

Figure 7-3. Alternative representation of Equation (6-3) and the first 24 equations in Figure 4-5.

differentials are set to zero, as in Equations (7-7) and (7-8). Table 7-15 presents the solution of Equation (7-16) found by using the *first* procedural algorithm in Figure 7-2 (using all scenarios). Table 7-16 presents the solution of Equation (7-16) found by using the *second* procedural algorithm in Figure 7-2 (maximizing use of meteorology outlooks). Checks reveal that  $\sum w_i^2 < 2n$  for both procedural algorithms; therefore, the

Index, i	Weight, w <sub>i</sub>	Index, i	Weight, w <sub>i</sub>	Index, i	Weight, w <sub>i</sub>
(1)	(2)	(3)	(4)	(5)	(6)
1	1.060475	17	1.766160	33	2.120593
2	2.768731	18	1.666037	34	2.734721
3	0	19	2.971981	35	1.360656
4	0	20	0.372594	36	0.453465
5	0	21	0.573141	37	1.114288
6	0.778769	22	3.028104	38	0.848485
7	0.743021	23	2.363683	39	0
8	0	24	2.020468	40	0.153503
9	0	25	0.086473	41	2.099952
10	0.007176	26	1.870353	42	1.595679
11	0	27	1.964278	43	0.351949
12	1.744236	28	0.966552	44	0
13	0	29	0.942909	45	0.424029
14	2.047606	30	0.275731	46	0.933164
15	0	31	0.087902	47	0
16	0	32	2.703132		

 Table 7-13. Outlook weights maximizing use of meteorology outlooks in Figure

 7-3 for the Lake Superior supply outlook example.

Month	Nonexceedance quantiles								
	3%	10%	20%	30%	50%	70%	80%	90%	97%
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sep 98	-21.1	-13.9	-3.39	2.38	31.7	43.2	52.2	57.3	74.8
Oct 98	-33.6	-2.10	11.2	18.7	24.2	54.7	68.1	93.2	112
Nov 98	-61.5	-24.9	-22.7	-19.9	-11.3	19.0	24.1	33.8	64.1
Dec 98	-77.6	-53.0	-43.1	-39.6	-34.2	-23.0	-15.8	-10.6	4.87
Jan 99	-67.3	-65.0	-53.1	-49.0	-36.4	-29.6	-21.2	-3.87	-1.93
Feb 99	-40.9	-38.9	-31.1	-28.1	-20.8	-9.11	-5.00	12.5	33.3
Mar 99	-28.9	-18.5	-8.15	-4.34	11.7	27.2	41.5	52.9	66.1
Apr 99	68.9	71.3	83.8	95.3	105	122	135	155	161
May 99	97.8	105	120	133	172	195	212	233	243
Jun 99	103	108	117	129	143	165	175	197	201
Jul 99	71.6	83.0	95.6	107	118	139	148	175	201
Aug 99	41.0	46.6	53.4	66.3	93.4	113	123	135	147
Sep 99	-12.7	0.41	33.8	50.2	64.9	93.6	107	113	156

 Table 7-14. September 15, 1998, Lake Superior probabilistic outlook of net

 basin supply (mm) maximizing use of meteorology outlooks.

solutions in both cases represent minimums. (All computations are with probabilities, both reference quantiles and forecasts, significant to three digits after the decimal point.)

The first procedural algorithm matches Equations (4-9a) through (4-9g) while using all of the meteorology time series segments from 1948 to 1994; see Table 7-15. The second procedural algorithm matches Equations (4-9a) through (4-9h) but has zero weights for years 1952, 1953, 1957, 1963, 1972, 1974, 1982, and 1994; see Table 7-16. (It is interesting to note that none of the years omitted are La Niña years in Table 4-2.)

Table 7-15. Outlook weights using all meteorology time series segments for Equation (7-16) for the La Niña Lake Superior supply outlook example

Index, i	Weight, w <sub>i</sub>	Index, i	Weight, w <sub>i</sub>	Index, i	Weight, w <sub>i</sub>
(1)	(2)	(3)	(4)	(5)	(6)
1	0.259214	17	2.167816	33	0.565279
2	1.780569	18	0.259214	34	1.314670
3	0.945037	19	1.468592	35	0.259214
4	1.264503	20	0.644569	36	0.945037
5	0.259214	21	0.964035	37	0.091891
6	0.259214	22	1.081345	38	0.945037
7	0.259214	23	1.780569	39	0.964035
8	0.565279	24	2.013894	40	0.259214
9	1.729262	25	0.245813	41	0.093670
10	0.259214	26	2.718715	42	0.945037
11	2.167816	27	0.259214	43	1.116178
12	0.413135	28	0.796713	44	0.413135
13	0.796713	29	1.883183	45	0.964035
14	1.030038	30	2.019491	46	2.167816
15	2.013894	31	2.167816	47	0.259214
<u> </u>	0.259214	32	0.964035		

example.						
Index, i	Weight, w <sub>i</sub>	Index, i	Weight, $w_i$	Index, i	Weight, w <sub>i</sub>	
(1)	(2)	(3)	(4)	(5)	(6)	
1	0.599708	17	2.209499	33	0.595273	
2	1.798767	18	0.599708	34	1.244683	
3	1.287485	19	1.495376	35	0	
4	1.309397	20	0.651246	36	1.287485	
5	0	21	1.340538	37	0.322668	
6	0	22	1.084643	38	0.620105	
7	0.599708	23	1.798767	39	0.673159	
8	0.595273	24	1.958807	40	0.599708	
9	1.937663	25	0	41	0.161108	
10	0	26	2.699638	42	0.620105	
11	2.209499	27	0	43	1.085412	
12	0.850400	28	1.063499	44	0.183021	
13	0.396120	29	1.520976	45	0.673159	
14	1.223539	30	1.985514	46	2.209499	
15	1.958807	31	2.209499	47	0	
16	0	32	1.340538	<u> </u>	l	

Table 7-16. Outlook weights maximizing use of meteorology event probabilities for Equation (7-16) for the La Niña Lake Superior supply outlook example

Using either set of weights allows probabilistic hydrology outlooks for Lake Superior net basin supply to be built from Table 7-12, in the way they were built for Table 7-14 in the previous example. See Exercise A2-7 in Appendix 2 for this example, which also illustrates the setting of arbitrary user-defined (non-agency) probabilities such as Equations (4-9).

## **ORDERING PRIORITIES**

There are several practical ways for ordering priorities. First, a practitioner would use meteorology probability forecasts of appropriate lead and length for the derivative forecasts at hand. Thus, one would place meteorology forecasts over the next few days at higher priority than a 1-month forecast if one desired the derivative hydrology forecast at the end of the week. Likewise, if a lake level outlook over the next 6 months was to be made, then the 3-month meteorology forecasts beginning with the present month, the following month, and the month after would be more important than the second-week meteorology forecast. Another consideration for the practitioner is to place the most important variables first, reflecting his or her goals or purposes. For example, February air temperatures may be much more important for snowmelt events than June-July-August precipitation. Users may also assign priorities according to their confidence in the meteorology outlooks. For example, an older meteorology forecast may have a better forecast success rate in the user's application area than another agency; this too can be reflected in the user's priority listing.

Very often, priorities do not change much in day-to-day forecasting activities. A lot of thought may go into selecting a reasonable set of priorities for the agency forecasts that are used to make a derivative hydrology outlook. As long as the meteorological outlooks that are being used are not removed (even though they are allowed to change) from day to day, their priorities may remain unchanged. It is necessary only to recalculate the weights when the meteorology forecasts or their priorities change. The same set of weights can be used in day-to-day updated hydrology forecasts in the interim, reflecting only updated initial conditions used in the model simulations.

### ADDITIONAL METHODOLOGY CONSIDERATIONS

Formulating an optimization, as described in this chapter, allows for a general approach in determining weights in the face of multiple outlooks. However, this formulation also involves arbitrary choices, the largest of which is the selection of a relevant objective function. As mentioned earlier, other measures of relevance of the weights to a goal are possible and could require reformulating the solution methodology. An earlier approach, not described in this book, was to minimize the sum of squared differences between the relative frequencies associated with the bivariate distribution of precipitation and temperature before and after application of the weights. The goal was to make the resulting joint distribution as similar as possible to that observed historically while making the marginal distributions match the climate outlooks. Unfortunately, that method was intractable for consideration of more than one climate outlook. Alternative formulations that use linear measures for comparing alternative solutions to determine which is "best" are described in Chapter 10.

Most significantly, the method allows joint consideration of multiple probabilistic meteorology outlooks of event probabilities. The next chapter extends the methodology to also include probabilistic meteorology outlooks of most-probable events.

This page intentionally left blank

# Chapter 8

# MIXING MOST-PROBABLE METEOROLOGY OUTLOOKS

Chapters 6 and 7 discussed restructuring of the operational hydrology future scenarios sample to match forecast event probabilities as given, for example, in the National Oceanic and Atmospheric Administration's (NOAA's) monthly *Climate Outlooks* or its 8–14 day outlook. However, that chapter did not address matching most-probable event forecasts such as the NOAA 6–10 day outlook or the Environment Canada (EC) 1month and 3-month outlooks. This chapter extends the approach to mix all probabilistic meteorology outlooks to generate hydrology outlooks.

#### MATCHING MOST-PROBABLE EVENTS

Consider matching most-probable event forecasts such as are available in NOAA's 6–10 day outlooks for average air temperature and total precipitation, EC's monthly outlooks for average air temperature, or EC's seasonal and extended seasonal outlooks for average air temperature and total precipitation. Most-probable event forecasts are a special case of a more general category of probability statements. Generally, r + 1 intervals for a variable's values are set by defining interval limits,  $z_1 < z_2 \cdots < z_r$ . The general form of the probability statement in which a most-probable event forecast can be cast is that the *j*th event (interval) has a probability in excess of a specified value; the probability can be written in terms of the relative frequencies to be matched:

$$\hat{P}\left[z_{j-1} < X \leq z_j\right] > \phi_j \tag{8-1}$$

where X may be average air temperature or total precipitation and  $\phi_j$  is a probability limit.  $z_0 = -\infty$  and  $z_{r+1} = +\infty$  are understood and for these cases, Equation (8-1) becomes

$$\hat{P}[z_0 < X \le z_1] = \hat{P}[X \le z_1] > \phi_1$$
(8-2a)

$$\hat{P}[z_r < X \le z_{r+1}] = \hat{P}[X > z_r] > \phi_{r+1}$$
(8-2b)

[In the NOAA forecast of most-probable air temperature and precipitation events and in the EC forecasts of most-probable precipitation events,  $z_k$  is defined as the  $\gamma_k$  quantile  $(\xi_k)$  estimated from the 1961–90 period. In the EC forecasts of most-probable air temperature events, the quantiles are estimated from the 1963–93 period. In general:

$$\hat{P}\left[X \leq \hat{\xi}_k\right] = \gamma_k \qquad 1 \leq k \leq r \tag{8-3}$$

where  $\gamma_1 < \gamma_2 < \cdots < \gamma_r$  and  $\phi_k$  is defined in terms of the quantile probabilities:

$$\phi_k = \gamma_k - \gamma_{k-1} \qquad 1 \le k \le r+1 \tag{8-4}$$

where  $\gamma_0 = 0$  and  $\gamma_{r+1} = 1$ . For the NOAA 6–10 day most-probable event temperature

forecast, r = 4,  $\gamma_1 = 0.1$ ,  $\gamma_2 = 0.3$ ,  $\gamma_3 = 0.7$ , and  $\gamma_4 = 0.9$  ( $\phi_1 = 0.1$ ,  $\phi_2 = 0.2$ ,  $\phi_3 = 0.4$ ,  $\phi_4 = 0.2$ , and  $\phi_5 = 0.1$ ); for the NOAA 6–10 day most-probable event precipitation forecast and all of the EC most-probable event temperature and precipitation forecasts, r = 2,  $\gamma_1 = 1/3$ , and  $\gamma_2 = 2/3$  ( $\phi_1 = \phi_2 = \phi_3 = 1/3$ ). However, the more general definitions of  $z_k$  and  $\phi_k$  are used in this chapter to allow for other outlooks that may be more broadly defined than either of the present NOAA or EC most-probable event forecasts.]

Many most-probable event forecasts are accompanied by the implicit assumption that *only* the most-probable interval has forecast probability exceeding its reference probability. Equation (8-1) would then become:

$$\hat{P}\left[z_{j-1} < X \le z_j\right] > \phi_j \tag{8-5a}$$

$$\hat{P}[z_{k-1} < X \le z_k] \le \phi_k$$
  $k = 1, ..., r+1; \quad k \ne j$  (8-5b)

Alternatively, Equations (8-5) can be written in terms of the complement for the first event as:

$$\hat{P}\left[\operatorname{not}\left(z_{j-1} < X \leq z_{j}\right)\right] < 1 - \phi_{j}$$
(8-6a)

$$\hat{P}[z_{k-1} < X \le z_k] \le \phi_k \qquad k = 1, \dots, r+1; \quad k \neq j$$
(8-6b)

If the user does not wish to make the assumption, then the r inequalities in Equations (8-6b) can be omitted.

Weights are determined by matching relative frequencies in the operational hydrology sample to the most-probable interval forecast of Equations (8-6) [as was done in Chapter 6 to replace Equation (6-1) with Equation (6-2) by using Equation (5-8)]:

$$\frac{1}{n} \sum_{i \mid \text{not}(z_{j-1} < x_i \le z_j)} w_i < 1 - \phi_j$$
(8-7a)

$$\frac{1}{n} \sum_{i \mid z_{k-1} < x_i \le z_k} w_i \le \phi_k \qquad k = 1, \dots, r+1; \quad k \neq j$$
(8-7b)

Alternatively, write Equations (8-7) as follows:

$$\sum_{i=1}^{n} \alpha_{j,i} w_i < e_j \tag{8-8a}$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i \leq e_k \qquad k = 1, \dots, r+1; \quad k \neq j$$
 (8-8b)

where the  $\alpha_{k,i}$  are defined, as they were for Equation (6-10), as 1 or 0, corresponding to the inclusion or exclusion, respectively, of each variable in the respective appropriate sets of Equations (8-7), and where  $e_k$  corresponds to the probability limits specified in the most-probable event forecast  $[e_j = n(1 - \phi_j)]$  and  $e_k = n\phi_k$ ,  $k \neq j$ ]. The r + 1inequalities in Equations (8-8) represent one most-probable event forecast; if there are multiple most-probable event forecasts (from different agencies, for different periods and lags, and for different variables), represent them by the p + q inequalities:

$$\sum_{i=1}^{n} \alpha_{k,i} w_i < e_k \qquad k = 1, ..., p$$
 (8-9a)

$$\sum_{i=1}^{n} \alpha_{k,i} w_i \leq e_k \qquad k = p+1, \dots, p+q \qquad (8-9b)$$

where p = the total number of strictly-less-than constraints and q = the total number of less-than-or-equal-to constraints to be considered. Note that while Equations (8-9) may refer to different variables over different periods with different lengths and lag times, the equations are written in terms of a single set of weights ( $w_i$ , i = 1, ..., n) as was done for Equation (6-10).

## MIXING PROBABILISTIC METEOROLOGY OUTLOOKS

By adding the constraints corresponding to most-probable event forecasts in Equations (8-9) to those of the event probability forecasts in Equation (6-10) and the requirement of Equation (6-3), the following set of equations is formed to be to solved simultaneously:

$$\sum_{i=1}^{n} \alpha_{k,i} w_i = e_k \qquad k = 1, \dots, m \qquad (8-10a)$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i < e_k \qquad k = m+1, \dots, m+p \qquad (8-10b)$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i \leq e_k \qquad k = m+p+1, \dots, m+p+q \qquad (8-10c)$$

Again defining an optimization problem and solving by searching for an "optimum" solution, as in Equations (7-5), the optimization becomes:

$$\min \sum_{i=1}^{n} (w_i - 1)^2 \text{ subject to}$$
(8-11a)

$$\sum_{i=1}^{n} \alpha_{k,i} w_i = e_k \qquad k = 1, \dots, m \qquad (8-11b)$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i < e_k \qquad k = m+1, \dots, m+p \qquad (8-11c)$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i \leq e_k \qquad k = m+p+1, \dots, m+p+q \qquad (8-11d)$$

The solution to Equations (8-11) may give positive, zero, or negative weights, but only nonnegative weights make physical sense. Again, two procedural algorithms are used for finding nonnegative weights without adding additional constraints to Equations (8-11), so that the solution is analytically tractable. These algorithms repeatedly eliminate the lowest-priority equation or inequality in Equations (8-11b), (8-11c), and (8-11d) until nonnegative weights are obtained. As before, the first algorithm guarantees that all scenarios in the operational hydrology sample are used and the second maximizes the number of equations or inequalities (meteorology outlooks) used.

Equations (8-11) are equivalent to:

$$\min \sum_{i=1}^{n} (w_i - 1)^2 \text{ subject to}$$
(8-12a)

$$\sum_{i=1}^{n} \alpha_{k,i} w_i = e_k \qquad k = 1, \dots, m \qquad (8-12b)$$

$$\sum_{i=1}^{n} \alpha_{k,i} w_i + w_{n+k-m} = e_k \qquad k = m+1, \dots, m+p+q \qquad (8-12c)$$

$$w_i > 0$$
  $i = n+1, ..., n+p$  (8-12d)

$$w_i \ge 0$$
  $i = n + p + 1, ..., n + p + q$  (8-12e)

where the  $w_i$ , (i = n + 1, ..., n + p + q) are "slack" variables added to change consideration of an inequality constraint to consideration of an equality constraint in the optimization. This, in turn, is equivalent to:

$$\min \sum_{i=1}^{n} (w_i - 1)^2 \qquad \text{subject to} \qquad (8-13a)$$

$$\sum_{i=1}^{n+p+q} \alpha_{k,i} w_i = e_k \qquad k = 1, \dots, \ m+p+q \qquad (8-13b)$$

$$w_i > 0$$
  $i = n+1, ..., n+p$  (8-13c)

$$\geq 0$$
  $i = n + p + 1, ..., n + p + q$  (8-13d)

where the additional coefficients are defined as follows:

W;

$$\alpha_{k,i} = 0$$
  $k = 1, ..., m$   $i = n+1, ..., n+p+q$  (8-14a)

$$\alpha_{k,i} = 1$$
  $k = m+1, ..., m+p+q$   $i = n+k-m$  (8-14b)

$$\alpha_{k,i} = 0$$
  $k = m+1, ..., m+p+q$   $i > n, i \neq n+k-m$  (8-14c)

If the non-negativity constraints  $(w_i > 0, i = n+1, ..., n+p \text{ and } w_i \ge 0, i = n+p+1, ..., n+p+q)$  are ignored for now, Equations (8-13) become:

$$\min \sum_{i=1}^{n} (w_i - 1)^2 \text{ subject to}$$
(8-15a)

$$\sum_{i=1}^{n+p+q} \alpha_{k,i} w_i = e_k \qquad k = 1, \dots, \ m+p+q$$
(8-15b)

which is similar to Equations (7-5) and may be solved as before (Croley 1996, 1997a) by defining the Lagrangian (Hillier and Lieberman 1969, pp. 603-08),

$$L = \sum_{i=1}^{n} (w_i - 1)^2 - \sum_{k=1}^{m+p+q} \lambda_k \left( \sum_{i=1}^{n+p+q} \alpha_{k,i} w_i - e_k \right)$$
(8-16)

where  $\lambda_k$  = the unit penalty of violating the *k*th constraint in the optimization, and by setting the first derivatives with respect to each variable to zero:

$$\frac{\partial L}{\partial w_i} = 2(w_i - 1) - \sum_{k=1}^{m+p+q} \lambda_k \alpha_{k,i} = 0 \qquad i = 1, ..., n$$
 (8-17a)

$$\frac{\partial L}{\partial w_i} = -\sum_{k=1}^{m+p+q} \lambda_k \alpha_{k,i} \qquad = 0 \qquad i = n+1, \dots, n+p+q \quad (8-17b)$$

$$\frac{\partial L}{\partial \lambda_k} = -\sum_{i=1}^{n+p+q} \alpha_{k,i} w_i + e_k = 0 \quad k = 1, \dots, m+p+q \quad (8-17c)$$

This is a set of necessary but not sufficient conditions for the minimization of Equation (8-16) or the problem of Equations (8-15). The solution represents a "critical" point and must be checked further to identify it as either a minimum or a maximum. Equations (8-17) are linear and solvable via the Gauss-Jordan method of elimination because there are m + n + 2p + 2q equations in m + n + 2p + 2q unknowns (same number of equations and variables). For this problem where one of the equations in Equations (6-10) and (8-11) is Equation (6-3), the solution of Equations (8-17) represents the minimum if  $\sum w_i^2 < 2n$  and the maximum if  $\sum w_i^2 > 2n$  (see Appendix 3). Note that these are the same sufficiency conditions as for Equations (7-7). Equations (8-17) can be written in vector form as in Figure 8-1.

The solution of Equations (8-15) may give positive, zero, or negative weights and slack variables, but only nonnegative or strictly positive weights (either  $w_i \ge 0$  or  $w_i \ge 0$ , i = 1, ..., n) and slack variables ( $w_i \ge 0$ , i = n + 1, ..., n + p and  $w_i \ge 0$ , i = n + p + 1, ..., n + p + q) make physical sense, and the optimization must be further constrained. Two cases arise here:

$$w_i > 0$$
  $i = 1, ..., n$  (8-18a)

$$w_i > 0$$
  $i = n+1, ..., n+p$  (8-18b)

$$w_i \ge 0$$
  $i = n + p + 1, ..., n + p + q$  (8-18c)

and

$$w_i \ge 0$$
  $i = 1, ..., n$  (8-19a)

$$w_i > 0$$
  $i = n+1, ..., n+p$  (8-19b)

$$w_i \ge 0$$
  $i = n + p + 1, ..., n + p + q$  (8-19c)

In both cases, there is a mixture of strictly positive  $(w_i > 0)$  and simply nonnegative  $(w_i \ge 0)$  weights and slack variables for the optimization. These additional constraints can result in infeasibility (meaning there is no solution), and equations must be eliminated from Equations (8-15) to allow a feasible solution. To facilitate this, the engineer or hydrologist must prioritize the probabilistic meteorology outlook equations [and, hence, the equations in Equations (8-15)] so that the least important ones (lowest priority) can be eliminated first. The equation in Equations (8-15b) corresponding to Equation (6-3) should always be given top priority.

A procedural algorithm of successive optimizations is depicted in Figure 8-2; it preserves as many of the probability equations as possible while yielding results identical to Figure 7-2 when no slack variables are present (Croley 1997b). In Figure 8-2, if simple nonnegativity conditions would be violated in an optimization, even though other positivity conditions may also be violated, the procedural algorithm adds a zero constraint  $(w_i = 0)$  for each negative variable  $(w_i < 0)$ , as long as the resulting equation set still represents a nonempty space, and it solves the optimization again. If the resulting equation set would represent an empty solution space, then the algorithm eliminates all ear-