〈論 文〉

수문자료 확충을 위한 다중상관계수의 한계최소치 유도
Derivation of the Critical Minimum Values of the Multiple Correlation Coefficient for Augmenting Hydrologic Samples

허 준 행*
HEO, Jun-Haeng

Abstract : The augmenting hydrologic data using a correlation procedue has been used to improve the estimates of the mean and variance at the site of interest with short record when one or more nearby sites with longer records are available. The variance of the unbiased maximum likelihood estimator of σ_y^2 derived by Moran based on the multivariate normal distribution is modified into the form of Matalas and Jacobs for the bivariate normal distribution to get the critical minimum values of the multiple correlation coefficient which give the improvement for estimating the variance at the site of interest. Those values are tabulated for various lengths of records and the number of sites.

요 지 : 주변 관측지점의 자료가 유용한 경우 관측자료가 짧은 지점의 평균과 분산 추정치를 개선하기 위하여 상관계수를 이용한 수문자료 확충을 이용하여왔다. 본 연구에서는 관측지점의 분산 추정치를 개선하기 위한 다중상관계수의 한계최소치를 얻기 위하여, 다변량 정규분포에 근거하여 Moran이 유도한 확충자료 분산(σ_y^2)의 불편최우도추정량의 분산식을 Matalas와 Jacobs가 2변량 정규분포에 근거하여 유도한 식의 형태로 변형하였으며, 다양한 자료수와 지점수에 따라 다중상관계수의 한계최소치를 도표화 했다.

1. Introduction

If hydrologic data is short at the site of interest, the estimates of population parameters may be subject to large sampling errors. In such a case, augmenting hydrologic data using correlation analysis has been used to improve the estimates of parameters when longer records are available at nearby sites. For instance, estimates of the mean and variance of short hydrologic records may be improved based on longer records at other sites by using bivariate or multivariate normal distribution models. By using correlation analysis, Rosenblatt (1959) gave the expression of mean square error of the estimator of the variance for the bivariate normal population and Fiering (1963) considered the case of three sites based on the trivariate normal distribution. Matalas and Jacobs (1964) added a random component (noise) into the regression model in order to obtain an unbiased estimator of the population variance σ_y^2. This noise term did not affect the reliability of the estimate of the mean, but led to an unbiased estimate of the variance for the extended data at the site of interest. They gave the variance of the unbiased estimator of σ_y^2 and the critical minimum values of the correlation coefficient, R (for the bivariate case). For improving the estimate of the mean and variance based on data of the longer site, more re-

* 정회원, 연세대학교 공과대학 토목공학과 조교수

cently, Vogel and Stedinger (1985) proposed the improved estimators for the mean and variance at the site with short records based on longer records at another site by using optimal weights for both the mean and the variance, which minimize the mean square errors of estimators (of the mean and variance) which are linear combinations of two estimators of the mean (or of the variance). They showed that their estimators are specially worth for small samples.

The multivariate case for augmenting hydrologic data has been developed by Gilroy (1970) and Moran (1974), when there are in general m sites with longer records available. Gilroy (1970) expanded the model from the bivariate to the multivariate case based on the same formulation as Matalas and Jacobs' bivariate model (1964). However, Gilroy's expression of the variance of the unbiased estimator of σ_y^2 was not correct as noted by Moran (1974). The latter author indicated some anomalies of Gilroy's results and derived the unbiased maximum likelihood estimator of σ_y^2. The purpose of this paper is to modify the variance of the unbiased maximum likelihood estimator of σ_y^2 given by Moran (1974) into the form of Matalas and Jacobs based on the multivariate normal distribution so that the critical minimum values of R (the multiple correlation coefficient) to improve the estimates of the mean and variance can be more easily derived.

2. Statistical Model

The multiple linear regression to improve the estimates of the mean and variance at a site with a short record is applied when there are m additional sites with longer records available. Consider a site with a short record and m other sites with longer records. Assume that the variable y represents the short record of size N1 at a given site and the variable x(i) represents a site i with longer record of size N1+N2. The observed hydrologic data are displayed by

$$\begin{array}{l}
y_1, \quad y_2, \quad \cdots\cdots, y_{N1} \\
x_1(1), x_2(1), \cdots\cdots, x_{N1}(1), x_{N1+1}(1), \cdots\cdots, \\
\qquad x_{N1+N2}(1) \\
x_1(2), x_2(2), \cdots\cdots, x_{N1}(2), x_{N1+1}(2), \cdots\cdots, \\
\qquad x_{N1+N2}(2) \\
\qquad \vdots \qquad\qquad\qquad \vdots \\
x_1(m), x_2(m), \cdots\cdots, x_{N1}(m), x_{N1+1}(m), \cdots\cdots, \\
\qquad x_{N1+N2}(m)
\end{array}$$

It is assumed that all series are independent in time and the concurrent series are drawn from a multivariate normal population with parameters μ_{y}, $\mu_{x(i)}$, σ_y^2, $\sigma_{x(i)}^2$ and R, where $\mu_{x(i)}$ and $\sigma_{x(i)}^2$ denote the population mean and variance of $x_t(i)$, respectively for $i=1, \cdots\cdots, m$; μ_y and σ_y^2 are the population mean and variance of y_t, respectively, and R is the population multiple correlation coefficient between series y_t and $x_t(i)$. The problem is to transfer information from m sites with record length N1+N2 to the site of interest with short record length N1 and to improve the estimates of the parameters, μ_y and σ_y^2. After replacing by the sample estimates, the regression model is defined as (Gilroy, 1970)

$$y_t = \hat{a} + \sum_{i=1}^{m} \hat{b}_i x_t(i) + \alpha\theta(1-\hat{R}^2)^{1/2}\,\hat{\sigma}_{yc}\varepsilon_t \qquad (1)$$

where ε_t is a normal random variable with zero mean and unit variance and

$$\hat{a} = y_c - \sum_{i=1}^{m} \hat{b}_i x_c(i) \qquad (2)$$

$$\hat{b}_i = \sum_{i=1}^{m} d(ij)c(j), \qquad i=1, \cdots\cdots, m \qquad (3)$$

where d(ij) is the inverse element of the following matrix elements

$$g(ij)=\sum_{t=1}^{N1}[x_t(i)-x_c(i)][x_t(j)-x_c(j)],$$

$$i, j=1, \cdots\cdots, m \qquad (4)$$

and

$$c(j)=\sum_{t=1}^{N1}[x_t(j)-x_c(j)][y_t-y_c],$$

$$j=1, \cdots\cdots, m \qquad (5)$$

$$x_c(i)=\frac{1}{N1}\sum_{t=1}^{N1}x_t(i),$$

$$i=1, \cdots\cdots, m \qquad (6)$$

$$\bar{y}_c=\frac{1}{N1}\sum_{t=1}^{N1}y_t, \qquad (7)$$

$$\hat{\sigma}_{yc}=\left[\frac{1}{N1-1}\sum_{t=1}^{N1}(y_t-y_c)\right]^{1/2} \qquad (8)$$

and the coefficient α given by

$$\alpha=\left[\frac{N2(N1-2m-2)(N1-1)}{(N2-1)(N1-m-2)(N1-m-1)}\right]^{1/2} \qquad (9)$$

is required to yield an unbiased estimator of $\sigma_y{}^2$. The coefficient θ is equal to 1 if the noise is added, and $\theta=0$ if not.

The variance of the unbiased maximum likelihood estimator of $\sigma_y{}^2$ was given by Moran (1974) as

$$\mathrm{Var}(\hat{\sigma}_y{}^2)=\frac{2R^2\sigma_y{}^4}{(N1+N2-1)}+\frac{2m\sigma_y{}^4(1-R^2)^2}{(N1+N2-1)^2}\left[1+\frac{2N2}{(N1-m-2)}+\frac{N2[(N1-2)(N2+m+1)-(m-1)(m+2)]}{(N1-m-1)(N1-m-2)(N1-m-4)}+\frac{N2(m-2)(N1+N2-m-2)}{(N1-m-1)(N1-m-2)^2(N1-m-4)}\right]$$
$$+\frac{4\sigma_y{}^4R^2(1-R^2)}{(N1+N2-1)^2}\left[N1+2N2-1+\frac{N2(N2+Nm+1)}{(N1-m-2)}\right]+\frac{2\sigma_y{}^4(1-R^2)^2}{(N1-m-1)}\left[1-\frac{m(N1+N2-m-2)}{(N1-m-2)(N1+N2-1)}\right]^2 \qquad (10)$$

For convenience, the above variance can be written in the form of Matalas and Jacobs (1964) as

$$\mathrm{Var}(\hat{\sigma}_y{}^2)=\frac{2\sigma_y{}^4}{(N1-1)}+\frac{N2\sigma_y{}^4}{(N1+N2-1)^2}(A_NR^4+B_NR^2+C_N) \qquad (11)$$

where

$$A_N=\frac{2}{N2}\{(N1+N2-1)+mC_1-2C_2+C_3\} \qquad (12a)$$

$$B_N=\frac{4}{N2}(-mC_1+C_2-C_3) \qquad (12b)$$

$$C_N=\frac{2}{N2}\left[mC_1+C_3-\frac{(N1+N2-1)^2}{(N1-1)}\right] \qquad (12c)$$

and

$$C_1=1+\frac{2N2}{(N1-m-2)}+\frac{N2[(N1-2)(N2+m+1)-(m-1)(m+2)]}{(N1-m-1)(N1-m-2)(N1-m-4)}+\frac{N2(N1-2)(N1+N2-m-2)}{(N1-m-1)(N1-m-2)^2(N1-m-4)} \qquad (13a)$$

$$C_2=N1+2N2-1+\frac{N2(N2+m+1)}{(N1-m-2)} \qquad (13b)$$

$$C_3=\frac{[(N1-m-2)(N1+N2-1)-m(N1+N2-m-2)]^2}{(N1-m-1)(N1-m-2)^2} \qquad (13c)$$

Thus, the variance of $\hat{\sigma}_y{}^2$ has a quadratic function of R^2 and the term $\cdot\ 2\hat{\sigma}_y{}^4/(N1-1)$ represents the variance of $\hat{\sigma}_{yc}^2$. If the second term on the right hand side of Eq. (11) is negative, then $Var(\hat{\sigma}_y^2) < Var(\hat{\sigma}_{yc}^2)$, which means that $\hat{\sigma}_y^2$ is a better estimator of $\hat{\sigma}_y^2$ than $\hat{\sigma}_{yc}^2$. This is satisfied if the following condition holds

$$|R| > \left[\frac{-B_N \pm \sqrt{B_N^2 - 4A_N C_N}}{2A_N} \right]^{1/2} \qquad (14)$$

where A_N, B_N and C_N are defined by Eq. (12) and the critical minimum value of R for improving the estimate of the variance, say Rv, is determined when R is equal to right hand side of Eq. (14).

In addition, the variance of the mean of the extended sequence, y is given by (Gilroy, 1970)

$$Var(y) = \frac{\sigma_y^2}{N1} \left[1 - \frac{N2[(N1-2)R^2 - m]}{(N1+N2)(N1-2-m)} \right] \qquad (15)$$

where $\sigma_y^2/N1$ is the variance of y_t, $t=1, \cdots\cdots, N1$ and

$$y = \frac{1}{N1+N2} \sum_{t=1}^{N1+N2} y_t \qquad (16)$$

Thus, there is an improvement for estimating the mean in Eq. (15) if

$$|R| > \left[\frac{m}{N1-2} \right]^{1/2} \qquad (17)$$

where N1 should be greater than m+2 and the critical minimum value of R for improving the estimate of the mean, say Rm, is determined when R is equal to right hand side of Eq. (17).

Tables 1 to 5 show the values of Rv for various sample sizes N1, N2 and the number of sites m. Some values of Rv given by Matalas and Jacobs (1964) were wrong for m=1, but those values were corrected in the IACWD Bulletin #17B(1982). The concurrent sample size N1 should be greater than m+4 (N1>m+4) to avoid indefinite values in Eq. (13). That is the reason why a blank column appears in Table 4 for N1=8 and m=4. Likewise, there is no solution for Eq. (14) if the term $\sqrt{B_N^2 - 4A_N C_N}$ is negative.

For instance, this occurs for N1=8 and m=5 as indicated with an asterisk in Table 5. As ex-

Table 1 The Critical Values of R for Improving the Estimate of the Variance for m=1

N1 / N2	6	8	10	12	14	16	18	20	25	30	35	40	45	50	55	60
3	.815	.712	.643	.592	.551	.518	.490	.467	.420	.385	.357	.335	.317	.301	.287	.275
6	.819	.714	.645	.593	.552	.519	.491	.467	.420	.385	.358	.335	.317	.301	.287	.275
8	.820	.715	.645	.594	.553	.519	.491	.467	.420	.385	.358	.335	.317	.301	.287	.275
10	.821	.716	.646	.594	.553	.520	.492	.468	.421	.385	.358	.335	.317	.301	.287	.275
12	.822	.716	.647	.594	.553	.520	.492	.468	.421	.385	.358	.335	.317	.301	.287	.275
14	.822	.717	.647	.595	.554	.520	.492	.468	.421	.386	.358	.335	.317	.301	.287	.275
16	.823	.717	.647	.595	.554	.520	.492	.468	.421	.386	.358	.335	.317	.301	.287	.275
18	.823	.717	.648	.595	.554	.520	.492	.468	.421	.386	.358	.336	.317	.301	.287	.275
20	.823	.718	.648	.595	.554	.521	.492	.468	.421	.386	.358	.336	.317	.301	.287	.275
25	.824	.718	.648	.596	.555	.521	.493	.469	.421	.386	.358	.336	.317	.301	.287	.275
30	.824	.719	.649	.596	.555	.521	.493	.469	.421	.386	.358	.336	.317	.301	.287	.275
35	.825	.719	.649	.596	.555	.521	.493	.469	.422	.386	.358	.336	.317	.301	.287	.275
40	.825	.719	.649	.597	.555	.521	.493	.469	.422	.386	.358	.336	.317	.301	.287	.275
45	.825	.719	.649	.597	.555	.522	.493	.469	.422	.386	.358	.336	.317	.301	.287	.275
50	.825	.720	.649	.597	.555	.522	.493	.469	.422	.386	.358	.336	.317	.301	.288	.275
55	.825	.720	.650	.597	.556	.522	.494	.469	.422	.386	.359	.336	.317	.301	.288	.276
60	.825	.720	.650	.597	.556	.522	.494	.469	.422	.386	.359	.336	.317	.301	.288	.276

pected, the values of Rv decrease as the record length N1 increases or the number of sites decreases as shown in Tables 1 to 5. The effect of N2 is not as much as the effects of N1 and m. The values of Rv increase only slightly as N2 increases. Table 6 displays the values of Rm and Rv for m=1, 2, 3, 4, 5 when N2 is equal to 60. Figures 1 and 2 show the relationships between Rm, Rv and m, respectively.

Table 2 The Critical Minimum Values of R for Improving the Estimate of the Variance for m=2

N1 N2	8	10	12	14	16	18	20	25	30	35	40	45	50	55	60
3	.824	.752	.697	.654	.617	.586	.560	.507	.466	.434	.408	.386	.367	.351	.337
6	.827	.755	.700	.655	.619	.587	.561	.507	.467	.434	.408	.386	.367	.351	.337
8	.829	.756	.701	.656	.619	.588	.561	.508	.467	.435	.408	.386	.367	.351	.337
10	.830	.757	.701	.657	.620	.589	.562	.508	.467	.435	.408	.386	.368	.351	.337
12	.831	.758	.702	.657	.620	.589	.562	.508	.467	.435	.409	.386	.368	.351	.337
14	.831	.758	.703	.658	.621	.589	.562	.508	.467	.435	.409	.387	.368	.351	.337
16	.832	.759	.703	.658	.621	.590	.563	.508	.468	.435	.409	.387	.368	.351	.337
18	.832	.759	.703	.659	.621	.590	.563	.509	.468	.435	.409	.387	.368	.351	.337
20	.833	.760	.704	.659	.622	.590	.563	.509	.468	.435	.409	.387	.368	.352	.337
25	.833	.760	.704	.659	.622	.591	.563	.509	.468	.436	.409	.387	.368	.352	.337
30	.834	.761	.705	.660	.623	.591	.564	.509	.468	.436	.409	.387	.368	.352	.337
35	.834	.761	.705	.660	.623	.591	.564	.510	.468	.436	.409	.387	.368	.352	.337
40	.835	.762	.706	.661	.623	.592	.564	.510	.469	.436	.409	.387	.368	.352	.337
45	.835	.762	.706	.661	.623	.592	.564	.510	.469	.436	.410	.387	.368	.352	.337
50	.835	.762	.706	.661	.624	.592	.565	.510	.469	.436	.410	.387	.368	.352	.338
55	.835	.762	.706	.661	.624	.592	.565	.510	.469	.436	.410	.387	.368	.352	.338
60	.835	.763	.707	.661	.624	.592	.565	.510	.469	.436	.410	.387	.368	.352	.338

Table 3 The Critical Minimum Values of R for Improving the Estimate of the Variance for m=3

N1 N2	8	10	12	14	16	18	20	25	30	35	40	45	50	55	60
3	.902	.828	.773	.728	.690	.658	.630	.573	.529	.494	.465	.441	.420	.402	.386
6	.906	.832	.776	.730	.692	.659	.631	.574	.530	.495	.466	.441	.420	.402	.386
8	.907	.833	.777	.731	.693	.660	.632	.574	.530	.495	.466	.441	.420	.402	.386
10	.908	.834	.778	.732	.694	.661	.632	.575	.530	.495	.466	.441	.421	.402	.386
12	.909	.835	.779	.733	.695	.662	.633	.575	.531	.495	.466	.442	.421	.402	.386
14	.910	.836	.780	.734	.695	.662	.633	.575	.531	.495	.466	.442	.421	.402	.386
16	.910	.837	.780	.734	.696	.662	.634	.576	.531	.496	.466	.442	.421	.403	.386
18	.911	.837	.781	.735	.696	.663	.634	.576	.531	.496	.467	.442	.421	.403	.387
20	.911	.838	.781	.735	.696	.663	.634	.576	.531	.496	.467	.442	.421	.403	.387
25	.912	.839	.782	.736	.697	.664	.635	.577	.532	.496	.467	.442	.421	.403	.387
30	.912	.839	.783	.737	.698	.664	.636	.577	.532	.496	.467	.442	.421	.403	.387
35	.913	.840	.783	.737	.698	.665	.636	.577	.532	.497	.467	.443	.421	.403	.387
40	.913	.840	.784	.738	.699	.665	.636	.578	.533	.497	.467	.443	.422	.403	.387
45	.913	.840	.784	.738	.699	.666	.637	.578	.533	.497	.468	.443	.422	.403	.387
50	.914	.841	.784	.738	.699	.666	.637	.578	.533	.497	.468	.443	.422	.403	.387
55	.914	.841	.785	.738	.699	.666	.637	.578	.533	.497	.468	.443	.422	.403	.387
60	.914	.841	.785	.739	.700	.666	.637	.578	.533	.498	.468	.443	.422	.404	.387

Table 4 The Critical Minimum Values of R for Improving the Estimate of the Variance for m=4

N1 N2	8	10	12	14	16	18	20	25	30	35	40	45	50	55	60
3		.886	.831	.786	.748	.715	.686	.627	.581	.544	.513	.487	.464	.445	.428
6		.890	.834	.789	.750	.717	.688	.628	.582	.544	.513	.487	.465	.445	.428
8		.892	.836	.790	.751	.718	.688	.628	.582	.545	.514	.487	.465	.445	.428
10		.893	.837	.791	.752	.719	.689	.629	.582	.545	.514	.488	.465	.445	.428
12		.894	.838	.792	.753	.719	.690	.630	.583	.545	.514	.488	.465	.445	.428
14		.895	.839	.793	.754	.720	.690	.630	.583	.545	.514	.488	.465	.446	.428
16		.895	.840	.794	.754	.721	.691	.630	.583	.546	.515	.488	.465	.446	.428
18		.896	.840	.794	.755	.721	.691	.631	.584	.546	.515	.488	.466	.446	.428
20		.896	.841	.795	.755	.721	.692	.631	.584	.546	.515	.488	.466	.446	.428
25		.897	.842	.796	.756	.722	.693	.632	.584	.547	.515	.489	.466	.446	.429
30		.898	.842	.796	.757	.723	.693	.632	.585	.547	.516	.489	.466	.446	.429
35		.898	.843	.797	.758	.724	.694	.633	.585	.547	.516	.489	.466	.446	.429
40		.899	.843	.797	.758	.724	.694	.633	.586	.548	.516	.489	.467	.447	.429
45		.899	.844	.798	.759	.724	.695	.633	.586	.548	.516	.490	.467	.447	.429
50		.899	.844	.798	.759	.725	.695	.634	.586	.548	.516	.490	.467	.447	.429
55		.899	.844	.798	.759	.725	.695	.634	.586	.548	.517	.490	.467	.447	.429
60		.899	.845	.799	.759	.725	.695	.634	.587	.548	.517	.490	.467	.447	.429

Table 5 The Critical Minimum Values of R for Improving the Estimate of the Variance for m=5

N1 N2	8	10	12	14	16	18	20	25	30	35	40	45	50	55	60
3	*	.933	.878	.833	.795	.762	.762	.672	.625	.586	.554	.527	.503	.482	.464
6	*	.937	.881	.836	.798	.764	.764	.673	.626	.587	.555	.527	.503	.482	.464
8	*	.938	.883	.838	.799	.765	.765	.674	.626	.587	.555	.527	.503	.483	.464
10	*	.940	.884	.839	.800	.766	.766	.675	.627	.588	.555	.528	.504	.483	.464
12	*	.940	.885	.840	.801	.767	.767	.675	.627	.588	.555	.528	.504	.483	.464
14	*	.941	.886	.841	.802	.768	.768	.676	.628	.588	.556	.528	.504	.483	.465
16	*	.942	.887	.841	.803	.768	.768	.676	.628	.589	.556	.528	.504	.483	.465
18	*	.942	.887	.842	.803	.769	.739	.677	.628	.589	.556	.528	.504	.483	.465
20	*	.942	.888	.843	.804	.770	.739	.677	.629	.589	.556	.529	.505	.484	.465
25	*	.943	.889	.844	.805	.771	.740	.678	.629	.590	.557	.529	.505	.484	.465
30	*	.944	.889	.844	.806	.771	.741	.679	.630	.590	.557	.529	.505	.484	.466
35	*	.944	.890	.845	.806	.772	.742	.679	.630	.591	.558	.530	.505	.484	.466
40	*	.944	.890	.846	.807	.773	.742	.680	.631	.591	.558	.530	.506	.484	.466
45	*	.945	.891	.846	.807	.773	.743	.680	.631	.591	.558	.530	.506	.485	.466
50	*	.945	.891	.846	.807	.773	.743	.680	.631	.591	.558	.530	.506	.485	.466
55	*	.945	.891	.847	.808	.774	.743	.681	.632	.592	.559	.530	.506	.485	.466
60	*	.945	.892	.847	.808	.774	.744	.681	.632	.592	.559	.531	.506	.485	.466

Table 6 The Values of Rv(N2=60) and Rm for M=1, 2, 3, 4, 5

	m=1		m=2		m=3		m=4		m=5	
N1	Rm	RV	Rm	RV	Rm	RV	Rm	RV	Rm	RV
8	.408	.720	.577	.835	.707	.914	.816		.913	*
10	.354	.650	.500	.763	.612	.841	.707	.899	.791	.945
12	.316	.597	.447	.707	.548	.785	.632	.845	.707	.892
14	.289	.556	.408	.661	.500	.739	.577	.799	.645	.847
16	.267	.522	.380	.624	.463	.700	.535	.759	.598	.808
18	.250	.494	.354	.592	.433	.666	.500	.725	.559	.774
20	.236	.469	.333	.565	.408	.637	.471	.695	.527	.744
25	.209	.422	.295	.510	.361	.578	.417	.634	.466	.681
30	.189	.386	.267	.469	.327	.533	.378	.587	.423	.632
35	.174	.359	.246	.436	.302	.498	.348	.548	.389	.592
40	.162	.336	.229	.410	.281	.468	.324	.517	.363	.559
45	.152	.317	.216	.387	.264	.443	.305	.490	.341	.531
50	.144	.301	.204	.368	.250	.422	.289	.467	.323	.506
55	.137	.288	.194	.352	.238	.404	.275	.447	.307	.485
60	.131	.276	.186	.338	.227	.387	.263	.429	.294	.466

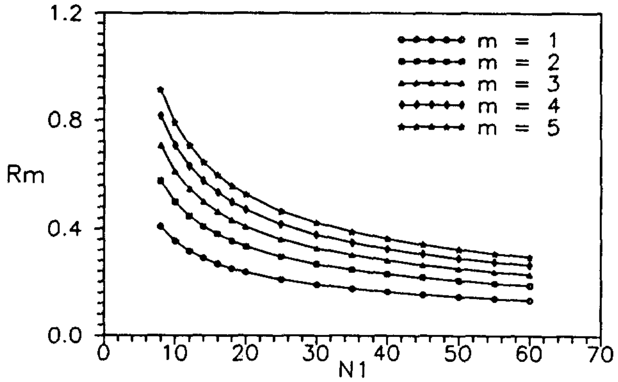

Fig. 1 The critical minimum values of R for improving the estimate of the mean (Rm) as a function of N1 for m=1, 2, 3, 4, 5

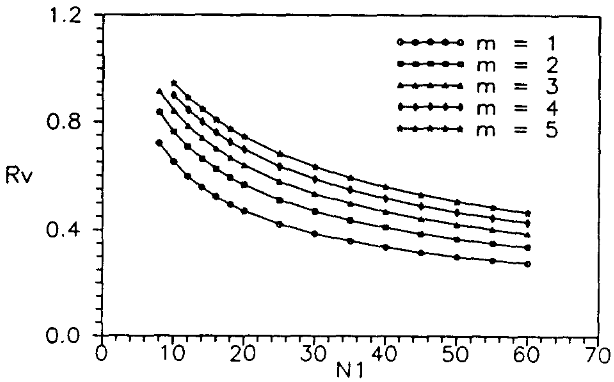

Fig. 2 The critical minimum values of R for improving the estimate of the mean (Rv) as a function of N1 for N2=60 and m=1, 2, 3, 4, 5

3. Summary and Conclusions

In this paper, the variance of the unbiased maximum likelihood estimator of σ_x^2 based on the multivariate normal distribution given by Moran (1974) was modified to the form of Matalas and Jacobs (1964). The critical minimum values of the multiple correlation coefficients for improving the estimates of the mean and variance at the site of interest with short records can be calculated easily from the modified equation as a function of the sample sizes (N1 and N2) and the number of neighboring sites m. The critical values to improve the estimates of the variance, Rv, were tabulated for various values of the sample sizes N1 and N2 for a given number of sites m=1, ……, 5. The Rv decreases as the short records N1 increases. On the other hand, the Rv is increased as the number of sites is increased. Similarly, the Rv increases as the records N2 increases. However, the effect of N2 on the Rv is very small with compare to N1 and m, especially when N1 is large.

Acknowledgement

The author would like to thank Dr. J. D. Salas, Professor and Program leader of Hydrologic Science and Engineering Program, Colorado State University for his motivation of this study.

References

1. Fiering, M. B. (1963), Use of Correlation to Improve Estimates of the Mean and Variance, U. S. Geological Survey Prof. Paper, 434-C.
2. Gilroy. E. J. (1970), Reliability of a Variance Estimate Obtained from a Sample Augmented by Multivariate Regression, Water Resour. Res., Vol. 6, No. 6, pp. 1595-1600.
3. Interagency Advisory Committee on Water Data. (1982), Guideline for Determining Flood Flow Frequency, Bulletin #17B, Hydrology Subcommittee, U.S. Geological Survey, Reston, Virginia.
4. Matalas, N. C., and B. Jacobs (1964), A Correlation Procedure for Augmenting Hydrologic Data, U. S. Geological Survey Prof. Paper, 434-E.
5. Moran. M. A. (1974), On Estimators Obtained from a Sample Augmented by Multiple Regression, Water Resour. Res., Vol. 10, No. 1, pp. 81-85.
6. Rosenblatt, J. R. (1959), On the Synthetic Record Problem, Natl. Bur. Standards Tech. Report 1.
7. Vogel, R. M., and J. R. Stedinger (1985), Minimum Variance Streamflow Record Augmentation Procedures, Water Resour.Res., Vol.21, No. 5, pp.715-723.