Disentangling climate change trends in Australian streamflow

Rutger Vervoort[1], Michaela Dolk[2], and Floris van Ogtrop[1]

[1]The University of Sydney
[2]Swiss Re Management (US) Corporation

April 28, 2020

Abstract

The effect of climate change on water resources has been an area of continued research, especially in Australia. Previous studies have suggested significant trends in rainfall, and these are amplified causing larger changes in streamflow. However, most of the previous analysis was based on annual time scales or modelled data and did not account for changes in land cover, which could interact with changes in climate. Climate data and streamflow data sourced from 13 fully forested small catchments (<250 km2) was analysed for trends. Non-parametric Mann-Kendall trend analysis, generalised additive mixed modelling and rainfall-runoff modelling were combined for the analysis. This indicates consistent increases in maximum temperature and varied decreases in rainfall. Limited to a small number of catchments in south eastern Australia there were small, but significantly amplified, decreases in streamflow. In general, overall decreases are much smaller than predicted in earlier research.

Introduction

The effects of climate change will influence water resource availability in the future (Huntington, 2006; Trenberth et al., 2014). This is particularly of importance in Australia where major droughts have drastically increased public concerns about water availability (Leblanc, Tweed, Van Dijk, & Timbal, 2012). As a result, there have been several studies focussing on climate change impacts on water resources (Betts et al., 2007; Chien, Yeh, & Knouft, 2013; Chiew & McMahon, 2002; Chiew et al., 2009).

Most of the earlier studies on future water availability have used rainfall-runoff models with climate data from downscaled global circulation models to highlight future scenarios (e.g. Chien et al., 2013; Chiew et al., 2009; Leblanc et al., 2012; Vaze, Davidson, Teng, & Podger, 2011). For Australia, most of these studies indicate a more variable future and a decreased water availability (van Dijk et al., 2013). The main explanation is decreased rainfall and increased evaporation and therefore decreased streamflow. A drawback of such studies is that they introduce at least two types of errors: the calibration uncertainty of the rainfall-runoff model and the associated uncertainty in the prediction; and uncertainty in future climate data, either from downscaling global models (Leblanc et al., 2012), or from scaling of historical data (Chiew, 2006; Fu, Charles, & Chiew, 2007).

Reductions in streamflow in Australia have also been attributed to increased CO_2 fertilisation of the vegetation due to changes in the atmospheric CO_2 concentrations (Trancoso, Larsen, McVicar, Phinn, & McAlpine, 2017; Ukkola et al., 2015). However, as Ajami et al. (2017) highlight, this effect varies along aridity gradients and depends on other factors such as nutrient availability. Biological work (Franks et al., 2013) suggests that tree stomata adapt to changes in climate, i.e. there is a certain level of bio-plasticity to reduce how water is transpired as a function of temperature. This highlights further other, possibly interacting and counteracting, factors that could influence future changes in streamflow, well beyond the traditional forcing variables.

1

Another issue with the earlier work, including some of the latest work on the changes in catchment responses with drought (Ajami et al., 2017; Booij, Schipper, & Marhaento, 2019; Saft, Western, Zhang, Peel, & Potter, 2015), is that very few of the studies have controlled for land cover and therefore might have stationarity issues (Montanari et al., 2013; Navas, Alonso, Gorgoglione, & Vervoort, 2019). Ajami et al. (2017) analysed vegetation cover and found that runoff sensitivity to precipitation varied with fractional vegetation cover. Moreover, they confirmed that increased vegetation cover decreases runoff sensitivity, confirming earlier work (Chiew, 2006). However, Ajami et al. (2017) also point out that fractional vegetation cover covaries with precipitation, therefore making the comparison difficult.

An alternative approach to biophysical modelling with future climate scenarios to understand changes in future water yield, would be to analyse historical data sets, as several older studies in Australia have done (Booij et al., 2019; Cai & Cowan, 2008; Chiew, Whetton, McMahon, & Pittock, 1995; Petrone, Hughes, Van Niel, & Silberstein, 2010; Potter, Petheram, & Zhang, 2011; Trancoso et al., 2017). However, there are several reasons why the issue of trends in streamflow might need revisiting. Many of the older studies were based on annual data, a limited geographical area (Petrone et al., 2010) and varied in terms data periods and data aggregation in space and time. Some studies focused on a few rivers (Petrone et al., 2010; Potter & Chiew, 2011), some focussing on the whole Murray Darling water balance (Cai & Cowan, 2008) and some focussing on a large data base of catchments (Booij et al., 2019; Chiew, 2006; Trancoso et al., 2017).

Finally, the models and types of analyses varied across the studies, including simple linear regression on un-transformed and transformed data (Cai & Cowan, 2008; Chiew, 2006; Chiew et al., 2014; Potter et al., 2011), non-parametric Mann-Kendall and step change methods (Petrone et al., 2010; Potter & Chiew, 2011), statistical analyses and using rainfall runoff models (Chiew, 2006; Potter, Chiew, & Frost, 2010) and applying an energy framework (Booij et al., 2019) and focussing on baseflow (Trancoso et al., 2017). As a result, despite a consistent view of a sharply declining streamflow in response to changing rainfall, it is difficult to compare across studies.

Rather than using conceptual or deterministic modelling approaches, data based statistical models allow quick identification of the main variables explaining variations in streamflow. If the variables in the statistical model represent real physical processes (Fu et al., 2007), the model removes all variance based on the physical processes and the remaining unexplained trend in the residuals can be analysed. Modern regression approaches also allow non-linearity and the identification of uncertainties (Wood, 2006), rather than being limited to linear relationships. The disadvantage of statistical models is that they are essentially a "black box" as they do not explicitly include any physical processes. This means that, while the model can show correlation, it does not necessarily prove causation, and a lack of understanding of processes within a system may lead to an incorrect interpretation of statistical relationships.

The above review highlights that it is therefore important to have consistent methodology that includes several different statistical analyses to cross check the trend results, as each analysis included methodological and data uncertainty.

Given the range of approaches, the aim of this paper is 1) to offer a combination of statistical approaches to analyse climate change trends in streamflow data in forested catchments; 2) to use the increased observed data length to revisit the work by Chiew (2006) and other past research to investigate climate change trends on a subset of the Australian hydrological reference stations across the continent; and 3) to highlight the spatial and temporal variation in these rainfall and temperature effects on streamflow across the Australian continent.

While the data in this paper focuses on Australia, the outcomes and methodology are relevant for other regions that are experiencing changes in rainfall and temperature and have predicted changes in streamflow.

In light of the complexity of the topic, and the number of existing papers with somewhat conflicting results, all the data, explanations and code for the analyses used in this paper have been stored as supplementary data in an on-line repository, which can be accessed via: https://doi.org/10.5281/zenodo.3757041

Methods

Site selection and data

A total of 13 hydrologic reference stations across Australia, focussing on stations with catchments areas that are less than 250km^2 were selected from the Bureau of Meteorology (http://www.bom.gov.au/water/hrs/) (Figure 1 and Table 1). Using a basic land use data interpretation, we focussed on catchments that are essentially fully forested to minimise effects of landuse variation. By focussing on relatively small catchments, we also aimed to control for catchment size which influences storage and routing.

Rainfall and temperature stations were selected for each streamflow station based on proximity, data length and data continuity. Data for the selected rainfall and temperature sites for the period 1970-2010 were obtained from the Bureau of Meteorology. However, since climate stations were not necessarily close to the catchments, we also extracted the ANUCLIM 1km gridded rainfall product via the eMast NCI server (http://dapds00.nci.org.au/thredds/catalog/rr9/ANUClimate/catalog.html) for each of the catchment outlets.

Analysis of annual trends and elasticities following Chiew (2006)

As a first step annual summaries of the data were analysed, following Chiew (2006). Both the non-parametric approach for calculating the rainfall elasticity, ε_P, (Fu et al., 2007) and the modelling approach (Chiew, 2006) were used. The non-parametric approach calculates the elasticity as (Fu et al., 2007):

$$\epsilon_p = median\left(\frac{Q_t-Q}{P_t-P}\frac{P}{Q}\right)(1)$$

For the modelling approach, both GR4J (Perrin, Michel, & Andréassian, 2003) and SimHyd including the Muskingum routing routine (Chiew et al., 2009) were used on the daily data. The models were all calibrated on the full 40 year daily dataset (1970 – 2010). The calibration used shuffled complex evolution optimisation with 20 complexes, and this was repeated 10 times to capture model uncertainty due to parameter equifinality. The residuals of the modelling were calculated as the difference between observed and predicted values.

Following the original work (Chiew, 2006), the daily rainfall series (P) was scaled by -15%, -10%, 0%, +10%, while the daily maximum temperature ($MaxT$) series, rather than the monthly Potential Evapotranspiration (PET) as in Chiew (2006), was scaled by 0%, +5% and +10%. The resulting predicted daily flow series (Q) were summarised to annual series and ε_P and the elasticity due to ET, ε_{ET}, were calculated using regression (Chiew, 2006):

$$\delta Q_{\mathrm{annual}} = \varepsilon_P \delta P_{\mathrm{annual}} + \varepsilon_{\mathrm{ET}} \delta \mathrm{MaxT}_{\mathrm{annual}}(2)$$

Where, $MaxT_{annual}$ is the average annual maximum daily temperature rather than the maximum annual daily temperature. The elasticity is a measure of amplification of the rainfall change in the streamflow, with elasticities greater than unity indicating amplification of the rainfall and ET signal (Chiew, 2006).

Mann-Kendall non-parametric trend test and linear trends

The presence of monotonic trends in weekly mean streamflow, rainfall and maximum temperature were initially assessed using non-parametric Mann-Kendall tests. The advantages of Mann-Kendall based tests are that there is no assumption that the data are normally distributed, nor does it assume that the trend is linear. However, there are two possible issues with simply using a Mann-Kendall test on the data. The first is outlined by Westra et al. (2012), which highlights that even in random data there is a probability of identifying a trend. As a result, we followed Westra et al. (2012) and ran a bootstrap with 500 samples to identify this error in the basic Mann-Kendall analysis. The second issue is outlined by Koutsoyiannis and Montanari (2007): A locally observed trend is not necessarily a true trend in the data due to scaling

properties. Under the assumption of scaling, the Mann-Kendall trend can be further tested to identify if the Hurst coefficient is significant and the Mann-Kendall test can be corrected for this (Hamed, 2008), we indicate this test as the LTPMK test (long time period Mann-Kendall).

The daily data were de-seasonalised and summarised to weekly before running the Mann-Kendall and LTPMK tests using the package "deseasonalize" (McLeod & Gweon, 2013). Calculations for the standard Mann-Kendall and LTPMK tests were performed using the "Kendall" package (Mcleod, 2011) and "HKProcess" (Tyralis, 2016) in R (R Core Team, 2013), respectively.

In addition to analysing the data series, the trend in the residuals of the modelled series from the previous rainfall-runoff modelling step was also analysed using the same LTPMK test. If the modelled series deviates from the observed series for the post calibration period with a consistent trend, then this would be another indication of a trend in the data series. For this analysis, the residuals were summarised to weekly data to match the observed data. All analyses were repeated with both the station rainfall data and the gridded rainfall data.

Generalized additive mixed model (GAMM)

The basic starting assumption in this analysis is that streamflow (Q), combining both surface runoff and groundwater inflows, in relatively small homogeneous catchments can be modelled as a simple water balance of rainfall (P) and evapotranspiration (ET):

$Q = P - ET$ (3)

Due to the relatively small size of the catchments, storage effects based on the weekly data can be considered minor and the majority of the storage delays can be captured (in other words, residence times of water moving towards the stream are considered to be smaller than a week for the majority of the flow).

Generalized additive mixed models (GAMMs) were defined to model the weekly streamflow at each station as a function of rainfall and weekly mean maximum temperature. GAMMs may be considered as an extension of generalized linear mixed models whereby the predictors are additive, but can also be non-linear (Chen, 2000). The mixed model addition to the standard Generalized Additive Model (GAM), allows modelling of the autocorrelation in the residuals. Details of the statistical basis of GAMMs can be found in Chen (2000) and Lin and Zhang (1999).

Mean weekly streamflow and rainfall are indicated by Q_w and P_w, respectively. A smoothing function, s(), was applied to the rainfall term to account for non-linear responses in the rainfall-runoff relationship. Weekly mean streamflow data was used to remove any effects of filled in missing data by the Bureau of Meteorology, and to remove most of the stream routing effects.

The models were applied in several steps (Table 2). Initially, the linear trends in both rainfall and temperature were analysed using generalised least squares using an autoregressive order 1 (AR1) correlation structure fitted to the errors (e_t). As an added benefit, the AR1 error structure accounts for further storage effects in the catchment, including the effect of past rainfall.

Subsequently a GAMM with a smooth function of P_w, a linear trend and autoregressive order 1 (AR1) correlation structure fitted to the errors (e_t) was evaluated.

$\text{Log}(Q_w) \sim s(P_w) + trend + \varepsilon_t$ (4)

Streamflow data were log transformed to stabilise the variance of the residuals. The trend in equation 3 can be interpreted as a change in the residual $log(Q_w)$ over time. In order, to identify $\underline{\delta X}$

greekt as a fractional change the trend needs to be back transformed using:

$\underline{\delta X}$

greekt$=(\exp(trend)-1)$. (5)

4

The first model (equation 4) removes the direct effect of weekly rainfall (P_w) on the weekly streamflow (Q_w). Amplification can be defined as $\frac{\delta X}{}$

greekP¿1, using the annual data. However, the trend in equation 4 can include other phenomena changing the rainfall-runoff response over time. As a second step, a GAMM with a smooth function of P_w and a smooth function of the interaction between $maxT_w$ and P_w was fitted:

$$\text{Log}\,(Q_w) \sim s\,(P_w) + s\,(\text{max}T_w, P_w) + trend + \varepsilon_t \quad (6)$$

The interaction between $maxT_w$ and P_w is a proxy for the actual evapotranspiration. Equation 6 identifies the residual trends in the streamflow data after removal of the effects of rainfall and actual evapotranspiration. The difference in the trend in the second model (equation 6) and the first model (equation 4) highlights the impact of actual evapotranspiration effects on streamflow. The final trend remaining in the model therefore includes any "other" phenomena in the rainfall-runoff relationship. Examples of these could be changes due to tree physiology, CO_2 fertilization or geomorphology and is the degree of amplification, i.e. the change in streamflow after changes in rainfall and evapotranspiration have been considered.

Using equation 6, a GAMM which removes the linear trend was also fitted, to test whether the assumption of a linear trend affected the results:

$$\text{Log}\,(Q_w) \sim s\,(P_w) + s\,(\text{max}T_w, P_w) + \varepsilon_t \quad (7)$$

The residuals (ϵ_τ) of this final GAMM model (equation 7) were subsequently analysed for a trend using a LTPMK test.

The GAMM modelling was conducted using the "mgcv" package in R (Wood, 2006).

Given the complexity of all the methodology a flow chart of the different approaches and comparisons is presented in Figure 2.

Results

Annual data trends following Chiew (2006)

The relationship between annual rainfall and runoff for the catchments (Figure 3), generally plots below the 1:1 line. Ignoring storage variations, the divergence of the 1:1 line is related to the removal of water due to evapotranspiration. Except for some single points all the data plot below the 1:1 line for the gridded rainfall data. For the station data this is not the case for a few of the catchments (supplementary data on zenodo https://doi.org/10.5281/zenodo.3757041), but the results generally show similar behaviour. One of the catchments, where a lot of the data plotted above the 1:1 line using the observed station rainfall data, is a small catchment (SOUT). Two other catchments (NIVE and HELL) occur in areas with steep elevation gradients, and as a result the nearest rainfall station might not be representative. In contrast the ANUCLIM gridded data appears more representative, as less of the data plot above the 1:1 line.

Non-parametric elasticities (ε_p, red triangles in Figure 4), calculated from the streamflow and gridded rainfall data, were in the same range as those observed by Chiew (2006). All elasticities were less than 3, with most about 2 or above, and some stations indicating elasticities very close to 1 (COCH, HELL and SOUT).

In contrast, elasticities calculated from the simulated data based on the rainfall runoff models were quite different (boxplots and large blue dots indicating mean values in Figure 4), and in fact almost constant. The variation between the 10 replications of the rainfall runoff model calibration was very narrow, particularly for GR4J, so most of the boxplots look like single straight horizontal lines. Overall the elasticities calculated from the rainfall runoff modelling results were close or smaller than 1. This suggest decreases in runoff that are smaller than the associated decreases in rainfall, considering changes in evapotranspiration.

The calibration of the rainfall models with the gridded rainfall data was generally good (NSE > 0.5) considering the 41 year data period, although the SimHyd performance was lower than the GR4J performance (Figure 5), and some catchments had low model performance (Chiew et al., 2009). In addition, the different replications of the calibration with SimHyd showed considerable variation (wide boxplots), while the replications in the GR4J calibrations were very consistent (resulting in the boxplots being plotted as single lines). Comparing the two rainfall sources. the calibration performance using the observed station rainfall data (document 6 supplementary data) was slightly worse for both models and the calculated elasticities were also slightly lower (document 6 supplementary data).

Mann-Kendall and LTPMK tests

The results of the standard Mann-Kendall tests on the de-seasonalised weekly data indicate that there is a consistent decreasing trend in streamflow (Table 3), with a significant (p-value < 0.05) trend for nine of the 13 sites, which are all located in south of the continent and Western Australia. At least five and maybe seven of these sites with significant trends had trends outside the bootstrap distribution (Figure 6).

There is a matching decreasing trend in most of the gridded rainfall records, with a significant (p-value < 0.05) trend for 8 sites. However, only 2 of these are outside the bootstrap distribution (supplementary material document 2). The station rainfall data had similar results to the gridded data. Basically, all the 13 de-seasonalised weekly average maximum temperature records have a significant increasing Mann Kendall trend, even though only 5 of these are falling outside the bootstrap distribution (supplementary material document 2), with locations in south eastern Australia.

In contrast, all the LTPMK results indicated highly significant Hurst coefficients and therefore, based on this analysis, the Man Kendall trends were all not significant for all three variables under the scaling hypothesis (Hamed, 2008). To check the overall test, we also ran the monthly and annual summaries of the three variables. The results (supplementary material document 2A) show that the significance of the LTPMK trends increases with an increase in the time aggregation. For example, the LTPMK average maximum temperature was significant and increasing for six stations at the monthly time scale, and significant and increasing for eight of the 13 stations at the annual time scale. Similar results were found with the time integration for the rainfall and streamflow data, but in this case all significant trends were decreasing in time.

Generalized additive mixed modelling

The results of the generalised least squares (GLS) modelling (model 1 in Table 2) suggest only four of the studied catchments have significant linear trends in the streamflow over the last 41 years (Table 4), which is fewer than in the earlier Mann Kendall results (Table 3 and Figure 6). The negative trends were very small in actual percentage value, in the order of 10^{-4} % of the streamflow (Table 4). Similarly, significant trends in rainfall were also all negative, and very small in actual percentage. For those stations that have significant trends in both rainfall and streamflow, "amplification" (calculated as the trend in Q/trend in P) between 0.95 and 4.29 occurs. This is a larger trend then calculated from the original data and the rainfall-runoff modelling in Figure 3. However, the trend in streamflow derived in the GLS analysis still incorporates the changes in climate in the streamflow, such as changes in rainfall and temperature. These need to be removed to identify the true "amplification" effect.

Comparing the streamflow model results without rainfall (model 2, Table 4) with the results of the model that removes the rainfall effect (model 3, Table 5), indicates a small reduction in the negative trend and the same number of significant stations. Overall the inclusion of rainfall in the model removes approximately < 10 to 30% of the original trends. In other words, the linear trend is the remaining trend in the streamflow after removing the trend in the rainfall. In terms of amplification, this is the additional decrease in streamflow on top of any reduction in the rainfall. However, some of the remaining trends could be due to trends in temperature affecting potential ET.

Overall variation explained by the final models that include both rainfall and evapotranspiration (model 4, Table 2 and Table 6) was low to medium (-0.03 < Adj r^2 < 0.43, supplementary material document 3C). The worst performing model was for station NIVE, but more generally these results suggest that there are other processes causing variation in the runoff that the statistical model does not include. This is not necessarily an issue, as the goal of the statistical modelling was to explain the maximum variation in the streamflow related to rainfall and ET, and not to find the best predictive model.

After rainfall and evapotranspiration effects are accounted for (model 4), the models identify significant trends in the weekly data in only six catchments (Table 6). These trends are once again very small and only in the order of $1.1 - 5.5 \times 10^{-4}$ % change. This suggests that the overall streamflow is indeed declining over time, even after accounting for changes in rainfall and evapotranspiration, but currently the overall change is very small. This remaining trend is the amplification.

Note that inclusion of the maximum temperature (evaporation) explains very little of the variation in streamflow, as the improvement in the performance measure AIC (Aikaike Information Criterium) between model 3 (Table 5) and model 4 (Table 6) is small. As a result, for those catchments still showing significant trends, the variance explained by including rainfall and temperature is small, from < 10% to 30%. As with the previous analyses (Chiew et al., 2009; Potter & Chiew, 2011), the catchments with significant trends are all located in the South and West of Australia.

For the Mann-Kendall tests on the remaining residuals, eight catchments have significant trends (Table 6, Figure 7), again more than with the regression modelling. However, the bootstrap results in Figure 7 suggest that even fewer of the sites have true trends, with a minimum of one and a maximum of three of the residual trends clearly falling outside the bootstrap distribution, including some on the outer edge. The overlap between the results from the Mann-Kendall residual analysis and the trend in the GAMM is clear for most of the locations, except for the very small YARR, for which the residual analysis suggests there is no significant trend, but the GAMM suggests the linear trend is significant. In this case, the actual trend might not be linear, but in general the assumption of a linear trend is not influencing the results. Moreover, there is clear overlap between the LTPMK analysis and the traditional Mann-Kendall analysis, with the same stations showing significance, despite the Hurst coefficient being significant for all the stations.

Model non-stationarity over long periods

This part of the analysis tested whether numerical models used for the analysis of trends and elasticities impact the analysis of the elasticities presented earlier. This is important as several climate change studies are based on model output analysis (Chiew et al., 2009; Potter & Chiew, 2011; Vaze et al., 2011) rather than observed data. As calibrated models essentially fit a stationary series, the residuals (between observed and simulated) should retain any existing trend in the data. This builds on earlier analyses (Buzacott, Tran, van Ogtrop, & Vervoort, 2019; Saft et al., 2015; Vaze et al., 2010), which contrasted calibration differences in dry and wet periods to investigate non-stationarity.

The result of the Mann-Kendall non-parametric analysis on the weekly residuals (observed – predicted) for the 41 years of the data series indicates significant slopes for only part of the stations (Figure 8). While there appears to be no consistent pattern, it indicates that for SimHyd (at least for some of the stations) there are stronger decreasing trends than for GR4J. In contrast, the residuals of GR4J had more significant trends in the standard Mann Kendall analysis, but also predicted some (very small) positive trends. Overall the identified catchments with significant slopes, CORA COTT, RUTH, SOUT and MURR, match the GAMM and earlier Mann-Kendall analysis on the data. Again, the small YARR catchment is an outlier, indicating a significant slope in the Mann-Kendall analysis, but positive for the GR4J residuals and negative for the SimHyd residuals. However, in both cases the actual slope value is very close to 0. Overall this matches the uncertainty for this catchment indicated in the earlier analyses.

The LTPMK tests however indicated that none of the identified trends were significant under this test (supplementary documents, document 6A). This essentially shows that identified slope in the model residuals

exists locally, but currently the overall variation in the data means we cannot yet affirm a long-term trend in the residuals under the scaling hypothesis (Hamed, 2008).

Overall, these results are broadly similar to the earlier analyses, suggesting trends in only some of the South-eastern Australian stations, but with many of the trends being very small. All analyses also indicate the largest negative trend for the RUTH catchment.

Discussion

Summarising, this study used multiple approaches to confirm that in general streamflow in south-eastern Australia is declining based on the 41 years of data, and that this decline tends to be greater than the associated decline in rainfall (Tables 4 and 6). In other words, there is significant "amplification" of the observed rainfall declines. However, the overall declines, for small forested catchments and considering 41 years of data, is (still) very small in percentage and mm year^{-1} terms, once variability of rainfall and temperature is accounted for. In addition, using the LTPMK approach, which accounts for the autocorrelation in the data and test for significance of the Hurst coefficient, most of the approaches indicated non-significant trends. This indicates that while there is a clear local trend in the 41 years of some of the data, we cannot (yet) say with confidence whether this represents a long-term declining trend in streamflow.

Comparison to other studies

In comparison, an analysis of Australian streamflow records for the latter half of the 20th Century using a Mann-Kendall test by Chiew and McMahon (1993) found significant declines at 7 of the 30 studied stations. That study points out that due to the high inter-annual variability of Australian streamflow, a greater change might be required to identify a statistically significant trend (Chiew & McMahon, 1993). A similar observation can be made in this study considering the significance under the LTPMK test. Significant larger declines in streamflow have been observed by Petrone et al. (2010), for catchments in southwest Western Australia between 1950-2008. A more recent paper identified significant trends in 20 of 199 catchments, concentrated in the South-East of Australia (Ajami et al., 2017). An analysis of global data (Milly, Dunne, & Vecchia, 2005) similarly indicated decreases in streamflow in Australia, but the database from that study overlaps with the other studies.

In general, decreasing trends in rainfall in the south east of Australia and increasing trends in temperature have been widely identified (Cai & Cowan, 2012; Chowdhury, Beecham, Boland, & Piantadosi, 2015; Murphy & Timbal, 2008). In contrast, positive trends in rainfall were found in the 20th Century in earlier studies (Collins & Della-Marta, 1999; Smith, 2004), although the data period was suggested to be an issue (Smith, 2004). Some other studies have found much larger decreases in annual rainfall (Gallant, Hennesy, & Risbey, 2007; Murphy & Timbal, 2008) in the order of 1 mm per year and greater decreases in streamflow (Petrone et al., 2010). The differences are most likely related to the difference in aggregation time scale (annual anomalies versus weekly data). Another complicating factor could have been that many of these earlier studies had timeseries to 2008, which was at the end of the millennium drought (van Dijk et al., 2013), while our study included the drought breaking year 2010. This "local trend" phenomena can also explain why many of our LTPMK tests showed no significant trends. This highlights once again the need for careful consideration of the trend tests applied to the data.

The low rainfall elasticities calculated from the rainfall runoff simulations (Figure 4) differ from earlier research (Chiew, 2006), which suggested much higher amplifications from their SimHyd modelling. There are however differences with the original work (Chiew, 2006), which could explain the outcomes. Our study simulates the streamflow on a daily scale and scales this to annual data, while the original work calibrated on monthly data and scaled to annual. As larger time scales are smoother and show more significant trends than weekly and daily data (see the monthly and annual Mann Kendall results in the supplementary documents, document 3) this could have also affected the amplification calculations.

There could be differences in the magnitude of the analysed trends in this study compared to other studies, due to the differences in time scale. A reduction in variance occurs when the data is summarised to annual and monthly data (Cai & Cowan, 2008; Petrone et al., 2010). In addition, there could be differences due to the translation of remotely sensed (Ukkola et al., 2015), filtered (Trancoso et al., 2017) or modelled data (Chiew et al., 2009) to streamflow declines. Several other studies investigated mainly the difference between decades or specific periods (Chiew et al., 2014; Potter et al., 2010; Saft et al., 2015).

Furthermore, our work concentrates on 13 single land cover (forested) small catchments, while several older studies did not control for land cover, but mainly selected "unimpaired" catchments (meaning no water resource development). The data based analyses (triangles in Figure 4), and the GAMM modelling supports the magnitude of the amplification (Chiew, 2006) for the few catchments that showed significant trends. However, the non-parametric rainfall elasticities (grey triangles in Figure 4) could be biased (Fu et al., 2007) as this does not take into account the full interaction between rainfall and temperature.

Remaining trends

What the remaining trends in the streamflow (after removing rainfall and temperature effects) exactly represent is still difficult to say. Amplification has been mainly explained as an additional drying effect (Chiew, 2006), mostly driven by the increases in temperature. Given that in most catchments the remaining trend (Table 6) represents up to 80% of the original trend in the streamflow (Table 4), this would correspond to amplifications around 4 (the data suggests 0.98 - 4.5).

There is no indication that streamflow in these forested catchments has decreased by a large amount (24 - 28%) due to increases in evapotranspiration as a result of CO_2 fertilisation (Ukkola et al., 2015; van Dijk et al., 2013). Because we selected homogeneous land cover (forestry), the larger declines found in earlier studies (Ukkola et al., 2015; van Dijk et al., 2013) might be due to trends in land management and land cover, such as changes in cropping patterns, introduction of conservation tillage, or terracing and crop variety effects. Furthermore, changes in the vegetation water use efficiencies and changes in the tree physiology (Franks et al., 2013; Ukkola et al., 2015) can also explain part of the remaining trends. Conceptually, if trees become more water use efficient and produce more leaf mass, then this could affect evapotranspiration and streamflow. In addition, larger leaf area could also result in increased interception, specifically for smaller rainfall events (Stoof et al., 2012). This clearly highlights that more work in the area of vegetation effects on streamflow would be useful, given the dominance of transpiration in the global water balance (Jasechko et al., 2013), to support investigations into climate change effects.

Distribution changes

Changes in the timing and distribution of rainfall could impact the amplification of the rainfall signal in the streamflow. Shifts in the main rainfall season from winter towards summer, or changes in the average storm depth and timing between rainfall events can all affect the runoff process (Ishak, Rahman, Westra, Sharma, & Kuczera, 2013; Westra et al., 2012) and shift the balance between surface runoff and groundwater flow. For example, Trancoso et al. (2017) recently found declines of about 1 mm/year focussing on baseflow. We compared streamflow and rainfall distribution curves across the four decades in the data (supplementary data, document 4) but found no obvious or clear patterns. But this analysis was essentially qualitative and a more detailed analysis of changes in the distribution of flows and rainfall is warranted.

Limitations

However, there are also some limitations to the present study. Small catchments might have smaller temperature effects on streamflow (as there is less time for evaporation to work on the catchment storage). However, the advantages of controlling for routing and land cover in smaller catchments were important for this study.

Some of the reasons for the low performance of the GAMM models could be that the weekly data is too coarse to pick up some of the non-linear rainfall responses, or that the AR1 model is not complex enough to model the remaining storage delays in the weekly data despite the small catchments. However, computational difficulties and missing data make working with daily data in GAMM challenging.

The low impact of rainfall and ET on explaining variation in streamflow contrasts the earlier work in this area using mainly linear regression and with annual and monthly data (Cai & Cowan, 2008; Potter & Chiew, 2011). We believe that this difference has to do with the differences in aggregation time scale and the non-linearity of the processes.

The value of using observed data to analyse trends is that the model bias (Figure 5) is not affecting the results. However, finding the right method of analysis is tricky (Fu et al., 2007). We chose to use statistical modelling to remove the need to make the assumptions related to mechanistic rainfall-runoff models. This also removes the need for water balance closure, which can affect the parameter calibration (Vervoort, Miechels, van Ogtrop, & Guillaume, 2014).

This study has concentrated on a relatively small group of Australian catchments as an example, because we controlled for catchment size and land cover. We identified about 5 more catchments that could be included, but these were either very small or would skew the spatial distribution of the catchments. However, future research could be to take advantage of other (global) long term databases of unimpaired and forested catchments, such as the Critical Zone Observatories in the US (Troch et al., 2015) or the well-publicized French catchment database (Oudin, Andréassian, Perrin, Michel, & Le Moine, 2008) and other European catchments (Euser et al., 2013; Fenicia et al., 2014; Kirchner, 2009). This would identify whether the trends in streamflow are unique to Australia, or link in with global changes in the hydrological cycle (Huntington, 2006). Another avenue of future research would be to run the same analysis on a comparison catchment dataset which is strictly agricultural. Based on this research, we would hypothesise that these catchments would demonstrate greater changes in streamflow and larger elasticities.

Conclusions

This work highlights that trends in streamflow can be difficult to detect, even using a range of methods. The analysis is sensitive to the aggregation scale, the noise in the data and the length of the data series. In this work, only very small downward trends in the streamflow for 13 hydrological reference stations in Australia were observed between $1970 - 2010$, and only four stations indicated consistent statistically significant trends across the different methods applied. This is despite evidence of increasing trends in temperature in all catchments, and declining rainfall in the southern part of the continent. While the analyses identified trends in the $1970 - 2010$ data, the Mann Kendall analysis under scaling hypothesis did not always identify significant trends, meaning that we cannot yet draw a definite conclusion on the long term decline in streamflow, but can be definite about trends in the 41 years under investigation.

In the catchments where a statistical downward trend was detected, only a small part of the trend could be directly explained by changes in rainfall and temperature. As a result, extrapolation of models calibrated on current datasets might bias future predictions, specifically if the assumption is that the runoff response to rainfall remains stationary.

Currently there is no clear biophysical explanation for the remaining negative trends after accounting for rainfall and temperature changes (the observed amplification). Given that the catchments were chosen based on homogeneous landcover, this suggests that possible vegetation effects impact the observed trends. An alternative explanation is that timing and distribution of rainfall affects the rainfall runoff response. Both these areas will require further research.

Data Availability Statement

The data that support the findings of this study are openly available in https://doi.org/10.5281/zenodo.3757041.

References

Ajami, H., Sharma, A., Band, L. E., Evans, J. P., Tuteja, N. K., Amirthanathan, G. E., & Bari, M. A. (2017). On the non-stationarity of hydrological response in anthropogenically unaffected catchments: an Australian perspective. *Hydrol. Earth Syst. Sci., 21* (1), 281-294. doi:10.5194/hess-21-281-2017

Betts, R. A., Boucher, O., Collins, M., Cox, P. M., Falloon, P. D., Gedney, N., . . . Webb, M. J. (2007). Projected increase in continental runoff due to plant responses to increasing carbon dioxide.*Nature, 448* (7157), 1037-1041. doi:http://www.nature.com/nature/journal/v448/n7157/suppinfo/nature06045_S1.html

Booij, M. J., Schipper, T. C., & Marhaento, H. (2019). Attributing Changes in Streamflow to Land Use and Climate Change for 472 Catchments in Australia and the United States. *Water, 11* (5), 1059.

Buzacott, A. J. V., Tran, B., van Ogtrop, F. F., & Vervoort, R. W. (2019). Conceptual Models and Calibration Performance—Investigating Catchment Bias. *Water, 11* (11), 2424.

Cai, W., & Cowan, T. (2008). Evidence of impacts from rising temperature on inflows to the Murray-Darling Basin. *Geophysical Research Letters, 35* (7), n/a-n/a. doi:10.1029/2008GL033390

Cai, W., & Cowan, T. (2012). Southeast Australia Autumn Rainfall Reduction: A Climate-Change-Induced Poleward Shift of Ocean–Atmosphere Circulation. *Journal of Climate, 26* (1), 189-205. doi:10.1175/JCLI-D-12-00035.1

Chen, C. (2000). Generalized additive mixed models. *Communications in Statistics - Theory and Methods, 29* (5-6), 1257-1271. doi:10.1080/03610920008832543

Chien, H., Yeh, P. J. F., & Knouft, J. H. (2013). Modeling the potential impacts of climate change on streamflow in agricultural watersheds of the Midwestern United States. *Journal of Hydrology, 491* (0), 73-88. doi:http://dx.doi.org/10.1016/j.jhydrol.2013.03.026

Chiew, F. H. S. (2006). Estimation of rainfall elasticity of streamflow in Australia. *Hydrological Sciences Journal, 51* (4), 613-625. doi:10.1623/hysj.51.4.613

Chiew, F. H. S., & McMahon, T. A. (1993). Detection of trend or change in annual flow of Australian rivers. *International Journal of Climatology, 13* (6), 643-653. doi:10.1002/joc.3370130605

Chiew, F. H. S., & McMahon, T. A. (2002). Modelling the impact of climate change on Australian streamflow. *Hydrological processes, 16* , 1235-1245.

Chiew, F. H. S., Potter, N. J., Vaze, J., Petheram, C., Zhang, L., Teng, J., & Post, D. A. (2014). Observed hydrologic non-stationarity in far south-eastern Australia: implications for modelling and prediction.*Stochastic Environmental Research and Risk Assessment, 28* (1), 3-15.

Chiew, F. H. S., Teng, J., Vaze, J., Post, D. A., Perraud, J. M., Kirono, D. G. C., & Viney, N. R. (2009). Estimating climate change impact on runoff across southeast Australia: Method, results, and implications of the modeling method. *Water Resources Research, 45* (10), W10414. doi:10.1029/2008WR007338

Chiew, F. H. S., Whetton, P. H., McMahon, T. A., & Pittock, A. B. (1995). Simulation of the impacts of climate change on runoff and soil moisture in Australian catchments. *Journal of Hydrology, 167* (1-4), 121-147.

Chowdhury, R. K., Beecham, S., Boland, J., & Piantadosi, J. (2015). Understanding South Australian rainfall trends and step changes.*International Journal of Climatology, 35* (3), 348-360. doi:10.1002/joc.3982

Collins, D., & Della-Marta, P. (1999). Annual climate summary 1998: Australia's warmest year on record. *Aust. Met. Mag, 48* , 273-283.

Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., & Savenije, H. H. G. (2013). A framework to assess the realism of model structures using hydrological signatures. *Hydrol. Earth Syst. Sci., 17* (5), 1893-1912. doi:10.5194/hess-17-1893-2013

Fenicia, F., Kavetski, D., Savenije, H. H. G., Clark, M. P., Schoups, G., Pfister, L., & Freer, J. (2014). Catchment properties, function, and conceptual model representation: is there a correspondence?*Hydrological processes, 28* (4), 2451-2467. doi:10.1002/hyp.9726

Franks, P. J., Adams, M. A., Amthor, J. S., Barbour, M. M., Berry, J. A., Ellsworth, D. S., . . . von Caemmerer, S. (2013). Sensitivity of plants to changing atmospheric CO2 concentration: from the geological past to the next century. *New Phytologist, 197* (4), 1077-1094. doi:10.1111/nph.12104

Fu, G., Charles, S. P., & Chiew, F. H. S. (2007). A two parameter climate elasticity of streamflow index to assess climate change effects on annual streamflow. *Water Resources Research, 43* , W11419. doi:doi:10.1029/2007WR005890

Gallant, A. J. E., Hennesy, K. J., & Risbey, J. (2007). Trends in rainfall indices for six Australian regions: 1910-2005. *Australian Meteorological Magazine, 56* , 223-239.

Hamed, K. H. (2008). Trend detection in hydrologic data: The Mann–Kendall trend test under the scaling hypothesis. *Journal of Hydrology, 349* (3), 350-363. doi:https://doi.org/10.1016/j.jhydrol.2007.11.009

Huntington, T. G. (2006). Evidence for intensification of the global water cycle: Review and synthesis. *Journal of Hydrology, 319* (1-4), 83-95.

Ishak, E. H., Rahman, A., Westra, S., Sharma, A., & Kuczera, G. (2013). Evaluating the non-stationarity of Australian annual maximum flood.*Journal of Hydrology, 494* (0), 134-145. doi:http://dx.doi.org/10.1016/j.jhydrol.2013.04.021

Jasechko, S., Sharp, Z. D., Gibson, J. J., Birks, S. J., Yi, Y., & Fawcett, P. J. (2013). Terrestrial water fluxes dominated by transpiration. *Nature, 496* (7445), 347-350. doi:10.1038/nature11983

http://www.nature.com/nature/journal/v496/n7445/abs/nature11983.html#supplementary-information

Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resour. Res., 45* , W02429.

Koutsoyiannis, D., & Montanari, A. (2007). Statistical analysis of hydroclimatic time series: Uncertainty and insights. *Water Resources Research, 43* (5). doi:doi:10.1029/2006WR005592

Leblanc, M., Tweed, S., Van Dijk, A., & Timbal, B. (2012). A review of historic and future hydrological changes in the Murray-Darling Basin.*Global and Planetary Change, 80–81* (0), 226-246. doi:http://dx.doi.org/10.1016/j.gloplacha.2

Lin, X., & Zhang, D. (1999). Inference in generalized additive mixed models by using smoothing splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 61* (2), 381-400. doi:10.1111/1467-9868.00183

Mcleod, A. I. (2011). Kendall: Kendall rank correlation and Mann-Kendall trend test. R package version 2.2. In: https://CRAN.R-project.org/package=Kendall.

McLeod, A. I., & Gweon, H. (2013). Optimal Deseasonalization for Monthly and Daily Geophysical Time Series. *Journal of Environmental Statistics, 4* (11).

Milly, P. C. D., Dunne, K. A., & Vecchia, A. V. (2005). Global pattern of trends in streamflow and water availability in a changing climate.*Nature, 438* (17 November 2005), 347-350.

Montanari, A., Young, G., Savenije, H. H. G., Hughes, D., Wagener, T., Ren, L. L., . . . Belyaev, V. (2013). "Panta Rhei—Everything Flows": Change in hydrology and society—The IAHS Scientific Decade 2013–2022.*Hydrological Sciences Journal, 58* (6), 1256-1275. doi:10.1080/02626667.2013.809088

Murphy, B. F., & Timbal, B. (2008). A review of recent climate variability and climate change in southeastern Australia. *International Journal of Climatology, 28* (7), 859-879.

Navas, R., Alonso, J., Gorgoglione, A., & Vervoort, R. W. (2019). Identifying Climate and Human Impact Trends in Streamflow: A Case Study in Uruguay. *Water, 11* (7), 1433.

Oudin, L., Andréassian, V., Perrin, C., Michel, C., & Le Moine, N. (2008). Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments. *Water Resources Research, 44* (3), W03413. doi:10.1029/2007wr006240

Perrin, C., Michel, C., & Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology, 279* (1–4), 275-289. doi:http://dx.doi.org/10.1016/S0022-1694(03)00225-7

Petrone, K. C., Hughes, J. D., Van Niel, T. G., & Silberstein, R. P. (2010). Streamflow decline in southwestern Australia, 1950–2008. *Geophysical Research Letters, 37* (11), n/a-n/a. doi:10.1029/2010GL043102

Potter, N. J., & Chiew, F. H. S. (2011). An investigation into changes in climate characteristics causing the recent very low runoff in the southern Murray-Darling Basin using rainfall-runoff models. *Water Resources Research, 47* (12).

Potter, N. J., Chiew, F. H. S., & Frost, A. J. (2010). An assessment of the severity of recent reductions in rainfall and runoff in the Murray–Darling Basin. *Journal of Hydrology, 381* (1–2), 52-64. doi:http://dx.doi.org/10.1016/j.jhydrol.2

Potter, N. J., Petheram, C., & Zhang, L. (2011). *Sensitivity of streamflow to rainfall and temperature in south-eastern Australia during the Millennium drought.* Paper presented at the 19th International Congress on Modelling and Simulation, Perth, Dec.

R Core Team. (2013). R: A language and environment for statistical

computing. Vienna, Austria.: R Foundation for Statistical Computing Retrieved from http://www.R-project.org/

Saft, M., Western, A. W., Zhang, L., Peel, M. C., & Potter, N. J. (2015). The influence of multiyear drought on the annual rainfall-runoff relationship: An Australian perspective. *Water Resources Research, 51* (4), 2444-2463. doi:10.1002/2014WR015348

Smith, I. (2004). An assessment of recent trends in Australian rainfall. *Australian Meteorological Magazine, 53* (3), 163-173.

Stoof, C. R., Vervoort, R. W., Iwema, J., van den Elsen, E., Ferreira, A. J. D., & Ritsema, C. J. (2012). Hydrological response of a small catchment burned by experimental fire. *Hydrol. Earth Syst. Sci., 16* (2), 267-285. doi:10.5194/hess-16-267-2012

Trancoso, R., Larsen, J. R., McVicar, T. R., Phinn, S. R., & McAlpine, C. A. (2017). CO2-vegetation feedbacks and other climate changes implicated in reducing base flow. *Geophysical Research Letters, 44* (5), 2310-2318. doi:10.1002/2017GL072759

Trenberth, K. E., Dai, A., van der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., & Sheffield, J. (2014). Global warming and changes in drought. *Nature Clim. Change, 4* (1), 17-22. doi:10.1038/nclimate2067

Troch, P. A., Lahmers, T., Meira, A., Mukherjee, R., Pedersen, J. W., Roy, T., & Valdes-Pineda, R. (2015). Catchment coevolution: A useful framework for improving predictions of hydrological change? *Water Resources Research, 51* (7), 4903-4922. doi:10.1002/2015WR017032

Tyralis, H. (2016). HKprocess: Hurst-Kolmogorov Process. R package version 0.0-2. . In: https://CRAN.R-project.org/package=HKprocess.

Ukkola, A. M., Prentice, I. C., Keenan, T. F., van Dijk, A. I. J. M., Viney, N. R., Myneni, R. B., & Bi, J. (2015). Reduced streamflow in water-stressed climates consistent with CO2 effects on vegetation. *Nature Clim. Change, advance online publication* . doi:10.1038/nclimate2831

http://www.nature.com/nclimate/journal/vaop/ncurrent/abs/nclimate2831.html#supplementary-information

van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., de Jeu, R. A. M., Liu, Y. Y., Podger, G. M., . . . Viney, N. R. (2013). The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resources Research, 49* (2), 1040-1057. doi:10.1002/wrcr.20123

Vaze, J., Davidson, A., Teng, J., & Podger, G. (2011). Impact of climate change on water availability in the Macquarie-Castlereagh River Basin in Australia. *Hydrological processes, 25* (16), 2597-2612. doi:10.1002/hyp.8030

Vaze, J., Post, D. A., Chiew, F. H. S., Perraud, J. M., Viney, N. R., & Teng, J. (2010). Climate non-stationarity – Validity of calibrated rainfall–runoff models for use in climate change studies. *Journal of Hydrology, 394* (3–4), 447-457. doi:http://dx.doi.org/10.1016/j.jhydrol.2010.09.018

Vervoort, R. W., Miechels, S. F., van Ogtrop, F. F., & Guillaume, J. H. A. (2014). Remotely sensed evapotranspiration to calibrate a lumped conceptual model: Pitfalls and opportunities. *Journal of Hydrology, 519, Part D* (0), 3223-3236. doi:http://dx.doi.org/10.1016/j.jhydrol.2014.10.034

Westra, S., Alexander, L. V., & Zwiers, F. W. (2012). Global Increasing Trends in Annual Maximum Daily Precipitation. *Journal of Climate, 26* (11), 3904-3918. doi:10.1175/JCLI-D-12-00502.1

Wood, S. (2006). *Generalized additive models: an introduction with R* . Boca Raton, FL: CRC Press.

Tables

Table 1 Overview of hydrologic reference stations and characteristics

Abbreviation	Gauge Number	State	Catchment area (km^2)	Latitude	Longitude	Mean annual flow
COTT	410730	ACT	130	-35.592	148.821	327.5
RUTH	219001	NSW	14	-36.593	149.442	451.3
CORA	215004	NSW	166	-35.147	150.033	305.5
ELIZ	G8150018	NT	96	-12.606	131.074	643.1
COCH	113004A	QLD	95	-17.74	145.629	1854.0
COEN	922101B	QLD	170	-13.956	143.175	685.2
SCOT	A5030502	SA	29	-35.099	138.673	125.9
HELL	312061	TAS	101	-41.419	145.67	1316.7
NIVE	304497	TAS	174	-42.031	146.42	1098.1
MURR	405205	VIC	106	-37.414	145.564	496.4
SOUT	225020A	VIC	10	-37.833	146.354	1330.7
YARR	614044	WA	80	-32.812	116.155	23.4
DOMB	607155	WA	116	-34.579	115.972	309.2

Table 2 Overview of different statistical models and trend analysis related to the models

No	Model	Trend	Analysis
1	$\text{Log}(Q) \sim trend + \epsilon_t$	linear	40-year trend in streamflow to compare with LTPMK analysis
2	$\text{Log}(P) \sim trend + \epsilon_t$	linear	40-year trend in rainfall to compare with LTPMK analysis

No	Model	Trend	Analysis
3	$\text{Log}(Q) \sim$ $s(P) + trend + \epsilon_t$ (equation 4)	linear	The trend in this model relative to model 1 indicates the importance of "other processes", while the comparison with model 2 indicates the "amplification".
4	$\text{Log}(Q) \sim s(P) +$ $s(maxT, P) + trend + \epsilon_t$ (equation 6)	linear	Difference between model 3 and 4 is the effect of evapotranspiration on the trend. The remaining trend is related to changes over time in the rainfall runoff response.
5	$\text{Log}(Q) \sim$ $s(P) + s(maxT, P) + \epsilon_t$ (equation 7)	Mann Kendall and LTPMK	Check if linear trend assumption is biased.

Table 3 Mann-Kendall and Long time period Man Kendall (LTPMK) test results on de-seasonalised weekly time series.

	Mann Kendall	Mann Kendall	LTPMK	
Catchment	tau	p-value	Hurst p-value	adj p-value
	Stream flow			
COTT	-0.17	0.00	0.00	0.70
RUTH	-0.25	0.00	0.00	0.39
CORA	-0.08	0.00	0.00	0.80
ELIZ	-0.01	0.71	0.00	0.89
COCH	-0.02	0.24	0.00	0.98
COEN	-0.02	0.18	0.00	0.98
SCOT	-0.02	0.10	0.00	0.91
HELL	-0.10	0.00	0.00	0.88
NIVE	-0.04	0.01	0.00	0.95
MURR	-0.22	0.00	0.00	0.63
SOUT	-0.15	0.00	0.00	0.79
YARR	-0.15	0.00	0.00	0.69
DOMB	-0.05	0.00	0.00	0.81
	Gridded Rainfall	Gridded Rainfall		
COTT	-0.05	0.00	0.00	0.04
RUTH	-0.02	0.09	0.00	0.52
CORA	-0.02	0.22	0.00	0.58
ELIZ	0.01	0.59	0.00	0.98
COCH	0.00	0.89	0.00	0.97
COEN	-0.01	0.43	0.00	0.99
SCOT	-0.04	0.01	0.00	0.53
HELL	-0.01	0.54	0.00	0.93
NIVE	-0.03	0.03	0.00	0.41
MURR	-0.05	0.00	0.00	0.22

	Mann Kendall	**Mann Kendall**	**LTPMK**	
SOUT	-0.08	0.00	0.00	0.01
YARR	-0.01	0.65	0.00	0.94
DOMB	-0.03	0.02	0.00	0.83
	Maximum Temperature	Maximum Temperature		
COTT	0.14	0.00	0.00	0.90
RUTH	0.14	0.00	0.00	0.90
CORA	0.07	0.00	0.00	0.93
ELIZ	0.07	0.00	0.00	0.85
COCH	0.07	0.00	0.00	0.93
COEN	0.13	0.00	0.00	0.85
SCOT	0.04	0.01	0.00	0.97
HELL	0.06	0.00	0.00	0.94
NIVE	0.06	0.00	0.00	0.94
MURR	0.10	0.00	0.00	0.92
SOUT	0.10	0.00	0.00	0.92
YARR	0.07	0.00	0.00	0.95
DOMB	0.04	0.00	0.00	0.97

Table 4 log linear trends in streamflow (Q) and rainfall (P), conversion to % change and associated p-values (models 1 and 2 in Table 2)

Catchment	Linear trend estimate in Q (log scale) $\times 10^{-4}$	p- value for trend in Q	Trend in Q ($\times 10^{-4}$ %)	Linear trend estimate in P $\times 10^{-4}$	p-value for trend in P	Trend in P ($\times 10^{-4}$ %)
COTT	-3.55	0.01	-3.54	-1.67	0.00	-1.67
RUTH	-7.99	0.00	-7.98	-1.86	0.01	-1.86
CORA	-3.23	0.01	-3.23	-1.50	0.02	-1.50
ELIZ	4.50	0.12		8.31	0.44	
COCH	-1.02	0.93		-0.73	0.40	
COEN	2.59	0.48		-2.98	0.06	-2.98
SCOT	-2.85	0.31		-0.53	0.44	
HELL	-1.68	0.17		-0.41	0.48	
NIVE	-0.63	0.79		-1.08	0.04	-1.08
MURR	-2.63	0.00	-2.63	-0.94	0.11	
SOUT	-1.58	0.00	-1.58	-1.67	0.00	-1.67
YARR	-2.41	0.53		-0.19	0.83	
DOMB	-2.25	0.57		-1.05	0.24	

Table 5 Results of the GAMM modelling with only P as a smooth term and a trend (model 3). The results indicate the AIC (Aikaike Information Criterion) and the estimate of the log transformed trend and the significance of the trend term, followed by (for those trends that were significant at p < 0.05) the trend as % change in streamflow.

Catchment	AIC	Trend estimate ($\times 10^{-4}$ %)	Trend p-value	Trend ($\times 10^{-4}$ %)
COTT	1347.3	-2.31	0.03	-2.31

16

Catchment	AIC	Trend estimate (×10⁻⁴ %)	Trend p-value	Trend (×10⁻⁴ %)
RUTH	1956.8	-5.37	0.00	-5.37
CORA	3801.0	-1.63	0.00	-1.63
ELIZ	2737.1	-0.30	0.89	
COCH	2157.5	-0.01	0.99	
COEN	2790.9	-0.13	0.95	
SCOT	1561.0	-0.64	0.45	
HELL	2901.6	-1.37	0.16	
NIVE	3394.7	-0.61	0.69	
MURR	774.9	-2.03	0.00	-2.03
SOUT	2239.9	-1.14	0.02	-1.14
YARR	-512.9	-1.10	0.02	-1.10
DOMB	2127.6	-1.14	0.52	

Table 6 Results of the GAMM modelling with both P and maximum Temperature as a smooth term (Model 4) including trend analysis. The first three columns relate to the results of the GAMM model which includes a trend (model 4), while the second three columns relate to the GAMM model without a trend (model 5), for which the residuals are analyzed use a Mann-Kendall test.

Catchment	AIC model 4[1]	Trend estimate (log scale) ×10⁻⁴	Trend p-value	Trend[2] (×10⁻⁴ %)	AIC model 5[3]	MK tau residuals	p-value MK	p-value Hurst coefficient	Ac... p-v...
COTT	997.8	-2.36	0.06	-2.36	1007.8	-0.03	0.02	0.00	0.0
RUTH	1746.2	-5.49	0.00	-5.48	1777.4	-0.12	0.00	0.04	0.0
CORA	3803.7	-1.63	0.00	-1.63	3812.3	-0.04	0.00	0.69	0.0
ELIZ	2676.5	-0.30	0.89		2674.7	-0.01	0.66	0.15	0.7
COCH	1693.1	-0.08	0.94		1695.2	0.01	0.40	0.00	0.2
COEN	2780.9	-0.16	0.93		2781.4	0.01	0.34	0.00	0.2
SCOT	1555.6	-0.66	0.43		1555.0	-0.02	0.28	0.00	0.0
HELL	2814.8	-1.44	0.20		2824.2	-0.02	0.15	0.00	0.0
NIVE	3291.4	-0.70	0.69		3296.6	-0.01	0.49	0.00	0.6
MURR	-404.6	-2.02	0.00	-2.02	-391.7	-0.05	0.00	0.00	0.0
SOUT	1460.1	-1.25	0.03	-1.25	1468.3	-0.03	0.07	0.00	0.0
YARR	-514.3	-1.10	0.02	-1.10	-511.7	-0.01	0.37	0.00	0.1
DOMB	2131.5	-1.14	0.52		2129.8	0.00	0.88	0.00	0.8

[1]Trend included in the GAMM model. [2]Backtransformed trend which is interpreted as a percentage of the flow. [3]no trend in the model, Mann-Kendall trend analysis on the residuals. The Mann-Kendall tau is similar to a correlation coefficient and varies between -1 to +1

Figure Legends

Figure 1 Location of hydrologic reference stations used in the study (http://www.bom.gov.au/water/hrs/)

Figure 2 Flowchart of the different approaches and comparisons of trends and significance used in the methodology

Figure 3 Annual rainfall runoff relationships for all the catchments based on the available 41 years of stream-flow data and associated gridded rainfall data.

Figure 4 Rainfall elasticities calculated from the original streamflow and gridded rainfall data using a non-parametric method (Chiew, 2006) (red triangles). Also plotted are the rainfall elasticities calculated from the streamflow simulations of the two different rainfall runoff models (SimHyd and GR4J) using the gridded rainfall data. As the model calibrations were replicated 10 times to estimate parameter equifinality, the elasticities from the rainfall-runoff models are actually distributions. This equifinality is in most cases very small and the boxplots show up as single horizontal gray lines. Outliers from the boxplots are shown as light gray dots, median values of the boxplots are shown as blue points.

Figure 5 Boxplots and median values (black dots) for goodness of fit of the GR4J and SimHyd calibrations on the 1970 - 2010 daily data using the gridded rainfall data. The goodness of fit is indicated by the Nash Sutcliffe Efficiency, which is equal to 1 for a perfect fit, and 0 for a fit equal to the mean daily streamflow. Because most of the catchments, and particularly the GR4J model, had very small equifinality (variability between repeated calibrations), the boxplots are only visible as horizontal lines.

Figure 6 Mann Kendall trends (tau) for deseasonalised streamflow for the different catchments, against the distribution of slope results of a bootstrap analysis. Black points that are located outside the bootstrap distribution can be confidently seen as highly significant and not related to possible trends in random data.

Figure 7 Mann-Kendall tau values (Dot points) and tau results of the bootstrapped residual trend analysis (distribution histograms) for the residuals of the GAMM analysis.

Figure 8 Analysis of residual trends between modelled and observed data for the models based on the 1970 – 2010 data. Red distribution boxes are significant (p value < 0.05), while blue boxes are not significant. The actual p-values can be found in the supplementary material (document 6A). Boxplots represent the 10 replications of the calibration. As in Figure 5, the variation between the calibration replications is very small, especially for GR4J, in some of the catchments resulting in the boxplots plotting as straight lines.

<figure>
Map of Australia with catchment locations labelled: ELIZ, COEN, COCH, YARR, DOMB, SCOT, COTT, CORA, RUTH, MURR, SOUT, HELL, NIVE
</figure>