Environment Protection Engineering

Vol. 5 1979 No. 1

F. J. HUMENIK, M. R. OVERCASH, L. F. HORNEY,
L. F. BLIVEN. F. A. KOEHLER*

AREAWIDE ASSESSMENT OF RURAL STREAM WATER QUALITY**

PART I

STATISTICAL SAMPLING METHODOLOGY

Stream water quality is often impacted by increased chemical loadings due to human activities, Comprehensive watershed water quality planning must quantify the impact of both point and nonpoint sources in conjunction with abatement costs to determine the most cost--effective plan to achieve water quality goals or standards. This paper series demonstrates the usefulness of probability (random) sampling and statistical data analysis techniques for watershed problem assessment and planning. The objectives are fulfilled using data obtained during a feasibility study which quantified the impact of nonpoint sources on rural water quality in a 12,900 km^2 watershed. Nonpoint sources fluctuate in space and time as a function of numerous processes which vary throughout a watershed and which may be subject to human control (agricultural practices, construction site activities, and urban storm water runoff treatment) or may not be subject to human control (soil type and weather factors). Therefore, adequate assessment of areawide stream water quality requires a sampling plan that accounts for both spatial and temporal variations. The potential of probability sampling compared to judgment sampling for quantifying new and more difficult water quality concerns are reviewed and a model probability sampling plan to assess areawide rural water quality are presented in Part I of this paper series. In the subsequent Part II article, data analysis techniques which are useful for quantifying areawide water quality are reviewed using data from a feasibility demonstration of the model probability sampling plan.

1. INTRODUCTION

The United States Congress in 1972 mandated through Public Law 92-500 [1] that the surface waters of the United States shall be "fishable and swimmable" where obtainable by the year 1983. The quality of receiving waters is impacted by both point and non-

* Department of Biological and Agricultural Engineering, North Carolina State University, Raleigh, North Carolina, USA.

** The use of trade names in this publication does not imply endorsement by the North Carolina Agricultural Experiment Station of the products names, nor criticism of similar ones not mentioned.

point sources. In the United States, all point sources are regulated by state governments within guidelines set by the U.S. Environmental Protection Agency. Discharge permits which are required for all point sources typically specify maximum effluent concentrations for low flow conditions of the receiving stream.

Many definitions have been proposed for nonpoint sources but none have achieved wide general acceptance. The definition presented here incorporates concepts which the authors feel provide a meaningful definition. Nonpoint sources (NPS) of pollution characteristically are diffuse inputs above a level normally measured for undeveloped land of a similar genesis. The dominant loading from NPS often results due to rainfall runoff transport of materials associated with human land activities. The common classes of NPS are (1) agriculture, (2) silviculture, (3) mining, (4) construction, (5) sources affecting groundwater, (6) urban and industrial water runoff, and (7) residuals. NPS reduction can often be achieved by changing management practices with the most cost-effective management practices referred to as best management practices (BMP's).

Nonpoint source impact must be quantified in order to develop cost-effective programs for water quality improvement because NPS have a significant effect on stream water quality. River basins, however, have such a large number of NPS inputs that complete spatial and temporal monitoring of streams draining these sources is impractical. Furthermore, judgment sampling, which employs monitoring at sites judged to be typical of an area by a qualified individual, may produce results which are uncharacteristic of the area due to unconscious investigator bias.

Probability (random) sampling utilizes statistical theory to estimate the mean and variance of a system. It provides several advantages over judgment sampling in assessing areawide water quality. It permits the use of statistical methods for making tests and inferences. The work is repeatable in that anyone can employ the same sampling at any time or place. Finally, because probability sampling methods are widely known to be objective, their use negates the potential criticisms of subjectivity and investigator bias.

The research study reported here presents the general methodology required to develop a statistically-based sampling plan and employs this methodology to develop a model sampling plan to assess the impact of rural nonpoint sources. The model plan illustrates procedures which have been successfully employed in a field monitoring program of rural water quality on an areawide basis.

2. GENERAL RANDOM SAMPLING PLAN DESIGN

Random sampling plans are designed according to well-established guidelines [2] but vary considerably because they are developed to provide the maximum information with the least effort (expense) for a given problem. The steps in plan development which lead to a random sample selection include setting objectives, defining a sampling unit, identifying the sampling universe, and drawing up a sampling frame. A review of these steps is presented because random sampling represents a methodology which can be applied to a broad range of water quality assessment problems.

Project goals need to be established first. This is often one of the more difficult tasks when divergent interests are involved and funds are limited. Next the items of interest or sampling unit must be defined and the sample universe which identifies the entire set of items to be investigated must be specified. These definitions need to be detailed enough to exclude unwanted data sources but not so restrictive as to severely limit the project conclusions. Listing all of the sampling units in the sample universe so that sample selection can be conducted with known probability is the step in the planning processes that defines the sampling frame. Finally, the sample which is a set of items selected at random from the sample universe to provide a basis for making inferences about the nature of the universe is drawn.

The method of drawing the sample which most effectively provides the required data needs to be specified. Two common sampling plans employ simple random sampling and stratified random sampling. Simple random sampling is selection of a sample in which the probability of drawing each possible sample of a given size is constant. With simple random sampling, the precision of the estimate may only be increased by an increase in the sample size.

Stratified random sampling is an alternative sample selection method for which the universe is divided into two or more inclusive but non-overlapping subsets. For each subset, a separate sample is selected and an estimate made. The advantages of stratification are numerous. For example, it permits the use of professional and technical knowledge to improve the efficiency of probability sampling. It also increases precision for the sample size (or allows the same precision at lower cost) because it reduces the variance of the estimate. This is achieved by reducing the variability due to average differences between parts of the sample universe and by allocation of increased sampling effort to the more important parts of the universe. Another advantage of stratification is that separate estimates for separate parts of the universe are provided. Finally, it can be employed to ensure a relatively even distribution of the sampling effort over all parts of the universe.

3. MODEL PLAN TO ASSESS RURAL NONPOINT SOURCES

To assess areawide rural water quality impacted by nonpoint sources, a sample site definition and statistical monitoring strategy were developed. Details of this project design are presented as an example for future studies to assess rural water quality. The model plan accounted for both spatial and temporal variations. The feasibility of employing it was tested by a demonstration study conducted in the 12,900 km^2 Chowan River Basin in the southeastern states of North Carolina and Virginia.

3.1. SAMPLING UNIT DEFINITION

A general sample site definition was developed to measure the impact of NPS on stream water quality. It employed maps of a specified scale to identify subbasins and their associated sampling site location. The identification of an NPS subbasin was as follows: start at the origin of a stream on a map and follow the stream to the road crossing that defines the single largest subbassin in the range of 1.3–44 km^2; do not count as a road crossing the situation where a bridge crosses an impoundment, but do count as a road crossing a road on top of a dam; omit a subbasin if the streams are marked "dry "or "intermittent." Subbasin area is estimated by outlining topographic boundaries. If topographic maps are not available, roads, railroads, or the mid-point between adjacent streams define subbasin boundaries. The sample site location for data collection is at the road or railroad crossing where surface drainage exits the subbasin.

A rural NPS subbasin must fulfill two further restrictions. It must be primarily rural and it must not contain any permitted point sources. Urban areas are differentiated from rural areas on the basis of city limits as indicated on maps.

This sampling subbasin definition defines heterogeneous land use subbasins which usually are large enough that the streams are flowing except during drought periods. Thus the proposed sample site definition is adequate for assessment of rural NPS impact on receiver stream quality, but a different definition could be developed to accomplish other goals. For example, smaller watersheds could be defined so that a particular land use could be examined in detail altough these streams may be dry except during and immediately after rainfall events.

The actual feasibility study employed 1 : 250,000 topographic maps of the Chowan watershed to identify NPS subbasins. The size distribution of NPS subbasins is summarized in table 1. The data show that the greatest number of subbasins were in the 5.2–20.8 km^2

Table 1

Small drainage basins* of the Chowan River system, with a preliminary stratification by area and whether rural or urban

Stratum	Range basin area (square km)	Stratum area percent of total	Number of drainage basins		
			Rural*	Urban*	Total
I	$\leqslant 1.6$	1.4	75	1	76
II	$1.6 < A \leqslant 12.9$	15.6	182	13	195
III	$12.9 < A \leqslant 24.1$	14.7	54	7	61
IV	$24.1 < A \leqslant 35.4$	13.3	27	8	35
V	$35.4 < A \leqslant 56.3$	16.1	23	7	30
Total		61.1	361	36	397
Remainder of Chowan Basin		38.9			

Total area of Chowan watershed basin 12,900 km^2

size class and that rural NPS subbasins accounted for 61 percent of the total watershed area. The data also showed that only about one-tenth of the subbasins were urban, indicating that the dominant land use was rural.

3.2. AREAL SAMPLING

Soil types and land use vary throughout the Chowan Basin so areal stratification of sampling sites was employed with four representative areas selected and sampled to allow spatial comparisons. These study areas included all typical land use and geoclimatic conditions present in the total watershed; but the sample universe represented only 25% of the total watershed area (fig. 1). There are two major physiographic regions located

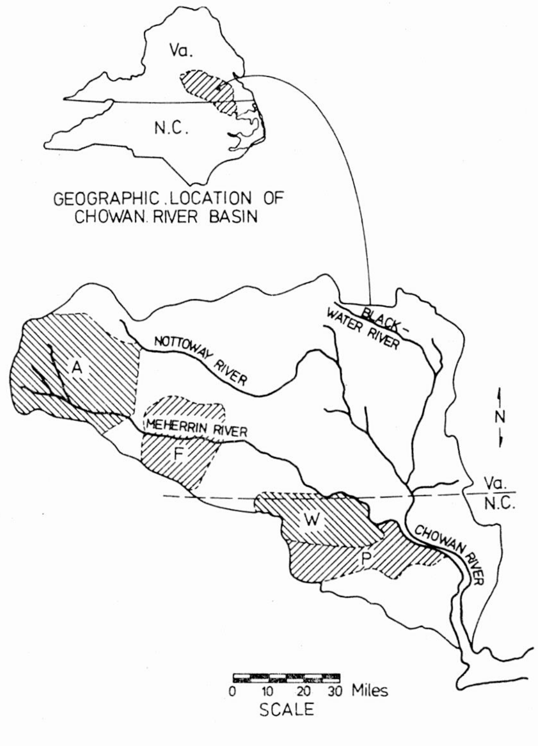

Fig. 1. Study location

A — agricultural piedmont, *F* — forested piedmont, *W* — well-drained coastal plain, *P* — poorly-drained coastal plain

Rys. 1. Miejsce badań

A — tereny podgórskie użytkowane rolniczo, *F* — zalesione tereny podgórskie, *W* — przybrzeżna równina dobrze odwadniana, *P* — przybrzeżna równina słabo odwadniana

in this basin — Piedmont and Coastal Plain. Within the Coastal Plain, the soils can be classified as either poorly-drained or moderately well-drained sands. Because the poorly--drained soils have a higher content of organic matter and conditions frequently favor denitrification, higher organic but lower nitrate-nitrogen loading were expected compared to the well-drained soils. The well-drained soils usually support a more intensive agricultural program and normally have greater slopes as compared to the poorly-drained soils.

For these reasons, the Coastal Plain was stratified into two areas where most of the soils were either poorly-drained or well-drained.

Nearly all the soils are loamy within the Piedmont area of the watershed. Thus there was no reason to stratify by soil type. Because water quality differences are likely to result with varying land use, a forested area was selected to represent background conditions and an agricultural area was selected to measure land use impact. Thus, the four strata were forested and agricultural Piedmont, and poorly- and well-drained Coastal Plain.

After assembling the spatial sampling frame (listing all sites which met the definitions of a rural NPS subbasin) for the study areas, 15 rural subbasins were randomly selected for monitoring. Three sites were selected from the agricultural Piedmont and four sites were selected from each of the other strata. Accurate watershed boundaries for each of the 15 subbasins were outlined on aerial photographs and field verified. Because land use was considered to be an important variable effecting stream quality, a summary of the subbasin land uses that were planimetered from the aerial photographs is presented in table 2. The land classification scheme employed row crops, pasture, forest, logged areas,

Table 2

Land use of Chowan study subbasins

Subbasin	Drainage Area sq km	Forest %	Crop %	Pasture %	Developed %	Logged %	Ponds %
			Silvicultural piedmont				
F-1	13.5	72.1	15.0	8.1	0.0	4.8	0.0
F-2	15.9	91.9	3.7	2.8	0.2	1.4	0.0
F-3	36.4	90.3	7.0	1.0	0.1	1.6	0.0
F-7	15.7	82.8	10.6	3.0	2.9	0.7	0.0
			Agricultural piedmont				
A-1	14.4	63.6	27.0	7.3	0.2	1.7	0.2
A-4	11.1	55.6	29.9	14.3	0.2	0.0	0.0
A-8	4.5	38.1	20.6	32.5	1.0	7.6	0.2
			Poorly-drained coastal plain				
P-8	11.7	77.1	17.4	4.2	1.0	0.3	0.0
P-10	9.7	72.1	25.9	2.0	0.0	0.0	0.0
P-11	12.7	69.0	23.0	3.0	4.0	1.0	0.0
P-13	98.5	66.2	26.4	4.7	1.5	1.1	0.0
			Well-drained coastal plain				
W-3	16.2	44.7	53.5	1.3	0.3	0.2	0.0
W-4	0.5	52.6	46.2	1.2	0.0	0.0	0.0
W-8	8.6	56.2	41.9	0.4	1.5	0.0	0.0
W-10	16.5	48.5	43.3	7.0	0.9	0.3	0.0

developed areas, and farm pond categories. To verify planimetering precision for each subbasin, the land use classes were summed and the total watershed was planimetered; all results agreed within 1%.

3.3. TEMPORAL SAMPLING

A sampling station was established at the surface outlet for each subbasin. Two common methods of monitoring streams are automated and grab sampling. Automated sampling at a site usually employs a stage recorder to allow flow estimates and a water sampling instrument which may be either stage or time activated. Although initial instrumentation costs may be relatively high, automated sampling provides a long term, continuous type record for interpretation of base flow and runoff conditions. Grab sampling is discrete sampling by field personnel designed to obtain a flow measurement and a water sample on each site visit. Either method may be employed in a statistical survey plan once the sites have been selected. Cost estimates for each sampling method with respect to data type and precision must be reviewed to determine the most effective scheme in terms of sampling method, number of sites, and samples from each site. Sufficient cost data is not generally available so investigator judgment must often be directive.

Grab sampling schedules should also utilize probability sampling. Again, stratification can be beneficial and random grab sampling can be time stratified to (1) provide data at a uniform rate during the study or (2) provide increased data retrieval during events with high variance. An example of the latter case is stratification to increase measurements during rainfall runoff events.

During the feasibility study, the 15 sites were grab sampled following a time stratified plan which ensured that data were obtained at a uniform rate for all seasons of the year. The time sampling universe consisted of all days including weekends and holidays, but the time of sampling was restricted to daylight hours for safety and practical reasons. The actual sampling days for each site were chosen by a restricted sampling method that permitted cost-effective travel routes and necessary rest intervals for equipment maintenance and data processing. Details of the sampling schedule design follow.

A 28-day interval was the basic time period for the feasibility study with two samples from each site per period. The 15 sites were assigned to three groups to reduce travel costs. For a sampling trip, each group was allocated four days with three days necessary for sampling and the fourth day for support type activities. From the 28-day period, one 4-day interval selected at random was set aside for rest. From the remaining days, the first three, 4-day time units were devoted to the first visit and the last three, 4-day time units to the second visit. Within a visit, the order of groups covered was selected by random permutation. Within a group, the order covering the sites was permutated among all reasonable travel routes.

4. SUMMARY AND CONCLUSIONS

Probability (random) sampling represents an effective methodology which can be used to quantify areawide water quality. It provides an unbiased estimate of the mean value and an estimate of the variance. It also permits statistical methods to be employed for making tests and inferences. Finally, the use of probability sampling negates possible criticism of investigator subjectivity impacting study results.

Probability sampling designs are developed to best meet the requirements of a particular problem, but some general principles are common to all sampling designs. These principles were reviewed to demonstrate the potential benefits of random sampling for water quality assessments. Thus the concepts of a sample universe, sampling frame, and drawing a sample (by either simple random and stratified random sampling) were discussed.

A model probability sampling plan to measure areawide rural water quality was presented that provided definitions and practical methods for basic concept implementation. This sampling plan was designed to permit comparison of four geoclimatic, land-use areas of the feasibility study watershed and to assess the temporal variation of the streams. The model plan can readily be employed or modified for similar or different needs to assess areawide water quality or specific point and nonpoint sources.

In conclusion, nonpoint sources can represent a significant impact on streams and impoundments. Sampling techniques traditionally employed to monitor point sources may not be the most cost-effective to assess nonpoint sources because of the diffuse characteristic and temporal variation associated with NPS. Probability sampling represents a valuable tool which can easily be employed to assess nonpoint source impact, and water quality planning agencies should consider developing and implementing statistical sampling plans to assess areawide water quality.

Data analysis techniques to quantify areawide water quality are reviewed utilizing data from the feasibility study in the subsequent Part II article.

REFERENCES

[1] Bliven L. F., Koehler F. A., Horney, L. F., Overcash M. R., Humenik, F. J. *Areawide Assessment of Rural Stream Water Quality*, Part II, *Data Analysis Techniques*, Environ. Prot. Engr., Vol pp.

[2] Snedecor G. W., Cochran, W. G., *Statistical Methods*, 6th edition, Iowa State University Press, Ames, Iowa, U.S.A. 1974.

OKREŚLENIE WPŁYWU ZANIECZYSZCZEŃ OBSZAROWYCH NA JAKOŚĆ WODY W ZLEWNIACH ROLNICZYCH

I. METODOLOGIA STATYSTYCZNA POBORU PRÓB

Wskutek działalności człowieka wody powierzchniowe są często zanieczyszczane przez związki chemiczne. Planowanie jakości wód w całej zlewni musi uwzględniać zarówno punktowe jak i niepunktowe źródła zanieczyszczeń oraz koszty ich unieszkodliwiania, po to aby określić najbardziej efektywny plan uzyskania stanu czystości wód w regionie. W artykule niniejszym przedstawiono przydatność techniki poboru prób zgodnie z teorią prawdopodobieństwa (próby wyrywkowe) oraz metody statystycznej analizy wyników — jako sposób szacowania i określenia jakości wód rzeki w zlewni o charakterze rolniczym.

Cel pracy jest przedstawiony na podstawie danych z badań nad ustaleniem wpływu niepunktowych źródeł zanieczyszczeń wód w zlewni rolniczej o powierzchni 12900 km^2. Niepunktowe źródła zanieczyszczeń zmieniają się w czasie i przestrzeni jako funkcja wielu procesów, zmiennych w obszarze zlewni, podlega-

jących kontroli człowieka (uprawy rolne, roboty budowlane, oczyszczanie wód deszczowych z miast) lub jej nie podlegających (rodzaj gleby i pogoda). Stąd adekwatne określenie jakości wód w regionie wymaga planowania uwzględniającego zarówno zmiany przestrzenne jak i chwilowe. W części I pracy omówiono możliwość poboru prób (na podstawie teorii prawdopodobieństwa) w celu ilościowego określenia nowych i dotąd nieuchwytnych elementów ochrony wód, oraz przedstawiono modelowy plan poboru prób dla określenia jakości wód w całej zlewni. W następnej części przedstawi się technikę analizy wyników otrzymanych zgodnie z omówionym wyżej planem statystycznego poboru prób oraz ich interpretację pod kątem ustalenia wpływu przestrzennych źródeł w zlewniach rolniczych na jakość wód.

DIE BESTIMMUNG DES EINFLUSSES VON FLÄCHENARTIGEN VERUNREINIGUNGEN AUF DIE WASSERQUALITÄT IN LANDWIRTSCHAFTLICHEN EINZUGSGEBIETEN

I. STATISTISCHE METHODOLOGIE DER PROBENAHME

Aufgrund der vielseitigen Tätigkeit der Menschen leidet die Qualität der Oberflächengewässer, die oft mit chemischen Substanzen überlastet werden. Die Planung der Wasserqualität in ganzen Einzuggebieten muß sowohl den Einfluß der punkt- und flächenartigen Verunreinigungen sowie die Kosten deren Beseitigung in Betracht ziehen. Dies its für die Bestimmung des best-wirtschaftlichen und effecktiven Planes der Erhaltung entsprechender Ziele und Gütenormen des Wassers in der ganzen Region von ausschlaggebender Bedeutung.

Der vorliegende Beitrag beschreibt die Brauchbarkeit der Technik der Probenahme gemäß den Gesetzen der Wahrscheinlichkeitsrechnung (Stichproben) und der statistischen Analyse der Ergebnisse — als Alternative zur Methodologie der Abschätzung und Bestimmung der Wasserqualität im Flußeinzugsgebiet landwirtschaftlichen Charakters.

Der Ziel der Arbeit wird am Beispiel eines Einzungsgebietes von 12 900 km^2 Fläche mit flächenartigen Verunreinigungsquellen veranschaulicht. Die nicht-punktartigen Verunreinigungsquellen verändern sich in Zeit und Raum als Funktion verschiedenartiger Prozeße. Dies sind nicht nur Variablen im Raum des Flußeinzugsgebiets die einer Kontrolle unterliegen (Ackerbau, Bautätigkeit, Abwasserreinigung) sondern auch solche, die einer Kontrolle nicht unterliegen (Bodenarten, Wetterverhältnisse u.a.).

Eine genaue Bestimmung der Wasserqualität ist planungsbedürftig, wobei sowohl Raumänderungen wie auch momentane Veränderungen berücksichtigt werden müssen.

Im ersten Teil der Arbeit wird die Möglichkeit einer Probenahme (gemäß der Wahrscheinlichkeitsrechnung) zur quantitativen Bestimmung neuer und bisher nicht erfaßten Elemente des Gewässerschutzes beschrieben. Dieser Teil beinhaltet auch einen Modellplan der Probenahme für die weiträumige Bestimmung der Wasserqualität eines ganzen Einzugsgebietes.

ОПРЕДЕЛЕНИЕ ВЛИЯНИЯ АРЕАЛЬНЫХ ЗАГРЯЗНЕНИЙ НА КАЧЕСТВО ВОДЫ НА СЕЛЬСКОХОЗЯЙСТВЕННОЙ ВОДОСБОРНОЙ ПЛОЩАДИ

I. СТАТИЧЕСКАЯ МЕТОДОЛОГИЯ ОТБОРА ПРОБ

В результате деятельности человека качество поверхностных вод часто находится под влиянием повышенной нагруженности химическими соединениями. В планировании качества воды на всей водосборной площади должно учитываться влияние как точечных, так и неточечных источников загрязнений, а также расходы по их обезвреживанию для определения наиболее экономически эффективного плана достижения конкретных целей и стандартов чистоты воды в регионе.

В настоящей статье показана пригодность техники отбора проб основанной на теории вероятностей (выборочный контроль); представлены также методы статистического анализа результатов — как альтернативная методология оценки и определения проблем качества воды в бассейне реки на сельскохозяйственной территории.

Цель работы заключается в определении влияния неточечных источников загрязнений на качество воды в сельскохозяйственном водосборном бассейне площадью в 12 900 км^2. Названные источники меняются во времени и пространстве, представляя собой функцию многих процессов, тоже изменяющихся на территории водосборной площади и поддающихся (обработка почвы, строительные работы, очистка городских дождевых вод) или не поддающихся контролю (вид почвы, метеорологические факторы). Поэтому адекватное определение качества воды в регионе требует планирования, учитывающего как пространственные, так и временные изменения. В I-й части работы обсуждена возможность отбора проб (по теории вероятностей) для количественного определения новых и до сих пор неуловимых элементов защиты воды от загрязнений; представлен также модельный план отбора проб для всестороннего определения качества воды на всей водосборной площади. Во II-й части описана техника анализа результатов, полученных по вышеобсужденному статистическому плану отбора проб, и их интерпретации с точки зрения определения влияния пространственных источников на сельскохозяйственной водосборной площади на качество воды.