

PROBLEM OF MISTAKES IN DATABASES, PROCESSING AND INTERPRETATION OF OBSERVATIONS OF THE SUN. I.

N. I. Lozitska

Astronomical Observatory of Taras Shevchenko National University of Kyiv
Observatorna 3, Kyiv, 04053, Ukraine, nloz@observ.univ.kiev.ua

ABSTRACT. In databases of observations unnoticed mistakes and misprints could occur at any stage of observation, preparation and processing of databases. The current detection of errors is complicated by the fact that the works of observer, databases compiler and researcher were divided. Data acquisition from a spacecraft requires the greater amount of researchers than for ground-based observations. As a result, the probability of errors is increasing. Keeping track of the errors on each stage is very difficult, so we use of cross-comparison of data from different sources. We revealed some misprints in the typographic and digital results of sunspot group area measurements.

Keywords: Sun: – Areas of sunspot groups, Observations, Identification of errors, correction of datasets

1. Introduction

Systematic and isolated errors can be found in various datasets. The difference of sunspot group areas from different observatories had been considered as systematic errors in publication by Gnevysheva (1968), Baranyi (1999, 2013). In present paper we identify and correct some isolated errors in databases of sunspot group areas.

2. Data of Observations

Data of sunspot groups area is available at the sites of Solar-Terrestrial Physics Division (STP) NOAA <http://www.ngdc.noaa.gov/stp/space-weather/solar-data/> (Solar Observing Optical Network USAF), Solar Physics Group (SPG) of NASA's Marshall Space Flight Center <http://solarscience.msfc.nasa.gov/>, Debrecen Observatory (Györi, 2011) <http://fenyi.solarobs.unideb.hu/>, Pulkovo Interactive database of solar activity (Miletsky, 2005) <http://www.gao.spb.ru/database/csa/>.

3. The Cross-comparison Method and Results

We made the comparison of sunspot group areas from databases of Pulkovo, Greenwich, and Debrecen Observatories. Our interest was focused on revealing especially large random errors. The greatest isolated errors were detected in Pulkovo's datasets.

Table 1 presents such comparison for whole group area and for leading spot of group SD 1939264 with same group RGO 13394 (transition of 1939, August 26 – September 7).

Table 2 contains the comparison for whole group area and for leading spot of group SD 1939-274 with same

group RGO 13405 (transition of 1939, September 4–16). From these tables we can see the significant excess of area of sunspot groups, up to 1500 millionth parts of hemisphere (m.p.h.) Even greater differences were found for the leading spots in some days. In first case, we can suppose the transfer printing slips from Catalogues of solar activity to electronic dataset. In second one, the probable explanation is contribution of heterogeneous data from different observatories in average value presented in SD catalogue without necessary alignment to same measuring system. Observations in the 1930th – 1940th were just being started. Likely, two parameters – provisional areas in millionth parts of a disk and corrected areas in millionth parts of a hemisphere – were mixed.

Table 3 demonstrates analogous data for the largest groups SD 194618 and SD 1947057 which were observed in 1946, January 29 – February 12, and in 1947, March 4 – 17, respectively.

Table 4 shows how significant errors in measurement result records of areas in SOON USAF observatories lead to overestimation of average daily group areas, calculated by SPG, using the primary data of SOON observatories. Our calculations of daily values USAF_{corr} which were made after removal of errors coincide well with DPD observatory values. These essential errors of measurements of USAF observatories we have identified as the values that go beyond 3 standard deviations. Some of them were misspellings during numbers input when the number of adjacent columns are shifted to the right so that added to the fourth rank in the value of the area. The rest seems to have been the case of records in the database in millionth parts of the disc, which are sometimes found among the correct values of the area, measured in millionth parts of the hemisphere. In this case, situation is analogous to Pulkovo catalogue. Measurements carried out on different observatories had contained the provisional data among corrected ones. The Solar Physics Group had used this data for calculation of daily averaged data without eliminating errors of each measurement. As a result, many areas of daily values of sunspot groups are significantly overestimated.

Acknowledgements. We express our gratitude to all the observers, whose measurements are used in the work.

References

- Gnevysheva R.N.: 1968, *Sov. Phys. Astron.*, **11**, 976.
- Baranyi T. et al.: 2013, *MNRAS*, **434**(2), 1713.
- Baranyi T. et al.: 1999, in Proc. ESA **SP-448**, 569.
- Györi, L. et al.: 2011, in Proc. IAU, **S273**, 2011, 403.
- Miletsky, E. et al.: 2005, in Proc. of Pulkovo conf., **9**, 47.

Table 1:

Data	S_{gr} SD, mph	S_{gr} , RGO, mph	S_{gr} SD - S_{gr} RGO, mph	S_{lider} SD, mph	S_{liter} RGO, mph	S_{lider}/S_{gr} SD,	S_{lider}/S_{gr} SD,
1939 09 03	2376	2295	81	1476	1125	0,62	0,49
1939 09 04	3666	2513	1153	1730	1472	0,47	0,59
1939 09 05	3272	2785	487	1362	1053	0,42	0,38
1939 09 06	6174	3054	3120	6174	1438	1	0,47

Table 2:

Data	S_{gr} SD, mph	S_{gr} , RGO, mph	S_{gr} SD - S_{gr} RGO, mph	S_{lider} SD, mph	S_{liter} RGO, mph	S_{lider}/S_{gr} SD,	S_{lider}/S_{gr} SD,
1939 09 08	2561	2282	279	425	404	0,17	0,18
1939 09 09	3045	2561	484	3045	431	1,00	0,17
1939 09 10	3810	2623	1187	3674	497	0,96	0,19
1939 09 11	3876	2927	949	3653	499	0,94	0,17
1939 09 12	4226	2723	1503	4114	513	0,97	0,19
1939 09 13	4442	2993	1449	2543	557	0,57	0,19

Table 3:

Data	S_{gr} SD, mph	S_{gr} RGO, mph	S_{gr} SD - S_{gr} RGO, mph	Data	S_{gr} SD, mph	S_{gr} RGO, mph	S_{gr} SD - S_{gr} RGO, mph
1946 02 02	6402	4940	1462	1947 03 08	5630	3615	2015
1946 02 03	7406	5080	2326	1947 03 09	6826	3886	2940
1946 02 04	5633	4952	681	1947 03 10	7500	4179	3321
1946 02 05	5692	4533	1159	1947 03 11	5091	4547	544
1946 02 06	6655	4799	1856	1947 03 12	6744	4554	2190
1946 02 07	5732	5202	530	1947 03 13	5208	4205	1003
1946 02 08	5763	4898	865	1947 03 14	5199	4052	1147
1946 02 09	5716	4596	1120	1947 03 15	6292	3989	2303

Table 4:

Data	NOAA №	S_{gr} , mph			Data	NOAA №	S_{gr} , mph		
		SPG	USAF _{corr}	DPD			SPG	USAF _{corr}	DPD
1982 09 13	3804	2240	228	299	1989 09 01	5670	250	206	241
1983 05 14	4173	2110	1782	1754	1989 09 07	5670	220	56	40
1983 12 05	4373	1010	28	8	1990 02 14	5927	800	504	427
1984 02 19	4421	1560	1464	1310	1990 02 15	5927	1160	564	446
1984 04 27	4474	2500	3174	3215	1990 02 16	5927	850	546	523
1988 07 18	5076	230	42	38	1990 02 22	5954	670	14	13
1988 10 10	5175	2540	1022	1043	1990 06 28	6129	170	70	59
1988 12 05	5261	290	249	193	1990 08 10	6204	680	42	36
1989 02 05	5350	1650	24	14	1990 08 18	6197	180	84	104
1989 02 09	5355	1380	252	383	1990 08 18	6224	80	50	3
1989 02 27	5378	1460	88	104	1990 08 19	6206	810	32	37
1989 03 07	5383	350	178	293	1990 08 25	6212	670	308	322
1989 06 30	5555	450	395	445	1990 09 27	6283	990	703	605
1989 07 02	5563	570	378	390	1990 10 18	6314	590	478	470
1989 06 29	5569	960	493	506	1990 11 06	6353	1060	428	20
1989 07 02	5569	520	322	427	1990 11 10	6361	870	322	349
1989 07 02	5572	740	145	186	1990 12 27	6427	290	238	228
1989 07 27	5612	660	248	234	1991 03 28	6560	120	67	89
1989 08 09	5638	110	48	58	2011 06 17	11234	100	70	200
1989 08 14	5644	210	59	57	2011 06 17	11236	260	424	522
1989 08 31	5668	270	91	146	2011 06 24	11236	180	109	267
1989 08 29	5669	2630	1403	663					