Climate-Biased Storm-Frequency Estimation

By Thomas E. Croley II[1]

Abstract: Storm frequencies for the future are often estimated directly from past historical records of sufficient length. The estimation requires no detailed knowledge of the area's meteorology, but presumes it is unchanged in the future. However, the climate seldom remains static. Numerous climate forecasts of meteorological probabilities over extended periods are now available. It is possible to use these meteorological forecasts directly in the estimation of storm frequencies from the historical record. A heuristic approach is defined here to estimate storm frequencies that recognize forecasts of extended weather probabilities. Basically, those groups of historical meteorological record segments matching forecast meteorological probabilities are weighted more than others, during the estimation of storm frequencies. (Affiliated groups of hydrologic record segments may be similarly weighted for hydrological estimation; e.g., flood frequency estimation.) An example of frequency estimation is made for maximum annual daily flow, using currently available agency meteorological forecasts in the United States and Canada.

INTRODUCTION

One of the important problems in hydrology deals with interpreting a past record of events, in terms of probabilities of occurrence, for the estimation of future behavior. Natural meteorological and hydrologic phenomena are variable but many times amenable to probabilistic interpretation and estimation analysis. Such interpretation requires that samples taken from the historical record be representative, unbiased, and independent. This direct estimation of event frequencies is simple and seemingly reliable. It enables engineers to obtain estimates of probabilities of events without detailed knowledge of the application area meteorology or hydrology. It is also theoretically satisfying since one is working with observations of real events and not abstract process models. However, it requires a statistically characteristic record and the assumption that the past is representative of the future. This assumption is often violated since climate varies.

Multiple long-lead forecasts of meteorological probabilities are now available to the water resource engineer and hydrologist and are reviewed elsewhere (Croley 1996, 1997b, 2000a). These forecasts are defined for different time scales and time periods at different lag times, spatial domains, event definitions, and meteorological variables, and they forecast either meteorological event probabilities or only most-probable meteorological events. Now, it is possible to estimate meteorological and hydrologic frequencies from the historical record by using these forecasts. (Various ''storm'' frequencies are considered here.) Such estimates embody all ''meteorological uncertainty'' inherent in both the record and in the forecasts.

Basically, those groups of meteorological segments from the historical record matching forecast meteorological probabilities are given more weight than those not matching, when estimating storm frequencies. The historical meteorological record is thus used in a nonrepresentative and biased manner to match meteorological forecasts; therefore, the requirements are removed for representative and unbiased records and methodology. Still, the methodology requires that the historical record contain an adequate range of independent possibilities.

The present paper first summarizes the methodology for estimation of storm frequencies from the historical record. Then, storm frequency estimates are modified to reflect available forecasts of meteorological probabilities. The adaptation is based on techniques used in operational hydrology approaches to transform forecasts of meteorological probabilities into forecasts of hydrologic probabilities (Croley 1996, 1997a,b, 2000a). Next, a comparison is made with a Bayesian framework, more flexible objectives are incorporated into the solution, and linear programming techniques are modified to eliminate infeasible meteorological probability forecasts. Finally, an example frequency estimate is made for annual maximum daily flow. It makes use of forecasts of meteorological probabilities, and includes agency forecasts in the United States and Canada.

STORM FREQUENCY ESTIMATION

Since storm frequencies are unknown, they are estimated from the historical record, assumed ergodic and treated as a ''random sample,'' wherein successive observations, (X_1, X_2, . . . , X_n), are considered identically distributed and equally likely to occur (both in the past and future). (For example, X may represent the annual maximum daily precipitation or the annual maximum flood flow.) Likewise, the observations must be defined so they can be considered as independent of each other. (For example, two successive storms occurring very closely may result in a high degree of dependence of the second on the first.) Temporal dependence can be minimized by defining long event interarrival times or record pieces. For example, annual maximum floods or rainfalls (interarrival time on the order of a year) are often taken as time independent, as are one-year record segments. Spatial independence must also be assured when multiple application areas are to be considered simultaneously. In practical parlance, the observation values, $x_1, \ldots, x_n$, of $X_1, X_2, \ldots, X_n$, respectively, sometimes are called a random sample also.

Storm frequencies or ''exceedance probabilities,'' $P[X \geq x]$ can be estimated directly from the historical record. Suppose all values, x_i, in a random sample of annual maximums (x_i, $i = 1, \ldots, n$) are ordered from largest to smallest to define the ordered variable values (y_l, $l = 1, \ldots, n$), where $y_l = x_{i(l)}$ and $i(l)$ is the number of the value in the unordered sample corresponding to the lth order. There are several methods to estimate exceedance probabilities from annual exceedance series (Chow 1964); without loss of generality, the popular ''Weibull'' method is used here as an example

$$\hat{P}[X \geq y_l] = \frac{l}{n+1}, \quad l = 1, \ldots, n \tag{1}$$

The caret (ˆ) denotes an estimate of the characteristic named under the caret. (Other methods also could be used.) Rewriting (1)

[1]Res. Hydrol., Great Lakes Environmental Research Laboratory, 2205 Commonwealth Blvd., Ann Arbor, MI 48105-2945.

Note. Discussion open until January 1, 2002. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on December 22, 1999; revised October 3, 2000. This paper is part of the ***Journal of Hydrologic Engineering***, Vol. 6, No. 4, July/August, 2001. ©ASCE, ISSN 1084-0699/01/0004-0275–0283/$8.00 + $.50 per page. Paper No. 22205.

$$\hat{P}[X \geq y_l] = \frac{1}{n+1} \sum_{i=1}^{l} 1, \quad l = 1, \ldots, n \tag{2}$$

If the values in the random sample are annual extremes, as are the storm frequencies considered here, then the ''recurrence interval'' or ''return period'' is defined as the average number of years during which an event may be expected to occur once. It is computed as the reciprocal of the exceedance probability.

This estimator is called ''nonparametric'' since knowledge of the underlying distribution and its parameters is not required. Other estimators (called ''parametric'') derive from knowledge (or supposition) of the type of underlying distribution. Functions of a random sample may be used as estimators of the parameters of the underlying distribution. Several of interest here are the ''sample mean,'' $\hat{\mu}$, ''sample variance,'' $\hat{\sigma}^2$, and ''sample skew coefficient,'' $\hat{\psi}$

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{3a}$$

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \hat{\mu})^2 \tag{3b}$$

$$\hat{\psi} = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} (x_i - \hat{\mu})^3/(\sqrt{\hat{\sigma}^2})^3 \tag{3c}$$

They are estimators of distribution mean, μ; variance, σ^2; and skew coefficient, ψ, respectively. Other estimators (Koutrouvelis and Canavos 1999) also could be used with no loss of generality.

EXAMPLE STORM FREQUENCIES

The daily flow records of the Maumee River at Waterville, Ohio (basin area = 16,390 km^2) were assembled and searched over 1948–1995 and the annual maximum daily flows are given in Table 1. The exceedance frequencies for the annual maximum flows were estimated with (2) and plotted in Fig. 1. The log-Pearson Type III distribution also was fit to the data set of Table 1. This distribution results from supposing the natural logarithms of the data in Table 1 [$Z = \ln(X)$] are distributed as a three-parameter gamma distribution

$$f_Z(z) = \frac{1}{|\beta|\Gamma(\alpha)} \left(\frac{z-c}{\beta}\right)^{\alpha-1} e^{-(z-c)/\beta}, \quad \begin{matrix} c \leq z < \infty \ (\beta > 0) \\ -\infty < z \leq c \ (\beta < 0) \end{matrix} \tag{4}$$

where $f_Z(z) = (\partial/\partial z)P[Z \leq z]$; $\Gamma(\alpha)$ = gamma function; and α, β, and c = distribution parameters. Estimates of the parameters are given in terms of (3) defined on the natural logarithms of the data (USWRC 1967), by replacing expected values with sample moments

$$\hat{\alpha} = (2/\hat{\psi})^2 \tag{5a}$$

$$\hat{\beta} = \sqrt{\hat{\sigma}^2}\hat{\psi}/2 \tag{5b}$$

$$\hat{c} = \hat{\mu} - 2\sqrt{\hat{\sigma}^2}/\hat{\psi} \tag{5c}$$

The estimated log-Pearson Type III distribution is shown also in Fig. 1. See Koutrouvelis and Canavos (1999) for other parameter estimators.

MATCHING A PROBABILITY FORECAST

The probability of any event A, $P[A]$, can be inferred with the estimator, $\hat{P}[A]$, defined as the number of observations in the random sample for which A occurs (i.e., for which the event A is true), n_A, divided by the total number of observations in the sample, n

TABLE 1. Annual Maximum Daily Maumee River Flow at Waterville, Ohio (Latitude 41:30:00, Longitude 83:42:46; Basin Area = 16,390 km^2)

Year	Flow (m^3 s^{-1})
1949	1280
1950	2620
1951	1500
1952	1500
1953	940
1954	663
1955	1300
1956	1210
1957	1770
1958	841
1959	2270
1960	1270
1961	1520
1962	1300
1963	1000
1964	1330
1965	1030
1966	2240
1967	1380
1968	1610
1969	1910
1970	943
1971	1100
1972	1330
1973	1130
1974	1970
1975	1400
1976	1940
1977	1810
1978	2450
1979	1510
1980	1260
1981	2420
1982	3200
1983	1530
1984	1450
1985	2580
1986	1030
1987	666
1988	649
1989	1210
1990	2320
1991	2460
1992	1530
1993	1840
1994	1810
1995	1440

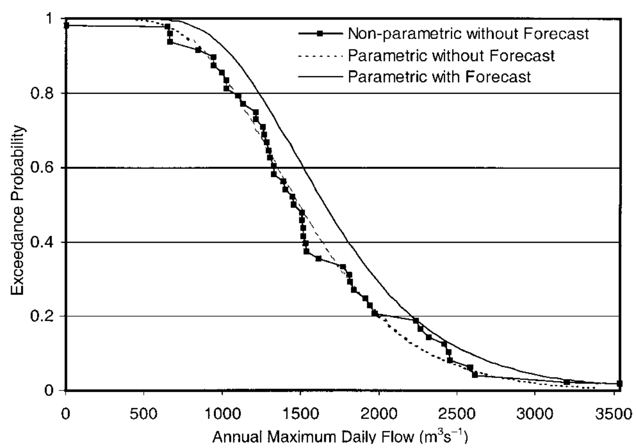

FIG. 1. Annual Maximum Daily Maumee River Flow Exceedance Frequency, Made in September 1999 for Period of September 1999–August 2000

$$\hat{P}[A] = \frac{n_A}{n} = \frac{1}{n}\sum_{i|A} 1 \qquad (6)$$

In (6), the count (sum) is taken over all i (members of the random sample) for which A occurs. The estimator in (6) is recognized as the ''relative frequency'' of A in the random sample. Other estimators, such as the one in the Weibull method, $\hat{P}[A] = n_A/(n + 1)$, do not have all of the properties associated with probability measures (Pfeiffer 1965) and cannot be used in the following derivations. However, (6) does, and derivations with (6) will behave as probabilities.

A forecast associates a probability with a value, reflecting the anticipated likelihood of an event defined in terms of that value. Given an event A for which a probability forecast is desired and an event A_k for which others have forecast a probability ($P[A_k]$), what is the forecast probability of A? By the theorem of total probability (Pfeiffer 1965)

$$P[A] = P[A|A_k]P[A_k] + P[A|A_k^C]P[A_k^C] \qquad (7)$$

If estimates of the conditional probabilities, $P[A|A_k]$ and $P[A|A_k^C]$, are taken from the historical record, then knowing $P[A_k]$ (and $P[A_k^C] = 1 - P[A_k]$) from others' forecasts enables estimation of the forecast probability of A by (7). Eq. (7) can be written in terms of probability estimates; by using the definition of conditional probability

$$P[A|A_k] = P[AA_k]/P[A_k] \qquad (8)$$

and by replacing with sample counts, n_A [number of scenarios in the sample for which event A occurs, as in (6)]

$$\hat{P}[A] = \frac{n_{AA_k}}{n_{A_k}}\hat{P}[A_k] + \frac{n_{AA_k^C}}{n_{A_k^C}}\hat{P}[A_k^C] \qquad (9a)$$

$$\hat{P}[A] = \sum_{i|AA_k}\frac{1}{n_{A_k}}\hat{P}[A_k] + \sum_{i|AA_k^C}\frac{1}{n_{A_k^C}}\hat{P}[A_k^C] \qquad (9b)$$

$$\hat{P}[A] = \frac{1}{n}\sum_{i|A} w_i \qquad (9c)$$

where

$$w_i = \frac{n}{n_{A_k}}\hat{P}[A_k] \quad \forall\ i|A_k \qquad (10a)$$

$$w_i = \frac{n}{n_{A_k^C}}\hat{P}[A_k^C] \quad \forall\ i|A_k^C \qquad (10b)$$

BAYESIAN COMPARISON

Note that in the example of (7)–(10), while $\hat{P}[A|A_k]$ and $\hat{P}[A|A_k^C]$ are unchanged from the historical record estimates, $\hat{P}[A]$ and (of course) $\hat{P}[A_k]$ are changed from the historical record estimates of (6), reflecting the biasing of the sample to match the forecast of $P[A_k]$. Furthermore, one can easily show that $\hat{P}[AA_k]$ and $\hat{P}[A_k|A]$ are also changed from the historical record estimates. This reveals an important distinction between the Bayesian statistic approach (Pfeiffer 1965) for estimation of conditional probabilities and the biased-sampling approach used here. In the Bayesian approach, the probability of an event A, conditioned on the occurrence of an event A_k (called ''*a posterior*,'' $P[A|A_k]$) is estimated from an unconditional probability (*a priori*, $P[A]$), a ''likelihood'' function ($P[A_k|A]$), both estimated from the record, and an experimental observation (in this case a forecast, $P[A_k]$). In terms of estimators, the Bayes theorem is

$$\hat{P}[A|A_k] = \frac{\hat{P}[A]\hat{P}[A_k|A]}{\hat{P}[A_k]} \qquad (11)$$

In the biased sample approach used here, one estimates new joint probabilities (for A and A_k) that preserve observed conditional probabilities ($P[A|A_k]$ and $P[A|A_k^C]$), conditioned on a key event A_k while matching that key event (forecast) probability $P[A_k]$. The theorem of total probability allows one to calculate a new probability for the event of interest (A) as in (7). Thus, the biased sample procedure uses conditional probabilities observed in the record rather than estimating them anew from forecasts of selected meteorological events. It estimates the probability of the event of interest from a new joint distribution that also matches the meteorological event forecast.

MATCHING MULTIPLE PROBABILITY FORECASTS

Croley (1996, 1997b, 2000a) biased samples, by multiplying sample observations by nonnegative weights, w_i, to calculate probabilities of any event A, as in (9), but where the weights are determined from matching others' multiple probability forecasts of events. These other probability forecasts are read from agency forecast maps for a point of interest. Croley (1996, 1997b, 2000a) reviews the basis for a variety of these forecasts and how to read the maps; he also derives El Niña forecast probabilities for an application area (Croley 2000a). Note from (9), for A selected so that $\hat{P}[A] = 1$

$$\sum_{i=1}^{n} w_i = n \qquad (12)$$

Consider, for example, that forecasts of event probability can be interpreted in $m - 1$ probability equations (Croley 1996) and forecasts of most-probable events can be interpreted in $p + q$ probability inequalities (Croley 1997b). They are expressed in terms of relative frequencies over a random sample as follows:

$$\hat{P}[A_k] = a_k, \quad k = 2, \ldots, m \qquad (13a)$$

$$\hat{P}[A_k] < a_k, \quad k = m + 1, \ldots, m + p \qquad (13b)$$

$$\hat{P}[A_k] \le a_k, \quad k = m + p + 1, \ldots, m + p + q \qquad (13c)$$

where a_k = forecast probabilities. Croley (2000a) illustrates how to interpret agency forecast probability maps to yield probability statements in one of the forms in (13). For now, replace the ''strictly less than'' inequalities in (13) with ''less than or equal to'' inequalities

$$\hat{P}[A_k] = a_k, \quad k = 2, \ldots, m \qquad (14a)$$

$$\hat{P}[A_k] \le a_k, \quad k = m + 1, \ldots, m + u \qquad (14b)$$

where $u = p + q$. Writing the forecasts of meteorological probabilities of (14) in terms of weights, as in (9), and adding to (12), yields a system of equations to be solved for the weights

$$\sum_{i=1}^{n} w_i = n \qquad (15a)$$

$$\sum_{i|A_k} w_i = na_k, \quad k = 2, \ldots, m \qquad (15b)$$

$$\sum_{i|A_k} w_i \le na_k, \quad k = m + 1, \ldots, m + u \qquad (15c)$$

Equivalently

$$\sum_{i=1}^{n} \alpha_{k,i} w_i = e_k, \quad k = 1, \ldots, m \qquad (16a)$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i \le e_k, \quad k = m + 1, \ldots, m + u \qquad (16b)$$

where $\alpha_{k,i} = 1$ for $k = 1$ or for $k \neq 1$ and $i|A_k$ [inclusion of the ith random sample value (ith event or segment of the historical record) in the event of the kth probability statement]; otherwise it is zero. Also, $e_k = n$ for $k = 1$ and $e_k = na_k$ for $k > 1$. Any set of weights that satisfies (16) yields weighted-sample relative frequencies of events that match forecasts of meteorological probabilities. These weights can also be used to yield other corresponding biased sample estimators; e.g., (2) and (3) become, for nonzero weights

$$\hat{P}[X > y_l] = \frac{1}{n+1}\sum_{k=1}^{l} w_{i(k)}, \quad l = 1, \ldots, n \tag{17a}$$

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} w_i x_i \tag{17b}$$

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n} w_i(x_i - \hat{\mu})^2 \tag{17c}$$

$$\hat{\psi} = \frac{n}{(n-1)(n-2)}\sum_{i=1}^{n} w_i(x_i - \hat{\mu})^3/(\sqrt{\hat{\sigma}^2})^3 \tag{17d}$$

Observations (x_i) corresponding to zero-valued weights are unused and the smaller sample size (d) replaces n in statistics such as (17). The weights are multiplied by d/n to sum to the smaller sample size as in (12).

This method of using observations from the historical record is similar to Extended Streamflow Prediction (ESP) methodology (Day 1985; Ingram et al. 1995; Smith et al. 1992). In ESP, meteorological observations from the historical record (time series segments) are transformed with hydrology models to yield a sample of possible ''futures'' that can then be described probabilistically. Croley (1996, 1997b, 2000a) describes how to bias this sample, to reflect weather forecasts, in this ESP approach (operational hydrology). Here, single observations from the historical record of either meteorology or hydrology (annual maximums instead of time series segments) are used directly (no model transformation) to estimate a simple exceedance probability distribution, which is then biased to reflect climatic forecasts.

Generally, some of the equations in (15) or (16) may be either redundant or infeasible (nonintersecting with the rest, resulting in no solutions) and must be eliminated. (If the number of equations is greater than the number of weights, then some of the equations must be either redundant or infeasible.) In practice, one could assign each equation in (15) or (16) a priority reflecting its importance. The highest priority is given to (15*a*) or (16*a*) corresponding to (12), guaranteeing that all relative frequencies sum to unity. Each equation (starting with the second highest priority equation) is compared to the set of all higher-priority equations and eliminated if redundant or infeasible. Thus (16) can always be reduced so that the allowed number of forecasts of meteorological probabilities is less than or equal to the number of historical record pieces (sample size). If less, then there are multiple solutions to (16), and a choice must be made as to which solution to use.

OPTIMUM SOLUTION

If there are multiple solutions to (16), the identification of the best requires a measure or objective function for comparing them. Solutions of (16) with larger values of this measure can be judged better than those with smaller values. One such measure is the probability of a selected event. Consider an example: (14) allows probabilities equal to a_k for $k = m + 1, \ldots, m + p$, while (13) requires them strictly smaller. One can use an objective function to minimize a probability to keep it under its limit (instead of equal to it). Likewise, one can use an objective function to maximize a probability (increase it above a limit) as in the following:

$$\max \hat{P}[T_{\mathrm{Sep}} > \hat{\tau}_{\mathrm{Sep},0.5}] \quad \text{subject to} \tag{18a}$$

$$\hat{P}[T_{\mathrm{Sep}} > \hat{\tau}_{\mathrm{Sep},0.5}] \geq 0.5 \tag{18b}$$

$$\hat{P}[T_{\mathrm{Sep}} \leq \hat{\tau}_{\mathrm{Sep},0.5}] \leq 0.5 \tag{18c}$$

where T_{Sep} = September air temperature; and $\hat{\tau}_{\mathrm{Sep},\gamma}$ is the γ-quantile for September air temperature estimated from a reference historical period (usually 1961–1990 or 1963–1993) such that

$$\hat{P}[T_{\mathrm{Sep}} \leq \hat{\tau}_{\mathrm{Sep},\gamma}] = \gamma \tag{19}$$

This allows us to work only with less-than-or-equal-to inequalities instead of strictly-less-than inequalities, or with greater-than-or-equal-to inequalities instead of strictly-greater-than inequalities. Therefore, the replacement of (13) with (14) is still useful. Also, consider the example

$$\max(\hat{P}[T_{\mathrm{Sep}} \leq \hat{\tau}_{\mathrm{Sep},0.333}] + \hat{P}[Q_{\mathrm{Sep}} > \hat{\theta}_{\mathrm{Sep},0.667}]) \quad \text{subject to} \tag{20a}$$

$$\hat{P}[T_{\mathrm{Sep}} \leq \hat{\tau}_{\mathrm{Sep},0.333}] \geq 0.333 \tag{20b}$$

$$\hat{P}[\hat{\tau}_{\mathrm{Sep},0.333} < T_{\mathrm{Sep}} \leq \hat{\tau}_{\mathrm{Sep},0.667}] \leq 0.334 \tag{20c}$$

$$\hat{P}[T_{\mathrm{Sep}} > \hat{\tau}_{\mathrm{Sep},0.667}] \leq 0.333 \tag{20d}$$

$$\hat{P}[Q_{\mathrm{Sep}} \leq \hat{\theta}_{\mathrm{Sep},0.333}] \leq 0.333 \tag{20e}$$

$$\hat{P}[\hat{\theta}_{\mathrm{Sep},0.333} < Q_{\mathrm{Sep}} \leq \hat{\theta}_{\mathrm{Sep},0.667}] \leq 0.334 \tag{20f}$$

$$\hat{P}[Q_{\mathrm{Sep}} > \hat{\theta}_{\mathrm{Sep},0.667}] \geq 0.333 \tag{20g}$$

where September precipitation, Q_{Sep}, and its quantiles, $\hat{\theta}_{\mathrm{Sep},\gamma}$, are defined similarly to (19). The optimization of (18) satisfies a most-probable event forecast for above-normal September air temperature by maximizing the associated probability. The optimization of (20) attempts to satisfy two most-probable September event forecasts (one for below-normal air temperature and one for above-normal precipitation) by maximizing a sum of associated probabilities.

The objective function formulation can be used to express many different goals, besides the two in (18*a*) and (20*a*) for giving extreme probabilities that match meteorological forecasts. Another example expresses the goal of retreating from high temperatures beyond the near term

$$\min\{\hat{P}[T_{\mathrm{Dec}} > \hat{\tau}_{\mathrm{Dec},0.667}]\} \quad \text{subject to} \tag{21a}$$

$$\hat{P}[T_{\mathrm{Sep}} \leq \hat{\tau}_{\mathrm{Sep},0.333}] = 0.233 \tag{21b}$$

$$\hat{P}[T_{\mathrm{Sep}} > \hat{\tau}_{\mathrm{Sep},0.667}] = 0.433 \tag{21c}$$

If the objective function is always a statement of maximizing or minimizing a probability, as in the examples of (18*a*), (20*a*), and (21*a*), then it can be added to the problem statement of (16) to yield an optimization problem solvable with a standard optimization technique. In particular, the objective function formulations of (18*a*), (20*a*), and (21*a*) can be expressed in terms of weights by matching relative frequencies in the sample of historical record segments, as was done to replace (12) and (14) with (15) by using (9). The reformulation of the objective functions of (18*a*), (20*a*), and (21*a*) become, respectively

$$\max\left[\frac{1}{n}\sum_{i|(t_{\mathrm{Sep}})_i > \hat{\tau}_{\mathrm{Sep},0.5}} w_i\right] \tag{22}$$

$$\max\left[\frac{1}{n}\sum_{i|(t_{\mathrm{Sep}})_i \leq \hat{\tau}_{\mathrm{Sep},0.333}} w_i + \frac{1}{n}\sum_{i|(q_{\mathrm{Sep}})_i > \hat{\theta}_{\mathrm{Sep},0.667}} w_i\right] \tag{23}$$

$$\max\left[\frac{1}{n}\sum_{i|\mathrm{not}[(t_{\mathrm{Dec}})_i > \hat{\tau}_{\mathrm{Dec},0.667}]} w_i\right] \tag{24}$$

where $(t_j)_i$ and $(q_j)_i$ are period j air temperature and precipitation sample observations i.

Each example, from (22)–(24), can in turn be equivalently expressed in the general form

$$\max \sum_{i=1}^{n} \alpha_{0,i} w_i \tag{25}$$

where $\alpha_{0,i}$ are defined similarly to (16) in which the objective function corresponds to $k = 0$. The problem of solving (16) can now be formulated as an optimization, maximizing the objective function subject to a ''constraint set'' of equations

$$\max \sum_{i=1}^{n} \alpha_{0,i} w_i \quad \text{subject to} \tag{26a}$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i = e_k, \quad k = 1, \ldots, m \tag{26b}$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i \le e_k, \quad k = m + 1, \ldots, m + u \tag{26c}$$

$$w_i \ge 0, \quad i = 1, \ldots, n \tag{26d}$$

LINEAR PROGRAMMING

Eqs. (26) are amenable to standard linear programming optimization techniques. An existing algebraic procedure, termed the Simplex method, has been developed that progressively approaches the optimum solution through a well-defined iterative process until optimality is finally reached; Hillier and Lieberman (1969) provide an explanation of the Simplex method and details of its use. The following is not an account of the Simplex method, but of an adaptation of it in face of potentially infeasible solutions and the need to further restrict allowable equations to allow feasible solutions. Prior to application of the Simplex method, the equations and inequalities in (26) are transformed into an equivalent two-stage problem. First, (26) may be written all in terms of inequalities,

$$\max \sum_{i=1}^{n} \alpha_{0,i} w_i \quad \text{subject to} \tag{27a}$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i \le e_k, \quad k = 1, \ldots, m \tag{27b}$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i \ge e_k, \quad k = 1, \ldots, m \tag{27c}$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i \le e_k, \quad k = m + 1, \ldots, m + u \tag{27d}$$

$$w_i \ge 0, \quad i = 1, \ldots, n \tag{27e}$$

where the equalities in (26*b*) have been replaced (equivalently) with two sets of inequalities in (27*b*) and (27*c*). The solution to (27) is identical to that of (26). Furthermore, the greater-than-or-equal-to inequalities in (27*c*) can be summed into a single equation without changing the solution

$$\max \sum_{i=1}^{n} \alpha_{0,i} w_i \quad \text{subject to} \tag{28a}$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i \le e_k, \quad k = 1, \ldots, m + u \tag{28b}$$

$$\sum_{k=1}^{m} \sum_{i=1}^{n} \alpha_{k,i} w_i \ge \sum_{k=1}^{m} e_k \tag{28c}$$

$$w_i \ge 0, \quad i = 1, \ldots, n \tag{28d}$$

Then, (28) can be written as equalities by adding ''slack'' variables

$$\max \sum_{i=1}^{n} \alpha_{0,i} w_i \quad \text{subject to} \tag{29a}$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i + w_{n+k} = e_k, \quad k = 1, \ldots, m + u \tag{29b}$$

$$\sum_{k=1}^{m} \sum_{i=1}^{n} \alpha_{k,i} w_i - w_{n+m+u+1} = \sum_{k=1}^{m} e_k \tag{29c}$$

$$w_i \ge 0, \quad i = 1, \ldots, n + m + u + 1 \tag{29d}$$

where w_i, ($i = n + 1, \ldots, n + m + u + 1$) are nonnegative slack variables. The reason that equalities in (26) were first eliminated and then restored is so that a slack variable is introduced for every equation. These slack variables enable an initial solution (set of values for w_i, $i = 1, \ldots, n + m + u + 1$ that satisfies the constraints) if one exists, from which to begin the Simplex search for the optimum. An initial solution can be obtained from an optimization similar to (29)

$$\max -v \quad \text{subject to} \tag{30a}$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i + w_{n+k} = e_k, \quad k = 1, \ldots, m + u \tag{30b}$$

$$\sum_{k=1}^{m} \sum_{i=1}^{n} \alpha_{k,i} w_i - w_{n+m+u+1} + v = \sum_{k=1}^{m} e_k \tag{30c}$$

$$w_i \ge 0, \quad i = 1, \ldots, n + m + u + 1 \tag{30d}$$

$$v \ge 0 \tag{30e}$$

where v = ''artificial'' variable introduced as a computational device. The maximization of $-v$ corresponds to the minimization of v. If the minimum occurs at $v = 0$, then the solution to (30) is feasible in (29), and the Simplex search [as described by Hillier and Lieberman (1969)] can begin in (29) from this solution. If the minimum v is not zero, or no feasible solution to (30) exists, then there is no feasible solution to (29). This means the constraint set must be changed by eliminating the lowest-priority equation or inequality in (26) and its corresponding member in (29) and (30). This cycle of optimization of (30) and elimination of the lowest-priority equation is repeated until a feasible solution is found to the problem of (30); i.e., the artificial variable in (30) equals zero. At this point, the Simplex method is applied to (29) from the solution of (30). Croley (2000a) describes the procedure in more detail. Fig. 2 depicts the algorithm of the entire linear programming optimization and successive elimination of infeasible equations.

NOAA and EC FORECAST EXAMPLE

The estimates of Fig. 1 are modified by incorporating selected forecasts from the National Oceanic and Atmospheric Administration (NOAA) event probability forecasts and Environment Canada (EC) most-probable event forecasts (Croley 1996, 1997b, 2000a). The forecasts are summarized in Table 2, in a priority order where the earliest-made forecasts are placed first (the earlier NOAA event-probability equations precede the later EC most-probable event inequalities), precipitation outlooks precede temperature, and shorter-lagged outlooks precede longer-lagged. Note that the precipitation forecasts in Table 2 are for high precipitation with only one exception (the EC September-October-November, or SON, forecast). The objective in matching these forecasts is therefore

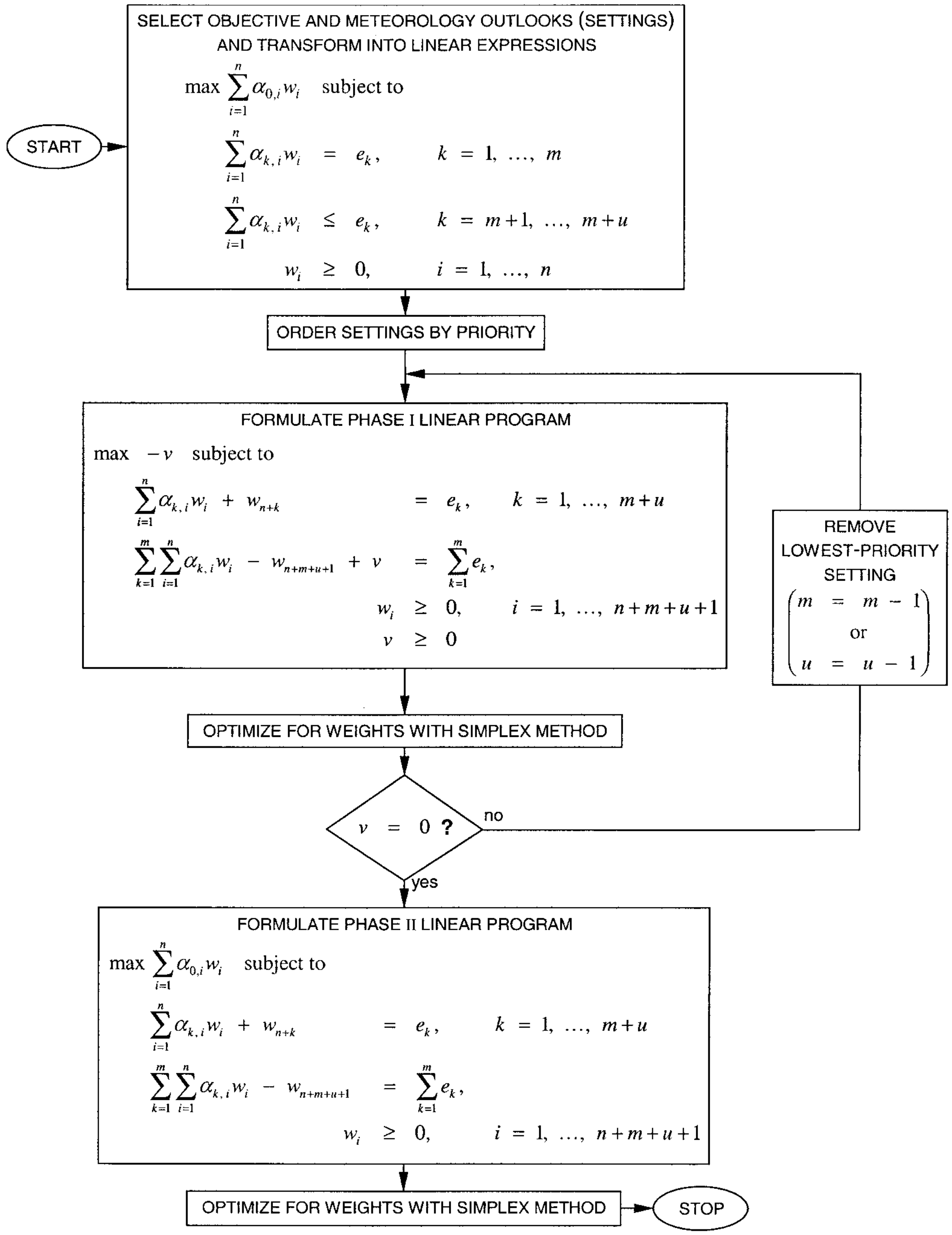

FIG. 2. Determining Physically Relevant Weights through Linear Programming

arbitrarily taken here as maximizing the probability that precipitation over the period November 1999–July 2000 will be in the upper third of its historical range (determined from 1961–1990).

$$\max \hat{P}[Q_{\text{Nov'99–Jul'00}} > \hat{\theta}_{\text{Nov–Jul},0.667}] \qquad (31)$$

The daily meteorological data over the entire Maumee River basin were assembled to determine the average air temperature and precipitation, for the periods shown in Table 2 and (31); only average daily precipitation for the October-November-December (OND) and November–July periods are presented here and given in Table 3. According to the agencies, the NOAA temperature and precipitation forecasts and the EC precipitation forecasts are defined relative to historical reference quantiles estimated over 1961–1990. Likewise, the EC temperature forecasts are defined relative to historical reference quantiles estimated over 1963–1993. By ordering data from these periods, the reference quantiles are estimated; example values are given only for OND and November–July precipitation in Table 4 as derived from the data in Table 3.

Suppose now that a hydrologist has acquired, in September 1999, the forecasts of Table 2 and wishes to make an estimate of storm frequencies (annual maximum Maumee River flows) at that time for the coming winter and following spring and summer (September 1999–August 2000). He or she would consider each of the possibilities in Table 1 as a possibility for this period. (The Maumee River annual maximum flow typically occurs as spring snowmelt.) The objective function of (31) and the forecasts of Table 2 would apply prior and through the beginning of each year in the sample, so that the time lag accounts for the meteorology driving the hydrology. In other words, each year of record is to be weighted to reflect the objective of (31) and the beginning winter as forecast in Table 2 (a total period from September of the year before through the following August). For example, the first value in Table 1 for 1949 corresponds to the objective and forecast values for September 1948–August 1949. The coefficients in

TABLE 2. Mixed NOAA and EC Probabilistic Meteorological Forecasts Made in September 1999 over Maumee River Basin

No.	Equation
1	$\hat{P}[Q_{\mathrm{OND'99}} \le \hat{\theta}_{\mathrm{OND},0.333}] = 0.283$
2	$\hat{P}[Q_{\mathrm{OND'99}} > \hat{\theta}_{\mathrm{OND},0.667}] = 0.383$
3	$\hat{P}[Q_{\mathrm{NDJ'99}} \le \hat{\theta}_{\mathrm{NDJ},0.333}] = 0.273$
4	$\hat{P}[Q_{\mathrm{NDJ'99}} > \hat{\theta}_{\mathrm{NDJ},0.667}] = 0.393$
5	$\hat{P}[Q_{\mathrm{DJF'99}} \le \hat{\theta}_{\mathrm{DJF},0.333}] = 0.273$
6	$\hat{P}[Q_{\mathrm{DJF'99}} > \hat{\theta}_{\mathrm{DJF},0.667}] = 0.393$
7	$\hat{P}[Q_{\mathrm{JFM'00}} \le \hat{\theta}_{\mathrm{JFM},0.333}] = 0.133$
8	$\hat{P}[Q_{\mathrm{JFM'00}} > \hat{\theta}_{\mathrm{JFM},0.667}] = 0.533$
9	$\hat{P}[Q_{\mathrm{FMA'00}} \le \hat{\theta}_{\mathrm{FMA},0.333}] = 0.273$
10	$\hat{P}[Q_{\mathrm{FMA'00}} > \hat{\theta}_{\mathrm{FMA},0.667}] = 0.393$
11	$\hat{P}[Q_{\mathrm{MAM'00}} \le \hat{\theta}_{\mathrm{MAM},0.333}] = 0.273$
12	$\hat{P}[Q_{\mathrm{MAM'00}} > \hat{\theta}_{\mathrm{MAM},0.667}] = 0.393$
13	$\hat{P}[T_{\mathrm{OND'99}} \le \hat{\tau}_{\mathrm{OND},0.333}] = 0.333$
14	$\hat{P}[T_{\mathrm{OND'99}} > \hat{\tau}_{\mathrm{OND},0.667}] = 0.333$
15	$\hat{P}[T_{\mathrm{NDJ'99}} \le \hat{\tau}_{\mathrm{NDJ},0.333}] = 0.333$
16	$\hat{P}[T_{\mathrm{NDJ'99}} > \hat{\tau}_{\mathrm{NDJ},0.667}] = 0.333$
17	$\hat{P}[T_{\mathrm{DJF'99}} \le \hat{\tau}_{\mathrm{DJF},0.333}] = 0.273$
18	$\hat{P}[T_{\mathrm{DJF'99}} > \hat{\tau}_{\mathrm{DJF},0.667}] = 0.393$
19	$\hat{P}[T_{\mathrm{JFM'00}} \le \hat{\tau}_{\mathrm{JFM},0.333}] = 0.263$
20	$\hat{P}[T_{\mathrm{JFM'00}} > \hat{\tau}_{\mathrm{JFM},0.667}] = 0.403$
21	$\hat{P}[T_{\mathrm{FMA'00}} \le \hat{\tau}_{\mathrm{FMA},0.333}] = 0.333$
22	$\hat{P}[T_{\mathrm{FMA'00}} > \hat{\tau}_{\mathrm{FMA},0.667}] = 0.333$
23	$\hat{P}[T_{\mathrm{MAM'00}} \le \hat{\tau}_{\mathrm{MAM},0.333}] = 0.333$
24	$\hat{P}[T_{\mathrm{MAM'00}} > \hat{\tau}_{\mathrm{MAM},0.667}] = 0.333$
25	$\hat{P}[Q_{\mathrm{SON'99}} \le \hat{\theta}_{\mathrm{SON},0.333}] \le 0.333$
26	$\hat{P}[\hat{\theta}_{\mathrm{SON},0.333} < Q_{\mathrm{SON'99}} \le \hat{\theta}_{\mathrm{SON},0.667}] \ge 0.334$
27	$\hat{P}[Q_{\mathrm{SON'99}} > \hat{\theta}_{\mathrm{SON},0.667}] \le 0.333$
28	$\hat{P}[Q_{\mathrm{JJA'00}} \le \hat{\theta}_{\mathrm{JJA},0.333}] \le 0.333$
29	$\hat{P}[\hat{\theta}_{\mathrm{JJA},0.333} < Q_{\mathrm{JJA'00}} \le \hat{\theta}_{\mathrm{JJA},0.667}] \le 0.334$
30	$\hat{P}[Q_{\mathrm{JJA'00}} > \hat{\theta}_{\mathrm{JJA},0.667}] \ge 0.333$
31	$\hat{P}[T_{\mathrm{SON'99}} \le \hat{\tau}_{\mathrm{SON},0.333}] \le 0.333$
32	$\hat{P}[\hat{\tau}_{\mathrm{SON},0.333} < T_{\mathrm{SON'99}} \le \hat{\tau}_{\mathrm{SON},0.667}] \le 0.334$
33	$\hat{P}[T_{\mathrm{SON'99}} > \hat{\tau}_{\mathrm{SON},0.667}] \ge 0.333$
34	$\hat{P}[T_{\mathrm{DJF'99}} \le \hat{\tau}_{\mathrm{DJF},0.333}] \le 0.333$
35	$\hat{P}[\hat{\tau}_{\mathrm{DJF},0.333} < T_{\mathrm{DJF'99}} \le \hat{\tau}_{\mathrm{DJF},0.667}] \le 0.334$
36	$\hat{P}[T_{\mathrm{DJF'99}} > \hat{\tau}_{\mathrm{DJF},0.667}] \ge 0.333$
37	$\hat{P}[T_{\mathrm{JJA'00}} \le \hat{\tau}_{\mathrm{JJA},0.333}] \ge 0.333$
38	$\hat{P}[\hat{\tau}_{\mathrm{JJA},0.333} < T_{\mathrm{JJA'00}} \le \hat{\tau}_{\mathrm{JJA},0.667}] \le 0.334$
39	$\hat{P}[T_{\mathrm{JJA'00}} > \hat{\tau}_{\mathrm{JJA},0.667}] \le 0.333$

(26), $\alpha_{k,i}$, have values of 1 or 0 corresponding to the inclusion or exclusion, respectively, of each variable in the sets indicated in the variable subscripts in (31) and Table 2. For (31), the reader can see from inspection of the third column in Table 3 that the relation, $q_{\mathrm{Nov-Jul}} > \hat{\theta}_{\mathrm{Nov-Jul},0.667}$ (or $q_{\mathrm{Nov-Jul}} > 2.62$ mm; see the third column in Table 4) is satisfied by the following indices: 1 (corresponding to 1948), 2, 3, 4, 11, 19, 20, 21, 25, 27, 32, 35, 38, 41, 42, and 45 (corresponding to 1992). Index 19 (corresponding to 1966) does not appear to satisfy this relation because of round-off error, but in fact does. Thus, (31) would be written, similar to (26*a*), as

$$\max(w_1 + w_2 + w_3 + w_4 + w_{11} + w_{19} + w_{20} + w_{21} + w_{25} + w_{27} + w_{32} + w_{35} + w_{38} + w_{41} + w_{42} + w_{45}) \quad (32)$$

Expressing (32) in vector form,

$$\max[1111000000010000000011100010100001001001001100100]\mathbf{w} \quad (33)$$

where $\mathbf{w}$ = column vector of the weights. Similarly, the reader can see from inspection of Tables 3 and 4 that (12) and the first two equations in Table 2 become the vector equations

$$[11111111111111111111111111111111111111111111111]\mathbf{w} = 47 \quad (34a)$$

TABLE 3. Average Maumee River Basin Daily Precipitation (mm)

Year	OND	November–July
1948	2.52	2.88
1949	2.03	3.24
1950	2.87	2.91
1951	2.86	2.82
1952	1.49	2.15
1953	0.96	2.29
1954	2.90	2.15
1955	2.61	2.49
1956	1.31	2.60
1957	2.84	2.53
1958	1.50	2.67
1959	2.74	2.34
1960	1.10	2.39
1961	1.58	1.98
1962	1.25	1.88
1963	0.86	2.24
1964	1.06	2.30
1965	2.42	2.00
1966	3.14	2.62
1967	3.38	2.81
1968	2.33	2.77
1969	2.01	2.49
1970	1.85	2.11
1971	2.13	2.48
1972	2.67	2.94
1973	2.58	2.43
1974	1.96	2.63
1975	2.10	2.51
1976	1.02	1.98
1977	2.28	2.37
1978	1.98	2.45
1979	2.46	2.71
1980	1.34	2.56
1981	2.04	2.46
1982	2.84	2.65
1983	3.81	2.62
1984	2.31	2.38
1985	3.17	3.00
1986	2.09	1.92
1987	2.27	1.71
1988	2.64	2.65
1989	1.50	2.81
1990	3.53	2.41
1991	2.55	2.51
1992	2.71	2.88
1993	1.85	2.10
1994	1.89	2.39

$$[00001100101010111110000010001010101000000000100011]\mathbf{w} = 0.283 \times 47 \quad (34b)$$

$$[10110011010100000011000011000000100110100101110 0]\mathbf{w} = 0.383 \times 47 \quad (34c)$$

Likewise, the remainder of the equations in Table 2 can be combined with inspection of the average air temperature and precipitation data and quantiles, not shown here, and used with (33) and (34) to construct the entire optimization problem statement equivalent of (26)

$$\begin{aligned} \max[\alpha_{0,1}\ \alpha_{0,2}\ \cdots\ \alpha_{0,47}]\ & \mathbf{w} \text{ subject to} \\ [\alpha_{1,1}\ \alpha_{1,2}\ \cdots\ \alpha_{1,47}]\ & \mathbf{w} = e_1 \\ [\alpha_{2,1}\ \alpha_{2,2}\ \cdots\ \alpha_{2,47}]\ & \mathbf{w} = e_2 \\ \vdots\qquad & \quad\vdots \\ [\alpha_{25,1}\ \alpha_{25,2}\ \cdots\ \alpha_{25,47}]\ & \mathbf{w} = e_{25} \\ [\alpha_{26,1}\ \alpha_{26,2}\ \cdots\ \alpha_{26,47}]\ & \mathbf{w} \le e_{26} \\ [\alpha_{27,1}\ \alpha_{27,2}\ \cdots\ \alpha_{27,47}]\ & \mathbf{w} \le e_{27} \\ \vdots\qquad & \quad\vdots \\ [\alpha_{40,1}\ \alpha_{40,2}\ \cdots\ \alpha_{40,47}]\ & \mathbf{w} \le e_{40} \end{aligned} \quad (35)$$

Note that the values of $\alpha_{k,i}$ and e_k for the first four lines of (35) are given in (33) and (34). In the ensuing optimization

TABLE 4. Historical Average Daily Precipitation Reference Quantiles (mm)

Quantile[a]	Period, i	
	OND	November–July
$\hat{\theta}_{i,0.333}$	1.98	2.37
$\hat{\theta}_{i,0.667}$	2.42	2.62

[a]Quantiles based on the period 1961–1990.

of (35) (free software is available over the web from http://www.glerl.noaa.gov/wr/outlookweights.html), 19 weights are zeroes, indicating that some of the historical record is not used. However, all but the last three equations in Table 2 and (35) are used [all forecasts except the EC most-probable June-July-August (JJA) air temperature forecast].

Climate-biased storm frequencies for the annual maximum daily flow can now be estimated by applying these weights to the data in Table 1 by using (17). Only results for the fitted log-Pearson Type III distribution are given in Fig. 1. Compare the log-Pearson Type III distribution derived from the parametric estimates without the forecasts to that made with the forecasts, in Fig. 1. There is a large shift, making all flows more likely to be exceeded.

The complete example, represented only partially here in (33), (34), and (35), is given by Croley (2000b) with all supporting information. In addition, Croley (2000b) also gives one other fully documented example of this type for ''water year'' annual maximum Maumee River flow exceedance frequencies. He also gives two other fully documented examples for annual maximum precipitation exceedance frequencies, which utilize specially derived El Niño and La Niña probability forecasts for the Maumee River basin, along with background for the derivations.

MULTIPLE SOLUTIONS

In the optimization of (26), all expressions are linear, including the objective function [(26*a*)]. This allows a linear programming optimization technique to be used as compared to earlier formulations (Croley 1996, 1997b). Those used the minimization of the sum of squared differences of each weight with unity, $\Sigma\ (w_i - 1)^2$, and employed classical differential calculus solutions for zero slope of the Lagrangian. The linear formulation proves superior in its ability to include alternative objectives (expressed as maximization of selected event probabilities).

Also, the formulation of (26) allows nonnegativity constraints on the weights to be explicitly included. This means that all solutions can be searched. The earlier formulations of Croley (1996, 1997b) lacked explicit inclusion of nonnegativity constraints for the weights. There, optimum solutions were considered and discarded (along with lowest-priority constraints) if nonnegativity constraints were unsatisfied. But that also discarded the many other possibilities that, while not optimum, might satisfy all constraints.

There is a trade-off however. Multiple optima solutions are now a possibility that did not exist before. In the search algorithms employed in the linear programming solutions, these multiple optima can be detected (that is, the existence of more than a single optimum can be discerned) but the systematic exploration of them can be extensive. These multiple optima (while infinite in extent) can always be described as weighted combinations of a finite number of solution points. In practical terms, there are three limitations to the search for multiple optima: computation storage (or computer memory extent), computation time, and the growth of numerical error. The first two are easily appreciated. As multiple solution points are discovered, they must be saved and compared to subsequent solution points to avoid repeating their discovery. The searching and comparisons also can take a great amount of time, depending upon the number of solution points. Numerical error is unavoidable on the computer and, while it can be minimized, can cause error growth. The linear programming algorithm involves a systematic search from one point in the constraint space to another, and an error in calculating one point can induce errors in the subsequent. While this appears manageable in most practical applications for finding the first solution point, it appears sometimes not to be manageable for finding a large number of additional solution points when multiple optima are present. In fact, in many such problems, error growth can limit searches for other optima more than computer memory or time constraints.

Since the linear programming solution of (26) may not be able to identify all multiple optima (solution points), the practitioner is faced with an interesting choice. He or she can either limit the number of optimum solution points returned by the optimization to the first few, or formulate the problem (choose an objective) so that the optimum is unique. In the first alternative, all optima are not generally found, so that other possibilities exist uninvestigated. Maybe some of these other possibilities would have been preferable (as determined from considerations outside of the problem as formulated), but remain unknown. In the second alternative, the practitioner may be facing the dilemma of electing not to solve the problem at hand, but instead (re)formulating the problem so that its solution behaves in a certain way. While this is a practical consideration, it is theoretically unappealing. Changing the problem so that it is solvable in a certain manner is not the same as solving the original problem. Nevertheless, there are many alternative, yet practical, problem formulations of interest and the dilemma is not further addressed here.

SUMMARY AND OBSERVATIONS

The methodology described herein allows one to recognize changing climate in the estimation of storm frequencies, removing one of the worst assumptions associated with this, which is that future probabilities are the same as the past. Existing forecasts of meteorological probabilities as well as other conditional probabilities [e.g., based on El Niño or La Niña events, see Croley (2000a,b)] can be used to bias storm frequency estimates for a changing climate. The methodology is adapted from earlier work that uses forecasts of meteorological probabilities to derive forecasts of consequent hydrologic probabilities in an operational hydrology approach, similar to ESP. Here, a linear objective function is used in the search for a solution, enabling both the incorporation of an event probability in the objective, and the use of existing linear programming optimization techniques. Because of the possibility of multiple optimums and more than one way to consider them, the practitioner is faced with an interesting choice. Either the problem can be reformulated (changing the objective probability statement or the forecast probabilities or order) to find an unique optimum, or the search can be restricted to the first few optima found.

The example presented here may be more representative of storm frequency estimation in an operational setting rather than in a design setting. Climate-biased storm frequencies were estimated, in effect, conditioned on meteorological forecasts. These conditions are current and are not generally regarded as applying over a very long time into the future. The resulting biased storm frequencies can only be considered applicable over the same time period as the meteorological forecasts used to condition them. The examples given here applied over the next several months are appropriate for use in an operational setting. If probabilities can be defined (estimated) corresponding to climate shifts expected from the present forward, then the resulting biased storm frequencies could be used in a design setting.

While the example presented here is for annual maximum daily flow rates, other frequency estimation problems can be similarly addressed. These include derivation of complete duration-area-intensity curves for precipitation, and annual minimum extreme events. In the latter case, estimation would be for the cumulative distribution function (probabilities of being less than or equal to given levels) rather than for exceedance frequencies; but they are handled similarly. Complete software, in the form of an easy-to-use interactive Windows graphical user interface, and all worked examples are available free of charge over the World Wide Web. The software, examples (including additional El Niño and La Niña exercises), and tutorial materials may be acquired in a self-installing file by visiting and downloading from the web site, http://www.glerl.noaa.gov/wr/outlookweights.html.

ACKNOWLEDGMENTS

This is GLERL contribution No. 1147.

REFERENCES

Chow, V. T. (1964). *Handbook of applied hydrology*, McGraw-Hill, New York, 8–29.

Croley, T. E., II (1996). Using NOAA's new climate outlooks in operational hydrology. *J. Hydrologic Engrg.*, ASCE, 1(3), 93–102.

Croley, T. E., II (1997a). ''Water resource predictions from meteorological probability forecasts.'' *Proc., Sustainability of Water Resour. under Increasing Uncertainty* (*Proc. Rabat Symp.*), IAHS Publication No. 240, D. Rosbjerg et al., eds., IAHS Press, Wallingford, U.K., 301–309.

Croley, T. E., II (1997b). ''Mixing probabilistic meteorology outlooks in operational hydrology.'' *J. Hydrologic Engrg.*, ASCE, 2(4), 161–168.

Croley, T. E., II. (2000a). *Using meteorology probability forecasts in operational hydrology*, ASCE, Reston, Va.

Croley, T. E. (2000b). ''Climate-corrected Storm-Frequency Examples.'' *NOAA Tech. Memo. ERL GLERL-118*, Great Lakes Environmental Research Laboratory, Ann Arbor, Mich.

Day, G. N. (1985). ''Extended streamflow forecasting using NWSRFS.'' *J. Water Resour. Plng. and Mgmt.*, ASCE, 111(2), 157–170.

Hillier, F. S., and Lieberman, G. J. (1969). ''Chapter 5: Linear programming.'' *Introduction to Operations Research*, Holden-Day, San Francisco, 127–171.

Ingram, J. J., Hudlow, M. D., and Fread, D. L. (1995). ''Hydrometeorological coupling for extended streamflow predictions.'' *Preprints of Conf. on Hydrology, 75th Annu. Meeting of the Am. Meteorological Soc.*, American Meteorology Society, Dallas, 186–191.

Koutrouvelis, I. A., and Canavos, G. C. (1999). ''Estimation in the Pearson type 3 distribution.'' *Wat. Resour. Res.*, 35(9), 2693–2704.

Pfeiffer, P. E. (1965). *Concepts of probability theory*, McGraw-Hill, New York.

Smith, J. A., Day, G. N., and Kane, M. D. (1992). ''Nonparametric framework for long-range streamflow forecasting.'' *J. Water Resour. Plng. and Mgmt.*, ASCE, 118(1), 82–92.

U.S. Water Resources Council (USWRC). (1967). ''A uniform technique for determining flood-flow frequency.'' *Bull. 15*, Hydrology Committee, Washington, D.C.

NOTATION

The following symbols are used in this paper:

A = event label;
A_k = event label for kth event;
a_k = probability in kth equation in set of forecasts of meteorological probabilities;
c = location parameter for three-parameter gamma distribution;
$\hat{c}$ = sample estimate of c;
e_k = value of kth equation in set, equivalent to forecast of meteorological probability;
$f_Z(z)$ = probability density function for three-parameter gamma distribution;
$i(l)$ = number of value in unordered random sample corresponding to lth order;
l = order of a value in sample of size n, ordered from largest (1) to smallest (n);
m = number of meteorological forecast equalities plus one;
n = number of random variables or observations in random sample (size of sample);
n_A = number of random sample observations for which A occurs (i.e., for which A is true);
$P[\cdot]$ = probability of event in brackets, representing its likelihood;
$\hat{P}[\cdot]$ = relative frequency in sample, of event in brackets, used to estimate probability;
p = number of most-probable meteorological event ''strictly-less-than'' inequalities;
Q_j = total precipitation over period j;
q = number of most-probable meteorological event ''less-than-or-equal-to'' inequalities;
$(q_j)_i$ = value of Q_j in observation i of random sample;
T_j = average air temperature over period j;
$(t_j)_i$ = value of T_j in observation i of random sample;
u = number of most-probable meteorological event inequalities;
v = ''artificial'' variable in Phase I of Simplex method for optimization;
$\mathbf{w}$ = column vector of weights, w_i; $i = 1, \ldots, n$;
w_i = weight applied to ith random sample observation;
X_i = ith random variable in random sample;
x_i = ith observation in random sample (value of X_i);
y_l = lth ordered value, from largest to smallest, for variable X in random sample;
Z = random variable, defined as natural logarithm of random variable X;
α = shape parameter for three-parameter gamma distribution;
$\hat{\alpha}$ = sample estimate of α;
$\alpha_{k,i}$ = integer coefficient equal to unity (1) for $k = 1$ or for $k \neq 1$ when ith random sample value (ith event or segment of historical record) is included in event of kth probability statement, and zero (0) otherwise;
β = scale parameter for three-parameter gamma distribution;
$\hat{\beta}$ = sample estimate of β;
$\Gamma(\alpha)$ = gamma function ($= \int_0^\infty e^{-x}x^{\alpha-1}\,dx$);
μ = population mean, defining central location of distribution;
$\hat{\mu}$ = sample mean, used as sample estimate of μ;
$\hat{\theta}_{j,\gamma}$ = reference total precipitation γ-probability quantile estimate for period j;
σ^2 = population variance, defining spread of distribution about its central location;
$\hat{\sigma}^2$ = sample variance, used as sample estimate of σ^2;
$\hat{\tau}_{j,\gamma}$ = reference average air temperature γ-probability quantile estimate for period j;
ψ = population skew coefficient, used to define asymmetry of distribution; and
$\hat{\psi}$ = sample skew coefficient, used as sample estimate of ψ.