Consequences of Land-cover Misclassification in Models of Impervious Surface

Gerard McMahon

Abstract

Model estimates of impervious area as a function of land-cover area may be biased and imprecise because of errors in the land-cover classification. This investigation of the effects of land-cover misclassification on impervious surface models that use National Land Cover Data (NLCD) evaluates the consequences of adjusting land-cover within a watershed to reflect uncertainty assessment information. Model validation results indicate that using error-matrix information to adjust land-cover values used in impervious surface models does not substantially improve impervious surface predictions. Validation results indicate that the resolution of the land-cover data (Level I and Level II) is more important in predicting impervious surface accurately than whether the land-cover data have been adjusted using information in the error matrix. Level I NLCD, adjusted for land-cover misclassification, is preferable to the other land-cover options for use in models of impervious surface. This result is tied to the lower classification error rates for the Level I NLCD.

Introduction

Urbanization often results in land being converted from a condition that is permeable by water to one that is impermeable, or impervious, to water. An increase in impervious surface, in turn, adversely affects terrestrial and aquatic environments. Negative effects may begin to occur at levels as low as 10 percent imperviousness overall in the watershed (Schueler, 1994; Arnold and Gibbons, 1996; Barnes *et al.*, 2000; American Planning Association, 2002). A report on the effects of impervious surfaces in watersheds of the Chesapeake Bay (Barnes *et al.*, 2000) indicates that development, in general, and impervious surface, in particular, affect water quantity (e.g., decrease in depression storage (Novotny and Chesters, 1981)), stream flashiness (McMahon *et al.*, 2003), water quality (e.g., conventional (Frenzel and Couvillion, 2002) and toxic (Johnson *et al.*, 2000) pollutants), stream biota (Kennen, 1999; Coles *et al.*, 2004), stream habitat (Booth and Jackson, 1997; Booth *et al.*, 2002), and the physical characteristics of streams (Dunne and Leopold, 1978; U.S. Environmental Protection Agency, 1997).

Investigating the relation among basin characteristics, including imperviousness and the physical, chemical, and biological effects in streams requires quantitative information about basin characteristics and environmental responses that can be linked to these characteristics (McMahon and Cuffney, 2000). Several protocols have been developed and widely implemented for characterizing stream response to human activities (e.g., stream chemistry (Shelton, 1994), stream biota (Barbour *et al.*, 1999; Moulton *et al.*, 2002), and stream habitat (Fitzpatrick *et al.*, 1998)). These data commonly are collected at a number of scales as part of local (Krug and Goddard, 1986), state (Davis and Simon, 1995; Yoder, 1995) and federal (U.S. Geological Survey, 2002c) water quality investigations. The widespread availability of relatively large-scale, mapped digital databases has simplified the development of extensive tables of basin characteristics that can be used in studying the association between terrestrial and aquatic environments (Van Sickle, 2003).

A direct study of the effects of impervious surface on water quality, however, is difficult because accurate measurement or estimation of impervious surface is time consuming and costly (Stankowski, 1972; Prisloe *et al.*, 2000; Slonecker and Tilley, 2004). The most accurate measures of imperviousness can be obtained from ground-based measurements or by digitizing impervious areas from very large-scale planimetric tax maps (Center for Watershed Protection, 1998). Relatively accurate impervious surface measurements also can be made by interpreting impervious surface area from aerial photographs or digital orthophotos and digitizing the information into a geographic information system (Sleavin, 1999). Capturing impervious surface data using these methods, however, is labor intensive and impractical for investigations where a study may have a cumulative area of hundreds or thousands of square kilometers.

Estimates of imperviousness in a watershed also can be developed by using models that estimate impervious surface as a function of other spatial data available for an area of interest. Complex models may be used to estimate impervious surface from satellite remote-sensing data. The Landsat Thematic Mapper data are being used at the U.S. Geological Survey's (USGS) Earth Resources Observation Systems (EROS) Data Center to quantify impervious surfaces as a continuous variable using multi-sensor and multi-source data in a regression tree model (Yang *et al.*, 2003). Civco and Hurd (1997) describe the use of high-resolution spatial data, such as digital orthophotos, to train neural network models, which are used with remotely sensed data to estimate impervious surface information for large areas. Perhaps the most important drawback in the use of satellite remote-sensing-based impervious surface estimates is the necessity for a great deal of expertise to develop and evaluate these frameworks.

A simpler, model-based approach for estimating impervious surface involves the use of land-cover data in conjunction with impervious surface coefficients (Schueler, 1994;

U.S. Geological Survey, 3916 Sunset Ridge Road, Raleigh, NC 27607 (gmcmahon@usgs.gov).

Photogrammetric Engineering & Remote Sensing
Vol. 73, No. 12, December 2007, pp. 1343–1353.

0099-1112/07/7312–1343/$3.00/0

Center for Watershed Protection, 1998; Sleavin, 1999; Jennings *et al*., 2004). For example, if there are three land-cover classes represented in a watershed, the proportion of a watershed covered by impervious surface is estimated as:

$$IS = (IS_1 * Area_1 + IS_2 * Area_2 + IS_3 * Area_3)/(Area_1 + Area_2 + Area_3) \quad (1)$$

where IS = impervious surface of a drainage basin (proportion of total watershed area); IS_i = impervious surface coefficient for land-cover class i (proportion of land-cover area for land-cover type i that is impervious; domain ranges from 0–1); and $Area_i$ = area of land-cover class i in the watershed. Coefficients specific to each of the land-cover classes in the spatial dataset (e.g., high-density urban or deciduous forest) may be obtained from the literature and applied to the area of each land-cover type in each study drainage basin to estimate the percentage of the basin area that is impervious.

If land-cover data are available for a study area, a coefficient-based impervious surface modeling approach is relatively straightforward and quick to implement. Impervious surface estimates using this approach, however, may be biased and imprecise for several reasons: (a) a review of the literature suggests a high degree of variability in impervious surface coefficients associated with land-cover classes; (b) land-cover specific coefficients that account for the physical and hydroclimatic conditions of a study area of interest may not be available; (c) it may not be possible to identify coefficients that match the land-cover categories in a particular land-cover classification; and (d) the land-cover classification may contain errors. For example, an accuracy assessment of the NLCD completed for the eastern United States indicates that the user accuracy of the general (Level I) developed land-cover class is 74 percent, and the user accuracy of the three more detailed (Level II) developed land-cover classes is 61 percent for low-density residential land-cover, 40 percent for high-density residential land-cover, and 48 percent for commercial/industrial/transportation land-cover (Zhu *et al*., 2000; Vogelmann *et al*., 2001; Yang *et al*., 2001; Stehman *et al*., 2003; Stehman *et al*., 2003; Wickham *et al*., 2003). These relatively low classification accuracy rates indicate a substantial amount of confusion, or misclassification, between the developed and other land-cover classes. Such misclassifications introduce error into models that use developed land-cover area estimates, including models of impervious surface. Because adverse effects of imperviousness can occur when a relatively small proportion of watershed area is impervious, even small biases or inaccuracies in estimating watershed impervious surface can limit the ability to anticipate and manage impervious-related effects.

The overall goal of the investigation reported here is to evaluate the effects of land-cover misclassification on models that use NLCD to predict impervious surface. The context for the study is an investigation in coastal New England of the relation between factors associated with urbanization, including impervious surface, and the physical, chemical, and biological aspects of water quality (Coles *et al*., 2004). The New England study, conducted under the auspices of the USGS National Water-Quality Assessment (NAWQA) Program, examined the effects of a gradient of urban-development intensity on stream ecology in 32 watersheds in the Gulf of Maine ecoregion in southern New England (Figure 1; McMahon and Cuffney, 2000; U.S. Geological Survey, 2003b). Four impervious surface models, described below and based on Equation 1, were used to assess a procedure for adjusting the NLCD to account for misclassification and determine the relative importance of the NLCD classes in predicting impervious surface area.

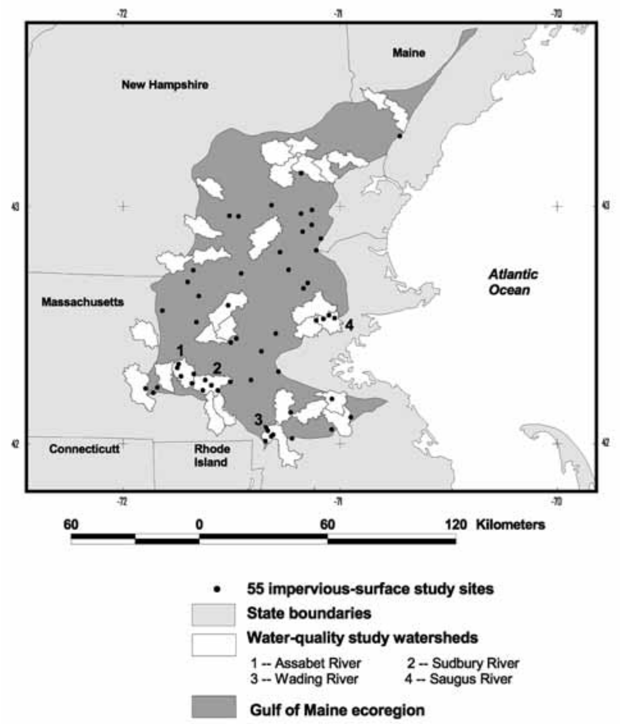

Figure 1. New England impervious surface study area.

The evaluation focuses on three questions:

1. What are the results of adjusting land-cover area in a watershed to reflect accuracy-assessment information?
2. Does explicit adjustment of land-cover values used in impervious surface models to account for land-cover classification error rates result in meaningful improvements of impervious surface predictions?
3. Do modeled impervious surface values have a stronger association with measures of stream ecological conditions than other measures of urbanization?

Methods

Four coefficient-based models for predicting impervious surface area were evaluated against measured impervious surface values at 55 study sites to assess the effect of accounting for classification error. The models, based on the approach of Equation 1, use adjusted and unadjusted land-cover for both Level I and Level II NLCD and literature review-derived impervious surface (Table 1). Models were used to predict impervious surface in the 32 New England study watersheds, and the resulting impervious surface estimates were compared with stream ecological conditions.

Measuring Impervious Surface Data

All models were assessed using impervious surface data measured from digital orthophotos (1-meter resolution) at 55 study sites, each having an area of 0.5 km × 0.5 km (Figure 1). These 55 study sites were selected to meet several study goals. In addition to being used to develop the impervious surface models, digital orthophoto data were needed for training purposes in 4 of the 32 study watersheds (Assabet, Sudbury, Wading, and Saugus Rivers shown in Figure 1) where satellite data were used in a separate

Table 1. General (Level I) and Detailed (Level II) Land-cover Classes in the Gulf of Maine Ecoregion from National Land-Cover Data (Vogelmann *et al.*, 2001)

	Impervious-Surface Coefficients (based on literature review)		Land Area (%) in the Gulf of Maine Ecoregion
	Mean	Standard Deviation	
Level I land-cover			
Developed	0.41	0.25	21
Barren	0.21	0.19	<1
Forest	0.02	0.02	59
Shrubland	0.02	0.01	<1
Agricultural	0.05	0.03	8
Wetland	0.02	0.02	7
Level II land-cover			
Low density residential	0.16	0.07	15
High density residential	0.39	0.15	2
comm./industrial	0.60	0.26	5
Bare rock	0.21	0.19	<1
Quarries	0.21	0.19	<1
Transitional	0.21	0.19	<1
Deciduous forest	0.03	0.02	29
Evergreen forest	0.02	0.01	11
Mixed forest	0.05	0.11	19
Deciduous shrub	0.02	0.01	<1
Pasture	0.05	0.03	2
Row crops	0.04	0.04	5
Urban/recreational grass	0.05	0.03	2
Woody wetland	0.02	0.01	5
Herbaceous wetland	0.02	0.01	2

(Notes: Impervious surface values for wetlands are assumed equal to literature values for forest (Level I) and deciduous forests (Level II). Literature review includes Stankowski, 1972; Dunne and Leopold, 1978; Northern Virginia Planning District Commission, 1980; U.S. Department of Agriculture, 1986; Bedient and Huber, 1988; Aqua Terra Consultants, 1994; Kluitneberg, 1994; Schueler, 1994; City of Olympia Public Works Department, 1995; Center for Watershed Protection, 1998.)

investigation of several image classification approaches. Because of the dual objectives of supporting the image classification research at the four basins and of focusing on modeling impervious surface, it was determined that approximately 30 percent of the 55 impervious surface study sites would be distributed among the four study watersheds, and that at least half of the study sites would be composed predominantly of developed land.

The site selection process began by creating a grid of 0.5 km × 0.5 km cells for the Gulf of Maine ecoregion, overlaying this grid onto a map that included both land-cover and the four study basins, and determining the dominant land-cover type within each grid cell. The ecoregion was then stratified into five sub-regions composed of the four NECB study basins and the rest of the Gulf of Maine ecoregion; small areas without existing digital orthophotos, primarily in Maine and New Hampshire, were excluded from consideration. The Assabet, Sudbury, and Wading River watersheds each were allocated two predominantly developed study sites (i.e., study sites where the land-cover is predominantly developed), two forested sites, and one agricultural site; the Saugus River basin, which is largely developed, was allocated as three developed and one forested site. In the Gulf of Maine sub-region, 18 sites were predominantly developed land, and nine sites each were predominantly forested and agricultural land (Table 2).

Every 0.5 km × 0.5 km cell in each sub-region was assigned a unique number, and a sampling period was determined by dividing the total number of cells in the region by the number of cells to be selected in the random sample (e.g., for the Wading River, there are 199 cells and 5 cells to be sampled; thus, a sampling period of 40 cells). For the four watersheds, a random sequence of desired land-cover types was generated (e.g., for the Wading River, the sequence was Developed-Forested-Developed-Forested-Agriculture). A random grid number was generated (less than or equal to the period) as a starting location. If the cell corresponding to this number was composed of the desired land-cover type (e.g., D for the first cell to be selected for the Wading River watershed), then the cell was selected. If not, then the closest cell with the desired land-cover type was selected. From this location, the cell located one sampling period away was identified, and the closest cell with the next desired land-cover was selected (e.g., F in the Wading example). This continued until all sites had been selected. For the Gulf of Maine subregion, the sampling periods for each land-cover type were 2,220 cells for developed land and 4,444 cells for agricultural and forested land. Three random grid values less than or equal to 2,220 cells were generated, with agriculture assigned to the first value, developed land to the second value, and forested land to the third value. Sampling proceeded as above, but independently for each land-cover type.

Once the 55 study sites were selected, impervious surface area within each cell was digitized with a geographic information system (GIS) using black and white digital orthophotos that were taken in 1995 (Mike Altshul, University of Connecticut, written communication, December 2001). The measured impervious surface area for each site was saved for further analysis.

Adjusting Land-cover Estimates Using Accuracy Assessment Information

A land-cover accuracy assessment is based on comparisons between the land-cover designation in the classification at each location and the actual land-cover. Reference data can be developed either from a large-scale source of actual land-cover information, such as an aerial photograph or from an onsite visit. Accuracy assessment results typically are reported in an error matrix, with columns representing the amount of land at the sample sites associated with the true land-cover type, as determined from ground truth or reference data, and rows indicating how the land at the sample sites was labeled in the land-cover classification (Congalton, 1991). The error matrix can be used to estimate producer's accuracy, which refers to the proportion of total ground-truth sites known to belong to class X that are correctly classified in the map product, and user's accuracy, which refers to the proportion of ground-truth sites mapped as X that actually belong to class X (McGwire and

Table 2. Dominant Land-cover Types in Randomly Selected Impervious Surface Study Sites, by Subregion

Subregion	Dominant Land-Cover			
	Developed	Forested	Agricultural	Total
Assebet River	2	2	1	5
Sudbury River	2	2	1	5
Wading River	2	2	1	5
Sudbury River	3	1		4
Rest of Gulf of Maine ecoregion	18	9	9	36
Total	27	16	12	55

Fisher, 2001). Because of the focus of applying NLCD in impervious surface modeling, the primary interest in this study is user's accuracy.

A number of methods can be used to adjust land-cover area estimates using information about classification accuracy (Card, 1982; Prisley and Smith, 1987; Hay, 1988; Czaplewski and Catts, 1992; Buckland and Elston, 1994; Hess and Bay, 1997). Following the procedure described in Prisley and Smith (1987) and Hess and Bay (1997), an adjusted land-cover estimate for an area (e.g., a study site or watershed), represented in vector **W**, can be obtained by multiplying the original land-cover data, represented by vector **V**, by the user-probability matrix, **U**, as indicated in the following equation:

$$\mathbf{W}^{T} = \mathbf{V}^{T}\mathbf{U}. \tag{2}$$

This calculation adjusts the pixel count based on an estimate of the number of misclassified pixels in a category (i.e., provided by the user probability matrix) and apportions the pixels that were confused in the error matrix to other categories. If the error matrix reflects the expected distribution of misclassifications, the adjusted land-cover estimates can be considered a more accurate representation of the site conditions than the unadjusted scores (Prisley and Smith, 1987).

An example of this adjustment calculation was presented in Hess and Bay (1997). The original classified land-cover area was composed of land-cover A (27,090 cells) and B (20,910 cells). The corrected scene was calculated as the product of a transposed vector of the original land-cover data and the user-probability matrix, or:

$$\begin{array}{ccc} \text{Adjusted data} & = & \text{Original land-cover data} \\ \begin{array}{cc} \text{A} & \text{B} \end{array} & & \begin{array}{cc} \text{A} & \text{B} \end{array} \\ [30{,}595 \quad 17{,}405] & = & 27{,}090 \quad 20{,}910 \\ & & \text{User-probability matrix} \\ & \times & \begin{array}{cc} .904 & .096 \\ .292 & .708 \end{array} \end{array}$$

The user-probability matrix indicates that a relatively large proportion of pixels classified as type B are actually type A; this misclassification pattern is reflected in the adjustment, in which additional land area is apportioned to land-cover A from the area of land-cover B. The total land area, after adjustment, remains the same.

A simulation approach can be used to determine the sensitivity of land-cover adjustments in this example to variations in the original land-cover composition and classification accuracy. The amount of land-cover A in the watershed is varied, in increments of 2,000 cells, between 2,000 and 48,000 cells. The amounts of land-cover B are varied accordingly so that the total watershed area remains 48,000 cells. Nine scenarios of land-cover classification accuracy are considered, with land-cover classes A and B having varying combinations of low (0.50), medium (0.75), or high (0.90) user accuracy.

By focusing on land-cover A, several patterns can be perceived in the results of the simulation (Figure 2). When the original (i.e., unadjusted) value for land-cover A is low (e.g., less than 10,000 cells), the difference between the adjusted and unadjusted value for land-cover A is always positive regardless of the accuracy of the classification for category A or B; that is, the adjustment process results in a larger amount of land-cover A. As the original amount of land-cover A becomes larger, less gain is derived from the adjustment. Beyond a certain original magnitude, the adjustment becomes negative, resulting in the adjusted value for land-cover A being smaller than the original value.

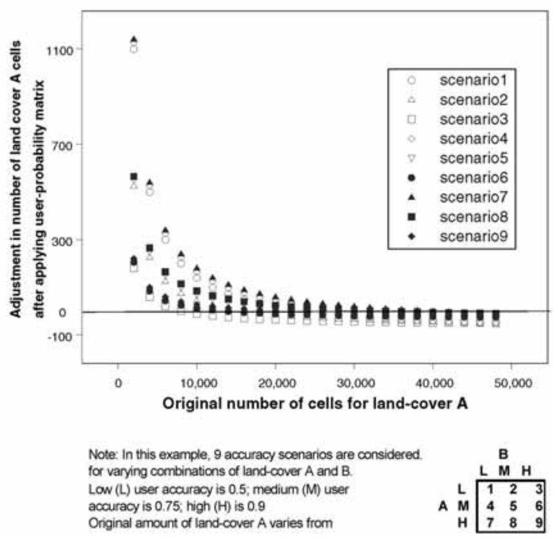

Figure 2. Effect of varying land-cover composition and classification accuracy for a hypothetical watershed in determining the adjustment percentages in land-cover composition.

While this pattern holds, regardless of the classification accuracy scenarios, the relative classification accuracies of A and B have an effect on the slope of the line that describes the transition between gaining and losing land in the adjustment process and on the location along the horizontal axis, representing zero percent adjustment, when the adjustment changes from positive to negative. When the classification accuracy for land-cover B is low (i.e., scenarios 1, 4, 7), the positive adjustment in the amount of land-cover A is always larger than when the accuracy of land-cover B is medium or high. The greater degree of confusion between classifying a cell as A or B results in the transfer of a larger amount of land classified originally as B to A in the adjustment process. In addition, when the classification accuracy for land-cover B is low, the adjustments in favor of land-cover A continue over a larger range of original values for class A. That is, when the classification accuracies for B are relatively low, adjustments based on the user probability matrix favor class A, even at relatively high original levels of class A.

Evaluating Whether Adjustment Improves Model Results

Equation 1 was used to predict impervious surface at the 55 study sites under four modeling scenarios. Mean literature-derived impervious surface coefficient values for each land-cover category were used in Equation 1 (Table 1). Models were distinguished by (a) the aggregation level of the land-cover data used (Level I or Level II), and (b) whether the original or adjusted land-cover data were used. Land-cover adjustments were made using original NLCD data for each study site and the user probability matrices (Yang *et al.*, 2001; Table 3). The user probability matrices for the NLCD Level I and II data, derived for NLCD data for the eastern United States, were assumed to be representative of classification accuracy conditions at the

Table 3. User Probability Matrix for Level I National Land-Cover Data (Vogelmann *et al.*, 2001; Yang *et al.*, 2001)

	Water	Developed	Barren	Forested	Shrub	Agriculture	Wetlands
Water	0.95	0.01	0.01	0.01	0.00	0.00	0.02
Developed	0.02	0.74	0.03	0.10	0.00	0.10	0.01
Barren	0.01	0.02	0.52	0.07	0.00	0.32	0.06
Forested	0.00	0.02	0.03	0.85	0.01	0.06	0.03
Shrub	0.00	0.00	0.15	0.05	0.80	0.00	0.00
Agriculture	0.01	0.06	0.03	0.11	0.00	0.77	0.01
Wetlands	0.02	0.01	0.04	0.10	0.00	0.04	0.78

55 study sites; ecological fallacy issues associated with the application of aggregate data to individual cases are not obviated by this assumption, but the assumption was deemed reasonable. Impervious surface predictions from each of the four scenarios were compared with the measured impervious surface data using several approaches.

The Wilcoxon signed-rank test was used to test whether the median differences between paired measured and predicted impervious surface areas at the 55 study sites under the four modeling scenarios were equal to zero (SAS, Incorporated, 1990; Helsel and Hirsch, 1992). Correlation coefficients, and the modeling efficiency coefficient (MEC) (Reckhow *et al.*, 1990; Stow *et al.*, 2003) also were used to evaluate the performance of the impervious surface models, inclusive. The correlation coefficient is a measure of the tendency of model predictions and the measured data to vary together linearly. A high correlation value, however, may mask a shift in the mean of the measured and predicted values, which may vary together, but the actual values may not match well. The MEC measures the prediction accuracy of a model relative to the mean of the measurements. The MEC is calculated by dividing the RMSE by the variance of the measured values and subtracting this quantity from one. A value near one indicates a close match between predicted and measured values and a value of zero indicates that the model predicts individual measured values no better than the mean of the measured impervious surface values. Finally, graphical evidence was used to examine the relation between observed and predicted impervious surface values under the four modeling scenarios.

Association of Impervious Surface Estimates and Stream Conditions

The four modeling approaches evaluated at the 55 study sites were used to estimate impervious surface data in the 32 New England study watersheds. Impervious surface estimates for each watershed then were compared with watershed-specific measures of stream ecological condition. Aquatic invertebrates were collected at the 32 NECB sites in August 2000 using both quantitative and qualitative methods (Coles *et al.*, 2004). Correspondence analysis was used to ordinate invertebrate community data, allowing sites to be differentiated based on the patterns and composition of the community data. Values from the first axis of this ordination, called site scores, represent the differences in community structure among the 32 sites with scores scaled so that high values represented good water-quality conditions. Site scores were correlated with the impervious surface values predicted from the four models to examine the relation between impervious surface area and invertebrate communities. For comparison, correlations also were derived between the site scores and several other measures of basin development, including the amount of developed land, 1999 population density, and road density (McMahon and Cuffney, 2000). It is expected that water-quality conditions, as measured by the site scores, will be inversely related to all of these development measures.

Results and Discussion

The Results of Adjusting Land-cover Using Accuracy Assessment Information

The results of applying the user probability matrix from published accuracy assessments to adjust the original Level I and Level II land-cover data for the 55 study sites are largely consistent with the land-cover adjustment example presented earlier, with the largest land-cover adjustments occurring in land-cover classes with the lowest classification accuracy. Across the 55 study sites, the land-cover adjustment procedure results in a decrease in the area of the two Level I NLCD classes with the largest land areas, forested (mean area of 42 percent before adjustment, with a mean loss of 0.6 percent; Table 4) and developed land (mean area of 33 percent before adjustment, with a mean loss of 7 percent). The third largest NLCD class, agriculture (mean area of 13 percent before adjustment), had a mean increase in area of 3.7 percent across the 55 sites. The pattern of adjustment losses for large land-cover classes also occurs for most of the larger Level II NLCD classes, including the low-density residential class (mean area of 20.2 percent before adjustment with a mean loss of 4.6 percent), the commercial/industrial class (mean area of 10.9 percent before adjustment with a mean loss of 3 percent), and the deciduous forest class (mean pre-adjustment area of 25 percent and a mean adjustment loss of 7.1 percent). Although the mixed forest class has a relatively large mean area across the 55 study sites (11.7 percent), the adjustment process results in a mean gain in land area of 4.7 percent.

The extent to which land-cover area is reallocated among classes is affected by the classification accuracy as well as by the relative amounts of each land-cover type within a study site. The percentage of the original area at which the adjustment becomes negative varies with the classification accuracy of the three classes (Table 4; Figure 3). Level I developed land-cover has the lowest user accuracy (74 percent), and adjustments begin resulting in a loss of developed land area when the area reaches about 10 percent. Agricultural land-cover has a slightly higher user accuracy (77 percent) than land classified as developed; land-cover adjustment results in a loss of agricultural land area when agricultural land exceeds 25 percent. Finally, forested land, which has the highest user accuracy (85 percent), does not begin losing land area in the adjustment process until the original area exceeds 45 percent. In addition, the mean loss of forested land (0.6 percent) is smaller than would be expected given the generally large amount of forested land at the study sites.

Similar results occurred for the Level II classes associated with developed and forested land (Figure 4). At sites where the original, unadjusted percentages of developed and forested land are low, the adjustment procedure increases the areas. The percentage of land-cover composition at which the adjustment becomes negative is, once again, related to classification accuracy. Adjustments remain positive for the land classified as low-density residential (user accuracy = 0.61)

TABLE 4. COMPARISON OF MEAN VALUE OF MEASURED LAND-COVER DATA, USER ACCURACY, AND PERCENT CHANGE FROM ORIGINAL VALUE AFTER USER PROBABILITY ADJUSTMENT PROCESS FOR LEVEL I AND LEVEL II NATIONAL LAND-COVER DATA (NLCD; VOGELMANN *ET AL.*, 2001) ACROSS 55 IMPERVIOUS SURFACE STUDY SITES

	Original Land-Cover Data (mean % over 55 study sites)	User Accuracy	Adjusted Land-Cover Data (% change from original values)
Level I			
Developed	33.3	0.74	–7.1
Barren	0.9	0.52	2.7
Forested	41.9	0.85	–0.6
Shrub	0.0	0.80	0.5
Agricultural	13.0	0.77	3.7
Wetland	9.0	0.78	–0.2
Level II			
Low-density residential	20.2	0.61	–4.6
High-density residential	2.1	0.40	0.0
Commercial/industrial	10.9	0.48	–3.0
Bare rock	0.2	0.47	1.2
Quarries	0.1	0.29	0.1
Transitional	0.6	0.45	2.2
Deciduous forest	25.0	0.60	–7.1
Evergreen forest	5.1	0.53	1.8
Mixed forest	11.7	0.67	4.7
Deciduous shrub	0.0	0.77	0.5
Pasture	1.4	0.33	2.4
Row crops	5.2	0.49	–0.2
Urban grass	6.4	0.61	2.0
Woody wetland	5.7	0.61	–0.8
Herbaceous wetland	3.2	0.76	0.1

across a greater range of original values (i.e., the intercept of the zero axis is farther to the right) than for land classified as high-density residential (user accuracy = 0.40) or commercial/industrial (user accuracy = 0.48) land. Adjustments for mixed forested land (user accuracy = 0.67) remain positive across a greater range of original class percentages than for deciduous (user accuracy = 0.60) and evergreen (user accuracy = 0.53) forested land. The preferential adjustment pattern for mixed forested land results in an average adjustment gain across the 55 sites.

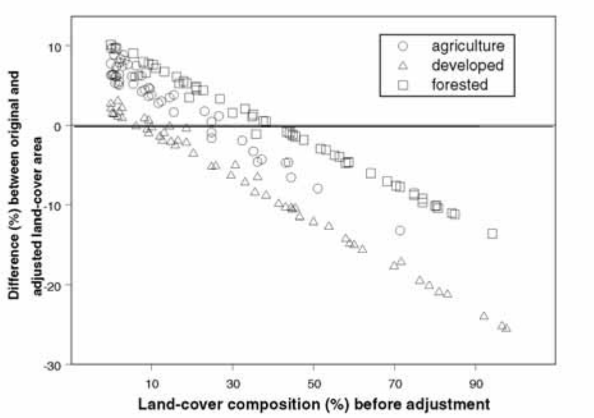

Figure 3. Effect of adjustment, with user-probability matrix, of three major Level I National Land-Cover Data (Vogelmann *et al.*, 2001) land-cover categories in 55 study sites in the Gulf of Maine ecoregion.

Adjusting land-cover using the user probability matrix has two effects on impervious surface predictions (Table 5). Median values for impervious surface predictions from models using Level I NLCD (adjusted and unadjusted) are larger than the median values predicted by using adjusted or unadjusted Level II NLCD, and not significantly different from the median value of the observed, or measured, impervious surface values at these sites. Level I-based median predictions are greater than predictions based on Level II classes because of the distribution of Level II developed land among the three Level II developed classes and the relative size of the impervious surface coefficients for Level I and II developed land. Level II developed land is primarily low-density residential, which has a smaller mean impervious surface coefficient (0.16 from Table 1) than the Level I developed land class (0.41). The median value of Level II predictions, adjusted and unadjusted, are larger than the median value.

Median predicted impervious surface values generated by models using unadjusted NLCD (Level I or Level II) are larger than the median predicted values for models using adjusted data (Table 5). The relatively low classification accuracy associated with developed land-cover, and associated large downward adjustment in the in the amount of developed land in many of the study sites (Figures 3 and 4), and the relatively high impervious surface coefficient of developed land together reduce the median predicted impervious-surface value across the study sites after adjustment. In the case of Level I data, the impervious surface model using unadjusted data under-predicts (relative to the models using adjusted data) up to an impervious surface prediction from unadjusted models of approximately 0.10; these are sites with low amounts of developed land and do not lose much developed land in the adjustment procedure. Beyond this point (i.e., for sites with greater amounts of developed land before adjustment), the losses of developed

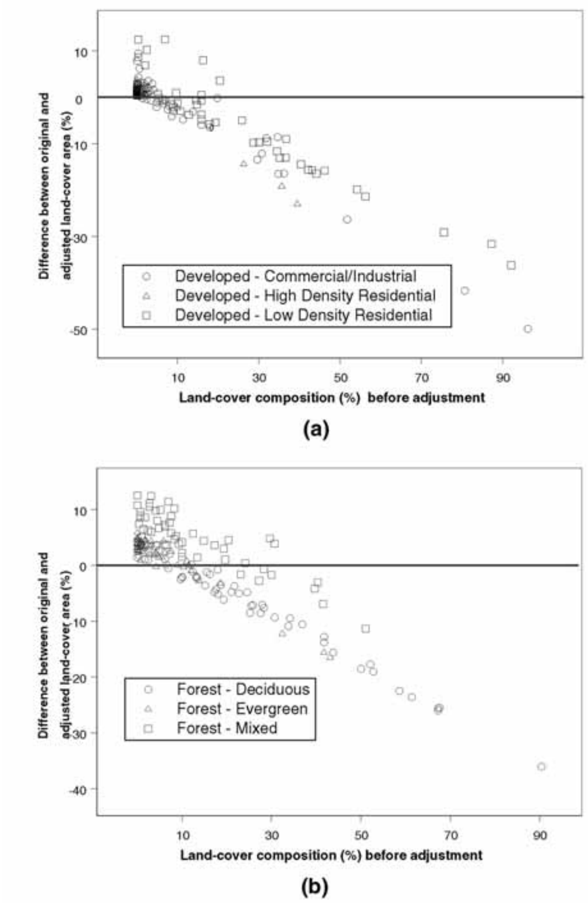

Figure 4. Effect of adjustment, with user-probability matrix, of National Land-Cover Data (NLCD; Vogelmann *et al*., 2001) (a) developed, and (b) forested Level II land-cover categories in 55 study areas.

land from the adjustment process are large enough that results from the unadjusted model predictions are always larger than those from the adjusted models. A similar pattern occurs in models using Level II NLCD. Beyond an impervious surface prediction from unadjusted models of approximately 0.15, large losses of developed land occur in the adjustment process, particularly from the commercial/industrial class, which has a high impervious surface coefficient.

Results of the Adjustment to Improve Impervious Surface Predictions

Multiple lines of evidence indicate that using error matrix information to adjust land-cover data used in impervious surface models does not substantially improve impervious surface predictions. The resolution of the land-cover data (Level I and Level II) is more important in predicting impervious surface accurately than whether the land-cover data have been adjusted using information in the error matrix.

Results from the Wilcoxon signed-rank test (which hypothesizes that the median difference between paired measured and predicted impervious surface values is zero) indicate that that there is no significant difference between the measured and predicted impervious surface values using either Level I adjusted or unadjusted data (Table 5; $\alpha = 0.05$); that is, regardless of whether the Level I NLCD land is adjusted or not, the predictions based on Level I NLCD are not significantly different from the measured values. This finding is supported by the modeling efficiency coefficient (MEC) results. The MEC measures how well a model predicts relative to the mean of the measured data, with values near one indicating a close match between predicted and measured values. MEC values for the Level I models are both relatively low, corroborating the Wilcoxon results. Predictions from both the adjusted and unadjusted Level II models are significantly different from the measured impervious surface value.

If the median difference between measured and predicted values is zero, data points in Figure 5 will lie along a 1:1 line and the Lowess line (locally weighted scatter plot smoothing, indicating the general trend in the relation among the data points; Helsel and Hirsch, 1992) will be coincident with the 1:1 line. Data points associated with the Level I (unadjusted) model are centered around this 1:1 line for predicted impervious surface values less than approximately 0.15 (corresponding with sites that are less than approximately 20 percent developed land; Figure 5a). For impervious surface values ranging from approximately 0.15 to 0.25, the model using unadjusted Level I data over-predicts relative to measured values; beyond impervious surface values of 0.25 the model consistently under-predicts relative to measured values. The overall pattern is similar for results from the model using adjusted Level I NLCD data (Figure 5b); the adjustment process reduces the amount of developed land and removes the area of over-prediction. Except at very low impervious surface levels, the model using unadjusted Level II data generally under-predicts relative to measured impervious surface values (Figure 5c). The adjustment process with Level II results in a similar pattern of model over-prediction as with the Level I data (Figure 5d).

These results reflect the effect of adjusting for classification error. When the percentages of developed land-cover in the study sites are low, the adjustment process does not introduce a large bias in the predictions relative to the measured values. As the amount of developed land increases, impervious surface predictions resulting from adjusted land-cover become increasingly lower than from the measured values. When amounts of developed land are greater, adjustments associated with land-cover misclassification become increasingly negative and predicted impervious surface values decrease. Points associated with the Level II model begin to diverge from the 1:1 line at lower impervious surface values than with the Level I data, reflecting the high classification-error rates associated with the Level II developed classes.

The Relation Between Impervious Surface and Stream Ecological Conditions

All four models were used to predict impervious surface areas at 32 New England watersheds used in the NAWQA water quality study. Modeled impervious surface values from all four models have a strong negative correlation with invertebrate condition site scores, indicating that as the amount of impervious surface increases, stream ecological conditions, as measured by invertebrate condition site scores, worsen (Table 6). The strength of this relation is similar regardless of the impervious surface model; again, the adjustment process does not seem to provide valuable additional information. The correlation between imperviousness and stream ecological condition has a similar magnitude and direction as the correlations between stream condition and other measures of urbanization. None of these development measures is clearly superior to the others in the strength of its association with this water quality measure.

Table 5. Impervious Surface Model Results Based on Level I and II National Land-Cover Data (NLCD; Vogelmann *et al.*, 2001) in 55 Study Sites in the New England Study Area: (i) Measured Values; (ii) Predicted Values; and (iii) Validation Results

(i) Observed Impervious-surface values at the 55 study sites

median	0.106
mean	0.166

(ii) Summary of predicted impervious surface values at the 55 study sites

		Predicted impervious surface values	
NLCD resolution	NLCD data used in model	Median impervious surface value	Mean impervious surface value
Level I	Unadjusted	0.125	0.157
	Adjusted	0.117	0.137
Level II	Unadjusted	0.094	0.131
	Adjusted	0.099	0.118

(iii) Validation results comparing predicted and observed impervious surface values

NLCD resolution	NLCD used in model	Wilcoxen sign rank test (p-value)	Correlation coefficient	Modeling efficiency (MEC)
Level I	Unadjusted	0.75	0.78	0.59
	Adjusted	0.68	0.79	0.49
Level II	Unadjusted	0.01	0.92	0.74
	Adjusted	0.03	0.93	0.49

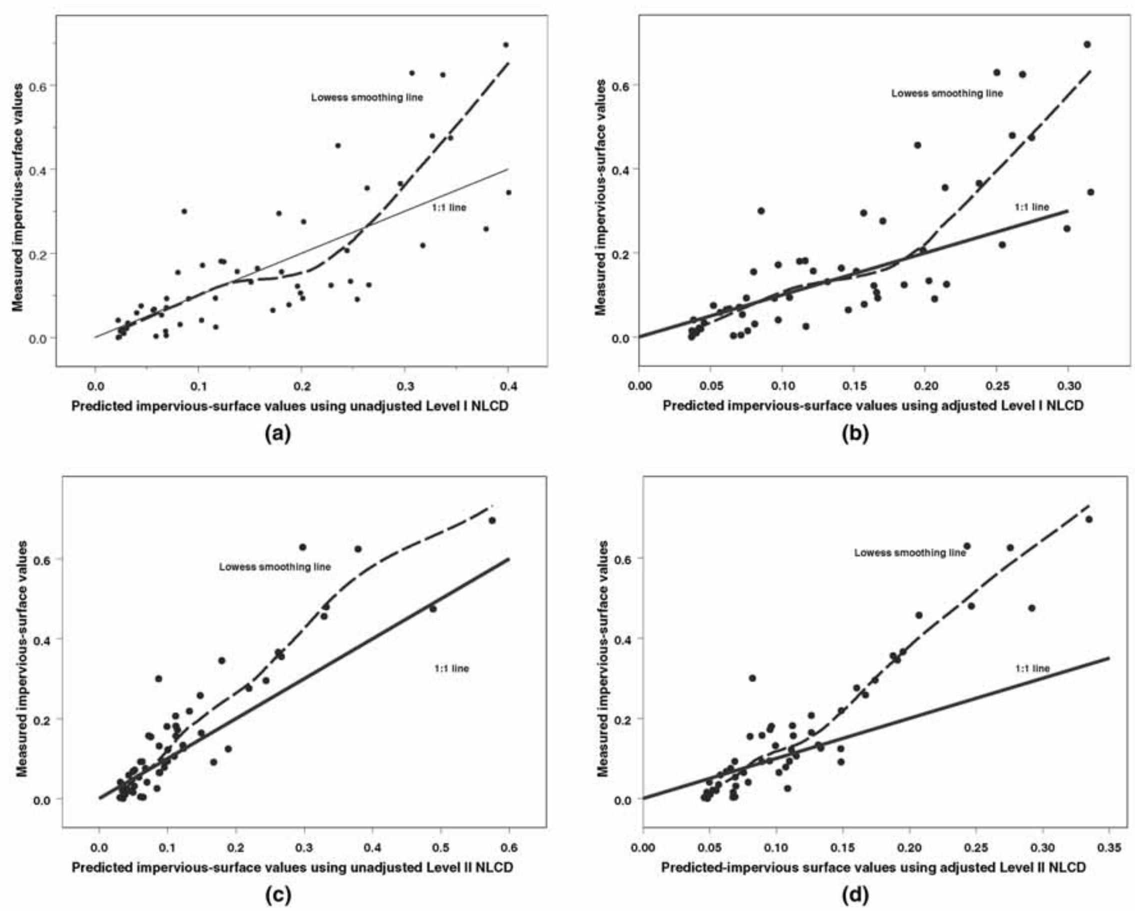

Figure 5. Measured and predicted impervious surface values at 55 New England study sites. Predictions are based on models using (a) unadjusted National Land-Cover Data (NLCD; Vogelmann *et al.*, 2001) Level I data, (b) adjusted Level I NLCD, (c) unadjusted NLCD Level II data, and (d) adjusted NLCD Level II data.

Table 6. Correlation Between Invertebrate Condition Ordination Site Scores (Coles *et al.*, 2004), Modeled Impervious Surface Predictions Using Level I and II National Land-Cover Data (NLCD; Vogelmann *et al.*, 2001) and Other Measures of Development in 32 New England Watersheds

	Invertebrate Condition Site Scores
Modeled impervious surface	
NLCD Level I, unadjusted	–0.87
NLCD Level I, adjusted	–0.87
NLCD Level II, unadjusted	–0.87
NLCD Level II, adjusted	–0.88
Other measures of developmental intensity	
Level I developed land cover (%)	–0.87
1999 population density	–0.85
Road density	–0.89

Conclusions

The evaluation of consequences of land-cover misclassification focused on three questions:

1. What are the results of adjusting land-cover area within a watershed to reflect accuracy-assessment information?
2. Does explicit adjustment of land-cover values used in impervious surface models to account for land-cover classification error rates result in meaningful improvements of impervious surface predictions over models that do not make these adjustments?
3. Do modeled impervious surface values have a stronger association with measures of stream ecological conditions than other measures of urbanization?

In the case examined above, when the original (i.e., unadjusted) amount of a specific land-cover is low, the difference between the adjusted and unadjusted value is always positive regardless of the classification accuracy; that is, the adjustment process results in a larger amount of the specific land-cover type. As the original area of the specific land-cover becomes larger, there is less gain from the adjustment; beyond a certain original magnitude, the adjustment becomes negative. Relative classification accuracy affects the point at which the adjustment changes from a gain in land-cover area to a loss. For Level I developed land-cover, which has a relatively low classification accuracy, adjustments begin resulting in a loss of developed land area when the amount of developed land is approximately 10 percent, much sooner than for the more accurately classified land-cover types, such as forest (beginning at 45 percent) and agriculture (beginning at 25 percent). A similar pattern occurs for Level II land classes.

As noted by DeFries and Los (1999), the consequences of adjusting for misclassification errors are most meaningfully viewed in the context of the land-cover application. Given an interest in improving impervious surface predictions and using impervious surface estimates to understand stream ecological condition, the model validation results indicate that using error-matrix information to adjust 1992 NLCD land-cover values used in impervious surface models does not substantially improve impervious surface predictions. In fact, the resolution of the land-cover data (Level I and Level II) is more important in predicting impervious surface accurately than whether the land-cover data have been adjusted using information in the error matrix. Level I NLCD is preferable to the other land-cover options for use in models of impervious surface. This result is tied to the lower classification error rates for the Level I NLCD. As indicated by the results in Figure 5, the lower classification-error rate for Level I developed land relative to Level II developed land results in a closer fit between predicted and measured impervious surface rates over a larger range of impervious surface values (and amounts of developed land-cover) than is the case for the Level II models.

A comparison of the association between modeled impervious surface values, other measures of urbanization, and ecological measures of stream condition indicates that there is little difference between imperviousness based on adjusted or unadjusted NLCD and other measures of urbanization in the strength of association with invertebrate condition site scores. All measures of urbanization have a negative association with water quality, as measured by invertebrate abundance.

In the context of developing impervious surface estimates and using these estimates to understand stream ecological conditions, it does not appear that the consequences of making the adjustments warrant the effort. Nevertheless, this evaluation exercise provides useful information. Given the relatively low classification accuracy of the Level I and II developed land data, the evaluation indicates that regardless of whether adjustment is done, the Level I NLCD are preferable to Level II NLCD for impervious surface models. Evaluation of land-cover adjustment based on the accuracy assessment also improves understanding of the consequences of classification accuracy for land-cover classes important in a modeling exercise. The loss of land-cover area owing to the adjustment process happens most quickly, and the average amount of change is higher when the classification accuracy is low. This indicates that classification accuracy is particularly important for land-cover classes that are important factors in modeling, such as developed land-cover for models of impervious surface.

Acknowledgments

The author is extremely grateful to Julie McIntyre and Steve Stehman for substantial assistance in conceptualizing the sampling approach and implementation of the land-cover adjustment procedure. Mike Altshul developed the measured impervious data at the 55 study sites, and Dan Civco and Sandy Prisloe provided guidance in the early stages of the project. Insights provided by three anonymous reviewers have substantially strengthened the paper.

References

American Planning Association, 2002. *Policy Guide on Water Resources Management*, URL: http://www.planning.org/policyguides/waterresources.htm (last date accessed: 15 August 2007).

Aqua Terra Consultants, 1994. *Chambers Watershed HSPF Calibration*, (D.C. Beyerlein and J.T. Brascher, editors), Thurston County Storm and Surface Water Program, Everett, Washington.

Arnold, C.L., and C.J. Gibbons, 1996. Impervious surface coverage: The emergence of a key environmental indicator, *Journal of the American Planning Association*, 62(2):243–258.

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling, 1999. *Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish*, U.S. Environmental Protection Agency, Office of Water, EPA 841-B-99–002, Washington, D.C.

Barnes, K.B., J.M. Morgan, III, and M.C. Roberge, 2000. Impervious surfaces and the quality of natural and build environments, *Synergy Project – Chesapeake Bay and Mid-Atlantic From Space*, Towson University, Department of Geography and Environmental Planning, URL: http://chesapeake.towson.edu/landscape/impervious, Towson, Maryland (last date accessed: 15 August 2007).

Bedient, P.B., and W.C. Huber, 1988. *Hydrology and Floodplain Analysis*, Addison-Wesley Publishing, Reading, Massachusetts.

Booth, D.B., D. Hartley, and R. Jackson, 2002. Forest cover, impervious surface area, and the mitigation of stormwater impacts, *Journal of the American Water Resources Association*, 38(3):835–845.

Booth, D.B., and C.R. Jackson, 1997. Urbanization of aquatic systems: Degradation thresholds, stormwater detection, and the limits of mitigation, *Journal of the American Water Resources Association*, 33(5):1077–1090.

Buckland, S.T., and D.A. Elston, 1994. Use of groundtruth data to correct land-cover area estimates from remotely sensed data, *International Journal of Remote Sensing*, 15(6):1273–1282.

Card, D.H., 1982. Using known map category marginal frequencies to improve estimates of thematic map accuracy, *Photogrammetric Engineering & Remote Sensing*, 48(3):431–439.

Center for Watershed Protection, 1998. *Rapid Watershed Planning Handbook: A Comprehensive Guide for Managing Urbanizing Watersheds*, Ellicott City, Maryland.

City of Olympia Public Works Department, 1995. *Impervious Surface Reduction Study*, Olympia, Washington.

Civco, D.L., and J.D. Hurd, 1997. Impervious surface mapping for the state of Connecticut, *Proceedings of the 1997 ASPRS/ACSM Annual Convention*, Seattle, Washington, 3:124–135.

Climate Source, 2003. Climate data where you need it, URL: http://www.climatesource.com/ (last date accessed: 15 August 2007).

Coles, J.F., T.F. Cuffney, Gerard McMahon, and K.M. Beaulieu, 2004. *The Effects of Urbanization on the Biological, Physical, and Chemical Characteristics of Coastal New England Streams*, U.S. Geological Survey Professional Paper 1695, 47 p.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 37(1):35–46.

Czaplewski, R.L., and G.P. Catts, 1992. Calibration of remotely sensed proportion or area estimates for misclassification error, *Remote Sensing of Environment*, 39:29–43.

Davis, W.S., and T.P. Simon, 1995. *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*, Lewis Publishers, Boca Raton, Florida.

DeFries, R.S., and S.O. Los, 1999. Implications of land-cover misclassification for parameter estimates in global land-surface models: An example from the Simple Biosphere Model (SiB2), *Photogrammetric Engineering & Remote Sensing*, 65(9):1083–1088.

Dunne, T., and L.B. Leopold, 1978. *Water in Environmental Planning*, W.H. Freeman and Company, San Francisco, California.

Fitzpatrick, F.A., I.R. Waite, P.J. D'Arconte, M.R. Meador, M.A. Maupin, and M.E. Gurtz, 1998. *Revised Methods for Characterizing Stream Habitat in the National Water-Quality Assessment Program*, U.S. Geological Survey Water Resources Investigations Report 98-4052, Reston, Virginia.

Frenzel, S.A., and C.S. Couvillion, 2002. Fecal-indicator bacteria in streams along a gradient of residential development, *Journal of the American Water Resources Association*, 38(1):265–273.

Hay, A.M., 1988. The derivation of global estimates from a confusion matrix, *International Journal of Remote Sensing*, 9(8):1395–1398.

Helsel, D.R., and R.M. Hirsch, 1992. *Statistical Methods in Water Resources*, Elsevier Publishers, Amsterdam, 522 p.

Hess, G.R., and J.M. Bay, 1997. Generating confidence intervals for composition-based landscape indices, *Landscape Ecology*, 12:309–320.

Jennings, D.B., S.T. Jarnagin, and D.W. Ebert, 2004. A modeling approach for estimating watershed impervious surface area from National Land Cover Data 92, *Photogrammetric Engineering & Remote Sensing*, 70(11):1295–1307.

Johnson, R., J.F. Pankow, D.A. Bender, C.V. Price, and J.S. Zogorski, 2000. MTBE-To what extent will past releases contaminate community water supply wells?, *Environmental Science & Technology*, 34(9):210A–217A.

Kennen, J.G., 1999. Relation of macroinvertebrate community impairment to catchment characteristics in New Jersey streams, *Journal of the American Water Resources Association*, 35(4):939–955.

Kluitneberg, E., 1994. *Determination of Impervious Surface Area and Directly Connected Impervious Area*, Memo for the Wayne County Rouge Program Office, Detroit, Michigan.

Krug, W.R., and G.L. Goddard, 1986. *Effects of Urbanization on Streamflow, Sediment Loads, and Channel Morphology in Pleasant Branch Basin near Middleton, Wisconsin*, U.S. Geological Survey Water Resources Research Investigations 85–4068.

McGwire, K.C., and P. Fisher, 2001. Spatially variable thematic accuracy: Beyond the confusion matrix, *Spatial Uncertainty in Ecology, Implications for Remote Sensing and GIS Applications* (C.T. Hunsaker, M.A. Friedl, M.F. Goodschild, and T.J. Case, editors), Springer, New York, pp. 308–329.

McMahon, G., J.D. Bales, J.F. Coles, E.M.P. Giddings, and H. Zappia, 2003. Use of stage data to characterize hydrologic conditions in an urbanizing environment, *Journal of the American Water Resources Association*. 39(6):1529–1546.

McMahon, G., and T.F. Cuffney, 2000. Quantifying urban intensity in drainage basins for assessing stream ecological conditions, *Journal of the American Water Resources Association*, 36(6):1247–1262.

Moulton, S.R. II, J.G. Kennen, R.M. Goldstein, and J.A. Hambrook, 2002. *Revised Protocols for Sampling Algal, Invertebrate, and Fish Communities as Part of the National Water-Quality Assessment Program*, U.S. Geological Survey Open-File Report 02-150, Reston, Virginia.

Northern Virginia Planning District Commission, 1980. *Guidebook for Screening Urban Nonpoint Pollution Management Strategies*, Metropolitan Washington Council of Governments, Falls Church, Virginia.

Novotny, V., and G. Chesters, 1981. *Handbook of Urban Nonpoint Pollution: Sources and Management*, Van Nostrand Reinhold Company, New York.

Prisley, S.P., and J.L. Smith, 1987. Using classification error matrices to improve the accuracy of weighted land-cover models, *Photogrammetric Engineering & Remote Sensing*, 53(9):1259–1263.

Prisloe, M.P., L. Giannotti, and W.J. Sleavin, 2000. Determining impervious surfaces for watershed modeling applications, *Proceedings of the 8th National Nonpoint Source Monitoring Conference*, 10–14 September, Hartford, Connecticut.

Reckhow, K.H., J.T. Clements, and R.C. Dodd, 1990. Statistical evaluation of mechanistic water-quality models, *Journal of Environmental Engineering*, 116(2):250–268.

SAS, Incorporated, 1990. *SAS/STAT User's Guide*, Version 6, Cary, North Carolina.

Schueler, Thomas, 1994. The importance of imperviousness, *Watershed Protection Techniques*, 1(3):100–111.

Shelton, L., 1994. *Field Guide for Collecting and Processing Stream-water Samples for the National Water-Quality Assessment Program*, U.S. Geological Survey Open-File Report 94–455, URL: http://ca.water.usgs.gov/pnsp/pest.rep/sw-t.html, Sacramento, California (last date accessed: 15 August 2007).

Sleavin, W.J., 1999. *Measuring Impervious Surfaces in Connecticut Using Planimetric GIS Data*, Masters Thesis, University of Connecticut, Storrs, Connecticut.

Spatial Climate Analysis Service, 2003. Spatial climate analysis service, Oregon State University, URL: http://prism.oregonstate.edu/ (last date accessed: 15 August 2007)

Slonecker, E.T., and J.S. Tilley, 2004. An evaluation of the individual components and accuracies associated with the determination of impervious area, *GIScience and Remote Sensing*, 41(2):165–184.

Stankowski, S.J., 1972. *Population Density as an Indirect Indicator of Urban and Suburban Land-Surface Modifications*, U.S. Geological Survey Professional Paper 800-B, Reston, Virginia, pp. B219–B224.

Stehman, S.V., T.L. Sohl, and T.R. Loveland, 2003. Statistical sampling to characterize recent United States land-cover change, *Remote Sensing of Environment*, 86(4):517–529.

Stehman, S.V., J.D Wickham, J.H. Smith, and L.Yang, 2003. Thematic accuracy of the 1992 National Land-Cover Data for the eastern United States: Statistical methodology and regional results, *Remote Sensing of Environment*, 86(4):500–516.

Stow, C.A., C. Roessler, M.E. Borsuk, J.D. Bowen, and K.H. Reckhow, 2003. Comparison of estuarine water quality models for total maximum daily load development in the Neuse River estuary, *Journal of Water Resources Planning and Management*, 129(4):307–314.

U.S. Department of Agriculture, 1986. *Urban Hydrology for Small Watersheds*, Natural Resources Conservation Service Technical Release 55, Washington, D.C.

U.S. Department of Agriculture, 1994. *State Soil Geographic (STATSGO) Data Base Data Use Information*, U.S. Department of Agriculture-Natural Resources Conservation Service Miscellaneous Publication 1492.

U.S. Environmental Protection Agency, 1997. *Urbanization and Streams: Studies of Hydrologic Impacts*, EPA841-R-97–009, Washington, D.C.

U.S. Environmental Protection Agency, 2002. Water discharge permits (Permit Compliance System), URL: http://www.epa.gov/enviro/html/pcs/ (last date accessed: 15 August 2007).

U.S. Environmental Protection Agency, 2003. Toxic Release Inventory (TRI) Program, URL: http://www.epa.gov/tri/tridata/tri01/index.htm (last date accessed: 15 August 2007).

U.S. Geological Survey, 2002a. The National Elevation Dataset, URL: http://ned.usgs.gov/ (last date accessed: 15 August 2007).

U.S. Geological Survey, 2002b. National Land-cover Characterization Project, URL: http://landcover.usgs.gov/ (last date accessed: 15 August 2007).

U.S. Geological Survey, 2002c. National Water-Quality Assessment Summary Reports, URL: http://water.usgs.gov/nawqa/nawqa_sumr.html (last date accessed: 15 August 2007).

U.S. Geological Survey, 2003a. National Hydrography Dataset, URL: http://nhd.usgs.gov/index.html (last date accessed: 15 August 2007).

U.S. Geological Survey, 2003b. National Water-Quality Assessment Program, URL: http://water.usgs.gov/nawqa/ (last date accessed: 15 August 2007).

Van Sickle, J., 2003. Analyzing correlations between stream and watershed attributes, *Journal of the American Water Resources Association*, 39(3):717–726.

Vogelmann, J.E., S.M. Howard, L. Yang, C.R. Larson, B.K. Wylie, and N. Van Driel, 2001. Completion of the 1990s National Land Cover Data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources, *Photogrammetric Engineering & Remote Sensing*, 67(6):650–652.

Yang, L., S.V. Stehman, J.H. Smith, and J.D. Wickham, 2001. Thematic accuracy of MRLC land-cover for the eastern United States, *Remote Sensing of Environment*, 76:418–422.

Yang, L., C. Huang, C.G. Homer, B.K. Wylie, and M.J. Coan, 2003. An approach for mapping large-area impervious surfaces: Synergistic use of Landsat 7 ETM+ and high spatial resolution imagery, *Canadian Journal of Remote Sensing*, 29(2):230–240.

Yoder, C., 1995. Biological criteria program development and implementation in Ohio, *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making* (W.S. Davis and T.P. Simon, editors), Lewis Publishers, Boca Raton, Florida, pp. 109–144.

Zhu, Z., L. Yang, S.V. Stehman, and R.L. Czaplewski, 2000. Accuracy assessment for the U.S. Geological Survey regional land-cover mapping program: New York and New Jersey Region, *Photogrammetric Engineering & Remote Sensing*, 66(12):1425–1435.

(Received 06 July 2005; accepted 28 November 2005; revised 05 May 2006)