Click here to return to USGS publications

PROCEDURES FOR ADJUSTING REGIONAL REGRESSION MODELS OF URBAN-RUNOFF QUALITY USING LOCAL DATA

By Anne B. Hoos and Joy K. Sisolak

U.S. GEOLOGICAL SURVEY

Open-File Report 93-39

science for a changing world

Nashville, Tennessee
1993

U.S. DEPARTMENT OF THE INTERIOR
BRUCE BABBITT, Secretary

U.S. GEOLOGICAL SURVEY
Dallas L. Peck, Director

For additional information write to:

District Chief
U.S. Geological Survey
810 Broadway, Suite 500
Nashville, Tennessee 37203

Copies of this report can be purchased from:

U.S. Geological Survey
Books and Open-File Reports Section
Federal Center
Box 25425
Denver, Colorado 80225

CONTENTS

	Page
Abstract	1
Introduction	1
Purpose and scope	2
Regional regression models of urban-runoff quality	3
Local urban-runoff quality data	4
Procedures for adjusting regional regression models of urban-runoff quality using local data	6
Model-adjustment procedures	6
Single-factor regression against regional prediction	7
Regression against regional prediction	8
Regression against regional prediction and additional local variables	8
Weighted combination of regional prediction and local-regression prediction	9
Selection of appropriate adjustment procedures	11
Model-adjustment procedure testing	13
Test procedures	13
Test results	16
Bellevue	17
Denver	18
Knoxville	20
Sensitivity analysis	20
Estimating the accuracy of model-adjustment procedures	23
Example application	24
Prediction of annual or seasonal urban-runoff quality	26
Summary	26
References cited	27

SUPPLEMENTAL INFORMATION

A. Program (MINITAB) of exploratory data analysis procedures applied to calibration data set to guide selection of model-adjustment procedures . . . 31
B. Program (MINITAB) of statistical procedures applied to calibration data set to derive coefficients for model-adjustment procedures . . . 33
C. Formulas for standard error of prediction for model-adjustment procedures . . . 36
D. Program (MINITAB) applied to data from an unmonitored site to calculate the prediction using model-adjustment procedures . . . 38

FIGURES

1. Graph showing observed and predicted chemical oxygen demand load in storm runoff for City X's local data base . . . 5
2. Flowchart guiding selection of model-adjustment procedure (MAP) based on exploratory data analysis of the calibration data set . . . 12

TABLES

1. Standard errors of estimate for regional regression models of storm-runoff loads and mean concentrations of selected constituents . . . 3
2. Exploratory data analysis of the calibration data sets from the data base for Bellevue, Washington . . . 14

Page

3. Exploratory data analysis of the calibration data sets from the data base for Denver, Colorado 14
4. Exploratory data analysis of the calibration data sets from the data base for Knoxville, Tennessee 15
5. Root mean square errors and associated rankings for model-adjustment procedures and other estimators for verification data sets, compared with rankings for standard error of estimate for corresponding calibration data sets, from the data base for Bellevue, Washington 17
6. Root mean square errors and associated rankings for model-adjustment procedures and other estimators for verification data sets, compared with rankings for standard error of estimate for corresponding calibration data sets, from the data base for Denver, Colorado 19
7. Root mean square errors and associated rankings for model-adjustment procedures and other estimators for verification data sets, compared with rankings for standard error of estimate for corresponding calibration data sets, from the data base for Knoxville, Tennessee 21
8. Effect of size of calibration data sets for model-adjustment procedures on root mean square errors for verification data sets taken from the Bellevue, Washington, data base 22
9. Sample of calibration data set and values for standard errors of estimate, bias-correction factors, and coefficients for the model-adjustment procedures for City X 25

CONVERSION FACTORS

Multiply	By	To obtain
inch (in.)	25.40	millimeter
pound (lb)	0.4536	kilogram
square mile (mi^2)	2.589	square kilometer

Procedures for Adjusting Regional Regression Models of Urban-Runoff Quality Using Local Data

By Anne B. Hoos *and* Joy K. Sisolak

ABSTRACT

Statistical operations termed model-adjustment procedures (MAP's) can be used to incorporate local data into existing regression models to improve the prediction of urban-runoff quality. Each MAP is a form of regression analysis in which the local data base is used as a calibration data set. Regression coefficients are determined from the local data base, and the resulting 'adjusted' regression models can then be used to predict storm-runoff quality at unmonitored sites. The response variable in the regression analyses is the observed load or mean concentration of a constituent in storm runoff for a single storm. The set of explanatory variables used in the regression analyses is different for each MAP, but always includes the predicted value of load or mean concentration from a regional regression model. The four MAP's examined in this study were: single-factor regression against the regional model prediction, P_u (termed MAP-1F-P), regression against P_u (termed MAP-R-P), regression against P_u and additional local variables (termed MAP-R-P+nV), and a weighted combination of P_u and a local-regression prediction (termed MAP-W).

The procedures were tested by means of split-sample analysis, using data from three cities included in the Nationwide Urban Runoff Program: Denver, Colorado; Bellevue, Washington; and Knoxville, Tennessee. The MAP that provided the greatest predictive accuracy for the verification data set differed among the three test data bases and among model types (MAP-W for Denver and Knoxville, MAP-1F-P and MAP-R-P for Bellevue load models, and MAP-R-P+nV for Bellevue concentration models) and, in many cases, was not clearly indicated by the values of standard error of estimate for the calibration data set. A scheme to guide MAP selection, based on exploratory data analysis of the calibration data set, is presented and tested.

The MAP's were tested for sensitivity to the size of a calibration data set. As expected, predictive accuracy of all MAP's for the verification data set decreased as the calibration data-set size decreased, but predictive accuracy was not as sensitive for the MAP's as it was for the local regression models.

INTRODUCTION

Urban land use has been shown to be a major source of nonpoint-source pollution. Recognizing this, the amendments of 1987 to the Clean Water Act require that cities with populations of more than 100,000 provide estimates of storm-runoff loads from urban areas to receiving streams (U.S. Environmental

Protection Agency, 1990, p. 48070). City engineers have a variety of options for developing these estimates, ranging from simple empirical techniques (Young and others, 1979; U.S. Environmental Protection Agency, 1983; Schueler, 1987) to more advanced statistical regression (Driver and Tasker, 1990) and conceptually-based models (reviewed in Huber, 1986; Nix, 1991). The Driver-Tasker models are regression models of storm-runoff quality (constituent load and mean concentration) on physical, land-use, and climatic characteristics from the data base of the Nationwide Urban Runoff Program (NURP). Separate sets of regression models were developed for mean-annual runoff quality and for single-storm runoff quality.

Regardless of the method selected, provision should be made for adjustment of the '*a priori*' prediction using local urban-runoff quality data currently being collected in each city to meet additional regulatory requirements (U.S. Environmental Protection Agency, 1990, p. 48069-48070). The local storm-load data base for each city will consist in most cases of about three storms at 5-10 sites, or about 15-30 load observations.

A procedure to adjust the regional single-storm models (Driver and Tasker, 1990) for a particular city, using a small data base from that city, was presented in a recent study by Hoos (1991). Although such a model adjustment procedure (MAP) may seem to be a reasonable approach, at least intuitively, several unanswered questions come to the fore about the validity of this procedure and of possible alternative procedures. For example:

- What are the assumptions for the several proposed MAP's, and can these be codified for potential adjustors as they examine their local data bases? For example, is there a minimum size for a local data base to be used in the various MAP's, below which size the assumptions in the procedures are not valid?

- Of all statistically valid MAP's, which will provide the most reliable predictions for unmonitored sites?

- Do the models for constituent load differ from the models for constituent mean concentration with respect to their suitability for MAP's?

- How can the uncertainty of an adjusted-model prediction for an unmonitored site be estimated?

Purpose and Scope

The purpose of this investigation is to provide information regarding appropriate statistical methods for combining or weighting regional model predictions of storm-runoff quality with local data. This report describes:

- the assumptions for four proposed MAP's, and how these assumptions translate into requirements for the local data base;

- a scheme for selecting the appropriate adjustment procedure based on exploratory data analysis of the local data base;

- results from split-sample tests of the four proposed MAP's and the selection scheme; and

- expressions for calculating standard errors of prediction and confidence intervals for unmonitored sites using each of the proposed MAP's.

REGIONAL REGRESSION MODELS OF URBAN-RUNOFF QUALITY

Urban-runoff quality at unmonitored sites is commonly estimated using either deterministic models of washoff and transport processes in the watershed or statistical models calibrated with observed data at other sites. Although in the case of estimating at unmonitored sites, neither type of model can be calibrated with at-site data, the statistical-model approach has the advantage of providing a measure of the uncertainty in the model predictions. This advantage could be an important consideration for city engineers or planners responsible for developing remedial water-quality management programs or designing additional data-collection programs.

Regression models were developed by the U.S. Geological Survey (Driver and Tasker, 1990) from regression analysis of the NURP national data base (Mustard and others, 1987; U.S. Environmental Protection Agency, 1983). Separate sets of regression models were developed for mean-annual runoff quality and for single-storm runoff quality. The single-storm regression models relate storm-runoff quality (constituent load and mean concentration, the response variables) from a single storm to easily measured physical, land-use, and climatic characteristics (the explanatory variables). Models were developed for 11 constituents: chemical oxygen demand (COD), suspended solids (SS), dissolved solids (DS), total nitrogen (TN), total ammonia plus organic nitrogen as nitrogen (TKN), total phosphorus (TP), dissolved phosphorus (DP), total recoverable cadmium (CD), total recoverable copper (CU), total recoverable lead (PB), and total recoverable zinc (ZN). A set of three models corresponding to three regional divisions was developed for each constituent load (Driver and Tasker, 1990, tables 1 and 3) and for each constituent mean concentration (Driver and Tasker, 1990, table 5). The basis for the regional divisions was mean annual rainfall (region I, less than 20 inches; region II, 20-40 inches; region III, greater than 40 inches), which provided the best results of seven bases tested for regionalization/stratification (Driver and Tasker, 1990, p. 5). Standard errors of estimate (*SE*) were generally smallest for region I models and largest for region III models (table 1), indicating that as mean annual rainfall increases, the ability to estimate storm-runoff quality decreases.

Table 1. Standard errors of estimate for regional regression models of storm-runoff loads and mean concentrations of selected constituents

[Values for standard error of estimate (SE) from Driver and Tasker, 1990, tables 2, 3, and 6; COD, chemical oxygen demand; TKN, total kjeldahl nitrogen; PB, total recoverable lead; SS, suspended solids; Lsa, stepwise-analysis regression model for storm-runoff load; Csa, stepwise-analysis regression model for storm-runoff mean concentration; L3, 3-variable regression model for storm-runoff load]

	Standard error of estimate					
	Region I		Region II		Region III	
Model	Percent	Log	Percent	Log	Percent	Log
COD.Lsa	86	0.324	97	0.355	169	0.505
COD.Csa	61	.245	79	.303	78	.300
COD.L3	116	.403	106	.376	186	.531
TKN.Lsa	71	.277	106	.377	165	.498
TKN.Csa	60	.242	85	.321	85	.321
TKN.L3	129	.431	107	.381	184	.529
PB.Lsa	141	.455	131	.435	227	.586
PB.Csa	88	.331	103	.371	179	.414
PB.L3	166	.500	135	.442	228	.586
SS.Lsa	230	.589	165	.498	265	.627
SS.Csa	131	.434	128	.427	178	.519
SS.L3	251	.613	173	.512	290	.651

Two separate sets of models of storm-runoff load were developed for each constituent and for each region. One set, referred to as the stepwise-analysis regression models, was developed from a stepwise regression analysis of 13 candidate explanatory variables; the number of explanatory variables selected as significant for a particular model ranged from three to six (Driver and Tasker, 1990, table 1). The second set included only the three most significant explanatory variables: total storm rainfall, total contributing drainage area, and impervious area (Driver and Tasker, 1990, table 3). For the purpose of this report, the stepwise-analysis load and concentration models will be referred to as Lsa and Csa, respectively, and the 3-variable load models as L3. The Lsa models fit the observed data better than L3 models (table 1). *SE* measures fit of observed data rather than predictive accuracy. The fit of the load and concentration models should not be compared on the basis of *SE*, because the response variable units in each case were different.

A final set of national regression models was developed to predict load from an average storm (response variable) based upon five explanatory variables (Driver and Tasker, 1990, table 10). Estimates from these models can be used in conjunction with an estimate of the average number of storms per year to yield an estimate of mean annual load.

LOCAL URBAN-RUNOFF QUALITY DATA

Faced with the need to develop estimates of storm-runoff quality for a large number of unmonitored sites, a city engineer might wish to employ the published regression models, provided the published standard errors of estimate are deemed acceptable (table 1). A separate option would be to test the published models by comparing regional single-storm model (henceforth termed regional model) estimates with available local urban-runoff quality data to appraise the predictive accuracy of the regional models for the particular city of interest. The magnitude of the model errors could indicate the relative accuracy and usefulness of these models for estimating loads and mean concentrations of constituents for watersheds in that city.

When regional-model results prove inaccurate for estimating storm-runoff quality in a particular city, the city engineer might wish to use local data to 'adjust' (through a partial recalibration procedure) the regional models and obtain more accurate results. Local data bases used for the adjustment of regional models should possess certain attributes if the adjustments are to result in more accurate estimates. Among these attributes are:

- *The monitoring sites in the local data base should represent a wide range of conditions of physical characteristics (size of drainage area, percent impervious area) and land-use characteristics.* This will ensure that the values for these explanatory variables at any unmonitored site for which an estimate is desired will fall within the range represented by the local data base. It may be useful to compare the range represented by the local data base with the range represented by the regional NURP data base (Driver and Tasker, 1990, table 4).

- *The monitored storms in the local data base should represent a wide range of storm characteristics (total storm rainfall, duration of each storm, and antecedent conditions),* for the same reason cited previously. Although explanatory variables related to antecedent conditions (for example, preceding number of dry days, amount of rainfall during the preceding day, 3 days, or 7 days) are not included in the regional models, such variables could account for some of the unexplained error in these models and, therefore, may be candidates for use in adjusting the models.

The following discussion illustrates the use of a local data base (for a hypothetical City X, located in region II) to test the validity of the regional models for a particular city. Data for storm-runoff load of COD have been collected during three storms at each of five sites in City X, with a resulting data base of 15 observations. For each of these observed loads, a corresponding predicted load can be computed by

evaluating the explanatory variables and applying the regional model for COD for region II. The observed and corresponding predicted values are shown in figure 1 for each of the 15 events. Examination of the pattern of correspondence (or lack of correspondence) between observed and corresponding predicted values, and knowledge of the local data base and NURP data base for region II, can lead to one of the following conclusions.

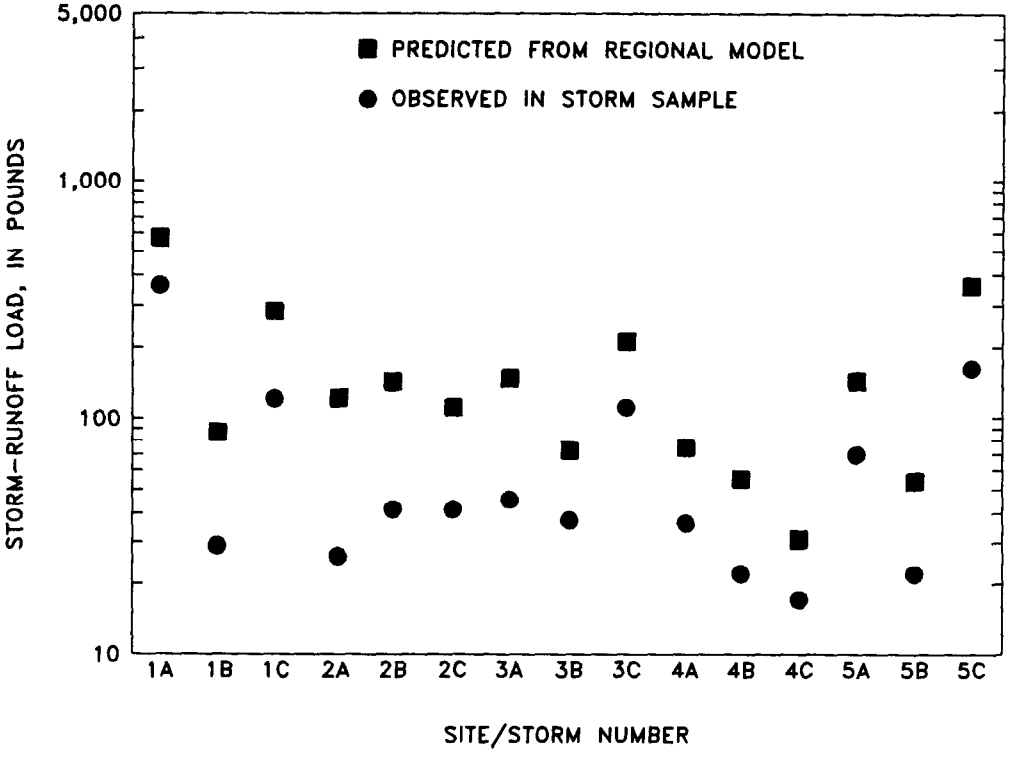


Figure 1. Observed and predicted chemical oxygen demand load in storm runoff for City X's local data base.

One possible conclusion is that the site or storm characteristics (explanatory variables) represented by the local data base are not representative of the full range of storm-runoff conditions in City X, whereas the characteristics of the calibration data set for the region II models are representative. Consequently, the regional-model predictions, although appearing inaccurate for estimating the local data, might be more accurate estimates for a typical unmonitored site and typical storm in City X. Explanations for drawing such a conclusion might include: knowledge that sites in the local data base might be influenced by point-source discharges, or knowledge that storms monitored for the local data base are atypical of average storm characteristics for City X.

A second possible conclusion is that the regional model predictions are biased relative to actual storm-runoff conditions in City X, and that the observations in the local data base are representative of local conditions. Conditions supporting this conclusion might include: (1) the values of the explanatory variables for watersheds in City X are consistently outside the range of values for explanatory variables in the NURP region II data base (for example, mean annual rainfall in City X is higher than for any city included in the

region II data base); and (2) a physical, land-use, or climatic variable not tested or included, but responsible for some of the unexplained error in the regional model, might have a different range of values for City X relative to cities included in the NURP region II data base (for example, drainage-structure construction materials or practices used in City X might be different from materials or practices used in any NURP region II city; the range of antecedent conditions occurring in City X might be different from the range of antecedent conditions occurring in any NURP region II city). If either of these two conditions can be shown to exist in City X, then it might be valid to adjust the regional models for application to unmonitored sites.

PROCEDURES FOR ADJUSTING REGIONAL REGRESSION MODELS OF URBAN-RUNOFF QUALITY USING LOCAL DATA

Before any particular adjustment procedure for a constituent model is considered, it is helpful to examine the pattern of correspondence between the observed and predicted values from the local data base. The pattern illustrated in figure 1 has the following characteristics, both of which tend to indicate that model adjustment is a valid approach:

(1) the direction of bias of predicted values relative to observed values is consistent (in the case of figure 1, it is a consistent positive bias), and

(2) the predicted and observed values are significantly and positively correlated, so that the variation in predicted values explains much of the variation in the observed values. This implies that the regional model explains or models the relation between the response variable and the explanatory variables.

Consistent direction of bias in the local data base (predicted and observed data pairs) can be determined by a signed rank test on the paired data (Iman and Conover, 1983, p. 256-260). Correlation of the predicted and observed data can be determined by the test for significance of the rank correlation coefficient, Spearman's rho (r_s) (Iman and Conover, 1983, p. 341). If the test statistic from each of these tests is significant at the selected level, then it might be concluded that a MAP is a valid approach.

Model-Adjustment Procedures

All of the MAP's considered in this report are in the form of a regression analysis (or, in one case, a weighting of the results of two separate regression analyses) in which local data are used for calibration. Regression coefficients are determined using local data, and the resulting 'adjusted' regression models are then used to predict storm-runoff quality at unmonitored sites. The response variable in the regression analyses is the observed load or mean concentration of a constituent in storm runoff for a single storm. The set of explanatory variables used in the regression analyses is different for each procedure, but always includes the predicted value of load or mean concentration from the regional single-storm model. The name for each procedure is an acronym describing the form of the procedure and the set of explanatory variables: for example, MAP-R-P denotes a model-adjustment procedure (MAP) in the form of a regression (R) on the single explanatory variable, predicted value (P) from the regional single-storm model.

Values for the response and explanatory variables were transformed to log units for the regression analysis. From the analysis by Driver and Tasker of the large NURP database, both response and explanatory variables most closely approximate a normal distribution when a log transformation is used (Driver and Tasker, 1990, p. 6). Because the response variables and most of the explanatory variables used in the adjustment procedures were also included in Driver and Tasker's analysis, it is appropriate to use the same transformation.

Single-Factor Regression Against Regional Prediction

Single-factor regression against the predicted value from the regional model, or MAP-1F-P, is a modification of simple linear regression. In this modification, the coefficient, β_1, shown in equation 1 below is forced to unity (suggested by Timothy A. Cohn and Gary D. Tasker, U.S. Geological Survey, oral commun., 1990; documented in Hoos, 1991). The log-transformed observed values of load or concentration in the calibration data set (the local data base) are regressed against the corresponding log-transformed predicted values from the unadjusted regional model using only one calibration coefficient:

$$\log O = \beta_0 + \beta_1 \log P_u, \tag{1}$$

where

- O is observed values of storm-runoff load or mean concentration;
- P_u is predicted values of storm-runoff load or mean concentration from the unadjusted regional model
- β_0 is the single calibration coefficient; and
- β_1 is the regression coefficient forced to unity.

Because MAP-1F-P is not a true regression procedure, the value for the calibration coefficient, β_0, is determined from the calibration data set (local data base) according to a simple formula rather than from the standard regression formula. Using equation 1, the value for β_0 can be computed as:

$$\beta_0 = \overline{\log O} - \overline{\log P_u}, \tag{2}$$

where the overbar denotes mean value.

An adjusted prediction at an unmonitored site i can then be calculated (from the detransformation of equation 1) as

$$P_{ai} = \beta'_0 * P_{ui} * BCF, \tag{3}$$

where

- P_{ai} is the adjusted-model predicted value of storm-runoff load or mean concentration at unmonitored site i;
- β'_0 is 10^{β_0};
- P_{ui} is the unadjusted-regional-model predicted value of storm-runoff load or mean concentration at unmonitored site i; and
- BCF is a bias correction factor.

The BCF must be included in the detransformed model if an unbiased estimate of the mean is to be obtained (Driver and Tasker, 1990; Miller, 1984; and Duan, 1983). The BCF is calculated for each adjustment procedure using a nonparametric method based on the average residuals in original units:

$$BCF = \frac{1}{n} \sum 10^{e_i}, \tag{4}$$

where

- e_i is the least-squares residual for observation i from the calibration data set, in log units; and
- n is the number of observations.

This procedure is appropriate under two sets of conditions: (1) a small calibration data set (the local data base might consist of only 15 data pairs) argues against attempting to calibrate more than one coefficient, and (2) the relation between explanatory variables and the response variable appears to be adequately modeled by the regional model (r_s is significant and positive) and the predicted values are biased in a consistent direction (test statistic from signed rank test is significant) and by a constant factor.

Regression Against Regional Prediction

In the second procedure (MAP-R-P), log-transformed observed values are regressed against a single independent variable (log-transformed predicted values from the unadjusted regional model) in a standard linear regression:

$$\log O = \beta_0 + \beta_1 * \log P_u, \tag{5}$$

where

β_0, β_1 are coefficients determined from a simple linear regression analysis of the calibration data set (local data base).

An adjusted prediction at an unmonitored site i (P_{ai}) can then be calculated (from the detransformation of equation 5) as

$$P_{ai} = \beta_0' * P_{ui}^{\beta_1} * BCF. \tag{6}$$

There are two cases in which the use of this MAP could be preferable to MAP-1F-P. In areas where the calibration data set is relatively large (more than 20 observations), calibration of two regression coefficients can be justified and might provide more accurate results. In other areas, adjustment by a single factor might not be adequate because the difference between the log-transformed observed and predicted values may be a function of the magnitude of the values. Inclusion of the additional β_1 regression coefficient could model this functionality (W.O. Thomas, Jr., U.S. Geological Survey, oral commun., 1991).

Regression Against Regional Prediction and Additional Local Variables

In the third procedure (MAP-R-P+nV), log-transformed observed values are regressed against several independent variables (including the log-transformed predicted values from the unadjusted regional model) in a multiple linear regression:

$$\log O = \beta_0 + \beta_1 * \log P_u + \beta_2 * \log V_1 + \ldots + \beta_{n+1} * \log V_n, \tag{7}$$

where

$\beta_0, \beta_1, \ldots, \beta_{n+1}$ are coefficients determined from multiple linear regression analysis of the calibration data set (local data base); and

$V_1, V_2, \ldots, V_n$ are values of additional explanatory variables from the calibration data set.

An adjusted prediction at an unmonitored site i (P_{ai}) can then be calculated (from the detransformation of equation 7) as

$$P_{ai}=\beta_0'*P_{ui}^{\beta_1}*V_1^{\beta_2}*\ldots*V_n^{\beta_{n+1}}*BCF. \tag{8}$$

This MAP (MAP-R-P+nV) might be appropriate when the pattern of correspondence between O and P_u indicates that a MAP based on P_u alone (MAP-1F-P or MAP-R-P) is not appropriate (when the test statistic from either the signed rank test or the test for significance of r_s is not significant). The most likely candidates for inclusion as additional explanatory variables are physical, land-use, or climatic variables not tested or included in the regional model, but suspected of being significant and a possible source of unexplained error. Antecedent dry days is presented by Driver and Tasker (1990, p. 11-12) as such a variable (although the evidence is contradictory). Because of its inconsistent appearance in the NURP data base, it was excluded from the regression analysis. Percent of drainage area under construction also was presented by Driver and Tasker as a potential variable, particularly for prediction of suspended sediment load or concentration. In cities where the calibration data set (local data base) is relatively large (more than 30 observations), calibration of three or more regression coefficients can be justified and might provide more accurate results.

Weighted Combination of Regional Prediction and Local-Regression Prediction

The fourth procedure (MAP-W) differs fundamentally from the other suggested MAP's. Rather than resulting from regression analysis of observed values against regional-model predicted values (and possibly other variables), the prediction at an unmonitored site i is calculated from an explicit weighting algorithm that weights the predicted value from the unadjusted regional model with a predicted value based only on the local monitoring data (D.R. Helsel, U.S. Geological Survey, oral commun., 1992):

$$P_{ai}=P_{ui}^{j_i}*P_{loci}^{(1-j_i)}*BCF. \tag{9}$$

where

j_i is a weighting factor (a fraction between 0 and 1), which has a unique value for each unmonitored site; and

P_{loci} is the predicted value at unmonitored site i based on local data.

The value for P_{loc} at the unmonitored site i might be derived from a regression model from the local data base (a regression analysis of observed values against values for selected physical, land-use, and climatic characteristics), or might be set as the mean value of the observed values. The weighting factor, j_i, is a function of the variances of prediction at the unmonitored site i (V_{pi}) resulting from the estimating procedures for P_{loc} and P_u (G.D. Tasker, U.S. Geological Survey, oral commun., 1992):

$$j_i=\frac{V_{pi\text{-}loc}}{(V_{pi\text{-}loc}+V_{pi\text{-}u})}, \tag{10}$$

$$V_{pi\text{-}loc}=SE_{loc}^2(1+x_i(X'X)^{-1}x_i'), \tag{11}$$

$$V_{pi\text{-}u}=SE_u^2(1+z_i(Z'Z)^{-1}z_i'), \tag{12}$$

where

- $V_{pi\text{-}loc}$ is variance of prediction at unmonitored site *i* for the local regression model;
- $V_{pi\text{-}u}$ is variance of prediction at unmonitored site *i* for the unadjusted regional model;
- SE_{loc} is standard error of estimate (in log units) for the local regression model;
- SE_u is standard error of estimate (in log units) computed from the regional (NURP) calibration data set for the unadjusted regional model;
- x_i is a (1 x p) row vector of the p-1 explanatory variables used in the local regression, evaluated (in log units) for unmonitored site *i*, augmented by a 1 as the first element;
- X is a (n x p) matrix of the p-1 explanatory variables used in the local regression, evaluated (in log units) for all n sites in the local calibration data set, augmented by a 1 as the first column;
- z_i is a (1 x k) row vector of the k-1 explanatory variables used in the regional regression, evaluated (in log units) for unmonitored site *i*, augmented by a 1 as the first element; and
- Z is a (m x k) matrix of the k-1 explanatory variables used in the regional regression, evaluated (in log units) for all m sites in the regional (NURP) calibration data set, augmented by a 1 as the first column.

SE_u is taken from the published values (Driver and Tasker, 1990, tables 2, 3, and 6) for the regional model; these values are included for selected constituents and model types in table 1 of this report (in columns titled 'Log').

SE_{loc} can be computed according to the general formula for SE:

$$SE=\sqrt{\frac{\sum(\log O_i-\log P_i)^2}{n-(k+1)}}, \tag{13}$$

where

- SE is standard error of estimate of a regression model for the calibration data set, in log units;
- O_i is i^{th} observed value for the response variable in the calibration data set;
- P_i is i^{th} fitted value for the response variable in the calibration data set;
- n is number of observations in the calibration data set; and
- k is number of explanatory variables in the regression model.

The matrix operations are factored into the formulas for V_{pi} to make j_i responsive to the difference between the explanatory-variable values for the unmonitored site and the mean values for the calibration data sets associated with P_u and P_{loc}. A simpler, although statistically less valid, formula for V_{pi} can be employed by dropping the term comprising the matrix operations from equations 11 and 12 giving:

$$V_{pi\text{-}loc}=SE_{loc}^2, \tag{14}$$

$$V_{pi\text{-}u}=SE_u^2. \tag{15}$$

In this case, the variance of prediction and the weighting factor are not calculated uniquely for an unmonitored site i, but rather are constants (V_p and j, rather than V_{pi} and j_i) for a particular city and constituent.

The MAP-W might be appropriate (as was the case for MAP-R-P+nV) when the pattern of correspondence between O and P_u indicates that a MAP based on P_u alone (MAP-1F-P or MAP-R-P) is not appropriate (when the test statistic from either the signed rank test or the test for significance of r_s is not significant). Selection of explanatory variables for the local regression analysis should be made using accepted statistical procedures; for example, a best-regression analysis. A list of candidate explanatory variables should be compiled based upon knowledge of processes controlling storm-runoff quality in the area of interest. A starting point for the compilation of this list might be the six or seven most significant variables from the regional regression analyses of Driver and Tasker. The absolute value of the standardized beta coefficient for an explanatory variable (Driver and Tasker, 1990, table 4) can be used as an indication of its significance in their analysis. The analyst can then add other explanatory variables believed to be controlling variables of urban runoff quality (for example, antecedent dry days, or percent of drainage area under construction). The best regression model for a set of k explanatory variables can then be determined by regression analysis of the 2^k possible subsets and comparison of an appropriate statistic from the regression (for example, the PRESS statistic or Mallows Cp; see Draper and Smith, 1981, for additional information on these methods). The analyst, however, might wish to restrict his choice to subsets with fewer than a certain number of variables depending upon the size of the calibration data set.

Selection of appropriate adjustment procedures

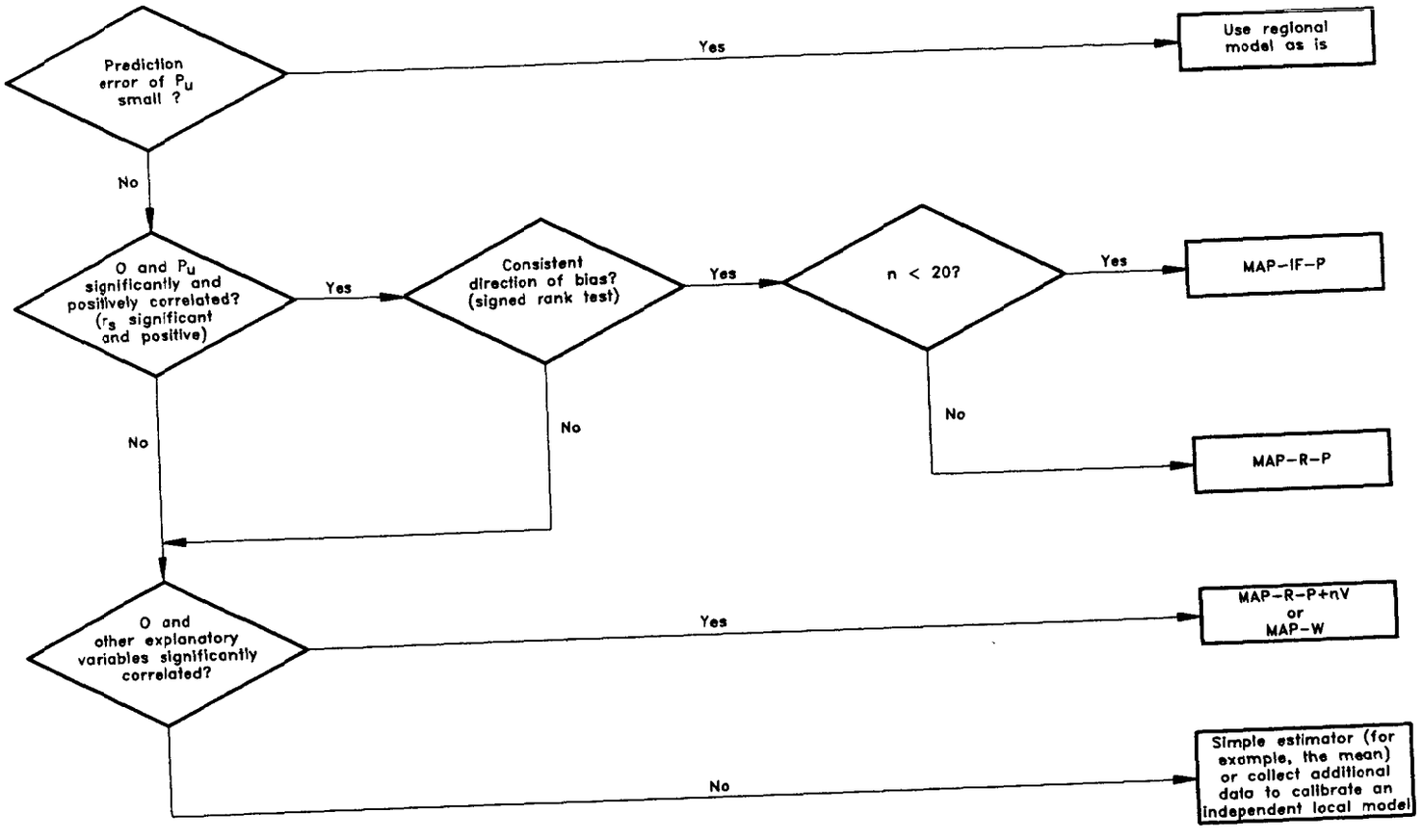
The conditions for application of each MAP cited in the preceding discussion are organized into a scheme (fig. 2) to select the most appropriate MAP for a selected constituent model and local data base. This scheme is based solely on exploratory data analysis (EDA) of the local data base.

In the first operation in this scheme, the analyst determines if any adjustment procedure is necessary, or if the regional model can be used without adjustment. Examination of data plots of P_u and O, similar to figure 1, and evaluation of an appropriate error statistic, such as root mean square error, can guide the data analyst in determining whether the prediction error of the unadjusted regional model is within acceptable limits.

Next the analyst performs the test for significance of r_s and the signed rank test. If the test statistic from each of these tests is significant at the selected level, then a MAP based on P_u alone [MAP-1F-P and MAP-R-P) is most appropriate. The choice between these two MAP's can be based on either the size of the calibration data set (as indicated in figure 2), or consideration as to whether the observed bias can be corrected by a constant factor (β_1 for the MAP-R-P is not significantly different from unity for the calibration data set).

If either of the test statistics is not significant at the selected level, the analyst continues the EDA, testing the correlation between the response variable and the candidate explanatory variables to be used in MAP-R-P+nV and MAP-W. If any of the correlations is significant, the analyst may select either MAP-R-P+nV or MAP-W. No basis is known for choosing between MAP-R-P+nV and MAP-W using EDA.

If none of the tested correlations are acceptably significant, then the analyst should reject the MAP approach for that constituent. Two possible alternatives are: (1) use a simple estimator, such as mean value of the response variables from the local data base, to estimate constituent load and mean concentration; or (2) collect sufficient local runoff quality data to allow for calibration of a completely independent, local regression model.



EXPLANATION

P_u	PREDICTED VALUES OF STORM-RUNOFF LOAD OR MEAN CONCENTRATION FROM THE UNADJUSTED REGIONAL MODEL	MAP-IF-P	SINGLE-FACTOR REGRESSION AGAINST REGIONAL PREDICTION
O	OBSERVED VALUES OF STORM-RUNOFF LOAD OR MEAN CONCENTRATION	MAP-R-P	REGRESSION AGAINST REGIONAL PREDICTION
r_s	SPEARMAN'S RHO	MAP-R-P+nV	REGRESSION AGAINST REGIONAL PREDICTION AND ADDITIONAL LOCAL VARIABLES
n	NUMBER OF OBSERVATIONS	MAP-W	WEIGHTED COMBINATION OF REGIONAL PREDICTION AND LOCAL-REGRESSION PREDICTION

Figure 2. Flowchart guiding selection of model-adjustment procedure (MAP) based on exploratory data analysis of the calibration data set.

Other logical schemes for selecting the appropriate MAP are possible. The analyst could calibrate and compute associated error statistics (for example, *SE* or PRESS) for all MAP's, then use relative values of error statistics to guide selection of the MAP. Using *SE* alone to guide MAP selection is shown later to be unreliable. The PRESS method (Draper and Smith, 1981) cross validates a calibration using a 1 and (n-1) data split of the calibration data set repeated n times, and therefore the PRESS statistic may be a more reliable indicator of predictive accuracy. No scheme based on comparison of calibration error statistics alone, however, can provide the basis for deciding whether the MAP approach is valid for a particular data base and constituent, or whether some alternative to model adjustment should be sought. The scheme presented in figure 2 does provide such a basis.

Model-Adjustment Procedure Testing

The four proposed MAP's were tested for relative predictive accuracy for unmonitored sites or storms, and for relative sensitivity to size of the calibration data set. The performance of each MAP was compared among each type of model (Lsa, Csa, L3) to determine whether the models differed in their suitability for a particular MAP. The results of these tests were used in turn to measure the success of the MAP selection scheme described in figure 2.

Test Procedures

Testing was accomplished using a split-sample analysis of three separate data bases; the 'local' data bases for the NURP study areas in Denver, Colorado (region I), Bellevue, Washington (region II), and Knoxville, Tennessee (region III). Each region was represented so that each set of regional models could be tested. Values for storm-runoff load (response variable) were read directly from archived data files for each city (Mustard and others, 1987, table 1). Values for storm-runoff mean concentration (response variable) were calculated by dividing storm-runoff load, in pounds, by average storm-runoff depth over the basin, in inches, and by total contributing drainage area, in square miles, multiplied by a conversion factor. Predicted values from the unadjusted regional model were computed from values for the basin and storm characteristics (explanatory variables) read from the archived data files.

For the split-sample analysis, the data base for each city was divided into two data sets; a calibration data set and verification data set. Division into two groups of about equal size was accomplished following a systematic procedure to avoid bias. Individual storms were ordered first by site number and multiple storms at each site were ordered chronologically. Storms on this master list were then assigned alternately to the calibration or verification set. This resulted in sample sizes for the calibration and verification sets of 56 each for the Denver data base, 41 each for the Bellevue data base, and 31 each for the Knoxville data base.

The EDA and MAP selection scheme prescribed in figure 2 were applied to the calibration data set from each data base to select the most appropriate MAP for each constituent model. Values for the test statistics and the selected MAP option are presented separately for the Bellevue, Denver, and Knoxville data bases (tables 2, 3, and 4, respectively). For the Bellevue data base, the MAP-1F-P or MAP-R-P were selected for most of the load models, whereas the MAP-R-P+nV or MAP-W were selected for two of the four concentration models (table 2). For most constituents in the Denver data base, the MAP-R-P+nV or MAP-W were selected for both load and concentration models (table 3). For most constituents in the Knoxville data base, the EDA suggested that the MAP approach should be rejected in favor of alternatives (table 4).

Following initial exploratory data analysis, observations in the calibration data set were used to derive coefficients ($\beta_0,\beta_1,\ldots,\beta_{n+1}$, defined in equations 1, 5, and 7; and SE_{loc}, defined in equations 11 and 13) for the MAP's. Two indications of predictive accuracy were computed and compared among the MAP's for the

Table 2. Exploratory data analysis of the calibration data sets from the data base for Bellevue, Washington

[RMSE, root mean square error between observed and predicted (from unadjusted regional model) values of the response variable, in log units; r_s, Spearman's rho; 0.005 is the selected level of significance for the test statistic; O, observed value of the response variable; P_u, predicted value of response variable from the unadjusted regional model; TRN, total storm rainfall; DA, total contributing drainage area; IA, impervious area; ADD, antecedent dry days; COD, chemical oxygen demand; TKN, total kjeldahl nitrogen; PB, total recoverable lead; SS, suspended solids; Lsa, stepwise-analysis regression model for storm-runoff load; Csa, stepwise-analysis regression model for storm-runoff mean concentration; L3, 3-variable regression model for storm-runoff load]

Constituent and model type	Prediction error		O and P_u positively correlated		Consistent direction of bias		Correlation of variable with O				O significantly correlated with any variable?	Best MAP
	RMSE	Acceptably small?	r_s	Significant at 0.005?	p-value	Significant at 0.005?	TRN	DA	IA	ADD		
COD.Lsa	0.459	N	0.893	Y	<0.0001	Y	0.760	0.358	-0.434	-0.130	Y	MAP-1F-P; MAP-R-P
COD.Csa	.440	N	.428	Y	<.0001	Y	-.469	-.099	-.020	.571	Y	MAP-1F-P; MAP-R-P
COD.L3	.433	N	.887	Y	<.0001	Y	.760	.358	-.434	-.130	Y	MAP-1F-P; MAP-R-P
TKN.Lsa	.345	N	.875	Y	<.0001	Y	.753	.409	-.460	-.142	Y	MAP-1F-P; MAP-R-P
TKN.Csa	.339	N	.239	N	<.0001	Y	-.322	.066	-.150	.513	Y	MAP-R-P+nV; MAP-W
TKN.L3	.449	N	.876	Y	<.0001	Y	.753	.409	-.460	-.142	Y	MAP-1F-P; MAP-R-P
PB.Lsa	.379	N	.806	Y	<.0001	Y	.632	.417	-.375	-.215	Y	MAP-1F-P; MAP-R-P
PB.Csa	.360	N	.327	N	.0002	Y	-.300	-.063	-.046	.506	Y	MAP-R-P+nV; MAP-W
PB.L3	.412	N	.792	Y	.002	Y	.718	.322	-.394	-.100	Y	MAP-1F-P; MAP-R-P
SS.Lsa	.495	N	.814	Y	<.0001	Y	.800	.210	-.296	-.266	Y	MAP-1F-P; MAP-R-P
SS.Csa	.435	N	.205	N	<.0001	Y	.007	-.200	-.022	.222	N	None
SS.L3	.711	N	.816	Y	<.0001	Y	.800	.210	-.296	-.266	Y	MAP-1F-P; MAP-R-P

Table 3. Exploratory data analysis of the calibration data sets from the data base for Denver, Colorado

[RMSE, root mean square error between observed and predicted (from unadjusted regional model) values of the response variable, in log units; r_s, Spearman's rho; 0.005 is the selected level of significance for the test statistic; O, observed value of the response variable; P_u, predicted value of response variable from the unadjusted regional model; TRN, total storm rainfall; DA, total contributing drainage area; IA, impervious area; ADD, antecedent dry days; COD, chemical oxygen demand; TKN, total kjeldahl nitrogen; PB, total recoverable lead; Lsa, stepwise-analysis regression model for storm-runoff load; Csa, stepwise-analysis regression model for storm-runoff mean concentration; L3, 3-variable regression model for storm-runoff load]

Constituent and model type	Prediction error		O and P_u positively correlated		Consistent direction of bias		Correlation of variable with O			O significantly correlated with any variable?	Best MAP
	RMSE	Acceptably small?	r_s	Significant at 0.005?	p-value	Significant at 0.005?	TRN	DA	IA		
COD.Lsa	0.336	N	0.741	Y	0.784	N	0.446	0.511	-0.078	Y	MAP-R-P+nV; MAP-W
COD.Csa	.216	Y	.754	Y	.28	N	-.788	.228	.020	Y	MAP-R-P+nV; MAP-W[1]
COD.L3	.344	N	.69	Y	.025	N	.446	.511	-.078	Y	MAP-R-P+nV; MAP-W
TKN.Lsa	.305	N	.83	Y	.245	N	.524	.604	-.117	Y	MAP-R-P+nV; MAP-W
TKN.Csa	.225	Y	.691	Y	.245	N	-.685	-.153	.145	Y	MAP-R-P+nV; MAP-W[1]
TKN.L3	.375	N	.768	Y	.0008	Y	.524	.604	-.117	Y	MAP-1F-P; MAP-R-P
PB.Lsa	.458	N	.797	Y	.0778	N	.144	.857	-.449	Y	MAP-R-P+nV; MAP-W
PB.Csa	.282	Y	.631	Y	.0117	N	-.539	.379	.354	Y	MAP-R-P+nV; MAP-W[1]
PB.L3	.499	N	.787	Y	.0092	N	-.144	.857	-.449	Y	MAP-R-P+nV; MAP-W

[1] The value for RMSE indicates, however, that the regional model could be used unadjusted.

Table 4. Exploratory data analysis of the calibration data sets from the data base for Knoxville, Tennessee

[RMSE, root mean square error between observed and predicted (from unadjusted regional model) values of the response variable, in log units; r_s, Spearman's rho; 0.005 is the selected level of significance for the test statistic; O, observed value of the response variable; P_u, predicted value of response variable from the unadjusted regional model; TRN, total storm rainfall; DA, total contributing drainage area; IA, impervious area; ADD, antecedent dry days; COD, chemical oxygen demand; TKN, total kjeldahl nitrogen; PB, total recoverable lead; Lsa, stepwise-analysis regression model for storm-runoff load; Csa, stepwise-analysis regression model for storm-runoff mean concentration; L3, 3-variable regression model for storm-runoff load]

Constituent and model type	Prediction error		O and P_u positively correlated		Consistent direction of bias		Correlation of variable with O			O significantly correlated with any variable?	Best MAP
	RMSE	Acceptably small?	r_s	Significant at 0.005?	p-value	Significant at 0.005?	TRN	DA	IA		
COD.Lsa	0.641	N	0.481	Y	<0.0001	Y	0.356	0.199	0.283	N	MAP-1F-P; MAP-R-P
COD.Csa	.497	N	.050	N	<.0001	Y	-.498	.012	.003	N	None
COD.L3	.625	N	.581	Y	.9999	N	.356	.199	.283	N	None
TKN.Lsa	.924	N	.320	N	<.0001	Y	.232	.245	.280	N	None
TKN.Csa	.481	N	.069	N	.0014	Y	-.449	.341	-.041	N	None
TKN.L3	.894	N	.425	Y	<.0001	Y	.232	.245	.280	N	MAP-1F-P; MAP-R-P
PB.Lsa	.639	N	.614	Y	<.0001	Y	.320	.314	.227	N	MAP-1F-P; MAP-R-P
PB.Csa	.296	Y	.181	N	.9999	N	-.449	.341	-.041	N	None[1]
PB.L3	.714	N	.614	Y	<.0000	Y	.320	.314	.227	N	MAP-1F-P; MAP-R-P

[1] The value for RMSE indicates, however, that the regional model could be used unadjusted.

calibration data set; the coefficient of determination (r^2) and the standard error of the estimate (*SE*). If the r^2 value is multiplied by 100, it represents the percentage of variation in the response variables that is explained by the explanatory variables. The *SE* is a measure of how well the estimated values (from the MAP) agree with the observed values for the calibration data set, and is computed, in log units, according to equation 13. The *SE*, in percent, can be calculated from the *SE*, in log units, according to the formula

$$SE(percent)=100[e^{SE^2*5.302}-1]^{1/2}. \tag{16}$$

The *SE* can be interpreted as follows: approximately two out of three observed values will fall within one *SE* of the estimated value, if the residuals are normally distributed. Computer programs used to perform the exploratory data analysis and MAP-calibration calculations for each calibration data set are given in Supplements A and B, respectively.

Log-transformed observations in the verification data set were used to measure how well the adjusted models estimated the response variables (log-transformed storm-runoff load and mean concentration) for an unmonitored site or storm. Predictive accuracy for the verification data set was measured using the root mean square error of the estimated response variable, calculated as:

$$RMSE_v=\sqrt{\frac{\sum(\log O_{iv}-\log P_{aiv})^2}{n}} \tag{17}$$

where

$RMSE_v$ is root mean square error for the verification data set, in log units;

O_{iv} is i^{th} observed value for the response variable in the verification data set;

P_{aiv} is i^{th} predicted value for the response variable in the verification data set; and

n is number of observations in the verification data set.

The relative predictive accuracy of the MAP's for the verification data set was used in turn to measure the success of the MAP selection scheme. This was accomplished by comparing the selected MAP (tables 2, 3, and 4) for a constituent model with the MAP with the smallest $RMSE_v$.

The Lsa, Csa, and L3 models for the constituents COD, TKN, PB, and SS were included in the testing. The regional models for TKN were among the most accurate developed by Driver and Tasker (1990, p. 32), whereas the regional models for SS were the least accurate. Consequently, the results for these selected constituents might be expected to provide an estimate of the range of results for all 11 modeled constituents.

Application of the MAP-W procedure requires development of a local regression model (using local basin and storm characteristics as explanatory variables and excluding the predicted value from the unadjusted regional model). Although in a real application, a best-regression analysis examining all possible combinations of a nominated list of explanatory variables should be performed, this was deemed neither feasible nor necessary for testing purposes. For these tests, best-regression analysis was performed using only four variables where they were available: total storm rainfall (TRN), drainage area (DA), percent impervious area (IA), and antecedent dry days (ADD). The first three variables in this list were most consistently found to be significant explanatory variables in the regression analysis by Driver and Tasker (1990, p. 17, 21).

The selection of the additional explanatory variable for the MAP-R-P+nV differed among cities. For the Bellevue analysis, the variable ADD was used. Because this variable was not present in the data base for Denver or Knoxville, the MAP-R-P+nV for these cities was tested using as the additional explanatory variable the most significant variable from the local regression analysis of the calibration data set.

Test Results

Comparison among MAP predictive accuracy for the verification data set was made to indicate the most accurate MAP for each constituent model for each of the test data bases. None of the MAP's emerged from the split-sample testing as clearly superior for all constituent models and data bases. These test results cannot, therefore, be used to indicate the most reliable MAP for any other local data base. These results can be used to evaluate proposed procedures for selecting a MAP for a particular constituent and data base, and in this way are of benefit to analysts working with other local data bases. The following discussion of test results for each data base emphasizes this evaluation process.

Bellevue

Results of the split-sample analysis are presented in table 5 for the Bellevue data base. For each constituent model, the $RMSE_v$ (in log units) and the relative ranking for each MAP are reported, along with the $RMSE_v$ and relative ranking for other estimators; the prediction from the unadjusted regional model, the prediction from local regression models, and the mean value of the response variable (in the calibration data set). When results for all models were aggregated, the MAP-R-P provided the best predictive accuracy for the verification data set, reducing the $RMSE_v$ from a mean value of 0.436 log units (or 132 percent) for the

Table 5. Root mean square errors and associated rankings for model-adjustment procedures and other estimators for verification data sets, compared with rankings for standard error of estimate for corresponding calibration data sets, from the data base for Bellevue, Washington

[Test results from split-sample analysis of calibration and verification data-set sizes of 41 each; COD, chemical oxygen demand; TKN, total kjeldahl nitrogen; PB, total recoverable lead; SS, suspended solids; Lsa, stepwise-analysis regression model for storm-runoff load; Csa, stepwise-analysis regression model for storm-runoff mean concentration; L3, 3-variable regression model for storm-runoff load; MEAN, mean value of response variable from calibration data set used as an estimator; Unadjusted regional model, the appropriate single-storm model from Driver and Tasker (1990, tables 1, 3, and 5); LOC, local regression model based on total storm rainfall, drainage area, impervious area, and antecedent dry days; model-adjustment procedures (MAP's) defined in explanation in text; MAP-R-P+nV used antecedent dry days as additional explanatory variable; MAP-W used local regression model defined above in LOC; $RMSE_v$, root mean square error for verification data set, in log units]

Constituent and model type	MEAN		Unadjusted regional model		LOC		MAP-1F-P		MAP-R-P		MAP-R-P+nV		MAP-W	
	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank
COD.Lsa	0.465	7.0	0.437	6.0	0.283	4.0	0.257	1.0	0.260	2.0	0.266	3.0	0.283	5.0
COD.Csa	.238	6.0	.397	7.0	.205	1.0	.225	4.0	.226	5.0	.213	2.0	.214	3.0
COD.L3	.465	7.0	.409	6.0	.283	5.0	.258	1.0	.260	2.0	.267	3.0	.279	4.0
TKN.Lsa	.498	7.0	.341	6.0	.262	4.0	.253	2.0	.248	1.0	.255	3.0	.265	5.0
TKN.Csa	.220	6.0	.289	7.0	.181	1.0	.208	5.0	.207	4.0	.182	2.0	.195	3.0
TKN.L3	.498	7.0	.442	6.0	.262	4.0	.251	2.0	.249	1.0	.255	3.0	.274	5.0
PB.Lsa	.582	7.0	.391	6.0	.325	5.0	.310	2.0	.299	1.0	.316	4.0	.313	3.0
PB.Csa	.331	6.0	.381	7.0	.317	2.0	.319	4.0	.318	3.0	.278	1.0	.322	5.0
PB.L3	.554	7.0	.439	6.0	.380	5.0	.360	3.0	.338	2.0	.329	1.0	.378	4.0
SS.Lsa	.643	7.0	.522	6.0	.454	5.0	.401	1.0	.402	2.0	.413	3.0	.443	4.0
SS.Csa	.373	4.0	.463	7.0	.377	6.0	.353	2.0	.356	3.0	.352	1.0	.376	5.0
SS.L3	.643	6.0	.721	7.0	.454	4.0	.402	1.0	.403	2.0	.414	3.0	.459	5.0
Mean	.459	6.4	.436	6.4	.315	3.8	.300	2.3	.297	2.3	.295	2.4	.317	4.3
Mean Lsa	.547	7.0	.423	6.0	.331	4.5	.305	1.5	.302	1.5	.313	3.3	.326	4.3
Mean Csa	.291	5.5	.383	7.0	.270	2.5	.276	3.8	.277	3.8	.256	1.5	.277	4.0
Mean L3	.540	6.8	.503	6.3	.345	4.5	.318	1.8	.313	1.8	.316	2.5	.348	4.5

Rankings of standard error of estimate for calibration data sets[1]

	MEAN Rank	Unadjusted regional model Rank	LOC Rank	MAP-1F-P Rank	MAP-R-P Rank	MAP-R-P+nV Rank	MAP-W Rank
Mean	6.3	6.6	1.2	4.4	3.3	2.8	3.4
Mean Lsa	6.8	6.3	1.5	4.0	3.0	3.0	3.5
Mean Csa	5.8	7.0	1.0	4.8	3.8	3.0	2.8
Mean L3	6.5	6.5	1.0	4.5	3.3	2.3	4.0

[1]Value ranked for unadjusted regional model is actually root mean square error for calibration data set.

unadjusted regional model, to 0.297 log units (or 77 percent). The MAP-1F-P provided almost the same $RMSE_v$ reduction, to 0.300 log units (or 78 percent). The MAP-W proved least effective in reducing $RMSE_v$.

When results were aggregated only by model type (Lsa, Csa, and L3), a different pattern of MAP performance emerged. The results for the Lsa and L3 models were similar to the total-aggregate results (MAP-R-P and MAP-1F-P providing the best predictive accuracy). For the Csa models, however, the procedures that included local explanatory variables (MAP-R-P+nV and local regression, both of which included antecedent dry days as an explanatory variable) gave the best results.

The success of the proposed MAP selection procedure for this data base is evaluated by comparing, for each constituent model, the MAP that was selected on the basis of EDA of the calibration data set (table 2) with the MAP that produced the smallest $RMSE_v$ (table 5). In the 11 cases for which a MAP selection was made, nine of the selections provided the most accurate MAP. These results support the validity of the MAP selection procedure. The support is somewhat weakened, however, by the fact that the procedure does not provide a basis for choosing between MAP-R-P+nV and MAP-W.

As an alternative to the EDA approach to MAP selection, the choice could be guided by relative values, among the MAP's, of *SE* for the calibration data set. As with the preceding approach, the success of this criterion is evaluated by comparing, for each constituent model, the MAP that was selected on the basis of minimum *SE* for the calibration data set with the MAP that produced the smallest $RMSE_v$. The relative rankings for *SE* for the calibration data sets, aggregated by model type, are presented in table 5 in order to make this comparison.

For the Bellevue data base, selection on the basis of this criteria would favor the local regression model, which was ranked first (smallest values of *SE* for calibration data set) for both load and concentration models. Application of the local regression model to the verification data set, however, yielded among the poorest results (largest value of $RMSE_v$) of all the tested procedures for the load models, and yielded the second-ranked results for the concentration model. Similarly, the top-ranked procedures for the verification data set for the load models, MAP-1F-P and MAP-R-P, were among the poorest ranked for the calibration data set. This mismatch suggests that, whereas it may be possible to calibrate a local regression model so that it fits the calibration data set more closely than any MAP, its predictive accuracy might be much smaller than the MAP's for an unmonitored site or storm. Clearly the MAP selection procedure based on EDA is a better guide to selection of an appropriate MAP than the relative magnitude of *SE* for the calibration data set.

Denver

Results of the split-sample analysis are presented in table 6 for the Denver data base. MAP-W provided the best predictive accuracy for almost all of the verification data sets, reducing the $RMSE_v$ from a mean value of 0.370 log units (103 percent), for the unadjusted regional model, to 0.312 log units (82 percent). The MAP-1F-P proved least effective in reducing $RMSE_v$. MAP performance did not differ significantly among model types (Lsa, Csa, and L3).

The MAP selection procedure based on EDA was successful for the Denver data base. The selected MAP (table 3) proved to be the most accurate (smallest $RMSE_v$, table 6) for seven of the nine models analyzed. The lack of consistent direction of bias between O and P_u prompted selection of the 'MAP-R-P+nV or MAP-W' option for almost every model. Although the choice between MAP-R-P+nV and MAP-W cannot be made based on EDA, this did not detract substantially because the two MAP's performed almost equally.

Table 6. Root mean square errors and associated rankings for model-adjustment procedures and other estimators for verification data sets, compared with rankings for standard error of estimate for corresponding calibration data sets, from the data base for Denver, Colorado

[Test results from split-sample analysis of calibration and verification data-set sizes of 56 each; COD, chemical oxygen demand; TKN, total kjeldahl nitrogen; PB, total recoverable lead; SS, suspended solids; Lsa, stepwise-analysis regression model for storm-runoff load; Csa, stepwise-analysis regression model for storm-runoff mean concentration; L3, 3-variable regression model for storm-runoff load; MEAN, mean value of response variable from calibration data set used as an estimator; Unadjusted regional model, the appropriate single-storm model from Driver and Tasker (1990, tables 1, 3, and 5); LOC, local regression model based on total storm rainfall, drainage area, and impervious area; model-adjustment procedures (MAP's) defined in explanation in text; MAP-R-P+nV used drainage area as additional explanatory variable in load models, total storm rainfall in mean concentration models; MAP-W used local regression model defined above in LOC; $RMSE_v$, root mean square error for verification data set, in log units]

Constituent and model type	MEAN		Unadjusted regional model		LOC		MAP-1F-P		MAP-R-P		MAP-R-P+nV		MAP-W	
	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank
COD.Lsa	0.543	7.0	0.358	5.0	0.306	2.0	0.365	6.0	0.357	3.0	0.358	4.0	0.296	1.0
COD.Csa	.343	7.0	.233	4.0	.228	3.0	.238	6.0	.233	5.0	.225	2.0	.220	1.0
COD.L3	.584	7.0	.379	5.0	.303	1.0	.381	6.0	.372	4.0	.367	3.0	.313	2.0
TKN.Lsa	.682	7.0	.377	5.0	.342	2.0	.403	6.0	.373	4.0	.372	3.0	.327	1.0
TKN.Csa	.303	7.0	.282	6.0	.275	4.0	.281	5.0	.266	1.0	.268	2.0	.272	3.0
TKN.L3	.682	7.0	.402	6.0	.342	1.0	.395	5.0	.363	4.0	.362	3.0	.346	2.0
PB.Lsa	1.035	7.0	.474	5.0	.379	2.0	.484	6.0	.453	4.0	.448	3.0	.361	1.0
PB.Csa	.420	7.0	.335	6.0	.285	2.0	.332	5.0	.329	4.0	.327	3.0	.278	1.0
PB.L3	1.035	7.0	.493	5.0	.379	1.0	.500	6.0	.417	4.0	.398	3.0	.392	2.0
Mean	.625	7.0	.370	5.2	.315	2.0	.375	5.7	.351	3.7	.347	2.9	.312	1.6
Mean	.753	7.0	.403	5.0	.342	2.0	.417	6.0	.394	3.7	.393	3.3	.328	1.0
Mean Csa	.355	7.0	.283	5.3	.263	3.0	.284	5.3	.276	3.3	.273	2.3	.257	1.7
Mean L3	.767	7.0	.425	5.3	.341	1.0	.425	5.7	.384	4.0	.376	3.0	.350	2.0

Rankings of standard error of estimate for calibration data sets[1]

	MEAN Rank	Unadjusted regional model Rank	LOC Rank	MAP-1F-P Rank	MAP-R-P Rank	MAP-R-P+nV Rank	MAP-W Rank
Mean	7.0	5.6	1.9	5.3	3.2	3.4	1.6
Mean Lsa	7.0	5.0	3.0	5.7	2.7	3.7	1.0
Mean Csa	7.0	6.0	1.7	5.0	3.7	3.3	1.3
Mean L3	7.0	5.7	1.0	5.3	3.3	3.3	2.3

[1]Value ranked for unadjusted regional model is actually root mean square error for calibration data set.

Selection of a MAP based on the relative ranking of *SE* for the calibration data set (table 6) would favor the MAP-W, so that selection guided by this criteria would have been successful (resulted in choosing the MAP with the greatest predictive accuracy for the verification data set) for this data base.

Knoxville

Results of the split-sample analysis are presented in table 7 for the Knoxville data base. MAP-W provided the best predictive accuracy for the verification data sets for the Knoxville data base (table 4), reducing $RMSE_v$ from a mean value of 0.674 log units, or 318 percent, for the unadjusted regional model, to a mean value of 0.475 log units, or 152 percent. The MAP's based on P_u alone performed poorly; for many models, $RMSE_v$ was larger than for estimation with a constant (the mean value of the response variable from the calibration data set). In addition, MAP-R-P+nV, MAP-W, and the local regression model were not as successful (compared with the results from the Bellevue and Denver data bases) in reducing $RMSE_v$, compared with the mean estimator.

The MAP selection procedure based on EDA had mixed success for the Knoxville data base. The MAP approach was deemed inappropriate (table 4) for five of the nine models analyzed, and so comparison with $RMSE_v$ (table 7) was not possible. The pattern of $RMSE_v$ described in the preceding paragraph validates the rejection by the EDA of the MAP approach, however. Such a rejection does provide the analyst with some useful information, warning the analyst that: (1) other explanatory variables should be sought and included in the analysis; or (2) the MAP approach should be abandoned in favor of a simple estimator or collection of additional monitoring data.

For the remaining four models, the selected MAP (MAP-1F-P or MAP-R-P) proved to be the poorest performer. The lower reliability of the MAP-selection procedure for the Knoxville data base may be due to the large difference (several orders of magnitude) between values of O and P_u for the calibration data set, as evident from the values of root mean square error (table 4). Thus, despite the apparently significant level of correlation and consistent bias between O and P_u, the MAP's based on P_u alone were not successful in reducing error compared with MAP's that included additional, although weakly correlated, explanatory variables.

Sensitivity analysis

To examine variance of MAP performance as a function of calibration data set size, split-sample analysis was repeated several times for the Bellevue data base, using different sizes for the calibration data set. Results from this sensitivity analysis are presented in table 8 for calibration data-set (CDS) sizes of 51, 41, 31, and 21 and for the Lsa and Csa models. Test bias, which might result from selecting biased subsets of the CDS, was avoided by random selection of observations for the CDS from the entire data base. For each constituent and model form, the random selection and testing was repeated 50 times and the results averaged.

As expected, $RMSE_v$ increased for all MAP's as CDS size decreased. Because $RMSE_v$ increased by different amounts for different procedures, however, the relative ranking among the procedures changed as the CDS size decreased. For the load models, the increase in $RMSE_v$ was larger for the local regression model than for the other procedures. The greater number of explanatory variables and calibration coefficients for the local regression model and MAP-R-P+nV, which causes a larger variance of prediction for these procedures, might cause the model to perform more poorly, compared to the other procedures, for the smaller CDS size. This also might explain why the MAP-1F-P and MAP-R-P reverse their relative ranking to first and second, respectively, as CDS size decreases. The single calibration coefficient in MAP-1F-P minimizes the variance of prediction. Although the relative ranking of MAP-W improved with decreasing CDS size, the best-performing MAP's for load models, regardless of CDS size, were the MAP-1F and MAP-R-P.

Table 7. Root mean square errors and associated rankings for model-adjustment procedures and other estimators for verification data sets, compared with rankings for standard error of estimate for corresponding calibration data sets, from the data base for Knoxville, Tennessee

[Test results from split-sample analysis of calibration and verification data-set sizes of 31 each; COD, chemical oxygen demand; TKN, total kjeldahl nitrogen; PB, total recoverable lead; SS, suspended solids; Lsa, stepwise-analysis regression model for storm-runoff load; Csa, stepwise-analysis regression model for storm-runoff mean concentration; L3, 3-variable regression model for storm-runoff load; MEAN, mean value of response variable from calibration data set used as an estimator; Unadjusted regional model, the appropriate single-storm model from Driver and Tasker (1990, tables 1, 3, and 5); LOC, local regression model based on total storm rainfall, drainage area, and impervious area; model-adjustment procedures (MAP's) defined in explanation in text; MAP-R-P+nV used impervious area as additional explanatory variable in load models, total storm rainfall in mean concentration models; MAP-W used local regression model defined above in LOC; $RMSE_v$, root mean square error for verification data set, in log units]

Constituent and model type	MEAN		Unadjusted regional model		LOC		MAP-1F-P		MAP-R-P		MAP-R-P + nV		MAP-W	
	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank
COD.Lsa	0.507	1.0	0.741	7.0	0.517	3.0	0.553	6.0	0.530	5.0	0.507	2.0	0.521	4.0
COD.Csa	.521	4.0	.545	7.0	.500	2.0	.527	5.0	.529	6.0	.516	3.0	.473	1.0
COD.L3	.507	1.0	.750	7.0	.517	2.0	.564	6.0	.556	5.0	.528	3.0	.528	4.0
TKN.Lsa	.373	5.0	.868	7.0	.325	1.0	.404	6.0	.369	4.0	.337	2.0	.360	3.0
TKN.Csa	.540	7.0	.491	1.0	.532	6.0	.495	2.0	.516	5.0	.507	4.0	.505	3.0
TKN.L3	.373	6.0	.824	7.0	.325	1.0	.367	5.0	.363	4.0	.341	3.0	.340	2.0
PB.Lsa	.574	6.0	.653	7.0	.545	2.0	.561	4.0	.549	3.0	.567	5.0	.519	1.0
PB.Csa	.540	7.0	.491	1.0	.532	6.0	.495	2.0	.516	5.0	.507	3.0	.512	4.0
PB.L3	.574	6.0	.706	7.0	.545	2.0	.561	4.0	.549	3.0	.567	5.0	.518	1.0
Mean	.501	4.8	.674	6.3	.482	2.8	.503	4.4	.497	4.4	.486	3.3	.475	2.6
Mean Lsa	.485	4.0	.754	7.0	.462	2.0	.506	5.3	.483	4.0	.470	3.0	.467	2.7
Mean Csa	.534	6.0	.509	3.0	.521	4.7	.506	3.0	.520	5.3	.510	3.3	.497	2.7
Mean L3	.485	4.3	.760	7.0	.462	1.7	.497	5.0	.489	4.0	.479	3.7	.462	2.3

Rankings of standard error of estimate for calibration data sets[1]

	MEAN Rank	Unadjusted regional model Rank	LOC Rank	MAP-1F- Rank	MAP-R-P Rank	MAP-R-P + nV Rank	MAP-W Rank
Mean	5.3	7.0	1.3	4.9	3.2	2.4	3.8
Mean Lsa	5.3	7.0	1.0	4.7	2.7	3.0	4.3
Mean Csa	4.7	7.0	1.7	6.0	4.3	2.3	2.0
Mean L3	6.0	7.0	1.3	4.0	2.7	2.0	5.0

[1]Value ranked for unadjusted regional model is actually root mean square error for calibration data set.

Table 8. Effect of size of calibration data sets for model-adjustment procedures on root mean square errors for verification data sets taken from the Bellevue, Washington, data base

[Test results from split-sample analysis of varying calibration and verification data-set sizes; COD, chemical oxygen demand; TKN, total kjeldahl nitrogen; PB, total recoverable lead; SS, suspended solids; Lsa, step-analysis regression model for storm-runoff load; Csa, step-analysis regression model for storm-runoff mean concentration; L3, 3-variable regression model for storm-runoff load; CDS, calibration data set; MEAN, mean value of response variable from calibration data set used as an estimator; Unadjusted regional model, the appropriate single-storm model from Driver and Tasker (1990, tables 1, 3, and 5); LOC, local regression model based on total storm rainfall, drainage area, impervious area, and antecedent dry days; model-adjustment procedures (MAP's) defined in explanation in text; MAP-R-P+nV used antecedent dry days as additional explanatory variable; MAP-W used local regression model defined above in LOC; $RMSE_v$, root mean square error for verification data set, in log units]

Constituent and model type	CDS size	MEAN		Unadjusted regional model		LOC		MAP-1F-P		MAP-R-P		MAP-R-P+nV		MAP-W	
		$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank
SS.Lsa	21	0.618	7.0	0.507	6.0	0.440	5.0	0.375	1.0	0.385	2.0	0.408	3.0	0.418	4.0
COD.Lsa	21	.444	6.0	.452	7.0	.280	4.0	.251	1.0	.259	2.0	.267	3.0	.287	5.0
TKN.Lsa	21	.482	7.0	.341	6.0	.289	5.0	.251	1.0	.258	2.0	.266	3.0	.275	4.0
PB.Lsa	21	.550	7.0	.389	6.0	.361	5.0	.315	1.0	.316	2.0	.322	3.0	.337	4.0
Mean rank	21		6.8		6.3		4.8		1.0		2.0		3.0		4.3
SS.Lsa	31	.620	7.0	.511	6.0	.340	1.0	.363	2.0	.372	3.0	.384	4.0	.406	5.0
COD.Lsa	31	.435	6.0	.450	7.0	.268	4.0	.243	1.0	.252	2.0	.258	3.0	.278	5.0
TKN.Lsa	31	.477	7.0	.341	6.0	.263	5.0	.247	2.0	.246	1.0	.249	3.0	.261	4.0
PB.Lsa	31	.551	7.0	.392	6.0	.340	5.0	.321	1.5	.321	1.5	.327	4.0	.326	3.0
Mean rank	31		6.8		6.3		3.8		1.6		1.9		3.5		4.3
SS.Lsa	41	.618	7.0	.522	6.0	.398	4.0	.376	1.0	.381	2.0	.389	3.0	.407	5.0
COD.Lsa	41	.444	6.0	.455	7.0	.255	4.0	.242	1.0	.246	2.0	.248	3.0	.275	5.0
TKN.Lsa	41	.476	7.0	.346	6.0	.255	4.0	.246	2.5	.245	1.0	.246	2.5	.258	5.0
PB.Lsa	41	.545	7.0	.402	6.0	.328	5.0	.314	2.0	.312	1.0	.317	3.0	.327	4.0
Mean rank	41		6.8		6.3		4.3		1.6		1.5		2.9		4.8
SS.Lsa	51	.622	7.0	.519	6.0	.387	4.0	.366	1.0	.370	2.0	.377	3.0	.399	5.0
COD.Lsa	51	.431	6.0	.459	7.0	.246	4.0	.236	1.0	.240	2.0	.243	3.0	.267	5.0
TKN.Lsa	51	.489	7.0	.355	6.0	.257	4.0	.252	3.0	.249	1.0	.250	2.0	.262	5.0
PB.Lsa	51	.549	7.0	.404	6.0	.315	4.0	.310	3.0	.306	1.0	.307	2.0	.318	5.0
Mean rank	51		6.8		6.3		4.0		2.0		1.5		2.5		5.0

Table 8. Effect of size of calibration data sets for model-adjustment procedures on root mean square errors for verification data sets taken from the Bellevue, Washington, data base--Continued

Constituent and model type	CDS size	MEAN		Unadjusted regional model		LOC		MAP-1F-P		MAP-R-P		MAP-R-P+nV		MAP-W	
		$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank	$RMSE_v$	Rank
SS.Csa	21	0.352	4.0	0.448	7.0	0.363	6.0	0.340	1.0	0.348	3.0	0.347	2.0	0.358	5.0
COD.Csa	21	.254	6.0	.409	7.0	.210	1.0	.230	4.0	.236	5.0	.216	2.0	.219	3.0
TKN.Csa	21	.252	6.0	.298	7.0	.231	3.0	.235	4.0	.243	5.0	.220	1.0	.226	2.0
PB.Csa	21	.328	6.0	.359	7.0	.308	3.0	.314	4.0	.324	5.0	.290	1.0	.297	2.0
Mean rank	21		5.5		7.0		3.3		3.3		4.5		1.5		3.0
SS.Csa	31	.351	6.0	.443	7.0	.342	3.0	.338	2.0	.343	5.0	.336	1.0	.343	4.0
COD.Csa	31	.253	6.0	.410	7.0	.202	1.0	.229	4.0	.233	5.0	.211	2.0	.217	3.0
TKN.Csa	31	.247	6.0	.298	7.0	.215	2.0	.230	4.0	.235	5.0	.213	1.0	.216	3.0
PB.Csa	31	.320	6.0	.351	7.0	.296	3.0	.316	4.0	.324	5.0	.288	1.0	.295	2.0
Mean rank	31		6.0		7.0		2.3		3.5		5.0		1.3		3.0
SS.Csa	41	.344	6.0	.448	7.0	.328	2.0	.330	3.0	.334	4.0	.326	1.0	.342	5.0
COD.Csa	41	.261	6.0	.414	7.0	.198	1.0	.234	4.0	.239	5.0	.212	2.0	.216	3.0
TKN.Csa	41	.245	6.0	.295	7.0	.209	3.0	.228	4.0	.232	5.0	.209	2.0	.212	1.0
PB.Csa	41	.322	6.0	.361	7.0	.282	2.0	.308	4.0	.312	5.0	.277	1.0	.287	3.0
Mean rank	41		6.0		7.0		2.0		3.8		4.8		1.5		3.0
SS.Csa	51	.343	6.0	.443	7.0	.326	2.0	.331	3.0	.334	4.0	.323	1.0	.340	5.0
COD.Csa	51	.250	6.0	.410	7.0	.191	1.0	.227	4.0	.230	5.0	.204	2.0	.208	3.0
TKN.Csa	51	.240	6.0	.294	7.0	.200	1.0	.225	4.0	.228	5.0	.201	2.0	.208	3.0
PB.Csa	51	.331	6.0	.364	7.0	.293	2.0	.317	4.0	.320	5.0	.285	1.0	.294	3.0
Mean rank	51		6.0		7.0		1.5		3.8		4.8		1.5		3.5

For concentration models, the increase in $RMSE_v$ as CDS size decreased was also larger for the local regression than for the other procedures. As with the load models, the relative ranking among the other procedures remained the same (MAP-R-P+nV was the best-performing MAP at any CDS size), indicating relative insensitivity to CDS size. Performance of the local regression models, however, did prove to be sensitive to CDS size.

Estimating the Accuracy of Model-Adjustment Procedures

The accuracy of a model-adjustment procedure, and the relative accuracy of each MAP, will be different for each local data base (calibration data set). Three estimates of accuracy can be computed and compared among the MAP's for a given local data base. These indices are the coefficient of determination (r^2), the standard error of the estimate (*SE*), and the standard error of prediction (*SEP*). The r^2 and *SE* (defined and discussed earlier) are computed from the calibration data, and the *SEP* is computed when a prediction is prepared for an unmonitored site.

Although it may be assumed that the MAP with the smallest value of *SE* and largest value of r^2 will produce the greatest predictive accuracy for an unmonitored site, the results of the split-sample testing (tables 5, 6, and 7) illustrate that this interpretation should be made with caution. For most of the constituents tested, the MAP with the smallest value of *SE* (reported in the lower part of tables 5, 6, and 7) did not also provide the smallest value of $RMSE_v$. Exploratory data analysis of the calibration data set and application of the MAP selection procedure illustrated in figure 2 is probably a better guide to selection of an appropriate MAP than the relative magnitude of *SE*.

The SEP_i is a measure of the predictive accuracy of the MAP for a particular unmonitored site *i*. The SEP_i is computed as a function of the *SE* of the MAP as well as the difference between explanatory-variable values for the unmonitored site and the mean values of the calibration data set. The equations for computing SEP_i (in log units) for each MAP are presented in Supplement C. The SEP_i, in percent, can be calculated from SEP_i, in log units, using the same conversion factors presented in equation 16 for *SE*.

Calculation of confidence intervals also can help evaluate the accuracy of the procedures. A 100(1-α) confidence interval for the true value of the response variable (storm-runoff load or mean concentration) for an unmonitored site *i* and for a selected MAP can be computed by:

$$(\frac{1}{T})P_{ai}<Y_i<(T)P_{ai} \tag{18}$$

where

Y_i is true (but unknown) value of the response variable at unmonitored site i;
P_{ai} is predicted value at unmonitored site i, from the adjusted model; and

T is calculated as follows:

$$\log T=t_{(\frac{\alpha}{2},n-p)}*SEP_i,$$

where

$t_{(\alpha/2,n-p)}$ is critical value of the t-distribution for n-p degrees of freedom;
n is number of observations in the calibration data set;
p is number of explanatory variables plus 1; and
SEP_i is expressed in log units.

Example Application

The following example illustrates the estimation of storm-runoff load for an unmonitored site and a single storm, using the four MAP's with a local data base consisting of 18 storms from five sites. A city engineer from City X would like to estimate a storm-runoff load for COD for any size storm and at any unmonitored site *i* in that city. Using the COD load model (Lsa) for region II (Driver and Tasker, 1990, table 1) and the determined values for the explanatory variables for that model (TRN; DA; industrial land use, LUI; commercial land use, LUC; nonurban land use, LUN; and mean annual rainfall, MAR), the engineer calculates a value for storm-runoff load (P_u) to correspond with each monitored storm in the local data base. The candidate basin- and storm-characteristic variables to be used as additional explanatory variables (for calibrating MAP-R-P+nV) and in local regression models (for calibrating MAP-W) are also evaluated. The hypothetical calibration data set is now assembled for City X (table 9). The engineer then follows the EDA and MAP selection scheme prescribed in figure 2. The root mean square error is 0.453 in log units, or 130 percent. The city engineer decides this is unacceptably large, and proceeds to evaluate the

MAP approach. P_u is significantly and positively correlated with O (r_s is 0.887) and biased in a consistent direction relative to O (p-value for the signed-rank test is less than 0.0001), suggesting that either MAP-1F-P or MAP-R-P would be appropriate for the COD load model for City X. Because of the small data set size (n=18), the engineer selects MAP-1F-P. Coefficients for MAP-1F-P are then determined by performing a set of regression calculations on the calibration data set such as those listed in Supplement B; the results for City X are listed in table 9, along with the results for the other MAP's.

Table 9. Sample of calibration data set and values for standard errors of estimate, bias-correction factors, and coefficients for the model-adjustment procedures for City X

[P_u, predicted load from unadjusted regional model; O, observed load; TRN, total storm rainfall; DA, total contributing drainage area; IA, impervious area; ADD, antecedent dry days; SE, standard error of estimate; BCF, bias-correction factor; β_0, β_1, β_2, β_3, β_4, j, coefficients for the MAP's; MAP-1F-P, single-factor regression against regional prediction; MAP-R-P, regression against regional prediction; MAP-R-P+nV, regression against regional prediction and local data; MAP-W, weighted combination of regional prediction and local-regression prediction; LOC, local regression model; –, additional data not shown]

P_u, in pounds	O, in pounds	TRN, in inches	DA, in square miles	IA, in percent	ADD, in days
578	360	1.45	0.15	36.1	6
87	29	.15	.15	36.1	5
285	120	.62	.15	36.1	5
122	26	.56	.04	56.5	4
142	41	.56	.04	56.5	2
–	–	–	–	–	–
–	–	–	–	–	–
–	–	–	–	–	–

Model-adjustment procedure	SE, log	BCF	β_0	β_1	β_2	β_3	β_4	j
MAP-1F-P	0.235	1.25	-0.328					
MAP-R-P	.233	1.14	-.397	1.02				
MAP-R-P+nV	.229	1.24	-.684	1.14	0.065			
MAP-W	.237	1.43						0.262
LOC	.224	1.12	4.44	.914	.265	-1.26	0.029	

The city engineer is now interested in estimating storm-runoff load for COD for a particular unmonitored site i (DA = 0.15 square miles, IA = 40 percent, LUI = 5 percent, LUC = 40 percent, LUN = 20 percent) for a particular storm of 0.2 inch rainfall (TRN = 0.2 inch) that followed 5 days of no rainfall (ADD = 5 days). The mean annual rainfall for City X is 25 inches (MAR = 25 inches). The engineer first calculates the value for unmonitored site i predicted from the unadjusted regional model (P_{ui}):

$$P_{ui}(COD)=36.6*(0.2)^{(.878)}*(0.15)^{(.696)}*(5+1)^{(.072)}*(40+1)^{(.261)}$$

$$*(20+2)^{(-.056)}*(25)^{(.866)}*1.389;$$

$$P_{ui}(COD) = 136 \textit{ pounds.}$$

Employing MAP-1F-P, P_{ui} is adjusted to P_{ai} using equation 3, and using the values listed for β_0 and BCF in table 9:

$$P_{ai} = 10^{(-0.328)} * 136 * 1.25 = 80 \textit{ pounds.}$$

The SEP (in log units) for unmonitored site and storm i for MAP-1F-P is computed using equation A in Supplement C, and using the value for $SE_{1F\text{-}P}$ listed in table 9, as:

$$SEP_i = \sqrt{0.235^2(1 + \frac{1}{18})} = 0.242.$$

The value for SEP_i expressed in percent units is 60.

The 95-percent confidence interval for the prediction is calculated as follows. The critical value for the t distribution for (18-2=16) degrees of freedom and $\alpha/2=0.025$ is determined (from a standard statistical table) to be 2.12. Then

$$T=10^{(2.12*0.242)}=3.26.$$

The values for the lower and upper bounds of the 95-percent confidence interval (L_{95} and U_{95}, respectively) are therefore

$$L_{95}=\frac{1}{3.26}*80 = 25 \textit{ pounds,}$$

$$U_{95} = 3.26*80 = 260 \textit{ pounds.}$$

A MINITAB program for calculating P_{ai} for each MAP is given in Supplement D.

Prediction of Annual or Seasonal Urban-Runoff Quality

A prediction of annual or seasonal urban-runoff load at an unmonitored site i can be obtained by applying the procedure described in the preceding example to a series of storms and producing a synthetic record of storm loads. Values of storm characteristics that are used as explanatory variables (for example, TRN; ADD; duration of each storm, DRN; maximum intensity during a 15-minute period, MI15) may be determined for the series of storms from the long-term rainfall record for a station near the unmonitored site. The synthesized record of storm loads may be reduced to an estimate of mean annual load by summing loads from each storm, then dividing by the number of years in the period of the synthetic record. Reduction to an estimate of mean seasonal load may be accomplished by summing loads only from the season of interest before dividing by the number of years of record.

SUMMARY

Water-quality management and load allocations from point and nonpoint sources in urban areas require city engineers, planners, and designers to estimate loads and mean concentrations of constituents in storm

runoff. Although many deterministic and statistical models of urban-runoff quality are available, these models were calibrated using either national, regional, or local data bases for only a few selected cities. When the city engineer can assemble data on urban-runoff quality from a local monitoring network, he may wish to adjust the '*a priori*' prediction from the model with local data. This report presents four statistical procedures, MAP's, by which the predictions of urban-runoff quality from existing regression models can be combined or weighted with information from local data.

Each MAP is a form of regression analysis, in which the local data base is used as a calibration data set. Regression coefficients are determined from the local data, and the resulting 'adjusted' regression models can then be used to predict storm-runoff quality at unmonitored sites. The response variable in the regression analyses is the observed load or mean concentration of a constituent in storm runoff for a single storm. The set of explanatory variables used in the regression analyses is different for each MAP, but always includes the predicted value of load or mean concentration from the regional single-storm models developed by Driver and Tasker (1990, tables 1, 3, and 5).

The MAP's were tested by means of split-sample analysis, using data from three cities included in the Nationwide Urban Runoff Program: Denver, Colorado; Bellevue, Washington; and Knoxville, Tennessee. The MAP that provided the greatest predictive accuracy for the verification data set differed among the three test data bases and among model types (MAP-W for Denver and Knoxville, MAP-1F-P and MAP-R-P for Bellevue load models, and MAP-R-P+nV for Bellevue concentration models) and, in many cases, was not clearly indicated by the values of *SE* for the calibration data set. This does not mean, however, that it is impossible for the analyst working without a verification data set to anticipate which MAP will provide the greatest predictive accuracy for an unmonitored site. A scheme to guide MAP selection based on exploratory data analysis of the calibration data set is presented and tested. When O and P_u in the calibration data set are not strongly correlated (as for Bellevue concentration models and for Knoxville models), or when the direction of bias between O and P_u is not consistent (as for Denver models), the MAP's based on P_u alone (MAP-1F-P and MAP-R-P) should be rejected in favor of either MAP-R-P+nV or MAP-W. If, however, correlation between response variable and any of the explanatory variables used in MAP-R-P+nV or MAP-W is not strong (as for Knoxville), then these MAP's cannot be expected to provide better predictive accuracy than a simple estimator such as mean value of the response variable in the calibration data set. When O and P_u in the calibration data set are strongly correlated and related according to a consistent direction of bias (as for Bellevue load models), then MAP-1F-P and MAP-R-P are the most reliable procedures.

The MAP's were tested for sensitivity to the size of a calibration data set. As expected, predictive accuracy of all MAP's for the verification data set decreased as the calibration data-set size decreased, but their performance was not as sensitive as for the local regression models.

REFERENCES CITED

Draper, N.R., and Smith, H., 1981, Applied regression analysis: New York, John Wiley and Sons, Inc., 709 p.

Driver, N.E., and Tasker, G.D., 1990, Techniques for estimation of storm-runoff loads, volumes, and selected constituent concentrations in urban watersheds in the United States: U.S. Geological Survey Water-Supply Paper 2363, 44 p.

Duan, Naihua, 1983, Smearing estimate: A nonparametric retransformation method: Journal of the American Statistical Association, v. 78, no. 383, p. 605-610.

Hoos, A.B., 1991, Storm-runoff quality and its relation to land use in the Nashville metropolitan area, Tennessee, *in* Jennings, M.E., ed., Symposium proceedings on urban hydrology: Bethesda, Maryland, American Water Resources Association Technical Publication Series TPS-91-4, p. 33.

Huber, W.C., 1986, Deterministic modeling of urban runoff quality, *in* Torno, H.C., Marsalek, Jiri, and Desbordes, Michael, eds., Urban runoff pollution, a NATO advanced workshop, Montpelier, France, 1985, Proceedings: Berlin, Springer-Verlag, NATO ASI Series G, Ecological Sciences, v. 10, p. 167-242.

Iman, R.L., and Conover, W.J., 1983, A modern approach to statistics: New York, John Wiley & Sons, Inc., 497 p.

Miller, D.M., 1984, Reducing transformation bias in curve fitting: The American Statistician, v. 38, no. 2, p. 124-126.

Mustard, M.H., Driver, N.E., Chyr, John, and Hansen, B.G., 1987, U.S. Geological Survey urban-stormwater data base of constituent storm loads: characteristics of rainfall, runoff, and antecedent conditions; and basin characteristics: U.S. Geological Survey Water-Resources Investigations Report 87-4036, 327 p.

Nix, S.J., 1991, Applying urban runoff models: Water Environment and Technology, v. 3, no. 6, p. 47-49.

Schueler, T.R., 1987, Controlling urban runoff: A practical manual for planning and designing urban BMPs: Washington Metropolitan Water Resources Planning Board, publication no. 87703, various pagination.

U.S. Environmental Protection Agency, 1983, Results of the nationwide urban runoff program: U.S. Environmental Protection Agency Final Report, 4 volumes.

___ 1990, National pollutant discharge elimination system permit application regulations for storm water discharges: U.S. Federal Register, v. 55, no. 222, p. 47990-48091.

Young, G.K., Bondelid, T.R., and Athayde, D.N., 1979, Urban runoff pollution method: Journal of the Water Resources Planning and Management Division, v. 105, no. WR2, Paper 14854, p. 353-369.

SUPPLEMENTAL INFORMATION

Supplement A. Program (MINITAB) of exploratory data analysis procedures applied to calibration data set to guide selection of model-adjustment procedures

```
#     'EDA.MTB'    MACRO
#     ('Exploratory Data Anlysis')
#
#     This macro performs several tests (exploratory data analysis procedures) on the calibration data set (a
#     local data base) to determine which MAP will provide the highest prediction accuracy for an
#     unmonitored site or storm in that city.
#
#     Input data for this macro are:
#
#     C1  -  value for prediction from the unadjusted regional model for a particular site and storm
#                (Pu), in real (not log-transformed) units
#     C2  -  observed value for that site and storm (O), in real (not log-transformed) units
#     C3  -  order number for site/storm (for bookkeeping purposes)
#
#     The next five variables are those to be tested (using a best-regression analysis) for inclusion as
#     explanatory variables in a local 5-variable regression model. The local regression model is then used as
#     part of the MAP-W procedure.  The variables are also tested for inclusion in the MAP-R-P+nV
#     procedure.
#
#     C4  -  total rainfall (in.)
#     C5  -  drainage area (acres)
#     C6  -  any explanatory variable, in real units
#     C7  -  any explanatory variable, in real units
#     C8  -  any explanatory variable, in real units
#
#     Log-transform all variables
#
LET C11 = LOGTEN(C1)
LET C12 = LOGTEN(C2)
LET C4 = LOGTEN(C4)
LET C5 = LOGTEN(C5)
LET C6 = LOGTEN(C6)
LET C7 = LOGTEN(C7)
LET C8 = LOGTEN(C8)
#
PLOT C11 C12
#
#     Calculate root mean square error (log units, K1), from applying the unadjusted regional model to the
#     calibration data set.  If RMSE is acceptably small, the analyst may wish to use the regional model
#     without any adjustment.  (Respond 'Yes' for 'Prediction Error of Pu Small?' in flowchart.)
#
LET C21 = (C12 - C11) ** 2
LET K1 = SUM(C21)/N(C21)
LET K1 = SQRT(K1)
PRINT K1
#
```

Supplement A. Program (MINITAB) of exploratory data analysis procedures applied to calibration data set to guide selection of model-adjustment procedures--Continued

```
#     Check to see if regional model captures relative variability among the observations:  calculate and test
#     Spearman's rho, rS.  Compare the result (the value for the correlation printed below) against T* listed
#     for selected alpha level (see, e.g., figure 11.9 of Iman and Conover):  if Spearman's rho is greater than
#     the listed T* for a given n, then respond 'Yes' for 'O and Pu Significantly and Positively Correlated?' in
#     flowchart.
#
RANK C11 C9
RANK C12 C10
CORRELATION C9 C10
#
#     Now test whether predictions (Pu) are consistently biased relative to observed values (O).  If so,
#     this would indicate the appropriateness of using the predicted value as the single explanatory variable in
#     the adjusted model.  Use the signed rank test (paired data) to test for bias.  If p-values are smaller than a
#     selected alpha, then respond 'Yes' for 'Consistent Direction of Bias?' in flowchart.
#
LET C15 = C12 - C11
STEST 0 C15
#
#     Check correlation between response variable (O) and each of the local explanatory variables.  If one or
#     more of the candidate explanatory variables are significantly correlated with the response variables,
#     respond 'Yes' for 'O and Other Explanatory Variables Significantly Correlated?' in flowchart.
#
CORRELATE C12 C4
CORRELATE C12 C5
CORRELATE C12 C6
CORRELATE C12 C7
CORRELATE C12 C8
#
#     First best regression test.  Check for best regression model from list of combinations of explanatory
#     variables.  Select from among all models with Cp ≤ p or high adjusted r2 values.  Make final selection
#     in favor of the simplest model with physically logical parameter values.  This model would then be used
#     in MAP-W.  If the local regression is to be used alone (independent of MAP-W) then it should include
#     total rainfall and drainage area, as a minimum.
#
BREG C12 C4 C5 C6 C7 C8;
INCLUDE C4-C5;
BEST 5.
#
#     Second best regression test.  The results of the following best regression should be used in determining
#     which variables should be used in the MAP-R-P+nV method.  Variables that are dropped from the
#     equation should not be used.
#
BREG C12 6 C11 C4 C5 C6 C7 C8;
INCLUDE C11;
BEST 5.
END
```

Supplement B. Program (MINITAB) of statistical procedures applied to calibration data set to derive coefficients for model-adjustment procedures

```
#     'CALIBRATE.MTB'    MACRO
#
#     This macro uses the local data base (calibration data set) to derive coefficients for each
#     model-adjustment procedure (MAP).  Although the user may have selected one MAP as a result of
#     exploratory data analysis of the calibration data set, this macro includes all procedures.
#
#     IMPORTANT!!!!!!
#     In this macro:
#     for MAP-R-P+nV, n=5
#     for MAP-W, the local regression is a 5-variable model
#     The user must revise the number of variables used if so indicated by the EDA.MTB results.
#
#     Input data for this macro are:
#
#     C1 -  value for prediction from the unadjusted regional model for a particular site and storm (Pu),
#                in real (not log-transformed) units
#     C2 -  observed value for that site and storm (O), in real (not log-transformed) units
#     C3 -  order number for site/storm (for bookkeeping purposes)
#
#     The local explanatory variables, chosen from using the best regression EDA.MTB results, are used in
#     the MAP-R-P+nV and MAP-W procedures.  This macro is written to use five variables, as listed below.
#
#     C4 -  total rainfall (in.)
#     C5 -  drainage area (acres)
#     C6 -  any explanatory variable, in real units
#     C7 -  any explanatory variable, in real units
#     C8 -  any explanatory variable, in real units
#
#     WARNING!!!  Do not attempt to use the data matrix that may be stored in the MINITAB worksheet as a
#     result of a preceding execution, during the current MINITAB session, of EDA.MTB.  The values input
#     for C1-C8 must be in real units.
#
#     K51  -  value for SE, in log units, for the regional regressions.  Taken from WSP 2363, table 2 (for Lsa
#                 models), table 6 (for Csa models) and table 3 (for L3 models)
#
#     Log-transform all variables
#
LET C11 = LOGTEN(C1)
LET C12 = LOGTEN(C2)
LET C4 = LOGTEN(C4)
LET C5 = LOGTEN(C5)
LET C6 = LOGTEN(C6)
LET C7 = LOGTEN(C7)
LET C8 = LOGTEN(C8)
#
NAME C52 'LOC', C53 'MAP-1F-P', C54 'MAP-R-P'
```

Supplement B. Program (MINITAB) of statistical procedures applied to calibration data set to derive coefficients for model-adjustment procedures--Continued

```
NAME C55 'MAP-R-P+', C56 'MAP-W'
#
#    Procedure 1.  MAP-1F-P
#    The no-exponent fitting of observed values against predicted values (recommended by Tasker and
#    Cohn, September 90).  Calculate Bo (K2), SE (K3), and BCF (K4) and store results in C53.
#
LET K2 = MEAN(C12) - MEAN(C11)
LET C53(1) = K2
LET C16 = C11 + K2
LET C17 = (C16 - C12)
LET K3 = SUM(C17 ** 2)/(N(C17)-2)
LET K3 = SQRT(K3)
LET K4 = SUM(10**(C17))/N(C17)
LET C53(10) = K3
LET C53(11) = K4
#
#    Procedure 2.  MAP-R-P
#    Straight regression of observed values against predicted values (recommended by Will Thomas,
#    October 91).  Store results in C54 for coefficients,  SE (K12), and BCF (K13).
#
REGRESS C12 1 C11;
COEFFICIENTS C54;
RESID C15;
MSE K12.
LET K12 = SQRT(K12)
LET K13 = SUM(10**(C15))/N(C15)
LET C54(10) = K12
LET C54(11) = K13
#
#    Procedure 3.  MAP-R-P+nV
#    Straight regression of observed values against predicted values and additional independent variables.
#    Store results in C55 for coefficients, SE (K16), and BCF (K17).
#
REGRESS C12 6 C11 C4 C5 C6 C7 C8;
COEFFICIENTS C55;
RESID C28;
MSE K16.
LET K16 = SQRT(K16)
LET K17 = SUM(10**(C28))/N(C28)
LET C55(10) = K16
LET C55(11) = K17
#
#    Procedure 4.  MAP-W
#    Weighting of prediction from unadjusted regional model with prediction from a local regression.
#
```

Supplement B. Program (MINITAB) of statistical procedures applied to calibration data set to derive coefficients for model-adjustment procedures--Continued

```
#     First, fit coefficients for the local (5-variable) regression model and store results in C52 for coefficients,
#     SE (K6), and BCF (K7).
#
REGRESS C12 5 C4 C5 C6 C7 C8 C99 C30;
COEFFICIENTS C52;
RESID C20;
MSE K6.
LET K6 = SQRT(K6)
LET K7 = SUM(10**(C20))/N(C20)
LET C52(10) = K6
LET C52(11) = K7
#
#     Next, compute and store results in C56 for the weighting factor 'j',  SE (K18), and BCF (K19).
#
LET C56(1) = C52(10)**2/(C52(10)**2+K51**2)
LET C23 = C56(1)*C11+(1-C56(1))*C30
LET C24 = (C23 - C12)
LET K18 = SUM(C24**2)/(N(C24)-2)
LET K18 = SQRT(K18)
LET K19 = SUM(10**(C24))/N(C24)
LET C56(10) = K18
LET C56(11) = K19
#
#     Printout results
#
PRINT C52-C56
WRITE 'COEFF.DAT' C52-C56
END
```

Supplement C. Formulas for standard error of prediction for model-adjustment procedures

MAP-1F-P

$$SEP_i = \sqrt{SE^2_{1F\text{-}P}\left(1+\frac{1}{n}\right)}, \tag{A}$$

where

SEP_i	is	standard error of prediction for unmonitored site i;
$SE_{1F\text{-}P}$	is	standard error of estimate (in log units) for the calibration of equation 1; and
n	is	number of observations in the calibration data set.

MAP-R-P

$$SEP_i = \sqrt{SE^2_{R\text{-}P}\left(1+u_i(U'U)^{-1}u_i'\right)}, \tag{B}$$

where

$SE_{R\text{-}P}$	is	standard error of estimate (in log units) for the calibration of equation 5;
u_i	is	a (1 x 2) row vector containing 1 as the first element, and the value for the single explanatory variable, P_u, evaluated (in log units) for unmonitored site i, augmented by a 1 as the first element; and
U	is	a (n x 2) matrix containing 1 as the first column, and the values for the single explanatory variable, P_u, evaluated (in log units) for all n sites in the R-P calibration set, in the second column.

MAP-R-P+nV

$$SEP_i = \sqrt{SE^2_{R\text{-}P+nV}\left(1+y_i(Y'Y)^{-1}y_i'\right)}, \tag{C}$$

where

$SE_{R\text{-}P+nV}$	is	standard error of estimate (in log units) for the calibration of equation 7;
y_i	is	a (1 x j) row vector of the j-1 explanatory variables (the variable P_u and the j-2 additional explanatory variables) used in the R-P+nV regression, evaluated (in log units) for unmonitored site i, augmented by a 1 as the first element; and
Y	is	a (n x j) matrix of the j-1 explanatory variables used in the local regression, evaluated (in log units) for all n sites in the R-P+nV calibration data set, augmented by a 1 as the first column.

Supplement C. Formulas for standard error of prediction for model-adjustment procedures--Continued

MAP-W

$$SEP_i = \sqrt{\frac{V_{pi\text{-}loc} * V_{pi\text{-}u}}{V_{pi\text{-}loc} + V_{pi\text{-}u}}}. \qquad \text{(D)}$$

where

$V_{pi\text{-}loc}$ and $V_{pi\text{-}u}$ are as defined in equations 11 and 12.

Supplement D. Program (MINITAB) applied to data from an unmonitored site to calculate the prediction using model-adjustment procedures

```
#     'PREDICT.MTB'    MACRO
#
#     This macro computes a predicted value for an unmonitored site/storm(s) using each MAP.  To do
#     this, it uses the output file generated from CALIBRATE.MTB, which contains coefficients
#     (determined using the local database) for each MAP.  This macro, like CALIBRATE.MTB, is
#     written for the inclusion of all five additional variables in MAP-R-P+nV (n=5) and use
#     of all five variables in the local regression used in MAP-W.  THE USER MUST CHANGE THE
#     FORMULAS IF EDA.MTB AND CALIBRATE.MTB SO INDICATE!!!
#
#     Input data for this macro are:
#
#     The output file from CALIBRATE.MTB, which is read into C51-C56 automatically if user does not
#     exit MINITAB.
#
#     C1  -  predicted value for unmonitored site/store from Driver-Tasker equations and reported
#                in real (not log-transformed) units
#     C3  -  order number for site/storm (for bookkeeping purposes)
#     C4  -  total rainfall (in.)
#     C5  -  drainage area (acres)
#     C6  -  any explanatory variable, in real units
#     C7  -  any explanatory variable, in real units
#     C8  -  any explanatory variable, in real units
#
NAME C30 'Pa-LOC', C36 'Pa-1F-P', C39 'Pa-R-P', C42 'Pa-R-P+'
NAME C45 'Pa-W'
#
#     Compute a predicted value using the MAP-1F-P procedure (the B1-forced-to-unity fit of observed
#     against predicted).
#
LET C36 = 10**(C53(1))*C1*C53(11)
#
#     Compute a predicted value using the MAP-R-P procedure (the 'regular' regression of observed
#     against predicted).
#
LET C39 = 10**(C54(1))*C1**(C54(2))*C54(11)
#
#     Compute a predicted value using the MAP-R-P+nV procedure (regression of observed against
#     predicted value and five explanatory variables)
#
LET C42 = 10**(C55(1))*C1**(C55(2))*C4**(C55(3))*C5**(C55(4))*C6**(C55(5))&
*C7**(C55(6))*C8**(C55(7))*C55(11)
#
#     Compute a predicted value using the MAP-W procedure.  First, compute a predicted value using
#     coefficients (derived from the calibration dataset) for the 5-variable regression model based on local
#     data alone.
```

Supplement D. Program (MINITAB) applied to data from an unmonitored site to calculate the prediction using model-adjustment procedures--Continued

```
#
LET C30 = 10**(C52(1))*C4**(C52(2))*C5**(C52(3))
LET C30 = C30*C6**(C52(4))*C7**(C52(5))*C8**(C52(6))*C52(11)
#
#     Now apply the MAP-W prediction equation:
#
LET C45 = C1**(C56(1))*C30**(1-C56(1))*C56(11)
#
#     Print results
#
PRINT C30,C36,C39,C42,C45
END
```