River Basin Export Reduction Optimization Support Tool (RBEROST) User Guide, v1.15

12/08/2021

Table of Contents

1 Contributors and Contacts

Catherine Chamberlin (ORISE - EPA): contributor to RBEROST v1 and Upper Connecticut Basin case study; chamberlin.catherine@epa.gov

Naomi Detenbeck (EPA): lead PI; detenbeck.naomi@epa.gov

Marilyn ten Brink (EPA): co-PI; tenbrink.marilyn@epa.gov

Alyssa Le (ICF): contributor to WMOST Scaled-up Optimization v1 and Upper Connecticut Basin case study

Kate Munson (ICF): contributor to WMOST Scaled-up Optimization v1 and Upper Connecticut Basin case study

Isabelle Morin (ICF): contributor to WMOST Scaled-up Optimization v1

Miranda Marks (ICF): contributor to WMOST Scaled-up Optimization v1

Yishen Li (ICF): contributor to Upper Connecticut Basin case study

2 List of Acronyms

ACESD – Atlantic Coastal Environmental Sciences Division

ACRE – Agricultural Conservation Reduction Estimator

AMPL – A Mathematical Programming Language

BMP – Best Management Practice

CEMM – Center for Environmental Measurement and Modeling

COMID – Common Identifier

CPLEX – Optimizer based on the simplex method as implemented in the C programming language

CSV – Comma Separated Values

DEL_FRAC – Delivery Fraction

EPA – Environmental Protection Agency

EQIP – Environmental Quality Incentives Program

HDMA – Hydrologic Derivatives for Modeling and Analysis

HSG – Hydrologic Soil Group

HUC – Hydrologic Unit Code

HUC10 – 10-digit Hydrologic Unit Code

HUC12 – 12-digit Hydrologic Unit Code

HUC8 – 8-digit Hydrologic Unit Code

LOESS – locally estimated scatterplot smoothing

MA – Massachusetts

MS4 – Municipal Separate Stormwater Sewer

N – Nitrogen

NATSGO – National Soil Survey Geographic Database

NEIWPCC – New England Interstate Water Pollution Control Commission

NH – New Hampshire

NHD – National Hydrography Dataset

NHDPlus – National Hydrography Dataset Plus

NLCD – National Land Cover Dataset

NPDES – National Pollution Discharge Elimination System

NRCS – Natural Resources Conservation Service

ORD – Office of Research and Development

ORISE – Oak Ridge Institute for Science and Education

P – Phosphorus

PI – Principal Investigator

R – Programming language

RBEROST – River Basin Export Reduction Optimization Support Tool

RShiny – R package providing graphical user interface for user inputs and outputs

Se – Standard Error

SPARROW – Spatially Referenced Regressions On Water

TDEP – Total Nitrogen Deposition

TN – Total Nitrogen

TP – Total Phosphorus

USGS – United States Geological Survey

UVM – University of Vermont

VT – Vermont

WEDB – Watershed and Estuarine Diagnostic Branch

WMOST – Watershed Management Optimization Support Tool

WQv – Water Quality Volume

WWTP – Wastewater Treatment Plant

XML – Extensible Markup Language

3 Model Framework

The River Basin Export Reduction Optimization Support Tool (RBEROST) is a decision support tool designed to support integrated, regional, watershed planning. The tool is designed to help managers reduce nutrient loading to targeted waterbodies for the least financial cost. This tool optimizes costs for meeting targets for nutrient export at the annual scale and is designed to be used as a screening tool for large watersheds (e.g., HUC 6 - HUC 8 scale). The tool is mathematically similar to the Watershed Management Optimization Support Tool, or WMOST (Detenbeck, ten Brink, et al., 2018; Detenbeck, Piscopo, et al., 2018). WMOST is built for optimization at the HUC10 or HUC12 scale and can be run on daily or monthly timesteps. RBEROST was developed as a regional screening tool to overcome computational challenges with running WMOST at larger spatial scales. All RBEROST scripts are written in R 4.0.5. and may display incompatibilities with other versions. Execution of the RBEROST application is recommended with RStudio version 1.4.1106 and may display incompatibilities with other versions. Additional R packages may need to be installed by the user. The R code and instructions on how to execute the code necessary for these installations are included in the documentation below.

RBEROST allows users to screen which locations and choices of Best Management Practices (BMPs) will meet annual loading targets for the least financial cost. There are three main steps within RBEROST including a preprocessing step, an interaction with an online server, and a postprocessing step. Additional work may be necessary before beginning RBEROST to collect and format the necessary data. The preprocessing step combines medium-resolution National Hydrography Dataset Plus (NHDPlus v2; McKay et al. (2012)) reach lengths, NHDPlus v2 catchment-level annual nutrient loading, land use data, hydrologic soil group data, nitrogen (N) deposition data, user-defined loading targets, user-defined agricultural, urban, point source and riparian buffer BMPs, and data on BMP-specific costs and nutrient removal efficiencies. The preprocessing step then uses this information to write three program files in A Mathematical Programming Language, or AMPL, including a model, data, and command file. There are two options available for the preprocessing step, i.e., creating AMPL files with or without uncertainty information included. The AMPL files describe a model that defines a cost-minimization optimization problem subject to meeting downstream annual loading targets. These files are then sent to a free online CPLEX server hosted by the Network-Enabled Optimization System (NEOS; University of Wisconsin in Madison (2021)). CPLEX is a linear solver that will solve the optimization problem defined by the user inputs of loading targets and selected BMPs for the least cost. Once the solution is optimized, the model outputs its decisions of which BMPs to implement, and where to implement them. The final step of RBEROST is the postprocessing step. This step parses the output from NEOS into a summary report describing which BMPs were implemented. It also provides csv files for download that describe the BMP choices on the NHDPlus v2 catchment-level scale. If the user chooses to run the model with uncertainty, the postprocessor will provide information on the expected cost range, as well as the likelihood of meeting each loading target. When run with uncertainty, the model can display multiple scenarios with increasing cost and likelihood for meeting targets. Figure 3.1 shows a diagram of the model framework.

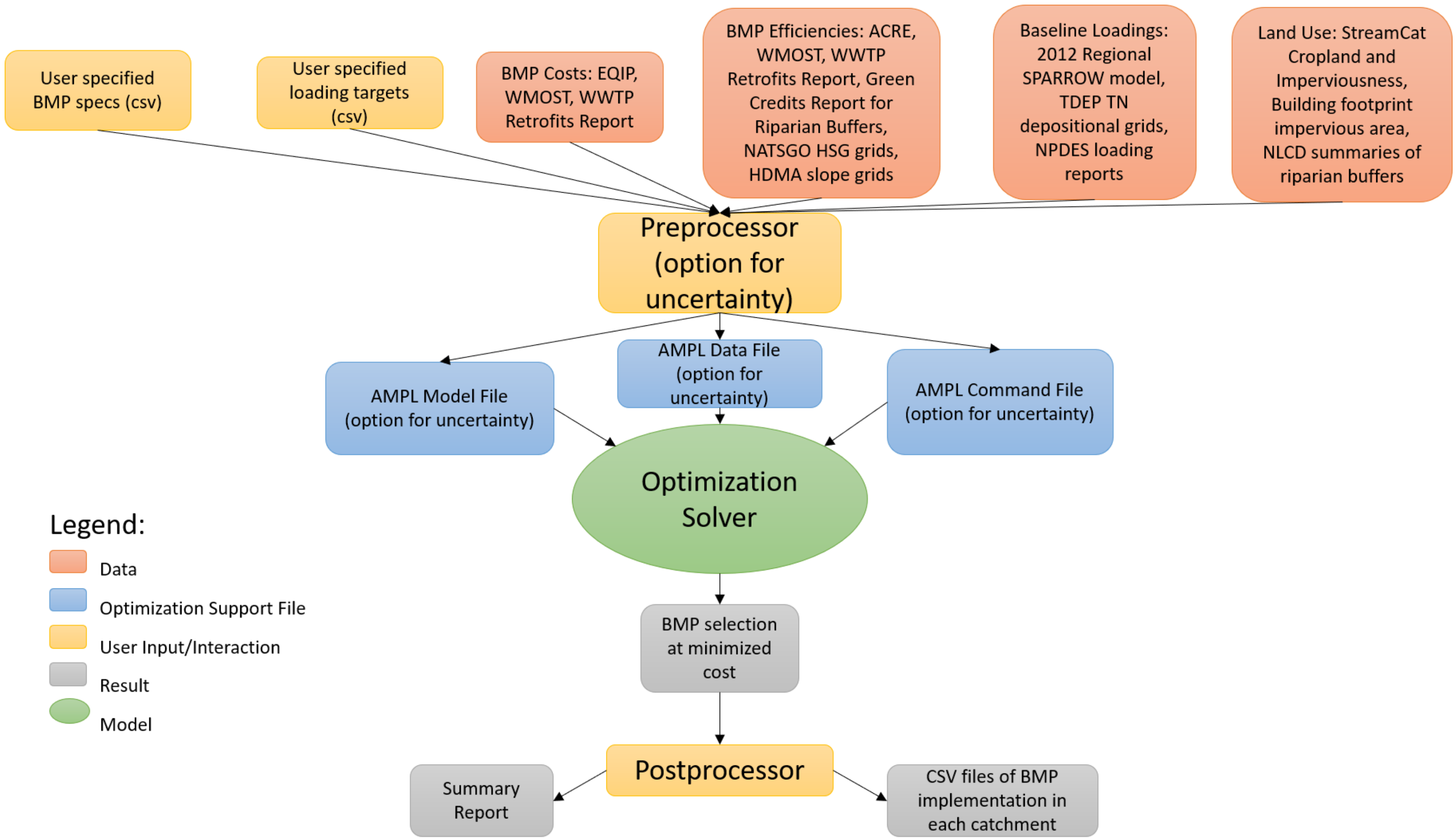


Figure 3.1: Schematic overview of RBEROST, made up of data, support file, user input, result, and model components

3.1 Optimization Method

RBEROST uses the IBM ILOG CPLEX Optimizer for linear programming that solves a mathematical problem written in AMPL to minimize the total annualized cost of selected BMPs. RBEROST interacts with this solver through the NEOS server, hosted by the University of Wisconsin-Madison. The CPLEX Optimizer was chosen because it accepts XML calls so that RBEROST can interact with it directly, but users also have the option of manually interacting with the server through the webpage at https://neos-server.org/neos/solvers/lp:CPLEX/AMPL.html.

3.2 Optimization Variables and Constraints

RBEROST includes four categories of BMPs:

1. **Point Source BMPs.** Point source BMP options include water treatment plant low-cost retrofits to facilitate nutrient removal in wastewater effluent.
2. **Urban BMPs.** Urban BMP options include practices applied to developed land and may serve functions such as increasing evaporation from standing surface water, infiltration of ponded water into soil media, percolation of infiltrated water into groundwater, filtration of particulate matter, denitrification, or outflow through an orifice or weir, among others.
3. **Agricultural BMPs.** Agricultural BMP options include practices applied to agricultural land and may serve functions such as slowing runoff flow velocities from cropland areas, increasing infiltration into underlying soils, routing runoff through pools and basins, and adjusting fertilizer application or other farming practices to slow and reduce nutrient transport to waterbodies, among others. RBEROST only treats row crop area with agricultural BMPs, and 'ag' throughout the code and documentation refers only to row crop area.
4. **Riparian Buffer BMPs.** Riparian Buffer BMP options include the conversion of land within riparian areas to either herbaceous/grassed or forested land. Such practices slow water as it approaches the stream or river and increases infiltration. Nutrients may then be removed by soil processes.

RBEROST can consider removal of both total nitrogen (TN) and total phosphorus (TP) simultaneously. Table 3.1 summarizes the optimization variables included in the model and their associated constraints. Constraints on the point source BMPs are binary, such that the model chooses to implement or not to implement point source BMPs (in this case, low-cost wastewater treatment plant retrofits) based on the associated cost and removal efficiency. Constraints on the urban and agricultural BMPs relate to the fraction of each land area that is treated. The model chooses to treat a fraction of urban or agricultural land area with each BMP based on the associated cost and removal efficiency, however the sum of all fractions must be between 0 and 1. No agricultural or urban land can be treated by two BMPs. Additional constraints exist on infiltration-based BMPs, as the model will not implement them in catchments with very low infiltration values, and on porous pavement BMPs, as the model will only implement these on roadways and parking lots. Constraints on riparian buffer BMPs relate to the length (in ft) of stream bank available to be treated, where riparian BMPs can only be implemented on currently un-buffered stream lengths. Initial conditions for the optimization model reflect current practices, where the selected point source, urban, and agricultural BMPs have not yet been implemented and the load is set to the baseline annual nutrient load delivered by the upstream reaches to the target waterbodies.

Table 3.1: A summary of RBEROST model variables, constraints, and initial conditions

Optimization Variable Description	Optimization Variable Name in AMPL Model File	Constraint Description	Constraint	Initial Conditions
Per-catchment agricultural BMP selection	agBMP_bin	Choose whether to implement	Binary (0 or 1)	0
Per-catchment urban BMP selection	urbanBMP_bin	Choose whether to implement	Binary (0 or 1)	0
Point Source BMP	point_dec	Choose whether to implement	Binary (0 or 1)	0
Urban BMP	urban_frac	Fraction of urban and treated	Fraction >= User Specified Min * Fraction urban area that is suitable. Fraction <= User Specified Max * Fraction urban area that is suitable. Sum Fractions <= 1. Sum Fractions for pavement BMPs <= Fraction urban area that is roads	0
Agricultural BMP	ag_frac	Fraction of row crop land treated	Fraction >= User Specified Min. Fraction <= User Specified Max. Sum Fractions <= 1	0
Riparian Buffer BMP	ripbuf_length	Length of stream reach treated	Length <= Unbuffered Stream length * User specified max fraction. Removal along all Length <= Riparian loads. Sum all lengths <= total unbuffered bank length	0

3.3 Objective Function

The objective function minimizes BMP implementation costs while achieving a reduction in baseline nutrient loadings to the specified targets. Minimized costs are a function of the costs to implement point source, urban, agricultural, and riparian buffer BMPs in the catchments included in the optimization model (Eq. 3.1).

$$Cost_{Minimized} = \sum_{i=1}^{n} (\sum_{j=1}^{p} (Cost_{Agricultural,i,j}) + \sum_{k=1}^{q} (Cost_{Urban,i,k}) + Cost_{PointSource,i} + \sum_{l=1}^{r} (Cost_{Riparian\ Buffers,i,l})) \quad (3.1)$$

In Eq 3.1 and following equations, $i = 1...n$ are NHDPlus catchments, $j = 1...p$ are agricultural BMPs, $k = 1...q$ are urban BMPs, $l = 1...r$ are Riparian Buffer BMPs, all costs are in 2019 USA dollars, and individual parts of Eq. 3.1 are defined below in Eq. 3.2, Eq. 3.3, Eq. 3.4, Eq. 3.5, and Eq. 3.6 .

$$Cost_{Agricultural,i,j} = (Fraction_{\ TreatedLand,i,j}) * (Agricultural\ Land\ Area_i) * Ag\ Cost\ Adjustment * (Capital\ Costs_{i,j} + O\&M\ Costs_{i,j}) \quad (3.2)$$

$Cost_{Agricultural,i,j}$ describes the costs of implementing BMP *j* in row crop agricultural fields. $Fraction_{\ TreatedLand,i,j}$ ranges 0-1 and is treated as a variable in the optimization problem, $Agricultural\ Land\ Area_i$ is the amount of rowcrop area in catchment *i*, $Ag\ Cost\ Adjustment$ reflects the difference in Environmental Quality Incentives Program (EQIP) base payments and actual costs of agricultural BMPs, $Capital\ Costs_{i,j}$ represent the annualized capital costs of BMP *j*, and $O\&M\ Costs_{i,j}$ represent the annual operations and maintenance costs of BMP *j*. Costs may differ by catchment *i*.

$$Cost_{Urban,i,k} = (Capital\ Costs_k + O\&M\ Costs_k) * Urban\ Cost\ Adjustment_i * WQv_k \quad (3.3)$$

$Cost_{Urban,i,k}$ describes the cost of treating stormwater runoff from urbanized land in catchment *i* with stormwater BMP *k*. $Captial\ Costs_k$ and $O\&M\ Costs_k$ are base costs for urban BMPs and do not differ by catchment. The $Urban\ Cost\ Adjustment_i$ value scales urban costs between 1- and 3-times base costs based on the expected amount of retrofitting as determined by the intensity of urban development in catchment *i* (Hill, et al., 2016; Voorhees, 2016; Yang, et al., 2018). WQv_k is the water quality treatment volume (in cubic feet) of runoff that can be treated by BMP *k* defined as

$$WQv_k = Urban\ design\ depth_k * Rv_i * (Urban\ Land\ Area_i) * (Fraction_{\ TreatedLand,i,k}) \quad (3.4)$$

$Urban\ design\ depth_k$ is user-specified and determines the treatment capacity of stormwater BMP *k*. $Urban\ design\ depth_k$ is defined by the user in inches and is converted to feet within RBEROST. Rv_i is the runoff coefficient for catchment *i* calculated as $0.05 + 0.009 * Percent\ Site\ Imperviousness_i$ (Schueler, 1987; Vermont Agency of Natural Resources, 2017). $Urban\ Land\ Area_i$ is the area of urbanized land in catchment *i* (in ft^2), and $Fraction_{TreatedLand,I,k}$ ranges 0-1 and is treated as a variable in the optimization problem.

$$Cost_{PointSource,i} = (0,1)_i * (Capital\ Costs_i + O\&M\ Costs_i) \quad (3.5)$$

$Cost_{PointSource,i}$ describes the total costs of retrofitting WWTPs given a binary decision $(0,1)_i$ of whether to implement. This binary decision is treated as a variable in the optimization model.

Capital Costs and *O&M Costs* refer to annual operations and maintenance costs, and differ by WWTP in catchment *i.*

$$Cost_{Riparian\ Buffers,i.l} = Length\ _{Treated\ Bank,i,l} * Ag\ Cost\ Adjustment * (Capital\ Costs_{i,l} + O\&M\ Costs_{i,l}) \qquad (3.6)$$

$Cost_{Riparian\ Buffers,i,l}$ describes the costs of converting riparian zones into forested or grassed buffers, $Length_{TreatedBank,i,l}$ is a variable in the optimization and ranges from 0 ft to twice the length of stream reach (reflecting total bank length in ft) in catchment *i.* $Capital\ Costs_{i,l}$ and $O\&M\ Costs_{i,l}$ are base costs from EQIP, and are adjusted with *Ag Cost Adjustment* as in Eq. 3.2.

3.4 Optimization Parameters

Sources for BMP costs (and other parameters) in the Upper Connecticut case study are given in Table 3.2. The water quality volume (WQ_v), or treated volume, reflects Vermont Water Quality Treatment Standards (Vermont Agency of Natural Resources, 2017). Table 3.2 summarizes parameters in RBEROST and their sources for the case study described later (Ator, 2019; Detenbeck, ten Brink, et al., 2018; Detenbeck, Piscopo, et al., 2018; Heris et al., 2020; Hill et al., 2015; Houle et al., 2019; Jin et al., 2019; JJ Environmental, 2015; McKay et al., 2012; National Atmospheric Deposition Program, 2021; New Hampshire Department of Environmental Services, 2020a, 2020b; Soil Survey Staff, 2020a, 2020b; U.S. Department of Agriculture, Natural Resources Conservation Service, n.d.; U.S. Department of Agriculture Staff, 2021; UVM Spatial Analysis Lab, 2019; Verdin, 2017; Vermont Agency of Natural Resources, 2017; Voorhees, 2016; White et al., 2019; Yang et al., 2018). Table 3.3 summarizes additional parameters used for uncertainty analysis and their data sources (Ator, 2019; Dell et al., 2016; Hill et al., 2015; Houle et al., 2019; Jin et al., 2019; McKay et al., 2012; New Hampshire Department of Environmental Services, 2020a, 2020b; Schueler, 1987; Soil Survey Staff, 2020a, 2020b; U.S. Department of Agriculture Staff, 2021; U.S. Geological Survey, National Geospatial Program, 2020; U.S. Geological Survey & U.S. Department of Agriculture, Natural Resources Conservation Service, 2013; Verdin, 2017; Vermont Agency of Natural Resources, 2017; White et al., 2019; Wickham et al., 2017).

Table 3.2: A summary of RBEROST parameters.

Optimization Parameter Description	**Optimization Parameter Name in AMPL Model and Data Files**	**Data Source**
Nitrogen baseline annual average loading data per catchment to each TN target	baseloads_N1…baseloadsNn	Northeastern Regional SPARROW model (Ator, 2019) annual average loadings output (ne_sparrow_model_output_tn.txt) modified with pollutant discharge data from NPDES Permit No. NH0100200 (New Hampshire Department of Environmental Services 2020a)
Phosphorus baseline annual average loading data per catchment to each TP	baseloads_P1…baseloadsPn	Northeastern Regional SPARROW model (Ator, 2019) annual average loadings output (ne_sparrow_model_output_tp.txt)
Nitrogen baseline annual riparian loading data per catchment to each TN target	riparianload_N1…riparianload_Nn	Methodology from Pollutant Removal Credits for Buffer Restoration in MS4 Permits Final Panel Report (Houle et al. 2019) using land cover data from the National Landcover Database (Jin et al. 2019, Yang et al. 2018) and river reach shape files from NHDPlus V2 (McKay et al. 2012)
Phosphorus baseline annual riparian loading data per catchment to each TP target	riparianload_P1…riparianload_Pn	Methodology from Pollutant Removal Credits for Buffer Restoration in MS4 Permits Final Panel Report (Houle et al. 2019) using land cover data from the National Landcover Database (Jin et al. 2019, Yang et al. 2018) and river reach shape files from NHDPlus V2 (McKay et al. 2012)
Available urban and row crop area per catchment	area	Urban area: Northeastern Regional SPARROW Model (Ator, 2019) input data (ne_sparrow_model_input.txt) - “urban_km2” field. Row crop/Agricultural area: Percent cropland data from the 2011 NLCD database (catchment-specific data per state downloaded from StreamCat; Hill et al. 2015) - “PctCrop2011Cat” field
Percent of Urban land that is	urban_bmp_implementationpotential	Hydrologic soil group: gNATSGO (Soil Survey Staff,

suitable for BMP implementation		2020a; Soil Survey Staff, 2020b) and USDA NRCS 2009. River Reach shape files from NHD plus v2 (McKay et al. 2012). Building footprints from a national dataset of rasterized building footprints for the U.S. (Heris et al., 2020). High resolution impervious area from the Vermont Base Landcover 2015 database (UVM Spatial Analysis Lab, 2019)
Length of streambank that is not already buffered per catchment	unbuffered_banklength	Land cover data from the National Landcover Database (Jin et al. 2019, Yang et al. 2018) and river reach shape files from NHDPlus V2 (McKay et al. 2012)
Total stream bank length (2 x reach length) per catchment	total_banklength	River reach shape files from NHDPlus V2 (McKay et al. 2012)
Capital costs associated with implementing agricultural BMPs	ag_costs_capital	U.S. Department of Agriculture; Environmental Quality Incentives Program state-by-state cost sheets from 2020. For the case study, Massachusetts costs were used when New Hampshire or Vermont costs were not available (U.S. Department of Agriculture Staff, 2021)
Operations and maintenance costs associated with implementing agricultural BMPs	ag_costs_operations	U.S. Department of Agriculture; Environmental Quality Incentives Program state-by-state cost sheets from 2020. For the case study, Massachusetts costs were used when New Hampshire or Vermont costs were not available (U.S. Department of Agriculture Staff, 2021)
Costs associated with implementing wastewater treatment plant retrofits for nutrient removal	point_costs	NEIWPCC Final Report - Low Cost Retrofits for Nitrogen Removal at Wastewater Treatment Plants in the Upper Long Island Sound Watershed (JJ Environmental, 2015)
Costs associated with implementing urban BMPs	urban_costs	WMOST Theoretical Documentation (Detenbeck, ten Brink et al., 2018; Table 7-2)
Capital costs associated with	ripbuf_costs_capital	U.S. Department of Agriculture; Environmental Quality

implementing riparian buffer BMPs		Incentives Program state-by-state cost sheets from 2020.
Operations and maintenance costs associated with implementing riparian buffer BMPs	ripbuf_costs_operations	U.S. Department of Agriculture; Environmental Quality Incentives Program state-by-state cost sheets from 2020.
Calculated urban volumetric runoff coefficient, based on percent impervious area per catchment	runoff_coeff_urban	Imperviousness data from National Land Cover Database, 2011 (catchment-specific data per state downloaded from StreamCat; Hill et al. 2015) - “PctImp2011Cat” field. Data used to calculate urban runoff coefficient, based on Vermont Stormwater Manual equation (0.05+0.009*PctImp2011Cat; Vermont Agency of Natural Resources, 2017)
Adjustments to urban BMP costs based on the expected amount of retrofitting (proportional to relative intensity of urban land per catchment)	urban_cost_adjustment_coef	Development intensity data by catchment from the National Land Cover Database, 2011 (catchment-specific data per state downloaded from StreamCat; Hill et al. 2015) - “PctUrbOp2011Cat,” “PctUrbLo2011Cat,” “PctUrbMd2011Cat,” and “PctUrbHi2011Cat” fields. Cost adjustment factors from Opti-Tool documentation (Voorhees, 2016)
Agricultural nitrogen removal efficiency per BMP option and HUC12	ag_effic_N	“Fert_20” and “Manure_Injection” BMPs: Efficiencies provided by EPA via NRCS. Remaining agricultural BMPs: Efficiencies summarized by HUC12 (or if unavailable, HUC10 and/or HUC8) based on ACRE database (White et al., 2019)
WWTP-specific nitrogen removal efficiency	point_effic_N	NEIWPCC Final Report - Low Cost Retrofits for Nitrogen Removal at Wastewater Treatment Plants in the Upper Long Island Sound Watershed (JJ Environmental, 2015)
Urban nitrogen removal efficiency per BMP option and catchment	urban_effic_N	New Hampshire MS4 Permit BMP Performance Curves (New Hampshire Department of Environmental Services, 2020b). Hydrologic soil group: gNATSGO

		(Soil Survey Staff, 2020a; Soil Survey Staff, 2020b; USDA NRCS 2009). Land cover data from the National Landcover Database (Jin et al. 2019, Yang et al. 2018)
Agricultural phosphorus removal efficiency per BMP option and HUC12	ag_effic_P	“Fert_20” and “Manure_Injection” BMPs: Efficiencies provided by EPA via NRCS. Remaining agricultural BMPs: Efficiencies summarized by HUC12 (or if unavailable, HUC10 and/or HUC8) based on ACRE database (White et al., 2019)
WWTP-specific phosphorus removal efficiency	point_effic_P	Not applicable
Urban phosphorus removal efficiency per BMP option and catchment	urban_effic_P	New Hampshire MS4 Permit BMP Performance Curves (New Hampshire Department of Environmental Services, 2020b). Hydrologic soil group: gNATSGO (Soil Survey Staff, 2020a; Soil Survey Staff, 2020b; USDA NRCS 2009). Land cover data from the National Landcover Database (Jin et al. 2019, Yang et al. 2018)
Total nitrogen removal by riparian buffers per catchment for each TN target	riparianremoval_N1…riparianremoval_Nn	Annual removal rates are a function of riparian loading and nutrient removal efficiency of riparian buffers. Loading is dependent on the amount of urbanized land within a 400 ft riparian zone, and efficiency is dependent on slope and hydrologic soil group (Houle et al. 2019). Land cover data from the National Landcover Database (Jin et al. 2019, Yang et al. 2018). River reach shape files from NHDPlus V2 (McKay et a.l 2012). Hydrologic soil group: gNATSGO (Soil Survey Staff, 2020a; Soil Survey Staff, 2020b). Slope data from Hydrologic Derivatives for Modeling and Analysis by the USGS (Verdin, 2017)
Total phosphorus removal by riparian buffers per catchment for each TP target	riparianremoval_P1…riparianremoval_Pn	Annual removal rates are a function of riparian loading and nutrient removal efficiency of riparian buffers. Loading is dependent on the amount of urbanized land

		within a 400 ft riparian zone, and efficiency is dependent on slope and hydrologic soil group (Houle et al. 2019). Land cover data from the National Landcover Database (Jin et al. 2019, Yang et al. 2018). River reach shape files from NHDPlus V2 (McKay et al. 2012). Hydrologic soil group: gNATSGO (Soil Survey Staff, 2020a, 2020b). Slope data from Hydrologic Derivatives for Modeling and Analysis by the USGS (Verdin, 2017)
Total nitrogen load allowed after reduction for all TN targets	loads_lim_N1…loads_lim_Nn	User supplied
Total phosphorus load allowed after reduction for all TP targets	loads_lim_P1…loads_lim_Pn	User supplied
The sum of SPARROW model annual nitrogen average loads delivered to each TN target that are not associated with point sources, urban area, or row crop area (e.g., loads from septic sources)	other_loads_N1…other_loads_Nn	Northeastern Regional SPARROW model (Ator, 2019) annual average loadings output (ne_sparrow_model_output_tn.txt) modified with N depositional data from the National Atmospheric Deposition Program (National Atmospheric Deposition Program, 2021)
The sum of SPARROW model annual phosphorus average loads delivered to each TN target that are not associated with point sources, urban area, or row crop area (e.g., loads from bedrock leaching)	other_loads_P1…other_loads_Pn	Northeastern Regional SPARROW model (Ator, 2019) annual average loadings output (ne_sparrow_model_output_tp.txt)
Agricultural BMP costs reflected in ag_costs	agcost_frac	Correspondence with EPA and NRCS state conservationist for Vermont

represent base payment costs (75% of actual costs). The agcost_frac parameter reflects 100% of costs (100/75 = 1.33).		
Conversion from acre-ft to cubic ft	acfttoft3	Not applicable
Conversion for precipitation: inches to ft	pcp	Not applicable
Parameter to ensure that agricultural BMPs are implemented on at least some minimum acreage. Currently set to 0.	agBMP_minarea	Not applicable
Minimum allowed percent implementation of urban BMPs per catchment	urban_frac_min	User supplied
Maximum allowed percent implementation of urban BMPs per catchment	urban_frac_max	User supplied
Minimum allowed percent implementation of agricultural BMPs per catchment	ag_frac_min	User supplied
Maximum allowed percent implementation of agricultural BMPs per catchment	ag_frac_max	User supplied
Minimum allowed percent implementation of riparian buffer BMPs per catchment	ripbuf_frac_min	User supplied

Maximum allowed percent implementation of riparian buffer BMPs per catchment	ripbuf_frac_max	User supplied
The specified design depth for each urban BMP	urban_design_depth	User supplied

Table 3.3: A summary of RBEROST parameters used for uncertainty analyses.

Optimization Parameter Description	**Optimization Parameter Name in AMPL Model and Data Files**	**Data Source**
Standard error of nitrogen baseline annual average loading data per catchment to each TN target	baseloads_N1_se…baseloadsNn_se	Northeastern Regional SPARROW model (Ator, 2019) annual average loadings output (ne_sparrow_model_output_tn.txt) modified with standard error of regression of pollutant discharge over time from NPDES Permit No. NH0100200 (New Hampshire Department of Environmental Services 2020a)
Standard error of nitrogen baseline annual average loading data per catchment to each TP target	baseloads_P1_se…baseloadsPn_se	Northeastern Regional SPARROW model (Ator, 2019) annual average loadings output (ne_sparrow_model_output_tp.txt)
Standard error of available urban and row crop area per catchment	area_se	Coefficient of variation for NLCD data and StreamCat values from Wickham et al., 2017 and Hill et al., 2015. Uncertainty in Incremental area from U.S. Geological Survey and U.S. Department of Agriculture, Natural Resources Conservation Service, 2013, McKay et al., 2012, and U.S. Geological Survey, National Geospatial Program, 2020.
Standard error of capital costs associated with implementing agricultural BMPs	ag_costs_capital_se	U.S. Department of Agriculture; Environmental Quality Incentives Program state-by-state cost sheets from 2018 - 2021. For the case study, Massachusetts costs were used when New Hampshire or Vermont costs were not available (U.S. Department of Agriculture Staff, 2021)
Standard error of operations and	ag_costs_operations_se	U.S. Department of Agriculture; Environmental Quality Incentives Program state-by-state cost sheets

maintenance costs associated with implementing agricultural BMPs		from 2018 - 2021. For the case study, Massachusetts costs were used when New Hampshire or Vermont costs were not available (U.S. Department of Agriculture Staff, 2021)
Standard error of capital costs associated with implementing riparian buffer BMPs	ripbuf_costs_capital_se	U.S. Department of Agriculture; Environmental Quality Incentives Program state-by-state cost sheets from 2018 - 2021 (U.S. Department of Agriculture Staff, 2021)
Standard error of operations and maintenance costs associated with implementing riparian buffer BMPs	ripbuf_costs_operations_se	U.S. Department of Agriculture; Environmental Quality Incentives Program state-by-state cost sheets from 2018 - 2021 (U.S. Department of Agriculture Staff, 2021)
Standard error of the calculated urban volumetric runoff coefficient, based on percent impervious area per catchment	runoff_coeff_urban_se	Standard error around the slope and intercept of the equation presented by Vermont Agency of Natural Resources, 2017 calculated from the data originally published by Schueler, 1987.
Standard error of adjustments to urban BMP costs based on the expected amount of retrofitting (proportional to relative intensity of urban land per catchment)	urban_cost_adjustment_coef_se	Sampled from development intensity data by catchment from the National Land Cover Database, 2011 (catchment-specific data per state downloaded from StreamCat; Hill et al. 2015) - "PctUrbOp2011Cat," "PctUrbLo2011Cat," "PctUrbMd2011Cat," and "PctUrbHi2011Cat" fields. Cost adjustment factors from Opti-Tool documentation (Voorhees, 2016)
Standard error of agricultural nitrogen removal efficiency per BMP option and	ag_effic_N_se	"Manure_Injection" BMP: Dell et al. 2016. Remaining agricultural BMPs: Standard error of efficiencies summarized by HUC12 (or if unavailable, HUC10 and/or HUC8) based on ACRE database

HUC12		(White et al., 2019)
Standard error of urban nitrogen removal efficiency per BMP option and catchment	urban_effic_N_se	Sampled from values calculated according to New Hampshire MS4 Permit BMP Performance Curves (New Hampshire Department of Environmental Services, 2020b). Hydrologic soil group: gNATSGO (Soil Survey Staff, 2020a; Soil Survey Staff, 2020b). Land cover data from the National Landcover Database (Jin et al. 2019, Yang et al. 2018)
Standard error of agricultural phosphorus removal efficiency per BMP option and HUC12	ag_effic_P_se	Manure_Injection" BMP: Dell et al. 2016. Remaining agricultural BMPs: Standard error of efficiencies summarized by HUC12 (or if unavailable, HUC10 and/or HUC8) based on ACRE database (White et al., 2019)
Standard error of urban phosphorus removal efficiency per BMP option and catchment	urban_effic_P_se	Sampled from values calculated according to New Hampshire MS4 Permit BMP Performance Curves (New Hampshire Department of Environmental Services, 2020b). Hydrologic soil group: gNATSGO (Soil Survey Staff, 2020a; Soil Survey Staff, 2020b). Land cover data from the National Landcover Database (Jin et al. 2019, Yang et al. 2018)
Standard error of total nitrogen removal by riparian buffers per catchment for each TN target	riparianremoval_N1_se…riparianremoval_Nn_se	Sampled values of annual removal rates. Annual removal rates are a function of riparian loading and nutrient removal efficiency of riparian buffers. Loading is dependent on the amount of urbanized land within a 400 ft riparian zone, and efficiency is dependent on slope and hydrologic soil group (Houle et al. 2019). Land cover data from the National Landcover Database (Jin et al. 2019, Yang et al. 2018). River reach shape files from NHDPlus V2 (McKay et al. 2012). Hydrologic soil group: gNATSGO (Soil Survey Staff, 2020a; Soil Survey Staff, 2020b). Slope data from Hydrologic Derivatives for Modeling and

		Analysis by the USGS (Verdin, 2017)
Standard error of total nitrogen removal by riparian buffers per catchment for each TP target	riparianremoval_P1_se…riparianremoval_Pn_se	Sampled values of annual removal rates. Annual removal rates are a function of riparian loading and nutrient removal efficiency of riparian buffers. Loading is dependent on the amount of urbanized land within a 400 ft riparian zone, and efficiency is dependent on slope and hydrologic soil group (Houle et al. 2019). Land cover data from the National Landcover Database (Jin et al. 2019, Yang et al. 2018). River reach shape files from NHDPlus V2 (McKay et al. 2012). Hydrologic soil group: gNATSGO (Soil Survey Staff, 2020a, 2020b). Slope data from Hydrologic Derivatives for Modeling and Analysis by the USGS (Verdin, 2017)

Uncertainty in several additional parameters are not included as parameters in the model, but rather are hard coded into the command scripts. These include uncertainty around point_costs, point_effic_N and point_effic_P, other_loads_N* and other_loads_P*, and agcost_frac. Uncertainty around point costs were estimated based on the predicted contractor markup, and assuming that additional engineering services or detailed designs would have a similar markup (JJ Environmental, 2015). Point efficiencies assumed a coefficient of variation of 20% (JJ Environmental, 2015). Other loads for nitrogen and phosphorus were derived from SPARROW models (Ator, 2019). As other nitrogen loads were adjusted by changes in N deposition, uncertainty in the change in N deposition was also accounted for according to a modified version of the qualitative measure presented by (Walker et al., 2019). The ratio of actual agricultural costs to base payments (agcost_frac) was modeled as a uniform distribution between 1 and 1.67 (i.e., base payment costs ranging from 60% - 100% of actual costs).

4 Model Use

In order to run RBEROST, the user first defines specifications, including the load reduction goals and the potential BMPs to be considered for the optimization model. The user specifications are then run through the R preprocessing code, which develops and formats the AMPL files required for the NEOS server optimization run. The user can choose to generate AMPL files that either include uncertainty or do not include uncertainty analyses. The user uploads the AMPL files to the NEOS server and accesses the optimization results using information e-mailed to the user. Finally, the user runs the R postprocessing code to summarize results of the NEOS run. Figure 4.1 summarizes this workflow.

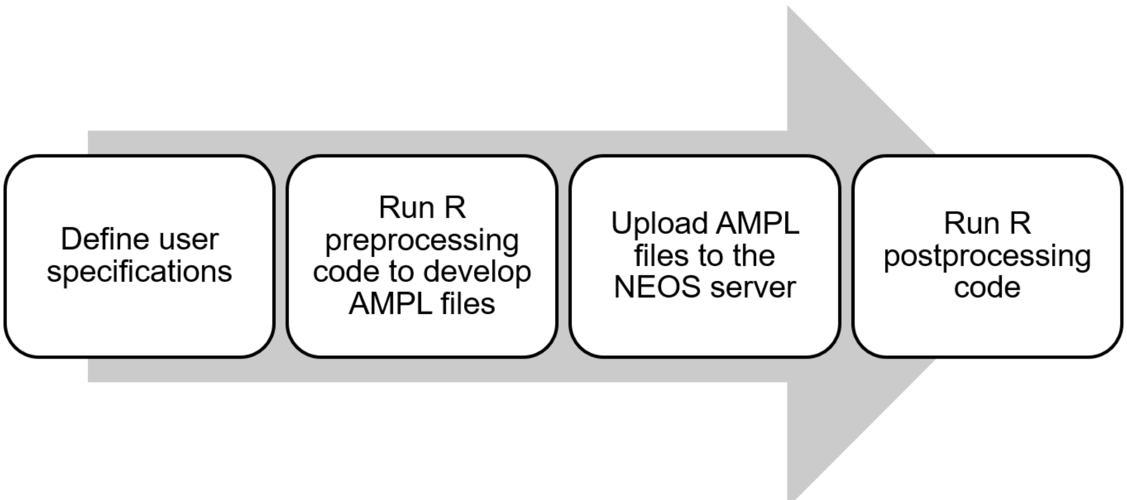


Figure 4.1: User workflow of RBEROST.

4.1 Getting Started

Users download a .zip file that contains RBEROST program files and data required for preprocessing (creation of AMPL files) and postprocessing (viewing of NEOS optimization results). It is necessary for the file structure of the zip file as well as naming conventions for Preprocessing inputs and R files to be maintained.

For the default pathways in RBEROST to work, you will need to open the file via the `RBEROST.Rproj` file. Double click the file to open the project in RStudio. Figure 4.2 shows the file in the unzipped folder. The project should open in RStudio, and the files can be seen on the bottom right. RBEROST can be executed through the `RunRBEROST.Rmd` file, which is opened by clicking on it in the `Files` pane of RStudio (Figure 4.3)

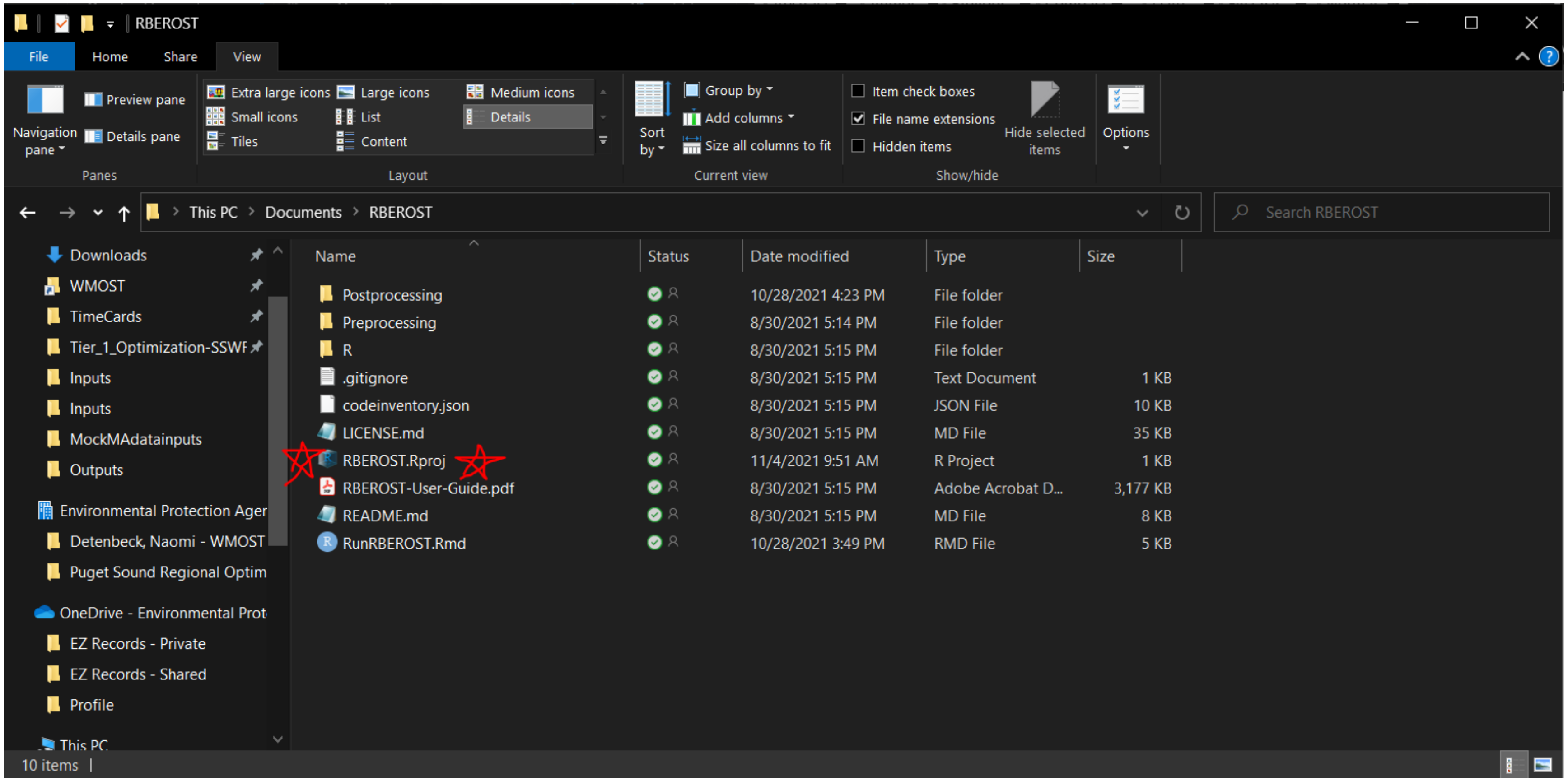


Figure 4.2: RBEROST Rproject file

RBEROST was developed in R version 4.0.5 and can be run through RStudio. RBEROST was developed with RStudio 1.4.1106. The folder structure includes R scripts to preprocess the data and postprocess optimization results. All files are accessed through the `RunRBEROST.Rmd` file. `RunRBEROST.Rmd` will source the necessary scripts to run each part of RBEROST. Scripts used in RBEROST are in the `./R` folder, where `./` refers to the file path of the unzipped folder. All scripts must retain their location and naming to be called by RBEROST. Files include `Optimization_HelperFunctions.R` which includes functions used in RBEROST, `01_Optimization_Preprocessing_gateway.R` that routes RBEROST to either preprocessing with or without uncertainty, `01_Optimization_Preprocessing.R` which creates AMPL files without uncertainty, and `01_Optimization_Preprocessing+Uncertainty.R` which creates AMPL files with uncertainty. Running RBEROST with uncertainty will also automatically produce AMPL scripts without uncertainty as well, and the user is free to ignore or save these files without uncertainty. R scripts for the postprocessor include `02_Optimization_RunShiny.R` which sources `Optimization_ServerFile.R`, `Optimization_ServerFunctions_Postprocessor.R`, `Optimization_UserInterfaceFile.R` and `Optimization_UI_Postprocessor.R` to build the Shiny application.

RBEROST relies on several R packages to run. The application was developed with tidyverse 1.3.1, reshape2 1.4.4, data.table 1.14.0, stringr 1.4.0, foreach 1.5.1, shiny 1.6.0, shinycssloaders 1.0.0, tidygraph 1.2.0, and bit64 4.0.5. If these packages are already installed, it is recommended to update. RBEROST provides code necessary to install and/or update all packages in the `RunRBEROST.Rmd` file. This code will only need to be run once, the first time RBEROST is opened, by clicking the green triangle on the first code chunk (indicated by the red circle in Figure 4.4). While the code is running, the green triangle will change to a red square (Figure 4.5). The icon will change back to a green triangle when the code is finished. During package installation, you may be prompted with a question "Do you want to install from sources the package which needs compilation?" (Figure 4.6). Select "Yes."

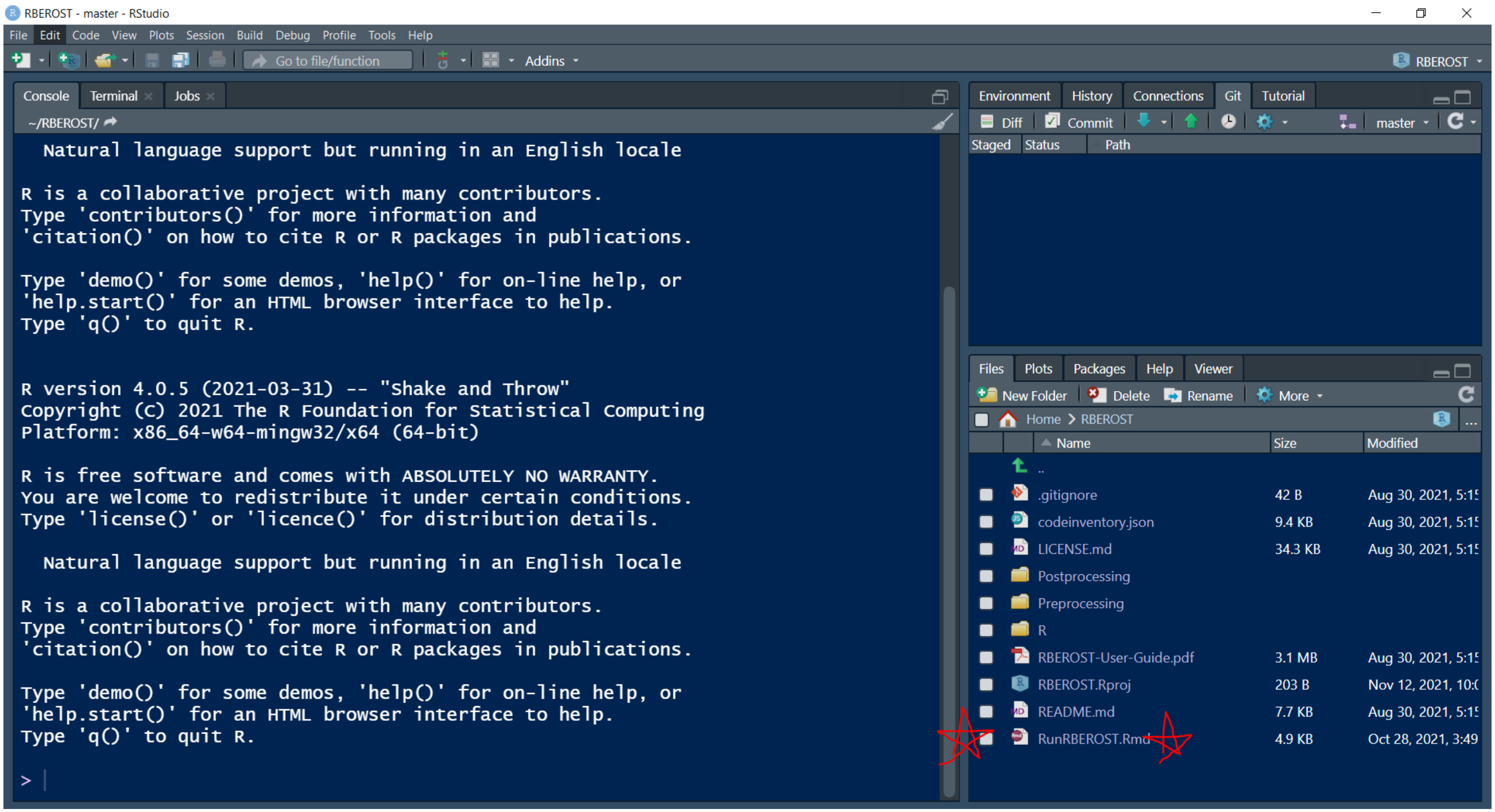


Figure 4.3: Where to find the file 'RunRBEROST.Rmd' to open RBEROST though RStudio

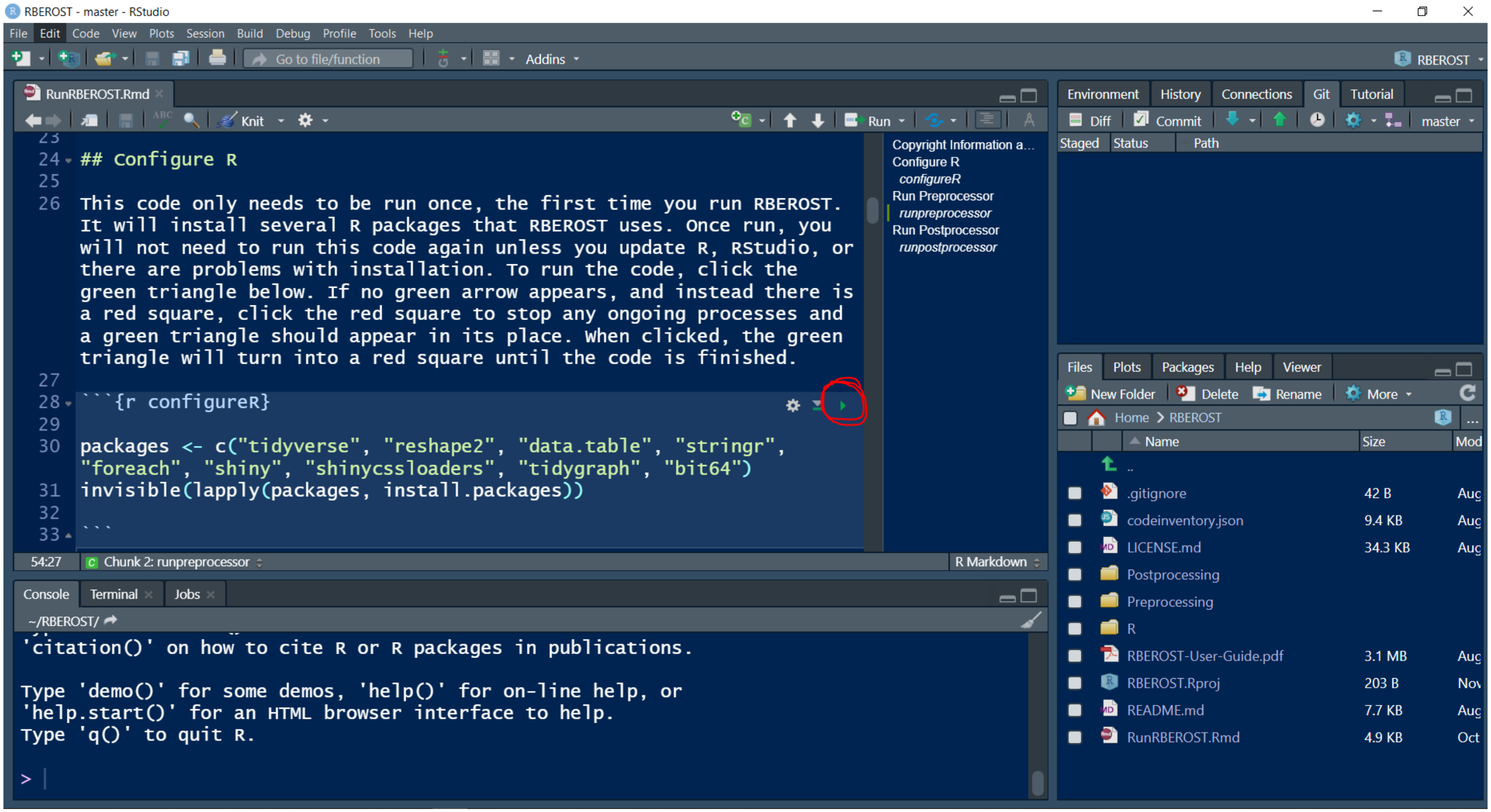


Figure 4.4: Run package installation the first time RBEROST is opened.

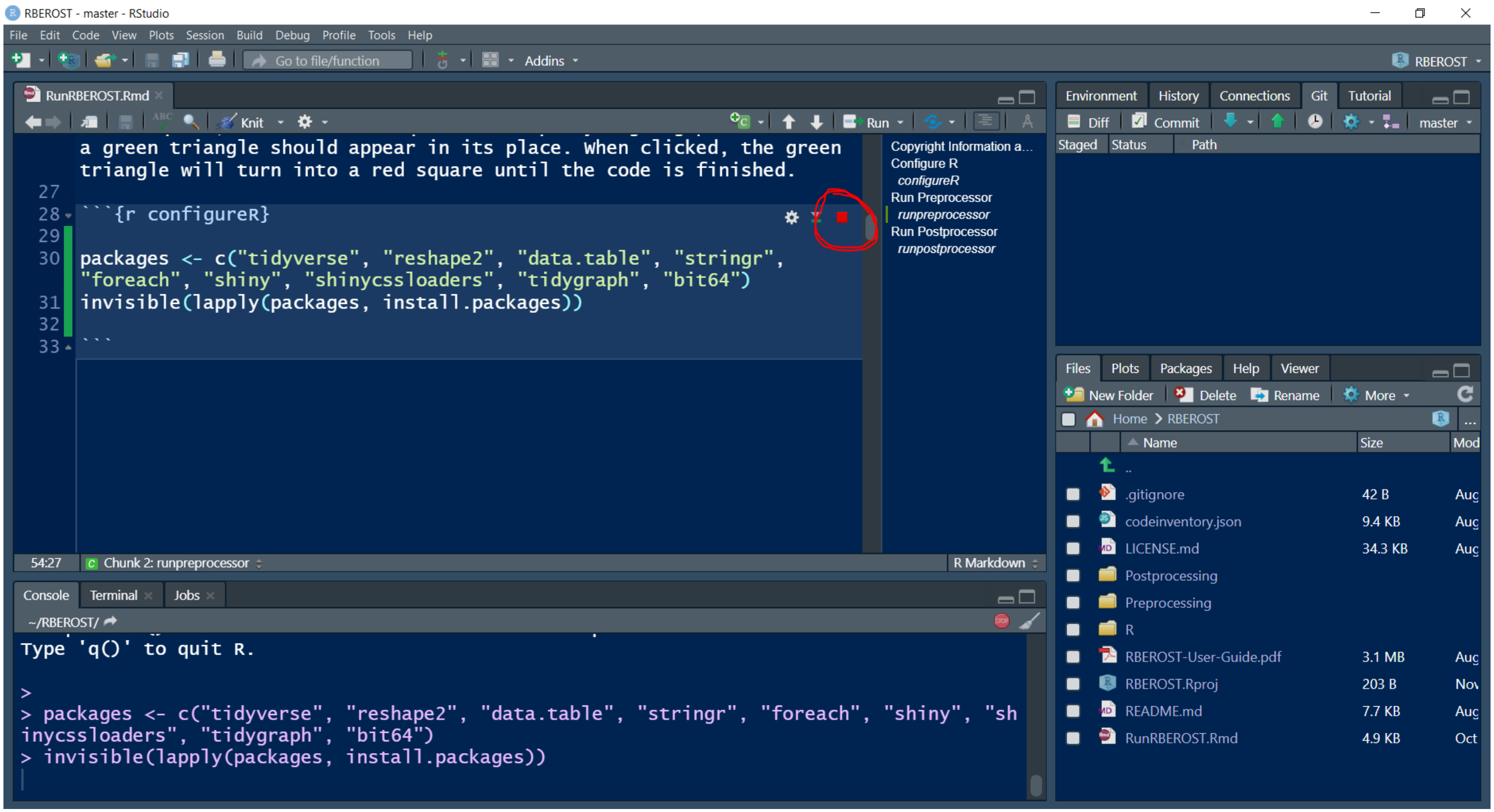


Figure 4.5: An image of what RStudio displays while code is running

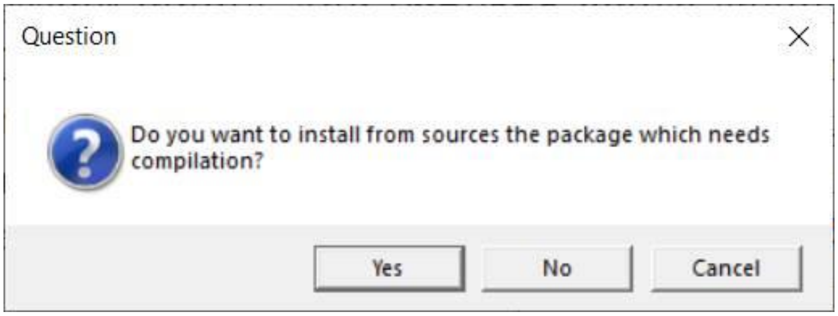


Figure 4.6: A message users may encounter while installing packages.

4.2 Preprocessing

This section details the processes required to develop the AMPL model files that are run through the NEOS server to solve the optimization. The user is required to update the Run Preprocessor section of the code in `RunRBEROST.Rmd` as well as the two csv files `01_UserSpecs_BMPs.csv` and `01_UserSpecs_loadingtargets.csv`, which are in the `./Preprocessing/Inputs` folder (Figure 4.7). To run the preprocessing step, first edit the UserSpecs files and the available options within `RunRBEROST.Rmd`. Then, click the green triangle in the upper right of the preprocessing code chunk. The green triangle will turn into a red square until the code is done running, at which point the red square will return to a green triangle.

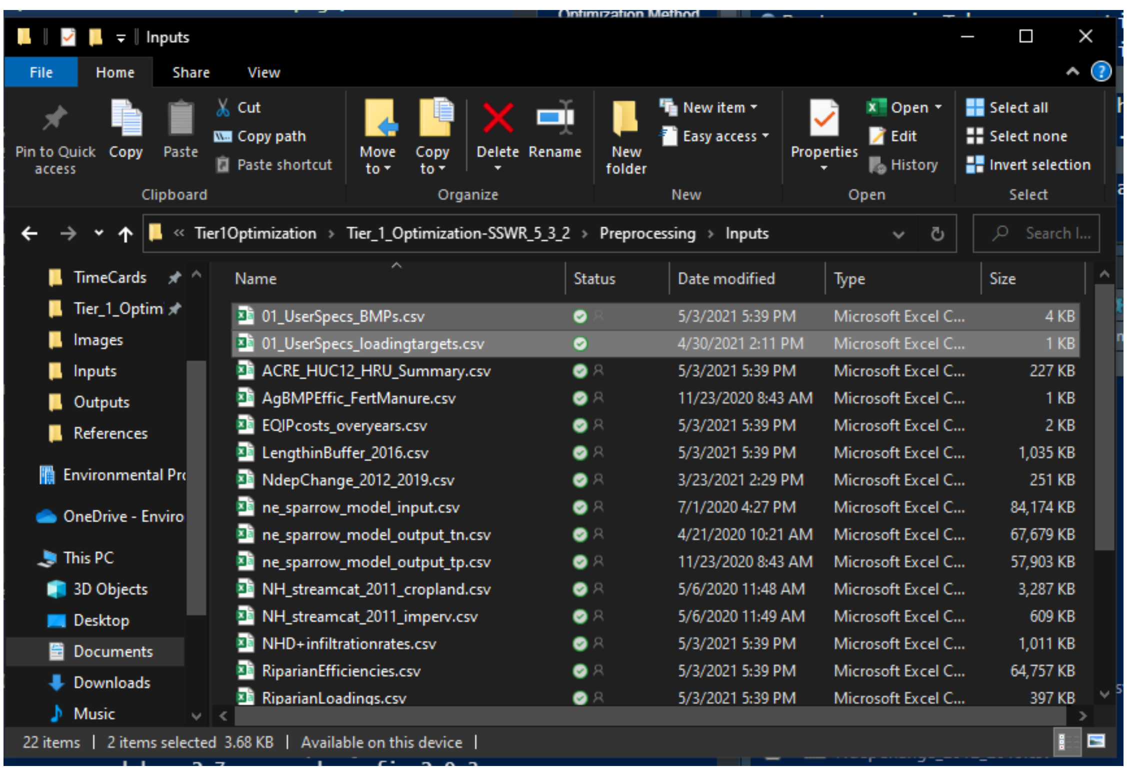


Figure 4.7: Location of the 2 User Specification csv files used in RBEROST.

4.2.1 User Specifications

There are eight parameters that users can change in the `RunRBEROST.Rmd` file. These include `InPath`, which should point to the folder with the Preprocessing inputs, and `OutPath` which should point to the folder that the AMPL files will be written to. If RBEROST is opened from the `RBEROST.Rproj` file, and if file structure is left intact, these parameters do not need to be changed. Other parameters include `horizon` which is the planning horizon used to annualize costs. The default value is 15 years. The variable `interest_rate` is the expected interest for the project, and the default value is 0.03, or 3%. `AgBMPcomparison` can be set as `“No Practice”` or `“Baseline”` and determines how the efficiencies of agricultural BMPs will be determined. `“No Practice”` represents the efficiency of a practice in isolation compared to conventional practices, and `“Baseline”` represents the efficiency of removing existing conservation practices and replacing them with a certain practice. `IncludeUncertainty` is a logical parameter and can be set as `TRUE` or `FALSE`. If `TRUE`, RBEROST will create AMPL scripts that include uncertainty calculations. If `FALSE`, it will not. Making AMPL scripts that include uncertainty may take several minutes. Variables `n.scenarios` and `scenariostepchange` only apply when running RBEROST with uncertainty. The variable `n.scenarios` sets the number of scenarios you would like to view, and `scenariostepchange` sets the relative difference between these scenarios. The different scenarios describe increasingly restrictive loading targets to be applied after the first scenario is solved with the user specifications from `01_UserSpecs_loadingtargets.csv`. Each successive scenario decreases the loading targets by the percent given in `scenariostepchange`. This variable `scenariostepchange` is equivalent to the additional margin of safety applied to each scenario. The default is to compute three scenarios with a 1% change in loading targets between scenarios. The lines of code to edit are marked with red stars in Figure 4.8.

Users specify the BMP options that will run through RBEROST within the `01_UserSpecs_BMPs.csv` file (Figure 4.9). BMP options are specified by placing a capitalized “X” in the `BMP_Selection` field of the `01_UserSpecs_BMPs.csv` file. Users can also choose a minimum and/or maximum implementation of each BMP as a fraction of total possible implementation. `01_UserSpecs_BMPs.csv` includes all the BMPs available for model optimization per BMP category (agricultural, urban, point source, and riparian buffer BMPs), state-specific capital and operations & maintenance costs to implement the BMP, units of the costs, user-specified design runoff depths for urban BMPs and buffer widths for riparian buffer BMPs.

The data provided covers the geographic extent of the Upper Connecticut River case study. To expand the geographic range of optimization, data files will need to be expanded to include ComIDs located in additional areas.

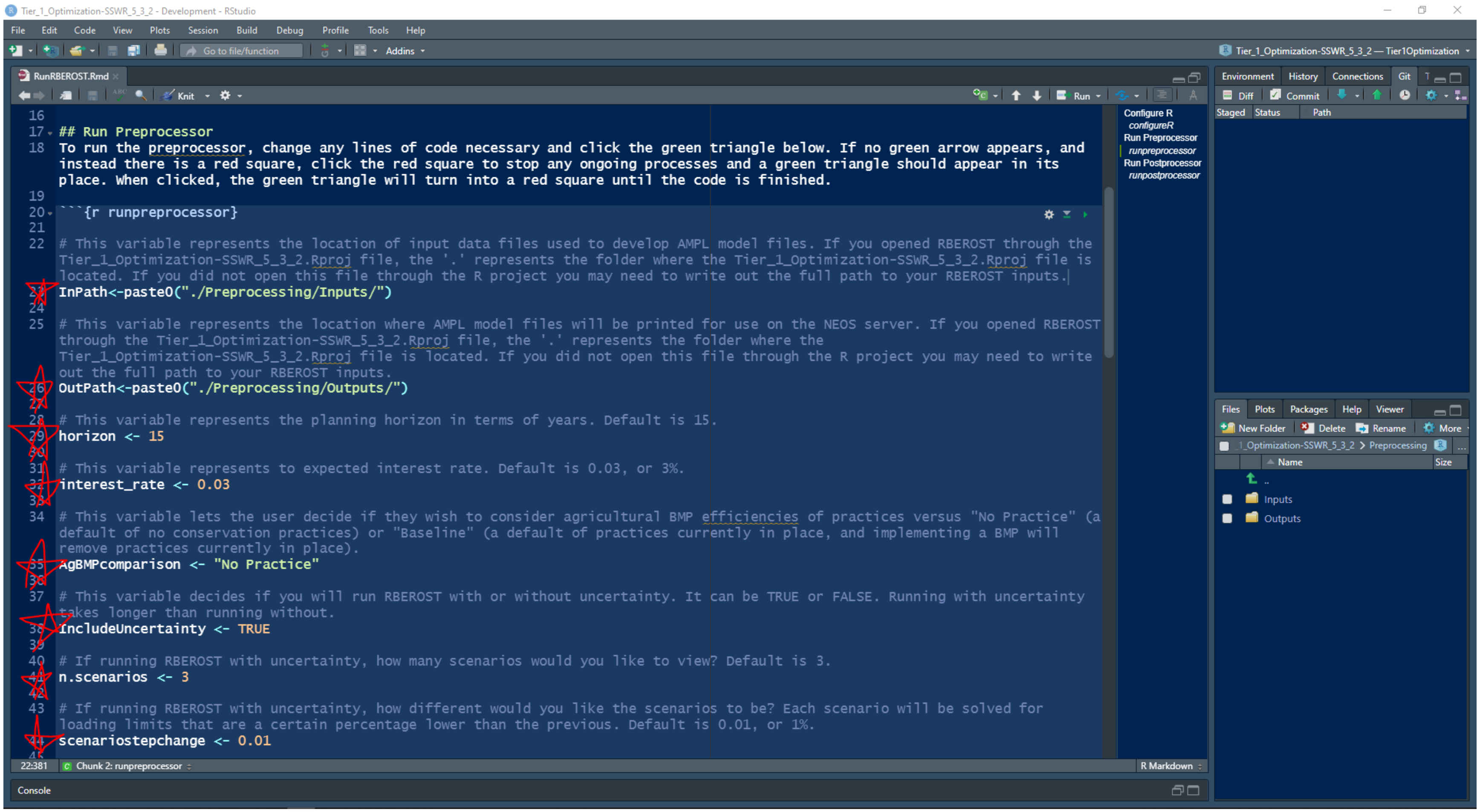
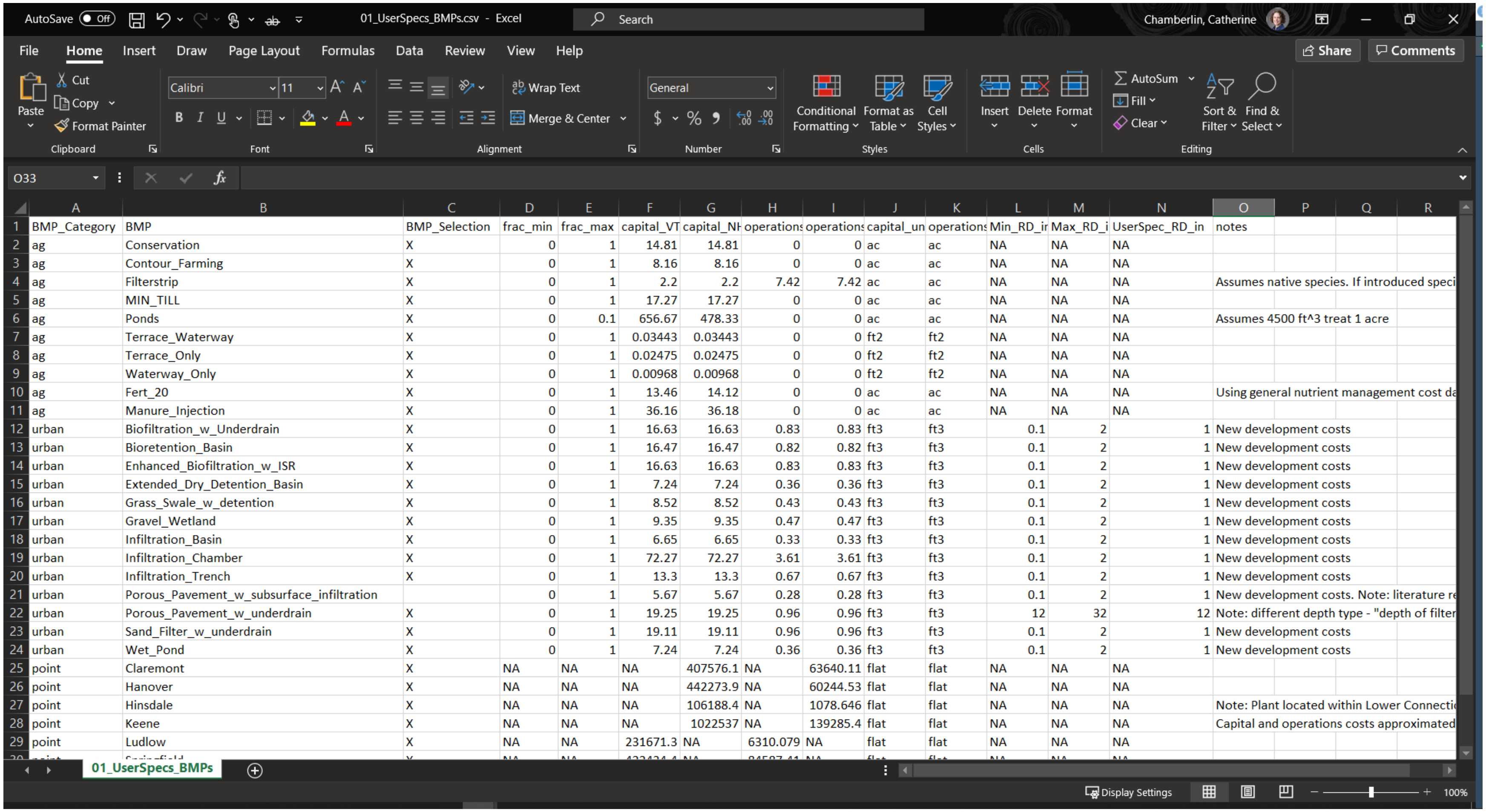


Figure 4.8: Lines of code for the user to edit in the preprocessing step.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	BMP_Category	BMP	BMP_Selection	frac_min	frac_max	capital_VT	capital_NH	operations	operations	capital_un	operations	Min_RD_in	Max_RD_i	UserSpec_RD_in	notes
2	ag	Conservation	X	0	1	14.81	14.81	0	0	ac	ac	NA	NA	NA	
3	ag	Contour_Farming	X	0	1	8.16	8.16	0	0	ac	ac	NA	NA	NA	
4	ag	Filterstrip	X	0	1	2.2	2.2	7.42	7.42	ac	ac	NA	NA	NA	Assumes native species. If introduced speci
5	ag	MIN_TILL	X	0	1	17.27	17.27	0	0	ac	ac	NA	NA	NA	
6	ag	Ponds	X	0	0.1	656.67	478.33	0	0	ac	ac	NA	NA	NA	Assumes 4500 ft^3 treat 1 acre
7	ag	Terrace_Waterway	X	0	1	0.03443	0.03443	0	0	ft2	ft2	NA	NA	NA	
8	ag	Terrace_Only	X	0	1	0.02475	0.02475	0	0	ft2	ft2	NA	NA	NA	
9	ag	Waterway_Only	X	0	1	0.00968	0.00968	0	0	ft2	ft2	NA	NA	NA	
10	ag	Fert_20	X	0	1	13.46	14.12	0	0	ac	ac	NA	NA	NA	Using general nutrient management cost da
11	ag	Manure_Injection	X	0	1	36.16	36.18	0	0	ac	ac	NA	NA	NA	
12	urban	Biofiltration_w_Underdrain	X	0	1	16.63	16.63	0.83	0.83	ft3	ft3	0.1	2	1	New development costs
13	urban	Bioretention_Basin	X	0	1	16.47	16.47	0.82	0.82	ft3	ft3	0.1	2	1	New development costs
14	urban	Enhanced_Biofiltration_w_ISR	X	0	1	16.63	16.63	0.83	0.83	ft3	ft3	0.1	2	1	New development costs
15	urban	Extended_Dry_Detention_Basin	X	0	1	7.24	7.24	0.36	0.36	ft3	ft3	0.1	2	1	New development costs
16	urban	Grass_Swale_w_detention	X	0	1	8.52	8.52	0.43	0.43	ft3	ft3	0.1	2	1	New development costs
17	urban	Gravel_Wetland	X	0	1	9.35	9.35	0.47	0.47	ft3	ft3	0.1	2	1	New development costs
18	urban	Infiltration_Basin	X	0	1	6.65	6.65	0.33	0.33	ft3	ft3	0.1	2	1	New development costs
19	urban	Infiltration_Chamber	X	0	1	72.27	72.27	3.61	3.61	ft3	ft3	0.1	2	1	New development costs
20	urban	Infiltration_Trench	X	0	1	13.3	13.3	0.67	0.67	ft3	ft3	0.1	2	1	New development costs
21	urban	Porous_Pavement_w_subsurface_infiltration		0	1	5.67	5.67	0.28	0.28	ft3	ft3	0.1	2	1	New development costs. Note: literature re
22	urban	Porous_Pavement_w_underdrain	X	0	1	19.25	19.25	0.96	0.96	ft3	ft3	12	32	12	Note: different depth type - "depth of filter
23	urban	Sand_Filter_w_underdrain	X	0	1	19.11	19.11	0.96	0.96	ft3	ft3	0.1	2	1	New development costs
24	urban	Wet_Pond	X	0	1	7.24	7.24	0.36	0.36	ft3	ft3	0.1	2	1	New development costs
25	point	Claremont	X	NA	NA	NA	407576.1	NA	63640.11	flat	flat	NA	NA	NA	
26	point	Hanover	X	NA	NA	NA	442273.9	NA	60244.53	flat	flat	NA	NA	NA	
27	point	Hinsdale	X	NA	NA	NA	106188.4	NA	1078.646	flat	flat	NA	NA	NA	Note: Plant located within Lower Connectic
28	point	Keene	X	NA	NA	NA	1022537	NA	139285.4	flat	flat	NA	NA	NA	Capital and operations costs approximated
29	point	Ludlow	X	NA	NA	231671.3	NA	6310.079	NA	flat	flat	NA	NA	NA	

Figure 4.9: A screenshot of the BMP UserSpecs file with the specifications for the Upper Connecticut River case study.

4.2.1.1 BMP Costs

Implementation of the majority of the BMPs requires both capital and operations & maintenance costs, except for several agricultural BMPs. All default costs are scaled to 2019$. The user can override these costs if they choose costs that are representative of costs per treated area for agricultural and riparian buffer BMPs, costs per treated volume for urban BMPs, and costs per retrofit implementation for point source BMPs. Urban costs are specified assuming implementation in new developments rather than with retrofitting. RBEROST will adjust these costs internally based on the expected amount of retrofitting and complexity of installations in each catchment as judged by development intensity.

If cost data from additional states beyond the geographic extent of the Upper Connecticut River case study must be included, these are added to the file `01_UserSpecs_BMPs.csv`. For each additional state, two columns must be added to this file, following the naming convention of "`captial_[two letter state code]`" and "`operations_[two letter state code]`". The units for these costs must be the same across all states included, and units are indicated in the "`capital_units`" and "`operations_units`" fields. Possible units are "`ac`", "`ft2`", "`yd2`" and "`km2`" for agricultural and riparian buffer BMPs, "`ft3`" for urban BMPs, and "`flat`" for point source BMPs. If optimizations that include uncertainty are performed for geographic ranges that extend to states beyond the ones in the case study, additional columns must also be added to the file `EQIPcosts_overyears.csv`. Column headings should follow the naming conventions of "`capital_[two letter state code]_[year]`" or "`operations_[two letter state code]_[year]`", and cost units must be consistent across all columns.

4.2.1.2 Urban BMP Runoff Depth and Riparian Buffer Width

Nutrient removal efficiency resulting from implementation of urban BMPs depends on the BMP design runoff depth. The `01_UserSpecs_BMPs.csv` file specifies default runoff depths for BMPs, but the user may specify alternate runoff depths in the `UserSpec_RD_in` field. The `Min_RD_in` and `Max_RD_in` fields specify the minimum and maximum runoff depths (i.e., upper and lower limit) that the user can specify (in inches) for the urban BMP `UserSpec_RD_in` field. RBEROST does not use the `Min_RD_in` and `Max_RD_in` fields for any calculations or data modifications, the fields are just for reference. Specification of riparian buffer width also affects the amount of nutrients removed by the buffers. Users can specify buffer widths in the same column, `UserSpec_RD_in`. A note is provided that, though presented to the user in the same column, the units for riparian buffer BMPs are different than for urban BMPs.

4.2.1.3 Load Reduction Goal

Users can specify loading targets in the `01_UserSpecs_loadingtargets.csv` file (Figure 4.10). Target reductions are described as a percent reduction from current loading. Users can specify whether targets are for Total Nitrogen (TN) or Total Phosphorus (TP). If targets are not within the river network, they can be flagged with a capital X in `OutofNetworkFlag_X`. If the target is the terminal end of a watershed, it can be flagged with a capital X in `TermFlag_X`. `COMID` refers to the NHDPlus V2 catchment that contains the loading target.

Waterbody_Name	ComID	Watershed_HUC	Percent_Reduction	TN_or_TP	OutofNetworkFlag_X	TermFlag_X
CT River at MA border	9332552	10801	0.1	TN		X
Back Lake (NHLAK801010203-01-01)	4592401	10801010203	0.12	TP		
Forest Lake (NHLAK802010401-01-01)	4594723	10801030101	0.02	TP		

Figure 4.10: A screenshot of the loading targets UserSpecs file with the specifications for the Upper Connecticut Rive case study.

4.2.2 Preprocess Data Inputs

The data preprocessing section of the code adjusts and formats the input data for AMPL model file development. Input data include baseline nutrient loads, land use information, watershed characteristics, BMP costs, and BMP nutrient removal efficiencies. The provided data cover the geographic extent of the Upper Connecticut River case study only.

4.2.2.1 Baseline Nutrient Loading and Land Use Data

Recent regional SPARROW model outputs from the U.S. Geological Survey provide the basis for RBEROST baseline nutrient loading conditions (Ator, 2019). These models provide catchment-level nutrient loads and land use data. Catchments are specified using common identifiers (COMIDs) based on the National Hydrography Dataset (NHD) Plus Version 2 reach network (McKay et al., 2012). RBEROST also includes 2011 cropland and imperviousness land use data from the National Land Cover Database (Yang et al., 2018) provided at the NHDPlus V2 catchment level via StreamCat (Hill et al., 2015). For modeling that includes states outside of the region presented in the case study, StreamCat files for additional states can be added to the Preprocessing/Inputs folder and renamed according to the same convention as the provided files.

Data modifications entail classifying incremental baseline SPARROW loadings as point source, urban, agricultural, or "other" loads. Incremental loads in SPARROW are stream loads that originate from the NHDPlus V2 catchment. This is differentiated from total loads in SPARROW which is the sum of incremental load for each catchment and delivered load from upstream. Other loads include loadings from sources that cannot be classified as point, urban, and agricultural (e.g., septic systems or atmospheric deposition). The model code adjusts baseline point source loading to account for changes in wastewater nutrient effluent and atmospheric nitrogen deposition since 2012. This is done within RBEROST so that the user can potentially include additional locations for baseline loadings adjustments where changes are known to have occurred. This would be done by editing the input files `WWTP_BaselineRemoval_Finergrain.csv`, `WWTP_COMIDs_BslnRemoval.csv`, and `NdepChange_2012_2019.csv` to include information for additional COMIDs. Details of these files are available in Table 6.1.

The optimization depends on agricultural and urban land area available for BMP implementation. Urban acreage data are derived from the SPARROW regional models, while agricultural acreage is calculated as the product of incremental acreage available per catchment (from SPARROW; Ator (2019)) and the percent of area defined as cropland per catchment (from National Land Cover Data/StreamCat; Hill et al. (2015)).

Urban BMP costs depend on the volume of water being treated. RBEROST relies on a water quality volume equation from the Vermont Stormwater Management Manual (Vermont Agency of Natural Resources, 2017). Water quality volume depends on a volumetric runoff coefficient, equal to $0.05 + 0.009 * Imperviousness$. The National Land Cover Dataset (NLCD) provides imperviousness data at the catchment level via StreamCat (Hill et al., 2015).

Riparian loading depends on the amount of impervious surface cover within the river corridor, as per Houle et al. (2019). These estimates must be derived prior to modeling with RBEROST. They are provided in the `RiparianLoadings.csv` file. Riparian loading is a subset of urban, ag, point, or other loads and does not describe an additional source. Users who wish to use different methodologies to estimate riparian loads (e.g., including agricultural loads) may do so and overwrite the `RiparianLoadings.csv` file with a new file with the same naming conventions and format.

4.2.2.2 Watershed and Target Specifications

The `01_UserSpecs_loadingtargets.csv` file describes the locations of loading targets that will be included in RBEROST. The COMID identifiers are used to identify upstream contributing reaches to each target. These contributing reaches are included in the input baseline loading and land use datasets used in calculating inputs to each target waterbody.

The SPARROW regional models specify a delivery fraction (DEL_FRAC) that reflects the fraction of the incremental nutrient loads that are delivered to a flowline's terminal reach (in most cases, the ocean). RBEROST revises this delivery fraction to recognize each loading target as the specified terminal reach. For instance, if the specified target COMID has a delivery fraction of 92% in the SPARROW regional model dataset, indicating that 92% of the incremental loads associated with the target reach are delivered to the terminal reach, the revised delivery fraction at this location is 100%, indicating that 100% of the incremental loads associated with the target are delivered to the corresponding target reach.

The delivery fractions for all reaches upstream of the target are revised based on the ratio of the revised target delivery fraction to the original terminal reach delivery fraction. For instance, if the SPARROW regional model data indicates that an upstream reach delivers 83% of its incremental load to the terminal reach and that the target reach delivers 92% of its incremental load to the terminal reach, then the revised delivery fraction for the upstream reach to the target reach is (100%/92%)*(83%) = 90%. Delivery fractions typically differ between TN and TP, and these calculations are performed separately for each nutrient target. RBEROST allows nested targets as well, in which case a stream reach that contributes to more than one loading target will be assigned more than one DEL_FRAC value.

4.2.2.3 BMP Costs and Efficiencies

BMP costs specified in the `01_UserSpecs_BMPs.csv` file are adjusted to per-acre units for agricultural BMPs, and per-square foot units for riparian buffer BMPs. Because point source BMPs are location-specific, the model code merges the wastewater treatment plant (WWTP) costs with each plants' reach location (COMID) prior to developing the AMPL model files. Default urban BMP capital costs represent new development costs. RBEROST will adjust costs for expected amounts of retrofits based on retrofit costs specified in Table 7-2 of the WMOST v3 Theoretical Documentation (Detenbeck, ten Brink, et al., 2018). This adjustment is a weighted average of the new development cost (assigned to open and low development), retrofit costs (assigned to medium development) and difficult retrofit costs (assigned to high development). The development data originate from the NLCD (Yang et al., 2018) and are summarized by catchment (Hill et al., 2015). As in the WMOST model, default urban BMP operations and maintenance costs are assumed to be 5% of capital costs, but this can be modified by the user if more accurate estimates are available.

RBEROST assumes a default planning horizon of 15 years and an interest rate for capital costs of 3%. The user can adjust the horizon and interest_rate variables within the `RunRBEROST.Rmd` file.

While some BMP efficiencies are BMP-specific, others are also location-specific. For agricultural BMPs with nutrient removal efficiencies that vary based on HUC12, the model code identifies the individual reaches that fall within each HUC12 and assigns the HUC12-specific efficiency to each reach. There are two available agricultural BMP cost datasets that differ based on how they compare the export of nutrients. “No Practice” compares the export of nutrients with a given BMP versus the export of nutrients under conventional practices. “Baseline” compares the export of nutrients with only one given BMP versus the export of nutrients occuring under existing conservation practices. The “Baseline” approach assumes existing conservation practices are removed and only the one BMP is implemented. Point source BMP efficiencies are specific to each WWTP upgrade and urban BMP efficiencies are calculated based on the user-specified runoff depth. As some urban BMP efficiencies are dependent on the infiltration rate of the soil underlying developed land, the urban BMP efficiency curves are selected based on the average infiltration rate. This information is provided in `NHD+infiltrationrates.csv` and is derived from gNATSGO soil layers (Soil Survey Staff, 2020b, 2020a).

Riparian buffer BMP efficiencies similarly depend on average infiltration rate, as well as the slope of the riparian buffer. The efficiency curves (as a function of buffer widths) of different riparian buffer BMPs for different nutrients at five representative buffer widths are provided in `RiparianEfficiences.csv`. These efficiency curves have been assigned based on infiltration rate and slope of different buffer widths in each COMID. RBEROST will use the efficiency curve based on the closest match to the user-selected buffer width.

4.2.3 Write AMPL Model Files

Running the optimization model through the NEOS server depends on the development of three AMPL model files: the command file (.amp), the data file (.dat) and the model file (.mod). The command file contains instructions for the optimization server to solve the mathematical problem and defines what results should be displayed afterwards. The data file contains all of the data used in the optimization, including values for each parameters across their geographic range. The model file defines the mathematical problem to be optimized and contains a list of constraints placed on the model during optimization. After running the preprocessing code, the AMPL model files will by default be written to the `./Preprocessing/Outputs` folder, unless otherwise specified in the `RunRBEROST.Rmd` file (Figure 4.11). If the user chooses to run the model with uncertainty, these command, data and model files will be designated by “`_uncertainty`” in their name.

Several messages appear while running RBEROST preprocessor (Figure 4.12, as highlighted in the red box). The first indicates when RBEROST has begun and when it has finished determining the list of COMIDs upstream from the loading targets. This step usually takes a considerable amount of time, so the messages are intended to assure the user that the code is running. The next few indicate whether there are more COMIDs in the StreamCat datasets or in the SPARROW dataset of the contributing reaches to each target. The number of messages will match the number of targets. SPARROW COMIDs that are not in StreamCat, or StreamCat COMIDs that are not in SPARROW contribute to the other_loads parameters. The next messages will indicate when RBEROST has completed writing AMPL scripts. When running RBEROST with uncertainty analysis, AMPL scripts without uncertainty will be written first as part of the preprocessing. These AMPL files without uncertainty may be ignored or saved for further use.

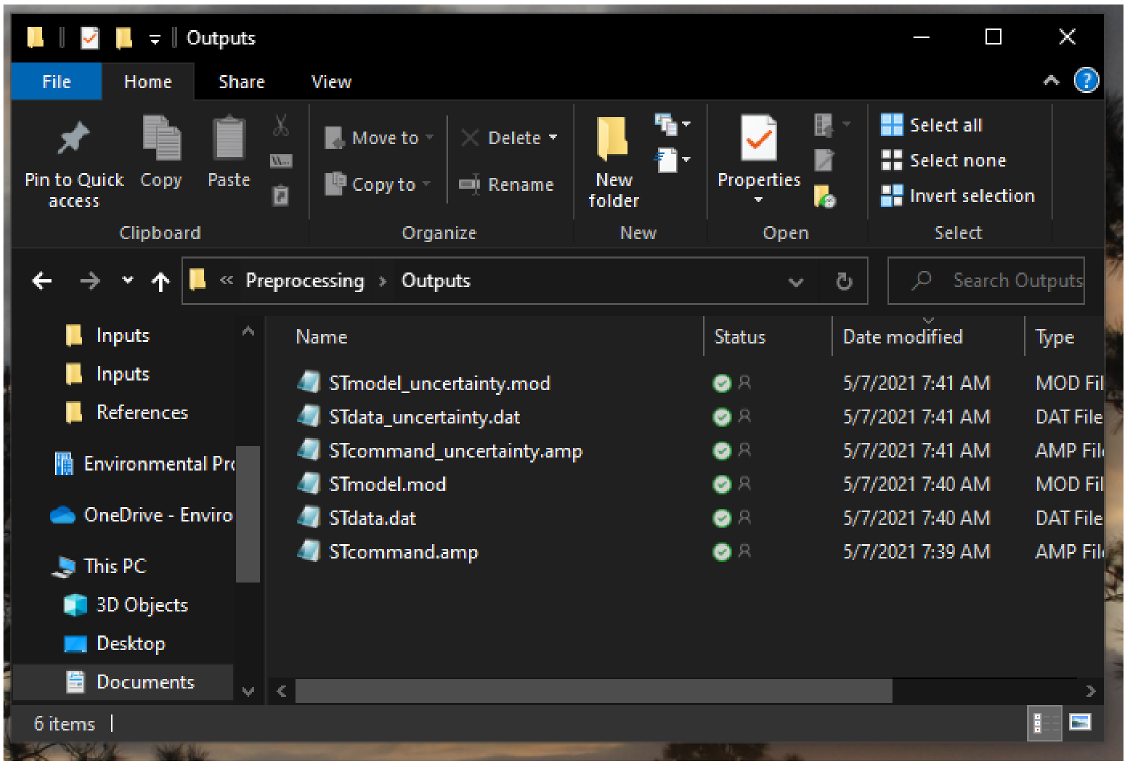


Figure 4.11: Outputs of running RBEROST preprocessor with uncertainty.

4.2.3.1 AMPL Command File

The command file specifies display characteristics of the NEOS optimization results. If the model is being run with uncertainty, the command file specifies the number of scenarios to run and alters the loading limits accordingly for each successive solve. The command file also calculates 200 ‘bootstraps,’ of cost and total loading based on the solved model’s suggested suite of BMPs. These bootstraps are 200 calculations of cost and total loading based on random sampling of values for all the necessary parameters in RBEROST. For most parameters, this is a normal distribution, though several parameters are described as uniform distributions. The mean value used for the resampling is the value used by RBEROST to solve the optimization problem, and the standard errors are additional parameters written to the `STdata_undertainty.dat` file.

4.2.3.2 AMPL Data File

The data file specifies the catchment- and BMP-specific data required for the NEOS optimization run, including:

- Baseline loadings, efficiency, and cost data formatted by the R code;
- Total reduced loading values (loads_lim parameters) calculated based on the specified load reduction percentage;
- Total loading value for loads specified as “other loads” (other_loads parameters) that are not available for BMP loading reduction;
- The fraction of agricultural costs (agcost_frac) that reflect base payment versus actual agricultural BMP costs;
- The adjustment factor of urban costs (urban_cost_adjustment_coef); and
- Conversion factors for urban BMP treated water volume (acfttoft3) and precipitation (pcp).

These parameters are described in greater detail in Table 3.2. Additional parameters written when modeling with uncertainty are described in detail in Table 3.3.

Depending on the geographic extent of optimization, the data file with uncertainty may become larger than the allowed upload size for NEOS. If this occurs, it is recommended to proceed using the AMPL scripts that do not contain the uncertainty analysis.

4.2.3.3 AMPL Model File

The model file specifies the following:

- COMIDs, BMPs, cost types, and load types that are included in the optimization;
- The bases on which parameters vary (e.g., agricultural efficiencies vary by both reach and BMP);
- Constraints on parameters and variables (e.g., whether a variable represents binary conditions or a fraction less than or equal to 1);
- The total load reduction functions; and
- The cost minimization objective function.

The model file with uncertainty includes the same specifications for parameters that describe standard errors.

4.3 NEOS Server

RBEROST uses the CPLEX Optimizer available through NEOS to solve the optimization. CPLEX will allow XML calls from RBEROST to submit AMPL files to the server, however at the time of current release, only a manual interaction option with the server is available. Users will click on the following link and upload the AMPL model, data, and commands files to the Web Submission Form: https://neos-server.org/neos/solvers/lp:CPLEX/AMPL.html (Figure 4.13). Before submitting, the user will specify the e-mail address to receive an update when the model run is complete. Users are also highly encouraged to include a descriptive note of the model in the comments section (Figure 4.14). Please refer to the WMOST User Guide (Detenbeck, Piscopo, et al., 2018) for additional information on how to run optimization models on the NEOS server. All AMPL files must be of the same type - either with uncertainty, or without uncertainty. Mixing scripts of the two types will produce errors and the model will not solve. If you are running RBEROST with uncertainty analysis, you only need to submit the AMPL files with "_uncertainty". The other files that were created, which describe the same problem without the uncertainty analysis, may be ignored or saved for reference.

Generally, NEOS can solve models without uncertainty in less than five minutes, while models with uncertainty may take nearer to 20 minutes. Exact times may vary, perhaps greatly, depending on the complexity of the specified problem and the current demands on NEOS and on the CPLEX optimizer. Once the optimization has solved, users will receive an email from the server (Figure 4.15) with a link, a job number, and a password. Users can retrieve their results from the server with this information (Figure 4.16). Users save the NEOS server optimization results by selecting all text within the results window (CTRL+A), pasting it into a text file (via text editor such as Notepad), and saving it as a text file (.txt) to a location that makes sense for them (Figure 4.17). Example output files are provided in `.\Postprocessor\Input`. If NEOS was unable to optimize the model, the returned result may be very short, simply stating that optimization failed. This result file may still be submitted to the RBEROST Postprocessor.

```
|--------------------------------------------------|
|==================================================|
[1] "RBEROST is now determining which reaches are included in the watersheds
of the identified loading targets."
[1] "RBEROST has now determined which reaches are included in each
watershed."
Note: For the 1st TN target, there are fewer reaches in the provided
StreamCat datasets than are included in SPARROW.
Only the reaches that are included in both datasets will be available
for BMP optimization. Loads the remaining reaches will be included in
the 'other_loads' parameter.
Note: For the 1st TP target, there are the same number of reaches in SPARROW
and Streamcat
subsetted datasets.
All reaches available for BMP optimization
Note: For the 2nd TP target, there are the same number of reaches in SPARROW
and Streamcat
subsetted datasets.
All reaches available for BMP optimization
[1] "Urban base costs are assumed to be the same across states. RBEROST is
using values from VT."
[1] "RBEROST has finished writing AMPL scripts without uncertainty at 2021-
11-12 12:36:45"
[1] "RBEROST is now creating AMPL files with uncertainty analysis at 2021-11-
12 12:36:46"
[1] "RBEROST has finished writing AMPL scripts with uncertainty at 2021-11-12
12:38:36"
```

Figure 4.12: Messages displayed by RBEROST preprocessor

Figure 4.13: The CPLEX Optimizer hosted by NEOS.

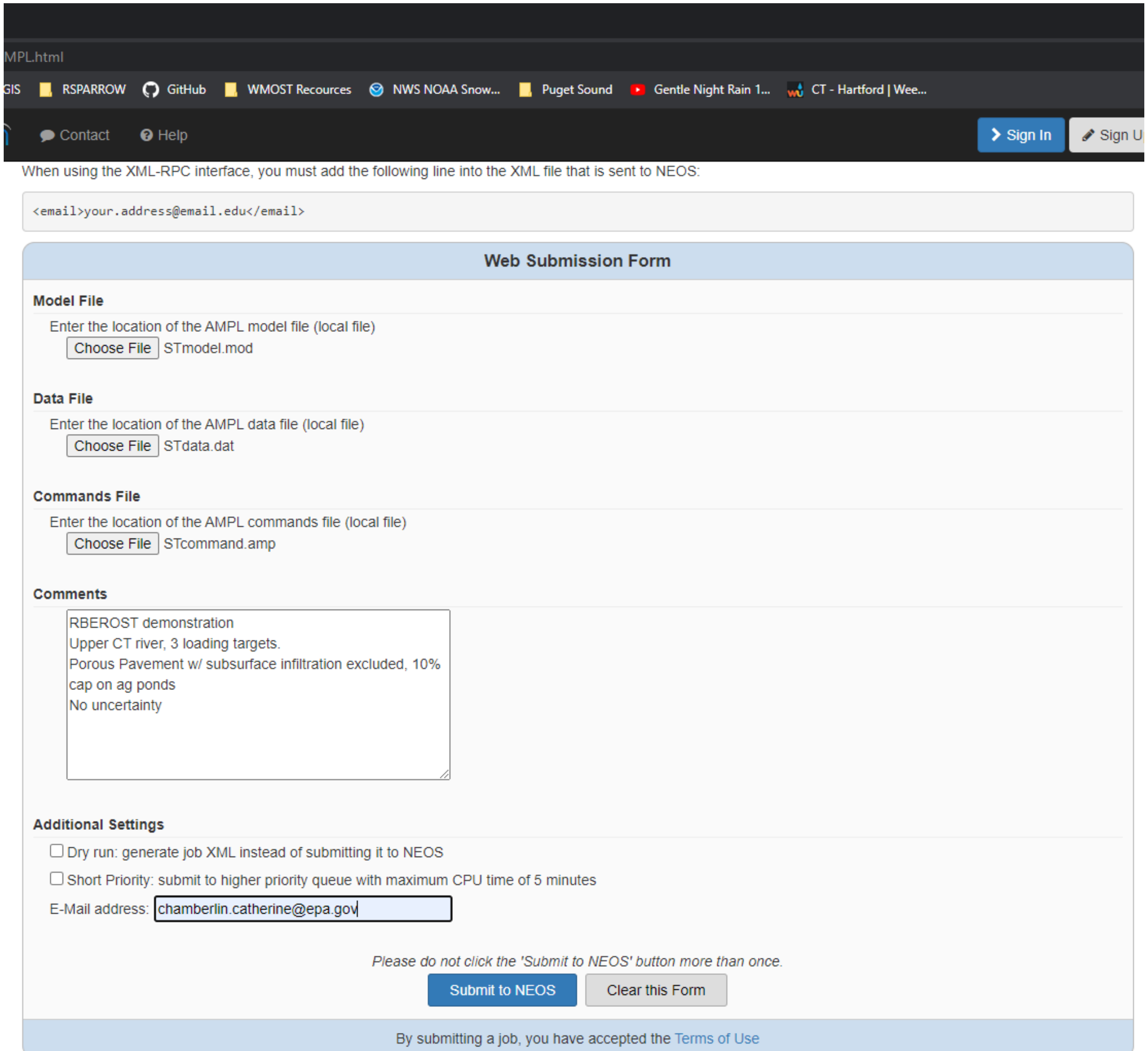


Figure 4.14: User inputs to CPLEX to run RBEROST.

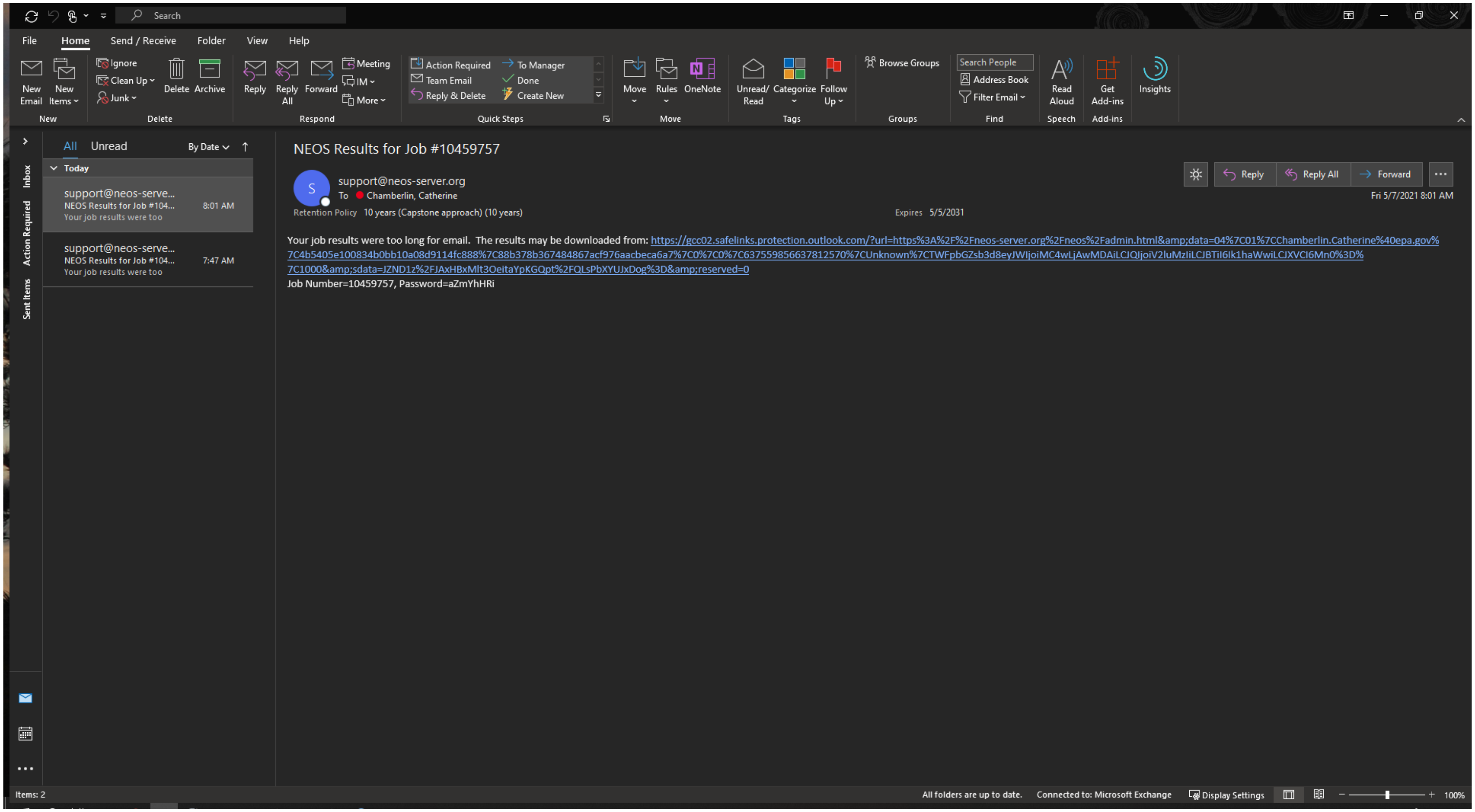


Figure 4.15: Example email from NEOS.

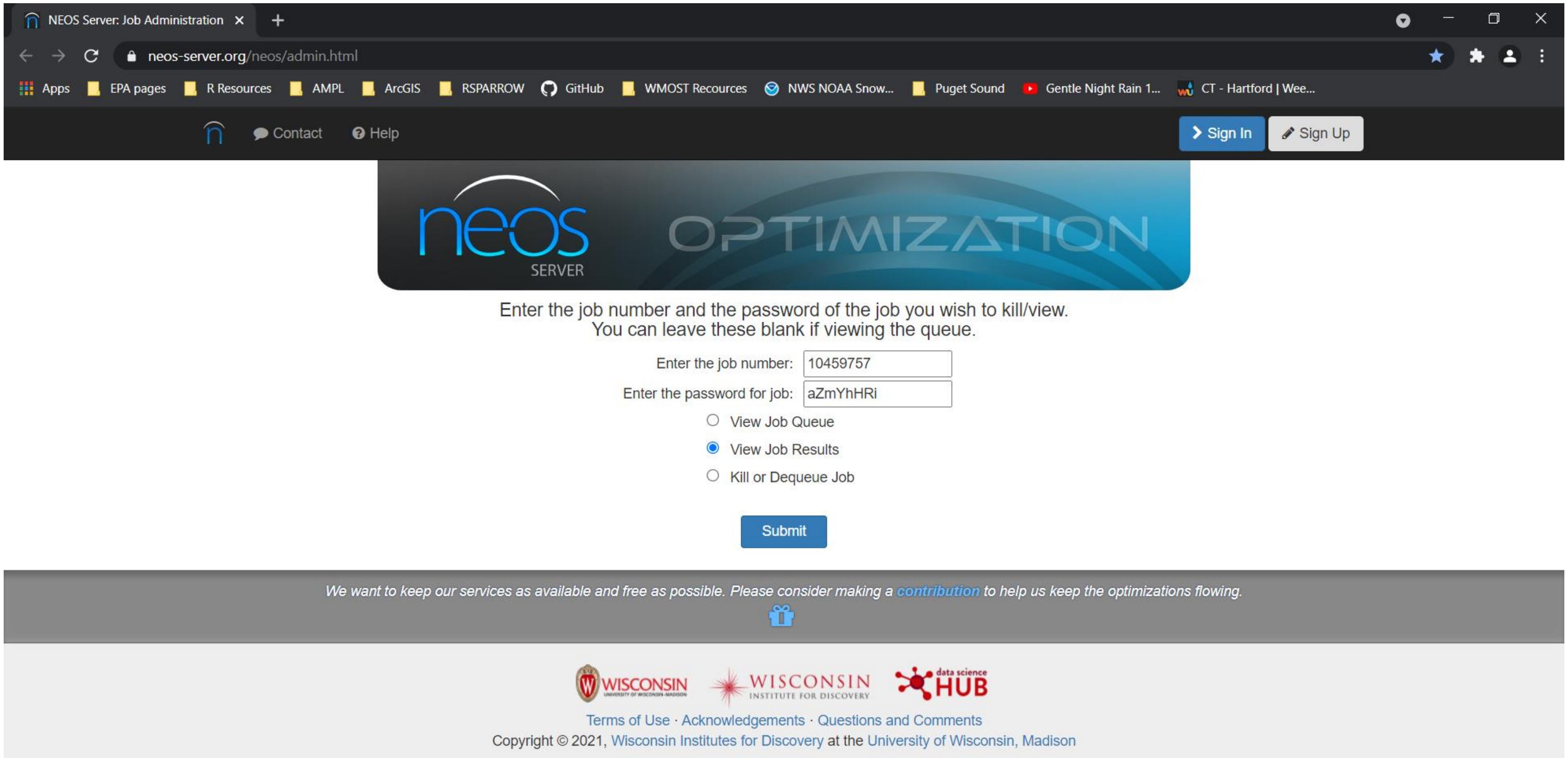


Figure 4.16: Retrieving results through the NEOS Job Administration page.

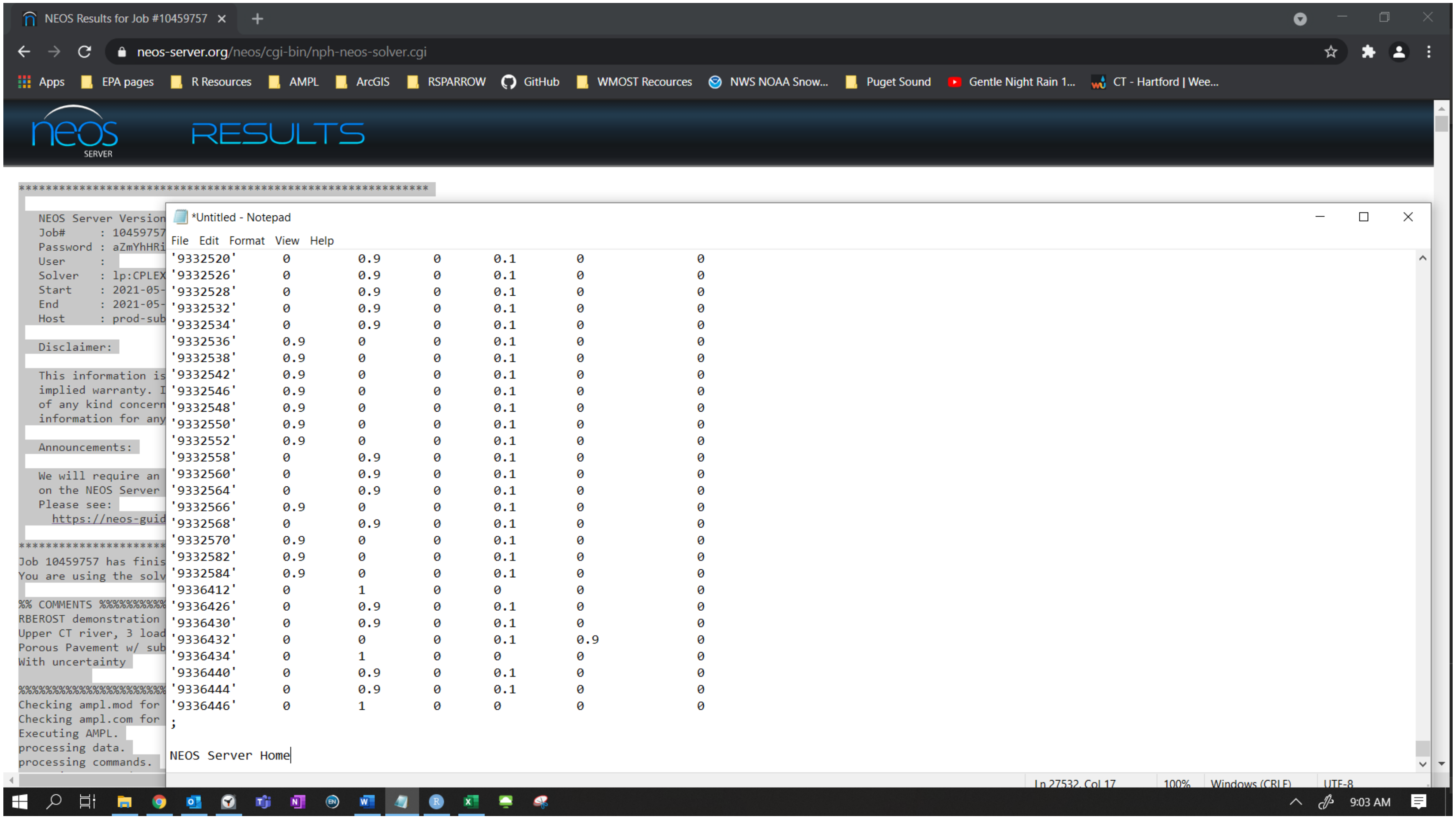


Figure 4.17: Saving NEOS results as a text file.

4.4 Postprocessing

The postprocessing R code combines optimization results from the NEOS server with input data to develop result summaries. The code uses R-Shiny, an R package that produces a summary of RBEROST either within the RStudio interface or in the default web browser.

To run the postprocessing code, users can click the green triangle in the `Run Postprocessor` section of the `RunRBEROST.Rmd` file (Figure 4.18). More than one result file can be viewed per session, however the Shiny App will not automatically close if the viewing window is closed. To stop the Shiny App after closing the browser window, click the red square in the `Run Postprocessor` section, and a green triangle will appear that will allow the user to launch the application again.

If RBEROST postprocessor is running within a browser, you may print the results screen to a pdf file for your records.

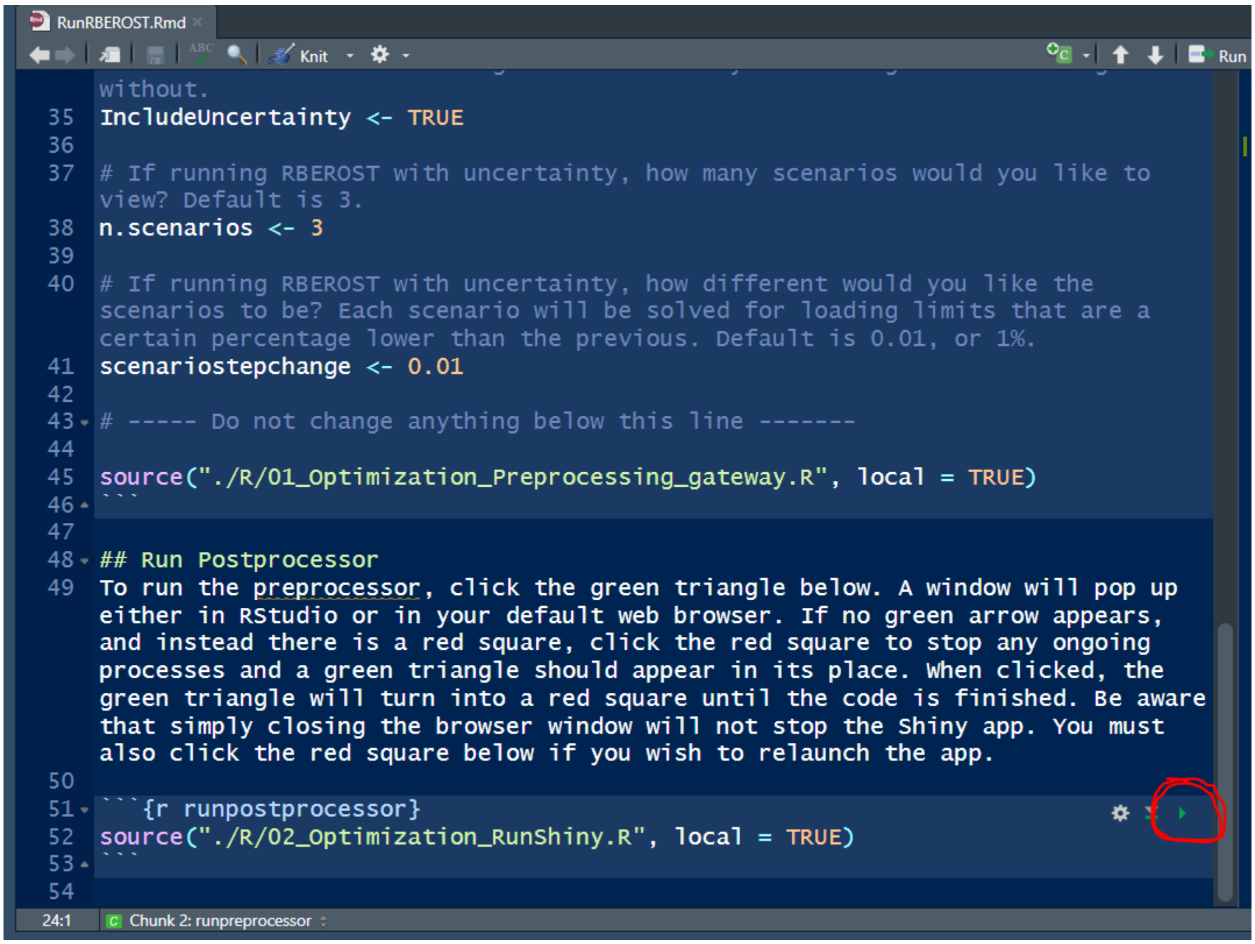


Figure 4.18: Run the RBEROST postprocessor

4.4.1 Necessary Files

The postprocessor requires upload of at least five files and will display warning messages if files are not uploaded to the correct options (Figure 4.19). The upload of StreamCat cropland files accepts all the state cropland files used in the optimization, which may be more than one file. Printing to pdf is the current recommended way of capturing metadata regarding which files are provided to the postprocessor. Required files include a result file from a NEOS optimization run, the list of WWTPs by their COMID locations, the SPARROW input file, and the StreamCat files for all states included in the optimization.

4.4.2 Preview Files

The preview step is not necessary but may be helpful for users to ensure they have selected and uploaded the intended files. To preview files, click the `Preview Uploads` button and navigate to the `File Preview` tab (Figure 4.20). At any point users can upload new files and preview them by selecting `Preview Uploads` again.

4.4.3 Display Results

To view results, select `View NEOS Results` and navigate to the `View Results` tab (Figure 4.21). Several panels will appear including an option to select the scenario the user wishes to view, the comments entered by the user to NEOS, a summary of the total cost calculated by RBEROST, and a summary of BMP implementations by category. For RBEROST runs without uncertainty, only one scenario is available. For RBEROST runs with uncertainty, more than one scenario may be available, and only one will be displayed at a time. Additional panels displayed for models with uncertainty include a panel showing the expected probability density distribution of costs (Figure 4.22) and a panel showing the total annual expected loading to the user-specified loading targets (Figure 4.23). If the model was unable to solve, a message will appear with the following explanation:

> "RBEROST may fail to optimize models for a variety of reasons. Often, this is a result of loading targets that are too restrictive, or a result of not including enough BMPs for the model to use. It may be informative to run RBEROST again with less restrictive loading targets and/or including more BMPs in the optimization. Please check your user inputs and refer to the Model Sensitivity section of the documentation for more information."

In addition, in the event of failure due to AMPL syntax, errors from CPLEX will be displayed. These errors can occasionally occur for a variety of reasons, including if the users have uploaded the AMPL files in the wrong order during job submission.

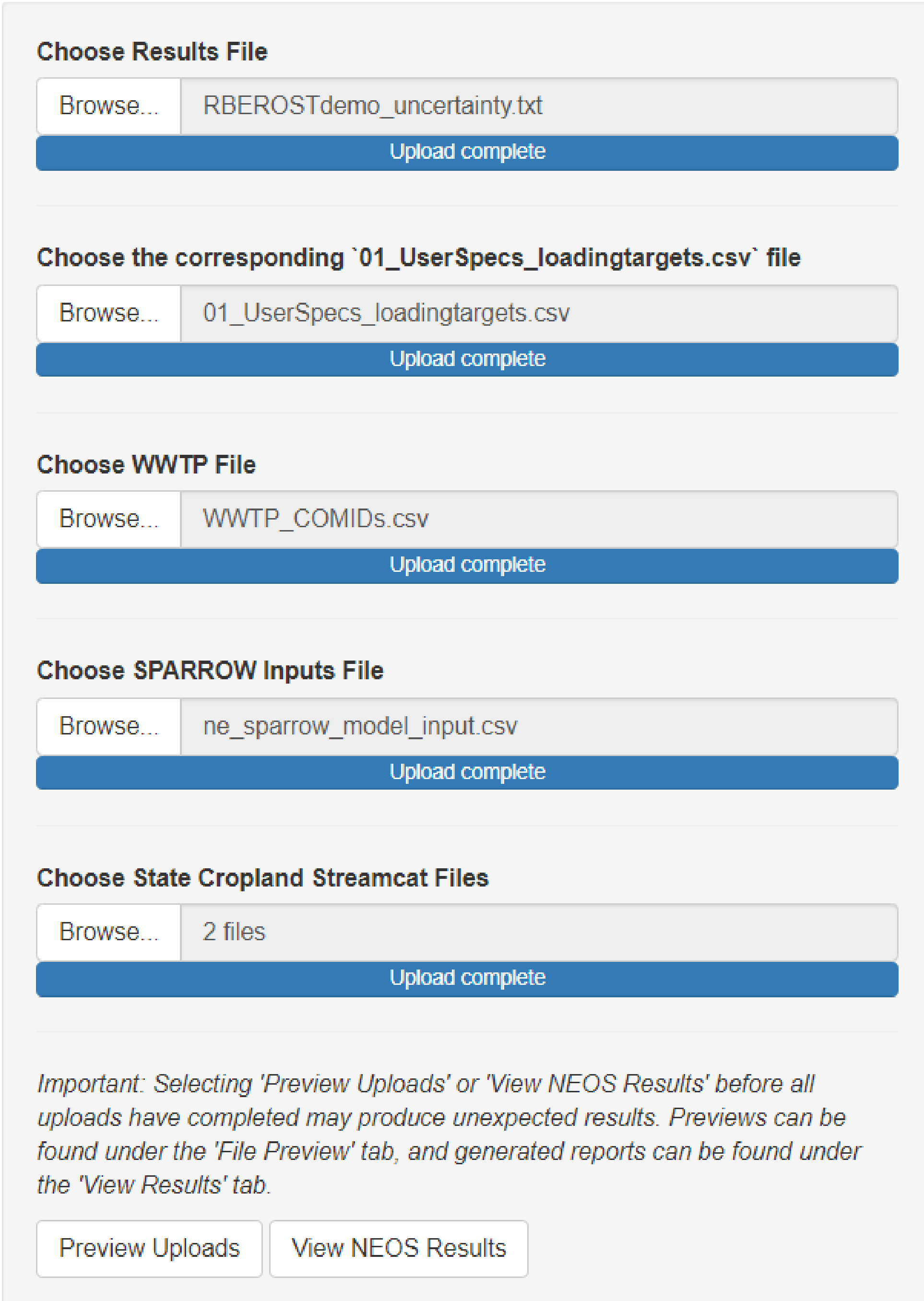


Figure 4.19: Files necessary to run RBEROST post-processor

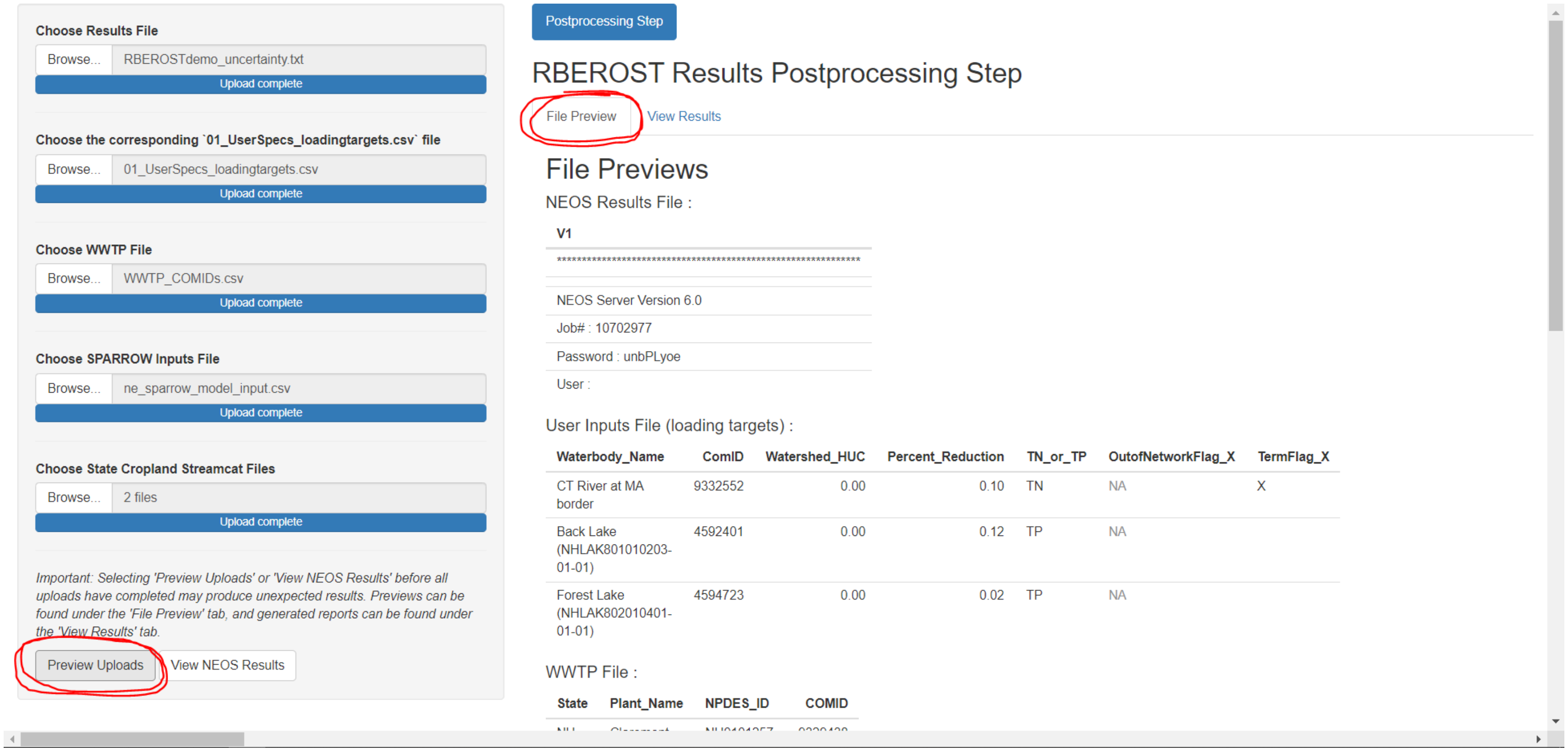


Figure 4.20: Previewing files for RBEROST postprocessor.

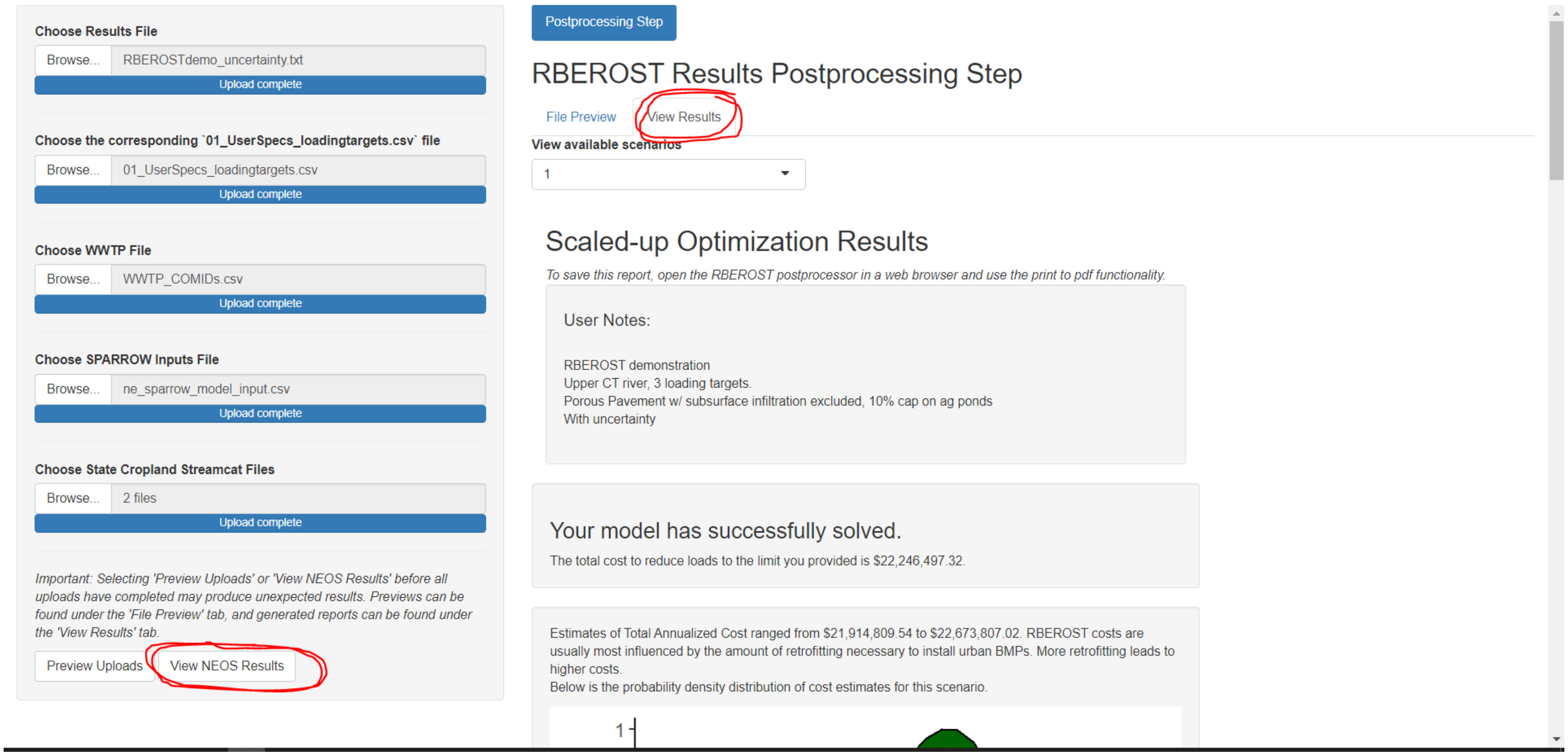


Figure 4.21: Viewing RBEROST results.

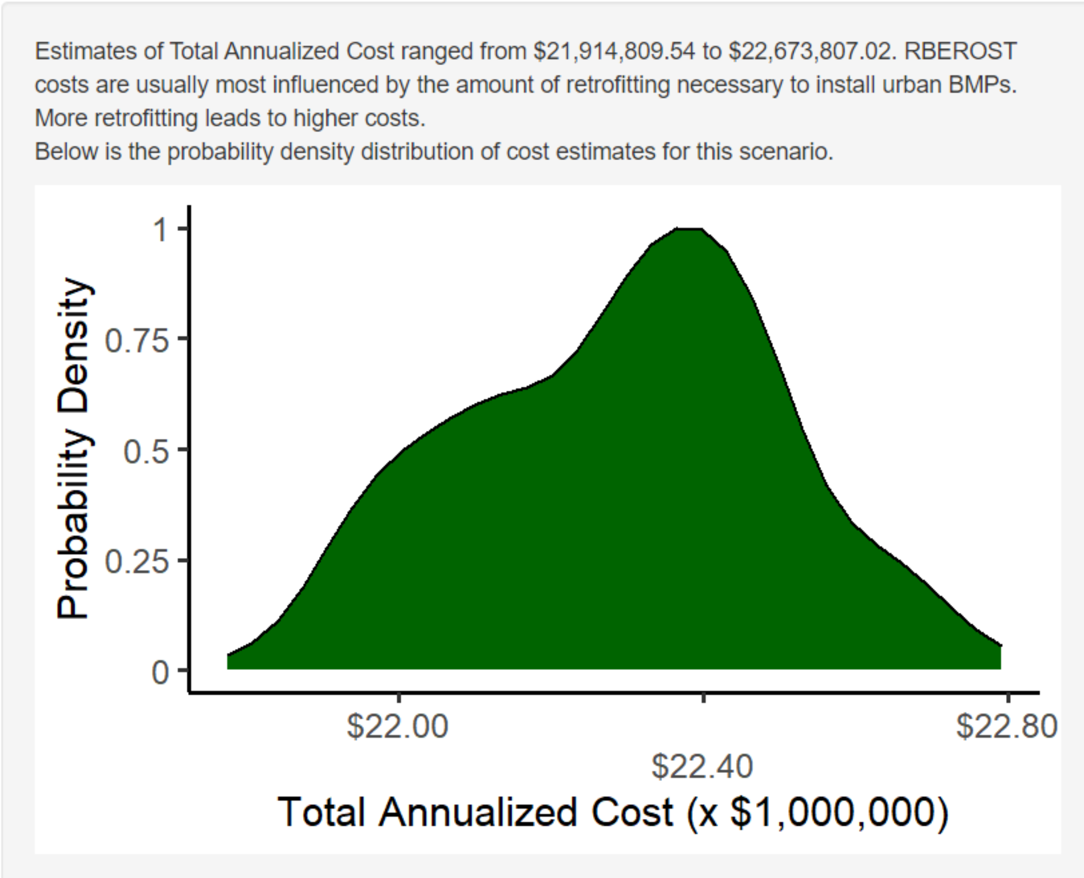


Figure 4.22: Viewing results of uncertainty analysis in cost. The estimated range reflects the 95 percent confidence interval of bootstrapped cost estimates. The probability density distribution of the bootstrapped cost estimates is shown in the figure.

This suggested plan has an estimated minimum of 20% likelihood of meeting each of the specified loading targets. Likelihoods of meeting each specified loading target individually are listed in the table below. 'Likelihood' refers to the percentage of the probability density of estimated total annual load that falls below the user specified loading target.

Loading Target	Your Specified Target Load	95% CI of Estimated Annual Total Loading	Likelihood of Meeting Your Loading Target
CT River at MA border	10% reduction in TN; or 6,218,700.0 kg N/yr	6,156,828.3 - 6,318,925.9 kg N/yr	42%
Back Lake (NHLAK801010203-01-01)	12% reduction in TP; or 3.8 kg P/yr	1.1 - 16.1 kg P/yr	20.5%
Forest Lake (NHLAK802010401-01-01)	2% reduction in TP; or 2.7 kg P/yr	0.7 - 7.4 kg P/yr	41.5%

Below are the probability density distributions of total annual load estimates at each of the loading targets specified by the user. The vertical dotted grey lines denote the specified target loads at each loading target. Probability densities to the left of the grey line are load estimates that meet the loading target, and probability densities to the right of the grey line are load estimates that exceed the loading target.

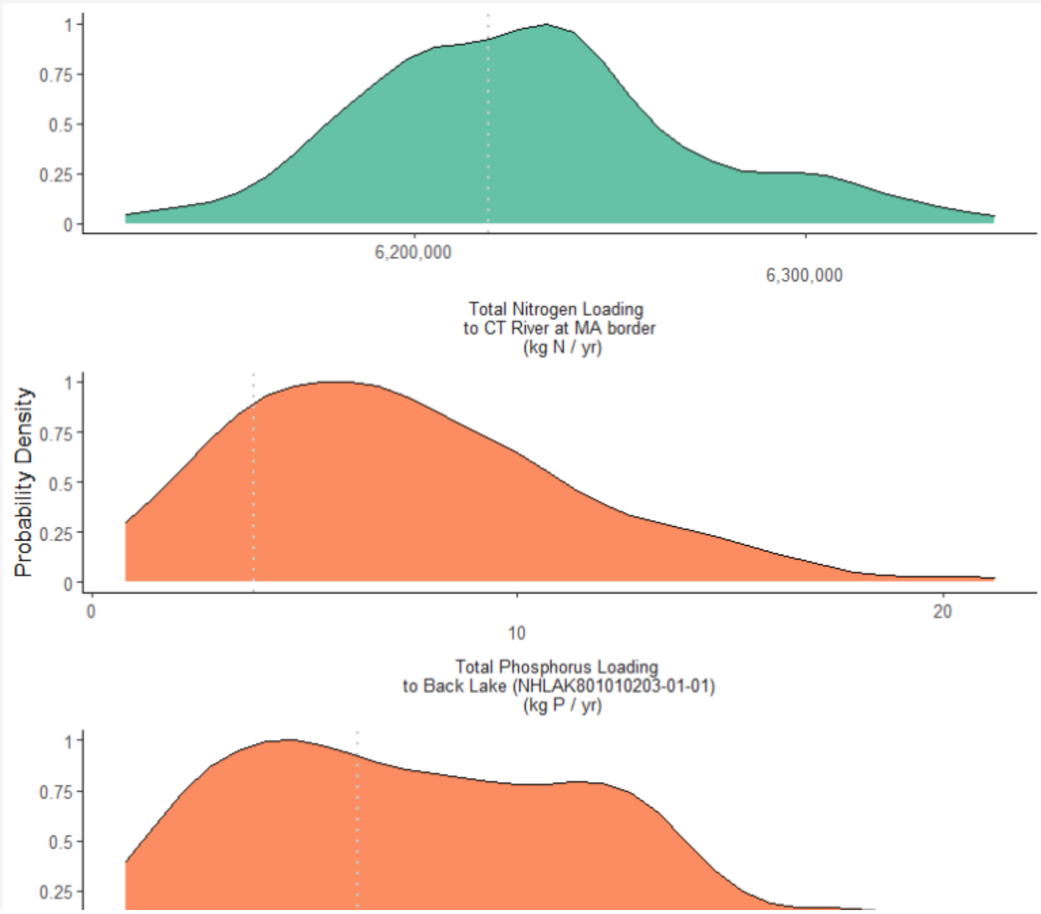
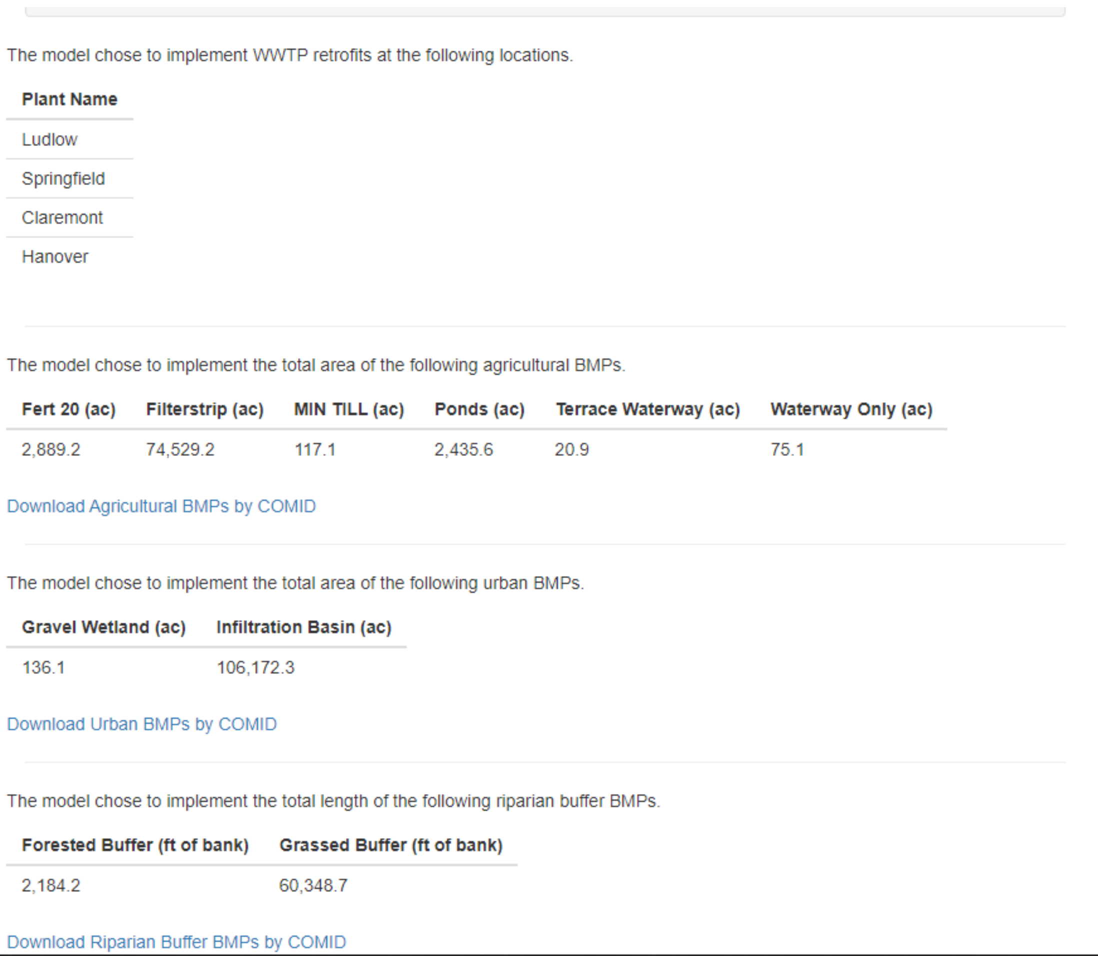


Figure 4.23: Viewing results of uncertainty analysis in total annual loading. The leading sentence describes the lowest likelihood of successfully meeting the loading targets described out of all the targets included. The table shows all of the included loading targets, their specified reductions, the 95 percent confidence intervals of bootstrapped estimates of total annual loading, and the likelihood of meeting loading targets. The likelihood of meeting loading targets refers to the proportion of the probability density distribution that falls below the specified target reduction. Below the table are plots of the probability density distributions of estimated annual total load at each specified loading target. Vertical grey dotted lines indicate the target reduction specified. Probability density to the left of this line meets the target, and probability density to the right of this line exceeds the target.

4.4.4 Download Detailed Results

COMID-specific BMP implementation decisions can be downloaded from the Shiny App for any single scenario (Figure 4.24). Point-source decisions are not available for download and are displayed within the Shiny interface. Additional contextual information is included in the downloadable csv files including the HUC12 code and state for each COMID, the size of the catchment, and the total area of agricultural land for ag BMPs, or urban land for urban BMPs, or the total river reach length for riparian buffer BMPs.



The model chose to implement WWTP retrofits at the following locations.

Plant Name
Ludlow
Springfield
Claremont
Hanover

The model chose to implement the total area of the following agricultural BMPs.

Fert 20 (ac)	Filterstrip (ac)	MIN TILL (ac)	Ponds (ac)	Terrace Waterway (ac)	Waterway Only (ac)
2,889.2	74,529.2	117.1	2,435.6	20.9	75.1

Download Agricultural BMPs by COMID

The model chose to implement the total area of the following urban BMPs.

Gravel Wetland (ac)	Infiltration Basin (ac)
136.1	106,172.3

Download Urban BMPs by COMID

The model chose to implement the total length of the following riparian buffer BMPs.

Forested Buffer (ft of bank)	Grassed Buffer (ft of bank)
2,184.2	60,348.7

Download Riparian Buffer BMPs by COMID

Figure 4.24: Downloading RBEROST results.

5 Model Sensitivity

RBEROST results are sensitive to provided data and user inputs. Model results are most sensitive to the user-defined loading TN target in the terminal reach of the watershed, the estimates of "other" TN loads to that terminal reach, and the TN removal efficiencies of several BMPs that are implemented extensively (specifically Filterstrip, Ponds, and Infiltration Basins). "Most sensitive" means that a 1% change or less in *both* the positive and negative direction cause either total costs or spending on individual BMPs to change by more than 10%. Other parameters may have been more sensitive in one direction than the other, or generally less sensitive in both directions. RBEROST sensitivity to each individual user input is given in Table 5.1, and sensitivity to each individual data parameter provided in the RBEROST preprocessing input csv files is given in Table 5.2. The `Parameter` columns in Tables 5.1 and 5.2 are the names of the parameters as they appear in the AMPL model and data scripts. The `Set member` columns generally describe the specific BMP being defined but may also reflect other groupings, such as capital or operations costs, urban or agricultural area, or the source of nutrient loading baselines.

Tolerances are defined as the relative amount of variation in the estimate of each parameter that produces RBEROST results with costs that are within 10% of the reference cost, and that produces results with relative amounts of spending on each BMP that are within 10% of the amount spent on that BMP in the reference result. The reference result is the RBEROST solution to a problem with all BMPs included and no limits on implementation (sometimes referred to as the 'unconstrained' model. Parameters are considered sensitive if a change of less than 10% produces results that are more than 10% different from the reference. Parameters are considered highly sensitive if a change of less than 1% produces results that are more than 10% different from the reference. The last two columns in Tables 5.1 and 5.2 describe the sensitivity to parameters as they are decreased from their default value and as they are increased from their default value. Parameters with tolerances marked as "not applicable" either have default values of 0, or already have default values that are at the minimum or maximum of their permitted range.

The user specified loading target at the terminal reach of the watershed was the single biggest predictor of whether RBEROST could succeed in solving the optimization problem. Figure 5.1 shows the model's success (1 on the y-axis) or failure (0 on the y-axis) as a function of terminal reach target loads for 1000 iterations of the same scenario. Each iteration differed in the exact estimate of each parameter. The estimate of each parameter was sampled from a distribution that described the likely values that the parameter may take. Most parameters were sampled from the normal distributions used for uncertainty analysis. User specified parameters were mostly drawn from uniform distributions encompassing the allowed range of that parameter. Maximum implementations were sampled between 70%-100% and loading targets were sampled between 2% and 20% reductions. Minimum implementations were not varied, and were kept at 0%, because it is unlikely that managers would wish to impose minimum implementations of multiple BMPs at once. Additionally, minimum implementations impose more constraints on the model than maximum implementations and more frequently prevent the optimizer from finding feasible solutions. The image in Figure 5.1 is from an optimization of the Upper Connecticut River watershed, which drains to the Massachusetts boarder. Occasionally optimization of scenarios fail even when target loading is less restrictive. This is a function of an overly constrained problem and may be alleviated by relaxing maximum restrictions on BMPs, by including more BMPs, or by relaxing intermediate loading targets.

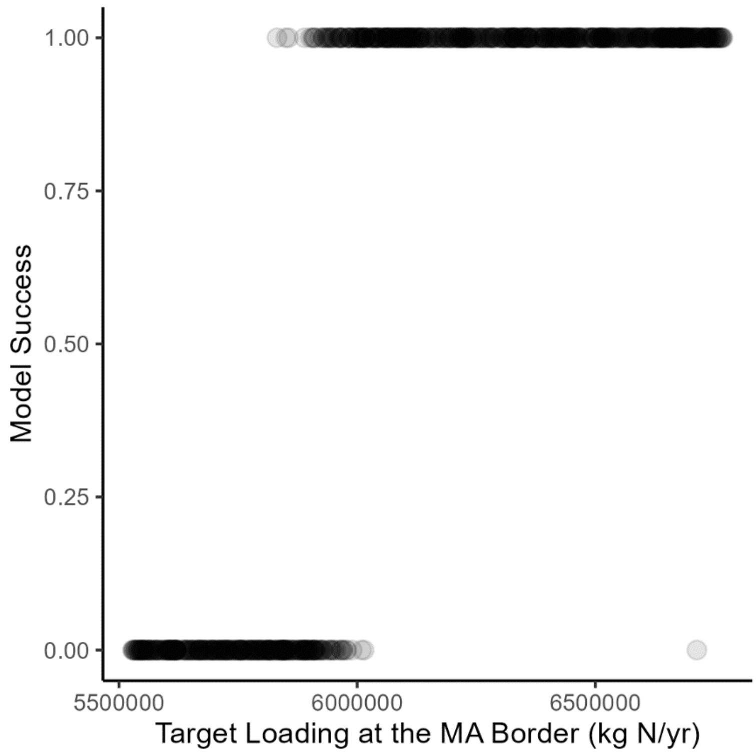


Figure 5.1: Sensitivity of RBEROST success in finding an optimized result as a function of loading target.

RBEROST predictions of annualized total cost increase exponentially as loading targets become more restrictive (Figure 5.2). Figures 5.3, 5.4, 5.5, and 5.6 show the annualized spending on each specific BMP. BMPs are differentiated by color. The clouds of points represent 635 of the 1000 scenarios described above which were optimized with feasible solutions. The solid lines show the LOESS smoothed curve of the data. RBEROST consistently implements the same WWTP upgrades (Figure 5.3) and reliably chooses the same agricultural BMPs dependent on target loading (Figure 5.4). RBEROST less clearly distinguishes between urban BMPs (Figure 5.5) and riparian zone BMPs (Figure 5.6). This means that the choice of which urban or riparian zone BMPs to implement depends on the exact estimates of efficiency and cost. To test if this affects a certain model scenario, users may wish to run RBEROST several times changing the costs in `01_UserSpecs_BMPs.csv` slightly for each run. If RBEROST subsequently chooses different urban BMPs, additional modeling for urban BMPs may be advantageous. Generally, users will benefit by having the most accurate estimates of cost, efficiency, suitability of sites for each BMP, and the expected difficulty of retrofitting existing urban infrastructure with urban BMPs.

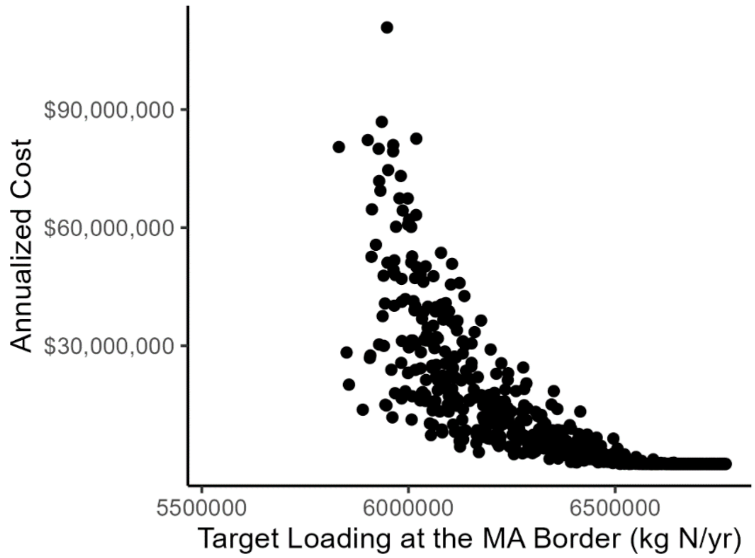


Figure 5.2: Predicted annual total cost as a function of loading target.

Figure 5.3: Annualized spending on point source BMPs as a function of loading target

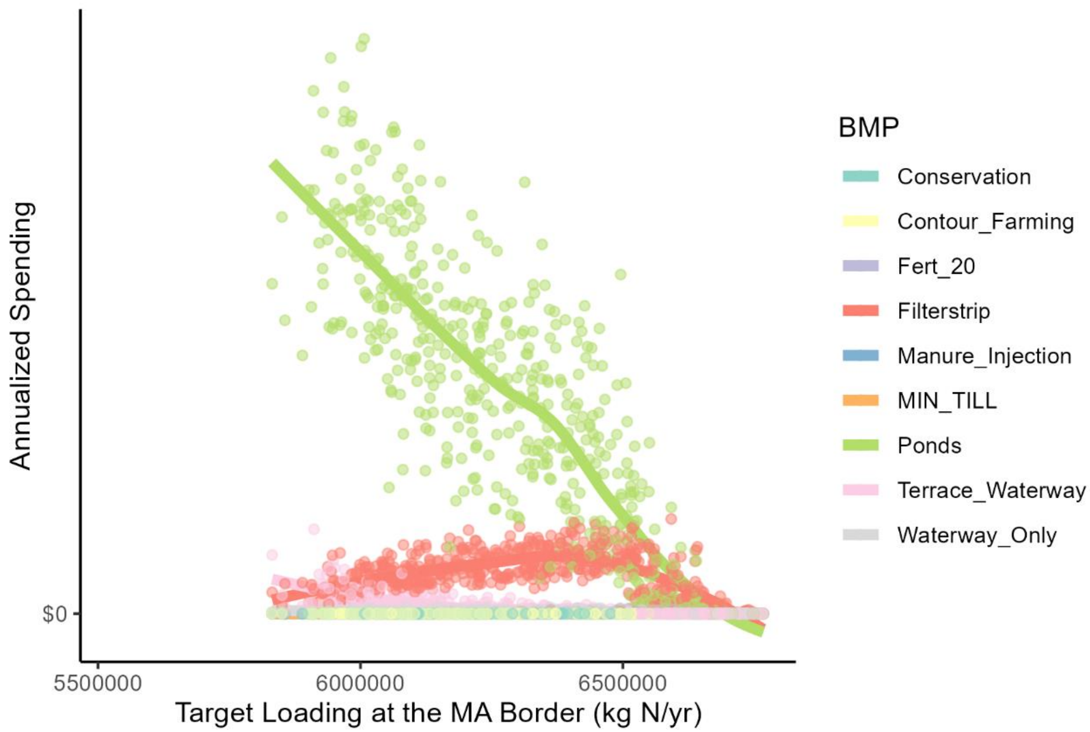


Figure 5.4: Annualized spending on agricultural BMPs as a function of loading target

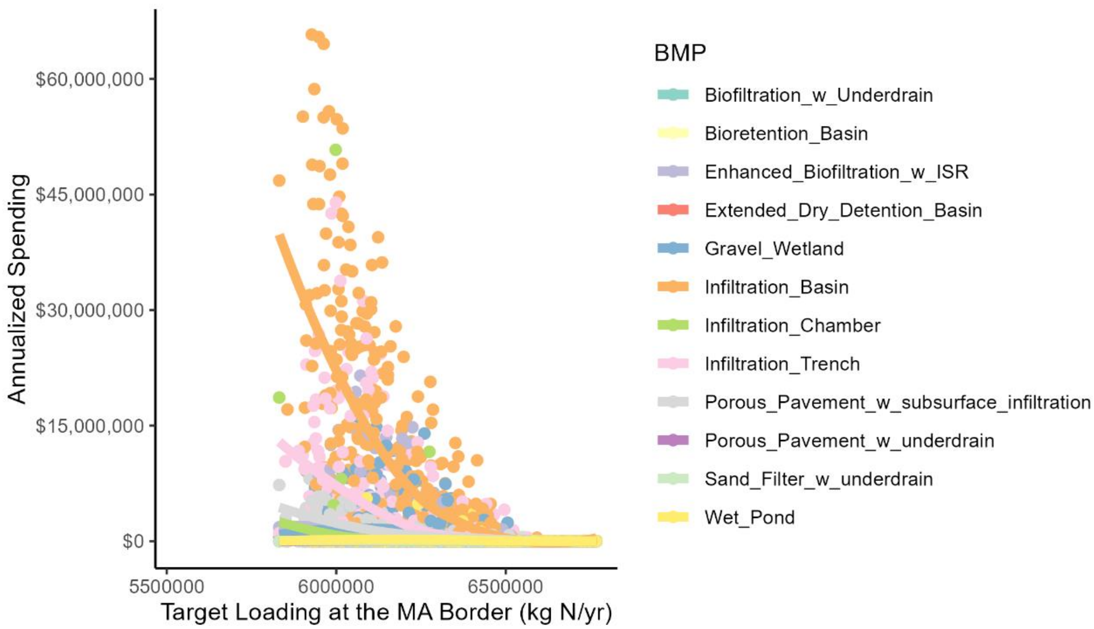


Figure 5.5: Annualized spending on urban BMPs as a function of loading target

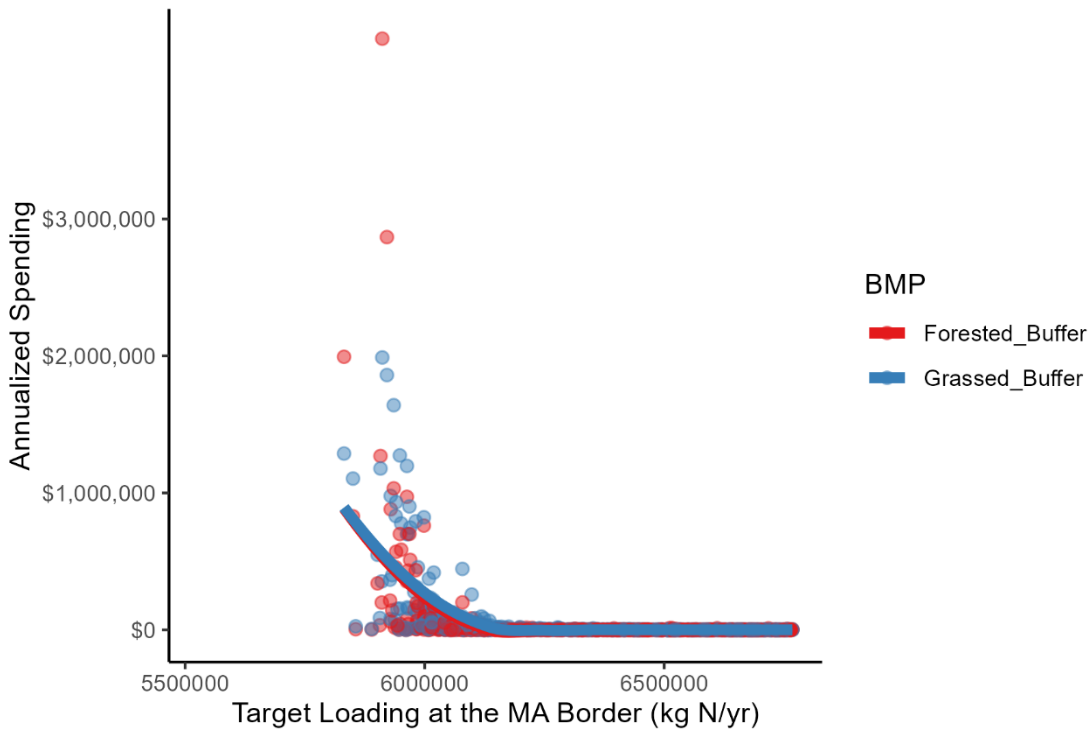


Figure 5.6: Annualized spending on riparian zone BMPs as a function of loading target

Table 5.1: RBEROST tolerance to variation in User Specifications

Parameter	Set member (BMP or other)	Tolerance to decreased values	Tolerance to increased values
ag_frac_max	Conservation	not sensitive	not applicable
ag_frac_max	Contour_Farming	not sensitive	not applicable
ag_frac_max	Fert_20	not sensitive	not applicable
ag_frac_max	Filterstrip	highly sensitive	not applicable
ag_frac_max	Manure_Injection	not sensitive	not applicable
ag_frac_max	MIN_TILL	not sensitive	not applicable
ag_frac_max	Ponds	highly sensitive	not applicable
ag_frac_max	Terrace_Only	not sensitive	not applicable
ag_frac_max	Terrace_Waterway	highly sensitive	not applicable
ag_frac_max	Waterway_Only	not sensitive	not applicable
ag_frac_min	Conservation	not applicable	not applicable
ag_frac_min	Contour_Farming	not applicable	not applicable
ag_frac_min	Fert_20	not applicable	not applicable
ag_frac_min	Filterstrip	not applicable	not applicable
ag_frac_min	Manure_Injection	not applicable	not applicable
ag_frac_min	MIN_TILL	not applicable	not applicable
ag_frac_min	Ponds	not applicable	not applicable
ag_frac_min	Terrace_Only	not applicable	not applicable
ag_frac_min	Terrace_Waterway	not applicable	not applicable
ag_frac_min	Waterway_Only	not applicable	not applicable
loads_lim_N1	NA	highly sensitive	highly sensitive
loads_lim_P1	NA	not sensitive	sensitive
loads_lim_P2	NA	highly sensitive	not sensitive

ripbuf_frac_max	Forested_Buffer	sensitive	not applicable
ripbuf_frac_max	Grassed_Buffer	highly sensitive	not applicable
ripbuf_frac_min	Forested_Buffer	not applicable	not applicable
ripbuf_frac_min	Grassed_Buffer	not applicable	not applicable
urban_design_depth	Biofiltration_w_Underdrain	not sensitive	not sensitive
urban_design_depth	Bioretention_Basin	not sensitive	not sensitive
urban_design_depth	Enhanced_Biofiltration_w_ISR	sensitive	not sensitive
urban_design_depth	Extended_Dry_Detention_Basin	not sensitive	not sensitive
urban_design_depth	Grass_Swale_w_detention	not sensitive	not sensitive
urban_design_depth	Gravel_Wetland	highly sensitive	sensitive
urban_design_depth	Infiltration_Basin	sensitive	highly sensitive
urban_design_depth	Infiltration_Chamber	not sensitive	not sensitive
urban_design_depth	Infiltration_Trench	not sensitive	not sensitive
urban_design_depth	Porous_Pavement_w_subsurface_infiltration	not sensitive	sensitive
urban_design_depth	Porous_Pavement_w_underdrain	not sensitive	not sensitive
urban_design_depth	Sand_Filter_w_underdrain	not sensitive	not sensitive
urban_design_depth	Wet_Pond	not sensitive	not sensitive
urban_frac_max	Biofiltration_w_Underdrain	not sensitive	not applicable
urban_frac_max	Bioretention_Basin	not sensitive	not applicable
urban_frac_max	Enhanced_Biofiltration_w_ISR	not sensitive	not applicable
urban_frac_max	Extended_Dry_Detention_Basin	not sensitive	not applicable
urban_frac_max	Grass_Swale_w_detention	not sensitive	not applicable
urban_frac_max	Gravel_Wetland	highly sensitive	not applicable
urban_frac_max	Infiltration_Basin	sensitive	not applicable
urban_frac_max	Infiltration_Chamber	not sensitive	not applicable
urban_frac_max	Infiltration_Trench	not sensitive	not applicable

urban_frac_max	Porous_Pavement_w_subsurface_infiltration	sensitive	not applicable
urban_frac_max	Porous_Pavement_w_underdrain	not sensitive	not applicable
urban_frac_max	Sand_Filter_w_underdrain	not sensitive	not applicable
urban_frac_max	Wet_Pond	not sensitive	not applicable
urban_frac_min	Biofiltration_w_Underdrain	not applicable	not applicable
urban_frac_min	Bioretention_Basin	not applicable	not applicable
urban_frac_min	Enhanced_Biofiltration_w_ISR	not applicable	not applicable
urban_frac_min	Extended_Dry_Detention_Basin	not applicable	not applicable
urban_frac_min	Grass_Swale_w_detention	not applicable	not applicable
urban_frac_min	Gravel_Wetland	not applicable	not applicable
urban_frac_min	Infiltration_Basin	not applicable	not applicable
urban_frac_min	Infiltration_Chamber	not applicable	not applicable
urban_frac_min	Infiltration_Trench	not applicable	not applicable
urban_frac_min	Porous_Pavement_w_subsurface_infiltration	not applicable	not applicable
urban_frac_min	Porous_Pavement_w_underdrain	not applicable	not applicable
urban_frac_min	Sand_Filter_w_underdrain	not applicable	not applicable
urban_frac_min	Wet_Pond	not applicable	not applicable

Table 5.2: RBEROST tolerance to variation in parameters in RBEROST data sets

Parameter	Set member (BMP or other)	Tolerance to decreased values	Tolerance to increased values
ag_costs_capital	Conservation	not sensitive	not sensitive
ag_costs_capital	Contour_Farming	not sensitive	not sensitive
ag_costs_capital	Fert_20	not sensitive	not sensitive
ag_costs_capital	Filterstrip	not sensitive	not sensitive
ag_costs_capital	Manure_Injection	not sensitive	not sensitive
ag_costs_capital	MIN_TILL	not sensitive	not sensitive
ag_costs_capital	Ponds	sensitive	sensitive
ag_costs_capital	Terrace_Only	not sensitive	not sensitive
ag_costs_capital	Terrace_Waterway	sensitive	not sensitive
ag_costs_capital	Waterway_Only	sensitive	not sensitive
ag_costs_operations	Conservation	not sensitive	not sensitive
ag_costs_operations	Contour_Farming	not sensitive	not sensitive
ag_costs_operations	Fert_20	not sensitive	not sensitive
ag_costs_operations	Filterstrip	sensitive	sensitive
ag_costs_operations	Manure_Injection	not sensitive	not sensitive
ag_costs_operations	MIN_TILL	not sensitive	not sensitive
ag_costs_operations	Ponds	not sensitive	not sensitive
ag_costs_operations	Terrace_Only	not sensitive	not sensitive
ag_costs_operations	Terrace_Waterway	not sensitive	not sensitive
ag_costs_operations	Waterway_Only	not sensitive	not sensitive
ag_effic_N	Conservation	not sensitive	not sensitive
ag_effic_N	Contour_Farming	not sensitive	not sensitive
ag_effic_N	Fert_20	not sensitive	not sensitive

ag_effic_N	Filterstrip	highly sensitive	highly sensitive
ag_effic_N	Manure_Injection	not sensitive	not sensitive
ag_effic_N	MIN_TILL	not sensitive	not sensitive
ag_effic_N	Ponds	highly sensitive	highly sensitive
ag_effic_N	Terrace_Only	not sensitive	not sensitive
ag_effic_N	Terrace_Waterway	highly sensitive	sensitive
ag_effic_N	Waterway_Only	not sensitive	sensitive
ag_effic_P	Conservation	not sensitive	not sensitive
ag_effic_P	Contour_Farming	not sensitive	not sensitive
ag_effic_P	Fert_20	not sensitive	not sensitive
ag_effic_P	Filterstrip	not sensitive	not sensitive
ag_effic_P	Manure_Injection	not sensitive	not sensitive
ag_effic_P	MIN_TILL	not sensitive	not sensitive
ag_effic_P	Ponds	not sensitive	not sensitive
ag_effic_P	Terrace_Only	not sensitive	not sensitive
ag_effic_P	Terrace_Waterway	not sensitive	not sensitive
ag_effic_P	Waterway_Only	not sensitive	not sensitive
agcost_frac	NA	sensitive	sensitive
area	ag	sensitive	sensitive
area	urban	sensitive	sensitive
baseloads_N1	ag	sensitive	highly sensitive
baseloads_N1	point	sensitive	highly sensitive
baseloads_N1	urban	sensitive	highly sensitive
baseloads_P1	ag	not sensitive	not sensitive
baseloads_P1	point	not sensitive	not sensitive
baseloads_P1	urban	not sensitive	not sensitive

baseloads_P2	ag	not sensitive	not sensitive
baseloads_P2	point	not sensitive	not sensitive
baseloads_P2	urban	not sensitive	not sensitive
other_loads_N1	NA	highly sensitive	highly sensitive
other_loads_P1	NA	sensitive	not sensitive
other_loads_P2	NA	not sensitive	highly sensitive
point_costs	capital	not sensitive	not sensitive
point_costs	operations	sensitive	sensitive
point_effic_N	point	sensitive	not sensitive
point_effic_P	point	not applicable	not applicable
riparianload_N1	NA	not sensitive	not sensitive
riparianload_P1	NA	not sensitive	not sensitive
riparianload_P2	NA	not sensitive	not sensitive
riparianremoval_N1	Forested_Buffer	highly sensitive	sensitive
riparianremoval_N1	Grassed_Buffer	sensitive	highly sensitive
riparianremoval_P1	Forested_Buffer	not sensitive	not sensitive
riparianremoval_P1	Grassed_Buffer	not sensitive	not sensitive
riparianremoval_P2	Forested_Buffer	not sensitive	not sensitive
riparianremoval_P2	Grassed_Buffer	not sensitive	not sensitive
ripbuf_costs_capital	Forested_Buffer	sensitive	sensitive
ripbuf_costs_capital	Grassed_Buffer	sensitive	sensitive
ripbuf_costs_operations	Forested_Buffer	not sensitive	not sensitive
ripbuf_costs_operations	Grassed_Buffer	not sensitive	not sensitive
runoff_coeff_urban	urban	sensitive	sensitive
total_banklength	NA	not sensitive	not sensitive
unbuffered_banklength	Forested_Buffer	sensitive	highly sensitive

unbuffered_banklength	Grassed_Buffer	highly sensitive	sensitive
urban_bmp_implementationpotential	Biofiltration_w_Underdrain	not sensitive	not sensitive
urban_bmp_implementationpotential	Bioretention_Basin	not sensitive	not sensitive
urban_bmp_implementationpotential	Enhanced_Biofiltration_w_ISR	not sensitive	not sensitive
urban_bmp_implementationpotential	Extended_Dry_Detention_Basin	not sensitive	not sensitive
urban_bmp_implementationpotential	Grass_Swale_w_detention	not sensitive	not sensitive
urban_bmp_implementationpotential	Gravel_Wetland	highly sensitive	not sensitive
urban_bmp_implementationpotential	Infiltration_Basin	sensitive	not sensitive
urban_bmp_implementationpotential	Infiltration_Chamber	not sensitive	not sensitive
urban_bmp_implementationpotential	Infiltration_Trench	not sensitive	not sensitive
urban_bmp_implementationpotential	Porous_Pavement_w_subsurface_infiltration	sensitive	sensitive
urban_bmp_implementationpotential	Porous_Pavement_w_underdrain	not sensitive	not sensitive
urban_bmp_implementationpotential	Sand_Filter_w_underdrain	not sensitive	not sensitive
urban_bmp_implementationpotential	Wet_Pond	not sensitive	not sensitive
urban_cost_adjustment_coef	NA	sensitive	sensitive
urban_costs	Biofiltration_w_Underdrain	not sensitive	not sensitive
urban_costs	Bioretention_Basin	not sensitive	not sensitive
urban_costs	capital	sensitive	sensitive
urban_costs	Enhanced_Biofiltration_w_ISR	sensitive	not sensitive
urban_costs	Extended_Dry_Detention_Basin	not sensitive	not sensitive
urban_costs	Grass_Swale_w_detention	not sensitive	not sensitive
urban_costs	Gravel_Wetland	highly sensitive	sensitive
urban_costs	Infiltration_Basin	sensitive	highly sensitive
urban_costs	Infiltration_Chamber	not sensitive	not sensitive
urban_costs	Infiltration_Trench	not sensitive	not sensitive
urban_costs	operations	sensitive	sensitive

urban_costs	Porous_Pavement_w_subsurface_infiltration	not sensitive	sensitive
urban_costs	Porous_Pavement_w_underdrain	not sensitive	not sensitive
urban_costs	Sand_Filter_w_underdrain	not sensitive	not sensitive
urban_costs	Wet_Pond	not sensitive	not sensitive
urban_effic_N	Biofiltration_w_Underdrain	not sensitive	not sensitive
urban_effic_N	Bioretention_Basin	not sensitive	not sensitive
urban_effic_N	Enhanced_Biofiltration_w_ISR	not sensitive	sensitive
urban_effic_N	Extended_Dry_Detention_Basin	not sensitive	not sensitive
urban_effic_N	Grass_Swale_w_detention	not sensitive	not sensitive
urban_effic_N	Gravel_Wetland	sensitive	highly sensitive
urban_effic_N	Infiltration_Basin	highly sensitive	highly sensitive
urban_effic_N	Infiltration_Chamber	not sensitive	not sensitive
urban_effic_N	Infiltration_Trench	not sensitive	sensitive
urban_effic_N	Porous_Pavement_w_subsurface_infiltration	sensitive	sensitive
urban_effic_N	Porous_Pavement_w_underdrain	not sensitive	not sensitive
urban_effic_N	Sand_Filter_w_underdrain	not sensitive	not sensitive
urban_effic_N	Wet_Pond	not sensitive	not sensitive
urban_effic_P	Biofiltration_w_Underdrain	not sensitive	not sensitive
urban_effic_P	Bioretention_Basin	not sensitive	not sensitive
urban_effic_P	Enhanced_Biofiltration_w_ISR	not sensitive	not sensitive
urban_effic_P	Extended_Dry_Detention_Basin	not sensitive	not sensitive
urban_effic_P	Grass_Swale_w_detention	not sensitive	not sensitive
urban_effic_P	Gravel_Wetland	not sensitive	not sensitive
urban_effic_P	Infiltration_Basin	not sensitive	not sensitive
urban_effic_P	Infiltration_Chamber	not sensitive	not sensitive
urban_effic_P	Infiltration_Trench	not sensitive	not sensitive

urban_effic_P	Porous_Pavement_w_subsurface_infiltration	not sensitive	not sensitive
urban_effic_P	Porous_Pavement_w_underdrain	not sensitive	not sensitive
urban_effic_P	Sand_Filter_w_underdrain	not sensitive	not sensitive
urban_effic_P	Wet_Pond	not sensitive	not sensitive

6 Data Dictionary

Table 6.1 lists the input files that are read into the preprocessing and postprocessing R code, describes the contents of the inputs and specific data fields used in the model, units of data, and data sources. Table 6.2 lists the output files that are created by the postprocessing R code, the data units, and the sources of associated data.

Table 6.1: Data Dictionary of Input Files

Input File Name	Input File Description	Data Units (If Applicable)	Source(s)
01_UserSpecs_BMPs.csv	User specification file for selection of BMPs to be included in optimization model, BMP capital and operations & maintenance costs, and urban BMP runoff depth specification. File includes default capital and operations & maintenance costs as well as default runoff depths. Minimum and maximum runoff depth fields (Min_RD_in and Max_RD_in) indicate the range between which the user can specify urban BMP runoff depths. Min_RD_in and Max_RD_in fields are not used in scaled-up optimization model code. Separate columns must be included for the capital and operations & maintenance costs for each state included in the geographic range of optimization. Column titles must follow the naming convention "capital_[two letter state code]" and "operations_[two letter state code]".	Per-area and per-volume BMP cost units described in file. Runoff depth in inches. Buffer width in feet.	BMP cost data sources described in Table 2. Minimum and maximum runoff depths for Urban BMPs from New Hampshire MS4 permit BMP performance curves (New Hampshire Department of Environment al Services, 2020b). Minimum and maximum buffer widths from Houle et al. (2019).
01_UserSpecs_loadingtargets.csv	User specification file for providing loading targets to include in RBEROST. Fields include Waterbody name, NHD+ COMID, Percent Reduction, choice of TN or TP, flags for out of network, and flags	Percent	Not Applicable

	for terminal reaches. If the user has loading targets on a mass basis, they should input the percent reduction from current loadings that the mass target represents.		
ACRE_HUC12_HRU_Summary_compareBaseline.csv	Agricultural BMP efficiencies summarized by HUC12 (or if unavailable, HUC10 and/or HUC8) based on ACRE database. “HUC” fields represent the hydrologic unit code (HUC12/HUC10/HUC8). “Scenario” field represents BMP scenarios, modeled by White et al. (2019) for the ACRE database. “MeanTN_Effic” field reflects average nitrogen removal efficiency per scenario, with respect to “Baseline” conditions.	Percent	Original ACRE database from White et al. (2019). Data summarized at HUC-levels by RBEROST associated R scripts.
ACRE_HUC12_HRU_Summary_compareNoPractice.csv	Agricultural BMP efficiencies summarized by HUC12 (or if unavailable, HUC10 and/or HUC8) based on ACRE database. “HUC” fields represent the hydrologic unit code (HUC12/HUC10/HUC8). “Scenario” field represents BMP scenarios, modeled by White et al. (2019) for the ACRE database. “MeanTN_Effic” field reflects average nitrogen removal efficiency per scenario, with respect to “No Practice” conditions.	Percent	Original ACRE database from White et al. (2019). Data summarized at HUC-levels by RBEROST associated R scripts.
AgBMPEffic_FertManure.csv	Agricultural BMP nitrogen removal efficiencies for 20% fertilizer reduction and manure injection (do not vary by location). “Category” field reflects the	Percent	EPA, NRCS

	agricultural BMP category within RBEROST, “BMP” field reflects the best management practice considered, and “N_Efficiency” field reflects the nitrogen removal efficiency.		
EQIPcosts_overyears.csv	EQIP costs for agricultural and riparian buffer BMPs for years 2018-2021. “BMP_Category” field reflects whether a BMP is implemented on agricultural (row crop) land or in riparian buffers. “BMP” field reflects the best management practice considered, “capital_units” reflect the cost units for capital costs as cost / unit area, “operations_units” reflect the cost units for operation costs as cost / unit area. The remaining fields describe the costs from EQIP payment schedules follow the naming convention "[capital or operations costs]_[state]_[year]”. At least one year of data must be provided for each state included in the geographic range of optimization.	Dollars per unit area	U.S. Department of Agriculture Staff. 2021
LengthinBuffer_2016.csv	The amount of each stream reach that is in riparian buffers of certain widths, as of 2016. Fields include “comid,” “totalbanklength_ft,” which is two times the stream reach, “totalbanklength_ft_se” describing the standard error associated with the total bank length, and fields describing the length of stream bank in buffer and the standard error around those values. Naming conventions for the	Feet	Stream reaches from McKay et al. 2012, landcover from Yang et al. 2018. Summaries provided following R

	rest of the fields are [Riparian buffer BMP]_[“buffer” for estimates, “buffer_se” for standard errors]_[minimum buffer width to be considered “in buffer”]_ft.		scripts and ArcPython model builder scripts associated with RBEROST.
NdepChange_2012_2019.csv	Changes in TN deposition. Fields include “comid,” “TDEP_TN_2012” (describing annual mass deposition in 2012), “TDEP_TN_2019 (describing annual mass deposition in 2019),” and “Change_2012_2019 (the percent change between years)”	kg-N/ha, kg-N/ha, percent	National Atmospheric Deposition Program (2021); COMID shapefiles from McKay et al. 2012; summarized with ArcMap and R scripts for RBEROST
ne_sparrow_model_input.csv	Northeastern Regional SPARROW Model input dataset. Fields used in scaled-up optimization model include “HUC_12,” “comid,” “IncAreaKM2” (incremental area per catchment), and “urban_km2” (urban area per catchment).	Incremental and urban areas in km2	Ator (2019)
ne_sparrow_model_output_tn.csv	Northeastern Regional SPARROW Total Nitrogen Model output dataset. Fields used in RBEROST include “comid,” “in_poin” (incremental point source load per catchment), “in_urb” (incremental urban source load), “in_fert”	Incremental loadings in kg/yr	Ator (2019)

	(incremental fertilizer load), “in_fix” (incremental load from direct fixation by crops), “in_manu” (incremental load from manure applications), “in_sept” (incremental load from septic systems), “in_atmo” (incremental load from atmospheric sources), and “DEL_FRAC” (fraction of incremental load delivered to terminal reach). Fields “sin_poin,” “sin_urb,” “sin_fert,” “sin_fix,” “sin_manu,” “sin_sept,” “sin_atmo,” and “SE_DEL_FRAC” describe the standard errors around each field that is used in RBEROST.		
ne_sparrow_model_output_tp.csv	Northeastern Regional SPARROW Total Phosphorus Model output dataset. Fields used in RBEROST include “comid,” “ip_poin” (incremental point source load per catchment), “ip_urb” (incremental urban source load), “ip_fert” (incremental fertilizer load), “ip_manu” (incremental load from manure applications), “ip_rock” (incremental load from bedrock weathering), and “DEL_FRAC” (fraction of incremental load delivered to terminal reach). Fields “sip_poin,” “sip_urb,” “sip_fert,” “sip_manu,” “sip_rock,” and “SE_DEL_FRAC” describe the standard errors around each field that is used in RBEROST.	Incremental loadings in kg/yr	Ator (2019)
NH_streamcat_2011_cropland.csv	New Hampshire percent cropland, Open development, low intensity	Percent	Hill et al. (2015)

	development, medium intensity development, and high intensity development data from National Land Cover Database, 2011 (catchment-specific data downloaded from StreamCat). Fields used in scaled-up optimization model include “comid,” “PctCrop2011Cat,” “PctOpUrb2011Cat,” “PctLowUrb2011Cat,” “PctMedUrb2011Cat,” and “PctHiUrb2011Cat.” *Note: if optimizing a geographic extent that includes states beyond those used in the Upper Connecticut Case study, additional state files may be included. Locating these files in the Preprocessor/Inputs folder and following the naming convention of “[two letter state code]_streamcat_2011_[cropland or imperv].csv” will allow them to be used automatically by RBEROST.		
NH_streamcat_2011_imperv.csv	New Hampshire percent imperviousness data from National Land Cover Database (catchment-specific data downloaded from StreamCat). Fields used in scaled-up optimization model include “comid” and “PctImp2011Cat.” *Note: if optimizing a geographic extent that includes states beyond those used in the Upper Connecticut Case study, additional state files may be included. Locating these files in the	Percent	Hill et al. (2015)

	Preprocessor/Inputs folder and following the naming convention of "[two letter state code]_streamcat_2011_[cropland or imperv].csv" will allow them to be used automatically by RBEROST.		
NHD+infiltrationrates.csv	The average infiltration rate in each COMID. Fields include "comid," "infiltrationrate_inperhr" (the average infiltration rate in inches per hour) and "infiltrationrate_inperhr_dist" (a resampling of infiltration rates used for RBEROST uncertainty analysis. Values are a character string of 10 samples that are parsed within RBEROST)	inches per hour, list of values in inches per hour	Hydrologic soil group values from Soil Survey Staff (2020a, 2020b). Infiltration rates for hydrologic soil groups from USDA NRCS 2009
RiparianEfficiencies.csv	Performance curves describing nutrient removal efficiency in riparian buffers. Fields include "comid," and fields with character entries of functions that RBEROST will parse to calculate efficiency. All functions are functions of buffer width, and the coefficients depend on slope and hydrologic soil group. Naming convention follows [N or P, designating which nutrient removal efficiency is being described]_[type of riparian buffer BMP]_[approximate width of buffer].	Percent	Methodology: Houle et al. 2019. River reach shape files: McKay et al. 2012. Hydrologic soil group: gNATSGO (Soil Survey Staff, 2020a; Soil Survey Staff, 2020b;).

	Though all equations are a function of buffer width, the slope and hydrologic soil group may differ based on the buffer width which may change the coefficients. Columns that describe uncertainty are designated with "_uncertainty" at the end. Uncertainty columns contain a character string of 10 different equations produced by sampling the slope and hydrologic soil groups of each buffer width. These character strings are parsed within RBEROST.		Slope data: Verdin, 2017
RiparianLoadings.csv	Nutrient loading from riparian zones. Fields include “comid,” “N_riparian_kgyr” “P_riparian_kgyr,” “N_riparian_kgyr_se,” and “P_riparian_kgyr_se,” where N or P designates the nutrient being described, and _se designates standard errors.	kg-N/yr; kg-P/yr; kg-N/yr; kg-P/yr	Methodology for calculating loading: Houle et al. 2019. Land cover data: Jin et al. 2019, Yang et al. 2018. River reach shape files: McKay et al. 2012
UrbanBMPPerformanceCurves.csv	Performance curves that describe the nutrient removal efficiency of urban BMPs. Fields include: “BMP” (the best management practice under consideration), “Pollutant” (designates if the performance curve relates to N or P), “InfiltrationRate_inperhr” (the	y in “Best.Fit.Curve” are in units percent, where x is the rating depth of the BMP in	New Hampshire Department of Environmental Service (2020b)

	infiltration rate that the performance curve assumes. BMPs that do not depend on infiltration rate have “NA.”), “Best.Fit.Curve” (a character string that lists the form of the performance curve), and “Coef.1,” “Coef.2” and “Coef.3.” The “Coef” columns list the coefficients that correspond with the function in “Best.Fit.Curve.” Standard error around these coefficients are given in “Coef.1_se,” “Coef.2_se” and “Coef.3_se.”	inches.	
VT_streamcat_2011_cropland.csv	Vermont percent cropland, Open development, low intensity development, medium intensity development, and high intensity development data from National Land Cover Database, 2011 (catchment-specific data downloaded from StreamCat). Fields used in scaled-up optimization model include “comid,” “PctCrop2011Cat,” “PctOpUrb2011Cat,” “PctLowUrb2011Cat,” “PctMedUrb2011Cat,” and “PctHiUrb2011Cat”. *Note: if optimizing a geographic extent that includes states beyond those used in the Upper Connecticut Case study, additional state files may be included. Locating these files in the Preprocessor/Inputs folder and following the naming convention of “[two letter state code]_streamcat_2011_[cropland or	Percent	Hill et al. (2015)

	imperv].csv" will allow them to be used automatically by RBEROST.		
VT_streamcat_2011_imperv.csv	Vermont percent imperviousness data from National Land Cover Database (catchment-specific data downloaded from StreamCat). Fields used in scaled-up optimization model include "comid" and "PctImp2011Cat." *Note: if optimizing a geographic extent that includes states beyond those used in the Upper Connecticut Case study, additional state files may be included. Locating these files in the Preprocessor/Inputs folder and following the naming convention of "[two letter state code]_streamcat_2011_[cropland or imperv].csv" will allow them to be used automatically by RBEROST.	Percent	Hill et al. (2015)
WWTP_BaselineRemoval_Finergrain.csv	An extension of WWTP_COMIDs_BslnRemoval.csv that includes loading for each year available in New Hampshire Department of Environmental Services 2020a. RBEROST uses the "load_lbday_2014/2015/2016/2017/2018" columns to calculate the standard error of change over time.	kg/yr or lb/day as designated in column headings	New Hampshire Department of Environmental Services (2020a)
WWTP_COMIDs_BslnRemoval.csv	Percentage change in WWTP effluent loadings from 2014-2018 based on Newport Wastewater Treatment Facility NPDES Permit No. NH0100200. Data used to scale point source loading	Percent	New Hampshire Department of Environment

	estimates from the Northeastern Total Nitrogen SPARROW Model (representing 2012 loadings) to current (as of 2018) loadings. “Rem_2014_2018_load_ch” field represents percent loading change from 2014-2018 per WWTP.		al Services (2020a)
WWTP_COMIDs.csv	Crosswalk between the WWTPs included in the scaled-up optimization model example application and their locations along the NHDPlus Version 2 flow network.	Not applicable	McKay et al. (2012), Ator (2019)
WWTP_RemovalEffic.csv	Estimated nitrogen removal efficiencies based on BioWin modeling. “Category” field reflects the point source BMP category within the scaled-up optimization model code, “BMP” field reflects the WWTP considering implementation of low-cost retrofits, and “N_Efficiency” field reflects the nitrogen removal efficiency.	Percent	JJ Environmental (2015)

Table 6.2: Data Dictionary of Output Files

Input File Name	Input File Description	Data Units (If Applicable)
AgBMP_by_ComID_Scenario1.csv (or other scenarios, if not renamed by user)	Agricultural (row crop) BMP implementation by COMID. Fields include "COMID," "HUC_12" (12 digit hydrologic unit code), "IncAreaKm" (the area of the catchment defined by COMID), "ag_km2" (the area within that COMID defined as ag), "PctCrop2011Cat" (the percent of NLCD 2011 pixels categorized as row crop in the COMID as summarized by StreamCat) and "StateAbbrev" (VT or NH). Any additional fields follow the naming convention of [BMP name]_[either FracImplement (the fraction of ag area treated with this BMP) or areaac (the total acreage of area treated with this BMP)]	_FracImplement fields as fractions, _areaac fields as acres
Urban_by_ComID_Scenario1.csv (or other scenarios, if not renamed by user)	Urban BMP implementation by COMID. Fields include "COMID," "HUC_12" (12 digit hydrologic unit code), "IncAreaKm" (the area of the catchment defined by COMID), "urban_km2" (the area within that COMID defined as urban), and "StateAbbrev" (VT or NH). Any additional fields follow the naming convention of [BMP name]_[either FracImplement (the fraction of urban area treated with this BMP) or areaac (the total acreage of area treated with this BMP)]	_FracImplement fields as fractions, _areaac fields as acres
RipBufBMP_by_ComID_Scenario1.csv (or other scenarios, if not renamed by user)	Riparian buffer BMP implementation by COMID. Fields include "COMID," "HUC_12" (12 digit hydrologic unit code), "IncAreaKm" (the area of the catchment defined by COMID), "totalbanklength_ft" (the distance of bank within the COMID, or 2 times "riverreachlength_ft"), "StateAbbrev" (VT or NH), and "riverreachlength_ft" (the length of river). Additional fields follow the naming convention of [BMP name]_LengthImplemented_ft.	Length, in ft

References

Ator, S. W. (2019*). Spatially referenced models of streamflow and nitrogen, phosphorus, and suspended-sediment loads in streams of the Northeastern United States.* U.S. Geological Survey Scientific Investigations Report 2019-5118, 57 p. Retrieved from `https://pubs.er.usgs.gov/publication/sir20195118`

Dell, C., Allen, A., Dostie, D., Meinen, R., & Maguire, R. (2016). *Manure Incorporation and Injection practices for Use in Phase 6.0 of the Chesapeake Bay Program Watershed Model*. Chesapeake Bay Program. Retrieved from `https://www.chesapeakebay.net/documents/Phase_6_FINAL_MII_Final_Report.pdf`

Detenbeck, N., ten Brink, M., Piscopo, A., Morrison, A., Stagnitta, T., Abele, R., et al. (2018). *Watershed Management Optimization Support Tool (WMOST) v3: Theoretical Documentation*. U.S. Environmental Protection Agency, Washington, DC EPA/600/R-17/220. Retrieved from `https://www.epa.gov/ceam/wmost-30-documentation`

Detenbeck, N., Piscopo, A., ten Brink, M., Weaver, C., Morrison, A., Stagnitta, T., et al. (2018). *Watershed Management Optimization Support Tool (WMOST) v3: User Guide*. U.S. Environmental Protection Agency, Washington, DC EPA/600/R-17/255. Retrieved from `https://www.epa.gov/ceam/wmost-30-documentation`

Heris, M. P., Foks, N. L., Bagstad, K. J., Troy, A., & Ancona, Z. H. (2020). A rasterized building footprint dataset for the United States. *Scientific Data*, *7*(207). `https://doi.org/10.1111/1752-1688.12372`

Hill, R. A., Weber, M. H., Leibowitz, S. G., Olsen, A. R., & Thornbrugh, D. J. (2015). The Stream-Catchment (StreamCat) Dataset: A Database of Watershed Metrics for the Conterminous United States. *Journal of the American Water Resources Association (JAWRA)*, *52*, 120–128. `https://doi.org/10.1111/1752-1688.12372`

Houle, J., Riley, C., & Leonard, D. (2019). *Pollutant Removal Credits for Buffer Restoration in MS4 Permits: Final Panel Report*. National Estuarine Research Reserve System Science Collaborative. Retrieved from `https://nerrssciencecollaborative.org/resource/pollutant-removal-credits-buffers`

Jin, S., Homer, C. G., Yang, L., Danielson, P., Dewitz, J., Li, C., et al. (2019). Overall methodology design for the United States National Land Cover Database 2016 products. *Remote Sensing*, *11*(24). `https://doi.org/10.3390/rs11242971`

JJ Environmental. (2015). *Low Cost Retrofits for Nitrogen Removal at Wastewater Treatment Plants in the Upper Long Island Sound Watershed*. N-2012-047. Retrieved from `https://neiwpcc.org/wp-content/uploads/2018/11/2012-047-Final-Report.pdf`

McKay, L., Bondelid, L., Dewald, T., Johnston, J., Moore, R., & Rea, A. (2012). *NHDPlus Version 2: User Guide*. U.S. Environmental Protection Agency, Washington, DC. Retrieved from `https://www.epa.gov/waterdata/learn-more#metadata`

National Atmospheric Deposition Program. (2021). *Total Deposition Maps. v2018.01*. Retrieved from http://nadp.slh.wisc.edu/committees/tdep/tdepmaps/

New Hampshire Department of Environmental Services. (2020a). *Authorization to Discharge Under the National Pollutant Discharge Elimination System; NPDES Permit No. NH0100200*. New Hampshire Department of Environmental Services. Retrieved from https://www.epa.gov/sites/production/files/2020-03/documents/draftnh0100200permit.pdf

New Hampshire Department of Environmental Services. (2020b). *General Permits for Stormwater Discharges from Small Municipal Storm Sewer Systems in New Hampshire*. New Hampshire Department of Environmental Services. Retrieved from https://www3.epa.gov/region1/npdes/stormwater/nh/2017-small-ms4-general-permit-nh-mod.pdf

Schueler, T. R. (1987). *Controlling Urban Runoff: A Practical Manual for Planning and Designing Urban BMPs*. The Metropolitan Washington Council of Governments, Washington Metropolitan Water Resources Planning Board.

Soil Survey Staff. (2020a). *The Gridded National Soil Survey Geographic (gNATSGO) Database for New Hampshire*. United States Department of Agriculture, Natural resources Conservation Service. Retrieved from https://nrcs.app.box.com/v/soils

Soil Survey Staff. (2020b). *The Gridded National Soil Survey Geographic (gNATSGO) Database for Vermont*. United States Department of Agriculture, Natural resources Conservation Service. Retrieved from https://nrcs.app.box.com/v/soils

U.S. Department of Agriculture, Natural Resources Conservation Service. (n.d.). Hydrologic Soil Group. Retrieved from https://directives.sc.egov.usda.gov/OpenNonWebContent.aspx?content=22526.wba

U.S. Department of Agriculture Staff. (2021). *Environmental Quality Incentives Program Fiscal Years 2018-2021*. United States Department of Agriculture. Retrieved from https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/programs/financial/?cid=nrcseprd1328426

U.S. Geological Survey, & U.S. Department of Agriculture, Natural Resources Conservation Service. (2013). Federal Standards and Procedures for the National Watershed Boundary Dataset (WBD) (4 ed.). Retrieved from http://pubs.usgs.gov/tm/tm11a3/

U.S. Geological Survey, National Geospatial Program. (2020). *USGS National Hydrography Dataset Best Resolution (NHD) for Hydrologic Unit (NH) 4 - 0410: Metadata*. U.S. Geological Survey. Retrieved from https://www.sciencebase.gov/catalog/item/5a58a47fe4b00b291cd68754

University of Wisconsin in Madison. (2021). NEOS Server: State-of-the-Art Solvers for Numerical Optimization. Wisconsin Institute for Discovery. Retrieved from https://neos-server.org/neos/

UVM Spatial Analysis Lab. (2019). *Vermont High-Resolution Land Cover: Final Report*. State of Vermont. Retrieved from `https://geodata.vermont.gov/pages/land-cover#documentation`

Verdin, K. L. (2017). *Hydrologic Derivatives for Modeling and Analysis - A new global high-resolution database*. U.S. Geological Survey Data Series 1053, 16 p. Retrieved from `https://doi.org/10.3133/ds1053.`

Vermont Agency of Natural Resources. (2017). *2017 Vermont Stormwater Management Manual Rule and Design Guidance*. State of Vermont. Retrieved from `https://dec.vermont.gov/sites/dec/files/wsm/stormwater/docs/Permitinformation/2017%20VSMM_Rule_and_Design_Guidance_04172017.pdf`

Voorhees, M. (2016). *Opti-Tool: Stormwater Nutrient Management Optimization Tool*. EPA. Retrieved from `https://www.epa.gov/tmdl/opti-tool-epa-region-1s-stormwater-management-optimization-tool#:~:text=Opti%2DTool%20is%20a%20spreadsheet,landscapes%20throughout%20the%20New%20England`

Walker, J. T., Bell, M. D., Schwede, D., Cole, A., Beachley, G., Lear, G., & Wu, Z. (2019). Aspects of uncertainty in total reactive nitrogen deposition estimates for North American critical load applications. *Science of the Total Environment*, *690*, 1005–1018. `https://doi.org/10.1016/j.scitotenv.2019.06.337`

White, M., DiLuzio, M., Gambone, M., Smith, D., McLellan, E., Bieger, K., et al. (2019). Development of Agricultural Conservation Reduction Estimator (ACRE), a simple field-scale conservation planning and evaluation tool. *Journal of Soil and Water Conservation*, *74*(6). `https://doi.org/10.2489/jswc.74.6.537`

Wickham, J., Stehman, S. V., Gass, L., Dewitz, J. A., Sorenson, D. G., Granneman, B. J., et al. (2017). Thematic accuracy assessment of the 2011 National Land Cover Database (NLCD). *Remote Sensing of the Environment*, *191*, 328–341. `https://doi.org/10.1016/j.rse.2016.12.026`

Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S., et al. (2018). A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. *ISPRS Journal of Photogrammetry and Remote Sensing*, *146*, 108–123. `https://doi.org/10.1016/j.isprsjprs.2018.09.006`