Aust. Met. Mag. 56 (2007) 93-103

A bivariate test for detecting inhomogeneities in pan evaporation time series

Dewi G.C. Kirono and Roger N. Jones
Climate Change Impacts and Risk, CSIRO Marine and Atmospheric Research, Australia

(Manuscript received April 2006; revised April 2007)

Homogeneous climatological time series are necessary for studying historical climate variability and change. This paper describes the use of the bivariate test for detecting and adjusting discontinuities in Class A pan evaporation time series for 28 stations across Australia, and illustrates the benefit of using corrected records in climate studies.

Ninety-two per cent of the inhomogeneities detected by the bivariate test are consistent with station metadata. Even though the test was designed to detect a single discontinuity in the mean, it can also be sensitive to multiple shifts in the mean. These show the suitability of the bivariate test as a tool for screening pan evaporation data. Having identified inhomogeneities, the adjustments were only applied to records which contained inhomogeneities that could be verified as having a non-climatic origin. The use of original and adjusted records in correlation analysis and in trend analysis produce very different conclusions. At Esperance and Woomera, for instance, unadjusted pan evaporation records do not correlate with potential evapotranspiration and are positively correlated with rainfall, whereas those of adjusted pan evaporation result in a more sensible inter-variable relationship (i.e. pan evaporation is negatively correlated with rainfall and is positively correlated with estimated potential evaporation). In a trend analysis, most unadjusted pan evaporation records show a statistically significant negative bias which, in most cases, is removed with adjustment. This is consistent with the effect of bird guard installation, early in the time series, that reduces pan evaporation. The trend in the original average of all stations was adjusted from -2.8 ± 1.7 to -0.7 ± 1.6 mm year^{-2} for 1970-2004, demonstrating the importance of screening the data before their use in climate studies.

Introduction

There is increasingly growing demand for evaporation data for studies of surface water and energy fluxes, especially for studies which address the impacts of global warming. Evaporation involves the transformation of water from its liquid state into a gas and the subsequent diffusion of water vapour into the atmosphere. However, the measurement of evaporation in the open environment is difficult and is usually done by proxy.

Corresponding author address: Dewi Kirono, CSIRO Marine and Atmospheric Research, Private Bag No 1, Aspendale, Vic. 3195, Australia.
Email: dewi.kirono@csiro.au

Potential evaporation is the variable most often used. Potential evaporation is a measure of the ability of the atmosphere to remove water from a surface assuming no limit to water availability, whereas actual evaporation is the quantity of water that is removed from that surface by evaporation (Brutsaert 1982). Therefore, actual evaporation is only equal to potential evaporation when a given surface is saturated.

The most widespread measurement method for potential evaporation uses a pan evaporimeter, which quantifies water loss from the instrument itself and not from the surrounding environment. The standard US Class A pan is the most commonly used instrument. It consists of a metal container usually covered by an open wire bird guard that is 1,207 mm across and 254 mm high. Evaporation is the amount of loss (gain) in mm depth with rainfall from an adjacent rain-gauge subtracted.

Pan evaporation records may contain many artefacts of measurement (inhomogeneities) caused by equipment changes, exposure changes and location changes (Jones 1992). More accurate estimates of potential evaporation can be obtained by applying other meteorological data to empirical, water budget, energy budget, and combination approaches. However, the most accurate approaches tend to be resource-intensive, site-specific and do not provide long-term estimates of change. Therefore pan evaporation records are the largest single source of data on historical evaporation trends.

A number of recent papers describe recorded trends in pan evaporation around the world, including Europe (Golubev et al. 2001), North America (Lawrimore and Peterson 2000), Australia (Roderick and Farquhar 2004) and New Zealand (Roderick and Farquhar 2005). Many of these trends show a decrease but interpreting and attributing these trends has been made difficult due to a lack of homogeneity within pan records. We maintain that, as for temperature and rainfall, reliable conclusions cannot be made about historical trends unless the data are shown to be largely free of inhomogeneities, particularly those that may bias regional signals.

Conrad and Pollak (1962) defined relative homogeneity for climate data as follows: ‘A climatological series is relatively homogeneous with respect to a synchronous series at another place if temperature differences (precipitation ratios) of pairs of homologous averages constitute a series of random numbers that satisfies the law of errors’. This permits the assumption that variations in average weather have similar tendencies over rather large regions (Karl and Williams 1987).

Testing the homogeneity of climatological time series is difficult because measurement errors and changes in observation site may be superimposed on the true climate variability. Therefore, it is impossible to verify that a climatological time series is truly homogeneous (Easterling and Peterson 1995). Ideally, inhomogeneity of data can be investigated from the history of the record (metadata). The metadata, however, may be unavailable. Even if it is available, the potential still exists that some undocumented station or instrument change may result in discontinuities. The alternatives in such a situation are to ignore the compromised part of the record or correct the record relative to another record that can be shown to be free of inhomogeneities.

Peterson et al. (1998) comprehensively reviewed a range of methods, approaches and philosophies involved in adjusting *in situ* climatological data. They categorised homogeneity testing techniques into direct and statistical approaches. The former uses metadata and/or side-by-side comparison of instruments, while the latter relies on the time series themselves. Even with the best possible metadata, however, some statistical inhomogeneity detection is advised (WMO 2003). In the statistical approach, the homogeneity test can either be based on the known data at the tested station alone, or a comparison with data from a reference station or series.

Numerous efforts to identify and/or adjust climatological time series have been conducted all over the world, with most generally focussed on rainfall and air temperature time series (e.g. Potter 1981; Buishand 1982; Bücher and Dessens 1991; Easterling and Peterson 1995; New et al. 1999; and New et al. 2000). In Australia, homogeneity testing has been underway since at least the late 1970s (Nicholls 1997). A high-quality historical rainfall dataset has been produced (Lavery et al. 1992; Lavery et al. 1997) along with a high-quality temperature dataset (Torok and Nicholls 1996; Trewin 1999; Della-Marta et al. 2004).

Unlike temperature and rainfall, pan evaporation studies in Australia are still relatively preliminary. An extended record of lake evaporation for western Victoria was constructed from homogenised climate records by Jones et al. (2001). At the time of the study there was no long-term pan evaporation record corrected for errors, but the National Climate Centre (NCC) is currently in the process of preparing a high-quality dataset for pan evaporation for the Australian region (Jovanovic et al. 2007). Until such a dataset is available, any analysis of historical change in pan evaporation is hindered unless one conducts homogeneity tests.

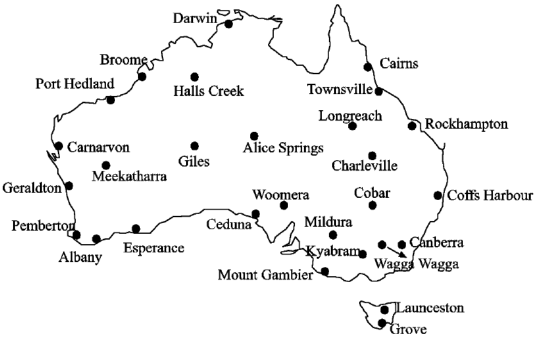
Here we describe the use of the bivariate test of Maronna and Yohai (1978), also known as the Potter method (WMO 2003), to detect inhomogeneities within the Australian record of Class A pan evaporation. Corrections were then applied to records where discontinuities were identified as a result of non-climatic origin with the aim of constructing a clean dataset suitable for climate studies. A description of the data, a short overview of the bivariate test, and the adjustment procedure are provided in the following section. The results are then presented along with an illustration of the implications of using the original and the adjusted records in climate studies.

Data and method

Data

Data used in this study include pan evaporation data, a number of climatic indices (see below), and station metadata. Monthly and annual pan evaporation data were obtained from the Australian Bureau of Meteorology archive (Climate Data: Australia, Version 2.2). The Bureau measures daily evaporation from standardised US Class A pans in a network established in the late 1960s that was relatively complete by the mid 1970s. Bird guards were placed in most of the stations in the late 1960s through early 1970s. According to van Dijk (1985), a bird guard can reduce pan evaporation by about seven per cent. About 250 stations are currently operating, but we only considered 28 stations with relatively long and complete records. These stations were inspected to determine their suitability for use in our study. Two basic requirements were necessary: records must be as long as possible (prior to 1970 to present) and must be as complete as possible (with less than one per cent missing data). Stations noticeably affected by an urban effect, such as Melbourne (86071), are also not considered. Figure 1 shows the location of the 28 stations analysed. Even though the distribution is relatively sparse and coastally biased, this is the most sensible distribution we can find that satisfies the requirements. Most of these stations have a complete record. We use linear regression to relate data from the few stations that do have gaps with data from the nearest neighbouring station (available from around 250 Bureau stations) in order to fill in the missing records. For example, Canberra (70014) has missing data in November 1980 and February 1991. In this case we filled in the missing data based on Canberra's relationship with Wagga Wagga (72150) which is located about 240 km west of Canberra. Both stations lie on the same climatic zone (as mapped by the Bureau), i.e. mild/warm summer and cold winter. The annual pan evaporation average in Canberra and Wagga Wagga is around 1600 mm and 1800 mm respectively.

Fig. 1 **Location of stations used in the study.**



Other selected climate records, i.e. rainfall, relative humidity and mean air temperature, for each station were also obtained from the Bureau archive. Most of these stations are included in the high-quality datasets of Lavery et al. (1997) (rainfall) and Della-Marta et al. (2004) (temperature). The all-Australian rainfall data series, developed based on the high-quality station network, was taken from the Bureau website (http://www.bom.gov.au/cgi-bin/silo/reg/cli_chg/timeseries.cgi). The estimated Morton's (1983) point potential evaporation for each station was also constructed to support the analysis as necessary. These estimates were developed from homogeneous climatological data series.

Station metadata were provided by the NCC (Jovanovic 2006, personal communication). These metadata summarise the known changes affecting pan evaporation measurement, their timing and comments for each station.

Overview of the bivariate test

The bivariate test, used by Potter (1981) for rainfall and Bücher and Dessens (1991) for temperature, aims to detect a single systematic change in mean in an independent time series, based on a second correlated series which is assumed to be stationary. The test not only provides knowledge of whether or not a change has occurred, but also provides maximum likelihood estimates as to the timing and the magnitude of the change, so that one may use the result to adjust the time series as needed. The test is relatively similar in approach to the Alexandersson (1986) technique. The differences are that Potter's test (1981) produces the magnitude of an inhomogeneity at the time of the break, whereas Alexandersson's test (1986) produces a change in ratio between the two series across the break.

The bivariate test assumes two serially independent sequences $\{x_i, y_i\}$ where $\{x_i\}$ is stationary. The statistic T_i, a measure of the difference between the two series at time i, is produced for all values over $i = 1, \ldots, i = \text{n-1}$. The statistic T_0 is the maximum value of T_i and is tested against levels of probability to test whether a significant change in mean has occurred (Potter 1981). T_0 occurs in the year before the change in mean and denotes a change in magnitude of D_0S_y (D_0 is the value of D_i, a statistic in test, when T_i equals T_0. S_y is the standard deviation; see Bücher and Dessens (1991)).

The assumptions of normality and independence may not always be satisfied in some hydro-climatological series. However, Potter (1981) has shown that the bivariate test performed very well in cases where data depart slightly from normal and where time series are not independent (e.g. auto-correlated rainfall records). Likewise, the assumption of stationarity is generally satisfied, particularly for an annual time series. If a time increment smaller than a year is required, then one may use one of the existing methods to remove seasonality. In addition, Potter also found that the results indicating inhomogeneities were consistent with station metadata, even in series where there are two or more shifts in the mean.

With regard to the length of the series, Vivès and Jones (2005) showed that the above robustness is conditional – the bivariate test is sensitive to deviations in mean near the beginning and the end of time series. Therefore, the test should be applied with caution in these situations. An alternative is to use metadata or another statistical technique to confirm the inhomogeneity. For this reason we also conducted another inhomogeneity test for each station using the RHTest of Wang and Feng (2004) applied with a test window of five years, which was used by the NCC to create a homogenised pan evaporation climatology (Jovanovic et al. 2007). The RHTest is based on the two-phase regression method introduced by Solow (1987) and later modified by Lund and Reeves (2002) and Wang (2003), and is designed to detect multiple step-change points in a record. Since the focus of this paper is to describe the bivariate test, the complete results of the RHTest are not presented, but are available from the authors. Overall, 58 per cent of the inhomogeneities detected by the bivariate test are also identified by the RHTest. For example, both tests identify the step-increase at Carnarvon in 1972 and the step-decrease at Longreach in 1972. The fact that the other 42 per cent of inhomogeneities are not consistently detected is due to different sensitivities of the two methods. In this case, the five-year window applied in the RHTest means that discontinuities within the first and last three years of the record may not be detected, and discontinuities occurring within three years of each other will be treated as one by the RHTest.

Test procedure

To satisfy the definition of relative homogeneity for climate data, it is advisable to test the candidate station against neighbouring stations or a regional mean (known as the reference series). This approach is often hindered by a lack of good and complete neighbouring data, and when the cause of that inhomogeneity occurs simultaneously across the region of interest, it cannot be detected using reference stations (Peterson et al. 1998). An alternative is to use a homogeneous reference series of a related variable that is closely correlated with the test variable.

In this study, the reference series was constructed from pan evaporation data of at least two adjacent sites with the highest correlation with the candidate series. The neighbouring station data were taken from the same source (i.e. the Bureau's full archive of around 250 stations) and were chosen to minimise the spatial distance from the candidate station. The correlation between the reference and the tested stations is normally higher than 0.6. Since the number of neighbours varies from one to another, the reference series also varies from one station to another.

In conducting each test, both the tested and reference series were standardised by their sample means and standard deviation, so that the critical values based on the null hypothesis explained in Potter (1981) could be used. A series is considered to be inhomogeneous when the change in mean is significant at least at the $p = 0.05$ level.

Adjustment procedure

Having identified discontinuities within the annual time series, both the annual and monthly time series were objectively adjusted. It must be noted that the adjustment was applied only to records which contain significant inhomogeneities with sufficient evidence showing that the discontinuity is non-climatic in origin. The supporting evidence was primarily sought from the available metadata. If the detected inhomogeneities were consistent with station metadata, adjustments were made. If the detected discontinuities were inconsistent with station metadata, following Jones (1995), a subsequent bivariate analysis was applied to other climate indices related to pan evaporation (i.e. mean air temperature, relative humidity and Morton's point potential evaporation) to check whether the discontinuity is also found in their records. In this case, we assumed that: (a) the phenomenon that could cause a simultaneous climatically consistent change in all variables would probably be a site change or other major change in the site environment, which would not go unrecorded; and (b) if there is an abrupt change in regional climate patterns causing a significant shift in pan evaporation, it is assumed that such a shift would also occur in related climatic indices.

The adjustments were made on a *pro rata* basis such that a single annual percentage change was applied evenly across all months. The potential disadvantage of this approach is that it may produce monthly biases since changes in the nature of the site (e.g. bird guard installation) may produce a stronger effect in windier months. However, van Dijk's (1985) analysis did not indicate that seasonal effects were present. Our test on Esperance station (not shown here) also indicates that the adjusted series based on a *pro rata* basis and on a monthly basis appears to be very similar, and both of them are homogeneous.

After the adjustments were made, we reran the bivariate analysis to test the adjusted series for homogeneity.

Results and discussion

Detection of inhomogeneity

As previously described, the statistic T_0, used to test whether or not a discontinuity in mean has occurred, is the maximum value of T_i, over all $i < n$. To illustrate this, Fig. 2 shows T_i, of selected series for which T_0 is significant at the 0.05 critical level and it occurs at the time when a bird guard was installed. In the plots of Esperance (9789) and Longreach (36031), the T_0 appears in the year preceding the documented change, whereas at Ceduna (18012) the T_0 occurs in the year following the documented change. Figure 3 shows T_i of selected series for which T_0 is significant but the time of the inhomogeneity is not near the documented bird guard installation. In the case of Rockhampton (39083), the bivariate test detects a discontinuity that is two years ahead of documented bird guard installation. In the case of Mildura (76031), the bivariate test detects a significant discontinuity in the year preceding the pan relocation (in 1989) but not the bird guard installation (in 1972). In the plots of T_i for Geraldton (8051), there are multiple peaks, which are all significant. The primary peak occurs in 1978 (a few years preceding some building alterations in 1983) while the secondary peak occurs at a time corresponding to the year when the bird guard was installed in 1969. This suggests that, although it was originally developed to detect a single discontinuity in the mean, the bivariate test can be sensitive to multiple shifts in the mean as has been demonstrated by Potter (1981).

The complete results of the bivariate tests along with documented changes at each station are summarised in Table 1. There are twenty-four stations that showed significant jumps (T_0) and 92 per cent of

Fig. 2 **Plots of T_i for selected stations for which the discontinuity (T_o) is significant and close to the time of the bird guard installation. The dotted line represents the 0.05 critical level, and the vertical lines mark the times when the bird guards were installed.**

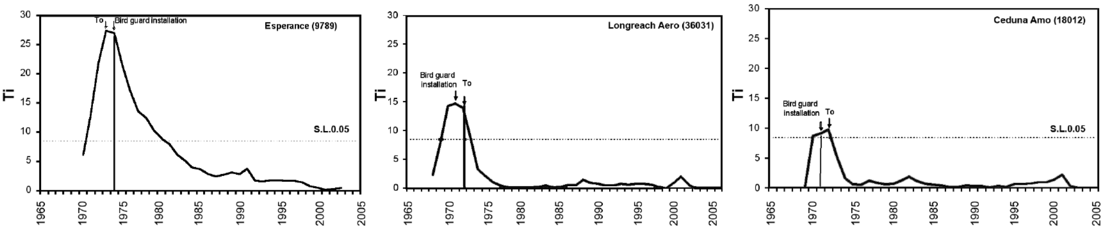


Fig. 3 **Plots of T_i for selected stations for which the discontinuity (T_o) is significant but is not near the time of the bird guard installation. The dotted line represents the 0.05 critical level, and the vertical lines mark the times when the bird guards were installed.**

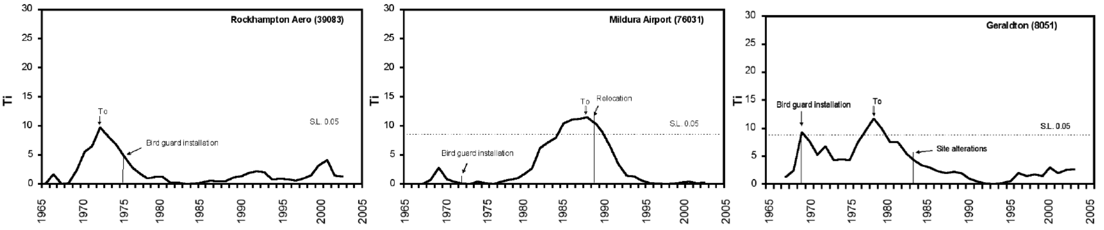


Table 1. Significant year of and estimated change in mean annual pan evaporation detected by the bivariate test for each station along with associated metadata.

Bureau site number	*Station*	*Period*	*Detected year and amount of change*			*Metadata*
			Year	*Change (mm)*	*Change (%)*	
2012	Halls Creek Airport	1969-2004	1992	-369	-11	BG* installed (1968)
						Problems with thermo. (1990)
						Dust in air from Mt Pinatubo (1991)
3003	Broome Airport	1967-2004	1967	-269	-10	BG installed (1968)
4302	Port Hedland Airport	1968-2004	1972	-372	-11	BG installed (1968)
						New float gauge (1972)
6011	Carnarvon Airport	1966-2004	1972	248	9	BG installed (1968)
7045	Meekatharra Airport	1967-2004	1974 a)	-485	-12	BG installed (1968)
			1981	-449	-12	Relocation (1974)
8051	Geraldton Airport	1967-2004	1969 a)	-339	-13	BG installed (1969)
			1978	-182	-7	Building alterations (1983)
9592	Pemberton	1968-2004	Homog	–	–	BG installed (1968)
9741	Albany Airport	1968-2004	1983	80	5	BG installed (1969)
						Pan replaced, site changed (1983)
9789	Esperance	1970-2004	1973	-459	-26	BG installed (1974)
13017	Giles Meteorological Office	1968-2004	Homog	–	–	BG installed (1967)
14015	Darwin Airport	1967-2004	1994	-289	-11	BG installed & site reloc. (1976)
						Enclosure changed (1990)
						AWS (1995)
15590	Alice Springs Airport	1967-2004	1984	410	12	BG installed (1966)
						Relocation (1974)
						BG wire broken (1985)
16001	Woomera Aerodrome	1967-2004	1978	355	11	BG installed (1970)
						Site moved (1980)
18012	Ceduna AMO	1969-2004	1972	-275	-12	BG installed (1971)
26021	Mount Gambier Aero	1967-2004	1972	-120	-9	BG installed (1973)
31011	Cairns Aero	1967-2004	1972	-303	-13	BG installed (1975)
32040	Townsville Aero	1970-2004	Homog	–	–	BG installed (1970)
36031	Longreach Aero	1968-2004	1971	-695	-22	BG installed (1972)
39083	Rockhampton Aero	1965-2004	1972	-247	-11	BG installed (1975)
			1992 b)	172	8	Instrument enclosure moved (1993)
44021	Charleville Aero	1968-2004	1972	-403	-15	Pan replaced (1973)
						BG installed (1974)
48027	Cobar MO	1969-2004	1972	-534	-22	BG installed (1972)
59040	Coffs Harbour MO	1968-2004	1974 a)	-288	-6	BG installed (1974)
			1991	-290	-9	New float assembly (1992)
70014	Canberra Airport	1967-2004	Homog	–	–	BG installed (1967)
72150	Wagga Wagga AMO	1967-2004	1971	-370	-20	BG installed (1972)
76031	Mildura Airport	1967-2004	1988	-174	-8	BG installed (1972)
						Relocation (1989)
80091	Kyabram (Inst. Sust. Ag.)	1967-2004	1985	-108	-7	BG installed (1978)
91104	Launceston Airport	1967-2004	1973	-150	-11	BG installed (1975)
94069	Grove (comparison)	1967-2004	1984	-110	-11	BG installed (1975)
						Relocation (1988-1990)

* BG= bird guard

a) shows a secondary, but significant, shift detected by the bivariate test

b) shows a discontinuity detected by the rerun of the bivariate test

those are associated with a documented change that could have been the cause of the discontinuity. As anticipated, stations with a bird guard installed on or before the beginning of the test period are free of inhomogeneities. The only exceptions are stations experiencing multiple changes, such as Port Hedland (4032) and Alice Springs (15590). The former had a new float gauge in 1972 while the latter had its bird guard wire broken in 1985; the bivariate test successfully detected these documented changes.

Of the twenty documented bird guard installations within the period of the series, twelve of them are successfully identified by the bivariate test. The other eight discontinuities related to bird guard installation are undetected mostly due to the existence of other documented changes such as instrument and site changes occurring after the bird guard installation (e.g. Port Hedland, Meekathara, Albany, Darwin, Alice Springs, Mildura and Grove). This finding suggests that, in the cases of multiple documented changes, the bivariate test tends to detect the largest jump as the primary discontinuity. Therefore, even though it was originally developed to detect a single discontinuity in the mean, the bivariate test can be used to screen a series with multiple discontinuities.

The only discontinuities identified by the bivariate test not supported by any documented change were a step-increase in Carnarvon (6011) in 1972 and a step-decrease in Kyabram (80091) in 1985 (Table 1). These discontinuities are also identified by the RHTest of Wang and Feng (2004) (not shown here), therefore the results obtained from the bivariate test are thought to be valid. To further investigate whether these are true changes or if the inhomogeneities are due to artificial, unrecorded causes, we followed the approach of Jones (1995), and tested the homogeneity of the other related climatic indices at Carnarvon and Kyabram. It is assumed that a simultaneous climatically consistent change in all variables could arise from a site change or major change in the site environment, but this would not be expected to go unrecorded. The other assumption is that a genuine climate change causing a significant increase in pan evaporation would be expected to occur in related climatic indices. Although a climate shift which only affects a single station is less probable, it is possible. Carnarvon, for instance, is an isolated station located on the coast where its climate may be dominated by local effects (Trewin, 2007, personal communication). Its mean monthly temperatures, for instance, have an insignificant correlation (correlation coefficients of 0.18) with those at Gascoyne Junction (6022) which is located about 155 km further inland.

The principal factors affecting pan evaporation are available energy, atmospheric humidity, and wind speed, hence the analysis is preferably conducted on those variables. Unfortunately, the wind speed observations (3 m height) for Carnarvon are only available from 1993 to the present, while those for Kyabram have a large gap from 1990 to 2000. We analysed the series of mean air temperature (as a surrogate of available energy), relative humidity (as a measure of atmospheric humidity) and Morton's point potential evaporation (as a measure of potential evaporation). For Carnarvon, the correlation coefficients of air temperature anomaly, relative humidity, and Morton's point potential evaporation, with the pan evaporation are +0.50, -0.60 and +0.55 respectively, while at Kyabram, these are +0.18, -0.66 and +0.61 respectively.

Bivariate testing of mean air temperature from Carnarvon against mean air temperature from a homogeneous regional reference series showed a significant discontinuity of +0.65°C in 1971. The test on relative humidity showed a significant discontinuity of -3.8% in 1972. Morton's point potential evaporation at Carnarvon tested against regional reference of potential evaporation also showed a discontinuity of +112 mm in 1971. These results provide evidence that the increase in pan evaporation at Carnarvon in 1972 is a climatic shift; therefore, no adjustment was conducted.

In the case of Kyabram, the bivariate test on temperature, relative humidity, and Morton's point potential evaporation showed no significant breaks, suggesting that the decrease in pan evaporation at Kyabram in 1985 is likely to be a non-climatic change affecting the pan only; therefore, an adjustment was conducted.

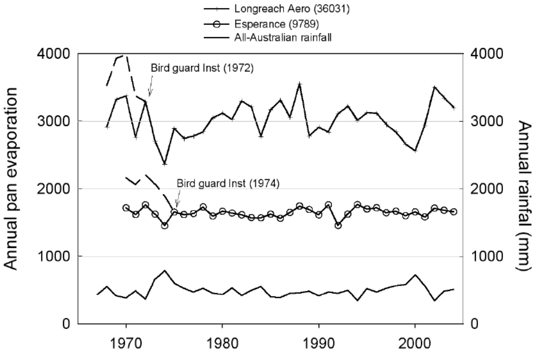
Adjustment

After all stations were screened, the detected discontinuities were then corrected using the approach described in the previous section. The corrections were applied only where inhomogeneities were identified as having a non-climatic origin. The corrections applied to the original records were mostly in the range -9 to -26% (-13.5% on average) of the mean annual pan evaporation, implying that a uniform correction of about -7% due to bird guard installation (van Dijk 1985) may not always hold. The results here show that such changes are not uniform between sites and are generally underestimated. The only positive corrections were those applied to Albany and Woomera, both as the result of a site movement.

Figure 4 shows plots of the original and adjusted series for Esperance and Longreach as examples. The all-Australian annual rainfall series, representing independent information, is also plotted. Apparently, the records of pan evaporations prior to the bird guard installation were biased towards a high value, and the adjustment seems to be quite reasonable. The interannual variations of the adjusted record closely mirror those of rainfall. The correlation coefficients between pan evaporation and all-Australian rainfall for Longreach were -0.60 (based on original data) and -0.75 (based on the adjusted data), whilst for Esperance they were -0.02 (based on original data) and -0.42 (based on adjusted data). Inspection and a rerun of the bivariate test on the adjusted series also confirm that the new series is homogeneous.

Retesting of all adjusted series with the bivariate test confirms that the new series are free from inhomogeneity. The only exception was for Rockhampton

Fig. 4 **Annual pan evaporation adjusted series for selected stations and plots of the all-Australian annual rainfall. The original pan evaporation series for each station is shown as a dashed line.**

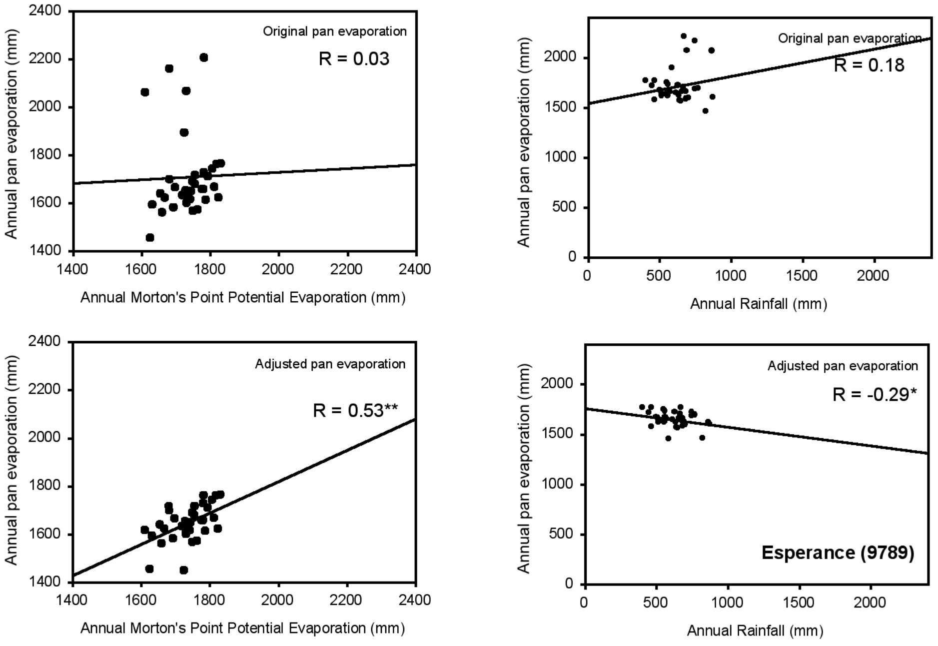


where a bird guard was installed in 1975 and the bivariate test initially detected an inhomogeneity in 1972 (see Table 1). After the adjustment was made, a rerun of the bivariate test still detected an inhomogeneity in 1992, the year preceding the instrument enclosure movement in 1993; therefore, a second adjustment was made for this station. Another rerun of the bivariate analysis then showed the series free of inhomogeneity.

Implications for climate studies

The implications of using the original and adjusted pan evaporation records in a climate study are investigated from two points of view: correlation analysis and trend analysis. The implications for correlation analysis are demonstrated by Fig. 5 which illustrates the relationship between pan evaporation and other climate indices (rainfall and estimated potential evaporation) at Esperance. In theory, measured pan evaporation and estimated potential evaporation should be positively and strongly related; however, Fig. 5 shows that the use of uncorrected pan evaporation data leads to a poor relationship, contrary to the accepted theory. When the corrected pan evaporation data are applied, however, the result is more consistent with the underlying logic. The relationship between pan evaporation and rainfall would be expected to be negative, because higher rainfall is associated with greater cloud cover (less available energy) and a higher atmospheric moisture content (less atmospheric demand) resulting in lower potential evaporation. Figure 5 demonstrates that one would have observed a weak positive relationship (based on the uncorrected data) instead of a

Fig. 5 **Relationship between pan evaporation and other climate indices (rainfall and Morton's estimated point potential evaporation) based on the original and adjusted pan evaporation data at Esperance. ** and * denote correlation coefficients which are significant at p<0.05 and p<0.1, respectively.**



strong negative relationship (based on the adjusted data). This is also observed, for instance, in the case of Woomera aerodrome where the uncorrected data series result in a correlation coefficient between pan evaporation and rainfall of +0.26 while using the corrected series results in a value of -0.68.

Implications for trend analysis are illustrated by calculating the linear trend of the original and adjusted pan evaporation data series. The results are summarised in Table 2. To obtain a fair comparison, the trends are calculated for the same period for all stations, i.e. 1970–2004. Thus, the table shows only stations where data later than 1970 were adjusted. The table shows distinct discrepancies between the trends calculated using the original and the corrected data series. Most of the trends calculated using the unadjusted data tend to be strongly negative. This is consistent with the effect of bird guard installation (almost all of which occurred in the first five years of the period) causing a reduction in pan evaporation (see van Dijk 1985) and producing a negative trend. As an illustration, the trends at Cobar are -10.3 ± 4.9 mm year^{-2} and -4.1 ± 4.5 mm year^{-2} for the original and adjusted series, respectively. In this case, the trend direction is consistent, but the trend magnitude is lessened and its statistical significance is reduced. Table 2 also indicates approximately half of the stations show large changes in trend direction, trend magnitude (beyond the margin of error) and trend significance when the trend analyses from the original and adjusted series are compared. For example, the original time series at Esperance shows a trend of -8.4 ± 2.5 mm year^{-2}, which is significant at the 0.05 level, while the adjusted series shows a trend of 0.9 ± 1.3 mm year^{-2}, which is not significant at the 0.05 level. This difference reflects bias in the earlier part of the record due to bird guard installation.

Table 2. Trends in pan evaporation for selected series based on the original and adjusted data series. The standard error is shown in brackets. ** and * denote trends which are significant at $p < 0.05$ and $p < 0.1$, respectively.

Station no.	*Station name*	*Trend in 1970–2004 (mm year^{-2})*	
		Original data	*Adjusted data*
2012	Halls Creek Airport	-9.8(5.4)*	3.4 (4.5)
3003	Broome Airport	-8.5 (2.8)**	-
4032	Port Hedland Airport	-3.0 (4.3)	2.7 (3.4)
6011	Carnarvon Airport	1.8 (1.7)	-
7045	Meekatharra Airport	-18.3 (3.5)**	-10.2 (3.9)**
8051	Geraldton Airport	-3.4 (2.5)	-
9592	Pemberton	-5.4 (1.2)**	-
9741	Albany Airport	4.9 (1.2)**	2.2 (1.1)*
9789	Esperance	-8.4 (2.5)**	2.8 (1.2)**
13017	Giles Meteorological Office	-0.4 (5.3)	-
14015	Darwin Airport	-12.1 (1.9)**	-1.4 (1.7)
15590	Alice Springs Airport	13.1 (7.0)*	8.5 (6.2)
16001	Woomera Aerodrome	21.5 (3.6)**	5.8 (5.1)
18012	Ceduna AMO	-1.6 (2.9)	2.5 (2.3)
26021	Mount Gambier Aero	-1.3 (1.0)	0.3 (1.1)
31011	Cairns Aero	0.5 (2.6)	4.7 (2.2)**
32040	Townsville Aero	-7.7 (3.3)**	-
36031	Longreach Aero	-1.0 (5.3)	6.5 (4.4)
39083	Rockhampton Aero	2.3 (2.9)	1.3 (2.2)
44021	Charleville Aero	0.1 (4.6)	-
48027	Cobar MO	-10.3 (4.9) **	-1.2 (4.5)
59040	Coffs Harbour MO	-13.4 (1.4)**	-0.45 (1.3)
70014	Canberra Airport	0.9 (2.7)	-
72150	Wagga Wagga AMO	-2.7 (2.7)	1.9 (2.7)
76031	Mildura Airport	-5.7 (3.0)*	2.0 (2.5)
80091	Kyabram (Ins. Sus. Ag.)	-4.8 (2.0)**	0.1 (1.7)
91104	Launceston Airport	-3.5 (1.5)**	-0.8 (1.4)
94069	Grove (comparison)	-3.5 (1.3)**	1.1 (1.1)

Figure 6 shows the all-sites time series, constructed by averaging the unweighted values of all 28 stations, along with the all-Australian high-quality rainfall time series, for 1970-2004. Note the caveat in having a simple unweighted all-stations average, particularly using a network that is relatively sparse. However, this exercise is not intended to create a national climatology but to show the danger of using unscreened datasets in a trend analysis.

The interannual variability of pan evaporation mirrors that of rainfall: the higher the rainfall the lower the pan evaporation, and vice versa. This figure also indicates that the use of adjusted time series produces a very different result compared to that of the original inhomogeneous time series.

Table 3 summarises trends in the all-Australian pan evaporation calculated using the original and adjusted time series. In our analysis, the original data show a trend of -2.8 ± 1.7 mm year^{-2}, whereas the adjusted dataset results in a not statistically significant trend of -0.7 ± 1.6 mm year^{-2}. In Roderick and Farquhar's (2004) analysis, the original data yielded a statistically significant trend of -4.3 ± 1.8 mm year^{-2}, whereas the adjusted dataset results in a trend of only -2.9 ± 1.7 mm year^{-2}. The homogenised data of Jovanovic et al. (2007), on the other hand, yield a very slight negative trend of 1.97 mm year^{-2} which is not statistically significant. The dissimilarities among trends obtained in different studies are likely due to different time periods, station selection, and methods of detection and adjustment. Our analysis uses data

Table 3. Trends in the all-Australian pan evaporation according to different studies. Note: in all studies, the all-station average was calculated using a simple unweighted average method.

Study	*Data period*	*Number of stations*	*Trends (mm/y^2)*	
			Original data	*Adjusted data*
Our analysis	1970–2004	28	-2.8±1.7	-0.7±1.6
Roderick and Farquhar (2004)	1970–2002	30	-4.3±1.8	-2.9±1.7
Jovanovic et al. (2007)	1970–2005	60	–	-1.97

for the period 1970-2004 and is averaged from unweighted values from 28 stations; Roderick and Farquhar's (2004) analysis is based on data from 1970 to 2002 and is averaged from unweighted values from 30 sites; whereas Jovanovic et al.'s (2007) analysis uses data from 1970 to 2005 and is averaged from unweighted values from 60 sites.

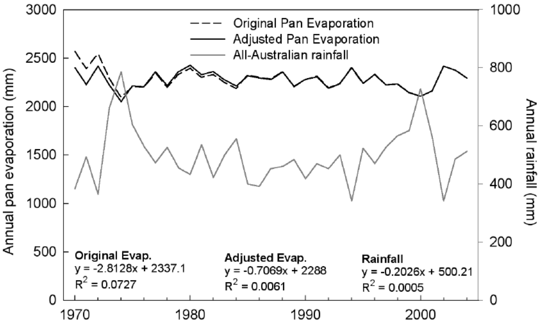
These results suggest that using inhomogeneous time series in trend analysis may produce a very different result than using homogeneous time series. In a forthcoming paper we present a more detailed trend analysis of the adjusted pan data on an annual and seasonal basis and compare this with trends in model estimates of potential evaporation.

Concluding remarks

In this paper, we describe the application of the bivariate test for detecting and adjusting discontinuities in a Class A pan evaporation time series from 28 stations across Australia. The results show:

- even though the bivariate test was designed to detect a single discontinuity in the mean, it can be sensitive to multiple shifts in the mean;
- in the presence of multiple changes, the bivariate test tends to detect the larger change as the primary break and identify the smaller change as the secondary break;
- 92 per cent of the breaks detected by the bivariate test are associated with documented changes that can be strongly associated with the discontinuity;
- most of the discontinuities were in one direction (downward), thereby biasing regional trends produced from unadjusted data;
- with the support of other independent evidence such as available metadata, the results from the bivariate test on other climate indices related to pan evaporation, and the results from inhomogeneity testing on pan evaporation using different techniques (i.e. RHTest), we were able to objectively adjust the records which contained inhomogeneities due to non-climatic causes.

Fig. 6 Time series of pan evaporation (before and after adjustment) averaged over 28 sites and the all-Australian high-quality rainfall dataset for 1970-2004. The equations for the linear trends are also shown.



There are distinct discrepancies between the results analysed using the original and those analysed using the corrected dataset. The results show that the trend (1970-2004) in the all-stations average was -2.8±1.7 for the original data but only -0.7±1.6 mm year^{-2} for the adjusted data. This demonstrates the importance of carefully scrutinising the data because of the risk in making far-reaching conclusions based on data that contains significant biases.

Finally, this study has made it possible to analyse historical pan evaporation data for comparison with modelled potential evaporation values (described in our forthcoming paper). This is important for constructing an evaporation climatology and future evaporation projections for Australia (CSIRO 2001). Future work, such as a comparison between the corrected data obtained in this study and those obtained in the independent study of NCC (Jovanovic et al. 2007) and an investigation of inhomogeneities in daily pan evaporation data is anticipated.

Acknowledgments

We appreciate the information provided by Dr David Jones and Dr Dean Collins of the Australian National Climate Centre (NCC) on the development of a high-quality pan evaporation dataset for Australia. Dr Branislava Jovanovic kindly provided metadata. Dr Ian Smith, Dr David Rayner, Dr Blair Trewin, and two anonymous reviewers are gratefully thanked for their valuable comments and suggestions for improving the manuscript. This study was conducted as part of the CSIRO Water for Healthy Country Flagship project (Project 1: Australia-wide climate baselines; and Project 2: Climate change impacts on rainfall and evaporation) and ACCSP project (Component 1: Understanding the key drivers of climate change in Australia).

References

Alexandersson, H. 1986. A homogeneity test applied to precipitation data. *J. Climatol., 6*, 661-75.

Bücher, A. and Dessens, J. 1991. Secular trend of surface temperature at an elevated observatory in the Pyrenees. *Jnl climate, 4,* 859-68.

Buishand, T.A. 1982. Some methods for testing the homogeneity of rainfall records. *J. Hydrol., 58*, 11-27.

Brutsaert, W. 1982. *Evaporation into the atmosphere: Theory, history, and application*. D. Reidel, 299 pp.

Conrad, V. and Pollak, C. 1962. *Methods in Climatology*. Harvard University Press, Cambridge, MA, 459 pp.

CSIRO 2001. *Climate change projections for Australia*. CSIRO Atmospheric Research: Melbourne.

Della-Marta, P.M., Collins, D.A. and Braganza, K. 2004. Updating Australia's high-quality annual temperature data set. *Aust. Met. Mag., 53*, 75-93.

Easterling, D.R. and Peterson, T.C. 1995. A new method for detecting undocumented discontinuities in climatological time series. *Int. J. Climatol., 15*, 369-77.

Golubev, V.S., Lawrimore, J.H., Groisman, P.Y., Speranskaya, N.A., Zhuravin, S.A., Menne, M.J., Peterson, T.C. and Malone, R.W. 2001. Evaporation changes over the contiguous United States and the former USSR: A reassessment. *Geophys. Res. Lett., 28*, 2665-68.

Jones, F.E. 1992. *Evaporation of Water: With Emphasis on Applications and Measurements*. Lewis, Chelsea, Michigan, 188 pp.

Jones, R.N. 1995. Modelling hydrologic and climatic controls of closed lakes, western Victoria. Unpublished PhD Dissertation, University of Melbourne.

Jones, R.N., McMahon, T.A. and Bowler, J.M. 2001. Modelling historical lake levels and recent climate change at three closed lakes, western Victoria, Australia (c. 1840-1990). *J. Hydrol., 246*, 158-79.

Jovanovic, B., Jones, D.A. and Collins, D. 2007. A high quality pan evaporation dataset for Australia. *Climatic Change*. In press.

Karl, T.R. and Williams, C.N. Jr 1987. An approach to adjusting climatological time series for discontinuous inhomogeneities. *Jnl Appl. Met., 27,* 1744-63.

Lavery, B.M., Kariko, A.P. and Nicholls, N. 1992. A historical rainfall data set for Australia. *Aust. Met. Mag., 40*, 33-9.

Lavery, B., Joung, G. and Nicholls, N. 1997. An extended high-quality historical rainfall dataset for Australia. *Aust. Met. Mag., 46,* 27-38.

Lawrimore, J.H. and Peterson, T.C. 2000. Pan evaporation trends in dry and humid regions of the United States. *J. Hydrometeorol., 1*, 543-6.

Lund, R. and Reeves, J. 2002. Detection of undocumented changepoints: A revision of the two-phase regression model. *Jnl climate, 15*, 2547-54.

Maronna, R. and Yohai, V.J. 1978. A bivariate test for the detection of a systematic change in mean. *J. Amer. Statis. Assoc., 73*, 640-5.

Morton, F.I. 1983. Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. *J. Hydrol., 66,* 1-76.

New, M., Hulme, M. and Jones, P.D. 1999. Representing twentieth-century space-time climate variability. Part I: Development of a 1961-90 mean monthly terrestrial climatology. *Jnl climate, 12*, 829-56.

New, M., Hulme, M. and Jones, P.D. 2000. Representing twentieth-century space-time climate variability. Part I: Development of 1961-90 monthly grids of terrestrial surface climate. *Jnl climate, 13,* 2217-38.

Nicholls, N. 1997. Developments in climatology in Australia: 1946-1996. *Aust. Met. Mag., 46,* 127-35.

Peterson, T.C., Easterling, D.R., Karl, T.R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullet, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Forland, E., Hanssen-Bauer, I., Alexandersson, H., Jones, P., Parker, D. 1998. Homogeneity adjustments of in situ atmospheric climate data: a review. *Int. J. Climatol., 18,* 1493-1517.

Potter, K. 1981. Illustration of a new test for detecting a shift in mean in precipitation series. *Mon. Weath. Rev., 109*, 2040-5.

Roderick, M.L. and Farquhar, G.D. 2004. Changes in Australian pan evaporation from 1970 to 2002. *Int. J. Climatol., 24,* 1077-90.

Roderick, M.L. and Farquhar, G.D. 2005. Changes in New Zealand pan evaporation since the 1970s. *Int. J. Climatol., 2*5, 2031-9.

Solow, A.R. 1987. Testing for climate change: an application of the two-phase regression model. *Jnl Clim. Appl. Met., 26*, 1401.

Torok, S.J. and Nicholls, N. 1996. A historical annual temperature data set for Australia. *Aust. Met. Mag., 45*, 251-60.

Trewin, B.C. 1999. The development of a high-quality daily temperature dataset for Australia, and implications for the observed frequency of extreme temperatures. *Proc. Sixth Nat. AMOS Conf.* Canberra, ACT, Australia, 87p.

van Dijk, M.H. 1985. Reduction in evaporation due to the bird screen used in the Australian Class A pan evaporation network. *Aust. Met. Mag., 33*, 181-3.

Vivès, B. and Jones, R.N. 2005. Detection of abrupt changes in Australian decadal rainfall (1890-1989). *CSIRO Atmospheric Research Technical Paper 73.*

Wang, X.L. and Feng, Y. 2004. *RHTest User manual*. Available from http://cccma.seos.uvic.ca/ETCCDMI/RHTestUserManual.doc

Wang, X.L. 2003. Comments on "Detection of undocumented changepoints: A revision of the two-phases regression model". *Jnl climate, 16*, 3383-5.

WMO 2003. Guidelines on climate metadata and homogenization. *WCDMP-No. 53, WMO-TD No. 1186*, World Meteorological Organization, Geneva, Switzerland, 51 pp.