Supplement of Adv. Stat. Clim. Meteorol. Oceanogr., 11, 123–132, 2025
https://doi.org/10.5194/ascmo-11-123-2025-supplement

Supplement of

Forecasting springtime rainfall in southeastern Australia using empirical orthogonal functions and neural networks

Stjepan Marčelja

Correspondence to: Stjepan Marčelja (stjepan.marcelja@anu.edu.au)

S1 Geographical regions and rainfall in this study

Figure S1
Murray-Darling Basin (green), SE Australia (violet) and the state of Victoria borders at the south of the continent are shown in the map provided by the Bureau of Meteorology.
http://www.bom.gov.au/climate/change/about/temp_timeseries.shtml

Region	Average Sept/Oct rainfall [mm] 1991-2020	Correlation with Murray-Darling	Correlation with SE Australia	Correlation with Victoria
Victoria	114.8	0.86	0.97	
Murray-Darling	73.4		0.91	0.86
SE Australia	108.5	0.91		0.97

Table S1
Average values and correlations between rainfall in the regions of this study.

S2 List of predictors

Victoria
Predictors with June data:
Niño3.4 June
Ei-Si June
Indian Ocean EOF2 June
Australia temperature June
Murray-Darling temperature June

Predictors with July data:
Niño3.4 July
Indian Ocean EOF2 July
Ei-Si July
Ei-Si June
Ei July

Murray-Darling Basin
Predictors with June data:
Ei-Si May
Ei-Si June
Niño3.4 January
Niño3.4 February
Niño3.4 June
Indian Ocean Dipole May
Indian Ocean Dipole June
Indian Ocean East May
Indian Ocean East June
Australia temperature June
SE Australia temperature June

Predictors with July data:
Niño3.4 July
Indian Ocean EOF1 July

SE Australia
Predictors with June data:
Niño3.4 June
Ei-Si June
Australia temperature June
Murray-Darling temperature June

Predictors with July data:
Niño3.4 July
Indian Ocean EOF1 July
Ei-Si June
Ei-Si July
Ei July
Si July

Predictors with August data (used for November hindcast/forecast):
Si August
Niño South August (170ºW to 120ºW, 30ºS to 20ºS)
Tahiti SLP June
South Pacific EOF1 August (140ºE to 100ºW, 40ºS to 5ºN)
South Pacific EOF2 August
South Pacific EOF3 August

S3 Testing longer temporal periods

For all hindcasting/forecasting, data from the period 2000-2024 were used as neural network inputs. This was a strictly empirical decision, based on the agreement in the validation periods. The inclusion of older data does not lead to large changes, but the errors increase with the increasing fraction of the older data.

Input data period	RMSE (Victoria with June data)
2000-2024	24.7 mm (see image main text)
1975-2024	29.5 mm
1950-2024	33.6 mm

Table S2
RMSE obtained with networks trained using different data periods.

Examples from the Table S1 are shown in the figure S2.

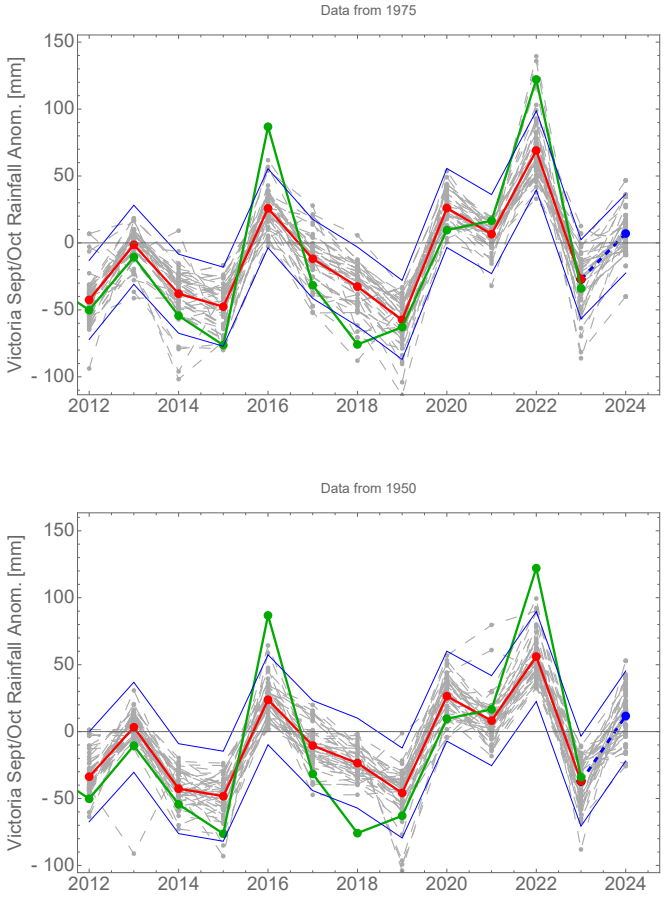


Figure S2
Hindcasting rainfall in September/October period in Victoria based on June data starting from 1975 (top) and 1950 (bottom). Result with the data starting from 2000 is shown as Fig. 4a in the main text.

Better performance of the hindcasts that use only relatively recent data is based on the changes in the correlation between the predictors and spring rainfall. Correlation between the best predictors and springtime rainfall has been gradually increasing since about 1975 (Fig. S3).

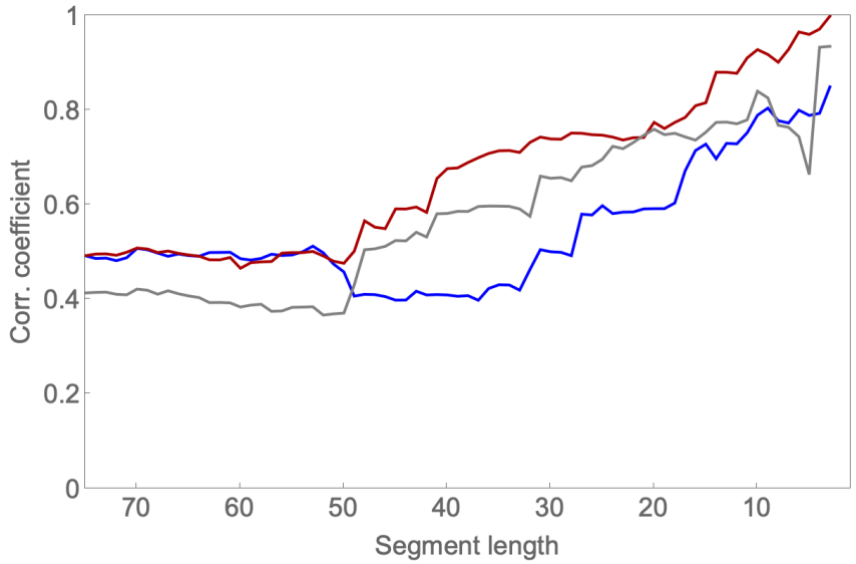


Figure S3
Correlation coefficient between the best predictors evaluated with segments of different length shown on the abscissa and the SE Australia September/October rainfall in the corresponding period. All segments end in 2024.
Red: Indian ocean EOF1;
Grey: IOD;
Blue: Niño 3.4.

Increasing correlation shown in Fig. S3 is also seen in a direct comparison between winter EOF1 in the Indian Ocean and SE Australia springtime rainfall (Fig. S4).

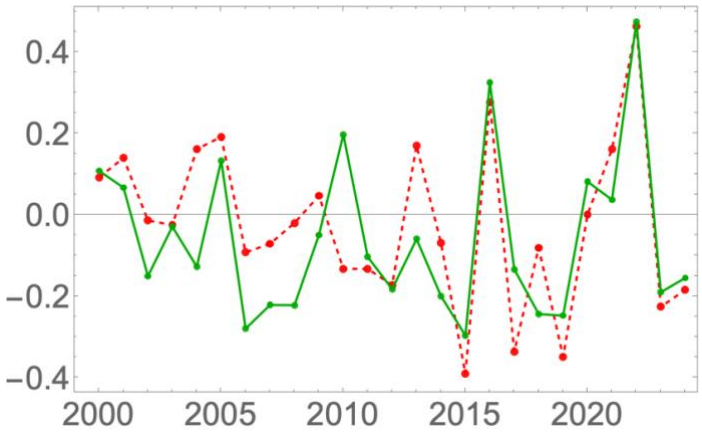


Figure S4
Normalised time series of the Indian Ocean EOF1 during July (red) and rainfall in SE Australia during September/October (green).

S4 Ocean map of correlations for Victoria and the Murray-Darling region

Search for correlation involves examination of many ocean maps that link SSTs with future rainfall. In the two examples shown below, information from the regions of strong correlation can be represented through EOFs and used in forecasting.

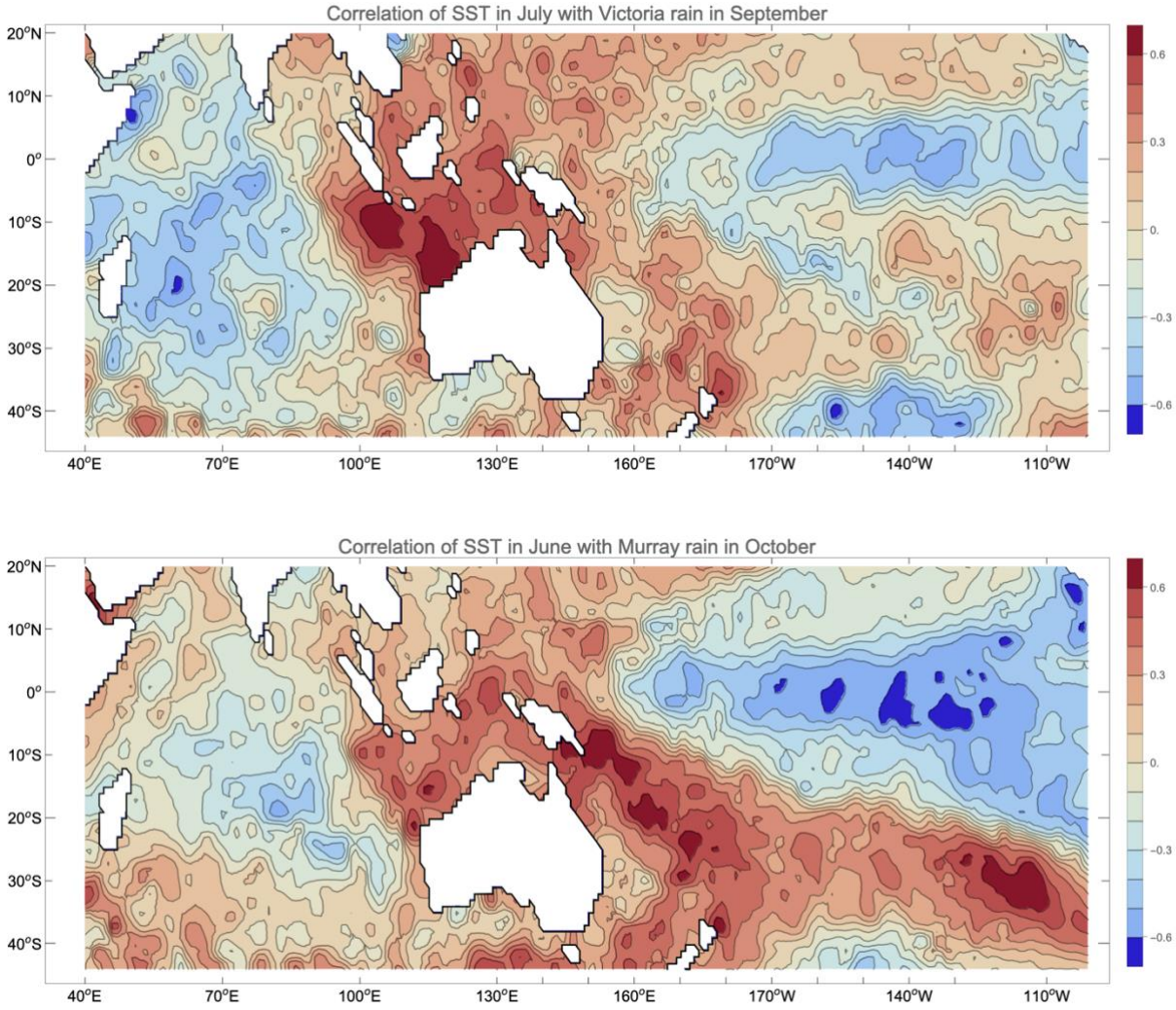


Figure S5
Top: Correlations between ocean SSTs during July with September rain in Victoria shows strong link with the Indian Ocean dipole.
Bottom: SSTs in June and Murray-Darling region rainfall in October are most strongly correlated in the Pacific Ocean east of the Australian continent.

S5 Mutual information

An alternative method to explore the information shared between two variables is a standard non-linear measure of *mutual information*. In the present case, mutual information between an ocean variable at the time t_1, $O(t_1)$, and rainfall at a different time $R(t_2)$ is defined as

$MI(O;R) = \Sigma_{t1,t2}\ P_{OR}(t_1, t_2)\ \log[P_{OR}(t_1, t_2)/(P_O(t_1)\ P_R(t_2))]$.s

The discrete individual distributions $P_O(t_1)$ and $P_R(t_2)$ and the joint distribution $P_{OR}(t_1, t_2)$ were approximated by fitting the data to continuous Gaussian distributions and the double sum was then evaluated as a double integral over the fitted distributions. Typical results are presented in Fig 4, which shows the mutual information between the meridional gradient and Southeast Australia rainfall at all times of a year.

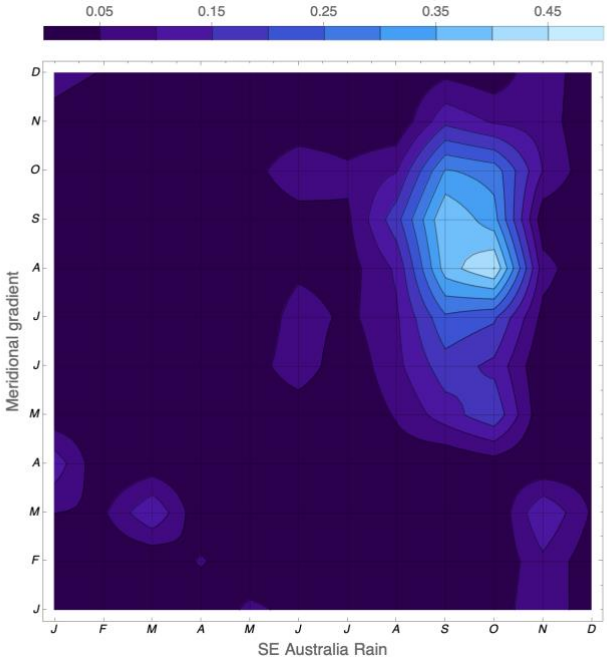


Figure S6.

Mutual information as a non-linear example of a link between the ocean and rainfall variables. Only in spring is the joint information between the SE Australia rainfall and the earlier meridional gradient significant, with the highest value of 0.45 obtained between meridional gradient in August and SE Australia rainfall in October.

S6 Accuracy of hindcasting

During the search, accuracy was monitored through RMSE over the validation period (Table S2). RMSEs are evaluated as the RMSE of the time series of average validation errors.

REGION	DATA MONTH	FORECAST SEPT/OCT RAINFALL (mm)	RMSE (mm)	AVERAGE RMSE (mm)	FORCAST (mm)	BOM MEASURED RAINFALL (mm)
Victoria	June	-11.0	24.7	22.5	-17.9 ± 22.5	-42.4
	July	-24.8	20.2			
Murray-Darling	June	-10.3	20.9	22.1	-25.5 ± 22.1	-32.1
	July	-40.7	23.3			
SE Australia	June	-31.6	16.9	16.8	-24.9 ± 16.8	-37.2
	July	-18.1	16.6			
SE Australia (forecast for November)	August			6.6	-7.3	3.7

Table S3
RMSE of the hindcasts for the combined September/October period, except for the last row which relates to the month of November. The last column, added during the revision, is using data for September/October rainfall that were not available at the time of original manuscript submission.

Alternative display of errors
An intuitive idea of the accuracy can be obtained from the bar graphs showing together the magnitudes of the deviation from the average and the error in hindcasting. The example in Figure S6 corresponds to the first example in the main text showing hindcasting of September/October rainfall in Victoria using June data.

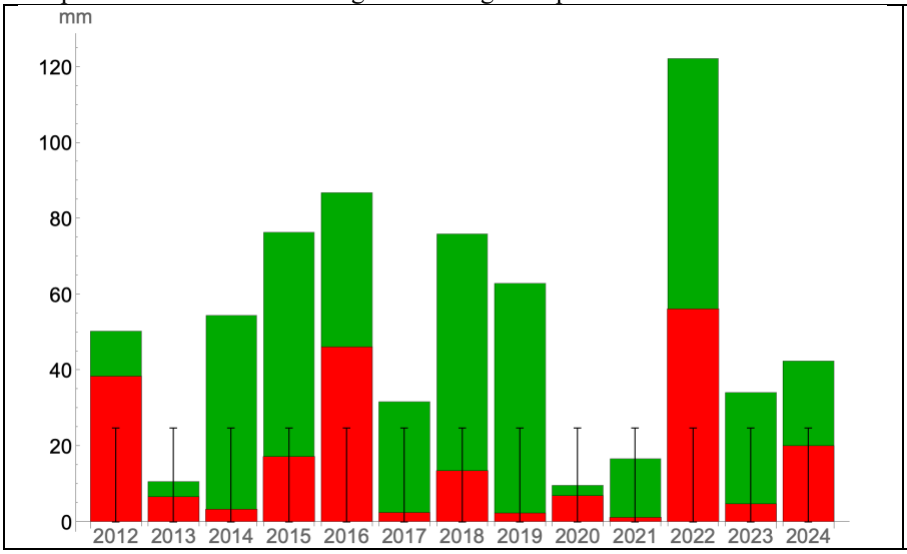


Figure S7
Accuracy of hindcasting of September/October rainfall in Victoria using June data.
Absolute values of rainfall deviation from the average are shown in green. Superimposed on the green bars are absolute values of the errors in hindcasts, shown in red together with RMSE bars. The last year was added during the revision.

An easy measure of accuracy is the prediction whether the month will receive more rain or less rain than the average. While our hindcasting is occasionally mistaken, in this example the sign of the deviation from the average was hindcast correctly for all 14 years shown in Fig S7. The largest errors occurred in the years when an unusually wet spring was not fully anticipated (2012, 2016 and 2022). Over the 13-year period those years were the only ones when the estimated RMSE was exceeded.

S7 Linear networks

Artificial neural networks used in this work can be run without a nonlinear element. In this case they act as an optimised linear predictor. We tested the performance of linear networks on the hindcasts with earlier June data, which is more interesting in view of agricultural applications.
For all regions, nonlinear networks performed better. The advantage of nonlinear analysis was very large for the Murray-Darling region and modest for Victoria.
The last column of the table shows performance of the standard linear regression using the same predictor input. As there is no validation period, all data until the last year were included in the evaluation of the prediction. While the errors are larger, the result is still useful as predictions and rainfall are strongly correlated. Neural networks advantage is significant in years with large deviations from average rainfall.

Region	RMSE linear (mm)	RMSE nonlinear (mm)	RMSE linear regression (mm)
Victoria	23.0	20.2	50.0
Murray-Darling	34.2	20.9	49.6
SE Australia	21.2	16.9	48.4

Table S4
Comparison of the performance of linear networks, nonlinear networks and linear regression in hindcasting with June data.

S8 2024 forecasts outcomes (added during the revision)

In the time since the initial submission, rainfall data for September and October were published by the *Bureau of Meteorology,* allowing us to compare the predictions from the manuscript with the actual outcomes. Briefly, the errors in forecasting are comparable to the errors in hindcasting shown in the ms. Numerical comparison is given in the Table S3 in the section S6, and Fig. 4 to Fig. 6 from the ms. for all three regions are repeated in Figs. S8 to S10 below with rainfall data (green) added to the last point and a few comments. Comparison of SE Australia forecast for November based on August data is shown in Fig. S11.

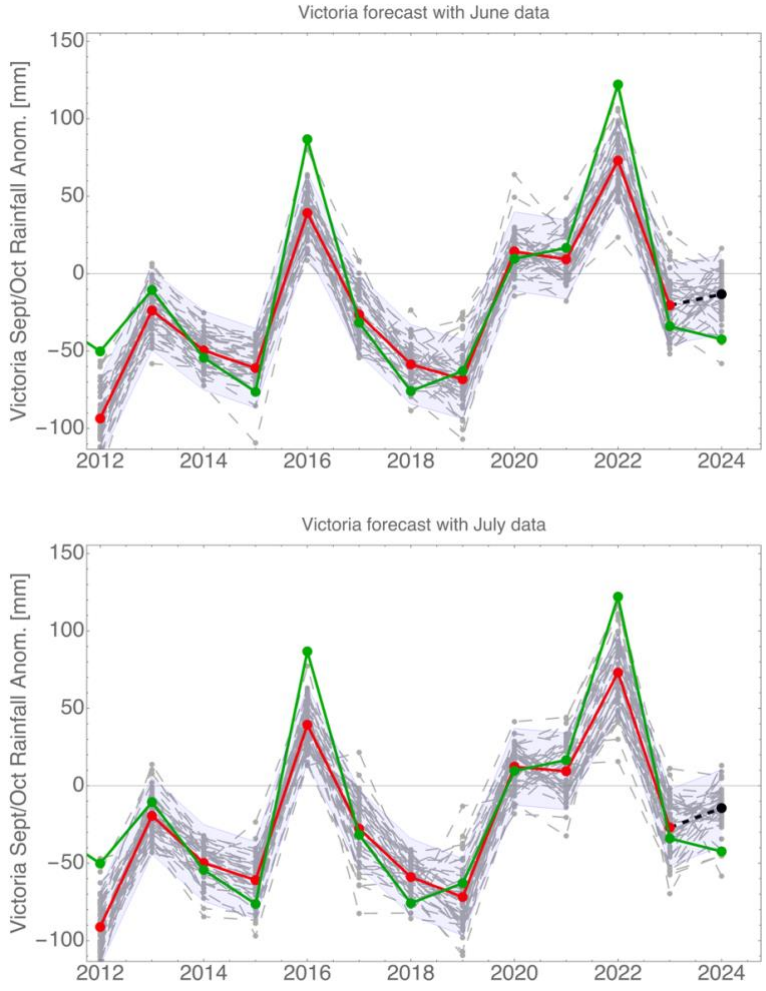


Figure S8
Forecast for Victoria was correct in foreseeing dry conditions, but more work is needed to improve accuracy.

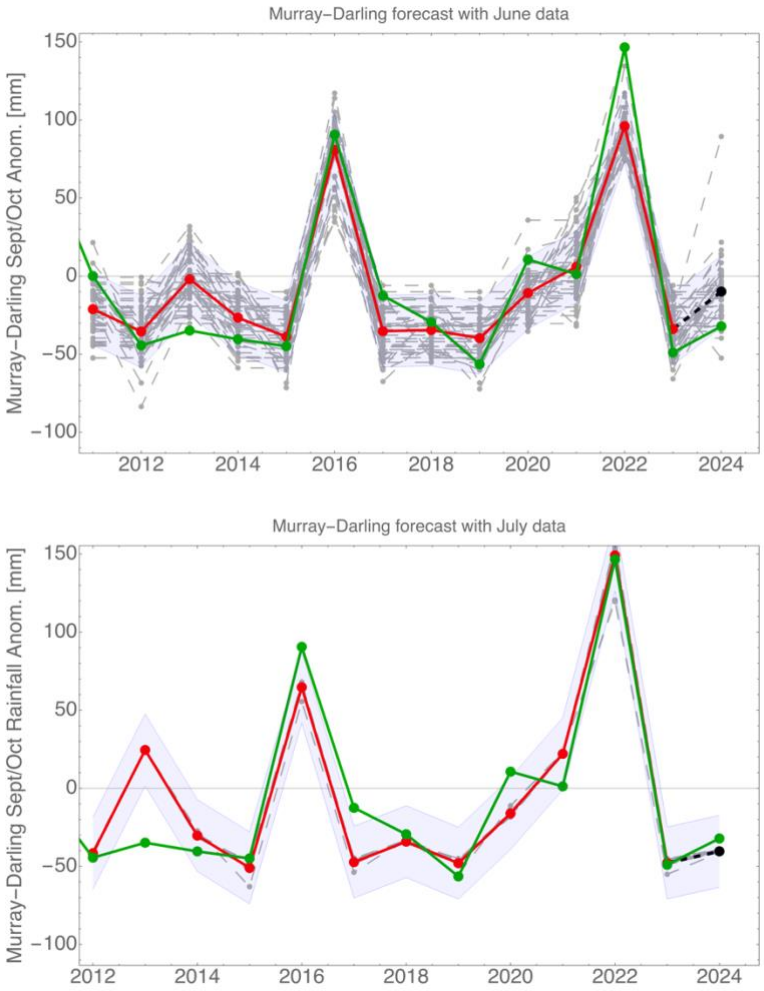
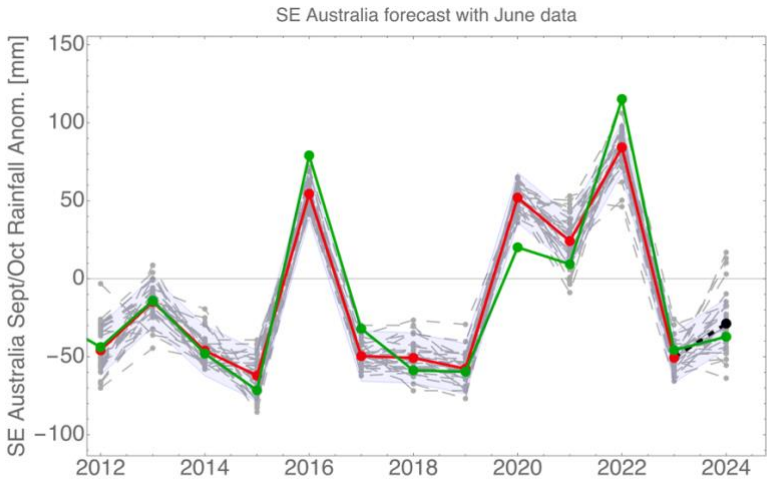


Figure S9
Murray-Darling forecast using July data and only two predictors (Indian Ocean EOF1 and Niño 3.4 SST was highly successful. In the same July data figure, two most recent hindcasts were also successful, perhaps due to the drift of best predictors towards stronger correlation with rainfall (Fig. S3).



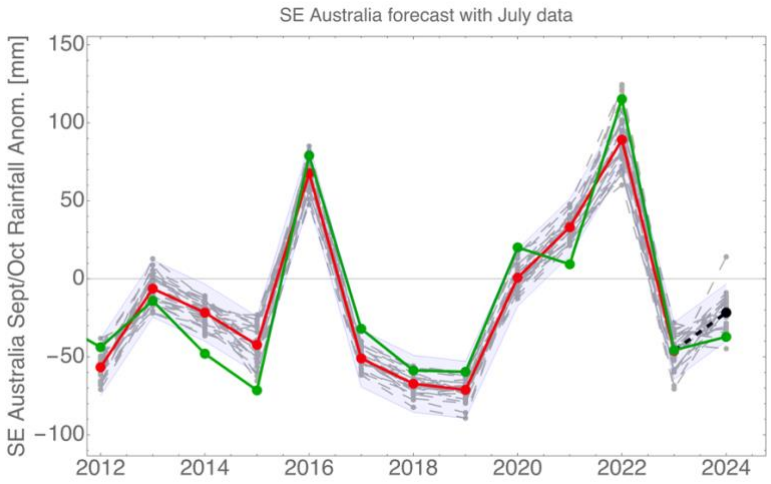


Figure S10
For SE Australia, June and July hindcasts accuracy is similar, and this year the forecast was better using June data.

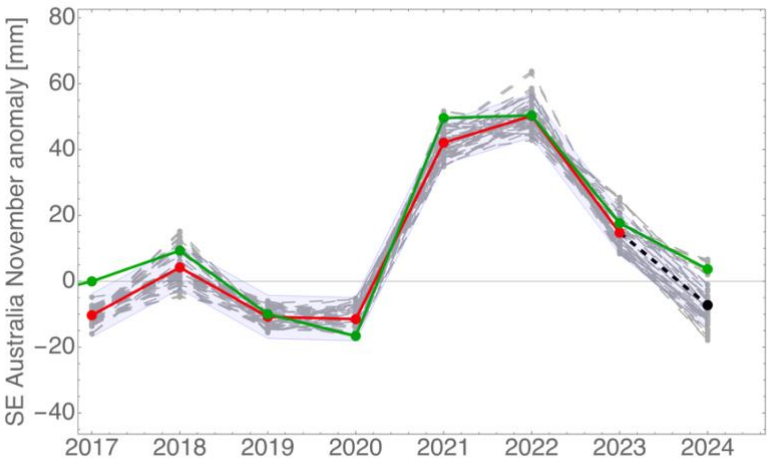


Figure S11
Good agreement obtained in this forecast was largely due to heavy rains on the last day of November. Had the rain arrived a day later our forecast would have been too high.