Assessing the predictability of MLR models for long-term streamflow using lagged climate indices as predictors: a case study of NSW (Australia)

Rijwana I. Esha and Monzur A. Imteaz

ABSTRACT

The current study aims to assess the potential of statistical multiple linear regression (MLR) techniques to develop long-term streamflow forecast models for New South Wales (NSW). While most of the past studies were concentrated on revealing the relationship between streamflow and single concurrent or lagged climate indices, this study intends to explore the combined impact of large-scale climate drivers. Considering their influences on the streamflow of NSW, several major climate drivers – *IPO (Inter Decadal Pacific Oscillation)/PDO (Pacific Decadal Oscillation), IOD (Indian Ocean Dipole) and ENSO (El Niño-Southern Oscillation)* are selected. Single correlation analysis is exploited as the basis for selecting different combinations of input variables for developing MLR models to examine the extent of the combined impacts of the selected climate drivers on forecasting spring streamflow several months ahead. The developed models with all the possible combinations show significantly good results for all selected 12 stations in terms of Pearson correlation (r), root mean square error (RMSE), mean absolute error (MAE) and Willmott index of agreement (d). For each region, the best model with lower errors provides statistically significant maximum correlation which ranges from 0.51 to 0.65.

Key words | MLR, *NINO3.4*, *PDO*, seasonal forecast, streamflow

Rijwana I. Esha
Monzur A. Imteaz (corresponding author)
Department of Civil and Construction Engineering,
Swinburne University of Technology,
Melbourne, VIC 3122,
Australia
E-mail: *mimteaz@swin.edu.au*

INTRODUCTION

The geographic location and extensive topographic variations present high climatic variability in Australia resulting in even higher inter-annual streamflow variability across the country, which is almost twice that of the rivers in any other part of the world (McMahon *et al.* 1992). As a consequence, the irrigators, agricultural producers, water managers, and planners have to undergo many problems to allocate irrigation water and environmental flows, manage and operate reservoirs, supply municipal water, estimate future hydroelectricity supply, etc. One such severe impact was the overall reduction of gross domestic production (GDP) by 1.6% during the 2002 to 2003 Australian drought (Horridge *et al.* 2005). According to Dutta *et al.* (2006), streamflow forecast is more significant compared to rainfall forecast as it can be predicted with longer lead times, therefore it enables the water users to make decisions at an earlier stage of the year, which ultimately increase the potential of financial benefits.

Two main sources of streamflow forecasting are initial catchment condition (antecedent streamflow, antecedent rainfall, soil moisture, etc.) and climate variables, i.e., rainfall, climate indices, etc. (Robertson & Wang 2009). While comparing to initial catchment condition, remote climate drivers have better predictability of streamflow as the climate indices fluctuate at very low frequencies which can impact the streamflow easily. Moreover, developing streamflow forecasting models incorporating initial catchment conditions is more complicated.

doi: 10.2166/nh.2018.171

Australia, being surrounded by the Pacific, Indian, and Southern Ocean, is greatly influenced by the climatic anomalies originating from the oceans. The impacts of climate indices including the sea level pressure (SLP) and sea surface temperature (SST) anomalies have spatial and seasonal variations. The climate of south-east Australia is influenced by four major climate drivers originating in the Pacific Ocean, the Indian Ocean, and the Southern Ocean, which are *ENSO (El Niño–Southern Oscillation)*, *IPO (Inter Decadal Pacific Oscillation)/PDO (Pacific Decadal Oscillation)*, *SAM (Southern Annular Mode)*, and *IOD (Indian Ocean Dipole)* (Duc *et al.* 2017).

The *ENSO* phenomenon, which results from the large-scale interactions between ocean and atmospheric circulation processes in the equatorial Pacific Ocean, has direct influences on the climate variability over many parts of the world (Ropelewski & Halpert 1987; Kiladis & Diaz 1989; Nicholls *et al.* 1991). El Niño and La Niña events are responsible for the different climatic conditions around the Pacific including eastern Australia (Stone & Auliciems 1992; Latif *et al.* 1997; Nazemosadat & Cordery 1997; Hoerling *et al.* 2001; Chiew 2006; CPTEC 2006). Several studies have revealed the influences of *ENSO* on streamflow throughout Australia (Chiew *et al.* 1998, 2003; Piechota *et al.* 1998; Dettinger & Diaz 2000; Dutta *et al.* 2006; Sharma *et al.* 2015). Chiew *et al.* (1998) and Piechota *et al.* (1998) found that *ENSO*-based (SOI and SST) streamflow predictions in northeast Australia are better than the forecasts from climatology. *ENSO* anomalies are found to be the strongest predictors of seasonal (spring) streamflow and rainfall in some studies (McBride & Nicholls 1983; Robertson & Wang 2009). The study of Chiew & Leahy (2003) explained that spring rainfall and runoff had high correlation (0.3 to 0.5) against winter *SOI* throughout eastern Australia.

The El Niño–Southern Oscillation Modoki events have significant influences on the climate of many parts of the world including Japan, New Zealand, western coast of United States (Ashok *et al.* 2007), Australia (Taschetto & England 2009), and South China (Feng & Li 2011). According to Taschetto & England (2009), classical *ENSO* causes reduction in precipitation in north-east and south-east Australia, whereas *EMI* (*El Niño–Southern Oscillation Modoki Index*) is responsible for reducing precipitation in north-west and northern Australia regions.

Although the dominant source of inter-annual variability in Australian rainfall and streamflow is believed to be the *ENSO* phenomenon, some recent evidence shows that Eastern Australia is also influenced by *IOD* and interdecadal modulation of *ENSO* as a result of the low frequency variability in the Pacific Ocean, which is referred to as *Pacific Decadal Oscillation or PDO* (Westra & Sharma 2009). Cai *et al.* (2011) and Risbey *et al.* (2009) found that *IOD* has an impact on austral winter (June to October) in the southern part of Australia, whereas *ENSO* has a strong influence on austral spring rainfall as a result of the strong covariation of *ENSO* and *IOD*.

Many researchers (e.g., Power *et al.* 1999; Kiem *et al.* 2003; Kiem & Franks 2004) have demonstrated the influence of the *IPO* to be significant on rainfall and streamflow variation on a decadal to multidecadal timescale. King *et al.* (2013) suggested that *IPO* played a significant role in the frequency of major floods during the 1950s, 1970s, and 2010–2011. Verdon *et al.* (2004) explained the enhanced rainfall and streamflow in eastern Australia as being the consequences of the combined impact of *ENSO* (La Niña) and *IPO* negative phase. *IPO* and *PDO* indices are suggested to be highly correlated and useful in explaining various warming and cooling phases in both northern and southern hemispheres (Power *et al.* 1999; Franks 2002b).

Whiting *et al.* (2003) studied the rainfall in Sydney and demonstrated the existence of a greater correlation of annual rainfall in Sydney with the *PDO* index than with *SOI*. A combination of correlation and wavelet-based methods was applied to identify the principal sources of variation in reservoir inflows of Sydney (Westra & Sharma 2009) which found *ENSO*, *PDO*, and *IOD* to be influential.

To date, most research has been focused on identification of suitable predictor variable(s) for forecasting rainfall and streamflow on daily or monthly scales (Sharma 2000; Kiem & Franks 2001; Ruiz *et al.* 2007; Robertson & Wang 2008, 2009; Heller *et al.* 2009; Kirono *et al.* 2010), while very few established the seasonal relationship in different parts of Australia (Piechota *et al.* 1998; Owens *et al.* 2003; Wang *et al.* 2009). The majority of the previous studies investigated the concurrent relation of single climatic variable with daily, monthly, or seasonal streamflow. Even though some studies considered the lagged climate modes, none explored the combined impact of different climate anomalies on seasonal streamflow of eastern Australia. Although a few of the previous studies considered lagged and combined impact of different climate anomalies on seasonal streamflow of eastern Australia, most of those were probabilistic approaches.

Some recent attempts were made by Abbot & Marohasy (2015), Mekanik *et al.* (2013), and Rasel *et al.* (2016) to forecast seasonal rainfalls for Queensland, Victoria, and South Australia, respectively, exploiting linear (MLR) and nonlinear artificial neural network (ANN) techniques and considering the combined influence of different climate modes. Kiem & Verdon-Kidd (2009) explored the relationships between large-scale climate drivers and Victorian streamflow and found that *ENSO* alone is able to explain a very small part of the variation in Victorian streamflow. Therefore, it is important to investigate the combined influence of multiple climate drivers on seasonal streamflow.

Various studies have intended to provide a probabilistic forecast (Piechota *et al.* 1998; Ruiz *et al.* 2007; Robertson & Wang 2009; Wang & Robertson 2011; Duc *et al.* 2017), whereas to solve the water management problems, a deterministic streamflow forecast is more useful as knowing the expected amount of future streamflow helps stakeholders to make more accurate decisions knowing the expected amount of future streamflow. It is to be noted that the Australian Bureau of Meteorology (http://www.bom.gov.au/water/ssf/index.shtml) provides seasonal streamflow forecasting using a Bayesian joint probability (BJP) method, which is again a probabilistic approach. Therefore, the present study aims to investigate the extent of interactions of large-scale multiple climate mode with seasonal streamflow of New South Wales (NSW) with a view to exploiting these relations to forecast seasonal streamflow with a deterministic output.

NSW, located in the south-eastern part of Australia, is affected by frequent droughts and floods, especially in the western and north-eastern regions of the state. Researchers (Erskine & Warner 1988; Franks 2002a; Franks & Kuczera 2002) have quantified the variability of flood risk and attributed these to *ENSO* and *IPO*. Climate variability has serious impacts on the yield of planted crops like wheat, rice, etc. in this agriculturally important region. Although efforts have been made to forecast streamflow and rainfall, none of the current practices provide reliable seasonal streamflow forecast, which can enable water stakeholders to make low-risk decisions at the early stage of the crop period (Khan *et al.* 2005).

This research is intended to provide a successful deterministic prediction of spring streamflow of NSW using the potential large-scale climate drivers with the help of the statistical MLR method. To accomplish this objective, NSW has been divided into four geographically distinct regions and for each region, at several streamflow measurement stations, investigations were conducted to identify the best predictor(s) for forecasting seasonal (spring) streamflow several months ahead in these regions. At this stage, current analysis is conducted for spring season only, considering the outcomes of the previous research works (McBride & Nicholls 1983; Robertson & Wang 2009). A preliminary study of this research on the concurrent analysis of seasonal streamflow and climate indices found spring streamflow provided the most significant correlations compared to other seasons (not shown in this paper). Thus, spring streamflow is expected to have better interactions with climate indices in the lagged correlation analysis.

STUDY AREA AND DATA

NSW situated on the east coast of Australia covers a land area of 880,0642 km^2 and is the most populous state of Australia with a population of 7.5 million, two-thirds of which live in the Greater Sydney Area. The state is bordered on the north by Queensland, on the west by South Australia, on the south by Victoria, and on the east by the Tasman Sea. The two most important features of NSW are the Great Dividing Range (GDR) and Murray Darling Basin (MDB), which accounts for nearly 40% of the value of agricultural production in Australia and 65% of the irrigated land. NSW possesses almost 61% of the water resources plan area of MDB (Department of Industry, NSW Government) while all of Australia's irrigated rice is produced by Murrumbidgee and NSW Murray irrigation regions Murray-Darling Basin Authority (2010).

Economically, NSW is the most important state of Australia as it contributes most of Australia's agricultural production, which is spread throughout the eastern two-thirds of the state. According to ABARE, 80.92% of the state is agricultural land which contributed 23% of the total gross value of agricultural production in Australia in the year 2015–16. Considering the geographical location, regional climatic variation, and agricultural importance, for the current study NSW was divided into four regions: Northern New South Wales (NNSW), Southern New South Wales (SNSW), Central West New South Wales (CWNSW), and Western New South Wales (WNSW), as shown in Figure 1. To explore the spatial variation of influences of different climatic variables for each region several stations were selected based on their long data records and fewer missing values (Table 1). A total of 12 stations were chosen with data records considered to be of appropriate length for the statistical analysis carried out in this study and also considering the predominance factor of coastal rivers in eastern Australia (Verdon *et al.* 2004) as well as the consumptive water use of the locations, where streamflow predictions are important. It can be seen from Figure 1 that the locations of the streamflow stations provide a good spatial coverage of NSW.

For this study, historical streamflow data were collected from the Australian Bureau of Meteorology (http://www.bom.gov.au/waterdata/). Observed monthly streamflow (in cumec) was collected for 102 years, ranging from 1914 to 2015 for nine stations while 99, 101, and 88 years of data were collected for North Cuerindi, Wee Jasper, and Mittangang Crossing stations, respectively. These stations have less than 0.5% missing values, which are filled by the series mean of the streamflow data. Using these data, seasonal mean discharge data are derived for the spring (September–October–November) season.

Considering the aforementioned research works on rainfall and streamflow in this region and the outcomes of single concurrent and lagged correlation analyses, four climate drivers: *ENSO*-based SST anomalies *NINO3.4*, *EMI*, *IPO/PDO*, and DMI (*IOD*) were selected for the MLR analysis.

The *ENSO* phenomenon has two components, SST and atmospheric pressure, which are strongly correlated and can be represented by two types of indicators, the SLP indicator

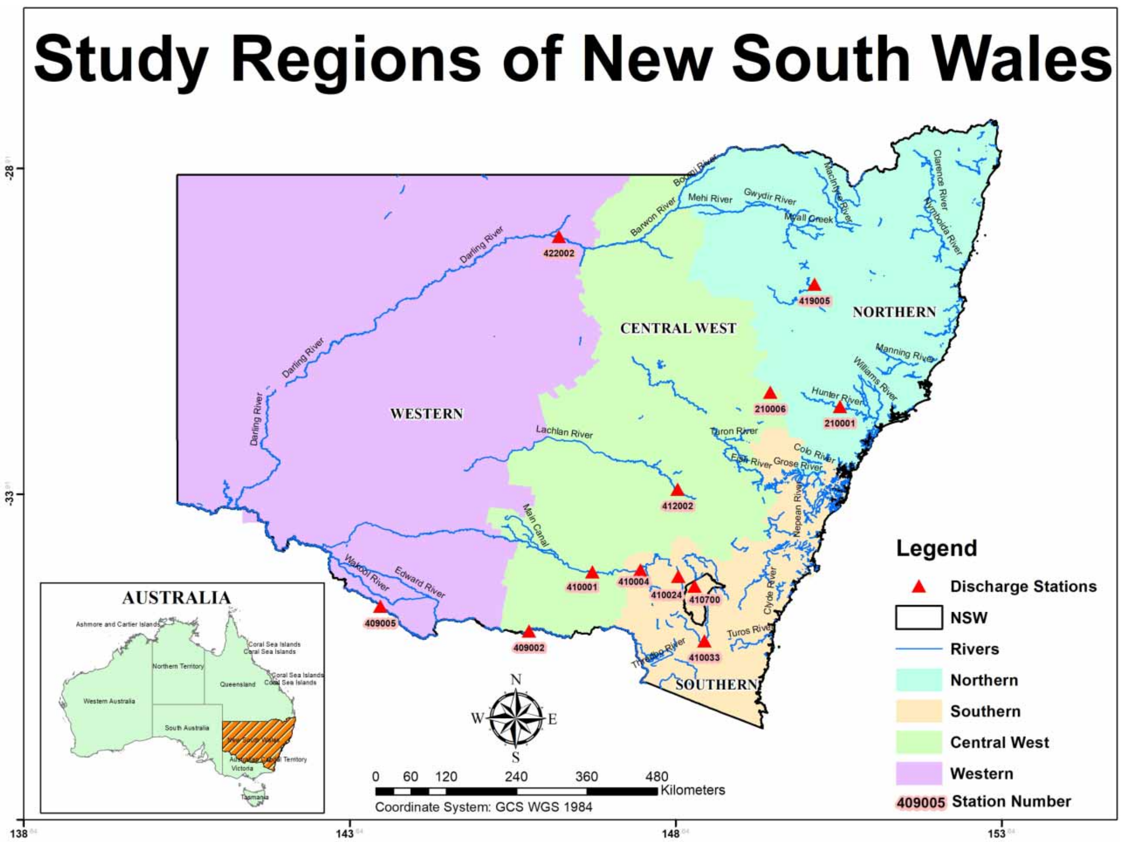


Figure 1 | Locations of the discharge stations in the four study regions of NSW.

Table 1 | Overview of the selected discharge stations

Study region	Station number	Latitude	Longitude	River name	Station name
Northern	210001	−32.56°S	151.17°E	Hunter	Singleton
	210006	−32.34°S	150.10°E	Goulburn	Coggan
	419005	−30.68°S	150.78°E	Namoi	North Cuerindi
Southern	410004	−35.07°S	148.11°E	Murrumbidgee	Gundagai
	410024	−35.17°S	148.69°E	Goodradigbee	Wee Jasper (Kashmir)
	410033	−36.16°S	149.09°E	Murrumbidgee	Mittagang Crossing
	410700	−35.32°S	148.94°E	Cotter	KIOSK
Central	409002	−36.01°S	146.40°E	Murray	Corowa
	410001	−35.10°S	147.37°E	Murrumbidgee	Wagga Wagga
	412002	−33.83°S	148.68°E	Lanchan	Cowra
Western	409005	−35.63°S	144.12°E	Murray	Barham
	422002	−29.95°S	146.86°E	Barwon	Brewarrina

and the SST indicator (Duc *et al.* 2017). Various SST anomalies are available which can be derived from different areas of the equatorial Pacific Ocean (Kiem & Franks 2001). Generally, the SST anomalies are monitored in three geographic regions (Figure 2) of the equatorial Pacific and defined as *NINO3* (5°S–5°N, 150°–90°W), *NINO3.4* (5°S–5°N, 170°–120°W), and *NINO4* (5°S–5°N, 160°–150°W) (Risbey *et al.* 2009) along with *NINO1 + 2* (0–10°S, 90°–80°W). Hanley *et al.* (2003) compared the response of pressure-based and SST-based anomalies to *ENSO* extreme events and found that *NINO3.4* and *NINO4* indices are equally sensitive to El Niño events, whereas SOI is less sensitive to La Niña events than others.

El Niño Modoki is an ocean–atmosphere coupled process, which results in unique tripolar SLP pattern during the evolution, similar to the Southern Oscillation phenomenon of El Niño (Ashok *et al.* 2007). Therefore, this phenomenon is named as El Niño–Southern Oscillation

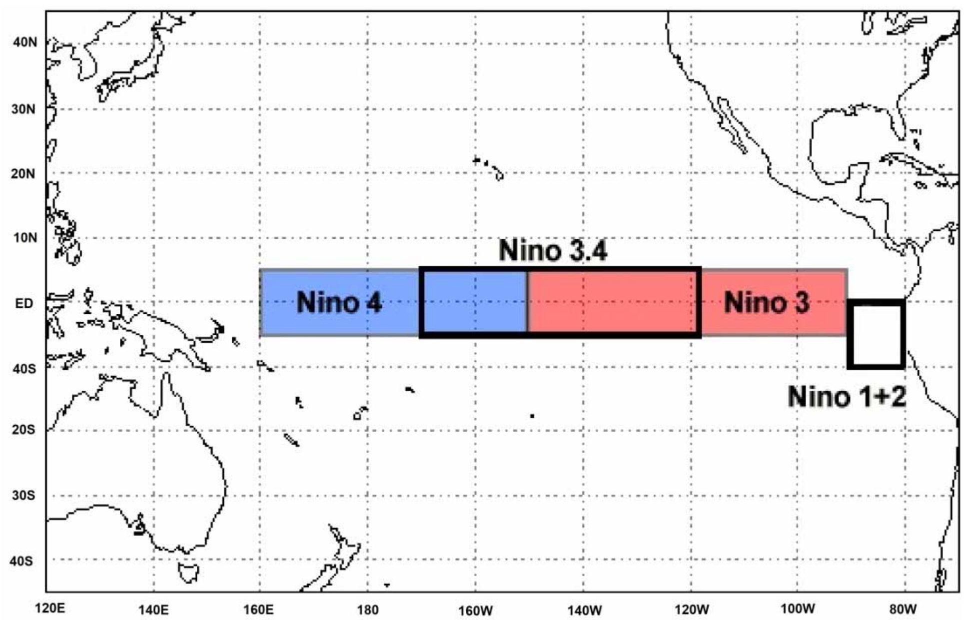


Figure 2 | Map showing *ENSO* regions (*Source:* https://wattsupwiththat.com/2014/11/19/axel-timmermann-and-kevin-trenberth-highlight-the-importance-of-natural-variability-in-global-warming/).

(*ENSO*) Modoki and expressed by the following equation (Ashok *et al.* 2007)

$$EMI = \mathrm{SSTX} - (0.5^{*}\mathrm{SSTY}) - (0.5^{*}\mathrm{SSTZ})$$

where,

$$\mathrm{X} = 165^{\circ}\mathrm{E} - 140^{\circ}\mathrm{W},\ 10^{\circ}\mathrm{S} - 10^{\circ}\mathrm{N},$$
$$\mathrm{Y} = 110^{\circ}\mathrm{W} - 70^{\circ}\mathrm{W},\ 15^{\circ}\mathrm{S} - 5^{\circ}\mathrm{N},$$
$$\mathrm{Z} = 125^{\circ}\mathrm{E} - 145^{\circ}\mathrm{E},\ 10^{\circ}\mathrm{S} - 20^{\circ}\mathrm{N}$$

IOD represents the coupled oceanic–atmospheric variability in the tropical Indian Ocean which is classified by SST anomalies of reverse sign in the east and west (Saji *et al.* 1999; Webster *et al.* 1999). The Dipole Mode Index (*DMI*), which is a measure of the *IOD*, is defined as the difference in SST anomaly between the tropical western Indian Ocean (10°S–10°N, 50°–70°E) and the tropical south-eastern Indian Ocean (10°S–equator, 90°–110°E).

The *IPO* is described as the Pacific *ENSO*-like pattern of SST which is found in the analysis of near-global inter-decadal SST (Folland *et al.* 1999). *IPO* has a cycle of 15–30 years and is characterized with two phases, namely, positive and negative (Salinger *et al.* 2001; Henley *et al.* 2015). While *IPO* is used for the whole Pacific Basin, *PDO* is used for the North Pacific, poleward of 20°N.

The five oceanic and atmospheric climate indices data were obtained from Climate Explorer website (http://climexp.knmi.nl), while the *EMI* data were collected from the website of JAMSTEC (http://www.jamstec.go.jp/frcgc/research/dl/iod/modoki) for a duration of 102 years (1914–2015). An overview of the used climatic variables for the current analysis is presented in Table 2.

METHODOLOGY

In the current study, linear relationships between the selected climatic variables and spring streamflow of NSW were explored using a multiple linear regression technique. At first, single concurrent and lagged correlation analysis was performed to explore dominating indices on spring streamflow. Subsequently, MLR analysis was performed by developing MLR models incorporating combined influences of two indices.

Multiple linear regression

There are several techniques for exploring relationships between two or more parameters. Regression analysis is one of the popular statistical approaches and is highly recommended for this kind of analysis (Ruiz *et al.* 2007; Mekanik *et al.* 2013; Pumo *et al.* 2016). The most commonly used form of linear regression is multiple linear regression analysis. MLR models establish a statistical relationship between two or more explanatory variables and a response variable and provide a linear equation as output which represents the significant correlation among the variables. In every equation, the value of every independent variable

Table 2 | Overview of climate indices and data source

Predictors	Predictor definition	Origin	Data period	Data source
PDO	SSTA anomaly in North Pacific Ocean, (north of 20°N latitude)	Pacific Ocean	1914–2015	ERSST (http://climexp.knmi.nl/)
IPO	SST anomaly in North and South Pacific Ocean (includes south of 20°N latitude)	Pacific Ocean	1914–2015	HadISST1 (http://climexp.knmi.nl/)
NINO3.4	Average SST anomaly over central Pacific Ocean (5°S–5°N, 120°–170°W)	Pacific Ocean	1914–2015	HadISST1 (http://climexp.knmi.nl/)
IOD	West Pole Index (10°S–10°N, 50°–70°E) -East Pole Index (10°S–0°, 90°–110°E)	Indian Ocean	1914–2015	HadISST (http://climexp.knmi.nl/)
EMI	Coupled ocean-atmosphere phenomenon in the tropical Pacific	Pacific Ocean	1914–2015	HadISST (http://www.jamstec.go.jp/frcgc/research/d1/iod/modoki)

(*NINO3.4*, *EMI*, *IOD*, or *IPO/PDO* for the current study) is associated with the value of dependent (spring streamflow for current study) variable. In many studies, climate forecasting has been undertaken using the MLR model, due to the fact that this model comprises many regressors to deal with the time series data base.

In the present study, to evaluate the goodness-of-fit of the models, F-test was used to verify the statistical significance of the overall fit. While developing MLR models, statistical significance of individual parameters of the combined model needs to be evaluated. Among the predictors, verification of multicollinearity is the key stage of MLR modeling. It occurs when predictors themselves are highly correlated, a small change in the data or the model results in remarkable change in parameter estimation. The variance inflation factor (VIF) is used to ascertain the multicollinearity among the predictors. In order to verify multicollinearity among the predictors' tolerance (T) and VIF are used,

$$tolerance = 1 - R^2, \; VIF = \frac{1}{tolerance}$$

where, R^2 is the coefficient of multiple determinations:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

where, SST is the total sum of squares, SSR is the regression sum of squares, and SSE is the error sum of squares. According to Quan *et al.* (2006), a tolerance of less than 0.20–0.10 or a VIF greater than 5–10 indicates a multicollinearity problem.

In order to ensure independence of the residuals error of the model, the Durbin–Watson (DW) test was performed, which assesses the serial correlation between errors. DW parameter has a range of 0 to 4; a value of less than 1 or greater than 3 is certainly a matter of concern (Field 2009).

The performance of the developed MLR models has been assessed by several statistical performance measures which are widely used for the evaluation of regression models. Statistical measures, namely, root mean square error (RMSE), mean absolute error (MAE), Pearson correlation coefficient (r), and Willmott index of agreement (d) are exclusively chosen for this study.

'd' is defined as follows:

$$d = 1 - \frac{\left[\sum |\hat{y}_i - x_i|^2\right]}{\left[\sum \left(|\hat{y}_i - \bar{x}_1| + |x_i - \bar{x}_1|\right)^2\right]}$$

where, $\hat{y}_i$ refers to the predicted value corresponding to i^{th} observation and x_i refers to i^{th} value of observation. The closer the 'd' value to 1, the better the model fits the observations. The development of MLR models and all the relevant statistical calculations were performed using the 'R Studio 3.3.1' software.

RESULTS

Pearson correlation analysis

A detailed study of past research works revealed that different climate anomalies have impacts on the seasonal streamflow of NSW which varies seasonally as well as spatially. Pearson correlation analysis was done for identifying the strength of relation between the seasonal climate anomalies. To analyze the extent of the influences of different climatic variables on seasonal streamflow of NSW, Pearson correlation analysis was applied in two different segments. The first segment took into account the concurrent relationship of climate indices and seasonal streamflow, while the second segment was conducted in order to investigate the lagged relationships of the climate indices and seasonal streamflow. Later on, the outcomes of the second phase served as the basis for selecting the suitable lagged climate indices to use as inputs while developing MLR models.

Concurrent relationships

The linear relationships between spring, summer, autumn, and winter streamflow and climate anomalies of the same seasons were explored by doing the Pearson correlation analysis for each station of the four selected regions of NSW.

In Table 3, concurrent correlations between seasonal streamflow and climate indices are presented, where it is observed that spring (Table 3(d)) shows stronger

Table 3 | Concurrent correlation analyses between seasonal streamflow and climate indices (a) summer, (b) autumn, (c) winter, (d) spring

Region	Station	NINO3	NINO4	SOI	IPO	PDO	EMI	IOD	NINO3.4
(a)									
NNSW	Singleton	−0.31**	−0.35**	0.31**	−0.27**	−0.26**	−0.18	0.18	−0.30**
	North Cuerindi	−0.33**	−0.38**	0.28**	−0.28**	−0.22*	−0.27**	0.1	−0.34**
	Coggan	−0.06	−0.03	−0.06	−0.03	0.14	−0.20*	0.25**	−0.1
SNSW	Mittagang Crossing	−0.19	−0.25*	0.12	−0.05	−0.14	−0.13	−0.12	−0.19
	Kiosk	−0.32**	−0.37**	0.34**	−0.05	−0.22*	−0.27**	0.12	−0.34**
	Gundagai	−0.12	−0.14	0.06	0.02	0.16	−0.31**	0.28**	−0.18
	Wee Jasper	−0.36**	−0.43**	0.33**	−0.12	−0.20*	−0.36**	0.1	−0.39**
CWNSW	Corowa	−0.06	−0.03	−0.06	−0.03	0.14	−0.20*	0.25**	−0.1
	Wagga Wagga	−0.14	−0.16	0.08	0	0.16	−0.32**	0.28**	−0.20*
WNSW	Cowra	−0.15	−0.11	0.11	−0.12	0.02	−0.12	0.27**	−0.15
	Barham	−0.38**	−0.42**	0.38**	−0.20*	−0.25**	−0.28**	0.16	−0.39**
	Brewarrina	−0.39**	−0.51**	0.48**	−0.35**	−0.33**	−0.44**	0.21*	−0.45**
(b)									
NNSW	Singleton	−0.18	−0.26**	0.23*	−0.3**	−0.2*	−0.13	−0.07	−0.22*
	North Cuerindi	−0.09	−0.19*	0.21*	−0.26**	−0.24*	−0.14	−0.11	−0.14
	Coggan	−0.14	−0.24*	0.36**	−0.2*	−0.18	−0.19	−0.07	−0.20*
SNSW	Mittagang Crossing	−0.18	−0.26**	0.31**	−0.22*	−0.33**	−0.13	−0.24*	−0.25*
	Kiosk	−0.18	−0.31**	0.33**	−0.25**	−0.31**	−0.23*	−0.19*	−0.26**
	Gundagai	−0.19	−0.3**	0.27**	−0.24*	−0.28**	−0.23*	−0.15	−0.27**
	Wee Jasper	−0.16	−0.29**	0.31**	−0.25**	−0.28**	−0.23*	−0.20*	−0.24**
CWNSW	Corowa	0.02	0	−0.04	−0.1	0.04	−0.11	0.05	−0.01
	Wagga Wagga	−0.18	−0.3**	0.28**	−0.25**	−0.27**	−0.23*	−0.16	−0.27**
WNSW	Cowra	−0.13	−0.19	0.28**	−0.14	−0.16	−0.16	−0.24*	−0.18
	Barham	−0.04	−0.12	0.12	−0.27**	−0.11	−0.06	−0.06	−0.06
	Brewarrina	−0.23*	−0.31**	0.32**	−0.38**	−0.31**	−0.17	−0.01	−0.28**
(c)									
NNSW	Singleton	−0.19*	−0.27**	0.3**	−0.12	−0.32**	−0.12	−0.15	−0.24*
	North Cuerindi	−0.2*	−0.27**	0.34**	−0.15	−0.24**	−0.21*	−0.18	−0.26**
	Coggan	−0.13	−0.27**	0.37**	−0.06	−0.04	−0.14	−0.32**	−0.18
SNSW	Mittagang Crossing	−0.18	−0.26**	0.31**	−0.22*	−0.33**	−0.13	−0.32**	−0.23*
	Kiosk	−0.2*	−0.34**	0.41**	−0.16	−0.17	−0.12	−0.32**	−0.25**
	Gundagai	−0.16	−0.32**	0.35**	−0.19*	−0.19	−0.13	−0.25**	−0.23*
	Wee Jasper	−0.22*	−0.3**	0.39**	−0.15	−0.08	−0.18	−0.32**	−0.28**
CNSW	Corowa	−0.12	−0.22*	0.34**	−0.17	−0.09	−0.07	−0.28**	−0.16
	Wagga Wagga	−0.15	−0.3**	0.34**	−0.18	−0.16	−0.15	−0.26**	−0.21*
	Cowra	−0.08	−0.23*	0.34**	−0.11	−0.18	−0.14	−0.14	−0.26**
WNSW	Barham	−0.04	−0.12	0.12	−0.27**	−0.11	−0.06	−0.21*	−0.24*
	Brewarrina	−0.15	−0.26**	0.33**	−0.06	−0.2*	−0.21*	−0.20*	−0.41**
(d)									
NNSW	Singleton	−0.35**	−0.42**	0.38**		−0.34**	−0.25**	−0.19*	−0.38**
	North Cuerindi	−0.43**	−0.54**	0.52**		−0.44**	−0.44**	−0.33**	−0.47**
	Coggan	−0.27**	−0.36**	0.33**		−0.28**	−0.3**		−0.32**
SNSW	Wee Jasper	−0.39**	−0.49**	0.46**	−0.23*	−0.26**	−0.36**	−0.50**	−0.43**
	Kiosk	−0.33**	−0.43**	0.42**		−0.37**	−0.22*	−0.46**	−0.34**
	Mittagang Crossing	−0.28**	−0.37**			−0.26**		−0.46**	−0.29**
	Gundagai	−0.26**	−0.36**	0.32**	−0.25**	−0.26**	−0.19*	−0.34**	−0.28**

(*continued*)

Table 3 | continued

Region	Station	NINO3	NINO4	SOI	IPO	PDO	EMI	IOD	NINO3.4
CWNSW	Corowa	−0.28**	−0.36**	0.31**	−0.21*	−0.3**	−0.08	−0.43**	−0.29**
	Wagga Wagga	−0.27**	−0.37**	0.33**	−0.24*	−0.26**	−0.21*	−0.36**	−0.28**
	Cowra	−0.19*	−0.32**	0.26**	−0.23*	−0.19	−0.21*	−0.24*	−0.22*
WNSW	Barham	−0.31**	−0.39**	0.34**		−0.28**		−0.51**	−0.33**
	Brewarrina	−0.36**	−0.47**	0.38**		−0.35**	−0.42**	−0.35**	−0.42**

*Correlation is significant at 5% level. **Correlation is significant at 1% level.

correlations with most of the indices compared to other seasons (Tables 3(a)–3(c)). It is evident from Table 3(d) that all the climate indices (except *IPO*) have significant correlations with spring streamflow for all the selected regions. *IPO* shows significant correlations only for CWNSW and for two stations of SNSW regions which are geographically close to each other. Therefore, it can be anticipated that only in the central-west and southern parts of NSW *IPO* has strong influence on spring streamflow.

Single lagged relationships

For each selected station of the four study regions of NSW, single lag correlation analysis was performed between spring streamflow at year ‘n’ and monthly (December$_{n-1}$ to August$_{n}$) values of each climate variable. The outcomes are presented in Tables 4 and 5.

It was observed from the single lagged analysis that different regions are influenced by different climatic

Table 4 | Pearson correlations (r) of lagged climate indices and spring streamflow NNSW and SNSW

			Lagged months								
Region	Stations	Indices	Dec$_{n-1}$	Jan$_{n}$	Feb$_{n}$	Mar$_{n}$	Apr$_{n}$	May$_{n}$	June$_{n}$	July$_{n}$	Aug$_{n}$
NNSW	Singleton	*NINO3.4*						−0.35**	−0.43**	−0.39**	−0.31**
		PDO		−0.26**	−0.33**	−0.29**	−0.30**	−0.35**	−0.30**	−0.37**	−0.34**
		IOD							−0.22*		
	North Cuerindi	*NINO3.4*						−0.28**	−0.31**	−0.35**	−0.26**
		PDO		−0.19*	−0.22*	−0.21*	−0.24**	−0.30**	−0.23*	−0.32**	−0.31**
		IOD							−0.21*		
	Coggan	*NINO3.4*						−0.34**	−0.45**	−0.46**	−0.39**
		PDO	− 0.21*	−0.27**	−0.26**	−0.22*	−0.29**	−0.33**	−0.31**	−0.41**	−0.36**
SNSW	Wee Jasper	*NINO3.4*						−0.27**	−0.40**	−0.42**	−0.40**
		PDO	− 0.21*	−0.23*		−0.19*	−0.21*			−0.19*	−0.21*
		IOD							−0.21*	−0.29**	−0.32**
		IPO	− 0.23*	−0.23*	−0.23*	−0.23*	−0.23*	−0.23*	−0.23*	−0.23*	−0.23*
		EMI								−0.26**	−0.32**
	Kiosk	*NINO3.4*						−0.19*	−0.25**	−0.35**	−0.33**
		PDO	− 0.29**	−0.31**	−0.30**	−0.32**	−0.26**				−0.24*
		IOD						−0.21*	−0.22*	−0.34**	−0.31**
	Mittagang Crossing	*NINO3.4*						−0.24*	−0.31**	−0.28**	−0.26**
		PDO	− 0.23*	−0.26**	−0.31**	−0.37**	−0.28**				−0.24*
		IOD	− 0.33**			−0.21*	−0.33**	−0.32**	−0.33**	−0.38**	−0.36**
	Gundagai	*NINO3.4*			−0.24*	−0.24**	−0.24**	−0.26**	−0.31**	−0.34**	−0.29**
		PDO	− 0.26**	−0.26**	−0.28**	−0.31**	−0.27**	−0.22*	−0.24**	−0.22*	−0.21*
		IPO		−0.25**	−0.25**	−0.25**	−0.25**	−0.25**	−0.25**	−0.25**	−0.25**
		IOD						−0.19*	−0.23*	−0.25**	−0.25**
		EMI								−0.20*	

*Correlation is significant at 5% level. **Correlation is significant at 1% level.

Table 5 | Pearson correlations (r) of lagged climate indices and spring streamflow CWNSW and WNSW

			Lagged months								
Region	**Stations**	**Indices**	**Dec$_{n-1}$**	**Jan$_n$**	**Feb$_n$**	**Mar$_n$**	**Apr$_n$**	**May$_n$**	**June$_n$**	**July$_n$**	**Aug$_n$**
CNSW	Corowa	*NINO3.4*						−0.22*	−0.25**	−0.33*	−0.29*
		PDO	– 0.24**	−0.24*	−0.27*	−0.31*	−0.25*				
		IOD							−0.22*	−0.35**	−0.40**
		IPO		−0.21*	−0.21*	−0.21*	−0.21*	−0.21*	−0.21*	−0.21*	−0.21*
	Wagga Wagga	*NINO3.4*			−0.23*	−0.23*	−0.22*	−0.25**	−0.30**	−0.35**	−0.29**
		PDO	– 0.25**	−0.25**	−0.27**	−0.29**	−0.25**	−0.20*	−0.22*	−0.20*	−0.19*
		IOD							−0.21*	−0.25**	−0.25**
		IPO		−0.24*	−0.24*	−0.24*	−0.24*	−0.24*	−0.24*	−0.24*	−0.24*
		EMI								−0.20*	v0.21*
	Cowra	*NINO3.4*	– 0.22*	−0.20*	−0.22*	−0.21*	−0.22*	−0.24**	−0.31**	−0.27**	
		PDO				−0.20*	−0.21*	−0.22*	−0.21*	−0.20*	
		IOD						−0.20*			
		IPO		−0.23*	−0.23*	−0.23*	−0.23*	−0.23*	−0.23*	−0.23*	−0.23*
		EMI								−0.19*	−0.22*
WNSW	Barham	*NINO3.4*						−0.20*	−0.26**	−0.35**	−0.33**
		PDO	– 0.24**	−0.25**	−0.26**	−0.28**	−0.20*				
		IOD							−0.30**	−0.44**	−0.46**
	Brewarrina	*NINO3.4*						−0.21*	−0.37**	−0.43**	−0.39**
		PDO					−0.20*	−0.26**	−0.26**	−0.32**	−0.28**
		IOD	– 0.25**							−0.19*	−0.23*
		EMI							−0.24*	−0.31**	−0.41**

*Correlation is significant at 5% level. **Correlation is significant at 1% level.

variables. Lagged *NINO3.4* and *PDO* have significant impacts on the spring streamflow of all four selected regions. Spring streamflows of CWNSW stations and two stations of SNSW (Gundagai and Wee Jasper stations, which are closely located to CWNSW stations) were also influenced by lagged *IPO* (Tables 4 and 5). Almost all the stations had significant correlations with lagged *IOD* indices, whereas lagged *EMI* had impacts on a very limited number of stations.

NINO3.4 shows statistically significant correlations up to a lagged period of four months (April to August) with the correlations ranging from −0.19 to −0.46 (Tables 4 and 5). These findings align with the study of Kirono *et al.* (2010), who found significant correlations for another two *ENSO* indices – *NINO3* and *SOI* as 0.35 and 0.36, respectively, in the same study regions. *SOI* was found to have a correlation of 0.51 with spring streamflow in the study of Chiew & Leahy (2003). Wang *et al.* (2009), also explored the strong impact of *NINO3.4* on spring rainfall up to a lag of two months. Duc *et al.* (2017) explained that *ENSO* indices have a strong impact on rainfall during austral spring. The lagged periods of *IOD* that had significant correlations were not consistent for all the stations, while most of the stations have significant correlations up to three months' lag. It is observed that *EMI* has significant correlation only up to two months' lag (Tables 4 and 5) for five stations (Wee Jasper, Gundagai, Wagga Wagga, Cowra, and Brewarrina) with a maximum significant correlation of −0.41 (Brewarrina station). The lagged relationship of *PDO* was quite different as for some stations it presents more significant correlations with longer lead times. The maximum lead times (up to eight months) with significant correlations for spring streamflow were also obtained for this climate variable of the Pacific Ocean (Tables 4 and 5). This is similar to the assessment of Whiting *et al.* (2003), who stated that *PDO* has greater correlation with annual rainfall of Sydney than that of *SOI*. Latif *et al.* (1997) also showed a strong relationship between *PDO* and Australian summer monsoon. Westra *et al.* (2008) evaluated correlation coefficients between seasonal inflows of a reservoir in Sydney and climate indices, where correlations of spring inflow with *NINO3.4* and *PDO* were found to be −0.17 and −0.19, respectively.

Multiple linear regression analysis

Various models with different combinations of lagged months' indices were analyzed for all 12 stations in order to find out the best forecasting model for each of the four study regions. Eighty-five years (from 1914 to 1998) of streamflow data were selected for the calibration of the models, while the remaining 17 years' (from 1998 to 2015) data were selected for validation in order to assess the future streamflow predictability of the developed MLR models. For all the stations, the best models with lower errors while satisfying the statistical limits were selected. A similar approach for validating the results was applied by Hossain *et al.* (2018), Rasel *et al.* (2016, 2017), and Mekanik *et al.* (2013) while predicting rainfall using climate indices. The Pearson correlation (r) values for both calibration and validation periods were satisfactory, while the correlation values in the validation period were found to be higher than that of the calibration period. F-test was performed to evaluate the best model that fit the population of the sample data, while the t-test was conducted to identify the significance of the individual parameters. The best model for each station with their regression coefficients, VIF, and DW statistics are presented in Table 6.

It can be seen from Table 6 that VIF values for the selected models are close to 1, which means that there is no multi-collinearity problem between the predictors. According to Field (2009), values less than 1 or greater than 3 for DW test will indicate the presence of serial correlations between the model errors. Thus, it can be concluded from the results of Table 6 that the DW test of each selected model satisfies the statistical limits, which also establishes the goodness-of-fit of the models.

DISCUSSION

The results of the MLR analysis depict a clear view of regional variation in influence of combined multiple models throughout the study area. A good number of models combining *PDO* and *NINO3.4* show statistically significant correlations with spring streamflow for the NNSW region which implies the strong impact of these two indices in the northern part of the state. Therefore, for this region the best model for forecasting spring streamflow was obtained with the combination of two months' lagged *NINO3.4* and five months' lagged *PDO* (at Singleton station) showing a significant correlation of 0.41. *PDO-NINO3.4* combined models were predominant in the southern part of the state

Table 6 | Performance test of the best MLR models for calibration and validation period

			Calibration period							Validation period		
Region	Station name	Model	Durbin-Watson	VIF	r	RMSE	MAE	d	r	RMSE	MAE	d
NNSW	Singleton	PDO_{Mar} $NINO3.4_{Jun}$	2.13	1.12	0.41	19.23	13.81	0.51	0.65	12.09	10.98	0.70
	North Cuerindi	PDO_{Jul} $NINO3.4_{Jul}$	1.47	1.16	0.46	8.92	5.98	0.55	0.62	6.92	6.12	0.70
	Coggan	PDO_{Jul} $NINO3.4_{Jul}$	1.99	1.17	0.33	2.15	1.26	0.38	0.61	1.25	1.16	0.67
SNSW	Wee Jasper	IOD_{Jul} $NINO3.4_{Jul}$	1.76	1.17	0.42	7.51	6.04	0.53	0.57	5.47	4.39	0.63
	Kiosk	PDO_{Aug} $NINO3.4_{Ju}$	1.73	1.30	0.45	3.98	3.04	0.57	0.41	4.24	3.74	0.52
	Mittagang Crossing	PDO_{Aug} $NINO3.4_{Jul}$	1.16	1.30	0.35	8.91	7.41	0.44	0.49	9.73	9.30	0.38
	Gundagai	IPO_{Jul} $NINO3.4_{Jul}$	1.83	1.05	0.43	71.40	58.0	0.55	0.51	72.37	66.03	0.40
CWNSW	Corowa	IPO_{Jun} IOD_{Jun}	1.71	1	0.30	139.33	114.87	0.40	0.48	131.01	126.17	0.30
	Wagga Wagga	IPO_{Jul} $NINO3.4_{JuL}$	1.91	1.05	0.43	85.84	69.26	0.55	0.55	80.53	71.44	0.44
	Cowra	PDO_{Mar} $NINO3.4_{Feb}$	1.62	1.44	0.27	35.01	28.91	0.37	0.44	28.04	24.75	0.44
WNSW	Barham	IOD_{Jun} $NINO3.4_{Jun}$	1.80	1.10	0.31	106.05	90.67	0.43	0.44	95.51	82.49	0.37
	Brewarrina	IOD_{Jul} $NINO3.4_{Jul}$	1.74	1.17	0.40	44.47	34.59	0.49	0.56	38.55	36.04	0.57

All correlations are significant at 5% level.

also, where two stations, Kiosk and Mittagang Crossing in SNSW, showed the highest significant correlations (0.41 and 0.49, respectively) for the same combination. On the other hand, the other two stations (Wee Jasper and Mittagang Crossing) of SNSW had maximum significant correlations (0.57 and 0.51, respectively) with *IOD_NINO3.4* and *IPO_NINO3.4* combinations, respectively. This is similar to the findings obtained in the western part of the state because except for Cowra (in CNSW), all the other four stations (Corowa, Wagga Wagga, Barham, and Brewarrina) are evidently influenced by either *IPO* or *IOD* interaction with *NINO3.4*. Both the two stations, Barham and Brewarrina in WNSW, are influenced by the same combination which is *IOD_NINO3.4* having maximum significant correlations of 0.44 and 0.56, respectively. *IPO* is found to be dominant in CWNSW for two stations, as Corowa and Wagga Wagga are influenced by *IPO_IOD* and *IPO_NINO3.4* combined models, respectively. Compared to the other two stations of CWNSW (Corowa and Wagga Wagga), Cowra is located further north, where the *PDO_NINO3.4* models were found to be dominant and satisfying this fact, Cowra also had maximum significant correlation (0.44) for the same combination. It is evident from the results that the *IOD* and *IPO* containing MLR models showed good performance with significant correlation for the southern and western parts (inland) of the state. However, in the coastal eastern part of the state *IOD–NINO3.4* combinations were observed to be less effective, which align with the findings of Pepler *et al.* (2014), who stated that eastern seaboard rainfall is less influenced by tropical SST variability such as *ENSO* and *IOD* than inland because the effect of the *IOD* opposes *ENSO* during the cool season. In the western part of the country *IOD–NINO3.4* and *PDO–NINO3.4* combined models can be confidently used to forecast spring streamflow. Thus, interactions of *NINO3.4*, *IOD*, and *IPO* can be influential in the inland of NSW, whereas *PDO* and *NINO3.4* interactions can be effective in the north-eastern and south-eastern coastal regions of the state.

In order to determine the accuracy of the developed MLR models, validation tests were performed. Table 6 shows the performance statistics of RMSE, MAE, and index of agreement (d) of the best developed models for the calibration and validation periods. It is evident from Figure 3 and Table 6 that there is significant increment (except for Kiosk station) of the correlation values from calibration to validation stage such as, for example, at Singleton station, where correlation value increased from 0.41 in the calibration stage to 0.65 in the validation stage for PDO_{March}–$NINO3.4_{June}$ combination. Again, the highest correlation (r = 0.71) from the current analysis was also obtained for the Singleton station with PDO_{Jan}–$NINO3.4_{May}$ combination in the validation period (not shown in Table 6). However, to get the best predictor model, a few unusual events which were outliers in a box-plot analysis were removed from the calibration and validation periods. Thereby, the ratio of the duration of calibration or validation period to the number of outliers may have affected the corresponding correlation (r) values.

The best predictor models for each of the 12 stations' regions are given below:

$$Q_{Singleton} = 18.81 - 1.81^{*}PDO_{\mathrm{Mar}} - 13.64^{*}NINO3.4_{\mathrm{Jun}}$$
$$Q_{North\ Cuerindi} = 9.58 - 1.67^{*}PDO_{\mathrm{Jul}} - 5.77^{*}NINO3.4_{\mathrm{Jul}}$$
$$Q_{Coggan} = 1.72 - 0.15^{*}PDO_{\mathrm{Jul}} - 1.11^{*}NINO3.4_{\mathrm{Jul}}$$
$$Q_{Wee\ Jasper} = 14.70 - 3.15^{*}IOD_{\mathrm{Jul}} - 4.80^{*}NINO3.4_{\mathrm{Jul}}$$
$$Q_{Kiosk} = 5.91 - 0.54^{*}PDO_{\mathrm{Aug}} - 2.68^{*}NINO3.4_{\mathrm{Jul}}$$
$$Q_{Mittagang\ Crossing} = 13.16 - 1.90^{*}PDO_{\mathrm{Aug}} - 2.54^{*}NINO3.4_{\mathrm{Jul}}$$
$$Q_{Gundagai} = 158.12 - 18.76^{*}IPO_{\mathrm{Jul}} - 42.94^{*}NINO3.4_{\mathrm{Jul}}$$
$$Q_{Corowa} = 262.05 - 37.29^{*}IPO_{\mathrm{Jun}} - 85.21^{*}IOD_{\mathrm{Jun}}$$
$$Q_{Wagga\ Wagga} = 176.47 - 20.15^{*}IPO_{\mathrm{Jul}} - 54.26^{*}NINO3.4_{\mathrm{Jul}}$$
$$Q_{Cowra} = 39.32 - 1.42^{*}PDO_{\mathrm{Mar}} - 11.11^{*}NINO3.4_{\mathrm{Feb}}$$
$$Q_{Barham} = 215.13 - 67.47^{*}IOD_{\mathrm{Jun}} - 39.48^{*}NINO3.4_{\mathrm{Jun}}$$
$$Q_{Brewarrina} = 50.27 - 2.74^{*}IOD_{\mathrm{Jul}} - 32.84^{*}NINO3.4_{\mathrm{Jul}}$$

The predictability of the best MLR models for each of the four regions has been explained through the time series plots of observed and simulated flows in Figure 4.

In Figure 4 some overestimations of the models can be observed during the validation stage which may result from the 'millennium drought' (Bond *et al.* 2008) periods that occurred from 1994 to 2010 over the continent. It was explored by Verdon-Kidd & Kiem (2009) that a combination of climate drivers in the Pacific Ocean (*ENSO*, *PDO*), *IOD* and *SAM* were responsible for the past three droughts in south-east Australia; the 'Federation' drought (1895–1902), the 'World War II' drought (1937–1945), and the 'Big Dry' (1994–2010). Again,

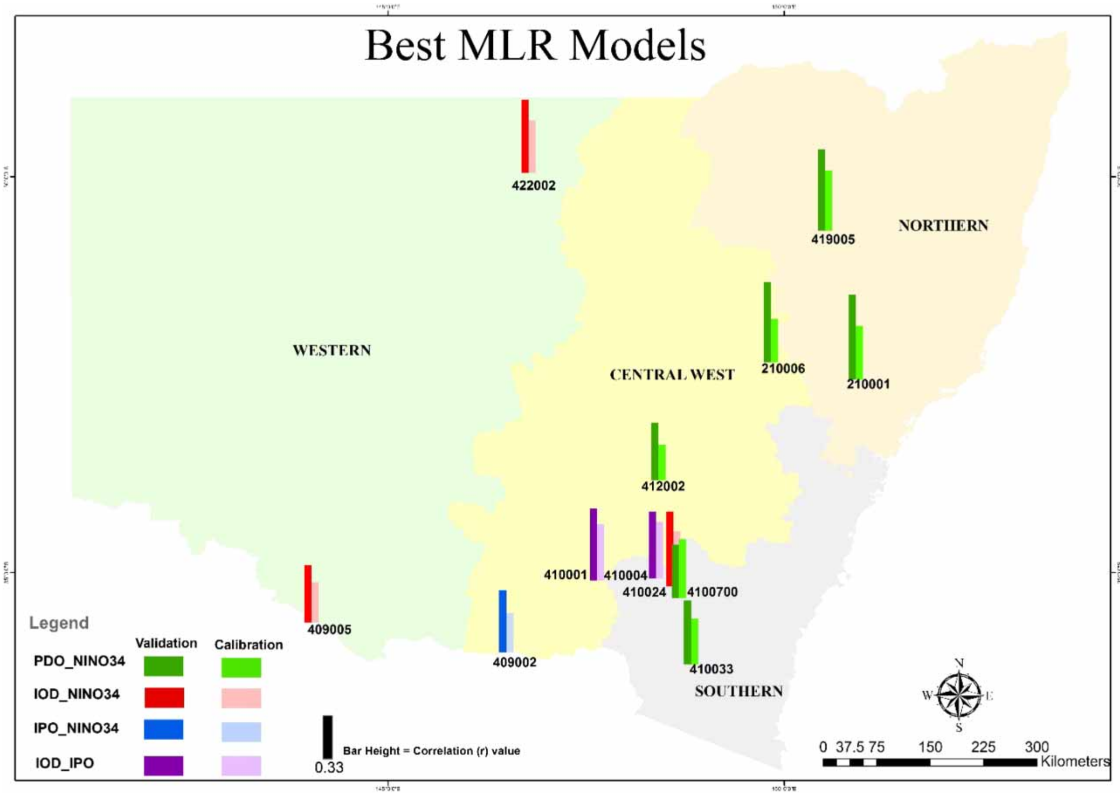


Figure 3 | Influence of best MLR models in terms of Pearson correlation values on the study region.

some streamflow events (e.g., 1950–1960, 1970–1980) are underestimated by the developed models. This may happen due to the failure of capturing the unusual extreme flood events, for instance, the floods of 1954 and 1974. One analysis by Lismore City Council (https://www.lismore.nsw.gov.au/cp_themes/default/page.asp?p = DOC-SVI-55-40-21) revealed that the flood of 1974 is considered to be a once-in-70-year event which resulted from the simultaneous occurrence of La Niña condition and negative *IOD* phase. Apparently, a simple MLR model consisting of only two climate indices (e.g., *NINO3.4* and *PDO*) is not likely to replicate an unusual phenomenon like 'millennium drought' or an extreme flood event like that of 1974. Another reason is that some other climate indices might have been more influential at that time rather than the selected indices for this study.

The capability of the developed models of forecasting spring streamflow with higher accuracy has been ensured as the values of RMSE, MAE, and d in the validation period showed good agreement with the calibration period. The index of agreement (d) for both the calibration and validation periods was close to 0.5, which ensured good forecasting ability of the models. Significant increase in the Pearson correlation values has proved that the combined models have greater skills for predicting streamflow than the single lagged indices. For instance, at Singleton station, during single correlation analysis, $NINO3.4_{June}$ and PDO_{March} showed correlations of -0.43 and -0.29, respectively (Table 4), whereas during MLR analysis, $NINO3.4_{June}$–PDO_{March} combination (Table 6) showed a higher correlation of 0.65 (in validation period). Moreover, while comparing the outcomes of the present research with the previous research studies on forecasting streamflow, it is evident that MLR models of this study showed higher correlations than any of the analyses that used single lagged index. Even for single lagged correlation analysis, in general, the present study showed higher correlations than the previous research (Table 7).

The variations of influences of different climate indices on different study regions of NSW is comparable with the recent study outcomes of Duc *et al.* (2017). They have

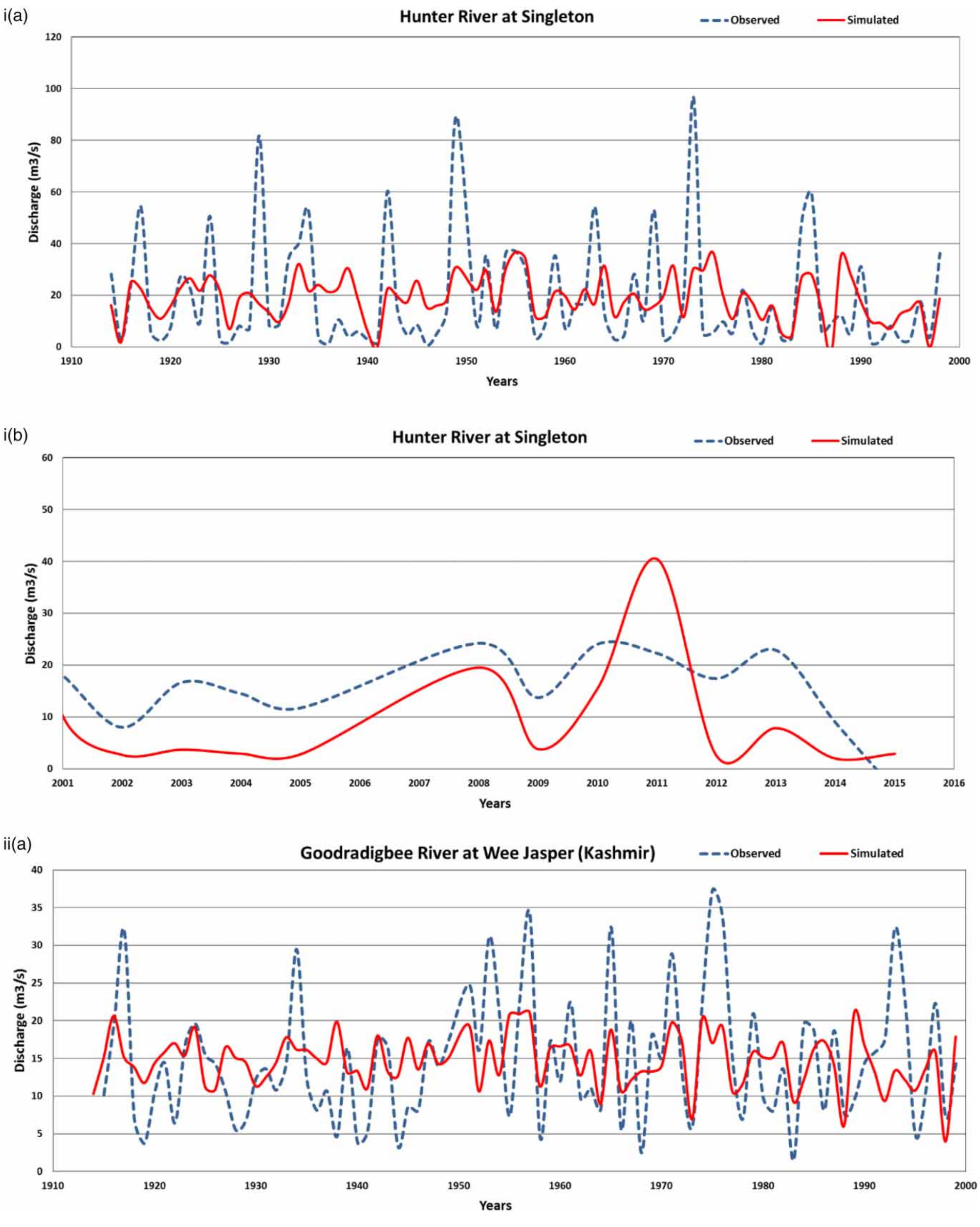
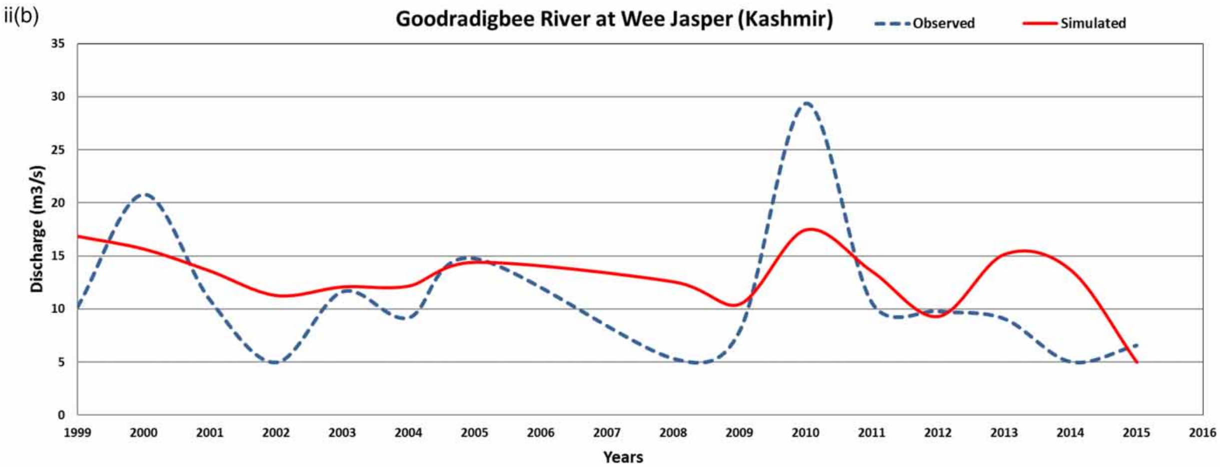
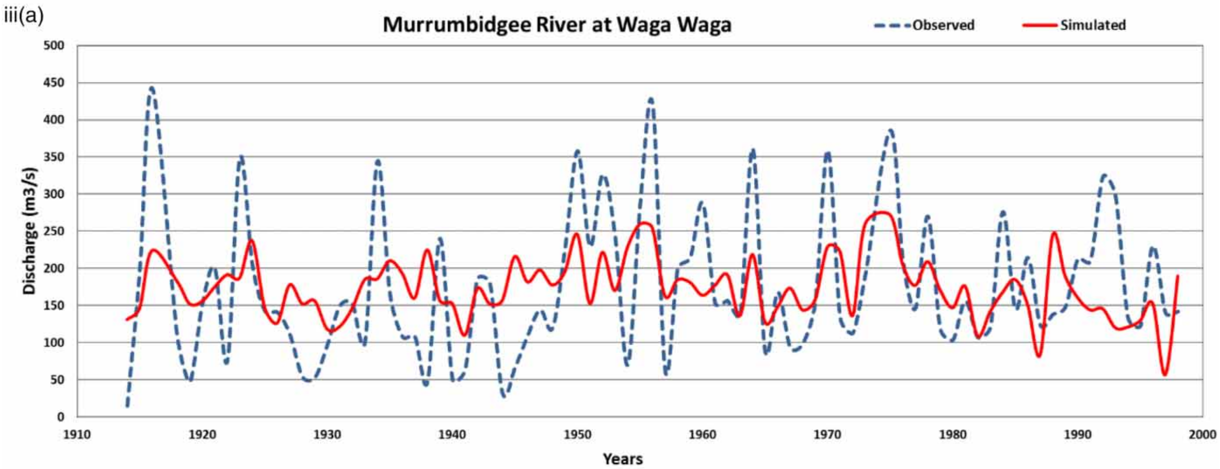


Figure 4 | Comparison between the observed and simulated streamflow during the (a) calibration (1914–1998) and (b) validation (1999–2015) periods for (i) Singleton (NNSW), (ii) Wee Jasper (SNSW), (iii) Wagga Wagga (CWNSW), (iv) Brewarrina (WNSW) stations. (*Continued.*)





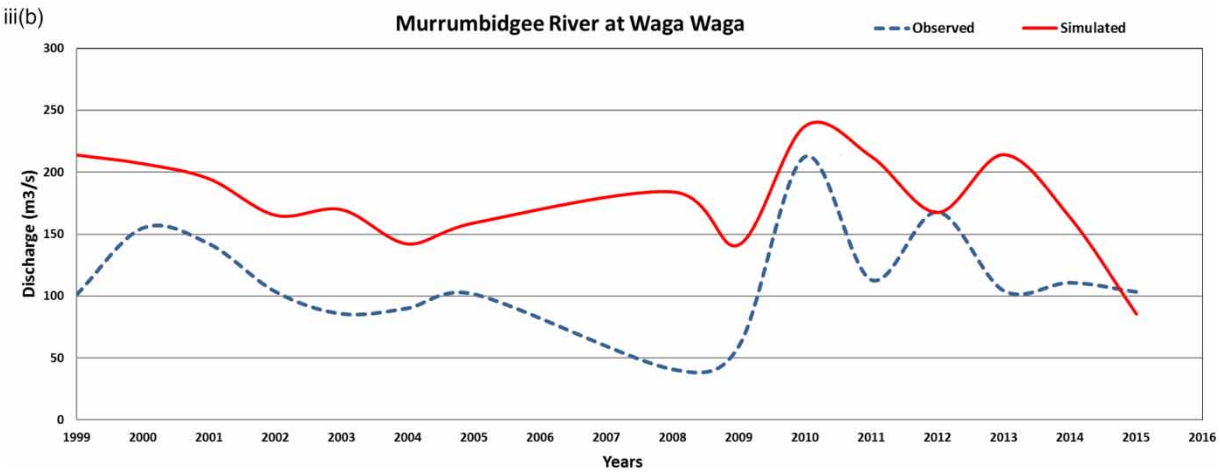


Figure 4 | Continued.

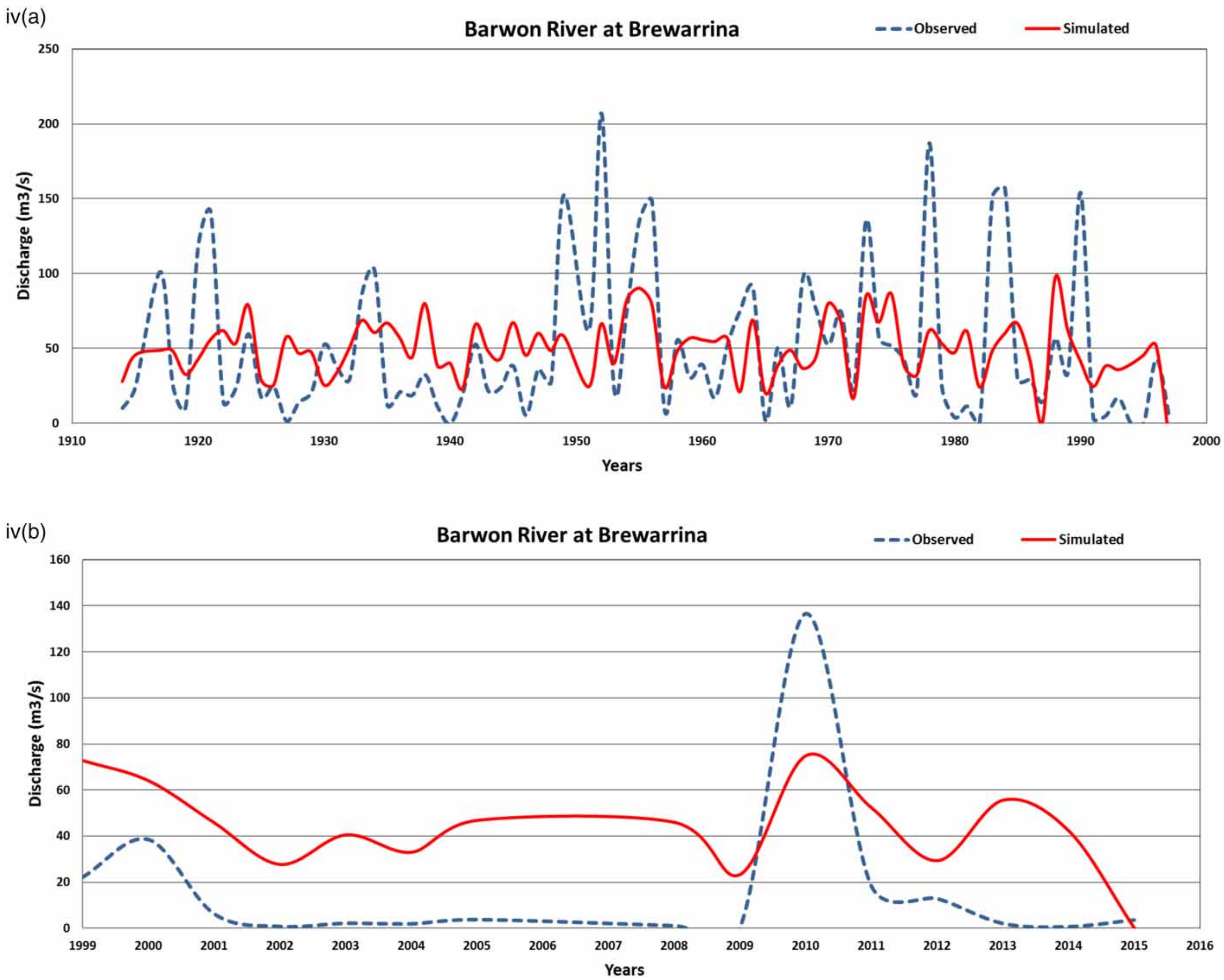


Figure 4 | Continued.

Table 7 | Comparison with the previous studies based on the highest correlations between indices and spring streamflow for south-east Australia

			Current study	
Indices	**Kirono *et al.* (2010)**	**Chiew & Leahy (2003)**	**Single lagged correlation**	**MLR correlation**
NINO3.4	–	–	−0.43[a]	0.65[b]
PDO	–	–	−0.41[c]	
NINO3	0.35[d]	–	0.36[e]	
SOI	0.36[f]	0.51[g]	0.51[h]	

[a]3 months' lagged *NINO3.4*.
[b]PDO_{MAR} and $NINO3.4_{JUN}$.
[c]2 months' lagged *PDO*.
[d]8 months' lagged *NINO3*.
[e]3 months' lagged *NINO3.4*.
[f]12 months' lagged *SOI*.
[g]Winter *SOI*.
[h]2 months' lagged *SOI*.

reported association of climate indices with NSW rainfall using Bayesian model averaging. Among their studied sites, outcomes of the sites which are within 160 km of our selected streamflow stations are similar to our findings. They have reported that single *IPO* cannot impact NSW rainfall significantly, however its association with *ENSO* is significantly influential on rainfall of almost the whole of NSW. The current study also found strong influence of *PDO–NINO3.4* on spring streamflow across almost the whole state (it is to be noted that *IPO* and *PDO* are similar as *IPO* acts on the whole Pacific basin and *PDO* is active in the North Pacific, poleward of 20°N). This finding is strongly supported by the findings of many past studies that suggested *IPO* or *PDO* phases modulate the frequency and magnitude of *ENSO* events (Power *et al.* 1999; Folland *et al.* 2002; Franks 2004; Verdon *et al.* 2004) which is influential on the streamflow volumes of many parts of the world (Kahya & Dracup 1993; Moss *et al.* 1994; Piechota & Dracup 1996; Chiew *et al.* 1998; Piechota *et al.* 1998; Dettinger & Diaz 2000; Kiem & Franks 2001; Wooldridge *et al.* 2001).

In Wagga Wagga, evidence of strong *IOD* influence has been found in the study of Duc *et al.* (2017), which is consistent with our study, as *IOD* combined models performed significantly near this area. Similar outcomes have been obtained for Mittagang Crossing station, where Duc *et al.* (2017) found *ENSO–IOD* combined impact on spring rainfall to be very strong (posterior probability = 1) and the current study has found (not shown here) significant correlation (r = 0.42) with the same combination of indices, i.e., *NINO3.4* and *IOD*.

CONCLUSION

In the current research, the multiple linear regression method was applied with a view to exploring the potential skills of combined multiple climate indicators to forecast the spring streamflow of NSW regions with a longer lead time than the usual practice. Before performing multiple regression analysis, first single correlation analyses were performed to identify potential climate predictors for the region. Through single correlation analysis, several indices (*NINO3.4*, *IOD*, *EMI*, *PDO*, and *IPO*) were found to have strong effects on spring streamflow of NSW with a lagged time of maximum three months. Some indices were found to give significant correlations with a lagged time of more than two months, however, in general, the correlation values decrease with the increase of lagged months. This study, through findings of five effective climate indices for the region, opened an opportunity to study with more than two indices which no one has ever done for this region. For the current study, to achieve better correlations (prediction capability) different combinations of two (out of the five significant) indices were tested in the multiple regression analysis. It was found that the same combination of indices did not turn out to be best for all the stations/regions, which is reasonable as we are dealing with a large region and the further the distance from a particular station the greater the likelihood of being influenced by other indexes. Also, the combined best models' lagged indices are not necessarily from the same month. In general, among the best combined dual indices, *NINO3.4* is found to be significant for all the stations except one (Corowa), *PDO* is more significant towards the north-eastern and south-eastern coastal region, *IPO* is more significant towards the central-south, whereas *IOD* is more significant towards the west of NSW. The best correlation is obtained for Singleton station in NNSW for the *PDO–NINO3.4* combined model with a correlation of 0.65 (in the validation period) for the prediction of spring streamflow with two months' lagged period. It is noteworthy that every time the combined model outperformed the models considering a single climate variable in terms of Pearson correlation (r), it was evidence of better predictive skills of the MLR models.

For this study, selections of the best models were based on the significant correlation values in both calibration and validation stages. However, while looking at the time series comparisons between the observed and simulated streamflow values, it is found that the developed models are unable to capture some unusual events like severe droughts or high floods. A simple multiple linear regression model consisting of only two climate indices is not expected to capture the complex relationships between streamflow and climate drivers very well, and thus is not anticipated to provide a very good match with observed values. Moreover, in fact, rainfall and streamflow are also influenced by some other local and/or regional factors (i.e., temperature, humidity, wind speed, soil moisture, etc.), which are not possible to consider in such regression models. Thus, the extension of this research work will include other non-linear techniques (i.e., ANN or Fuzzy logic) as some researchers (Mekanik *et al.* 2013; Abbot & Marohasy 2015; Rasel *et al.* 2016) successfully explained the non-linear relationship between rainfall and climate drivers using this technique, although they could not provide any output model which could be useful to stakeholders. Since the relationship between streamflow and remote climate drivers is likely to be non-linear (Piechota *et al.* 1998), a non-linear model is expected to give better results than a linear model. Also, an extension of the current research can be to explore the effectiveness of incorporating more than two indices in a multiple regression model. Nevertheless, the developed MLR models have the potential to provide indications on the possibility of getting increased or decreased amounts of streamflow and expected magnitude in the future season. Currently, water users in Australia get the seasonal predictions of streamflow just at the beginning of the season, which do not give them enough time for prudent decision making. Moreover, those predictions are stochastic, i.e., the users do not get any estimation of expected magnitude. The developed MLR models for the study area are expected to provide water users and planners with some insight which will enable them to take tactical cropping decisions three months in advance. This sort of study is mainly based on regional climate index/indices applicable for a region. However, a similar concept can be applied to other regions if any such index/indices are found to be effective for other regions.

REFERENCES

Abbot, J. & Marohasy, J. 2015 Using lagged and forecast climate indices with artificial intelligence to predict monthly rainfall in the Brisbane Catchment, Queensland, Australia. *International Journal of Sustainable Development and Planning* **10** (1), 29–41.

Ashok, K., Behera, S. K., Rao, S. A., Weng, H. & Yamagata, T. 2007 El Niño Modoki and its possible teleconnection. *Journal of Geophysical Research: Oceans* **112** (C11). doi:10.1029/2006JC003798.

Australian Bureau of Agricultural and Resource Economics and Sciences http://www.agriculture.gov.au/abares/research-topics/aboutmyregion/nsw#fisheries-sector (retrieved 25 April 2018).

Bond, N. R., Lake, P. S. & Arthington, A. H. 2008 The impacts of drought on freshwater ecosystems: an Australian perspective. *Hydrobiologia* **600** (1), 3–16.

Cai, W., Van Rensch, P., Cowan, T. & Hendon, H. H. 2011 Teleconnection pathways of ENSO and the IOD and the mechanisms for impacts on Australian rainfall. *Journal of Climate* **24** (15), 3910–3923.

Chiew, F. H. 2006 Estimation of rainfall elasticity of streamflow in Australia. *Hydrological Sciences Journal* **51** (4), 613–625.

Chiew, F. H. S. & Leahy, M. J. 2003 Inter-decadal Pacific oscillation modulation of the impact of El Niño/Southern Oscillation on Australian rainfall and streamflow. In: *MODSIM 2003: International Congress on Modelling and Simulation*, 1, 100–105.

Chiew, F. H., Piechota, T. C., Dracup, J. A. & McMahon, T. A. 1998 El Nino/Southern Oscillation and Australian rainfall, streamflow and drought: links and potential for forecasting. *Journal of Hydrology* **204** (1–4), 138–149.

CPTEC (Centro de Previsão de Tempo e Estudos Climáticos) 2006 El Niño e La Niña. http://wwwcptecinpebr/enos/artigosshtml.

Department of Industry, New South Wales Government https://www.water.nsw.gov.au/water-management/catchments-old/murray-darling-basin-old (retrieved 25 April 2018).

Dettinger, M. D. & Diaz, H. F. 2000 Global characteristics of stream flow seasonality and variability. *Journal of Hydrometeorology* **1** (4), 289–310.

Duc, H. N., Rivett, K., MacSween, K. & Le-Anh, L. 2017 Association of climate drivers with rainfall in New South Wales, Australia, using Bayesian Model Averaging. *Theoretical and Applied Climatology* **127** (1–2), 169–185.

Dutta, S. C., Ritchie, J. W., Freebairn, D. M. & Abawi, G. Y. 2006 Rainfall and streamflow response to El Niño Southern Oscillation: a case study in a semiarid catchment, Australia. *Hydrological Sciences Journal* **51** (6), 1006–1020.

Erskine, W. D. & Warner, R. F. 1988 Geomorphic effects of alternating flood-and drought-dominated regimes on NSW coastal rivers. In: *Fluvial Geomorphology of Australia* (R. F. Warmer, ed.). Academic Press, Sydney, pp. 223–244.

Feng, J. & Li, J. 2011 Influence of El Niño Modoki on spring rainfall over south China. *Journal of Geophysical Research: Atmospheres* **116**, D13102. doi:10.1029/2010JD015160.

Field, A. 2009 *Discovering Statistics Using SPSS*. Sage publications, London.

Folland, C. K., Parker, D. E., Colman, A. & Washington, R. 1999 *Beyond El Nino: Decadal and Interdecadal Climate Variability*. Springer Science & Business Media, Berlin, pp. 73–102.

Folland, C. K., Renwick, J. A., Salinger, M. J. & Mullan, A. B. 2002 Relative influences of the interdecadal Pacific oscillation and *ENSO* on the South Pacific convergence zone. *Geophysical Research Letters* **29** (13), 21–24.

Franks, S. W. 2002a Identification of a change in climate state using regional flood data. *Hydrology and Earth System Sciences* **6** (1), 11–16.

Franks, S. W. 2002b Assessing hydrological change: deterministic general circulation models or spurious solar correlation? *Hydrological Processes* **16** (2), 559–564.

Franks, S. W. 2004 Multi-decadal climate variability, New South Wales, Australia. *Water Science and Technology* **49** (7), 133–140.

Franks, S. W. & Kuczera, G. 2002 Flood frequency analysis: evidence and implications of secular climate variability, New South Wales. *Water Resources Research* **38** (5), 20–27.

Hanley, D. E., Bourassa, M. A., O'Brien, J. J., Smith, S. R. & Spade, E. R. 2003 A quantitative evaluation of *ENSO* indices. *Journal of Climate* **16** (8), 1249–1258.

Heller, G. Z., Stasinopoulos, D. M., Rigby, R. A. & Van Ogtrop, F. 2009 Randomly stopped sum models: a hydrological application. In: *International Workshop on Statistical Modelling* (James G. Booth, ed.). Cornell University, New York, pp. 156–163.

Henley, B. J., Gergis, J., Karoly, D. J., Power, S., Kennedy, J. & Folland, C. K. 2015 A tripole index for the Interdecadal Pacific Oscillation. *Climate Dynamics* **45** (11–12), 3077–3090.

Hoerling, M. P., Whitaker, J. S., Kumar, A. & Wang, W. 2001 The midlatitude warming during 1998–2000. *Geophysical Research Letters* **28** (5), 755–758.

Horridge, M., Madden, J. & Wittwer, G. 2005 The impact of the 2002–2003 drought on Australia. *Journal of Policy Modeling* **27** (3), 285–308.

Hossain, I., Rasel, H. M., Imteaz, M. A. & Mekanik, F. 2018 Long-term seasonal rainfall forecasting: efficiency of linear modelling technique. *Environmental Earth Sciences* **77** (280), 1–10.

Kahya, E. & Dracup, J. A. 1993 US streamflow patterns in relation to the El Niño/Southern Oscillation. *Water Resources Research* **29** (8), 2491–2503.

Khan, S., Robinson, D., Beddek, R., Wang, B., Dassanayake, D. & Rana, T. 2005 *Hydro-climatic and Economic Evaluation of Seasonal Climate Forecasts for Risk Based Irrigation Management*. CSIRO Land and Water Technical report No. 5/04.

Kiem, A. S. & Franks, S. W. 2001 On the identification of *ENSO*-induced rainfall and runoff variability: a comparison of methods and indices. *Hydrological Sciences Journal* 46 (5), 715–727.

Kiem, A. S. & Franks, S. W. 2004 Multi-decadal variability of drought risk, eastern Australia. *Hydrological Processes* **18** (11), 2039–2050.

Kiem, A. S. & Verdon-Kidd, D. C. 2009 Climatic drivers of Victorian streamflow: is *ENSO* the dominant influence? *Australian Journal of Water Resources* **13** (1), 17–29.

Kiem, A. S., Franks, S. W. & Kuczera, G. 2003 Multi-decadal variability of flood risk. *Geophysical Research Letters* **30** (2).

Kiladis, G. N. & Diaz, H. F. 1989 Global climatic anomalies associated with extremes in the Southern Oscillation. *Journal of Climate* **2** (9), 1069–1090.

King, A. D., Alexander, L. V. & Donat, M. G. 2013 Asymmetry in the response of eastern Australia extreme rainfall to low-frequency Pacific variability. *Geophysical Research Letters* **40** (10), 2271–2277.

Kirono, D. G., Chiew, F. H. & Kent, D. M. 2010 Identification of best predictors for forecasting seasonal rainfall and runoff in Australia. *Hydrological Processes* **24** (10), 1237–1247.

Latif, M., Kleeman, R. & Eckert, C. 1997 Greenhouse warming, decadal variability or El Nino? An attempt to understand the anomalous 1990s. *Journal of Climate* **10** (9), 2221–2239.

McBride, J. L. & Nicholls, N. 1983 Seasonal relationships between Australian rainfall and the Southern Oscillation. *Monthly Weather Review* **111** (10), 1998–2004.

McMahon, T. A., Finlayson, B. L., Haines, A. T. & Srikanthan, R. 1992 *Global Runoff: Continental Comparisons of Annual Flows and Peak Discharges*. Catena Verlag, Germany.

Mekanik, F., Imteaz, M. A., Gato-Trinidad, S. & Elmahdi, A. 2013 Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. *Journal of Hydrology* **503**, 11–21.

Moss, M. E., Pearson, C. P. & McKerchar, A. I. 1994 The Southern Oscillation index as a predictor of the probability of low streamflows in New Zealand. *Water Resources Research* **30** (10), 2717–2723.

Murray–Darling Basin Authority 2010 *Guide to the Proposed Basin Plan: Technical Background*. Murray–Darling Basin Authority, Canberra. Vol 2–0-12.

Nazemosadat, M. J. & Cordery, I. 1997 The influence of geopotential heights on New South Wales rainfall. *Meteorology and Atmospheric Physics* **63** (3), 179–193.

Nicholls, N., Glantz, M. H. & Katz, R. W. 1991 *Teleconnections Linking Worldwide Climate Anomalies: Scientific Basis and Societal Impact*. Cambridge University Press, Cambridge.

Owens, D. T., Clarkson, N. M. & Clewett, J. F. 2003 Persistence of Australian streamflow and its application to seasonal forecasts. MODSIM 2003. In: *International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand*, July 2003, 01 (A02/09).

Pepler, A., Timbal, B., Rakich, C. & Coutts-Smith, A. 2014 Indian Ocean Dipole overrides *ENSO*'s influence on cool season rainfall across the Eastern Seaboard of Australia. *Journal of Climate* **27** (10), 3816–3826.

Piechota, T. C. & Dracup, J. A. 1996 Drought and regional hydrologic variation in the United States: associations with the El Niño-Southern Oscillation. *Water Resources Research* **32** (5), 1359–1373.

Piechota, T. C., Chiew, F. H., Dracup, J. A. & McMahon, T. A. 1998 Seasonal streamflow forecasting in eastern Australia and the El Niño–Southern Oscillation. *Water Resources Research* **34** (11), 3035–3044.

Power, S., Casey, T., Folland, C., Colman, A. & Mehta, V. 1999 Inter-decadal modulation of the impact of *ENSO* on Australia. *Climate Dynamics* 15 (5), 319–324.

Pumo, D., Viola, F. & Noto, L. V. 2016 Generation of natural runoff monthly series at ungauged sites using a regional regressive model. *Water* **8** (5), 209.

Quan, X., Hoerling, M., Whitaker, J., Bates, G. & Xu, T. 2006 Diagnosing sources of US seasonal forecast skill. *Journal of Climate* **19** (13), 3279–3293.

Rasel, H. M., Imteaz, M. A. & Mekanik, F. 2016 Investigating the influence of remote climate drivers as the predictors in forecasting South Australian spring rainfall. *International Journal of Environmental Research* **10** (1), 1–12.

Rasel, H. M., Imteaz, M. A. & Mekanik, F. 2017 Multiple regression modelling approach for rainfall prediction using large-scale climate indices as potential predictors. *International Journal of Water* **11** (3), 209–225.

Risbey, J. S., Pook, M. J., McIntosh, P. C., Wheeler, M. C. & Hendon, H. H. 2009 On the remote drivers of rainfall variability in Australia. *Monthly Weather Review* **137** (10), 3233–3253.

Robertson, D. E. & Wang, Q. J. 2008 *An Investigation Into the Selection of Predictors and Skill Assessment Using the*

Bayesian Joint Probability (BJP) Modelling Approach to Seasonal Forecasting of Streamflows. Water for a Healthy Country flagship report, CSIRO Land and Water, Canberra.

Robertson, D. E. & Wang, Q. J. 2009 A Bayesian joint probability approach to seasonal prediction of streamflows: Predictor selection and skill assessment. In: *H2009: 32nd Hydrology and Water Resources Symposium, Newcastle: Adapting to Change*. Engineers Australia, p. 1545.

Ropelewski, C. F. & Halpert, M. S. 1987 Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Monthly Weather Review* **115** (8), 1606–1626.

Ruiz, J. E., Cordery, I. & Sharma, A. 2007 Forecasting streamflows in Australia using the tropical Indo-Pacific thermocline as predictor. *Journal of Hydrology* **341** (3), 156–164.

Saji, N. H., Goswami, B. N., Vinayachandran, P. N. & Yamagata, T. 1999 A dipole mode in the tropical Indian Ocean. *Nature* **401** (6751), 360–363.

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.

Sharma, A. 2000 Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: part 3–A nonparametric probabilistic forecast model. *Journal of Hydrology* **239** (1), 249–258.

Sharma, S., Srivastava, P., Fang, X. & Kalin, L. 2015 Long-range hydrologic forecasting in El Niño Southern Oscillation-affected coastal watersheds: comparison of climate model and weather generator approach. *Journal of Hydrologic Engineering* **20** (12), 06015006.

Stone, R. & Auliciems, A. 1992 SOI phase relationships with rainfall in eastern Australia. *International Journal of Climatology* **12** (6), 625–636.

Taschetto, A. S. & England, M. H. 2009 El Niño Modoki impacts on Australian rainfall. *Journal of Climate* **22** (11), 3167–3174.

Verdon, D. C., Wyatt, A. M., Kiem, A. S. & Franks, S. W. 2004 Multidecadal variability of rainfall and streamflow: eastern Australia. *Water Resources Research* **40** (10).

Verdon-Kidd, D. C. & Kiem, A. S. 2009 Nature and causes of protracted droughts in southeast Australia: comparison between the Federation, WWII, and Big Dry droughts. *Geophysical Research Letters* **36** (22).

Wang, Q. J. & Robertson, D. E. 2011 Multisite probabilistic forecasting of seasonal flows for streams with zero value occurrences. *Water Resources Research* **47** (2).

Wang, Q. J., Robertson, D. E. & Chiew, F. H. S. 2009 A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites. *Water Resources Research* **45** (5).

Webster, P. J., Moore, A. M., Loschnigg, J. P. & Leben, R. R. 1999 Coupled ocean-atmosphere dynamics in the Indian Ocean during 1997–98. *Nature* **401** (6751), 356–360.

Westra, S., Sharma, A., Brown, C. & Lall, U. 2008 Multivariate streamflow forecasting using independent component analysis. *Water Resources Research* **44** (2).

Westra, S. & Sharma, A. 2009 Probabilistic estimation of multivariate streamflow using independent component analysis and climate information. *Journal of Hydrometeorology* **10** (6), 1479–1492.

Whiting, J. P., Lambert, M. F. & Metcalfe, A. V. 2003 Modelling persistence in annual Australia point rainfall. *Hydrology and Earth System Sciences Discussions* **7** (2), 197–211.

Wooldridge, S. A., Franks, S. W. & Kalma, J. D. 2001 Hydrological implications of the Southern Oscillation: variability of the rainfall-runoff relationship. *Hydrological Sciences Journal* **46** (1), 73–88.

First received 17 October 2017; accepted in revised form 25 May 2018. Available online 22 June 2018