

AGENDA

1	Timeline
2	Overview
3	From West Virginia to Chesapeake Bay
4	Goals and Objectives
5	Interactive Optimization and Decision-Making
6	Next steps
7	2024 Chesapeake Bay Optimization Webinars



Timeline of the Project

Calendar Year				2021	021				022			2023 2024			2024	ļ			2025	2025 2			2026	
Calendar Quarter	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1
Project Year		Ye	ar 1			Yε	ar 2			Ye	ar 3			Ye	ar 4			Ye	ar 5			Yea	ar 6	
Task 1: Development of an efficient single-objective																								
optimization procedure for cost-effective BMP allocation																								\perp
1.1: Understanding CAST modules and effect of BMPs on objectives and constraints																								
1.2: Development of a simplified point-based structured single- objective optimization procedure																							ı	
1.3: Development of a hybrid customized single-objective optimization procedure																								
1.4: Verification and validation with CBP users and decision-makers and update of optimization procedure																								
			<u> </u>			-											╢——	<u> </u>			-			+-
Task 2: Development of an efficient multi-objective (MO) optimization procedure for cost-loading trade-off BMP																								
allocation																	_		<u> </u>					₩
2.1: Develop generative MO optimization using hybrid optimization procedure developed at Task 1																								
 2.2: Develop simultaneous MO customized optimization using population-based evolutionary algorithms 2.3: Comparison of generative & simultaneous procedures and validation with CBP users & decision-makers 2.4: Develop an interactive multi-criterion decision-making aid for choosing a single preferred solution 																								
																							ı	
Task 3: Multi-state implementation using machine learning and parallel computing platforms																							ı	
3.1: Comparative study to choose a few best performing methods																								
3.2: Scalability to State and Watershed level Scenarios																								ــــــ
3.3: "Innovization" approach for improving scalability						_																		┼
4.4: Distributed computing approach for improving scalability			-			_			-								-	-			-			+-
Task 4: Interactive optimization and decision-making using																								
user-friendly dashboard		-	-			-	-			-								+	-					
4.1: User-friendly optimization through a dashboard 4.2: Surrogate-assisted optimization procedures					-		<u> </u>		-								-	+						+
4.3: Robust optimization method for handling uncertainties in					-				-								1	1						
variables and parameters																								
4.4: Sustainable watershed management practices																								

We are here



Task 3: Multi-state implementation using machine learning and parallel computing platforms

- 3.1 Comparative study to choose a few best performing methods
- 3.2 Scalability to state and watershed level scenarios
- 3.3 "Innovization" approach for improving scalability
- 3.4 Distributed computing approach for improving scalability

Timeline of the Project





Problem Statement

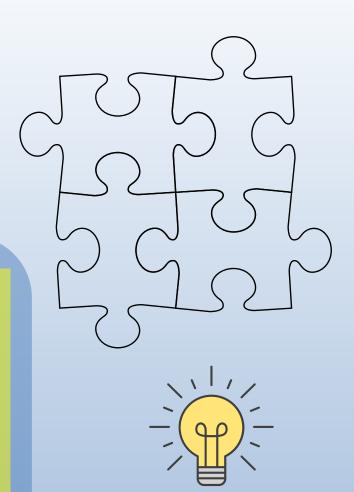
One of the most challenging issues for a real-world optimization problem is the **Search Space Complexity** (variables and constraints), especially for the CBW management problem.





Beating the "Curse of Dimensionality"

- Due to the large dimension of the problem, there is no off-the-shelf optimization algorithm capable of handling this problem.
- ☐ Therefore, we developed a customized optimization approach to speed up computational time and reduce the size of optimization variables to make the problem solvable in a reasonable time frame.





Innovation through Optimization

The major problem we are facing is the large number of optimization variables. These variables originate from three major components:

- Type of BMPs
- BMP implementation location
- Size of BMPs



Innovization

By understanding the common characteristics of the group of BMPs or locations of implementation, we can reduce the number of BMPs and ultimately reduce the total number of variables.



Study Area



The number of variables and constraints depends on the scale of the county or state under study.

West Virginia



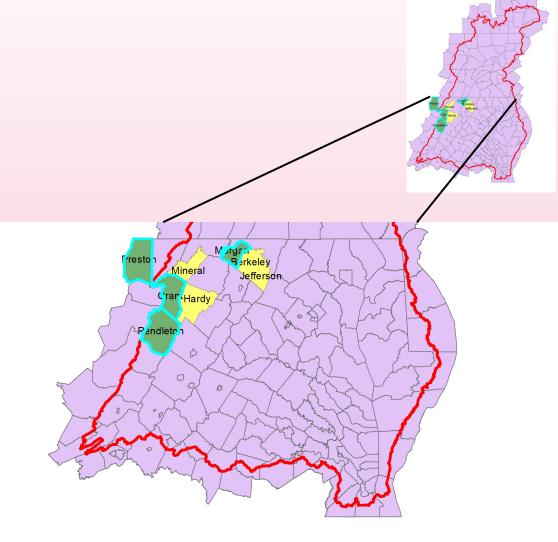
Four county: 65,260 variables (Berkeley, Jefferson, Mineral, and Hardy)



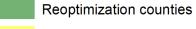
Entire state: the number of variables is about **153,818.**



For any optimization algorithm, these numbers are regarded as substantial and computationally challenging.

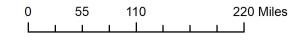


Legend

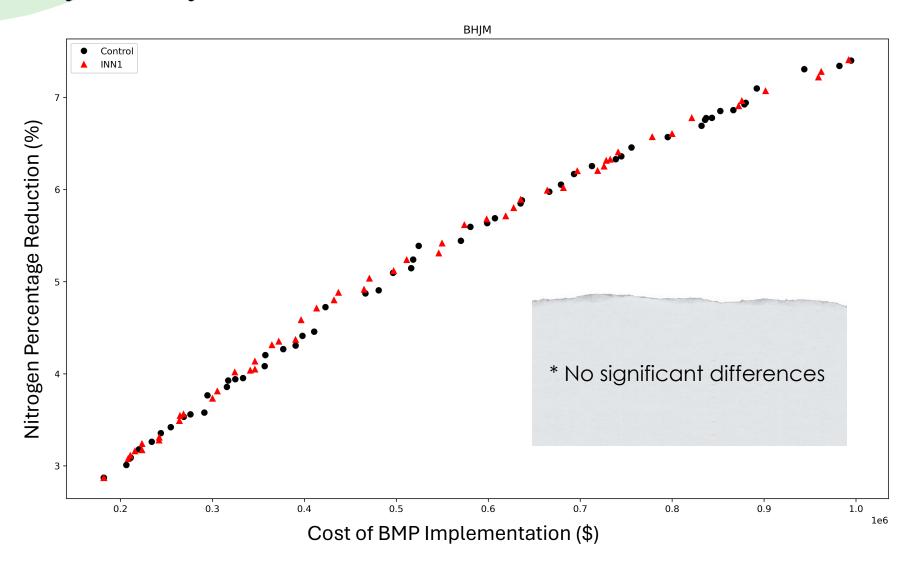


Counties for original optimization





Compare Original vs. Innovized Optimization Berkeley, Hardy, Jefferson, and Mineral Counties







Reoptimization



Evaluate the performance of original innovization technique other counties



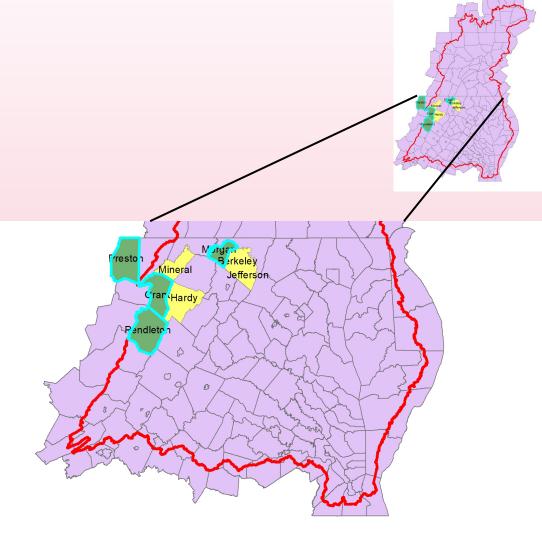
West Virginia



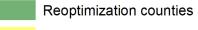
Four counties: 71,661 variables (Grant, Morgan, Pendleton, and Preston)

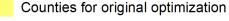


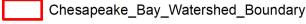
The best innovization strategy (10 BMPs) outperforms the control (>200 BMPs).

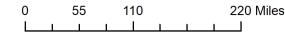


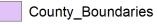
Legend





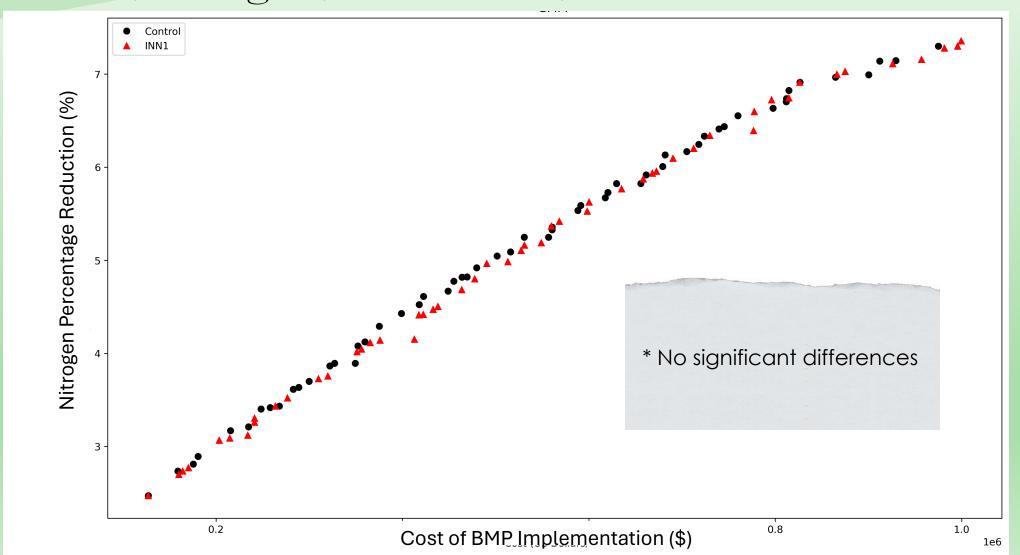








Reoptimization Compare Original vs. Innovized Optimization Grant, Morgan, Pendleton, and Preston Counties





Innovizationfor efficiency BMPs-



By reducing the number of BMP from about 200 types of efficient BMPs to only 10 BMPs, the size of the problem was significantly reduced.



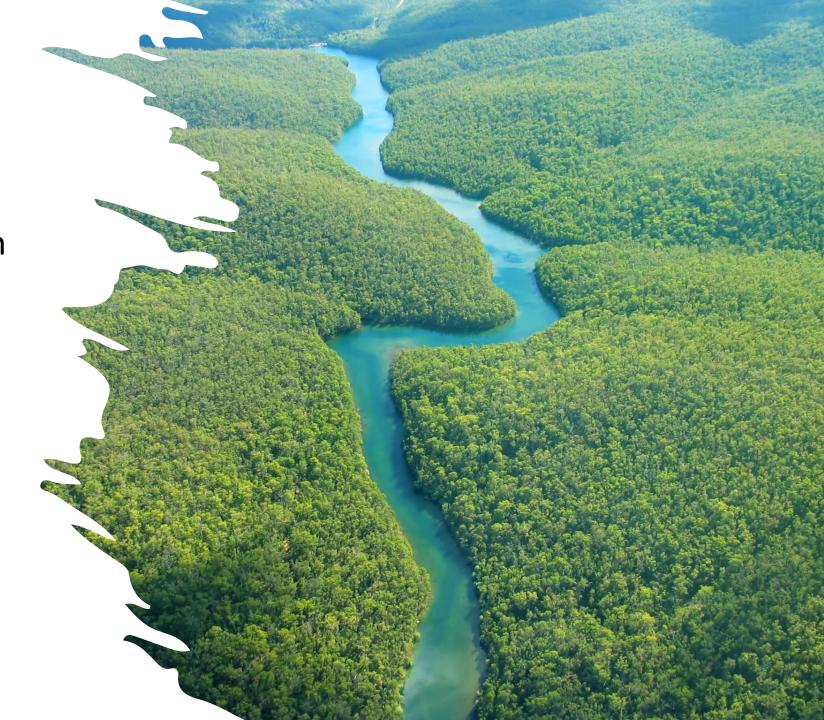
The innovizaed approach can reduce the number of variables by **about 96%**, which is a promising result.



Challenge

Innovization methods in WV effectively:

- reduce optimization problem size
- lower computational time and resources

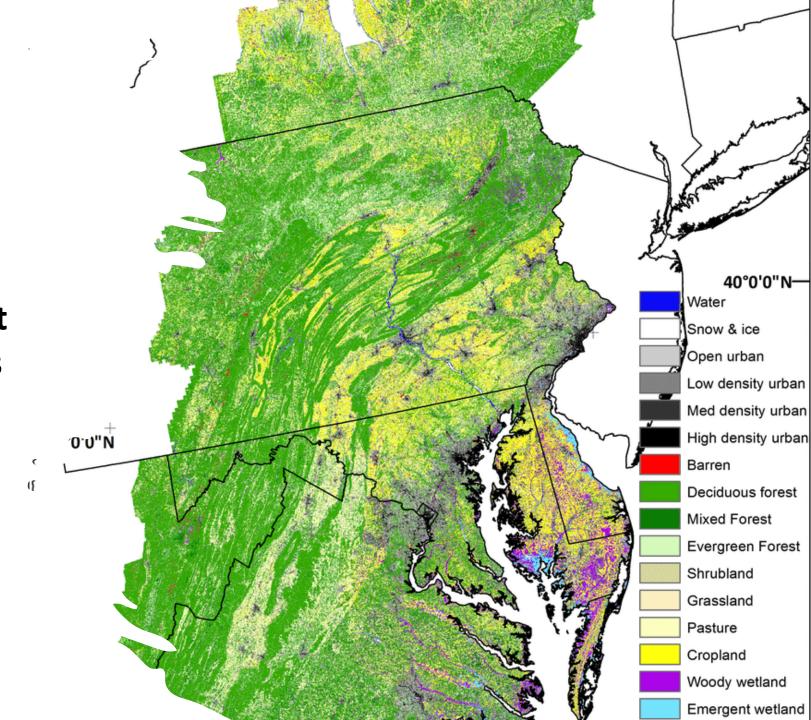


Study Area



Challenge

Innovization methods in WV don't fully represent the diverse load sources and land use traits of the entire Chesapeake Bay watershed.



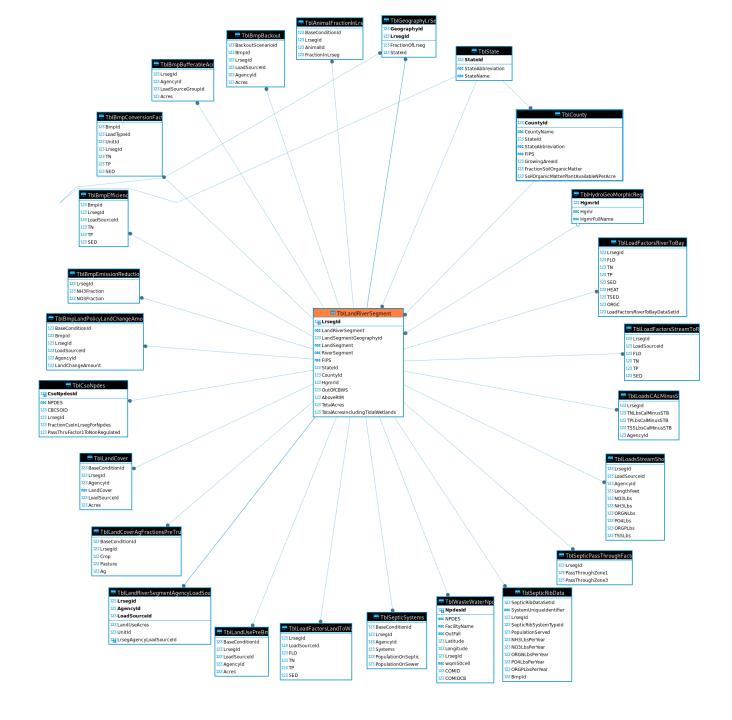


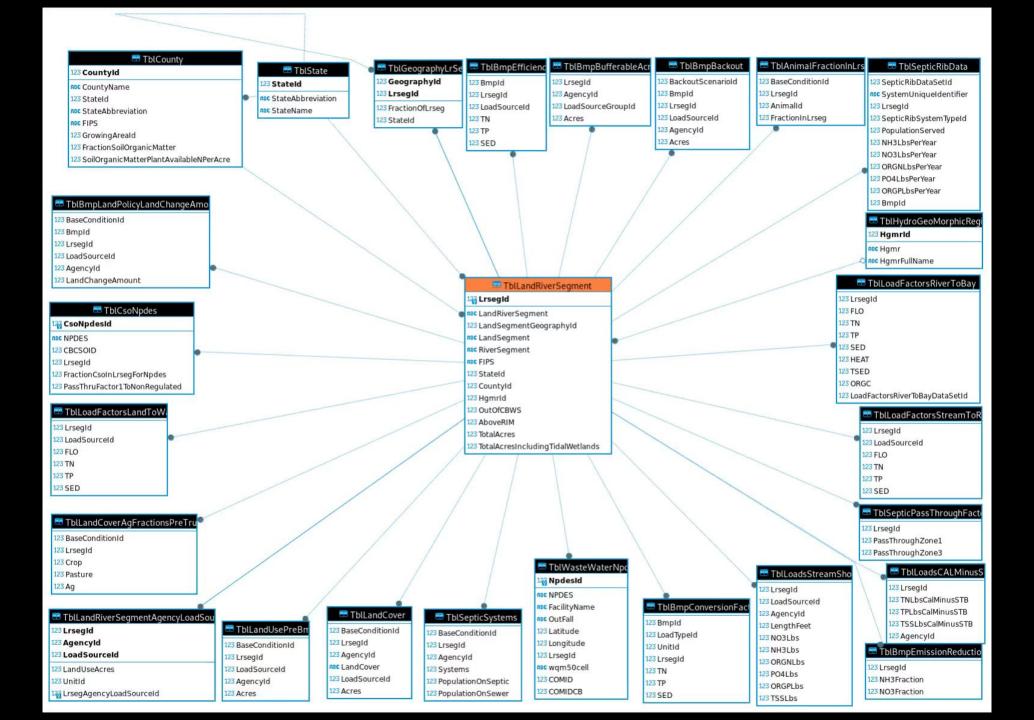
Objectives

Identify	Identify the most relevant variables.
Recognize	Recognize areas with similarities concerning the watershed features established in the first objective.
Determine	Determine the most frequently chosen BMPs for various land river segments.
Offer	Offer recommendations on selecting the most appropriate BMPs for optimization through the application of artificial intelligence methods.
Examine	Examine the reliability of the proposed strategy within the Chesapeake Bay Watershed.



Step 1: Variable Selection





Variable selection

COLUMN_NAME	DATA_TYPE
AboveRIM	bit
Acres	decimal
Ag	real
Agencyld	tinyint
AnimalId	tinyint
BackoutScenariold	int
BaseConditionId	int
Bmpld	smallint
CBCSOID	smallint
COMID	int
COMIDCB	int
Countyld	smallint
Crop	real
CsoNpdesId	smallint
FacilityName	varchar
FIPS	char
FLO	real
${\sf FractionCsoInLrsegForNpdes}$	real
FractionInLrseg	real
real	
Geographyld	int
HEAT	real
Hgmrld	tinyint
LandChangeAmount	real
LandCover	varchar
LandRiverSegment	char
LandSegment	char
LandSegmentGeographyId	int
LandUseAcres	bit



Step 2: Variable Reduction

VARIABLE REDUCTION TECHNIQUE CLUSTERING

How do you simplify your data analysis?



Variable Reduction

 Variable reduction streamlines the modeling process by minimizing the necessary inputs for constructing robust predictive or segmentation models, enhancing efficiency and manageability.

 By reducing redundant data, variable clustering clarifies the fundamental patterns and associations within the dataset's input variables.

Various Techniques for Variable Reduction

1. Principal Component Analysis (PCA): PCA is a statistical technique that transforms the original variables into a new set of uncorrelated variables ordered so that the first few retain most of the variation present in the original variables.

2. Bayesian variable reduction:

Bayesian variable reduction is a method that leverages Bayesian statistics to pinpoint the most relevant variables, incorporating prior knowledge and data evidence. It is useful for managing large variable sets and model uncertainty.

3. LASSO (Least Absolute Shrinkage and Selection Operator)

LASSO is a method that combines variable selection and regularization to improve model accuracy and simplicity by penalizing the absolute size of coefficients, reducing some to zero, ideal for datasets with more variables than observations.

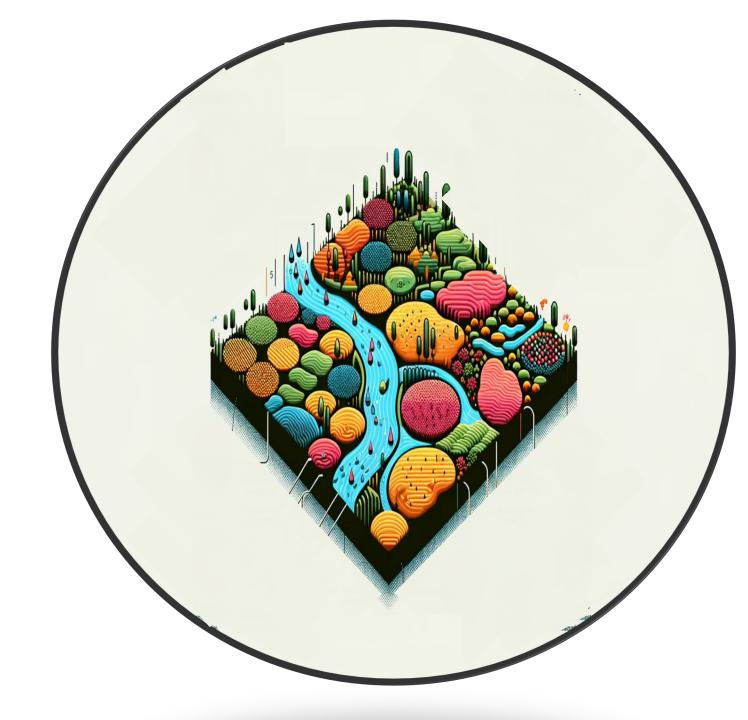
4. Ridge Regression

Ridge regression is a technique for analyzing multiple regression data that suffer from multicollinearity. Multicollinearity occurs when predictor variables in a model are highly correlated and can lead to unstable estimates of the regression coefficients. Ridge regression stabilizes these coefficients.

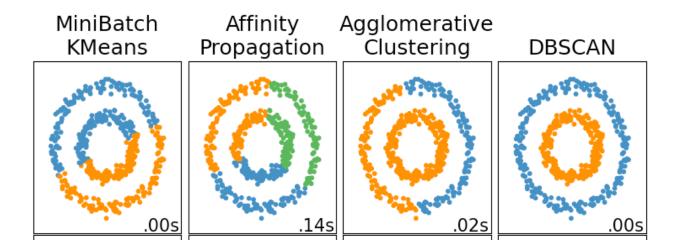




Step 3: Clustering



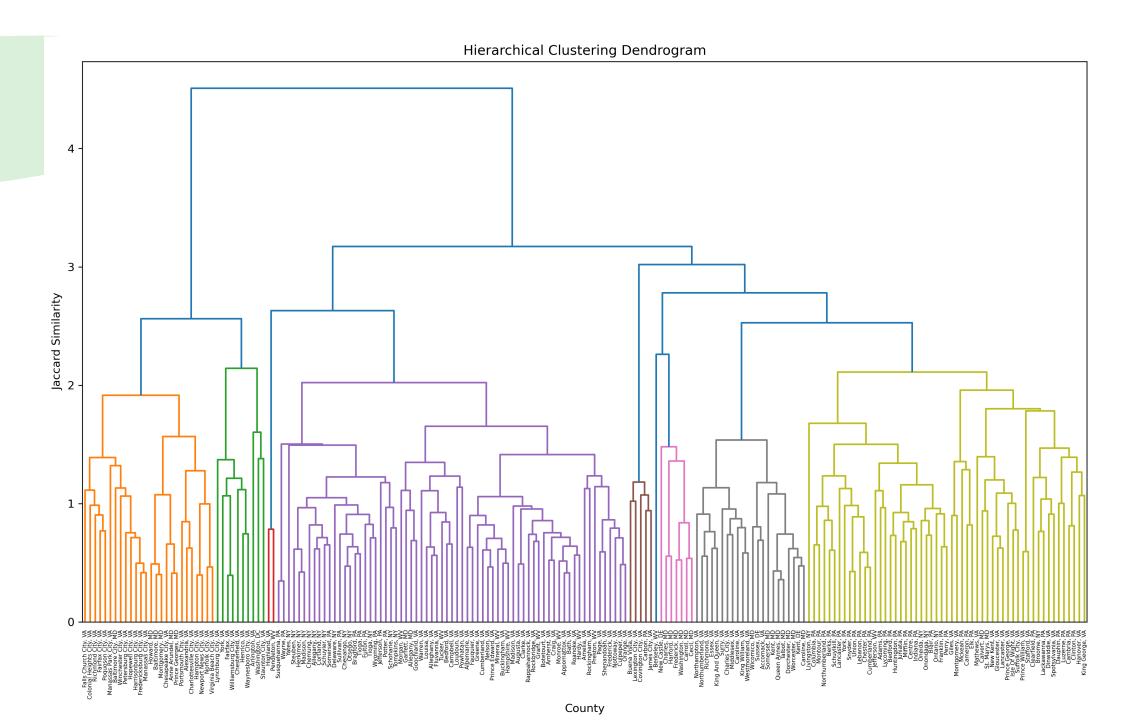
Clustering is an unsupervised/supervised learning method in machine learning and data analysis for grouping similar objects into clusters, widely used for statistical analysis across various fields.

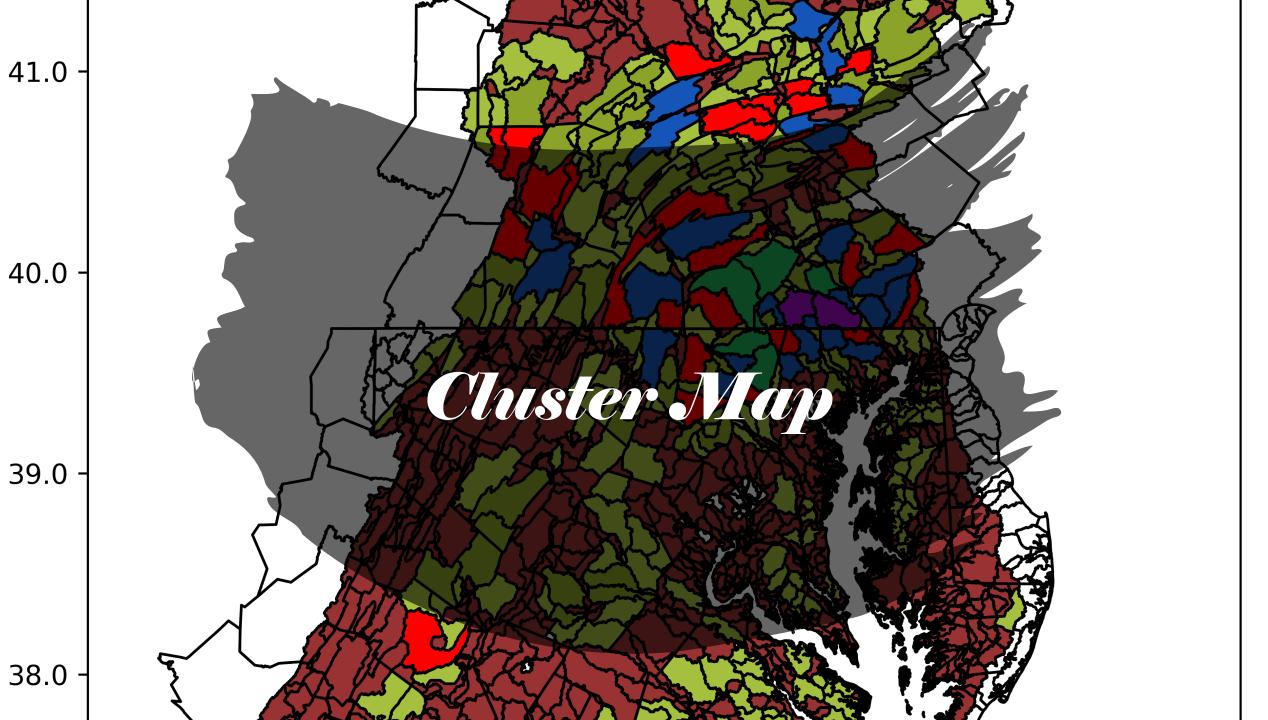


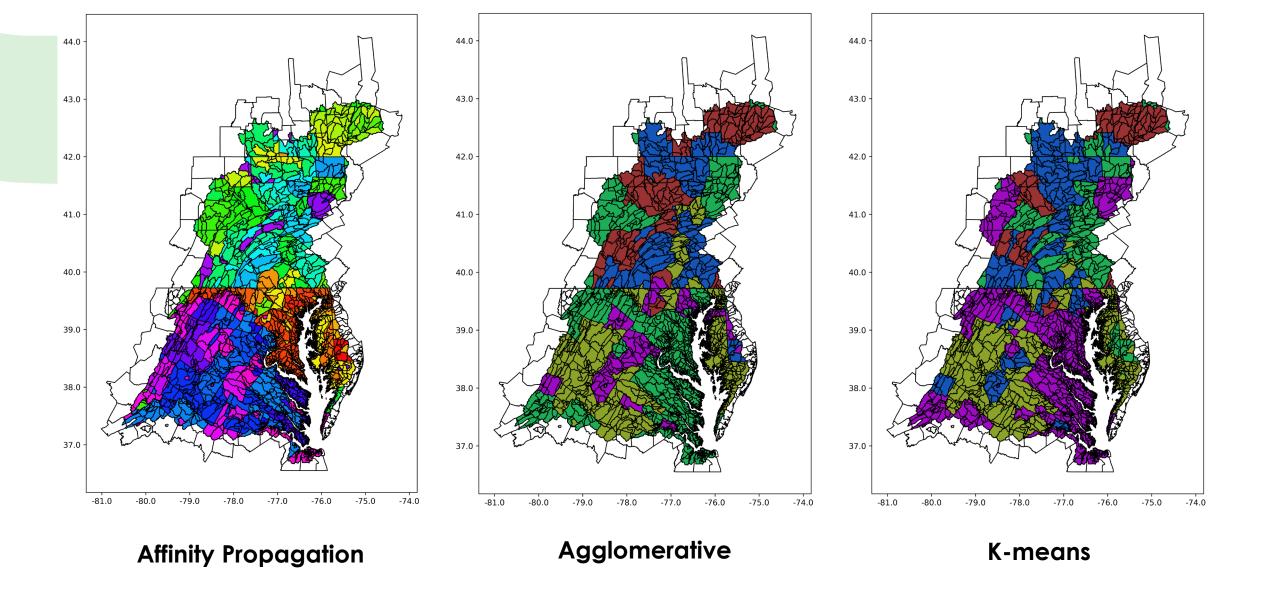


Common Clustering Techniques:

- K-Means Clustering: It partitions the dataset into K pre-defined distinct non-overlapping subgroups (clusters) where each data point belongs to only one group.
- <u>Hierarchical Clustering</u>: This method creates a tree of clusters called a dendrogram, which shows the arrangement of the clusters produced by the corresponding analyses.
- DBSCAN (Density-Based Spatial Clustering of Applications with Noise): It groups together points that are closely packed together, marking as outliers the points that lie alone in low-density regions.
- <u>Affinity Propagation</u> is determining clusters by identifying representative 'exemplars' through data point messaging.







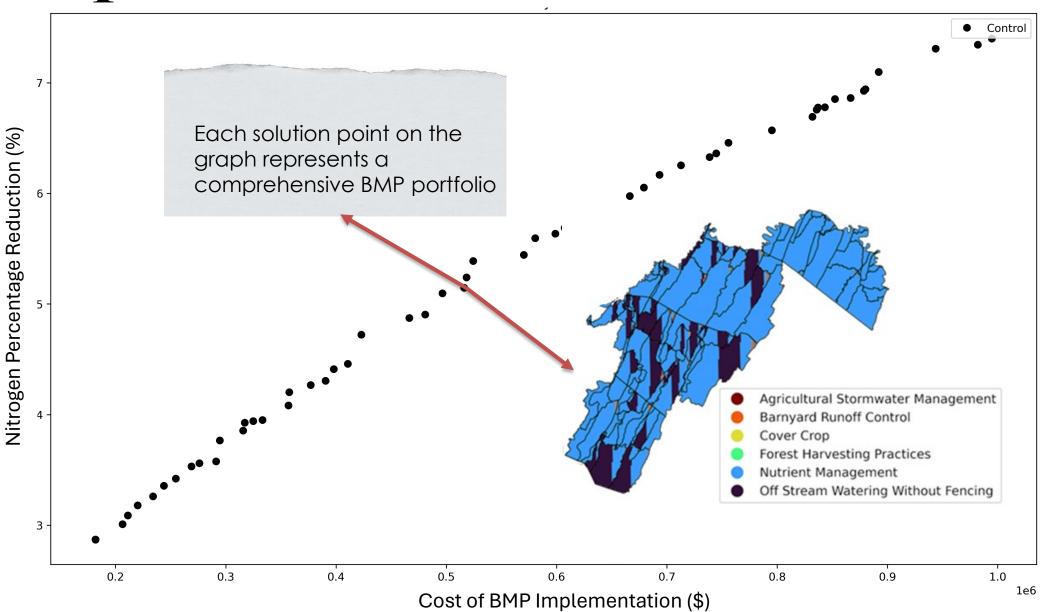
Clustering and Variable Reduction

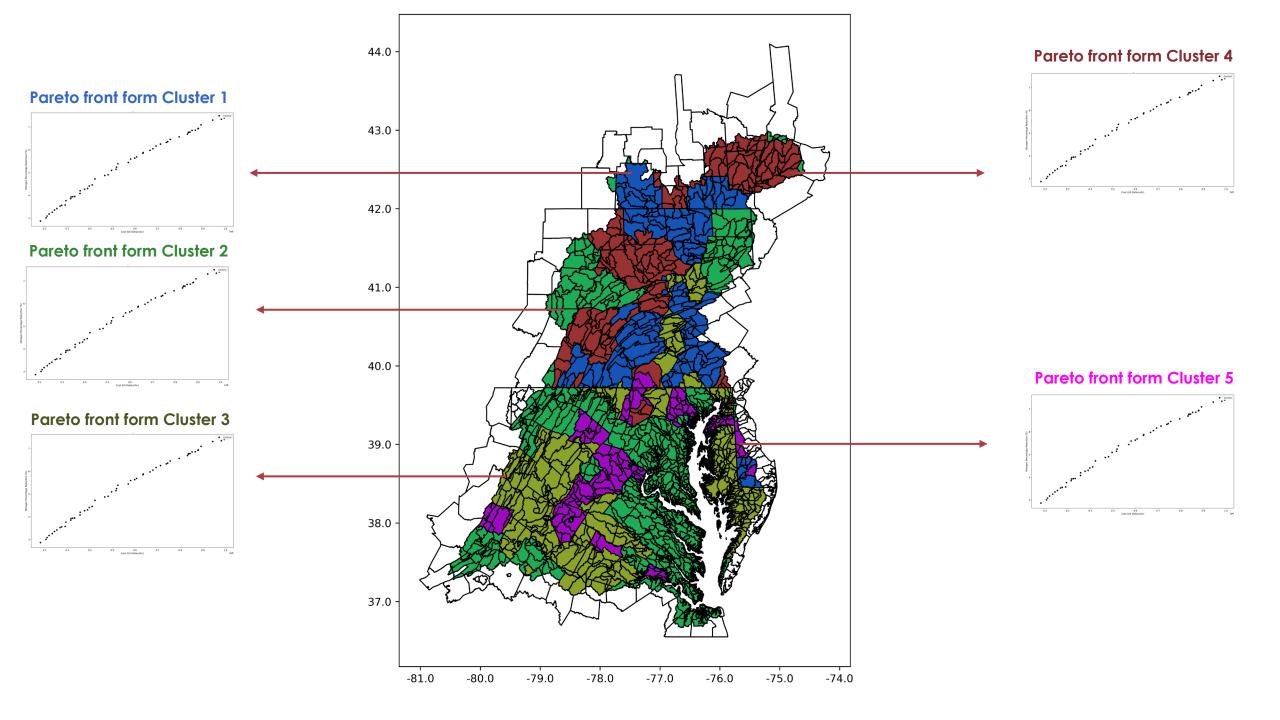


Optimization

 Perform an analysis using optimization techniques to determine the BMPs that are selected most often in land river segment and identify their specific features or characteristics.

Optimization

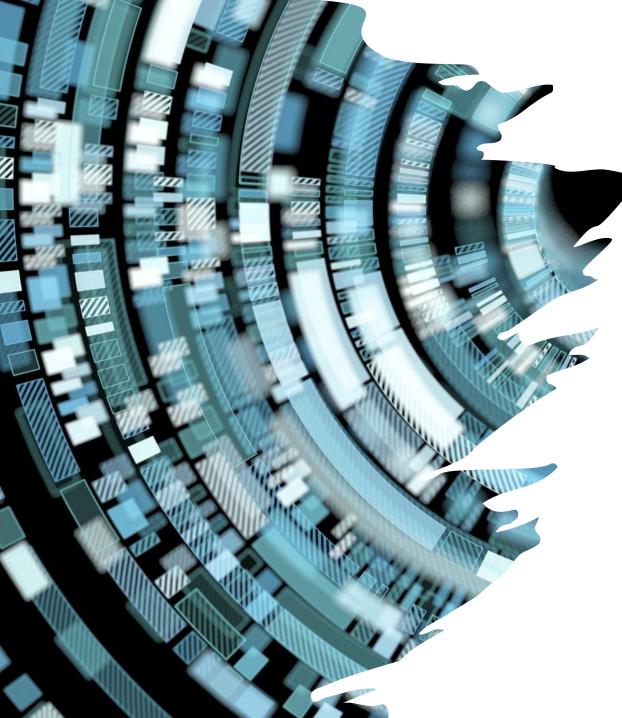






Data Discovery

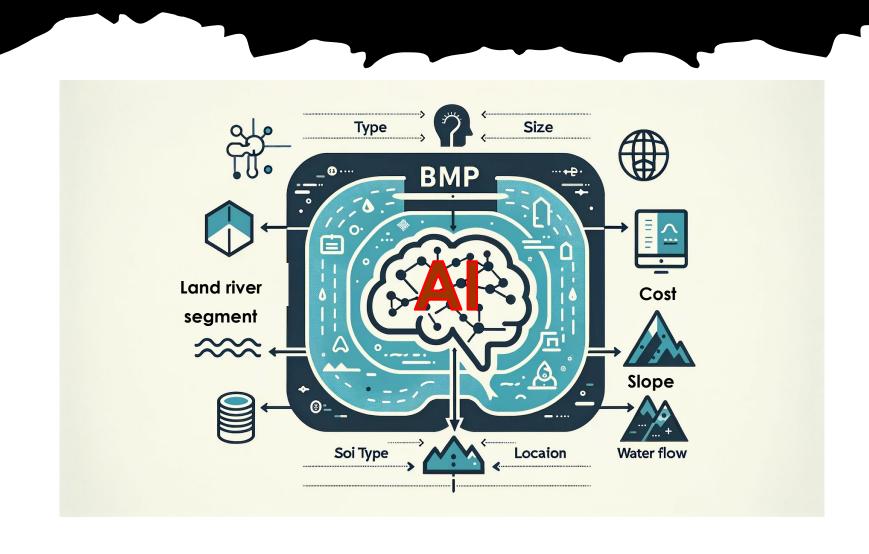
Employ artificial intelligence methods to recognize the relationships between clustering, variable combinations, land river segments, and BMP selection.



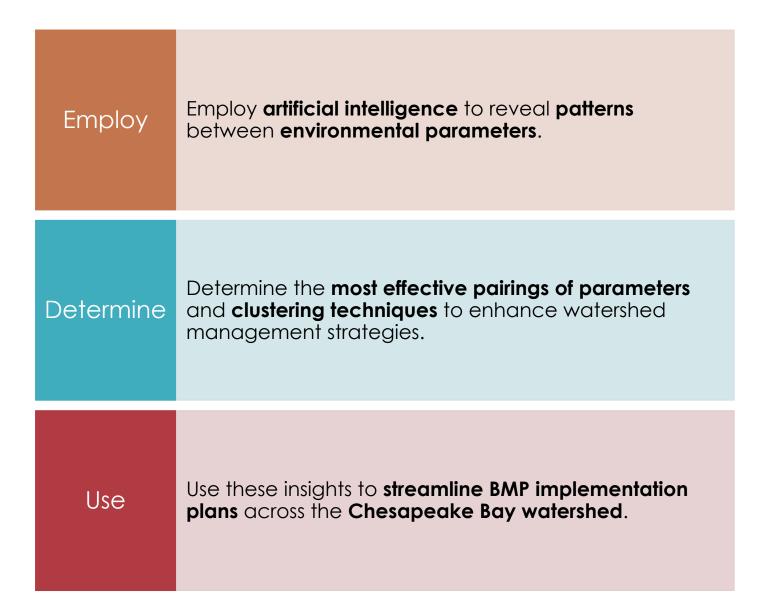
Artificial Intelligence

- Machine Learning Approaches (e.g., Random Forests, support vector machines, deep learning).
- Optimization and Search Algorithms (e.g., Genetic Algorithms).
- Bayesian Networks and Probabilistic Models (e.g., Markov chain Monte Carlo and Gaussian Processes).
- **Decision Support Systems** (e.g., Automated Decision Making and Cognitive Computing).

Connecting BMP and Land River Segment Characteristics



Connecting BMP and Land River Segment Characteristics



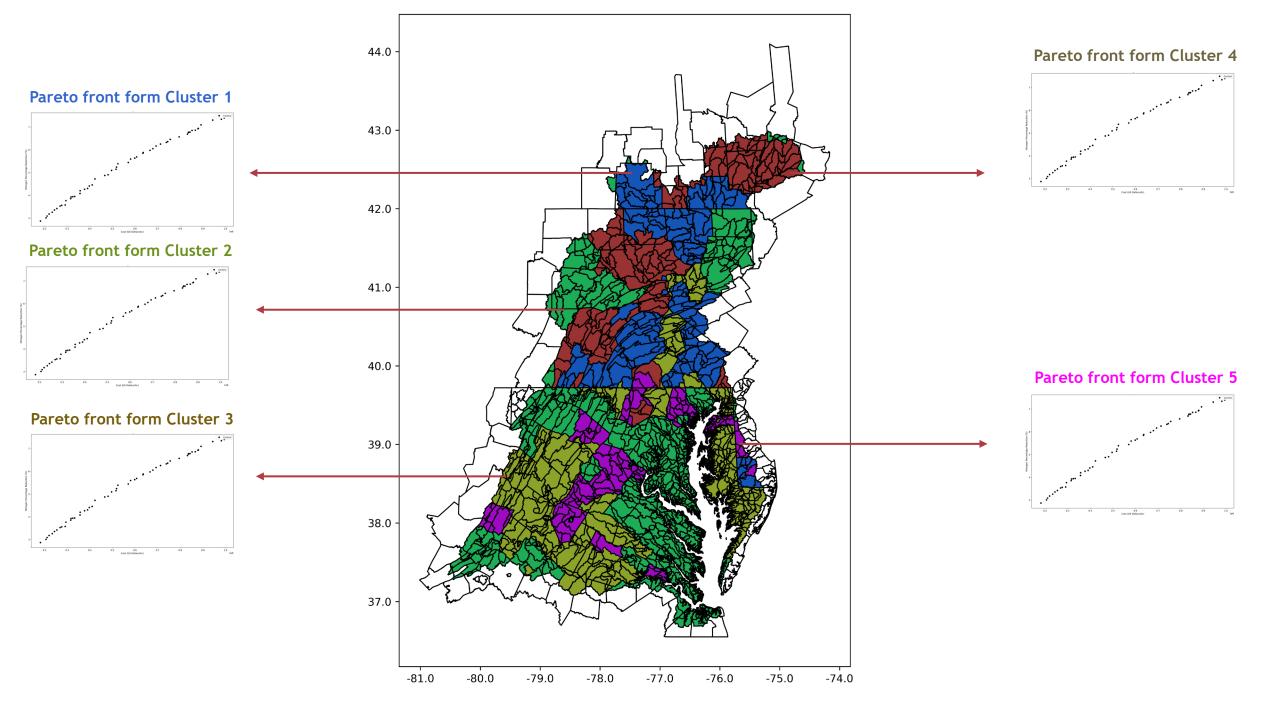


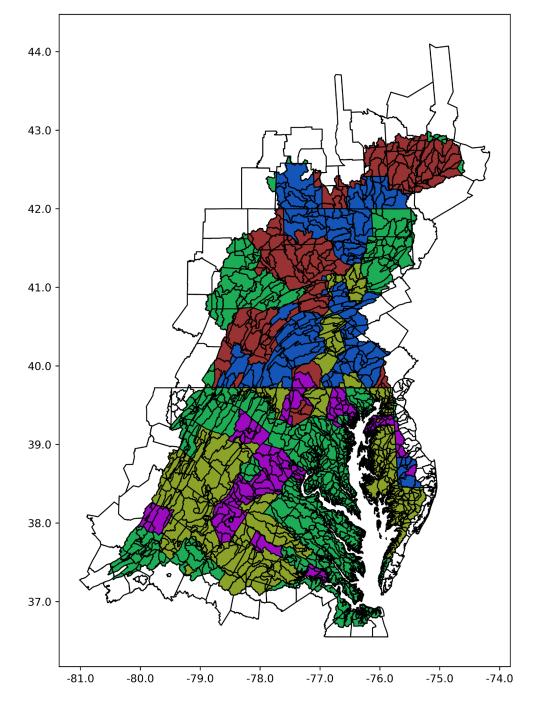
Step 7: Test and Validation

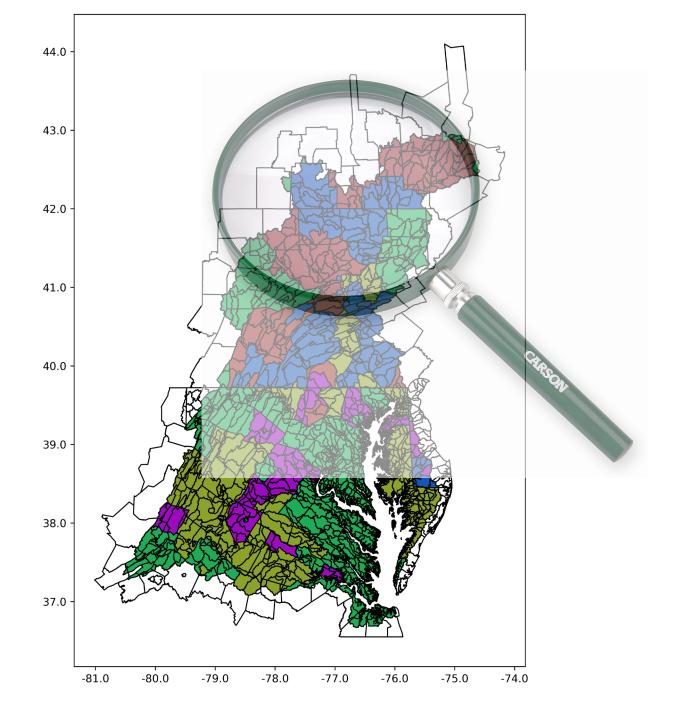


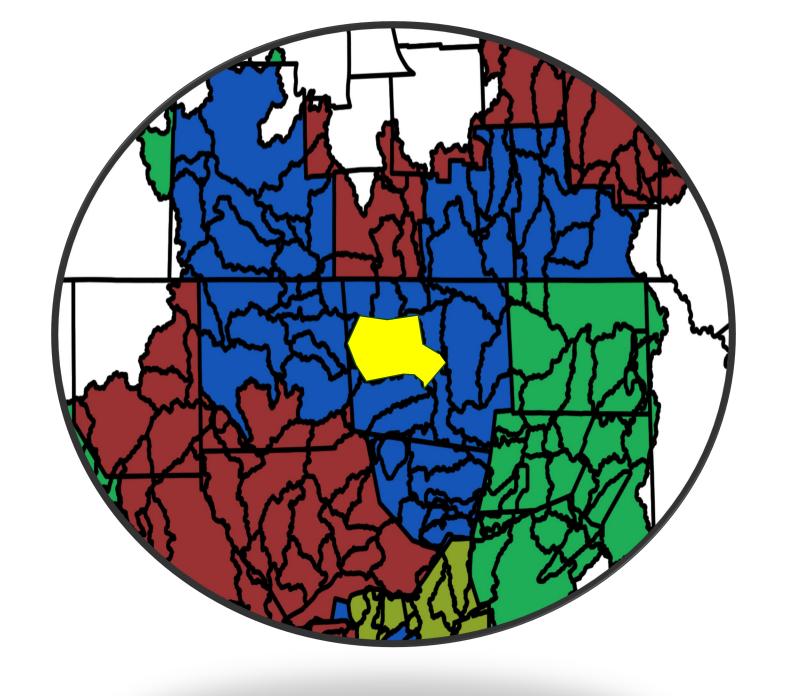
Test & Validation

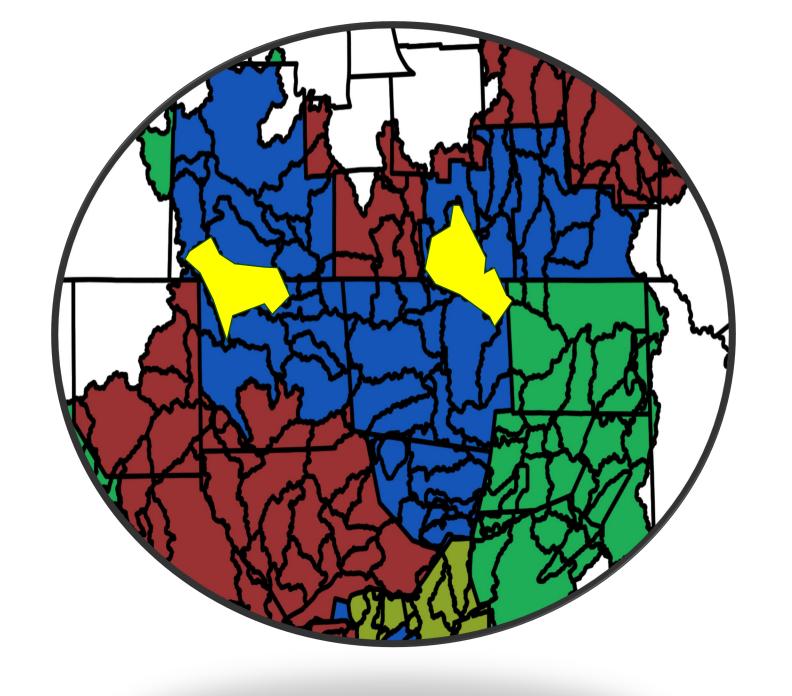
The outcomes from this section will be tested in different areas within the Chesapeake Bay watershed to confirm their reliability and applicability.











Interactive Optimization & Decision making



Scenario Info

SCENARIO NAME: JEFFERSON

COUNTIES: JEFFERSON, WV

BASE SCENARIO: NO BMP

BASE CONDITION: 2019 14

SCENARIO TYPE: OFFICIAL BMPS

COST PROFILE: WATERSHED

HISTORICAL CROP NEED SCENARIO: 6608

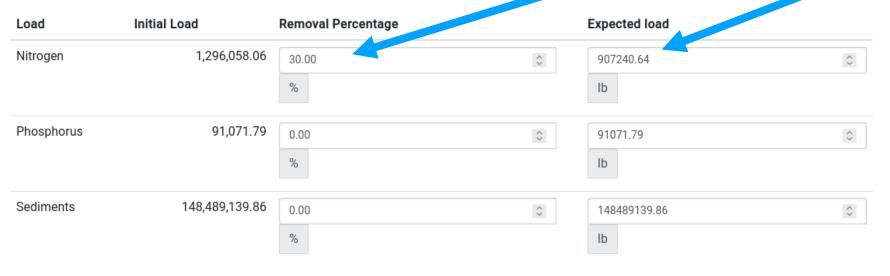
POINT SOURCE DATA SET: NO ACTION

The users can fine-tune their expected load targets, BMPs, costs, or constrain the use of BMPs using a 6 steps input flow. The first step shows information regarding the scenario.



The user can indicate a maximum removal percentage for loads. The user can modify any of the percentage or the actual load.

Loads



TOTAL DOLLARS FOR BMP IMPLEMENTATION:										
TOTAL ACRES ALLOCATED TO BMP:	145,790.88 ACRES	100.00	\$	%	145790.88	\$	Acres			



Controls to select BMPs

Save Selected BMPs

BMP Selection

Available BMPs

Search...

Agriculture BMPs

Alternative Crops
Animal Waste Management System

Barnyard Runoff Control

Biofilters

Cover Crop Traditional Annual Legume Early Aerial

Cover Crop Traditional Annual Legume Early Drilled

Cover Crop Traditional Annual Legume Early Other

Cover Crop Traditional Annual Legume Normal Other

Cover Crop Traditional Annual Ryegrass Early Aerial

Cover Crop Traditional Annual Ryegrass Early Drilled

Cover Crop Traditional Annual Ryegrass Early Other

Cover Crop Traditional Annual Ryegrass Normal Drilled

Cover Crop Traditional Annual Ryegrass Normal Other

Cover Crop Traditional Barley Early Aerial

Cover Crop Traditional Barley Early Drilled

Cover Crop Traditional Barley Early Other

Cover Crop Traditional Barley Normal Drilled

Cover Crop Traditional Barley Normal Other

*
*
*
*
*

Selected BMPs

Search..

Agriculture BMPs

Agricultural Stormwater Management

Broiler Mortality Freezers

Cover Crop Commodity Early

Cover Crop Commodity Late

Cover Crop Commodity Normal

Cover Crop Traditional Annual Legume Normal Drilled

Forest Buffer

Forest Buffer - Narrow

Forest Buffer Nitrogen Upland Acres

Forest Buffer Phosphorus and Sediment Upland Acres

Forest Buffer-Narrow with Exclusion Fencing

Forest Buffer-Streamside with Exclusion Fencing

Forest Buffer-Streamside with Exclusion Fencing Nitrogen Upland

Forest Buffer-Streamside with Exclusion Fencing Phosphorus and

Grass Buffer

Grass Buffer - Narrow

Grass Buffer Nitrogen Upland Acres

Animal BMPs

Animal Waste Management System

Riofilter

Dairy Precision Feeding and/or Forage Management

Lagoon Covers

Mortality Composters

Poultry Litter Amendments (alum, for example)

Poultry Nutrient Reduction

Riparian Fence

Manure Treatment BMPs

Manure Transport

Efficiency BMPs

Land Conversion BMPs

Animal BMPs

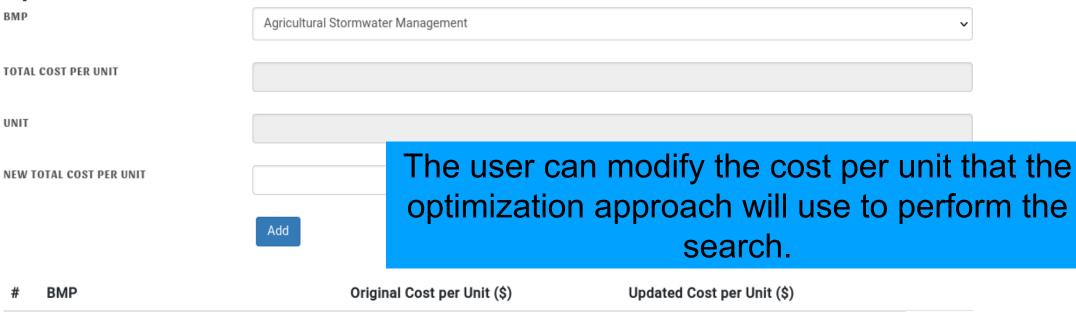
Manure Transport

When we detect that the user has selected Land Conversion, Animal or Manure BMPs, the system triggers specific optimization procedures regarding these BMPs



Update BMP Total Cost Per Unit

Cover Crop Commodity Early

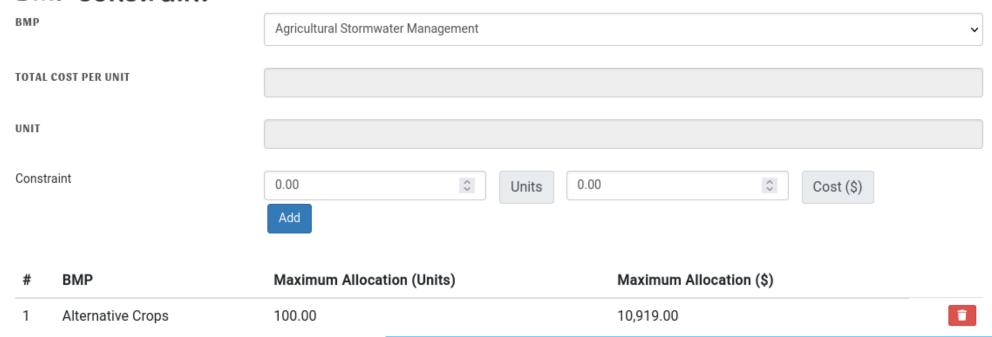


10.0

77.87



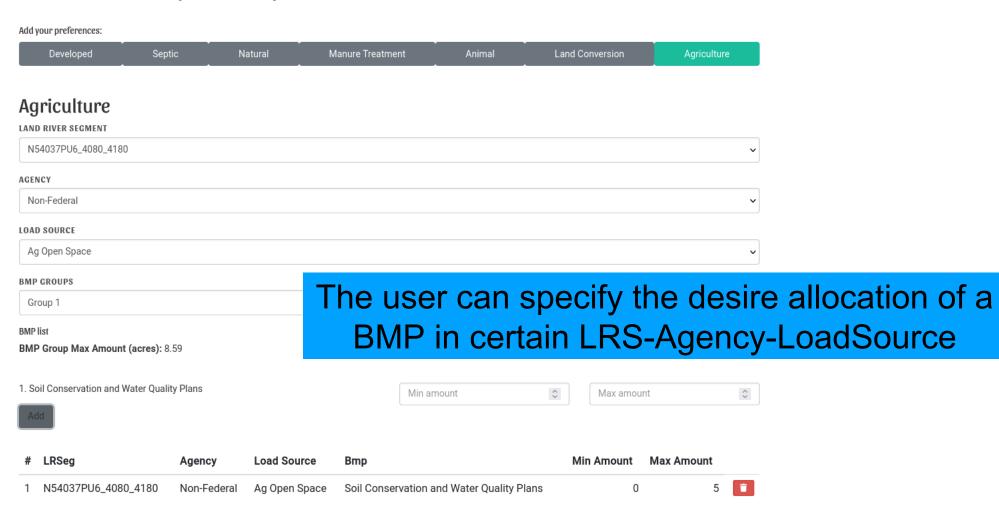
BMP Constraint



The user can indicate the maximum allocation of BMPs



BMP Constraint (Advanced)



Executions (Results)

+ Search Clear

ld	Cost	Nitrogen	Phosphorus	Sediments	
205	\$ 635,868	1,186,825 lbs	72,233 lbs	134,106,494 lbs	
206	\$ 1,120,260	1,164,530 lbs	72,036 lbs	132,555,760 lbs	
207	\$ 718,151	1,179,864 lbs	74,280 lbs	134,535,442 lbs	
208	\$ 1,167,	The optim	ization ap		
209	\$ 724,83	finds a set	of scena		
210	\$ 784,74	single exe	cution. Th		
211	\$ 1,314,	can explo	ore the ob		
212	\$ 543,83	S	olutions.		
213	\$ 819,504	1,168,267 lbs	74,826 lbs	134,406,563 lbs	
214	\$ 482,654	1,195,935 lbs	79,396 lbs	139,129,245 lbs	
215	\$ 1,006,390	1,164,717 lbs	72,440 lbs	132,520,354 lbs	
216	\$ 702,553	1,184,550 lbs	78,024 lbs	137,501,963 lbs	

BMP Results BMP

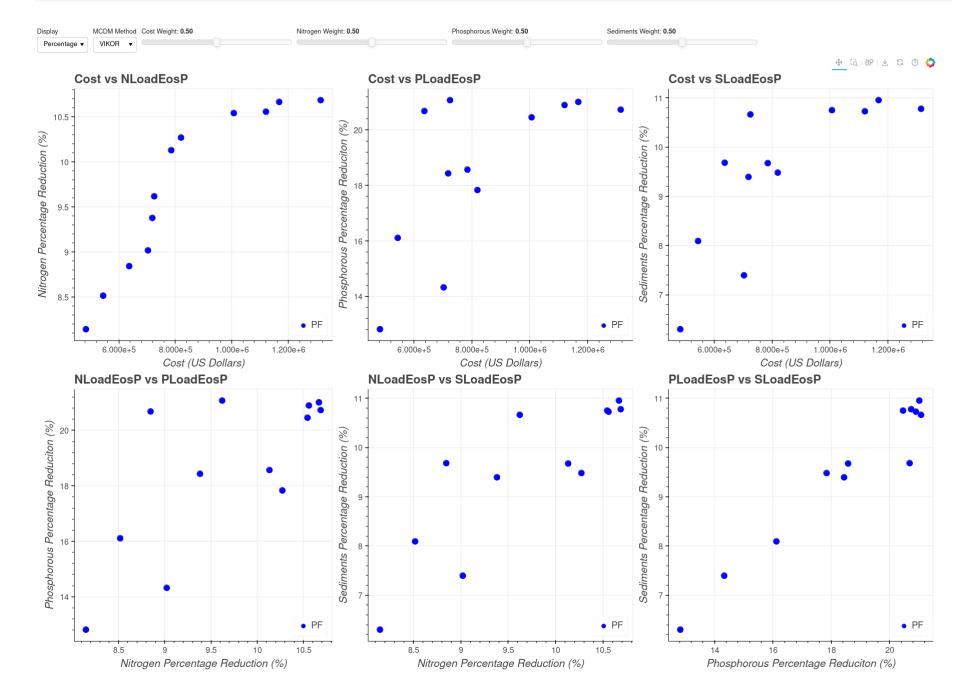
BMP Cost Mode Min Amount Max Amount Acres Off Stream Watering Without 12,827.78 12,956.06 ACRES $\hat{\ }$ $\hat{\ }$ 0 12,827.78 Fencing COST Nutrient Management Plan 0.00 0.00 ACRES 0 $\hat{\ }$ $\hat{\ }$ 0.00 O COST Cover Crop Traditional Rye Early 897.37 69,878.26 ACRES 0 $\hat{\ }$ $\hat{\mathbf{v}}$ 897.37 Drilled O COST Nutrient Management N Timing 6.45 $\hat{\ }$ It is possible to inspect the results, and to set up new Nutrient Management N Rate boundaries for BMPS to use in Barnyard Runoff Control $\hat{\ }$ future executions. Nutrient Management Plan High 26,538.60 52,811.82 ACRES 0 $\hat{\ }$ 26.538.60 $\hat{\ }$ Risk Lawn COST **Denitrifying Ditch Bioreactors** 6,010.65 193,422.72 ACRES 0 **\$** 6,010.65 $\hat{\ }$ COST

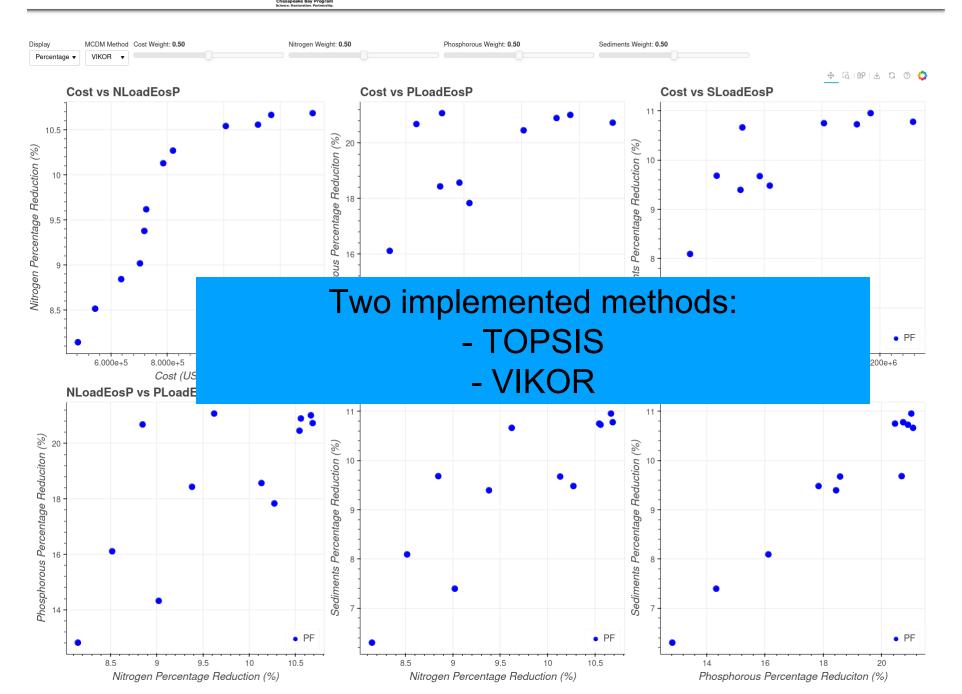
SEARCH:

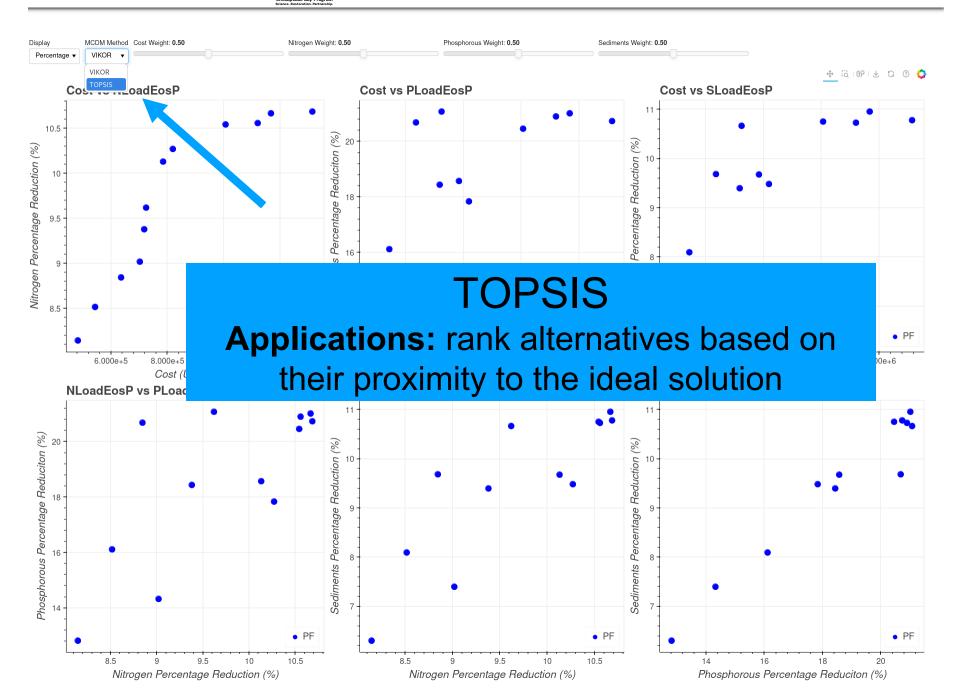
Detailed BMP Results

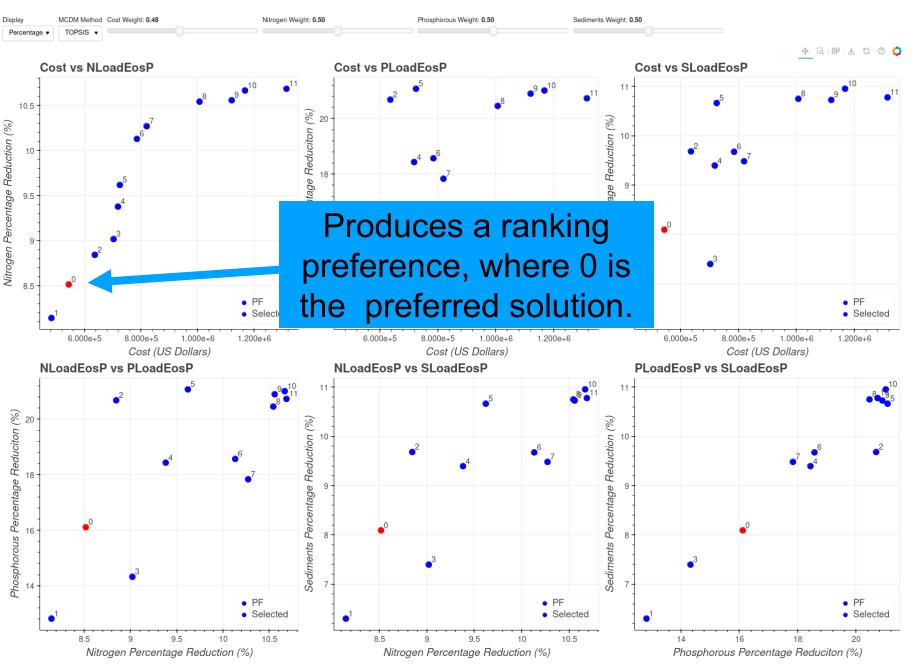
#	Land-River-Segment	Agency	Load Source	Bmp	Acres	Cost (\$)	Mode	Min Amount	Max Amount
1	N54037PU6_4180_4150	Non- Federal	Grain without Manure	Denitrifying Ditch Bioreactors	1,300.10	41,837.26	• FIXED • ACRES • COST	0 0	1,300.1 🗘
2	N54037PU2_3900_3750	Non- Federal	Grain without Mar	Denitrifying Ditch Users	50.66	1,630.32	⊚ FIXED ○ ACRES	0 0	50.66 🗘
3	N54037PU6_3750_3752	Non- Federal	Man	oscis opose onstrai	d BMI	P, and	d set	up	259.31 🗘
4	N54037PU6_4080_4180	Non- Federal	Grain without Manure	Denitrifying Ditch Bioreactors	123.85	3,985.36	FIXED ACRES COST	0 0	123.85 🗘
5	N54037PU2_4220_3900	Non- Federal	Grain without Manure	Denitrifying Ditch Bioreactors	1,299.34	41,812.71	• FIXED • ACRES • COST	0 0	1,299.3 🗘
6	N54037PU6_3752_4080	Non- Federal	Grain without	Denitrifying Ditch	8.27	266.27	FIXEDACRES	0 0	8.27 🗘

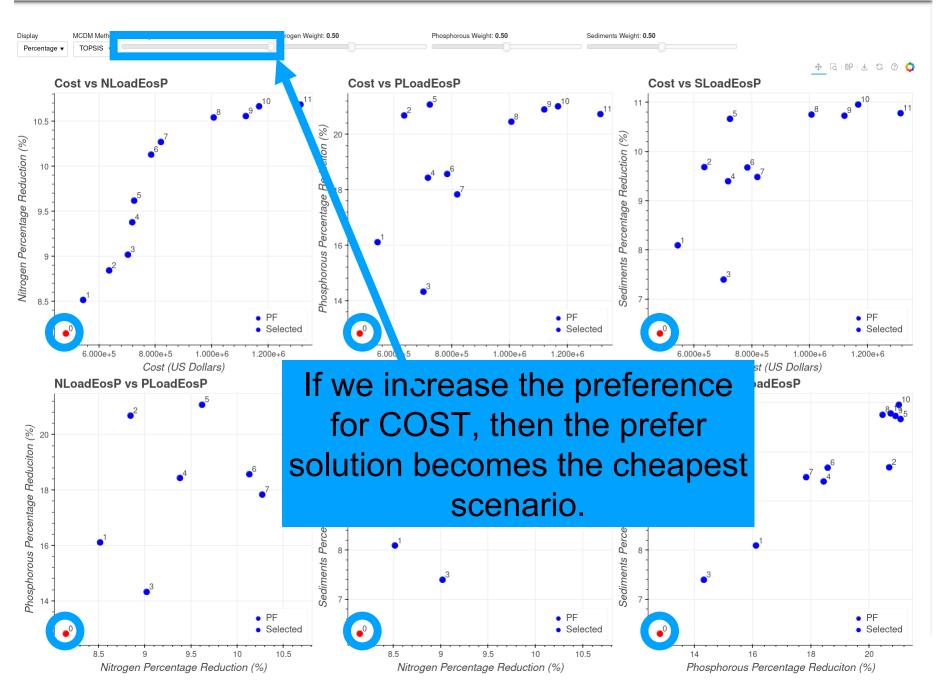
SEARCH:

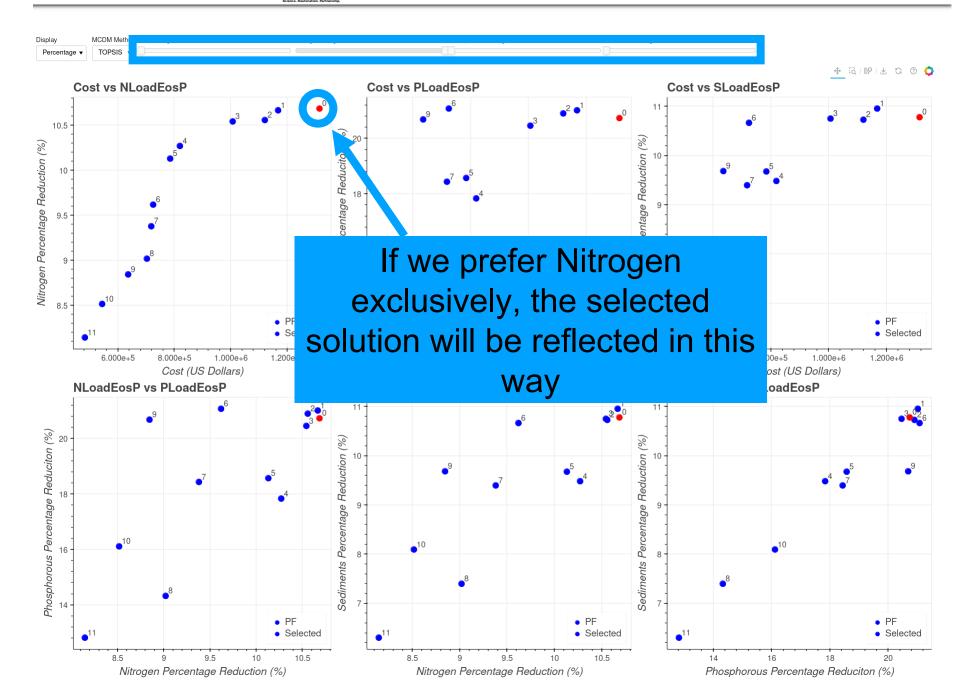


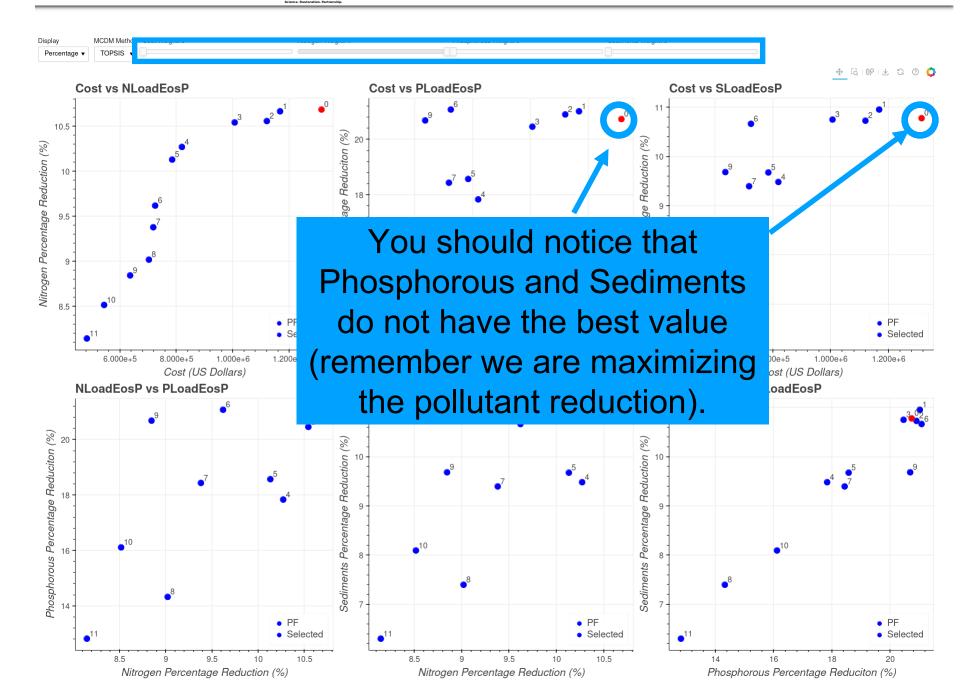


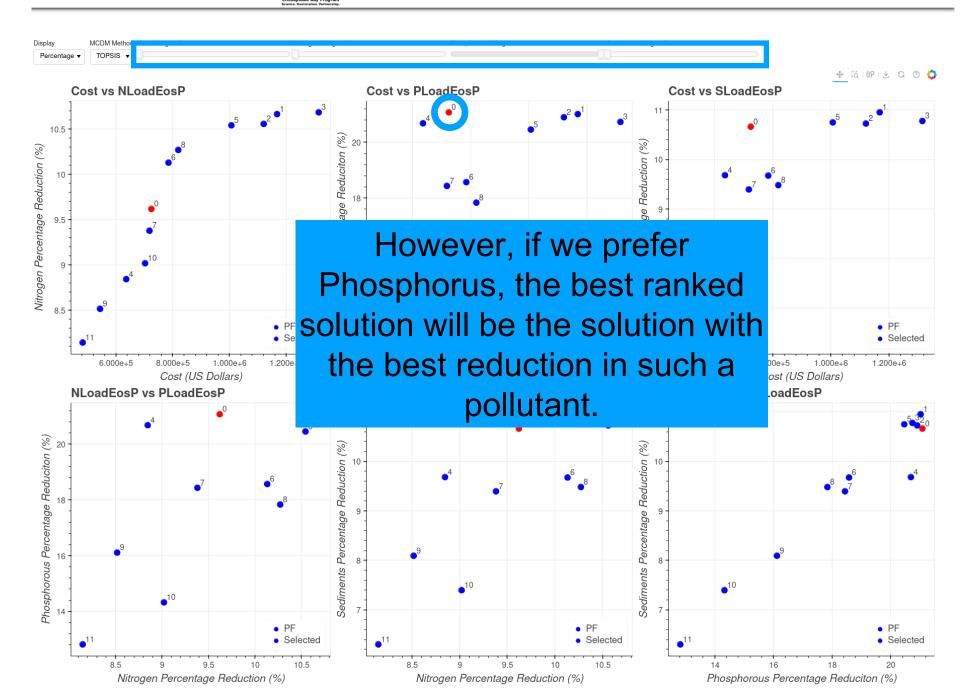


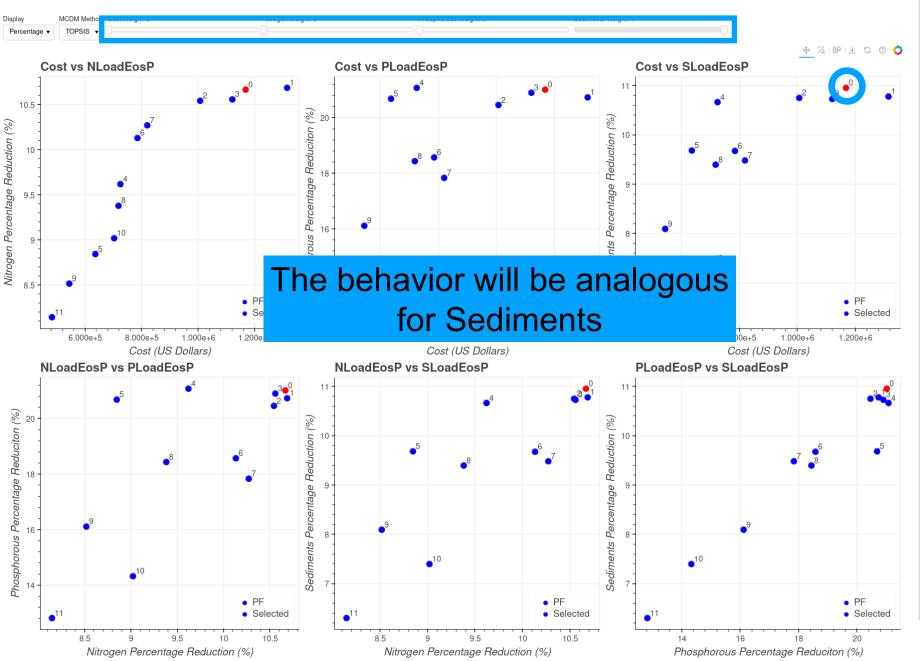


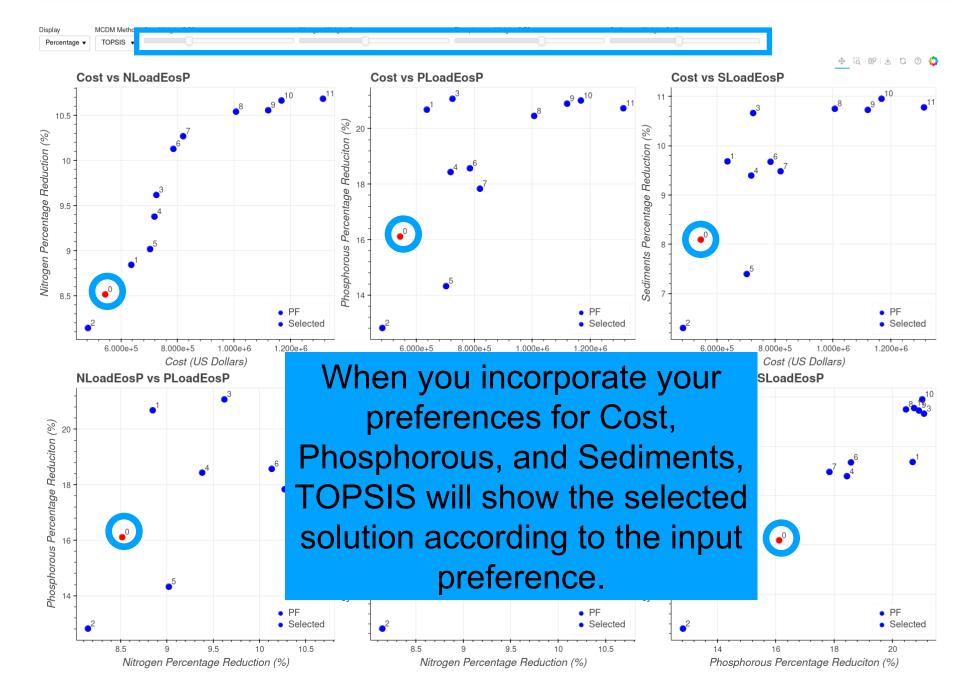


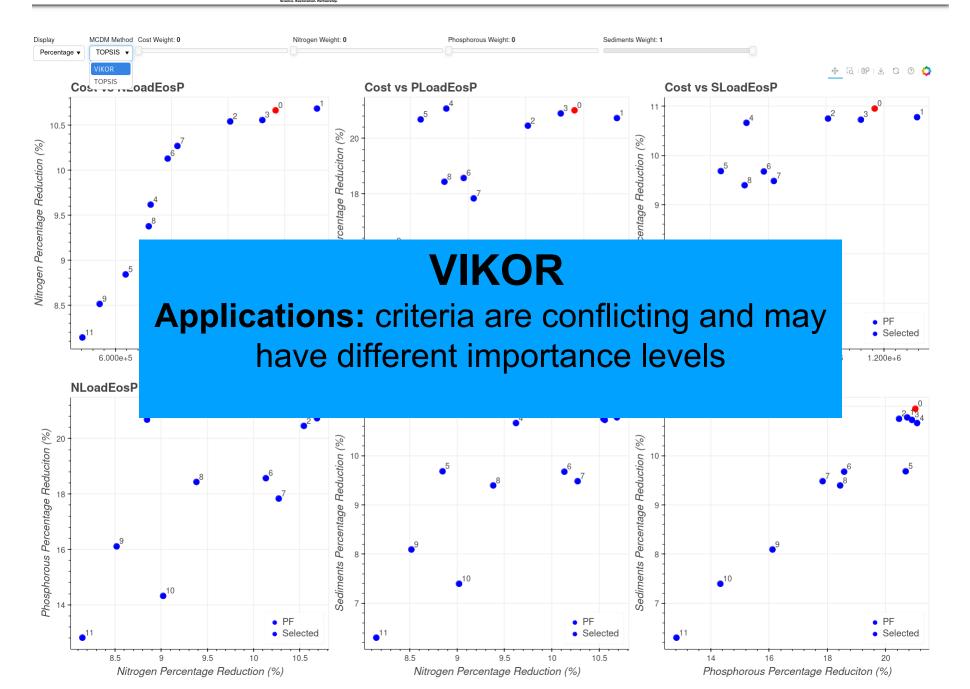


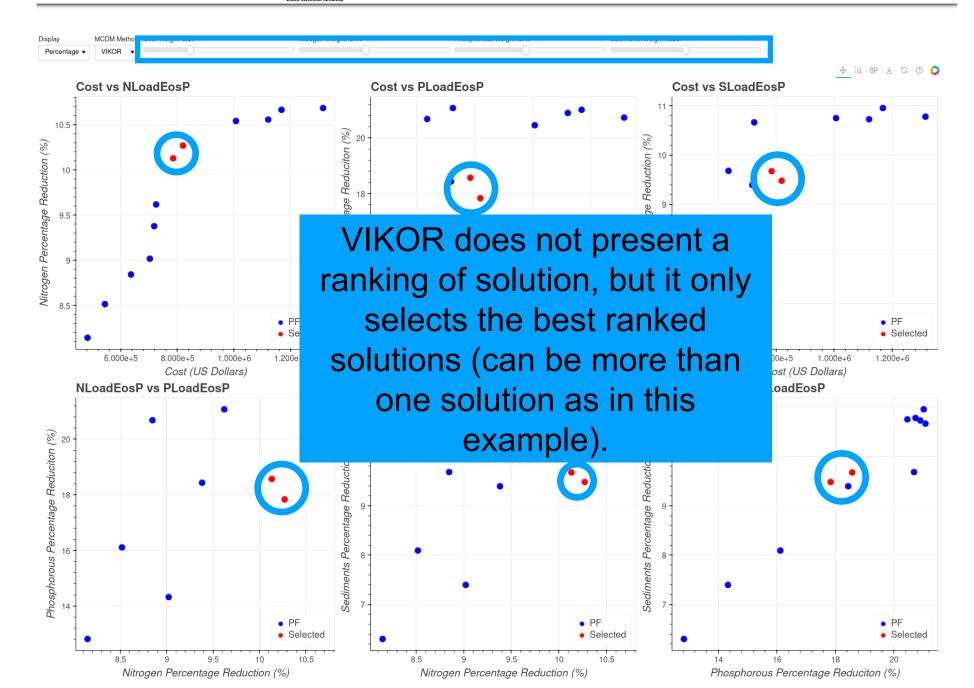


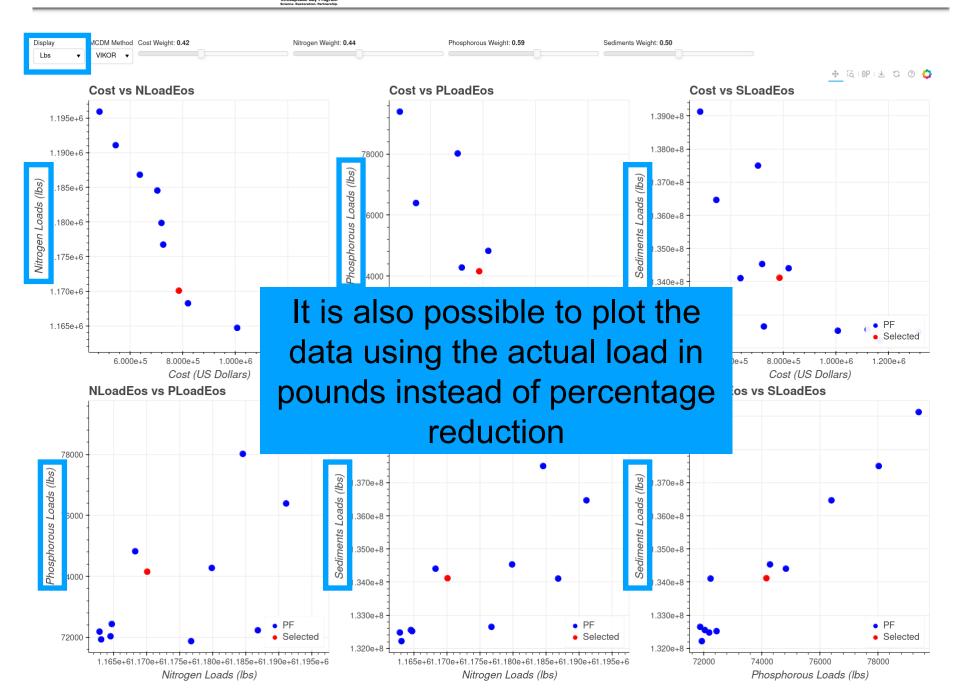












NEXT STEP

- Using artificial intelligence to enhance optimization
- Extensions to more counties and states
- Validation of our results on some critical county/state cases
- Parallel computing platform for faster execution
- Uncertainty and other practicality handling
- Workshops with CBP users for feedback and improvement of our approaches

Chesapeake Bay Optimization Webinars

Join experts in the discussion of bay preservation.





2024 Chesapeake Bay Optimization Webinars



Objective:

* Raise awareness about our BMP optimization framework.



Features:

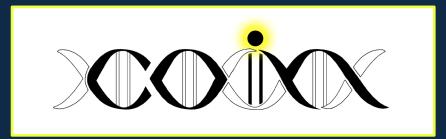
* Interactive show of the framework capabilities



Benefits:

- * Enhanced decision-making
- * Foster community collaborations





Computational Optimization and Innovation

Thankyou



