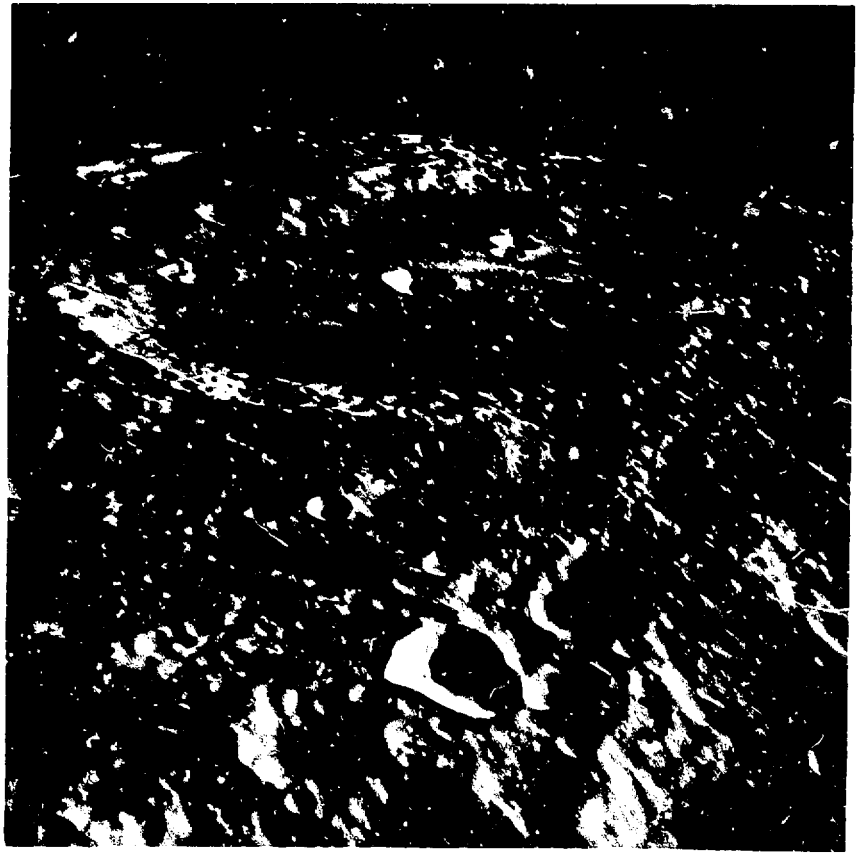
UCRL-52000-76-11

NOV. 1976.

Energy and Technology Review

Prepared for U.S. Energy Research & Development Administration under Contract No. W-7405-Eng-48



LAWRENCE
LIVERMORE
LABORATORY

LUNAR TEMPERATURE STUDY

Printed in the United States of America
Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, Virginia 22161
Price: Printed Copy $; Microfiche $2.25

Page Range	Domestic Price
001-025	$3.50
026-050	$4.00

Distribution Category UC-2

UCRL-52000-76-11
NOVEMBER 1976

Energy and Technology Review

ABOUT THE COVER

Surface features on the far side of the Moon as seen from orbit. Magnetic-field measurements were taken from orbit and on the surface to determine the electrical conductivities of deeply buried layers of rock. However, internal temperatures derived from the conductivities conflicted violently with values from other, usually reliable methods of estimating the subsurface temperatures. Independent of this problem, LLL geophysicists discovered a serious experimental error in former methods of measuring the electrical conductivities of minerals under extreme temperatures and pressures. When their attention was focused on the Moon temperature disagreement, they resolved the discrepancy, shedding some new insights on the structures and histories of both the Moon and the Earth. (NASA photo, Lunar Science Institute, Houston, Texas.)

LAWRENCE LIVERMORE LABORATORY
University of California
Livermore, California 94550

Briefs

The short items on this page announce recent developments of importance. Some of these items may be amplified in future issues; none of this material is reported elsewhere in this issue.

CONSTRUCTION BEGINS ON FUSION RESEARCH FACILITY

Construction has begun at LLL on a facility for studying how high-energy neutrons produced in fusion power reactors may affect vital reactor components, such as vacuum vessels, insulators, and magnets. Known as the Rotating Target Neutron Source (RTNS), this $5 million, 1600-$m^2$ facility is scheduled to begin operation in March 1978. It will be available under ERDA auspices to researchers from around the world for experiments crucial to the engineering design of prototype fusion power plants.

Researchers will expose small samples of metals to 14-MeV neutron bombardment one-fifth as intense as that expected in proposed fusion reactors. (Fusion reactors are expected to produce about 50 trillion 14-MeV neutrons per square centimetre per second.) Metals to be tested range from stainless steel, a structural material, to superconducting alloys of niobium and tin. Future magnetic-fusion power plants are almost certain to require superconducting magnets to minimize electrical consumption.

After exposure, the samples will be examined for weakening, embrittlement, blistering from helium bubble formation, and in the case of the superconductors changes in conductivity. Such effects are important in fusion reactor design because of their potential impact on the longevity of parts.

The RTNS facility will replace a similar but smaller research facility at Livermore that for the past 10 years has provided high-energy neutrons at about one-fiftieth the intensity of proposed fusion reactors. Both facilities are based on a rotating target concept for producing an intense flux of fusion neutrons. A narrow stream of deuterium bombards a high-speed rotating target containing tritium; the D-T reaction produces 14-MeV neutrons. The target spins and wobbles to vary the impact point of the deuterium beam, which otherwise would heat the target enough to release the tritium. The target is water-cooled.

In addition to this neutron source, the new facility will contain apparatus to irradiate reactor materials with charged particles and x rays, allowing researchers to investigate the combined effects of a variety of fusion reaction products.

TAKING THE MOON'S INTERNAL TEMPERATURE

LLL geophysicists have been instrumental in resolving a serious discrepancy between lunar magnetic-field data and melting studies of lunar basalts brought back from the Moon by Apollo astronauts. Estimates of the subsurface temperatures, based on lunar electrical conductivity measurements and laboratory experiments, were hundreds of degrees below those given by models using known melting points of various minerals. Our work uncovered a basic flaw in previous measurements. New measurements under more realistic conditions brought the electrical-conductivity temperature estimates into agreement with temperatures derived from melting experiments. This same work has also contributed to our *in situ* coal gasification studies, to ERDA's dry, hot-rock geothermal effort, and to our program of monitoring for seismic evidence of clandestine nuclear testing.

Is the interior of the Moon hot like the Earth's interior? How does the electrical conductivity of the mineral olivine vary with temperature at pressures to 5 GPa? There is a connection between these two questions; by answering the latter, LLL geophysicists were able to answer the former and resolve a conflict in lunar internal temperature calculations troubling NASA.

The physical properties of rocks and minerals at high temperatures and pressures have long been of interest to the test, seismic monitoring, and energy programs at LLL. Consequently we have extensive capabilities both for producing extremes of pressure and temperature and for measuring these properties. There seemed to be no immediate application for measurements of the electrical properties of rocks, however, and thus support for these measurements received a low priority – although theoretical studies continued.

NASA, at the same time, was facing a dilemma. Among the data brought back from the Moon by the Apollo flights were two apparently contradictory measurements from which lunar subsurface temperatures could be calculated. Estimates based on the melting properties of surface minerals gave a high value; measurements involving the lunar conductivity variation with depth, the solar wind, and the electric conductivities of minerals as measured in the laboratory gave values hundreds of degrees lower.

This was no trivial matter. The subsurface temperature profile is one of the prime pieces of evidence in all theories of how the Moon was formed. If the melting point measurements were right, it would mean revising most of the previous ideas about the Moon's origin. If the electrical conductivity interpretations were right, there would still be the question of how the melting point estimate could be so far off.

Both temperature and pressure increase with depth in the Earth and Moon. It is relatively easy to calculate the pressure at any depth from the densities of the overlying rocks. But temperature is not so easily ascertained. There is no way to probe the region directly; all subsurface temperatures are estimates based on indirect evidence. Temperature is thus one of the major uncertainties in the study of planetary interiors.

NASA's method for inferring the lunar interior temperature involved measuring the magnetic field induced in the Moon by sudden changes in the solar wind. The Moon's remnant field is very weak and varies greatly from place to place. To monitor the magnetic field, the astronauts set up a series of flux-gate magnetometers (Fig. 1) at the various landing sites. Comparing the readings from these magnetometers with those of magnetometers in orbit around the Moon produced a coherent picture of the magnetic field's response to changes in the solar wind. These responses, in turn, could be interpreted to determine the conductivity of the underlying rocks.

Although most rocks are good insulators at room temperature (especially when they are very dry, as they are on the Moon), their electrical conductivities depend strongly on the temperature. In some cases the conductivity may increase more than a billion-fold between room temperature and 1000°C. Hence the electrical conductivity of a dry rock should be a sensitive indicator of its temperature.

This temperature scale must first be calibrated, however, to be useful. For calibration, specimens of the appropriate minerals must be obtained, subjected to extreme temperatures and pressures, and their

Contact Alfred G. Duba (Ext. 7598) for further information on this article.

conductivities measured. To the extent that laboratory conditions duplicate the conditions in the interior of the Moon, such measurements should produce an accurate temperature scale.

According to the best present information, the outer 400 km of the Earth and the bulk of the Moon consist of two minerals: olivine and pyroxene. Olivine predominates in the Earth, pyroxene in the Moon. These minerals contain mainly oxygen, silicon, magnesium, and iron. The major compositional difference between them is that pyroxene contains more silicon, in proportion to its oxygen content, than olivine.

When NASA scientists compared the existing laboratory data on the temperature variation of electrical conductivity for olivine and pyroxene with their data from magnetometer measurements, their calculations showed temperatures hundreds of degrees below those necessary to produce the vast lava flows that fill the Moon's great maria plains. If the Moon was really hot enough to have produced large amounts of molten rock, why did the conductivity measurements indicate such low interior temperatures? Where had the measurements gone wrong?

Meanwhile, independent of the NASA difficulty, LLL geophysicists were beginning to have their doubts about laboratory measurements of electrical conductivity. They thought of examining the assumption that the laboratory conditions exactly duplicated subsurface conditions. If this assumption were faulty, the conductivity measurements might be in error.

The study's first clue resulted from conductivity measurements on a pair of olivine specimens of apparently identical composition. One came from the San Carlos Indian Reservation in Arizona and the other from St. John's Island in the Red Sea. The San Carlos olivine had a thousand times the conductivity of the Red Sea olivine at the same temperature.

Such a discrepancy clearly indicated that the specimens couldn't be truly identical. There had to be some subtle difference to account for the disparity in conductivity.

A careful reanalysis turned up a possible lead. Although both specimens contained the same amount

Fig. 1. The lunar surface magnetometer deployed at the Apollo 16 landing site. The boot prints indicate scale. The golden ribbon is a power and signal cable. This and similar instruments at several other landing sites, together with an orbiting magnetometer aboard Explorer 35, provided a continuous record of changes in the magnetic field on the Moon over many months. (NASA photo, Lunar Science Institute, Houston, Texas.)

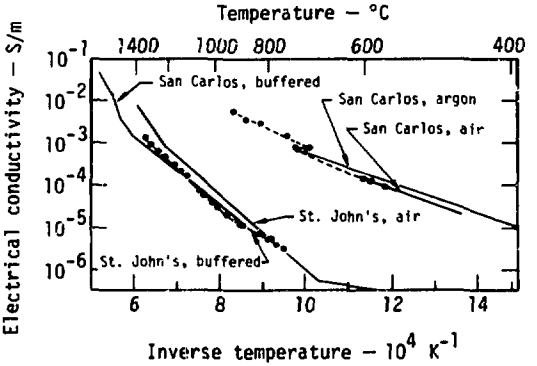


Fig. 2. Electrical conductivity measurements on two specimens of olivine under a variety of circumstances. The specimen from St. John's Island in the Red Sea has a conductivity less than a thousandth that of the specimen from the San Carlos Indian Reservation in Arizona when both are measured in air. Replacing the air with argon had almost no effect on the San Carlos specimen. The conductivities agree only when both are measured in a buffering mixture of carbon dioxide and carbon monoxide. This buffering mixture controls the oxygen fugacity, simulating conditions in the deep interior of the Earth and Moon.

of total iron, about 2% of the iron atoms in the San Carlos olivine were in the ferric oxidation state. The Red Sea olivine contained no trace of ferric iron; all its iron was in the more reduced ferrous state.

Once this difference in the olivines had been detected, it was fairly straightforward to deduce the reason for the conductivity dilemma. The olivine deep underground must contain no ferric iron; the very existence of olivine in planetary interiors at 1000°C is evidence that at that depth the oxygen fugacity (proportional to oxygen's partial pressure and the chemical reactivity of oxygen at a given temperature) is less than one-hundred-millionth of that at the Earth's surface. The San Carlos olivine specimen (and the vast majority of other olivine specimens) must have been exposed to oxygen while it was still hot enough to oxidize. The gem-quality Red Sea olivine, on the other hand, must have cooled below about 600°C before oxygen reached it.

This being the case, it follows that all previous measurements of olivine conductivity (and, by analogy, that of pyroxene) at high temperatures must be in error, because they started with partly oxidized material and because they failed to prevent further oxidation during the measurements. Even measurements conducted in vacuum or in an inert argon atmosphere (which would contain a few parts per million of oxygen as an impurity) would expose the olivine to thousands of times more oxygen than exists 100 km underground.

When word of these results reached NASA scientists, they immediately enlisted our help in unravelling their measurement mystery. The objective of our study was to find a way to simulate the conditions deep beneath the lunar surface closely enough to arrive at meaningful electrical conductivity measurements. Our conductivity measurements were to concentrate on pyroxene, the major constituent of the lunar interior.

As hinted above, simply excluding oxygen wasn't good enough. Going to the other extreme, that is, actively removing the oxygen, as with a hydrogen furnace, would be just as bad. In a hydrogen furnace the oxygen fugacity is about 10^{-35} Pa. Under these conditions oxygen would leave the olivine and metallic iron would appear, completely invalidating the conductivity measurements. Furthermore, the right value of fugacity was not a constant, but changed with the temperature.

The solution was to buffer the oxygen fugacity with a mixture of carbon dioxide and carbon monoxide. With this mixture the oxygen fugacity closely followed the olivine stability interval. Neither ferric nor metallic iron was produced regardless of temperature.

The proof of the pudding came in conductivity measurements on San Carlos olivine using the carbon dioxide/carbon monoxide buffer system. Once the specimen came into equilibrium with the buffer, its conductivity dropped a factor of a thousand and agreed closely with that of the Red Sea olivine measured under the same conditions (Fig. 2).

Naturally, the equipment for making these conductivity measurements works equally well with other materials. In fact, we have measured conductivities for pyroxene, garnet, albite, and basalt.

among others. These rocks and minerals are also present on the Moon and can make minor contributions to the overall conductivity.

When our values for the conductivity of lunar rocks are combined with the NASA magnetometer data, they yield a temperature profile that agrees substantially with the values derived from mineral melting points. Figure 3 compares the resulting selenotherm (a temperature-vs-depth curve for the Moon) with the previous uncertain estimates. The new selenotherm comes close to the solidus, the line representing melting point vs depth, and is consistent with a softening in the velocity-depth profile for moonquake waves observed in recent seismometer data.

This confirmation of a high-temperature interior for the Moon has also revived interest in the mascon problem. A mascon (from "mass concentration") is a gravitational anomaly — any of several regions on the Moon where gravity is measureably higher than on the rest of the surface. Our meager data on the subject indicate that mascons are flat, pancake-shaped regions with densities higher than their surroundings, lying somewhere below the surface of several of the basalt-filled maria. Mascon diameters are up to hundreds of kilometres.

From their shape and their location beneath maria, it appears that each mascon may be the result of an impacting meteor forming a crater that subsequently filled with lava to form the relatively smooth, uncratered surface above. The problem comes in accounting for the details of such a process. Why, for instance, did the mascons stay just below the surface? If we postulate rocks strong enough to support the extra weight of the mascon, this implies temperatures far below the solidus. Where, then, did the lava originate? If we postulate hotter interior conditions, capable of supplying lava, what prevented the mascon from coalescing into a ball and sinking down to the center of the Moon?

The new temperature data furnish a definite base point on which to test future hypotheses of mascon origin. Combined with information on the Moon's cooling rate, it could help to date the events that formed the mascons by placing an upper limit on their ages.

The improved electrical conductivity information and the techniques we developed for measuring conductivities under extreme conditions have several applications here on Earth. The electrical conductivities of weapons materials are important in studying the behavior of these materials under extreme conditions. Our measurements and methods have made significant contributions in this field.

In energy resource development, one of the problems in controlling *in situ* coal gasification is to define the shape and position of the burn front as it proceeds underground. Our studies of the temperature-induced changes in coal's electrical conductivity suggest that electromagnetic probing may prove a feasible way to trace the burn front. Similarly, our data on the conductivities of basalt and such minerals as albite are being used in ERDA's program to map "dry, hot-rock" geothermal resources.

Finally, some of the same considerations that apply to the Moon can be applied to the Earth as well. Figure 4 shows where the geotherm (temperature-vs-depth curve for the Earth) calculated from our

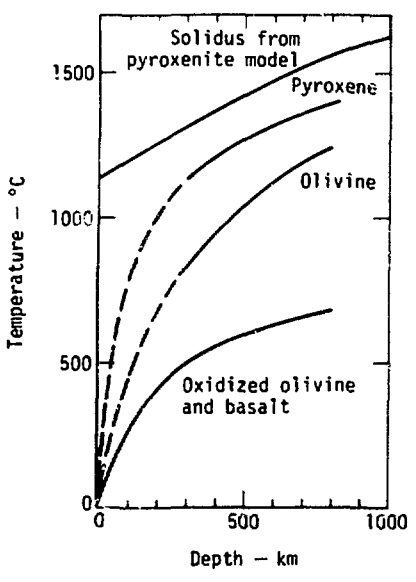


Fig. 3. Selenotherms (temperature-vs-depth curves for the Moon). The olivine and pyroxene curves are calculated from our laboratory conductivity measurements and NASA conductivity-vs-depth profiles. The previous olivine and basalt selenotherm was based on faulty electrical conductivity measurements. The pyroxene curve comes within about 100°C of the solidus, the variation of melting point with depth. Uncertainties in the data interpretation amount to about ±150-200°C.

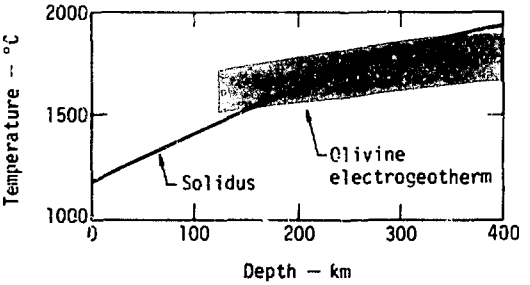


Fig. 4. The variation of temperature with depth in the Earth, calculated from our data on olivine electrical conductivity and on conductivity-vs-depth data from a variety of sources. The solidus is shown for comparison. Interstitial water interferes with the measurements down to about 320 km; only below that depth is the conductivity a reliable indication of temperature.

laboratory data on olivine conductivity and from conductivity-depth profiles obtained in many ways lies in relation to the solidus. This information improves our interpretation of seismic velocities in the Earth by refining our estimates of the states of the materials likely to be present. Such interpretation is of vital importance to the seismic monitoring program by helping us to verify compliance with treaties limiting the size of nuclear tests.

Key Words: albite, conductivity, basalt, electrical conductivity, effects of temperature, geothermal energy, exploration, moon exploration, olivine, pyroxenes.

RECOGNIZING THE PATTERN OF CRIME

Under ERDA's technology utilization program, LLL is helping the San Diego Police Department explore the usefulness of automated crime analysis. Using an LLL computer code, we have demonstrated experimentally that computer pattern recognition is a feasible method for analyzing crime. It now remains for us to complete the technology transfer cycle by assisting in the transition from the experimental system to an operational one. Then, the San Diego police can use it as a time-saving, cost-effective tool for allocating the department's available manpower more effectively to solve and prevent crime.

Police departments today are feeling the economic pinch of having to fight increasing crime with limited manpower. As a result, they are looking for ways to automate time-consuming administrative functions performed by their officers.

The Federal Law Enforcement Assistance Administration, in recognizing this problem, last year gave the San Diego Police Department a grant to construct an Automated Regional Justice Information System for use throughout San Diego County. One of the goals of this system is the implementation of an automated crime analysis capability.

"Crime analysis" is not a precisely defined or well understood term, but is generally interpreted to be the extraction of useful information from the study of historical crime data. Crime analysis techniques used today range from simple "brainstorming sessions" between senior police personnel to extensive statistical analyses of crime reports. To date, no technique has proved completely satisfactory, although occasional successes have been encouraging.

Because of the data-analysis problems of several LLL research projects, a computer pattern recognition capability has been developed at the Laboratory. The potential demonstrated by pattern recognition techniques in various physical and social sciences suggested that substantial advantages might be realized through applying these methods to crime analysis (see box on p. 6). LLL was requested to participate in a technology transfer effort on crime analysis with the City of San Diego.

Up to that time, little or no pattern recognition analysis had been done with real crime data. Our participation was thus to evaluate the capabilities of pattern recognition methods to perform crime analysis.

Contact Lyle A. Cox, Jr. (Ext. 3654) for further information on this article.

A brief study of the nature of the problem was performed, and the planned project was broken down into the following basic phases: feasibility, operational potential survey, extended variable selection, and algorithm choice.

Feasibility

The first phase – finding out if pattern recognition would work in crime analysis – was experimental. Our goal was to determine if some relationships existed that would allow us to predict the likelihood of a crime being solved based on certain variables: nature of crime, location, time, etc.[1]

Our first task was to review the crime statistics provided by the San Diego Police Department. They covered the first quarter of 1975 and were in the form of a magnetic tape containing 21 056 lines of data. Each line represented one crime report and its related variables, or information fields. After rejecting lines with errors or uncertainties, we were left with 11 645 case records. From these records we were able to summarize statistically the different crimes totaling 26 and ranging from petty theft to homicide.

This review gave us a better understanding of the data we would have to deal with in searching for relationships which would allow us to predict that a particular crime could be solved. Information fields from the San Diego Police Department tape and their reformatted structure for PATTER use (see box on p. 9) are given in Table 1. Note that, in reformatting, we assigned one of three properties to each case: closed, open, or material recovered. For our purposes, we considered the San Diego Police term "closed" to mean solved.

Additionally, the data were displayed graphically in a number of ways to facilitate visual interpretation. One display used digitized crime locations to produce a three-dimensional plot of crime frequency against an outline of the city itself (Fig. 5).

Initially, we used PATTER to find out how efficient various techniques were in predicting susceptibility to solution of newly reported crimes. We based this

COMPUTER PATTERN RECOGNITION

Computer pattern recognition combines computer science with the problem-solving techniques of applied mathematics. Over twenty years ago, as the first digital computers became available, researchers realized that their usefulness was not confined to numerical calculations but that computers could play a much broader information-processing role. Some of the first programs in this area were designed to automate decision-making processes, such as recognizing printed alpha-numeric characters. From these early efforts, the field of pattern recognition has evolved.

Some patterns may be physically representable, such as printed letters on a page or the measured echoes of sonar and radar units. Other patterns may be only abstractions, such as patterns discerned from psychological or social data. Regardless of their type, it is by recognizing patterns that we interpret the world about us. Consider how a young child learns to differentiate between mother and father. He learns to interpret the patterns of his senses and to make decisions based upon his interpretations. Clearly the child's learning process requires two steps: (1) observing patterns and determining a decision rule and (2) applying this rule.

Similarly, computers can be programmed to recognize patterns by first representing a set of data (e.g., crimes) of known properties (e.g., solved or unsolved) as a training set. A similar set of data whose properties are unknown are then analyzed to see if they fit the pattern of the training set. This allows the scientist to find or predict a property of the variables that cannot be measured directly but is thought to be related to the variable by some unknown connection.

Recently, amid discussions of the "information explosion," the need for automated information systems became clear. The emergence of such systems accentuated another problem – that of making intelligent decisions based on the vast quantities of data available through these systems. Using computerized pattern recognition techniques, we are able to consider large amounts of data in relatively short periods of time and to extract and apply the desired information efficiently.

Pattern recognition techniques today are used for solving data-processing problems in an increasing number of diverse areas: optical character recognition, speech recognition, spectrographic analysis, system design, weather prediction, medical diagnosis, and many others. These computerized techniques give today's scientist a means of handling the tedious task of reducing large amounts of data.

Table 1. Crime information fields received from the San Diego Police Department and their restructured format for computer use at LLL: phase 1

Information fields	Computer variables
Crime type (San Diego Police Dept. code)	1. Hour of day
Census tract number for the area of occurrence (geographical location)	2. Digitized year/day
Year of occurrence	3. East/west coordinates
Julian data	4. North/south coordinates
SDPD case number	5. Predicted property (characteristic):
Applicable penal code section	1=case closed
Street address where crime occurred	0.5=property (material) recovered
Hour of day	0=case open
Status indicator, giving status of case	

prediction on inferences from a historical training set, which contained a representative sample of crimes whose solution status we knew. Our preliminary results indicated that there was insufficient information to support accurate predictions.

We used four basic variables in these first experiments: time of day, time of year, east/west coordinates (X), and north/south coordinates (Y). But our accuracy in predicting whether a crime was solvable was essentially the same as that represented by the statistics in the training set (equivalent to a random guess). We had hoped for a higher value.

After reexamining the original crime data, we found a number of useful relationships that, in terms of added variables, might give us a higher predictive accuracy. For example, we realized that although we had included the north/south and east/west coordinates from a tract map of San Diego, they were not truly independent. San Diego has an irregular boundary and its population is not distributed uniformly. We had no linear measure of interpoint distances between crime locations. Consequently, as a simple mathematical technique to introduce the inherent nonlinearity, we added a fifth variable which was proportional to the crossproduct of the X and Y coordinates.

Also, we prepared a new data set, containing 100 crime cases – 20 each of rape, robbery, assault, burglary, and grand theft. Two other variables were then added: percentage occurrence and percentage of closed cases for each of the five crimes.

Our results from analyzing the 100 crimes using 7 variables are given in Table 2. Note that the multiclass technique achieved a 15% better predictive accuracy than one would expect from pure random guesswork. This increased accuracy had a twofold significance. First, it indicated that we were able to predict the probability of solving a crime with better accuracy than that suggested by the crime statistics. This carried with it the implication that with this predictive technique, the San Diego Police Department could optimize its manpower resources by assigning officers to crime cases that had a better chance of being solved.

The other significance of the 15% increase in accuracy was that it resulted from the three added variables. If we could isolate and measure additional variables, we might further increase our predictive accuracy.

Although these results were promising, our work was essentially an experiment designed from the point of view of scientists involved in an arbitrary data-analysis problem. However, we felt we had accomplished what we had set out to do: show that pattern recognition was feasible for automated crime analysis. Also, we had gained a better understanding of the property and variable relationships. Based on this progress, we went on to the next phase.

Potential

The goal of our second-phase research was to select an operationally reasonable model and to optimize the pattern recognition performance of the currently available variables within the context of the model. This meant that we would have to optimize the model's variables to yield that information best suited to the San Diego Police Department in allocating manpower.

Our work in phase 1 indicated that we might increase PATTER's predictive capability by isolating and measuring factors in the crime data whose relationship to a basic variable was not apparent. We discussed these factors with the San Diego Police Department and then reviewed the original 21 056 data records. This time we formatted 10 variables:

1. East/west coordinates (X).
2. North/south coordinates (Y).
3. XY (crossproduct).

4. Hour of day when crime occurred.
5. Date of crime (day of week).
6. Time certainty (indicating whether a crime was committed in daylight, at night, or if the time was uncertain).
7. Type of crime (indicating whether a crime was against a person or against property).
8. Closure rate (A), a crime's probability of being closed, as indicated in the original sample.
9. Occurrence rate (B), a crime's relative rate of occurrence, as indicated in the original sample.
10. AB (product of variables 8 and 9).

In the data set for this second phase, we again excluded entries with obvious data errors and entries that were "unsolvable" (e.g., natural deaths and suicides), but we retained entries with known uncertainties. This left us with 19 935 cases comprising 24 different crimes, as opposed to 26 in the first phase.

Instead of the 100 crimes selected for the training set in the first phase, this time we randomly picked 500 that had occurred before a given date. Based on this training set, we then developed classification rules.

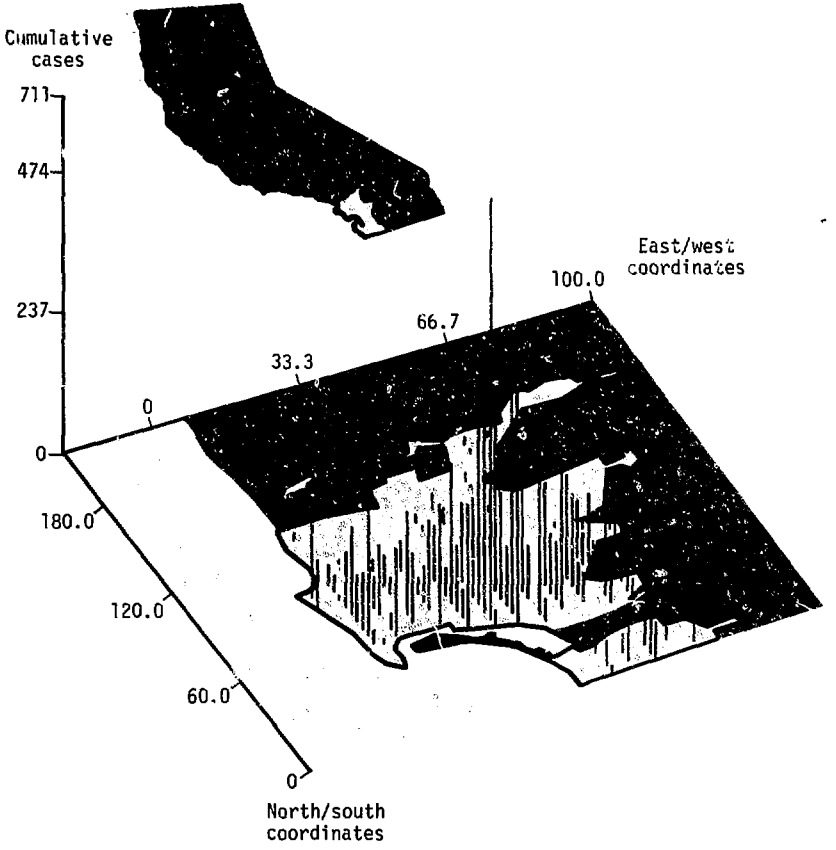


Fig. 5. Computer display of crime data: digitized crime locations plotted in three dimensions on a map outline of San Diego.

Table 2. Phase 1 results of pattern recognition analysis (PATTER), with 7 variables for 100 crime cases (20 each of rape, robbery, assault, burglary, and grand theft: 10 cases open, 10 closed)

Variables	PATTER accuracy (in predicting a crime is solvable)
1. Hour of day	K-nearest-neighbor technique:
2. Time of year	K 1 3 5 7
3. East/west coordinates (X)	% correct 63 58 57 56
4. North/south coordinates (Y)	Multiclass technique: 65% correct
5. XY (crossproduct)	
6. Percent closed originally of each type of crime	
7. Percent of each type of crime in original sample	

PATTER AND ITS TECHNIQUES

PATTER[2] was developed at LLL to solve scientific problems with pattern recognition techniques. Written in FORTRAN IV, this computer code can be used in an interactive mode, especially where fast graphic devices like CRT displays are available. It is specifically suited for modeling analytic processes, for use in spectral analyses, and for solving complex problems in data analysis in which the physical properties of large amounts of data are not well understood.

For a given collection of data, the system will try to determine if prediction of an unmeasured property of the data is possible based upon the information at hand. In constructing the system, we have striven to preserve and foster the capability for PATTER to answer not only "YES" but "NO" whenever such a response is warranted. When properly employed, PATTER will act upon a data set in such a way that it becomes apparent if useful information beyond that already discerned is contained in the data.

The input data consist of a number of patterns (NP), or objects, for which a number of variables (NV) have been recorded. This data tabulation constitutes a "pattern space" whose structure is a matrix of NV rows and NP columns, accommodating the application of linear algebra. Mathematically, each pattern can be thought of as a single point in NV space.

A well known technique that can be used for predicting an obscure or unknown property of an object is to apply the "K-nearest neighbor" rule. This rule classifies unknowns on the basis of the common properties of their K-nearest neighbors. If we accept that each pattern is a point in NV space, we can reasonably assume that in some cases points close together will be more similar than those more widely spaced. With proper data presentation, we can expect to see clusters or groups of points in which all members of a cluster exhibit some similarity or common property.

If we accept that closeness is related to similarity, then most of a point's neighbors can be expected to possess characteristics the same as those of the point in question. The K-nearest-neighbor technique is based on this concept. To predict the characteristics of a given (unknown) point, we poll one, three, five, or seven of its nearest neighboring points. The unknown point's membership in a cluster is determined by majority vote. Since each cluster is associated with a characteristic, the process of deciding cluster membership in effect predicts the point's characteristics, e.g., whether a crime can be solved.

Another useful technique is the multiclass linear classifier.[3] This rule also relies on the clustering properties of the input data. But rather than deciding a point's membership according to majority rule, it draws lines or planes in the pattern space, effectively isolating each cluster. We then use these artificial boundaries for predicting the properties of unknown points. Given an unknown point, we need only see within which boundaries it lies. Because the boundaries confine one cluster to one region, we can assign the unknown point to the cluster whose region it shares. This assignment allows us to predict that the point's properties are similar to those of the cluster's other members.

In contrast to these two discrete techniques, a number of linear and nonlinear continuous techniques exist, ranging in complexity from a conventional least-squares fit to other more sophisticated techniques. We have incorporated some of these options in the code, too.

Since pattern recognition techniques are sensitive to data representation, PATTER includes numerous statistical and mathematical routines for transforming the processed data into their most usable form. These range from autoscaling, which normalizes the pattern space, to sophisticated routines that make points more easily discernible and enhance their classification.[4]

To judge the accuracy of our predictive rules, we prepared a test-data set. Our test data consisted of 200 crimes, also picked at random, that had occurred at later dates than those in the training set. We felt that this would guarantee a true test of PATTER's capabilities to predict which crimes were the most solvable. Two hundred cases spanning the entire range of criminal activity represent a valid example of a single day's case load for the San Diego Police Department. Every day, supervisory personnel are faced with similar case loads and must decide how best to allocate their limited manpower resources to resolve these crimes. If pattern recognition techniques could establish a priority list for case assignments, police supervisors would be freed from a large part of their routine administrative burden.

A subroutine of PATTER ranked the test cases from 1 to 200 in order of probability of solution. The 50% point occurred between numbers 76 and 77 on the list. In looking at the records for the 200 cases, we found that 34 had been solved. Of these cases, PATTER had listed 20 in the first 76 of the listing. The relative cost effectiveness of solving 20 out of 76 cases (26% success rate) by automated crime analysis, as compared with solving 34 out of 200 cases (17% success rate) by current nonautomated methods, is encouraging but may not show the full leverage available if manpower is concentrated where it can do the most good – those cases which are most susceptible to solution. We have not considered the police department's method of assigning investigators to cases, and this could affect whether or not a crime will be resolved.

In the absence of specific data on case assignments, we were unable to compare results that would yield accurate cost-effective values. The availability of such data in the future will make such comparisons meaningful and let us judge how effective PATTER can be in an operational environment.

Extended Variables

Our use of PATTER in phase 1 indicated that, with a limited number of variables, pattern recognition techniques can help analyze crime statistics with predictive results that are 65% accurate. Phase 2 revealed that PATTER's predictive accuracy can be increased by including additional variables, and that this predictive capability appears to be useful in police operations.

Both phases underscored the significance of choosing appropriate variables when applying pattern recognition to crime analysis. We need still more information. As part of phase 3, we will tap data that the San Diego police routinely collect but that they never store in the department's automated data-processing system: information relating to suspects, witnesses, *modus operandi*, nature of property involved, etc., in crimes. Then we can analyze the role of these variables in the pattern recognition process, and determine if the Automated Regional Justice Information System should include these additional data fields in its computerized retrieval system.

Advanced Algorithms

If it is to yield the best results, any automated data-processing system must operate according to a well defined set of rules in a finite number of steps. In the final part of our study, we will select the more efficient processing algorithms that make use of the variables selected in the previous phase. These algorithms can then be incorporated into the San Diego system. Our last responsibility will be to close the technology transfer loop to assist the San Diego Police Department in implementing the appropriate pattern recognition techniques, thereby providing San Diego County with automated crime analysis.

Key Words: automated crime analysis; pattern recognition; computer simulation technology.

GETTING THE FACTS ABOUT OZONE

The new Satellite Ozone Analysis Center at LLL soon will be assimilating the flood of data provided by a new series of Air Force satellites for facts about ozone. Among other things, we will be producing daily high-resolution global maps of ozone concentration. Correlated with ground observations, these maps will help us to establish the current average ozone level, to check how solar flares, volcanic eruptions, and atmospheric nuclear tests change that level, to detect long-term trends in ozone concentration, and eventually to predict the effects of various forms of atmospheric pollution.

Contact James E. Lovill (Ext. 8811) for further information on this article.

The advent of aircraft capable of flying in the stratosphere – e.g., the Concorde, the USSR-SST, and the 747 SP – raises the serious question of how their exhausts will affect the ozone layer. There are also other ways by which the ozonosphere may be significantly altered, from both man-made and natural sources. Fluorocarbon releases, for example, may someday amount to enough to affect the ozonosphere, and atmospheric nuclear explosions release large quantities of NO_x, which also destroys ozone. The long-term influences of all these effects are unknown, and this lack of knowledge is in itself a cause for concern.

Land-based observations indicate that there has been a world-wide increase of ozone during the past decade.

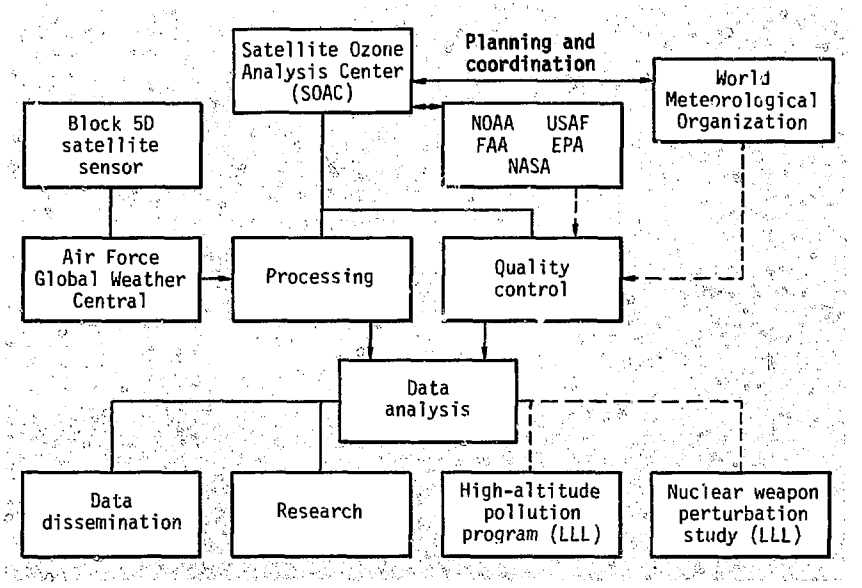


Fig. 6. Operational diagram for the Satellite Ozone Analysis Center. Coordinating with national and international agencies, the Center will accept data from the Block-5D satellite sensor through the Air Force Global Weather Central, process and compare them with independent data (quality control), and analyze and disseminate the data both for outside agencies and for our own research.

Studies of the available satellite data suggest that these observations may have been biased by a poor distribution of the stations. We will need much more satellite data to resolve this issue.

Table 3. Spectral centers, widths, and noise-equivalent spectral radiances (NESR) of the ozone, temperature, and water-vapor channels of the Block-5D, cross-track-scanning, multifilter radiometer

Center, μm	Center, cm^{-1}	Width, cm^{-1}	Species	NESR pW·sr·m (10^{-3} erg/sec-cm^2-sr-cm^{-1})
9.8	1022	12.5	O_3	50
12.0	835	9	Window	110
13.4	747	10	CO_2	120
13.8	725	10	CO_2	110
14.1	708	10	CO_2	110
14.4	695	10	CO_2	100
14.8	676	10	CO_2	90
15.0	668.5	3.5	CO_2	300
18.7	535	16	H_2O	150
24.5	408.5	12	H_2O	140
22.7	441.5	18	H_2O	90
23.9	420	20	H_2O	120
26.7	374	12	H_2O	180
25.2	397.5	10	H_2O	160
28.2	355	15	H_2O	250
28.3	353.5	11	H_2O	330

LLL's new Satellite Ozone Analysis Center was established to provide atmospheric ozone data to the national and international scientific communities and to perform research directed toward a more complete understanding of the variability of the ozonosphere. We will be analyzing ozone data from a new series of satellites and measuring the Earth's ozone variability. We will also be providing daily maps of global ozone distribution for input to computer models of the atmosphere. Figure 6 outlines how these various tasks will be organized.

The sensor used in the program will be aboard a U.S. Air Force satellite, the Block 5D, in an 835-km circular polar orbit. Four Block-5D satellites have been built, each with an estimated lifetime of at least two years. Data will be received from the first of them early in 1977.

The Block 5D contains a new scanning sensor: a crosstrack-scanning, multifilter radiometer. The sensor reads and reports 16 spectral radiance values. These include one at 9.8 μm for measuring ozone absorption, one at 12 μm for determining surface radiance, six in the 13- to 15-μm CO_2 band for delineating how temperature varies with altitude, and eight between 18 and 28 μm for measuring vertical and total water-vapor distribution.

Table 3 lists the spectral centers, widths, and noise equivalents of all these radiance channels. Note that the noise-equivalent spectral radiance (NESR) for the 9.8-μm ozone measurement is only 50 pW·sr·m (0.05 erg/sec-cm^2-sr-cm^{-1}). This allows it to detect very small changes in ozone.

The Block-5D satellites are controlled from sites in Maine and Washington. As a satellite passes overhead,

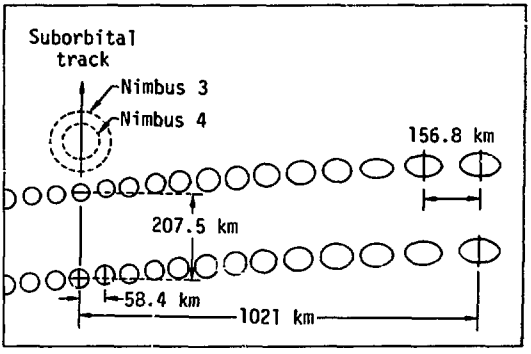


Fig. 7. Calculated Earth projection of the right half of the Block-5D sensor's scan geometry. The scan extends symmetrically out of the picture to the left. Each spot in succession is viewed by all 16 of the sensor's channels. The field of view is a minimum directly under the satellite and spreads out into an ellipse on either side. The minimum field of view is only 39.3 km diam. Dotted circles indicate the minimum fields of view of the IRIS sensors on Nimbus 3 and 4. In a day the Block-5D satellite will scan about 68 400 fields of view, 30 times as many as any previous satellite.

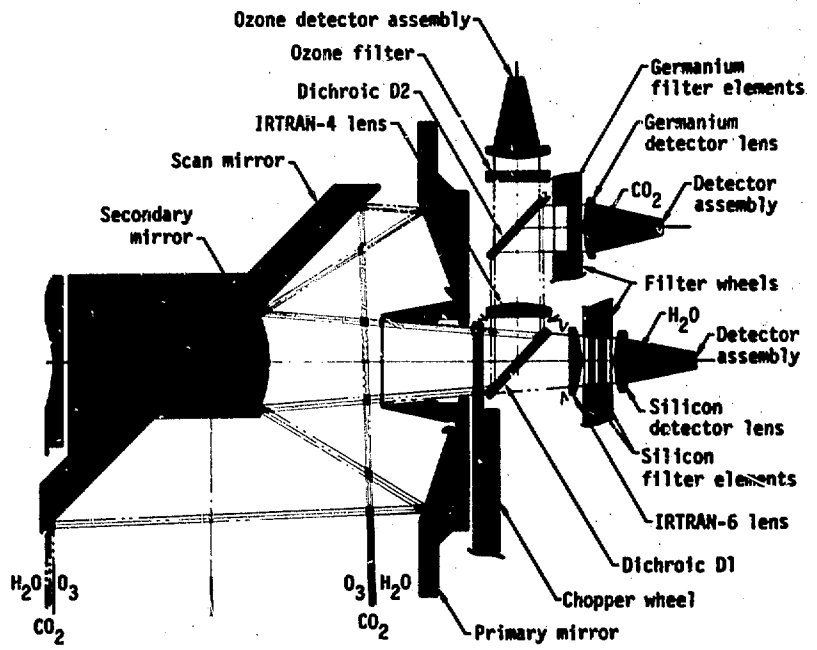


Fig. 8. Optical schematic of the ozone, temperature, and water-vapor sensors in the Block-5D satellite. Infrared radiation from the field of view is focused by the primary and secondary mirrors onto a pair of dichroic mirrors, D1 and D2, which direct the three main bands of radiation to the appropriate detectors. The filter wheels further subdivide the radiation bands into the various CO_2 and H_2O channels. The CO_2 detector with its germanium filters supplies the temperature information. The chopper wheel converts the inherently steady input signal into a rapid series of pulses, making it possible to amplify and measure the signal.

a tape recorder on board transmits the stored data back to the control station. A communications satellite then relays these data to the Air Force Global Weather Central in Nebraska.

The Weather Central transforms the satellite's 16-channel, digitized, calibrated, and Earth-located data into a more convenient format: an 8-word (36-bit) packet for each sensor sample point. There are 25 sample points per crosstrack scan, 190 crosstrack scans per orbit, and 14.4 orbits per day. This yields one ozone, six temperature, and eight water-vapor measurements for each of 68 400 locations on Earth each day.

Figure 7 shows the sensor's scan geometry. At the suborbital point, the field of view resolved is 39.3 km. This is an improvement over the fields of view for earlier satellites: 94 and 150 km for the IRIS sensors on Nimbus 4 and Nimbus 3. The instrument data-return rate is also 30 times faster than that of any previous ozone sensor, accomplishing in a day what formerly took a month or more. This greater information-gathering rate can produce rapid-fire images to capture short-term changes. It may also be allowed to accumulate to form time-exposure images that bring out in greater detail the structures of the more permanent features of the ozonosphere.

Figure 8 indicates the placement of the radiometer's optical components. The primary and secondary mirrors focus the incoming infrared radiation onto a pair of dichroic mirrors, D1 and D2. These mirrors, by selective reflection and transmission, divide the infrared radiation into three main wavelength bands and direct each band into a different detector.

The CO_2 and H_2O detectors have filter wheels that further subdivide these radiation bands into seven and eight narrow channels, respectively. Neither of these detectors can look at more than one channel at a time, however. This means that the ozone detector, with its single channel, is gathering energy continuously, while the other detectors are switching from channel to channel.

Between the secondary mirror and the first dichroic mirror there is a chopper wheel with evenly spaced teeth that breaks the beam into a rapid series of pulses. This converts the slowly varying incoming light into rapidly fluctuating detector signals that are easy to

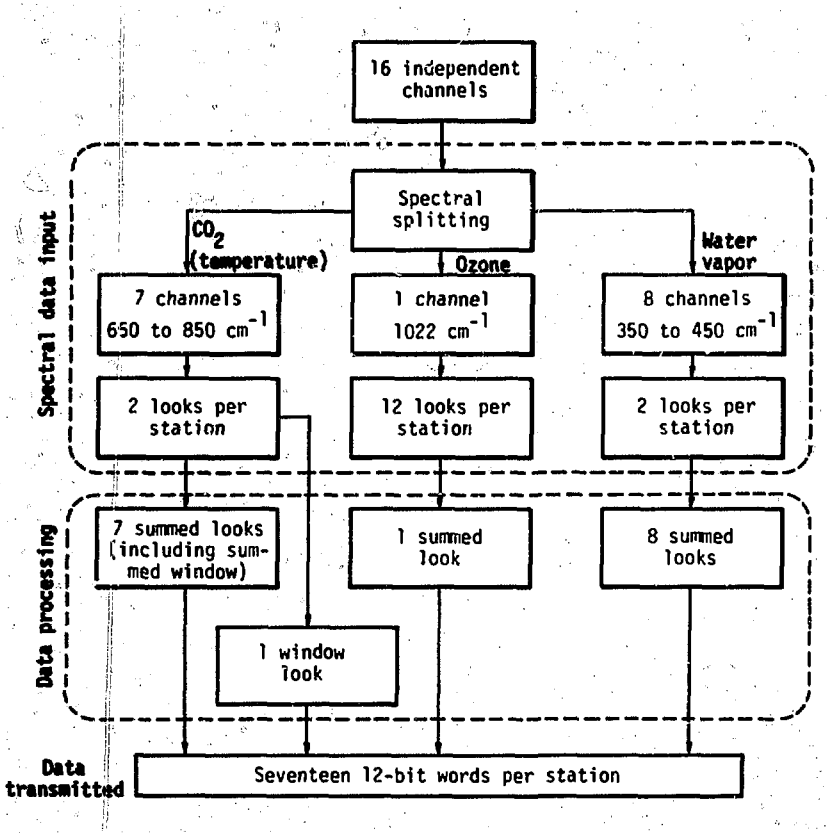


Fig. 9. Information flow diagram for the sensor in the Block-5D satellite. The ozone detector monitors only a single channel, whereas the temperature (CO_2) and the H_2O detectors must divide their time among seven or eight channels. This means that the ozone channel handles several times as much energy as the other channels.

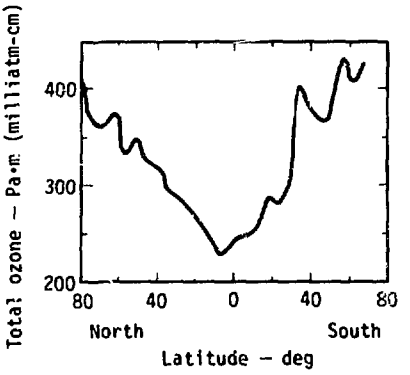


Fig. 10. Total-ozone variation during one orbital pass from western Siberia, across the western Pacific and Australia, to Antarctica. The data from which this curve was constructed were gathered by the Nimbus-3 satellite. There is a range of nearly a factor of 2 in ozone concentration between the polar and the equatorial regions.

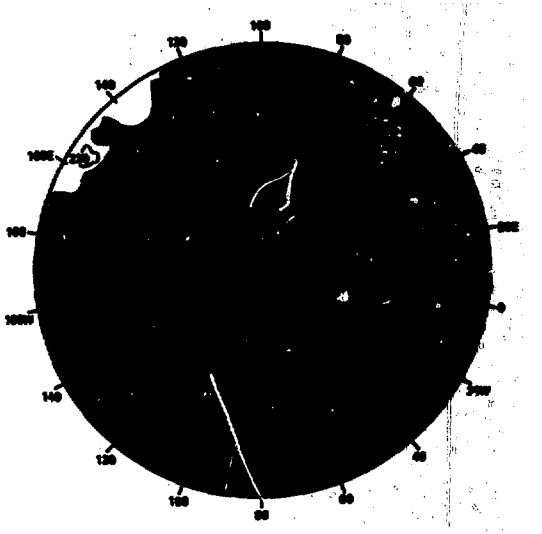


Fig. 11. Total-ozone distribution over the northern hemisphere averaged over a 93-day period, April to July, 1969. The contours indicate ozone concentration in Pascal-metres (milliatmosphere-centimetres). This map, drawn from Nimbus-3 data, is an example of the kind of map that we intend to produce daily. The grey area, outlined by the 240-Pa·m contour, indicates a semipermanent region of low ozone concentration located west of the Phillipine Islands and north of Australia. This phenomenon is not currently understood; we hope that additional data from the Satellite Ozone Analysis Center will help to explain it.

amplify and digitize. Each of the light pulses is termed a "look."

Figure 9 outlines the data flow in the sensor logic. The output from the data processing section is stored in an onboard tape recorder for transmission whenever the satellite comes within range of one of the ground control stations. The 17 12-bit words transmitted for each "station," i.e., sample location, contain the entire readout from the 16 channels of information.

Once we receive this information, we must first convert the radiance values into measurements of ozone concentration, temperature, and water-vapor distribution. This is a matter of mathematical inversion, using radiative transfer equations and

spectroscopic transmission functions to deduce the vertical temperature and compositional profiles. We have developed inversion techniques that provide solutions to the radiative transfer equations, and we have compiled transmission functions by combining molecular theory with field and laboratory measurements.

We will be comparing the satellite-derived total-ozone values with measurements from selected total-ozone ground stations at various locations in each hemisphere. For example, our total-ozone value will be compared with the Dobson Spectrophotometer value whenever a software program computes that the satellite ozone value comes from a spot within 50 km of a ground station. In addition, weekly and monthly Dobson average values at stations in the ground network will be compared with the satellite-derived average values.

As a further check, we will be comparing the Block-5D ozone values with those obtained from the NASA TOMS ozone sensor, which is scheduled for a 1978 launch date. When the second of the four Block-5D satellites is launched later in 1977, we will be comparing the ozone channels of the two satellites with one another and with the total-ozone ground network. (After the second 5D system is launched, two satellites will be maintained in orbit into the 1980's.)

An additional feature of these comparisons is that the polar orbits of the two Block-5D satellites will permit simultaneous observation of widely separated ozone responses to transient events, like volcanic eruptions and solar flares. They will also yield data on day-to-night ozone changes for particular locations all over the world.

Figure 10 shows the total-ozone measurement along a single north-south pass of the Nimbus-3 satellite. This particular pass started in western Siberia, crossed the western Pacific Ocean and Australia, and ended in Antarctica. The total ozone atmospheric column varies by almost a factor of 2 between the polar regions and the equator.

We will be performing a computer spline analysis on the 68 400 daily ozone measurements from the Block-5D satellite and displaying the data in a global 1° and 2° grid (as well as a larger 5° and 10° grid). Figure 11, compiled from Nimbus-3 data, indicates the form of these global displays. This figure summarizes about 83 000 data points gathered in the course of 93 days. The Block-5D system will provide a similar map *every day*.

In addition to daily global and hemispheric ozone maps, we will be compiling monthly analyses that average data over the preceding 30-day period. These analyses will be published and distributed to interested scientists throughout the world at cost. Similar analyses and maps of CO_2 and H_2O radiance will also be provided.

In summary, the Satellite Ozone Analysis Center will:

- Monitor the Earth's ozone variability from early 1977 into the 1980's, establishing a baseline against which regional and global trends can be measured.
- Provide global total-ozone distribution data as input to computer weather-and-ozone analysis and forecasting.
- Analyze ozone variability within the mesoscale (e.g., study with high-resolution satellite-ozone data the heavily travelled North Atlantic air corridor between the U.S. and Europe).
- Analyze the diurnal variability of total ozone with a two-satellite ozone-sensor system.
- Interact with the global ozone monitoring and research project of the World Meteorological Organization.

Key Words: ozone; ozone - depletion; ozone - distribution; Satellite Ozone Analysis Center; SOAC.

Notes and References

1. Results of this preliminary phase were reported in C. Bender, L. Cox, Jr., and G. Chappell, *An Application of Pattern Recognition Techniques to Crime Analysis*, Lawrence Livermore Laboratory, Rept. UCID-17224 (1976).
2. L. Cox, Jr., and C. Bender, *PATTER: A Polyalgorithm for the Analysis of Generalized Data Sets*, Lawrence Livermore Laboratory, Rept. UCID-16915 (1975).
3. C. Klopfenstein and C. Wilkins, *Computers in Chemical and Biochemical Research* (Academic Press, New York, 1972), vol. 1.
4. Details of the theory and application of classification techniques for pattern recognition can be found in W. Meisel, *Computer-Oriented Approaches to Pattern Recognition* (Academic Press, New York, 1972).

RBC/KLC/BDJ/JKP/lh/wl/lmc