

SYSTEM DEVELOPMENT OF THE SCREWWORM  
ERADICATION DATA SYSTEM (SEDS) ALGORITHM

NASA CR-  
147552

Job Order 92-105

(NASA-CR-147552) SYSTEM DEVELOPMENT OF THE  
SCREWWORM ERADICATION DATA SYSTEM (SEDS)  
ALGORITHM (Lockheed Electronics Co.) 85 p  
HC \$5.00 CSCL 06C

N76-22876

Unclas  
27396  
G3/51

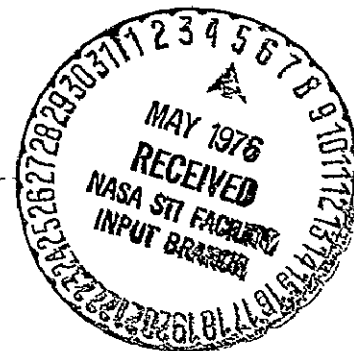
Prepared By

Lockheed Electronics Company, Inc.  
Aerospace Systems Division  
Houston, Texas

Contract NAS 9-12200

For

HEALTH APPLICATIONS GROUP  
LIFE SCIENCES DIRECTORATE



*National Aeronautics and Space Administration*  
**LYNDON B. JOHNSON SPACE CENTER**

*Houston, Texas*

January 1976

LEC-7646

SYSTEM DEVELOPMENT OF THE SCREWORM  
ERADICATION DATA SYSTEM (SEDS) ALGORITHM

Job Order 92-105


PREPARED BY

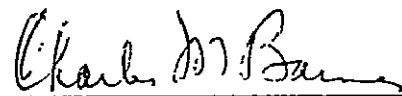
G. Arp  
F. Forsberg  
L. Giddings  
D. Phinney

APPROVED BY

LEC

NASA

  
Frank C. Forsberg, Manager  
Life Sciences Applications  
Department

  
Charles M. Barnes, Manager  
Health Applications Office

Prepared By

Lockheed Electronics Company, Inc.

For

Bioengineering Systems Division

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

January 1976

**TECHNICAL REPORT INDEX/ABSTRACT**  
(See instructions on reverse side.)

1. TITLE AND SUBTITLE OF DOCUMENT  SYSTEM DEVELOPMENT OF THE SCREWWORM ERADICATION DATA SYSTEM (SEDS) ALGORITHM	2. JSC NO.  JSC-10965
--	-----------------------------

3. CONTRACTOR/ORGANIZATION NAME  Lockheed Electronics Company, Inc.	4. CONTRACT OR GRANT NO.  NAS 9-12200
---	---

5. CONTRACTOR/ORIGINATOR DOCUMENT NO.  LEC-7646.	6. PUBLICATION DATE (THIS ISSUE)  January 1976
--	--

7. SECURITY CLASSIFICATION  Unclassified	8. OPR (OFFICE OF PRIMARY RESPONSIBILITY)  C. M. Barnes
--	---

9. LIMITATIONS GOVERNMENT HAS UNLIMITED RIGHTS <input checked="" type="checkbox"/> YES <input type="checkbox"/> NO  IF NO, STATE LIMITATIONS AND AUTHORITY  N/A	10. AUTHOR(S)  G. Arp F. Forsberg L. Giddings D. Phinney
--	---

11. DOCUMENT CONTRACT REFERENCES WORK BREAKDOWN STRUCTURE NO. N/A	12. HARDWARE CONFIGURATION SYSTEM N/A
---	---

CONTRACT EXHIBIT NO. N/A	SUBSYSTEM N/A
-----------------------------	------------------

DRL NO. AND REVISION N/A	MAJOR EQUIPMENT GROUP
-----------------------------	-----------------------

DRL LINE ITEM NO. N/A	N/A
--------------------------	-----

13. ABSTRACT

This document reports the results of a study to use remotely sensed data in the eradication of the screwworm and in the study of the role of the weather in the activity and development of the screwworm fly. As a result, the Screworm Eradication Data System (SEDS) Algorithm was developed.

**ORIGINAL PAGE IS  
OF POOR QUALITY**

14. SUBJECT TERMS

Remote Sensing Applications \_\_\_\_\_

Radiometric data \_\_\_\_\_

\_\_\_\_\_

## CONTENTS

Section	Page
1: INTRODUCTION. . . . .	1-1
1.1 <u>BACKGROUND</u> . . . . .	1-1
1.2 <u>PURPOSE</u> . . . . .	1-2
1.3 <u>DEVELOPMENT OF THE SYSTEM</u> . . . . .	1-2
1.3.1 DATA AVAILABLE FOR ANALYSIS. . . . .	1-3
1.3.2 DATA USED TO DEVELOP SCREWORM MODEL . . . . .	1-4
1.4 <u>EVALUATION OF THE SYSTEM</u> . . . . .	1-9
2. CALCULATION OF SEDS SYSTEM PARAMETERS . . . . .	2-1
2.1 <u>DAILY MEAN AIR TEMPERATURE (DMAT)</u> <u>COEFFICIENTS</u> . . . . .	2-1
2.1.1 BACKGROUND MATERIAL. . . . .	2-1
2.1.1.1 <u>Theory and Development</u> . . . . .	2-3
2.1.1.2 <u>SEDS Utilization</u> . . . . .	2-6
2.1.2 METHOD OF CALCULATION. . . . .	2-8
2.1.2.1 <u>Data Used</u> . . . . .	2-8
2.1.2.2 <u>Method of Analysis</u> . . . . .	2-12
2.1.3 VERIFICATION . . . . .	2-21
2.2 <u>RAINFALL COEFFICIENTS</u> . . . . .	2-24
2.2.1 BACKGROUND MATERIAL. . . . .	2-24
2.2.1.1 <u>Theory</u> . . . . .	2-25
2.2.1.2 <u>SEDS Utilization</u> . . . . .	2-26
2.2.2 <u>METHOD OF CALCULATION</u> . . . . .	2-27
2.2.2.1 <u>Data Used</u> . . . . .	2-28
2.2.2.2 <u>Analysis Used</u> . . . . .	2-28
2.2.3 VERIFICATION . . . . .	2-33

Section	Page
2.3 <u>THERMAL THRESHOLDS</u> . . . . .	2-34
2.3.1 BACKGROUND MATERIAL. . . . .	2-34
2.3.2 METHOD OF CALCULATION. . . . .	2-36
2.3.2.1 <u>Data Used</u> . . . . .	2-37
2.3.2.2 <u>Analysis Used</u> . . . . .	2-37
2.3.3 VERIFICATION . . . . .	2-40
2.4 <u>COMBINED SCREWORM PRODUCT COEFFICIENT</u> . . . . .	2-40
2.4.1 BACKGROUND MATERIAL. . . . .	2-40
2.4.1.1 <u>Independent Variables</u> . . . . .	2-41
2.4.1.2 <u>Model Selection</u> . . . . .	2-45
2.4.1.3 <u>Model Verification</u> . . . . .	2-47
2.4.1.4 <u>Developmental Phase Coefficients</u> . . . . .	2-48
2.4.2 COEFFICIENT CALCULATION AND VERIFICATION . . . . .	2-50
3. CALCULATION OF CMI. . . . .	3-1
3.1 <u>BACKGROUND MATERIAL</u> . . . . .	3-1
3.1.1 SOIL MOISTURE MODEL. . . . .	3-1
3.1.2 SEDS UTILIZATION . . . . .	3-2
3.2 <u>DATA HANDLING</u> . . . . .	3-5
3.3 <u>VERIFICATION</u> . . . . .	3-5
3.4 <u>PREPARATION OF CMI TAPES</u> . . . . .	3-6
4. PROCESSING OF EMISSIVITY ZONE MAPS. . . . .	4-1
4.1 <u>INTRODUCTION TO THE PROBLEM</u> . . . . .	4-1
4.2 <u>DEVELOPMENT OF THE EMISSIVITY ZONE MAP</u> . . . . .	4-1
4.3 <u>METHODS OF PRODUCTION OF THE SEDS EMISSIVITY MAP</u> . . . . .	4-3

Section	Page
4.3.1 SOURCES OF EMISSIVITY MEASUREMENTS. . . . .	4-3
4.3.2 SOURCES OF RADIOMETRIC DATA . . . . .	4-3
4.3.3 SELECTION OF IMAGES FOR USE IN COMPOSITING . . . . .	4-5
4.3.4 CREATION OF COMPOSITED IMAGES TAPE, . . . . .	4-5
4.3.5 CLUSTER ANALYSIS OF COMPOSITED DATA . . . . .	4-6
4.3.6 PRODUCTION OF THE FIRST HARD COPY OF THE EMISSIVITY ZONES, . . . . .	4-7
4.3.7 PRODUCT DISPLAY . . . . .	4-7
4.3.8 ASSIGNMENT OF EMISSIVITY VALUES TO ZONES. . . . .	4-7
4.3.9 CURRENT STATUS. . . . .	4-8
5. REFERENCES . . . . .	5-1

## TABLES

Table		Page
2-I	IMPROVEMENT OF DMAT ESTIMATES UPON ADDITION OF ELEVATION AS A REGRESSION VARIABLE TO JULY 1975 SEDS DATA. . . . .	2-6
2-II	PARTIAL LIST OF METEOROLOGICAL STATION STATISTICS PRINTOUT FOR 31 MARCH 1975. . . . .	2-9
2-III	SEDS PRIME VERIFICATION NETWORK. . . . .	2-10
2-IV	DAILY SURFACE DATA REPORT. . . . .	2-11
2-V	GROUND TRUTH ZONE CONTROL STATIONS . . . . .	2-13
2-VI	SAMPLE RADIOMETRIC DATA AS PREPARED FOR ANALYSIS . . . . .	2-14
2-VII	SAMPLE RADIOMETRIC DATA FROM TABLE 2-VI AFTER THERMAL THRESHOLDING AND CLOUD DISCRIMINATION . . . . .	2-17
2-VIII	MULTIPLE REGRESSION RESULTS FOR 15-29 MARCH 1975 FOR CASE 3 - ONLY NIGHT RADIOMETRIC DATA USABLE. . . . .	2-22
2-IX	15-29 MARCH 1975 DMAT COEFFICIENTS IN BOTH CODED AND UNCODED FORMS. . . . .	2-23
2-X	METEOROLOGICAL STATION STATISTICS PRINTOUT R CODE INTERPRETATION. . . . .	2-27
2-XI	TYPICAL STATISTICS PRODUCED BY ISOCLS/PRINT FOR 15-29 MARCH 1975 SEDS DATA (SELECTED STATIONS ONLY) . . . . .	2-30
2-XII	SAMPLE OUTPUT OF DISCRM FOR 15-29 MARCH 1975 SEDS DATA . . . . .	2-31
2-XIII	SCALED AND UNSCALED SEDS RAINFALL CLASSIFICATION COEFFICIENTS. . . . .	2-32
2-XIV	COMPARISON OF SURFACE REPORTS FROM TEXAS WITH SEDS CLOUD DETERMINATION FOR 5-11 JUNE 1975 DATA. . . . .	2-34

Table		Page
2-XV	TEMPERATURE OF FIRST CLOUD-FREE FREQUENCY MAXIMUM AND ESTIMATED THERMAL THRESHOLD FOR 28 SEPTEMBER TO 9 OCTOBER 1975 SEDS DATA . . . . .	2-37
2-XVI	F TEST OF SIGNIFICANCE FOR THREE POSSIBLE SCREWWORM REGRESSION MODELS. . . . .	2-48
2-XVII	COEFFICIENTS DERIVED FOR MULTIPLICATIVE MODEL. . . . .	2-49
2-XVIII	DEVELOPMENTAL PHASE COMBINED SCREWWORM PRODUCT COEFFICIENTS . . . . .	2-49
3-I	SURFACE METEOROLOGICAL STATIONS USED IN DEVELOPMENT OF CMI . . . . .	3-3
3-II	INTERPRETATION OF CMI VALUES IN TERMS OF FIELD CONDITIONS . . . . .	3-6



## FIGURES

Figure		Page
1-1	Average temperature during adult phase prior to first oviposition (C). . . . .	1-5
1-2	Generalized diagram of the life cycle of the screwworm fly, <i>Cochliomyia hominivorax</i> . . . . .	1-7
1-3	Data plot showing the number of screwworm cases and the weather. . . . .	1-8
2-1	SEDS System Parameter Control Form . . . . .	2-2
2-2	Comparison of satellite data with surface air temperature in the Fortín de las Flores test area . . . . .	2-4
2-3	Cloud discrimination decision field for March 1975 data. . . . .	2-16
2-4	Day radiometric data taken 1-10 August 1975 over the meteorological station network. . . . .	2-35
2-5	Histogram of day radiometric data 28 September - 9 October 1975 . . . . .	2-38
2-6	Histogram of night radiometric data 28 September - 9 October 1975 . . . . .	2-39
2-7	Relation between mean air temperature and growth per generation . . . . .	2-43
2-8	Relationship between CMI and growth potential. . . . .	2-44
2-9	Relationship between generation length and growth potential . . . . .	2-46
3-1	CMI station network. . . . .	3-4
4-1	Temperature error (°C) associated with an incorrect assumption of emissivity at 300° K . . . . .	4-2
4-2	Steps necessary for the creation of an emissivity zone map of Mexico. . . . .	4-4

## ACRONYMS AND ABBREVIATIONS

CMI	Crop moisture index
DAS	Data Analysis Station
DDSUM	Degree day sum
DMAT	Daily mean air temperature
GDS	Ground Data System Division
HAG	Health Applications Group
IR	Infrared
ITOS	Improved TIROS Operational Satellite
JSC	Johnson Space Center
LEC	Lockheed Electronics Company, Inc.
LTMAT	Long term mean air temperature
MPT	Mean pupation time
MSS	Multispectral scanner
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
SEDS	Screwworm Eradication Data System
STMAT	Short term mean air temperature
rms	Root mean square

## 1. INTRODUCTION

In 1973, the Health Applications Group (HAG) at the National Aeronautics and Space Administration (NASA) Johnson Space Center (JSC) began a program to evaluate the use of remotely sensed data as an additional tool in existing and projected insect eradication efforts. The need to relate remotely sensed data to screwworm infestations resulted in a two-part developmental study. First, the ability to use remotely sensed data to estimate weather conditions would be evaluated. Second, the effect of weather on screwworm populations would be modeled.

The first study, concerning the ability to use remotely sensed data, resulted in several reports indicating that the Improved TIROS Operational Satellite (ITOS) appeared to be the most desirable remote sensing satellite and that the Johnson Space Center had the necessary capability to process and register the data. The second study, concerning the screwworm and the role of the weather in the activity and development of the screwworm fly, also resulted in several reports. Information and illustrations from these reports will be included in the following text. Concurrently with the developmental studies, a production processor was designed and built.

### 1.1. BACKGROUND

The initiative for the joint NASA-U.S. Department of Agriculture (USDA) project on the eradication of screwworms was derived from Dr. Charles Barnes, PH. D, D.V.M., of NASA Johnson Space Center's Health Applications Group, Life Sciences Directorate. He was intensively familiar with the screwworm problem, having served earlier as an inspector with the USDA in Mexico. In addition, he was familiar with the untapped potential of using satellite data for health applications work, such as in support of screwworm eradication.

The Lockheed Electronics Company, Inc. (LEC), a support contractor, hired a group of multidisciplined scientists to provide space technology techniques to support an eradication program. Processing techniques were developed and test sites were established in Veracruz, Mexico and in Texas for purposes of collecting ground truth data. A production processing system, Screwworm Eradication Data System (SEDS), was designed and built by the NASA JSC Ground Data System Division (GDSD) and their support contractor Aeronutronic Ford. The basic algorithms were developed by the HAG/LEC and incorporated with some modifications into the SEDS design by GDSD/Aeronutronic Ford. Due to concern for operational problems, there was little active participation in the project by the USDA until after the initial algorithms were developed and the production system was completed and ready for evaluation.

## 1.2 PURPOSE

This report documents the development of the basic algorithms and explains how the various data are prepared for input into SEDS, as well as how operational coefficients are calculated. This report provides a knowledgeable programmer the necessary and available information with which to build a stand-alone system. It is intended to be used in conjunction with the basic SEDS documentation for purposes of technology transfer.

## 1.3 DEVELOPMENT OF THE SYSTEM

The development of the SEDS system was limited by the lack of reliable biological data for modeling. It was also limited by lack of funds, time, and availability of major pieces of computer equipment at JSC. Probably the most limiting factor in the SEDS processing system was the limited amount of computer memory, a factor which hampered the number of functions SEDS could perform.

The development of SEDS was and still is a cooperative effort; the operation of SEDS is also cooperative in that Aeronutronic Ford operates the SEDS on a production basis and LEC furnishes certain periodic operating coefficients and background correction maps which are produced for SEDS on other computers. It must be stressed that SEDS could be a stand-alone system provided that it had sufficient peripherals, memory core, and sufficient time for programming these other functions into SEDS.

### 1.3.1 DATA AVAILABLE FOR ANALYSIS

To quantify the effects of a parameter affecting screwworm growth, a statistical relationship between that parameter and the screwworm population must be shown. Among the important parameters that affect screwworm population dynamics are the availability of hosts, incidence of wounds, predators, efforts of the eradication program, daily temperatures, and soil moisture budget. Of these parameters, data on only the latter two were available and statistically reliable enough to warrant detailed analysis and comparison with screwworm population changes. Although screwworm populations could not be determined, the population changes were inferred from the number of screwworm cases.

The following data was used in the development of the screwworm growth potential functions.

<u>Year</u>	<u>State</u>	<u>Meteorological Divisions</u>
1968	Texas	5,6,7,9,10
1973	Arizona	1,3,4,6,7
1973	Texas	5,6,7,9,10
1974 (to 31 July)	Texas	5,9

Data was not analyzed outside of Texas except for Arizona (1973). The number of screwworm cases reported in 1967, 1969, 1970, and 1971 was low and was not used. The 1972 data represented a

large anomalous outbreak that remains largely unexplained, and was not used to develop the growth potential functions. A later test of Division 9 (1972) revealed that at least that division could have been used without affecting the final results. The data was censored to ten cases or more. Below ten cases, the data was not considered reliable, and therefore, was not used.

The meteorological data consisted of weekly mean air temperatures, total precipitation and crop moisture indexes for each of the divisions. The values used represent the average of data from all of the reporting stations within each division.

The screwworm cases were converted to growth per generation, that is, growth is defined as the ratio of the number of cases of a given week to the number of cases reported during the week in which the parent generation occurred.

A weekly plot was made of the meteorological data versus growth per generation as shown on figure 1-1. A curve was then hand drawn in a manner which, in the analyst's opinion, most closely fit the relationship of the two parameters being plotted. Each curve defined the growth potential for that meteorological parameter. Growth potential of a parameter is defined as the maximum likely growth, if that parameter is the only factor restricting screwworm growth.

### 1.3.2 DATA USED TO DEVELOP SCREWWORM MODEL

Extensive literature reviews of the effect of environmental conditions on the growth and development of the screwworm were used as the point of departure in producing a screwworm model. A brief review and summary of the surveys is contained in the

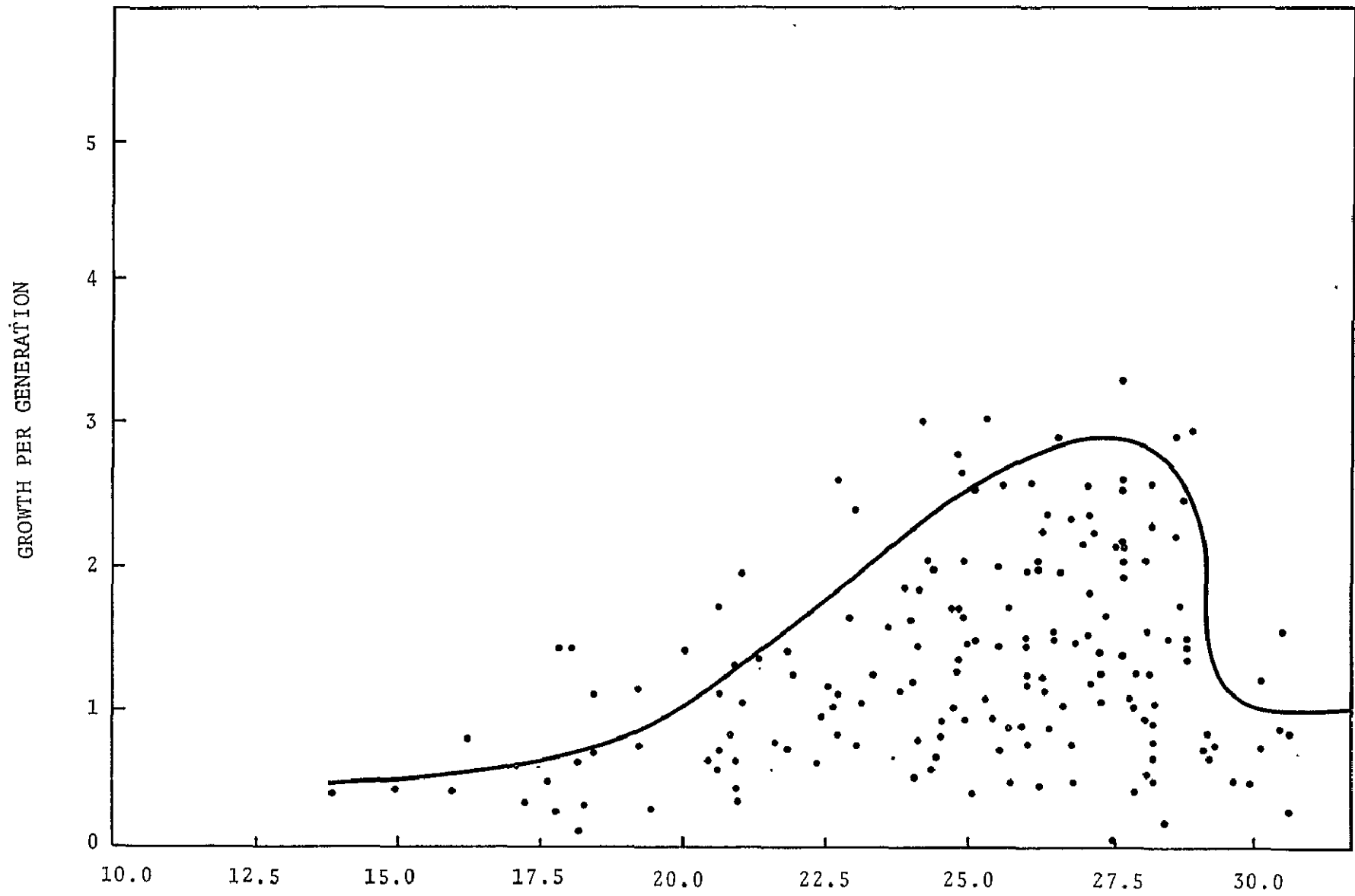


Figure 1-1. - Average temperature during adult phase prior to first oviposition (°C).

technical memorandum, "Selected Parameters Influencing the Growth Survival of *Cochliomyia hominivorax* (Coquerel)."

The generalized life cycle of the screwworm fly is presented in figure 1-2. Each stage of the life cycle has key factors which influences its development rate and which would have been desirable to incorporate into the model. However, the only statistically reliable data was on weather. A more complete description of the effects of environmental factors at each stage and some consideration for modeling is contained in the above-referenced technical memorandum.

The relationship of crop moisture index (CMI) and daily mean air temperature (DMAT) is clearly illustrated in figure 1-3. It shows that both temperature and CMI are the primary limiting factors affecting screwworm growth at different times of the season. One of the modeling considerations is the fact that the daily air temperature affects both the adult fly and the pupae in the ground. The newly hatched or adult fly are immediately affected by the existing air temperature, whereas the pupae are not, or only slightly, affected by the then existing air temperature. Pupae are affected more so by the daily mean air temperature over an extended period due to their slow development. Therefore, specifications for SEDS included multiple output products: long-term growth potential, short-term growth potential, CMI and degree day summation. These output products were to be used by trained USDA epidemiologists to assist them in planning fly drop missions. The epidemiologists were to have had the detailed knowledge of their area and other factors affecting screwworm growth, so that they could intelligently use the additional tools depicting environmental conditions especially in remote areas.



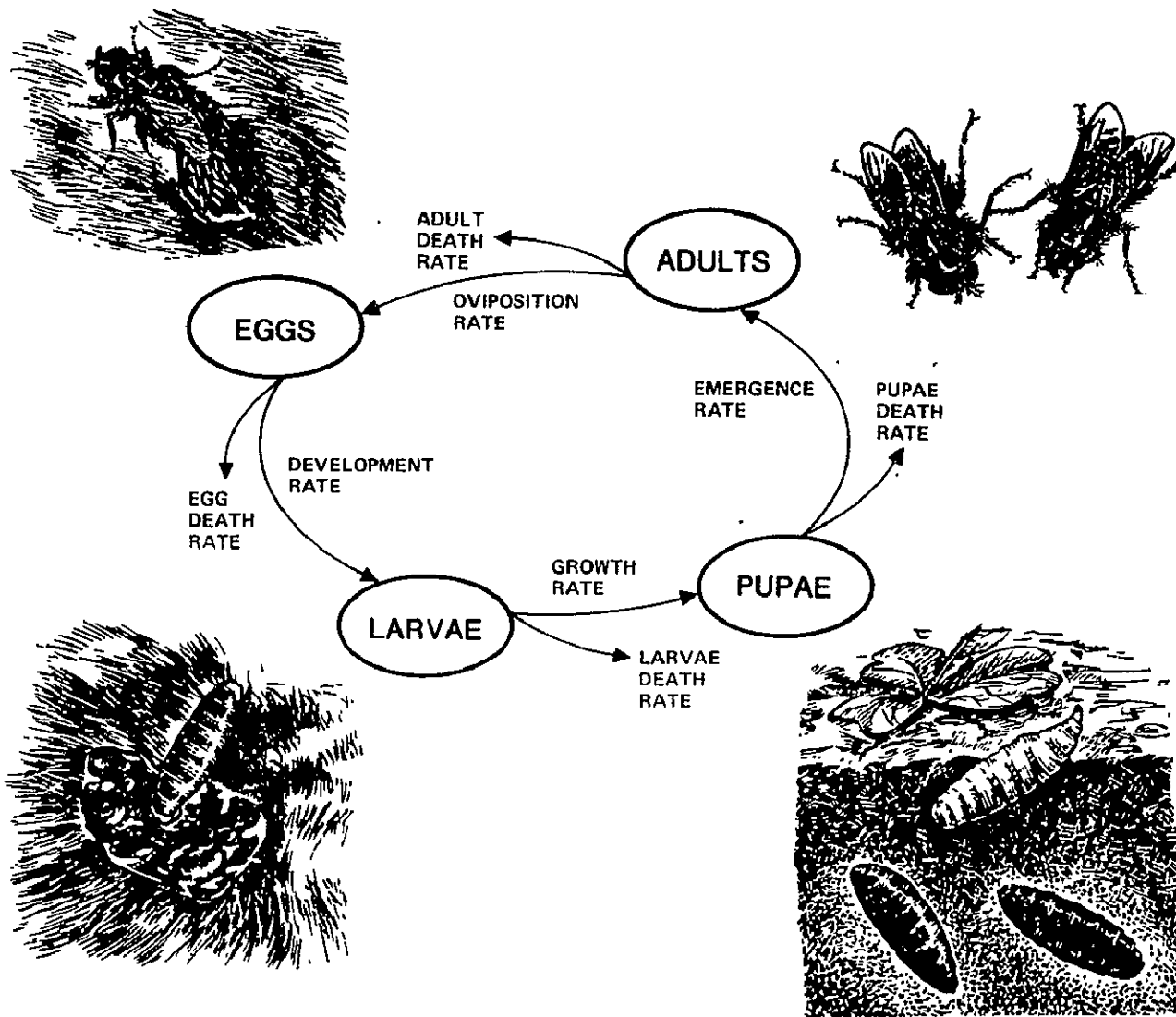


Figure 1-2. — Generalized diagram of the life cycle of the screwworm fly, *Cochliomyia hominivorax*.

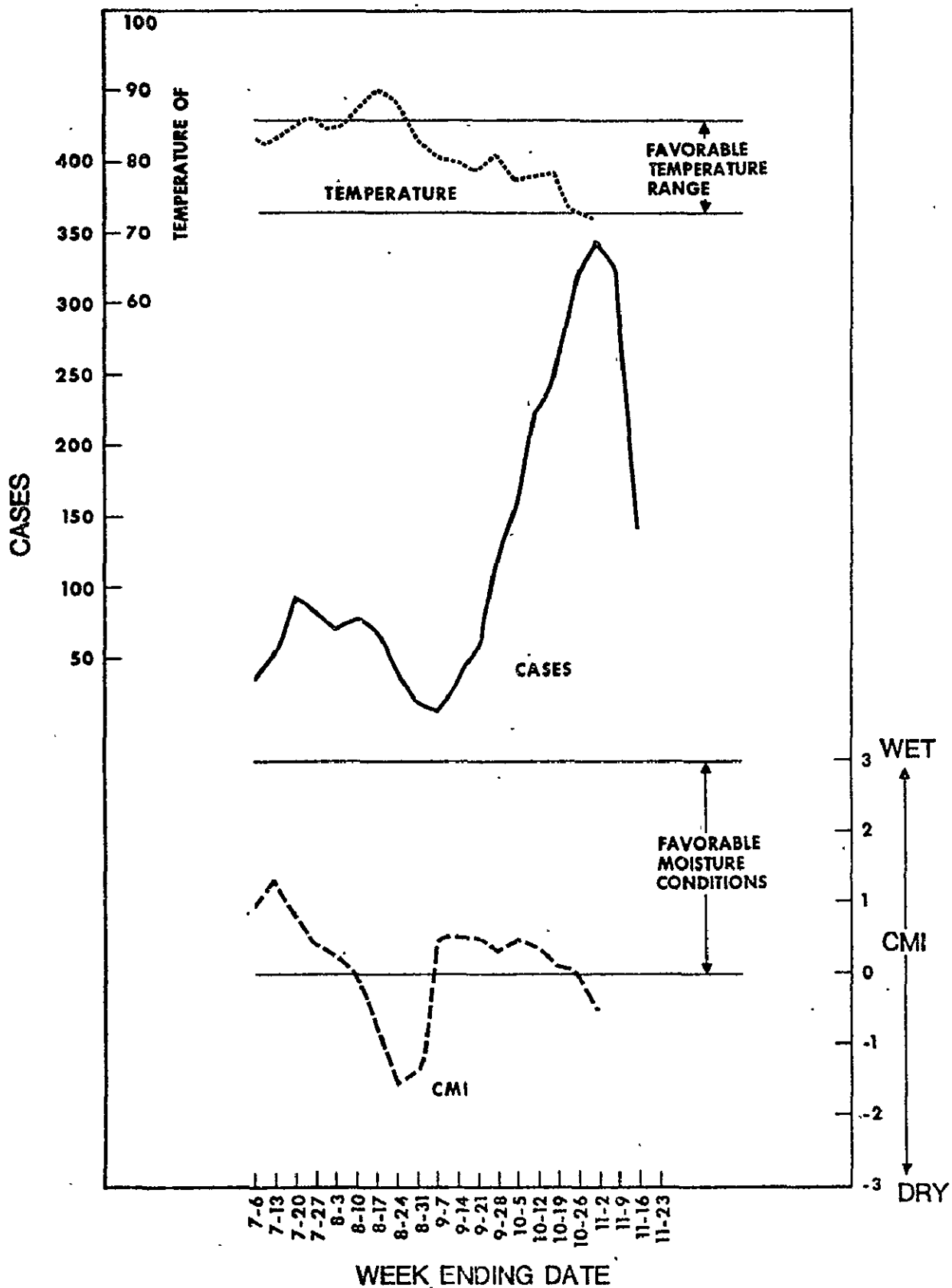


Figure 1-3.— Data plot showing the number of screwworm cases and the weather in Texas Climatological Division 9, 1968.

At a later stage in the development of SEDS, a decision was made that the four output products should be combined. A regression analysis of the four products was performed, and a preliminary set of coefficients was calculated for combining the four products. It is expected that these coefficients will change for different seasons of the year. It may also be possible after some experience to eliminate one or more output products without a loss of overall results.

#### 1.4 EVALUATION OF THE SYSTEM

SEDS first produced data in March 1975, and there was an immediate attempt to evaluate its effectiveness. During this initial evaluation, it became readily apparent that there needed to be a temperature fill-in scheme incorporated into SEDS to provide data when clouds obscured the ground. It was decided, therefore, to incorporate a delta temperature field and other fixes into SEDS and reprocess the spring 1975 data. An evaluation team of NASA and USDA representatives began their evaluation on the data as it was produced weekly.

Completion of spring data reprocessing occurred on January 9, 1976. The ongoing evaluation was accelerated, so that most of it could be reported at the following meeting of the full evaluation team.

The results of the spring evaluation are contained in separate documents or reports and are beyond the scope of this document. There were two factors associated with the evaluation which are worthy of listing since they could apply to any future evaluations. The first major limiting factor in the evaluation was the relatively small number of cases in the early spring from which to calculate actual growth potential for use in statistically comparing with SEDS growth potentials. For this reason,

much of the evaluation had to be subjective. It would have been preferable to have had sufficient data points from which to do a statistical evaluation.

The evaluation of the reprocessed spring data was also adversely affected by the speed at which it was produced. Since the registration coefficients had already been previously recalculated, that portion of the data did not need to be repeated. Consequently, four days of spring data could be reprocessed in each production day, leaving insufficient time to adequately analyze the data and compute the necessary operating coefficients for SEDS to perform to its capacity during the spring evaluation period.

## 2. CALCULATION OF SEDS SYSTEM PARAMETERS

The daily operation of SEDS requires the availability of several sets of parameters. These parameters are input to the system from a list of currently valid values. The SEDS system parameters form shown in figure 2-1 is also used to initiate any changes in system coefficients.

The following sections provide details on the calculation and utilization of each set of coefficients. Whenever possible, information has been provided on the theory and development of the current configuration of SEDS. However, emphasis is on developing a complete understanding of the analysis procedure required to maintain the highest possible system performance. Examples are provided both for calculation of the coefficient sets and for determining when recalculation is needed.

### 2.1 DAILY MEAN AIR TEMPERATURE (DMAT) COEFFICIENTS

SEDS uses a set of coefficients derived from multiple linear regression studies to convert radiometric data obtained by the National Oceanic And Atmospheric Administration (NOAA) satellite into estimates of mean air temperature. Computed daily, the mean air temperature is used by the various SEDS screwworm population dynamics algorithms to estimate the effect of temperature on different phases of the screwworm's life cycle. The performance of the SEDS screwworm growth potential predictions depends largely upon the accuracy of the estimated daily mean air temperature.

#### 2.1.1 BACKGROUND MATERIAL

The DMAT estimate is calculated by SEDS each day for each pixel in the data base image. The exact method of calculation depends upon the amount of radiometric data available. Data

SEDS SYSTEM PARAMETERS

Disk Number \_\_\_\_\_

Date \_\_\_\_\_

DMAT Coefficients

A <sub>0</sub> _____	B <sub>0</sub> _____	C <sub>0</sub> _____
A <sub>1</sub> _____	B <sub>1</sub> _____	C <sub>1</sub> _____
A <sub>2</sub> _____	B <sub>2</sub> _____	B <sub>2</sub> _____
A <sub>3</sub> _____		

Rainfall Classification Coefficients

A <sub>10</sub> _____	A <sub>20</sub> _____	A <sub>30</sub> _____
A <sub>11</sub> _____	A <sub>21</sub> _____	A <sub>31</sub> _____
A <sub>12</sub> _____	A <sub>22</sub> _____	A <sub>32</sub> _____
A <sub>40</sub> _____	A <sub>50</sub> _____	
A <sub>41</sub> _____	A <sub>51</sub> _____	
A <sub>42</sub> _____	A <sub>52</sub> _____	
T <sub>2</sub> _____	Day Thresh _____	Night Thresh _____

Screwworm Growth Coefficients

C _____	A <sub>1</sub> _____	A <sub>3</sub> _____
A <sub>0</sub> _____	A <sub>2</sub> _____	A <sub>4</sub> _____

Ground Truth Constants

KDAY _____	MINDLT _____
KNIGHT _____	DDT _____

Isothermal Bias and Offset

A _____	B _____
---------	---------

\*Parameters in parenthesis indicate those which are not in converted (internal) units.

Programming Support Section \_\_\_\_\_

Earth Resources Operations Section \_\_\_\_\_

Figure 2-1. -- SEDS System Parameter Control Form.

availability is largely a function of satellite coverage and of cloud cover.

The radiometric data must survive a number of filters which strive to eliminate all cloud contaminated data. Although these filters are described elsewhere, it is worth emphasizing that the quality both of the coefficients determination and of the DMAT estimates is dependent upon the use of cloud-free data. The quality of the registration of the radiometric data to the SEDS reference image is also extremely important. Without accurate registration, no meaningful comparison can be made between surface mean air temperature and the satellite observed radiometric temperatures.

#### 2.1.1.1 Theory and Development

Two assumptions are involved in attempting to relate radiometric temperatures to daily mean air temperatures. The first is that the instantaneous radiometric temperature, which is a surface temperature, is related to the air temperature measured at some fixed height above the ground. The second assumption is that one or two estimates of air temperature each day will be sufficient data to estimate the daily mean air temperature. Although in reality neither assumption is completely valid; as approximations of the complex, real situation, they appear to provide usable data.

On an experimental level, it can be seen from figure 2-2 that in data taken from a test area in Veracruz, Mexico, there was a good (.811) correlation between radiometric and air temperatures taken at the time of satellite passage. Simulations using surface observed air temperatures at the time of satellite passage, indicated that the potential for predicting daily mean air temperatures was quite good. Based on these data from ten Texas stations from the early summer of 1973, it was found that by

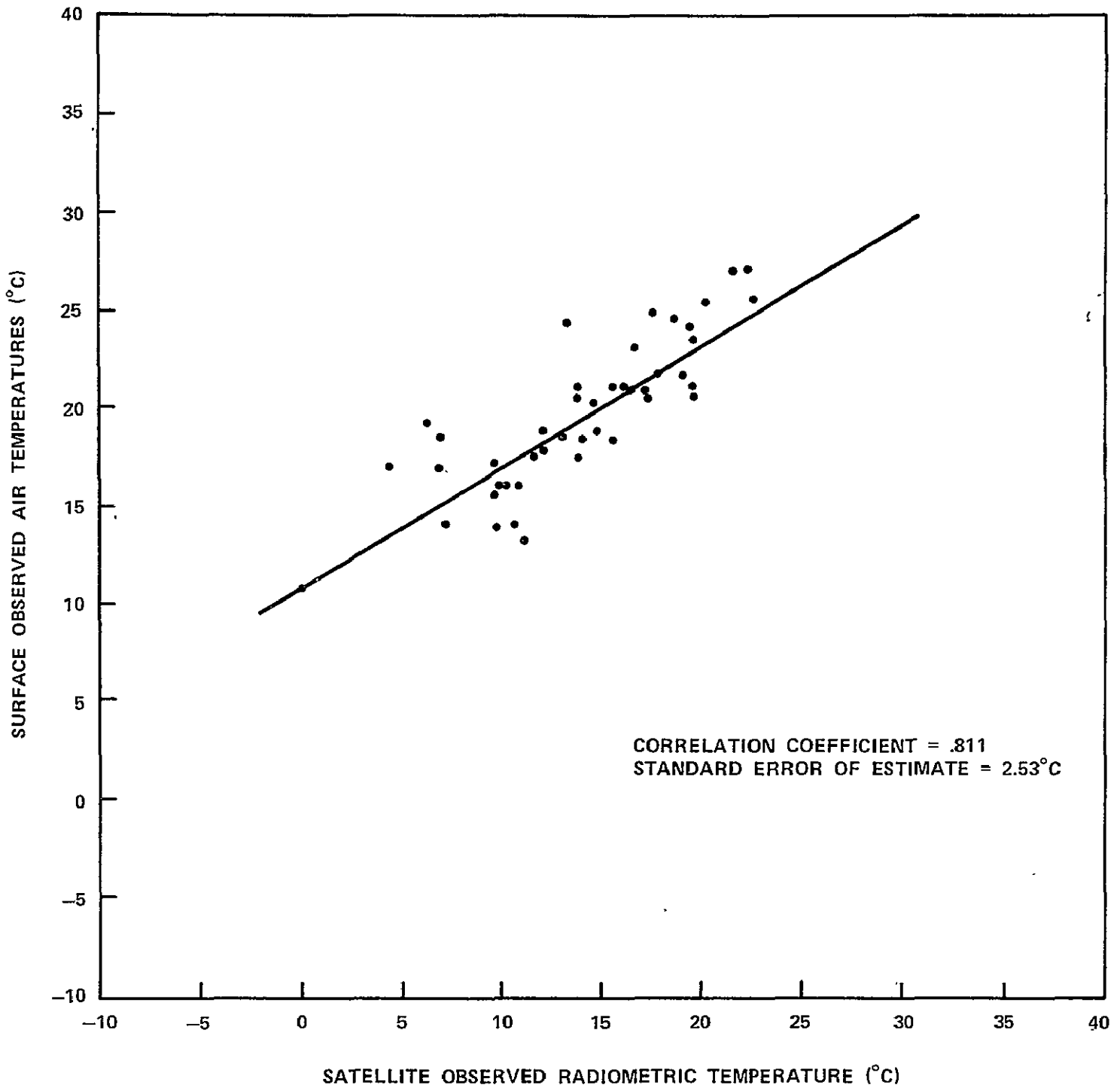


Figure 2-2. — Comparison of satellite data with surface air temperature in the Fortín de las Flores test area.



knowing the air temperature perfectly the daily mean could be predicted with error ranging from 1.07 to 3.43°C, depending upon the availability of data.

It can be shown that the effect of the air-ground temperature difference dominates the error produced by estimating actual surface net radiation flux from data taken in a weather shelter above the ground. In this case, the actual surface temperature, which is a function of the surface net radiation flux, is measured and an estimate is made of the net radiation flux (or temperature) of the air. If the air and surface are in thermal equilibrium, the temperature difference remains much the same over relatively short periods of time and under the influence of similar weather conditions.

Thus, it should be expected that radiometric data taken at night under relatively stable conditions would provide more reliable air temperature estimates than data taken in the early morning when the surface is undergoing rapid solar heating. Additionally, since the air-ground temperature difference has a strong seasonal dependence, it should be expected that any transformation between radiometric and air temperature would change with time.

A series of regression studies performed on July 1975 SEDS data revealed that the addition of terrain elevation as a variable could improve the errors in estimating mean air temperature on the order of 1°C. Table 2-I shows the relative improvement gained by including the additional variable. The data agrees with the expected results quite well. The DMAT estimates based solely on day radiometric data are of a lower quality than those based on night data. The elevation correction has little effect on the errors associated with the high quality night data.

TABLE 2-I. - IMPROVEMENT OF DMAT ESTIMATES UPON ADDITION OF ELEVATION AS A REGRESSION VARIABLE TO JULY 1975 SEDS DATA

	<u>Radiometric Data Only</u>	<u>Altitude Corrected</u>
Day Radiometric Only	6.5	3.4
Night Radiometric Only	2.6	2.0
Both Day and Night Radiometric Data	3.6	2.1
Overall (All Radiometric Data)	3.48	2.39

#### 2.1.1.2 SEDS Utilization

Depending upon the availability of radiometric data, there are three possible cases as shown below. Usable data is radiometric data which has survived all cloud detection algorithms.

- Case 1: Both night and day radiometric data usable  

$$\text{DMATK} = A_0 + A_1\text{TSNK} + A_2\text{TSDK} + A_3\text{ALT} \quad (2-1)$$
- Case 2: Only day radiometric data usable  

$$\text{DMATK} = B_0 + B_1\text{TSDK} + B_2\text{ALT} \quad (2-2)$$
- Case 3: Only night radiometric data usable  

$$\text{DMATK} = C_0 + C_1\text{TSNK} + C_2\text{ALT} \quad (2-3)$$

Where the symbols are defined as follows:

DMATK - Daily mean air temperature in degrees Kelvin  
TSDK - Day radiometric temperature in degrees Kelvin  
TSNK - Night radiometric temperature in degrees Kelvin  
ALT - In meters (terrain elevation)  
 $A_i, B_i, C_i$  - Regression coefficients.

Computationally SEDS uses the following relationships:

- Case 1:  $DMAT = A0C + A1C(TSN) + A2C(TSD) + A3C(PALT)$  (2-4)

- Case 2:  $DMAT = B0C + B1C(TSD) + B2C(PALT)$  (2-5)

- Case 3:  $DMAT = C0C + C1C(TSN) + C2C(PALT)$  (2-6)

Where the following symbols are used:

- DMAT — Eight-bit coded value of DMATK. DMATK and DMAT are related by the following

$$DMATK = \frac{DMAT}{4} + 260$$

- TSD — Eight-bit coded value of TSDK. TSDK and TSD are related by the following

$$TSDK = \frac{TSD}{2} + 202$$

- TSN — Eight-bit coded value of TSNK. TSNK and TSN are related by the following

$$TSNK = \frac{TSN}{2} + 202$$

- PALT — Eight-bit coded value of ALT. ALT and PALT are related by the following

$$ALT = 25(PALT) - 100.$$

The coded coefficients are floating point numbers which satisfy the following:

$$A0C = 4A_0 + 808A_1 + 808A_2 - 400A_3 - 1040 \quad (2-7)$$

$$A1C = 2A_1 \quad (2-8)$$

$$A2C = 2A_2 \quad (2-9)$$

$$A3C = 100A_3 \quad (2-10)$$

$$B0C = 4B_0 + 808B_1 - 400B_2 - 1040 \quad (2-11)$$

$$B1C = 2B_1 \quad (2-12)$$

$$B2C = 100B_2 \quad (2-13)$$

$$C0C = 4C_0 + 808C_1 - 400C_2 - 1040 \quad (2-14)$$

$$C1C = 2C_1 \quad (2-15)$$

$$C2C = 100C_2 \quad (2-16)$$

### 2.1.2 METHOD OF CALCULATION

The general approach to calculation is to develop a data set for each of the three cases. The data sets should contain cloud-free radiometric data which cover the entire range of expected values. Multiple linear regression analysis is used to calculate the coefficients for equations (2-1), (2-2), and (2-3). The coefficients are transformed into the form expected by SEDS and entered into the systems parameter list for implementation. A given set of coefficients remains valid for 4-6 weeks. In general, a continuing error analysis indicates when recalculation is necessary.

#### 2.1.2.1 Data Used

The radiometric data source is the SEDS meteorological station statistics printout. A typical printout is shown in table 2-II. The values for TSDK, TSNK, and VIS for those stations in the SEDS prime verification network are extracted. The stations in the prime verification network are shown in table 2-III. In general, at least two calendar weeks of SEDS data are required to be sure of obtaining a sufficient amount of data.

The primary source of surface meteorological data is the daily report provided by the National Weather Service. An example of this report is shown in table 2-IV. The daily mean air

ORIGINAL PAGE IS  
OF POOR QUALITY

TABLE 2-II. - PARTIAL LIST OF METEOROLOGICAL STATION STATISTICS  
PRINTOUT FOR 31 MARCH 1975

ID	Y	X	TSD	TSN	DMAT	R	CMI	DAY	TMET	DELT	TGT	ALT	FLAG1
CODE	GRID	GRID	K	K	K	CODE	CODE	VIS	K	K	K	M	
43	4	314	258.0	0.0	288.25	H	0.12	112	287.00	1.25	288.25	750.0	0
45	12	384	287.5	0.0	287.75	CO	0.12	44	286.50	1.25	287.75	150.0	0
48	31	338	269.5	0.0	290.00	H	0.12	89	287.00	3.00	290.00	450.0	0
47	39	385	277.0	0.0	286.75	L	0.12	80	286.50	2.25	288.75	150.0	0
48	39	284	287.5	0.0	286.25	CO	0.12	49	286.50	-0.25	286.25	750.0	0
43	14	253	290.0	0.0	286.75	CO	0.12	41	285.50	1.25	286.75	750.0	0
42	39	441	274.5	0.0	282.75	H	0.12	143	283.25	-0.50	282.75	150.0	0
49	19	482	282.5	0.0	283.75	H	1.50	150	283.25	0.50	283.75	150.0	0
44	42	300	268.5	0.0	287.50	L	-0.37	53	286.50	1.00	287.50	750.0	0
46	42	460	242.5	0.0	283.00	H	1.50	136	283.25	-0.25	283.00	150.0	0
4	40	207	298.5	0.0	286.50	CO	-0.12	44	285.50	1.00	286.50	1225.0	0
48	49	447	276.5	0.0	282.75	H	0.12	117	283.25	-0.50	282.75	150.0	0
52	53	339	293.5	0.0	287.50	H	0.12	35	286.50	2.50	289.25	450.0	0
73	95	105	290.0	0.0	284.00	CO	-0.37	48	282.50	1.50	284.00	1225.0	0
44	54	382	271.5	0.0	286.25	L	-0.37	67	287.00	-0.75	286.25	150.0	0
74	62	142	293.5	0.0	286.50	CO	-0.37	55	285.50	1.00	286.50	1225.0	0
39	65	141	295.0	0.0	286.50	CO	-0.37	52	285.50	1.00	286.50	1225.0	0
80	65	41A	267.5	0.0	287.00	H	0.12	112	287.00	0.00	287.00	150.0	0
40	67	257	297.0	0.0	286.75	CO	-0.75	41	285.50	1.25	286.75	1225.0	0
70	74	21	0.0	0.0	287.25	S	0.12	0	286.50	0.75	287.25	750.0	0
0	76	174	297.0	0.0	287.00	CO	-0.37	48	285.50	-1.50	287.00	1225.0	0
41	79	429	278.5	0.0	288.75	H	1.50	143	286.25	0.50	288.75	150.0	0
48	83	293	292.0	0.0	288.75	CO	-1.25	43	289.25	-0.50	288.75	450.0	0
93	84	36A	271.5	0.0	287.75	L	0.12	49	287.00	0.75	287.75	150.0	0
64	87	389	274.0	0.0	287.50	H	-0.12	123	288.25	-0.75	287.50	150.0	0
71	88	39	0.0	0.0	283.00	O	0.12	0	281.00	2.00	283.00	1225.0	0
72	89	60	293.0	0.0	283.25	CO	-0.12	45	282.00	1.25	283.25	1225.0	0
37	90	259	292.5	0.0	288.50	CO	-1.25	49	280.50	8.00	288.50	750.0	0
75	93	215	297.0	0.0	287.75	C	-0.37	35	280.50	5.00	285.25	1225.0	0
34	94	237	294.5	0.0	285.25	CO	-0.75	40	280.50	4.75	285.25	1225.0	0
97	94	415	254.0	0.0	288.50	H	-0.12	183	288.25	0.25	288.50	150.0	0
53	102	24	0.0	0.0	287.75	O	0.12	0	286.50	1.25	287.75	600.0	0
76	105	304	298.0	0.0	289.25	CO	-1.25	44	289.25	0.00	289.25	450.0	0
68	112	397	279.0	0.0	289.50	H	-0.75	140	288.25	1.25	289.50	150.0	0
13	120	210	300.5	0.0	288.00	CO	-0.37	46	280.50	-7.50	288.00	1225.0	0
77	122	356	282.0	0.0	288.50	CO	-0.37	51	288.25	0.25	288.50	150.0	0
27	123	241	297.0	0.0	288.75	C	-0.75	39	280.50	5.50	285.50	750.0	0
56	123	321	299.5	0.0	291.25	C	-0.75	33	289.25	1.50	290.75	150.0	0
96	131	40A	278.5	0.0	290.50	H	-0.75	99	289.25	1.25	290.50	150.0	0
89	152	357	302.5	0.0	292.75	C	-0.75	38	288.25	0.50	288.25	150.0	0
52	153	356	300.0	0.0	291.50	C	-0.75	37	288.25	0.25	288.00	150.0	0
90	153	380	300.0	0.0	291.50	C	-0.75	33	289.25	1.50	290.50	150.0	0
2	158	171	297.5	0.0	288.50	CO	-0.37	52	289.25	-0.75	288.50	2500.0	0
91	158	406	246.5	0.0	290.50	H	-1.25	119	289.25	1.25	290.50	150.0	0
1	169	36	0.0	0.0	291.00	O	-0.12	0	290.25	0.75	291.00	600.0	0
92	179	416	248.0	0.0	288.25	H	-1.25	128	288.75	-0.50	288.25	150.0	0
78	184	400	254.0	0.0	288.75	H	-1.25	105	289.75	-1.00	288.75	150.0	0
79	187	424	257.0	0.0	289.00	H	-1.25	138	288.75	0.25	289.00	150.0	0
47	188	423	256.0	0.0	289.00	H	-1.25	142	288.75	0.25	289.00	150.0	0
57	189	402	246.0	0.0	288.25	H	-1.25	126	289.75	-1.50	288.25	150.0	0
3	202	39	0.0	0.0	289.25	O	-1.50	0	289.25	0.00	289.25	0.0	0
6	207	34A	302.5	0.0	292.00	C	-0.75	40	288.25	0.25	288.00	600.0	0
38	213	72	297.0	0.0	298.75	CO	-0.37	41	289.25	-0.50	288.75	100.0	0
59	229	9	0.0	0.0	289.75	O	1.50	0	289.25	0.50	289.75	0.0	0
5	238	262	301.5	0.0	291.75	CO	-0.75	47	290.25	1.50	291.75	1500.0	0
12	257	427	248.0	0.0	289.25	L	-1.25	75	292.25	-3.00	289.25	100.0	0
11	264	394	240.0	0.0	285.25	H	-1.25	106	290.25	-5.00	285.25	600.0	0
8	282	157	298.0	0.0	293.25	CO	-0.37	41	292.25	1.00	293.25	600.0	0
9	291	238	300.5	0.0	292.75	CO	-0.37	47	290.25	2.50	292.75	600.0	0
18	301	442	295.0	0.0	290.50	CO	-1.25	47	292.25	-1.75	290.50	100.0	0
62	314	41A	254.5	0.0	292.75	L	-1.25	75	294.25	-1.50	292.75	100.0	0
7	317	7A	297.5	0.0	288.50	CO	-0.37	49	289.25	-0.75	288.50	100.0	0
58	325	355	304.5	0.0	291.50	CO	-0.75	47	294.25	-2.75	291.50	1500.0	0
17	325	364	304.0	0.0	291.00	C	-0.75	37	290.25	-3.25	291.00	1500.0	0
10	326	194	301.5	0.0	292.50	C	-0.37	33	292.25	0.00	292.25	100.0	0
20	330	465	290.0	0.0	290.50	H	-0.75	104	292.25	-1.75	290.50	100.0	0
36	342	319	297.5	0.0	291.00	CO	-0.75	51	294.25	-3.25	291.00	600.0	0
51	393	489	290.0	0.0	290.75	H	-0.75	122	292.25	-1.50	290.75	0.0	0
16	361	342	311.5	0.0	296.75	C	-0.75	35	296.25	-2.50	293.25	600.0	0
15	369	248	265.0	0.0	292.00	L	-0.37	52	292.25	-0.25	292.00	600.0	0
60	369	444	304.5	0.0	289.50	C	-0.75	42	0.00	-2.50	0.00	2500.0	0
19	370	433	298.0	0.0	286.00	C	-0.75	30	0.00	-2.00	0.00	2500.0	0
43	375	49A	253.5	0.0	291.75	H	-1.25	92	297.25	-5.50	291.75	600.0	0
26	380	518	281.5	0.0	294.25	CO	-0.75	70	297.25	-3.00	294.25	0.0	0

TABLE III. - SEDS PRIME VERIFICATION NETWORK

<u>Station Name</u>	<u>SEDS ID</u>	<u>Latitude</u>	<u>Longitude</u>
Hermosillo	01	29:04	110:57
Chihuahua	02	28:38	106:04
Guaymas	03	27:55	110:53
La Paz	07	29:04	110:20
Durango	09	24:01	104:40
Mazatlan	10	23:12	106:25
Soto la Marina	12	23:46	98:12
Tampico	14	22:13	97:51
Tepic	15	21:30	104:53
Rio Verde	17	21:56	100:00
Guadalajara	18	20:36	103:23
Pachuca	19	20:08	98:44
Tuxpan	20	20:57	97:23
Puebla	25	19:03	98:12
Veracruz	26	19:09	96:07
Oaxaca	30	17:03	96:43
Acapulco	31	16:50	99:55
Salina Cruz	32	16:10	95:12
Tucson	70	32:07	110:56
Douglas	72	31:27	109:36
Deming	73	32:15	107:42
El Paso	74	31:48	106:24
Marfa	75	30:22	104:01
Del Rio	76	29:22	100:55
Cotulla	77	28:27	99:13
McAllen	78	26:11	98:14
Brownsville	79	25:54	97:26
Midland-Odessa	83	31:57	102:11
San Angelo	84	31:22	100:30
Abilene	85	32:25	99:41
San Antonio	93	29:32	98:28
Austin	94	30:18	97:42
Waco	95	31:37	97:13
Corpus Christi	96	27:46	97:30
Victoria	97	28:51	96:55
Houston	98	29:58	95:12
Port Arthur	99	29:57	94:01

TABLE 2-IV. - DAILY SURFACE DATA REPORT

DH04 MEXICO SCREWORM DATA FOR SUN 26 OCT 1975

ID STN	TT	RRRR	ID STN	TT	RRRR	ID STN	TT	RRRR
MEXICAN STATION DATA FOR 10/27 0753								
00 122	99	9999	12 499	21	0027	24 680	18	0000
01 160	23	9999	13 518	99	9999	25 685	17	0000
02 225	18	0000	14 548	22	0000	26 692	28	0000
03 255	22	0000	15 556	22	0000	27 727	99	9999
04 342	99	0000	16 575	21	9999	28 741	27	0132
05 382	19	9999	17 581	19	0039	29 762	22	0000
06 393	19	9999	18 613	19	0000	30 775	20	0000
07 405	23	0000	19 632	15	0000	31 805	27	0000
08 412	26	9999	20 640	19	0118	32 833	27	0000
09 423	19	9999	21 654	99	9999	33 845	16	0004
10 458	25	0000	22 662	99	9999	34 855	99	0000
11 491	16	0000	23 679	18	0000	35 903	26	0070

36 AGS	17		48 MLG	99		60 TCG	14	
37 AGU	99		49 MLM	20		61 TGZ	25	
38 CEN	21		50 MTT	29		62 TMN	17	
39 CJS	18		51 NAU	25		63 UPN	17	
40 GRN	99		52 NLD	22		64 VCT	99	
41 IMA	22		53 NDG	22		65 VSA	27	
42 IZT	28		54 PTJ	99		66 ZIH	27	
43 JAL	18		55 PVR	27		80 TIJ	16	
44 LMM	99		56 PZA	99		81 MXL	24	
45 LTO	99		57 REX	99		82 PPE	15	
46 LZC	99		58 SLP	20				
47 MAM	99		59 SRL	20				

DH05 NOTE: TT, DEGS C; RRRR, TENTHS OF MM; 99 & 9999, MISSING DATA.

WW22 SMUS & SAUS 26 OCT 75

ID STN	XXX NN	TT	RRRR	ID STN	XXX NN	TT	RRRR
67 290		59	0	74 270		60	0
68 IPL		64	0	75 MRF		52	0
69 280		64	0	76 261		59	0
70 274		64	0	77 CDT		63	0
71 FHU		60	0	78 MFE		61	0013
72 DUG		58	0	79 250		60	T
73 DMN		56	0				

TEMPS DEGS F, 2 GRP SCRN (HUNDREDTHS OF IN.).

SUNDAY 26 OCT 1975  
DAY 1 OF WEEK 35

ORIGINAL PAGE IS  
OF POOR QUALITY

temperature is reported for each meteorological station under the TT heading. Missing data is indicated by 99.

The surface meteorological data for some stations are contained in the meteorological station statistics printout. For all stations in the prime verification network, the surface data are reported under TMET. Some care must be used when extracting surface data from this source. TMET will be the surface station data only for the ground truth control stations.

A list of these stations is given in table 2-V. For any station not in this list, the only source of data is the daily National Weather Service report. As an example, the data for Brownsville, Texas for 15-29 March 1975, has been extracted and is shown in table 2-VI.

#### 2.1.2.2 Method of Analysis

The raw data is used to determine the appropriate thermal thresholds as described in section 2.3 and to calculate the rainfall coefficients as described in section 2.2. The results of those analyses are used to screen the raw radiometric data for cloud contamination.

For the 15-29 March 1975, time period used in the previous example, the following thermal thresholds were used:

- Day Threshold = 289.5°K
- Night Threshold = 280.0°K

Any radiometric data less than the threshold is considered cloudy and is deleted.



TABLE 2-V. — GROUND TRUTH ZONE CONTROL STATIONS

<u>STATION NAME</u>	<u>SEDS ID</u>	<u>LATITUDE</u>	<u>LONGITUDE</u>	<u>ELEVATION</u>
Hermosillo	01	29:04	110:57	211 m
Chihuahua	02	28:38	106:04	1354 m
Guaymas	03	27:55	110:53	16 m
La Paz	07	29:04	110:20	10 m
Durango	09	24:01	104:40	1889 m
Mazatlan	10	23:12	106:25	3 m
Ciudad Victoria	11	23:44	99:08	321 m
Soto la Marina	12	23:46	98:12	25 m
Tampico	14	22:13	97:51	12 m
Tepic	15	21:30	104:53	915 m
Rio Verde	17	21:56	100:00	991 m
Guadalajara	18	20:36	103:23	1567 m
Pachuca	19	20:08	98:44	2426 m
Tuxpan	20	20:57	97:23	14 m
Mexico City	23	19:24	99:11	2298 m
Puebla	25	19:03	98:12	2162 m
Veracruz	26	19:09	96:07	13 m
Oaxaca	30	17:03	96:43	1500 m
Acapulco	31	16:50	99:55	82 m
Salina Cruz	32	16:10	95:12	70 m
Las Casas	33	16:44	92:38	2128 m
Tucson	70	32:07	110:56	778 m
Fort Huachuca	71	31:35	110:20	1426 m
Douglas	72	31:27	109:36	1250 m
Deming	73	32:15	107:42	1316 m
El Paso	74	31:48	106:24	1192 m
Marfa	75	30:22	104:01	1479 m
Del Rio	76	29:22	100:55	313 m
Cotulla	77	28:27	99:13	146 m
McAllen	78	26:11	98:14	34 m
Brownsville	79	25:54	97:26	6 m
Midland-Odessa	83	31:57	102:11	868 m
San Angelo	84	31:22	100:30	579 m
Abilene	85	32:25	99:41	543 m
San Antonio	93	29:32	98:28	240 m
Austin	94	39:18	97:42	182 m
Waco	95	31:37	97:13	153 m
Corpus Christi	96	27:46	97:30	12 m
Victoria	97	28:51	96:55	32 m
Houston	98	29:58	95:12	29 m
Port Arthur	99	29:57	94:01	5 m

TABLE 2-VI. - SAMPLE RADIOMETRIC DATA AS PREPARED FOR ANALYSIS

Station Name:	Brownsville, Texas			
Station ID:	79			
Station Altitude:	6 m			
<u>DATE</u>	<u>TSDK</u>	<u>VIS</u>	<u>TSNK</u>	<u>TI</u>
3-15-75	270.5	68	295.5	293.15
3-16-75	243.5	82	-	294.27
3-17-75	290.5	66	-	297.60
3-18-75	294.0	40	-	298.71
3-19-75	311.0	35	295.0	292.04
3-20-75	301.0	39	-	293.15
3-21-75	302.0	52	-	297.04
3-22-75	-	-	-	295.93
3-23-75	298.5	97	302.0	299.82
3-24-75	-	-	-	298.15
3-25-75	283.0	34	-	292.60
3-26-75	251.0	99	-	298.71
3-27-75	238.5	74	234.0	299.82
3-28-75	-	-	-	300.93
3-29-75	269.0	113	254.5	285.38

The rainfall coefficient defines a decision field which assigns all possible combinations of TSD and VIS to one of five classes. Only class 1 is cloud free. A part of the decision field used during the time period of the example is shown in figure 2-3.

Consider the data for 17 March 1975:

$$\text{TSDK} = 290.5$$

$$\text{VIS.} = 66$$

The value of TSDK is greater than the day threshold of 289.5 but the VIS, TSDK combination entered in figure 2-3 does not fall into the clear category. Table 2-VII shows the usable data that remained after thresholding and cloud discrimination. In this way, a data set can be assembled suitable for regression analysis.

The initial data sets should contain at least 50 points for each of the three cases. If insufficient data are available for one or more cases, either the time period or the number of stations should be expanded.

Multiple linear regression analysis assumes a model of the following form:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

The data are collected and may be written as

$$\begin{array}{l} X_{11} , X_{21} , \dots , X_{n1} , Y_1 \\ X_{12} , X_{22} , \dots , X_{n2} , Y_2 \\ \cdot \\ \cdot \\ \cdot \\ X_{1m} , X_{2m} , \dots , X_{nm} , Y_m \end{array}$$

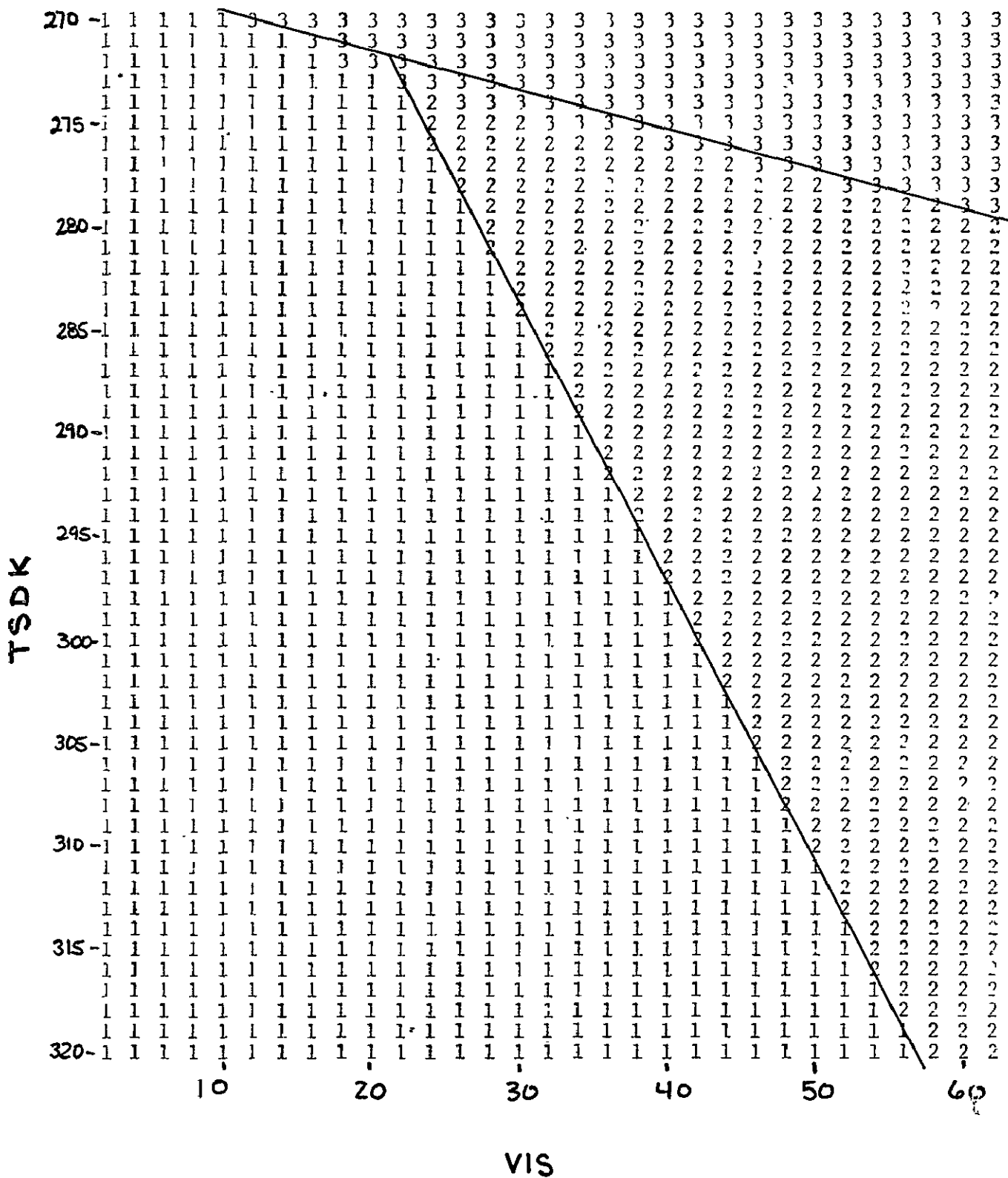


Figure 2-3. - Cloud discrimination decision field for March 1975 data.

TABLE 2-VII. — SAMPLE RADIOMETRIC DATA FROM TABLE 2-VI  
AFTER THERMAL THRESHOLDING AND CLOUD DISCRIMINATION

Station Name:	Brownsville, Texas		
Station ID:	79		
Station Altitude:	6 m		
<u>DATE</u>	<u>TSDK</u>	<u>TSNK</u>	<u>TT</u>
3-15-75	—	295.5	293.15
3-16-75	—	—	—
3-17-75	—	—	—
3-18-75	—	—	—
3-19-75	311.0	295.0	292.04
3-20-75	301.0	—	293.15
3-21-75	—	—	—
3-22-75	—	—	—
3-23-75	—	302.0	299.82
3-24-75	—	—	—
3-25-75	—	—	—
3-26-75	—	—	—
3-27-75	—	—	—
3-28-75	—	—	—
3-29-75	—	—	—

Assume that the  $X_i$  are measured exactly, but that  $Y$  has random, normally distributed error with constant variance. Least squares analysis provides a method of finding estimates  $b_i$  of the parameters  $\beta_i$ . These estimates are the regression coefficients.

The procedure is to minimize the sum of squares

$$S = \sum (Y - \beta_0 - \beta_1 X_1 - \beta_2 X_2 - \dots - \beta_n X_n)^2$$

by setting  $\frac{\partial S}{\partial \beta_0} = 0$ ,  $\frac{\partial S}{\partial \beta_i} = 0$ , and replacing  $\beta_0, \beta_i$  by their estimates  $b_0, b_i$ .

Unsubscripted summations are over the number of observation  $m$ ;  $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, n$ .

Defining,

$$c_{ij} = \sum x_i x_j = \sum X_i X_j - \frac{\sum X_i \sum X_j}{m}$$

$$c_{iy} = \sum x_i y = \sum X_i Y - \frac{\sum X_i \sum Y}{m}$$

$$c_{yy} = \sum y^2 = \sum Y^2 - \frac{(\sum Y)^2}{m}$$

leads to the  $n$  simultaneous linear normal equations

$$\left. \begin{aligned} c_{11}b_1 + c_{12}b_2 + \dots + c_{1n}b_n &= c_{1y} \\ c_{21}b_1 + c_{22}b_2 + \dots + c_{2n}b_n &= c_{2y} \\ \vdots & \\ \vdots & \\ c_{n1}b_1 + c_{n2}b_2 + \dots + c_{nn}b_n &= c_{ny} \end{aligned} \right\} I$$

Together with

$$b_0 = \bar{Y} - \sum_i b_i \bar{X}_i$$

The system I is conveniently solved by inverting the  $n \times n$  coefficient matrix  $C = (c_{ij})$  to give the matrix  $C^{-1} = (c'_{ij})$  (The program checks to see if  $C$  is singular and if so, it will print out.) We then have

$$b_i = \sum_j c'_{ij} c_{jy} .$$

The variance  $\sigma^2$  is estimated by

$$s^2 = \frac{(c_{yy} - \sum_i b_i c_{iy})}{(m - n - 1)}$$

there now being only  $(m - n - 1)$  degrees of freedom since  $(n + 1)$  coefficients have been estimated from the data. The variances and covariances of the  $b_i$  are given by

$$\begin{aligned} \text{var } (b_i) &= s^2 c'_{ii} \\ \text{cov } (b_i, b_j) &= s^2 c'_{ij} . \end{aligned}$$

It may also be shown that

$$\text{var } (b_0) = s^2/m + s^2 \sum_i \sum_j \bar{X}_i \bar{X}_j c'_{ij}$$

and

$$\text{cov } (b_0, b_i) = - s^2 \sum_j \bar{X}_j c'_{ij} .$$

The coefficient of multiple determination  $R^2$  is computed as follows:

$$R^2 = \frac{\sum_i b_i c_{iy}}{c_{yy}}$$

Its significance, and hence the overall significance of the regression, is tested by analysis of variance, and thus by obtaining the ratio

$$F = \frac{(m - n - 1)R^2}{n(1 - R^2)}$$

and testing it against the F-distribution with  $n$  and  $m - n - 1$  degrees of freedom.

An adjusted coefficient of multiple determination, corrected for sample size and number of independent variable often useful in comparing results, is computed as follows:

$$\bar{R}^2 = 1 - \frac{(m - 1)(1 - R^2)}{m - n - 1} .$$

The quality of the final regression fit can be improved by examining the residuals from the first run for each case. In general, any data set which has an error greater than twice the standard error of estimate should be deleted and the regression rerun. The final coefficient set can be obtained in, at most, three passes.

The correlation matrix, multiple correlation coefficients, standard error of estimate, analysis of variance, and regression coefficients should be retained for future reference.



Table 2-VIII shows the results of the final regression analysis for the 15-29 March 1975 SEDS night only radiometric data. In this example, the resulting regression is highly significant (F-Value = 179.1) with an excellent standard error of estimate of 1.90°C. The correlation matrix shows strong correlation between mean air temperature and both of the independent variables. When evaluated over all data during the 15-29 March 1975 time period, the performance of this coefficient was very good with a root mean square (rms) error of 2.41°C.

The SEDS DMAT coefficients are found by transforming the regression coefficients from equations (2-1) to (2-3) into the SEDS computational form shown in equations (2-4) to (2-6). The transformations are given in equations (2-7) to (2-16). Table 2-IX gives the coded and uncoded coefficients as found in the 15-29 March 1975 DMAT calculation.

### 2.1.3 VERIFICATION

The verification of the coefficients' set is a continuing process ultimately resulting in the recalculation of the DMAT coefficients. The performance of the DMAT estimates should be evaluated over the prime verification network shown in table 2-III.

The verification uses a measure of the mean and standard deviations of the observed errors. The error is defined as

$$\text{ERROR} = \text{Observed} - \text{Predicted}$$

The Bias error is then the arithmetic mean of the errors. A positive error means that SEDS underpredicted with the DMAT estimate. The rms error is the standard deviation of the errors. Approximately 67 percent of the values will be within one standard deviation of the mean and 95 percent will be within two standard deviations.

TABLE 2-VIII. - MULTIPLE REGRESSION RESULTS FOR 15-29 MARCH 1975  
 FOR CASE 3 - ONLY NIGHT RADIOMETRIC DATA USABLE

ORIGINAL PAGE IS  
 OF POOR QUALITY

<u>Correlation Matrix:</u>				
	TSNK	ALT	TT	
TSNK	1.000	-.670	.968	
ALT	-.670	1.000	-.714	
TT	.968	-.714	1.000	
<u>Multiple Correlation Coefficient:</u> .972				
<u>Standard Error of Estimate:</u> 1.901°C				
<u>Analysis of Variance for the Multiple Linear Regression:</u>				
	<u>Degrees Freedom</u>	<u>Sum Squares</u>	<u>Mean Squares</u>	<u>F Value</u>
Due to Regression	2	1294.86217	647.43108	179.12646*
Deviation about Regression	21	75.90197	3.61437	
Total	23	1370.76415		
*Significant at the .99 level.				
<u>Regression Coefficients:</u>				
Constant = 47.89565				
X(TSNK) = .83225				
X(ALT) = -.00155				

TABLE 2-IX. - 15-29 MARCH 1975 DMAT COEFFICIENTS  
IN BOTH CODED AND UNCODED FORMS

Coded:

$$\text{DMAT} = \text{A0C} + \text{A1C (TSN)} + \text{A2C (TSD)} + \text{A3C (ALT)}$$

$$\text{DMAT} = \text{B0C} + \text{B1C (TSD)} + \text{B2C (ALT)}$$

$$\text{DMAT} = \text{C0C} + \text{C1C (TSN)} + \text{C2C (ALT)}$$

$$\text{A0C} = -126.69744 \quad \text{B0C} = -82.66244 \quad \text{C0C} = -175.0970$$

$$\text{A1C} = 001.23440 \quad \text{B1C} = 01.07400 \quad \text{C1C} = 001.6651$$

$$\text{A2C} = -000.13490 \quad \text{B2C} = -00.19000 \quad \text{C2C} = -000.1550$$

$$\text{A3C} = 000.01000$$

Uncoded:

$$\text{DMATK} = \text{A}_0 + \text{A}_1 \text{TSNK} + \text{A}_2 \text{TSDK} + \text{A}_3 \text{ALT}$$

$$\text{DMATK} = \text{B}_0 + \text{B}_1 \text{TSDK} + \text{B}_2 \text{ALT}$$

$$\text{DMATK} = \text{C}_0 + \text{C}_1 \text{TSNK} + \text{C}_2 \text{ALT}$$

$$\text{A}_0 = 90.03634 \quad \text{B}_0 = 130.67039 \quad \text{C}_0 = 47.89565$$

$$\text{A}_1 = 00.61720 \quad \text{B}_1 = 000.53700 \quad \text{C}_1 = 00.83225$$

$$\text{A}_2 = 00.06745 \quad \text{B}_2 = -000.00190 \quad \text{C}_2 = -00.00155$$

$$\text{A}_3 = 00.00010$$

The Bias and rms errors for the original data set upon which the DMAT coefficients were calculated should be found. These values form the base line to which future errors should be compared.

In the example from 15-29 March 1975, the Bias error for Case 3 (night radiometric data only) was  $+0.52^{\circ}\text{C}$  and the rms error was  $2.41^{\circ}\text{C}$ . By 10-23 May 1975, the Bias error had changed to  $1.43^{\circ}\text{C}$  and the rms error to  $3.63^{\circ}\text{C}$ . At that time the DMAT coefficients were recalculated. For the same case, the new Bias error was  $0.0^{\circ}\text{C}$  with a rms error of  $2.16^{\circ}\text{C}$ . Thus, providing the base-line figures for the next evaluation.

In general, the errors for at least a week's data should be calculated to avoid random sampling errors. When either of the errors has increased more than  $1.0^{\circ}\text{C}$ , the DMAT coefficients should be recalculated.

## 2.2 RAINFALL COEFFICIENTS

The primary method used by SEDS for detecting cloud contaminated day radiometric data is the rainfall classifier. The day pass infrared and visible data are classified into one of five classes by a set of discriminant functions. Only the data which falls into the "clear" category is used by SEDS for calculating the daily mean air temperature estimates.

### 2.2.1 BACKGROUND MATERIAL

The name of the rainfall classification algorithm is misleading. As used by SEDS it is actually a cloud classification algorithm. This feature of SEDS was intended to serve the dual purpose of providing information about cloud contamination in the radiometric data and of providing a research tool for making estimates of precipitation. The research into estimating precipitation probability or amount from SEDS data was not carried out. The

algorithm provides a reliable means of eliminating cloud contamination from the day radiometric data.

#### 2.2.1.1 Theory

Briefly stated, the problem is one of differentiating between two or more populations on the basis of multivariate measurements. Given the existence of  $N$  populations, there exists a two-part problem. The first is to correctly distribute a set of samples among the  $N$  populations; the second is to determine a function based on these samples, which will allow the assignment of a new sample to the correct population. The first class of problems is generally called classification analysis, and the second is called discriminant analysis.

Classification analysis has been highly developed for applications as diverse as numerical taxonomy and land use classification. The problem is to classify a series of samples into groups which are as distinct as possible.

In SEDS, the day pass data consists of a pair of multispectral measurements for each point. The infrared and visible data when used as a pair provide more information on cloud cover than either data could by itself. Through classification analysis, it is possible to distribute a series of samples into the most appropriate groups. By controlling classification procedure, it is possible to create a single cloud-free class and several cloudy classes.

Discriminant analysis takes the samples placed in each group and derives a set of functions that enables the assignment of any new sample to the group which has the highest probability of containing the sample. These functions serve as the cloud classification algorithm in SEDS. The resulting decision field has been shown previously in figure 2-3.

### 2.2.1.2 SEDS Utilization

The algorithm used in SEDS is a maximum likelihood classification using linear discriminant functions. The algorithm assigns any possible input sample to one of five class regions or to an "unclassifiable" class. A thresholding capability allows the operator to set the probability level for inclusion in a class. To be used in the DMAT calculation an input sample must be classified clear. The classification algorithm may be defined as follows:

$$K_1 = a_{10} + a_{11}VIS + a_{12}TSD \quad (2-17)$$

$$K_2 = a_{20} + a_{21}VIS + a_{22}TSD \quad (2-18)$$

$$K_3 = a_{30} + a_{31}VIS + a_{32}TSD \quad (2-19)$$

$$K_4 = a_{40} + a_{41}VIS + a_{42}TSD \quad (2-20)$$

$$K_5 = a_{50} + a_{51}VIS + a_{52}TSD \quad (2-21)$$

Where the symbols are defined as follows:

$a_{ij}$  - The rainfall estimation coefficients

VIS - Eight-bit value of the visible detector voltage

TSD - Eight-bit coded value of the day radiometric temperature

The sample is assigned to the class which has the largest  $K_i$  for the (VIS, TSN) vector of that sample. Through the use of a threshold value, it is possible to emphatically set the probability level for inclusion in each class. When a given vector's probability of inclusion in any class is low, the vector is assigned to a null class. To date, this feature has not been used. Although the classes have no meaning beyond the clear/cloudy classification, the full results are given in the Meteorological Station Statistics Printout under "R CODE". Table 2-X gives the interpretation of each of the symbols in the printout.

TABLE 2-X. - METEOROLOGICAL STATION STATISTICS  
PRINTOUT R CODE INTERPRETATION

Symbol	Class	Interpretation
C	Dry	Clear
CD	Cloudy Dry	Cloudy
L	Light Rain	Cloudy
M	Moderate Rain	Cloudy
H	Heavy Rain	Cloudy
O	Null Class	Cloudy

For computational purposes, SEDS requires that the discriminant coefficients  $a_{ij}$  be scaled. The scaled coefficients,  $A_{ij}$ , are input to SEDS via the SEDS System Parameter Control Form shown in figure 2-1. The coefficients are scaled integers subject to the following constraints:

- The largest multiplicative term  $\leq 255$
- The absolute value of the largest constant term  $\leq 2^{15}$  -

### 2.2.2 METHOD OF CALCULATION

The approach consists of collecting a set of data from SEDS which is classified into five different groups by the classification program ISOCLS/PRINT. It identifies groups or clusters in the input data that have similar elements within clusters that are different from the elements of other clusters. The program uses builtin default values for all classification parameters. The samples from each group are then used to calculate the coefficients for equations 2-17 through 2-21. The discriminant function coefficients are calculated by the program DISCRM.

Program tapes and listings of these programs along with user documentation are available from Dr. C. M. Barnes, Health Applications Officer, Life Sciences Directorate, NASA, JSC. Details of these programs will not be repeated here.

### 2.2.2.1 Data Used

The radiometric data source is the SEDS meteorological station statistics printout. A typical printout was shown previously in table 2-II. The paired values for TSDK and day VIS for all stations are extracted. In general, two weeks of SEDS data are required to be sure of obtaining a sufficient amount of data. The data are key punched for the program ISOCLS/PRINT in the following format:

- Col 1-3 – Sequence number of the data card 1, 2, 3... (not required by program)
- Col 9-12 – TSDK as four digits without decimal
- Col 25-27 – VIS as three digits without decimal.

The primary source of data for verification is a report of sky conditions supplied by the National Weather Service.

### 2.2.2.2 Analysis Used

The analysis used is a two-step procedure. The raw data is classified into five classes by ISOCLS/PRINT. The samples in each class are then analyzed by DISCRM to provide the unscaled discriminant coefficients.

The program ISOCLS/PRINT performs an unsupervised classification, often called clustering, of the input data. The program will accept as input a set of cluster means as initialization values. The number of iterations required is greatly reduced by this initialization, and in this case the quality of the final result is better. The input cluster means are one of the results of the previous calculation of classification coefficients.

For example, the analysis for the spring evaluation was performed on 15-29 March 1975 SEDS data. The program ISOCLS/PRINT was initialized with the following set of class means which had to be calculated for late summer data.



MEANS		
<u>Cluster</u>	<u>Visible</u>	<u>Infrared</u>
1	36.34	304.44
2	44.52	291.26
3	70.42	264.58
4	115.60	278.77
5	116.17	236.83

Table 2-XI shows the statistical results of ISOCLS/PRINT after six iterations. As can be seen, the class numbers have been reordered and the class means have in some cases undergone considerable change. The clear class is the one with the highest TSDK value and the lowest VIS value. As TSDK decreases and VIS increases, the classes represent successively higher and more dense clouds.

ISOCLS/PRINT also outputs a list of the members of each class. The data cards are reordered on the basis of this list and input to the program DISCRM. DISCRM corrects the TSDK data to the eight-bit coded TSD used in SEDS. The VIS data is unchanged. A sample of the program output is shown in table 2-XII. The mean scores are the resulting class means for each variable. The discriminant coefficients are given for each class. Additionally, the program produces a table of the classification results for all input data. If the ISOCLS/PRINT program worked correctly, only a few outlying points will be classified incorrectly.

The unscaled and scaled coefficients as used by SEDS are shown in table 2-XIII. The scale factor is found by dividing the largest multiplicative term into 255. In this example  $255/.98661 = 258.46079$ . This algorithm is the primary cloud detector for the day radiometric data. A portion of the decision field has

TABLE 2-XI. -- TYPICAL STATISTICS PRODUCED BY  
ISOCLS/PRINT FOR 15-29 MARCH 1975 SEDS DATA  
(SELECTED STATIONS ONLY)

INTERMEDIATE PRINTOUT FOR ITERATION 6

TOTAL NUMBER OF CLUSTERS = 5  
TOTAL NUMBER OF POINTS = 105

CLUSTER	SYMBOL	POINTS IN CLUSTER
1	1	26
2	2	9
3	3	18
4	4	13
5	5	39

MEANS

CLUSTER	CH( 1)	CH( 2)
1	296.25	47.08
2	261.28	59.22
3	278.75	114.56
4	235.77	107.85
5	303.22	34.74

STANDARD DEVIATIONS

CLUSTER	CH( 1)	CH( 2)
1	6.80	7.95
2	12.96	7.55
3	12.80	15.98
4	10.59	20.21
5	6.43	4.34

DISTANCES BETWEEN CLUSTERS

CLUSTER	1	2	3	4	5
1	.00	4.04	6.27	8.59	2.35
2	4.04	.00	5.22	4.50	6.27
3	6.27	5.22	.00	3.71	9.95
4	8.59	4.50	3.71	.00	11.30
5	2.35	6.27	9.95	11.30	.00

ORIGINAL PAGE IS  
OF POOR QUALITY

ORIGINAL PAGE IS  
OF POOR QUALITY

TABLE 2-XII. — SAMPLE OUTPUT OF DISCRM FOR 15-29 MARCH 1975 SEDS DATA .

GROUP	1	2	3	4	5	TOTAL
SAMPLE	26	9	18	13	39	105
MEAN SCORES						
1	188.50000	118.55556	153.50000	67.53846	202.43590	
2	47.07692	59.22222	114.55556	107.84615	34.74359	
FUNCTION	1	2	3	4	5	
COEFFICIENT						
1	.57092	.37460	.50486	.25223	.60485	
2	.47117	.53059	.98661	.88849	.38091	
CONSTANT	-64.90030	-37.91668	-95.25861	-56.42783	-67.83849	

CLASSIFICATION MATRIX

FUNCTION	1	2	3	4	5	TOTAL
GROUP						
1	23	0	0	0	3	26
2	1	8	0	0	0	9
3	0	0	18	0	0	18
4	0	1	0	12	0	13
5	1	1	0	0	37	39

already been shown in figure 2-3. The uses of this field are also discussed in section 2.1.2.1.

TABLE 2-XIII. — SCALED AND UNSCALED SEDS  
RAINFALL CLASSIFICATION COEFFICIENTS

Unscaled		
$a_{10} = -67.83849$	$a_{30} = -37.91668$	$a_{50} = -95.25861$
$a_{11} = .38091$	$a_{31} = .53059$	$a_{51} = .98661$
$a_{12} = .60485$	$a_{32} = .37460$	$a_{52} = .50486$
$a_{20} = 64.90030$	$a_{40} = -56.42783$	
$a_{21} = .47117$	$a_{41} = .88849$	
$a_{22} = .57092$	$a_{42} = .25223$	
Scale factor: 258.46079		
Scaled		
$A_{10} = -17533$	$A_{30} = -98000$	$A_{50} = -24621$
$A_{11} = 98$	$A_{31} = 137$	$A_{51} = 255$
$A_{12} = 157$	$A_{32} = 97$	$A_{52} = 130$
$A_{20} = -16774$	$A_{40} = -14584$	
$A_{21} = 122$	$A_{41} = 230$	
$A_{22} = 148$	$A_{42} = 65$	

### 2.2.3 VERIFICATION

SEDS produces an image of the results of the rainfall classification algorithm. This product is extremely useful in verifying the quality of the classification. Clear and cloudy areas can often be identified with certainty in the isothermal imagery. Any conflict with the rainfall classification is apparent.

One of the key problem areas with this algorithm has been its classification of the highly reflective desert region. Often an otherwise satisfactory set of coefficients will consistently classify the bare soil and sand of the desert as cloud. The rainfall image is vital to identifying this problem when it occurs.

Numerical verification depends upon surface observations of cloud cover. This type of observation is difficult to interpret since it is an integrated average of what the observer sees over the entire hemisphere of the sky. On the other hand, the radiometer views a single kilometer square and may well find the only cloud in an otherwise clear sky.

Table 2-XIV shows an example of the verification of performance of this algorithm with surface data. As shown, the performance of the algorithm can be considered satisfactory. This is an example, of course, any real evaluation would require more data.

TABLE 2-XIV. — COMPARISON OF SURFACE REPORTS  
 FROM TEXAS WITH SEDS CLOUD DETERMINATION  
 FOR 5-11 JUNE 1975 DATA

Surface cloud cover reported in 10ths	SEDS Classification	
	Clear	Cloudy
0-2	7	0
3-5	3	0
6-8	2	7
9-10	2	27

### 2.3 THERMAL THRESHOLDS

The SEDS thermal thresholds are temperature limits used to screen radiometric data for cloud contamination. Any radiometric temperature less than the applicable thermal threshold is not used in the calculation of DMAT estimates as described in section 2.1. The thermal thresholds for the day and night radiometric data are independently calculated and used in SEDS.

#### 2.3.1 BACKGROUND MATERIAL

The problem is to identify cloud-free radiometric data on the basis of the temperature of the emitting surface. This approach makes use of the fact that the temperature of the lower atmosphere decreases with increasing height above the earth's surface. Since the tops of clouds are at the temperature of the air at that level, it is reasonable to expect the radiometric temperature to be less when viewing clouds than when viewing the ground.

The infrared radiometric data processed by SEDS shows temperatures ranging from less than 230°K to over 320°K. Figure 2-4

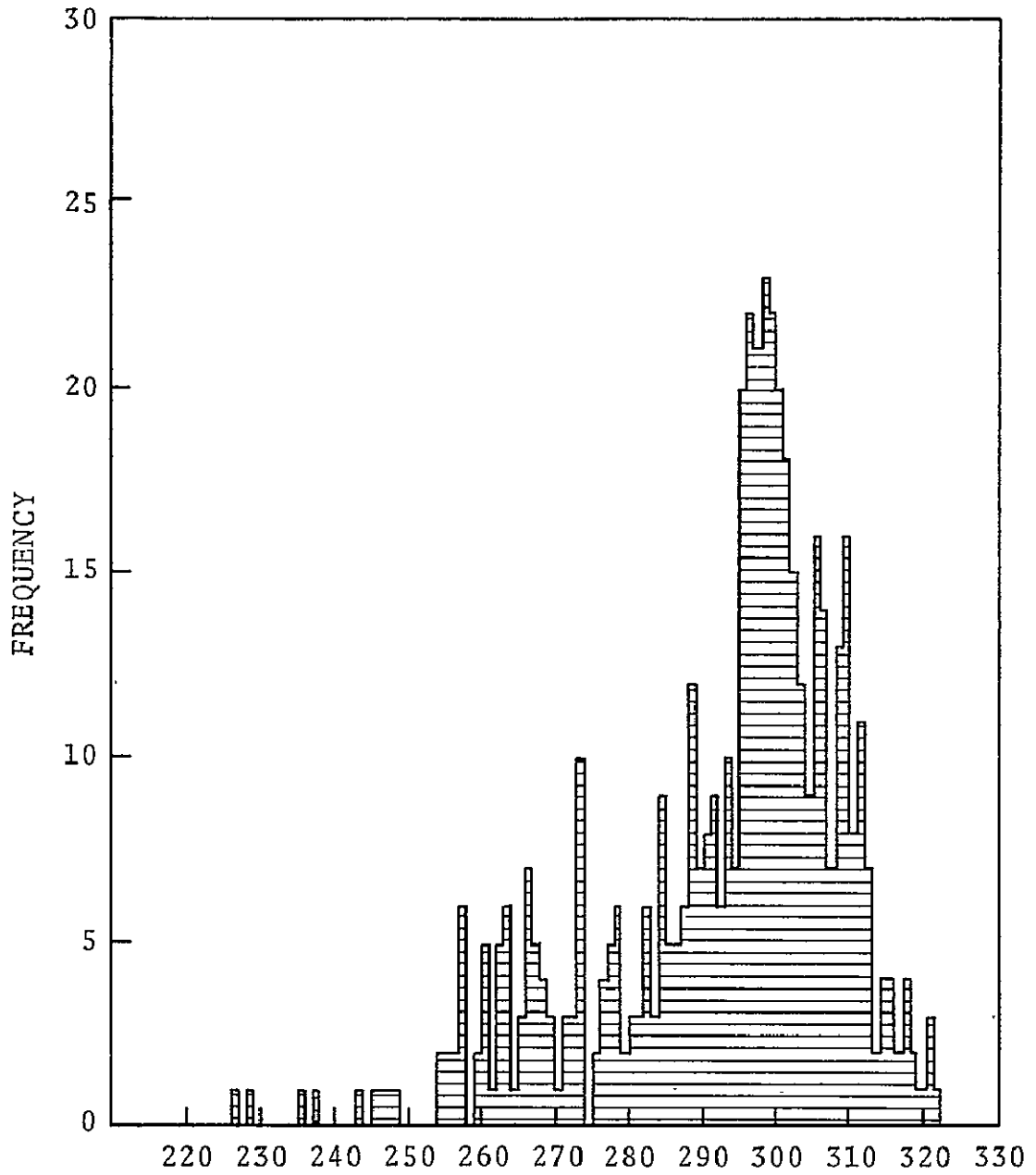


Figure 2-4. — Day radiometric data taken 1-10 August 1975 over the meteorological station network.

shows a histogram of the day radiometric temperatures over meteorological stations in the SEDS image during the 1-10 August 1975 time period. The coldest temperatures are from the tops of the highest clouds, while the warmest come from the desert regions of Mexico and the southwestern U.S.

However, the differences between cloud top and surface temperatures are not always great. In addition, the surface temperature in mountainous areas may well be less than the temperature of clouds occurring at lower altitudes. Careful analysis of histograms of the radiometric data provide estimates of the thermal thresholds.

SEDS software compares each pixel of data with predetermined temperature limits. Any radiometric temperature value greater than the threshold value is considered to be cloud free. The day thermal threshold is generally several degrees warmer than the night threshold. The threshold values change seasonally with the highest values occurring during the fall. All changes in the thermal thresholds are initiated via the SEDS System Parameter Control Form shown previously in figure 2-1.

### 2.3.2 METHOD OF CALCULATION

The basic method of calculation uses a histogram constructed with data taken from the SEDS meteorological station statistics printout. Fine tuning of the threshold values may be carried out by examining the residual errors of the DMAT estimates.

A given set of thermal thresholds remain valid for 4-6 weeks. Recalculation is generally carried out in conjunction with determination of the DMAT coefficients.



### 2.3.2.1 Data Used

The data source is the daily SEDS meteorological station statistics printout shown previously in table 2-II. The TSD and TSN data are extracted for all stations over a one-week period. During extremely cloudy periods, it may be necessary to use two weeks of data in order to have a significant number of cloud-free points.

### 2.3.2.2 Analysis Used

An analysis of histograms of the raw radiometric data is the approach used. Complete histograms are not required since extremely cold temperatures may be assumed to represent clouds. Separate histograms are plotted for the day and night data. Figures 2-5 and 2-6 show examples of these plots for the 28 September to 9 October 1975 time period.

The best thermal threshold value has been found to be 5-7 degrees less than the first major frequency maximum occurring for temperatures greater than 280°K. Table 2-XV shows these values for our example.

Table 2-XV. - TEMPERATURE OF FIRST CLOUD-FREE FREQUENCY MAXIMUM AND ESTIMATED THERMAL THRESHOLD FOR 28 SEPTEMBER TO 9 OCTOBER 1975 SEDS DATA.

	<u>Temperature of Frequency Maximum</u>	<u>Thermal Threshold</u>
Day	303°	296°
Night	291°	285°

The histogram itself indicates the correct choice by a noticeable discontinuity.

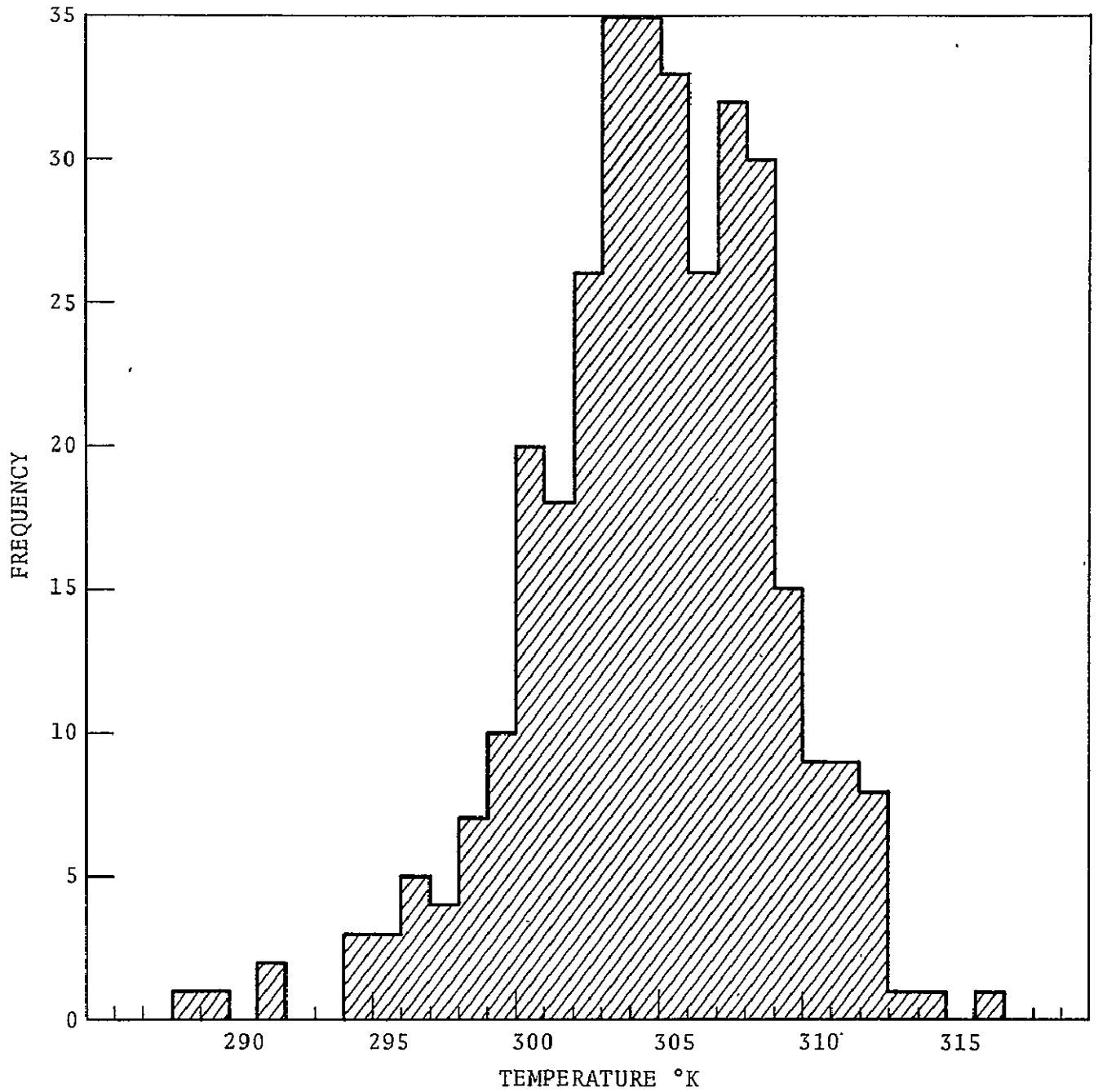


Figure 2-5. - Histogram of day radiometric data  
28 September - 9 October 1975.

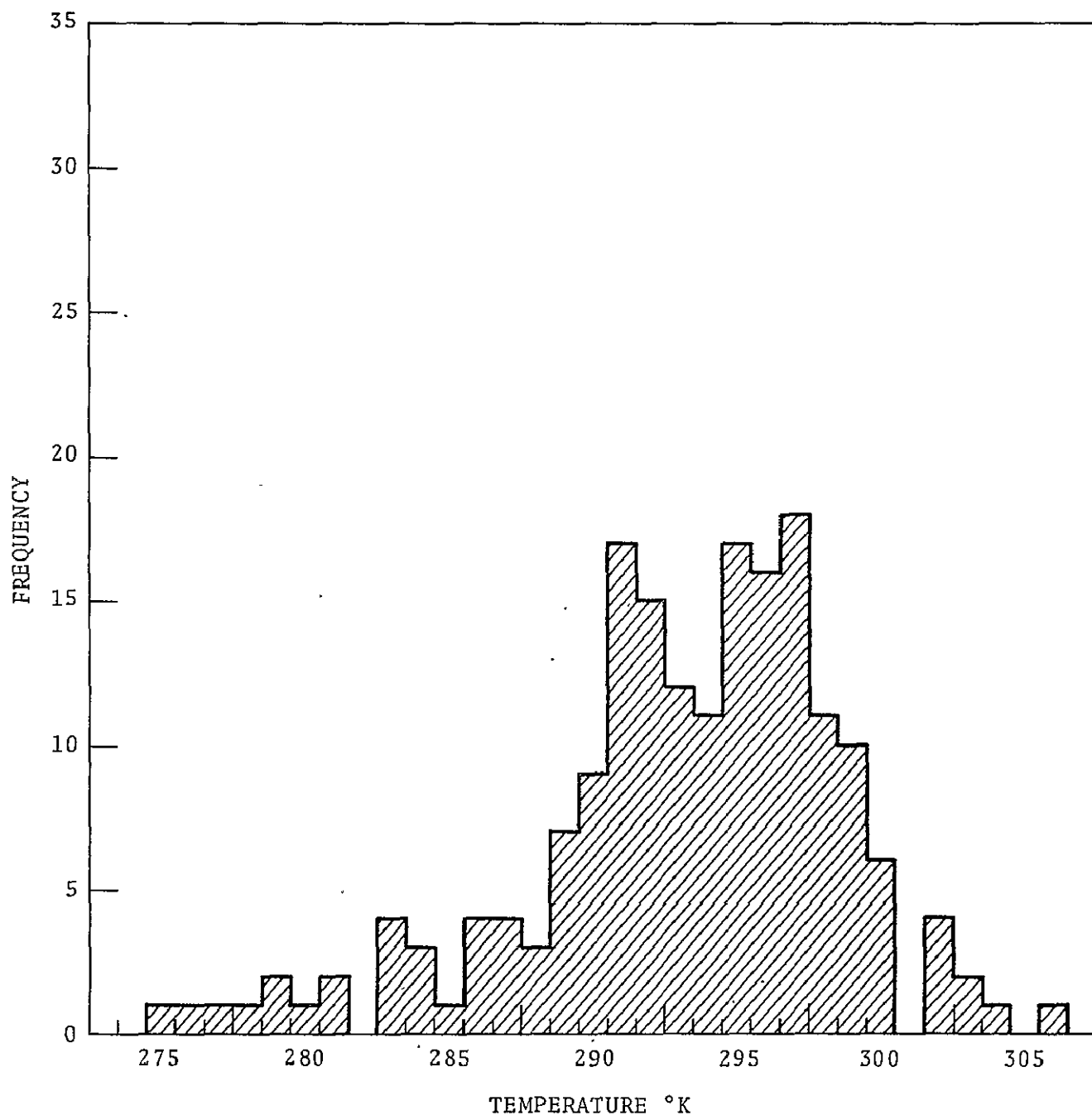


Figure 2-6 - Histogram of night radiometric data  
28 September - 9 October 1975.

### 2.3.3 VERIFICATION

The evaluation of the thermal thresholds are linked to the evaluation of the DMAT estimates as described in section 2.1.3. When large rms errors are noted in the DMAT estimates, the radiometric data should be histogrammed in order to determine if the thermal thresholds should be changed.

When the temperatures are increasing in the spring through late summer, an analysis of the residual errors associated with the DMAT estimates provides additional information. When the residual error is consistently large for temperatures just above the existing threshold, it is a clear indication that recalculation is necessary.

## 2.4 COMBINED SCREWORM PRODUCT COEFFICIENT

The SEDS combined screwworm product is a relatively late developmental product. Its purpose is to provide a single test estimate of the potential growth of a generation of screwworms as a function of weather conditions. SEDS was designed to provide four separate estimates of the potential population growth detailing the effect of temperature on adult and pupae survival, of soil moisture on pupae survival, and of mean pupation time. The combined product was implemented to provide a regression based combination of the four potential growth estimates.

### 2.4.1 BACKGROUND MATERIAL

This section is intended to give an overview of the development leading up to the combined screwworm product, which is still very much in the developmental phase. Although the mathematical form of the algorithm implemented in SEDS has been statistically verified, the regression coefficients are derived from simulations based on historical surface data. In addition, the

well known limitations on applying regression results outside of their developmental area make the general applicability of the combined product a matter for further research.

#### 2.4.1.1 Independent Variables

There are four variables calculated by SEDS that serve as the basic information from which the combined screwworm product is calculated. These variables are recalculated and updated daily for each pixel in the 625 element by 550 line data base array. The following is a brief summary of the manner in which SEDS handles each variable.

The short term mean air temperature (STMAT) depicts the effect of daily mean air temperature on the growth and survival of the adult screwworm fly. Computationally, the STMAT is determined by the following low pass filter.

$$\text{STMAT} = .44 \text{ DMAT} + .56 [\text{STMAT}(-1)]$$

where DMAT - Eight-bit coded values of the daily mean air temperature estimate as described in section 2.1.1.2.

STMAT(-1) - The STMAT value calculated the previous day.

The long term mean air temperature (LTMAT) shows the effect of daily mean air temperature on the growth and survival of the screwworm pupae. Computationally, the LTMAT is the running sum of the DMAT estimates over the last N days.

The crop moisture index (CMI) is a measure of soil moisture deficit as described in section 3.1. The CMI variable used in SEDS is an N-day sum of the daily CMI values. This variable is intended to depict the effect of soil moisture on the screwworm pupae.

The degree day sum (DDSUM) variable is an estimate of the screw-worm's mean pupation time. This variable is used directly as an independent variable and to estimate generation length. Computationally, the DDSUM is an N-day accumulation of the screw-worm degree day values.

$$\text{DDSUM} = \sum_{L=1}^N (\text{DMAT}(i) - 284) \text{ when } \text{DMAT}(i) > 284^{\circ}\text{K}.$$

The application of the degree day sum to screwworms is given elsewhere (ref. 1).

The STMAT, LTMAT, CMI, and DDSUM values are subject to additional calculation and coding prior to the determination of the combined product.

The STMAT is converted to a growth potential value using a look-up table based on figure 2-7. The development of this figure is given by Phinney (ref.1) and discussed briefly in section 1.1. The growth potential is encoded by

$$X_1 = \text{integer} [(100 \ln \text{GP}) + 128.5]$$

The LTMAT sum is divided by N to determine the mean and converted to a growth potential again using figure 2-7. The growth potential is encoded by

$$X_2 = \text{integer} [(100 \ln \text{GP}) + 128.5]$$

The CMI sum is divided by N to determine the mean and converted to a growth potential using a look-up table based on figure 2-8. The growth potential is encoded by

$$X_3 = \text{integer} [(100 \ln \text{GP}) + 128.5]$$

The DDSUM is used to calculate mean pupation time (MPT) using the following formula.

$$\text{MPT} = N(139.1/\text{DDSUM})$$

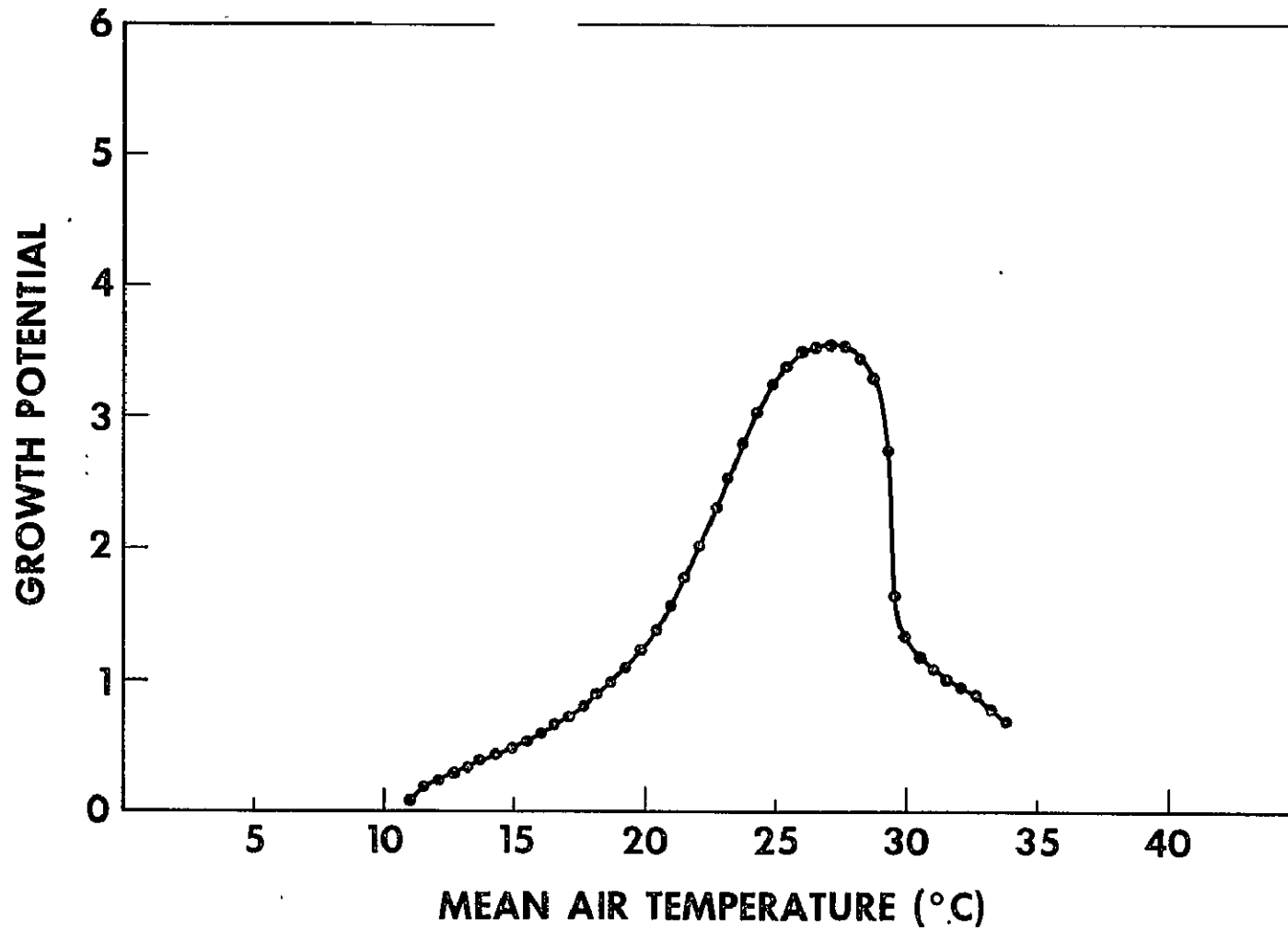


Figure 2-7.— Relation between mean air temperature and growth per generation.

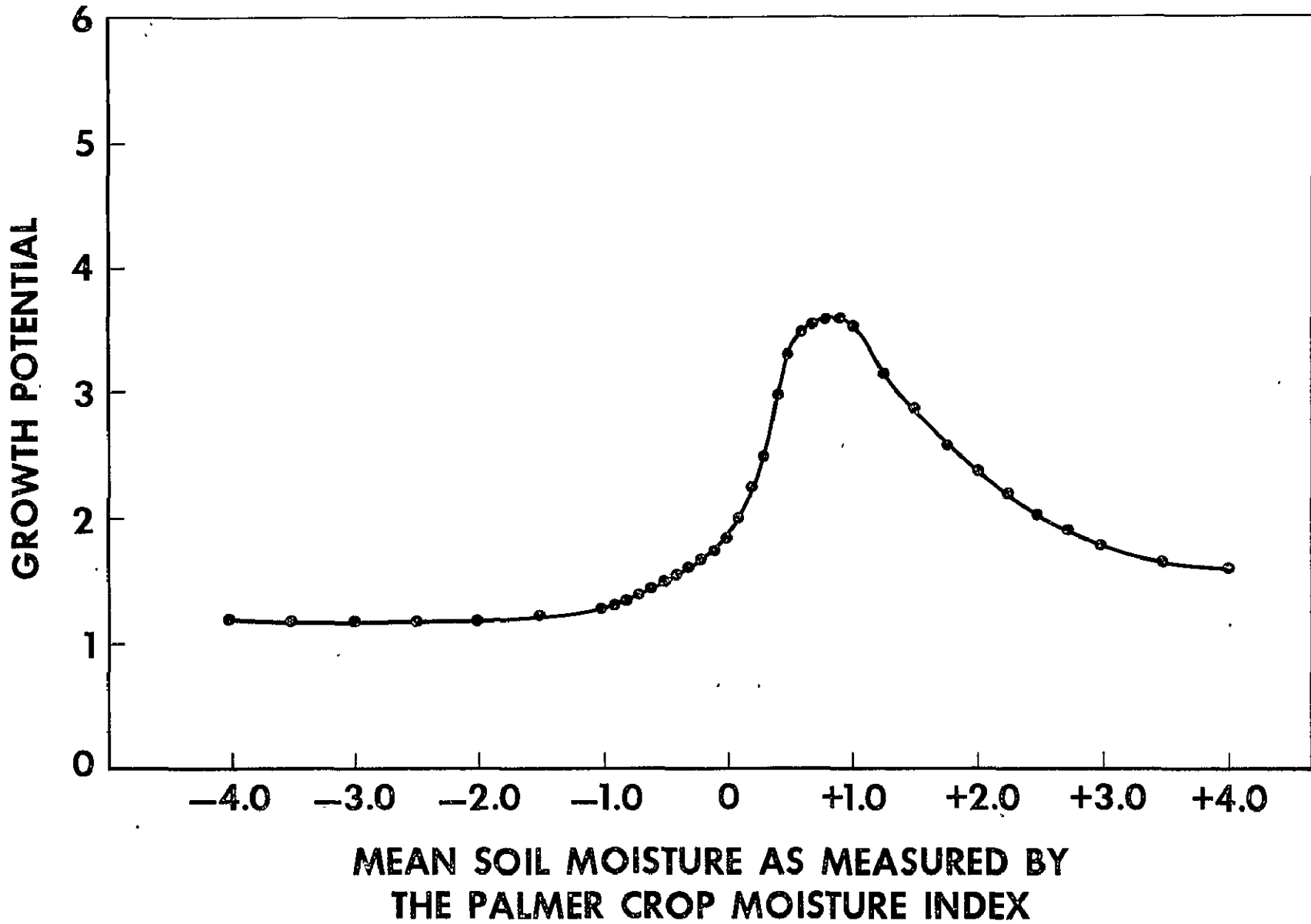


Figure 2-8.— Relationship between CMI and growth potential.



The estimated generation length is  $MPT + 15$ . The generation length is converted to a growth potential value using a look-up table based on figure 2-9. The growth potential is encoded by

$$X_4 = \text{integer} [(100 \ln GP) + 128.5]$$

The variables available for calculating a combined screwworm growth potential are

$X_1$  - Eight-bit coded growth potential  
calculated from the STMAT

$X_2$  - Eight-bit coded growth potential  
calculated from the LTMAT

$X_3$  - Eight-bit coded growth potential  
calculated from the CMI

$X_4$  - Eight-bit coded growth potential  
calculated from the DDSUM

#### 2.4.1.2 Model Selection

The selection of a multiplicative model for the SEDS combined screwworm model was based on computation and statistical expediency as well as mathematical plausibility. Statistically, it was desirable to have an intrinsically linear model so that multiple linear regression analysis could be used. The multiplicative model may be written in the following form

$$Y = \alpha_0 X_1^{\alpha_1} X_2^{\alpha_2} X_3^{\alpha_3} X_4^{\alpha_4} \epsilon \quad (2-22)$$

where  $\alpha_0$ ,  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$ , and  $\alpha_4$  are parameters and  $\epsilon$  is the random error. Taking the natural logarithms of Eq.(2-22) converts the model into a linear form.

$$\ln Y = \ln \alpha_0 + \alpha_1 \ln X_1 + \alpha_2 \ln X_2 + \alpha_3 \ln X_3 + \alpha_4 \ln X_4 + \ln \epsilon \quad (2-23)$$

The transformed model given in Eq.(2-23) is linear. Linear multiple regression may be used to calculate estimates of the parameters.

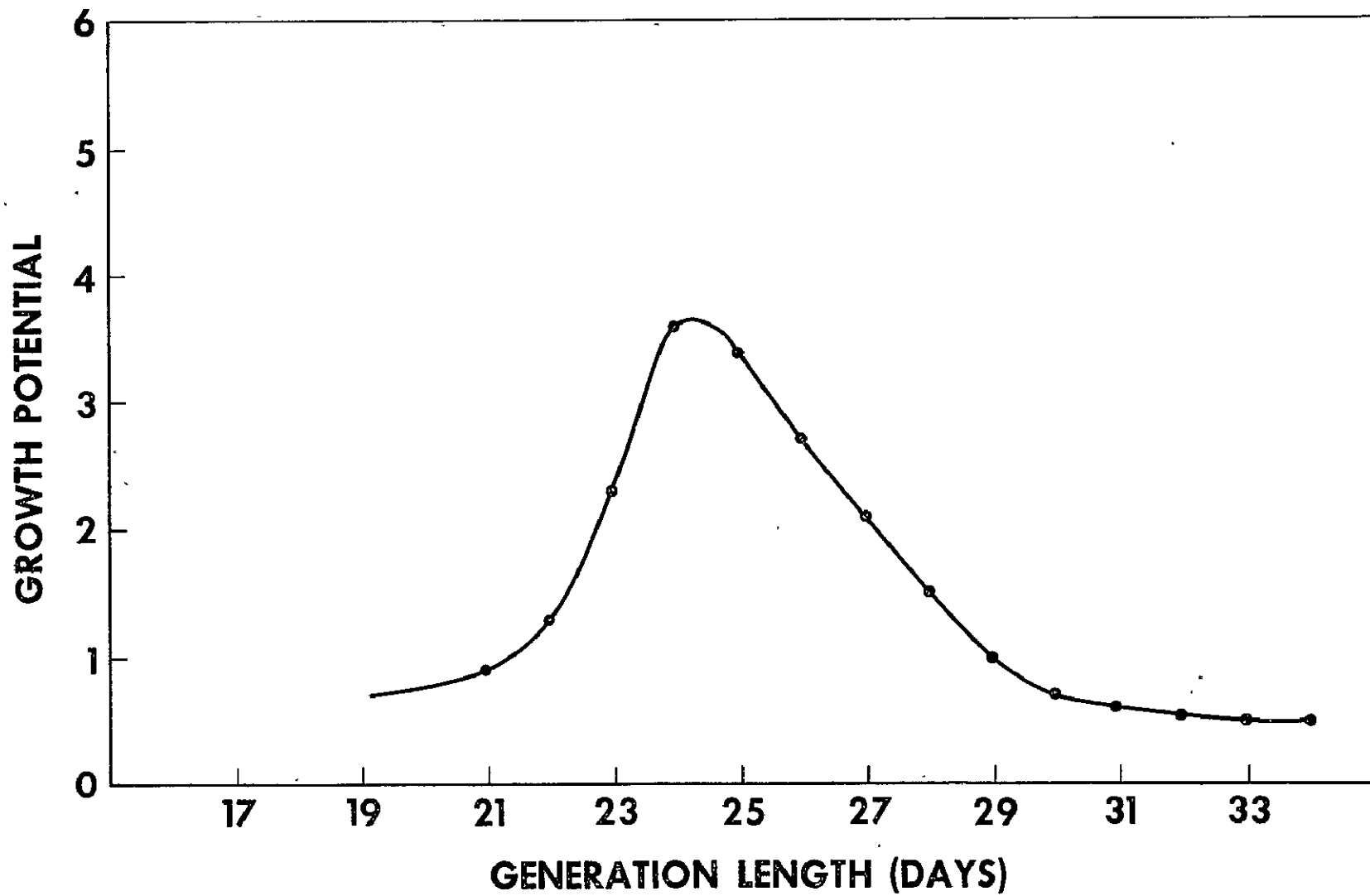


Figure 2-9.— Relationship between generation length and growth potential.

Computationally, this form was desirable in that the data base values to be used as independent variables were contained in SEDS in logarithmic form.

#### 2.4.1.3 Model Verification

In an effort to determine the validity of the mathematical form of the combined screwworm product algorithm, three basic models were examined. First was the linear model which may be written

$$Y = \alpha_0 + \alpha_1 GP(\bar{T}_A) + \alpha_2 GP(\bar{T}_p) + \alpha_3 GP(\overline{CMI}_p) + \alpha_4 GP(GEN) + \epsilon$$

where  $GP(\bar{T}_A)$  - Growth potential of the mean air temperature

$GP(\bar{T}_p)$  - Growth potential of the mean air temperature during the pupae phase

$GP(\overline{CMI}_p)$  - Growth potential of the mean CMI value during the pupae phase

$GP(GEN)$  - Growth potential of the estimated generation length.

Second was an exponential model which may be written

$$Y = \exp [\alpha_0 + \alpha_1 GP(\bar{T}_A) + \alpha_2 GP(\bar{T}_p) + \alpha_3 GP(\overline{CMI}_p) + \alpha_4 GP(GEN)] \epsilon.$$

In linear form, this model may be written

$$\ln Y = \alpha_0 + \alpha_1 GP(\bar{T}_A) + \alpha_2 GP(\bar{T}_p) + \alpha_3 GP(\overline{CMI}_p) + \alpha_4 (GEN) + \ln \epsilon.$$

The third was the multiplicative model implemented in SEDS which was given in section 2.4.1.2.

These models were evaluated using the set of historical data used to develop the growth potential functions as described in Phinney (ref.1). Case reports less than ten per week were eliminated to reduce the effect of sampling error. It should be emphasized that although similar, the variables are not exactly the same as the SEDS data base variables. The significance of these models as evaluated by the F test described in section 2.1.2.2 may be seen in table 2-XVI.

TABLE 2-XVI. — F TEST OF SIGNIFICANCE FOR THREE POSSIBLE SCREWORM REGRESSION MODELS

<u>Model</u>	<u>df</u>	<u>F</u>
Linear	(4,185)	8.5
Exponential	(4,185)	15.0
Multiplicative	(4,185)	16.7

These results suggest that for relatively simple models utilizing these variables that the choice of SEDS algorithm was correct.

#### 2.4.1.4 Developmental Phase Coefficients

During the developmental phase of this algorithm insufficient SEDS data were available to derive estimates of the parameters for Eq.(2-23). Rather than use arbitrary coefficients, it was decided to use those derived during the model evaluation described in section 2.4.1.3. Table 2-XVII gives the coefficients derived during that study.

TABLE 2-XVII. - COEFFICIENTS DERIVED FOR  
MULTIPLICATIVE MODEL

<u>Parameter</u>	<u>Coefficient</u>	<u>Computed t-test value</u>
$\ln\alpha_0$	-1.58553	
$\alpha_1$	.37046	2.644*
$\alpha_2$	-.17782	-.893
$\alpha_3$	.29818	2.660*
$\alpha_4$	.32869	2.050

\*Significant at the .99 level.

For implementation into SEDS  $\alpha_0$  was set equal to one. Additionally, the coefficients in table 2-XVII were scaled to take advantage of the full range of available colors in the output image. The scaling was carried out by dividing the individual coefficient by the sum of all four coefficients. Table 2-XVIII shows the coefficients used during the developmental phase.

TABLE 2-XVIII. - DEVELOPMENTAL PHASE COMBINED SCREWORM  
PRODUCT COEFFICIENTS

<u>Variable</u>	<u>Parameter</u>	<u>Coefficient</u>
Intercept	$\alpha_0$	1
STMAT	$\alpha_1$	.46
LTMAT	$\alpha_2$	-.22
CMI	$\alpha_3$	.36
DDSUM	$\alpha_4$	.40

#### 2.4.2 COEFFICIENT CALCULATION AND VERIFICATION

Both the calculation and verification of the SEDS combined screwworm product coefficients are currently under study. Data is currently being summarized in Texas by county for both SEDS and USDA case reports. When sufficient data has been accumulated for statistical significance, multiple regression studies will be carried out.

### 3. CALCULATION OF THE CMI

The crop moisture index (CMI) is a measure of short-term anomalies of soil moisture. Based on a relatively simple two-layer soil moisture model using weekly rainfall and temperature, the CMI is included in SEDS to provide a measure of the effect of soil moisture on the screwworm's pupae phase.

#### 3.1 BACKGROUND MATERIAL

The CMI values used by SEDS are calculated by the same basic program used operationally in the United States by the National Weather Service and published in the Weekly Weather and Crop Bulletin. Historical data for 36 stations were analyzed to provide information on the normal conditions required for weekly operation of the CMI for SEDS. The existing program is implemented on the JSC Univac 1110 system. Operating instructions for the current program along with sufficient documentation for implementation on other machines may be obtained from C. M. Barnes, Health Applications Officer, Life Sciences Directorate, NASA, JSC.

##### 3.1.1 SOIL MOISTURE MODEL

The basic CMI model assumes a hydrological balance for the soil. The amount of water in the soil can be calculated from a balance which includes precipitation, evapotranspiration, storage, runoff, and percolation through the soil.

The model assumes that the first inch of water is in the uppermost layer of the soil, the depth of which varies from one to eight inches depending upon the soil type. The remaining water is in a second layer. Precipitation fills the first layer followed by the second. Amounts over the holding capacity of the soil are disposed of by runoff and percolation.

Water loss through evapotranspiration is calculated by the Thomthwaite method in which potential evapotranspiration is a function of mean air temperature and day length. Losses are at full potential rate until the first inch of water is depleted. The second layer gives up water at a slower rate which is an empirical function of the ratio of water held to the water capacity.

The Palmer Drought Index (ref. 2) was developed to give a measure of long-term moisture anomalies which would be useful for determining the scope, severity, and frequency of prolonged droughts. This program determines the climatological norms from which deviations are measured. These norms are also required for the CMI analysis.

The historical analysis was performed for 32 Mexican and 4 United States stations shown in table 3-I. The required data included a historical set of monthly mean temperature and total precipitation along with station latitude and water holding capacity of the soil. The set of norms can be extrapolated in a linear manner to nearby stations which lack sufficient climatological records to allow direct computation.

A master data file was established on the JSC Univac 1110 for the stations shown in figure 3-1. This file contains the historical averages against which current weather is compared.

### 3.1.2 SEDS UTILIZATION

The weekly run of the CMI program provides CMI values for the 38 stations shown in figure 3-1. These values are analyzed by the National Weather Service. The land analysis is then prepared for digitization, digitized, and input into SEDS as a magnetic tape, as described in section 3.4.



TABLE 3-1. - SURFACE METEOROLOGICAL STATIONS  
USED IN DEVELOPMENT OF CMI

<u>STATION NAME</u>	<u>STATION NUMBER</u>	<u>LATITUDE</u>	<u>LONGITUDE</u>	<u>ELEVATION</u>
ACAPULCO	76805	16 50N	99 56W	3 M
CAMPECHE	76695	19 51N	90 32W	25 M
CHIHUAHUA	76225	28 38N	106 04W	1423 M
CIUDAD LAS CASAS	76845	16 44N	92 38W	2128 M
CIUDAD LERDO	76381	25 30N	103 32W	1140 M
CIUDAD OBREGON	76746	18 32N	92 39W	2 M
COATZACOALCOS	76741	18 09N	94 25W	14 M
ENSENADA	76050	31 53N	116 38W	15 M
GUADALAJARA	76612	20 41N	103 20W	1589 M
GUANAJUATO	76577	21 01N	101 15W	2073 M
GUAYMAS	76255	27 55N	110 53W	4 M
JALAPA	76687	19 32N	96 55W	1399 M
LA PAZ	76405	24 10N	110 12W	18 M
LEON	76575	21 07N	101 41W	1809 M
MANZANILLO	76654	19 04N	104 20W	8 M
MAZATLAN	76458	23 11N	106 25W	78 M
MERIDA	76644	20 58N	89 38W	22 M
MONCLOVA	76342	26 55N	101 25W	587 M
MONTERREY	76393	25 40N	100 18W	534 M
MORELIA	76665	19 42N	101 07W	1923 M
OAXACA	76775	17 04N	96 43W	1563 M
PIEDRAS NEGRAS	76243	28 42N	100 32W	220 M
PROGRESO	76593	21 17N	89 40W	14 M
PUEBLA	76685	19 02N	98 11W	2150 M
SALINA CRUZ	76833	16 12N	95 12W	56 M
SALTILLO	76390	25 26N	101 00W	1609 M
TACUBAYA	76679	19 24N	99 11W	2309 M
TAMPICO	76548	22 12N	97 51W	20 M
TAPACHULA	76903	14 54N	92 16W	168 M
TORREON	76382	25 34N	103 25W	1150 M
VERACRUZ	76692	19 12N	96 08W	16 M
ZACATECAS	76525	22 47N	102 34W	2612 M
BROWNSVILLE	72250	25 54N	97 26W	10 M
EL PASO	72270	31 48N	106 24W	1200 M
SAN DIEGO	72290	32 44N	117 10W	11 M
YUMA	72280	32 40N	114 36W	62 M

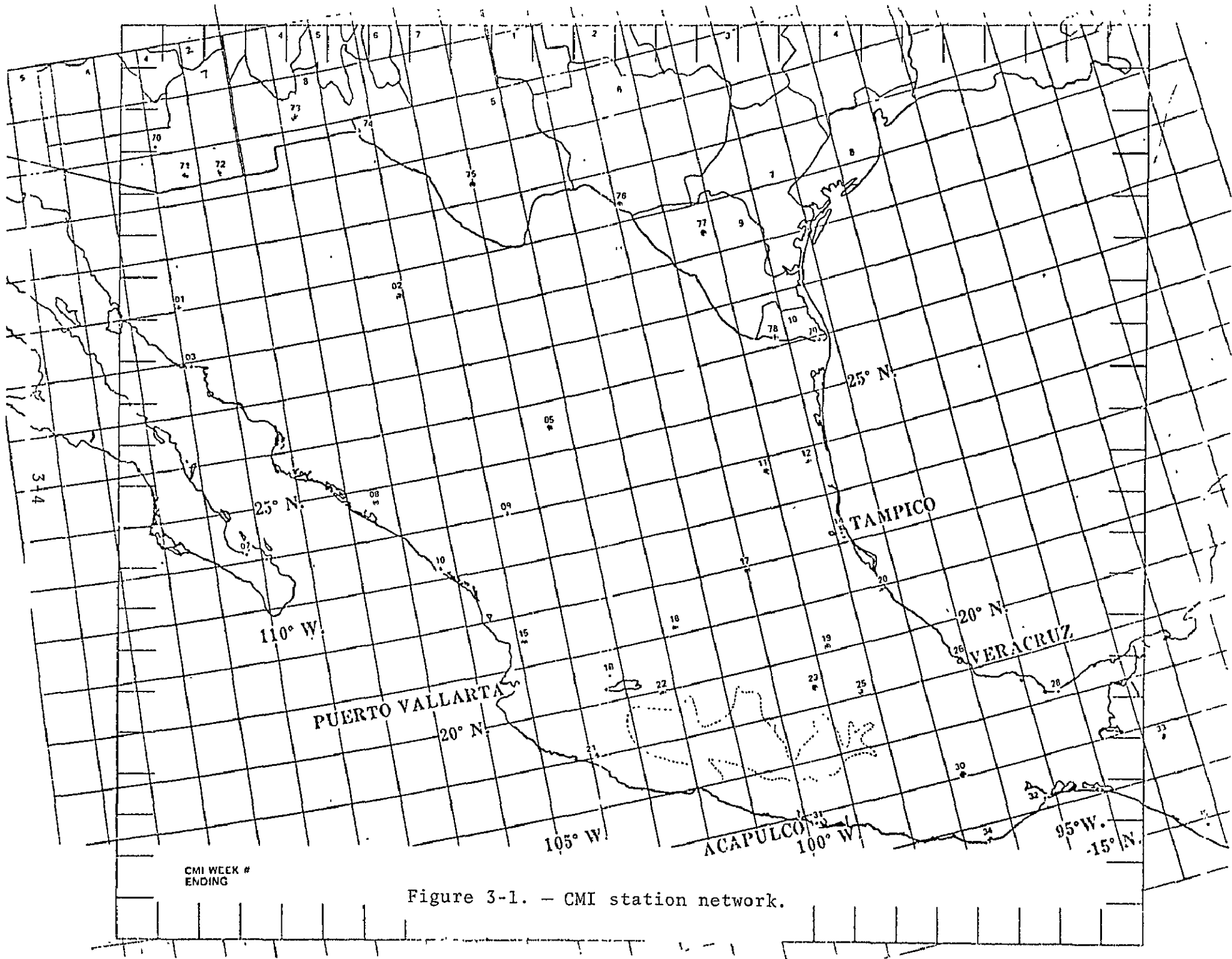


Figure 3-1. - CMI station network.

The digital CMI values are used on a point-by-point basis in the SEDS data base to calculate a growth potential image and to prepare the combined product as described in section 2.4.1.1.

### 3.2 DATA HANDLING

The basic source of input data is the daily SEDS Surface Observations shown in table 2-IV. The data is summarized for Sunday through Saturday each week in the form of average temperature and total precipitation. The data is keypunched and submitted to the JSC Univac 1110 for processing.

Detailed documentation of the data handling procedures may be obtained from C. M. Barnes, Ph. D., D.V.M.

### 3.3 VERIFICATION

Verification of the CMI product depends upon indirect methods; of primary importance is feedback from field observations. Table 3-II giving a generalized interpretation of the CMI in terms of field conditions can be invaluable in both original analysis and verification.

The CMI values change relatively slowly. Therefore, continuity in the CMI patterns should be expected. Discontinuity from the previous weeks analyses should be expected immediately following major occurrences of rainfall.

The U.S. analysis of CMI is useful in the Mexican-U.S. border regions. The U.S. operational analysis is calculated from meteorological data that has been averaged over climatological divisions, representing area averaged CMI values as opposed to the SEDS point values. However, in the regions of overlap between the two analyses good agreement in analysis should be expected.

TABLE 3-II. — INTERPRETATION OF CMI VALUES  
IN TERMS OF FIELD CONDITIONS

<u>Index</u>	<u>Interpretation</u>
Above 3.0	Excessively wet, some flooding
2.0 to 3.0	Too wet, some standing water
1.0 to 2.0	Wet in spots
0 to 1.0	Normal
0 to -1.0	Topsoil moisture short
-1.0 to -2.0	Abnormally dry
-2.0 to -3.0	Too dry
Below -3.0	Exceptionally dry

#### 3.4 PREPARATION OF CMI TAPES

There were several candidate methods for preparing digital tapes containing CMI data. All were designed to take raw CMI data for certain stations in Mexico calculated as described in section 3.2 to produce a tape containing CMI values for every pixel.

The first alternative was the most general. It involved the interpretation of these data by a meteorological technician, who prepared an isoline map, which assigned CMI values over the entire SEDS grid. This image was then colored, read into a digital computer, the colored zones classified by standard techniques, and each pixel assigned a CMI value according to its class. This procedure is documented extensively in L. Giddings and E. Thompson, "An Interactive Method for Digitizing Zone Maps," (LEC-6498; JSC-09809).

This method was chosen for development since it was the most general. It allowed the preparation of digital tapes of any data that could be represented as zonal maps. To eliminate some dependence on a trained operator this method was largely automated, using standard LARSYS classification program. This required some modifications which are described in L. Giddings and S. Boston, "Digitizing Zone Maps using a Modified LARSYS Program," (LEC-7498; JSC-10757). This modification converted long operation into a relatively short one.

Other methods could also be used. The most feasible one would involve direct interpolations of the original data, or direct interpolation of data extracted from an interpreted image. For example, raw CMI values for selected stations (grid points) could be interpolated in two dimensions to fill the SEDS grid; alternatively, the raw values could be interpreted by a meteorologist, and data points extracted from the interpretation could be interpolated to fill the SEDS grid. Development of these methods has not yet been completed.

## PROCESSING OF EMISSIVITY ZONE MAPS

### 4.1 INTRODUCTION TO THE PROBLEM

The SEDS screwworm algorithms are dependent upon the accuracy of the NOAA satellite measured ground surface temperatures. For accurate measurement, radiometric temperatures taken by the satellite must be corrected for certain errors, including the emissivity characteristics of the measured surface (figure 4-1). This section describes the emissivity correction utilized in SEDS and elaborates on the procedures for calculating the correction.

The thermal infrared emissivity,  $\epsilon$ , of a homogeneous surface is defined (ref. 3) as the ratio of the emittance of the surface at a specified wavelength and emitting temperature to the emittance of an ideal blackbody at the same wavelength and temperature. Values for emissivity may range from zero to unity.

To correct the problem, it is necessary to measure the emissivity of each surface and use the measured value to correct any radiometric measurements of that surface. A complete analysis of the emissivity problem and procedures for analysis are found in Phinney and Arp (ref. 4).

### 4.2 DEVELOPMENT OF THE EMISSIVITY ZONE MAP

A map of emissivity values for every pixel of SEDS is necessary to provide the correct emissivity for every portion of Mexico. Since it is not possible to measure the emissivity of each portion of Mexico, it is necessary to have an efficient mapping technique.

Analysis of the distribution of emissivity values and vegetation previously obtained in Veracruz reveals the following correlation.

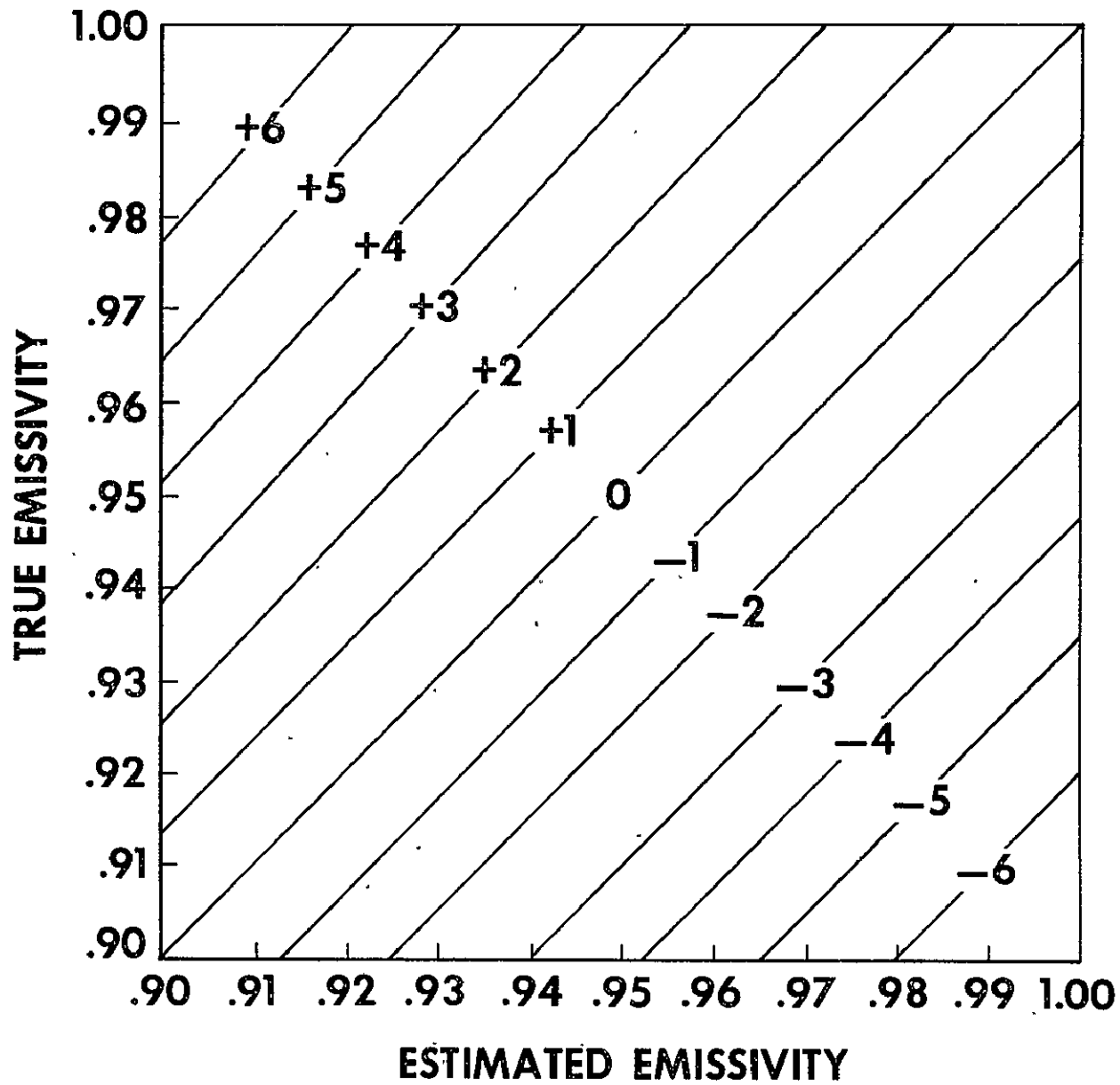


Figure 4-1. — Temperature error ( $^{\circ}\text{C}$ ) associated with an incorrect assumption of emissivity at  $300^{\circ}\text{K}$ .

Variation in emissivity is primarily in response to vegetation and soils. In a study of emissivity by Phinney and Arp (ref. 5), it was shown that visible and infrared data can be composited to produce a zonation of ground data. These zones proved to correspond to the distribution of vegetation in the Veracruz test site. By knowing the distribution of various plant communities (figure 4-2) and their calculated community emissivity values, it is possible to synthesize a map of the emissivity zones of Mexico.

#### 4.3 METHODS OF PRODUCTION OF THE SEDS EMISSIVITY MAP

##### 4.3.1 SOURCES OF EMISSIVITY MEASUREMENTS

In accordance with procedures set forth in Phinney and Arp (ref. 4), emissivity data were field gathered from each major vegetation community and soil type in Mexico. Determinations of the important communities of plants in Mexico were provided by analysis of the works of G. Flores-Mata et al. (ref. 6); J. S. Beard (ref. 7 & 8); F. Miranda and E. Hernandez (ref. 9) 1963; A. S. Starker (ref. 10); and S. R. Eyre (ref. 11).

##### 4.3.2 SOURCES OF RADIOMETRIC DATA

Each day NOAA sends two raw data tapes in analog form to NASA JSC. One tape contains the day visible and day infrared (IR) data; the other tape contains the night IR data. These tapes are registered and converted by Aeronutronic Ford from analog to SEDS digital form to produce three registered data tapes, day IR, day visible, and night IR. The three daily tapes are transferred to the Health Applications Office in Universal format suitable for direct read on the Aerojet Data Analysis Station (DAS). Day IR data are received at full resolution, while day visible and night IR data are compressed. At least two weeks of data are amassed to assure enough data to create a cloud-free image in all three channels.



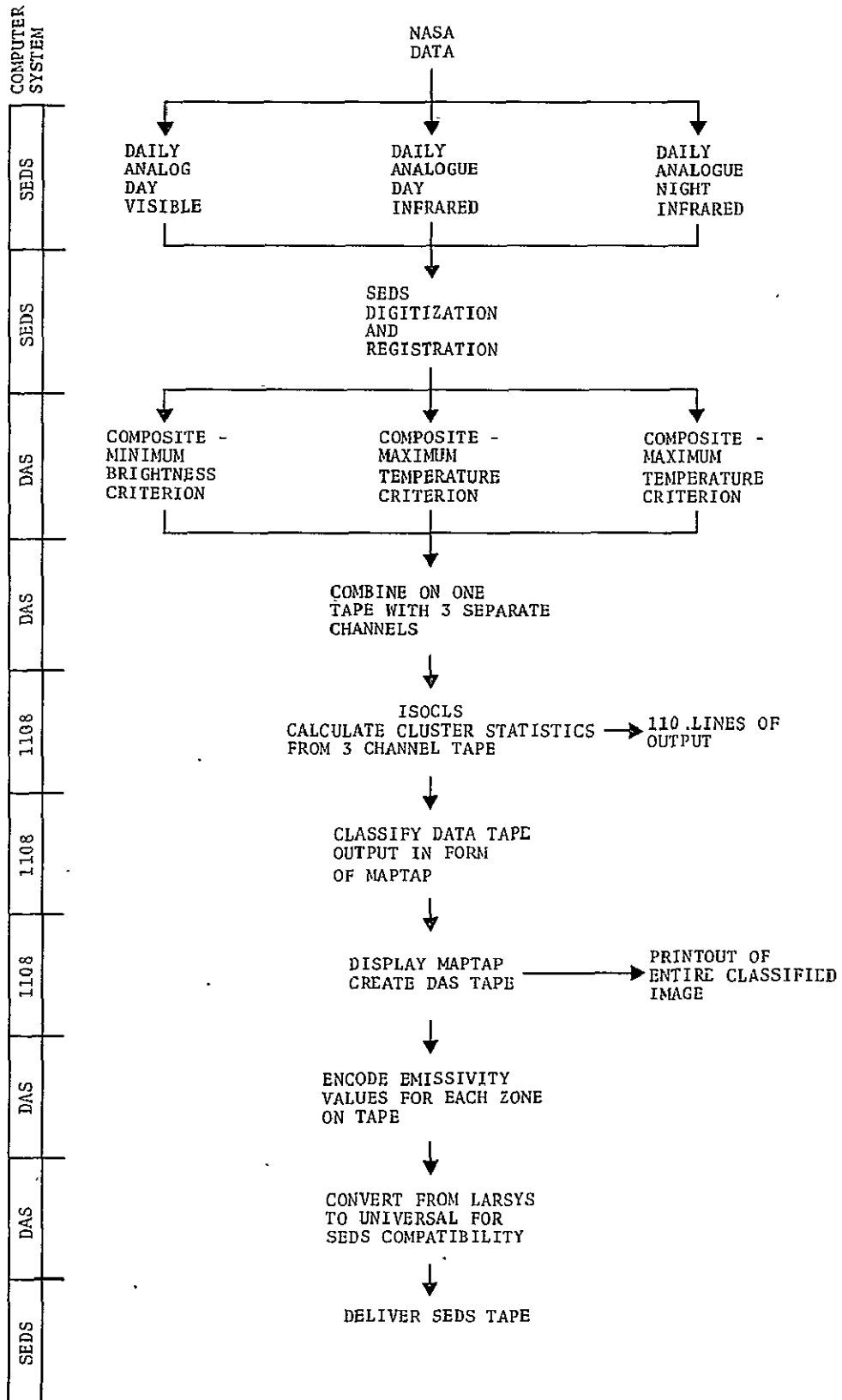


Figure 4-2. - Steps necessary for the creation of an emissivity zone map of Mexico.

ORIGINAL PAGE IS  
OF POOR QUALITY

Tapes are reformatted on the Aerojet DAS from Universal to LARSYS format using the standard system program U2L3AS and combined to make a multifile tape.

#### 4.3.3 SELECTION OF IMAGES FOR USE IN COMPOSITING

The best images for use in compositing are selected by either screening each image on the Aerojet DAS and photographing with a polaroid camera, or by viewing lantern slides of the SEDS data prepared by Aeronutronic Ford. In either case, images are selected so that no two images have clouds in the same regions. In non-cloudy periods, as few as two images may be used to create a cloud-free composite. In cloudy periods, five raw images may be needed to composite each data channel.

#### 4.3.4 CREATION OF COMPOSITED IMAGES TAPE

After the desired images are selected, they are composited onto a disk of the Aerojet DAS by use of the E-DISK routine. For program documentation, contact Dr. C. M. Barnes, Health Applications Office. The visible data is composited using a minimum brightness test, whereby each pixel of the second image is composited with its complementary pixel on the first image. The pixel that has the minimum brightness is recorded, thereby deleting the brightest cloud-contaminated pixels. Infrared data is similarly composited on separate files of the disk. In this case, each pixel of the second image is compared with the complementary pixel of the first image using a maximum temperature criterion. In this case, the warmer of the two pixels is recorded. This process removes any pixels contaminated by cold clouds. Completion of the compositing yields a 3-channel disk which is immediately read onto a 9-track tape, yielding a 3-channel, 9-track tape in LARSYS format for analysis.

#### 4.3.5 CLUSTER ANALYSIS OF COMPOSITED DATA

The composited data tape is treated as a 3-channel multispectral scanner (MSS) tape. The data tape is analyzed using the ISOCLS processor on the UNIVAC 1108 to define those areas of similar temperature and reflectivity. The clustering procedure used in ISOCLS (ref. 12) is an iterative procedure, which assigns each data sample to a specific cluster by determining the nearest cluster center (in absolute distance) and assigning the sample to it.

To begin clustering, all program defaults are used. For the radiometric data under study, about 20 classes will be created by the ISOCLS processor. Of the 20 classes, several will be water classes and several will be cloud classes. To determine the identity of the remaining classes, the interaction of a trained botanist is necessary. Frequently, two or more computer defined classes will correspond to various phases of a single vegetational zone. On occasion, the computer will not distinguish separate vegetative zones. This is frequently due to seasonality effects, but careful study is necessary to avoid serious misclassification of the data.

This process results in the creation of several clusters or regions which have similar temperature and reflectivity values. These regions correspond to vegetation zones and are ultimately presented in a map-like format to which emissivity values are assigned. The clustered results return as a statistics deck and 110 lines of graymap output. Initial verification of the clustered data based upon ground truth evidence may proceed at this point using the graymapped output. However, to receive a complete output of the classified image, the composited data must be processed again.

#### 4.3.6 PRODUCTION OF THE FIRST HARD COPY OF THE EMISSIVITY ZONES

The ISOCLS processor only produces a minimal 110 lines of output for use in interpretation of the clustered data and only analyzes every second or fourth element for purposes of efficiency. To receive all the clustered data, the composited data tape (9-track) plus the previously computed statistics deck are run on the UNIVAC 1108 through the CLASSIFY processor (ref. 12), which classifies the entire image data based upon the statistics computed from the ISOCLS processor. This procedure produces a 7-track tape "MAPTAP". containing every pixel classified according to the statistics generated by ISOCLS.

#### 4.3.7 PRODUCT DISPLAY

The 7-track MAPTAP is run through the 1108 again to produce the first hard copy of the entire image by using a program called DISPLAY. Results of the DISPLAY routine include a 9-track DAS tape for use on the Aerojet DAS, in addition to the printouts of the entire study area (ref. 12).

#### 4.3.8 ASSIGNMENT OF EMISSIVITY VALUES TO ZONES

The DAS tape (9-track) product from the DISPLAY routine is displayed on the Aerojet DAS using the Color Study program. Each zone defined in the ISOCLS routine is displayed and analyzed against known ground truth information. In some cases, a cluster represents a discrete zone on the ground; in other cases, it comprises only part of a larger zone. Identification of zones made from field work and references, and computation of the community emissivity for each zone are described in Phinney and Arp (ref. 4). Assignment of emissivity values is made to each zone by use of the E-MISS program using the formula:

$$f(e) = \text{Integer} \left[ \left( \frac{e \text{ ave}}{e \text{ true}} - 1 \right) (853.3333) + 80.5 \right]$$

$$e \text{ ave} = .975$$

Numbers for each community and the average emissivity value are encoded. The 9-track LARSYS tape is then converted to Universal format and transferred to Aeronutronic Ford for use in processing of the daily radiometric data. For program documentation, contact Dr. C. M. Barnes, Health Applications Group.

#### 4.3.9 CURRENT STATUS

Two tapes have been created to date. The spring tape uses selected data from 15 to 29 March and the fall tape uses selected data from 28 September to 9 October. In theory, no new emissivity tapes need to be made since the vegetation of Mexico will not change measurably in the next year or two.

In practice, however, it is not certain whether or not new tapes will need to be made. Very wet seasons or extended droughts will affect the vegetation strongly. There is no guarantee that the current tapes are typical for Mexico, or that this year was normal in all respects. In all probability, at least two more tapes should be generated for 1976 to determine what changes, if any, may result from year to year.

## 5. REFERENCES

1. Phinney, D. E.: The Influence of Meteorological Conditions on the Growth and Development of Screwworm Populations. LEC-5987, 25 pages, April 1975.
2. Palmer, E. C.: Meteorological Drought, U.S. Weather Bureau Research Paper 45, 45 pages, 1965.
3. Huschke, R. E.: Glossary of Meteorology. American Meteorological Society, Boston, Mass., p. 200, 1959.
4. Phinney, D.; and Arp, G. K.: Technical Memorandum, The Measurement of Thermal Infrared Emissivities. Lockheed Electronics Co., Inc., LEC-5621, p. 45, 1975.
5. Phinney, D.; and Arp, G. K.: The Potential for Mesoscale Climatic Classification Through Remotely Sensed Data. Bulletin of the American Meteorological Society, Vol. 56 (1), p. 118, 1957.
6. Flores - Mata, G.; et al.: Tipos de Vegetación de la República Mexicana. Secretaría de Recursos Hídricos Map+, 1971.
7. Beard, J. S.: Climax Vegetation in Tropical America. Ecology 25(2), p. 127-158, 1944.
8. Beard, J. S.: The Classification of Tropical American Vegetation Types. Ecology, 36(1), p. 89-100, 1955.
9. Miranda, F.; and E. Hernandez X.: Los Tipos de Vegetación de México y su Clasificación. Boletín de la Sociedad Botánica de México, No. 28, p. 29-59, 1963.
10. Starker, A. S.: Vegetation Zones of Mexico. Ecology 31(4), p. 507-518, 1950.
11. Eyre, S. R.: Vegetation and Soils; A World Picture. Aldine Pub. Co., Chicago, p. 324, 1963.
12. Minter, R.: Earth Observation Division's Version of the Laboratory for Applications of Remote Sensing System Program Q019. Lockheed Electronics Co., Inc., LEC-3984 Rev 3, 1975.