Southern Horizons in Time-Domain Astronomy
Proceedings IAU Symposium No. 339, 2019
E. Griffin, ed.

doi:10.1017/S1743921318002739

Accessing Data for Long-Term Variability

Workshop 12

E. Griffin[1] and L. Balona[2]

[1]Herzberg Astronomy & Astrophysics Research Centre, NRC, Victoria, BC, Canada
email: elizabeth.griffin@nrc-cnrc.gc.ca

[2]South African Astronomical Observatory, Cape Town, South Africa.

Abstract. Astronomy's heritage, comprised mainly of several million photographic plates (both images and spectra), extends backwards in time for many decades. It constitutes a unique and irreplaceable resource for research into long-term variations. Unfortunately, to date rather few of the plates can be accessed in digital formats and cannot therefore be used in modern research. That lack of use encourages the attitude that 'old' data have no value and might as well be destroyed. This Workshop discussed ways and means to avert such threats, and prioritised a list of actions that need to be taken as soon as possible.

Keywords. Astronomical Data Bases: catalogues

1. Background

Studies of the time domain need to involve observations that include a sufficient time-span to support the analyses usefully. For many of the examples that are presently being researched, data from the recent past (i.e., a few years, so within the gamut of our archives of electronic observations) are deemed adequate to provide at least a starting point for studies of the changes that are seen, and if the relevant observations are outside the reach of ground-based equipment and techniques then the electronic archive is all that the researcher can probe. But many other phenomena that display changes still await study, for the simple reason that the data, though archived, cannot be accessed electronically. Thus, astrophysics-rich as it may well prove to be, the information which they contain remains as good as lost to science. However, with an increasing interest in aspects of archival science, the opportunities are more ripe now than ever before to tackle this serious problem of accessing the valuable research resource which our heritage data constitute. But time is not on our side, and the Workshop discussed and prioritised the relevant issues: digitising log-books, the accessibility and location of suitable digitising equipment, whom to involve for training (whether as volunteers or employees), and how to find funding.

2. A Growing Urgency

Although the Workshop did not need to rehearse the scientific losses which threaten if plates are destroyed *en masse*, a few horror stories set the scene vividly as reminders that such such steps can be, *and have been*, taken, and with too little notice to organize a rescue. Feeling – quite rightly – that nothing in this particular domain is safe until all the various pieces of information are correctly converted into electronic files and archived by a major astronomy repository, the Workshop determined to explore all avenues in the search for practicable solutions.

Three distinct stages to tackling the challenges were recognised:

- advertise the contents by digitising log-books or contents catalogues,
- carry out appropriate digitisation of plates, and
- seek funding from supportive sources.

Advertising the contents of a plate-vault is indisputably an essential pre-requisite. Few customers will go into a high-street shop that has an empty window.

3. Step 1. The Essential Pre-requisite: Advertise the Contents!

The most common reason given for wishing to discard a plate collection is "no-one uses it nowadays". That is scarcely surprising since almost none of the plate stores has its contents on-line and in searchable form, or has a custodian of the plate archive ('plate librarian') on-site, so it can be difficult (and time-consuming) to discover what a given plate-store contains without visiting in person.

However, provided the logbooks or equivalent note-books are extant, then to convert that information into an electronic catalogue is not technically demanding. It requires a universally-shared template showing which are the basic elements to include (e.g.) object or central plate co-ordinates, observatory, date, time, exposure length and emulsion type, plus guidance as to whether, and what, is available by way of calibrations. For direct images, if there are no sensitometer 'spots' (or similar), a record of the emulsion type can constitute a starting guide for extracting the photometry usefully, provided a sequence of standard stars can also be established.

For spectra, two kinds of calibration are essential for full spectral analysis: (1) wavelength calibration (usually supplied as a laboratory emission spectrum beside the stellar one), and (2) photometric calibration (often generated by exposing a continuous source through a number of apertures of different but known sizes and giving rise to continuous strips in the direction of the dispersion). A photometric calibration exposure need not be on the same physical piece of glass as the star spectrum; it is better if it is, but some spectrographs do not have the space to do that, and if it is separate then the log-book entry must say where to find it, and what its record number is, if distinct from that of the stellar spectrogram.

The Labour Force. Digitising information of this nature requires quite simple training. Keying-in log-books may involve some background knowledge in order to clarify the shorthand notations that specify which equipment or set-up mode was used, but it can be learned. Printed or typed entries can be read with OCR, but experiments have shown that much depends on the contrast of the printing ink, the shapes of the characters and the state of the paper, and is by no means the answer to prayer. Hand-writing is very unlikely to be manageable satisfactorily by OCR. An alternative approach is to image or photocopy pages or catalogue cards, and invite volunteers to key in the descriptive information manually. Errors are checked by asking several people to handle the same pages. These tasks are well suited to 'citizen scientists', provided a satisfactory level of accuracy can be maintained, and that there is a will to persevere.

4. Step 2: Digitising Plates

'Scanning' a photograph for its scientific content is not the same procedure as 'scanning' it in order to see what it shows, or to convert it into a pretty image, and our language fails to honour that distinction. A flat-bed scanner is designed to prepare a pretty image, or an acceptable copy of a document. Many flat-bed scanners have built-in software for rendering the output visually acceptable (e.g., unsharpening abrupt changes of contrast), but which are disastrous for images of stars that are *intended to be* in sharp focus. The same machines are also rarely rigid enough to avoid positional distortions. High-end scanners are being built that can minimise (though not avoid) positional errors, and residual errors can be further reduced by scanning twice orthogonally. For accurate spectrophotometry, unfortunately, there is no easy solution, and flat-bed scans of spectra will suffer seriously from internally scattered light coming from the light-source. A photographic emulsion is gently rough and scatters light in all directions, though to an extent that depends upon the local plate-density and cannot be corrected since the amount is intrinsically unknown. The *PDS* microdensitometer, built in the 1970s, was specifically designed to avoid such problems; it has two narrow slits, one before the analysing beam passes through the plate, and a narrower one (the analysing slit) afterwards which passes the light onto the recording photomultiplier. It is accurate, but it is necessarily slow: each step has to be recorded in turn, whereas a flat-bed scanner images directly onto a CCD camera that records individual pixels simultaneously.

The great majority of stellar spectrograms show at least three spectra: that of the star is flanked on either side by a wavelength-calibration one (usually a laboratory arc of some kind). Some spectrographs have room for an intensity calibration on the same plate, but if not then an auxiliary calibration exposure will be needed. Not infrequently, there will be multiple stellar spectra ranged vertically down a plate, and each should have its own pair of wavelength calibrations.

The *PDS* works well for spectra, where aligning the sample can take at least as long as the actual scanning, but is prohibitively slow for scanning a direct plate. A flat-bed scanner may suffice in special cases, e.g., if a spectrogram has low dispersion, such as an objective-prism spectrum, or the emulsion is particularly grainy and the scan is required mostly for viewing purposes rather than measurement, but otherwise there is no happy medium between the two approaches. This is an inevitable cost of scientific quality.

In an effort to circumnavigate the speed problem for direct plates, three new scanners have been designed and built specifically for handling large collections of astronomical plates: at Harvard College Observatory, the Royal Observatory of Belgium at Brussels, and the Shanghai Astronomical Observatory. (See Note 1).

Sending plates to a willing scanning team is a possibility to consider, though it is expensive and challenging because of the weight of glass, and it also constitutes a physical risk to the plates themselves. Could the top ends of those machines (i.e., without their massive support structures) be cloned and transported to other observatories?

One significant advantage of distributing cloned scanners would be the introduction of an important degree of standardisation into the digitising project, since the driving software would also be identical. It could then be particularly informative to compare scans of the same plates performed by different observatories.

Who might do the scanning? Digitising photographic plates for scientific purposes is a skilled operation, and requires appropriate training, even if not a background of experience in similar tasks. For efficient throughput, a scheme will involve a carefully designed plan, from the initial handling of plates, listing a unique identifier for each, and filing away (perhaps in new plate envelopes) the ones that have been finished. Digitising spectroscopic plates requires up to five scans: (a + b) ‘clear plate’ both above and below the star exposure and as close as possible to it, (c + d) both wavelength-calibration exposures, and (e) the star spectrum itself. In order to automate the reduction processes by which one extracts an intensity-corrected version of the star spectrum, it will be helpful to perform the five steps in the same order. Orthogonal scans through the calibration exposures, if present or available, must also be made.

Executive decisions may be needed, e.g., when a plate is too weak or too dense, or if there are no wavelength calibration exposures. Is it worth digitising a plate if some of its potential science is thus limited?

https://doi.org/10.1017/S1743921318002739 Published online by Cambridge University Press

5. Step 3: Actions by the Community

(1) *Funding.* Labour is not cheap, particularly the skilled labour required for operating plate-digitising equipment. Suitable storage also incurs costs; a relative humidity of 30–50 % and a temperature around 10 °C or less are specified as *ANSI* standards. A more stringent requirement is for a lack of variability in either: daily changes of no more than 1% in RH and 2 °C in T are strongly recommended, and both should work in tandem so that there is minimal risk of condensation when a storage container is opened. Devoting a room to a plate archive means denying it to other needs – another cost. The output from digitising is mostly relatively small in volume compared to databases of modern astronomical observations, but a reliable channel for transferring electronic files into a responsibly-managed data archive still has to be designed and maintained.

However, it is also fair to compare those costs with the equivalent of the great many telescope-years of observing time and observer-hours which the plates represent. Observations were not made just to be tossed aside once examined; they are better regarded as investments in astrophysics, as applicable today as when they were first made, and (in some cases of time-variability) unique sources of information for solving knotty problems. When viewed against the price of knowledge, the costs for cataloguing and digitising are not in fact large, and it is an argument that should be put forward strongly when applying for funds to establish, maintain or operate any (if not all) of the tasks described above. The cost of a purpose-built scanner is large compared to that of even a high-end flat-bed scanner, and the case for why the latter is inappropriate needs to be made firmly and persuasively. It is also important to mention the value of heritage observations for studies of history and culture, where they certainly occupy a unique niche.

(2) *Revising a significant IAU Resolution.* In 2000 the IAU adopted Resolution B3: *Safeguarding the Information in Photographic Observations*†. It rehearsed the impact of positive action, and the dangers of doing nothing. It recalled sterling efforts to catalogue and preserve information about the *Carte du Ciel* plates and to create a *Wide Field Plates Database*, and recommended a supportive attitude towards programmes that would preserve historic information by transferring it to digital media as a public research resource. Some of the wording of the Resolution would now benefit from being re-cast so as to reflect modern technology more closely and to be more aligned with recent developments in other reaches of time-domain astronomy and database management. The Workshop therefore suggested that it be re-visited, updated and strengthened. (See NOTE 2).

(3) *Establishing a Committee or Working Group.* A committee that brings together people with like concerns has the power to achieve objectives, and the Workshop strongly supported setting up a Committee to speak on behalf of the issues which had been discussed. (See NOTE 3).

(4) *Advertising the need for action* could be addressed by a *Charter for Astronomical Heritage*, which would be circulated to astronomical observatories; it would state the truths that the Workshop expressed, and in much the same wording as in the IAU Resolution, and invite its endorsement by the Director or equivalent at each observatory that houses a collection of photographic plates. An endorsement would form an overture for engaging those responsible in conversation and creative discussion.

6. Updates

Since this report was submitted, two of the points made above have been addressed, and we felt it would enhance the usefulness of this report if that further information were 'added in proof'.

† https://www.iau.org/publications/iau/information_bulletins/ (No. 88, p. 40)

https://doi.org/10.1017/S1743921318002739 Published online by Cambridge University Press

Note 1. (March 2019): A successor to the MAMA digitiser, built by the NAROO project at Meudon Observatory, Paris, is to be in service shortly.

Note 2. At its General Assembly in 2018 the IAU adopted a new Resolution (Resolution B3): "on preservation, digitization and scientific exploration of historical astronomical data". It concludes:

> *Recognizing that the data accumulated over the past decades and even centuries will be lost unless a concentrated action is taken to identify and preserve all significant records, recommends that a concerted effort be made to ensure the preservation, digitization, and scientific exploration of all of astronomy's historical data, both analogue and primitive digital, and associated records.*

A Category B Resolution is binding upon the whole Union, not just upon the Commission(s) by which it was proposed, but does not commit the IAU to putting money to support its endeavours.

Note 3. 'AstroPlate-III', the third in the biennial 'AstroPlate' series of Workshops, was held in 2019 March in Bamberg. It agreed to adopt the IAU WG for the Preservation and Digitisation of Photographic Plates (PDPP; Division B) as its formal Committee to speak on behalf of the issues raised in this report.

https://doi.org/10.1017/S1743921318002739 Published online by Cambridge University Press