CAREERS

MONROE COUNTY PUBLIC LIBRARY

Archived photographs such as this image of trophy fish, taken in Florida in the 1950s, can help scientists to investigate long-term ecosystem changes.

HISTORICAL DATA

Hidden in the past

Old photos, logbooks and papers are a gold mine for fields such as ecology and climatology.

BY ROBERTA KWOK

In 2012, Ruth Thurstan turned to an unconventional source to study fishing: old newspapers. She wanted to know when people had started catching substantial numbers of snapper (*Pagrus auratus*), a fish species abundant off Australia's coast, and how much effort was needed at the time to catch them. But available detailed data stretched back only to the late 1980s. Thurstan, a marine historical ecologist now at Deakin University in Warrnambool, Australia, noticed that today's fishers of snapper often recount their experiences in magazine articles and blog posts. She wondered where fishers from the past would have published such descriptions.

A search through Australian newspaper archives revealed about 270 articles containing quantitative information about snapper-fishing trips from 1871 to 1939. "The amount of detail was astounding," Thurstan says. One account[1] she found read: "... down went about twenty-four lines; in two minutes the cry rose 'schnapper'; in three minutes more, at least a dozen splendid fish were flapping on the deck … For four hours and a quarter the sport was sustained with undiminished ardor and success … When our respective plunder was totalled up we were found to have lessened that particular tribe of schnapper by about 735 individuals."

By combining numbers gleaned from the articles with information from early-twentieth-century government fishery surveys and from a fisherman's book, Thurstan's team estimated that the average number of snapper caught per person per hour was about nine times higher in the late 1800s and early 1900s than it was by the end of the twentieth century[1]. She is working with fisheries modellers to incorporate the data into snapper-stock assessments and to refine estimates of sustainable fishing levels.

Thurstan's study is one of many examples ▶

of the value of historical materials for investigating scientific questions about long-term environmental changes. Much exploration of historical archives is concentrated in ecology and climate science, owing partly to rising concerns about the effects of climate change, although ongoing digitization efforts are also making it easier to conduct historical studies in fields such as astronomy and epidemiology.

Researchers who delve into archived materials often find surprises. "There are so many stories that are locked away in historical data," says Andrew Trant, an ecologist at the University of Waterloo in Canada. These accounts can also add a human dimension to the data. "If you have a narrative saying: 'We had such a difficult time and some people died of hunger,'" says Astrid Ogilvie, a climate historian at the Stefansson Arctic Institute in Akureyri, Iceland, "it puts flesh on the bones."

CAUTIOUS APPROACH

But old materials don't give up their secrets easily, say those who have taken this route (see 'Time-travel tips'). Useful documents may be buried in obscure archives — or may be so extensive that researchers need citizen scientists to help process them. Biased and incomplete historical data require careful interpretation, and researchers sometimes struggle to find funding for this kind of work. Scientists experienced in the search suggest partnering with historians, amassing many sources of such data and crafting grant proposals that emphasize the broad impact of their research.

Even when researchers conduct studies cautiously, they may encounter scepticism from colleagues. Scientists tend to value historical material less than scientific data, says Miguel Clavero, an ecologist at the Spanish National Research Council's Doñana Biological Station in Seville. But if the historical evidence is strong, he says, "it is as valuable as any other approach". Loren McClenachan, a historical ecologist at Colby College in Waterville, Maine, recalls that when she started her work in the early 2000s, it wasn't clear whether old materials would yield scientifically useful results. But now that a track record of these kinds of study exists, she says, it is less risky to enter the discipline. In February, researchers declared in *PLoS ONE* that historical ecology "has come of age", and listed 50 questions that the field should investigate, such as how humans have influenced extinctions[2].

Some ecologists worry that scientists have been measuring ecosystem changes over the course of their careers instead of across history. Because researchers then compare changes to already lowered baselines, they may underestimate the real magnitude of decline over generations. This concept, known as 'shifting-baseline syndrome', was defined in 1995 (ref. 3); since then, the number of citations of papers containing 'marine' and 'shifting baselines' has risen to about 200 per year, signifying the development of the marine historical-ecology field, according to a 2016 review[4].

> ***"There are so many stories that are locked away in historical data."***

Attitudes towards archived data have changed in other fields, too. Astronomers used to look down on colleagues who analysed old data, says Alyssa Goodman, an astronomer at the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts. They were seen as "some sort of second-rate scientist who couldn't get telescope time", says Goodman, who is also the founder of Astronomy Rewind, a citizen-science project to gather information from astronomy-journal images dating to the mid-1800s. "That is what has completely changed," she says — astroinformatics, a field that involves analysing previously gathered data, has emerged in the astronomy community.

Digitization has made historical materials such as newspaper archives more accessible. The programme 'Digital Access to a Sky Century @ Harvard' is scanning roughly 450,000 of the Harvard College Observatory's astronomy images recorded on glass photographic plates from 1885 to 1992. Elsewhere, as part of a project called the Venice Time Machine, researchers are digitizing Venetian documents from the tenth to the twentieth centuries in Italy, which may help scientists to study the spread of the plague (see *Nature* **546,** 341–344; 2017).

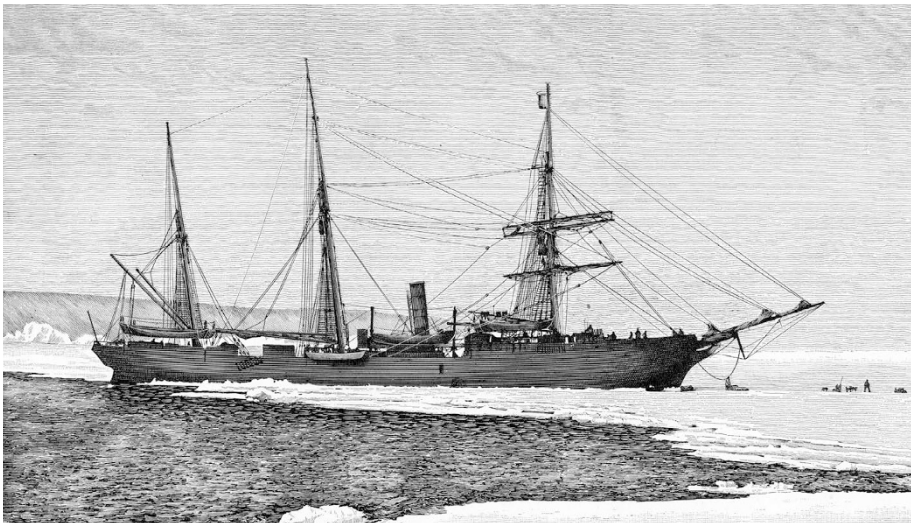
Public libraries, government archives and historical societies are potential sources of material. But these searches may not yield immediate results, so researchers should be patient and persistent, McClenachan says. When she asked a Florida public library for anything related to long-term changes in local coral reefs and fishing history, the archivist gave her articles about fisheries regulations, but they didn't contain the information that she needed. Finally, after three weeks, he brought out photos of trophy fish dating to the 1950s. She analysed the images to determine whether and how fish sizes had changed in five decades[5].

Other researchers have procured useful materials through a historian. In 2013, Clavero was studying whether white-clawed crayfish (*Austropotamobius italicus*) were introduced or native in Spain. He learnt from another ecologist that sixteenth-century Spanish-court letters suggested that the animals had been imported. Clavero contacted Susanne Kubersky-Piredda, an art historian at the Max Planck Institute for Art History in Rome who had studied the letters, and she provided copies. The documents showed that Spain's King Philip II had asked a member of his court to arrange for crayfish to be brought from Italy, and that the animals were shipped in 1588 (ref. 6) .

Sometimes, researchers come across a trove so enormous that they need the public's help to process it. Kevin Wood, a historical climatologist at the University of Washington in Seattle, found out that the US National Archives and Records Administration branches in College Park, Maryland, and Washington DC had tens of thousands of ship logbooks containing weather observations from 1801 to 1983. Local graduate students in information studies helped his team to scan more than 1 million pages. The researchers added the photos to the existing online citizen-science project Old Weather (www.oldweather.org). So far, volunteers have transcribed about 1.3 million US weather records from the logs, and Wood will use the data to improve climate models.

Historical materials can be tricky to interpret. Scientists can't estimate species abundance by looking at the number of records, because rare and valuable animals and plants were noted more frequently, McClenachan says. The documents' authors may have distorted the truth; for instance, hunters might have under-reported harvests to avoid taxes.

And sometimes, descriptions are unclear. Marcel Salathé, a digital epidemiologist who is studying health records in the Venice Time



NOAA

Logs from ships such as the *Bear* (pictured) contain weather data that citizen scientists are helping to extract.

ALEX NEWMAN/PRI

Inspection of old images on historical glass plates can often uncover valuable data.

Machine project, can guess whether a patient died of the plague on the basis of their symptoms and death-rate patterns, but can't be certain. "You're looking at very old data with very little context," says Salathé, who is at the Swiss Federal Institute of Technology in Lausanne. "You have to be extremely cautious."

Checking multiple historical sources and incorporating data from other fields can increase confidence in a finding, experts say. In her snapper study, Thurstan found that catch rates derived from newspaper articles and from survey data gathered by an Australian government vessel were fairly consistent. However, catch rates from a 1905 fisherman's book were higher[1], suggesting that the author might have described only his best trips. Similarly, Clavero's team backed up its finding[6] from the Spanish-court letters with evidence from fields such as archaeology, linguistics, genetics and ecology.

BRING IN THE EXPERTS

A historian's input is crucial for evaluating a source's accuracy, says Ogilvie. She points to *Setbergsannáll*, an eighteenth-century Icelandic manuscript containing weather information. The document's descriptions span 1201–1713, but much of the information it contains is thought to be inaccurate, because it is second-hand or possibly even invented, Ogilvie says. A historian can help to distinguish reliable from unreliable sources — for instance, by comparing the material to accounts written closer in time to the event.

Historians can also provide context, such as whether numbers are typical of that time period. A scientist reading old fishing reports for the first time might dismiss high catch rates as unrealistic, but a historian might say that they were normal, Thurstan says. When materials are written in another language or in medieval handwriting, a historian may be able to translate and clarify the text.

Even qualitative information can prove useful. Thurstan found that some newspaper articles about snapper didn't include data such as time spent fishing. However, she was able to estimate possible ranges by extrapolating numbers from other expeditions of the same boat, because she knew roughly how long the vessel typically spent on each trip at the fishing grounds and how many fishers it carried, and she used statistical methods to determine levels of uncertainty.

Old images present other challenges. In 2015, Trant's team started analysing nineteenth-century and early-twentieth-century photos of mountains in Canada to study forest changes. Because the photographer had not pointed the camera straight down towards the treetops from above, the researchers couldn't count pixels to measure tree cover. Instead, the team compared treelines to those in modern photos taken from the same position, and then located those points on Google Earth to estimate how far forest edges had moved.

Historical projects can be a tough sell to science-funding agencies if the research does not seem cutting-edge. Given the materials' age, applicants must convince reviewers that no one has done the analysis before. Winning grants from humanities funding agencies can be tricky, too, in part because they tend to have less money. Because historians generally value outputs such as monographs, they may find it difficult to evaluate scientists' paper publications, which reduces the chances of receiving funding, Thurstan warns.

Some researchers have found support from non-governmental organizations and philanthropic organizations that focus on conservation. Other scientists have had luck with programmes related to archived data, such as NASA's Astrophysics Data Analysis Program and data-rescue efforts at the North Pacific Research Board in Anchorage, Alaska.

Expanding the project's remit can also help. Wood's team is collaborating with Joseph Reidy, a historian at Howard University in Washington DC, on a potential project to analyse records from US Civil War Navy ships; Wood hopes to extract weather data from logbooks, whereas Reidy is interested in the role of African-American sailors. "It broadens the value of what we're doing beyond our discipline," says Wood, whose team is seeking funding from the non-profit Council on Library and Information Resources in Washington DC. Emphasizing outreach may improve applicants' chances, because the public likes historical data, Trant says.

Historical research can help scientists to establish a niche. Ashley Pagnotta, an astrophysicist at the College of Charleston in South Carolina, found a nova eruption captured on a 1900 glass plate in the Harvard College Observatory's collection[7]. Because few astronomers study glass plates, other scientists have asked Pagnotta to collaborate on similar analyses.

The work also gave her new perspective. Female astronomers analysed the Harvard images in the 1800s and early 1900s, and seeing their contributions — including their handwriting on the plates — made Pagnotta feel connected to these pioneers. "You feel like you're the next in a line of women astronomers," she says.

Scientists are accustomed to precise data, not to messy manuscripts or blurry images. But historical researchers say that if these sources are ignored, valuable details may go undiscovered. "Why," Ogilvie asks, "wouldn't we use every scrap of information we can get?" ■

DIG DEEP

Time-travel tips

- Partner with a historian, who can translate text, assess sources' reliability and provide context.
- If initial requests for materials don't yield results, continue working with archivists to refine the search.
- Be aware of each source's biases, and compare data from multiple sources.
- Focus on a specific goal, such as looking for a phenomenon or testing an algorithm to detect patterns.
- To improve funding chances, use new technology to make materials more accessible or to analyse text. **R.K.**

Roberta Kwok *is a freelance science writer in Kirkland, Washington.*

1. Thurstan, R. H. *et al. Fish Fish.* **17,** 210–225 (2016).
2. Armstrong, C. G. *et al. PLoS ONE* **12,** e0171883 (2017).
3. Pauly, D. *Trends Ecol. Evol.* **10,** 430 (1995).
4. Engelhard, G. H. *et al. ICES J. Mar. Sci.* **73,** 1386–1403 (2016).
5. McClenachan, L. *Conserv. Biol.* **23,** 636–643 (2009).
6. Clavero, M., Nores, C., Kubersky-Piredda, S. & Centeno-Cuadros, A. *Biol. Rev.* **91,** 1036–1049 (2016).
7. Pagnotta, A., Schaefer, B. E., Xiao, L., Collazzi, A. C. & Kroll, P. *Astron. J.* **138,** 1230–1234 (2009).

CORRECTION

The Careers feature 'Hidden in the past' (*Nature* **549,** 419–421; 2017) gave the wrong details for reference 7. It should have read Pagnotta, A., Schaefer, B. E., Xiao, L., Collazzi, A. C. & Kroll, P. *Astron. J.* **138,** 1230–1234 (2009).