Assessment of apparent nonstationarity in time series of annual inflow, daily precipitation, and atmospheric circulation indices: A case study from southwest Western Australia

Bryson C. Bates,[1] Richard E. Chandler,[2] Stephen P. Charles,[3] and Edward P. Campbell[4]

[1]   The southwest region of Western Australia has experienced a sustained sequence of low annual inflows to major water supply dams over the past 30 years. Until recently, the dominant interpretation of this phenomenon has been predicated on the existence of one or more sharp breaks (change or jump points), with inflows fluctuating around relatively constant levels between them. This paper revisits this interpretation. To understand the mechanisms behind the changes, we also analyze daily precipitation series at multiple sites in the vicinity and time series for several indices of regional atmospheric circulation that may be considered as drivers of regional precipitation. We focus on the winter half-year for the region (May to October) as up to 80% of annual precipitation occurs during this "season". We find that the decline in the annual inflow is in fact more consistent with a smooth declining trend than with a sequence of sharp breaks, the decline is associated with decreases both in the frequency of daily precipitation occurrence and in wet-day amounts, and the decline in regional precipitation is strongly associated with a marked decrease in moisture content in the lower troposphere, an increase in regionally averaged sea level pressure in the first half of the season, and intraseasonal changes in the regional north-south sea level pressure gradient. Overall, our approach provides an integrated understanding of the linkages between declining dam inflows, declining precipitation, and changes in regional atmospheric circulation that favor drier conditions.

1.   Introduction

[2]   Surface water management systems have customarily been designed and operated under the assumption of statistical stationarity of inflow series. That is, the probability distribution of observations does not change with time and hence the statistics of future inflow series will replicate those of the past. *Milly et al.* [2008] argue that stationarity should no longer serve as the basis for water resources risk assessment and planning given the growing evidence of hydroclimatic change. Recent studies reporting nonstationarity include those by *Mauget* [2003], *Kahya and Kalayci* [2004], *Stewart et al.* [2005], *Dixon et al.* [2006], *Zheng et al.* [2007], *Adam and Lettenmaier* [2008], *Burn et al.* [2008], and *Miller and Piechota* [2008].

[3]   Southern Australia is experiencing an extended period of drought, with southwest Western Australia (SWA) in a state of hydrological drought since the mid-1970s [*Power et al.*, 2005; *Bates et al.*, 2008; *Murphy and Timbal*, 2008]. The decline in aggregated annual inflows to the 11 major dams of the Integrated Water Supply Scheme (IWSS) in SWA is illustrated in Figure 1. Historically, the dominant interpretation of this inflow series has been predicated on the existence of one or more change points, with inflows fluctuating around relatively constant levels between them [see, e.g., *IPCC*, 2007, Figure 11.3]. Detection of change points is often based on statistical hypothesis testing techniques such as those reviewed by *Jarušková* [1997]. Other approaches to the assessment of nonstationarity in hydrological time series include techniques such as linear regression and nonparametric tests to establish or refute the existence of trends in individual series [see, e.g., *Kundzewicz and Robson*, 2004]. However, all of these techniques are necessarily limited in the understanding that they can deliver: they may provide evidence of nonstationarity, but they do not provide insight into the causal mechanisms responsible. Moreover, they rely on simplifying assumptions that are often unrealistic. For example, for most standard trend tests and change point detection techniques the null hypothesis is that the observations are independent and identically distributed [*Kendall and Ord*, 1990, section 2.1; *Jarušková*, 1997,

[1]Climate Adaptation Flagship, CSIRO Marine and Atmospheric Research, Wembley, Western Australia, Australia.
[2]Department of Statistical Science, University College London, London, UK.
[3]Climate Adaptation Flagship, CSIRO Land and Water, Wembley, Western Australia, Australia.
[4]Climate Adaptation Flagship, CSIRO Mathematics, Informatics and Statistics, Wembley, Western Australia, Australia.
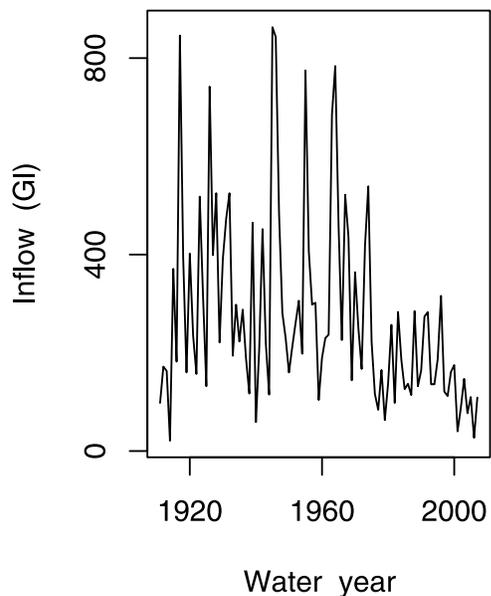
Figure 1. Aggregated annual inflow series for 11 major dams in the Integrated Water Supply Scheme (IWSS). Water year is May to April.

section 4]. The rejection of such a hypothesis implies that the observations are either dependent or have different distributions (or both), but it does not follow that a trend or change point is present. One could imagine testing for a change point in a series with a linear trend: given enough data, any such test will correctly reject the null hypothesis, but it would be wrong to infer that the series contains an abrupt break. The effective use of test procedures therefore relies on the ability to construct null hypotheses in such a way that if they are rejected, the alternative of interest provides the most plausible explanation. Broadly speaking, there are two ways to achieve this: use either an appropriately designed experiment to collect the data or more sophisticated statistical methods. In environmental applications, designed experiments are rarely feasible except when computer models are used to generate data (see section 4.2.3 below for an example); therefore more sophisticated methods are needed.

[4] In this paper we use a suite of modern, model-based statistical approaches to build an explicit representation of the changes in water availability and the associated climatic drivers in SWA. This empirical approach provides insight into potential causal pathways from changes in regional atmospheric circulation to changes in at-site daily precipitation (occurrence and amounts), and from at-site daily precipitation to dam inflow. In section 2 we briefly describe the study area, review previous research, and present our approach. The data and methods used are described in sections 3 and 4, respectively. Results are presented in section 5, where we demonstrate that the IWSS inflow data are in fact more consistent with a smooth declining trend than with a sequence of sharp breaks, and we subsequently seek to understand the climatic drivers of this trend by analyzing daily time series of at-site precipitation and several indices of regional atmospheric circulation (of demonstrated relevance to regional precipitation) for the period from the

middle of the 20th century to the present. Finally, a discussion and our findings are given in section 6.

2. Study Area, and Previous and Present Research

2.1. Physical and Climatic Setting

[5] SWA extends approximately from 30° to 35°S and from 115° to 120°E (Figure 2). It is bounded by the Indian Ocean to the west and the Southern Ocean to the south. A 320 km escarpment runs parallel to and some 25–50 km from the coast with rain shadow effects eastward. The average and maximum heights of the escarpment are about 300 and 582 m at Mount Cooke (32°25′S, 116°18′E), respectively.

[6] The region experiences a Mediterranean climate with hot dry summers and mild wet winters. During autumn (March to May) the subtropical belt of high pressure moves northward, and it lies almost outside the region during winter (June to August). Owing to the breakup of the belt into anticyclonic cells, the prevailing winds are anticlockwise, with moist westerly flow during winter. The winds attain greater speeds than in similar latitudes in the Northern Hemisphere due to the low orography of the region [*Gentilli*, 1972]. During winter the mean track of low pressure systems is always south of the region and moves away to the southeast. Thus precipitation decreases from west to east and from south to north [see, e.g., *Bates et al.*, 2008, Figure 3]. For the 30 daily precipitation stations depicted in Figure 2, the percentage of annual precipitation that falls in the winter half-year (May to October) varies from 66% to 86%: 25 stations have percentages greater than 71%. Two fundamental types of winter half-year precipitation have been identified [*Wright*, 1974]: continuous precipitation due to uplift in the midtroposphere associated with surface winds north of west (dominant in May to July), and showers associated with convection in a moist unstable airstream from between west and southwest that are enhanced by topography and coastal convergence (dominant in August to October). Precipitation of the first type is closely related to the intensity of the westerlies, but precipitation for the second type is not similarly related to atmospheric circulation features.

[7] Climate change has been evident since the 1970s. Annual mean temperatures have increased at the rate of $+0.15°C$ decade^{-1}, with increases occurring in all seasons except summer, where the rate of change is $-0.1°C$ decade^{-1}. Simultaneously, there has been a reduction in winter half-year precipitation, particularly for May to July, where the mean precipitation for 1975–2004 is about 14% less than that for the mid-1900s to 1974. Another notable feature is the absence of very wet years since the mid-1970s [*Bates et al.*, 2008].

2.2. Integrated Water Supply Scheme (IWSS)

[8] The IWSS supplies water for 1.5 million people in the city of Perth (31°57′S, 115°52′E) and surrounding areas. The IWSS extends more than 600 km east-west and 200 km north-south. Despite the extensive use of demand management strategies, substantial investment has been required due to reductions in annual inflows and groundwater recharge coupled with population growth. Between 1996 and 2006, the Water Corporation (which provides water, sewerage,
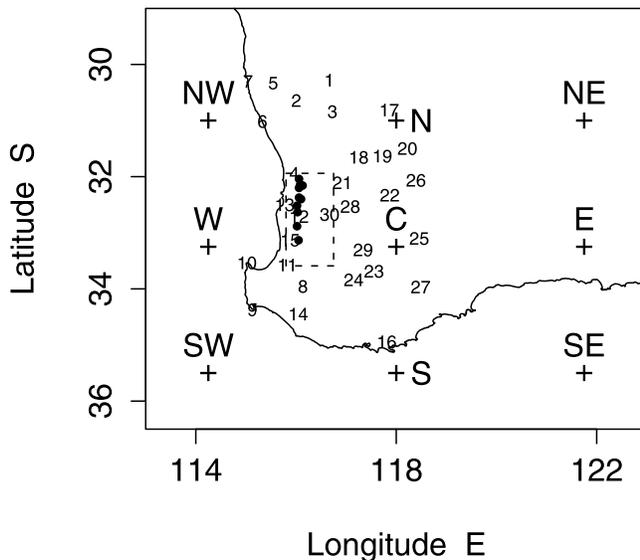
Figure 2. Map of study area. Dots denote locations of major dams in the IWSS and numerals denote the 30 daily precipitation gauges used by *Charles et al.* [1999a, 1999b, 2004] and *Bates et al.* [2001]. Crosses denote grid points for atmospheric circulation data. "C", "E", "N", "S", and "W" denote central, east, north, south, and west grid points. Dashed box indicates the subregion (roughly 150×100 km^2) over which the at-site precipitation analysis is performed.

and drainage services to 97% of Western Australia's population) invested $A921 million in source development for the IWSS. At the start of 2007, the water supply to the IWSS was derived from 11 main dams (Figure 2) with total storage capacities ranging between 2 and 138 GL, 136 groundwater bores located in confined and unconfined aquifers underlying the northern metropolitan area, and a new 45 GL yr^{-1} seawater desalination plant.

2.3. Related Studies

[9] There are now several articles related to the observed precipitation decline in SWA. *IOCI* [2002] and *Bates et al.* [2008] discuss the issue broadly, although the latter also discusses the impacts of the research outcomes on policy development across Australia. Key findings from these previous studies include the following: (1) winter precipitation has declined by about 15%–20% [*Hennessy et al.*, 1999; *IOCI*, 2002; *Smith*, 2004]; (2) declines in the number of wet days and precipitation amounts in extreme events are evident [*IOCI*, 1999; *Y. Li et al.*, 2005]; (3) most of the precipitation decline is confined to the first half of the winter half-year (May to July) [*IOCI*, 2002; *Bates et al.*, 2008]; (4) there is a statistically significant inverse relationship between regional to near global scale mean sea level pressure (SLP) and winter precipitation on interannual to decadal time scales [*Allan and Haylock*, 1993; *Ansell et al.*, 2000; *F. Li et al.*, 2005]; (5) while the strength of the direct linkage between the precipitation decline and changes in sea surface temperatures (SSTs) is less prominent, SSTs affect regional precipitation primarily through their relationship with SLP [*Ansell et al.*, 2000; *Smith et al.*, 2000; *F. Li et al.*, 2005; *England et al.*, 2006]; (6) an increase in SLP in winter

and decreasing trends in atmospheric moisture in winter and spring are important in explaining the precipitation decline [*Charles et al.*, 1999b; *Bates et al.*, 2001; *Timbal*, 2004]; (7) while average SLP has increased across the region, the frequency and intensity of troughs associated with wet conditions has declined markedly since the mid-1970s, and this is responsible for almost half the reduction in observed precipitation from 1958–1975 to 1976–2003 [*Simmonds and Keay*, 2000; *Hope et al.*, 2006]; (8) since the mid-1970s there has been a 20% reduction in the strength of the subtropical jet over Australia, a reduction in the intensity of cyclogenesis across southern Australia, and a southward deflection of some storms [*Frederiksen and Frederiksen*, 2007]; (9) the precise attribution of the precipitation decline to natural forcing and anthropogenic factors such as increasing atmospheric concentrations of trace greenhouse gases, land clearing, and changes in forest and forest fire management remains elusive [*IOCI*, 2002; *Cai et al.*, 2005; *Timbal et al.*, 2006; *Bates et al.*, 2008].

2.4. Contribution and Approach

[10] This paper differs from previous works on observed precipitation and inflow declines in SWA in four major respects: (1) our analysis introduces a comprehensive methodological framework consisting of a suite of modern, model-based statistical techniques, with due attention to the underlying assumptions; (2) considerable attention is focused directly on the annual inflow series for the IWSS, and on precipitation changes in the vicinity of its main dams; (3) a stochastic downscaling model is used to provide a synoptic weather-typing scheme that relates spatial patterns in daily precipitation occurrence to a more comprehensive set of indices of regional-scale atmospheric circulation ("atmospheric predictors") and subsequently to determine the role of these indices, individually and in combination, in explaining changes in weather type frequencies; (4) rather than using exploratory analysis and simple linear trends, we provide a formal and rigorous analysis of the temporal changes in seasonality as well as interannual variations in precipitation, atmospheric predictors, and weather types.

[11] Figure 3 illustrates the data and methods used. Three data sets are examined: the annual inflow series for the IWSS dams (section 3.1); daily precipitation occurrence and wet-day amount series (section 3.2); atmospheric predictor series derived from data provided by the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) Reanalysis project (section 3.3). Trend detection and estimation were carried out using a suite of parametric (sections 4.1.1–4.1.3) and nonparametric (section 4.1.4) regression techniques. Subsequently, a stochastic downscaling model (the nonhomogeneous Markov model, NHMM) was used to provide an explicit, objective, and quantitative representation of the linkage between indices of regional atmospheric circulation and daily precipitation occurrence at multiple sites (section 4.2.1). A useful by-product of the NHMM is an objective weather typing that links the atmospheric predictors (section 4.2.2) to a finite number of spatial patterns of daily precipitation occurrence. The frequencies of the resulting weather types (or "states") were subjected to trend analysis using local linear regression (LLR), as was the annual inflow series for the IWSS. Finally, to investigate the
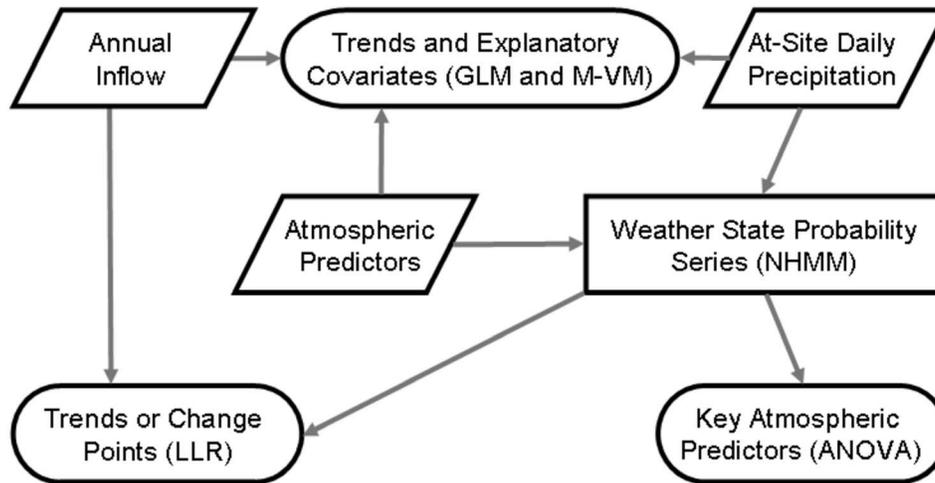
Figure 3. Data flow diagram for the analysis. Arrows depict data flow lines, parallelograms depict input data, rectangles depict computational steps, and ovals depict end results. ANOVA, analysis of variance; GLM, generalized linear model; LLR, local linear regression; M-VM, joint mean-variance model; NHMM, nonhomogeneous hidden Markov model.

roles of the different atmospheric circulation indices in controlling trends in weather state frequencies, a factorial experiment was carried out in which different configurations of the indices were used to drive NHMM simulations (section 4.2.3).

3. Description of Data

3.1. Annual Inflow Series

[12] We used aggregated annual inflow data for the main IWSS dams for the period 1911–2007. Following local convention, the water year is May to April and is designated by the calendar year in which it starts. The inflow series extends further back in time than the existence of most IWSS dams: only two dams were operational in 1911. Thus inflows prior to the 1950s were largely reconstructed using a precipitation-runoff model and precipitation data from a network of gauges having limited coverage. The remainder of the series was derived from gauging station data or estimated using a water balance method for the short periods when stations were nonoperational. The inflow series from the 1950s to the present is regarded as reliable (C. Jeevaraj, personal communication, 2009).

3.2. At-Site Precipitation

[13] The locations of the 30 stations used in previous downscaling studies are shown in Figure 2 [see, e.g., *Charles et al.*, 1999a, 1999b, 2004; *Bates et al.*, 2001]. The stations have reasonably complete high-quality records from 1958 onward, although not all stations were operational throughout this period. They provide a nearly uniform spatial coverage of the region and hence the opportunity to characterize synoptic-scale precipitation structures.

[14] Long-term, high-quality precipitation records in the vicinity of the main IWSS dams are scant. We used daily precipitation series for the winter half-year recorded at stations 4, 11, 12, and 15 in Figure 2 and seven additional stations within the dashed box shown in Figure 2 (locations

not shown for clarity) to characterize precipitation trends. The record lengths at some of the additional stations are shorter than those for the first group (the shortest runs from 1974 to 2002), and two contained a relatively large number (around 15%) of missing observations.

[15] The data from most stations are recorded to a nominal resolution of 0.1 mm; however, there is a strong preference to record even-numbered decimal digits. Differences in recording resolution can lead to spurious trends that are detectable by the methods used below [*Yang et al.*, 2006]. To avoid this, the data were rounded to a resolution of 0.2 mm for the analyses reported in section 5.2. Moreover, at some sites, changes in recording practice from 2001 onward have led to the underrecording of nonzero values at weekends. The effect of this is to exaggerate drying trends in the records. To eliminate this effect, all post-2000 weekend observations from the affected sites, as well as a few other suspect observations, were discarded from our analyses. For consistency with the previous studies noted above, any day with a recorded precipitation amount less than 0.3 mm was considered "dry."

3.3. NCEP/NCAR Reanalysis

[16] We use data from the NCEP/NCAR Reanalysis project [*Kistler et al.*, 2001] for the period 1958–2007. Data are available from 1948 onward. However, the earliest decade is generally considered less reliable due to a scarcity of upper-air observations in the Southern Hemisphere. The data are 6 hourly and interpolated to a 2.5° latitude-longitude grid. We used the data at 1200 UT on the basis that they are close to the midpoint of the daily precipitation recording period for the study region (2400–0100 UT).

[17] Variables in the reanalysis are classified into four categories [*Kalnay et al.*, 1996; *Kistler et al.*, 2001], from the most reliable (type "A") to the least reliable (type "D"). We used four type "A" variables (SLP, geopotential heights at 500 and 850 hPa, and air temperature at 850 hPa) and one type "B" variable (specific humidity at 850 hPa, which was

used to estimate dew point temperature at that level). The atmospheric fields were interpolated onto the 2.25° latitude by 3.75° longitude grid used in previous downscaling studies (Figure 2).

4. Methods

4.1. Regression Techniques

[18] To understand the climatic mechanisms responsible for the clear trend in Figure 1, the first step is to quantify the nature of this trend, along with trends in precipitation and indices of regional-scale atmospheric circulation. Regression techniques are a natural tool for this purpose.

4.1.1. Multiple Linear Regression (MLR)

[19] When analyzing an individual series $Y_1,\ldots, Y_T$, the starting point for the methods considered herein is the linear trend model:

$$Y_t = \mu_t + \varepsilon_t, \qquad (1)$$

where $\mu_t = \beta_0 + \beta_1 t$ (the "trend function") is a linear function of the time index t and the "errors" (ε_t) form a sequence of independent, normally distributed random variables with zero mean and a common variance. For the analysis of environmental time series at subannual time resolution, however, such a model is rarely adequate: most series exhibit seasonality and autocorrelation, for example, as well as potential relationships with other variables of interest. Such features can be accommodated straightforwardly in the regression framework by writing $\mu_t = \beta_0 + \sum_{j=1}^{p}\beta_j x_{jt}$, so that the expected value of Y_t is now a linear function of several covariates, of which one might be t. Seasonality can be incorporated via covariates defined as sine and cosine functions representing the annual cycle and, if necessary, its harmonics. Autocorrelation can be accounted for either by specifying an appropriate time series structure for ε_t or by including an appropriate number of lagged values $Y_{t-1}, Y_{t-2},\ldots$ as extra covariates in the model. The latter approach suffers from the disadvantage that interpretation of model coefficients can be complicated when lagged values are included [e.g. *Greene*, 2003, section 19.4]. However, it offers increased flexibility in other aspects. One such aspect is the possibility of modeling interactions involving the lagged values: in a regression model, two covariates are said to interact if one of them modulates the effect of the other or, equivalently, if the regression coefficient of one of them can be written as a function of the other. This can be accommodated by adding an extra term, involving the product of the interacting covariates, to the model [*Chandler and Wheater*, 2002, section 3.2]. Interactions involving lagged values imply that the autocorrelation structure, which at fine time scales reflects the dynamics of the underlying processes, changes in response to other variables; interactions between the time index and seasonal covariates suggest that the seasonal cycle is changing over time.

4.1.2. Generalized Linear Models (GLMs)

[20] A key assumption of the linear regression model is that the observations are normally distributed given the covariates. We find that this assumption is reasonable for the logarithms of the dam inflow series considered in this paper, and for the atmospheric predictors. However, the distribution of precipitation at a daily time scale is far from normal,

so that other techniques are required to analyze precipitation records. GLMs extend the linear regression model by allowing multiple nonnormal response variables and covariates, and hence provide the required flexibility. For the analysis of an individual series, a GLM specifies a distribution for Y_t with mean μ_t, linked to a vector x_t of covariates via an equation of the form

$$g(\mu_t) = \beta_0 + \sum_{j=1}^{p}\beta_j x_{jt}, \qquad (2)$$

for some monotonic function $g(\cdot)$, known as the link function. Note that the standard linear regression model can be regarded as a GLM in which the link function $g(\cdot)$ is taken as the identity and the distribution of the $\{Y_t\}$ is normal with constant variance. Seasonality and interactions can be represented in the same way as before. In general, however, autocorrelation is most easily dealt with by including lagged values $Y_{t-1}, Y_{t-2},\ldots$ as extra covariates, possibly after appropriate transformation. Estimation of the coefficients $\beta_0,\ldots, \beta_p$ is usually carried out using maximum likelihood, which has the potential to detect weak and complex signals in noisy records such as daily precipitation sequences [see, e.g., *Chandler and Wheater*, 2002]. If required, the importance of several terms in a model can be assessed simultaneously by comparing the log likelihoods of models with and without the terms of interest, providing the models have been fitted to the same data. This procedure is referred to as likelihood ratio testing [*Davison*, 2003, section 4.5].

[21] If a GLM is to be fitted simultaneously to data from several sites as in our precipitation analyses, in general it is necessary to include covariates accounting for systematic regional variation; for example, precipitation occurrence may be altitude dependent. Such structure is often difficult to parameterize, in which case it may be modeled via an appropriate orthogonal basis representation [see *Chandler*, 2005]. In addition, it is necessary to account for intersite correlation. We have done this by fitting models as though sites are independent and then making appropriate adjustments to the standard errors of coefficient estimates [see *Chandler*, 2005], as well as to likelihood ratio test statistics [*Chandler and Bate*, 2007].

[22] Following *Coe and Stern* [1982] the probability p_{st} of precipitation at site s on day t is modeled using logistic regression, as

$$\log[p_{st}/(1-p_{st})] = \beta_0 + \sum_{j=1}^{p}\beta_j x_{st}^{(j)}, \qquad (3)$$

where the $\{x_{st}^{(j)}\}$ are the corresponding covariate values. If site s experiences rain on day t, the nonzero precipitation amount is taken to be gamma distributed with mean μ_{st}, such that

$$\log \mu_{st} = \gamma_0 + \sum_{j=1}^{p}\gamma_j \xi_{st}^{(j)}, \qquad (4)$$

where now the $\{\xi_{st}^{(j)}\}$ are covariate values and the $\{\gamma_j\}$ are coefficients. The gamma distributions are all assumed to have a common shape parameter, ν, say. Throughout the work described below, we use a range of diagnostics to

check the assumptions underlying our GLMs; for details of these diagnostics, see *Chandler and Wheater* [2002].

4.1.3. Joint Mean-Variance Models (M-VMs)

[23] From an exploratory analysis (not shown) of the atmospheric drivers considered in this paper, the assumption of normality seems reasonable, but there is seasonal structure in the variance. This cannot be handled directly using the methods discussed so far; however, the variance structure is of some interest, particularly if a formal analysis indicates systematic long-term changes in the variability of one or more drivers. The situation can be dealt with using a joint mean-variance model under the assumption of normality, as described by *Yang et al.* [2006], for example. Such models consist essentially of a pair of linked GLMs describing the mean and variance structure, respectively. Fitting is again carried out using maximum likelihood, and the usual procedures are available for comparing and checking models and for testing hypotheses. Care is required, however, if the normality assumption fails, because this is required for accurate inference about the factors driving changes in variability.

4.1.4. Local Linear Regression (LLR)

[24] In the techniques described above, it has been assumed implicitly that any temporal trends can be represented adequately by incorporating the time index t as a covariate. For most of the analyses reported below, this yields a perfectly acceptable, albeit sometimes approximate, summary of overall change. However, analysis of the aggregated inflow series requires more care given the implications of different trend scenarios for water resource management [see, e.g., *Brennan*, 2008]. In view of this, for the inflow analysis we have adopted a nonparametric approach which imposes minimal assumptions on the nature of the trend function and hence allows the data to "speak for themselves" as far as possible. Several such approaches are possible; we have used local linear regression [*Bowman and Azzalini*, 1997], as implemented via the "sm" library in the R programming environment [*Bowman and Azzalini*, 2007]. In this framework, the trend function μ_t in (1) is assumed only to be smooth in the first instance, although the errors (ε_t) are assumed to be independent and normal with zero mean and constant variance. Smoothness implies that the trend function can be approximated by a straight line in the neighborhood of any time point τ: $\mu_t \approx \mu_\tau + \beta_\tau(t - \tau)$, say. Estimates of the coefficients μ_τ, β_τ are found by solving the weighted least squares problem:

$$\min_{\mu_\tau, \beta_\tau} \sum_{t=1}^{T} [Y_t - \mu_\tau - \beta_\tau(t - \tau)]^2 w(t - \tau; h), \qquad (5)$$

where the weight function $w(\cdot; h)$ is the normal probability density function with mean 0 and standard deviation h so that observations distant from τ are downweighted. The estimate of μ_τ is then taken as an estimate of the underlying trend at time τ. This procedure is repeated over a range of values of τ, to build up a complete picture of the estimated trend function. The quantity h in (5) must be chosen by the user: it is known as the smoothing parameter or bandwidth, as it controls the smoothness of the resulting trend estimate. It is expressed in the same units as t. The sm library offers three different automatic methods for the selection of h: cross validation (h_{cv}), an approximate degrees of

freedom criterion (h_{df}), and a corrected Akaike information criterion (h_{aicc}) [*Bowman and Azzalini*, 1997; *Hurvich et al.*, 1998]. We compared the results from all of these to ensure that our conclusions are not sensitive to the precise choice of h. In addition, residual checks were carried out to check the assumptions of normality, constant variance, and independence.

[25] A key issue is whether the fitted regression curve represents a real long-term trend or whether it can be attributed to random variation. This can be addressed by testing the null hypothesis $H_0: \mu_t = \mu$ (i.e., the null model of a constant mean) against the alternative H_1 that the mean changes over time (i.e., that a genuine trend is present). Another null hypothesis that is perhaps of more interest is $H_0: \mu_t = \beta_0 + \beta_1 t$. Acceptance of this hypothesis indicates that any trend is represented reasonably by a linear function. Rejection indicates that the mean changes nonlinearly over time. The sm library contains routines, based on procedures described in detail in *Bowman and Azzalini* [1997], that enable formal tests of either hypothesis. The tests are constructed from pseudo-likelihood ratio statistics of the form $F = (\text{RSS}_0 - \text{RSS}_1)/\text{RSS}_1$, where RSS_i denotes the residual sum of squares under hypothesis H_i ($i = 0, 1$). For the null hypothesis of no change, RSS_0 is the sum of squared deviations from the mean of the observations; for the hypothesis of a linear trend, it is the sum of squared residuals from the least-squares trend line; and in both cases RSS_1 is the sum of squared residuals from the local linear regression fit. Under the null hypothesis, the observed significance level (p-value) for the test can be obtained from the quantiles of a scaled and shifted chi-square distribution. A plot of the p-value as a function of the bandwidth h is referred to as a significance trace, and it assesses the sensitivity of the test results to the choice of h. So-called variability bands indicating the size of two standard errors above and below $\hat{\mu}_t$ can also be constructed, to provide an informal assessment of uncertainty in the estimated trend function.

[26] The local linear regression methodology has also been extended by *Bowman et al.* [2006], to allow for the possibility of discontinuities in the regression function. This is particularly relevant here, given that conflicting interpretations of the trend in inflows can have substantial implications for water infrastructure planning [*Brennan*, 2008]. The nonparametric approach allows us to evaluate the evidence for discontinuities in a data-driven way, without imposing assumptions (such as linearity) on the nature of any smooth underlying regression function: we are trying to be as objective as possible. Moreover, the approach acknowledges that a series may contain both discontinuities and a smooth trend, which is a substantial advantage over standard techniques for discontinuity detection that implicitly assume that any long-term change must be due to a discontinuity (see section 1 above). The possibility that a series contains only discontinuities and no additional trend is, of course, not ruled out: in this case, the "smooth trend" is just a constant function.

[27] The essence of the *Bowman et al.* [2006] discontinuity test is as follows: denoting by τ a potential time at which a discontinuity may have occurred, two smooth estimates of the regression function at τ are obtained. The first is obtained from observations before τ, and the second is obtained from observations afterward. A large difference between the two estimates indicates a discontinuity at τ.

To test the null hypothesis that the trend function is smooth throughout (i.e., that no discontinuities are present), a test statistic is constructed as a sum of squared standardized differences between left and right estimators of the regression function at a grid of τ values spanning the entire series. Here we have used the grid $\tau \in \{5.5, 6.5,\ldots, T - 4.5\}$, where T is the series length: this ensures that at least five observations are available from which to estimate the regression function on each side of every candidate discontinuity location. Large values of this test statistic provide strong evidence against the null hypothesis and hence suggest that one or more discontinuities are present, and p-values can be calculated under the assumption of independent normal errors with zero mean and constant variance. As before, a significance trace can be constructed to assess the sensitivity of the results to the choice of bandwidth.

4.2. Stochastic Downscaling

[28] The second step in our analysis uses an NHMM to characterize the linkage between indices of regional atmospheric circulation and precipitation occurrence. Here we briefly describe the model and its use to explore the way in which atmospheric predictors have combined to produce changes in precipitation regimes.

4.2.1. Model Description

[29] The NHMM has been chosen for this analysis because it has been subjected to an extensive series of tests and applications in SWA over the past decade [see, e.g., *Bates et al.*, 1998, 2001; *Charles et al.*, 1999a, 1999b, 2004, 2007]. It is particularly noteworthy that a NHMM fitted to the winter half-years in the period 1978–1992 is able to reproduce several interannual to interdecadal at-site precipitation occurrence statistics for the wetter epoch from 1958 to 1977 and the drier epoch from 1993 to 2007 [*Charles et al.*, 2007]. The model structure and assumptions therefore seem robust to marked changes in regional atmospheric circulation over the study period.

[30] The NHMM relates synoptic-scale, atmospheric circulation variables to multisite daily precipitation occurrence via a finite number N of "hidden" (unobserved) weather states [*Charles et al.*, 1999a, 1999b; *Hughes et al.*, 1999]. For day t, let $\mathbf{R}_t = \{R_t^{(1)},\ldots, R_t^{(n)}\}$ denote the precipitation occurrence pattern at n stations, with observed values $\mathbf{r} = \{r_t^{(1)},\ldots, r_t^{(n)}\}$ where $r_t^{(i)} = 1$ if precipitation occurs at station i and 0 otherwise. Moreover, let S_t denote the unobserved weather state and $\mathbf{X}_t$ denote a vector of atmospheric circulation data (atmospheric predictors). In the analyses below, we used centered predictors obtained by subtracting long-term means from the reanalysis series of interest.

[31] The NHMM is based on two assumptions:

$$P\left(\mathbf{R}_t | S_1^T, \mathbf{R}_1^{t-1}, \mathbf{X}_1^T\right) = P(\mathbf{R}_t | S_t); \tag{6a}$$

$$P\left(S_t | S_1^{t-1}, \mathbf{X}_1^T\right) = P(S_t | S_{t-1}, \mathbf{X}_t), \tag{6b}$$

where the notation $\mathbf{X}_1^T$ means all values of $\mathbf{X}_t$ from day 1 to day T (similarly for S_1^T). Thus, the precipitation occurrence pattern on a given day is conditioned on the weather state for that day and the daily weather state sequence forms a first-order Markov process. The process is nonhomogeneous as

the transition probabilities depend on $\mathbf{X}_t$ and hence vary from day to day. A particular NHMM is defined by models for the precipitation occurrence distribution, $P(\mathbf{R}_t | S_t)$, and the weather state transition matrix, $P(S_t | S_{t-1}, \mathbf{X}_t)$. Parameterizations for $P(\mathbf{R}_t | S_t)$ and $P(S_t | S_{t-1}, \mathbf{X}_t)$, and model fitting procedures, are detailed by *Charles et al.* [1999a]. Weather state and daily precipitation occurrence sequences are generated using the procedure described by *Hughes and Guttorp* [1994, Appendix 2].

4.2.2. Atmospheric Predictors

[32] In this study we consider the NHMM fitted by *Bates et al.* [2001] to the NCEP/NCAR Reanalysis and daily precipitation data from the 30 precipitation stations depicted in Figure 2 for the winter half-years in the period 1978–1992. This model has six distinct and physically realistic weather states and uses four atmospheric predictors: (1) the mean of SLP over five grid points in Figure 2 [(N + W + C + E + S)/5], hereafter referred to as MSLP; (2) north-south SLP gradient (PG) across four grid points [(N + NW)/2 − (S + SW)/2]; (3) dew point temperature depression at 850 hPa (a measure of how close the atmosphere is to saturation at 850 hPa), defined by DTD = $T^{850} - T_d^{850}$, where T^{850} and T_d^{850} respectively denote the 850 hPa air temperature (K) and dew point temperature (K) at grid point C; (4) the first variate derived from a canonical correlation analysis of precipitation occurrence residuals [*Bates et al.*, 2001], which is defined approximately by

$$CV1 \approx 0.04 z(\text{MSLP}) - 0.05 z\left(\text{GPH}^{500}\right)$$
$$+ 0.02[z(\text{PG}_{\text{EW}}) + z(\Delta \text{GPH}) + z(\text{DTD})], \tag{7}$$

where $z(\cdot)$ denotes a standardized variable with mean zero and unit variance, GPH^{500} is the mean 500 hPa geopotential height (m), ΔGPH is the 850–500 hPa thickness (the difference in geopotential height between the 850 and 500 hPa levels) at grid point C, and PG_{EW} is the east-west SLP gradient across four grid points [(N + C)/2 − (NW + W)/2]. The selection of MSLP, PG, and DTD was based on the use of professional judgment, correlation analysis, classification trees, and sequential model fitting guided by the Bayes Information Criterion. The inclusion of CV1 improves the performance of the NHMM by removing a systematic bias in daily precipitation occurrence probabilities at intraseasonal scales.

[33] Equation (7) contrasts surface pressure, its east-west gradient, 850–500 hPa thickness, and the closeness to saturation at 850 hPa with midlevel geopotential height. Thus CV1 will be high when a strong midlevel trough over the region is accompanied by a dry lower troposphere, high 850–500 hPa thickness, high surface pressure, and high east-west surface pressure gradient. Inspection of the results of the canonical correlation analysis [*Bates et al.*, 2001] indicates that if precipitation occurs over the study region on a day when CV1 is high, it is more likely to occur in the northeast corner and well away from the catchments of the IWSS dams.

4.2.3. Sensitivity Analysis

[34] To investigate the controls of the different atmospheric predictors upon trends in the weather state probability series generated by the NHMM, we have carried out a full factorial experiment in which we assess the effect of re-ordering the predictor sequences. Separate analyses were
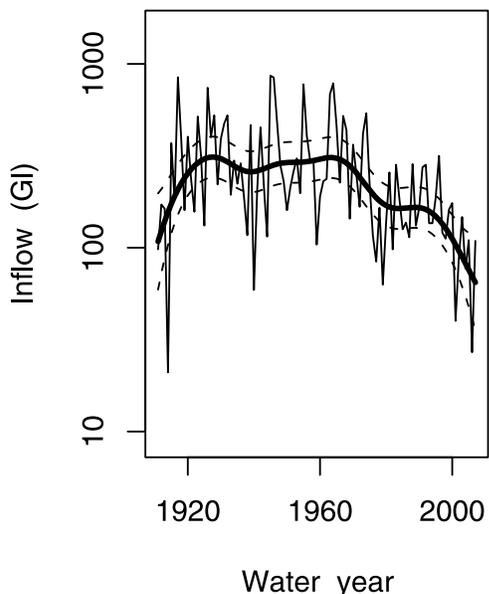
Figure 4. Local linear regression curve with $h = h_{cv} = 6.01$ (thick line) and variability bands (dashed lines) for log-transformed, aggregated annual inflow series. Water year is May to April.

carried out for three "seasons": May to October (MJJASO), May to July (MJJ), and August to September (ASO). The idea is that if a simulated trend is predominantly due to one, or a particular combination, of these predictors, then reversing the relevant input sequences should reverse the sign of the trend. In our experiment, each "factor" corresponds to one of the predictors and has two levels, coded as "forward" and "backward". For the "forward" level, the temporal ordering of the predictor is preserved. For the "backward" level the sequence of "seasons" is reversed, but the temporal ordering within each "season" is retained in recognition of the different mechanisms that dominate precipitation generation in early and late winter (section 2.1). There are 2^4 possible factor combinations in total: two for each of the four predictors. For each of these combinations, we have run three NHMM simulations to estimate the factorial effects with reasonable precision. We studied the effects upon several response variables representing different indices of change: first, the changes in the mean marginal probabilities of each of the weather states $[\overline{P}(S = j): j = 1,...N]$ between the periods 1958–1967 and 1998–2007, and second, the corresponding changes for the mean permanence probabilities $[\overline{P}(S_t = j|S_{t-1} = j)]$.

[35] The analysis of a factorial experiment is usually done via a linear model in which the coefficients (β) represent the effects of the factors individually and in combination [*Davison*, 2003, section 9.2]. For each season, we use half-normal plots [*Davison*, 2003, p. 444] of the coefficient estimates to show the combinations of atmospheric predictors with the strongest effects on the responses defined above (the half-normal distribution is the distribution of $|\beta|$ with $\beta \sim N(0, \sigma^2)$). If none of the combinations has any effect, the balanced design of the factorial experiment ensures that the coefficient estimates will behave like a random sample from a normal distribution centered on zero; hence they

will lie on an approximately straight line on a half-normal plot. Conversely, if a coefficient estimate appears as a substantial outlier on such a plot, it can be concluded that the corresponding combination of predictors is an influential driver of long-term changes in the weather state probability series of interest.

5. Results

5.1. Annual Inflow Series

[36] Inspection of Figure 1 indicates that the variability of the annual inflow series for the IWSS decreases with time. We therefore work with the logarithms of the inflows, to stabilize the variance as required by LLR. Figure 4 displays the log-transformed series, with the LLR trend estimate (bandwidth $h_{cv} = 6.01$) superimposed. Diagnostic checks (not shown) indicate that the model assumptions (independent, normally distributed errors with constant variance) are not unreasonable. Formal tests were carried out, as described in section 4.1.4, to determine whether the trend in log-transformed annual inflows can be regarded as linear. A significance trace (not shown) indicates that, for all reasonable values of the smoothing bandwidth h, the null hypothesis of a linear trend is convincingly rejected (p-values below 0.005). We conclude that the trend in log inflows is not linear.

[37] Figure 5 displays the significance trace for the *Bowman et al.* [2006] discontinuity test applied to the logarithms of the inflow series. For all values of h, the null hypothesis of no discontinuity is comfortably accepted (p-values above 0.05). This suggests that the decline in annual inflows is more consistent with a smooth trend than with an underlying step function.
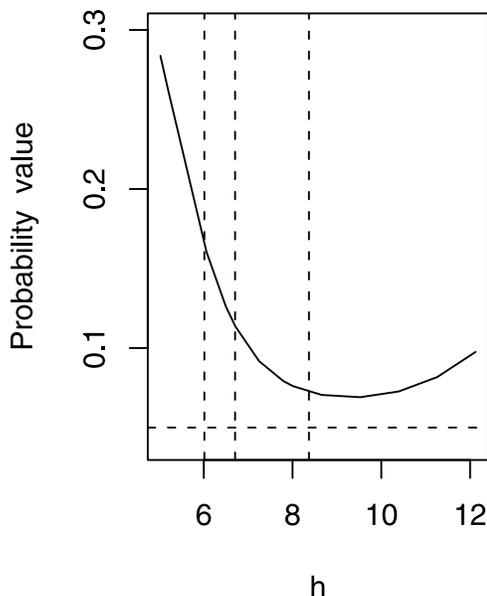

Figure 5. Significance trace to assess the evidence for the presence of discontinuities in the log-transformed, aggregated annual inflow series. Horizontal dashed line indicates the p-value 0.05, and dashed vertical lines indicate h_{aicc}, h_{cv}, and h_{df}.

Table 1. Characteristics of Weather States Obtained From the NHMM

Weather State	Spatial Precipitation Occurrence Pattern	Synoptic Situation[a]	Mean Probability		
			MJJASO	MJJ	ASO
1	Wet north and hinterland	Low pressure trough associated with midlevel moisture source or spring thunderstorms	0.08	0.07	0.09
2	Wet everywhere	Ridging low pressure system, frontal westerly winds	0.18	0.22	0.13
3	Wet west coast and central	Coastal convergence, light winds	0.18	0.20	0.17
4	Wet southwest corner	Postfrontal, southwesterly winds	0.21	0.21	0.21
5	Dry everywhere	High pressure system centered east of SWA, east to northeast winds	0.25	0.22	0.29
6	Wet south coast	Ridging high pressure system, moist southerly winds	0.10	0.08	0.11

[a]*Charles et al.* [1999a].

5.2. At-Site Precipitation

[38] For the generalized linear modeling of daily precipitation in the vicinity of the IWSS dams, the starting point was, separately for occurrence and amounts, to build a "baseline" model representing the overall intraseasonal and regional structure in the data. The model-building strategy followed that outlined by *Chandler* [2005], and was informed both by physical understanding and by examination of model diagnostics. The baseline models were then expanded by adding covariates representing temporal trends and their interactions. Two different trend representations were considered: the first was a linear function of time and the second was the trend estimate obtained for the inflow series logarithms (Figure 4). The results indicated overwhelmingly significant (p-value $< 10^{-8}$ according to dependence-adjusted likelihood ratio tests) trends in both precipitation occurrence and amounts. For all practical purposes, the fitted models were the same for both trend representations; we therefore focus on the linear version here, since it is easier to interpret. Diagnostic checks for the extended models were satisfactory.

[39] The extended occurrence model contained terms representing interactions between lagged occurrences and both seasonal and regional covariates; the amounts model contained interactions between seasonal and regional covariates. This suggests that persistence in precipitation occurrence is seasonally and regionally varying, and that intraseasonal variation in intensity is regionally varying. However, the main purpose of the models is to understand the nature of precipitation trends in the region. As discussed by *Chandler and Wheater* [2002], the effect of any covariate can be studied in detail by extracting all of the terms (including interactions) involving that covariate in the estimated linear predictors (3) and (4); the result is a function of other covariates. With time measured in decades, the respective contributions of the linear trend function to the linear predictors for occurrence and amounts are

$$[-0.0747 + (0.0312 \cdot \text{Long1}) + (0.0338 \cdot \text{Elevation})] \cdot \text{time} \quad (8)$$

and

$$[-0.0289 - (0.0222 \cdot \text{Lat1}) - (0.0498 \cdot \text{Long1})] \cdot \text{time}. \quad (9)$$

[40] Here, elevation is measured in hundreds of meters and "Lat1" and "Long1" are Legendre polynomials of degree 1 in latitude and longitude; these are linear functions taking the value −1 at the southern (respectively western) edge of the dashed rectangle in Figure 2 and increasing to +1 at the northern (respectively eastern) edge. Note that the models are intended to approximate the regional structure of changes only within this rectangle, and should not be taken as a basis for extrapolation over a wider area. At the 11 sites used for model fitting, the value of (8) ranges from −0.093 to +0.036. Unfortunately, this range is difficult to interpret directly because the occurrence model also contains covariates representing lagged occurrences: the modeled trend in precipitation occurrence thus has a direct component given by (8), along with an indirect component resulting from the trend in the lagged occurrences themselves. The net effect of these components can be calculated [*Chandler et al.*, 2011], although the technical details are well beyond the scope of the present paper. The qualitative impression from (8), that negative trends are strongest at locations with low values of Long1 and Elevation, i.e., in low-lying westerly locations, is correct, however.

[41] Fortunately, the final amounts model contained no terms representing previous days' rainfalls since these were found to be insignificant. Interpretation of (9) is much more straightforward therefore: since this is a contribution to the log mean wet day rainfall, exponentiating gives an average decadal change. At the sites used for model fitting, this change is negative everywhere except in the southwest corner of the dashed box in Figure 2: the largest change, at the easternmost site in this box, corresponds to a decrease of 6% per decade. Combining the modeled changes in both occurrence and amounts as *Chandler et al.* [2011] did, we find small increases (up to 0.7% per decade) at three sites but net decreases of up to 5% per decade at the remainder. In general, the fitted models suggest substantial decreases in precipitation, associated with changes in both occurrence and intensity, over most of the main water supply catchments. These results are consistent with, but add more detail to, the conclusions of previous studies reviewed in section 2.3. In particular, (8) and (9) for the first time provide insight into the spatial structure of regional precipitation trends.

5.3. Stochastic Downscaling

5.3.1. Model Application

[42] Table 1 provides qualitative descriptions of the spatial daily precipitation occurrence patterns and associated synoptic situations for the six states of the fitted NHMM, and lists their mean probability across 1000 simulations for the period 1958–2007. The patterns and situations are essentially the same as those depicted by *Charles et al.* [1999a, Figure 2], albeit for an earlier version of the model. The descriptions show that the weather states have a

Table 2. Summary of p-Values for Local Linear Regression Model Fits to Mean Marginal Probability Series for NHMM Weather States

Probability Series	Model Setting	Season								
		MJJASO			MJJ			ASO		
$P(S = 3)$	h	4.30	29.2	29.2	4.30	8.28	29.2	4.30	6.97	29.2
	H_0: null	0.099	0.015	0.015	0.004	0.0005	0.00005	0.510	0.562	0.700
	H_0: linear	0.517	1.00	1.00	0.381	0.290	0.527	0.428	0.452	0.766
$P(S = 4)$	h	0.988	2.13	4.3	1.72	3.55	4.30	4.30	29.2	29.2
	H_0: null	0.009	0.004	0.002	0.011	0.004	0.003	0.422	0.192	0.192
	H_0: linear	0.011	0.005	0.004	0.013	0.004	0.003	0.575	0.650	0.650
$P(S = 5)$	h	4.30	29.2	29.2	4.30	21.8	29.2	4.30	29.2	29.2
	H_0: null	0.434	0.056	0.056	0.015	0.0002	0.0002	0.905	1.00	1.00
	H_0: linear	0.885	0.388	0.388	0.424	0.174	0.164	0.826	0.847	0.847
$P(S = 6)$	h	4.30	25.3	29.2	1.92	4.30	20.1	4.30	29.2	29.2
	H_0: null	0.212	0.106	0.103	0.075	0.315	0.415	0.264	0.032	0.032
	H_0: linear	0.365	0.284	0.286	0.053	0.222	0.155	0.799	0.866	0.866

high degree of physical realism. States 2–5 are the most (and almost equally) frequent for each season considered. States 1 and 6 are the least frequent, each occurring about one third to one half as often as each of states 2–5.

[43] To explore changes in weather state occurrence during the 1958–2007 period, LLR-based trend tests were applied to annual series of mean marginal and permanence probabilities, for each of the "seasons" described in section 4.2.3. Table 2 shows the results for the weather state, season, $h = (h_{aicc}, h_{cv}, \text{and } h_{df})$, and reference model combinations that exhibit trends significant at the 0.05 level. The conclusions from this table are supported by plots of the data, trend estimates, and significance traces (not shown). There are decreasing linear trends in the state 3 probabilities over a wide range of h values for MJJASO and MJJ. The trend estimates for state 4 probabilities are nonmonotonic, with a higher frequency of occurrence since 1980. Over a range of values of h there is evidence of an increasing linear trend in state 5 probabilities for MJJ. States 1 and 2 are not included in Table 2 as their trends are not significant at the 0.10 level. For ASO, the only significant test result is for state 6, where there is some evidence for a decreasing linear trend. While

none of the permanence probabilities showed any significant trends at the 0.05 level, we note for example that all of the series $[\overline{P}(S_t = 5|S_{t-1} = j)]$, $j = 1,\ldots, 6$, for MJJ exhibit increasing trends.

[44] To summarize these results in the context of the overall precipitation decline in the region, the NHMM weather states have been aggregated on the basis of whether the corresponding spatial patterns of precipitation occurrence reflect "dry," "mixed," or "wet" conditions. A pattern was categorized as "mixed" if stations with a high probability of precipitation occurrence are concentrated in particular subregions of SWA. Thus state 5 was classified as "dry," states 2 and 3 were classified as "wet," and states 1, 4, and 6 were classified as "mixed" (Table 1). Figure 6 shows the mean marginal probability series for each of these aggregated classes for MJJ. Broadly, these show an increase in the frequency of "dry" conditions, a decrease in "wet" conditions, and little change in the frequency of "mixed" conditions. These conclusions are supported by the results of hypothesis tests (not shown). However, there is little evidence for trends in any of the aggregated series for ASO and hence MJJASO (p-values ≥ 0.10).
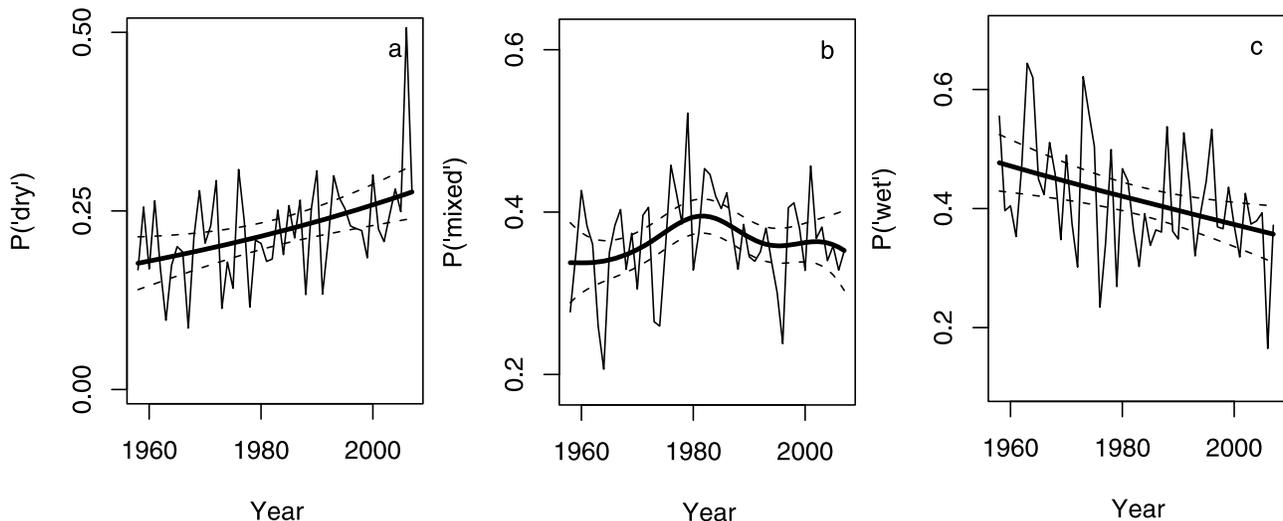

Figure 6. Local linear regression curves (thick lines) and variability bands (dashed lines) with $h = h_{cv}$ for aggregated weather state probability series (thin lines) for the MJJ season: (a) $P(\text{"dry"}) = \overline{P}(S = 5)$, (b) $P(\text{"mixed"}) = \overline{P}(S = 1) + \overline{P}(S = 4) + \overline{P}(S = 6)$, and (c) $P(\text{"wet"}) = \overline{P}(S = 2) + \overline{P}(S = 3)$.

5.3.2. Modeling of Atmospheric Predictors

[45] Exploratory analysis of the atmospheric predictor series produced a richer level of detail compared with previous efforts to characterize changes in regional atmospheric circulation (see section 2.3). Over the period 1958–2007 there have been slight increasing trends in MSLP and PG and stronger increasing trends in DTD and CV1; there is some suggestion that the trend in CV1 is nonlinear, but the departure from linearity is slight. There are also increasing trends in the interquartile ranges of the predictors suggesting that variability is increasing as well as the mean. Intraseasonally, there are increasing trends in the dates of the maxima for MSLP and for PG in particular. There is also an increasing date of minimum DTD, indicating that maximum atmospheric moisture at the 850 hPa level is tending to occur later in the winter half-year. Moreover, there are apparent linear relationships between the means and standard deviations for all four predictors, with the strongest being for DTD and CV1.

[46] Regression modeling was used to put the above results on a more formal footing. For each of the variables, covariates considered in the models included a linear trend function, sine and cosine functions to represent intraseasonal variability, lagged values to account for autocorrelation, and interactions. Furthermore, it is possible that trends in DTD and CV1 can be explained via the dependence of these variables upon MSLP and PG, which between them provide a crude summary of synoptic conditions over the region. To account for this, therefore, concurrent and lagged values of MSLP and PG were considered as potential covariates when modeling DTD and CV1; DTD was also considered as a potential covariate for CV1. Moreover, to account for the fact that MSLP and DTD are part of the definition of CV1 at (7), the quantity $0.04z(\text{MSLP}) + z(0.02DT_d^{850})$ was included as a fixed and known contribution to the model for this variable; such a contribution is referred to as an offset [e.g., *Davison*, 2003, p. 498]. On the basis of the exploratory analysis, all four variables were taken to be normally distributed but with nonconstant variance; the methods of section 4.1.3 were thus used to fit joint mean-variance models. In the first instance, for each variable the standard deviation was taken as a linear function of the mean. Subsequent diagnostics suggested that this gave an adequate representation of the variance structure for DTD and CV1, but not for the pressure variables. Therefore, extended models were subsequently fitted for MSLP and PG in which the standard deviation was explicitly represented as a linear function of the other covariates. As elsewhere, the choice of covariates to retain in each model was based on a combination of diagnostics, formal techniques, and physical understanding.

[47] Overall, our results confirmed that there have been statistically significant changes in the behavior of all four variables over the past 50 years. However, the complex structure of the models makes it difficult to infer the exact nature of these changes from the regression coefficients. Therefore, having identified where statistically significant changes have occurred, we return to the original data to illustrate them. Figure 7 presents the differences between the monthly means and standard deviations of each variable in the first and second halves of the data record. Six features are apparent. First, there is a very clear seasonal structure to the changes in MSLP, with increases in MJJ and a decrease in September. The increase in MSLP for MJJ in particular suggests either a decrease in the intensity of low pressure troughs or an increase in the intensity of high pressure systems. Second, mean PG has decreased in MJJ and increased in ASO, particularly during September. The decrease in PG in MJJ suggests a decrease in the strength of the prevailing westerly winds during the season and hence a reduction in moisture advection. Third, the increase in mean DTD during May to August indicates a reduction in the moisture content of the lower troposphere over the centre of the region during that part of the winter half-year: there is little change in September and October. Fourth, there has been a fairly uniform increase in CV1 between May and August, and little change in September and October. Fifth, the variability (standard deviations) of the predictors has increased over much if not all of MJJASO; according to the fitted models, these increases are significant at the 5% level for all variables except PG. Sixth, relative to the intraseasonal changes in the atmospheric predictors for MJJ, the corresponding changes for ASO are less conducive to drier conditions. This may explain why the observed decline in ASO precipitation is markedly less than that for MJJ [*Bates et al.*, 2008, Figure 4].

5.3.3. Sensitivity Analysis

[48] As reported above, the changes in NHMM weather state frequencies are consistent with the observed precipitation decline in the region. The factorial experiment described in section 4.2.3 enables us to determine the combinations of predictors that are driving these changes in the simulations. Figure 8 shows half-normal plots corresponding to each of the probability series with trends significant at the 0.05 level (Table 2). Overall five features are apparent. First, there are only three large effects: MSLP, DTD, and PG. None of the interaction effects is important. Thus each predictor has the same effect on simulated precipitation changes regardless of the state of the other predictors. Second, MSLP and DTD are the primary drivers of the decreasing trends in state 3 probabilities for MJJ and MJJASO (Figures 8a and 8c). This reflects the importance of changes in the intensities of low pressure (decrease) or high pressure (increase) systems and the reduction in moisture supply to precipitation generated by coastal convergence (Table 1). Third, MSLP is the primary driver of the increasing frequency of state 4 for MJJ and MJJASO (Figures 8b and 8d). This reflects the importance of changes in low and high pressure systems to the generation of postfrontal precipitation (Table 1). Further, PG is an important effect for MJJASO but not for MJJ (cf. Figures 8b and 8d). Thus the contrasting changes in PG and the strength of the prevailing westerly winds for the first and second halves of the winter half-year are needed to explain the trend for the entire winter half-year. Fourth, DTD closely followed by MSLP both control the increasing trend in state 5 probabilities for MJJ (Figure 8e). This reflects the importance of the reduction in moisture supply and an increase in the number of days with a high pressure system centered over the region (Table 1) to a reduction in precipitation occurrence. Finally, DTD and possibly PG are responsible for the comparatively mild decreasing trend in the frequency of state 6 for ASO (Figure 8f). Although the high probabilities of precipitation occurrence within state 6 are situated along the south coast (Table 1), local river flows have declined only slightly (J. K. Ruprecht, personal communication, 2009).
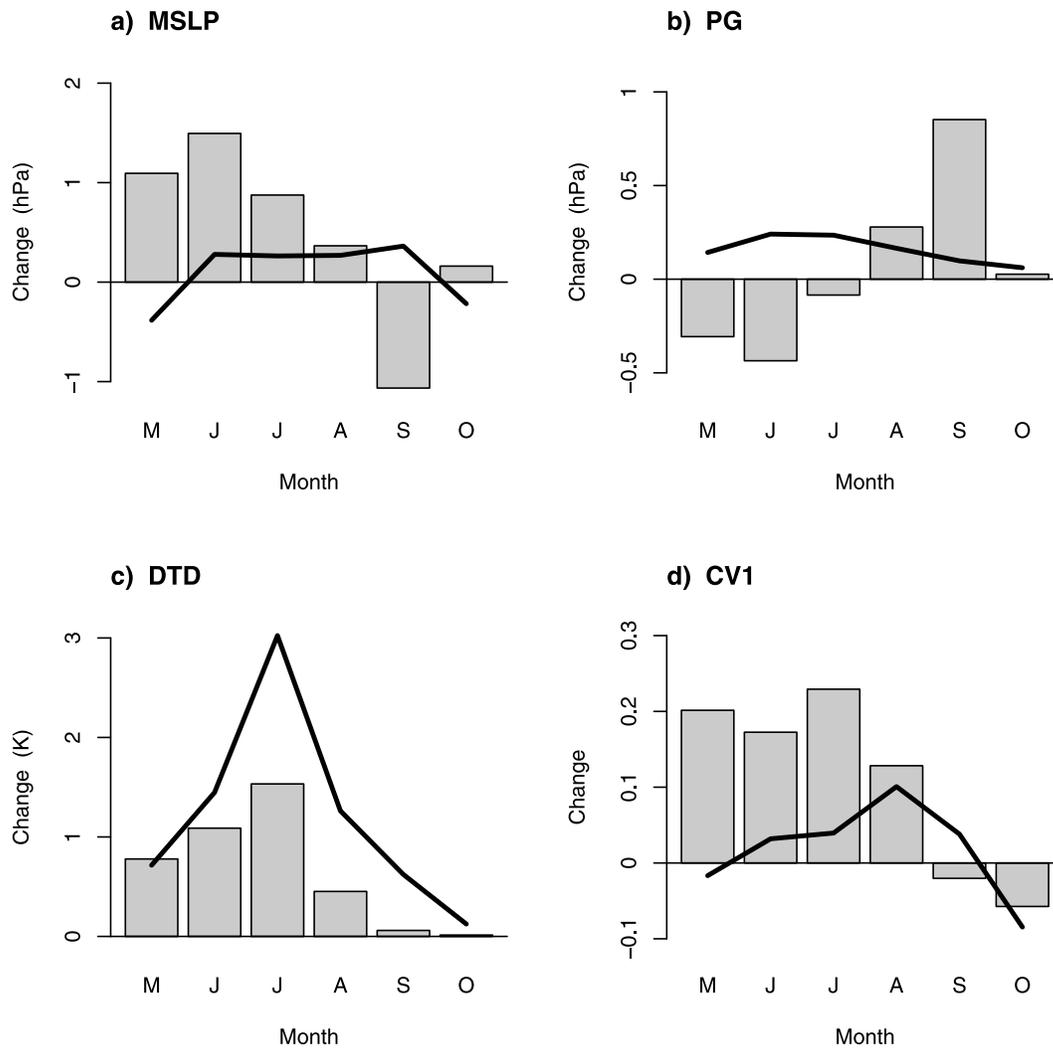
Figure 7. Changes in monthly means (bars) and standard deviations (lines) for each atmospheric predictor (1983–2007 versus 1958–1982).

6. Discussion and Conclusions

[49] A common and central assumption in the design and operation of surface water supply systems is the statistical stationarity of inflow series. Our study of the Integrated Water Supply Scheme (IWSS) was motivated by the need for a comprehensive understanding of a hydrologic series that appears to exhibit nonstationarity through a thorough investigation of the associated climatic drivers. Our research uses a comprehensive modeling approach to characterize the nonstationarity and to identify and describe possible causal pathways from changes in regional atmospheric circulation to changes in surface water availability; this provides a richer level of detail than previous studies. Some of our results are in agreement with what is already known (section 2.3), but many others relate to aspects that have not been previously or comprehensively investigated. They include the following: the nature of the trend in the annual inflow series for the IWSS; a thorough analysis of trends in both precipitation occurrence and amounts in the vicinity of the IWSS dams; a detailed analysis of trends in the frequencies and temporal persistence of weather types that are linked directly to spatial patterns of precipitation occurrence;

a comprehensive understanding of intraseasonal as well as interannual changes in indices of regional atmospheric circulation that are key to trends in spatial patterns of precipitation occurrence; and for the first time an assessment of the relative importance of the indices, and combinations of indices, in explaining the trends in the frequencies of weather types. Overall, our results indicate a more complicated reality than that suggested by previous studies. Owing to the paucity of long, high-quality precipitation records over the period of interest (1958–2007), however, we were unable to include precipitation-runoff modeling to assess the relative influences of climatic drivers and anthropogenic factors such as changes in land use and management. While this makes precise attribution difficult, the available evidence suggests that the impact of the rainfall decline is greater than the combined impact of other natural and anthropogenic factors [*Croton and Reed*, 2007].

[50] Our main findings in relation to the case study are as follows.

[51] 1. The decline in annual inflows for the 11 major dams in the Integrated Water Supply Scheme (IWSS) over the past three decades is more consistent with a smooth
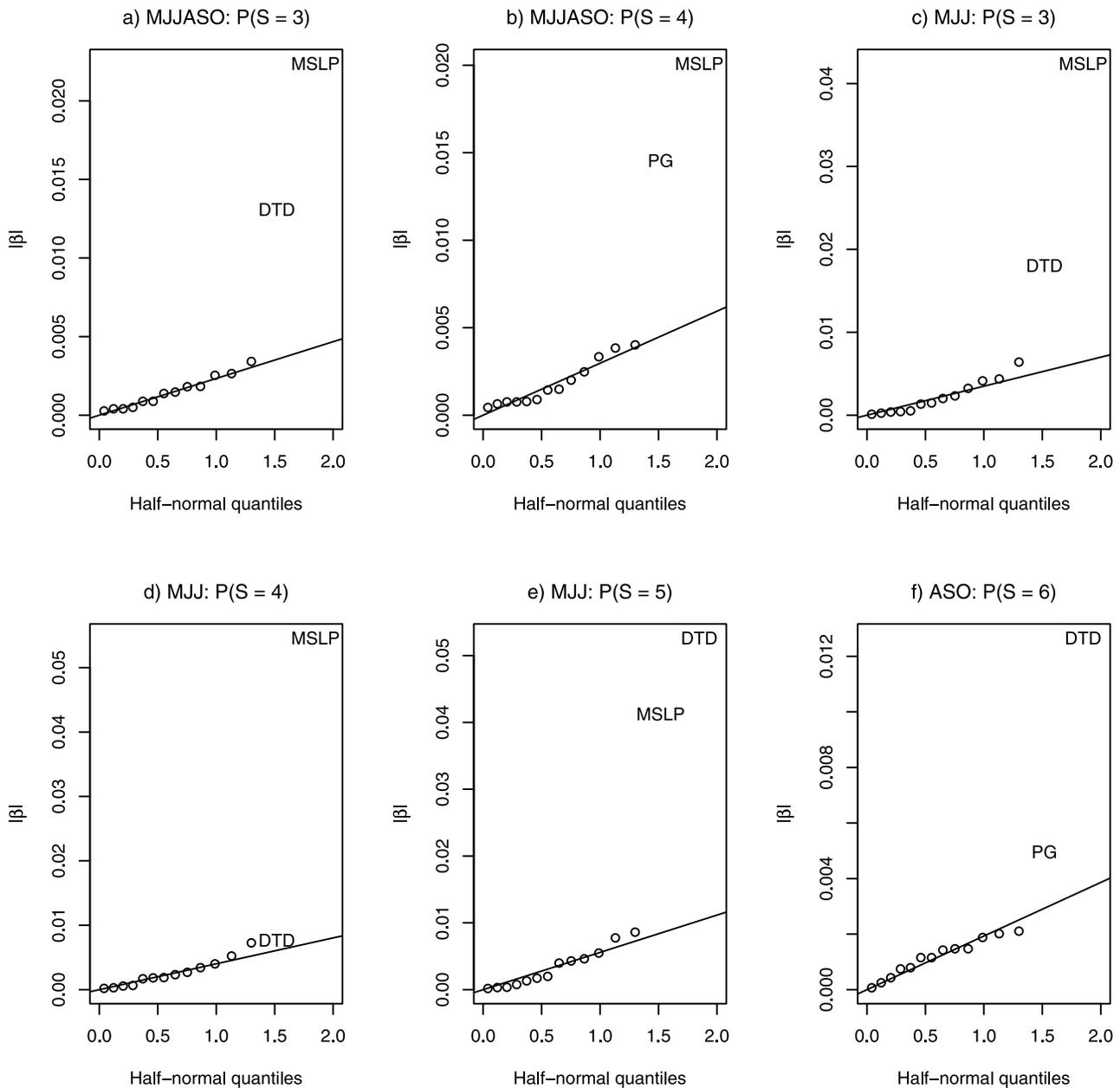
Figure 8. Half-normal plots of selected main and interaction effect estimates from the factorial experiment. Lines in each plot are least squares fits to the 12 lowest values of $|\beta|$ and the corresponding half-normal quantiles. MSLP, PG, and DTD denote estimated effects for mean sea level pressure, north-south SLP gradient, and dew point temperature depression at 850 hPa, respectively (see section 4.2.2 for further details).

nonlinear trend than with successive periods of relatively constant levels separated by abrupt changes. This conclusion, based as it is on a comparison of credible models for the inflow series rather than upon rejection of an implausible null hypothesis, challenges the conventional interpretation of the series. The apparent trend shown in Figure 4 has implications for future infrastructure investment.

[52] 2. Substantial decreases in precipitation, associated with changes in both occurrence and intensity, have occurred over most of the main water supply catchments. Combining the modeled changes in both occurrence and amounts, there have been net decreases of up to 5% per decade over the past half-century. Moreover, persistence in precipitation occurrence is seasonally and regionally varying, and intraseasonal variation in intensity is regionally varying.

[53] 3. During the first half of the winter half-year (May to July), there is evidence of an increasing trend in the probability of the weather type associated with dry conditions across the study region. This increase is accompanied by a decrease in the aggregated probabilities of weather types associated with widespread wet conditions across SWA and within the vicinity of the main dams of the IWSS in particular. These trends are strongly associated with a decrease in

atmospheric moisture and an increase in regionally averaged sea level pressure. There is little evidence of any changes in the persistence of the weather types.

[54] 4. Analysis of time series for four indices of regional-scale atmospheric circulation have revealed marked changes in their seasonality and variability. The changes for May to July (e.g., higher MSLP, lower north-south MSLP gradient (PG), and lower moisture in the lower troposphere (i.e., higher DTD)) are highly consistent with the dry conditions observed over the past 30 years. The changes for August to October are mixed (higher and lower MSLP in September and August, respectively, increased PG in August and September in particular, and lower atmospheric moisture in August but only marginal changes in September and October) and are consistent with previous analyses that indicate only comparatively small reduction in August to October precipitation. The variability of three out of the four indices (MSLP, DTD, and CV1) has increased over much if not all of the winter half-year.

[55] 5. The temporal orderings of three of the four indices examined (MSLP, PG, and DTD) have a demonstrable impact on trends in the weather type probability series. The effects of the 11 possible interaction terms between the indices on trends in the weather type probability series are small by comparison, suggesting that the effects of the indices can be considered individually, at least as far as the NHMM is concerned.

[56] 6. While the changes in regional atmospheric circulation and precipitation described herein may only partially contribute to the observed decline in dam inflow, the magnitudes of these changes are large and they must have a substantial impact.

[57] 7. The results of our study demonstrate the benefits of a multifaceted approach to the study of the nature and drivers of nonstationarity in hydrologic series. They were obtained through a collaborative research effort involving multiple disciplines (hydrology, climatology, and statistics), an approach foreshadowed by *Lettenmaier and Burges* [1978].

References

Adam, J. C., and D. P. Lettenmaier (2008), Application of new precipitation and reconstructed streamflow products to streamflow trend attribution in Northern Eurasia, *J. Clim.*, *21*, 1807–1828.

Allan, R. J., and M. Haylock (1993), Circulation features associated with the winter rainfall decrease in southwestern Australia, *J. Clim.*, *6*, 1356–1367.

Ansell, T. J., C. J. C. Reason, I. N. Smith, and K. Keay (2000), Evidence for decadal variability in southern Australian rainfall and relationships with regional pressure and sea surface temperature, *Int. J. Climatol.*, *20*, 1113–1129.

Bates, B. C., S. P. Charles, and J. P. Hughes (1998), Stochastic downscaling of numerical climate model simulations, *Environ. Modell. Software*, *13*, 325–331.

Bates, B. C., S. P. Charles, and E. P. Campbell (2001), Stochastic downscaling experiments for southwest Western Australia, in *Towards Understanding Climate Variability in South Western Australia: Research Reports on the Second Phase of the Indian Ocean Climate Initiative*, pp. 97–148, Indian Ocean Climate Initiative Panel, East Perth, Australia.

Bates, B., P. Hope, B. Ryan, I. Smith, and S. Charles (2008), Key findings from the Indian Ocean Climate Initiative and their impact on policy development in Australia, *Clim. Change*, *89*, 339–354, doi:10.1007/s10584-007-9390-9.

Bowman, A. W., and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*, 193 pp., Oxford Univ. Press, New York.

Bowman, A. W., and A. Azzalini (2007), *Applied Smoothing Techniques for Data Analysis: the Kernel Approach with S-Plus Illustrations*, version 2.2, Oxford Sci. Publ., Oxford, U. K. (Available at http://www.stats.gla.ac.uk/~adrian/sm.)

Bowman, A. W., A. Pope, and B. Ismail (2006), Detecting discontinuities in nonparametric regression curves and surface, *Stat. Comput.*, *16*, 377–390, doi:10.1007/s11222-006-9618-y.

Brennan, D. (2008), Will we all be rooned without a desal Plant? Hanrahan's lament and the problem of urban water planning under climate change, *Agenda*, *15*, 5–20. (Available at http://epress.anu.edu.au/agenda/015/03/pdf/15-3-AN-1.pdf.)

Burn, D. H., L. Fan, and G. Bell (2008), Identification and quantification of streamflow trends on the Canadian Prairies, *Hydrol. Sci. J.*, *53*, 538–549, doi:10.1623/hysj.53.3.538.

Cai, W. J., G. Shi, and Y. Li (2005), Multidecadal fluctuations of winter rainfall over southwest Western Australia simulated in the CSIRO Mark 3 coupled model, *Geophys. Res. Lett.*, *32*, L12701, doi:10.1029/2005GL022712.

Chandler, R. E. (2005), On the use of generalized linear models for interpreting climate variability, *Environmetrics*, *16*(7), 699–715.

Chandler, R. E., and S. M. Bate (2007), Inference for clustered data using the independence log-likelihood, *Biometrika*, *94*, 167–183.

Chandler, R. E., and H. S. Wheater (2002), Analysis of precipitation variability using generalized linear models: A case study from the west of Ireland, *Water Resour. Res.*, *38*(10), 1192, doi:10.1029/2001WR000906.

Chandler, R. E., B. C. Bates, and S. P. Charles (2011), Rainfall trends in southwest Western Australia, in *Statistical Methods for Trend Detection and Analysis in the Environmental Sciences*, chap. 7, edited by R. E. Chandler and E. M. Scott, Wiley, Chichester, in press.

Charles, S. P., B. C. Bates, and J. P. Hughes (1999a), A spatio-temporal model for downscaling precipitation occurrence and amounts, *J. Geophys. Res.*, *104*(D24), 31,657–31,669.

Charles, S. P., B. C. Bates, P. H. Whetton, and J. P. Hughes (1999b), Validation of a downscaling model for changed climate conditions in southwestern Australia, *Clim. Res.*, *12*, 1–14.

Charles, S. P., B. C. Bates, I. N. Smith, and J. P. Hughes (2004), Statistical downscaling of daily precipitation from observed and modelled atmospheric fields, *Hydrol. Process.*, *18*(8), 1373–1394.

Charles, S. P., M. A. Bari, A. Kitsios, and B. C. Bates (2007), Effect of GCM bias on downscaled precipitation and runoff projections for the Serpentine catchment, Western Australia, *Int. J. Climatol.*, *27*, 1673–1690, doi:10.1002/joc.1508.

Coe, R., and R. D. Stern (1982), Fitting models to daily rainfall, *J. Appl. Meteorol.*, *21*, 1024–1031.

Croton, J. T., and A. J. Reed (2007), Hydrology and bauxite mining on the Darling Plateau, *Restor. Ecol.*, *15*(Suppl.), S40–S47.

Davison, A. C. (2003), *Statistical Models*, Cambridge Univ. Press, Cambridge, U. K.

Dixon, H., D. M. Lawler, and A. Y. Shamseldin (2006), Streamflow trends in western Britain, *Geophys. Res. Lett.*, *33*, L19406, doi:10.1029/2006GL027325.

England, M. H., C. C. Ummenhofer, and A. Santoso (2006), Interannual rainfall extremes over southwest Western Australia linked to Indian Ocean variability, *J. Clim.*, *19*, 1948–1969.

Frederiksen, J. S., and C. S. Frederiksen (2007), Interdecadal changes in Southern Hemisphere winter storm track modes, *Tellus*, *59A*, 599–617.

Gentilli, J. (1972), *Australian Climate Patterns*, 285 pp., Nelson, Melbourne, Australia.

Greene, W. H. (2003), *Econometric Analysis*, 5th ed., 1026 pp., Pearson, Upper Saddle River, N. J.

Hennessy, K. J., R. Suppiah, and C. M. Page (1999), Australian rainfall changes, 1910–1995, *Aust. Meteorol. Mag.*, *48*(1), 1–13.

Hope, P. K., W. Drosdowsky, and N. Nicholls (2006), Shifts in the synoptic systems influencing south west Western Australia, *Clim. Dyn.*, *26*, 751–764.

Hughes, J. P., and P. Guttorp (1994), A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena, *Water Resour. Res.*, *30*, 1535–1546.

Hughes, J. P., P. Guttorp, and S. P. Charles (1999), A non-homogeneous hidden Markov model for precipitation occurrence, *Appl. Stat.*, *48*, 15–30.

Hurvich, C. M., J. S. Simonoff, and C.-L. Tsai (1998), Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion, *J. R. Stat. Soc., Ser. B*, *60*, 271–293.

Indian Ocean Climate Initiative (IOCI) (1999), *Towards Understanding Climate Variability in South Western Australia: Research Reports on the First Phase of the Indian Ocean Climate Initiative*, Bur. of Meteorol. Res. Cent., East Perth, WA, Australia. (Available at http://www.ioci.org.au/pdf/IOCI_FPR_1.pdf.)

Indian Ocean Climate Initiative (IOCI) (2002), *Climate Variability and Change in Southwest Western Australia*, Dep. of Environ. Water and Catchment Prot., East Perth, WA, Australia.

Intergovernmental Panel on Climate Change (IPCC) (2007), Climate Change 2007: Impacts, Adaptation and Vulnerability, in *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by M. L. Parry et al., 976 pp., Cambridge Univ. Press, Cambridge, U. K.

Jarušková, D. (1997), Some problems with application of change-point detection methods to environmental data, *Environmetrics*, *8*, 469–483.

Kahya, E., and S. Kalayci (2004), Trend analysis of streamflow in Turkey, *J. Hydrol.*, *289*, 128–144.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471.

Kendall, M., and J. K. Ord (1990), *Time Series*, 3rd ed., Edward Arnold, London.

Kistler, R., et al. (2001), The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and Documentation, *Bull. Am. Meteorol. Soc.*, *82*, 247–267.

Kundzewicz, Z. W., and A. J. Robson (2004), Change detection in hydrological records: A review of the methodology, *Hydrol. Sci. J.*, *49*, 7–19.

Lettenmaier, D. P., and S. J. Burges (1978), Climate change: detection and its impact on hydrologic design, *Water Resour. Res.*, *14*, 679–687.

Li, F., L. E. Chambers, and N. Nicholls (2005), Relationships between rainfall in the southwest of Western Australia and near-global patterns of sea-surface temperature and mean sea-level pressure variability, *Aust. Meteorol. Mag.*, *54*(1), 23–33.

Li, Y., W. Cai, and E. P. Campbell (2005), Statistical modeling of extreme rainfall in southwest Western Australia, *J. Clim.*, *18*, 852–863.

Mauget, S. A. (2003), Multidecadal regime shifts in U.S. streamflow, precipitation, and temperature at the end of the twentieth century, *J. Clim.*, *16*, 3905–3916.

Miller, W. P., and T. C. Piechota (2008), Regional analysis of trend and step changes observed in hydroclimatic variables around the Colorado River basin, *J. Hydrometeorol.*, *9*, 1020–1034, doi:10.1175/2008JHM988.1.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer (2008), Stationarity is dead: whither water management, *Science*, *319*, 573–574, doi:10.1126/science.1151915.

Murphy, B., and B. Timbal (2008), A review of recent climate variability and climate change in southeastern Australia, *Int. J. Climatol.*, *28*, 859–879, doi:10.1002/joc.1627.

Power, S., B. Sadler, and N. Nicholls (2005), The influence of climate science on water management in Western Australia, *Bull. Am. Meteorol. Soc.*, *86*, 839–844.

Simmonds, I., and K. Keay (2000), Variability of Southern Hemisphere extratropical cyclone behaviour, 1958–97, *J. Clim.*, *13*, 550–561.

Smith, I. (2004), An assessment of recent trends in Australian rainfall, *Aust. Meteorol. Mag.*, *53*(3), 163–173.

Smith, I. N., P. McIntosh, T.J. Ansell, and C. J. C. Reason (2000), Southwest Western Australian winter rainfall and its association with Indian Ocean climate variability, *Int. J. Climatol.*, *20*, 1913–1930.

Stewart, I. T., D. R. Cayan, and M. D. Dettinger (2005), Changes toward earlier streamflow timing across western North America, *J. Clim.*, *18*, 1136–1155.

Timbal, B. (2004), Southwest Australia past and future rainfall trends, *Clim. Res.*, *26*, 233–249.

Timbal, B., J. M. Arblaster, and S. Power (2006), Attribution of the late-twentieth-century rainfall decline in southwest Australia, *J. Clim.*, *19*, 2046–2062.

Wright, P. B. (1974), Seasonal rainfall in southwestern Australia and the general circulation, *Mon. Weather Rev.*, *102*, 219–232.

Yang, C., R. E. Chandler, V. S. Isham, and H. S. Wheater (2006), Quality control of daily observational rainfall series in the UK, *Water Environ. J.*, *20*, 185–193.

Zheng, H., L. Zhang, C. Liu, Q. Shao, and Y. Fukushima (2007), Changes in stream flow regime in headwater catchments of the Yellow River basin since the 1950s, *Hydrol. Process.*, *21*, 886–893.

B. C. Bates, Climate Adaptation Flagship, CSIRO Marine and Atmospheric Research, Wembley, WA 6014, Australia. (bryson.bates@csiro.au)

E. P. Campbell, Climate Adaptation Flagship, CSIRO Mathematics, Informatics and Statistics, Wembley, WA 6014, Australia.

R. E. Chandler, Department of Statistical Science, University College London, London WC1E 6BT, UK.

S. P. Charles, Climate Adaptation Flagship, CSIRO Land and Water, Wembley, WA 6014, Australia.