Development and implementation of a climate data management system for western Pacific small island developing states

David J. Martin,* Andrew Howard, Rod Hutchinson, Simon McGree and David A. Jones

Bureau of Meteorology, Climate Information Services, Docklands, Victoria, Australia

ABSTRACT: This study describes the development and deployment of a new climate data management system aimed at improving climate data management and associated climate services in Pacific island countries and East Timor. The system is called Climate Data for the Environment (CliDE). Installed locally, it provides each country with a central relational database and web-based user interface that includes customizable key entry forms, quality control tools, station maintenance forms, meteorological and climate reports, and data file extracts. It has been deployed as free and open source software in 15 countries (East Timor, Papua New Guinea, Solomon Islands, Vanuatu, Palau, Federated States of Micronesia, Nauru, Marshall Islands, Kiribati, Tuvalu, Fiji, Tonga, Samoa, Niue and Cook Islands).

In developing CliDE, the project team sought to develop a database capable of providing a robust method of managing an individual country’s meteorological observations from networks typically consisting of several to hundreds of stations. In many Pacific countries significant data remain only in paper form; therefore, a key consideration was in providing a secure and efficient means for digitizing paper records. As CliDE has developed, it has become the central hub of a multitude of climate and meteorological services of benefit to small national meteorological services, such as statistical reports, graphical analyses, data extractions, climate summaries, and products that can provide input to public works planning, agriculture and health sectors. It has helped improve significantly the work flow, data integrity and consistency beyond previous practices in the western Pacific.

KEY WORDS climate data management system; Pacific region; data rescue; data management; climate record

Received 30 October 2013; Revised 16 January 2014; Accepted 26 February 2014

1. Introduction

High-quality, accessible climate data have widespread application, such as monitoring climate variability and change, supporting decisions around the risk management of natural hazards and disaster risk reduction, and supporting future climate predictions and projections (Australian Bureau of Meteorology and CSIRO, 2011). In the context of climate change, historical climate data define the range of climate variability experienced as well as provide context for the interpretation of projected climate changes for the future (e.g. Jones et al., 2013).

Managing climate data presents many challenges and these do not scale necessarily with the size of a meteorological service or the number of stations in networks. National meteorological services (NMSs) in the Pacific region (the focus of this work) have particular challenges owing to their small operational budgets, such as difficulty in maintaining suitable expertise, the large quantity of data that still remain only on paper records, and the requirement that data management solutions be low cost and sustainable (e.g. Page et al., 2004). Furthermore, the use of climate data is becoming more sophisticated, meaning that more data are required more frequently and rapidly (e.g. for the calculation of sub-daily Intensity-Frequency-Duration design rainfalls) and often combined with other environmental data to inform decisions (e.g. dengue incidence, Hales et al., 1999). This creates added complications in that near-real time high frequency observations are only feasible using automatic electronic instrumentation, which comes at a higher cost both financially and technologically compared to the traditional manual methods.

Recognizing the vulnerability of island states in the western Pacific to climate change, the Australian government established a programme of climate research called the Pacific Climate Change Science Program/Pacific-Australia Climate Change Science and Adaptation Planning Program (PCCSP/PACCSAP) to meet the high-priority science needs of supporting climate change adaptation in countries in the region (http://www.pacificclimatechangescience.org/). The countries participating in this science programme included the Cook Islands, East Timor, Federated States of Micronesia, Fiji, Kiribati, Marshall Islands, Nauru, Niue, Palau, Papua New Guinea, Samoa, Solomon Islands, Tonga, Tuvalu and Vanuatu (‘partner countries’ Figure 1). The programme sought to improve the understanding of past and future climate change in the western Pacific, as well as build local capacity through the provision of education, training, and awareness of climate change science more generally (Power et al., 2011; Australian Bureau of Meteorology and CSIRO, 2011).

This study describes the development and deployment of a new Climate Data Management System (CDMS) called Climate Data for the Environment (CliDE) in support of East Timor and Pacific NMSs. The goal of the CliDE project was to improve data availability for climate monitoring leading to an improved management of the risks posed by climate variability and change. CliDE was developed with the core principles of portability, free and open source software, and minimizing complexity whilst still maintaining all the required functionality for routine climate services. These principles were followed
to ensure the viability and sustainability of the system beyond the life of the PCCSP/PACCSAP.

CliDE stores primarily meteorological (or weather) time-series observations. However, the term climate data is used in this study to include a broader set of data such as corrected or adjusted values, monthly aggregates, station metadata, quality flags, and lineage information, all of which can be stored in CliDE.

2. Recent climate data management practices in the Pacific

2.1. Historical background

Historically, the partner countries of the programme, with the exception of Tonga have been administered by other nations: Britain, France, Spain, Portugal, Germany, Indonesia, Japan, United States, Australia, and New Zealand (Table 1). National independence or self-governance has only recently been gained, with a majority of countries becoming independent between the 1960s and 1980s. For some partner countries, defence and foreign policy remain the responsibility of the former administrator, which to some extent influences present day meteorological operations.

Partner country meteorological observations date back more than a century with the first known observations beginning in 1851 at the Malua Theological College, Samoa, which was the first theological college to be established in the South Pacific in 1848 by the London Missionary Society (Turner, 1861). These observations were typically set up by enthusiastic settlers (e.g. Smythe, 1864; Holmes 1877; 1881) and industries with the purpose of exploiting the agricultural or mineral potential of the colony. Colonial governments also introduced meteorological observation practices and procedures but this was largely limited to the chief administrative centres. The establishment of the southwest Pacific as the fifth region of the International Meteorological Organization in 1937 and the Second World War prompted further expansion of observation networks from the early 1940s onwards. In the following decades, observation and forecasting needs for military and civil aviation, weather forecasting, tropical cyclone warnings, and further economic activity (e.g. agriculture, forestry, mining, and tourism) provided the main impetus for network development (Krishna, 2009).

2.2. Climate data practice

For the most part, the partner country NMSs are well placed to operate their observation networks. However, many do not have the finance to provide all services to World Meteorological Organisation (WMO) standards (e.g. WMO, 2004; 2011). For many countries, external contributions in the form of grants and foreign aid are required to ensure the NMSs are sustainable. NMSs were observed to be competing with sectors such as health, education, and infrastructure for the limited government funding that is available and resource pressures are clearly apparent in the form of station closures, persistent instrument breakdown and less than optimal climate services including the storage of paper records.

Climate data are the basis for many climate services, including urban, flood and building design, seasonal prediction (forecast verification and model calibration) and assessment of weather risk. Decision making during events such as droughts depends on climate data. Climate data and simple climate summaries, such as monthly and yearly averages and totals, have been a hallmark of climate services for decades. In recent years, there has been an increase in demand for climate services particularly in weather sensitive industries such as agriculture, fishing, health and infrastructure. CliDE was developed to help meet this burgeoning demand for improved data and services.

‘The nature of climatology has changed from a discipline that emphasized the collection and archiving of data to one that retained this dimension, but added new dimensions of service and fast response’ (Slatyer and Bonner, 1996).

Best practice climate data management should be an enabler of both basic and more complex derived products and services. Meteorological data are at their most useful if they fulfil the following two requirements. First, the meteorological element should be observed at the same time and location over a long period of time, at least 30 years, to eliminate year-to-year variations and provide statistically significant patterns and trends. In addition, having a number of appropriately spaced stations recording the same observations at the same time provides suitable data to enable scientists to garner an understanding of the weather systems affecting an area of interest. Both spatial and temporal distribution of meteorological observations are
### Table 1. Summary timeline of pre-independence administration (adapted from Macellan, 2000).

| Country                  | Year of independence | First European colonization | Previous colonial power(s) and status changes                                                                 |
|--------------------------|----------------------|-----------------------------|-------------------------------------------------------------------------------------------------------------|
| Cook Islands             | 1965                 | 1888                        | British Protectorate with New Zealand 1901 1965: autonomy Portugal Annexed by NZ in 1901                   |
| East Timor (Timor-Leste) | 2002 (restored)      | 1769 (Dili capital of Portuguese Timor) | Portuguese (1942–1945: Japanese occupation) After WW2: Portugal 1975: Indonesia 1999: UN until independence in 2002 Spain 1899: Germany 1947: US-administered UN Trusteeship Current compact to be reviewed in 2001 Germany |
| Federated States of Micronesia | 1987                 | 1885                        | Spain 1899: Germany 1913: Japan 1946: US-administered UN Trusteeship                                        |
| Fiji                     | 1970                 | 1874                        | After WW2: US-administered UN Trusteeship Great Britain 1881: Rotuma becomes dependency 1970: independence 1987: Fiji becomes a republic British Protectorate 1915: British Colony of the Gilbert and Ellice Islands to independence Treaty with Germany 1886: German Protectorate WW1 to WW2 : Japan 1947: US-administered UN Trusteeship Current compact to be reviewed in 2001 Germany |
| Kiribati                 | 1979                 | 1892                        | 1914: Australian occupation 1919: joint GB, NZ and Australian administration (1942–1945: Japanese occupation) New Zealand |
| Marshall Islands         | 1986                 | 1878                        | 1946: US-administered UN Trusteeship 1981: Constitution German and British After WW1: Australia German 1914: annexed by New Zealand (First Pacific country to achieve independence) Northern Solomons German protectorate from 1885 to 1899 |
| Nauru                    | 1968                 | 1888                        | British Protectorate until independence in 1978 First missionaries in the 1830’s Constitution in 1875 Close ties with Great Britain until emancipation in 1970 |
| Niue                     | 1974                 | 1901                        | Spain 1899: Germany 1913: Japan 1946: US-administered UN Trusteeship                                        |
| Palau (Belau)            | 1994                 | 1885                        | German and British After WW1: Australia German 1914: annexed by New Zealand (First Pacific country to achieve independence) Northern Solomons German protectorate from 1885 to 1899 |
| Papua New Guinea Samoa   | 1975                 | 1884                        | British Protectorate until independence in 1978 First missionaries in the 1830’s Constitution in 1875 Close ties with Great Britain until emancipation in 1970 |
| Solomon Islands          | 1978                 | 1893                        | British Protectorate until independence in 1978 First missionaries in the 1830’s Constitution in 1875 Close ties with Great Britain until emancipation in 1970 |
| Tonga                    | Never                |                              | British Protectorate until independence in 1978 First missionaries in the 1830’s Constitution in 1875 Close ties with Great Britain until emancipation in 1970 |
| Tuvalu                   | 1978                 | 1877                        | Great Britain 1892: British Protectorate 1916: British Colony of the Gilbert and Ellice Islands until 1975 France and Great Britain French and British Condominium of the New Hebrides until 1980 |
| Vanuatu                  | 1980                 | 1887                        | Great Britain 1892: British Protectorate 1916: British Colony of the Gilbert and Ellice Islands until 1975 France and Great Britain French and British Condominium of the New Hebrides until 1980 |

Important for scientists to learn about and understand the nature of the weather and the climate.

For any NMS, operating a sound meteorological observational network and having the means by which all the data can be well managed to ensure its veracity are critical for maintaining accurate records. The records themselves have limited value, but when analysed using appropriate mathematical theory, the understanding of the nature of large scale atmospheric patterns, which drive the climate of a particular area, become well understood (e.g. Trewin, 2013).

Good data management involves a range of activities aimed at ensuring that the values derived from the instruments and the techniques used to observe meteorological phenomena are the most accurate estimates available. The installation of the CliDE software and the associated training supplied by the project team improved substantially the management practices for climate.
data in the western Pacific region, providing a central, secure location for each country’s climate record with quality flags, ability to record lineage, and regular backups.

Climatologists from an NMS should provide consultation and advice regarding:

- planning of station networks;
- location or relocation of climatological stations;
- care and security of observation sites;
- regular inspection of stations;
- selection and training of observers; and
- instruments and observing systems so as to ensure that representative and homogeneous records are obtained (WMO, 2011) (WMO-No. 100).

A well-maintained instance of CliDE can assist the climate section staff in many of those tasks.

Before the development of CliDE, a range of climate databases and electronic storage mechanisms were being used by countries in the Pacific and East Timor. For more than a decade, the most successful of these was the CLImate COMputing (CLICOM) system (WMO/NOAA, 1988; Lianso, 1994; Stuber et al., 2011).

CLICOM is a CDMS that was initiated in 1985 by the WMO, with the assistance of the National Oceanic and Atmospheric Administration (NOAA) to help small countries with (the) climate data management. CLICOM was developed as a PC-based database running under the Disk Operating System (DOS). Over several decades, CLICOM began to lose reliability on newer Microsoft (MS) Windows operating systems (particularly following the release of Windows 2000), with the last official upgrade to CLICOM being Version 3.1 Release 2, released in English in January 2000 (a French translation was released in February 2001).

Over the years, many computer backup files were damaged though disk crashes, PC failures and so on, and the loss of historical electronic data was notable in countries around the world. For example, a data rescue effort was led in Papua New Guinea (PNG) to recover data from a magnetic tape containing CLICOM data that could no longer be read using the available technology in that country. The disk was taken to Melbourne, Australia, where the data files were recovered onto modern storage media. Without such efforts, the electronic data would have been lost permanently, resulting in considerable time and money required to re-digitize those observations from available paper records, to the extent that these still existed.

Across the partner countries in PACCSP, only the Solomon Islands, Fiji and to some extent Niue were found to have maintained operational CLICOM systems at the start of the CliDE project in 2009. However, at the time of writing, only Fiji continues with an operational standalone CLICOM database running on MS Windows 98. As CLICOM became increasingly unreliable under later versions of the Windows operating system, many countries began to use electronic spreadsheets, most notably Microsoft Excel, as a readily available and user-friendly software application to record data and produce reports and graphs. However, spreadsheets provide few or no data validation tools, pose difficulty around quality assurance and version control, and have capacity constraints, particularly for the ingest of high frequency data that can exceed quickly the number of records supported by such software. Spreadsheets also pose significant issues for product generation as the ability to fashion lineage, and regular backups. The dependence on and reluctance to give up easy-to-use spreadsheets meant that the CliDE team had to spend a considerable amount of time explaining the benefits of collecting and storing the climate data in a relational database. The NMSs in many partner countries have limited climate service sections and the focus has been on collecting daily observations for weather monitoring and aeronautical applications.

Finding a suitable replacement for CLICOM proved difficult. Problems such as significant capital and re-current costs, a lack of ease-of-use, poor support and labour-intensive maintenance were all encountered when evaluating alternative CDMS.

2.3. The CliDE project

In October 2009, the PCCSP commenced the development of the CliDE software, which has since been installed in all 15 of the programme’s partner countries. The total budget for the data rescue activities undertaken during the combined PCCSP/PACCSP spread over 4 years was ~US$2.3 million, covering software development, hardware purchases, travel, workshops and training. The total included the development of CliDE and seeding the database, through data rescue and digitization. The motivation for including the production of a new CDMS and associated data rescue activities in a science programme was to support climate change adaptation. Ensuring improved data management, increased data security and increased data access were seen as important long-term legacies that would support and expedite climate research in the current as well as future science programmes.

The circumstances found in the Pacific region are by no means unique, meaning that CliDE and associated practices may have broader applicability. Specifically:

- digital data availability was slow and limited as many records were stored in spreadsheets;
- large volumes of data remained only on paper records and/or offshore and the efficient digitization and/or extraction of these data was not feasible;
- data security was a recurring issue, with limited disaster recovery plans and backups;
- data audit trails were limited, for example, knowing whether a piece of data was as-read or in some way modified was not possible with existing practices; and
- considerable time was being taken up with maintaining existing datasets, rather than allowing the staff to focus on value adding, including derived products and services.

The development of the CliDE software was only one aspect of the project. Issues such as records management, repatriation of climate data, in-country user training, technical capacity building, suitable hardware, and building personal relationships and trust were all considered and resourced appropriately during the planning stages.

The CliDE project has built on earlier data rescue work in the Pacific region including the ‘Climate Data Rescue in the Pacific: a first step’ project, which commenced in July 2005. The project undertook data rescue activities in five Pacific island countries: Papua New Guinea, the Solomon Islands, Vanuatu, Fiji, and Kiribati. A second and related project, ‘Building robust and reliable data monitoring infrastructure for climate change monitoring’, commenced in June 2006 and helped develop a robust infrastructure for the collection, storage, quality control, and access to climate data, which are essential for adapting to climate variability and change. These projects were undertaken as a partnership between the Australian Bureau of Meteorology, NOAA, and New Zealand’s National Institute of Water and Atmospheric Research (NIWA) under the Australian Greenhouse Office – Bilateral Climate Change partnership programme.
3. Climate data rescue

Data rescue is the ongoing process of preserving data at risk of being lost due to deterioration of the medium and digitizing current and past data into computer compatible form for easy access (WMO, 2013). Meteorological data have been stored on paper manuscripts for over a hundred years, in some cases, and are in need of preservation because paper records decay over time.

The partner countries have a meteorological history that has resulted in different ways of recording observations and, in a few cases, different units of measurement. CliDE needed to be able to handle the climate data storage needs of each of the partner countries. The WMO has produced Guidelines on Climate Data Rescue (Tan et al., 2004), which has been used to guide data rescue efforts in PCCSP/PACCSAP.

All countries should understand and provide for the climate-related information needs of the public and more specialized users. This understanding requires meteorological observations, management and transmission of data, various data services, climate monitoring, practical applications and services for different user groups, forecasts on intraseasonal to seasonal time scales, climate projections, policy relevant assessments of climate variability and change, and the research priorities that increase the potential benefits of all these activities. Many countries in the western Pacific region do not have sufficient individual capacity to perform all of these services.

The World Climate Conference 3 in Geneva in 2009 led to the creation of a Global Framework for Climate Services (GFCS) to strengthen the production, availability, delivery, and application of science-based climate prediction and services. The Framework intends to provide a mechanism for developers and providers of climate information as well as climate sensitive sectors around the world to work together to help the global community better adapt to the challenges of climate variability and change (WMO, 2011).

3.1. Recording booklets for manual observations

The meteorological observations conducted in the partner countries are influenced by New Zealand, Australia, the United States of America, Indonesia, Portugal, and, to a lesser extent, France and Germany (Shepherd, 1926; Macellan, 2000). Many observation registers in the partner countries reflect the practices of the pre-independence governing bodies. Most of these standards meet international practice in the observing and recording of meteorological observations, but there are sufficient differences between them to make it difficult to define logical connections between all the observations and storage locations in CliDE. Those differences include different observation times, different instrument shelters and different meteorological element definitions.

Colonial administrations managed NMSs by providing funding and expatriates who usually administered the office and local staff. This continues to influence the structure of most NMSs and has left a legacy of work flows (i.e. meteorological observation and recording practices) that are common among clusters of countries.

The largest influence that the administering countries have had is in the design and content of the observation registers and office work flow. Observation of meteorological data is recorded in field books (registers) and across the partner countries. There is a handful of major variations in the booklets used for recording synoptic observations: The Recording of Meteorological Observations (A8), Australia; Field Book for Weather Observations (MET801), New Zealand; Field Book for Climatological Observations (MET303), New Zealand; Observations Register, Vanuatu; Meteorological Observations (MET301), New Zealand. Variations in these are in current use and have been identified by the CliDE team. While most meteorological parameters observed are common across the partner countries, the units of measurement and manuscript layouts differ between countries, meaning that efficient data entry requires variation in data entry forms.

For example, in those countries influenced by New Zealand there is a mixture of Douglas and Beaufort Codes, Imperial (IMP) and the International System of Units (SI) that have changed over time, but most have moved to SI in recent years. Exceptions to this are the countries from northwestern Oceania, which follow USA imperial measurements (e.g. continuing to use inches as the units for rainfall and pressure).

As another example of the different units of measurement being used by partner countries, New Zealand practice uses the Beaufort letters and symbols for recording wind velocity, cloud and weather phenomena and, thus, has continued in those countries where New Zealand has had an influence. Countries influenced by Australia generally use WMO code form 4677 (WMO, 2012a) to record weather phenomena. The Beaufort weather notation is difficult to map accurately (i.e. difficult to define an exact logical connection) to WMO code 4677. For CliDE in this case, it was decided to record simply the scale in its original form without translating it to WMO Code 4677. In this way, the intent of the observer is recorded, although the data cannot be standardized.

3.2. Implications for CliDE development

From early in CliDE’s design phase, the differing requirements of all 15 partner countries were taken into account. The CliDE development team considered the commonality of meteorological observations made globally, which allowed a basic data definition to be captured from all participating partners. A selected set of non-SI units was also captured and those units were converted automatically to SI when entered via the CliDE data entry forms. For file ingest though, the non-SI units must be included explicitly in the input file. Both the original non-SI values and the converted SI values are kept in the CliDE database. When generating outputs, users can select to develop many products in SI or non-SI units as required.

Across the different nations, observers and climate staff record meteorological observations using different observation register layouts. The keyboard data entry pages therefore allow a degree of customization to ensure that data entry is familiar and rapid. The data entry forms include synoptic observations, climate station observations, and METAR/SPECI (aviation weather reports (WMO, 2004)).

Figures 2 and 3 show an example of the sub-daily keyboard data entry form in CliDE and how the field names resemble closely those on New Zealand’s MET801 manuscript, ensuring that CliDE’s interface is familiar to trained meteorological observers.

4. Software architecture

Development, analysis and construction were undertaken as a collaboration between an information technology (IT) expert and functional climate and climate data management experts, which allowed for rapid development of a working product. With a constrained scope and a small, agile development team, it was less than 12 months from the start of software development (including requirement gathering, hardware selection and early...
Design and development was assisted by early feedback from end users at a project workshop held in Darwin in June 2010. The initial emphasis was on developing CliDE to address the functional foundations: an applicable data model and an easy-to-use graphical user interface (GUI) that supported data rescue of paper transcripts using a web-based interface that would be intuitive to users. Additional functionality and features would come later.

Simplicity was favoured over functionality so as to enable technical support in-country by staff without technical training. Some general technical and network administration skills are well developed in the partner countries: technical support staff are experienced generally in providing computer desktop support, developing and maintaining small office networks, and monitoring and maintaining meteorological equipment. However, software development skills are rather limited and unlikely to improve substantially in the near future.

A key to designing the CliDE software was to form a functional separation between the data ingest/input (including keyboard data entry), the climate database, and finally the product generation or services layer (Figure 4). This provided a separation of concerns, dividing the application into distinct features with little overlap, promoting extensibility, and minimizing maintenance conflicts. The separation allows for future independent upgrades of these components, noting, for example that some components of the technology such as web services remain in a state of rapid international development, whereas database technologies are rather more mature. It also gave CliDE the security provided by separation of duties that allows larger meteorological services the ability to restrict the amount of power held by any one individual and hence limit the potential damage to data and services.

4.1. Design considerations and constraints

There were numerous inputs and constraints that were identified during the design phase for CliDE:

- computer literacy of users in the NMSs was sometimes limited and resources were limited: software complexity and training requirements needed to be minimized for basic functions, whilst still allowing more complex use by experienced users;
- CliDE needed to cater to several sovereign nations and so an emphasis was placed on providing a foundation capability that all members could use;
- ongoing costs such as licensing, software patches, and upgrades had to be low or preferably zero: the system could not be a financial burden or commit partner countries to future licence costs;
- power supplies are less reliable and interruptions inevitable in many of the partner countries that rely on small scale
electricity generation (often supplied from diesel generators) and so the design needed to be robust and data integrity ensured when the system went down due to electrical power issues;

- user interface design needed to be lightweight in a technical sense so as not to overload local computer networks or web servers;
- a reliable Internet connection could not be assumed: the system needed to operate without external network access, being able to run on the NMS’s Local Area Network (LAN);

- the need for many countries was urgent: as noted previously in PNG and as subsequently found in other countries, data security was a concern and many paper records were at risk of being lost permanently before they could be digitized.

The lack of funding for climate controlled computing environments combined with the limited experience and expertise in computer systems management has resulted in operating environments that contribute to increased rates of failure in power supplies and data storage devices such as hard disk drives. So,
the choice of hardware and where to place it became important decisions for the stability and long-term success of CliDE.

In addition, there was a desire to build a system that was independent of specific choices of hardware and operating system. This arose out of the need to be sustainable beyond the life of PCCSP/PACCSAP, with individual countries having varying degrees of access to new hardware and replacement parts, and local system administrators having different operating system preferences.

CliDE was developed as free and open source software, released under the terms of the GNU General Public License (GPL) (Free Software Foundation, 2011), with a familiar web-based user interface and highly reliable relational database management system. In this context, it is free for the international community to use for improving climate data management and services.

4.2. Web application

An early design decision was to build CliDE using web technologies; in particular the user interface was written in HTML, JavaScript and the PHP programming language served by the Apache HTTP Server (web server).

Figure 5 shows the typical network setup of CliDE in the partner countries, with the web server and database server running on a single piece of hardware accessed over the LAN by several client computers located around the NMS office.

The reasons for using web technologies rather than develop CliDE as a desktop application included:

- Web languages and tools are written necessarily to global standards, meaning that the skill set required to expand, enhance, or alter any logic is commonly available and cost effective.
- The application can scale up from one user to more than 100 users without changes to the application, extra software installations, or licence fees.
- Deployment and upgrades are easier with a single central location for the application software (i.e. the CliDE server), providing for a single version of CliDE and associated data in each NMS.
- High security: total control over access to climate data by database administrators and owners.
- A web server makes automatic input and output feeds easier to implement, with message handling, file transfers, and asynchronous processes easily catered for in a web server environment.
- Separation of product generation from the database: while CliDE provides products such as graphs and summaries, the development of more sophisticated services can occur in a separate service layer that communicates to the CliDE database using open standards.
- Independent of the client’s operating system: as long they have compatible web browser software, they will be able to use CliDE, including on mobile devices such as tablets (this also means that the client’s operating system can be upgraded without an impact on the use of CliDE).
- If someone wants a ‘stand-alone’ application running on a single desktop computer, that can still be accomplished by installing all the software dependencies on the same computer and used locally with no network connections.

A feature of the choice of this web architecture is that CliDE can be run across a range of operating systems including GNU/Linux, Microsoft Windows, and UNIX, and has also been demonstrated to work on a cloud computing service. CliDE installations can have a heterogeneous hardware environment and still can run securely and efficiently.

4.3. The CliDE data model

4.3.1. Relational database

Data storage techniques have evolved over time to reflect the increased capabilities of hardware and the increasing demands placed on software. In recent years, relational databases have become the de facto standard for managing large amounts of well-defined, structured, transactional data. This method stores...
data in ‘tables’ that organize the data to minimize duplication and maximize data integrity. A relational database management system provides:

- a standard interface via SQL (or similar) that allows different tools to access the data in a consistent way;
- multituser sharing of data, meaning that multiple users can access the same information without a need for data duplication;
- security systems to stop unauthorized access and to reduce the risk of unintentional changes to data or loss of data;
- search and retrieval;
- data validation;
- referential integrity;
- no repeated data; and
- rules describing how data can be applied.

4.3.2. Transforming old data for ingestion into CliDE

As part of the process of loading data into CliDE, transforming data held in spreadsheets across the partner countries into a common file format was a time-consuming task. Due to the use of user-defined fields in CLICOM, the team encountered problems when migrating data from old databases across different countries. Country-specific mapping tables and rules needed to be developed to ingest data from CLICOM databases into new CliDE installations. Each installation of the CLICOM database was different, with different countries using different database elements for storing observations. CLICOM provides the user with a very flexible approach to element description, identifiers, and units. Each element in CLICOM is given a number by which a meteorological parameter is identified and that number also includes a name, description, and the units of measurement. Those CLICOM numbers were used to map data to fields in the CliDE database. In some countries, users used the standard set of CLICOM element descriptions and identifiers, but at times a different identifier and/or units of measurement were used for the same element.

For example, one country stored the same element in two datasets in CLICOM with different identifier numbers:

- Maximum air temperature had two IDs as 2 and 177, in two different datasets in CLICOM.
- Minimum air temperature had two IDs as 3 and 178, in two different datasets in CLICOM.
- 24 hour accumulated rainfall had two IDs as 5 and 176, in two different datasets in CLICOM.

CliDE has only one location for each element in question and thus, an awareness of such ambiguities and a policy to deal with them appropriately was required on a country by country basis.

Spreadsheet holdings proved more complex as spreadsheet layouts differed from country to country and these structures had tended to change through time.

Spreadsheets by nature have limited structure so elements located in cells in one country were unlikely to be in the same cell in another country. Extracting data held in spreadsheets was one of the most time consuming tasks the CliDE team embarked on in moving PACCSAP partner countries to a CliDE installation.

4.3.3. A well-defined data model

Practice at the Australian Bureau of Meteorology has shown that data entry and management are more efficient when the observed meteorological data are stored in relational table structures and these were adopted in CliDE. So, in contrast to CLICOM, CliDE uses strictly defined elements, using the Australian Data Archive for Meteorology (ADAM) database schema as a basis (see Australian Bureau of Meteorology, 2010).

Linking observations to a particular location was an important design issue, which continues to pose a challenge for climate data management. Almost all of the partner countries had a local numbering system for stations that provided a unique identifier for each station, though this number was not usually the same one used for international reporting through the WMO numbering system (WMO, 2012b).

Following the design for ADAM, in the current version of CliDE observations are linked directly to stations using the local station number as the unique identifier. Such an approach provided the closest match to the existing partner country practice and supplied the database with the required referential integrity between stations and observations. This design should be revisited with a more flexible proposal to link observations to instruments and a station in the database should then be composed of multiple instruments. This may be a necessary extension to cater for multiple instruments (for example from Automatic Weather Stations, AWSs, in the same enclosure) and parallel observations. Currently, there is difficulty in preserving ‘duplicate’ measurements such as those that can occur when manual and automatic weather stations are trialled concurrently.

There are several different observation tables in the CliDE database, each one recording a different category or frequency of observation: daily, sub-daily (for synoptic observations), monthly, AWS, upper air, and aero (for METAR message data predominantly used for aviation purposes) (see the table names in the database diagram in Figure 6). If necessary, database views and CliDE reports can be used to bring data together from different tables and display it in a coherent way to the user.

4.4. PostgreSQL database management system

DBMS software organises data into logical table structures and manages the data on the computer ‘disk’ system. PostgreSQL is an open source object relational database management system that has had more than 15 years of active development (PostgreSQL, 2012). It is a mature software that is applied widely including the Australian Bureau of Meteorology for operational climate services. It is used world-wide in small to medium businesses and can handle the expected data volumes and load anticipated for CliDE.

PostgreSQL was selected as the DBMS for CliDE because of the existing expertise in using it within the Australian Bureau of Meteorology, availability for multiple operating systems, its free and open source software licence, and its geospatial functionality when combined with PostGIS for a possible future map-based user interface for data editing and reporting. PostgreSQL has proved to be fast, reliable and secure, requiring very little administration or performance tuning for operational deployment. It has been tested under load in Australia with positive results.

4.5. GNU/Linux operating system

CliDE is not dependent on any particular operating system. Any operating system that can run the chosen database system (PostgreSQL), programming language (PHP), and web server software (Apache HTTP Server) can be used. However, the team did need to make a decision about which operating system to install on the supplied hardware when deploying it to the partner countries.
Computer viruses were observed as a major problem on computers running Microsoft Windows in the partner countries, to the point that memory sticks and laptop hard drives used in the countries had to be discarded because of massive infestation. Prior experience at workshops and on training visits was that encountering multiple viruses on operational computers was common place. A particular issue was the lack of institutional level security, meaning that while single computers may be disinfected, reinfection was almost inevitable over the LAN. Therefore, other operating systems that were less susceptible to computer viruses were investigated and GNU/Linux provided an excellent free and open source alternative. GNU/Linux can run on low-cost PCs but can also scale to clusters of high-end servers and even supercomputers.

The Ubuntu 10.04 LTS Linux distribution (Canonical Ltd, 2012) was selected as it was the most popular GNU/Linux distribution available at the time (Zachte, 2011; Noyes, 2012) with good support available through both community and commercial channels. This distribution of GNU/Linux (referred to commonly as Linux, without the 'GNU') contains all of the technical advantages of commercial UNIX with the added advantage of modern software package managers to simplify software administration.

In addition to computer viruses, other factors that were considered before deciding to use Ubuntu Linux were documentation, community support, and ease of administration. For the small amount of system administration that may be required, Ubuntu 10.04 LTS provides a graphical desktop interface that makes Microsoft Windows users comfortable and can be tailored and used with minimal training. Command line UNIX commands can still be used in Ubuntu, providing powerful and flexible options for system administration. If necessary, the system can also be administered remotely provided a suitable network is available.

In terms of online support and documentation, the user community for Linux in general is large and Ubuntu Linux has excellent tutorials, discussion fora and ‘how to’ pages.

As the CliDE servers were being installed in many offices with limited server administration expertise, the decision was made to select the desktop edition instead of the server edition of the Ubuntu distribution. Moreover, it was about making the technical staff in the partner countries feel as comfortable as possible with the new server even if they had previously little or no experience with Linux.

5. Deployment to partner countries

High-end workstation hardware for the server plus an uninterruptible power supply (UPS) was freighted to each of the 15 partner countries ahead of in-country trips by the CliDE project team to install the new servers and provide both user and technical training. To minimize the risk of installation problems and avoid large and expensive Internet downloads in-country, the server software was installed on the workstations prior to shipping from Australia. For consistency and ease of future upgrades, the same hardware was selected for all the partner countries. As many of the countries did not have dedicated server rooms, a mini-tower form-factor server was chosen rather than rack-mount servers. Existing desktop computers in each NMS office were used as clients by the local staff.

Typically, two members of the CliDE project team travelled together to each of the partner countries for installation and training. Travelling to each country had the benefit of giving more local staff access to the training courses and to build an appreciation of the local workplace context in the trainers. Building personal relationships with local staff was another important advantage of these in-country visits, developing trust and a closer working relationship between the CliDE project team and local staff. A second round of in-country visits to assist with upgrading the CliDE software and provide updated user training followed. These follow-up visits also concentrated on finding ways to fit CliDE into the meteorological service work flow to help improve office productivity and strengthen the move away from past practices. With CliDE now meeting the basic requirement of supporting the security of climate records within countries, the emphasis is on improving and supporting the uptake and use of these data and the generation of higher value products and services. One motivation for supporting productivity is to raise awareness of the value derived from CliDE and data contained therein, and therefore encouraging its use and subsequent support for improved climate data management in partner countries.

6. Outcomes and benefits

CliDE now provides a stable, consistent database of meteorological observations and related data for all 15 partner countries. The development and installation of CliDE is already starting
to provide benefits in many of the partner countries. It has contributed to improve the productivity of staff in the NMSs, with tasks such as manual encoding of messages now able to be carried out automatically. The time taken to produce data files needed for climate research has been reduced greatly, with output formats from CliDE written specifically for use in products that are commonly used by the NMSs such as SCOPIC (for seasonal forecasts), RClimDex (for climate extreme indices), RHTests (homogenization tests), and the Pacific Climate Change Data Portal.

The following are examples of the range of situations in which CliDE has been used since its deployment.

6.1. Data digitization in Vanuatu

The Vanuatu Meteorology and Geohazard Department (VMGD), known formerly as the Vanuatu Meteorological Services (VMS), wished to digitize all historical sub-daily weather observations and with the deployment of CliDE, it became the destination for these data. By then 70% of Vanuatu’s climate data (i.e. all known historical observations of meteorological/climate elements) had already been digitized and the remaining 30% of known records were digitized under the PACCSAP digitization project into CliDE (Figure 7).

Vanuatu climate data had been recorded previously in CLIMatic SOFTware (CLIMSOFT) (Mhanda, 2002). Data were generally entered into an Electronic Monthly Register (EMR) and software macros used to pull data together and upload it into CLIMSOFT. As part of PACCSAP, those data were converted from the CSV text files from EMR into the CliDE Native file ingest format (CSV files with a well-defined format) for loading and storage in CliDE.

Eight students were employed for 4 months to digitize the remaining data by entering data directly into CliDE using the web-based user interface. Keyboard data entry took place on a network of thin client devices in the Climate section of the VMGD.

6.2. Data digitization for Fiji

Over the period from 1880 to 1973, the Colonial Sugar Refining Company, Australia, managed sugar production and milling in Fiji. To gain an understanding of the local climate, meteorological observations were taken at the five sugar mills and sugar sector offices. Copies of the original observation forms were sent to Australia and are archived currently at the Noel Butlin Archives, at the Australian National University in Sydney.

In 2010, the observation forms were scanned in Sydney and the images transferred to the Australian Bureau of Meteorology where some of the data have been digitized using CliDE. Particularly noteworthy is the digitization of the Nausori Mill record from 1882 to 1961 (daily data from 1905). A rainfall composite has been created with a neighbouring, currently operational, Nausori Airport record to create a 130 year time-series that is assessed as homogeneous (Figure 8). With 99.8% of the monthly record available, this is perhaps the longest most complete rainfall record available in the Pacific Islands. Copies of both the images and digitized data have been transferred to the Fiji Meteorological Service.

6.3. Storing high frequency AWS data in Samoa and Fiji

NIWA was contracted to install four automatic weather stations (AWSs) and develop a Climate Early Warning System (CLEWS) in Samoa. NIWA chose to use CliDE for the long-term storage of these automated high frequency weather observations.

Every hour, NIWA’s FloSys (NIWA, 2013) outputs CSV text files in the CliDE Native file format that are then sent using file transfer protocol (FTP) to the CliDE server. A scheduled job on the CliDE server then ingests any files it finds into the database, updating the climate data holdings automatically. Figure 9 shows a simplified flow of the data from the AWS network into CliDE. Manual observations are still entered directly into the CliDE user interface via the keyboard data entry forms.

The result is that CliDE in Samoa is a near real-time database of meteorological observations. Quality assurance checks on the
data can be readily undertaken and preliminary reports generated the same day.

A similar system has recently been installed in the Fiji Meteorological Service, automatically ingesting AWS data into Fiji’s CliDE database.

6.4. Data digitization for East Timor

Climate data for East Timor have been difficult to locate by the PCCSP/PACCSP team. For 400 years, the Portuguese administered East Timor, finally leaving in 1975 and replaced subsequently by an Indonesian administration. Formal independence came to East Timor in May 2002.

As a result of the changes in administration, meteorological observations have not been continuous. Manuscripts produced by the Portuguese administration have been located in the National Archives of East Timor (NATL) and some records from the 1970s and 1980s are held by the Indonesian Meteorological, Climatological, and Geophysical Agency (BMKG) in Jakarta, Indonesia.

The Portuguese era manuscripts covering the period 1951–1974 are currently being digitized in Dili at the offices of...
7. Future work

CliDE has proven itself as a robust and affordable solution in support of the climate data management requirements of partner countries in the western Pacific and East Timor. In this regard, the aims set down by PACCSAP and a priority list was created. It is hoped that the DNMG will continue with the digitizing after PACCSAP is completed.

Meanwhile, Dili records from between 1981 and 1999 that were sent to Darwin for safe keeping in 1999 (during a period of civil strife) were digitized into a local copy of CliDE at the Australian Bureau of Meteorology in Melbourne, Australia, and this work was completed prior to PACCSAP finishing. The manuscripts and data were repatriated to East Timor in May 2013 and presented by Australia’s ambassador to East Timor to the East Timorese Minister of Infrastructure. These data in combination will allow for the first analysis of climate change in East Timor noting that previous studies have found digital data holdings to be insufficient to document climate trends (Jones et al., 2013).

7.1. Homogenized time-series data

CliDE, as with the Australian Bureau of Meteorology’s ADAM database, was developed to store raw meteorological observations. However, deriving products from raw data is generally discouraged because quality control through data entry can be inadequate. Furthermore, for some towns and cities there are multiple station records through time due to site moves, meaning that long time-series often do not exist in the raw record.

To produce high quality products, there is a need to store homogenized time-series data in CliDE. Homogenized data archived in a CDMS is not currently being practised elsewhere. Homogenized datasets have their own metadata storage requirements, recording the concatenation of time-series from multiple stations, data break-points, and adjustments. In addition, there is a need to develop a flexible, extensible station numbering scheme for the homogenized datasets because they can be composed of observations from multiple observing stations. Currently, homogenized data are held in flat files outside of databases in the partner countries and, indeed, in Australia (e.g. Jones et al., 2013; Trewin, 2013).

If the complexities discussed above can be managed, storing homogenized data in CliDE would provide many new opportunities to develop products and would do away with the need for record duplication. The Expert Team on Climate Change Detection and Indices (ETCCDI) has produced indices that are currently being considered as potential products (Peterson et al., 1998; Karl et al., 1999). Other opportunities include linking software tools such as SCOPIC (Seasonal Climate Outlooks in Pacific Island Countries, http://www.bom.gov.au/cosppac/comp/scopic/) to CliDE. The two main components of SCOPIC are seasonal statistical climate outlooks and a drought monitoring tool. Both rely on historical monthly scale climate data archived on individual PCs or a connected file server, thereby requiring manual updating and risking differences to the main NMS database. These manual processes and risks could be avoided if SCOPIC were to use homogenized data directly from CliDE as its data source.

7.2. New database schema linking observations to instruments

Currently in CliDE, observations are linked directly to a station through the use of a local station numbers. If there are multiple instruments at the station’s location recording the same element (e.g. precipitation) then only one, primary, observation is stored. To store the observations for the other instruments requires the creation of additional station records in the database.

A proposed new database schema would link observations to instruments in the database and then stations are composed of a collection of instruments. However, this would require a major rewrite of parts of the CliDE user interface to allow the recording of observations against instruments. For example, there may be multiple rain gauges within a station’s enclosure, each requiring a place on the keyboard data entry forms and all being recorded against a particular station number within the database.

With the increasing use of AWS equipment, this change to the database schema and associated user interface changes would give CliDE important flexibility as a climate database into the future.

7.3. Disaster recovery and data backup

The workstation hardware provided to the partner countries included three hard disk drives that allowed the configuration of automated nightly internal data back-ups between drives. Establishing regular off-site backups still needs considerable time and effort in most partner countries for training and the development of disaster recovery plans. Best practice requires that a climate database is backed up frequently, that copies are held off-site (i.e. in another building, and preferably in an entirely different physical location) and that the reconstitution of data is tested and achievable within a defined tolerance window.

This future activity would establish an off-site disaster backup for each CliDE installation including remote support. Keeping in mind the meteorological and geohazards faced by all the partner countries, the off-site backups would ideally reside in another geographical location and from a technical point of view perhaps in another country. However, the backup would need strict security arrangements and governance to ensure that data ownership is preserved.

7.4. Interactive maps

The need to develop CliDE’s core functionality to work well on low-bandwidth networks, possibly with old client computers, meant that the user interface of CliDE was kept simple and utilitarian.

However, several partner countries are well equipped to handle a richer, interactive user experience. Therefore, a major new user interface can be developed, primarily map-based for viewing and
selecting stations. Keeping accessibility compliance in mind, the limited bandwidth and old computers being used in some of the partner countries, such a rich user interface could not be the only way to interact with CliDE, but instead would be an optional addition in environments with the necessary IT infrastructure and computing performance.

7.5. Support for multiple languages

The user interface of CliDE has been written exclusively in English. The need to develop rapidly a functional system meant that some localization and internationalization flexibility is currently missing, particularly with regard to interface languages.

Significant changes to the software are required to facilitate native language support in the interface and to work towards full internationalization of CliDE. This is particularly needed in countries such as East Timor where the use of English as a language is limited.

8. Conclusion

Climate Data for the Environment (CliDE) has already proved to be an important tool for climate data rescue activities and climate data management in small national meteorological services (NMSs) in the Pacific and East Timor. Built on readily available open source software tools, it can be deployed on a variety of operating systems and hardware from small netbooks to multi-server cluster environments. CliDE is available on request from the authors of this study with further information available through the project website at http://www.bom.gov.au/climate/pacific/about-clide.shtml.

The decision to develop CliDE using free and open source software tools has proved to be the right one, giving the partner countries high quality software with no ongoing licence costs and the confidence that they have all the necessary code available to fix problems and add features in the future if they wish. CliDE has been well received in the partner countries and is being integrated into the daily and monthly work flows of the local NMSs. In countries such as Vanuatu, Samoa, Fiji and East Timor, it proved to be a vital tool for digitization activities that have contributed to a major expansion of digital data available for subsequent use and analysis, across both climate services and research. Combined with the data rescue efforts of the Pacific Climate Change Science Program/Pacific-Australia Climate Change Science and Adaptation Planning Program (PCCSP/PACCSAP) returning data to the country of origin, CliDE has helped improve the standing of the NMSs as they receive attribution as the source of climate data for their countries.

Some remaining challenges include ongoing user training, disaster recovery systems and the potential fragmentation of the developer community. Almost all of the partner countries have received two or more visits from the members of the CliDE project team for user training. Partner Country climate staff also attended two workshops at which there were CliDE activities and discussions. While the second round of in-country user training greatly improved people’s use and understanding of how to get the most benefit out of CliDE, there remain strong calls from the partner countries for additional in-country training visits. Ongoing support to users of the CliDE system is strongly recommended to allow the region to reap all the benefits of having a modern climate data management system (CDMS).

Disaster recovery planning is critical in locations that are vulnerable to natural disasters. However, it is often difficult to find suitable locations for secure off-site data backups, particularly in the coral atoll nations of the Pacific Ocean where past tropical cyclones and storm surge events have seen inundation of entire islands.

The long-term future of CliDE as a useful climate database in the partner countries may well rely on local technical staff being able to understand and improve the software beyond the life of PACCSAP. Alternatively, if sufficient open source development infrastructure and documentation can be established, funding for short-term contractors may also provide a stream of improvements to the software that can benefit all users.

It would be disappointing if the goodwill and focus of people wanting to see the development and success of a completely free and open source climate data management system was divided between different versions of CliDE. It is easy to copy the source code and start independent development of (i.e. fork) an open source project and split potentially the community supporting the project (Wheeler, 2007). Important though they may be, there is only a small global market for CDMS and to have that community split between competing and incompatible versions of CliDE would diminish the resources available for improving the installation of CliDE of any one country.

Acknowledgements

This work was undertaken under the International Climate Change Adaptation Initiative (ICCAI), in which Australia is providing assistance to vulnerable countries in the Asia-Pacific region to meet high priority climate change adaptation needs. The Pacific Climate Change Science Program (PCCSP) and the Pacific-Australia Climate Change Science and Adaptation Planning Program (PACCSAP) were components of the ICCAI project, improving understanding of the physical climate system to inform effective adaptation.

Thanks are due to colleagues in the partner countries who have worked closely with this study to improve CliDE. Other project team members and staff at the Australian Bureau of Meteorology whose input has influenced its architecture, William Wright, Yuriy Kuleshov and Janita Pahalad, are also acknowledged. Thanks go to Andrew Charles and Terry Johnson for reviewing this study and to the anonymous reviewers whose comments have greatly improved the study.

References

Australian Bureau of Meteorology. 2010. ADAM (Australian Data Archive for Meteorology) Manual Version 2.5. Australian Bureau of Meteorology: Melbourne.

Australian Bureau of Meteorology and CSIRO. 2011. Climate change in the Pacific: scientific assessment and new research. Volume 1: Regional overview. 257. http://www.pacificclimatechange.science.org/publications/reports/(accessed 10 January 2013).

Canonical Ltd. 2012. About Ubuntu. The Ubuntu story. http://www.ubuntu.com/project/about-ubuntu (accessed 23 November 2012).

Free Software Foundation. 2011. Licenses – GNU project – Free Software Foundation (FSF). http://www.gnu.org/licenses/ (accessed 19 December 2012).

Hales S, Weinstein P, Souares Y, Woodward A. 1999. El Niño and the dynamics of vectorborne disease transmission. Environ. Health Perspect. 107: 99–102.

Holmes RL. 1877. The climate of Fiji: results of meteorological observations made at Delanasau, Bay of islands, North Coast of the province of Bua, Vanuatera, Fiji, during the five years ending December 31st, 1875. Q. J. R. Meteorol. Soc. 3: 260–277, DOI: 10.1002/qj.4970032103.
Holmes RL. 1881. The climate of Fiji: result of meteorological observations taken at Delanaua, Bua, Vanua Levu. Fiji. 1871–1880. Q. J. R. Meteorol. Soc. 7: 222–243. DOI: 10.1002/qj.4970074004.

Jones DA, Collins DA, McGree S, Trewin B, Skillling E, Diamond H, Fa'amanu O, Hiriasis D, Hugony S, Hutchinson R, Inape K, Jacklick L, Malsale P, Moniz T, Ngemaes M, Porteous A, Seuseru S, Tahani L, Teimutiti F, Toorua U, Vaimene M, Vunilagi V, Vavae H. 2013. An updated analysis of homogeneous temperature data at Pacific Island stations. Aust. Meteorol. Mag. 61: 285–302.

Karl TR, Nicholls N, Ghazi A. 1999. CLIVAR/GCOS/WMO workshop on indices and indicators for climate extremes: workshop summary. Clim. Change 42: 3–7.

Krishna R. 2009. Brief History of the Fiji Meteorological Service, Information Sheet No.2. Climate Services Division, Fiji Meteorological Service: Nadi.

Lianos H. 1994. CLICOM: Overview and U.S. Contributions and Plans. U.S. CLICOM Project. NOAA/National Weather Service: Silver Spring, MD.

Macellan N. 2000. Colonial history. Australian Broadcasting Commission. http://www.abc.net.au/radio/colonies (accessed 27 June 2013).

Mhanda A. 2002. CLIMSOFT: CDMS evaluation workshop, WMO: Geneva. http://www.wmo.int/pages/prog/wcp/wcdmp/wcdmp_series/documents/Annex9.pdf (accessed 23 November 2012).

NIWA. 2013. Flosys. http://www.niwa.co.nz/software/flosys

Peterson TC, Easterling DR, Karl TR, Groisman P, Nicholls N, Plummer N, Peterson TC, Rahimzadeh F. 2004. Guidelines on climate data rescue. WMO WCDMP-55. http://www.wmo.int/pages/prog/wcp/wcdmp/documents/WCDMP-55.pdf (accessed 8 October 2013).

Teimutiti F, Toorua U, Vaimene M, Vunilagi V, Vavae H. 2013. A daily homogenised temperature data set for Australia. Int. J. Climatol. 33: 1510–1529. DOI. 10.1002/joc.3530.

Turner G. 1861. Nineteen years in Polynesia: missionary life, travels, and researches in the islands of the Pacific. John Snow, Pater-noster Row: London. Digital version available at Google Books. http://books.google.com/.

Wheeler DA. 2007. Why open source software/free software (OSS/FS, FOSS, or FLOSS)? Look at the numbers! http://www.dwheeler.com/oss_fs_whyno.html#forking (accessed 23 November 2012).

WMO. 2004. WMO technical regulations volume II, Meteorological Service for International Air Navigation. http://www.wmo.int/pages/prog/www/WIS/Publications/49%20%28Technical%20Regulations%29/49%20-%20Volume%2049%20_I_E.pdf (accessed 10 January 2012).

WMO. 2011. Guide to Climatological Practises, 3rd edn. World Meteorological Organization, Geneva. http://www.wmo.int/pages/prog/wcp/ccl/documents/WMO_100_en.pdf (accessed 18 May 2013).

WMO. 2012a. WMO manual on codes volume I, part A – alphanumeric codes. http://www.wmo.int/pages/prog/www/WMOCodes/WMO306_v11/VolumelI.html (accessed 30 November 2012).

WMO. 2012b. WMO publication No. 9, Volume A, Observing stations and WMO catalogue of Radiosondes. http://www.wmo.int/pages/prog/www/om/omVolume-A/vola-home.htm (accessed 30 November 2012).

WMO. 2013. Data rescue projects and initiatives (DARE). http://www.wmo.int/pages/prog/wcp/wcdmp/CDM_2.php (accessed 10 June 2011).

WMO/NOAA. 1988. CLICOM System Reference Manual: Data Management Module, Version 2.1, USA: World Climate Data Program, NOAA.

Zachte E. 2011. Wikimedia traffic analysis report – operating systems. 11 January, 2011. http://stats.wikimedia.org/archive/squid_reports/2010-12/SquidReportOperatingSystems.htm (accessed 23 November 2012).

Software Glossary/References

Apache HTTP Server: open source web server software. http://httpd.apache.org/

CLICOM. http://www.wmo.int/pages/prog/wcp/wcdmp/clicom/index_en.html

PHP: a general-purpose server-side scripting language originally designed for web development to produce dynamic web pages. http://www.php.net/

PostgreSQL: an open source object-relational database management system (ORDBMS). http://www.postgresql.org/

Ubuntu: a computer operating system based on the Debian Linux distribution and distributed as free and open source software. http://www.ubuntu.com/