

# Development of an enhanced tropical cyclone tracks database for the southwest Pacific from 1840 to 2010

H. J. Diamond,<sup>a,\*†</sup> A. M. Lorrey,<sup>b</sup> K. R. Knapp<sup>c†</sup> and D. H. Levinson<sup>d†</sup>

<sup>a</sup> NOAA's National Climatic Data Center (NCDC), Silver Spring, MD, USA

<sup>b</sup> National Institute of Water and Atmospheric Research (NIWA), Auckland, New Zealand

<sup>c</sup> NOAA's National Climatic Data Center, Asheville, NC, USA

<sup>d</sup> US Forest Service, Redding, CA, USA

**ABSTRACT:** The ecosystems and economies of small island nation states and territories of the tropical southwest Pacific region are widely agreed to be among the most vulnerable to climate variability and weather extremes anywhere in the world. Tropical Cyclones (TCs) are capable of exacerbating existing hazards and those made more severe by climate change (e.g. local sea level rise). In order to properly understand TC impacts in this region, a comprehensive database of TC tracks is required. This work has collated TC best track data from forecast centres around the globe with the aim of producing a unified global best TC track dataset for the historical period. Data from the International Best Tracks for Climate Stewardship (IBTrACS) project information for the southwest Pacific (135 E–120 W, 5 –25 S) is built upon and included in this effort. We document the construction of an enhanced TC database for the southwest Pacific, the quality controls needed to construct the database, and discuss how it has enhanced the chronology of region-wide historical TC activity in light of newly discovered data. We suggest this enhanced dataset can be used in forthcoming climate and weather studies to better characterize the climatology and behaviour of TCs in the southwest Pacific. Copyright © 2011 Royal Meteorological Society

**KEY WORDS** tropical cyclone; best tracks; data rescue; southwest Pacific; IBTrACS; visualisation tools; GoogleEarth®

Received 3 May 2011; Revised 25 June 2011; Accepted 28 June 2011

## 1. Introduction

The ecosystems and economies of small islands of the tropical southwest Pacific region (see Figure 1) are vulnerable to climate variability and change (Mimura, *et al.*, 2007). The region regularly experiences climate extremes such as droughts and deluge related to motions of the South Pacific Convergence Zone (modulated by the El Niño-Southern Oscillation (ENSO)) and also Tropical Cyclones (TCs). TCs are capable of exacerbating existing hazards driven by climate change (e.g. eustatic sea level rise; storm surge etc. Terry and Etienne, (2010)). The small island nations and territories in this region share a number of distinctive characteristics that increase their sensitivity and vulnerability. Apart from their small physical size, they are: (1) surrounded by large expanses of ocean; (2) highly prone to natural disasters, particularly extreme events such as TCs and droughts; and (3) relatively isolated from world economic markets. All

these factors make the region highly sensitive to natural disasters, and the impacts due to altered frequencies and intensities of extreme weather, climate and sea-level events are very likely to change (Parry *et al.*, 2007). Since 1950, for example, natural disasters have affected more than 3.4 million people and caused over 1700 reported fatalities in the region, excluding Papua New Guinea, and TCs accounted for 76% of the reported disasters from 1950 to 2004 (World Bank, 2006).

While there have been many TC studies in other ocean basins, there has been a dearth of research in the southwest Pacific Basin due to: (1) the sheer vastness of size and the low population density of the basin; (2) a very limited amount of observational data; and (3) the lack of a collated, comprehensive, and homogeneous database of TC information. Given the limited amount of South Pacific region data, and the devastating effects that TCs can have on the development of small island nations, it is vital to develop a more definitive history of TC activity so that climatologies can be established. Therefore, the effort described in this paper will clearly benefit the scientific community that is striving to assist regional aid coordination authorities and Pacific Island National Meteorological Services (PINMS) to better prepare for the impacts of TCs.

\*Correspondence to: H. J. Diamond, NOAA/NCDC, 1100 Wayne Avenue, Suite 1202, Silver Spring, MD 20910, USA.  
E-mail: howard.diamond@noaa.gov

† The contributions of these authors to this article were prepared as part of their official duties as United States Federal Government employees.

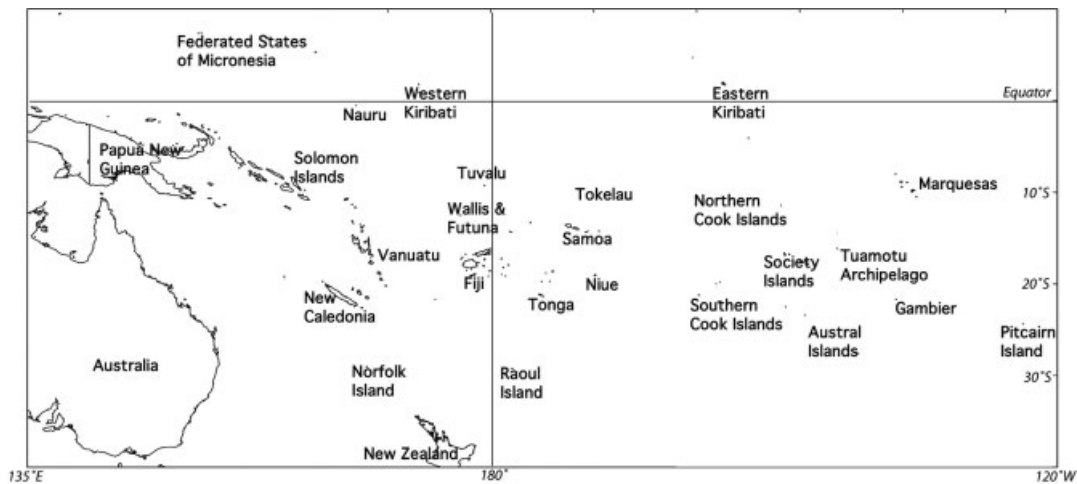


Figure 1. Base map of the southwest Pacific, showing the location of countries, territories, and islands in the region.

While there has been some considerable work done in documenting TCs in the southwest Pacific Basin (Visher, 1925; Giovannelli, 1952; Kerr, 1976; Lourenz, 1981; Revell, 1981; Thompson *et al.*, 1992; Maunder, 1995; Kuleshov, *et al.*, 2008, and Terry and Gienko, 2010), we do not necessarily believe that this has resulted in a consolidated and quality controlled database of TCs in the region. The work done to date has been very good; however, we believe that the integration of previous work needs to additionally: (1) tie these various sources together; (2) digitize track information along with post-1969 satellite data allowing for further statistical and spatial analysis; and (3) make the data easily available in convenient formats and at a centralized, secure location in line with the work of the International Best Tracks for Climate Stewardship (IBTrACS; Knapp *et al.*, 2010). The 1970 TC season is generally considered to be the first reliable year due to the availability of geostationary satellite data (Holland, 1981). In the Southern Hemisphere, the TC season straddles the change of the calendar year; therefore, the 1970 season extends from November 1969 through April 1970, and the season is thus labelled as 1970. This will be the seasonal convention utilized throughout this paper.

We approached this problem by collecting data from the sources listed above to enhance the spatio-temporal coverage of the IBTrACS database. The resulting integrated data set has been termed the Southwest Pacific Enhanced Archive of Tropical Cyclones (SPEARTC). Here we describe the steps leading up to the SPEArTC database, including data rescue and scanning, and the quality control method referred to as Graphical Interpretation of Tracks (GrIT), undertaken by using readily available GoogleEarth<sup>®</sup> software. We discuss the benefits of using this interactive visualisation approach to make judgements about historical track data, and discuss how the refined characteristics of the overall SPEArTC dataset have improved on the IBTrACS data.

While the SPEArTC dataset is built upon the work accomplished by IBTrACS, the dataset itself is different in three primary areas: (1) in addition to adding a

number of new records, the QC process described in Section 3 has corrected a number of errors (Table AII gives a documentation of corrected errors from seasons 1970 to 2010) and concatenated a number of storms that were originally thought to be separate ones; (2) the data structure for SPEArTC focuses on a core set of 13 variables that are documented in Table AIII; and (3) while IBTrACS moved to carrying the actual wind and pressure data for all storms in its latest version, SPEArTC carries the averaged wind and pressure data through the 2008 season, and then begins using the IBTrACS convention for wind and pressure data beginning with data from the 2009 season. While IBTrACS has done a good job of collecting existing information, there are two things that the SPEArTC effort has brought to bear on improving the database of information as follows: (1) a detailed track-by-track quality control of the data for the entire record; and (2) the analysis and inclusion of TC information from other sources not previously considered that were provided by the PINMSs. As noted in Table AII, a total of 13 track errors were corrected in the era from 1970 to 2010, and while there may not be a great number of corrections, they are important in making for a better and more reliable database for use by researchers interested in this region.

## 2. Description and Geographic Extent of the Database

The SPEArTC database consists of existing IBTrACS TC tracks and new information recovered from historical archives of the regional PINMSs. The dataset covers the period from 1840 to 2010 in a discontinuous fashion and extends the record temporally by nearly 50 years. In addition, the dataset improves the spatial coverage of existing data (Trewin, 2008 and Knapp *et al.*, 2010) by adding data from supplementary sources; correcting errors via a thorough quality control of the data as described in Section 3. The TC seasons utilized for SPEArTC include 1841, 1848, 1865,

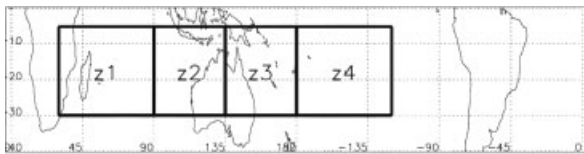


Figure 2. Zonal extent of the SPEArTC area of coverage includes regions z3 and z4 (from Leroy and Wheeler, 2008).

1871, 1873, 1875, 1877–1878, 1880–1883, 1886, 1888, 1890–1891, 1893–1895, 1897–1899, 1902–1903, and 1905–2010.

The SPEArTC dataset includes TCs from zones Z3 and Z4 (Figure 2) and ranges from 135° E in the Gulf of Carpentaria to 120° W, which includes French Polynesia. The latitude range is from 5° S to 65° S, which encompasses both the southern latitudinal extent of TCs (5° S–25° S) in addition to the latitudinal range over which extra-tropical transition (ETT) occurs (25° S–35° S; Sinclair, 2002). However, the SPEArTC information in the database is not restricted to the tropics and many tracks in the dataset extend south of 25° S. From 1970 to 2010, a total of 270 and 189 storms, respectively, are tracked south of 25° and 35° S in the SPEArTC dataset. The SPEArTC dataset only has one storm which extends as far south as 65° S. (The Coriolis force is generally too weak at very low latitudes to cause air to rotate around an area of low pressure; therefore, classical meteorological theory is that TCs do not form any closer to the equator than about 5° latitude (Gray, 1979)). This extratropical aspect to SPEArTC is an important element of the database, and is consistent with procedures at the Regional Specialised Meteorological Centre (RSMC) in Wellington, New Zealand (Ready, 2011). Extratropical storms often produce severe impacts and retain some tropical characteristics (e.g. high moisture content, warm core structure, and strong surface winds) for several days following ETT, and as such are important data for climatological studies in the region.

The region covered by SPEArTC is an informally defined area, different from World Meteorological Organization (WMO)-defined regions of responsibility for tracking TCs, but one that makes geographic sense. The formal area of coverage for southwest Pacific TCs as defined by the WMO is based upon the forecast areas of responsibility from the RSMCs in Nadi, Fiji, and Wellington (Figure 3), which indicates that the southwest Pacific extends from longitude 160° E and then east to 120° W. It turns out that many TCs which affect the southwest Pacific actually form over the Coral Sea (area IX of the map in Figure 3), thus extending the area of coverage for this study well to the west of the more artificial WMO boundary of 160° E in the Gulf of Carpentaria is much more relevant in constructing the SPEArTC dataset for the southwest Pacific.

### 3. Dataset Construction

This section describes the rescue and digitisation of TC data from across the region, the management of the

digital data in Excel®, and the quality control of the data using a newly developed technique known as GrIT as described in Section 3.3 below. GrIT was devised by the authors as an easy-to-use method for the visual inspection of TC tracks which employs off-the-shelf GoogleEarth® software. Inspection of the collective set of tracks using this visualisation approach can readily assist in the quality control of any basin's TC tracks. Compilation of track data for the SW Pacific showed at least three track morphology variants that could be identified. They were: (a) non-replicated tracks; (b) replicated tracks; and (c) erroneous or extraneous tracks with divergent points. Non-replicated tracks were single storms represented in the dataset that are not duplicated by any other storm track or partial track segment. Replicated tracks could be identified when TC tracks having similar spatial and temporal characteristics occur (e.g. starting and stopping within a seven-day window; having similar cyclogenesis points or trajectories). Erroneous tracks were simply storms that were obviously bogus, such as systems that followed the International Date Line or had digitising errors that included suspicious right angle turns or were digitized from south to north instead of from north to south.

#### 3.1. Rescuing and digitising TC data

Beginning in 2005, a survey of possible TC data sources was initiated in concert with the PINMSs in the region to identify old paper synoptic charts or data which were not incorporated into any digital TC database. The collection of source data builds upon excellent working relationships developed between scientists in Australia, New Zealand, the US, and a variety of Pacific Island states (Diamond *et al.*, 2008). Annual TC track maps were acquired from the meteorological services of Fiji, New Caledonia, New Zealand, Tonga, Solomon Islands, and Vanuatu. In addition, a book published in 1925 (Visher, 1925; Figure 4) contained old TC maps with tracks dating back to the 1841 season (1840–1841). This archive survey resulted in obtaining several hundred maps that were rescued and scanned into electronic jpeg-formatted graphics. The scans were then imported into ArcGIS® where georegistration was employed to ascribe, extract, and digitally record the correct latitude and longitude for points along each TC track. The ArcGIS® coding for this georegistration procedure (along with all the scanned images that were digitized) can be found on the resources section of the SPEArTC website at <http://apdrc.soest.hawaii.edu/projects/speartc>.

The assumption was made that the SPEArTC dataset could contain documented tropical storms or deep depressions that were tracked prior to achieving Named Storms (NS) TC status. NSs are defined as having sustained winds that are  $\geq 34$  kts (Diamond, 2010). It is impossible to determine if the final part of a historical storm track is the ETT phase when the storm would have transitioned from a TC to an extratropical storm or tropical depression (Sinclair, 2002). The associated

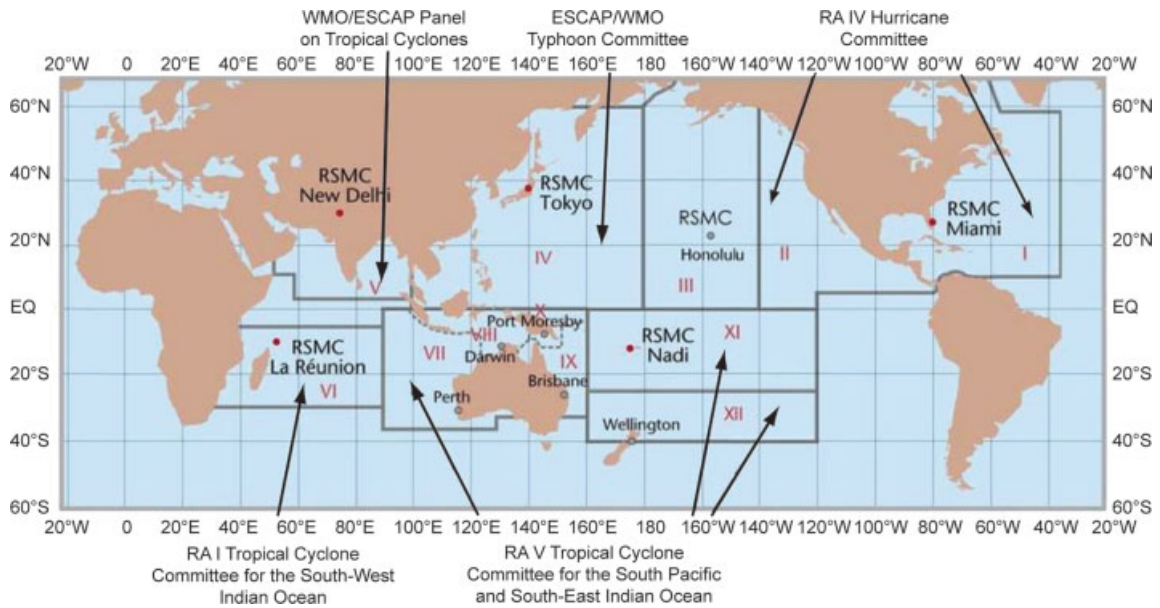


Figure 3. Map of RSMCs and their coverage areas (Source: WMO Tropical Cyclone Programme – <http://www.wmo.int/pages/prog/www/tcp/Advisories-RSMCs.html>).

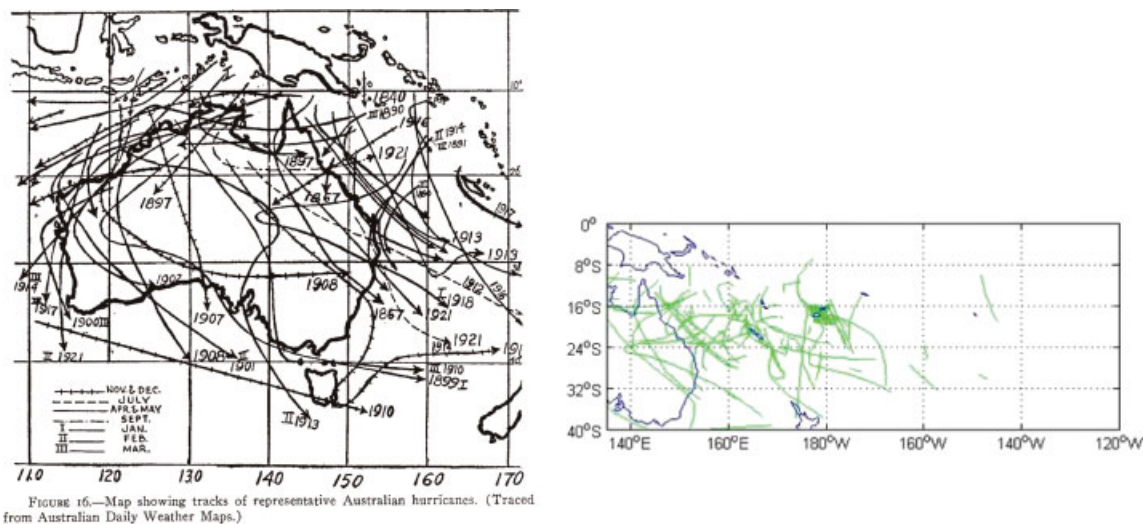


Figure 4. Tracks from Visher, 1925 (left) and from digital database (right).

metadata for storm strength is highly variable throughout the historical record, and the occurrence of this information would depend on the nature of how the storm was tracked and recorded during the pre-satellite era. Even today that procedure differs between RSMCs that track TCs with highly sophisticated remote sensing systems. However, the primary purpose of the SPEArTC dataset is to assess the occurrence and geographic variability of tropical storm characteristics through time in the south-west Pacific. While the magnitude of the TC at any given point along the track is not highly relevant for this exercise, such intensities when available (particularly in the satellite era after 1969) have been incorporated as input to the IBTrACS database and are included in the SPEArTC dataset. Therefore, for each track in the dataset, the storm is assumed to have been tropical in origin, and achieved at least an NS status at some time during its life cycle.

Up until the 1950s or so, the data mostly provided only track information and are missing intensity (e.g. wind and pressure data) that would be critically important to studies related to climate change. However, the existence of past track morphologies does provide important insights into the temporal and spatial aspects of cyclogenesis in the region. Such information is vital for assisting in the development of seasonal outlooks related to the occurrence and frequency of TCs during periods of fluctuating climate conditions primarily related to El Niño and La Niña, and the risks associated with their occurrence.

### 3.2. Data management

All the storm tracks in SPEArTC were arranged in an MS-Excel® spreadsheet to include serial number (IBTrACS identifier (see <http://www.ncdc.noaa.gov/oa/>

ibtracs/index.php?name=numbering), associated TC season (Austral spring to Austral autumn; November–April), and a time and date stamp in six-hourly intervals. The date and time stamp was sometimes applied artificially to tracks recovered from a synoptic chart, and as such can have an error of  $\pm 12$  h. For the purposes of constructing track morphologies this is not considered to be a major problem. The need to artificially time stamp some storms also explains why some replicated track segments appear temporally offset. The temporal uncertainty is due to differences in reporting times used by some of the aforementioned meteorological services.

For ease of file handling, the data were split into three distinct time periods as follows: (1) 1841–1899; (2) 1902–1969; and (3) 1970–2010. These time periods are also significant because the first set of TC seasons are the most poorly observed of the three. While it was beneficial to obtain as much data from that period as possible, we recognize that the observational record in the 19th century will be incomplete. The second period had a considerable volume of data, and overlapping observations from across the region. While these data were considered to be more reliable than the first period, they are not as reliable as the most third and most recent period of time (1970–2010) which is coincident with the beginning of the satellite era of TC monitoring.

A series of formulas in an MS-Excel<sup>®</sup> spreadsheet provided a template to translate track information into Keyhole Markup Language (.KML) file scripts that contained a string of coordinates, dates, and time information for each storm track (a copy of the template can be found on the resources section of the website at <http://apdrc.soest.hawaii.edu/projects/speartc>). The KML script derived from the template was copied and pasted into a file with a .KML file extension, which could then be opened in GoogleEarth<sup>®</sup> to enable application of the GrIT quality control procedure.

### 3.3. Quality controlling the new dataset

Quality control was accomplished using a technique known as GrIT. In brief, GrIT employs a set of objective and subjective criteria to aid quality control by visually inspecting TC track data, and takes advantage of capabilities for displaying multiple TC tracks and animating the tracks using chronology information. Coupled with a knowledge of how TCs behave (in some cases erratically), GrIT affords a rapid way to spot duplicate tracks, identify observational information from different monitoring sources that represent segments of the same storm, replicates of the same storm system, and obvious errors that have been introduced either during digital scanning of paper data or keying processes.

GrIT is a subjective approach that uses human operators and the GoogleEarth<sup>®</sup> interface to assess the number of annual storm tracks that were contained in the unedited SPEArTC dataset. The following criteria and steps were used to verify and consolidate historical storm track data. For non-replicated storms, a single track was selected

from the SPEArTC dataset to represent a TC. Identification of mutually overlapping tracks within a narrow time window (7–10 days) typically indicates duplication of a storm in the dataset. For this situation, when multiple tracks cover all or part of the same storm, the collection of time steps and positions (latitude and longitude) for points along the length of the TC track need to be amalgamated. The first step in amalgamation is to identify the set of tracks that go together (Figure 5 shows the characteristics used to group tracks). The end points of the storm track are then established. Subsequently, replicated track segments and divergent pathways between the start and end point of each storm track are identified. In the case where one part of a storm track diverges from other tracks that are spatially corroborated, the corroborated version is chosen. In the case where the ETT for two tracks in the same storm differ (e.g. the end points do not overlap) the longest storm track is preferentially chosen, as it usually indicates the storm was being monitored by one centre for a longer period of time.

In order to better vet the qualitative aspects of the GrIT method, it was decided to employ three independent analysts to examine each season's set of storm tracks. In this respect, each analyst independently employed the GrIT method for each season with duplicated storm tracks and when completed, an intercomparison of the GrIT results was discussed. In this way, only after consensus was reached, were specific storm tracks within a season selected for deletion and/or modification within the dataset for final incorporation into SPEArTC. During the intercomparison process, the total number of storms for any one season is then checked with the 'baseline' data from IBTrACS to ensure first that the results of the GrIT analysis are within reason, and second, to look at flagging possible additions, modification, or deletions for IBTrACS to consider implementing as well.

Differences in TC track start point, end point, and trajectories in the enhanced dataset arise because the collated information is derived from digitized synoptic charts from multiple sources across the region. These synoptic charts originated from more than one regional meteorological service in the southwest Pacific, and also from redrawn published historical accounts. The independent sets of historical accounts that document past storms, coupled with the fact that synoptic charts were hand drawn, means there is some incorporated subjectivity in the pictorial display of tropical depression centres and TCs, storm tracks, and also when and where storms started and stopped (which depended on how long storms were being monitored).

Erroneous tracks in the dataset were commonly identified when track pathways directly aligned with the International Dateline or with a line of latitude (an obvious digitising error from automated routines). Alternatively, an erroneous track could be attributed to extraneous track segments that were added as a best guess upon ETT, as the result of incomplete synoptic analyses at the time the storm occurred. Erroneous tracks are also compounded



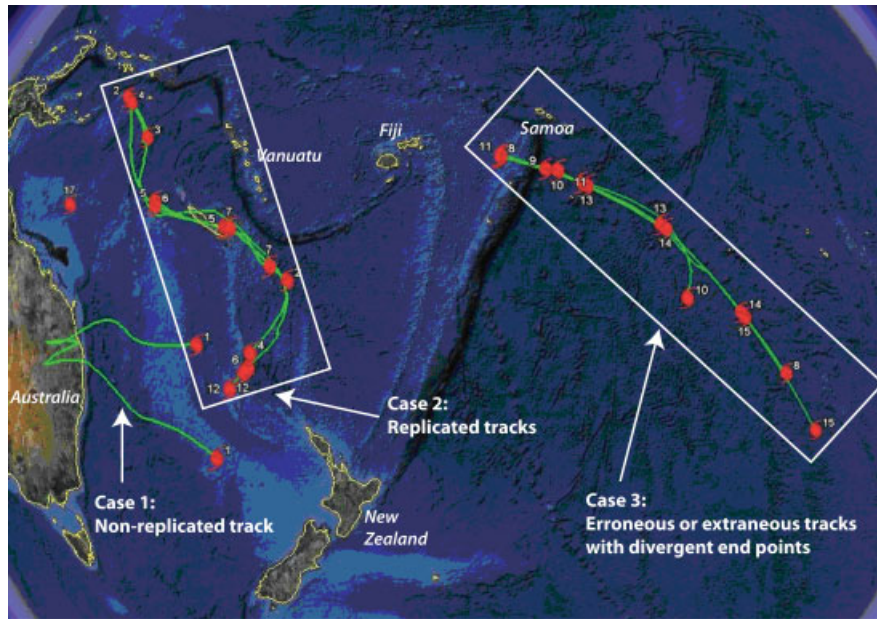


Figure 5. Examples of track types seen in the SPEArTC dataset using an example from the 1955–1956 TC season. The small numbers alongside the red cyclone icons refer to the track number in the enhanced dataset for this particular TC season that were assigned to each track after the IBTrACS and rescued historical information from regional meteorological services were merged. Green lines indicate trajectories that were merged.

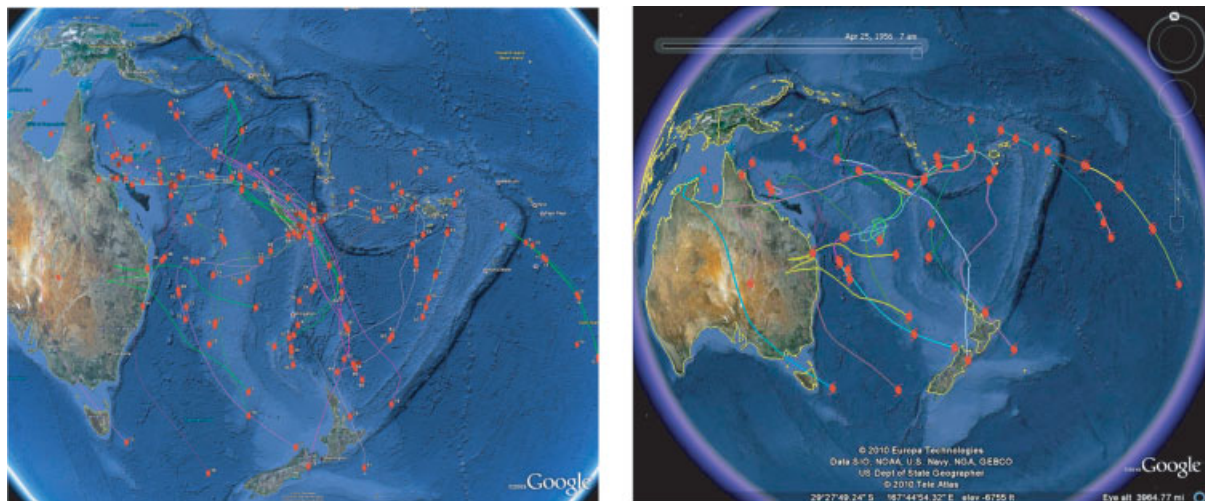


Figure 6. Before (left) and after (right) depictions of TCs in SPEArTC for the 1955–1956 season.

by artifacts of manually tracing tracklines off digital scans of ‘spaghetti’ synoptic charts (seasonal or climatology maps that contain multiple criss-crossing storm tracks) that result in obvious outlying points that are best detected graphically via the GrIT method or by simply looking at the numeric positions in a spreadsheet. Finally, there are other artifacts from digitising or existing data in the IBTrACS dataset that are also best found via the GrIT methodology. Since TC tracks usually have distinctive curves, kinks, and loops that are inherent to the nature of TCs, when one comes across a pure straight-line track across the basin, it at least flags the track as being potentially erroneous and warranting future investigation. It was further found that digitisation errors such as storms

originating in the southern portion of the basin and moving north (something that just does not occur with TCs in the Southern Hemisphere) were easily spotted using GrIT and corrected in the dataset. Figure 6 depicts TCs for the 1955–1956 season before and after the application of GrIT.

A summary of the number of TCs in SPEArTC after the initial compilation and GrIT QC process are noted in Table I below. When the initial SPEArTC database was constructed (Table I), there were a total of 2576 tracks which came from both IBTrACS as well as the hundreds of digitized charts from the various PINMSs across the region, and so there were many duplicate storms to examine. Of the 2576 original number of tracks, 1447 were deleted due to duplication; 148 were concatenated

Table I. Summary of GrIT QC results.

Original # of TCs in SPEArTC	# of TCs Deleted from SPEArTC	# of TCs Concatenated in SPEArTC	# of TCs Corrected in SPEArTC	# of TCs Added in SPEArTC	Original # of TCs in IBTrACS	Final # of TCs in SPEArTC Dataset
2576	1447	148	28	114	1124	1095

Table II. Summary and comparison of TC counts by era.

Era	Total # of storms in IBTrACS	Total # of storms in SPEArTC	Avg # of storms in IBTrACS	Avg # of storms in SPEArTC <sup>a</sup>	Avg # of days per TC in SPEArTC
1841–1899	9	31	0.41	1.41	7.00
1902–1969	549	532	8.19	7.94	7.31
1970–2010	566	532	13.80	12.98	8.20
1981–2010	391	370	13.03	12.33	8.19

<sup>a</sup> The temporal coverage of the era from 1841 to 1899 was very sparse; as such, average calculations were based on a smaller number of years for which data was found (a total of 22 years).

due to being able to piece together tracks from across the region; and 114 new tracks were added resulting in a final total of 1095 tracks in the SPEArTC database from 1840–2010.

For the seasons from 1841 to 2010, there were 55 seasons that had at least one storm deleted from IBTrACS which totalled 138 deleted tracks, while for another 56 seasons, at least 1 storm per season was added to IBTrACS. These changes amounted to the addition of 114 new tracks to the database (18 in the era from 1970 to 2010). Furthermore, another 176 storm tracks were either corrected or were enhanced by concatenating previously identified individual storms into existing storms (49 of these in the era from 1970 to 2010), thus lowering the total number in the database. In addition to better storm numbers for the region, each track was for the first time carefully quality-controlled. The GrIT process has produced a much higher quality best tracks dataset for the southwest Pacific basin. The resources section of the SPEArTC web site noted in Section 6 will, in addition to hosting the final quality controlled SPEArTC database, also have the original pre-quality-controlled SPEArTC dataset available for examination by those who may wish to replicate the quality control work documented here.

#### 4. Implications for Regional TC Climatology

The TC season in the southwest Pacific basin typically extends from November to April (although storms occurring outside that window are not out of the ordinary) and, therefore, crosses over the calendar year. The long-term average number of TCs in the basin (as defined by the WMO) for the 30-year period from 1977 to 2007 is 9 TCs per year, with the peak of the season being from February to March (Sinclair, 2002). However, the greater geographic coverage including in the SPEArTC

dataset which expands the western extent of the basin out to 135 E longitude results in a 30-year average of 12.33 TCs per season (Table II) from 1981 to 2010. The discrepancies in that number result from a combination of (1) the difference between the area of study for SPEArTC and the formal WMO-defined forecast area; (2) the fact that SPEArTC uses named and unnamed storms in its counts; and (3) actual new storms that SPEArTC acquired during the research phase of this paper. A follow-up paper will use the SPEArTC dataset to better characterize the TC climatology for the region will delve deeper into this aspect of TCs in the region for the period from 1981 to 2010.

As noted in Section 3.2, the significance of the four eras noted in Table II relate to the confidence level of the data in the dataset. The period from 1841 to 1899 is an era with the least amount of confidence in the data as there was no real systematic manner to collect track data during that period. The period from 1902 to 1969 is more contiguous in nature and during that time the tracking of TCs was better documented. The period from 1970 to 2010 is the era with the highest confidence due to advent of satellite monitoring for TCs. For climatological purposes, a standard 30-year period from 1981 to 2010 was chosen from which to set baselines as to TC track and intensity; this is also a period with the best available data for the entire dataset. While it is difficult to draw climatological conclusions from track data prior to the satellite era beginning with the 1970 season, the improved track data prior to that time, resulting from this study, still has considerable value in being able to contribute to studies on the impact of individual tropical cyclones on other areas of study including coral reefs, coastal erosion, deposition, slope stability, hydrology and fluvial geomorphology.

An analysis of storm numbers for the first and second halves of the TC season (Nov–Jan and Feb–Apr),

was also conducted to better inform the generation of seasonal outlooks and Table III summarizes seasonal activity based on which portion of the TC season is generally busiest. A detailed listing of the number of NSs by season from 1970–2010 can be found in the Table AI. The results of this analysis indicate that the latter half of the TC season for the southwest Pacific is nearly 50% more active than the first half for the modern era. A monthly analysis of the data from 1970 to 2010 as depicted in Table IV clearly shows that the height of the TC season in the basin is from January through March, which is a period in which over 2/3 of the storms occur.

**5. Availability**

With the construction of the SPEArTC dataset, further work has been initiated to relate the enhanced information there to formulate a more complete climatological study of TCs in this region of the world. Further study of TCs in the southwest Pacific using the SPEArTC dataset will aid in furthering knowledge of TC trends in the southern hemisphere, and will complement previous studies that have done so for the south Indian Ocean and Australian regions (30 E to 135 E) and the south Pacific Ocean (135 E to 120 W) (Kuleshov *et al.*, 2010) by extending an enhanced climatology of TCs across the entire southern hemisphere basin. Additionally, all SPEArTC TC track adjustment will be incorporated into future releases of IBTrACS. Finally, the SPEArTC dataset is available upon request by registering for it at the resources section of the website

Table III. Summary of NS occurrences within each season from 1970 to 2010.

Total # of storms	Total # of NSs	# of NSs (Nov–Jan)	# of NSs (Feb–Apr)
532	411	166	245

Table IV. TC activity by month from 1970 to 2010.

Month	Total	Average	Monthly (%)
November	21	0.51	3.95
December	65	1.59	12.21
January <sup>a</sup>	114	2.78	21.43
February <sup>a</sup>	130	3.17	24.44
March <sup>a</sup>	121	2.95	22.74
April	54	1.32	10.15
May	15	0.37	2.82
June	5	0.12	0.94
July	1	0.02	0.19
August	1	0.02	0.19
September	0	0.00	0.00
October	5	0.12	0.94
Annual	532	12.98	100.00

<sup>a</sup> The peak months for seasonal TC activity are from January to March.

at <http://apdrc.soest.hawaii.edu/projects/speartc>, and the documentation is posted there as well as in Table AIII (along with a sample listing of data). For those interesting more about IBTrACS, that information is available at the following site at <http://www.ncdc.noaa.gov/oa/ibtracs/>.

**6. Concluding Remarks**

The results of this systematic analysis of tropical cyclone data for the region resulted in a number of new tracks being added to several seasons, a larger number of

Table AI. Detailed listing of NSs by season portion from 1970 to 2010.

Season	Total # storms	Total NSs	Named from Nov–Jan	Named from Feb–Apr
1970	14	6	1	5
1971	11	8	3	5
1972	19	14	6	8
1973	17	12	6	6
1974	18	12	7	5
1975	12	6	3	3
1976	14	13	4	9
1977	17	14	6	8
1978	13	10	5	5
1979	15	10	4	6
1980	12	10	3	7
1981	16	11	3	8
1982	11	10	5	5
1983	17	14	6	8
1984	13	10	3	7
1985	17	12	6	6
1986	16	10	2	8
1987	15	13	6	7
1988	8	6	2	4
1989	15	13	4	9
1990	8	8	2	6
1991	9	5	2	3
1992	13	11	4	7
1993	16	12	4	8
1994	9	6	3	3
1995	8	5	2	3
1996	8	8	3	5
1997	18	16	7	9
1998	24	19	11	8
1999	14	10	5	5
2000	10	9	2	7
2001	11	9	2	7
2002	12	7	4	3
2003	13	11	5	6
2004	8	7	2	5
2005	13	11	3	8
2006	8	8	3	5
2007	11	9	4	5
2008	9	6	6	0
2009	10	10	3	7
2010	10	10	4	6
Totals	532	411	166	245
Averages	12.98	10.02	4.05	5.98



duplicate or erroneous tracks deleted, and most importantly, a comprehensive repair to the overall dataset that will at least help to normalize the number of TCs that are studied for climatological purposes. Given the very time consuming and detailed QC of the TCs in this region, the authors have a high confidence level in the resultant dataset, and suggest that the methods in this paper be considered for application to other basin data in the IBTrACS dataset. That coupled with a continued search for new data sources should only set to improve the overall quality of global TC data.

The resultant new quality-controlled SPEArTC dataset should also have positive impacts on TC and climatology studies in the region. For example, a paper currently in preparation (Lorrey *et al.*, 2011 (in preparation)) is examining precipitation patterns for the Pacific Islands region during an extreme La Niña (1955–1956). Information about past TC tracks is required in order to put regional rainfall anomalies at individual sites into context over time, and this will contribute to several case

studies encompassed in the South Pacific Rainfall Atlas (SPRAT).

This study builds on prior work done in documenting TCs in the region (Lourenz, 1981); the New Zealand Meteorological Service (Kerr, 1976; Revell, 1981; and Thompson *et al.*, 1992); Meteo France (Giovannelli, 1952); and the WMO (Maunder, 1995). However, up until this study, there does not appear to have been a consolidated TC database that (1) tied these various sources together; and (2) digitized this information along with post-1969 satellite data allowing for further statistical analysis. This paper accomplished those two goals, and now further work can be undertaken to better relate various TC elements together over time to relationships with other phenomena such as the ENSO, including the oceanic sea surface temperature and atmospheric Southern Oscillation Index components of ENSO, as well as the coupled nature of these phenomena. Finally, the SPEArTC provides the baseline information necessary to apply that information towards a seasonal outlook tool

Table AII. Documentation of GRIT corrections made from 1970 to 2010.

Season	Storm ID	Storm name	Correction	Comments
1998	1997319S09180	Nute	Extraneous points deleted	25 positions deleted that made storm loop abruptly to the north erroneously
1985	1985061S13170	Gavin	Extraneous points deleted	42 positions deleted that made storm loop abruptly to the north erroneously
1984	1983343S14148	Fritz	Extraneous points deleted	8 positions deleted that made storm move abruptly to the east in a 36-h period
1983	1983042S11158	Elinor	Extraneous points deleted	25 positions deleted that made storm abruptly loop to the north and east in a four-day period.
1983	1983056S11193	Prema	Extraneous points deleted	46 positions deleted that made storm abruptly loop to the north and repeat its track.
1983	1983097S10222	Veena	Extraneous points deleted	57 positions deleted that made storm abruptly loop to the north and begin repeating its track.
1980	1980001S13173	Paul	Extraneous points deleted	26 positions deleted that made storm abruptly loop abnormally to the west.
1979	1979093S08156	Stan	Extraneous points deleted	20 positions deleted that made storm abruptly loop abnormally to the east.
1977	1977034S17175	Miles	Extraneous points deleted	14 positions deleted that made storm abruptly loop abnormally to the north and begin repeating its track.
1974	1974017S20150 and 1974021S17148	Vera and Wanda	Storm tracks properly identified	The tracks for Vera and Wanda were confused and via the GrIT process were able to be properly sorted and separated into two distinct storm tracks both just off the eastern coast of Australia in January 1974.
1974	1973029S16185	Glenda	Duplicate track	A duplicate track of Glenda was identified and removed from the database.
1974	1973052S13151	Kirsty	Erroneous points deleted and smoothed with an orphan track	8 positions deleted that made storm abruptly loop abnormally to the north and begin repeating its track; and in addition, an extraneous track identified as a separate storm was merged with Kirsty, thus cleaning up the database.
1972	1972001S12159 and 1972005S10152	Carlotta	Duplicate storms	Two separate storms both noted as Carlotta turned out to be the same storm and were merged as one storm in the database

that can be used to aid the nations of the region in both disaster preparation and mitigation. Not only will the work of SPEArTC (Section 4) be incorporated into the global IBTrACS dataset, but it will stand on its own for those who wish to use it as a resource for TC studies in the region.

### Acknowledgements

The authors would like to recognize Professor Scott Shipley from George Mason University in Fairfax, Virginia, USA, in association with Mr Ernest Daddio from Riverside Technology, Inc. Dr Shipley is a leading GIS expert who assisted in the development of the necessary ArcGIS<sup>®</sup> techniques (and coding) necessary to digitize all the paper TC tracking maps that formed the

basis for this paper. Dr Lorrey's work on this paper was partially supported by the Foundation for Research Science and Technology contract CO1X0701, 'Adaptation to Climate Variability and Change'. In addition, we would like to thank Mr Craig Stanton at NIWA Auckland for his assistance with formatting data into GoogleEarth<sup>®</sup> for input to the GrIT process; Ms Petra Chappel at Auckland University for her assistance in analysing data via the GrIT process; and Mr Michael Klatt at the University of Oklahoma who assisted in the initial database formatting of the digitized tracks that were used as input into the GrIT process. We are also thankful to Dr Jim Potemra and Ms Sharon DeCarlo of the Asia Pacific Data Research Center at the University of Hawaii for assisting in hosting the research web page noted in the paper. Finally, we are indebted to and

Table AIII. SPEArTC CSV file data documentation.

Column	Further Details <sup>a</sup>
A	Unique Storm Serial Number
B	Season; e.g., 1956 equates to 1955–56 (November 1955–April 1956)
C	Sequential Storm Number within the Season
D	IBTrACS Basin Code – SP (South Pacific); SI (South Indian)
E	IBTrACS Sub-Basin Code – EA (E. Australia); WA (W. Australia)
F	Storm Name (if named) or Source Filename
G	Synoptic Date and Time of the Storm's Position (MM/DD/YYYY hh:mm:ss) <sup>b</sup>
H	NR – Not Rated; TS – Tropical Storm; ET – Extratropical
I	Latitude in tenths of a degree ('-' indicates degrees South)
J	Longitude in tenths of a degree ('-' indicates degrees West)
K	Mean Wind Speed in Knots <sup>c</sup> (–999 indicates missing)
L	Mean Central Pressure in hectopascals or millibars <sup>d</sup> (–999 is missing)
M	Tracking Centre or Source <sup>e</sup>

SPEArTC data sample

A	B	C	D	E	F	G	H	I	J	K	L	M
2009025S12177	2009	8	SP	MM	HETTIE	1/28/2009 12:00	NR	–21.7	–177.4	35	995	nadi
2009025S12177	2009	8	SP	MM	HETTIE	1/28/2009 18:00	NR	–22.1	–177.5	35	995	nadi
2009025S12177	2009	8	SP	MM	HETTIE	1/29/2009 0:00	NR	–22.6	–177.7	35	995	nadi
2009025S12177	2009	8	SP	MM	HETTIE	1/29/2009 6:00	NR	–23.2	–177.8	30	997	nadi
2009025S12177	2009	8	SP	MM	HETTIE	1/29/2009 12:00	NR	–23.5	–177.9	25	1000	nadi
2009025S12177	2009	8	SP	MM	HETTIE	1/29/2009 18:00	NR	–24.1	–177.9	25	1000	nadi
2009011S16139	2009	9	SP	EA	CHARLOTTE	1/10/2009 18:00	NR	–15.4	138.8	0	995	bom
2009011S16139	2009	9	SP	EA	CHARLOTTE	1/10/2009 21:00	NR	–15.7	139.1	25	994	bom
2009011S16139	2009	9	SP	EA	CHARLOTTE	1/11/2009 0:00	NR	–15.9	139.4	35	993	bom
2009011S16139	2009	9	SP	EA	CHARLOTTE	1/11/2009 3:00	TS	–16	139.7	40	990	bom
2009011S16139	2009	9	SP	EA	CHARLOTTE	1/11/2009 6:00	TS	–16.1	139.9	40	990	bom

<sup>a</sup> MM indicates missing data.

<sup>b</sup> From 1840 to 1899 the format of this field is YYYY-MM-DD hh:mm:ss.

<sup>c</sup> Through the 2008 season, this is the IBTrACS averaged value based on IBTrACS Format v02r01; beginning in 2009 this is actual value from IBTrACS v03r02.

<sup>d</sup> Through the 2008 season, this is the IBTrACS averaged value based on IBTrACS Format v02r01; beginning in 2009 this is actual value from IBTrACS v03r02.

<sup>e</sup> This is either a Forecast Centre (e.g., Australian BoM; Wellington, Nadi; or Joint Typhoon Warning Center); or a supplementary data source (Neumann; NCDC TD9636; or SPEArTC data rescue as described in this paper).

would like to thank the directors and staff of the regional Meteorological Services in Fiji, French Polynesia, New Caledonia, New Zealand, Tonga, Solomon Islands, and Vanuatu for assisting us by sharing their paper-based TC maps that were critical to ensuring that the SPEArTC database had as much source information as possible in its construction.

## References

- Diamond HJ, (ed). 2010. The tropics [in "State of the Climate in 2009"]. *Bulletin of the American Meteorological Society* **91**(7): 579–106.
- Diamond HJ, Jones D, Kuleshov Y. 2008. A bilateral project between the U.S. and Australia to provide a consolidated high quality database of Southern Hemisphere tropical cyclones. *Eos Transactions AGU* **89**(23): Western Pacific Geophysics Meeting Supplement, Abstract U33A-06.
- Giovannelli PJ. 1952. *Les Cyclones Tropicaux En Nouvelle-Caledonia, Au Cours D'un Siecle (1852–1951)*, Service De La Meteorologie, Noumea, New Caledonia.
- Gray WM. 1979. Hurricanes: Their formation, structure and likely role in the tropical circulation. In *Meteorology Over Tropical Oceans*, Shaw DB, (ed). Royal Meteorological Society: Bracknell, UK, 155–218.
- Holland GJ. 1981. On the quality of the Australian tropical cyclone data base. *Australian Meteorological Magazine* **29**: 169–181.
- Kerr IS. 1976. Tropical Storms and Hurricanes in the Southwest Pacific, November 1939 to April 1969. NZ Ministry of Transport, Wellington, New Zealand.
- Knapp KR, Kruk MC, Levinson DH, Diamond HJ, Neumann CJ. 2010. The International Best Track Archive for Climate Stewardship (IBTrACS). *Bulletin of the American Meteorological Society* **91**: 363–376, DOI: 10.1175/2009BAMS2755.1.
- Kuleshov Y, Fawcett R, Qi L, Trewin B, Jones D, McBride J, Ramsay H. 2010. Trends in tropical cyclones in the South Indian Ocean and the South Pacific Ocean *Journal of Geophysics Research* **115**: D01101, DOI: 10.1029/2009JD012372.
- Kuleshov Y, Qi L, Fawcett R, Jones D. 2008. On tropical cyclone activity in the Southern Hemisphere: Trends and the ENSO connection. *Geophysical Research Letters* **35**: L14S08, DOI: 10.1029/2007GL032983.
- Leroy A, Wheeler MC. 2008. Statistical prediction of weekly tropical cyclone activity in the Southern Hemisphere. *Monthly Weather Review* **136**: 3637–3654.
- Lourenz RS. 1981. Tropical Cyclones in the Australian Region, July 1909 to June 1980. Bureau of Meteorology, October 1981 Melbourne: Australia.
- Maunder J. 1995. An Historical Overview Regarding the Intensity, Tracks, and Frequency of Tropical Cyclones in the South Pacific During the Last 100 Years, and an Analysis of any Changes in These Factors. WMO/TD-No. 692. Tropical Cyclone Programme, Report TCP-37. World Meteorological Organization, Geneva, Switzerland.
- Mimura N, Nurse L, McLean RF, Agard J, Briguglio L, Lefale P, Payet R, Sem G. 2007. *Small islands. Climate Change 2007: Impacts, Adaptation and Vulnerability*. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Parry ML, Canziani OF, Palutikof JP, van der Linden PJ, Hanson CE (eds). Cambridge University Press: Cambridge, UK, 687–716.
- Parry ML, Canziani OF, Palutikof JP, van der Linden PJ, Hanson CE (eds), *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007*. Cambridge, United Kingdom and New York, NY, USA.
- Ready S. 2011. A personal communication with Mr. Ready from the New Zealand MetService, Ltd, regarding *TC data monitoring and best tracks data at RSMC Wellington*.
- Revell CG. 1981. *Tropical cyclones in the southwest Pacific, November 1969 to April 1979*. NZ Ministry of Transport: Wellington, New Zealand.
- Sinclair MR. 2002. Extratropical transition of Southwest Pacific tropical cyclones. Part I: Climatology and mean structure changes. *Monthly Weather Review* **130**: 590–609.
- Terry JP, Etienne S. 2010. Tempestuous Times in the South Pacific Islands. *Science* **23**: 428–429, DOI: 10.1126/science.328.5977.428.
- Terry JP, Gienko G. 2010. Climatological aspects of South Pacific tropical cyclones, based on analysis of the RSMC-Nadi (Fiji) regional archive. *Climate Research* **42**: 223–233.
- Thompson CS, Ready S, Zheng X. 1992. *Tropical cyclones in the Southwest Pacific: November 1979 to May 1989*. New Zealand Meteorological Service: Wellington, New Zealand.
- Trewin BC. 2008. An Enhanced Tropical Cyclone Data Set for the Australian Region: Tropical Meteorology Special Symposium; 20th Conference on Climate Variability and Change; 88th Annual Meeting of the American Meteorological Society, New Orleans, 20–24 January 2008.
- Visher S. 1925. *Tropical Cyclones of the Pacific*. Bulletin 20, Honolulu, Bernice P. Bishop Museum, 163 p.
- World Bank. 2006. *Not if, but when. Adapting to natural hazards in the Pacific Islands region; A policy note*, The World Bank, East Asia and the Pacific Region, Pacific Islands Country Management Unit: Washington, DC, 46 pp.