**Astronomy
&
Astrophysics**

Analysis of full-disc Ca II K spectroheliograms

III. Plage area composite series covering 1892–2019[⋆]

Theodosios Chatzistergos[1,2], Ilaria Ermolli[1], Natalie A. Krivova[2], Sami K. Solanki[2,3], Dipankar Banerjee[4,5],
Teresa Barata[6], Marcel Belik[7], Ricardo Gafeira[6,8], Adriana Garcia[6], Yoichiro Hanaoka[9], Manjunath Hegde[4],
Jan Klimeš[7], Viktor V. Korokhin[10], Ana Lourenço[6], Jean-Marie Malherbe[11,12], Gennady P. Marchenko[10],
Nuno Peixinho[6], Takashi Sakurai[9], and Andrey G. Tlatov[13]

[1] INAF Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monte Porzio Catone, Italy
e-mail: theodosios.chatzistergos@inaf.it
[2] Max Planck Institute for Solar System Research, Justus-von-Liebig-weg 3, 37077 Göttingen, Germany
e-mail: chatzistergos@mps.mpg.de
[3] School of Space Research, Kyung Hee University, Yongin, Gyeonggi 446-701, Republic of Korea
[4] Indian Institute of astrophysics, Koramangala, Bangalore 560034, India
[5] Aryabhatta Research Institute of Observational Sciences, Nainital 263 001, Uttarakhand, India
[6] Univ Coimbra, CITEUC - Center for Earth and Space Research of the University of Coimbra, Geophysical and Astronomical
Observatory, 3040-004 Coimbra, Portugal
[7] Observatory Upice, U lipek 160, 542 32 Úpice, Czech republic
[8] Instituto de Astrofísica de Andalucía (CSIC), Apartado de Correos 3004, 18080 Granada, Spain
[9] National Astronomical Observatory of Japan 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[10] Astronomical Institute of Kharkiv V.N. Karazin National University, 35 Sumskaya St., Kharkiv 61022, Ukraine
[11] LESIA, Observatoire de Paris, 92195 Meudon, France
[12] PSL Research University, Paris, France
[13] Kislovodsk Mountain Astronomical Station, Central Astronomical Observatory, Russian Academy of Sciences, Kislovodsk,
Russia

ABSTRACT

Context. Studies of long-term solar activity and variability require knowledge of the past evolution of the solar surface magnetism.
The archives of full-disc Ca II K observations that have been performed more or less regularly at various sites since 1892 can serve
as an important source of such information.
Aims. We derive the plage area evolution over the last 12 solar cycles by employing data from all Ca II K archives that are publicly
available in digital form, including several as-yet-unexplored Ca II K archives.
Methods. We analysed more than 290 000 full-disc Ca II K observations from 43 datasets spanning the period between 1892–2019.
All images were consistently processed with an automatic procedure that performs the photometric calibration (if needed) and the
limb-darkening compensation. The processing also accounts for artefacts affecting many of the images, including some very specific
artefacts, such as bright arcs found in Kyoto and Yerkes data. Our employed methods have previously been tested and evaluated on
synthetic data and found to be more accurate than other methods used in the literature to treat a subset of the data analysed here.
Results. We produced a plage area time-series from each analysed dataset. We found that the differences between the plage areas
derived from individual archives are mainly due to the differences in the central wavelength and the bandpass used to acquire the
data at the various sites. We empirically cross-calibrated and combined the results obtained from each dataset to produce a composite
series of plage areas. The 'backbone' approach was used to bridge the series together. We have also shown that the selection of the
backbone series has little effect on the final composite of the plage area. We quantified the uncertainty of determining the plage areas
with our processing due to shifts in the central wavelength and found it to be less than 0.01 in fraction of the solar disc for the average
conditions found on historical data. We also found the variable seeing conditions during the observations to slightly increase the plage
areas during the activity maxima.
Conclusions. We provide the most complete so far time series of plage areas based on corrected and calibrated historical and modern
Ca II K images. Consistent plage areas are now available on 88% of all days from 1892 onwards and on 98% from 1907 onwards.

Key words. Sun: activity – Sun: photosphere – Sun: chromosphere – Sun: faculae, plages

1. Introduction

There is a need for obtaining a better understanding of long-term solar magnetic activity, which is also important for studies of the Earth's climate (Haigh 2007; Gray et al. 2010; Ermolli et al.

2013; Solanki et al. 2013). These goals require extensive and reliable solar activity indices (e.g. Kopp et al. 2016; Yeo et al. 2017; Shapiro et al. 2017; Wu et al. 2018).

One such dataset comes from the collection of full-disc Ca II K observations. Numerous Ca II K archives, recorded and stored at various observatories all over the world, have been digitised over the past three decades. Presenting their data in either physical or digital form, such archives have been used

---

to derive information about the evolution of plage regions (e.g. Antonucci et al. 1977; Ermolli et al. 2009b; Dorotovic et al. 2010; Chatzistergos et al. 2016, 2019b; Barata et al. 2018; Tlatov & Tlatova 2019, and references therein), the solar radius variations (Meftah et al. 2018; Hiremath et al. 2020), network cell properties (e.g. Berrilli et al. 1999; Ermolli et al. 2003; Chatterjee et al. 2017; Raju 2018), photometric properties of disc features over the solar cycle (e.g. Ermolli et al. 2007, 2010), as well as for the purpose of reconstructing of irradiance variations (e.g. Ermolli et al. 2011; Chapman et al. 2012; Fontenla & Landi 2018) and studies of the relation between the Ca II K brightness and the photospheric magnetic field (e.g. Loukitcheva et al. 2009; Chatzistergos et al. 2019d, and references therein).

The results presented in the literature on the plage area evolution show considerable discrepancies (Chatzistergos 2017; Chatzistergos et al. 2019b; Ermolli et al. 2018). Indeed, a critical aspect of such studies is that the accuracy of the processing applied to the data has not been evaluated. Also, only few individual archives have been used until now, and the results from various archives have not been combined. This is largely because the techniques used for the data analysis were specifically developed for each single archive and could not be directly applied to different data. The most employed archives are those from the Arcetri (Ermolli et al. 2009a), Kodaikanal (Chatterjee et al. 2016), Mt Wilson (Lefebvre et al. 2005), and Sacramento Peak (Tlatov et al. 2009) observatories.

To overcome these limitations, in our previous paper (Chatzistergos et al. 2018b, Paper I, hereafter) we introduced a novel approach to process the historical and modern Ca II K observations, to perform their photometric calibration, to compensate for the intensity centre-to-limb variation (CLV, hereafter), and to account for various artefacts. By using synthetic data, we also showed that our method can perform the photometric calibration and account for image artefacts with higher accuracy than other methods presented in the literature. More importantly, we showed that, as long as the archives are consistent with each other, for example, they are centred at the same wavelength and employing the same bandwidth for the observations, the method can be used to derive accurate information on the evolution of plage areas without the need of any adjustments in the processing of the various archives (Chatzistergos et al. 2019b, Paper II, hereafter). In Paper II, we applied our method to 85 972 images from 9 Ca II K archives to derive plage areas and produce the first composite of plage areas over the entire 20th century. In particular, we analysed the Ca II K archives from the Arcetri, Kodaikanal (8-bit digitisation), McMath-Hulbert, Meudon, Mitaka, Mt Wilson, Rome/PSPT, Schauinsland, and Wendelstein sites. Five out of the nine analysed archives were amongst the most studied and prominent ones, while the remaining archives were from less studied data sources. There are, however, many other Ca II K archives that are available and still remain largely unexplored. These archives harbour the potential to fill gaps in the available plage series as well as to address inconsistencies among the various archives and within individual archives (e.g. change in data quality, or in the measuring instrument with time). Moreover, since the work presented in Paper II, more data from various historical and modern archives became available in digital form. In particular, historical data that have been made available in the meantime include those from the latest 16-bit digitisation of the Kodaikanal archive, Catania, Coimbra, Kenwood, Kharkiv, Kyoto, Manila, Rome, Sacramento Peak, and Yerkes observatories, as well as additional data from the Meudon and Mt Wilson archives. In this light, here we present results from the most comprehensive analysis to date

of historical and modern Ca II K observations taken between 1892 and 2019 from 43 different datasets for the purposes of producing a composite plage area series.

The paper is organised as follows. In Sect. 2, we present the data analysed in our study and the methods applied on the data. Our results for the plage areas from individual archives, as well as the composite series are presented in Sect. 3. In Sect. 4, we discuss our results. In Sect. 5, we summarise the results of our study and present our conclusions.

2. Data and methods

2.1. Ca II K observations

We analysed solar full-disc Ca II K observations from 43 datasets, which include series of photographic images and data series acquired with CCD cameras. Table 1 summarises information on the datasets analysed in our study and their main characteristics. For the sake of clarity, we also list here all the datasets that formed the basis of the analysis and the corresponding abbreviations used in the text to refer to the various series: Arcetri (Ar), Baikal (Ba), Brussels (Br), Calern (CL), Catania (CT), Coimbra (Co), Kanzelhöhe (Ka), Kenwood (Ke), Kharkiv (Kh), Kislovodsk (Ki), Kodaikanal taken with the spectroheliograph (Ko), Kodaikanal taken with the Twin telescope (KT), Kodaikanal taken with the White-Light Active Region Monitor (WARM) telescope (KW), Kyoto (Ky), Manila (Ma), Mauna Loa (ML) taken with the Precision Solar Photometric Telescope (PSPT), McMath-Hulbert (MM), Mees (MS), Meudon taken with the spectroheliograph (MD1), Meudon taken with an interference filter (MD2), Mitaka taken with the spectroheliograph (Mi1), Mitaka taken with the Solar Flare Telescope with an interference filter (Mi2), Mt Wilson (MW), Pic du Midi (PM), SOlar Diameter Imager and Surface Mapper (SODISM) telescope on board the PICARD spacecraft (PS), Rome taken with the equatorial bar at Monte Mario (Ro), Rome taken with the PSPT (RP1), Rome taken with the PSPT with narrow bandwidth (RP2), Sacramento Peak (SP), San Fernando taken with the Cartesian Full-Disk Telescope (CFDT) 1 (SF1), San Fernando taken with the CFDT2 (SF2), Schauinsland (Sc), Teide (Te) taken with the Chromospheric Telescope (ChroTel), Upice (UP), Valašské Meziříčí (VM), Wendelstein (WS), and Yerkes (YR). Figure 1 shows examples of observations from all datasets except for MM and Sc, examples of which can be found in Chatzistergos et al. (2018a, 2019b). In addition to the 38 datasets listed in Table 1, we also analysed the five datasets included in Table 2. These five datasets include observations centred at different locations of the wing of the Ca II K line.

The majority of the analysed datasets stem from observatories located in Europe. However, there are datasets from Asia and North America that provide an overall good temporal coverage. All in all, there are 290 147 images taken between 25/06/1892 and 31/12/2019 covering 41 163 days. Figure 2 shows the fraction of days within a year with at least one observation among all datasets considered in our study. We also show the total annual coverage by the sum of all the datasets analysed here, the coverage for the case when only Ko and MW are used as well as for the plage area composite presented in Paper II. Figure 2 reveals that the data analysed in our study offer a nearly complete coverage, with the exception of the period before 1925, 1944–1946, and 1986–1987. The coverage is on average 88% for the whole period of time since 1892. However, it is on average 98% and is above 76% for all years if only the period after 1907 is considered. In contrast, the annual coverage when only

Table 1. Ca II K datasets analysed in this study.

Observatory	Acronym	Detector	Instrument	Period	Images	SW [Å]	Pixel scale [″/pixel]	Ref.
Arcetri[a]	Ar	Plate	SHG	1931–1974	4871	0.3	2.5	1
Baikal	Ba	CCD	Filter	2004–2019	846	1.2	2.7	2
Big Bear	BB	CCD	Filter	1982–2006	5027	3.2, 1.5[b]	4.2, 2.4[c]	3
Brussels	Br	CCD	Filter	2012–2019	14699	2.7	1.0	4
Calern	CL	CCD	Filter	2011–2019	1560	7	1.0	5
Catania	CT	Plate	SHG	1908–1977	1008	–	1.1–5	6
Coimbra	Co	Plate/CCD[d]	SHG	1925–2019	19758	0.16	2.2	7
Kanzelhöhe	Ka	CCD	Filter	2010–2019	8550	3.0	1.0	8
Kenwood	Ke	Plate	SHG	1892	5	–	3.1	9
Kharkiv	Kh	Plate/CCD[e]	SHG	1952–2019	564	3.0	3.3	10
Kislovodsk	Ki	Plate/CCD[f]	SHG	1960–2019	9738	–	1.3, 2.3[f]	11
Kodaikanal[a]	Ko	Plate	SHG	1904–2007	45047	0.5	0.9	12
Kodaikanal Twin	KT	CCD	Filter	2008–2013	3059	1.2	1.2	13
Kodaikanal WARM	KW	CCD	Filter	2017–2019	585	1.0	2.4	14
Kyoto	Ky	Plate	SHG	1928–1969	3119	0.74	2.0	15
Manila	Ma	Plate	SHG	1968–1978	162	0.5	1.2	16
Mauna Loa PSPT	ML	CCD	Filter	1998–2015	31933	2.7	1.0	17
McMath-Hulbert[a]	MM	Plate	SHG	1948–1979	4932	0.1	3.1	18
Mees	MS	CCD	Filter	1982–1998	1519	1.2	5.5	19
Meudon[a]	MD1	Plate/CCD[g]	SHG	1893–2019	20117	0.15, 0.09[h]	2.2, 1.5, 1.1[i]	20
Meudon	MD2	CCD	Filter	2007–2014	1519	1.4	0.9	21
Mitaka[a]	Mi1	Plate	SHG	1917–1974	4193	0.5	0.9, 0.7[j]	22
Mitaka	Mi2	CCD	Filter	2015–2019	897	4.5	1.0	23
Mount Wilson[a]	MW	Plate	SHG	1915–1985	39545	0.2	2.9	24
Pic du Midi	PM	CCD	Filter	2007–2019	3794	2.5	1.2	25
PICARD/SODISM	PS	CCD	Filter	2010–2014	1218	7	1	26
Rome Monte Mario	Ro	Plate	Filter	1964–1979	5826	0.3	5.0	27
Rome PSPT[a]	RP1	CCD	Filter	1996–2019	3449	2.5	2.0[k]	28
Rome PSPT	RP2	CCD	Filter	2008–2019	1298	1.0	2.0[k]	28
Sacramento Peak	SP	Plate	SHG	1960–2002	7750	0.5	1.2	29
San Fernando CFDT1	SF1	CCD	Filter	1988–2015	4986	9	5.1	30
San Fernando CFDT2	SF2	CCD	Filter	1992–2013	4065	9	2.6	30
Schauinsland[a]	Sc	Plate	SHG	1958–1965	18	–	1.7, 2.6[l]	31
Teide ChroTel	Te	CCD	Filter	2009–2019	1843	0.3	1.0	32
Upice	UP	CCD	Filter	1998–2019	3234	1.6	4.0, 2.4[m]	33
Valašské Meziřičí	VM	CCD	Filter	2011–2018	318	2.4	1.8	34
Wendelstein[a]	WS	Plate	SHG	1947–1977	422	–	1.7, 2.6[l]	31
Yerkes	YR	Plate	SHG	1903–1907	7	–	2.4	35

Notes. Columns are: name of the observatory, abbreviation used in this study, type of detector, type of instrument, period of observations, total number of images (including multiple images on a single day when available) analysed in this study, spectral width of the spectrograph/filter, average pixel scale of the images, and the bibliography entry. [a]These archives were considered in Paper II, although in the case of the Ko data the earlier 8-bit digitisation was used. [b]The two values correspond to the period before and after 10/09/1996. [c]The two values correspond to the period before and after 08/11/1995, when the CCD camera was upgraded. [d]The CCD camera was installed in January 2007, but observations were stored on photographic plates up to December 2007. [e]The CCD camera was installed on 01/09/1994. [f]The CCD camera was installed on 18/12/2002. [g]The observations were stored on photographic plates up to 27/09/2002, while observations with a CCD camera started on 13/05/2002. [h]The values refer to the periods before and after 15/06/2017. [i]The values refer to the periods [24/10/1893–27/09/2002], [28/09/2002–14/06/2017], and since 15/06/2017. [j]These data derive from two digitisations and the two values correspond to the earlier and more recent digitisation, respectively. See Chatzistergos et al. (2019b) for more information. [k]The pixel scale is for the resized images to match the seeing conditions of the observing location. [l]These data were stored in TIFF and JPG files with different spatial resolution, the values correspond to the TIFF and JPG files, respectively. [m]The two values correspond to the period before and after 01/01/2018, when the CCD camera was upgraded.

References. (1) Ermolli et al. (2009a); (2) Golovko et al. (2002); (3) Naqvi et al. (2010); (4) http://www.sidc.be/uset/; (5) Meftah et al. (2018); (6) Zuccarello et al. (2011); (7) Garcia et al. (2011); (8) Hirtenfellner-Polanec et al. (2011); (9) Hale (1893); (10) Belkina et al. (1996); (11) Tlatov et al. (2015); (12) Priyal et al. (2014); (13) Singh & Ravindra (2012); (14) Pruthvi & Ramesh (2015); (15) Kitai et al. (2013); (16) Miller (1965); (17) Rast et al. (2008); (18) Mohler & Dodson (1968); (19) http://kopiko.ifa.hawaii.edu/; (20) Malherbe & Dalmasse (2019); (21) http://bass2000.obspm.fr/data_guide.php; (22) Hanaoka (2013); (23) Hanaoka & Solar Observatory of NAOJ (2016); (24) Lefebvre et al. (2005); (25) Koechlin et al. (2019); (26) Meftah et al. (2014); (27) Chatzistergos et al. (2019a); (28) Ermolli et al. (2007); (29) Tlatov et al. (2009); (30) Chapman et al. (1997); (31) Wöhl (2005); (32) Bethge et al. (2011); (33) Klimeš et al. (1999); (34) Lenza et al. (2014); (35) Hale & Ellerman (1903).
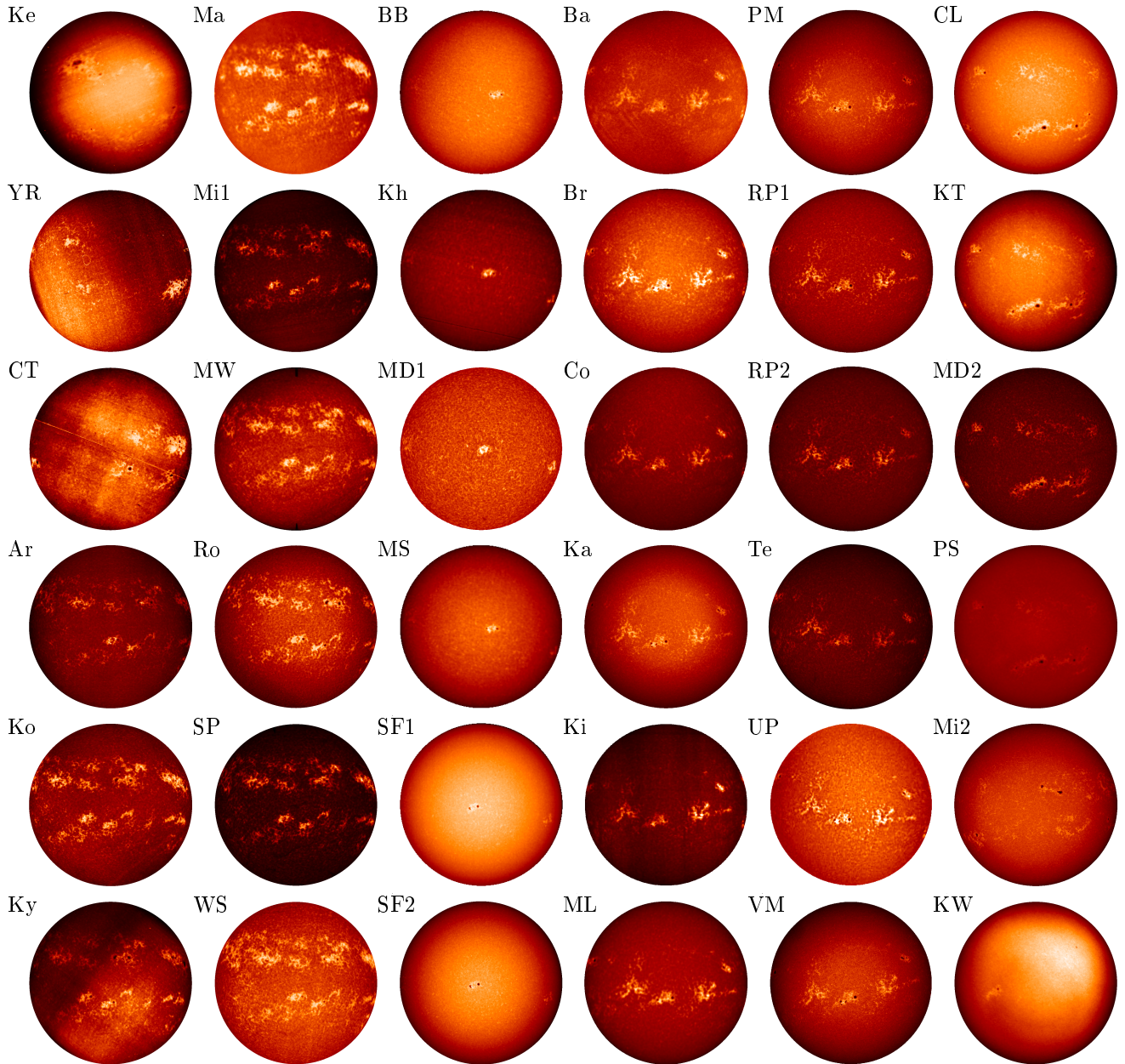
Fig. 1. Examples of observations from the various archives analysed in this study, with the exception of MM and Sc. Aside from the images in the *first and last column*, the images within each column correspond to roughly the same day. Within a column, the images are shown in alphabetical order according to the name of the observatory, given by a 2-letter abbreviation, with a numeral added in some cases (see Table 1 for the corresponding observatory name). The specific dates of the observations are: 14/07/1892 for Ke; 16/02/1907 for YR; 04/09/1908 for CT; 03/02/1968 for Ma; 04/02/1968 for Ko, Ky, Mi1, MW, and SP; 05/02/1968 for Ar, Ro, and WS; 16/07/1995 for BB, Kh, MD1, and MS; 15/07/1995 for SF1 and SF2; 13/03/2014 for Ba, Br, Co, Ka, Ki, ML, PM, RP1, RP2, Te, UP, and VM; 01/08/2012 for CL, KT, MD2, and PS; 10/07/2015 for Mi2; 23/04/2018 for KW; respectively. The images are shown after pre-processing to identify the disc and re-sample them to account for the disc's ellipticity (when applicable) and convert the historical data to density values. The images have been roughly aligned to show the Solar north pole at the top.

Ko and MW are considered is on average 80%, while it drops down to 5% in the 1990's. The composite provides full annual coverage since 2010 and for 21 more years. In contrast, the Ko and MW series together provide a full coverage only for 1967. This illustrates the substantial benefit of using multiple datasets to achieve a better coverage over the entire 20th century, but also the need to recover more historical data. The missing data from Abastumani (Khetsuriani 1981), Anacapri (Antonucci et al. 1977), Baikal, Cambridge (Moss 1942), Catania, Crimea, Ebro (Curto et al. 2016), Huairu (Suo 2020), Kandilli (Dizer 1968),

Kenwood, Kharkiv, Kislovodsk, Locarno (Waldmeier 1968), Madrid (Vaquero et al. 2007), Manila, Meudon, Yerkes, Wendelstein, and Schauinsland would be invaluable for this purpose.

Tables 1 and 2 illustrate the diversity of the analysed data in terms of, for example, instruments, bandwidth used, central wavelength, and the resulting pixel scale of the images. The observations from 14 of these datasets were stored on photographic plates (we will refer to all physical photographs as plates, even though celluloid film was used by some archives), a CCD camera was exclusively used for 25 datasets, while

Table 2. Off-band Ca II K datasets analysed in this study.

Observatory	Acronym	Instrument	Central wavelength [Å]	Period	Images	SW [Å]
Coimbra[a]	CoW	SHG	3932.3	2008–2018	3113	0.16
Mauna Loa PSPT	MLW	Filter	3936.3	2004–2015	9552	1.0
Meudon	MDV	SHG	3933.4	2002–2017	5717	0.15
Meudon	MDR	SHG	3934.0	2002–2017	5652	0.15
Meudon[a]	MDW	SHG	3932.3	2002–2018	4632	0.15

Notes. Columns are: name of the observatory, abbreviation used in this study, type of instrument, central wavelength, period of observations, number of images, and the spectral width of the spectrograph/filter. [a]Here we restricted our analysis only to the CCD-based data centred at the wing of the line from the CoW and MDW series. Note, however, that off-band data from these sources extend back to 1925 and 1893, respectively.
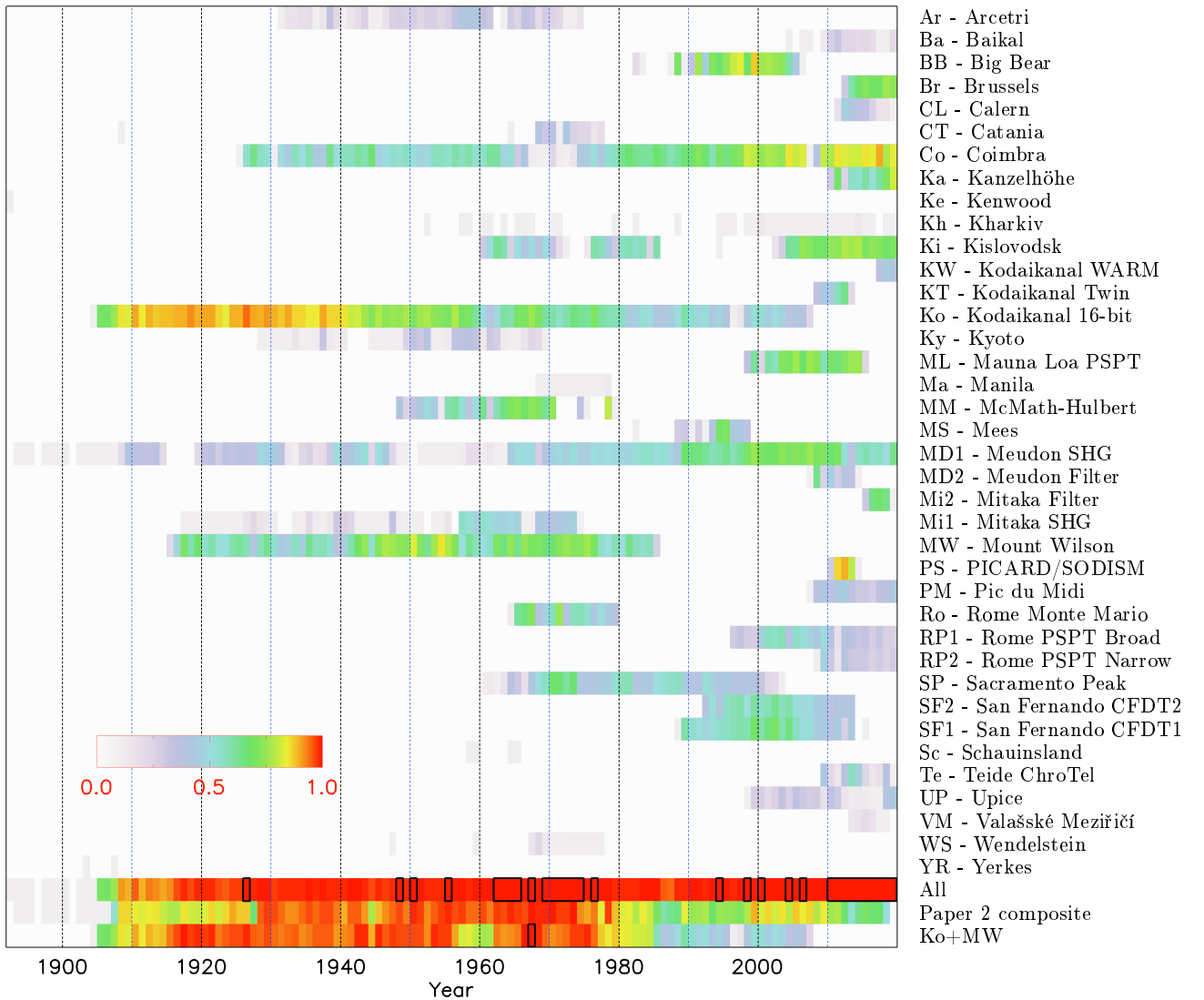
Fig. 2. Annual coverage of the various Ca II K archives analysed in this study, except from those taken off-band (see Sect. 2 for details). Also shown is the annual coverage of all the archives combined, the one for the Chatzistergos et al. (2019b, Paper II) composite series, and the annual coverage when only Ko and MW are considered. The annual coverage is colour-coded as shown by the colour bar plotted in the bottom left corner of the panel. Years with full coverage are outlined with black rectangles.

four datasets include data taken with a CCD camera as well as stored on photographic plates. We note that a few datasets include images obtained at the same observatory, but with different telescopes or instruments, for example, with Kodaikanal and Kodaikanal Twin or Meudon spectroheliograms and Meudon fil-

tergrams. However, we do not make a distinction with regard to the medium used to store the original data, whether it was a CCD camera or photographic plates. We note that this categorisation is merely to simplify the discussions in this manuscript. The arrangement of the datasets for the calibration procedure
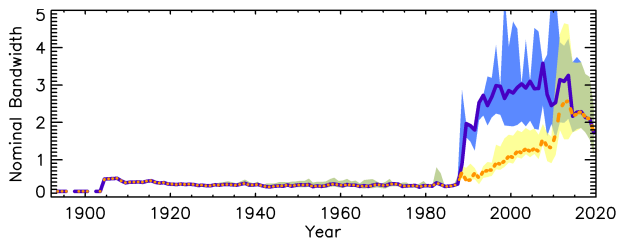
Fig. 3. Average nominal bandwidth from all 38 archives included in this study (solid purple) as well as from all archives excluding SF1 and SF2 (dashed orange). The shaded areas show the 1σ uncertainty.

to produce the plage area composite is different and is outlined in Appendix B. The majority of the data stored on photographic plates were acquired with a spectroheliograph, with only Ro including images taken with an interference filter. In contrast to that, most data taken with a CCD camera were obtained with interference filters, with only Co, MD1, Kh, and Ki images resulting from a spectroheliograph. These four datasets include data taken with a CCD but also stored on photographic plates. We point also that Meudon data over 2002–2017 are provided in data cubes with observations taken at the core of the Ca II K line as well as centred at four different wavelengths away from the core. In our derivation of the plage area series, we included only the observations taken at the core of the line (referred to as MD1) and the images taken at two extreme offsets (MDV and MDR, hereafter, for the data centred at the violet and red wing of the line). To discuss our results, we also analysed Meudon and Coimbra observations centred beyond the K1 violet wing of the line (MDW and CoW, respectively), as well as Mauna Loa data centred at the red wing of the line (MLW, hereafter). All the off-band observations are summarised in Table 2.

The bandwidth of the analysed observations ranges between 0.1 and 9 Å, thus sampling substantially different heights in the solar atmosphere. These differences can be seen in Fig. 1, where the images with relatively narrow bandwidth appear to have higher contrast in the plage regions, while the CLV is reduced compared to those with broader bandwidths. In particular, the observations from MD1 and SF1 are quite indicative, having been taken on the same day with nominal bandwidth of 0.15 and 9 Å, respectively. Consequently, the network regions are enhanced in the MD1 image, while they are barely discernible in the one from SF1. Conversely, sunspots are clearly visible in the SF1 image, but barely hinted at the MD1 one. Comparing RP1 and RP2, taken with a bandwidth of 2.5 and 1 Å, respectively, the sunspots are only minutely reduced in size in RP2, while the CLV has also been reduced.

It is noteworthy that modern datasets include observations taken in general with broader bandwidths than the ones used for the historical data. Figure 3 shows the annual mean value of the nominal bandwidth from all datasets included in this study for which we have information on the bandwidth and for which the observations were centred at the core of the line. It is interesting that there is a slight decrease in the mean bandwidth between 1905 and 1920 while it remains roughly constant at around 0.3 Å up to the late 1980's. The variations over that period are rather low, with average bandwidths being in the range 0.2–0.5 Å. However, the variations are more extreme since 1987, with a range of average bandwidths between 0.4 and 3.6 Å. To some degree this increase is due to the broad bandwidth of SF1 and SF2 data, but not entirely. Figure 3 also shows the mean nominal bandwidth from all archives excluding SF1 and SF2. In

this case there is still some, though weaker, increase of ~0.3 Å in 1988 and an almost steady increase in the mean bandwidth after that.

Since the data availability during the mid 1980's is poorer than over other periods and the bandwidths change significantly, the uncertainty of cross-calibrating results from Ca II K data over that period is also higher. However, we note that the central wavelength also affects the brightness of the magnetic features in the images and consequently their disc coverage. We also note that some datasets appear to include observations taken slightly offset from the core of the Ca II K line. This affects, for instance, the observations from ML, which have a central wavelength of 3934.15 Å instead of 3933.67 Å. Unfortunately, the precise values of the central wavelength used for the observations from most archives are not available, so we were not able to show its change over time. Furthermore, the observations from Ka, PM, and VM have almost the same nominal bandwidth as RP1 (3.0, 2.5, 2.4, and 2.5 Å, respectively), however, the CLV is stronger in the Ka, PM, and VM observations compared to the RP1 ones, hinting that Ka, PM, and VM observations might have been taken outside of the line centre or that the actual bandwidth is broader. A similar evaluation is more difficult for the historical data, which suffer from more artefacts than the modern data. We only mention here that SP observations, taken with a nominal bandwidth of 0.5 Å, have the lowest CLV among all datasets analysed in our work. This suggests that the actual bandwidth used at SP might be narrower than the nominal value.

Furthermore, we note that in addition to a bandwidth different from the nominal one, there are other parameters of the observation that can affect the data. Indeed, parameters such as stray light contribution, blurring due to atmosphere seeing, over- or underexposure of photographic plates, vignetting, instrument- and setup-specific filter transmission profiles, potential contamination from secondary lobes in the filter transmission profile may have affected the CLV. We point out that also modern data are affected by problems. For instance, a few ML images taken with a CCD were found to be saturated. Finally, observations from various sites have been copied and shared with other observatories. Therefore, it could also happen that images have been scanned and erroneously attributed to a wrong observatory.

It is also worth mentioning that some of the datasets listed in Table 1 underwent multiple digitisations, for example, those from the Ko, Mi1, MW, and SP observatories, which were digitised with 8 and 16-bit devices. Compared to the analysis by Paper II, in this study we included 34 more datasets, as well as a new version of the Ko, MD1, and MW datasets. In particular, we used the Ko data from the more recent 16-bit digitisation, MD1 data over the period 1939–1948 and 1964–1967, which had not been available before, along with recently recovered data from the MW dataset. The MW dataset stems from the 16-bit digitisation by Lefebvre et al. (2005), which is the same as the one used in Paper II. The complete original dataset was, however, considered lost due to a failure of the storage hardware. Luckily, it was recently recovered, and we found that it includes 3463 images which were missing from the dataset considered in Paper II. However, 164 images from the dataset that were considered in Paper II are still missing in the new series. In this study, we analysed the recently recovered series of 16-bit MW data, but included the missing 164 images from the analysis by Paper II. Furthermore, observations from CT, Ko, Ma, Ro, and WS were found in 35 mm celluloid films, which were distributed as the Photographic journal of the Sun. These were produced by the observatory of Rome over the period 1967–1978 as

supplementary material to their monthly bulletins. We digitised, with 8-bit accuracy, all the observations missing from our collection with the reflecta RPS 10M commercial film scanner, which is the same scanner previously used for the Ro observations (Chatzistergos et al. 2019a). The datasets from CT, Ke, Kh, Ki, Ma, MD1, Sc, WS, and YR have only been partially digitised. Considering the large gaps in observations in the BB and MS datasets, it is possible that more data were taken over the 1980's, which, however, we were so far unable to recover.

Finally, although the datasets from Ba, CL, Ka, PS, Te, and VM have multiple observations per day, for our analysis we used either the 'best' observation of the day, as selected by the observers of the BK and VM datasets, or for the CL, Ka, PS, and Te datasets, we used an automatic process to randomly select between one and three images per day which were unaffected by cloud coverage. We did make one exception for the first seven days in June 2014. For this period, we analysed all available observations from all datasets to study the sensitivity of our results to daily variations in seeing (see Sect. 4 for more details). The period was chosen randomly, with the only requirements being that it ought to be an active period and in the summer to ensure improved seeing conditions. In this way, we provide a lower estimate of the uncertainty in the derived plage areas due to seeing variations. Within these seven days there are 36, 62, 89, 1640, 79, and 561 images in the Ba, Br, CL, Ka, ML, and Te datasets, respectively.

2.2. Methods

We consistently processed all images with the methods described by Chatzistergos et al. (2018b, 2019b, 2020). To describe our actions in brief, we started by identifying the solar disc to extract the information on the coordinates of the disc centre and the radius (Chatzistergos et al. 2019b, 2020). We applied the calibration for the digitisation device, where the relevant data were available. In particular, the Ko, KW, and MS datasets include the appropriate information to make it possible to perform the calibration of the CCD employed for the digitisation and observation, respectively. The CCD calibration has also been applied to BB (only over the period between 07 July 2000 and 21 September 2006), Co, KT, MD2, Mi2, RP1, and RP2 data, but it has not been applied, or it is unclear whether it has been applied, to the images from the other datasets. However, the flat-field images taken at KW were found to exhibit large saturated areas, hence, we decided not to use those flat-field files. Here, we stress that the image calibration of the CCD recording device improves the accuracy of the analysed images and allows for removal of artefacts due to the device and its use, for example, the dust on the detector. Figure 4 shows examples of the calibration of the CCD recording device for Ko, MS, and RP1 images. All the calibration data shown here exhibit intensity variations across the different quadrants of the CCD. Additionally, numerous dark small-scale round artefacts are evident in the MS observation. The flat image for the Ko observation also shows some scratch-like patterns. Such artefacts are accounted for by using the calibration images, thus reducing the uncertainties of analysing these data. Besides, for all datasets we applied a data selection (Chatzistergos et al. 2019b) merely to exclude pathological cases characterised by severely distorted discs, missing parts of the disc, or strong artefacts over the disc. Moreover, the solar disc in BB, Co, Kh, Ki, MD1, MW, SP, and Te images was re-sampled to account for its ellipticity following Chatzistergos et al. (2020).

Then the images from the datasets with data stored on photographic plates were photometrically calibrated (Chatzistergos
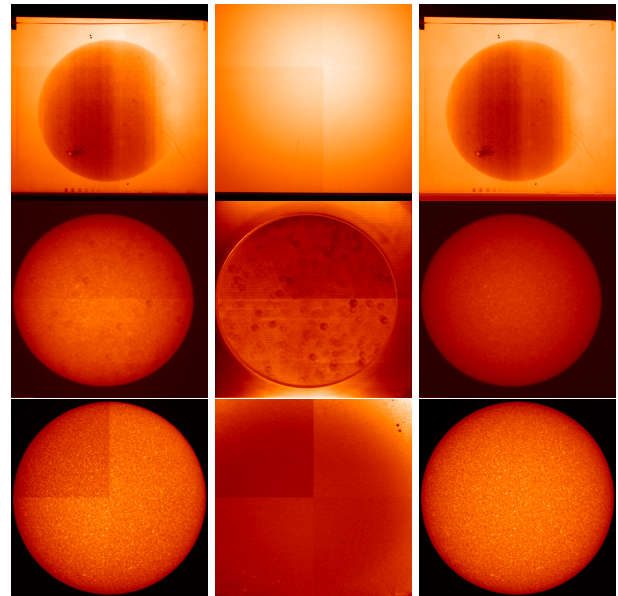
Fig. 4. Examples of raw images (*left column*), data employed for the calibration of the CCD recording device (*middle column*), and calibrated images (*right panels*) from the Ko (*first row*), MS (*middle row*), and RP1 (*bottom row*) archives taken on 19/08/2006, 19/12/1995, and 28/02/2020, respectively. The images have not been compensated for ephemeris and are shown in their entire range of values.

et al. 2018b) to account for the non-linear response of the photographic material. All images, historical and modern, were compensated for the limb-darkening as described in Paper I. In particular, we define a contrast image as $C_i = (I_i - I_i^{QS})/I_i^{QS}$, where C_i, I_i are the contrast and intensity values of the calibrated image at pixel i, while I_i^{QS} is the intensity of the quiet Sun (QS, hereafter) at the pixel i. Ky and YR data suffer from a very specific artefact manifesting itself as bright or dark arcs on the solar disc. Therefore, to improve the accuracy of the analysis of the Ky and YR data we added one further processing step. The process applied on these data is described in detail in Appendix A. Figure 5 displays examples of the calibrated and limb-darkening compensated contrast images after the preprocessing to correct for the elliptical discs (where needed) and to convert the historical images to density values for the same observations as shown in Fig. 1. More details on: the processing of the Ar, MM, MD1, Mi1, MW, RP1, Sc, and WS datasets can be found in Paper II; on processing the CT and Ro datasets in Chatzistergos et al. (2019a); Ko in Chatzistergos et al. (2019c); and Ky and SP in Chatzistergos et al. (2020).

All the processed images were segmented to identify plage areas with a multiplicative factor, $m_p = 8.5$, to the standard deviation of the QS regions (Chatzistergos et al. 2019b). The segmentation was applied consistently with the same multiplicative factor to all the datasets. The observation time for all archives was converted to Universal standard time (UT). Figure 6 shows the corresponding segmentation masks of the observations shown in Fig. 1, singling out the plage regions, which are the solar features that are mainly considered in this study.

3. Analysis of the Ca II K series

3.1. Individual Ca II K series

Figure 7 shows the plage areas series derived from each dataset that we analysed. For clarity, we split the results into three
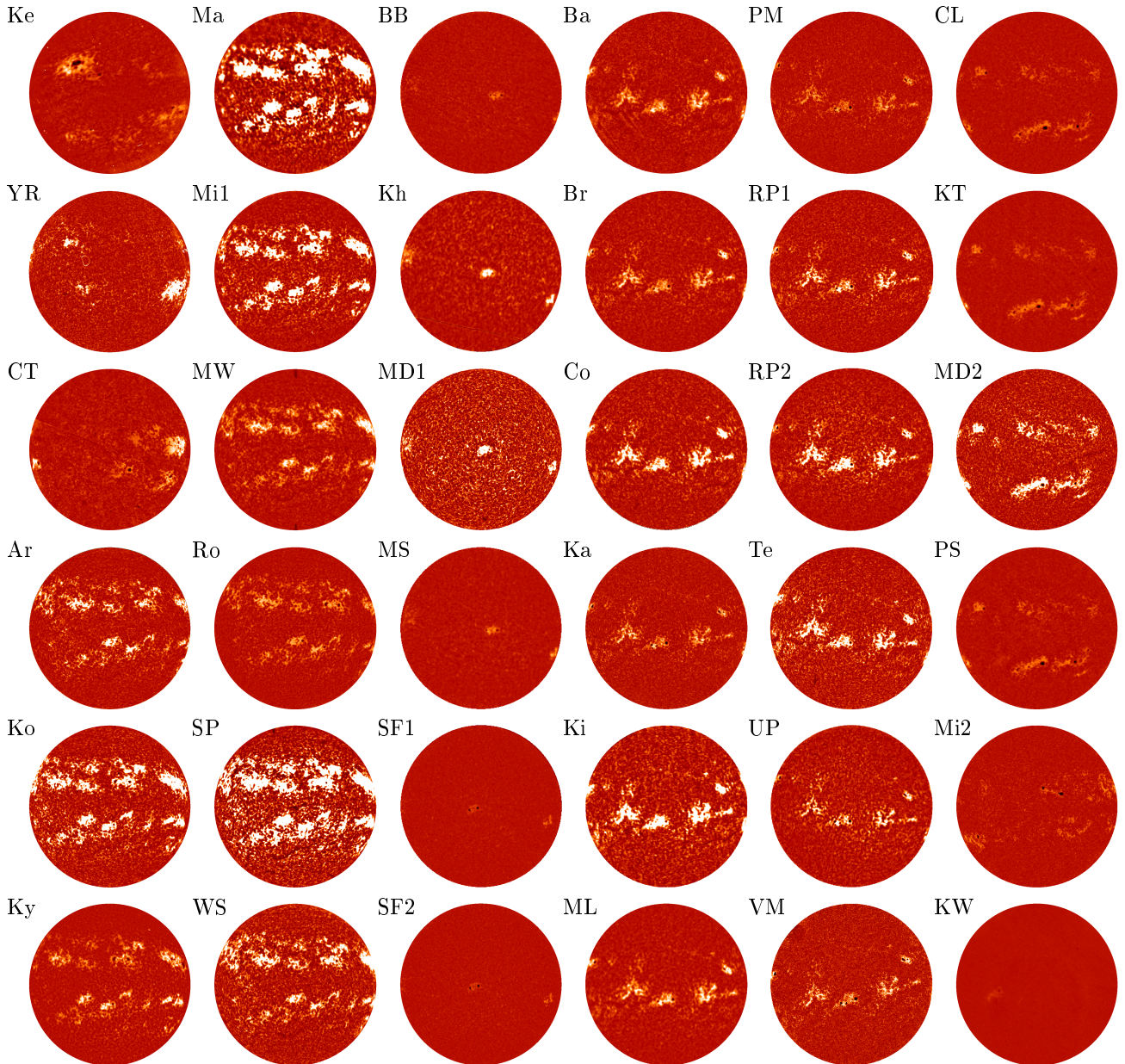
Fig. 5. Calibrated and CLV-compensated contrast images of the observations shown in Fig. 1. Plotted are contrast values in the fixed range of [−0.5,0.5] for all images.

panels, each one showing periods of roughly four solar cycles (SC). The temporal profile of the evolution of plage areas is, in general, similar. Thus, we can recognise the same features in all datasets during various periods; for example, the period 1971–1975 or the double peaks of SC 22 maximum over the period 1989–1992. However, there are also some obvious differences.

For instance, the plage areas from Ka, KW, Mi2, PM, UP, and VM are considerably lower than those of all other datasets. This is in agreement with the comment in Sect. 2 that these data were probably taken off-band or with a broader bandwidth. The plage areas from Co are considerably lower than from other datasets over SC 20. However, the annual values from Co data are not representative over that period due to the low number of Co images because of the relocation of the observatory (Lourenço et al. 2019). We also note that the Co plage areas over SC 19 are lower than SC 17 and 18, hinting at a potential issue with the data over SC 19. Plage areas derived from Kh images are greater than

those from Ko data over SC 23. This is contrary to the expectation considering that the bandwidth used at Kh is double the one used at Ko. Given that both observatories employed spectroheliographs[1], this lends support to our suggestion that the actual bandwidth of the Ko observations is broader than reported or there is an offset in the central wavelength (Chatzistergos et al. 2019b,c).

We also compare the SF1 and SF2 series. These data have the same nominal observational characteristics except for the spatial resolution, which in SF1 data, is half of that in SF2 data. We find a linear correlation factor of 0.9 and RMS differences of 0.005 between the determined plage areas of the two series when considering only the 3821 days for which observations with both telescopes exist. These differences are at least partly due to the

---

[1] We remind that Kh used a CCD camera after 1994, while Ko used only photographic plates.
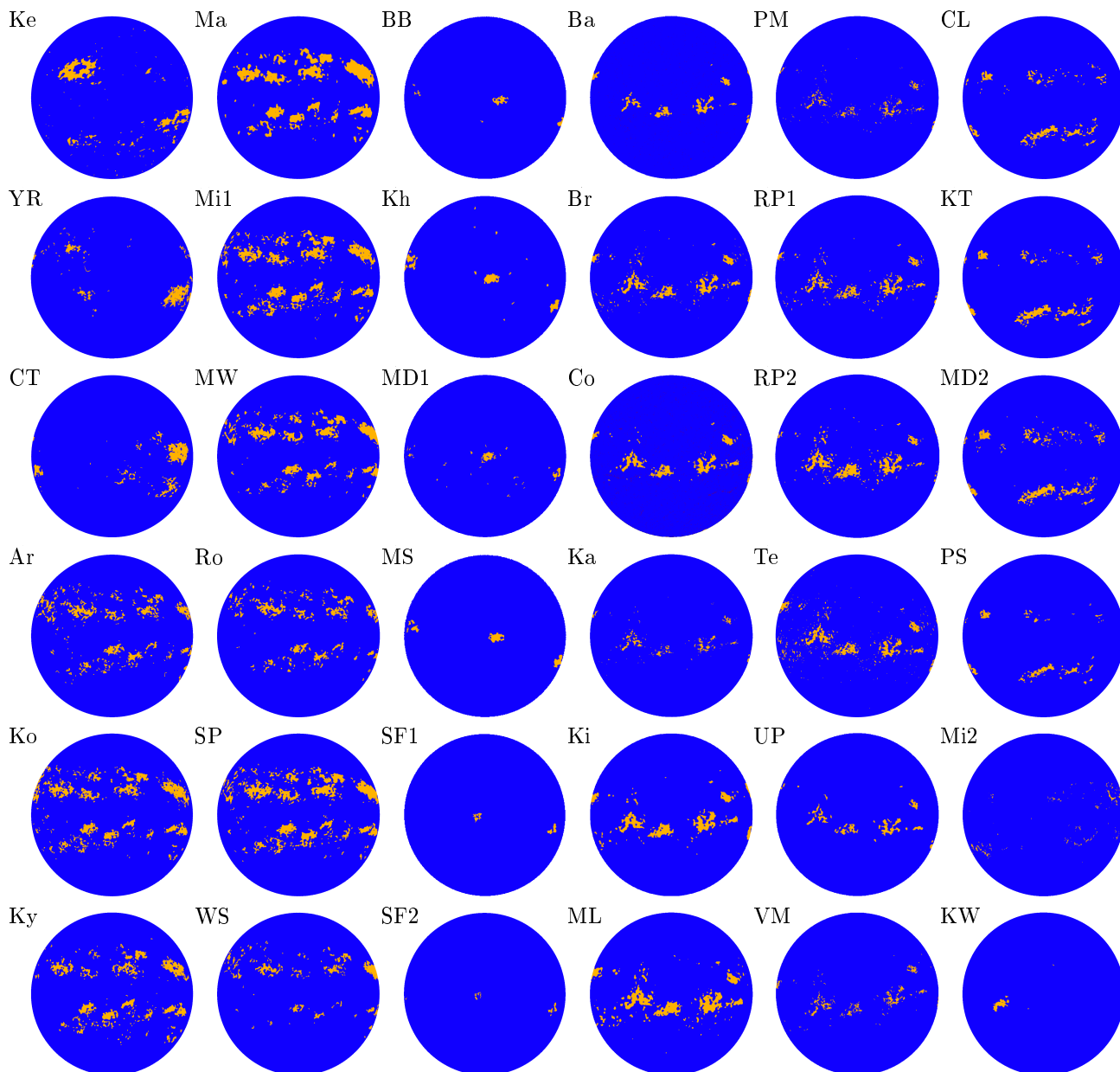
Fig. 6. Segmentation masks derived from the observations shown in Figs. 1 and 5. Plage are highlighted in yellow, while quiet Sun and network regions together form the blue background. We note that the same threshold was used for all datasets to identify the plage regions, which is why the different datasets seem to give rather different plage coverage depending on the employed bandwidth or central wavelength.

lower resolution of SF1 data compared to the SF2 ones, which results in some smearing of the features. However, this discrepancy might also be due to potential issues with SF1 data during 1997–2001. This is evidenced by a sharp increase of plage areas over 1997 and a decrease between 2000–2001, around the activity maximum of SC 23. When excluding the data between July 1997 and December 2001 we find a linear correlation factor of 0.95 and RMS differences of 0.003.

3.2. Plage areas composite

In Paper II, we presented a plage area composite derived from the analysis of nine Ca II K archives. The composite of plage areas was the average series of those obtained from using the results from the eight-bit Ko and 16-bit MW series as the references. In this study, we present a plage area composite

obtained on the basis of the results from 38 datasets. These include the data from the 16-bit digitisation of the Ko dataset (Chatzistergos et al. 2019c) as well as the recently recovered data from the 16-bit digitisation of MW. Furthermore, we use a different methodology to create our plage area composite series, employing the "backbone" approach (Svalgaard & Schatten 2016; Chatzistergos et al. 2017). In particular, we split the datasets into two backbones, roughly representing the historical and modern data separately. We took the RP1 series as the reference for the modern data backbone, while we used the Ko and MW series as references for the historical data. Appendix B describes the assignment of the various analysed series to the backbones.

Following Paper II, we cross-calibrated the individual series to the ones entering the backbone by using the daily statistics of the determined plage areas. In particular, we started
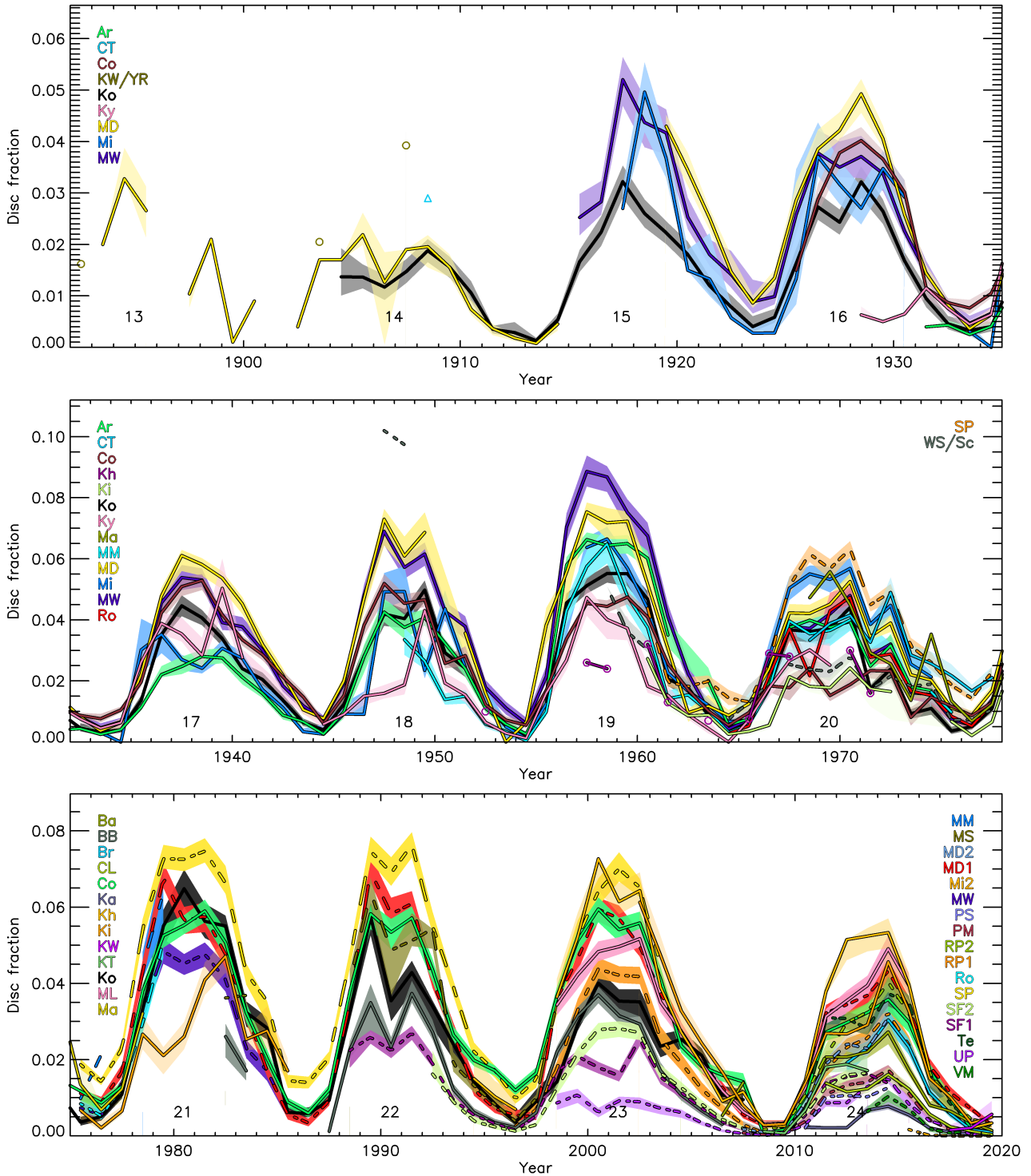
Fig. 7. Evolution of plage areas given as fraction of the solar disc derived from analysis of the 38 datasets considered in this study. Shown are annual median values (lines) along with the asymmetric 1σ interval (shaded surfaces) for each dataset as specified in the legend. To improve visibility, the archives listed in the legend in the right side of the plot are represented by dashed lines. Due to the scarcity of observations, the plage areas derived from the Ke/YR and CT observations in the top panel, as well as the Kh in the middle panel are represented by circles, triangles, and circles, respectively. The conventional solar cycle (SC) numbers are given below the curves.

by computing daily mean plage areas for all series. Then, we constructed a probability distribution function (PDF, hereafter) matrix between each individual series and the corresponding backbone one. To create this matrix we first identified the days for which both series have a plage area measurement. Then we separated these days into arrays for which the secondary series reported plage areas in bins of 0.001 in disc fraction. For each of these arrays, we computed the histogram of the reported plage areas from the backbone series in bins of 0.001 in disc fraction. For each of those selected arrays, we normalised the histogram with the total number of data within that array, thus creating a PDF. See Chatzistergos et al. (2017) or Paper II for
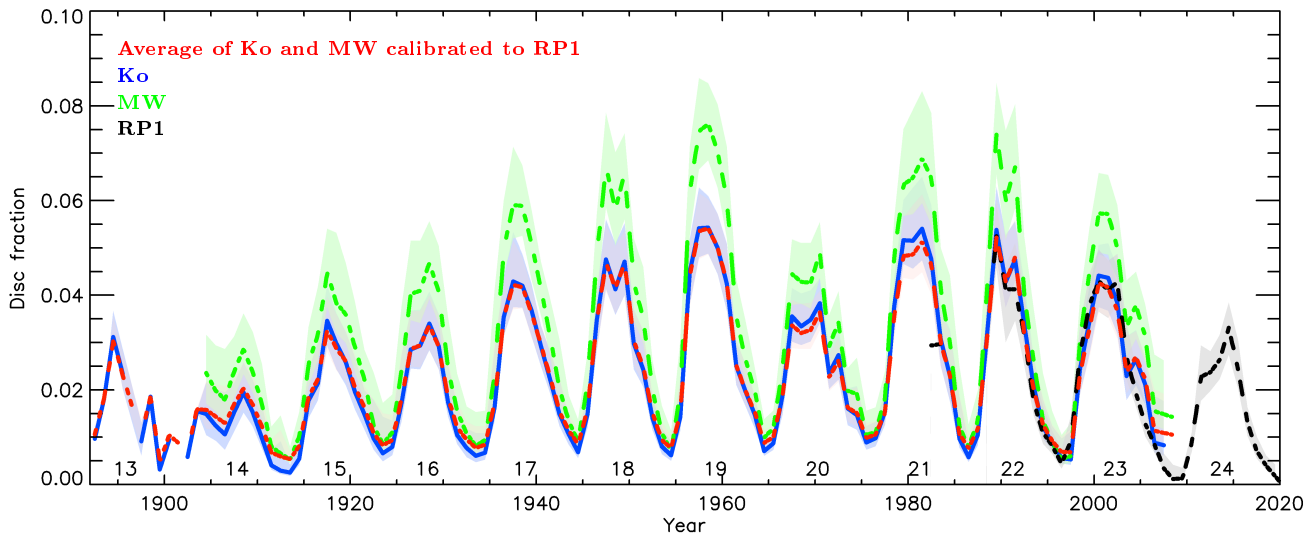
Fig. 8. Backbone series of plage areas by using as reference the series from RP1 (dashed black), Ko (solid blue), MW (dashed green), and the average backbone of the MW and Ko series after their cross-calibration to the RP1 one (dashed red). Shown are annual median values (lines) along with the asymmetric 1σ interval (shaded surfaces). The solar cycle numbers are given below the curves.
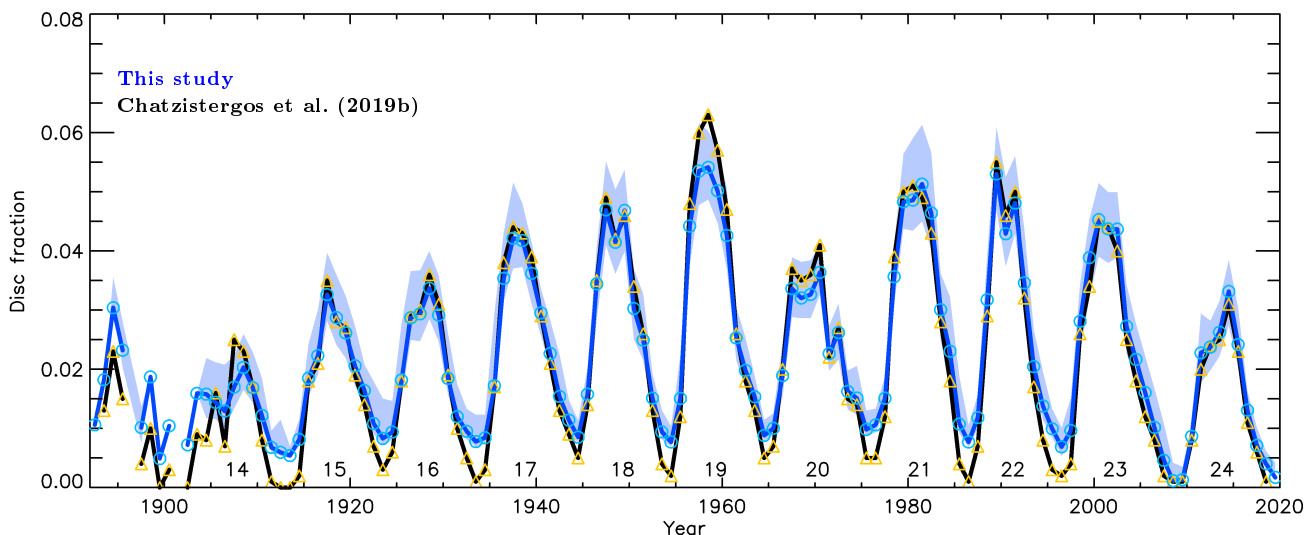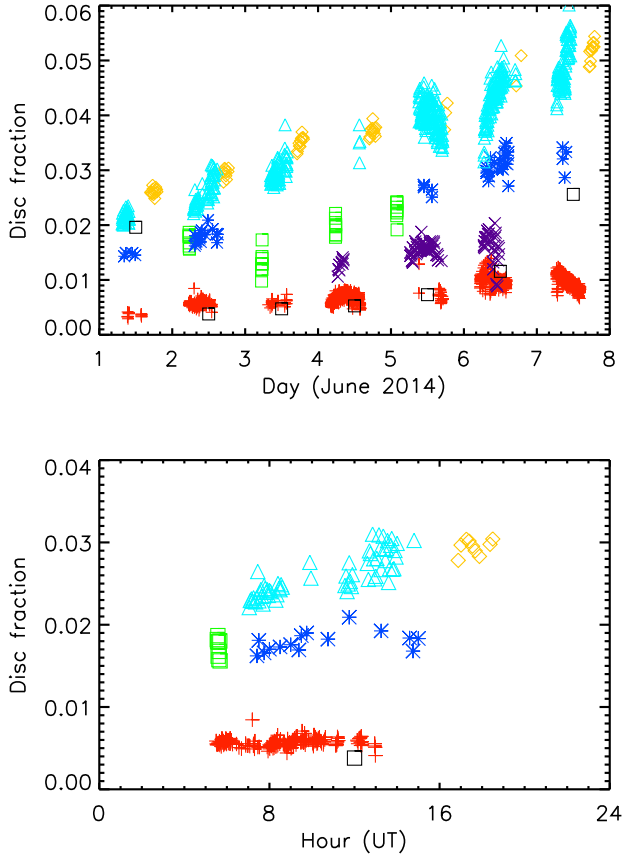

Fig. 9. Composite of plage areas in disc fraction derived in this study (blue line and ciel circles) along with the composite by Chatzistergos et al. (2019b, black line and yellow triangles). Shown are annual median values (solid lines and symbols) along with the asymmetric 1σ interval (shaded surfaces). The solar cycle numbers are given below the curves.

more details on this process. We computed the mean value of each PDF along with the asymmetric 1σ interval, which we used to perform a weighted fit of a power law function with an offset (Chatzistergos et al. 2019b) and a linear fit. These relationships were used to scale the plage areas of the secondary series to the level of the backbone one. Following Paper II, we used the linear relation for the Ko and RP1 backbones, while we used the power law for the MW ones. All calibrated series were appended to the backbone one to create the backbone composite series. This way we construct one backbone for the modern data and two for the historical ones. We cross-calibrated the two historical backbone series and the RP1 backbone series in the same way as done for the individual series. Then the two historical backbone series were averaged to create the average historical backbone series.

Figure 8 shows the RP1, Ko, MW, and the average historical backbone. We note that the Ko and RP1 backbones are at similar levels over SCs 22 and 23, with the latter being slightly lower. The variation of the plage areas in the MW backbone has

higher amplitude than in the other backbone series by ~0.15 in disc fraction. The average historical backbone (average backbone series of Ko and MW after their cross-calibration to RP1) appears very similar to the raw Ko backbone, with most SCs after SC 20 being slightly reduced in amplitude. We note that over SC 23 the plage areas in all backbone series except the RP1 exhibit a secondary peak around 2004. This is attributed to the Ko data over that period as was also mentioned by Chatzistergos et al. (2019c).

Figure 9 shows the final composite[2] produced by merging the RP1 and the average historical backbones as well as the plage area composite obtained in Paper II for comparison. The two composites agree on the absolute level, which is expected since in this study we used RP1 as the reference, while the composite presented in Paper II had a scaling factor of 1 for RP1. However,
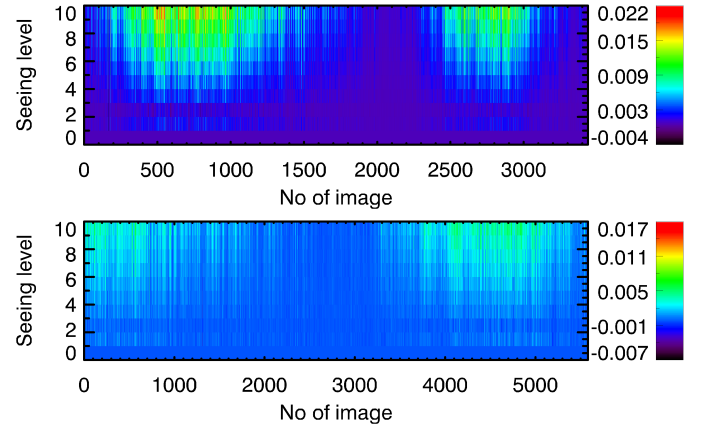
---

[2] Available at http://www2.mps.mpg.de/projects/sun-climate/data.html

Fig. 11. Colour-coded difference in fractional plage areas derived from RP1 (*top*) and MD1 (*bottom*) images downgraded to simulate effects of seeing and those derived from the original RP1 and MD1 images, respectively. The MD1 data used here are only those taken with a CCD camera between 2002 and 2017. Each row (column) of boxes shows results derived from a given observation (width of the smoothing Gaussian function). See Sect. 4 for details.

Fig. 10. Plage areas in disc fraction as a function of time from the datasets of Ba (green squares), Br (blue asterisks), CL (purple x symbols, only in the *top panel*), Ka (red plus signs), ML (orange rhombuses), and Te (ciel triangles). Results are shown for individual images over the course of the first week in June 2014 (*top panel*) and over the course of 02 June 2014 (*bottom panel*). Also shown are the sunspot areas by Balmaceda et al. (2009, black squares) multiplied by ten to bring them to roughly the same level as the plage areas for the sake of comparison.

we notice that the plage areas at the maxima of SC 14, 19, 20, 22, and 23 are slightly reduced in the new composite with respect to those in Paper II, while most minima are slightly elevated. The increase of the values over activity minima is partly due to the inclusion of more data taken with a relatively narrow bandwidth compared to the composite in Paper II. Similarly, over SC 19, the newly added data favour lower activity level compared to that of MW or MD1 data. However, we note that the majority of MD1 data over SC 19 are still not digitised, which may affect our composite series.

4. Discussion

Here we focus on the uncertainties of our results due to characteristics of the analysed data, that is, the relevant seeing conditions, or central wavelength used for the observations, as well as those due to our method of producing the composite plage area series. The accuracy of the methods applied to the processing of the observations was tested and discussed by Chatzistergos et al. (2018b, 2019b,c).

First, we estimate the sensitivity of daily plage areas from individual datasets to the observational conditions. These include the varying seeing conditions at different times of day and different locations, as well as stray-light. To this end, we

analysed all available observations of the first week of June 2014. Figure 10 shows the derived plage areas over that period. Over those days, the plage areas from all datasets as well as the sunspot areas show a roughly steady increase. We also notice the derived plage areas to increase through the course of each day, which is consistent with the worsening seeing conditions towards local noon, that smears the plage regions. However, at least part of this increase is due to the increase of solar activity over the course of that week. This increase is expected to be greater for the archives with observations taken with narrow bandwidths, such as Te. We find the areas from Ka to show more stable values within a given day, with only a slight increase of the areas around 06:00 UTC followed by a slight decrease after that. The results in Fig. 10 allow a rough estimate of the uncertainty of the derived plage areas due to the seeing conditions during the acquisition of the images. We find the areas from Ba, Br, and CL observations to exhibit a typical daily variation of ∼0.01 in disc fraction, with a somewhat lower variation for Te and Ka data (about 0.005), and a greater variation (up to about 0.02) for ML data. We cannot judge whether the bandwidth of the observations has any effect on these results.

However, the passage of plage regions over the disc, part of which can go behind the limb or appearance of new plage regions at the limb within a day, also affects these results. To remove this uncertainty, we also simulated the effects of seeing on RP1 and MD1 images by applying on them a Gaussian smoothing filter with varying widths. We used ten values for the width uniformly distributed in the range [0,2] σ, where σ is the standard deviation of the Gaussian function. Figure 11 shows the residual plage fractional areas derived from the smoothed and the original (not smoothed) images. Similarly to the results of the actual archives, we find a variation in the plage areas with generally higher values for the smoothed images. This might come as a surprise considering that the plage areas are smoothed and hence their areas are expected to get reduced. However, the smoothing also decreases the standard deviation of the QS regions, hence our threshold to isolate the plage areas is lower compared to the original images. The variation in the residual plage areas follows the SC, with the highest values occurring during activity maxima. The variations reach a value of 0.02 for the most severe case considered here.
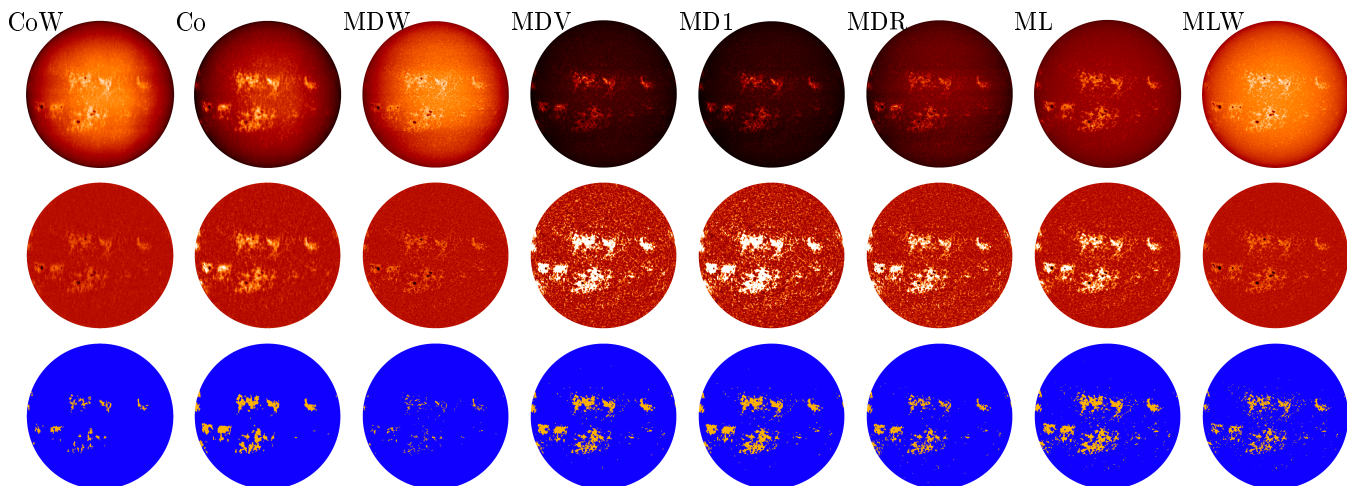
Fig. 12. Examples of observations taken on the same date (04 July 2014) from the various datasets analysed in this study with images not centred at the core of the Ca II K line along with those from the same observatories centred at the core of the Ca II K line. Shown are images from the CoW, Co, MDW, MDV, MD1, MDR, ML, MLW datasets after the preprocessing to identify the disc and re-sample them to account for the disc ellipticity (*top row*), after CLV-compensation (*middle row*), and after the identification of plage (*bottom row*). The images have been aligned to show the solar north pole at the top. The raw images are shown to their entire range of values, the CLV-compensated images are shown in the range [−0.5,0.5], while in the masks, the plage regions are shown in orange and the quiet Sun and network regions in blue. We stress that the same threshold was used for all datasets to identify the plage regions, which is why the different datasets seem to provide rather different ranges of plage coverage.

The results for the MD1 and RP1 data are almost identical. The only differences we identified are that MD1 data have marginally lower errors (up to 0.017 instead of 0.019 found in RP1 results when considering the common observations to MD1) and that there are images for which the plage areas decrease slightly more than RP1 ones (minimum value of −0.007 instead of −0.002 found in RP1 data for the common days to MD1). The values for the differences in the plage areas are consistent with the results of the actual datasets.

We also estimate the uncertainties in the derived plage areas due to variations in the central wavelength of the observations. All Ca II images can exhibit such variations, irrespectively of whether they were taken with a spectroheliograph, or with an interference filter. In the case of the spectroheliograph, the variations are due to the positioning of the slit, for example, because the observers intentionally or unintentionally took observations off-centred from the line core. In general, cases where the observers intentionally took observations off-centred are documented, as in the archives from MD1 and Co for which observations centred at K1 and K3 are separated (the former are labelled MDW and CoW, respectively, in this paper, while the latter are abbreviated as MD1 and Co). However, this practice was not always documented consistently for all of the datasets considered in our study. In the case of the interference filters, the variations are more consistent and mostly due to the replacement of the filters or filter degradations, manifesting as a single offset or a drift in time, respectively.

In order to test the effects of the different central wavelength on our derived plage areas we used the off-band observations from Coimbra, Meudon, and Mauna Loa. Meudon has observations over the period 2002–2017 taken at four different wavelengths around the core of the Ca II K line along with the data taken at the core of the line. All of those observations were taken with the same bandwidth. Here we use only those taken at the centre of the line and the two extreme cases which correspond to offsets of ±0.3 Å. We refer to those series as MDV and MDR for the data taken in the violet and red part of the wing of the line,

respectively. Coimbra and Meudon also have observations centred beyond the K1V minimum, with an offset of −1.4 Å of the Ca II K line (CoW and MDW, hereafter), taken with the same bandwidth as those centred at the core of the line. The Mauna Loa dataset includes data taken at the core of the Ca II K line as well as data taken 2.6 Å offset to the red wing (MLW, hereafter) of the line with a narrower bandwidth of 1 Å compared to those taken at the core of the line (which have a bandwidth of 2.7 Å).

Figure 12 shows examples of these data for observations taken on 4 July 2014. We notice that the contrast of plage areas decreases for all data taken off-centred, and unsurprisingly it is lower for data taken further away from the core of the line. Moreover, the network regions are diminished in contrast and the sunspots are enlarged. The MLW images are quite similar to the CoW and MDW ones, even though MLW is supposed to be taken further away from the core of the line than the other two. However, the contrast of the MLW images is slightly greater than those from CoW and MDW. This might be an effect of the differences in the bandwidth of these observations, with the one used for MLW being considerably broader than those used for CoW and MDW.

Figure 13 gives the absolute difference in the derived plage areas between the datasets centred at the core of the line and those taken off-centred. The same segmentation method was applied to all the data. The plage areas derived from such datasets also decreases compared to the values we get for the data centred at the core of the line. The difference for the ML-MLW data is lower than for the MD1-MDW or Co-CoW ones, reaching a value of 0.02 during activity maximum. The differences show variations following the SC, but a small offset is also noticed during activity minimum, being on average between 0.001 and 0.002 for MD1-MDV and MD1-MDW, respectively. The differences are greatest for the cases MD1-MDW and Co-CoW reaching values of 0.04 during activity maxima. We expect the typical variations of the central wavelength in the historical data to be similar to those for the cases of MD1-MDR and MD1-MDV and, hence, they can provide a very approximate
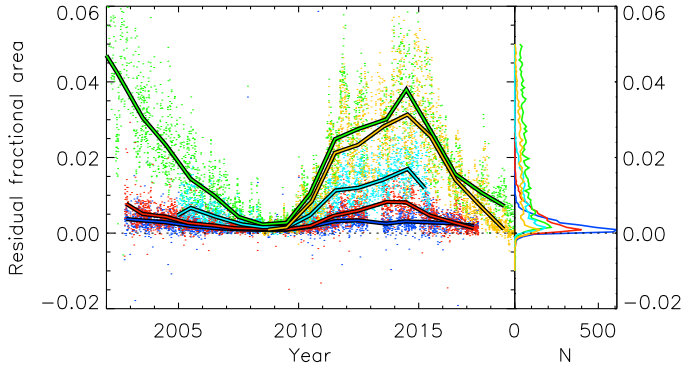
Fig. 13. *Left*: difference of fractional plage area between those derived from centred and off-centred Ca II K observations as a function of time. The differences shown are for the Co-CoW (orange), MD1-MDW (green), MD1-MDV (blue), MD1-MDR (red), ML-MLW (ciel). Dots show daily values, while the solid lines show annual median values. The dotted horizontal line denotes a difference of 0. *Right*: distributions of the differences.

uncertainty level in the determined plage areas from the historical data due to shifts in the central wavelength. For MD1-MDR and MD1-MDV, we notice that the distribution of the differences to be quite narrow when compared to the other cases tested here. The differences in the derived plage areas for these cases are on average (RMS difference) less than 0.003 (0.003). However, we also notice that even though the wavelength offsets for MDR and MDV are equal in absolute value, the results for the plage areas are not exactly the same. The difference in the plage areas for MDR are greater than for MDV. This is unsurprising considering the similarity of MDV and MD1 images compared to the MDR ones (see Fig. 12).

Next, we tested to which degree the composite series is affected by our choice of individual backbone datasets. In this process, we used all historical datasets with sufficiently long periods to act as backbone references. These are the Ar, Ko, Ky, MD1, MM, MW, Ro, and SP datasets. Figure 14 shows the resulting plage area composite series. We did not consider Co, Mi1, or Mi2 in this test due to the poor overlap with many of the other datasets. Table B.1 lists the assignment of the various plage series to the individual backbones. For consistency, all series are normalized to the level of the RP1 backbone. All produced composites agree almost perfectly over SCs 22–24, owing to the RP1 backbone. However, there are disagreements for the remaining cycles. The differences are greatest for the composites created with MM, Ky, and SP as the backbones for the historical data. This is consistent with the rather low amount of data within those datasets, rendering the calibration of the various other (non-backbone) datasets to their level more uncertain. The remaining reconstructions show results that are very similar to our proposed series. The differences are typically below 0.005, but increase over SC 19 to 0.02 for daily values. This gives a rough estimate of the uncertainty in our official composite series due to the selection of individual datasets to act as the reference. We note, however, that the overlap between most of the series used as backbones for this test is not optimum and is always worse compared to that of the Ko and MW series. Furthermore, by averaging the Ko and MW backbone series we reduce the uncertainty due to the choice of the reference series.

Figure 15a shows the composite based on using RP1 as a backbone series (blue). It is plotted along with the RP1 series on its own (black). These two series match almost perfectly, with only small differences (RMS difference of 0.002 in disc frac-

tion) which are greatest in 1998 (reaching 0.008 in disc fraction). Also shown in Fig. 15a is the composite using the RP1 backbone series when only the days included in the individual RP1 series are considered (red). In this case the differences are reduced, with an RMS difference of 0.001 and a maximum difference in 1999 of 0.003 in disc fraction. This illustrates the accurate cross-calibration of the individual series to the RP1 backbone.

Finally, we also test the effect of including data taken with different bandwidths in the backbone series. For this purpose, we use the RP1 backbone and produce three different versions of it: (i) keeping only datasets with either bandwidths narrower than 2 Å or broader than 3 Å (Ba, CL, Co, Kh, PS, RP2, SF1, SF2, Te, Up). Since the RP1 bandwidth is 2.5 Å, these limits imply that the chosen bandwidths differ by at least ±0.5 Å from that of RP1. (ii) including only datasets with the nominal bandwidth between 2 Å and 3 Å, that is, within ±0.5 Å of the bandwidth of RP1. In this case, we further subdivide the datasets according to whether the bandwidths that have been assigned to them; ii (a) appear to be consistent with their actual behaviour (BB, Br, ML) ii (b) or their assigned bandwidth does not appear to be consistent with their actual behaviour (Ka, PM, VM). For this test, we considered the BB and ML datasets only over the periods when the final instrumentation was used (see Appendix B), in order to avoid inconsistencies due to instrumental changes. Figure 15b displays the three test backbone series in comparison to the one used in our composite. All three reconstructions of the RP1 backbone series lie within the uncertainties. There are generally small differences, which are greater in 2011, before 2000, and after 2018, reaching up to 0.01 in disc fraction.

5. Summary and conclusions

We processed 43 datasets of full-disc Ca II K observations spanning the period 1892–2019 to derive the evolution of plage areas over the last 12 solar cycles. We processed the data consistently with the method developed by Chatzistergos et al. (2018b, 2019b). An extra step in the processing was added to ensure accurate results from the analysis of images from the Ky and YR datasets, which suffer from specific artefacts along arcs on the solar disc. We adapted our processing such that these artefacts can be precisely accounted for and showed that we can obtain accurate results for those datasets as well.

We used our results for the 38 datasets with observations centred at the Ca II K line to create a plage area composite by applying the backbone approach employed by Chatzistergos et al. (2017) to create a sunspot group number composite. We created two backbones, one mostly for the modern CCD-based data and another for the historical data, which were mainly stored on photographic plates. We considered the plage area series from Rome (RP1) observations as the overall reference series and to act as the backbone for the modern data, while both Kodaikanal (Ko) and Mt Wilson (MW) acted as the references for the historical data. The obtained composite on the whole reasonably is consistent with the one we presented in Paper II, although small differences exist. The composite derived in this study has an average annual coverage of 98% for the periods after 1906, with observations for only 672 days missing. The coverage, however, remains rather low for the period 1892–1906 with 4917 days without any observations recorded. Previous results in the literature were based on considerably poorer temporal coverages. We also illustrated the importance of using multiple datasets to improve the annual coverage in comparison to the case when the results derive from observations from the Ko and MW datasets
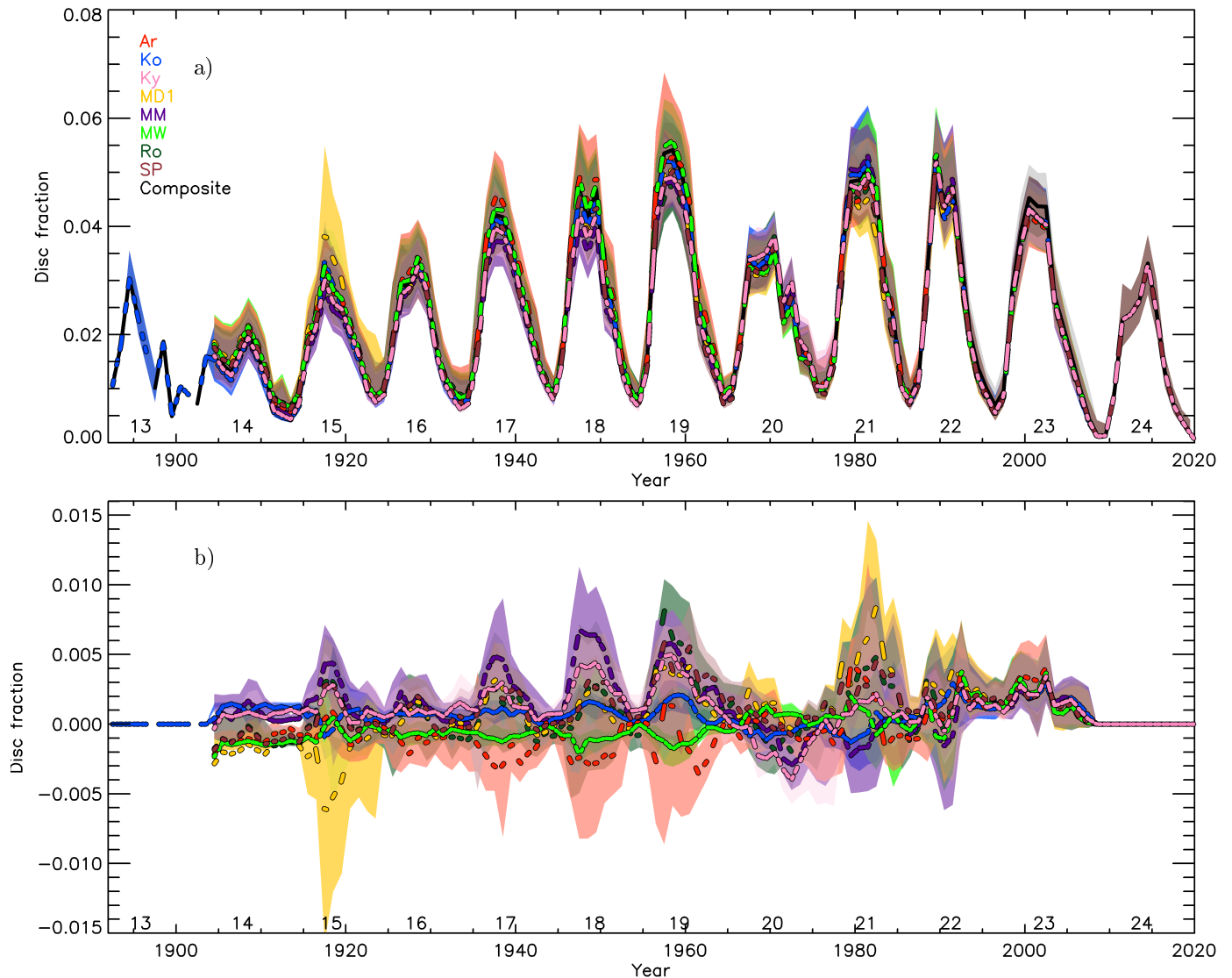
Fig. 14. Plage area composites *(a)* and their differences in comparison to our main composite series *(b)* produced when the historical backbone series was computed with various individual historical datasets as the backbone reference instead of using the average of the calibrated Ko and MW ones to the RP1 one. RP1 is always taken as the overall reference for all composite series. The datasets used as backbones are those from Ar (red), Ko (blue), Ky (pink), MD1 (orange), MM (purple), MW (green), Ro (dark green), and SP (brown) sites. The composite plage area series derived in this study is also displayed in black. Shown are the annual median values (lines) of the final plage area composites along with the asymmetric 1σ interval (shaded surfaces). The numbers below the curves denote the conventional SC numbering.

alone, which are the ones most employed in studies of the plage areas evolution so far.

Many observatories, whose data have been analysed here, have stopped the solar monitoring in the Ca II K line. However, observations in the Ca II K line continue at the Ba, Br, CL, Co, Ka, Kh, Ki, KW, MD1, Mi2, PM, RP1, RP2, SF1, SF2, Te, UP, and VM sites to this day. A combination of the data from all those sites provides a full annual coverage. There is no day without an observation in our composite series since 2010. However, there are still gaps in our composite. Hopefully, more historical archives, such as those from Abastumani (Khetsuriani 1981), Anacapri (Antonucci et al. 1977), Cambridge (Moss 1942), Crimea, Ebro (Curto et al. 2016), Huairu (Suo 2020), Kandilli (Dizer 1968), Locarno (Waldmeier 1968), Madrid (Vaquero et al. 2007), and the remaining data from the Baikal, Catania, Kenwood, Kharkiv, Kislovodsk, Manila, Meudon, Yerkes, Wendelstein, and Schauinsland sites will be digitised in the near

future, which can potentially further increase the daily coverage of the data entering our composite series.

In this paper, we have aimed to shed light on various issues affecting individual Ca II K archives. We accounted for some of these in a simple manner by splitting the series into different parts and performing their cross-calibration to the backbone series separately. However, there are more issues that remain unaccounted for in our analysis, such as the variable bandwidth and central wavelength of the observations taken with a spectroheliograph. For these, we present an estimate of the uncertainty in our results. However, we plan to further address the effects due to variable bandwidth and central wavelength of the observations with a machine learning approach, considering that such methods have shown great potential on image-to-image conversion (e.g. Kim et al. 2019; Galvez et al. 2019; Park et al. 2019). In addition, we plan to continue processing the data from the currently operating programs in the Ca II K line to update the
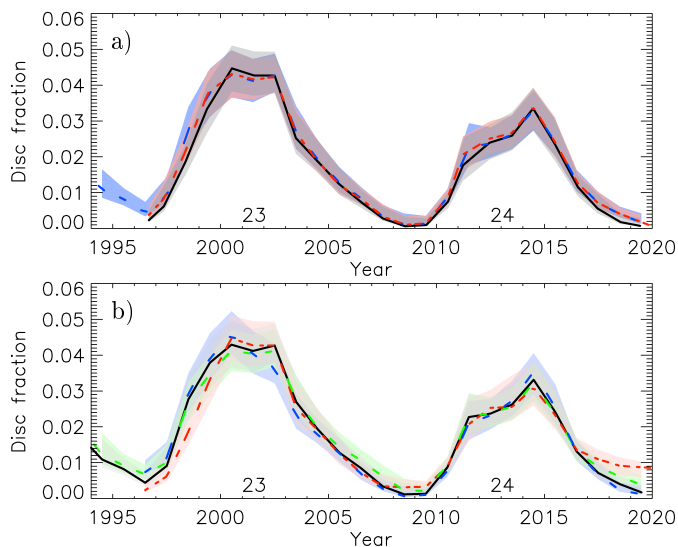
Fig. 15. *Top*: RP1 plage area series (solid black) plotted along with the RP1 backbone composite series used in our plage area composite (dashed blue) and RP1 backbone composite series keeping only the days common with the RP1 individual series (dashed red). *Bottom*: RP1 backbone composite series when varying the included individual datasets based on their bandwidth. The different curves show the RP1 backbone composite by keeping Ba, Co, Kh, RP2, SF1, SF2, Te, Up (dashed green), Br, ML, BB (dashed blue) and Ka, PM, VM (dashed red). The RP1 backbone series used for our composite is shown in solid black. Depicted are annual median values (solid lines) along with the 1σ asymmetric intervals (shaded surfaces). The SC numbers are given below the curves.

composite at regular intervals. Finally, we plan to continue improving and updating the composite by including further historical data whenever these are digitized and made available.

## References

Antonucci, E., Azzarelli, L., Casalini, P., & Cerri, S. 1977, Sol. Phys., 53, 519
Balmaceda, L. A., Solanki, S. K., Krivova, N. A., & Foster, S. 2009, J. Geophys. Res.: Space Phys., 114,
Barata, T., Carvalho, S., Dorotovic, I., et al. 2018, Astron. Comput., 24, 70
Belkina, I. L., Beletskij, S. A., Gretskij, A. M., & Marchenko, G. P. 1996, Kinematics Phys. Celest. Bodies, 12, 55
Berrilli, F., Ermolli, I., Florio, A., & Pietropaolo, E. 1999, A&A, 344, 965
Bethge, C., Peter, H., Kentischer, T. J., et al. 2011, A&A, 534, A105
Chapman, G. A., Cookson, A. M., & Dobias, J. J. 1997, ApJ, 482, 541
Chapman, G. A., Dobias, J. J., & Arias, T. 2011, ApJ, 728, 150
Chapman, G. A., Cookson, A. M., & Preminger, D. G. 2012, Sol. Phys., 276, 35
Chatterjee, S., Banerjee, D., & Ravindra, B. 2016, ApJ, 827, 87
Chatterjee, S., Mandal, S., & Banerjee, D. 2017, ApJ, 841, 70
Chatzistergos, T. 2017, Analysis of Historical Solar Observations and Long-term Changes in Solar Irradiance PhD thesis (Uni-edition)
Chatzistergos, T., Ermolli, I., Solanki, S. K., & Krivova, N. A. 2016, in Coimbra Solar Physics Meeting: Ground-based Solar Observations in the Space Instrumentation Era, eds. I. Dorotovic, C. E. Fischer, & M. Temmer (San Francisco), ASP Conf. Ser., 504, 227
Chatzistergos, T., Usoskin, I. G., Kovaltsov, G. A., Krivova, N. A., & Solanki, S. K. 2017, A&A, 602, A69
Chatzistergos, T., Ermolli, I., Solanki, S. K., & Krivova, N. A. 2018a, A&A, 609, A92
Chatzistergos, T., Ermolli, I., Krivova, N. A., & Solanki, S. K. 2018b, in Long-term Datasets for the Understanding of Solar and Stellar Magnetic Cycles, eds. D. Banerjee, J. Jiang, K. Kusano, & S. Solanki (Cambridge, UK: Cambridge University Press), IAU Symp., 340, 125
Chatzistergos, T., Ermolli, I., Falco, M., et al. 2019a, Il Nuovo Cimento, 42C, 5
Chatzistergos, T., Ermolli, I., Krivova, N. A., & Solanki, S. K. 2019b, A&A, 625, A69
Chatzistergos, T., Ermolli, I., Solanki, S. K., et al. 2019c, Sol. Phys., 294, 145
Chatzistergos, T., Ermolli, I., Solanki, S. K., et al. 2019d, A&A, 626, A114
Chatzistergos, T., Ermolli, I., Krivova, N. A., & Solanki, S. K. 2020, IOP J. Phys.: Conf. Ser., 1548, 012,007
Curto, J. J., Solé, J. G., Genescà, M., Blanca, M. J., & Vaquero, J. M. 2016, Sol. Phys., 291, 2587
Dizer, M. 1968, Sol. Phys., 3, 491
Dorotovic, I., Rybák, J., Garcia, A., & Journoud, P. 2010, 20th National Solar Physics Meeting, 20, 58
Ermolli, I., Berrilli, F., & Florio, A. 2003, A&A, 412, 857
Ermolli, I., Criscuoli, S., Centrone, M., Giorgi, F., & Penza, V. 2007, A&A, 465, 305
Ermolli, I., Marchei, E., Centrone, M., et al. 2009a, A&A, 499, 627
Ermolli, I., Solanki, S. K., Tlatov, A. G., et al. 2009b, ApJ, 698, 1000
Ermolli, I., Krivova, N., & Solanki, S. K. 2010, in 38th COSPAR Scientific Assembly, COSPAR Meeting, 38, 133
Ermolli, I., Criscuoli, S., & Giorgi, F. 2011, Contrib. Astron. Obs. Skalnate Pleso, 41, 73
Ermolli, I., Matthes, K., Dudok de Wit, T., et al. 2013, Atm. Chem. Phys., 13, 3945
Ermolli, I., Chatzistergos, T., Krivova, N. A., & Solanki, S. K. 2018, in Long-term Datasets for the Understanding of Solar and Stellar Magnetic Cycles, eds. D. Banerjee, J. Jiang, K. Kusano, & S. Solanki (Cambridge, UK: Cambridge University Press), IAU Symp., 340, 115
Fontenla, J. M., & Landi, E. 2018, ApJ, 861, 120
Galvez, R., Fouhey, D. F., Jin, M., et al. 2019, ApJS, 242, 7
Garcia, A., Sobotka, M., Klvana, M., & Bumba, V. 2011, Contrib. Astron. Obs. Skalnate Pleso, 41, 69
Golovko, A. A., Golubeva, E. M., Grechnev, V. V., et al. 2002, Solar Variability: From Core to Outer Frontiers (ESA Publications Division), 506, 929
Gray, L. J., Beer, J., Geller, M., et al. 2010, Rev. Geophys., 48, 4001
Haigh, J. D. 2007, Liv. Rev. Sol. Phys., 4, 2
Hale, G. E. 1893, Mem. Soc. Spettr. It., 21, 68
Hale, G. E., & Ellerman, F. 1903, Publ. Yerkes Obs., 3, I.1
Hanaoka, Y. 2013, J. Phys. Conf. Ser., 440, 2041
Hanaoka, Y. 2016, Solar Observatory of NAOJ 2016, in Coimbra Solar Physics Meeting: Ground-based Solar Observations in the Space Instrumentation Era, eds. I. Dorotovic, C. E. Fischer, & M. Temmer (San Francisco), ASP Conf. Ser., 504, 313
Hiremath, K. M., Rozelot, J. P., Sarp, V., et al. 2020, ApJ, 891, 151

Hirtenfellner-Polanec, W., Temmer, M., Pötzi, W., et al. 2011, Cent. Eur. Astrophys. Bull., 35, 205

Khetsuriani, T. S. 1981, Sol. Phys., 69, 405

Kim, T., Park, E., Lee, H., et al. 2019, Nat. Astron., 1

Kitai, R., Ueno, S., Maehara, H., et al. 2013, Data Sci. J., 12, WDS213

Klimeš, J. J., Belik, M., Klimeš, J. S., & Marková, E. 1999, in 8th SOHO Workshop: Plasma Dynamics and Diagnostics in the Solar Transition Region and Corona, eds. J. C. Vial, & B. Kaldeich-Schü, ESA SP, 446, 375

Koechlin, L., Dettwiller, L., Audejean, M., Valais, M., & Ariste, A.L. 2019, A&A, 631, A55

Kopp, G., Krivova, N., Wu, C. J., & Lean, J. 2016, Sol. Phys., 291, 2951

Lefebvre, S., Ulrich, R. K., Webster, L. S., et al. 2005, Mem. Soc. Astron. It., 76, 862

Lenza, L., Srba, J., Gregorova, B., Exnerova, M., & Lenzova, N. 2014, System for Simultaneous Observation of Solar Flares in Spectral Lines of H-alpha and CaII K

Loukitcheva, M., Solanki, S. K., & White, S. M. 2009, A&A, 497, 273

Lourenço, A., Carvalho, S., Barata, T., et al. 2019, Open Astron., 28, 165

Malherbe, J.-M., & Dalmasse, K. 2019, Sol. Phys., 294, 52

Meftah, M., Hochedez, J.-F., Irbah, A., et al. 2014, Sol. Phys., 289, 1043

Meftah, M., Corbard, T., Hauchecorne, A., et al. 2018, A&A, 616, A64

Miller, R. A. 1965, Appl. Opt., 4, 1085

Mohler, O. C., & Dodson, H. W. 1968, Sol. Phys., 5, 417

Moss, W. 1942, MNRAS, 102, 86

Naqvi, M. F., Marquette, W. H., Tritschler, A., & Denker, C. 2010, Astron. Nachr., 331, 696

Park, E., Moon, Y.-J., Lee, J.-Y., et al. 2019, ApJ, 884, L23

Priyal, M., Singh, J., Ravindra, B., Priya, T. G., & Amareswari, K. 2014, Sol. Phys., 289, 137

Pruthvi, H., & Ramesh, K. B. 2015, Int. Conf. Opt. Photonics, 2015, 9654 96540I

Raju, K. P. 2018, MNRAS, 478, 5056

Rast, M. P., Ortiz, A., & Meisner, R. W. 2008, ApJ, 673, 1209

Shapiro, A. I., Solanki, S. K., Krivova, N. A., et al. 2017, Nat. Astron., 1, 612

Singh, J., & Ravindra, B. 2012, Bull. Astron. Soc. India, 40

Solanki, S. K., Krivova, N. A., & Haigh, J. D. 2013, ARA&A, 51, 311

Suo, L. 2020, Adv. Space Res., 65, 1054

Svalgaard, L., & Schatten, K. H. 2016, Sol. Phys., 291, 2653

Tlatov, A. G., & Tlatova, K. A. 2019, Geomag. Aeron., 59, 6

Tlatov, A. G., Pevtsov, A. A., & Singh, J. 2009, Sol. Phys., 255, 239

Tlatov, A. G., Dormidontov, D. V., Kirpichev, R. V., Pashchenko, M. P., & Shramko, A. D. 2015, Geomag. Aeron., 55, 961

Vaquero, J. M., Gallego, M. C., & Acero, F. J. 2007, in The Physics of Chromospheric Plasmas, eds. P. Heinzel, I. Dorotovic, & R. J. Rutten, ASP Conf. Ser., 368

Vogler, F., Brandt, P., Otruba, W., Pötzi, W., & Hanslmeier, A. 2008, Cent. Eur. Astrophys. Bull., 32, 141

Waldmeier, M. 1968, Sol. Phys., 5, 423

Wöhl, H. 2005, Hvar Obs. Bull., 29, 319

Wu, C.-J., Krivova, N. A., Solanki, S. K., & Usoskin, I. G. 2018, A&A, 620, A120

Yeo, K.L., Krivova, N.A., & Solanki, S.K. 2017, J. Geophys. Res.: Space Phys., 2016JA023733

Zuccarello, F., Contarino, L., & Romano, P. 2011, Contrib. Astron. Obs. Skalnate Pleso, 41, 85

Appendix A: Processing of Kyoto and Yerkes data
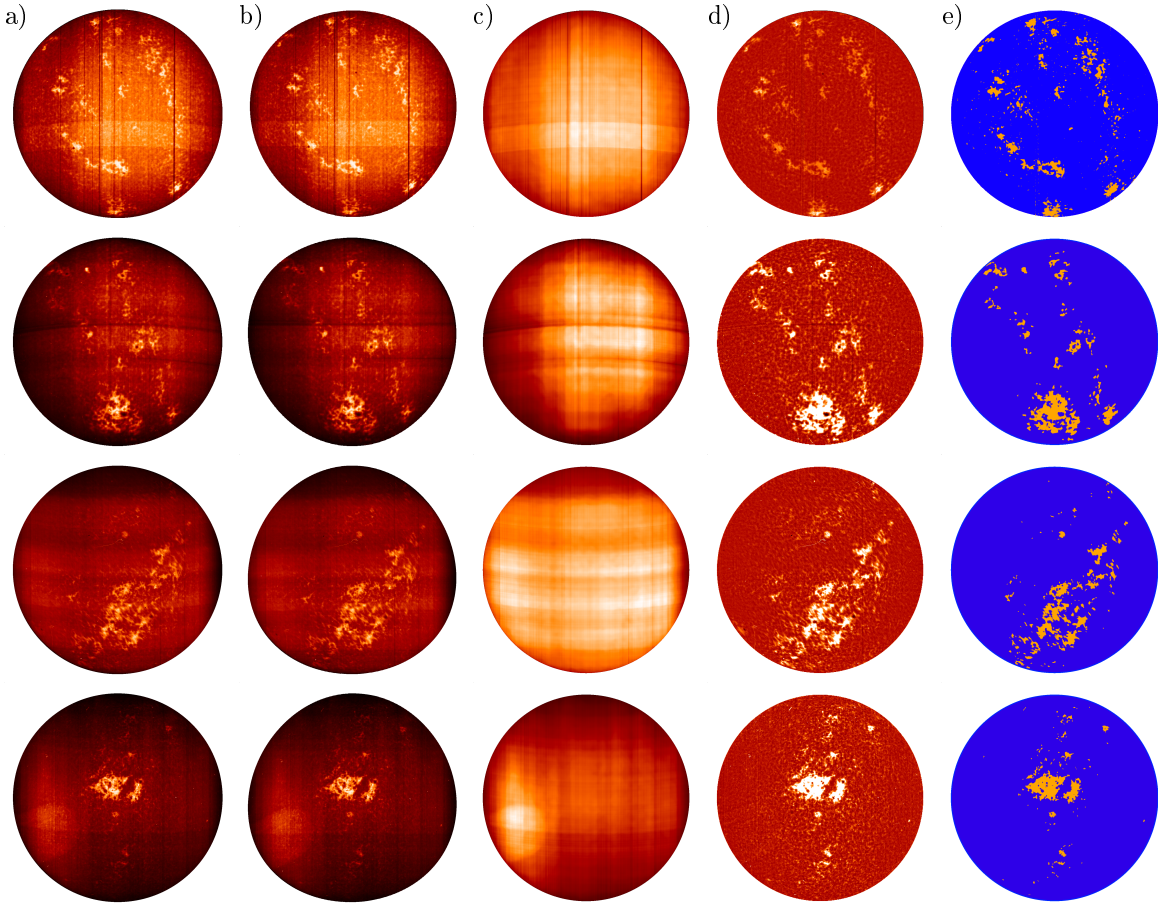
Fig. A.1. Examples of processing steps applied on Ky observations taken on 22/10/1937 (*first row*), 28/10/1939 (*second row*), 21/01/1959 (*third row*), and 30/04/1961 (*fourth row*). Columns are: original raw density image *(a)*, density image resampled to straighten the arcs *(b)*, computed background of the original image *(c)*, photometrically calibrated and limb-darkening-compensated image *(d)*, and segmentation mask *(e)*. The raw density images and the backgrounds are shown to the entire range of values within the disc, the calibrated images are shown in the range $[-0.5, 0.5]$, while the masks show plage regions in orange and QS and network regions in blue. The images are not compensated for ephemeris.

Images from all sites that used spectroheliographs suffer from artefacts introduced by the motion of the slit of the spectroheliograph. Most of these follow the slit's shape, which is linear in almost all instruments. For the Ky and YR archives, however, these artefacts appear to be curved rather than linear. To make things more complicated, there are linear artefacts roughly perpendicular to the curved ones as well (see Fig. A.1), likely introduced by other instrumental issues. The curved artefacts are more evident in the Ky data than in YR ones and for that reason we focus on Ky data here, even though the same process was applied on both series. To account for the curved artefacts in the data, we adopted the following processing.

During the pre-processing of the data, we identified the centre and the radius of the disc as well as the orientation and the curvature of the arcs within each raw image. The identification of the linear and curved segments was initially done automatically. The linear segments in most images are aligned with the frame of the photograph. The frame, however, is not always aligned in the digital image. Therefore, we applied Sobel filtering to identify the frame and then determine the angle needed to align the linear segments vertically. The curved segments were identified by singling out bright or dark regions in the image after it was divided by a map constructed with a running window median filter with width of ten pixels.

However, both the linear and the curved segments are not always clearly visible and the code was not always able to detect them. For that reason, the results were afterwards manually inspected and corrected when deemed necessary. We identify pixels belonging to the same arc and for each pixel n we determine its horizontal distance from the left side of the image, y_n, and its vertical distance from the bottom of the image, x_n. In order to get the parameters of the arcs, we fit a parabolic function of the form:

$$y_n = b_0 + b_1 (x_n - b_2)^2 , \qquad (A.1)$$

where, b_0 and b_2 are the vertical and horizontal distances of the centre of the parabola, while b_1 is its curvature. The values of b_1, b_2 and the angle to orient the linear brightenings or darkenings on the vertical direction were stored in the header of the raw files, while b_0 was not stored as it has a different value for each arc. Furthermore, we make the assumption that all arcs can be described with the same parabola with different offsets in the vertical direction. This is justified, unless there are other distortions of the image affecting the shape of the artefacts. We also note that our processing can return consistent results with small deviations in the determination of the curvature. The angle of the linear segments is used at the beginning of the calibration process to rotate the image so that the linear brightenings or darkenings are on the

vertical direction and the arcs are on the horizontal direction. The image is temporarily re-sampled prior to performing the polynomial fits in the horizontal direction, so that the arcs are straightened (see 2nd column in Fig. A.1). The re-sampling is done by applying a transformation of $y_{ij} = -b_1(x_{ij} - b_2)^2$, where y_{ij} and x_{ij} is the distance in the vertical and horizontal direction of the i, j pixel, respectively. After each fit in the horizontal direction, we apply the inverse transformation on the result of the fit, so to reintroduce the curvature of the arcs. Hence, we apply the polynomial fits across linear segments, but due to the aforementioned transformation, the result follows the curvature of the arcs. The image resulting after the first transformation is somewhat egg-shaped (see second column of Fig. A.1), while the circular shape of the disc is recovered after the inverse transformation. Both the direct and its inverse transformations were applied every time a polynomial fit on the horizontal direction had to be performed. We note that this transformation affects only the result of the fits on the horizontal direction, the original image remained unchanged by this procedure. In addition, we adapted the width of the running window median filter to be $R/20$ instead of $R/6$ which was used for the processing of data from all the other datasets. Examples of the processing applied to Ky data can be seen in Fig. A.1, showing that our processing could accurately remove the severe artefacts affecting the raw data. More such examples are shown in Chatzistergos et al. (2020).

Appendix B: Assignment of datasets in the backbone series

Table B.1 lists the various datasets used in this study and the backbone that they were assigned to. The annual coverage of the datasets in the individual backbone series is shown in Figs. B.1–B.3.

RP1 is the overall reference series for our composite. All series with a sufficient direct overlap to RP1 were assigned to that backbone. This is not the case for the series from Ko and SP, as well as the early periods of BB, Co, Kh, MD1, and SF1 which were not included in the RP1 backbone due to insufficient overlap.

Following Paper II, we split the Ar series at 25/05/1953, the Mi series at 01/02/1966, the MW series at 21/08/1923 and 01/01/1976, the Kh series at 01/01/1994, and the SP series at 01/01/1963 to account for instrumental changes affecting the coherence of the series. Similarly, we split the BB series at 01/07/1992 and 10/09/1996 (Naqvi et al. 2010), the Ka series at 24/11/2012, the SF1 series at 05/02/2002, and the SF2 series at 28/07/1998 when the employed filters were replaced (Chapman et al. 2011). We also split the ML series at 01/01/2005, because prior to that date there were many instrumental issues as evidenced by the log of the observations[3] (see also Vogler et al. 2008). Due to the inhomogeneities of the ML dataset, results from the data prior to 2005 had not been presented prior to this study. We also split the Co series at 01/01/1992 to account for the change of the grating system and 01/01/2008 when a CCD camera started being used for the regular observations. Similarly, we split the MD1 series at 01/01/1919, 17/09/2002, and 15/06/2017 to account for instrumental changes, the introduction of a CCD camera, and the change in bandwidth used for the observations, respectively. Finally, we split the Ki series at 2002 to account for the introduction of the CCD.

We merged the series by Ke[4] and YR[5] due to the low number of recovered images from these datasets and the fact that they were performed with the same telescope and spectroheliograph (Hale & Ellerman 1903). Similarly we merged the series by Sc and WS. We will refer to these two series as Ke/YR and WS/Sc, respectively.

---

[3] http://lasp.colorado.edu/media/projects/pspt_access/PSPT_Final_Release_Notes.pdf
[4] The data were downloaded from http://peabody.yale.edu/collections
[5] The data were downloaded from http://photoarchive.lib.uchicago.edu

Table B.1. Observatories within each backbone.

Observatory	Backbone	
	Our series	Individual
Ar (before 25/05/1953)	Ko, MW	Ky, MD1, MM
Ar (after 25/05/1953)	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
Ba	RP1	
BB (before 01/07/1992)	Ko	MD1, SP
BB (01/07/1992–10/09/1996)	Ko	MD1, SP
BB (after 10/09/1996)	RP1	
Br	RP1	
CT	Ko, MW	Ar, MD1, MM, Ro, SP
CL	RP1	
Co (before 01/01/1992)	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
Co (01/01/1992–01/01/2008)	RP1	
Co (after 01/01/2008)	RP1	
Ka (before 24/11/2012)	RP1	
Ka (after 24/11/2012)	RP1	
Ke/YR	Ko	
Kh (before 01/01/1994)	Ko, MW	Ar, Ky, MM, Ro
Kh (after 01/01/1994)	RP1	
Ki (before 01/01/2002)	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
Ki (after 01/01/2002)	RP1	
Ko	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
KT	RP1	
KW	RP1	
Ky	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
Ma	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
ML (before 01/01/2005)	RP1	
ML (after 01/01/2005)	RP1	
MM	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
MS	RP1	
MD1 (before 01/01/1919)	Ko	
MD1 (01/01/1919–26/09/2002)	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
MD1 (27/09/2002–14/06/2017)	RP1	
MD1 (after 15/06/2017)	RP1	
MD2	RP1	
Mi1 (before 01/02/1966)	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
Mi1 (after 01/02/1966)	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
Mi2	RP1	
MW (before 21/08/1923)	Ko	MD1
MW (21/08/1923–31/12/1975)	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
MW (after 31/12/1975)	Ko	MD1, MM, Ro, SP
PM	RP1	
PS	RP1	
Ro	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
RP1	RP1	
RP2	RP1	
SF1 (before 05/02/1996)	Ko	MD1, SP
SF1 (after 05/02/1996)	RP1	
SF2 (before 28/07/1998)	RP1	
SF2 (after 28/07/1998)	RP1	
SP (before 01/01/1963)	Ko, MW	Ar, Ky, MD1, MM
SP (after 01/01/1963)	Ko, MW	Ar, Ky, MD1, MM, Ro, SP
Te	RP1	
UP	RP1	
VM	RP1	
WS/Sc	Ko, MW	Ar, Ky, MD1, MM, Ro, SP

Notes. Columns present: the name of the observatory and backbone to which it has been assigned in our composite series as well as the individual backbone series discussed in Sect. 4.

Fig. B.1. Annual coverage of datasets included in the RP1 backbone.
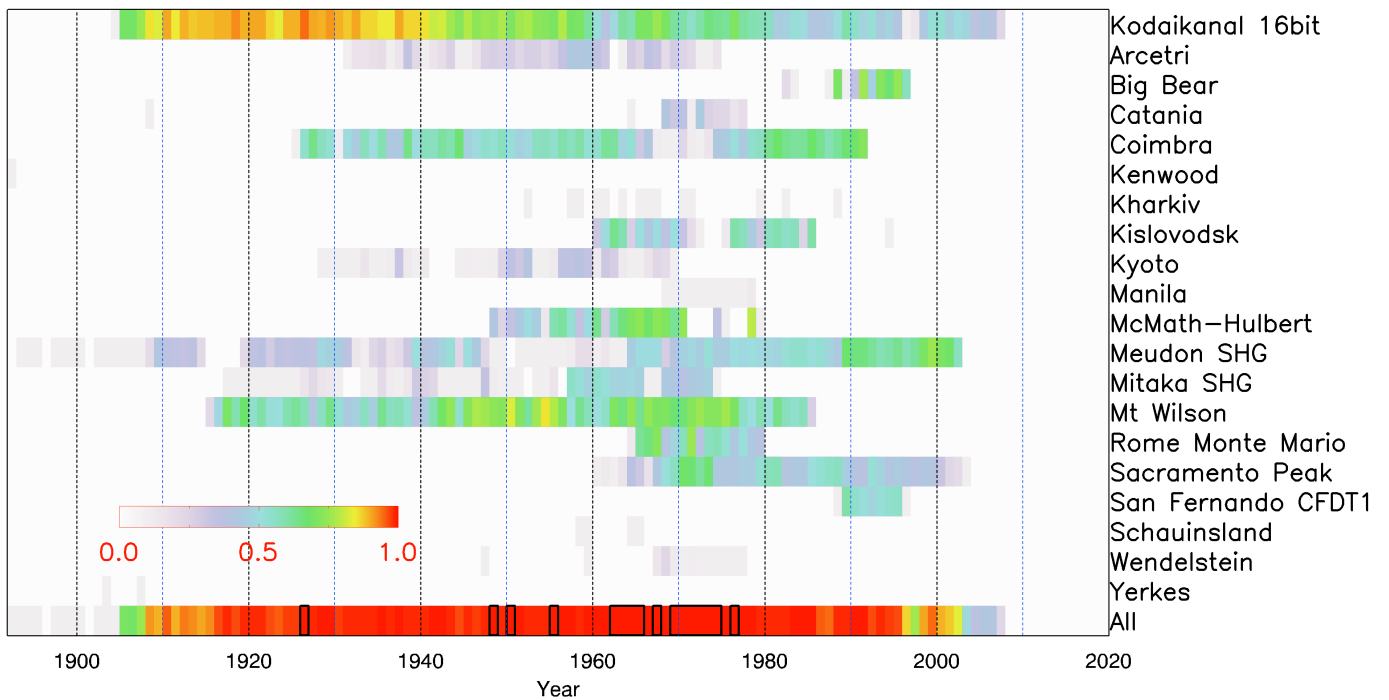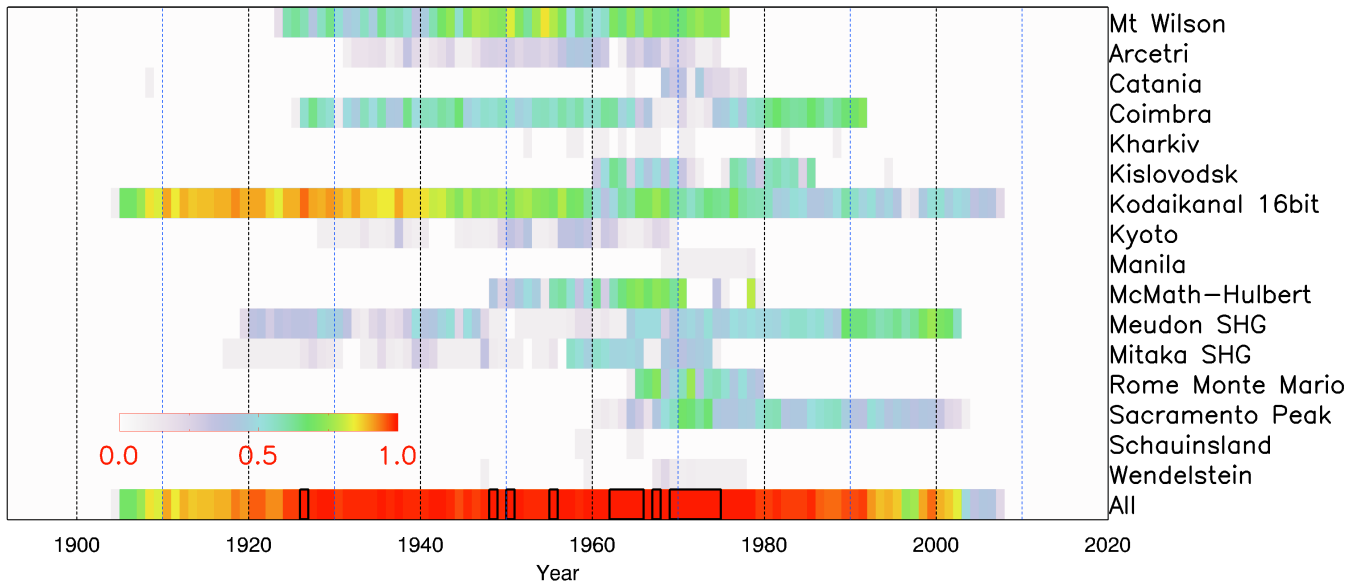

Fig. B.2. Annual coverage of datasets included in the Ko backbone.

Fig. B.3. Annual coverage of datasets included in the MW backbone.