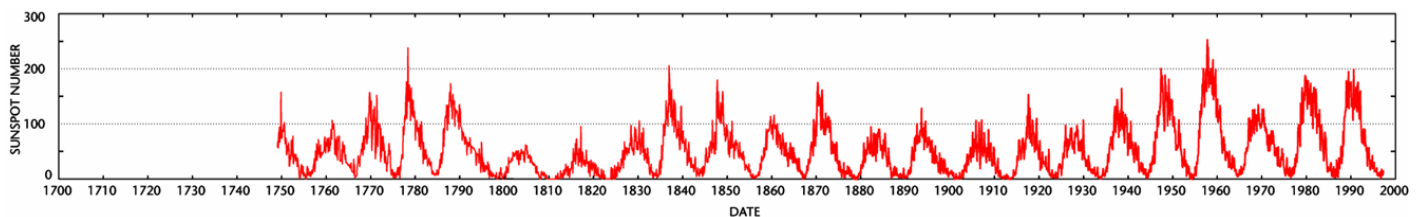


Abstract for *SORCE Decadal Variability in the Sun and Climate*
Meredith, New Hampshire, October 27-29, 2004

NOAA NESDIS Data Rescue Solar Image Scanning Project

Helen E. Coffey
NOAA National Geophysical Data Center
National Environmental Satellite, Data, and Information Service
325 Broadway, Boulder, CO 80305
Helen.E.Coffey@noaa.gov

The National Oceanic and Atmospheric Administration (NOAA) Climate Database Modernization Program (CDMP) funds some data rescue efforts within the NOAA line organizations to make climate and environmental data more accessible and easier to use. Data restricted to file cabinets and Warehouse storage are becoming accessible via the World Wide Web. NOAA's National Geophysical Data Center (NGDC) holds solar imagery back to the 1920s, available on 35 mm microfilm, as well as printed photos and drawings. A CDMP project was funded to scan many solar images to digital format and make them available on the NGDC Website. The scanned daily images document many phases of solar activity, from decadal variation to rotation variation to daily changes. The images include Calcium K, Hydrogen Alpha, and white light photos, as well as sunspot drawings and the comprehensive drawings of a multitude of solar phenomena on one daily map (Fraunhofer maps and Wendelstein drawings). An overview of the project and the current status will be given.



Introduction

For a number of years, NOAA NESDIS has had a Climate Database Modernization Program (CDMP) that supports all five NOAA line organizations with services intended to make climate environmental data and information more accessible and easier to utilize. Proposals are submitted and awards depend on available funding. Tasks generally involved digitizing historical climate records. It can involve creation of a digital image from paper, microfilm or microfiche, or keying data from these records. It may also involve creation of a web system for storing and accessing data or creation of a new database to support on-line data access. Current funded tasks include keying lightship data from Finland and Sweden, surface observations from Uruguay, NOAA Central Library Rare Books, scanning historical coast pilot editions, sea surface water temperature and density data, Great Lakes Ice Thickness, keying SeaCat/Bongo stations, scanning DMSP film, keying ionospheric data, Moravian data in Germany, scanning glacier photos, keying radiosonde data, etc.

Several years ago, NGDC obtained **NOAA data rescue funds** to key in "ancient" solar data – including

- Sunspot Region data (<http://www.ngdc.noaa.gov/stp/SOLAR/ftpsunspotregions.html>) from US Naval Observatory data 1927-1971, Boulder 1966-1981, USAF 1975-1981, and USSR Solnechnye Dannyue (Solar Data) region data;

- White Light faculae data from Greenwich 1874-1955 and summary data 1956-1976 (<http://www.ngdc.noaa.gov/stp/SOLAR/ftpwhitelightfaculae.html>); and
- Airglow -- Annals of the IGY Vol XXIV by I.G. YAO (pseudonym for International Geophysical Year Airglow Observers) (<http://www.ngdc.noaa.gov/stp/SOLAR/ftpairglow.html>).

Another **joint data rescue program with NASA** funded the conversion of analog to digital form of a number of “ancient” solar databases – including

- IAU QBSA solar radio daily flux data from 35 stations 1949-1986 (<http://www.ngdc.noaa.gov/stp/SOLAR/ftpsolarradio.html>);
- HAO IGY solar filaments 1957-1959;
- Carte Synoptiques limb (1919-1957), filaments (1957-1989) and centres of activity (1957-1989) (<http://www.ngdc.noaa.gov/stp/SOLAR/ftpfilaments.html>);
- McMath Calcium plage data (1942-1960) (<http://www.ngdc.noaa.gov/stp/SOLAR/ftpcalcium.html>), cosmic ray hourly data 1957-1979 (<http://spidr.ngdc.noaa.gov>); and
- auroral observations 1000-1900 (<http://www.ngdc.noaa.gov/stp/SOLAR/ftpaurorae.html>).

Current Project

This past year, NGDC submitted a proposal to scan in daily solar images from past years. This was to update the data product of daily solar imagery scanned to microfilm in the 1970s -- HISTORICAL SOLAR OBSERVATIONS ON 35MM MICROFILM – <http://www.ngdc.noaa.gov/stp/SOLAR/microfilm.html>. These databases include

- **McMath-Hulbert Calcium Plage Regions 1942-1979** -- *Full disk daily drawings. Calcium drawings from Mt Wilson and Big Bear observatories complete the set after McMath Observatory closed in September 1979,*
- **Sunspot Daily Drawings** -- *Full disk drawings of sunspot areas for April 1948-December 1992. Drawings from Boulder, Sacramento Peak and Holloman observatories were used to make the set. There are no observations for July-December 1964.*
- **Fraunhofer Daily Maps of the Sun** - *Detailed drawings of sunspot groups; of calcium plages and flare positions; and of corona, prominences and filaments for the years 1956-1973. Data reduction program discontinued in 1973. and*
- **H-Alpha Solar Patrol Films** - *Daily solar patrol films from nine stations.*

An in-house scanning program is on-going and entails scanning the Wendelstein daily drawings. Fraunhofer daily maps will also be scanned at NGDC. The scanned images are to reside on the NGDC ftp site or be accessible via web interface.

Logistics

The logistics of handling this large project are complex, from locating the original daily images to supplying these to the contractor LASON in Maryland, to determining the best format for the scan (tif with LTZ compression, 600 to 800 dpi for photos, 300 to 400 dpi for drawings), to deciding what

the orientation of the images are (North at top and East on left), to quality controlling the scanned output, to developing a web interface for data access. The Discovery Phase included scanning on several different scanners at a number of different resolutions to obtain the optimum result. Boxes of solar patrol scout photos were shipped to LASON for scanning, as well as IGY composite drawings and McMath drawings. A web interface for the images is being developed using Adobe Photoshop Web Photo Gallery.

In addition, sample scans of film strips taken at 600, 900, 1200 and 1800 dpi have been done – see pixelation images. A choice has to be made about the resolution. Higher resolution costs a lot more and takes a lot longer to scan. The tradeoff is better quality digital images versus a smaller number of images scanned. We are asking the community to help with this decision, in the expectation that this scanning project will be funded in the next year. It is noted that Debrecen Heliophysical Observatory continues to use film recording for their white light imagery until the digital imagery resolution reaches 8000x8000 dpi when both film and analog will be similar resolution.

Staff available to handle the scanning project is minimal – a senior staff member has been scouring the local archives and the Warehouse archives for the images. In addition, the difficulty in providing microfilm solar patrol films is proving problematic and will not be done this year. We are hopeful that the scanning project will continue next year and these issues with films will be resolved. At stake is cutting daily strips out of the patrol film to be able to scan the images. This means cutting and splicing the patrol film which would probably destroy at least two solar images per cut.

Current Status

The contractor LASON has sent 25 DVDs with 8054 images of McMath Calcium K and H-alpha scanned from scout photos – see the two Web Photo Gallery pages for samples at <http://www.ngdc.noaa.gov/stp/SOLAR/WebPhotoGallery/McMath1960CPMCsep/index.htm> and <http://www.ngdc.noaa.gov/stp/SOLAR/WebPhotoGallery/McMath1960HPMCsep/index.htm> -- see Figure 1 below. Each image is scanned to tif format with LZW compression at 600 dpi or 800 dpi resolution. In addition there are 2 DVDs with 410 images of US Naval Observatory White Light scans. The current scanning project runs through March 2005. We expect LASON to scan in the daily Calcium drawings from McMath and the daily sunspot drawings from Boulder and other observatories.

Years covered by the 25 DVDs of McMath scouts include 1948, 1952-1970 and 1974-1975. Some years are incomplete or haven't been scanned yet. An effort is being made to locate missing years and months. Generally there is one Calcium K and one H-alpha image per day when observing was possible.

In-house, NGDC has scanned many of the Wendelstein daily drawings. These cover the years 1950-1987. 1968-70, 1973, 1980-1981 have been quality controlled. Stations providing data for these comprehensive daily maps include: H-alpha observations from Anacapri, Athens, Burbank, Catania, Freiburg, Haleakala, Kodaikanal, Sac Peak, Teheran, Tonantzintla and Wendelstein; Calcium K3 observations from Anacapri, Arcetri, Catania, Kodaikanal, Manila, Rome and Wendelstein; and Corona 5303A observations from Mt. Norikura, Pic du Midi, Sac Peak and Wendelstein – from Apr-Jun 1969 inventory listing.

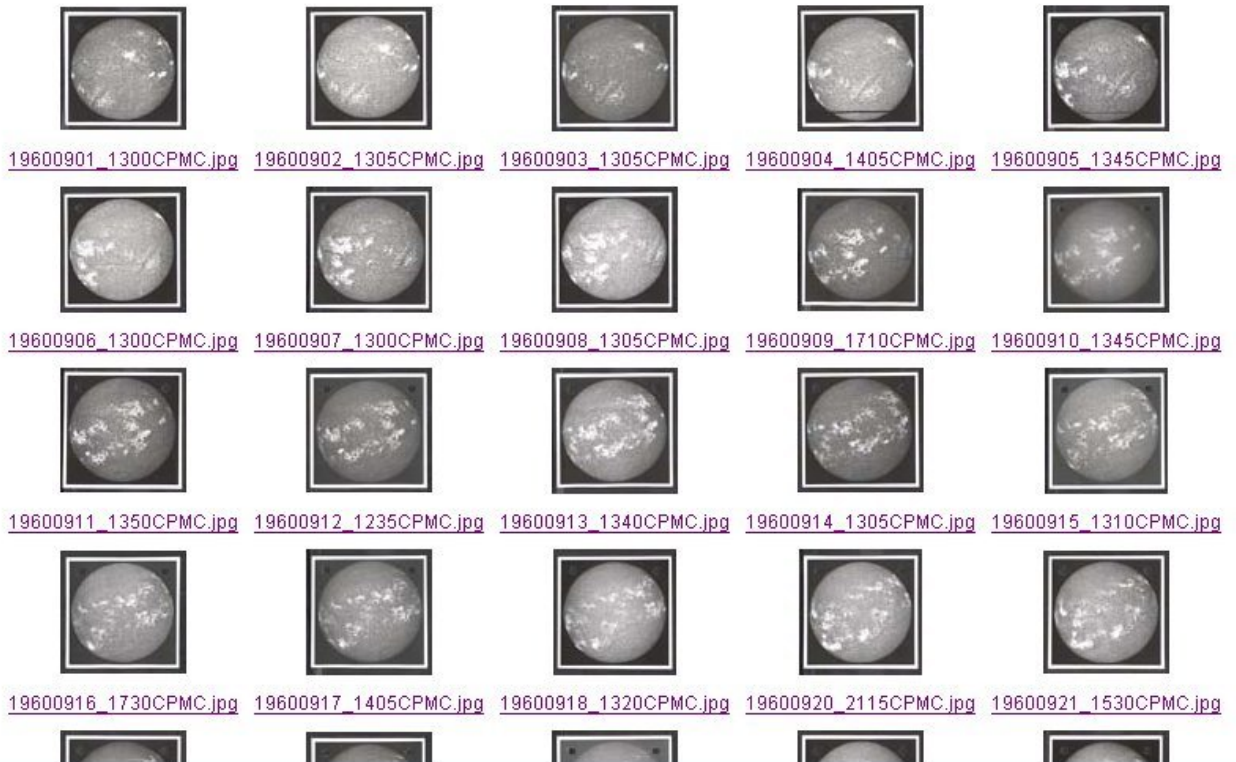


Fig. 1 <http://www.ngdc.noaa.gov/stp/SOLAR/WebPhotoGallery/McMath1960HPMCSep/index.htm>

As an interesting aside, we found that the early McMath Calcium drawings in 1942 have listed sunspot areas, not calcium plage areas. These sunspot areas are listed in the ASCII McMath Calcium plage database as calcium plage areas. Changes need to be made to this online database and errata issued at <http://www.ngdc.noaa.gov/stp/SOLAR/ftpcalcium.html#mcmath>. Records found with the old data indicate that “areas of prominences are given in prominence units. One prominence unit is very nearly equal to three millionths of the sun’s visible hemisphere.” More investigation is needed to determine whether these might be the spurious values that appear in the 1947-1948 data as noted by P. Foukal (GRL, Vol. 23, No. 16, pages 2169-2172, 1996) or perhaps the spurious data are an activity index A_{xI} (area x intensity).

At LASON, waiting to be scanned, are the International Geophysical Year (IGY) oversized drawings that were published in Annals of the International Geophysical Year Volume XXI and XXII. This significant contribution from the IGY era is notable, especially with the 50th anniversary of the IGY occurring in 2007. LASON also has some Beograd (Belgrade) white light images from the IGY period that are quite good and show the enormous sunspot regions developing and decaying during this outstanding solar cycle that peaked in 1957. Another selection of good observations found in the archives is some Mt. Wilson Solar White Light Photos 1946-1947 – these were sent to LASON for scanning.

The question arises about the importance of scanning films versus photos or drawings. As many contacts have emphasized, for research efforts, the scanned film images are best. Because of the problems involved with getting the films to the contractor, NGDC has first tackled the easier option of providing daily photos and drawings. Much observer time was put into the daily drawings. Wendelstein and Fraunhofer Observatories collected many different kinds of observations and produced daily hand-drawn solar maps showing all of the different solar phenomena on one map. Wendelstein maps include H-alpha filaments, Calcium K plages and coronal observations. The IGY Volumes XXI and XXII are similar hand-drawn daily maps. These comprehensive and long time interval maps of solar activity can be studied for the changing behaviors of the different solar phenomena. They are a record of what the Sun looked like over many years.

Scope of work proposed for FY05

Continue scanning, putting emphasis on scanning 35 mm film imagery which will give the best results for scientific research. An inventory of images to be scanned follows:

- 67 years of daily H-alpha images from 35 mm high quality film -- 24,000 images
- White light images from USNO original plates – 12,000 images
- Calcium K images from 35 mm high quality film – 10,000 images
- Missing years of scout photos for McMath and USNO will be located for scanning – several thousand images.
- Some early sunspot drawings, even back to 1860, have been located in the NGDC archives – several thousand images.

Work on developing a web interface using SPIDR (<http://spidr.ngdc.noaa.gov>) or WIST (<http://wist.ngdc.noaa.gov>) needs to be done.

Find out what the accepted best solar image processing procedures are to date and implement those when possible to best prepare a final product for the user community. This can include correcting the image orientations so that North is at the top and East is on the left. Basic image corrections should also be included. The original scans will be kept as a level zero archive and the corrected images made available via the web.

We currently have a disk space problem on our ftp site due to the large sizes of solar images. This ftp site is expected to go away sometime in the near future due to security issues. We are working on remote storage of the original images, in Tivoli or ADICS library system, with something like a Photoshop Web Photo Gallery web interface that will connect to a downloadable zipped file for the original image.

Conclusions

Many solar images from years past are available only in analog form, as paper drawings, photos, on 35 mm film, etc. Several institutions worldwide are scanning in old archives of imagery, including Mt. Wilson and Greenwich. NGDC is making an attempt to transfer archived analog solar images to digital versions. Using image processing tools, researchers will be able to do more detailed studies with the daily digital images. However, until 8000x8000 pixel digital imagery are available, photographs and films are still superior to the digital imagery.