

Manifest to Accompany Images of Sir Charles Todd Weather Folios

National Archives of Australia	
Series number	D1348
Title	Sir Charles Todd Weather Folios
Accumulation dates	01 Jan 1879 – 30 Jun 1909
Agency/person recording	Australian Meteorological Association hosted by the Australian Government Bureau of Meteorology
Physical format	Electronic Images (JPG)
Imaging Dates	2008-2010

Function and purpose

These images are a complete record of all pages in the weather folios created by Sir Charles Todd (and his employees), Government Astronomer and later Meteorologist for the Colony of South Australia. The folios contain daily national weather maps, sometimes encompassing New Zealand, and clippings of newstext, telegraphic and handwritten information about the weather mainly in east and south-eastern Australia, with occasional reference to other parts of Australasia and the world. A more extensive listing of the contents of these folios is outlined below.

The folios represent a detailed contemporary view of weather of the day as determined by South Australia's weather professionals and as detailed mainly in the press of the day. The folio series cannot be considered a complete record of weather as not all contemporary weather observations were included in the folios (eg. wet dry bulb, ground temperatures, etc) and Todd's employees were not equally diligent in their capture of relevant news clippings through the whole of the 31 year period. However, the 25,000+ images are a valuable starting point for weather researchers interested in east and south eastern Australian weather between 1879 to mid 1909, when the old colonial weather offices were federalised.

Advice to researchers

- From 1 Jan 1879 to 31 Oct 1881, the maps and text were kept in bound folios with pages measuring 290mm wide and 460mm high and from 1 Nov 1881 onwards, the folio size changed to pages measuring 440mm wide and 330mm high. From Nov 1881 each folio covered six months. *See image file Readme2.jpg.*
- Weather maps occurred on a random basis from mid 1879 to late April 1881, with no weather maps created in 1880 (except maps sourced from newspapers).
- Images of news text typically follow the sequence of the pages (or articles) in the weather folios. Thus 18930204T01.tif is the first page (article) of text for 4 Feb 1893 and 18930204T03.tif is the third page (article). However, there are exceptions where many images are associated with a single date. The multiple images may exist because articles were folded over and/or pasted over each other, making the recreation of a single, comprehensive image of the page impossible. Thus, the images are a close but not a true representations of the folios. However they do contain all the visual material contained therein.
- When searching for an event, researchers should hunt the complete series of files for that date PLUS check either side of the target date as news articles may have been pasted up on the date of newspaper publication rather than the date of occurrence of the weather event.
- Occasionally there may be some missing text for reasons including:
 - article (or part) lost/degraded prior to folios being imaged,
 - rare impossibility of scanning complete articles because there were too many article fold-outs or fold-overs, or portion of article obscured by a glued item.
 - items glued over articles.
- There are some cases of orphan and part-articles that could not be matched to their date.
- Occasionally it may appear that an article is hidden behind other articles. The article will typically appear elsewhere on the image or on other images for that date.

- On some weather maps, the isobaric lines appear faded when other portions of the map are clear. This reflects the occasional random fading of the ink used.
- Those files noted as “script” usually contain useful meteorological data but some may just be simple handwritten annotations of no meteorological significance.
- In the handwritten script, “SA” means South Australia.

Physical characteristics

The original data capture used two methods: Approximately half the records are in .jpg colour format while the remainder are 300 dot per inch greyscale .tiff's. Many pages were both scanned and photographed (see Image Workflow).

Metadata

Metadata was automatically applied by the process of imaging (for example, camera settings as EXIF metadata) In addition, metadata has been applied to each image using Metadata tags that adhere to current Dublin Core and IPTC guidelines. Full details of the schema are detailed in a separate document Sir Charles Todd Weather Folios Metadata Schema. The following metadata fields were applied to each image post-production:

Name	*fixed/var	Metadata field name
<i>Caption**</i>	V	<i>Iptc.Application2.Caption</i>
<i>Copyright Notice</i>	F	<i>Iptc.Application2.Copyright</i>
<i>Country</i>	F	<i>Iptc.Application2.CountryName</i>
<i>Creator</i>	F	<i>Iptc.Application2.Writer</i>
<i>Description Headline</i>	F	<i>Iptc.Application2.Headline</i>
<i>Instructions</i>	F	<i>Iptc.Application2.SpecialInstructions</i>
<i>Intellectual genre</i>	F	<i>Xmp.iptc.intellectualgenre</i>
<i>Keywords</i>	F	<i>Iptc.Application2.Keywords</i>
<i>Country Code</i>	F	<i>Iptc.Application2.CountryCode</i>
<i>Phone</i>	F	<i>Xmp.iptc.CiTelWork</i>
<i>Original filename</i>	V	<i>Iptc.Application2.TransmissionReference</i>
<i>Subject</i>	V	<i>Iptc.Application2.Subject</i>
<i>Subject Code</i>	F	<i>Xmp.iptc.subjectcode</i>
<i>Title**</i>	V	<i>Xmp.dc.title</i>
<i>Website(s)</i>	F	<i>Xmp.iptc.CiUrlWork</i>
<i>World region (wldreg):</i>	F	<i>Xmp.iptcExt.WorldRegion</i>

*F fixed for all images V – variable, dependant on image characteristics

** dates in metadata fields are expressed in little endian form, for example 01/07/1889.

System of arrangement and control

This series of electronic images is virtually a complete record of the National Archives of Australia physical series D1348. The original electronic images are held by the Archives, while the images here are thumbnail derivatives of lesser quality.

References

Provenance, THORPE, Jas. (1886), 'Meteorology In The Southern Colonies', Brisbane Courier, Sat. 2 Jan. p.6, cols.3-6 is a contemporary reference to Charles Todd's office constructing the weather folios.

Project Team Leader's Contextual View of Todd Folios, BENOY, Mac, The Birth of a Familiar, Everyday Map, The Globe #67, 2011, pp.9-22

Advice on technical issues

- Some newsprint image files (filename ends in "T") appear as incomplete page scans. These folio pages had news print pasted on only part of the page thus not all the page was fully scanned. In most cases, the imaging team endeavoured to capture each page to its four edges and beyond to assure the researcher that the full image was captured.
- There is occasional variation in the apparent quality of images. These occurred as a result of the imaging team fine-tuning the workflow and technologies as the project progressed. The variations can include:
 - Image size variation.
 - Varying image intensities and clarity. As part of the quality assurance phase, some folio pages were re-photographed. As this was done later in the project when a high quality camera and lighting were available, the image quality is better than the original run.
 - "Spotty", "Streaky" and changing intensity of grey/black background colours.
 - The backgrounds of some of the images were software enhanced resulting in minor modifications to the document image, primarily in the loss of very small portions of the edges of the paper.
 - Very few images (under 10) have been digitally enhanced.
- The folio covering the period January to June 1897 was imaged differently. This folio was not de-bound (to be kept for posterity). Imaging was done using a higher quality camera only and not the scanner/camera combination as used for the other folios.
- There may be some redundancies in newstext. This occurred for 2 reasons:
 - newstext, which was captured in the scanned text images also appears on the photographed chart, rainfall or script images for the same date.
 - in stitching multiple images together to recreate a page of news articles that fold-over or fold-out, sections of articles may re-occur.

Imaging Workflow

As the folios were fragile, judged by the conservators of the State Library of South Australia as too fragile to be handled, a case was made to de-bind the material to facilitate the handling of individual pages. This was especially important as a large number of the documents were scanned using a flat bed scanner, where handling a heavy large bound volume would have been highly destructive. One volume covering Jan-Jun 1897 was left in its original binding.

Based on contemporary equipment used in the regional office of the National Archives of Australia, a 12 megapixel Canon camera was acquired for imaging. Test camera-captured images were compared with the equivalent scanner-captured images, and though to a naked eye they looked similar in detail, an Optical Character Recognition (OCR) software suite (ABBYY) had recognition rates up to 90% of news print in the scanned images and only 5-10% in their equivalent camera images.

Thus an A3 Microtek ScanMaker 9800XL scanner was acquired to scan all printed text (mainly news print) in greyscale 300 dpi. Since the scanner was much slower than photography, the camera was used to image all folio pages that contained non news-text material (typically weather map, rainfall map and handwritten items).

As the images were created, an enduring identifier was applied as a filename using the following schema:

First six characters identify year, month day as `yyyymmdd`

Next character used as T for newstext, C for chart (map), R for rainfall map and S for pages containing script (handwriting)

Optional character denoting multiple page number (1...n), where required.

In the case of the scanned images, many folio pages required multiple scans as the folio pages were larger than the A3 scanner, and/or the page contained news articles that were folded over.

These multiple images were then stitched together to recreate each page as close as possible to the folio . The software used was Rasterstitch 2.0 which allowed images to be stitched by choosing individual matching pixels in each image giving near seamless aggregation of images. At times the images required manipulation (typically cropping, cut and paste) and Irfanview 4.27 was the main software used for this purpose, but occasional use was made of Photoshop and it's freeware cousin, GIMP.

As a result of a quality review process that examined every image, some post production work was carried out on many images, including straightening, background re-colouring, cropping, re-sizing, blanking, etc.

A final step was completed using software designed by one of the volunteers. This identified suspicious file nomenclature and inconsistent file formats, renamed files from their working monikers, then applied tailored metadata and calculated a check sum for each image file.