Physica A 442 (2016) 91–99

Contents lists available at ScienceDirect

Physica A

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

Maximum entropy spectral analysis for streamflow forecasting

Huijuan Cui [a,b,*], Vijay P. Singh [c,d]

[a] Institute of Geographic Science and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China
[b] Water Management and Hydrological Science, Texas A&M University, College Station, TX, 77843-2117, USA
[c] Department of Biological & Agricultural Engineering, Texas A&M University, College Station, TX, 77843-2117, USA
[d] Zachry Department of Civil Engineering, Texas A&M University, College Station, TX, 77843-2117, USA

HIGHLIGHTS

- Configurational entropy spectral analysis is developed with spectral power as a random variable.
- The proposed spectral analysis yields Burg's maximum entropy spectral analysis.
- The maximum entropy spectral analysis encompasses the Burg entropy spectral analysis and two configurational entropy spectral analyses.

ARTICLE INFO

Article history:
Received 22 April 2015
Received in revised form 31 July 2015
Available online 9 September 2015

Keywords:
Burg entropy
Configurational entropy
Spectral analysis
Streamflow forecasting

ABSTRACT

Configurational entropy spectral analysis (CESAS) is developed with spectral power as a random variable for streamflow forecasting. It is found that the CESAS derived by maximizing the configurational entropy yields the same solution as by the Burg entropy spectral analysis (BESA). Comparison of forecasted streamflows by CESAS and BESA shows less than 0.001% difference between the two analyses and thus the two entropy spectral analyses are concluded to be identical. Thus, the Burg entropy spectral analysis and two configurational entropy spectral analyses form the maximum entropy spectral analysis.

1. Introduction

The entropy theory comprising the Shannon [1] entropy and the principle of maximum entropy (POME) [2,3] has been widely applied in hydrology [4–8]. The advantage of using the entropy theory is that it combines statistical information with physical characteristics and provides least-biased estimation. However, it was not used for forecasting until Burg [9,10] developed the maximum entropy spectral analysis (MESA) which is called the Burg entropy spectral analysis (BESA). The Burg entropy is defined in terms of frequency f as a random variable:

$$H_B(p) = \int_{-W}^{W} \ln[p(f)]\mathrm{d}f \tag{1}$$

where frequency f is considered as a random variable, W is the Nyquist frequency, and $p(f)$ is the normalized spectral density taken as the probability density function (PDF) of f.

* Corresponding author at: Institute of Geographic Science and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China. Tel.: +86 13466547017.
E-mail address: cuihj@igsnrr.ac.cn (H. Cui).

http://dx.doi.org/10.1016/j.physa.2015.08.060
0378-4371/

For a stationary random process BESA computes spectral power from the autocorrelation of given lags, without assuming autocorrelation of unknown lags as zero [11]. It has an advantage over classical method in terms of computational ease, short and smooth spectra with a high degree of resolution, and robustness of estimates and their stability. As a result, BESA has been widely applied to spectral analysis of geomagnetic fields, climate indices, surface air temperature, geophysical exploration, tide levels, precipitation, and runoff [12–21]. BESA has also been employed for long-term streamflow forecasting and real-time flood forecasting [22–25,6,7] and has been shown to have an advantage in long-term streamflow forecasting over traditional stochastic methods, but has not been found to be superior for short-term flow forecasting.

The maximum entropy spectral analysis can be derived using the configurational entropy introduced by Frieden [26] and Gull and Daniell [27], which is defined as

$$H_{CF}(f) = -\int_{-W}^{W} p(f)\ln[p(f)]\mathrm{d}f. \tag{2}$$

It is noted from Eq. (2) that the configurational entropy is defined in the same form as the Shannon entropy. The configurational entropy spectral analysis with frequency as a random variable (CESAF) is shown to be preferred over BESA for autoregressive moving average (ARMA) and moving average (MA) processes [28]. CESAF has been applied to monthly streamflow forecasting, and has been found to perform better than BESA [29].

On the other hand, configurational entropy spectral analysis can be derived with spectral power as a random variable (CESAS). The streamflow time series $y_t,\ t = 1, 2, \ldots, T$ can be transferred to spectral powers $x_k,\ k = 1, \ldots, n$, in the frequency domain by the Fast Fourier transform. For each frequency f_k there is one associated spectral power x_k. Let $\vec{x} = (x_1, x_2, \ldots, x_n)$ and let it be assumed that each probability density function $p\ (x_k)$ is considered independent identically distributed. Then the joint probability density function can be noted as $p(\vec{x}) = p(x_1)\cdots p(x_n)$. Now assuming each spectral power x_k as a random variable, the configurational entropy is defined as

$$H_{CS}(p) = -\int_{D} p(\vec{x})\ln[p(\vec{x})]\mathrm{d}\vec{x} = -E\{\ln[p(\vec{x})]\}. \tag{3}$$

However, it was shown by Gray [30] that if x_k came from an N-dimensional Gaussian distribution, then the joint distribution can be given by

$$p(\vec{x}) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}}\left(\frac{1}{\det\Re}\right)^{\frac{1}{2}}\exp\left(-\frac{1}{2}\vec{x}^{t}\Re^{-1}\vec{x}\right) \tag{4}$$

where $\det\Re$ is the determinant of the autocorrelation matrix defined by

$$\Re = E[Y^{T}Y] = \begin{bmatrix} \rho_0 & \rho_1 & \cdots & \rho_{n-1} & \rho_n \\ \rho_1 & \rho_0 & \cdots & \rho_{n-2} & \rho_{n-1} \\ \vdots & & & & \vdots \\ \rho_{n-1} & \rho_{n-2} & \cdots & \rho_0 & \rho_1 \\ \rho_n & \rho_{n-1} & \cdots & \rho_1 & \rho_0 \end{bmatrix} \tag{5}$$

where ρ_n is the autocorrelation of the nth lag. [Define matrix Y.] Substitution of Eq. (4) into Eq. (3) yields

$$H_{CS}(p) = \ln[(2\pi e)^{\frac{N}{2}}(\det\Re)^{\frac{1}{2}}]. \tag{6}$$

It is noted that the autocorrelation is linked to the spectral density. Thus, replacing the autocorrelation in Eq. (6) with spectral density, the result is [give the intermediate steps] $H(f) = \int_{-W}^{W}\ln[p(f)]\,\mathrm{d}f$, which is the Burg entropy.

The objective of this paper therefore is to derive the configurational entropy spectral analysis with spectral power as a random variable, and to show how it yields Burg entropy spectral analysis.

2. Review of Burg entropy spectral analysis

Using the principle of maximum entropy, Burg [9,10] developed BESA for a stationary random process, which provides a basis to connect the spectra with the autoregressive (AR) process. By maximizing the Burg entropy in Eq. (1) with the use of the method of Lagrange multipliers, he obtained the spectral density as

$$P(f) = \frac{1}{\sum_{n=-N}^{N}\lambda_n \mathrm{e}^{-\mathrm{i}2\pi f n\Delta t}} \tag{7}$$

where λ_n are the Lagrange multipliers. To extend the autocorrelation and forecast time series, Burg [9,10] showed that Eq. (7) is equivalent to the spectral density of the AR process, which can be expressed as

$$P(f) = \frac{\sigma^2}{\left|1 + \sum_{n=-N}^{N} a_n z^n\right|^2} \tag{8}$$

where σ^2 is the variance of streamflow values, $z = \exp[-\mathrm{i}2\pi f \Delta t]$, and a_n are the prediction coefficients.

Then, Burg [9,10] modified the Levinson algorithm [31] for estimating the prediction coefficients by using both forward and backward forecasting errors. The original Levinson algorithm can be written as

$$a_s^N = \begin{cases} 1, & s = 1 \\ a_s^{N-1} + c_N a_{N-s}^{*N-1}, & 1 < s < N \\ c_N, & s = N \end{cases} \tag{9}$$

where c_N is called the reflection coefficient. By minimizing the backward and forward prediction error, Burg [9,10] computed the reflection coefficient as

$$c_N = \frac{-2 \sum_{m=1}^{M} w_m (b_m f_m)}{\sum_{m=1}^{M} w_m (b_m^2 + f_m^2)} \tag{10}$$

where w_m are the weight factors such that $\sum_{m=1}^{M} w_m = 1$, and f_m and b_m are the forward and backward prediction errors, respectively, which are defined as

$$f_m = a_{m-1} y_2 + \cdots + a_1 y_m + y_{m+1} \tag{11}$$

and

$$b_m = y_1 + a_1 y_2 + \cdots + a_{m-1} y_m. \tag{12}$$

It has been shown that the Levinson–Burg algorithm is equivalent to the least-squared fitting of a discrete-time all-pole model to streamflow data series [32,33].

3. Development of configurational entropy spectral analysis

The configurational entropy spectral analysis is developed with spectral power as a random variable. Therefore, unlike the original Burg entropy, the probability density function of spectral power is first obtained by maximizing entropy, then the spectral density. With the entropy defined in Eq. (3), the alternative way to derive the maximum entropy theory consists of the following steps: (1) construct the constraints, (2) determine the probability density function, (3) determine the Lagrange multipliers, (4) determine the spectral power, (5) extend the autocorrelation or autocovariance, and (6) forecast streamflow.

3.1. Specification of constraint

The probability density function of spectral power must satisfy:

$$\int_D p(\vec{x}) \mathrm{d}\vec{x} = 1. \tag{13}$$

The other constraints are constructed from the relationship between spectral power and autocovariance. Let S_k denote the expected value of x_k, written as

$$S_k = \int_D x_k p(\vec{x}) \mathrm{d}\vec{x}. \tag{14}$$

It is known that the mean spectral power S_k is the Fourier transform of autocovariance R_r. On the other hand, autocovariance can be expressed as the inverse Fourier transform of the spectral power S_k as

$$R_r = \sum_{k=1}^{n} S_k \exp(2\pi \mathrm{i} r \Delta t f_k) = \sum_{k=-n}^{n} S_k c_{rk} \quad -N \le r \le N \tag{15}$$

where $c_{rk} = \exp(2\pi \mathrm{i} r \Delta t f_k)$, r is the lag, Δt is taken as 1 month for monthly streamflow, $f_k = k/N$, and N is the largest lag for given autocorrelation or autocovariance, usually taken as 1/4 or 1/2 of the streamflow length T. Substituting Eq. (14) into

Eq. (15), the autocovariance is determined from the probability density function as

$$R_r = \int_D \sum_{k=1}^{n} x_k c_{rk} p(\vec{x}) \mathrm{d}\vec{x}, \quad -N \le r \le N. \tag{16}$$

Thus, the autocovariance function from lag $-N$ to N with Eq. (16) is considered to consist of $2N+1$ constraints for applying the maximum entropy spectral analysis

3.2. Determination of distribution of spectral power

The probability density function is computed by maximizing entropy using the method of Lagrange multipliers. Using constraints in Eqs. (13) and (16), the Lagrangian function $L(p)$ can be written as

$$L(p) = -\int p(\vec{x}) \ln[p(\vec{x})] \mathrm{d}\vec{x} + (\lambda_0 - 1)\left[\int_D p(\vec{x}) \mathrm{d}\vec{x} - 1\right] + \sum_{r=-N}^{N} \lambda_k \left[\int_D \sum_{k=1}^{n} x_k c_{rk} p(\vec{x}) \mathrm{d}\vec{x} - R_r\right] \tag{17}$$

where λ_r are the Lagrange multipliers. Taking the partial derivative of the Lagrangian function with respect to $p(\vec{x})$, and equating the derivative to zero, one obtains

$$\frac{\partial L(p)}{\partial p(\vec{x})} = \ln[p(\vec{x})] + \lambda_0 + \sum_{r=-N}^{N} \lambda_r \sum_{k=1}^{n} x_k c_{rk} = 0. \tag{18}$$

Thus, the least-biased probability distribution of spectral power becomes

$$p(\vec{x}) = \exp\left[-\lambda_0 - \sum_{r=-N}^{N} \lambda_r \sum_{k=1}^{n} x_k c_{rk}\right]. \tag{19}$$

3.3. Determination of Lagrange multipliers

It is noted from Eq. (19) that the spectral power estimated by maximizing the entropy involves Lagrange multipliers. To compute the Lagrange multipliers, Eq. (19) is substituted into constraints Eqs. (13) and (16). Then, one obtains

$$\int_D \exp\left[-\lambda_0 - \sum_{r=-N}^{N} \lambda_r \sum_{k=1}^{n} x_k c_{rk}\right] \mathrm{d}\vec{x} = 1 \tag{20}$$

and

$$R_r = \int_D \exp\left[-\lambda_0 - \sum_{r=-N}^{N} \lambda_r \sum_{k=1}^{n} x_k c_{rk}\right] \sum_{k=1}^{n} x_k c_{rk} \mathrm{d}\vec{x}. \tag{21}$$

It is shown that $2N+2$ nonlinear equations need to be solved for computing the Lagrange multipliers.

3.4. Determination of spectral power

The expected spectral power can be determined by

$$S_k = \int x_k p(x_k) \mathrm{d}x_k = \int x_k \exp\left(-\frac{\lambda_0}{n} - \sum_{r=-N}^{N} \lambda_r x_k c_{rk}\right) \mathrm{d}x_k. \tag{22}$$

Let $A = \exp(-\lambda_0)$. Then, Eq. (19) can be written as

$$p(\vec{x}) = A \exp\left[-\sum_{r=-N}^{N} \lambda_r \sum_{k=-n}^{n} x_k c_{rk}\right] = A \exp\left[-\sum_{k=-n}^{n} \left(\sum_{r=-N}^{N} \lambda_r c_{rk}\right) x_k\right]. \tag{23}$$

Integrating Eq. (23) for x_k over 0 to infinity yields

$$\begin{aligned} 1 &= \int_0^{\infty} p(\vec{x}) \mathrm{d}\vec{x} = \iiint A \exp\left[-\sum_{k=-n}^{n} \left(\sum_{r=-N}^{N} \lambda_r c_{rk}\right) x_k\right] \mathrm{d}x_k \\ &= A \prod_{k=1}^{n} \int \exp\left(-\left(\sum_{r=-N}^{N} \lambda_r c_{rk}\right) x_k\right) \mathrm{d}x_k = A \prod_{k=1}^{n} \frac{1}{\left(\sum_{r=-N}^{N} \lambda_r c_{rk}\right)}. \end{aligned} \tag{24}$$

$$\text{Thus, } A = \prod_{k} \left(\sum_{r=-N}^{N} \lambda_r c_{rk}\right). \tag{25}$$

Inserting Eq. (25) in Eq. (23) yields a multi-variate exponential distribution as

$$p(\vec{x}) = \prod_{k=-n}^{n} \left(\sum_{r=-N}^{N} \lambda_r c_{rk} \right) \exp \left[- \left(\sum_{r=-N}^{N} \lambda_r c_{rk} \right) x_k \right]. \tag{26}$$

Thus, the spectral power can be solved for by inserting Eq. (26) into Eq. (22), which yields

$$S_k = \frac{1}{\sum_{r=-N}^{N} \lambda_r c_{rk}} \tag{27}$$

where the Lagrange multipliers are computed from solving Eqs. (20) and (21).

3.5. Extension of autocovariance

It is seen from Eq. (27) that the spectral power derived by CESAS is in the form of inverse polynomials similar to BESA shown in Eq. (7). The only difference is that Eq. (7) was normalized to spectral density, while Eq. (27) is in the form of spectral power. Thus, Eq. (27) is also equivalent to the spectral power of the AR process in Eq. (8), which can be written as

$$S_k = \frac{1}{\sum_{r=-N}^{N} \lambda_r c_{rk}} = \frac{1}{\sum_{r=-N}^{N} \lambda_r z^{-r}} = \frac{\Delta^2}{\left| 1 + \sum_{n=0}^{N} a_n z^{-n} \right|^2} \tag{28}$$

where $z = \exp(-2\pi \mathrm{i} r \Delta t f_k)$, a_n are the forecasting coefficients, and Δ^2 is the gain satisfying the Yule–Walker equation [34]. Eq. (28) shows that the spectral power obtained by CESAS satisfies the form of linear prediction as BESA did. Thus, the denominator in Eq. (28) becomes

$$\begin{aligned} \sum_{r=-N}^{N} \lambda_r z^{-r} &= \Delta^{-2} \left[1 + a_1 z + a_2 z^2 + \cdots + a_N z^Z \right] \left[a + a_1^* z^{-1} + a_2^* z^{-2} + \cdots + a_N^* z^{-N} \right] \\ &= \Delta^{-2} \sum_{s=0}^{N} a_s z^s \sum_{s=0}^{N} a_s^* z^{-s} = \Delta^{-2} \sum_{r=-N}^{N} \sum_{k=0}^{N-r} a_{k+r} a_k^* z^r. \end{aligned} \tag{29}$$

From Eq. (29), the Lagrange multipliers can be expressed by the convolution of forecasting coefficients as

$$\lambda_r = \frac{1}{\Delta^2} \sum_{k=0}^{N-r} a_{k+r} a_k^*. \tag{30}$$

Thus, the autocorrelation can be extended using the coefficients as

$$R_{N+1} = a_1 R_N + a_2 R_{N-1} + \cdots + a_m R_{N+m-1}. \tag{31}$$

It is noted that the extension of the autocorrelation is the same as BESA. However, the computation of forecasting coefficients has two procedures for CESAS. First, Lagrange multipliers have to be solved from nonlinear equations (19) and (20). Second, the forecasting coefficients are determined by solving Eq. (30). As a result, it has lower computation efficiency than BESA using the Levinson–Burg algorithm.

3.6. Forecast

It is known that using the coefficients for extending the autocorrelation in Eq. (31) to weigh the time series yields the least squared prediction [10,23]. Thus, for stationary normalized time series, forecasting follows the way of extending the autocorrelation or the autocovariance function as:

$$y_{T+1} = a_1 y_T + a_2 y_{T-1} + \cdots + a_m y_{T+m-1} \tag{32}$$

where m is the order of forecasting model, and is identified by the Akaike information criterion (AIC) or Bayesian information criterion (BIC) [35,36].

4. Application

4.1. Data description

The proposed derivation of configurational entropy spectral analysis was verified using observed streamflows from the US Geological Survey (USGS) website. Five stations were selected from Minnesota River, Upper Mississippi River, Iowa River, Des Moines River and Illinois River, which are listed in Table 1 along with their drainage areas and locations.

Table 1
Selected stations from Mississippi River watershed.

Location	Station	Area (km^2)	Latitude	Longitude
Minnesota River	05301000	10 489	45°01′17″	95°52′05″
Upper Mississippi River	05420500	221 704	41°46′50″	90°15′07″
Iowa River	05449500	1 111	42°45′36″	93°37′18″
Des Moines River	05476000	3 237	43°37′06″	94°59′05″
Illinois River	05543500	21 391	41°19′37″	88°43′03″

Table 2
Streamflow characteristics.

Location	Period	Mean (m^3/s)	Standard deviation (m^3/s)	Peak (m^3/s)	Peak/mean
Minnesota River	2005–2010	67.0	15.8	105.4	1.6
Upper Mississippi River	1994–2000	1693.8	566.1	2877.5	1.7
Iowa River	2002–2009	9.3	4.8	26.8	2.9
Des Moines River	2002–2009	13.7	9.1	37.0	2.7
Illinois River	2006–2011	375.1	86.2	531.4	1.4

Table 3
Prediction coefficient estimated using BESA and CESA.

Prediction coefficient	BESA	CESAS	Difference
a_0	1.000	1.000	0.000
a_1	0.462	0.462	0.000
a_2	0.182	0.181	−0.001
a_3	−0.043	−0.043	0.000
a_4	0.153	0.153	0.000
a_5	−0.238	−0.237	0.001
a_6	0.039	0.039	0.000
a_7	0.068	0.068	0.000
a_8	−0.078	−0.078	0.000
a_9	−0.033	−0.034	−0.001
a_{10}	−0.055	−0.054	0.001
a_{11}	0.036	0.036	0.000
a_{12}	0.686	0.686	0.000

The selected five stations are distributed over the whole Mississippi River watershed, and the drainage area varies from 1111 km^2 to 221,704 km^2. Thus, streamflow characteristics are quite different from station to station. Basic statistics of streamflow of Mississippi River are listed in Table 2 that show that averaged monthly streamflow discharge varies from 9 m^3/s to over 1700 m^3/s and the standard deviation varies from 4.8 m^3/s to 570 m^3/s. The peak streamflow represents the average of yearly maximum monthly streamflow, which is shown to be 1.4–2.9 times the mean streamflow.

4.2. Parameter estimation

The coefficients of prediction by CESAS and BESA theories are computed for the Minnesota River, and are listed in Table 3. The prediction coefficients of BESA were estimated from the Levinson–Burg algorithm, while the prediction coefficients of CESAS were obtained by solving Eq. (30) numerically. As shown in the table, the difference was minimal for the two methods, which is less than 0.001. However, the computing speed was faster using the Levinson–Burg algorithm for BESA. The recursive Levinson-algorithm is more efficient, as it involves the order of N^2 operations with memory storage on the order of N compared to the order of N^3 operations by the Newton–Raphson method for solving the nonlinear equations.

4.3. Results and comparison

Streamflow was forecasted using Eq. (30) by CESAS. Fig. 1 plots the forecasted streamflow in the Mississippi watershed with 90% confidence intervals. The forecasting lead year varied from 1 year to 3 years based on the characteristics of streamflow. For rivers like the Upper Mississippi River with sharp repeated peaks every 12 months, CESAS was capable to forecast with high r^2 over 0.9 for up to a 3 year lead time. But for the Minnesota River, where peak streamflow was less significant, CESAS only forecasted for a 1 year lead time with r^2 of 0.766.

Though forecasted streamflow did not fit the observed values exactly, most of the observed values fell inside the 90% confidence intervals. For example, the mean of forecasted streamflow of Minnesota River turned out as 63.1 m^3/s, which was 5.7% less than the observed value. The peak flow in April was 93.4 m^3/s, which was 11.3% lower than the observed value. Nevertheless, all the observations fell between the upper and lower 90% confidence intervals, as shown in Fig. 1. However, there is an exception for Iowa River. Streamflow in Iowa had an unexpected peak streamflow of 26.8 m^3/s during the second

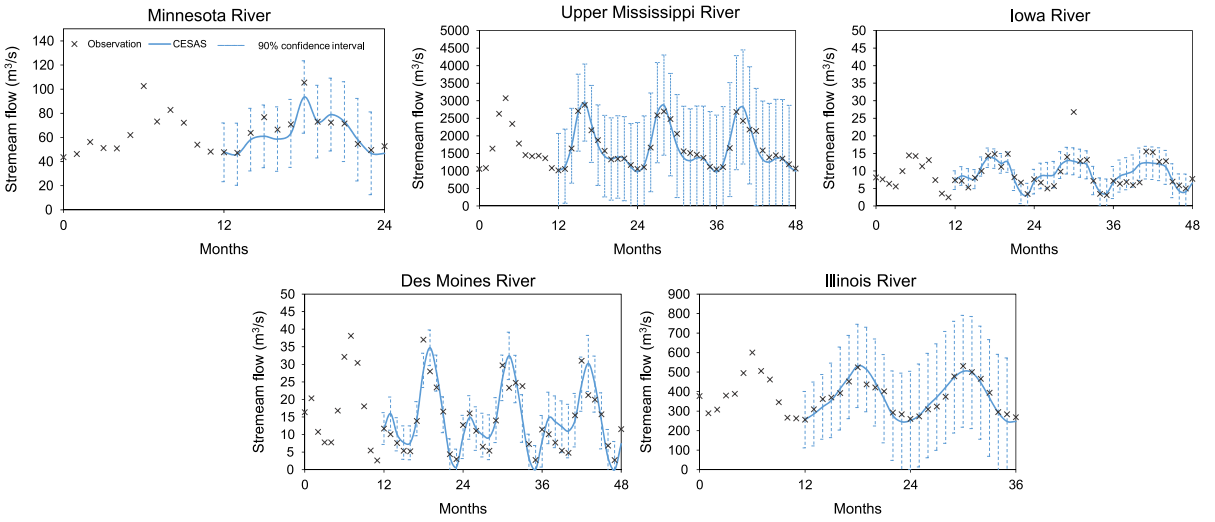

Fig. 1. Forecasted streamflow with 90% confidence intervals.

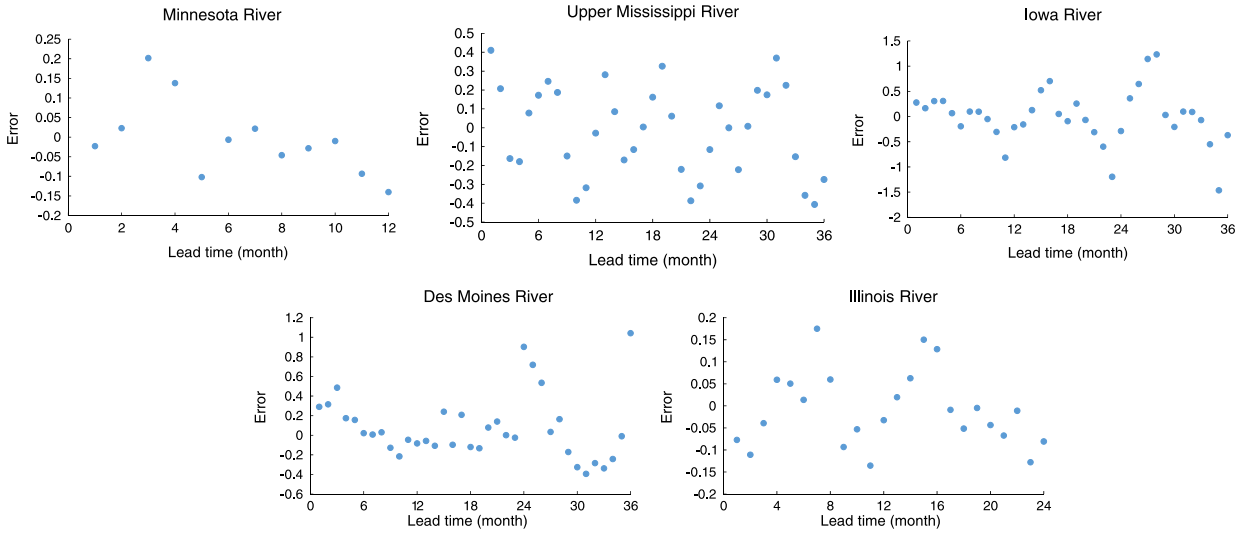

Fig. 2. Forecasted errors using CESAS.

lead year, which exceeded by 38.5% over the past peak streamflow. In this case, the forecasted streamflow was 12.8 m^3/s and the upper 90% was 17.1 m^3/s, still smaller than the observed value. It implies that forecasted streamflows with CESAS were not able to capture the irregular changes in the time series and may miss some unexpected large peaks.

On the other hand, CESAS was not good at forecasting streamflow in low flow season. It can be seen from Des Moines River plotted in Fig. 1 that streamflow in this river does not monotonically decrease after the peak, but there is another small peak during the low flow season. In this case, CESAS forecasted streamflow higher than observation, and the differences between the observed values and the forecasted values became larger as the lead time increased. As a result, the observed values fitted the lower 90% confidence intervals for the third lead year as shown in Fig. 2.

Forecasted errors were computed and plotted versus the lead time in Fig. 2. It is seen from the figure that errors for the Upper Mississippi River had the most random pattern, which suggested that forecasting by CESAS for this river was consistently satisfactory during the lead time of 3 years. However, forecasted errors for Iowa River and Des Moines River increased over time, which suggests that CESAS would not be valid for longer lead time forecasts for these rivers.

The forecasted results are verified using the relative error (*RE*), root mean squared error (*RMSE*), the coefficient of determination (r^2) and Nash–Sutcliffe efficient (*NSE*) for all five stations, and are summarized in Table 4. Results showed good forecasting by the proposed CESAS. The *RE* values were around 0.1, which means that the forecasting error was around 10%. The r^2 values for all the rivers were above 0.7, and r^2 was even higher than 0.9 for Upper Mississippi River. Besides, the *NSE* values for all cases were higher than 0.4.

Table 4
Measures of forecasting results for five stations.

Location	*RE*	*RMSE*	r^2	*NSE*
Minnesota River	0.083	7.320	0.766	0.484
Upper Mississippi River	0.072	163.578	0.914	0.737
Iowa River	0.129	3.091	0.872	0.675
Des Moines River	0.106	4.661	0.729	0.576
Illinois River	0.069	30.218	0.872	0.658

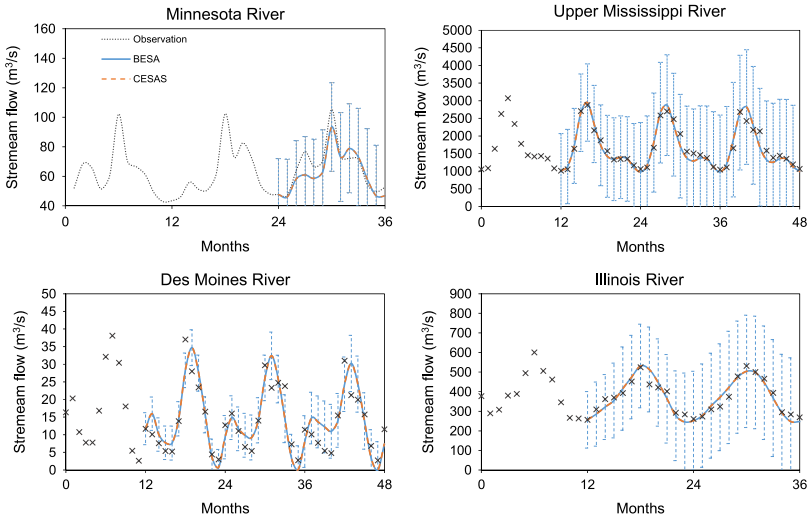

Fig. 3. Streamflow forecasted using the CESAS and BESA methods.

The prediction coefficients obtained by BESA and CESAS methods were about the same, with differences less than 0.001%. As a result, streamflows forecasted by the BESA and CESAS coincided, as shown in Fig. 3. No difference was found in forecasting streamflow for an accuracy of 0.001 m^3/s for rivers from the Mississippi watershed and BESA and CESAS shared the same confidence intervals as shown in Fig. 3. The drawbacks of CESAS in forecasting irregular flow of dry season in Des Moines River and non-normal streamflow peak in Iowa River also occurred in the BESA forecasting. This suggests that the derived CESAS was identical to BESA.

5. Conclusion

The configurational entropy spectral analysis was developed with spectral power as a random variable. Thus, the maximum entropy spectral analysis encompasses the Burg entropy spectral analysis (BESA) and two configurational entropy spectral analyses (CESAS and CESAF). The developed CESAS is shown to be identical to the entropy spectral analysis developed by Burg [9,10]. The derived CESAS is examined using streamflow data obtained from the Mississippi Watershed, and yields the same results as those estimated from BESA. Besides, CESAS has the same drawbacks as BESA in forecasting streamflow in low flow season and unexpected peak flows. However, the prediction coefficients are estimated by solving nonlinear equations, and the computing speed is slower than the Levinson–Burg algorithm. To sum up, CESAS and BESA can be seen as similar, and the Levinson–Burg algorithm for computing the parameters is recommended.

References

[1] C.E. Shannon, A mathematical theory of communication, At&T Tech. J. 27 (4) (1948) 623–656.
[2] E.T. Jaynes, Information theory and statistical mechanics. 2, Phys. Rev. 108 (2) (1957) 171–190.
[3] E.T. Jaynes, Information theory and statistical mechanics. 1, Phys. Rev. 106 (4) (1957) 620–630.
[4] V.P. Singh, The use of entropy in hydrology and water resources, Hydrol. Process. 11 (6) (1997) 587–626.
[5] V.P. Singh, Hydrologic synthesis using entropy theory: review, J. Hydrol. Eng. 16 (5) (2011) 421–433.
[6] V.P. Singh, Entropy Theory and its Application in Environmental and Water Engineering, John Wiley & Sons, 2013, pp. 436–491.
[7] V.P. Singh, Entropy Theory in Hydrologic Science and Engineering, McGraw-Hill Education, New York, 2015, pp. 557–598.
[8] V.P. Singh, S.K. Jain, A. Tyagi, Entropy theory and its applications in risk analysis, Risk Reliab. Anal. (2007) 356–391.
[9] J.P. Burg, Maximum entropy spectral analysis, in: Proceedings of 37th Meeting of Society Exploration Geophysics, 1967, pp. 34–41.
[10] J.P. Burg, Maximum entropy spectral analysis (Ph.D. thesis), Stanford University, 1975, 123 pp..
[11] J.A. Edward, M.M. Fitelson, Notes on maximum-entropy processing, IEEE Trans. Inform. Theory 19 (2) (1973) 3.
[12] R.G. Currie, Geomagnetic line spectra-2 to 70 years, Astrophys. Space Sci. 21 (1973) 14.
[13] N.R. Dalezios, P.A. Tyraskis, Maximum-entropy spectra for regional precipitation analysis and forecasting, J. Hydrol. 109 (1–2) (1989) 25–42.

[14] M. Ghil, M.R. Allen, M.D. Dettinger, K. Ide, D. Kondrashov, M.E. Mann, A.W. Robertson, A. Saunders, Y. Tian, F. Varadi, P. Yiou, Advanced spectral methods for climatic time series, Rev. Geophys. 40 (1) (2002).
[15] H.M. Hasanean, Fluctuations of surface air temperature in the Eastern Mediterranean, Theor. Appl. Climatol. 68 (1–2) (2001) 75–87.
[16] G. Padmanabhan, A.R. Rao, Maximum-entropy spectral-analysis of hydrologic data, Water Resour. Res. 24 (9) (1988) 1519–1533.
[17] E. Pardo-Iguzquiza, F.J. Rodriquez-Tovar, Maximum entropy spectral analysis of climatic time series revisited: Assessing the statistical significance of estimated spectral peaks, J. Geophys. Res.-Atmos. 111 (D10) (2006).
[18] Y.F. Sang, D. Wang, J.C. Wu, Q.P. Zhu, L. Wang, The relation between periods' identification and noises in hydrologic series data, J. Hydrol. 368 (1–4) (2009) 165–177.
[19] Y.F. Sang, Z.G. Wang, C.M. Liu, Period identification in hydrologic time series using empirical mode decomposition and maximum entropy spectral analysis, J. Hydrol. 424 (2012) 154–164.
[20] I. Tosic, M. Unkasevic, Analysis of precipitation series for Belgrade, Theor. Appl. Climatol. 80 (1) (2005) 67–77.
[21] D. Wang, Y.-F. Chen, G.-F. Li, Y.-H. Xu, Maximum entropy spectral analysis for annual maximum tide levels time series of the changjiang river estuary, J. Coast. Res. 43 (2004) 101–108.
[22] P.F. Krstanovic, V.P. Singh, A univariate model for long-term streamflow forecasting. 1. Development, Stoch. Hydrol. Hydraul. 5 (3) (1991) 173–188.
[23] P.F. Krstanovic, V.P. Singh, A univariate model for long-term streamflow forecasting. 2. Application, Stoch. Hydrol. Hydraul. 5 (3) (1991) 189–205.
[24] P.F. Krstanovic, V.P. Singh, A real-time flood forecasting model based on maximum-entropy spectral analysis: I. Development, Water Resour. Manag. 7 (2) (1993) 109–129.
[25] P.F. Krstanovic, V.P. Singh, A real-time flood forecasting model based on maximum-entropy spectral analysis: II. Application, Water Resour. Manag. 7 (2) (1993) 131–151.
[26] B.R. Frieden, Restoring with maximum likelihood and maximum entropy, J. Opt. Soc. Amer. 62 (4) (1972) 511-518.
[27] S.F. Gull, G.J. Daniell, Image-reconstruction from incomplete and noisy data, Nature 272 (5655) (1978) 686–690.
[28] C. Nadeu, E. Sanvicente, M. Bertran, A new algorthm for spectral esimation, in: International Conference on Digital Signal Processing, Florence, Italy, 1981, pp. 463–470.
[29] H. Cui, V.P. Singh, Configurational entropy theory for streamflow forecasting, J. Hydrol. 521 (0) (2015) 1–17.
[30] D.A. Gray, Maximum entropy spectrum analysis technique—A review of its theoretical properties, Weapons Research Establishment, 1977.
[31] N. Levinson, The Wiener rms (root mean square) error criterion in filter design and prediction, J. Math. Phys. Camb. 25 (4) (1946) 261–278.
[32] A. Papoulis, Maximum-entropy and spectral estimation—a review, IEEE Trans. Acoust. Speech 29 (6) (1981) 1176–1186.
[33] A. Van Den Bo, Alternative interpretation of maximum entropy spectral analysis, IEEE Trans. Inform. Theory 17 (4) (1971) 493–494.
[34] G.U. Yule, On a method of investigating periodicities in disturbed series, with special reference to Wofer's sunspot numbers, Philos. Trans. R. Soc. Lond. 226 (1927) 267–298.
[35] G.E.P. Box, G.M. Jenkins, Time Series Analysis; Forecasting and Control. Holden-Day Series in Time Series Analysis, Holden-Day, San Francisco, 1970, 553 pp..
[36] K.W. Hipel, A.I. McLeod, Time series modelling of water resources and environmental systems, in: Developments in Water Science, vol. 45, Elsevier, Burlington, 1994, p. 1. online resource (1053 p.).