water

MDPI

Article

Machine Learning Assessment of the Impact of Global Warming on the Climate Drivers of Water Supply to Australia's Northern Murray-Darling Basin

Milton Speer *, Joshua Hartigan and Lance Leslie

School of Mathematical and Physical Sciences, University of Technology Sydney, Ultimo, NSW 2007, Australia
* Correspondence: milton.speer@uts.edu.au

Abstract: Droughts and long dry spells, interspersed with intense rainfall events, have been characteristic of the northern Murray-Darling Basin (NMDB), a major Australian agricultural region. The NMDB precipitation results from weather systems ranging from thunderstorms to larger scale events. The larger scale events exhibit high seasonal and annual rainfall variability. To detect attributes shaping the NMDB precipitation patterns, and hence net water inflows to the vast Darling River catchment area, numerous (45) possible attributes were assessed for their influence on rainfall trends. Four periods were assessed: annual, April–May (early cool-season), June–September (remaining cool-season), and October–March (warm-season). Linear and non-linear regression machine learning (ML) methods were used to identify the dominant attributes. We show the impact of climate drivers on the increasingly dry April–May months on annual precipitation and warmer temperatures since the early 1990s. The NMDB water supply was further reduced during 1992–2018 by the lack of compensating rainfall trends for the April–May decline. The identified attributes include ENSO, the Southern Annular Mode, the Indian Ocean Dipole, and both local and global sea surface temperatures. A key finding is the prominence of global warming as an attribute, both individually and in combination with other climate drivers.

Keywords: northern Murray-Darling Basin; southeast Australia; river water management; climate change and drivers; precipitation and temperature trends; machine learning techniques

Citation: Speer, M.; Hartigan, J.; Leslie, L. Machine Learning Assessment of the Impact of Global Warming on the Climate Drivers of Water Supply to Australia's Northern Murray-Darling Basin. *Water* **2022**, *14*, 3073. https://doi.org/10.3390/w14193073

Academic Editors: Ana-Maria Ciobotar and Fi-John Chang

Received: 21 August 2022
Accepted: 24 September 2022
Published: 29 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

The northern Murray-Darling Basin (Figure 1), hereafter NMDB, has a variable and intermittent rainfall pattern, with long dry periods interspersed with intense rainfall events. The NMDB occupies a large geographical area situated in the northern subtropical latitudes (between 25° S and 34° S) of Australia. It is influenced by tropical climate drivers such as the El Niño-Southern Oscillation phenomenon (ENSO) [1] more than the southern MDB (hereafter, SMDB) (Figure 1; located between 34° and 38° S), which is more influenced by mid-latitude weather systems in the cool-season months of April–September [2,3], so the two areas have different climate regimes.

Historically, both NMDB and SMDB floods have occurred as a result of winter/spring rains. However, floods also can occur in summer, particularly in La Niña years, for example, in 2010–2012 [4] and, more recently, in 2020–2022. The NMDB summer flood rainfall typically results from tropical low-pressure systems, including landfalling tropical cyclones [5].

Figure 1. Map of northern and southern Murray-Darling Basin in southeast Australia. The MDB lies within subtropical latitudes (25° S–38° S) of the Australian continent. Observation stations used for precipitation, TMax and TMin averaging that represent the NMDB are marked and indicated in a legend. (Source: Murray-Darling Basin Authority, G.P.O. Box 1801, Canberra City, ACT 2601 Australia. https://www.mdba.gov.au/sites/default/files/pubs/Murray-Darling_Basin_Boundary.pdf (accessed 19 August 2022). Reproduced with some place name deletions and insertions via license: Creative Commons Attribution-Non Commercial-NoDerivatives4.0 International Public License (CC BY-NC-ND 4.0).

Globally, each decade since 1980 has been warmer than the previous decade, with 2010–2019 being around 0.2 °C warmer than 2000–2009 [6], indicating an acceleration in global warming in recent decades. In addition, the mean annual near-surface air temperature since 1910 in Australia has increased by nearly 1.5 °C, relative to the 1910–1950 mean; this is about 30% greater than the global average [6].

Since the early 1990s, climate change research has revealed a reduction of approximately 12% in April–October precipitation over southeast Australia [6]. Notably, no trend is apparent annually because there has been an increase in summer precipitation [7] and a decrease in autumn (MAM) precipitation [8]. For the SMDB, ref. [9] found a reduction in late autumn (April–May) precipitation since the 1990s.

Rainfall in the NMDB is generated by weather systems on a range of time and space scales, resulting from climate drivers including the ENSO the southern annular mode (SAM), the Indian Ocean Dipole (IOD), and the Interdecadal Pacific Oscillation (IPO), or its Tripole Index [10]. These modes drive the high variability and unreliability in seasonal and annual rainfall amounts and hence in stream inflows that eventually reach the Darling River. However, contributions to the significant precipitation reduction in the SMDB in late autumn (April–May) were found by [9].

Farms and other local communities rely on river water for human consumption and for irrigation of agricultural crops, underlining the social and economic importance of river water for the NMDB. There is a wide variety of crops requiring irrigation, the most important being cotton, with the NMDB producing over 90% of Australia's total cotton crop [11]. In addition, key NMDB environmental regions require an adequate water supply to sustain important NMDB ecosystems and wetlands, and to maintain the many river water holes in which fish and other species survive long interruptions of water flow [12]. Mass fish deaths in the Lower Darling River in 2019 were caused by decreased river flows, due mainly to poor river regulation combined with recent drought and higher temperatures [13]. Small flushing flows improve water quality along the river by reducing salinity levels, breaking up stratification in pools, mitigating algal blooms, and are essential for ensuring the diversity and abundance of aquatic life such as prawns, mussels and aquatic insects which form a vital food resource for fish and waterbirds [12]. While acknowledging that there are known management issues affecting river levels, 94% of the gauges in the NMDB show a declining trend in streamflow since records began in 1970. Moreover, the headwaters of the NMDB are reporting declining streamflow trends [7]. Rainfall extremes have continued after the Millennium Drought (1997–2009). Following a wetter-than-average 2016, most of the SMDB received decile 2–3 or decile 1 rainfall over the 4-year period from January 2017 to December 2020. During the same period, almost all the NMDB fared even worse, receiving either decile 1 rainfall or its driest 4 years on record (Figure 2).

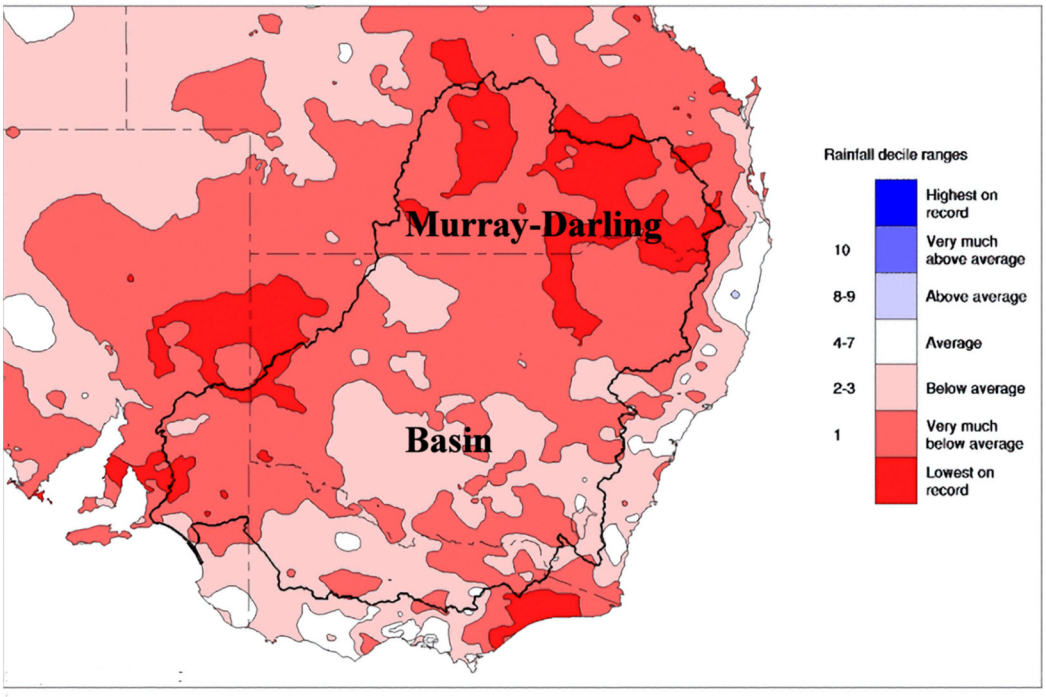


Figure 2. Murray-Darling Basin rainfall deciles. Rainfall deciles for the 48 months January 2017 to December 2020 in southeast Australia focusing on the MDB defined by the area within the solid black line. Note the lowest on record in the north of the basin and the very much below or below average rainfall in the rest of the basin. (Reproduced with permission under Creative Commons Attribution Licence 3.0 from the Australian Bureau of Meteorology). Available at: http://www.bom.gov.au/climate/maps/rainfall/?variable=rainfall&map=decile&period=48month&region=md&year=2020&month=12&day=31 (accessed 19 August 2022).

While flood events do not occur regularly in the NMDB, there were two major floods events during 2010–2012 and, more recently, 2020–2022. However, due to the general drying trend in the NMDB [14], it is important to analyze potential trends in variables that impact water availability. These include decreased NMDB catchment precipitation, and increased evaporation rates due to higher temperatures resulting from global warming. For example, rainfall was very low across the NMDB during January–February 2019, typically the wetter months of the year, resulting in 2018–2019 being the driest since 2002–2003. Consequently, area-averaged actual evapotranspiration deciles have become the lowest on record, as shown in Figure 3. Notably, a study covering 1975–2016 found that prior to 1994 wind speed dominated evaporation rates across southern Australia, but after 1994 temperature became dominant [15]. It is notable that evapotranspiration in the NMDB during 2018–2019 was even greater than in the southern MDB (Figure 3). Moreover, the water storage capacity in the NMDB dams is about 1/3 that of the SMDB dams. Hence, when the NMDB dams fill in the warm season (October–March), as they did in 2020/2021 and 2021/2022, irrigation usage and high evaporation rates can rapidly reduce dam levels. It also indicates that an excessively high proportion of the NMDB inflow reaching the Darling River, the main river in the NMDB, is the direct result of rainfall inflows rather than releases from dams. Because most tributaries in the NMDB are unregulated, historically, this is why most farms have found it necessary to construct on-farm dams for water storage [16].

The future development of more accurate seasonal and annual prediction models on a range of time scales, would benefit from climate drivers identified in this study, which we investigate using ML techniques. Hence, valuable guidance could be provided to state and local water authorities, thereby assisting future water allocation decisions.

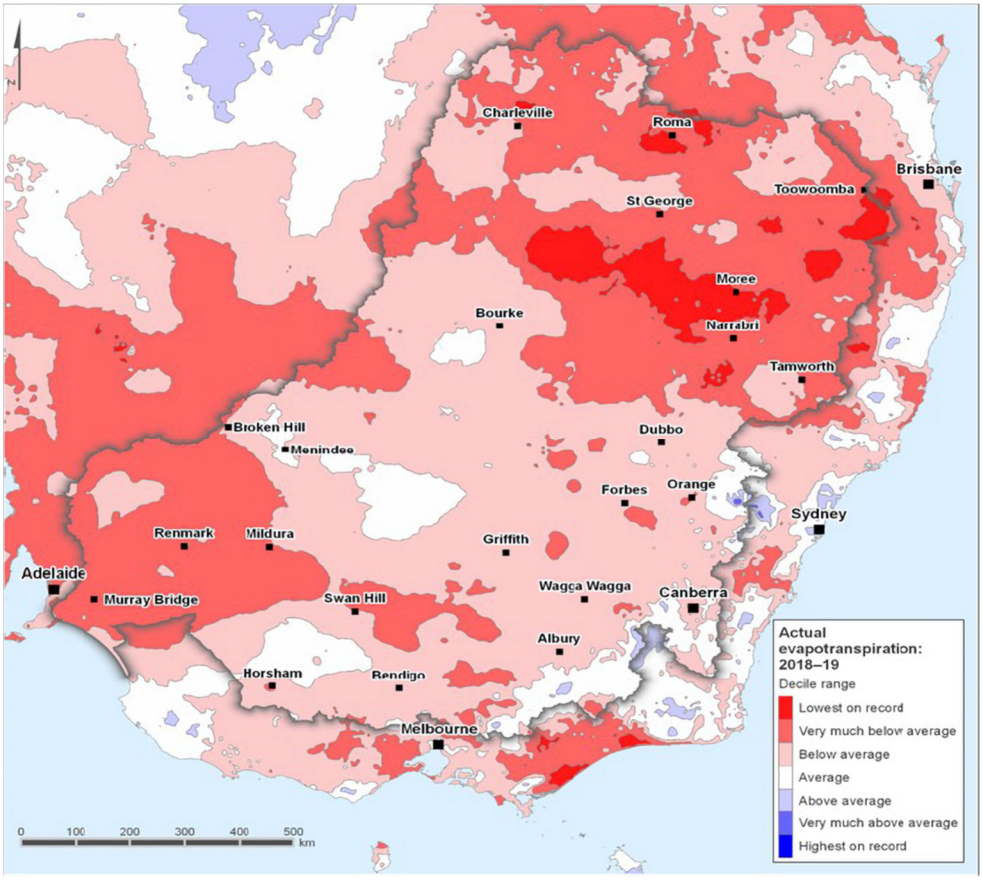


Figure 3. Annual deciles of actual evapotranspiration and soil moisture 2018–2019. Map of southeast Australia showing deciles for the MDB region during 2018–2019, for annual area-averaged actual evapotranspiration calculated using The Australian Landscape Water Balance Model (AWRA-L v6) [17]. Note the decile area of lowest on record in the NMDB. (Reproduced with permission under Creative Commons Attribution Licence 3.0 from the Australian Bureau of Meteorology. Available at: http://www.bom.gov.au/water/nwa/2019/mdb/climateandwater/climateandwater.shtml (accessed 19 August 2022).

The major aims of this study are as follows. First, is to identify trends in precipitation and temperature from sections of the vast NMDB catchment area that determine water availability in the Darling River, and from the numerous river systems and tributaries that feed into it. Next is a detailed attribution analysis of climate drivers affecting annual and seasonal rainfall using machine learning techniques. Such an approach to rainfall stations in this region of southeast Australia has not previously been attempted. Here, it is performed by ranking the numerous possible attributes in importance, and by using a range of linear and non-linear ML regression techniques, described in Sections 2.4 and 3.4, fitted to the observed precipitation amounts, for the period 1965–2018. The attributes are assessed both individually and in combinations, with the five most effective attributes chosen on their percentage appearance in 10-fold cross-validation (to avoid overfitting) of annual precipitation, precipitation in late autumn (April–May), in the remaining cool season (JJAS), and in the warm season (October–March), to highlight warm season and cool season influences. The regression techniques chosen, have proven to be successful in previous studies by the authors, in identifying attributes in previous climate change studies (e.g., [18,19]). In addition, a wavelet analysis was applied to investigate the possible contributions from a range of climate drivers.

2. Data and Methodology

2.1. Data

Monthly precipitation data were obtained from the Bureau of Meteorology's homogeneous climate record through its climate change site network [20]. These data have undergone complex quality control to address inconsistencies and errors. There are 8 precipitation stations comprising Augathella, Cunnamulla, Normandy, Miles, Surat, Bellata, Bingara and Curlewis. Additionally, monthly mean maximum temperature (TMax) and mean minimum temperature (TMin) data were obtained for Charleville, Miles, Thargomindah, St. George, Moree, Inverell, Walgett and Gunnedah from the Bureau of Meteorology's climate change site network. These sites were selected as they cover a broad area of the NMDB (Figure 1) and to contain a long, continuous record beginning in 1910, in a similar manner to an earlier MDB study [21], in order to approximate the precipitation and temperature characteristics across the region. The precipitation data were summed to obtain a proxy to the total precipitation falling across the entire NMDB catchment area, while the mean of the temperature time series data was calculated to obtain the mean of TMax and TMin across the NMDB. The available time series data cover the period from 1910–2019.

2.2. Statistical Analysis

To analyze trends in the data, the time series were plotted with their percentiles displayed in four panels: annually, late autumn (April–May), the four remaining cool season months (JJAS), and the warm season (October–March). The data were then grouped into four 27-year periods (1911–1937, 1938–1964, 1965–1991 and 1992–2018), to obtain a representative climatology that covers almost all of the time series. Additionally, the decision on the groupings was influenced by the fact that the last three 27-year periods mentioned above coincide with the main time span from 1950 when global warming (GW) became prominent [22], then from the 1970s when GW began to accelerate [23,24], and particularly from the early to mid-1990s [6]. Change point analysis of the data is not applicable because the acceleration in GW occurred over several years in the early to mid-1990s. If years other than 1991 had instead been chosen to demarcate the last two periods, the findings would have remained almost unchanged. Bootstrap resampling with replacement was applied [25], with 5000 resamples, to the mean and variance of the time series data to develop box plots that can be used to identify potential trends, including the direction of such trends, between the time periods. Two-sided permutation testing then was applied to test for statistical significance of any potential trends. A value of $\alpha < 0.1$ is a marginally significant result, $\alpha < 0.05$ is significant, and $\alpha < 0.01$ is highly significant.

2.3. Wavelet Analysis

Wavelet analysis [26,27] was applied to each time series to detect potential climate drivers such as the El-Niño Southern Oscillation. This approach provides both the local wavelet power spectrum and the global power spectrum, as shown in Section 3.3. The local wavelet power spectrum shows how the influence of climate drivers changes over time, while the global power spectrum provides an overview of which drivers are dominant in the time series. In this study, we used the Morlet wavelet as the mother wavelet.

2.4. Attribute Selection

Numerous climate drivers were considered as potential attributes for annual, April–May, JJAS, and October–March precipitation. The drivers assessed were the Atlantic Multi-decadal Oscillation (AMO), the Indian Ocean Dipole (IOD), global sea surface temperature anomalies (GlobalSSTA), global temperature anomalies (GlobalT), Niño3.4, TPI, the Southern Annular Mode (SAM), the Southern Oscillation Index (SOI), and the Tasman Sea surface temperature anomalies (TSSST). Time series of these climate drivers were obtained from the Earth System Research Laboratory

(http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/ -accessed 19 August 2022), except for TSSST which was obtained from the Australian Bureau of Meteorology (http://www.bom.gov.au/climate/change/?ref=ftr#tabs=Tracker&tracker=timeseries – accessed 19 August 2022).

Two-way interaction terms between the above predictors, where one variable is multiplied by another (e.g., AMO*IOD), were also considered potential predictors. These two-way interaction terms are considered as possible attributes as one attribute might reinforce another. Additional relationships between possible attributes, such as additive and quotient relationships, although assessed by Richman and Leslie [28] for a single station, were not considered in this study, because some of these are highly correlated with other attributes considered here, and all ocean basins already are represented. Furthermore, we restricted the combinations to those involving two potential attributes, because the number of combinations of the 9 possible predictors (AMO, SOI, Niño3.4, SAM, IOD, TPI, Global temperature anomalies, and Global and Local SST anomalies) would have been prohibitive to assess, as explained in recently published work (see Table 1 in ref. [29]).

As relationships between climate drivers and precipitation can vary between weak and strong, linear and non-linear, and in combinations, this study develops both linear and non-linear statistical models of annual, April–May, JJAS and October–March precipitation, for a range of climate drivers. The non-linear models considered are support vector regression (SVR) [30], with either the polynomial (Poly) or radial basis function (RBF) kernels, and random forests (RF) [31,32]. When the two-way interaction combinations are included, there are 45 potential attributes considered for the prediction of precipitation. If all attributes were used, the models would be subject to overfitting, which reduces physical understanding and is likely to lead to large errors if the models are applied to a test data set in prediction mode. To select attributes that generalize well from the set of 45 attributes, ten-fold cross-validation was applied to the data set, using both forward and backward selection through the space of potential attributes as in Maldonado and Weber [33]. The comprehensive, detailed set of results from all attributes in all models are provided in Table 2 in Section 3.4.

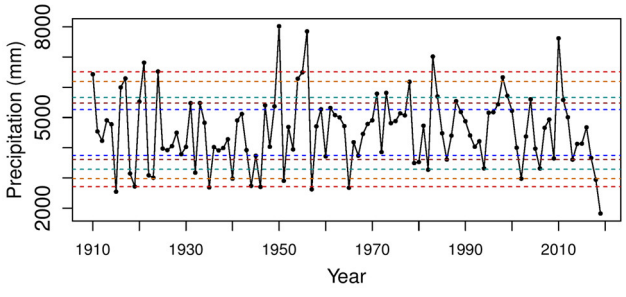
3. Results and Discussion

First, precipitation and temperature time series are discussed, followed by the *p*-values and box–whisker plots for precipitation, TMax and TMin. Then, wavelet analyses of precipitation, TMax and TMin are discussed to highlight the periodicity in the climate drivers. Finally, attribute selection of climate drivers is described for each time period.

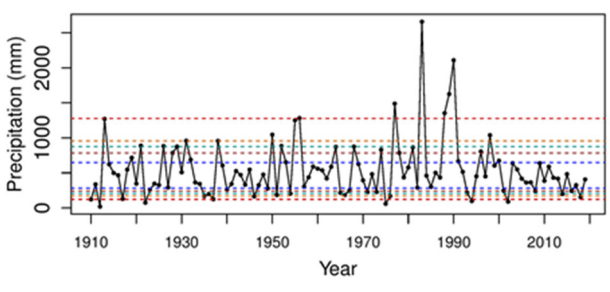
3.1. Precipitation and Temperature Time Series in the Northern Murray-Darling Basin

3.1.1. Precipitation

April–May is the only period of the year exhibiting long-term changes, with an apparent decrease in both the mean and variance from the 1990s (Figure 4a). The annual time series shows no long-term trend (Figure 4b), which also is the case for JJAS and October–March. However, the steep decrease in the annual times series after 2011, to well below the 5th percentile in 2019, has contributions from all of the periods.



(**a**)



(**b**)

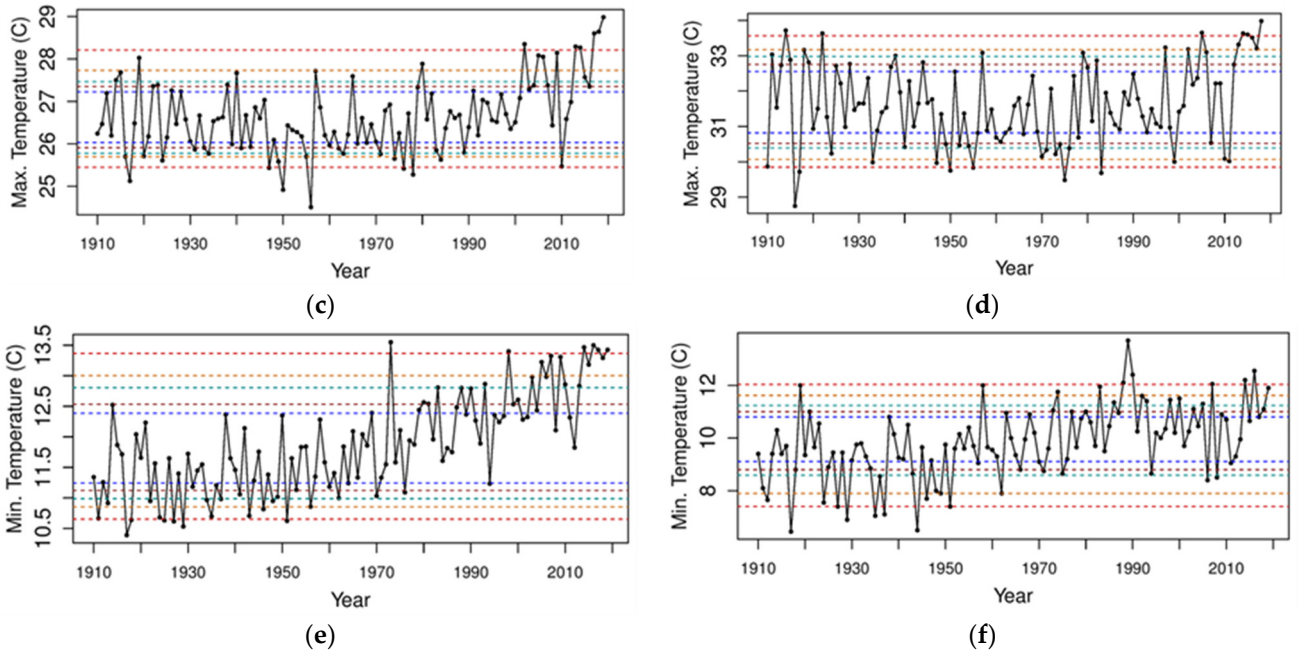


Figure 4. Precipitation, TMax and TMin time series in the NMDB. Precipitation time series in the NMDB for, (**a**) annual, (**b**) April–May. Dashed lines indicate percentiles 5th and 95th (red); 10th and 90th (orange); 15th and 85th (light blue); 20th and 80th (brown); and 25th and 75th (dark blue). Note the apparent decrease in mean and reduction in values greater than the 75th percentile for April–May since the 1990s and the decrease from 2011 to well below the 5th percentile in 2019 for the annual time series; the time series of TMax in the NMDB for, (**c**) annual, (**d**) October–March. Note the steep increase approaching the 100th percentile in both October–March and annual time series which mirrors the steep decrease in 2019 annual precipitation (Figure 4b); the time series of TMin in the NMDB for, (**e**) annual, (**f**) April–May. Note the almost linear increase in the annual times series since the 1960s and the step increase with much less variation since 2014, while there is no clear trend in April–May apart from one step increase after the 1950s and another from 2014.

3.1.2. TMax

The annual times series of TMax shows an increasing trend from the early 1990s (Figure 4c). That is due to the strong contributions from JJAS in particular, but also from April–May. However, there is a steep increase approaching the 100th percentile in the period October–March (Figure 4d) which mirrors the steep decrease in 2019 precipitation.

3.1.3. TMin

The annual TMin time series shows an almost linear increase since the 1960s, with a step increase and much less variation since 2014 (Figure 4e). The increase is due mostly to similar increases in JJAS and October–March. The step increase and lower variability are present also in JJAS and October–March. However, after an initial step increase in April–May from the early 1950s (Figure 4f), there appears to be no clear trend after that, aside from a step increase and lower variability from 2014.

3.2. p-Values and Box–Whisker Plots for Precipitation, TMax and TMin

The *p*-values in Table 1 allow comparisons of the four 27-year intervals. Statistically significant differences (i.e., *p*-value < 0.1) are shown, in bold italics, for means and variances of precipitation, TMax and TMin. There are comparisons involving the four intervals 1911–1937, 1938–1964, 1965–1991 and 1992–2018.

Table 1. *p*-values from permutation testing differences in interval means and variances. *p*-values from permutation testing differences in interval means and variances for, April-May, JJAS, October-March and annual precipitation, TMax and Tmin, based on area averages of observing stations in the northeast part of the northern Murray-Darling Basin. Significant values ($p < 0.10$) are italicised in bold. Note that the *p*-value for each variance test is calculated after one sample has had bias correction in the mean. Key points to note are the highly significant *p*-values ($p < 0.05$) for April-May mean and variance precipitation decreases from 1965-1991 to 1992-2018; and highly significant increases generally in mean TMin and also generally in mean TMax for most of the periods.

Period	Observation	Mean & Variance	Interval *p*-Values			
			1911–1937 vs. 1965–1991	1938–1964 vs. 1965–1991	1911–1937 vs. 1992–2018	1965–1991 vs. 1992–2018
April-May	Precip.	Mean	**0.1**	0.289	0.48	**0.0318**
		Variance	**0.0168**	**0.0154**	**0.0946**	**<0.01**
	TMax	Mean	0.194	**0.0542**	**<0.01**	**<0.01**
		Variance	0.934	0.599	0.677	0.73
	TMin	Mean	**<0.01**	**<0.01**	**<0.01**	0.847
		Variance	0.726	0.897	0.318	0.584
JJAS	Precip.	Mean	0.593	0.788	0.388	0.679
		Variance	0.23	0.815	0.581	0.585
	TMax	Mean	0.689	0.577	**<0.01**	**<0.01**
		Variance	0.602	0.474	0.569	0.935
	TMin	Mean	**0.0230**	**0.0142**	**<0.01**	**<0.01**
		Variance	0.913	0.989	0.86	0.795
Oct-Mar	Precip.	Mean	0.225	0.705	**0.0424**	0.404
		Variance	0.664	0.1	0.573	0.909
	TMax	Mean	0.157	0.698	0.373	**0.0194**
		Variance	0.251	0.718	0.864	**0.0626**
	TMin	Mean	**<0.01**	**<0.01**	**<0.01**	**<0.01**
		Variance	0.759	0.94	0.556	0.211
Annual	Precip.	Mean	0.295	0.933	0.498	0.688
		Variance	0.22	0.12	0.584	0.644
	TMax	Mean	0.628	0.304	**<0.01**	**<0.01**
		Variance	0.619	0.614	0.42	0.207
	TMin	Mean	**<0.01**	**<0.01**	**<0.01**	**<0.01**
		Variance	0.859	0.416	0.825	0.97

3.2.1. Precipitation

The most important precipitation finding is the significant decrease in the April–May mean and highly significant decrease in variance, between the intervals 1965–1991 and 1992–2018, with *p*-values of 0.0318 and <0.01, respectively (Table 1). It is noteworthy that the significant decrease in the mean is an abrupt reversal of the increasing trend in the mean in the periods prior to 1992–2018. The decrease in variance between 1992 and 2018 and previous intervals, including 1965–1991, is also highly significant, with a *p*-value < 0.01, as shown in Table 1 and is very apparent in the box–whisker plots (Figure 5a,b). Possible climate influences are discussed in the following wavelets section.

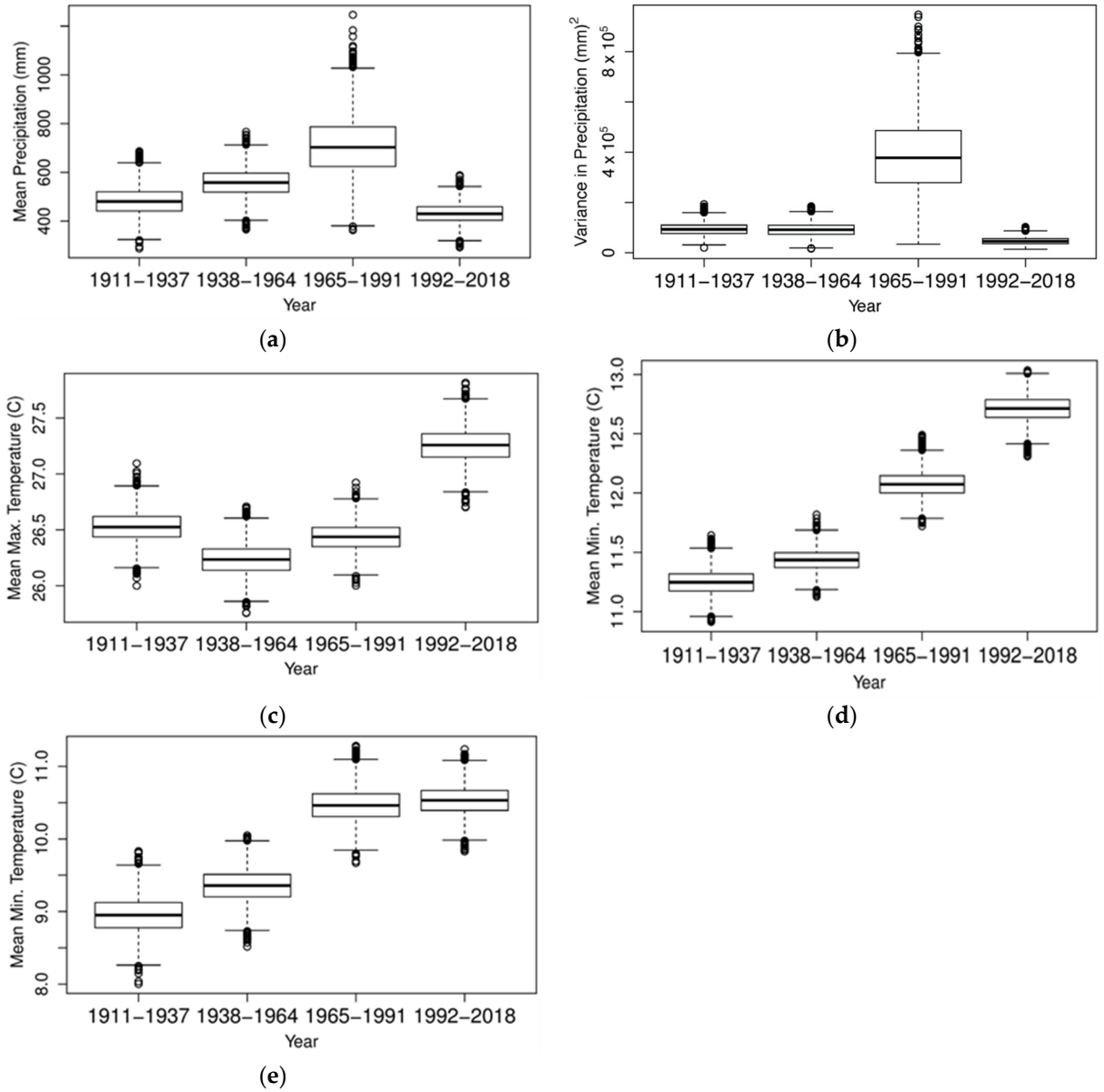


Figure 5. Box and whisker plots of NMDB precipitation, TMax and TMin. Box and whisker plots of NMDB for, (**a**) mean April–May precipitation, and (**b**) variance; (**c**) mean annual TMax; (**d**) mean annual TMin; (**e**) mean April–May TMin.

3.2.2. TMax

Increasing mean TMax values are highly significant throughout the year between the intervals 1965–1991 and 1992–2018, with a *p*-value of <0.01 (Table 1) and confine clearly the accelerated GW over the last 50 years [23,24]. The annual mean box–whisker plot illustrates this significance (Figure 5c).

3.2.3. TMin

There also is high significance in mean TMin increases between all periods with all intervals and 1992–2018, apart from the notable exception of April–May between 1965–1991 and 1992–2018, where the *p*-value is 0.847 (Table 1). This highlights the lack of change shown in the box–whisker plots (Figure 5d,e). The most likely explanation is that clearer nighttime and early morning skies due to the highly significant decrease in precipitation, and possibly lighter winds, from the 1990s, has increased radiational cooling relative to 1965–1991. The very high significance in the TMin increase from 1938–1964 to 1965–1991

for all periods (Table 1) corresponds approximately to 1950, when the importance of GW is documented [22].

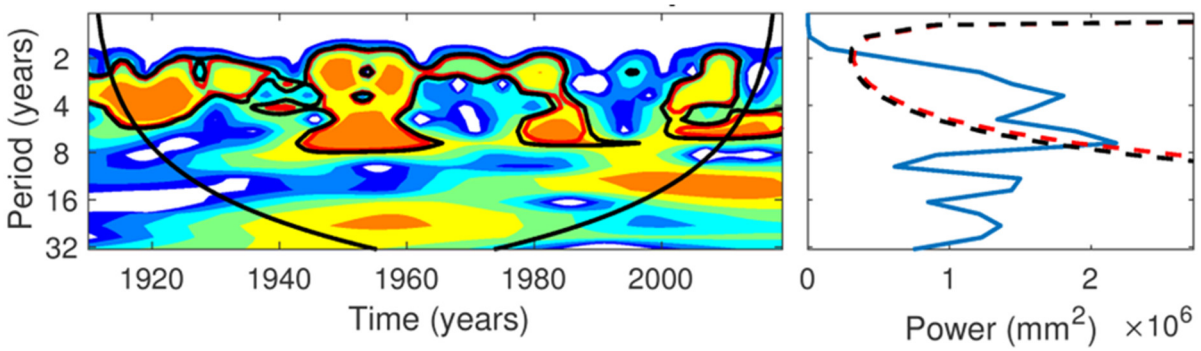
3.3. Wavelet Analysis of Temperature and Precipitation 1911–2018

It is noted that oscillations in atmospheric and oceanic phenomena such as ENSO are expected to appear in wavelets of their time series. However, many climate drivers, such as the SAM, IPO, IOD, and the Atlantic multi-decadal oscillation (AMO), are modes that operate on variable time scales. As such, they are not as well-represented in wavelet analyses as ENSO. Following [28], wavelet analysis was applied to detrended anomalies of the precipitation and temperature time series. Wavelet analysis produces both a local and a global spectrum. The local wavelet power spectrum reveals the temporal evolution of those periodic signals located in the time series, allowing the detection of oscillatory climate drivers such as ENSO, and shows their temporal evolution. The global power spectrum also is generated, displaying the major amplitudes present in the time series.

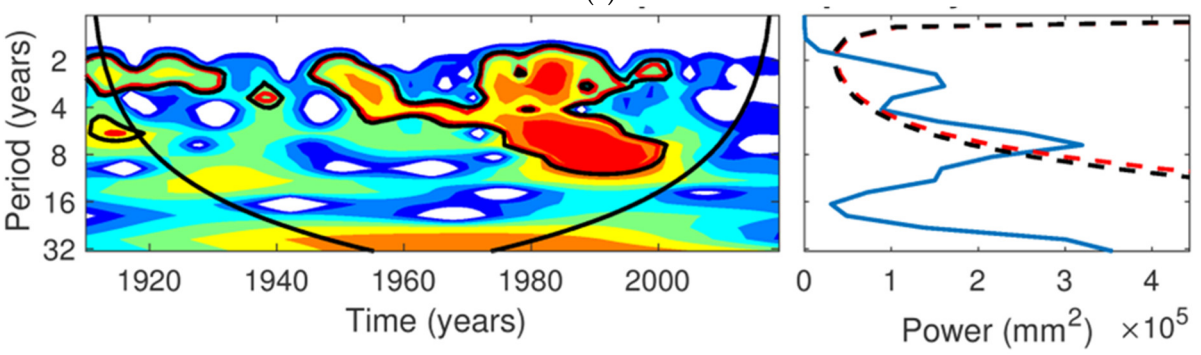
The following annual, April–May, JJAS and October–March wavelets for the NMDB are interpreted for precipitation. The wavelets for TMax and TMin were very similar to precipitation, emphasizing the 2–8-year ENSO periodicity. Hence, they are not shown here.

3.3.1. Precipitation

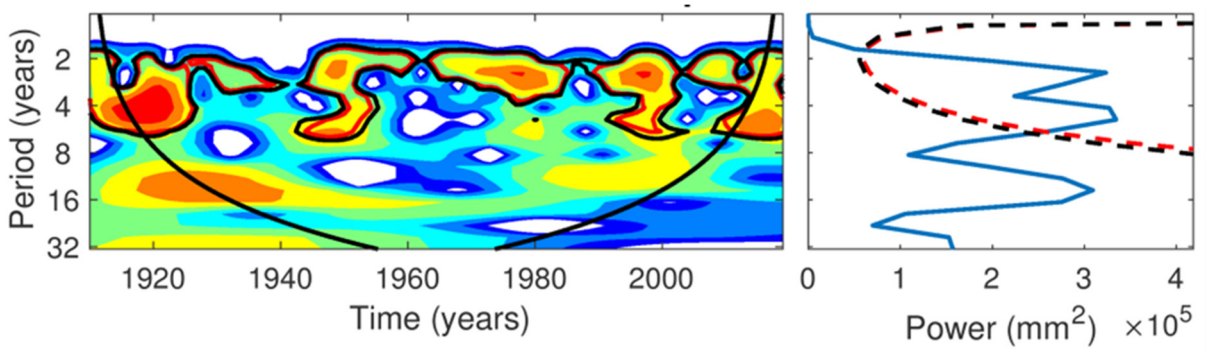
The annual precipitation wavelet plot shows ENSO significance at both the 90th and 95th percentiles in the 2–8-year period (Figure 6a), dominated by October–March (Figure 6b) and JJAS (Figure 6c), and a significant period from 1950 to 2000 in April–May (Figure 6d). There is a notable non-significant, but possible, IPO periodicity in the 15–30-year time frame. It is noteworthy that although it is not part of this study, the recent La Niña of late 2020 and early 2021 provided increased precipitation in the NMDB, highlighting the ongoing impact of ENSO, when two of its catchments had two of their three largest one-day increases in water storage levels since 1993 [34].



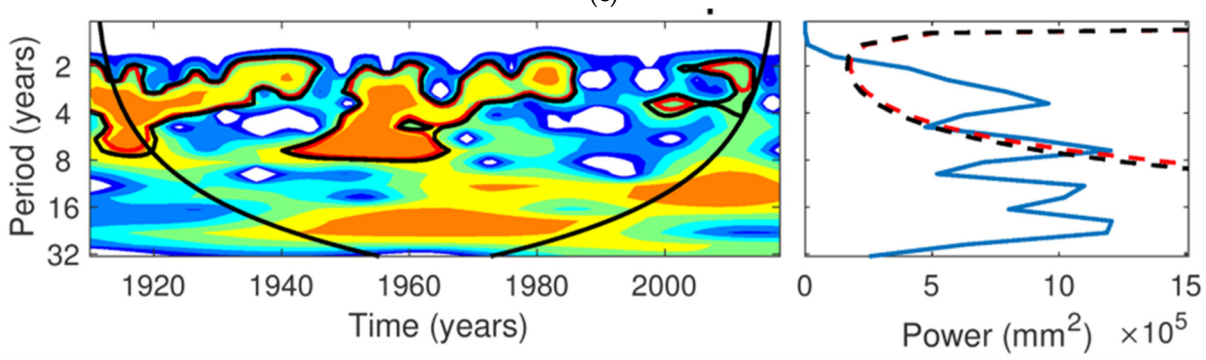
(**a**)



(**b**)



(**c**)



(**d**)

Figure 6. Wavelets for precipitation. Wavelets representing NMDB precipitation for the periods, (**a**) annual, (**b**) April–May, **(c)** JJAS, and (**d**) October–March. The local wavelet power spectra are displayed in the left sub-figure, while the global wavelet power spectra that represent globally dominant periodicities in the time series are displayed in the right sub-figure. The dashed black and red lines are the 90th and 95th confidence percentiles, respectively. The black cone shape on the local wavelet power spectra displays the cone of confidence.

3.3.2. TMax

The wavelets for TMax (not shown) also indicate that the ENSO periodicity is the dominant influence in all four periods. Interestingly, its significance in approximately 12-year periodicity starts from the 2000s, outside the 90th and 95th percentile cones of confidence in October–March and appears annually.

3.3.3. TMin

As for TMax, ENSO is the only periodicity detected in the TMin wavelets (not shown) and is apparent in all four time periods.

3.4. Attribute Selection

As described in Section 2.4 above, the possible attributes considered in this study are the Atlantic Meridional Oscillation (AMO), the Indian Ocean Dipole Mode Index (DMI), global sea surface temperature anomalies (GlobalSSTA), global temperature anomalies (GlobalT), the Niño3.4, the Interdecadal Pacific Oscillation (IPO), the Southern Annular Mode (SAM), the Southern Oscillation Index (SOI), and the Tasman Sea surface temperature anomalies (TSSST). Two-way interaction terms between these predictors also were considered (e.g., AMO*DMI). The total number of potential attributes was 45. The precipitation data covers the period 1965–2018 and 10-fold cross-validation was applied to limit overfitting with the above techniques. Those attributes which appeared in at least 50% of the folds across the eight training methods are considered likely attributes of NMDB precipitation. However, because many attributes appeared in at least 50% of folds, only the five attributes with the highest percentages above 50% were selected here as the key attributes of NMDB precipitation. For annual precipitation, these attributes were

GlobalSSTA*TSSST, AMO*SAM, AMO*Niño3.4, TPI and DMI*Niño3.4 (Table 2). Two-way interactions were selected more often than either attribute in isolation, highlighting the importance of considering two-way interactions as potential attributes. For example, GlobalSSTA*TSSST occurs in 60% of folds, whereas GlobalSSTA occurs only in 46.25% and TSSST in 55%. For April–May precipitation, the key attributes were SOI, GlobalSSTA*TSSST, TPI, GlobalT and DMI. For JJAS precipitation those selected were SOI, Niño3.4, SAM, TPI and DMI*TSSST. Finally, the key drivers of October–March precipitation were SOI, DMI*SAM, SAM, GlobalT*GlobalSSTA and SOI*TPI (Table 2).

Notably, for each precipitation period, the selected attributes are related to GW, GlobalT, GlobalSSTA, and TSSST, highlighting the profound influence that GW now has on precipitation in the NMDB. This is particularly relevant for April–May precipitation, where it was shown in Table 1 that there is a statistically significant decrease between periods 1965 and 1991 and 1992 and 2018, as two of the climate drivers are direct attributes to accelerated GW during 1992–2018. Additionally, Table 2 reveals that drivers related to the atmospheric based SOI, the ENSO based tropical Pacific Ocean indices, the TPI and DMI are known to influence precipitation across all precipitation groupings. Finally, AMO appears to play a role in annual precipitation in the NMDB by modulating the influence of SAM and Niño3.4.

Table 2. Major precipitation attributes identified for each time period. The five major precipitation attributes identified, for each time period. They had the highest percentages of appearances in the 10-fold, cross-validation of the machine learning schemes, applied to the 1965–2018 observed precipitation data set.

Annual	April + May	JJAS	Oct-Mar
GlobalSSTA*TSSST	SOI	SOI	SOI
AMO*SAM	GlobalSSTA*TSSST	Niño3.4	DMI*SAM
AMO*Niño3.4	TPI	SAM	SAM
TPI	GlobalT	TPI	GlobalT*GlobalSSTA
DMI*Niño3.4	DMI	DMI*TSSST	SOI*TPI

4. Conclusions

Australia's northern Murray-Darling Basin (NMDB), which occupies a vast area of subtropical eastern Australia, is found to be experiencing the impacts of accelerated GW on its catchment area rainfall seasonality and effectiveness. Observations of precipitation, maximum temperatures (TMax) and minimum temperatures (TMin) from 1910 to 2018 were divided into four consecutive 27-year periods: 1911–1937, 1938–1964, 1965–1991, and 1992–2018. They revealed that mean April–May precipitation over the NMDB, for the period 1992–2018, decreased significantly (p-value = 0.0318) from that for the period 1965–1991. April–May are the late Autumn season months that moisten the NMDB catchment area prior to the cool season rainfall events that flow into the NMDB river systems. However, there is no compensating increasing trend in the remaining cool-season months of the year (JJAS) to offset the drying in April–May. In the warm season months (October–March), La Niña phases have produced record rainfall in parts of southeast Australia in the years 2010–2012 and, more recently, in 2020–2022. Nevertheless, significant increases in mean TMax and TMin through JJAS and October–March further disrupt the hydrological cycle and reduce the rainfall effectiveness as the plant evapotranspiration increased in the period 1992–2018 due to rising values of TMax. Hence, this combination of decreased April–May precipitation and associated higher daytime temperatures decreases run-off into streams that feed into the Darling River and its tributaries, thereby reducing NMDB water availability, frequently with drastic consequences which are expected to worsen further in the future, as global warming continues.

The precipitation wavelets highlight the importance in 2–8-year periodicity of ENSO for the annual, JJAS (late cool season) and October to March (warm season) periods. The

climate drivers selected as likely attributes for the observed annual and seasonal precipitation patterns varied between the four different periods. For precipitation, the main attributes contributing to the trends in rainfall seasonality variability patterns were identified by using the time series in 10-fold machine learning regression models. It was found that the key attributes differed between the different periods. However, they had in common an emphasis on the role of GW both individually and in combination with known local and large-scale climate drivers, notably ENSO, SAM, IOD, TPI, and Tasman Sea SSTs. Four climate drivers, namely ENSO, IOD, TPI and Tasman Sea SSTs, being ocean-based time series, are highly correlated with the recent quasi-linearly increasing trend in the Global SST time series.

The methods used in this study consist of first investigating the statistical significance of trends in precipitation and temperature between four 27-year time periods spanning 1911 to 2018. Then, ML techniques are applied to identify climate driver attributes of precipitation in the NMDB from nine potential climate drivers. This has the advantage of focusing the importance of many possible combinations of variables down to the most important ones and provides a link to developing hybrid approaches of combining high resolution climate model projections with machine learning models.

Further work, beyond attribute identification, is planned, with the aim of developing sub-seasonal and seasonal outlooks, and annual prediction schemes, likely by combining high resolution climate model projections with machine leaning models. This hybrid approach is expected to be necessary, as traditional machine learning approaches, using training and test data sets, can rapidly lose predictive skill if there is a high level of non-stationarity in the training (observational) data set, especially if it continues into the test data set.

Author Contributions: Conceptualization, M.S. and L.L.; methodology, J.H.; software, J.H.; validation, M.S., J.H. and L.L.; formal analysis, M.S. and L.L.; investigation, M.S.; resources, J.H.; data curation, J.H.; writing—original draft preparation, M.S. and J.H.; writing—review and editing, M.S. and L.L; visualization, M.S. and J.H.; supervision, L.L.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data used is freely available via the web links in the text and references. The precipitation and temperature data are available at: https://doi.org/10.5281/zenodo.6068913.

Acknowledgments: The authors acknowledge the University of Technology Sydney for supporting this research. J.H. acknowledges support from the Australian Government Research Training Program Scholarship.

Conflicts of Interest: The authors declare no conflict of interest.

Code: The software used for the time series, box plots, permutation tests and machine learning, was R (Version 3.5.1). To run forward/backward selection, code was written as in Maldonado and Weber [31]. For the wavelet analysis, MATLAB R2018a was used by adapting the code from [26]. The code can be made available upon request to the second author (J.H.).

References

1. McBride, J.L.; Nicholls, N. Seasonal Relationships between Australian Rainfall and the Southern oscillation. *Mon. Wea. Rev.* **1983**, *111*, 1998–2004.
2. Pook, M.J.; Risbey, J.S.; McIntosh, P.C. A comparative synoptic climatology of cool-season rainfall in major grain-growing regions of southern Australia. *Theor. Appl. Climatol.* **2014**, *117*, 521–533. https://doi.org/10.1007/s00704-013-1021-y.
3. Risbey, J.S.; Pook, M.J.; McIntosh, P.C.; Ummenhofer, C.C.; Meyers, G. Characteristics and variability of synoptic features associated with cool season rainfall in southeastern Australia. *Int. J. Climatol.* **2009**, *29*, 1595–1613. https://doi.org/10.1002/joc.1775.
4. Record-Breaking La Niña Events. Available online: http://www.bom.gov.au/climate/enso/history/La-Nina-2010-12.pdf ISBN: 978 0 642 70621 8 (accessed on 19 August 2022).
5. Flooding in the Murray-Darling Basin. Available online: https://www.mdba.gov.au/issues-facing-basin/flooding-murray-darling-basin (accessed on 19 August 2022).

6. State of the Climate 2020. Available online: https://bom.gov.au/state-of-the-climate/ (accessed on 19 August 2022).
7. Science Solutions for the Murray-Darling Basin. Managing Today's Resources for the Future. Available online: https://publications.csiro.au/rpr/download?pid=csiro:EP206574&dsid=DS1 (accessed on 19 August 2022).
8. Cai, W.; Cowan, T. Southeast Australia Autumn Rainfall Reduction: A Climate-Change-Induced Poleward Shift of Ocean–Atmosphere Circulation. *J. Clim.* **2013**, *26*, 189–205 https://doi.org/10.1175/JCLI-D-12-00035.1.
9. Speer, M.; Leslie, L.M.; MacNamara, S.; Hartigan, J. From the 1990s Climate Change Has Decreased Cool Season Catchment Precipitation and Reduced River Heights in Australia's Southern Murray-Darling Basin. *Sci. Rep.* **2021**, *11*, 16136. Available online: https://www.nature.com/10.1038/s41598-021-95531-4 (accessed on 19 August 2022).
10. Australian Climate Influences. Available online: http://www.bom.gov.au/climate/about/australian-climate-influences.shtml (accessed on 19 August 2022).
11. Murray Darling Basin. Available online: http://www.murrayriver.com.au/about-the-murray/murray-darling-basin/#agriculture (accessed on 19 August 2022).
12. Characterising the Ecological Effects of Changes in the 'Low-Flow Hydrology' of the Barwon-Darling River. Advice to the Commonwealth Environmental Water Holder Office. Available online: https://www.environment.gov.au/system/files/resources/bb774e1f-d7fa-4825-8851-cb5e5f1b3f51/files/characterising-eco-effects-changes-low-flow-barwon-darling.pdf (accessed on 19 August 2022).
13. The Australian Institute. Available online: https://australianinstitute.org.au/report/a-fish-kill-qanda/ (accessed on 19 August 2022).
14. Holland, J.E.; Luck, G.W.; Finlayson, C.M. Threats to food production and water quality in the Murray-Darling Basin of Australia. *Ecosyst. Serv.* **2015**, *12*, 55–70. https://doi.org/10.1016/j.ecoser.2015.02.008.
15. Stephens, C.M.; McVicar, T.R.; Johnson, F.M.; Marshall, L.A. Revisiting pan evaporation trends in Australia a decade on. *Geophys. Res. Lett.* **2018**, *45*, 11,164–11,172. https://doi.org/10.1029/2018GL079332.
16. Independent Assessment of the 2018-19 Fish Deaths in the Lower Darling. Available online: https://www.mdba.gov.au/sites/default/files/pubs/Final-Report-Independent-Panel-fish-deaths-lower%20Darling_4.pdf (accessed on 19 August 2022).
17. The Australian Landscape Water Balance Model (AWRA-L v6). Available online: http://www.bom.gov.au/water/landscape/assets/static/publications/AWRALv6_Model_Description_Report.pdf (accessed on 19 August 2022).
18. Hartigan, J.; MacNamara, S.; Leslie, L.M. Application of machine learning to attribution and prediction of seasonal precipitation and temperature trends in Canberra, Australia. *Climate* **2020**, *8*, 76. https://doi.org/10.3390/cli8060076.
19. Hartigan, J.; MacNamara, S.; Leslie, L.M.; Speer, M. Attribution and prediction of precipitation and temperature trends within the Sydney Catchment using machine learning. *Climate* **2020**, *8*, 120. https://doi.org/10.3390/cli100120.
20. Australian climate change site networks. Available online: http://www.bom.gov.au/climate/change/?ref=ftr#tabs=Tracker&tracker=site-networks (accessed on 19 August 2022).
21. Hydrological Assessment of Flow Changes in the Northern Basin. Available online: https://www.mdba.gov.au/sites/default/files/pubs/1209-Hydrologic-assessment-of-flow-changes-in-the-northern-basin.pdf (accessed on 19 August 2022).
22. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Pachauri et al. Available online : https://epic.awi.de/id/eprint/37530/ (accessed on 19 August 2022).
23. Morice, C.P.; Kennedy, J.J.; Rayner, N.A.; Jones, P.D. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 dataset. *J. Geophys. Res.* **2012**, *117*, D08101. http://doi.org/10.1029/2011JD017187.
24. National Centers for Environmental Information, State of the Climate: Global Climate Report for 2019, Published Online January 2020. Available online: https://www.ncdc.noaa.gov/sotc/global/201913/supplemental/page-3 (accessed on 19 August 2022).
25. Good, P.I. *Resampling Methods: A Practical Guide to Data Analysis,* Birkhäuser: Boston, MA, USA, 2006; ISBN 978-0-8176-4386-7.
26. Lau, K.M.; Weng, H. Climate signal detection using wavelet transform: How to make a time series sing. *Bull. Amer. Meteorol. Soc.* **1995**, *76*, 2391–2402. https://doi.org/10.1175/1520-0477(1995)076<2391:CSDUWT>2.0.CO;2.
27. Torrence, C.; Compo, G.P. A Practical Guide to Wavelet Analysis. *Bull. Amer. Meteor. Soc.* **1998**, *79*, 61–78. https://doi.org/10.1175/15200477(1998)079<0061:APGTWA>2.0.CO;2.
28. Richman, M.B.; Leslie, L.M. Machine Learning for Attribution of Heat and Drought in Southwestern Australia. *Procedia Comput. Sci.* **2020**, *168*, 3–10. https://doi.org/10.1016/j.procs.2020.02.244.
29. Ramsay, H.A.; Richman, M.B.; Leslie, L.M. Seasonal tropical cyclone predictions using optimized combinations of ENSO Regions: Application to the Coral Sea region. *J. Clim.* **2014**, *27*, 8527–8542. https://doi.org/10.1175/JCLI-D-14-00017-1.
30. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 1995; ISBN 978-1-4757-3264-1.
31. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. https://doi.org/10.1023/A:1010933404324.
32. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer Science & Business Media: New York, NY, USA, 2009; ISBN 978-0-387-84858-7.
33. Maldonado, S.; Weber, R. A wrapper method for feature selection using Support Vector Machines. *Inf. Sci.* **2009**, *179*, 2208–2217. https://doi.org/10.1016/j.ins.2009.02.014.
34. Special Climate Statement 74—Extreme Rainfall and Flooding in Eastern and Central Australia in March 2021. Available online: http://www.bom.gov.au/climate/current/statements/scs74.pdf?20210621 (accessed on 19 August 2022).