A HISTOGRAM MATCHING APPROACH FOR ASSESSMENT OF FLOW REGIME ALTERATION: APPLICATION TO ENVIRONMENTAL FLOW OPTIMIZATION

JENQ-TZONG SHIAU[a] and FU-CHUN WU[b*]

[a] *Department of Hydraulic and Ocean Engineering, National Cheng Kung University, Tainan 701, Taiwan*
[b] *Department of Bioenvironmental Systems Engineering, Hydrotech Research Institute and Center for Ecological Engineering, National Taiwan University, Taipei 106, Taiwan*

ABSTRACT

In this paper we present a Histogram Matching Approach (HMA) for assessment of the flow regime alteration. The HMA uses the degree of histogram dissimilarity as a metric for impact assessment, which is based on the quadratic-form distance between the frequency vectors of the pre- and post-impact histograms weighted by a specified similarity matrix. The HMA is coupled with an aggregated multiobjective optimization genetic algorithm and applied to a case study on the Kaoping diversion weir (Taiwan) for determining the optimal environmental flow scheme that balances the ecosystem and human needs objectives. Two key issues are addressed in this study. First, we compare the performances of the HMA and existing Range of Variability Approach (RVA). Second, we employ three types of similarity function to investigate their effect on the outcomes of the HMA. The results reveal that the HMA consistently outperforms the RVA in preserving the natural flow variability regardless of what type of similarity function is used. No single type of similarity function can be found that would simultaneously best preserve the natural patterns of 32 Indicators of Hydrologic Alteration (IHA). For the situations where the water-supply reliability is of critical concern, the pulse similarity is recommended because it would assure the smallest water-supply deficit. If, however, minor degradation in the water-supply reliability may be overlooked, the linear similarity is suggested because it would generally result in the post-impact flows that most satisfactorily resemble to the natural flow regime. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: Histogram Matching Approach (HMA); Range of Variability Approach (RVA); Indicators of Hydrologic Alteration (IHA); natural flow regime; quadratic-form distance; dissimilarity metric

Received 14 May 2007; Revised 28 October 2007; Accepted 5 December 2007

INTRODUCTION

Determining the instream flow requirements (or alternatively termed 'environmental flows') for riverine ecosystems continues to be a challenge for contemporary scientists and natural resource managers. A key issue regarding the definition of environmental flows is to determine how much of the original flow regime should continue to flow down a river and onto its floodplains in order to maintain the valued features of an ecosystem. Over the last decade, the concept of 'natural flow regime' has been emerging as a paradigm for river conservation and restoration (Poff *et al.*, 1997), which recognizes the full range of natural flow variability as a primary driving force for sustaining the ecological health of a river. More recently, researchers have started to assess the ecologically important components of the flow regime (e.g. Wood *et al.*, 1999; Lake, 2003; Monk *et al.*, 2006). Olden and Poff (2003) reviewed over 200 hydrologic indices for characterizing natural and altered flow regimes. Among many others, Richter *et al.* (1996) adopted 32 hydrologic parameters to develop a suite of Indicators of Hydrologic Alteration (IHA). The 32 IHA are categorized by five groups of hydrologic features, that is, flow magnitude, duration, timing, frequency and rate of change (Table I). Each of the IHA is ecologically relevant when considering rivers in general. For a particular river, species or period of time, however, some of the IHA may be more relevant and some may be not. For example, Monk *et al.* (2006), based on their paired hydroecological data series, pointed

---

Table I. Indicators of Hydrologic Alteration (IHA) used in the RVA and HMA

IHA group	Hydrologic parameter [units]
Group 1: Magnitude of monthly flows	Mean flow of each calendar month [m^3/s]
Group 2: Magnitude and duration of annual extreme flows, and base flow condition	Annual 1-day minimum flow [m^3/s]
	Annual 1-day maximum flow [m^3/s]
	Annual 3-day minimum flow [m^3/s]
	Annual 3-day maximum flow [m^3/s]
	Annual 7-day minimum flow [m^3/s]
	Annual 7-day maximum flow [m^3/s]
	Annual 30-day minimum flow [m^3/s]
	Annual 30-day maximum flow [m^3/s]
	Annual 90-day minimum flow [m^3/s]
	Annual 90-day maximum flow [m^3/s]
	Base flow condition (Annual 7-day minimum flow divided by annual mean flow) [–]
Group 3: Timing of annual extreme flows	Julian date of annual 1-day minimum flow
	Julian date of annual 1-day maximum flow
Group 4: Frequency and duration of high and low pulses[*]	Number of high pulses in each year
	Number of low pulses in each year
	Mean duration of high pulse [days]
	Mean duration of low pulse [days]
Group 5: Rate and frequency of flow changes	Rise rate (Mean of all positive differences between consecutive daily flows) [$m^3/s/day$]
	Fall rate (Mean of all negative differences between consecutive daily flows) [$m^3/s/day$]
	Number of flow reversals

A total of 32 IHA are categorized by five groups of hydrologic features.
[a]High and low pulses are those periods in which the daily flows are above the 75th- and below the 25th-percentile pre-impact daily flows, respectively.

out that variables from IHA groups 1 and 2 are the best predictors of macroinvertebrate community response to flow. Thoms and Parsons (2003) presented a multidimensional temporal–spatial approach to advancing the application of the natural flow paradigm. Poff *et al*. (2006) suggested a hierarchical approach for comparing hydrologic variability across different spatial scales (i.e. basins, regions and continents).

The IHA have become increasingly popular tools for river research and management. Abundant examples can be found where researchers have employed the IHA to assess the hydrologic changes induced by anthropogenic and natural processes, as demonstrated by the IHA Applications Database (The Nature Conservancy, 2005). Despite the popularity of IHA, studies that incorporated the natural flow regime to optimizing water release strategies were rarely reported, primarily due to lack of a widely accepted suite of methods for quantifying the ecological fitness of environmental flows. Although previous attempts have been made to quantify ecological fitness using physical habitat or population sizes (Sale *et al*., 1982; Cardwell *et al*., 1996; Jager and Rose, 2003), they were mainly concerned with fisheries and not the maintenance of biodiversity and ecosystem integrity associated with the natural flow regime. Only recently, has the incorporation of the regime-based environmental flows in water resources and ecosystem management become practicable with the aid of the Range of Variability Approach (RVA) (Shiau and Wu, 2004a, 2004b, 2006, 2007a, 2007b).

The RVA employs the natural (or pre-impact) flow series to establish the IHA target ranges (Richter *et al*., 1997). The management goal is to recommend environmental flow schemes that would attain the target ranges as frequently as the natural flow series, and promotions of the ecological health are expected as such a goal is approached. Richter *et al*. (1998) further suggested an IHA target range bracketed by the 25th- and 75th-percentile values, an implication of which is that 50% of the pre-impact years would have the values of the hydrologic
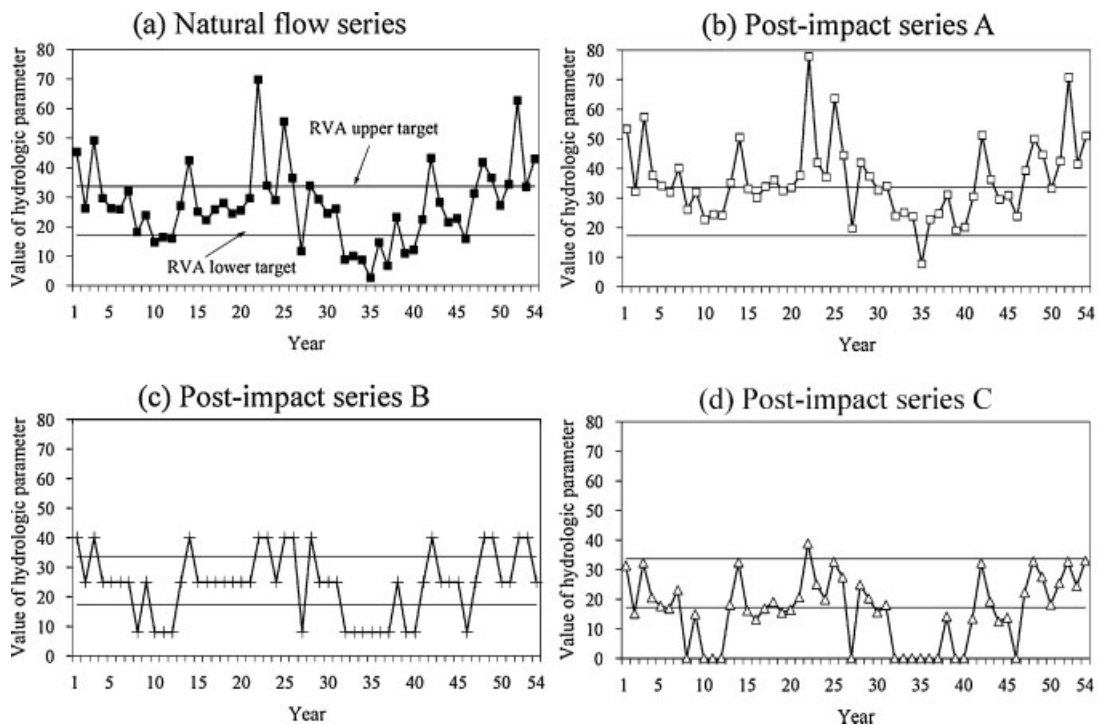
Figure 1. Natural flow series and three hypothetical post-impact flow series

parameter within the target range. As an illustrative example, we show in Figure 1a the natural time series of a hydrologic parameter (in this instance, the monthly flows of January during 1951–2004 at Lilin Bridge gauge station upstream of the Kaoping diversion weir, southern Taiwan). As shown, the 25th- and 75th-percentile values are 17.3 and 33.7 m³/s, respectively, and by definition there are 50% (=27 years) of the parameter values falling in the target range, while the remaining 50% are evenly distributed on both sides of the target range.

To evaluate the deviation of the post-impact flow regime from the natural conditions, Richter *et al.* (1998) defined a 'degree of hydrologic alteration' as follows:

$$D_{R,m} = \left| \frac{N_{o,m} - N_e}{N_e} \right| \times 100\%, \, m = 1, \ldots, 32 \tag{1}$$

where $D_{R,m}$ is the degree of alteration for the mth IHA; $N_{o,m}$ is observed number of years whose post-impact values of the mth IHA are within the target range and N_e is expected number of years whose IHA values would fall in the target ranges $p \cdot N_T$, here N_T is total number of post-impact years and $p = 50\%$ by definition. The RVA has been employed in a series of water allocation studies by the authors (Shiau and Wu, 2004a, 2004b, 2006, 2007a, 2007b) to assess the flow regime alteration induced by weir diversions and specify the optimal environmental flows balancing the ecosystem and human needs objectives.

Although the RVA is one of the first approaches aiming to preserve the natural flow regime, it is subject to potential limitations (Gippel, 2001; Richter *et al.*, 2006). Specifically, the RVA only concerns the frequency of a hydrologic parameter falling in the target range. Variations of the parameter value within the target range are not explicitly taken into account. Moreover, the value and frequency of the hydrologic parameter falling beyond the target range, be it above the upper- or below the lower-target, are totally ignored. These could potentially result in false evaluations of the flow regime. To demonstrate this, three hypothetical post-impact flow series are given in Figure 1b–d. All these three have 50% (=27 years) of their parameter values in the target range. According to

Table II. Degree of hydrologic alteration D_R and degrees of histogram dissimilarity D_Q for three hypothetical post-impact flow series shown in Figure 1

Post-impact flow series		D_Q (%) based on different types of similarity function		
	D_R (%)	Linear	Exponential ($\alpha = 5$)	Pulse
Series A	0	25.4	19.9	17.9
Series B	0	46.0	36.7	30.9
Series C	0	23.4	18.7	17.4

The values of D_Q are based on three types of similarity function.

Equation (1), the degree of hydrologic alteration would be zero for each of these three post-impact series (Table II), which appears to indicate a close resemblance to the natural regime. However, series A (Figure 1b) is an upward-shifted version of the natural series, whereas series C (Figure 1d) is a downward-shifted version. For these two, the data ranges have been considerably altered. In series B (Figure 1c), the timing and frequencies of the data falling above the upper-target, within the target range and below the lower-target are identical with those of the natural series, but the full range of natural variability is completely missing given only three constant values of the hydrologic parameter.

The deviations of these post-impact series from the natural flow regime become clear as we compare their frequency histograms (Figure 2). For the natural series, more than 55% of the data are in the third and fourth classes (18–36), which have been shifted to the fourth and fifth classes (27–45) for the post-impact series A; while for the post-impact series C, the data have been redistributed such that nearly 80% are in the first three classes (0–27). The post-impact series B only fall in the first (0–9), third (18–27) and fifth (36–45) classes, with the corresponding relative frequencies being 26%, 50% and 24%, respectively.

In this paper we present a novel Histogram Matching Approach (HMA) for resolving the limited flow variability resulting from the RVA. The HMA uses a dissimilarity metric to assess the flow regime alteration, which is based on the quadratic-form distance between the frequency vectors of the pre- and post-impact histograms weighted by a specified similarity matrix. Using a case study we employ the HMA to determine the optimal environmental flow scheme for the Kaoping diversion weir in Taiwan. The performances of the HMA and RVA are evaluated. A sensitivity analysis is also performed on different similarity functions.
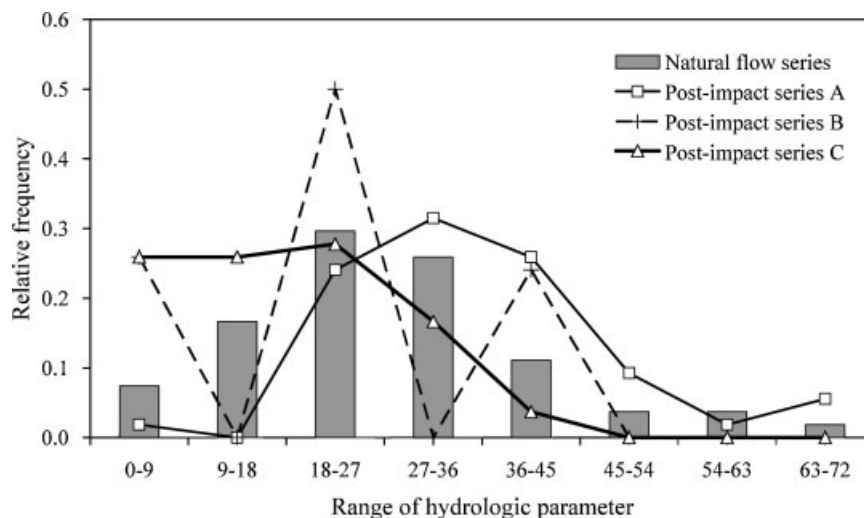

Figure 2. Frequency histograms of natural flow series and three hypothetical post-impact flow series

METHODS

Histogram matching approach (HMA)

The distribution of hydrologic data is typically expressed by a continuous probability function or a discrete histogram. The former is suitable for the situation where a sufficiently large amount of data is available for deriving an unbiased probability distribution. For situations where available data are limited, however, the latter is adopted by partitioning the data space into a predefined number of classes showing the frequency of occurrence in each class. The central idea behind the HMA is that two flow regimes would be similar if their frequency histograms of the 32 IHA resembled to each other, this resemblance can be measured using the 'statistical distances' between the pre- and post-impact frequency histograms. As such, alteration of flow regimes can be assessed with a distance-based dissimilarity metric. The procedure of the HMA is described as follows.

The number of classes must be determined before a histogram can be constructed. A histogram with too large or small a number of classes would be inadequate for exhibiting the pattern of data distribution. Several methods to determine the proper number of classes can be found in the statistics literature (e.g. Scott, 1979; Kottegoda and Rosso, 1997). The following formula is used in this study because it takes into account the total number of data available and characteristics of data distribution:

$$n_c = \frac{r \cdot n^{1/3}}{2 \cdot r_{iq}} \qquad (2)$$

where n_c is number of classes; r is data range which is the difference between the largest and smallest data values; n is total number of data and r_{iq} is inter-quartile range, defined as the difference between the third- and first-quartile values.

A quadratic-form distance is employed herein to measure the dissimilarity between two histograms H and K (Figure 3). This metric, originally proposed by Niblack *et al*. (1993) for colour-based image retrieval, accounts for both the class-by-class correspondence and cross-class information. The quadratic-form distance d_Q is defined as

$$d_Q(H,K) = \sqrt{(|\mathbf{h} - \mathbf{k}|)^T \mathbf{A}(|\mathbf{h} - \mathbf{k}|)} \qquad (3)$$

in which $\mathbf{h} = (h_1, h_2, \ldots, h_{n_c})^T$ and $\mathbf{k} = (k_1, k_2, \ldots, k_{n_c})^T$ are frequency vectors of the histograms H and K, and $|\mathbf{h} - \mathbf{k}|$ is statistical distance vector. The cross-class correspondence is incorporated via a similarity matrix
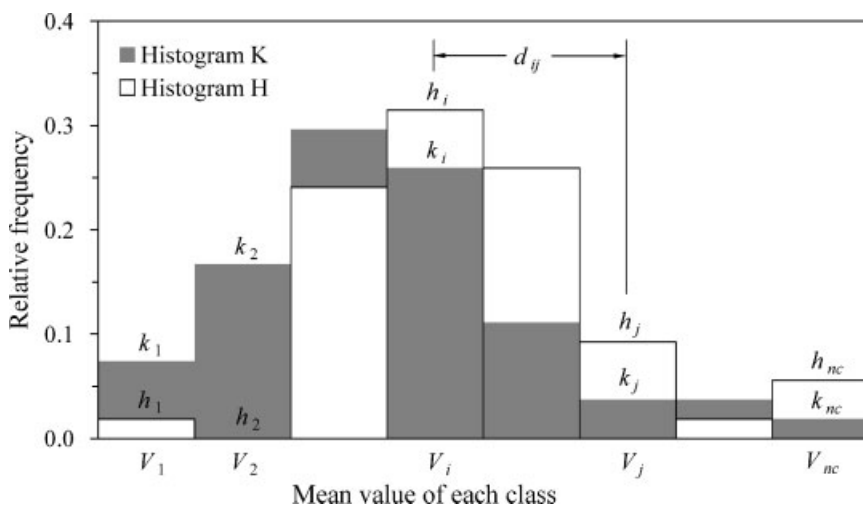

Figure 3. A schematic diagram showing the notations used in the Histogram Matching Approach (HMA)
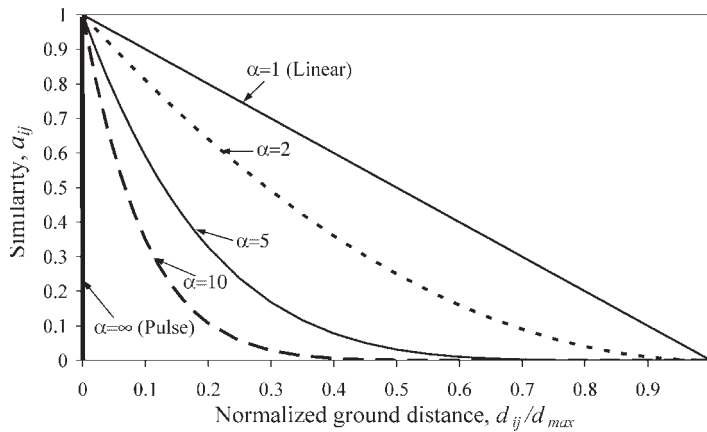
Figure 4. Variations of similarity with normalized ground distance for various α

$\mathbf{A} = [a_{ij}]$, where a_{ij} is similarity between classes i and j. The values of a_{ij} vary between 0 and 1 as a function of the ground distance between classes i and j. A general expression of a_{ij} is given by

$$a_{ij} = \left(1 - \frac{d_{ij}}{d_{\max}}\right)^{\alpha} \tag{4}$$

in which $d_{ij} = |V_i - V_j|$ is ground distance between classes i and j, where V_i and V_j are mean values of classes i and j; $d_{\max} = \max(d_{ij}) = |V_1 - V_{n_c}|$. Values of α are specified in a range between 1 and ∞, a linear similarity function is obtained with $\alpha = 1$, whereas a diagonal pulse similarity matrix that ignores the cross-class correspondence is obtained with $\alpha = \infty$. Variations of a_{ij} with $d_{ij}/d_{\max}$ for a variety of α are illustrated in Figure 4. As shown, the linear and pulse similarities are the envelops of the similarity curve family.

Degree of histogram dissimilarity

To be consistent with the definition of $D_{R,m}$ given in Equation (1), the quadratic-form distance is scaled with its maximum value to define a 'degree of histogram dissimilarity':

$$D_{Q,m} = \frac{d_{Q,m}}{\max(d_{Q,m})} \times 100\%, \quad m = 1, \ldots, 32 \tag{5}$$

where $D_{Q,m}$ is the degree of histogram dissimilarity for the mth IHA; $d_{Q,m}$ is quadratic-form distance of the mth IHA and $\max(d_{Q,m})$ is the maximum value of each $d_{Q,m}$, evaluated by

$$\max(d_{Q,m}) = \sqrt{2 + 2\left(1 - \frac{1}{n_{c,m} - 1}\right)^{\alpha_m}} \tag{6}$$

where $n_{c,m}$ and α_m are values of n_c and α adopted for the mth IHA.

The degrees of histogram dissimilarity for those three post-impact series shown in Figures 1 and 2 are summarized in Table II, where the results associated with the linear, exponential and pulse similarities are obtained with $\alpha = 1$, 5 and ∞, respectively. As it is shown, the values of D_Q for the series B are consistently the greatest no matter which type of similarity is used, indicating the most severely altered flow regime among the three. The values of D_Q for series A and C are very close, with the former being slightly greater than the latter. These results coincide with our intuitive notion of similarity, and eliminate the oddity associated with the corresponding values of D_R (=0).

Overall degree of flow regime alteration

Because individual values of $D_{Q,m}$ may exhibit different levels of dissimilarity, an integrative index is used to evaluate the overall degree of flow regime alteration, that is,

$$D_{OQ} = \left(\frac{1}{32} \sum_{m=1}^{32} D_{Q,m}^2 \right)^{1/2} \tag{7a}$$

where D_{OQ} is the overall degree of 'hydrologic dissimilarity'. Similarly, individual values of $D_{R,m}$ can be integrated as an overall degree of 'hydrologic alteration', that is,

$$D_{OR} = \left(\frac{1}{32} \sum_{m=1}^{32} D_{R,m}^2 \right)^{1/2} \tag{7b}$$

Minimizing D_{OQ} or D_{OR} is regarded as equivalent to best preserving the natural flow regime, thus is taken to be a surrogate objective of ecosystem needs in our case study, as described in the following.

STUDY SITE

Overview of Kaoping diversion weir

The Kaoping diversion weir is located in the midstream of Kaoping Creek, southern Taiwan. The length of the Kaoping Creek is 171 km, with a drainage area total of 3257 km². The alluvial plain of Kaoping Creek is a major agricultural area. To mitigate the impacts of groundwater overdraft and provide an alternative source of water supply, construction of the Kaoping diversion weir was initiated in 1992 and completed in 1999, with a design capacity of 35 m³/s. The monthly flow characteristics (1951–2004) of the Lilin Bridge gauge station, located immediately above the weir site, are given in Table III, where a highly fluctuating and uneven flow pattern is demonstrated. The water-supply objectives of the diversion weir are to meet agricultural and domestic water demands. The projected monthly diversions are also shown in Table III, which withdraw a total of 1064 million m³ each year.

Currently a minimum flow (=9.5 m³/s) is released from the Kaoping diversion weir for protection of downstream water quality (WCA, 2000). However, this environmental flow release does not resemble the natural flow regime (Shiau and Wu, 2006). Since the post-diversion flows vary as a function of the environmental flow prescriptions, a weir operation model is used to simulate the flow series diverted for water supplies and released for ecosystem preservation.

Weir operation model

The system of flows in the weir operation model is depicted in Figure 5, where two flow criteria are to be met at time t, that is, the projected diversions Q_{PD}^t and environmental flows Q_{EF}^t; Q_I^t denotes the natural (or pre-diversion) inflow; Q_{AD}^t denotes the amount of flow actually diverted for water supplies and Q_O^t denotes the post-diversion

Table III. Monthly flow characteristics (1951–2004) and projected monthly diversions at the Kaoping diversion weir (units in m³/s)

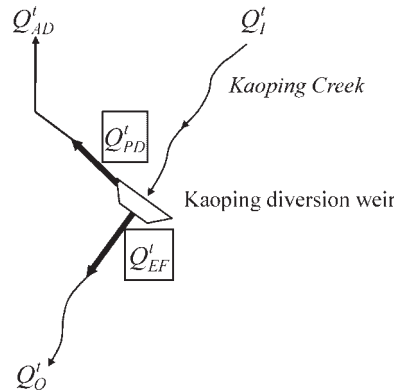|	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Max flow	69.8	131.6	278.8	357.3	786.3	1812.8	1385.4	1958.0	1381.4	544.0	350.2	171.1
Mean flow	26.4	27.3	37.3	57.7	184.9	546.2	466.1	715.4	481.1	182.5	81.9	44.5
Min flow	2.7	1.2	2.3	8.2	13.8	25.1	25.1	31.0	87.7	23.5	18.6	17.5
Projected monthly diversion	22.6	22.5	22.1	22.1	28.1	45.3	45.4	47.5	47.0	47.4	28.4	25.3

Figure 5. Flow system of the Kaoping diversion weir. Flows in the boxes are projected diversions (Q_{PD}^t) and environmental flow prescriptions (Q_{EF}^t); Q_I^t and Q_O^t denote inflows and outflows, respectively; Q_{AD}^t denotes the amount of flow actually diverted for water supplies; superscripts t denote time

outflow. The projected monthly diversions Q_{PD}^t are summarized in Table III, and the values of Q_{EF}^t are the only decision variable to be specified herein. A monthly varying environmental flow scheme is adopted in this study, thus a total of 12 Q_{EF}^t values are to be prescribed. The operational rules of the Kaoping diversion weir are given by

$$
\begin{cases}
Q_O^t = Q_I^t, \ Q_{AD}^t = 0 \ \text{if} \ Q_I^t \leq Q_{EF}^t \\
Q_O^t = Q_{EF}^t, \ Q_{AD}^t = Q_I^t - Q_{EF}^t \ \text{if} \ Q_{EF}^t < Q_I^t \leq Q_{EF}^t + Q_{PD}^t \\
Q_O^t = Q_I^t - Q_{PD}^t, \ Q_{AD}^t = Q_{PD}^t \ \text{if} \ Q_I^t > Q_{EF}^t + Q_{PD}^t
\end{cases}
\tag{8}
$$

where the human demands will be supplied only if the environmental flow criteria are fully met. These operational rules are currently implemented by the Water Resources Agency of Taiwan with a constant value of Q_{EF}^t (=9.5 m^3/s). Here we modify the operational rules by allowing the values of Q_{EF}^t to vary monthly. The daily flows at the Lilin Bridge gauge station (1951–2004) are used in our simulation as the inflow series Q_I^t. The post-diversion series, Q_O^t, are then used to assess the degree of flow regime alteration. The flow series actually diverted for human demands, Q_{AD}^t, is used to evaluate the water supply deficit.

Index of water supply deficit

The shortage ratio, SR, is employed in this study as an index of water supply deficit, which is defined as follows (Cancelliere *et al.*, 1998):

$$
\text{SR} = \frac{\sum\limits_{t=1}^{N} \left| \min(Q_{AD}^t - Q_{PD}^t, 0) \right|}{\sum\limits_{t=1}^{N} Q_{PD}^t} \times 100\%
\tag{9}
$$

where N is total number of days in the simulation period. The value of SR represents a human needs objective to be minimized with the aggregated multiobjective optimization algorithm described below.

Aggregated multiobjective optimization

The operational goal of the Kaoping diversion weir is to supply human demands while retaining the natural flow variability, which formulates a typical multiobjective optimization problem. The objective function can be expressed by

$$
\text{Min}\{D_O, \text{SR}\}
\tag{10}
$$

where D_O is overall degree of flow regime alteration, herein either D_{OQ} or D_{OR} is used. Optimization problems involving multiple conflicting objectives introduce Pareto tradeoff solutions rather than a single optimal solution. More recently, multiobjective evolutionary algorithms (MOEAs) have emerged and become increasingly popular. These MOEAs make use of the population-based approaches to find the Pareto-optimal solutions in one run (Deb, 2001). The authors have employed the nondominated sorting genetic algorithm II (NSGA-II) (Deb *et al.*, 2002) to seek the Pareto-optimal environmental flow schemes that incorporate the inter- and intra-annual variability of the natural flow regime (Shiau and Wu, 2007b).

Because the purpose of this study is to demonstrate the proposed HMA and compare the performances of the HMA and RVA, rather than fully explore the tradeoffs between Pareto solutions, an aggregated multiobjective optimization genetic algorithm (AMOGA) is used to find the optimal solution of a rescaled and aggregated objective function, that is

$$\text{Min}\left[\left(\frac{D_O - D_{O,\text{min}}}{D_{O,\text{max}} - D_{O,\text{min}}}\right)^2 + \left(\frac{\text{SR} - \text{SR}_{\text{min}}}{\text{SR}_{\text{max}} - \text{SR}_{\text{min}}}\right)^2\right]^{1/2} \qquad (11)$$

where $D_{O,\text{max}}$ and $D_{O,\text{min}}$ are maximum and minimum values of D_O and SR_{max} and SR_{min} are maximum and minimum values of SR. The values of $D_{O,\text{max}}$ and SR_{min} are obtained with an extreme condition that $Q_{\text{EF}}^t = 0$ at any time t, whereas $D_{O,\text{min}}$ and SR_{max} correspond to the condition that $Q_{\text{AD}}^t = 0$ at any time t. The optimal set of Q_{EF}^t is obtained using a simple genetic algorithm (Haupt and Haupt, 2004), with the population size of 1000, and crossover and mutation rates 0.8 and 0.05, respectively. The selection, crossover and mutation operators are used to iteratively evolve a population toward the true optimal solution. The procedure is repeated until a stable optimal solution is obtained. The solution so obtained, as revealed by Equation (11), is a compromise between the human and ecosystem needs objectives.

RESULTS

In this section, we present the results obtained with the HMA and RVA. The emphases of our discussion are placed upon two key issues. First, the performances of the HMA and RVA are compared. Second, different types of similarity function are used in the calculation of the quadratic-form distance to examine their effect on the outcomes of the HMA.

Comparing the results of HMA and RVA

The results obtained with the RVA and HMA, including the optimal sets of Q_{EF}^t, optimal values of D_O and SR, and statistics of D_Q and D_R are summarized in Table IV, where the values of D_O correspond to the RVA and HMA are D_{OR} and D_{OQ}, respectively. Note that the values of D_{OR} and D_{OQ} are not to be compared since different formulas and similarity functions are employed in the calculation, but are listed here for the sake of completeness. As shown, these values of D_O are largely on the order of 10% (i.e. between 9.2 and 11.1%), indicating that the overall degrees of flow regime alteration associated with all of these four AMOGA-based optimal environmental flow schemes are relatively low, as compared to those earlier results obtained with the corridor-searching algorithm (Tables III–V, Shiau and Wu, 2007a). Though such a result may have to imply that the RVA is as useful as HMA for finding the optimal environmental flows, the strength of the HMA becomes apparent as we look at the post-optimal series of IHA. Two examples are given below to demonstrate this.

Figure 6a shows the natural and post-optimal series of monthly flows in December, obtained using both the RVA and HMA with a pulse similarity (referred to as HMA-pulse hereinafter). As shown, the post-optimal series based on the HMA-pulse closely resembles to the pattern of natural series, while the post-optimal series based on the RVA exhibits much less variability despite that 50% of the data points are within the target range. Most of these within-target data points obtained with the RVA have values that are just passing the lower target. Moreover, the percentage above the upper target is reduced to 11% while that below the lower target is increased to 39%. The RVA-based post-optimal series shown in Figure 6a is similar to that obtained earlier with the three-class overall degree of hydrologic alteration (Figure 8a, Shiau and Wu, 2007a). The frequency histograms given in

Table IV. Optimal sets of environmental flows Q_{EF}^t obtained with the RVA and HMA, the associated outcomes of D_O and SR, and statistics of D_Q and D_R

				RVA	HMA with different types of similarity function		
					Linear	Exponential ($\alpha = 5$)	Pulse
D_O (%) [D_{OR} or D_{OQ}]				9.3	11.1	9.3	9.2
SR (%)				29.2	34.1	32.1	30.9
Q_{EF}^t (m^3/s)	January			39.7	59.0	40.4	40.5
	February			34.4	52.5	59.5	50.3
	March			27.3	66.0	28.2	30.7
	April			22.8	40.7	28.4	38.3
	May			23.4	27.5	41.0	27.6
	June			26.2	19.3	19.1	26.2
	July			9.6	14.9	13.3	26.2
	August			30.0	26.2	11.8	25.9
	September			9.6	13.3	13.3	22.0
	October			28.3	26.7	26.7	26.7
	November			47.5	65.0	48.7	21.5
	December			27.9	38.9	39.0	55.0
No. of IHA with	$D_Q \leq 5\%$	4[a]	7[b] 9[c]		10	12	14
	$5\% < D_Q \leq 10\%$	10	11 11		10	10	11
	$10\% < D_Q \leq 15\%$	7	7 7		6	7	4
	$15\% < D_Q \leq 20\%$	5	2 1		4	1	1
	$D_Q > 20\%$	6	4 4		2	2	2
No. of IHA with	$D_R \leq 33.3\%$			32	32	32	31
	$33.3\% < D_R \leq 66.7\%$			0	0	0	1
	$D_R > 66.7\%$			0	0	0	0

Superscripts a, b and c associated with the RVA-based results denote the values of D_Q calculated with the linear, exponential and pulse similarities, respectively.

Figure 6b further reveal that the post-optimal series based on the HMA-pulse closely follows the distribution of the natural series, whereas 65% of the RVA-based post-optimal series are concentrated in the second class, which is 37% over the corresponding natural frequency and 24% over the post-optimal frequencies resulting from the HMA-linear and -exponential.

As the second example, the natural and post-optimal series of annual mean low-pulse durations are shown in Figure 7a. The post-optimal series obtained with the HMA-linear exhibits a reasonably good resemblance to the natural series. The RVA-based post-optimal series of low-pulse durations, as it is shown, is not as disturbed as that of monthly flows in December, with most of the RVA-based results being slightly greater than the corresponding values obtained with the HMA-linear. As a consequence, the frequency of the RVA-based post-optimal series in the fourth class (i.e. 120–160 days) is higher than the corresponding HMA-based post-optimal frequencies (Figure 7b), while the RVA-based post-optimal frequency in the first class (i.e. 0–40 days) is consistently lower than the corresponding HMA-based post-optimal frequencies. Nevertheless, both of the AMOGA-based low-pulse duration series shown in Figure 7a follow more closely the natural series than those earlier results obtained with the corridor-searching algorithm (Figure 8b, Shiau and Wu, 2007a).

The above two examples demonstrate that the HMA outperforms the RVA in preserving the natural flow variability, which is achieved at the cost of greater water-supply deficits, as revealed by the slightly greater values of SR associated with the HMA (i.e. ranging from 30.9 to 34.1%) than the RVA-based value of SR (=29.2%). Such minor degradation in human needs fitness, however, may be overlooked as we compare the statistics of D_Q derived from the RVA- and HMA-based post-optimal series of 32 IHA. We show in Table IV the numbers of IHA with the corresponding values of D_Q in five different levels, where the superscripts a, b, and c associated with the RVA-based results denote the values of D_Q calculated with the linear, exponential and pulse similarities, respectively. Comparison of D_Q should be made between the values obtained with the same type of similarity. For $D_Q \leq 5\%$ (the

Table V. Post-optimal values of D_Q for all IHA obtained with HMA using different types of similarity function

IHA group	Hydrologic parameter	D_Q (%) based on different types of similarity		
		Linear	Exponential ($\alpha = 5$)	Pulse
1	Monthly flow of January	1.9	4.2	4.1
	Monthly flow of February	5.1	3.6	3.7
	Monthly flow of March	9.9	10.4	5.9
	Monthly flow of April	9.8	11.6	5.9
	Monthly flow of May	15.1	8.2	11.0
	Monthly flow of June	8.8	5.9	5.2
	Monthly flow of July	12.9	9.3	8.5
	Monthly flow of August	6.1	4.9	4.9
	Monthly flow of September	13.8	10.5	10.5
	Monthly flow of October	30.3	24.1	23.9
	Monthly flow of November	17.1	20.5	25.7
	Monthly flow of December	15.5	11.5	4.9
2	Annual 1-day minimum flow	3.5	1.5	2.6
	Annual 1-day maximum flow	2.2	1.9	1.9
	Annual 3-day minimum flow	1.9	1.5	1.9
	Annual 3-day maximum flow	2.2	1.9	1.9
	Annual 7-day minimum flow	0.0	1.6	2.6
	Annual 7-day maximum flow	6.3	5.0	4.9
	Annual 30-day minimum flow	2.1	3.1	3.2
	Annual 30-day maximum flow	9.2	6.7	6.7
	Annual 90-day minimum flow	4.8	6.1	6.1
	Annual 90-day maximum flow	15.6	11.9	11.3
	Base flow condition	10.2	8.5	7.2
3	Date of annual 1-day maximum flow	0.0	0.0	0.0
	Date of annual 1-day minimum flow	5.4	4.0	3.2
4	Number of high pulses in each year	8.1	6.1	5.9
	Number of low pulses in each year	10.5	8.4	11.0
	Mean duration of high pulse in each year	20.6	15.2	15.6
	Mean duration of low pulse in each year	7.0	7.9	9.3
5	Flow rise rate	3.9	3.0	3.2
	Flow fall rate	14.0	10.4	9.1
	Number of flow reversals	12.7	13.0	8.9

most similar level), the HMA-based post-optimal numbers of IHA are consistently greater than the corresponding RVA-based values regardless of which similarity function is used, while for $D_Q > 20\%$ (the most dissimilar level), the RVA-based post-optimal numbers of IHA are consistently greater than the corresponding HMA-based values. These statistics of D_Q clearly indicate that, compared to the outcomes of the RVA, the post-optimal series of IHA obtained with the HMA preserve more of the natural flow regime via reducing the dissimilarity to the histograms of the natural series, thus offering the improved ecosystem needs fitness with only minor increases in SR.

HMA with different types of similarity function

The effect of similarity function on the outcomes of the HMA is explored in this section. Three types of similarity, that is, linear, exponential ($\alpha = 5$) and pulse, are discussed, among which only the pulse similarity ignores the cross-class correspondence. From the frequency histograms of the monthly flows in December (Figure 6b), we can see that the post-optimal distribution obtained with the HMA-pulse is most similar to the natural pattern, while the identical results obtained with the linear and exponential similarities are inferior. The best outcome associated with the HMA-pulse is attributed to the largest value of Q_{EF}^t prescribed (Table IV), which,
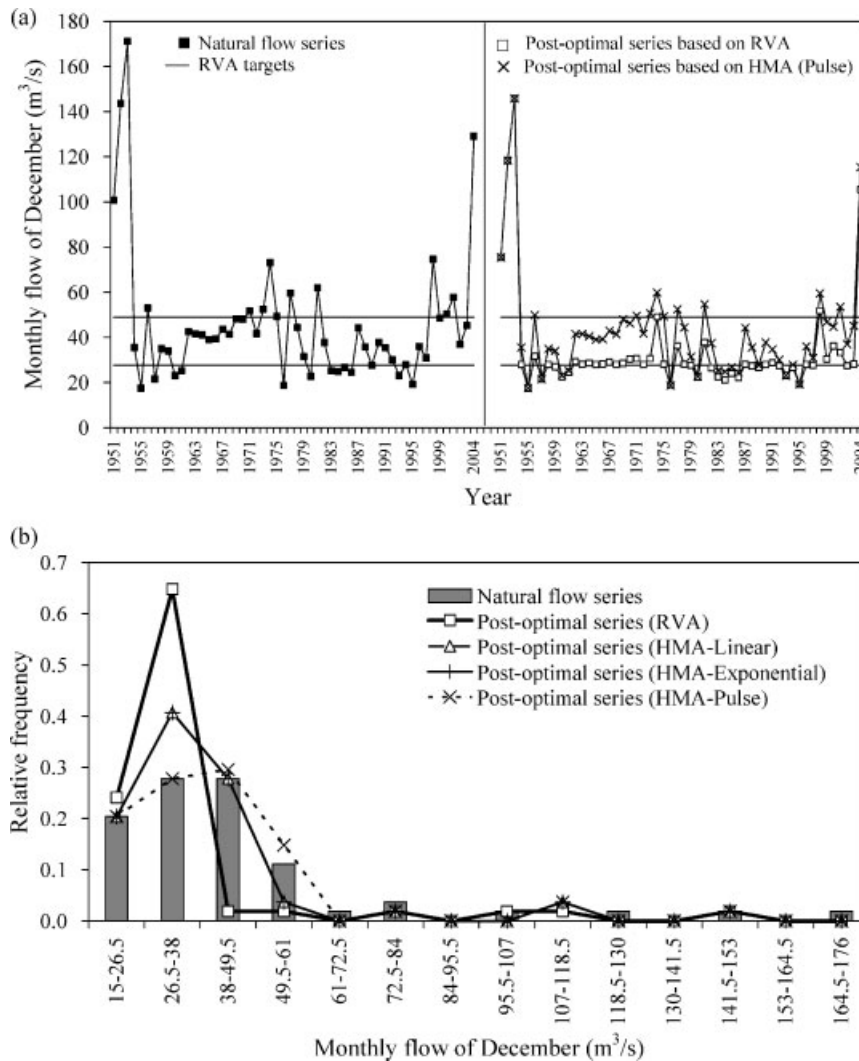
Figure 6. (a) Natural and post-optimal series of monthly flows in December; (b) frequency histograms of natural and post-optimal series of monthly flows in December

however, does not necessarily mean that the environmental flows prescribed with the HMA-pulse are greater than those obtained with other types of similarity function. For example, the value of Q_{EF}^t prescribed for November using the pulse similarity is smallest, resulting in the only value of D_R ranging between 33.3% and 66.7% while others are below 33.3% (Table IV). The frequency histograms of the low pulse durations (Figure 7b), on the other hand, reveal that the post-optimal histogram associated with the HMA-linear is most similar to the natural pattern, whereas that associated with the HMA-pulse is least similar to the natural one.

To further assess the outcomes resulting from different types of similarity, the 32 sets of post-optimal D_Q values obtained with the linear, exponential and pulse similarities are summarized in Table V. Although each set of D_Q values derived for a specific item of IHA cannot be used for ranking the outcomes associated with different types of similarity, these data provide information from which a general conclusion can be drawn. Specifically, there is no single type of similarity function that would make the post-optimal series of 32 IHA simultaneously best retain the natural flow regime. Nevertheless, practical guidelines may be obtained from the finding of this study. For the situations where water-supply reliability is of critical concern, the pulse similarity is recommended because it would assure the smallest water-supply deficit. However, if minor degradation in the water-supply reliability may
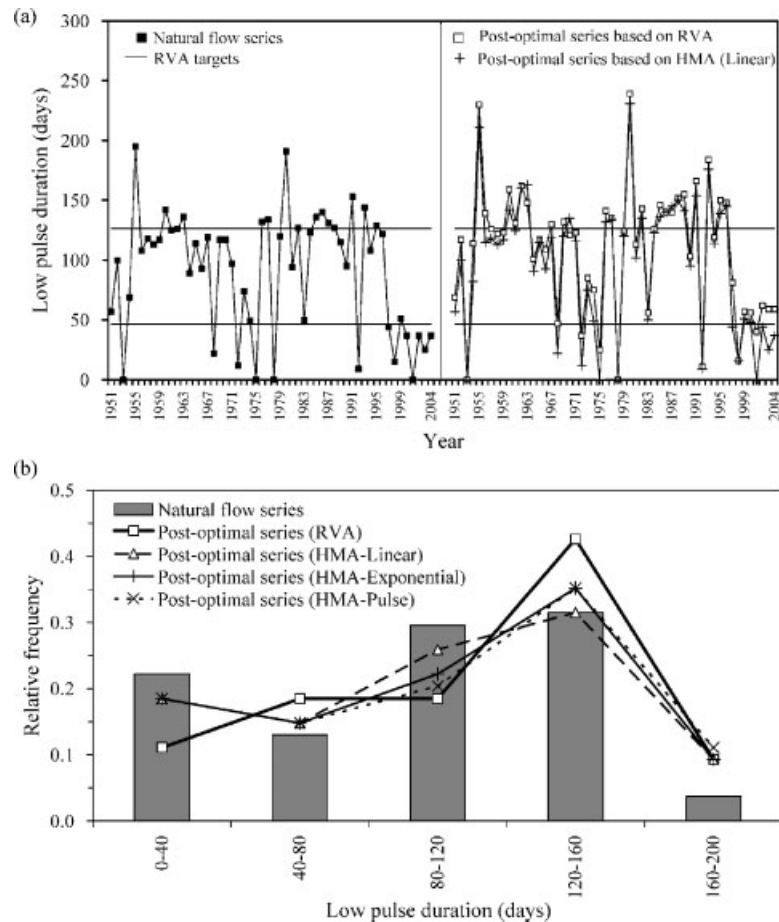
Figure 7. (a) Natural and post-optimal series of mean low-pulse durations; (b) frequency histograms of natural and post-optimal series of mean low-pulse durations

be overlooked, the linear similarity is suggested because it would generally result in the post-impact flows that most satisfactorily resemble to the natural flow regime.

DISCUSSION

In this work we present a novel HMA for assessment of flow regime alteration. The proposed HMA is applied to a case study on the Kaoping diversion weir for determining the optimal environmental flow scheme that balances ecosystem and human needs objectives. The results suggest that the HMA eliminates some shortcomings of the existing RVA, and consistently outperforms it in preserving natural flow variability regardless of which type of similarity function is used. These AMOGA-based results are also superior to those earlier results obtained with the corridor-searching algorithm. Such performances of the HMA are achieved via reducing the dissimilarities to the pre-impact frequency histograms of 32 IHA.

The current approach is aimed at preserving the biodiversity and ecosystem integrity associated with the full range of natural flow regime, thus is of a broader context than those earlier approaches using the physical habitat defined for a specific species or fish population size to quantify the ecological fitness. For situations where some aspects of the flow regime are of vital importance and cannot be compromised, weighting factors may be included in Equation (7) to come up with an optimal solution in favour of some specific IHA. No single type of similarity

function can be found that would simultaneously best retain the natural patterns of 32 IHA. For the situations where water-supply reliability is of critical concern, the pulse similarity can be used to assure the smallest water-supply deficit. If, however, minor degradation in the water-supply reliability may be overlooked, the linear similarity is an ideal candidate because it would generally result in the post-impact flows that most satisfactorily resemble to the natural flow regime.

The proposed HMA may be coupled with the regional frequency analysis (RFA) methodology (Hosking and Wallis, 1997) to evaluate the alteration of a flow regime at ungauged sites. Furthermore, the temporal variations may be incorporated in the HMA to account for the inter-annual variability of the natural flow regime (Shiau and Wu, 2007b). In this paper we demonstrated an HMA-based assessment framework applied to determine the optimal environmental flow scheme. The simulation and optimization of weir operations presented here were based on the historical flow records. For the proposed framework to be used as a planning tool for making environmental flow recommendations into the future, a module for the real-time prediction of inter-annual hydrologic variations should be included. These are potential topics that can be addressed in future studies.

NOTATIONS

$\mathbf{A}$	similarity matrix
a_{ij}	similarity between histogram classes i and j
d_{ij}	ground distance between histogram classes i and j
$d_{\max}$	maximum value of d_{ij}
d_Q	quadratic-form distance
D_O	overall degree of flow regime alteration
D_{OQ}	overall degree of hydrologic dissimilarity
D_{OR}	overall degree of hydrologic alteration
$D_{Q,m}$	degree of histogram dissimilarity for the mth IHA
$D_{R,m}$	degree of hydrologic alteration for the mth IHA
$\mathbf{h}$ and $\mathbf{k}$	frequency vectors of histograms H and K
m	IHA index
n	total number of data
n_c	number of histogram classes
N	total number of days in the simulation period
N_e	expected number of years with IHA values falling in the target ranges
$N_{o,m}$	number of post-impact years with the mth IHA in the target range
N_T	total number of post-impact years
Q_{AD}^t	amount of flow actually diverted for water supplies at time t
Q_{EF}^t	environmental flow at time t
Q_{I}^t	natural (or pre-diversion) inflow at time t
Q_{O}^t	post-diversion outflow at time t
Q_{PD}^t	projected diversion at time t
r	data range
r_{iq}	inter-quartile range
SR	water-supply shortage ratio
α	parameter of similarity function.

REFERENCES

Cancelliere A, Ancarani A, Rossi G. 1998. Susceptibility of water supply reservoirs to drought conditions. *Journal of Hydrologic Engineering* **3**: 140–148.

Cardwell H, Jager HI, Sale MJ. 1996. Designing instream flows to satisfy fish and human water needs. *Journal of Water Resources Planning and Management* **122**: 356–363.

Deb K. 2001. *Multi-objective Optimization using Evolutionary Algorithms*. John Wiley: Hoboken, NJ.

Deb K, Pratap A, Agarwal S, Meyarivan T. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* **6**: 182–197.

Gippel CJ. 2001. Hydrological analyses for environmental flow assessment. In *Proceedings MODSIM 2001, International Congress on Modelling and Simulation, Modelling & Simulation Society of Australia & New Zealand*, Ghassemi F, Whetton P (eds). The Australian National University: Canberra, Australia; 873–880.

Haupt RL, Haupt SE. 2004. *Practical Genetic Algorithms*. John Wiley: Hoboken, NJ.

Hosking JRM, Wallis JR. 1997. *Regional Frequency Analysis: an Approach Based on L-Moments*. Cambridge University Press: New York, NY.

Jager HI, Rose KA. 2003. Designing optimal flow patterns for fall chinook salmon in a central valley, California, river. *North American Journal of Fisheries Management* **23**: 1–21.

Kottegoda NT, Rosso R. 1997. *Statistics, Probability, and Reliability for Civil and Environmental Engineers*. McGraw-Hill: New York, NY.

Lake PS. 2003. Ecological effects of perturbation by drought in flowing waters. *Freshwater Biology* **48**: 1161–1172.

Monk WA, Wood PJ, Hannah DM, Wilson DA, Extence CA, Chadd RP. 2006. Flow variability and macroinvertebrate community response within riverine systems. *River Research and Applications* **22**: 595–615.

Niblack W, Barber R, Equitz W, Flickner MD, Glasman EH, Petkovic D, Yanker P, Faloutsos C, Taubin G, Heights Y. 1993. Querying images by content, using color texture, and shape. In *SPIE Conference on Storage and Retrieval for Image and Video Databases*, 1908: 173–187.

Olden JD, Poff NL. 2003. Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. *River Research and Applications* **19**: 101–121.

Poff NL, Allan JD, Bain MB, Karr JR, Prestegaard BD, Richter BD, Sparks RE, Stromberg JC. 1997. The natural flow regime: a paradigm for river conservation and restoration. *Bioscience* **47**: 769–784.

Poff NL, Olden JD, Pepin DM, Bledsoe BP. 2006. Placing global stream flow variability in geographic and geomorphic contexts. *River Research and Applications* **22**: 149–166.

Richter BD, Baumgartner JV, Powell J, Braun DP. 1996. A method for assessing hydrologic alteration within ecosystems. *Conservation Biology* **10**: 1163–1174.

Richter BD, Baumgartner JV, Wigington R, Braun DP. 1997. How much water does a river need? *Freshwater Biology* **37**: 231–249.

Richter BD, Baumgartner JV, Braun DP, Powell J. 1998. A spatial assessment of hydrologic alteration within a river network. *Regulated Rivers: Research and Management* **14**: 329–340.

Richter BD, Warner AT, Meyer JL, Lutz K. 2006. A collaborative and adaptive process for developing environmental flow recommendations. *River Research and Applications* **22**: 297–318.

Sale MJ, Brill ED, Herricks EE. 1982. An approach to optimizing reservoir operation for downstream aquatic resources. *Water Resources Research* **18**: 705–712.

Scott DW. 1979. On optimal and data-based histograms. *Biometrika* **66**: 605–610.

Shiau JT, Wu FC. 2004a. Assessment of hydrologic alterations caused by Chi-Chi diversion weir in Chou-Shui Creek, Taiwan: opportunities for restoring natural flow conditions. *River Research and Applications* **20**: 401–412.

Shiau JT, Wu FC. 2004b. Feasible diversion and instream flow release using Range of Variability Approach. *Journal of Water Resources Planning and Management* **130**: 395–404.

Shiau JT, Wu FC. 2006. Compromise programming methodology for determining instream flow under multiobjective water allocation criteria. *Journal of the American Water Resources Association* **42**: 1179–1191.

Shiau JT, Wu FC. 2007a. A dynamic corridor-searching algorithm to seek time-varying instream flow releases for optimal weir operation: comparing three indices of overall hydrologic alteration. *River Research and Applications* **23**: 35–53.

Shiau JT, Wu FC. 2007b. Pareto-optimal solutions for environmental flow schemes incorporating the intra-annual and interannual variability of the natural flow regime. *Water Resources Research* **43**: W06433. Doi:10.1029/2006WR005523

The Nature Conservancy. 2005. *IHA Applications Database*. http://www.nature.org/initiatives/freshwater/files/iha_apps.pdf

Thoms MC, Parsons M. 2003. Identifying spatial and temporal patterns in the hydrological character of the Condamine-Balonne River, Australia, using multivariate statistics. *River Research and Applications* **19**: 443–457.

Water Conservancy Agency (WCA). 2000. *Feasibility Planning on Conjunctive Use of Tsengwen Reservoir and Nanhua Reservoir*. WCA: Taichung, Taiwan (in Chinese).

Wood PJ, Armitage PD, Cannan CE, Petts GE. 1999. Instream mesohabitat biodiversity in three groundwater streams under base-flow conditions. *Aquatic Conservation: Marine and Freshwater Ecosystems* **9**: 265–278.