PART 2

~~~~~~~~~~~~~~~~

Aquatic Biodiversity Information

~~~

Freshwater Flora and Fauna: Species Diversity, Community Structure and Ecology

~~~

Classification of Watersheds and Freshwater Systems

~~~

5. AQUATIC BIODIVERSITY INFORMATION IN MAINE

This chapter reviews the information that contributed to MABP's data compilation and which, in turn, was used to produce the data summaries and analyses presented in this report. The chapter also addresses the issue of how sampling effort influences what we know about biodiversity in different parts of the State, and where major data gaps exist today. The chapter concludes with a brief description of how the MABP database can be accessed by anyone interested in further use of the data.

5.1 Information Resources: An Overview

A central task of MABP was data compilation and standardization. At the outset of the project, all of the information eventually used during the project was highly dispersed and stored in a bewildering array of database formats, hardcopy reports and field data sheets. Data were compiled into a relational database in order to facilitate integration, analysis and interpretation (see Appendix 11.1 for a review of accessed data sets). Many of the older data sets had to be digitized in order to include them in the MABP database. Examples include (i) the Cooper and Fuller surveys of Maine lakes from the 1930s and 1940s (of particular value in terms of evaluating species introductions over the past half century), (ii) several studies of stream macroinvertebrates conducted in northern and eastern Maine during the 1970s and 1980s (e.g. Mingo et al. 1979, Mingo and Gibbs 1980), and (iii) stream electrofishing surveys conducted by MDIFW. Regardless of whether or not data were newly computerized during MABP, all data sets had to be "massaged" to ensure that they would be compatible to the unifying design of the MABP database. Standardization of key fields was a major challenge since there was considerable variation in how taxa were named and how townships and other spatial identifiers were labelled. The range of data analyses and information syntheses presented in this report would have been impossible without this process of data standardization and integration.

All data were geo-referenced to the extent possible. Ideally, source data contained coordinates or waterbody names and/or codes. Where they did not, we generated estimated site coordinates if the site descriptions permitted this. In some cases, it was not possible to reference data at a level finer than, for example, township. Variation in the scale at which data were geo-referenced influences the scale at which different data sets are summarized and presented in this report. For example, the amphibian and reptile data are summarized by township since the primary data source for this taxonomic group identified sites only by township. Other amphibian and reptile data, and for which point coordinates were available, were entered into the database with these coordinates, but were integrated into data summaries at the township level.

The MABP database includes over 15,500 distinct "sites" from across Maine. A site may be a point, associated with data from one or more studies. Alternatively, the site may be a polygon, again associated with data from one or more studies. Examples of polygons include lakes and townships (available for mapping either as a polygon or as a centroid, i.e. center point). Figure 5.1 shows the distribution of sites associated with data from the four major taxonomic groups covered by MABP. All polygon-based sites are shown as points for the purpose of this figure.

MABP used four broad categories of data (Table 5.1). Just under 1/3 of the data sets were derived from survey and/or monitoring studies. These included several extensive data sets made available by state and federal agencies, including (i) the MDIFW lake fish data, (ii) macroinvertebrate data from MDEP's stream biomonitoring program, (iii) fisheries data from EPA's Environmenatal Monitoring and Assessment Program (EMAP), (iv) MDIFW's odonate, mussel and amphibian/reptile survey data, and (v) MNAP rare plant database. A second data category included student dissertations and theses (mainly from the University of Maine). A third category consisted of existing data compilations focusing on, for example, gastropods, stoneflies and various dipteran groups. Many of these compilations appear in the scientific literature; a few were obtained via on-line access. We also include in this category data sets obtained (or accessed) from museum collections, including the University of Maine's herbarium and the University of New Hampshire's insect collection. Museum records were not examined individually during MABP, although some of the data that we used resulted from reviews of museum records by other researchers, for example P. Brunelle for odonates (Brunelle 1999, and later unpublished data), and S. Burian for mayflies (Burian and Gibbs 1991 and later unpublished data). Relatively few environmental impact studies were used as data sources during MABP. This was for two reasons. First, these studies often use data that are available from other (primary) data sources. Second, it was difficult to find reports generated from many of these studies and extremely time-consuming to extract the data.

The primary criterion for accessing a data set during MABP was that it contained "useful" taxonomic data, i.e. species- or genus-level records. Generally, data sets in which the fauna or flora were identified only at the family level (or above) were not used. However, some data sets contain records spanning a range of taxonomic resolution. In this case, records identified only at a supra-generic level were also included. Some data sets were included in the MABP database even though they did not provide any new taxonomic information for a particular waterbody. For example, a number of fishery-focused dissertations did not contribute any additional species to the checklists provided through the main MDIFW lake fish database. However, these dissertations did address ecological issues that were relevant to MABP's theme of examining biodiversity in Maine. As with the number of data sets, the number of individual records in the MABP database is skewed toward fish, with aquatic insects being the second most-represented group (Table 5.2). Approximately 1/3 of all invertebrate records consisted of taxa identified only to the level of genus.

Table 5.1: Synopsis of data sets contributing to the MABP database

Taxonomic Group	Number of Data Sets (1)
Vascular plants	16
Macro-invertebrates	79
Fish	113
Amphibians	14
Reptiles	5
Study Category	**Number of Data Sets**
Survey / monitoring	79
Thesis (2)	93
Compilation of data from other sources	20
Environmental impact study	3
Other research	14
# Location IDs Contained in Data Set (4)	**Number of Data Sets**
< 10	124
10 – 99	57
100 – 999	18
1000 +	5

(1) Some data sets contain information on > 1 taxonomic group.
(2) Some theses contain survey-type data.
(3) Each taxon record in the MABP database is assigned a Location Identifier which, depending on the spatial information available in the data source, refers to a point location, a lake or stream segment, or undefined site(s) in a town, county or the state. The number of Location IDs in a data set thus reflects both the spatial resolution of the location identifiers and the number of defined sampling sites.

Table 5.2: Summary, by major taxonomic group, of records in the MABP database where taxa have been identified at the level of genus or species.
In this table, an individual record consists of one taxon at one location in one data set. Records are summed over all data sets, locations and species.

Taxonomic Group	# Records at Level of:	
	Genus	Species
Platyhelminthes (flat worms)	56	139
Annelids (segmented worms & leeches)	384	1,066
Malacostraca (crayfish & allies)	280	622
Bivalvia (freshwater mussels & allies)	222	3,650
Gastropoda (snails)	671	541
Insecta	27,011	43,141
Arachnida (mites)	209	3
Fish	210	101,281
Amphibians	--	2,831
Reptiles	--	1,581
Bryophytes (mosses)	11	--
Chlorophyta (green macroalga Chara)	83	--
Vascular plants	179	11,525

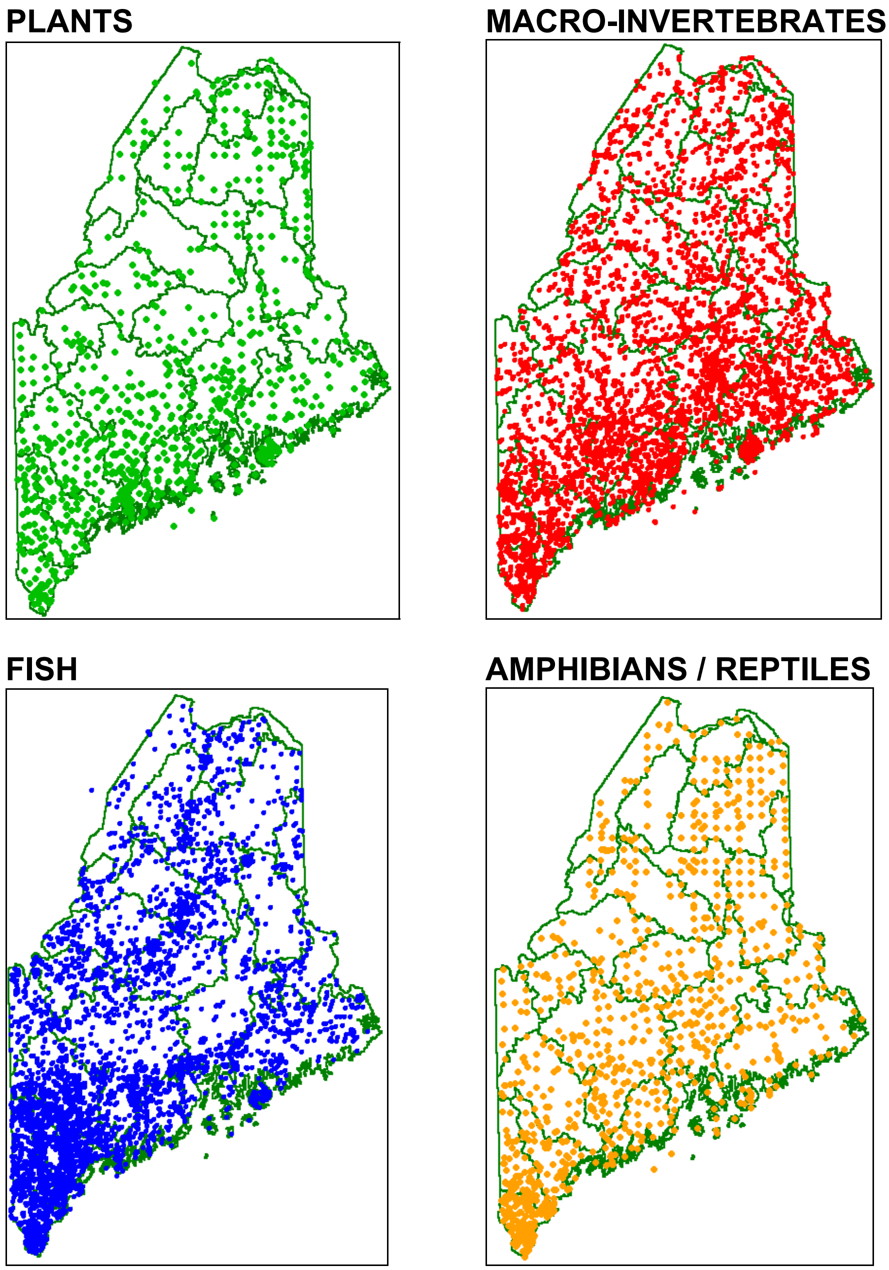


Figure 5.1: Sites with taxon records in the MABP database, by major group.
Dots indicate true point locations or centers of polygons (e.g. lake, town) depending on spatial referencing available in source data sets. Green lines indicate HUC-8 watersheds.

5.2 Data for Assessing Status and Trends

A number of data sets are particularly relevant from the perspective of documenting status and trends in Maine's freshwater biodiversity These data sets either represent broad-based surveys and/or derive from sampling designs that have produced quality assured, quantitative data. Some examples follow – all data are discussed in greater detail in Chapter 6.

Amphibians and Reptiles: The Maine Amphibian and Reptile Assessment Project (MARAP) amassed a wealth of information relating to species distributions and relative abundance (Hunter et al. 1999). Although completed in 1992, the database continues to grow as a result of ongoing MDIFW ecoregional surveys and other studies. Monitoring of frogs and toads by the Maine Amphibian Monitoring Project (MAMP) is currently yielding quantitative data that will provide an important tool for evaluating current status of populations and trends through time.

Fish: There are more data on fish than any other taxonomic group in Maine. As of early 2004, a total of 2154 lakes had been surveyed by MDIFW (Figure 5.2). The data used by MABP consist of species lists for each lake (updated every 1-2 years), together with an assignment of "fishery importance" for each species in each system. Additional lakes information collected by MDIFW include length-weight data, age determinations and population estimates for target fish species, for example brook trout and landlocked salmon. For some waters, there are extended time series for such data. For example, in western Maine, the database for nine lakes spans more than 10 years (F. Bonney, MDIFW, pers. comm.). While of great importance from a fisheries management perspective, these population-level fisheries data were not central to the core mission of MABP; the data were not included in the MABP database and are used in only a few of the analyses presented in this report. In addition to the lake fish inventory, MDIFW maintains an extensive database of stocking records for both lakes and streams, extending back to 1937. MABP has accessed these data – they are reviewed in Chapter 6.4.

The amount of survey effort that serves as the base for generating the MDIFW lake fish species lists varies among lakes. Approximately 50% of lakes smaller than 100 acres have been fully surveyed only once (Figure 5.3). For most of the 1042 lakes with only one survey, these surveys occurred in the 1950s and 1960s (Figure 5.3). However, it is important to note that MDIFW adds fish species to their inventory list as new records become available, i.e. even without a full survey. Therefore lakes that have only old surveys often have a relatively current species list. However, the extent to which this updating process occurs is variable and depends on a variety of factors, such as how familiar biologists are with a particular lake and/or how much 'anecdotal' information they receive from anglers.

While most of the larger lakes in Maine have some fisheries data, unsurveyed lakes are more common among the smaller size classes (Table 5.3). Generally 40% or more of lakes smaller than 100 acres have not been surveyed at all. This analysis is based on MIDAS-numbered lakes which under-represent the smallest size classes (Table 2.1). Thus ponds smaller than 10 acres are even less-well sampled than the data in Table 5.3 would appear to indicate.

A limitation of the MDIFW lakes fish species database is that it stores limited time series information. Species lists are simply updated on an annual basis and, prior to about 2000, new species records were not tagged with dates (T. Obrey, MDIFW, pers. comm.). This means that it is difficult to evaluate trends through time (in species composition) without reviewing individual hard-copy field data sheets in regional office archives (which was not attempted during MABP, although is being done for several lakes as part of at least two on-going student dissertations at the University of Maine). Lakes sampled by G. Cooper and colleagues during the 1930s and 1940s represent an exception to the problem of few time-series data. Cooper surveyed 205 lakes, primarily in southern and central Maine, for fish, benthic invertebrates, plankton and water quality (Figure 5.4) (Cooper 1939, 1940, 1941, 1942; Cooper and Fuller 1945; Fuller and Cooper 1946). While these surveys probably did not fully document all non-game fish species, they do represent an invaluable source of information to evaluate species changes over the past 60 years. Cooper survey data were digitized by MABP and integrated into the MABP database.

As mentioned above, the MDIFW lake fish species lists derive from variable amounts of survey effort and do not include any measure of relative abundance. The only study that has attempted to sample lakes with a standardized protocol, using multiple collection methods in a defined sampling "window", was EPA's Environmental Monitoring and Assessment Program (EMAP), conducted during the 1990s. EMAP covered the northeastern U.S., and selected survey lakes on a probabilistic basis, which permitted the extrapolation of data to regional status assessments. Fifty four EMAP lakes are in Maine (Figure 5.6). The EMAP data are included in the MABP database.

Access to stream fish data was more of a challenge during MABP than was the case for lakes data. For most of the duration of MABP, the MDIFW stream survey data were available electronically for only part of the State (Figure 5.5). Data from southern Maine were already digitized at the start of MABP, although georeferencing had not been completed at that time. Data from other regions were digitized by MABP in collaboration with MDIFW --this process was incomplete as of mid-2004 but is scheduled for completion in 2005. MDIFW stream fish data tend to be quite heterogeneous, with various levels of sampling effort and a variable emphasis on recording non-game species. An exception to this pattern is MDIFW's brook trout monitoring study which has sampled a series of stream sites across the State since 1990 (Figure 5.6). The period of record is variable among sites and non-game species were not always fully recorded in the early years of the study. This study provides an excellent data resource for documenting status and trends in stream fish populations because data are collected using a standardized protocol and parallel habitat / water quality information is obtained.

None of the stream sites sampled by MDIFW were selected probabilistically and thus it is important to consider the degree to which they are representative of the overall population of streams in various parts of the State. Figure 5.7 attempts this comparison using data from four MDIFW regions: south, central, downeast and west. MDIFW sites tended to be skewed toward smaller streams (smaller watershed areas). In particular, very few of the sampled stream segments drained an area of > 100 square miles (Figure 5.7).

Large rivers have, as suggested above, been under-surveyed in Maine. A recent study, part of an effort designed to develop an Index of Biotic Integrity for rivers in the northeastern U.S., has begun to correct this disparity. Sponsored by EPA, this study has used electrofishing to survey the Androscoggin, Presumpscot, Kennebec, Sebasticook and Penobscot Rivers. Final data are not available at this time, but Figure 5.8 depicts sampling sites (for all rivers except for the Penobscot).

The Atlantic Salmon Commission also does some sampling on larger (as well as smaller) rivers. However, this sampling effort focuses on salmon passage through fishway traps located on dams and fish weirs. While other trapped species are generally recorded, this work does not provide an assemblage-level picture of river fish (nor is it intended to do so). The ASC also conducts extensive electrofishing surveys of the smaller salmon rivers, but the extent to which non-salmon species are recorded in the data has been highly variable.

Overall, most of the MDIFW and ASC stream fish data derive from a single year at any one site (Table 5.4). Data records extending over 5 years are available from 73 sites in the State. Eight sites have been monitored in the MDIFW brook trout monitoring study for 15 or more years.

The EPA's New England Wadeable Streams (NEWS) Project is an EMAP-type study focusing on streams in which sites are selected probablistically and sampled with standardized protocols. A partial suite of sites sampled for this study appears in Figure 5.6. Locations of the most recently sampled sites were not available prior to preparation of this report. It is likely that the NEWS project will provide valuable data for rigorously comparing stream fish assemblages, both within Maine and across the northeast.

Macroinvertebrates: Freshwater mussels, odonates and (to a lesser degree) mayflies have been extensively surveyed in Maine (see Chapter 6.5-6.7). The odonate and mayfly surveys have incorporated in-depth reviews and compilations of historical material. Species lists for odonates and mayflies will likely continue to grow as additional locations are sampled.

Lake benthos data are available from four studies (Figure 5.9), although the resolution at which the taxa are identified is variable – the Cooper data, in particular, are of relatively coarse taxonomic resolution. Three of the four studies tended to focus on larger lakes, whereas the EMAP lakes were probabilistically selected and thus more representative of the overall lake population. Unfortunately, the EMAP benthos data were unavailable to MABP. Courtemanch's (1982) study focused on profundal chironomids, not the entire suite of benthic invertebrates.

Twenty three lakes in Maine, located entirely in the southern half of the State, have assemblage-level data for all of three major taxonomic groups: benthic invertebrates, fish and plants (Figure 5.10).

The primary source of information for stream macroinvertebrates is MDEP's biomonitoring program. This program uses macroinvertebrate communites to evaluate the extent to which water quality goals are being attained in various stream and river segments. Sampling sites are not selected probabilistically, but rather on the basis of site/regional prioritization. Different regions in the State are generally sampled on a five-year rotational basis (Davies et al. 1999). Through 2002, data are available from a total of 540 sites (Figure 5.11). Most of the sites have been sampled for only one year, although many others have data for up to five years. Longer time series are available for about ten sites (Figure 5.11, Table 5.4). The value of these data is enhanced by the fact that they are collected with standardized methodology (Davies et al. 1999) and samples are processed with an emphasis on taxonomic consistency. In addition to the MDEP data, stream macroinvertebrate data in the MABP database derive from a number of sources, including student dissertations. The studies of Siebenmann (1995) and Davies (1987) are particularly worth noting here since both address temporal changes in stream invertebrate assemblages over approximately 10-year periods for the Narraguagus and Penobscot Rivers, respectively.

Plants: Plant data are available from a relatively small number of studies, most of which have focused on lakes (e.g. Greene et al. 1997, Cameron 2000, Dieffenbacher-Krall 1998). MNAP provided rare plant data from the Biological Conservation Database, and D. Cameron (MNAP) supplied data from a series of "rapid bioassessment" surveys of Maine lakes.

***Biodiversity Information From Conservation Lands*:** A key issue for evaluating the representativeness of conservation lands (both existing and planned) is the extent to which these lands include regional species and communities. GAP analysis (Scott et al., 1993) was developed to address this need. GAP analyses generally rely on using habitat-based models to predict species distributions (e.g. Krohn et al. 1998). Habitat modeling was beyond the scope of MABP and thus we are currently restricted to evaluating aquatic biodiversity "coverage" in conservation lands from the perspectives of (i) how much data derives from conserved areas, and (ii) what proportions of regional species pools are documented as occurring in conserved areas. The extent to which lakes are included in conservation lands was addressed in Chapter 2 (Figure 2.8).

In terms of data resources, we focus on lakes. The majority of information from lakes in conservation lands is on fish – either formal fish-survey data or data derived from non-survey information sources (Table 5.5). Mirroring the statewide lakes population, more of the larger lakes in conservation lands have fisheries data than do the smaller lakes. In general, between

40% and 60% of lakes in the 10-100 acres size class have fish data. For all the other taxonomic groups, there is a paucity of data from lakes within conservation lands.

To evaluate species representation, we adopted used a different approach, focusing on watersheds associated with the GAP 1 and 2 conservation lands. Using GIS, we first identified all HUC-12 watersheds that intersected with conservation lands. We then extracted from the MABP database all records that fell within these watersheds. Finally, for each region, we compared the numbers of species in the conservation-land watersheds with the numbers recorded from the entire region. The total area covered by the target HUC-12 watersheds is larger than the area of conservation lands since many watersheds extend beyond the borders of these lands. However, from an ecological perspective, this over-estimate of land area would seem not to invalidate the approach – species present in a particular watershed are potentially present in the associated conservation land, regardless of whether or not there are actual records from the managed area. An exception to this would be for very small conservation parcels that possess little aquatic habitat. It seems unlikely that these will significantly influence the results of the overall analysis because data are here aggregated by region.

Across all taxonomic groups and regions, between about 30% and 100% of taxa within the regional species pools have been documented to occur in watersheds associated with conservation lands (Table 5.6). Highest representations occur in the amphibians, odonates and fish (both total species and Maine natives). Least representation occurs in the stoneflies, caddisflies, chironomids, crayfish and aquatic snails [9]. For the other groups, the degree of representation varies with region. Highest representation occurs as follows:

Aquatic plants: south and downeast;
Mayflies: central and west;
Aquatic beetles: south and northeast;
Mussels: central and downeast;
Turtles: south and central.

Again, note that these data refer to documented species occurrences, not modeled distributions. Further, the spatial extent of sampling effort in each region presumably influences the degree of representation – more collections from outside conservation lands will tend to increase the regional species pool and thus decrease the *proportional* richness of conservation lands.

[9] In part, this reduced representation derives from the fact that much of the data for these groups – crayfish excepted – come from MDEP's biomonitoring program, which focuses on impacted waterbodies, along with associated reference sites.

Table 5.3: Percentage of MIDAS-numbered lakes that have been surveyed for fish assemblages, by lake size class and region.
See Figure 2.3 for regions. Source data: MDIFW; data are current through 2004.

Region	Region # (Fig. 2.3)	SizeClass(acres)	Total Lakes	% Lakes With Data
SOUTH	1	<1	70	2.9
	1	1-9.9	290	12.4
	1	10-99	204	58.8
	1	100-999	75	93.3
	1	1000+	10	100.0
CENTRAL	2	<1	159	0.0
	2	1-9.9	432	6.0
	2	10-99	322	57.1
	2	100-999	167	91.6
	2	1000+	42	100.0
DOWNEAST	3	<1	110	0.9
	3	1-9.9	307	10.7
	3	10-99	264	53.0
	3	100-999	134	88.8
	3	1000+	38	100.0
NORTHEAST	4	<1	85	1.2
	4	1-9.9	258	8.9
	4	10-99	182	49.5
	4	100-999	53	81.1
	4	1000+	15	100.0
WEST	5	<1	330	0.3
	5	1-9.9	965	12.3
	5	10-99	601	67.6
	5	100-999	208	88.9
	5	1000+	54	90.7
NORTHWEST	6	<1	67	0.0
	6	1-9.9	196	8.7
	6	10-99	171	59.1
	6	100-999	56	98.2
	6	1000+	13	100.0

Table 5.4: Period of record for stream fish and stream macroinvertebrate data.
Table shows the number of sampling sites by period of record. Stream fish data are from MDIFW and ASC. Invertebrate data are from multiple sources (primarily MDEP).

# YEARS IN RECORD	FISH	MACROINVERTEBRATES
1	1,996	382
2 – 5	268	174
6 – 10	48	17
11 – 15	17	0
15 – 20	5	1
20 +	3	0

Table 5.5: Aquatic biodiversity data resources for lakes within conservation lands.

Table entries are numbers of lakes, by region and size class, that have data on fauna and flora. A lake was considered to be in a conservation land if more than 80% of its perimeter buffer is located in land with Gap status of 1 or 2. Total lakes refers to the total number of waterbodies of a size-region class that are within conservation land.
Regions: 1 = South; 2 = Central; 3 = Downeast; 4 = Northeast; 5 = West; 6 = Northwest (see Figure 2.3 for map).
Lake size classes: 1 = < 1 ac; 2 = 1-9 ac; 3 = 10-99 ac; 4 = 100-999 ac; 5 = ≥ 1000 ac. Data source: MABP database. Conservation land data provided by TNC, Boston.

			# Lakes With Data for Taxonomic Group:							
Region	Size Class	Total Lakes	Fish: Survey Data	Fish: Non-Survey Data	Benthic Inverts.	Macro-phytes	Zoo-plankton	Phyto-plankton	Mussels	Crayfish
1	1	7	0	0	0	0	0	0	0	0
1	2	15	1	0	0	0	0	0	0	0
1	3	13	7	3	0	0	0	1	0	0
1	4	5	2	3	0	1	0	0	0	0
1	5	0	0	0	0	0	0	0	0	0
2	1	1	0	0	0	0	0	0	0	0
2	2	14	0	0	0	0	0	0	0	0
2	3	10	5	2	1	1	1	2	1	0
2	4	2	1	0	0	0	0	0	0	0
2	5	1	1	1	1	0	0	1	1	0
3	1	12	0	1	0	1	0	0	0	0
3	2	38	5	6	0	5	3	0	0	0
3	3	34	21	13	4	9	6	4	1	0
3	4	8	6	2	4	2	2	2	2	0
3	5	0	0	0	0	0	0	0	0	0
4	1	7	0	0	0	0	0	0	0	0
4	2	30	3	1	0	0	1	0	0	0
4	3	20	7	0	1	0	0	0	1	0
4	4	2	0	0	1	0	0	0	1	0
4	5	0	0	0	0	0	0	0	0	0
5	1	96	0	0	0	0	0	0	0	0
5	2	274	53	3	1	0	1	0	1	0
5	3	182	127	19	4	1	6	4	4	0
5	4	54	48	21	12	0	4	2	11	0
5	5	10	10	8	8	1	4	4	5	1
6	1	23	0	0	0	0	0	0	0	0
6	2	54	7	1	1	0	1	0	1	0
6	3	49	32	10	7	0	2	0	3	3
6	4	10	10	9	3	0	0	0	3	1
6	5	4	4	4	4	0	1	1	4	0

Table 5.6: Species documented from conservation lands and associated watersheds, by region, for selected taxonomic groups. Data are numbers of species and *(% of regional species pools)*.

Numbers in red and in parentheses are species numbers expressed as % of total number of species, of the taxonomic group, documented from the region. Conservation lands were defined as managed areas in GAP status 1 or 2 (see text for description of GAP categories). Using GIS, all HUC-12 watersheds intersecting conservation lands were selected. Biodiversity data from these watersheds were extracted from the MABP database. Note that species totals reflect documented records, not inferred species richness. Regions are shown in Figure 2.3. See text for additional information.

GROUP	REGION					
	South	***Central***	***Downeast***	***NE***	***West***	***NW***
Aquatic Plants	81 *(81)*	84 *(76)*	90 *(97)*	33 *(48)*	53 *(66)*	21 *(45)*
Aquatic Plants: Tracked	12 *(80)*	13 *(68)*	11 *(79)*	2 *(33)*	6 *(67)*	3 *(43)*
Mayflies	28 *(58)*	85 *(81)*	58 *(52)*	28 *(43)*	114 *(86)*	38 *(62)*
Stoneflies	9 *(56)*	9 *(53)*	9 *(18)*	11 *(69)*	10 *(63)*	7 *(54)*
Odonates	138 *(98)*	133 *(90)*	124 *(89)*	81 *(74)*	103 *(79)*	83 *(91)*
Odonates: Tracked	19 *(100)*	8 *(57)*	9 *(75)*	2 *(33)*	4 *(40)*	5 *(100)*
Caddisflies	34 *(74)*	36 *(35)*	87 *(40)*	32 *(35)*	100 *(68)*	64 *(46)*
Chironomids	87 *(77)*	73 *(70)*	44 *(63)*	61 *(74)*	38 *(53)*	15 *(58)*
Blackflies	6 *(86)*	19 *(61)*	5 *(42)*	4 *(100)*	33 *(92)*	4 *(80)*
Water beetles	14 *(88)*	11 *(27)*	6 *(24)*	14 *(82)*	4 *(31)*	4 *(31)*
Crayfish	3 *(43)*	7 *(70)*	2 *(25)*	2 *(22)*	5 *(45)*	1 *(25)*
Aquatic snails	4 *(31)*	13 *(46)*	2 *(50)*	11 *(69)*	2 *(50)*	0 *(0)*
Mussels	4 *(57)*	10 *(100)*	7 *(88)*	7 *(78)*	8 *(73)*	3 *(75)*
Fish	39 *(87)*	43 *(74)*	39 *(83)*	31 *(79)*	44 *(98)*	32 *(97)*
Fish: ME natives	32 *(86)*	35 *(78)*	34 *(83)*	30 *(86)*	36 *(95)*	31 *(94)*
Amphibians	16 *(100)*	17 *(94)*	15 *(94)*	14 *(93)*	16 *(100)*	12 *(86)*
Turtles	7 *(100)*	7 *(100)*	3 *(60)*	2 *(67)*	4 *(67)*	1 *(50)*
TOTAL ACRES (x 10^3) *	814.9	1255,8	885.7	649.7	2592.2	1694.2

* Total area of the HUC-12 watersheds that intersect with the conservation lands. These watersheds were used as the spatial framework for extracting the biological data summaries from the MABP database.

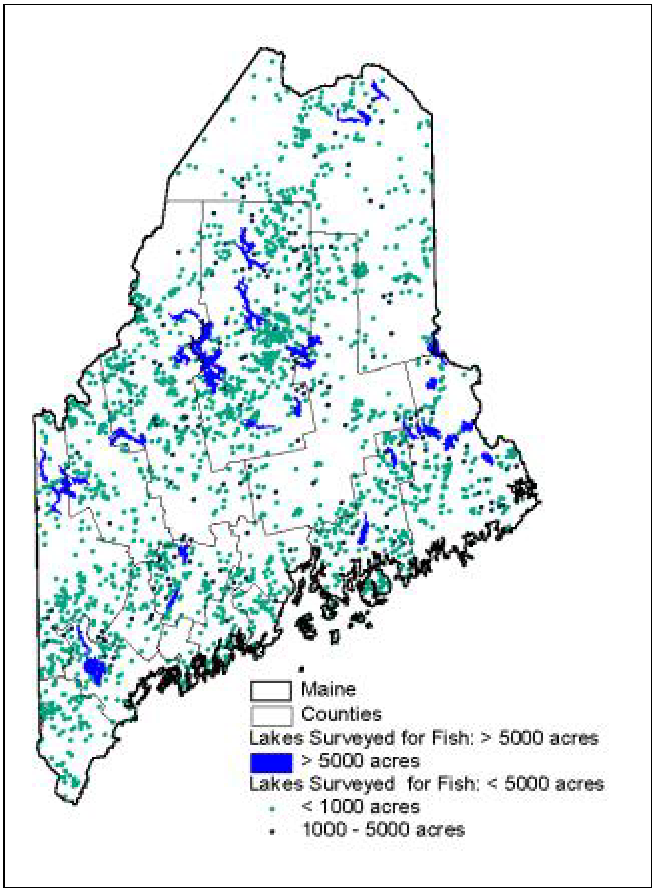


Figure 5.2: Map showing lakes that have been surveyed for fish assemblages (as of 2004). Large lakes (> 5,000 acres) are shown as polygons, smaller lakes as points. Note that a few river impoundments do not show on this map, even though they have been surveyed, since they are not included in the lakes & ponds GIS coverage. Data source: MDIFW.

Region	Size Class (acres)	% Lakes Surveyed Once
1	<1	50
1	1-9.9	69
1	10-99	54
1	100-999	19
1	1000+	10
2	<1	--
2	1-9.9	73
2	10-99	59
2	100-999	39
2	1000+	5
3	<1	0
3	1-9.9	58
3	10-99	57
3	100-999	50
3	1000+	34
4	<1	0
4	1-9.9	61
4	10-99	57
4	100-999	42
4	1000+	13
5	<1	100
5	1-9.9	45
5	10-99	61
5	100-999	57
5	1000+	51
6	<1	--
6	1-9.9	53
6	10-99	33
6	100-999	27
6	1000+	8

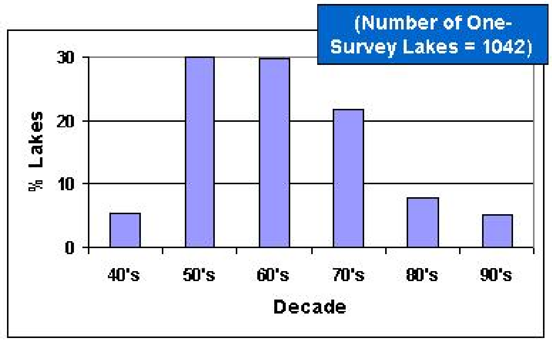


Figure 5.3: Percentage of lakes that have been surveyed only once for fish assemblages by MDIFW, by size class and region. Inset graph shows the decade of survey for one-survey lakes.
Note that "survey" refers to a full MDIFW lake survey; additional species records may occur in years subsequent to the last complete survey. See map in Figure 2.3 for regions. Data source: MDIFW, 2004.

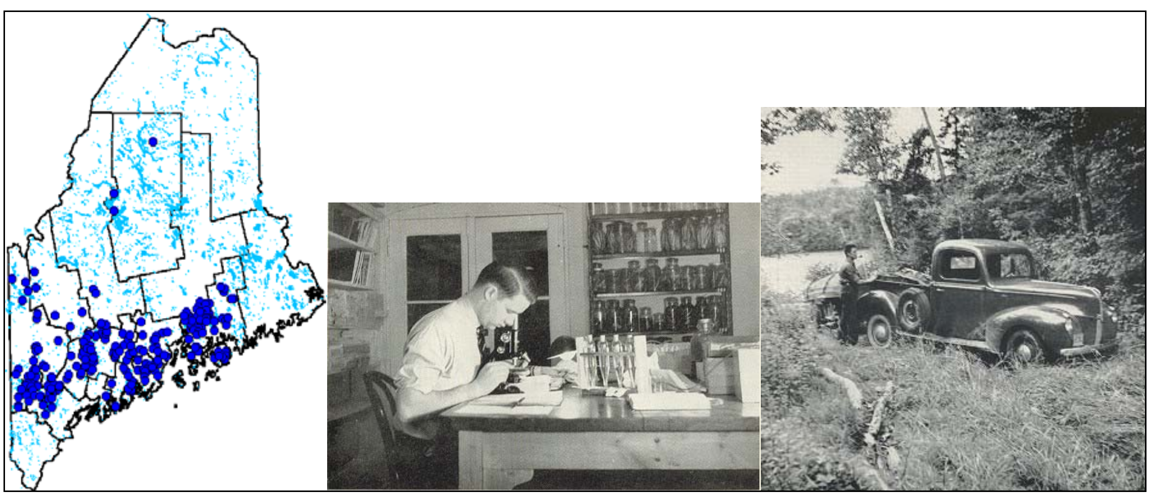


Figure 5.4: Lakes sampled by G. Cooper (photo) and colleagues between 1938 and 1944. These lake surveys collected data on fisheries, plankton, benthos and water quality. Lakes sampled by Cooper are shown as dark blue dots; lake statewide are in light blue.

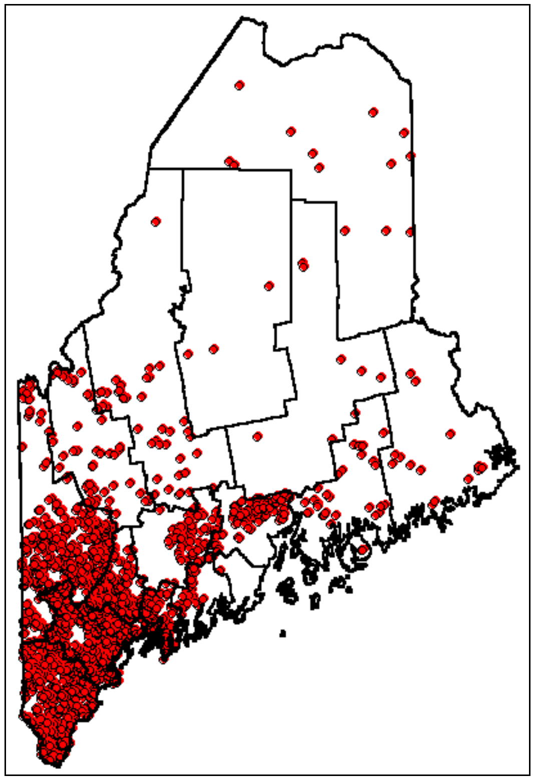


Figure 5.5: MDIFW stream electrofishing sites with computerized data (as of March, 2004). Data source: MABP database. (Note: data from additional sites in central and northern Maine were computerized later in 2004.)

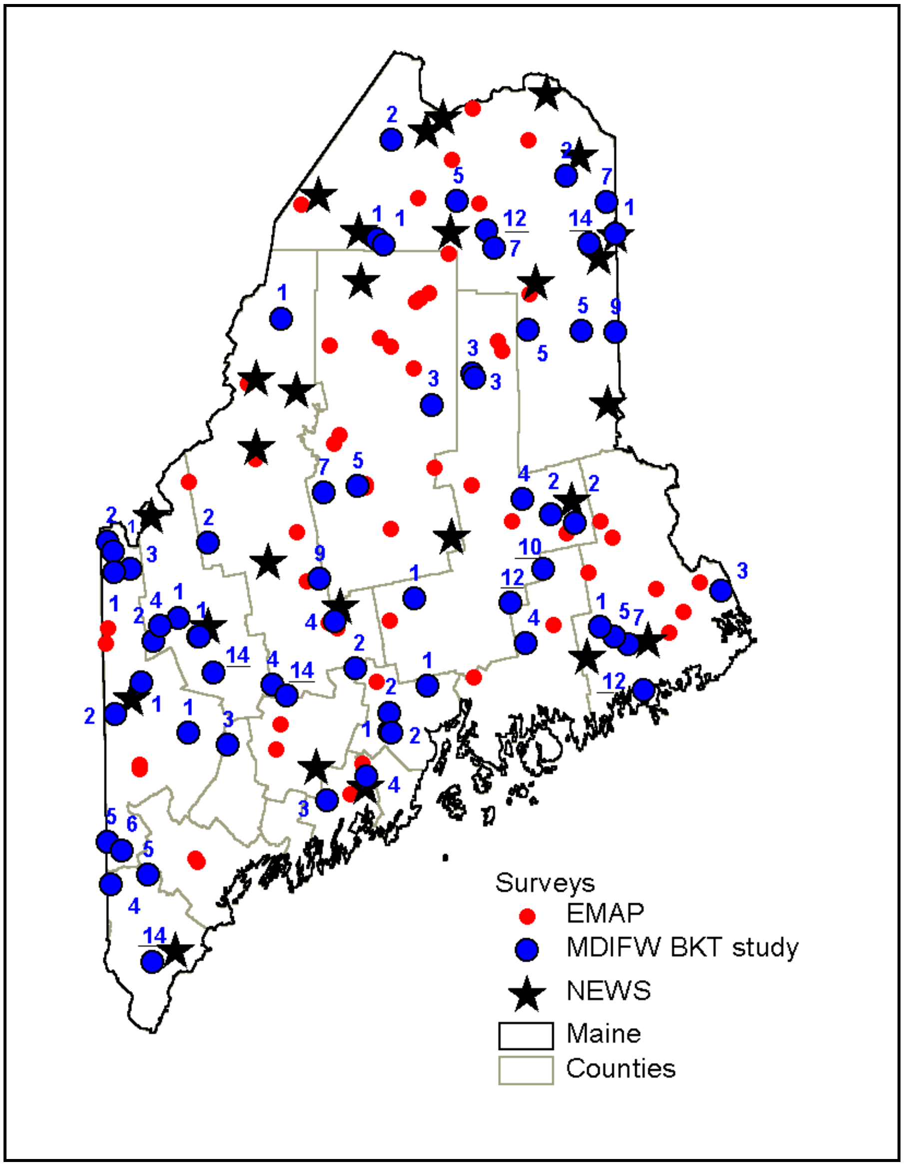


Figure 5.6: Sites sampled for fish assemblages using standardized protocols and effort levels.
Note that standardization refers to within each study, but not across the three studies. EMAP sampled lakes (1 or 2 years' data at each lake) during 1990s. The MDIFW Brook trout survey and the New England Wadeable Streams (NEWS) project sample streams – both surveys were on-going as of 2004. The number of years in the MDIFW survey are shown for each site (through 2004) – sites with >9 years' of data are underlined. Data sources: USEPA (EMAP and NEWS) and MDIFW as compiled in MABP database (2004).

(A)

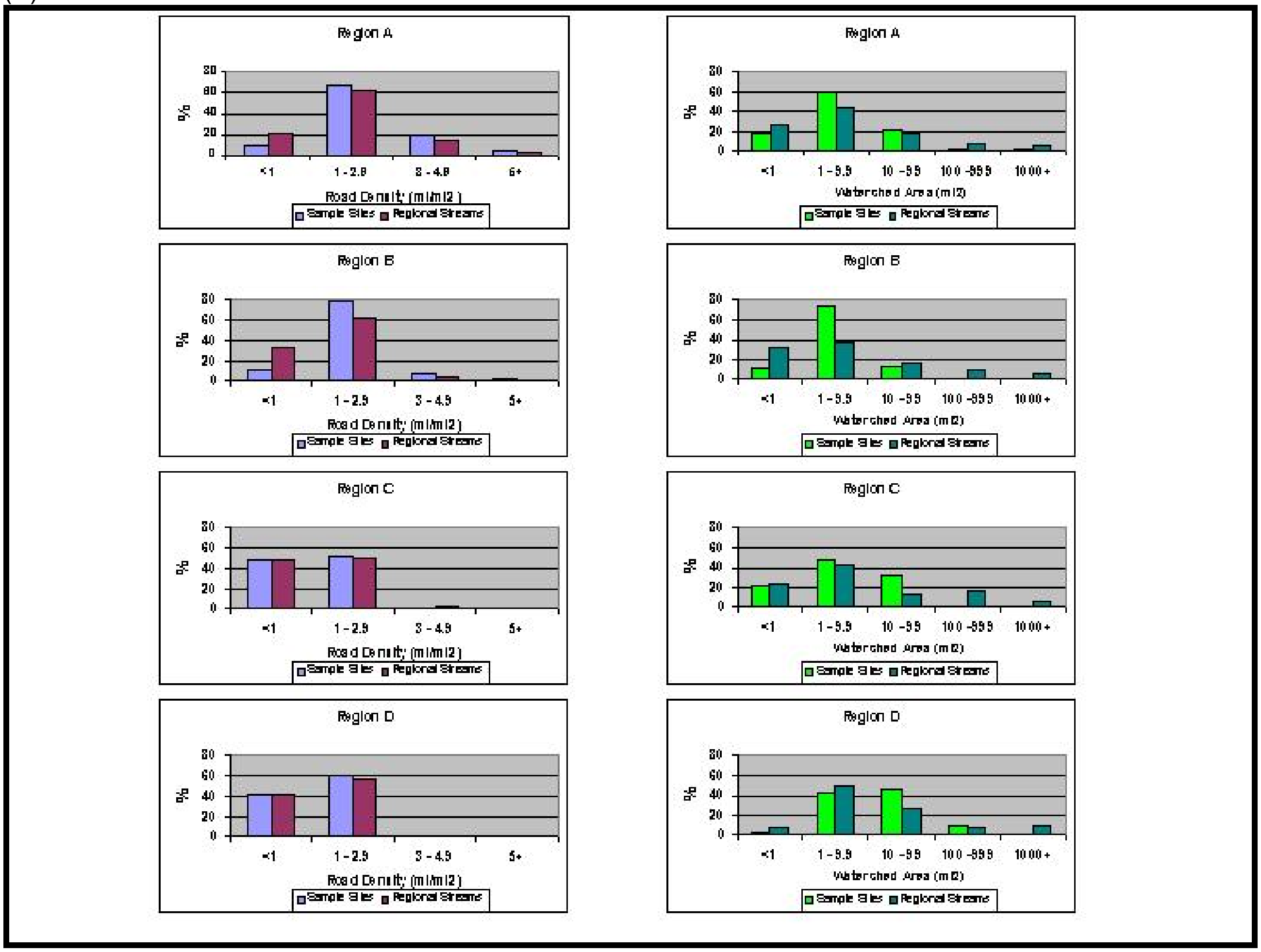


Figure 5.7: Four landscape-level characteristics of stream sites sampled for fish by MDIFW.
Sites are grouped by MDIFW region: A = south, B = central, C = Downeast, D = western Maine. Graphs show landscape attributes for (a) watersheds upstream of sampling sites and (b) all stream segments in the region. Data (based on 1992 satellite imagery) were derived by TNC using an ArcInfo script to calculate watershed-level landscape characteristics for stream segments mapped at the 1:100K scale. Data sources: MDIFW and TNC, Boston.

(B)

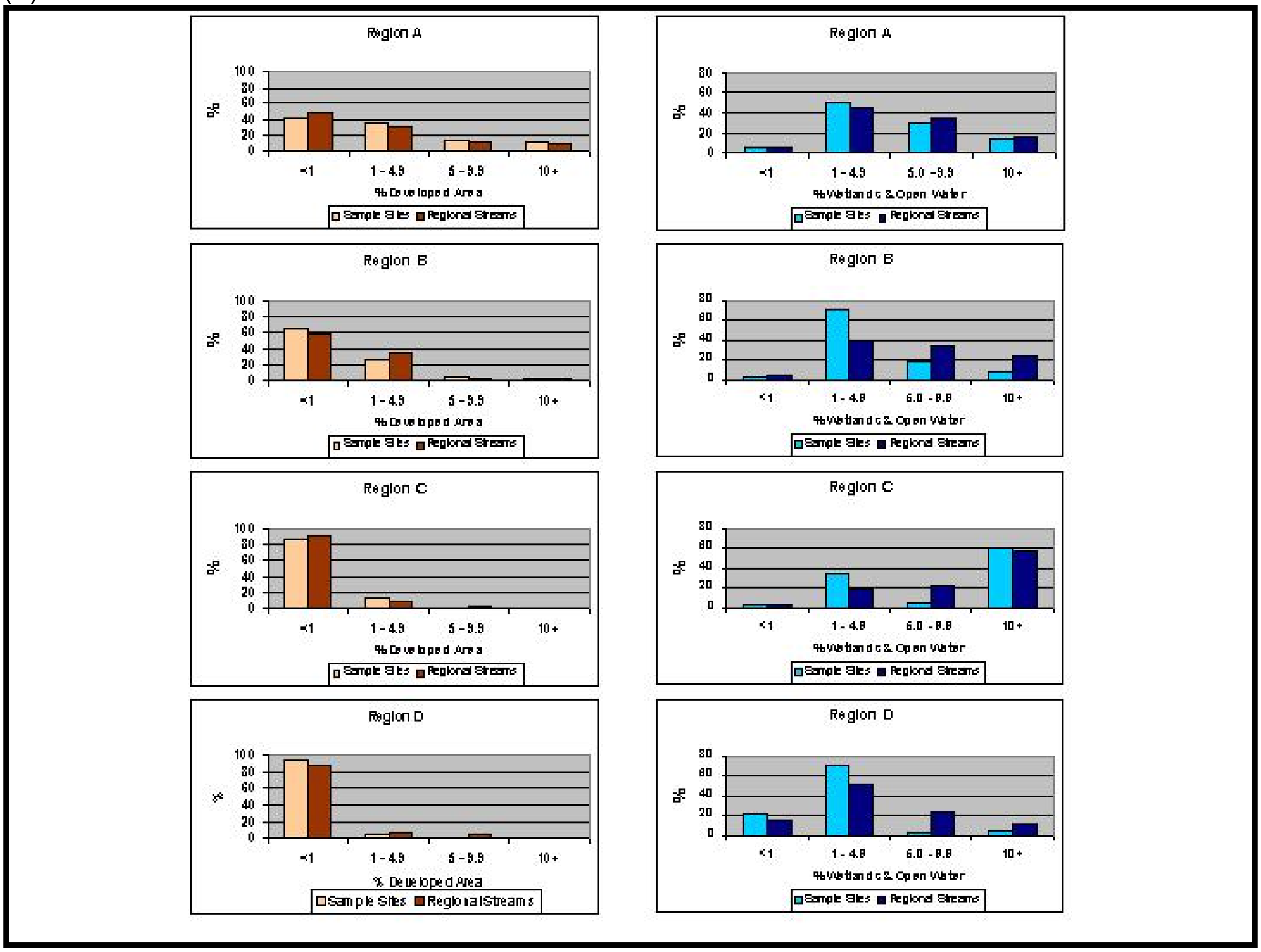


Figure 5.7 (cont.)

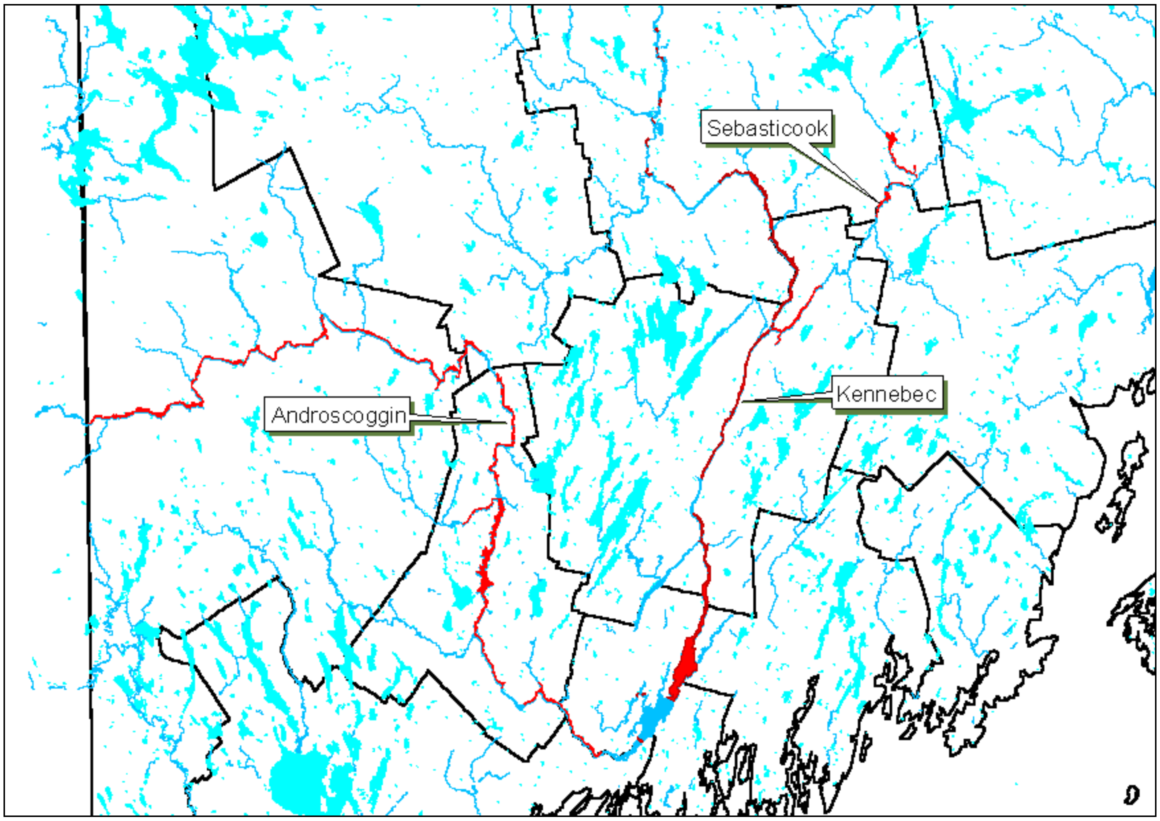
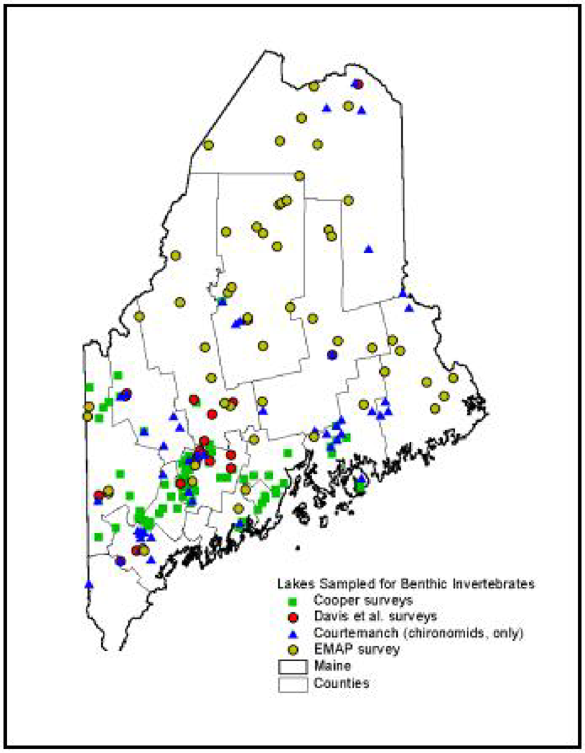


Figure 5.8: Large river surveys of fish assemblages.
Red-shading indicates reaches sampled by electrofishing during 2002 and 2003 by Yoder et al. Not shown is sampling of the Penobscot River during 2004. Data source: D. Halliwell (MDEP) and M. Gallagher (MDIFW), obtained originally from B. Kulik, (Kleinschmidt Associates, Pittsfield, ME).



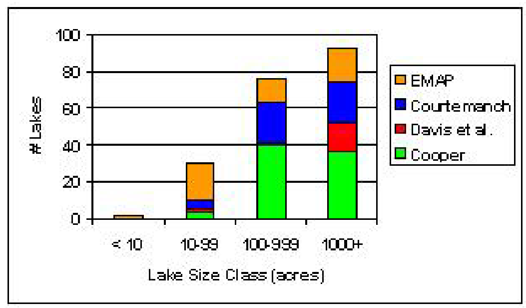
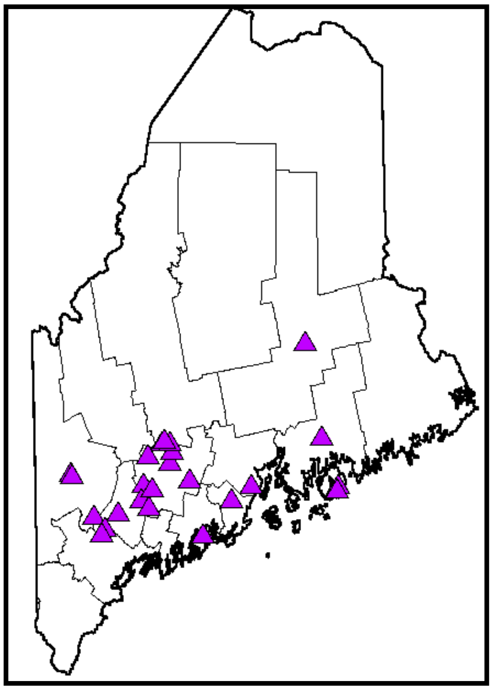


Figure 5.9: Lakes that have been sampled for benthic macroinvertebrates. Graph at right shows the frequency distribution of lake size classes, by study.
Some lakes were sampled in more than one study. Data sources: EMAP data provided by USEPA; Courtemanch (1982); Davis et al. (1978); Cooper (1939, 1940, 1941, 1942; Cooper and Fuller 1945, Fuller and Cooper 1946).



MIDAS	LAKE NAME	TOWN
0098	NARROWS P (UPPER)	WINTHROP
0447	LONG P	MOUNT DESERT
2146	COLD STREAM P	ENFIELD
3688	RANGE P (UPPER)	POLAND
3748	AUBURN L	AUBURN
3760	RANGE P (LOWER)	POLAND
3770	HOGAN P	OXFORD
3772	WHITNEY P	OXFORD
3814	COCHNEWAGON P	MONMOUTH
3828	BERRY P	WINTHROP
4350	GRAHAM L	MARIAVILLE
4608	JORDAN P	MOUNT DESERT
4798	ALFORD L	HOPE
4848	PITCHER P	NORTHPORT
5182	FLYING P	VIENNA
5238	SAND P (TACOMA LKS)	LITCHFIELD
5280	MESSALONSKEE L	BELGRADE
5344	NORTH & LITTLE PONDS	ROME
5348	MCGRATH P	OAKLAND
5349	EAST P	SMITHFIELD
5416	THREEMILE P	CHINA
5710	BISCAY P	DAMARISCOTTA
9683	SOUTH & ROUND PONDS	GREENWOOD

Figure 5.10: Lakes that have assemblage-level data for three major taxonomic groups: fish, macroinvertebrates and vascular plants.
Data source: MABP database, from multiple sources.

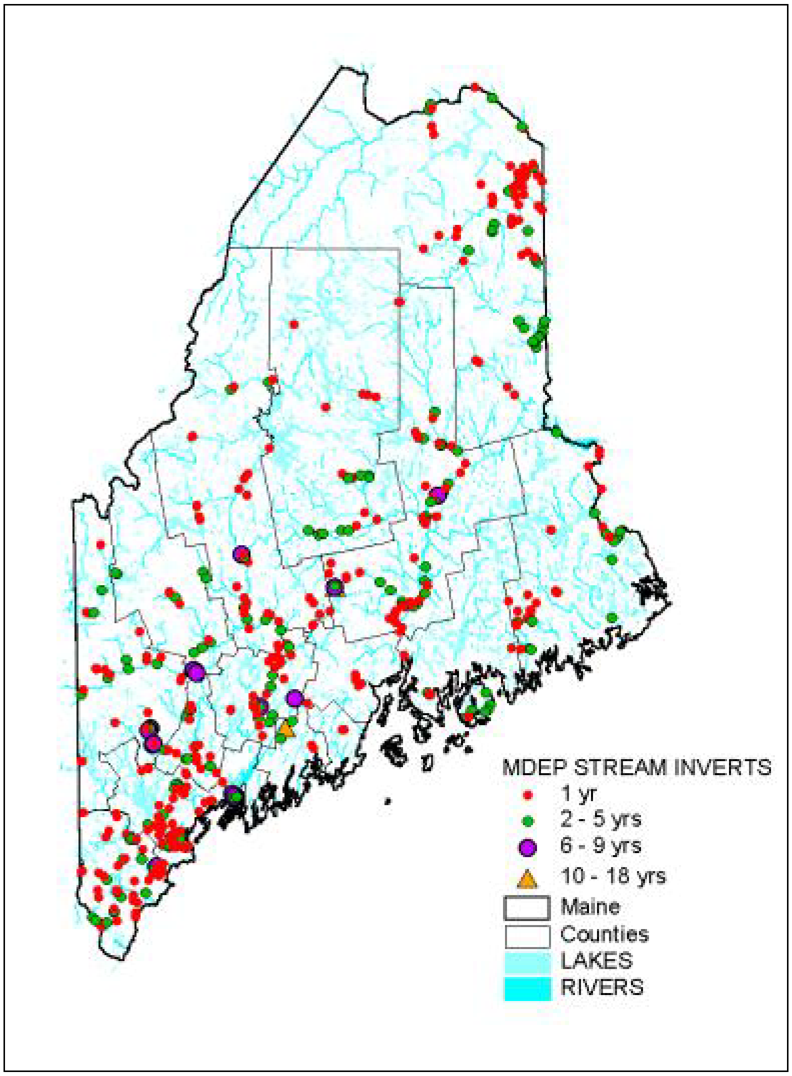


Figure 5.11: Locations of MDEP stream macroinvertebrate assemblage sampling sites. Sites are coded by number of years in the data record. Data source: MDEP, 2004 (data current through 2003).

5.3 Sampling Effort Issues

How much we know about ecological communities depends in part on how well they are sampled – in terms of both the quantity and quality of the sampling effort. Much has been written about the relationship between sampling effort and the completeness with which plant and animal assemblages are documented. Numerous studies have evaluated various solutions to the problem of inadequate sampling (e.g. Fisher et al.1943, Palmer 1990, Soberon and Llorente 1993, Rosenzweig et al. 2003). Other studies have focused on the issue of conservation planning in the context of sparse biological data (e.g. Gaston and Rodrigues 2003). We make no attempt here to summarize this research. Instead we use specific examples from the database on Maine's freshwater biodiversity to illustrate the sampling effort issue and how important it is to keep it in mind when examining species data[10].

- **Repeat sampling of lake fish**: Data from EMAP illustrate the fact that a single survey of a lake – even with multiple gear types, as was the case with the EMAP study – generally does not record all fish species present (Table 5.7). Sixteen lakes were re-sampled, either within the same year and/or in different years. In general, the number of species recorded from a single visit was between 80% and 90% of the cumulative number of species recorded during multiple visits. It is possible that some of this variation in species number reflects transient species in the lake, e.g. stream fish moving between lake (sampled) and tributaries (not sampled). However, it seems much more likely that it is simply a sampling artifact.

- **Annual variation in documented stream fish richness:** An analogous data summary to that developed from the EMAP data derives from the MDIFW brook trout monitoring program. Nine sites have been sampled for over 10 years, with standardized sampling protocols. Since 1994, all fish species present in the samples have been recorded. The number of species captured in any one year is always less that the cumulative number of species recorded over the period from 1994 through 2002 (Table 5.8). For some streams, the single-year totals are less than half of the cumulative totals. As with the lake data, it is likely that this results from a combination of sampling effort and fish mobility.

- **Stream fish diversity by watershed:** One way to summarize biodiversity data is as the number and identity of species present in a watershed (rather than at an individual site). Multiple sites within a watershed are sampled and a cumulative species list is developed. Using stream electrofishing data from southern and central Maine, we examined the relationship between cumulative species richness in watersheds (HUC-10) and the number of sites that were sampled in each watershed (Figure 5.12). There is evidence of the classic asymptotic species-area curve. However, it appears that, across this region as a whole, about 20 sites are sufficient to obtain a reasonably complete count of the total number of (stream) fish species present in a watershed.

- **Mayfly and odonate diversity by watershed:** A similar approach was taken with mayfly and odonate data (Figure 5.13). As the measure of sampling effort at the HUC-10 watershed level, we used either the number of distinct sampling sites (mayflies) or the number of site-date combinations (odonates – some sites were re-visited). Although survey intensity presumably varied among sites and dates, these are the only available indicators of sampling effort. For odonates (Figure 5.13), the relationship between effort and number of species is striking.

[10] For an additional illustration of the species-effort relationship for aquatic plants, see Figure 6.2.5 in Chapter 6.2.

The relationship is less clear for mayflies, but more species in a watershed were identified as more sites were sampled.

- **Mussel diversity by watershed:** Analogous data from the MDIFW mussel survey are shown in Figure 5.14. Since there are only 10 freshwater mussel species in Maine, the combined lake and stream data reach an asymptote after about 20 sites.

- **Species, effort and watershed area:** One possible explanation for the patterns illustrated above is that larger watersheds contain more species and were sampled via a larger number of sites. However, data summarized in Figure 5.15 suggest that this was not the case. For mussels, more sites clearly were sampled in larger watersheds (Figure 5.15). However, the number of species per watershed was not strongly associated with watershed area, in itself. For both mayflies and odonates there was no relationship between number of species and watershed area and, for mayflies, only a weak relationship between number of sites and watershed area (Figure 5.15).

- **Tracked species and sampling effort:** MDIFW tracks a subset of 23 rare odonate species. In contrast to the full species complement (Figure 5.13), there is no obvious relationship between species and effort for tracked species (Figure 5.16), except possibly at very low levels of sampling effort. There is also only a very weak association between the number of tracked species recorded from a watershed and the total number of species (Figure 5.16, lower panel).

- **Rare plant species richness by watershed:** The association between the number of rare plant species documented in a watershed and the number sites visited was discussed in Chapter 3 (see Figure 3.9).

- **Sampling effort by watershed:** Figure 5.17 summarizes the level of sampling effort, by watershed, for four invertebrate groups. In an effort to portray different levels of sampling "adequacy", the effort categories were defined based on the species-effort curves discussed earlier (Figures 5.12 – 5.16). The dark-shaded watersheds appear to have been sampled sufficiently intensively to assure relatively complete species lists. These summaries can be used as a first approach to identify data gaps (see below).

5.4 Data Gaps

There are several approaches to defining, and subsequently identifying, data gaps. From a biodiversity perspective, the most basic approach addresses whether or not there exist reasonably complete species lists for any of several spatial frameworks (state, watershed, conservation land parcel, waterbody type, etc.). Species lists may address the entire species pool, or may focus on a subset of species such as rare and high-value species. From a conservation planning perspective, data gaps might be addressed through keystone, focus or umbrella species, rather than the entire species pool. At a finer resolution, it is possible to identify data gaps in a regional context and by ecosystem type, for example lake size or trophic class. A common form of data gaps relates to time series records. Allied to this are gaps in baseline data, or more specifically, baseline data that are collected in a sufficiently consistent manner to permit rigorous spatial and temporal comparisons.

An important class of data gaps relates to multi-taxon information from one or more habitats or collection of habitats. Ecosystem level data such as the series of plant, invertebrate and fish assemblages within the same lake or stream system are particularly important for developing biologically based systems of aquatic classification. Similarly, the composition of aquatic communities along different environmental and stressor gradients can provide an important tool for quantifying environmental impacts and assessing compliance. Gaps in multi-taxon data from reference and impacted sites influence the extent to which we are able to interpret and utilize biological information.

Table 5.9 provides a summary of key data gaps identified from the analyses presented in this report. This information is intended only as a general characterization of what is needed to improve our overall knowledge of the status and trends in Maine's freshwater biodiversity. Over-arching needs are

(i) More data on the distribution, relative abundance and habitat of rare taxa.
(ii) Improved documentation of biodiversity in relatively uncommon aquatic systems, such as high elevation lake, fishless ponds, acidic seepage lakes, headwater streams.
(iii) More species-level data on the least-well surveyed taxonomic groups (as noted in Table 5.9).
(iv) Rigorously collected multi-taxon assemblage-level data at a series of reference stream and lake sites.
(v) Allied to (iv), establishment of network of long-term monitoring sites, based where possible on existing data-rich sites. In the case of stream fish, some of the sites in the brook trout monitoing study represent an excellent base from which to develop this network.
(vi) Multi-taxon data to improve our understanding of how key stressors influence aquatic biota: for example, mercury and other toxins, habitat deterioration, timing and extent of water level fluctuations, species introductions.
(vii) Improved understanding of among-population genetic variation, especially in rare species.

5.5 Access to MABP information

Starting in late 2005, most of the biodiversity data that have been compiled in the MABP database will be made available on-line at the PEARL website (**www.pearl.maine.edu**). Data can be searched by topic and spatially, by waterbody, town or watershed. In addition, a number of the data syntheses and graphics presented in this report will be available on PEARL, beginning in late 2005.

Table 5.7: Number of fish species documented at 16 Maine lakes with repeated sampling using standardized effort.
Data are from EPA Environmental Monitoring and Assessment Program (EMAP). Lakes were sampled with multiple gears; gear mix was generally consistent across all visits to a lake.

LAKE	**MIDAS**	**# YEARS** (a)	**# VISITS** (b)	**Total Species** (c)	**# Species / Visit: Mean** (d)	**# Species / Visit: SD** (d)
Alder Brook P.	0055	2	4	11	9.0	1.4
Bog P.	2586	2	4	5	3.8	1.0
Chaffin P.	3718	2	4	7	4.8	1.0
Cranberry P.	7509	2	4	7	5.3	0.5
Long P.	9861	1	2	15	12.5	3.5
Machias L. (Fourth)	1148	2	3	12	11.0	1.0
Mattanawcook L.	2226	2	4	10	9.3	0.5
Mud P.	1600	2	3	5	4.0	1.7
Otter P.	3074	2	4	5	4.3	0.5
Papoose P. (Little)	3268	1	2	7	7.0	0.0
Reed P. (Big)	2842	1	2	9	8.0	0.0
Reed P. (Little)	2838	1	2	5	5.0	0.0
Spring (Muddy) P.	4904	2	4	7	5.3	0.5
Sysladobsis L. (Upper)	4688	2	4	16	12.5	1.9
Thurston P.	4321	2	4	13	12.3	0.5
Twin Island P.	5084	2	3	4	4.0	0.0

(a) Total number of years lake was sampled. (b) Total number of sampling visits, across all years. (c) Total number of species recorded from lake. (d) Mean and standard deviation of number of species per visit.

Table 5.8: Temporal variation in stream fish assemblages – 1994-2002.
Data are number of species recorded annually for each stream, and cumulative number of species over the entire sampling period. All sites were sampled with standardized methodology.

YEAR	Alder Brook (Perkins Twp.)	Branch Brook (Sanford)	Clark Brook (Presque Isle)	Greenlaw Stream (T12 R7)	Indian River (Addison)	Rome Trout Brook (Rome)
1994	3	4	7	8	4	3
1995	4	3	6	8	2	3
1996	4	3	7	8	2	3
1997	4	3	7	8	2	3
1998	4	4	6	9	2	6
1999	4	3	7	8	2	3
2000	3	3	6	8	3	4
2001	4	3	8	8	n.s.	7
2002	4	3	7	n.s.	3	4
TOTAL SPP (1994-2002)	5	6	11	11	6	9

(n.s. = not sampled)

[(1)] Data source: MDIFW brook trout monitoring project.

Table 5.9: Review of key data gaps.

Group	Under-surveyed regions / resources	Key Data Types Needed
Vascular plants	West, Northwest, Downeast. Ponds with existing records of invasive species (to quantify impacts from invasives).	Assemblage-level and coverage data. Periodic re-surveys of selected lakes with existing quantitative data.
Amphibians & Reptiles	West, Northwest, Downeast are least surveyed. Southern Maine, to better understand effects of urbanization / sprawl.	Quantitative data at baseline monitoring sites (e.g. MAMP)
.Fish	Lakes <50 acres – all regions. Remote ponds. High elevation lakes. Conservation land ponds and streams. Predicted fishless ponds. Streams, North and Downeast. Large rivers, esp St. John and St. Croix basins. Coastal rivers. Penobscot (w.r.t dam removal). .	Multi-gear surveys to obtain relatively complete information on full species assemblages and relative abundance. Base line data for selected reference systems. Extended time series data for selected stream sites, especially in western and NW Build on BKT monitoring and/or NEWS studies. Stream fish / habitat assocations – to quantify impacts of habitat disturbance.
Mussels	Generally good coverage statewide. West and Northwest have fewest collection sites. Penobscot basin upstream of Veazie. Lakes, in general.	Habitat data. Fish host data (via genetics studies).
Clams	All regions.	Species inventories: lakes and streams.
Snails	All regions, especially to south of Aroostook County.	Species inventories: lakes and streams. Habitat data. Focused searches for 3 globally imperiled / vulnerable species: *Amnicola decisa, Stagnicola mighelsi and S. oronensis.*
Crayfish	West, Downeast and northern Washington County. Central Maine to evaluate status of introduced red swamp crayfish.	Species inventories: lakes and streams.
Amphipods and isopods	All regions.	Species inventories.
Mayflies	West and Northwest. Washington County. High richness watersheds (Figure 6.7.1) to evaluate reasons for diversity. Western mountain streams to search for Roaring Brook mayfly.	Species inventories and target searches.
Odonates	Generally good coverage statewide. Least surveyed watersheds shown in Fig. 6.7.5. Wetlands.	Species inventories.
Stoneflies	All regions, especially West, South and Central.	Species inventories. Habitat data.
Caddisflies	Most regions, especially West, North and southern Washington County.	Species inventories.
Fishflies / alderflies	All regions.	Species-level data.
Hemiptera	All regions.	Species-level data.
Beetles	All regions.	Species-level data.
Chiromonids	West and Downeast.	Species-level data.
Blackflies	North and coastal lowlands.	Species-level data.
Macro-invert. assemblages.	West and Northwest.	Assemblage-level data, collected with standard protocols (e.g. rock baskets). Broader assemblage-level data, collected with multiple methods.

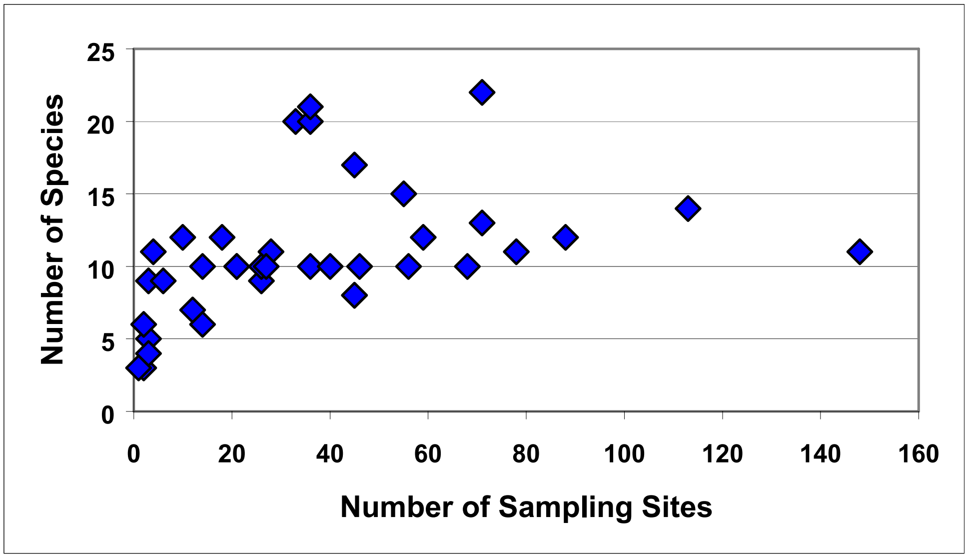


Figure 5.12: Relationship between number of sampling sites and cumulative stream fish species richness, by watershed (HUC-10).
Data are from southern and central Maine (Regions A and B). Data source: MDIFW.

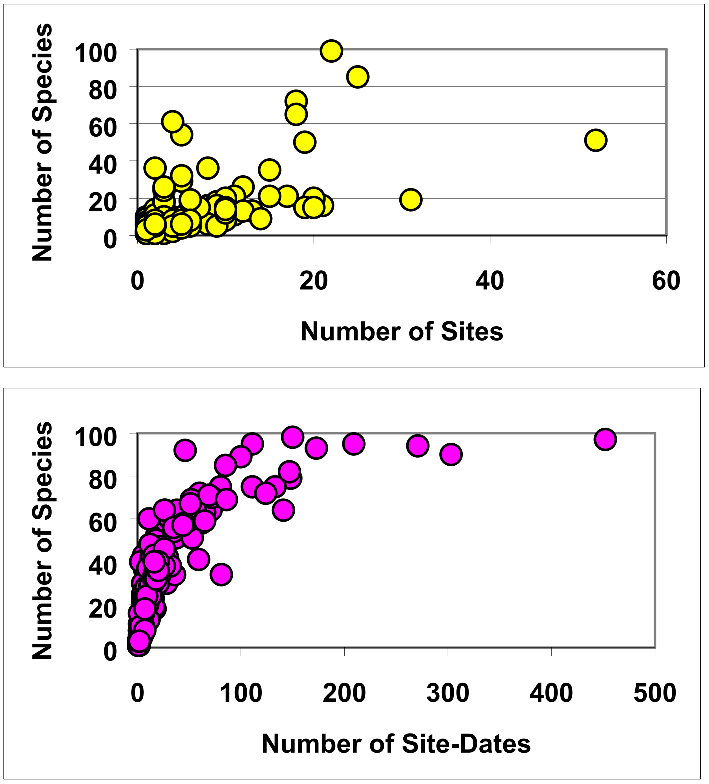
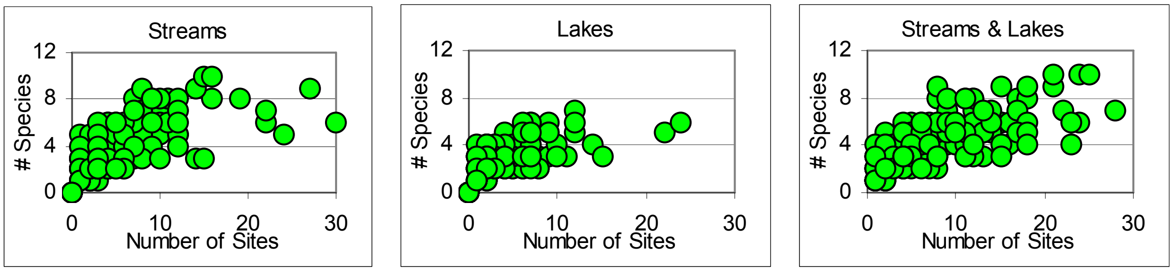


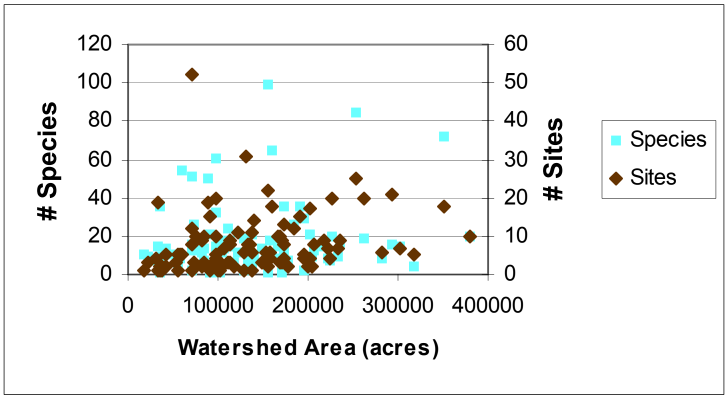
Figure 5.13: Relationship between sampling effort and documented species richness by watershed (HUC-10) for mayflies (upper panel) and odonates (lower panel).
For mayflies, sampling effort is quantified as number of sampling sites (unique coordinate pairs) in a watershed; for odonates, as number of site-date combinations. It is unlikely that either of these effort measures provides a full indication of the amount of search (sampling) effort employed for each sample. See text for more information. Data sources: Mayflies – Burian and Gibbs (1991), Mack (1988) and other sources compiled in MABP database. Odonates – Brunelle/DeMaynadier, unpublished data, MDIFW odonate survey and historical data compilation (current through 2003).



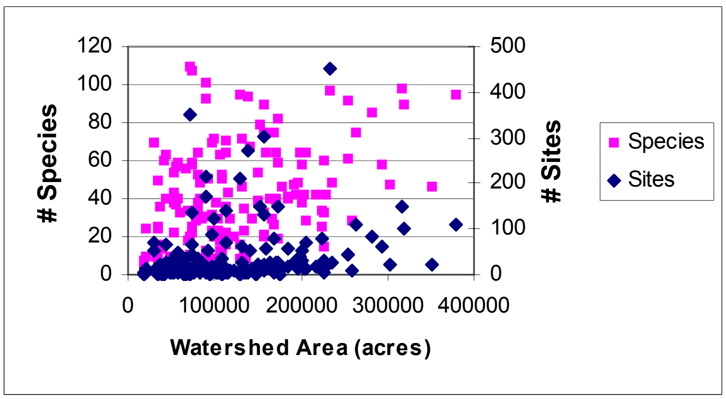
(A) (B) (C)

Figure 5.14: Relationship between sampling effort and species richness for mussels by watershed (HUC-10): (A) lakes, (B) streams, and (C) lakes and streams combined. Data source: MDIFW mussel survey and other sources as compiled in MABP database**.**

A. Mayflies



B. Odonates



C. Mussels

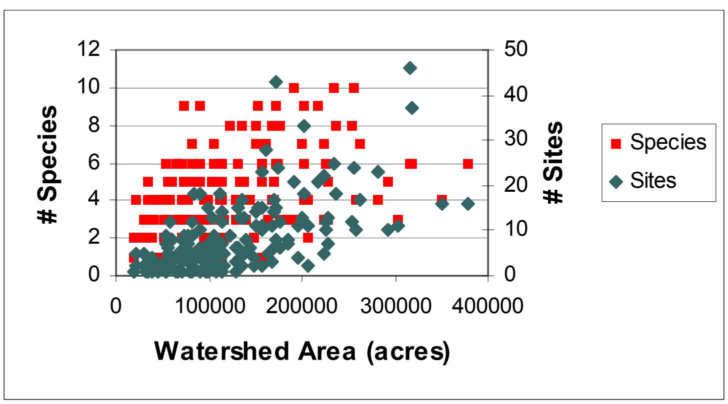
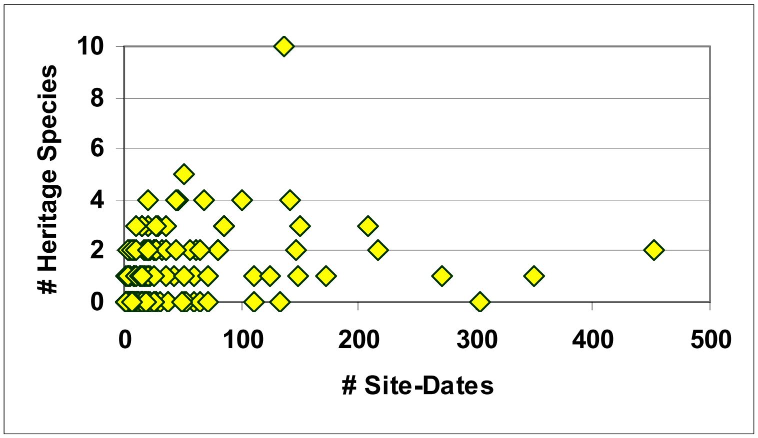


Figure 5.15: Relationship between watershed area (HUC-10) and (i) number of species, and (ii) number of collection sites for (A) mayflies, (B) odonates, and (C) mussels.

(A)



(B)

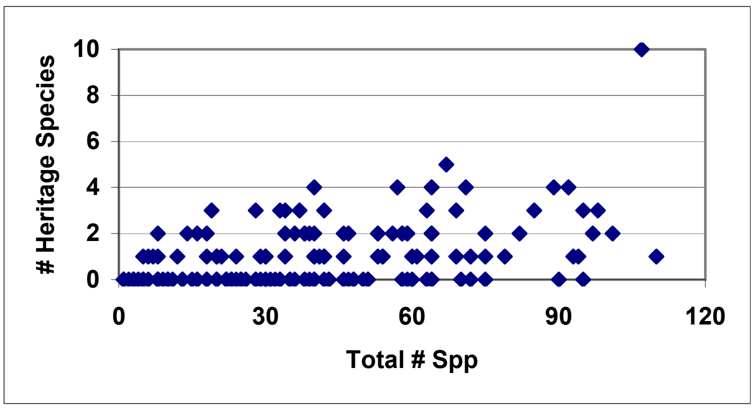


Figure 5.16: Two views of species number vs. sampling effort for Natural Heritage ("tracked") odonate species, by watershed (HUC-10). (A) Number of tracked species vs. number of site-data combinations. (B) Number of tracked species vs. total number of species.

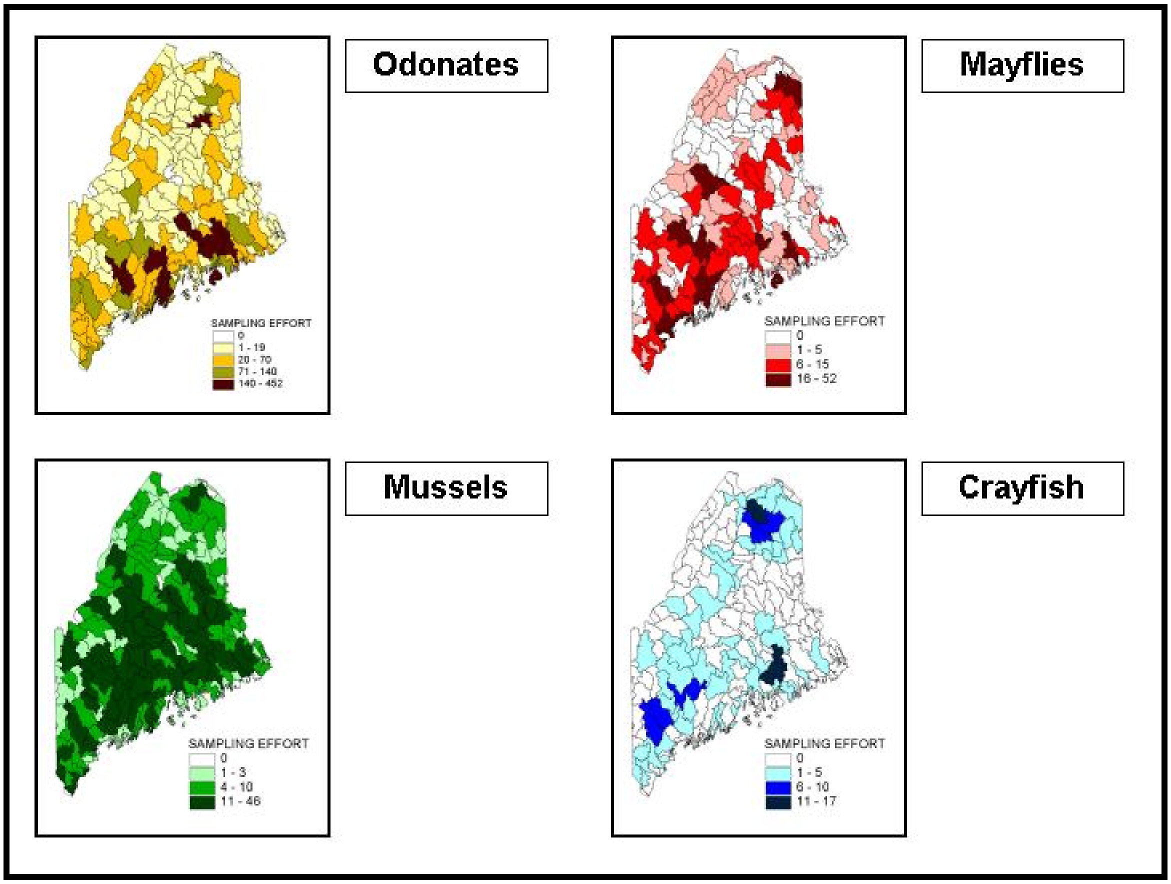


Figure 5.17: Sampling effort by watershed for four invertebrate groups.
Watersheds are HUC-10s. Sampling effort for odonates is # of site-date combinations; for the other groups, it is the number of sites. Effort-level groupings were derived from visual examination of species-effort plots (Figures 5.12 – 5.15) in an attempt to characterize apparent levels of sampling adequacy.
Data sources: Multiple, as compiled in MABP database.