Application of Multivariate Techniques for Spatial Drought Modelling using Satellite Rainfall Estimate

Anshuka Anshuka (✉ anshuka92@gmail.com)

University of Sydney <https://orcid.org/0000-0002-2434-8701>

Alexander JV Buzacott

University of Sydney

Floris van Ogtrop

University of Sydney

Research Article

Keywords: Forecasting Hydrological Extreme, Canonical Correlation Analysis, Effective Drought Index, Satellite-Derived Rainfall Estimates, Spatial drought maps

Posted Date: March 21st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1255720/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Monitoring hydrological extremes are essential for developing risk-mitigation strategies. One of the limiting factors for this is the absence of reliable on the ground monitoring networks that capture data on climate variables, which is highly evident in developing states such as Fiji. Fortunately, increasing global coverage of satellite-derived datasets facilitates the utilisation of this information to monitor dry and wet periods in sparse data regions. This study evaluated three global satellite rainfall datasets (CHIRPS, PERSIANN-CDR and CPC) for Fiji. All satellite products had reasonable correlations with station data, and CPC had the highest correlation with minimum error values. The Effective Drought Index (EDI), a useful index for understanding hydrological extremes, was then calculated. Thereafter, a canonical correlation analysis (CCA) was employed to forecast the EDI using sea surface temperature anomaly (SST) data. A high canonical correlation of 0.98 was achieved between the mean SST and mean EDI PCs, showing the influence of ocean-atmospheric interactions on precipitation regimes in Fiji. CCA was used to perform a hindcast and a short-term forecast. The training stage produced a coefficient of determinant (R^2) value of 0.83 and mean square error (MSE) of 0.11. The results in the testing stage for the forecast were more modest, with an R^2 of 0.45 and MSE of 0.26. This easy-to-implement system can be a helpful tool used by disaster management bodies to enact water restrictions, provide aid, and make informed agronomic decisions such as planting dates or extents.

1 Introduction

Island states normally suffer from a range of geological and hydrological hazards (Davis et al. 2013). Fiji, an island country is highly exposed to extreme events, with evidence indicating that many local communities are sensitive to various hydro-meteorological hazards such as cyclones, floods and drought (Anshuka et al. 2021). The 1997–1998 El Niño-induced drought cost the Fijian government approximately FJ\$1 million in emergency supplies. The country experienced a decline of 3.7% in the gross domestic product (GDP) due to the collapse of the sugar industry (Terry and Raj 1999). Hazard modelling can be a means to increase preparedness among the government and local communities to reduce risks.

However, models' development relies on a long series of continuous and homogeneous data for various parameters, often lacking in small island developing states (SIDS). The restricted data availability in Fiji may be due to a lack of instrumental hydrological records and equipment (Grimes et al. 2003; McAneney et al. 2017; McGree et al. 2010; Page et al. 2004). Although recorded radar and gauge data have advantages such as higher accuracy for point measurements, poor accessibility, lack of dense data networks, and uneven distribution of gauges are often a deterrent for carrying out studies using recorded data (Sun et al. 2018). Fortunately, satellite data can be used to overcome some of these challenges. Satellite products have many advantages, including broader spatial coverage, equidistant data points and availability at multiple time resolutions (Kidd 2001). Satellite products are widely applied in global and regional climate trend analysis and monitoring droughts and floods (Funk et al. 2015; Sun et al. 2018). The method of estimating rainfall by measuring Cloud –Cloud– duration, such as the Climate Hazards Group Infrared Precipitation with Station (CHIRPS) data and *Climate Prediction Centre* (CPC) US

unified precipitation, is claimed to be better for longer timescales, while passive microwave sensors such as the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks–Climate Data Record (PERSIANN-CDR) are better for instantaneous rain in well-defined regions (Kidd 2001). The appropriateness of using a satellite product depends on the regional topography, seasonal weather patterns, number of heavy rain events, and influence of orographic effect (Xu et al. 2015). Importantly, where possible, satellite data should be checked for reliability and bias to test the suitability of the products in a given geographical location (Anshuka 2019).

The Fiji Meteorological Services presently uses the Standardised Precipitation Index (SPI) to monitor the country's wet and dry climate periods (Pahalad and McGree 2002). Several studies have used SPI for drought-related research on Fiji (Deo 2011; Rhee and Yang 2018b). Deo (2011), performed a Mann Kendal test using SPI and concluded that most stations indicated negative values depicting reduction in rainfall. Recently, SPI has also been successfully used to conduct drought monitoring in Fiji using a hybrid approach of multi-model ensemble seasonal climate forecast and machine learning method (Rhee and Yang 2018b). However, one of the drawbacks in SPI calculation is that it assumes stationary climate conditions (Kim et al. 2009). Additionally, a common practice has been to use timescales on a monthly or annual basis (Heim Jr 2002). Although the use of monthly scales in other drought indices yields a good judgement of drought for long-term planning, it provides minimum information on the status within a short time span. SPI is known to cause errors when calculated at shorter timescales (1-month) (Wu et al. 2007). Therefore, in this study, we used the Effective Drought Index (EDI) by Byun and Wilhite (1999), which takes into account the decay of water resources in the soil over time and employs a daily timescale giving a better indication of hydrological extremes (Kim et al. 2009). Studies testing different drought indices have concluded that EDI has superior performance in determining the early onset of drought (Jain et al. 2015; Morid et al. 2006). In addition, EDI's rationale of depreciation of precipitation has also been used to create a flood index, indicating the versatility of the index (Deo et al. 2018).

Large-scale sea surface temperature (SST) gradients are known to displace the South Pacific Convergence Zone (SPCZ) in the South West Pacific, consequently affecting precipitation regimes (Griffiths et al. 2003; McGree et al. 2014). Statistical techniques such as canonical correlation analysis (CCA) and principal component analysis (PCA) yield information on essential climate features and precipitation, such as El Niño–Southern Oscillation (ENSO), which is derived from SST (Barnett and Preisendorfer 1987). Correlation patterns in SST are one of the large-scale features driving the precipitation patterns in the Pacific and can be used in the area of hydrological extreme monitoring (Salinger et al. 2001). CCA has been successfully applied to predict precipitation from SST and generate forecasts (Chen et al. 2017; Rana et al. 2018).

Unlike floods, which are typically caused by point source locations such as rivers affecting an area at a time (Anshuka et al. 2022), droughts have a larger spatial extent, which is why it is essential to understand the spatial distribution of droughts (Ionita et al. 2012). CCA can be used to understand the spatial distribution of droughts through spatial mapping. Maps containing geographical information are an essential communication tool to the broader public and the stakeholders because of the ease of

presenting and understanding the information. Importantly, climatic variations are increasingly recognised in the Pacific, and traditional coping mechanisms are no longer sufficient to deal with the adversities ([Kuleshov et al. 2014](#)). Therefore, seasonal forecast maps can enhance the coping capacity and reduce the vulnerability of communities in SIDS ([Anshuka et al. 2020](#); [Mercer et al. 2007](#)). Drought modelling in Fiji has been previously done with time-series ARIMA model for point based forecasts ([Anshuka et al. 2018](#)).

Therefore, this research aimed to obtain spatial drought forecasts in Fiji by employing CCA using satellite-derived rainfall estimates.

2 Materials And Methods

Study Region

The study area, Fiji, is located at 17.7134° S and 178.0650° E. It occupies 18,270 square kilometres, with a maximum elevation of 1324 metres ([Neall and Trewick 2008](#)). The nation consists of 322 islands with two main islands, Viti Levu and Vanua Levu. The region falls within the tropics and therefore experiences wet and dry seasons. The mean annual temperature is 25 degrees Celsius ([Neall and Trewick 2008](#)), and the mean precipitation ranges from 1676 to 3544 mm at different locations ([Kumar et al. 2014](#)). The Southern winter season observes the movement of the SPCZ northward, away from Fiji; therefore, rainfall during winter results from regional features such as orographic lift. Fiji consists of mountainous topography with prevailing south-easterly trade winds, resulting in an orographic lift ([Terry 2005](#)). This is highly evident in the dry season, whereby the leeward side (Western and Northern Divisions) receives 20% of the annual rainfall. In comparison, the windward side (Central Division) receives 33% of the yearly rainfall ([Terry 2005](#)). Most of the rain-fed agriculture is located in the Western and Northern Divisions, including the sugar industry. The intra-seasonal movement of the SPCZ northward and southward results in a warm-phase El Niño and a cool-phase La Nina event, respectively, causing unexpected rainfall anomalies ([Juillet-Leclerc et al. 2006](#)). Precipitation overall in the country is driven by the South Pacific Convergence Zone, El Niño–Southern Oscillation (ENSO) and tropical cyclones, which are accompanied by heavy rain events during the wet season.

Data

Satellite Rainfall Data

The satellite precipitation data were selected based on the start date of the record and the spatial resolution (Table 1). Because Fiji occupies a small area, we needed to ensure that products with satisfactory resolutions were selected. Rainfall can be highly variable in different areas despite being in the same zone; therefore, the most miniature spatial scale was chosen for validation ([Xu et al. 2015](#)).

Following the methods of [Dembélé and Zwart \(2016\)](#), satellite data were extracted using the point to pixel method based on the precise observation station coordinates.

Table 1 Information summary of the satellite precipitation products

Product	Spatial Resolution	Temporal Resolution	Start Date
PERSIANN-CDR	0.25°	Daily	1983
CPC-Precipitation	0.5°	Daily	1979
CHIRPS	0.05°	Daily	1981

PERSIANN-CDR: PERSIANN-CDR data ([Ashouri et al. 2015](#); [Hsu et al. 1997](#)) were retrieved from <http://chrsdata.eng.uci.edu/>. The dataset has a resolution of 0.25° with a spatial extent of 60° S to 60° N, and is available at hourly, monthly and daily time steps. One of the reasons for selecting this dataset was its long record, starting from 1983 to the near-present, making it ideal for the calculation of the hydrological indices. This product uses GEO infrared imaging technology and microwave sensors to compute an estimate of rainfall at each grid pixel ([Ashouri et al. 2015](#)). Since its development, the PERSIANN-CDR has been used to carry out streamflow modelling ([Liu et al. 2017](#)), monitor rainfall patterns ([Arvor et al. 2017](#)) and perform drought modelling ([Rhee and Yang 2018a](#)).

CPC unified data: *CPC US unified precipitation data (referred to as CPC data in the paper) were obtained from the NOAA/OAR/ESRL Physical Sciences Division (PSD), Boulder, Colorado, United States, and their website <https://www.esrl.noaa.gov/psd/>*. The temporal coverage of the CPC daily dataset started from January 1979 to the present time with a spatial resolution of 0.5°. The dataset is created using the optimal interpolation objective analysis technique derived by [Chen et al. \(2008\)](#). The dataset was validated using 16,000 global gauge stations but now uses reports from 30,000 stations ([Sun et al. 2018](#)).

CHIRPS: This dataset, starting from 1981 with a spatial extent ranging from 50° S to 50° N, was retrieved from <http://chg.ucsb.edu/data/>. CHIRPS data have a resolution of 0.05° and use thermal infrared bands and in situ station data to create the final gridded product. The satellite information accounts for sparsely gauged locations, and precipitation estimates are available on daily, pentad and monthly scales ([Funk et al. 2015](#)). CHIRPS data have not been explored before for Pacific island topographies; however, they have been tested in other areas such as the Caribbean. Evaluation of the CHIRPS in Mozambique has indicated an excellent ability to detect rainfall during hurricanes ([Toté et al. 2015](#)), which is especially useful given Fiji's proneness to flooding during cyclones.

In Situ Rainfall Data

The daily and monthly precipitation dataset was obtained from Fiji Meteorological Services, which operates over 20 manual and automatically telemetered rain gauges over the Fiji Islands. Six high-quality stations were used in the current evaluation: Labasa, Nadi, Navua, Ba, Matei and Nausori (Fig. 1). These meteorological stations were selected considering the variation in local climate, topography and orographic effects observed in Fiji (Table 2). The Central Division stations, Navua and Nausori, are in the windward region characterised mainly by wet weather conditions (See Table 2 and Fig. 2 for Station climatology). The Western Division is on the drier side of the island, represented by the stations Rarawai and Nadi. Labasa station is the main station on the island of Vanua Levu, while Matei is a town situated on Taveuni (an outer smaller island); both serve to provide datasets for the Northern Division of Fiji. The data were checked for homogeneity, and missing values were replaced with the long-term monthly averages. The missing data ranged from 0 – 5 points for all the stations, while Matei had 31 missing data points.

Table 2 Station climatology for monthly rainfall in Fiji for the wet and dry seasons

Station	Longitude	Latitude	Wet Mean (mm)	Wet Max (mm)	Wet Min (mm)	Dry Mean (mm)	Dry Max (mm)	Dry Min (mm)
Nausori	178.5591	−18.0464	313.9	1216.8	27.9	175.1	913.6	12.1
Nadi	177.4439	−17.7543	246.6	1180.6	3.6	71.5	341.8	0.0
Labasa	179.3396	−16.4689	298.7	1068.2	15.2	77.7	490.9	0.0
Matei	179.8667	−16.6833	299.2	918.1	23.1	151.4	728.4	7.6
Rarawai	177.6814	−17.5564	282.6	1907.8	0.0	73.8	402.4	0.0
Navua	178.1700	−18.2186	343.1	843.6	78.8	237.1	593.6	41.1

Sea Surface Temperature Data

The Kaplan SST data used in this analysis were retrieved from <https://climatedataguide.ucar.edu/climate-data/kaplan-sea-surface-temperature-anomalies>. The dataset has a temporal coverage from 1856 and a global spatial extent with a spatial resolution of 5.0 x 5.0 units. The data are produced by various steps, such as Empirical Orthogonal Functions (EOF) projection, optimal interpolation (OI), the Kalman Filter (KF) forecast, KF analysis and an optimal smoother (OS), which ultimately minimise the number of missing data and error values (Reynolds and Smith 1994). A bounding box was created in Arc GIS, ranging from 150° E to 80° W between the Tropic of Cancer and the Tropic of Capricorn, which resulted in 259 grid points. This ensured that the SST would be extracted for the region surrounding Fiji. The series length extracted was the same for the EDI and SST, from 1980 to 2017 (Fig. 3). The data were checked for homogeneity, and grids with missing data were omitted.

Methods

Satellite Rainfall Validation

Continuous verification statistics were applied to test the satellite-derived precipitation products (Jolliffe and Stephenson 2003) for the six stations (Fig. 1). The following performance measures were used to evaluate the products: coefficient of determination (R^2), root-mean-square error (RMSE) and Lin's concordance. The R^2 measures the extent of association between the actual recorded data and the satellite data in this case. The RMSE assesses the variance of errors independently and indicates the inconsistency between the recorded and the satellite values, with lower values indicating minimum error difference between the recorded and satellite data (Adamowski et al. 2012). Lin's concordance created by Nickerson (1997) is calculated based on the degree to which two variables fall on the 45° line that passes through the origin. The calculation was undertaken in R using the DescTools package (Signorell et al. 2016). Lin's concordance has values between 0 and 1 without a unit, and a higher value indicates a strong relationship similar to R^2 . The unit for RMSE is millimetres. The daily data were aggregated to monthly and tested for the identified performance metrics. A complimentary check was also made on the Nadi station to test for differences in the wet and dry seasons, respectively for the CPC product (Results shown in Table 5). The eyeball method is a subjective form of verification (Ebert 2007). In this study, we made scatter plots of observed versus satellite data for Nadi station and created kernel density plots to determine instances of overestimation and underestimation.

Index Calculation

CPC data were used for further analyses because these data rendered the best correlation results with local station data (Table 4). Therefore, following point extraction, the dataset was checked for quality and homogeneity. The zoo package (Zeileis and Grothendieck 2005) in R was used to carry out linear interpolation in the case of missing precipitation data. Point linear interpolation is calculated using the values measured on either side of the missing data (North and Livingstone 2013). The drought index was calculated based on the principles identified by Byun and Wilhite (1999) for each of the grid cells retrieved from the CPC data. The daily Effective Drought Index (EDI) values were aggregated into monthly EDI values because the climate index data were available at monthly timescales (Fig. 4).

Multivariate Analysis

CCA successfully establishes interrelations to identify linear combinations between more than one response and explanatory variable (Ouarda et al. 2001), while PCA employs the same approach on a single multi-dimensional dataset. The relationships between the weights of the response and explanatory data are known as loading patterns, and these can be used to visualise potential physical mechanisms in large-scale processes (Shabbar and Barnston 1996). The multivariate analysis was performed according to the guidelines provided by Wilks (2011). SST was used as the explanatory variable and the gridded EDI as the response variable, from which a probabilistic forecast was generated. If a dataset $x(t)$ leads to $y(t)$, then CCA can be used to forecast the y variable by using the lagged x variable in a gridded form. A lag

of 2 months was introduced in the data matrix between the EDI and the SST; consequently, a forecast of EDI could be provided in short term (2 months ahead). Short-term forecasts are accurate and adequate to assist managers and farmers in making sound management and agronomic decisions (Anshuka et al. 2019). As a first step, the data were standardised and centred to remove any trend. This ensures that the principal components (PCs) are mutually orthogonal (Westra and Sharma 2010). Previous studies have also applied pre-processing of data using empirical orthogonal analysis (Landman and Mason 1999; Tang et al. 2000; Yu et al. 1997). This has many advantages, such as reducing dimensionality and noise in the data and giving equal opportunity for predictors to contribute to forecasting models (Shabbar and Barnston 1996).

We selected the first 10 PCs from the SST and EDI to perform the CCA. Forecast evaluation was performed by splitting the data into two parts, whereby a hindcast was generated, which entailed model training on data from 1980 to 2004, and the forecast was tested using the data from 2005 to 2017. The forecast skill was verified using the coefficient of determination (R^2) and mean square error (MSE). A spatial plot was generated to determine how well the forecast and the hindcast predicted the EDI classes, as shown in Table 3.

Table 3 EDI drought classes

Value	Category	EDI
3	Extremely Wet	≥ 2
2	Very Wet	1.5 to 1.99
1	Moderately Wet	1.0 to 1.49
0	Normal	-0.99 to 0.99
-1	Moderately Dry	-1.0 to -1.49
-2	Severely Dry	-1.5 to -1.99
-3	Extremely Dry	≤ -2

3 Results

Satellite Rainfall Validation

First, the satellite data products were validated for the Fiji region. From the results (Table 4), the best satellite product was selected, which was used to calculate the EDI.

Table 4 Summary statistics for CPC, CHIRPS and PERSIANN-CDR for monthly precipitation with observed station measurements in Fiji

Dataset	Station	Correlation (R^2)	RMSE	Lin's Concordance
CPC	Nausori	0.8104	80.2851	0.8356
	Nadi	0.8473	75.0368	0.8770
	Labasa	0.5251	129.2369	0.6466
	Matei	0.5016	132.4340	0.5746
	Rarawai	0.6948	125.0123	0.7423
	Navua	0.4514	154.4701	0.5139
CHIRPS	Nausori	0.6720	87.5506	0.8182
	Nadi	0.8430	69.7970	0.8945
	Labasa	0.6891	102.4270	0.7949
	Matei	0.4023	150.9371	0.4501
	Rarawai	0.7066	118.2916	0.7752
	Navua	0.3724	141.8225	0.5818
PERSIANN-CDR	Nausori	0.6108	102.8872	0.7350
	Nadi	0.7817	77.6224	0.8607
	Labasa	0.7118	99.0153	0.7857
	Matei	0.5886	113.7052	0.6676
	Rarawai	0.6739	115.5148	0.7706
	Navua	0.3348	165.5948	0.4801

The CPC had the best results for most stations (Table 4). Nadi station had the highest correlation of 0.85, and respectively a high Lin's concordance of 0.88 with the lowest RMSE value of 75mm. All the datasets had the same best-correlated and poorly correlated stations; that is, Nadi had the best performance in all three datasets. Similarly Navua had the poorest correlation in the three datasets. There was a marginal difference in the performance measures for CPC and CHIRPS data. Similarly, the PERSIANN-CDR data, although having the lowest performance of the three datasets, still had modest results. Navua had the weakest correlation and the highest error, followed by Matei station.

The density plot (Fig. 5) indicated that the distribution was skewed to the right; most precipitation values ranged from 0 to 200 mm. Out of the three datasets, the CPC most closely matched the distribution of the gauge data. The datasets showed instances of overestimation, especially for low rainfall and, to a lesser extent, underestimation of heavy rainfall events. However, not all events were captured; notably, some of the high rainfall events were under predicted. None of the datasets was able to detect extreme heavy rainfall events, but the CPC was able to follow the gauge data closely up to values as high as 1000 mm.

Table 5 Summary statistics for Nadi station versus CPC data in the wet and dry seasons

	Correlation	RMSE	Lin's Concordance
Wet Season	0.7878	101.6735	0.8243
Dry Season	0.8237	30.3896	0.8870

The scatterplot for the Nadi station (Fig. 6) shows that CPC data represents the rain gauge data well. Additionally, validation for Nadi station was performed separately by considering the two seasons in Fiji: the wet season including January, February, March, April, November and December and the dry season including May, June, July, August, September and October. Data validation for the wet and dry seasons at Nadi showed that the correlations were higher in the dry season with low error values (Table 5). Based on satellite validation performance, CPC data was selected to calculate the Effective Drought Index.

Spatial Drought Modelling

Principal Component Analysis

A principal component analysis was performed on EDI and SST. The first five Principal Components (PCs) explained approximately 86% of the variation within the SST data, while the first five PCs explained 98% of the variation in the EDI data (Table 6). The first 10 PCs were retained for further analysis.

Table 6 Variation for the first five PCs performed on SST and EDI.

	PC1	PC2	PC3	PC4	PC5
SST					
Standard deviation	5.99	3.50	2.38	1.83	1.64
Proportion of variance	0.52	0.18	0.08	0.05	0.04
Cumulative proportion	0.52	0.69	0.77	0.82	0.86
EDI					
Standard deviation	2.81	0.84	0.73	0.60	0.41
Proportion of variance	0.80	0.07	0.05	0.04	0.02
Cumulative proportion	0.80	0.87	0.92	0.96	0.98

Canonical Correlation Analysis

CCA was performed on the leading principal components derived from Table 6. The first canonical variate pair shows that 98.6% of the variation in the EDI was explained by the variation in the SST (Table 7), indicating a strong relationship between the two variables.

Table 7 Pearson-product-moment correlations between the pairs of canonical variates of EDI and SST

Pair	Canonical Correlation
1	0.9860
2	0.9620
3	0.8753
4	0.8484
5	0.7076
6	0.6284
7	0.5942
8	0.4372
9	0.2376
10	0.0009

Spatial Modelling

The predictive skill of the model was assessed for both the hindcast and the forecast (Fig. 6). The hindcast was the model training stage (Fig. 7), which obtained an R^2 score of 0.83 and an MSE value of 0.12. The model was tested in the forecasting stage (Fig. 7) with a noted decrease in the R^2 skill and an increase in the MSE value by almost 50%. The colour sequence revealed that class error predictions in the hindcast (Fig. 8) and the forecast (Fig. 9) were generally well predicted. The grid cells that fall in the Central Division, mainly affected by orographic rainfall, were not correctly predicted in both the training and the testing stages. The time series plot (Fig. 7) shows instances of under-prediction in the testing stage and, to a lesser extent, in the training stage. Comparisons between observed mean EDI showed an increase in precipitation trend across the Fiji island group. In 2004, several dry events were observed; however, in 2017, no dry conditions were observed across Fiji.

4 Discussion

Satellite Rainfall Verification

All the satellite-derived precipitation products were reasonably well correlated with the observed precipitation data and, therefore can be used in further hydrological studies. The CPC and CHIRPS gridded precipitation products showed high correlations with observed data, except for Navua and Matei stations (Table 4). The high correlation is likely due to the data products being calibrated using the same local rain gauge data we used to verify the products. Specifically, in CPC, gauge reports are collected and compared with historical records, other nearby stations, numerical forecasts and subsequently created

into a field of precipitation by factoring for orographic effect (Xie et al. 2010). Regardless, the quality of the gridded data is partially dependent on the availability of data on the ground, as poor availability on the ground will result in a poor satellite-derived product (Toté et al. 2015). Therefore, it may be helpful to identify gauges not used in the calibration process to verify the products as future research.

All three data sources showed overestimation and underestimation of the frequency of rainfall volumes as indicated by the density plot (Fig. 4). Previous studies have found similar results with various satellite products, whereby overestimation of low rainfall events and underestimation of high rainfall events were reported (Kubota et al. 2009; Nair et al. 2009; Paredes-Trejo et al. 2017; Toté et al. 2015). Gridded datasets are derived from gauge-based products that undergo interpolation, resulting in smoother extreme values (Sun et al. 2018). This is an important consideration; in drought monitoring studies, one needs to ensure that the satellite-derived product does not overestimate rainfall; similarly, models that underestimate rainfall may not be suitable for flood monitoring studies (Toté et al. 2015). Again, the seasonal correlations for Nadi were better in the dry season than in the wet season (Table 5). This may result in better forecasts for the dry season than those for the wet season in hydrological-modelling studies. This could also be a possible explanation for low correlations achieved for Navua, which is the wettest station in the country with a mean annual rainfall of 3573 mm. Because performances are generally better in the dry season than in the wet season, the datasets, therefore, are less able to accurately make estimates for stations that exhibit extremely wet characteristics, such as Navua. A bias correction specific to the study site can also be undertaken for low rainfall or high rainfall, depending on the focus of the study (Vernimmen et al. 2012; Yin et al. 2008).

The satellite products use algorithms that cannot be equally applied to different terrains. The CHIRPS's cloud-top temperature method of rain detection may fail to identify warm orographic clouds, which rise higher than other rain/non-rain producing clouds in mountainous and coastal regions (Dinku et al. 2010). Matei, situated in Taveuni (an outer island in Fiji), comprises mountainous topography as it houses Mount Uluigalau (situated 1214 m above sea level and the second-highest mountain in Fiji), and this may explain the poor results of Matei station. Similar results were reported by Kubota et al. (2009), whereby poor results over snow-covered mountains and small islands were reported. In contrast, a comparison between TAMSAT and the CPC for Southern Africa showed better outcomes for the CPC over mountainous regions due to adjustment of the final product for orographic rainfall (Thorne et al. 2001). This could also explain the better overall performance for CPC data, which considers rainfall caused by the orographic effect in parts of the Fiji island group. While CHIRPS was marginally less well correlated with local rainfall data, it has a finer spatial resolution, and so it or potentially a hybrid between the products should also be investigated.

Spatial Drought Modelling

The analyses revealed potential to develop spatial monitoring tools using multivariate techniques with reasonable accuracy. An approach that identifies physical relationships between precipitation and

climate drivers can yield helpful information to understand the risks better (Ntale et al. 2003; Shabbar and Barnston 1996; Singh et al. 2012). The accuracy of the forecast through CCA depends on the feature responsible for rain events in a given geographic location. The high correlation between the first canonical pair suggests a significant influence of the SST on precipitation regimes in Fiji. Therefore, it is also expected that the rainfall is well correlated to the ENSO patterns. This is in agreement with other studies that have previously shown the dominance of ENSO on climate in Fiji (Kostaschuk et al. 2001; Terry et al. 2001). For example, Fiji's 1997–1998 drought has been linked to a strong El Niño event (Terry and Raj 1999). However, it is important to consider that ENSO events do not occur every year, which may cause low predictability in the years when ENSO events are not recorded (Landman and Mason 1999). In the South-West Pacific, precipitation is also influenced by underlying climate drivers occurring at decadal scales (Folland et al. 2002; Salinger et al. 2001). In a European study by Ionita et al. (2012), long time series of data from the 1900s were used to show associated quasi-periodic trends for the self-calibrated Palmer Drought Severity Index (PDSI) and the Pacific Decadal Oscillation (PDO) and Atlantic Multi-decadal Oscillation. The study also noted a significant drop in the EDI every 10–15 years. In our study, the relatively short data series (37 years) was sufficient to capture shorter-scale oscillations such as ENSO, which occurs semi-periodically with a period of 2-7 years. Conversely, one phase of PDO can last from 20 to 30 years (Mantua and Hare 2002); therefore, more data are required to capture the effects of longer-scaled decadal phenomena.

The overall skill of the model is similar to that obtained by Landman and Mason (1999) and Chu and He (1994). The spatial drought variability between the period 1980–2004 and 2004–2017 revealed an increase in precipitation trend across Fiji. The forecasting model captured the increasing trend in precipitation in 2017, with few out-of-phase predictions. Boulanger et al. (2005) highlighted the importance of understanding non-linear relationships between precipitation and the SST for statistical forecasts. While techniques such as machine learning may outperform CCA, CCA has the added advantage to identify spatial patterns that can suggest physical climate-related processes such as ENSO, hence can be used with other techniques. For example, CCA models have been used to improve results in global climate models (GCMs) (Shabbar and Barnston 1996; Singh et al. 2012). Additionally, the CCA model forecasts are robust and easy to compute and implement, making CCA a good first choice option where limited expertise and no hydrological extreme early warning system exist, particularly in developing nations such as Small Island Developing States. The forecasts generated from CCA are skilful but not necessarily perfect. Therefore, the method defined here can be extended further according to farmers or stakeholders' targeted needs and can be incorporated through a reflective decision framework among the different parties (Moeller et al. 2008; van Ogtrop et al. 2014).

Inconsistencies in the results can also be attributed to the effect of topographical features (Lal et al. 2008). This was highlighted in an African-based study where exceptions in the dry season were noted due to moist westerly Congo air masses (Ntale et al. 2003). Similarly, in Fiji, the South Easterly trade winds are responsible for much of the rain in the south-eastern side of the island (Suva, Navua and Nausori) and less rain in the north-western side during the dry (winter) season is observed. The Western Division in Fiji has a lower correlation with the Southern Oscillation Index during the winter (dry season) (Walsh et al.

2001). This may indicate that winter precipitation has no systematic link to large-scale circulations controlled by SST mainly and is influenced by regional features such as orographic lift. Tropical cyclone activity can also affect forecasts as heavy rains associated with cyclones are not usually reflected in the SST (Yu et al. 1997). Therefore, SST alone may not be the best predictor of rainfall in Fiji, and features such as spatial atmospheric pressure data (Chu and He 1994) and ocean thermocline (Ruiz et al. 2006) warrant further investigation.

5 Conclusion

The results suggest that all satellite products reasonably correlate with observed rainfall for the Fiji region. However, CPC data gives the best correlation with the recorded rainfall data. Therefore, it is recommended that CPC data are used for future hydrological studies. The mean SST and annual mean EDI series in Fiji indicate an increasing trend in temperature and precipitation, respectively. The study has also identified a strong link between the SST and the EDI, with a correlation of 0.98. A CCA model shows reasonable skill in predicting EDI over the Fiji region; hence, this model can be used as a 2-month advanced spatial drought forecasting tool for Fiji. Future studies can incorporate sea level pressure and ocean thermocline as explanatory variables of the EDI to further improve the forecasts. The loading patterns can also be explored to assess the effect of major climate drivers in different parts of the island group in Fiji and the South-West Pacific region. In the absence of an operational early-warning system, this relatively easy-to-implement and straightforward system can be available locally in Fiji to various stakeholders.

Declarations

Compliance with ethical standards

Conflict of interest: The authors declare that they have no conflict of interest.

References

- Adamowski J, Fung Chan H, Prasher SO, Ozga-Zielinski B, Sliusarieva A (2012) Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada Water Resources Research 48
- Anshuka (2019) Forecasting Drought Indices using Machine learning and Spatial Modelling Approaches: An Application in Fiji University of Sydney eScholarship doi:<http://hdl.handle.net/2123/20582>
- Anshuka A, Buzacott AJV, Vervoort RW, van Ogtrop FF (2020) Developing drought index-based forecasts for tropical climates using wavelet neural network: an application in Fiji Theoretical and Applied Climatology 143:557-569 doi:10.1007/s00704-020-03446-3

- Anshuka A, Ogtrop F, Vervoort W (2018) Drought Modelling in Small Island Developing States: A Case Study in Fiji EGUGA:3251
- Anshuka A, van Ogtrop FF, Sanderson D, Leao SZ (2022) A systematic review of agent-based model for flood risk management and assessment using the ODD protocol Natural Hazards doi:10.1007/s11069-022-05286-y
- Anshuka A, van Ogtrop FF, Sanderson D, Thomas E, Neef A (2021) Vulnerabilities shape risk perception and influence adaptive strategies to hydro-meteorological hazards: A case study of Indo-Fijian farming communities International Journal of Disaster Risk Reduction 62:102401
doi:<https://doi.org/10.1016/j.ijdr.2021.102401>
- Anshuka A, van Ogtrop FF, Willem Vervoort R (2019) Drought forecasting through statistical models using standardised precipitation index: a systematic review and meta-regression analysis Natural Hazards 97:955-977 doi:10.1007/s11069-019-03665-6
- Arvor D, Funatsu BM, Michot V, Dubreuil V (2017) Monitoring Rainfall Patterns in the Southern Amazon with PERSIANN-CDR Data: Long-Term Characteristics and Trends Remote Sensing 9:889
- Ashouri H et al. (2015) PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies Bulletin of the American Meteorological Society 96:69-83
- Barnett T, Preisendorfer R (1987) Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis Monthly Weather Review 115:1825-1850
- Boulanger J-P, Leloup J, Penalba O, Rusticucci M, Lafon F, Vargas W (2005) Observed precipitation in the Paraná-Plata hydrological basin: long-term trends, extreme conditions and ENSO teleconnections Climate dynamics 24:393-413
- Byun H-R, Wilhite DA (1999) Objective quantification of drought severity and duration Journal of Climate 12:2747-2756
- Chen M, Shi W, Xie P, Silva VB, Kousky VE, Wayne Higgins R, Janowiak JE (2008) Assessing objective techniques for gauge-based analyses of global daily precipitation Journal of Geophysical Research: Atmospheres 113
- Chen X, You Q, Sielmann F, Ruan N (2017) Climate change scenarios for Tibetan Plateau summer precipitation based on canonical correlation analysis International Journal of Climatology 37:1310-1321
- Chu PS, He Y (1994) Long-range prediction of Hawaiian winter rainfall using canonical correlation analysis International Journal of Climatology 14:659-669

- Davis S, Anshuka A, Aoki K, Cook M, Duley S, Huff M, Logan C (2013) Managing risk and allure at volcanoes: In Hawaii: How close is too close? *Geojournal of Tourism and Geosites* 12:85-93
- Dembélé M, Zwart SJ (2016) Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa *International Journal of Remote Sensing* 37:3995-4014
doi:10.1080/01431161.2016.1207258
- Deo RC (2011) On meteorological droughts in tropical Pacific Islands: time-series analysis of observed rainfall using Fiji as a case study *Meteorological Applications* 18:171-180
- Deo RC, Adamowski JF, Begum K, Salcedo-Sanz S, Kim D-W, Dayal KS, Byun H-R (2018) Quantifying flood events in Bangladesh with a daily-step flood monitoring index based on the concept of daily effective precipitation *Theoretical and Applied Climatology* doi:10.1007/s00704-018-2657-4
- Dinku T, Connor SJ, Ceccato P (2010) Comparison of CMORPH and TRMM-3B42 over mountainous regions of Africa and South America. In: *Satellite Rainfall Applications for Surface Hydrology*. Springer, pp 193-204
- Ebert EE (2007) Methods for verifying satellite precipitation estimates. In: *Measuring precipitation from space*. Springer, pp 345-356
- Folland C, Renwick J, Salinger M, Mullan A (2002) Relative influences of the interdecadal Pacific oscillation and ENSO on the South Pacific convergence zone *Geophysical Research Letters* 29:21-21-21-24
- Funk C et al. (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes *Scientific Data* 2:150066 doi:10.1038/sdata.2015.66
- Griffiths G, Salinger M, Leleu I (2003) Trends in extreme daily rainfall across the South Pacific and relationship to the South Pacific Convergence Zone *International Journal of Climatology: A Journal of the Royal Meteorological Society* 23:847-869
- Grimes D, Coppola E, Verdecchia M, Visconti G (2003) A neural network approach to real-time rainfall estimation for Africa using satellite data *Journal of Hydrometeorology* 4:1119-1133
- Heim Jr RR (2002) A review of twentieth-century drought indices used in the United States *Bulletin of the American Meteorological Society* 83:1149-1166
- Hsu K-I, Gao X, Sorooshian S, Gupta HV (1997) Precipitation estimation from remotely sensed information using artificial neural networks *Journal of Applied Meteorology* 36:1176-1190
- Ionita M, Lohmann G, Rimbu N, Chelcea S, Dima M (2012) Interannual to decadal summer drought variability over Europe and its relationship to global sea surface temperature *Climate Dynamics* 38:363-377

- Jain VK, Pandey RP, Jain MK, Byun H-R (2015) Comparison of drought indices for appraisal of drought characteristics in the Ken River Basin Weather and Climate Extremes 8:1-11
- Jolliffe IT, Stephenson DB (2003) Forecast verification: a practitioner's guide in atmospheric science. John Wiley & Sons,
- Juillet-Leclerc A, Thiria S, Naveau P, Delcroix T, Le Bec N, Blamart D, Corregge T (2006) SPCZ migration and ENSO events during the 20th century as revealed by climate proxies from a Fiji coral Geophysical Research Letters 33
- Kidd C (2001) Satellite rainfall climatology: A review International Journal of Climatology: A Journal of the Royal Meteorological Society 21:1041-1066
- Kim D-W, Byun H-R, Choi K-S (2009) Evaluation, modification, and application of the Effective Drought Index to 200-Year drought climatology of Seoul, Korea Journal of hydrology 378:1-12
- Kostaschuk R, Terry J, Raj R (2001) Tropical cyclones and floods in Fiji Hydrological sciences journal 46:435-450
- Kubota T, Ushio T, Shige S, Kida S, Kachi M, Okamoto Ki (2009) Verification of high-resolution satellite-based rainfall estimates around Japan using a gauge-calibrated ground-radar dataset Journal of the Meteorological Society of Japan Ser II 87:203-222
- Kuleshov Y et al. (2014) Extreme weather and climate events and their impacts on island countries in the Western Pacific: cyclones, floods and droughts Atmospheric and Climate Sciences 4:803
- Kumar R, Stephens M, Weir T (2014) Rainfall trends in Fiji International Journal of Climatology 34:1501-1510
- Lal M, McGregor JL, Nguyen KC (2008) Very high-resolution climate simulation over Fiji using a global variable-resolution model Climate Dynamics 30:293-305
- Landman WA, Mason SJ (1999) Operational long-lead prediction of South African rainfall using canonical correlation analysis International Journal of Climatology: A Journal of the Royal Meteorological Society 19:1073-1090
- Liu X, Yang T, Hsu K, Liu C, Sorooshian S (2017) Evaluating the streamflow simulation capability of PERSIANN-CDR daily rainfall products in two river basins on the Tibetan Plateau Hydrology and Earth System Sciences (Online) 21
- Mantua NJ, Hare S (2002) Pacific-Decadal Oscillation (PDO) Encyclopedia of global environmental change 1:592-594

- McAneney J, van den Honert R, Yeo S (2017) Stationarity of major flood frequencies and heights on the Ba River, Fiji, over a 122-year record *International Journal of Climatology*
- McGree S et al. (2014) An updated assessment of trends and variability in total and extreme rainfall in the western Pacific *International Journal of Climatology* 34:2775-2791
- McGree S, Yeo SW, Devi S (2010) Flooding in the Fiji Islands between 1840 and 2009 *Risk Frontiers*
- Mercer J, Dominey-Howes D, Kelman I, Lloyd K (2007) The potential for combining indigenous and western knowledge in reducing vulnerability to environmental hazards in small island developing states *Environmental Hazards* 7:245-256
- Moeller C, Smith I, Asseng S, Ludwig F, Telcik N (2008) The potential value of seasonal forecasts of rainfall categories—case studies from the wheatbelt in Western Australia's Mediterranean region *Agricultural and Forest Meteorology* 148:606-618
- Morid S, Smakhtin V, Moghaddasi M (2006) Comparison of seven meteorological indices for drought monitoring in Iran *International Journal of Climatology: A Journal of the Royal Meteorological Society* 26:971-985
- Nair S, Srinivasan G, Nemani R (2009) Evaluation of multi-satellite TRMM derived rainfall estimates over a western state of India *Journal of the Meteorological Society of Japan Ser II* 87:927-939
- Neall VE, Trewick SA (2008) The age and origin of the Pacific islands: a geological overview *Philosophical Transactions of the Royal Society B: Biological Sciences* 363:3293-3308
- Nickerson CA (1997) A note on "A concordance correlation coefficient to evaluate reproducibility" *Biometrics*:1503-1507
- North RP, Livingstone DM (2013) Comparison of linear and cubic spline methods of interpolating lake water column profiles *Limnology and Oceanography: Methods* 11:213-224
- Ntale HK, Gan TY, Mwale D (2003) Prediction of East African seasonal rainfall using simplex canonical correlation analysis *Journal of Climate* 16:2105-2112
- Ouarda TB, Girard C, Cavadias GS, Bobée B (2001) Regional flood frequency estimation with canonical correlation analysis *Journal of Hydrology* 254:157-173
- Page CM et al. (2004) Data rescue in the southeast Asia and south Pacific region: challenges and opportunities *Bulletin of the American Meteorological Society* 85:1483-1489
- Pahalad J, McGree S Rainfall forecasting and its applications (Fiji Case Study). In: Case study presented as part of Theme, 2002.

- Paredes-Trejo FJ, Barbosa H, Kumar TL (2017) Validating CHIRPS-based satellite precipitation estimates in Northeast Brazil *Journal of arid environments* 139:26-40
- Rana S, Renwick J, McGregor J, Singh A (2018) Seasonal Prediction of Winter Precipitation Anomalies over Central Southwest Asia: A Canonical Correlation Analysis Approach *Journal of Climate* 31:727-741
- Reynolds RW, Smith TM (1994) Improved global sea surface temperature analyses using optimum interpolation *Journal of climate* 7:929-948
- Rhee J, Yang H (2018a) Drought Prediction for Areas with Sparse Monitoring Networks: A Case Study for Fiji Water (20734441) 10
- Rhee J, Yang H (2018b) Drought Prediction for Areas with Sparse Monitoring Networks: A Case Study for Fiji Water 10:788
- Ruiz JE, Cordery I, Sharma A (2006) Impact of mid-Pacific Ocean thermocline on the prediction of Australian rainfall *Journal of Hydrology* 317:104-122
- Salinger M, Renwick J, Mullan A (2001) Interdecadal Pacific oscillation and south Pacific climate *International Journal of Climatology* 21:1705-1721
- Shabbar A, Barnston AG (1996) Skill of seasonal climate forecasts in Canada using canonical correlation analysis *Monthly weather review* 124:2370-2385
- Signorell A, Aho K, Alfons A, Anderegg N, Aragon T (2016) DescTools: Tools for descriptive statistics. R package version 0.99. 18 R Found Stat Comput, Vienna, Austria
- Singh A, Kulkarni MA, Mohanty U, Kar S, Robertson AW, Mishra G (2012) Prediction of Indian summer monsoon rainfall (ISMR) using canonical correlation analysis of global circulation model products *Meteorological Applications* 19:179-188
- Sun Q, Miao C, Duan Q, Ashouri H, Sorooshian S, Hsu KL (2018) A review of global precipitation data sets: Data sources, estimation, and intercomparisons *Reviews of Geophysics* 56:79-107
- Tang B, Hsieh WW, Monahan AH, Tangang FT (2000) Skill comparisons between neural networks and canonical correlation analysis in predicting the equatorial Pacific sea surface temperatures *Journal of Climate* 13:287-293
- Terry J (2005) Hazard warning! Hydrological responses in the Fiji Islands to climate variability and severe meteorological events *Proceedings of 7th IAHS Scientific Assembly, IAHS publication* 296:33-40
- Terry JP, Raj R The 1997–98 El Niño and drought in the Fiji Islands. In: *Hydrology and water management in the humid tropics. Proceedings of the second international colloquium, 1999.* pp 22-26

- Terry JP, Raj R, Kostaschuk RA (2001) Links between the Southern Oscillation Index and hydrological hazards on a tropical Pacific island *Pacific Science* 55:275-283
- Thorne V, Coakeley P, Grimes D, Dugdale G (2001) Comparison of TAMSAT and CPC rainfall estimates with raingauges, for southern Africa *International Journal of Remote Sensing* 22:1951-1974
- Toté C, Patricio D, Boogaard H, van der Wijngaart R, Tarnavsky E, Funk C (2015) Evaluation of Satellite Rainfall Estimates for Drought and Flood Monitoring in Mozambique *Remote Sensing* 7:1758
- van Ogtrop F, Ahmad M, Moeller C (2014) Principal components of sea surface temperatures as predictors of seasonal rainfall in rainfed wheat growing areas of Pakistan *Meteorological Applications* 21:431-443
- Vernimmen R, Hooijer A, Aldrian E, Van Dijk A (2012) Evaluation and bias correction of satellite rainfall data for drought monitoring in Indonesia *Hydrology and Earth System Sciences* 16:133-146
- Walsh KJ, Bettio L, Power S, Fawcett R, Pahalad J (2001) Extended seasonal prediction of precipitation in Fiji *Australian Meteorological Magazine* 50:195-203
- Westra S, Sharma A (2010) An upper limit to seasonal rainfall predictability? *Journal of Climate* 23:3332-3351
- Wilks DS (2011) *Statistical methods in the atmospheric sciences* vol 100. Academic press,
- Wu H, Svoboda MD, Hayes MJ, Wilhite DA, Wen F (2007) Appropriate application of the standardized precipitation index in arid locations and dry seasons *International Journal of Climatology* 27:65-79
- Xie P, Chen M, Shi W CPC unified gauge-based analysis of global daily precipitation. In: Preprints, 24th Conf. on Hydrology, Atlanta, GA, Amer. Meteor. Soc, 2010.
- Xu W, Zou Y, Zhang G, Linderman M (2015) A comparison among spatial interpolation techniques for daily rainfall data in Sichuan Province, China *International Journal of Climatology* 35:2898-2907
- Yin Z-Y, Zhang X, Liu X, Colella M, Chen X (2008) An assessment of the biases of satellite rainfall estimates over the Tibetan Plateau and correction methods based on topographic analysis *Journal of Hydrometeorology* 9:301-326
- Yu Z-P, Chu P-S, Schroeder T (1997) Predictive skills of seasonal to annual rainfall variations in the US affiliated Pacific islands: Canonical correlation analysis and multivariate principal component regression approaches *Journal of Climate* 10:2586-2599
- Zeileis A, Grothendieck G (2005) zoo: S3 infrastructure for regular and irregular time series *arXiv preprint math/0505527*

Figures

Figure 1

Map of the Fiji islands showing the location of the six climatic stations

Figure 2

Climographs of the six selected stations, Nausori, Rarawai, Nadi, Matei, Labasa and Navua, showing average monthly temperature and precipitation

Figure 3

Spatially averaged annual time series of mean SST from 1980 to 2017 showing an increase in temperature from 2011 onwards

Figure 4

Spatial averaged annual time series of mean EDI in Fiji from 1980 to 2017 showing an increase in EDI values.

Figure 5

Kernel density plots for CPC, CHIRPS and PERSIANN-CDR data, and rain gauge data showing the underlying distribution of the different datasets for the Nadi station

Figure 6

Scatterplot for observed monthly Nadi station rainfall data versus CPC satellite data

Figure 7

Graphical representation of a hindcast and a forecast for the predicted and observed mean EDI time series for Fiji.

Figure 8

Spatial Plot of observed and Hindcast of mean EDI in 2004 in model training stage with $R^2 = 0.83$ and MSE = 0.12 (colour sequence representing the different EDI classes)

Figure 9

Spatial plot of observed and forecast of mean EDI in 2017 in model validation stage with $R^2 = 0.45$ and MSE = 0.26 (colour sequence representing the different EDI classes)