

Patterns

Meteorological data rescue: Citizen science lessons learned from Southern Weather Discovery

Highlights

- Data rescue on Zooniverse allowed rapid transcription of historic weather observations
- Eight replicate data entries can be used to obtain consensus with minimal errors
- Transcribed weather observations can dramatically expand OCR character libraries

Authors

Andrew M. Lorrey, Petra R. Pearce, Rob Allan, ..., Ed Hawkins, Patrick Quesnel, Gilbert P. Compo

Correspondence

a.lorrey@niwa.co.nz

In brief

Southern Weather Discovery is a citizen science project on Zooniverse that captured handwritten historical weather observations. This descriptor article outlines how we ran that citizen science project, which can be adapted to a wide range of disciplines. We highlight replicated data keying requirements to minimize transcription errors, some common pitfalls to avoid, and the importance of a good communications strategy. Our partnership with industry on optical character recognition shows potential to harness computer vision to accelerate historical scientific data capture.

Descriptor

Meteorological data rescue: Citizen science lessons learned from Southern Weather Discovery

Andrew M. Lorrey,^{1,8,*} Petra R. Pearce,¹ Rob Allan,² Clive Wilkinson,³ John-Mark Woolley,¹ Emily Judd,^{1,4} Stuart Mackay,¹ Sudhir Rawhat,⁵ Laura Slivinski,⁶ Sally Wilkinson,³ Ed Hawkins,⁷ Patrick Quesnel,⁵ and Gilbert P. Compo⁶

¹National Institute of Water and Atmospheric Research LTD, 41 Market Place, Auckland 1011 New Zealand

²Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK

³RECLAIM, Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK

⁴Now based at Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, Washington DC 20560, USA

⁵Microsoft New Zealand LTD, Microsoft House, Level 5/22 Viaduct Harbour Avenue, Auckland CBD, Auckland 1010, New Zealand

⁶Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, NOAA Physical Sciences Laboratory, Boulder, USA

⁷National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading RG6 6BB, UK

⁸Lead contact

*Correspondence: a.lorrey@niwa.co.nz

<https://doi.org/10.1016/j.patter.2022.100495>

THE BIGGER PICTURE Citizen science has the potential to capture historical handwritten scientific tabulated data that are not held in digital databases. However, undertaking a citizen science campaign for that purpose is not well described, which we address here. Our citizen science data rescue approach constrained data keying targets, developed participant instructions using clear examples, established replication levels to maximize completeness and confidence of data transcription, and demonstrated common data rescue pitfalls. We highlight how an effective communications strategy helps to maintain project momentum. Collaborating with industry to enhance optical character recognition (OCR) capability has the benefit of accelerating data rescue progress that can rapidly augment scientific data repositories. The resulting improvements to comprehensive historical weather datasets with global coverage can support models and predictive capabilities that help mitigate impacts on society from extreme weather.



Proof-of-Concept: Data science output has been formulated, implemented, and tested for one domain/problem

SUMMARY

Daily weather reconstructions (called “reanalyses”) can help improve our understanding of meteorology and long-term climate changes. Adding undigitized historical weather observations to the datasets that underpin reanalyses is desirable; however, time requirements to capture those data from a range of archives is usually limited. Southern Weather Discovery is a citizen science data rescue project that recovered tabulated handwritten meteorological observations from ship log books and land-based stations spanning New Zealand, the Southern Ocean, and Antarctica. We describe the Zooniverse-hosted Southern Weather Discovery campaign, highlight promotion tactics, and replicate keying levels needed to obtain 100% complete transcribed datasets with minimal type 1 and type 2 transcription errors. Rescued weather observations can augment optical character recognition (OCR) text recognition libraries. Closer links between citizen science data rescue and OCR-based scientific data capture will accelerate weather reconstruction improvements, which can be harnessed to mitigate impacts on communities and infrastructure from weather extremes.

INTRODUCTION

The importance of meteorological data rescue

Historical climate research has significantly bolstered global reconstructions of daily weather, also known as reanaly-

ses.^{1–4} Reanalyses are valuable tools for visualizing and contextualizing local weather patterns and extreme weather events,^{3,5,6} as well as investigating climate variability and teleconnections.^{7–9} These long weather reconstructions also improve understanding of broad climate change trends in

regions where observation depths are robust and can be cautiously interpreted.¹⁰

Broadly, reanalyses rely on incorporating historical observations into estimates of past conditions using modern weather models. The performance of centennial-length reanalyses, like the 20th Century Reanalysis,^{3,5} can be highly dependent on the density and accuracy of such observations throughout time. Uncertainties in historical daily weather patterns from these reanalyses can arise from a diminished spatiotemporal coverage of near-surface terrestrial and marine observations that were assimilated into the reconstruction. However, international surface observation databases (e.g., International Combined Ocean-Atmosphere Dataset [ICOADS]; International Surface Pressure Databank [ISPD])^{11–16} that underpin reanalyses are continually expanding as new data are recovered and digitized by ongoing meteorological data rescue efforts.¹⁷ Data rescue therefore creates new pathways to improve reanalyses, like 20CR, but those opportunities are heavily dependent on (1) the ability to locate missing meteorological records for areas where spatial coverage is weak, and (2) the capability and capacity to rapidly capture, transcribe, and efficiently quality control numerical observations contained in archives.

The first data rescue dependency is being addressed in parallel by individual research efforts and coordinated international initiatives. Both approaches have uncovered new historical meteorological observation sources that have led to improved visibility, management, and curation of those data.¹⁸ Examples of international coordination efforts include the international Atmospheric Circulation Reconstructions over Earth (ACRE) initiative,^{19,20} the International Data Rescue (I-DARE) portal hosted by the World Meteorological Organization (WMO; <https://www.idare-portal.org/>), and the Copernicus C3S Data Rescue Service (<https://data-rescue.copernicus-climate.eu/>). The WMO I-DARE and Copernicus portals (<https://datarescue.climate.copernicus.eu/>) are currently being integrated into the same framework. Collectively, these efforts have improved the quality of recovered data, helped with resource sharing when capturing digital surrogates of original data sources, and reduced replication when obtaining archived meteorological data resources.¹⁷ The second data rescue dependency has been addressed either by individual researchers or research groups who manually transcribe historical data into digital format, or by using computer-aided recovery of text and numeric data (e.g., optical character recognition [OCR]).^{21,22}

The efficacy of the latter approach to date has, in a handful of trials, shown some promise but with significant limitations.^{21,22} However, progress to speed up transcription has been made using citizen science, which relies on individuals that are willing to voluntarily transcribe historical analogue meteorological data. There are many projects that have used this approach in recent years across a range of document types (see Ashcroft et al., 2016 for some Australasian examples). Pioneering efforts for handwritten tabulated observations are exemplified by OldWeather (www.oldweather.org), Meteororum ad Extremum Terrae (<http://Zooniverse.org/projects/acre-ar/meteororum-ad-extremum-terrae>), and Weather Rescue (www.weatherrescue.org).

In this descriptor article, we summarize our experiences from Southern Weather Discovery (SWD) (www.southernweatherdiscovery.org), a citizen science initiative hosted on

the Zooniverse web platform, to show how southern hemisphere meteorological time series have been generated and quality controlled from historical ship logbooks. This case study builds on prior work that has documented preparation of historical documents and transcription tactics,^{23–27} but also adds detail by explaining key elements of publicity and a media strategy plan that engendered public support for meteorological data rescue. We provide a stepwise account of our methods, highlighting some successes and pitfalls, that other researchers may benefit from to improve citizen science data rescue efforts for the geosciences. We provide details on the use of a data transcription interface on Zooniverse, preparation of historical documents for transcription, requirements for retrieving data from Zooniverse, and tactics to form a comprehensive observation dataset with minimal transcription errors. We also discuss serendipitous outcomes from SWD citizen science, where replicate keying of meteorological observations can be harnessed to improve artificial intelligence (AI) transcription of tabulated scientific data.

Launching a data rescue mission from the antipodes

A citizen science data rescue effort was launched as a component of the New Zealand Deep South National Science Challenge (DeepSouthChallenge.co.nz [DSC]) in 2015. The DSC's main aim is to understand the role of the Antarctic and Southern Ocean in determining New Zealand's future climate conditions and environmental outcomes from climate changes. The focus of DSC data rescue work was to recover undigitized weather observations and use them to help assess and evaluate the New Zealand Earth System Model (NZESM).²⁸ Many late 19th and early 20th century historical weather and climate events caused damage and disruption to New Zealand's civil infrastructure and economy (e.g., significant snowfall, floods, droughts).²⁹ Ensemble uncertainty in the 20CR analysis during the late 19th/early 20th century is large around New Zealand and the South Pacific (Figure 1), providing little insight into the atmospheric conditions leading up to these important past episodes. Thus, evaluating the quality of the NZESM using a reanalysis during those times is hampered. However, a few examples of full transcriptions and analyses of early handwritten scientific observations show potential to address this shortcoming.³⁰ Improving the efficacy of long-range reanalyses like 20CR with newly rescued data could enable further testing and more detailed validation of the NZESM, with direct applications toward improving our understanding of weather events, climate variability, and long-term changes.

To achieve data rescue aims that could contribute to the DSC, ACRE Antarctica (a chapter of ACRE) was created to focus on data rescue of historical meteorological observations within the high-latitude region bounded by Australia, South America, and Antarctica.²⁰ Within that geographic domain, there are key atmospheric and oceanic centers of action that are linked to modes of climate variability including the Pacific South American Mode (PSA),³¹ Zonal Wave 3 (ZW3),³² and the Southern Annular Mode (SAM)^{33,34} that directly and indirectly (via teleconnections) impinge on New Zealand's weather and regional climate conditions (Figure 1). In contrast to the broad continental expanse of the northern hemisphere and the tropics with numerous land-based weather observation stations, the ACRE

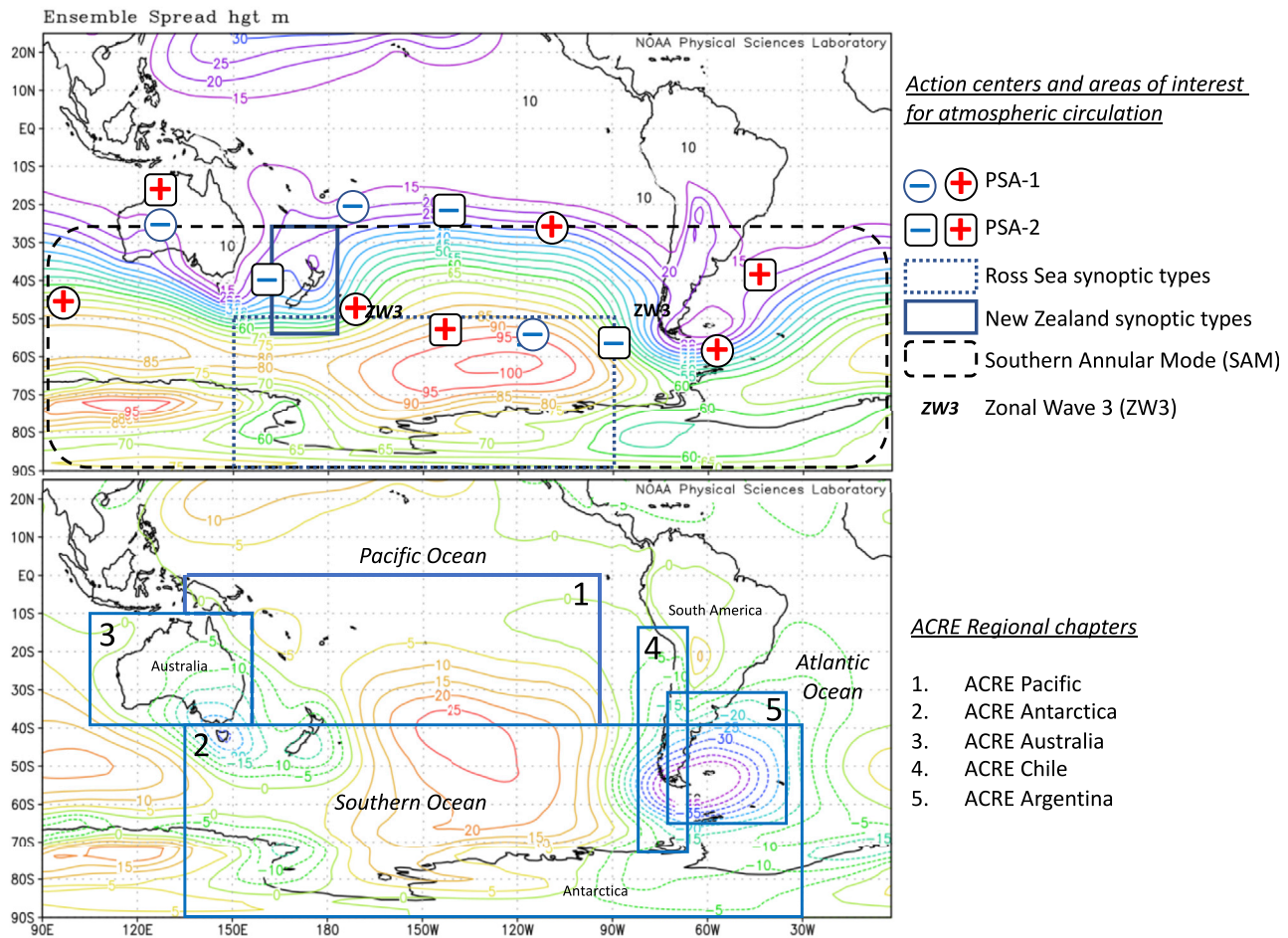


Figure 1. Twentieth Century Reanalysis ensemble spread in late 1800s and early 1900s

The Twentieth Century Reanalysis version 2c (20CRv2c) mean ensemble spread for the 1,000 hPa geopotential height (top) from 1891 to 1910 is contrasted with a spread anomaly plot (bottom) where the zonal (latitude average) mean for the same interval has been subtracted. This was the most recent version of 20CR that existed when Southern Weather Discovery (SWD) began. These plots show the effects of areas where there are relatively concentrated (e.g., New Zealand, Australia) and diminished (e.g., Amundsen Sea) observations. Outside of the south west Pacific tropics, and especially around Antarctica, there is higher uncertainty in the 20CR daily weather reconstruction. Areas where centers of action for modes of variability that affect New Zealand (including PSA, SAM, ZW3) are indicated. Plots are courtesy of National Oceanic and Atmospheric Administration Physical Science Laboratory (NOAA PSL). The spatial extent of the ACRE Antarctica regional chapter domain in the bottom panel was the focus of SWD data rescue.

Antarctica region is dominated by the South Pacific Ocean and Southern Ocean. Thus, historical observations are sparse and significant gaps in observation coverage over the southern hemisphere oceans produce large weather reconstruction uncertainties (Figures 1 and 2). This geographic predicament means any data rescue efforts need to consider limitations of coastal and maritime scientific data resources from lighthouses and seasonal coastal stations, and should harness the benefits of observations from harbor-based and ocean-going vessels. For the latter type of resource, ship log books have previously shown great potential to bolster maritime instrumental observations for the 19th and 20th centuries.^{35,36} Essential climate variables (ECVs), including atmospheric pressure, air temperature, sea surface temperature, and sea ice extent, were targeted for recovery in ACRE Antarctica’s initial funding support from the DSC. Additional financial support from the Copernicus Climate Change Service (C3S) further allowed us to refine citizen sci-

ence meteorological data rescue methods that have helped to streamline citizen science data capture of historical maritime weather observations.

Data Navigating happy hunting grounds for historical maritime weather data

Significant efforts have been made to photograph logbooks from ships that visited the southern hemisphere during the early to mid-20th century. Those resources are widely dispersed across a broad range of archives (see Teleti et al.³⁷ and Chappell et al.³⁸). For example, in the archives of New Zealand’s National Institute of Water and Atmospheric Research (NIWA), there are copies of published historical scientific expeditions to the Antarctic region written in English, French, Spanish, Portuguese, Russian, Norwegian, Finnish, and Swedish. Many of the original historical ship logbooks supporting those publications are held in European

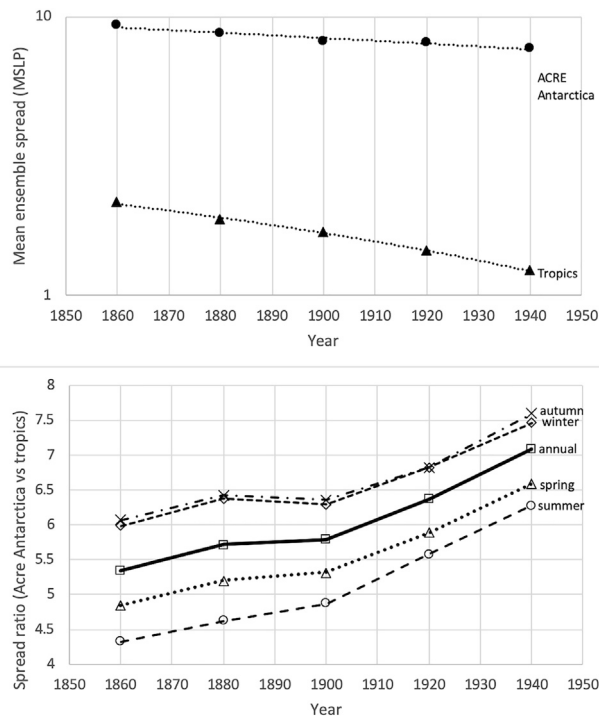


Figure 2. Change in selected 20CR uncertainty metrics through time (Top) The inter-region mean ensemble spread (uncertainty in daily reconstruction of weather) for the tropics and the ACRE Antarctica domain in the 20CR version 3 (20CRv3) January to March. It shows progressive improvement for both regions through time from the mid-19th to mid-20th century. (Bottom) The mean ensemble spread ratio (ACRE Antarctica mean ensemble spread divided by the tropics mean ensemble spread) is a dimensionless index indicating that, despite overall 20CR improvement, there is still lower uncertainty for past daily weather in the tropics (possibly a result of greater density or greater consistency of observations in that region) relative to the high southern latitudes. It also shows that distinct seasonal differences for the mean ensemble spread uncertainty are lowest for the southern high latitudes in summer and worst for autumn and winter.

archives and include British merchant and immigration ships that visited New Zealand, Australia, and the South Pacific.^{38,39} In addition, significant numbers of ship logbooks exist in Scandinavia^{40,41} related to whale hunting in the southern hemisphere.

An initial assessment spanning the period 1900–1960 indicates a minimum of seven million unique ship logbook weather observations for the high southern latitudes.^{40,41} In addition to regional data richness, data consistency is an important element to consider, given the sheer scale of processing and formatting citizen science data (see explanations below). Many logbooks have a standard printed table format that shipboard observers completed while at sea, and some shipboard expedition reports also include land-based observations from stationary and overland traverses.⁴² Our primary focus was then honed and directed at ship logbook data rescue from 1900 to 1950, a time frame that encompasses several severe weather events that affected New Zealand. Some 150,000 logbook images from more than 300 individual voyages were carefully photographed⁴³ and passed to NIWA in support of SWD to bolster mid-19th to mid-20th century sample depth

and reduce ensemble uncertainty in future 20CR iterations (Figure 2). Table 1 shows details for the log books that were digitized. The experimental procedures section outlines the process of data collection for this study. It describes how we established the SWD project identity and set up a data transcription platform on Zooniverse. This section also highlights progressive changes in data-rescue tactics and the approaches we deployed to recruit personnel who transcribed data from ship logbooks and land-based meteorological registers.

RESULTS

A total of 210 logs from voyages of 85 unique ships were transcribed in phase I of SWD (see Table 1 for details including years of coverage). Over 2,500 log book images were collectively obtained for those voyages, and 1,521 of those images were then selected for transcription. From the log book images that were used, 18,490 clips containing multiple meteorological observations were loaded to Zooniverse (taking into account that 16.6% of a grand total of 22,180 clips were blank and did not need to be transcribed). A grand total of 150,690 meteorological observations were recovered through replicate keying ($n_{\text{uncorrected barometer}} = 32,747$; $n_{\text{attached thermometer}} = 31,399$; $n_{\text{corrected barometer}} = 32,196$; $n_{\text{air temperature}} = 27,330$; $n_{\text{sea temperature}} = 27,018$). The total time for SWD phase I data capture was 9 months (running from October 2018 to July 2019), with a majority of transcribed observations obtained within the first 2 months from the project launch.

Charting a new course for streamlined data transcription

Determining what transcription retirement limit to use for individual observations was still an open-ended question when we launched SWD and after completing phase I. Replicated keying of logbook segments is designed to provide a majority consensus (and a measure of confidence through replication) that defines what numeric value exists in each table cell. Replicated keying levels for numeric values from an individual cell is proportional to time, but the effort to repeat keying as a way to increase confidence should have a functional limit. A choice of too few replicate keying attempts places the onus back on the researcher more frequently to re-classify questionable values that are not resolved via consensus. In turn, that can also generate re-work in terms of reposting logbook clips online to obtain additional transcriptions. We initially chose to have entries transcribed by 10 different volunteers during the first phase of SWD. This limit was increased initially from five entries after we discovered some problems with respect to the general data format returned by Zooniverse (see issues outlined below).

In SWD phase II, a goal was to determine optimal transcription and image clip retirement limits. Tabulated historical weather observations for eight ECVs (attached thermometer, uncorrected barometer, corrected barometer, maximum temperatures, minimum temperatures, wind direction, wind force, wind run) for the austral winter of 1939 (June, July, August) on original meteorological Form 301 paper copies taken held in NIWA's archive were digitally scanned from 63 stations spread across New

Table 1. Ship logbook meteorological observations recovered in SWD phase I hosted on Zooniverse

Unique ship	Unique logbook file name	Ship name	Log images	Images clipped	Number of clips	Clip not loaded	Year of data	Barometer uncorrected	Attached thermometer	Barometer corrected	Air temperature	Sea temperature	Total observations
1	MF911_44460_Athel Chief	Athel Chief	10	5	180	6	1946	226	200	229	0	0	655
2	MF911_19702_Bullyses	Bullyses	5	3	40	5	1930	77	78	77	59	59	350
2	MF911_19703_Bullyses	Bullyses	4	2	30	12	1930	47	47	47	47	47	235
3	MF911_32842_Cambridge	Cambridge	6	3	54	0	1935	53	53	53	53	52	264
4	MF911_27139_Canonosa	Canonosa	4	2	36	3	1933	36	35	31	36	36	174
4	MF911_28085_Canonosa	Canonosa	10	5	90	9	1933	172	173	157	173	173	848
4	MF911_29126_Canonosa	Canonosa	8	4	72	6	1934	82	82	82	81	80	407
4	MF911_29949_Canonosa	Canonosa	8	4	72	9	1934	72	71	71	72	72	358
5	MF911_24993_Coptic	Coptic	6	3	54	3	1932	66	65	66	65	65	327
5	MF911_27188_Coptic	Coptic	6	3	54	9	1933	59	60	59	60	60	298
5	MF911_32788_Coptic	Coptic	6	3	54	0	1935	61	63	63	60	61	308
5	MF911_34004_Coptic	Coptic	6	3	108	15	1936	58	58	58	58	58	290
5	MF911_37438_Coptic	Coptic	6	3	108	9	1937	64	61	63	0	0	188
6	MF911_27819_Cumberland	Cumberland	8	4	72	15	1933	65	64	64	66	66	325
6	MF911_37610_Cumberland	Cumberland	8	4	144	30	1937	71	72	72	2	2	219
6	MF911_39392_Cumberland	Cumberland	8	4	144	24	1938	68	68	68	65	66	335
7	MF911_17513_Deucalion	Deucalion	1	1	15	0	1929	21	0	37	37	37	132
7	MF911_17626_Deucalion	Deucalion	1	1	15	0	1929	20	0	27	27	27	101
8	MF911_17770_Devon	Devon	2	2	30	12	1929	42	42	38	42	41	205
8	MF911_22989_Devon	Devon	4	4	72	0	1931	173	176	169	176	175	869
9	ML911_2798_Discovery II	Discovery II	29	29	342	0	1950	450	451	450	0	0	1,351
10	MF911_19398_Dorington Courier	Dorington Courier	2	2	30	0	1930	73	73	73	72	73	364
11	MF911_35639_Dunedin Star	Dunedin Star	3	3	108	24	1936	53	55	53	51	47	259
12	MF911_39566_Durham	Durham	2	2	48	0	1938	107	107	107	107	107	535
12	MF911_41980_Durham	Durham	3	3	54	9	1939	117	117	116	116	117	583
12	ML911_900_Durham	Durham	11	11	231	27	1948	258	258	258	0	0	774
13	MF911_44536_Empire Victory	Empire Victory	11	11	198	0	1946	351	352	352	0	0	1,055
14	MF911_40649_Essex	Essex	3	3	108	18	1938	54	48	54	55	55	266
15	MF911_33390_Fordsdale	Fordsdale	3	3	54	0	1935	84	85	85	79	79	412
15	MF911_37896_Fordsdale	Fordsdale	8	4	105	39	1937	65	66	66	64	63	324
16	MF911_17310_Gloxinia	Gloxinia	4	2	30	0	1929	63	0	61	64	64	252
16	MF911_17474_Gloxinia	Gloxinia	4	2	30	0	1929	76	0	78	75	75	304
16	MF911_19208_Gloxinia	Gloxinia	2	1	15	2	1929	38	0	38	38	33	147
16	MF911_19287_Gloxinia	Gloxinia	8	3	45	12	1930	85	0	85	85	78	333
17	MF911_33108_Hertford	Hertford	8	4	72	9	1935	75	75	74	75	75	374

(Continued on next page)

Table 1. Continued

Unique ship	Unique logbook file name	Ship name	Log images	Images clipped	Number of clips	Clip not loaded	Year of data	Barometer uncorrected	Attached thermometer	Barometer corrected	Air temperature	Sea temperature	Total observations
17	MF911_37308_Hertford	Hertford	4	2	72	15	1937	39	37	38	0	0	114
18	MF911_20476_Hororata	Hororata	4	2	30	12	1930	48	48	48	48	48	240
19	MF911_25026_Huntingdon	Huntingdon	6	3	54	0	1932	67	68	64	68	68	335
19	MF911_25770_Huntingdon	Huntingdon	8	4	72	15	1932	72	70	73	73	73	361
19	MF911_26744_Huntingdon	Huntingdon	5	2	36	0	1932	40	43	43	43	43	212
19	MF911_27605_Huntingdon	Huntingdon	6	3	54	6	1933	56	55	58	59	60	288
19	MF911_30667_Huntingdon	Huntingdon	8	4	72	18	1934	68	68	67	68	68	339
19	MF911_35480_Huntingdon	Huntingdon	8	4	144	24	1936	72	72	72	69	69	354
19	MF911_37975_Huntingdon	Huntingdon	8	4	144	21	1937	78	79	80	78	78	393
20	MF911_34739_Hurunai	Hurunai	8	4	144	144	1936	148	0	146	147	146	587
21	MF911_34145_Ionic	Ionic	8	4	144	144	1936	115	117	115	112	112	571
22	ML_17955_Junie	Junie	26	20	160	160	1929	442	442	443	444	444	2,215
23	MF911_27575_Karamea	Karamea	6	3	54	54	1933	59	59	59	59	59	295
23	MF911_28249_Karamea	Karamea	8	4	72	72	1933	64	63	64	64	64	319
23	MF911_37224_Karamea	Karamea	6	3	54	54	1937	134	133	134	135	135	671
23	MF911_40779_Karamea	Karamea	7	3	108	108	1938	60	60	59	60	60	299
23	ML_18569_Karamea	Karamea	31	20	160	160	1932	441	439	441	434	435	2,190
23	ML_18678_Karamea	Karamea	29	21	168	168	1932	443	444	443	422	426	2,178
24	MF911_9415_Kiaora	Kia Ora	6	3	45	45	1925	90	0	0	90	89	269
25	MF911_44413_Lafonia	Lafonia	8	3	108	108	1946	41	41	41	0	0	123
26	MF911_39550_Loriga	Loriga	6	3	54	54	1938	121	121	120	121	118	601
27	MF911_39459_Losada	Losada	2	1	18	18	1938	30	30	30	30	30	150
28	MF911_32584_Mahana	Mahana	6	3	108	108	1935	60	53	53	59	58	283
29	MF911_26079_Mahia	Mahia	8	4	72	72	1932	74	75	74	75	75	373
29	MF911_27728_Mahia	Mahia	6	3	54	54	1933	52	52	52	52	52	260
29	MF911_37618_Mahia	Mahia	8	4	144	144	1937	71	71	71	0	0	213
29	MF911_43608_Mahia	Mahia	10	3	108	108	1946	55	55	52	50	49	261
30	ML_17885_Maimoa	Maimoa	29	21	168	13	1929	451	450	442	451	451	2,245
30	ML_18480_Maimoa	Maimoa	35	24	192	0	1931	529	523	527	407	356	2,342
30	ML_18660_Maimoa	Maimoa	32	23	184	10	1932	496	496	494	478	475	2,439
30	MF911_27208_Maimosa	Maimosa	8	4	72	6	1933	73	73	72	70	71	359
30	MF911_34685_Maimosa	Maimosa	8	4	144	16	1936	78	78	78	76	72	382
31	MF911_12243_Mamari	Mamari	4	2	30	0	1926/27	77	0	0	78	77	232
31	MF911_13623_Mamari	Mamari	6	3	45	0	1927	92	0	93	93	93	371
32	MF911_25506_Matakana	Matakana	6	3	54	0	1932	68	68	66	68	68	338
32	MF911_27290_Matakana	Matakana	6	3	54	0	1933	71	71	71	71	71	355

(Continued on next page)

Table 1. Continued

Unique ship	Unique logbook file name	Ship name	Log images	Images clipped	Number of clips	Clip not loaded	Year of data	Barometer uncorrected	Attached thermometer	Barometer corrected	Air temperature	Sea temperature	Total observations
32	MF911_28274_Matakana	Matakana	8	4	72	9	1933	81	81	80	81	81	404
32	MF911_33110_Matakana	Matakana	8	4	72	15	1935	71	71	71	72	71	356
32	MF911_9992_Matakana	Matakana	4	2	30	0	1925	76	0	0	77	77	230
32	ML_17869_Matakana	Matakana	30	23	184	20	1929	470	463	465	467	465	2,330
33	MF911_27403_Middlesex	Middlesex	10	5	90	33	1933	138	137	138	138	138	689
34	ML_18676_Norfolk	Norfolk	29	21	168	0	1932	442	446	442	436	429	2,195
34	ML911_958_Norfolk	Norfolk	36	10	210	21	1948	233	234	231	0	0	698
35	MF911_44234_Northumberland	Northumberland	13	7	252	21	1947	292	291	296	0	0	879
36	MF911_23534_Opawa	Opawa	6	3	54	3	1931	64	64	64	64	64	320
36	MF911_26440_Opawa	Opawa	8	4	72	0	1932	62	64	61	64	64	315
36	MF911_27357_Opawa	Opawa	8	4	72	24	1933	63	64	63	64	64	318
36	MF911_28505_Opawa	Opawa	8	4	72	15	1933	70	71	70	70	71	352
36	MF911_29526_Opawa	Opawa	6	3	54	6	1934	63	63	63	63	63	315
36	MF911_30783_Opawa	Opawa	8	4	72	21	1934	66	66	66	66	66	330
36	ML_18575_Opawa	Opawa	31	19	152	0	1932	428	424	428	428	426	2,134
37	MF911_27680_Orari	Orari	6	3	54	3	1933	56	56	50	54	54	270
37	MF911_30480_Orari	Orari	4	2	36	0	1934	33	34	34	34	34	169
37	ML911_527_Orari	Orari	22	12	252	30	1947	262	262	257	0	0	781
37	ML911_81_Orari	Orari	31	12	252	30	1947	240	240	238	0	0	718
38	ML_18115_Otaki	Otaki	39	29	232	36	1929	562	563	563	456	426	2,570
39	MF911_10588_Otira	Otira	4	2	30	0	1926	40	40	39	40	40	199
39	MF911_26306_Otira	Otira	6	3	54	0	1932	77	77	72	79	79	384
39	MF911_28021_Otira_DUP	Otira	8	4	72	9	1933	79	79	79	79	79	395
40	MF911_29047_Pakeha	Pakeha	10	5	90	9	1933-34	88	89	89	89	89	444
40	ML_17655_Pakeha	Pakeha	29	22	176	14	1928	467	467	467	459	462	2,322
40	ML_17804_Pakeha	Pakeha	37	29	232	20	1928	578	581	577	580	578	2,894
40	ML_18410_Pakeha	Pakeha	32	21	168	10	1931	453	455	455	336	333	2,032
41	MF911_25239_Piako	Piako	8	4	72	6	1932	78	78	78	77	77	388
41	MF911_27502_Piako	Piako	8	4	72	9	1933	74	75	75	74	75	373
42	MF911_25348_Port Adelaide	Port Adelaide	9	5	90	1	1932	167	163	168	153	153	804
42	MF911_25349_Port Adelaide	Port Adelaide	2	1	18	0	1932	25	29	29	0	0	83
42	MF911_33271_Port Adelaide	Port Adelaide	8	4	72	0	1935	69	70	71	69	71	350
42	MF911_35042_Port Adelaide	Port Adelaide	14	7	252	23	1936	140	141	140	134	134	689
42	ML_18174_Port Adelaide	Port Adelaide	35	19	152	0	1930	417	420	414	410	383	2,044
43	MF911_25998_Port Alma	Port Alma	8	4	72	0	1932	75	76	76	76	76	379
43	MF911_27851_Port Alma	Port Alma	8	4	72	15	1933	70	69	69	71	62	341

(Continued on next page)

Table 1. Continued

Unique ship	Unique logbook file name	Ship name	Log images	Images clipped	Number of clips	Clip not loaded	Year of data	Barometer uncorrected	Attached thermometer	Barometer corrected	Air temperature	Sea temperature	Total observations
43	ML_18499_Port Alma	Port Alma	32	20	160	0	1931	422	421	425	376	382	2,026
43	ML_18587_Port Alma	Port Alma	37	21	168	0	1932	463	465	463	452	445	2,288
44	MF911_32068_Port Auckland	Port Auckland	6	3	108	0	1935	66	66	66	66	52	316
44	MF911_41786_Port Auckland	Port Auckland	12	6	108	0	1938	130	130	129	131	131	651
44	ML_17895_Port Auckland	Port Auckland	30	23	184	18	1929	473	474	475	475	475	2,372
44	ML_18144_Port Auckland	Port Auckland	30	21	168	0	1930	452	455	453	368	369	2,097
45	MF911_16080_Port Bowen	Port Bowen	4	2	30	0	1928	76	0	0	76	76	228
45	MF911_32187_Port Bowen	Port Bowen	8	4	136	34	1935	58	53	57	59	60	287
45	MF911_33308_Port Bowen	Port Bowen	6	3	54	0	1935	65	65	65	65	61	321
45	MF911_34293_Port Bowen	Port Bowen	8	4	144	38	1936	64	64	64	64	64	320
45	MF911_35486_Port Bowen	Port Bowen	4	2	72	0	1936	47	47	47	47	47	235
46	ML_17849_Port Campbell	Port Campbell	28	22	176	16	1929	461	460	457	458	459	2,295
47	MF911_26975_Port Caroline	Port Caroline	10	4	72	6	1933	82	83	83	83	83	414
47	ML_18260_Port Caroline	Port Caroline	32	21	168	0	1930	436	433	434	400	401	2,104
47	ML_18565_Port Caroline	Port Caroline	38	23	184	0	1932	479	478	482	468	467	2374
48	MF911_31626_Port Chalmers	Port Chalmers	8	4	117	0	1935	71	59	73	73	72	348
48	MF911_37654_Port Chalmers	Port Chalmers	6	3	108	11	1937	62	62	62	60	61	307
49	MF911_25515_Port Darwin	Port Darwin	8	4	72	9	1932	70	69	71	68	66	344
49	MF911_35675_Port Darwin	Port Darwin	6	3	108	3	1936	65	65	65	65	48	308
49	MF911_37099_Port Darwin	Port Darwin	8	4	144	0	1937	70	70	71	0	0	211
49	MF911_39424_Port Darwin	Port Darwin	8	4	144	30	1938	68	68	68	68	68	340
50	MF911_13131_Port Denison	Port Denison	4	2	30	0	1927	79	0	78	79	79	315
50	MF911_29212_Port Denison	Port Denison	8	4	72	9	1934	76	76	76	76	76	380
50	MF911_34150_Port Denison	Port Denison	8	4	144	27	1936	76	67	74	75	74	366
50	MF911_35293_Port Denison	Port Denison	8	4	144	18	1936	81	66	80	84	79	390
51	MF911_41942_Port Dunedin	Port Dunedin	14	7	126	12	1939	136	131	130	137	135	669
51	ML_18673_Port Dunedin	Port Dunedin	29	20	160	0	1933	416	415	415	365	360	1,971
52	MF911_36019_Port Fremantle	Port Fremantle	6	3	114	0	1936	114	114	114	0	0	342
52	ML_18558_Port Fremantle	Port Fremantle	35	19	152	0	1932	422	419	420	406	410	2,077
52	ML_18630_Port Fremantle	Port Fremantle	31	21	168	8	1932	470	470	468	459	441	2,308
52	ML_18680_Port Fremantle	Port Fremantle	30	21	168	0	1932	453	452	453	437	441	2,236
53	ML_18476_Port Gisborne	Port Gisborne	33	18	144	0	1931	422	421	422	274	252	1,791
53	MF911_32744_Port Gisborne	Port Gisborne	6	3	108	0	1935	70	70	70	67	68	345
53	MF911_34915_Port Gisborne	Port Gisborne	6	3	108	6	1936	54	54	56	54	54	272
53	MF911_39080_Port Gisborne	Port Gisborne	6	3	108	9	1938	62	64	64	63	66	319
53	MF911_40057_Port Gisborne	Port Gisborne	6	3	108	3	1938	66	65	64	66	65	326

(Continued on next page)

Table 1. Continued

Unique ship	Unique logbook file name	Ship name	Log images	Images clipped	Number of clips	Clip not loaded	Year of data	Barometer uncorrected	Attached thermometer	Barometer corrected	Air temperature	Sea temperature	Total observations
54	MF911_33928_Port Hobart	Port Hobart	6	3	54	0	1936	70	71	71	71	71	354
54	MF911_34904_Port Hobart	Port Hobart	6	3	108	0	1936	70	71	67	64	65	337
54	MF911_35996_Port Hobart	Port Hobart	6	3	108	0	1936	70	69	70	66	66	341
55	ML_18639_Port Hunter	Port Hunter	34	24	192	20	1932	498	503	504	345	365	2,215
56	MF911_40199_Port Jackson	Port Jackson	10	5	180	12	1938	94	98	98	95	96	481
57	ML_17977_Port Melbourne	Port Melbourne	32	24	192	22	1929	474	477	477	479	479	2,386
58	MF911_11208_Port Napier	Port Napier	6	2	30	0	1926	43	41	43	43	42	212
59	ML_17873_Port Nicholson	Port Nicholson	31	23	184	18	1929	479	480	479	478	480	2,396
59	ML_18399_Port Nicholson	Port Nicholson	30	21	168	12	1931	448	446	439	418	407	2,158
60	ML_18155_Port Sydney	Port Sydney	33	22	176	0	1930	443	446	443	417	416	2,165
61	MF911_41432_Port Townville	Port Townsville	6	3	108	18	1938	56	56	53	0	0	165
62	ML_17860_Port Victor	Port Victor	35	28	224	36	1929	541	540	540	539	537	2,697
63	MF911_23307_Port Wellington	Port Wellington	8	4	72	12	1931	151	1	152	152	153	609
63	MF911_27086_Port Wellington	Port Wellington	8	4	72	9	1933	150	150	150	152	152	754
63	MF911_28051_Port Wellington	Port Wellington	8	4	72	12	1933	84	84	86	86	85	425
63	MF911_37821_Port Wellington	Port Wellington	7	3	108	6	1937	51	51	51	45	45	243
64	MF911_33984_Port Wyndham	Port Wyndham	6	3	108	0	1936	55	55	55	2	2	169
64	MF911_36181_Port Wyndham	Port Wyndham	6	3	108	21	1936	47	47	47	46	46	233
64	MF911_39352_Port Wyndham	Port Wyndham	6	6	3	108	1938	44	44	44	37	32	201
64	MF911_40313_Port Wyndham	Port Wyndham	6	3	108	15	1938	44	44	43	41	41	213
65	MF911_41898_Reina del Pacifico	Reina del Pacifico	4	2	36	3	1939	5	5	66	66	66	208
66	ML_17827_Rimutaka	Rimutaka	42	34	272	26	1929	705	689	705	708	701	3,508
67	ML_17998_Runpenu	Ruapehu	36	25	200	26	1929	496	495	495	458	452	2,396
68	ML_18579_Somerset	Somerset	35	22	176	0	1932	458	459	460	435	439	2,251
68	ML_18646_Somerset	Somerset	38	21	168	6	1932	457	458	456	274	242	1,887
69	MF911_44525_Southern Harvester	Southern Harvester	4	2	72	27	1947	51	52	53	0	0	156
70	MF911_18465_Southern King	Southern King	2	1	15	0	1929	38	38	0	38	11	125
70	MF911_18548_Southern King	Southern King	2	1	15	9	1929	13	0	0	13	0	26
70	MF911_18819_Southern King	Southern King	2	1	15	9	1929	13	0	0	13	0	26
70	MF911_18978_Southern King	Southern King	2	1	15	9	1929	13	0	0	13	12	38
70	MF911_18979_Southern King	Southern King	2	1	15	9	1929	16	15	0	15	10	56
70	MF911_21761_Southern King	Southern King	6	3	54	16	1930	95	79	0	21	20	215
71	ML911_1542_Struan	Struan	24	14	294	55	1947	15	0	142	0	0	157
72	MF911_43636_Suffolk	Suffolk	8	3	108	12	1946	121	120	121	0	0	362
73	MF911_10646_Tairoa	Tairoa	4	2	30	0	1926	69	69	69	69	69	345

(Continued on next page)

Table 1. Continued

Unique ship	Unique logbook file name	Ship name	Log images	Images clipped	Number of clips	Clip not loaded	Year of data	Barometer uncorrected	Attached thermometer	Barometer corrected	Air temperature	Sea temperature	Total observations
73	MF911_30690_Tairoa	Tairoa	8	4	72	6	1934	78	78	78	78	78	390
73	MF911_35417_Tairoa	Tairoa	8	4	144	33	1936	66	69	66	66	69	336
73	MF911_37729_Tairoa	Tairoa	8	4	144	24	1937	77	77	76	76	75	381
73	MF911_8723_Tairoa	Tairoa	6	3	45	9	1925	90	90	89	90	90	449
74	MF911_25999_Taranaki	Taranaki	6	3	54	0	1932	143	143	143	142	141	712
74	MF911_26898_Taranaki	Taranaki	6	3	54	0	1933	67	67	65	67	67	333
74	MF911_27674_Taranaki	Taranaki	6	3	54	0	1933	69	69	68	69	69	344
74	MF911_30161_Taranaki	Taranaki	6	3	54	0	1934	64	64	63	64	64	319
74	MF911_42691_Taranaki	Taranaki	12	6	216	3	1939	120	133	125	134	133	645
74	ML_18585_Taranaki	Taranaki	30	19	152	0	1932	379	383	385	331	319	1,797
75	MF911_22425_Tasmania	Tasmania	8	4	72	9	1931	78	78	78	78	78	390
75	MF911_26731_Tasmania	Tasmania	9	4	72	0	1932	107	107	107	107	107	535
75	MF911_27816_Tasmania	Tasmania	10	5	90	15	1933	122	122	122	123	123	612
76	ML911_2149_Thule	Thule	37	28	333	0	1950	430	434	436	0	0	1,300
77	MF911_37834_Tongariro	Tongariro	8	4	144	36	1937	68	68	68	69	69	342
77	MF911_42032_Tongario	Tongariro	12	6	108	3	1938	134	134	134	134	134	670
77	ML_18641_Tongariro	Tongariro	28	19	152	1	1932	428	429	429	332	353	1,971
78	MF911_44607_Trepassey	Trepassey	8	4	72	6	1946	76	75	9	0	0	160
79	MF911_35419_Tuscan_Star	Tuscan Star	6	3	108	18	1936	56	57	55	57	59	284
80	MF911_12609_Verbania	Verbania	4	2	30	0	1927	47	0	0	47	47	141
81	MF911_10751_Waimana	Waimana	6	3	45	0	1926	80	0	0	80	80	240
82	MF911_32663_Waipawa	Waipawa	4	2	72	0	1935	42	42	42	41	41	208
82	MF911_34853_Waipawa	Waipawa	10	5	180	12	1938	105	102	103	0	0	310
82	ML911_988_Waipawa	Waipawa	39	12	252	48	1948	232	232	232	0	0	696
83	MF911_35087_Waiwera	Waiwera	6	3	108	9	1936	64	64	64	63	63	318
83	MF911_36265_Waiwera	Waiwera	6	3	108	12	1936	61	61	61	61	61	305
83	MF911_39360_Waiwera	Waiwera	6	3	108	17	1938	58	59	57	56	59	289
84	MF911_32763_Westmoreland	Westmoreland	6	3	108	0	1935	135	135	135	135	135	675
85	MF911_24120_Zealandic	Zealandic	6	3	54	0	1931	142	142	142	142	142	710
85	MF911_25112_Zealandic	Zealandic	6	3	54	0	1932	139	139	139	139	139	695
85	MF911_25805_Zealandic	Zealandic	6	3	54	0	1932	129	129	129	129	129	645
85	MF911_30311_Zealandic	Zealandic	6	4	72	3	1934	81	81	81	85	85	413

A total of 150,690 observations from 85 unique ships that embarked on 210 voyages were successfully captured by replicate keying from citizen scientists. Log images are the total number of digital files that correspond to the unique logbook file, images clipped are the total number of images that had data within the ACRE Antarctica domain (see [Figure 1](#)) that were selected for processing, the number of clips are the total segment number that were extracted from all pages selected for processing, and blanks not loaded are the number of clips that had no data. Total number of recovered observations for each category for each unique voyage (barometric pressure, air temperature, and sea temperature) are shown. Total number of unsuccessful transcriptions not shown.

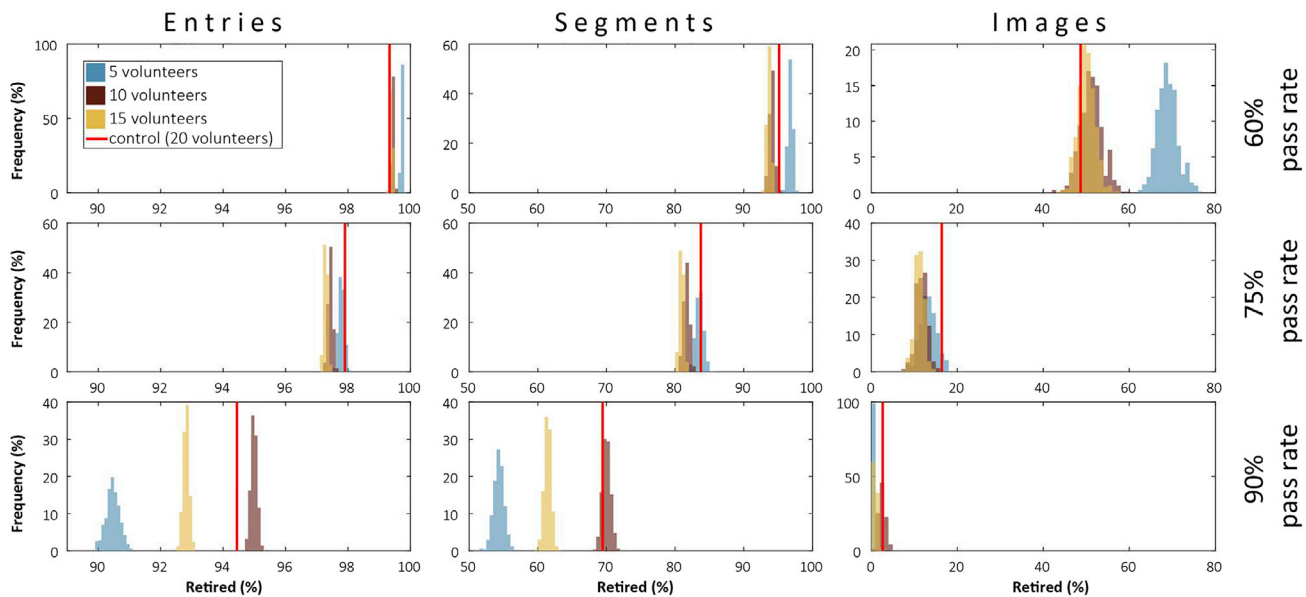


Figure 3. WISE consensus results

Frequency of successful consensus classification using different thresholds of agreement for individual entries, meteorological form segments (clips), and entire meteorological form images pooled from unique land-based stations that were transcribed the WISE phase of work on SWD. See [supplemental information](#) for more details about the number of data points, segments, and images that comprise these statistics.

Zealand to cover the winter season when the 1939 Week it Snowed Everywhere (WISE) event occurred.

A brute-force approach was employed by setting SWD transcription retirement limits at 20 for WISE, which was twice the sample pool of SWD phase I transcription. Using these data, we were able to use hierarchical degradation that progressively lowered replicate transcription sample depth of keyed values in order to evaluate optimal data keying retirement limits. Retirement statistics (completed successful transcription) were assessed for individual entries (each individual observation recorded in a log book clip), segments (the log book clips), and entire logbook images (with multiple segments that contain multiple entries). We considered results from our entire pool of volunteers ($n = 20$), the control dataset, to evaluate the effects of transcription sample depth degradation. The 20-volunteer sample depth also allowed us to gather a large enough dataset to evaluate type 1 (consensus acceptance of an incorrect value; false-positive/acceptance) and type 2 (non-consensus and rejection of a value that was legitimate; false-negative/rejection) errors. We also used the WISE dataset to evaluate minimum number of repeat classifications needed to obtain a 100% complete dataset via majority consensus with minimal transcription errors. Some examples of log books that had the most common errors are provided in the [supplemental information](#).

The percentage of entries, segments, and images that were “retired” (i.e., consensus reached, with citizen science transcription considered a success) decreased for all replicate keying tests conducted on each of the hierarchical transcription classes (5, 10, 15, and 20 volunteers) when the pass rate threshold was raised progressively from 60% to 90% (Figure 3). Results for all the hierarchical classes appeared most similar for entries, segments, and images for the 75% pass rate threshold and were the most different for the 90% pass rate test. The difference between suc-

cess of the five-volunteer class within the 60% pass rate test and the 90% pass rate test resulted from the fact that all five answers need to align for the latter to be considered a success, while the former only requires three out of five to be right. The results for 10 versus 20 volunteers in the 90% pass rate test also appeared similar. Few appreciable differences were also observed in the 60% pass rate test for the 10, 15, and 20 volunteer classes.

We evaluated the probability of type 1 and type 2 errors by comparing expert-guided transcriptions of original logbook entries with the consensus values obtained through WISE. This experiment used tests based on several draws of 20 entries at random for each of the eight WISE data entry tasks, and just over 46,200 values constituted the pool that could be analyzed to assess errors associated with data entry. Across the entire dataset, 56% of entries had a low risk of error, 7% had a medium risk, 1% had a high risk, and 36% were blank. Blank and failed consensus entries were automatically excluded from these random draws. Each draw was evaluated with respect to consensus keying based on either a threshold of agreement (termed O75, 75% consensus, 15 of 20 values; O60, 60% consensus, 12 of 20 values) or by selecting the first five or 10 keyed responses (O5, O10) out of the 20 selected values. An additional test, termed “output resampled” (ORS), added a step to the O60 consensus processing with a random draw for entries that failed to reach consensus as a way to reach a definitive result. Each of the failed entries from this test had a statistical mode calculated from 500 iterations that individually pulled a five-sample random draw from the pool of 20 entered values (Table 2). We further classed type 1 and type 2 errors in each of these tests across three categories of keying success with respect to whether there was an increased likelihood of either error occurring (with the *a priori* assumption this would be strongly linked to the quality of the uploaded image on SWD). These

Table 2. Type 1 and type 2 errors associated with the WISE hierarchical degradation and resampling tests (O75, O60, O5, O10, and O-RS)

	Low risk			Medium risk			High risk			Blank cells			Whole set		
	T1	T2	Correct	T1	T2	Correct	T1	T2	Correct	T1	T2	Correct	T1	T2	Correct
O75	0	0	100	0	13	87	0	89	11	0	0	100	0.00	1.83	98.17
O60	0	0	100	0	5	95	5	62	34	0	0	100	0.05	0.96	98.99
ORS	0	0	100	0	0	100	16	0	84	0	0	100	0.16	0.00	99.84
O5	0	0	100	3	0	97	26	30	44	0	0	100	0.51	0.30	99.19
O10	0	0	100	0	3	97	12	49	39	0	0	100	0.12	0.69	99.19

The percentages for each of these risk categories was calculated by weighting by the proportion of composition for the entire dataset by the percentage correct in that particular category (low, medium, high) in order to obtain a percentage error and percentage correct whole-set results. These results represent the aggregate for all entry types (pressure, temperature, and wind) across eight tasks. More details about this experiment can be found in the [supplemental information](#).

categorical tests spanned low-risk images (consensus pass rate = 100%; clear penmanship, no edits in the original cell), medium-risk images (consensus pass rate <80%), and high-risk images (consensus not reached; often associated with edited original tabulated entries or messy penmanship).

Blank cells were identified correctly in all tests. For the low-risk image category, type 1 and type 2 errors were absent, but type 1 errors slightly increased and more so for type 2 errors in medium-risk images (Table 2). For high-risk images, all of the tests except O75 revealed type 1 errors, and there were no type 2 errors associated with the ORS analysis. The most common incorrect transcription issues had to do with (1) confusion between 4s and 7s and 4s and 6s; (2) omission of decimal points or other delimiters; and (3) extraneous notes, arrows, or values in cells where original data had been manually corrected (i.e., crossed out and rewritten).

Training machines to guide the data rescue ship

The WISE dataset was also used to independently test Microsoft Read API (Figure 4) using digital photograph surrogates of the 1939 Form 301s that contained the original analogue data. One advantage with Microsoft Read API is the ability to transcribe an entire sheet using computer vision, which can save research preparation time related to clipping and uploading segments of a page onto Zooniverse for citizen science transcription. Six high-resolution scans of full original Form 301 sheet data sheets from two stations were used for a Microsoft Read API preliminary test, which draws on an OCR engine based on deep learning algorithms.^{44–46} A Microsoft Excel template indicating the position of data on the page (row cell and column) was also provided to the Microsoft team for supplying values back to NIWA for validation.

The results from OCR using Microsoft Read API indicate variable efficacy between different observing sites and for different observation types (Table 3). Across five quantitative observation categories (attached thermometer, barometer uncorrected, barometer corrected, maximum temperature, minimum temperature), the Microsoft Read API validation grand strike rate was 69% ± 15% (n = 920). Results for transcribing ECVs were also site dependent (related to penmanship of the observer who filled in the data table).

Extraneous formatting and errors related to decimals and dashes were not considered when validating the Microsoft Read API because they are minor (see Table 3; difference be-

tween uncorrected and potential corrected strike rate for machine learning transcription). The most common issues identified where the Microsoft Read API auto-transcription did not validate related to incorrect transcription of the first digit of a numeric string, and designation of a letter where a number actually occurred. The most common digits that were not transcribed correctly were 4s and 7s (often swapped). Both of those shortcomings are similar to issues that we experienced on SWD for citizen scientists keying in data for the WISE experiment. Additional simplified guidance for unsupervised machine learning algorithms could be applied in those cases (e.g., pressure values recorded in inches of mercury must begin with a 2 or 3) to improve strike rate results for the Microsoft Read API (Table 3).

DISCUSSION

Consolidating lessons learned from the SWD data rescue journey

Improving our understanding of past weather events and the roles that modes of variability have played in guiding extreme conditions requires better reanalyses, and in particular the coverage for the southern high latitudes needs to be dramatically augmented (Figure 1). There is massive potential to improve global reanalyses using the troves of historical meteorological data that are stored in a wide range of archives.^{17,18,35,47,48} These observations can be digitized by volunteers with assistance from scientists who can prioritize and arrange data rescue activities. A major advantage to using Web-based citizen science for data rescue efforts is that the human resource can be drawn from all regions on Earth, volunteer time is free, and progress is made more or less continuously. In addition, citizen science data rescue provides an opportunity to engage and educate the general public about the importance of long-term meteorological observations and climate change.⁴⁹ In SWD, we learned that when data rescue is conducted under the auspices of a global effort like ACRE,^{19,50} and with support from agencies like the World Meteorological Organization⁵¹ and Copernicus Climate Change Service,^{52,53} it engenders increased regional responsibility for data stewardship and archives while raising the profile of the science. This typically has a positive feedback for conducting additional data rescue activities, particularly in remote and under-resourced regions.⁵⁴ In addition, there are improvements for transparency of nation- and archive-specific data holdings that engenders wider data

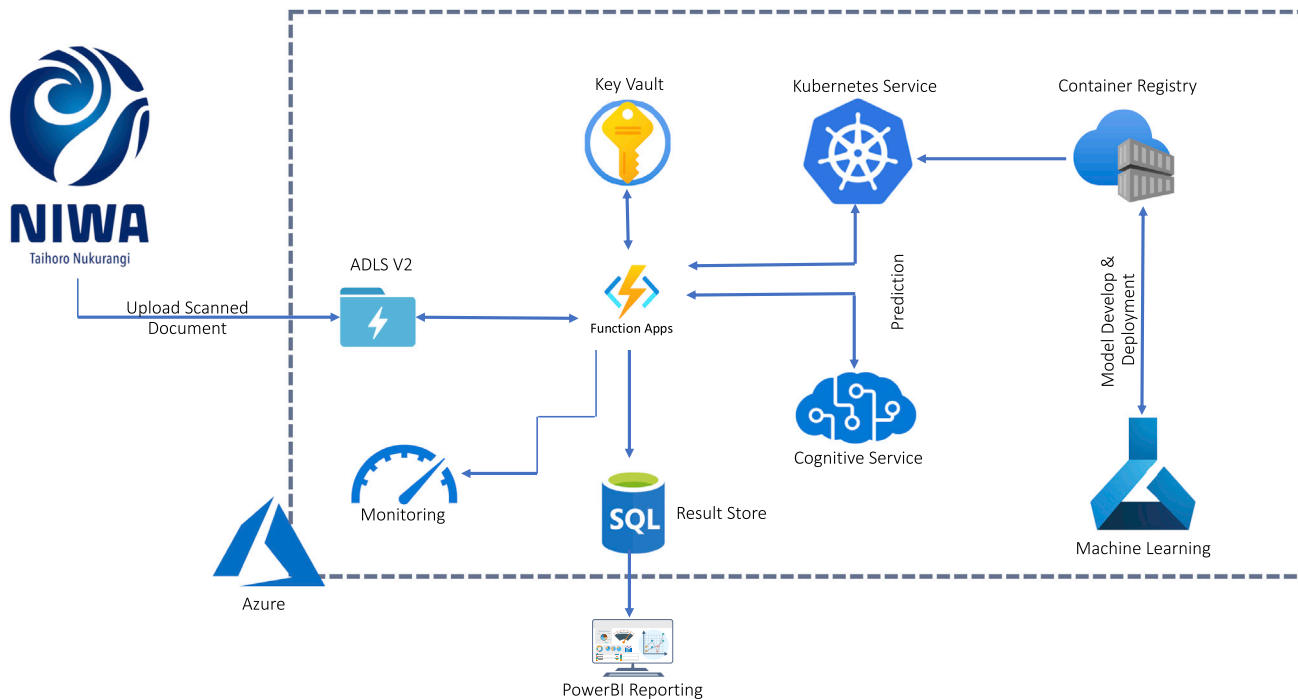


Figure 4. Azure OCR pipeline

Generalized architecture of the automated Azure cloud computing pipeline hosted by Microsoft that was used for the WISE OCR and transcription experiment. Handwritten meteorological tables in portable document file (PDF) format were transferred to Microsoft and loaded to the Azure Data Lake Storage (ADLSv2), where a Function Apps code forwarded them for text extraction. The Read API Azure Cognitive Service was used to extract handwritten digits from each PDF, in conjunction with custom machine learning models deployed using the Azure Kubernetes service via the Azure Container Registry. The custom model removed noise from the digital surrogates and located cells with digits in them. The extracted components from each page were further processed and the final outcome from OCR analysis was stored in the Azure SQL database (Result Store) where they were accessed, analyzed, and visualized using Power BI. In addition, capabilities for inter-service communication were securely held in Key Vault.

sharing that can exceed what *ad hoc* efforts undertaken by isolated researchers have achieved in the past. It is also clear that automated OCR approaches, like those we tested using Microsoft Read API, could be greatly improved with using the vast data captured through citizen science efforts like SWD.

The SWD core team that undertook the tasks required to capture handwritten observations using Zooniverse consisted of nine people. Our team members collectively found and captured digital twins of data sheets in multiple archives, prepared them for transcription on the Web platform, retrieved/parsed replicate keyed observations, and undertook statistical analyses of the results. Each of these data rescue tasks does not constitute an equivalent time investment or skill level. We also obtained external support from national and international collaborators to achieve many of our aims (e.g., finding ship log books in archives, testing machine learning OCR transcriptions). We divided basic data rescue tasks between senior scientists, casual staff, and students in order to maximize the use of limited funding. Overall, the foundation for a data rescue project like ours could be run on 1.0 full-time equivalent (FTE) employment. However, it is likely that multiple years would be required if one person were to do all of the associated tasks, including the field work. This type of effort also requires a broad enough skill set that includes development, adaptation, or augmentation of code that automates tasks through scientific programming. In addition, support from professional media ex-

perts would be required to attain the level of external project promotion we achieved.

The benefits of crowd-sourcing labor to key historical observations are partially offset by some unique challenges. A significant investment of time is required to train personnel in how to manually clip the logbook images or to set up different workflows for logbooks that are printed in different formats. This echoes findings learned from citizen science efforts to key United Kingdom Met Office daily weather reports, where it was noted that the effort required to clip segments of images and provide them using consistent formatting for end-user context places an additional time burden on the research team.⁵⁵ Automated clipping routines we tested reduced the time investment for this specific data rescue step, but success is highly dependent on the quality of photography and the types of scientific data tables being rescued. We are aware that the efforts from ACRE Argentina at present are using clipping approaches that focus on single cells and providing them in Zooniverse without formatting, which is a potential time-saving measure (<https://www.zooniverse.org/projects/acre-ar/meteororum-ad-extremum-terrae>). For SWD phase one, all the logbooks that were not in a consistent format were omitted.

An issue related to consistency of data transcription from international audiences can also arise. We noted that dashes and decimals were commonly substituted with commas or used as a delimiter, making our post-transcription data retrieved

Table 3. Results from Microsoft Read API for the WISE

	Attached thermometer	Barometer	Barometer corrected	Minimum temperature	Maximum temperature
Grand strike rate, uncorrected (%)	65.1	77.1	69.0	64.2	70.7
Grand strike rate, potential correction (%)	81.5	80.1	76.0	78.7	80.8

Strike rate (percentage correct) across five meteorological variables transcribed by Microsoft Read API for Albert Park (A64871) and Christchurch (H32561) spanning June to August 1939. The potential corrected grand strike rate is corrected for any miss related to a decimal or a dash that was not captured in the automated transcription.

from native Zooniverse outputs difficult. There were also significant discrepancies related to the citizen science transcription of ship coordinates that led us to eventually input that category manually using an expert team. Significant time was also required to respond to questions from volunteers (particularly in the early stages following initial project launch).

Scanning the horizon for fair winds and smooth data rescue sailing

Based on the outcome of the WISE experiment tests, we consider a compromise can be reached between time spent keying by citizen science volunteers and achieving completeness and accuracy of a transcribed dataset when eight replicate entries are employed. To achieve that, we recommend initially setting a minimum 60% pass rate threshold (five out of eight in agreement) and then using a resampling scheme for any values that did not reach consensus. Using that scheme, we would expect that type 2 errors would be absent from the transcribed data, and type 1 errors would be, on average, less than two in 1,000. In addition, the transcribed dataset will be 100% complete and close to 99.5% accurate.

It is also worth noting that these results are dependent upon the nature of the data being transcribed. The specific retirement limit and broader strategy employed for scientific data transcription may need to be adjusted based on the type of observations being rescued. Integer values with no decimal points are the most straightforward to key and require little repetition to ensure a correct consensus value. Conversely, alphanumeric values and values with many significant figures introduce more complexity or opportunity for variation among responses from the citizen scientists (e.g., representing a decimal with a period, a comma, a space, or ignoring the decimal altogether). For example, within our dataset, we observed significantly more errors in the temperature fields, which generally include decimal points, than the pressure (integer-only values) and wind run (alphabetic-only values) fields.

As such, ironing out idiosyncrasies that can make data rescue efforts through Zooniverse universal and successful requires the following minimum requirements:

- Prepare scans of data tables in a way that enables efficient keying and that is easy to understand.
- Test and re-test workflows to ensure they are simple to follow (heeding participant feedback).
- Design tasks so that citizen scientists have the best chance of entering a correct result.
- Evaluate initial inputs and data retrievals with a small dataset before launching a full data rescue campaign.
- Optimize replicate keying levels to balance confidence of results with time invested from citizen scientists.

- Prepare enough material in advance to ensure momentum can be continually maintained.

Promotion of our project and engagement with media and the public was strongly connected to the rate of retirement for logbook segments and the overall success of completing the recovery of meteorological data via SWD. Our approach kept the following in mind:

- A strong communications strategy with a “hook” to get people involved.
- Willingness to engage with the media and the project participants.
- Promotion of the project on multiple social media platforms.
- Repeated contact with the citizen science community using emails and updates as tasks progressed.

Data rescue on Zooniverse has a proven successful track record for several projects that have focused on the recovery of historical weather observations. Our approach for SWD is something that can be easily replicated for other disciplines where tabulated scientific data need to be transcribed. We recently provided training to assist the launch of the Climate History Australia project (<https://climatehistory.com.au>) using the lessons we learned via SWD. It is important to note that inter-project knowledge sharing for meteorological data rescue has largely been by word of mouth and interpersonal relationships (having been helped by colleagues in the Weather Rescue and Old Weather projects that came prior to our project). There are relatively few references in the literature that describe exactly how data rescue that engages the general public is undertaken. Hence, we hope that this study provides a basic roadmap for novice practitioners that highlights insights about success and challenges for data rescue, and that the scientific community can build upon these lessons to accelerate the rapid acquisition of historical scientific data for wider societal benefits.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Andrew Lorrey (a.lorrey@niwa.co.nz).

Materials availability

Digital twins of the original log books and meteorological forms used in this study are held by NIWA. They can be made available on reasonable request.

Data and code availability

The data that were generated in this study are held by the [lead contact](#) and are currently being analyzed. They will be made available in the future through ACRE

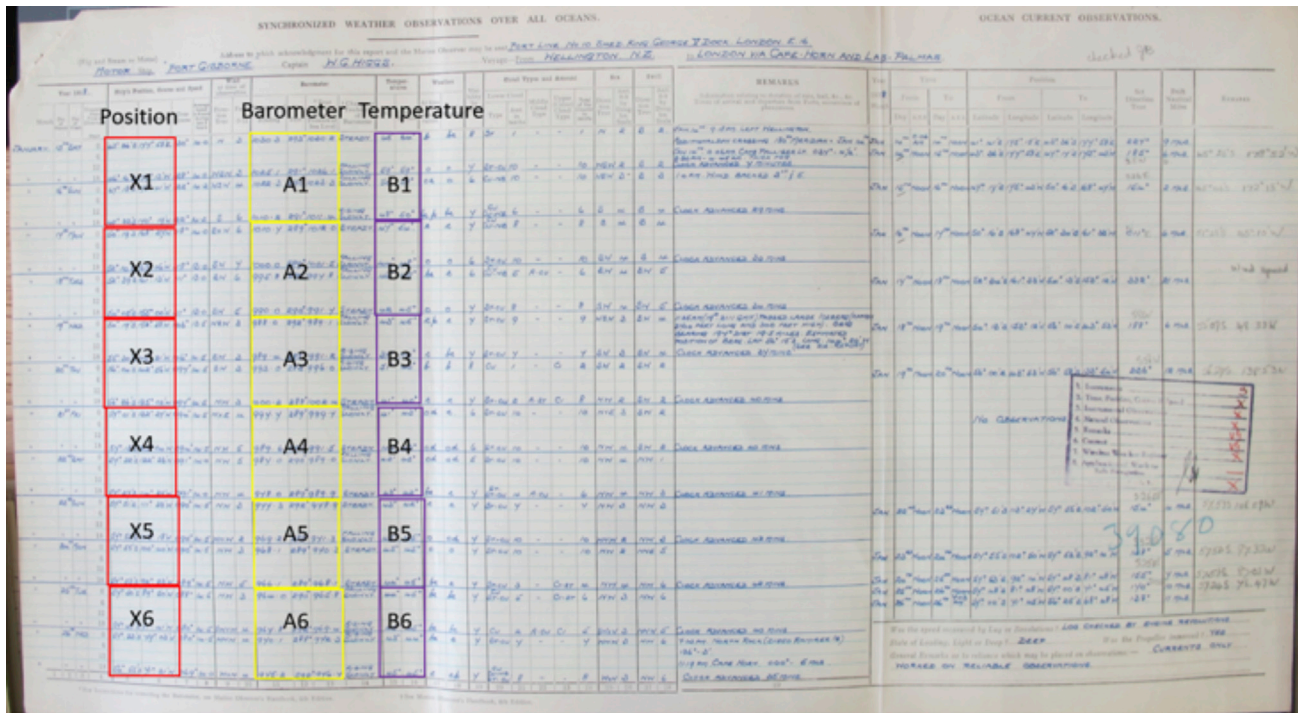


Figure 5. A log book page used in SWD

This example shows a standard weather observation register that was transcribed by citizen scientists in SWD. Clipping masks were placed over the digital version of the register, with alphanumeric labels placed on to ship position (X1–X6), barometric pressure (A1–A6), and temperature (B1–B6). The original file name contains a unique sample identifier, the name of the ship (in this case the MS Port Gisborne), and an image number, to which the clipping mask alphanumeric code was added before uploading to Zooniverse (e.g., MF9111_39,080_Port Gisborne_IMG_6247_B5.jpg for the clip of the register corresponding to box B5). This scheme facilitated ease of data retrieval and reparsing the data into a continuous time series for replicate quality assurance and further analysis. Image supplied by C. Wilkinson, RECLAIM.

or on reasonable request. Codes that were developed and/or augmented for this study are available in the links found in the [supplemental information](#).

Establishing a citizen science identity for our data rescue crew

When our work began, leading exemplars for historical meteorological data rescue harnessing citizen science were OldWeather^{56–58} and Weather Rescue.^{59,60} The latter project was built on the free-to-use Zooniverse Web platform (www.zooniverse.org) and demonstrated a capability to recover millions of observations keyed in replicate. Based on the global success of Weather Rescue, both in terms of public engagement and the great speed and volume of historical weather data transcribed, our research team decided to employ a similar design. We registered our project on Zooniverse, and simultaneously created a project identity. Our project description included a name and icon connected to the southern hemisphere region where we wanted to generate “discovery” science about weather and climate with historical meteorological observations and reanalyses. The heavy focus on rescuing maritime data in our project led us to use a ship as a project icon, including a sail with an Antarctica logo that was embellished with a thermometer and sun in the background. The name SWD arose out of testing word combinations we thought were reflective of the project work and regional focus. It is also a subtle play on words with respect to the well-known RRS Discovery Antarctic expeditions (from which we have obtained data). A website domain name was purchased in order to make a shortcut (via redirection) to SWD (www.southernweatherdiscovery.org) to make it easier for the general public to find us on Zooniverse (instead of directing them to find the project at the Zooniverse URL <https://www.zooniverse.org/projects/drewdeepsouth/southern-weather-discovery>).

Guiding citizen scientists through an ocean of data

The Zooniverse Web platform is designed to accommodate novice citizen science practitioners who have no prior knowledge of website design or Web

development. The build-a-project instructions (<https://help.zooniverse.org/getting-started/>) guide the completion of a project setup leading to two basic website components: a front end, which the general public can see and work with, and a back end that contains the design and content elements required to organize workflows and create data entry fields. There are several Web page hierarchical elements that can be viewed on the SWD front end, which include primary navigation tabs labeled About, Classify, Talk, and Collect. We discuss the first three of these tabs below.

Under the About tab, there are subsidiary tabs for Research, The Team (biographic information), Results, and Frequently Asked Questions (FAQ). We felt it was important to complete details for the Research and Team tabs under the About heading in order to establish our project identity upon launching SWD. We used the Team tab to outline biographic information; this element of the website humanizes the project by providing a face behind the science, as well as key points of contact. Additional considerations for providing personal details need to be weighed by each research team; we included the ability for citizen scientists to contact us to engender a better connection between our role as researchers and the public who we were trying to engage with for participating in data transcription. The Research tab provided an opportunity to outline more in-depth reasons for doing citizen science data transcription. Many of the citizen scientists using the Zooniverse platform are genuinely excited about the research, and providing these additional details helps them to engage with the project.

The Talk tab included conversations between our research team and citizen scientists. It was used to engage with participants who initiated questions or discussions, with most of the queries related to general data entry issues that were not pre-emptively thought of for the tutorial (mostly uncommon problems). In rare cases, submitted questions were related to reiteration of instructions when an occasional user did not understand our tutorial. More detailed questions about experimental design, including reasons for retirement limits

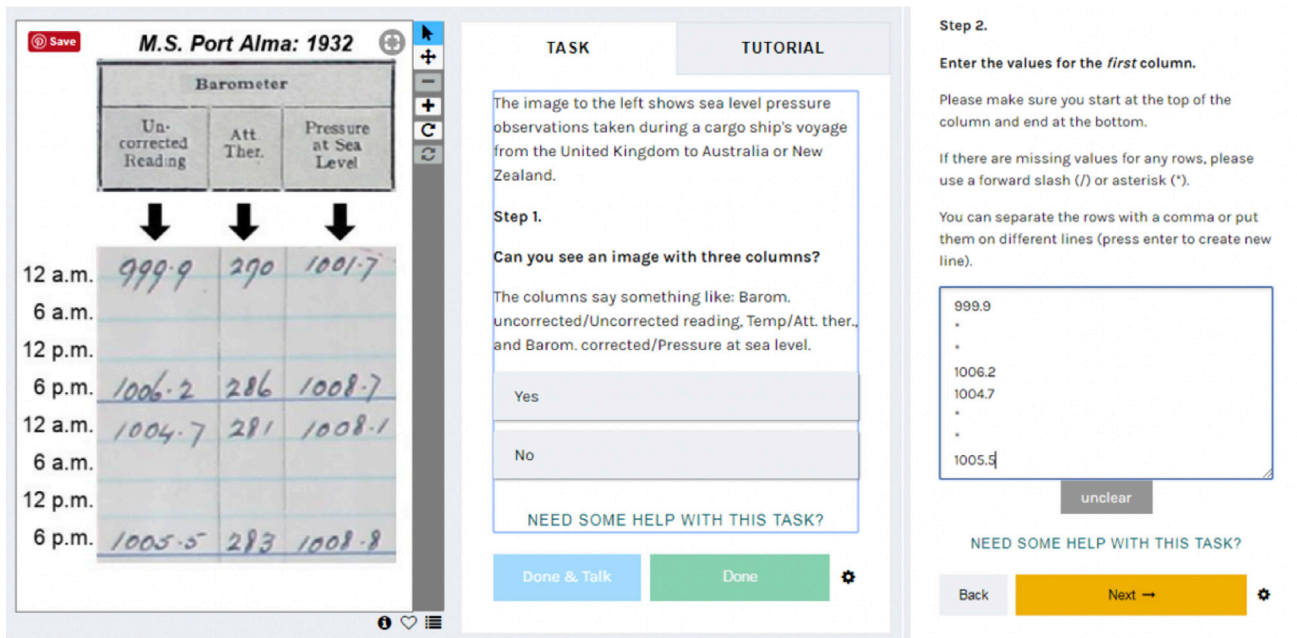


Figure 6. Example of ship log segment in SWD

(Left) Formatted clipped segment taken from the MS Port Alma in 1932 that shows 2 days of handwritten regimented observations in tabulated format for uncorrected atmospheric pressure, attached thermometer, and corrected pressure (reduced to sea level). (Center) The task description for keying these observations (step 1) serves as a check that the correct image clip was uploaded, while the example for the data entry field instructions (right) indicate to the citizen scientist which column to key and how to separate the values.

for each image, and how to deal with missing data were popular topics. We also used the Talk tab to occasionally provide new instructions for data keying, and in one case we specifically asked citizen scientists to change their transcription on the fly (to not use commas as a numeric separator due to a data formatting issue with Zooniverse). The Classify tab will be discussed below in more detail when we outline how workflows for data transcription were made.

Advanced preparations for a long data rescue voyage

Historical ship logbook observations and land-based meteorological registers were handwritten on standardized printed table forms (Figure 5), making them ideal for Zooniverse platform transcription. We undertook two main transcription tranches, each with a slightly different approach for uploading digital copies of meteorological registers for transcription. Ahead of volunteers keying data online, the architecture of a basic workplan needs to be considered to determine how the division of labor should proceed. This helps to maximize efficiency, minimize transcription errors, and reduce preparation time. Consideration about the data types that are keyed and preparation related to both SWD transcription tranches are provided below.

Most of the logbooks used in the first phase of SWD (Table 1) were sourced from UK merchant and immigration ships that visited New Zealand and Australia via the South Pacific, and were acquired from international archives through an extension of the Recovery of Logbooks and International Marine data (RECLAIM) project.^{39,40} Many of those logs used a standard printed register that arranged multiple observations in columns containing a unique variable (e.g., 9 a.m. temperature, pressure) and discrete entries in rows corresponding to a common date, time, and location (Figure 5). In the second phase of SWD (see section “shore leave for the WISE”), New Zealand land-based observations from meteorological registers containing a broader range of observations were drawn on, with nine discrete variables to key. For both phases of SWD, individual meteorological register pages were subdivided into small parts to provide a segment for an individual volunteer to transcribe rather than providing the whole page. This choice was based on discussions with colleagues and feedback from volunteers, and helped to (1) ensure data keying contributions could

be completed in short bursts rather than taking up lengthy intervals of time, (2) minimize the risk of volunteers abandoning a data entry form before submitting a full transcription, (3) reduce mistakes that are associated with transcribing the wrong column or row, and (4) decrease the probability of widespread error propagated across an entire log, if, for example, a specific volunteer had a difficult time deciphering the handwriting on a specific page. Our project's contractual requirements for the DSC also meant we prioritized certain observations on a logbook page, and therefore only a subset of meteorological logbook segments for each page and logbook were targeted for transcription.

A task fit for a clipper

To accommodate the structure of Zooniverse workflows that lead citizen science volunteers through keying (covered in detail below), we created subsets of each logbook page that were cropped and uploaded to SWD in a standard format. Adobe Illustrator software was used to crop segments of each logbook page using the artboard function. Logbook segments usually covered 2 days of a voyage, with four rows for observations per day (Figure 6). Clipping each segment out of the entire page was initially a semi-manual process, because many logbook images were not positioned identically each time a digital surrogate was created in archive (e.g., pages were inconsistently positioned when captured). This meant artboards used for clipping had to be iteratively adjusted to ensure the observations contained in the 2-day logbook segments were not truncated. Eventually, our team fixed this in pre-processing so the entire process of clipping could be automated. Once artboards were adjusted and aligned to the standardized logbook table dimensions, we adapted existing JavaScript to automate image labeling and cropping to produce logbook segments. A labeling convention was devised to identify where each segment clip was located on the original logbook page, with column A assigned to barometric pressure, column B assigned to temperature, and column X for ship position (Figure 5). The image names of each clip contained information about the archive folder, the ship name, the original image name, and the position of the clip on the logbook page attached as a file name suffix (see Figure 5).

Prior to uploading each clipped image to Zooniverse, a Jupyter notebook (a Web-based interactive computing platform) script run in Python was



Figure 7. Snow during the Week it Snowed Everywhere

Snowfall evidence for the Week it Snowed Everywhere (WISE) during late July 1939 at (left) Pukekohe, Auckland (credit: Huia Mitchell via Auckland Libraries Heritage collections, Footprints 03,956), and (right) in the streets of Dunedin, Otago (credit: Evening Star, reproduced by the Otago Daily Times).

used to add the name of the ship, the year of the voyage, the hours of observation, and column headings (see Figure 6). Our decision to use a Jupyter notebook for this step enabled research team members to generate logbook segment clips regardless of their scientific computing experience. In SWD phase II, we also initiated automation of logbook segment clipping using MATLAB to help streamline this stage of the data rescue process. This labeling system also makes reassembling data after transcription easier. Links to code for the aforementioned steps are provided in the [supplemental information](#).

Changing tack with specialized workflows

Three specialized workflows were created on Zooniverse for volunteers to take part in transcribing data for the SWD first tranche: ship position, temperature, and barometric pressure. The workflows were designed to be as simple as possible and utilized the logbook clips discussed above rather than displaying a whole logbook page. In the first tranche, we used an open entry field and asked volunteers to key a small column of data, with values separated by a range of delimiters (e.g., space, comma). In the SWD second phase, we upgraded the data entry forms to provide an individual entry box for each observation and adjusted the subdivisions of the logbook pages to ensure only one column of data was keyed in a step. This was assisted by Zooniverse via the Combo Task feature, which was experimental during early 2020, having been trialed through the Weather Rescue project. Although further customization of a Zooniverse-hosted citizen science website is possible, our team only used minimal special requirements like this that were facilitated by the Zooniverse staff.

The workflow questions were designed to lead the volunteers through the image with handwritten meteorological data: first, we asked if the image related to what the workflow task indicated (to potentially eliminate images that had been loaded into the wrong workflow). This was followed by a number of sequential questions that asked the volunteer to transcribe columns of numbers (Figure 6). A workflow task also asked the volunteers to transcribe the latitude and longitude so the historical weather observations could be ascribed to a location, date, and time. A separate workflow for temperature observations asked volunteers to transcribe air, sea, dry bulb, and wet bulb temperatures. Finally, a barometric pressure workflow asked volunteers to transcribe uncorrected pressure, the attached thermometer (required for correcting raw pressure measurements), and the corrected pressure at sea level.

Alongside each workflow, Zooniverse requires tutorials and a field guide to guide volunteers through each workflow step by step and address any idiosyncratic tasks for that workflow. As such, each separate workflow has a unique tutorial. The field guide addresses more general questions from the workflows and about the project in general, and it can be found on the side of any page of the Zooniverse project (see more details on southernweatherdiscovery.org).

Conscripting data rescue participants

Our maiden voyage into the deep south

Our team used a multiphase communication plan to introduce and promote SWD, including video and print media campaigns to garner participation and maintain interest in our meteorological data rescue project. An initial step was to create an introductory video for the SWD website that would encourage people to participate in digital keying of historical weather observations. The video content was crafted with the assumption that the audience had never heard of the project or previously participated in a citizen science effort. This portion of our strategy, in addition to parallel strategies for social media, print media, and radio, was designed by the SWD team and the NIWA Communications team across several months of work to ensure the data rescue scientific content was robust and that delivery to multiple media outlets would be ready in time for launching our SWD project on Zooniverse.

In pre-production for the SWD introductory video, we noted an obvious limitation related to visual content being restricted to historical ship logbooks. However, we were fortunate to find historical footage shot by Herbert Ponting of Robert Falcon Scott's British Antarctic Expedition to the South Pole in 1910.⁶¹ The black and white video footage of Scott's expedition showcases the conditions under which the scientific observations were made during the "heroic age of exploration." Weaving several segments from this historical video into our messaging was central to the strategy of initiating and maintaining engagement with SWD data rescue. To reflect the nautical elements of SWD data rescue, key segments were filmed at the Auckland Maritime Museum. We also framed the central issue around the difficulty of transcribing handwriting, and demonstrated how the audience could be a part of the solution. The general progression of the video also highlighted the importance of recovering historical meteorological observations to provide insights about our current and future climate (refer to the statement "Using their legacy to help ours" at the 2-minutes-and-13-seconds mark in our first SWD video; <https://vimeo.com/297007476>). The mixture of contemporary and historical video footage engenders ties to the golden age of exploration, with the idea to "breadcrumb" prospective citizen scientists toward participating.

For the SWD launch, observations that were taken at temporary encampments and during overland sledging missions during Scott's expedition were added onto the SWD website. These workflows were used to entice members of the public to take part in the transcribing effort, and for the media to create a story around. Some of these data came from printed tables and had already been transcribed by other researchers (without our prior knowledge). However, it was considered a minimal time sacrifice to copy and clip those images to garner significant public interest in the project.

SWD was launched on 30 October 2018. The introductory SWD video was promoted by NIWA and reached 127,200 people on Facebook, resulting in 288

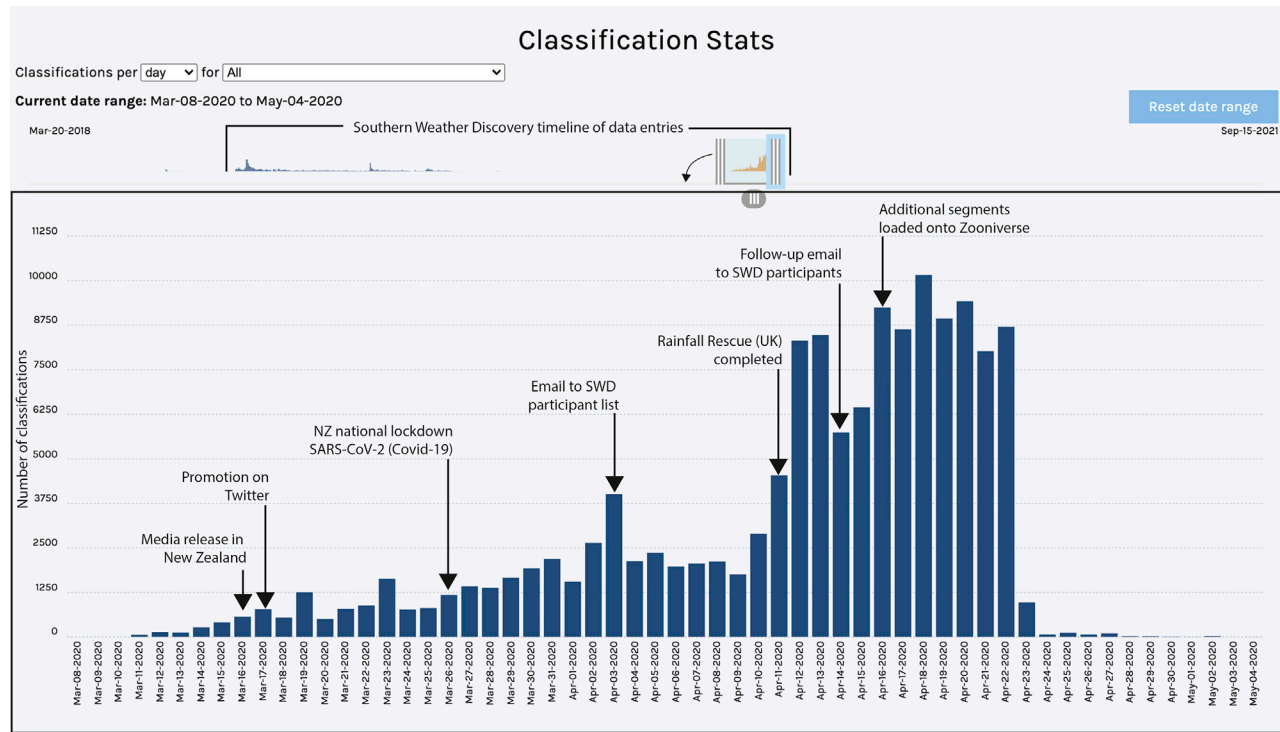


Figure 8. Zooniverse classification daily progress

Classification statistics for SWD, highlighting phase II WISE activity. Each classification is an instance where a citizen scientist has undertaken a keyed data transcription for a small section of a log book uploaded to Zooniverse. A noticeable boost in keying and project participation by citizen scientists coincided with media and social media advertising, emails to participants, and new material being uploaded to the site. The largest keying increase was associated with the completion of the UK Rainfall Rescue project, which bolstered international participation in our project.

comments, likes, or shares. It was viewed 55,000+ times on Facebook, YouTube, and Vimeo. The project was also promoted through the NIWA Communications team to New Zealand media, with a story featured on primetime television news (TV One), which has a nightly audience of ~600,000 national viewers (>10% of New Zealand's population) (<https://www.tvnz.co.nz/one-news/new-zealand/robert-scotts-weather-logs-give-kiwi-scientists-new-insight-climate-change>).

A story about SWD also featured on the front page of *The New Zealand Herald* website, New Zealand's second largest online news outlet with a monthly reach of 1.3 million subscribers (https://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=12151407). SWD also had significant coverage in provincial and regional newspapers, with an additional estimated 100,000+ audience reach. A parallel social media campaign was also launched through NIWA's social channels (Facebook and Twitter) and by members of the research team, with subsidiary re-promotion of materials to reach the Weather Rescue participants (who were largely based overseas and at the time were waiting for more data to key). NIWA's Twitter promotion about the project reached an audience of 7,955 people with 165 comments, likes, or retweets. Posts on Twitter about the project were shared by climate scientists, international and New Zealand science organizations, and hundreds of members of the public (and even by Chelsea Clinton to her ~2.4 million followers). The project was also promoted in an e-mail newsletter to all Zooniverse volunteers.

The metrics and progress components supplied from Zooniverse also allowed us to track the progress of data transcription, feedback from participants, and also opportunities to push social media to re-energize and draw in more people. The uptake of data keying by volunteers was swift, and over 50,000 observations, including all of the ice sledging data from Scott's expedition, were initially transcribed in replicate over the first 2 days after the launch of the SWD project. In total, 167,914 unique meteorological observations were successfully captured in replicate through phase I of the SWD project.

Shore leave for the WISE

The second phase of SWD focused on a project called The Week it Snowed Everywhere (WISE). This was a phrase that we coined to describe a significant snowfall event that affected most of New Zealand during the austral mid-winter of 1939 (Figure 7). A primary goal for this phase of SWD was to evaluate transcription retirement limits (replicate keying) and how those limits relate to optimal accuracy of citizen science transcription when dealing with different levels of replicate transcription. We also wanted to highlight the serendipitous benefit of SWD citizen science data rescue that comes from high levels of keying replication, including the ability to augment training libraries that underpin computer vision transcription of handwritten tabulated numbers. Automated transcription using computer vision techniques commonly relies on a standardized digital library called the Modified National Institute of Standards and Technology database (MNIST),⁶² which is used to train AI approaches.^{63,64} However, the MNIST dataset is relatively limited in terms of exemplary forms for handwritten digits compared with available contemporary resources and offerings in old texts.

A promotional video for WISE homed in on the technological connection between citizen science-driven data rescue and AI-based handwriting transcription (<https://vimeo.com/374313908>). A primary goal for this promotion was to communicate to the citizen scientists how their assistance could accelerate technology improvements and our scientific goals. In this case, having humans contribute to deep datasets that can train AI for handwriting transcription would result in more rapid realization of the benefits of weather reconstructions on a global scale.

We began the WISE video by establishing the value of the ship log observations for understanding past weather events (see quote at 20 seconds in the video, which states "We cannot go back. This is our time machine"). Then, we highlighted the problem that OCR has for transcribing tabulated handwritten digits. We coupled both concepts with the idea that combining scientific knowledge of meteorological data with citizen science and partnering with

a global leader in software provision (Microsoft) could help to rapidly overcome a significant problem.

Despite OCR technology being used for decades, there are limited video exemplars that demonstrate exactly how it works. To get around this shortfall for communicating to the target audience, our team made a suite of visual animations depicting what OCR software basically does (including a mock visualization of the MNIST training dataset). These connections helped to bring two project elements together: historical handwritten logbooks and OCR technology. The understated message is that the 1939 snowfall event provides data that can lead to improved OCR technology, which in turn can help to surmount present limits on rapid acquisition of historical scientific observations. The example from the snowfall event of 1939 also connects the importance of studying past extreme weather events with understanding global change from a relatively isolated location in the antipodes.

The WISE video launched at the November 2019 Microsoft Envision Forum NZ held in Auckland. In parallel, there was a promotional media campaign driven by NIWA Communications, with uptake of the story by all major print media, television, and radio outlets in New Zealand (Figure 8). The connection between Microsoft New Zealand and their parent organization also meant shared Twitter reach presenting the link to the promotional video exceeded 400,000 impressions in under 1 month. We also received significant help by “piggy-backing” off of Rainfall Rescue, a UK-based data rescue project running on Zooniverse, which supplied a volunteer corps to SWD directly after their project was completed. This influx of citizen scientists saw our project classifications increase by about 500%, and that level was maintained through completion, dramatically reducing the time for data capture (Figure 8).

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.patter.2022.100495>.

ACKNOWLEDGMENTS

This work was partly funded by the New Zealand Government's Deep South National Science Challenge Assessing and Validating the New Zealand Earth System Model Using Modern and Historical Observations and Strategic Science Investment Fund support from NIWA for Climate Present and Past contract CAO2101. The European Union Copernicus Climate Change Data Rescue Service (2017/C3S_311a_Lot1_Met Office) awarded to the United Kingdom Met Office also supported this work, and the World Meteorological Organization kindly provided co-funding for this study to be published. Support for the 20CR project version 3 dataset is provided by the US Department of Energy, Office of Science Biological and Environmental Research (BER); by the National Oceanic and Atmospheric Administration Climate Program Office; and by the NOAA Physical Sciences Laboratory. Support for the 20CR version 2c dataset is provided by the US Department of Energy, Office of Science BER and by the National Oceanic and Atmospheric Administration Climate Program Office. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the US Department of Energy under contract no. DE-AC02-05CH11231 using NERSC award BER-ERCAP0020982. Contributions by G.P.C. were partially supported by the NOAA Physical Sciences Laboratory, NOAA Climate Program Office, and the NOAA Cooperative Agreement with CIRES, NA17OAR4320101. We thank Lohit Batra, Bob Glancy, and Lucas Joppa for facilitating support to NIWA for SWD via a Microsoft AI for Earth grant in aid and for arranging a showcase of our efforts at the Microsoft Envision Forum in Auckland. We thank Peer Hechler from the WMO and Philip Brohan from the UKMO for comments that improved this manuscript. We thank all researchers from NIWA in New Zealand that have contributed to rescuing historical weather observations. We would also like to thank the team at Zooniverse and our SWD volunteers, without whom our project would not have been possible.

AUTHOR CONTRIBUTIONS

Supervising the work, A.M.L.; responsibility for all data, figures, and text, A.M.L., P.R.P., R.J.A., J.-M.W., E.J., L.S., S.M., S.R., P.Q., E.H., S.W., G.C.;

ensuring that authorship is granted appropriately to contributors, A.M.L.; ensuring that all authors approve the content and submission of the paper, as well as edits made through the revision and production processes, A.M.L.; ensuring adherence to all editorial and submission policies, A.M.L.; identifying and declaring competing interests on behalf of all authors, A.M.L.; identifying and disclosing related work by any co-authors under consideration elsewhere, A.M.L.; archiving unprocessed data and ensuring that figures accurately present the original data (see [Data and code availability](#) section), A.M.L., E.J., J.-M.W., C.W., and R.J.A.; arbitrating decisions and disputes and ensuring communication with the journal (before and after publication), sharing any relevant information or updates with co-authors, and being accountable for fulfilling requests for reagents and resources, A.M.L.; supply of data and analysis, A.M.L., E.J., J.-M.W., S.R., and L.S.; computational resources, S.R., P.Q., G.C., and L.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 9, 2020

Revised: March 8, 2022

Accepted: March 28, 2022

Published: May 27, 2022

REFERENCES

1. Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* 77, 437–471.
2. Kistler, R., Collins, W., Saha, S., White, G., Woollen, J., Kalnay, E., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., et al. (2001). The NCEP–NCAR 50-year reanalysis: monthly means CD–ROM and documentation. *Bull. Am. Meteorol. Soc.* 82, 247–268.
3. Slivinski, L.C., Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Giese, B.S., McColl, C., Allan, R., Yin, X., Vose, R., Titchner, H., et al. (2019). Towards a more reliable historical reanalysis: improvements for version 3 of the Twentieth Century Reanalysis system. *Q. J. R. Meteorol. Soc.* 145, 2876–2908. <https://doi.org/10.1002/qj.3598>.
4. Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., Bechtold, V.D.C., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., et al. (2005). The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.* 131, 2961–3012. <https://doi.org/10.1256/qj.04.176>.
5. Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Matsui, N., Allan, R.J., Yin, X., Gleason, B.E., Vose, R.S., Rutledge, G., Bessemoulin, P., et al. (2011). The Twentieth century reanalysis project. *Q. J. R. Meteorol. Soc.* 137, 1–28. <https://doi.org/10.1002/qj.776>.
6. Slivinski, L.C., Compo, G.P., Sardeshmukh, P.D., Whitaker, J.S., McColl, C., Allan, R.J., Brohan, P., Yin, X., Smith, C.A., Spencer, L.J., et al. (2021). An evaluation of the performance of the Twentieth century reanalysis version 3. *J. Clim.* 34, 1417–1438. <https://doi.org/10.1175/JCLI-D-20-0505.1>.
7. Gallant, A.J.E., Phipps, S.J., Karoly, D.J., Mullan, A.B., and Lorrey, A.M. (2013). Nonstationary Australasian teleconnections and implications for paleoclimate reconstructions. *J. Clim.* 26, 8827–8849. <https://doi.org/10.1175/JCLI-D-12-00338.1>.
8. Jiang, N., Griffiths, G., and Lorrey, A. (2013). Influence of large-scale climate modes on daily synoptic weather types over New Zealand. *Int. J. Climatol.* 33, 499–519. <https://doi.org/10.1002/joc.3443>.
9. Liu, Z., and Alexander, M. (2007). Atmospheric bridge, oceanic tunnel, and global climatic teleconnections. *Rev. Geophys.* 45, RG2005. <https://doi.org/10.1029/2005RG000172>.
10. Thorne, P.W., and Vose, R.S. (2010). Reanalyses suitable for characterizing long-term trends. *Bull. Am. Meteorol. Soc.* 91, 353–362. <https://doi.org/10.1175/2009BAMS2858.1>.
11. Freeman, E., Kent, E.C., Brohan, P., Cram, T., Gates, L., Huang, B., Liu, C., Smith, S.R., Worley, S.J., and Zhang, H.-M. (2019). The international

- comprehensive Ocean-Atmosphere data set – meeting users needs and future priorities. *Front. Mar. Sci.* 6. <https://doi.org/10.3389/fmars.2019.00435>.
12. Freeman, E., Woodruff, S.D., Worley, S.J., Lubker, S.J., Kent, E.C., Angel, W.E., Berry, D.I., Brohan, P., Eastman, R., Gates, L., et al. (2017). ICOADS Release 3.0: a major update to the historical marine climate record. *Int. J. Climatol.* 37, 2211–2232. <https://doi.org/10.1002/joc.4775>.
 13. Cram, T.A., Compo, G.P., Yin, X., Allan, R.J., McColl, C., Vose, R.S., Whitaker, J.S., Matsui, N., Ashcroft, L., Auchmann, R., et al. (2015). The international surface pressure Databank version 2. *Geosci. Data J.* 2, 31–46. <https://doi.org/10.1002/gdj3.25>.
 14. Compo, G.P., Sliwinski, L.C., Whitaker, J.S., Sardeshmukh, P.D., McColl, C., Brohan, P., Allan, R., Yin, X., Vose, R., Spencer, L.J., et al. (2019). The international surface pressure databank version 4. *Res. Data Arch. Natl. Cent. Atmos. Res. Comput. Inf. Syst. Lab.* <https://doi.org/10.5065/9EYR-TY90>.
 15. Woodruff, S.D., Worley, S.J., Lubker, S.J., Ji, Z., Eric Freeman, J., Berry, D.I., Brohan, P., Kent, E.C., Reynolds, R.W., Smith, S.R., et al. (2011). ICOADS Release 2.5: extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.* 31, 951–967. <https://doi.org/10.1002/joc.2103>.
 16. Worley, S.J., Woodruff, S.D., Reynolds, R.W., Lubker, S.J., and Lott, N. (2005). ICOADS release 2.1 data and products. *Int. J. Climatol.* 25, 823–842. <https://doi.org/10.1002/joc.1166>.
 17. Brönnimann, S., Brugnara, Y., Allan, R.J., Brunet, M., Compo, G.P., Crouthamel, R.I., Jones, P.D., Jourdain, S., Luterbacher, J., Siegmund, P., et al. (2018a). A roadmap to climate data rescue services. *Geosci. Data J.* 5, 28–39. <https://doi.org/10.1002/gdj3.56>.
 18. Thorne, P.W., Allan, R.J., Ashcroft, L., Brohan, P., Dunn, R.J.H., Menne, M.J., Pearce, P.R., Picas, J., Willett, K.M., Benoy, M., et al. (2017). Toward an integrated set of surface meteorological observations for climate science and applications. *Bull. Am. Meteorol. Soc.* 98, 2689–2702. <https://doi.org/10.1175/BAMS-D-16-0165.1>.
 19. Allan, R., Brohan, P., Compo, G.P., Stone, R., Luterbacher, J., and Brönnimann, S. (2011a). The international atmospheric circulation reconstructions over the Earth (ACRE) initiative. *Bull. Am. Meteorol. Soc.* 92, 1421–1425. <https://doi.org/10.1175/2011BAMS3218.1>.
 20. Allan, R., Wood, K., Freeman, E., Wilkinson, C., Andersson, A., Lorrey, A., Brohan, P., Stendel, M., and Kennedy, J. (2021). Learning from the past to understand the future: historical records of change in the ocean. *WMO Bull.* 70, 36–42.
 21. Brohan, P. (2019). Testing Google Vision for weather data rescue. <https://brohan.org/Google-Vision/>.
 22. Brohan, P. (2020). Testing AWS Textract for weather data rescue. <https://brohan.org/AWS-Texttract/>.
 23. Kaspar, F., Tinz, B., Mächel, H., and Gates, L. (2015). Data rescue of national and international meteorological observations at Deutscher Wetterdienst. *Adv. Sci. Res.* 12, 57–61. <https://doi.org/10.5194/asr-12-57-2015>.
 24. Ashcroft, L., Gergis, J., and Karoly, D.J. (2014). A historical climate dataset for southeastern Australia, 1788–1859. *Geosci. Data J.* 1, 158–178. <https://doi.org/10.1002/gdj3.19>.
 25. Bridgman, H., Ashcroft, L., Thornton, K., Di Gravio, G., and Oates, W. (2019). Meteorological observations for eversleigh station, near Armidale, New South Wales, Australia: 1877–1922. *Geosci. Data J.* 6, 174–188. <https://doi.org/10.1002/gdj3.80>.
 26. Slonosky, V., Sieber, R., Burr, G., Podolsky, L., Smith, R., Bartlett, M., Park, E., Cullen, J., and Fabry, F. (2019). From books to bytes: a new data rescue tool. *Geosci. Data J.* 6, 58–73. <https://doi.org/10.1002/gdj3.62>.
 27. Ashcroft, L., Coll, J.R., Gilabert, A., Domonkos, P., Brunet, M., Aguilar, E., Castella, M., Sigro, J., Harris, I., Unden, P., et al. (2018). A rescued dataset of sub-daily meteorological observations for Europe and the southern Mediterranean region, 1877–2012. *Earth Syst. Sci. Data* 10, 1613–1635. <https://doi.org/10.5194/essd-10-1613-2018>.
 28. Williams, M., Varma, B., Hayek, O., and Dean, M. (2016). Development of the New Zealand Earth system model. *Weather Clim.* 36, 25.
 29. Brenstrum, E. (1998). *The New Zealand Weather Book* (Craig Potton Publishing).
 30. Lorrey, A.M., and Chappell, P.R. (2016). The ‘dirty weather’ diaries of Reverend Richard Davis: insights about early colonial-era meteorology and climate variability for northern New Zealand, 1839–1851. *Clim. Past* 12. <https://doi.org/10.5194/cp-12-553-2016>.
 31. Mo, K.C., and Paegle, J.N. (2001). The Pacific-South American modes and their downstream effects. *Int. J. Climatol.* 21, 1211–1229. <https://doi.org/10.1002/joc.685>.
 32. Raphael, M.N. (2004). A zonal wave 3 index for the Southern Hemisphere. *Geophys. Res. Lett.* 31. <https://doi.org/10.1029/2004GL020365>.
 33. Ummenhofer, C.C., and England, M.H. (2007). Interannual extremes in New Zealand precipitation linked to modes of Southern Hemisphere climate variability. *J. Clim.* 20, 5418–5440. <https://doi.org/10.1175/2007JCLI1430.1>.
 34. Kidston, J., Renwick, J.A., and McGregor, J. (2009). Hemispheric-scale seasonality of the southern annular mode and impacts on the climate of New Zealand. *J. Clim.* 22, 4759–4770. <https://doi.org/10.1175/2009JCLI2640.1>.
 35. Brönnimann, S., Allan, R., Atkinson, C., Buizza, R., Bulygina, O., Dahlgren, P., Dee, D., Dunn, R., Gomes, P., John, V.O., et al. (2018b). Observations for reanalyses. *Bull. Am. Meteorol. Soc.* 99, 1851–1866. <https://doi.org/10.1175/BAMS-D-17-0229.1>.
 36. Wheeler, D., García Herrera, R., Koek, F., Wilkinson, C., Können, G., del Rosario Prieto, M., Jones, P., and Casale, R. (2007). *CLIWOC, Climatological Database for the World’s Oceans: 1750 to 1850; Results of a Research Project EVK1-CT-2000-00090* (European Commission).
 37. Teleti, P.R., Rees, W.G., Dowdeswell, J.A., and Wilkinson, C. (2019). A historical Southern Ocean climate dataset from whaling ships’ logbooks. *Geosci. Data J.* 6, 30–40. <https://doi.org/10.1002/gdj3.65>.
 38. Chappell, P.R., and Lorrey, A.M. (2014). Identifying New Zealand, Southeast Australia, and Southwest Pacific historical weather data sources using Ian Nicholson’s log of logs. *Geosci. Data J.* 1, 49–60. <https://doi.org/10.1002/gdj3.1>.
 39. Wilkinson, C., Woodruff, S.D., Brohan, P., Claesson, S., Freeman, E., Koek, F., Lubker, S.J., Marzin, C., and Wheeler, D. (2011). Recovery of logbooks and international marine data: the RECLAIM project. *Int. J. Climatol.* 31, 968–979. <https://doi.org/10.1002/joc.2102>.
 40. Wilkinson, C., and Vásquez, M. (2016). Report on the imaging of historic ice. Vestfold archive, Sandefjord, Norway: meteorological and oceanographic data in Antarctic waters (Historic sea-ice and meteorological data sources for Southern Ocean and Antarctic). <https://doi.org/10.13140/RG.2.2.21903.38569>.
 41. Wilkinson, C., and Vásquez, M. (2017). Report on the Imaging of Sources of Historic Ice, Meteorological and Oceanographic Data in the Southern Ocean – Åland Maritime Museum, Mariehamn, Finland (Historic sea-ice and meteorological data sources for Southern Ocean and Antarctic). <https://doi.org/10.13140/RG.2.2.24494.61763>.
 42. Kidson, E. (1947). *Daily Weather Charts Extending from Australia and New Zealand to the Antarctic Continent* (Government Printer).
 43. Wilkinson, C., and Freeman, E. (2021). *Best Practice Guidelines for Keying Data from Historic Marine Documents* (Copernicus Climate Change Data Rescue Service).
 44. Ding, H., Chen, K., Yuan, Y., Cai, M., Sun, L., Liang, S., and Huo, Q. (2017). A compact CNN-DBLSTM based character model for offline handwriting recognition with tucker decomposition. In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR) (IEEE), pp. 507–512.
 45. Zhong, Z., Sun, L., and Huo, Q. (2019). An anchor-free region proposal network for Faster R-CNN-based text detection approaches. *Int. J.*

- Doc. Anal. Recognit. 22, 315–327. <https://doi.org/10.1007/s10032-019-00335-y>.
46. Ma, C., Zhong, Z., Sun, L., and Huo, Q. (2019). A relation network based approach to curved text detection. In 2019 International Conference on Document Analysis and Recognition (ICDAR) (IEEE), pp. 707–713. <https://doi.org/10.1109/ICDAR.2019.00118>.
47. Brönnimann, S., Allan, R., Ashcroft, L., Baer, S., Barriendos, M., Brázdil, R., Brugnara, Y., Brunet, M., Brunetti, M., Chimani, B., et al. (2019). Unlocking pre-1850 instrumental meteorological records a global inventory. *Bull. Am. Meteorol. Soc.* 100, ES389–ES413. <https://doi.org/10.1175/BAMS-D-19-0040.1>.
48. Allan, R., Compo, G., and Carton, J. (2011b). Recovery of global surface weather observations for historical reanalyses and international users. *Eos, Trans. Am. Geophys. Union* 92, 154. <https://doi.org/10.1029/2011EO180008>.
49. Ashcroft, L., Allan, R., Bridgman, H., Gergis, J., Pudmenzky, C., and Thornton, K. (2016). Current climate data rescue activities in Australia. *Adv. Atmos. Sci.* 33, 1323–1324. <https://doi.org/10.1007/s00376-016-6189-5>.
50. Allan, R., Endfield, G., Damodaran, V., Adamson, G., Hannaford, M., Carroll, F., Macdonald, N., Groom, N., Jones, J., Williamson, F., et al. (2016). Toward integrated historical climate research: the example of Atmospheric Circulation Reconstructions over the Earth. *Wires Clim. Chang.* 7, 164–174. <https://doi.org/10.1002/wcc.379>.
51. Brunet, M., and Jones, P. (2011). Data rescue initiatives: bringing historical climate data into the 21st century. *Clim. Res.* 47, 29–40. <https://doi.org/10.3354/cr00960>.
52. Wilkinson, C., Brönnimann, S., Jourdain, S., Roucaute, E., Crouthamel, R., Brohan, P., Valente, A., Brugnara, Y., Brunet, M., Compo, G.P., et al. (2019). Best Practice Guidelines for Climate Data Rescue v1, of the Copernicus Climate Change Service Data Rescue Service (ECMWF). <https://doi.org/10.24381/x9rm-mp92>.
53. Brunet, M., Brugnara, Y., Noone, S., Stephens, A., Valente, M.A., Ventura, C., Jones, P., Gilbert, A., Brönnimann, S., Luterbacher, J., et al. (2020). Best practice guidelines for climate data and metadata formatting, quality control and submission of the Copernicus climate change service data rescue service. <https://doi.org/10.24381/kctk-8j22>.
54. Page, C.M., Nicholls, N., Plummer, N., Trewin, B., Manton, M., Alexander, L., Chambers, L.E., Choi, Y., Collins, D.A., Gosai, A., et al. (2004). Data rescue in the Southeast Asia and South Pacific region: challenges and opportunities. *Bull. Am. Meteorol. Soc.* 85, 1483–1490. <https://doi.org/10.1175/BAMS-85-10-1483>.
55. Craig, P.M., and Hawkins, E. (2020). Digitizing observations from the Met Office daily weather reports for 1900–1910 using citizen scientist volunteers. *Geosci. Data J.* 7, 116–134. <https://doi.org/10.1002/gdj3.93>.
56. Eveleigh, A., Jennett, C., Blandford, A., Brohan, P., and Cox, A.L. (2014). Designing for dabblers and deterring drop-outs in citizen science. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (ACM)*, pp. 2985–2994. <https://doi.org/10.1145/2556288.2557262>.
57. Brohan, P. (2014). *Citizen Science for Data Rescue: Recovering Historical Climate Records with a Network of 20,000 Volunteers (American Geophysical Union, Fall Meeting 2014)*.
58. Brohan, P. (2012). *oldWeather. Org: citizen science for climate reconstruction. In AGU Fall Meeting Abstract (American Geophysical Union), ED53A-0922*.
59. Hawkins, E., Burt, S., Brohan, P., Lockwood, M., Richardson, H., Roy, M., and Thomas, S. (2019). Hourly weather observations from the Scottish Highlands (1883–1904) rescued by volunteer citizen scientists. *Geosci. Data J.* 6, 160–173. <https://doi.org/10.1002/gdj3.79>.
60. Burt, S., and Hawkins, E. (2019). Near-zero humidities on Ben Nevis, Scotland, revealed by pioneering 19th-century observers and modern volunteers. *Int. J. Climatol.* 39, 4451–4466. <https://doi.org/10.1002/joc.6084>.
61. Read J. Scott's Last Journey. 1964. United Kingdom: British Broadcasting Corporation; <https://archive.org/details/scottslastjourney/scottslastjourneyreel1.mov>.
62. LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324.
63. Ahlawat, S., and Choudhary, A. (2020). Hybrid CNN-SVM classifier for handwritten digit recognition. *Proced. Comput. Sci.* 167, 2554–2560. <https://doi.org/10.1016/j.procs.2020.03.309>.
64. Kussul, E., and Baidyk, T. (2004). Improved method of handwritten digit recognition tested on MNIST database. *Image Vis. Comput.* 22, 971–981. <https://doi.org/10.1016/j.imavis.2004.03.008>.

Patterns, Volume 3

Supplemental information

Meteorological data rescue: Citizen science

lessons learned from Southern Weather Discovery

Andrew M. Lorrey, Petra R. Pearce, Rob Allan, Clive Wilkinson, John-Mark Woolley, Emily Judd, Stuart Mackay, Sudhir Rawhat, Laura Slivinski, Sally Wilkinson, Ed Hawkins, Patrick Quesnel, and Gilbert P. Compo

Meteorological data rescue – citizen science lessons learned from Southern Weather
Discovery : Supplemental experiment procedures

S1. Code for automated clipping of log book images

JavaScript code that was adapted and used to automate image labelling and cropping to produce logbook segments can be found at the following address:

http://kelsocartography.com/scripts/scripts/nvkelso/PlaceMultipleFiles_v1_images.js

Additional code used in the WISE experiment to produce clips of F301 meteorological logs can be found at the following address:

https://www.dropbox.com/s/ocktsq85ifnfuli/NZ_MF301_clipping.zip?dl=0

S2. Common failure examples from Southern Weather Discovery.

Image segment
(G22581_1939_01_s3)

11	41	29 05
12	31	29 19
13	36	29 63
14	32	29 44
15	41	29 48

Figure S1. Example of confusion between 4, 6, and 7. In this example, the penmanship is clear, but there are similarities between how these numbers looked to a range of end users.

Image segment
(A53291_1939_05_s1)

	° F.	Inches.
1	47.4	29.668
2	51.3	29.924
3	52.4	29.918
4	52.3	29.946
5	52.2	29.892

Figure S2. A wider view of the entire table for the example seen in Figure S2.1, which citizen scientists did not see in SWD, suggest two pen strokes were used to write the number 4, while one stroke was used for the number 6 and 7.

Image segment
(D05960_1939_03_s3)

11	390	33016
12	402	303018
13	442	2981
14	455	2990
15	442	2994

Figure S3. Missing decimals with a string of numbers that had either no space or a space delimiter was one of the common examples where consensus was not reached. The first column is attached thermometer and the second column is uncorrected barometer. Values were transcribed as written, so decimal points were omitted. Values in column 1 reached consensus in all but one case (day 14), but converged on the incorrect solution. Values in column 2 did not reach consensus.

Image segment
(I58061_1939_01_s6)

26	44	28	682
27	47	28	460
28	47	28	560
29	48	28	702
30	48	28	614
31			702

Figure S4. Pressure values that have a space delimiter instead of a decimal were transcribed as 28, 28 682, and 28.682. To a degree, a post-hoc correction can be made in this situation, but not the former because temperatures are sometimes in Fahrenheit, Celsius, or Kelvin.

Image segment
(I58061_1939_01_s6)

26	44	28	682	29	742
27	47	28	460	29	530
28	47	28	560	29	620
29	48	28	702 6x4	29	762 6x4
30	48	28	614 702	29	674 762
31					

Figure S5. Corrected entries often posed a problem that limited consensus being reached. Despite asking citizen scientists to enter original values, instead of corrected values, in some cases the original numbers cannot be seen. The default position is the citizen scientist enters the corrected value. In this situation, entries consisted of 29.674, 29.162, and 29.762 (the correct original value).

S3. WISE metrics for different experiments

	Completed entries	% completed entries	Segments completed	% completed segments	Completed images	% completed images
O75	45231	97.90	2850	84.07	32	16.93
O60	45892	99.33	3192	94.16	93	49.21
ORS	46200	100.00	3390	100.00	189	100.00
O5	46070	99.72	3294	97.17	124	65.61
O10	45936	99.43	3206	94.57	102	53.97

Table S1. The number and percent of completed entries, logbook segments and logbook images for the WISE experiments. Note, the ORS experiment results in a 100% complete dataset, and small variations in the percentage completed entries has a larger knock on effect for the number of completed segments and images.

	O75	O60	ORS	O5	O10
Attached thermometer	64 / 1.1%	15 / 0.3%	0 / 0%	11 / 0.2%	13 / 0.2%
Barometer uncorrected	66 / 1.1%	21 / 0.4%	0 / 0%	17 / 0.3%	16 / 0.3%
Barometer corrected	51 / 0.9%	9 / 0.2%	0 / 0%	9 / 0.2%	10 / 0.2%
Max Temp	204 / 3.5%	79 / 1.4%	0 / 0%	36 / 0.6%	76 / 1.3%
Min Temp	164 / 2.8%	68 / 1.2%	0 / 0%	28 / 0.5%	75 / 1.3%
Wind direction	173 / 3.0%	55 / 1.0%	0 / 0%	21 / 0.4%	37 / 0.6%
Wind force	223 / 3.9%	55 / 1.0%	0 / 0%	8 / 0.1%	32 / 0.6%
Wind run	24 / 0.4%	6 / 0.1%	0 / 0%	0 / 0%	5 / 0.1%
Pressure (all)	181 / 1.0%	45 / 0.3%	0 / 0%	37 / 0.2%	39 / 0.2%
Temperature (all)	368 / 3.2%	147 / 1.3%	0 / 0%	64 / 0.6%	151 / 1.3%
Wind (all)	420 / 2.4%	116 / 0.7%	0 / 0%	29 / 0.2%	74 / 0.4%

Table S2. The number and percent of entries per task (top 8 rows) and workflow (bottom three rows) that did not reach a consensus value for the WISE experiments. Note the overall higher failure rate for temperature tasks compared to pressure and wind tasks.