CE 440
Basic Concepts of Diffuse Pollution
Lecture Notes on Chapter 3

Text Reading: pp. 104-132

Watersheds and Drainage Networks

Define watershed, stream order, thalweg, floodplain

Important table from Leopold, Wolman, and Miller (1964) on stream orders in US:

Strahler Order	Number	Average Length (km)	Total Length (km)	Mean Drainage Area (sq km)	% Total Length	Cumulative %total length	% Total number	Cumulative %total number
1	1570000	1.6	2526547	2.5897353	48%	48%	78%	78%
2	350000	3.7	1295459	12.171756	25%	73%	17%	95%
3	80000	8.5	682329	59.563913	13%	86%	4%	99%
4	18000	19.3	347601	282.28115	7%	93%	1%	100%
5	4200	45.1	189250	1341.4829	4%	97%	0%	100%
6	950	103.0	97843	6370.749	2%	98%	0%	100%
7	200	236.6	47312	30299.904	1%	99%	0%	100%
8	41	543.9	22301	143989.29	0%	100%	0%	100%
9	8	1250.4	10003	683690.13	0%	100%	0%	100%
10	1	2896.7	2897	3237169.2	0%	100%	0%	100%

Perennial vs. intermittent vs. ephemeral flow

Geomorphic floodplain vs. regulatory 100-yr floodplain

Effluent dominated streams

Types of Diffuse Pollution Loads and Transport Routes

Water pollution requires both a source of a contaminant and moving water to transport the contaminant. The moving water can be either surface water or ground water or both.

See bullets describing sources on pp. 108-109

To characterize pollutant loadings from a given source, both surface runoff and percolation to ground water must be estimated.

Mass load per time = (concentration) x (flow rate)

Generally, flows are estimated using a **mass balance** for a defined **system** or **control volume,** for example a watershed, aquifer, agricultural field, mine, parking lot, or any region with carefully defined system boundaries.

Δ storage = inflow - outflow

Unit Loads (Export Coefficient Concept), starts on page 110

A unit load (UL) is a mass or weight of pollution per unit area, per capita*day, or per unit length of curb (CL).

Field measurements and "typical" unit loads from urban areas are given in Figure 3.5 and on page 115 (references differ appreciably).

"Edge of field" or "edge of site" loads vs. delivery
Be careful – some export coefficients include background

Sediment Delivery Ratio (DR)

Definition

Enrichment Ratio (ER)

Definition

The overall enrichment is a combination of the mineral soil and organic matter enrichment ratios. We do not have well established ERs for organic matter.

For **annual loads**, a rule-of-thumb type number to use for overall enrichment is ER = 2.

Loading Calculations

Load of adsorbed pollutant = Concentration on sediment x Sediment load

Load of dissolved pollutant = Dissolved concentration x Water flow rate

Concentration on sediment = Concentration on soil x ER

Mass of pollutant in runoff = pollutant emissions from source areas
x enrichment during overland flow
- attenuation in the watershed
+ background sources

Load of pollutant in runoff = DR x UL (if UL doesn't include background)

Load of adsorbed pollutant in runoff = DR x ER x soil erosion (mass/time) x $C_{s\text{-}soil}$

Statistical Quality Characteristics of Urban Runoff

The NURP study (USEPA, 1993) produced data on 28 urban sites around the country.

Runoff is often characterized by Event Mean Concentrations, EMC's.

EMC = Mass of pollutant transported during event/Total flow during event

$$EMC = \frac{\sum Q_i C_i}{\sum Q_i}$$

So each storm has a particular EMC for each water quality variable.

At a given site, one can monitor EMC's for a few seasons to determine their statistical characteristics. Such as

- shape of distribution, histogram
- summary statistics
- percentiles
- return periods of various EMC's

Sample Mean EMC $\bar{x} = \frac{\sum x_i}{N} \approx m$

Sample Standard Deviation $S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N-1}} \approx s$

For normally distributed data, 2/3 of the population (or of random observations) are within ± 1 standard deviation of the mean, 95% of the population are within ± 2σ of the mean.

The coefficient of variation $CV = \frac{S}{\bar{x}} \approx \frac{s}{m}$

The median is the 50th percentile of the distribution.

Estimating Percentiles (from data)

First rank the data from smallest to largest. The rank of the (100p)th percentile is

M=p(N+1)

where p is the fraction of the population below the percentile of interest. The value of p is sometimes called the **non-exceedence probability**. 1-p would be the **exceedence probability**.

Example: For the 90th percentile, p=0.9. If we have N=100 observations then M=(0.9)(101)=90.9. To find the 90th percentile, rank the data and interpolate between the 90th and 91st largest observations.

Try p=0.9 and n=10. Then M=9.9.

Try p=0.95 and n=10. Then M=10.4. (won't work)

Oops, we can't estimate the 95th percentile with only 10 data points using this estimator. This actually makes sense, we should not be able to estimate extreme percentiles (p close to 0.0 or 1.0) unless you have a sufficiently large sample size. Extreme percentiles are more difficult to estimate than those close to the median, but estimates of all percentiles will benefit from larger sample sizes.

An alternative estimator is M=pN, which will never result in M>N. This sounds good, but it will not provide estimates which are as good as those from the other estimator, and it still has the problem that N can be < 1. This is bad because you can't interpolate below the smallest value.

Return Periods

The return period of a given event, such as a rainfall depth or EMC observed from a single storm, is defined as 1/the exceedence probability of that event based on **annual maximum data**.

$$T_r = \frac{1}{1-p}$$

Let's say, for example, that 1-p=0.1, meaning that there is a 10% chance that the event in question will be exceeded in any given year. (Also, p= 0.9 meaning that there is a 90% chance that the event will not be exceeded in any given year.)Then T_r = 1/0.1 = 10 years.

To calculate return periods from a data set of rainfall amounts or EMC's, first find the annual maximum for each year, rank those and then find

$p = \frac{M}{N+1}$ from the ranked data. This equation is known as the Weibull plotting position. Then calculate the return period T_r as above. Interpolate as necessary.

Values of p and return periods can also be calculated based on an assumed distribution such as normal or lognormal distribution.

Define a lognormal distribution. What does the histogram look like?

This would be called a parametric approach since a particular distribution is assumed. The above method based on ranks is a nonparametric approach. No particular distribution is assumed.

NURP Results:

- NURP results suggest that EMC's are often lognormally distributed
- Geographic location, land use, runoff volume are poor predictors of EMC's
- There is a significant difference between urban and nonurban or open sites
- Statistical results overall sites are given in tables 8.14 (medians) and 8.15 (means). For lognormal data the mean>median. This shows up here.

CSO's:

It is useful to know the critical rainfall intensity at which overflow occurs. This can then be used to determine how often a CSO will occur. Typical values are 1 mm/hr.

Calculation of percentiles based on normal or lognormal distribution (parametric method):

For the normal distribution, the (100p)th percentile is given by

$$\mu + Z_p\sigma$$

where Z_p is the value of the standard normal distribution with nonexceedence probability or left-tail area =p which can be found from tables.

To use the lognormal distribution, first convert each observation to its log and then proceed as with the normal distribution. That is, find the mean and standard deviation of the logs, etc.

<u>Study Boxes 3.1 through 3.4.</u>

Management and Targeting Critical Areas

Source controls vs. change runoff characteristics vs. interception of pollutants

Standard Normal Distribution

m=0, s = 1.0

Commonly Used Values of Z_p

p	Z_p
0.99	2.327
0.975	1.96
0.950	1.645
0.900	1.282
0.85	1.037
0.84	1.00
0.80	0.842
0.75	0.675
0.50	0.000
0.25	-0.675
0.20	-0.842
0.16	-1.00
0.15	-1.037
0.10	-1.282
0.05	-1.645
0.025	-1.96
0.01	-2.327