FY2007 NGDC Climate Database Modernization Program Proposal
(New Task)

A. Organization and Task Lead:
NESDIS / National Geophysical Data Center
Helen Coffey

B. Name or title of task:
Historical Cosmic Ray Ionization Chamber Data - Digitization of historical cosmic ray observations (photographic records – linear traces) to digital 2-minute values format.

C. How It Will Contribute to the Usefulness of the Data:
Dr. Scott Forbush was a pioneer in studying cosmic radiation which was discovered in 1911. He maintained a worldwide network of ionization chamber monitors at Cheltenham, Maryland; Christchurch, New Zealand; Fredericksburg, Virginia; Godhavn, Greenland; Climax, Colorado; Huancayo, Peru; and Mexico City, Mexico for the time period 1936-1959. An ionization chamber measures the total ionization produced by secondary cosmic rays capable of passing through about 10 cm of lead. Until recently, these "legacy data" archives were kept at the Carnegie Institute, Washington, DC. Dr. Ken McCracken, cosmic ray physicist, spent time during the last 2 years culling out the important data in the Forbush archives and inventorying the contents. Carnegie shipped these to NGDC in July 2006. These original photographic records contain important information never made public and which are highly relevant to modern day scientific studies. They have considerable significance for space weather studies. McCracken recommends that the photographic records for all 5 ionization chambers be recognized as being of high scientific value to be preserved until an electronic form is developed and proven, with rigorous quality control.

The Forbush legacy data are the only source of continuous cosmic ray data for the period of increasing solar activity prior to the 1950s and it is certain that those data will be a vital input to studies of the manner in which cosmic ray phenomena respond to changes in the level of solar activity. The Forbush legacy data are the only source of higher time resolution data for the four Ground Level Enhancements (GLEs – cosmic rays produced by the Sun) that occurred prior to 1956. Obtaining higher resolution data for these four events could provide a very significant increase in our knowledge of the cosmic ray events and of the interplanetary magnetic fields. These GLEs can damage our terrestrial and spaceborne infrastructure and constitute a hazard for manned space travel. There is major value remaining in these primary records that has not been utilized to date.

The data consist of nine boxes of photographic records of daily strips annotated with date and time, joined to make a continuous strip for each month, folded at each midnight, and made into a concertina like stack with a full day visible at a time. The contrast on most strips is excellent and should permit machine based reading. Dr. McCracken is preparing a brief "handbook" that will allow others to understand the nature of the photographic record and the pitfalls therein.

D. Describe the task in detail:
Nine boxes of data exist containing the following:

- Huancayo photographic records 1936-1948
- Hunacayo photographic records 1949-1960
- Cheltenham/Fredericksburg photographic records 1935-1946
- Cheltenham/Fredericksburg photographic records 1947-1960
- Christchurch photographic records 1936-1950
- Christchurch records 1951-1959 and Mexico records 1954-1959

- Godhavn photographic records 1938-1950
- Godhavn photographic records 1951-1959
- Climax photographic records 1949-1959

In each box are annual bundles of strip charts, broken down with monthly bundles of one big long strip chart folded at midnight each day. Each day is about 24 inches long by 2 inches wide, with sprocket holes on the edges. 1 hour is about 1 inch. Vertical and horizontal line scaling are included on the graphs. A sample image of the data is attached. The ionization chamber adds up the ionization on a capacitor, i.e., it gives a time integral of the cosmic ray flux. Every hour the capacitor is shorted out- that is the step you see every hour. Occasionally there are fast downwards steps- they are "bursts"- a sudden pulse of ionization due to a nuclear reaction in the wall of the chamber. They have to be subtracted from the total hourly deviation to give the ionization for that hour.

These charts need to be digitized in 2-minute intervals. This implies a data record of 500,000 bytes per station per year (station-year). There are 114 station years total, thus 70 Mbytes of data total. For evaluation, Dr. McCracken suggests a sample time period of a single month of Cheltenham data after 1957 be used for test purposes since it has good contrast. The time calibration of the record will require careful consideration.

Data output should be returned to NGDC as daily digital files of x-y format (time vs total ionization flux) written to DVDs. Each file should contain one day of data from a particular station. Filenames should be in the format YYYYMMDDXX for year (YYYY), month (MM), day of month (DD) and station ID (XX), e.g., 19510116CC for a Christchurch (CC) photographic digitized file on Jan 16, 1951.

In addition, the bihourly data compilations published by Forbush and his coworkers are an additonal major component of this legacy data. This 4-volume set should be digitized (key entered) and made available via the web. There are about 800 pages of data. A simple data entry format will be developed.

E. Estimate of time needed to prepare the data for the CDMP contractors:
The nine boxes of the early Forbush ionization chamber cosmic ray charts currently reside in the NGDC Warehouse. These original records will be shipped to the CDMP contractor by FedEx. Quality control (QC) includes plotting the data, comparing it with the originals, making corrections, verifying times, etc. A computer program needs to be developed to convert the data to useable values and to do basic QC of the data. For the conversion of the two minute data, we will measure the position of the trace every two minutes; subtract successive measurements and subtract any bursts therein. A web page needs to be developed. At least one month is needed to complete these tasks. Dr. McCracken has indicated he is willing to assist in setting up the digitizing process.

Funding for NGDC Salaries for this project for this contract period will require $35K. This includes preparation of the data sets for submission to the contractors, quality assurance efforts, translating the digitized data into useable values, and making the data accessible via the web.

F. Data Availability: The digitized data will be available online at the NGDC ftp site. NGDC currently has some early cosmic ray ionization chamber data available at ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/COSMIC_RAYS/Annals_of_the_IGY_Data. Information about the

Forbush data can be found at ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/COSMIC_RAYS/ForbushArchives1935-1960.