


Subseasonal prediction framework for tropical cyclone activity in the Solomon Islands region

Alick Haruhiru¹  | Savin S. Chand¹ | Nargiz Sultanova¹ | Hamish Ramsay² | Krishneel K. Sharma¹ | Lloyd Tahani³

¹Institute of Innovation, Science and Sustainability, Federation University Australia, Ballarat, Victoria, Australia

²CSIRO, Environment, Aspendale, Victoria, Australia

³Solomon Islands Meteorological Service, Honiara, Solomon Islands

Correspondence

Alick Haruhiru, Institute of Innovation, Science and Sustainability, Federation University Australia, Mt Helen Campus, Ballarat, VIC 3350, Australia.
Email: a.haruhiru@federation.edu.au and alikey@gmail.com

Funding information

Department of Foreign Affairs and Trade, Australian Government

Abstract

Recently, we developed seasonal prediction schemes with improved skill to predict tropical cyclone (TC) activity up to 3 months in advance for the Solomon Islands (SI) region (5°–15°S, 155°–170°E) using sophisticated Bayesian regression techniques. However, TC prediction at subseasonal timescale (i.e., 1–4 weeks in advance) is not being researched for that region despite growing demands from decision makers at sectoral level. In this paper, we first assess the feasibility of developing subseasonal prediction frameworks for the SI region using a pool of predictors that are known to affect TC activity in the region. We then evaluate multiple predictor combinations to develop the most appropriate models using a statistical approach to forecast weekly TC activity up to 4 weeks in advance. Predictors used include indices of various natural climate variability modes, namely the Madden–Julian Oscillation (MJO), the El Niño–Southern Oscillation (ENSO), the Indian Ocean Dipole (IOD) and the Interdecadal Pacific Oscillation (IPO). These modes often have robust physical and statistical relationships with TC occurrences in the SI region and the broader southwest Pacific territory as shown by preceding studies. Additionally, we incorporate TC seasonality as a potential predictor given the persistence of TCs occurring more in certain months than others. Note that a model with seasonality predictor alone (hereafter called the “climatology” model) forms a baseline for comparisons. The hindcast verifications of the forecasts using leave-one-out cross-validation procedure over the study period 1975–2019 indicate considerable improvements in prediction skill of our logistic regression models over climatology, even up to 4 weeks in advance. This study sets the foundation for introducing subseasonal prediction services, which is a national priority for improved decision making in sectors like agriculture and food security, water, health and disaster risk mitigation in the Solomon Islands.

KEYWORDS

El Niño–Southern Oscillation, Madden–Julian Oscillation, subseasonal prediction, tropical cyclones

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *International Journal of Climatology* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

1 | INTRODUCTION

Tropical cyclone (TC) prediction at various timescales is a crucial component of the ongoing Solomon Islands (SI) community preparedness effort to minimize potential risks associated with anticipated TC events. Currently, the Solomon Islands Meteorological Service (SIMS) is providing cyclone early warnings up to 3 days ahead using forecasts produced by advanced numerical weather prediction (NWP) models. Forecasting TCs at seasonal to interannual timescales has also improved substantially over the SI region due to our increased understanding of the patterns, drivers and predictability of slowly evolving components of the global climate system such as the El Niño–Southern Oscillation (ENSO) phenomenon (e.g., Magee et al., 2020; Haruhiru et al., 2023). In between these two timescales is a subseasonal range, which has long been considered a “predictability desert” and received much less attention globally compared to short-range and seasonal TC forecasts (e.g., Camargo et al., 2019; Robertson et al., 2020). At this timescale, even the more advanced high-resolution NWP models often have substantial limitations to skilfully forecast TC development due to their lack of ability to explicitly resolve essential small-scale processes that affect genesis development (e.g., Molteni & Palmer, 1988; Robertson et al., 2020). While few studies have primarily investigated subseasonal predictability of TC activity for the wider southwest Pacific (SWP) basin (0°–35°S, 135°E–120°W; Figure 1) (e.g., Lee et al., 2018; Leroy & Wheeler, 2008; Vitart et al., 2010), none of them focussed on country-level predictions. Therefore, our emphasis here is to predict TCs in the SI region at subseasonal timescale

(i.e., up to 4 weeks in advance in our case) using sophisticated statistical techniques, noting that SI region is among the most vulnerable island nations in the SWP in terms of TC impacts (e.g., Britton, 1987; Haruhiru et al., 2022; Maru et al., 2018; Solomon Islands Government, 2014).

Over the past decades, our understanding of climate variability modes that operate at intraseasonal timescale, and hence modulate subseasonal predictability of TCs activity (e.g., Zhang, 2005, 2013), has improved significantly (Robertson et al., 2020). The Madden–Julian Oscillations (MJO) is well documented as the primary driver of TC occurrences across the globe at intraseasonal timescale (Klotzbach, 2014), and therefore serves as a main predictor of subseasonal TC activity (e.g., Chand & Walsh, 2010; Hall et al., 2001; Leroy & Wheeler, 2008). It operates at 30–90-day cycle and is characterized by an eastward propagation of unstable weather conditions (e.g., rainfall, convective clouds and strong winds) across the tropical Pacific region (Hendo & Salby, 1959; Zhang, 2005). As MJO progresses eastward, it impacts TCs by modulating large-scale environmental conditions such as vertical wind shear, low-level relative vorticity and mid-level moisture that are known to affect TC activity. In the SWP, when convectively active phases of the MJO are over the region, TC numbers can be enhanced by up to ~36% (e.g., Diamond & Renwick, 2015). In contrast, when convectively inactive phases of the MJO passes over the region, TC numbers can be significantly suppressed by up to 17% as demonstrated by Diamond and Renwick (2015). The number of cyclones reaching hurricanes or major hurricane intensity can also undergo significant enhancement in active phases of the MJO when compared to inactive phases (Chand & Walsh, 2010; Klotzbach, 2014). Moreover,

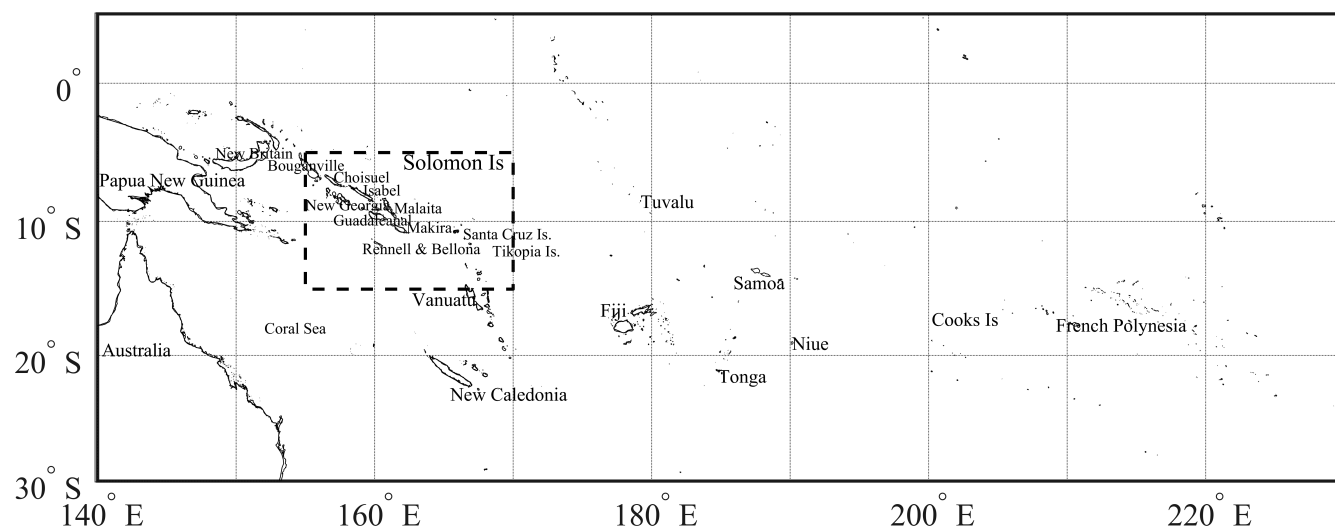


FIGURE 1 Map of the southwest Pacific. The Solomon Islands region (study area) is bordered with dashed line

the relationship between TC activity and the MJO can be further strengthened during ENSO events. For example, if the enhanced phases of the MJO co-occur with El Niño events, more TCs are likely to form in the SWP basin when compared to the enhanced phases of the MJO co-occurring with La Niña events (Chand & Walsh, 2010; Hall et al., 2001). Other modes of natural variability such as the Indian Ocean Dipole (IOD) (e.g., Magee et al., 2017) and Interdecadal Pacific Oscillation (IPO) (e.g., Magee & Verdon-Kidd, 2018) also modulate TCs in the broader SWP basin, as well as in the SI region (Haruhiru et al., 2023, unpublished); however, the extent to which their co-occurrence with the MJO can affect TCs in the SWP is not known.

The scope of the present investigation is to explore the predictability of TCs in the SI region at subseasonal timescale using combinations of predictors that modulate climate variability in the SWP. The multiweek TC prediction product developed in this work will have a significant impact on disaster preparedness and emergency procedures for the SI region. It is anticipated the product will enable SIMS to extend its forecasts beyond 3 days. Extending the prediction to multiweeks can have considerable socioeconomic benefits in terms of planning and preparation for cyclone events. For example, emergency management sectors within the island country may incorporate extended forecast information into their long-term emergency evacuation plan to make informed decisions and take appropriate steps when preparing for and responding to TCs and associated hazards. Also, the methodology developed in this project will be very cost-effective, and therefore can be applied to other SWP countries that are impacted by TCs.

2 | TC DATA AND DEFINITIONS

TC data are acquired from the Southwest Pacific Enhanced Archive for Tropical Cyclone (SPEARTC) database. SPEArTC is developed using the International Best Track Archive for Climate Stewardship (IBTrACS) data for the region between the latitudes 5°–65°S and the longitudes 135°–120°W (Diamond et al., 2012). Hence, it has all the qualities attributed to IBTrACS but with some crucial updates that incorporate removing duplicate records and adding TCs that were not previously documented for the SWP (Diamond et al., 2012). The SPEArTC data are available from the 1840s onward with mostly 6-hourly track points and the conventional World Meteorological Organisation (WMO) 10-min wind averaging period. However, we only utilized TC records from 1974/1975 to 2018/2019 to be consistent with the period when satellite monitoring of TCs became routine. This period also

coincides with the availability of the updated version of the MJO indices used in this study (e.g., Wheeler & Hendon, 2004).

Consistent with our earlier studies (Haruhiru et al., 2022, 2023, unpublished), a TC season is defined from July 1st to June 30th in the following year to comply with the WMO classification for the SI TCs. This period encompasses the active cyclone season from November 1st through April 30th. The second calendar year is used to refer to a particular season (e.g., the 1975 TC season is from July 1st, 1974, to June 30th, 1975). Altogether, there were 45 cyclone seasons in the analysis. Note that TCs that achieve sustained windspeed of at least 34 knots (or $\sim 62 \text{ km}\cdot\text{h}^{-1}$) and entered the SI region are considered in this investigation (hereafter, referred to as TC activity). Using these criteria, a total of 147 TCs were recorded for the period 1975–2019 and are considered for the development of a subseasonal prediction framework in this study.

3 | CONSTRUCTING OBSERVED WEEKLY TC PROBABILITIES

We constructed a weekly index of TC probability comprising of a binary indicator variable P such that $P = 0$ when no TC is observed during the week of interest and $P = 1$ if any TC is observed. This statistical approach is considered optimal for predicting weekly TC activity using logistic regression where dependent variables are binary (see section 4). A week is defined here as a 7-day period over each cyclone season starting from July 1st to June 30th. This definition gives a total of 52 weeks per season, noting some overlapping of days in the last week of adjacent seasons, and a total of 2348 weeks over the study period 1975–2019.

Climatological seasonal cycle (CSC) of weekly TC probabilities is created by compositing weekly TC probabilities over the study period such that

$$\text{CSC}_k = \frac{\sum_{i=1}^n P_k}{n}, \quad (1)$$

where P_k is the probability of any cyclone in the week of interest k (noting $P_k = 0$ if no cyclone, and $P_k = 1$ if any cyclone), n is the total number of seasons (i.e., 45 in our case) and k ranges from 1 to 52. Figure 2 provides the climatological seasonal cycle of weekly TC activity in the SI region. The smoothed seasonality is obtained using kernel density estimates (e.g., Bowman & Azzalini, 1997). As expected, enhanced weekly TC activity is found over the months of December–March (i.e., smoothed weekly

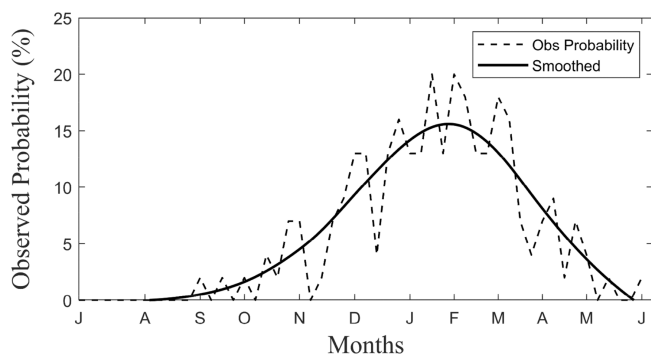


FIGURE 2 Climatological seasonal cycle of weekly (7 days) tropical cyclone occurrence (TC seasonality) in the SI region. Actual and smoothed values are represented by dashed and solid lines, respectively

probabilities ranging from 10% to 15%); the peak weekly activity occurs during January and February. Note that TCs rarely occur outside of the active cyclone season (i.e., November–April) for the SI region. TC *Raquel* in June 2015 and TC *Namu* in May 1985 are the two known examples of cyclone events recorded outside of the active TC season in the SI region over the study period.

4 | POTENTIAL CLIMATE PREDICTORS

Two major factors were considered in the process of choosing appropriate TC predictors. First, the predictor must have a robust physical and statistical relationship with weekly TC activity in the SI region (or the wider SWP basin) and second, these predictors must be available at least 1 week in advance to predict weekly TC activity. Our earlier studies focused on the SI region (e.g., Haruhiru et al., 2022, 2023, unpublished) showed that cyclones affecting that region are strongly modulated by different modes of natural climate variability. For example, at intraseasonal timescale, MJO is a primary driver of weekly TC activity in the SI region, whereas at interannual timescale, ENSO has a dominating influence on TC activity. At interdecadal timescale, IPO can also modulate TCs in the SI region (e.g., Haruhiru et al., 2022, 2023, unpublished) and in the broader SWP region (Magee et al., 2017; Sharma et al., 2019). Other studies have shown that IOD and other modes of Indian Ocean SST variability can also affect the interannual predictability of TC activity in the SWP region (e.g., Ramsay et al., 2017). Based on the information from these studies, we consider MJO, ENSO, IPO and IOD—in addition to the climatological seasonal cycle discussed above—as

potential predictors of weekly TC activity in the SI region. Below we discuss how each of these modes of variability are incorporated as weekly predictors of the SI region TC activity.

4.1 | The MJO

We used the daily Real-Time Multivariate MJO index that was developed by Leroy and Wheeler (2008) as the leading predictor of TC activity at intraseasonal timescale. The choice of the MJO as a potential predictor is based on recent studies that have established the significant role of MJO modulation on TC occurrences in the SI region (Haruhiru et al., 2022; Maru et al., 2018). The index comprises a pair of empirical orthogonal functions (EOFs) derived from the 850- and 250-hPa zonal wind fields and outgoing long-wave radiation (OLR) around the equatorial region (Wheeler & Hendon, 2004). Briefly, the progression of the MJO across the equatorial region is described using the coefficient time series of the two EOFs known as Real-Time Multivariate MJO 1 (RMM1) and Real-Time Multivariate MJO 2 (RMM2) (Wheeler & Hendon, 2004). The index has eight MJO phases plus a weak phase (i.e., when the magnitude of RMM1 and RMM2 is less than 1). The RMM1 and RMM2 values were first available since 1974 onwards and are obtained from the Australian Bureau of Meteorology (BOM) website (Australian Bureau of Meteorology, 2022a).

The genesis position of each TC track was initially stratified into the MJO phases (and the weak phase) to understand its degree of modulation on TC formation. Because the individual bins have very few TC counts, we combined the MJO phases as 2 + 3, 4 + 5, 6 + 7, 8 + 1, and a weak phase following the classification by Ramsay et al. (2012). The daily genesis rate (DGR) was then calculated for each of the combined MJO phases by dividing the total number of TCs (NTCs) by the respective total number of MJO days (N); the results are expressed in percentages (Figure 3).

A total of 13,337 MJO days was found during the study period, but only 147 of these days are associated with TC activity. As shown in Figure 3c, phase 7 + 6 has the highest DGR of $\sim 2.43\%$, followed by the MJO phases 4 + 5 at $\sim 1.09\%$ (Figure 3b). In contrast, the MJO phase 2 + 3 has a relatively lower DGR of $\sim 0.93\%$, whereas the weak phase has the smallest value of DGR (i.e., $\sim 0.86\%$). A comprehensive analysis of the MJO modulations on TC genesis in the SI region is found in Haruhiru et al. (2022). Essentially, the MJO significantly modulates TC genesis in the SI region, as well as the broader SWP territory (e.g., Chand & Walsh, 2010; Deo et al., 2021; Ramsay et al., 2012), and hence, is regarded as the primary predictor of weekly TC activity in our study.

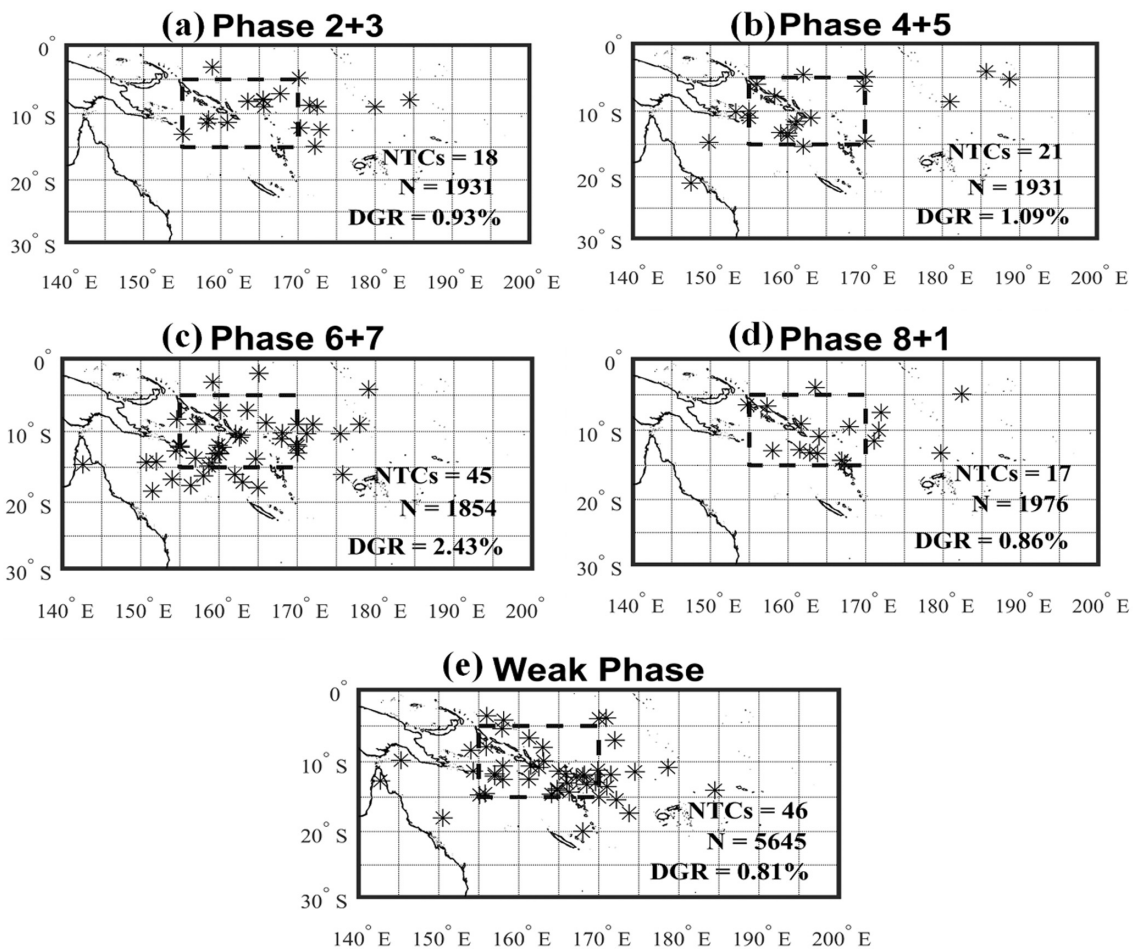


FIGURE 3 Genesis positions denoted as asterisks (*) of TCs affecting the SI region for various MJO phases during the period July-1974 to June-2019. The MJO phases are combined: (a) The MJO phases 2 + 3, (b) 4 + 5, (c) 6 + 7, (d) 8 + 1, and (e) the weak phase. NTC, N and DGR represent total TC counts, total MJO days and daily genesis rate, respectively

For ease of interpretation and analysis, we extracted the mid-week RMM1 and RMM2 values as predictors of the MJO phase for the week of interest. This method assigns each week with the appropriate MJO indices over the study period. For predictions of weekly TC activity (W_k), respective lagged values are utilized such that

$$W_k = \text{MJO}_{k-1}. \quad (2)$$

That is, to predict TC activity 1 week in advance, the MJO indices of the previous week (MJO_{k-1}) are utilized, and so on.

4.2 | El Niño–Southern Oscillation

The Niño3.4 index (e.g., Trenberth, 1997) is used as the ENSO based predictor of TCs in the SI and wider SWP region. Note that Niño3.4 is preferred because it is slightly more skilful than alternative predictors such as

Niño3, Niño4, Southern Oscillation Index (SOI) and the ENSO longitude index (for predictor evaluations, see, e.g., Haruhiru et al., 2023, unpublished). The ENSO–TC connection is well-reported in the SI region (Haruhiru et al., 2022; Maru et al., 2018), the wider SWP (e.g., Diamond et al., 2013; Dowdy et al., 2012; Ramsay et al., 2012), the Australian region (e.g., McDonnell & Holbrook, 2004; Nicholls, 1985; Ramsay et al., 2008), and the Fiji, Samoa and Tonga region (Chand & Walsh, 2011b; Tu'uholoaki et al., 2022). Synergistic relationship between the MJO and ENSO can also modulate TC activity in the South Pacific (e.g., Chand & Walsh, 2010; Hall et al., 2001).

To demonstrate the effects of ENSO on weekly TC activity in the SI region, and hence its feasibility as a predictor, we used the same procedures as with MJO to construct the observed probability of TC genesis during the three ENSO phases: El Niño, La Niña and neutral phases (Figure 4a–c). El Niño and La Niña years were defined using thresholds of SST anomaly values of $+0.5^\circ\text{C}$ and

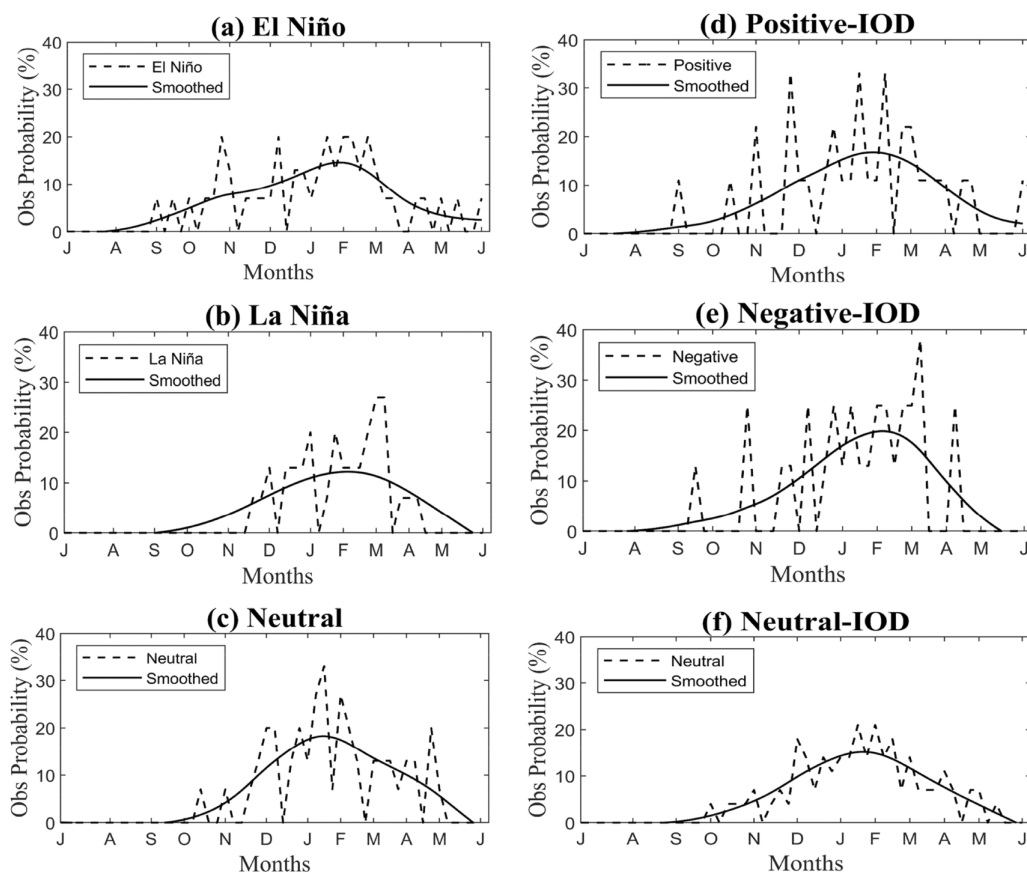


FIGURE 4 Observed probability of TCs forming in the SI region during ENSO phases: (a) El Niño, (b) La Niña and (c) ENSO-neutral years, and during the IOD phases: (d) positive IODs, (e) negative IOD and (f) IOD-neutral years. Dashed lines represent the observed probabilities and in solid black are smoothed curves

above, and -0.5°C and below, respectively, in the Niño3.4 region using 1981–2010 as the base period (e.g., Haruhiru et al., 2022). Values that fall between these two thresholds were marked as ENSO-neutral years. We found that there were 15 years each for the three ENSO phases from 1975 to 2019. During El Niño events, the smoothed weekly probability of TCs occurred during November–April ($\sim 10\%$ to 15%), with the peak weekly activity during January and February (Figure 4a). For La Niña years, the observed smoothed probability is slightly reduced during November–April when compared with the El Niño climatology (i.e., $\sim 5\%$ – 12%) (Figure 4b). However, an interesting difference between the El Niño and La Niña climatologies is that for El Niño, weekly TC activity is often enhanced outside the usual cyclone season (i.e., August–November and May–June). This distinction in TC activity between El Niño and La Niña years can serve an important role in predicting weekly TC activity in the SI region. For ENSO-neutral years, it appears that the magnitude of the smoothed weekly TC activity is enhanced ($\sim 10\%$ – 18%) during the peak seasons of November–April (Figure 4c). Based on these assessments,

we include ENSO as a potential predictor of the weekly TC activity in the SI region.

Note that daily or weekly values of ENSO indices can be very “noisy,” which may lead to misleading results. As Niño3.4 index is readily available at monthly timescale, we use previous month’s value to predict weekly TC activity for the subsequent month. Since all 4 weeks often fall within a particular month, each week (W_k) is assigned the same value of the ENSO predictor from the previous month such that

$$W_k = \text{ENSO}_{m_{k-1}}, \quad (3)$$

where m_k is the month that spans the week k for which the prediction is made.

4.3 | IOD and IPO

We also included the IOD and IPO as potential predictors of weekly TC activity in the SI region. The Dipole Mode Index (DMI; e.g., Saji et al., 1999) is often used as an

indicator of IOD, whereas the IPO index (Chand et al., 2023; Mantua et al., 1997; Salinger et al., 2001) as a measure of the IPO phases. Both IOD and IPO are proven as potential predictors of TC activity in the SI region (e.g., Haruhiru et al., 2022, 2023 unpublished) and in the wider SWP region (Magee et al., 2017; Magee & Verdon-Kidd, 2018; Ramsay et al., 2017). Refer to Haruhiru et al. (2023, unpublished) for comprehensive evaluation of climate-based predictors for seasonal TC activity in the SI region. The monthly IPO dataset are obtained from the National Oceanic and Atmospheric Administration website (NOAA ESRL, 2022) and were extracted for the period June 1974–December 2019 to coincide with our TC dataset from the SPEArTC. IPO phases are consistent with the ENSO thresholds definitions. The IOD phases are defined according to the standard Australian Bureau of Meteorology (BOM) definitions that are accessible at their website and updated in the BOM climate archives (Australian Bureau of Meteorology, 2022b).

As with the MJO and ENSO, we demonstrated the observed weekly probability of TCs occurring in the SI region during the three IOD phases: positive, negative and neutral (Figure 4c,d). Weekly TC activity during the positive and negative IOD phases are generally similar for each other, with peak activity occurring during January and February (Figure 4d,e). However, in both phases, we note the presence of the off-season weekly TC activity, associated with El Niño events. The magnitude of the weekly TC activity during the active cyclone season for IOD-neutral years is relatively low compared with positive and negative IOD (Figure 4f). Indices of DMI and IPO are available at monthly timescale, and therefore, we use their previous month's values to predict weekly TC activity for the subsequent month as with ENSO.

5 | MODEL DEVELOPMENT PROCEDURES

5.1 | Logistic regression model structure

We used a logistic regression model to predict weekly TC activity in the SI region following the approach pioneered by Leroy and Wheeler (2008) for the broader SWP region. Logistic regression is an effective technique for making probabilistic forecasts to demonstrate whether a TC will develop or not (Wilks, 2006). The logistic regression model takes the following structure:

$$\hat{P} = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m}}, \quad (4)$$

where \hat{P} is the predicted probability of TC formation and (x_1, x_2, \dots, x_m) are the predictors. In our case, as discussed

previously, the weekly probability of observed TCs takes the value of 0 if no TC is observed during the week of interest, and the value of 1 otherwise. The model coefficients $(\beta_0, \beta_1, \beta_2, \dots, \beta_m)$ are calculated using the least squares procedure (Leroy & Wheeler, 2008; Marquardt, 1963).

5.2 | Predictor selections and model development

Three logistic regression models were developed and assessed, and each model incorporated different sets of predictors (Table 1). The first model (hereafter referred to as Model 1) is a single predictor model which only consisted of the climatological seasonal cycle (or seasonality). The second model, Model 2, consisted of the five climate-derived predictors, namely RMM1, RMM2, Niño3.4, DMI and IPO indices. The third model, Model 3, comprised all predictors (i.e., seasonality from Model 1 and climate-derived predictors from Model 2).

Questions may also arise on the need to apply objective statistical procedures such as the stepwise method (e.g., Leroy & Wheeler, 2008; Vitart et al., 2010) for selection of predictors. Often, a large number of predictors in a multiple regression model can lead to multicollinearity, where independent variables are highly correlated with each other (e.g., Wilks, 2006). In our case, the focus is essentially on physical reasonings for predictor selection, and so it is important to check for multicollinearity in the model, a situation where the predictors are highly correlated with each other (e.g., Chatterjee & Simonoff, 2013). Multicollinearity is a matter of degree, so there is no irrefutable test that it is or is not a problem. However, there are several warning signals for multicollinearity (e.g., Shrestha, 2020), and for the purpose of this study, we used the “Variance Inflation Factor (VIF)” to investigate any multicollinearity issue (e.g., Marcoulides & Raykov, 2019; Shrestha, 2020). Values of $VIF > \sim 10$ are considered problematic, but in our case, none of the models shows any signs of multicollinearity (see Appendix A). Hence, all the predictors are retained justifiably.

TABLE 1 The logistic models and the associated TC predictors

Models	Predictor selection
Clim model	Probability of any TC formed ($\sim 6\%$)
1	Seasonality only
2	RMM1, RMM2, Niño3.4, DMI, IPO
3	Seasonality, RMM1, RMM2, Niño3.4, DMI and IPO

Note: Clim model is the climatology model.

Finally, the performance of the three models was compared with a baseline (or reference model), which is the climatology model (Clim) in this study. The climatology model simply predicts the probability that any cyclone will form in a given week over the period 1975–2019. In our case, we have a total of 2348 weeks and only 147 of those are associated with at least any TC event. From a frequentist perspective, the probability that a cyclone will form in each week based on these statistics is 6%. Table 1 shows a summary of the models with their respective predictors.

5.3 | Observed and modelled weekly TC probabilities

Figure 5 shows the observed and modelled weekly TC probabilities for the SI region. Model 1 (i.e., the model with seasonality predictor) and Model 3 (i.e., the model with seasonality, RMM1, RMM2, Niño3.4, DMI and IPO) can capture the general pattern of the observed weekly TC probabilities relatively well. However, there appears to be a slight overprediction of weekly TC probabilities during January and February, the two peak months of TC activity in the SI region. In contrast, Model 2 (i.e., the model with RMM 1, RMM2, Niño3.4, DMI

and IPO), significantly underpredicts weekly TC probabilities during the peak season. Additionally, it is worth noting that all the models can also capture the observed weekly probabilities during the off-season as well (i.e., September–November and April–June). A potential reason for improved performance by Models 1 and 3 could be attributed to the inclusion of TC seasonality as an input predictor, consistent with findings by Leroy and Wheeler (2008) in the wider Southern Hemisphere and by Maier-Gerber et al. (2021) in the North Atlantic, while the poor prediction by Model 2 is due to the absence of TC seasonality as a predictor. This suggests that the observed TC seasonality alone is a valuable probabilistic indicator or predictor, just as Model 3, which combines TC seasonality and all input predictors of TCs. This is an important result as, to date, no subseasonal prediction model exists for the SI region, and a model as simple as seasonality alone can provide much-needed information for decision-making purposes.

So far, our assessment of models was at zero lead time as we wanted to assess the performance of models relative to the observation. In practice, one would like to use these models to make predictions several weeks in advance. In the next section, we examine the performance of these logistic regression models in predicting

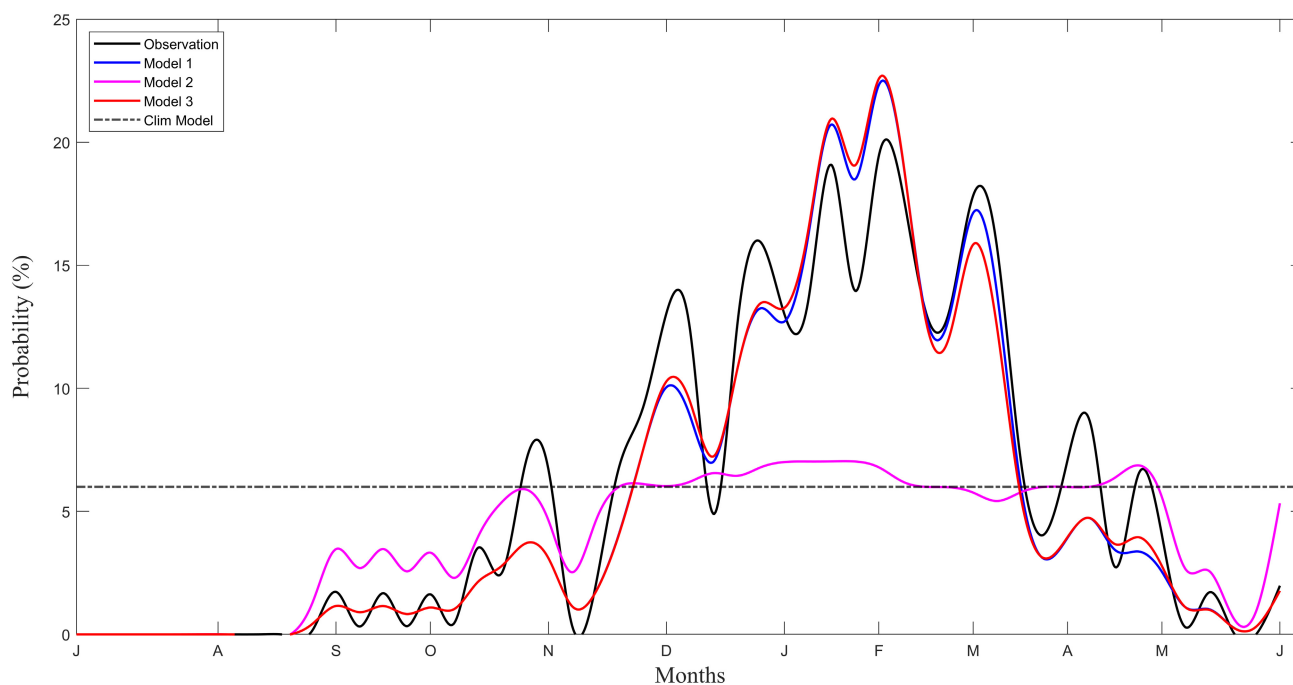


FIGURE 5 Observed probability vs model predictions of TC genesis in the SI region. The black curve indicates observed probability, while blue, magenta and red lines represent predictions by Model 1 (TC seasonality), Model 2 (model with predictors: RMM1, RMM2, Niño3.4, DMI and IPO) and Model 3 (model with predictors: TC seasonality, RMM1, RMM2, Niño3.4, DMI and IPO), respectively. Horizontal dashed black line marks the Climatology model at 6%. The curves are smoothed using the spline function [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

weekly TC probabilities at 4-weekly lead times: 1-week (W1), 2-week (W2), 3-week (W3) and 4-week (W4).

5.4 | Model validation

A crucial element of our work is to validate the developed models by determining how good they are at estimating weekly TC activity in practice for the SI region. This is accomplished by testing the three models on an independent dataset (i.e., those datasets that were not used to construct the models), noting that independent dataset for model assessments can be constructed by different statistical approaches. One such widely used approach is the leave-one-out cross-validation (LOOCV; e.g., Leroy & Wheeler, 2008) technique, which we have applied here to evaluate the prediction skill of the three models (i.e., Models 1, 2 and 3) and assess their performances with respect to the reference model (Clim). In summary, the LOOCV technique operates by sequentially omitting a weekly observation from the dataset and using the remaining data (also known as the training dataset) to develop a model, which is then used to predict the omitted observation. The resulting “prediction” is normally termed as the “hindcast.”

Results from the hindcast predictions are verified using three verification metrics: the mean square error (MSE), Brier skill scores (BSS) and reliability diagrams. These are commonly used verification methods for assessing probabilistic forecasts (Camp et al., 2018; Chand & Walsh, 2011a; Leroy & Wheeler, 2008; Maier-Gerber et al., 2021; Vitart et al., 2010; Wilks, 2006).

Following Elsner and Jagger (2006) and Chand and Walsh (2011a), the MSE is defined as the averaged squared difference between the predicted probability of TCs and the observed probability (which is either $k = 1$ or $k = 0$ depending on whether a cyclone is observed in the week of interest or not, respectively) such that

$$MSE = \frac{1}{N} \sum_{i=1}^N \sum_{k=0}^1 P_i(k) [k - y_i(k)]^2, \tag{5}$$

TABLE 2 The hindcast mean squared error (MSE) and Brier skill score (BSS) for the climatology model, and the three logistic models (Models 1, 2 and 3) for W1

Models	MSE	BSS
Clim model	0.113	
1	0.108	0.045
2	0.112	0.008
3	0.106	0.058

where $P_i(k)$ is the predicted weekly probability associated with no TC ($k=0$) or any TC ($k=1$), y_i is the observed weekly probability during the study period ($i=1, 2, \dots, N$). The nearer the hindcasts' weekly probabilities are to the actual observed weekly probabilities, the lesser the MSE, hence the better the model. Since the models provide probability forecasts, the MSE values are also used to compute the BSS with reference to the Clim model such that

$$BSS = 1 - \left(\frac{MSE_{model}}{MSE_{Clim}} \right). \tag{6}$$

A value of BSS = 1 signifies a perfect forecast, while BSS = 0 is for a model that performs relatively similar to the Clim model. In contrast, a negative BSS value indicates model that performs poorer than the Clim model.

Table 2 shows the verification scores of each model for predicting weekly TC activity at one-week lead time (i.e., W1 forecast). The MSEs for all models are small, indicating that these models do well overall. To objectively assess these models' performance with respect to the reference Clim model, we look at the Brier skill score. The BSS of all models at W1 lead time are positive, indicating they all perform better than the Clim model. In addition, the BSS for Model 3 (BSS = 0.058) is greater than those for Model 1 (BSS = 0.045) and Model 2 (BSS = 0.008), suggesting that Model 3 appears to be more skilful than Models 1 and 2 and that Model 1 has more skill than Model

TABLE 3 As in Table 2, but for W2 forecasts

Models	MSE	BSS
1	0.108	0.040
2	0.113	0.002
3	0.108	0.045

TABLE 4 As in Table 2, but for W3 predictions

Models	MSE	BSS
1	0.109	0.037
2	0.113	-0.002
3	0.108	0.039

TABLE 5 As in Table 2, but for W4 predictions

Models	MSE	BSS
1	0.110	0.031
2	0.113	-0.002
3	0.109	0.035

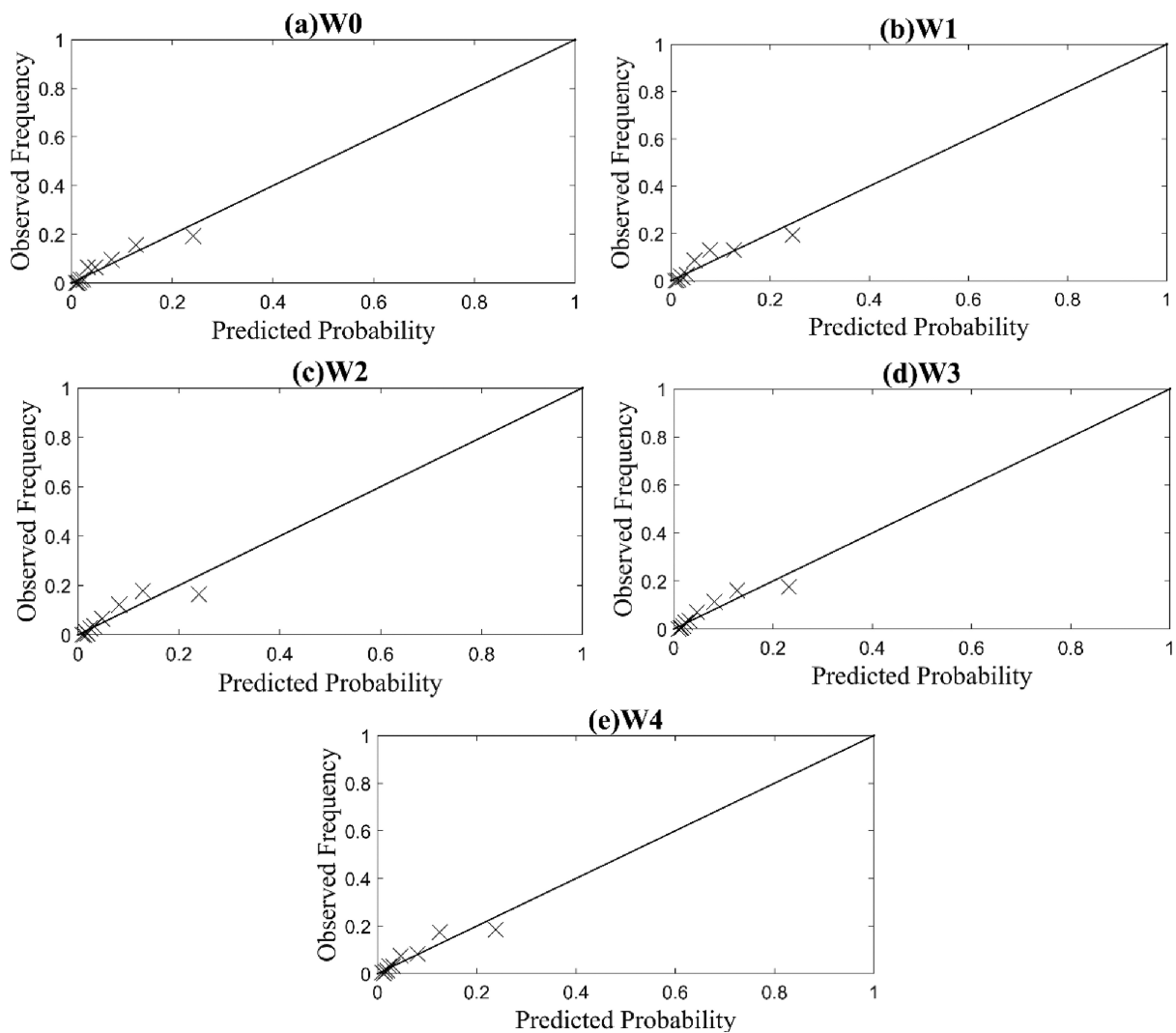


FIGURE 6 Reliability diagrams for weekly TC activity forecasts produced by Model 3. The diagonal line indicates a perfect forecast. The average of predicted and observed probabilities are indicated by the symbol “x.” W0 represents zero prediction and W1, W2, W3 and W4 are the respective forecast weeks

2. Table 3 displays the MSE and BSS values for the W2 forecasts. As in W1 forecast, while all models perform better than Clim model, Model 3 (BSS = 0.045) seems to be the best model compared with Models 1 (BSS = 0.040) and 2 (BSS = 0.002) for the W2 forecasts. Likewise, Model 3 (BSS = 0.039) outperformed Models 1 and 2 at W3 and W4 forecast lead-times (Tables 4 and 5). However, at these W3 and W4 lead times, Model 2 is outperformed by the Clim model, suggesting that this model may not be appropriate for the prediction of weekly TC activity at more than 3 weeks lead time. Apart from individual performance, the verification scores also indicate models' decreasing skill from W1 to W4 forecasts. Overall, our Model 3 appears to be the most suitable model for making predictions at all four lead-times.

Finally, reliability diagrams are used to demonstrate how often the probability forecast may actually

materialize for the cross-validated forecasts (Wilks, 2006). The reliability diagrams were constructed by plotting the observed weekly TC probability against the hindcast weekly probability. Briefly, the hindcasts acquired from the LOOCV procedure were binned into K -subsets (in our case, $K = 234$ such that each bin has 10 realizations). Next, the average values from the corresponding groups were plotted against each other to produce the reliability diagrams. Reliability diagrams for different lead times are shown in Figure 6 (note that the diagonal lines signify a perfect forecast). The closer the plotted points are to the diagonal line, the more reliable the model. For ease of interpretation, we only demonstrated the reliability diagrams for Model 3 (i.e., the model that consists of TC seasonality, RMM1, RMM2, Niño3.4, DMI and IPO) as it is the most representative model for predicting weekly TCs activity in the SI region. Overall, it is

clear from these plots that the observed and hindcast values are concentrated towards lower probabilities, indicating that TCs are rare events. These values are also aligned along the diagonal line, further suggesting that Model 3 is well-calibrated for the forecasts at different lead times and there seems to be no conditional biases between the observed and hindcast probabilities.

6 | RESULT AND SUMMARY

This study presents the development of a statistical prediction framework to make subseasonal probabilistic forecasts of weekly TC activity for the SI region (5°–15°S and 155°–170°E). At a country level, we found that there is no subseasonal prediction framework in the literature for any of the Pacific Island Countries (PICs), a region that is often affected and exposed to TC-related hazards (e.g., Deo et al., 2021, 2022; Kuleshov et al., 2020; Magee et al., 2016). Leroy and Wheeler (2008) demonstrated the feasibility of predicting TC genesis on weekly basis in the Southern Hemisphere using statistical approach. Here, we utilized the methodology developed by Leroy and Wheeler (2008) to predict weekly TC activity for the SI region up to 4 weeks in advance.

Over the period 1975–2019, 147 cyclones are found to affect the SI region. Climatologically, the observed weekly probability of TC activity peaks during the months of January–March. However, we note that TC formations in the SI region can also cluster in time, primarily associated with the MJO events. Clearly, enhanced cyclone activity is noticed in the active phases of the MJO (i.e., phases 6 + 7) compared with inactive and weak phases. We also note that off-season occurrences of TCs in the SI region are characteristic of El Niño when compared with La Niña years and that IOD and IPO also modulate weekly TC activity in the region at interannual and decadal timescales, respectively. Differences in the seasonality of TC activity, as well as different effects of climatic variables (particularly the MJO) on TCs, provide us with the basis for the development of weekly TC prediction schemes over the SI region.

The MJO is the key climate driver of TC occurrences at intraseasonal timescale, especially during the active MJO phases 6 and 7 over the SI region (Haruhiru et al., 2022). Hence, the daily MJO Index developed by Wheeler and Hendon (2004) is used as predictors at inter-seasonal timescale. At longer timescales, ENSO, IOD and the IPO have shown robust connections with TC genesis over the SI region and wider SWP territory (Haruhiru et al., 2022; Magee & Verdon-Kidd, 2018). Hence, the monthly anomaly values of Niño3.4, DMI and IPO are used as predictors. These monthly predictors were also effectively

used as cyclone predictors in developing a seasonal TC prediction framework for the SI region (Haruhiru et al., 2023, unpublished). Additionally, the constructed observed TC seasonality from July 1974 to June 2019 is found to be an effective TC predictor; consequently, it is considered an important input predictor in our study.

Three Logistic regression models were then developed with different input predictors to forecast weekly TC activity in the SI region at four lead-times: W1, W2, W3 and W4. Model 1 had only one predictor which is constructed from TC seasonality, while Model 2 comprised RMM1, RMM2, Niño3.4, DMI and IPO indices. Model 3 combined the input predictors in Models 1 and 2. The LOOCV technique was employed to determine hindcast probabilities of weekly TC activity for each model. The MSE, BSS and reliability diagrams were then used to compare the skill of different models. Overall, we found that Model 1 and Model 3 performed comparatively well over the reference model (Clim) for all lead times, but Model 3 appears to be more optimal and has shown substantial ability to predict weekly TC activity in the SI region.

These results provide a potentially useful framework for operational subseasonal TC outlook services at the Solomon Islands Meteorological Service. This study can also be replicated and further improved by other National Meteorological and Hydrological Services in the SWP to enhance their TC advisory services given that subseasonal prediction of TCs is one of the priority subjects by disaster managers in PICs. It is important to note that our methodology is purely statistical in nature and utilizes readily available climate indices as predictors to make subseasonal predictions of TC activity up to 4 weeks in advance. We acknowledge that results from dynamical models (such as those from the subseasonal to seasonal prediction project (S2S) by the World Weather Research Program and the World Climate Research Program; Vitart et al., 2017; Vitart & Robertson, 2018) are now available. It is anticipated that incorporating predictors from the S2S project into our statistical models as part of future work can lead to improved forecasts of TC activity in the SI region.

AUTHOR CONTRIBUTIONS

Alick Haruhiru: Funding acquisition; writing – original draft; methodology; validation; visualization; writing – review and editing; software; formal analysis; project administration; data curation; investigation; conceptualization. **Savin S. Chand:** Validation; supervision; writing – review and editing; methodology; visualization. **Nargiz Sultanova:** Writing – review and editing; supervision; validation. **Hamish Ramsay:** Writing – review and editing; supervision; methodology; validation. **Krishneel K. Sharma:**

Writing – review and editing. **Lloyd Tahani:** Writing – review and editing; conceptualization.

ACKNOWLEDGEMENT

Alick Haruhiru is grateful to the Australian Government, for funding his PhD, through the Australian Award Scholarship, at Federation University Australia. We also acknowledge funding from the Climate Systems Hub of the Australian Government's National Environmental Science Program (NESP). Open access publishing facilitated by Federation University Australia, as part of the Wiley - Federation University Australia agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ORCID

Alick Haruhiru  <https://orcid.org/0000-0001-7799-5622>

REFERENCES

- Australian Bureau of Meteorology. (2022a) *Madden–Julian Oscillation (MJO)*. Melbourne: Australian Bureau of Meteorology (BOM) website.
- Australian Bureau of Meteorology. (2022b) *Indian Ocean Dipole*. Melbourne: Australian Bureau of Meteorology (BOM) website.
- Bowman, A.W. & Azzalini, A. (1997) *Applied smoothing techniques for data analysis: the kernel approach with S-PLUS illustrations*. London: Oxford University Press, p. 1997.
- Britton, N.R. (1987) Disaster in the South Pacific: impact of tropical cyclone “Namu” on the Solomon Islands, May 1986. *Disasters*, 11(2), 120–133.
- Camargo, S.J., Camp, J., Elsberry, R.L., Gregory, P.A., Klotzbach, P.J., Schreck, C.J. et al. (2019) Tropical cyclone prediction on subseasonal time-scales. *Tropical Cyclone Research and Review*, 8(3), 150–165. Available from: <https://doi.org/10.1016/j.tcr.2019.10.004>
- Camp, J., Wheeler, M.C., Hendon, H.H., Gregory, P.A., Marshall, A.G., Tory, K.J. et al. (2018) Skilful multiweek tropical cyclone prediction in ACCESS-S1 and the role of the MJO. *Quarterly Journal of the Royal Meteorological Society*, 144, 1337–1351.
- Chand, S., Power, S., Walsh, K., Holbrook, N., McInnes, K., Tory, K. et al. (2023) Climate processes and drivers in the Pacific and global warming: a review for informing Pacific planning agencies. *Climatic Change*, 176, 5. Available from: <https://doi.org/10.1007/s10584-022-03467-z>
- Chand, S. & Walsh, K. (2010) The influence of the Madden–Julian Oscillation on tropical cyclone activity in the Fiji region. *Journal of Climate*, 23(4), 868–886. Available from: <https://doi.org/10.1175/2009JCLI3316.1>
- Chand, S.S. & Walsh, E.J.K. (2011a) Forecasting tropical cyclone formation in the Fiji region: a Probit regression approach using Bayesian fitting. *Weather and Forecasting*, 26(2), 150–165. Available from: <https://doi.org/10.1175/2010WAF2222452.1>
- Chand, S.S. & Walsh, K.J.E. (2011b) Influence of ENSO on tropical cyclone intensity in the Fiji region. *Journal of Climate*, 24(15), 4096–4108. Available from: <https://doi.org/10.1175/2011JCLI4178.1>
- Chatterjee, S. & Simonoff, J.S. (2013) *Handbook of regression analysis*. Hoboken, NJ: John Wiley & Sons.
- Deo, A., Chand, S.S., McIntosh, R.D., Prakash, B., Holbrook, N.J., Magee, A. et al. (2022) Severe tropical cyclones over southwest Pacific Islands: economic impacts and implications for disaster risk management. *Climatic Change*, 172(3–4), 1–23. Available from: <https://doi.org/10.1007/s10584-022-03391-2>
- Deo, A., Chand, S.S., Ramsay, H., Holbrook, N.J., McGree, S., Magee, A. et al. (2021) Tropical cyclone contribution to extreme rainfall over southwest Pacific Island nations. *Climate Dynamics*, 56, 3967–3993. Available from: <https://doi.org/10.1007/s00382-021-05680-5>
- Diamond, H.J., Lorrey, A.M. & Renwick, J.A. (2013) A southwest Pacific tropical cyclone climatology and linkages to the El Niño–Southern Oscillation. *Journal of Climate*, 26(1), 3–25. Available from: <https://doi.org/10.1175/JCLI-D-12-00077.1>
- Diamond, H.J., Lorrey, M.A., Levinson, H.D., 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. *International Journal of Climatology*, 32(14), 2240–2250. Available from: <https://doi.org/10.1002/joc.2412>
- Diamond, H.J. & Renwick, J.A. (2015) The climatological relationship between tropical cyclones in the southwest Pacific and the Madden–Julian Oscillation. *International Journal of Climatology*, 35(5), 676–686. Available from: <https://doi.org/10.1002/joc.4012>
- Dowdy, A.J., Qi, L., Jones, D., Ramsay, H., Fawcett, R. & Kuleshov, Y. (2012) Tropical cyclone climatology of the South Pacific Ocean and its relationship to El Niño–Southern Oscillation. *Journal of Climate*, 25(18), 6108–6122. Available from: <https://doi.org/10.1175/JCLI-D-11-00647.1>
- Elsner, J.B. & Jagger, T.H. (2006) Prediction models for annual U.S. hurricane counts. *Journal of Climate*, 19(12), 2935–2952. Available from: <https://doi.org/10.1175/JCLI3729.1>
- Hall, J.D., Matthews, A.J. & Karoly, D.J. (2001) The modulation of tropical cyclone activity in the Australian region by the Madden–Julian Oscillation. *Monthly Weather Review*, 129(12), 2970–2982. Available from: [https://doi.org/10.1175/1520-0493\(2001\)129<2970:TMOTCA>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2970:TMOTCA>2.0.CO;2)
- Haruhiru, A., Chand, S.S., Turville, C. & Ramsay, H. (2022) Tropical cyclone activity in the Solomon Islands region: climatology, variability and trends. *International Journal of Climatology*, 43(1), 593–614. Available from: <https://doi.org/10.1002/joc.7797>
- Haruhiru, A., Chand, S.S., Sultanova, N., Ramsay, H. & Hiriasia, D. H. (2023). A Bayesian framework for seasonal prediction of tropical cyclones affecting the Solomon Islands region. *Weather and Climate Extremes (accepted)*.
- Hendo, H.H. & Salby, M.L. (1959) The life cycle of the Madden–Julian Oscillation. *Journal of the Atmospheric Sciences*, 51(15), 2225–2237.
- Klotzbach, P.J. (2014) The Madden–Julian oscillation's impacts on worldwide tropical cyclone activity. *Journal of Climate*, 27(6), 2317–2330. Available from: <https://doi.org/10.1175/JCLI-D-13-00483.1>
- Kuleshov, Y., Gregory, P., Watkins, A.B. & Fawcett, R.J.B. (2020) Tropical cyclone early warnings for the regions of the Southern Hemisphere: strengthening resilience to tropical cyclones in small Island developing states and least developed countries.

- Natural Hazards*, 104(2), 1295–1313. Available from: <https://doi.org/10.1007/s11069-020-04214-2>
- Lee, C.Y., Camargo, S.J., Vitart, F., Sobel, A.H. & Tippett, M.K. (2018) Subseasonal tropical cyclone genesis prediction and MJO in the S2S dataset. *Weather and Forecasting*, 33(4), 967–988. Available from: <https://doi.org/10.1175/WAF-D-17-0165.1>
- Leroy, A. & Wheeler, M.C. (2008) Statistical prediction of weekly tropical cyclone activity in the Southern Hemisphere. *Monthly Weather Review*, 136(10), 3637–3654. Available from: <https://doi.org/10.1175/2008MWR2426.1>
- Magee, A.D., Lorrey, A.M., Kiem, A.S. & Colyvas, K. (2020) A new Island-scale tropical cyclone outlook for Southwest Pacific nations and territories. *Scientific Reports*, 10(1), 1–13. Available from: <https://doi.org/10.1038/s41598-020-67646-7>
- 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. *International Journal of Climatology*, 38, e774–e795. Available from: <https://doi.org/10.1002/joc.5406>
- 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. *International Journal of Climatology*, 37, 1118–1137. Available from: <https://doi.org/10.1002/joc.5070>
- Magee, A.D., Verdon-Kidd, D.C., Kiem, A.S. & Royle, S.A. (2016) Tropical cyclone perceptions, impacts and adaptation in the southwest Pacific: an urban perspective from Fiji, Vanuatu and Tonga. *Natural Hazards and Earth System Sciences*, 16(5), 1091–1105. Available from: <https://doi.org/10.5194/nhess-16-1091-2016>
- Maier-Gerber, M., Fink, A.H., Riemer, M., Schoemer, E., Fischer, C. & Schulz, B. (2021) Statistical-dynamical forecasting of sub-seasonal North Atlantic tropical cyclone occurrence. *Weather and Forecasting*, 36, 2127–2142. Available from: <https://doi.org/10.1175/waf-d-21-0020.1>
- Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M. & Francis, R.C. (1997) A Pacific Interdecadal climate oscillation with impacts on Salmon production. *Bulletin of the American Meteorological Society*, 78(6), 1069–1079. Available from: [https://doi.org/10.1175/1520-0477\(1997\)078<1069:APICOW>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2)
- Marcoulides, K.M. & Raykov, T. (2019) Evaluation of variance inflation factors in regression models using latent variable modeling methods. *Educational and Psychological Measurement*, 79(5), 874–882. Available from: <https://doi.org/10.1177/0013164418817803>
- Marquardt, D.W. (1963) An algorithm for least-squares estimation of nonlinear parameters. *Society for Industrial and Applied Mathematics*, 11(2), 431–441.
- Maru, E., Shibata, T. & Ito, K. (2018) Statistical analysis of tropical cyclones in the Solomon Islands. *Atmosphere*, 9(6), 1–13. Available from: <https://doi.org/10.3390/atmos9060227>
- McDonnell, K.A. & Holbrook, N.J. (2004) A Poisson regression model approach to predicting tropical cyclogenesis in the Australian/southwest Pacific Ocean region using the SOI and saturated equivalent potential temperature gradient as predictors. *Geophysical Research Letters*, 31(20), 1–5. Available from: <https://doi.org/10.1029/2004GL020843>
- Molteni, F. & Palmer, T.N. (1988) An experimental scheme for the prediction of forecast skill at ECMWF. In: *Workshop on predictability in the medium and extended range, 16–18 May 1988*. Reading: ECMWF.
- Nicholls, N. (1985) Predictability of interannual variations of Australian seasonal tropical cyclone activity. *Monthly Weather Review*, 113, 1144–1149.
- NOAA ESRL. (2022) *Climate indices*. NOAA Working Group on Surface Pressure. Silver Spring.
- Ramsay, H., Richman, M.B. & Leslie, L.M. (2017) The modulating influence of Indian Ocean sea surface temperatures on Australian region seasonal tropical cyclone counts. *Journal of Climate*, 30(13), 4843–4856. Available from: <https://doi.org/10.1175/JCLI-D-16-0631.1>
- Ramsay, H.A., Camargo, S.J. & Kim, D. (2012) Cluster analysis of tropical cyclone tracks in the Southern Hemisphere. *Climate Dynamics*, 39(3), 897–917. Available from: <https://doi.org/10.1007/s00382-011-1225-8>
- Robertson, A.W., Vitart, F. & Camargo, S.J. (2020) Subseasonal to seasonal prediction of weather to climate with application to tropical cyclones. *Journal of Geophysical Research: Atmospheres*, 125(6), e2018JD029375. Available from: <https://doi.org/10.1029/2018JD029375>
- Saji, N., Goswami, B., Vinayachandran, P. & Yamagata, T. (1999) A dipole mode in the Tropical Ocean. *Nature*, 401(6751), 360–363.
- Salinger, M.J., Renwick, J.A. & Mullan, A.B. (2001) Interdecadal Pacific Oscillation and South Pacific climate. *International Journal of Climatology*, 21(14), 1705–1721. Available from: <https://doi.org/10.1002/joc.691>
- Sharma, K.K., Verdon-Kidd, D.C. & Magee, A.D. (2019) Decadal variability of tropical cyclogenesis and decay in the southwest Pacific. *International Journal of Climatology*, 40(5), 2811–2829. Available from: <https://doi.org/10.1002/joc.6368>
- Shrestha, N. (2020) Detecting multicollinearity in regression analysis. *American Journal of Applied Mathematics and Statistics*, 8(2), 39–42. Available from: <https://doi.org/10.12691/ajams-8-2-1>
- Solomon Islands Government. (2014) Rapid assessment of the macro and sectoral impacts of flash floods in the Solomon Islands, April 2014, Honiara.
- Trenberth, K.E. (1997) The definition of El Niño. *Bulletin of the American Meteorological Society*, 78(12), 2771–2777. Available from: [https://doi.org/10.1175/1520-0477\(1997\)078<2771:TDOENO>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2)
- Tu'uholoaki, M., Singh, A., Espejo, A., Chand, S. & Damlamian, H. (2022) Tropical cyclone climatology, variability, and trends in the Tonga region, southwest Pacific. *Weather and Climate Extremes*, 37, 100483. Available from: <https://doi.org/10.1016/j.wace.2022.100483>
- Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C. et al. (2017) The subseasonal to seasonal (S2S) prediction project database. *Bulletin of the American Meteorological Society*, 98(1), 163–173. Available from: <https://doi.org/10.1175/BAMS-D-16-0017.1>
- Vitart, F., Leroy, A. & Wheeler, M.C. (2010) A comparison of dynamical and statistical predictions of weekly tropical cyclone activity in the Southern Hemisphere. *Monthly Weather Review*, 138(9), 3671–3682. Available from: <https://doi.org/10.1175/2010MWR3343.1>
- Vitart, F. & Robertson, A.W. (2018) The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. *Climate and Atmospheric Science*, 1(1), 1–7. Available from: <https://doi.org/10.1038/s41612-018-0013-0>

- Wheeler, M.C. & Hendon, H.H. (2004) An all-season real-time multivariate MJO index: development of an index for monitoring and prediction. *Monthly Weather Review*, 132(8), 1917–1932.
- Wilks, D. (2006) *Statistical methods in the atmospheric sciences*. London: Academic Press.
- Zhang, C. (2005) Madden–Julian Oscillation. *Reviews of Geophysics*, 43, RG2003. Available from: <https://doi.org/10.1029/2004RG000158.1>. **INTRODUCTION**
- Zhang, C. (2013) Madden–Julian oscillation: bridging weather and climate. *Bulletin of the American Meteorological Society*, 94(12), 1849–1870. Available from: <https://doi.org/10.1175/BAMS-D-12-00026.1>

How to cite this article: Haruhiru, A., Chand, S. S., Sultanova, N., Ramsay, H., Sharma, K. K., & Tahani, L. (2023). Subseasonal prediction framework for tropical cyclone activity in the Solomon Islands region. *International Journal of Climatology*, 43(12), 5763–5777. <https://doi.org/10.1002/joc.8173>

APPENDIX A

A.1 | Multicollinearity

To check for multicollinearity, we used the variance inflation factor (VIF). For each predictor variable x_j , we fit the model,

$$x_j = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_{j-1} x_{j-1} + \beta_{j+1} x_{j+1} + \dots + \beta_k x_k + \epsilon, \tag{A1}$$

where β_0 is the model parameter intercept, $\beta_1, \beta_2, \dots, \beta_k$ are the regression coefficients, and ϵ is the variance of the residual term.

Then we compute R_j^2 , the coefficient of determination for this model. If x_j is correlated with the other regressor variables, then R_j^2 will be large as

$$VIF_j = \frac{1}{1 - R_j^2}. \tag{A2}$$

As a rule of thumb, VIFj values greater than 10 are considered evidence of multicollinearity which is problematic. Observing $VIF_j > 10$ is equivalent to observing $R_j^2 > 0.9$ (Chatterjee & Simonoff, 2013; Shrestha, 2020).

A.2 | Model 2: Regression equation

Below is the regression equation for Model 2 and the associated coefficients values (see Table A1),

$$P(1) = \exp(Y') / (1 + \exp(Y')), \tag{A3}$$

$$Y' = -2.855 - 0.0059 \text{ RMM} + 0.3992 \text{ RMM} + 0.088 \text{ Niño3.4} + 0.492 \text{ DMI} + 0.0701 \text{ IPO}. \tag{A4}$$

TABLE A1 Coefficients and VIF for Model 2

Term	Coef	SE Coef	VIF
Constant	-2.855	0.101	
RMM1	-0.0059	0.0871	1.00
RMM2	0.3992	0.0867	1.01
Niño3.4	0.088	0.121	1.57
DMI	0.492	0.299	1.27
IPO	0.0701	0.0948	1.31

TABLE A2 Coefficients and VIF for Model 3

Term	Coef	SE Coef	VIF
Constant	-4.261	0.202	
TC seasonality	15.83	1.45	1.02
RMM1	-0.0006	0.0876	1.01
RMM2	0.3146	0.0842	1.02
Niño3.4	0.136	0.118	1.46
DMI	0.680	0.344	1.14
IPO	0.011	0.104	1.38

A.3 | Model 3: Regression equation

Below is the regression equation for Model 3 and the associated coefficients values (see Table A2),

$$P(1) = \exp(Y') / (1 + \exp(Y')), \tag{A5}$$

$$Y' = -4.261 + 15.83 \text{ TC seasonality} - 0.0006 \text{ RMM} + 0.3146 \text{ RMM} + 0.136 \text{ Niño3.4} + 0.680 \text{ DMI} + 0.011 \text{ IPO}. \tag{A6}$$