Advances in Water Resources 134 (2019) 103430

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Assessing the sensitivity of hydro-climatological change detection methods to model uncertainty and bias

Ze Jiang, Ashish Sharma*, Fiona Johnson

School of Civil and Environmental Engineering, University of New South Wales, Sydney, New South Wales, Australia

ARTICLE INFO

Keywords:
Detection & attribution
Sensitivity
Model uncertainty and bias
Soil moisture

ABSTRACT

Detection of systematic changes in the climate system resulting from anthropogenic forcing is a critical area of research. Detection and attribution of hydro-climatological change has been limited by model uncertainty and bias as well as the poor spatial-temporal coverage of observational data. This study assesses a routinely adopted detection methodology and its sensitivity to model uncertainty and bias within a hydro-climatological context. Using a synthetic case study, we establish the sensitivity of detection approaches to the magnitude and consistency of trend and variance along with the length of data available. It is found that the extent of uncertainty (as measured by the variance) plays a critical role in changing the detection outcome. Another important factor is the consistency of trend between simulations and observations. A case study of soil moisture in select locations within Australia shows that averaging over multiple years (e.g., five years to a decade) improves the detection of the climate change signal as long as consistency in the trends exists. Our results also demonstrate that there are substantial differences in simulated trends across climate models. Therefore, even though ensemble averaging is effective in modulating variance, it has the risk of canceling out the signal over models with markedly different responses.

1. Introduction

Detection and Attribution (D&A) of hydro-climatological change due to human-induced climate changes is a significant area of research across the world (Kaufmann and Stern, 1997; Min et al., 2011; Mondal and Mujumdar, 2015; Gedney et al., 2006; Pall et al., 2011; Wan et al., 2015; Zhang et al., 2007). Following the definition of Intergovernmental Panel on Climate Change (IPCC), detection of change is defined as the process of demonstrating that climate or a system affected by climate has changed in some defined statistical sense without providing a reason for that change. Attribution seeks to determine whether a specified set of external forcings and/or drivers are the cause of an observed change in a specific system (Bindoff et al., 2013). However, there is still considerable uncertainty about changes in the hydro-climatological cycle under potential future warming because it is difficult to separate the effects of rising greenhouse gases from multi-year/decadal climate variability (Haerter and Berg, 2009; Stocker et al., 2013b). Additional uncertainty is introduced through the model simulations which are a vital input of D&A studies, whether these are physical, mathematical, or statistical models.

Most D&A assessments have been conducted by comparing observed trends to those simulated by a large number of General Circulation Models (GCMs) (Jones et al., 2013; Wan et al., 2015; Najafi et al., 2015; Zhang et al., 2007). There are multiple climate models available as part of the Coupled Model Intercomparison Program (CMIP) which contain multiple types of simulations. For the purposes of D&A, three groups of simulations are relevant. These are simulations driven with all relevant anthropogenic and natural forcings (ALL), simulations driven with natural forcings (e.g., solar and volcanic activity) only (NAT), and simulations driven with historical anthropogenic forcings (mainly rising greenhouse gases) only (ANT). Significant differences in the statistical attributes characterizing these simulations is used as the basis for most D&A approaches. The simplest of these approaches (more details to follow later) entails ascertaining the temporal trend or drift in each of these series and statistically ascertaining whether they represent different population distributions. Given the uncertainty and variability across model simulations, model ensemble means are often used. Even though the mean of multi-model simulations has good skill in estimating the climate responses to external forcings (Stocker et al., 2013a) and a reliable screening method and skill scores have been used for filtering climate models based on the quality and performance of these models (Pierce et al., 2009; Johnson et al., 2011),the uncertainty and bias in model-simulated trends remains high (Zhang et al., 2007). As a consequence of these discrepancies in response patterns, the use of trend or regression-based methods may lead to physically inconsistent D&A results such as negative trends when they are expected to be positive as

* Corresponding author.
E-mail address: a.sharma@unsw.edu.au (A. Sharma).

https://doi.org/10.1016/j.advwatres.2019.103430
Received 11 February 2019; Received in revised form 27 September 2019; Accepted 30 September 2019
Available online 2 October 2019

per the laws of physics (such as a negative regression slope for global surface temperature in response to rising greenhouse gases over time (Ribes and Terray, 2013)).

The simplest way of fitting the model-simulated responses to observations is to assume that the responses to different forcings are linearly additive, so the response to any one forcing can be scaled up or down without affecting any of the others. Additionally, climate variability is assumed independent of the response to external forcing (Bindoff et al., 2013). Under these conditions, attribution can be expressed as a variant of linear regression. Meehl et al. (2003) and Gillett et al. (2004) have tested the linear additivity assumption and suggested that it might hold for large-scale temperature changes but may not apply to hydro-climatological variables such as precipitation (Hegerl et al., 2007; Hegerl and Zwiers, 2011; Shiogama et al., 2013; Ghosh et al., 2012), nor to regional temperature changes (Terray, 2012). However, so far non-additive approaches have not been widely used, and it takes a longer time for the science community to accept new methods especially when the current approach is simple and practical. We also believe that there is no "right" or "wrong" approach to D&A since it is all context-dependent when it comes to how people view errors of a particular type (Lloyd and Oreskes, 2018).

Here, we aim to ascertain the sensitivity of a routinely adopted D&A methodology with the underlying linear additivity assumption to model uncertainty and bias within a hydro-climatological context. With a combination of global climate models and sophisticated statistical models that allow us to simulate underlying trends in a controlled environment, we assess the reliability of the simple regression method that has been widely used in D&A. Based on the outcomes of this assessment, we proceed to demonstrate the implications of our findings in the context of soil moisture data simulated using a range of CMIP5 GCMs, using select locations in Australia where interesting changes have been identified.

The remainder of the paper is organized as follows. Section 2 reviews the regression-based D&A methods (e.g., simple least squares, weighted and generalized linear regression method) along with their drawbacks and sources of uncertainties, given the need for a good understanding of the fundamentals of D&A methodology and their possible sensitivity to model uncertainty and bias investigated in the examples that follow. In Section 3, synthetic data is introduced to evaluate the sensitivity of the simple least squares approach to various sources of uncertainty and bias. Section 4 consists of a case study focusing on soil moisture change in select locations within Australia followed by discussion in Section 5. We conclude with a summary of the main outcomes in Section 6.

2. Background

2.1. Detection and Attribution

There are three core elements in D&A studies. First, observations of climate indicators contain spatiotemporal information. Second, an estimate of external forcing of the climate system (e.g., natural and/or anthropogenic forcings) using climate models is generally taken as an ensemble mean of forced model runs. Last, an estimate of climate variability is often derived from a physical-based model, generally obtained from unforced preindustrial model runs. D&A methodologies have evolved to include simple non-optimal fingerprint techniques to more complex regression-based methods. We illustrate the philosophy behind the D&A methodology with the non-optimal fingerprint, and summarize both non-optimal and optimal fingerprints in the form of linear regression.

In D&A studies, empirical orthogonal function (EOF) analysis is often used to study possible spatial modes (i.e., patterns) of variability and how they change with time. In statistics, EOF analysis is known as Principal Component Analysis (PCA), which is also used for dimensionality reduction. The "Optimal Fingerprint" terminology was first introduced by Hassellmann (1979). Regardless of non-optimal or optimal fingerprint, the fingerprint $F(i, j)$ represents a pattern of change in space, generally a function of latitude i and longitude j, characterizing the climate system response $X(i, j, t)$ to external forcing (Marvel et al., 2019). The fingerprint is often defined as the leading EOF (i.e., the unit-norm eigenvector $\boldsymbol{v}_1$ corresponding with the highest eigenvalue λ_1) by computing the eigenvalues and eigenvectors of a spatiotemporal covariance matrix. It often results from first averaging over members of each CMIP5 historical model ensemble and then over models (Hasselmann, 1993; Marvel et al., 2019; Santer et al., 1995). The leading EOF can not only estimate a "form" of response how the variable of interest responds to an external forcing but also reduces the dimensionality of the spatiotemporal data by projecting it onto the fingerprint.

Given a dataset of observations $Y(i, j, t)$ or model simulations $X(i, j, t)$, we estimate the projection $P(t)$ which is the amplitude of the fingerprint in the dataset, by projecting it onto the fingerprint $F(i, j)$,

$$P_Y(t) = \sum_{i,j} Y(i,j,t)F(i,j)/\sqrt{\lambda_1}$$
$$P_X(t) = \sum_{i,j} X(i,j,t)F(i,j)/\sqrt{\lambda_1} \quad (1)$$

where $\sqrt{\lambda_1}$ is used for normalization by the fingerprint. Physically, the projection $P(t)$ as a function of time indicates the covariance between the fingerprint $F(i, j)$ and the observational or model data in space. The projection time series $P(t)$ will, hence, show an upward trend if the fingerprint is increasingly present in the data (Marvel et al., 2019).

In the end, we compare the amplitude of fingerprint in observations with both amplitude of signal in different forced model runs (e.g., ALL and NAT forcings) by regression (Eq. (2)) and demonstrate that the response patterns of alternative external forcings (e.g., NAT forcings) are unlikely to explain the observed change.

$$P_Y(t) = \beta_i P_{X_i}(t) + u \quad (2)$$

where β_i is the scaling factors that adjusts the amplitudes of those patterns, and u is noise.

With the non-optimal detection approach summarized above, it has been recognised that it can be cast as a regression-based problem with respect to generalised multivariate regression (Hasselmann, 1997; Allen and Tett, 1999; Allen and Stott, 2003). The regression model has the form

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (3)$$

where vector $\boldsymbol{y}$ is a filtered version of the observations, matrix $\boldsymbol{X}$ contains the estimated response patterns to the external forcings (signals) that are under investigation, $\boldsymbol{\beta}$ is a vector of scaling factors that adjusts the amplitudes of those patterns, and $\boldsymbol{\epsilon}$ is noise that represents climate variability (Hegerl et al., 2010). When the response $\boldsymbol{X}$ is noise-free, vector $\boldsymbol{\beta}$ is estimated with

$$\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^T\boldsymbol{C}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{C}^{-1}\boldsymbol{y} \quad (4)$$

where $\boldsymbol{C}$ is the covariance matrix of the noise (Hassellmann, 1979; Allen and Tett, 1999; Mitchell, 2001). The most commonly used assumption about vector $\boldsymbol{\epsilon}$ is that it follows a Gaussian distribution and the associated covariance matrix $\boldsymbol{C}$ becomes an identity matrix. Thus, scaling factors $\boldsymbol{\beta}$ is simplified as

$$\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{y}, \quad (5)$$

which is equivalent to the non-optimal method. In essence, the covariance matrix $\boldsymbol{C}$ gives a somewhat greater weight to information in the low variance parts of the observations. Generally, in terms of how we generalize the covariance matrix $\boldsymbol{C}$ (i.e., measure the distance between the data points and the regression line), the detection technique can be generally divided into three methods as outlined in Table 1.

It should be noted that the Simple least squares approach (Table 1) is adopted generally when alternate models are expected to represent independent realizations of the process, and are expected to have equal residual error variances. This, however, is often too constraining an assumption especially when the process exhibits a chaotic pattern, leading

Table 1
Variations of D&A regression-based methodologies.

Covariance matrix	Method
$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$	Simple least squares (non-optimal)
$\mathbf{C} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_3^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$	Weighted least squares (partially optimal)
$\mathbf{C} = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2}^2 & \sigma_{1,3}^2 & \cdots & \sigma_{1,n}^2 \\ \sigma_{2,1}^2 & \sigma_2^2 & \sigma_{2,3}^2 & \cdots & \sigma_{2,n}^2 \\ \sigma_{3,1}^2 & \sigma_{3,2}^2 & \sigma_3^2 & \cdots & \sigma_{3,n}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n,1}^2 & \sigma_{n,2}^2 & \sigma_{n,3}^2 & \cdots & \sigma_n^2 \end{bmatrix}$	Generalized linear regression (fully optimal)

to the Weighted and the Generalized approaches being adopted. Under heteroscedasticity (the residual error variances of the observations are unequal) or autocorrelation, Simple least squares is no longer the optimal estimate whereas the other two estimators can be unbiased and more efficient (Allen and Stott, 2003; Ribes and Terray, 2013).

Fitting the regression model requires an estimate of the covariance matrix $\boldsymbol{C}$ (i.e., the climate variability), which is usually obtained from an additional, independent sample of simulated unforced variation (e.g., long control simulations) because the instrumental record is too short to provide a reliable estimate and may be affected by external forcings. However, GCMs may not simulate climate variability accurately (Johnson et al., 2011), casting doubt over the estimates of $\boldsymbol{C}$. Even with good estimates of $\boldsymbol{C}$ from GCMs control simulations, studies (Wan et al., 2015; Santer et al., 1995) have showed that optimal approaches have no clear advantages over non-optimal methods. Hence, in this study, we assess the sensitivity of D&A approaches to model uncertainty and bias using the Simple least squares approach.

2.2. Systematic biases in climate model simulations, and their implications for D&A

A failure to detect a particular response includes the possibility that the responses are collinear or weak relative to climate variability (Ribes et al., 2017), or that the metric used to measure the change is insensitive to the expected change (Pachauri and Reisinger, 2007). Aside from being able to distinguish climate variability in observations (Hegerl et al., 2007; Ropelewski and Halpert, 1987; Mantua et al., 1997; Newman et al., 2003; Verdon and Franks, 2006; Zhang et al., 2007) and the selection of proper metrics, the key to assessing the reliability of D&A methods is to understand the uncertainty and bias in climate models. Climate models have errors that consist of unsystematic and systematic components (Teutschbein and Seibert, 2013). Unsystematic model errors occur due to the chaotic nature of the climate system and the unconstrained nature of the GCMs simulations (Christensen et al., 2001; Eden et al., 2012). For instance, global scale feedbacks like water vapor and cloud feedback and possible unknown feedbacks within the climate system could contribute substantially to uncertainties on the response patterns and magnitudes (Ribes et al., 2017; Mehrotra and Sharma, 2006). Due to uncertainties in the forcings themselves, the estimated responses to these external forcings fluctuate significantly. The greenhouse gases forcing has substantial uncertainties if effective radiative forcings are considered rather than radiative forcings (Myhre et al., 2013). Various studies have quantified the extent of these errors, and attributed them to model structure, emission scenario and initialization (Woldemeskel et al., 2016). The extent of these errors varies depending on the variable of interest, with hydro-climatological variables (such as precipitation and soil moisture) expected to exhibit greater error due to model structure, as compared to other factors. The systematic error component is commonly termed as model bias (Teutschbein and Seibert, 2013) and is a result of a range of various factors. They include the coarse scale of the GCMs, incomplete model structures, feedback loops including the albedo and land-atmosphere interactions and in the case of precipitation the parameterisation of clouds and convection (Allen et al., 2002; Haerter et al., 2011; Masson and Knutti, 2011; Randall et al., 2007; Sun et al., 2006; Maurer et al., 2013; Mehrotra et al., 2004).

A number of bias correction methods have been developed to address these biases prior to D&A exercises, including scaling or quantile matching, empirical-statistical correction (Gobiet et al., 2015), nesting logic-based approaches (Johnson and Sharma, 2012), and multivariate bias correction models (Mehrotra and Sharma, 2015). Bias correction is popular because it can be applied directly and easily to climate model outputs, and is able to correct the GCMs simulations for the parameters of interest (Johnson and Sharma, 2012). On the other hand, there is a suspicion that bias correction lacks a sound physical basis (Ehret et al., 2012) since it does not necessarily preserve the dynamic relationships between different variables (Haerter et al., 2011). The fundamental assumption of these methods is that if the biases can be removed from the model simulations, then the corrected GCMs outputs will properly represent the expected responses in the climate system (Nahar et al., 2017). However it has been found that bias correction can sometimes alter the climate change signal, which in some cases is considered undesirable (Hagemann et al., 2011) and in others is considered beneficial for example by improving model consensus on the direction of changes (Gobiet et al., 2015; Johnson and Sharma, 2015). Therefore, this is a strong assumption and we should be very cautious when making an inference.

In a nutshell, where models are used in D&A, the assessment of model uncertainty and bias should be considered. Bias correction is a powerful tool to remove model biases prior to D&A analysis, and it should be applied with great care.

3. Synthetic data

The aim of the first part of the study is to formulate realistic time series representing observed and model-simulated sequences such that the impact of known uncertainties and biases can be assessed on the D&A outcomes. The synthetic series are formulated to broadly represent the types of biases seen in hydro-climatological data. The synthetic case study is followed by an application to gridded soil moisture data in Australia, described in the next section.

The base model for synthetic time series generation is the random walk with drift model (Shumway and Stoffer, 2011) given by

$$x_t = \delta + x_{t-1} + w_t \tag{6}$$

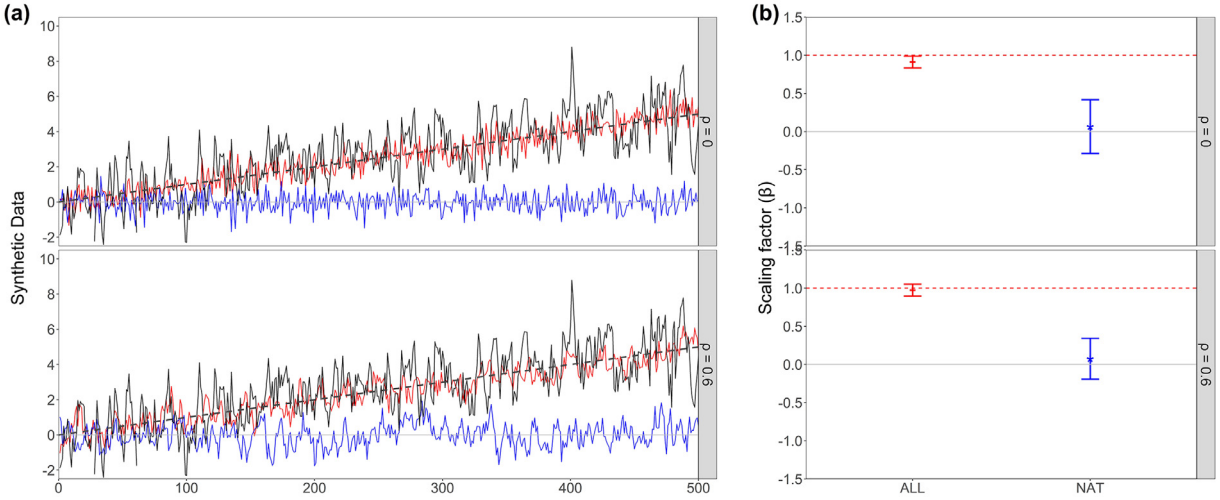

Fig. 1. A synthetic example: (a) Generated observations (black line) y and the true trend (dashed black line) in observations, simulated response to ALL forcings (red line) x_1, and simulated response to NAT forcings (blue line) x_2; (b) The comparison between observations and ALL forcings case by regressing y on x_1 (red short line with error bar, 95% confidence interval); the comparison between observations and NAT forcings case by regressing y on x_2 (blue star with error bar, 95% confidence interval). In each figure, the two subplots represent the cases with ($\rho = 0.6$) or without ($\rho = 0$) autocorrelation in the white noise. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

for $t = 1, 2, \ldots,$ n with initial condition $x_0 = 0$, and where w_t is Gaussian white noise, $w_t \sim N(0, \sigma_w^2)$. The constant δ is called the drift, and when $\delta = 0$, the Eq. (6) is called simply a random walk. The term random walk comes from the fact that when $\delta = 0$, the value of the time series at time t is the value of the series at time t-1 plus a completely random movement determined by w_t. Note that we may rewrite the Eq. (6) as a cumulative sum of white noise variates. That is,

$$x_t = \delta t + \sum_{j=1}^{t} w_t. \tag{7}$$

From the rewritten Eq. (7), the drift δ in the model can be seen as the slope (or trend) of the time series. As previously described, there are three core elements in D&A studies, consisting of observations $\boldsymbol{y}$, estimated response pattern to external forcing $\boldsymbol{x}$, and estimated climate variability ϵ. Also, to assess the effects of different degrees of persistence in the climate system, two different autocorrelations ($\rho = 0$ and $\rho = 0.6$) was introduced in Gaussian white noise using a first-order autocorrelation (AR(1)) model to replace the cumulative sum of white noise variates in the Eq. (7). To summarize, in this synthetic analysis, they are generated by

$$\begin{aligned} y_t &= \delta_{obs} t + \varepsilon_t \qquad w_t \sim N(0, \sigma_{obs}^2) \\ x_t &= \delta_{sim} t + \varepsilon_t \qquad w_t \sim N(0, \sigma_{sim}^2) \\ \varepsilon_t &= \rho \varepsilon_{t-1} + w_t \end{aligned} \quad . \tag{8}$$

To simplify the synthetic case, we assume the spatial structure of the signals does not change over time and only one grid cell is considered, thus the generated data is equivalent to projected and normalized data. Fig. 1 presents an example of the synthetic dataset to illustrate the implementation in this statistical model, with $\delta_{obs} = \delta_0 = 0.01$, $\sigma_{obs} = \sigma_0 = 1$ and $n = 500$ for both simulated "observations" and "response pattern to external forcing". We compare observations $\boldsymbol{y}$ with both response patterns to ALL ($\delta_{sim} = 0.01$ and $\sigma_{sim} = 0.5$) and NAT ($\delta_{sim} = 0$ and $\sigma_{sim} = 0.5$) forcings $\boldsymbol{x}$ by regression. In this synthetic dataset, it is noted that the observed changes are unlikely to be explained by NAT, given corresponding scaling factor is round zero. However, the estimated response pattern to ALL forcings captures the change in observations well, given associated scaling factor is close to one.

To assess the sensitivity of D&A methods to model uncertainty and bias, we divide the experiment into four groups, varying drift δ, drift ratio $\frac{\delta_{sim}}{\delta_{obs}}$, noise ratio $\frac{\sigma_{sim}}{\sigma_{obs}}$ and sample size n respectively. For both simulated "observed" and generated "ALL forcings model" time series, all three statistics (δ, σ_w and n) were varied. The Group-I series were generated varying drift δ ($\delta_{sim} = \delta_{obs}$) ranging from $0.1\delta_0$ to $1.5\delta_0$, the Group-II were generated with inconsistent drift δ between ALL and observations (fixed drift of observations as $\delta_{obs} = \delta_0$) as indicated by drift ratio $\frac{\delta_{sim}}{\delta_{obs}}$ varying from 0.1 to 1.5, the Group-III series were generated by altering the white noise standard deviation ratio $\frac{\sigma_{sim}}{\sigma_{obs}}$ ranging from 0.01 to 1.5 (fixed standard deviation of observations as $\sigma_{obs} = \sigma_0$), and the Group-IV series were generated using the base model with different length n varying from 10 to 1000. Similarly, another set of four group series was generated by replacing ALL with NAT forcing scenario using simple random walk (i.e., $\delta_{sim} = 0$) model to simulate its response. Group-I and Group-II were used to investigate the influences of trend magnitude and consistency on signal detection. Group-III was used to investigate the influences of model uncertainty (as measured by variance) on signal detection. Group-IV was used to investigate the influence of series length on the power of the simple regression method. The AR(1) variation was introduced to study the sensitivity of the D&A method to serial correlation.

Since the true signal in the synthetic series are clearly known and identical in simulated ALL forcings case in Group-I, Group-III and Group-IV, the theoretically derived true scaling factor β equals one, while in Group II, the theoretically derived true scaling factor β equals the inverse of their ratio ranging from ~0.67 to 10. Since the drift in NAT forcings case is zero, the theoretical true value of scaling factor β for synthetic NAT forcings case is always zero.

As shown in Fig. 2, the estimated scaling factors β of NAT forcings case are all near the theoretical true value of zero. The simulated ALL forcings case provides guidance on the issues that may hamper D&A studies. The Group-I and Group-II results show how the magnitude and consistency of trend influence the signal detection. When the trend is strong and consistent with the observed trend, it is easier to correctly detect the change ($\beta \approx 1$). However, when the signal is weak and inconsistent with observations, the estimated scaling factors tend to become imprecise. The Group-III results show the presence of noise (e.g., climate variability in observations and uncertainty in model simulations) influences the signal detection, and it suggests that reducing the variance in model simulations will lead to more precise estimates of the scaling

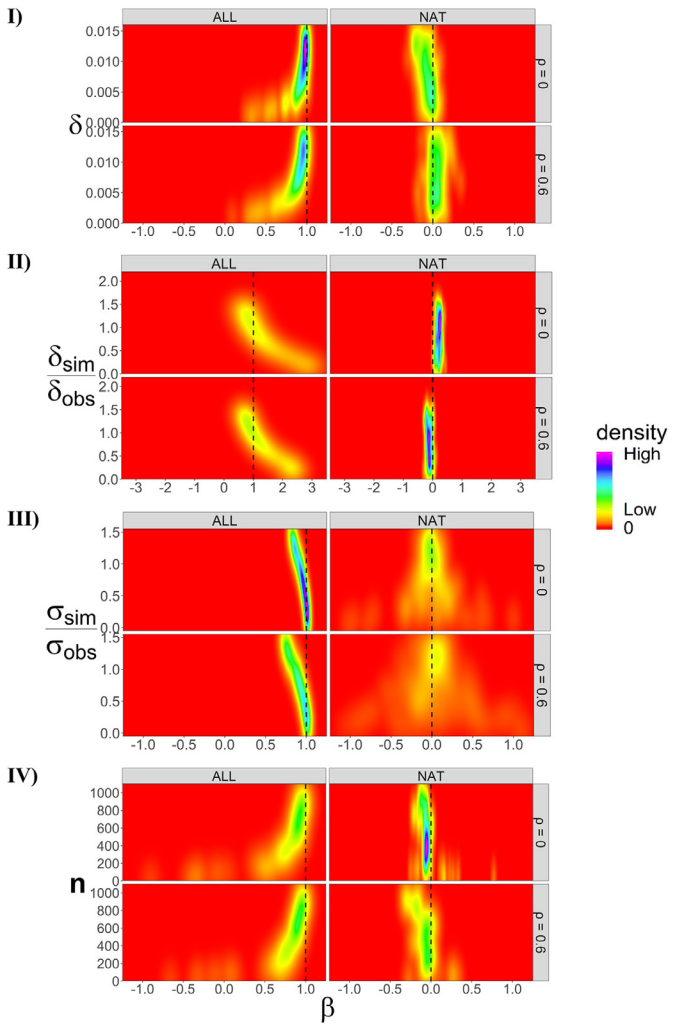

Fig. 2. Sensitivity analysis of the D&A method corresponding with four groups of synthetic series with the black dashed line indicating the theoretical true values of scaling factors β. (Group I): Varying trend in ALL forcings simulations consistent with observations; (Group II): Varying trend in ALL forcings simulations inconsistent with observations; (Group III): Decreasing variance in simulations; (Group IV): The impact of data length. In each figure, the columns give the synthetic ALL forcings case and NAT forcings case; the rows represent the cases with or without autocorrelation in the white noise. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

factors. Group-IV shows that signals are easier to detect when the time series are long.

What is most notable about these results are the cases where β deviates from 1. This appears most clearly when the drift (δ) is low (less than 0.007), where the drift ratio is not equal to one (most clearly when this ratio is less than unity), and when the ratio of the error standard deviations is greater than 0.25. This implies, for detection to be successful, one needs statistically significant trends that are consistent with observed (resulting in a drift ratio of 1), and, most importantly, model residuals that exhibit 25% of the variability exhibited through observations. Scenarios where these conditions can be met within a hydro-climatological setting may be difficult to identify. This issue is investigated further using soil moisture simulations in the real case study presented next.

4. Soil moisture change detection case study

Understanding the behavior of soil moisture is essential due to its key role in the hydrological cycle (Owe et al., 2008). Soil moisture influences rainfall-runoff processes, infiltration, groundwater recharge, and constrains evapotranspiration and photosynthesis (Holgate et al., 2016). Thus, it is directly involved in water and energy exchanges between the land, vegetation and the atmosphere (Taylor et al., 2012). As an antecedent condition of various environmental forecasts, it has a range of applications including assimilation into land surface models for numerical weather forecasting (Dharssi et al., 2011), national water accounting (Viney et al., 2014), agricultural planning and bushfire control (van Dijk et al., 2013; van Dijk et al., 2015) as well as flood prediction (Wanders et al., 2014).

The biggest challenge of soil moisture study is the sparsity of in situ observations. To overcome this difficulty, soil moisture information can also be obtained from satellite remotely sensed estimates and hydrological/land surface model predictions. In situ measurements are at the point scale and cover limited areas with hourly or sub-daily time steps. Remotely sensed soil moisture estimates have a larger spatial scale (tens or hundreds of square kilometers) and are available from a growing number of satellites on daily or longer basis. However, there are concerns with the representativeness of these measurements and there is no long time series available. The accuracy of modeled soil moisture varies significantly depending on the accuracy and spatial coverage of the input precipitation and soil hydraulic properties data, in addition to the model structure and purpose. Previous studies have compared these data within Australia and showed that the Australian Water Resource Assessment landscape model (AWRA-L), a grid-based distributed biophysical model of the water balance between the atmosphere, soil, groundwater and surface water stores (Viney et al., 2015), yields a strong agreement with in situ observations (Holgate et al., 2016). Hence, we adopted AWRA-L soil moisture as a surrogate of observed soil moisture for this D&A case study. AWRA-L estimates a daily running water balance on a $0.05^\circ \times 0.05^\circ$ grid across Australia from 1911 to present.

The model-simulated monthly soil moisture is taken from CMIP5 archive to estimate soil moisture responses to ALL and NAT forcings. In total, four GCMs are selected to carry out the D&A analysis, as summarized in Table 2. These GCMs were chosen due to the availability of soil moisture simulations and their larger number of initial condition ensemble members. As is common practice in this field, the ensemble mean of multiple model realisations has been used meaning that the uncertainty resulting from ensemble members of different models has not been considered in this study. This is firstly because, as shown in Table 2 the same number of realizations is not available for all GCMs, thus it is hard to compare across GCMs. Secondly, the variation across realizations of a single GCM is generally much smaller than the variations in the entire multimodel ensemble (Johnson et al., 2011) and can thus be neglected for the purposes of this investigation. Hawkins and Sutton (2009) also suggest that the model structure plays a more important role in model uncertainty rather than the model realizations of different initial conditions. More importantly, no weighting or screening is applied across models (Bhowmik et al., 2017), which ensures that the average is not dominated by any model in CMIP5, and the result should give us a common response (Marvel et al., 2019).

Due to the differences in the grid resolution between AWRA-L and selected GCMs, all the data were re-gridded to $2.5^\circ \times 2.5^\circ$ over Australia, using weighted area interpolation for the AWAR-L simulations and bilinear interpolation for the GCMs. Anomalies of annual, non-overlapping five-year and decadal averages were computed from 1911 to 2000.

Two regions in Australia were investigated, one in Northern Australia which is expected to have increasing soil moisture trends ($\delta_{obs} = 0.02383$) due to increasing rainfall, and the other one in Southwest Australia that experienced decreased trends ($\delta_{obs} = -0.0022$) due to the strong decreases in available water in this region over the last three to four decades (CSIRO and Bureau of Meteorology, 2016).

Fig. 3 shows the time series of soil moisture from AWRA-L and four GCMs at two grid points in these two regions. The study locations were chosen to illustrate both significant trends as well as the difference in factors between the ALL and NAT forcings model simulations. The

Table 2
List of models, variables, experiments and grid resolution for this analysis.

Model	Variable name	Long name (water content at top 10 cm)	Grid size (latitude × longitude degree)	No. of model realizations		Control simulation (years)
				ALL simulation	NAT simulation	
CSIRO-Mk3.6.0	*mrsos*	Moisture in upper portion of soil column	1.865 × 1.875	10	10	500
GFDL-CM3			2×2.5	4	3	500
GISS-E2-R			2×2.5	6	5	3175
IPSL-CM5A-LR			1.89 × 3.75	6	3	1000
AWRA-L	*szsm*	Upper soil moisture	0.05 × 0.05	–	–	–

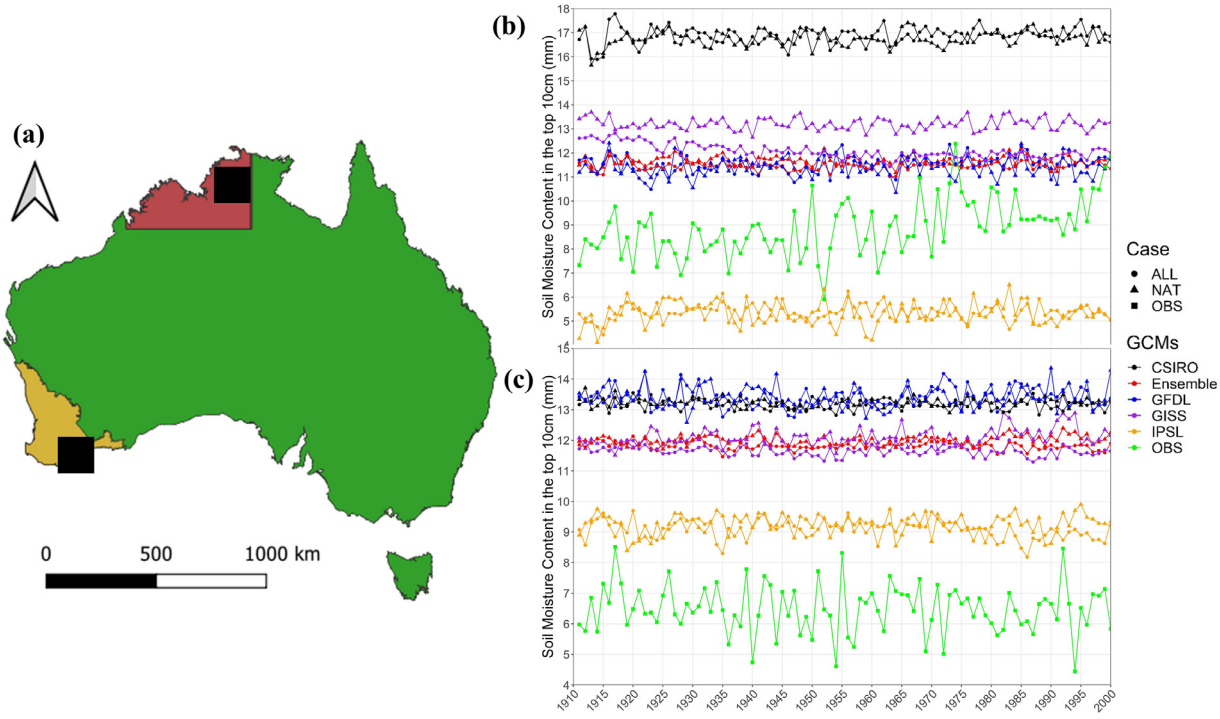

Fig. 3. Time series of the annual soil moisture from AWRA-L (green line, given as OBS and indicated by square) and four GCMs (colour lines) and their ensemble (red line) including ALL (indicated by circle) and NAT (indicated by triangle) forcings model at two grid points in two regions: (a) Two grid points (black box) in two study regions, Northern Australia (brown colour) and Southwest Australia (gold colour); (b) Time series of grid point in Northern Australia; (c) Time series of grid point in Southwest Australia. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ensemble time series is the mean of all the four GCMs simulations. The magnitude of the trend in all GCM simulations and their consistency (as measured by trend ratio) relative to observations at two grid points in these two regions are given in the table below. It is noted that there are substantial differences of simulated mean soil moisture between climate models. Differences in climate change patterns, the diversity of model soil schemes and soil and vegetation parameterizations in each model contribute to the difference of simulated soil moisture (Orlowsky and Seneviratne, 2013; Berg et al., 2017).

Detection analysis was carried out at these two grid points. Fig. 4 shows the detection result of the grid point in Southwest Australia for each GCM and their ensemble average. As per the synthetic data analysis, when the variance is reduced by temporal averaging, the trend is generally easier to detect since the signal-to-noise ratio is enhanced. The GISS model performs well even at the annual time scale and further temporal averaging does not provide great improvements. However, for CSIRO and IPSL the signal is better detected (scaling factors larger than zero) by taking longer averaging periods. The climate model GFDL perform poorly since it has an inconsistent trend with observed soil moisture, seen in Table 3. Importantly, their ensemble seems to be the best example of illustrating the implication of variance modulation. The detection outcome improves gradually over longer averaging periods. It is known that taking ensemble mean is a way to further reduce the variance of all the model simulations.

Fig. 5 shows the results of the analysis for Northern Australia, and it emphasizes how the consistency of trend influences the D&A outcome. As per the synthetic data analysis, given that the trend is weak and inconsistent with observations, the estimated scaling factors tend to become imprecise. According to Table 3, compared with observations, the GISS has the opposite trend, and their ensemble average has the weakest trend. Importantly, compared to the southwest region, the trends of all GCMs in the northern region are significantly inconsistent with observations. Thus, the estimated scaling factors in the northern region have larger confidence intervals (note that in Fig. 5 the *y*-axis limits are wider than the *y*-axis limits in Fig. 4). Model GISS has negative scaling factors in ALL forcings case because of its opposite trend relative to the observed soil moisture trend. Their ensemble has the largest confidence interval due to its lowest trend consistency. As variance modulation has been applied in both northern and southwest region, the analysis suggests that variance is not the only factor that affects D&A performance but also the consistency of trend. With inconsistent trends between simulations and observations, modulating variance may lead to high uncertainty in D&A

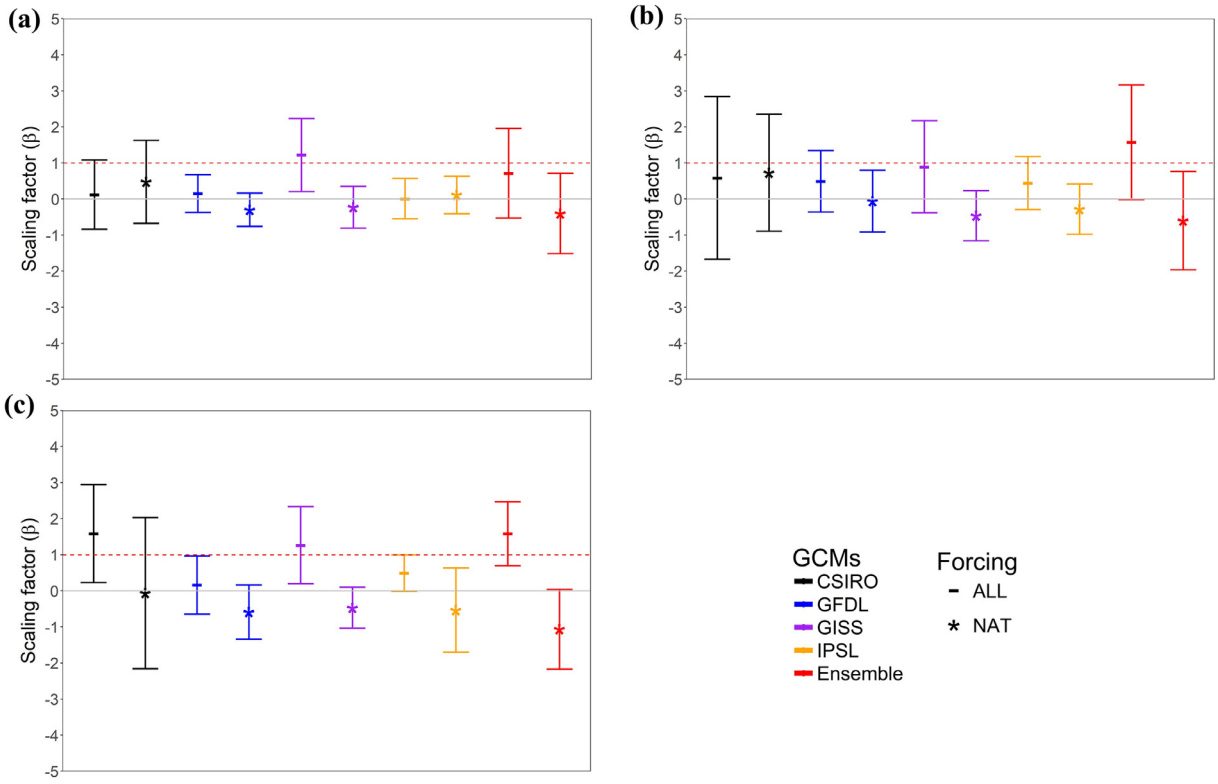

Fig. 4. Detection analysis in the grid point in Southwest Australia, and the error bars are the 95% confidence interval: (a) D&A analysis with the annual data; (b) D&A analysis with the five-year average data; (c) D&A analysis with the decadal average data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
The magnitude and consistency of trend in all GCMs relative to observations.

	Region	Northern Australia		Southwest Australia	
	GCMs	Trend (δ_{sim})	Trend Ratio ($\delta_{obs}/\delta_{sim}$)	Trend (δ_{sim})	Trend Ratio ($\delta_{obs}/\delta_{sim}$)
ALL	CSIRO	0.00251	9.48	−0.0010	2.27
	GFDL	0.00399	5.97	0.0007	−3.11
	GISS	−0.00767	−3.11	−0.0017	1.32
	IPSL	0.00125	19.13	−0.0031	0.70
	Ensemble	0.00002	1118	−0.0013	1.73
NAT	CSIRO	0.00224	10.65	0.0002	−12.42
	GFDL	−0.00005	−473.58	0.0001	−38.12
	GISS	−0.00018	−130.48	0.0039	−0.56
	IPSL	0.00005	475.82	0.0013	−1.71
	Ensemble	0.00051	46.39	0.0014	−1.62

analysis. Taking ensemble averaging is a way of reducing variance over multiple simulations, but it also has the risk of canceling out the signal over models with markedly different responses as given here. Thus, the D&A outcome totally changes.

It should also be pointed out that the results above are largely consistent with the synthetic case study presented in the previous section. For instance, the estimated mean scaling factor is close to one in ALL forcings case while for the case of NAT forcings the estimated scaling factor is distributed around zero. The magnitude of the trends is smaller than 0.007 and drift ratio is considerably different from 1 in ALL forcings case. In terms of NAT forcings case, when the standard deviation or trend of simulated data is much smaller than that of observations, its scaling factor might be far away from zero (including both negative and positive), which is shown in the synthetic case of Group II and III (note that in Group II the *x*-axis limits are wider than the *x*-axis limits of Group III). These observations are loosely consistent with what could be observed with other hydro-climatological variables such as precipitation, although trends for temperature derived variables (such as evaporation) are likely to be more significant.

5. Discussion

5.1. Use of variance stabilization strategies in previous D&A applications

In past hydro-climatological D&A applications, a majority of studies have applied averaging over time or ensemble means of multiple simulations for reasons that are apparent from the synthetic test results presented before. Jones et al. (2013) and Najafi et al. (2015) both attributed the temperature variations to human-induced greenhouse gases through averaging using the decadal mean and five-year mean, respectively. Zhang et al. (2007) and Wan et al. (2015) investigated the anthropogenic impact on northern high-latitude precipitation and used five-year mean precipitation anomalies. Willett et al. (2007) identified a significant global-scale increase in surface specific humidity that is at-

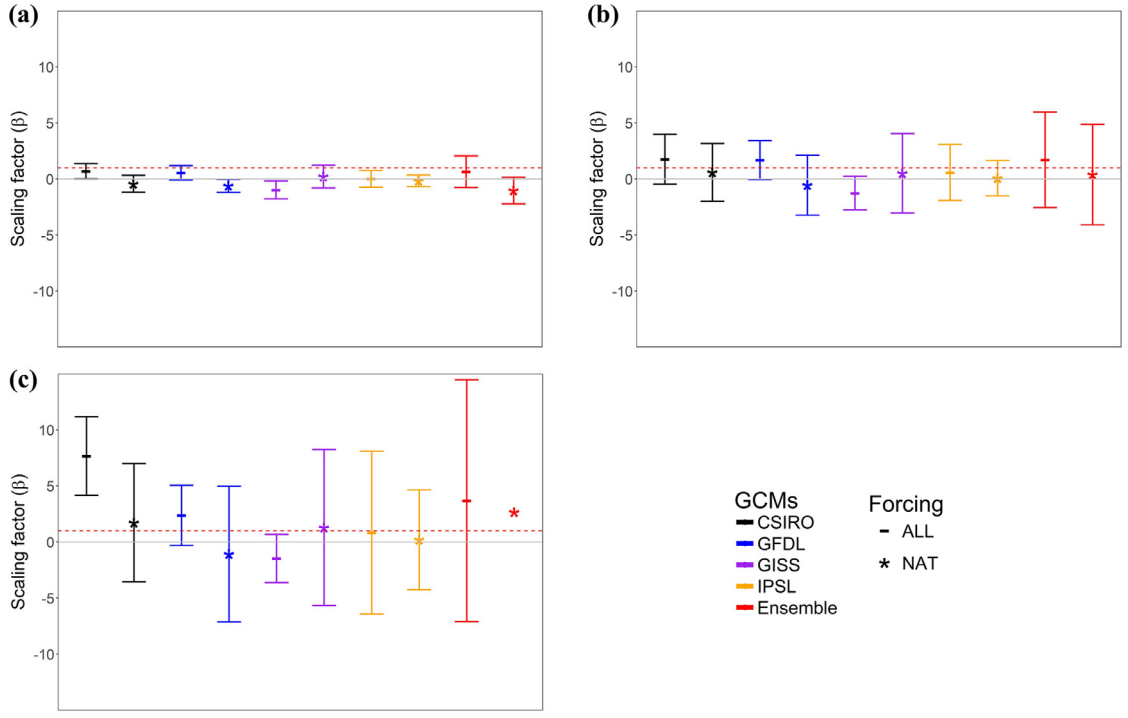

Fig. 5. Detection analysis in the grid point in Northern Australia; the error bars are the 95% confidence interval, and if there are no error bars shown, it means the confidence intervals are beyond the maximum or minimum value of *y*-axis: (a) D&A analysis with the annual data; (b) D&A analysis with the five-year average data; (c) D&A analysis with the decadal average data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tributable mainly to human influence using non-overlapping three-year means. All the studies used the ensemble mean of the GCMs available for the variables of interest. It is clear that by averaging over time and/or multiple ensemble members we can filter out the seasonal, interannual variations and/or multiple year phenomena and improve the potential climate change signal-to-noise ratio. Additionally, the argument that different climate model simulations offer time synchronicity (even though they share only a common initialization and anthropogenic forcings) is made more tenable when the time averaging extends to decadal time scales. However, given the influences of variance reduction as demonstrated in our synthetic experiments, one questions whether the conclusions drawn are an artifact of this reduction or a genuine signal present across models. One needs to be especially careful of the fact that variance modulation can lead to low precision when the simulated trend is significantly different from the observed trend, something that occurs often even across multiple datasets of the same hydro-climatological variable (Anabalón and Sharma, 2017).

5.2. Implications of variance stabilization for other hydro-climatological variables

The chaotic nature of many hydro-climatological variables will have a significant impact on the D&A studies and the smoothing and filtering method used to pre-process the data also plays an important role. For instance, Zhang et al. (2007) and Wan et al. (2015) not only used five-year mean precipitation anomalies at temporal scale but also averaged rainfall over a much larger spatial domain compared with original GCMs grid resolutions. They found that with different spatial configurations averaging over one, two, three and six sub-regions, the non-optimal method is able to provide consistent detection outcomes. However, when averaging over short time periods (e.g., using 3-year or 1-year and half-year means), it fails to detect NAT. Longer averaging periods including 8-year and 10-year means precipitation anomalies led to similar results to those of 5-year mean. These results suggest that the detection method is not sensitive to the spatial resolution, however identifying an optimal temporal moving window is critical and generally longer moving windows of precipitation tend to give better results. As discussed earlier in Section 2, GCMs share a number of sources of uncertainty and bias such as numerical approximations, a limited description of geophysical fields, and parameterization of physical processes (Saini et al., 2015). Rainfall process is a complex physical process at local scale, thus it is parameterized in GCMs to replace processes that are too small-scale or complex to be physically represented in the model. Soil moisture, particularly surface zone soil moisture, largely depends on temperature, precipitation as well as soil hydraulic properties. Hence, it also has high temporal and spatial variability, and the effect of averaging over a larger domain and temporal period is unknown since the land cover and soil types can differ significantly.

5.3. Possible ways forward

To detect subtle changes and account for the various sources of uncertainty and bias in climate models, D&A methodologies have evolved gradually. So far, the priority of the D&A methods has been to assess change over a short period of time (e.g., 30 to 50 years) mainly due to the limited observational records, especially in hydrology. However, this focus on temporal trends may be the issue of the current D&A methodology. Some studies approach this issue from a different angle. For instance, direct atmospheric carbon dioxide (CO_2) effects on plant transpiration were identified by Field et al. (1995), and Gedney et al. (2006) also pointed out that the increase of continental runoff trend through the twentieth century is consistent with a suppression of plant transpiration due to CO_2-induced stomatal closure. Therefore, a possible way forward in D&A applications is to confirm the presence or absence of temporal change with corresponding change with reference to an alternate variable such as CO_2 or equivalent. While

the choice of this variable may depend on the process being studied, such a framework could hold greater promise than the alternatives used currently. The framework could also include the option of global sensitivity analysis (GSA), attributing the uncertainty of the D&A outcome to various sources of model uncertainty and bias as considered here (Sheikholeslami et al., 2019). This could assist in ensuring that the variable selected presents a stable detection outcome across multiple climate models.

6. Conclusions

This study assesses the sensitivity of a routinely adopted detection methodology to model uncertainty and bias within a hydro-climatological context. Several inferences can be drawn from the results presented in our study. The analysis of synthetic series indicates that the extent of uncertainty (as measured by the variance) plays a critical role in changing the detection outcome in any hydro-climatological experiment conducted. An additional factor that contributes equally to the outcome is the consistency in the sign and magnitude of the trend between simulations and observations. As hydro-climatological simulations drawn from climate models exhibit considerable uncertainty due to the uncertainty associated with precipitation simulations, our assessment used soil moisture simulations from multiple models with locations identified to help illustrate the main findings noted above. For the case study of soil moisture in Australia, taking multiple years (e.g., five years) or decadal averaging is found to improve the detection of the climate change signal given that there is a high degree of consistency in the trend between simulations and observations. On the contrary, with inconsistent trends between simulations and observations, modulating variance may lead to high uncertainty in D&A analysis. Hence, even though averaging over larger temporal moving windows of climate variables of interest tends to give better detection outcome, it could fail at some cases as given here. Also, our results demonstrate that there are significant discrepancies among various climate models with markedly different responses. Thus, taking ensemble averaging, often adopted to ensure the response is not dominated by any model in CMIP5, may result in canceling out the climate change signal.

In closing, we argue that blind use of existing D&A approaches with applications to hydro-climatological problems can lead to erroneous conclusions. Given the influences of variance reduction as demonstrated in our synthetic experiments, whether the conclusions drawn are an artifact of this reduction or a genuine signal present across models is in doubt. It is essential to be especially careful of the fact that variance modulation can lead to low precision when the simulated trend is significantly different from the observed trend. Before any such application, it is recommended users ascertain the magnitude of the simulated trend across modeling scenarios, allowing assessments of the consistency in the sign of individual trends, as well as the overall variability exhibited. Following this, a variance stabilization strategy should be employed to create greater consistency in trends, and a variance that is suitably reduced taking into account the sample data length in use. While the above approach will result in clearer D&A assessment outcomes for any variable of interest, we also feel the assessment should include additional variables and spatial locations where one would expect a consistency in the outcome as per physical reasoning and observational data. The above steps may reduce the mis-detection of change in hydro-climatological simulations, increasing confidence in mitigating actions that may be under consideration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by the Australian Research Council linkage grant (LP150100548) and Crown lands & Water Division, Department of Industry, NSW, Australia. AWRA-L soil moisture data can be obtained from Bureau of Meteorology (BOM), Australia (http://www.bom.gov.au/water/landscape/). The authors are grateful to the editors and anonymous reviewers for their constructive comments that have significantly improved the quality of this work.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.advwatres.2019.103430.

References

Allen, M., Stott, P., 2003. Estimating signal amplitudes in optimal fingerprinting, Part I: theory. Clim. Dyn. 21, 477–491.
Allen, M.R., Kettleborough, J., Stainforth, D., 2002. Model error in weather and climate forecasting. ECMWF Predictability of Weather and Climate Seminar. European Centre for Medium Range Weather Forecasts http://www.ecmwf.int/publications/library/do/references/list/209.
Allen, M.R., Tett, S.F., 1999. Checking for model consistency in optimal fingerprinting. Clim. Dyn. 15, 419–434.
Anabalón, A., Sharma, A., 2017. On the divergence of potential and actual evapotranspiration trends: an assessment across alternate global datasets. Earth's Fut. 5, 905–917.
Berg, A., Sheffield, J., Milly, P.C., 2017. Divergent surface and total soil moisture projections under global warming. Geophys. Res. Lett. 44, 236–244.
Bhowmik, R.D., Sharma, A., Sankarasubramanian, A., 2017. Reducing model structural uncertainty in climate model projections—a rank-based model combination approach. J. Clim. 30, 10139–10154.
Bindoff, N.L., et al., 2013. Detection and attribution of climate change: from global to regional.
Christensen, O., Gaertner, M., Prego, J., Polcher, J., 2001. Internal variability of regional climate models. Clim. Dyn. 17, 875–887.
CSIRO & Bureau of Meteorology 2016. Technical Report (online http://www.csiro.au/resources/State-of-the-Climate.html). Commonwealth Scientific and Industrial Research Organization/Australian Bureau of Meteorology, Melbourne.
Dharssi, I., Bovis, K., Macpherson, B., Jones, C., 2011. Operational assimilation of ASCAT surface soil wetness at the MET office. Hydrol. Earth Syst. Sci. 15, 2729–2746.
Eden, J.M., Widmann, M., Grawe, D., Rast, S., 2012. Skill, correction, and downscaling of GCM-Simulated precipitation. J. Clim. 25, 3970–3984.
Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., Liebert, J., 2012. HESS opinions "Should we apply bias correction to global and regional climate model data?". Hydrol. Earth Syst. Sci. 16, 3391–3404.
Field, C.B., Jackson, R.B., Mooney, H.A.J.P, 1995. Stomatal responses to increased CO_2: implications from the plant to the global scale. lant Cell Environ. 18, 1214–1225.
Gedney, N., Cox, P.M., Betts, R.A., Boucher, O., Huntingford, C., Stott, P.A., 2006. Detection of a direct carbon dioxide effect in continental river runoff records. Nature 439, 835–838.
Ghosh, S., Das, D., Kao, S.-C., Ganguly, A.R., 2012. Lack of uniform trends but increasing spatial variability in observed Indian rainfall extremes. Nat. Clim. Change 2, 86.
Gillett, N., Wehner, M., Tett, S., Weaver, A., 2004. Testing the linearity of the response to combined greenhouse gas and sulfate aerosol forcing. Geophys. Res. Lett. 31.
Gobiet, A., Suklitsch, M., Heinrich, G., 2015. The effect of empirical-statistical correction of intensity-dependent model errors on the temperature climate change signal. Hydrol. Earth Syst. Sci. 19, 4055–4066.
Haerter, J.O., Berg, P., 2009. Unexpected rise in extreme precipitation caused by a shift in rain type? Nat. Geosci. 2, 372–373.
Haerter, J.O., Hagemann, S., Moseley, C., Piani, C., 2011. Climate model bias correction and the role of timescales. Hydrol. Earth Syst. Sci. 15, 1065–1079.
Hagemann, S., Chen, C., Haerter, J.O., Heinke, J., Gerten, D., Piani, C., 2011. Impact of a statistical bias correction on the projected hydrological changes obtained from three GCMs and two hydrology models. J. Hydrometeorol. 12, 556–578.
Hassellmann, K., 1979. On the signal-to-noise problem in atmospheric response studies. Meteorol. Trop. Oceans.
Hasselmann, K., 1993. Optimal fingerprints for the detection of time-dependent climate change. J. Clim. 6, 1957–1971.
Hasselmann, K., 1997. Multi-pattern fingerprint method for detection and attribution of climate change. Clim. Dyn. 13, 601–611.
Hawkins, E., Sutton, R., 2009. The Potential to Narrow Uncertainty in Regional Climate Predictions. Bulletin of the American Meteorological Society 90, 1095–1108.
Hegerl, G., Zwiers, F., 2011. Use of models in detection and attribution of climate change. Wiley Interdiscip. Rev. 2, 570–591.
Hegerl, G.C., Hoegh-Guldberg, O., Casassa, G., Hoerling, M.P., Kovats, R., Parmesan, C., Pierce, D.W., Stott, P.A., 2010. Good practice guidance paper on detection and attribution related to anthropogenic climate change. Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Detection and Attribution of Anthropogenic Climate Change. IPCC Working Group I Technical Support Unit, University of Bern.

Hegerl, G.C., Zwiers, F.W., Braconnot, P., Gillett, N.P., Luo, Y., Marengo Orsini, J., Nicholls, N., Penner, J.E., Stott, P.A., 2007. Understanding and Attributing Climate Change.
Holgate, C., de Jeu, R., van Dijk, A., Liu, Y., Renzullo, L., Dharssi, I., Parinussa, R., van der Schalie, R., Gevaert, A., Walker, J., 2016. Comparison of remotely sensed and modelled soil moisture data sets across Australia. Remote Sens. Environ. 186, 479–500.
Johnson, F., Sharma, A., 2012. A nesting model for bias correction of variability at multiple time scales in general circulation model precipitation simulations. Water Resour. Res. 48.
Johnson, F., Sharma, A., 2015. What are the impacts of bias correction on future drought projections? J. Hydrol. 525, 472–485.
Johnson, F., Westra, S., Sharma, A., Pitman, A.J., 2011. An assessment of GCM skill in simulating persistence across multiple time scales. J. Clim. 24, 3609–3623.
Jones, G.S., Stott, P.A., Christidis, N., 2013. Attribution of observed historical near–surface temperature variations to anthropogenic and natural causes using CMIP5 simulations. J. Geophys. Res. 118, 4001–4024.
Kaufmann, R.K., Stern, D.I., 1997. Evidence for human influence on climate from hemispheric temperature relations. Nature 388, 39.
Lloyd, E.A., Oreskes, N., 2018. Climate change attribution: when is it appropriate to accept new methods. Earth's Fut. 6, 311–325.
Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M., Francis, R.C., 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. Bull. Am. Meteorol. Soc. 78, 1069–1080.
Marvel, K., Cook, B.I., Bonfils, C.J.W., Durack, P.J., Smerdon, J.E., Williams, A.P., 2019. Twentieth-century hydroclimate changes consistent with human influence. Nature 569, 59–65.
Masson, D., Knutti, R., 2011. Spatial-scale dependence of climate model performance in the CMIP3 ensemble. J. Clim. 24, 2680–2692.
Maurer, E.P., Das, T., Cayan, D.R., 2013. Errors in climate model daily precipitation and temperature output: time invariance and implications for bias correction. Hydrol. Earth Syst. Sci. 17, 2147–2159.
Meehl, G.A., Washington, W.M., Wigley, T., Arblaster, J.M., Dai, A., 2003. Solar and greenhouse gas forcing and climate response in the twentieth century. J. Clim. 16, 426–444.
Mehrotra, R., Sharma, A., 2006. A nonparametric stochastic downscaling framework for daily rainfall at multiple locations. J. Geophys. Res. 111.
Mehrotra, R., Sharma, A., 2015. Correcting for systematic biases in multiple raw GCM variables across a range of timescales. J. Hydrol. 520, 214–223.
Mehrotra, R., Sharma, A., Cordery, I., 2004. Comparison of two approaches for downscaling synoptic atmospheric patterns to multisite precipitation occurrence. J. Geophys. Res. 109.
Min, S.-K., Zhang, X., Zwiers, F.W., Hegerl, G.C., 2011. Human contribution to more-intense precipitation extremes. Nature 470, 378.
Mitchell, J.F., 2001. Detection of climate change and attribution of causes. Climate Change 2001: The Scientific Basis.
Mondal, A., Mujumdar, P.P., 2015. On the detection of human influence in extreme precipitation over India. J. Hydrol. 529, 1161–1172.
Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., 2013. Anthropogenic and natural radiative forcing. Clim. change 423, 658–740.
Nahar, J., Johnson, F., Sharma, A., 2017. Assessing the extent of non-stationary biases in GCMs. J. Hydrol. 549, 148–162.
Najafi, M.R., Zwiers, F.W., Gillett, N.P., 2015. Attribution of Arctic temperature change to greenhouse-gas and aerosol influences. Nat. Clim. Change 5, 246.
Newman, M., Compo, G.P., Alexander, M.A., 2003. ENSO-forced variability of the Pacific decadal oscillation. J. Clim. 16, 3853–3857.
Orlowsky, B., Seneviratne, S.I., 2013. Elusive drought: uncertainty in observed trends and short-andlong-term CMIP5 projections. Hydrol. Earth Syst. Sci. 17, 1765–1781.
Owe, M., de Jeu, R., Holmes, T., 2008. Multisensor historical climatology of satellite-derived global land surface moisture. J. Geophys. Res. 113.
Pachauri, R.K., Reisinger, A., 2007. Synthesis report. In: Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 151–165.
Pall, P., Aina, T., Stone, D.A., Stott, P.A., Nozawa, T., Hilberts, A.G., Lohmann, D., Allen, M.R., 2011. Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. Nature 470, 382.
Pierce, D.W., Barnett, T.P., Santer, B.D., Gleckler, P.J., 2009. Selecting global climate models for regional climate change studies. Proc. Natl. Acad. Sci. 106, 8441–8446.
Randall, D.A., Wood, R.A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., 2007. Climate models and their evaluation. Climate change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC (FAR). Cambridge University Press.
Ribes, A., Terray, L., 2013. Application of regularised optimal fingerprinting to attribution. Part II: application to global near-surface temperature. Clim. Dyn. 41, 2837–2853.
Ribes, A., Zwiers, F.W., Azaïs, J.-M., Naveau, P., 2017. A new statistical approach to climate change detection and attribution. Clim. Dyn. 48, 367–386.
Ropelewski, C.F., Halpert, M.S., 1987. Global and regional scale precipitation patterns associated with the El Niño/Southern oscillation. Mon. Weather Rev. 115, 1606–1626.
Saini, R., Wang, G., Yu, M., Kim, J., 2015. Comparison of RCM and GCM projections of boreal summer precipitation over Africa. J. Geophys. Res. Atmos. 120 (9), 3679–3699. https://doi.org/10.1002/2014jd022599.
Santer, B.D., Mikolajewicz, U., Brüggemann, W., Cubasch, U., Hasselmann, K., Höck, H., Maier-Reimer, E., Wigley, T.M.L., 1995. Ocean variability and its influence on the detectability of greenhouse warming signals. J. Geophys. Res. 100, 10693–10725.
Sheikholeslami, R., Razavi, S., Gupta, H.V., Becker, W., Haghnegahdar, A., 2019. Global sensitivity analysis for high-dimensional problems: how to objectively group factors and measure robustness and convergence while reducing computational cost. Environ. Modell. Softw. 111, 282–299.
Shiogama, H., Stone, D.A., Nagashima, T., Nozawa, T., Emori, S., 2013. On the linear additivity of climate forcing-response relationships at global and continental scales. Int. J. Climatol. 33, 2542–2550.
Shumway, R.H., Stoffer, D.S., 2011. Time series regression and exploratory data analysis. Time Series Analysis and Its Applications. Springer.
Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. 2013a. Climate change 2013. The physical science basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change-Abstract for decision-makers. Groupe d'experts intergouvernemental sur l'evolution du climat/Intergovernmental Panel on Climate Change-IPCC.
Stocker, T.F., Qin, D., Plattner, G.K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, B., Midgley, B.M. 2013b. IPCC, 2013: climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.
Sun, Y., Solomon, S., Dai, A., Portmann, R.W., 2006. How often does it rain? J. Clim. 19, 916–934.
Taylor, C.M., de Jeu, R.A., Guichard, F., Harris, P.P., Dorigo, W.A., 2012. Afternoon rain more likely over drier soils. Nature 489, 423.
Terray, L., 2012. Evidence for multiple drivers of North Atlantic multi-decadal climate variability. Geophys. Res. Lett. 39.
Teutschbein, C., Seibert, J., 2013. Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? Hydrol. Earth Syst. Sci. 17, 5061–5077.
van Dijk, A., Yebra, M., Cary, G., 2015. A model-data fusion framework for estimating fuel properties. Vegetation Growth Carbon Storage and the Water Balance at Hillslope Scale: Feasibility Study in Namadgi National Park ACT. Bushfire and Natural Hazards CRC.
van Dijk, A.I., Beck, H.E., Crosbie, R.S., de Jeu, R.A., Liu, Y.Y., Podger, G.M., Timbal, B., 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 Resour. Res. 49, 1040–1057.
Verdon, D.C., Franks, S.W., 2006. Long-term behaviour of ENSO: interactions with the PDO over the past 400 years inferred from paleoclimate records. Geophys. Res. Lett. 33.
Viney, N., Vaze, J., Crosbie, R., Wang, B., Dawes, W., Frost, A., 2015. AWRA-L v5. 0: Technical Description of Model Algorithms and Inputs. CSIRO.
Viney, N.R., Vaze, J., Vleeshouwer, J., Yang, A., van Dijk, A., Frost, A., 2014. The AWRA modelling system. In: Hydrology and Water Resources Symposium 2014. Engineers Australia, p. 1018.
Wan, H., Zhang, X., Zwiers, F., Min, S.-K., 2015. Attributing northern high-latitude precipitation change over the period 1966–2005 to human influence. Clim. Dyn. 45, 1713–1726.
Wanders, N., Karssenberg, D., Roo, A.D., de jong, S., Bierkens, M., 2014. The suitability of remotely sensed soil moisture for improving operational flood forecasting. Hydrol. Earth Syst. Sci. 18, 2343–2357.
Willett, K.M., Gillett, N.P., Jones, P.D., Thorne, P.W., 2007. Attribution of observed surface humidity changes to human influence. Nature 449, 710.
Woldemeskel, F., Sharma, A., Sivakumar, B., Mehrotra, R., 2016. Quantification of precipitation and temperature uncertainties simulated by CMIP3 and CMIP5 models. J. Geophys. Res. 121, 3–17.
Zhang, X., Zwiers, F.W., Hegerl, G.C., Lambert, F.H., Gillett, N.P., Solomon, S., Stott, P.A., Nozawa, T., 2007. Detection of human influence on twentieth-century precipitation trends. Nature 448, 461.