do using DON data (already collected) on conflict behavior between nations 1962-1968. These data will be divided by month so that we can analyze the trends in dyadic conflict behavior through time and the particular

time trends of each dyad, such as US → USSR and USSR → China.

Secondly, we plan to test our ability to forecast behavior from the time trends defined by TMFA. We are after all considering TMFA as a solution to our forecasting problems (given our data) and we should have an acid test. Our approach will be to take the by-month conflict data 1962-1968 analyzed above and to partition them into by-month data 1962-1965 and 1966-1968. Then, we plan to apply TMFA to the 1962-1965 data to define the dyadic conflict time trends. Given these time trends, we will then forecast conflict behavior between nations for the months 1966-1968. We can then compare our forecasts with what actually occurred. Moreover, since the whole period 1962-1968 will have been previously analyzed, we will have a background knowledge of 1966-1968 trends as a continuation of those before. Against this knowledge we can view our ability or inability to actually forecast conflict during this period.

Finally, to understand and have confidence in a method applied to a new area, it is helpful to experiment with the method on known data with known relationships. This approach has been used with great success in understanding factor analysis. For example, artificial experiments on factor analysis have involved the dynamics of balls (Cattell and Dickman, 1962), the dimensions of cups of coffee (Cattell and Sullivan, 1962), the dimensions of boxes (Thurstone, 1947), and the dimensions of cylinders (Thurstone, 1947). In a similar vein, we plan to construct three or four such experiments on TMFA. So that we can compare our results with those experiments involving factor analysis, we plan to select artificial data already involved in these experiments, but with the time mode added in.

These three test of TMFA—comparison with alternative methods, forecastability, and artificial experiments—should give us a sound basis for evaluating TMFA. We will then have some ability to gauge how valuable it might be for generally forecasting national and international environments and behavior.

REFERENCES

1. Ahmavaara, Yrjo, 1954, "Transformation Analysis of Factorial Data," *Annales Academiæ Scientiarum Fennicæ*, Ser. B, Tom 88, 2 Helsinki.
2. ———, 1957, "On the Unified Factor Theory of Mind," *Annales Academiæ Scientiarum Fennicæ*, Ser. B, Tom 106, Helsinki.

3. Cattell, Raymond B. and K. Dickman, 1962, "A Dynamic Model of Physical Influences Demonstrating the Necessity of Oblique Simple Structure," *Psychological Bulletin*, 59, 389-400.
4. ——— and William Sullivan, 1962, "The Scientific Nature of Factors: A Demonstration by Cups of Coffee," *Behavioral Science*, 7, 184-93.
5. Hall, Dennis, 1968, "Computer Program Profile," Dimensionality of Nations Research Report No. 14.
6. ——— and R. J. Rummel, 1968, "The Dynamics of Dyadic Foreign Conflict Behavior 1955-1963," Dimensionality of Nations Research Report No. 16.
7. Harman, Harry, 1967, *Modern Factor Analysis*, 2nd Edition, Chicago: University of Chicago Press.
8. Levin, Joseph, 1963, *Three-Mode Factor Analysis*, Ph.D. dissertation, University of Chicago.
9. Mills, David H. and L. R. Tucker, 1965, *A Three-Mode Factor Analysis of Clinical Judgement of Schizophrenicity*, ONR (Grant Nonr 1834(39)) Technical Report.
10. Phillips, Warren, 1968, "Dynamic Patterns of International Conflict: A Dyadic Research Design," Dimensionality of Nations Project, Research Report No. 17.
11. Rummel, R. J., 1967, "Understanding Factor Analysis," 1967, *The Journal of Conflict Resolution*, Vol. XI, No. 4, Dec. 1967, pp. 444-480.
12. ——— 1968a, "Lecture Notes from L. R. Tucker's Workshop in Multivariate Methods," University of Hawaii, Summer 1968.
13. ——— 1968b, "The DON Project: A Five Year Research Program," Research Report No. 9, The Dimensionality of Nations Project, Department of Political Science.
14. ——— 1969a, *Applied Factor Analysis*, Vol. I-II, Evanston: Northwestern University Press, forthcoming.
15. ——— 1969b, "Some Empirical Findings on Nations and Their Behavior," *World Politics*, Vol. XXI, January, pp. 226-241.
16. ——— 1970, *Dimensions of Nations*, forthcoming.
17. Sokal, R. R., and P. H. Sneath, 1963, *Principles of Numerical Taxonomy*, New York: Freeman.
18. Thurstone, L. L., 1947, *Multiple-Factor Analysis*, Chicago: University of Chicago Press.
19. Tucker, L. R., 1963, "Implications of Factor Analysis of Three-Way Matrices for Measurement of Change," in C. W. Harris (ed.) *Problems in Measuring Change*, Madison: University of Wisconsin Press, pp. 122-137.
20. ———, 1964, "The Extension of Factor Analysis to Three-Dimensional Matrices," in N. Frederikson and H. Gulliksen (eds.), *Contributions to Mathematical Psychology*, New York: Holt, Rinehart, and Winston, pp. 109-127.
21. ———, 1965a, *Some Mathematical Notes on Three-Mode Factor Analysis*, ONR (Grant No. 1834 (39)) Technical Report.
22. ———, 1965b, *Three-Mode Factor Analysis of Parker-Fleishman Complex Teaching Behavior Data*, ONR (Grant 46-32-66-357) Technical Report.

DISCUSSION FOLLOWING PAPER GIVEN BY DR. RUDOLPH J. RUMMEL, DIMENSIONALITY OF NATIONS PROJECT, UNIVERSITY OF HAWAII, "FORECASTING INTERNATIONAL BEHAVIOR AND ENVIRONMENTS: A PROPOSED INVESTIGATION OF THREE-MODE FACTOR ANALYSIS"

(The following statement and Dr. Rummel's reply were stated in the middle of the prepared speech.)

MR. FRESHMAN: Taking into consideration the environment of the future is the job of the Air Staff or the Chief of Naval Operations and so on. That is not our bailiwick. It is their bailiwick to do it. It is supposed to be taken into consideration when they make up the operational capability objectives that they wish for 1985. They take into consideration international behavior and environment, the stuff that you are talking about now.

DR. RUMMEL: This I quite realized and I am not making the point that international environment is not taken into account. What I am trying to say is that the manner in which it is taken into account, the scientific basis for such projections, is not the same as with your technological forecasts. You have a heavy investment of science on one side and a

heavy investment of judgmental and intuitive factors on the other. This is the kind of balance of which I speak.

(Discussion following the prepared presentation):
DR. WARD: Have you ever worked with multiple discriminant analysis? I wonder why you would have chosen this method which seems to add a certain amount of complication. I am not sure that you buy something commensurate with added work. Multiple discriminant analysis is a bit more straightforward. I am not sure you would lose too much with it.

DR. RUMMEL: Multiple discriminant analysis is a variety of factor analysis. It is equivalent to carrying out a factor analysis such that the dimensions of two subsets of data have a maximum fit to each

other. Well, we do use this kind of methodology in linking up international behavior to nation attributes.

DR. WARD: I guess what is bothering me is you have gone the empirical orthogonal function route first, which certainly reduces the total number of data you have to work with. It has that attribute, but you end up with orthogonal functions which can be very, very difficult to interpret in terms of anything in the real world.

DR. RUMMEL: Economic development? Size?

DR. WARD: I would like to see what went into economic development. I have a feeling it is a very complicated thing which *you* might label economic development but *I* might not.

DR. RUMMEL: GNP per capita, telephones per capita, vehicles per capita, proportion of agricultural population, literacy.

DR. WARD: Economic development? You do not end up with a mixture of some odds and ends in there?

DR. RUMMEL: Yes, we do, as a matter of fact. Having obviously noneconomic development variables come into this pattern gives us clues as to what variables may be causing or be a consequence of economic development. Incidentally, on this plane, a similar analysis was done on international relations by myself. The variables measured the total behavior of nations, such as total exports, total aid, etc., rather than the behavior of one nation to another. These variables were analyzed to determine the major orthogonal components of behavior, and it was determined how they related to the attributes of nations. A very interesting result came out: if you want to predict the volume of participation of a nation in the international system,—that is total exports, total embassies, total number of international organizations of which it is a member,—you can do so on the basis of the attribute dimensions of economic development and size and with, I would say, about eighty percent reliability.

DR. WARD: I guess what is bugging me is: I cannot believe you are lucky enough to get empirical orthogonal functions out of your analysis, all of which can be named without a considerable amount of bugging, either in the force fitting of certain functions or computational factors. This is really what I was trying to get at. I am not accusing you of illegitimacy. I was just interested in how you got the data to fit so nicely.

DR. RUMMEL: It really does not fit so nicely. We do bugger around in one specific way. Let me explain it. We first determine the principal axes

of our data. Each axis measures the maximum orthogonal variance in the data. These axes, however, are not easily interpretable. The reason for this can be seen quite easily. You can have two uncorrelated variables forming two orthogonal vectors in the space of your data. Now, a principal axis—or orthogonal function as you term it—can fall between the two variable-vectors such that the projections of both on the principal axis are indistinguishable. Because of this possibility, the projections of variables on the principal axes are difficult to interpret. What we do, therefore, is linearly transform our principal axes to new coordinates such that each defines a highly inter-correlated cluster of variables. The new coordinates are still orthogonal functions, but now they become interpretable. Economic development, you see, is really a cluster of variables that co-vary together.

DR. WARD: Why could you not do some simple multiple discriminant analysis, find out which things were correlated and in what way, and then construct your functions? They might not be orthogonal at that point, but I do not see that that is tremendously crucial. These functions would be totally meaningful and you could require them to be the same from one year to the next and hence eliminate certain of your difficulties in trying to make interpretable orthogonal functions from year to year.

DR. RUMMEL: This is certainly an alternative approach and is one that has been applied by a good number of people. The problem occurs in trying to deal with total systems. In order to find conflict functions, I believe you have to deal with the variety of behavior in the international systems, that is, with the major dimensions or patterns of behavior. In order to do this, you have got to begin with a fairly large number of variables covering at least the major behavior of the nations. Given this problem, I am bound into a methodology allowing me to deal with many variables. There is another point here. When you analyze a small number of variables, you do not know whether any of the relationships found between them are due to variables outside of the system. For example, you might find two variables that are quite orthogonal. The reason they are orthogonal, however, is that there is a third variable hiding their relationship. Remove—partial out—the influence of the third variable and the relationship between the two will become evident. In an area like international

relations with a dearth of good theory and empirical work, how are we to know when such a third variable exists? We often cannot. Therefore, we employ a large number of diverse background variables as a backstop against which we assess what bivariate relationships we find.

DR. WARD: But you have a problem in the approach you have taken: you have to assume a certain kind of relationship to generate orthogonal functions to start with. You have to assume that certain things are higher or lower on some scale. In other words you are measuring something. You have to say two is twice as good as one, and who knows where two is?

DR. RUMMEL: You are saying that I assume my data to be measured on a metric scale. But this assumption is not involved and is not required by the mathematics.

DR. WARD: Using this methodology, how can you have empirical construction unless you have numbers?

DR. RUMMEL: The orthogonal functions enable you to reproduce the variance in the original set of measures. You can pick any measures such that there is an ordinal continuum of some sort. This requirement would be satisfied, for example, by a dichotomous variable like sex: male = 1, non male = 0. This is what the economists call a dummy variable.

DR. WARD: But the assumption in itself generates not only an order but a scale in the process.

DR. RUMMEL: That is right, ordinal functions are scales.

DR. WARD: You lose a certain amount on that, you know.

DR. RUMMEL: This is a fascinating question. You have often heard the expression "garbage in, garbage out." Our methodology now enables us to begin with data that are measured in a very loose way, data like censorship measured on a three point scale with 1 = no censorship, 2 = some, and 3 = censorship. This is very loose data for the scientists to work with. We now have a methodology which takes this data and squeezes out the metric information in the data—forms the data into metric scales. It is somewhat amazing.

DR. WARD: I am wondering where this happens. I know you can get it out of multiple discriminant analysis but I cannot see where you got it out of this.

DR. RUMMEL: The secret is dealing with a large number of variables. You can then build up in the space of your data constraints on the location

of your variables. The first variable you locate in the space can be located anywhere. The second variable you locate in this space must be relative to the first in terms of the data. It has some constraint built on it by the first variable's location. A third variable has additional restraints on its location in terms of the first two. Finally by the tenth or twentieth variable, there are absolutely no degrees of freedom left. The last variables must be located in a definite place within that space in spite of originally loose measurement. The dimensions of this space then define all these constraints and provide by the projections of the variables on them a set of metric measurements. These new measurements then become a new set of variables to which very demanding mathematics such as calculus can be applied.

DR. KNAUSENBERGER: I was wondering about this marriage between the linear algebra and what psychologists use. You are using linear algebra and you use some concepts from the analysis in psychology. Is this the quantitative or qualitative evaluation of attributes? How does this combination occur?

DR. RUMMEL: It is a very interesting combination which, except for a few, was ignored in psychology a long time. The psychologist's interest was in finding the space of attitudes, abilities, or intelligence among a sample of individuals in terms of their answers on tests. Now, a method called multi-factor analysis was developed by L. L. Thurstone at the University of Colorado which tried to reduce the matrix of test answers to the major factors influencing the test responses. These factors were originally solved for the principal axes themselves. These axes were then the sought after factors. However, the solution for the principal axes involved computing the eigenvalues and eigenvectors of the data. With the concepts of eigenvector—value functions you merge right into linear algebra. Over this bridge you can draw on the whole paraphernalia of linear algebra. You begin to interpret your tests, as the psychologists now do, as vectors in space. You can conceptualize test data on individuals as a psychological space, an intelligence space, or an attitude space and begin to work theoretically with vector concepts, visually, in a way that those dependent upon symbolic manipulations alone cannot do. As I try to develop a theory relating attributes to the behavior of nations, this ability to move from the methodology to vector spaces and draw on the whole structure of linear algebra is very helpful.

DR. LINSTONE: I wonder how discontinuity is dealt with. I can visualize discontinuities arising from, let's say, new technology in 1945, from charismatic leadership which was introduced in one country and shifted the country's behavior very suddenly from new ideology, and so forth. Cuba, for example, experienced a sudden drop in 1951-53 of the per capita GNP, which Tanter and Midlarsky relate to the Castro revolution which followed. Can you predict such behavior?

DR. RUMMEL: If the discontinuities exist, we can observe them. Then we can try to fit functions to such discontinuities and project them into the future. The question is, of course, whether these are unique discontinuities or whether they really occur regularly in some fashion. If so, we can project them. This becomes then a scientific question, an empirical question.

DR. WARD: Just to follow on this thing: what do you do about lags? Do you test for any of the involved times lags which would bear, I think, on what he is trying to get at.

DR. RUMMEL: The particular theory I am working with assumes that time operates simultaneously. Whatever occurred in the past is assumed bound into the present. Regardless of what is assumed about time, however, the value of this explicit methodology is that if somebody is dissatisfied with my particular approach, they can redo it their way. They can easily lag the computer cards, for example, and run them through the computer again to see what happens. By the way, the question always comes up in the area of international relations: what would happen if. . . ? Well, once you have the basic set of data and methodology, it is possible to experiment with the data to provide answers to such questions. You do not have to completely leave these questions to intuition.

MR. ELLISON: I am generally all for your factor analysis procedures. However, if you are trying to predict criteria (one of the long range goals of such research) and if you have constructed a criterion through factor scores—call it a composite—which you are halfway satisfied with and which you are working with and trying to understand, then would it not be worthwhile at some time to go back in and work with the unique variance which you dropped out through all these procedures? In other words, might you not have very crucial variables which you by and large have lost through factor analysis?

DR. RUMMEL: This is a question we are very concerned about, and you are quite correct. By omit-

ting the unique variance we may well have lost some crucial variance. This is always a problem in any research endeavor, however. What one tries to do is account for as much of the variance as possible in some dependent variable by the major factors that have been defined. Then, as those factors are unable to explain variance in the dependent variable, we try to add back some of the variables that were dropped along the way in delineating the factors. So far we have been able to account for about 25 to 50 percent of the variance in international behavior by factors delineating attribute variation among nations, such as economic development, size, and political orientation. The technique employed to determine this variance has been canonical analysis. Being able to account for 25-50 percent of the variation in behavior would be pretty good for the psychologist.

MR. FRESHMAN: It is not very good at all.

DR. RUMMEL: For psychology it would be good. If you have a natural science background, it appears poor. From the point of view of the policy maker who is interested in making—indeed, often must make—projections that are reliable, one can say, "But 25-50 percent variance explained means 50-75 percent unreliability." Now that I have laid our error on the line, so to speak, I should ask you as a policy maker, "What variance can you account for in international behavior?"

MR. FRESHMAN: I do not think that is the question. I think you are asking me to take a chance and believe in your system when you only give me thirty percent.

DR. RUMMEL: Do you know how much your consultants on international relations are in error?

MR. FRESHMAN: I do not know but I am taking your word for it.

DR. WARD: The thirty percent is a very conservative estimate. If there are errors in the data and we are still generating thirty percent, he may in fact be giving you a forecast that could be as good as sixty, eighty, or even a hundred percent. He does not say the other seventy is unexplained.

MR. FRESHMAN: Let me ask an important question in my mind: are you predicating here that after you got your attributes and were able to glean them from past behavior, you are projecting future behavior based on past?

DR. RUMMEL: Not in terms of these analyses. Building on them, however, we should eventually have this potential.

MR. FRESHMAN: Are you projecting on the basis of any change or predicting change?

DR. RUMMEL: I think it is wise in this area where we are just beginning a science—we do not have it yet and I would not advise any policy maker to base policy on what we have done so far—that you should take these quantitative analyses as only points in the continuing growth of scientific work. Eventually, I have faith, sufficient knowledge upon which to base policy will be accumulated.

MR. FRESHMAN: I am not trying to discourage; I am just trying to understand your rationale. You are taking past periods and projecting future behavior based on that.

DR. RUMMEL: Here we are just trying to determine the relationship between behavior and attributes at a point in time. This tells us nothing about change. Our results only say that for this period, which happens to be 1955, there is a linkage of this sort. Now, this is a beginning: in order to eventually make predictions—forecasts—you have to know what happened at different points in time. While we are now working on only one point in time, we do intend to extend the analyses over a number of time points. We then can begin to develop time functions.

MR. FRESHMAN: You are not at the point where you can do any time projection in the future?

DR. RUMMEL: No.

MR. FRESHMAN: Even if you would be, are you still going to be basing it on past behavior?

DR. RUMMEL: Yes, that is correct. If you would ask me about the level of participation in the international systems of Egypt in 1975, I think I will eventually be able to give you a better answer than could any consultant. Our results are beginning to show a very neat lock-in between participation in the international system and the economic development and size of nations. Using subject area experts to project indicators of economic development and size for Egypt, I think I should then be able to project what Egypt's volume of participation would be in the international system.

MR. FRESHMAN: Leaving out these continuities, would that give me a prediction on the conflicts, which ally they would be based with?

DR. RUMMEL: No, not in terms of what research I am involved in. The question is however, a feasible one to do research on with existing methods.

MR. FRESHMAN: That is very important to us because at the beginning you were saying to us how we were projecting orthogonally without knowledge of things in the future. You are not going to give us anything better.

DR. RUMMEL: Again this is the beginning. What

we need is an army of workers to do research on these various questions. There are some quantitative studies now being done on alliance behavior, however, at Yale and Pennsylvania. They are trying to determine how alliance structures relate to the attributes and behavior of nations. They are trying to get systematic evidence of the dependence of alliances on other kinds of bonds, such as their exports, treaties, mail flows, etc. Here again we need an army to do all the work that is necessary. It is frustrating to have the tools and masses of data from open sources to really answer some of these questions, but not to have the manpower. We need trained students.

LT COL HAIDLER: Are you able to feed in any attitudes of population or things of that nature? How do you measure?

DR. RUMMEL: In one analysis we employed 236 variables to delineate the attribute patterns of nations. Given this mass of data, we tried not to concentrate too much on any one type variable—such as economic—in collecting data. Three of the variables were from David McClelland's work, *The Achieving Society*, and comprised content analysis data on the drive for power, for achievement, and for affiliation. These data were for some forty nations, and were based on analysis of children's readers. Given this data, we then put it into the analysis to see if motivations, which had a lot of variance across the nations of this type would tell us anything about their behavior. Drive toward power, which seems to be high in some nations and low in others: would it tell anything about their participation in international relations. The upshot of including these variables in the analysis along with measures of—talk about unique variance—GNP per capita, population, defense budget, conflicts, etc., was that they came out really unique. That is, McClelland's variables did not tell us much about other attributes of nations or their behavior.

DR. KNAUSENBERGER: You certainly know a book which was published in Germany some years ago: *Formulae Concerning Power* by Fucks*. He uses practically only three dominant factors. The question is here, what are the essential and leading variables out of the possible multitude. Do you have any mechanisms, so to speak, to discern or define these leading variables?

DR. RUMMEL: Since my field is international re-

* Title of Book: Wilhelm Fucks, *Formeln zur Macht, Prognosen über Völker, Wirtschaft, Potentialc.* Deutsche Verlags-Anstalt Stuttgart. 1965.

lations, I had some interest in the intercorrelations between indicators of power. Therefore, a large number of such indicators were included in the 236 variable factor analysis of nation attributes. I selected indicators of power from the dominant textbooks in the field of international relations which suggested such indicators as area, population, national income, energy consumption times population, number of men in the military, defense budget, resource availability, political centralization, and so on. It was fascinating to see how all these variables came out of the analysis. All the indicators of power, really power capability, were highly intercorrelated and formed a dimension quite independent of economic development and political orientation. Because many of the power indicators also measure the size of a nation, such as area and population, the dimension can be called "size" or "power capability," depending on your interest.

DR. KNAUSENBERGER: Fucks says essentially, I think I am right, that the historic situation, the geographic location, the population and economic development are the parameter and his prediction is that there will be a dominance of Red China.

DR. RUMMEL: China is very high on the size or power capability dimension which would be in accord with Fucks's position. I had mentioned that economic development has also been formed as a dimension of nations and that it is independent from power capability. Also, I think I mentioned that participation in international relations can be accounted for by a nation's economic development and size. Consider now that size also measures power potential and you can imagine how meaningful these results are to the student of international relations. It means then that the two most important variables for understanding international relations in terms of volume of participation are economic development and power. From the same analysis giving us these results, we can also say that conflict behavior is independent of the level participation in the international system. That is, the degree of conflict a nation manifests in international relations is not related to its other kinds of behavior. We can predict participation in international relations from economic development and power capability, but we cannot at all predict conflict from these two variables. Since we found that foreign conflict comes out as an independent dimension, the question is: what does predict levels of conflict between nations? The result across a number of attribute

dimensions are consistent in indicating that it is not the level or value of a nation on an attribute, such as amount of economic development, that relates to foreign conflict behavior. The evidence now accumulating is that it is relative levels that count. That is, it is the differences and similarities—the socio-economic and political distances between two nations that appear to well predict their conflict behavior.

MR. FRESHMAN: Did you test whether the ability to predict conflict between two nations rejected economic competition in foreign trade and affected the market?

DR. RUMMEL: Let me go on. The relationship between conflict and socio-economic and political distances has been investigated in the following manner. Let me make this as clear as possible by putting it on the blackboard. . .

MR. FRESHMAN: The reason I asked this is because many people have said the reason for conflict between two nations is basically economics, conflict on the most vital source of materials and the trade.

DR. RUMMEL: So far our results are unpublished on this point. We have found, however, that differences in economic development and power capability account for about sixty-three percent of the variation in one nation's behavior towards another in terms of translations, tourists, treaties and memberships in international organizations.

DR. WARD: To each other?

DR. RUMMEL: To each other.

(An explanation was put on the blackboard by the speaker and is not reproduced here).

DR. RUMMEL: Another finding was that conflict and cooperation are linked together. We often account for more behavior when we deal with a linear combination of conflict and cooperation.

DR. LINSTONE: Does this correspond to what is brought out in "The Intimate Enemy"? Past studies (I believe Richardson, for example) which tried to be quantitative although not too sophisticated, looked at statistics and concluded that neighbors are much more likely to have conflict than non-neighbors. That means that geography is important.

DR. RUMMEL: Our correlation for the number of boundaries of a nation with its conflict behavior has been low, about 0.30. Lewis Fry Richardson, however, found such a correlation to be 0.78. The difference between our two sets of findings may be due to sampling variance, however.

MR. BOBROW: How many variables did you com-

bine into your conflict measure to get your figure? If you just took wars, at some level, what would that do with this particular finding?

DR. RUMMEL: The correlation involved both military violence and negative communications. Military violence is a pattern comprising number killed, clashes, number of discrete military actions, such as shooting down a spy plane, and number of wars. Negative communications involves threats, protests, accusations, and denunciations.

MR. BOBROW: If you were just to look at the first half and leave out the negative communications and if you then try and work with conflict chronologies, dividing it into those where the parties are contiguous and not contiguous, would you still get your point three?

DR. RUMMEL: Sixty percent?

MR. BOBROW: No, your point seven there. It gets closer to the Intimate Enemy hypothesis.

DR. RUMMEL: The answer here is "No." Richardson used "number of killed" as his measure of war for his correlation. In the international system war is the extreme. There is a spectrum of conflict behavior between ignoring another country and finally being involved in matters of violence. What I am trying to do, in contrast, is to tap the whole spectrum of conflict.

MR. ELLISON: Just a point of clarification here: the things you have listed on the left and right and the sixty percent: which of these are factors and what is the sixty percent?

DR. RUMMEL: Sixty percent (making reference to a figure drawn on the blackboard) refers to the canonical correlation squared between the be-

havior of one nation toward another on the one hand and differences in economic development and power capability on the other.

MR. ELLISON: Those are factors?

DR. RUMMEL: The differences, or distances, are measured on the economic development and power capability factors—dimensions. Consequently, we are not bound into an arbitrarily selected indicator like GNP per capita or population. We know that the factors employed in the analysis tap the major variance among dozens of measures of economic development and power capability. Therefore, when I talk about economic development, I am speaking about a whole cluster of measures.

MAJOR THOENY: In the discovery of your pattern of size, you mentioned you searched through international relations textbooks to look at the usual list of factors, but the examples you gave were all quantitative factors. Did you make any attempt to measure the qualitative factors that one finds?

DR. RUMMEL: There were some qualitative variables employed. For example, variables were included that defined whether a nation had a federal or military political organization, whether the political system was centrally organized, whether there was censorship, whether groups were free to organize and oppose the government and whether a nation played an active role in bloc politics.

MAJOR THOENY: Another factor which is frequently quoted is qualitative military measure as well as the metric measure of military power.

DR. RUMMEL: Yes it is.

FORECASTING SOLAR EVENTS

by FREDERICK W. WARD, JR.*

INTRODUCTION

The title of the talk as you see it printed on the program certainly tells the general area; but it misses the main point of my remarks. As I was preparing the talk, a number of sub-titles came to mind which more accurately expressed the philosophy of what we've done. The three most descriptive sub-titles that came to mind were: "Technique Wise and Data Foolish", or "Spare the Data and Spoil the Forecast", or even catchy things like "Better Technique Do Not a Better Forecast Make—Necessarily"! But I guess the one that's a little longer, not quite as catchy, but really describes what I'm going to be talking about, is the one that goes, "The Best Way To Improve Forecasts Is Not Always To Improve Forecasting Techniques". This isn't to say that when we started we didn't plan to put most or all of our effort into improving the forecasting techniques—because we did—that is, we naively assumed that a concentration on the techniques would yield the best results. Most of this paper, however, will be devoted to subjects other than the improvement in forecasting techniques, the seemingly unrelated efforts into which a large fraction of our concentrated effort finally went. It became obvious to us that these efforts would produce a far greater improvement in forecast usefulness and accuracy than any reasonable work on forecasting techniques.

I probably should give you a hint as to what the difficulties are and why we went in the direction we did before I outline the problem. I guess it can best be summed up by saying that the available data both for use in existing forecast techniques and for the purposes of developing new forecasting techniques were totally inadequate—inadequate in almost every way imaginable. They were inconsistent. Measurements were made by a variety of techniques, dependent seemingly on the whims of each observer. They were incomplete to the extent that even some data that were recorded couldn't be used because of other data that weren't recorded. The data were plagued by all kinds of errors, many correctible—but with very great difficulty—many uncorrectible on the basis of the available information. The instruments used to collect the data were badly suited toward this end, and constantly undergoing changes and modifications in the name of so-called "improvements". This catalog could be extended, but these were the main difficulties.

Faced with this situation, and the responsibility for improving the accuracy and usefulness of forecasts of solar events, I believe that one must take a very hard look at the areas for potential improvement. The choices we made were dictated by many pressures; scientific, operational, and human. I hope that this description of our approach may be of some limited value to others faced with a similar situation.

* Dr. Frederick W. Ward, Jr. was born in Cambridge, Massachusetts, 28 February 1930. He attended Cambridge public schools and the Massachusetts Institute of Technology, from which he received his S.B. (1952), S.M. (1955), and Ph.D. (1957) in meteorology. In 1952 he joined the Air Force Cambridge Research Laboratories and worked in the Meteorology Laboratory on problems of the general circulation of the earth's atmosphere. In 1963 he joined the Space Physics Laboratory to work on the general circulation of the sun's atmosphere and space environmental forecasting. He is Chief of the Space Forecasting Branch of that Laboratory.

SOLAR ACTIVITY AND ITS GEOPHYSICAL EFFECTS

But before we get into the space-weather forecasting business, it is necessary to understand the physical picture of what's going on with the sun, between the sun and the earth, and in the upper atmosphere of the earth. It is also necessary to

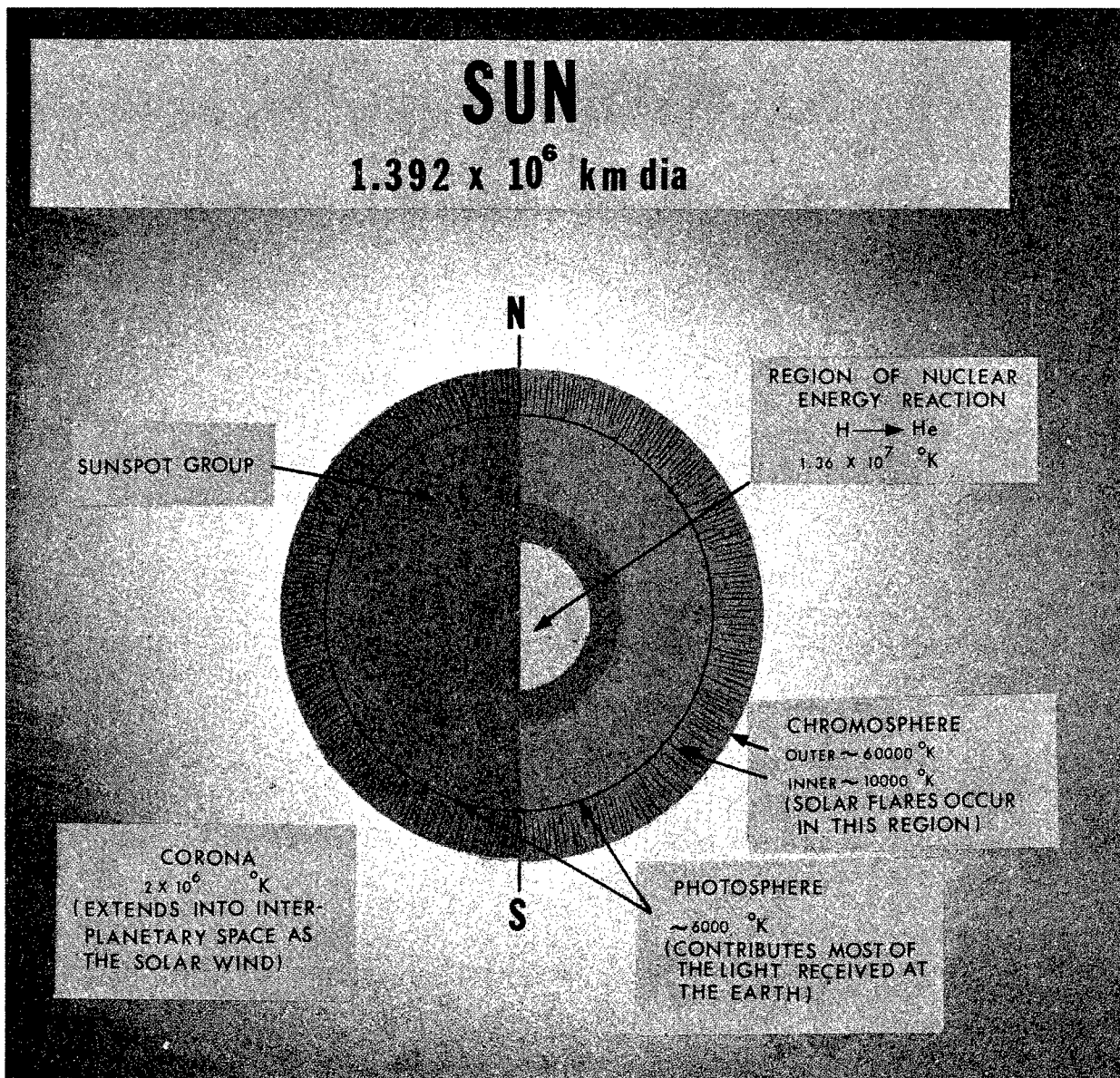
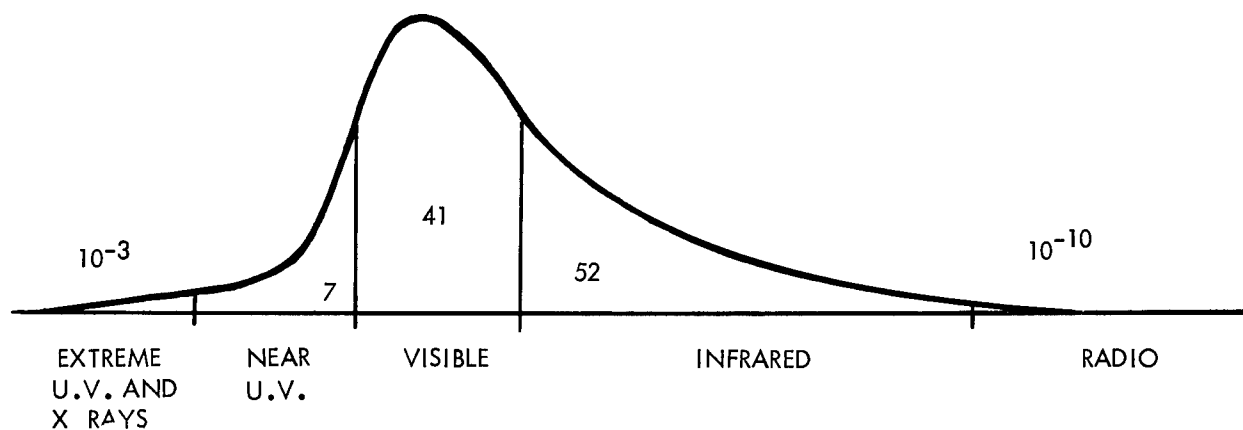


FIGURE VIII-1.

understand how certain parts of this picture affect Air Force operations, and hence, what is being forecast, and why these forecasts are important. Figure VIII-1 shows two views of the sun put together, the right half being a cut-away section of the sun extending from the core out to the high atmosphere, and the left half of the picture including a view of the disk as we usually think of it, plus some of the high atmosphere of the sun which we usually don't see, except with specially equipped telescopes. Returning to the right half of the slide, there is a small section way down in the solar interior where the temperature is over

ten million degrees, and where all of the energy that heats the sun is produced in nuclear reactions. Between this section and the part that we usually consider the surface of the sun (which is labeled the photosphere) is a large region usually considered to be a zone of violent convection. The top of this zone is the photosphere, where the temperature has dropped to around six thousand degrees kelvin.

In the photosphere, that is the visible surface as we know it, we see such things as the sunspots, the little dark groups of spots which locate the centers of all of the interesting solar activity. It is the activity in and around the sunspots which



SCHMATIC DIAGRAM OF THE DISTRIBUTION OF ENERGY IN THE SOLAR SPECTRUM (NOT TO SCALE). THE NUMBERS ARE PERCENTAGES OF THE SOLAR CONSTANT, $H = 2.00$ CALORIES/SQUARE CENTIMETER/MINUTE = 1390 WATTS/SQUARE METER. THE FIGURE FOR THE RADIO ENERGY IS FOR THE OBSERVED BAND FROM 15 TO 30,000 MHZ.

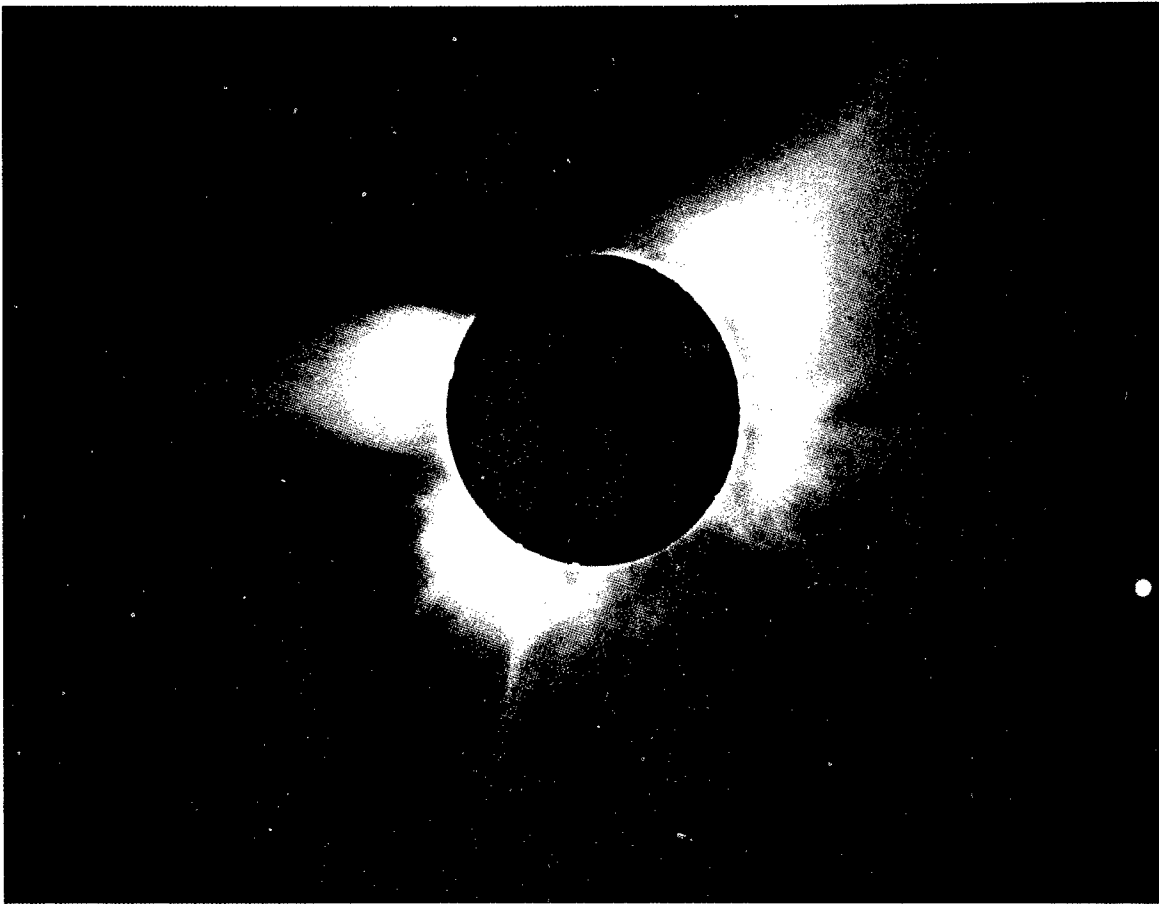
FIGURE VIII-2.

produces the radiations which affect Air Force (space-weather) operations. These sunspots extend below the visible surface of unknown distance, and above it up into the area which is labeled on this picture the "chromosphere". Most of the visible light passes right through the chromosphere with very little interference or loss. The temperature of the chromosphere increases steadily from the photosphere out through the chromosphere and then out into the corona, which is the halo-like ring we see at the times of total solar eclipse. It should be noted at this point that the photosphere (which contributes most of the light received at the earth) is also the coolest layer of the sun. The temperature increases steadily both below the photosphere into the solar interior, and also above it up into the chromosphere and the corona. The chromosphere, as the name implies, is noted for strong emission lines in various places in the solar spectrum. One particular line is in the red part of the solar spectrum (an emission line of hydrogen) and it is in this line, the so-called H-alpha line, that we monitor a large amount of the solar activity. By taking pictures of the sun in this line continuously we watch for the most significant occurrences of solar activity, which are called solar flares, and about which we'll have more later.

Solar energy is radiated over the entire electro-

magnetic spectrum, extending from the very shortest waves (x-rays) up through the ultraviolet (of slightly longer wavelength) into the visible (the next longest wavelength part of the spectrum) to the infrared, and up into radar and radio waves which have very long wave lengths. The x-rays and ultraviolet radiation are absorbed in the earth's high atmosphere and never reach the surface. Most of the "visible", part of the infrared radiation and much of the radar and radio radiation on the other hand, reach the earth's surface. The total amount of radiated energy from the sun is very nearly constant with time, with the bulk of the radiation coming in the visible and infrared part of the spectrum. Significant increases in the radiation though do occur in some of the frequencies and are associated with solar activity, especially with solar flares. These increases in radiation come primarily in the radio, the ultraviolet, and x-ray parts of the spectrum where the total amount of radiated energy from the sun is very small. Hence, even though the radiation increases as a result of solar flares may be as much as a thousand percent over the background, the contribution to the total emission of the sun is quite small.

Figure VIII-2 is a schematic diagram of the distribution of energy from the sun over the entire



SOLAR CORONA - 12 NOVEMBER 1966. COURTESY HIGH ALTITUDE OBSERVATORY, BOULDER COLORADO.

FIGURE VIII-3.

solar spectrum, and while it shows the general picture, it is not to scale. The numbers underneath the heavy line are percentages of the total solar energy emitted, that is percentages of what is called the solar constant. In this figure the wavelength of the light increases from the very shortest wavelengths, the extreme ultraviolet and x-ray radiation on the left side of chart, through the near ultraviolet, the visible, the infrared, and then out to the radio wavelengths extending up to 30,000 MHz.

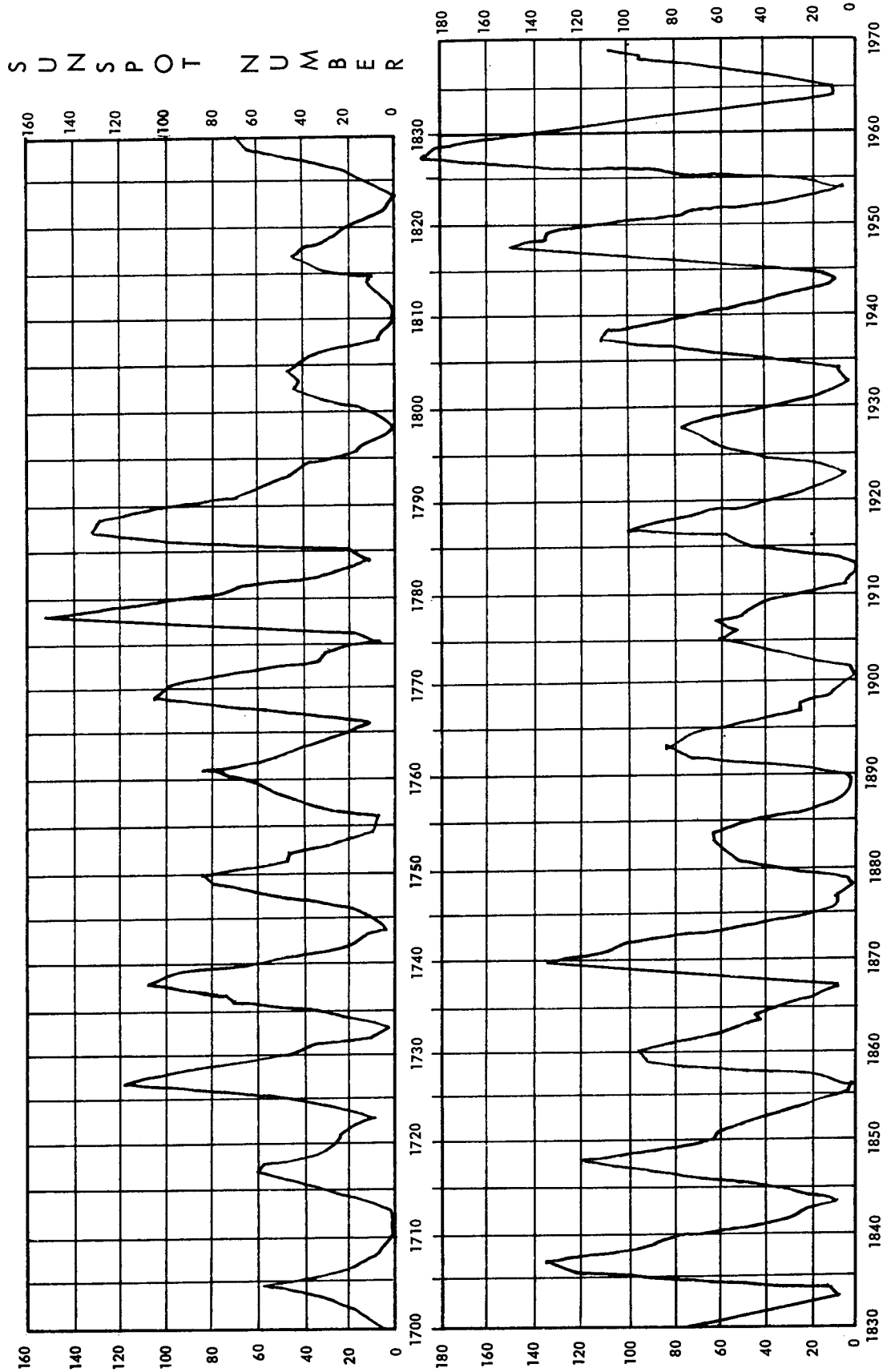
Speaking only from an energy standpoint, this picture shows most of the solar radiation, that is, it covers the wave radiation. The sun does give out a different kind of radiation in addition, called particle radiation. In fact, it emits a steady stream of particles which is called the solar wind. The total number of these particles is extremely small and the total amount of energy emitted from the sun in the form of particles is very minor compared

to the emitted wave radiation. These particles, primarily protons and electrons, are emitted from the corona and travel outward through the entire solar system.

In the regions of activity on the sun, where the sunspots are located, there are from time to time very rapid enhancements in emission, which are called solar flares. These solar flares, in addition to giving off the wave radiation in the form of a brightening, also hurl out streams of particles, some of which arrive near the earth and affect the earth's magnetosphere and upper atmosphere. Some of these particles cause the aurora borealis, the northern lights. There are other solar particle streams, not associated with solar flares, which are also hurled out from the sun to produce the aurora borealis.

Figure VIII-3 is a picture of the solar corona—this is the part of the sun's atmosphere which extends well out from the solar disk. The black circle

SOLAR CYCLE

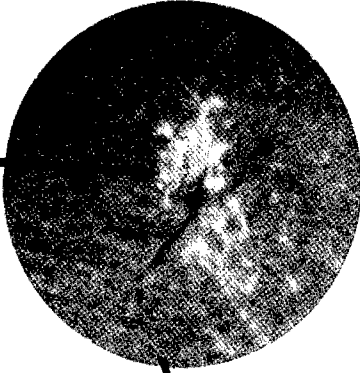


VARIATION OF THE SUNSPOT NUMBER
DURING THE YEARS 1700 TO 1968

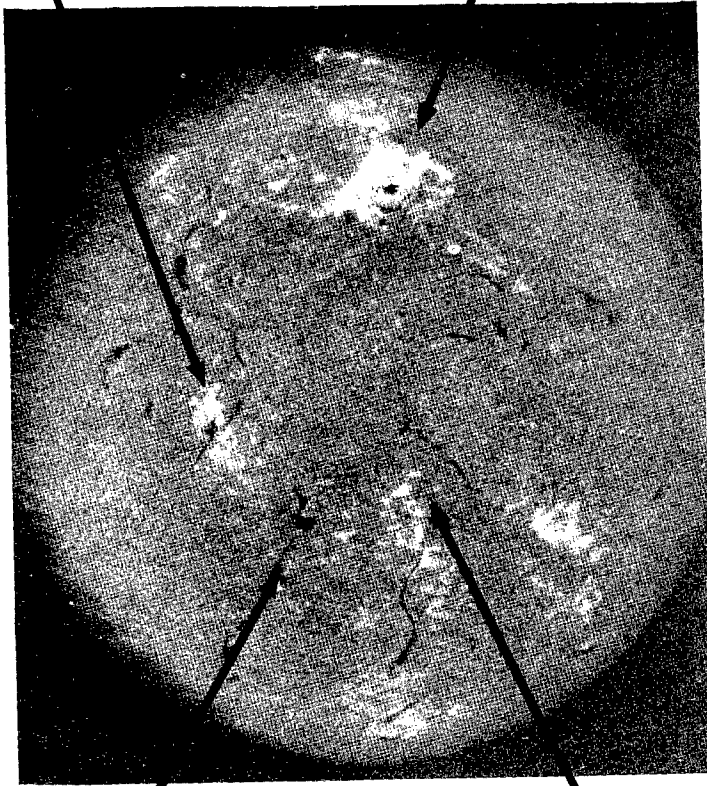
FIGURE VIII-4.

REGIONS OF SOLAR ACTIVITY

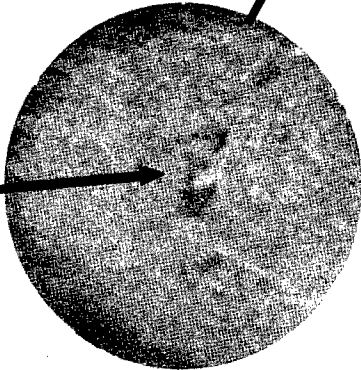
PLAGE



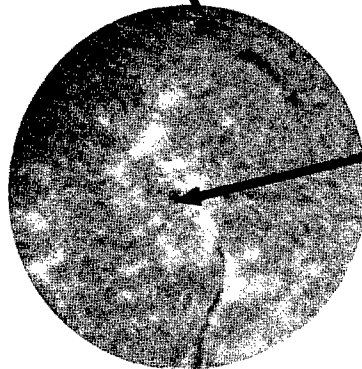
FLARE



FILAMENT



SUNSPOT



THE SOLAR DISK AS OBSERVED IN H-ALPHA.

FIGURE VIII-5.

in the center is an occulting disk built into the telescope to block out the intense light from the face of the sun, performing the same job as the moon does at times of total solar eclipses. The ray features extending out around the occulting disk vary not only during the sunspot cycle itself, but also as a function of the sunspot regions and associated activity near the limb at the base of these streamers. While it's easy to state that the shape, intensity and extent of the coronal streamers is related to the solar activity at their base, attempts to correlate a particular set of streamers with a particular active center on the disk have not been very successful.

Solar activity occurring on the face of the sun, and related to sunspots, goes up and down over a period of about eleven years. For example, back in 1957, we had the greatest amount of solar activity that had ever been observed. It had climbed up from 1954 when there were periods with almost no sunspots on the disk. After 1957, the activity slowly declined toward the next sunspot minimum in 1964. Since 1964, solar activity has been going up and seems to be leveling off at a maximum either last year or this (1969), and it will decline for the next six or seven years.

Along with the sunspots, we observe other phenomena on the face of the sun and in the chromosphere just above. For example, the flares, which I've already mentioned, happen over a period of minutes or hours. In any one sunspot region there may be many dozens of such flares, some large and many others quite small. Around the sunspots themselves there'll be a bright area which will appear before the sunspots themselves appear, cover an area to enclose all of the sunspots, and it will persist long after the last sunspot has faded out. These bright areas are called plages. This entire conglomerate of activity is considered an active region. However, because sunspots are the easiest phenomena to observe, and have been measured the longest, they're usually considered to define the solar cycle.

Figure VIII-4 is a trace of the sunspot number from the year 1700 through the year 1968. Some cycles are very much more active than others, and during many sunspot minima the activity does not really stop. Even the length of time between the successive maxima of sunspot activity can vary anywhere from about eight to about fifteen years. The highest sunspot cycle on record, as I previously noted, was the one in the late 1950's. In the realm of sunspot cycle forecasting, many

people have spent many hours drawing conclusions out of this chart. (I should state, many *erroneous* conclusions out of this chart.) There have been attempts to project solar activity into the future, with little or no success, and there have been even more ambitious attempts to forecast solar activity extending fifty or a hundred years into the future with even less success.

In Figure VIII-5 we see a picture of the sun as observed in an emission line in the red part of the solar spectrum (H-alpha). There are many regions of solar activity shown, the most outstanding one being in the right center part of the picture, a blowup of which is shown in the inset in the lower right. The dark spot in the center is the sunspot (with a good picture even more spots could be seen). The bright area around the spot is a solar flare which has covered most of the plage area underlying it. A typical plage area, at a time when no flare is occurring, can be seen in the inset at the upper right. It surrounds the darker areas of sunspots and persists with little change in structure over periods of many weeks. There are scattered over the face of the sun in this picture a number of other active regions, one down in the lower left, and another one just above it.

From here on, the emphasis of my paper will be on solar flares and their associated wave and particle emissions. It is the emissions from the flares which produce almost all of the variations in space weather which are important for Air Force operations, and their travels and effects are shown schematically in Figure VIII-6. Flares are sudden, relatively short-lived brightenings in the plage region around the sunspot centers. The flares start suddenly, and have lifetimes of a few minutes to a few hours (the average being about half an hour). The energy of one of the large flares is equivalent to the explosion of a billion H-bombs. Enhanced electromagnetic and particle radiation emitted during a flare affects the entire area between the earth and sun, the geomagnetic field of the earth, and the earth's upper atmosphere. The particle streams from the flare distort the earth's magnetic field. They also raise the radiation levels in the Van Allen belts and heat the earth's high atmosphere. They disturb the ionosphere, and as a result, long-range radio communications which depend on reflections in the ionosphere, are considerably distorted, and sometimes completely disrupted.

In Figure VIII-7 some of the characteristics of solar flares are shown. Flares are classified according to their size or importance, and their

ALL FLARES PRODUCE ENHANCED ELECTROMAGNETIC RADIATION
 BUT NOT ALL YIELD PARTICLE STREAMS DETECTABLE AT THE
 EARTH

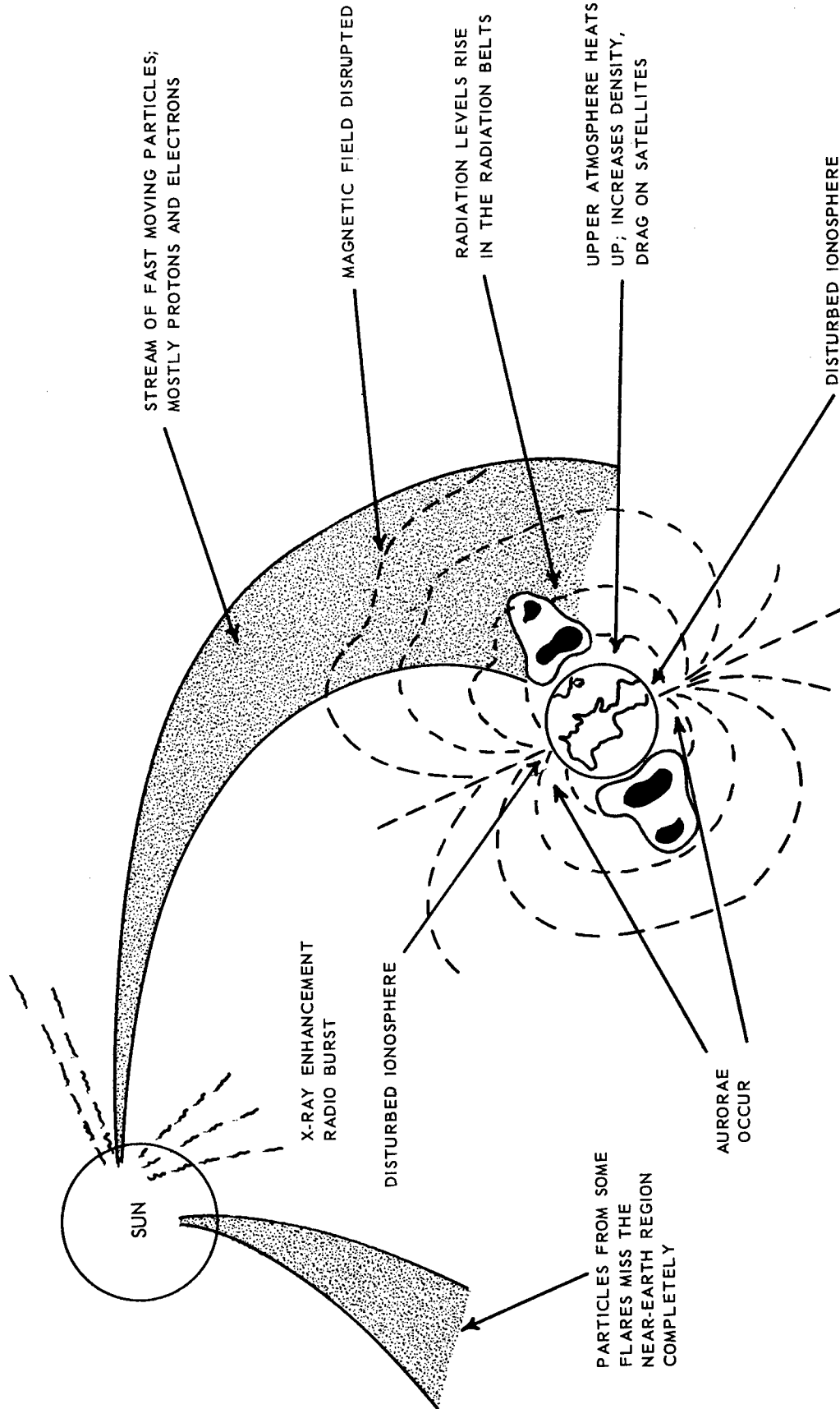


Figure VIII-6.

SOLAR FLARE CLASSIFICATION

IMPORTANCE	AREA* (Millionths of Solar Hemisphere)	AVERAGE DURATION
0	Less than 100	17 min
1	100 to 249	32 min
2	250 to 599	69 min
3	600 to 1200	145 min
4	Greater than 1200	145 min

BRIGHTNESS CATEGORIES: FAINT (F)
 NORMAL (N)
 BRILLIANT (B)

* The area of the earth's disk is approximately equivalent to the area of an Importance one flare.

FIGURE VIII-7.

brightness. The importance of a flare is defined by the flare's area at maximum brightness, and the brightness itself is a qualitative assessment given by the observer at the telescope. In general, the higher the importance classification the stronger the geophysical effects will be. But this general rule has many, many exceptions. Sub-flares (importance 0) occur most often (probably seventy-five percent of the time) but since they are the most difficult to observe, any accurate assessment of their frequency is difficult. Less than one percent of all flares are of importance 4—the really big ones. The effects on the earth produced by solar flares depend on whether the effect is produced by the wave radiation from the sun (which travels with the speed of light and starts its effect coincidentally with the observation of the flare) or by particles which travel at less than the speed of light, and which may take anywhere from some minutes to a few days to arrive in the vicinity of the earth.

The effects of the wave radiation from the flare are the radio bursts observed in the radar and radio wavelengths, and sudden ionospheric disturbances which are the result of the X-radiation from the flare. The delayed effects depend on the speed of

the particles emitted from the flare. The aurora borealis (the northern lights) is the most well-known effect of solar flares and is produced from two to three days after the flare by some of the lowest energy particles emitted. Coincident with the aurora borealis are magnetic and ionospheric storms. Going up the ladder, higher energy particles, arriving about twelve hours after a flare, can cause a complete blackout in radio communications in the polar regions. Some of the most immediate effects of solar particles are produced by the fastest, most energetic particles from the sun and these can happen within less than an hour.

We'll talk first about what are called "sudden ionospheric disturbances", as shown in Figure VIII-8. These are sudden, in that they occur simultaneously with the solar flare itself, being caused by the solar x-ray and ultraviolet radiation. The radiation affects the ionosphere by increasing the ionization of the lower ionospheric regions. The SID's, as they are called, affect radio wave propagation by absorbing high-frequency radio signals, by causing high frequency deviations, by causing very low frequency—and low frequency phase changes, and by enhancing the very low frequency—and low frequency signals. The ef-

SUDDEN IONOSPHERIC DISTURBANCES (SID)

Occur simultaneously with the flares.

Caused by increase of solar X-ray and ultraviolet radiation which increases the ionization in the lower ionospheric regions.

SID's affect radio wave propagation by:

Absorbing HF signals.

Causing HF frequency deviations.

Causing VLF and LF phase changes.

Enhancement of VLF and LF signals.

Effects endure on a time scale of minutes to several hours.

FIGURE VIII-8.

ffects of the x-ray and ultraviolet radiation last from a few minutes to several hours.

Figure VIII-9 shows some of the phenomena, and the frequency ranges in the radio communication spectrum that are involved. For example, there are classes of events called sudden frequency

deviations (SFD) which effect the high frequency communications, and short wave fade-outs (SWF), which are complete or partial fade-outs of the signal (either high frequency or very high frequency). There are also sudden phase anomalies (SPA) in the very low frequencies, sudden atmos-

SUDDEN IONOSPHERIC DISTURBANCES

PHENOMENON

FREQUENCY RANGES INVOLVED

Sudden Frequency Deviation (SFD)

HF

Short-Wave Fadeout (SWF)

HF and lower VHF

Sudden Phase Anomaly (SPA)

VLF

Sudden Enhancement of Atmospherics (SEA)

VLF

Sudden Enhancement of Signal (SES)

VLF

Sudden Cosmic Noise Absorption (SCNA)

HF (Cosmic radio noise) and lower VHF

FIGURE VIII-9.

SOLAR PARTICLE EVENTS

EVENT	CAUSE (Particles)	ARRIVAL TIME AFTER FLARE
Ground Level Event	Very High Energy Protons	15 min - 1 hr
Polar Cap Absorption	High Energy Protons	20 min - 20 hrs
Magnetic and Ionospheric Storms; Aurorae	Low to Medium Energy Protons and Electrons	20 hrs - 72 hrs

FIGURE VIII-10.

pheric enhancement (SEA), sudden enhancement of the signal (SES), and sudden cosmic noise absorption events (SCNA). In Figure VIII-10 we can see some of the effects of the arrival of solar particles in the earth's vicinity. For example, there are ground level events (GLE) caused by very high energy protons which are little affected by either the earth's magnetic field or by the atmosphere, and hence make it all the way down to the ground. Since they are of very high energy, they arrive ahead of the other particles, anywhere from fifteen minutes to an hour after the flare. There are somewhat less energetic (but still very high energy) protons which cause the polar cap absorption events, shown in Figure VIII-11. These polar cap absorption events (PCA) completely black out radio communications in the Arctic and Antarctic regions. They start twenty minutes to twenty hours after the flare is observed. The laggard (low energy) particles, those stopped mainly by the high atmosphere and the earth's magnetic field, cause the magnetic and ionospheric storms and the aurora borealis. These arrive in the earth's vicinity from a day to three days after the solar flare.

The most extreme of these events are the ground level events. These are caused by extremely high energy solar particles incident upon the earth's upper atmosphere. The particles travel from the sun at velocities approaching the speed of light. Fortunately, they don't occur very often. There

were only sixteen such events observed during the period between 1942 and 1968. These are normally followed by, but don't cause a polar cap absorption event.

The polar cap absorption event is caused by high energy protons deflected toward the polar regions by the earth's magnetic field, producing increased ionization in the D-region. High frequency radio waves are severely absorbed as they pass across the polar regions, and this attenuation can last for several days. It's been called the Polar Cap Event or a Polar Cap Blackout or a Polar Blackout, but whatever term is used, there is very severe disruption of polar communications. Polar cap absorption events are normally followed by disturbances in the geomagnetic field and in the ionosphere.

The third class of events are caused by the lowest energy particles which reach the earth anywhere from one to three days following their emission from the sun. These cause the geomagnetic and ionospheric storms and auroral activity. There are three phases to a geomagnetic storm, normally comprising an initial phase which includes a sudden commencement, followed by a quiet period lasting several minutes or hours. After the lull, the main phase begins, reflected by an increase in the disturbance, persisting for several hours or days. This disturbance is measured by rapid variations in the magnetic field. The stronger the fluctuations in the

POLAR CAP ABSORPTION

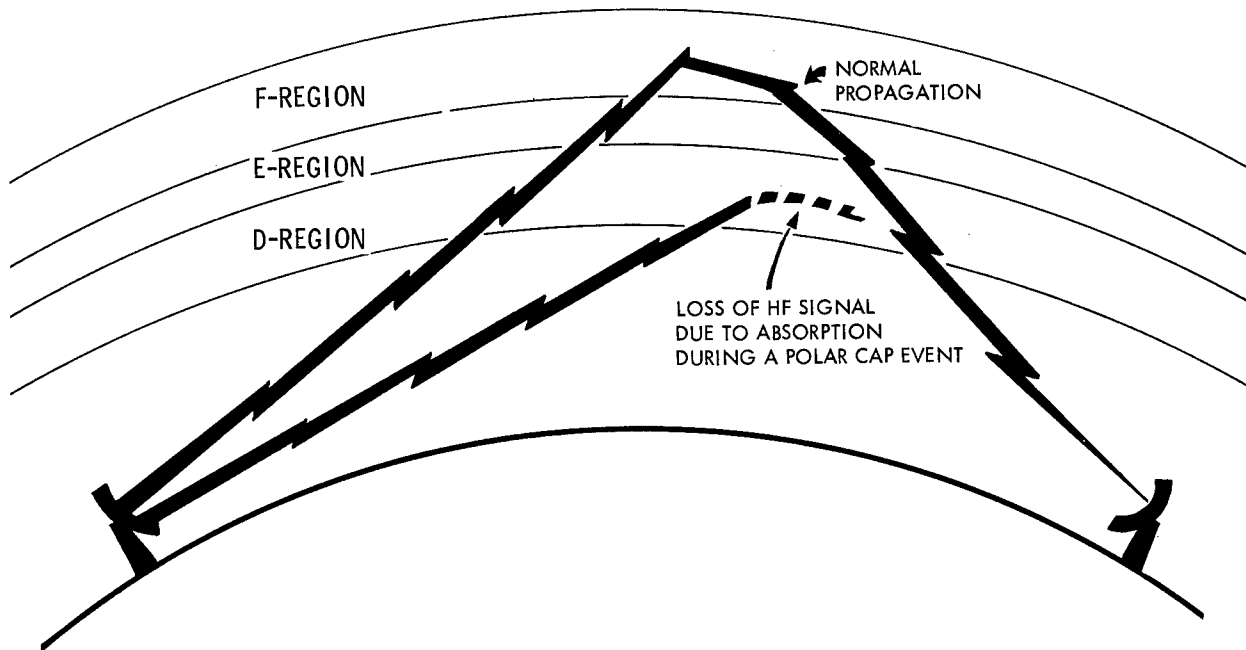


FIGURE VIII-11.

magnetic field, the higher the disturbance. The main phase is followed by a gradual return to normal, the recovery phase.

Various indices are used to describe the magnitude of the geomagnetic disturbance, the most common ones are the K-index and the A-index, usually followed with a subscript "p" to indicate a planetary-wide measure of the disturbance. The occurrence of the *major* geomagnetic disturbances are quite rare.

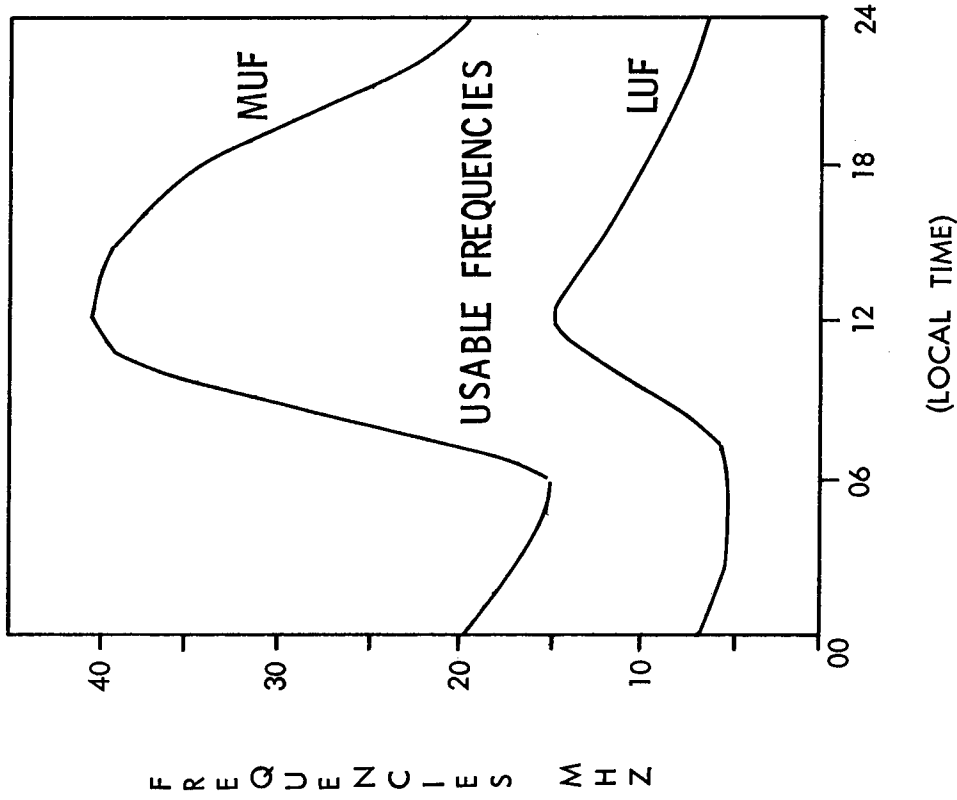
Ionospheric storms usually go along with the geomagnetic storms, because the same enhanced flux of particles that deflects the earth's magnetic field also produces significant changes in the ionosphere. The F_2 region, which is the highest ionospheric layer in normal use, is affected in such a way that the critical frequencies are reduced. This results in a reduction in the highest frequency that is usable for long distance radio communications. On the low end of the spectrum, the increased absorption in the D-region means that there's an increase in the lowest usable frequency. The net result of the maximum usable frequency coming down, and the lowest usable frequency going up, is a compression in the range of usable frequencies for long-distance radio communication. The effects

of these storms are more pronounced at higher latitudes.

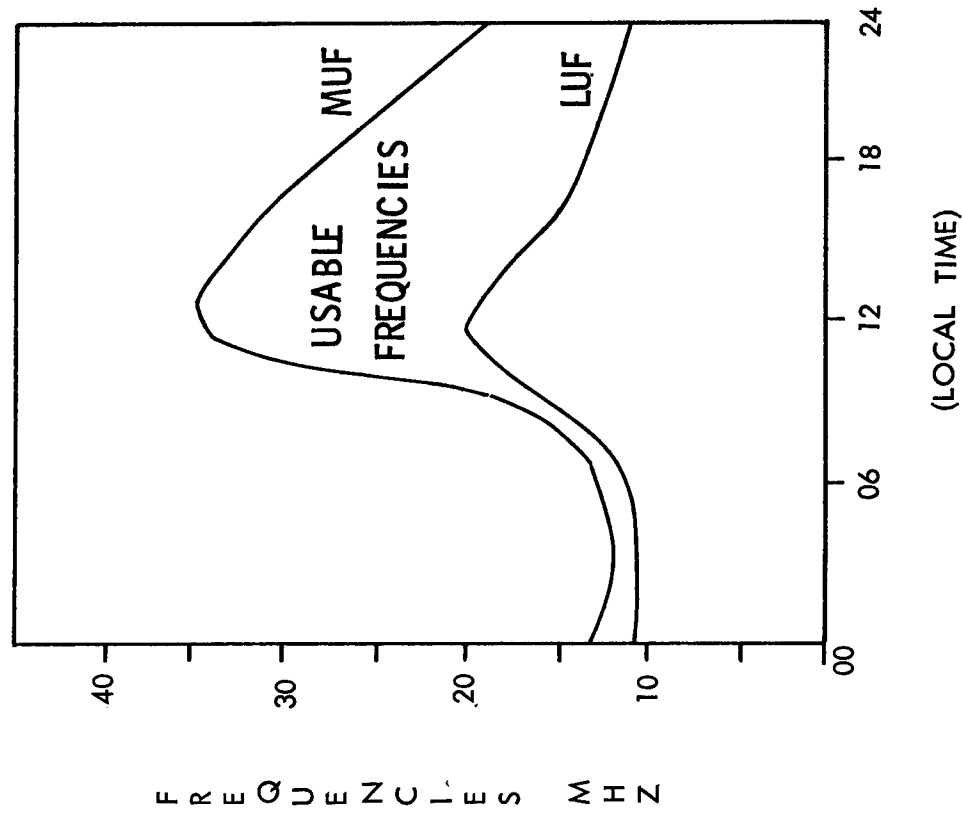
In Figure VIII-12 we can see two typical conditions for the propagation of radio waves. The left side of the chart shows the range of usable frequencies as a function of the local time of day under quiet ionospheric conditions. The maximum range of usable frequencies occurs around local noon, but there is always a reasonably good range of about 10 MHz or more at all times of the day. Under the disturbed ionospheric conditions shown on the chart at the right, there are times when the band of usable frequencies is very restricted. The lowest usable frequency rises as a result of increased absorption in the D-region and the highest usable frequency goes down as a result of significant changes in the F_2 region.

The aurora borealis is a visual indication of the disturbances going on in the upper atmosphere. There are aurorae observed on almost every clear night of the year in the auroral zones. In Figure VIII-13 the auroral zone is marked in heavy shading, running from just north of Russia on the upper side of the North Pole down over Alaska, through central and southern Canada, across the Atlantic between Iceland and the British Isles,

HF RADIO PROPAGATION

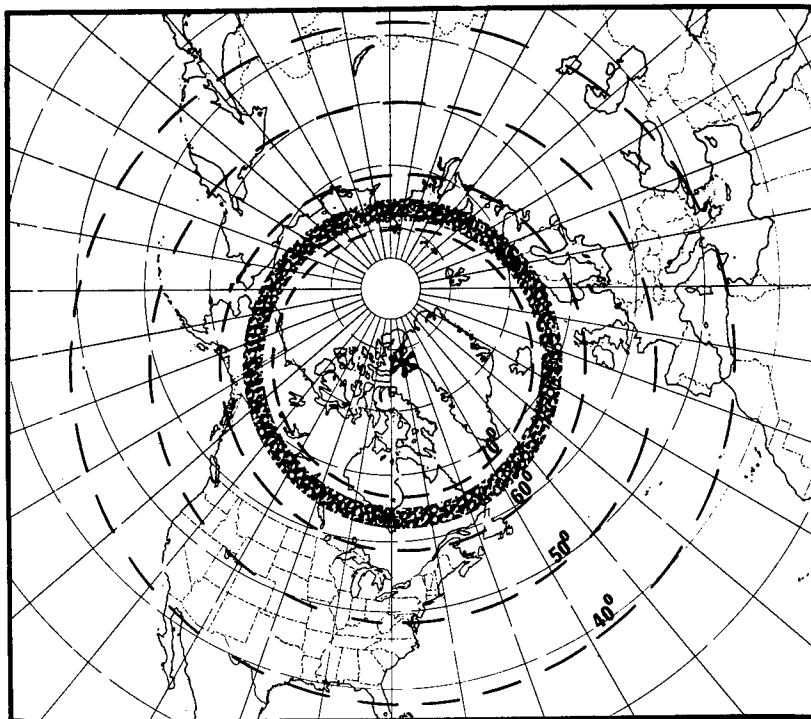


QUIET IONOSPHERIC CONDITIONS



DISTURBED IONOSPHERIC CONDITIONS

FIGURE VIII-12.



 AURORAL ZONE

 GEOMAGNETIC NORTH POLE

 GEOMAGNETIC LATITUDES

FIGURE VIII-13.

and then off the north coast of Norway. Departing from this zone in either direction, the frequency of aurorae decreases. During the largest geomagnetic storms, the auroral particles will be seen quite far south. Whenever there are reports of northern lights in the southern part of the United States, you can be very sure that there's a very intense geomagnetic storm in progress. Reports from the northern United States indicate at least moderate storminess. As far as Air Force operations are concerned, the most severe radio blackouts occur in the auroral zones especially during times of enhanced magnetic activity.

EFFECT ON AIR FORCE OPERATIONS

With this as background, the next subject is the effect of solar activity on military operations. The effects can be put into three general categories: satellite surveillance, radio communications

or lack thereof, and effects on the manned-space program—the manned-orbiting laboratory (MOL).

Increased solar activity from solar flares, which produces magnetic storms and large values of the A-index, etc., increases the atmospheric density at certain levels. The increased density in the high atmosphere increases the drag on satellites. As a result of the increased drag, the orbital speed of the satellite decreases. The variations in orbital velocity introduce difficulties in tracking, and in determining where certain satellites or parts of satellites are located. The increased density results in lowering orbits, and when a satellite is close to the point where its whole orbit is going to decay, a change in solar activity can produce a change of many days or weeks in the time at which this satellite will re-enter or burn up.

Some of the most important and complex problems in space-weather forecasting revolve around radio communications. And in the communications business, the problems resulting from solar activity

RADIO COMMUNICATIONS

FREQUENCY RANGES	PROPAGATION MODE
Very Low Frequency (VLF) - 3-30 KHz	Reflection from the D-region
Low Frequency (LF) - 30-300 KHz	Reflection from the D-region
Medium Frequency (MF) - 300-3000 KHz	Day: line-of-sight; night: reflection from the E-region
High Frequency (HF) - 3-30 Mhz	Reflection from the E- or F-regions
Very High Frequency (VHF) - 30-300 Mhz	Line-of-sight or scattering by D- and E-region inhomogeneities
Ultra High Frequency (UHF) - 300-3000 Mhz	Line-of-sight or scattering by tropospheric irregularities
Super High Frequency (SHF) - 3000-30000 Mhz	Line-of-sight or scattering by tropospheric irregularities

FIGURE VIII-14.

MILITARY RADIO COMMUNICATION SYSTEMS

SYSTEM	FREQUENCY RANGES
High Frequency	HF
Ionospheric Scatter	Lower VHF
Tropospheric Scatter	UHF to lower SHF
Microwave	UHF - SHF
Satellite	VHF - UHF - SHF

FIGURE VIII-15.

TRANSMITTED FREQ TOO HIGH - PASSES INTO SPACE

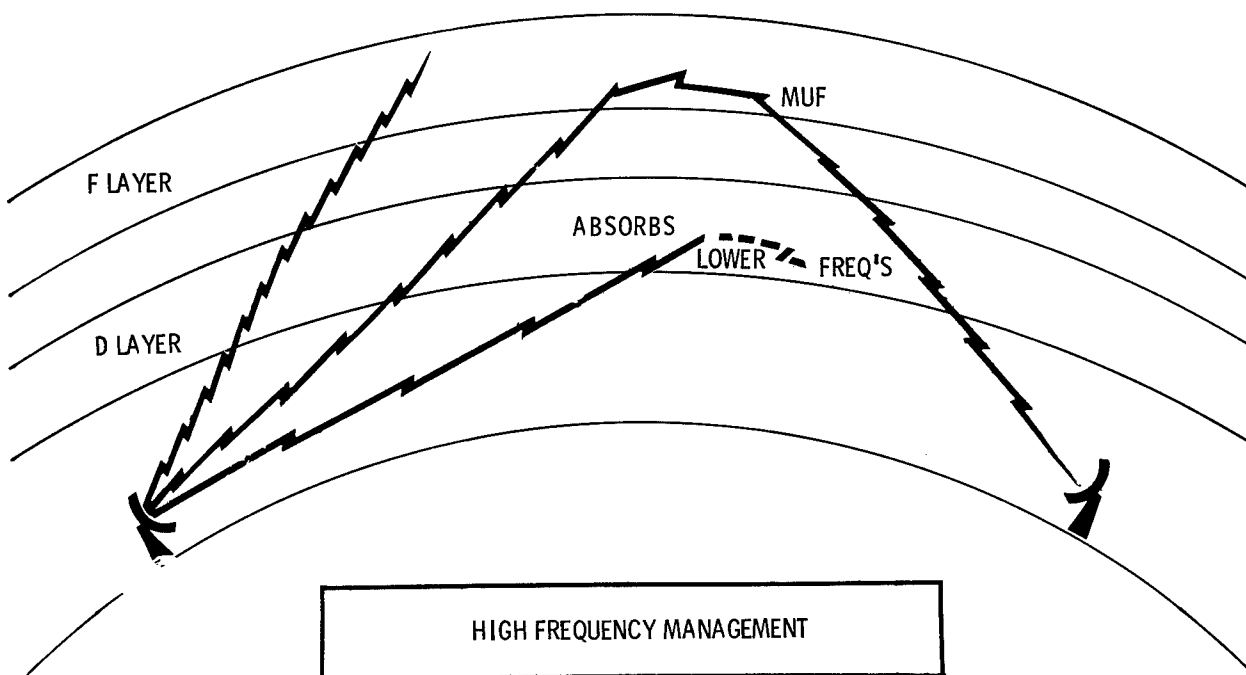


FIGURE VIII-16.

depend on the propagation mode, that is, what part of the ionosphere is used for propagating the signal over long distances. These are shown in Figure VIII-14. For the very low frequency (VLF) from three to thirty kilohertz, or the low frequency (LF) mode from thirty to three hundred kilohertz, we depend on reflections from the lowest layer of the ionosphere, the D-region. The next highest frequencies, the medium frequency (MF) radio waves in the three hundred to three thousand kilohertz range, use line of sight transmission in daytime and reflection from the E-region at night. The high frequency (HF) three to thirty megahertz is propagated over long distances by one or more reflections from the E or F regions. The very high frequency (VHF) thirty to three hundred megahertz is propagated by line of sight, or scattering by D and E region inhomogeneities, and the ultra-high frequency (UHF) three hundred to three thousand megahertz and superhigh frequency (SHF) three thousand to thirty thousand MHz by line of sight or scattering by tropospheric irregularities (the troposphere being the region that we live in, from the surface up to ten or twenty kilometers).

The military radio communications systems can be divided into five different types as shown

in Figure VIII-15: the high frequency system, which of course is in the high frequency range; the ionospheric scatter system, which uses the lower part of the very high frequency range; the tropospheric scatter system, which uses ultra-high frequency to super-high frequency ranges; the microwave system, which uses ultra-high and super-high frequencies; and satellite communications, which use VHF, UHF and SHF. In the high frequency system, we have the problem of frequency management as shown in Figure VIII-16. The high frequencies normally propagate by reflection from the F-layer for single hops, or back and forth between the ground and the F-layer for multiple hops. Frequencies too high will pass into space; frequencies too low will be absorbed in the D-layer. Forecasts are required of the range of usable frequencies between these extremes (between the maximum usable and the lowest usable frequencies) so that operators can know what frequency ranges are available at particular times.

Both the highest and lowest usable frequencies vary with solar activity, the season, and the time of day; and at times the available range of usable frequencies may be reduced to zero. Major flares decrease the strength of the transmitted signal by absorbing the signal as it passes through the D-

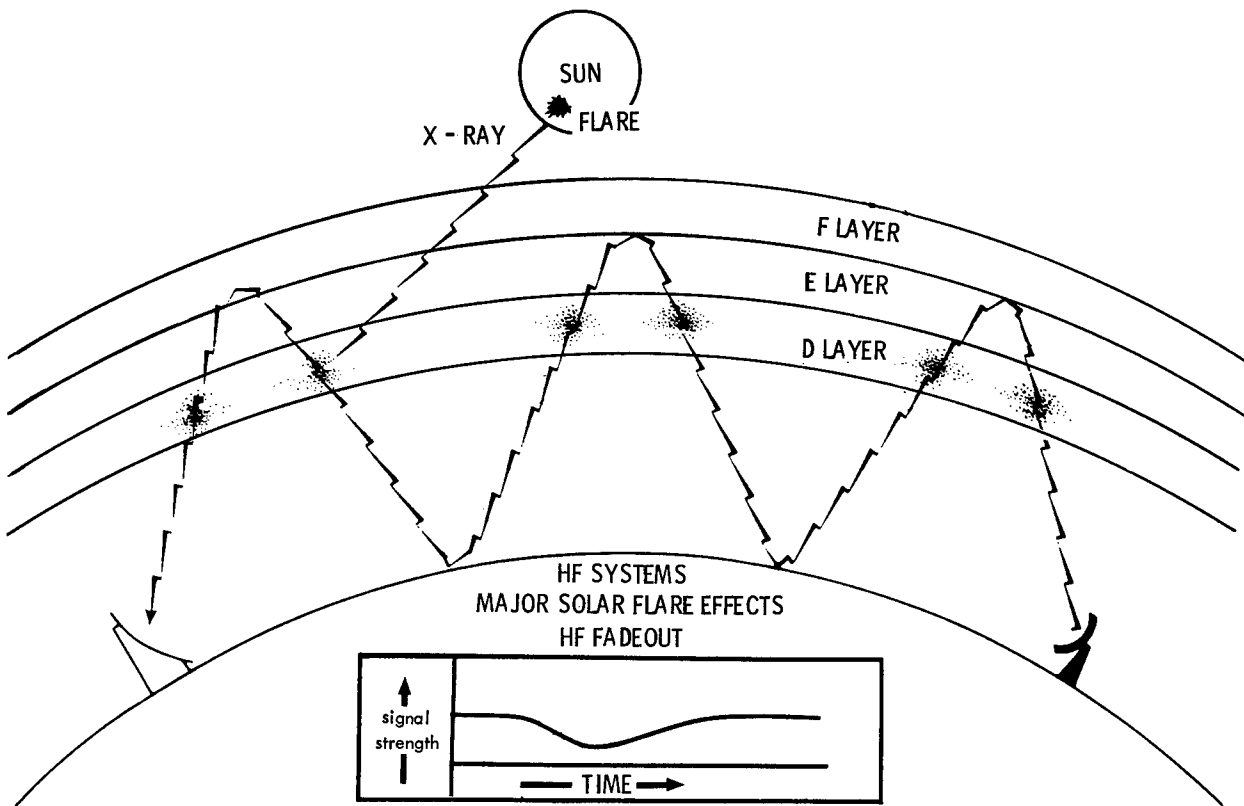


FIGURE VIII-17.

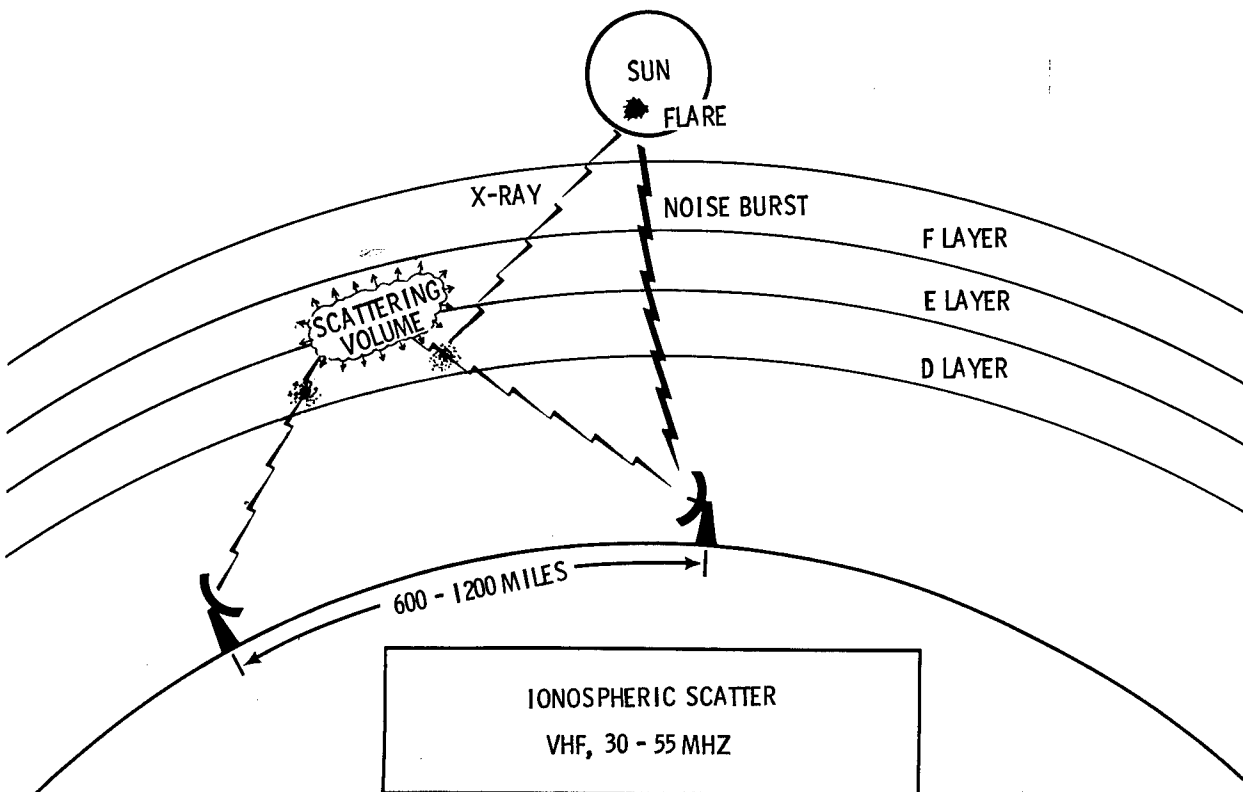


FIGURE VIII-18.

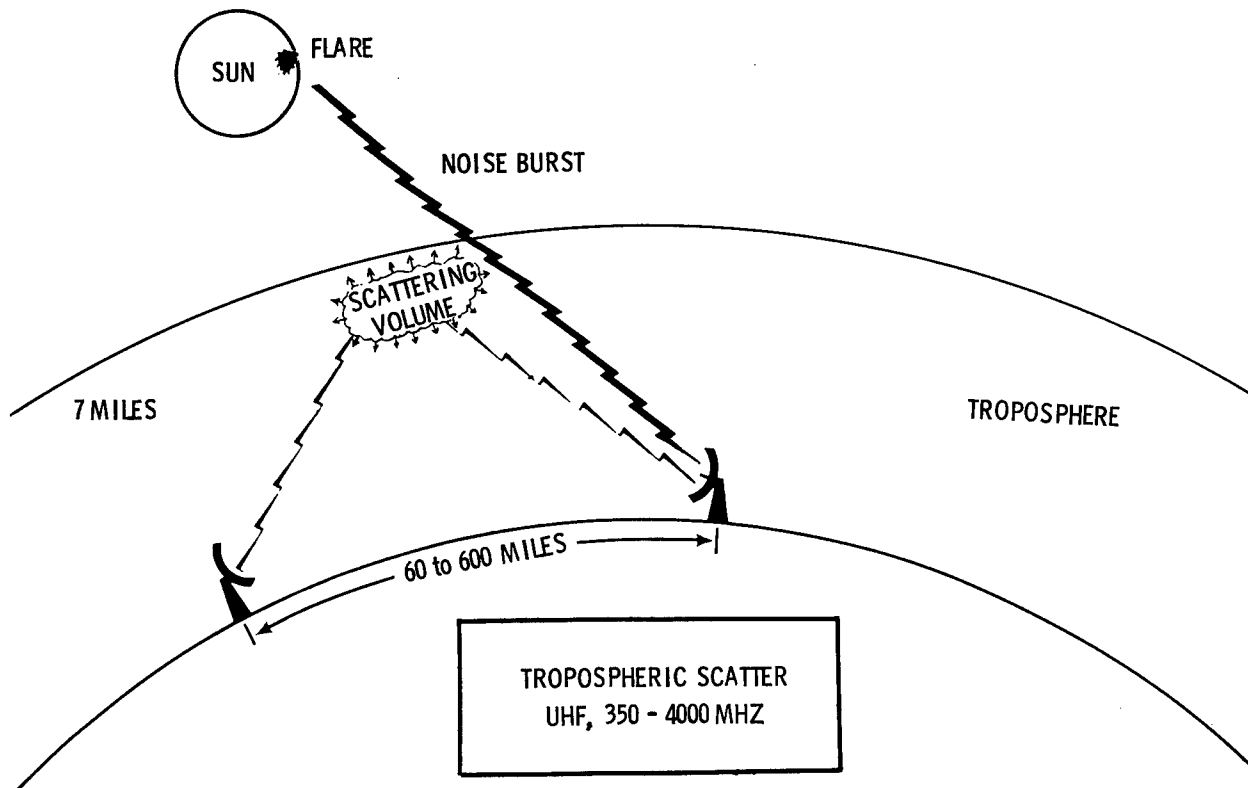


FIGURE VIII-19.

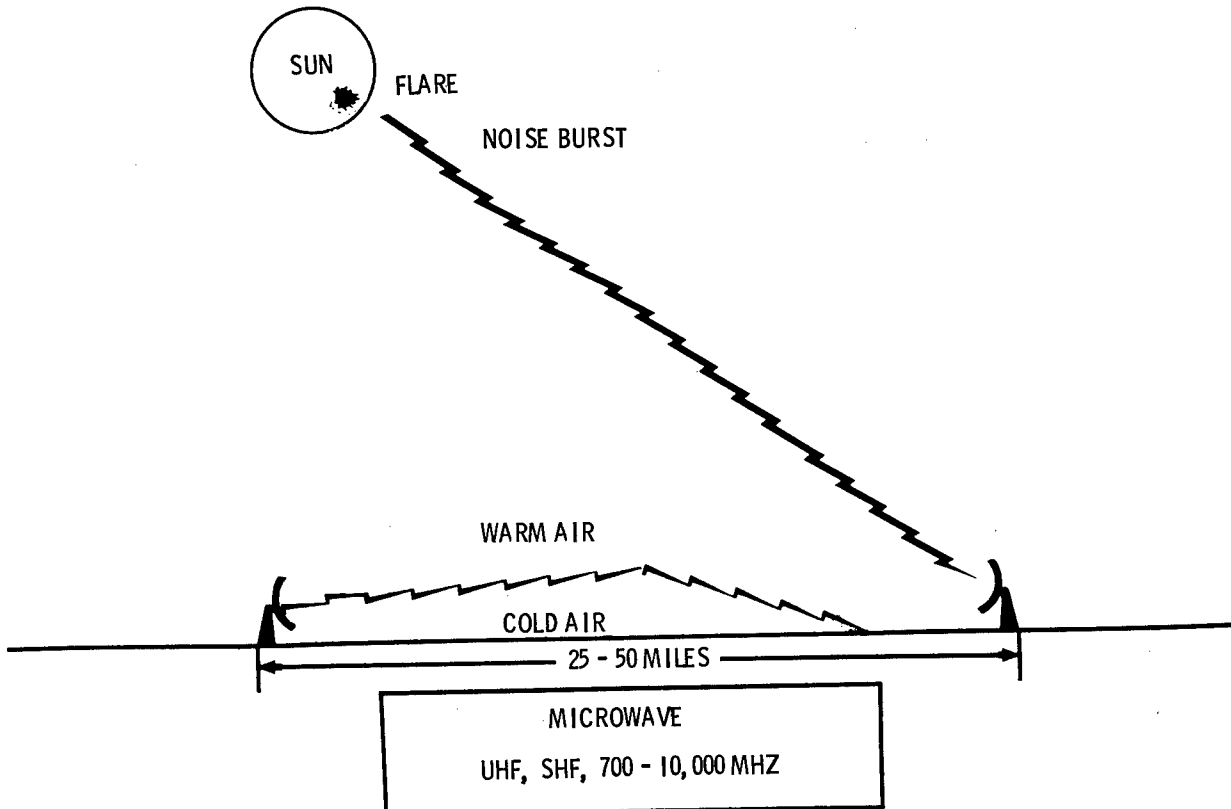
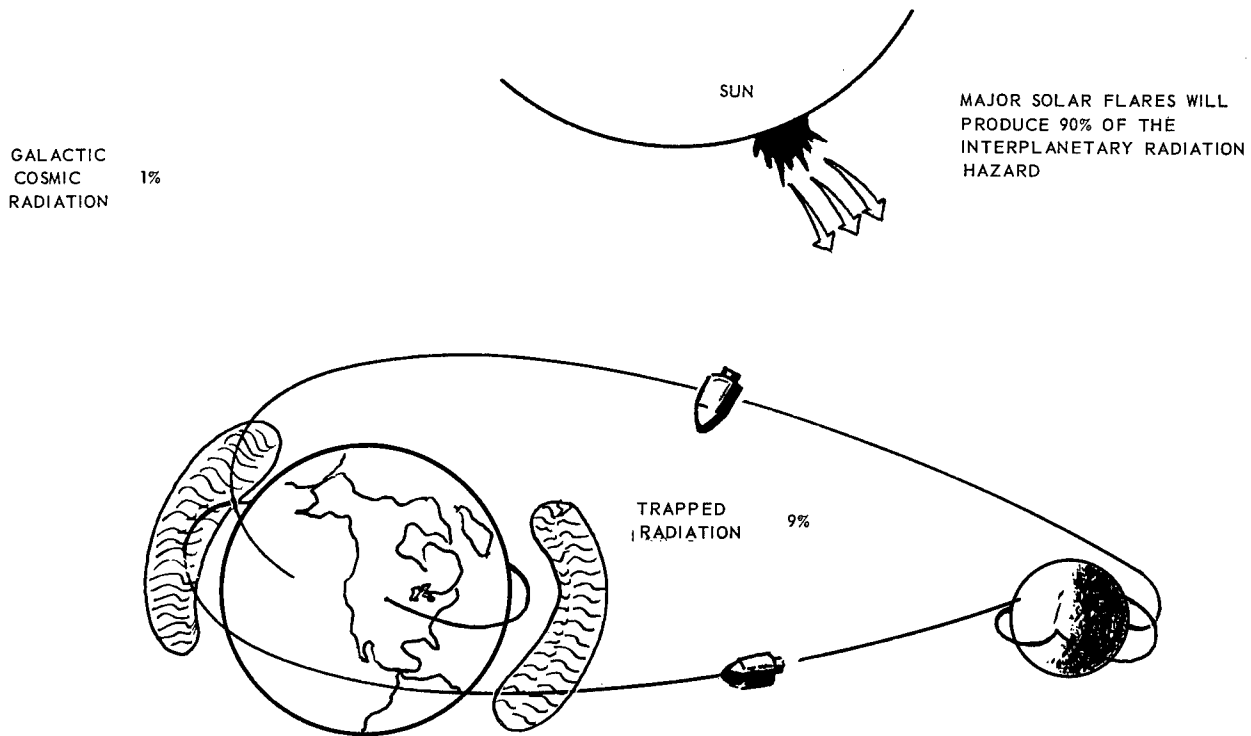


FIGURE VIII-20.

SOLAR—GEOPHYSICAL ACTIVITY EFFECTS ON THE MANNED—SPACE PROGRAM



THE NORMAL INTERPLANETARY DISTRIBUTION OF RADIATION DOES NOT PRESENT ANY SERIOUS HAZARDS, BUT TROUBLE-SOME RADIATION MAY COME FROM THE SUN DURING A SOLAR FLARE.

FIGURE VIII-21.

layer, as shown in Figure VIII-17. This absorption is caused by x-ray ionization of the D-layer by the flares. The amount of absorption from the D-layer varies with the particular flare, the position of the radio path relative to the sun, and the design characteristics of the radio system.

The second of the five military radio communications systems is the ionospheric scatter system, and in Figure VIII-18 we can see how this works. Propagation is by scattering from irregularities in the D and E layers. The received signal strength is very low, and powerful amplifiers are needed to produce a usable signal. The flare-associated radio noise burst may cause interference at the receiver if the sun is in a significant lobe of the detection

pattern of the receiver. Also, during a flare, the strength of the received signal may decrease due to absorption in the D-layer.

The third system, the tropospheric-scatter communications system, shown in Figure VIII-19, propagates by scattering from the troposphere, the layer we live in, and the layer we normally associate with weather. The received signal strength is very low and powerful amplifiers are again needed to produce a usable signal, making it also sensitive to flare-associated radio noise bursts, again, if the sun is in a significant lobe of the receiver detection pattern. Weather systems may also affect receipt of the scattered signal.

The fourth military radio communications system

is the microwave system, shown in Figure VIII-20. It propagates by line-of-sight transmission, with flare-associated radio noise bursts a problem. In addition, air masses of different temperatures cause bending of the transmitted signal, the beam being bent toward the colder air.

In the communications satellite system, which is the fifth and last of the military radio communications systems that I mentioned, radio signals propagate by line-of-sight, to and from satellites. Flare-associated radio noise bursts can cause interference at the receiver.

We mentioned earlier that the effects of solar activity on military operations fell into three areas having to do with satellite surveillance, radio communications, and effects on the manned-space program. In the military sense, we're talking about the manned-orbiting laboratory program, where military astronauts will spend extended periods in a space capsule orbiting the earth. Our problem here is in the radiation hazards produced by large solar flares. The normal interplanetary radiation does not really present any serious hazards, but during some large solar flares there may be enough damaging radiation on occasions so that the dosage is unacceptably high for the astronauts. This is shown in Figure VIII-21. It will be important when these astronauts are up there, to be able to take avoidance measures at certain points if the wrong kinds of solar flares occur.

THE FORECASTING PROBLEM

Now that I've outlined the background of solar activity and how it affects military operations, I hope that it is obvious that the real culprit is the solar flare. These flares give off bursts of energy lasting anywhere from a few minutes to a few hours. There are large ones, and small ones, and every size in between. They give off both wave radiation and particles. The energy spectrum of the wave radiation is different from every flare, and as best we can tell, the particle spectrum, that is, the amounts of particles in different energy levels, is different also. In general, the largest flares give off the largest amounts of wave and particular radiation, but this isn't always true. And most flares, with very, very few exceptions, occur in well-developed sunspot regions. Different energy ranges in the waves and the particles affect different parts of the earth and its magnetic field, and they arrive at different times. The wave

radiation arrives first, the highest energy particles shortly thereafter, and the rest dribble in over a period of hours or days, the last to arrive being the lowest energy particles. These low energy particles cause the geomagnetic and ionospheric storms and the aurora.

Perhaps at this point it would be helpful to outline the breakdown of responsibilities, and how my group at the Air Force Cambridge Research Laboratories fits into the picture. The collection of observations, and the day-to-day forecasting are the responsibilities and problems of the Air Weather Service—which is the military counterpart to the United States Weather Bureau. This is an operating agency, and by hook or by crook, they must make the forecasts that are required for Air Force operations. It's also their responsibility when they feel *either* that their forecasts are not adequate, *or* that a lot can be done to improve them to come to the research arm of the Air Force and tell them what they need. In turn, since the Air Force Cambridge Research Laboratories is the solar and geophysical part of this research arm, this problem came to our doorstep some five or six years ago. It was then that my group got involved in trying to see what we could do in the way of improving the procedures by which the Air Weather Service was making solar and geophysical forecasts, and to recommend the directions for future improvements. The Air Weather Service efforts started in the early 60's, when they organized the Solar Forecast Center, and began setting up a number of observatories to provide the data which were felt to be necessary for the forecasts. I think then, as now, we can sub-divide their basic forecasting problems into three different areas. One is the problem of forecasting some days in advance which of the sunspot regions on the disk will produce flares, or more importantly, some very significant flares. The second problem involves a forecast some hours in advance:—given a particular region that's likely to be active, when will a flare occur? And then the third problem:—after a flare has been observed, what are likely to be its geophysical effects? Will it produce a lot of hard radiation or will it be very minor in terms of its radiation and effects on the earth environment?

The AWS Solar Forecast Center was faced with solving these problems and they took the only reasonable course under the circumstances. They developed some rules of thumb, they went around and talked with experts in the field, and they had lots of discussions. But basic to all of this was very

little hard information on what the useful predictors were, how they should be put together, how the data should be organized, what data was necessary, etc. You would guess from this listing that their forecasts weren't perfect; as a matter of fact, I think an analysis would show (and I think the Forecast Center itself would admit) that the forecasts were far from perfect. The AWS was concerned about what they should do, and AFCRL was brought into the picture. We're supposed to be experts in solar activity and in the geophysical effects of solar activity, and in 1964 AFCRL set up a Space Forecasting Program to find out what was the matter with the AWS forecasts, and how they could be improved.

I could give you a chronological listing of what we did and why, and how different things worked out and why I think some things worked and others didn't. But it would take too much time, and the lessons we learned might not be obvious. I do think though that you can get a pretty good feeling if I discussed just a few of the problems that we ran into—why we ran into them—what we *proposed* to do about them—what we *finally* did about them, and what the results were. But first there's some necessary background which is important to set the stage.

Back in the early 1950's, there were people making solar and geophysical forecasts. They were forecasting when major activity would occur on the sun, which regions would develop, which wouldn't, which ones would give geomagnetic storms and aurorae, and which ones would affect the radio propagation over long distances. Basically, the same kinds of problems, with a few exceptions, that the AWS is now faced with. If we look at the operation as it existed in those days, it could best be described as a sincere effort by the best available people to solve a very difficult problem, with resources which were totally inadequate to the magnitude of the task. The people who were making the forecasts were neither adequately supported nor of the right disposition to tackle this large and complex problem. I speak of the right disposition not in a disparaging sense, but to drive home the point that those involved were primarily research and development people. For many of them this was not their primary job, but they derived a certain amount of satisfaction out of it. They also derived data which was usable in their own more interesting research and development. Gradually, over a period of time, the requirements—the pressures—of the forecasting increased.

Commensurate with this, the amount of data available and the number of customers to be serviced also increased. There were many attempts made to try to make a research type operation the second purpose of operational forecasting. There were attempts to make the observations as intercomparable as possible, and to get twenty-four hour coverage of the sun and the other events which were happening. These attempts were, however, only partially successful.

When the AWS got involved in the early 60's it was a very natural thing (under the circumstances the only reasonable thing) to continue and to extend the same general philosophy of operation as existed at that time. By the time AFCRL got involved, a year or two later, the time was ripe for some rather marked changes in direction. And I should note that AWS contributed many of the ideas and suggestions for changes. These changes of direction can be summarized best by saying that a completely different order of organization was necessary. At first, our thoughts on organization extended naively to the question of putting the observations together, carefully and with adequate corrections, so that they could be used and processed to yield the basic forecasts which were required.

Most of our efforts during the early years were concerned with acquiring a good base of observations, with massaging the observations to make them intercomparable, and with combining the different types of observations in such a way that the maximum amount of information could be sifted out for use in forecasting. After all, there had been many years of effort directed toward making international observations directly intercomparable. Much of it had gone into designing the procedures, or the rules by which observations were to be made, the formats by which they were to be reported, and the procedures to be followed in reporting them. It was natural to assume that not too much effort would be necessary in accumulating and utilizing these observations, to which much apparent care had been devoted. We decided that we would accumulate the data from one full sunspot cycle, extending from the sunspot minimum in 1955 through the following sunspot minimum in 1964. This also included the international geophysical years of 1957 and 1958, around sunspot maximum. Lots of effort was expended in the accumulation and checking of this data.

At this point, it should be noted that some of

our most serious problems were the direct result of the care devoted to the international organization of the observations. If this seems surprising, one example might serve as illustration. The International Astronomical Union had set up rules for making observations of solar flares. But these rules were constantly being modified. Observations from 1955 did not conform to those in 1958, and neither conformed to those of 1964. This would have been sufficiently troublesome, except for the fact that when the rules changed, not all observatories changed with them. So that at any point in time, two or three different sets of rules were being followed simultaneously. And a few observatories knew "better" and followed their own private rules—which of course changed as observers changed. I think you can see from this that the attempts to produce a certain amount of organization of the data were in fact working in just the opposite direction. Despite the difficulties, we charged ahead, trying to make enough sense out of the data to determine when changes were made at different observatories, which observatories were following which rules, and what had to be done to make all the data intercomparable. I'd like to tell you that now, some five years later, we had finished the task, but this wouldn't be true. By and large, we know what's happened, we know which data are usable, and certainly the large bulk of the data has been made intercomparable, but other large fractions of it have had to be discarded because there seemed no way, at reasonable cost, to make the data intercomparable.

If the previous data were difficult and required much effort, what then about the present data? And now I'm thinking about the present in terms of 1966. There were more than fifty observatories around the world making observations of solar flares. There were lots of reports. We had close to one hundred thousand flare reports in the ten year period from 1955 to 1964. And had the reporting been up to current standards, there probably would have been twice this number. But even in 1966, a careful look at the data, that is, the flare reports themselves, would have led to the conclusion that there was more than one sun. When I say that the data didn't match from different observatories supposedly watching the sun at the same time, I mean the mismatch was so bad in many cases, as to make one wonder which aspects of the observations were wrong. Was the time, the day, the month, even the year wrong? Or were the measurements themselves grossly in

error? And yet, with few exceptions, every observatory sincerely claimed that *its data were right*. Now if you have data which do not match, but are claimed to be error-free, what do you do? Our reaction at this point was to talk with the experts. More than a year was consumed in meetings and discussions, without any real agreement on who was right, and without agreement on what it was that was producing the differences. At this point, we must have had a list of fifty or a hundred possible reasons for the discrepant results. If we were to sit down and make a reasonable analysis of these sources of error, and try to determine the magnitude of their effect in producing discrepancies, we would have had a job which probably would last for at least one more full sunspot cycle.

So we decided instead to run a test with two telescopes and two observers side by side, prevent them from talking, and try to see what kinds of results they get. But we weren't going to pick just any two telescopes, because during our frustrating years it became blatantly obvious that, despite the existence of many dozens of telescopes dedicated to observations of the sun, there were no two which matched in their detailed characteristics. We decided that we would construct the first matched pair of solar telescopes, matched as closely as we felt was reasonable within the existing state of the telescope art. Further, that we would outfit them with identical accessories, put them side by side, and try to make the first matched solar observations under controlled conditions. We are optimistic at the present time that we will get data to test the effects of different observing procedures, and of varying accessory equipment. Almost one year was spent in getting agreement on how close we could match telescopes for any reasonable price, and what were the important criteria to hold constant. Right now, we have two telescopes (with accessories) under fabrication, and we hope to start the test by the end of this summer.

You might ask what we expect to get out of such a test. As a minimum we expect to get data which matches to a reasonable degree of accuracy, and if we do, we intend to press for international standardization on a matched set of instruments. We have arranged these instruments such that we will have electronic equipment to help make the observations, to extract the measurements, and to prepare the data for reporting over the teletype. We expect to have the instruments set up side by side to check out the telescopes and to

FEATURES OF THE NEW CODES

FOR OBSERVING

1. Encoding simplified (Data groups arranged in same order as observation).
2. Messages can be truncated for rapid reporting.
3. Little or no computations required.

FOR COMMUNICATIONS

1. Message identification simplified (Standard Formats).
2. Message loss minimized (Hallmarks).

FOR PROGRAMMING

1. Decoding and error checking programs simplified (Hallmarks).
2. Correction and deletion of data made easier (Status Code & Serial No.).

FOR FORECASTING

1. Data type readily identified by code name.
2. Scanning of key data groups simplified.
3. Additional descriptive information included in new codes.

EXAMPLE OF FLARE FORMAT

FLARE

BKSTA YMDD 3//nn

11111 gsJJJ SSmmL QXXYY 5IBCC MMmmL 7AAAA EEmmL 99999

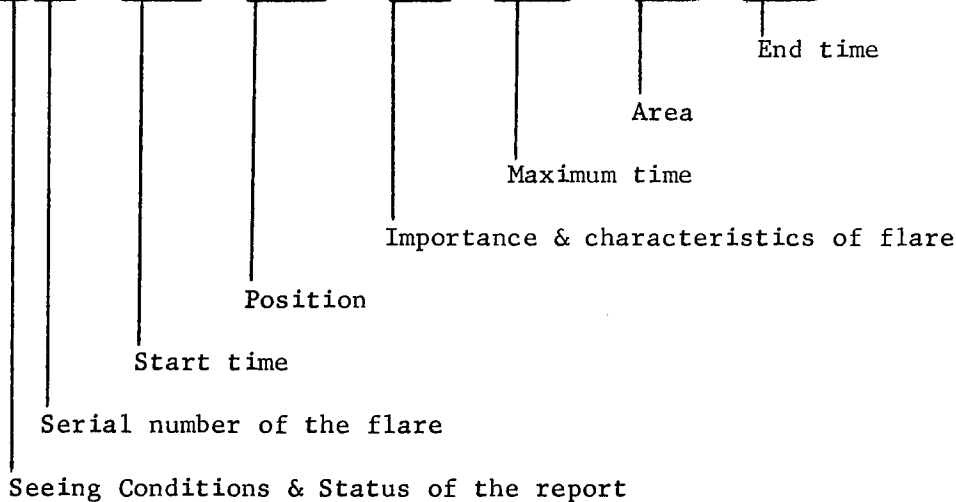


FIGURE VIII-22.

insure that differences in seeing conditions do not produce discrepancies. If this part of the test is successful, then we plan to move one of the telescopes a thousand miles away into completely different seeing conditions, and determine the effect of changes of seeing conditions on the matching—or mismatching—of the observations.

Our hope is that by the end of this calendar year we will have been sufficiently successful so that we can recommend to the AWS the specifications for a set of telescopes to completely reequip the AWS solar observatories around the world. We expect as a by-product of this to determine what one should try to get out of solar observations, so that

other observatories around the world (not part of the AWS) can make observations toward a common end.

Our second problem, which is quite different from the first, but still related to the forecasting problem, is in the transmission of the observations from where they are made, to the place where they are used, the AWS Solar Forecast Center in Colorado Springs. An idea of the problems involved can be gleaned by considering just the optical observations coming in to the Forecast Center from five to ten sites around the world. We studied the various sources of errors in the observations as received, the types of errors, who was making them, and what has to be done in order to correct them. In this case I'm not talking about situations where the teletype itself garbles messages, because these errors basically are not controllable. What I am talking about are the problems in the errors that are made at the observing sites themselves in the recording and sending of the observations. Now if you take a look, as we did, at the data written down by the observer, there are lots of errors. There had been rules set down by the international community on the way that observations are to be sent, so that other people can use them. The codes that were prescribed by the international community, codes in the sense of prescribing the order, the format, etc., can best be described by noting that they apparently were designed to produce a maximum number of errors. You have all had experience with procedures which were set up to reduce the chance of an error by a human being to an absolute minimum, and you've seen some which allowed lots of errors to happen, and you've probably also seen them set up so that the likelihood of an error is actually enhanced. The third situation (the worst of all) was the one which existed with regard to the international codes back in 1966. Our analysis showed that if we were going to improve forecasting, such obvious sources of error must be corrected. So we redesigned the codes with two things in mind. We tried to minimize the errors in recording what the observer sees, and we tried to arrange the codes so that when errors did creep in we would be able to check and correct them with a minimum amount of work, and with a maximum likelihood of reconstructing the actual observation. Over the last couple of years we have redesigned almost all of the international codes with these two things in mind. We have reasonable acceptance amongst the observatories that report in real-time, and we are

optimistic that at some point in the distant future these will become the accepted international standard.

I think just one example would suffice to show the difference between the old and new codes, one of the basic difficulties in the old reporting codes and the general philosophical approach that we took reporting codes and the general philosophical approach that we took with the new codes. The new code is shown in Figure VIII-22. There were a number of changes made, but I'm only going to concentrate on one particular change, which was the source of the largest fraction of the errors in the old solar flare reports. There are certain bulletin headings for the teletype to specify that this is a flare, where it's from, and data, etc. I'm primarily concerned with the long line, the observation itself. That line of observation has been laid out so that the flare *report* follows the flare *development*. The first information available is that a flare has started at a certain point in time, in a particular place on the sun. The first two (five-letter) words in the code given the start time (SSmmL) and the flare position. The next item of information which is available to the observer watching this flare is that it does reach a maximum at some time, therefore the next item of intelligence in the code is its importance (5IBCC), the time of maximum (MMmmL), and its size (7AAAA). Continuing along in the process of the flare during the slow decline, the observer can continue to watch it and then record the last time he can see it (EEmmL), end the bulletin, and transmit the data. In all cases, he has been recording the information in the same order in which it became available. I imagine at this point you are wondering what the old code could have looked like. Without going into all of the details, I can point out one particular difference, which was the source of a large number of errors. The old code recorded the start time first, in its logical sequence, and then the end time, followed by the time of maximum and the associated data.

I'm sure you've all guessed by now that in many cases the maximum time was recorded where the end time should be and vice versa—and you're right! We've done an analysis of the number of errors in some flare reports for the period 1955 through 1964, and these are shown in Figure VIII-23. You'll note that we went over two particular sources of published flare data, those published by what used to be the National Bureau of Standards in Boulder in the left-hand column

Number of Erroneous Flare Reports (excluding sub-flares) in Flare Data (1955-1964)		
Erroneous Entry	Discrepancies	
	CRPL-PART B (1955-1964)	QB (1955-1962)
Date	8	1
Station Identifier	68	13
Time	2604	158
Position	570	219
Importance	335	10
Areas	1475	3128
Total Discrepancies	5060	3529
Erroneous Reports	4440 (12.5%)	3506 (11.2%)

Figure VIII-23.

(CRPL stood for the Central Radio Propagation Laboratory), and on the right is the International Astronomical Union Quarterly Bulletin on Solar Activity. The CRPL report is put out within a few months after the observations. The Quarterly Bulletin comes out a few years later, and hence there's time to make many more corrections. You'll notice in the CRPL reports that the largest source of error is in the time groups, and the major contribution to this was the transposition of the maximum and end times in the reports. In the Quarterly Bulletin someone has apparently taken care of most of the transposed times.

There are however, a lot of discrepancies in the areas in the Quarterly Bulletin which do not appear in the CRPL-I' series. Most of these errors occurred because the old code required the observatories to report areas twice. The first was the measured area, that is, the area as the observer saw it, and the other area was corrected for the geometrical foreshortening near the limb of the sun. An area will appear to look smaller and smaller as it moves closer and closer to the limb. All of the problems that one imagines could happen—did happen. There were a lot of mistakes made in arithmetic, the measured areas were reported in the place where the corrected areas should be, and vice versa. And last, but not least, every observatory had its own way of making the corrections, and many changed their procedures from time to time. In the new code there is only one space for an area measurement and that is the area as the observer sees it. At least we'll know what he saw, and we can make any and all corrections, in any form or forms as required. The bottom part of the slide shows the disastrous results of all of these sloppy reporting procedures, that is, over ten percent of the data with important discrepancies. I don't want to belabor the problem of reporting codes too much, but *all* errors tend to degrade the forecast accuracy, and the forecast accuracy will rise in proportion to the rate at which any noise is reduced.

The third group of problems that I'd like to talk about has to do with the combination of reports from different observatories of the same event and from the same observatories of the same event with different kinds of data. As an example, consider the problem of determining the history of an active region; how the sunspot has been growing, whether there have been new spots born, how the pattern of spots has been changing, how the configuration and strength of the magnetic

fields on the face of the sun has been changing, what the radio brightness temperature in the chromosphere and corona above the region is, and how they've been changing over the past few hours or few days. We've found that all of this information is of predictive value, but only when organized in a particular form. If you read a textbook on the sun, there is never any doubt in the author's mind but that a particular sunspot region has certain phenomena associated with it. This is the textbook picture; the real world is different! As just one example, five or ten percent of the flare observations are not obviously associated with *one* particular region on the disk, but occur between *two* regions, and you're not sure whether it should be associated with the first region or the second. In many cases we find upon analysis that there is an *error* in the position report, that there is a certain amount of uncertainty in the position measurements, because flares are a very complicated phenomenon and their exact position is difficult to specify. But even allowing for all of these discrepancies, there are still many reports confirmed by a number of observatories which place a flare, sometimes a very large flare, in a position which is not easily associated with just one active region. Therefore, when we talk about an active region life history, we're talking about a somewhat arbitrary grouping of data. Historically, this association of flares, spots, plages, magnetic configurations, radio brightness temperatures, etc., was done by hand. Which would have been alright if two experts doing the same thing came up with the same results. Alas, this was not always the case. There were difficulties when a sunspot group was spread out over a large area, Was it one sunspot group or two? If you combine such difficulties with the difficulties of associating each flare with a particular active region, you can see that there are a large number of ambiguous associations. As a result, the statistics of active regions are dependent to a certain extent on the person who associated all of the relevant data.

If we're going to develop forecasting schemes, the forecast values and their accuracies are going to depend on the accuracy of the data association. Our forecasting schemes are going to be dependent on how we put the data together, a job for which science and technology are not going to be very of much help. Our objective has been to devise an objective scheme, which is reproducible, and easily modifiable. We ran a number of tests to find the magnitude of the problems, and the ex-

tent to which a set of relatively simple rules would work. The big advantages of such a scheme are its independence of the person using it, or his prejudices. It can be undone and redone if a better method is developed. It can be programmed for a computer; be done faster and cheaper. Here again I should point out that we're still not talking about forecasting, but about making changes that are bound to improve the accuracy of the forecasts, because we have the flexibility to check to see which way gives us the best final product.

At this point I'd like to get into the center of the forecasting problem, namely the development of improved forecasting techniques. While this may be the direct approach to improving the forecasts, there is always the question of whether a given amount of work directly in the area of forecasting will yield as much improvement in forecast accuracy as the same amount of effort in any or all of the three areas that I've talked about so far. There are certain basic considerations when dealing with the space forecasting problem. The most important is that the present state of our knowledge precludes forecasting on the basis of the physical interrelationships, or the dynamics and magnetohydrodynamics of the sun's atmosphere. In such a situation, one resorts to statistics. Not blind statistics, because our knowledge of the physics and dynamics is not zero, but statistics tempered with a degree of physical insight.

You all know that there are lots of very powerful statistical procedures for forecasting. We planned, for example, to use various forms of multiple regression analysis in order to filter the best forecasting techniques from the available data. But these statistical methods work best when the data is well standardized and very consistent over long periods of time. To the extent that these two criteria are not fulfilled, the methods are less and less useful in filtering out the necessary information. One of the first things that we looked into was the consistency of the data, the consistency of observations from different places at the same times. We were trying to determine the length of the data samples necessary to get meaningful results. Our studies showed that a minimum of some many years of data on a daily basis are necessary for the statistics to settle down to a point at which meaningful conclusions can be drawn from them. We were then concerned with what types of data and what samples of data were sufficiently consistent over these lengths of time so that they could be subjected to a meaningful statistical analysis, and

what other data, with adequate massaging, could be made to conform at least roughly to the required standards. We spent a considerable amount of effort in removing just the basic errors. We were left with about a dozen possible parameters which could be tested in combination to determine their utility in just one of the required forecasting problems, namely that of forecasting the probability of occurrence of a solar flare during the following twenty-four hours. The predictor data included the obvious parameters such as sunspots, plages (the bright areas around the spots), the past history of flares in that sunspot region, the radio brightness temperatures in the chromosphere and corona above it, the magnetic configurations, etc.

We hoped to analyze all of the data simultaneously and get a reasonably fast answer from the filtering of the maximum amount of data. This approach didn't work because there was no firm basis on which to make an a priori judgment about the form of the relationship. There were many possibilities, and this necessitated a certain amount of ground-work. The final approach that we made to this problem was a piecemeal one, with a lot of time spent on single parameters to determine the form of the relationship. After each parameter was analyzed, it was incorporated in a stepwise fashion into the analysis.

A very interesting side result came out of this procedure. Despite thirty years of observations, there had been no objective determination of the frequency of occurrence of flares per unit of time—a vital ingredient in the forecast. This isn't to say that there weren't a lot of ideas as to what probability of occurrence was, but the cold hard fact that nobody had ever calculated it. I'm not implying that this could be done easily; it involved some very interesting problems. For example, if an observatory does not say it has a flare at a particular point in time, does that mean that there is in fact no flare on the sun, or does it mean that the observatory wasn't looking? There isn't always an observatory looking at the sun, and some observatories did not report when they were not looking. One of our major observatories, in particular, didn't feel that there was any intelligence in the information about whether they were looking or not. When it was hinted that their flare reports were of little value without observational status information, there was considerable agitation. But it is absolutely necessary for determining the probability of flare occurrence to know how many in-

dependent flare events there were, and how much time the sun was under observation.

Even with this data, the problem isn't exactly straightforward, because flares last a finite amount of time. Therefore, a complete catalog of the larger flares can be compiled from observations made once every half hour. On the other hand, half-hourly peeks at the sun will miss most of the smallest flares. So that there were a number of adjustments necessary to modify the reported hours during which observatories were looking at the sun. The calculations have now been done in a reasonably accurate way, and for the first time we can state that the probability of a flare is a certain number. The effect of this determination was very interesting. The Solar Forecast Center had been over-forecasting the flare probability by a factor of between two and four; that is, when they were forecasting a sixty percent chance, they should have been forecasting fifteen-thirty percent. When informed of the results of this study, they adjusted all of their forecasting probabilities downward to the factors indicated by our study. After not too many weeks they noticed that their forecast values were too low by something approaching a factor of two. A crisis of confidence developed; we were pretty confident that our data was not faulty, and they were pretty confident that their forecasts were not coming out the way they should have been. Many more weeks were consumed before we figured out what the difficulty was. Our data were right, and they were also under-forecasting. They were not forecasting for every single active region on the sun; the very smallest, the least active ones were ignored because nobody was concerned about them. They had been preselecting active regions on the basis of certain other criteria, namely, the radio brightness temperature in the corona above the region. If you're still with me and you've had a bit of elementary statistics, you'll know that pre-selection on the basis of a parameter, which in itself adds independent predictive information, affects the probability markedly. As long as the Forecast Center continued to pre-select on the basis of radio brightness temperatures, there was an upward adjustment necessary to all of our probability figures. The very next parameter that we put into our forecasting scheme was just this radio brightness temperature. It did turn out to be of predictive value, and took into account the pre-selection criteria which were, and are, routinely used at the Solar Forecast Center. I'd like to em-

phasize a point which is in some ways humorous, and in other ways discouraging; that is, the adjustment in the flare probability forecast values by this elementary probability computation has made the largest increase in forecast accuracy of anything else that we have done in the interim, or probably will do in the next five or ten years. It's very unlikely that the state-of-the-art will allow us to make any similar quantum jumps in forecast accuracy equivalent to this very simple and very basic change.

Before we leave the problem of the development of forecasting schemes, I'd like to point out a by-product. There are a lot of people who are skeptical about all of the money that's spent in collecting what appears to them to be very redundant data, and it's very tempting under such circumstances to say "you really don't need all of it". So we started looking very carefully in our analysis not only to develop forecasting schemes, but in the process to determine the relative values of different kinds of data in combination with other forms of data. Fortunately, the multiple regression schemes are very well suited to both purposes. I think it would be reasonable to say that our value to the Forecast Center has been only partially in the generation of forecasting schemes. We have been of at least equal value in providing the necessary groundwork, direction, and justification for them to continue to get certain kinds of observations and to expand their observational capability in many areas. This is another case where we have made contributions to the accuracy of their forecasts in an indirect way, in this case by enabling them to retain certain kinds of observations which would otherwise be discontinued or to expand their observational capabilities in a number of different directions. These additional data result directly in more accurate forecasts.

The last special topic that I want to discuss is the question of forecast verification. There are at least three reasons for pursuing a vigorous verification program. First, there must be as many crackpots in the space-weather forecasting business as there are serious workers. These range from the types who always seem to have the answer, down to people who are in reputable positions and who base their forecasts on apparently careful analysis, but who never temper their statistics with the slightest bit of common sense. In a government organization, crackpots can be especially troublesome, because they have congressmen too! And frustrated crackpots can be the most indignant

people in the entire universe. Nothing quiets a crackpot faster than a test of his forecast accuracy. The second reason for a verification program is to determine what the present state-of-the-art is. It's also very instructive, however, from a third point of view in that it does hint at the direction of future improvements.

SUMMARY

In summary, I'd like to restate my original point—the best way to improve the accuracy of many forecasts isn't always to improve the forecasting techniques. The overall accuracy and usefulness of a forecast is determined by many factors, only some of which can be manipulated at any particular time. In the domain of forecasting solar events and their effects on the earth, we found that the subjective forecasting *techniques* already in

use were not so bad, and that some simple modifications would raise them to temporarily acceptable levels. By contrast, many other factors, directly influencing the forecast accuracy in a major and deleterious way, needed drastic revision. Changes were necessary in the basic observational equipment, in the procedures for recording and transmitting the data, in processing and reduction, in the ways in which they are applied to the analysis, forecasting and verification, and in their treatment and preparation for archiving. Every one of these factors directly directly contributes either to the present forecast accuracy, or to the groundwork for future improvements.

It's been said many times that a good observation is the best short-range forecast, and in too many instances it's true. The reverse is even more true; that is, a bad observation is not only a bad short-range forecast, it generates bad long-range forecasts, and absolutely guarantees that the forecasts will not improve.

DISCUSSION FOLLOWING PAPER GIVEN BY DR. FREDERICK WARD, AIR FORCE CAMBRIDGE RESEARCH LABORATORIES "FORECASTING SOLAR EVENTS"

DR. RUMMEL: Obviously you are working in a poorly developed area.

DR. WARD: This is the physical sciences.

DR. RUMMEL: Let me say that we in international relations have often heard about our field from physicists and I am happy at this opportunity to reverse roles. The problem of errors in data, especially if one intends to use various forms of regression analysis, has been studied extensively in the mathematical statistical literature. A number of ways have been developed to get approximations to a least squares solution, given that you have errorful data, by either grouping techniques or the use of control variables. I am not sure whether in your situation these kinds of techniques might be useful or not. Recently a book by John R. Wolberg called *Prediction Analysis* has been published. It tries to allow you to estimate regression results from a knowledge or approximate knowledge of the error variance around an observation. This may fit your situation in which you have redundant data on the same event, from which you could calculate the variance of the particular observation. From these variances you could try to get estimates of your parameters.

DR. WARD: I guess my answer goes in two parts. Number one: there are ways of getting around a lot of errors in the data, and the first thing we decided was that we would not spend too much time doing this because it would not cure the basic problem. If we did not do something about the observations, two or three years from now, we would still have the same problems. Secondly, the errors in these data are very peculiar. Some of them are random errors, but most are not. And for the non-random ones we do not even know the direction of the bias in most cases. If you try to use the simplest procedures in the regression analysis to try to fix these up, you get into more difficulties than you get around. It would be all right if we did not know very much and we could assume the data errors were unbiased with respect to the observations. But it is a little difficult when you know that the error is badly biased. For that reason and because we have a limited amount of effort, we have concentrated basically in the other areas and not in the forecasting techniques. I cannot justify this approach in the sense of knowing that out of a number of possible courses, this was the best one. But I do know that the existing situation is never

going to improve unless we do something about it. You look at this problem and you say to yourself, "What should I do?" The conclusion we came to (I hope correctly) was that we should make a big push to fix up the observations. Right now we have our fingers crossed that we will get good observations.

MR. IRWIN: I was very interested by the different approach which you and Rudy used. You both had a body of data which were inaccurate. You took the approach of improving the data, whereas Rudy used the data as they stood and manipulated them mathematically. I am not disagreeing with the point about cleaning up the data. My particular field is trying to get better migration data. If Rudy thinks that only foreign statistics are bad, he should examine the migration statistics I am working on. I certainly agree about the necessity of fixing up the data, but the question I ask is: could the mathematical transformations which Rudy used and which he felt dealt with the errors of data be used in your field, or is your field so different that they are not usable?

DR. WARD: Some of the things which Rudy has used could be used. I do not want to give the impression that we could not make progress by fixing up the errors—that is, fixing up the errors after the fact.

MR. IRWIN: But it does not seem to me that Rudy fixed up the error. He used the data as is, but treated them in special ways.

DR. WARD: I meant to get around the error problem in the analysis.

MR. IRWIN: That is the route I take. I try to take the numbers and fix them up. Rudy did not do that.

DR. WARD: But Rudy did what he had to do, in that he could not fix up the data. He cannot go to these countries and tell them to make better data.

MR. IRWIN: I think you can try to make an estimate. For instance, you think the population count is incomplete. You can try to estimate the degree of incompleteness and boot up the number. But Rudy did not do that, right? At least, I did not get that impression from your talk.

DR. WARD: I suspect he did a certain amount of it.

DR. RUMMEL: I did a bit of error analysis on population for my analyses. I was able to estimate from UN sources an error margin within which we felt the population for each country would lie. But once I had that, I was faced with the problem of what to do with it since we did not know in which way the error was within the estimated range.

MR. IRWIN: You analyzed the data to see how it was incorrect. So I make the same point: that you did use the bad numbers and then used mathematical techniques to try and deal with this. I am just wondering whether that works.

DR. WARD: We have used the bad numbers in all of the routines which we have developed and as far as possible with a reasonable amount of effort, we have tried to minimize the effect of the bad numbers on the result. As Rudy points out, you can always do this, spending more or less time on it, depending on what other things you have to do. I think if you are in a position (which neither you nor Rudy are in) to actually get in and ensure that the data are fixed up from now on, you have to make a very serious judgment about how to expend your effort. And when I say expend effort, this is a matter of going back to lots of people and having lots of discussions. It was less a technical problem that a human problem: trying to decide in what direction you should finally move. But I think almost anybody in these circumstances, where they have the option of doing something about the measurements, is sorely tempted to take that attitude: "Damn it, I have got a certain amount of effort and even if I can only get a little bit out of the present data, I am going to fix it, so that the next guy that comes along will have the kind of data that he needs." It is a touchy problem.

MR. IRWIN: I really don't argue that point. Let me put the question in a different way. It would seem to me from listening that you would be in a position now to use the previous data and do better forecasts than are being done—I'm not commenting on the point of fixing up the differences today—I'm trying to do that in the '70 census—I'm trying to get more detail in migration statistics so they will be better and get better tabulations—but couldn't you, as you stand there now, using these techniques perhaps go after the forecasts that are now made using the past data?

DR. WARD: The answer is yes, and it is just a question: which way do you want to go? The Air Weather Service has to make the new observations and so if we invest time now, in getting the Air Weather Service in the position where they can start to make good observations, then it is their problem. We can then sit back and say, "Now we have the time," to either use good new data or fix up and use some of the old data. So it is the order in which you want to approach things.

MR. FRIEDMAN: The question I have is not a

question concerning the statistical treatment of the data. But rather let us suppose that all your observations and all the data is completely accurate. What would you have?

DR. WARD: What would you have in the way of accuracy?

MR. FRIEDMAN: Weather forecast data? If all your data was absolutely reliable, what would you have?

DR. WARD: What you are asking is a question: effectively: what is the present state of the art?

MR. FRIEDMAN: What art?

DR. WARD: The art of forecasting solar and geophysical events.

MR. FRIEDMAN: If you were able to forecast that accurately, what would you have?

DR. WARD: If you were able to forecast it or the observations accurately? It is two different things. . .

MR. FRIEDMAN: If you could observe it accurately, would you then get a proper forecast?

DR. WARD: There is no way to make hundred percent perfect forecasts. The accuracies are at some intermediate level. The exact value depends on the particular forecast. What we are striving for is to determine what the state of the art is; given perfectly accurate data from all the things that we know: how well can we do. This, I think, is what most people would define as the state of the art. This has not yet been determined, and that is what we have been trying to do because it is important. The people who want forecasts concerning radiation, density changes and things like that say, "We will use your forecast, but only if it is good." They ask, "how good is it?" It is difficult at this stage to put a number on it. I think it is very important that you be able to put a number on it; and, more important than that, that you be able to estimate the increased accuracy if you were able to get all good data. Then you can make a judgment as to how much time and money you should expend fixing up the data.

MR. FRIEDMAN: If I were a pilot in the Air Force and I was looking to you for certain kinds of data, what kind of data would I get from you? Weather forecast?

DR. WARD: I am only concerned with things above a certain level, say, sixty kilometers and above. That excludes weather.

MR. FRIEDMAN: Who uses the data? I am not really trying to pin you to the wall.

DR. WARD: I am trying to avoid certain classified problems. What I was going to say was: imagine

a person in the position of deciding to commit a certain number of planes to an operation. Getting these planes up and flying them around costs a lot of money. In some cases it costs millions of dollars, and a decision has to be made on the basis of whether a peculiar effect was a natural event or whether it was due to somebody else starting something. If you can reduce the number of times when you wastefully send a number of planes out, you save money, to say nothing of the problem you have if you send the planes out and they use up all their fuel and a few hours later you need them but they are out of gas. There are a lot of little problems like that.

DR. OBERBECK: Taking us back to the weather problem: what is the value of precise weather predictions; you look at this and it seems to be either in terms of what does it mean to the operations to be conducted on the basis of knowledge or forecast, or you can look at it simply in terms of providing weather forecasts, which is an object in itself; it is not your problem to worry about what it means to a potential user. What now is your objective? Well, here, what accuracy can you buy with so much money. I think the question you are trying to ask comes to this paradox: what is the value of the forecast? Are you going to answer that question in terms of the users' evaluation of what it means or are you going to look at it in terms of, let's say, an end in itself?

MR. FRIEDMAN: There is some talk now about man's ability, say, in the year two thousand, to manipulate the weather. If man, in fact, is able to do that, will you virtually be out of business as far as weather forecasting is concerned?

DR. WARD: I think as far as the weather is concerned, the meteorologists would be performing a different role because there will always be the question of what you want to do and then if you deal in a certain way, what kind of results you produce. I think the weather forecaster, under the circumstances you have mentioned, would be in the position not of saying what is going to happen if nothing is done, but what will happen if I do or do not do certain types of things. So there would be just that many more meteorologists. Now as far as the space-weather problem is concerned, I do not think the situation is materially different except that in trying to do something with the sun, it is going to take a lot more ingenuity.

MR. BENNETT: First of all, were there two or three or four observatories that quite consistently produced similar or nearly identical variances?

DR. WARD: I think I can state that it depends on exactly what you mean by identical. Out of the observatories which reported in real time, which were the ones we went very avidly after as an experiment, there were no two whose results were generally consistent with one another. There were a hundred others which did not report in real time, but I do know whether by chance a few match. It was almost amazing that out of ten observatories not one matched with any other one.

MR. BENNETT: My second question is getting into regression equations. How would you have handled or how do you handle the intercorrelation in the predicted variances?

DR. WARD: If you do multiple regressions, you have a matrix of all correlations. The final result is the regression equation which says: weight certain predictors with specific constants to yield the forecast, again with all the usual restrictions. This takes into account all of the correlations.

MR. BENNETT: Because of the higher correlation?

DR. WARD: You will end up in a lot of cases with very high correlations. The correlation with the information may be high, and the noise correlation can be anything.

DR. RUMMEL: You have quite a problem if there is correlation among the independent variables. One assumption in carrying out the regression is that the matrix of data on the independent variables is nonsingular. As the matrix approaches singularity, a zero determinant, you get increasing instability in your parameters. So if you have a fairly high degree of correlation among the independent variables, it can produce poor estimates.

DR. WARD: This is a big advantage in having a lot of noise in it. It guarantees a nonsingular mat-

rix. One thing I can think of that the noise does for us: it keeps us in work.

DR. RUMMEL: This may be true. Certainly you have a nonsingular matrix if you have noise in the data. Still, you can also have a degree of intercorrelation among the independent variables which will effect the stability of the regression coefficients.

DR. WARD: But as a practical matter, the degree of intercorrelation has to be very high before a reasonable amount of instability comes in. And the amount of intercorrelation has to be much higher than any that we ever have. You can put in things like the variable itself squared, you know, where both the noise and the information are intercorrelated and the correlation is 0.7, and you still do not get any problems with the result. Actually what we have been using in most cases is a regression where we pick out one variable first, the one with the highest correlation with the predictand. Then we take out the effect of that variable on all the others. This method eliminates certain problems while producing others; but we ran a lot of tests of various and sundry kinds of multiple regression and the results are not materially different, no matter what kind of buggering you do with the data.

DR. RUMMEL: Are you talking about forecast results or about least squares fits to the data?

DR. WARD: We tried both because once you get a large data sample, you end up with the forecast, the fit being almost the same. In fact that is the fastest way to find out how much data you need for a reasonably stable relationship. I will not say all of the possibilities have been tried, but enough have been so that one has a reasonable feeling for how much data is needed.