John E. Estes
Jeffrey L. Star
Geography Department
University of California
Santa Barbara, CA 93106
Philip J. Cressy
NASA Goddard Space Flight Center
Greenbelt, MD 20771
Michael Devirian
NASA Headquarters
Washington, DC 20546

Pilot Land Data System

PLDS will be a limited-scale distributed information system to explore scientific, technical, and management approaches to satisfy land science research needs.

Introduction

Satellite remote sensing can provide information of tremendous value for the Earth Sciences. However, realization of this potential requires information systems not currently available. Technological advances now make it possible to design a data system to meet the land scientists' most critical information systems needs, and prepare the community for the Space Station era.

Abstract: *Beginning in 1983 and continuing until the present, the National Aeronautics and Space Administration (NASA) coordinated a series of meetings to develop initial plans for a Pilot Land Data System (PLDS). PLDS is intended to improve the ability of NASA and NASA-sponsored researchers to conduct research on land processes. The meetings have coordinated planning, concept development, and implementation activities, and examined research and information science requirements, and strategies for system evaluation. PLDS will be a limited-scale distributed information system to explore scientific, technical, and management approaches to satisfy land science research needs. Implementation is beginning in FY85. PLDS can pave the way for a Land Data System, and possibly the Earth Observing Information System of the space station era, by improving access to data and analysis capabilities, fostering an environment in which information synthesis can occur at scales not previously possible.*

Under the sponsorship of the NASA Information Systems Office, the Universities Space Research Association (USRA) assembled a working group to examine the need for a Pilot Land Data System (PLDS). The working group included discipline scientists, information scientists, and management personnel from universities, private industry, and the federal government.

Participants at meetings determined that the goal of the pilot program should be to establish a limited-scale, distributed information system to explore approaches to satisfy the needs of the land science research community. Accomplishing this goal will require land and information scientists working closely together to understand the needs of the users of scientific data. The system must support the full spectrum of functions needed to conduct land science investigations, including data location, acquisition, processing, and transfer. Properly developed, PLDS will provide a sound technical basis for a future, fully operational Land Data System. Such a system could serve as a key component of an Earth Observing Information System (EOIS), which will support science in the space station era of the 1990's.

Study of environmental processes on the Earth's surface requires a multidisciplinary approach. This has been recognized in the definition of several major programs, such as the International Geosphere Biosphere Program (IGBP), Global Habitability and Global Biology, and the International Satellite Land Surface Climatology Program

Photogrammetric Engineering and Remote Sensing,
Vol. 51, No. 6, June 1985, pp. 703-709.
0099-1112/85/5106-0703$02.25/0

(ISLSCP; see Waldrop (1984), NASA (1983a), NASA (1983b), NASA (1984b)). This recognition has given rise to the concept of multidisciplinary information systems. Integration of the PLDS (as the precursor to a Land Data System) with other discipline data systems can provide a foundation for such integrated systems as EOIS and the Global Resources Information System (Billingsley *et al.*, 1984).

It was recognized from the outset that developing a PLDS is a complex task. PLDS must be a distributed system, because both the data and the users are geographically distributed. This complicates network and communications designs dramatically, compared with other NASA pilot data systems (for example, the Pilot Ocean Data System at the Jet Propulsion Laboratory, or the Pilot Climate Database System at Goddard Space Flight Center). Also, due to the multi-disciplinary and inter-institutional nature of land sciences research, this proposed system must be based on cooperation among NASA Centers and other institutions.

A number of principles were adopted for PLDS planning and design:

- Data bases tend to remain most viable when maintained by active researchers with a long term commitment to the use and sharing of the data (CODMAC, 1982);
- PLDS will serve the data and information systems needs of NASA and NASA-related scientists working on land science projects;
- PLDS represents a research and "proof-of-concept" tool;
- Long-term goals must be defined, both in terms of major Earth science issues to be examined and feasible tools for the task;
- PLDS must exploit components in place at participating institutions, and testing of the elements of the PLDS must build upon ongoing research programs; and
- System development should be based on available, well-understood technology. Close coordination with NASA computer science and communications research must provide the mechanism for incorporation of new technology and upgrades of the PLDS/LDS.

The material which follows details the science objectives, whose pursuit (let alone successful accomplishment) requires the development of an advanced information system. An abridged science scenario is included to provide an example of the type of requirements which could be levied on a Land Data System. A conceptual overview of a future Land Data System is presented next. This is followed by a brief description of the proposed development of the Pilot Land Data System. Finally, a brief discussion of the conclusions of the PLDS Working Group is presented and the current status of PLDS is described.

At present, PLDS has been funded for the initial planning phase. Management responsibilities for technological areas within PLDS have been assigned to different organizations within NASA, and a science steering group has been assembled. System implementation is beginning in 1985.

Background

The launch of Landsat 1 stimulated major advances in the science and technology of remote sensing. These, as well as comparable advances in information sciences, are changing the nature of land science. Traditionally, field studies in land sciences have been limited in focus to a few variables in a small geographic area. This was due in large part to the problems of obtaining and working with large volumes of data.

Through the Landsat program, the land science community now has observational tools at scales appropriate to examine the critical processes that define a "real world" system. However, full realization of the scientific potential of satellite remote sensing has been handicapped by inadequate information systems. The ability to access and exchange both data and software is hindered by lack of both communications and standards. Appropriate computational resources are often lacking. Scientists are now required to devote a significant portion of their efforts to data acquisition and preparation. Better integration of remote sensing and information technologies can overcome these barriers.

In recognition of the need for improved understanding of large scale Earth processes, there is a movement in scientific research in general (and in NASA Earth science programs) to ask research questions which are both multidisciplinary and global in scale (Gwynne, 1982; NASA, 1983a; NASA, 1983b; NASA, 1984b). The resolution of such large-scale science issues requires interdisciplinary research teams and sophisticated technologies. It was intended from the beginning that these kinds of science issues drive the evolution of the PLDS, and that new advances in information science will, in turn, create a new perspective for looking at critical problems in the land sciences.

There are many environmental problems with important economic, human health, and environmental impacts. An ultimate objective of the scientific community is to understand correctly the factors involved in land processes and to provide a sound predictive modeling capability. Some of the important goals identified (NASA, 1983a) in the land science area are to

- establish methods by which a global carbon budget model may be developed and monitored;
- detect the presence and amounts of pollutants;
- establish the relationship between the energy balance and biophysical conditions on land, and their interrelationships with climate;
- improve the accuracy of models used for prediction of the availability and quality of water, snow, and ice;

- identify the early indicators of change in global element cycles, climate, and hydrology;
- advance the understanding of global and regional geologic and geomorphic structure and process; and
- develop improved methods for assessing and monitoring geologic hazards.

In order to achieve goals of this scale, an efficient processing and information management system is essential. While it is relatively easy to conceive of the general operation of a Land Data System, there is no existing prototype for it. The technology for each element of a Land Data System is understood, but experience in the integration of technologies must be developed before NASA can proceed toward implementing a global scale system (NASA, 1984a). A well-defined pilot data system, serving a relatively small group of users, is a necessary first step.

SCIENCE SCENARIOS

As a pilot program, PLDS cannot meet all possible needs. A limited number of science scenarios were selected from both existing and proposed research projects within the NASA land science community. Through these scenarios, the working group derived the generic information system functions and requirements while keeping an appropriately narrow focus.

The science projects used as examples were

- Vegetation Biomass and Productivity, and Large Area Inventory;
- Biogeochemical Cycling in Forests;
- Land Surface Climatology;
- Hydrologic Modeling and Soil Erosion/Productivity Modeling;
- Multispectral Analysis of Sedimentary Basins; and
- Monitoring Environmental Change

The land surface climatology and sedimentary basins projects were later selected by NASA Office of Space Science and Applications Code E personnel for initial incorporation in PLDS. These projects, in particular, represent a wide mix of requirements from both science and information systems perspectives.

The following is an abridged version of the Land Surface Climatology scenario, developed primarily by researchers at NASA Goddard Space Flight Center. This scenario (as well as the others used to drive PLDS design) was not intended to be a complete research description, but to highlight data access and processing requirements.

The objective of the Land Surface Climatology Project is to develop a better understanding of the interactions among the Earth's biospheric, edaphic, hydrologic, and atmospheric systems, and to determine their roles in influencing climate. An improved understanding of these processes and interactions can best be achieved through the development and validation of terrestrial and climatological process models which require many diverse types of data.

Investigations in land-surfaces climatology are being supported through a new international program, the international Satellite Land Surface Climatology Project (ISLSCP). ISLSCP is conducted under the auspices of COSPAR and the International Association of Meteorology and Atmospheric Physics. Goddard Space Flight Center and a number of other institutions (including universities, other federal agencies, and international organizations) will be participating in this project.

This research program requires facilities to move information between the data bases and computational systems at physically separate facilities. The development of preprocessing capabilities and interchange standards will greatly facilitate this research. The definition of generic data formats, projections, and file structures could lead towards greater compatability among institutions. From an evaluation of this and other science scenarios, a number of information systems' support needs were developed.

REQUIRED FUNCTIONS

Figure 1 shows a general model for an information system to support land sciences research. This diagram generalizes the steps in this type of research, and identifies the functions which could be supported by a PLDS. Table 1 summarizes the results of analyzing each project in this manner. Priorities were assigned by discipline scientists as follows:

1—enable the scientists to do the research.

2—enhance the scientists' ability to do the research.

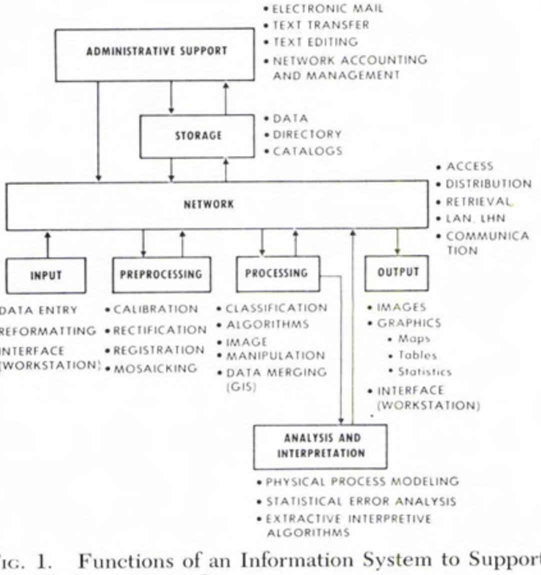

FIG. 1. Functions of an Information System to Support Land Science Research.

TABLE 1.

Processing Functions	Scenario 1	2	3	4	5	6	Total
Input							
Data Encoding	1	2	2	1	2	2	10
Data Reformatting	1	2	1	1	2	2	9
Preprocessing							
Data Calibration	2	3	1	1	2	2	11
Image Registration	1	1	1	1	1	2	7
Image Mosaicking	1	1	3	2	1	2	10
Processing							
Multi-source geocoded data overlay	1	1	1	1	1	1	6
Image and statistical processing (software sharing)	1	1	2	1	3	1	9
Analysis							
Statistical analysis	2	2	4	2	4	4	18
Modeling	1	2	4	2	4	4	17
Output							
Image	1	2	3	1	3	4	14
Statistical (tabular)	2	2	3	2	4	4	17
Tables and figures (graphic)	2	2	3	2	4	4	17
Storage media—CCT, disc	2	3	3	2	4	2	16
Network Storage							
Directory	1	1	3	1	2	2	10
Catalog	2	1	2	1	2	2	10
Data	2	2	1	2	3	2	12
Network Distribution							
Access to archive data	1	1	1	1	1	2	7
Networking of processing	1	1	1	1	3	2	9
Shared peripherals for output	2	2	2	1	4	2	13
Network Administrative Support							
Electronic mail	2	1	3	3	2	4	15
Text transfer	2	1	3	3	2	4	15

Legend: 1—Enable Scientific Research
2—Enhance scientist ability to do research
3—Research could be accomplished now but Support Service would be useful
4—PLDS support not required

3—research can already be accomplished but support would be useful.
4—PLDS support is not required.

Table 2 shows the functions in order of priority. The science scenarios require PLDS for data storage, input, preprocessing, and distribution on a high priority basis. Lower priority is assigned to support for analysis and output as well as network administration (although some of these items may be implicitly required to support the functions that were assigned higher priority).

CHARACTERISTICS OF A LAND DATA SYSTEM

Consideration of the science scenarios led the working group to a description of an information system for the future, a Land Data System (LDS). The overall goal of LDS would be to provide a powerful and responsive system to support land science research, facilitate understanding of the land resource complex, and provide general access to relevant data sets and processing capabilities.

Important characteristics of LDS include

- User-friendly interfaces for novice and expert users;
- Systematic archiving, maintenance, and access to relevant data;
- Data management and manipulation tools;
- Simple acccess to existing bibliographic information systems; and
- Maintenance of the history of data sets (origin, calibration information, and so on).

A system exhibiting these characteristics could change the character of land science research. Such a system would enable multi-disciplinary, multi-institutional research which is not now practical, and could allow experiments to be conducted in near-real time when required.

As currently envisioned, the LDS would consist of five major subsystems: Data Management, Communications and Networking, Intense Computational Processing, User Interface, and Input/Output Interface.

The Data Management Subsystem would provide both data and information about the data to the scientists. Ultimately, users would communicate with the subsystem using natural language. The subsystem would also store and update large amounts of data and support many users concurrently.

The Communications and Networking Subsystem

TABLE 2. ORDERED RANKING OF INFORMATION SYSTEM FUNCTIONS

	Total
Multi-source geocoded overlay	6
Access to archived data	7
Image registration	7
Data reformatting	9
Software sharing	9
Networking of processing	9
Mosaicking of images	9
Directory of information	10
Calibration of data	10
Data encoding	10
Data storage	12
Shared peripherals for output	13
Image output production	14
Electronic mail	15
Text transfer (compatible text editing)	15
Output storage media	16
Tabular output production	17
Graphic output production	17
Modelling	17
Statistical analysis	18

would support a near-real-time interface between the other subsystems. Such communications would be supported by several technologies, including packet switching network communications, local area networking, and satellite communications.

In an LDS, data manipulation and analysis would occur on a number of the subsystems. The Intensive Computational Processing Subsystem would provide the power of large-scale computers to users around the network for technically demanding tasks such as image interpretation and pattern recognition.

The User Interface Subsystem would consist of a range of microprocessor and minicomputer workstation types with differing capabilities. Workstations will be connected to the other subsystems by means of the Communication and Networking Subsystem and will either interface to local processing facilities or serve as free-standing processing stations.

The Input/Output Interface Subsystem would connect the overall LDS with outside computer systems and data sources. This subsystem will perform reformatting, modification, and data manipulation, and allow the overall LDS to communicate efficiently.

A functional overview of the LDS concept is seen in Figure 2, and possible structure of a node is shown in Figure 3. The detailed node in the latter could be a NASA Center, but nodes are also expected to be located at universities and other institutions or agencies.

PILOT LAND DATA SYSTEM DEVELOPMENT

PLDS development is based on three fundamental principles. First, the system will build on existing

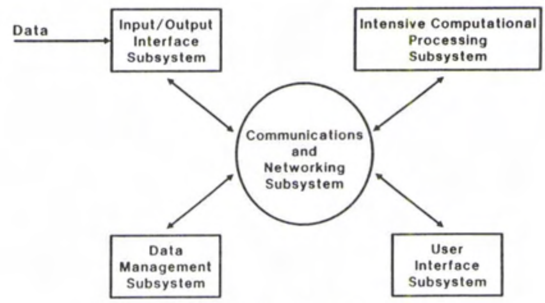

FIG. 2. LDS Subsystems Overview.

capabilities where possible, minimizing costs and permitting rapid concept testing. Second, a structured system engineering effort at the onset of the project helps to keep long-range goals in view. Finally, new technologies should be regularly reviewed for integration where appropriate. It is important to note that researchers will continue their ongoing efforts to improve data management and communications; the PLDS provides a formalism and a focus for further progress.

Functional requirements for PLDS can be summarized as goals by the end of fiscal year 1987:

- Establish communications capabilities;
- Build directories and catalogs of data sets at NASA Centers, participating universities, and other agencies;
- Develop an efficient data management system;
- Demonstrate remote access and use of data;
- Demonstrate that remote requests for value-added services (calibration and rectification, for example) can be answered in a timely way; and
- Demonstrate the expandability of the system.

In any pilot study, there must be periodic benchmarks to measure progress. Technical measures (e.g., data transmission volumes) can be used to evaluate some aspects of the system. While scientific achievements may not lend themselves to similar quantification, periodic evaluations and peer review, as well as publications in reviewed journals, can serve to evaluate program progress.

When fully implemented, PLDS will be capable of supporting a subset of the NASA-sponsored land science community. Systems concepts for the phases of PLDS, and the transition to an operational LDS (which could begin in 1989), are illustrated in Figure 4. Because technologies used in the PLDS are in a state of rapid development, development must proceed with a full awareness of the volatility of these technologies in order to prevent built-in obsolescence.

Implementation of PLDS has begun. A science working group has been convened under Dr. Robert Price of NASA Goddard Space Flight Center (GSFC). Dr. Paul Smith (also of GSFC) heads the technology working group. The sedimentary basin study at Jet Propulsion Laboratory and the land surface clima-

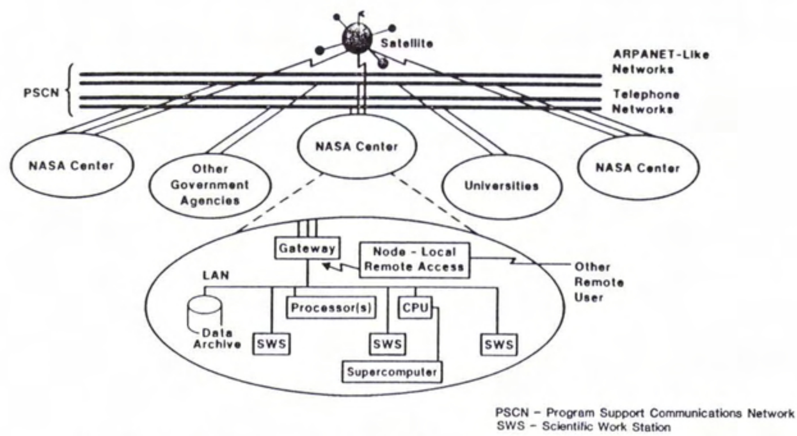

FIG. 3. LDS Communications and Information Analysis Overview.

tology project at GSFC are being used as initial science scenarios to drive pilot development. PLDS implementation is under the coordinated management of the Earth Science and Applications Branch and the Information Systems Office of NASA Headquarters Code E.

In addition, a science steering group under Dr. Ray Avidson (Washington University, St. Louis) has been convened. Specific development is occurring in the areas of systems engineering, communications, work stations, image processing, and intensive computation. Early efforts are directed towards a demonstration of the ability to link scientists at a number of institutions, and facilitate their research through improved communications.

CONCLUSION AND SUMMARY

Satellite remote sensing is a unique tool, providing data of a type and on a scale previously unobtainable. Yet, particularly with Space Station and the Earth Observing System (EOS) on the horizon, the applications of satellite remote sensing are handicapped by inadequate information systems. Future systems must not stop at the ground receiving station, but must fully integrate the flow of information to meet the user's ultimate needs.

There is a need to improve the ability of NASA and NASA-sponsored scientists to use remotely sensed and other land resource science data. Unless the ability to handle these and other land science data is established now, effective use of data from future systems (e.g., Moderate Resolution Imaging Spectrometer, High Resolution Imaging Spectrometer, High Resolution Multifrequency Microwave Radiometer, and Synthetic Aperture Radar) will be severely impacted. PLDS can permit researchers to better address important, multi-disciplinary science questions, and can lead to improved understanding of many land processes. PLDS can be a means of increasing scientific productivity through better use of information science technology.

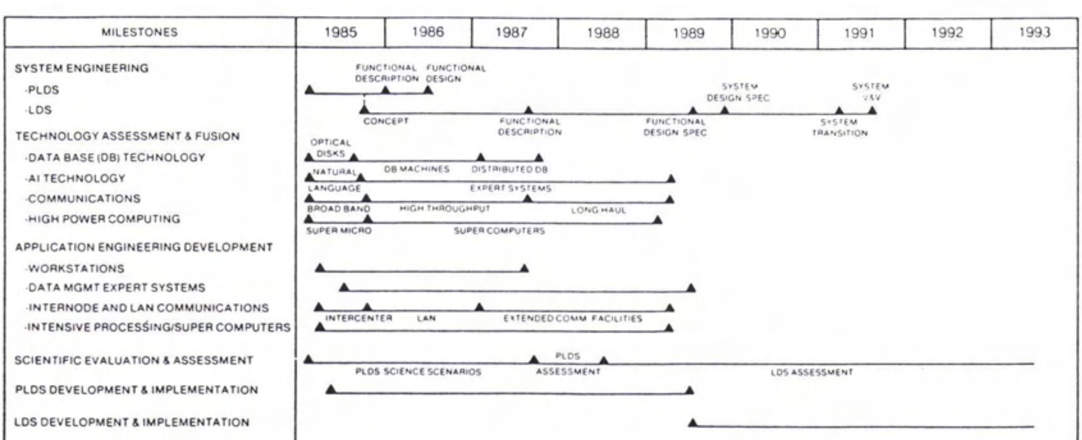

FIG. 4. PLDS/LDS Development Schedule.

Based upon these conclusions and the recommendations of the PLDS Working Group, NASA has begun a detailed planning phase for PLDS. The initial system will be a limited scale, distributed information system, directed towards supporting NASA and NASA-sponsored land science researchers, to permit them to function more effectively as scientists rather than librarians and communications experts. Work has begun on PLDS . The effort involves the cooperation of both discipline- and technology-oriented scientists. The degree to which it can improve the quality of research will be its ultimate measure of success.

The working group established that such a system is needed and will facilitate research. We encourage all land science researchers to monitor the progress of this system, and we encourage your comments.

ACKNOWLEDGMENTS

We gratefully acknowledge the sponsorship and encouragement of the Information Systems Office, NASA Headquarters and Dr. Caldwell McCoy; the Office of Space Science and Applications, Mr. Alexander Tuyahov; Ms. Janet Franklin and Universities Space Research Association; and especially the members of the Pilot Land Data System Working Group, whose contributions and synthesis we have summarized here.

REFERENCES

Billingsley, F. C., J. L. Urena, J. E., Estes, and J. L. Star, 1984. *Global Resources Information System: A Concept Paper*. NASA Jet Propulsion Laboratory Publication D-1524.

CODMAC (Committee on Data Management and Computation), 1982. *Space Science Board, National Academy of Sciences, Data Management and Computation, Vol 1: Issues and Recommendations*, National Academy Press.

Gwynne, M. D., 1982. The Global Environment Monitoring System (GEMS) of UNDP. *Environmental Conservation* 9(1):35-41.

NASA (National Aeronautics and Space Administration), 1983a. *Land-Related Global Habitability Science Issues*. Land-Related Global Habitability Sciences Working Group. NASA Technical Memorandum 85841. 112 p.

NASA (National Aeronautics and Space Administration), 1983b. *Global Biology Research Program; Program Plan*. NASA Technical Memorandum 85629. 112 p.

NASA (National Aeronautics and Space Administration), 1984a. *The Pilot Land Data System: Report of the Program Planning Workshops*. Pilot Land Data Systems Working Group. NASA Technical Memorandum 86250. 170 p.

NASA (National Aeronautics and Space Administration), 1984b. *Earth Observing System: Science and Mission Requirements Working Group Report*. NASA Technical Memorandum 86129. 107 p.

Waldrop, M. M., 1984. An inquiry into the State of the Earth. *Science* 226:33-35.

(Received 13 October 1984; revised and accepted 11 February 1985)

Forum

December 1984 Cover Photo

I WOULD LIKE to make a comment on the explanation accompanying the cover photo in the December 1984 issue of *Photogrammetric Engineering and Remote Sensing*. The explanation states that "red areas were dominated by green vegetation in May, but not in August," which is a very strange interpretation for the tropical rain forest in West Africa, especially as August is in the middle of the wet season. One would have expected the West African coastal rain forest to be colored yellow as in the neighboring hinterland, signifying green vegetation in both May and August, but it would appear that the coastal rain forest actually has a higher vegetation index at the start of the wet season than in the middle.

One possible explanation is that the rain forest has a higher productivity rate at the beginning of the wet season than in the middle. Carneggie *et al.* (1974) have shown that near infrared reflectance is more dependent upon productivity rate than upon the green biomass of the grassland they were investigating; but if this was the reason for the apparent greater vegetation index in May than in August for the West African coastal rain forest, why is the hinterland not also red?

The reason may be that rain forest actually appears to have a lower vegetation index than wet season grassland, probably because of the relatively smooth grassland surface and high internal reflection within the forest. By the middle of the wet season (August), the forest could have an even denser forest canopy, producing an even stronger internal reflection, and so apparently lowering the vegetation index compared with that at May at the beginning of the wet season.

REFERENCE

Carneggie, D. M., S. D. de Gloria, and R. D. Colwell, 1974. *Usefulness of ERTS 1 and supporting aircraft data for monitoring plant development and range conditions in California's annual grasslands*. University of California, Berkeley, School of Forestry and Conservation, Final Report to the U.S. Department of the Interior, Bureau of Land Management.

—*R. B. King*