We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Analysis of Hydrologic Alteration Due to River Diversion

Giancarlo Principato and Giuseppe Viggiani
Department for Soil Conservation, University of Calabria, Rende, Italy

1. Introduction

Relevant alteration of rivers is most often caused by water diversions for irrigation, hydropower, industry and/or domestic uses. The operation of weirs and reservoirs has resulted in severe modification of the hydrologic regime (Rosenberg et al., 2000; Postel & Richter, 2003; World Commission on Dams [WCD], 2000; World Conservation Union [IUCN], 2000). As consequence, the discharges released into or left in a river for ecological purposes - the so-called “environmental flow requirements” - have received great attention in the last few decades from scientists, technicians and water managers (Richter et al., 2003; Tharme, 2003).

Among ecologists, the importance of flow variability in determining the structure and the function of the river ecosystem has been widely recognized (Richter et al., 1996; Poff et al., 1997; Bunn & Arthington, 2002). Richter et al. (1996) stated the natural flow paradigm as ‘the full range of natural intra- and inter-annual variation of hydrological regimes, and associated characteristics of timing, duration, frequency and rate of change, are critical in sustaining the full native biodiversity and integrity of aquatic ecosystems’.

The above “variation of hydrological regime” should be calculated statistically, i.e. as the change in the frequency distribution of a set of hydrologic variables useful to describe the overall flow regime (Arthington et al., 2006) or specific flow events (Stewardson & Gippel, 2003).

According to this environmental flow approach, the first step is to define ecologically-meaningful flow variables that capture natural flow variability. A great number of variables has been suggested for this purpose (Olden & Poff, 2003) and the 33 Indicators of Hydrologic Alteration (IHA) (Richter et al., 1996) are intended to represent each of the major facets of the flow regime (Olden & Poff, 2003; Monk et al., 2007). The IHA method assumes that, if these indicators are significantly altered, this will result in a decline in overall ecosystem health.

Therefore, the second step of the analysis involves the estimation of the alteration due to a water diversion through comparison of the probability distributions of the hydrologic variables (the IHA or other variables) in the pre- and post-impact conditions.
Unfortunately, the probability distributions can be difficult to obtain for several variables. Data may show skewed, scattered or multi-modal distribution (Principato & Viggiani, 2006) and non-normal distribution is recurrent (The Nature Conservancy, 2009).

Consequently, simplified statistical analyses have been proposed.

The Range of Variability Approach (RVA) (Richter et al., 1997) was formulated to quantify the modification of the IHA by comparing the frequencies within three fixed intervals. The RVA is a milestone in the hydrologic alteration assessment and has been widely utilised (e.g.: Galat & Lipkin, 2000; Irwin & Freeman, 2002; Shiau & Wu, 2004, 2007a, 2007b). Zolezzi et al. (2009) combined the RVA with a wavelet transform analysis in order to separate the scales of variability and investigate their alterations independently.

The method of Suen & Eheart (2006) integrated the natural flow paradigm with the “intermediate disturbance hypothesis” (Connel, 1978), which states that species diversity will be highest at sites that have had an intermediate frequency of disturbance and lower at sites that have experienced extremely high or extremely low disturbance frequencies. Based on this assumption, they suggested that the range of intermediate disturbance levels for six eco-hydrological indicators could be used as a management target.

Shiau and Wu (2008) developed a Histogram Matching Approach (HMA) for resolving the limited flow variability resulting from RVA. For each IHA, histograms with a greater number of intervals (to RVA) are first obtained. Then, the histograms are compared using a metric proposed by Niblack et al. (1993).

Botter et al. (2010) avoided the histogram analysis through the estimation of the probability density functions of flow (in the pre- and post-impact conditions) by means of a few climate, soil and vegetation parameters. The relevant parameters of the probability density functions (like mean, mode, skewness, peak probability) are the solely used to estimate the hydrologic alteration.

As final step, a flow release that minimizes hydrologic alteration has to be defined. To this end, great effort has been focused on the use of optimization models (Shiau & Wu, 2004, 2007a, 2007b; Suen & Eheart, 2006).

Despite the accuracy of such optimization models, the robustness and the statistical accuracy of the current methods used for hydrologic alteration analysis have not been extensively evaluated and the second step of the overall analysis, i.e. the alteration assessment, remains vague.

This chapter deals with the second step. First, the drawbacks of few existing methods are highlighted in a conceptual framework. Then, a real case is selected as the test case, in which a 40 years release scenario downstream of a diversion is compared to an unaltered condition. More precisely, the frequency distributions of the hydrologic variables in the pre- and post-impact conditions are compared using two of the methods mentioned above, namely RVA and HMA. Then, the need for a more accurate frequency analysis of hydrologic variables and the importance of capturing the full range of variation of the flow regime are incorporated in a novel Frequency-Based Approach (FBA). The FBA results are analysed and compared to the results of other methods.
To this aim, a reference solution cannot be obtained. Therefore, a proper analysis technique is developed, based on the physical correctness of diversion-alteration relationships. Comparison of methods is performed both for single indicators (IHA) and in terms of overall hydrologic alteration.

2. Existing methods

Once the natural flow regime (or another condition) is chosen as reference condition, the hydrologic alteration has to be evaluated by taking into account the modification of flow descriptors (IHA or other variables) from pre- to post-impact conditions. The variables in pre-impact conditions are usually computed using daily hydrological measurements from a relevant gauge station from a period prior that reflects the unaltered condition. The post-impact condition can be computed using either post-impact measurements (i.e., downstream of an existing diversion) at the same gauging station or synthesised release scenarios (i.e. downstream of a planned diversion).

Said $n_1$ and $n_2$ the number of years in the pre-impact and post-impact conditions, each hydrologic variable can be regarded as a stochastic variable for which there are two samples $a_i (i = 1, ..., n_1)$, $b_j (j = 1, ..., n_2)$.

2.1 Range of Variability Approach (RVA)

In the RVA, the alteration is evaluated by comparing the frequency with which observed (pre-impact) and post-impact variables (usually the IHA) fall within three intervals (categories). The standard interval limits are equal to 0, 33$^{rd}$, 67$^{th}$ and 100$^{th}$ percentiles, represented by a histogram with three bins. The same number of values ($n_1/3$) is expected to fall in each interval. However, if some values in the set are equal, the values in each category can be different from $n_1/3$.

Discrepancies can also occur because values that are equal to the category boundaries are placed in the middle category (The Nature Conservancy, 2009).

Hydrologic alteration is assumed to occur if the number of the post-impact values falling in the central interval (33$^{rd}$÷67$^{th}$ percentiles) differ from the expected ones (i.e. the number of the pre-impact values). To assist in the evaluation of this, an Hydrologic Alteration Factor (HAF) is calculated for each variable in the central category

$$HAF = \frac{F - F_0}{F_0}$$

where $F$ is the observed frequency and $F_0$ is the expected frequency. A positive HAF means that the frequency of values in the category has increased from the pre-impact to the post-impact period (with a maximum value of 2), while a negative value means that the frequency of values has decreased (with a minimum value of -1). If the other categories are not taken into account, however, important information can be lost. For example, when dealing with minimum 1-day annual flow, the value $HAF=-1$ may indicate that all values in the post-impact condition fall in the low category (extreme drought decrease in magnitude) or all values in the post-impact conditions fall in the high category as well (extreme drought increase in magnitude).
Thus, in order to more accurately evaluate the hydrologic alteration, HAF should be computed for the other two categories and all three values of HAF evaluated (latest version of the RVA, reported in The Nature Conservancy, 2009).

Information may also be lost if the distribution of the hydrologic variable is skewed, as the analysis may be reduced to a comparison of histograms in which most values fall in a single bin. This is more likely to arise for variables describing the number of events (number of zero flow days, number of high/low pulses, number of reversals).

Most of the drawbacks of the RVA are due to the low number of intervals (or histogram bins) used to analyse frequency modification, that is unrelated to $n_1$, $n_2$. Moreover, since the pre-impact data set determines the intervals size, the resultant intervals can be unsuitable to describe the post-impact data set or to compare the two sets.

Limitations of the RVA were discussed in the work of Shiau & Wu (2008) and are further detailed in the application below.

### 2.2 Suen and Eheart (2006) method

The method introduced by Suen & Eheart (2006) is based on six eco-hydrological indicators (coefficient of efficiency of the yearly trend of the hydrograph, dry season 10-day minimum, wet season 3-day maximum, number of high-flow events, mean duration of low-flow events and mean of all positive differences between consecutive values in wet season, i.e. the rising rate). Fuzzy theory was applied to represent the degree of disturbance levels and a Gaussian shape membership function has been used to describe indicators variability. In accordance to the intermediate distance hypothesis (Connel, 1978), disturbance occurs if indicators values are far from intermediate level (i.e. membership values are not close to 1). The ecosystem needs objective are coupled with human needs objective and incorporated in an optimization model.

However, the intermediate disturbance hypothesis partially contrasts the conceptual statement of the natural flow paradigm, according to which the full range of natural intra- and inter-annual variation of hydrological regimes has to be taken into account. Moreover, biodiversity concepts, initially based only on species diversity, have been progressively extended to integrate biotic and abiotic patterns and processes across scales, encompassing structural and functional processes, which are crucial in riverine context (Ward & Tockner, 2001). In fact, the importance of supra-seasonal extreme events has been assessed (Poff et al., 1997), especially for highly variable flow rivers (Lake, 2003), for which efforts to reduce the flow variability in order to increase biodiversity or to “restore” the river system to one that better fits a perception of a “healthy” river may not be the best ecological option (Boulton et al., 2000).

Finally, as noted above, the Gaussian distribution can be inappropriate for several hydrologic variables.

For these reasons, the Suen & Eheart (2006) method can be weak both from conceptual and statistical point of view.

### 2.3 Histogram Matching Approach (HMA)

In the HMA (Shiau & Wu, 2008), pre- and post-impact values of each IHA variable are represented with histograms whose number of bins is calculated taking into account all

www.intechopen.com
data and the characteristics of data distribution. As is the case of RVA, dissimilarity of histograms (in the pre- and post-impact condition) denotes the extent of the hydrologic alteration. The dissimilarity is evaluated using the quadratic-form distance proposed by Niblack et al. (1993), which accounts for both the class-by-class correspondence (differences of frequencies in each bin, like in the RVA) and cross-class information (distance between central values of bins).

The outcome is a ‘degree of histogram dissimilarity’, $D_Q$, for each variable.

Different options can be adopted depending on the value of $\alpha$, which is the exponent of the similarity function ($1 < \alpha < \infty$).

The HMA is incorporated in an aggregated multi-objective optimization genetic algorithm that minimizes both alteration ($D_Q$) and a shortage ratio that accounts for human needs - see Shiau & Wu (2008) for details.

One drawback of the HMA is that the metric used by Niblack et al. (1993) underestimates the distances between bins because it tends to accentuate the similarity of distributions without a pronounced mode (Rubner et al., 1998). Serratosa & Sanfeliu (2006) observed that most distance measures considers the overlap or intersection between two histograms as a function of the distance value, but they do not take accurately into account the similarity of the non-overlapping parts of the two histograms. The effects of these drawbacks are examined through the test case reported below.

3. Frequency Based Approach (FBA)

The Frequency Based Approach (FBA) is proposed here. It is an analysis of the frequency modification of each hydrologic variable from pre- to post-impact condition based on the Earth Mover’s Distance (EMD) algorithm (Rubner et al., 1998, 2000) applied to “extended signatures” (Serratosa & Sanfeliu, 2006).

Three key elements are considered in the FBA, due to the necessity to encompass:

i. **Histograms that better represent an empirical estimate of the probability density function**

The histogram of a set of univariate data ($n_1$ and $n_2$ data points of each hydrologic variable) provides the basis for an empirical estimate of the probability density function, that is required for comparing pre- and post- impact conditions. The formal definition of a histogram with $N$ bins used for estimating the density function of a stochastic variable $x$ with $n$ data (Silvermann, 1986) is

$$f(x) = \frac{h_i}{n \Delta x}, \quad (x_0 + j \Delta x \leq x < x_0 + (j + 1) \Delta x)$$  \hspace{1cm} (2)

where $h_i$ is the number of data points, $x_i$ within the $j^{th}$ bin, whose limits are $x_0 + j \Delta x$ and $x_0 + (j+1) \Delta x$, (in which $x_0$ is an origin of the range of interest). It is worth noting that the count is also ratioed to the bin width $\Delta x$ and therefore the definition satisfies the condition of histogram total area equal to 1, that matches the probability definition. In effect, the information conveyed is the area of the plot - not only the height of the column representing the bin.

ii. **Histograms that avoids to compute dissimilarity caused by histograms structure**
This is because dissimilarity between histograms can be influenced by arbitrary choice of bin centres and limits of the two histograms, as well as by bin number.

iii. a different metric that overcomes the weakness of other existing metrics

If the ordering of the elements in the set is unimportant, a histogram is a lossless representation of the set itself (Serratosa & Sanfeliu, 2006). Thus, the “distance” between two sets can be computed in an efficient way by computing the distance between their histograms. For this reason, a number of measures of similarity between histograms have been proposed and used in other fields, especially in computer vision and pattern recognition (e.g.: Protein classification). These include: the quadratic-form distance proposed by Niblack et al. (1993), the B-distance, the Matusita approach and the “K–L distance” (Cha & Srihari, 2002). As noted above, many methods do not accurately take into account the similarity of the non-overlapping parts of the two histograms (Serratosa & Sanfeliu, 2006). This problem was overcome by Rubner et al. (1998, 2000), who presented a new definition of the distance measure between histograms called Earth Mover’s Distance (EMD), that is more robust than histogram matching techniques.

3.1 Histograms construction

The above described points (a), (b) are taken in account with a proper construction of histograms.

A histogram \( h_j \) is defined as a mapping from a set of \( d \)-dimensional integer vectors \( j \) to the set of nonnegative integers. These vectors typically represent bins (or their centres) in a fixed-size partitioning of the relevant region of the underlying space, while the associated integers are a measure of the mass of the distribution that falls into the corresponding bin.

On the other hand, a signature is a variable-size description of a distribution, aimed to reach a balance between expressiveness and efficiency of representation, thus overcoming deficiency of fixed-size structure like histograms (Rubner et al., 2000). A element of a signature is defined as

\[
\{ s_i \equiv (m_i, w_i) \}\]

in which \( m_i \) is the mean (or mode) of the cluster \( j \) and \( w_i \) is the number of elements that belong to that cluster.

In the present application, fixed-size histograms can be still adequate and can be considered as a special case of signatures, i.e. the histogram \( h_j \) with \( N \) bins can be viewed as a signature in which \( m_i \) is the central value of the bin \( j \) of the histogram \( (j=1,\ldots, N) \) and \( w_j \) is equal to \( h_j \). Therefore, in such case, the signatures

\[
\{ s_a = (m_a, w_a) \}\]

\[
\{ s_b = (m_b, w_b) \}\]

- in which \( w_a = w_{a,j} \) \( (j=1,\ldots,N_a) \), \( w_b = w_{b,j} \) \( (j=1,\ldots,N_b) \), are the number of elements of \( a, b \) in clusters \( m_a = m_{a,j} \) \( (j=1,\ldots,N_a) \), \( m_b = m_{b,j} \) \( (j=1,\ldots,N_b) \) - are only a formal definition of the histograms of the variables \( a, b \).
The variables \(a, b\) and the vector \(m_a\) can be adimensionalised by dividing for the mean of \(a_i\) (\(i = 1, \ldots, n_1\))

\[
a' = a / \text{mean}(a) \tag{5a}
\]

\[
b' = b / \text{mean}(a) \tag{5b}
\]

\[
m_{a}' = m_a / \text{mean}(a) \tag{6}
\]

As subsequent step, the histograms can be represented as "extended signatures" (Serratosa & Sanfeliu, 2006), in which the minimum number of empty bins is added to assure that the number of bins for both pre- and post-impact data are the same. To this aim, the vector \(m'\) is obtained adding equal size bins to \(m_a'\) in order to include all values of \(b'\). The size of \(m'\), \(N'\), is equal or greater than \(N\). As result, the same \(m'_j\) (\(j = 1, \ldots, N'\)) are used for both adimensionalised variables \(a', b'\).

The extended signatures

\[
\{s'_a = (m', w'_a)\} \tag{7a}
\]

\[
\{s'_b = (m', w'_b)\} \tag{7b}
\]

(in which \(w'_a, w'_b\) are the number of elements of \(a', b'\) in each cluster of \(m'\)) are obtained.

Finally, the vectors \(W_{a} = W_{a,j}\) (\(j = 1, \ldots, N'\)), \(W_{b} = W_{b,j}\) (\(j = 1, \ldots, N'\)) are calculated using the equations

\[
W_{a} = \frac{w'_a}{n_1 \Delta m'} \tag{8a}
\]

\[
W_{b} = \frac{w'_b}{n_2 \Delta m'} \tag{8b}
\]

(in which \(\Delta m'\) is the \(m'\) cluster size) in order to satisfy the conditions related to frequency

\[
\sum W_{a} \Delta m' = 1 \tag{9a}
\]

\[
\sum W_{b} \Delta m' = 1 \tag{9b}
\]

The final result consists of two signatures

\[
\{S_{a} = (m', W_{a})\} \tag{10a}
\]

\[
\{S_{b} = (m', W_{b})\} \tag{10b}
\]

corresponding to histograms with same bins (same centres and bin limits) and total area equal to 1. Each bin includes a non-negative number of elements.

3.2 The Earth Mover’s Distance (EMD) for hydrologic variable signatures

The Earth Mover’s Distance was proposed by Rubner et al. (1998, 2000). Given two distributions, one can be seen as a mass of earth properly spread in space (supplier), the
other as a collection of holes in that same space (consumer). The EMD measures the least amount of work needed to fill the holes with earth. A unit of work corresponds to the cost of transporting a unit of earth by a unit of ground distance.

The supplier and the consumer can be also two signatures. In such case, the EMD is defined as the minimum amount of work that must be performed to transform one signature into the other by moving distribution mass.

The cost $c_{ij}$ is the ground distance between element $i$ in the first signature and element $j$ in the second signature (in the present case, $i=1,\ldots,N', j=1,\ldots,N'$). $c_{ij}$ can be any distance, e.g. an Euclidean distance. The matrix $C$ of elements $c_{ij}$ is the cost matrix.

For signatures 10a, 10b, the problem is finding a set of flows $f_{ij}$ that minimize the overall cost

$$
\sum_{i=1}^{N'} \sum_{j=1}^{N'} c_{ij} f_{ij} = \text{min} \tag{11}
$$

subject to the following constraints

$$
f_{ij} \geq 0, \ (i = 1, \ldots, N', j = 1, \ldots, N') \tag{12}
$$

$$
\sum_{j=1}^{N'} f_{ij} \leq W_{ai} \quad (1 < i < N') \tag{13}
$$

$$
\sum_{i=1}^{N'} f_{ij} \leq W_{b,i} \quad (1 < j < N') \tag{14}
$$

$$
\sum_{i=1}^{N'} \sum_{j=1}^{N'} f_{ij} = \text{min} \left( \sum_{i=1}^{N'} W_{a,i}, \sum_{j=1}^{N'} W_{b,j} \right) \tag{15}
$$

Constraint (12) allows shipping of supplies from a supplier to a consumer and not vice versa. Constraint (13) limits the amount of supplies that can be sent to the clusters of the consumer to their weight. Constraint (14) limits the clusters of the consumer to receive no more supplies than their weight. Constraint (15) is related to total flow and forces to move the maximum amount of supplies possible.

The optimal flow matrix $F=[f_{ij}]$ (whose size, in the present case, is $N' \times N'$) is obtained solving the classic transportation problem by the simplex algorithm (Dantzig, 1951).

Once the transportation problem has been solved and the optimal flow has been computed, the Earth mover’s distance is defined as

$$
\text{EMD} = \frac{\sum_{i=1}^{N'} \sum_{j=1}^{N'} c_{ij} f_{ij}}{\sum_{i=1}^{N'} \sum_{j=1}^{N'} f_{ij}} \tag{16}
$$

where the denominator is a normalization factor that coincides with the smaller signature, because of constraint (15). This factor is needed when the two signatures have different total weight and it avoids favoring signatures with smaller total weights.

If the ground distance is a metric and the total weights of two signatures are equal, the EMD is a true metric.

3.3 The FBA algorithm

FBA comprises the following steps:

i. in order to compare the hydrologic alteration of different variables (whose values can show different order of magnitude and even different units), the values of each variable (in both conditions) are adimensionalised dividing by the mean of each pre-impact variable set a (eq. 5a, 5b);

ii. the histograms and their signatures are defined for each variable using the above described procedure, in order to obtain eq. 10a, 10b;

iii. the Earth Mover’s Distance (16) is computed for each variable using the algorithm available in C++ at http://www.cs.duke.edu/~tomasi/software/emd.htm, that is incorporated in a Matlab routine.

In the present case, total weight of the two signatures are coincident (eq. 9a, 9b) as well as their cluster number $N'$. Therefore, constraints 13, 14, 15 become

$$\sum_{j=1}^{N'} f_{ij} = W_{a,i} \quad (1 < i < N')$$  \hspace{1cm} (17)

$$\sum_{i=1}^{N'} f_{ij} = W_{b,i} \quad (1 < j < N')$$  \hspace{1cm} (18)

$$\sum_{i=1}^{N'} \sum_{j=1}^{N'} f_{ij} = \sum_{i=1}^{N'} W_{a,i} = \sum_{j=1}^{N'} W_{b,j}$$  \hspace{1cm} (19)

The number $N$ of bins in the signatures 4a, 4b is chosen according Sturges rule

$$N = 1 + \log_2 n_1$$  \hspace{1cm} (20)

Though this rule is not applicable to sets of data that have strongly non-Gaussian distribution, for moderate $n_1$ (less than 200) it gives similar results to other alternative rules (Scott, 1992) and thus it is assumed for FBA. The Euclidean distance

$$c_{ij} = |m'_i - m'_j| = || - || \Delta m'$$  \hspace{1cm} (21)

is chosen as ground distance between central points of $i$-bin in signature $\{s_a=(m'_a, W_a)i\}$ and $j$-bin in signature $\{s_b=(m'_b, W_b)j\}$, in order to obtain the $N'\times N'$ cost matrix $C$ of elements $c_{ij}$ ($i=1,...,N', j=1,...,N'$).

$C$, $W_a$, $W_b$ are the input of the algorithm, whose output is the scalar EMD and the $N'\times N'$ flow matrix $F$.

The algorithm is applied to each hydrologic variable, for which the EMD is a dimensionless measure of the dissimilarity of the two histograms, i.e. measures of the frequency modification (and hence the hydrologic alteration).

It should be noted that the FBA is entirely based on sample(s), while statistical modelling aimed to estimate the probability distribution is not performed. In other words, the registered flows are considered the most reliable information.

Finally, it should be stressed that the full range of variation of the flow regime is adequately incorporated in the FBA, while in the RVA extreme events (minimum and maximum values of each variable) are included in histogram bins that contain also ordinary events (first and third bins). In the FBA and HMA, extreme events are more accurately represented using histograms.
with a greater number of bins. However, as noted above, the similarity of the non-overlapping parts of the two histograms (mainly the external bins) is more accurately computed using the EMD algorithm rather than Niblack et al. (1993) metric adopted in the HMA.

3.4 Case study

The Crati River is the main watercourse of Calabria (South Italy), with catchment area of 2431 km², mean altitude of 600 m and length of 81 km. It is a typical Mediterranean river, characterized by an irregular and perennial flow regime - that strongly depend on rainfall. Periods of zero-flow are rare, while severe summer droughts and autumn-winter floods are fairly regular. Crati estuary area is a natural reserve.

A time-series of 40-years records (1927-1966) of mean daily (near) natural discharge is available from “Conca” gauge station. Relevant data for Conca gauge station are:

- distance from estuary: 23 km, altitude: 35 m;
- catchment area: 1332 km², with a mean altitude of 664 m and moderate permeability;
- mean annual rainfall: 1260 mm (rainfall concentrated from November to May);
- mean annual flow: 26.0 m³/s, mean February flow: 56.5 m³/s, mean August flow: 3.8 m³/s, annual coefficient of variation (standard deviation of all the daily flow values, divided by the mean annual flow): 1.27.

The alluvial plain of the Crati River (downstream of Conca gauge station) is a major agricultural area of Calabria and, as consequence, the Tarisio diversion weir (near Conca gauge station) was completed before 1970 to provide water for off-channel irrigation.

In the present application, a variable release scenario was hypothesized to occur for a 40 years period (1967-2006), during which time the flow regime represented by the recorded registered mean daily discharge (Q) in the period 1927-1966 was assumed to occur upstream of the flow diversion, while releases downstream of the diversion (Q_r) were assumed to be in accordance to the operational rule

\[
\begin{cases}
Q \leq Q_{\text{min}} \rightarrow & Q_r = Q \\
Q_{\text{min}} \leq Q \leq Q_{\text{min}} + Q_d \rightarrow & Q_r = Q_{\text{min}} \\
Q > Q_{\text{min}} + Q_d \rightarrow & Q_r = Q - Q_d
\end{cases}
\]  

(22)

where Q is the flow upstream the diversion, Q_{min} is the value of flow corresponding to some environmental flow prescription and Q_d is the projected diversion. Eq. (22) is the typical flow release rule for an unregulated weir.

In this application, according to prescription of the Basin Authority of Calabria, Q_{min}=4.5 m³/s for the entire year was assumed (Principato & Viggiani, 2009). Different (monthly or daily) assumptions can be done for Q_{min}.

Q_d was set equal to 2, 5, 10, 15, 20, 25, 30, 35, 40 m³/s (larger values were not considered because mean annual flow is 26.2 m³/s).

4. Results analysis

The hydrologic alteration due to different values of Q_d was calculated for the 33 IHAs using RVA (with standard interval limits), HMA (with α=1) and FBA. The 33 IHA are considered,
although other variables can also be used. The Suen & Ehart (2006) method has been discussed above, but is based on other variables and thus not used here for comparison.

Thus, the Alteration Indexes AI (HAF$^k$, D$^k$, and EMD$^k$, $k=1,\ldots,33$) were calculated. The $\mathbf{a}^r$, $\mathbf{b}^r$, and $\mathbf{c}^r$ for March mean flow (with a diversion of $Q_d=20\,\text{m}^3/\text{s}$) and the corresponding histograms and signatures (with $n_1=n_2=40$, $N'=8$, mean flow: $40.16\,\text{m}^3/\text{s}$) are reported in Fig. 1-2.

The March mean flow signatures of $\mathbf{a}$ and $\mathbf{b}$ are first defined by (eq. 4a, 4b)

\[
\mathbf{w}_a = (13, 13, 5, 6, 1, 2)
\]

\[
\mathbf{m}_a = (26.35, 36.65, 46.95, 57.25, 67.55, 77.85)
\]

\[
\mathbf{w}_b = (1, 10, 5, 5, 1, 2)
\]

\[
\mathbf{m}_b = (9.37, 19.12, 28.87, 38.62, 48.37, 58.12)
\]

Then, the adimensional vector $\mathbf{m}_a'$ is calculated by dividing each element of $\mathbf{m}_a$ for the mean of $\mathbf{a}$ (eq. 5a) and the vector $\mathbf{w}_b'$ is obtained adding two equal size bins to $\mathbf{m}_b'$ in order to include all values of $\mathbf{b}'$ (eq. 6). $\mathbf{m}'$ contains the central values of bins in the signatures of adimensional variables $\mathbf{a}'$, $\mathbf{b}'$, for which (eq. 7a, 7b)

\[
\mathbf{w}_a' = (0, 1, 12, 13, 5, 6, 1, 2)
\]

\[
\mathbf{w}_b' = (13, 13, 5, 6, 1, 1, 1, 0)
\]

Finally, the vectors

\[
\mathbf{W}_a = (0, 0.098, 1.170, 1.267, 0.487, 0.585, 0.098, 0.195)
\]

\[
\mathbf{W}_b = (1.267, 1.267, 0.487, 0.585, 0.098, 0.098, 0.098, 0)
\]

are calculated using equations 8a, 8b. Total area of each signature $\{S_a=(\mathbf{m}',\mathbf{W}_a)\}$, $\{S_b=(\mathbf{m}',\mathbf{W}_b)\}$ is 1. The cost matrix $\mathbf{C}$ of elements $c_{ij}$ is given by eq. 21 (in which $\Delta m'=0.257$).

The optimal flow matrix $\mathbf{F}$, whose values are positive (eq. 12), is

\[
\begin{array}{cccccccccc}
\text{F}\text{=} & & & & & & & & & \\
\text{j}=1,\ldots,\text{N}'=8 & & & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
0.098 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.098 \\
0.682 & 0 & 0.487 & 0 & 0 & 0 & 0 & 0 & 0 & 1.170 \\
0.487 & 0.195 & 0 & 0.585 & 0 & 0 & 0 & 0 & 0 & 1.267 \\
0 & 0.487 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.487 \\
0 & 0.390 & 0 & 0 & 0 & 0.098 & 0.098 & 0 & 0 & 0.585 \\
0 & 0.098 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.098 \\
0 & 0.098 & 0 & 0 & 0 & 0 & 0.098 & 0.098 & 0 & 0.195 \\
\end{array}
\]

\[
\begin{array}{cccccccccc}
\text{W}_a & & & & & & & & & \\
\text{j}=1,\ldots,\text{N}'=8 & & & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
0.098 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.098 \\
0.682 & 0 & 0.487 & 0 & 0 & 0 & 0 & 0 & 0 & 1.170 \\
0.487 & 0.195 & 0 & 0.585 & 0 & 0 & 0 & 0 & 0 & 1.267 \\
0 & 0.487 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.487 \\
0 & 0.390 & 0 & 0 & 0.098 & 0.098 & 0 & 0 & 0 & 0.585 \\
0 & 0.098 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.098 \\
0 & 0.098 & 0 & 0 & 0.098 & 0.098 & 0 & 0 & 0 & 0.195 \\
\end{array}
\]
The flow matrix indicates the amounts of $W_a$ to ship to locations specified by $m'$ to obtain $W_b$, or vice versa (each signature can be a supplier or a consumer because of coincident total weights).

Therefore, sum of rows is equal to $W_a$, while sum of column is equal to $W_b$. E.g., the amounts in first column have to be shipped from 2nd, 3rd and 4th elements of $W_a$ to obtain first element of $W_b$.

It should be noted that amounts of $W_a$ are not added to $W_b$, but to empty locations (clusters) of $W_b$ (or vice versa), while the total cost is minimum (eq. 11).

The correspondence between histograms centres and bin size in the pre- and post-impact conditions has to be stressed. In fact, it ensures that histograms dissimilarity is due to flow alteration rather than histograms structure.

Fig. 1. March mean flow in the pre-impact ( — ) and post-impact ( - - ) conditions in the $n_1=n_2=40$ year periods ($Q_d=20$ m$^3$/s).

Fig. 2. Histograms for March mean flow in the pre-impact (a) and post-impact (b) conditions and corresponding extended dimensionless signatures (c, d).
4.1 “Expected variations” technique for evaluating methods accuracy

Once the RVA, HMA and FBA results were obtained, the problem to evaluate and compare their accuracy was considered. Preferably, results should be compared to a reference solution, that can usually be an analytical solution or a set of experimental data. Unfortunately, when analysing hydrologic alteration, the reference solution is not known neither analytically nor experimentally.

Moreover, the indexes HAF<sub>k</sub>, D<sub>Q,k</sub> and EMD<sub>k</sub> are all influenced by natural variability of flow regime, that cannot easily be distinguished from hydrologic alteration. In order to overcome this problem and to focus the analysis on methods accuracy, a post-impact flow (upstream of diversion) equal to the pre-impact flow is hypothesized in the present application. This option ensures that all alteration is attributable to the flow diversion. In other words, if Q<sub>d</sub>=0 (no diversion), all indexes of alteration are equal to zero.

When Q<sub>d</sub>≠0, different values of alteration indexes AI are obtained (-1< HAF<2, 0< D<sub>Q</sub><1, 0< EMD<1). These values cannot be compared each other. Therefore, an accuracy analysis based on indexes values cannot be conducted (what is the exact value of alteration index?).

Relevant conclusions on indexes variation can be formulated instead. In fact, for many variables, the expected variation of AI due to Q<sub>d</sub> increase can be qualitatively stated. For the IHA, such expected variations are detailed below and summarized in Tab. 1.

Since the operational rule (22) causes the flow to be reduced (if Q>Q<sub>min</sub>) or unaltered (if Q<Q<sub>min</sub>):

- for variables related to the magnitude of flow, alteration is expected to increase (or remain constant) when diverted discharge increases (Not Decreasing condition-ND);
- Number of zero-flow days is not influenced by operational rule (22) – (No alteration condition - N).
- Base flow is computed as 7-day minimum flow/mean flow for year. Mean flow decreases with Q<sub>d</sub> while, according to (22), 7-day minimum flow is not modified if it is less than Q<sub>min</sub>. Therefore, base flow alteration is generally expected to increase with Q<sub>d</sub> (Increasing condition - IN), but the condition 7-day minimum flow<Q<sub>min</sub> has to be verified in each application.
- Julian date of each annual 1-day maximum is not influenced by operational rule (22), because each annual maximum is reduced by the same quantity (Q<sub>d</sub>) in the post-impact condition (N), except in the years when the annual 1-day maximum flow is less than the sum of Q<sub>min</sub> and Q<sub>d</sub> (this condition has to be verified in each application). On the contrary, Julian date of each annual 1-day minimum can be altered because, if there are multiple days in the water year with the same minimum flow value, the earliest date is considered in the computation (The Nature Conservancy, 2009). In this case, alteration should either increase or remain constant (ND).
- Number and duration of high/low pulses, fall/rise rate can increase or decrease owing to (22), depending on the specific sequence of natural flow. General conclusions cannot be drawn.
- Number of reversals decreases when Q<sub>min</sub><Q<sub>Q</sub><Q<sub>min</sub>+Q<sub>d</sub> and Q<sub>d</sub>=Q<sub>min</sub>; otherwise it is not modified. Being periods with Q<sub>min</sub><Q<sub>Q</sub><Q<sub>min</sub>+Q<sub>d</sub> more long and frequent for large Q<sub>d</sub>, alteration increases with Q<sub>d</sub> (IN).
As result, the variation of AI due to \( Q_d \) increase/decrease is known for 25 to 27 IHA (depending on the conditions to be tested), while for remaining 6 variables conclusions cannot be drawn. Furthermore, three kinds of variation can occur (IN, ND, N). Therefore, the technique here developed allows to evaluate the accuracy of RVA, HMA and FBA through the analysis of 25-27 IHA.

### 4.2 Comparison of results for single variables

The results for indexes HAF, \( D_Q \), EMD - expressed in terms of variation due to \( Q_d \) increase - are summarized in Tab. 1.

If a minimum exists in the \( AI(Q_d) \) relationship, results are not physically correct, since alteration cannot decrease if diverted flow \( Q_d \) increases. In such case, Decrease (D) is indicated in Tab. 1 (it means that the corresponding method fails). In summary:

- **Monthly mean flows** (ND expected; Fig. 3): ND condition does not occur both for HAF (February, March, April, October, December) and \( D_Q \) (February, August, September). Moreover, relationships HAF(\( Q_d \)) and \( D_Q(Q_d) \) are irregular also for other monthly flow. Such conditions are not observed for index EMD, that shows a regularly not decreasing relationship to \( Q_d \).

- **Annual minima** - 1, 3, 7, 30, 90 days mean (ND expected; Fig. 4): HAF is not sensitive to alteration of 1, 3, 7 days mean, while \( D_Q \) values for 30, 90 days mean are irregularly related to \( Q_d \). ND condition is satisfied only using EMD.

- **Annual maxima** - 1, 3, 7, 30, 90 days mean (ND expected; Fig. 5): HAF(\( Q_d \)) relationship is very irregular (two minima are observed for all but one variables). \( D_Q(Q_d) \) irregularity is less evident, but slight decrease occurs for 1, 3, 7, 30 days mean. ND condition is satisfied using EMD.

- **N. of zero-flow days** (N expected): all indexes are equal to zero for all values of \( Q_d \).

- **Base flow** (IN expected; Fig. 6a): a highly irregular relationship is observed both for HAF and \( D_Q \) while EMD satisfies IN condition.

- **Julian dates of each annual 1-day minimum** (ND expected; Fig. 6b): most of alteration is due to 1954 winter drought, that preceded the rainy summer of the same year (in which \( Q>Q_{min} \)). So, in 1993, minimum flow release (\( Q=Q_{min} \)) would occur 1\(^{st} \) January. HAF is not sensitive to such alteration, while \( D_Q \) does not satisfy the ND condition. On the contrary, index EMD satisfies the ND condition.

- **Julian dates of each annual 1-day maximum** (N expected): alteration is not expected nor observed for the three indexes.

- **Number and duration of low/high pulses**: conclusions cannot be drawn.

- **Rise/fall rate**: conclusions cannot be drawn.

- **Number of reversals** (IN expected): alteration decreases with \( Q_d \) both for HAF and \( D_Q \), while it correctly increases when using EMD (Fig. 8).

As result, only FBA matches IN, N and ND conditions for the 27 variables for which qualitative relationships are known. Discrepancies are significant both for RVA (17 mismatches) and HMA (13 mismatches) and therefore a fictitious alteration (due to method accuracy) is added to (or subtracted from) real alteration.

In accordance with Shiau & Wu (2008), it should be concluded that the RVA can lead to misleading outcome. However, also HMA reveals unphysical response to diverted flow.
increase/decrease and should be discarded or carefully evaluated as hydrologic alteration assessment method.

Fig. 3. $Q_d$ - Alteration indexes HAF (●), $D_Q$ (▲), EMD (■) relationships for January mean flow (a), August mean flow (b), September mean flow (c), October mean flow (d)

Fig. 4. $Q_d$ - Alteration indexes HAF (●), $D_Q$ (▲), EMD (■) relationships for annual minima-7 days mean (a), annual minima-90 days mean (b)
Fig. 5. Qd - Alteration indexes HAF (♦), DQ (▲), EMD (■) relationships for annual maxima-7 days mean (a), annual minima-30 days mean (b)

Fig. 6. Qd - Alteration indexes HAF (♦), DQ (▲), EMD (■) relationships for base flow (a) and Julian date of each annual one-day minimum (b)

Fig. 7. Qd - Alteration indexes HAF (♦), DQ (▲), EMD (■) relationships for number of reversals
expected results | RVA | HMA | FBA
---|---|---|---
IHA Group 1: Discharge for each cal. month (m$^3$/s)
January | ND | ND | ND | ND
February | ND | D | D | ND
March | ND | D | ND | ND
April | ND | D | ND | ND
May | ND | ND | ND | ND
June | ND | ND | ND | ND
July | ND | N | ND | ND
August | ND | N | D | ND
September | ND | ND | D | ND
October | ND | D | ND | ND
November | ND | ND | ND | ND
December | ND | D | ND | ND
IHA Group 2: Discharge (m$^3$/s)
Annual min. - 1 day mean | ND | N | ND | ND
Annual min. - 3 days mean | ND | N | ND | ND
Annual min. - 7 days mean | ND | N | D | ND
Annual min. - 30 days mean | ND | ND | D | ND
Annual min. - 90 days mean | ND | ND | D | ND
Annual max. - 1 day mean | ND | D | D | ND
Annual max. - 3 days mean | ND | D | D | ND
Annual max. - 7 days mean | ND | D | D | ND
Annual max. - 30 days mean | ND | D | D | ND
Annual max. - 90 days mean | ND | D | ND | ND
N. of zero-flow days | N | N | N | N
Base flow | IN | D | D | IN
IHA Group 3: timing of annual extreme water conditions
Julian date of each annual 1-day minimum | ND | N | D | ND
Julian date of each annual 1-day maximum | N | N | N | N
IHA Group 4: frequency and duration of high/low pulses
Number of low pulses each year | - | IN | D | D
Duration of low pulses within each year (days) | - | D | D | D
Number of high pulses each year | - | D | D | IN
Duration of high pulses within each year (days) | - | D | D | D
IHA Group 5: rate/frequency of water condition changes
Rise rate | - | ND | D | ND
Fall rate | - | ND | D | IN
Number of hydrologic reversals | IN | D | D | IN

Table 1. Analysis of Al-$Q_d$ relationships: computed and expected results for RVA, HMA and FBA (ND: not decreasing, IN: increasing, D: decreasing, N: no variation)
4.3 Comparison of results for overall hydrologic alteration

Once alteration indexes \( AI \) are computed for each \( N \) variables, overall hydrologic alteration can be estimated using the index proposed in Shiau & Wu (2007a). The Index

\[
L_2 = \left( \frac{1}{N} \sum_{k=1}^{N} AI_k^2 \right)^{0.5}
\]  \hspace{1cm} (23)

is adopted here for the purpose of obtaining overall alteration-\( Q_d \) relationships for HAF, \( D_Q \) and EMD (Fig. 8). The suitability of \( L_2 \) is not discussed here. Other indexes could be used for comparison.

Adopting HAF, the pattern is quite irregular and \( L_2 \) for \( Q_d=35 \text{ m}^3/\text{s} \) is less than \( L_2 \) for \( Q_d=30 \text{ m}^3/\text{s} \). For \( D_Q \), pattern is less irregular, while more regular variations of overall alteration are observed for EMD. These patterns are obviously due to irregularities in many of the 33 \( AI(Q_d) \) relationships.

It should be concluded that the scarce accuracy of RVA and HMA (for single variables) also affects the estimation of overall hydrologic alteration.

![Fig. 8. \( L_2-Q_d \) relationships from RVA (†), HMA (▲) and FBA (■)](image)

5. Conclusions

The main characteristics and the accuracy of few methods for hydrologic alteration assessment have been analysed.

First, few theoretical drawbacks are highlighted for few existing methods, which are due to poor statistical analysis (RVA), inadequate metric (HMA) and weak ecological hypothesis (Suen & Eheart, 2006 method).

Then, quantitative evaluations have been conducted for a test case study using RVA and HMA, revealing weakness of both methods.

As consequence, a more robust method has been proposed here for the first time, namely the FBA. It considers the full range of flow regime variations (in accordance to natural flow paradigm) and ensures that a statistically-based analysis is conducted (comparison of the frequency distributions of a proper set of hydrologic variables), as required in the ecological context.
The FBA allows to obtain physically correct results (AI(Qd) relationships) for all hydrologic variables (IHA) for which qualitative relationships are known. Consequently, also overall alteration index are more accurately computed.

The hydrologic alteration analysis performed with FBA has many potential applications, like definition of water release plans, evaluation of river status (and correlation to ecological status), analysis of flow variability (natural variability or the effects of climate change).

However, hydrologic alteration analysis is far from being conducted with the same accuracy in river reaches for which adequate registered flow series are not available. Accurate applications in non-gauged river reaches remain to be defined both for predictive purposes (water release plans for future diversions) and existing diversions.

It is worth noting that, if alteration-diverted flow relationships are not physically and statistically meaningful, optimization efforts aimed to define release plans are not correctly addressed. Consequently, the integration of FBA in the existing optimization models - aimed to perform the third step of the overall analysis - is encouraged.

Notations

- $a_i$: values of $a$
- $a$: vector of hydrologic variable in the pre-impact condition
- $AI$: alteration index
- $b_i$: values of $b$
- $b$: vector of hydrologic variable in the post-impact condition
- $c_i$: cost to ship a unit of supply
- $C$: matrix of elements $c_i$ (cost matrix)
- $D$: decrease
- $D_Q$: degree of histogram dissimilarity in the HMA
- $EMD$: Earth Mover’s Distance
- $f_{ij}$: set of flows from supplier to consumer
- $F$: matrix of elements $f_{ij}$ (flow matrix)
- $F$: observed frequency in the RVA
- $F_0$: expected frequency in the RVA
- $h_i$: number of data within the $j$th bin of the histogram
- $HAF$: Hydrologic Alteration Factor in the RVA
- $IC$: Increasing Condition
- $L_2$: overall hydrologic alteration index
- $m_j$: central value of the bin $j$ of the histogram or signature
- $m'_j$: central value of the bin $j$ of the adimensionalised histogram or signature
- $m, m'$: vector of $m_j, m'_j$
- $n_1, n_2$: number of year in the pre-impact and post-impact conditions
- $N$: number of histogram bins
- $N'$: number of extended signature clusters
- $ND$: Not Decreasing condition
- $Q$: flow upstream the diversion (m$^3$/s)
- $Q_{min}$: flow corresponding to some environmental flow prescription (m$^3$/s)
- $Q_d$: projected diversion (m$^3$/s)
- $Q_r$: released discharges (m$^3$/s)
6. Acknowledgements

The authors acknowledge Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas for letting their C code available on the web.

7. References


The ecosystems present a great diversity worldwide and use various functionalities according to ecologic regions. In this new context of variability and climatic changes, these ecosystems undergo notable modifications amplified by domestic uses of which it was subjected to. Indeed the ecosystems render diverse services to humanity from their composition and structure but the tolerable levels are unknown. The preservation of these ecosystemic services needs a clear understanding of their complexity. The role of the research is not only to characterise the ecosystems but also to clearly define the tolerable usage levels. Their characterisation proves to be important not only for the local populations that use it but also for the conservation of biodiversity. Hence, the measurement, management and protection of ecosystems need innovative and diverse methods. For all these reasons, the aim of this book is to bring out a general view on the biogeochemical cycles, the ecological imprints, the mathematical models and theories applicable to many situations.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
