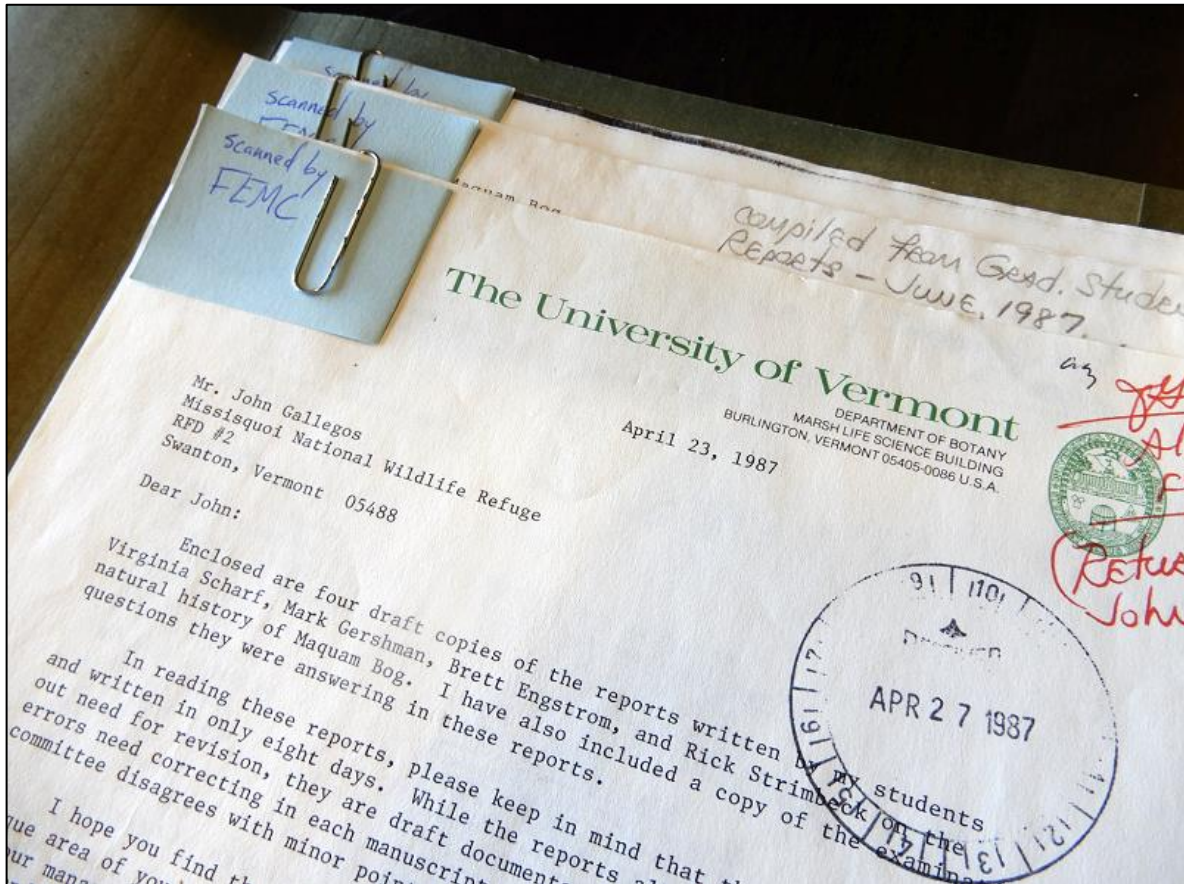

DATA RESCUE

FINDING AND PRESERVING DATA AT RISK

Final Report



Last updated: February 12, 2020

Data Rescue: Finding and Preserving Data at Risk

Published February 12, 2020

Forest Ecosystem Monitoring Cooperative

South Burlington, VT, USA

femc@uvm.edu

(802) 656-2975

Matthias Sirch^a and James Duncan^{a,b}.

^a Forest Ecosystem Monitoring Cooperative, 705 Spear Street, South Burlington, VT, USA

^b University of Vermont, Rubenstein School of Environment and Natural Resources, 81 Carrigan Drive, Burlington, VT, USA

DOI: <https://doi.org/10.18125/mslslh>

Cover image: Missisquoi National Wildlife Refuge 1987 correspondence of the Maquam Bog Missisquoi National Wildlife Refuge plant and animal inventories, scanned and archived for the Data Rescue project. Photo taken by Matthias Sirch at the Missisquoi National Wildlife Refuge visitor center in Swanton, Vermont, in 2019.

Preferred Citation

Sirch, MW and JA Duncan. 2020. Data Rescue: Finding and preserving data at risk. Forest Ecosystem Monitoring Cooperative: South Burlington, VT. <https://doi.org/10.18125/mslslh>

Report available online at <https://www.uvm.edu/femc/file/info/10377>

Data Rescue inventory available online at www.uvm.edu/femc/data_rescue

Acknowledgements

The Forest Ecosystem Monitoring Cooperative (FEMC) acknowledges FEMC's committees and the many partners who contributed their time and resources to this project. A number of FEMC interns and staff worked to compile this material, including Brenna Christensen, Gene Desideraggio, Jake Van Deursen, Alexandra Kosiba, Lucyanna Labadie, Yoshiya Ohno, Julia Pupko, Emma Tait, and John Truong.

We are appreciative of the long-term funding from the U.S. Department of Agriculture, Forest Service State & Private Forestry, Vermont Agency of Natural Resources and the University of Vermont that made this report possible.



Table of Contents

Table of Contents.....	3
Summary.....	4
Methods for Inventory and Rescue	4
Developing a Data Rescue Inventory	4
Priority Rescues.....	5
Inventory and Rescue Outcomes.....	6
Data Rescue Inventory.....	6
Priority Rescues.....	7
Summary, Gaps and Next Steps	9
Appendix – Organizations contacted for Data Rescue	11

Summary

A wealth of monitoring, inventory, and research data exist that are currently inaccessible to the broader community. These data range from field survey sheets stored in file cabinets to electronic files stored on hard drives of personnel who do not have the time or resources to archive this material. In each case, data are at risk of being lost to natural disasters, staff turnover or the loss of institutional memory. The Forest Ecosystem Monitoring Cooperative (FEMC) has worked to alleviate some of this risk by engaging with natural resources organizations in New England and New York to develop a Data Rescue inventory of at-risk datasets and documents.

The FEMC has additionally worked to rescue several high priority projects chosen for their contribution to long-term monitoring efforts relating to forested ecosystems. These rescue projects will be stored on the FEMC archive to preserve them in perpetuity and to make the datasets and documents available to interested researchers, land managers and other interested professionals.

Methods for Inventory and Rescue

Developing a Data Rescue Inventory

To include a broad range of institutions and forest-related disciplines in the Data Rescue effort, we contacted personnel focused on natural resources from federal and state agencies, academic institutions, and non-government organizations (NGO) (see Appendix, Table A1). Disciplines related to key components of the forested ecosystem beyond vegetation, including wildlife, water, air, and soil. Although the boundaries of the rescue were New England and New York, we did not contact organizations from Connecticut or Rhode Island as they were not yet participating in the FEMC in April, 2019. Maine, though not officially incorporated, was included as partners from that state have been participating in the FEMC in other ways.

For any respondents interested in learning more about the Data Rescue project, we scheduled meetings designed to provide a brief background of the FEMC as well as brainstorm to identify at-risk material. These meetings were conducted primarily over the phone, and to a lesser extent over email or in person. As an alternative to these meetings, participants were also encouraged to fill out an online form to provide information about the format, volume, ecosystem focus, spatial and temporal extents of the data, and what would be required to rescue each dataset or document.

We then developed a Data Rescue inventory of identified datasets and documents at risk of being lost across the region. This included all information gathered from the meetings and email responses. Metadata, including data descriptions, physical locations, and keywords, were developed for each item on the Data Rescue inventory as well as who to contact for more information. We then built a web interface at www.uvm.edu/femc/data_rescue to make this information available to the broader community for easy discovery and communication of the resources available that could be at risk.

Priority Rescues

The FEMC staff developed an index to choose several priority datasets to rescue between April and December, 2019. Priority rescues were selected using an index that weighted each project by parameters relating to temporal coverage, spatial coverage, and institutional value of the data (Table 1). Priority was given to projects with a high cumulative score and an additional effort was taken to equally represent all primary components of the forested ecosystem (e.g. soil, air, wildlife, etc.).

Table 1: Index developed to select priority rescues from the Data Rescue inventory list.

Index Parameter	Description	Weight
Temporal coverage	5 years	0.2
	10 years	0.5
	15 years	0.8
	20+ years	1
Spatial coverage	Localized site	0
	Regional within state	0.5
	Statewide	0.8
	Regional across states	1
Institutional value	Low	0
	Medium	0.5
	High	1

Depending on the needs of the organization that had the data, a priority rescue encompassed anything from scanning, digitizing, and archiving paper copies of data to converting outdated file types to contemporary formats. Documents were scanned to PDF files using Brother ADS-1700W and HP Photosmart C4599 feeder and table scanners at 300 dots per inch (dpi) resolution. Collections of 35mm slides were digitized by Green Mountain Camera.

To convert outdated files into CSV (comma separated value) tabular format, we used LibreOffice Calc 6.2.5.2. Files were retrieved from floppy disks using a LaCie External Floppy Disk Drive 1x provided by the University of Vermont Howe Memorial Library. Many of the tabular files were reformatted, namely one table per file with one row heading to fit the requirements of FEMC archive.

All rescued items were also included on the final Data Rescue inventory with methodology for how the material was rescued. Links to view the data are available which direct to the FEMC archive when applicable.

Inventory and Rescue Outcomes

Data Rescue Inventory

To date, more than 300 datasets and documents were identified as at-risk from respondents across the region. The material was 63% paper files to be scanned and 35% digital files to be formatted, ultimately to be uploaded to the FEMC archive. Three organizations had binders of 35mm slides to be scanned of tree canopy and vernal pool photos. In all, this material dated from 1626 to 2019 with the majority of at-risk data derived from wildlife and forestry studies conducted between 1986–2000 (Figure 1).

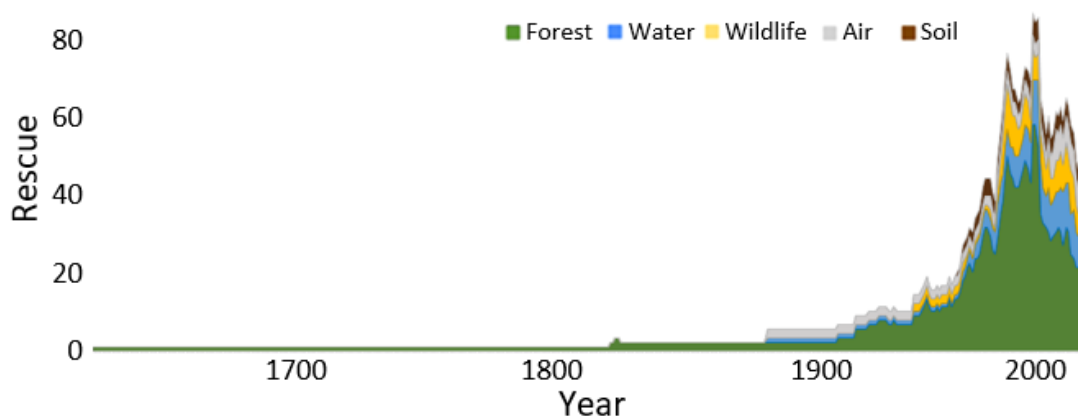


Figure 1: Years relating to the content of at-risk material categorized by key components of the forested ecosystem.

Of the 126 organizations that we contacted (Table 2), 38 responded with interest in preserving potentially at-risk material. All respondents opted to convey information about their data over the phone or in person rather than filling out the online form. From scheduled phone calls and in-person visits, we accumulated approximately 26 meeting hours.

Table 2: Location and type of organizations contacted for Data Rescue between April – November, 2019.

State	Federal Agency	State Agency	Academic	NGO	Total
MA		9	3	3	15
ME	3	15	2	1	19
NH	7	9	3	4	23
NY	1	10	16	7	34
VT	6	10	3	3	22
Regional	7			9	16
Total	24	53	27	27	126

To visualize and efficiently search for at-risk data, we developed a Data Rescue inventory web interface (Figure 2; https://www.uvm.edu/femc/data_rescue) which was reviewed by 12 environmental PhD students for usability and content requirement. This interface includes an interactive map displaying where the physical data were collected or are currently stored, and tables containing metadata relating to the data and documents including the rescue status of the material. A “Suggest a Data Rescue” link allows our stakeholders to recommend additional material to include on the Data Rescue inventory. This inventory is not only a way to identify communal resources at risk, but can serve as a first step in new monitoring, assessment, and research efforts, as interested parties can more easily discover and learn about material that could be utilized in novel ways.

Rescue Name	Status	Organization	State	Year	Theme(s)
1942 Report on Piscataquis County, Maine	Needs Rescuing	Baxter State Park	Maine	1942	Forest
1955 Map on Piscataquis County, Maine	Needs Rescuing	Baxter State Park	Maine	1955	Forest
Baxter State Park Historical Inventory Data	Needs Rescuing	Baxter State Park	Maine	1964 - 2016	Forest
Baxter State Park Inventory	Needs Rescuing	Baxter State Park	Maine	1996 - 2018	Forest, Soil
Baxter State Park Map	Needs Rescuing	Baxter State Park	Maine	1972	Forest
Estimate of Timber Volumes for Piscataquis County, Maine	Needs Rescuing	Baxter State Park	Maine	1971	Forest
Map of Forest Types in Piscataquis County, Maine	Needs Rescuing	Baxter State Park	Maine	1972	Forest
Stock and Stand Table for Piscataquis County, Maine	Needs Rescuing	Baxter State Park	Maine	1972	Forest
Webber Survey Remarks	Needs Rescuing	Baxter State Park	Maine	1832 - 1833	Forest
Catskill Bogs	Needs Rescuing	Cary Institute of Ecosystem Studies	New York	1995 - 2019	Water, Soil, Forest

Figure 2: Data Rescue inventory web interface developed to present at-risk data from across the Northeast. This web interface will be updated as more potentially at-risk data are identified.

Priority Rescues

We selected 21 organizations (Table 3) to participate in actual data rescue because they were rated as a high priority due to their risk of loss. These resources included 215 documents, datasets, surveys, and aerial images. The FEMC staff chose these rescues based on the Data Rescue index which weighted the spatial and temporal range of the data as well as their potential contribution of knowledge to a variety of agencies studying a diversity of disciplines. Priority rescues were also chosen with a practical understanding of what could be feasibly accomplished between April and December, 2019. To this effect, an organization was also selected if the rescue required minimal effort and did not interfere with the work of larger projects, or supported other FEMC regional

projects (see www.uvm.edu/femc/cooperative/regional_projects). Priority rescues related primarily to forests, wildlife, and water. Despite our efforts to choose priority rescues that equally examined all primary forest ecosystem components, 75% of the rescued projects related only to trees and other forms of forest vegetation.

Table 3: Priority rescues between April – November, 2019.

Organization	State	Contact	Rescue Material	Action Needed
Eastern Native Tree Society	MA	Bob Leverett	Trail guides	Archive
MA DCR, Bureau of Forest Fire Control and Forestry	MA	Ken Gooch	Misc. computer files	Archive relevant datasets
MA DCR, Bureau of Forest Fire Control and Forestry	MA	Bill VanDoren	Continuous Forest Inventory, 1960-2002	Scan and archive
MA DCR, State Archives	MA	Sean Fisher	Annual reports and defoliation maps, 1930-2010	Scan and archive
Baxter State Park	ME	Mike Pouch	Forest inventory and historic documents, 1626-2019	Scan, organize, and archive
Holt Research Forest	ME	Jack Witham	Salamander and small mammal monitoring, 1984-2019	Scan and archive
University of Maine	ME	Ivan Fernandez	Tree regeneration data, 1997-1999	Digitize and archive
Society for the Protection of NH Forests	NH	Gabe Roxby	Aerial imagery and forest inventory, 1953-2019	Scan and archive
US Forest Service	NH	Bob Cooke	North American Maple Project, 1988-1999	Convert outdated files
Cary Institute of Ecosystem Studies	NY	Jamie Deppen	Michael Kudish field notes from the Catskills, 1965-2019	Scan, archive, design portal
Stony Brook University	NY	Sharon Pochron	SBU Hotel Project environmental assessment, 2010	Archive
Herp Atlas	VT	Jim Andrews	Reptile and amphibian habitat reports and photo records, 1988-2010	Scan and archive
Missisquoi National Wildlife Refuge	VT	Ken Sturm	Species and land type inventories, 1966-2000	Scan and archive
VT DEC, Watershed Management Division	VT	Heather Pembrook	Benthic macroinvertebrate and vernal pool surveys, 1981-2010	Scan and archive
VT Fish and Wildlife	VT	Bob Popp	Development related projects with threatened or endangered plant species, 1990-2016	Develop new database and transfer data
VT Forests, Parks and Recreation	VT	Barbara Schultz	Maple decline study, 1977-1989	Convert files and scan, archive data

During the working session discussion, FEMC partners refined the Data Rescue framework by reexamining the priority index and exploring ways to improve outreach. The working session also connected the FEMC staff with several individuals interested in participating in the Data Rescue project or creating resources to help others rescue their own data.

Suggestions for modifying the priority index included elevating the weight of a rescue's temporal coverage as well as adding an 'immediacy' criteria that accounts for the condition of the data source or immediacy of a particular forest issue. Qualitatively, the FEMC staff would also consider the risk of losing institutional knowledge from a single site and account for how many institutions might benefit from rescuing the data.

Several opportunities were presented for the Data Rescue project to connect with other archiving organizations. Examples included the Digital Preservation Coalition which is an international Data Rescue effort, the Research Data Alliance which is a data conservation interest group, the Ecological Society of America, Mass Audubon, and the Maine State Museum. Collaboration with these organizations can reduce any redundant effort across the region and more efficiently address rescues by calling upon the strengths of each organization.

The FEMC is interested in maintaining the Data Rescue inventory and continuing to discover at-risk data. To maximize outreach of the Data Rescue project, the FEMC will organize a webinar on the topic, disseminate additional social media and newsletter material, and continue conversations with partners interested in exploring more ways to rescue and make use of data at risk.

Appendix – Organizations contacted for Data Rescue.

Table A1: Organizations contacted for participation in the Data Rescue effort between April – November, 2019.

Maine

Massabesic Experimental Forest
Penobscot Experimental Forest
USFS, retired forester
Office of Information Technology
Bureau of Parks and Lands
Department of Agriculture, Conservation and Forestry
Maine Forest Service
Maine Geological Survey
Maine Natural Areas Program
Bureau of Air Quality
Bureau of Water Quality
Department of Environmental Protection
The Biological Monitoring Program
Department of Inland Fisheries and Wildlife, Fisheries
Department of Inland Fisheries and Wildlife, Wildlife
Lands Management Program
Baxter State Park
Unity College
University of Maine, Holt Research Forest
The Nature Conservancy, ME

Massachusetts

Department of Conservation and Recreation (DCR)
DCR Forest Health Program
DCR Office of Watershed Management
Department of Environmental Protection (DEP)
DEP Division of Conservation Services
DEP Office of Research and Standards
Department of Fish and Game (MassWildlife)
MassWildlife Natural Heritage & Endangered Species Program
MassGIS, Bureau of Geographic Information
Hopkins Memorial Forest, Williams College
Smith College, MacLeish Field Station

Table A1 - continued: Organizations contacted for participation in the Data Rescue effort between April – November, 2019.

Massachusetts, continued

UMass Amherst, Department of Environmental Conservation
Berkshire Natural Resources Council
Mass Forest Alliance
The Nature Conservancy, MA

New Hampshire

USFS, Durham, NH Field office
USFS, Forest Entomologist
USFS, Forest Health Protection
USFS, Forest Watershed Specialist
USFS, Geologist
USFS, GIS analyst
USFS, White Mountain National Forest
DES Bureau of Air Resources
DES Bureau of Water
DNCR NH State Parks (Parks and Recreation)
DNCR NH Geological Surves
DNCR Division of Forests and Lands
DNCR Forest Health Program
DNCR Forest Management Bureau
Fox Research Forest
NH Fish and Game Department
Antioch University New England, Environmental Studies
Dartmouth College, Woodlands
University of New Hampshire, extension
Mount Washington Observatory
NH Audubon
Northern Forest Center
Society for the Protection of New Hampshire Forests

New York

USFS Fingerlakes National Forest
Department of Agriculture and Markets (NYSDAM)
Catskills Environmental Research and Monitoring (CERM)
DEC Division of Lands and Forests

Table A1 - continued: Organizations contacted for participation in the Data Rescue effort between April – November, 2019.

New York, continued

DEC NY Natural Heritage Program
DEC NYC Department of Protection
DEC Partners for Regional Invasive Species Management
Adirondack Park Agency
Geographic Information Systems Clearinghouse
NY State Museum
Ward Pound Ridge Reservation
Bard College
Cary Institute of Ecosystem Studies
Cornell, Arnot Forest
Cornell, NY Invasive Species Research Institute
Cornell, Uihlein Maple Research Forest
Fordham University, Louis Calder Center
Paul Smiths, Adirondack Watershed Institute
Siena College
St. Lawrence University
SUNY, Albany, Arts and Sciences
SUNY, Cobleskill
SUNY, ESF Huntington Wildlife Forest
SUNY, Stony Brook
Syracuse University, Civil and Environmental Engineering
Teatown Lake Reservation
Vassar College, Ecological Preserve
Buffalo Museum of Science
Empire State Forest Productions Association
Environmental Monitoring and Management Alliance
Huyck Preserve and Biological Research Station
Mianus River Gorge
Mohonk Preserve
The Nature Conservancy, NY

Vermont

Missisquoi National Wildlife Refuge
Marsh-Billings-Rockefeller National Historical Park
USFWS Fisherie

Table A1 - continued: Organizations contacted for participation in the Data Rescue effort between April – November, 2019.

Vermont, continued

USFS, GIS Coordinator
USFS, Green Mountain National Forest
USFS, Research Forester
ANR Acid Lakes Monitoring Program
ANR Air Quality and Climate Division
ANR Vermont Acid Precipitation Monitoring Program
ANR VT Geology Survey
ANR DEC Watershed Management Division
ANR Department of Forests, Parks and Recreation
ANR Private Lands Program
ANR Lands and Habitat Program
ANR VT Fish and Wildlife, Botanist
ANR VT Fish and Wildlife, Wildlife
Green Mountain College
Gund Institute for Environment
University of Vermont
The Nature Conservancy, VT
Vermont Center for Ecostudies
Vermont Family Forests
Calfee Woodland

Regional

US Fish and Wildlife Service
Forest Health Assessment and Applied Sciences Team
State & Private Forestry - Region 9
USFS, Forest Health Protection, Washington Office
USFS, Green Mountain and Finger Lakes National Forests
USGS, New England Water Science Center (NEWSC)
USGS, NEWSC, Sleepers River Research Watershed
Lake Champlain Basin Program
New England Association of Aquatic Biologists
New England Governors and Eastern Canadian Premiers
North Atlantic Landscape Conservation Cooperative
Northeast States for Coordinated Air Use Management
Northeastern Soil Monitoring Cooperative

Table A1 - continued: Organizations contacted for participation in the Data Rescue effort between April – November, 2019.

Regional, continued

Northeastern States Research Cooperative

Northern Forest Center

The Native Tree Society

Trust for Public Land



FEMC

Forest Ecosystem Monitoring Cooperative



The University of Vermont

The U.S. Department of Agriculture (USDA) prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, and where applicable, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, reprisal, or because all or part of an individual's income is derived from any public assistance program. (Not all prohibited bases apply to all programs.) Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotape, etc.) should contact USDA's TARGET Center at 202-720-2600 (voice and TDD). To file a complaint of discrimination, write to USDA, Director, Office of Civil Rights, 1400 Independence Avenue, S.W., Washington, D.C. 20250-9410, or call 800-795-3272 (voice) or 202-720-6382 (TDD). USDA is an equal opportunity provider and employer.

Providing the information needed to understand, manage, and protect the region's forested ecosystems in a changing global environment

This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

