Volume 5

Long-Term Climatic Variations in the Almati Oblast in Central Asian Kazakhstan: Correlations Between National Centers For Environmental Prediction (NCEP) Reanalysis II Results and Oblast Meteorological Station Data From 1949 to the Present

J.E. Welker, A.Y. Au

The NASA STI Program Office … in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and mission, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results . . . even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at http://www.sti.nasa.gov/STI-homepage.html

- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA Access Help Desk at (301) 621-0134

- Telephone the NASA Access Help Desk at (301) 621-0390

- Write to:
  NASA Access Help Desk
  NASA Center for AeroSpace Information
  7121 Standard Drive
  Hanover, MD 21076–1320
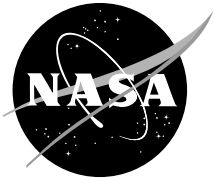
Volume 5

Long-Term Climatic Variations in the Almati Oblast in Central Asian Kazakhstan: Correlations Between National Centers For Environmental Prediction (NCEP) Reanalysis II Results and Oblast Meteorological Station Data From 1949 to the Present

J.E. Welker, Goddard Space Flight Center, Greenbelt, MD

A.Y. Au, Raytheon ITSS and Goddard Space Flight Center, Greenbelt, MD

LONG-TERM CLIMATIC VARIATIONS IN THE ALMATI OBLAST IN CENTRAL ASIAN KAZAKHSTAN: CORRELATIONS BETWEEN NATIONAL CENTERS FOR ENVIRONMENTAL PREDICTION (NCEP) REANALYSIS II RESULTS AND OBLAST METEOROLOGICAL STATION DATA FROM 1949 TO THE PRESENT

J. E. Welker

Terrestrial Information Systems Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland

A. Y. Au

Raytheon ITSS, and Space Geodesy Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland

Abstract

As part of a larger analysis of country systems described elsewhere, named a Crop Country Inventory, CCI, large variations in annual crop yield for selected climate sensitive agricultural regions or sub-regions within a country have been studied over extended periods in decades. These climate sensitive regions, principally responsible for large annual variations in an entire country's crop production, generally are characterized by distinctive patterns of atmospheric circulation and synoptic processes that result in large seasonal fluctuations in temperature, precipitation and soil moisture as well as other climate properties. The immediate region of interest is drought prone Kazakhstan in Central Asia, part of the Former Soviet Union, FSU. As a partial validation test in a dry southern region of Kazakhstan, the Almati Oblast was chosen. The Almati Oblast, a sub-region of Kazakhstan located in its southeast corner, is one of 14 oblasts within the Republic of Kazahstan. The climate data set used to characterize this region was taken from the results of the current maturely developed Global Climate Model, GCM. In this paper, the GCM results have been compared to the meteorological station data at the station locations, over various periods. If the empirical correlation of the data sets from both the GCM and station data is sufficiently significant, this would validate the use of the superior GCM profile mapping and integration for the climatic characterization of a sub-region. Precipitation values interpolated from NCEP Reanalysis II data, a global climate database spanning over five decades since 1949, have been statistically correlated with monthly-averaged station data from 1949 through 1993, and with daily station data from April through August, 1990 for the Almati Oblast in Kazakhstan. The resultant correlation is significant, which implies that the methodology may be extended to different regions globally for Crop Country Inventory studies.

Introduction

Global Climate Model, GCM, results over a fifty year or greater period for the climatic characterization of a sub-region by model profile mapping and integration over the sub-regional boundaries should be able to be empirically correlated with recorded meteorological station data in the sub-region in order to validate the use of the GCM climate characterization. The GCM's individual reading or time averaged data, e.g., averaged monthly values, evaluated at the latitude and longitude of the meteorological station location over the fifty year or greater period can be used as an empirical correlation with the recorded meteorological station data in the sub-region over the same period. If the empirical correlation of the data sets from both the GCM and station data is sufficiently significant, this could validate the use of the superior GCM profile mapping and integration for the climatic characterization of a sub-region. Then, the climatic variability of the sub-region and its possible cyclic nature can be tested using the GCM data alone. This was done for the Almati Oblast in Central Asian Kazakhstan.

The methodology described above has general global applicability. The use of the superior GCM profile mapping and integration over the digitized boundaries of a sub-region for the climatic characterization of the sub-region is important because the sub-regions produce reporting statistics, e.g., crop production, area, yield, etc. These statistics then could be correlated and modeled with climate parameters as long as the complicated agricultural system as well as the agro-climatic system of the sub-region is fully considered and realized. The significance of applying this analysis to the Almati Oblast in Kazakhstan is indicated below.

A global climate data set based on dynamical general circulation (GCM) models, the NCEP (National Centers for Environmental Prediction) Reanalysis data, spanning over 50 years will be used for climatic characterization of a sub-region. In our current study, interpolated values derived from this global data set should be able to be empirically correlated with recorded meteorological station data in a sub-region. If the resultant correlation between the global data set and the station data is sufficiently significant, the use of modeled data sets such as the NCEP Reanalysis II data for the climatic characterization of a sub-region can be justified. Then, the climatic variability of the sub-region and its possible cyclic nature can be studied based on modeled data alone.

Regional Selection

Kazakhstan, centrally located in the largest landmass on the Earth, is an extremely dry region, especially in its southern half, with a high frequency of periodic drought periods. Although Kazakhstan grain production is typically three-quarters wheat, and spring wheat at that, its grain output in 1992 was 15% of the total grain output of the total Former Soviet Union, FSU.[1] The significance of Kazakhstan was greatly increased suddenly in 1954, when Krushchev inaugurated the New or Virgin Lands Program that caused some 40 million acres of flat, treeless, fine soil areas in the very dry steppe area east of the Volga, mainly in Northern Kazakhstan and western Siberia, to be plowed and cultivated over the next four years. The principal crop grown was spring wheat, and although yields were low with the dry climate, the total Soviet wheat cultivated area increased by about 30% by 1958.[2] The correlation between cyclic variations in climate, especially dry periods and drought, and the resultant fluctuations in Russian/Soviet/CIS spring grain yields in Central Asia, especially from the Kazakhstan Republic, is well known from observation as well as from experimental agricultural research, and is part of a long and documented historical record as well.[3] This climatic characterization of the oblast of Almati is the beginning of a much larger study of Kazakhstan and its drought conditions,[4] but the methodology developed in this study will have general applicability to climatic characterization to all sub-regions globally.

The current choice of the Almati Oblast, in the southeast corner of Kazakhstan, is extremely dry region and not an important oblast region of Kazakhstan for the production of spring wheat. In a very good year, 1992, both the old Almati and Taldy-Kurgan Oblasti together produced only 0.6% of the total spring wheat output of Kazakhstan.[5] However, the NCEP data should be evaluated in both good and bad neighboring regions for climatic characterization comparison.

Almati is only one of the formerly 19 (until 1995), 17 (until 1999) and currently only 14, oblasti[6] or sub-regions of the Kazakhstan Republic. The Kazakhstan Republic was formerly part of the Soviet Union and is now one of the Russian Confederation of Independent States (CIS) that has replaced the Former Soviet Union in the 1990s. Under the current 14 oblasti configuration, shown in Figure 1, the Almati Oblast is a composite of both the old Almati Oblast plus the old Taldy-Kurgan Oblast, from the old 19 oblasti configuration. Almati City, with a population of over a million, is the capital of Almati Oblast, with a population over a million and a half. The area of Almati Oblast is listed as 224,000 square miles, that is located in the south-east corner of Kazakhstan.[7]

---

[1] International Agriculture and Trade Reports Former Soviet Union Situation and Outlook Series, Table 30, page 51, ERS-USDA WRS-94-1, May 1994.

[2] For a very brief description of the New or Virgin Lands Program, see Lydolph, Paul E., *Geography of the USSR*, 2nd Edition, pp. 414-415. Specific statistics of production, area and yield for that period can be found in *The Economics of the Soviet Wheat Industry*, p. 9 and *Strana Sovetov za 50 let*, pp. 128-129 and 132-133, also quoted in Lydolph, p. 416.

[3] Ibid.

[4] It is claimed that certain climatologists had pointed out that the New Lands were separated from the old grain lands of the Ukraine by a distance approximately equal to half a wave length of the standing waves of the upper atmosphere, that broadly control seasonal weather. Thus, the theory was that a bad grain crop in the Ukraine would be compensated by a good grain crop in Kazakhstan in the same year and vice-versa. Lydolph, p. 415.

[5] Kazakhstan oblasti data from a private communication with the Economic Research Service(ERS) of the United States Department of Agriculture (USDA)

[6] Oblasti in Republics of the Former Soviet Union are somewhat analogous to counties within States in the United States. Almati Oblast is now a composite of the former oblasts of Taldy-Kurgan and Almati.

[7] ???

[8] Retrieved from the Web site of the NCEP/NCAR CDAS/Reanalysis Project of the National Oceanic and Atmospheric Administration (NOAA).

However, there are a number of meteorological stations in the current Almati Oblast that can be tested against the GCM data at the same latitude and longitude locations as the stations, in order to validate the use of the GCM's climate characterization of oblasti in Kazakhstan.

GCM and Station Data Correlations

NCEP Reanalysis II[8] data spanning over five decades since 1949 are used to test correlation between observed precipitation data at meteorological stations at the old Almati and Taldy-Kurgan Oblasti. Two independent sets of observed data are used in our comparison. One is the Former Soviet Union Monthly Precipitation Archive, which consists of monthly precipitation measurements from 622 stations located within the former Soviet Union for the years 1891 to 1993. There are, unfortunately, data gaps for some stations in this archive. The other data set is daily precipitation taken from the Part 1 of the USSR Meteorological Monthly Reports from April through August for 1990.

Eight stations have been selected from the Monthly Archive (MA), whereas four stations have been taken from the Monthly Reports (MR) (See Table 1). The geographical locations of these stations are shown in Figures 1a and 1b.

Table 1. Pearson Linear Correlation Coefficients and their Significance of Coefficient for stations selected from Former Soviet Union Monthly Precipitation Archive (MA) and Part 1 of the USSR Meteorological Monthly Reports (MR).

Station #	Station Name	Corr. Coeff.	Sig. of Coeff.	Source
337	Kuygan	0.36	8.4	MA
338	Alma-Ata	0.58	16.5	MA
339	Ili, Rail. Stn.	0.44	10.6	MA
340	Podgornoye	0.53	13.6	MA
341	Taldy-Kurgan	0.56	15.6	MA
342	Panphilov, AS	0.43	11.1	MA
343	Uch-Aral	0.38	9.4	MA
344	Ush-Tobe	0.44	9.3	MA
172	Balhash	0.02	0.2	MR
180	Ucharal	0.34	4.4	MR
190	Panphilov	0.24	3.0	MR
196	Alma-Ata	0.29	3.7	MR
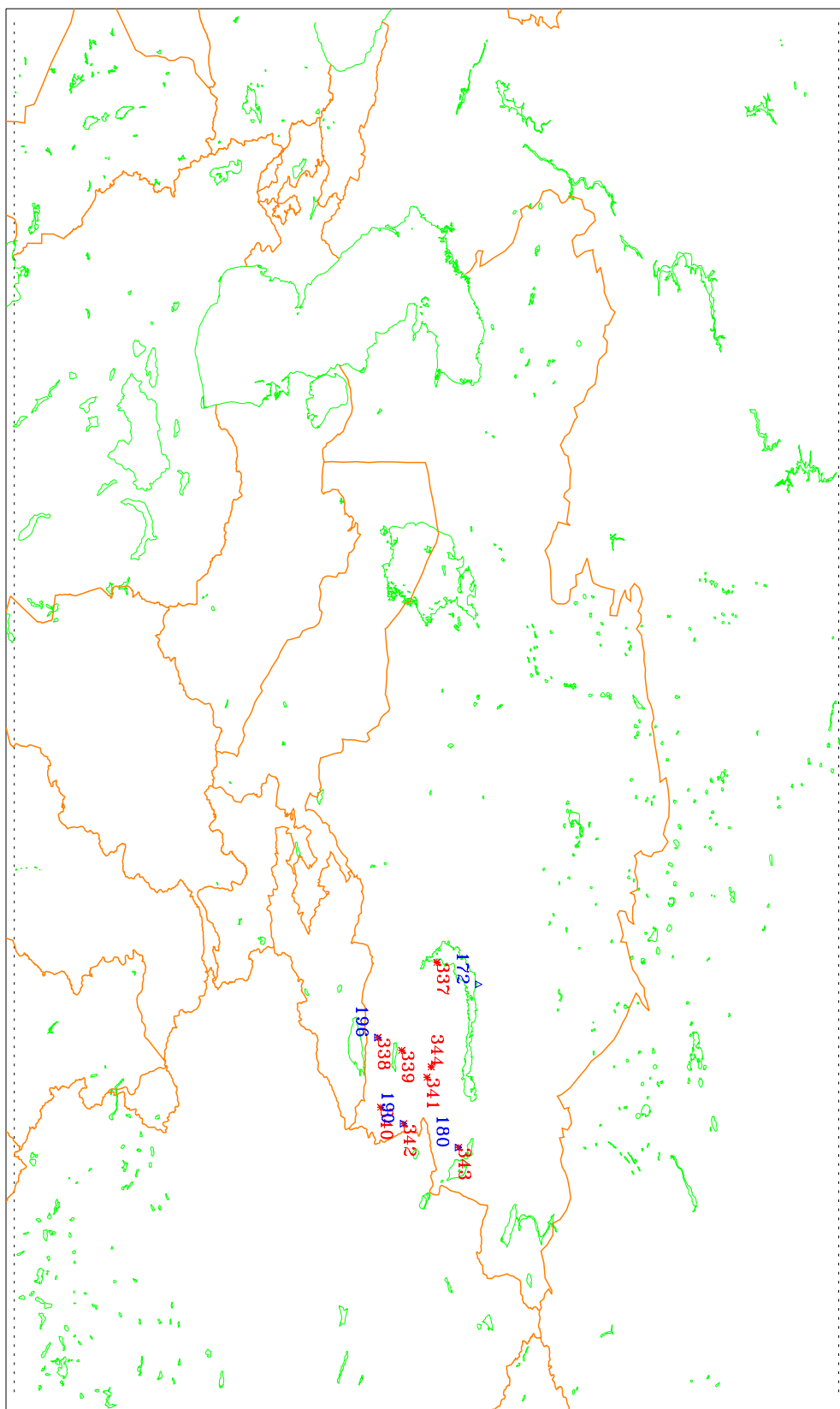
Figure 1a. Locations of the stations in this current study in Kazakhstan. Stations from the Former Soviet Union Monthly Precipitation Archive (MA) are in red, whereas stations from Part 1 of the USSR Meteorological Monthly Reports (MR) are in dark blue.
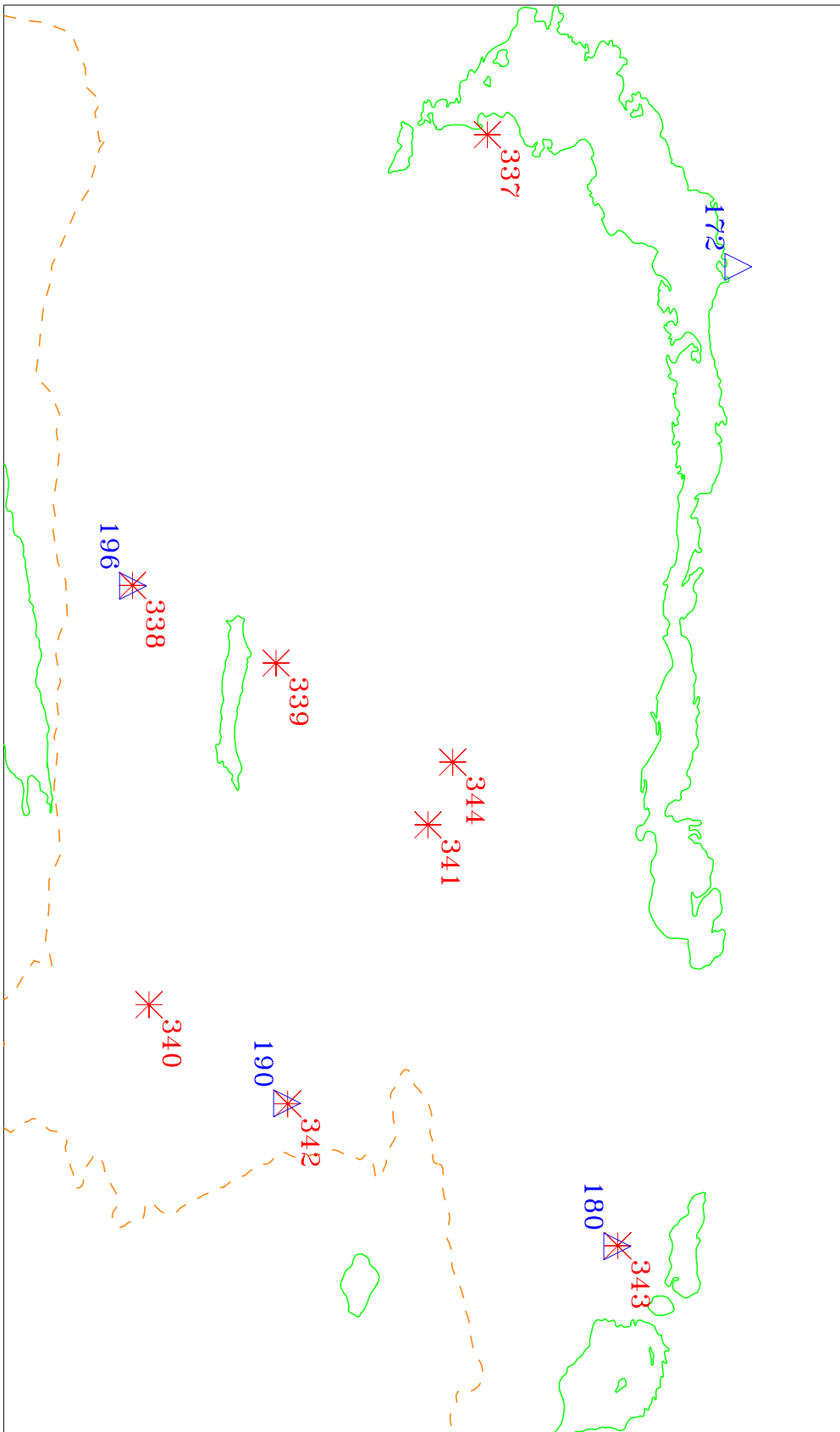
Figure 1b. Enlargement of the study area (Almati Oblast).

The NCEP Reanalysis data sets are arranged in regular data grid 2.5 degrees in longitude and about 2 degrees in latitude for both daily-averaged and monthly-averaged data. Precipitation values are interpolated via the minimum curvature surface method for the geographical locations of the selected stations in our current study. A comparison of NCEP data with the Archive monthly data from 1949 through 1993 for the eight selected stations are shown in Figure 2. Comparison of NCEP data with the daily data from the Monthly Reports for four stations are shown in Figure 3. The Pearson linear correlation coefficients and the significance of the coefficient between the NCEP Reanalysis data and the observed monthly and daily data are shown in Table 1. Pearson linear correlation is a popular test when one wishes to see if there is a relationship between two variables, which are measured at the interval or ratio levels. For the sampling size of our current data set (on the order of a few hundred data points), any correlation with coefficients above 0.2 (the critical value) is considered to be statistically significant. It is not necessary to have linear correlation coefficient with values greater than 0.9 before the correlation between two variables are considered statistically significant.

Close examination of the match between observed and NCEP reanalysis data for the eight selected stations from the Former Soviet Union Monthly Precipitation Archive from 1949 through 1993 in Figure 2 suggests that although the correlations are fair, the NCEP results are generally a factor of two to three too high in absolute amount of precipitation for the stations, especially for stations 339, 341, 342, and 344. Meanwhile, the absolute amount of observed precipitation among the eight stations are quite uniform. Given that the region is a semi-desert, the NCEP results appear anomalous in this case.

Correlations of observed daily precipitation with NCEP results shown in Figure 3 are considerable poorer that of the monthly data. It should not, in fact, be surprising because of the short data span (only from April through August, 1990), and the inherent noisiness of the daily data. Nonetheless, only station 172 fails the 0.2 criterion of a statistical significant correlation.

In light of the low Pearson linear correlation coefficient between the observed data from the Meteorological Monthly Reports and the NCEP Reanalysis data, a further statistical test has been devised to confirm the statistical significance of the comparison, except for station 172. We would like to test of a null hypothesis that the observed precipitation, as a data set, is no better than any set of randomly selected values from an ensemble of possible measurements. In another word, we hypothesize that the observed data sets for these stations are as good as any random numbers!

In this test, we have created an ensemble of all the observed daily precipitation measurements from April through August, 1990 taken from about 90 stations in Kazakhstan and neighboring area. Then, a hypothetical "measured" data set is created by random selection of 153 observed daily precipitation from the ensemble. The Pearson linear correlation coefficient between this hypothetical data set and the real data set is determined. This process is repeated 1000 times, for the sake of statistical significance, for every selected station. The results for all four stations are shown in Figure 4 and tabulated in Table 2.

It is observed that the mean correlation coefficients of the 1000 randomly selected data sets are virtually zero for all these four stations, with a half-height width of less than 0.1. The actual correlation coefficients of the observed data and NCEP Reanalysis data lie well beyond the approximately 0.1 half-height width, except for station 172. The shape of the simulation population centered about zero in the form of a log-normal distribution. The deviation from an ideal Gaussian shape is expected because of the bias of the ensemble C there is no negative value of precipitation.

We may, therefore, state confidently that the correlation between the observed daily precipitation of the Meteorological Monthly Reports and the NCEP Reanalysis data are real, not a fluke of statistical fluctuation, in spite of the relatively low value of their linear correlation coefficients.
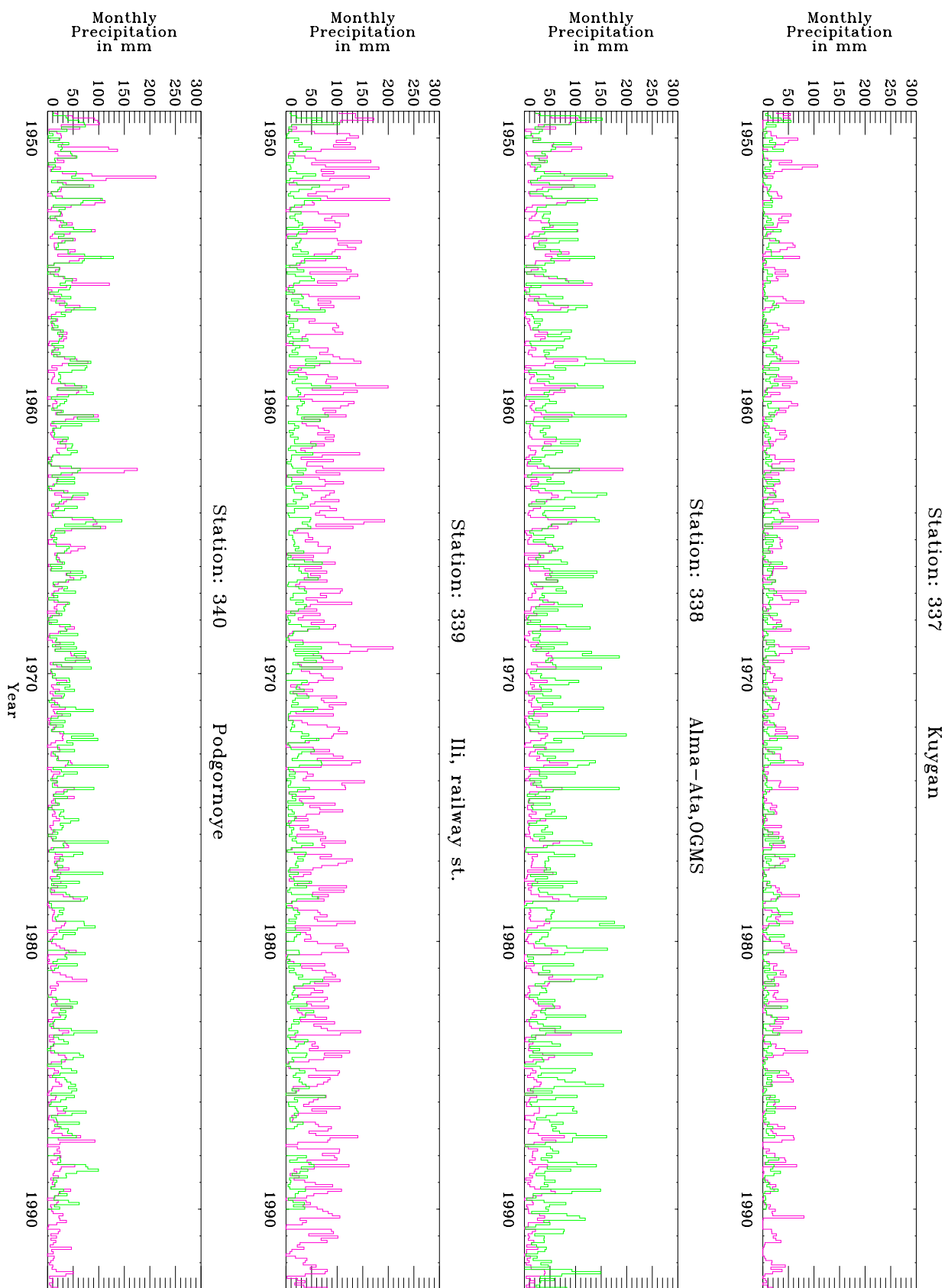
Figure 2. Comparison of NCEP data (pink) with the Archive monthly data (green) from 1949 through 1993 for eight selected stations.
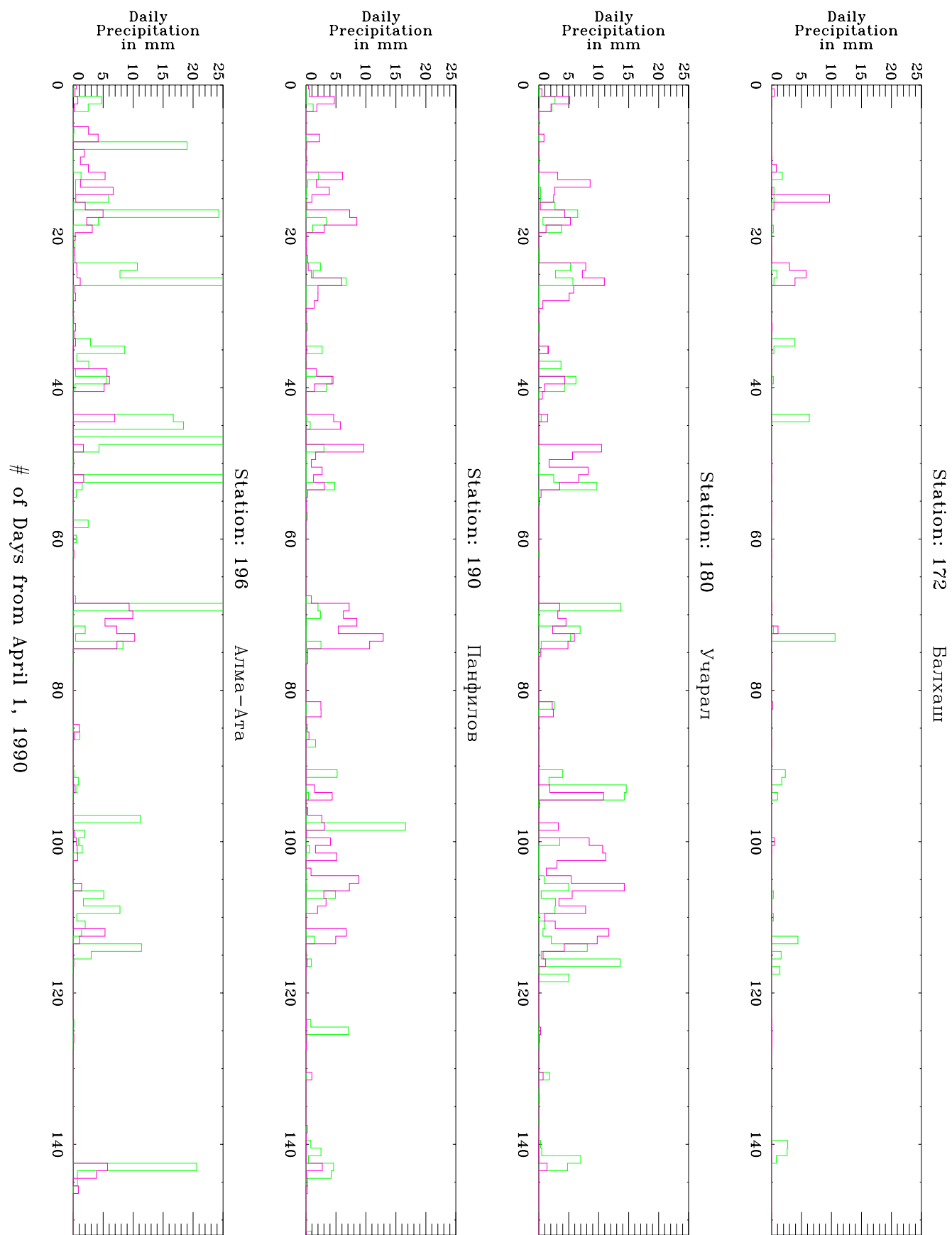
Figure 3. Comparison of NCEP data (pink) with the daily data from the Monthly Reports (green) for four selected stations.
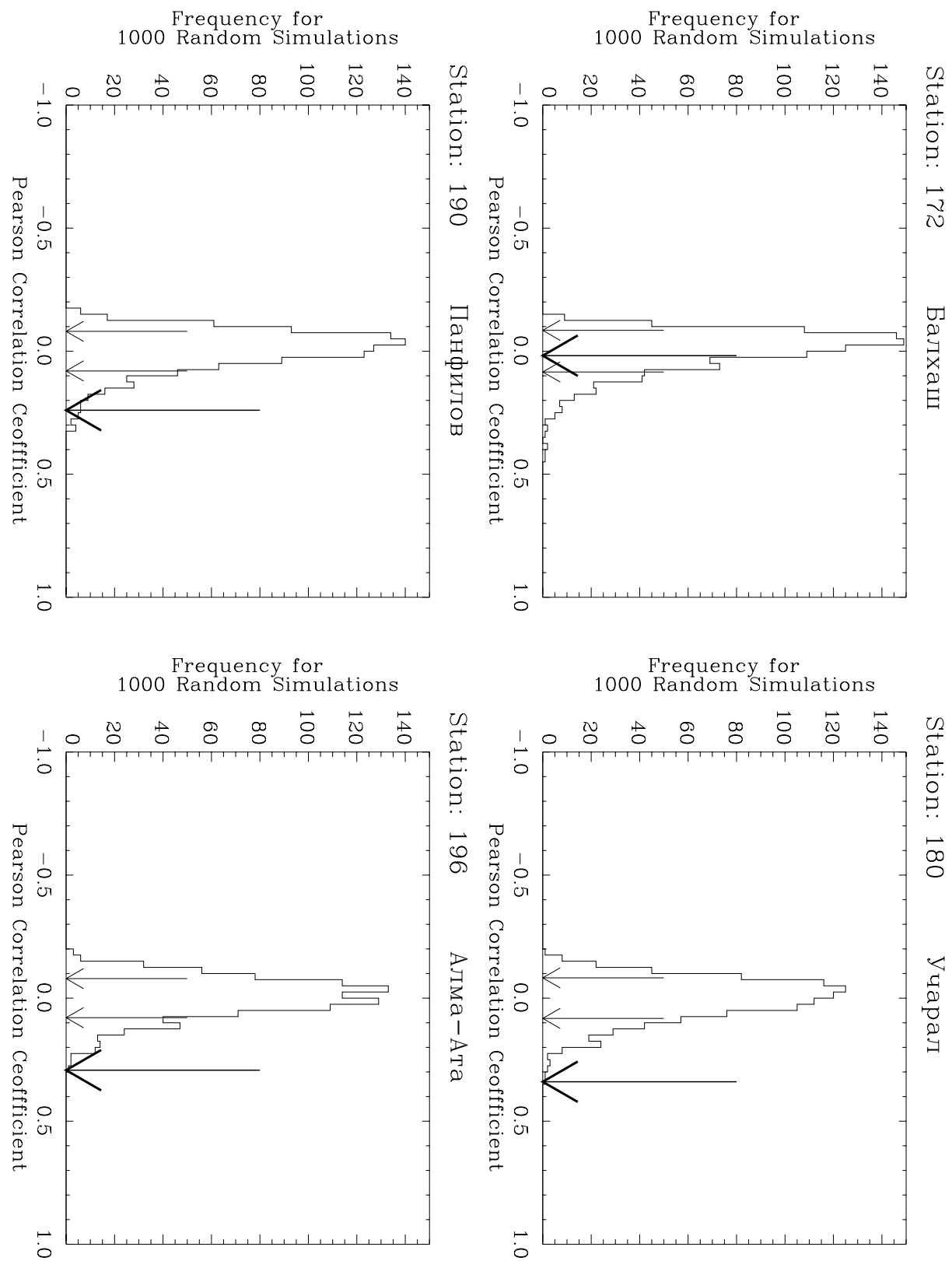
Figure 4. Test of nonrandomness of the observed data from the Monthly Reports for four selected stations. The pair of thin arrows defined the half-height width of the simulated distribution, whereas the large thick arrows represent the actual correlation coefficients for the observed data.

Table 2. Bootstrap Test of Statistical Significance of Correlation Coefficients

Station #	Station Name	Avg. Hypo. Corr.Coeff.	Half-height width of Hypo. Corr. Coeff.	Actual Corr. Coeff.
172	Balhash	0.0012	0.085	0.02
180	Ucharal	0.0052	0.082	0.34
190	Panphilov	-0.0041	0.080	0.24
196	Alma-Ata	-0.0019	0.079	0.29

Discussion

In this preliminary study, we have limited our sampling to only a few stations in the Almati Oblast. It should be appreciated that station data are inherently noisy, and they may be biased. The correlation of the station data with modeled NCEP data, unfortunately, does not come out as high as one may wish it to be. Nonetheless, it may be argued that it is hard to construct a coherent and reliable picture of precipitation pattern over a period of time just based on scattered station data. On the other hand, since NCEP results are based on dynamical modeling, they represent good averages over large area where there are no reliable data. Of course, fair correlation with strong statistical significance does not necessarily mean that the NCEP model is a good model for the region. Other independent studies may be needed to ensure that we may use NCEP results in lieu of station observations.

We would like to conclude that the eventual product of such correlation studies is to validate the use of dynamically modeled global data sets for the climatic characterization of any or all sub-regions globally. With its initial concentration on Kazakhstan in Central Asia, we will be able to acquire a better understanding of the cyclic drought pattern everywhere but especially in Central Asia. As a second product, an improved capability to predict the Central Asian grain crop yield as the crop is grown is envisioned.

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE December 2003	3. REPORT TYPE AND DATES COVERED Technical Memorandum

4. TITLE AND SUBTITLE Long-Term Climatic Variations in the Alamti Oblast in Central Asian Kazakhstan: Correlations Between National Centers For Environmental Prediction (NCEP) Reanalysis II Results and Oblast Meteorological Station Data From 1949 to the Present

5. FUNDING NUMBERS

Code 922

6. AUTHOR(S)

J.E. Welker

A.Y. Au

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS (ES)

Goddard Space Flight Center
Greenbelt, MD 2077

8. PEFORMING ORGANIZATION REPORT NUMBER

2004-00660-0

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS (ES)

National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

TM-2003-212252
Vol. 5

11. SUPPLEMENTARY NOTES

A.Y. Au, Raytheon ITSS and Goddard Space Flight Center, Greenbelt, MD

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Unclassified–Unlimited
Subject Category: 47
Report available from the NASA Center for AeroSpace Information,
7121 Standard Drive, Hanover, MD 21076-1320. (301) 621-0390.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)* As part of a larger analysis of country systems described elsewhere, named a Crop Country Inventory, CCI, large variations in annual crop yield for selected climate sensitive agricultural regions or sub-regions within a country have been studied over extended periods in decades. These climate sensitive regions, principally responsible for large annual variations in an entire country's crop production, generally are characterized by distinctive patterns of atmospheric circulation and synoptic processes that result in large seasonal fluctuations in temperature, precipitation and soil moisture as well as other climate properties. The immediate region of interest is drought prone Kazakhstan in Central Asia, part of the Former Soviet Union, FSU. As a partial validation test in a dry southern region of Kazakhstan, the Almati Oblast was chosen. The Almati Oblast, a sub-region of Kazakhstan located in its southeast corner, is one of 14 oblasts within the Republic of Kazakhstan. The climate data set used to characterize this region was taken from the results of the current maturely developed Global Climate Model, GCM. In this paper, the GCM results have been compared to the meteorological station data at the station locations, over various periods. If the empirical correlation of the data sets from both the GCM and station data is sufficiently significant, this would validate the use of the superior GCM profile mapping and integration for the climatic characterization of a sub-region. Precipitation values interpolated from NCEP Reanalysis II data, a global climate database spanning over 5 decades since 1949, have been statistically correlated with monthly-averaged station data from 1949 through 1993, and with daily station data from April through August, 1990 for the Almati Oblast in Kazakhstan. The resultant correlation is significant, which implies that the methodology may be extended to different regions globally for Crop Country Inventory studies.

14. SUBJECT TERMS
Global Climate Model, GCM; Crop Country Inventory, CCI; Almati Oblast; Kazakhstan

15. NUMBER OF PAGES
10

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL