
AI translation · View original & related papers at
chinaxiv.org/items/chinaxiv-201711.01269

Yunnan Observatories Hand-Drawn Sunspot Data System Postprint

Authors: Zhu Gaofei, Zheng Sheng, Lin Ganghua, Zeng Xiangyun, Feng Yongli, Tao Jinping, Shu Yao

Date: 2017-09-26T00:00:00+00:00

Abstract

The Yunnan Observatories, Chinese Academy of Sciences has accumulated nearly 16,000 hand-drawn sunspot observation data over the years, necessitating the establishment of a comprehensive query and statistical system for scientific management and statistical analysis of these massive datasets. The digitized data comprises nearly 90,000 sunspot records and over 1 million valid data entries. The system provides a platform for data management, retrieval, and statistical analysis. Through the data management system, long-term statistical analysis of sunspot-related parameters can be conducted.

Full Text

The Yunnan Observatories Hand-Drawn Sunspot Data System

Zhu Gaofei^{1,2}, Zheng Sheng¹, Lin Ganghua², Zeng Xiangyun^{1,2}, Feng Yongli³, Tao Jinping³, Shu Yao¹

¹College of Science, China Three Gorges University, Yichang 443002, China

²Key Laboratory of Solar Activity, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China

³Yunnan Observatories, Chinese Academy of Sciences, Kunming 650011, China

Abstract: The Yunnan Observatories of the Chinese Academy of Sciences has accumulated nearly 16,000 hand-drawn sunspot observation records over the years. Establishing a comprehensive query and statistical system for the scientific management and analysis of these massive datasets has become essential. The digitized data includes approximately 90,000 sunspot records and over 1 million valid data entries. The system provides a platform for data management, retrieval, and statistical analysis, enabling long-term statistical studies of sunspot-related parameters.

Keywords: Sunspot; Data management; Sunspot information retrieval; Sunspot statistical analysis

Among various solar activities, sunspots are the most conspicuous and easily observable phenomena. Telescopic observations of sunspots began sporadically in 1610 abroad, but it was not until 1818 that relatively regular daily sunspot observations commenced, yielding more reliable data [1]. Early sunspot observations were typically hand-drawn on paper, resulting in a vast collection of physical records. Due to the fragility of paper and the difficulty of information retrieval, data integration and sharing among observatories have been hindered. Consequently, converting these paper records into digital format is necessary to facilitate analysis and access.

China's solar physics observation data possess a time-zone advantage, making them internationally valuable and relatively scarce. The hand-drawn sunspot images from Yunnan Observatories contain rich information and represent invaluable historical data for long-term solar activity research. The sunspot area data observed by Yunnan Observatories from 1981 to 1992 has been used multiple times to supplement and improve global sunspot datasets¹. Recognizing the unique value of China's hand-drawn sunspot observation materials, a systematic scanning, organization, and digital image storage project for nearly a century's worth of historical hand-drawn sunspot observations from Yunnan Observatories and Purple Mountain Observatory was initiated in 2014, led by the Huairou Solar Observing Station of the National Astronomical Observatories. To better leverage the scientific value of these historical observations for studying long-term variations in solar activity, it is essential to accurately and reliably extract sunspot information from the scanned digital images and enable online sharing. This digitization is achieved through image analysis techniques for segmenting handwritten characters and deep learning methods [5] for character recognition to automatically extract hand-drawn sunspot information [8].

Systematic management of the raw digitized sunspot data enables database retrieval of sunspot information and statistical analysis of various variations, offering convenience and ease of use. Most importantly, it allows for the aggregation of observational data from multiple observatories and effective data organization to facilitate comprehensive and in-depth research on sunspot activity and its effects on Earth. Therefore, systematic management and analysis of data following the digitization of hand-drawn sunspot images are particularly crucial. This paper describes the construction of the Yunnan Observatories hand-drawn sunspot data system using the MVC pattern combined with EF (Entity Framework) technology, enabling integrated data management and clear, intuitive statistical analysis.

1 Data Acquisition

The digital scanning of hand-drawn sunspot images from Yunnan Observatories for the period 1958-2015 has been completed. Each hand-drawn sunspot image

contains extensive and diverse information types, primarily including routine records of the hand-drawn sunspot images and detailed sunspot information, as shown in [Figure 1: see original paper].

Box 1 contains: Rotation number (the moment when the prime meridian passed the center of the solar disk on November 9, 1853, is defined as the start of the first solar rotation, with continuous numbering of rotation cycles; the rotation numbers and start dates for each cycle in a given year can be found in astronomical almanacs), observation date, Beijing time (standard time for 120°E), and Universal Time.

Box 2 contains: Position angle (the azimuth angle of the solar rotation axis north pole measured from the solar north point), and (the latitude and longitude of the central point of the solar disk at 0:00 UT on the observation date), and (representing the longitude of the central point of the solar disk at the observation time).

Box 3 contains: Number of sunspot groups in the northern hemisphere, number of sunspot groups in the southern hemisphere, total number of sunspot groups on the entire solar disk, number of sunspots in the northern hemisphere, number of sunspots in the southern hemisphere, total number of sunspots on the entire solar disk, Wolf number for the northern hemisphere, Wolf number for the southern hemisphere, and Wolf number for the entire solar disk.

Box 4 contains: Weather conditions, visibility, and remarks.

Ellipse contains: Sunspot group number, longitude, latitude, sunspot group structure type, total area of individual sunspot group, area of the largest sunspot in the group (in millionths of the solar disk area, abbreviated as area of the largest sunspot), and straight-line distance from the centroid of the sunspot group to the center of the solar projection disk (in millimeters, abbreviated as radius). , observatory factor, and Wolf number.

Through computer-aided manual processing, the information described above is digitized from the scanned hand-drawn sunspot images. Reference [11] details the application of deep learning methods for digitizing information from hand-drawn sunspot images. Through image segmentation, handwritten characters in the drawings are isolated and then recognized using deep learning methods. The digitized information is stored in two data tables with formats shown in and , containing nearly 90,000 sunspot information records and approximately 1 million valid data entries.

2 System Design

2.1 Function Design

The system has been fully constructed and provides not only retrieval and download services for sunspot observation data but also multi-faceted statistical analysis capabilities, enabling researchers to intuitively understand variation trends

in sunspot-related parameters. The system is primarily divided into three major functional modules and numerous sub-modules, as illustrated in [Figure 2: see original paper].

2.2 Architecture Design

The system is built using the MVC (Model-View-Controller) design pattern, which separates business logic, data, and interface display to organize code. This approach consolidates business logic into a single component, allowing interface improvements and user interaction customization without rewriting the business logic. The system structure is shown in [Figure 3: see original paper].

2.2.1 Model Class Design The model class design aims to create convenient and accessible entity classes for controller operations and view invocation. The system includes two primary model classes and multiple auxiliary view classes:

(1) Basic Information Class

This class primarily maps one-to-one relationships in the hand-drawn sunspot images, including data such as date, serial number, standard time, universal time, weather conditions, visibility, and remarks.

(2) Sunspot Information Class

This class maps various sunspot parameter information for each sunspot (group) in the hand-drawn images, including sunspot number, total area of individual sunspot group, area of the largest sunspot, longitude, latitude, radius, and other information.

2.2.2 Controller Design Controllers handle user requests and invoke appropriate views for display. The system includes three main controllers and multiple auxiliary controllers:

(1) Sunspot Master Controller

This controller primarily manages the display of various information from the hand-drawn sunspot image repository, including list displays of drawing information and detailed displays of sunspot information.

(2) Sunspot Query Controller

This controller processes data retrieval tasks based on user requests with different parameters, such as year, sunspot group number, total area of individual sunspot group, area of the largest sunspot, longitude, latitude, and other information, returning the relevant sunspot data to users through corresponding views.

(3) Sunspot Statistics Controller

This controller functions as an analytical processor, primarily responsible for statistical analysis of individual sunspot group total area, largest sunspot area, radius, and variations in total sunspot area and group numbers, presenting results visually to users through charts.

2.2.3 View Design View design, also known as UI interface design, is primarily responsible for receiving data from controllers and presenting it to users in a friendly manner. The system includes main views, shared views, partial views, and various display views corresponding to controllers.

3 System Functions

3.1 Sunspot Information Repository

The sunspot information repository organizes all digitized information from Yunnan Observatories' hand-drawn sunspot images by year, displaying all observed drawings for each year in list format to users. [Figure 4: see original paper] shows the data display interface for 1991. Double-clicking any drawing in the query results displays the image and its corresponding digitized information, as shown in [Figure 5: see original paper], with interactive mouse functionality to show local image details in the right display panel.

3.2 Sunspot Information Query

The sunspot information query module includes time range queries, sunspot group number queries, and comprehensive queries based on parameters such as date, sunspot group number, longitude, latitude, total area of individual sunspot group, and area of the largest sunspot. After processing by the sunspot query controller on the server, the system returns sunspot information matching the filter criteria. As an example, using comprehensive query with date range from January 1, 1990, to January 1, 2000, and total area of individual sunspot group between 0.5 and 1.5 yields results shown in [Figure 6: see original paper].

3.3 Sunspot Information Statistics

The sunspot information statistics module includes: variation curves for sunspot group area (including total area of individual sunspot group and area of the largest sunspot), variation curves for sunspot group radius length, total area variation, and sunspot number variation. The system can process annual observational data as well as statistical analyses spanning multiple years or even a century. Using 1991 data as an example, the total sunspot area variation is shown in [Figure 7: see original paper], and sunspot number statistics are shown in [Figure 8: see original paper].

3.4 Long-Term Sunspot Information Statistical Analysis

Currently, the system has imported hand-drawn sunspot information from 1993 to 2015. Using the system, researchers can clearly and intuitively observe annual variations in sunspot-related physical parameters. Statistical analysis of sunspot-related attributes from 1993 to 2015 yields results shown in . Through data visualization, the total area of individual sunspot groups, area of the largest

sunspot, and sunspot numbers from are plotted by year, as shown in [Figure 9: see original paper].

As [Figure 9: see original paper] demonstrates, the system's statistics on sunspot area and sunspot (group) numbers exhibit clear periodic peaks and troughs. The number of sunspots and sunspot groups correlates with the variation trend of sunspot area, with the magnitude of sunspot area and number directly indicating the intensity of solar activity. This pronounced characteristic phenomenon aligns with solar cycle variation patterns, which is significant for solar activity research and demonstrates that the data stored in the system is scientifically valid.

Conclusion

The Yunnan Observatories hand-drawn sunspot data system has completed construction of the hand-drawn sunspot image repository, sunspot information query, and sunspot information statistical analysis functions. The system enables researchers to conveniently query sunspot information under various conditions and conduct visual studies of sunspot-related activities. The system functionality can be further extended to provide convenient services for solar physics researchers and holds significant importance for advancing sunspot activity research. Future work will focus on expanding system capabilities such as sunspot motion trajectory plotting and three-dimensional sunspot image display to facilitate more convenient, intuitive, and clear application in long-term sunspot activity research.

We thank Yunnan Observatories of the Chinese Academy of Sciences for providing the hand-drawn sunspot image data for this study and acknowledge the suggestions and support from relevant staff members.

References

- [1] Li Kejun, Su Tongwei, Liang Hongfei. Periodicity of sunspot activity observed by modern sunspots [J]. Chinese Science Bulletin, 2004, 49 (24): 2511-2516.
- [2] Baranyi T, Gyori L, Ludmány A, et al. Comparison of sunspot area databases[J]. Monthly Notices of the Royal Astronomical Society, 2001, 323(1):223-230.
- [3] Győri L. Study of differences between sunspot and white light facular area data determined from SDO/HMI and SOHO/MDI observations[J]. Solar Physics, 2012, 280(2):365-378.
- [4] Balmaceda L A, Solanki S K, Krivova N A, et al. A homogeneous database of sunspot areas covering more than 130 years[J]. Journal of Geophysical Research Atmospheres, 2009, 114(A7):431-433.

- [5] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks[J]. Journal of Machine Learning Research, 2011, 15:315-323.
- [6] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [7] Yu D, Deng L. Deep learning and its applications to signal and information processing [Exploratory DSP][J]. IEEE Signal Processing Magazine, 2011, 28(1):145-154.
- [8] Niu X X, Suen C Y. A novel hybrid CNN-SVM classifier for recognizing handwritten digits[J]. Pattern Recognition, 2012, 45(4):1318-1325.
- [9] Zheng S, Zeng X Y, Lin G H, et al. Sunspot drawings handwritten character recognition method based on deep learning[J]. New Astronomy, 2016, 45:54-59.

¹ <https://www.ngdc.noaa.gov/stp/solar/sunspotregionsdata.html>

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.