JSC 09388

CROP IDENTIFICATION TECHNOLOGY ASSESSMENT FOR REMOTE SENSING (CITARS)

VOLUME V

THE FIRST EARTH RESOURCES TECHNOLOGY SATELLITE (ERTS-1) DATA PREPARATION

National Aeronautics and Space Administration

LYNDON B. JOHNSON SPACE CENTER

Houston, Texas

March 1975

ii is blank

JSC 09388

CROP IDENTIFICATION TECHNOLOGY ASSESSMENT FOR REMOTE SENSING (CITARS)

VOLUME V

THE FIRST EARTH RESOURCES TECHNOLOGY SATELLITE (ERTS-1) DATA PREPARATION

PREPARED BY

William R. Simmons
W. R. Simmons
(LARS)

David M. Freeman
D. M. Freeman
(LARS)

Marvin E. Bauer
M. E. Bauer
(LARS)

APPROVED BY

Robert M. Bizzell
Robert M. Bizzell, Project Manager

Andrew E. Potter
Andrew E. Potter, Chief
Research, Test, and Evaluation Branch

R. B. MacDonald
R. B. MacDonald, Chief
Earth Observations Division

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

March 1975

iv is blank

V

TECHNICAL REPORT INDEX/ABSTRACT

(See instructions on reverse side.)

1. TITLE AND SUBTITLE OF DOCUMENT Crop Identification Technology Assessment for Remote Sensing (CITARS) - Volume V, The First Earth Resources Technology Satellite (ERTS-1) Data Preparation	2. JSC NO. JSC-09388
3. CONTRACTOR/ORGANIZATION NAME Lockheed Electronics Company, Inc.	4. CONTRACT OR GRANT NO. NAS9-12200
5. CONTRACTOR/ORIGINATOR DOCUMENT NO. LEC-4326D	6. PUBLICATION DATE (THIS ISSUE) March 1975
7. SECURITY CLASSIFICATION Unclassified	8. OPR (OFFICE OF PRIMARY RESPONSIBILITY)
9. LIMITATIONS GOVERNMENT HAS UNLIMITED RIGHTS ☐ YES ☒ NO IF NO, STATE LIMITATIONS AND AUTHORITY	10. AUTHOR(S) W. R. Simmons, D. M. Freeman, and M. E. Bauer
11. DOCUMENT CONTRACT REFERENCES	12. HARDWARE CONFIGURATION
WORK BREAKDOWN STRUCTURE NO. Job Order 71-645	SYSTEM LARSYS
CONTRACT EXHIBIT NO.	SUBSYSTEM
DRL NO. AND REVISION	MAJOR EQUIPMENT GROUP
DRL LINE ITEM NO.	

13. ABSTRACT

This document records the data preparation activities and their results carried out in support of the Crop Identification Technology Assessment for Remote Sensing. The major objective of this task was quantification of the crop identification performance resulting from the remote identification of corn, soybeans, and wheat by using several automatic data processing techniques developed within a controlled environment at the National Aeronautics and Space Administration - Lyndon B. Johnson Space Center/Earth Observations Division, the Laboratory for Applications of Remote Sensing at Purdue University, and the Environmental Research Institute of Michigan. For this objective, subtasks utilizing data collected by the Earth Resources Technology Satellite, the Bendix M^2S, and the multispectral scanner systems of the Environmental Research Institute of Michigan were defined. The primary data used for analysis were collected by the Earth Resources Technology Satellite. This document describes the results of the acquisition and preparation of multispectral scanner data for classification. Data quality evaluation, reformatting, multitemporal registration, geometric correction, and section and field coordinate locations are discussed.

14. SUBJECT TERMS

Crop identification	Geometric correction	
Reformatting	Section and field coordinate location	
Data quality evaluation	Multitemporal registration	

JSC Form 833 (Rev Sep 74) NASA-JSC

V

vi is blank

GLOSSARY

Along-track – in the direction of the spacecraft ground track; sometimes called vertical when referring to output product coordinates.

ASCS – Agricultural Stabilization and Conservation Service, an agency of the U.S. Department of Agriculture.

CCT – computer-compatible tapes (standard 1/2-inch wide magnetic tapes in nine-track or seven-track format) containing digital ERTS-1 data.

CITARS – Crop Identification Technology Assessment for Remote Sensing.

DAS – data analysis station, a computer system consisting of tape drives and computer, a display and control console, and film recorder used to reformat, analyze, and review remotely sensed digital data.

EOD – Earth Observations Division of the Lyndon B. Johnson Space Center, Houston, Texas.

ERIM – Environmental Research Institute of Michigan, Ann Arbor, Michigan.

ERTS-1 – the first Earth Resources Technology Satellite, renamed LANDSAT-1 in January 1975.

GSFC – Goddard Space Flight Center, Greenbelt, Maryland.

Image skew — image distortion caused when the scan of the sensor is not perpendicular to the plane formed by the spacecraft and the instantaneous ground-track velocity vector.

JSC — Lyndon B. Johnson Space Center, Houston, Texas.

LARS — Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana.

LARSYS — the name designating the remote sensing data analysis programs developed at LARS.

Least squares — a statistical procedure that involves minimizing the sum of squared differences.

MSS — multispectral scanner system.

NASA — National Aeronautics and Space Administration.

Nearest neighbor rule — the point nearest the desired sample location used to represent the value at the desired location.

Pixel — picture element (refers to one instantaneous field of view as recorded by the multispectral scanning system, which is equivalent to approximately 0.44 hectare, or 1.09 acres).

rms – root mean square.

Spectral bands – the division of the visible and near-infrared portions of the electromagnetic spectrum into discrete segments.

MSS channel	ERTS-1 band	Wavelength, micrometers	Spectrum
1	4	0.5 - 0.6	Green
2	5	0.6 - 0.7	Red
3	6	0.7 - 0.8	Reflective infrared
4	7	0.8 - 0.11	

Standard deviation – a measure of the dispersion of data points around their mean value represented by the symbol σ .

Threshold – the boundary in spectral space beyond which a data point (pixel) has a sufficiently low probability of being included in a given class and, therefore, is purposely excluded from that class.

X is blank

CONTENTS

xii is blank

TABLES

xiv is blank

FIGURES

1.0 INTRODUCTION

Preparation of data from the first Earth Resources Technology Satellite (ERTS-1) for the Crop Identification Technology Assessment for Remote Sensing (CITARS) project consisted of data quality evaluation, geometric correction, multitemporal registration, and location of section and field coordinates. Geometric correction was performed to facilitate accurate locations of section and field coordinates. Registration of the data from two or more ERTS-1 passes over the same scene was required for multitemporal data analysis procedures and for reduction of the number of times which section and field coordinates had to be located. With registered data, the desired coordinates needed to be found only once.

The procedures were designed to allow each institution participating in CITARS to use common training and test field boundaries and to duplicate ERTS-1 data tapes. Such a procedure was followed to permit more meaningful performance comparisons and to eliminate needless duplication of tasks and resources at each institution.

This report summarizes the ERTS-1 data acquisition and its quality; describes the procedures followed for reformatting, geometric correction, multitemporal registration, and location of section and field coordinates; and evaluating those procedures.

Page 2 is blank

2.0 DATA ACQUISITION SUMMARY

The data acquisition plan specified that ERTS-1 data from cycles 18 through 24 should be requested from the National Aeronautics and Space Administration/Goddard Space Flight Center (NASA/GSFC) and reformatted at the Laboratory for Applications of Remote Sensing (LARS) into the LARSYS data format for each of the six 8- by 32-kilometer (5- by 20-mile) test sites used in CITARS. Figure 1 shows the test site locations, and table I illustrates the dates of coverage of each test site. Each test site was purposely located in the ERTS-1 coverage overlap zone. Thus, each test site is covered on 2 consecutive days of each ERTS-1 cycle.

The ERTS-1 multispectral scanner (MSS), system-corrected computer-compatible tapes (CCT's) including coverage of the test area segments for which NASA/GSFC reported cloud coverage of 70 percent or less were placed on standing order for shipment to NASA/Lyndon B. Johnson Space Center (JSC). Frames reported as including more than 70 percent cloud cover were screened on microfilm copy at NASA/JSC, and CCT's were ordered for frames including test segments that were significantly cloud free.

Results of ERTS-1 MSS data acquisition efforts are shown in table II. From the 84 segment-period-passes or chances for collecting an acceptable data set, 26 data sets from 20 segment-periods were determined as being acceptable, leaving 22 segment-periods unrepresented by good data. Data sets were not included for analysis for one of the following reasons: (1) the ERTS-1 MSS did not operate during the overpass, (2) the test segment was divided between two frames of an orbit, and data from the other pass of the segment-period were acceptable, (3) the test segment coverage was incomplete (off the eastern or western frame edge), or (4) the test segment had greater than 30 percent cloud cover.

Reasons for data rejections are indicated in table II. Fifty-three data sets were rejected because of cloud cover, two because of incomplete segment coverage, two because of data unavailability from NASA/GSFC, and one because of its being split between two consecutive frames. The split-frame data set could have been handled, but the work required was not warranted because good-quality data were available from the preceding pass of the segment-period.

One deviation from the *Task Design Plan* (ref. 1) involved delivery of ERTS-1 CCT's to LARS. Shipment of CCT copies from NASA/JSC to LARS was not required because many of the required CCT's were being shipped directly to LARS from NASA/GSFC by another project contract involving the same geographical test areas as the ones in CITARS. This modified procedure for receiving ERTS-1 CCT's caused no problems.

TABLE I.– ERTS-1 COVERAGE SCHEDULE FOR CITARS SEGMENTS

ERTS-1 cycle	Month, 1973	Period	Date of overflight along-track[a]				
			L	M	N	O	P
18	June	I	8	9	10	11	12
19	June	II	26	27	28	29	30
20	July	III	14	15	16	17	18
21	Aug.	IV	1	2	3	4	5
22	Aug.	V	19	20	21	22	23
23	Sept.	VI	6	7	8	9	10
24	Sept.	VII	24	25	26	27	28

Counties covered:

Huntington
Shelby
Indiana

White
Livingston
Fayette
Lee
Illinois

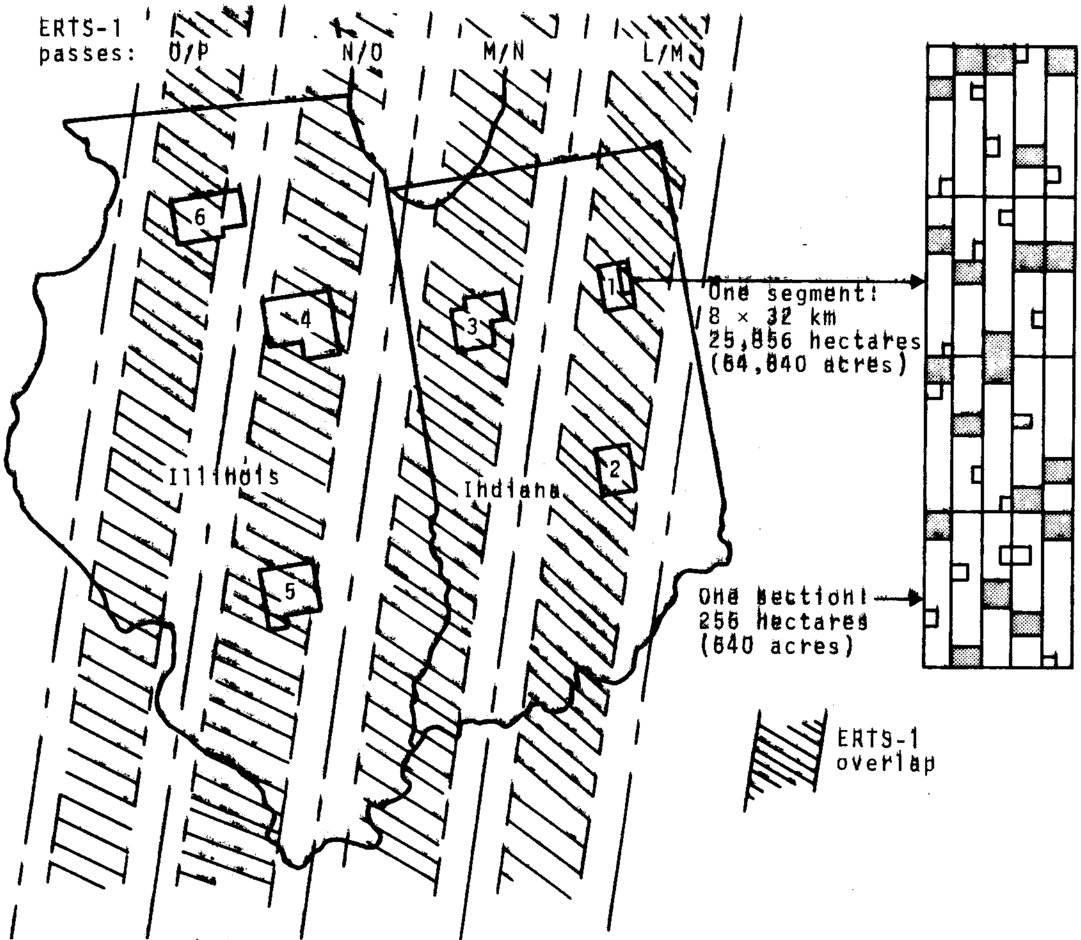
[a]The first pass, L, occurs on the first day of the period; M, the second; N, the third; O, the fourth; and P, the fifth. Passes L through P lie alphabetically side by side from east to west over Indiana and Illinois. Adjacent passes overlap slightly.

TABLE II.— SUMMARY OF ERTS-1 MSS DATA ACQUISITION

Segment	Pass	Period, 1973						
		I 6/08-12/73	II 6/26-30/73	III 7/14-18/73	IV 8/01-05/73	V 8/19-23/73	VI 9/06-10/73	VII 9/24-28/73
Huntington	1	*1320-15534 73037601(0)	(100)	(20)	(100)	No data collected	(80)	1428-15520 73087602(5)
	2	*1321-15593 73033707(0)	(100)	1357-15590 73046407(0)	(90)	(50)	(40)	(100)
Shelby	1	*1320-15541 73037701(0)	(90)	(80)	(100)	No data collected	(80)	1428-15523 73087402(0)
	2	Incomplete coverage	(80)	Incomplete coverage	(80)	1393-15581 73061203(30)	1411-15581 73067007(5)	(70)
White	1	*1321-15593 73033705(0)	(100)	(70)	(90)	(75)	1411-15581 73067005(20)	(100)
	2	Incomplete coverage	(80)	(70)	(80)	1394-16042 73060705(0)	(90)	(80)
Livingston	1	*1322-16051 73033601(0)	(80)	1358-16045 73047602(0)	1376-16043 73052902(20)	1394-16042 73060707(30)	(90)	(100)
	2	(30)	1341-16104 73047802(20)	Segment broken (5)	(100)	(50)	(100)	(80)
Fayette	1	1322-16054 73037402(0)	(50)	1358-16051 73046592(0)	(80)	1394-16044 73060802(0)	(90)	(50)
	2	*1323-16112 73039101(5)	1341-16111 73052002(0)	1359-16105 73054202(0)	(100)	(50)	(100)	(90)
Lee	1	(35)	(80)	*1359-16100 73052102(5)	(100)	(100)	(100)	(100)
	2	(100)	(40)	1360-16155 73066301(0)	1378-16153 73120202(0)	(90)	(50)	(100)

Legend:

xxxx-xxxxx	ERTS-1 scene identification number
xxxxxxxx	LARSYS run number
( )	Percent cloud cover
*	Base run for image registration



Study Area Counties:

Indiana	Illinois
1. Huntington	4. Livingston
2. Shelby	5. Fayette
3. White	6. Lee

Data Acquisition Periods:

0 – 5/21-25/73	IV – 8/01-05/73
I – 6/08-12/73	V – 8/19-23/73
II – 6/26-30/73	VI – 9/06-10/73
III – 7/14-18/73	VII – 9/24-28/73

Ground Truth:

ASCS – 20 quarter sections (white) each ERTS-1 pass

Photointerpretation – 20 sections (black) each ERTS-1 pass

Figure 1.– Location of test sites.

Page 8 is blank

3.0 DATA QUALITY EVALUATION

A majority of the selected data was of good quality. However, a few problems affecting either data analysis procedures or results were observed: (a) occasional erratic data throughout individual scan lines or portions of lines, (b) detector-to-detector differences among the mean values obtained from the six detector channels that comprise each spectral band as averaged over a large sample of the data, and (c) differences in the variances observed from the detector channels over the same data sample.

3.1 BAD SCAN LINES

The quality of ERTS-1 data was assessed by several different methods. First, each spectral band was visually inspected on a digital display to determine the presence of any lines of bad data through one of the 8- by 32-kilometer (5- by 20-mile) segments. More than half (14) of the 26 sets had no bad lines, and the worst sets were one with 8, two with 19, one with 25, and one with 40 bad lines (table III).

The data set with 40 bad scan lines was analyzed using only bands 5, 6, and 7. For the remainder of the data sets, entire fields were deleted if the bad scan line points made up a majority of the data, and bad scan line points were deleted from the field coordinate cards which represented mostly good data.

3.2 VARIATIONS AMONG DETECTORS

Next, histograms and sample statistics (mean and standard deviation) were computed for samples of the data — every line and every thirtieth point for all cases, every line and every sixth point for many, and every line and every point for a few. These statistics were calculated separately for each detector in each spectral band; unrotated ERTS-1 data were utilized for these tests. The largest differences in mean values of any two detector outputs are shown in table IV.

One would expect some variation between values in the various detectors because each detector is calibrated separately. To evaluate the degree of similarity between these statistics, a mean, m_μ , of the six detector means and the sample standard deviation, s_μ , of the individual values from that overall mean were calculated for each spectral band. The ratio, s_μ/m_μ, was computed for each data set; a histogram of these values is presented in figure 2. All values lie below 3 percent, except for one of 9.3 percent, which corresponds to the data set with 40 lines of bad data. No clear relationship could be found between the number of bad lines and s_μ/m_μ values below 3 percent. The number of good lines present was sufficient to mask effects of a few bad lines, and channel-to-channel variations existed in all data.

Similarly, the detector value standard deviations were analyzed; a histogram of s_σ/m_σ is presented in figure 2. In this instance, the s/m values exhibit considerably more spread than they do for the detector means. Most values are

$\leq$5 percent, but they scatter up to 24 percent, with an extreme of 87 percent for the 40-bad-line case. Here again, except for extreme values, no direct correlation between s_σ/m_σ and the number of bad lines present was apparent.

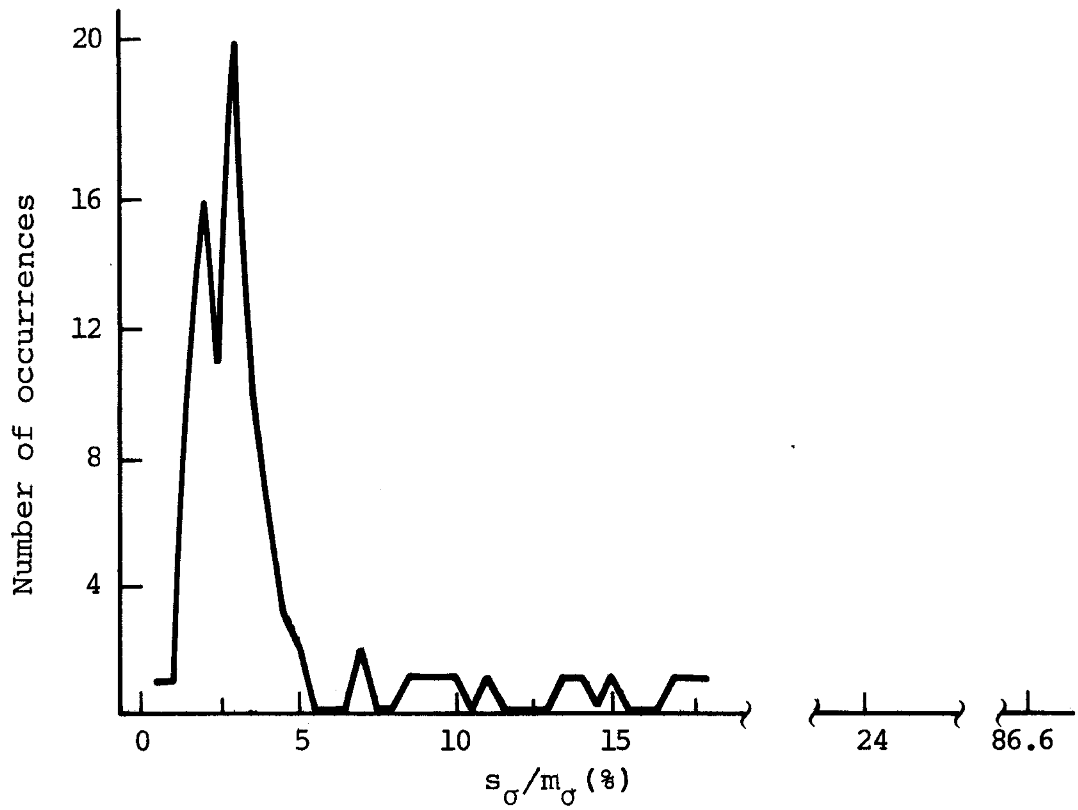
TABLE III.— REPORT OF BAD SCAN LINES FOR ERTS-1 MSS DATA

LARSYS run	Segment	Period	LARSYS data tape channel number			
			1	2	3	4
73033601	Livingston	I	—	—	—	—
73037402	Fayette	I	—	1	—	—
73039101	Fayette	I	1	—	—	—
73033705	White	I	—	—	—	—
73037601	Huntington	I	—	—	—	—
73033707	Huntington	I	—	—	—	—
73037701	Shelby	I	—	—	—	—
73047802	Livingston	II	—	1	9	40
73052002	Fayette	II	1	—	7	11
73052102	Lee	III	—	—	—	—
73066301	Lee	III	40	—	—	—
73047602	Livingston	III	—	—	3	—
73046502	Fayette	III	—	1	10	1
73054202	Fayette	III	—	—	—	—
73046407	Huntington	III	—	1	—	—
73052902	Livingston	IV	—	—	—	—
73120202	Lee	IV	—	—	—	—
73060707	Livingston	V	—	—	—	—
73060802	Fayette	V	—	—	—	—
73060705	White	V	—	—	—	—
73061203	Shelby	V	—	—	—	—
73067005	White	VI	—	—	—	—
73067007	Shelby	VI	—	—	—	—
73087602	Huntington	VII	—	—	—	—
73087402	Shelby	VII	—	—	—	—

TABLE IV.— PEAK DETECTOR MEAN DIFFERENCES FOR ERTS-1 MSS DATA

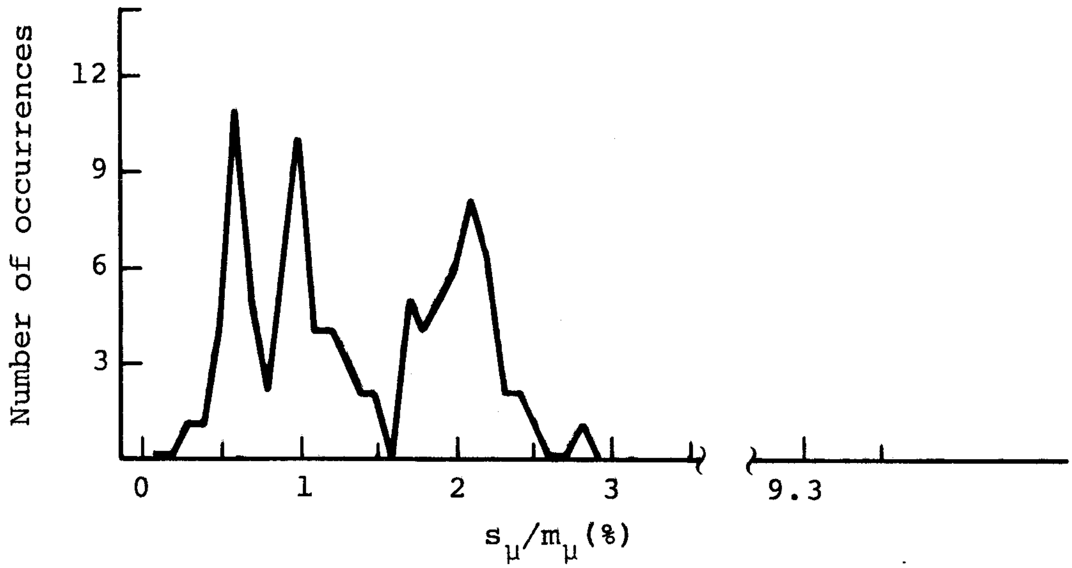
LARSYS run[a]	Segment	Period	Peak detector mean difference			
			Channel 1	Channel 2	Channel 3	Channel 4
73033600(01)	Livingston	I	0.511	0.445	1.590	1.080
73037400(02)	Fayette	I	.766	.859	1.497	1.487
73039100(01)	Fayette	I	1.058	1.232	2.017	1.493
73033702(05)	White	I	.889	.755	1.397	1.307
73037600(01)	Huntington	I	.618	.589	2.122	1.390
73033701(07)	Huntington	I	.857	.784	1.931	1.433
73037700(01)	Shelby	I	.692	.961	1.543	1.428
73047800(02)	Livingston	II	.998	.662	1.548	1.669
73052000(02)	Fayette	II	.735	.637	2.207	1.760
73052100(02)	Lee	III	.557	.252	2.226	1.795
73066300(01)	Lee	III	8.832	.424	1.960	1.721
73047600(02)	Livingston	III	.853	.578	2.374	1.700
73046500(02)	Fayette	III	.709	.313	1.532	1.777
73054200(02)	Fayette	III	1.139	.370	2.911	1.942
73046402(07)	Huntington	III	.767	.668	2.832	1.639
73046401(10)	Shelby	III	.909	.368	2.742	1.845
73052900(02)	Livingston	IV	.955	.574	3.471	1.928
73120200(02)	Lee	IV	.805	.489	2.301	1.834
73060700(07)	Livingston	V	.598	.472	2.484	1.856
73060800(02)	Fayette	V	.798	.437	2.722	1.892
73060701(05)	White	V	.508	.451	2.340	2.106
73061200(03)	Shelby	V	.542	.807	2.870	1.809
73067002(05)	White	VI	.607	.390	2.213	1.898
73067001(07)	Shelby	VI	1.038	.564	2.223	1.784
73087400(02)	Shelby	VII	1.207	.608	.630	1.295
73087600(02)	Huntington	VII	1.334	.941	1.178	1.400

[a]Numbers within parentheses represent the percentage of cloud cover during each data acquisition period.



(a) Differences in signal standard deviations.

Figure 2.— Variations among the six detector channels in single bands of ERTS-1 MSS data. [Originally prepared by the Environmental Research Institute of Michigan (ERIM), ref. 2]



(b) Differences in signal means.

Figure 2.– Concluded.

Page 16 is blank

4.0 REFORMATTING

The standard ERTS-1 CCT to LARSYS III procedures described in the *Task Design Plan* were followed for all CITARS data. Only the portion of the ERTS-1 frame containing the segment was reformatted. This aspect of the reformatting was initiated by supplying latitude and longitude coordinates to the reformatting process. Coordinates used (given in table V) exceed actual test site boundaries 8 to 19 kilometers (5 to 12 miles) to allow for tolerances in the ERTS-1 frame center coordinate values given on the ERTS-1 CCT's and effects of frame rotation. After reformatting of test site area data selected as base data, test site boundaries were located with a 3-kilometer (2-mile) outside margin. This boundary location became the definition of data included in the final version of each selected data set. One exception was the Fayette data set, which was defined with larger than 3-kilometer margins because of the urgency in processing it.

TABLE V.— TEST SITE COORDINATES USED FOR REFORMATTING DATA FROM CCT TO LARSYS III FORMAT

Test site	Northwest corner		Southeast corner	
	Longitude	Latitude	Longitude	Latitude
Huntington	85°42'	41°05'	85°21'	40°34'
Shelby	85°57'	39°47'	85°37'	39°17'
White	87°03'	41°02'	86°42'	40°28'
Livingston	88°48'	41°12'	88°24'	40°38'
Fayette	89°19'	39°16'	88°59'	38°46'
Lee	89°35'	41°59'	89°13'	41°29'

5.0 MULTITEMPORAL REGISTRATION

Data preparation plans called for selection of a base data set for each of the six test sites to which all other later acceptable data sets would be registered. Thus, for each test site one ERTS-1 line/column coordinate grid became standard. This procedure was adopted so that field boundary coordinate selection would only be required only one time for each test site. Registrations are referred to as "virtual" because they were stored as separate LARSYS four-channel data runs rather than being combined as one run having 16 to 24 channels.

Image registration procedures given in the CITARS *Task Design Plan* and reference 3 were followed, except in certain difficult or problem cases in which one or both of two procedural modifications were implemented. A problem registration case arose when the image being registered did not numerically correlate satisfactorily with the base image. The first recourse taken was to enhance both images by a principal components transformation and then to correlate the best suited components of the transformation. Severe cases, however, required registration or correlation of the image with another registered image rather than with the base image itself. This, of course, was a last resort since registration errors are accumulative. Problem cases occurred most frequently when late August (Period V) data were registered with early June (Period II) base data.

In March 1974, an error was discovered in the registration system program. The error caused a periodic duplication of data lines. That is, data which would have been in line N were being replaced with data of line N-1. This error affected 11 completed data sets. The data sets were reregistered, except for the Shelby Period III run, which had been deleted from the analysis schedule. The reruns did not cause a major time delay because replications of the image correlations were not required and because field coordinates had not yet been found satisfactory.

Root-mean-square (rms) values for checkpoint residuals of each overlay (shown in table VI) were computed. These values do not explicitly denote registration accuracy but can be used as accuracy indicators. The values were computed as the rms value of the residuals of the least squares second-order polynomial fit as applied to the checkpoint sets (discussed in appendix A). These values are actually indicators of how well the derived polynomial fits the checkpoints. Values of 0.5 are considered acceptable, and values above 0.8 are considered marginally acceptable. In practice, new checkpoints were derived for cases with larger than 0.5 values to achieve the 0.5 standard when possible.

The effect of registration on classification performance was also evaluated in five segment-period combinations by locating coordinates in nonregistered data and by comparing classification results to those obtained from the registered data. These results will be presented in the final report on classification results.

TABLE VI.— RMS RESIDUALS OF THE LEAST SQUARES POLYNOMIAL FIT CALCULATION FOR REGISTRATION PROCESSING

Segment	Period	Number of checkpoints	Residual, rms	
			Lines	Columns
Huntington	I	30	0.44	0.36
	III	27	.31	.39
	VII	51	.30	.43
Shelby	III	23	.63	.38
	V	9	.27	.27
	VI	43	.30	.47
	VII	59	.31	.43
White	V	61	.32	.39
	VI	16	.28	.14
Livingston	II	32	.31	.44
	III	9	.44	.37
	IV	9	.16	.87
Fayette	I	41	.39	.33
	II	52	.37	.29
	III	53	.44	.34
	V	19	.57	.39
Lee	III	100	.34	.58
	IV	84	.32	.41

Page 22 is blank

6.0 GEOMETRIC CORRECTION

The digital form of the ERTS-1 data (CCT's) contains several geometric distortions. These distortions include scale differential, altitude and attitude variations, Earth rotation skew, orbit velocity change, scan time skew, nonlinear scan sweep, scan angle error, and frame rotation. The major errors concern the scale and skew. Also, rotation of the data to north orientation is highly desirable. A two-step process, developed by LARS for the geometric correction of ERTS-1 data over small areas, was applied to all CITARS data.

Briefly, the procedure uses four linear transformations to correct or adjust for horizontal and vertical scale differences, rotation, skew caused by Earth rotation, and output scale factors. The process assigns radiance values in a rescaled, rotated, deskewed coordinate system with data from the existing grid; that is, the raw ERTS-1 data. Because a fixed grid output device (the printer or digital display screen) is used, some interpolation is required to produce new samples. The point nearest the desired sample location is used to represent the value at the desired location (nearest neighbor rule). The procedures are fully described in the *Task Design Plan.*

When the data are printed on an 8-line-per-inch, 10-column-per-inch computer line printer, the resulting scale is approximately 1:24,000 and the image is north-oriented. Comparisons to topographic maps indicate about a 1- to 2-percent scale error. This format facilitated locating section and field boundaries since rectified photograph overlays of the same scale could be matched to the ERTS-1 imagery.

The actual matrices used to perform the geometric correction of respective CITARS test sites were applied to the data sets as discussed in the CITARS *Task Design Plan* and reference 4. Data values were selected from input images for each output element as follows.

$$\begin{bmatrix} C_{in} \\ L_{in} \end{bmatrix} = A \begin{bmatrix} C_{out} \\ L_{out} \end{bmatrix}$$

where

C_{in} = input image column

L_{in} = input image line

C_{out} = output image column

L_{out} = output image line

and

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

with values given below

$$\begin{bmatrix} 1.04722 & 0.34713 \\ -0.15539 & 0.92163 \end{bmatrix} \text{Fayette}$$

$$\begin{bmatrix} 1.03357 & 0.35246 \\ -0.15953 & 0.93160 \end{bmatrix} \text{Huntington}$$

$$\begin{bmatrix} 1.04530 & 0.35544 \\ -0.16181 & 0.91990 \end{bmatrix} \text{Lee}$$

$$\begin{bmatrix} 1.03349 & 0.35277 \\ -0.15977 & 0.93153 \end{bmatrix} \text{Livingston}$$

$$\begin{bmatrix} 1.03448 & 0.34855 \\ -0.15650 & 0.93240 \end{bmatrix} \text{Shelby}$$

$$\begin{bmatrix} 1.03360 & 0.35230 \\ -0.15941 & 0.93163 \end{bmatrix} \text{White}$$

Page 26 is blank

7.0 SECTION AND FIELD COORDINATE LOCATIONS

Locating section and field coordinates was a major task preparatory to the classification of ERTS-1 data. To ensure that registration errors did not interfere with identification performance of field centers, no picture elements (pixels) were chosen from a one-pixel guard row and column around the perimeter of each field. That is, at least one entire pixel lay between the true boundary of any field and the pixels chosen from that field. The effect of this conservative but necessary criterion for selecting field pixels greatly reduced the number of pixels available for training and testing the classifiers.

Selecting pixels under these conditions was the direct cause of the several iterations of locating field coordinates (described in the following paragraphs). On the one hand, more pixels were wanted; but on the other hand, it was necessary not to sample too closely to a field boundary. The result was that there was no way of leaving sufficient guard pixels unsampled so as to eliminate any question or doubt that selected pixels were truly field center pixels. Changing the placement of the overlay on the ERTS-1 imagery by a fraction of a pixel would considerably change the pixels included in the sample. As long as this was a subjective decision by the analyst, there was never complete agreement among all experiment participants that the correct pixels had been chosen. The solution to this problem was to move by steps toward the use of more objective and quantitative methods of determining the location of field boundaries. In retrospect, it would have been better to have started with the method finally used; but all the difficulties involved in using the manual methods were not predicted at the start of the task. In summary, the real problem was that experiment participants were trying to locate and select pixels more precisely than the ERTS-1, 80-meter spatial resolution allowed.

This task was first attempted with a manual method for location of fields in ERTS-1 data displayed as single-band, gray-scale, line printer maps. This required that field boundaries be easily distinguished in the imagery. When minimal spectral contrast existed among crop fields, non-supervised classifications were performed to produce an enhanced image. Whether gray-scale or computer-enhanced images are used, reasonably large fields are required to assure that the selection of pixels from within the field boundaries.

With the CITARS data, little contrast among fields of interest was evident because the first data were collected early in the growing season (June 8 through 12). It was important, however, to use the registration base data because any registration errors would be cumulative. For instance, at this time of the year, corn and soybeans were only a few centimeters tall, the spectral response was primarily from the soil, roads were not as visible in the imagery as they generally are in data collected later in the season, and many fields were small ($\leq$8 hectares, or $\leq$20 acres). Therefore, when individual fields could not be clearly seen in the imagery, procedures for accurately locating fields were required to meet project requirements.

To improve the accuracy of the manual location method, ERTS-1 images were geometrically corrected and rescaled to produce a nominal 1:24,000-scale map on a line printer.

This product alone made the location of fields more precise and more rapid than it would have been on uncorrected data. Photograph overlays with outlines of the section and field boundaries were prepared by the Earth Observations Division (EOD). The initial overlays, made from photographs enlarged to a nominal scale of 1:24,000, were helpful but not completely satisfactory because of distortions in the photographs. Rectified photographs that could be manually overlaid to the line printer maps of the ERTS-1 data were then produced at scales of 1:24,000.

After all field and section coordinates in the ERTS-1 data were manually located, the precision was still not adequate to meet the requirement of a maximum error of one pixel. Therefore, ERIM used a previously developed computer-assisted procedure to locate section corners and to define ERTS-1 data coordinates for sections (ref. 5). A map transformation from Earth coordinates on a rectified aerial photograph to ERTS-1 data coordinates was calculated for each segment with approximately 30 control points for each calculation. The control points were located visually in the rotated and geometrically corrected ERTS-1 data and by coordinate digitization on the photograph. A map transformation was then computed by the method of least squares; ERTS-1 coordinates of the few control points with large residuals (>1 pixel) were checked and modified or deleted, as appropriate, and the transformation was recomputed. Next, the transformation was applied to all section corners of interest (whose locations on the photograph had been digitized at the same time as the control points) to find their fractional line and column coordinates in the ERTS-1 data. Final standard errors of estimate (for control points) were less than 0.5 and typically between 0.2 and 0.4 ERTS-1 pixels; that is, 15 to 30 meters on the ground. The rms error in digitizing the location of the individual points was about 3 meters on the ground (errors of approximately 0.013 centimeter or less on a photograph at a scale of 1:24,000).

These section corner coordinates (calculated in fractional ERTS-1 line and column coordinates) were used to locate field boundaries of individual fields within the sections. A major advantage of the procedure is that it preserves the relative positions of all points considered with an accuracy that cannot be matched manually. Another feature of the ERIM procedure was utilized to generate ERTS-1 data coordinates for each outlined section. All pixels whose centers fell inside lines connecting the vertices (again, located by fractional coordinates) were automatically included on coordinate definition cards.

After section and field coordinates had been located in the base period data set, imagery for the later ERTS-1 data was examined for cloud cover. The primary images were the three-band color composites produced by EOD. Clouds and shadows showed up very clearly on this imagery. If that imagery was not available for a particular data set, ERTS-1 band-5 imagery displayed on the video display screen at LARS was examined. If clouds were found, gray-scale printouts of ERTS-1 band 5 were produced with all pixels having a response less than a threshold shown as blanks and pixels with response greater than the threshold value with printed. The threshold response for clouds was typically 115 but was varied until the size, shape, and location of the clouds matched that observed on the hard-copy imagery. Shadows were drawn manually on the printer output maps to match the size, shape, and location of the shadows seen on the three-band color imagery. All fields and sections having clouds or shadows in them were deleted from the deck of section and field coordinate cards for that data set.

Page 32 is blank

8.0 CONCLUSIONS AND RECOMMENDATIONS

Two conclusions for future MSS data preparation can be made from the CITARS experience. One result is outstanding; 25 ERTS-1 overpasses of six test sites were reformatted into a single grid system for each test site and made available to analysts in a form that necessitated locating field boundaries in each test site only once. The success of this method for data preparation demonstrates the feasibility and practicality of the method. Two recommendations are suggested for future similar tasks: (1) base data sets for the registration process should be selected from the middle of the data collection time frame rather than from the beginning, and (2) registration and geometric correction processing should be combined into a single operation.

The second major conclusion to be drawn from the CITARS results is that manual methods are generally inadequate for precise and accurate location of field coordinates in ERTS-1 imagery. Improved results can be obtained by transforming digitized map or photograph coordinates to ERTS-1 coordinates.

Page 34 is blank

9.0 REFERENCES

1. CITARS Task Design Team: *Task Design Plan for the Crop Identification Technology Assessment for Remote Sensing (CITARS) Phase of the Large Area Land Resources and Crop Acreage Inventory Applications System Verification Test Project.* Oct. 1973.

2. Hall, F. G.; Bauer, M. E.; and Malila, W. A.: "First Results from the Crop Identification Technology Assessment for Remote Sensing (CITARS)." Paper presented at 9th Internat. Symp. on Remote Sensing of Environment (Ann Arbor, Mich.), Apr. 1974.

3. Anuta, P. E.: "Spatial Registration of Multispectral and Multitemporal Digital Imagery Using Fast Fourier Transform Techniques." *IEEE Trans. Geoscience Electronics*, vol. 8, Oct. 1970, pp. 353-368.

4. Anuta, P. E.: "Geometric Correction of ERTS-1 Digital Multispectral Scanner Data." LARS Information Note 103073, 1973.

5. Malila, W. A.; Hieber, R. H.; and McCleer, A. P.: "Correlation of ERTS MSS Data and Earth Coordinate Systems." Proc. of Conference on Machine Processing of Remotely Sensed Data. ERIM note no. 193300-18SA/J, Oct. 1973.

6. Second Monthly Status Report to EOD Management on the CITARS Phase of the LAP ASVT as of September 1, 1973.

7. Third Monthly Status Report to EOD Management on the CITARS Phase of the LAP ASVT as of October 4, 1973.

8. Fourth Monthly Status Report to EOD Management on the CITARS Phase of the LAP ASVT as of November 3, 1973.

APPENDIX A

SUMMARY OF DATA PREPARATION BY PERIOD

APPENDIX A

SUMMARY OF DATA PREPARATION BY PERIOD

In the following paragraphs, procedures, problems, and results are discussed for each data collection period.

A.1 PERIOD I (JUNE 8 THROUGH JUNE 12, 1973)

Both passes of the Fayette county test site yielded acceptable data. Pass 2 data were selected for analysis and as the base data set for the test site. Selection was based on the facts that the pass 2 data set had a bad data line in channel 2 and that the ERTS-1 frame was generally more cloudy. Both passes of the Huntington test site were acceptable, and pass 1 was arbitrarily selected as the base data set. Pass 2 was later registered and made available for analysis. Pass 1 of the Livingston site was selected for analysis of cloud cover considerations, and passes 1 of the Shelby and White test sites were selected because passes 2 incompletely covered the sites.

Test site corner boundary coordinates for base data sets were determined such that the final reformatted product would have a 3-kilometer margin around the perimeter of the 8- by 32-kilometer sites. Following coordinate selection, the base data sets were geometrically corrected for systematic errors. Fayette and Lee county data sites were corrected using a slightly different correction matrix than was used for White, Shelby, Livingston, and Huntington sites. After correction of the latter test sites, the correction program was modified, reversing the sequence of two of the five correction matrix factors.

A.2 PERIOD II (JUNE 25 THROUGH JUNE 30, 1973)

Because of severe cloud cover, acceptable data were collected for only two test sites, Livingston and Fayette. The Livingston data set had numerous bad data lines in channels 3 and 4 in addition to 30 percent cloud cover. This data set was registered to Period I base counterpart with difficulty because of the "popcorn" cloud formations, which necessitated extensive editing of image correlation results. The Fayette data set included several bad data lines in channels 3 and 4 but did not have significant cloud cover.

A.3 PERIOD III (JULY 14 THROUGH JULY 18, 1973)

Period III data sets were registered to base counterparts with difficulty caused by greatly changing scene characteristics. To achieve registered data sets, images were correlated with a decreased interval between correlation blocks. Thus, more correlation attempts were made; and in one case, correlation of linear combinations of data set channels was required. Finally, images were registered with only marginally adequate numbers of acceptable correlation checkpoints.

The Shelby data set required correlation of linear combinations of channels to achieve adequate registration. Combinations were made by the principal components method. This data set was later dropped from analysis because of a lack of complete site coverage. For the first time, acceptable data were collected over the Lee test site, and a base data set was established. Actually, both passes were relatively cloud free, and some question was raised as to which pass should be selected for analysis.

A.4 PERIOD IV (AUGUST 1 THROUGH AUGUST 5, 1973)

Two test sites were acceptably free of clouds, Livingston and Lee. Registration procedures required principal components image transformation before successful correlations could be achieved.

A.5 PERIOD V (AUGUST 19 THROUGH AUGUST 25, 1973)

Four data sets were reformatted for Period V: Fayette, Shelby, Livingston, and White counties. Correlation difficulties were most severe for these data sets. Fayette was registered by correlating with the Period III registered data set (a last resort measure). Registrations for the other test sites were achieved by using principal components transformed data.

A.6 PERIOD VI (SEPTEMBER 6 THROUGH SEPTEMBER 10, 1973)

Data for two test sites, White and Shelby, were processed. Data of this period correlated well with the base data, and no transformations were required.

A.7 PERIOD VII (SEPTEMBER 24 THROUGH SEPTEMBER 28, 1973)

Two segments, Huntington and Shelby, were cloud free during this period; and the data were processed by standard procedures.

APPENDIX B

AIRCRAFT SCANNER DATA PROCESSING

APPENDIX B

AIRCRAFT SCANNER DATA PROCESSING

Aircraft data were scheduled to be collected by the Bendix M^2S and ERIM M-7 scanner systems with flight dates coordinated with the ERTS-1 overpasses. To take advantage of improved weather conditions, the schedules were made more flexible than those given in figure 2. Six data acquisition missions were scheduled for the M^2S system and two for the M-7 system. The missions were to be flown as coincidently as possible to ERTS-1 cycles 19 through 24 for M^2S and to ERTS-1 cycles 19 and 22 for M-7. Aircraft coverage within 4 days of the ERTS-1 overpass with less than 10 percent cloud cover and sun angle greater than 40° was specified as highly desirable. Data would be acceptable if acquired 5 to 8 days after the overpass with less than 30 percent cloud cover and sun angle greater than 30°.

The Bendix M^2S scanner system collected data on six missions. As reported in references 6, 7, and 8, the coverage and quality of the data were not acceptable, and use of the data would require redesign of the CITARS task. No Bendix M^2S data have been received at LARS. Reported problems in the data include misregistration of scanner channels, noisy scanner channels, intermittent saturated signals, and scanner data dropout.

The ERIM M-7 scanner system data were collected for two missions as planned. Table B-1 shows the dates and flight lines of the missions. All data shown in the table have been received at LARS in ERIM 2.54-centimeter analog tape format. As noted (ref. 6), portions of three test segments were not covered for Mission 1. Portions of the data reviewed at LARS appeared to be good quality data. Five flight times not shown in table B-1 were collected and received at LARS. Flight line 6 of the Shelby test site of Mission 1 was incorrectly flown on two attempts on July 6, 1973, because of navigational difficulties. Because both flight lines of a segment must be collected in the same time frame, line 5 of Shelby was repeated on July 7, 1973, when line 6 was correctly flown. In addition, because cloud cover was more favorable over the Huntington site on July 7, 1973, it was repeated, making a total of five extra runs.

Because of problems in the Bendix M^2S aircraft data and in other aspects of the CITARS task, analysis of aircraft data was placed at a low priority. Preparation of ERIM scanner data was therefore halted before significant results were achieved. One flight line, however, was reformatted for demonstration. Flight line 8, flown over the Fayette County test area on August 21, 1973, was digitized and computer reformatted in a special manner. The objective was to demonstrate effects of line averaging. For this data set, the digitizing line interval for zero line-to-line overlap was 11 (every twelfth line used). To test effects of averaging, all scan lines were digitized in blocks of 12. Each block was input for generation of one output data line. The blocks were processed three times, producing three LARSYS data sets, each in geometric registration. The first was produced by using line 6 of each block (LARS run 73035800); the result is conventionally processed data. The second run was produced by averaging with equal weight lines 4 through 9 of each block to generate each output line (LARS run 73035801). The third run was generated by averaging all 12 lines of each block with equal weight (LARS run 73035802).

TABLE B-I.— SUMMARY OF ERIM M-7 SCANNER DATA ACQUISITION

Segment	Line	Date	Time G.m.t.[a]	Absolute altitude, (10^3 ft)	Ground heading	LARS run	Date	Time G.m.t.[a]	Absolute altitude, (10^3 ft)	Ground heading	LARS run
Huntington	1	7/07/73	14:38	13.7	South	730365	8/21/73	17:56	13.7	North	730564
	2	7/07/73	14:51	13.7	North	730366	8/21/73	18:12	13.7	South	730565
Shelby	5	7/07/73	13:58	13.7	North	730364	8/21/73	16:18	13.7	North	730561
	6	7/07/73	13:46	13.7	South	730363	8/21/73	16:07	13.7	South	730560
White	3	7/06/73	17:09	13.8	South	730360	8/21/73	17:06	13.8	North	730562
	4	7/06/73	16:52	13.8	North	730354	8/21/73	17:20	13.8	South	730563
Livingston	9	7/05/73	16:35	13.8	North	730353	8/20/73	20:11	13.8	North	730557
	10	7/05/73	16:22	13.8	South	730352	8/20/73	19:59	13.8	South	730556
Fayette	7	7/05/73	20:13	14.0	North	730355	8/21/73	14:43	14.0	North	730559
	8	7/05/73	19:59	14.0	South	730354	8/21/73	14:30	14.0	South	730558
Lee	11	7/05/73	15:42	13.7	North	730350	8/20/73	19:22	13.7	North	730555
	12	7/05/73	15:54	13.7	South	730351	8/20/73	19:09	13.7	South	730554

[a]Greenwich mean time.