A Proposed Method for Assessing the Vulnerability of Road-Stream Crossings to Climate Change:

Deerfield River Watershed Pilot

UMassAmherst

Paula L. Sturdevant Rees
Scott D. Jackson
Stephen B. Mabee

massDOT
Massachusetts Department of Transportation
Highway Division

Katherin M. McArthur

DECEMBER 2018

A Proposed Method for Assessing the Vulnerability of Road-Stream Crossings to Climate Change:

Deerfield River Watershed Pilot

DECEMBER 2018

Paula L. Sturdevant Rees
Scott D. Jackson
Stephen B. Mabee

UNIVERSITY OF MASSACHUSETTS AMHERST

Katherin M. McArthur

MASSACHUSETTS DEPARTMENT OF TRANSPORTATION

UMassAmherst | massDOT Massachusetts Department of Transportation Highway Division

WRRC University of Massachusetts Amherst | MASSACHUSETTS GEOLOGICAL SURVEY | The Center for Agriculture, Food and the Environment

ACKNOWLEDGMENTS

Design: Penny Michalak

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

ACOE	Army Corps of Engineers
AOGCMs	Atmospheric-ocean general circulation models
AEP	Annual Exceedance Probability
AMS	Annual Maxima Series
AR	Assessment Report
AUC	Area under the curve
BCSD	Bias-Correction Spatial Disaggregation
CAPS	Conservation Assessment and Prioritization System
CCAP	Climate Change Adaptation Plan
CMIP	Coupled Model Intercomparison Project
D	Height of culvert barrel
DA	Drainage area
DCHP	Downscaled Climate and Hydrology Projections
DEM	Digital Elevation Model
DOE	Department of Energy
DOI	Department of the Interior
DRW	Deerfield River Watershed
EEA	Massachusetts Executive Office of Energy and Environmental Affairs
EHP	Environmental and Historic Preservation
EMS	Emergency Medical Services
EPA	Environmental Protection Agency
ESRI ArcGIS	A commercial software used for mapping and spatial reasoning
FGDC-STD	Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata
FEMA	Federal Emergency Management Agency
FGM	Fluvial geomorphology
FHWA	Federal Highway Administration
FRCOG	Franklin Regional Council of Governments
GCM	Global Climate Model
GEV	Generalized Extreme Value
GOF	Goodness of Fit
GPS	Global Positioning System
HBV	Hydrologiska Byrans Vattenbalansavdelning model
HDS-5	Hydraulic Design Series 5
HDSC	Hydrometeorological Design Studies Center
HRU	Hydrological Response Unit
HSPF	Hydrologic Simulation Program Fortran

HW	Headwater depth
ICE	Interactive Catchment Explorer
IEI	Index of Ecological Integrity
LiDAR	Light Detection and Ranging
LRFD	Load Resistance Factor Design
IPCC	Intergovernmental Panel on Climate Change
IWR	Institute for Water Resources
LOOCV	Leave-one-out cross validation
MA	Massachusetts
MaPIT	MassDOT Project Intake Tool
MassDOT	Massachusetts Department of Transportation
MassDEP	Massachusetts Department of Environmental Protection
MC-FRM	Massachusetts Coastal Flood Risk Model
MassGIS	Massachusetts Department of Geographic Information Systems
MEMA	Massachusetts Emergency Management Agency
MMI	Milone & MacBroom, Inc.
MOVIT	Mapping Our Vulnerable Infrastructure Tool
NASA	National Aeronautics and Space Administration
NAACC	North Atlantic Aquatic Connectivity Collaborative
NARCCAP	North American Regional Climate Change Assessment Program
NALCC	North Atlantic Landscape Conservation Cooperative
NBIS	National Bridge Inspection Standards
NCA3	3rd National Climate Assessment report
NCAR	National Center for Atmospheric Research
NED	National Elevation Dataset
NEX-GDDP	National Earth Exchange Global Daily Downscaled Projections
NFIP	National Flood Insurance Program
NHDHRD	National Hydrography Dataset High Resolution Delineation
NLCD	National Land Cover Database
NOAA	National Oceanic and Atmospheric Administration
NPCC	New York City Panel on Climate Change
NRCC	Northeast Regional Climate Center
NRCS	Natural Resources Conservation Service
NSF	National Science Foundation
NWIS	National Water Information System

OLS	Ordinary Least Squares
PCMDI	Program for Climate Model Diagnosis and Intercomparison
PE	Professional Engineer
PRISM	Parameter-elevation Regression or Independent Slopes Model
Q$_{critical}$	Critical flow
RCPs	Representative Concentration Pathways
RCM	Regional Climate Model
RI	Return Interval
RMSE	Root Mean Square Error
RPA	Regional Planning Authority
RPFE	Regional Peak Flow Equations
SCE	Stream Crossing Explorer
SimCLIM	A climate change analysis software
SHEDS	Spatial Hydro-Ecological Decision System
SRES	Special Report on Emissions Scenarios
SSP	Specific Stream Power
SSURGO	Soil Survey Geographic
TI	Topographic Index
TNC	The Nature Conservancy
TU	Trout Unlimited
TW	Tailwater
UMass	University of Massachusetts Amherst
U.S.	United States
U.S. DOT	United States Department of Transportation
USFWS	United States Fish and Wildlife Service
USGCRP	United States Global Change Research Program
USGS	United States Geological Survey
VT	Vermont
WCRP	World Climate Research Programme
WGCM	Working Group on Coupled Modeling
WRFH	WRF-Hydro model

Executive Summary

Over the last decade, both total rainfall and the number of heavy precipitation events have increased across the Northeast, and the general consensus of the scientific community is that this trend will continue. In 2011, the destruction associated with Tropical Storm Irene highlighted the vulnerability of the transportation system to flooding and erosion. This federally recognized major disaster served as a wakeup call for communities across northwestern Massachusetts, emphasizing the need for increased resilience to extreme weather events.

Road-stream crossings are a critical component of the transportation system, and are particularly vulnerable to flooding damage and failure. The goal of the Deerfield River Vulnerability Pilot Project (Deerfield Project) was to develop a rapid assessment and prioritization methodology for evaluating and ranking the vulnerability of road-stream crossings to extreme weather and climate change. This methodology was to be of use to both local communities and the Commonwealth and have the potential to be implemented beyond the original study area. Understanding the range and underlying causality of uncertainties in the methodology was a second major goal.

The pilot study was located in the Deerfield River watershed, which straddles the border between northwestern Massachusetts and southern Vermont and included approximately 1,002 road-stream crossings in Massachusetts. The watershed experienced multiple culvert failures, damage to bridges, and associated transportation disruptions as a result of runoff volume, bank erosion, undermining, and landslides caused by Tropical Storm Irene. This history of documented damage and stakeholder awareness made the Deerfield River watershed a good location for methodology development.

Project Deliverables

There are five main deliverables for the project. The first is this report. Second, a large amount of very diverse data were collected and generated over the course of the Deerfield Project. The majority of these files have been organized in a single, relational Access database; additional data are provided as separate zipped Excel files, and all data are included in the deliverables. Third, GIS files are provided as part of the project database. Metadata for the Deerfield Project were prepared for all spatial data layers, rasters and summary spreadsheets. Fourth, a data visualization and decision support tool was developed to assist with locating and prioritizing stream crossings that meet user-defined criteria for vulnerability. This tool, known as the Stream Crossing Explorer (SCE), can be

seen at http://sce.ecosheds.org. SCE conveys information related to the risk of failure of road-stream crossings and associated disruption of emergency services, along with aquatic connectivity and ecological aspects of river systems. This tool was designed for use by state and municipal agencies, local decision-makers, regional planners, conservation organizations, and natural resource managers. Lastly, to advertise and share the Deerfield Project results through the SCE, an outreach effort was conducted.

Assessing Vulnerability and Prioritization

As part of the project, methodologies were developed to evaluate linked components of road-stream crossing transportation vulnerability and ecological disruption.

Transportation Vulnerability includes three factors:

1. **Risk of Failure**, based on the integrated risk posed by three potential road-stream crossing failure mechanisms (structural failure, hydraulic failure, and geomorphic failure),

2. **Climate Change**, which incorporates potential increases in hydraulic failure risk due to changes in precipitation and temperature caused by climate change, and

3. **Criticality** of each crossing based on its potential to disrupt emergency medical services.

Ecological Disruption accounted for two factors:

1. **Connectivity Loss**, which assesses the amount of aquatic connectivity lost due to the barrier effects of each road-stream crossing, and

2. **Connectivity Restoration Potential**, which combines Connectivity Loss with a measure of habitat quality (ecological integrity).

The scores for Transportation Vulnerability and Ecological Disruption were combined to develop a scoring system for overall prioritization of crossings.

Road-Stream Crossing Data

Existing data sets that were used included Laser Imaging Detection and Ranging (Lidar) data previously collected for a USGS flood-inundation mapping study, Massachusetts Department of

Transportation (MassDOT) structure inspection field reports for bridges in the basin, and information collected from local municipalities about the history of flooding and related failures in the transportation network. Data for both the Vermont and Massachusetts portions of the watershed necessary to support basin-wide geomorphic and hydraulic analyses were obtained from publically available GIS data.

The project utilized field assessment methodologies developed by the Massachusetts River and Stream Continuity Partnership (a precursor to the North Atlantic Aquatic Connectivity Collaborative or NAACC). Trout Unlimited-trained technicians were contracted to conduct the field assessments of road-stream crossings based on the NAACC protocols for assessing aquatic passability, as well as supplemental information necessary to evaluate culvert condition, geomorphic vulnerability, and hydraulic capacity. Field assessment of non-road crossings such as railroad bridges and dams was not included in the scope of this project. Aquatic passability and culvert condition data collected in support of the project are stored in the NAACC database at naacc.org.

A licensed professional engineer from Milone & MacBroom Associates (MMI) was contracted to collect identical field data at a random 10% of the road-stream crossings. While a qualitative comparison between TU and MMI collected field data indicates some differences, these differences had little impact on the overall analysis. The Deerfield Project demonstrates that 1) field technicians can meet the data collection quality control needs for this and similar projects with sufficient training, and 2) the NAACC rapid field assessment protocol provides sufficient data for the assessment of multiple factors contributing to crossing vulnerability.

Climatic Data

In addition to predictions of future stream flows, an understanding of how uncertainty in future climate condition translates into uncertainty in streamflow is needed to effectively evaluate the future vulnerability of the transportation network. This requires the evaluation of multiple climatic models and atmospheric condition scenarios to determine which ones are most applicable. In addition, different types of streamflow models require different types of data. In particular, deterministic or physically based hydrologic models require higher resolution data than statistical hydrologic models. In order to develop useful climate projections, the data generated by Global Climate Model (GCM) simulations had to be downscaled to a scale appropriate to the hydrological analyses performed as part of the project.

Climate prediction data from four sources were considered for the Deerfield Project. These data included:

1. GCM data produced through the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project phase 3 (CMIP3) and phase 5 (CMIP5), available through a United States Department of Transportation (U.S. DOT) CMIP Climate Data Processing Tool,

2. Dynamically downscaled GCM-Regional Climate Model (RCM) climate predictions from the North American Regional Climate Change Assessment Program (NARCCAP) based on CMIP3,

3. Dynamically downscaled GCM-RCM climate predictions from the NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) based on CMIP5, and

4. Climate simulations by the National Center for Atmospheric Research (NCAR) Community Climate System Model (CCSM4) downloaded from the NCAR Climate Inspector.

Data suitability was assessed both based on intended use and model needs, and on accuracy based on comparison with observed station data.

Linked Component Evaluation

The scoring system for overall prioritization of crossings was developed based on the integration of assessments of transportation vulnerability on one side, accounting for risk of failure and criticality of the crossing, and ecological disruption on the other, taking into consideration barriers to aquatic organism passage, and potential benefits if the crossing were to be improved. The methodology for developing the various scoring regimes created as part of the project is described below.

1. Transportation Vulnerability Sub-score: Risk of Failure

Risk of Failure was determined as the combination of three component scores (structural, hydraulic and geomorphic risk) into a single score representing the overall potential for crossing failure.

Individual Component Risk Score Calculation

1A. Structural Failure Risk (Chapter 5.1)
Structural Failure Risk scores were developed with the assistance of a technical advisory committee made up of Jim MacBroom, Roy Schiff and Matthew Gardner (MMI) and Scott Civjan (UMass Amherst Department of Civil and Environmental Engineering). Two sources of data were used to score crossings: (1) data collected in the field for culverts using the culvert condition assessment protocol and (2) MassDOT bridge inspection data for bridged crossings in the watershed.

The structural risk of failure for culverts was determined based on the maximum of three subscores, assuming that the vulnerability of a given crossing is based on its weakest "link." The first subscore identified culvert crossings which appeared to be structurally deficient enough to be at risk of imminent failure (super critical); the second subscore identified culvert crossings where up to three less severe but still significant structural deficiencies were noted (critical); and the third subscore identified culverts where an accumulation of moderately concerning deficiencies posed an increased risk of failure (poor).

The structural risk of failure for bridges was determined based

on two metrics from MassDOT's bridge inspection data, metric (60) – "Substructure" and (61) – "Channel & Channel Protection." A crosswalk was developed to convert the MassDOT bridge inspections scores, which range from 0 (bad) to 9 (good), to the 0 (good) to 1 (bad) Deerfield Project structural risk scoring system.

Structural Failure Risk scores ranged from zero (lowest risk) to one (highest risk). For crossings with multiple culverts, each culvert was individually assessed and the highest (worst) score was assigned to the crossing. Of the 801 crossings that were assessed for Structural Risk of Failure, ~80% received risk scores <0.5 while ~18% were scored as being high risk (scores > 0.8) and 6.4% fell in the highest risk category (score = 1.0).

1B. Geomorphic Failure Risk (Chapter 5.2)

Geomorphic Failure Risk was based on an assessment of the geomorphic (landscape) context of each crossing within the watershed as well as the vulnerability of each crossing to geomorphic stresses. The assessment was completed in two phases. The first consisted of a desktop based, watershed scale analysis of specific stream power at the stream reach level. Stream power is a measure of the potential for the stream to change the geomorphic form of the bed and banks of a channel. Specific or unit stream power is power per unit area of a channel. The second phase was a detailed local-scale assessment of the geomorphic conditions at each crossing to evaluate both the evidence of and propensity for the crossing to experience scour and blockage. Four scoring categories were developed based on data resulting from the phased analyses. The scoring categories and contributing factors were:

- Woody debris, based on structure-reach alignment and absolute structure width,
- Sedimentation, based on stream power, structure width ratio and level of stream aggradation,
- Scour, based on stream power, level of stream erosion, level of footing scour, and extent of a downstream scour pool, and
- Blockage, based on observed blockage reported on the conditions survey.

Each category score ranged in value from 0 (poor condition) to 4 (good condition), which was normalized to a 0 to 1 range. The overall geomorphic score was determined from the lowest condition score of the four categories, which was advanced as the overall score for geomorphic risk of failure after inverting it from a condition score to a risk of failure score ranging from 0 (low) to 1 (high).

A total of 811 crossings were assessed for geomorphic risk. The median score for all 811 crossings was 0.50 and the average score was 0.52. A total of 48 crossings had a score of 1.0, suggesting a greater risk of failure, and 25 crossings had a score between 0.1 and 0.2, indicating a minimal risk of failure. Crossings with the highest risk of failure are almost universally due to observed blockage of the crossing structure.

1C. Hydraulic Failure Risk (Chapter 5.3)

Hydraulic Failure Risk was evaluated based on the perceived ability of a road stream crossing to pass a critical flow, defined as the maximum streamflow a road-stream crossing can accommodate before damaging the road subsurface or overtopping the road. Hydraulic risk determination consists of two parts: (1) calculation of the critical flow for a given location, and (2) determination of the likelihood, relative to the other crossings in the watershed, that the critical flow will be exceeded.

Estimates of critical flow were based on the evaluation of each individual crossing structure's size, shape, and construction material. Based on that information, allowable water levels were defined (for instance, concrete culverts were considered at risk of failure if water reached a level of greater than 1 foot below the road surface). Whenever sufficient data were available, Federal Highway Administration Hydraulic Series 5 (HDS-5) methodologies were utilized to determine culvert capacity, using the commercially available CulvertMaster software by Bentley. Manning's Equation was utilized to estimate culvert capacity whenever CulvertMaster could not be applied, typically due to size constraints (e.g., the maximum culvert dimensions allowable in CulvertMaster are exceeded) or lack of data.

Likelihood of exceedance of the critical flow was determined under current, mid-century (2041 – 2070), and end-century (2071 – 2090) climatic conditions utilizing a multiple model framework for estimating several return interval streamflow discharges, specifically the annual peak discharge exceeded on average once in a 10-, 25-, 50-, and 100-year period at each of the road stream crossings in the watershed. The framework included two types of models, process-based (commonly referred to as physical or physically-based models) and statistical models (also referred to as empirical or stochastic models). Four regression type statistical models were included: the United States Geological Survey regional peak flow equations for Massachusetts, for Vermont, and for New Hampshire, as well as the Jacobs equation. Three physical models were incorporated, including the Hydrologic Simulation Program Fortran (HSPF), the Hydrologiska Byrans Vatten-balansavdelning (HBV), and the WRF-Hydro (WRFH) models. Output from the physical models consisted of a daily peak flow, from which return interval flood flow estimates were extrapolated using statistical techniques. Lower, most-likely, and upper confidence intervals (CIs) were defined in order to capture uncertainty.

A wide range of potential future climate scenarios were considered to develop the return interval flood flow estimates. Multipliers derived from applicable scientific literature, the U.S. DOT CMIP processing tool, and the NCAR Climate Inspector tool were applied to the historical data utilized as input to the statistical models. In contrast to the statistical models, which only require annual or monthly precipitation data, the physically based models require daily climatic data, which are available from only two of the climatic data sets considered for the project (NARCCAP CMIP3 and NEX-GDDP CMIP5). Comparisons of the NARCCAP CMIP3 and NEX-GDDP CMIP5 annual and monthly precipitation totals against observed historical data indicated that NEX-GDDP precipitation totals were erroneously low across the board, and NEX-GDDP data were thus excluded from further use. The nine

sets of GCM-RCM projections based on CMIP3 from the NARCCAP study were utilized after bias correction as input to drive the physically based models to estimate mid-century (2041-2070) streamflow. Only mid-century flow estimates were directly available for the physically based models because end-of-century precipitation estimates were not available from the NARCCAP study. Physical model end-of-century estimates for streamflow were developed by extrapolating the mid-century results.

A logistic equation was utilized to define hydraulic risk for each crossing based on the critical flow and likelihood of exceedance data. The equation provides a continuous (versus discrete) range of hydraulic risk values from 0 (low risk of failure) to 1 (high risk of failure) for the scoring system. The equation assigns the midpoint of the curve (i.e., a risk score of 0.5) based on a value of 1.0 for the ratio of the crossing's critical flow value and the estimate for the 25-year (Q_{25}) return interval flow for that location. Steepness of the logistic function curve is set such that crossings with a critical flow value twice that of Q_{25} have essentially no risk of hydraulic failure, while crossings where the critical flow is very small in comparison to Q_{25} get scores close to 1.0. Hydraulic risk scores were developed for the individual models for current climate and each future climate scenario. In addition, hydraulic risk scores were calculated for the low-, most-likely, and upper-CI physical model estimates.

Hydraulic risk scores take into consideration uncertainty with regards to model selection, streamflow estimation, and future climate. Results show that the physically based models generally assign lower risk scores than the statistically based models do, but the models tended to identify the same crossings as having relatively higher or lower risk. The statistical models predict only minor shifts in hydraulic risk scores from current to mid-century, while the physically based models predict a larger shift to a higher hydraulic risk of failure at mid-century. The physical model risk scores vary considerably across hydrologic and climate models, but there is greater uncertainty in terms of risk scores at mid-century due to the uncertainty surrounding streamflow estimates than due to the range of climate projections.

An ensemble hydraulic risk score was developed to incorporate all of the results into a single score. The ensemble score combines the results from the physical and statistical models to provide a more balanced prediction of risk.

Approximately 19% of the road-stream crossings in the Deerfield River watershed have a hydraulic risk score > 0.8 under current conditions. This number is estimated to increase to ~29% by mid-century, and to ~39% by end-century.

Total Risk of Failure Score Calculation (Chapter 6.2)

Both hydraulic and geomorphic risk scores are based on the likelihood that a crossing will fail during or after a significant storm event, and it is assumed that there is little risk of hydraulic or geomorphic failure in the near future without a storm. The more severe the storm, the more likely that a particular structure will fail. However, the nature of the relative scoring system used to evaluate

geomorphic and hydraulic risk means that we can't say much about the exact probability of failure, only that there is some vulnerability. In contrast, structural risk is based on an assessment of bridge or culvert condition at each crossing. Although storms may increase the risk of structural failure, crossings rated as being at high risk are at risk under all weather conditions.

Because the mechanism of failure with the highest score is the most likely mechanism to cause a crossing failure, the overall risk of failure score for each crossing was computed using the highest of the three component scores. An Overall Risk of Failure score was calculated even if data were available for only one or two of the three metrics.

Of the 830 crossings for which an Overall Risk of Failure score could be calculated, a large percentage of crossings (57.5%) had overall risk scores ≥ 0.6 and 184 crossings (22.2%) had risk scores ≥ 0.9. The results of this risk of failure analysis suggest that a small percentage of crossings (51 crossings) may be at high risk of structural failure and should be inspected by a qualified engineer in the near future. These results also suggest that a significant percentage of crossings in the Deerfield River watershed may be vulnerable to storm-related failure due to issues related to geomorphic vulnerability or hydraulic capacity and, given the expectation that climate change will produce future storms of increasing severity as well as more frequent severe storms, these issues are anticipated to get worse over time. Many crossings with overall risk of failure scores ≥ 0.9 (107 crossings under current conditions) received those high scores due to hydraulic risk. Note that hydraulic failures may or may not be catastrophic failures. In some cases, such a failure could result in the complete loss of a crossing; in other cases, it might result in a temporary closure due to water overtopping the road.

The spatial viewer includes data for each component of risk as well as the total risk of failure score.

2. Transportation Vulnerability Sub-score: Criticality (Chapter 5.4)

Crossing failures during extreme storms and flooding events can block essential transportation routes and severely disrupt the ability of communities to provide critical emergency services. For this component of the project, emergency management and network analysis experts collaborated to identify critically important road-stream crossings based on the impacts their failure would have on emergency medical services (EMS), specifically the response times for ambulances and subsequent transport to hospitals.

First a network model based on actual EMS call data from the Shelburne Communication Center from 2011-2015, population data, digital road maps, speed limit data, ambulance dispatch locations, and hospital locations was developed and utilized to establish a probabilistic distribution of EMS trips. The model was then used to simulate the effect of crossing failures on the transportation system. This methodology assessed only one crossing failure at a time. In working on this project, it became clear that the ability to analyze multiple failures would be extremely beneficial. Emergency

managers identified four metrics of interest, including average delay in terms of hospital arrival (which includes trips unaffected by the failure), average affected delay (excludes trips unaffected by the failure), maximum delay, and an integrated metric for overall delay accounting for both the number of trips affected by the crossing failure and the magnitude of delay for each affected trip.

Results for the various Disruption of EMS metrics paint very different pictures of which crossings are most important for maintaining these critical services. The Average EMS Delay metric results in only a handful of crossings, most located on Route 2, scattered among the higher disruption scores with the remaining crossings clustered in the low score categories. Average Affected Delay scores are also strongly skewed toward the low end of the disruption scale. The crossings with the highest Average Affected Delay scores occur on small roads, presumably dead end roads that lack alternative routes, while many of the crossings with average delays of 10-20 minutes are on Route 2. The Maximum EMS delay scores show the same skewed distribution as for Average Delay and Average Affected Delay scores. Twenty-five crossings on small roads with few or no alternative routes had the highest possible maximum delay scores of 60 minutes. Thirty crossings had maximum delay scores of between 10 and 20 minutes; again, these crossings tended to occur on major routes such as Routes 2, 8A and 112. The Overall EMS Delay metric was intentionally set up (transformed) to avoid strongly skewed results, yielding a broad range of scores distributed throughout the watershed. Many of the crossings with the highest disruption scores occurred on highways and larger roads, with many moderately high scores occurring on smaller roads.

The integrated metric for overall delay was subsequently rescaled from 0 (low disruption) to 1 (maximum disruption) for incorporation with the other linked components to set priorities for upgrading the transportation infrastructure.

3. Transportation Vulnerability Score (Chapter 6.3)

As defined for the Deerfield Project, vulnerability is a combination of risk and criticality. In the transportation domain, Risk of Failure (calculated from the three risk factors - structural, geomorphic, and hydrologic) was combined with Criticality (based on Overall EMS Delay) to create Transportation Vulnerability scores. The vulner-ability score was calculated as the product of the risk and criticality scores. Transportation Vulnerability scores range from zero (low risk) to one (high risk). If data were available for only one of the two scores, then no Transportation Vulnerability Score was calculated.

Transportation Vulnerability scores were computed for 910 crossings. Just over 35 percent of crossings have low Transportation Vulnerability scores (scores ≤ 0.1) indicating that they had low scores for risk, criticality, or both. The other 65 percent of crossings have scores that are distributed throughout the scoring range, providing a relative ranking that can be easily used for transportation decision-making. It is important to remember that these scores are

based on a measure of criticality that only considered EMS. It is likely that some of these low scoring crossings could score higher if other elements of criticality were included.

4. Ecological Disruption Maps: Connectivity Restoration Potential (Chapter 5.5.2)

Dams and road crossings can disrupt aquatic connectivity, resulting in significant impacts on river and stream ecology. Dams generally present more severe barriers to the movement of wood, sediment and aquatic organisms than road crossings, but crossings can also affect these ecological processes and are much more numerous than dams. The collective impact of road crossings can significantly disrupt stream ecology and reduce the viability of populations of fish and wildlife. The connectivity restoration potential analysis for this project considered both connectivity loss and habitat quality in terms of Index of Ecological Integrity (IEI, a measure of habitat quality) scores computed by the Conservation Assessment and Prioritization System (CAPS) which could be evaluated separately or in combination. Note that these factors are combined with cold water restoration potential (described in the next section) for the ecological disruption score, described in Section 6.

Connectivity Loss in this project is defined as the loss of aquatic connectivity without consideration of habitat quality. Aquatic passability was assessed at road stream crossings using a protocol developed by the UMass Stream Continuity Project (a precursor to the NAACC) to identify to what extent a structure is blocking the movement of aquatic organisms. Data on aquatic passability were then utilized to assess the loss of aquatic connectivity using Critical Linkages, a component of CAPS. Critical Linkages is used to assess the effects of culvert upgrades and dam removals. It uses the CAPS dataset and one of the CAPS metrics, aquatic connectedness, to assess each crossing replacement or dam removal in turn. In order to maintain a consistent system of scoring for data used in the Deerfield Project and the spatial data viewer, we converted the passability scores to impassability scores using the simple mathematical function (1 – passability). Connectivity Loss maps are available in the spatial data viewer so that users can use the score as is or combine it with measures of habitat quality.

Connectivity Restoration Potential scores are also provided in the spatial data viewer. These scores are based on a standard Critical Linkages analysis of all stream reaches, combining Connectivity Loss and CAPS IEI.

5. Ecological Disruption Maps: Cold water Restoration Potential (Chapter 5.5.1)

When evaluating the disruptive influence of road-stream crossings on stream ecology, it is important to take into account the extent to which crossings disrupt aquatic access to high quality habitat. Beyond aquatic connectivity, habitat suitability must be assessed to address the conservation needs of species associated with cold water

habitats. Towards this end, a new stream temperature model was developed and applied throughout the Northeastern U.S.

Input for the stream temperature model was derived for each stream reach from GIS data. Predictor variables for temperature included total drainage area, riparian forest cover, daily precipitation, upstream impounded area, percent agriculture, and percent high intensity development. The model was applied to the portion of the year between the spring and autumn breakpoints in order to model the non-winter, ice-free stream temperatures in each reach. Overall, the model fit the historic temperature data quite well. The estimated error was 0.59°C when all data were used and 2.03°C for the model validation data set.

Metrics derived from the model were used to identify stream reaches that would qualify as cold water streams. Specifically, the temperature model was used to identify stream segments that would qualify as cold water streams based on mean summer temperature thresholds of 14°C, 16°C, 18°C, 20°C, and 22°C. The stream temperature model results were used in a special application of Critical Linkages for assessing aquatic connectivity restoration potential for cold water streams under different climate scenarios, such as how the amount and distribution of cold water streams, defined using a 16°C threshold, would change in the Deerfield River watershed with a 2°C rise in water temperatures. The cold water Critical Linkages analyses were funded by a separate project (a USFWS Hurricane Sandy Recovery and Mitigation grant).

Results of the stream temperature model for the Deerfield River watershed are included in the spatial data viewer so that users can evaluate the effects of climate change on water temperatures and draw connections with aquatic connectivity.

Habitat quality was measured by using the stream temperature model along with other GIS data, to model eastern brook trout occupancy (probability of occurrence) for stream reaches throughout the region. The occupancy model for brook trout was developed based on presence/absence data from agencies and landscape data. Predictions were made under current environmental conditions and for future increases in stream temperature.

Input for the brook trout model was derived for each reach from GIS data. Predictor variables included total drainage area, riparian forest cover, the mean of the summer daily precipitation record, the mean July stream temperature, upstream impounded area, percent agriculture, and percent high intensity development.

The model was then utilized to predict brook trout occupancy, including estimates of the probability of occupancy for each stream reach with increases in stream temperature of 2, 4 or 6°C. This was done by simply increasing input values for mean July stream temperature by 2, 4, or 6°C and estimating occupancies. Results of the probability of brook trout occupancy model for the Deerfield River watershed are included in the spatial data viewer so that users can evaluate the effects of road-stream crossings on brook trout populations using both aquatic connectivity and habitat quality.

6. Ecological Disruption Vulnerability Score (Chapter 6.4)

As previously defined for the Deerfield Project, vulnerability is a combination of risk and criticality. In the ecological domain, vulnerability takes into account the potential for restoring aquatic connectivity via crossing replacement, regardless of temperature, and for cold water streams with mean summer temperatures ≤ 16°C. The score for Ecological Disruption was determined as the maximum of the two component scores. If data for only one of the two component metrics were available, then the Ecological Disruption score was the same as the score for the available metric. About 25 percent of crossings offered little restoration potential (low scores of 0.0 to 0.1) and the remaining crossings were well distributed throughout the scoring range. Ecological Disruption scores were more influenced by the Connectivity Restoration Potential sub-score (dictates 79.5% of scores) than by Cold water Restoration Potential (dictates 20.5% of scores).

Crossing Prioritization (Chapter 6.5)

The Crossing Priority score was designed to combine the Ecological Disruption and Transportation Vulnerability scores. A raw priority score was first calculated as the maximum of the two vulnerability scores added to an average of both scores. The final Crossing Priority score was the raw score rescaled to a range from zero to one by dividing the raw score for each crossing by the maximum across all crossings. The objective was to create a combined score that ensures that a high score for one of the domains is not canceled out by a low score for the other domain, and that can identify as high priorities those crossings that rate highly in both domains. If data were available for only one of the two scores, then no Priority score was calculated.

A total of 770 crossings received priority scores. The scores are well distributed through the range, which is important for priority setting and decision-making.

Overview of Findings

There were three broad objectives for the Deerfield Project.

1. Develop methodologies for assessing structural, geomorphic and hydraulic risk of failure for road-stream crossings and the associated disruption of emergency medical services.

2. Incorporate climate change into hydraulic analyses using down-scaled climate models to predict future flows and hydraulic risk of failure for road-stream crossings.

3. Use these new methodologies, along with existing method-ologies for assessing ecological disruption, to assess the vulnerability of road-stream crossings in the Massachusetts portion of the Deerfield River watershed to extreme weather and climate change.

The development of methodologies associated with hydraulic risk proved to be the most complex and difficult task. Because many crossings (especially culverted crossings) occur on small waterways that lack stream gauges, we lack data necessary to validate the model predictions at all basin scales. Therefore, we have no way to judge whether physical models or statistical models are more accurate in predicting flows at road-stream crossings with drainage basin areas less than those for which gaged data are available. Although scores derived from statistical models were quite different from those from those generated by physical models, crossings that scored highest for risk as predicted by physical models were generally among the crossings that had the highest risk as predicted by statistical models. Thus, data from any of these models can be used to evaluate the relative risk of failure due to hydraulic capacity. The results from the Deerfield Project give people who want to evaluate the risk of crossing failure based on hydraulic capacity a sense of how the models used in this study compare: physical models produce lower risk estimates while statistical models produce higher ones. However, we cannot know for sure which of these estimates is more likely to be correct.

At the outset of the Deerfield Project, we planned to determine hydraulic risk based on whether crossings would pass a 10, 25, or 50-year storm. In implementing the project and analyzing the results, we decided that this type of evaluation would suggest a level of accuracy that we cannot reasonably claim. Instead, we decided to score hydraulic risk of failure based on relative rather than absolute risk. Our scores range from zero (low risk) to 1.0 (high risk). Ensemble averages of Q_{25} estimates from the hydrologic models were compared to crossing capacity to calculate a risk score. We chose Q_{25} for this analysis because these storms are unusually severe, but common enough that it is likely that many crossings will experience Q_{25} flows at some point in their design life. At this time, it is impossible to say what scores equate to moderate, high or extremely high risk categories in an absolute sense. However, we are confident that a crossing with a score of 0.9 is at higher risk of hydraulic failure than a crossing with a score of 0.7. Future storms may provide opportunities to calibrate these scores to actual failure risk.

The data collected and hydrological and hydraulic analyses conducted as part of the Deerfield Project were intended for use in a qualitative assessment and are not sufficient for use in the design of road-stream crossings. Hydrological models may have been updated since the work for the Deerfield Project was completed. While future hydrology and hydraulic risk predictions were considered in this report, extrapolating recent climate trends may not be effective for predicting future conditions because climate change can create new states that no longer adhere to past patterns. Climate models are still in early stages of development, and the various models used for the Deerfield Project yielded highly variable results.

The geomorphic vulnerability score is assigned as the highest of four component scores, three related to separate mechanisms for geomorphic failure (sediment blockage, blockage by woody debris, and scour), and a fourth based on the presence/absence and severity of blockages documented at the crossing at the time of assessment.

The scoring system results in a reasonable distribution of risk scores. Forty-six crossings received geomorphic vulnerability scores of 1.0 because field assessments revealed that they were already significantly blocked by sediment and/or woody debris.

Structural Risk of Failure was based on MassDOT bridge inspection results or, if those were not available, an ends-only rapid assessment of culvert condition. Most crossings fell into low risk categories, with only about 20 percent scoring 0.5 or higher. Theoretically, high-risk crossings could fail at any time. However, it is conceivable that these structures would be at even higher risk during severe storm events.

The calculation of an overall risk of failure score by taking the maximum score for structural, hydraulic and geomorphic risk yields a distribution skewed toward higher scores (higher risk). That is because there is little correlation among the three components of risk. As it turns out, many crossings are at relatively high risk for at least one of these components. At this point, we cannot relate our risk of failure scores to a probability of failure. Although we attempted to create scoring systems that provide comparable scores for structural, geomorphic and hydraulic risk (meaning that similar scores would equate to similar probability of failure) only with time and monitoring will we be able to identify numeric scores most indicative of high, moderate and low probability of failure.

We used the Deerfield Project to develop computer heuristics and algorithms that would allow us to measure the criticality of each structure by identifying alternative routes for the provision of EMS when faced with crossing failures. A unitless metric was developed based on a logistic weighting function to account for the number of trips affected and the magnitude of delays. This "Overall Delay" metric tends to spread the scores more evenly throughout the range of zero to 1.0. When crossing failure risk, which is skewed high, is combined with criticality, the resulting Transportation Vulnerability scores are reasonably distributed. Disruption of EMS was the only aspect of criticality assessed in the Deerfield Project. Other elements of criticality that were not assessed include access to critical infrastructure (water supplies, wastewater treatment facilities, electrical substations, gas compressor stations, etc.) and the broader need to facilitate vehicular traffic.

The ecological component of the Deerfield Project used existing assessment protocols and applied existing connectivity models to evaluate how each crossing affects aquatic connectivity. We also ran a specialized version of Critical Linkages focusing on connectivity for cold water streams. A score for Ecological Disruption was calculated as the maximum score between restoration potential for the base analysis or for cold water streams using a 16°C threshold. The number of crossings affecting cold water streams is significantly less than the total number of crossings in the watershed. As a result, the Ecological Disruption scores are controlled by the base assessment for about 80 percent of crossings, with 20 percent of crossing scores dictated by cold water restoration potential. It is not possible at this time to relate our ecological disruption scores to passability for particular species of fish or other aquatic organisms.

One question that we wanted to answer with the Deerfield

Project was to what degree the results of crossing assessments for aquatic passability could be used to evaluate risk of failure. Impassability is fairly strongly correlated with overall risk of failure. The relatively high correlations probably stem from the fact that undersized crossing structures tend to both be highly disruptive of aquatic organism passage and vulnerable to geomorphic and hydraulic failure. This suggests that impassability scores derived from NAACC data could serve as an imperfect but reasonable stand-in for risk of failure if data from geomorphic and hydraulic risk assessments are not available.

Our assessments should not be used on their own to make decisions about crossing repair or replacement. They are intended to be used as a screening tool to draw attention to crossings that warrant further, more detailed analysis for risk of failure, restoration of aquatic connectivity, or criticality for provision of emergency services.

Project Overview

1.1 Rationale for Project

Tropical Storm Irene, which hit Massachusetts in August 2011, had a tremendous effect on Western Massachusetts and Vermont, with many hundreds of road-stream crossings failing or suffering damage due to the storm. Irene demonstrated the vulnerability of the transportation network to storm damage and the role of road-stream crossings as potential weak links in that network. Aging infrastructure and the anticipated impacts of climate change necessitate the assessment of vulnerabilities in, and resiliency of, state transportation networks.

Rivers and streams, like transportation networks, are widely spread across the landscape providing many opportunities for intersections with roads. These crossings are points of potential vulnerability for transportation infrastructure and riverine ecosystems. Many crossings are partial or nearly complete barriers to aquatic organism passage. Of 5,143 Massachusetts crossings in the North Atlantic Aquatic Connectivity Collaborative (NAACC) database, 13.4 percent are severe barriers and 30.7 percent range from moderate to severe. If we look only at culverted crossings, 18 percent are severe barriers and 42 percent are moderate to severe barriers to aquatic passage.

Road-stream crossings are a critical, and sometimes vulnerable, component of the transportation system. Flooding and erosion associated with severe storms can disrupt transportation networks and thus the ability to provide essential services. Crossing failures can be more than an inconvenience; they can threaten public safety and result in significant economic impacts. For example, damage to road-stream crossings during Irene left some Vermont towns completely cut off from their neighbors. This storm also resulted in the closure of Route 2 in Western Massachusetts, a major east-west thoroughfare, causing significant transportation disruption, economic impact and inconvenience. Both the severity and frequency of severe storms are expected to increase through the end of the century across New England. Developing methods for identifying crossings more likely to fail/be damaged due to extreme events will enable the Massachusetts Department of Transportation (MassDOT) to more efficiently manage the vulnerability of transportation networks in Massachusetts.

Methods for assessing the vulnerability of road-stream crossings can leverage rapid assessments and metrics developed for other sectors. In particular, natural resource agencies and organizations have focused on stream crossings for a number of years in order to better understand the role of these crossings in disrupting river and stream networks and set priorities for restoring aquatic connectivity. Consideration of aquatic connectivity is important for two reasons: (1) stream crossing design standards for new roads are based on fish and other aquatic organism passage and stream continuity, and (2) the same data collected as part of the field assessments for these projects can be used to evaluate risk of structural failure as well as the barrier effects of the crossing. For example, a study of crossings in two watersheds, one in western Massachusetts and one in southern Vermont, that had experienced multiple culvert failures during Tropical Storm Irene determined that it was possible to use data collected for the assessment of aquatic organism passage to predict crossing failure due to storm damage (Jospe, 2013). Similarly, several methods for geomorphic assessment have been considered across the region to assess vulnerability of road crossings. Leveraging existing methodologies benefits the project by providing historical insight as well as multi-stakeholder collaborative opportunities for working towards resiliency.

The body of aligned work also provides insight on the characteristics of an effective, efficient screening protocol. Due to the large number of road-stream crossings in Massachusetts, it is impractical to conduct detailed engineering and hydrological studies at every road-stream crossing in order to assess its vulnerability to structural or storm-related failure. Therefore, it is desirable to create coarse screening procedures based on rapid field assessment of crossings and GIS analyses to identify those crossings most likely to fail. More intensive, follow-up assessments can subsequently be conducted at priority locations.

The purpose of The Deerfield River Vulnerability Pilot Project ("Deerfield Project") was to develop a credible rapid assessment and prioritization methodology. This assessment was informed by intensive methods for assessing risk of failure, yet simplified sufficiently to minimize cost without sacrificing reliability. Similarly, the Deerfield Project aimed to develop a methodology holistic enough to be of value to a wide-range of stakeholders beyond MassDOT, in part by piggybacking on the data collection protocols, network of cooperators, and programmatic infrastructure (online database, online training, electronic data collection, and distributed coordination) of projects such as the NAACC. Leveraging and aligning with other on-going projects across the region provides MassDOT with opportunities to collaboratively implement the protocol beyond the pilot area and helps identify potential partners with a shared goal of replacing or upgrading structures at priority crossings.

1.1.1 GOAL AND OBJECTIVES

The goal of the Deerfield Project was to develop an innovative systems-based approach to improve the assessment, prioritization, planning, protection and maintenance of roads and road-stream crossings that:

- Complements existing MassDOT project development and bridge design processes;

- Provides a decision-making tool that can be used at the state or local level during transportation project planning and development phases; and

- Engages other agencies, such as the Massachusetts Department of Environmental Protection (MassDEP), the U. S. Army Corps of Engineers (ACOE), the Federal Emergency Management Agency (FEMA), the Massachusetts Emergency Management Agency (MEMA), and municipal governments and Departments of Public Works, and familiarizes them with this approach.

A proactive approach for upgrading structures to account for climate change may be more cost effective than responding to the road and crossing failures which may occur due to inaction. An integrated approach - accounting for culvert condition, geomorphic vulnerability, hydraulic capacity, climate change, river and stream continuity (fish and other aquatic organism passage), and the potential disruption of critical services in the decision making process - will reduce uncertainties and improve prioritization schemes compared to vulnerability assessments that focus solely on any one aspect.

There were three broad objectives for the Deerfield Project:

1. Develop methodologies for assessing structural, geomorphic and hydraulic risk of failure for road-stream crossings and the associated disruption of emergency medical services

2. Use these new methodologies, along with existing methodologies for assessing ecological disruption, to assess road-stream crossings in the Massachusetts portion of the Deerfield River watershed

3. Incorporate climate change into hydraulic analyses using downscaled climate models to predict future flows and hydraulic risk of failure for road-stream crossings.

1.1.2 GENERAL APPROACH

The general approach used for this project was to build on field assessment methodologies developed by the Massachusetts River and Stream Continuity Partnership (a precursor to the NAACC) to create assessment methodologies to evaluate the risks of various mechanisms of crossing failure, the potential for crossing failures to disrupt emergency services, and the role of crossings in disrupting the aquatic connectivity and ecological integrity of rivers and streams. The Deerfield Project involved field data collection, compilation of GIS data, hydrological, climate and ecological

modeling, and the development of scoring systems for use in prioritizing crossings for more detailed evaluation or action.

Several linked components were included to meet the project goals, including:

- **Risk of Failure.** Assessment methodologies were developed to evaluate three mechanisms for crossing failure: structural failure (failure due to deficiencies in the structural integrity of crossings), hydraulic failure (inability to pass enough water to avoid structural damage, washouts, or water overtopping the road), and geomorphic failure (loss or substantial damage due to clogging with woody debris, sediment blockage, or scour). Results from these three assessments were combined to generate an overall risk of failure score for each crossing.

- **Climate Change and Associated Impacts on Hydraulic Risk.** A methodology was developed for identifying crossings that would be vulnerable to failure during storm events due to changes in precipitation and temperature patterns as the result of climate change.

- **Disruption of Emergency Medical Services (EMS).** Using network theory, an approach was developed for evaluating the impact of crossing failures on the delivery of emergency medical services (EMS).

- **Ecological Disruption.** A process was implemented for assessing crossings that serve as barriers to aquatic and wildlife movement, and sites where mitigation of barriers would do the most good for fish, other aquatic organisms, and wildlife populations were identified.

- **Transportation Vulnerability.** A scoring system was developed for transportation vulnerability that combined Disruption of EMS (criticality) and Overall Risk of Failure (risk).

- **Crossing Prioritization.** A scoring system for overall prioritization of crossings for attention was developed combining Transportation Vulnerability and Ecological Disruption.

1.1.3 SCORING AND PRIORITIZATION

Chapter 5 includes details on the models and methods used to assess the individual risks of failure based on structural, geomorphic, and hydraulic mechanisms. The chapter also describes how climate change associated impacts on hydraulic risk and ecological disruption was evaluated, and how disruption of continuity, stream temperatures, and brook trout habitat were accounted within the project. The scoring systems developed for each element of risk using a scale from zero (low risk) to one (high risk) are also described in Chapter 5. These three risk factors are combined to create an overall risk of failure score.

Criticality was used as a measure of how critical a failure would be, a combination of the magnitude of impact and adaptive capacity (ability to avoid or minimize impacts through alternatives or backup systems). The criticality impact assessed as part of this project was the level of delay in delivery of emergency medical

services. Risk and criticality scores were then combined to create Vulnerability scores for the Transportation domain.

In the ecological domain, Ecological Disruption scores provide a measure of ecosystem vulnerability. Ecological disruption is based on connectivity restoration potential for 1) all streams and 2) cold water streams. Combining the Vulnerability scores for the two domains (ecological and transportation) yielded an overall priority score for each crossing, as seen in Figure 1-1.

1.2   The Deerfield River Watershed

The site for the pilot study is the Deerfield River watershed, Figure 1-2, which straddles northwestern Massachusetts and southern Vermont and has a drainage-area of approximately 1720 km². It is a major subbasin of the Connecticut River. The largest tributary of the Deerfield River is the North River, with a total drainage area of approximately 230 km², Table 1-1. There is extensive hydroelectric-power generation (ten major dams) in the watershed. The watershed is mostly undeveloped, with only about 5.3% of the total area classified as developed and about 82% as forest according to the 2011 National Land Cover Dataset (Homer et al., 2015). There are three USGS stream gauges operated in the watershed on the North, South, and Green Rivers that are not impacted by the reservoir and dam operations. The catchments of these gauges represent about 23% of the total drainage area of the Deerfield River watershed, Table 1-1.

Elevations in the Deerfield River watershed range from about 35 meters above sea level in the Connecticut Valley Lowlands to about 1,202 meters in the ridges of the Berkshire Hills, with a mean altitude of about 475 meters (1,558 feet). Average annual precipitation in the watershed is 104 – 112 cm (41 – 44 inches) in the low altitudes and 127 – 188 cm (50 – 74 inches) in the higher altitudes (PRISM Climate Group, 2004; Knox and Nordenson, 1955). Snowmelt in spring and evapotranspiration in summer and fall cause annual cyclical trends in mean monthly runoff, even though mean monthly precipitation is evenly distributed throughout the year (Gay et al., 1974).

There are approximately 1,002 road-stream crossings in the Deerfield River watershed. The watershed experienced multiple culvert failures, bridge damage, and associated transportation disruptions as a result of Tropical Storm Irene. Observed damage/ failure was attributed to a range of factors, including bank erosion, undermining due to stream erosion, and landslides. With its history of documented damages, lack of development, and range of elevation and precipitation regimes, the Deerfield River watershed provided a unique location for the pilot study.

1.3   Benefits to/from Other Projects

The Deerfield Project is aligned with existing MassDOT programs and complements other concurrent studies/policy changes.
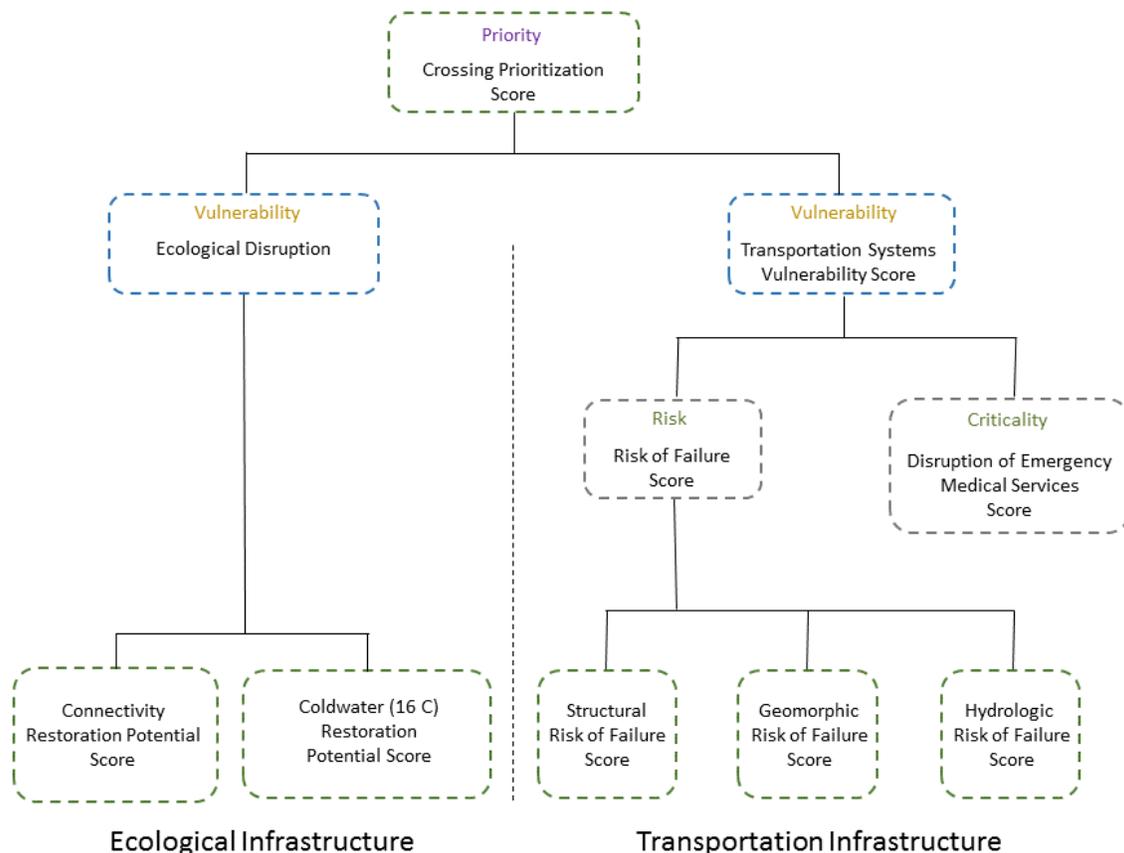
Figure 1-1. Scoring scheme for the Deerfield Project.
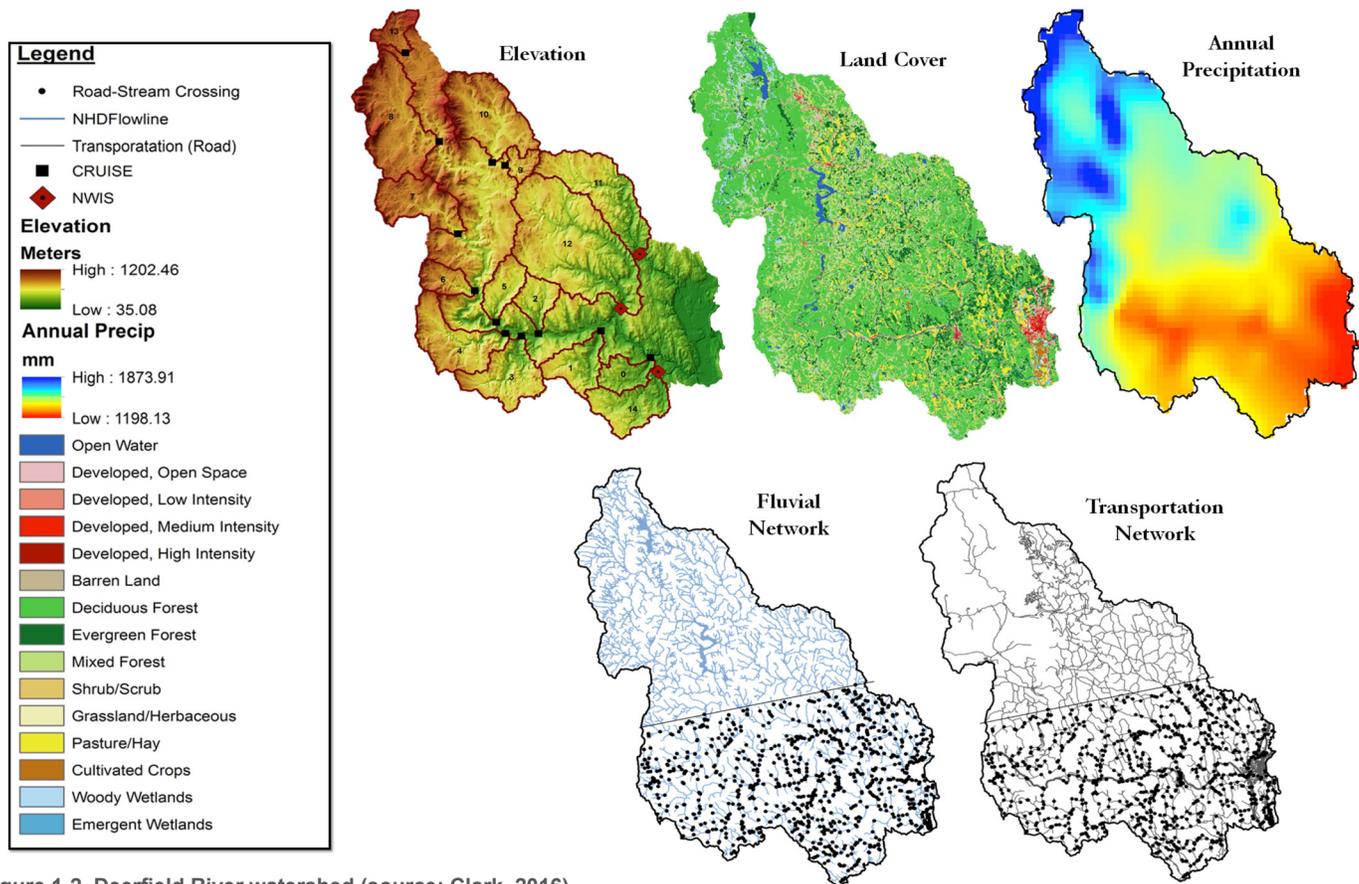
Figure 1-2. Deerfield River watershed (source: Clark, 2016).

MassDOT already has an effective statewide Bridge Inspection program that provides rigorous hands-on bridge structural and site safety inspection coverage for almost 9,000 bridges and culverts[1]. The intent of this project was not to replace the existing statewide inspection program, but rather to collect additional data, in particular for closed bottomed structures less than 10 feet in length. The project team deferred to the existing MassDOT overall condition categorization of poor, critical, or not at risk when available. The Deerfield Project also aligns with the MassDOT-Parsons Brinkerhoff Climate Change Vulnerability project, which will provide vulnerability estimates for road-stream crossings across the state. Additionally, the Deerfield Project incorporates a much deeper look at variability and uncertainty due to choice of hydrologic model and future climate scenario than that provided by current practice. The results of this pilot study provide context to potential error bounds inherent in studies that incorporate a more limited range of models, climate scenarios, and factors that influence vulnerability.

The Deerfield Project is synergistic with three other existing or recently completed projects, which also focused on the watershed. RiverSmart Communities, funded by the UMass Center for Food, Agriculture and the Environment as well as the U.S. ACOE Institute for Water Resources (IWR), aims to achieve ecologically restorative flood prevention and remediation based on fluvial geomorphological science and collaborations from local municipalities to federal agencies. Farms, Floods and Fluvial Geomorphology, funded by the U.S. Department of Agriculture National Institute of Food and Agriculture, was an aligned project providing factsheets and strategies for land management to maximize overall watershed/river health. A third initiative, a series of fluvial geomorphological assessments overseen by State Geologist Steve Mabee, was focused on reducing flood and erosion hazards in the Deerfield River watershed. These projects mutually benefited by sharing of progress, data, modeling results, and work with stakeholders.

Recent projects in New York and New Hampshire also aim to identify culverts at risk of hydraulic failure under current and future climate conditions. These include Cornell University's *Determining Peak Flow Under Different Scenarios and Assessing Organism Passage Potential: Identifying and Prioritizing Undersized and Poorly Passable Culverts* and Trout Unlimited's web-based tool developed for the Piscataquog River watershed.

The NAACC is a program for development and deployment of road-stream crossing assessments in the 13-state North Atlantic region of the United States. Rapid assessment tools and landscape models have been developed and deployed to evaluate the barrier effects of culverts and bridges and the potential for restoring aquatic connectivity via culvert replacement. The Deerfield Project

---

1 1,590 with 10 foot – 20 foot span, and 5,230 with a 20 foot or greater span

Table 1-1. Selected catchment characteristics in the Deerfield River watershed and at the three unimpaired USGS streamflow gauges (Source: Clark, 2016).

Catchment Property	Deerfield Watershed	01170100 Green River	01169000 North River	01169900 South River
Drainage Area (km^2)	1718.2	107.7	231.2	62.8
Mean Annual Precipitation (mm)[a]	1374.8	1384.0	1378.5	1289.1
Mean Temperature (°C)[a]	6.3	6.6	6.6	7.3
Max Temperature (°C)[a]	12.1	12.4	12.4	13.2
Mean Elevation (m)[b]	475.1	413.5	430.8	343.2
Mean Slope (deg)[b]	9.0	9.8	8.6	8.8
North Facing (%)[b]	8.8	7.9	9.3	12.3
East Facing (%)[b]	17.3	16.9	17.6	17.9
Developed (%)[c]	5.3	3.0	4.4	6.8
Forest (%)[c]	82.0	90.3	84.0	78.6
Agriculture (%)[c]	5.9	3.8	7.8	10.0
Hydrological Group B (%)[d]	23.3	20.8	22.1	16.3
Hydrological Group B (%)[d]	1.9	0.7	0.7	0.8
Hydrological Group B (%)[d]	5.4	1.3	10.1	9.5
Stream Density (km/km^2)[e]	1.48	1.7	1.4	1.31

Notes: [a] PRISM; [b] National Elevation Dataset; [c] NLCD (2011); [d] NRCS SSURGO Dataset; [e] NHD High Resolution Dataset

benefitted from several aspects of the NAACC, including the xy coding system for identifying road-stream crossings, an aquatic passability assessment methodology, culvert condition assessment methodology, and the NAACC online database used to store and score data from these two assessment modules. There are also ways in which the NAACC is expected to benefit from the Deerfield Project. There is interest throughout the 13-state North Atlantic region in developing rapid field assessment and GIS approaches that can be used to evaluate the risk of crossing failure due to structural deficiency or storm damage and flooding. Experience gained from the Deerfield Project will be valuable for the development of such methodologies.

A related project, funded through Department of the Interior (DOI) Hurricane Sandy Mitigation Funds, considers multiple aspects of road-stream crossing assessment for climate resilience and aquatic connectivity in the Sandy-Impacted Northeastern US. It was conducted by a project team from the University of Massachusetts Amherst, with subcontractors including The Nature Conservancy. The project included three components: Passage Assessment (aquatic organism passage through road-stream crossings via the North Atlantic Aquatic Connectivity Collaborative), Hydrologic Risk Assessment (modeling exercise to estimate peak streamflows associated with extreme events in twenty Sandy-impacted watersheds), and Ecological Benefits Synthesis (calculating metrics for ecological considerations such as cold water habitats, diadromous fish). It included development of the Stream

Crossings Explorer Tool (in conjunction with the Deerfield Project), which will be expanded to all North Atlantic states for data that are available.

The Franklin Regional Council of Governments (FRCOG) has funded two river-corridor mapping projects, both conducted by John Fields of Fields Geology Services. The first focused on the South River watershed within the Deerfield River watershed and followed protocols developed by the state of Vermont. More recently, Fields is developing an alternative approach for the North River watershed, distinct from both the Vermont fluvial geomorphic and Milone & MacBroom Associates (MMI) stream power based protocols. The approach under development instead visually defines the river corridor based on LIDAR and aerial photographs, supplements these data with field work to look for traces of old stream channels and identify the height of terraces above the current floodplain, and then couples the desktop observations and data to determine where the river could meander in a reasonable future timeframe. The North River project aims to identify river areas most susceptible to erosion at a parcel-level scale through a cost-effective tool that can be used by towns, state and federal agencies.

The Federal Highway Association (FHWA) recently funded a MassDOT project to bring flood resiliency into asset management, building on the Deerfield Project. Specifically, MMI will provide stream power and bankfull width estimates for road-stream crossings across the state. They will utilize this information and existing structure width data to 1) establish short-term statewide

resiliency tools to predict the vulnerability of structures and roads, 2) update MassDOT data collection methods and forms to synchronize evaluations of bridges and culverts, and bring in resiliency data for current and future use, and 3) create long-term resiliency tools that build on steps 1 and 2.

1.4 Report Organization

The remainder of this report summarizes the Deerfield Project as follows. Chapter 2 provides an overview of current FEMA reimbursement policies, as well as federal and state efforts to address vulnerability and resiliency of transportation infrastructure. Chapter 3 describes the road-stream crossings evaluated, and provides a summary of the existing data utilized to support the project. The field and geomorphic data collected by the project team are also described in this chapter. Chapter 4 summarizes the climate data utilized for the project, detailing the range of predictions and comparing project data against data utilized in other DOT projects. Chapter 5 details the technical approach for evaluating structural, geomorphic, and hydraulic risk, as well as determining road-stream crossing criticality and ecological disruption. Chapter 6 describes the prioritization of road-stream crossings by combining ecological disruption and transportation vulnerability to "score" each crossing. Deerfield Project deliverables are briefly described in Chapter 7. Chapter 8 provides a detailed discussion summarizing the key findings, limitations on use of the data, lessons learned, and recommendations. Chapter 9 provides a listing of references from each chapter. Additional technical details are provided in the appendices.

Policies and Regulatory Programs Related to Streams

This chapter provides a brief overview of current policies and programs. The first section summarizes FEMA replacement policies, examples of alternative solutions, and suggested changes. The second section summarizes on-going efforts at the state level to address climate change adaptation plans of the Commonwealth.

2.1 Federal Emergency Management Agency Policies

FEMA aid is triggered by federal declaration at the request of the state. FEMA typically approves replacement of "in-kind" crossings after a disaster, meaning the replacement of a failed bridge or culvert with an identical structure. However, as part of mitigation measures post-disaster recovery[2], it is within FEMA's discretion to provide funding for the upgrade or relocation of a structure to reduce the risk of future damage under certain conditions. For example, FEMA will consider hazard mitigation measures that aim to prevent the same damage from occurring again, particularly at sites where damages have occurred repeatedly[3]. FEMA allows funding of upgrades after a disaster if the upgrade is fully consistent with a code or standard that requires it (Gillespie et al., 2014).

Logistical constraints and analytical complexity associated with crossing upgrades are factors that contribute to FEMA funding being utilized for replacement "in-kind." FEMA requires that communities take fish habitat, wetlands, and other downstream impacts into consideration during both the design and construction of upgraded culverts. FEMA's Environmental and Historic Preservation (EHP) specialists review projects to ensure that they meet all relevant environmental laws and regulations, such as the Endangered Species Act, the National Historic Preservation Act, and the Clean Water Act. It can be difficult for towns and communities to meet FEMA's design and feasibility documentation requirements. On top of this, FEMA requires that no physical actions occur before they complete environmental compliance review and identify historic structures and archaeological resources. Grant funds are not available until 6 months have passed since a disaster, and it can take one year before the amount of money available is known. In post-disaster situations, many communities resort to replacement in-kind,

resulting in crossings remaining vulnerable and increasing long-term and downstream hazards[4].

There are examples where communities have successfully worked with FEMA to fund upgrades. In New Hampshire, the state received FEMA reimbursement for the repair and upgrade of 18 culverts (17 were repaired with an upgrade between 8/30/2010 and 5/23/2014 and one more is underway), all paid for with FEMA Hazard Mitigation Assistance Program funding[5]. In the town of Townshend, VT, Tropical Storm Irene blew out culverts that were subsequently replaced with upgraded structures with FEMA funding. In Buckland, MA, the town was also successful in working with FEMA to approve reimbursement for the upgrade of several culverts versus replace-in- kind after Tropical Storm Irene[6].

There are several common themes to these success stories:

- Existence of statewide policies/regulations that can be uniformly and universally implemented,

- Assessments that identify crossings as vulnerable as currently designed and which would require an upgrade to no longer be considered vulnerable to failure,

- Standard designs designated for at risk sites based on analysis of what would be necessary to make these sites no longer at risk of failure, and

- An economic analysis of the benefit-costs for an upgrade.

In the case of New Hampshire, the state has been proactively working with communities to develop improved culvert designs that include a benefit-cost analysis for the installation of flood resilient stream crossings. In the case of Vermont, FEMA at first refused to pay for the upgrades, but agreed to do so after the state appealed the decision and revised its culvert standards to make them enforceable.

The Nature Conservancy (TNC) and American Rivers met with FEMA in Washington to discuss these issues, and are committed to working with FEMA and Regional Planning Agencies/towns to include by default consideration of ecosystem services, such as incorporating aquatic organism passage concerns, and cost avoidance valuations as part of the crossing replacement and upgrade decision process. It is clear that cost-effectiveness will be a key factor. FEMA has published a benefit-cost analysis methodology for re-engineering (FEMA, 2009) and TNC has published an economic analysis discussing the cost savings that communities can

---

2 *Refer to FEMA's Environmental and Historic Preservation (EHP) Fact Sheet*
3 https://www.fema.gov/public-assistance-frequently-asked-questions
4 *Refer to Vogel et al. (2016) "Supporting New England Communities to Become River-Smart" for further discussion*
5 *Personal communication, 6/27/17 email from Shane Csiki, NH DES*
6 *Personal communication with Ms. Cheryl Dukes, then chair of the Buckland Select Board*

expect if they build culverts the right size and angle (Levine, 2013). While FEMA's 2014 – 2018 Strategic Plan resolves to "build a solid foundation for working smarter through analytics by establishing standards, policies, governance structures, tools, training, and an authoritative inventory of data and source systems" (strategy 5.2.1), there is currently no regulatory change in place at FEMA to replace its preference for "in-kind emergency repair to upgrading road-crossing structures based on flood resiliency" (Gillespie, 2014).

There is still work to be done at both the state and federal levels to facilitate increased flood resiliency of the transportation network by simplifying the design, permitting, and funding processes to replace vulnerable and damaged crossings quickly and efficiently (Vogel et al., 2016; Gillespie, 2014).

2.2   Highlights of Federal and State Efforts

2.2.1   ALIGNMENT WITH COMMONWEALTH CLIMATE CHANGE ADAPTATION GOALS

In their 2011 Climate Change Adaptation Report, the Massachusetts Executive Office of Energy and Environmental Affairs (EEA) and the Adaptation Advisory Committee stated that the frequency of extreme weather events in Massachusetts is expected to increase, resulting in an increased risk of damage to transportation systems. Several general strategies to prepare for increased extreme weather events were proposed, including:

•   Bolstering of ongoing efforts by state agencies to factor climate change into future design, permitting, and building requirements,

•   Formulation of risk-based methods to evaluate the service life of infrastructure assets against adverse climate change,

•   Identifying, developing and implementing solutions such as reconstruction, removal, or relocation of vulnerable infrastructure, and

•   Consideration of future storm event characteristics in the sizing of stormwater management structures (e.g., pipes, culverts, outfalls) by highway agencies in order to balance the upfront costs of incrementally installing larger structures today against the future costs of replacing an entire drainage system.

The Deerfield Project helped to fulfill these goals by laying the groundwork to assess the vulnerability of infrastructure at stream crossings and creating a tool to prioritize replacement or upgrade of such infrastructure on a watershed basis.

This project also fulfills section 3.d of Governor Charlie Baker's Executive Order No. 569 (Establishing an Integrated Climate Change Strategy for the Commonwealth) to "provide technical assistance to Cities and Towns to complete vulnerability assessments, identify adaptation strategies, and begin implementation of these strategies."

2.2.2   ALIGNMENT WITH OTHER COMMONWEALTH EFFORTS

Other ongoing or completed MassDOT projects to fulfill Climate Change Adaptation goals in Massachusetts include:

•   The Central Artery/Tunnel (CA/T) Vulnerability and Adaptation Assessment Pilot Project (completed in 2015) created the Boston Harbor Flood Risk Model to identify the risk of flooding in Boston under current and future storms and sea level rises.

•   The Coastal Transportation Vulnerability Assessment (underway) is creating the Massachusetts Coastal Flood Risk Model (MC-FRM), refining the Boston Harbor Flood Risk Model to extend it to the entire coastline of Massachusetts in order to assess transportation infrastructure vulnerable to storm surges and sea level rises. The MC-FRM will predict tidal epochs for 2030, 2050, 2070 and potentially also for 2100.

•   The MassDOT statewide transportation asset climate change vulnerability assessment study aims to provide a prioritized list of MassDOT's critical assets that are most likely to be at risk to riverine flooding under future extreme precipitation events as projected by climate models. A prototype methodology for mapping future 100-year flood plains and flood depths at HU-8 watershed level has been developed and tested in a sub-watershed. A methodology for assessing stream channel stability (erosion and aggradation potential) under extreme events is currently under development. MassDOT plans to implement both methodologies statewide, and will map out future 100-year flood plains and stream channel stability and quantify critical assets' potential exposure to such extreme flood events.

•   The Bringing Flood Resiliency in MassDOT Asset Management: Stream Power and Structure Bankfull Width Estimation for the Commonwealth of Massachusetts project is explicitly incorporating flood resiliency into MassDOT assets management by establishing short-term statewide resiliency tools to predict vulnerability of structures and roads and to update data collection methods to synchronize evaluations of bridges and culverts. The project builds directly from the geomorphic stream power work conducted for the Deerfield Project.

•   The MassDOT Rivers and Roads Training Program is a collaboration between MassDOT, Milone & MacBroom, the Nature Conservancy, and VTDEC Rivers to use fluvial geomorphology (FGM) to reduce conflicts between transportation assets and rivers. There are three tiers.

    ◦   Tier one is an online course to introduce the fundamentals of FGM and raise awareness (https://anrweb.vt.gov/DEC/RoadsTraining/Default.aspx).

- Tier two is a one-day course for advanced understanding of FGM.

- Tier three consists of one-day subject matter courses to enable project implementation.

More information is available through the Baystate Roads One Center which coordinates the trainings.

- The MassDOT Mapping Our Vulnerable Infrastructure Tool (MOVIT) is a web-based application that will capture geo-spatial data on existing vulnerabilities based on agency staff institutional knowledge to help minimize impacts of extreme weather events on the transportation system. The data collection will be on an on-going basis as well as through exit interviews with staff prior to them leaving the agency.

- The MassDOT Project Intake Tool (MaPIT), completed in Fall 2017, has modernized the Highway Division project planning process through an online project submittal form with automated GIS analysis against transportation, safety, environmental and vulnerability data.

2.2.3 ALIGNMENT WITH CURRENT MASSDOT ENGINEERING PRACTICE

The most definitive discussion of hydrologic computational methods MassDOT currently endorses for bridge and culvert design is presented in Section 1.3 of the 2013 revision of the MassDOT Bridge Load and Resistance Factor Design (LRFD) Manual. The LRFD states that "To the extent practicable, the design of new and replacement bridge waterway openings shall conform to applicable sections of 2011 Massachusetts Stream Crossing Standards" (page 8, MassDOT Bridge LRFD Manual). It should be stressed that bridge and culvert design were NOT elements of this pilot program. Rather, the Deerfield Project aimed to provide DOT with a vulnerability assessment, which they may choose to utilize when prioritizing structures for replacement or upgrade. Once a structure is identified for replacement or upgrade, it is expected that MassDOT will implement the approved computational methods for bridge and culvert design. These accepted computational methodologies were utilized as the basis of the vulnerability assessment. The five key methodologies are:

1. At crossings with relatively unregulated upstream watersheds, Section 1.3 recommends use of either the existing USGS Massachusetts Regionalized Peak Flow Equation (RPFE) system or NRCS TR-55 procedures. At crossings with upstream watersheds regulated by built natural impoundments or diversions of runoff flow, Section 1.3 recommends the use of the ACOE HEC-HMS Hydrologic Modeling System.

2. Use of the USGS StreamStats for Massachusetts web application and/or the resources of the Massachusetts Department of Geographic Information Systems (MassGIS) is recommended to support development of watershed design variables for all the computational methods described above.

3. Utilization of the National Oceanographic and Atmospheric Administration (NOAA) Hydrometeorological Design Studies Center (HDSC) Hydrologic Atlas 14 web application is recommended for extreme precipitation frequency and duration data[7].

4. Designers are directed to review the results generated by several different computational methods, apply professional engineering judgment to identify the output set that best reflect local and regional hydrologic conditions, and then document the rationale for that selection in the project's hydraulic study report.

5. Currently effective National Flood Insurance Program (NFIP) 10-, 50-, base (100-year), and 500-year flood discharges must be employed within hydraulic studies performed for proposed replacement bridges that cross waterways with either existing NFIP Regulatory Floodway delineations or published flood elevation profiles.

2.2.4 UMASS RIVERSMART COMMUNITIES PROJECT

A simultaneous project at UMass, "RiverSmart Communities and Federal Collaborators," investigated ways for New England's communities, and the governments that help them, to better deal with and adjust to river floods. As part of this project, meetings were held with town officials and community members across the Deerfield River watershed to gain an understanding of their concerns and challenges in terms of flood preparation and mitigation. A consistent problem mentioned was that after Tropical Storm Irene, towns were forced to replace culverts with functional ones that were the same size as the ones that had failed.

In their report "Supporting New England Communities to Become River-Smart" (Vogel et al, 2016), the authors developed five policy recommendations to support communities' future safety and well-being. Recommendation Number Two is to "Upgrade Vulnerable Stream Crossing Infrastructure." More detailed elements of this recommendation are:

1. *Improve stream crossing regulatory standards to support upgrades, be consistent across agencies, and allow site-specific flexibility.* In Massachusetts, stream crossing standards have been updated in collaboration with the ACOE to account for aquatic organism passage to the extent practicable, providing some flexibility in application on a site-by-site basis.

2. *Streamline permit and funding processes and requirements, and incentivize replacing vulnerable and damaged crossings with upgrades.*

---

7 *At the time of project development, MassDOT endorsed use of the regional extreme precipitation frequency data set developed by the NRCS-funded Northeast Regional Climate Center (NRCC) as an interim measure pending availability of Atlas 14, which is now available.*

3. *Develop and make available easy-to-follow design templates and guidelines for upgraded crossings which will receive quick permitting and funding review and high likelihood of approval.*

4. *Develop and support an accessible inventory and database of stream crossings that identifies vulnerable crossings.*
    In New England, the NAACC provides such a database (naacc.org). The Deerfield Project leveraged the NAACC database to store crossing survey data and additionally developed a crossing condition assessment that was added to the database. Furthermore, a major deliverable of the Deerfield Project is the Deerfield Watershed Stream Crossing Explorer (SCE), which provides a GIS based tool to locate road-stream crossings of interest based on any of the risk and vulnerability metrics developed for the project, as described in the subsequent sections of this report (sce.ecosheds.org).

5. *Increase and diversify funding for stream crossing upgrades.*

Data Overview

3.1 Road-Stream Crossings in the Deerfield River Watershed

The original project estimate of 842 road-stream crossings in the Deerfield River watershed (DRW) was based on a stream network considered for purposes of aquatic passage. To ensure that no crossing structures were missed, in November-December 2014 the Massachusetts Geological Survey revised estimates of road-stream crossings in the watershed by consulting additional sources of data. GIS-based intersections between the MassDOT roads layer and the 1:25,000 USGS hydrology layer (or NALLC-generated streamlines layer) did a better job predicting a full list (994) of road-stream crossings. However, some of these are errant intersections of arcs on a map instead of actual crossings; alternatively, actual crossings were excluded in some of the uppermost reaches of streams where ephemeral streams were not mapped to intersect with a road. In addition, the NAACC Stream Continuity Database was consulted. This database was considered to be highly accurate because at that time 80-90% of the watershed's identified road-stream crossings had been field checked. In November-December 2014, the NAACC database included 1,004 crossings in the Deerfield River watershed. Field visits identified an additional 13 road-stream crossings for a total of 1017 that were considered during the Deerfield Project (Figure 3-1). Each of these crossings was assigned an "xycode" by NAACC. The xycode consists of the latitude and longitude of the geographical location of the crossing, preceded by the letters xy. For example, a crossing located at latitude 42.123456 and longitude -73.987654 has an xycode of xy4212345673987654. The DRW Pilot adopted the NAACC xycode as the primary identifier of road-stream crossings for the project.

3.2 Existing Data

3.2.1 DISTRICT AND TOWN DATA

The project team worked with MassDOT Headquarters and Districts 1 and 2 to obtain road-stream crossing data of relevance to the project. These data include original culvert and bridge crossing design specifications, inspection reports, and information on reported failures and overtoppings. Data compiled by the USGS (Parker, 1997) that assessed stream stability at 2,361 MassDOT crossings were reviewed and incorporated into the project databases as applicable. The Parker (1997) study yielded geomorphic information (scour versus aggradation) for 140 of the watershed's

bridges. In addition, UMass requested copies of the input and output files for the National Flood Insurance Program (NFIP) HEC-2 hydraulic models for the waterways that were studied in detail as part of the NFIP studies performed for communities located within the Deerfield River watershed boundaries. The USGS flood-inundation maps for the Deerfield River (USGS SIR 2015-5104) include data for the top width of the Deerfield River floodplain and updated hydrologic and hydraulic analyses, but were published after the start of the project. Due to timing of the analyses, the project was unable to leverage these data. However, the USGS flood-inundation mapping study was based on Laser Imaging Detection and Ranging (LIDAR) data collected in 2012. These same data were utilized by MMI for determining slope for the Deerfield Project. Data for dams and non-road crossings were collected as part of this task to support the modeling effort.

3.2.2 BRIDGE INSPECTION DATA

Structure inspection field reports were obtained from MassDOT for all bridges in the Massachusetts portion of the Deerfield River watershed. These inspection reports consist of the following information:

- Location and structure number
- Date of inspection and name of inspector
- Structure type, owner, maintainer
- Structure deck details
- Structure superstructure details
- Structure substructure details
- Channel & channel protection details
- Traffic safety details
- Accessibility details
- Detailed remarks.

3.2.3 LOCAL MUNICIPALITY DATA COLLECTION

Residents and officials of the relatively remote Deerfield River watershed towns maintain a close relationship with the natural world, with special interest in its intersection with the built environment. For example, in our experience surveying bridges, curious abutters would commonly offer, without prompting, the same critical observation—that Hurricane Irene's waters had filled the space under their bridge.
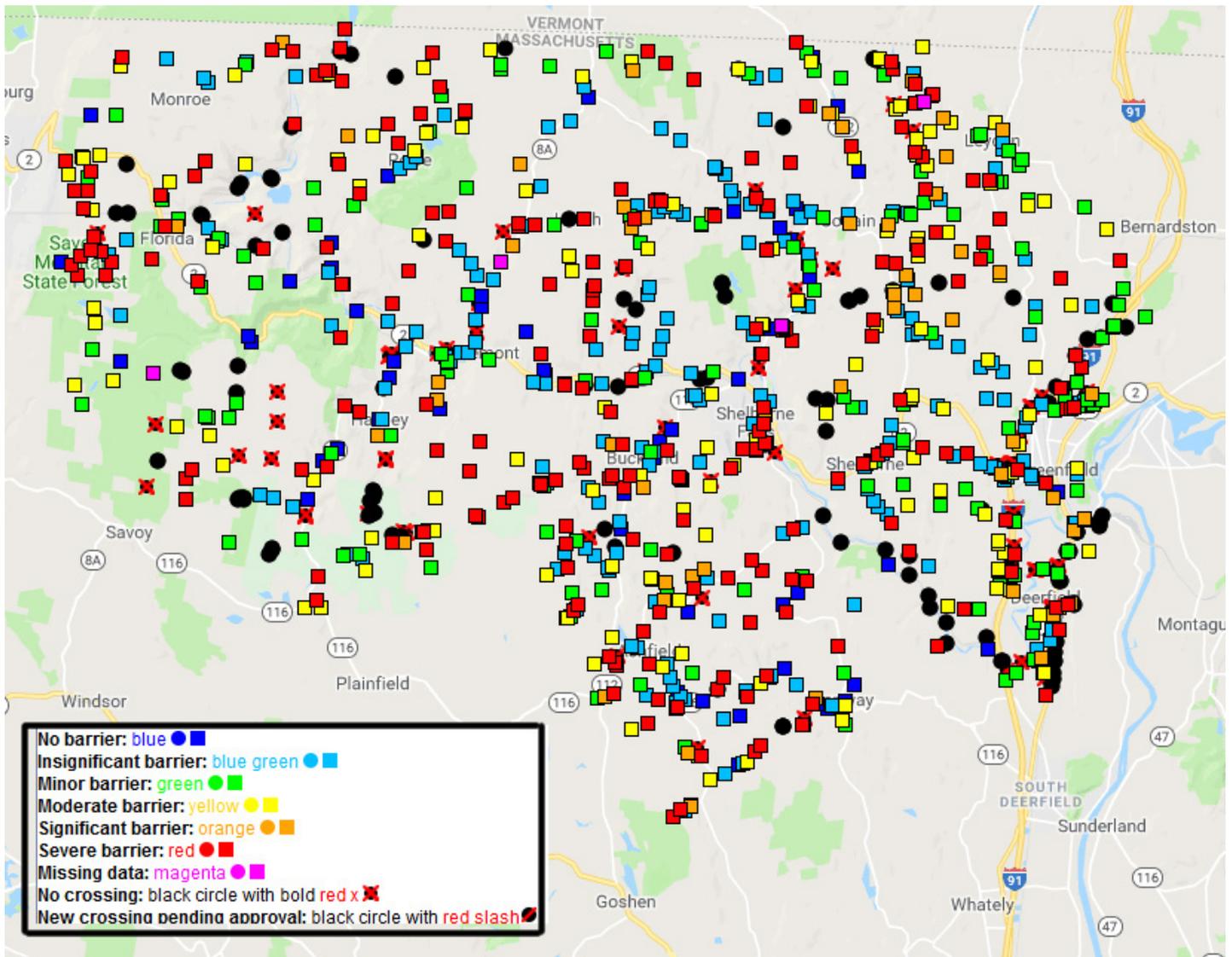
Figure 3-1: Map of crossings in the Deerfield River watershed in the NAACC Database Map Viewer. Squares and dot show the location of stream crossings.

In order to include this rich observational record, we asked municipal highway department officials about the history of flooding and related failures in the transportation network in a February 2015 letter to 17 municipalities (Appendix A). The letter briefly introduced our project goals, provided contact information for team members for more information, and asked a series of specific questions regarding the town's road and stream networks. We also enclosed a color topographic map for each town, 11" x 17", with all of the town's known (Stream Continuity Database) road-stream crossings plotted and numbered, and including labeled road and stream networks (see Appendix A). Town boundaries, as well as a scale and graticule[8] overlay, uniquely described each crossing in space. In companion tables, each road-stream crossing and its identifying properties as described above were also listed.

Maps and tables were prepared for each town identifying the location of known road-stream crossings. Based on these maps, the town officials were asked to provide:

1. Information about road-stream crossing failures by

   A. Circling known failures of road-stream crossings within the past ten years,

   B. Marking whether these had occurred through overtopping, embankment failure, blockage by debris, structural failure, washout, or another process, and

   C. Marking locations characterized by multiple failures.

2. Information about flooding at road-stream crossings by

   A. Marking locations of roadways flooded in the past ten years,

---

8 *Grid used in a geographic coordinate system*

B. Marking locations characterized by multiple flood events.

3. Information about locations of known side slope failures and/or roadway washout by

A. Circling locations of known side slope failures and/or roadway washouts within the past ten years, and

B. Marking locations of repeated failure.

4. Information about on-going or planned road-crossing related work by

A. Noting the phase of any road-crossing related work in preparation or progress and its timescale for completion, and

B. Noting priority road-stream or roadway restoration sites at these crossings and cause for concern at these sites.

5. Any road-stream crossings not included on the map, and

6. Additional information regarding areas of concern and failures.

Responses were received from eight towns (Ashfield, Bernardston, Florida, Heath, Leyden, Rowe, and Savoy), with information on 387 crossings, summarized in Table 3-1:

Table 3-1: Deerfield River watershed towns survey results.

Crossing Issue	Number of Crossings Affected
Overtopping	22
Embankment failure	9
Structural failure	9
Blocked by debris	12
Roadway flooding	4
Fluvial erosion	8
Repeated failures	13
Washed out	6

3.3 Stream Crossing Data Collected by Trout Unlimited and UMass

3.3.1 FIELD ASSESSMENT AND CONDITION SURVEY

Using the procedures developed by NAACC, we coordinated a comprehensive aquatic passability assessment of most state and local stream crossing structures in the Massachusetts portion of the Deerfield River watershed. Although state highways were the primary focus of this pilot project, an assessment of local stream crossing structures was necessary to effectively model aquatic connectivity and accurately predict restoration potential for upgrading crossing structures on state highways. These data were also critical for identifying channel constrictions that impact stream velocity and the associated upstream-downstream impacts of such constrictions.

In aggregate, field efforts in support of these aspects along with geomorphic context and hydraulic capacity are perhaps the study's most ambitious element. Over 1,000 individual structures, each unique, were thoroughly and systematically surveyed over a two-year period. Basic navigation through the watershed and confident identification of road-stream crossings were complicated by the rugged terrain and damage to the transportation network caused by Tropical Storm Irene. Many roads were unimproved, unpaved, and/or unlabeled; steep valley walls obstructed GPS-satellite communications and darkened photographs; and thick vegetation during summer months obstructed sight lines and inhibited movement of the surveyors. Culverts are all unlabeled, and larger rivers and highways pose their own obvious hazards and difficulties. In anticipation of these challenges, field surveys of culverts began shortly after the project start date, but before grants and contracts were awarded and before an exhaustive inventory of road-stream crossings in the study area could be assembled. As familiarity with the watershed's infrastructure increased, however, field maps and routes were improved, and survey crews were able to work with greater efficiency.

3.3.2 GENERAL CROSSING DATA

We contracted with Trout Unlimited (TU), which used trained technicians to conduct the field assessments of road-stream crossings. Assessment efforts focused on culverts, but bridge crossings were also evaluated when they could safely be accessed. Only fieldwork in the Massachusetts portions of the river was conducted as part of the Deerfield Project. Detailed crossing data were not needed for the Vermont portions of the watershed because site-specific vulnerability was not assessed in Vermont. Instead, data for the Vermont portions of the watershed necessary to support watershed-wide geomorphic and hydraulic analyses were inferred from GIS data. Field assessment of non-road crossings like railroad bridges and dams was not included in the scope of this project. Data necessary to support geomorphic and hydraulic calculations for the Vermont portion of the watershed were collected via GIS "desk top collection."

The field assessment included collection of aquatic passability, structure condition, some geomorphic data, and data necessary for hydraulic calculations. At each road-stream crossing, detailed information was collected for identification purposes. In addition to routine information (e.g., names and roles of the surveyors, the date of the survey, and the weather) road-stream crossings were described by their watercourse and road, their town, and unique identification numbers. These unique identification numbers were either assigned by the Conservation Assessment and Prioritization System (CAPS) or the Stream Continuity Database hosted by the NAACC (as described in Section 3.1 above). GPS coordinates were measured at each site, and the type of road and water level were noted. Photographs of each bridge/culvert from upstream (inlet) and from downstream (outlet) were taken from the stream bed, and from the crossing facing upstream and/or downstream in some cases. Bridges were sketched in plan view. MassDOT bridge inspection

data for bridge crossings were used preferentially over data collected by TU.

To evaluate the accuracy of data collected by TU technicians, MMI was contracted to have a licensed professional engineer (PE) collect identical data at a random 10% of the road-stream crossings. Section B4 in Appendix B is a Quality Control Chapter that describes qualitatively the comparison between MMI collected field data and TU/UMass collected data on a subset of crossings in order to inform the methodologies suggested for potential future implementation in other watersheds. For structures that were field inspected by MMI staff, MMI prepared National Bridge Inspection Standards (NBIS) forms for items 62 (culverts) or 113 (bridges), plus item 61 (channel protection) and item 71 (waterway adequacy).

The aquatic passability, culvert condition, and geomorphic data are stored in the NAACC database at naacc.org. Note that to view the data, one must chose the UMass Stream Continuity Project (2005-2017) Data Set, and that the culvert condition data are not visible to those who have not been granted access to those data.

3.3.3 AQUATIC PASSABILITY DATA

Field data in support of aquatic passability were collected by TU field teams as quantitative and visual observations, as well as inferences of conditions during low flow (Appendix A, Field Data Sheets). The aquatic passability protocol used in the study was developed by UMass for the project and has subsequently been revised and adopted by the NAACC. Evidence and observations of fish were noted on the form. Each crossing was examined for support of fish by answering questions such as:

- Was the stream flowing?

- Did the crossing constrict the channel?

- If so, how severely?

- What type of crossing was it and what general condition was it in?

- Was it aligned to the stream?

- Did it have a scour pool at the outlet?

Each cell of every bridge and culvert was examined in detail for aquatic passability. Physical barriers to fish and wildlife passage were noted, as well as the extent to which the structure offered dry passage to terrestrial wildlife by answering questions such as:

- Was the structure embedded into some alluvial substrate?

- How did any substrate within the structure compare to the adjacent stream bottom?

- Did the structure contain internal features such as weirs, baffles, or supports?

- Was there a clear line of sight through the structure?

The broader context of the wildlife corridor that the structure served was also described. The team looked upstream and

downstream of the structure for riparian buffers and signs of wildlife and noted the type of vegetation on the banks. They looked along the road for wildlife killed by motor vehicle traffic.

3.3.4 CULVERT CONDITION DATA

The TU survey teams also collected field data to support an assessment of culvert condition or integrity (Appendix A, Field Data Sheets). Consideration of condition of bridges and railroad crossings was beyond the scope of this study, as bridges are already inspected thoroughly by MassDOT and railroad crossings are privately owned. Culvert condition was assessed using a new rapid condition assessment protocol for culverts based on a Federal Highway Administration (FHWA) publication (Hunt et al., 2010). It has since been adopted by NAACC in a slightly modified version. Structures were photo-documented at the same time crossings were assessed for aquatic passability. One page of the form was used to describe the inlet; the other described the outlet. Items considered with respect to each end of the culvert complement each other rather than duplicating exactly, though most of the same areas of focus appear for both inlet and outlet condition description. Observations were grouped as acceptable, poor (having suffered some damage or deterioration), or critical (failure of structure likely or imminent).

At the culvert inlet, basic properties were noted, including material and entrance type. An inlet was classified as "poor" or "critical" if any of the following were observed:

- Invert deterioration, characterized as consisting of holes, abrasion and scour or scaling, or displaced masonry blocks.

- Poor joints and seams, which might allow soil or water to pass out of the culvert and result in accompanying damage to the roadway or embankment.

- Cross-sectional deformation, which might be accompanied by longitudinal cracking (in concrete), and become critical when flow is interrupted by blockage, soil infiltration or voids.

- Deficiencies in structural integrity of the barrel, which might be marked by cracks in concrete, displaced masonry blocks, multiple tears, cracks, or deformation in plastic, and become critical when these holes allow excessive soil infiltration, failure, voids, and/or road/embankment failure.

- Longitudinal misalignment, which becomes critical when the pipe, embankment, or roadway accompany horizontal or vertical misalignment.

- Exposure of footing tops, which indicates deterioration and becomes critical when footings begin to crack, break or chip, or their bottoms become exposed.

- Damage to headwalls and wingwalls, marked by pervasive cracking, breaking, and/or chipping, gaps between the barrel and the wall, and exposure of footings. Such damage is considered critical if it includes collapse of the headwall/wingwall and damage to the embankment/roadway.

- For flared end sections, pervasive cracks, piping, or an end crushed or separated from the barrel indicate poor conditions, with conditions considered critical when this affects performance or causes roadway/embankment damage.

- Substantial blockage restricting at least a third of the area at the inlet, with conditions considered critical when sediment restricts more than half of the area.

- An inlet barrel raised above the streambed due to buoyancy, considered critical if embankment piping is observed.

Outlets were classified as "poor" or "critical" in the same manner as inlets if any of the following were observed: Invert deterioration, poor joints and seams, cross-sectional deformation, deficiencies in structural integrity of the barrel, longitudinal (mis) alignment, exposed footings, buoyancy or embankment piping, damage to headwalls/wingwalls, or cracking, piping, or separation of flared end sections. Additional factors also evaluated for outlets resulted in a "poor" or "critical" classification if either of the following was observed:

- Apron deterioration, such as pervasive cracking or piping/ undermining, with conditions considered critical if the apron is collapsed or affects performance to the extent that it is causing roadway/embankment damage.

- Damaged armoring, resulting in undermining or similar deterioration that reduces culvert performance, with conditions considered critical if armoring is failing and causing more pronounced negative impacts on culvert performance.

The above-described field data including information on both passability and structural integrity were entered into the online Crossings Database, where crossings were automatically scored for passability using an algorithm developed by the NAACC.

3.3.5 GEOMORPHIC DATA

General geomorphic context - including evaluation for potential structural loss or substantial damage due to clogging with woody debris, sediment blockage, or scour - was examined at each structure, for the upstream and downstream reaches. Conditions at the structure, upstream, and downstream were compared. Geomorphic failure risk was inferred from data collected specifically for other purposes (e.g., for passability and structural integrity), as well as based on a discrete block of questions specifically regarding geomorphic concerns (Appendix A, Field Data Sheets).

Data collection designed specifically to examine the fit of the crossing within the broad geomorphic context included:

- Width of the floodplain, and to what extent it is filled by the roadway.

- Water depth, velocity, and slope through the structure, compared to that of the stream. Was one deeper, faster, or steeper than the other?[9]

- Measurements of bankfull width, taken at regular intervals immediately upstream and downstream of each structure, and at regular intervals across reference sections upstream.

- Linear distance to hydraulic controls upstream and downstream.

- Depth of the pool downstream.

- Representative sampling of the streambed substrate[10], utilized as a quantitative measure of grain size, compared to the sediment within the structure.

- Accumulation and depth of sediment with respect to structure, the channel, and banks.

Comparison between conditions immediately upstream and downstream of the structure provided a second framework for examination, and included evaluating factors such as:

- Streambed erosion or aggradation,

- Breaks in the slope of the valley,

- Potential impacts on the road if the channel were to avulse (escape its banks),

- Dominant substrate grain size,

- Extent of bank erosion and condition of hardbank armoring, if any,

- Extent of scour at the structure inlet and outlet, including noting if bedrock is exposed,

- Type of material providing tailwater control, including armoring if noted,

- Height of banks upstream compared to downstream,

These geomorphic data were noted on field sheets, then entered into a spreadsheet for analysis. As already noted, the condition assessment provided further detail into the extent of scour and blockage at inlet and for scour at the outlet.

UMass Geosciences team members subsequently visited some 30 sites in the spring and summer of 2016 to collect missing geomorphic data, including dominant particle size, constriction ratio, alignment, and sediment continuity (aggradation/erosion upstream and downstream). Geomorphic scoring had been simplified to eliminate slope from the parameters considered by that point.

---

9   *Additional hydraulic data information is now available from the USGS flood-inundation maps (Lombard and Bent, 2015), but were not available at the time of analysis for the Deerfield Project.*

10 *The "pebble count" procedure was modified from Wolman (1954) and consisted of measuring the axes of fifty randomly selected individual samples collected from riffles, where sediment was neither accumulating nor eroding. The median axis of the samples was calculated as the quantitative representation of grain size.*

3.3.6 HYDRAULIC DATA

Field data in support of hydraulic calculations were also collected by TU in field data sheets (Appendix A). Measurement of dimensions began with length of the stream through the structure. At the inlet, they noted the shape of the crossing, and measured its cross-sectional height and width. They measured water depth at the inlet and any drop there. Similarly, at the outlet, crossing shape was indicated and cross-sectional height and width measured. Again, water depth and any drop at the outlet were measured. Information about the dominant substrate and the material on the walls was noted to constrain roughness.

To determine slope of the stream and through each culvert, TU measured elevations relative to the road using a level-sight and stadia rod. They measured the elevations of the upstream hydraulic control, at the inlet invert, the top of the pipe, the upstream road surface, the downstream road surface, the top of the pipe at the outlet, the outlet invert, downstream hydraulic control, and any pool bottom between hydraulic control and the outlet invert.

For bridges (some 130 crossings), cross-sectional area and wetted perimeter, slope, and roughness data were collected by the Massachusetts Geological Survey in fall 2015 and spring and summer 2016. Aiming to characterize the greatest constriction at the crossing, cross-sectional area and wetted perimeter were typically surveyed with a total station at the upstream edge of the bridge. In some instances, however, a laser range finder and stadia rod were able to more efficiently capture these dimensions. Where possible, a tape measure was also used to describe the cross-sectional widths between abutments and piers. The total station was often preferable, however, as it also provided an efficient way to capture slope of the river within the visible local reach and through the infrastructure. Roughness of the streambed at the structure was characterized by qualitative description of substrate grain size, and description of materials composing bridge abutments and piers, in addition to any embankment and/or vegetation along channel walls. Through the natural reach, we compared channel dimensions, substrate, and banks to empirically measured values in natural channels (Barnes, 1967) and to FHWA tables (Schall et al., 2008).

3.4   Geomorphic Data Collected by MMI

MMI developed a field data collection form to access relevant data previously collected by TU that would be useful to have in the field and to provide a location for recording field observations. Rather than develop a unique form, MMI leveraged existing data spreadsheets provided by the project team, which were retained in their original format, and a new spreadsheet was made that accessed existing data using the xycode. As described more fully in Section A.2 of Appendix A, the field data form included:

- Identification information such as CAPS ID, xycode, Culvert ID, stream name, road name, descriptive location, and latitude and longitude.

- Information about the structure such as number of culvert cells; culvert shape; width, height, and length;

- Culvert slope relative to channel slope;

- Culvert width as percent of bankfull channel width;

- Culvert alignment with channel; and

- Stream data, including

  ◦ drainage area,

  ◦ channel bankfull width,

  ◦ upstream channel slope,

  ◦ estimated bankfull flow,

  ◦ specific stream power (SSP),

  ◦ D50 (median grain size), and

  ◦ upstream and downstream channel substrate type (e.g. gravel vs. cobble).

Selection criteria for field assessment and past observations of damage were also included. The reverse side of the form included blank fields for data to be collected during each assessment.

MMI ultimately selected 197 bridges and culverts for field assessment in the Deerfield River watershed. Structures were selected for field assessment if they were reported or observed to be damaged in previous assessments by MassDOT, TU, or MMI. Structures with past damages were assigned a code by the towns to indicate the type of the problem (Table 3-2). Other structures were included that could potentially be damaged based on specific stream power (SSP) derived from GIS analysis in conjunction with dominant stream bed particle size.

Table 3-2: Structure damage codes.

Damage Code	Description
O	Overtopping
E	Embankment Failed
B	Blocked by Debris
S	Structural Failure
W	Washed Out
F	Roadway Flooding
L	Fluvial Erosion
*	Repeated Failures

Structures selected for assessment were located in the field using field maps developed by MMI and the existing location data. Once located, existing data were used to confirm the identity of the structure to be assessed. During each assessment, MMI measured the structure inlet width and height, structure length and slope, channel bankfull width, and local channel slope upstream and downstream of the structure. Local channel slope, structure slope, and measurements too large to make with a folding survey rod were

made using a laser rangefinder[11]. Pebble counts (Wolman, 1954) were performed at any structure lacking existing median grain size (D50) data using the size bins previously established by TU and UMass. Qualitative observations made in the vicinity of the structure included dominant particle size, channel bedforms, hydraulic features, stream channel geomorphic type, and channel stability. Collected field data are provided in the data repository (see Appendix O).

3.5 GIS Data

Road-stream crossings were analyzed for flow following the Stream Continuity Database list as of November 2014. Several of the hydraulic models utilized, as described in Section 5.3 of Chapter 5, required cumulative upstream drainage areas calculated for each crossing. Thus, this exercise first required a way to match location data for road-stream crossings to appropriate flow accumulation values in the 30 m x 30 m raster. The USGS National Elevation Dataset (NED) 30 m x 30 m Digital Elevation Model (DEM) was utilized. This assignment of flow accumulation data was performed in ArcGIS using the North Atlantic Landscape Conservation (NALCC) Streamlines and flow accumulation layer and the latitude and longitude for crossings recorded in the Stream Continuity Database. Those streamlines were taken from 1:25,000 National Hydrography Database (NHD) streams, uncorrected for flow[12]. Detailed command-line descriptions of this process (Crossing Flow Accumulation Assignment) are found in Appendix A.5.

Fifty-nine (6%) crossings needed to be reassigned latitude and longitude values to correspond to the appropriate 30 m x 30 m tile of the flow accumulation layer. The NALCC provided a useful raster file that characterized the fit of the flow accumulation layer to the streamlines. Tiles along the streamline that reflected those cumulative flow accumulation values were designated "type 1." In places where the river arcs crossed corners of tiles that were characterized by the much lower flow accumulation values of the surrounding terrain, the tiles were designated "type 2."

To reassign appropriate flow accumulation values to stream crossings, we first created points from the stream raster and snapped the crossing locations to the stream centerline. Crossings that plotted on a "type 2" tile were assigned a new flow accumulation value by drawing from the center point of the nearest "type 1" tile. Another eight culverts did not plot on a tile near a stream centerline. These were up in the headwaters of watersheds, but nonetheless within the

drainage. These were visually examined on the flow accumulation layer and an appropriate value was chosen for them.

A few hundred crossings were then visually examined to confirm that the method results were robust. Mis-assigned flow accumulation values of crossings were typically lower than they should have been because points plotting off the stream network tiles had much lower flow accumulation values. Thus, the visual check of crossings started with the ones assigned the smallest flow accumulation values.

Culvert and stream gauge locations in the Deerfield River watershed were delineated creating polygons of the contributing drainage area to each site. Select watershed attributes were calculated for each of the sites. The GIS processing utilizes the existing version 2 of the National Hydrography Dataset High Resolution Delineation (NHDHRDV2) layers and may serve as a template for future work on point delineation. This work was completed as a part of the hydrologic assessment completed by Gordon Clark for this project. More detail on the source of data used can be found at https://conte-ecology.github.io/projectRoadmap/

---

11 *Laser Technology, Inc.; Truepulse 360 Model; Centennial, CO; Accuracy: Inclination ±0.25 degrees = ±0.4% = ± 0.004 ft/ft; Distance ±1 foot for a reflective target, and ±3 feet for a poorly reflecting target.*

12 *The decision to utilize NHD streamline data rather than MassGIS hydrography data was based on the advice of Brad Compton, UMass, who has worked with both data sets and suggested the NHD and NED data sets for consistency throughout the MA and VT portions of the watershed. This also provided consistency with data utilized for development of the Stream Crossing Explorer.*

Future climate assessment

4.1 Introduction

Over the past several years, intense local storms as well as larger events such as tropical storms Irene and Sandy have resulted in considerable damage to both roads and streams, particularly within the Deerfield River watershed. An understanding of how uncertainty in future climate condition translates into uncertainty in streamflow is needed to effectively evaluate future vulnerability of the transportation network. This chapter summarizes the climate predictions evaluated as part of this study and describes the range and type of data subsequently utilized to predict streamflow under future conditions. Section 4.2 provides a general overview of available types of climate model predictions. Section 4.3 presents the range of predictions suggested by two frequently cited studies of relevance to the northeast. Section 4.4 summarizes climate model predictions specific to the Deerfield River watershed, that were downloaded, accessed and processed for this project, compares them, and notes differences to data utilized in another MassDOT funded study. Section 4.5 details what climate data were subsequently utilized for stream-flow prediction. Section 4.6 acknowledges the modeling groups responsible for the climate predictions discussed in Section 4.4.

4.2 Overview of Climate Models

4.2.1 GLOBAL CLIMATE MODELS

A wide range of climate model predictions and methodologies have been developed by the atmospheric modeling community. The Intergovernmental Panel on Climate Change (IPCC) is the international body charged with assessing the science related to climate change. Since 1988 they have released five assessments of published literature on climate change, the most recent completed in 2014. There is a lack of consensus on the models and scenarios most applicable for various regions. Most studies draw from Global Climate Model (GCM) simulations produced through the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset and phase 5 (CMIP5) multi-model ensemble[13]. Details on the CMIP3 and CMIP5 data are available from Maurer (2007) and Reclamation (2011, 2013, 2014).

---

13 *Archived at: http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/*

There is not yet a consensus among the climate science community on whether CMIP5 climate projects are more reliable than CMIP3 (U.S. DOT, 2014; Brekke et al., 2013). The CMIP3 model ensemble was utilized for the IPCC Fourth Assessment Report (AR4), released in May 2013. The IPCC Fifth Assessment Report (AR5) draws more heavily from CMIP5, but also considers CMIP3 modeling results. Data are available from the U.S. Bureau of Reclamation's Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections (DCHP) website (online at: http://gdo-dcp.ucllnl.org/downscaled_cmip_projections) (Maurer et al., 2007; Reclamation, 2013). Direct comparisons between the datasets are not possible as they utilize different emission scenarios. The CMIP3 data available through the DCHP website include 9 climate models utilizing emission scenarios B1, A1B, and A2, published in the IPCC Special Report on Emissions Scenarios associated with the IPCC Fourth Assessment Report (Reclamation, 2011). CMIP3 simulation results are available for only three specific time periods: 1961 – 2000, 2046 – 2065 (mid-century), and 2081-2099 (end-of-century). The CMIP5 data available through the DCHP website include 21 climate models utilizing Representative Concentration Pathways (RCPs), specifically RCP2.5, RCP4.5, RCP6.0, and RCP8.5 (Reclamation, 2014). CMIP5 simulation results are available from 1950 – 2099.Table 4-1 and Table 4-2 provide brief descriptions of the emission scenarios and representative concentration pathways for CMIP3 and CMIP5, respectively. Please refer to Table 4-8 in Section 4-6 for details on the modeling groups that contributed to CMIP3 and CMIP5.

4.2.2 REGIONAL CLIMATE PREDICTIONS

Both CMIP3 and CMIP5 results have been utilized to generate more regionally specific summaries of future climate predictions. Local-scale climate projections can be derived from coarse Global Climate Model (GCM) fields through dynamical or statistical downscaling. Dynamical downscaling involves the use of a high-resolution model to simulate physical climate processes at finer spatial scales. In these simulations, a high-resolution regional climate model (RCM) is forced at the boundaries of a given region with the coarse-scale GCM data. In contrast to dynamical downscaling, statistical downscaling is designed to capture historical relationships between large-scale weather features and local climate, with the estimated relationships then used to translate future projections down to the desired fine spatial scale, provided adequate observations exist to inform the downscaling. An advantage of statistical downscaling over dynamical downscaling is numerical cost; high-resolution

Table 4-1: CMIP3 emissions scenarios.[14]

	Description	Global Surface Temp. Change by 2100	Global Mean Sea Level Rise by 2011
B1	Low emissions. Declining global population after mid-century, transition to lower emission technologies and economies.	0.54 - 1.62 °F (0.3 - 0.9 °C)	0.59 – 1.25 ft (0.18 - 0.38 m)
A1B	Medium-High emissions. Rapid economic growth, declining global population after mid-century, more efficient technologies.	3.06 - 7.92 °F (1.7 - 4.4 °C)	0.69 – 1.57 ft (0.21 – 0.48 m)
A2	Medium-High emissions. High population growth, regional economic development, slower technology change.	3.6 – 9.72 °F (2.0 – 5.4 °C)	0.75 – 1.67 ft (0.23 – 0.51 m)

Table 4-2: CMIP5 representative concentration pathways.[15]

Scenario Name	Description	Concentrations (ppm CO2 equiv.) by 2100	Global Surface Temp. Change by 2100	Global Mean Sea Level Rise by 2100
RCP2.6	Substantial and sustained emissions reductions	475	0.5 – 3.0 °F (0.3 - 0.9 °C)	0.85 – 1.8 ft (0.26 - 0.55 m)
RCP4.5	Stabilization	630	2.0 – 4.7 °F (1.1 – 2.6 °C)	1.0 – 2.1 ft (0.32 - 0.63 m)
RCP6.0	Stabilization	800	2.5 – 5.6 °F (1.4 – 3.1 °C)	1.1 – 2.1 ft (0.33 - 0.63 m)
RCP8.5	High emissions continue	1313	4.7 – 8.6 °F (2.6 – 4.8 °C)	1.5 – 2.7 ft (0.45 - 0.82 m)

modeling often takes months of computer time to complete each simulation. However the advantage of the dynamically downscaled data is the resulting high spatial-resolution data. Dynamically downscaled climate projections are generally preferred over statistically downscaled data if they are to be used as the forcing data for hydrologic models.

Among the available downscaled climate projections for New England suitable for use with physically based models, the following are of particular note:

1. Hayhoe et al. (2007) applied statistical downscaling to a set of five GCMs participating in the CMIP3 program. The Climate Change Adaptation Report (2011) released by EEA references the Hayhoe data. The Hayhoe et al. (2007) simulations are based on GCMs from CMIP3.

2. The North American Regional Climate Change Assessment Program (NARCCAP) produced high-resolution simulations across North America from different combinations of atmosphere-ocean general circulation models (AOGCMs) and regional climate models (RCMs). NARCCAP utilizes dynamical downscaling. Spatial resolution for each model is

approximately 50 km and daily precipitation and air temperature data are available for two 30-year periods, 1971-2000 (present) and 2041-2070 (future). The NARCCAP study is also based on CMIP3 data.

3. Downscaled climate scenarios for the globe at a resolution of 0.25 x 0.25 degrees are available from the National Aeronautics and Space Administration (NASA) Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) website for 42 climate projections from 21 CMIP5 GCMs and two RCP scenarios (RCP4.5 and RCP8.5) for the period from 2006 to 2100 (Thrasher et al., 2012). The Bias-Correction Spatial Disaggregation (BCSD) method was used to produce final grids.

Two other studies of note were not complete in time to be of use for this project. The United States Global Change Research Program (USGCRP) is producing a National Climate Assessment that will contain a section on changes across the Northeast U.S., including estimates derived from the NARCCAP regional climate model suite. Additionally, the Northeast Climate Adaptation Science Center is beginning a study to identify the most appropriate climate model and scenario to utilize for future studies.

Dynamically downscaled data available from the NARCCAP and NEX-GDDP programs were considered for this project and are discussed further in Section 4.4.3.

---

14 *Sourced from U.S. DOT (2014); Original source: UN IPCC Working Group I: The Scientific Basis (https://www.ipcc.ch/report/ar3/wg1/)*

15 *Sourced from U.S. DOT (2014); Original source: UN IPCC, Climate Change 2013: The Physical Science Basis (http://www.ipcc.ch/report/ar5/wgl/)*

4.3 Summary of Climate Prediction Utilized in Two Studies of Note

Several studies have documented observed and predicted climatic changes in the Northeast. While a complete literature review is beyond the scope of this project, brief summaries from two studies are provided in the following section. These studies were utilized to help identify mid- and end-century multipliers for the statistical models utilized in the project, as well as for comparison against the data utilized for the physically based models.

4.3.1 NATIONAL CLIMATE ASSESSMENT REPORT

In 2014 the U.S. Global Change Research Program published the National Climate Assessment report Our Changing Climate (referred to as the NAC3 report), which summarizes the impacts of climate change on the United States (Melillo et al., 2014). A team of more than 300 experts contributed to the report, which was compiled by a 60-member Federal Advisory Committee and included extensive public and expert review. Chapter 2 (Walsh et al., 2014) provides an overview of observed and predicted climate change, with further details provided in Appendix 3 of the report. Chapter 3 (Georgakakos et al., 2014) covers the water sector, and Chapter 16 (Horton et al., 2014) focuses more specifically on the Northeast region. The NAC3 Report incorporates projections from three sets of model simulations, including:

- 25 representations of different models from CMIP3, completed for the Fourth IPCC assessment, which were utilized as the foundation for most of the report findings,

- 30 representations of different models from CMIP5, completed for the Fifth IPCC assessment, which were utilized primarily for comparison purposes, and

- Six regional climate model analyses from the North American Regional Climate Change Assessment Program (NARCCAP) regional climate model analysis for the continental U.S. run for current/past (1971-2000) and projected (2041-2070) time periods.

The spatial resolution of CMIP3 ranges from 125 to 187 miles, while that of CMIP5 ranges from 62 to 125 miles depending on latitude. NARCCAP, in comparison, runs simulations at a ~30-mile horizontal resolution based on coarser resolution results from CMIP3 models used as the boundary conditions.

The following key observed changes in climate noted in the NAC3 report for the Northeast are:

- Heavy downpours have increased across the nation over the last three to five years,

- Over the Northeast region specifically,

  - The occurrence of very heavy precipitation – classified as the 2-day, 5-year return interval rainfall total (two-day rainfall rate that is exceeded 20% of the time) - increased 50% from 2001 – 2012 compared to the period 1901 – 1960,

  - The one-day, 100-year return interval rainfall total (daily rainfall rate that is exceeded 1% of the time) has increased by 70% over the period 1958 – 2011 (Walsh et al., 2014).

Based on the projections for the Northeast included in the NAC3 Report, higher emissions (A2) scenario results predict a warming of 4.5°F – 10°F by the 2080s, while reduced emission (B1) scenario results predict a warming of 3°F to 6°F. In addition to net increases in annual temperature, the frequency, intensity, and duration of heat waves is anticipated to increase. Precipitation predictions are more uncertain, but precipitation, particularly in winter and spring, and the frequency of heavy precipitation are expected to increase. Seasonal drought risk is also expected to increase, particularly in the summer and fall. End-century predictions for the Northeast range as follows:

- 2070 – 2099 compared to 1971 – 2000. CMIP 3 A2:

  - Annual maximum precipitation increase from 10 – 30%
  - Winter mean precipitation, 10 - 20%
  - Spring mean precipitation, 0 – 10%
  - Summer mean precipitation, no change
  - Fall mean precipitation, 0 – 10%

- 2070 – 2099 compared to 1971 – 2000. CMIP5 RCP 2.6:

  - Winter mean precipitation, 0 - 10%
  - Spring mean precipitation, 0 – 10%
  - Summer mean precipitation, 0 – 10%
  - Fall mean precipitation, no change

- 2070 – 2099 compared to 1971 – 2000. CMIP5 RCP 8.5:

  - Winter mean precipitation, 10 - 20%
  - Spring mean precipitation, 10 - 20%
  - Summer mean precipitation, 0 – 10%
  - Fall mean precipitation, 0 – 10%

The NCA3 suggested a heavy rain multiplier for the Northeast at the end-of -century is 1 – 2 for RCP 2.6 and 3 – 4 for RCP 8.5.

4.3.2 2015 NEW YORK CITY PANEL ON CLIMATE CHANGE

The 2015 New York City Panel on Climate Change (NPCC, 2015) provides an overview of several relevant observed changes and predictions that are drawn from data that are generally applicable across the northeast. It is useful to see the types of data other government agencies consider important for regional planning, as well as how they attempt to bound uncertainty. Example changes and predictions presented in the NPCC (2015) report include:

- Annual average precipitation in New York City (Central Park) has increased at a rate of approximately 0.8 inches/decade from 1900 to 2013; Year-to-year variability, as measured by the standard deviation, has increased from 6.1 inches from 1900 – 1956 to 10.3 inches from 1957 to 2013. Data graphics in the report cover the New England area, indicating that annual average precipitation in the Deerfield River watershed from

1991 – 2012 was 5 to 15% greater relative to the 1901 – 1960 average.

- Mean annual temperature in New York City (Central Park) increased at a rate of 0.3oF per decade from 1900 to 2013.

- Six of the top 10 wettest days on record in New York City have occurred since 1972.

- Based on 25th and 75th percentile results of 35 global climate models (GCMs) and two Representative Concentration Pathways (RCPs),

  - Mean annual temperatures in New York City are projected to increase by 4.1 to 5.7oF by the 2050s and by 5.3 to 8.8oF by the 2080s;

  - Mean annual precipitation is projected to increase by 4 to 11% by the 2050s and by 5 to 13% by the 2080s.

- The number of extreme events is also expected to increase slightly as follows, based on the 25th and 75th percentile results of the same 35 GCMs and two RCPs:

  - Number of days per year with rainfall at or greater than 1" - 14 to 15 days in the 2020s, to 14 to 16 days in the 2050s and 15 – 17 days in the 2080s

  - Number of days per year with rainfall at or greater than 2" – 3 to 4 days in the 2020s, 4 days in the 2050s, and 4 to 5 days in the 2080s

  - Number of days per year with rainfall at or greater than 4" – 0.3 to 0.4 days in the 2020s, 0.3 – 0.4 days in the 2050s, and 0.3 – 0.5 days in the 2080s.

In summary, New York City is considering impacts due to changes in daily and annual average precipitation as well as mean annual temperature based on the 25th and 75th percentile predictions of 35 GCMs for two RCPs.

4.4 Climate Prediction Data Downloaded for the Deerfield Project

Climate prediction data from four sources were downloaded specifically for the Deerfield River watershed for further evaluation as part of this project. These data included (1) GCM data for CMIP3 and CMIP5 available through a United States Department of Transportation (U.S. DOT) CMIP Climate Data Processing Tool, (2) dynamically downscaled GCM-RCM climate predictions from

the North American Regional Climate Change Assessment Program (NARCCAP), (3) dynamically downscaled GCM-RCM climate predictions from the NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) and (4) climate simulations by the NCAR Community Climate System Model (CCSM4) downloaded from the National Center for Atmospheric Research (NCAR) Climate Inspector. In this section each data set is first briefly described and then the range of predictions is summarized.

4.4.1 GCM PREDICTIONS FROM DOT CMIP PROCESSING TOOL

In 2014, the U.S. DOT released the U.S. DOT CMIP Climate Data Processing Tool[16]. The tool was designed to make statistical summary information based on downscaled climate data readily available to transportation planners for use at the local level. The tool utilizes data from CMIP3 and CMIP5, downloaded via the U.S. Bureau of Reclamation's Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections (DCHP) website[17]. Processing of the data is done through two Excel files, one specific to CMIP3 data and one specific to CMIP5 data. The tool provides a variety of temperature and precipitation statistical summaries at mid- and end-century, as well as estimates of change compared to modeled "current" climate conditions.

For the purpose of the UMass-DOT project, the U.S. DOT CMIP tool provided the opportunity to examine the variability of CMIP3 and CMIP5 GCM predictions available specifically for the Deerfield River watershed. Data were downloaded for four different grids:

- **Grid 1:** (latitude: 42.5625, longitude: -72.9375)

- **Grid 2:** (latitude: 42.5625, longitude: -73.0625)

- **Grid 3:** (latitude: 42.6875, longitude: -73.0625)

- **Grid 4:** (latitude: 42.6875, longitude: -72.9375)

These four grid cells were selected as being representative of the variation in climate conditions across the watershed, constrained by the capacity of the CMIP Climate Data Processing Tool to accept data from multiple grids. CMIP3 data were available for three time periods (1961-2000, 2046-2065, and 2081-2099), while CMIP5 data were available from January 1950 to December 2099.

The data for these four grids were combined through the processing tool to generate "composite" mid- and end-century summary statistics. Although the DCHP website was designed to facilitate the download of multiple model and emissions path scenarios for each grid to develop ensemble prediction, for this comparison, data were downloaded for each CMIP – Emissions Scenario combination individually. However, data for all runs of the model were included. Different runs represent slight changes in the initial conditions utilized for the model. CMIP3 data were down-loaded as daily data, referred to as "climate daily" by the tool, using 1/8th degree freedom with variables of precipitation rate as well as minimum and maximum surface air temperature. The available CMIP3 emission scenarios and climate models are listed in Table 4-3. CMIP5 data were also downloaded as "climate daily", using 1/8th

---

16 *Prepared by ICF International for the U.S. Department of Transportation Center for Climate Change and Environmental Forecasting under* The Gulf Coast Study, Phase 2, Impacts of Climate Change and Variability on Transportation Systems and Infrastructure.

17 *Sourced from CMIP user guide, http://gdo-dcp.ucllnl.org/ downscaled_cmip_projections/*

degree freedom with the same variables. The available CMIP5 emission scenarios and climate models are listed in Table 4-4.

The data were then run through the U.S. DOT processing tool to provide current, mid- and end-century statistics for the following specific temperature and precipitation variables[18]:

- Average Total Annual Precipitation

- "Very Heavy" 24-hr Precipitation Amount (defined as 95th percentile precipitation)

- "Extremely Heavy" 24-hr Precipitation Amount (defined as 99th percentile precipitation)

- Average Number of Baseline "Very Heavy" Precipitation Events per Year (defined as >0.7 inches in 24 hours)

- Average Number of Baseline "Extremely Heavy" Precipitation Events per Year (1.3 inches in 24 hours)

- Average Total Monthly Precipitation January to December

- Average Total Seasonal Precipitation Winter to Fall

- Largest 3-Day Precipitation Event per Season Winter to Fall

- Average Annual Mean Temperature

- Average Annual Maximum Temperature

- Average Annual Minimum Temperature.

The summary statistic data were then transferred into an Excel sheet.

---

18 Sourced from climate data processing tool

4.4.2 REGIONAL CLIMATE PREDICTIONS FROM NARCCAP AND NEX-GDDP

Two sets of dynamically downscaled projections were considered for this project, data from the NARCCAP and NEX-GDDP projects. NARCCAP data are based on CMIP3. The spatial resolution for each model in the NARCCAP ensemble is approximately 50 km, and daily precipitation and air temperature data are available for two 30-year periods, 1971-2000 (present) and 2041-2070 (future). NARCCAP data are downloadable from http://www.narccap.ucar.edu/. In contrast, NEX-GDDP data are based on CMIP5. The spatial resolution for each is approximately 0.25 x 0.25 degrees, and data are available for the period from 1950 – 2100. NEX-GDDP data are downloadable from https://cds.nccs.nasa.gov/nex-gddp/.

A number of GCM-RCM model outputs from each data set were downloaded and compared against observed station data at the annual and seasonal scale, Appendix C. Comparisons of the NARCCAP and NEX-GDDP model outputs against observed station data at the seasonal scale indicated that NEX-GDDP data were biased low for the Deerfield River watershed region. NEX-GDDP data were thus excluded from further use by the project. The NARCCAP data are described further below.

Air temperature and precipitation data archived under the NARCCAP project represent some of the highest resolution data currently available for climate impact studies. The NARCCAP high-resolution simulations across North America result from different combinations of AOGCMs and RCMs. Output from four AOGCM models (GFDL, CGCM3, HADCM3 and CCSM3) based on the SRES A2 emissions scenarios provide the boundary conditions for six regional models (MM5 – Iowa State; RegCM3 – UC Santa Cruz; CRCM – Quebec, Ouranos; HADRM3 – Hadley Centre; RSM – Scripps; and WRF-PNNL). Further details are available from Mearns et al. (2009) and Mearns et al. (2007, updated 2014).

Table 4-3: CMIP3 model and emission scenario results for the Deerfield River watershed processed through the DOT CMIP Processing Tool.

GCM	Emissions Scenario		
	A1b	A2	B1
cccma_cgcm3_1	three runs	three runs	three runs
cnrm_cm3	one run	one run	one run
gfdl_cm2_0	one run	one run	one run
gfdl_cm2_1	one run	one run	one run
ipsl_cm4	one run	one run	one run
miroc3_2_medres	two runs	two runs	two runs
miub_echo_g	three runs	three runs	three runs
mpi_echam5	NA	one run	one run
mri_cgcm2_3_2a	five runs	five runs	five runs

Table 4-4: CMIP5 model and emission scenario results for the Deerfield River watershed processed through the DOT CMIP Processing Tool.

GCM	Emissions Scenario			
	RCP2.6	RCP4.5	RCP6.0	RCP8.5
access1-0	NA	one run	NONE	one run
bcc-csm1-1	one run	one run	one run	one run
bnu-esm	NA	NA	NA	NA
canesm2	five runs	five runs	NA	five runs
ccsm4	two runs	two runs	two runs	two runs
cesm1-bgc	NA	one run	NA	one run
cnrm-cm5	NA	one run	NA	one run
csiro-mk3-6-0	ten runs	ten runs	ten runs	ten runs
gfdl-cm3	one run	NA	one run	one run
gfdl-esm2g	one run	one run	one run	one run
gfdl-esm2m	one run	one run	one run	one run
inmcm4	NA	one run	NONE	one run
ipsl-cm5a-lr	three runs	four runs	one run	four runs
ipsl-cm5a-mr	one run	one run	one run	one run
miroc-esm	one run	one run	one run	one run
miroc-esm-chem	one run	one run	one run	one run
miroc5	three runs	three runs	one run	three runs
mpi-esm-lr	three runs	three runs	NA	three runs
mpi-esm-mr	one run	three runs	NA	one run
mri-cgcm3	one run	one run	one run	one run
noresm1-m	one run	one run	one run	one run

Table 4-5: NARCCAP data utilized in project.

GCM	RCM
CGCM3	RCM3
CGCM3	CRCM
CGCM3	WRFG
CCSM	CRCM
CCSM	MM5I
CCSM	WRFG
GFDL	ECP2
GFDL	RCM3
GFDL	HRM3

Rawlins et al. (2012) examined the reliability of NARCCAP projections across the northeast U.S. compared to observed data and found good agreement in spatial variability and pattern for temperature, but only moderate agreement for precipitation. They observed a modest seasonal cold bias for temperature, and a wet bias for precipitation in winter, spring, and summer.

For impacts analysis or use as forcing inputs for modeling (e.g. hydrologic models), biases in the RCM estimates need to be removed before the data can be applied. Bias-correction of the NARCCAP data for the Deerfield River watershed is described in Appendix D. The final daily air temperature and precipitation data utilized for the project are provided as part of the data repository. Nine of the NARCCAP GCM-RCM model outputs were included, Table 4-5. The other two available GCM-RCM combinations were discarded due to remaining biases compared to observed data. In

addition, one of the hydrologic models was not able to run the CCSM_WRFG or CGCM3_WRFG projections. Because NARCCAP climate predictions are available only through mid-century, estimation of end-century hydraulic vulnerability based on NARCCAP data is not possible.

4.4.3 NATIONAL CENTER FOR ATMOSPHERIC RESEARCH CLIMATE INSPECTOR

The National Center for Atmospheric Research (NCAR) Climate Inspector is an interactive web application that enables visualization and download of climate simulations by the NCAR Community Climate System Model (CCSM4) prepared for the 5th Assessment

Report of the Intergovernmental Panel on Climate Change. It is available online at: https://gisclimatechange.ucar.edu/inspector). The inspector tool enables the user to explore climate anomalies, variability, and uncertainty in space and time. The range of results predicted for the Deerfield River watershed is summarized in Table 4-6.

4.4.4 COMPARISON OF DOWNLOADED DATA

4.4.4.1 COMPARISON OF GCM AND GCM-RCM DATA
Figure 4-1 and Figure 4-2 summarize the range of mid- and end-century, respectively, percent changes in mean annual temperature and precipitation predicted across the CMIP3 and CMIP5 GCM and emission scenarios for the Deerfield River watershed, downloaded

Table 4-6: Summary of NCAR Community Climate System Model precipitation change predictions for the Deerfield River watershed.

	Low (RCP 2.6)		Medium (RCP 4.5)	
	Annual change (in)	April total (in)	Annual change (in)	April total (in)
2000 - 2019	+ 0.3	3.9	2.2	3.9
2041 - 2060	+2.3 (5%)	4.3 (10%)	2.5 (5.6%)	4.2 (7.7%)
2071 - 2090	+2.9 (6%)	4.3 (10%)	3.5 (8%)	4.3 (10%)

	Medium High (RCP 6.0)		High (RCP 8.0)	
	Annual change (in)	April total (in)	Annual change (in)	April total (in)
2000 - 2019	+ 1.3	3.9	1.0	3.9
2041 - 2060	+2.3 (5%)	4.2 (7.7%)	3.4 (7.6%)	4.4 (13%)
2071 - 2090	+4.1 (9%)	4.6 (18%)	5.3 (12%)	4.5 (15%)

Figure 4-1: Percent change from current to mid-century of mean annual temperature and precipitation based on CMIP3 and CMIP5 climate predictions download from the DCHP website and processed through the DOT tool.

End–Century Change in Temperature vs Change in Precipitation
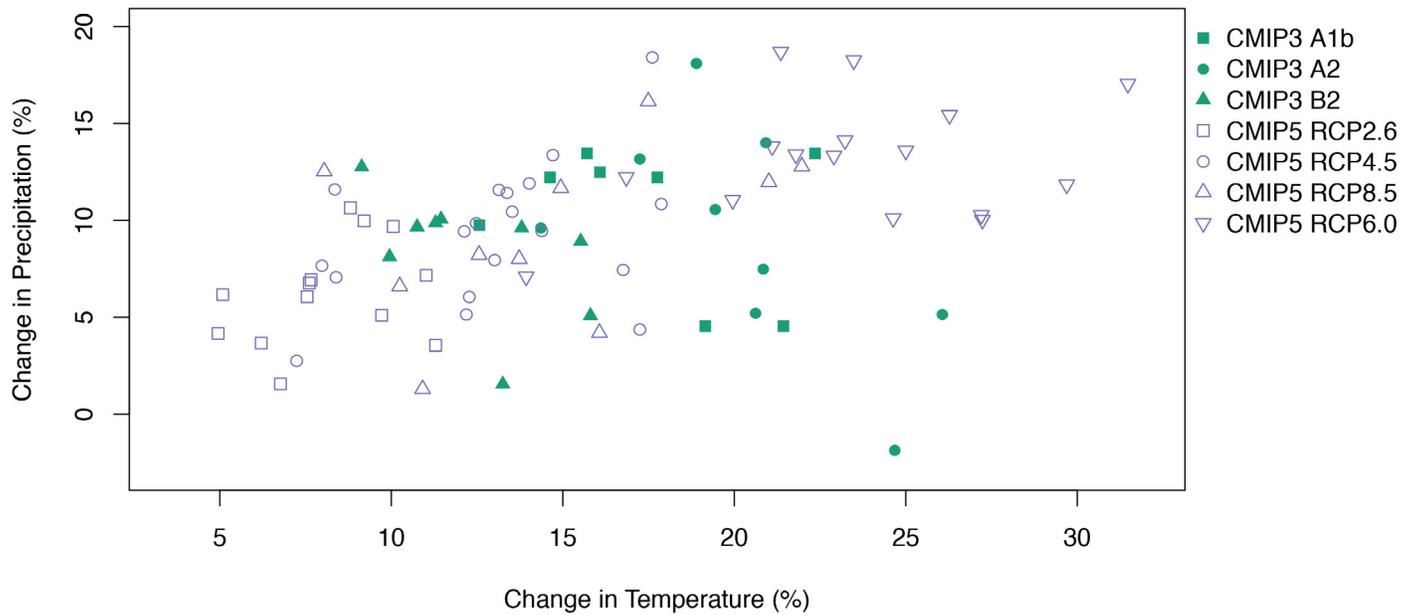
Figure 4-2: Percent change from current to end-century of mean annual temperature (T) and precipitation (P) based on CMIP3 and CMIP5 climate predictions download from the DCHP website and processed through the DOT tool.

from the DCHP website, as processed by the U.S. DOT Climate Processing Tool. CMIP3 data are shown as shaded symbols while CMIP5 data are shown as open symbols. Changes predicted by the 9 NARCCAP GCM-RCM models utilized in the project are included on the figure for mid-century, Figure 4-1, shown as stars.

As discussed further in the next section, GCM-RCM data may provide some advantage over GCM data at the local scale (Rawlins et al., 2012), and regional studies in the northeast have previously used GCM-RCM climate predictions over GCM data (Massachusetts EEA, 2011). In general, the NARCCAP data for the Deerfield River watershed fall in a narrower band in terms of predictions for changes in temperature, ranging from a 7.5 to 12.5% change in annual temperature by mid-century compared to a range from about 7% to 17% for CMIP3 data and from about 5% to 20% for CMIP5 data. Some of the NARCCAP GCM-RCM data indicate slight decreases in annual precipitation, and one predicts a 20% increase, but in general predictions range from slight (1 – 2%) to 10% increases in annual precipitation. CMIP3 and CMIP5 GCM predictions both range from a 0% change in annual precipitation to ~ 16% increase at mid-century. Predicted changes in precipitation at end-century fall across the same range, but both the CMIP3 and CMIP5 predictions are more variable in terms of temperature, ranging from increases of ~9% to 27% and 5% - 34% respectively.

4.4.4.2 COMPARISON TO MASSDOT CLIMATE ADAPTATION VULNERABILITY ASSESSMENT STUDY DATA

Development of downscaled climate projections is underway for the MassDOT statewide Climate Change Adaptation Plan[19] (CCAP). The plan will provide three sets of climate projection maps for four future time periods and emissions scenarios, including:

- Projected percent change in future 24-hour 100-year return interval precipitation,

- Projected future 24-hour 100-year return interval precipitation depth, and

- Projected annual maximum number of consecutive days > 95°F.

The data sets are based on climate projections for three emission scenarios from CMIP5, specifically RCP4.5, RCP6.0, and RCP8.5. The four future time periods considered are 2030, 2050, 2070, and 2100. For each data set and RCP, "low," "median," and "high" projections are provided based on the 10th, 50th, and 90th percentile of the CMIP5 model-ensemble. The raster data for the maps were developed through the CLIMSystem's SimCLIM 2013 climate change analysis software, which processes the raw GCM outputs with the pattern scaling method[20].

A direct comparison between the precipitation projections utilized in this project and those available through the MassDOT statewide CCAP study was not feasible. The CCAP analyses focused on precipitation data utilized for design storm calculations (e.g., 24-hour duration 100-year storm precipitation), while the UMass study focused on precipitation data necessary for generating

---

19 Further details available online at: http://gis.massdot.state.ma.us/cpws/
20 Based on the technical document published as part of the website.

streamflow estimates from physically based models (e.g., daily precipitation accumulation data) or statistical models (April or annual total precipitation).

The scientific community has not yet determined which climate projection data are most reliable for the northeastern U.S. It is thus important to place the data utilized for any project in context with other available climate projection data. In this light, a few differences between the two project approaches are worth noting:

- GCM predictions are generally considered too coarse in terms of spatial scale for direct use at the local scale. Some form of downscaling is recommended. In addition, even once downscaled, a bias-correction based on local observed data is recommended.

- Daily temperature and precipitation data are required input for the physically based models utilized in the DRW Pilot. Daily climate predictions for mid-century were downscaled specifically for the Deerfield River watershed as follows (refer to Appendix D):

  ◦ Data were downscaled from GCM models by utilizing the GCM data as boundary input conditions for RCMs (e.g., GCM-RCM model results were utilized). The GCM-RCM data were leveraged from the NARCCAP and NEX-GDDP programs.

  ◦ The GCM-RCM model projections were further bias-corrected to local observed data based on a monthly algorithm.

  ◦ GCM-RCM model precipitation projections based on CMIP5 specific to the Deerfield River watershed (NEX-GDDP data) were omitted from the UMass project as they were biased low compared to observed data in terms of the average number of annual days with greater than 1-inch of precipitation.

These downscaled data were used as input to the physically based models to provide return interval (RI) flow estimates at mid-century.

- The SimCLIM 2013 utilized by the MassDOT statewide CCAP does not include downscaling or bias-correction. It is based on the ensemble 0.5 degree x 0.5 degree resolution data for three emission scenarios (RCP4.5, 6.0, and 8.5) and four time periods (2030, 2050, 2070, 2100) across 22 GCMs.

- To address downscaling concerns for empirical (statistical) model applications (refer to Section 4.5), the DRW Pilot took into consideration CMIP3, CMIP5, and GCM-RCM data to set a range of percent changes in precipitation and temperature that may occur at mid- and end-century across the Deerfield River watershed.

  ◦ The percent changes were then paired with local observed data.

  ◦ The local observed data, adjusted by a range of potential change factors, were then utilized by the models.

It is beyond the scope of this project to explicitly calculate either the percent change or depth of the 24-hour 100-year return interval precipitation for direct comparison against the MassDOT statewide CCAP data. The daily data projections generated for the UMass project have not been converted statistically to return interval level data. MassDOT statewide CCAP return interval data may, however, be compared against other data evaluated for this project. As noted earlier, the NCA3 report suggests a heavy rain multiplier for the end of the century in New England of 1 to 2 for RCP2.6 and 3 to 4 for RCP8.5. The NCAR Climate Inspector indicates that the 24-hour duration 100-year storm precipitation depth through the end of the century (10th, 50th, and 90th ensemble member) is anticipated to be 6 to 8 inches for RCP4.5, 6 to 10 inches for RCPs 6.0 and 8.5. The current value for the region is 6.0 to 6.5 inches (available online: precip.eas.cornell.edu). Thus in terms of percent change, the viewer suggests a 0 to 10% change through the end of the century based on RCP4.5, 0 to 40% based on RCP6.0, and 0 to 50% change based on RCP8.5. Projections of extremely heavy 24-hr, 1-day precipitation from CMIP3 and CMIP5 projections, as processed through the DOT CMIP Processing Tool, are summarized for mid-century (2046 – 2065) and end-century (2079 – 2099) on Figure 4-3 and Figure 4-4, respectively. These rainfall depths are not the same as presented by the MassDOT statewide CCAP, but they do indicate similar trends of increases from 10 to 50%. In summary, there is a wide range of uncertainty associated with precipitation projections.

4.5   Climate Data Utilized for Streamflow Predictions

This section of the report describes how the climate predictions presented in Section 4.3 (Studies of Note) and Section 4.4 (Climate Predictions Downloaded for this Study) serve as model input to predict future streamflows and associated uncertainty.

4.5.1   OVERVIEW – UTILIZATION OF CLIMATE DATA TO ESTIMATE FUTURE STREAMFLOW

For the purposes of this project, hydrologic conditions are considered to be processes that control the translation of rainfall to runoff across a watershed, while hydraulic conditions refer to the resulting streamflows that have the potential to disrupt the transportation network. A range of methods may be considered when conducting a vulnerability assessment for road-stream crossings due to future extreme flows, including:

a. Utilizing specific GCM or GCM-RCM precipitation and climate model predictions to drive empirically based or physically based models that require such forcing data,

b. Adjusting historical precipitation and climate data by set percentages or values to encompass the projected changes in climate, and utilizing these adjusted data to drive empirically based or physically based models that require such forcing data, or

Extremely Heavy 24–hr Precipitation (99th Percentile) Current vs Mid–Century
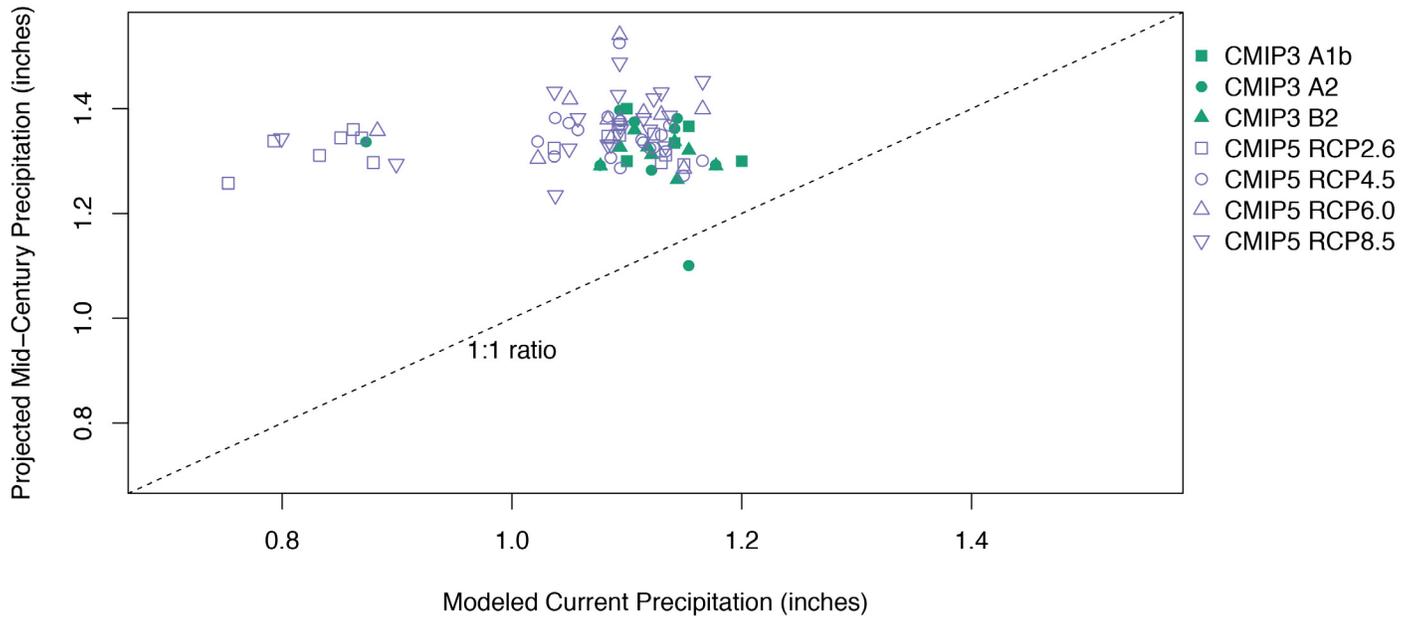
Figure 4-3: DOT CMIP Processing Tool summary of mid-century extremely heavy (99th percentile) 24-hr precipitation depth predictions.

Extremely Heavy 24–hr Precipitation (99th Percentile) Current vs End–Century
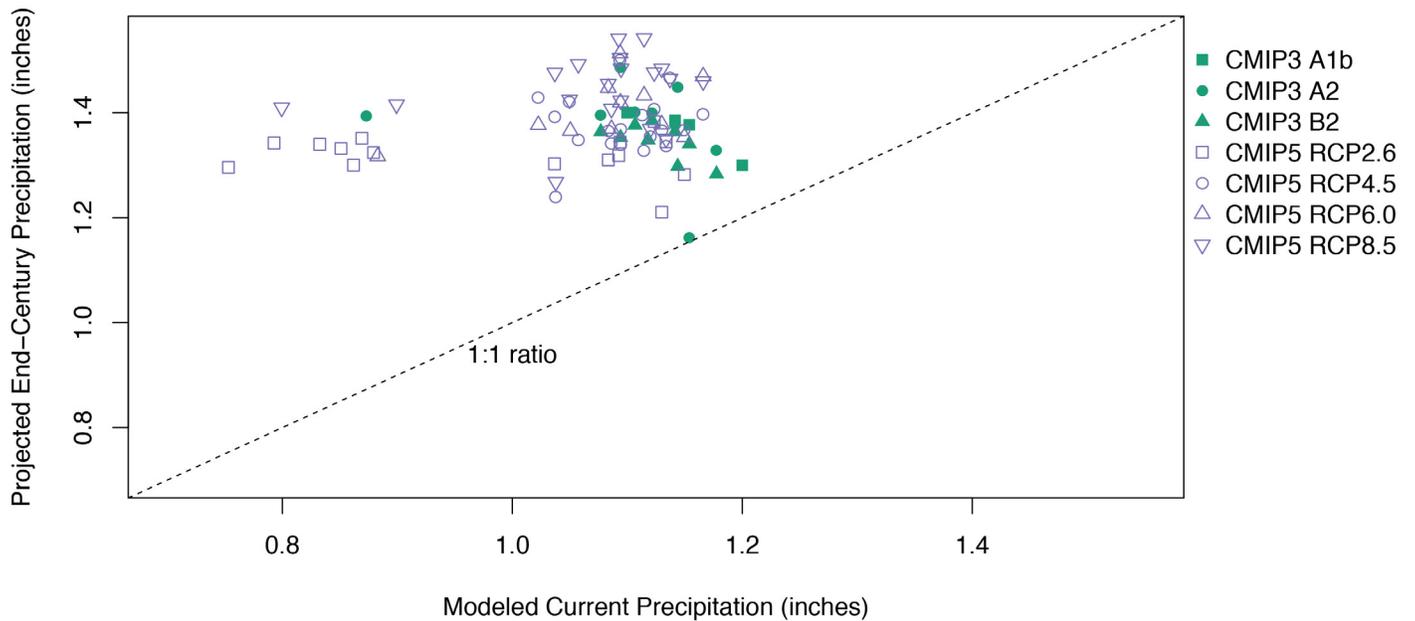
Figure 4-4: DOT CMIP Processing Tool summary of end-century extremely heavy (99th percentile) 24-hr precipitation depth predictions.

c. Utilizing statistical methods and historical data to extrapolate what streamflow conditions may be under future climate conditions, then placing these changes in the context of the various climate models and scenarios.

For this project, approach (a) is utilized for the physically based models while approach (b) is utilized for the statistically based streamflow models. MassDOT does not currently endorse the use of forecasted precipitation frequency data for hydrologic computations. It is important to note that in this project, forecasted data are utilized for estimating vulnerability, rather than predicting streamflows as a basis for design.

Climate data requirements vary across hydrologic and hydraulic models, as described in detail in Chapter 5.3. Statistical modeling approaches (also referred to as empirical models) require aggregate data, such as total annual precipitation, monthly precipitation, or multi-day extreme precipitation. Mathematical approximations translate summary statistics of watershed characteristics and, in some cases, rainfall properties to hydraulic condition (e.g., discharge magnitude) associated with specific probabilities of occurrence as defined by historical records. In contrast, physically based models require sub-daily climate data, including estimates of temperature and precipitation, to directly simulate both hydrologic and hydraulic processes with varying levels of detail. Streamflow estimates from physically based models may then be analyzed statistically to determine the probability of occurrence of a given flow.

The different climatic forcing requirements for statistical versus physically based models provide different insights as well as uncertainties. Most statistical models do not incorporate rainfall as a predictor variable and thus are not able to predict changes in flow due to climate. Statistically based models that include precipitation as a predictor variable are typically more suited to capturing the influence of seasonal or annual precipitation on streamflows (e.g., riverine floods). Physically based hydrologic models are explicitly designed to simulate the translation of short-term rainfall to streamflow conditions (e.g., flash floods), but are limited by the spatial and temporal resolution of available precipitation data. In addition, the general consensus of the scientific climate research community is that projected changes in precipitation are highly uncertain, with uncertainty increasing with decreasing spatial

(e.g., global to regional to watershed) and temporal (e.g., decadal to yearly to seasonal to daily) scales. Because of their need for at least daily precipitation estimates, the impact of climate prediction uncertainties is greater for physically based models (compared to statistical models requiring climate data). However, in the northeast U.S., climate observations indicate that extreme rainfall rates are increasing more than annual precipitation rates. Statistical models that don't account for the impact of short-term rainfall rates on streamflow may under-predict future flows. The trade-offs between statistical and physically based streamflow modeling approaches are due to both mechanistic and climatic uncertainties.

4.5.2 DATA UTILIZED IN REGIONAL PEAK FLOW EQUATIONS

The Regional Peak Flow Equations (RPFEs) derived for Massachusetts (MA RPFE) do not include precipitation as an explanatory variable. However, the RPFEs developed for New Hampshire (NH RPFE) include the basinwide mean of average April precipitation, in inches, as a variable, while the RPFEs for Vermont (VT RPFE) include the basinwide mean of the average annual precipitation. The NH and VT RPFE models were included in this project due to similarities between the steeper terrain of the Deerfield River watershed and the watersheds in NH and VT from which their respective RPFE relationships were derived. Both the NH and VT RPFE equations utilize Parameter-elevation Regression on Independent Slopes Model (PRISM) climatic data for current conditions.

Multipliers drawn from the literature described in Sections 4.3 and 4.4 were applied to the PRISM data in order to estimate streamflows at mid- and end-century based on the NH and VT RPFEs. The mid- and end-century precipitation multipliers utilized for the project are summarized in Table 4-7. Mid-Century multipliers were based on the range of mid-century predictions compared to current conditions from the 9 NARCCAP models utilized for the physically based modeling (Section 4.4.2) after monthly bias adjustment. The NARCCAP based multipliers were compared to values in the literature (Sections 4.3 and 4.4.3), and adjusted as follows:

* Mid-century annual multipliers were selected towards the lower range of values based on NARCCAP data, to be more consistent with literature values,

Table 4-7: Multipliers applied to PRISM spatial data for current climate for calculation of empirical streamflow estimates.

	Mid-century (2041 – 2060)		End – century (2071 – 2090)	
	Annual	April	Annual	April
Low Estimate	1.03	1.07	1.06	1.14
Best Estimate	1.05	1.10	1.10	1.15
High Estimate	1.07	1.13	1.14	1.20

- Mid-century April multipliers were selected towards the higher range of values based on NARCCAP data, to be more consistent with literature values.

End-Century annual multipliers were determined by linearly extending (e.g., doubling) the mid-century values. However, when doubled, end-century April multipliers were significantly higher than literature values. To better capture the range of values predicted in the literature, the range of end-century values were set to span the range of 2x the mid-century low estimate and 2x the mid-century best estimate.

4.5.3   DATA UTILIZED FOR PHYSICALLY-BASED MODELS

GCM predictions are considered too coarse in terms of spatial scale for use with physically based models, and both downscaling and bias adjustment of climate predictions to the local scale are necessary. As discussed in Appendix D, dynamically downscaled climatic predictions are preferred for use in physically based models. For this project, the nine sets of GCM-RCM projections from the NARCCAP study, Section 4.4.2, were utilized as input to drive the physically-based models after bias correction (Appendix D).

4.6   Acknowledgments

Table 4-8: CMIP5 models and modeling groups data downloaded through DCHP tool and processed through the U.S. DOT CMIP Processing Tool.

Modeling Center (or Group)	Institute ID	Model Name
Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia	CSIRO-BOM	ACCESS1.0 ACCESS1.3
Beijing Climate Center, China Meteorological Administration	BCC	BCC-CSM1.1 BCC-CSM1.1(m)
College of Global Change and Earth System Science, Beijing Normal University	GCESS	BNU-ESM
Canadian Centre for Climate Modelling and Analysis	CCCMA	CanESM2 CanCM4 CanAM4
National Center for Atmospheric Research	NCAR	CCSM4
Community Earth System Model Contributors	NSF-DOE-NCAR	CESM1(BGC) CESM1(CAM5) CESM1(CAM5.1,FV2) CESM1(FASTCHEM) CESM1(WACCM)
Centre National de Recherches Météorologiques / Centre Européen de Recherche et Formation Avancée en Calcul Scientifique	CNRM-CERFACS	CNRM-CM5 CNRM-CM5-2
Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence	CSIRO-QCCCE	CSIRO-Mk3.6.0
NOAA Geophysical Fluid Dynamics Laboratory	NOAA GFDL	GFDL-CM2.1 GFDL-CM3 GFDL-ESM2G GFDL-ESM2M GFDL-HIRAM-C180 GFDL-HIRAM-C360
Institute for Numerical Mathematics	INM	INM-CM4
Institut Pierre-Simon Laplace	IPSL	IPSL-CM5A-LR IPSL-CM5A-MR IPSL-CM5B-LR
Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	MIROC	MIROC-ESM MIROC-ESM-CHEM
Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	MIROC	MIROC4h MIROC5
Max-Planck-Institut für Meteorologie (Max Planck Institute for Meteorology)	MPI-M	MPI-ESM-MR MPI-ESM-LR MPI-ESM-P
Meteorological Research Institute	MRI	MRI-AGCM3.2H MRI-AGCM3.2S MRI-CGCM3 MRI-ESM1
Norwegian Climate Centre	NCC	NorESM1-M NorESM1-ME

Modeling and Analysis

5.1 Structures

5.1.1 APPROACH

The Massachusetts Department of Transportation (MassDOT) has a well-developed methodology and program for assessing bridges, but no such assessment program had been implemented for culverts. Therefore, UMass Extension's Stream Continuity Project (now the NAACC) developed a methodology for assessing the structural condition of culverts. This ends-only rapid assessment (culverts were assessed from outside looking in at each end of the culvert) was based, to a large degree, on a FHWA publication "Culvert Assessment and Decision-Making Procedures Manual" (Hunt et al., 2010).

We used the UMass Extension protocol to assess culverted crossings and small bridges in the Deerfield River watershed and convened a technical advisory committee to use the data from those assessments to score these crossings for structural risk of failure. If information about structure condition was available from MassDOT inspection reports, these were used to assess crossings for Structural Risk; otherwise, we used data from the UMass Extension protocol.

It is important to keep in mind when using Structural Risk scores that the culvert condition assessment is a non-technical, rapid assessment protocol intended for use by trained lay volunteers and technicians. This assessment protocol is intended as a coarse screening tool to draw attention to crossings that should be assessed by qualified engineers and/or highway personnel to determine if further action is required.

5.1.2 METHODS - WHICH MODELS/METHODS WERE USED, AND WHY

Technicians under the supervision of Trout Unlimited (TU) collected field data at culverted crossings and small bridges using the UMass Extension Culvert Condition Assessment protocol. The field data form used for the condition assessments is in Appendix A. The form is formatted as a table of crossing elements for both the culvert inlet and outlet along with narrative descriptions of conditions that would result in a rating of either "poor" or "critical." Any element for which conditions didn't qualify for either of these two categories was documented as "not poor or critical," "unable to observe," or "not applicable."

Data were entered into the NAACC online database. The database was programmed to implement the Structural Risk scoring algorithm so that data were scored automatically as they were entered.

5.1.3 DATA SOURCES

Two sources of data were used to score crossings for Structural Risk of Failure:

- Data collected in the field using the UMass Extension Culvert Condition Assessment protocol and entered into the NAACC online database, and

- MassDOT bridge inspection data for bridged crossings in the Deerfield River watershed.

5.1.4 SCORING

A scoring system for converting culvert condition assessment data into Structural Risk scores was developed with the assistance of a technical advisory committee made up of Jim MacBroom, Roy Schiff and Matthew Gardner (Milone & MacBroom, Inc.) and Scott Civjan (UMass Amherst Department of Civil and Environmental Engineering).

Below is the scoring system for culverts, yielding scores ranging from zero (low risk) to one (high risk). For crossings with multiple culverts, each culvert was individually assessed and the highest (worst) score was assigned to the crossing. The scoring system assumes that the vulnerability of a given crossing is based on its weakest "link."

Structural Risk of Failure Score = max (V1, V2, V3)
V1 Structural Deficiency – Super Critical

Variables marked "Critical"	Score
Any 1 of the following	*1.0*
Cross-Sectional Deformation Excessive deformation resulting in significant reduction of available flow area, and/or extensive infiltration of soil, voids, structural failure or embankment/roadway damage	Inlet or outlet
Structural Integrity of barrel Cracks, tears, splits, bulges, holes or section loss have led to extensive infiltration of soil, structural failure, voids and embankment/roadway damage	Inlet or outlet
Footings Bottom of footing exposed and/or undercut	Inlet or outlet
Performance Problems (entire crossing) Embankment piping: settlement, deep cracks or holes in roadway or embankment outside of culvert	Entire crossing

V2 Structural Deficiency - Critical

Variables marked "Critical"	Score
Any three of the following *Any two of the following* *Any one of the following*	*1.0* *0.9* *0.8*
Invert Deterioration	Inlet or outlet or both counts as 1
Joints & Seams	Inlet or outlet or both counts as 1
Longitudinal Alignment	Inlet or outlet or both counts as 1
Footings Footing exposed with signs of cracking or breaking off of flakes or chips	Inlet or outlet or both counts as 1
Headwall/Wingwalls	Inlet or outlet or both counts as 1
Flared End Section	Inlet or outlet or both counts as 1
Apron (Critical)	Outlet
Armoring (Critical)	Outlet

V3 Structural Deficiency - Poor

Variables marked "Poor"	Score
For each of the following identified as "Poor"	*Addition of 0.1 pt. starting with 0.0 up to a maximum score of 0.7*
Invert Deterioration	Inlet or outlet or both counts as 1
Joints & Seams	Inlet or outlet or both counts as 1
Longitudinal Alignment	Inlet or outlet or both counts as 1
Cross-Sectional Deformation	Inlet or outlet or both counts as 1
Structural Integrity of barrel	Inlet or outlet or both counts as 1
Longitudinal Alignment	Inlet or outlet or both counts as 1
Footings	Inlet or outlet or both counts as 1
Headwall/Wingwalls	Inlet or outlet or both counts as 1
Flared End Section	Inlet or outlet or both counts as 1
Performance Problems	Entire crossing
Apron	Outlet
Armoring	Outlet

Two metrics from MassDOT's bridge inspection data were used to score inspected bridges: (60) "Substructure" and (61) "Channel & Channel Protection[21]." For a few crossings, metric (62) "Culverts" was also used. The MassDOT scoring system ranges from 0 (bad) to 9 (good). These scores were converted to the 0-1 Deerfield Structural Risk scoring system using the crosswalk in Table 5-1.

5.1.5   SUMMARY OF RESULTS

801 crossings were assessed for Structural Risk of Failure. The distribution of scores is presented in Figure 5-1 and Table 5-2. The distribution is distinctly bimodal. A large percentage (~80%) of crossing received relatively low scores for risk (< 0.5). Most of the remaining crossings (~18%) were scored as being of high risk (scores > 0.8). 6.4% of crossings were in the highest risk category (score = 1). The geographic distribution of Structural Risk of Failure scores is shown in Figure 5-4.

Table 5-1: Crosswalk for converting MassDOT bridge inspection scores to Structural Risk scores.

MassDOT Bridge Score	Deerfield Structural Risk Score
0-Failed	1
1-Imminent failure	1
2-Critical	0.9
3-Serious	0.7
4-Poor	0.5
5-Fair	0.3
6-Satisfactory	0.1
7-Good	0
8-Very good	0
9-Excellent	0

Only four crossings (2%) were assigned to the highest risk category using MassDOT bridge inspection data (Figure 5-2 and Table 5-3); six other crossings (3%) received moderately high scores (score = 7). In comparison, 7.8 percent of culverted crossings and small bridges evaluated using the UMass Extension protocol fell into this highest category with another 15.8 percent of these crossings receiving scores of 8 or 9 (Figure 5-3 and Table 5-4). It is not surprising that the percentage of inspected bridges is lower than for culverts given that these crossings are regularly inspected and the consequences of bridge failures are generally much higher than for failed culverted crossings. Alternatively, it could be that the rapid assessments implemented by lay technicians (UMass protocol) tend to yield higher risk scores than those conducted by trained professionals.

---

21 *Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges. Federal Highway Administration Report No. FHWA-PD-96-001. December 1995.*
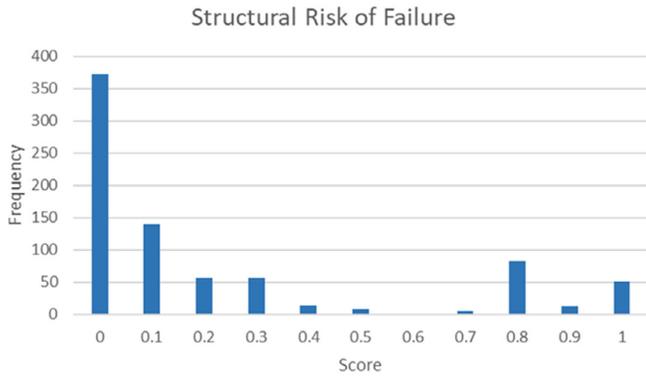
Figure 5-1: Distribution of Structural Risk scores.

Figure 5-2: Distribution of Structural Risk scores for MassDOT inspected bridges.

Table 5-2: Distribution of Structural Risk scores.

Score	Count	Percent
0	372	46.4
0.1	139	17.4
0.2	57	7.1
0.3	57	7.1
0.4	14	1.8
0.5	9	1.1
0.6	1	0.1
0.7	6	0.8
0.8	82	10.2
0.9	13	1.6
1	51	6.4
Total	801	

Table 5-3: Distribution of Structural Risk scores for MassDOT inspected bridges.

Score	Count	Percent
0	97	48.7
0.1	50	25.1
0.2		
0.3	35	17.6
0.4		
0.5	7	3.5
0.6		
0.7	6	3.0
0.8		
0.9		
1	4	2.0
Total	199	

Figure 5-3: Distribution of Structural Risk scores for culverts and small bridges assessed with the UMass Culvert Condition Assessment protocol.

Table 5-4: Distribution of Structural Risk scores for culverts and small bridges assessed with the UMass Culvert Condition Assessment protocol.

Score	Count	Percent
0	275	45.7
0.1	89	14.8
0.2	57	9.5
0.3	22	3.6
0.4	14	2.3
0.5	2	0.3
0.6	1	0.2
0.7		
0.8	82	13.6
0.9	13	2.2
1	47	7.8
Total	602	
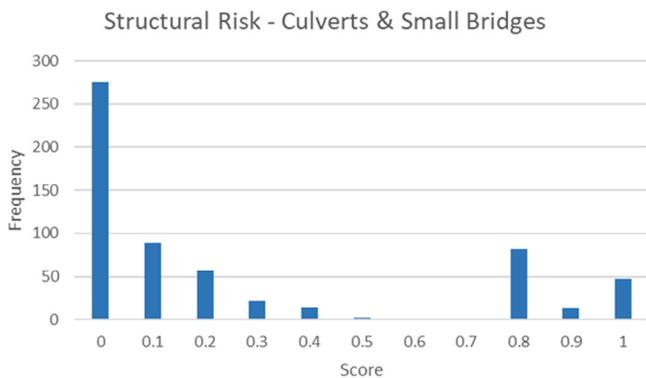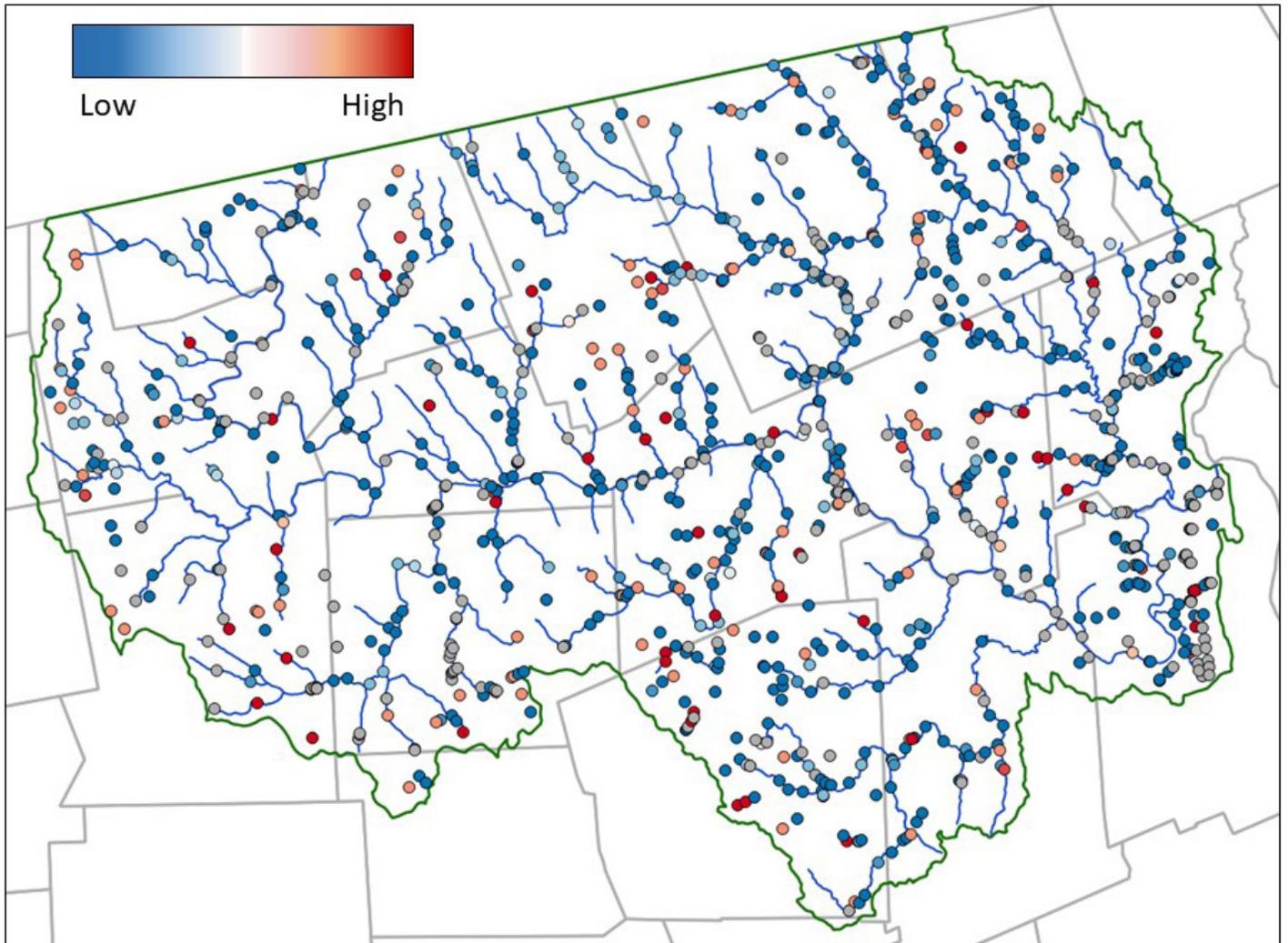
Figure 5-4: Geographic distribution of Structural Risk of Failure scores. High scores on a 0 to 1 scale are in red, low scores in blue.

5.2 Geomorphology

The critical factor in evaluating the geomorphic condition of a specific structure along a channel is determining the direction the channel is changing (aggrading vs. incising) and how the channel may respond under annual flooding and extreme events. The key is placing each structure in the context of the entire watershed as well as understanding river dynamics at each specific crossing. What happens within the river system upstream and downstream of a structure can have an influence on a structure that is just as important as the impact that the structure may have on the river.

The purpose of this task is to assess the geomorphic context of each crossing within the watershed, as well as to evaluate the vulnerability of each crossing to geomorphic stresses, and, in turn, the risk of geomorphic failure.

5.2.1 APPROACH

The geomorphic assessment was completed in two phases. Phase 1 involved a desktop, watershed scale analysis that calculated the specific stream power (SSP) for every distinct stream reach in the Massachusetts portion of the Deerfield River watershed. The SSP map produced in this desktop analysis is a stand-alone GIS-based

product that can be used to identify areas of exceptionally high energy where scour is likely and areas of low energy where deposition can occur. This map provides the regional context for each crossing and can be used as a screening tool.

Phase 2 was a more detailed local-scale assessment of the geomorphic conditions at each individual crossing. The data from this assessment were used, along with SSP, to develop a scoring system that evaluates the propensity for woody debris production at the crossing, and susceptibility of the crossing to scour or sedimentation. Other factors included in the scoring were evidence of scouring of footings, existence of a downstream scour pool, and evidence of blockage. These were combined to determine a risk of failure score due to geomorphic stress.

5.2.2 METHODS - WHICH MODELS/METHODS WERE USED, AND WHY

Milone & MacBroom, Inc. (MMI) was hired as a subcontractor to develop a vulnerability screening tool for bridges and culverts to help gauge the potential risk of failure. MMI was charged with three tasks: 1) conduct a watershed-scale stream power analysis that can be used as an indicator of erosion or deposition risk along specific stream reaches, 2) prepare a vulnerability screening tool for bridges and culverts that incorporates stream power and other physical features of the river system, and verify its suitability on a subset of 200± crossings with data describing past damage in the watershed, and, 3) check field data acquired by TU and UMass on a subset of 20 crossings for quality assurance and provide recommendations for method improvements.

The screening tool provided by MMI was then used with other geomorphic datasets collected by TU and UMass to produce a fluvial geomorphological scoring system (hereafter, scoring system) for assigning a risk of failure score to each culvert. The scoring system includes four categories, each expressing a specific geomorphic condition at a crossing that could lead to a crossing failure under extreme flooding conditions. These include the propensity for woody debris accumulation, susceptibility to sedimentation, susceptibility to scour, and evidence of blockage.

Details describing basin-wide calculation of SSP and development of the scoring system are included in Appendix B and in the "River and Stream Power Assessment Report Including Culvert and Bridge Vulnerability Analysis" prepared by MMI dated April 4, 2017 (hereafter, MMI final report), presented in Appendix E.

5.2.2.1 SPECIFIC STREAM POWER ANALYSIS

Specific stream power is sometimes used to describe the erosive power of a river because it is easier to compute than other formulas that rely on stream velocity and shear stress. While the concept of stream power is not new, this is the second major application of stream power as a tool for assessing potential risk to stream crossings in New England.

Specific stream power is defined as $\omega = (\rho g Q_2 S)/W$, where ω is specific stream power in Watts per square meter (W/m^2), ρ is the density of water (kg/m^3), g is acceleration of gravity (m/sec^2), Q_2 is the bankfull (2-year frequency) discharge (m^3/sec), S is the channel slope in m/m and W is the channel bankfull width (m).

Discharge was computed using the Jacobs (2010) equation:

$Q_2 = 0.01601 A^{0.889} P^{2.12}$ (1)

where Q_2 is the 2-year flood frequency in cubic feet per second, also assumed to be the bankfull channel forming flood, A is the drainage area in square miles and P is mean annual precipitation in inches. The Jacobs equation was developed to estimate the magnitude of peak flows for steep gradient streams in New England. Comparison of the Jacobs equation and the USGS Regional Regression equations for Massachusetts (Bent and Waite, 2013) with statistical analysis of the 2-year frequency flood at 5 gauging stations in the Deerfield River watershed showed that the USGS equations underestimate significantly the 2-year peak discharge. The Jacobs equation did a better job of predicting the 2-year peak discharge and therefore was chosen for the SSP computation (see Appendix F).

Slope was calculated using LiDAR data collected for the Hudson, Deerfield and Hoosic River watersheds following Tropical Storm Irene in 2011. Bankfull width was computed using the equation:

$W = 3.68 \, Q_2^{0.5}$ (2)

(Soar and Thorne, 2001), where W is bankfull width (m) and Q_2 is the bankfull discharge (m^3/sec). Comparisons of field measured bankfull widths along the Deerfield River and selected tributaries with hydraulic geometry relations and regime equations (at gauging stations only) show that the use of the Jacobs equation to estimate Q_2 combined with the Soar and Thorne equation is a better predictor of bankfull width when compared to field measurements than the Massachusetts statewide hydraulic geometry regression equations (2013) developed by the USGS (Bent and Waite, 2013) (see Appendix B.2.1.6). The primary reasons for this discrepancy are the steeper gradients and gravel/cobble dominated substrates typical of the Deerfield River watershed. The Massachusetts regional regression equations underestimate the field measured bankfull widths.

The SSP was calculated using equation (1) for each stream reach within the watershed. A stream reach is a section of a stream or river along which there are similar hydrologic conditions, such as discharge, depth, cross sectional area, slope, substrate type (e.g., bedrock vs. cobble) or degree of valley confinement. Stream reaches were delineated using the Vermont Stream Geomorphic Assessment protocol (Vermont Agency of Natural Resources, 2007). Within the Massachusetts portion of the Deerfield River watershed, there are 1,960 reaches.

The SSP map is shown in Appendix G.

5.2.2.2 SCORING SYSTEM DEVELOPMENT

The scoring system was developed through a trial and error process using detailed field data collected at a subset of 197 culverts, 51 of which had known damage, and other geomorphic data collected by TU and UMass Amherst. The first step was to develop a vulnerability screen. The screen was developed using a combination of previous work by MMI in New England, reviewing the literature, and plotting the distribution of damaged and undamaged structures within the watershed as a function of several geomorphic variables.

These geomorphic variables include:

1. Specific stream power versus bed resistance – It is assumed that more damage to the streambed adjacent to infrastructure and the infrastructure itself will occur as the stream power increases due to mobility of larger and more numerous particles in the water column and along the bed. However, to begin moving particles, the stream must overcome the streambed's critical shear stress. Critical shear stress is a function of particle size. MMI combines stream power and bed resistance as a proxy for degree of scour or deposition.

2. Structure Width Ratio – This is the ratio of the structure width to the channel width. The lower the ratio (expressed as a %), the more the opening to the structure constricts channel flow. The assumption is a smaller opening can promote blockage and/or ponding with subsequent overtopping or other damage.

3. Structure slope – This is the difference between the local channel slope and the slope of the structure (ft/ft). A change in slope, either because the structure is flatter or steeper than the channel, can result in deposition or erosion at either end of the structure.

4. Sediment continuity – This is the ability of the structure to transmit sediment and is based on field observations of either scour or deposition at the down or upgradient ends of the structure.

5. Structure alignment – This is the alignment of the structure to flow in the channel. Skewed alignments are assumed to produce greater erosion of embankments, which can compromise structural integrity or block structure openings with large woody debris.

Success or failure of the vulnerability screen was determined by comparing the number of damaged structures (n=51) with undamaged structures (n=146). These comparisons were made with respect to each variable and with respect to a combined, single vulnerability score. In other words, if a crossing with high stream power is more likely to experience damage, then a higher percentage of damaged culverts should occur at culverts associated with high stream power than at those with low stream power. Similarly, if skewed structures are more vulnerable to failure, a higher percentage of damaged culverts should fall in the skewed category than in the aligned category.

The second step was to combine the vulnerability screen with other condition data to develop a final scoring system. Four categories, each expressing a specific geomorphic condition at a crossing that could lead to failure, were selected. These categories include propensity for woody debris accumulation, susceptibility to sedimentation, susceptibility to scour, and evidence of blockage. Individual parameters from the vulnerability screen and condition surveys were then selected and used as components in assigning scores for the four geomorphic categories (Table 5-5).

Channel alignment and absolute structure width were considered to be the primary factors that cause blockage by woody debris. Thus, the woody debris category score was based on a combination of these two variables.

Stream power, structure width ratio and sediment continuity were considered key factors affecting the propensity for blockage via sedimentation. Literature review conducted by MMI showed that low stream power or a sudden drop in stream power causes sedimentation. Degree of deposition is a function of stream power, but not erodibility of the bed. Therefore, only the specific stream power, not bed resistance, was considered when assigning values. Accordingly, we flag crossings with stream power ≤ 100 W/m^2, low structure width ratio, or field evidence for aggradation of sediment as being at substantial risk of sedimentation.

High stream power, evidence of erosion in the field, and evidence of footing scour or a large downstream scour pool suggest a susceptibility to scour. Evidence of footing scour or a downstream scour pool came from a different dataset that included only structure condition data. The condition identifiers for the footing score were a "poor" or a "critical" judgement by the observer, which were assigned values of 3 and 4, respectively. A "Not Poor or Critical" judgment was given a score of "0." Condition identifiers for the downstream scour pool score were a "small" or a "large" judgment and given a score of 1 and 2, respectively. A "None" judgment was given a score of "0."

Blockage is provided as a separate category based on similar field evidence from the structure condition dataset. Since this category had only one factor, no averaging or weighting was necessary in assigning values to score components. As such, if blockage was considered critical, it was given a score of 1. This score takes precedence over other scores and means the crossing is highly vulnerable. If it is poor, the score is 0.5, otherwise zero. In this case, blockage is not specified in the condition field sheet and can mean blockage by woody debris or sediment.

The overall geomorphic score is determined by taking the maximum score determined in each of the four categories. Taking the average of the category scores is not valid, because a poor scour score will be cancelled out by a good sedimentation score and vice versa.

Table 5-5: Scoring system organizing the geomorphic parameters into four scoring categories: Woody debris, sedimentation, scour and blockage.

Category	Parameter
Woody Debris	Structure Alignment, Absolute Structure Width
Sedimentation	Stream Power ≤100 W/m2, Structure Width Ratio, Sediment Continuity (Aggradation)
Scour	Stream Power, Sediment Continuity (Erosion), Footing Scour Score, Downstream Scour Pool Score
Blockage	Blockage reported on the condition survey

The final values assigned in the scoring system were determined two ways. First, a series of photographs of culverts were examined by two geologists and scored independently for each category. Values were modified systematically until the computed scores were comparable to the scores estimated by the two geologists. Second, the geomorphic, structural and hydraulic risks of failure score distributions were compared. It was clear from this comparison that the geomorphic scores skewed the overall risk of failure score on the high side. This warranted adjusting the values assigned to the scoring system.

The MMI final report (Appendix E) and Appendix B provide a detailed description of how the vulnerability screen and scoring system were developed.

5.2.3 DATA SOURCES

5.2.3.1 PRECIPITATION FOR JACOBS EQUATION
The Jacobs equation requires mean annual precipitation, which was obtained from the 1961 to 1990 PRISM maps developed by the U.S. Department of Agriculture with Oregon State University. The annual map is produced by summing the 12 monthly maps. This approach accounts for orographic effects by distributing the rainfall totals spatially over the entire watershed.

5.2.3.1 AREAS FOR JACOBS EQUATION
For the Deerfield Project, one standard digital elevation model (DEM) was used for area determinations so that hydraulic calculations would be consistent across methods. Though 2-meter LiDAR and 5-meter DEM data are available, it was decided to use the North Atlantic Landscape Conservation Cooperative (NALCC) 30 m DEM. The reason for selecting this lower resolution DEM is because it is consistent with the resolution of DEMs from other parts of New England. Many other regions do not have high resolution DEMs. This will enable researchers to compare data from the Deerfield with other parts of New England using the same base map. In addition, a flow accumulation layer was already developed for this DEM that has been checked for quality assurance. Drainage areas for each reach exit point were determined using the flow accumulation layer. Using the precipitation and area calculations, the two-year discharge (Q_2) was computed in GIS for each reach and tabulated.

5.2.3.3 LIDAR DATA FOR SLOPE MEASUREMENTS
Slope measurements were derived from LiDAR imagery collected in April 2012 for the Hudson, Hoosic and Deerfield River watersheds following Tropical Storm Irene in August 2011. The LiDAR imagery was flown and processed by Northrop Grumman, Advanced GeoINT Solutions Operations Unit, Huntsville, AL. Horizontal resolution is 2 meters and vertical resolution is 0.15 meters. The horizontal datum is NAD83 and vertical datum is NAVD88. Stream reaches were laid over the bare earth model of the topography in GIS. The difference between the elevation at the start and end of each reach was calculated and divided by the length of the reach to determine the slope of the reach.

5.2.3.4 OTHER DATA SOURCES
Structure alignment (skewed vs. aligned), dominant substrate type (either qualitatively described or determined via pebble counts), streambed condition (eroding, aggrading, not aggrading or eroding), upstream bankfull width measurements (when measured) were provided by TU from field data collected on the TU Stream Crossing Assessment form. Structure dimensions, downstream scour pool condition and the structure width ratio (severe constriction, mild constriction, spans bank to bank, spans bank and channel) were obtained from data recorded by TU on the Road-Stream Crossing Inventory field data form. Finally, the condition of structure footings was obtained from data collected by TU on the MassDOT culvert condition assessment field form. These field forms are provided in Appendix A.

5.2.4 SCORING

Figure 5-5 provides the final iteration for the geomorphic scoring system. Each category (woody debris, sedimentation, scour and blockage) is scored separately. The scores are normalized between 0 and 1 by taking the total and dividing by the maximum possible score for that category. The maximum value from the category scores is advanced as the overall score for geomorphic risk of failure.

Structure width ratio (sedimentation category) was scored two ways depending on available data. If upstream bankfull widths were provided in the data set, the structure width was divided by the average bankfull width to determine the actual ratio of structure width to bankfull width, expressed as a percentage. However, if bankfull widths were not measured in the field, then the qualitative descriptions provided on the field sheets were used instead as follows:

1. Structure width/channel bankfull width (%) <50% is equivalent to culverts described as a severe constriction

2. Structure width/channel bankfull width (%) between 50 and <100% is equivalent to culverts described as a mild constriction

3. Structure width/channel bankfull width (%) between 100 and <125% is equivalent to culverts described as spanning bank to bank

4. Structure width/channel bankfull width (%) >125% is equivalent to culverts described as spanning channel and banks.

In addition, slope was not included in the scoring system. Originally, MMI included a variable in their vulnerability screen that compared the channel slope with the structure slope. However, local channel slope was not measured as a matter of course at all locations. Accordingly, the slope criterion could not be applied to all structures in the watershed.

A detailed analysis was conducted on the subset of 197 structures that MMI evaluated as part of their work to determine the effect of removing slope from their vulnerability screen (see Appendix B).

Results suggest that removal of slope from the vulnerability screen did change results some, but not significantly. Removal of slope tended to move culverts in the direction of greater risk (i.e., more conservative direction). However, there was still a clear

Woody Debris Score

Structure Alignment	
Skewed	Aligned
3	0

Absolute Structure Width (ft)					
<1	1-2	2-3	3-5	5-10	>10
4	3.5	3	2	1	0

Sedimentation Score

Structure Width Ratio			
Structure Width/Channel Width (%)			
< 50%	50 - <100%	100 - <125%	≥125%
3	2	1	0
Severe	Mild	Spans Bank to Bank	Spans Channel and Bank

Sediment Continuity				
		Downstream		
		Aggradation	None	Erosion
Uplstream	Aggradation	3	3	3
	None	2	1	1
	Erosion	2	1	1

Specific Stream Power versus Bed Resistance

		Dominant Particle Size (Bed Resistance)					
		Silt	Sand	Gravel	Cobble	Boulder	Bedrock
Specific Stream Power (W/m^2)	0-60	3	3	3	3	3	3
	60-100	2	2	2	2	2	2
	100-300	0	0	0	0	0	0
	300+	0	0	0	0	0	0

Scour

Specific Stream Power versus Bed Resistance

		Dominant Particle Size (Bed Resistance)					
		Silt	Sand	Gravel	Cobble	Boulder	Bedrock
Specific Stream Power (W/m^2)	0-60	1	1	1	1	0	0
	60-100	1	1	3	2	1	1
	100-300	1	2	4	3	2	2
	300+	1	2	4	4	2	2

Sediment Continuity				
		Downstream		
		Aggradation	None	Erosion
Upstream	Aggradation	1	1	1
	None	3	1	3
	Erosion	4	3	4

Footing Score

Not Poor/Critical	Poor	Critical
0	3	4

Downstream Scour Pool

None	Small	Large
0	1	2

Blockage

Not Poor/Critical	Poor	Critical
0	0.5	1

Figure 5-5. Geomorphic scoring system with final scoring values.

separation between the percentage of damaged culverts indicating minimal risk of failure scores and the percentage of damaged culverts indicating medium or elevated risk of failure scores. A higher percentage of damaged culverts is associated with higher risk of failure scores.

Obtaining slope data in the field is time consuming depending on the accuracy desired. It must be done using a folding rule and level, laser rangefinder or total station. For this study, MMI used a rangefinder and folding rule. Slope was evaluated by determining the slope on the upstream side of the culvert and the slope on the downstream side of the culvert (typical distance from the culvert of ~100 feet), combining those two measurements to obtain an average slope and then subtracting the structure slope from the average channel slope.

One consideration in measuring slope is determining over what horizontal distance the measurement should be taken. It matters how far upstream or downstream from the culvert one goes before making a measurement. Should the distance be 50 feet from the culvert or 100 feet, how is this chosen? In addition, it matters where in the river the measurement is made. Should it be the river bottom in the thalweg, at the river's edge, or the water surface at the edge of a cobble? The ranges used by MMI to distinguish slope categories are quite small, in some cases only 0.01 ft/ft. It can be argued that this is within the uncertainty associated with the measurements. Accordingly, if slope measurements are instituted in any protocol, a consistent, repeatable field procedure needs to be developed. The data obtained so far do not appear to justify the labor needed to collect the data in that removal of the change in slope variable does not affect the overall vulnerability score significantly. If slope is desired for any reason, reach scale slope determined from the specific stream power analysis can be substituted reasonably well (see Appendix B) and provides a much faster, cost effective means of obtaining slope.

The geomorphic scores are intended to be used as quantitative measures of relative risk. They cannot be used to predict how imminent a failure is likely to be. Crossings that score high for geomorphic risk are more likely to fail than those that have lower scores. However, reasonable predictions as to what type of storm is likely to cause the failure cannot be made. There is no basis for saying one culvert is likely to fail in the next 10 years, while another might not fail for 25 to 50 years.

5.2.5 SUMMARY OF RESULTS

A total of 811 crossings were assessed for geomorphic risk (Figure 5-6). The median score for all 811 crossings was 0.50 and the average score was 0.52. A total of 48 crossings had a score of 1.0 suggesting a greater risk of failure whereas 25 crossings had a score between 0.1 and <0.2 indicating a lower risk of failure. Of the crossings with a score of 1.0, 47 culverts showed evidence of blockage during site visits and one showed evidence of scour. Table 5-6 shows the distribution of scores by geomorphic category.

Crossings with the highest risk of failure are due to blockage. The majority of culverts with risk of failure scores between 0.6 and <1.0 are due to woody debris, followed by sedimentation and scour.

The score distributions for woody debris, sedimentation and scour are also shown as histograms (Figure 5-7, Figure 5-8 and Figure 5-9).

5.3 Hydraulics

The purpose of this section is to describe how hydraulic risk is evaluated for each crossing within the watershed, and, in turn, how the risk of hydraulic failure is scored for inclusion in the ranking scheme for determining transportation vulnerability and crossing prioritization.

For this project, Hydraulic Risk refers to the ability of a road-stream crossing to accommodate streamflow without overtopping and flooding of the road surface. Culvert "blowouts" are not explicitly accounted for, however the failure criteria for corrugated metal culverts is set to avoid conditions which have been observed to cause blowouts[22]. MassDOT is interested in identifying crossings that are vulnerable to overtopping during storm events under current climatic conditions, as well as in the future due to changes in precipitation and temperature patterns anticipated as the result of climate change. Hydraulic risk is evaluated based on the perceived ability of a road stream crossing to pass a critical flow ($Q_{critical}$), defined as the maximum flow a road-stream crossing can accommodate before potentially damaging the road subsurface or causing road overtopping. Based on these criteria for defining critical flow, the allowable headwater elevation for the project is defined as 1 foot below the road surface. Hydraulic risk determination consists of two parts: calculation of the critical flow for a given location and determination of the likelihood, relative to the other crossings in the watershed, that the critical flow will be exceeded. Both elements of hydraulic risk determination are subject to uncertainty and as such need to be well defined in the context of stakeholder needs and expectations.

This project is unique in that it aims to place into context the impact of different methodologies and changing climate on risk categorization. Questions examined include:

- What is the impact of streamflow estimation method on risk categorization?

- What is the impact of climate uncertainty on risk categorization at mid-century?

- How does the magnitude of uncertainty due to streamflow estimation method compare to the uncertainty associated with climate?

- What is the potential impact on risk categorization of assumptions made in the determination of critical flow for each crossing?[23]

---

22 MMI has observed that corrugated metal culverts are most likely to collapse and cause a blowout when piping in the soil between the road surface and culvert occurs. See criteria in 5.3.2 set to avoid such conditions.

23 This question is beyond the original scope of the project and as such is examined with less detail.

Table 5-6: Distribution of scores by geomorphic category.

Score Category	Woody Debris	Sedimentation	Scour	Blockage
0 (0 to <0.1)	218	84	17	543
1 (0.1 to <0.2)	111	61	109	
2 (0.2 to <0.3)	170	146	125	
3 (0.3 to <0.4)	97	195	238	
4 (0.4 to <0.5)	31	103	127	
5 (0.5 to <0.6)	61	162	139	35
6 (0.6 to <0.7)	65	39	35	
7 (0.7 to <0.8)	42	21	14	
8 (0.8 to <0.9)	16	0	6	
9 (0.9 to <1.0)	0	0	0	
10 (1.0)	0	0	1	47
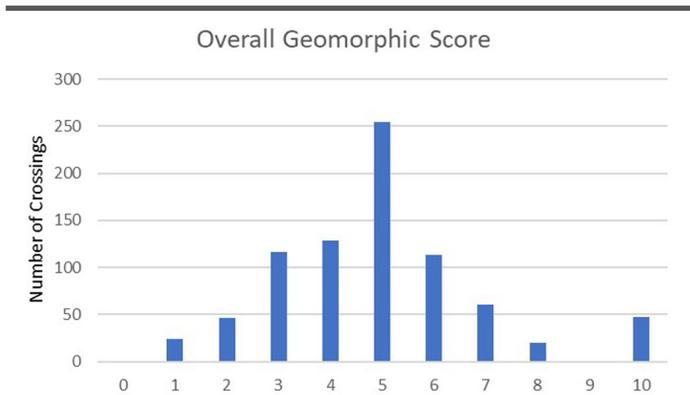
Figure 5-6: Overall geomorphic score for risk of failure.
A zero on the x-axis means a score between 0 and <0.1. A 1 means score between 0.1 and <0.2, and so on. A 10 means a score of 1.0 Y-axis is the number of crossings. N=811.
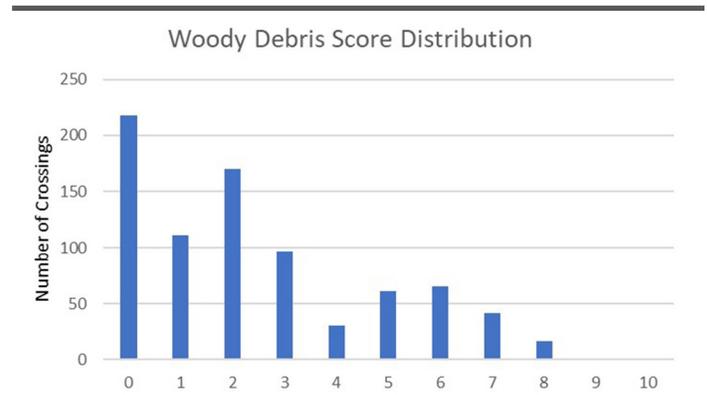
Figure 5-7: Woody debris score distribution.
A zero on the x-axis means a score between 0 and <0.1. A 1 means a score between 0.1 and <0.2, and so on. A 10 means a score of 1.0. Y-axis is the number of crossings. Average score is 0.35 and median score is 0.38. N=811.
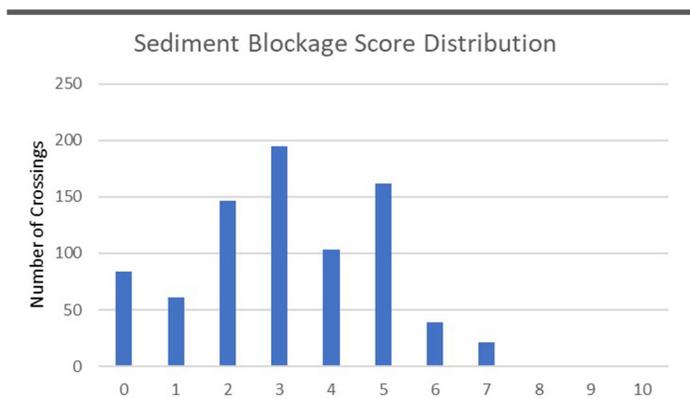
Figure 5-8: Sedimentation score distribution.
A zero on the x-axis means a score between 0 and <0.1. A 1 means a score between 0.1 and <0.2, and so on. A 10 means a score of 1.0. Y-axis is the number of crossings. Average score is 0.38 and median score is 0.42. N=811.
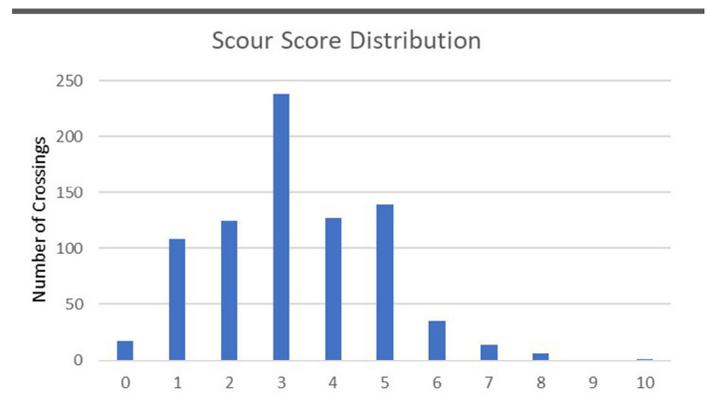
Figure 5-9: Scour score distribution.
A zero on the x-axis means a score between 0 and <0.1. A 1 means a score between 0.1 and <0.2, and so on. A 10 means a score of 1.0. Y-axis is the number of crossings. Average score is 0.35 and median score is 0.38. N=811.

This section is organized as follows to provide an overview of how hydraulic risk is categorized. The overall approach is described in Section 5.3.1. Methods utilized to calculate $Q_{critical}$ and determine the likelihood, relative to other crossings in the watershed, that critical flow will be exceeded are described in Section 5.3.2. Section 5.3.3 provides a short discussion of how the methods utilized in the Deerfield Project align with MassDOT engineering practice. Section 5.3.4 similarly describes alignment with concurrent studies. Sources of data utilized for critical flow and streamflow estimations are described in Section 5.3.5. The development of a hydraulic risk score based on the critical flow and likelihood estimates at each crossing is described in Section 5.3.6. Results are summarized in Section 5.3.7.

5.3.1   APPROACH

As noted above, for this project, Hydraulic Risk is determined based on the perceived ability of a road stream crossing to pass a critical flow. It consists of two parts, including calculation of the $Q_{critical}$ for a given location and determination of the likelihood of that flow to be exceeded relative to other crossings in the watershed.

The magnitude and probability of occurrence of streamflows in excess of $Q_{critical}$ at each crossing form the basis for assessing the likelihood of hydraulic risk of failure. This project considers four Return Interval (RI) discharges, specifically, estimates of the annual peak discharge exceeded on average once in 10-, 25-, 50-, and 100-years at each crossing (Q_{10}, Q_{25}, Q_{50} and Q_{100}). The annual exceedance probability (AEP) is defined as the inverse of the return storm year (i.e., 0.1, 0.04, 0.02 and 0.01). A variety of statistical and physically based hydrologic models are utilized to estimate flows at the 1000+ road stream crossings throughout the Deerfield River watershed.

The project utilizes a multiple model framework for estimating RI discharges, Figure 5-10. Two types of models are utilized, process-based (commonly referred to as physical or physically-based models) and statistical models (also referred to as empirical or stochastic models in the literature).

Statistically based models directly estimate the streamflow associated with a given probability of occurrence, or RI, in any given year. On the other hand, the output from physically based models consists of estimates of daily flow for each day of the simulation period. This time series of daily discharge estimates must be translated into RI estimates. To do so, during each time period of interest (e.g., current conditions or mid-century), the highest daily discharge estimated in each year of the period is utilized to develop an annual exceedance probability curve from which the discharges exceeded statistically once in 2-, 10-, 25-, and 50-, and 100-years are estimated. Because the model run periods are less than 100 years, the RI values must be statistically derived, and confidence intervals (CIs) can be determined. For this project, three estimates

for each RI were considered for the physically based models: low, most-likely, and high estimates. While some of the statistical models can also provide CI estimates, they were not included as part of this project.

Return interval discharge estimates are derived for two time-periods of interest for the project, current (typically 1971 – 1999) and mid-century (typically 2041 – 2070)[24]. The original project plan was to also include estimates for end-of-century. However, climate predictions for the end-of-century were found to be highly unreliable. Because of this, the physically based hydrologic models are run only for current and mid-century. These results may be extrapolated to end-of-century, but confidence in the results would be low.

5.3.2   METHODS - WHICH MODELS/METHODS WERE USED, AND WHY

5.3.2.1   *CULVERT FLOW CAPACITY CALCULATIONS*
Approach
The capacity of a culvert to pass flow can be controlled by conditions at the outlet as well as conditions at the inlet. For this project, estimates of critical flow were based on both inlet and outlet control. These controls were set specific to the structure, as summarized in Figure 5-11, and were set in consultation with MassDOT. Allowable headwater (HW) elevations for inlet control are summarized in Table 5-7.

- Stone masonry culverts are considered to be at risk of failure when the headwater depth (HW) becomes greater than the height of the culvert barrel (D). Thus the allowable headwater elevation for calculation of $Q_{critical}$ is set at HW/D = 1.

- Corrugated metal pipes and corrugated plastic pipes are considered to be at risk of failure if the headwater depth is 1.2 times the culvert diameter.

- Concrete culverts are designed to accommodate pressurized flow, and as such failure is considered to be the point at which the road may be subject to overtopping. For concrete culverts, the allowable headwater elevation is considered to be one foot below the road sag to avoid compromise of the road substrate materials.

- The $Q_{critical}$ for structures classified as bridges is determined based on a headwater elevation one foot below the bottom of the bridge deck. In some cases, the approach to a road stream crossing is such that passability of the structure could be compromised at a lower elevation by flow diverting around the structure and over the road surface at another point. In such cases, this sag elevation is preferentially utilized as the allowable headwater.

The capacity of a culvert to pass flow can be controlled by conditions at the outlet as well as conditions at the inlet. For culvert locations where sufficient data were available to estimate flow based on both inlet and outlet control, allowable tailwater (TW) conditions are specified based on slope as 0.75 times the culvert diameter for

---

24 *The exact years representing current and mid-century conditions varies slightly depending on the source of climate data.*
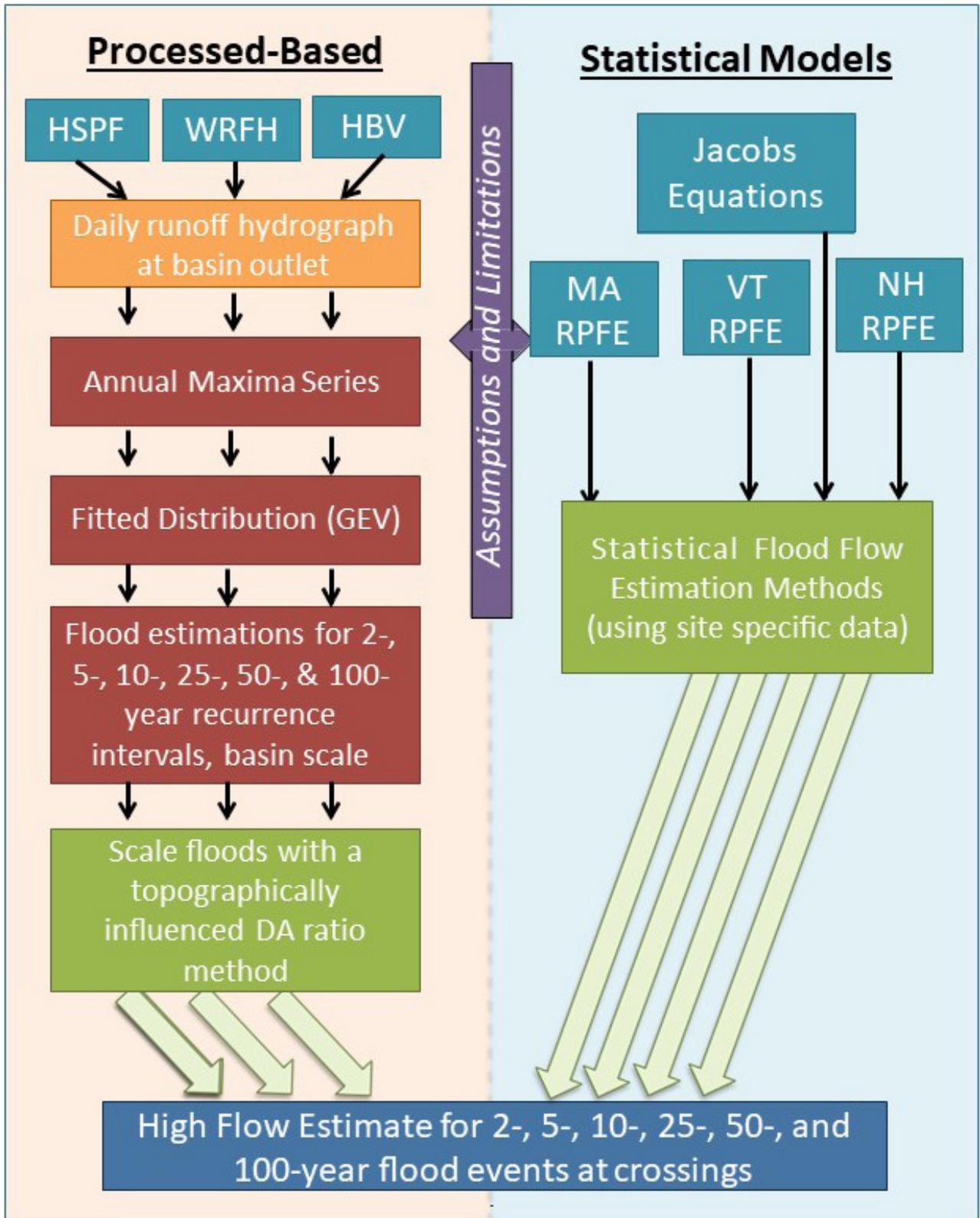
Figure 5-10: Multiple model framework (Source: Clark, 2016).

slopes > 2% and 1.0 times the culvert diameter for slopes <2%, as seen in Table 5-8

Bentley's CulvertMaster

Whenever sufficient data were available, FHWA Hydraulic Design Series 5 (HDS-5) (Schall, 2012) methodologies were utilized to determine $Q_{critical}$. HDS-5 describes equations developed for various possible culvert hydraulic conditions: inlet control, outlet control, submerged and unsubmerged. The equations for unsubmerged culverts generally apply to a Headwater to Interior Rise (culvert height) Ratio of 1, while the equations for submerged culverts apply from about a ratio of 1.5 and higher. Between ratios of 1 and 1.5, the HW depth can be approximated using a linear interpolation between the submerged and unsubmerged equations. Above a ratio of 3, the standard orifice equation for a submerged culvert under inlet control can be used. Below a value of 0.5, the headwater is not calculated, as the culvert is considered successful in passing the flow (outlet control). Each culvert type is described by the shape of the culvert, the material it is made of, the inlet edge configuration, and the inlet

end type. The commercially available CulvertMaster software by Bentley was used to implement the HDS-5 equations for the project. CulvertMaster has the advantage over other methods of being able to directly calculate flow through the culvert for specific HW and TW elevations[25]. Details on the theory and application of CulvertMaster are provided in Appendix H.

Manning's Equation

Manning's Equation was utilized to estimate culvert capacity whenever CulvertMaster could not be applied. Manning's Equation was typically applied due to size constraints (e.g., the maximum culvert dimensions allowable in CulvertMaster are exceeded) or lack of data. Manning's equation is also described in Appendix H.

The most common instance of Manning's equation implementation for critical discharge calculation was in the case of

---

25 *This is the reason that HY-8, the freely available FHWA Culvert Hydraulic Analysis Program was not utilized. HY-8 would have required interpolation of the flow from a graph generated for a range of HW and TW elevations.*
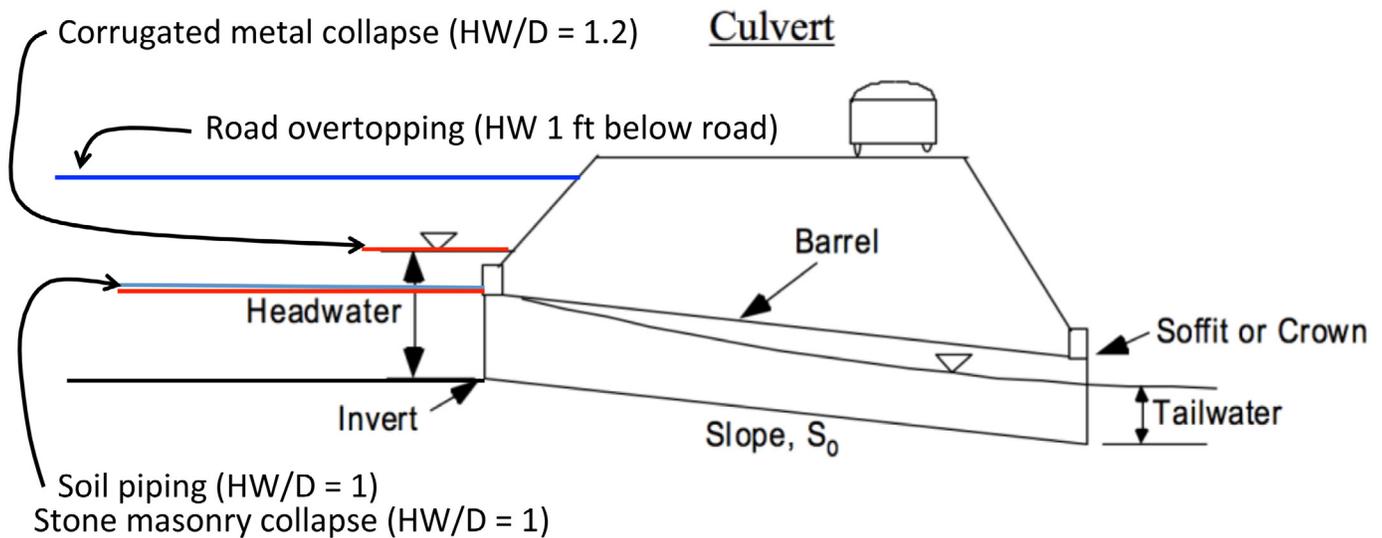
Figure 5-11: Illustration of culvert headwater elevation utilized for determining $Q_{critical}$ based on the type of material.

Table 5-7: Headwater elevations utilized in determination of $Q_{critical}$

Road Stream Crossing Type	Allowable Headwater Elevation[1]
Stone Masonry Culvert	HW/D = 1.0
Corrugated Metal Culvert	HW/D = 1.2
Concrete Culvert	HW = 1 foot below road
Bridge	1 foot below bottom of bridge deck

[1] In some cases a lower elevation in the approach to a road stream crossing was utilized instead as the allowable headwater elevation

Table 5-8: Tailwater elevations utilized in determination of $Q_{critical}$ where appropriate.

Culvert Slope	Allowable Tailwater Elevation
> 2%	TW = 0.75 D
< 2%	TW = 2.0 D

bridges with data collected by Massachusetts Geological Survey field crews, typically due to size constraints. The failure criteria for bridges was calculated based on a maximum water surface elevation of one foot below the bridge deck, maintaining open channel flow. Field crews collected data necessary to capture the most conservative cross-sectional area and wetted perimeter of the structure. In addition, an estimate of Manning's roughness coefficient was made in the field and slope data were collected. Area and perimeter were calculated based on the maximum water surface elevation of one foot below the bridge deck.

Manning's roughness coefficient was estimated for the streambed and any banks underneath the structure in the field, and field crews made note of any bridge materials that were part of the wetted perimeter. A weighted average roughness coefficient was then calculated by averaging the estimated values and values that were assumed based on material type over their respective portion of the wetted perimeter.

Multiple water surface elevations were measured in the field parallel to the stream. The two points with the longest reach were used to estimate channel slope, as long as the distance between the points was an accurate measure of the along-stream distance. If the channel was particularly sinuous on either side of the structure, points that marked the bounds of a smaller, straighter reach may have been chosen for slope calculations.

Manning's open channel flow equation was also used in cases where TU data were collected, but Culvert Master could not provide a $Q_{critical}$ solution. Generally the reason for this was that the structure had some important characteristic that CulvertMaster could not accommodate, like having opening dimensions that exceeded the software's upper size threshold. Calculating open channel flow based on Manning's requires only the Manning's roughness coefficient, cross-sectional area, hydraulic radius and slope. Cross-sectional areas and wetted perimeter were calculated up to one foot below the top of the structure in this case as well, again to maintain open channel flow conditions. Dimensions used for area and perimeter calculations were acquired from the database. Manning's roughness coefficient was based on structure material data from the condition data field sheet, and slope was based on the difference between the inlet invert elevation and the outlet invert elevation from the TU field data form.

Multiple Culvert Locations

Multiple culverts located at the same crossing were treated as separate single culverts in performing critical discharge calculations, and then their individual culvert flows were summed. In almost all cases, capacity calculations were possible with CulvertMaster, but when not feasible, the same protocol was followed as described above, and capacity was estimated instead based on Manning's equation. The primary complication in performing these calculations was finding data for the same culvert in all three field data sheets used. There is an NAACC protocol for numbering structures, and these calculations rely on this protocol being followed correctly across all field crews. The data collection field sheets, protocols and database were described in Chapter 3.

In general, individual culverts in multiple culvert crossings are extremely similar, in almost all cases having the same material and entrance type and extremely similar opening dimensions, slopes and invert elevations. As a result, our multiple culvert calculations are judged to be good estimates of critical discharge, with only slightly higher uncertainties in measurements than single culverts.

Multiple culverts where individual openings were dramatically different from each other were usually too complicated for field data collection by TU, and ended up being surveyed by Massachusetts Geological Survey field crews. The critical discharge for these crossings was calculated using Manning's Equation for each individual culvert and then summed.

Summary of Results

Originally 1,019 road stream crossings in the Massachusetts Deerfield River watershed were identified for inclusion in this project. Another 38 potential crossings were identified during the course of data collection for a total of 1,057. Of the 1,057 crossings considered, 902 road-stream crossings were identified for $Q_{critical}$ calculation. The remaining 155 potential crossings were found to either have been removed (4), were duplicates of other crossings in the system (24), were actually buried streams (7), no crossing was found upon visiting (71), or were not quantifiable because they were inaccessible both physically and visually (49).

$Q_{critical}$ values were calculated for 762, or 84.5%, of the 902 "viable" crossings. A mix of reasons was responsible for the inability to calculate critical flows at the other 140 identifiable crossings, as summarized in Table 5-9. These included inaccessibility (61), data issues (30), missing data (7), not possible to calculate (1 – bridge over dam), and unknown reasons related to complex structures (41). Some of the more common reasons for the inability to calculate a $Q_{critical}$ due to inaccessibility were that the crossing was unsafe to survey at the time of the field crew visit, under Interstate 91, or on private land or railroad owned. Common reasons for data issues included missing a critical elevation point that field crews were unable to capture.

5.3.2.2 ESTIMATION OF STREAMFLOWS AT ROAD-STREAM CROSSINGS

Three common statistical streamflow estimation methods are regression equations, index methods, and geostatistical methods. This project incorporates four regression type statistical approaches:

- USGS regional peak flow equations for Massachusetts (MA RPFE)

- USGS regional peak flow equations for Vermont (VT RPFE)

- USGS regional peak flow equations for New Hampshire (NH RPFE), and

- Jacobs equation.

The equations derived for Massachusetts (MA RPFE and Jacobs) do not include precipitation as an explanatory variable. However, the RPFEs developed for New Hampshire (NH RPFE) include the basinwide mean of average April precipitation, in inches, as a

Table 5-9: Summary of ability to calculate Q$_{critical}$ based on road stream crossings type.

Road Stream Crossing Type	Master List	Critical Flow	
		Calculated	Can't Calculate
Bridge	**286**	**272**	**14**
Bridge – no problems	272	272	0
Inaccessible	8	0	8
Bridge over dam	1	0	1
Data issues	5	0	5
Culvert	**596**	**490**	**106**
Multiple culvert	36	30	6
Open bottom arch	15	15	0
Single culvert	480	445	35
Data issues	25	0	25
Inaccessible	12	0	12
Removed	4	0	4
Duplicate	24	0	24
Other	**175**	**0**	**175**
Railroad	41	0	41
Buried stream	7	0	7
Blank	7	0	7
No crossing	71	0	71
Unknown-Inaccessible	49	0	49
Sub Total		762	295
Total	**1057**	**1057**	

variable, while the RPFEs for Vermont (VT RPFE) include the basinwide mean of the average annual precipitation. The NH and VT RPFE models were included in this project due to their inclusion of precipitation as an explanatory variable and because of the similarities between the steeper terrain of the Deerfield River watershed and the watersheds in NH and VT from which their respective RPFE relationships were derived. The USGS RPFE equations for MA, VT, and NH are briefly summarized in Appendix I and the Jacobs equation is summarized in Appendix F. Trout Unlimited was subcontracted to provide results for the MA RPFE, VT RPFE, and NH RPFE equations through their web-based tool, described in Appendix J. As an internal check for the project, UMass also ran the USGS MA RPFEs. As discussed in the results section, the UMass MA RPFE and TU MA RPFE results were inconsistent. It is believed that these differences are due to how the TU tool developed the flow accumulation grid, and then identified the location of crossings with respect to this flow grid, compared to the process used by other models in the project even though a consistent DEM was utilized.

Physically-based models mathematically represent the processes governing transformation of rainfall to runoff through overland, interflow, and baseflow with equations describing the conservation of mass and momentum. In addition, physical models simulate the accumulation of flow as streamflow, and the translation/magnification of that flow downstream. Physical models vary in their complexity, including how they represent spatial variability of the landscape. Common spatial representations, in order of increasing complexity and data needs, include lumped conceptual models, distributed hydrological response units (HRUs), and distributed grid-based discretization. Data requirements for setting up physical models include land-use, soil properties, elevation and slope data, and information about the stream channel. Climate data are necessary to "drive" the models, including at a minimum temperature and precipitation, but often additional data such as potential evapotranspiration, cloud cover, and solar energy are required. Physical models require both a calibration step, to set model parameters, and validation utilizing a separate data period, to show that the model adequately represents physical processes for a

given gauge location. This project incorporates three physical models:

- Hydrologic Simulation Program Fortran (HSPF),

- Hydrologiska Byrans Vattenbalansavdelning (HBV), and

- WRF-Hydro (WRFH).

Overviews of the three physically-based models used in the project are provided in Appendix K (HBV and HSPF) and Appendix L (WRFH). Output from the physical models consists of estimates of daily streamflow. Additional models reviewed for the project (HEC-HMS, PRMS, SWAT, VIC) were dropped either due to poor calibration (e.g., differences between predicted and observed flows in the calibration data set were unacceptably high) or other complicating factors such as data needs that could not be met.

The three physically-based models HSPF, HBV, and WRFH utilized in this project are applied at different spatial scales. The subbasin outlets and meteorological climate stations used for the HSPF and HBV models are shown on Figure 5-12 while the subbasins that form the basis of the WRFH model are shown on Figure 5-13. The locations for which model results are directly simulated by HSPF, HBV, and WRFH do not correspond with the road-stream crossings due to the data requirements needed for model development, calibration, and verification. For comparison, the locations of the road-stream crossings are shown on Figure 5-14.

5.3.2.3   ESTIMATING FLOOD FLOWS FROM DAILY STREAMFLOW ESTIMATES

While statistical models provide direct estimates of extreme flow at set RIs, the physical models instead provide estimates of average streamflow for each day of the simulation period. Estimates of the 2-, 5-, 10-, 25-, 50-, and 100-year Return Interval flood flows for the physical models were developed based on these daily average streamflow predictions. First the daily data are converted to an annual maxima series (AMS) by extracting the maximum average daily discharge for each year. A Generalized Extreme Value (GEV) distribution is then fit to the AMS in order to extrapolate flood flows from the continuous daily record. The GEV distribution is preferred over the log Pearson type 3 (LP3) distribution in the northeastern U.S. (Villarini and Smith, 2010; Vogel and Wilson, 1996). The GEV cumulative distribution function is:

$$F(x; \mu;\ \sigma;\ \xi) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \quad \text{(Equation 5.3.1)}$$

where μ, σ, and ξ are the location, scale, and shape parameters, respectively. Confidence intervals for the GEV distribution model fit can be readily predicted.

The parameters of the GEV distribution were estimated using the method of linear moments (L-Moments). Hosking and Wallis (1997) developed L-moments, which are linear combinations of order statistics. Compared to conventional moments, L-moments are considered to be an unbiased, more robust estimate less sensitive to outliers (Hosking, 1990; Kochanek et al., 2010; Millington et al., 2011; Kuczera and Franks, 2006).

The GEV distribution is sensitive to the characteristics of the
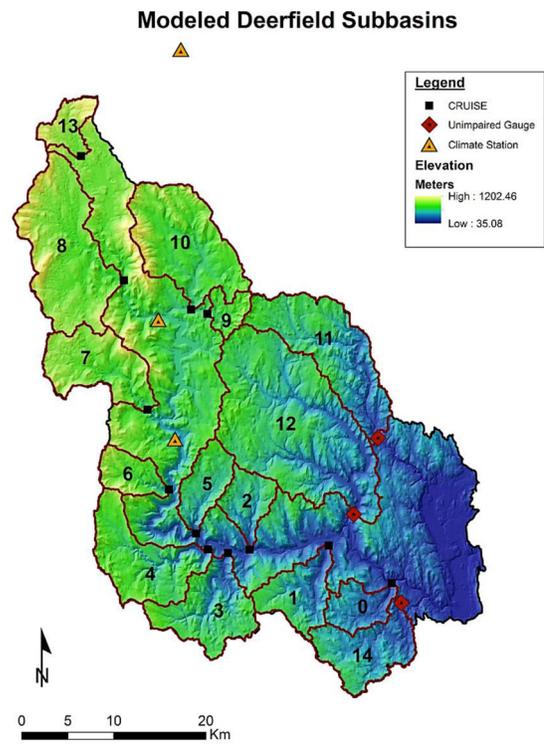
Figure 5-12: The subbasins outlets and the meteorological climate stations used in the Deerfield River watershed for the HSPF and HBV models. From Clark (2016).
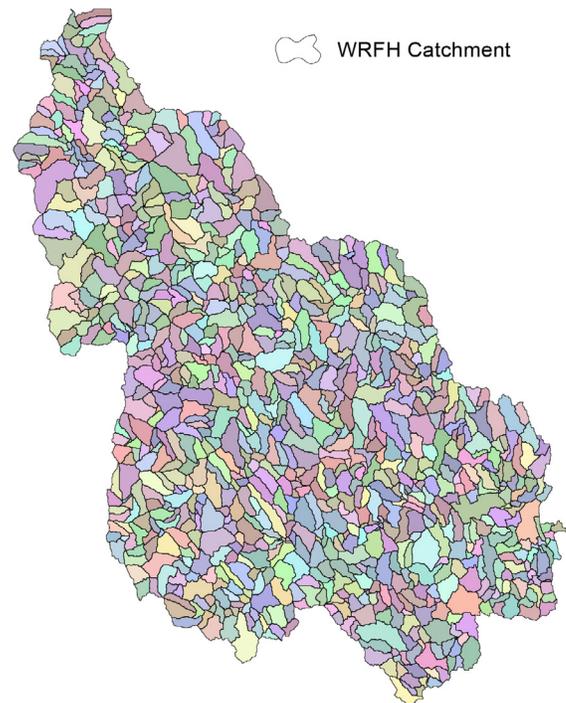
Figure 5-13: The subbasins in the Deerfield River watershed that are the basis of the WRFH model.
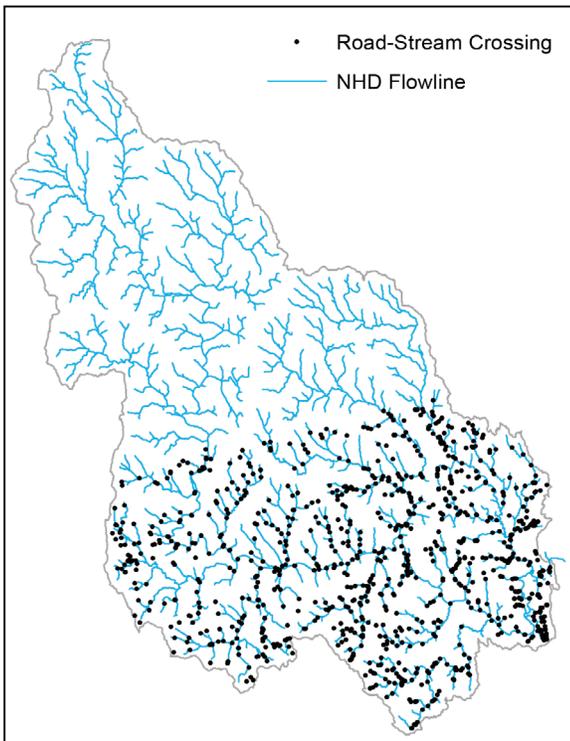
Figure 5-14: Road-stream crossing locations included in this project (Note: Only crossings in MA were assessed for vulnerability. The demarcation between MA and VT is clearly visible as no crossings are shown for VT).

historical record it is based on. Clark (2016) demonstrated this in his work for the project by splitting the historical record for the North River stream gauge in the Deerfield River watershed into two periods, 1940 – 1975 and 1976 – 2015. Figure 5-15 shows the fitted GEV distribution for the two periods (solid lines) along with the 90% confidence intervals for both curves.

This analysis is included to highlight two points:

- Uncertainty in the magnitude of high-flow events increases as the storm return period increases, and

- The streamflow historical records suggest that the magnitude and frequency of high-flow events is increasing in the watershed.

It should also be noted that the width of the confidence interval tends to decrease as the period of record upon which the RI calculations are based increases.

5.3.2.4 DOWNSCALING OF BASIN SCALE STREAMFLOW TO CROSSING LOCATIONS

As noted in Section 5.3.2.2, the locations for which model results are directly simulated by HSPF, HBV, and WRFH do not correspond with the road-stream crossings due to the data requirements for model development, calibration, and verification. A topographically influenced drainage-area scaling approach – a modified statistical index method of drainage area (DA) ratio – was developed for estimating flood flows at crossing locations within the HSPF, HBV and WRFH subbasins.
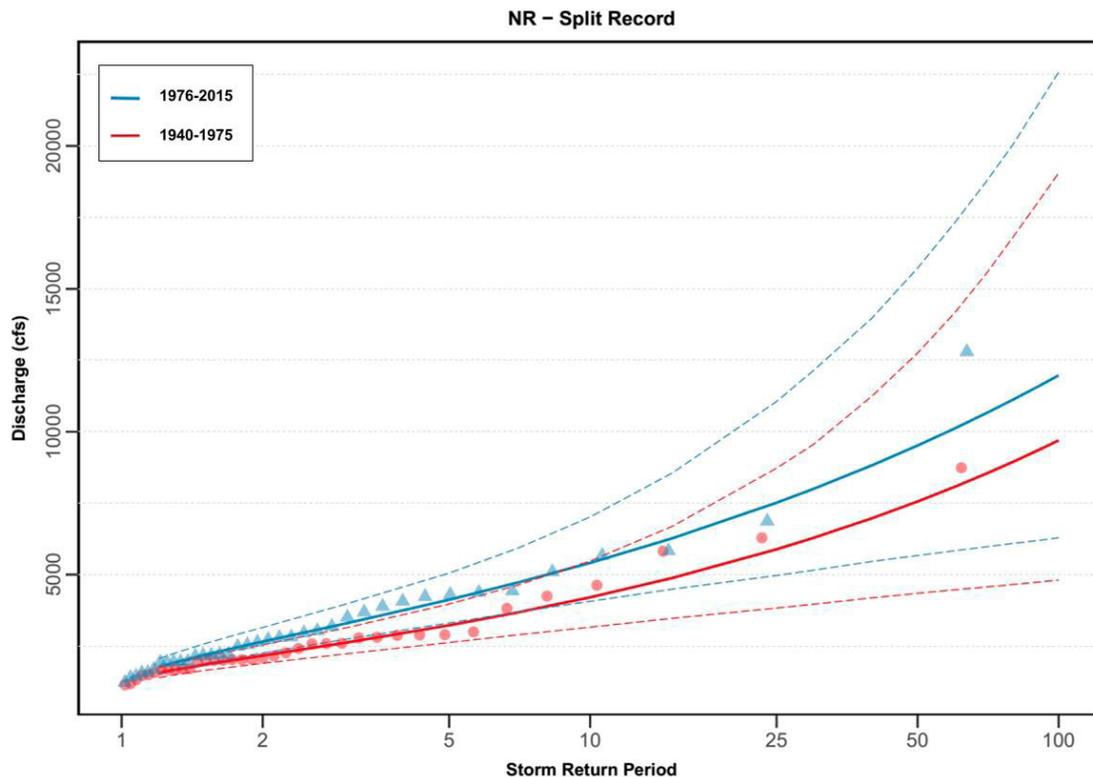
Figure 5-15: Split record at the North River USGS stream gauge (ID: 01169000). The solid lines represent the fitted GEV distribution for each half of the 75 years of record and the hashed-lines represent the confidence interval for the GEV distribution model fit. From Clark (2016).

The simplest and most common method for estimating flow between similar catchments is the index method. Index methods assume that the time series of runoff, once normalized by the mean flow, is identical between the donor catchment and the ungaged catchment. The drainage area ratio method (Stedinger et al., 1993) is one of these index methods and is the most widely used. The drainage area (DA) ratio method assumes that the runoff at the donor and recipient ungaged catchments only differ because the sizes of the drainage areas at the respective catchments are different and that for a given time the runoff per unit area at the donor and recipient catchments are equal (Stedinger et al., 1993).

The statistical index method of DA ratio was evaluated for use in estimating flood flows at road-stream crossing locations within the HSPF, HBV and WRFH subbasins. To evaluate the appropriateness of the method for the project, a set of 24 U.S. Geological Survey (USGS) gauges in the northeast to mid-Atlantic that are unimpaired, meaning they are minimally impacted by human influence, and nested, meaning they are located above (or nested within) another stream gauge on the same stream, were identified, Table 5-10. The dataset for each gauge includes 20+ complete years of discharge record since 1950. Streamflow data were also downloaded for the 24 downstream "donor" gauges.

Catchments for each stream gauge were delineated in ESRI ArcGIS (ver 10) based on the National Elevation Dataset (NED) digital elevation model (DEM) data. Physical and climate related characteristics were calculated for all of the delineated catchments including mean elevation, catchment area, mean annual precipitation, mean temperature, and mean minimum temperature. Additional calculated characteristics include land-use types from the 2011 National Land Cover Database (percent developed, forest, agricultural land), topography through DEM raster calculations (mean slope, standard deviation of slope, percent eastern aspect), hydrological variables from the National Hydrography Dataset,(total stream length, stream density), soil characteristics derived from the NRCS SSURGO data (hydrologic soil groups B, C, and D), and climate characteristics from the 2004 PRISM dataset (mean maximum, minimum, and average temperature).

These characteristics were selected based on suggested influential characteristics from the literature (Singh et al., 2014) and also because these attributes could can be readily calculated for any catchment in the entire northeast region from publically sourced GIS data layers allowing for easy translation to any other catchment.

In addition the mean topographic index (TI) was determined. The topographic index provides a numerical representation of hydrological similarity based on topographical data (Beven, 1987) and is used to derive a representation of hydrologically important topographical features in the landscape. The 'topographic index' was calculated using the following raster calculation:

$$\lambda = \ln\left(\frac{\alpha}{\tan\beta}\right) \qquad \text{(Equation 5.3.2)}$$

where λ is the topographic index, α is the upslope contributing area per unit length of contour (flow accumulation), and β is the topographic slope of the cell. The topographic index was originally developed as part of the TOPMODEL fully-distributed hydrologic model framework (Beven, 1995) which was one of the first attempts to model distributed hydrological response based on spatial and topographical patterns in a catchment (Beven et al., 2012). Singh et al. (2014) suggest that for correct process-based model parameter transfer in the northeast, the topographic index is a useful metric, which when used in tangent with soil and climate characteristics results in higher probability of successful parameter transfer.

The developed data were utilized to develop coefficients and evaluate suitability of the DA index method. The DA index method can be represented by the following:

$$Q_u = \left(\frac{A_u}{A_g}\right)^b Q_g \qquad \text{(Equation 5.3.3)}$$

where Q_u is the estimated flow statistic for the ungaged site, A_u is the drainage area for the ungaged site, A_g is the drainage area for the stream gaged site, Q_g is the flow statistic for the stream gaged station, and b value was estimated for the purposes of this study using regression from catchment attributes[26].

Optimal b-values for the 24 "pseudo-ungaged" nested catchments were determined by minimizing the Root Mean Square Error (RMSE) of the 2-, 5-, 10-, 25-, 50-, and 100-year flood flows over the historical period of record. The b-value in the DA ratio index model was treated as a single unknown parameter and a generic algorithm was used to search for the optimal b-value defined as the value that provided the lowest RMSE of the flood flows.

From the twenty-four gauges that were identified in the northeast region as being unimpaired natural flows and nested in another unimpaired catchment with overlapping records, two were removed from the subsequent analysis. USGS streamflow gauge #04233286 was removed because it was the only catchment that was located in the 04 HUC2 basin group and had b-values that were significantly different than the others in the group. For this reason, it was deemed that the catchment must have had significant hydrological landscape differences between the other gauges in the analysis, so it was removed. The USGS streamflow gauge #01643700 was also removed because the streamflow record for this gauge has been reported to be inaccurate above 500 cfs and there were other notable discrepancies in the historical record in the higher flows. The remaining twenty-two gauges, listed in Table 5-10, were carried through for the rest of the analysis.

A Kruskal analysis was performed at the 0.05 significance level to test whether the catchment characteristics are a derivative of identical populations without assuming the data have a normal distribution. The results from the Kruskal-Wallis test performed on the twenty-two catchments indicated that most of the catchment characteristics were from identical populations with p-values that were greater than 0.05. Three characteristics – mean temperature, mean minimum temperature and mean topographic index (TI) value – differed statistically across the basins as indicated by p-values that

---

26 See (Reis et al., 2005) for further discussion on the exponent b.

Table 5-10: Gauges used in DA Ratio Index scaling analysis.

USGS Station ID	HUC8 Code	Base Gage	Stream	Town	Drainage Area (sq. km)	Period of Record	DA Ratio	Overlapping Years
01022330	01050002	01022500	Narraguagus River	Deblois, ME	249.934	Sep 2002 to Mar 2007	0.425	8
01031300	01020004	01031500	Piscataquis River	Blanchard, ME	305.619	Oct 1996 to Current	0.396	20
01134500	01080102	01135000	Moose River	Victory, VT	194.767	Jan 1947 to Current	0.588	37
01135150	01080102	01135300	Pope Brook	North Danville, VT	8.417	Dec 1990 to Current	0.076	26
01137500	01080101	01138000	Ammonoosuc River	Bethlehem Junction, NH	226.883	Aug 1938 to Current	0.228	42
01174600	01080204	01174900	Cadwell Creek	Pelham, MA	1.554	Jul 1961 to Sep 1994	0.235	34
01349711	02020005	01349810	West Kill	Spruceton, NY	12.872	Oct 1997 to Current	0.184	19
01364959	02020007	01365000	Rondout Creek	Peekamoose, NY	13.882	May 1996 to Sep 2010	0.140	15
01413408	02040102	01413500	Dry Brook	Arkville, NY	212.897	Oct 1996 to Current	0.504	20
01421618	02040101	01423000	Town Brook	Hobart, NY	37.037	Oct 1997 to Current	0.043	19
01422389	02040101	01423000	Coulter Brook	Bovina Center, NY	1.968	Oct 1997 to June 2009	0.002	13
01422500	02040101	01423000	Little Delaware River	Delhi, NY	128.982	Jan 1997 to Current	0.150	40
01422738	02040101	01423000	Wolf Creek	Mundale, NY	1.580	Oct 1998 to June 2009	0.002	12
01422747	02040101	01423000	East Brook	Walton, NY	63.973	Oct 1998 to Sept 2013	0.074	16
01434017	02040104	01435000	East Branch Neversink River	Claryville, NY	59.311	July 1991 to Current	0.344	25
01434021	02040104	01435000	West Branch Neverwink River	Frost Valley, NY	1.994	Jan 1991 to Current	0.012	23
01434025	02040104	01435000	Biscuit Brook	Frost Valley, NY	9.635	June 1983 to Current	0.056	33
01510000	02050102	01510500	Otselic River	Cincinnatus, NY	380.729	Jun 1928 to Sep 1964	0.677	27
01613050	02070004	01613095	Tonoloway Creek	Needmore, PA	27.713	Oct 1965 to Current	0.096	11
01634500	02070006	01635090	Cedar Creek	Winchester, VA	264.179	Jun 1937 to Current	0.662	10
02059485	03010101	02059500	Goose Creek	Bunker Hill, VA	323.749	Dec 2006 to Jun 2014	0.665	9
0143400680	02040104	01435000	East Branch Neversink River	Denning, NY	23.129	Oct 1990 to Current	0.134	24

were less than the 0.05 threshold for the Kruskal analysis. The null hypothesis that these factors are from identical populations was rejected for these variables.

Levene's test for homogeneity of variance was calculated across the 2-digit hydrologic unit code groups in the dataset at a 0.05 significance level. The 2-digit Hydrologic Unit Code (HUC2) level is indicated by the first two digits of the HUC8 code listed in Table 5-10 and represents the largest aggregation of watersheds known as hydrologic regions. The 'New England' hydrologic region is represented by 01, 'Mid Atlantic' by 02, and the 'South Atlantic-Gulf' by 03. The results from Levene's test for homogeneity of variance looking across the three HUC2 groups represented in Table 5-10 suggested that there was no significant variance across the HUC2s with p-values also greater than 0.05. In other words, it is reasonable to group the basins despite their location in different hydrologic regions.

The optimal b-values for each catchment were then correlated to the catchment characteristics and the most significant relationships between these two assumed independent and normally distributed values were determined. Ordinary-least-squares (OLS) regression was used to relate the significantly correlated catchment physical and climate characteristics for each of the subbasins to the optimal b-values. Results from the OLS regression for the DA scaling are shown in Table 5-11. The topographic parameters for the minimum TI value and the standard deviation of the slope measures were highly correlated to the optimal b-value and serve as the independent variables in the linear regression. These variables were selected based on significant R^2 values (>0.424) with N-22 and a p-value = 0.05 significance level. Through this approach, b-values are estimated.

Goodness-of-fit measures including the R^2, NSE, KGE, and VE values were applied to the comparison of the simulated DA scaled daily flow estimations with the historical records, Figure 5-16. The DA scaling model seems to perform reasonably well. The KGE criterion has the most variance across all the locations. The mean R^2 value of 0.9 and a mean NSE value of about 0.89 suggest that scaling by the drainage area is a reasonable method for estimating the daily flows at an ungaged site.

Figure 5-17 summarizes the DA ratio method results for the set of 22 subbasins utilized in the development of the method. Estimated versus observed flood flows for the 2-, 5-, 10-, 25-, 50-, and 100-year floods for three different scenarios are shown:

- Scaling with just the DA ratio (b=1),

- Scaling with the DA ratio with a b-parameter estimated from the topographic characteristics of the catchment (Table 5-11), and

- Leave-one-out cross validation (LOOCV) estimations where a single catchment was left out, the topographic characteristics of the other catchments were used to develop the OLS regression equation for the b-parameter, then the b-parameter was estimated for the catchment that was left out and the flood flows were estimated.

Table 5-11: Results from OLS regression.

Regional DA Ratio Analysis Regression	
Parameter	b_value
min_TI	0.066* (0.030)
std_slope	-0.017 (0.025)
Constant	0.831** (0.211)
N	22
R^2	0.497
Adjusted R^2	0.444
Residual Std. Error	0.110 (df = 19)
F Statistic	9.380** (df = 2; 19)

Notes: **Significant at the 1 percent level. *Significant at the 5 percent level.
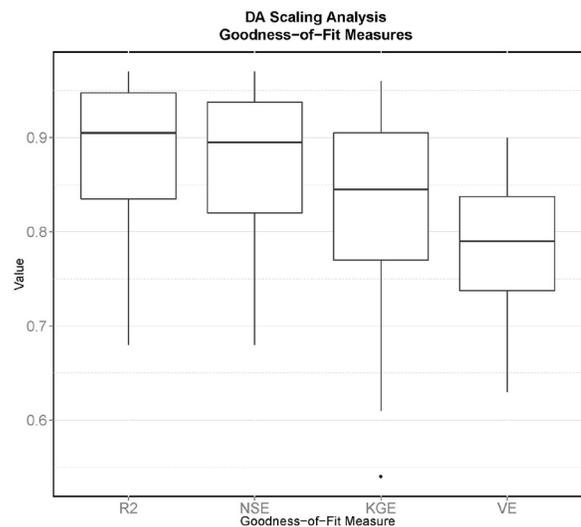
Figure 5-16: Distribution of goodness-of-fit (GOF) measures across all 22 unimpaired, "nested" catchments in the Northeastern US. GOF values were calculated at a daily time-step.

Results suggest that the DA method for scaling flood flows, with the inclusion of the topographically influenced estimate of the b-exponent, provides reliable and reasonable estimates of the 2-, 5-, 10-, 25-, 50-, and 100-year floods for ungaged sites nested within a basin for which flood flows may be directly estimated from a GEV distribution fitted to the observed (or modeled) AMS series. Variability in prediction error generally increases as the storm size increases, as indicated by the increased scatter along the one-on-one line. There is a slight negative bias, with the DA ratio tending to under-predict actual storm flows. LOOCV results suggest that the method for estimating the b-parameter in the DA ratio model is significantly robust and does a fairly consistent job at providing estimates that minimize the RMSE error of the flood flow predictions. These results suggest that using a topographically influenced b-parameter may reduce error in estimates of larger floods.

In summary, 22 unimpaired basins (Table 5-10) spanning the northeast to mid-Atlantic were utilized to develop a method for

estimating streamflows at road-stream crossing locations within the subbasins for which daily streamflow values are directly simulated by the physically-based models HSPF, HBV, and WRFH. Each of the 22 basins utilized in method development had available USGS streamflow data at their outlet as well as at an interior, or "nested" gauging site (the site listed in Table 5-10). The ratio of the drainage area of the "nested" to the "downstream" gauging site provided a reasonable scaling factor for estimating streamflow at the nested catchment, which can be improved by accounting for additional topographic characteristics (the b-parameter estimate in Table 5-11). The importance of including the b-factor increases with RI. The OLS regression equation developed for estimating the b-parameter based on topographic characteristics is robust, meaning that the equation was not more strongly influenced by data associated with any one of the 22 basins utilized in its derivation. This provides confidence that the drainage area ratio method incorporating a factor based on basin characteristics (equation 5.3.3) may be applied to develop b-factors for nested catchments within the Deerfield River watershed.

5.3.2.5 APPLICATION OF CLIMATE DATA FOR FUTURE STREAMFLOW ESTIMATION

The climate scenarios utilized for prediction of flood flows at mid- and end-of-century are fully described in Chapter 4. Application of these scenarios for streamflow estimation is briefly summarized in this section.

The RPFEs developed for New Hampshire (NH RPFE) include

the basinwide mean of average April precipitation, in inches, as a variable, while the RPFEs for Vermont (VT RPFE) include the basinwide mean of the average annual precipitation. Mid- and end-century precipitation multipliers (Table 4-7) were applied to PRISM spatial data for current climate to provide mid- and end-century RI streamflow estimates based on the NH and VT RPFEs. As discussed in Chapter 4, the multipliers were based on a range of predictions drawn from the literature as well as by the nine NARCCAP GCM-RCM predictions, bias adjusted specifically for the Deerfield River watershed as part of this project.

Nine of the dynamically downscaled projections (GCM-RCMs) available from NARCCAP were bias adjusted for the Deerfield River watershed and utilized as input to the physically based models to predict flows at mid-century [CGCM3-RCM3, CGCM3-CRCM, CGCM3-WRFG , CCSM-CRCM, CCSM-MM5I,CCSM-WRFG, GFDL-ECP2, GFDL-RCM3, GFDL-HRM3]. One of the hydrologic models, WRFH, was not able to run the CCSM_WRFG or CGCM3_WRFG projections. Daily climate predictions at end-century that were evaluated for the project were deemed too uncertain to produce meaningful end-century flood flow estimates.

5.3.3 ALIGNMENT WITH MASSDOT ENGINEERING PRACTICE

The most definitive discussion of hydrologic computational methods MassDOT currently endorses for bridge and culvert design is presented in Section 1.3 of the 2013 revision of the MassDOT Bridge Load and Resistance Factor Design (LRFD) Manual. It
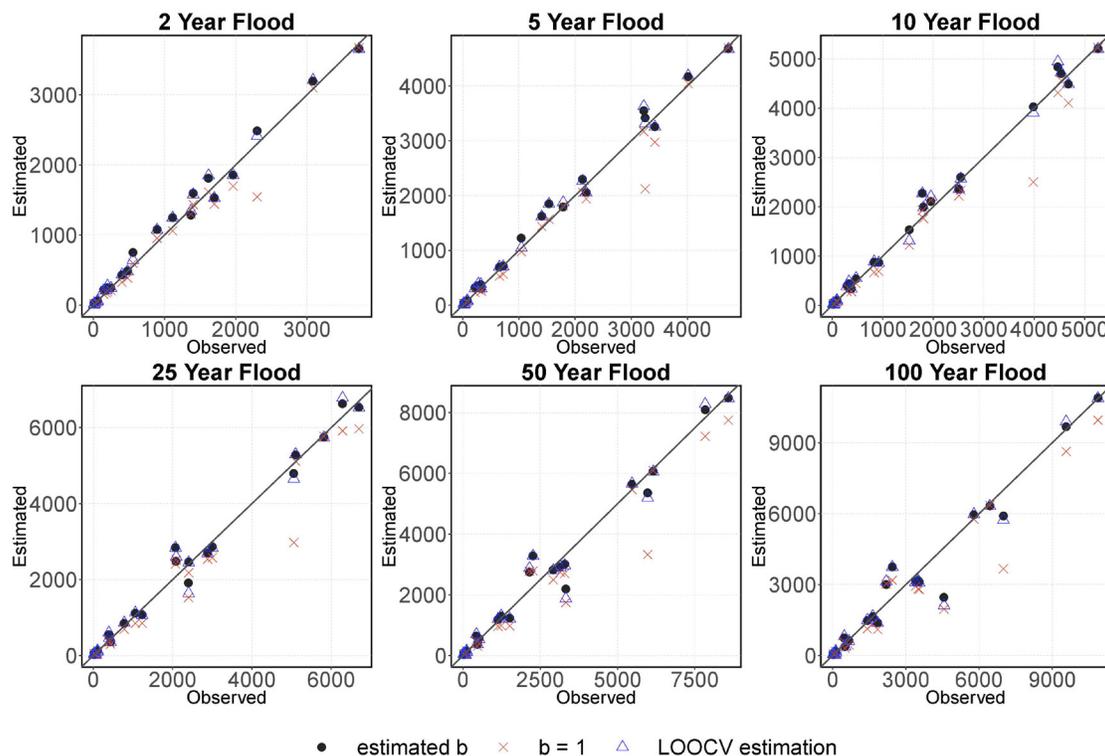
Figure 5-17: Comparison of DA ratio method for predicting flood flows for the 2-, 5-, 10-, 25-, 50-, and 100-year floods for three different b value scenarios.

should be stressed that bridge and culvert design are NOT elements of the pilot program described in this report. Rather, this project provides MassDOT with a vulnerability assessment, which they may choose to utilize when prioritizing structures for replacement or upgrade. Once a structure is identified for replacement or upgrade, it is expected that MassDOT will implement the approved computational methods for bridge and culvert design. However, the following five accepted computational methodologies were taken into consideration during the development of this project:

1. At crossings with relatively unregulated upstream watersheds, Section 1.3 recommends use of either the existing U.S. Geological Survey (USGS) Massachusetts (MA) Regionalized Peak Flow Equation (RPFE) system or NRCS TR-55 procedures. At crossings with upstream watersheds regulated by built natural impoundments or diversions of runoff flow, Section 1.3 recommends the use of the USACOE HEC-HMS Hydrologic Modeling System. While the MA RPFEs were included as one of the project's streamflow estimating methodologies, the modeling and data requirements for HEC-HMS were beyond the project scope.

2. Use of the USGS StreamStats for Massachusetts web application and/or the resources of the Massachusetts Department of Geographic Information Systems (MassGIS) is recommended to support development of watershed design variables for all the computational methods described above. StreamStats and MassGIS data were leveraged for calculation of subbasin characteristics for this project.

3. The National Oceanographic and Atmospheric Administration (NOAA) Hydrometeorological Design Studies Center (HDSC) Hydrologic Atlas 14 web application was not available at the start of this project or when the VT, NH, and MA RPFEs were updated by the states. The use of regional extreme precipitation frequency maps developed by the NRCS-funded Northeast Regional Climate Center (NRCC) was endorsed by MassDOT as input data for hydrologic model simulations in the interim.

4. Designers are directed to review the results generated by several different computational methods, apply professional engineering judgment to identify the output set that best reflect local and regional hydrologic conditions, and then document the rationale for that selection in the project's hydraulic study report.

5. Currently effective National Flood Insurance Program (NFIP) 10-, 50-, base (100-year), and 500-year flood discharges must be employed within hydraulic studies performed for proposed replacement bridges that cross waterways with either existing NFIP Regulatory Floodway delineations or published flood elevation profiles.

5.3.4 ALIGNMENT WITH RELEVANT CONCURRENT STUDIES

5.3.4.1 MASSDOT CONCURRENT STUDIES/POLICY CHANGES
Development of the climate projections for the MassDOT Statewide CCAP were completed about two-thirds of the way through the Deerfield Project. Differences between the climate data utilized by the CCAP and Deerfield Project were described in Chapter 4. It should be noted that the CCPA study utilizes revised RPFE equations released by USGS in 2017 (Zarriello, 2017). UMass was not able to obtain the revised equations in time for inclusion in this project.

The selection of exceedance probabilities utilized as the basis of the hydraulic condition categorization and risk scoring for this project was made prior to these changes. Results from this project may help prioritize crossings for update. While it is beyond the scope of the current project, it may be possible to develop an alternative hydraulic risk score based on the Q_{100}. In the current methodology, crossings with a score of zero will not successfully pass the 100-year flood flow.

5.3.4.2 COMPARISON TO HYDRAULIC RISK METHODS USED IN RELEVANT CONCURRENT STUDIES
Recent projects in New York and New Hampshire also aim to identify culverts at risk of hydraulic failure under current and future climate conditions. These are described briefly below.

Cornell's Project Culvert Capacity Calculations
Cornell University's Determining Peak Flow Under Different Scenarios and Assessing Organism Passage Potential: Identifying and Prioritizing Undersized and Poorly Passable Culverts project determines peak flow capacity under head conditions at the inlet up to the height of the road surface. To determine peak flow capacity, Cornell's method uses standard engineering equations for circular pipe flow based on the culvert's size, shape, inlet type, length, slope, and culvert material. It is assumed that maximum capacity is when the inlet is fully submerged under a water elevation equal to that of the road, but not spilling over. This leaves three flow conditions: submerged outlet control, unsubmerged outlet control, and inlet control; each with a unique equation to calculate flow. The equations utilized in the Cornell method are provided in Appendix M.

Trout Unlimited
As noted earlier, for this project TU was subcontracted to modify a web-based tool they had developed for other projects in order to provide the MA, VT, and NH RPFE equation estimates for this project. The TU tool was specifically developed for the evaluation of the potential of culverts to fail under current and future climate conditions. TU's method for assessing failure differs significantly from that utilized in this project. Specifically, rather than calculating culvert capacity based on a given HW elevation and comparing capacity against RI estimates (as done in the Cornell and UMass methods), TU estimates the HW elevation associated with each RI estimate. Culvert "status" is assigned based on the resulting ratio of HW to interior rise of the culvert. The culvert is considered "passing" if the ratio is under 0.85, the culvert is considered "failing" if the ratio is over 1.15, and the culvert is considered

"transitional" if the ratio is between 0.85 and 1.15. Output consists of a score (passing, transitional, failing) for each RI considered. TU used the same DEM as this project, but their tool developed the flow accumulation grid and identified crossing locations differently. A summary of their methods, prepared and provided by Joel Ballestero, is provided in Appendix J.

Risk Assessment Comparison

In summary, the three approaches (UMass, Cornell, and TU) for estimating hydraulic risk of failure are as follows:

UMASS

- Road-stream crossing capacity is calculated and compared against RI estimates for current and future climate scenarios,

    ○ Capacity of culverts and small bridges that function as culverts is based on HDS-5 equations considering both inlet and outlet control, calculated with CulvertMaster® software. Capacity of structures larger than can be calculated in CulvertMaster® is estimated via Manning's equation

    ○ Three HW conditions are utilized to determine capacity, selected based on the culvert material

    ○ Output consists of a $Q_{critical}$ value, the maximum predicted based on inlet or outlet control, for the HW permitted based on culvert material

- A single hydraulic risk of failure score is assigned to the crossing based on an empirical equation developed specifically for the project (described below) based on capacity and the Q_{25} RI estimate.

CORNELL

- Road-stream crossing capacity is calculated and compared against RI estimates for current and future climate scenarios,

    ○ Capacity is determined based on pipeflow equations considering inlet and outlet control. Only circular culverts are evaluated

    ○ It is assumed that maximum capacity is when the inlet is fully submerged under a water elevation equal to that of the road, but not spilling over. No consideration of possible failure due to culvert material at lower HW elevations is made

    ○ Output consists of a $Q_{critical}$ value based on the HW condition

- A single hydraulic risk of failure is determined based on the maximum return period that can be accommodated based on the calculated capacity.

TU

- HW depth associated with a range of RI discharge estimates is calculated,

    ○ HW depth for a given RI discharge value based on a parabolic equation estimate of the HDS-5 equations

    ○ RI discharge values are calculated based on one of the USGS RPFEs except for very small watersheds, where the SCS method is used

- Multiple hydraulic risk of failure estimates are identified, one for each RI,

    ○ The culvert is rated based on the ratio of the back-calculated HW elevation to the Interior Rise of the culvert

    ○ The culvert is considered Passing if the ratio is under 0.85, the culvert is considered Failing if the ratio is over 1.15, and the culvert is considered Transitional if the ratio is between 0.85 and 1.15

    ○ Output consists of a score (passing, transitional, failing) for each RI considered.

Due to the underlying differences in the three approaches, it is difficult to directly compare hydraulic risk of failure estimates. However, the UMass and Cornell approaches are similar enough that $Q_{critical}$ results based on the Cornell method may be calculated and then run through the UMass empirical equation for comparison purposes. The UMass and Cornell hydraulic risk estimates may then be compared against the TU estimates qualitatively. Although beyond the scope of the project, such a comparison is of interest and was completed for a subset of road stream crossings in Appendix N.

5.3.5   DATA SOURCES

5.3.5.1   CLIMATE DATA

Jacobs and RPFE Equations

The Jacobs equation requires mean annual precipitation, which was obtained from the 1961 to 1990 PRISM maps developed by the U.S. Department of Agriculture with Oregon State University. The annual map is produced by summing the 12 monthly maps. This approach accounts for orographic effects by distributing the rainfall totals spatially over the entire watershed.

Physical Models

Historical climate used for the physical models included precipitation, temperature, and potential evapotranspiration. A total of four stations were selected based on the time-period of record and quality of data, two of which were located in the Deerfield River watershed and two of which were adjacent to the watershed, Figure 5-18. These observations were selected based on their proximity to the watershed as well as the continuity and temporal overlap of the historical records.

5.3.5.2   STREAMFLOW

There are seven streamflow gauges in the watershed as part of the USGS National Water Information System (NWIS) network,

Figure 5-19. However, only three of these gauges are considered unimpaired (Falcone et al., 2010), as there are several major dams in the watershed. These three unimpaired gauges represent approximately 23% of the total drainage area of the Deerfield River watershed. Forest cover is the dominant land-use type in each of these catchments and there is relatively little impervious or developed landscape. The North River is the Deerfield River's largest gauged unimpaired tributary with a historical observation record of about 73 years. The Green River and South River gauges (1170100 and 01169900) have periods of record of about 49 and 48 years, respectively. The USGS streamflow gauge data were utilized for model calibration and validation as described in the appendices

5.3.5.3 AREAS AND OTHER WATERSHED CHARACTERISTICS
For this project one standard digital elevation model (DEM) was used for area determinations so that hydraulic calculations would be consistent across methods. Though 2-meter LiDAR and 5-meter DEM data are available it was decided to use the NALCC 30 m DEM. The NALCC 30 m DEM is based on the NED DEM for the Deerfield River watershed.

The reason for selecting this lower resolution DEM is because it is consistent with the resolution of DEMs from other parts of New England. Many other regions do not have high resolution DEMs. In addition, a flow accumulation layer was already developed for this DEM that has been checked for quality assurance. Drainage areas for each reach exit point were determined using the flow accumulation layer. It is important to note that the DEM was not flow-corrected. Instead, the flow accumulation grid was based on burning the 1:25k NHD streams into the flow grid. While TU utilized this DEM in their tool to provide RI flows for the MA, VT, and NH RPFEs, there are differences in methodologies between how the flow accumulation grid was determined by their tool, compared to the other models utilized in the project.

Catchments for each stream gauge were delineated in ESRI ArcGIS (ver 10) based on the NED DEM. Physical and climate related characteristics were calculated for all of the delineated catchments including mean elevation, catchment area, mean annual precipitation, mean temperature, and mean minimum temperature. Additional calculated characteristics include land-use types from the 2011 National Land Cover Database (NLCD) (percent developed, forest, agricultural land), topography through DEM raster calculations (mean slope, standard deviation of slope, percent eastern aspect), hydrological variables from the National Hydrography Dataset (total stream length, stream density), soil characteristics derived from the National Resources Conservation Service (NRCS) Soil Survey Geographic (SSURGO) database (hydrologic soil groups B, C, and D), and climate characteristics from the 2004 PRISM dataset (mean maximum, minimum, and average temperature). These characteristics were selected based on suggested influential characteristics from the literature (Singh et al., 2014) and also because these attributes can be readily calculated for any catchment in the entire northeast region from publically sourced GIS data layers, allowing for easy translation to any other catchment.
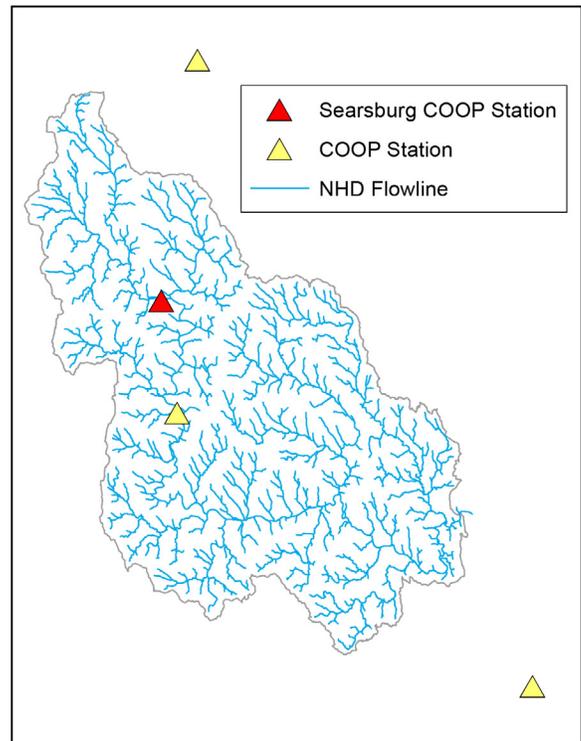
Figure 5-18: Climate stations utilized for the physical models.
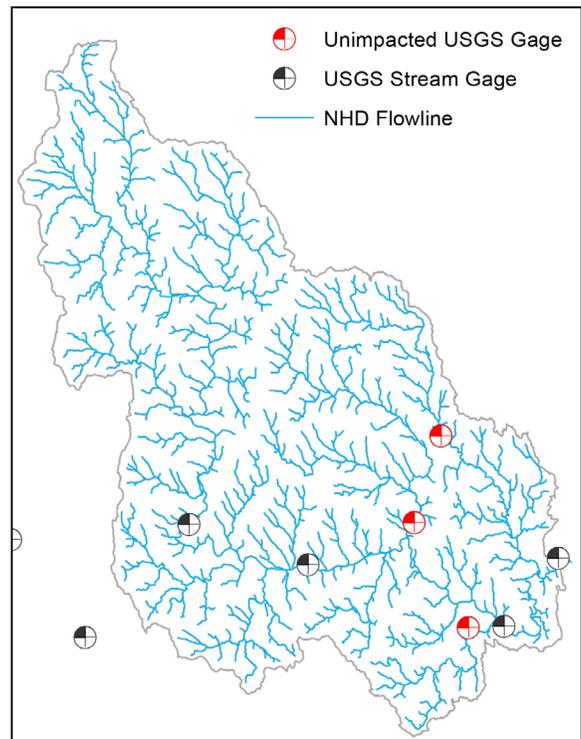
Figure 5-19: USGS streamflow gauges in the watershed.

Field data in support of the hydraulic calculations were collected by Trout Unlimited and the UMass Geological Survey as described in Chapter 3.

5.3.6   SCORING

Critical flow and streamflow estimates are combined to provide a Hydraulic Risk score for each structure as described in this section. The objective is to provide an estimate of hydraulic risk of failure that:

a.   Ranges from a low risk score of zero to a high risk score of 1,

b.   Can provide a continuous rather than discrete range of values from 0 to 1,

c.   Is consistent from crossing to crossing (i.e., a score of 0.5 at one crossing is equivalent to a score of 0.5 at another crossing), and

d.   Is roughly analogous to the scales utilized for structural and geomorphic risk of failure (i.e., a hydraulic risk of 0.7 has roughly the same relative risk of failure compared to other culverts hydraulically as a geomorphic risk of 0.7 would have).

Development of the scoring system is described in this section, including presentation of intermediate results.

RI flow estimates vary widely based on the RI, choice of model, and climate data utilized. Assessment of this variability is a critical first step for developing an effective hydraulic risk scoring system. Rather than directly compare differences in RI estimate magnitudes, the crossings were assigned a hydraulic condition category based on each model by comparing the $Q_{critical}$ against the range of RI flow estimates predicted for the crossing. For example, crossings for which the 10-year return interval is greater than $Q_{critical}$ (e.g., $Q_{critical} < Q_{10}$) are classified as "Poor," while crossings that can pass the Q_{10} but not the Q_{25} (e.g., $Q_{critical} < Q_{10}$) are classified as "Med-Low," Table 5-12.

Relative classification of the culverts across the models for current conditions is shown on Figure 5-20, with the models arranged from left to right based on the number of culverts predicted to be able to safely pass the Q_{50}. Only the median, or "most likely," results for the physical models are included on this figure. Note that both sets of MA RPFE equation results are included, one calculated by UMass and the other by TU utilizing the same set of equations. Figure 5-21 shows the classifications for mid-century for the nine climate scenarios applied to the three physically based models (HSPF, HBV, WRFH) and the three applied to the two statistical models that include precipitation as a predictor variable (TU NH RPFE, TU VT RPFE). Results under current conditions are included for comparison purposes. Figure 5-22 shows the lower, most-likely,

and upper confidence interval (CI) estimates[27] for HBV, HSPF, and WRFH under current conditions as well as for each of the mid-century climate projections utilized in the project.

Similarities and differences in results were further explored by calculating Pearson's and Spearman's (ranked) correlation coefficients for Q_{25} estimates from the eight models, Table 5-13. Rank correlations close to 1.0 suggest that while absolute values predicted by the models may differ, the same crossings tend to come out high, low, or moderate. Pearson's correlation coefficients indicate that Q_{25} estimates based on the TU VT RPFE model did not correlate well (coefficient <0.3) with any of the other models. Ranked correlations were also poor (<0.3) for the TU VT RPFE model, except for with the other two TU models, which had Spearman's correlations of 0.8 and 0.7 respectively. Spearman's correlations among the UMass models, both statistical and physical, were all >0.8, indicating that while the actual magnitude of the Q_{25} estimates predicted across the models varied, the relative magnitude of flows predicted at the crossings was very similar. Spearman's correlations between the UMass models (both statistical and physical) and the TU models were consistently <0.67. Further review suggests that the basin characteristics derived by TU to generate their RPFE estimates appear to have been off due to inconsistencies in the GIS interpretation between the road stream locations and the flow accumulation grids.

The following observations are drawn from this analysis on hydraulic condition category:

1.   Physical models (HBV, HSPF, and WRFH) identify fewer crossings as vulnerable than statistical models (MA RPFE, VT RPFE, Jacobs, TU MA RPFE, and NH RPFE).

   -   The percent of crossings in the worst condition category for physical models range from 8.4 to 11.1 percent; for statistical models the range is 22.8 to 56.5 percent.

   -   There are no failure data available for validation in order to decide which is more accurate. Differences between the statistical and physical models are explored further in the next section, where the hydraulic risk score is developed, as well as in Chapter 8.

2.   Within the statistical models there is considerable variability, with UMass model (MA RPFE and Jacobs) results more conservative, assigning 22.8 and 34.8 % of crossings to the worst condition category, and TU model results assigning 46.1 and 56.5 % to the worst category.

---

27 *The CI are derived statistically. There is 90% confidence that the actual flow associated with a given RI will fall between the .05 (lower) and .95 (upper) CI. The .05 CI is analogous to a 95% exceedance probability, the 0.5 CI to a 50% exceedance probability, and the 0.95 CI to a 5% exceedance probability.*

Table 5-12: Hydraulic condition category.

Hydraulic Condition Category						
Poor	Exceedance = 0.1 10 year RI	**Medium-Low**	Exceedance = 0.04 25 year RI	**Medium-Good**	Exceedance = 0.02 50year RI	**Good**
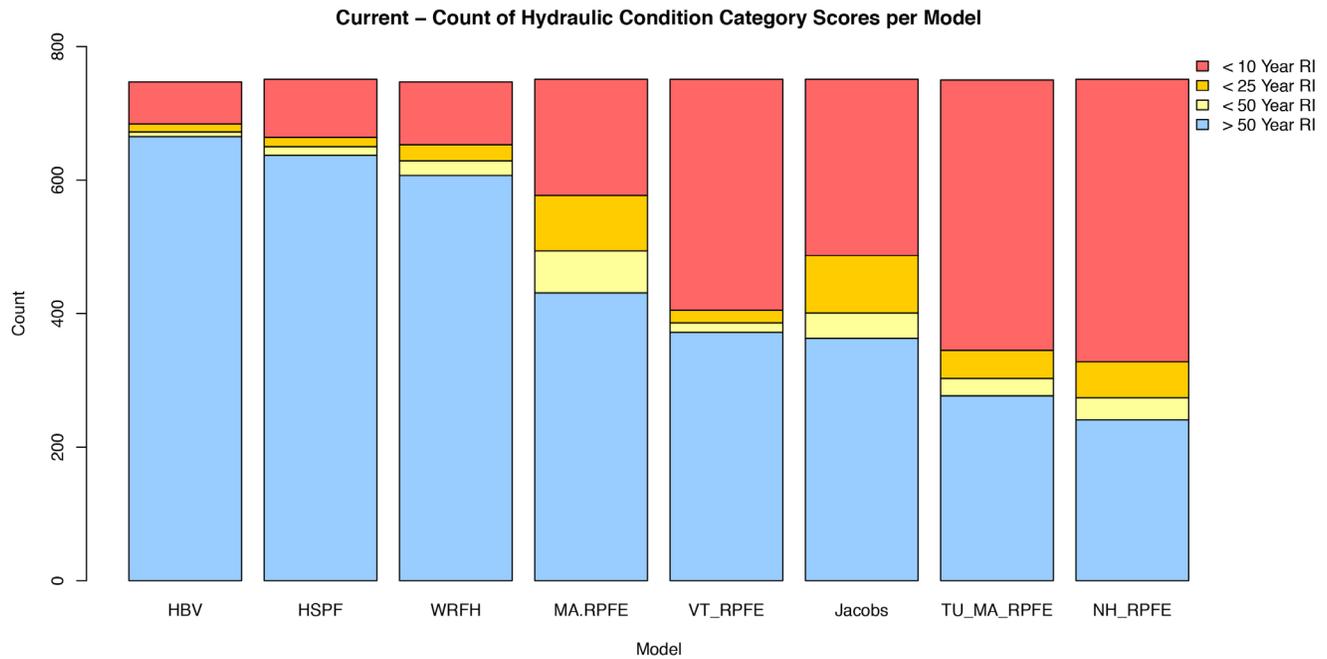
Figure 5-20: Hydraulic condition category classification across models for current conditions (only "most likely" results for physical models shown).
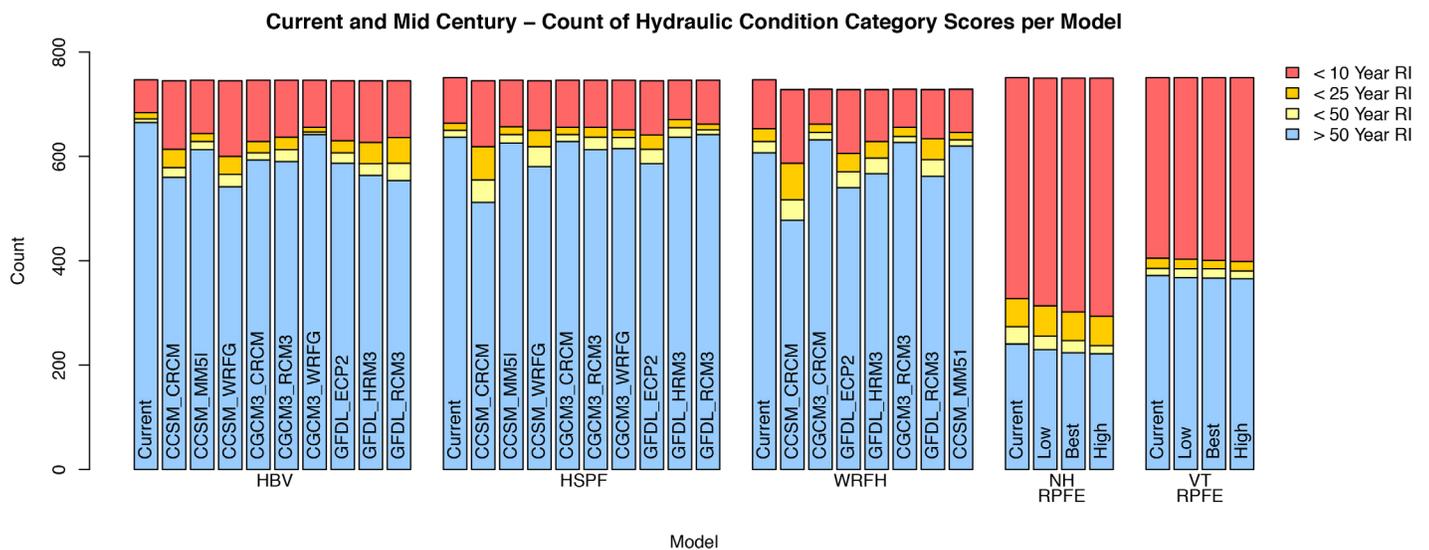
Figure 5-21: Hydraulic condition category classification across models for range of mid-century climate predictions (only "most likely" results for physical models shown; current condition categorization shown for reference).
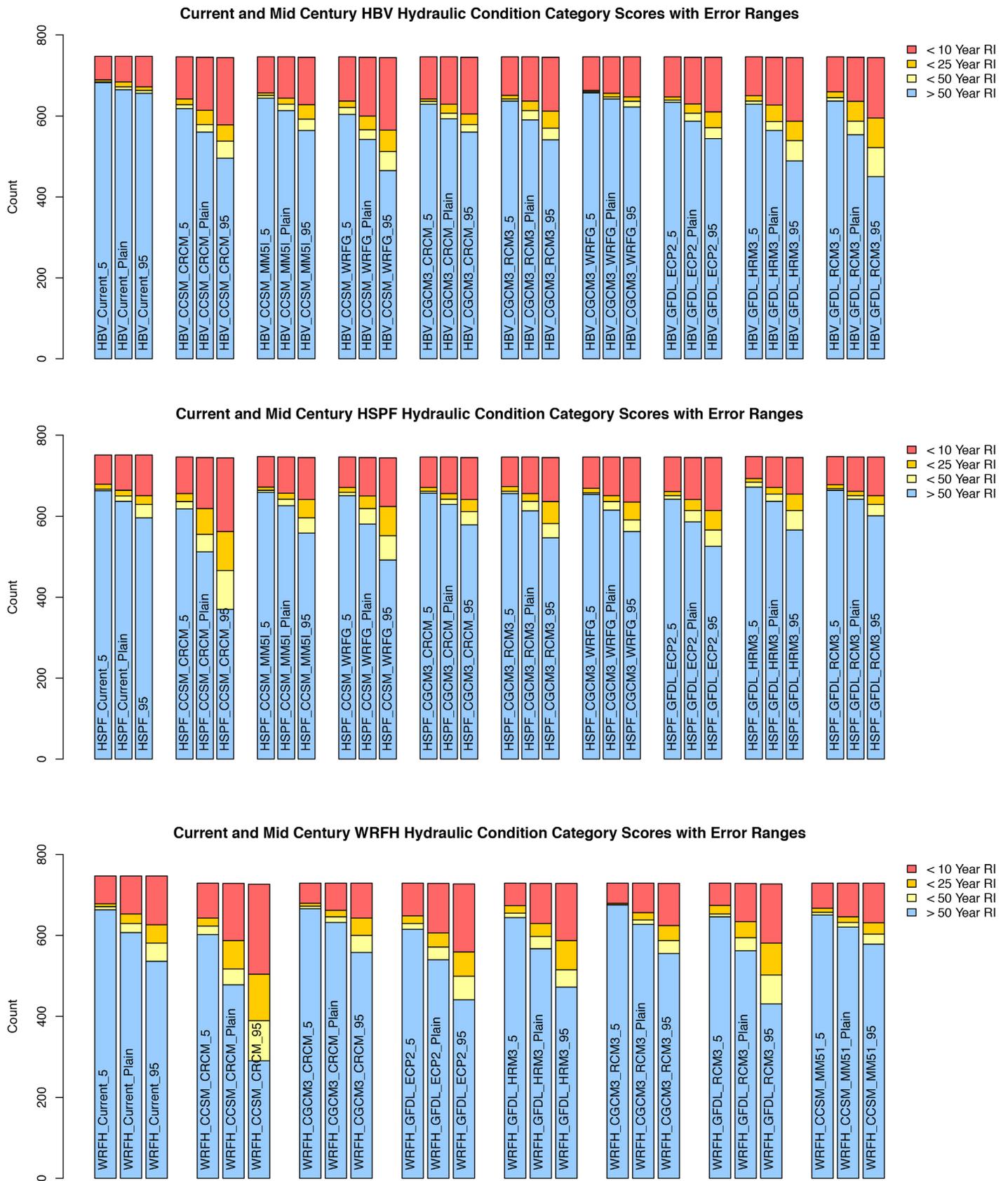
Figure 5-22: Hydraulic condition category classification for physical models HBV (top), HSPF (middle) and WRFH (bottom) (low, best, and high estimates; current and future conditions).

3. Validity of the TU model results is questionable.

   - The TU estimates result in very high percentages of crossings being assigned to the worst condition category for all three TU models (MA RPFE, NH RPFE and VT RPFE). These don't seem credible; roughly half of crossings are not expected to be able to handle a 10-year storm. If that were true, many of these crossings would have failed by now.

   - The TU model results for the MA RPFE equations differ significantly from the UMass model results for the MA RPFE equations. The UMass estimates are deemed more credible for two reasons: (a) They correlate better with other model results and (b) the UMass estimates result in a more reasonable distribution of risk of failure based on current conditions.

4. For the physical models, the within-model variability (5 to 95% confidence interval results) is comparable to the among-model variability under current conditions. The range between the 5 and 95% CI results (e.g. CI uncertainty) is greater for WRFH. Under current conditions, it seems unlikely that the choice of CI for creation of ensemble scores will make much difference if values are averaged (e.g., average central results, average the 5% results, average the 95% results, or average all three across the three models).

5. CI uncertainty increases for the mid-century results and in some instances results in greater uncertainty than that due to the climate projections.

6. Because there are unequal numbers of physical and statistical models that can be combined, it is best to average/combine the results separately for physical and statistical models, and then combine (average) the mean statistical and mean physical results for combined scores. TU estimates should be considered separately and not included in any ensemble scores to avoid skewing results.

7. One of the more interesting observations is that, for all models, very few crossings fall into the intermediate categories for condition: 19.5% (MA RPFE), 16.7% (Jacobs), 11.6% (NH RPFE), 6.3% (WRFH), 4.4% (VT RPFE), 3.6% (HSPF) and 2.5% (HBV). This suggests that the differences among the different return storms (10, 25, 50) are relatively small compared to the uncertainties inherent in the models.

Table 5-13: Correlations across model estimates of Q_{25}.

Q_{25} Pearson's Correlation Coefficients

	Jacobs	MA RPFE	TU MA RPFE	TU NH RPFE	VT RPFE	HSPF	HBV	WRF Hydro
Jacobs	1							
MA RPFE	0.999468	1						
TU MA RPFE	0.923034	0.922443	1					
TU NH RPFE	0.970641	0.971716	0.899662	1				
TU VT RPFE	0.273428	0.277092	0.111968	0.204172	1			
HSPF	0.998748	0.998411	0.924079	0.969784	0.271192	1		
HBV	0.998748	0.998411	0.924079	0.969784	0.271192	1	1	
WRF Hydro	0.936739	0.939077	0.856159	0.888243	0.297072	0.93531	0.93531	1

Q_{25} Spearman's Correlation Coefficients

	Jacobs	MA RPFE	TU MA RPFE	TU NH RPFE	VT RPFE	HSPF	HBV	WRF Hydro
Jacobs	1							
MA RPFE	0.985077	1						
TU MA RPFE	0.668335	0.660345	1					
TU NH RPFE	0.623875	0.630661	0.938446	1				
TU VT RPFE	0.257679	0.243576	0.814585	0.69386	1			
HSPF	0.97438	0.951311	0.667385	0.612503	0.278	1		
HBV	0.97438	0.951311	0.667385	0.612503	0.278	1	1	
WRF Hydro	0.875117	0.849206	0.658882	0.545681	0.321946	0.870233	0.870233	1

5.3.6.2 BASE HYDRAULIC RISK SCORING SYSTEM

The hydraulic condition classification results suggest that a scoring system based on these four categories is not sufficient to meet the objectives for the risk of failure score, as it does not give a continuous rating. Because of the tendency of the results to fall in either the poor or good categories, the ratio of $Q_{critical}$ to Q_{25} was evaluated as a potential alternative basis for a scoring system. From a logic standpoint, developing a scoring system around the Q_{25} makes sense. Q_{25} is the RI estimate demarking the "middle" of the hydraulic condition categories. Road-stream crossing structures tend to be designed for a 10 to 50-year life span, so 25 captures a middle ground.

Another advantage of a continuous scoring system based on the ratio of $Q_{critical}$ to Q_{25} is that it better accommodates the uncertainty that is inherent in estimating hydraulic risk. This crossing assessment project relies on field data collected using a rapid assessment approach that covered over 1,000 crossings in the Deerfield River watershed. These assessments do not include the type of detailed surveys of crossing dimensions and stream elevations that would allow precise calculations of hydraulic capacity. Further, nearly all crossings are on ungaged streams, requiring flows to be estimated using models. The uncertainty around the estimates of flow and hydraulic capacity makes it difficult to have high confidence in binary determinations as to whether a crossing will fail for certain sized storms. However, the magnitude of the ratio of $Q_{critical}$ to Q_{25} provides a relative probability of failure (e.g. a ratio of 1.25 is less likely to fail than a ratio of 0.8) that is normalized and comparable across different crossings. Over time, as data about culvert failures accumulate, a better sense as to what scores are truly associated with high risk will be developed. The scoring system will continue to be useful independent of the threshold utilized for denoting high risk, such as scores > 0.6 or scores > 0.8.

A logistic equation is utilized to provide a range of hydraulic risk values from a low score of zero to a high score of 1 as well as to provide a continuous (versus discrete) range of hydraulic risk values from 0 to 1 for the scoring system. The equation assigns a value of 1.0 for the ratio of $Q_{critical}$ to Q_{25} as the midpoint of the curve. Comparisons between Q_{25}, Q_{50} and Q_{100} estimates from the various models indicate that Q_{100} is never more than 2.2 times the Q_{25} (Table 5-14 and Table 5-15).

Based on these results, the steepness of the logistic function curve was set such that crossings with a $Q_{critical}$ twice the Q_{25} have essentially no risk of hydraulic failure. The final formula for calculating the hydraulic risk of failure score is:

$$Hydraulic\ Risk = \frac{1}{1+e^{6\times(\frac{Qcritical}{Q25}-1)}} \qquad \text{(Equation 5.3.4)}$$

where the ratio of $Q_{critical}$ to Q_{25} represents hydraulic condition. With this equation, a ratio of $Q_{critical}$ to Q_{25} close to zero (i.e., $Q_{critical}$ is very small compared to the Q_{25}) gets a score close to 1.0 (i.e., highest risk) while a ratio of $Q_{critical}$ to Q_{25} close to 2 (i.e., can not only pass Q_{25} safely but also in most instances Q_{100}) gets a score close to 0.0
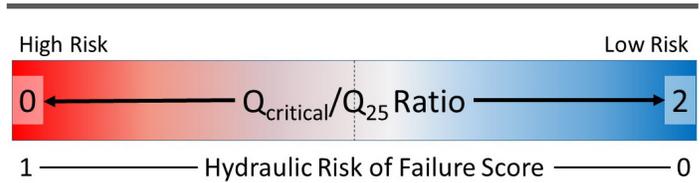
Figure 5-23: Visual of how the ratio $Q_{critical}$ /Q_{25} translates into a hydraulic risk score, and in turn how that translates into high (red) or low (blue) hydraulic risk of failure.

(i.e., lowest risk). A visual of how the ratio $Q_{critical}$/Q_{25} translates into a hydraulic risk score, and in turn how that translates into high or low hydraulic risk of failure, is provided in Figure 5-23.

Hydraulic risk scores for the individual models are shown in Figure 5-24 and Figure 5-25 for current and mid-century, respectively, based on the 0.5 CI level estimates of Q_{25} for the physical models. The data presented in Figure 5-23 are also presented in Table 5-16. Figure 5-26 shows the low-, most-likely, and high-CI hydraulic risk scores for HBV, HSPF, and WRFH under current conditions as well as for each of the mid-century climate projections utilized in the project. Table 5-17 summarizes variability in the shift to higher risk of hydraulic failure at mid-century from current conditions across the 9 climate projections, 3 models, and 3 CIs. This is summarized as the *difference* in the percent of culverts that fall in the lowest risk category. A positive number indicates fewer culverts fall in the lowest risk of failure category (i.e., a shift to higher risk) while a negative number indicates more culverts fall in the lowest risk of failure category (i.e., a shift to lower risk). Table 5-18 focuses only on mid-century, again summarizing the *difference* in the percent of culverts that fall in the lowest risk category. The following observations may be drawn:

1. While the logistic equation utilized for the hydraulic risk score does produce a continuous range of scores between 0.0 and 1.0, the largest number of culverts still score in either the lowest (i.e., 0 to 0.1) and highest (i.e., 0.9 to 1.0) risk category. Refer, for example, to Figure 5-24, as indicated by the predominance of blue (low) and red (high) scores. A wider range of scores is predicted by two of the UMass generated statistical model (MA RPFE and Jacobs) results (e.g., more scores in the 0.1 to 0.9 range).

2. The difference between TU results and the other models is even more apparent based on the hydraulic risk classification system (Figure 5-24). Table 5-19 presents the Pearson's and Spearman's correlation coefficients between the model risk scores, emphasizing how different the TU results are compared to the other models.

3. As noted previously for condition category, scores for the physically based models are more conservative while those for the two statistical models[28] predict higher levels of risk (Figure 5-24); an ensemble score that combines the two types

---

28 *Refer back to Section 5.3.6.1 for further discussion on why the MA RPFE and Jacobs statistical results are considered more reliable than the TU results.*

Table 5-14: Summary of Q_{50} to Q_{25} ratios.

	High Bound	Plain Bound	Low Bound	All CI Estimates	All Estimates	No TU
Ratio	*Frequency*	*Frequency*	*Frequency*	*Frequency*	*Frequency*	*Frequency*
0.6 - 0.8	0	0	0	0	0	0
0.8 - 1.0	0	0	0	0	0	0
1.0 - 1.2	441	995	1004	947	21	25
1.2 - 1.4	557	9	0	57	982	979
1.4 - 1.6	6	0	0	0	16	0
1.6 - 1.8	0	0	0	0	0	0
1.8 - 2.0	0	0	0	0	0	0
2.0 - 2.2	0	0	0	0	0	0
2.2 - 2.4	0	0	0	0	0	0
>2.4	0	0	0	0	0	0

Table 5-15: Summary of Q_{100} to Q_{25} ratios.

	High Bound	Plain Bound	Low Bound	All CI Estimates	All Estimates	No TU
Ratio	*Frequency*	*Frequency*	*Frequency*	*Frequency*	*Frequency*	*Frequency*
0.6 - 0.8	0	0	0	0	0	0
0.8 - 1.0	0	0	0	0	0	0
1.0 - 1.2	0	95	1004	0	0	0
1.2 - 1.4	290	891	0	864	3	8
1.4 - 1.6	534	18	0	124	661	996
1.6 - 1.8	148	0	0	16	324	0
1.8 - 2.0	28	0	0	0	28	0
2.0 - 2.2	4	0	0	0	3	0
2.2 - 2.4	0	0	0	0	0	0
>2.4	0	0	0	0	0	0

of models would be beneficial to provide a more balanced prediction of risk. Although no data are available to validate the absolute risk scores in order to decide which models are most accurate, the relative risk of failure predicted by physical and statistical (Jacobs and UMass MA RPFE only) models is similar (Table 5-19).

4. The two statistical models which can be utilized[29] to provide mid-century estimates (NH RPFE and VT RPFE) predict just slight increases in the number of crossings at high risk under future climate scenarios; differences between the low, best, and high estimates of precipitation at mid-century resulted in only minor shifts in scores (Figure 5-25).

5. In general, the physically based models predict a shift to higher hydraulic risk of failure at mid-century compared to current conditions (e.g., fewer scores <0.1, or blue, and more > 0.9, or red. However, the results vary considerably across hydrologic and climate models Figure 5-25 and Figure 5-26).

a. The CCSM_CRCM climate projections produce the largest shift towards increasing hydraulic risk (more scores 0.9 – 1.0) for the HSPF and WRFH models, and the second largest shift for HBV (Figure 5-25).

b. Some of the model - CI - climate projection combinations result in a slight increase in the number of crossings in the lowest risk category (< 0.1), as indicated by a negative number in Table 5-17.

---

29 *They include rainfall as a predictor variable*

Table 5-16: Hydraulic risk scores for current conditions in tabular format, shown based on the 0.5 CI level Q$_{25}$ estimates for the physical models. Values shown are percent of total scores in a given category.

Model	HBV	HSPF	WRFH	MA RPFE	Jacobs	TU VT RPFE	TU MA RPFE	TU NH RPFE
0.0 – 0.1	86.21	81.12	80.17	54.29	44.31	38.88	36.18	28.30
0.1 – 0.2	1.47	2.68	1.31	4.69	3.35	1.44	1.34	1.24
0.2-0.3	0.80	1.34	1.90	2.95	2.41	0.32	0.53	1.51
0.3 – 0.4	0.40	0.67	1.31	1.61	1.87	0.80	0.93	1.79
0.4 – 0.5	0.94	1.07	0.73	2.68	1.61	0.16	1.2	1.51
0.5 – 0.6	0.80	0.80	0.73	2.41	2.81	1.12	1.47	1.65
0.6 – 0.7	1.2	0.94	1.31	2.82	3.08	0.48	1.07	1.79
0.7 – 0.8	0.94	1.20	2.04	3.75	5.22	0.48	2.00	2.06
0.8 – 0.9	1.87	2.14	2.77	5.63	6.96	1.76	1.60	3.57
0.9 – 1.0	5.35	8.03	7.73	19.17	28.38	54.56	53.67	56.59

Table 5-17: Summary of relative impact of model, CI and climate projection at mid-century on change in risk from current to mid-century as summarized by differences in the percent of culverts that fall in the lowest risk category (<0.1).

Model	Current – Mid-Century	CI		
		0.05	0.50	0.95
HBV	Maximum	12.43	19.13	26.23
	Minimum	3.62	5.37	7.49
HSPF	Maximum	8.29	15.64	25.33
	Minimum	-1.97	-1.88	-.51
WRFH	Maximum	8.87	17.27	28.57
	Minimum	-1.02	-1.66	-3.91

Table 5-18: Summary of relative impact of model, CI and climate projection at mid-century as summarized by differences in the percent of culverts that fall in the lowest risk category (<0.1).

Model	Max Difference across CI Range at Mid-Century	Dif in Max and Min risk across 9 climate scenarios at Mid-Century for given model and CI	
HBV	19.4	8.8	@ 0.05 CI
		13.8	@ 0.50 CI
		18.7	@ 0.95 CI
HSPF	26.6	10.3	@ 0.05 CI
		17.5	@ 0.50 CI
		25.8	@ 0.95 CI
WRFH	36.2	9.9	@ 0.05 CI
		18.9	@ 0.50 CI
		32.5	@ 0.95 CI

c. The HBV model predicts at least a slight increase in risk (e.g., fewer crossings with a score <0.1) for all climate projects across all three CI estimates.

d. The GFCL_HRM3 (0.5 and 0.05 CI) and GFDL_RCM3 (all CIs) climate predictions translate into slightly lower Q$_{25}$ values from the HSPF model at mid-century, and thus more crossings with a risk score <0.1.

e. The CCRM_CGCM3 and RCM_CGCM3 climate predictions translate into slightly lower Q$_{25}$ values from the WRFH model at all CIs for mid-century, and thus also more crossings with a risk score <0.1.
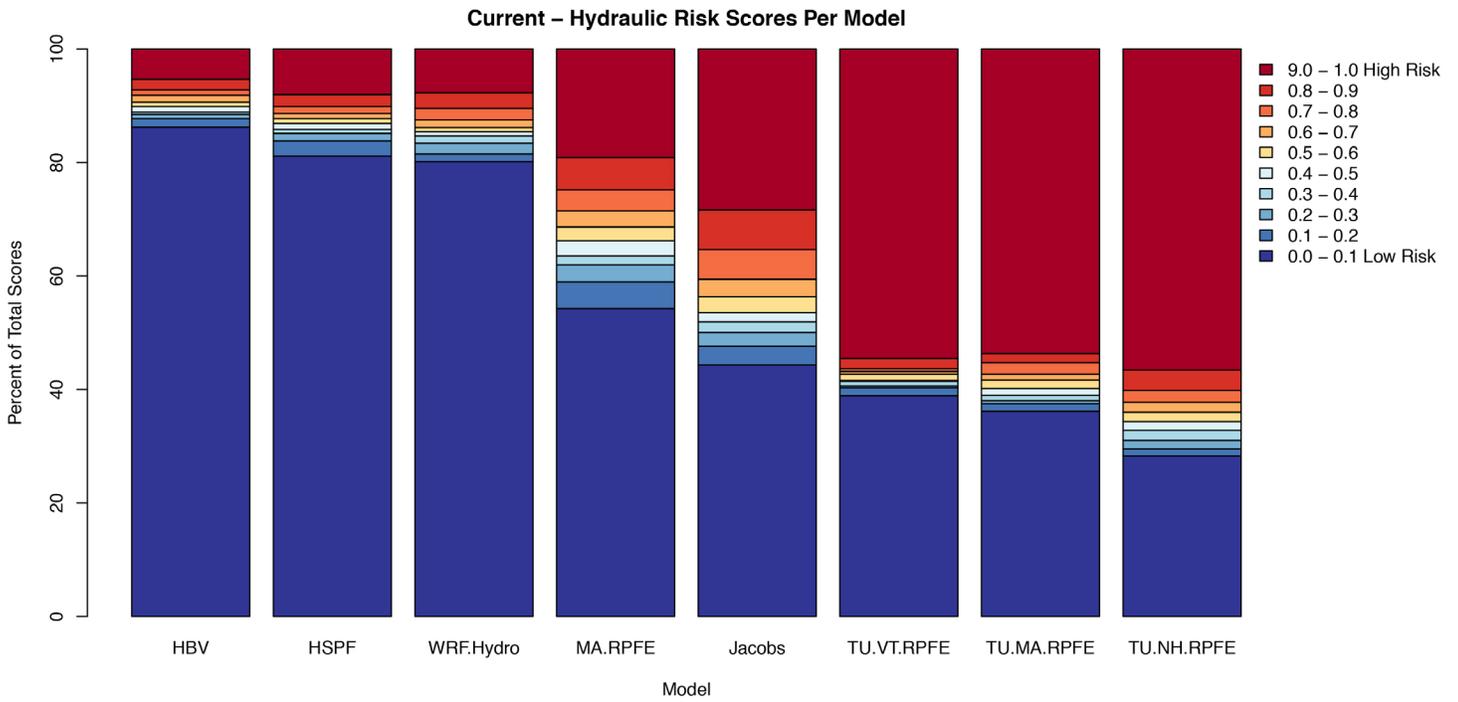
Figure 5-24: Comparison of hydraulic risk scores for current conditions, shown based on the 0.5 CI level Q$_{25}$ estimates for the physical models.
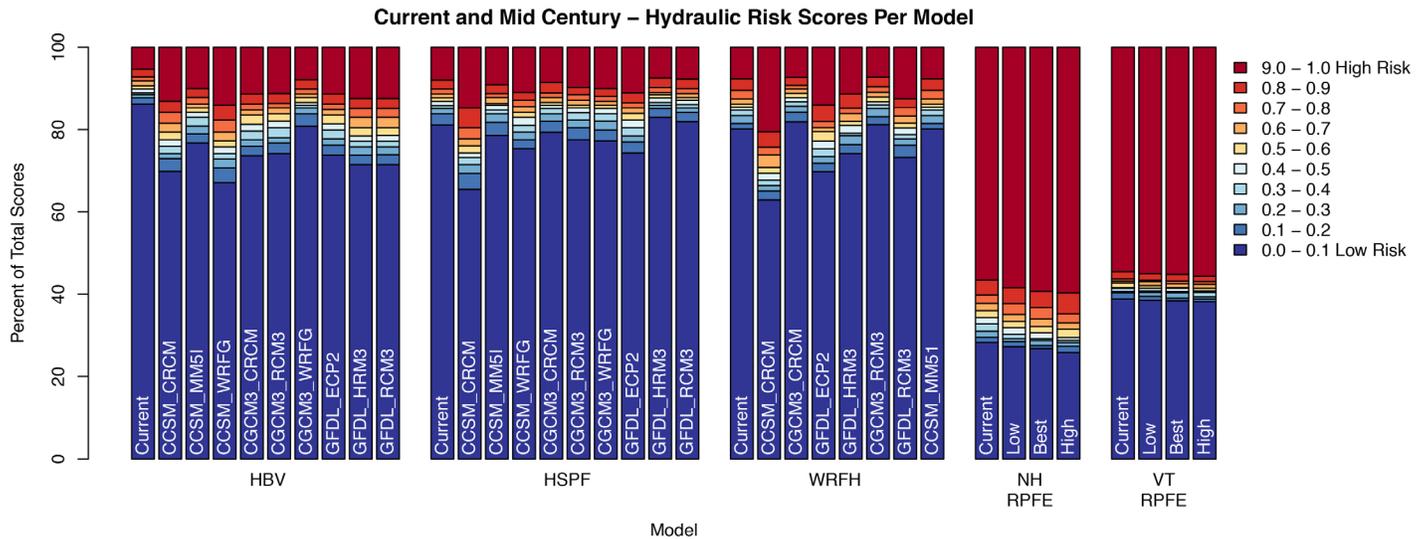

Figure 5-25: Comparison of hydraulic risk scores for mid-century, shown based on the 0.5 CI level Q$_{25}$ estimates for the physical models.
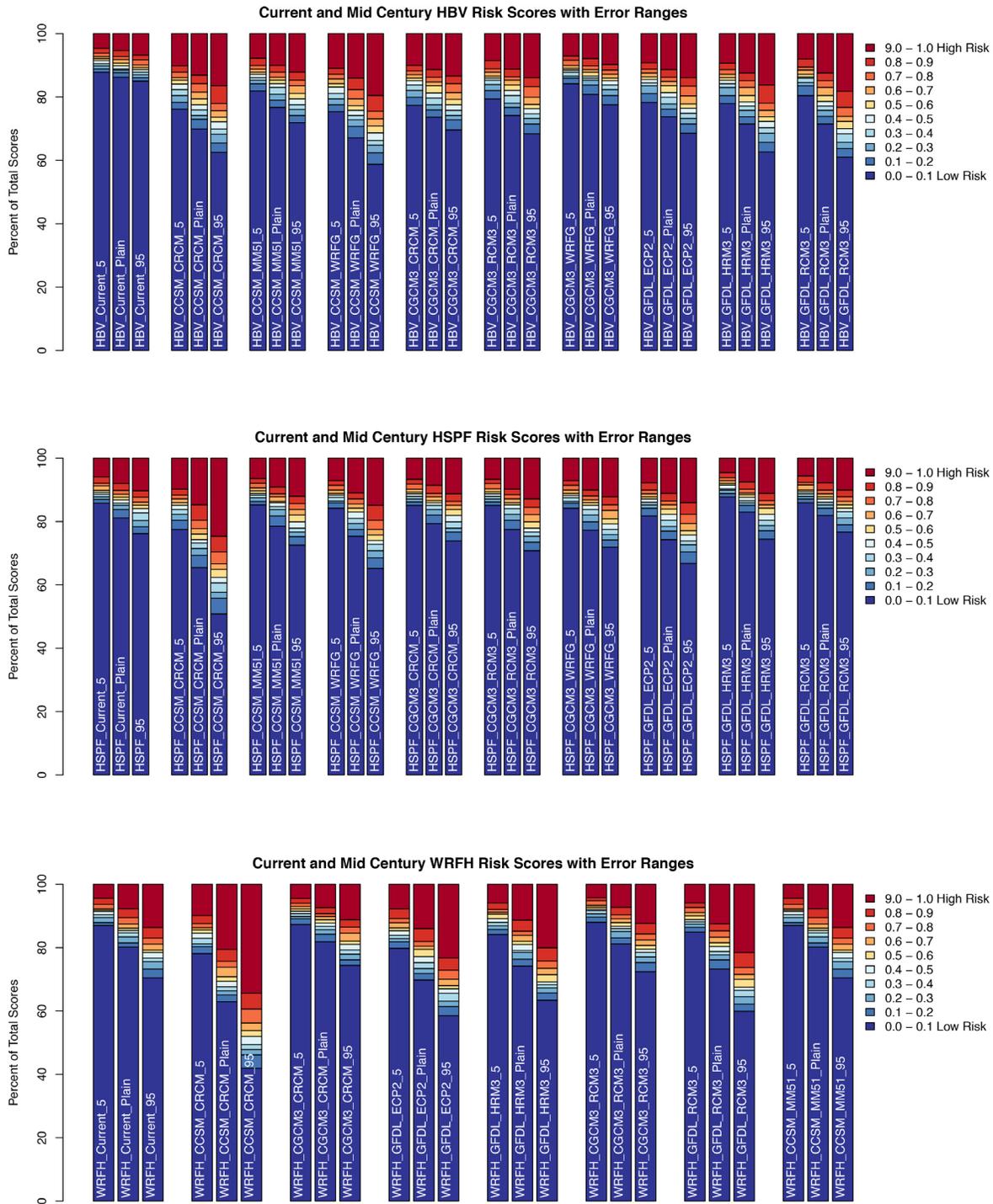
Figure 5-26: Hydraulic condition category classification for physical models HBV (top), HSPF (middle) and WRFH (bottom) (low, best, and high estimates; current and future conditions).

6. There is greater uncertainty in terms of hydraulic risk of failure at mid-century due to the error bounds around the Q_{25} estimate[30] than due to the range of climate predictions[31], Table 5-17 and Table 5-18.

   a. The difference in the percent of culverts that fall in the lowest category of risk ranges from 8.8 to 18.7 for HBV, 10.3 to 25.8 for HSPF, and 9.9 to 32.5 for WRFH across the 9 mid-century climate scenarios depending on which CI is considered.

   b. The maximum difference between the number of culverts in the lowest risk category between the .95 and .05 CI results for a given climate projection is 19.4 for HBV, 26.6 for HSPF, and 36.2 for WRFH.

7. The uncertainty due to both the error bounds around the Q_{25} estimate and mid-century climate are greatest for the WRFH model result, followed by HSPF, followed by HBV, Figure 5-25 and Figure 5-26.

8. WRFH predicts the largest shift in risk from current to mid-century, followed by HSPF, followed by HBV, Figure 5-26.

A more compact summary of the impact of the CI predictions of Q_{25} on hydraulic risk across the three physical models (HSPF, HBV and WRFH) and nine climate scenarios at mid-century is provided in Figure 5-27. Each panel compares the hydraulic risk scores at mid-century based on the minimum (A), average (B), or maximum (C) Q_{25} from the three models and nine climate projections. The top panel is based only on the 0.05 CI estimates of Q_{25} (result in the lowest risk), the middle panel is based only on the 0.5 CI estimates of Q_{25}, and the lower panel is based on only the 0.95 CI estimates of Q_{25} (result in the highest risk). Comparing horizontally, for example looking at columns A, B and C for a given CI, provides insight on sensitivity due to the combined differences in the models and climate predictions. Comparing vertically, for example looking at column A results from top to bottom, provides insight on sensitivity due to uncertainty in the Q_{25} estimates. Sensitivity to model and climate prediction choice increases as the potential risk of underestimating[32] Q_{25} decreases.

5.3.6.3 ENSEMBLE HYDRAULIC RISK SCORES

The individual model-climate-CI result summaries presented in Sections 5.3.6.1 and 5.3.6.2 form the basis for calculating final ensemble hydraulic risk scores for the project. Guiding principles utilized to develop scores are as follows:

---

30 *As inferred from the maximum difference in the percent of culverts that fall in the lowest risk category for the 0.95 CI model results compared to the 0.05 CI model results for a given model and climate predictions, 2nd column in Table 5-17.*

31 *As inferred from the difference in the percentage of culverts that fall in the lowest risk category for a given CI across the climate predictions, 3rd column in Table 5-18.*

32 *By calculating risk scores based on the upper bounds of estimates for Q25 (Figure 5 27 row 3, column C – maximum of 0.95 CI estimates) compared to the lowest (row 1, column A – minimum of 0.05 CI estimates).*

- This project considers three types of uncertainty with regards to hydraulic risk, including that due to:

  ○ Model selection,

  ○ Confidence Intervals (CIs), resulting from the extrapolation of an annual maxima series of streamflows to estimate flood flows with a given RI, and

  ○ Climate projections.

It is important to understand the relative impacts of each.

- TU model predictions are omitted due to concerns they provide an exaggerated view of hydraulic risk. There are no evident reasons for omitting any of the other statistical or physical models.

- Since the statistical models tend to predict higher levels of risk than the physical models, an ensemble score that blends the two modeling methods will provide a more balanced estimate of hydraulic risk. A method for extrapolating the statistical model results to mid-century is thus needed.

- Uncertainty in hydraulic risk resulting from the three CI estimates of Q_{25} adds complexity that is difficult to interpret. Scores based on the 0.95 CI estimates of Q_{25} likely overestimate risk, while scores based on the 0.05 CI estimates likely underestimate risk. Final scores will focus solely on the 0.5 CI level estimates of Q_{25} to provide a more balanced view of risk. End-users should, however, be made aware of the underlying uncertainty due to CIs.

- Lacking consensus on the most probable future climate conditions, it is important to include all of the mid-century climate predictions utilized in this project.

Averaging for calculation of ensemble scores is done at the Q_{25} level. When blended scores including both physical and statistical models are determined, the average Q_{25} for the physical models and statistical models are first calculated separately, then combined as a simple average to give both model types equal weight, rather than the individual models.

Five ensemble scores are summarized on Figure 5-28 based on:

A. Average 0.5 CI level Q_{25} estimates for the statistical models (Jacobs and UMass MA RPFE) under current conditions,

B. Average 0.5 CI level Q25 estimates for the physical models (HBV, HSPF, and WRFH) combined with the average Q_{25} estimates for the statistical models (Jacobs and UMass MA RPFE) under current conditions,

C. Average 0.5 CI level Q_{25} estimates for the physical models (HBV, HSPF, and WRFH) under current conditions,

D. Average 0.5 CI level Q_{25} estimates for the physical models (HBV, HSPF, and WRFH) at mid-century averaged across all available climate projections (9 for HSPF and HBV, 7 for WRFH),

Table 5-19: Correlations across the risk scores predicted by the different models.

Pearson's Correlation Coefficients Risk Score Pearson's Correlation Coefficients.

	Jacobs	MA RPFE	TU MA RPFE	TU NH RPFE	VT RPFE	HSPF 50%CI	HBV 50%CI	WRFH 50%CI
Jacobs	1							
MA RPFE	0.910	1						
TU MA RPFE	0.574	0.524	1					
TU NH RPFE	0.646	0.552	0.862	1				
TU VT RPFE	0.1741	0.141	0.704	0.574	1			
HSPF 50%CI	0.5501	0.641	0.359	0.325	0.083	1		
HBV 50%CI	0.471	0.561	0.319	0.286	0.079	0.960	1	
WRF Hydro 50%CI	0.5142	0.552	0.372	0.345	0.180	0.681	0.669	1

Risk Score Spearman's Correlation Coefficient.

	Jacobs	MA RPFE	TU MA RPFE	TU NH RPFE	VT RPFE	HSPF 50%CI	HBV 50%CI	WRFH 50%CI
Jacobs	1							
MA RPFE	0.983	1						
TU MA RPFE	0.856	0.834	1					
TU NH RPFE	0.877	0.859	0.984	1				
TU VT RPFE	0.740	0.714	0.952	0.933	1			
HSPF 50%CI	0.952	0.936	0.791	0.815	0.660	1		
HBV 50%CI	0.952	0.936	0.791	0.815	0.660	1	1	
WRF Hydro 50%CI	0.863	0.838	0.751	0.760	0.658	0.852	0.852	1

E.   Case D, averaged with an extrapolation of the statistical model results for the current climate (Case A) to mid-century, described below.

Case B, the ensemble score that blends the statistical and physical model results, is considered the most robust estimate of hydraulic risk under current conditions. In order to provide a similar blended ensemble score for mid-century, the percent change in the average Q_{25} across the physical models under current (case C) and mid-century (case D) climate was determined. This percent change was then applied to the current climate statistical model Q_{25} estimates and combined with the physical model mid-century results to produce a blended mid-century hydraulic risk estimate, case E in Figure 5-28.

The "best" hydraulic scores for current, mid-, and end-century are summarized on Figure 5-29. To summarize, these ensemble scores are based on:

B.   Average 0.5 CI level Q_{25} estimates for the physical models (HBV, HSPF, and WRFH) combined with the average Q_{25} estimates for the statistical models (Jacobs and UMass MA RPFE) under current conditions,

E.   Average 0.5 CI level Q_{25} estimates for the physical models (HBV, HSPF, and WRFH) at mid-century, averaged across all climate projections (9 for HSPF and HBV, 7 for WRFH) combined with a linear extrapolation of the statistical model results (Jacobs and UMass MA RFPE) to mid-Century, and

F.   Linear extrapolation of the blended ensemble of physical and statistical models at mid-century (Case E) to end-century.

5.3.7   SUMMARY OF RESULTS

Hydraulic risk scores take into consideration uncertainty with regards to model selection, CIs resulting from the extrapolation of an AMS to estimate RI flood flows, and future climate. Scores close to 1.0 are indicative of a high relative risk of failure, while scores close to 0.0 indicate a low relative risk of failure.

Table 5-20 summarizes the percent of crossings with risk scores greater than a given level for the three time periods current, mid-century, and end-century. Based on these results, for example, approximately 19% of the road-stream crossings in the Deerfield River watershed have a hydraulic risk score > 0.8 under current conditions. This number is estimated to increase to ~29% by mid-century, and to ~39% by end-century. The same data are summarized for cumulative risk less than a given level, in Table 5-21.
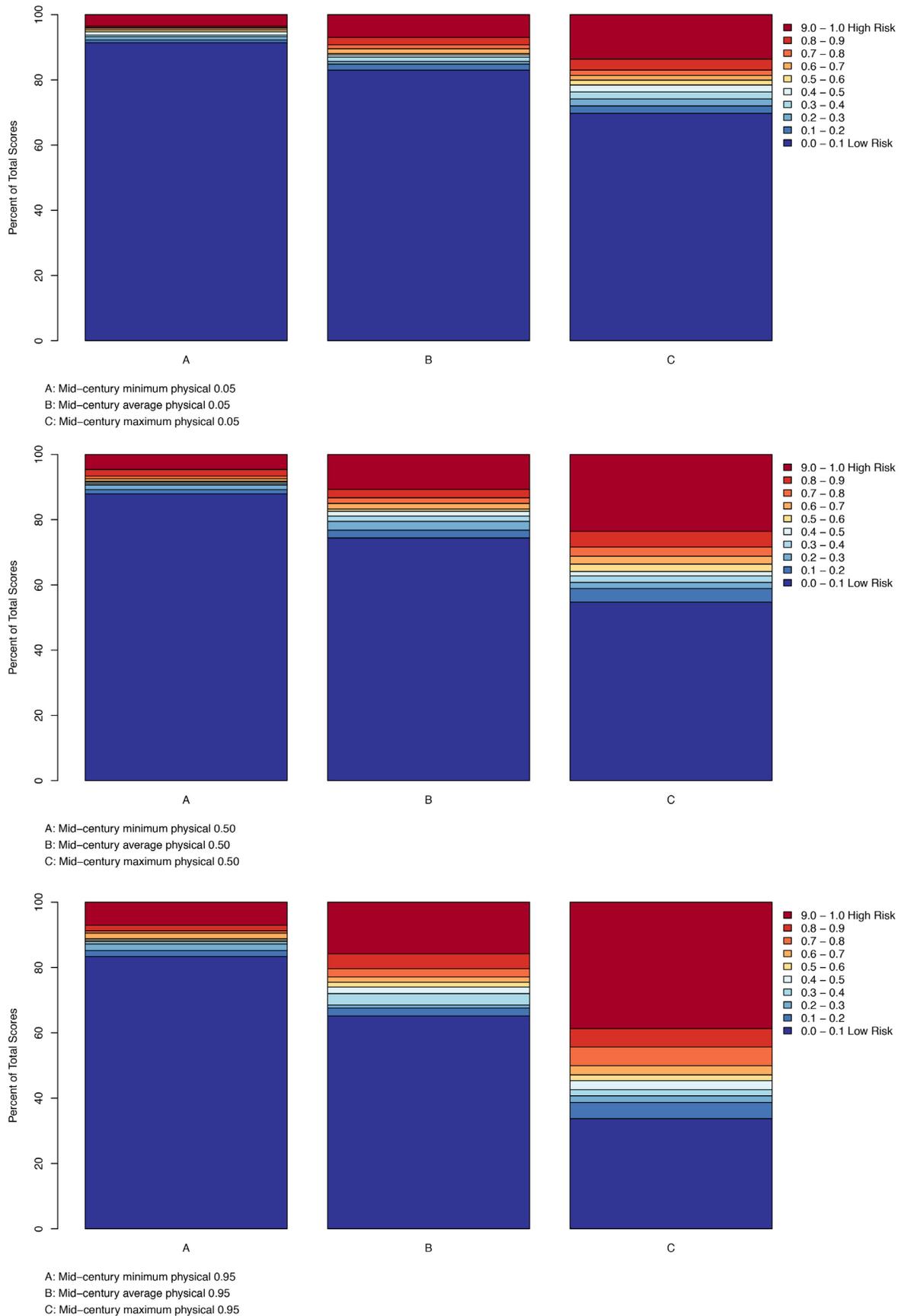
Figure 5-27: Comparison of hydraulic risk scores at mid-century based on the minimum (A), average (B), or maximum (C) of the physical model predictions for Q$_{25}$ at the .05 CI (top), 0.5 CI (middle), and 0.95 CI (bottom).
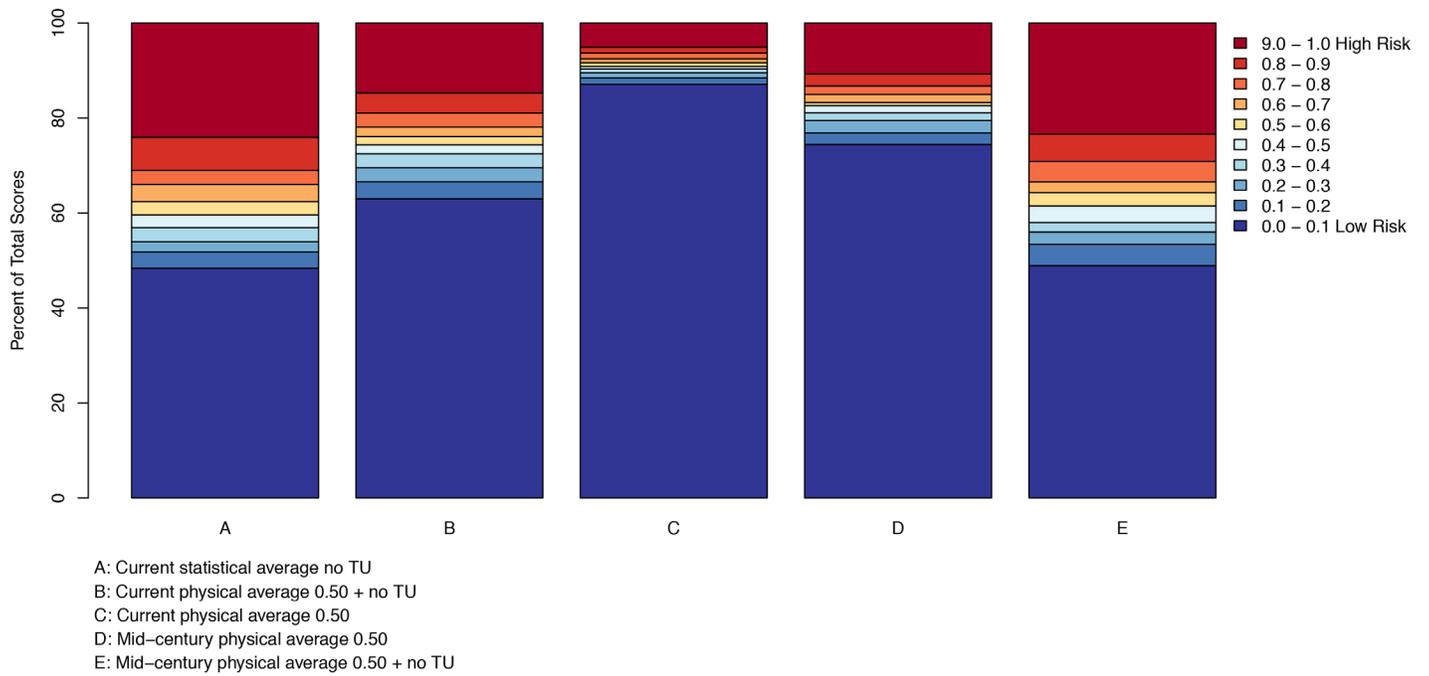
Figure 5-28: Comparison of simple and blended ensemble hydraulic risk scores for current and mid-century climate.

A: Current statistical average no TU
B: Current physical average 0.50 + no TU
C: Current physical average 0.50
D: Mid–century physical average 0.50
E: Mid–century physical average 0.50 + no TU

B: Current physical 0.50 + No TU
E: Mid-century physical 0.50 + No TU
F: End-century physical 0.50 + No TU
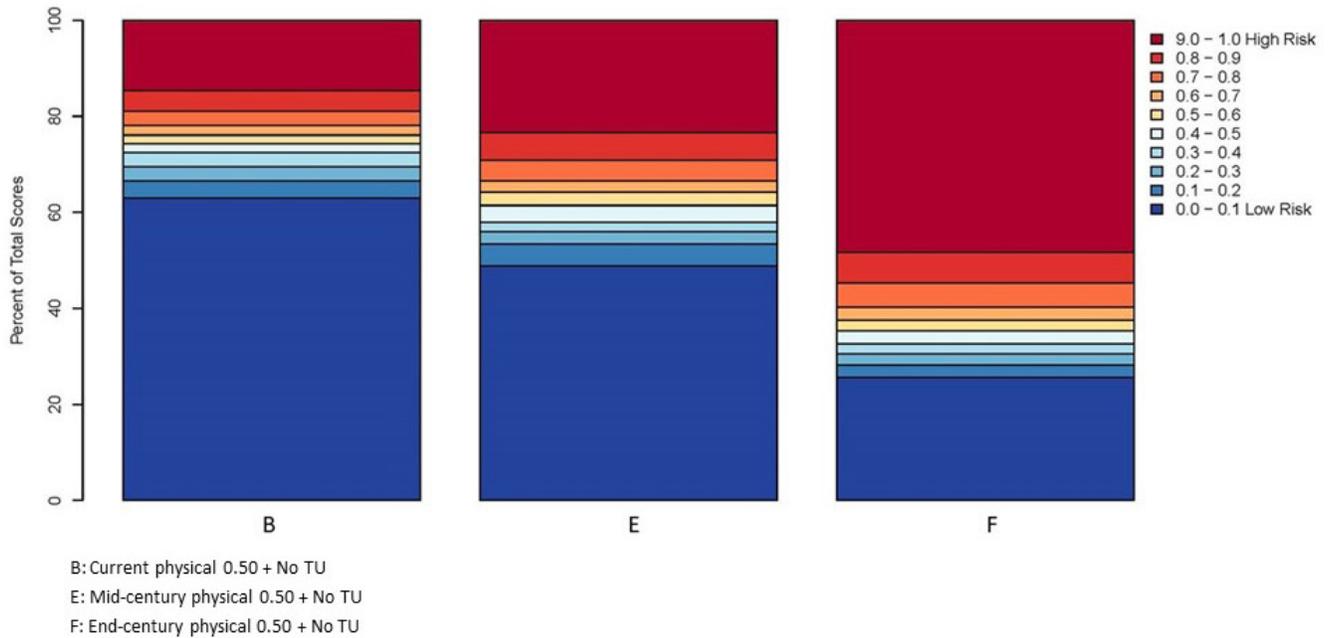
Figure 5-29: Summary of "best" hydraulic risk scores for current, mid-, and end-century.

Table 5-20: Percent of crossings exceeding a given hydraulic risk level, based on "best" scores for current, mid, and end-century.

Cumulative Risk	Current	Mid-Century	End-Century
>0.0	100.00	100.00	100.00
>0.1	37.03	51.14	60.78
>0.2	33.42	46.59	55.96
>0.3	30.48	44.04	53.55
>0.4	27.54	42.03	51.14
>0.5	25.67	38.55	49.00
>0.6	23.93	35.74	46.05
>0.7	21.93	33.47	43.24
>0.8	18.98	29.18	38.96
>0.9	14.71	23.43	31.86

Table 5-21: Cumulative hydraulic risk as a percent, based on "best" scores for current, mid-, and end-century.

Cumulative Risk	Current	Mid-Century	End-Century
<0.1	62.97	48.86	39.22
<0.2	66.58	53.41	44.04
<0.3	69.52	55.96	46.45
<0.4	72.46	57.97	48.86
<0.5	74.33	61.45	51.00
<0.6	76.07	64.26	53.95
<0.7	78.07	66.53	56.76
<0.8	81.02	70.82	61.04
<0.9	85.29	76.57	68.14
<1.0	100.00	100.00	100.00

5.4   Criticality

5.4.1   APPROACH

Crossing failures during extreme storms and flooding events can impede critical routes and severely disrupt the ability of communities to provide critical emergency services. This project component was a pilot to identify critically important road-stream crossings based on the impacts their failure would have on emergency medical services, and resulted in a scoring system that can be used to set priorities for upgrading transportation infrastructure. This phase focused on response times for ambulances and subsequent transport to hospitals. Network modeling was used to generate metrics. The approach essentially used scenario analysis to assess the interconnected network under current conditions and compare it to a scenario with a road-stream crossing failure. This component required collaboration with experts on emergency management and network analyses (computer scientists).

The potential to disrupt emergency medical services (EMS) in the Deerfield River watershed was evaluated for each road-stream crossing represented in the source data needed to support the analysis. There were two elements to this analysis.

1. Determination of origination and destination points for emergency response trips and identification of the most likely route to be taken to respond to medical emergencies and transport of victims to a local hospital.

2. Calculation of delay for each trip due to culvert or bridge failure. For each trip and each crossing, an alternative route that provided the shortest time delay was selected.

5.4.2   METHODS – WHICH MODELS/METHODS WERE USED, AND WHY

As part of the Deerfield Project, potential disruption of emergency medical services resulting from single crossing failures was evaluated as one component of Criticality (a measure of the impact of failure). The analyses were conducted by scientists from the UMass Amherst College of Information and Computer Sciences using methodologies developed specifically for this project.

In summary: First, a set of past EMS trips of ambulances and other vehicles in the study area was characterized. Each trip consisted of a starting location (i.e., address of the responding ambulance dispatch center) and a target location (e.g. the address of a patient). For each crossing location, a road closure due to culvert or bridge failure was simulated, the most time efficient alternative route was identified, and the amount of delay that each EMS trip experienced due to the failure was calculated. Then, a score was computed based on the delays.

EMS trips in the Deerfield River watershed were synthesized using a model derived from real emergency response call data. Data were obtained from the Shelburne Communications Center for 3,144 EMS response calls to target locations in the Deerfield River watershed (Figure 5-30). These trips occurred over a five-year period, from 2011 through 2015. Based on these data, a model was created to synthesize EMS response trips throughout the entire watershed. Population density was used as a measure of the likelihood that a patient call would originate in a certain area (if more people live in an area, it is more likely that a patient call will be made there). A probabilistic distribution of patient locations throughout the watershed was created from the population density data. With a patient location, the closest ambulance center can be determined as the ambulance dispatch location and the closest hospital as the target location. Digital road maps with speed limit data were used to identify the most probable route for each EMS trip. In this way, a probabilistic distribution of EMS trips was established. A total of 5,000 synthesized trips were used for this analysis (Figure 5-31).
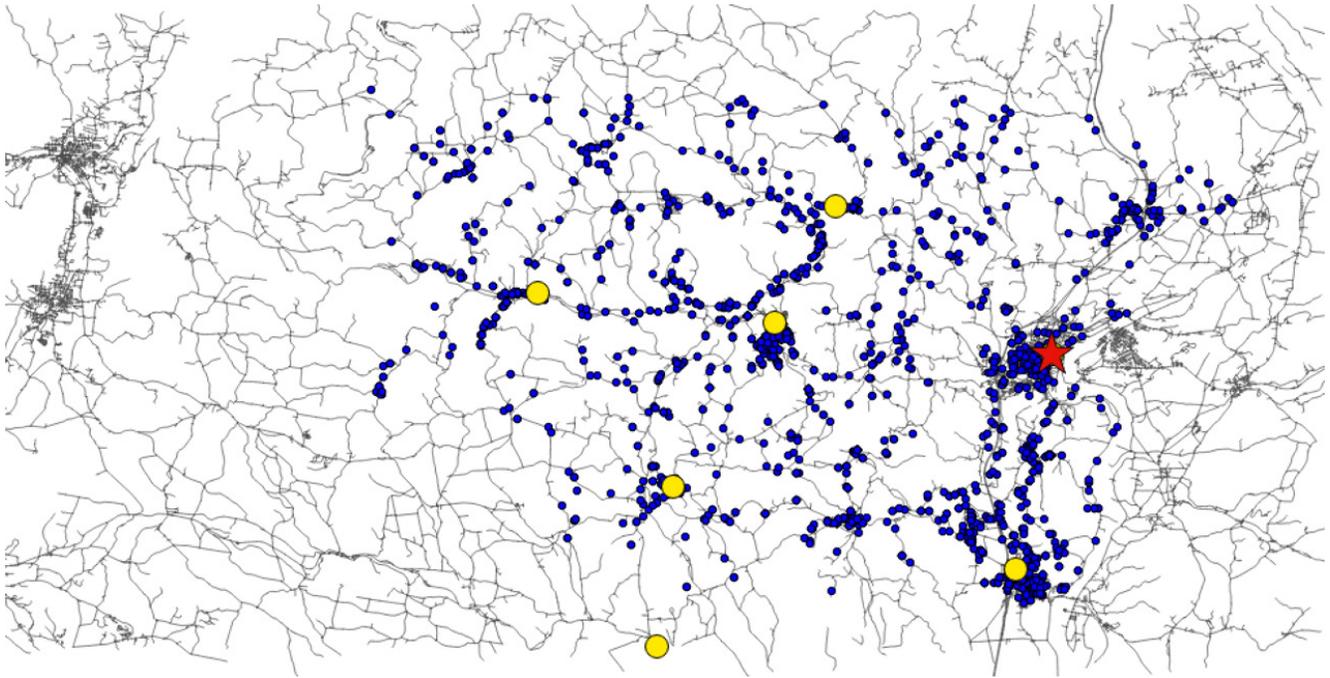
Figure 5-30: Location of 3,144 EMS calls (blue dots) handled by the Shelburne Communication Center, 2011-2015. Yellow dots are origination points for EMS response; the red star is the closest hospital.
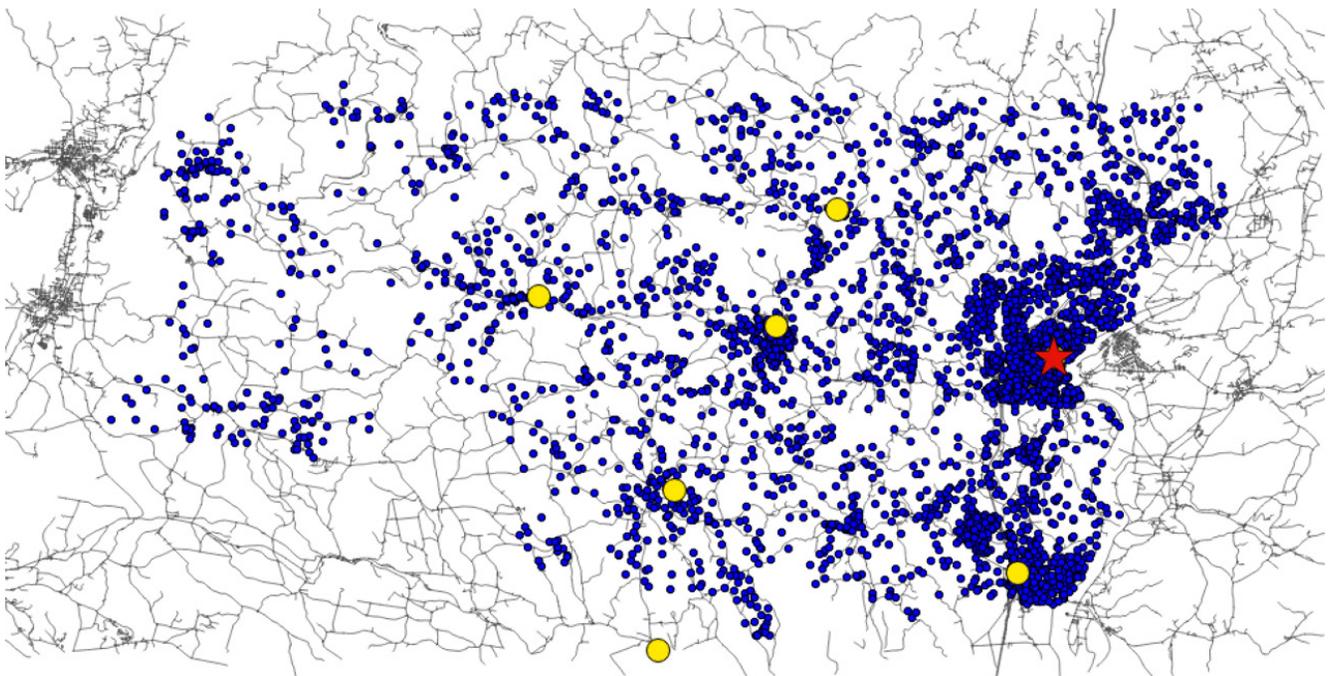
Figure 5-31: Location of 5,000 EMS calls (blue dots) synthesized by a model based on actual call data. Yellow dots are origination points for EMS response; the red star is the closest hospital.

To evaluate the effect of each crossing failure, the following procedure was used:

1. Calculate the shortest path lengths of all EMS trips based on the road map. Path length is based on distance and speed of travel (based on speed limits) and is a measure of time, not distance. Note the assumption that an ambulance always chooses the shortest path.

2. Remove the crossing from the map (simulate failure), making that road segment impassable.

3. Recalculate the shortest path lengths of all EMS trips. As a road segment becomes impassable, some shortest path lengths will

increase causing a delay in EMS response. Delays of over an hour (addresses on dead end roads with no alternative routes) were truncated to 60 minutes.

4. Combine delay data for all trips affected by a crossing failure to assess the impact on EMS response.

5. Various metrics were computed based on the number of affected trips and the magnitude of the delays (details in scoring section below).

5.4.3 DATA SOURCES

- Road maps: A map was downloaded from OpenStreetMap that consisted of a set of road segments and their joint locations. Each segment was associated with its speed limit.

- Population Density: U.S. Census data available from MassGIS.

- Crossings data: Crossing locations from the NAACC database.

- EMS trips: With assistance from our emergency management consultant (Josh Shanley), data from the Shelburne Communications Center of "first responder" services from private ambulance companies, fire stations, and hospitals in the area were compiled. Emergency response data were collected from seven communities resulting in 3,144 trip samples from actual calls received from 2011 to 2015. These data were acquired as a large text file and required a lot of parsing before they could be used in this analysis. This could be a barrier to using real data in the future depending on how those data are collected, stored and assessed by the emergency call centers. Use of synthesized EMS response data from models representative of different areas of the state would be one way of addressing this constraint.

5.4.4 SCORING

Based on stakeholder input it was clear that agencies/organizations have different preferences as to how best to characterize the impact of crossing failure on EMS. Four metrics were calculated, providing a variety of options that can be used to understand the effect of crossing failures on EMS.

- **Average delay:** Sum of trip delays, in minutes, of all trips affected by crossing failure, divided by the total number of trips in the watershed (includes trips unaffected by the failure).

- **Average affected delay:** Sum of trip delays, in minutes, of all trips affected by crossing failure, divided by the total number of trips affected by the failure (excludes trips unaffected by the failure).

- **Maximum delay:** The maximum delay, in minutes, encountered for any trip affected by a crossing. Because individual trip delays are capped at 60 minutes, maximum delay can never be greater than 60 minutes.

- **Overall delay:** An integrated metric that accounts for both the number of trips affected by a crossing failure and the magnitude of the delay for each affected trip. Creation of the integrated metric was a two-stage process. First, time delays were scored (each delay affecting one trip due to failure of one crossing) on a scale of 0 to 1 using a logistic equation. The raw score for each crossing is the sum of the delay scores for all the trips affected by that crossing. These raw scores were then logarithmically transformed and rescaled to range from 0 to 1.

When investigating the disruption of EMS, all four metrics will be available in the spatial data viewer for users to consider. However, when combining the Disruption metric with scores from other factors (risk of failure, ecological disruption) one metric (Overall EMS Delay) was chosen and rescaled from 0 (low disruption) to 1 (maximum disruption).

Each of the metrics has its shortcomings when used on its own.

Maximum delay: crossings affecting a single residence on a dead end road would always be identified as among the highest priorities.

Average delay: will tend to emphasize crossings on routes that received many trips even if the amount of delay per trip is quite small.

Average affected delay: removes the bias toward busy roads, but suffers from the same limitation as for maximum delay. Crossings that have profound effects on a single or small number of trips will be prioritized over crossings with shorter delays but that affect a much greater number of trips.

Overall delay: by integrating the magnitude of delay and number of trips, this metric addresses, to some degree, the limitations of the other metrics. However, the results are not in units of time (minutes) and therefore cannot be intuitively interpreted.

Although the results of the integrated metric are harder to intuitively understand, this metric does address some of the weaknesses of the other metrics. Because it will be combined with other metrics (risk of failure, ecological disruption), a process that obscures the original units in the metrics, it was chosen as the most appropriate one to use for a prioritization scheme based on a number of factors.

5.4.5 SUMMARY OF RESULTS

Results for the various Disruption of EMS Services metrics paint very different pictures of which crossings are most important for maintaining these critical services.

The Average EMS Delay metric results in only a handful of crossings scattered among the higher disruption scores with the remaining crossings clustered in the low score categories (Table 5-22 and Figure 5-32). The geographic distribution of higher scores shows that most occur on Route 2, reflecting the influence of large numbers of trips on the scores (Figure 5-33, Figure 5-34, and Figure 5-35). Although these scores are in units of minutes, due to the large number of trips modeled for the watershed (5,000) the highest delay score for these high Average EMS Delay crossings is less than half a minute. Although these scores are useful in a relative sense they are not particularly useful in absolute terms.

Much like the Average EMS Delay, Average Affected Delay scores are strongly skewed toward the low end of the disruption scale (Table 5-23, Figure 5-36 and Figure 5-37). Thirteen crossings received the highest score of 60 minutes, likely because of dead end roads with no alternative routes. Although they cluster toward the low end of the data distribution (Table 5-23), eight crossings had average affected delays of between 10 and 20 minutes; delays that could be significant for health emergencies that require rapid responses.

The crossings with the highest average affected delay scores occur on small roads, presumably dead end roads that lack alternative routes, while many of the crossings with average delays of 10-20 minutes are on Route 2 (Figure 5-38 and Figure 5-39).

The Maximum EMS delay scores show the same skewed distribution as for Average Delay and Average Affected Delay scores (Table 5-24, Figure 5-40 and Figure 5-41). Twenty-five crossings had the highest possible maximum delay scores of 60 minutes. These occurred on small roads with few or no alternative routes (Figure 5-42 and Figure 5-43). Thirty crossings had maximum delay scores of between 10 and 20 minutes; again, these crossings tended to occur on major routes such as Routes 2, 8A and 112 (Figure 5-42 and Figure 5-43).

The Overall EMS Delay metric was intentionally set up (transformed) to avoid strongly skewed results, yielding a broad range of scores distributed throughout the watershed (Table 5-25, Figure 5-44 and Figure 5-45). Many of the crossings with the highest disruption scores occurred on highways and larger roads, yet many moderately high scores occurred on smaller roads (Figure 5-46 and Figure 5-47).

Overall delay scores are unit-less and designed to provide relative rankings of crossings according to their disruption potential. It is important to keep in mind that those crossings that received

high scores may appear to be highly disruptive, and relative to other crossings in the watershed, they are. However, because the Deerfield River watershed is one of the most rural watersheds in Massachusetts, these scores may not be representative of crossings in other regions of Massachusetts. Until this methodology and scoring system are applied in other watersheds it is not known whether high scoring crossings in the Deerfield River watershed are of higher or lower concern than high scoring crossings in other watersheds.

This component generated a lot of interest from stakeholders, including MEMA, Regional Planning Authorities (RPAs), and municipalities. This seems like an aspect of the Deerfield Project that is ripe for future work. This methodology assessed only one crossing failure at a time. In working on this project, it became clear that the ability to analyze multiple failures would be extremely beneficial. Combining a multiple crossing failure analysis with probability of failure distributions derived from risk of failure scores will allow the use of probabilistic modeling with large numbers of replicates. This would provide a robust approach for identifying crossings with the highest potential to disrupt emergency services during storms of varying sizes that result in multiple crossing failures. A recommendation is to expand to multiple failures as one potential next phase of work.

The Disruption of EMS Services analysis is just one component of criticality. For the Deerfield Project, the use of network modeling to quantify criticality was piloted. To assess criticality fully it will be necessary to develop similar methodologies for assessing other elements of criticality, such as access to important infrastructure (water treatment plants, power plants, electrical substations, gas compressor stations) and the core functionality of the road network (transport of materials and people).
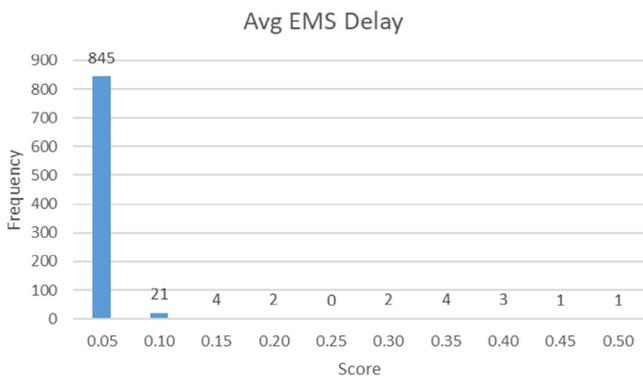
Figure 5-32: Distribution of Average EMS Delay scores (minutes).

Table 5-22: Distribution of Average EMS Delay scores (minutes).

Bin	Count	Percent
0.00-0.05	845	95.70%
0.05-0.10	21	2.38%
0.10-0.15	4	0.45%
0.15-0.20	2	0.23%
0.20-0.25	0	0.00%
0.25-0.30	2	0.23%
0.30-0.35	4	0.45%
0.35-0.40	3	0.34%
0.40-0.45	1	0.11%
0.45-0.50	1	0.11%
Total	883	
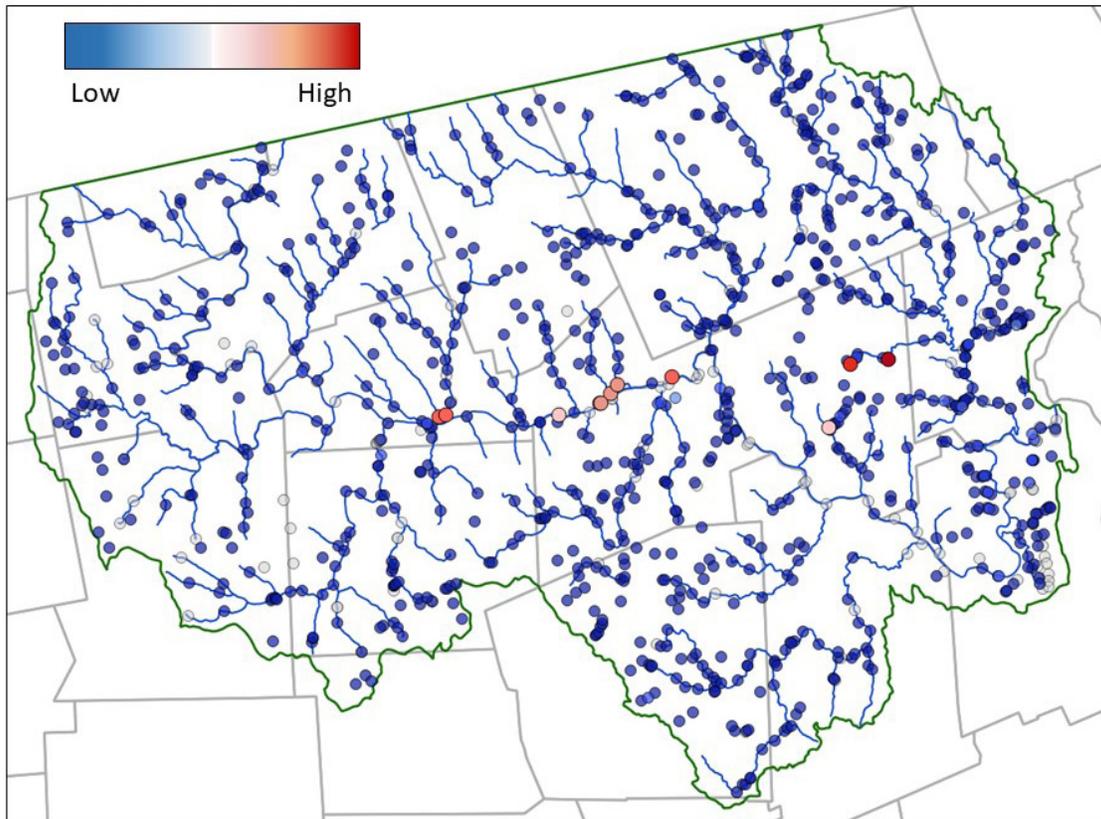
Figure 5-33: Geographic distribution of Average EMS Delay scores (minutes). The range is 0 to 0.5 minutes; high scores are in red; low scores in blue.
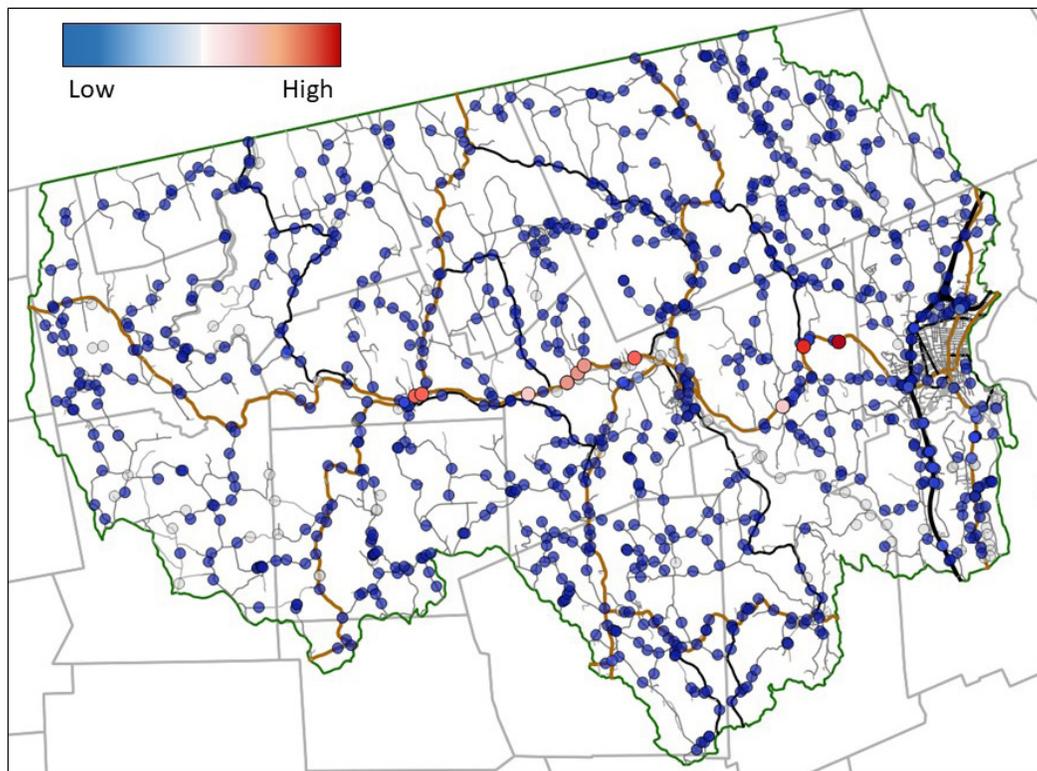
Figure 5-34: Geographic distribution of Average EMS Delay scores with roads. The range is 0 to 0.5 minutes; high scores are in red; low scores in blue.
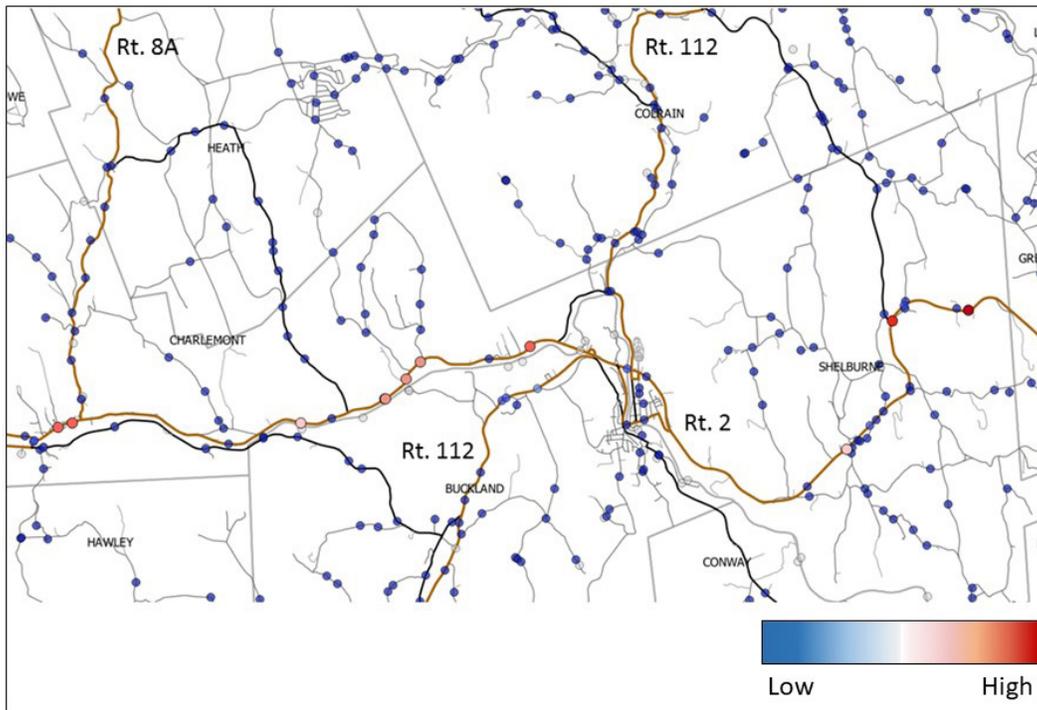
Figure 5-35: Geographic distribution of Average EMS Delay scores with roads (close up). The range is 0 to 0.5 minutes; high scores are in red; low scores in blue.
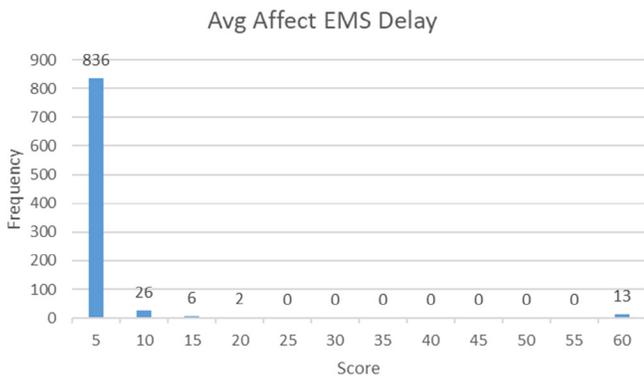
Figure 5-36: Distribution of Average Affected EMS Delay scores (minutes).

Table 5-23: Distribution of Average Affected EMS Delay scores (minutes).

Bin	Count	Percent
0-5	836	94.68%
5-10	26	2.94%
10-15	6	0.68%
15-20	2	0.23%
20-25	0	0.00%
25-30	0	0.00%
30-35	0	0.00%
35-40	0	0.00%
40-45	0	0.00%
45-50	0	0.00%
50-55	0	0.00%
55-60	13	1.47%
Total	883	
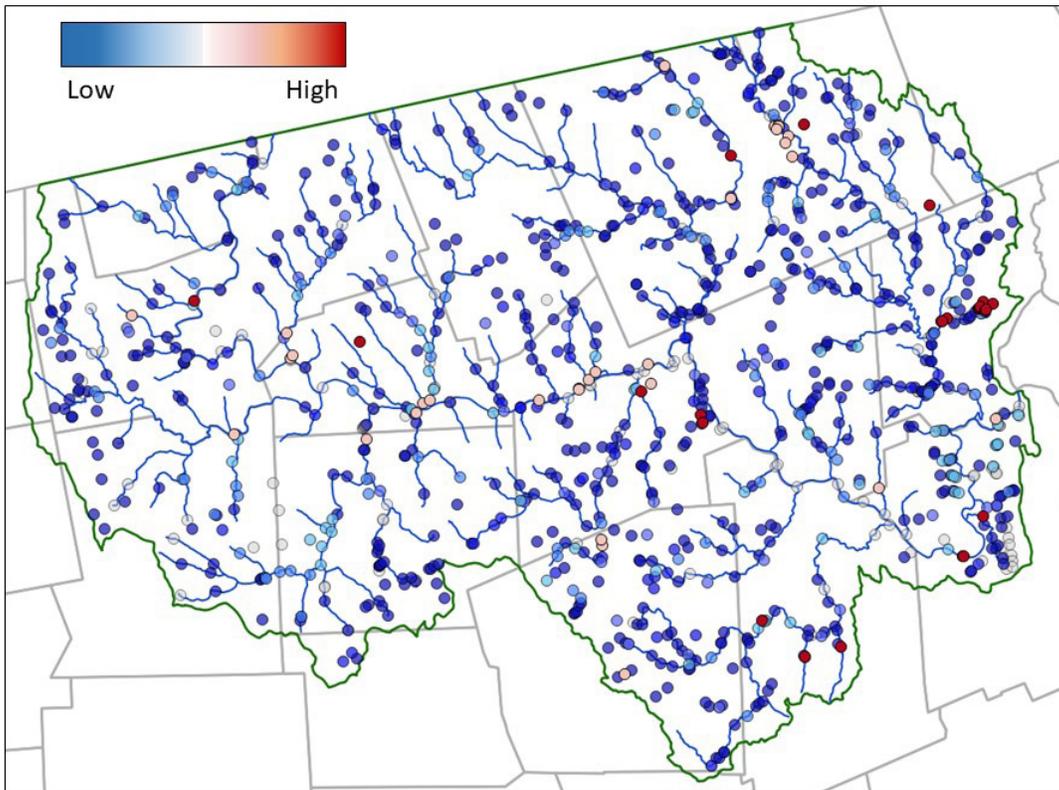
Figure 5-37: Geographic distribution of Average Affected EMS Delay scores (minutes). The range is 0 to 60 minutes; high scores are in red; low scores in blue.
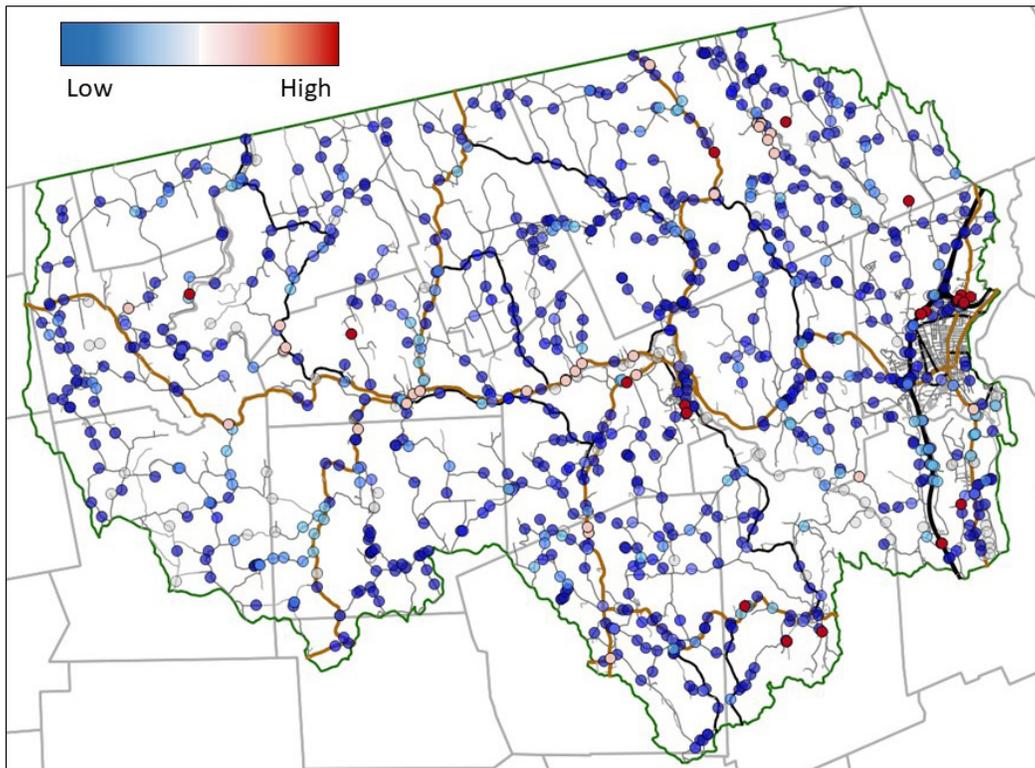
Figure 5-38: Geographic distribution of Average Affected EMS Delay scores with roads. The range is 0 to 60 minutes; high scores are in red; low scores in blue.
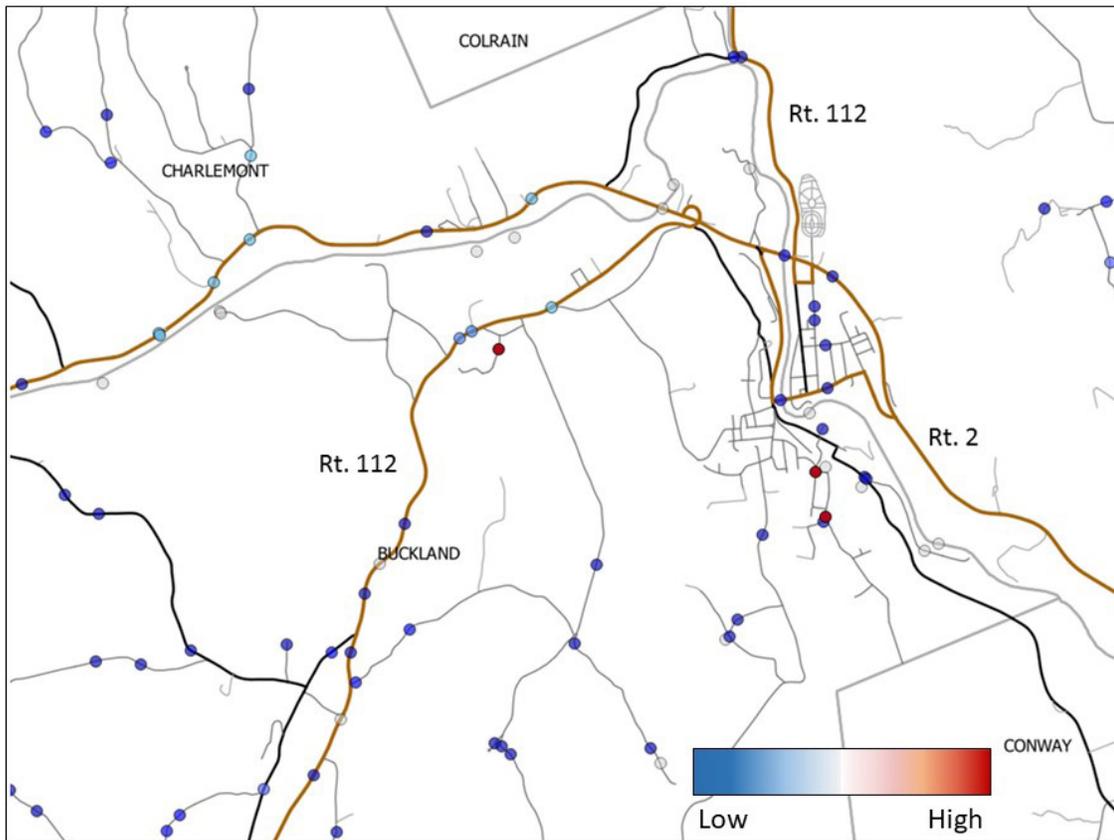
Figure 5-39: Geographic distribution of Average Affected EMS Delay scores with roads (close up). The range is 0 to 60 minutes; high scores are in red; low scores in blue.
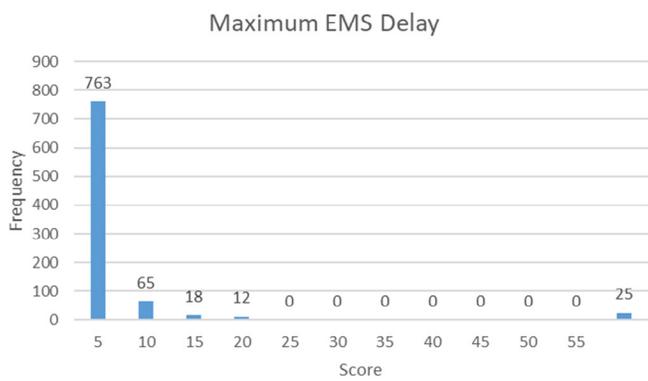
Figure 5-40: Distribution of Maximum EMS Delay scores (minutes).

Table 5-24: Distribution of Maximum EMS Delay scores (minutes).

Bin	Count	Percent
0-5	763	86.41%
5-10	65	7.36%
10-15	18	2.04%
15-20	12	1.36%
20-25	0	0.00%
25-30	0	0.00%
30-35	0	0.00%
35-40	0	0.00%
40-45	0	0.00%
45-50	0	0.00%
50-55	0	0.00%
55-60	25	2.83%
Total	883	
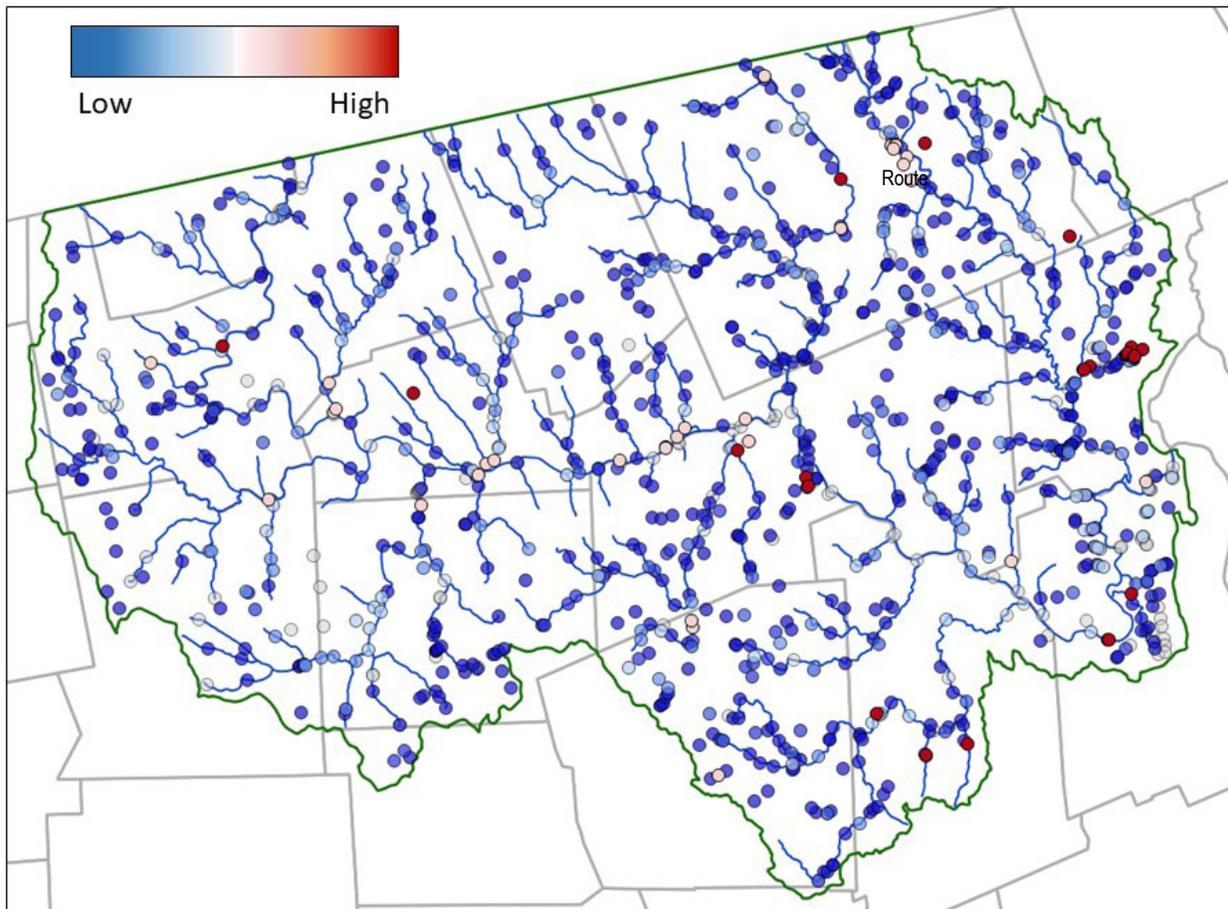
Figure 5-41: Geographic distribution of Maximum EMS Delay scores (minutes). The range is 0 to 60 minutes; high scores are in red; low scores in blue.
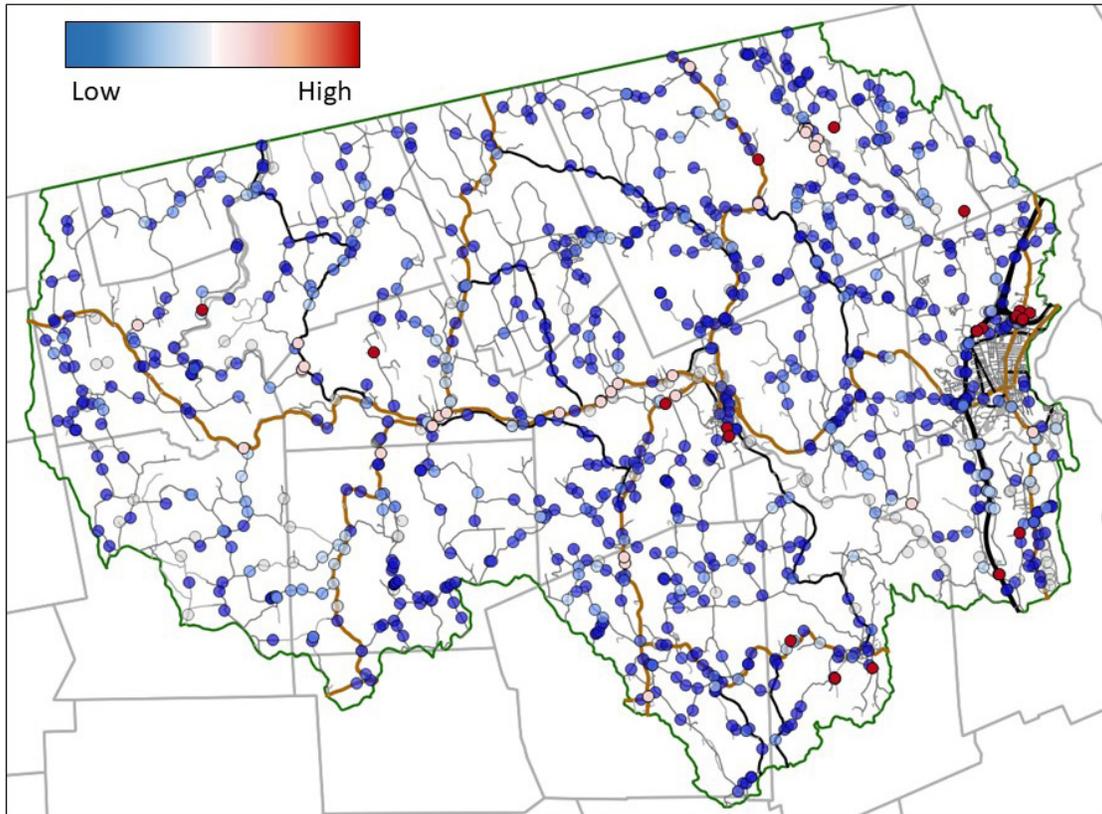
Figure 5-42: Geographic distribution of Maximum EMS Delay scores with roads. The range is 0 to 60 minutes; high scores are in red; low scores in blue.
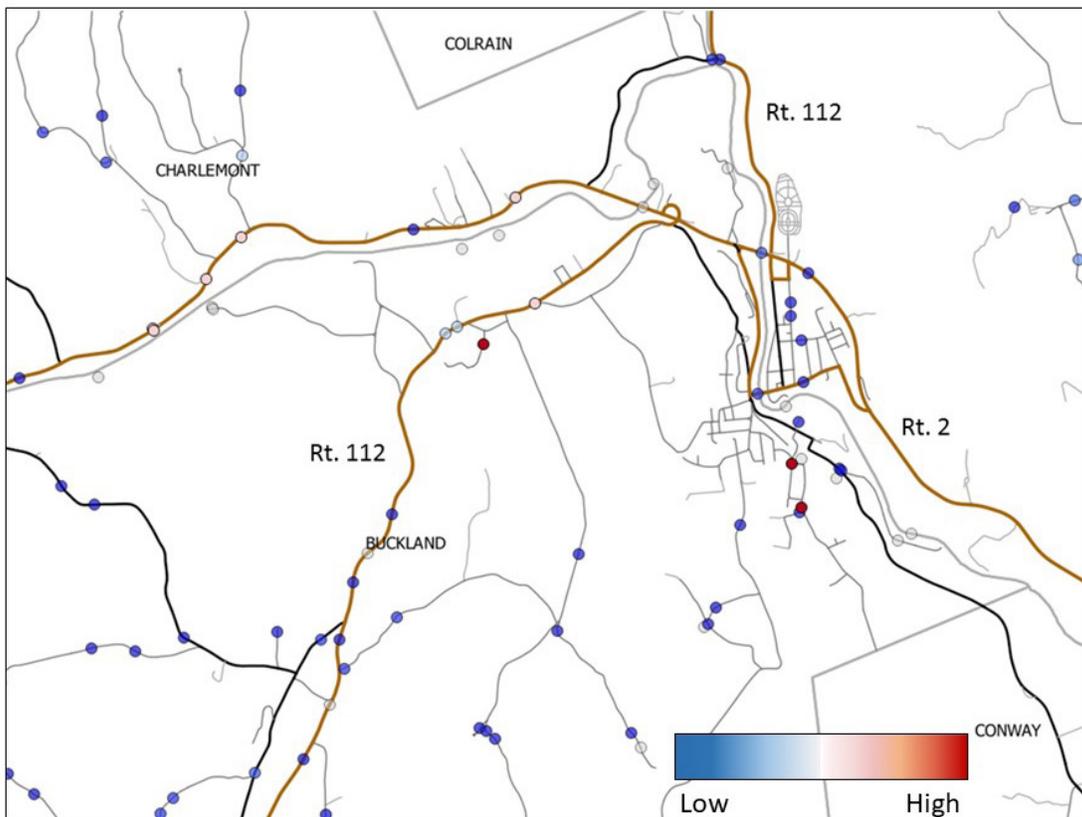
Figure 5-43: Geographic distribution of Maximum EMS Delay scores with roads (close up). The range is 0 to 60 minutes; high scores are in red; low scores in blue.

Figure 5-44: Distribution of Overall EMS Delay scores (0-1 scale).

Table 5-25: Distribution of Overall EMS Delay scores (0-1 scale).

Bin	Count	Average
0 - 0.1	295	33.41%
0.1 - 0.2	4	0.45%
0.2 - 0.3	13	1.47%
0.3 - 0.4	26	2.94%
0.4 - 0.5	83	9.40%
0.5 - 0.6	123	13.93%
0.6 - 0.7	162	18.35%
0.7 - 0.8	100	11.33%
0.8 - 0.9	54	6.12%
0.9 - 1.0	23	2.60%
Total	883	
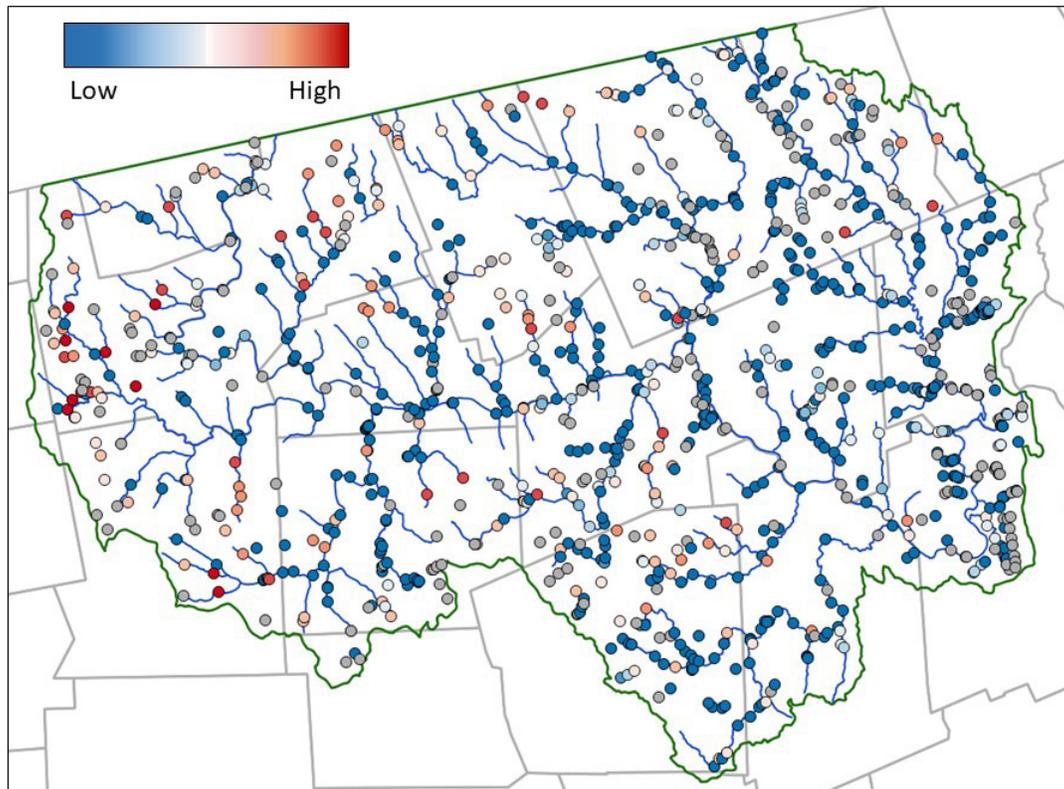
Figure 5-45: Geographic distribution of Overall EMS Delay scores. High scores on a 0 to 1 scale are in red; low scores in blue.
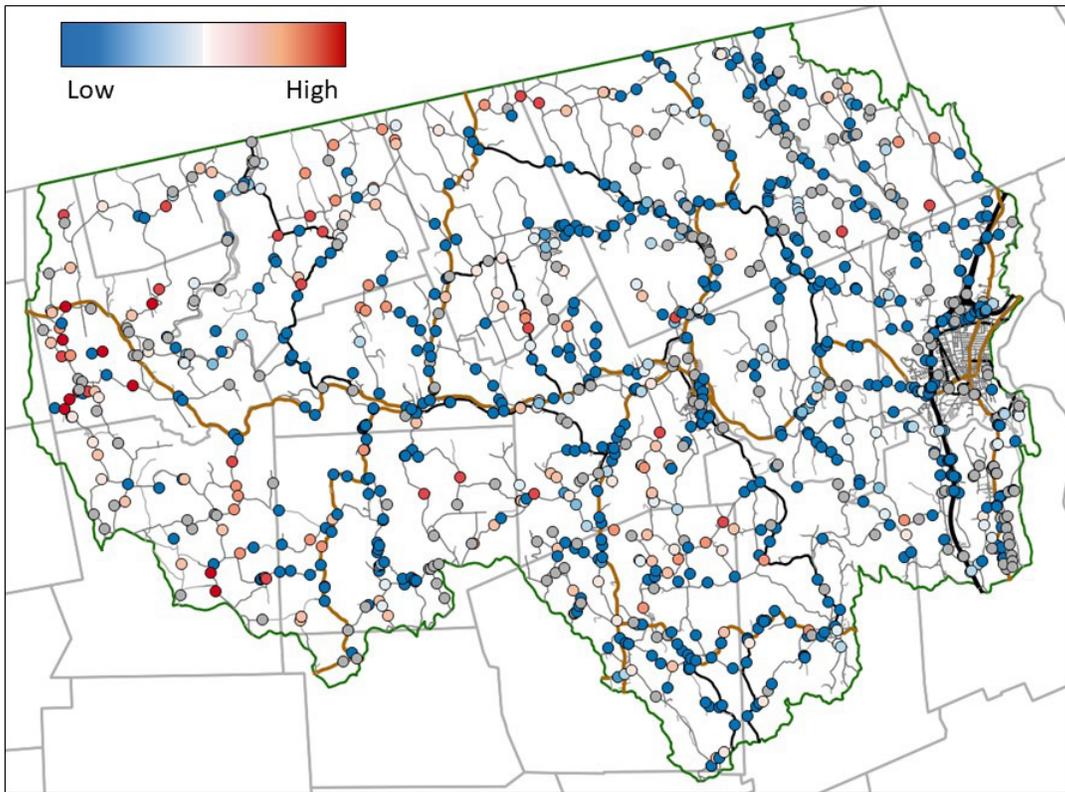
Figure 5-46: Geographic distribution of Overall EMS Delay scores with roads. High scores on a 0 to 1 scale are in red; low scores in blue.
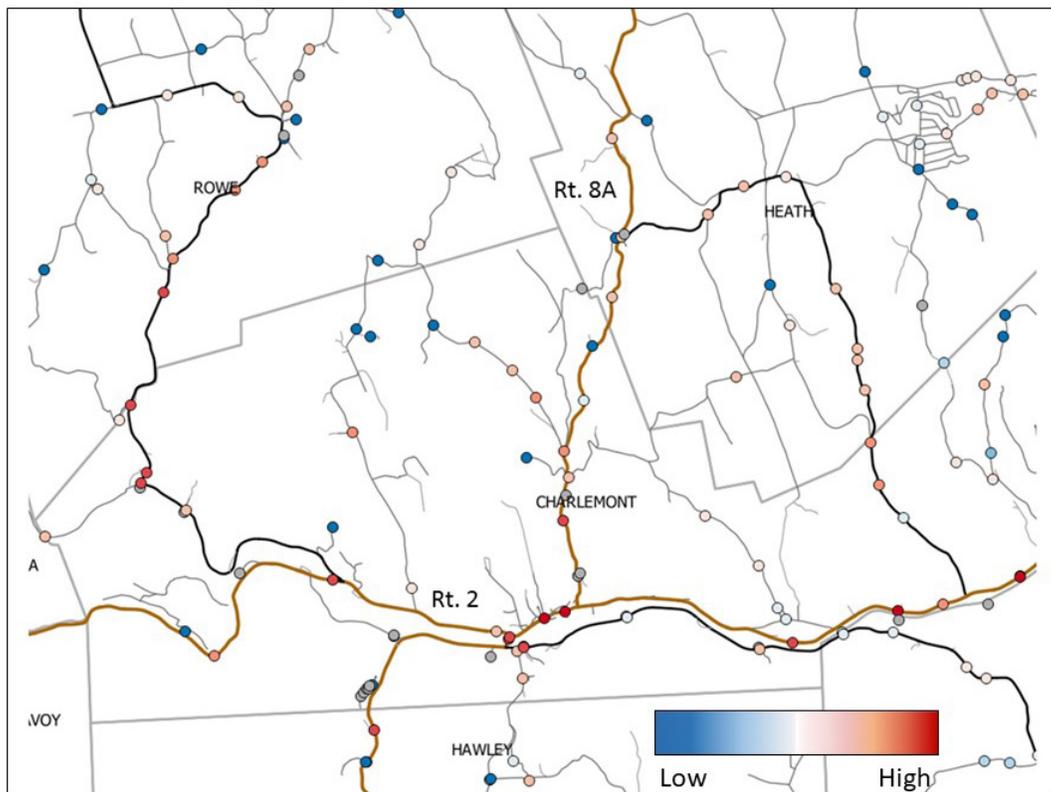
Figure 5-47: Geographic distribution of Overall EMS Delay scores with roads (close up). High scores on a 0 to 1 scale are in red; low scores in blue.

5.5 Ecological Disruption

Dams and road crossings can disrupt aquatic connectivity, resulting in significant impacts on river and stream ecology. Dams generally present more severe barriers to the movement of wood, sediment and aquatic organisms, but road crossings can also affect these ecological processes and are much more numerous. The collective impact of road crossings can significantly disrupt stream ecology and reduce the viability of populations of fish and wildlife.

When evaluating the disruptive influence of road-stream crossings on stream ecology it is important to take into account the extent to which crossings disrupt aquatic connectivity and the quality of habitat affected. For the Deerfield Project, the loss of aquatic connectivity was assessed using Critical Linkages, a component of the Conservation Assessment and Prioritization System (CAPS) (see Section 5.5.3).

Habitat quality was assessed in a variety of ways. As part of the Critical Linkages analysis, the loss of aquatic connectivity is multiplied by the Index of Ecological Integrity (IEI), an output of CAPS analyses. Two other aspects of habitat quality were analyzed as part of this project so that they can be combined with changes in aquatic connectivity to address specific needs. To address the conservation needs of cold water habitat and associated species, a stream temperature model was developed and applied throughout the Northeastern U.S. This stream temperature model was then used, along with other GIS data, to model Eastern brook trout occupancy (probability of occurrence) for stream reaches throughout the region.

5.5.1 HABITAT QUALITY - STREAM TEMPERATURE

5.5.1.1 APPROACH

Stream temperature is an important factor influencing populations of stream organisms such as fish, amphibians and invertebrates. Many streams lack data on water temperature that can be used to assess habitat suitability for cold water species or project changes in habitat quality as climate change affects the Northeastern U.S. Stream temperature models can play an important role in conservation by estimating thermal regimes for streams that lack temperature data. To meet this need, a statistical model of daily stream temperature was developed and applied across the eastern U.S. This model incorporates features of stochastic models and extends the Letcher et al. (2016) framework to large geographic areas.

5.5.1.2 METHODS - WHICH MODELS/METHODS WERE USED, AND WHY

Statistical models of stream temperature often rely on the close relationship between air temperature and water temperature. However, this relationship breaks down during the winter in temperate zones, particularly as streams freeze. The winter period, when phase change and ice cover alter the air-water relationship, differs both spatially and temporally (annually). An index of air-water synchrony was developed specifically to model the portion of the year that is not affected by freezing temperatures. The index

Table 5-26: Summary of stream temperature data by state and locations (dependent variable).

State	$N_{records}$	N_{years}	$N_{locations}$	$N_{reaches}$
CT	5,007,479	19	515	418
DE	294,591	10	1	1
MA	3,212,204	20	628	546
MD	258,076	13	497	402
ME	5,522,845	22	274	189
NH	17,191,459	9	151	124
NJ	247,974	4	61	42
NY	6,357,709	20	292	266
PA	17,280,353	10	162	142
RI	2,615	3	4	4
VA	159,334	2	41	41
VT	21,161	13	54	53
WV	835,882	8	214	185
Totals:	56,391,682	22	2894	2413

is the difference between air and observed water temperatures divided by the water temperature.

The index was calculated for each day of the year at each reach for each year with temperature data. The 99.9% confidence interval of index values was then calculated for days between May 5 and October 2. Moving from the middle of the year (day 180) to the beginning of the year, the first occurrence of 10 consecutive days that were outside the 99.9% CI was identified. This was designated as the spring breakpoint. Similarly, moving from the middle to the end of the year, the first event with fewer than 16 consecutive days within the 99.9% CI was designated as the autumn breakpoint. Independent breakpoints were determined for each reach-year combination. More details regarding the identification of the synchronized period can be found in Letcher et al. (2016). The portion of the year between the spring and autumn breakpoints was used to model the non-winter, ice-free stream temperatures.

A generalized linear mixed model was used to account for correlation in space (stream reaches nested within HUC8 watersheds). This allowed the incorporation of short time series and long time series from different reaches, as well as discontinuous time series from the same reaches, without risk of pseudoreplication (Hurlbert, 1984). A linear model could be used by limiting the stream drainage area to <200 km² and modeling only the synchronized period of the year, avoiding the non-linearities that occur at very high temperatures due to evaporative cooling and at freezing temperatures due to phase change (Mohseni and Stefan, 1999). A total of 248,517 stream temperature records from 1,352 streams were used to fit the model and 100,909 records were withheld for model validation.

Table 5-27: GIS data used in modeling stream temperature (predictor variables).

Variable	Description	Source	Processing
Total Drainage Area	The total contributing drainage area from the entire upstream network	The SHEDS Data project	The individual polygon areas are summed for all of the catchments in the contributing network
Riparian Forest Cover	The percentage of the upstream 200ft riparian buffer area that is covered by trees taller than 5 meters	The National LandCover Database (NLCD)	All of the NLCD forest type classifications are combined and attributed to each riparian buffer polygon using GIS tools. All upstream polygon values are then aggregated.
Daily Precipitation	The daily precipitation record for the individual local catchment	Daymet Daily Surface Weather and Climatological Summaries	Daily precipitation records are spatially assigned to each catchment based on overlapping grid cells using the zonalDaymet R package
Upstream Impounded Area	The total area in the contributing drainage basin that is covered by wetlands, lakes, or ponds that intersect the stream network	U.S. Fish & Wildlife Service (FWS) National Wetlands Inventory	All freshwater surface water bodies are attributed to each catchment using GIS tools. All upstream polygon values are then aggregated.
Percent Agriculture	The percentage of the contributing drainage area that is covered by agricultural land (e.g. cultivated crops, orchards, and pasture) including fallow land.	The National LandCover Database	All of the NLCD agricultural classifications are combined and attributed to each catchment polygon using GIS tools. All upstream polygon values are then aggregated.
Percent High Intensity Developed	The percentage of the contributing drainage area covered by places where people work or live in high numbers (typically defined as areas covered by more than 80% impervious surface)	The National LandCover Database	The NLCD high intensity developed classification is attributed to each catchment polygon using GIS tools. All upstream polygon values are then aggregated.

Table 5-28: Derived metrics based on model predictions from 1980-2015.

Metric	Mean	Min	Max	Description
Mean maximum temperature	20.57	12.61	34.11	Maximum daily mean water temperature (°C) averaged over 36 years (1980 - 2015)
Max maximum temperature	22.30	14.05	35.25	Maximum over years of the maximum daily mean temperature
Mean July temperature	18.25	8.83	32.34	Mean daily July temperature over years
Mean August temperature	17.74	8.52	31.76	Mean daily August temperature over years
Mean summer temperature	17.49	7.92	31.77	Mean daily summer temperature over years
Mean 30-day maximum temperature	18.76	9.68	32.71	Maximum 30-day temperature for each year averaged over years
Mean number of days over 18°C	47.73	0.00	194.00	Mean number of days per year the mean daily temperature exceeds 18°C
Mean number of days over 22°C	5.17	0.00	194.00	Mean number of days per year the mean daily temperature exceeds 22°C
Annual frequency of exceeding 18°C	0.86	0.00	1.00	Frequency of years the mean daily temperature ever exceeds 18°C
Annual frequency of exceeding 22°C	0.28	0.00	1.00	Frequency of years the mean daily temperature ever exceeds 22°C
Mean annual resistance	311.95	69.96	789.80	Mean annual resistance of water temperature to peak (summer) air temperature
Thermal sensitivity	0.61	0.35	0.98	Thermal sensitivity of water temperature to changes in air temperature
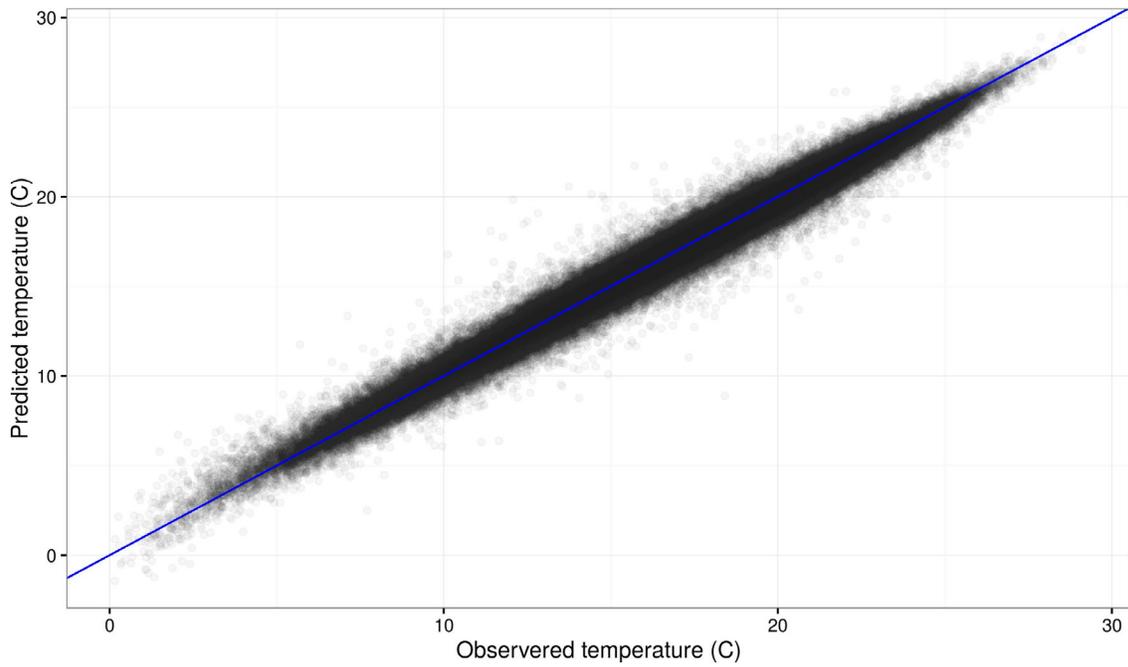
Figure 5-48: Relationship between observed and predicted water temperature for all data.
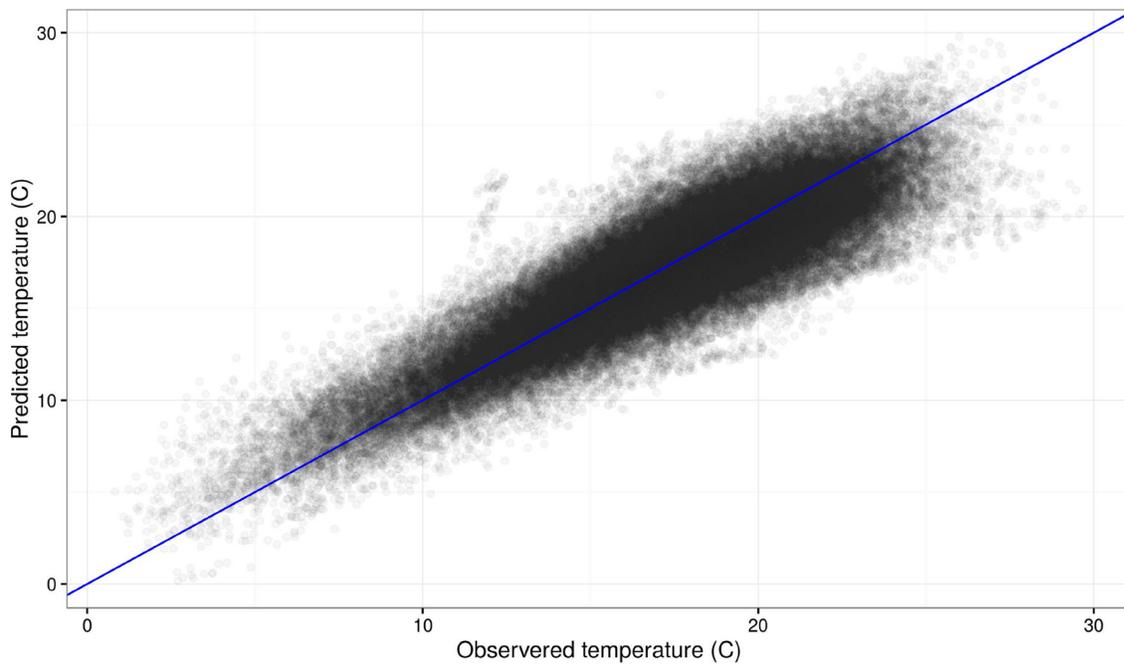
Figure 5-49: Relationship between observed and predicted water temperature for data held out for validation.
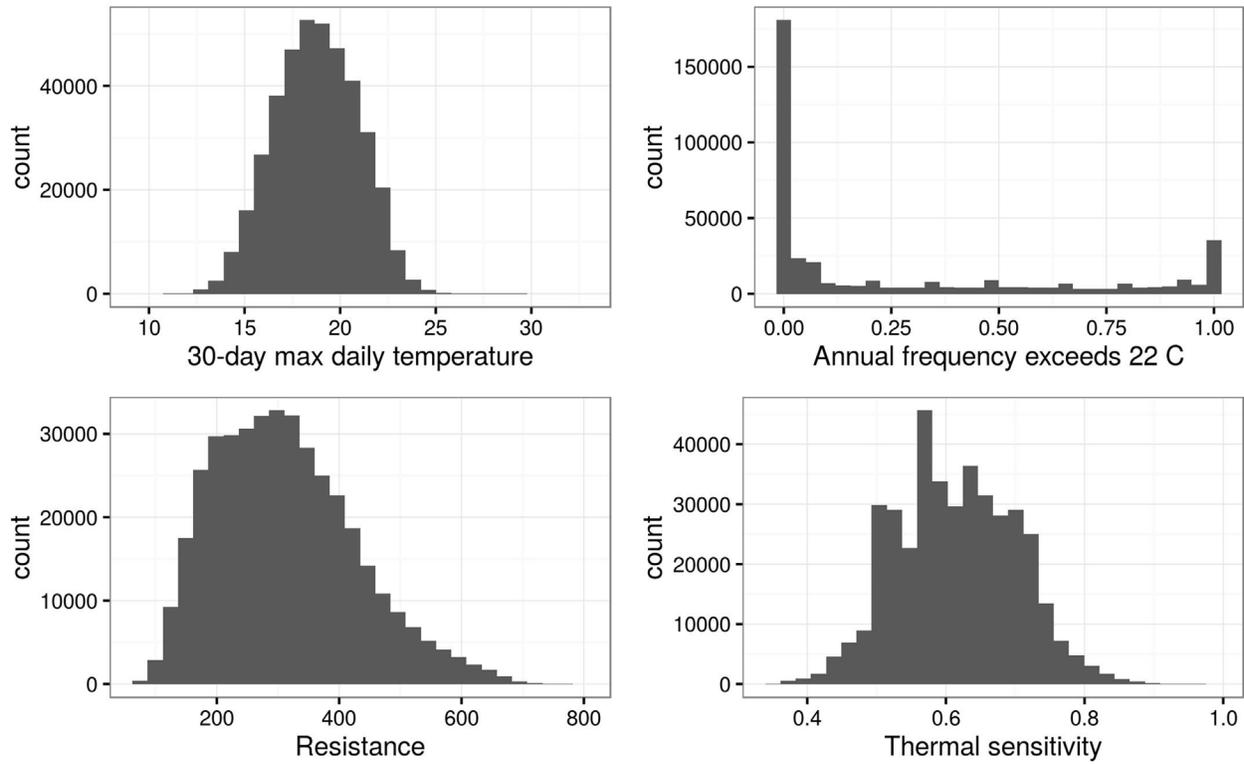
Figure 5-50: Distributions of model predictions for four derived metrics.

Table 5-29: Model results. A positive coefficient means that an increase in the variable leads to an increase in stream temperature and vice-versa.

Parameter	Mean	SD	LCRI	UCRI
Intercept	16.69	0.135	16.4182	16.949
AirT	1.91	0.022	1.8620	1.950
7-day AirT	1.36	0.029	1.3015	1.417
2-day Precip	0.06	0.002	0.0546	0.063
30-day Precip	0.01	0.006	0.0005	0.026
Drainage Area	0.04	0.096	-0.1452	0.232
Impounded Area	0.50	0.095	0.3181	0.691
Forest Cover	-0.15	0.047	-0.2455	-0.059
AirT x 2-day Precip	0.02	0.002	0.0195	0.028
AirT x 30-day Precip	-0.01	0.004	-0.0224	-0.007
AirT x Drainage	-0.06	0.029	-0.1170	-0.006
AirT x Impounded Area	0.02	0.029	-0.0345	0.077
AirT x Forest	-0.02	0.015	-0.0508	0.009
2-day Precip x Drainage	-0.04	0.002	-0.0424	-0.034
30-day Precip x Drainage	-0.06	0.006	-0.0709	-0.046
AirT x 2-day Precip x Drainage	-0.01	0.002	-0.0156	-0.008
AirT x 30-day Precip x Drainage	-0.01	0.004	-0.0193	-0.004
AR1	0.77	0.002	0.7681	0.776
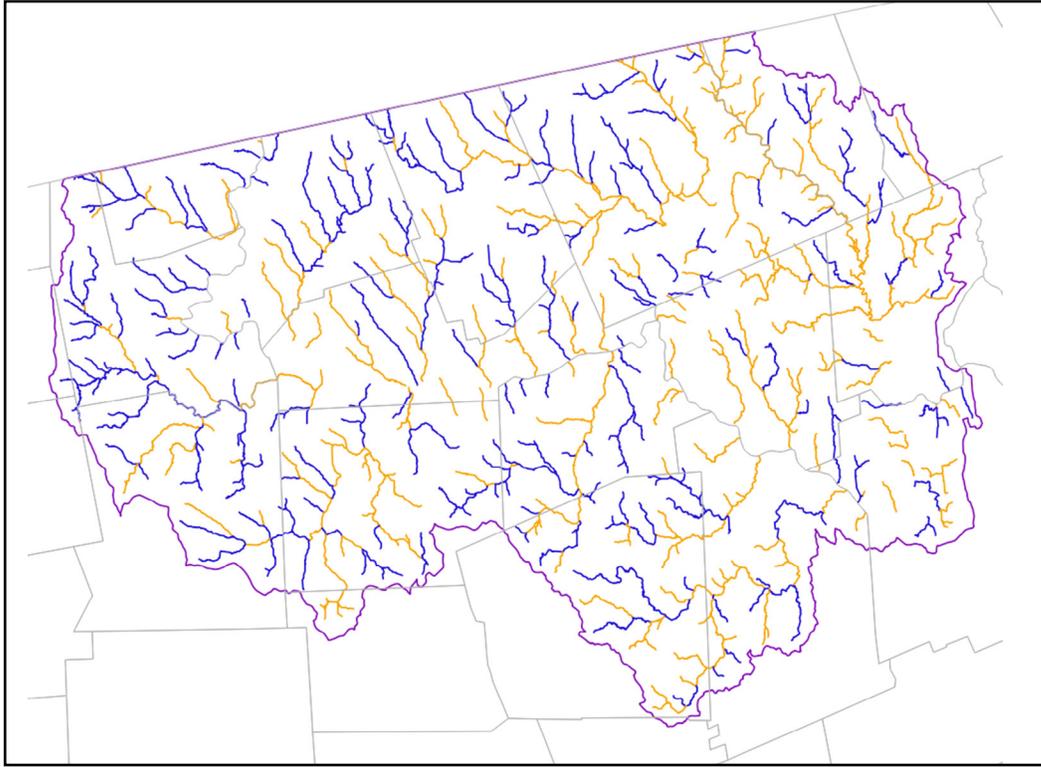
Figure 5-51: Streams with summer mean temperatures < 16°C (blue), > 16°C (orange) and not modeled (gray).

Figure 5-52: Streams with summer mean temperatures < 18°C (blue), > 18°C (orange) and not modeled (gray).

Figure 5-53: Streams with summer mean temperatures < 20°C (blue), > 20°C (orange) and not modeled (gray).

5.5.1.3 DATA SOURCES

Temperature data from the 13-state, North Atlantic region of the U.S. (Table 5-26) were used, along with a variety of GIS data (Table 5-27), to create and test the model.

5.5.1.4 SUMMARY OF RESULTS

Overall, the model fit the temperature data quite well. The estimated error was 0.59°C when all data were used (Figure 5-48) and 2.03°C when applied to data held out for validation (Figure 5- 49).

Table 5-28 contains a summary and description of derived metrics for each stream reach summarized for predictions from 1980-2015. Four of these derived metrics are summarized in Figure 5-50. The mean number of days over 18°C and 22°C were calculated only for predictions in the middle 194 days of the year to avoid problems outside the synchronized period of the year while keeping the length consistent among reaches across the region. The model results are summarized in Table 5-29.

The temperature model was used to identify stream segments that would qualify as cold water streams based on mean summer temperature thresholds of 14°C, 16°C (Figure 5-51), 18°C (Figure 5-52), 20°C (Figure 5-53) and 22°C. These results were used in a special application of Critical Linkages for assessing aquatic connectivity restoration potential for cold water streams (see Section 5.5.3). Figure 5-54 shows how the amount and distribution of cold water streams, defined using a 16°C threshold, would

change in the Deerfield River watershed with a 2°C rise in water temperatures.

5.5.2 HABITAT QUALITY - BROOK TROUT OCCUPANCY.

5.5.2.1 APPROACH

The USGS Conte Anadromous Fish Research Laboratory developed an occupancy model for brook trout based on presence/absence data from agencies and landscape data housed in the SHEDS web environment (ecosheds.org). The aim of the model was to provide predictions of occupancy (probability of presence) for catchments smaller than 200 km² in the northeastern US from Virginia to Maine. Predictions were made under current environmental conditions and for future increases in stream temperature.

5.5.2.2 METHODS - WHICH MODELS/METHODS WERE USED, AND WHY

The probability of brook trout occupancy in stream reaches was estimated using a logistic mixed effects model that included landscape, land-use, and climate variables. We also added terms to the model to account for spatial covariance - the possibility that estimates of occupancy were similar due simply to proximity. These terms are called random effects. They allowed estimates for catchments within HUC10 basins to be similar to each other (derived from a common statistical distribution). Fish data came from state and federal agencies that sample streams for brook trout as part of regular monitoring. A stream was considered occupied if
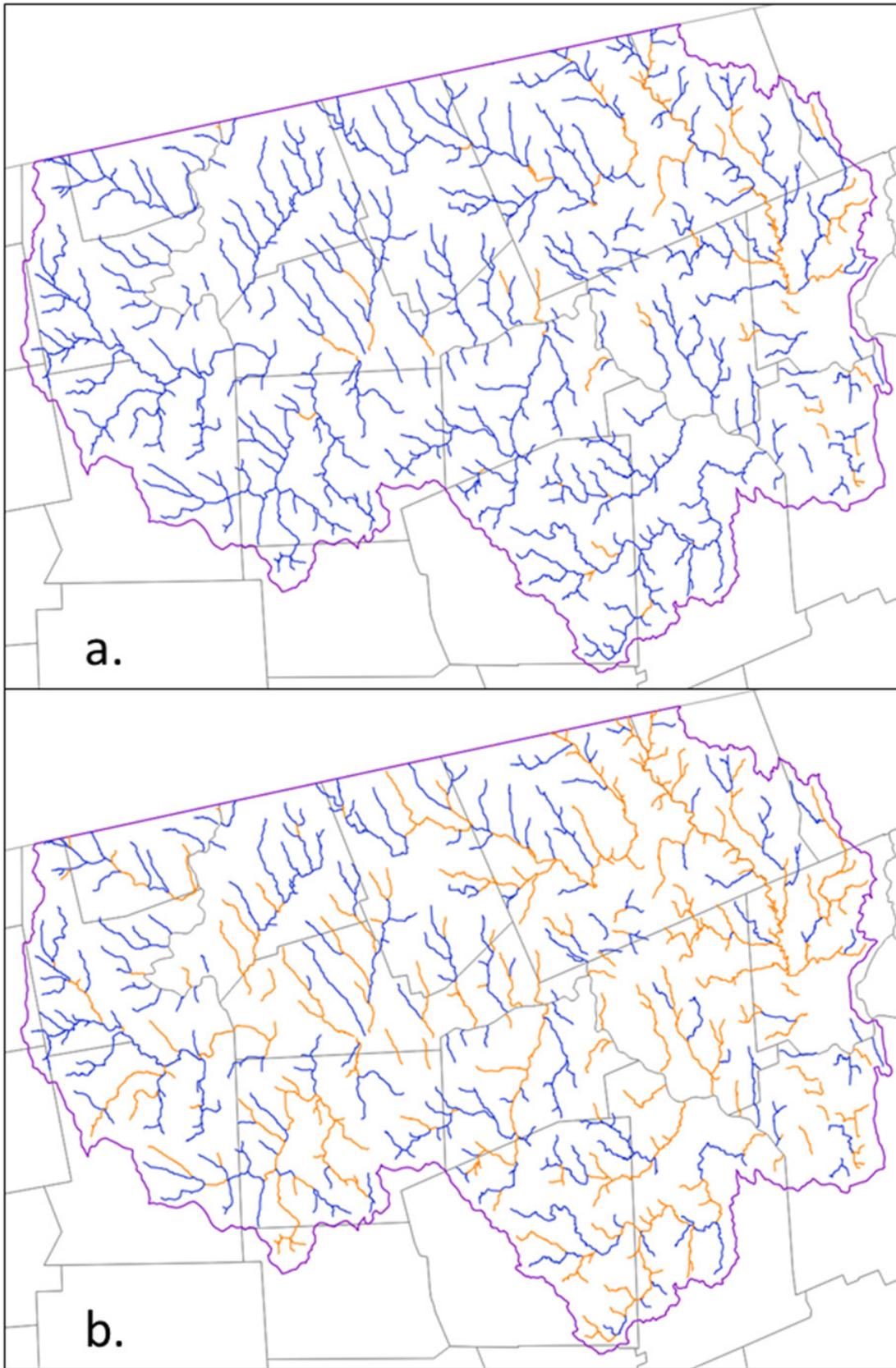
Figure 5-54: Streams meeting a summer mean temperature threshold of 16°C now (a) and with a 2°C rise in water temperatures (b).

brook trout had ever been caught during an electrofishing survey between 1991 and 2010.

Environmental data were characterized and used as predictor variables to predict brook trout occupancy. Model predictions are based on presence/absence data but expressed as the probability of occupancy. Therefore, it is difficult to evaluate how well the model predicts the data. The probabilities of occupancy must be converted to presence-absence data for comparison. This is done over a range of thresholds (= cutoffs). A threshold is the probability above which the stream is assumed to be occupied (brook trout = present). For example, if the probability of occupancy for a stream is 0.45 and we set a threshold = 0.50, the stream would be assigned as unoccupied (absent). However, if a threshold of 0.4 was used, then this same stream would be assigned as occupied (present). If the true (observed) state of the stream was occupied, then using a threshold of 0.5 would result in a false absence (predicted absent when brook trout are really present) but if a threshold of 0.4 was used, the stream would be correctly assigned as occupied (true positive). Assigning a threshold is a balance of trade-offs between false positives and false negatives. The balance is based on the risk tolerance to the consequences of type I and type II errors. False positive and false negative rates were examined and the Area Under the Receiver Operating Characteristic ROC curve (AUC) was used to assess the model fit (Zipkin et al. 2012).

To assess the model's predictive power, data from 1933 stream reaches that were withheld from model fitting were used. The term "fitted data" is used to refer to the data used to fit (estimate) the model. For comparison, others use the terms "training data" or "calibration data" synonymously. Validation data are the independent data withheld from model fitting for the purpose of understanding how well a model predicts to unobserved space and time. To evaluate this predictive power, the false positive rate (1-specificity) was plotted vs. the true positive rate (sensitivity) and the AUC was calculated.

The AUC when predicting for the validation data was 0.75, which indicates that the model has good ability to discriminate between occupied and unoccupied stream reaches for locations without survey data.

Estimates of the probability of occupancy for each catchment with increases in stream temperature of either 2, 4 or 6°C are provided. This was done by simply increasing input values for mean July stream temperature by 2, 4, or 6°C and estimating occupancies. Maps of current and predicted future occupancies are available at ice.ecosheds.org.

5.5.2.3   DATA SOURCES
Brook trout presence-absence data from the 13-state, North Atlantic region of the U.S. (Table 5-30), along with a variety of GIS data (Table 5-31), were used to create and test the model.

5.5.2.4   SUMMARY OF RESULTS
Results of the probability of brook trout occupancy model for the Deerfield River watershed are shown in Figure 5-55. These data are included in the spatial data viewer (see Section 7.3) so that users
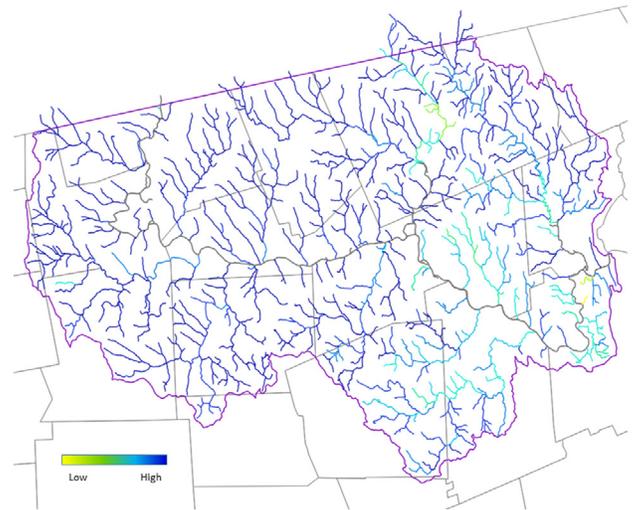
Figure 5-55: Probability of brook trout occupancy modeled for the Deerfield River watershed under current conditions.

can evaluate the effects of road-stream crossings on brook trout populations using both aquatic connectivity (see 5.5.3) and habitat quality.

5.5.3   DISRUPTION OF CONNECTIVITY

5.5.3.1   APPROACH
To evaluate how road-stream crossings disrupt aquatic connectivity, data were collected and analyzed using a variety of existing methodologies.

- Aquatic passability was assessed at road stream crossings using a protocol developed by the UMass Stream Continuity Project (a precursor to the NAACC).

- Data on aquatic passability were fed into CAPS models and assessed using Critical Linkages to evaluate the impact of crossings on aquatic connectivity and the potential for restoring ecological integrity via crossing upgrades or replacements.

- Using data on stream temperature developed as part of the Deerfield Project (see Section 5.5.1) a customized version of Critical Linkages was developed and implemented to assess the impact of crossings on aquatic connectivity for cold water streams. The cold water Critical Linkages analyses were funded by a separate project (a USFWS Hurricane Sandy Recovery and Mitigation grant); data from these analyses were included in the calculation of Ecological Disruption (see Section 6.2) and in the spatial data viewer (see Section 7.3).

5.5.3.2   METHODS - WHICH MODELS/METHODS WERE USED, AND WHY
Data used to evaluate aquatic connectivity were collected using a field data form (see Appendix A) and protocol developed by the

Table 5-30: Summary of brook trout presence-absence data by state and locations (dependent variable).

State	Number of samples	Number of catchments	Min year	Max year	Number of years
CT	1535	1268	1991	2010	20
DE	1	1	1991	2010	20
MA	630	608	1991	2010	20
MD	225	224	1991	2010	20
ME	2167	1875	1991	2010	20
NH	12	12	1991	2010	20
NJ	9	9	1991	2010	20
NY	6461	4355	1991	2010	20
PA	3850	3804	1991	2010	20
RI	4	3	1991	2010	20
VA	422	422	1991	2010	20
VT	457	320	1991	2010	20
WV	233	233	1991	2010	20

Table 5-31: GIS data used to model brook trout probability of occurrence (predictor variables).

Variable	Description	Source	Processing
Total Drainage Area	The total contributing drainage area from the entire upstream network	The SHEDS Data project	The individual polygon areas are summed for all of the catchments in the contributing network
Riparian Forest Cover	The percentage of the upstream 200ft riparian buffer area that is covered by trees taller than 5 meters	The National LandCover Database (NLCD)	All of the NLCD forest type classifications are combined and attributed to each riparian buffer polygon using GIS tools. All upstream polygon values are then aggregated.
Precipitation	The mean of the summer daily precipitation record for the individual local catchment	Daymet Daily Surface Weather and Climatological Summaries	Daily precipitation records are spatially assigned to each catchment based on overlapping grid cells using the zonalDaymet R package
Mean July Stream Temperature	Estimated stream temperature from the SHEDS regional model	SHEDS stream temperature model	Daily stream temperature estimates were aggregated to a mean July value for each catchment
Upstream Impounded Area	The total area in the contributing drainage basin that is covered by wetlands, lakes, or ponds that intersect the stream network	U.S. Fish & Wildlife Service (FWS) National Wetlands Inventory	All freshwater surface water bodies are attributed to each catchment using GIS tools. All upstream polygon values are then aggregated.
Percent Agriculture	The percentage of the contributing drainage area that is covered by agricultural land (e.g. cultivated crops, orchards, and pasture) including fallow land.	The National LandCover Database	All of the NLCD agricultural classifications are combined and attributed to each catchment polygon using GIS tools. All upstream polygon values are then aggregated.
Percent High Intensity Developed	The percentage of the contributing drainage area covered by places where people work or live in high numbers (typically defined as areas covered by more than 80% impervious surface)	The National LandCover Database	The NLCD high intensity developed classification is attributed to each catchment polygon using GIS tools. All upstream polygon values are then aggregated.

UMass Stream Continuity Project, now the NAACC. This is a rapid assessment methodology conducted during typical low-flow conditions, and data include dimensions of crossing structures and information about substrate, water depth and velocity, evidence of scour, physical barriers, and other characteristics of road-stream crossings. Crossings are automatically scored for aquatic passability using an algorithm developed by the NAACC (scores range from zero to one) when data are entered into the NAACC database. Aquatic connectivity disruption was evaluated using the Critical Linkages methodology developed by the Landscape Ecology Lab at UMass Amherst as part of the CAPS (www.umasscaps.org). Critical Linkages is a specific application of CAPS, used to assess the effects of culvert upgrades and dam removals. It uses the CAPS dataset and one of the CAPS metrics, aquatic connectedness, to assess each crossing replacement or dam removal in turn. For more information about Critical Linkages see McGarigal et al. 2012, revised July 2017. Critical Linkages data used in the Deerfield Project were from an analysis conducted in 2016.

Description of the Critical Linkage Analysis

1. Field data were collected by TU using a road-stream crossing assessment protocol developed as part of the UMass Stream Continuity Project and scored for passability using methods developed by the NAACC (streamcontinuity.org/naacc).

2. Aquatic passability scores from step 1 served as the basis for resistance values used in a resistant kernel analysis to calculate "aquatic connectedness" in CAPS. Aquatic connectedness is a quantitative assessment of the connectivity of aquatic and wetland ecosystems. Because it is a quantitative metric is can be used in scenario analyses to model changes resulting from different management actions.

3. As part of the Critical Linkages analysis, the passability score for each crossing is individually reset to 1.0 (full passage) and the change in the aquatic connectedness metric calculated. For each cell of river/stream the change in aquatic connectedness (delta scores) was then multiplied by the index of ecological integrity (IEI) scores from CAPS to yield "effect" scores.

4. "Effect" scores are used in this project as the degree to which each crossing disrupts aquatic connectivity. It is a function of diminished connectivity (delta aquatic connectedness) and habitat quality (IEI).

A special application of Critical Linkages focusing on cold water streams was developed and implemented as part of a project funded by a USFWS Hurricane Sandy Recovery and Mitigation grant. In that analysis cold water streams defined by various mean summer temperature thresholds (14°C, 16°C, 18°C, 20°C and 22°C) were identified using stream temperature data developed as part of the Deerfield Project (see Section 5.5.1). For each temperature threshold, Critical Linkages was used to evaluate crossings and calculate the "effect" of crossing upgrades or replacements on aquatic connectivity for cold water streams. Effect is calculated as the change in aquatic connectedness for cold water stream reaches

multiplied by the IEI values for those stream reaches. All crossings were included in the analyses because it is conceivable that a stream reach that does not meet the definition of cold water (based on chosen temperature thresholds) might still be important for linking together various cold water stream reaches.

5.5.3.3   DATA SOURCES
Data from the following sources were used for the analysis of aquatic connectivity disruption.

*   Temperature data as described in Section 5.5.1.

*   Land use and ecological settings data from CAPS (see McGarigal et al. 2011, revised July 2017)

*   Aquatic Passability scores from the UMass Stream Continuity Project dataset in the NAACC database (NAACC.org). These scores range from zero (bad) to one (good). In order to maintain a consistent system of scoring for data used in the Deerfield Project and the spatial data viewer, we converted the passability scores to impassability scores using the simple mathematical function (1 – passability). The result was impassability scores that ranged from zero (good) to one (bad). The frequency distribution of impassability scores is presented in Figure 5-56 and Table 5-32. The geographic distribution of these scores is shown in Figure 5-57.

5.5.3.4   SCORING
Various scores from NAACC and Critical Linkages were modified and renamed for use in the Deerfield Project (Table 5-33) To combine the various Connectivity Restoration Potential scores into an overall Ecological Disruption score (see Section 6.2) we rescaled the Critical Linkages "effect" scores to range from zero-to-one. We found that the distribution of effect scores was heavily skewed and determined that conversion of these scores to a zero-to-one scale resulted in a very high proportion of low scores (> 98 % between 0.0 and 0.1) and only a very small fraction of scores higher than 0.5 (< 0.05 %) (Figure 5-58).

Two options were considered for creating a more useful distribution of scores: quantile rescaling and logarithmic transformation. We considered quantile rescaling to be inappropriate because it would redistribute low raw scores evenly through the zero-to-one range. Logarithmic transformation serves to compress the range of scores, creating something approaching a Gaussian distribution, without elevating low scores to unreasonable levels (Figure 5-59).

The result of log transformation prior to rescaling is that most of the crossings now fall in the middle range of disruption scores (> 82 % between 0.2 and 0.8), but with relatively few crossings standing out as being of very low or very high importance (Figure 5-59). Log transformed "effect" scores are rescaled between zero and one based on statewide results (i.e. maximum score of 1.0 is based on the highest "effect" score statewide, rather than the highest "effect" score in the Deerfield River watershed).
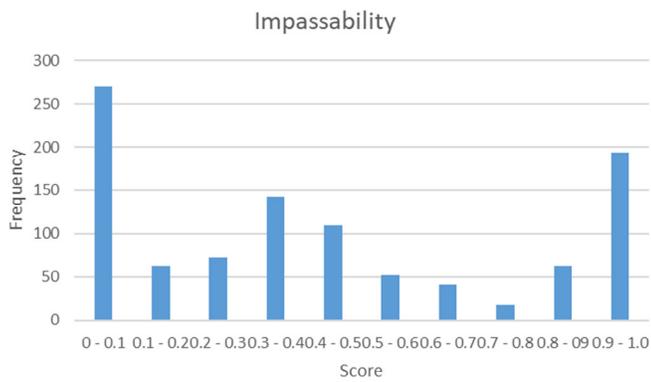
Impassability

Figure 5-56: Distribution of Impassability scores.

Table 5-32: Distribution of Impassability scores.

| Bin	Count	Percent
0 - 0.1	270	26.37%
0.1 - 0.2	63	6.15%
0.2 - 0.3	72	7.03%
0.3 - 0.4	142	13.87%
0.4 - 0.5	109	10.64%
0.5 - 0.6	52	5.08%
0.6 - 0.7	41	4.00%
0.7 - 0.8	18	1.76%
0.8 - 0.9	63	6.15%
0.9 - 1.0	194	18.95%
Total	1024	
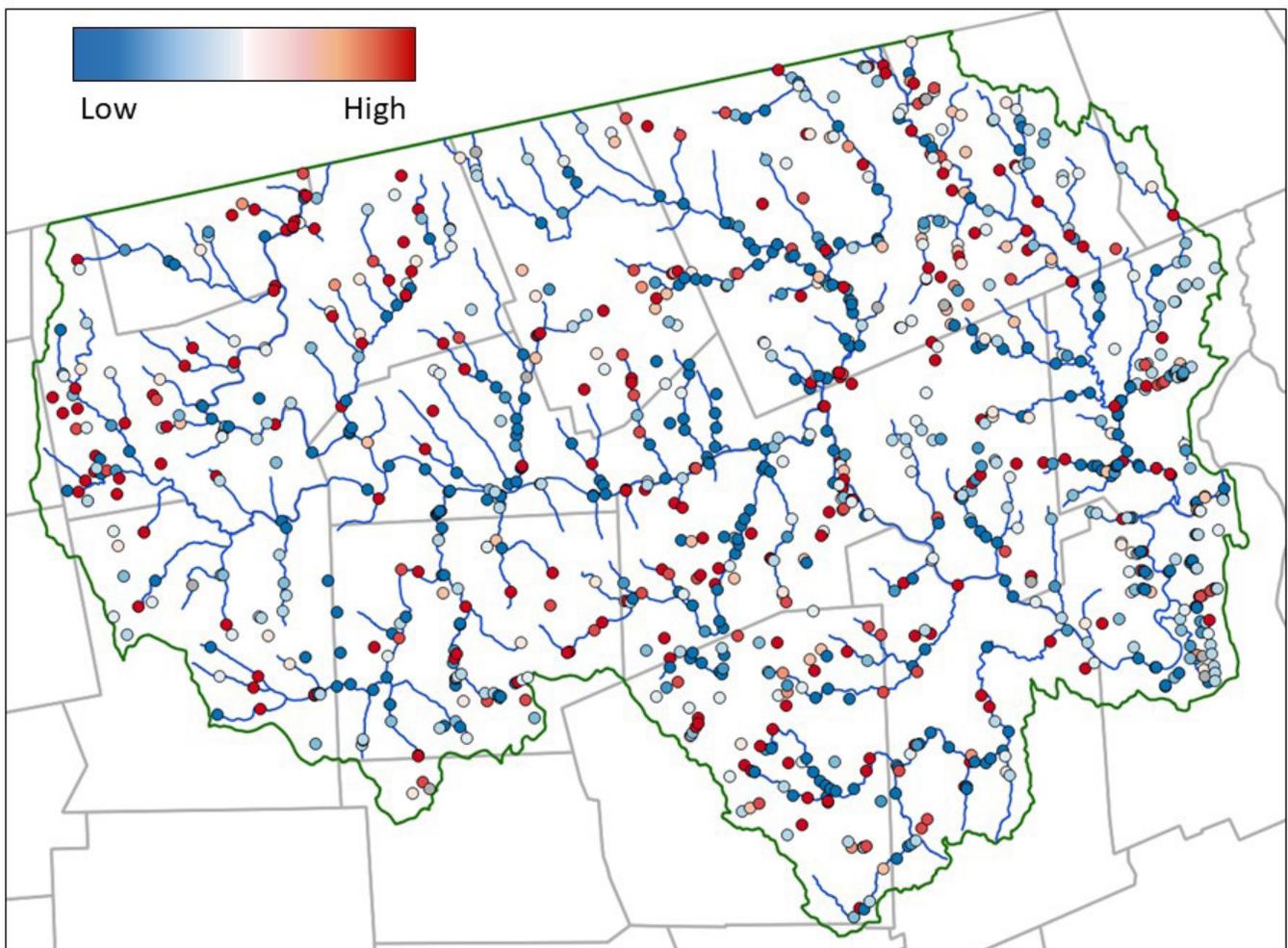
Figure 5-57: Geographic distribution of Impassability scores derived from NAACC data (1 - aquatic passability). High scores on a 0 to 1 scale are in red; low scores in blue.

**Summary of the Proposed Scoring Process using the Integrated Metric**

1. Field data were collected and scored for passability using a NAACC scoring algorithm.

2. Critical Linkages analysis used field-based passability scores from step 1 to calculate a raw aquatic connectivity disruption ("effect") score for each road-stream crossing.

3. The "effect" scores from step 2 are logarithmically transformed using the equation ln (x + 1).

4. The log transformed scores from step 3 are rescaled to range from zero (little disruption) to one (much disruption) based on statewide results.

5.5.3.5   SUMMARY OF RESULTS

A total of 834 crossings were included in the Critical Linkages analyses. The vagaries of stream mapping resulted in map blocks with varying densities of streams in the hydrography source data, presumably as an artifact of photointerpretation (Figure 5-60). It was necessary to trim stream networks using a 30 ha watershed

Table 5-33: Scores used in the Deerfield Project that were renamed from the source data.

Name used in the Deerfield Project	Original name	Comments
Impassability	Aquatic passability	NAACC Aquatic passability scores were converted to impassability scores via the equation impassability = 1 – passability.
Connectivity loss	Delta	Connectivity loss scores are Critical Linkages delta scores that have been logarithmically transformed using the equation ln (x + 1) and then rescaled to range between zero and one.
Connectivity restoration potential	Effect	Connectivity restoration potential scores are Critical Linkages effect scores that have been logarithmically transformed using the equation ln (x + 1) and then rescaled to range between zero and one.
Cold water 14°C restoration potential	Effect	Same as for connectivity restoration potential except only for cold water streams defined by a summer mean temperature of 14°C.
Cold water 16°C restoration potential	Effect	Same as for connectivity restoration potential except only for cold water streams defined by a summer mean temperature of 16°C.
Cold water 18°C restoration potential	Effect	Same as for connectivity restoration potential except only for cold water streams defined by a summer mean temperature of 18°C.
Cold water 20°C restoration potential	Effect	Same as for connectivity restoration potential except only for cold water streams defined by a summer mean temperature of 20°C.
Cold water 22°C restoration potential	Effect	Same as for connectivity restoration potential except only for cold water streams defined by a summer mean temperature of 22°C.
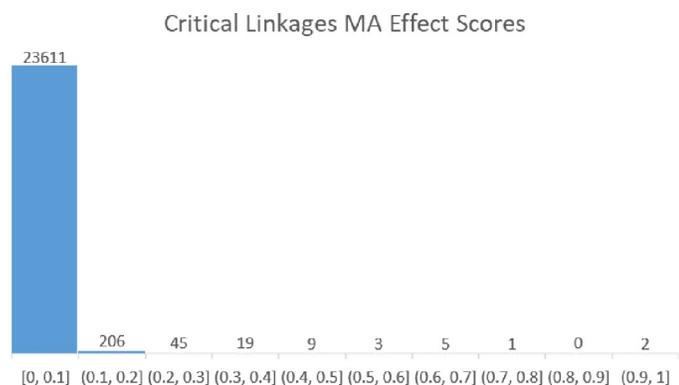
Figure 5-58: Rescaled Critical Linkages "Effect" scores for all of Massachusetts.
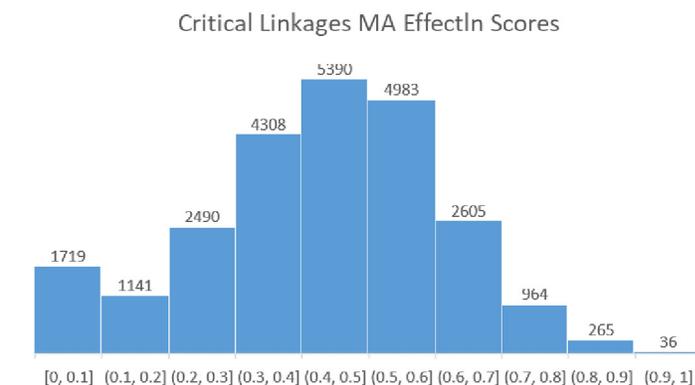
Figure 5-59: Rescaled natural log transformed "Effect" scores for all of Massachusetts.

threshold to create consistent stream densities across Massachusetts. As a result, crossings on very small streams (with watershed areas < 30 ha) were excluded from Critical Linkages analyses.

Connectivity Loss is the loss of aquatic connectivity without consideration of habitat quality. It is available in the spatial data viewer so that users can use it as is, or combine it with measures of habitat quality (e.g. stream temperature or brook trout occupancy) other than CAPS IEI. The frequency distribution of Connectivity Loss scores is presented in Figure 5-61 and Table 5-34. The geographic distribution of scores is shown in Figure 5-62.

Connectivity Restoration Potential scores are based on a standard Critical Linkages analysis of all stream reaches, combining Connectivity Loss and CAPS IEI (a measure of habitat quality). Experience suggests that these weighted scores are more useful for prioritizing restoration options than Connectivity Loss on its own. Most approaches for evaluating potential for restoring aquatic connectivity via crossing replacement or dam removal use a combination of some measure of improved connectivity (often expressed in stream miles opened up) and a measure of habitat quality. The Critical Linkages analysis uses change in the aquatic connectedness metric as the measure of improved connectivity and

CAPS IEI as a measure of habitat quality. The frequency distribution of Connectivity Restoration Potential scores is presented in Figure 5-63 and Table 5-35. The geographic distribution of these scores is shown in Figure 5-64.

The frequency distributions for cold water Critical Linkages analyses are presented in Figure 5-65 and Table 5-36 (16°C), Figure 5-67 and Table 5-37 (18°C), and Figure 5-69 and Table 5-38 (20°C). Geographic distributions of these scores are shown in Figure 5-66 (16°C), Figure 5-68 (18°C), and Figure 5-70 (20°C).

Figure 5-71 shows how crossing replacement priorities for restoring cold water stream connectivity change depending on how water temperatures change over time. Water temperatures are expected to rise over the next several decades in response to climate change. Setting priorities for long-term conservation will require decision-making not only considering current conditions but also the anticipated future distribution of cold water resources. The spatial data viewer (Section 7.3) will provide users with an opportunity to work with data reflecting both current conditions and plausible future conditions with regard to water temperature and crossing replacement prioritization.

Figure 5-60: Examples of unequal stream densities in hydrography source data, presumably due to differences in photointerpretation.

Figure 5-61: Distribution of Connectivity Loss scores.

Table 5-34: Distribution of Connectivity Loss scores.

Bin	Count	Percent
0 - 0.1	188	22.54%
0.1 - 0.2	20	2.40%
0.2 - 0.3	21	2.52%
0.3 - 0.4	29	3.48%
0.4 - 0.5	61	7.31%
0.5 - 0.6	133	15.95%
0.6 - 0.7	166	19.90%
0.7 - 0.8	134	16.07%
0.8 - 0.9	66	7.91%
0.9 - 1.0	16	1.92%
Total	834	
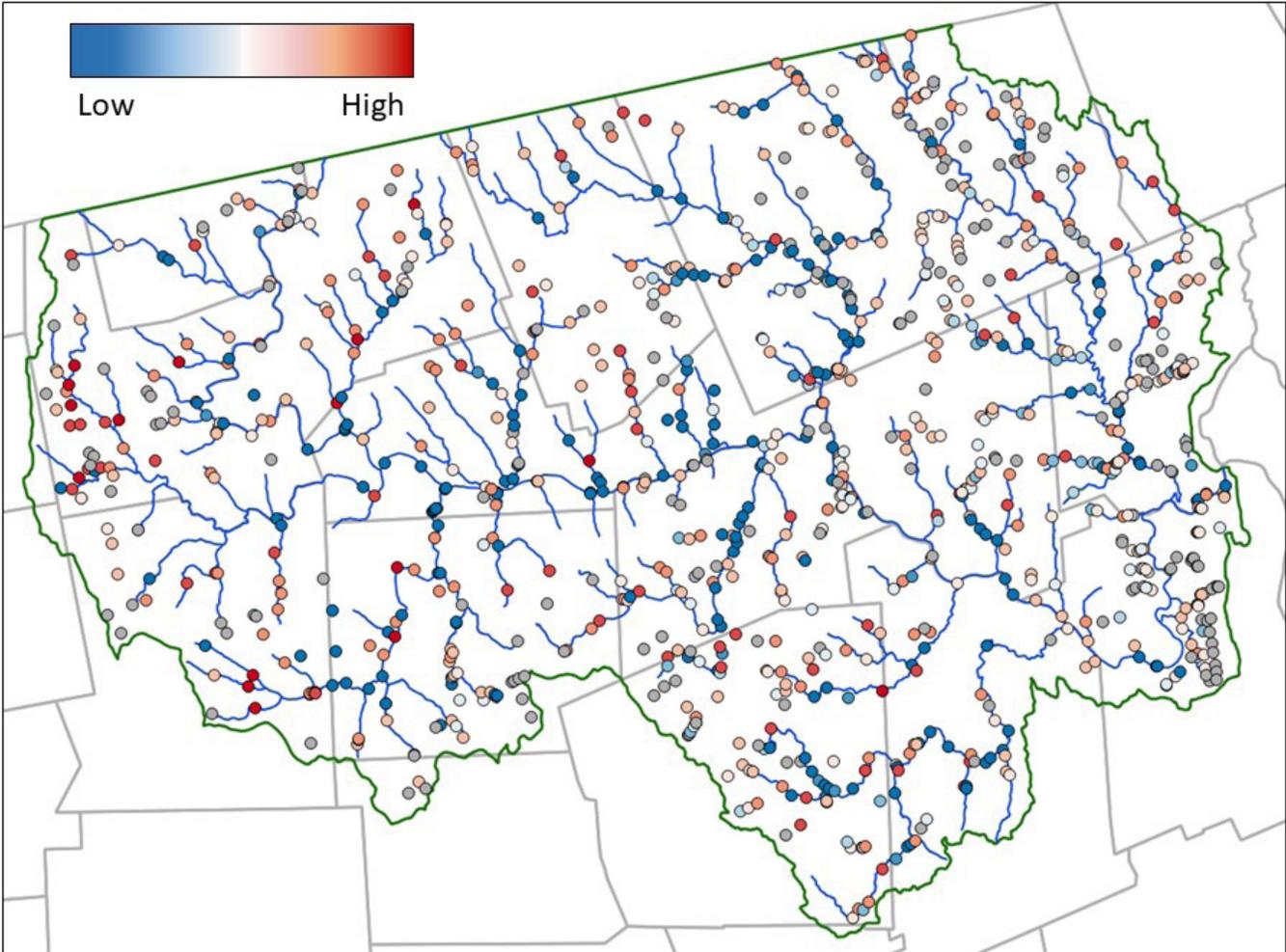
Figure 5-62: Geographic distribution of Connectivity Loss scores, calculated as the change in the CAPS Aquatic Connectedness metric rescaled to a range of 0-1. High scores are in red; low scores in blue.
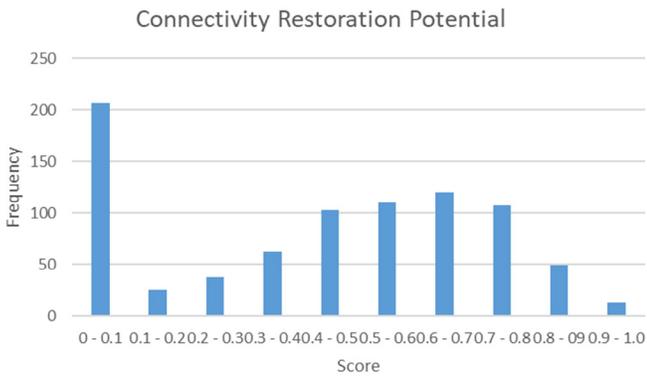
Figure 5-63: Distribution of Connectivity Restoration Potential Scores.

Table 5-35: Distribution of Connectivity Restoration Potential scores.

Bin	Count	Percent
0 - 0.1	206	24.70%
0.1 - 0.2	25	3.00%
0.2 - 0.3	38	4.56%
0.3 - 0.4	63	7.55%
0.4 - 0.5	103	12.35%
0.5 - 0.6	110	13.19%
0.6 - 0.7	120	14.39%
0.7 - 0.8	107	12.83%
0.8 - 0.9	49	5.88%
0.9 - 1.0	13	1.56%
Total	834	
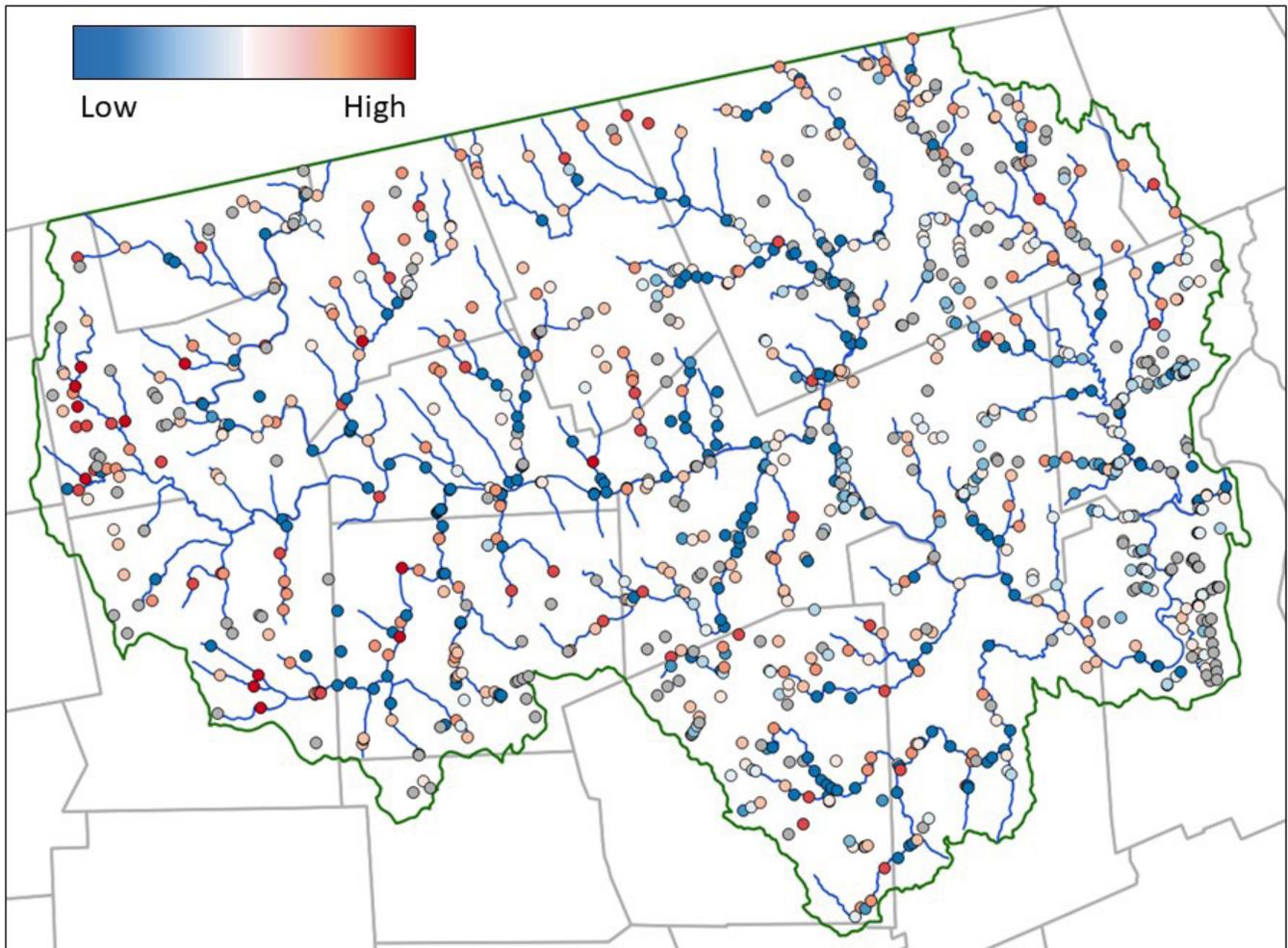
Figure 5-64:Geographic distribution of Connectivity Restoration Potential, calculated as Connectivity Loss multiplied by CAPS IEI scores rescaled to a range of 0-1. High scores are in red; low scores in blue.

Figure 5-65: Distribution of cold water (16°C) Connectivity Restoration Potential scores.

Table 5-36: Distribution of cold water (16°C) Connectivity Restoration Potential scores.

Bin	Count	Percent
0 - 0.1	416	63.61%
0.1 - 0.2	14	2.14%
0.2 - 0.3	17	2.60%
0.3 - 0.4	26	3.98%
0.4 - 0.5	41	6.27%
0.5 - 0.6	39	5.96%
0.6 - 0.7	44	6.73%
0.7 - 0.8	28	4.28%
0.8 - 0.9	20	3.06%
0.9 - 1.0	9	1.38%
Total	654	
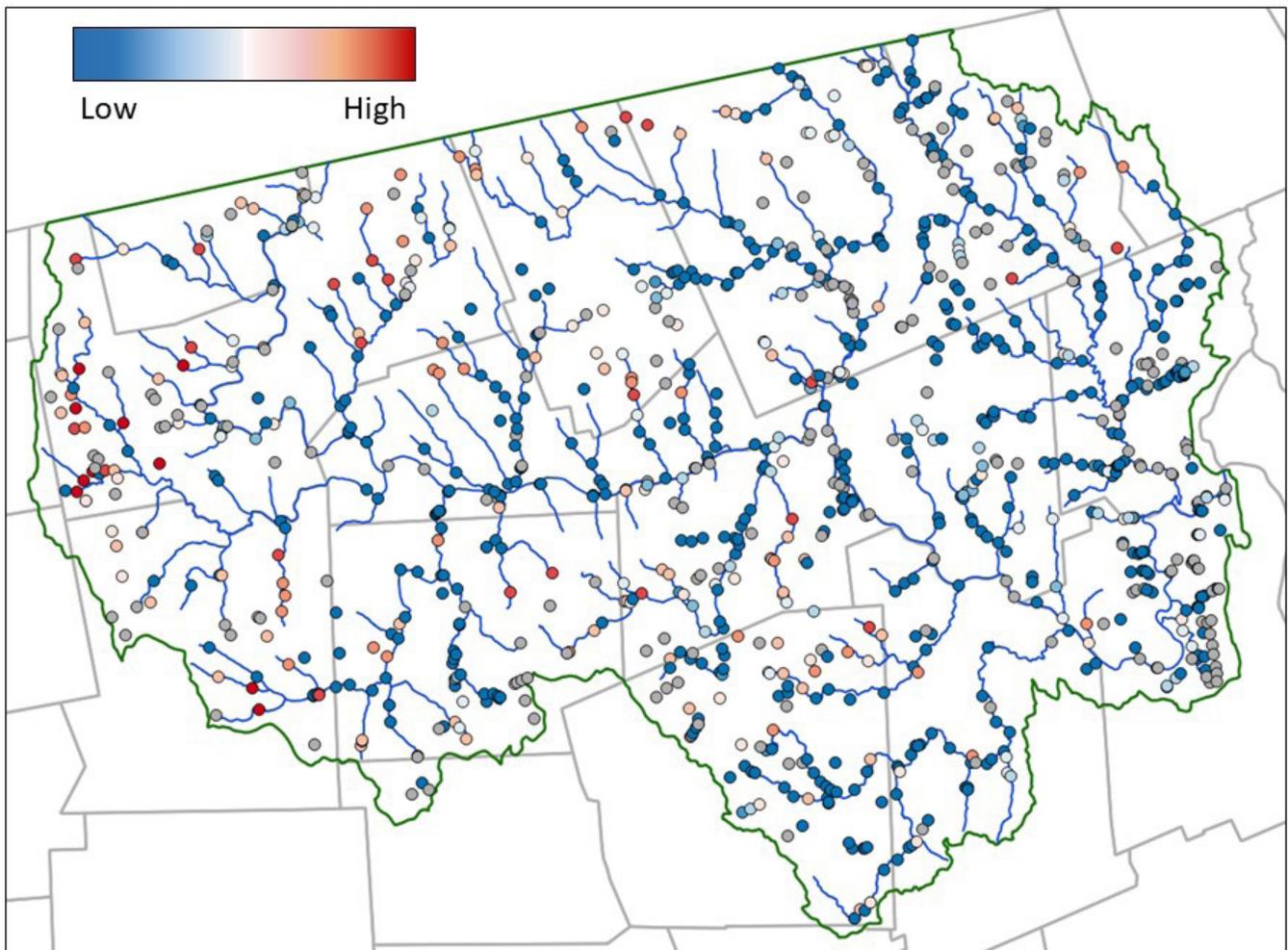
Figure 5-66: Geographic distribution of cold water Connectivity Restoration Potential using a 16°C temperature threshold. High scores on a 0 to 1 scale are in red; low scores in blue.

Figure 5-67: Distribution of cold water (18°C) Connectivity Restoration Potential scores.

Table 5-37: Distribution of cold water (18°C) Connectivity Restoration Potential scores.

Bin	Count	Percent
0 - 0.1	211	32.26%
0.1 - 0.2	18	2.75%
0.2 - 0.3	30	4.59%
0.3 - 0.4	42	6.42%
0.4 - 0.5	59	9.02%
0.5 - 0.6	73	11.16%
0.6 - 0.7	80	12.23%
0.7 - 0.8	74	11.31%
0.8 - 0.9	48	7.34%
0.9 - 1.0	19	2.91%
Total	654	
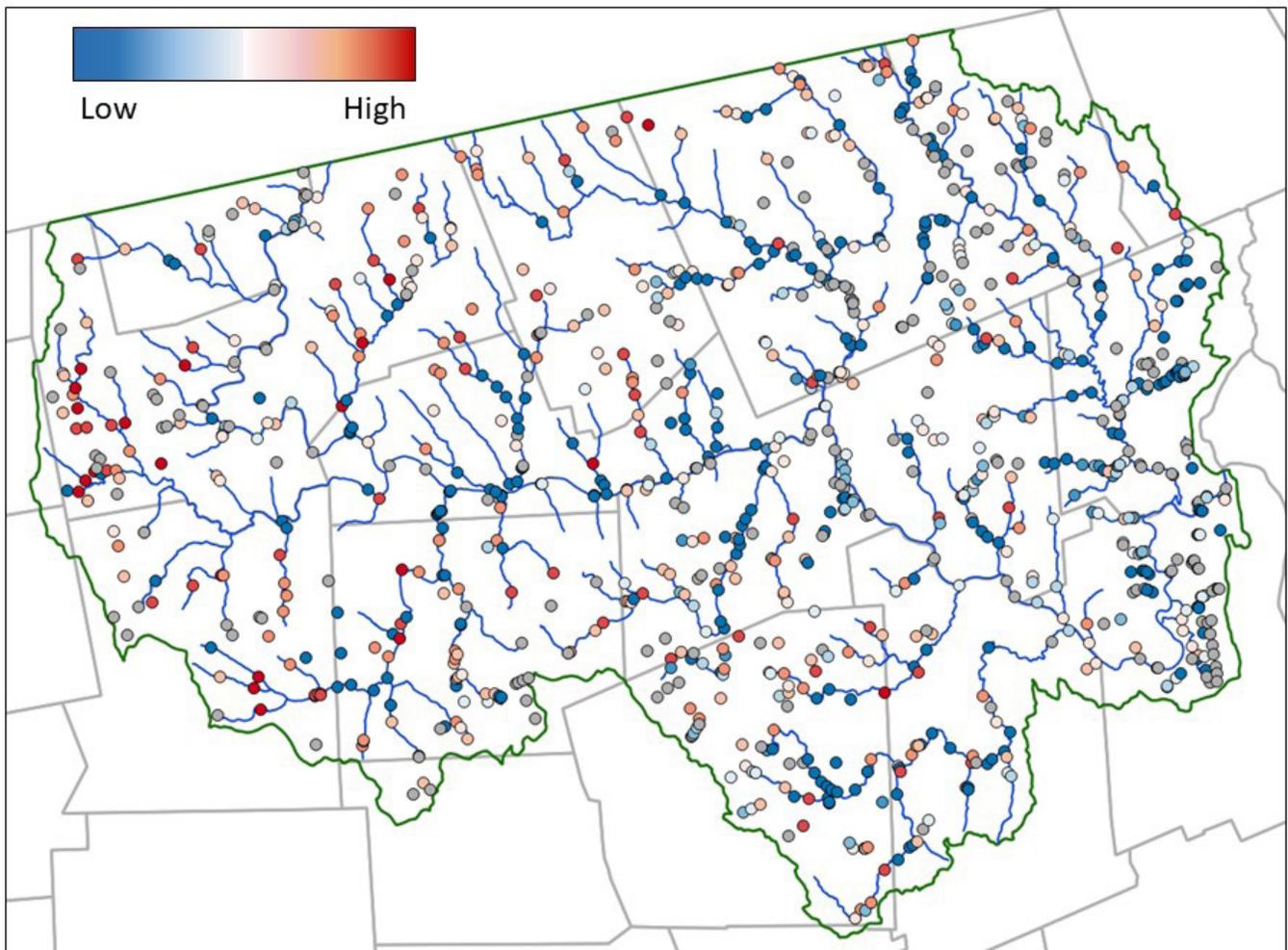
Figure 5-68: Geographic distribution of cold water Connectivity Restoration Potential using an 18°C temperature threshold. High scores on a 0 to 1 scale are in red; low scores in blue.

Figure 5-69: Distribution of cold water Connectivity Restoration Potential using a 20°C temperature threshold.

Table 5-38: Distribution of cold water (20°C) Connectivity Restoration Potential scores.

Bin	Count	Percent
0 - 0.1	145	22.17%
0.1 - 0.2	23	3.52%
0.2 - 0.3	36	5.50%
0.3 - 0.4	51	7.80%
0.4 - 0.5	72	11.01%
0.5 - 0.6	85	13.00%
0.6 - 0.7	85	13.00%
0.7 - 0.8	84	12.84%
0.8 – 0.9	54	8.26%
0.9 - 1.0	19	2.91%
Total	654	
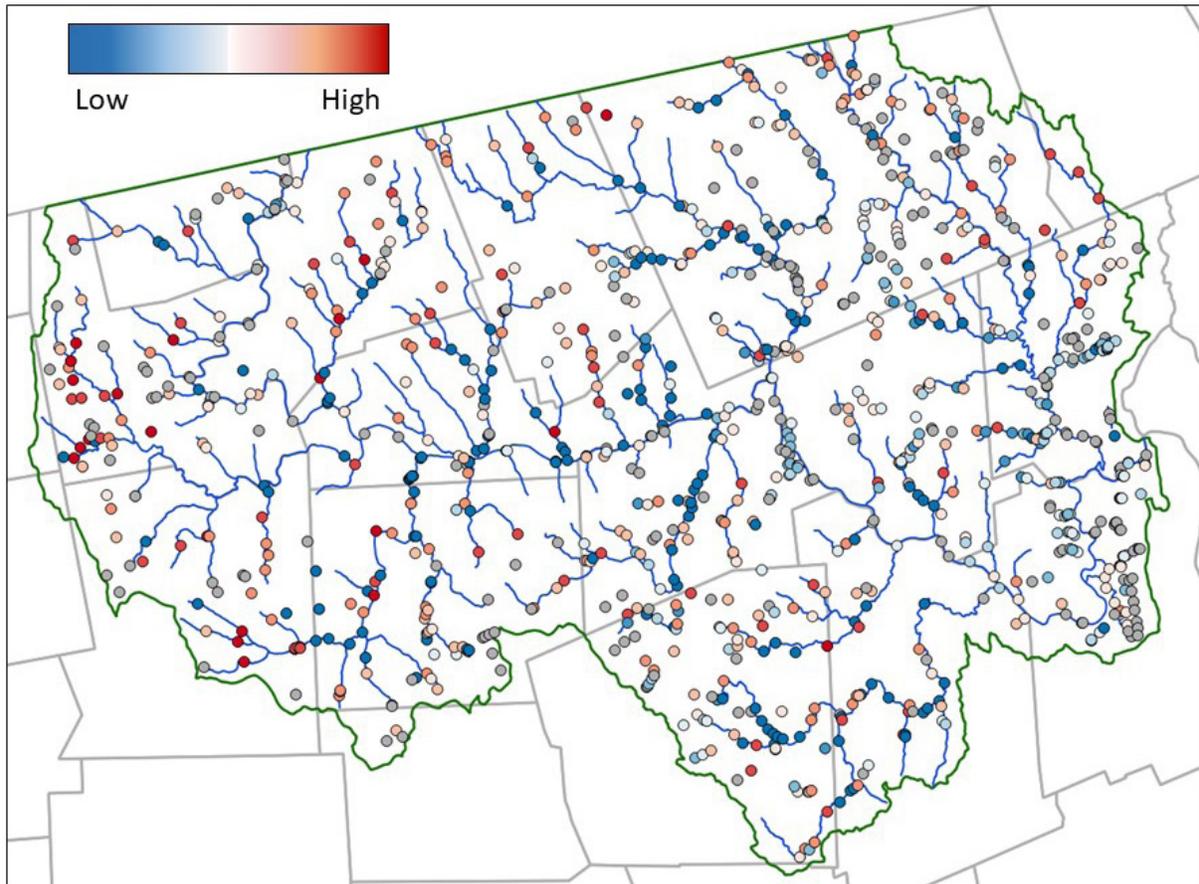
Figure 5-70: Geographic distribution of cold water Connectivity Restoration Potential using a 20°C temperature threshold. High scores on a 0 to 1 scale are in red; low scores in blue.
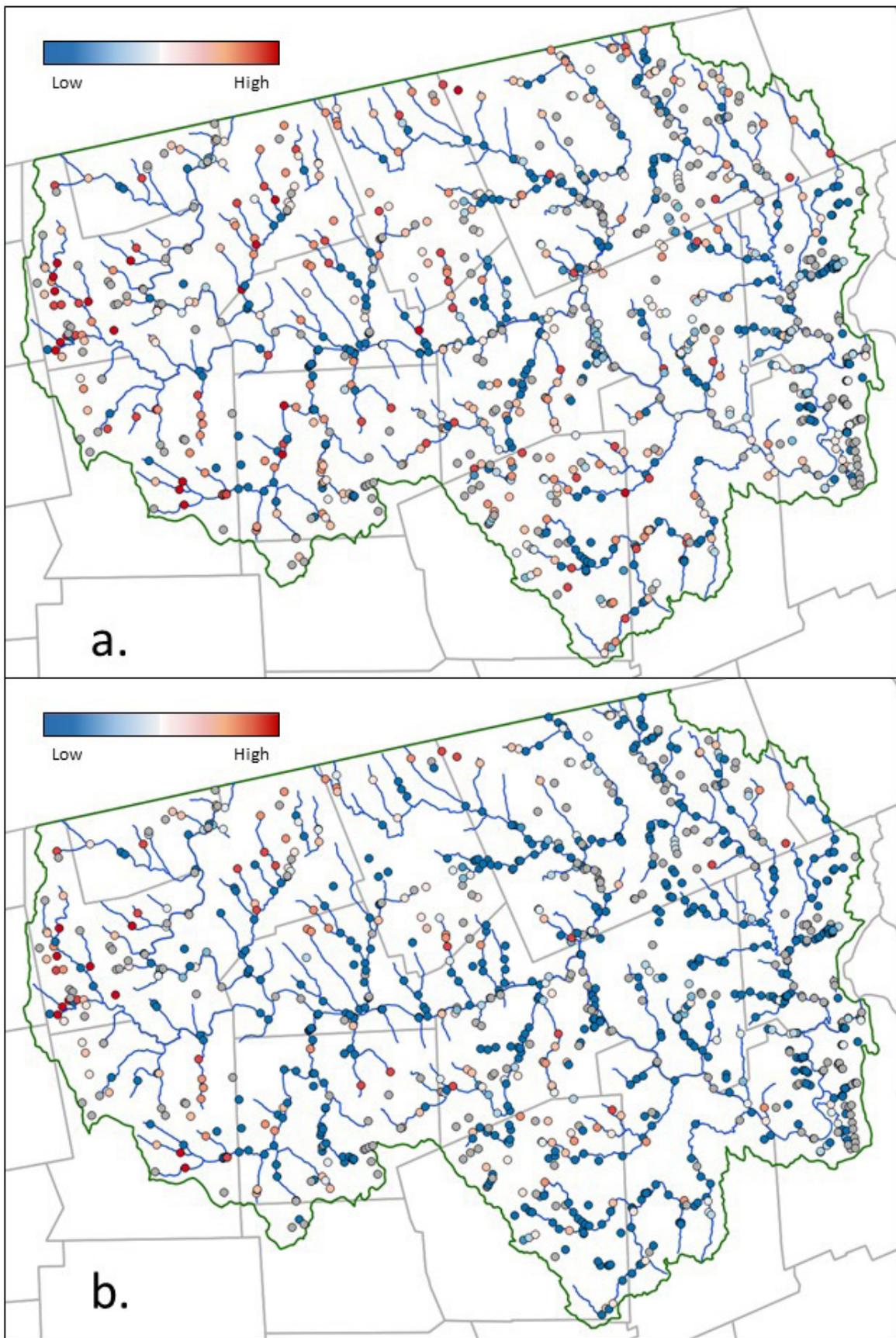
Figure 5-71: Cold water Connectivity Restoration Potential using an 18°C temperature threshold now (a) and with a 2°C rise in water temperatures (b). High scores on a 0 to 1 scale are in red; low scores in blue.

Prioritization

6.1 Overview

The Deerfield Project conducted a detailed analysis of road-stream crossings to evaluate the associated risks and vulnerabilities to ecological and transportation infrastructure in the watershed for current and future conditions, taking into account climate change. To facilitate prioritization of crossings for upgrade or replacement, scoring systems were developed for the various elements of risk and vulnerability using a scale from zero (low risk) to one (high risk).

In the transportation domain, three risk factors are combined to create an overall risk of failure score. Criticality was used as a measure of how disruptive a crossing failure would be, combining the magnitude of the disruption and the transportation system's adaptive capacity (ability to avoid or minimize impacts through alternatives or backup systems). The criticality impacts assessed as part of the Deerfield Project were delays in delivery of EMS. Risk and criticality scores were then combined to create Vulnerability scores for the Transportation domain.

In the ecological domain, we created Ecological Disruption scores as a measure of ecosystem vulnerability. Ecological disruption is based on connectivity restoration potential for 1) all streams and 2) cold water streams. Combining the Vulnerability scores for the two domains (ecological and transportation) yielded an overall priority score for each crossing.

6.2 Risk of Failure

6.2.1 APPROACH/METHODS

Overall Risk of Failure is a combination of the three component scores (structural, hydraulic and geomorphic risk) into a single score representing the overall potential for crossing failure. It is important to recognize one key difference between structural risk on one hand, and hydraulic and geomorphic risk on the other. Both hydraulic and geomorphic risk scores are based on the likelihood that a crossing will fail during or after a significant storm event, and it is assumed that there is little risk of hydraulic or geomorphic failure in the near future without a storm. The more severe the storm, the more likely that a particular structure will fail, and during a severe storm those structures with the highest risk scores are the ones most likely to fail. In contrast, structural risk is based on an assessment of bridge or culvert condition at each crossing. Although storms may increase

the risk of structural failure, crossings rated as being at high risk are at risk under all weather conditions.

6.2.2 SCORING

The Overall Risk of Failure score is derived from the three component risk scores (structural, hydraulic, and geomorphic). The overall score for each crossing was computed using the highest of the three component scores (see Sections 5.1, 5.2, and 5.3 for details on how these component scores were calculated).

Overall Risk of Failure = max [structural risk, hydraulic risk, geomorphic risk]

The rationale for this approach is that the mechanism of failure with the highest score is the most likely mechanism to cause a crossing failure. The metaphor of the weak link in the chain seems appropriate. Certain characteristics of a crossing may make it more vulnerable to failure from multiple mechanisms. For example, an undersized crossing would be more vulnerable to failure for hydraulic reasons, and may make it more likely to be plugged by sediment or woody debris and fail for geomorphic reasons. However, whichever mechanism reaches a critical stage first will determine whether a crossing fails under specified conditions. Therefore, the highest score among the three components was used as the Overall Risk of Failure for the crossing.

An Overall Risk of Failure score was calculated even if data were available for only one or two of the three metrics. If only two metrics were available then the Overall score was the maximum of the two metric scores. For crossings with only one available metric, the Overall score was the same as the score for the available metric.

6.2.3 SUMMARY OF RESULTS

830 crossings in the Deerfield River watershed received Overall Risk of Failure scores. The frequency distribution for those scores is presented in Figure 6-1 and Table 6-1. The geographic distribution of the scores is shown in Figure 6-2. A large percentage of crossings (57.5 %) had overall risk scores ≥ 0.5 (Table 6-1). Of these, 184 crossings (22.2 % of all assessed crossings) had risk scores ≥ 0.9. It is important to note that only 51 crossings (27.7 % of crossings with overall risk of failure scores ≥ 0.9 and 6.1 percent of all crossings assessed) have risk scores that high due to potential structural deficiencies (Table 6-2). Storm-related risk (hydraulic or geomorphic risk) accounts for 144 crossings with overall risk of
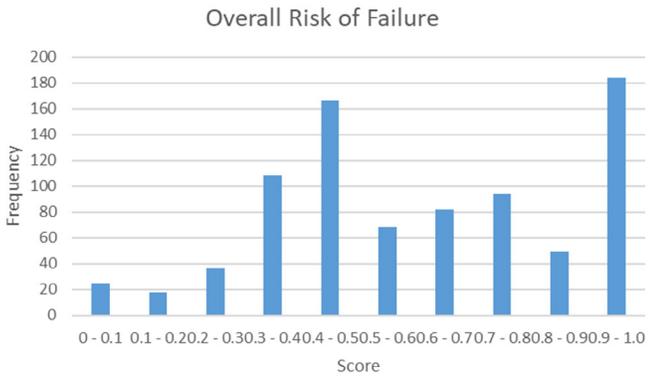
Figure 6-1: Distribution of Overall Risk of Failure scores.

Table 6-1: Distribution of Overall Risk of Failure scores.

| Bin	Count	Percent
0 - 0.1	25	3.0
0.1 - 0.2	18	2.2
0.2 - 0.3	36	4.3
0.3 - 0.4	108	13.0
0.4 - 0.5	166	20.0
0.5 - 0.6	68	8.2
0.6 - 0.7	82	9.9
0.7 - 0.8	94	11.3
0.8 - 0.9	49	5.9
0.9 - 1.0	184	22.2
Total	830	
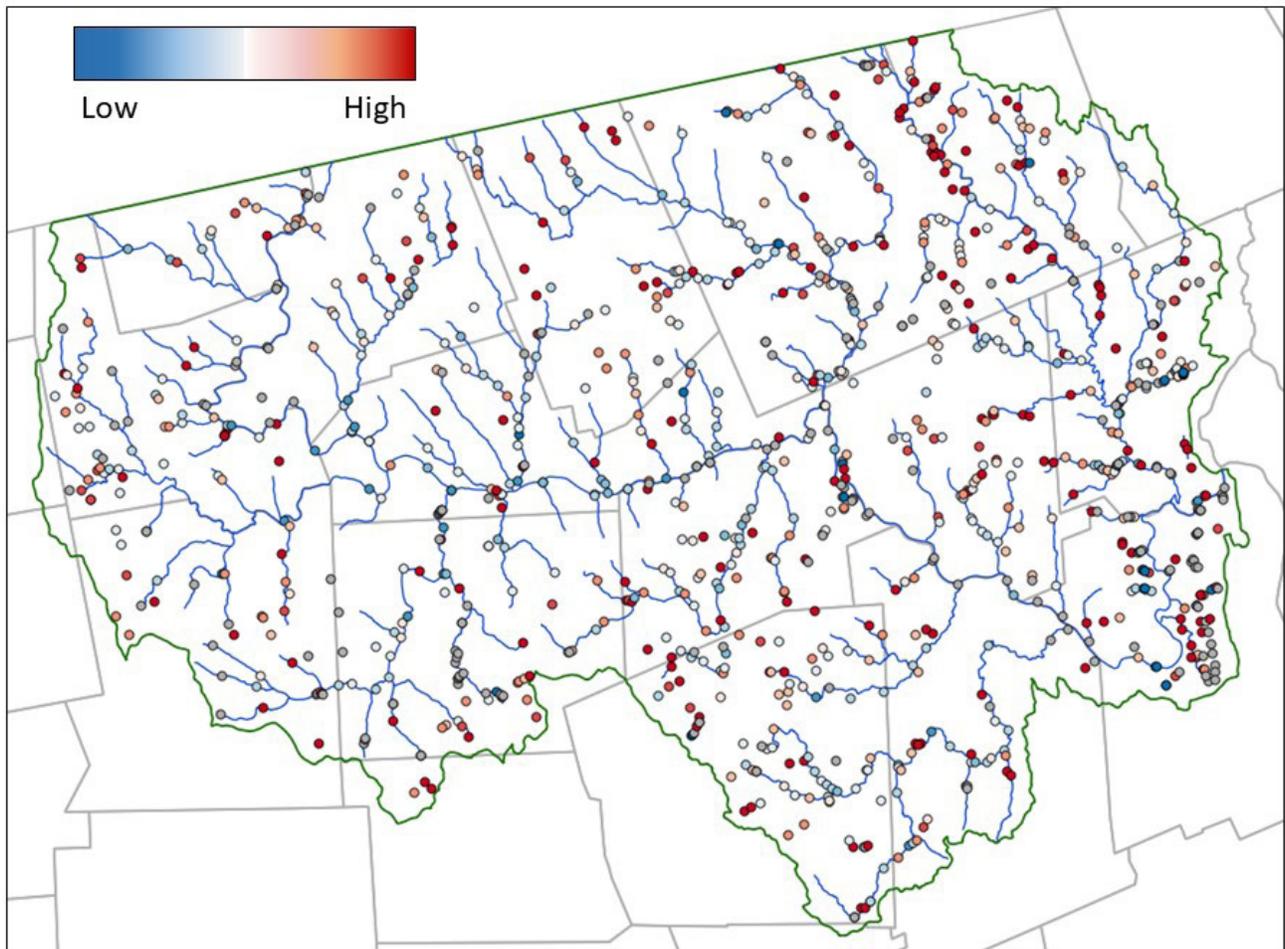
Figure 6-2: Geographic distribution of Overall Risk of Failure scores. High scores on a 0 to 1 scale are in red; low scores in blue.

Table 6-2: Component risk of failure metrics that resulted in Overall Risk of Failure scores ≥ 0.9.

Metric	# Crossings	Percent
Geomorphic risk	34	18.5
Hydraulic risk	89	48.4
Structural risk	40	21.7
Geomorphic & structural risk	3	1.6
Hydraulic & structural risk	7	3.8
Geomorphic & Hydraulic risk	10	5.4
All three failure mechanisms	1	0.5
Total	184	

failure scores ≥ 0.9. Many crossings with scores in this category (n = 107) had hydraulic risk scores ≥ 0.9. Hydraulic failures may or may not be catastrophic failures. In some cases, such a failure could result in the complete loss of a crossing; in other cases, it might result in a temporary closure due to water overtopping the road. The results of this risk of failure analysis suggest that a small percentage of crossings may be at high risk of structural failure and should be inspected by a qualified engineer in the near future. These results also suggest that a significant percentage of crossings in the Deerfield River watershed may be vulnerable to storm-related failure due to issues related to geomorphic vulnerability or hydraulic capacity. The relative scoring system used to evaluate geomorphic and hydraulic risk means that we can't say much about the probability of failure, only that there is some vulnerability. Given the expectation that climate change will produce future storms of increasing severity, as well as more frequent severe storms, these issues are likely to get worse over time.

Table 6-3: Distribution of Transportation Vulnerability scores.

Bin	Count	Percent
0 - 0.1	322	35.4
0.1 - 0.2	70	7.7
0.2 - 0.3	124	13.6
0.3 - 0.4	124	13.6
0.4 - 0.5	101	11.1
0.5 - 0.6	70	7.7
0.6 - 0.7	49	5.4
0.7 - 0.8	31	3.4
0.8 - 0.9	10	1.1
0.9 - 1.0	9	1.0
Total	910	

6.3 Transportation Vulnerability

6.3.1 APPROACH/METHODS

As defined for the Deerfield Project, transportation system vulnerability is a combination of risk and criticality. For risk, the Overall Risk of Failure was used. The aspect of criticality addressed in this project was Disruption of EMS, and Overall EMS Delay was used as the second component for calculating Transportation Vulnerability. See Sections 6.2 and 5.4 for methods used to calculate these component scores.

6.3.2 SCORING

A vulnerability score was calculated as the product of the risk and criticality scores.

Vulnerability = (Overall Risk of Failure) x (Overall EMS Delay)

As a combination of risk and criticality, Transportation Vulnerability should be highest when both component scores are high, and should be zero if either of the component scores is zero. It would not matter (vulnerability score = 0) if the risk of failure is high as long as the consequences of failure (criticality) are nonexistent. Likewise, a crossing where criticality is high but there is essentially no risk of failure (risk score = 0) should have a Transportation Vulnerability score of zero. If data were available for only one of the two scores then no Vulnerability score was calculated, unless the available score was zero. In these cases, the Vulnerability score was also zero.

6.3.3 SUMMARY OF RESULTS

Transportation Vulnerability scores were computed for 910 crossings. The frequency distributions for those scores are presented in Figure 6-3 and Table 6-3. The geographic distribution of scores is shown in Figure 6-4. Just over 35 percent of crossings had low Transportation Vulnerability scores (scores ≤ 0.1) indicating that they had low scores for either risk or criticality, or both. It is important to remember that these scores are based on a measure of criticality that only considered EMS. It is likely that some of these low scoring crossings could score higher if other elements of criticality were included. The other 65 percent of crossings have scores that are distributed throughout the scoring range, providing a relative ranking that can be easily used for transportation decision-making.
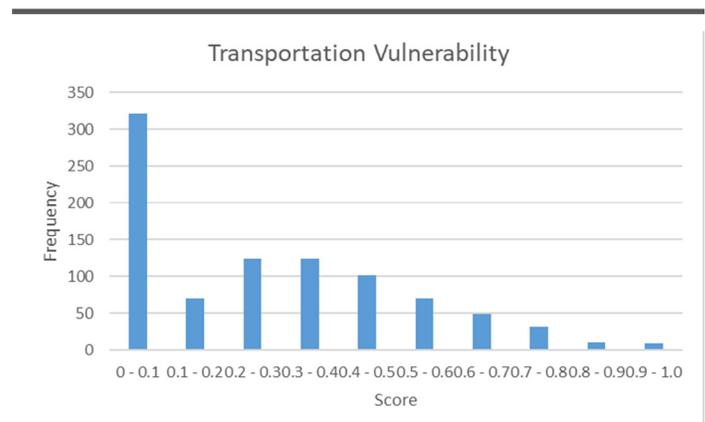
Figure 6-3: Distribution of Transportation Vulnerability scores.
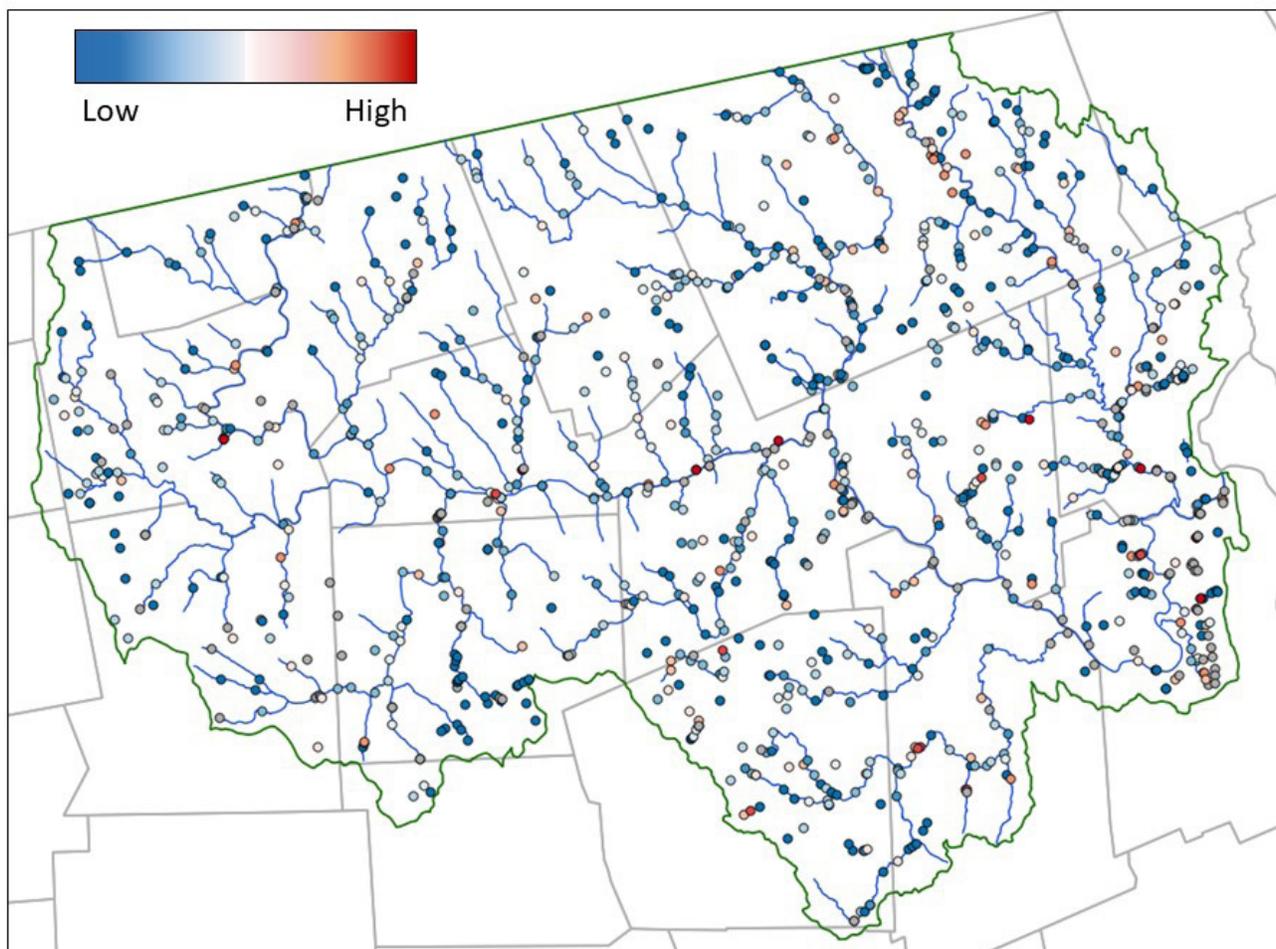
Figure 6-4: Geographic distribution of Transportation Vulnerability scores. High scores on a 0 to 1 scale are in red; low scores in blue.

6.4 Ecological Disruption

6.4.1 APPROACH/METHODS

Ecological Disruption takes into account the potential for restoring aquatic connectivity via crossing replacement for all streams (regardless of temperature) and cold water streams (defined as having mean summer temperature ≤ 16°C). The components for the Ecological Disruption score are Connectivity Restoration Potential and cold water (16°C) Restoration Potential. See Section 5.5.3 for methods used to calculate these component scores.

6.4.2 SCORING

The score for Ecological Disruption is simply the maximum of the two component scores.

Ecological Disruption = max [Connectivity Restoration Potential, cold water (16°C) Restoration Potential]

The rationale for this scoring approach is that both ecological objectives are equally valid and therefore, the highest score between the two components was assigned as the Ecological Disruption score for the crossing. If data for only one of the two component metrics were available, then the Ecological Disruption score was the same as the score for the available metric.

6.4.3 SUMMARY OF RESULTS

The distribution of Ecological Disruption scores is presented in Figure 6-5 and Table 6-4. This distribution lends itself well to conservation decision-making, with about 25 percent of crossings offering little restoration potential and the remaining crossings well-distributed throughout the scoring range.

The geographic distribution of these scores is shown in Figure 6-6. Ecological Disruption scores were dictated by Connectivity Restoration scores from the general Critical Linkages analysis (79.5 %) with 20.5 percent dictated by cold water Critical Linkages scores (Table 6-5).

Table 6-4: Distribution of Ecological Disruption scores.

Bin	Count	Percent
0 - 0.1	206	24.70%
0.1 - 0.2	22	2.64%
0.2 - 0.3	41	4.92%
0.3 - 0.4	61	7.31%
0.4 - 0.5	100	11.99%
0.5 - 0.6	110	13.19%
0.6 - 0.7	116	13.91%
0.7 - 0.8	110	13.19%
0.8 - 0.9	52	6.24%
0.9 - 1.0	16	1.92%
Total	834	
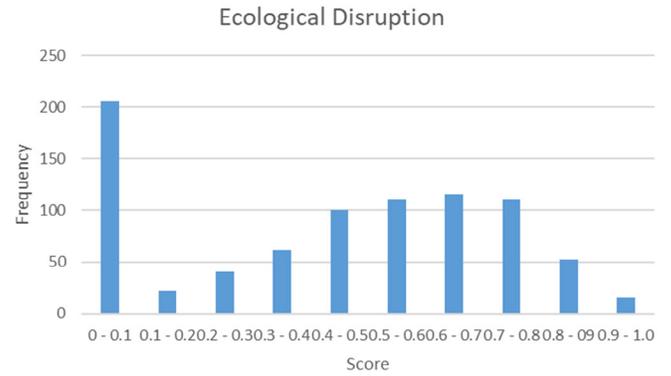
Figure 6-5: Distribution of Ecological Disruption scores.
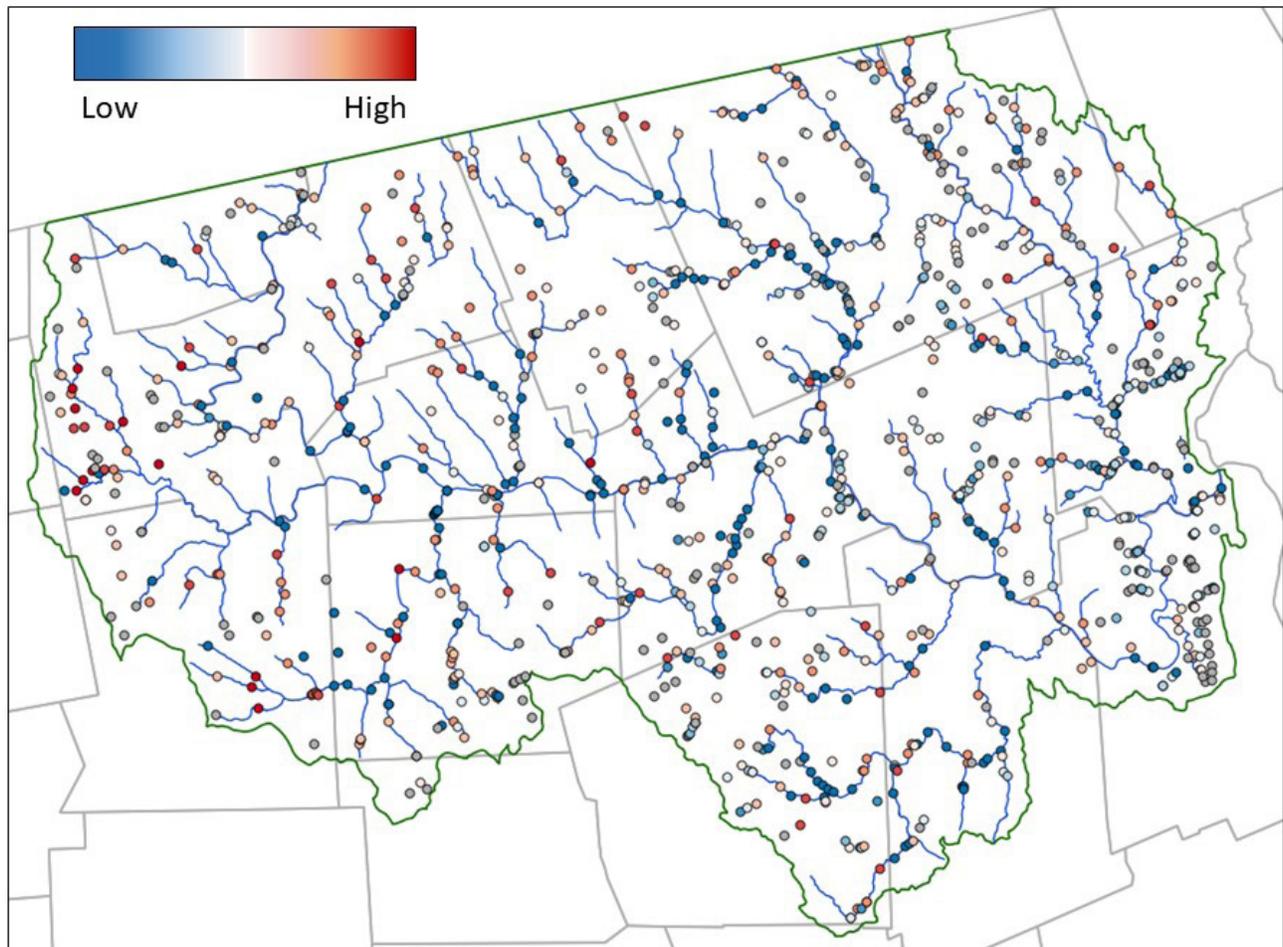
Figure 6-6: Geographic distribution of Ecological Disruption scores. High scores on a 0 to 1 scale are in red; low scores in blue.

Table 6-5: Connectivity restoration metrics that determined overall Ecological Disruption scores.

Metric	# Crossings	Percent
Connectivity Restoration Potential	663	79.50
Cold water (16°C) Restoration Potential	171	20.50
Total	834	100.00

Table 6-6: Distribution of Crossing Priority scores.

Bin	Count	Percent
0 - 0.1	49	6.36%
0.1 - 0.2	61	7.92%
0.2 - 0.3	79	10.26%
0.3 - 0.4	65	8.44%
0.4 - 0.5	106	13.77%
0.5 - 0.6	136	17.66%
0.6 - 0.7	126	16.36%
0.7 - 0.8	102	13.25%
0.8 - 0.9	35	4.55%
0.9 - 1.0	11	1.43%
Total	770	

6.5 Crossing Priority

6.5.1 APPROACH/METHODS

The Crossing Priority score was designed to combine vulnerability scores from the two domains (ecological and transportation) used in this analysis. The component metrics used for this calculation are Ecological Disruption (see Section 6.4) and Transportation Vulnerability (see Section 6.3).

6.5.2 SCORING

The Crossing Priority score was calculated as the maximum score added to an average of both scores, rescaled to range from zero to one.

Raw Priority = (max [Ecological Disruption, Transportation Vulnerability]) + (average [Ecological Disruption, Transportation Vulnerability])

Crossing Priority (rescaled) = Raw Priority ÷ max [Raw Priority]

The objective is to create a combined score that ensures that a high score for one of the domains is not canceled out by a low score for the other domain, and that can identify as high priorities those crossings that rate highly in both domains. If data were available for only one of the two scores then no Priority score was calculated.

6.5.3 SUMMARY OF RESULTS

A total of 770 crossings received priority scores. The frequency distribution of the scores is presented in Figure 6-7 and Table 6-6. The scores are well distributed through the range, which is important for priority setting and decision making. The geographic distribution of scores is shown in Figure 6-8.


Figure 6-7: Distribution of Crossing Priority scores.
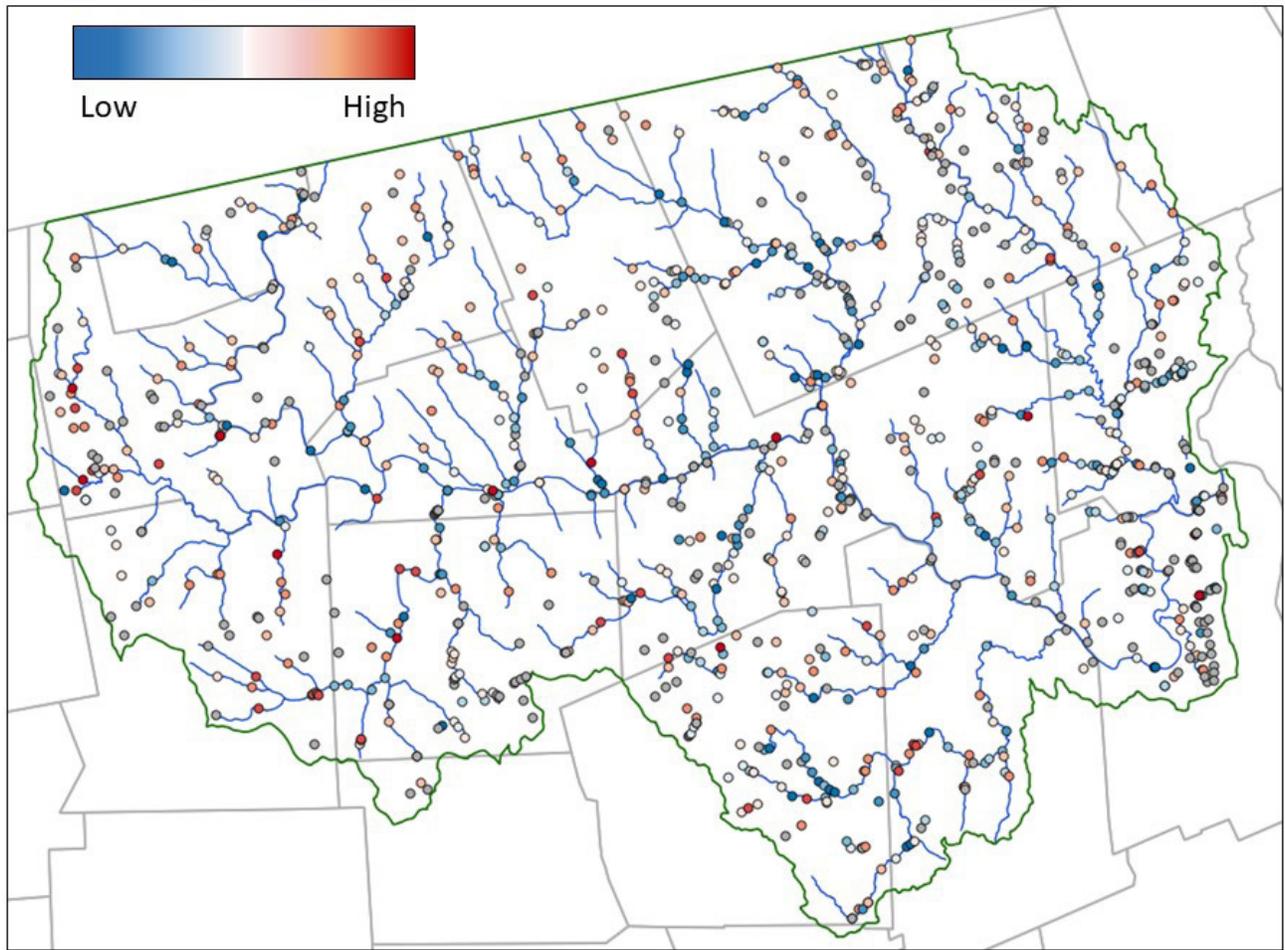
Figure 6-8: Geographic distribution of Crossing Prioritization scores. High scores on a 0 to 1 scale are in red; low scores in blue.

Project Deliverables

7.1 Database

A large amount of very diverse data have been collected or generated over the course of the Deerfield Project. The majority of these files have been organized in a single, relational Access database. The database includes raw data, such as the data collected by TU on stream continuity and culvert condition, climate projection data, and derived data, such as return interval flows and $Q_{critical}$ results for each crossing. A full list of the data included in the Access database is provided in Appendix O. Some additional data are provided as separate zipped Excel files, as also detailed in Appendix O. All electronic files were provided separately to MassDOT.

7.2 GIS files

GIS files are also provided as part of the project database. These were described previously in Section 3.5, and a complete listing is provided in Appendix O. Metadata for the Deerfield Project were prepared for all spatial data layers, rasters and summary spreadsheets. Metadata were created using the Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata (FGDC-STD-001-1998) and are provided in two formats, xml and html. Xml formatted metadata were created and error-checked in XML- Notepad and converted to html format in ArcGIS 10.4.x using the USGS metadata wizard stylesheet. Note: not all xml metadata were converted to html format. The conversion needs to be done after MassDOT determines where the data from the Deerfield Project will be housed. Without knowing where the data will be housed, the onlink[33] field in the metadata that tells the user where to find the data cannot be filled in.

7.3 Data Viewer

A data viewer was created to visualize and deliver some of the results of the Deerfield Project as well as a related project undertaken with Department of Interior Hurricane Sandy mitigation funds. This viewer, the Stream Crossings Explorer (SCE), can be seen at: http://sce.ecosheds.org.

The SCE is a data visualization and decision support tool that was developed to assist with locating and prioritizing stream crossings that meet user-defined criteria. SCE conveys information related to aquatic connectivity and ecological aspects of river systems as well as risk of failure of road-stream crossings and associated disruption of emergency services. This tool was designed to be used by state and municipal agencies, local decision-makers, regional planners, conservation organizations, and natural resource managers.

Data for this tool come from a variety of sources and were developed in partnership with other efforts, including CAPS and the NAACC. The SCE currently focuses on the Deerfield River watershed, but continues to be updated and will expand to the 13-state region with limited data. This tool is based on the Interactive Catchment Explorer (ICE) and is part of the Spatial Hydro-Ecological Decision System (SHEDS), a collection of innovative data visualization and decision support tools for exploring and better understanding dynamic relationships in stream ecosystems.

The ecological component of the Deerfield Project used existing assessment protocols and applied existing connectivity models to evaluate how each crossing affects aquatic connectivity. As a base assessment, the existing Critical Linkages model was run using pre-existing NAACC data and data collected as part of this project, and the results were used to score crossings for connectivity restoration potential. We also ran a specialized version of Critical Linkages focusing on connectivity for cold water streams. These analyses used a variety of temperature thresholds (14°C, 16°C, 18°C, 20°C, and 22°C) to define "cold water streams," and cold water connectivity restoration potential was scored using each of these thresholds. The SCE contains data from both the base analysis and the cold water analyses, allowing users to focus on connectivity for all streams or for cold water streams, as well as assess priorities for crossing replacement based on current or future temperatures. SCE users will have the flexibility to prioritize crossings for replacement depending on whether the objective is to restore connectivity for all streams (Connectivity Loss), for high-quality (high IEI) streams (Connectivity Restoration Potential), or cold water streams (Cold water Restoration Potential).

SCE is currently supported on the latest versions of all major web browsers. Google Chrome is highly recommended for the best user experience. SCE is not intended for use on mobile devices. To achieve feature filtering in a highly responsive way, SCE was developed as a client-side web application, which means all computations are performed within the user's web browser (as

---

33 *"Onlink" is the field name in the Federal Geospatial Data Committee metadata content standard schema where the metadata creator fills in the URL or internet link where the user can find the data that the metadata describes.*

opposed to remotely on the web server). All options and features on the SCE tool have associated titles that pop up on mouse hover and are accessible to screen readers.

Our team incorporated feedback from potential users and a web designer consultant to develop a user-friendly and intuitive tool. Two user feedback sessions were held in May 2016 and March 2017. Participants were asked for input on tool navigability, functionality, and overall usability, including language. Participants came from a variety of organizations, including the Connecticut River Conservancy, The Nature Conservancy, US Fish and Wildlife Service, MA Division of Ecological Restoration, MA Dept. of Environmental Protection, MA DOT, UMass Transportation Center, Franklin Regional Council of Governments, and Pioneer Valley Planning Commission. A web designer consultant from Common Media Inc. was hired through the related Sandy project to provide input into the design and usability of the tool, including input towards the screen real estate and use of icons.

7.4   Outreach

To advertise and share the Deerfield Project results through the SCE, an outreach effort was initiated in late summer 2017. We began contacting potentially interested stakeholders from non-profit groups, state agencies, regional planning authorities, and all municipalities in the Deerfield River watershed via email to apprise them of the availability of the SCE and to offer that a team member present the project and demonstrate how to use the tool. Table 7-1 lists the various stakeholders identified and Table 7-2 lists the presentations made, where and by whom.

The presentations included a PowerPoint slideshow summarizing the Deerfield Project, followed by an interactive demonstration of the SCE on the Internet. The length of the presentations varied between 10 minutes and 1.5 hour, depending on the audience.

In November 2017, this outreach effort was put on hold, to be resumed when the present report was finished and approved by MA DOT.

Table 7-1: Organizations identified in initial outreach effort.

Organization	Location
American Rivers	Northampton, MA
Connecticut River Conservancy	Greenfield, MA
Consultants (MMI)	CT, VT
Deerfield Creating Resilient Communities	Deerfield, MA
FEMA	Maynard, MA
Franklin Conservation District	Greenfield, MA
Franklin Regional Council of Governments	Greenfield, MA
Friends of Conte Refuge (includes many organizations such as The Nature Conservancy)	Brattleboro, VT
North Atlantic LCC	Hadley, MA
MA DER	Boston, MA
MA DOT	Boston, MA
MEMA	Framingham, MA
Natural Resources Conservation Service	
Pioneer Valley Planning Commission	Springfield, MA
Regional DOT offices	
Selectboard, Planning Board, Open Space Committee, Conservation Commission, Highway Department	Ashfield, Buckland, Charlemont, Conway, Florida, Greenfield, Hawley Heath, Leyden, Monroe, Rowe, Savoy, Shelburne –all MA
Silver Jackets Flooding Group	Statewide, mostly Boston-based
Trout Unlimited	Brattleboro, VT
US Fish and Wildlife Service	Region 5 Hadley, MA

Note: (groups in italics were not directly contacted, many were part of collaborative groups such as Friends of the Conte Refuge).

Table 7-2: Presentations given.

Date &Time	Location	Audience	Presenter
9/13/2017 10:00	Springfield MA	PVPC Transportation	McArthur
9/19/2017 10:00	Brattleboro VT	Friends of Conte Refuge	Ocana
9/28/2017 9:30	Deerfield MA	Creating Resilient Communities	Jackson
10/12/2017 12:00	Hadley MA	USFWS	Jackson
10/23/2017 19:00	Shelburne, MA	Open Space Committee + more	Jackson
10/25/2017 10:00	Greenfield MA	Planning Dept staff	Hatte
11/2/2017 18:45	Greenfield MA	FRCOG Planning Board	Hatte
11/20/2017 11:00	Framingham MA	MEMA	Jackson, Dan Sheldon, Mabee

8.1 Key Findings

There were three broad objectives for the Deerfield Project.

1. Develop methodologies for assessing structural, geomorphic and hydraulic risk of failure for road-stream crossings and the associated disruption of emergency medical services.

2. Use these new methodologies, along with existing methodologies for assessing ecological disruption, to assess road-stream crossings in the Massachusetts portion of the Deerfield River watershed.

3. Incorporate climate change into hydraulic analyses using down-scaled climate models to predict future flows and hydraulic risk of failure for road-stream crossings.

The development of methodologies associated with hydraulic risk proved the most complex and difficult task, involving the calculation of critical flows ($Q_{critical}$) for various bridge and culvert types and the use of five hydrologic models and nine climate models to determine predicted current and future flows. Two types of hydrologic models were utilized, physical models and statistical models. Recall that physical models mathematically represent the processes governing transformation of rainfall to runoff with equations describing the conservation of mass and momentum. In addition they simulate the accumulation of flow as streamflow, and the translation/magnification of that flow downstream. Statistically based models, on the other hand, directly estimate the streamflow associated with a given probability of occurrence, or RI, in any given year based on regression equations relating streamflow to explanatory variables. Of the five hydrologic models, the three physical models resulted in flows and risk scores that were considerably lower than the two statistical models. The hydrologic models were validated by comparing streamflow estimates against historical data. However, because many crossings (especially culverted crossings) occur on small waterways that lack stream gauges, we lack data necessary to validate the model predictions at all basin scales. Therefore, we have no way to judge whether physical models or statistical models are more accurate in predicting flows at road-stream crossings with drainage basin areas less than those for which gaged data are available. In order to incorporate the results of both types of model, we combined the physical and statistical results to create a default score that fell in between the average physical and average statistical flows. Averaging for calculation of ensemble scores was done at the Q_{25} level. In other words, when blended scores including both physical and statistical models were determined, the average Q_{25} for the physical models and statistical models were first calculated separately, then combined as a simple average. The ensemble Q_{25} was then utilized to calculate a risk score. We chose Q_{25} for this analysis because, although storms of this size are not frequent, it is likely that many crossings will experience Q_{25} flows at some point in their design life.

Although absolute correlation coefficients for hydraulic risk scores (Table 8-1) were not as high as we might have hoped, especially for statistical versus physical models (0.575), ranked coefficients were quite strong (Table 8-2). In other words, although scores derived from statistical models were quite different from those from those generated by physical models, crossings scored highest for risk by physical models were generally among the crossings that had the highest risk as predicted by statistical models (Spearman's coefficient = 0.906). Thus, data from any of these models can be used to evaluate the relative risk of failure due to hydraulic capacity.

These results give people who want to evaluate the risk of crossing failure based on hydraulic capacity a sense of how the models used in this study compare: physical models produce lower risk estimates while statistical models produce higher ones.

Table 8-1: Pearson's correlation coefficients for hydraulic risk scores.

	Hydraulic Risk-All	Hydraulic Risk-Statistical
Hydraulic Risk-Statistical	0.888	
Hydraulic Risk-Physical	0.761	0.575

Table 8-2: Spearman's rank correlation coefficients for hydraulic risk scores.

	Hydraulic Risk-All	Hydraulic Risk-Statistical
Hydraulic Risk-Statistical	0.983	
Hydraulic Risk-Physical	0.933	0.906

However, we cannot know for sure which of these estimates is more likely to be correct.

We selected nine climate models that predicted past flows in our study area that were best correlated with historical data. Despite this agreement, the nine climate models were quite variable in their predicted temperature and precipitation values, and thus in the associated predicted flows at each crossing. We have no basis for deciding which model or models among the nine were "correct" or best, and therefore we used an ensemble of predicted conditions to estimate default values for hydraulic risk. The variability in predicted hydraulic risk scores based on individual climate models suggests that one should be wary of predictions or assessments based on a single climate model. In this study, certain models tended to predict higher levels of risk than others did. When using a limited number of models for assessing climate vulnerability, it is important to consider whether the model or models used tend to predict flows at the low, middle or high portion of the range of possibilities.

At the outset of the Deerfield Project, we planned to determine hydraulic risk based on whether crossings would pass a 10, 25, or 50-year storm. In implementing the project and analyzing the results, we decided that this type of evaluation would suggest a level of accuracy that we cannot reasonably claim. Some of the uncertainties involved include:

• Uncertainty associated with the calculation of hydraulic capacity ($Q_{critical}$) based on data collected during rapid field assessments,

• Uncertainty based on the variability of outputs from the various hydrologic models,

• Uncertainty based on the variability of outputs from the various climate models, and

• Uncertainty about what constitutes a 10, 25 or 50-year storm, as indicated by the confidence interval range calculated for the physical hydrologic models, due to the available data record as well as other factors such as changes in precipitation patterns that have already been experienced in New England, or changes in land-use.

Instead, we decided to score hydraulic risk of failure based on relative rather than absolute risk. Our scores range from zero (low risk) to 1.0 (high risk). At this time, it is impossible to say what scores equate to moderate, high or extremely high risk categories in an absolute sense. However, we are confident that a crossing with a score of 0.9 is at higher risk of hydraulic failure than a crossing with a score of 0.7. This decision also brings our hydraulic scoring into line with the methods used to score structural and geomorphic risk. Future storms may provide opportunities to calibrate these scores to actual failure risk.

Geomorphic vulnerability is much like hydraulic vulnerability in that it is storm-related, including periods of rapid snowmelt and multiple storms occurring within a short period of time. That means that it may be many years before failures occur, or they could occur next month depending on when the next storm of sufficient severity

hits the area. They differ from structural risk of failure in that structural deficiencies could result in crossing failure even in the absence of storms. Thus, it is conceivable that we could combine hydraulic and geomorphic risk into a single, storm-related risk of failure score for assessing impacts of future, Irene-like storm events.

The geomorphic vulnerability score is assigned as the highest of four component scores, three related to separate mechanisms for geomorphic failure (sediment blockage, blockage by woody debris, and scour), and a fourth based on the presence/absence and severity of blockages documented at the crossing at the time of assessment. The two primary metrics used to evaluate geomorphic vulnerability are structure width relative to bankfull width (woody debris and sediment blockage) and a combination of specific stream power and bed resistance (sediment blockage and scour). Other information used to score geomorphic vulnerability came from field assessments, and include the observation of aggradation or scour, sediment continuity upstream and downstream of the crossing, condition of footings (if present), and the presence/absence and size of any downstream scour pool. Component scores for woody debris, sedimentation and scour generally ranged between zero and 0.8 with mean scores ranging from 0.38 to 0.42. Thus, the scoring system results in a reasonable distribution of risk scores. Forty-six crossings received geomorphic vulnerability scores of 1.0 because field assessments revealed that they were already significantly blocked by sediment and/or woody debris.

Structural Risk of Failure was based on MassDOT bridge inspection results or, if those were not available, an ends-only rapid assessment of culvert condition. Most crossings fell into low risk categories, with only about 20 percent scoring 0.5 or higher. The vast majority of high scoring crossings (≥ 0.8) were culverts or small bridges; only four bridge crossings assessed by MassDOT scored ≥ 0.8. Theoretically, high-risk crossings could fail at any time. However, it is conceivable that these structures would be at even higher risk during severe storm events.

Table 8-3 contains Spearman's Rank Correlations for ecological disruption, overall risk of failure, and criticality. The correlations are not particularly strong, ranging from 0.37 (ecological disruption and criticality) to 0.57 (ecological disruption and risk of failure). This suggests that these factors are somewhat independent of each other and would need to be evaluated separately in order to comprehensively evaluate the ecological and transportation vulnerabilities associated with road-stream crossings.

We used the Deerfield Project to develop computer heuristics and algorithms that would allow us to measure the criticality of each structure by identifying alternative routes for the provision of EMS when faced with crossing failures. These heuristics and algorithms allow the efficient screening of alternatives to yield near-optimal results when seeking best alternative routes (a form of adaptability). The approach used to model EMS trips, and the computer program and scoring systems, are available to be adapted and used in other watersheds and applied to other elements of criticality (e.g. access to critical infrastructure, emergency response to disasters).

The criticality analysis used for the Deerfield River watershed created four metrics for use in prioritizing crossings for upgrade or

Table 8-3: Spearman's rank correlation of risk of failure, EMS Delay and Ecological Disruption scores.

	Ecological Disruption	Overall Risk
Overall Risk	0.571	
Overall EMS Delay	0.372	0.492

replacement. "Average Delay" is the total delay (in minutes) caused by the failure of a crossing for all trips affected by that failure divided by all the trips modeled (n = 5,000). "Average Affected Delay" is summed delay (in minutes) for all trips affected by a failure divided by the total number of trips affected. "Maximum Delay" is the maximum delay (in minutes) caused by a crossing failure; the maximum delay for any crossing is capped at 60 minutes. All three of these tend to highlight only a handful of crossings with high delays. "Average Delay" tends to identify crossings located on frequently traveled routes (e.g. Route 2) as priorities. "Average Affected Delay" and "Maximum Delay" are more likely to prioritize crossings that affect a smaller number of emergency trips, but where at some of these trips result in long delays. A fourth (unitless) metric uses a logistic weighting function to account for the number of trips affected and the magnitude of delays. This metric accounts for both the number of trips affected (even if only with small delays) like the "Average Delay" metric, and trips with long delays (even if these are relatively few) such as the "Maximum Delay" and "Average Affected Delay" metrics. This "Overall Delay" metric tends to spread the scores more evenly throughout the range of zero to 1.0.

The calculation of an overall risk of failure score by taking the maximum score for structural, hydraulic and geomorphic risk yields a distribution skewed toward higher scores (higher risk). That is because there is little correlation among the three components of risk. As it turns out, many crossings are at relatively high risk for at least one of these components. When risk is combined with criticality (disruption of EMS), the resulting Transportation Vulnerability scores are no longer skewed high. Approximately 40 percent of crossings are in the lowest score category and the remaining scores are well distributed across the range with few crossings scored in the highest categories.

The ecological component of the Deerfield Project used existing assessment protocols and applied existing connectivity models to evaluate how each crossing affects aquatic connectivity. As a base assessment, the existing Critical Linkages model was run using pre-existing NAACC data and data collected as part of this project, and the results were used to score crossings for connectivity restoration potential. We also ran a specialized version of Critical Linkages focusing on connectivity for cold water streams. These analyses used a variety of temperature thresholds (14°C, 16°C, 18°C, 20°C, and 22°C) to define "cold water streams," and cold water connectivity restoration potential was scored using each of these thresholds. A score for Ecological Disruption was calculated

as the maximum score between restoration potential for the base analysis or for cold water streams using a 16°C threshold. The number of crossings affecting cold water streams is significantly less than the total number of crossings in the watershed. As a result, the Ecological Disruption scores are controlled by the base assessment for about 80 percent of crossings, with 20 percent of crossing scores dictated by cold water restoration potential. Ecological Disruption scores contain many crossings in the lowest category (~ 25 %) with the remaining scores distributed in something approaching a Gaussian distribution for the range 0.1 through 1.0.

Combining Ecological Disruption and Transportation Vulnerability scores yielded Crossing Priority scores that were well distributed throughout the range with few crossings falling in the lowest or highest categories and many receiving scores in the middle of the range.

One question that we wanted to answer with the Deerfield Project was to what degree the results of crossing assessments for aquatic passability could be used to evaluate risk of failure. In other words, how correlated are impassability scores with risk of failure scores as estimated by this project? Table 8-2 contains a correlation matrix showing Spearman's Rank correlation coefficients for Impassability, Ecological Disruption, and the various risk of failure scores.

Impassability is fairly strongly correlated (0.698) with overall risk of failure (Table 8-4 and Figure 8-1). The relationship is nearly as strong for Impassability and Geomorphic Risk (0.687) and is


Figure 8-1: Impassability vs. Overall Risk of Failure. The degree to which a crossing represents a barrier to aquatic organism passage is somewhat strongly correlated with a crossing's overall risk of failure, Spearman Rank Correlation = 0.698.

Table 8-4: Spearman's rank correlation of risk of failure and Impassability scores.

	Impassability Score	Overall Risk	Structural Risk	Geomorphic Risk
Ecological Disruption	0.761			
Overall Risk	0.698			
Structural Risk	0.566	0.715		
Geomorphic Risk	0.687	0.827	0.523	
Hydraulic Risk-All	0.613	0.791	0.543	0.561
Hydraulic Risk-Statistical	0.619	0.791	0.541	0.558
Hydraulic Risk-Physical	0.607	0.795	0.543	0.577

weakest for Structural Risk of Failure (0.566). The relatively high correlations probably stem from the fact that undersized crossing structures tend to both be highly disruptive of aquatic organism passage and vulnerable to geomorphic and hydraulic failure. This suggests that impassability scores derived from NAACC data could serve as an imperfect but reasonable stand-in for risk of failure if data from geomorphic and hydraulic risk assessments are not available.

Spearman's rank coefficients for overall risk of failure when compared to the three components of risk (structural, geomorphic and hydraulic) are all moderately high and consistent. This suggests that the contributions of these three components to overall risk scores are comparable.

The relationship between Impassability and Ecological Disruption is strong (0.761) but not very strong. This is due to the role that habitat quality (CAPS IEI) and presence of other nearby barriers plays in translating impassability to ecological restoration potential.

8.2   Limitations on Use of Data

There are a number of considerations or limitations that should be kept in mind when using data and analyses generated by the Deerfield Project.

- Hydrological models may have been updated since the work for the Deerfield Project was completed.

- The data collected and hydrological and hydraulic analyses conducted as part of the Deerfield Project were intended for use in a qualitative assessment and are not sufficient for use in the design of road-stream crossings.

- At this point, we cannot relate our risk of failure scores to a probability of failure. Although we attempted to create scoring systems that provide comparable scores for structural, geomorphic and hydraulic risk (meaning that similar scores

would equate to similar probability of failure), only with time and monitoring will we be able to identify numeric scores most indicative of high, moderate and low probability of failure.

- It is not possible at this time to relate our ecological disruption scores to passability for particular species of fish or other aquatic organisms.

- Disruption of EMS is the only aspect of criticality assessed in the Deerfield Project. Other elements of criticality that were not assessed include access to critical infrastructure (water supplies, wastewater treatment facilities, electrical substations, gas compressor stations, etc.) and the broader need to facilitate vehicular traffic.

- Our assessments should not be used on their own to make decisions about crossing repair or replacement. They are intended to be used as a screening tool to draw attention to crossings that warrant further, more detailed analysis for risk of failure, restoration of aquatic connectivity, or criticality for provision of emergency services.

- While future hydrology and hydraulic risk predictions were considered in this report, extrapolating recent climate trends may not be effective for predicting future conditions because climate change can create new states that no longer adhere to past patterns. Climate models are still in early stages of development, and the various models used for the Deerfield Project yielded highly variable results.

8.3   Lessons Learned

In the process of completing the Deerfield Project, we encountered unexpected circumstances and difficulties, and learned some lessons that will help in the application of our approach in other watersheds.

- We would have benefited from designing and creating a

database early in the project and then populating it with data as they were collected. It can get confusing when multiple versions of datasets are in circulation among project personnel.

- Having time to adequately identify all the road-stream crossings for assessment prior to the field season and development of standard operating procedures for field crews would have made field work more efficient and helped to avoid repeat trips to field sites.

- Although we were able to collect data at a number of crossings that had failed during tropical storm Irene, these data were not as useful as they could have been because most structures had already been replaced. We, therefore, lacked information necessary to calculate risk for structures that had actually failed during that storm.

- Gauge data on streamflows were generally lacking on the small streams where many culverted crossings are located.

- The Deerfield Project would have benefited from more detailed elevation data. A relatively high-resolution digital elevation model (DEM) that was available was not very good. Therefore, we used a coarser DEM that was consistent with other data used in the project.

- Data on crossing slope collected may have discrepancies due to inadequate surveying equipment and personnel trained to use it. It is not clear to what degree our results would have differed with better quality slope data.

- We made changes in our approaches to scoring hydraulic and geomorphic risk mid-project.

  ○ For hydraulic risk, we abandoned a scoring system that was based on discrete return intervals in favor of an approach based on relative risk. We also changed the scoring system from a low-medium-high rating to a continuous numeric score ranging from zero to one.

  ○ For geomorphic risk, we moved from our original plan for a single geomorphic scoring approach to one that involved the calculation of separate scores for each mechanism of geomorphic failure. These component scores were combined to create a single geomorphic risk of failure score for each crossing.

8.4 Recommendations for Consideration

Following are our recommendations for future work on the assessment of road-stream crossings and decision-making about crossing upgrades and replacement.

- Risk of failure assessments are essentially predictions that can be validated over time as crossings fail. It would be useful to have a system for tracking crossing failures and replacements.

- Further research is needed to create an effective model for predicting crossings most vulnerable to plugging by woody debris.

- The rich data set compiled as part of the Deerfield Project should be explored to see if it would be possible to create a simple but credible model to predict storm-related risk of failure (e.g. combined geomorphic and hydraulic risk).

- Risk of failure assessments should be implemented for road-stream crossings across Massachusetts. Having data in hand prior to future storms and large numbers of crossing failures, will allow validation, improvement and/or creation of new risk assessment methodologies. It would have been substantially easier for us to create risk assessment protocols and scoring systems had we collected crossing data collected prior to tropical storm Irene.

- Expand the criticality component of the Deerfield Project by applying the methodology to other aspects of criticality in the transportation system.

- Build on EMS disruption modeling to evaluate criticality for large storms that involve multiple crossing failures; incorporate risk of failure as a stochastic element in probabilistic modeling.

- Build on EMS disruption modeling to create the capacity to respond to large storms with multiple crossing failures by identifying optimum sequencing for crossing repair.

- A logical next step in the process of developing and implementing a methodology for comprehensively assessing road-stream crossings in Massachusetts would be to apply a version of the Deerfield Project methodology in another watershed and evaluate how robust the methodologies are when applied to dissimilar landscapes.

- Given all the effort that went into estimating flows and hydraulic risk at road-stream crossings, one might expect that one outcome of this research would be a recommendation about what standards should be adopted (e.g. what design storm to use) for future crossings. However, the Deerfield Project revealed that there is a tremendous amount of uncertainty involved in predicting future streamflows. These uncertainties are related to:

  ○ A lack of stream gauge data,

  ○ A high level of variability in the outputs of different hydrological models,

  ○ A high level of variability in the predictions from different climate models, and

  ○ A lack of confidence in our ability to use data from past storms/flows to predict the return frequency of future storms.

It doesn't mean much to say that a certain size storm is a 50 or 100-year event. Current predictions of storm return intervals are based on a limited amount of data collected many years ago. An effort to calculate new storm return interval predictions based on the longer record now available would still be suspect, because it relies on an assumption of stationarity, an assumption that past conditions can reliably be used to predict future conditions.

- An alternative approach for establishing design standards for road-stream crossings would be to shift away from design standards based on storm recurrence rates that are increasingly difficult to predict, and instead adopt the MA River and Stream Crossings Standards, originally developed to avoid creating barriers to aquatic organism passage and maintain aquatic connectivity.

- Explore expanding bridge and culvert design using fluvial geomorphic principles to reduce risk of failure.

REFERENCES

Barnes Jr., H. 1967. *Roughness characteristics of natural channels. USGS water supply paper 1849*, U. S. Government Printing Office, Washington, D.C., 213p.

Bent, G.C. and A.M. Waite. 2013. *Equations for estimating bankfull channel geometry and discharge for streams in Massachusetts*. U.S. Geological Survey Scientific Investigations Report 2013-5155, 61p.

Beven, K. 1987. Towards the use of catchment geomorphology in flood frequency predictions. *Earth Surface Processes and Landforms*. January/February, Volume 12, Issue 1, pp. 69-82.

Beven, K. 1995. Linking parameters across scales: subgrid parameterizations and scale dependent hydrological models. *Hydrological Processes*. Volume 9, Issue 5-6, pp. 507-525.

Beven, K., P. Smith, I. Westerberg, and J. Freer. 2012. Comment on "Pursuing the method of multiple working hypotheses for hydrological modeling" by P. Clark et al. Water Resources Research, 48(11).

Bicknell, B. R., J.C. Imhoff, J.L. Kittle Jr., A.S. Donigian Jr., and R.C. Johanson. 1996. *Hydrological simulation program-FORTRAN*. User's manual for release 11. US EPA.

Brekke, L., B.L. Thrasher, E.P. Maurer, and T. Pruitt. 2013. Downscaled CMIP3 and CMIP5 Climate Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with Preceding Information, and Summary of User Needs. Prepared for users of the "Downscaled CMIP3 and CMIP5 Climate Hydrology Projections: Release of Downscaled CMIP5 Climate Projections". Online: *http://gdo-dcp.ucllnl.org/ downscaled_cmip_projections/.*

Clark, G. 2016. Hydrologic Modeling at Ungauged Locations in Support of the Development of a Vulnerability Ranking Protocol System for Road-Stream Crossing Infrastructure, University of Massachusetts Amherst Environmental & Water Resources Engineering Masters Project, 77. *https://doi.org/10.7275/97th-gp81*

Commonwealth of Massachusetts Governor Executive Order No. 569. 2016. "Establishing an Integrated Climate Change Strategy for the Commonwealth" *http://www.mass.gov/governor/legislationexecorder/ execorders/executive-order-no-569.html*

Falcone, J.A., D.M. Carlisle, D.M. Wolock, and M.R. Meador. 2010. GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States: Ecological Archives E091-045. Ecology, 91(2), 621-621.

FEMA. 2009. Benefit-Cost Analysis Re-engineering (BCAR) version 4.5 May 2009 *https://www.fema.gov/media-library-data/20130726-1738-25045-2254/floodfulldata.pdf*

Gay, F.B., L.G. Toler, and B.P. Hansen. 1974. *Hydrology and water resources in the Deerfield River watershed, Massachusetts* (No. 506).

Georgakakos, A., P. Fleming, M. Dettinger, C. Peters-Lidard, T.C. Richmond, K. Reckhow, K. White, and D. Yates. 2014. Ch. 3: Water Resources. *Climate Change Impacts in the United States: The Third National Climate Assessment*, J. M. Melillo, Terese (T.C.) Richmond, and G. W. Yohe, Eds., U.S. Global Change Research Program, doi:10.7930/ J0G44N6T, pp. 69-112.

Gillespie, N., A. Unthank, L. Campbell, P. Anderson, R. Gubernick, M. Weinhold, D. Cenderelli, B. Austin, D. McKinley, S. Wells, J. Rowan, C. Orvis, M. Hudy, A. Bowden, A. Singler, E. Fretz, J. Levine, and R. Kirn. 2014. Flood Effects on Road–Stream Crossing Infrastructure: Economic and Ecological Benefits of Stream Simulation Designs.

Fisheries, Vol 39 No 2, pp. 62-76.

Hayhoe, K., C.P. Wake, T.G. Huntington, L. Luo, M.D. Schwartz, J. Sheffield, E. Wood, B. Anderson, J. Bradbury, A. DeGaetano, T.J. Troy, and D. Wolfe. 2007. Past and future changes in climate and hydrological indicators in the US Northeast. *Climate Dynamics*, 23:4, pp. 381 – 407.

Homer, C.G., J.A. Dewitz, L. Yang, S. Jin,, P. Danielson, G. Xian, J. Coulston, N.D. Herold, J.D. Wickham, and K. Megown. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. P*hotogrammetric Engineering and Remote Sensing*, v. 81, no. 5.

Hortness, J. E. 2006. *Estimating low-flow frequency statistics for unregulated streams in Idaho*. US Department of the Interior, US Geological Survey.

Horton, R., G. Yohe, W. Easterling, R. Kates, M. Ruth, E. Sussman, A. Whelchel, D. Wolfe, and F. Lipschultz. 2014. Ch. 16: Northeast. *Climate Change Impacts in the United States: The Third National Climate Assessment* (NCA3), J. M. Melillo, Terese (T.C.) Richmond, and G. W. Yohe, Eds., U.S. Global Change Research Program, 16-1-nn.

Hosking, J.R.M. 1990. L-Moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society*, Series B, 52(2), 105-124.

Hosking, J.R.M. and J.R. Wallis. 1997. *Regional frequency analysis: an approach based on L-moments*. Cambridge: Cambridge UP.

Hunt, J.H., S.M. Zerges, B.C. Roberts, and B. Bergendahl. 2010. *Culvert Assessment and Decision-making Procedures Manual for Federal Lands Highway*. FHWA-CFL/TD-10-005. Central Federal Lands Highway Division, Lakewood CO, 222p.

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments, *Ecol. Monogr.*, 54:2, 187–211, doi:10.2307/1942661

Jacobs, J. 2010. *Estimating the magnitude of peak flows for steep gradient streams in New England. New England Transportation Consortium*, Project No. NETC 04-3, 49 p.

Jospe, A. 2013. *Aquatic barrier prioritization in New England under climate change scenarios using fish habitat quality, thermal habitat quality, aquatic organism passage, and infrastructure sustainability.* M.S. Thesis, University of Massachusetts Amherst. 84+xiip.

Knox, C.E. and T.J. Nordenson. 1955. *Average annual runoff and precipitation in the New England-New York area* (No. 7).

Kochanek, K., I. Markiewicz, and W.G. Strupczewski. 2010. "*On Feasibility of L-Moments method for distributions with cumulative distribution function, and its inverse inexpressible in the explicit form*". International Workshop: Advances in Statistical Hydrology, Taormina, Italy.

Kuczera, G. and S. Franks. 2006. Australia Rainfall and Runoff, Book IV, Estimation of Peak Discharge-Draft. *Engineers Australia*.

Letcher, B.H., D.J. Hocking, K. O'Neill, A.R. Whiteley, K.H. Nislow, and M.J. O'Donnell. 2016. A hierarchical model of daily stream temperature using air-water temperature synchronization, autocorrelation, and time lags, PeerJ, 4: e1727, 1–26, doi:10.7717/peerj.1727.

Levine, J. 2013. *An Economic Analysis of Improved Road-Stream Crossings*. The Nature Conservancy, Adirondack Chapter, Keene Valley, NY.

Lombard, P.J. and G.C. Bent. 2015. *Flood-inundation Maps for the Deerfield River, Franklin County, Massachusetts, from the Confluence with the Cold River Tributary to the Connecticut River*. Scientific Investigations Report 2015-5104. U.S. Geological Survey: Reston, VA., 22 p.

Massachusetts Climate Change Adaptation Report, September 2011. *http://www.mass.gov/eea/docs/eea/energy/cca/eea-climate-adaptation-report.pdf*

MassDOT. June 2013. *Bridge Load and Resistance Factor Design (LRFD) Manual*, Section 1.3.

Maurer, E. P., L. Brekke, T. Pruitt, and P. B. Duffy. 2007. *Fine-resolution climate projections enhance regional climate change impact studies*, Eos Trans. AGU, 88(47), p. 504.

McGarigal, K., B.W. Compton, S.D. Jackson, E. Plunkett, K. Rolih, T. Portante, and E. Ene. 2011 (revised July 2017). Conservation Assessment and Prioritizatioin System (CAPS) Statewide Massachusetts Assessment:2015 Landscape Ecology Program, UMass Amherst. 57+ip.

McGarigal, K., B. W. Compton, S. D. Jackson, E. Plunkett, and E. Ene. 2012 (revised July 2017). Critical Linkages Phase 1: Assessing Connectivity Restoration Potential for Culvert Replacement, Dam Removal, and Construction of Wildlife Passage Structures in Massachusetts. Landscape Ecology Program, UMass Amherst. 27+i pp.

Mearns, L.O. et al. 2007, updated 2014. *The North American Regional Climate Change Assessment Program dataset*, National Center for Atmospheric Research Earth System Grid data portal, Boulder, CO. Data downloaded 2017-06-26. [doi:10.5065/D6RN35ST]

Mearns, L.O., W.J. Gutowski, R. Jones, L.-Y. Leung, S. McGinnis, A.M.B. Nunes, and Y. Qian. September 2009. A regional climate change assessment program for North America. EOS, Vol. 90, No. 36, pp. 311-312. [doi:10.1029/2009EO360002]

Melillo, J.M., T.C. Richmond, and G. W. Yohe, Eds. 2014. *Climate Change Impacts in the United States: The Third National Climate Assessment* (NCA3). U.S. Global Change Research Program, 16-1-nn. On the Web: *http://nca2014.globalchange.gov/report/regions/northeast*

Millington, N., S. Das, and S.P. Simonovic. 2011. *The comparison of GEV, log-Pearson type 3 and Gumbel distributions in the Upper Thames River watershed under global climate models*.

Mohseni O. and H.G. Stefan. 1999. Stream temperature/air temperature relationship: a physical interpretation, *Journal of Hydrology*. 218:3–4, 128–141, doi:10.1016/S0022-1694(99)00034-7

NCAR GIS Program. 2012. *Climate Change Scenarios, version 2.0. Community Climate System Model, June 2004 version 3.0. http://www.cesm.ucar.edu/models/ccsm3.0/* was used to derive data products. Downscaling procedure in: T. Hoar and D. Nychka. 2008. Statistical Downscaling of the Community Climate System Model (CCSM) monthly temperature and precipitation projections, URL: *http://www.gisclimatechange.org*. Data Access Date: July 2016.

NLCD. 2011 – See Homer et al. 2015, preferred citation.

New York City Panel on Climate Change (NPCC) 2015 Report. *Annals of the New York Academy of Sciences*. Building the Knowledge Base for Climate Resiliency: Volume 1336. Appendix I: Climate Risk and Projections NPCC 2015 Infographics.

Parker, G.W., L. Bratton, and D.S. Armstrong. 1997. *Stream Stability and scour assessments at bridges In Massachusetts:* U.S. Geological Survey Open File Report 97-588, 2 CD-ROMs.

PRISM Climate Group. 2004. Oregon State University, *http://prism.oregonstate.edu*.

Rawlins, M.A., R.S. Bradley, and H.F. Diaz. 2012. *Assessment of regional climate model simulation estimates over the northeast United States.* Journal of Geophysical Research, 117, D23112, [doi:10.1029/2012JD018137].

Reclamation. 2011. *West-Wide Climate Risk Assessments: Bias-Corrected and Spatially Downscaled Surface Water Projections*, Technical Memorandum No. 86-68210-2011-01, prepared by the U.S. Department of the Interior, Bureau of Reclamation, Technical Services Center, Denver, Colorado, 138 p.

Reclamation. 2013. *Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with preceding Information, and Summary of User Needs*, prepared by the U.S. Department of the Interior, Bureau of Reclamation, Technical Services Center, Denver, Colorado, 47p.

Reclamation. 2014. *Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections: Release of Hydrology Projections, Comparison with preceding Information, and Summary of User Needs*, prepared by the U.S. Department of the Interior, Bureau of Reclamation, Technical Services Center, Denver, Colorado, 110 p.

Reis, D.S. Jr., J.R. Stedinger, and E.S. Martins. 2005. Bayeseian generalized least squares regression with application to log Person type 3 regional skew estimation. *Water Resources Research*, Volume 41, W10419 (14 pages).

Schall, J.D., E.V. Richardson, and J.L. Morris. 2008. Introduction to highway hydraulics, hydraulic design series number 4, fourth edition, Federal Highway Administration, 204p. *https://www.fhwa.dot.gov/engineering/hydraulics/pubs/08090/HDS4_608.pdf*

Schall, J.D., P.L. Thompson, S.M. Zerges, R.T. Kilgore, and J.L. Morris. 2012. *Hydraulic Design of Highway Culverts*, Third Edition. Hydraulic Design Series Number 5, FHWA-HIF-12-026. U.S. Department of Transportation, Federal Highway Administration: Washington, D.C.

Singh, R., S.A. Archfield, and T. Wagener. 2014. Identifying dominant controls on hydrologic parameter transfer from gage to ungage catchments–A comparative hydrology approach. *Journal of Hydrology*, 517, pp. 985-996.

Soar P., Thorne C. 2001. Channel restoration design for meandering channels, U.S. Army Corp of Engineers, Vicksburg, MS, Engineer Research and Development Center, ERDC CR-01-1, 454p.

Stedinger, J. R., R. M. Vogel and E. Foufoula-Georgiou. 1993. Frequency Analysis of Extreme Events, in: D. R. Maidment, Ed., *Handbook of Hydrology*, McGraw-Hill, New York.

Thrasher, B., E.P. Maurer, C. McKellar, and P.B. Duffy. 2012. *Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping*. Hydrology and Earth System Sciences, 16(9), pp. 3309-3314.

U.S. DOT. 2014. *U.S. DOT CMIP Climate Data Processing Tool User's Guide*. Prepared by ICF International for the U.S. Department of Transportation Center for Climate Change and Environmental Forecasting under *The Gulf Coast Study, Phase 2, Impacts of Climate Change and Variability on Transportation Systems and Infrastructure.*

Vermont Agency of Natural Resources. 2007. *Vermont Stream Geomorphic Assessment Phase 1 Handbook: Watershed Assessment*, 89p. Last accessed 5/31/2017 *http://dec.vermont.gov/sites/dec/files/wsm/rivers/docs/rv_SGA_Phase1_Protocol.pdf*

Villarini, G. and J.A. Smith. 2010. *Flood peak distributions for the eastern United States*. Water Resources Research, Vol. 46, Issue 6.

Vogel, E., N. Gillett, C. Hatch, B. Warner, J. Schoen, L. Payne, D. Chang, P. Huntington, J. Gartner, and N. Slovin. 2016. *Supporting New England Communities to Become River-Smart: Policies and Programs that can Help New England Towns Thrive Despite River Floods*. Water Reports, 3. *http://scholarworks.umass.edu/water_reports/3*

Vogel, R.M. and I. Wilson. 1996. Probability distribution of annual maximum, mean, and minimum streamflows in the united states. *Journal of Hydrological Engineering*. 1, pp. 69–76.

Walsh, J., D. Wuebbles, K. Hayhoe, J. Kossin, K. Kunkel, G. Stephens, P. Thorne, R. Vose, M. Wehner, J. Willis, D. Anderson, S. Doney, R. Feely, P. Hennon, V. Kharin, T. Knutson, F. Landerer, T. Lenton, J. Kennedy, and R. Somerville. 2014. Ch. 2: Our Changing Climate. *Climate Change Impacts in the United States: The Third National Climate Assessment*, J. M. Melillo, Terese (T.C.) Richmond, and G. W. Yohe, Eds., U.S. Global Change Research Program, pp. 19-67. doi:10.7930/J0KW5CXT.

Wolman, M.G. 1954. A Method of sampling coarse river-bed material. *Transactions of American Geophysical Union* 35, pp.951-956.

Zarriello, P.J. 2017. *Magnitude of flood flows at selected annual exceedance probabilities for streams in Massachusetts*: U.S. Geological Survey Scientific Investigations Report 2016–5156, 54 p. *https://doi.org/10.3133/sir20165156.*

Zipkin, E.F., E.H.C. Grant, W.F. Fagan. 2012. Evaluating the predictive abilities of community occupancy models using AUC while accounting for imperfect detection. *Ecological Applications* 22:1962–1972.

UMassAmherst

massDOT
Massachusetts Department of Transportation
Highway Division