A&A manuscript no.
(will be inserted by hand later)

Your thesaurus codes are:
04.01.1, 04.03.1

ASTRONOMY
AND
ASTROPHYSICS
3.12.1996

The General Catalogue of Photometric Data (GCPD). II.

J.-C. Mermilliod, M. Mermilliod, and B. Hauck

Institut d'Astronomie de l'Université de Lausanne, CH-1290 Chavannes-des-Bois, Switzerland

Received: September 15; accepted: September 15

Abstract. We describe the new version of the General Catalogue of Photometric Data. Our database of photometric photoelectric data has been updated and completely rebuilt. Most collections contain the data published to the end of 1995, taken from more than 2500 papers. We have updated the Master Index which now contains information for over 207500 stars and components. Public access to the data is provided through the World Wide Web at http://obswww.unige.ch/gcpd/gcpd.html. A connection has been established with the ADS Abstract Service which allows to retrieve data directly from the ADS output pages.

Key words: Database - photometry

1. Introduction

For twenty-five years the Institute of Astronomy of Lausanne University, Switzerland, which is a participating institute in the Strasbourg Data Centre, has been engaged in a long-term effort to collect published photoelectric photometric data in most existing systems. These collections certainly form the most complete photometric data bank ever realised. One of our main contributions concerns the homogeneization of stellar identifications. The data for the major systems (UBV, uvby, Geneva 7-colour) are available in Simbad, but no public and easy access has so far been provided for the wealth of existing photometric data.

The first database was coded by C. Nitschelm on a VAX computer and it has been described by Hauck et al. (1990). We have now completely rebuilt the database on Sun Sparc workstation under the /rdb package (Manis et al. 1988). In addition, we have updated all systems for which new data have been published before the end of 1995. The real new feature is the public access to the data collections, the Master Index, and the bibliographic references through the World Wide Web.

Send offprint requests to: Jean-Claude Mermilliod

Among the numerous systems recognised, less than 20 are in fact used more or less regularly. If several systems with wide bands (UBV and RI systems) or intermediate-width bands (uvby, DDO for example) became very popular, many others, mostly with narrow or intermediate-width bands designed to measure specific features in the stellar energy distribution, were not used very much. Collecting data for such small systems may seem of purely historical interest, but owing to the numerous features observed and the sensitivity of the filters or colour indices to various astrophysical quantities, each system offers specific interests which may prove useful in particular applications if one knows about their existence. The proliferation of photometric system resulted also from the realisation of similar systems with nevertheless slightly different filters and standard stars.

Our long-term work aimed at providing valuable solutions to the main questions: (1) in which systems has any star been observed and (2) how to get the best values among those published.

2. The photometric database

The photometric database consists of several directories containing the photometric datafiles, the references, the co-ordinates and various useful files, like cross-indexes and application programmes and scripts. The package used is /rdb (Manis et al. 1988), which is a Unix relational database management system. The advantage of /rdb is its simplicity and the possibility to include the database commands within Unix shell programming or command lines by using pipes and input/output redirection because this software uses simple ASCII files with a tabulator as field separator.

2.1. The photometric datafiles

We have prepared one file for each photometric system. Each one has the same structure and contains: the Lausanne identificator (LID), which uses the code numbering system designed by Mermilliod (1978), a remark on duplicity (component observed) or variability, the data, and the data source. We have found some 78 photometric systems, by considering that the use of different filters results in a different system. We have restored the full information on the components observed in visual double stars, like AB, BC or ABC. It is however clear that due to the frequent lack of information in the publications it is sometimes difficult to decide what has been really observed. Below 7 to 10 arcsec, we have considered that the double stars have usually not been separated, while above 25 to 30 arcsec, they should be separated. For separation between 10 and 25 arcsec, it is often difficult to decide what has been done, depending on the magnitude difference between the components. In case of doubts no information has been indicated when nothing was specified in the original paper.

Table 1 gives the number of entries for the photometric systems in which more than 1000 stars have been observed: the system designation, the total number of measurements, the real number of stars and the number of papers where data were found. For basic references on these photometric systems, please refer to those given by Hauck et al. (1990) paper or on the WWW site described below.

Table 1. Systems with more than 1000 stars

No	System	N_{Meas}	N_{Star}	N_{Ref}
01	UBV	169826	107996	1633
04	uvby	105891	66398	520
13	Geneva		39900	1
78	WBVR	13588	13588	1
11	Walraven	12872	12115	55
14	UBV Eggen	10489	10489	80
54	RI Cousins	10811	9506	153
12	DDO	10277	8077	100
08	UBVRI	9008	7397	118
02	UBV Cape	8309	7352	43
17	RI Eggen	7887	6777	83
21	Vilnius	7559	6763	81
16	uvby Eggen	6980	5861	45
36	Celescope		5719	1
19	RI Kron	5427	4489	45
09	IJHKLMN	3358	2353	44
77	Oja	2752	2752	2
81	CaII H&K	1990	1990	2
75	Maitzen Δa	1837	1780	20
18	13-Color	1911	1730	8
61	Andrews	1715	1715	3
20	Gyldenkerne	1453	1453	5
45	Washington	1445	1318	30
06	UVBGRI	1704	1299	26
63	Alexander	1218	1218	1
72	Moffet	1161	1152	11

The total number of entries is now about 400000, coming from 2526 papers. There are however some redundancies, because the UBV catalogue collects data from systems including UBV filters, like the various UBVRI systems. The corresponding UBV data are usually reduced to Johnson's standard UBV system.

When several observers are measuring stars in one particular system, they generally use differing instruments and techniques for observations and data reduction. This inevitably results in disparities in the published data (systematic errors, different accuracies). The main sources of these disparities are to be found in the photomultipliers used, the filter sets, the method of reduction outside the atmosphere, the choice and use of standard stars and the optical properties of the telescope. We have computed average data for two of the most used systems: the UBV (Mermilliod & Mermilliod 1994) and uvby (Hauck & Mermilliod 1996). The method developed for this purpose is explained in the introduction to the UBV catalogue (Mermilliod & Mermilliod 1994) and its main goal is to attribute a lower weight to discrepant observations. Briefly, a first weighted mean is computed, where the weights are determined by the number of observations in each data source. Then the differences between the mean value and the individual observations are computed. A second mean value is then computed, where the weights are proportional to the inverse of these differences. In this way, discrepant data receive a smaller weight and the average value is centred around the most common values.

The computation of mean photometric data has sometimes been criticised (Manfroid & Heck 1985), but we think that averaged data are very useful for many purposes. As we do not have access to the original measurements, it is of course impossible to re-reduce the data. The processing of the published data is also sometimes problematic in the absence of stars in common with other lists or because the list of the observed standards has not been published. We therefore think that the method used is the best possible in the present circumstances. In any case, the data for each source are fully available through the WWW server and anybody can retrieve them and decide which data he or she prefers to use.

Following an agreement with Geneva Observatory, only the mean values of the Geneva 7-colour system are stored in the database and accessible through the public WWW server described below.

2.2. Data sources

As a rule all astronomical journals are surveyed for new photometric data. The bibliographic references (data sources) are also collected in individual files, one for each system. The reference numbers are generally attributed sequentially in each system. The information entered in the files is the source number, the list of authors' names, the journal information (year, name, volume, page), the complete title and the ADS/CDS bibliographic code to enter the ADS Abstract Service system.

2.3. The Master Index

As described in Hauck et al. (1990), we have built a Master Index, which indicates in which systems each star has been observed. This is important for locating more easily the available information on each star. The new Index is different because it does not contain the photometry code as before, but the number of independent data sources in each photometry. The corresponding fields are left blank if no data at all exist for some photometric systems.

According to the relational database structure adopted, the Master Index contains the Lausanne identificator (LID), the remarks on duplicity or variability, and the number of data sources in each photometric system. This file contains more than 207500 stars. The merging of all these data has of course produced a considerable number of problems. Many of them have already been examined, the other ones will be looked at gradually.

2.4. Additional files

The database contains a number of additional files collecting general information for the stars (cross-identifications) and especially co-ordinates. Basic astrophysical catalogues containing non-photometric data have also been transformed into /rdb files to have additional on-line information and the possibility to perform selection on various criteria and later obtain the corresponding photometric data.

2.5. Stellar identifications

One of our main contributions concerns the stellar identifications. We have transformed the published identifications into a uniform system: various names are used in the literature, generally depending on the kind of sample investigated (proper motion stars, luminous stars, nearby stars, and so on) and the most time-consuming work is related to the necessity of collecting all the data for each star under the same identification, otherwise the advantage of working on large compilations partly vanishes. The basic principles adopted for this work have been explained in the description of the code numbering system (Mermilliod 1978) and the solution designed for star clusters is described by Mermilliod (1976).

2.6. Coordinates

Coordinates are very important for performing selections on the position in the sky. We have made a big effort to collect co-ordinates from the literature, the HST GSC, or to determine them from photographs of cluster or field areas. In spite of our efforts, several thousands of faint field stars (generally fainter than the limit of the GSC) still lack co-ordinates and we are not in a position to obtain them easily. This may result in the loss of faint stars that could be used, for exemple, to calibrate Schmidt plates or CCD frames.

2.7. Application software

Adapted software in the form of Unix shell scripts and Perl programmes has been developed to facilitate the handling of new data and database update. An effort will be made to automate the treatment of data files and the detection of errors or discrepancies for all sets of new data before they are introduced in the catalogue files. Graphics facilities are also being implemented to plot photometric diagrams or compare indices from different photometric systems. Eventually existing calibration codes will be implemented on the database to determine the astrophysical parameters of stars.

3. The World Wide Web server

The real new effort done is the development of a server on the World Wide Web, which permits an easy data retrieval in various ways. The present interface has been designed to retrieve data for small numbers of stars or prepare samples based on magnitudes or colour indices. The possibility to enter the database with a list (file) of stellar identifications will be investigated. For intensive use of the database, i.e. a large sample of predefined stars, it is still better to contact the authors. There are basically three forms to query (1) the Master Index, (2) each photometric system directly, and (3) the bibliographic reference files.

3.1. The Master Index

The query of the Master Index is aimed at displaying the existing information for one or several stars. It uses the power of hypertext to link the displayed information with CGI scripts that query the database and display further information. The query of the Master Index is possible with either a basic star identification, or co-ordinates. The output page presents the number of available data and a simple click on the link displays the real data in the selected photometric system. The query by coordinates (equinox 1950 or 2000) will display the same information for the stars falling in the area defined by the central co-ordinates and the given limits in right ascension and declination.

3.2. The photometric files

When the observations for one or several stars are desired in a specific system, the second form offers a direct access to each of the 78 photometric systems. Each photometric catalogue may be queried by star identification or co-ordinates. Selection on the photometric indices and parameters can also be made

3.3. The references

Access to the references make it possible to look for the papers published by an author or to search for some words in the titles of the papers. The link added to the output permits to display the whole dataset corresponding to the publication selected. This facility allows to easily get on-line data from 2526 papers published from 1953 to 1995. The stellar identifications are generally decoded and displayed with the usual acronyms, although many small lists do not have any acronyms and the code numbers are given. The Web site also contains the description of the code numbering system where all definitions and references can be found. A cross-link has been established with the ADS Abstract Service (Eichorn et al. 1996). We have added in our database the bibliographical code used by the ADS services and one can thus directly query the Abstract Service from our pages. Conversely, when papers are obtained from a query of the ADS, data can be directly obtained with the datalink provided, which is connected to our database.

3.4. The Web site

In addition to the form to query the database, the Web site offers varied information, as for example, a technical description of the photometric systems, or selected references to calibration papers.

4. Distribution of the data

We have transfered to the Strasbourg Data Centre all the photometric data files and the CDS will distribute them in full through their anonymous ftp facilities and the catalogues will also be accessible through the new CDS/ESRIN facilities called VizieR.

5. Discussion and Conclusion

The feasibility and the quality of such a catalogue strongly depends on the amount of work invested in rendering stellar identifications homogeneous. Stars can be called under a rather large variety of identifications, and many of them are in fact used. Therefore an important part of the action of cataloguing the data is devoted to the recovery of the best possible (primary) identification. The same problem is also encountered for stars in the region of star clusters and has been solved in an original way (Mermilliod 1979; 1986).

The existence of the database allows also to perform numerous checks on the data, by comparing the V magnitudes, plotting similar indices like (B-V, b-y) to search for anomalies and discrepant data. This may reveal interesting astrophysical questions, but more often uncovers a more or less trivial error. The origin may be difficult to trace back because it may happen at any stage of the chain, from the work at the telescope to the inclusion in the database. The first thing we check is that the data entered into the database are identical to those published.

Owing to its completeness, the GCPD offers wide possibilities for various applications. Usage may concern a statistical approach: (1) to locate the existing data for any star, (2) to prepare an observing campaign and identify the stars with or without data, (3) to investigate the completeness of available data for specific samples, like the nearby or bright stars, (3) to form a sample of stars measured in two or more photometric systems and study transformation relations, (4) to do any astrophysical research programme.

Our future plan is to implement existing calibration programmes on the database to allow users to get directly the astrophysical quantities through the Web server. Codes for the uvby and Geneva systems are available, one for the DDO photometry has been written and one for the Washington system has been received from J.J. Claria. We would be pleased to start any collaboration to implement additional software on the database and make the facilities available through the WWW server.

The main requirement of data archiving is to contribute to the improvement of the data quality and to the rejection of the poorest ones. It is the only way for disseminating reliable data and will provide future generations of astronomers with data that retain some interest. The effort made during long nights of observations by so many observers warrants some supplementary work to preserve the results and transmit them to posterity. Any correction or new information is welcomed to improve the database.

Acknowledgements. This long-term work has been supported by continuous grants from the Swiss National Foundation for Scientific Research (FNRS). We are pleased to thank Dr Juan Claria (Cordoba, Argentina) for sending a copy of his code on the Washington system and Mrs B. Wilhelm for reading and correcting the manuscript.

References

Eichhorn G., Accomazzi A., Grant C.S., Kurtz M.J., Murray S.S. 1996, ASP Conf. Ser. 101, 569
Hauck B., Mermilliod M. 1996, in preparation
Hauck B., Nitschelm C., Mermilliod M., Mermilliod J.-C. 1990, A&AS 85, 989
Manfroid J., Heck A. 1985, in "Calibration of Fundamental Stellar Quantities", Eds D.S. Hayes, L.E. Pasinetti and A.G.D. Philip (Reidel, Dordrecht), p. 565
Manis R., Schaffer E., Jørgensen R. 1988, "Unix Relational Database Management", (Prentice Hall, New Jersey)
Mermilliod J.-C. 1976, A&AS 24, 159
Mermilliod J.-C. 1978, Bull. Inform. CDS 14, 32
Mermilliod J.-C. 1979, A&AS 36, 163
Mermilliod J.-C. 1986, A&AS 63, 293
Mermilliod J.-C., Mermilliod M. 1994, "Catalogue of Mean UBV Data on Stars", Springer Verlag (New-York)

This article was processed by the author using Springer-Verlag LaTeX A&A style file *L-AA* version 3.