Weather and Climate Extremes 34 (2021) 100405

Contents lists available at ScienceDirect

Weather and Climate Extremes

journal homepage: www.elsevier.com/locate/wace

A decision tree approach to identify predictors of extreme rainfall events – A case study for the Fiji Islands

Krishneel K. Sharma [a,*], Danielle C. Verdon-Kidd [a], Andrew D. Magee [b]

[a] School of Environmental and Life Sciences, College of Engineering, Science and Environment, The University of Newcastle, Callaghan, NSW, 2308, Australia
[b] Centre for Water, Climate and Land (CWCL), College of Engineering, Science and Environment, The University of Newcastle, Callaghan, NSW, 2308, Australia

ARTICLE INFO

Keywords:
Extreme rainfall
Fiji Islands
Tropical cyclone tracks
Classification modelling
Decision tree

ABSTRACT

Extreme rainfall events often lead to excessive river flows and severe flooding for Pacific Island nations. Fiji, in particular, is often exposed to extreme rainfall events and associated flooding, with significant impacts on properties, infrastructure, agriculture, and the tourism sector. While these occurrences are often associated with tropical cyclones (TCs), the specific characteristics of TCs that produce extreme rainfall are not well understood. In particular, TC intensity does not appear to be a useful guide in predicting rainfall, since weaker TCs are capable of producing large rainfall compared to more intense systems. Therefore, other TC characteristics, in particular TC track morphology and background climate conditions, may provide more useful insights into what drives TC related extreme rainfall. This study aimed to address this problem by developing a decision tree to identify the most important predictors of TC related extreme rainfall (*i.e.,* 95th percentile) for Fiji. TC attributes considered include; TC duration, the average moving speed of TCs, the minimum distance of TCs from land, seasonality, intensity (wind speed) and the geometry of TCs (*i.e.,* geographical location, shape and length via cluster and sinuosity analyses of TC tracks). In addition, potential predictors based on the phases of Indo-Pacific climate modes were input to the decision tree to represent large scale background conditions. It was found that a TC's minimum distance from land was the most important influence on extreme rainfall, followed by TC cluster grouping, seasonality and duration. The application of this model could result in improved TC risk evaluations and could be used by forecasters and decision-makers on mitigating TC impacts over the Fiji Islands.

1. Introduction

The Fiji group (Fig. 1) comprises more than 300 islands and is centrally located over the International Date Line (180°) within the south-west Pacific (SWP) region (Yeo and Blong, 2010; Kumar et al., 2014). The weather pattern in Fiji is mostly described as a tropical marine (only slight seasonal temperature variations), where the climate features are influenced by South Pacific Convergence Zone (SPCZ), trade winds, sub-tropical highs, tropical cyclones (TCs) and land topography (CSIRO Australian Bureau of Meteorology and SPREP, 2015). Fiji is susceptible to regular flooding events (McKenzie et al., 2005) associated with orographic rainfall, which is attributed to the topography of its larger islands (Viti Levu and Vanua Levu), with a maximum elevation of up to 1300 m above sea level, as well as prevailing southeast trade winds (Yeo, 2013; Kuleshov et al., 2014; Kumar et al., 2014). These features result in a wet zone on Fiji's eastern side and a drier leeward zone on the western side (Terry, 2007).

During the wet season (*i.e.,* TC season between November and April), extreme or heavy rainfall events often lead to excessive river flows and severe flooding. Fiji has experienced extreme rainfall and associated flooding with significant damage to properties, public infrastructure, agriculture and the tourism sector (Yeo, 2013; Kuleshov et al., 2014). In particular, these events affect the fertile land along rivers, which impacts the agricultural sector, a source of income for Fijian farmers. Extreme rainfall events in this region are often associated with TCs (Terry, 2007; McGree et al., 2010; Kuleshov et al., 2014; Deo and Walsh, 2017), with four per year on average passing through Fiji's Exclusive Economic Zone (EEZ; Flanders Marine Institute, 2018; Sharma et al., 2021). Some examples of TCs that produced extreme rainfall for Fiji includes; TC Ami (January 2003; Category 3) with a peak 24-hour rainfall of 311 mm, TC Winston (February 2016; Category 5), with a peak 24-hour rainfall of 358 mm and TC June (May 1997; Category 2), which delivered 341 mm of rainfall within a 24-hour period (Terry and Raj, 1999; Terry et al., 2004; Fiji Meteorological Service, 2017).

* Corresponding author. GG03, Earth Sciences Building, The University of Newcastle, University Drive, Callaghan, NSW, 2308, Australia.
E-mail address: Krishneel.Sharma@uon.edu.au (K.K. Sharma).

https://doi.org/10.1016/j.wace.2021.100405
Received 1 February 2021; Received in revised form 10 November 2021; Accepted 2 December 2021
Available online 6 December 2021

Importantly, these three examples highlight that TC category is not necessarily the best indicator of extreme rainfall since weaker (non-severe) systems such as TC June (Category 2) recorded comparable 24-hour rainfall than the more intense Category 5 TC Winston. This indicates that the magnitude of a TC's impacts (in terms of extreme rainfall and associated flooding) are potentially associated with other TC characteristics and/or background climatic conditions. This hypothesis is supported by Terry (2007), who noted that the concept of TC intensity as a useful guide for expected rainfall is often misunderstood since weaker systems are also capable of producing large rainfall compared to intense systems. Indeed, intense TC systems are more often associated with wind-driven impacts, rather than rainfall.

An evaluation of typical TC track characteristics (in terms of shape, location and sinuosity) for the SWP region was recently carried out by Sharma et al. (2021) (hereafter, referred to as SMV21). The authors grouped historical TC tracks from 1948 to 2017 into five distinct clusters, with Fiji being largely impacted by TCs from Clusters 1, 2 and 4 (Fig. 2). Further, the authors grouped historical TCs into four different sinuosity categories from straight through to highly sinuous. However, it is not clear if/how these differing TC characteristics (or others such as distance to land, TC duration, *etc.*) may subsequently impact the likelihood of TC induced extreme rainfall. Research has also shown that Indo-Pacific climate modes (*e.g.,* El Niño-Southern Oscillation (ENSO), ENSO Modoki, Madden-Julian Oscillation (MJO), Indian Ocean Dipole (IOD) and Southern Annular Mode (SAM)) play an important role in modulating TC genesis and tracks in the SWP region and its sub-regions (Chand and Walsh, 2009, 2010; Diamond and Renwick, 2015; Magee et al., 2017; Magee and Verdon-Kidd, 2018; SMV21). The findings of these studies provide implications for understanding and improving seasonal TC outlooks and short-term projections of TC pathways once a TC is established. However, the impact of Indo-Pacific climate modes on TC induced extreme rainfall has yet to be investigated. This paper aims to identify how/what characteristics of TCs (including the background climate conditions in which they occur) contribute to extreme rainfall in Fiji and how this information can be modelled to improve TC outlooks. The objectives of this paper are:

1. Quantify the contribution of TCs to extreme rainfall events across Fiji.
2. Explore the relationship between TC characteristics, Indo-Pacific climate modes and the incidence of extreme rainfall events in Fiji to identify potential predictors.
3. Rank the importance of each predictor variable via classification modelling (*i.e.,* a decision tree).

2. Data

2.1. TC track data

This study uses the South Pacific Enhanced Archive of Tropical Cyclones (SPEArTC) database for best track TC data that are mostly available at 6-hour intervals for the entire TC lifecycle (Diamond et al., 2012). The SPEArTC database is built using several existing TC databases including the International Best Tracks for Climate Stewardship (IBTrACS; Knapp et al., 2010), Joint Typhoon Warning Centre (JTWC; Chu et al., 2002), Southern Hemisphere Tropical Cyclone Data Portal (Australian Bureau of Meteorology, 2018), Regional Specialised Meteorological Centre Nadi, Tropical Cyclone Warning Centre Wellington, TC tracking maps from Fiji, New Caledonia, New Zealand, Tonga, Solomon Islands and Vanuatu, and old historical records such as Visher (1922) (Diamond et al., 2012; Magee et al., 2016).

The first point of each unique track in the SPEArTC dataset is considered as the TC genesis point (TC formation), and the last point as the TC decay point (*i.e.,* where TC ended/terminated). This dataset was chosen for two reasons. Firstly, it is considered as the most complete repository of TC data for the SWP region (Magee et al., 2016; Sharma et al., 2020). Secondly, to maintain consistency with SMV21 where the same dataset was used to evaluate and characterise SWP TC tracks (*see* Fig. 2). As such, TC tracks between 1948 and 2017 that have been allocated to each of the five clusters and sinuosity index (SI) groups (*in* Figure 11 of SMV21) are utilised here.

In brief, these clusters consist of SWP TC tracks, separated and grouped via a probabilistic curve clustering technique (Gaffney, 2004; Gaffney et al., 2007) into distinct groups, based on their geographical

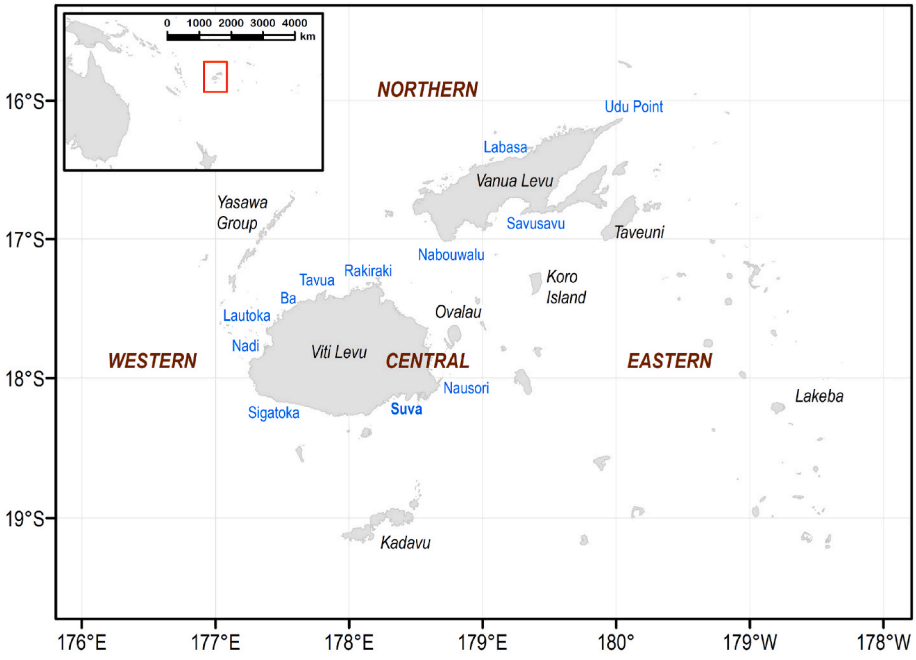

Fig. 1. Location map of the Fiji Islands within the SWP region (indicated in the inset map by a red box). Major towns/centres are listed in blue font (capital city in bold), island names in black italicised font and the four divisions in brown uppercase. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

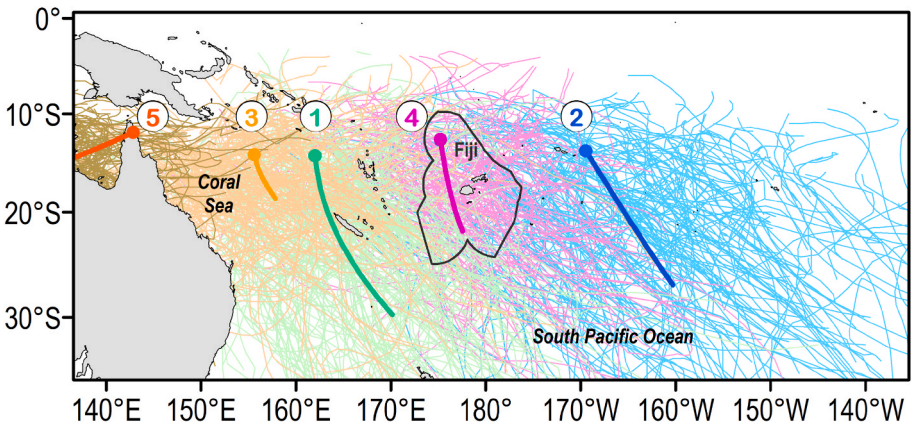

Fig. 2. TC tracks for all five clusters (characterised by their geographical location and shape) and their numbered mean regression curves (MRCs) from 1948 to 2017 adapted from SMV21. Also shown is the EEZ of the Fiji Islands.

location, shape and length. SMV21 also classified the SWP TC tracks into four SI categories: straight, quasi-straight, curving, and sinuous. Collectively, SMV21 reported continued dominance of straight TCs within the eastern SWP, while the western SWP region is typically exposed to highly sinuous tracks. Further, a baseline climatology of SWP island nations and territories, particularly of the Fiji Islands was developed, *i.e.,* the number of TCs that traversed Fiji's EEZ (Flanders Marine Institute, 2018). It was demonstrated that the Fiji Islands was most vulnerable to Cluster 4 tracks (*see* Fig. 2) that largely comprise of intricate TC tracks (established in Figure 11 of SMV21); however, were also occasionally impacted by Cluster 1 and 2 tracks. SMV21 highlighted that intricate (or highly sinuous) TC tracks could be more impactful given the risk of multiple landfalls due to looping and backtracking. As such, the analyses of this paper builds on these initial findings.

The SPEArTC database was filtered to identify only those TCs that traversed the Fiji EEZ between November and April (Fiji's TC season). Note that if the duration of a TC within the Fiji EEZ was very brief (*e.g.,* < 6 hours), or if a 6-hourly observation was not recorded, the event was not included in this analysis. The corresponding wind speed data for each TC from the SPEArTC database, which is the maximum 10-min sustained winds in knots, was also used in this analysis following the conversion 1 knot = 1.85 km/h.

2.2. Rainfall data

The rainfall data used in this study were obtained directly from the Fiji Meteorological Service (FMS). Daily rainfall was available from 1960-onwards for selected stations in Fiji (*see* Table 1; Fig. 1).

2.3. Climate indices

The following climate indices are used in this analysis:

- ENSO index: monthly Oceanic Niño Index (ONI) utilised in this study was obtained from the National Oceanic and Atmospheric Administration, Climate Prediction Centre (Kousky and Higgins, 2007). ONI is based on the 3-month running mean of Extended Reconstructed Sea Surface Temperature Version 5 (ERSSTv5) (Huang et al., 2017a, 2017b) sea surface temperature (SST) anomalies in the Niño3.4 region (5°N–5°S, 120°–170°W). The anomaly calculation of ONI is based on the multiple centred 30-year base period, which is updated every five years. This approach ensures that the classification of El Niño (EN) and La Niña (LN) events are defined by their contemporary climatology and remain fixed over most of the historical period. A threshold of ±0.5 °C is used to define the EN and LN phases when the periods of below-/above-average SSTs occur for at least five consecutive overlapping seasons, while SSTs in-between are classified as ENSO neutral events.
- ENSO Modoki Index (EMI): differences in the monthly SST anomalies (using ERSSTv5) between Modoki A (10°N–10°S, 165°E–140°W), Modoki B (5°N–15°S, 110°W–70°W) and Modoki C (20°N–10°S, 125°E–145°E) regions and calculated using equation (1) after Ashok et al. (2007):

$$\mathrm{EMI} = (\mathrm{Modoki\ A}) - 0.5 \times (\mathrm{Modoki\ B}) - 0.5 \times (\mathrm{Modoki\ C}) \quad (1)$$

  The three distinct phases of EMI are the El Niño Modoki (ENM), La Niña Modoki (LNM) and Modoki Neutral.
- IOD E (eastern Indian Ocean dipole): monthly ERSSTv5 SST anomalies within the south-eastern Indian Ocean (0°–10°S, 90°E–110°E), whereby positive (negative) IOD E phase indicates warmer (cooler) SSTs (Saji et al., 1999).

Table 1
List of stations from the FMS database with their respective locations in Fiji (*see* Fig. 1). Also shown is a summary of the raw daily rainfall data for all months from 1st January 1960 to 30th April 2017. See the table footnote for actual names of some stations as per FMS dataset.

Station	Location Lat / Lon	% of missing data
Nadi[a]	-17.75°S / 177.45°E	0.01%
Suva[b]	-18.15°S / 178.45°E	0.02%
Labasa[c]	-16.47°S / 179.33°E	2.38%
Lakeba	-18.23°S / 178.80°W	1.66%
Rakiraki[d]	-17.37°S / 178.17°E	0.88%
Nabouwalu	-16.98°S / 178.70°E	1.52%
Udu Point	-16.43°S / 179.38°E	5.49%

[a] Nadi Airport.
[b] Laucala Bay.
[c] Labasa Airfield.
[d] Penang Mill.

- Indonesian Index (II): monthly SST anomalies (using ERSSTv5) within the Indonesian region (0°–10°S, 120°E–130°E), whereby positive (negative) II phase indicates warmer (cooler) SSTs (Nicholls, 1989; Verdon and Franks, 2005).
- SAM: the station-based SAM index (used by the British Antarctic Survey which uses an empirical definition of SAM) is the monthly differences in the mean sea level pressure observed from six *in-situ* stations close to 40°S and 65°S (Marshall, 2003; Ho et al., 2012; Diamond and Renwick, 2015; Fogt and Marshall, 2020). Based on the recommendations of Ho et al. (2012) and its application on SWP TCs (Diamond and Renwick, 2015), the Marshall Index is selected, which is available from 1957. The Marshall Index's zonal mean was standardised using the 1981–2010 climatological baseline. SAM is calculated using the following equation after Gong and Wang (1999):

$$\mathrm{SAM} = \mathrm{P}^{*}_{40^{\circ}\mathrm{S}} - \mathrm{P}^{*}_{65^{\circ}\mathrm{S}} \quad (2)$$

All monthly anomaly values for all climate indices were calculated using the 1981–2010 climatological baseline (World Meteorological Organization, 2017). Positive, negative and neutral phases of each climate mode were defined using a threshold of ±0.5 standard deviations from the mean. TC tracks were then stratified (grouped) according to positive, negative and neutral phases of the monthly index values of each climate mode.

3. Methods

3.1. Homogeneity analysis of rainfall datasets

An initial quality test of the rainfall data was carried out by evaluating the statistics derived from the raw rainfall data obtained from the FMS (Table 1). This homogeneity test included an examination of the proportion of missing data (*i.e.,* examining the number of expected and missing data from each station), based on the time period. The RStudio software package, RHtests_dlyPrcp, was employed to test the dataset's completeness (Wang and Feng, 2013). RHtests_dlyPrcp software is designed explicitly for homogenisation of daily precipitation data, which are non-negative, non-continuous and non-normally distributed (Wang et al., 2010). The package uses a number of statistical tests such as transPMFred algorithm (Wang et al., 2010), which integrates a data adaptive Box-Cox transformation procedure into the PMFred algorithm (Wang, 2008a) to make the data approximate to a normal distribution. The PMFred algorithm is based on the penalised maximal *F* test (Wang, 2008b) that is embedded in a recursive testing algorithm (Wang, 2008a). With the benefit of using without a reference series, this package is used to detect changepoints or jumps in the dataset (Wang, 2008a, 2008b; Wang et al., 2010). Daily rainfall data for each station (indicated in Table 1) were tested to detect any changepoints. Consequently, another quality test was carried out to examine the available and missing rainfall data within each time period identified from (if any) changepoints until 2017 to warrant if the rainfall data within a station were reliable for further analysis. A consistent time period across all stations was then chosen for further analysis.

3.2. Identifying extreme rainfall events and related TCs

As per the World Meteorological Organisation (WMO), extreme rainfall events can be considered in two ways: (i) using a fixed threshold that has a certain associated impact, *e.g.,* riverine flooding—exemplified in Nguyen-Thi et al. (2012), or (ii) percentile-based thresholds (TT-DEWCE & WMO, 2016). For example, the Australian Bureau of Meteorology uses 95th and 99th percentiles to define very wet and extremely wet days, respectively (Haylock and Nicholls, 2000; Australian Bureau of Meteorology, 2020a). In this study, the 95th percentile threshold was chosen and applied to the rainfall data to identify extreme rainfall events for each station (identified after performing homogeneity tests). The rainfall values above (below) this threshold were categorised as extreme (non-extreme) rainfall events, and the dates on which this occurred were extracted. These extreme rainfall event dates were then cross-checked with TC events that passed within the Fiji EEZ (taking into account any time difference between rainfall and TC datasets). TCs that corresponded to extreme rainfall event dates were defined as an 'extreme rainfall TC event' and all others were labelled as a 'non-extreme rainfall TC event'. Further, to examine the temporal variability between extreme rainfall events that were either TC related or TC unrelated, the total number of events within each set (*i.e.,* extreme rainfall events related to TCs or not related to TCs) was tallied, followed by evaluation of their rainfall distribution (*i.e.,* 24-hour rainfall in mm).

3.3. Spatio-temporal analysis

The spatial variability between extreme rainfall TC events and non-extreme rainfall TC events across each station was assessed by computing difference density plots over a $1^{\circ} \times 1^{\circ}$ grid box. The temporal analysis included evaluation of TC track attributes, such as duration, average moving speed, seasonality, cluster and SI groupings. The latter two attributes were derived from SMV21, where the SWP TC tracks (inclusive of Fiji TC tracks) were classified into respective clusters and SI groups (*see* Fig. 2). Except for clusters and SI categories, all the TC attributes evaluated were based on the TC tracks portion within the Fiji EEZ (*e.g.,* duration within the EEZ rather than the TCs entire duration).

In addition to the above parameters, each TC's proximity to land was investigated, as this has previously been noted as an important factor in rainfall occurrence (Waylen and Harrison, 2005; Byun and Lee, 2012; Khouakhi et al., 2017). In this case, the minimum distance between each station and the TC's track position, when closest to the land, was measured. Further, TC intensity was included to assess its relationship with extreme rainfall (building on the observations made in the Introduction). All TCs were classed into five categories, in addition to a tropical depression category, following the Australian TC intensity scale (Table 2; Australian Bureau of Meteorology, 2020b). The classification scheme uses the maximum 10-min sustained wind speed (Table 2). TCs with no information on wind speed (and/or pressure estimates) in the SPEArTC database were not included in this analysis. Further, the influence of climate modes on both extreme rainfall TC events and non-extreme rainfall TC events was also investigated.

3.4. Decision tree analysis using C5.0

A decision tree analysis was performed to identify the relative importance of each TC track characteristic (and climate mode) on extreme rainfall events. This is a classification procedure that utilises a tree structure to model the relationships between input features (predictors) and potential outcomes (target/diagnosis variable) (Brodley and Friedl, 1997; Lantz, 2019). The decision tree approach is advantageous given it works well with a variety of data (*i.e.,* does not require the linearity assumption), and is not sensitive to outliers (Nam et al., 2018). Several algorithms have been implemented over time to compute decision trees. However, the most well-known is the C5.0 algorithm, which other studies have successfully applied to generate decision trees to detect TC genesis or analyse TC impact risks (*e.g.,* Park et al., 2016; Nam et al., 2018; Zhang et al., 2019).

Table 2
TC classification based on the Australian intensity scale. Maximum mean wind refers to 10-min sustained wind speed.

Category	Abbreviation	Maximum Mean Wind (km/h)
Tropical Depression	TD	< 63
One	Cat 1	63–88
Two	Cat 2	89–117
Three	Cat 3	118–159
Four	Cat 4	160–200
Five	Cat 5	> 200

Initially developed as C4.5, the C5.0 algorithm is an improved version that is efficient and performs well with large datasets, which can be continuous, categorical or binary (Quinlan, 1993; Kuhn and Johnson, 2013). The core algorithm of C5.0 (based on C4.5) uses a recursive binary split system to develop a decision tree from training data (Quinlan, 1993; Park et al., 2016; Rulequest Research, 2019). At each node of the decision tree, the most efficient attribute is selected, based on the entropy concept, to split the training samples into two groups by class (Kuhn and Johnson, 2013; Park et al., 2016; Lantz, 2019; Packt Editorial Staff, 2020). In other words, it partitions the dataset into smaller and homogenous groups such that each node of the split contains a larger proportion of one class, which is known as purity (Kuhn and Johnson, 2013). Purity is defined as having maximum accuracy or minimal misclassification errors and is evaluated through entropy/information gain (Quinlan, 1993; Kuhn and Johnson, 2013). In information theory, entropy assesses the uncertainty associated with a random variable (Shannon, 1948; Fan et al., 2011) and is calculated through the following equation:

$$\text{Entropy}\ (S) = -\sum_{i=1}^{c} p_i \log_2 p_i \tag{3}$$

Thus, for a given segment of data (S), p_i refers to the proportion of values falling into the class level i, and the term c refers to the number of different class levels. A low entropy with a minimum zero value indicates a completely homogenous sample, while a maximum value of one indicates higher uncertainty (Fan et al., 2011). In addition, information gain (for feature F) is associated with the changes in the homogeneity resulting from a split on each possible feature. This is computed by subtracting the weighted entropies of each branch resulting from the split (S_2), from the original entropy before the split (S_1), demonstrated in equation (4) (Packt Editorial Staff, 2020).

$$\text{InfoGain}(F) = \text{Entropy}\ (S_1) - \text{Entropy}\ (S_2) \tag{4}$$

After a split, a complication arises given that the data is divided into more than one partition. To accommodate this, the total entropy across all the partitions needs to be considered when computing Entropy (S_2). This is attained by weighting each partition's entropy according to the proportion of all records falling into that partition, demonstrated in equation (5) (Packt Editorial Staff, 2020).

$$\text{Entropy}\ (S) = \sum_{i=1}^{n} w_i\ \text{Entropy}(P_i) \tag{5}$$

Equation (5) ensures that the resultant total entropy from a split is the sum of entropy of each n partitions, weighted by the proportion of cases falling in the partition w_i. In training a decision tree, the best split is selected by maximising information gain, which is achieved by eliminating more entropies.

One of the C5.0 algorithm outcomes is the variable importance/attribute usage, presented either as 'splits' or 'usage'. The 'splits' metric calculates the percentage of splits associated with each variable or predictor. The 'usage' metric or the importance of predictors is measured by determining the proportion of the training set samples that fall into all the terminal nodes after the split (Quinlan, 1993; Kuhn and Johnson, 2013). The information gained from this metric enables identification of important input variables (in terms of their contribution) to predict the target/diagnosis variable (Park et al., 2016).

Consistent with previous studies (*e.g.,* Hsieh and Tang, 1998; Clark, 2004), the model was validated to overcome the problem of over-fitting (Kuhn and Johnson, 2013). The validation process assures the accuracy of the predictions that the model generates. In this study, this process was carried out in two parts. First, through the retrospective pruning process to ensure that the minimum number of samples (training data) supplied for at least two branches of a split did not result in an under-/over-fitted model (Rulequest Research, 2019). Next, a 10-fold cross-validation was performed, whereby the training dataset was randomly partitioned into 10 folds (*i.e.,* split into 10 groups or partitions). Around ~90% of the dataset (nine folds) were trained using the C5.0 model and then tested/validated on the remaining ~10% of the dataset to estimate the model's performance. This process was repeated until every fold was used as a test dataset (Quinlan, 1993; Kuhn and Johnson, 2013; Rulequest Research, 2019).

4. Results

4.1. Analysis of rainfall data

The FMS daily rainfall data's homogeneity was assessed to identify a reliable time period for further analyses. An initial quality check revealed inconsistencies (shown as proportions of missing data) in the dataset for the majority of the stations (Table 1). The homogeneity tests confirmed that, while a few stations showed no/minor indications of non-homogeneity (*e.g.,* Nadi and Suva), some stations either had a proportion of data missing—such as Lakeba (4.01%), Udu Point (17.83%), and/or exhibited a step jump late in the data series, *e.g.,* Nabouwalu with six changepoints and start date within the 2017 TC season (Table 3). Therefore, to obtain a reliable time period that included quality daily rainfall records for the longest overlapping period possible (1983–2017), with the least number of changepoints, Nadi, Suva and Labasa stations were chosen for the remainder of this paper's analyses. While Labasa station does contain 2.25% of missing data; it is one of the longest records (start date/season as 1978). In addition, these three stations largely represent the western (Nadi), northern (Labasa) and central/eastern (Suva) division or regions of Fiji, which are referred to throughout this paper.

The timeseries of the quality checked daily rainfall records for each station within the TC seasonal months is shown in Fig. 3. There is a high degree of spatial and temporal variability observed in the timeseries of daily rainfall data across all three stations from 1983 until 2017. For instance, on average, Suva recorded ~1793 mm/TC season, which is greater than Nadi (~1520 mm/TC season) and Labasa (~1650 mm/TC season). The higher mean value observed for Suva is likely due to its location within the wet zone (eastern side of Fiji), where rainfall events are mostly influenced by a range of other climate features than TCs

Table 3
Summary of evaluated (*i.e.,* after homogeneity analysis) daily rainfall data from the FMS. The table shows the number and years of changepoints, that is, the step jumps in the daily rainfall dataset. The start date and proportion of missing data (based on the TC season months) follows the most recent changepoint between each station's identified time period until 30th April 2017.

Station	No. of changepoints	Changepoint years	Start Date	% of missing data
Nadi	0	–	1/01/1960	0.00%
Suva	1	1982	1/11/1982	0.03%
Labasa	1	1977	1/11/1977	2.25%
Lakeba	2	1992, 1999	1/11/1999	4.01%
Rakiraki	2	1982, 2013	1/11/2013	2.07%
Nabouwalu	6	1970, 1987, 1999, 2002, 2007, 2016	8/12/2016	1.39%
Udu Point	3	2001, 2003, 2014	1/11/2014	17.83%

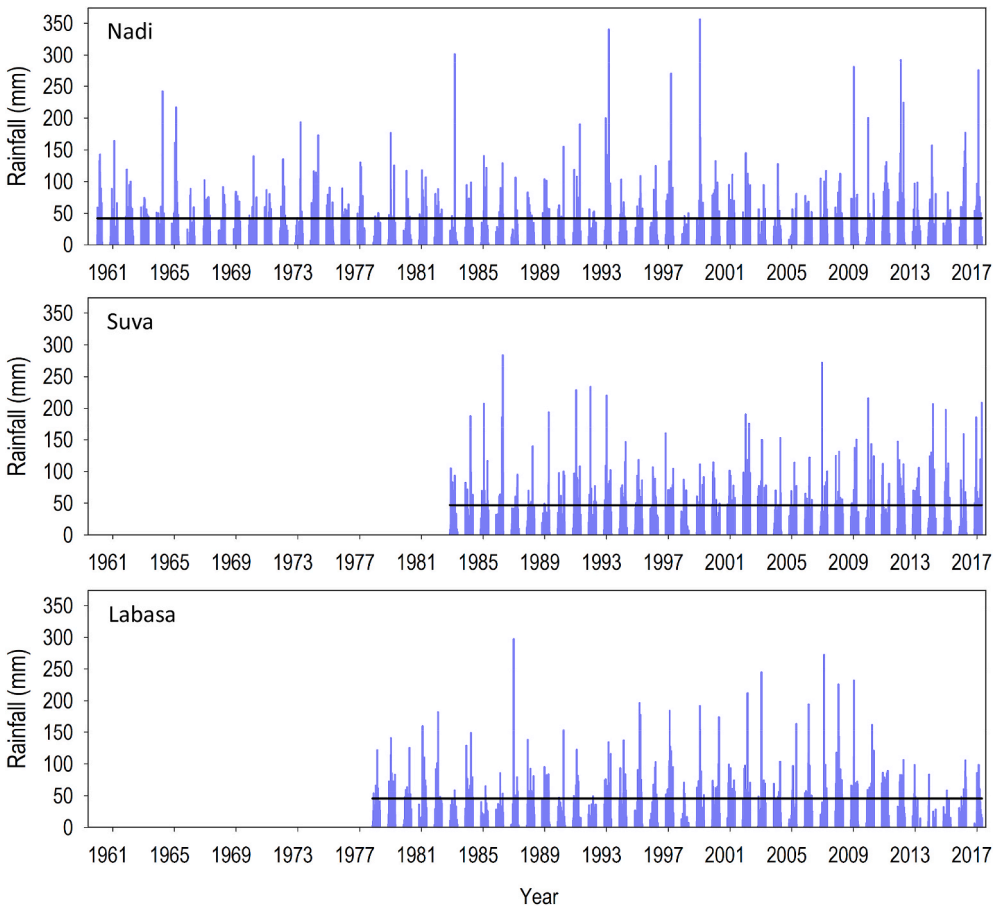

Fig. 3. Timeseries of daily rainfall data (in mm) within TC seasonal months (November–April) for Nadi (top), Suva (middle), and Labasa (bottom) stations, based on the start dates summarised in Table 3. The horizontal scale is the same for all timeseries (1960–2017). The black line represents the 95th percentile threshold.

(summarised in the Introduction). However, when the maximum daily rainfall values are evaluated (*i.e.,* 24-hour rainfall), both Nadi (~356 mm) and Labasa (~297 mm) experience higher maximum daily rainfall values compared to Suva (~284 mm). Notably, the daily peak values for Nadi and Labasa exceed the climatological average of their corresponding months (~253 and ~275 mm, respectively). This indicates much higher variability in the daily accumulated rainfall, as both Nadi and Labasa are located on the drier leeward zone (Fig. 1), where large proportions of rainfall or extreme rainfall are associated with TCs.

Fig. 3 also shows the 95th percentile threshold used to identify the extreme daily rainfall events across each station. The 95th percentile rainfall thresholds used to define extreme rainfall are 42 mm for Nadi, 46 mm for Labasa and 47 mm for Suva. The extreme rainfall events were also cross-checked with the SPEArTC dataset to explore the influence of TCs (Fig. 4). It was found that TCs account for approximately 14%, 16% and 18% of extreme (95th percentile) rainfall at Suva, Nadi and Labasa, respectively (Fig. 4, top row). Interestingly, the mean magnitude of rainfall is higher for TC related extreme events (average of 99 mm for Nadi, 105 mm for Suva and 87 mm for Labasa) than non-TC related events (74 mm for Nadi, 78 mm for Suva and 75 mm for Labasa) (Fig. 4, bottom row). The difference between extreme rainfall totals are statistically significant for Nadi and Suva, while for Labasa, the mean difference is minimal.

4.2. Evaluation of TC track attributes associated with extreme rainfall events

A total of 104 TCs traversed the Fiji EEZ between 1983 and 2017 season (Fig. 5). Of those, 35, 34 and 30 TCs were associated with extreme rainfall events over the northern (Labasa), western (Nadi), and central/eastern (Suva) regions of Fiji, respectively (Fig. 5, left panels). This is consistent with Terry (2007) who emphasised that the northern and western sides of high islands (within the South Pacific) are typically more vulnerable to the impacts of TCs. Density plots (where varying grid sizes were trialled, with the $1^\circ \times 1^\circ$ providing the best resolution for this analysis) were constructed to examine in more detail the differences in TC activity for both extreme and non-extreme rainfall events (Fig. 5, right panels). The density plots demonstrate that TCs resulting in extreme rainfall events generally pass close to the mainland, as shown by positive density values (*i.e.,* the number of TC tracks per $1^\circ \times 1^\circ$ grid) concentrated near each station. In particular, positive values observed along the western-southern side of Fiji highlight that extreme rainfall TC events associated with Nadi and Suva usually track from the western side. Interestingly, positive values associated with Labasa station are also observed beyond the Fiji group, potentially resulting from TCs traversing from the northern and eastern side.

TCs that did not result in extreme rainfall (classed as non-extreme rainfall TC events) in Fig. 5 (central panels) are clearly shown to be located farther away from each station (also shown by the density plots). These differences indicate that one of the key characteristics of TCs to influence extreme rainfall events could be their proximity to the station (or landmass in general). To investigate this further, the minimum distance from each station to the TC tracks of both extreme rainfall TC events and non-extreme rainfall TC events was measured, as illustrated in Fig. 6 (left panels). The minimum distance needed for TC events to deliver extreme rainfall across all stations significantly differs from the non-extreme rainfall TC events (Table 4). The average minimum distance of extreme rainfall TC events tends to vary with the lowest value

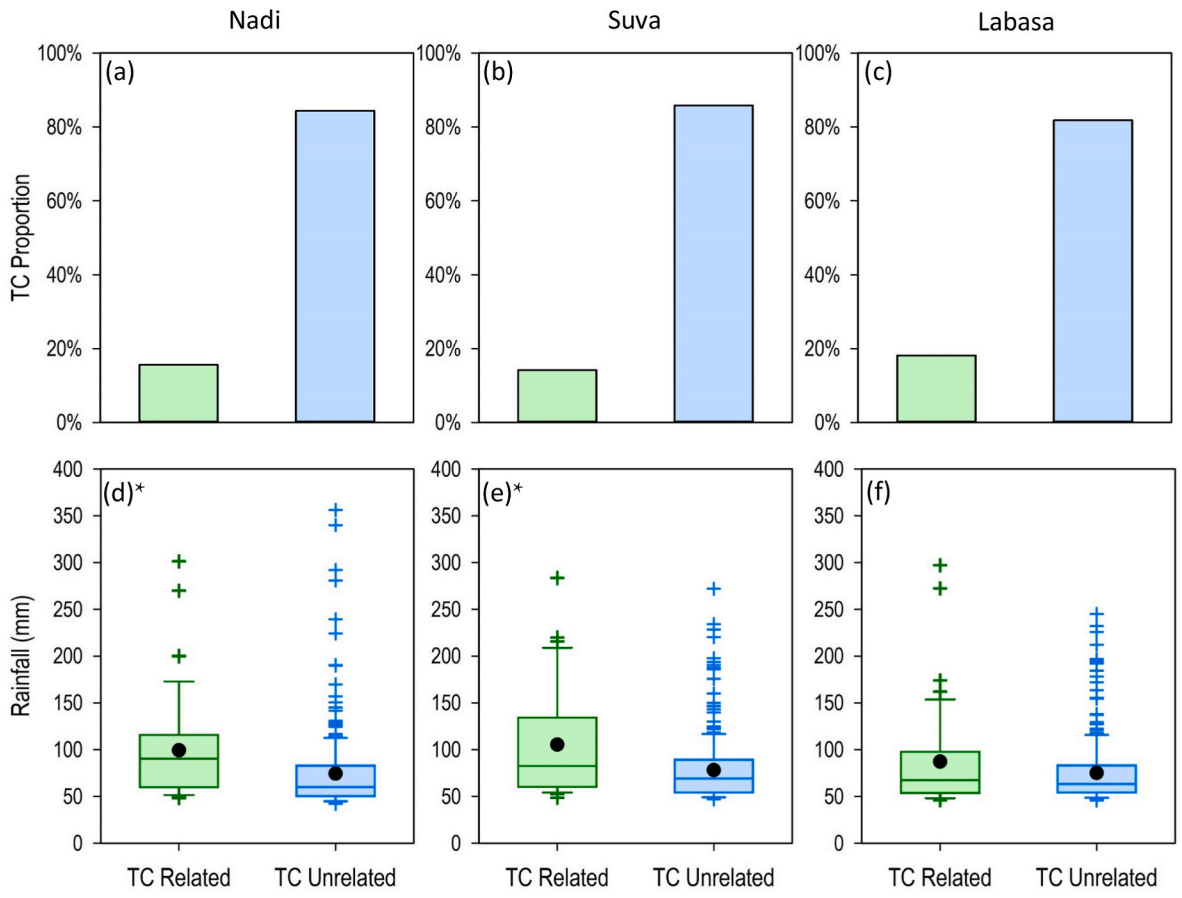

Fig. 4. Top row: Proportions of extreme rainfall events that are TC related and TC unrelated for Nadi (left panels), Suva (central panels) and Labasa (right panels) stations between 1983 and 2017. Bottom row: Rainfall distributions for extreme events (above 95th percentile threshold) that are TC related and TC unrelated. The boxes show the 25th and 75th percentiles, the lines inside the box mark the median, dots mark the mean and crosses marks the outliers (lower 5th and upper 95th percentiles). The asterisks in panels d and e denote that differences in the TC related/TC unrelated rainfall distributions are statistically significant at the 95% level, as per Mann-Whitney *U* test.

for Suva (~229 km), highest for Labasa (~274 km) and in-between for Nadi (~244 km).

Fig. 6 (central panels) shows that TC duration also plays a critical role in influencing extreme rainfall events, especially over the central/eastern and northern regions (Table 4). The most common TC duration among all three stations is ~2–3 days (for extreme rainfall TC events), while the duration of non-extreme rainfall TC events is generally short (< 2 days). In fact, the distributions for non-extreme rainfall TC events are skewed, where the median values are close to the lower quartiles for all stations, yet demonstrate extended durations (greater than ~3 days) shown through the outliers above the upper 95th percentiles (Fig. 6, central panels). Furthermore, the results displayed in the right panels of Fig. 6 reveal that extreme rainfall TC events have a lower average speed across all three stations, when compared to non-extreme rainfall TC events; however, this wasn't statistically significant based on the Mann-Whitney *U* test results in Table 4. Also, in both cases, the most common speed of TCs within Fiji EEZ fluctuates between ~15 km/h and ~30 km/h.

In line with the outcomes of SMV21 (*also* Fig. 2), the extreme rainfall TC events/non-extreme rainfall TC events for each station were also characterised and explored by cluster (column A), SI classification (column B) and seasonality (column C), shown in Fig. 7. This was to assess if variability in extreme rainfall events was related to the spatial location and geometry of the TCs and how often these events occurred within the TC season. The intensity (column D) of the TCs, based on their maximum wind speed within the Fiji EEZ and intensity scale outlined in Table 2, were also evaluated to see if TC category was related to their associated rainfall impacts. For all cases, a two-sample proportion test was performed by comparing the differences in the proportions of the two samples with the null hypothesis that the proportions are the same (Ashley et al., 2008; Rogerson, 2011).

The cluster grouping analysis (Fig. 7, column A) showed that the majority of the extreme rainfall TCs are associated with Cluster 4 (illustrated in Fig. 2). In particular, the impact of extreme rainfall TC events is statistically significant over the central/eastern region (Suva) when compared to the non-extreme rainfall TC events. In addition, the prevalence of sinuous type tracks for Suva (column B) is complementary to central SWP TCs (Cluster 4 type TCs as established in Figure 11 of SMV21). Furthermore, quasi-straight type tracks (column B) largely influence the extreme rainfall events, mainly on the drier side of Fiji, *i.e.,* Nadi (significant at 95%) and Labasa (not statistically significant but noteworthy).

The seasonality of both extreme rainfall TC events and non-extreme rainfall TC events (Fig. 7, column C) demonstrates that within the November–April TC season, Fiji experiences extreme rainfall events mostly from December to March. Extreme rainfall TC events significantly (at the 90% level based on a two-sample proportion test) impact Labasa early in the TC season (in December), while Nadi receives the majority of the extreme rainfall (> 30%) during the late TC season (in March). Further, the rainfall events influenced by both extreme rainfall TC events/non-extreme rainfall TC events are common (almost similar in proportion) during January and February (for Nadi and Suva), and March (for Suva and Labasa).

Many of the extreme rainfall events across all three stations are associated with both non-severe (TD and Cat 1–2) and severe (Cat 3–5)

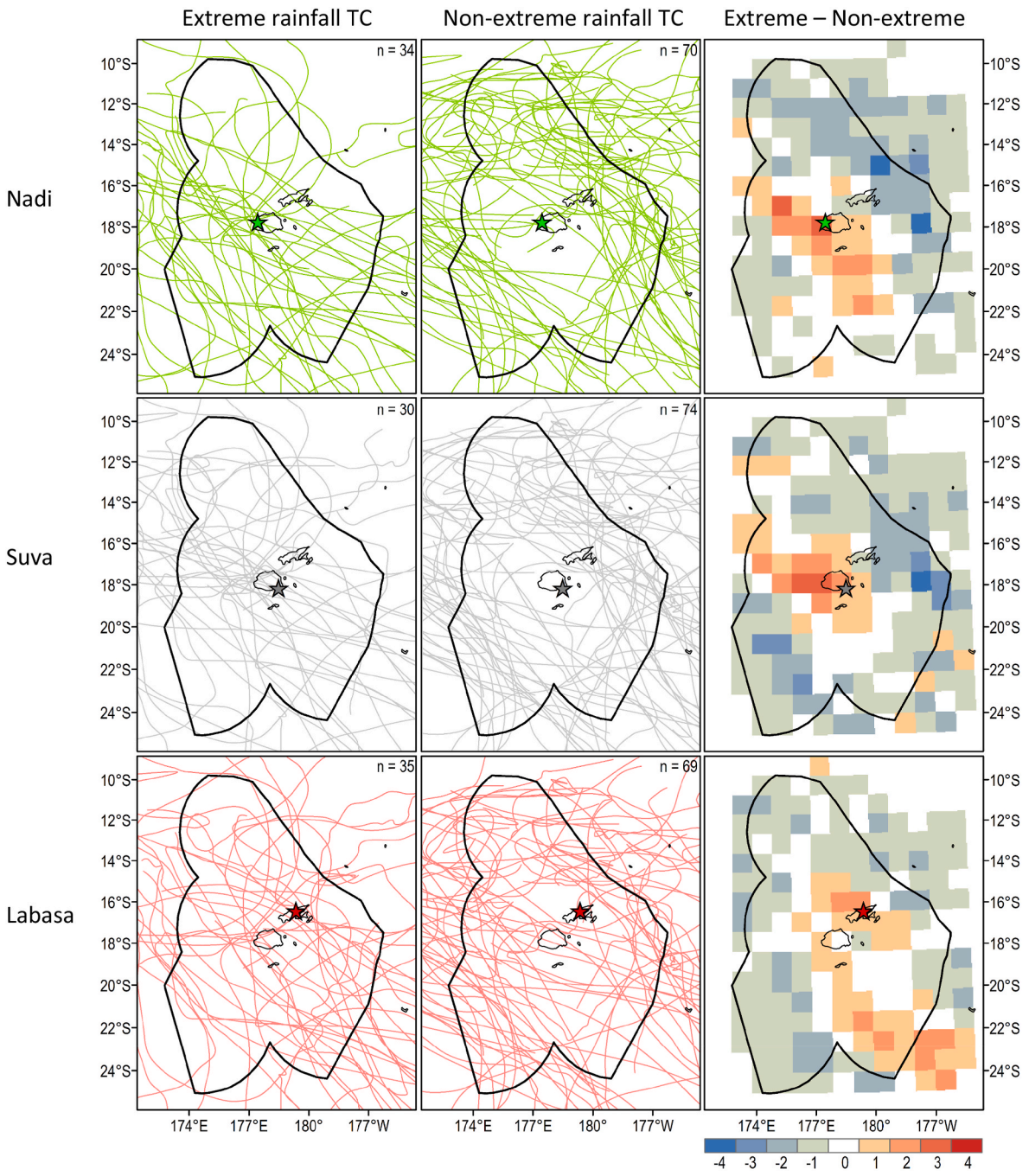

Fig. 5. Distribution of TC tracks within Fiji EEZ that resulted in extreme (left panels) and non-extreme (central panels) rainfall events across Nadi (top row; green star), Suva (middle row; grey star) and Labasa (bottom row; red star) stations. The number of TC tracks (n) is shown on the top-right of left and central panels. Right panels: TC track density differences for each station between extreme rainfall TC events and non-extreme rainfall TC events at 1° × 1° grid. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

TCs, illustrated in Fig. 7 (column D). For instance, the intensity range for extreme rainfall TC events across western and central/eastern regions includes Cat 1–2 and Cat 1, respectively, while those over the northern region correspond to Cat 2–3. In general, Cat 4 TCs significantly influence extreme rainfall across all three regions (Nadi and Labasa significant at the 90% and Suva at the 95% level). Although in very small proportion, Cat 5 TCs are also seen to modulate extreme rainfall events across all three regions. TDs tend to occur more frequently over the Fiji group, but do not often produce extreme rainfall when compared to non-extreme rainfall TC events (significant at the 95% level).

4.3. *Climate mode influence on TCs that produce extreme rainfall*

The influence of climate drivers on extreme/non-extreme rainfall TC events across all three stations were also evaluated (Fig. 8 and Table 5). The results highlight that ENSO is a primary driver of TCs and associated rainfall; however, the relationship is not spatially consistent between stations. In particular, extreme rainfall TC events tend to occur more during the LN phase, evident over the western region (Fig. 8, row 1, left panel). Conversely, EN may have influenced many of the extreme rainfall TC events over the central/eastern region while those associated with Labasa generally correspond to ENSO neutral (not statistically significant). Although not as convincing as the observed ENSO response,

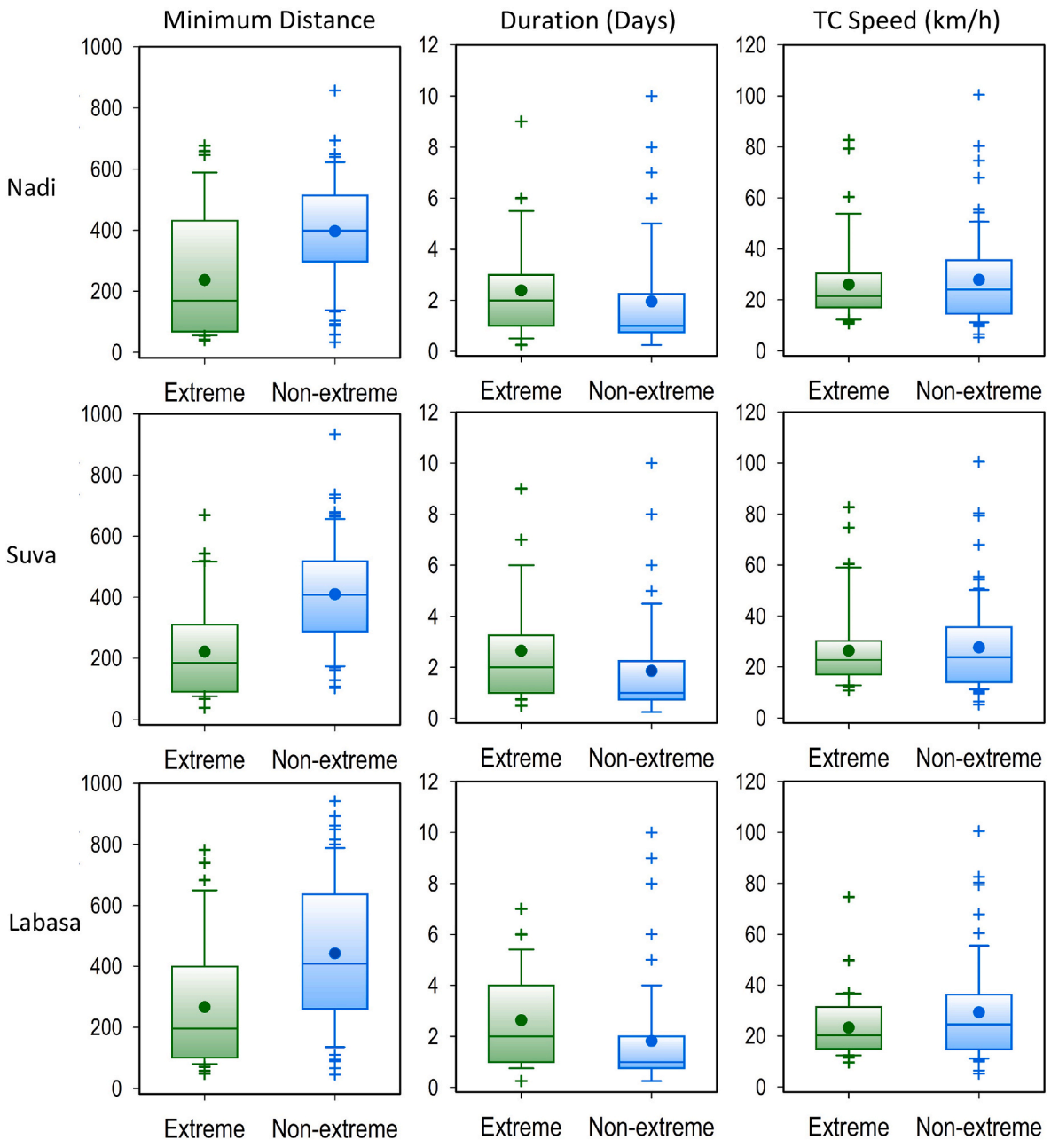

Fig. 6. Boxplots illustrating the distribution of extreme rainfall TC events (labelled as extreme) and non-extreme rainfall TC events (labelled as non-extreme) in terms of minimum distance from the land (left panels), TC duration (central panels) and average speed (right panels) across Nadi (top row), Suva (middle row) and Labasa (bottom row) stations. The boxes show the 25th and 75th percentiles, the lines inside the box mark the median, dots mark the mean and crosses marks the outliers (lower 5th and upper 95th percentiles).

Table 4
Mann-Whitney *U* test *p*-values comparing the track attributes (shown in Fig. 6) of extreme rainfall TC events and non-extreme rainfall TC events for each station. Bold values denote statistically significant results at the 95% level.

TC track attributes	Nadi	Suva	Labasa
Minimum distance	**0.00**	**0.00**	**0.00**
Duration	0.05	**0.01**	**0.00**
Average speed	0.79	0.97	0.34

the analysis shows extreme rainfall TC events over the northern region occur more often during the negative Modoki phase, *i.e.,* LNM (Fig. 8, row 2, right panel). The influence of Indian Ocean SST variability (represented by IOD E and II) was not found to be statistically significant (Table 5) for both extreme rainfall TC events and non-extreme rainfall TC events. However, extreme rainfall TC events across Fiji's western region appear to be influenced by the positive phase of SAM (Fig. 8, row 5, left panel).

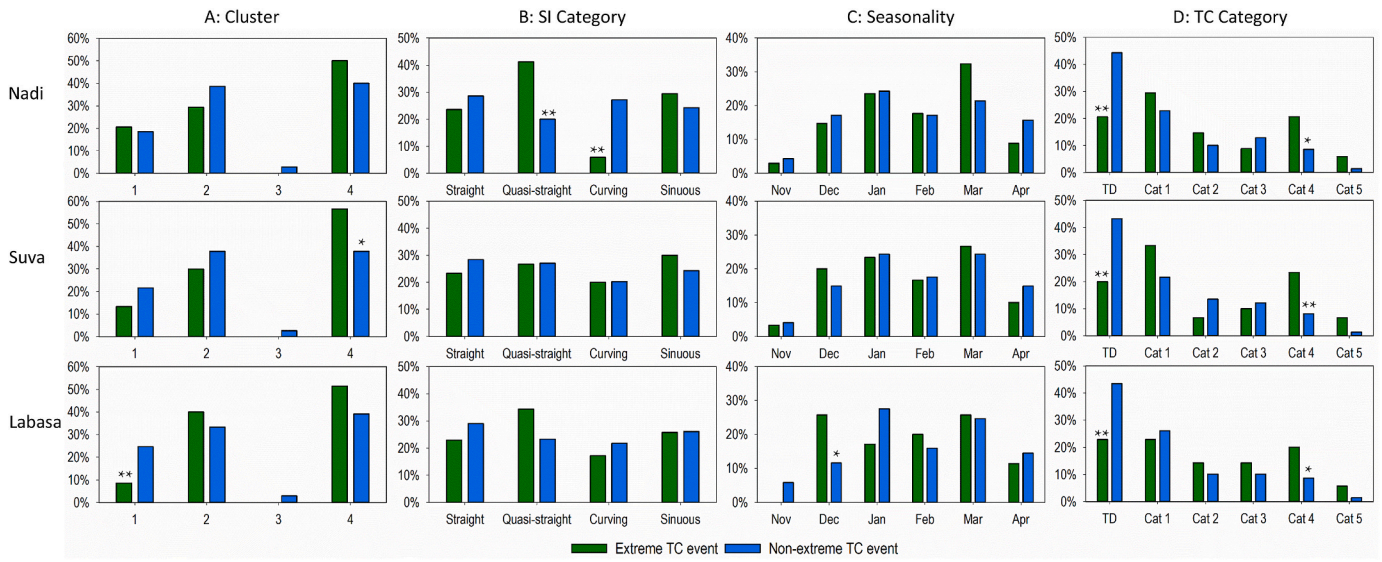

Fig. 7. Bar graphs illustrating the distribution of TCs that result in extreme and non-extreme rainfall events, grouped by clusters (column A), sinuosity (column B), seasonality (column C) and intensity/category (column D) for Nadi (top row), Suva (middle row) and Labasa (bottom row). Single (*) and double (**) asterisks indicate that the proportion of TCs associated with either event (extreme rainfall TC events/non-extreme rainfall TC events) for each attribute grouping is significantly less at the 90% and 95% levels, respectively.

4.4. Classification of TC track attributes using C5.0

The analysis presented so far has confirmed that extreme rainfall events across Fiji are modulated by varying TC track characteristics as well as by the surrounding climatic conditions, in particular ENSO and SAM. Therefore, to achieve the third objective of this paper (*i.e.,* to develop a decision tree model to rank the importance of each predictor that can assist in modelling the potential risk impacts), a decision tree model was constructed using the C5.0 classification algorithm described in Section 3.4. The purpose of the decision tree model was to objectively classify whether a TC will result in an extreme rainfall event or not. For modelling purposes, all TC events resulting in extreme rainfall (from all three stations) were categorised collectively as "Extreme". Similarly, all non-extreme rainfall TC events (that resulted in non-extreme rainfalls) were categorised collectively as "Non-extreme". Such organisation of the TC data allowed the model to diagnose the input variables decisively as either "Extreme" or "Non-extreme". An "Extreme" output refers to an extreme rainfall event, while "Non-extreme" refers to all other TCs. Altogether 312 effective cases were categorised including 99 "Extreme" and 213 "Non-extreme" cases. The TC track attributes (and climate indices) explored in this paper (across each station) were used as input variables to the model.

4.4.1. Calibration and validation of the model

The minimum number of cases (to be supplied at each node for splitting into two branches) determined through the retrospective pruning process was three. The results of 10-fold cross-validation, presented in Table 6, demonstrate the model's overall stability and consistency in terms of accuracy, error, tree sizes, and attribute usage at each fold. The attribute usage demonstrates the relative importance of the predictors (input variables) used in the construction of the tree model at each fold. Although some folds exhibit a broad range of tree sizes or error rates (*e.g.,* in Folds 2 and 3), their dependency on input attributes (*i.e.,* the predictors) remains consistent in the majority of cases.

4.4.2. Decision tree model

The decision tree model generated using the C5.0 algorithm and based on 312 effective cases is illustrated in Fig. 9. The overall accuracy of the model is 81%, where 52 and 202 events were classified correctly as "Extreme" and "Non-extreme", respectively (Table 7). This is supplemented by the sensitivity (53%), and specificity (95%) results in Table 7, which show the model's ability to measure the proportions of correctly classified TCs that resulted in extreme and non-extreme rainfall events, respectively. The model's performance is also shown by the positive predictive value (83%) and negative predictive value (81%)—a performance indicator demonstrating the portion of "Extreme" and "Non-extreme" cases that were predicted correctly, respectively. Further, the error rate of the model is 19%, as 47 "Extreme", and 11 "Non-extreme" cases were classified incorrectly (Table 7).

Fig. 9 presents the output process of the model. The first splitting criterion delivers all the TC tracks with a minimum distance $>$ 217.13 km to the end nodes classed as "Non-extreme", while the remaining tracks (*i.e.,* $<$ 217.13 km) are evaluated further according to the clusters. TCs occurring from the Cluster 1 domain (western SWP, *see* Fig. 2) are linked directly to "Extreme". Conversely, those associated with Clusters 2 and 4 (central-to-eastern SWP; Fig. 2) are subsequently forwarded to the seasonality criteria. TCs of early season (November) and late season (April) generally result in non-extreme rainfall events, while those occurring during December, February, and March are more likely to produce extreme rainfall events. This also includes January, where (according to the decision tree model) TCs need to occur for less than three days while maintaining a minimum distance of at most 153.81 km.

The relative importance of the predictors used in the tree model is shown in Table 8. The usage metric indicates that when an attribute is the most related variable to the target variable, the decision tree model frequently uses this attribute for classification (Nam et al., 2018). The usage statistics (*in* Table 8) indicate that the minimum distance attribute was used in 100% of the cases, followed by cluster groupings, seasonality, and duration as the least used attributes. This is also consistently shown across many folds (Table 6). These predictors correspond to the TC track attributes explored in Figs. 6 and 7 that showed significant distinctions between extreme and non-extreme rainfall events induced by TCs. The 'split' metric (Table 8) depicts the prevalence of minimum distance (as 40% – twice more than the other attributes), which is reflected twice in Fig. 9.

5. Discussion

The objectives of this paper were to (1) quantify the contribution of TCs to extreme rainfall events across a number of stations across Fiji; (2)

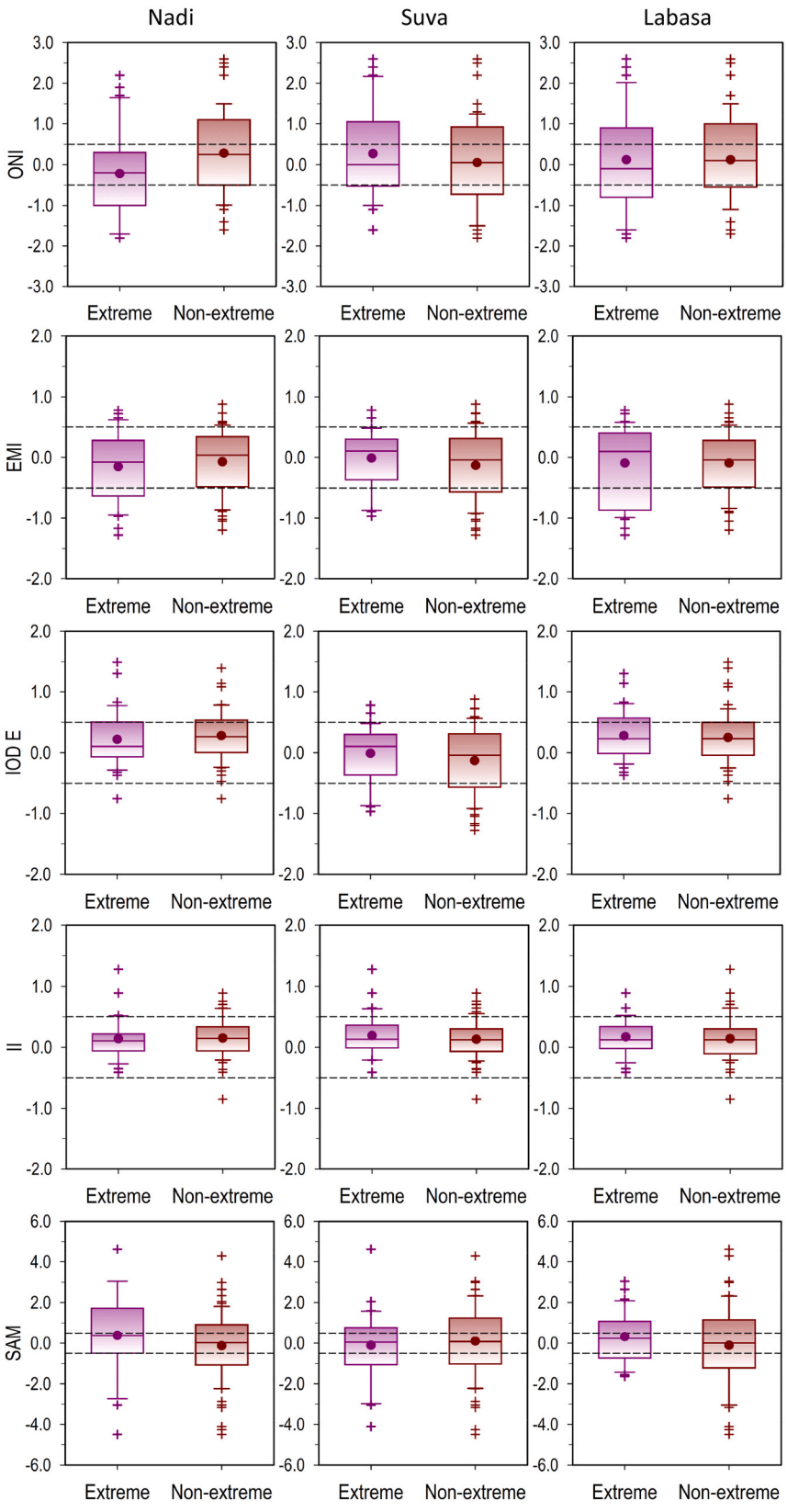

Fig. 8. Classification of TCs that result in extreme and non-extreme rainfall events according to the climate indices ONI, EMI, IOD E, II (*in* °C) and SAM across Nadi (left panels), Suva (central panels) and Labasa (right panels) stations. The horizontal dashed lines display the ±0.5 SD thresholds from the mean, where values above (below) represent positive (negative) phases of each climate mode. The values in-between represent the neutral phase. The boxes show the 25th and 75th percentiles, the lines inside the box mark the median, dots mark the mean and crosses marks the outliers (lower 5th and upper 95th percentiles).

Table 5
p-values of Student's *t*-test comparing climate mode indices (illustrated as boxplots in Fig. 8), associated with extreme rainfall TC events and climate mode indices during non-extreme rainfall TC events for each station. Asterisk (*) denotes a statistically significant result at the 95% level.

Climate modes	Nadi	Suva	Labasa
ENSO	0.02*	0.37	0.92
EMI	0.39	0.34	0.86
IOD E	0.47	0.45	0.90
II	0.67	0.49	0.81
SAM	0.18	0.54	0.17

explore the relationship between TC characteristics, Indo-Pacific climate modes and occurrence of extreme rainfall events to identify potential predictors, and; (3) use these findings to rank the importance of each predictor through a tree-based model that could be used for risk impact analysis and decision making. A homogeneity test was performed across all stations (listed in Tables 1 and 3) to assess the FMS daily rainfall data quality. The results suggested that rainfall data from 1983-onwards were consistent across Nadi, Suva and Labasa stations with minimal missing data (Table 3 and Fig. 3).

5.1. Characteristics of extreme rainfall TC events/non-extreme rainfall TC events

The validated rainfall data were quantified into extreme and non-extreme rainfall events based on the 95th percentile threshold. The results demonstrated that although the number of TCs associated with extreme rainfall events was less, compared to events not related to TCs, the magnitude of rainfall delivered was, on average, higher (Fig. 4). The large proportions of extreme rainfall events unrelated to TCs indicate that the Fiji region is influenced by other synoptic weather and climate features delivering large amounts of rainfall. For example, the Asian-Australian monsoon extends to the tropical SWP in summer, reaching as far east as Fiji and bringing additional rainfall (Stephens et al., 2018). Various studies have also reported on the existence of other surrounding climate features (*e.g.,* SPCZ, southeast trade winds, and MJO) that influence rainfall variability (Basher and Zheng, 1998; Folland et al., 2002; Mataki et al., 2006; Chand et al., 2020; Deo et al., 2021). The MJO was not considered since intraseasonal phenomena were not the focus of this study; however, it could be an avenue for future research, along with the influences of other climate features. Also, the sub-daily rainfall could be considered which may enable quantification of sub-daily rainfall extremes or assessment of those events that were not categorised as extreme rainfall (but had significant impacts). In this study, a simple binary approach was taken (*i.e.,* extreme or non-extreme); however, this method could be expanded to additional categories that represent an increasing scale of rainfall amounts (*e.g.,* 1–50th percentile, 50–75th percentile, 75–95th percentile, 95–100th percentile). The choice of thresholds could be tailored to the application if a specific rainfall amount is of interest.

For those TCs that did result in extreme rainfall across Fiji, it was established that track morphology plays a critical role. It was established in SMV21 that the Fiji region is mostly impacted by Cluster 4 TCs (those occurring within central SWP), which typically have quasi-straight track morphologies (as per Table 4 of SMV21). Our analyses demonstrated that the majority of extreme rainfall TC events (across all three stations) are typically longer and quasi-straight and have a slower average moving speed than non-extreme rainfall TC events. This is consistent with Chen et al. (2011) and Hernández Ayala and Matyas (2016), who also reported that slow moving TCs often cause heavy/extreme rainfall with extended duration.

Additional factors that may influence TC related rainfall, including the minimum distance of the TC from land, average moving speed, seasonality and intensity were also explored in this paper. Several studies have reported that as a TC approaches a region, the amount of rainfall delivered also increases (Hernández Ayala and Matyas, 2016; Lonfat et al., 2004; Ng et al., 2015). We show that this is also the case for Fiji (Fig. 6, left panels). In addition, slower moving systems are more likely to be associated with extreme rainfall than faster moving systems. The possible link between TC intensity and rainfall has been explored by several researchers and, although there is some debate in the literature, it has been commonly reported that a linear relationship exists between TCs daily maximum wind speeds and daily total rainfall (Rao and Macarthur, 1994; Cerveny and Newman, 2000; Shepherd et al., 2007; Hernández Ayala and Matyas, 2016). However, our results suggest that this relationship does not hold true (for Fiji at least) since both Cat 1 and Cat 4 rated TCs influenced extreme rainfall events across Fiji. One of the factors that could account for this is the TC system's size, which is based on the TC's radius measured from its centre (Terry, 2007). This feature was not investigated in the present study since the required data was not available for all TC events analysed; however, future studies could explore it.

Indo-Pacific climate variability was shown to influence the likelihood of extreme TC related rainfall under two primary conditions. That is, extreme rainfall TC events were found to be influenced mainly by LN and positive SAM for the western part of the country. This is complementary to the typical influence of the LN phase, which modulates TC genesis southwestwards (Magee et al., 2017), and post-formation, TCs traverse Fiji from the western side (Chand and Walsh, 2009). Conversely, EN influenced extreme rainfall TC events are prevalent over Fiji's eastern division, also demonstrated in Deo et al. (2021).

Table 6
Results of 10-fold cross-validation based on 312 effective cases with all TC track attributes and climate indices as predictors.

Model	Tree size	Accuracy (%)	Error rate (%)	Attribute usage (%)	
Fold 1	4	71	29	Minimum distance	100
				Cluster	29
				Seasonality	25
Fold 2	9	87	13	Minimum distance	100
				Cluster	33
				Seasonality	29
				Duration	8
				SAM Index	6
Fold 3	2	66	34	Minimum distance	100
Fold 4	7	77	23	Minimum distance	100
				Category	30
				ENSO Index	3
				Duration	3
Fold 5	2	61	39	Minimum distance	100
Fold 6	7	74	26	Minimum distance	100
				Category	30
				Duration	3
				Cluster	3
Fold 7	3	72	28	Minimum distance	100
				Cluster	29
Fold 8	6	81	19	Minimum distance	100
				Seasonality	30
				Duration	8
Fold 9	6	81	19	Minimum distance	100
				Seasonality	31
				Duration	8
Fold 10	4	71	29	Minimum distance	100
				Cluster	29
				Seasonality	26

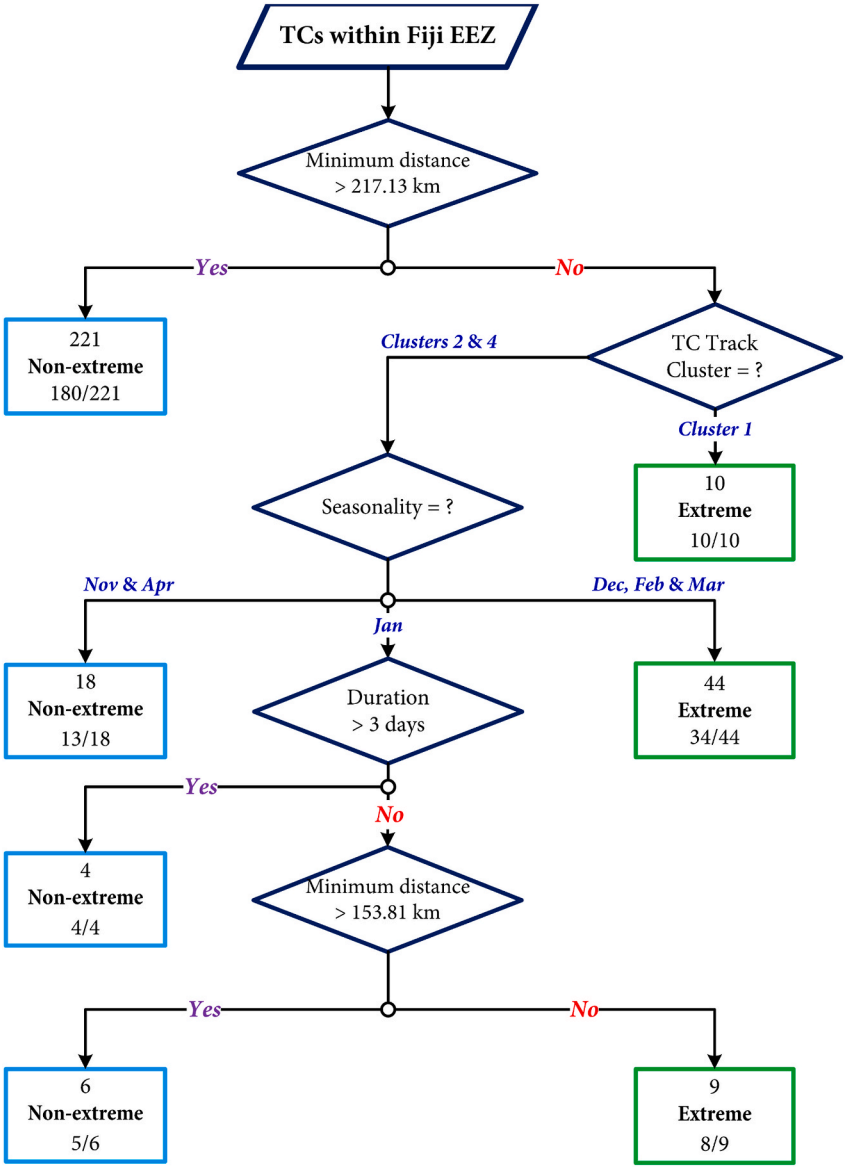

Fig. 9. Decision tree model for extreme and non-extreme rainfall occurrences generated using 312 cases (*i.e.,* the sum of all the extreme rainfall TC events and non-extreme rainfall TC events across Nadi, Suva, and Labasa) and nine TC attributes (explored in Figs. 6–8) as predictors/input variables. The parallelogram (supplied with predictors) specifies the start of the C5.0 algorithm. Each split node follows a rhombus box which contains questions. The corresponding answers include Yes (purple), No (red) and listings for the cluster and seasonal months (*in* blue font). The final diagnosis boxes (end nodes or the leaf) indicate the decisive outcomes of TCs as extreme (green) or non-extreme (blue) rainfall events. Numbers on the top of each final diagnosis box correspond to the number of cases diagnosed at each node or mapped to the leaf. Statistics below (as fractions) represent the precision of the diagnosis where the numerator is the number of correctly classified cases. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 7
Confusion matrix of the final model based on 312 cases and nine TC attributes. Note: "Extreme" and "Non-extreme" in this table refer to TCs that have either resulted in extreme or non-extreme rainfall events.

		Observation		Sum of classification
		Extreme	Non-extreme	
Prediction	**Extreme**	52	11	63
	Non-extreme	47	202	249
Sum of observation		99	213	312
Accuracy		81%		
Error rate		19%		
Sensitivity (Hit Rate)		53%		
Specificity (Selectivity)		95%		
Positive Predictive Value		83%		
Negative Predictive Value		81%		
Tree size		7		

Table 8
Relative importance of the predictors (input variables) represented as splits at the nodes and attribute usage in the development of a decision tree model.

Attributes	Splits (%)	Usage (%)
Minimum distance	40	100
Cluster	20	29
Seasonality	20	26
Duration	20	6

5.2. Importance of TC track attributes in potential risk analysis

The range of TC track attributes and climate conditions were used as model predictors to develop a decision tree with extreme rainfall as the desired output. The results demonstrated that the minimum distance of a TC from land is the most important factor when evaluating whether or not it will result in extreme/non-extreme rainfall for Fiji. Notably, the model identified two quantitative criteria for minimum distance (*i.e.,* minimum distance < 153.81 km and minimum distance < 217.13 km) that can be applied on a broader scale to classify extreme and non-extreme rainfall events across Fiji. While the former criterion falls within the latter, seemingly, the former tends to be more decisive for January TCs, that are closer to the land with shorter durations. This distance range falls within the average size of TCs in the South Pacific, in terms of radius, which is usually 250–350 km on average that are classed as medium to average in size (Terry, 2007). Also, these thresholds (*i.e.,* minimum distance) indicate that extreme rainfall events are mostly influenced by TCs that often pass by the Fiji group than those making landfalls. Indeed, Fiji has encountered some TCs in recent years that have caused substantial damage without making landfall (*e.g.,* TC Evan in 2012 and TC Gita in 2018).

The model output also confirmed that TC tracks cluster grouping (based on the outcomes of SMV21 and Fig. 2) and/or seasonality could be used to identify extreme and non-extreme rainfall events. Although TCs typically occur within the November–April season, the model distinctly identified in which months TCs that influence extreme rainfall could be expected. However, TCs occurring in January are not as definitive as other TC seasonal months because, as per the model, extreme rainfall events during this month are also modulated by TC duration and minimum distance.

By applying the decision tree methodology in the context of filtering important TC parameters that impact extreme rainfall events, some important new insights were gained. The model output implies that while a range of TC attributes influences extreme rainfall events across Fiji, only a few could be used as potential indicators/predictors for risk impact analysis or decision making. For example, the TC intensity parameter was not selected in generating the tree model (Fig. 9), indicating that TC intensity is not always a useful guide for determining the rainfall-related impacts of TCs (in Fiji at least) – a misconception commonly adopted for preparedness during TC events (Terry, 2007). However, care should still be taken in dismissing attributes (*i.e.,* that were not used by this model) particularly if applied to other regions, and further refinements of this process, such as the boosting application of C5.0 (Kuhn and Johnson, 2013) and the random forest model (Breiman, 2001) may further improve the predictive capability. Also, with these additional tools, more variables such as sub-daily rainfall, location of the SPCZ, southeast trade winds, MJO, TC size (mentioned in the previous sections), wind speed as well as economic losses due to TCs could be considered as a scope for future investigation.

6. Conclusions

Extreme or heavy rainfall events are of great concern as they often lead to excessive river flows and severe flooding for Pacific Island nations. The severity of their impact on society has been mostly felt by the small island developing states (Gero et al., 2011) of the SWP that have limited adaptive capacity. Fiji is a prime example of a nation that regularly experiences extreme rainfall impacts, with close to 20% of these events associated with TCs. There is a clear need for improved prediction of TC impacts to increase the general public's resilience to these events' destructive impacts (specifically TC induced flooding). Indeed, a range of meteorological warnings are provided by the Regional Specialised Meteorological Centre Nadi once a TC has formed (or is approaching the group); these warnings include strong winds, heavy rainfall, flooding, and storm surges, and are intended to promote extra precaution and preparedness. However, refinement of the likely severity of the approaching TCs impact will no doubt improve the effectiveness of the response. In this study, we demonstrate the potential to use TC parameters and climate indices as predictive tools in potential impact assessment (in terms of extreme rainfall). Importantly, we show the advantages of generating tree-based models based on these predictor variables to assess potential TC impacts (in terms of extreme rainfall). On a broader scale, the application of this model could result in improved TC outlooks and risk evaluations, and similar models could be used by forecasters and decision-makers including environmental managers and city planners, to assist in mitigating TC impacts over the Fiji Islands.

CRediT authorship contribution statement

Krishneel K. Sharma: Data curation, Formal analysis, Investigation, Validation, Software, Visualization, Writing – original draft, Project administration, Conceptualization, Methodology. **Danielle C. Verdon-Kidd:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Andrew D. Magee:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Krishneel K. Sharma would like to acknowledge the financial contribution from the University of Newcastle (UoN) which provided an International Postgraduate Research Scholarship (UNIPRS) and Research Scholarship Central (UNRSC50:50) for funding his Research Higher Degree (RHD) program at UoN. The authors would also like to acknowledge the Fiji Meteorological Service for providing the rainfall data used in this study. Also, thanks to Olivier Rey-Lescure, formerly at the UoN for technical assistance with spatial analysis. The authors declare that there are no conflicts of interest regarding the publication of this article.

References

Australian Bureau of Meteorology, 2020b. What is a Tropical Cyclone? Available at: http://www.bom.gov.au/cyclone/tropical-cyclone-knowledge-centre/understanding/tc-info/. (Accessed 27 September 2020).

Ashley, W.S., Krmenec, A.J., Schwantes, R., 2008. Vulnerability due to nocturnal tornadoes. Weather Forecast. 23 (5), 795–807. https://doi.org/10.1175/2008WAF2222132.1.

Ashok, K., Behera, S.K., Rao, S.A., Weng, H., Yamagata, T., 2007. El Niño Modoki and its possible teleconnection. J. Geophys. Res. 112 (C11), 1–27. https://doi.org/10.1029/2006JC003798.

Australian Bureau of Meteorology, 2018. Southern Hemisphere Tropical Cyclone Data Portal. Available at: http://www.bom.gov.au/cyclone/history/tracks/. (Accessed 10 April 2018).

Australian Bureau of Meteorology, 2020a. About the climate extremes analyses. Available at: http://www.bom.gov.au/climate/change/about/extremes.shtml. (Accessed 25 September 2020).

Basher, R.E., Zheng, X., 1998. Mapping rainfall fields and their ENSO variation in data-sparse tropical south-west Pacific Ocean region. Int. J. Climatol. 18 (3), 237–251. https://doi.org/10.1002/(SICI)1097-0088(19980315)18:3<237::AID-JOC250>3.0.CO;2.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324.

Brodley, C.E., Friedl, M.A., 1997. Decision tree classification of land cover from remotely sensed data. Remote Sens. Environ. 61 (3), 399–409. https://doi.org/10.1016/S0034-4257(97)00049-7.

Byun, K.Y., Lee, T.Y., 2012. Remote effects of tropical cyclones on heavy rainfall over the Korean peninsula - statistical and composite analysis. Tellus, Series A: Dynamic Meteorology and Oceanography 64 (1), 14983. https://doi.org/10.3402/tellusa.v64i0.14983.

Cerveny, R.S., Newman, L.E., 2000. Climatological relationships between tropical cyclones and rainfall. Mon. Weather Rev. 128 (9), 3329–3336. https://doi.org/10.1175/1520-0493(2000)128<3329:CRBTCA>2.0.CO;2.

Chand, S.S., Walsh, K.J.E., 2009. Tropical cyclone activity in the Fiji region: spatial patterns and relationship to large-scale circulation. J. Clim. 22 (14), 3877–3893. https://doi.org/10.1175/2009JCLI2880.1.

Chand, S.S., Walsh, K.J.E., 2010. The influence of the Madden-Julian oscillation on tropical cyclone activity in the Fiji region. J. Clim. 23 (4), 868–886. https://doi.org/10.1175/2009JCLI3316.1.

Chand, S.S., Dowdy, A., Bell, S., Tory, K., 2020. A Review of South Pacific Tropical Cyclones: Impacts of Natural Climate Variability and Climate Change. Springer Climate, pp. 251–273. https://doi.org/10.1007/978-3-030-32878-8_6.

Chen, X., Wu, L., Zhang, J., 2011. Increasing duration of tropical cyclones over China. Geophys. Res. Lett. 38 (2), L02708. https://doi.org/10.1029/2010GL046137.

Chu, J., Sampson, C.R., Levine, A.S., Fukada, E., 2002. The joint typhoon warning center tropical cyclone best-tracks, 19452000, Ref. NRL/MR/754002. http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/best_tracks/TC_bt_report.html, 16-21.

Clark, T.E., 2004. Can out-of-sample forecast comparisons help prevent overfitting? J. Forecast. 23 (2), 115–139. https://doi.org/10.1002/for.904.

CSIRO Australian Bureau of Meteorology and SPREP, 2015. Climate in the Pacific: a regional summary of new science and management tools. In: Pacific-Australia Climate Change Science and Adaptation Planning Program Summary Report.

Deo, Anil, Walsh, Kevin J.E., 2017. Evaluation of TRMM multi-satellite precipitation analysis during the passage of tropical cyclones over Fiji. Journal of Southern Hemisphere Earth Systems Science 66 (4), 442–456. https://doi.org/10.22499/3.6604.005.

Deo, A., Chand, S.S., Ramsay, H., Holbrook, N.J., McGree, S., Magee, A., Bell, S., Titimaea, M., Haruhiru, A., Malsale, P., Mulitalo, S., Daphne, A., Prakash, B., Vainikolo, V., Koshiba, S., 2021. Tropical cyclone contribution to extreme rainfall over southwest Pacific Island nations. Clim. Dynam. https://doi.org/10.1007/s00382-021-05680-5.

DEWCE, T.T.-, WMO, 2016. Guidelines on the Definition and Monitoring of Extreme Weather and Climate Events: Draft. *Version - First Review by TT-DEWCE December 2015*. Avialable at: https://www.wmo.int/pages/prog/wcp/ccl/opace/opace2/documents/DraftversionoftheGuidelinesontheDefinitionandMonitoringofExtremeWeatherandClimateEvents.pdf. (Accessed 25 September 2020).

Diamond, H.J., Renwick, J.A., 2015. The climatological relationship between tropical cyclones in the southwest pacific and the southern annular mode. Int. J. Climatol. 35 (4), 613–623. https://doi.org/10.1002/joc.4007.

Diamond, H.J., Lorrey, A.M., Knapp, K.R., Levinson, D.H., 2012. Development of an enhanced tropical cyclone tracks database for the southwest Pacific from 1840 to 2010. Int. J. Climatol. 32 (14), 2240–2250. https://doi.org/10.1002/joc.2412.

Fan, R., Zhong, M., Wang, S., Zhang, Y., Andrew, A., Karagas, M., Chen, H., Amos, C.I., Xiong, M., Moore, J.H., 2011. Entropy-based information gain approaches to detect and to characterize gene-gene and gene-environment interactions/correlations of complex diseases. Genet. Epidemiol. 35 (7), 706–721. https://doi.org/10.1002/gepi.20621.

Fiji Meteorological Service, 2017. Fiji Annual Climate Summary 2016. Available at: https://www.met.gov.fj/index.php?page=climatedataold#2016annualSum2018.09.25 00.41.46.pdf. (Accessed 21 November 2019).

Flanders Marine Institute, 2018. *Maritime Boundaries Geodatabase: Maritime Boundaries and Exclusive Economic Zones (200NM), Version 10*. Available at: https://www.marineregions.org/. (Accessed 24 July 2019).

Fogt, R.L., Marshall, G.J., 2020. The southern annular mode: variability, trends, and climate impacts across the southern hemisphere. WIREs Climate Change 11 (4), 1–24. https://doi.org/10.1002/wcc.652.

Folland, C.K., Renwick, J.A., Salinger, M.J., Mullan, A.B., 2002. Relative influences of the interdecadal pacific oscillation and ENSO on the south pacific convergence zone. Geophys. Res. Lett. 29 (13), 21–24. https://doi.org/10.1029/2001GL014201.

Gaffney, S., 2004. Probabilistic Curve-Aligned Clustering and Prediction with Regression Mixture Models. University of California, Irvine.

Gaffney, S.J., Robertson, A.W., Smyth, P., Camargo, S.J., Ghil, M., 2007. Probabilistic clustering of extratropical cyclones using regression mixture models. Clim. Dynam. 29 (4), 423–440. https://doi.org/10.1007/s00382-007-0235-z.

Gero, A., Méheux, K., Dominey-Howes, D., 2011. Integrating disaster risk reduction and climate change adaptation in the Pacific. Clim. Dev. 3 (4), 310–327. https://doi.org/10.1080/17565529.2011.624791.

Gong, D., Wang, S., 1999. Definition of antarctic oscillation index. Geophys. Res. Lett. 26 (4), 459–462. https://doi.org/10.1029/1999GL900003.

Haylock, M., Nicholls, N., 2000. Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910-1998. Int. J. Climatol. 20 (13), 1533–1541. https://doi.org/10.1002/1097-0088(20001115)20:13<1533::AID-JOC586>3.0.CO;2-J.

Hernández Ayala, J.J., Matyas, C.J., 2016. Tropical cyclone rainfall over Puerto Rico and its relations to environmental and storm-specific factors. Int. J. Climatol. 36 (5), 2223–2237. https://doi.org/10.1002/joc.4490.

Ho, M., Kiem, A.S., Verdon-Kidd, D.C., 2012. The southern annular mode: a comparison of indices. Hydrol. Earth Syst. Sci. 16 (3), 967–982. https://doi.org/10.5194/hess-16-967-2012.

Hsieh, W.W., Tang, B., 1998. Applying neural network models to prediction and data analysis in Meteorology and oceanography. Bull. Am. Meteorol. Soc. 79 (9), 1855–1870. https://doi.org/10.1175/1520-0477(1998)079<1855:ANNMTP>2.0.CO.

Huang, B., Thorne, P.W., Banzon, V.F., Boyer, T., Chepurin, G., Lawrimore, J.H., Menne, M.J., Smith, T.M., Vose, R.S., Zhang, H.-M., 2017a. Extended Reconstructed Sea surface temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons. J. Clim. 30 (20), 8179–8205. https://doi.org/10.1175/jcli-d-16-0836.1.

Huang, B., Thorne, P.W., Banzon, V.F., Boyer, T., Chepurin, G., Lawrimore, J.H., Menne, M.J., Smith, T.M., Vose, R.S., Zhang, H.-M., 2017b. NOAA Extended Reconstructed Sea Surface Temperature (ERSST), Version 5. NOAA National Centers for Environmental Information. https://doi.org/10.7289/V5T72FNM. (Accessed 25 November 2017).

Khouakhi, A., Villarini, G., Vecchi, G.A., 2017. Contribution of tropical cyclones to rainfall at the global scale. J. Clim. 30 (1), 359–372. https://doi.org/10.1175/JCLI-D-16-0298.1.

Knapp, K.R., Kruk, M.C., Levinson, D.H., Diamond, H.J., Neumann, C.J., 2010. The international best track archive for climate stewardship (IBTrACS) unifying tropical cyclone data. Bull. Am. Meteorol. Soc. 91 (3), 363376. https://doi.org/10.1175/2009BAMS2755.1.

Kousky, V.E., Higgins, R.W., 2007. An alert classification system for monitoring and assessing the ENSO cycle. Weather Forecast. 22 (2), 353–371. https://doi.org/10.1175/waf987.1.

Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer, New York, NY. https://doi.org/10.1007/978-1-4614-6849-3.

Kuleshov, Y., McGree, S., Jones, D., Charles, A., Cottrill, A., Prakash, B., Atalifo, T., Nihmei, S., Seuseu, F.L.S.K., 2014. Extreme weather and climate events and their impacts on island countries in the western pacific: cyclones, floods and droughts. Atmos. Clim. Sci. 4 (5), 803–818. https://doi.org/10.4236/acs.2014.45071.

Kumar, R., Stephens, M., Weir, T., 2014. Rainfall trends in Fiji. Int. J. Climatol. 34 (5), 1501–1510. https://doi.org/10.1002/joc.3779.

Lantz, B., 2019. Machine Learning with R: Expert Techniques for Predictive Modeling, third ed. Packt Publishing. Avilable at: https://www.packtpub.com/product/machine-learning-with-r-third-edition/9781788295864. (Accessed 19 August 2020).

Lonfat, M., Marks, F.D., Chen, S.S., 2004. Precipitation distribution in tropical cyclones using the Tropical Rainfall Measuring Mission (TRMM) microwave imager: a global perspective. Mon. Weather Rev. 132 (7), 1645–1660. https://doi.org/10.1175/1520-0493(2004)132<1645:PDITCU>2.0.CO;2.

Magee, A.D., Verdon-Kidd, D.C., 2018. On the relationship between Indian Ocean sea surface temperature variability and tropical cyclogenesis in the southwest Pacific. Int. J. Climatol. 38, E774–E795. https://doi.org/10.1002/joc.5406.

Magee, A.D., Verdon-Kidd, D.C., Kiem, A.S., 2016. An intercomparison of tropical cyclone best-track products for the southwest Pacific. Nat. Hazards Earth Syst. Sci. 16 (6), 1431–1447. https://doi.org/10.5194/nhess-16-1431-2016.

Magee, A.D., Verdon-Kidd, D.C., Diamond, H.J., Kiem, A.S., 2017. Influence of ENSO, ENSO Modoki, and the IPO on tropical cyclogenesis: a spatial analysis of the southwest Pacific region. Int. J. Climatol. 37, 1118–1137. https://doi.org/10.1002/joc.5070.

Marshall, G.J., 2003. Trends in the southern annular mode from observations and reanalyses. J. Clim. 16 (24), 4134–4143. https://doi.org/10.1175/1520-0442(2003)016<4134:Titsam>2.0.Co;2.

Mataki, M., Koshy, K.C., Lal, M., 2006. Baseline climatology of viti Levu (Fiji) and current climatic trends. Pac. Sci. 60 (1), 49–68. https://doi.org/10.1353/psc.2005.0059.

McGree, S., Yeo, S.W., Devi, S., 2010. Flooding in the Fiji islands between 1840 and 2009. Risk Frontiers.

McKenzie, E., Prasad, B.C., Kaloumaira, A., 2005. *Economic Impact of Natural Disasters on Development in the Pacific: Economic Assessment Tools* University of the South Pacific and AusAid Report. AusAid, Canberra, Australia.

Nam, C.C., Park, D.S.R., Ho, C.H., Chen, D., 2018. Dependency of tropical cyclone risk on track in South Korea. Nat. Hazards Earth Syst. Sci. 18 (12), 3225–3234. https://doi.org/10.5194/nhess-18-3225-2018.

Ng, B., Walsh, K., Lavender, S., 2015. The contribution of tropical cyclones to rainfall in northwest Australia. Int. J. Climatol. 35 (10), 2689–2697. https://doi.org/10.1002/joc.4148.

Nguyen-Thi, H.A., Matsumoto, J., Ngo-Duc, T., Endo, N., 2012. A climatological study of tropical cyclone rainfall in Vietnam. Scientific Online Letters on the Atmosphere 8 (1), 41–44. https://doi.org/10.2151/sola.2012-011.

Nicholls, N., 1989. Sea Surface Temperatures and Australian Winter Rainfall. Journal of Climate 2 (9), 965–973. https://doi.org/10.1175/1520-0442(1989)002<0965:Sstaaw>2.0.Co;2.

Packt Editorial Staff, 2020. Brett Lantz on Implementing a Decision Tree Using C5.0 Algorithm in R. Available at: https://hub.packtpub.com/brett-lantz-on-implementing-a-decision-tree-using-c5-0-algorithm-in-r/. (Accessed 19 August 2020).

Park, M.S., Kim, M., Lee, M.I., Im, J., Park, S., 2016. Detection of tropical cyclone genesis via quantitative satellite ocean surface wind pattern and intensity analyses using decision trees. Remote Sens. Environ. 183, 205–214. https://doi.org/10.1016/j.rse.2016.06.006.

Quinlan, J.R., 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers Inc.

Rao, G.V., Macarthur, P.D., 1994. The SSM/I estimated rainfall amounts of tropical cyclones and their potential in predicting the cyclone intensity changes. Mon. Weather Rev. 122 (7), 1568–1574. https://doi.org/10.1175/1520-0493(1994)122<1568:TSERAO>2.0.CO;2.

Rogerson, P.A., 2011. Hypothesis testing and sampling. In: Statistical Methods for Geography. SAGE Publications, Ltd, pp. 43–64. https://doi.org/10.4135/9781849209953.

Rulequest Research, 2019. C5.0: an informal tutorial. Avaliable at: https://www.rulequest.com/see5-unix.html#XVAL. (Accessed 29 September 2020).

Saji, N.H., Goswami, B.N., Vinayachandran, P.N., Yamagata, T., 1999. A dipole mode in the tropical Indian Ocean. Nature 401 (6751), 360–363. https://doi.org/10.1038/43854.

Shannon, C.E., 1948. A mathematical theory of communication. Bell System Technical Journal 27 (4), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x, 623–656.

Sharma, K.K., Verdon-Kidd, D.C., Magee, A.D., 2020. Decadal variability of tropical cyclogenesis and decay in the southwest Pacific. Int. J. Climatol. 40 (5), 2811–2829. https://doi.org/10.1002/joc.6368.

Sharma, K.K., Magee, A.D., Verdon-Kidd, D.C., 2021. Variability of Southwest Pacific tropical cyclone track geometry over the last 70 years. Int. J. Climatol. 41 (1), 529–546. https://doi.org/10.1002/joc.6636.

Shepherd, J.M., Grundstein, A., Mote, T.L., 2007. Quantifying the contribution of tropical cyclones to extreme rainfall along the coastal southeastern United States. Geophys. Res. Lett. 34 (23) https://doi.org/10.1029/2007GL031694 n/a-n/a.

Stephens, M., Lowry, J.H., Ram, A.R., 2018. Location-based environmental factors contributing to rainfall-triggered debris flows in the Ba river catchment, northwest Viti Levu island, Fiji. Landslides 15 (1), 145–159. https://doi.org/10.1007/s10346-017-0918-4.

Terry, J.P., 2007. Tropical Cyclones: Climatology and Impacts in the South Pacific. Springer Science & Business Media, New York, NY. Springer.

Terry, J.P., Raj, R., 1999. Island environment and landscape responses to 1997 tropical cyclones in Fiji. Pac. Sci. 53 (3), 257–272.

Terry, J.P., McGree, S., Raj, R., 2004. The exceptional flooding on Vanua Levu island, Fiji, during tropical cyclone Ami in january 2003. J. Nat. Disaster Sci. 26 (1), 27–36. https://doi.org/10.2328/jnds.26.27.

Verdon, D.C., Franks, S.W., 2005. Indian Ocean sea surface temperature variability and winter rainfall: eastern Australia. Water Resour. Res. 41 (9) https://doi.org/10.1029/2004WR003845.

Visher, S., 1922. Tropical cyclones in Australia and the south pacific and Indian oceans. Mon. Weather Rev. 50 (6), 288295. https://doi.org/10.1175/1520-0493(1922)50<288:Tciaat>2.0.Co;2.

Wang, X.L., 2008a. Accounting for autocorrelation in detecting mean shifts in climate data series using the penalized maximal t or F test. Journal of Applied Meteorology and Climatology 47 (9), 2423–2444. https://doi.org/10.1175/2008JAMC1741.1.

Wang, X.L., 2008b. Penalized maximal F test for detecting undocumented mean shift without trend change. J. Atmos. Ocean. Technol. 25 (3), 368–384. https://doi.org/10.1175/2007JTECHA982.1.

Wang, X.L., Feng, Y., 2013. *RHtestsV4 User Manual.* Climate Research Division, Atmospheric Science and Technology Directorate, Science and Technology Branch. Environment Canada (28 pp). Available at: http://etccdi.pacificclimate.org/software.shtml. (Accessed 22 September 2020).

Wang, X.L., Chen, H., Wu, Y., Feng, Y., Pu, Q., 2010. New techniques for the detection and adjustment of shifts in daily precipitation data series. Journal of Applied Meteorology and Climatology 49 (12), 2416–2436. https://doi.org/10.1175/2010JAMC2376.1.

Waylen, P.R., Harrison, M., 2005. The coincidence of daily rainfall events in Liberia, Costa Rica and tropical cyclones in the Caribbean basin. Int. J. Climatol. 25 (12), 1665–1674. https://doi.org/10.1002/joc.1241.

World Meteorological Organization, 2017. WMO Guidelines on the Calculation of Climate Normals. WMO/TD-No. 1203, Geneva, Switzerland. World Meteorological Organisation.

Yeo, S., 2013. A Review of Flood Resilience in Fiji. International Conference on Flood Resilience: Experiences in Asia and Europe. September 2013.

Yeo, S.W., Blong, R.J., 2010. Fiji's worst natural disaster: the 1931 hurricane and flood. Disasters 34 (3), 657–683. https://doi.org/10.1111/j.1467-7717.2010.01163.x.

Zhang, T., Lin, W., Lin, Y., Zhang, M., Yu, H., Cao, K., Xue, W., 2019. Prediction of tropical cyclone genesis from mesoscale convective systems using machine learning. Weather Forecast. 34 (4), 10351049. https://doi.org/10.1175/WAF-D-18-0201.1.