# A New Approach to Stochastically Generating Six-Monthly Rainfall Sequences Based on Empirical Mode Decomposition

THOMAS A. MCMAHON
*Department of Civil and Environmental Engineering, University of Melbourne, Victoria, Australia*

ANTHONY S. KIEM
*Sinclair Knight Merz, Armadale, Victoria, Australia*

MURRAY C. PEEL
*Department of Civil and Environmental Engineering, University of Melbourne, Victoria, Australia*

PHILLIP W. JORDAN
*Sinclair Knight Merz, Armadale, Victoria, Australia*

GEOFFREY G. S. PEGRAM
*Civil Engineering Programme, University of KwaZulu-Natal, Durban, Republic of South Africa*

(Manuscript received 5 November 2007, in final form 17 April 2008)

ABSTRACT

This paper introduces a new approach to stochastically generating rainfall sequences that can take into account natural climate phenomena, such as the El Niño–Southern Oscillation and the interdecadal Pacific oscillation. The approach is also amenable to modeling projected affects of anthropogenic climate change. The method uses a relatively new technique, empirical mode decomposition (EMD), to decompose a historical rainfall series into several independent time series that have different average periods and amplitudes. These time series are then recombined to form an intradecadal time series and an interdecadal time series. After separate stochastic generation of these two series, because they are independent, they can be recombined by summation to form a replicate equivalent to the historical data. The approach was applied to generate 6-monthly rainfall totals for six rainfall stations located near Canberra, Australia. The cross correlations were preserved by carrying out the stochastic analysis using the Matalas multisite model. The results were compared with those obtained using a traditional autoregressive lag-one [AR(1)], and it was found that the new EMD stochastic model performed satisfactorily. The new approach is able to realistically reproduce multiyear–multidecadal dry and wet epochs that are characteristic of Australia's climate and are not satisfactorily modeled using traditional stochastic rainfall generation methods. The method has two advantages over the traditional AR(1) approach, namely, that it can simulate nonstationarity characteristics in the historical time series, and it is easy to alter the decomposed time series components to examine the impact of anthropogenic climate change.

## 1. Introduction

Many water resources agencies now require stochastically generated data to supplement a system simulation analysis that is based on a single realization of the historical record. In the case of the multireservoir water supply system in Canberra Australia, some of us were invited to provide generated sequences of seasonal 6-monthly rainfalls at six locations within the Canberra catchments. In a separate project, the 6-monthly generation model would be developed further to incorporate potential evapotranspiration, and the generated sequences of rainfall and evaporation would be disaggregated into daily values for input into a rainfall–runoff model. This paper is the result of the first phase of the combined project, that is, the generation of 6-monthly rainfall totals at six locations near Canberra.

As part of the specification of the project, we were required to take into account recent research that has

*Corresponding author address:* Thomas A. McMahon, Department of Civil and Environmental Engineering, The University of Melbourne, VIC, 3010, Australia.
E-mail: t.mcmahon@civenv.unimelb.edu.au

DOI: 10.1175/2008JHM991.1

shown that climatic patterns in eastern Australia are strongly influenced by at least two natural phenomena, the El Niño–Southern Oscillation (ENSO) and the interdecadal Pacific oscillation (IPO). ENSO refers to the anomalous warming (El Niño) and cooling (La Niña) that periodically (every 1–3 yr) occurs in the central and eastern tropical Pacific Ocean as a result of the Southern Oscillation [refer to Philander (1990) or Diaz and Markgraf (2000) for a comprehensive description of ENSO and a review of research into its causes and effects]. The IPO is defined by low-frequency (15–30 yr) anomalous warming and cooling between the tropical and extratropical Pacific Ocean (Power et al. 1999; Folland et al. 1999; Allan 2000) and is similar to the Pacific decadal oscillation (PDO; Mantua et al. 1997). El Niño years are typically dry in Australia, and La Niña years are usually wet, with the correlations between the ENSO signal and eastern Australian precipitation being some of the strongest globally. Compounding this is the IPO, which modulates both the magnitude and frequency of ENSO effects on multidecadal time scales (Power et al. 1999; Kiem et al. 2003).

Flood and drought dominated the sequences of streamflow—and by inference, rainfall—have been observed in New South Wales (Erskine and Warner 1988). Although this observation has been challenged by Kirkup et al. (1998), any model should be programmed to recognize the potential drivers of pseudocyclicity in rainfall sequences. Peel et al. (2004) argued that, on the basis of runs of annual rainfalls below the median, most regions in Australia can be adequately represented by an autoregressive lag-one [AR(1)] model, although the method adopted is considered to be a simplistic representation of reality. Not withstanding the latter comments and whether the runs of high and low rainfall are statistically significant, it is known that traditional AR(1) models are unable to simulate the long runs of wet and dry sequences that are observed in the historical record, for example, an 11-yr run (1997–2007) of below-median annual rainfall for Melbourne, Australia, which is unprecedented in the 151 yr of record.

In Australia, it is now a requirement that all investigations into future water supply must consider the effect of potential climate change in a region (e.g., Erlanger and Neal 2005). In addition to the need to generate long runs of seasonal (i.e., 6 months) rainfall that incorporate natural variability on intradecadal and interdecadal time scales (i.e., long runs of high and low rainfalls similar to what has occurred in the past), it was also important that the stochastic model to be used in the project be amenable to incorporating projections of how anthropogenically induced climate change may affect future rainfall sequences.

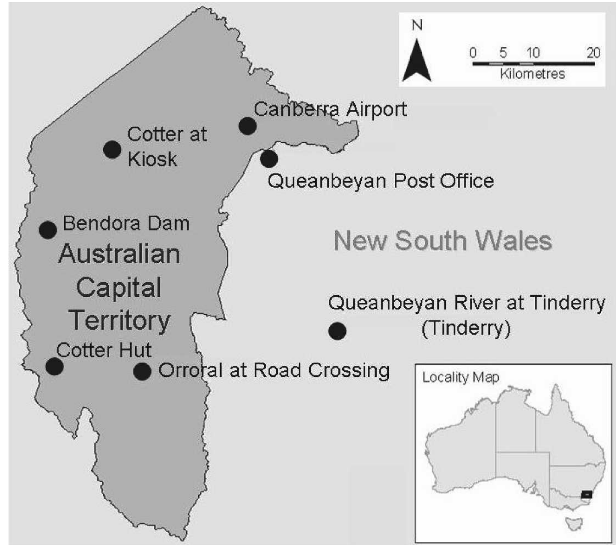


FIG. 1. Location of rainfall stations.

Following this introduction, we describe in section 2 the characteristics of the historical rainfall data used herein. A relatively new technique in hydrology known as empirical mode decomposition (EMD; Huang et al. 1998; Peel et al. 2005), which is the basis of the proposed stochastic model and recently introduced as a tool in hydrologic analysis, is outlined in section 3. Section 4 shows by using a computer experiment how a modification can be made to the traditional AR(1) stochastic model to generate wet and dry spells that are longer, and therefore more realistic, than those generated by the traditional model. In section 5, justification of the adopted approach is discussed. EMD analysis is applied in section 6 to data from the six rainfall stations in the Canberra catchments, and the outputs are used to generate 6-monthly rainfalls. The results of the EMD stochastic model are assessed in section 7 by comparing them with those from a traditional Matalas multisite autoregressive lag-one model (Matalas 1967). Some conclusions are recorded in section 8.

## 2. Historical rainfall data

Six rainfall stations [Canberra Airport (570926), Cotter Hut (570946), Cotter at Kiosk (570956), Bendora (570958), Tinderry (570965) and Orroral at Road Crossing (570966, designated as Orroral)] are used in this analysis. The location of these stations is shown in Fig. 1; continuous data are available from June 1871 to May 2006 (135 yr). It should be noted, however, that

TABLE 1. Statistical characteristics of annual rainfall for Canberra Airport and Orroral at Road Crossing stations based on 135 yr of data. Here, $\mu$ is mean, Cv is coefficient of variation, $\gamma$ is coefficient of skewness, $\rho$ is lag-one serial correlation, $a$ is annual, $w$ is winter–spring, and $s$ is summer–autumn.

| | Annual | | | | Winter–spring | | | Summer–autumn | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Rainfall stations | $\mu_a$ (mm) | $Cv_a$ | $\gamma_a$ | $\rho_a$ | $\mu_w$ (mm) | $Cv_w$ | $\gamma_w$ | $\mu_s$ (mm) | $Cv_s$ | $\gamma_s$ |
| Canberra Airport | 598 | 0.27 | 0.70 | 0.13* | 299 | 0.34 | 0.26 | 299 | 0.40 | 1.03 |
| Orroral at Road Crossing | 660 | 0.27 | 0.41 | 0.18 | 343 | 0.34 | 0.29 | 317 | 0.40 | 0.86 |

* Not different from zero at the 5% level of significance based on Yevjevich (1972).

62% (averaged across the six sites) of the data are infilled, of which half is by direct correlation with the Queanbeyan Post Office rain gauge (570911). The infilling was performed based on regression relationships developed between each of the sites using monthly data and then applying the "best" regression relationships on a daily basis to infill, and/or extrapolate, missing data. The quality of the relationships was generally poorer prior to 1917—the lowest coefficient of determination used prior (after) to 1917 was 0.46 (0.57). Further details on the infilling methodology, the regression relationships, and the amounts of the missing data are available in Sinclair Knight Merz (2004).

The statistical characteristics of the annual and 6-monthly rainfall series for two of the six rainfall stations are presented in Table 1. The values are typical of the parameters across the study area. In terms of stochastic modeling, two features are noted: 1) the autocorrelations at the annual level ($\rho_a$) are relatively high, presumably partly driven by ENSO; and 2) the relatively high coefficient of skewness for the summer–autumn 6-month rainfalls ($\gamma_s = 1.03$) suggests a normal distribution would be inappropriate.

In the analyses that follow, we have used the Canberra Airport and Orroral rainfall stations to illustrate the results. Canberra Airport station was chosen because it is a key station in the network and together with Orroral, the stations are located at the northern and southern extremities, respectively, of the study area.

## 3. Empirical mode decomposition

EMD analysis (Huang et al. 1998) is a form of adaptive time series decomposition that can handle both nonlinear and nonstationary time series. This is a major advantage over other decomposition techniques, such as Fourier analysis or wavelets (Huang et al. 1999; Torrence and Compo 1998), when the rainfall series that are being analyzed exhibit nonlinear and nonstationary characteristics such as amplitude and frequency modulation with time. EMD allows us to compute the proportion of variation in a time series that can be attributed to fluctuations (both low and high frequency) at different time scales. In the analysis presented here, we have adopted, because it was required by the project specification, a time interval of 6 months; in the analysis described later, we have available 135 yr of 6-monthly rainfalls, that is, time sequences of 270 values at each of the six stations. As described in section 6, seasonality is taken into account by considering separately each 6-month series through a multisite procedure.

Using the EMD procedure, fluctuations within a time series are automatically and adaptively selected, which results in the time series of observed data being decomposed into a set of independent intrinsic mode functions (IMFs), or time series, and a residual (trend) component. Figure 2, for example, is an analysis of the 6-month data for the Canberra Airport rainfall station and shows the observed record disaggregated into six IMFs and a trend component. If the IMFs and residual time series are summed together, they recombine to form the original time series. Theoretically, the individual IMFs and the residual are orthogonal—hence, uncorrelated—a characteristic that will be used in the stochastic data generation model being proposed. The residual might be a constant, a monotonic trend, or an incomplete ($\leq$3 extrema) fluctuation with an average period longer than the period of record. An analysis of IMF time series allows us to compute the amount of variance that is associated with each IMF and the residual along with the average period of each IMF. Theoretically, the variance of the IMFs and residual should sum to the variance of the original time series, or a value close to the observed variance when the IMF covariance terms are small. Details of the EMD methodology used in this analysis can be found in Peel et al. (2005) and, more generally, in Huang et al. (1998).

In the proposed stochastic rainfall generation model, which we have designated as the EMD stochastic model, we assume (taking the Canberra Airport data as an example calculation) that (1) the first three IMFs (average period of each $<$10 yr) are associated with intradecadal drivers (e.g., ENSO), and (2) IMF4–IMF6

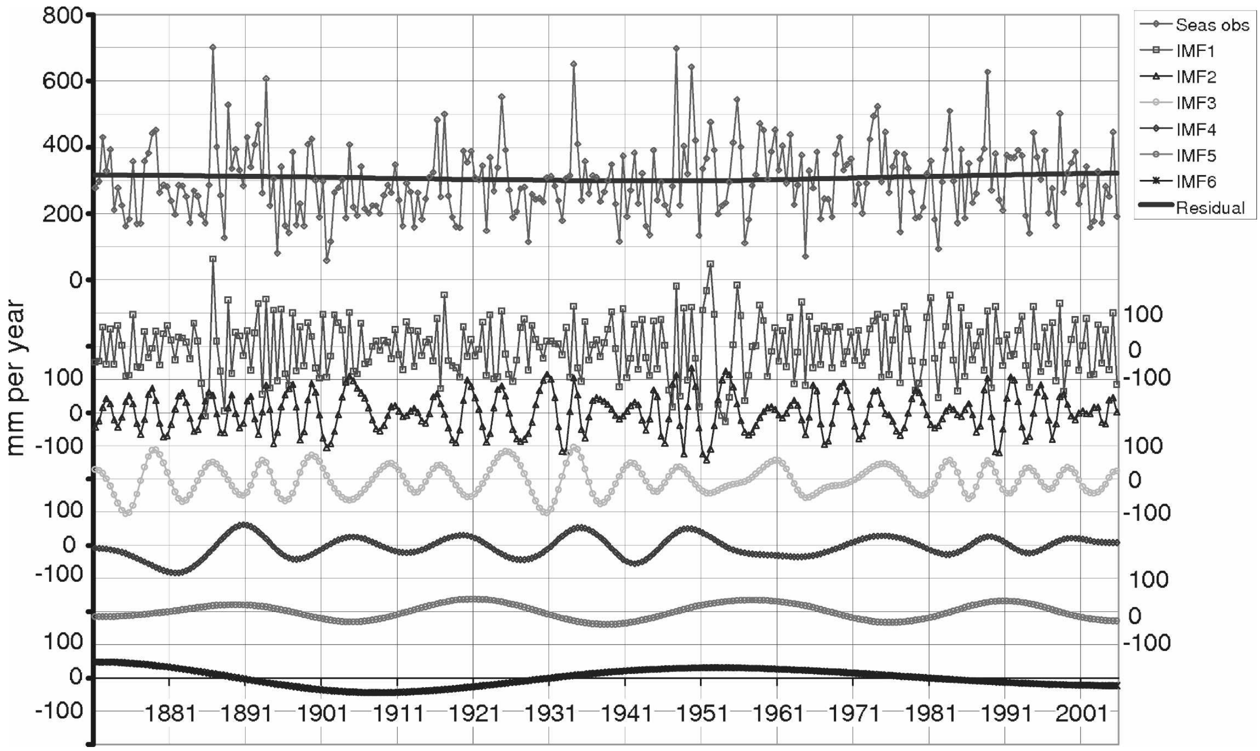


FIG. 2. IMFs and residual for the Canberra Airport rainfall station.

(average period of each ≥10 yr) and the residual are associated with interdecadal drivers (e.g., IPO). The EMD stochastic model for the Canberra Airport record is, therefore, based on two time series: the first would be IMF1 + IMF2 + IMF3, which we will denote as the intradecadal series; and the second would be based on IMF 4 + IMF 5 + IMF6 + residual and called the interdecadal series. As will be described in the next section, the two series can be independently used to stochastically generate the sequence of rainfalls and then combined to form generated data.

Summing the IMFs as described above, effectively using quasifilters, has been demonstrated to be a valid procedure by Gloersen and Huang (2003) and Peel and McMahon (2006). It should also be noted that, on the basis of the Canberra Airport 6-month rainfall data as an example, our analysis showed that the correlation between the intradecadal series and Southern Oscillation index (SOI) data is 0.23 and that between the interdecadal series and IPO is −0.49 (both measures were statistically significant at the 1% level, respectively). This observation is important because it supports our later contention that the EMD stochastic approach takes into account, in an appropriate way, the effects of large-scale climate variability such as ENSO when measured by the SOI, and the IPO when measured by the IPO index defined by Power et al. (1999). Additional analysis showed that the correlation between the intradecadal component and the IPO is −0.14 (not significant), but the correlation of 0.16 between the interdecadal component and SOI is significant at the 1% level. The lack of significant correlation between 6-month rainfall in the intradecadal component and the IPO is expected, given the persistent nature of the IPO and that within either IPO phase both wet years (i.e., La Niña events) and dry years (i.e., El Niño events) are possible, hence the significant correlation with SOI and 6-month rainfall in the interdecadal component.

EMD analysis was applied to the seasonal time series for the six rainfall stations in the Canberra water supply system. In the Canberra Airport example (Fig. 2), the time series for the six IMFs display the variance associated with high to progressively lower frequency fluctuations, finally leaving the lowest frequency fluctuation, or trend, in the residual. In this example, we note the residual shows little variation over the past 135 yr. The so-called Federation drought (years around 1901) is a significant feature in IMF6.

To check that the IMFs are independent of each other and the residual, values of the cross-correlation matrix between pairs of IMFs and residual time series were computed for the Canberra Airport rainfall station. Only one value, 0.246 (the correlation between

TABLE 2. Variances and average periods associated with IMFs and residual for Canberra Airport and Orroral at Road Crossing.

| | | row/ column | IMF 1 | IMF 2 | IMF 3 | IMF 4 | IMF 5 | IMF 6 | Residual | Sum |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Canberra Airport | Avg period (1/2 yr) | 1 | 3.1 | 7.0 | 15.9 | 36.0 | 67.5 | 180 | | |
| | Variance proportion | 2 | 0.617 | 0.261 | 0.148 | 0.079 | 0.042 | 0.055 | 0.004 | 1.206 |
| | variance | 3 | 51.1% | 21.7% | 12.3% | 6.5% | 3.5% | 4.6% | 0.3% | 100.0% |
| Orroral at Road Crossing | Avg period (1/2 yr) | 4 | 2.9 | 5.7 | 11.0 | 22.5 | 45.0 | 77.1 | | |
| | Variance proportion | 5 | 0.524 | 0.287 | 0.140 | 0.131 | 0.057 | 0.022 | 0.071 | 1.232 |
| | variance | 6 | 42.6% | 23.3% | 11.4% | 10.7% | 4.6% | 1.7% | 5.7% | 100.0% |
| Mean values for the six rainfall sites | Mean period (1/2 yr) | 7 | 2.9 | 6.3 | 13.6 | 28.5 | 58.2 | 118 | | |
| | Mean period (years) | 8 | 1.5 | 3.1 | 6.8 | 14.2 | 29.1 | 59.1 | | |
| | Mean variance | 9 | 47.3% | 21.9% | 11.6% | 8.8% | 4.3% | 1.9% | 4.2% | 100.0% |
| | | 10 | | ← 80.8% → | | | ← 19.2% → | | | |

IMF3 and IMF5) of 15 cross correlations, was found to be statistically significant; all other values were not significantly different from zero at the 5% level of significance.

An analysis of the set of IMF time series allows us to compute the amount of variance that is associated with each IMF and the residual together with the average period of each IMF. This information for two of the six rainfall time series and the mean values of the six is tabulated in Table 2. For the two stations and for each IMF, the average period (in units of ½ yr) is tabulated in rows 1 and 4. In each station set of results, the second line (Table 2) shows the variance of each IMF as a proportion of the observed variance (rows 2 and 5), and the third line shows the variance of each IMF as a percentage of the total IMFs and residual variance. As noted earlier, the variance of the IMFs and residual should sum to a value close to the variance of the original time series. The current methodology for extracting the IMFs appears to inadvertently introduce some additional variance into the IMFs (sum of variance proportions greater than 1 at all stations), and we are developing an alternative technique to decompose each IMF more efficiently. Typically, the first IMF, which is largely a noise component, contains about 50% of the variance with an average time period of 1½ yr. IMF2 and IMF3 with average periods of about 3–6 yr can be considered to be affected by ENSO and when combined account for a further 30% of the variance. In Table 2, it is observed (row 10) that the first three IMFs contain about 80% of the variance in the seasonal rainfall averaged over the six rainfall stations. The table also shows that the remaining variance is associated with IMF4–IMF6 and the residual. Because these have an average period greater than 10 yr, it is postulated that the variance contained within these components is related to interdecadal variability (e.g., IPO).

## 4. Using EMD to model multiyear wet and dry spells

The experiment reported in this section was carried out to demonstrate how a modification can be made to the traditional AR(1) time series model using EMD components to generate sequences in which the 2-yr, and longer, rainfall totals are more extreme than those generated by the traditional AR(1) model.

The time series A and B, each having 1000 hypothetical annual rainfalls, were generated using an AR(1) model with the assumption that the data were normally distributed with those characteristics listed in rows 1 and 2 of Table 3. The hypothetical parameters adopted in rows 1 and 2 of Table 3 are typical of two time series representing the EMD-based intradecadal and interdecadal components, respectively, which together sum to form an annual rainfall time series for a rainfall station located in a temperate climate zone.

In Table 3, the values of the parameters of the gen-

TABLE 3. Values of parameters for time series A and B.

| Reference | Row | | Series represents | Mean (mm) | Std dev (mm) | Autocorrelation |
|---|---|---|---|---|---|---|
| Series A | 1 | Input data | Intradecadal | 0.0 | 100 | 0.0 |
| Series B | 2 | Input data | Interdecadal | 500 | 50 | 0.9 |
| Series A | 3 | Generated (1000 values) | Intradecadal | 0.228 | 98.4 | 0.004 |
| Series B | 4 | Generated (1000 values) | Interdecadal | 497.3 | 55.6 | 0.917 |
| Series C | 5 | Generated (1000 values) | Original | 497.6 | 113.3 | 0.211 |

TABLE 4. Statistics computed from 25 000 variates generated for series A′ and B′ and the combination A′ + B′, using the parameter values in Table 3 (rows 3 and 4); and series C′, using the values in Table 3 (row 5).

| Generated sequence | Series represents | Mean (mm) | Std dev (mm) | Coefficient of skewness | Autocorrelation |
|---|---|---|---|---|---|
| Series A′ | Intra | 0.640 | 98.4 | 0.004 | 0.003 |
| Series B′, | Inter | 495.9 | 55.3 | 0.043 | 0.917 |
| Series A′ + B′* | Intra + Inter | 496.6 | 113.5 | 0.012 | 0.229 |
| Series C′ | Traditional | 496.6 | 113.2 | 0.038 | 0.216 |

* Statistics for series A′ + B′ are based on the time series obtained by summing the A′ and B′ values for each time step.

erated sequences (1000 values) are presented for series A and B (rows 3 and 4). Compared with the values in rows 1 and 2, the parameter values show that the AR(1) model using normally distributed random numbers preserves the parameters satisfactorily for the purpose of this experiment.

The next step was to combine by summing for each time step the generated series A and B, which are independent (the cross correlation is −0.021), to form the combined series, which is labeled series C in Table 3. For the experiment that follows, series C can be considered to be the original annual time series of rainfall, which can be decomposed into series A and B, which typically represents an intradecadal (high frequency) time series and an interdecadal (low frequency) time series, respectively. The significance of this step will become clear in the following sections. We will compare the properties emanating from the analysis of two different approaches in modeling the time series: 1) modeling A and B separately (designated as A′ and B′) then recombining them to form a time series A′ + B′ (EMD model) and 2) modeling C′ as a single time series in the conventional manner (traditional model).

In the “simulation,” 25 000 annual rainfall values were generated for the three series A′, B′, and C′ based on the parameters in Table 3 (rows 3–5), again using an AR(1) model and assuming the noise component is normally distributed. In the EMD stochastic model procedure (proposed herein), series A′ and B′ are combined by summation to form the generated sequences for the EMD stochastic model. Series C′ represents the sequences generated by the traditional AR(1) model method.

The statistics of the generated sequences resulting from the experiment are listed in Table 4, and cross-comparisons of their empirical distribution functions for annual rainfalls and 10-yr sums are presented in Fig. 3. It should be noted that the generated series A′ and B′ listed in Table 4 preserve the input parameters (mean, standard deviation, and autocorrelation in Table 3); furthermore, their sum (A′ + B′) preserves the original parameters (mean, standard deviation, and autocorrelation) in Table 3 (row 5). In addition, series C′ in Table 4 preserves the original parameters in Table 3.

Consider now Fig. 3, in which the empirical frequency distributions of 1) annual values and 2) 10-yr sums derived from 25 000 values of the traditional AR(1) model are compared with the EMD stochastic model. In Fig. 3, the slope of the line of best fit for the annual rainfalls is 1.0025, so for all practical purposes the distributions of the annual rainfalls for the two series are indistinguishable. This is expected because the same statistics (the means, standard deviations, and autocorrelations derived from series A′ + B′ (EMD model) or series C′ (traditional model) are used in both cases. However, as a result of the high autocorrelation in the interdecadal component (series B′) of the EMD stochastic model (in this example, a correlation of 0.917 was used), more extreme runs of high and low flows compared to the traditional model are expected. This is confirmed in Fig. 3 by the 10-yr sum series. The slopes for 2-, 5-, and 10-yr sums are 1.008, 1.143, and 1.323. (Plots for 2- and 5-yr sums are not shown). A separate analysis (also not shown) was carried out as above except the autocorrelation in row 2 of Table 3 for series B was set at 0.95 rather than 0.9. This resulted in the slopes for 5- and 10-yr sums lines (similar to the plot in

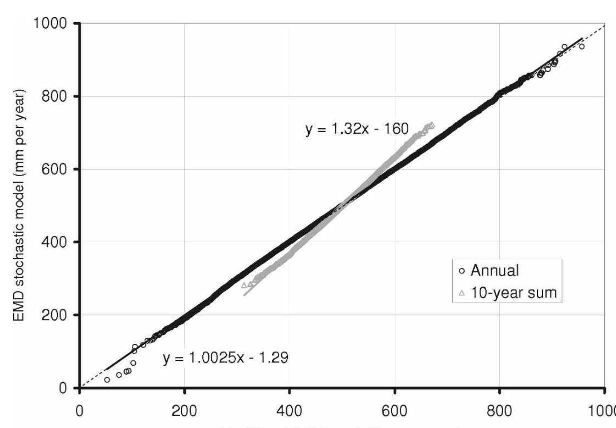


FIG. 3. Comparison between the empirical frequency distributions of the EMD stochastic model and the traditional model based on 25 000 ranked annual and 10-yr sums of annual rainfalls.

Fig. 3) increasing marginally to 1.148 and 1.363, respectively.

In summary, this analysis shows that the EMD stochastic model is able to preserve the input parameters (for normally distributed data), while also replicating quasistationarity (i.e., multiyear dry or wet spells) more realistically than the traditional AR(1) model. The experiment confirms that the more extreme events are driven through the high autocorrelation contained within the interdecadal series.

## 5. Justification of the EMD stochastic modeling procedure

For each of the six Canberra sites, the intradecadal and interdecadal IMFs were extracted for each season (June–November and December–May), resulting in 24 time series. The Matalas (1967) multisite procedure was used to generate the 6-monthly seasonal rainfall totals for each component. The Matalas model uses an AR(1) model that preserves the lag-zero and the lag-one cross correlations between the two seasons and the serial correlations of the EMD components across the six sites. Although each season and each EMD component is considered a separate "station" in the Matalas method, the autocorrelation between 6-month rainfall totals for a component at a station is preserved by the correlation structure between the two seasons in the Matalas method.

This procedure to stochastically generate data by decomposing the historical data into a series of independent simpler time series, which are then stochastically generated and recombined, has similarities with the approach recently proposed by Kwon et al. (2007). They disaggregated their time series by Wavelet decomposition into a series of orthogonal components and used several different stochastic models to generate the time series. They reported that ". . . the proposed model yields better results than a traditional linear autoregressive (AR) time-series model in terms of reproducing the time-frequency properties of the observed rainfall, while preserving the statistics usually reproduced by the AR models."

To ensure the most appropriate time series structure of the EMD stochastic model, model fitting by maximizing the likelihood function was employed (see, e.g., Hipel and McLeod (1994)). As a matter of interest, Granger and Morris (1976) show that the sum of an AR(1) model and white noise is equivalent to a combined autoregressive model of order one and moving average model of order one, that is, ARMA(1, 1). In the EMD stochastic model, we are summing two time series—an AR(1) for the interdecadal component and effectively white noise (the intradecadal component)—which when combined equate to ARMA(1, 1). The efficacy of an ARMA(1, 1) representing the winter and summer 6-monthly rainfall time series for the Canberra catchments was examined by applying standard time series analysis as set out in Hipel and McLeod (1994) to determine the most appropriate model. With the Akaike information criterion (AIC; Akaike 1974) and incorporating the Hurvich and Tsai (1989) correction for small sample sizes (denoted as AICc), we applied time series modeling as set out in Basson et al. (1994). The results for the six rainfall stations are summarized in Table 5 and suggest that the summer series is best represented by either an AR(1) or a MA(1) model because they have the same score. On the other hand, the AICc measure suggests an ARMA(1, 1) best fits the winter series. It is further noted that the magnitude of the average AICc values across both tables would indicate the ARMA(1, 1) model is at least as satisfactory as either AR(1) or MA(1). Consequently, the results in Table 5 imply that combining the two components in the EMD stochastic model is technically a valid time series procedure because it is equivalent to modeling the individual time series with an ARMA(1, 1) model.

## 6. Stochastic simulation comparing EMD and a traditional AR(1) model

As noted earlier, a key feature of the EMD process is that the IMFs are independent of each other. This means that the seasonal intradecadal and interdecadal time series at each station can be generated independently and then summed to produce a time series of 6-month rainfall totals at each rainfall station. In this section, we describe the application of EMD analysis to the seasonal rainfall data for the six rainfall stations in the Canberra water supply systems.

If we assume that the first three IMFs are associated with intradecadal drivers (e.g., ENSO) and IMFs 4–6 and the residual are associated with interdecadal drivers (e.g., IPO), a stochastic rainfall generation model can be proposed. The model would be based on two time series, "intradecadal" (which comprises IMF1 + IMF2 + IMF3) and "interdecadal" (which comprises IMF 4 + IMF 5 + IMF6 + residual). These two series have been plotted by way of example in Fig. 4 for the Canberra Airport rainfall station. Theoretically, these time series are independent—indeed, the cross correlation between the two series for the Canberra Airport data is −0.039—and represent the high (e.g., ENSO) and the low (e.g., IPO) frequency components in the time series. We propose, therefore, a stochastic model

TABLE 5. Values of AICc for parsimonious selection of appropriate time series models of 6-monthly summer and winter rainfalls. Values in bold indicate the minimum AICc values in each column, which define the optimum model. Here, $p$ is order of autoregressive process; $q$ is order of moving average process.

| ARMA ($p$, $q$) | Canberra Airport | Cotter Hut | Cotter at Kiosk | Bendora | Tinderry | Orroral at Road Crossing | Average AICc |
|---|---|---|---|---|---|---|---|
| 6-month summer rainfalls | | | | | | | |
| 0, 1 | **138.6** | **138.8** | **138.7** | **139.1** | **138.9** | **139.1** | **138.9** |
| 0, 2 | 140.4 | 140.6 | 140.5 | 140.9 | 140.5 | 141.1 | 140.7 |
| 1, 0 | 138.7 | **138.8** | **138.7** | **139.1** | **138.9** | **139.1** | **138.9** |
| 1, 1 | 140.2 | 140.6 | 140.5 | 141.2 | 141.2 | 141.1 | 140.8 |
| 1, 2 | 142.5 | 142.7 | 142.6 | 142.7 | 142.6 | 143.1 | 142.7 |
| 2, 0 | 140.4 | 140.6 | 140.5 | 140.9 | 140.5 | 141.1 | 140.7 |
| 2, 1 | 142.3 | 143.8 | 142.4 | 142.7 | 139.2 | 143.1 | 142.3 |
| 2, 2 | 143.8 | 144.7 | 144.6 | 143.3 | 141.9 | 144.8 | 143.9 |
| 6-month winter rainfalls | | | | | | | |
| 0, 1 | 138.4 | 136.3 | 138.6 | 137.5 | 137.1 | **138.8** | 137.8 |
| 0, 2 | 139.5 | 137.5 | 140.6 | 138.7 | 138.8 | 140.9 | 139.3 |
| 1, 0 | 138.5 | 135.8 | 138.6 | 137.3 | 137.3 | **138.8** | 137.7 |
| 1, 1 | **136.9** | **127.2** | **136.7** | **129.8** | **130.5** | 140.5 | **133.6** |
| 1, 2 | 138.9 | 129.3 | 142.7 | 131.7 | 132.1 | 142.9 | 136.3 |
| 2, 0 | 139.6 | 136.9 | 140.6 | 138.5 | 138.8 | 140.9 | 139.2 |
| 2, 1 | 138.9 | 129.1 | 142.8 | 131.6 | 132.7 | 142.7 | 136.3 |
| 2, 2 | 142.2 | 131.2 | 138.4 | 132.8 | 133.8 | 143.2 | 136.9 |

in which each series is stochastically generated separately and then combined, by addition, to form the stochastically generated time series of seasonal (i.e., 6-month) rainfall data.

To execute this model several practical issues need to be addressed.

1) The basic dataset, used to stochastically generate the 6-month rainfall totals at the six stations, consists of four time series for each station, a winter/spring (June–November) and a summer/autumn (December–May) series for both intradecadal and interdecadal components. The statistical characteristics of

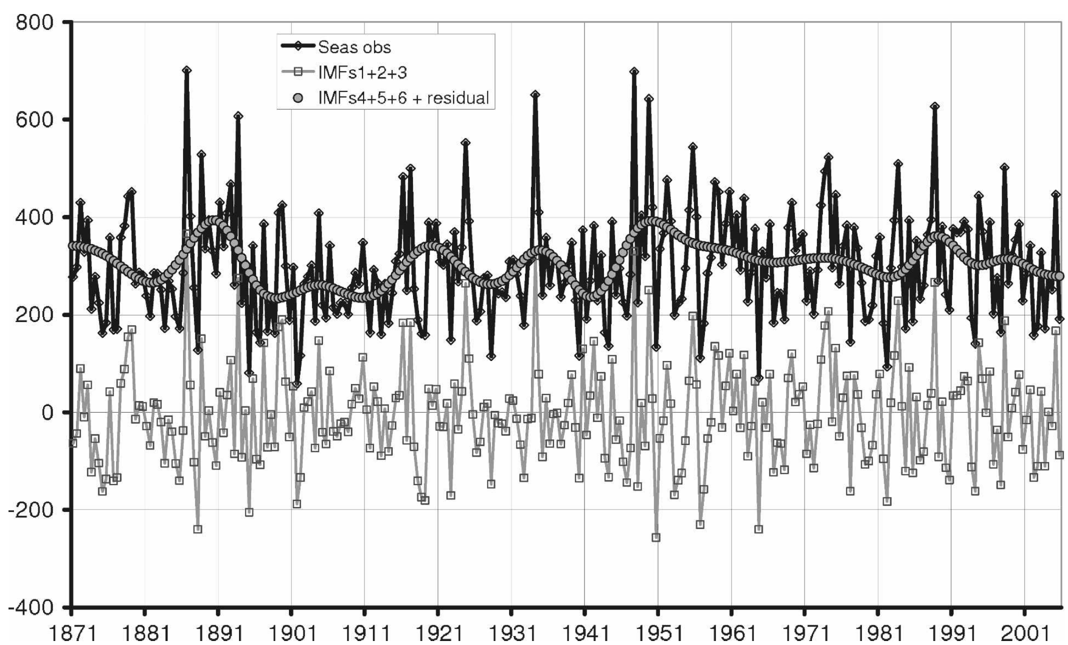


FIG. 4. Intradecadal (IMFs 1 + 2 +3) and interdecadal (IMFs 4 + 5 +6 + residual) time series components for the Canberra Airport rainfall station.

TABLE 6. Statistical characteristics of 6-month rainfall totals for intradecadal and interdecadal time components for Canberra Airport and Orroral at Road Crossing rainfall stations. Here, $\mu$, $\sigma$, $\gamma$, $\rho$ are the mean, std dev, coefficients of skewness, and lag-one serial correlation, respectively, of 6-monthly rainfalls.

| Rainfall stations | Decadal period | Season | $\mu$ (mm) | $\sigma$ (mm) | $\gamma$ | $\rho$ |
|---|---|---|---|---|---|---|
| Canberra Airport | Intradecadal | Winter–spring | −6.4 | 98.1 | 0.214 | 0.032 |
| | | Summer–autumn | −5.5 | 111.2 | 0.756 | |
| | Interdecadal | Winter–spring | 305.1 | 39.8 | 0.138 | 0.989 |
| | | Summer–autumn | 304.8 | 39.8 | 0.152 | |
| Orroral at Road Crossing | Intradecadal | Winter–spring | 5.5 | 107.1 | 0.149 | 0.000 |
| | | Summer–autumn | −19.7 | 116.4 | 0.645 | |
| | Interdecadal | Winter–spring | 337.1 | 51.2 | 0.210 | 0.982 |
| | | Summer–autumn | 337.0 | 51.2 | 0.216 | |

each of the four time series are listed in Table 6 for Canberra Airport and Orroral.

2) For the intradecadal winter/spring series and both seasons of the interdecadal series, the coefficient of skewness ($\gamma$) is small, suggesting that a normal probability density function (PDF) maybe appropriate. However, for the intradecadal summer/autumn series, $\gamma$ is moderate in magnitude, suggesting a normal PDF would not be suitable. Considering these differences, we transformed the data using a Wilson–Hilferty (W–H) transformation (Wilson and Hilferty 1931). [Prior to choosing the W–H transformation, we tested a Box–Cox (B–C; Box and Cox 1964) transformation; we found the W–H preserved the input parameters more closely.]
3) As expected, a significant contrast is evident in the lag-one serial correlations. Those for the intradecadal component are not significantly different from zero, whereas those for the interdecadal series are extremely high [approximately 0.98 and statistically significant as a result of the incorporation of the residual (trend) term in the series].

To generate 6-month rainfall totals, we apply the Matalas (1967) multisite procedure in which each of the 24 time series is considered to be a site. The Matalas procedure, which assumes an AR(1) model, ensures that the basic parameters and the lag-zero and lag-one cross correlations are preserved. The model is defined as follows:

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\boldsymbol{\varepsilon}_t, \tag{1}$$

where $\mathbf{x}_t$ and $\mathbf{x}_{t-1}$ are $24 \times 1$ vectors whose elements are standardized residuals for the $t^{\text{th}}$ and $(t-1)^{\text{th}}$ consecutive periods; $\boldsymbol{\varepsilon}_t$ is a $24 \times 1$ vector whose elements are random normal variates (resulting from either the B–C or W–H transformation); and $\mathbf{A}$ and $\mathbf{B}$ are $24 \times 24$ matrices whose elements consist of a combination of the lag-zero and lag-one cross correlations between the time series defined as follows:

$$\mathbf{A} = \mathbf{M}_1\mathbf{M}_0^{-1} \quad \text{and} \tag{2}$$

$$\mathbf{B}^{\mathrm{T}} = \mathbf{M}_0 - \mathbf{M}_1\mathbf{M}^{-1}\mathbf{M}^{\mathrm{T}}, \tag{3}$$

where $\mathbf{M} = E(\mathbf{x}_t\mathbf{x}_{t-1}^{\mathrm{T}})$, $\mathbf{M}_0 = E(\mathbf{x}_{t-1}\mathbf{x}_{t-1}^{\mathrm{T}})$, and $E(\cdot)$ is the expectation operator. Details for solving for $\mathbf{A}$ and $\mathbf{B}$ are given in Kuczera (1987), Salas (1993), and Basson et al. (1994).

Using the device of treating the seasonal components as different "sites" in the multisite AR(1) model in Eq. (1), the autocorrelations in and between the 6-month rainfall series are automatically accounted for in formulating $\mathbf{A}$ and $\mathbf{B}$; thus, the correlation structures embedded in those matrices ensure the intercorrelations (space and time) in the 24 individual time series are preserved.

The key steps to stochastically generating 100 replicates of 6-month rainfall totals, each of length 135 yr for the six rainfall stations in the Canberra catchment, are listed and discussed below.

Step 1: For each station, time series of historical 6-month rainfall totals are prepared, where the length of record is 135 yr (270 half-years).

Step 2: An EMD analysis is carried out on each time series. For the six Canberra catchments stations, five or six IMFs were found at each station.

Step 3: For each station, prepare separately the intradecadal time series by summing IMF1 + IMF2 + IMF3 time series (average period <10 yr) and the interdecadal time series by summing IMF4 + IMF5 + IMF6 (average period ≥10 yr) + residual (resulting in 270 half-years).

Step 4: Split the intradecadal time series into two time series (each 6-month rainfall consisting of 135 values) representing winter/spring and summer/autumn values, respectively.

Step 5: Split the interdecadal time series into two time series (each 6-month rainfall consisting of 135 values) representing winter/spring and summer/

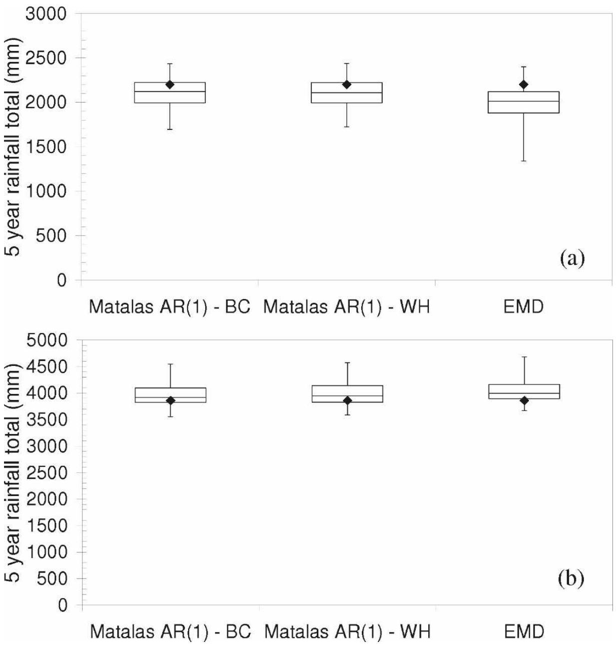


FIG. 5. Box plots of (a) minimum and (b) maximum 5-yr totals for historical data (diamond) and for generated rainfalls based on traditional AR(1) (using BC and WH transformations) and EMD stochastic models for Canberra Airport rainfall station.

autumn values, respectively. Note that Steps 1–5 are carried out separately for each rainfall station.

Step 6: Transform the 24 time series into normally distributed values and collate each into a single dataset or file ready for multisite stochastic data generation.

Step 7: Following Matalas (1967), solve for **A** and **B** in Eq. (1) based on the data files established in Step 6, and stochastically generate 100 replicates of the 24 time series (each consisting of 135 values) listed in Step 6 and apply inverse transforms to the generated values.

Step 8: For each replicate, station and season, add together the intradecadal and interdecadal time series to form 12 stochastically generated seasonal (i.e., 6 months) time series (i.e., for each station and replicate 135 half-years of winter/spring and summer/autumn rain).

Step 9: For each station and each replicate, interleave each pair of seasons to generate 100 replicates of 135 annual totals (June–May).

## 7. Assessment of adequacy of EMD stochastic model

To assess the adequacy of the EMD stochastic model, not only were the generated statistics reviewed but also a Matalas AR(1) model using the 6-month rainfall totals (not decomposed) at the six rainfall stations in a traditional analysis (see, e.g., Srikanthan et al. 1984; Nalbantis et al. 1992) was carried out and both sets of statistics were compared. The results of the comparisons are presented in Table 7 (Canberra Airport and Orroral) and Table 8 (all stations) and Fig. 5 (Canberra Airport). Note that when the traditional AR(1) approach was applied, it was found that a Box–Cox transformation satisfactorily normalized the historical data; however, a Wilson–Hilferty transformation was found to achieved similar results (Fig. 5). The statistics shown in Tables 7 and 8 for the traditional AR(1) model were obtained when the Box–Cox transformation was employed. By contrast, as noted earlier, the two components of the EMD model were better served by the Wilson–Hilferty transform.

Table 7 shows that both the EMD and traditional models preserve the mean ($\mu_a$), the coefficient of variation ($Cv_a$), the coefficient of skewness ($\gamma_a$), and the lag-one serial correlation coefficient ($\rho_a$) of the historical annual rainfalls. Also, at the annual time step Table 8 shows that the lag-zero and lag-one cross correlations for the generated data satisfactorily reproduce the historical values.

One of the objectives in developing the new stochastic model of rainfall using the EMD approach is to provide replicates of rainfall time series that will be

TABLE 7. Comparison between key annual statistics (mean, coefficients of variation and skewness, and autocorrelation) of historical data with those generated by the EMD stochastic model and the traditional Matalas AR(1) model for Canberra Airport and Orroral at Road Crossing rainfall stations. Values for EMD and traditional AR(1) models are means of the 100 replicates (numbers in parenthesis are standard deviations of the 100 simulated sequences). Here, $\mu_a$, $Cv_a$, $\gamma_a$, $\rho_a$ same as Table 6 but for annual rainfalls.

| Rainfall station | | $\mu_a$ (mm) | $Cv_a$ | $\gamma_a$ | $\rho_a$ |
|---|---|---|---|---|---|
| Canberra Airport | Historical | 598 | 0.269 | 0.698 | 0.134 |
| | EMD | 597 (12) | 0.271 (0.019) | 0.258 (0.232) | 0.124 (0.100) |
| | Traditional | 597 (16) | 0.264 (0.016) | 0.426 (0.244) | 0.127 (0.086) |
| Orroral at Road Crossing | Historical | 660 | 0.268 | 0.414 | 0.167 |
| | EMD | 658 (16) | 0.269 (0.019) | 0.172 (0.213) | 0.175 (0.098) |
| | Traditional | 659 (19) | 0.266 (0.016) | 0.430 (0.219) | 0.139 (0.093) |

TABLE 8. Comparison between annual lag-zero cross correlations of historical data with those generated by the EMD stochastic model and the traditional AR(1) model. Values for EMD and traditional AR(1) models are means of the 100 replicates.

| | | Rainfall stations | | | | |
|---|---|---|---|---|---|---|
| | | 2 | 3 | 4 | 5 | 6 |
| 1 | Historical | 0.893 | 0.889 | 0.903 | 0.915 | 0.867 |
| | EMD | 0.894 | 0.886 | 0.904 | 0.912 | 0.865 |
| | Traditional | 0.893 | 0.892 | 0.911 | 0.914 | 0.868 |
| 2 | Historical | 1 | 0.959 | 0.948 | 0.859 | 0.937 |
| | EMD | | 0.958 | 0.946 | 0.859 | 0.935 |
| | Traditional | | 0.960 | 0.951 | 0.856 | 0.940 |
| 3 | Historical | | 1 | 0.946 | 0.863 | 0.936 |
| | EMD | | | 0.945 | 0.864 | 0.936 |
| | Traditional | | | 0.949 | 0.867 | 0.942 |
| 4 | Historical | | | 1 | 0.838 | 0.900 |
| | EMD | | | | 0.836 | 0.898 |
| | Traditional | | | | 0.850 | 0.908 |
| 5 | Historical | | | | 1 | 0.910 |
| | EMD | | | | | 0.912 |
| | Traditional | | | | | 0.908 |

used (after disaggregating to a shorter time step) to estimate runoff through a rainfall–runoff model as input into the simulation of the Canberra water supply system. The combination of reservoirs (three in series and one in parallel) is a carryover system, and drawdown periods can last up to five years. Thus, reproducing the characteristics of potentially low flows cumulated over several years is important. However, although it is not possible to directly assess whether a model is performing well, it is possible to assess whether it is performing as expected. Thus, we would expect that the EMD model should be producing multiyear low (and high) rainfalls that, overall, are more extreme than those generated by the traditional model. With the 100 replicates of the stochastically generated rainfalls for the Canberra Airport rainfall station, Figs. 5a and 5b show the median, 25th and 75th percentiles, and the range for the minimum and maximum, respectively, 5-yr rainfall totals. The figure compares the results based on the EMD stochastic model with the traditional AR(1) model. As noted above, results from the two versions of the traditional AR(1) model are shown, one using a Box–Cox transformation and one using Wilson–Hilferty transformation.

It is observed in Figs. 5a and 5b that the EMD stochastic model exhibits more extreme rainfalls than the traditional models. Plots (not shown here) for the other rainfall stations exhibit similar features. In particular, the median value of the minimum 5-yr rainfall total based on the EMD stochastic model is 107 mm (95 mm) lower than the traditional model using a Box–Cox (Wilson–Hilfery) transformation. This difference in minimum 5-yr rainfall totals corresponds to approximately 22 mm, on average, per year (3.7% of mean annual rainfall; Fig. 5). Assuming a rainfall elasticity of runoff of 2.4—southeastern Australia values are probably larger than this; refer to Chiew et al. (2005)—we would expect the estimated 5-yr minimum runoff from the EMD stochastic rainfall to be at least 8.9% less than a value based on the traditional stochastic model. In terms of planning for extreme drought conditions (like those that the Canberra region has experienced in recent years), this difference would have considerable implications for water resources planning in the region.

The observation that the EMD stochastic model generates lower 5-yr rainfall sums than the traditional model is consistent with our expectation. In the traditional AR(1) model, the low-frequency fluctuations (probably driven by interdecadal scale climate phenomena such as the IPO) are not preserved, whereas by separating the historical time series into the intradecadal and the interdecadal components and generating each component separately but maintaining autocorrelation, the EMD stochastic model allows the nonstationary impacts of multiyear climate variability to be captured through the very high autocorrelation (Table 6). Hence, lower extreme rainfalls are produced in the EMD model compared with the traditional model.

### *Model attributes and limitations*

A key attribute of the EMD approach is that for a given time series, the IMFs and the trend are independent, which allows the IMFs to be combined as appropriate. In the application we have described above, it was important for the project objectives that the current known climate drivers for eastern Australia were modeled. By choosing the IMFs with periods $< 10$ yr and separately $\geq 10$ yr, the two most dominant large-scale climate influences on eastern Australia (i.e., ENSO at intradecadal time scales and IPO at interdecadal time scales, respectively) are implicitly accounted for, without overconditioning the model or excluding other possible climate influences. (Recall from our earlier analysis of 6-monthly rainfalls that the correlation between the intradecadal series for Canberra Airport and SOI was 0.23 and that between the interdecadal series and IPO was $-0.49$.)

An important feature of this formulation is that as our understanding of the natural climate processes improves and, importantly, our understanding of how anthropogenic influences are likely to change the future climate, our approach allows projected changes to various aspects of the climate drivers (e.g., ENSO, IPO) to be easily incorporated into the stochastic generating process through the relevant IMF. For example, Peel and McMahon (2006) observed that the interannual variability of precipitation has marginally decreased since 1970, but the proportion of total variance as a result of the intradecadal component of interannual variability has increased substantially relative to the interdecadal component. If such a trend were confirmed—for example, by global ocean/atmosphere circulation models—it would be straightforward to incorporate such a feature into the new EMD-based data generation process.

We also believe the model has application where the trend needs to be explicitly taken into account, for example, in estimating the impact of climate change on future water resources. This can be done by separately modeling the residual (trend) as a single component. After generation, the two or more time series (one incorporating the modified residual and the others handled separately) would be summed to produce the generated replicate.

As a general rule, stochastic data generation is not based on a causal model, but its justification is pragmatic (Fiering and Jackson 1971) in that it relies on the observed time series structure of the historical data. However, in the approach developed here, we have attempted to introduce some meteorological perspective into the method by observing that intradecadal sequences are related to SOI and interdecadal sequences are related to IPO, both relationships being statistically significant.

A minor limitation in the approach relates to the increased degree of analysis required to determine the IMFs via EMD. Future research into the efficient extraction of IMFs from a time series using EMD is expected to resolve the minor issue of extra variance introduced into the recombined sequence.

Our experience with EMD analysis (Peel et al. 2005; Peel and McMahon 2006) suggests that the method is applicable to any hydrologic time series (e.g., rainfall, streamflow, and temperature) in which there are sufficient data to define several IMFs.

## 8. Conclusions

In this paper, we introduced a new approach to stochastically generate rainfall, and applied the method to a multisite situation at a 6-month time step. The results were compared with those generated from a traditional AR(1) multisite model.

The first step in the procedure is to use empirical mode decomposition (EMD) to disaggregate the rainfall time series into a set of time series known as intrinsic mode functions (IMFs) and a residual (trend), which theoretically are independent of one another. These time series (IMFs), obtained when EMD analysis is applied to hydrometeorological series, display linear and/or nonlinear behavior in both frequency and amplitude modulation. However, because the IMFs and the residual time series are mutually uncorrelated, each series can be stochastically generated independently and then recombined to form a replicate of the original series.

In the example application described in the paper, the IMFs (5 or 6, depending on the data) were combined after the initial decomposition into two time series, one representing intradecadal drivers (average period of the IMFs $<10$ yr) such as ENSO and the other representing interdecadal drivers (average period of the IMFs $\geq 10$ yr) such as the IPO.

The method was applied within a project setting to six rainfall stations located in the Canberra (Australia) water supply catchments. The project required that rainfall be stochastically generated as 6-month totals at six rainfall stations as a multisite situation. The Matalas (1967) multisite AR(1) model was used to carry out this phase of the analysis. For comparison, rainfall sequences were generated using a traditional Matalas multisite AR(1) model applied to the original data without decomposition and the EMD stochastic model. For the various time series, a Box–Cox or Wilson–Hilferty transformation was used to normalize the rainfall data.

Both the EMD stochastic and the traditional models preserved the historical input parameters, but the EMD model was able to generate more extreme multiyear maximum and minimum rainfalls. This feature makes the EMD approach an attractive procedure to generate extremes in a climate change environment. It can simulate nonstationarity characteristics in the historical time series, and it is easy to alter the decomposed time series components to examine the effect of anthropogenic climate change.

*Acknowledgments.* The authors wish to thank ActewAGL, Canberra, for providing the rainfall data used in this study and for partially funding the analysis. The authors acknowledge Stephen Impey for preparing Figure 1. They also acknowledge Rory Nathan for his advice on, and contribution to, the development of this methodology and earlier drafts of this paper. The authors thank the anonymous reviewers for their comments that helped to improve the paper.

REFERENCES

Akaike, H., 1974: A new look at the statistical model identification. *IEEE Trans. Autom. Control,* **19,** 716–723.

Allan, R. J., 2000: ENSO and climatic variability in the last 150 years. *El Niño and the Southern Oscillation: Multiscale Vari-*

*ability and Global and Regional Impacts,* H. F. Diaz and V. Markgraf, Eds., Cambridge University Press, 3–56.

Basson, M. S., R. B. Allen, G. G. S. Pegram, and J. A. van Rooyen, 1994: *Probabilistic Management of Water Resource and Hydropower Systems.* Water Resources Publications, 424 pp.

Box, G. E. P., and D. R. Cox, 1964: An analysis of transformations. *J. Roy. Stat. Soc., Ser. A,* **B26,** 211–243.

Chiew, F. H. S., R. N. Jones, and W. C. Boughton, 2005: Modelling hydrologic sensitivity to climate conditions. *Proc. 29th Hydrology and Water Resources Symp*., Canberra, Australia, Institution of Engineers, 1–8.

Diaz, H. F., and V. Markgraf, Eds., 2000: *El Niño and the Southern Oscillation: Multiscale Variability and Global and Regional Impacts.* Cambridge University Press, 496 pp.

Erlanger, P. D. and B. Neal, 2005: Framework for urban water resource planning. Water Services Association of Australia Occasional Paper 14, 32 pp.

Erskine, W. D., and R. F. Warner, 1988: Geomorphic effects of alternating flood-and drought-dominated regimes on NSW coastal rivers. *Fluvial Geomorphology of Australia,* R. F. Warner, Ed., Academic Press, 373 pp.

Fiering, M. B., and B. B. Jackson, 1971: *Synthetic Streamflows. Water Resources Monogr.,*Vol. 1, Amer. Geophys. Union, 98 pp.

Folland, C. K., D. E. Parker, A. W. Colman, and R. Washington, 1999: Large scale modes of ocean surface temperature since the late nineteenth century. *Beyond El Niño: Decadal and Interdecadal Climate Variability,* A. Navarra, Ed., Springer, 73–102.

Gloersen, P., and N. Huang, 2003: Comparison of interannual intrinsic modes in hemispheric sea ice covers and other geophysical parameters. *IEEE Trans. Geosci. Remote Sens.,* **41,** 1062–1074.

Granger, C. W. J., and M. J. Morris, 1976: Time series modelling and interpretation. *J. Roy. Stat. Soc., Ser. A,* **139,** 246.

Hipel, K. W., and A. I. McLeod, Eds., 1994: *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, 1013 pp.

Huang, N. E., and Coauthors, 1998: The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. Roy. Soc. London,* **454A,** 903–995.

——, Z. Shen, and S. R. Long, 1999: A new view of nonlinear water waves: The Hilbert spectrum. *Annu. Rev. Fluid Mech.,* **31,** 417–457.

Hurvich, C. M., and C.-H. Tsai, 1989: Regression and time series model selection in small samples. *Biometrika,* **76,** 291–307.

Kiem, A. S., S. W. Franks, and G. Kuczera, 2003: Multi-decadal variability of flood risk. *Geophys. Res. Lett.,* **30,** 1035, doi:10.1029/2002GL015992.

Kirkup, H., G. Brieley, A. Brooks, and A. Pitman, 1998: Temporal variability of climate in south-eastern Australia: A reassessment of flood- and drought-dominated regimes. *Aust. Geogr.,* **29,** 241–255.

Kuczera, G., 1987: On maximum likelihood estimators for the multisite lag-one streamflow model: Complete and incomplete data cases. *Water Resour. Res.,* **24,** 641–645.

Kwon, H.-H., U. Lall, and A. F. Khalil, 2007: Stochastic simulation model for nonstationary time series using an autoregressive wavelet decomposition: Applications to rainfall and temperature. *Water Resour. Res.,* **43,** W05407, doi:10.1029/2006WR005258.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.,* **78,** 1069–1079.

Matalas, N. C., 1967: Mathematical assessment of synthetic hydrology. *Water Resour. Res.,* **3,** 937–945.

Nalbantis, I., D. Koutsoyiannis, and Th. Xanthopoulos, 1992: Modeling the Athens water supply system. *Water Resour. Manage.,* **6,** 57–67.

Peel, M. C., and T. A. McMahon, 2006: Recent frequency component changes in interannual climate variability. *Geophys. Res. Lett.,* **33,** L16810, doi:10.1029/2006GL025670.

——, G. G. S. Pegram, and T. A. McMahon, 2004: Global analysis of runs of annual precipitation and runoff equal to or below the median: Run length. *Int. J. Climatol.,* **24,** 807–822.

——, G. E. Amirthanathan, G. G. S. Pegram, T. A. McMahon, and F. H. S. Chiew, 2005: Issues with the application of empirical decomposition analysis. *Proc. MODSIM 2005 Int. Congress on Modelling and Simulation,* Melbourne, Australia, Modelling and Simulation Society of Australia and New Zealand, 5 pp.

Philander, S. G., 1990: *El Niño, La Niña, and the Southern Oscillation.* International Geosphysics Series, Vol. 46, Academic Press, 293 pp.

Power, S., T. Casey, C. Folland, A. Colman, and V. Mehta, 1999: Inter-decadal modulation of the impact of ENSO on Australia. *Climate Dyn.,* **15,** 319–324.

Salas, J. D., 1993: Analysis and monitoring of hydrologic time series. *Handbook of Hydrology,* D. R. Maidment, Ed., McGraw-Hill, 1424 pp.

Sinclair Knight Merz, 2004: Infilling of rainfall and evaporation data for Canberra Future Water Supply Options Study. Sinclair Knight Merz File Note WC02677.400, 17 pp.

Srikanthan, R., T. A. McMahon, G. P. Codner, and R. G. Mein, 1984: Theory and application of some multisite streamflow generation models. *Trans. Inst. Eng. Aust.,* **CE26,** 272–279.

Torrence, C., and G. P. Compo, 1998: A practical guide to wavelet analysis. *Bull. Amer. Meteor. Soc.,* **79,** 61–78.

Wilson, E. B., and M. M. Hilferty, 1931: The distribution of chi–square. *Proc. Natl. Acad. Sci. USA,* **17,** 684–688.

Yevjevich, V., 1972: *Stochastic Processes in Hydrology*. Water Resources Publications, 276 pp.