

MASSEE International Congress on Mathematics MICOM'2009

Workshop Astroinformatics: Computer Science and Mathematical Methods in Astronomy

Ohrid, September 16-20, 2009

Direct access to digitized logbooks of astronomical Wide-Field Plate Collections

N. Kirov*, M. Tsvetkov**

*Institute of Mathematics and Informatics, BAS and Department Informatics, NBU

*Institute of Astronomy, Bulgarian Academy of Sciences

Introduction

Logbooks for astronomical observations

Logbooks contain metadata for plates, written by the observers at the time of observation.

Tasks:

- Link from a record of WFPDB to the corresponding logbook page image
 - valuable information in **Remark** section.
- Check the data in WFPDB for correctness – the data are inserted manually (from the keyboard)
- Insert new data in WFPDB directly (from logbook files).

Harvard College Observatory

22

INSTRUMENT, # ROSS-RB7

No.	Class	Object	R. A.	Dec.	Started	Obs. H. A.	Obs. Dec.	Tel. E or W.	Load	Focus	Prisms	Sky at Start	Stopped	Exp.	Crds	Crk	Dew	Ref.	Obstr	REMARKS
✓ 1917	L	-	19 00 0	17 30 1	30E. 0	-	88	-	-	-	-	0	19	00 90	-	-	-	V.		
✓ 1918	"	-	19 30 -60	19 02 0	28E. -60	-	110	-	-	-	-	0	20	32 90	-	-	-	V+S		
✓ 1919	"	-	21 00 0	20 34 0	24E. 0	-	90	-	-	-	-	0	22	04 90	-	-	-	S.		
✓ 1920	"	-	22 30 -60	22 06 0	24E. -60	-	110	-	-	-	-	0	23	36 90	-	-	-	S. Dew on lens.		
July 19-20 (Sat.-Sun.) 1931																				
✓ 1921	L	-	20 00 -45	20 15 0	18W. -45	-	104	-	-	-	-	.1	21	45 90	-	-	-	V. Lighting right round in E.		
✓ 1922	"	-	22 00 -45	21 47 0	13E. -45	-	105	-	-	-	-	0	23	17 90	-	-	-	V. Damp. Lens wet with dew.		
July 20-21 (Sun.-Mon.) 1931																				
✓ 1923	L	-	14 00 -85	14 09 0	0 09W. -85	-	105	-	-	-	-	.1	15	39 90	-	-	-	D.		
✓ 1924	L	-	16 30 -60	15 41 0	0 49E. -60	-	110	-	-	-	-	.2	17	11 90	-	-	-	D. Cirri in South and West.		
✓ 1925	L	-	17 00 0	17 $\frac{1}{2}$ 0	0 $\frac{1}{2}$ N	0	88	-	-	-	-	.5	18	16 $\frac{1}{2}$ 0	-	-	-	D. Cirrus forming, passing over, disappearing again etc. → Yellow & continual - sky practically overcast with cirri.		
July 21-22 (Tue.-Wed.) 1931																				
✓ 1926	L	-	14 00 -45	14 55 0	55W. -45	-	103	-	-	-	-	.2	16	25 90	-	-	-	V. blonde in N. Moon 6 ^d .		
✓ 1927	"	-	16 00 -45	16 27 0	27W. -45	-	104	-	-	-	-	.2	17	57 90	-	-	-	V. " " "		
✓ 1928	"	-	18 00 -45	17 59 0	01E. -45	-	104	-	-	-	-	.1	19	29 90	-	-	-	V. " " "		
✓ 1929	"	-	19 00 0	19 31 0	31W. 0	-	88	-	-	-	-	0	21	01 90	-	-	-	V ² Moon 6 ^d . Moon set = 20:02		
✓ 1930	"	-	21 00 0	21 03 0	03W. 0	-	88	-	-	-	-	0	22	33 90	-	-	-	F.		
✓ 1931	"	-	0 00 25°	22 35 1	25E. 75°	-	118	-	-	-	-	0	24	08 90	-	-	-	F.		

23

The Royal Observatory of Belgium, located in Uccle

NUMÉRO du cliché.	ASCENSION DROITE approchée du centre.	B. D.	G.	α pour 1900,0.		δ pour 1900,0.	PRÉCESS
				α pour 1900,0.	δ pour 1900,0.		
34	$5^h 1^m$	$32^{\circ} 29' 0''$	8.1	$5^h 0^m 56^s 6$	$32^{\circ} 46' 7''$	$+ 3^{\circ} 905$	-
				$32^{\circ} 29' 1''$	$56^s 7$		
35	$5^h 10^m$	$32^{\circ} 93' 1''$	8.8	$9^m 56^s 4$	$33^{\circ} 3' 3''$	$+ 3^{\circ} 912$	-
36	$5^h 19^m$	$33^{\circ} 10' 40''$	8.9	$19^m 56^s 1$	$33^{\circ} 3' 52''$	$+ 3^{\circ} 930$	-
37	$5^h 28^m$	$32^{\circ} 10' 28''$	6.4	$26^m 54^s 8$	$32^{\circ} 43' 51''$	$+ 3^{\circ} 923$	-
38	$5^h 37^m$	$33^{\circ} 11' 34''$	7.7	$37^m 57^s 8$	$33^{\circ} 3' 16''$	$+ 3^{\circ} 939$	-
39	$5^h 46^m$	$33^{\circ} 11' 21''$	8.8	$45^m 9^s 4$	$33^{\circ} 19' 10''$	$+ 3^{\circ} 950$	-
40	$5^h 55^m$	$33^{\circ} 12' 09''$	6.9	$55^m 39^s 9$	$33^{\circ} 7' 47''$	$+ 3^{\circ} 945$	-
41	$6^h 4^m$	$33^{\circ} 12' 65''$	8.2	$6^h 4^m 12^s 2$	$33^{\circ} 0' 56''$	$+ 3^{\circ} 921$	-
42	$6^h 13^m$	$33^{\circ} 12' 92''$	8.8	$11^m 44^s 7$	$33^{\circ} 12' 23''$	$+ 3^{\circ} 947$	-
43	$6^h 22^m$	$33^{\circ} 13' 26''$	8.5	$21^m 28^s 3$	$33^{\circ} 0' 52''$	$+ 3^{\circ} 938$	-
				$21^m 50^s 33$	$13' 49''$	$+ 3^{\circ} 944$	
44	$6^h 31^m$	$33^{\circ} 13' 47''$	8.8	$29^m 56^s 0$	$32^{\circ} 58' 56''$	$+ 3^{\circ} 933$	-

Leander McCormick Observatory, Charlottesville, VA

					Mid	E.P.O.								Cause	
					E.S.T.	Hours									
38	May 23	Vega	18° 34'	+38° 42'	23	14 06	0 36 E	10	M	2-3	2-3	clear	b6	20	-Plan 8024 g
39	"	"	"	"	22	14 31	11 E	10	M						g
40	"	IC	19 34	+50 4	12	15 17	0 26 E	10	M						g Trail & Vega
41	"	IC	19 34	+50 4	13	16 02	13 E	10	M						Wrong field
42	May 24	& Comae Ber	18 6	+17 57	21	9 0	0 13	20	M	1-2	2	Thin Clouds	68	20	" "
43	"	Virgo	12 38	-1 1	20	9 37	0 28 W	10	M	1-2	2	Hazy	74	20	" " N.G. Clock driving poorly
44	May 25	"	"	"	21	9 52	0 43	10	M	1-2	2	Hazy	74	20	" " N.G. Set new guidance star
45	"	Comae Ber	18 6	+17 57	20	9 43	1 0 W	20	M						g Trail 18 45 17 57
46	"	Boötis	13 01	+18 48	23	10 18	8 W	20	M						p Trail y Bootis
47	"	Vega	18 34	+38 42	24	10 38	1 28	20	M						g
48	"	IC	19 34	+50 4	15	11 22	29 W	10	M	2	2	Clear	68	20	g
49	"	"	"	"	6	11 37	1 44	10	M						g
50	"	IC	19 34	+50 4	22	14 3	0 32 E	10	M						g
51	May 30	Virgo	12 38	-1 1	20	14 16	21	10	M						g
52	"	Cephe b	19 10	+49 42	12	15 29	4 E	10	M						g
53	"	"	"	"	0	16 6	0 36 E	10	M	1-2	2	Clouds	75	20	" " f Lamp turned out.
54	"	IC	19 34	+50 4	1	18	23	10	M	3	2	Clear	70	"	g
55	May 31	Virgo	12 38	-1 1	5	8 53	33	10	M	3	3	Thin Clouds	72	20	" " f Trail & Vega
56	"	"	"	"	6	9 10	0 50 W	10	M						g
57	"	"	"	"	19	9 55	0 57 E	10	M						g
58	"	"	"	"	20	14 7	45	10	M						g
59	"	"	"	"	20	14 25	0 24 E	10	M						g
60	"	IC	19 34	+50 4	17	15 2	0 18 E	10	M						g
61	"	Virgo	12 38	-1 1	10	12	5	10	M	2	2	Thin Clouds	72	20	" " f Trail & Vega
62	"	"	"	"	5	8 55	0 16 W	10	M						g
63	"	"	"	"	6	9 12	0 53	10	M						g
64	"	"	"	"	5	9 24	1 8 W	10	M						g

Bonn Observatory

Platten Nr.	Datum	Objekt	Aufnahmewinkel (WZ)	Struktur winkel	Fokus	Plattenort Filter	Entwicklung	Emulsionsart	Temp	Blatt	Bemerkungen
	A 1970										
2439	November 23/24	M4C 7023	20 ¹³ -20 ⁴³	+5 ⁵⁰ "	4.6	103a-0 16x16 mm	Reel, 1:20	B:2, Ci	+5°C		Flackung
2440	"	Pcr OB2	21 ⁴⁷ -22 ¹²	+1 ²⁰ "	"	"	"	DS 2, R 3-4	"	"	
2441	"	M4C 225	22 ²⁰ -22 ⁵⁰	+2 ⁰⁹ "	"	"	"	B:2, 0	"	"	
2442	"	JC 1848	22 ⁵⁴ -23 ²⁴	+0 ⁴⁵ "	"	"	"	DS 1, R 3	"	"	
2443	November 24/25	M4C 225	23 ¹⁵ -23 ⁴⁵	+3 ⁰⁸ "	"	"	"	B:2, 0, DS 2, Ci	+6°C	"	
2444	"	JC 1848	23 ⁵⁰ -00 ²⁰	+1 ⁴³ "	"	"	"	DS 3-4, R 3	"	"	
2445	"	Pcr OB2	00 ²⁵ -00 ⁵⁵	+1 ²⁴ "	"	"	"	DS 3	"	"	
2446	"	M4C 225	02 ¹⁶ -02 ³¹	+6 ⁰⁹ "	4.7!	"	"	B:2, 0 (Ci?)	+5°C	"	
2447	"	M4C 225	02 ³⁴ -03 ⁰⁴	+6 ²⁷ "	4.6	"	"	DS 1-2, R 3	"	"	
2448	"	Pcr OB2	03 ⁰⁸ -03 ³⁸	+4 ⁰² "	"	"	"	R 3-4	"	"	
2449	"	JC 1848	03 ⁴³ -04 ¹³	+5 ³⁷ "	"	"	"	R 3	"	"	
2450	November 25/26	Cep T1	17 ⁴⁰ -18 ¹⁰	+1 ³⁶ "	"	"	"	B:2, 0, DS 2	+9°C	"	
2451	"	Lac CB1	18 ¹³ -18 ⁴³	+0 ²⁰ "	"	"	"	R 2	"	"	
2452	"	Cep OB2	18 ⁴⁷ -19 ¹⁷	+1 ⁵⁰ "	"	"	"	R 3	"	"	

Overview of accessible logbooks

Observatory	WFPDB ID	Plates	Plates in WFPDB	Logbooks No	Data GB	Digitized Logbooks
Harvard	HAR	484668	111665	75	44	Yes
Potsdam	POT	8415	1661	11	1	Yes
Jena	JEN	1100	0	1	4	Yes
Bonn	BON	10530	0	9	0	No
Rozhen	ROZ	9152	9332	2	—	Partially
Bamberg	BAM	30535	24914	32	—	Partially
Vatican	VAT	7936	0	4	0	No
Copenhagen	COP	1364	0	1	0	No

Methods and technologies for data extraction

- Handwritten recognition
 - Low quality of images
 - + Structured pages (tables)
 - + Number recognition
 - + Constraints in columns
1. Separation of the text region on the page
 2. Binarization, rotation
 3. Rows detection
 4. Columns and cells detection
 5. Numbers spotting
 6. Numbers recognition

As a result, a digital copy of the logbook will be produced!

Example: The Harvard College Observatory Astronomical Plate Stacks

http://tdc-www.cfa.harvard.edu/plates/a/logs/01b/a01b_0212.jpg (2048 x 1536 Pixels)

212

INSTRUMENT, 24 = Bruce.

No.	Class.	Object.	R. A.	Dec.	Started.	Obs. H. A.	Obs. Dec.	Tel. E. or W.	Load.	Focus.	Prisms.
6881	-	X Serpentis	15 12 +2°5	14 47 0	26°E	+2°5	-	-	-	-	0
6891	-	Nova Cygni	16 54 -12°7	15 58 0	56°E	-12°7	-	-	-	-	0
6901	L th	-	16 10 +2°5	17 2 0	52°W	+2°5	-	-	-	-	0
6911	L	-	16 50 +2°5	17 22 0	32°W	+2°5	-	-	-	-	0
6921	L	-	17 30 +2°5	17 45 0	15°N	+2°5	-	-	-	-	0
6931	L	-	18 10 +2°5	18 03 0	07°E	+2°5	-	-	-	-	0

* = Wunder 8

213

DATE, Wednesday, June 20 1894

Est.	Sky at Start.	Stopped.	Clouds	Clouds	Dew.	Ref.	Obs'r.	REMARKS.
15 17 1"	15 47 0 1	-	X	W	X 688	Followed,		.586
16 28 42	16 58 0 0	-	X	W	X 689	"		.636
17 10 22	17 19 0 0	-	W					.665
17 32 10	17 42 0 0	-	C					.680
17 52 14	18 00 0 0	-	C					.694
18 10	18 17 5 0	-	C					.706

Text area

212

INSTRUMENT, 24 = Breee.

No.	Class.	Object.	R. A.	Dec.	Started.	Obs. H. A.	Obs. Dec.	Tel. E. or W.	Load.	Focus.	Prisms.
6881	-	X Sagittario	5 librae	15 12 +2°5	14 47	0 26 E	+2°5	-	-	-	0
689	-	Nova (Ghementi)	16 54 -12°7	15 58	0 56 E	-12°7	-	-	-	-	0
690	Ltt	-	16 10 +2°5	17 2	0 52 W	+2°5	-	-	-	-	0
691	L	-	16 50 +2°5	17 22	0 32 W	+2°5	-	-	-	-	0
692	L	-	17 30 +2°5	17 45	0 15 W	+2°5	-	-	-	-	0
693	L	-	18 10 +2°5	18 03	0 07 E	+2°5	-	-	-	-	0

* - Mission 5

DATE, Wednesday, June 20, 1894

213

+008
-005
1.003

Digitized by srujanika@gmail.com

.586

50
631

656
29
115

.665
15

.680

-694

- 1 -

• 106

100

100

100

100

Rows detection

Mark names

Go 7 rows Directory Harvard File p11112 Next Previous Save All: Go-Save

Threshold 141 Left m. 2 Right m. 1150 Bot.m. 200 Top m. 112 W.value 10 W.height 5 Row height 14

212

INSTRUMENT, 24 = Bress.

No.	Class.	Object.	R. A.	Dec.	Started,	Obs. H. A.	Obs. Dec.	Tel. E. or W.	Load.	Focu,	Prisms,	End.	Sky at Start.	Stopped.	CPds.
6881	*	Sejanus	-5° Librae	15 12 +2°5	14 47 0	26°E	+2°5	-	-	-	0	15 17	, "	15 47 0	,
689	*	Nova (Ghini)	16 54	-12°7	15 58 0	56°E	-12°7	-	-	-	0	16 28	42	16 58 0	,
690	L+	-	16 10	+2°5	17 2 0	52W	+2°5	-	-	-	0	17 10	22	17 19 0	,
691	L	-	16 50	+2°5	17 22 0	32W	+2°5	-	-	-	0	17 32	20	17 42 0	,
692	L	-	17 30	+2°5	17 45 0	18W	+2°5	-	-	-	0	17 52	14	18 00 0	,
693	L	-	18 10	+2°5	18 03 0	07E	+2°5	-	-	-	0	18 10	-	18 17 5	,

* - Meissier 5

Columns and sells determination

Mark names

Go Directory Harvard File p11112 Next Previous Save All: Go-Save

Threshold 142 Left.m. 4 Right.m. 1400 Bot.m. 50 Top.m. 112 W.value 10 W.height 5 Row height 14

No.	Class.	Object	R.A.	Dec.	Started	Obs. H. A.	Obs. Dec.	Tel. E. or W.	Load.
688	*	Sculptor	15 12 +2°5	14 47 0	26°E +2°5	-	-	-	-
689	*	Nova Ceph	16 54 -2°7	15 58 0	56°E -12°7	-	-	-	-
690	L	-	16 10 +2°5	17 2 0	52W +2°5	-	-	-	-
691	L	-	16 50 +2°5	17 22 0	32W +2°5	-	-	-	-
692	L	-	17 30 +2°5	17 45 0	15W +2°5	-	-	-	-
693	L	-	18 10 +2°5	18 03 0	07E +2°5	-	-	-	-
		* - Murier							

Digital copy of the page

212

INSTRUMENT, 24 = Buse.

No.	Class.	Object.	R. A.	Dec.	Started.	Obs. H. A.	Obs. Dec.
6881	-	X Serpentis	5 Librae 15 12 +2°5	14 47 0	26°E	+2°5	
689	*	- Noval Ghindji	16 54 -12°7	15 58 0	56°E	-12°7	

DATE, *Wednesday, June 20, 1894*

INSTRUMENT 24 = B.... DATE Wednesday, June 20, 1894

No.	Class.	Object.	R.A.	Dec.	Started.	Obs. H.A.	Obs. Dec.	Enl.	Stopped.
688	-	5 ...	15 12	+2° 0.5	14 47	0 25e	+2° 0.5	15 17	15 47
689	-	Nova...	16 54	-12° 0.7	15 58	0 56e	-12° 0.7	16 28	16 58

Connection to Wide-Field Plate Database – Sofia

Link to a page

One-to-one correspondence:

WFPDB Instrument <-> Logbook (a set of page images)

Index of /plates/a/logs/01b

Name	Last modified	Size
[DIR] Parent Directory	07-Apr-2007 09:24	-
[IMG] a01b_0000.jpg	05-Apr-2007 16:33	675k
[IMG] a01b_0001.jpg	05-Apr-2007 16:33	557k
[IMG] a01b_0002.jpg	05-Apr-2007 16:33	557k
[IMG] a01b_0004.jpg	05-Apr-2007 16:33	546k
[IMG] a01b_0006.jpg	05-Apr-2007 16:33	546k
[IMG] a01b_0008.jpg	05-Apr-2007 16:33	560k
[IMG] a01b_0010.jpg	05-Apr-2007 16:33	573k
[IMG] a01b_0012.jpg	05-Apr-2007 16:33	566k
[IMG] a01b_0014.jpg	05-Apr-2007 16:33	578k
...		

Instrument ≡ (observatory, telescope, site)

<http://tdc-www.cfa.harvard.edu/plates/a/>



Telescope
SMITHSONIAN ASTROPHYSICAL OBSERVATORY
Data Center



Harvard-Smithsonian Center for Astrophysics

Harvard Observatory Plate Stacks A Series Characteristics

Basics

Plate Size	Plate Scale	Telescope	Focal Length
8x10, 14x17	60"/mm	24-inch Bruce Doublet	135 inches

Observatory

Plate nos.	Location	Dates
1- 1730	Cambridge, Mass.	1893-11-09 - 1895-10-21
1731-14110	Arequipa, Peru	1896-04-02 - 1926-11-06
14111-27504	Bloemfontein, South Africa	1929-07-15 - 1950-10-11

Instrument ≡ (observatory, telescope, site)

Wide-Field Plate Database – Sofia

Details for archive: HAR061A

Location of the Archvie:

Site: **Cambridge**

Country: **USA MA**

Observatory:

Name: **Harvard**

Site: **Cambridge**

Country: **USA MA**

Time zone: -5 h

East longitude: -71° 52.2'

Latitude: 42° 22.8'

Altitude: 24 m

Clear aperture: 0.61 m

Mirror diameter:

Focal length: 3.34 m

Scale: 60 "/mm

Type: Rfr

Field size: 7.2°

Years of operation:

From: **1893**

To: **1895**

P/F:

Many-to-one: WFPDB record (plate) -> Logbook page

INSTRUMENT 24 = B.... DATE Wednesday, June 20, 1894

No.	Class.	Object.	R.A.	Dec.	Started.	Obs.H.A.	Obs.Dec.	Enl.	Stopped.
688	-	5 ...	15 12	+2 ⁰ .5	14 47	0 25e	+2 ⁰ .5	15 17	15 47
689	-	Nova...	16 54	-12 ⁰ .7	15 58	0 56e	-12 ⁰ .7	16 28	16 58

Add a list of plate numbers to a page (manually or as a result of first column recognition)

page 212 -> 688-693 (6 records)

catalog Harvard A Series Plates, equinox 1900, radecsys B1900

id	ra	dec	exp	epoch
00688	15:12	+02 : 50	60	1894-06-21T02:03:50.4
00689	16:54	-12 : 70	60	1894-06-21T03:15:50.4

Wide-Field Plate Database – Sofia

IDobs	IDins	IDsuf1	IDno	RAJ2000	DECJ2000	Date	Time
HAR	061	A	000688	15 17 02	02 27 55	1894 06 21	02 03 50
HAR	061	A	000689	16 59 37	-13 19 05	1894 06 21	03 15 50

ID of a plate in WFPDB: [IDobs IDins IDsuf1 IDno]

page number is contained in the image filename

[HAR 061 A 000688] -> http://tdc-www.cfa.harvard.edu/plates/a/logs/01b/a01b_0212.jpg

Numbers recognition

Interactive software system specialized for logbook recognition.

Conclusions

- Direct access to digitized logbooks of astronomical Wide-Field Plate Database – Sofia plate -> image of logbook page: manually or automatic (first column numbers recognition)
- OCR of the numbers in logbooks allows automatic insertion of new records in WFPDB

This talk has been supported by Grant No. DO02-275/2008
Astroinformatics, Bulgarian NSF, Ministry of Education and
Science.

Thank you for your attention.