Utility of LANDSAT-Derived Land Use Data for Estimating Storm-Water Pollutant Loads in an Urbanizing Area

Mi-Hyun Park[1]; Irwin H. (Mel) Suffet[2]; and Michael K. Stenstrom, F.ASCE[3]

Abstract: In many watersheds located in southern California, efforts are being focused on urban runoff because of its adverse impact on receiving water quality. The Sweetwater River watershed is a good example, where the drainage area is rapidly urbanizing and deteriorating reservoir water quality. Contaminated storm water is captured and diverted but as urbanization increases, additional runoff will be generated which will overload the existing infrastructure. To better manage the diversion systems and minimize future construction, storm-water volumes and pollutant loadings need to be estimated. Due to the lack of real-time storm-water runoff monitoring data, pollutant loadings must be estimated from land use information. We used satellite imagery to estimate selected storm-water pollutant loads and compared the results to predictions using land use information from public records. Satellite imagery was useful in estimating storm-water pollutant loads and identifying high loading areas. Satellite imagery with appropriate classification is a promising tool for watershed management and for prioritizing best management practices.

DOI: 10.1061/(ASCE)0733-9372(2007)133:2(203)

CE Database subject headings: Stormwater management; Urban areas; Remote sensing; Artificial intelligence; Geographic information systems; Data analysis.

Introduction

The Sweetwater River watershed in southern California is a rapidly urbanizing watershed facing many challenges. The drainage basins to Sweetwater Reservoir, an important potable water source, are experiencing urban growth. Because urban runoff degrades the water quality, current management efforts focus on storm-water runoff. For effective storm-water treatment, it is important to capture the first flush runoff because the first flush runoff delivers most of the contaminants to the reservoir during the beginning of a storm event (Geiger 1987; Thornton and Saul 1987; Lau et al. 2002). In order to divert the first flush runoff and to protect the reservoir, urban runoff diversion systems (URDS) were constructed immediately adjacent to the reservoir (Sweetwater Authority ⟨http://www.sweetwater.org⟩). Monitoring and modeling storm-water pollution is important for proper designing and managing of the diverted systems. However, it is difficult to obtain site and event specific data throughout the drainage basins or watershed (Duncan 1995; Vaze and Chiew 2003). Many researchers have suggested different monitoring and modeling methods to estimate storm-water pollution (Vaze and Chiew 2003). Utilizing land use and rainfall data are an alternative approach (Stenstrom et al. 1984; Wong et al. 1997; Chiew and McMahon 1999; Ackerman and Schiff 2003).

However, land use data may not be available or affordable in many areas. In addition, acquiring land use information can be time consuming and labor intensive. Recently, several researchers developed urban land use and land cover classification from satellite imagery, i.e., LANDSAT series imagery using several different approaches (Ridd and Liu 1998; Stefanov et al. 2001; Pal and Mather 2003; Park and Stenstrom 2006a,b), because satellite imagery provides inexpensive and archived information of the entire earth. Recent developments of high resolution satellite imagery, i.e., IKONOS (1 m for panchromatic and 4 m for multispectral) and QUICKBIRD (0.6 m for panchromatic and 2.4 m for multispectral), provides more detailed urban information, although these data are still expensive in comparison to LANDSAT imagery.

This research estimated storm-water pollutant loading to the Sweetwater Reservoir using land use classified from satellite imagery. Bayesian networks were used for satellite image classification, which are excellent prediction and diagnosis tools in artificial intelligence (Charniak 1991; Russell and Norvig 1995). The land use and pollutant loading results were compared to the corresponding results obtained from land use information from public records. The goal of this research was to estimate annual mass pollutant loading to the reservoir and to identify the main area of high pollutant loading. The approach will assist in effectively designing and managing the URDS to protect the Sweetwater Reservoir.

[1]Postdoctoral Researcher, Dept. of Civil and Environmental Engineering, Univ. of California, Los Angeles, 5732 Boelter Hall, Los Angeles, CA 90095 (corresponding author). E-mail: mhpark@seas.ucla.edu

[2]Professor, Environmental Science and Engineering Program, Univ. of California, Los Angeles, CA 90095. E-mail: msuffet@ucla.edu

[3]Professor, Dept. of Civil and Environmental Engineering, Univ. of California, Los Angeles, 5714 Boelter Hall, Los Angeles, CA 90095. E-mail: stenstro@seas.ucla.edu

Note. Discussion open until July 1, 2007. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on June 22, 2005; approved on January 12, 2006. This paper is part of the ***Journal of Environmental Engineering***, Vol. 133, No. 2, February 1, 2007. ©ASCE, ISSN 0733-9372/2007/2-203–210/$25.00.

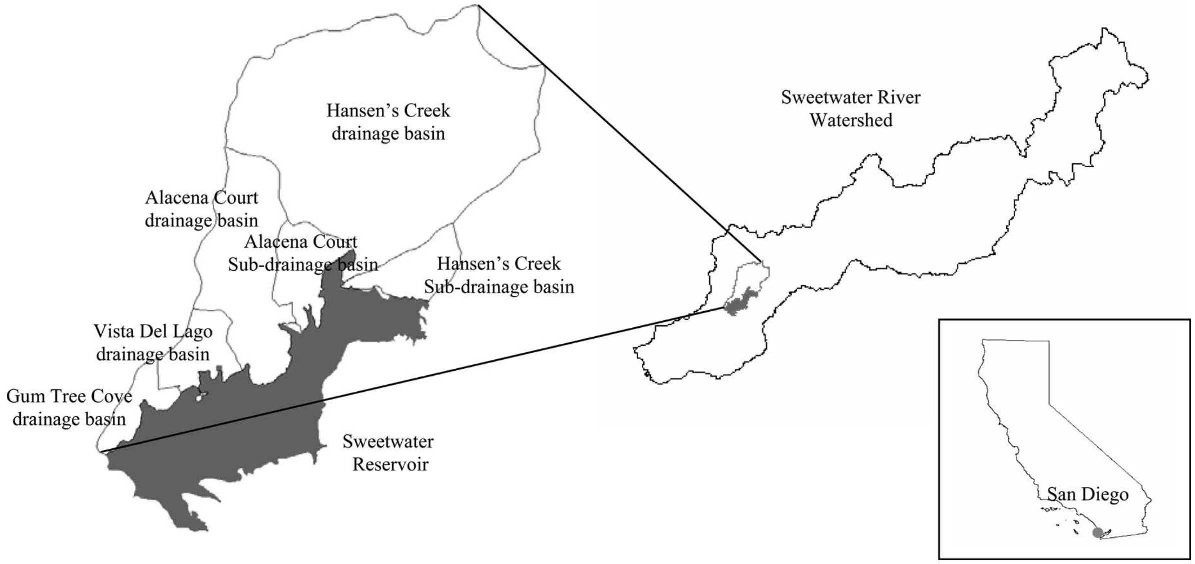


Fig. 1. The study area of six drainage basins adjacent to Sweetwater Reservoir

Methods and Materials

Study Area

The Sweetwater Reservoir is located 9 miles southeast of San Diego and is the primary drinking water source in the Sweetwater River watershed. The reservoir supplies its water to nearby cities, such as Chula Vista, National City, and Bonita, Calif. (Sweetwater Authority). In order to protect the reservoir from storm-water runoff pollution, URDSs were constructed in 1991. Six diversion ponds were constructed to capture the first flush runoff from the adjacent drainage basins. Fig. 1 shows the six drainage basins adjacent to the reservoir. The six basins compose our study area, which are the most developed areas in the watershed and undergoing rapid urbanization. The area is 12 km^2 corresponding to 2% of the entire Sweetwater River watershed. Hansen's Creek drainage basin is 7.6 km^2 and is the largest, which is 64% of the study area, followed by Alacena Court drainage basin with an area of 2.1 km^2 or 17% of the study area.

Data

The land use information from public records of the study area was obtained from San Diego's Regional Planning Agency (SANDAG 2000). The data were developed from color infrared satellite images, black and white digital orthophotography, the SanGIS landbase, the County Assessor Master Property Records file, and other ancillary information. The land use data were updated from information collected in 1990 with improved positional accuracy. The land use information was reviewed for accuracy by the local jurisdictions and the County of San Diego. Their land use system was defined with a four-digit code, as shown in Table 1, which is slightly different from USGS land cover/land use classification system (Anderson et al. 1976).

We used ArcGIS version 9.0 (ESRI, Redlands, Calif.) to reclassify the land use data into seven land uses, i.e., single-family residential (SFR), multiple-family residential (MFR), commercial (C), public (P), light industrial (I), transportation (T), and vacant (V) land uses. The study area was mostly dominated by residential and vacant land uses. Most drainage basins, except Hansen's Creek sub-drainage basin, were urbanized with approximately 60% in developed categories. Table 2 shows the distribution of land uses in each basin.

We used a LANDSAT ETM$^+$ image obtained on December 18, 1999 (path 40, row 37) for classification. The LANDSAT ETM$^+$ is the latest sensor among the LANDSAT series (Goward et al. 2001). The ETM$^+$ sensor has eight bands: Blue (Band 1: 0.45–0.52 μm), green (Band 2: 0.52–0.60 μm), red (Band 3: 0.63–0.69 μm), near infrared (Band 4: 0.76–0.90 μm), middle infrared

Table 1. Example of SANDAG Land Use Definitions

Code	Land use
1000	Spaced rural residential
1100	Single family residential
1200	Multi family residential
1300	Mobile home parks
1400	Group quarters residential
1500	Hotel/motel/resorts
2000	Heavy industry
2100	Light industry
2200	Extractive industry
2300	Junkyard/dumps/landfills
4100	Airports
5000	Commercial
6000	Office
6100	Public services
6500	Hospitals
6700	Military use
6800	Schools
7200	Commercial recreation
7600	Parks
8000	Agriculture
9100	Vacant and undeveloped land
9200	Water
9400	Public/semipublic

Note: From SANDAG.

Table 2. Land Use Percent in Each Drainage Basin from SANDAG Land Use Data (%)

Land use	Hansen	Hansen-sub	Alacena	Alacena-sub	Gum tree	Vista
SFR	26	6	46	35	27	38
MFR	8	0	0	11	2	0
C	1	0	1	0	2	0
P	5	0	2	0	14	2
I	4	0	0	0	0	0
T	13	2	18	16	14	30
V	43	92	34	39	41	31

Note: SFR=single-family residential; MFR=multiple-family residential; C=commercial; P=public; I=light industrial; T=transportation; and V=vacant.

(Band 5: 1.55–1.75 μm), thermal infrared (Band 6: 10.4–12.5 μm), another middle infrared (Band 7: 2.08–2.35 μm), and panchromatic (Band 8: 0.52–0.90 μm). All bands have spatial resolution of 30 m except the thermal infrared (60 m) and panchromatic bands (15 m). The satellite revisits all areas every 16 days. All spectral signatures are recorded in 8-bit information, which provides a range of 256 digital numbers. For this study, the thermal band was resampled to match the resolution of the other bands using the nearest neighbor method and the panchromatic band was excluded due to its overlapping spectral bandwidth with other bands. We also used spatial ancillary data, i.e., X and Y coordinate values of each pixel in the LANDSAT ETM$^+$ image because our previous work (Park and Stenstrom 2004) showed that spatial information improved classification accuracy.

The samples for training and test data were collected from all classes in approximate proportion to their abundance. The test data set was exclusive to the training data set in order to avoid biased results. The study area did not include major freeways, although a new freeway is planned near the study area. In order to include transportation land use in the classification exercise, sampling from the LANDSAT ETM$^+$ image was expanded to include transportation land use near the drainage basins.

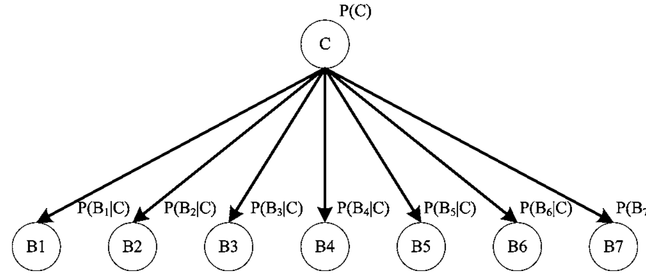


Fig. 2. An example of naïve Bayesian Classifier

Bayesian Network Classification

Bayesian networks provide knowledge representation by combining Bayes' theorem and graph theory (Pearl 1988). Fig. 2 shows graphical representation of the network with one class node, C, which is called a naïve Bayesian classifier. The nodes corresponding to the variables are connected by arcs, which illustrate dependencies among nodes associated with the conditional probability table (Pearl 1988; Neapolitan 1990; Jensen 2001). Bayesian networks can compute the posterior probability of a target variable from prior probability and likelihood using the conditional independence assumption. If we are given a class node C, which corresponds to a target variable, such as land use, and seven band nodes B_i, which correspond to measured variables, we can write Bayes' theorem in the following form

$$P(C|B_1,B_2,B_3,B_4,B_5,B_6,B_7) = \frac{P(B_1,B_2,B_3,B_4,B_5,B_6,B_7|C)P(C)}{P(B_1,B_2,B_3,B_4,B_5,B_6,B_7)} \tag{1}$$

If we assume that band nodes are conditionally independent, we can write

$$P(C|B_1, \ldots ,B_7) = \frac{P(B_1|C)P(B_2|C)P(B_3|C)P(B_4|C)P(B_5|C)P(B_6|C)P(B_7|C)P(C)}{P(B_1)P(B_2)P(B_3)P(B_4)P(B_5)P(B_6)P(B_7)} \tag{2}$$

Because we are dealing with probability distributions that sum to 1, we can write,

$$P(C|B_1, \ldots ,B_7) = \alpha P(B_1|C)P(B_2|C)P(B_3|C)P(B_4|C)P(B_5|C)P(B_6|C)P(B_7|C)P(C) \tag{3}$$

where α is set to normalize the probabilities to 1. We can easily estimate the prior probability of $P(C)$ and the likelihoods of $P(B_1|C)$ to $P(B_7|C)$ from a training data set and then we can calculate the posterior probability distribution over C for any values of any band.

The classification accuracy was assessed using tenfold cross validation in which the data were divided into ten subsets of equal size. The training and testing of the network were conducted ten times with nine subsets as training data and one subset as test data for each evaluation. The average of overall accuracies from each test was used to assess the accuracies

$$\text{overall accuracy} = \frac{\sum_k n_{kk}}{N} \tag{4}$$

where n_{kk}=number of pixels with row k; and column k and N=total number of test pixels in the confusion matrix.

Pollutant Loading Estimates

Storm-water pollutant loads for different land uses were calculated using runoff coefficients and event mean concentrations (EMCs) and were determined by each land use as following equation

Table 3. Runoff Coefficients and EMCs Based on Land Uses in LA County

LU	RC	COD	BOD_5	TSS	TKN	NH_3-N	NO_3-N	NO_2-N	TP	Cu	Pb	Zn
SFR	0.39	89	16	95	2.9	0.34	0.86	0.10	0.39	0.02	0.01	0.08
MFR	0.58	60	11	46	2.0	0.39	1.10	0.10	0.19	0.01	0.01	0.15
C	0.74	98	27	66	3.4	1.04	0.48	0.16	0.39	0.04	0.02	0.24
P	0.66	37	13	95	1.6	0.28	0.51	0.09	0.31	0.02	0.00	0.14
I	0.74	80	20	240	3.0	0.48	0.87	0.09	0.41	0.03	0.02	0.64
T	0.66	50	21	78	1.9	0.24	0.70	0.09	0.44	0.06	0.01	0.29
V	0.10	17	12	186	0.8	0.11	1.05	0.05	0.16	0.02	n/m	0.05

Note: The unit of RC, runoff coefficient, is dimensionless (after Stenstrom and Strecker 1993) and the unit of all EMCs of each storm-water quality parameter is milligram per liter (after LADPW 2000).

$$PL_i = \beta(RC)(RF)(EMC_i)A \tag{5}$$

where PL_i=annual storm-water pollutant loading for constituent i (kg/year); β=conversion factor that depends on units; RC=runoff coefficient that is the average ratio of runoff to rainfall (dimensionless); RF=annual rainfall (mm/year); EMC_i=annual event mean concentration for constituent i (mg/L); and A=area of each land use (m^2).

As shown in Table 3, runoff coefficients and EMCs of each constituent were based on the data from Los Angeles County Department of Public Works. Site specific EMCs are not available and may be measured later as a part of future projects. The 60-year average annual rainfall from local gauges of 320.5 mm was used as the annual rainfall National Oceanic and Atmospheric Administration ⟨http://www.weather.gov/climate/local-date.php?wfo=sgx⟩. Selected constituents were chemical oxygen demand (COD); biochemical oxygen demand (BOD_5); total suspended solids (TSS); nutrients, i.e., total Kjeldahl nitrogen (TKN), ammonia (NH_3-N), nitrite (NO_2-N), nitrate (NO_3-N), and total phosphorus (TP); and total copper (Cu), total lead (Pb), and total zinc (Zn).

Results and Discussion

The overall classification accuracy for satellite imagery was 84%, which demonstrates effective performance of Bayesian networks with satellite imagery. Individual land use showed varied levels of classification accuracy. We also evaluated the land use classification using satellite imagery compared to the land use information from public records (SANDAG). Fig. 3 illustrates both land use maps using SANDAG land use information and the classification from satellite imagery.

The land uses in the classification [Fig. 3(b)] visually agree with the SANDAG land use information [Fig. 3(a)], except that the classified map has a "salt and pepper" effect in land use corresponding to single-family residential land use shown through the SANDAG land use information. This occurs because the SANDAG land use information is often lumped into larger areas based upon ownership and other factors not related to spectral signature or land cover. In contrast to the SANDAG land use information, the satellite classification is able to discern these differences and assigns pixels within larger areas into their appropriate category (e.g., playgrounds or vacant land uses in school yards are classified as public in the SANDAG information).

Transportation land uses are generally missing in the classified map due to the spatial resolution of the LANDSAT ETM^+ image (30 m). Transportation land uses in the study area were mostly local streets and driveways, and were included with other land uses as mixed pixels. Therefore, the percentage of residential, commercial, and light industrial land uses predicted from satellite imagery was greater than the percentage of those land uses in SANDAG land use information as shown in Fig. 4.

The utility of the satellite image classification for estimating storm-water pollutant loading was compared to the results from SANDAG land use information. Fig. 5 illustrates comparison of the pollutant loads calculated from both approaches. Many constituents, especially BOD_5, TSS, NO_2-N, TP, and Zn show high degree of agreement that is less than 10%. The estimated loadings using satellite imagery were greater than the estimated loadings from SANDAG land use information for COD, BOD_5, TSS,

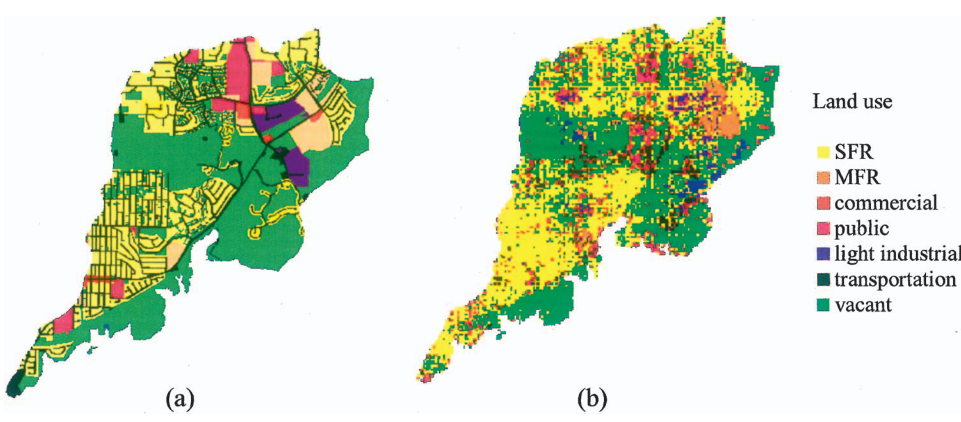


Fig. 3. (Color) Land use map in drainage basins to Sweetwater Reservoir (a) reclassified from SANDAG land use information; (b) classified from satellite imagery using a naïve Bayesian Classifier

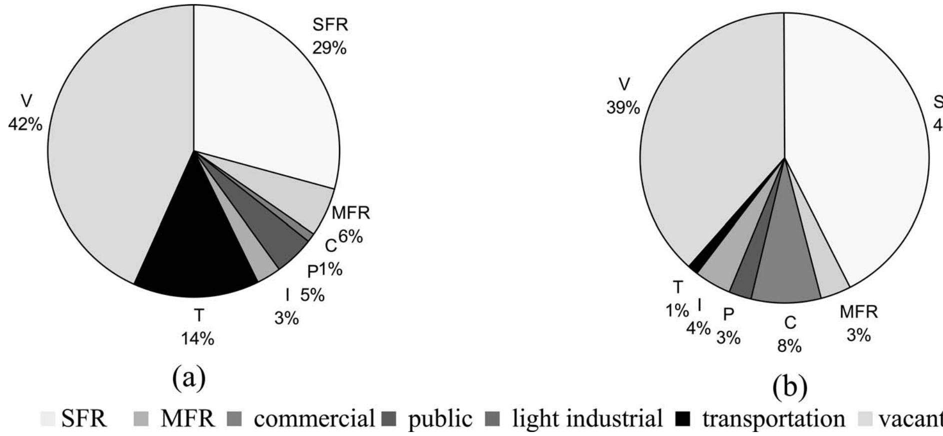


Fig. 4. Percentage of each land use in the drainage basins (a) reclassified map from SANDAG land use data; (b) classified map from satellite imagery using a naïve Bayesian Classifier

TKN, NH_3-N, NO_3-N and Pb and less for NO_2-N, TP, Cu, and Zn. The result demonstrates that using satellite imagery is useful and comparable to using the information from public records.

Fig. 6 shows that 61–71% pollutant loadings were discharged from Hansen's Creek drainage basin, which is due to its greatest area among the six drainage basins. Alacena drainage basin discharged the second greatest pollutant loadings (12–22%). Pollutant loadings from Hansen's Creek sub-drainage basin using SANDAG and satellite image classification land use were different. Estimates from SANDAG land use information ranged from 0.6 to 3% of the total loading, depending upon pollutants, whereas the predicted loading using satellite imagery was 4–7% of the total loading. Loadings for Pb and Zn were the most different between two approaches (0.6 versus 5.2% and 0.8 versus 7.0% for SANDAG and satellite image classification, respectively). The difference occurs because the land use estimated by satellite imagery contains greater amounts of commercial, public, and light industrial land uses. The three other drainage basins show similar estimates from both approaches.

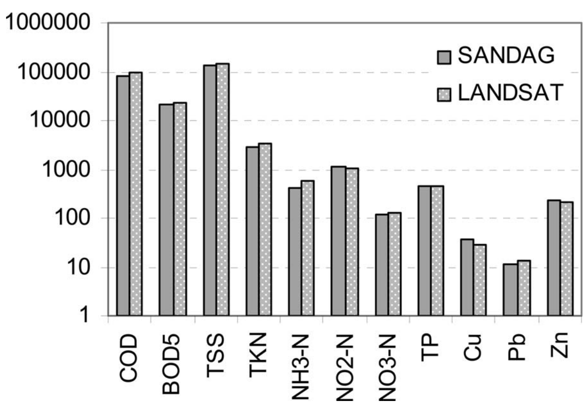


Fig. 5. Estimated annual loading for each pollutant type (kg/year)

Among the different land uses, those generating high stormwater pollutant loading were identified. Table 4 shows that single-family residential and transportation land uses generated the greatest pollutant loadings when using SANDAG land use information. Single-family residential land use discharged the greatest COD, TSS, nutrients, and Pb loadings while transportation discharged the greatest Cu and Zn loadings. When using satellite image classification, single-family residential land uses generated the greatest pollutant loadings. The exceptions were NH_3-N and Zn loadings, which were greatest from commercial land use. Pollutant loadings from transportation land use were less than other land uses because the area of transportation land use identified from satellite imagery was small. The results show that many pollutant loads are associated with the largest areas, with the exception of vacant land use. The very small runoff coefficient and

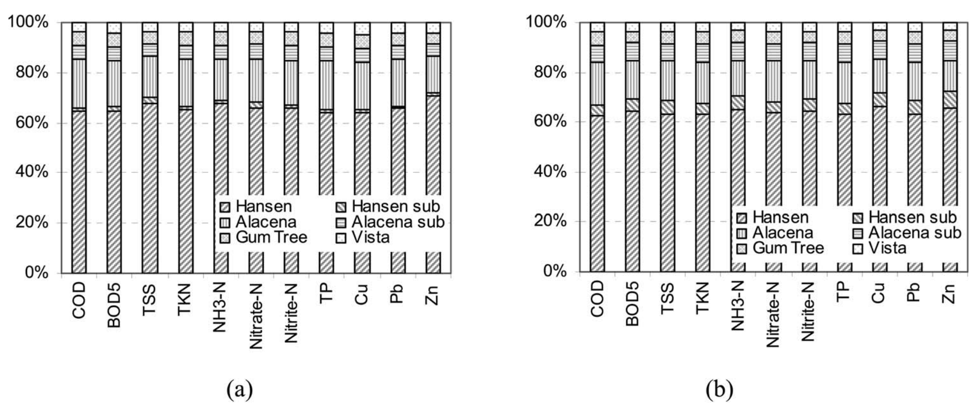


Fig. 6. Estimated annual loading percent of each pollutant type from each drainage basin (%) using (a) SANDAG land use information; (b) land use classification from satellite imagery

Table 4. Estimated Annual Loading of Each Pollutant Type in the Drainage Basins (kg/year)

Land use	COD	BOD_5	TSS	TKN	NH_3-N	NO_3-N	NO_2-N	TP	Cu	Pb	Zn
(a) Using SANDAG land use information											
SFR	40000	7100	42000	1300	150	380	45	170	6.7	4.5	35
MFR	7700	1400	5900	260	50	142	13	24	1.6	0.8	19
Commercial	2300	620	1500	78	24	11	3.7	9.0	0.9	0.4	5.5
Public	4300	1500	11000	190	33	60	11	36	2.8	0.6	16
Light industrial	6300	1600	19000	240	38	69	7.1	32	2.5	1.3	51
Transportation	18000	7400	28000	670	85	250	32	160	19.0	3.6	100
Vacant	2900	2000	31000	130	19	180	8.4	27	2.5	n/m	7.8
(b) Using land use classification from satellite imagery											
SFR	57000	10000	61000	1900	220	550	64	250	9.7	6.4	51
MFR	4400	800	3400	150	29	82	7.4	14	0.9	0.4	11
Commercial	22000	6000	15000	750	230	110	35	86	8.6	4.0	53
Public	2400	840	6100	100	18	33	5.8	20	1.5	0.3	8.9
Light industrial	9500	2400	29000	360	57	100	11	49	3.8	2.0	76
Transportation	1500	610	2300	55	7.0	20	2.6	13	1.6	0.3	8.5
Vacant	2500	1800	28000	120	16	160	7.4	24	2.2	n/m	6.8

EMCs (except TSS) for vacant land use reduce their contribution to overall loading, despite their large area.

When smaller areas contribute high pollutant loads, it may present an opportunity for storm-water management. For example, the contribution of Zn from light industrial land use was 21% of the total mass discharge even though its area is only 3% of the total area. Conversely, single family residential land use contributes only 15% of the Zn mass discharge, even though it is 29% of the total area.

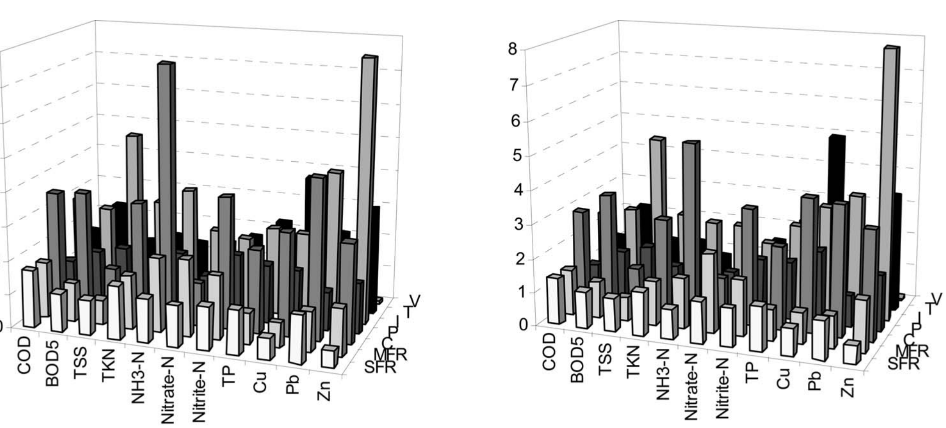
High contributions from small areas present opportunities because the cost of BMPs is generally proportional to land area treated (Devinny et al. 2004). If greater pollutant mass can be reduced from smaller areas, then a cost saving is likely. This concept has been previously called "leverage" and is quantified by the pollutant load in a particular watershed or sub watershed divided by the area of the particular land use contributing the pollutant. For example, Stenstrom et al. (1984) showed that parking lots and the associated commercial property, although composing only 6% of the area, contributed 28% of the pollutant load.

To illustrate the effects of the leverage on storm-water pollutant loading i, it was calculated for all land uses as follows:

$$\text{leverage}_i = \frac{PL_{ij}}{A_j} \tag{6}$$

where PL_{ij}=percentage of the pollutant loading i of land use j and A_j=percentage of the area of land use j. This equation represents the quantitative effect of storm-water runoff from each land use on the reservoir water quality regardless of the area of each land use. Fig. 7 and Table 5 show the leverage for each constituent, which ranged from 0 to 8.

In both approaches using SANDAG land use information and satellite image classification, commercial land use shows the highest leverage for COD, BOD, TKN, NH_3-N, NO_2-N, and Pb, whereas light industrial land use shows the highest leverage for TSS, NO_3-N, TP, and Zn. Transportation land use shows the highest leverage for Cu. Both approaches show similar leverage although the exact magnitude for each land use varied. The leverage results show that highly developed urban land uses, such as commercial, light industrial, and transportation land uses are important to manage storm-water pollution despite their small area. If the cost of applying best management practices (BMPs) is proportional to land area, the greatest pollutant reduction for a



(a) (b)

Fig. 7. Leverage: Percent pollutant loading divided by percent area of each land use in the drainage basins using (a) SANDAG land use information; (b) land use classification from satellite imagery

Table 5. Leverage: Percent Pollutant Loading Divided by Percent Area of Each Land Use in the Drainage Basins

Land use	COD	BOD5	TSS	TKN	NH_3-N	Nitrate	Nitrite	TP	Cu	Pb
(a) Using SANDAG land use information										
SFR	1.69	1.13	1.05	1.56	1.31	1.21	1.29	1.31	0.63	1.39
MFR	1.67	1.14	0.74	1.57	2.19	2.27	1.89	0.93	0.73	1.18
Commercial	3.52	3.61	1.38	3.45	7.55	1.28	3.90	2.47	3.07	4.72
Public	1.18	1.54	1.76	1.44	1.80	1.21	1.95	1.74	1.68	1.14
Light industrial	2.84	2.65	4.97	3.02	3.45	2.30	2.17	2.57	2.50	4.41
Transportation	1.59	2.49	1.45	1.71	1.55	1.65	1.95	2.47	3.91	2.32
Vacant	0.08	0.22	0.52	0.11	0.11	0.38	0.16	0.14	0.16	0.00
(b) Using land use classification from satellite imagery										
SFR	1.38	1.09	1.02	1.32	0.91	1.26	1.15	1.32	0.82	1.15
MFR	1.36	1.09	0.72	1.33	1.53	2.36	1.69	0.94	0.95	0.97
Commercial	2.88	3.47	1.34	2.92	5.26	1.33	3.49	2.49	4.01	3.89
Public	0.96	1.48	1.71	1.22	1.26	1.25	1.74	1.76	2.19	0.94
Light industrial	2.33	2.55	4.82	2.55	2.41	2.38	1.94	2.59	3.26	3.64
Transportation	1.30	2.39	1.40	1.45	1.08	1.72	1.74	2.49	5.10	1.92
Vacant	0.07	0.21	0.51	0.09	0.07	0.39	0.15	0.14	0.21	0.00

fixed total cost will be realized by addressing high leverage areas first.

In order to evaluate the impact of urban growth on the storm-water pollution, we calculated storm-water pollutant loading using SANDAG's projected changes in land use. Fig. 8 shows that urbanization will continue in the drainage basins, especially in Hansen's Creek drainage basin, Alacena Court and its sub-drainage basins, and Gum Tree Cove drainage basin. The increase of urban land uses will be significant in that single-family residential land use will increase by 44%, commercial land use by 225%, and light industrial land use by 38%. Conversely, vacant land uses will be reduced by 40%. Assuming that the urban land uses will be impervious without proper treatment facilities, the land use changes will increase storm-water pollutant loading. The estimated increase of pollutant loading will vary for each constituent and range from 13 to 28%. The result demonstrates that increasing urban land uses that are highly developed, such as commercial and light industrial, will further degrade reservoir water quality by increasing storm-water pollution.

The land use models can be used to demonstrate the potential for BMPs to reduce storm-water pollution. For example, recent efforts to reduce runoff from developed area have used bioswales to treat and infiltrate runoff. Horner et al. (2004) describes examples that retained 94% of the runoff. The impact of such efforts can be simulated by the GIS model if assumptions are made about bioswales. For instance, suppose a swale can store 0.15 m of runoff without flooding and the average runoff coefficient for the proposed land use changes for the Sweetwater Reservoir drainage basin is 0.6. As such, a 20 mm rainfall will produce 12 mm of runoff, which would need to be stored and infiltrated by the bioswale. The ratio of the land area to the bioswale area would be 150/12 or 12.5. Therefore, the proposed land development of 23 km^2 would require only 1.84 km^2 of bioswales. Alternatively, only 92% of the area could be developed and 8% could be dedicated to bioswales. The use of bioswales in this fashion would avoid the growth of storm-water pollution at a cost of only 8% of the proposed development, or even less if deeper bioswales can be used. Horner also notes that bioswales can provide benefits in

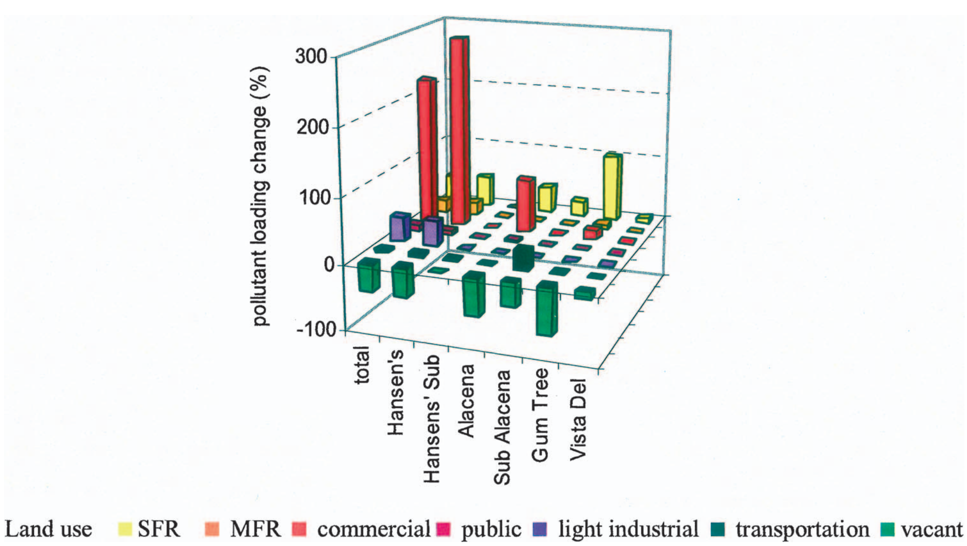


Fig. 8. (Color) Pollutant loading change due to the land use change in each drainage basin

addition to storm-water pollution prevention, such as increasing habitat and beautifying neighborhoods. Such approaches can provide a guideline to local regulation and planning agencies to protect the receiving water quality. The methods we describe in this paper can be used to model the potential impact of such BMPs.

Conclusions

This study has shown that land use data are useful in estimating storm–water pollutant loading in a watershed. Our approach simplifies estimating storm-water pollutant loading where event and site-specific storm-water monitoring data are not available. Land use classified from satellite imagery was especially useful because the result was comparable with best available land use information. The use of satellite imagery is advantageous where land use information is not available, not affordable, or not classified using environmentally important categories. In addition, the public records that provide land use information are updated only every 10 years because they require considerable amount of time and effort. Therefore, information of land use and the change in land use are not collected frequently enough for growth monitoring or modeling. This problem can be resolved by using LANDSAT series imagery because they are collected on a more frequent and regular basis (every 16 days). LANDSAT ETM^+ images can also be useful for watershed scale analysis because they are inexpensive. Bayesian networks provided accurate classification, although classifying transportation land use was not easy due to the spatial resolution of the LANDSAT ETM^+ imagery.

The results show that land use area is an important factor contributing to pollutant loading in a watershed. In the case of the Sweetwater Reservoir, which still has large undeveloped areas and a large percentage of single-family residential land use, the commercial, light industrial, and transportation land uses were minor contributors to the total load even though they contribute much more pollutant load per unit area. However, the impact of commercial, light industrial, and transportation land uses can be significant despite their small areas, and the impact will increase with development and growth. Therefore, these land uses should be prioritized for storm-water pollution management. Management of a small region of high pollutant loading land use can achieve significant reduction of storm-water pollution. In the specific case of the Sweetwater Reservoir, the future growth or modification of the existing urban runoff diversion systems can be guided by the land use modeling.

Acknowledgments

The writers would like to thank the U.S. EPA Region IX and the Sweetwater Authority for their partial support of this work.

References

Ackerman, D., and Schiff, K. (2003). “Modeling storm water mass emissions to the southern California bight.” *J. Environ. Eng.*, 129(4), 308–317.

Anderson, J. R., Hardy, E., Roach, J., and Witmer, R. (1976). “A land-use and land-cover classification system for use with remote sensor data.” *Professional Paper No. 964, U.S. Geological Survey*.

Charniak, E. (1991). “Bayesian network without tears.” *AI Mag.*, 12(4), 50–63.

Chiew, F. H. S., and McMahon, T. A. (1999). “Modelling runoff and diffuse pollution loads in urban areas.” *Water Sci. Technol.*, 39(12), 241–248.

Devinny, J. S., Kamieniecki, S., and Stenstrom, M. K. (2004). “Alternative approaches to stormwater quality control.” *Rep. to the Los Angeles Regional Water Quality Control Board*, Los Angeles.

Duncan, H. P. (1995). “A review of urban stormwater quality processes.” *Rep. No. 95/9*, Cooperative Research Centre for Catchment Hydrology, Melbourne, Australia.

Geiger, W. F. (1987). “Flushing effects in combined sewer systems.” *Proc., 4th Int. Conf. on Urban Storm Drainage*, International Association of Hydraulic Runoff Research, Lausanne, Switzerland, 40–46.

Goward, S. N., Masek, J. G., Williams, D. L., Irons, J. R., and Thompson, R. J. (2001). “The Landsat 7 mission terrestrial research and applications for the 21st century.” *Remote Sens. Environ.*, 78(1–2), 3–12.

Horner, R. R., Lim, H., and Burges, S. J. (2004). “Hydrologic monitoring of the Seattle ultra-urban stormwater management projects: Summary of the 2000–2003 water years.” *Water Resources Series Technical Rep. No. 181*.

Jensen, F. V. (2001). *Bayesian networks and decision graphs*. Springer, New York.

Los Angeles County Department of Public Works (LADPW). (2000). “Los Angeles County 1994–2000 integrated receiving water impacts report.” ⟨http://ladpw.org/wmd/NPDES/IntTC.cfm⟩.

Lau, S.-L., Ma, J.-S., Kayhanian, M., and Stenstrom, M. K. (2002). “First flush of organics in highway runoff.” *Proc., 9th Int. Conf. on Urban Storm Drainage*, ASCE, Portland, Ore.

Neapolitan, R. E. (1990). *Probabilistic reasoning in expert systems: Theory and algorithms*, Wiley, New York.

Pal, M., and Mather, P. M. (2003). “An assessment of the effectiveness of decision tree methods for land cover classification.” *Remote Sens. Environ.*, 86(4), 554–565.

Park, M., and Stenstrom, M. K. (2006a). “Using satellite imagery for stormwater pollution management with bayesian networks.” *Water Res.*, 40(18), 3429–3438.

Park, M., and Stenstrom, M. K. (2006b). “Spatial estimates of stormwater pollutant loading using bayesian networks and geographic information systems.” *Water Environ. Res.*, 78(4), 421–429.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*, Morgan Kaufmann, San Mateo, Calif.

Ridd, M. K., and Liu, J. (1998). “A comparison of four algorithms for change detection in an urban environment.” *Remote Sens. Environ.*, 63(2), 95–100.

Russell, S., and Norvig, P. (1995). *Artificial intelligence: A modern approach*, Prentice Hall, Upper Saddle River, N.J.

San Diego’s Regional Planning Agency (SANDAG). (2000). “2000 land use.” ⟨http://www.sandag.cog.ca.us/⟩.

Stefanov, W. L., Ramsey, M. S., and Christensen, P. R. (2001). “Monitoring urban land cover change; an expert system approach to land cover classification of semiarid to arid urban centers.” *Remote Sens. Environ.*, 77(2), 173–185.

Stenstrom, M. K., Silverman, G. S., and Bursztynsky, T. A. (1984). “Oil and grease in urban stormwaters.” *J. Environ. Eng.*, 110(1), 58–72.

Stenstrom, M. K., and Strecker, E. (1993). “Assessment of storm drain sources of contaminants to Santa Monica Bay.” *Annual pollutants loadings to Santa Monica Bay from stormwater runoff, UCLA-ENG-93-62*, Vol. I, Los Angeles.

Thornton, R. C., and Saul, A. J. (1987). “Temporal variation of pollutants in two combined sewer systems.” *Proc., 4th Int. Conf. on Urban Storm Drainage*, International Association of Hydraulic Runoff Research, Lausanne, Switzerland, 51–52.

Vaze, J., and Chiew, F. H. S. (2003). “Comparative evaluation of urban storm water quality models.” *Water Resour. Res.*, 39(10), 1280–1289.

Wong, K., Strecker, E. W., and Stenstrom, M. K. (1997). “GIS to estimate storm-water pollutant mass loadings.” *J. Environ. Eng.*, 123, 737–745.