

# Contents

Abstract .....	1
Acknowledgment .....	1
Part 1: A theoretical and practical introduction to SPARROW.....	2
1.1 Introduction.....	2
1.2 SPARROW modeling concepts.....	3
1.2.1 Model objectives .....	3
1.2.1.1 Water-quality description.....	3
1.2.1.2 Contaminant source analysis.....	4
1.2.1.3 Water-quality simulation.....	5
1.2.1.4 Hypothesis testing: examining the importance of explanatory factors and processes .....	5
1.2.1.5 Design of sampling networks.....	6
1.2.2 Mass balance approach.....	7
1.2.3 Time and space scales of the model.....	9
1.2.4 Accuracy and complexity of SPARROW models.....	11
1.2.5 Comparison of SPARROW with other watershed models .....	15
1.3 SPARROW model infrastructure .....	18
1.3.1 Monitoring station flux estimation .....	19
1.3.1.1 Model specification for monitoring station flux estimation .....	21
1.3.1.2 Monitoring station flux prediction (advanced).....	24
1.3.1.3 Mean detrended flux (advanced).....	26
1.3.1.4 Tools for flux estimation.....	28
1.3.1.5 Guidance on station and record selection.....	30
1.3.1.6 Guidance for specifying monitoring station flux models .....	31
1.3.1.7 Related topics .....	32
1.3.1.8 Mean flow-weighted concentration (advanced).....	32
1.3.1.9 Mean time-weighted concentration (advanced).....	33
1.3.2 Stream network topology.....	34
1.3.3 Watershed sources and explanatory variables .....	37
1.4 Model specification.....	39
1.4.1 Model equation and specification of terms .....	39
1.4.2 Contaminant sources .....	42
1.4.3 Landscape variables .....	46
1.4.4 Stream transport.....	48
1.4.5 Reservoir/lake transport.....	54
1.4.6 Regional model coefficients and nested model designs .....	57
1.5 Model estimation .....	59
1.5.1 A guide to nonlinear estimation .....	59
1.5.1.1 Brief review of ordinary least squares .....	59
1.5.1.2 Nonlinear weighted least squares .....	61
1.5.1.3 The asymptotic covariance matrix.....	64
1.5.1.4 Estimation of leverage in the nonlinear model (advanced).....	64
1.5.1.5 Estimation of gradients (advanced) .....	65
1.5.2 Asymptotic properties of the estimators (advanced).....	66
1.5.2.1 Implications of asymptotic normality (advanced).....	68
1.5.2.2 Asymptotic properties within a single basin—three examples (advanced).....	69
1.5.3 Coefficient bias and uncertainty—additional issues .....	74
1.5.3.1 Bootstrap estimate of coefficient bias (advanced).....	75
1.5.3.2 Bootstrap estimate of the coefficient covariance matrix (advanced) .....	76

1.5.3.3	Bootstrap coefficient confidence interval (advanced).....	76
1.5.3.4	Discussion of bootstrap methods for coefficient estimation.....	77
1.5.3.5	Measurement error (advanced) .....	78
1.5.4	Evaluation of the model parameters.....	80
1.5.4.1	Statistical evaluations .....	80
1.5.4.2	Physical interpretations.....	82
1.5.4.3	Statistical insignificance and multicollinearity.....	84
1.5.5	Evaluation of model errors .....	90
1.5.5.1	Heteroscedasticity.....	90
1.5.5.2	Spatial biases.....	94
1.5.5.3	Statistical outliers .....	95
1.5.6	Measures of model performance and fit.....	96
1.5.7	Non-nested tests (advanced).....	98
1.6	Model predictions.....	99
1.6.1	Prediction equation .....	99
1.6.2	Parametric predictions .....	101
1.6.3	Bias-corrected predictions based on bootstrapping (advanced).....	102
1.6.4	Prediction standard errors based on bootstrapping (advanced).....	105
1.6.5	Prediction intervals based on bootstrapping (advanced).....	107
1.6.6	Discussion.....	111
1.6.6.1	The effect of error in measured flux (advanced).....	112
1.6.7	Share of flux delivered to a target reach.....	113
1.7	Summary.....	115
1.8	References.....	116
Part 2: SPARROW User's Guide.....		123
2.1	Introduction.....	123
2.2	System requirements .....	123
2.3	Obtaining and installing software .....	123
2.4	Input/output structure .....	125
2.5	Navigating in SAS for Windows .....	127
2.5.1	The basic workspace.....	127
2.5.2	Active windows and menus.....	128
2.5.3	Opening SAS program files.....	128
2.5.4	Viewing SAS data files.....	130
2.5.5	Moving around in the SAS Explorer window .....	130
2.6	Model input.....	131
2.6.1	Data file .....	131
2.6.1.1	Reach topology.....	133
2.6.1.2	Reach attributes .....	133
2.6.1.3	Contaminant flux.....	133
2.6.2	Geographic Information System (GIS) base maps (optional).....	134
2.6.3	Control file.....	134
2.6.3.1	Directory and input data .....	135
2.6.3.2	Bootstrap iterations and seeds (advanced) .....	136
2.6.3.3	Model specification .....	138
2.6.3.4	Process specification .....	144
2.6.3.5	Advanced process specification .....	150
2.6.3.6	Additional variable definitions .....	155
2.6.3.7	Options for model execution .....	158
2.6.3.8	Data modifications .....	162
2.7	Executing the model.....	164

2.8	Model output.....	166
2.8.1	Estimation output.....	167
2.8.1.1	Nonlinear optimization results and diagnostics.....	167
2.8.1.2	Coefficients and diagnostic statistics for the nonlinear weighted least squares model.....	172
2.8.1.3	Graphical output.....	173
2.8.1.4	Estimation output data files.....	179
2.8.1.5	Summary estimation output.....	182
2.8.2	Prediction output.....	182
2.8.2.1	Summary results.....	182
2.8.2.2	Prediction output data files.....	184
2.8.3	Other file output.....	187
2.8.3.1	Ouput file "comments_all".....	187
2.8.3.2	Text file output (optional).....	187
2.8.3.3	Bootstrap intermediate files (optional).....	187
2.8.4	GIS maps (optional).....	188
2.8.4.1	Residuals map.....	188
2.8.4.2	Reach map.....	191
2.9	Common execution errors and diagnostic tests.....	193
2.9.1	Errors during data preparation.....	193
2.9.2	Estimation execution errors.....	195
2.9.2.1	Estimation execution errors caused by systematic errors in input data.....	195
2.9.2.2	Estimation execution errors caused by numerical overflow—using the test-calibration mode.....	197
2.9.2.3	Estimation execution errors related to bootstrap analysis.....	198
2.9.3	Prediction execution errors.....	199
2.9.3.1	Test-prediction mode.....	199
2.9.3.2	Evaluation of summary table of reach predictions.....	200
2.10	References.....	202

## Appendix

A.	Determination of the Bootstrap Confidence Interval Quantiles.....	203
B.	Hydrologic Network Development.....	204
C.	SAS/GIS Mapfile Creation.....	208
D.	Descriptions of Output Files.....	217
D.1	Estimation Output File "summary_betaest".....	217
D.2	Estimation Output File "cov_betaest".....	221
D.3	Estimation Output File "resids".....	222
D.4	Estimation Output File "boot_betaest_all".....	226
D.5	Estimation Output File "test_resids".....	227
D.6	Prediction Output File "predict".....	228
D.7	Prediction Output File "summary_predict".....	239
D.8	Prediction Output File "lu_yield_percentiles".....	239
D.9	Prediction Output File "test_data".....	239
D.10	Prediction Output File "test_predict".....	243
D.11	Prediction Output Files with Bootstrap Intermediate Results ("store_[variable_name]").....	247

## List of Tables

1.1	Station characteristics and model performance data for national and regional SPARROW total nitrogen models .....	12
1.2	Monte Carlo evaluation of Fluxmaster and LOADEST 2000 annual flux estimates .....	29
1.3	Factors affecting the percent standard error of total nitrogen (TN) and total phosphorus (TP) flux estimates at 700 National Stream Quality Accounting Network monitoring stations.....	30
1.4	Potential types of contaminant-source variables in SPARROW models.....	44
1.5	SPARROW estimates of model statistics for the United States national total nitrogen illustration data set.....	81
1.6	Model coefficient results for two correlated temperature land-to-water delivery factors in the national total nitrogen model .....	87
2.1	Input variables required for SPARROW model execution, evaluated for each stream reach.....	132
2.2	Estimation output files .....	180
2.3	Prediction output files .....	185

## List of Figures

1.1	Characteristics of the national and regional SPARROW total nitrogen models .....	12
1.2	Spatially interpolated mean-annual estimates of wet nitrate deposition, based on measurements from National Atmospheric Deposition Program (NADP) sites.....	13
1.3	A simple continuum of model types based on the level of statistical and mechanistic descriptions of contaminant sources and biogeochemical processes.....	15
1.4	Hypothetical illustration of the relation of model accuracy to model complexity .....	17
1.5	Schematic of the major SPARROW model components .....	18
1.6	A graphical depiction of a detrended series.....	20
1.7	Monte Carlo analysis comparing flux estimates from the programs Fluxmaster and LOADEST 2000.....	29
1.8	Schematic illustrating a vector stream reach network with node topology and water/contaminant reach-node routing table.....	34
1.9	Digital river reach networks for the northeastern United States .....	36
1.10	Schematic illustrating digital overlay of stream reach drainage area and polygonal areas associated with diffuse sources.....	38
1.11	Schematic illustrating digital overlay of stream reach drainage area and polygonal areas associated with landscape properties.....	38
1.12	Conceptual illustration of a reach network for five incremental watersheds.....	40
1.13	Regional nitrogen budget for watersheds of the northeastern United States .....	42
1.14	Examples of intensive and extensive measures of nitrogen sources in United States watersheds.....	43
1.15	Mapped hydrologic soils groups from the STATSGO database reflecting the water infiltration and saturation properties of soils .....	46
1.16	Estimates of in-stream nitrogen removal from SPARROW and experimental studies.....	49
1.17	Estimates of in-stream total nitrogen removal from SPARROW .....	51
1.18	Nutrient removal in lakes and reservoirs of the Waikato River Basin, New Zealand, expressed as a fraction of the nutrient inputs to the water body in relation to the areal hydraulic load for total phosphorus and total nitrogen.....	55
1.19	Completed and active SPARROW regional modeling applications.....	58
1.20	A graphical depiction of a simple hydrologic system consisting of a single reach segment of length $d_r$ divided into $n$ sub-segments of length $d_s$ , each having a measured logarithm of flux $y_s$ .....	70
1.21	The arrangement of reaches, sources and monitoring stations in a two-dimensional hydrologic model.....	71

1.22	Estimated municipal wastewater treatment coefficient in the national and regional SPARROW models .....	84
1.23	SAS output showing the eigensystem from an example SPARROW model .....	89
1.24	Diagnostic plots for evaluating SPARROW model errors and adherence of the residuals to the model assumptions .....	91
1.25	Map of prediction errors for the SPARROW total nitrogen model plotted for the stream monitoring stations in the North Carolina coastal watersheds.....	94
1.26	Map of prediction errors for the SPARROW national total nitrogen model plotted for the stream monitoring stations, as displayed by the SAS-based SPARROW modeling software .....	95
1.27	Plots of predicted and observed total nitrogen flux and yield for the New Zealand Waikato River Basin SPARROW model.....	96
1.28	SPARROW model estimates of in-stream nitrogen delivered to the Gulf of Mexico.....	114
2.1	SPARROW directory structure .....	125
2.2	SPARROW input/output structure .....	126
2.3	Basic features of the workspace, SAS v8 for Windows .....	128