Optimization Update: ***Transfer Learning***

Kalyanmoy Deb, Pouyan Nejadhashemi, Gregorio Toscano, and Hoda Razavi

Agenda

- Our current status
- "Innovization" concept for knowledge discovery
- Re-optimization using learned knowledge
- Transfer Learning to solve similar other problems quickly
- Results on above concepts on some counties
- Conclusions

Timeline of the Project

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

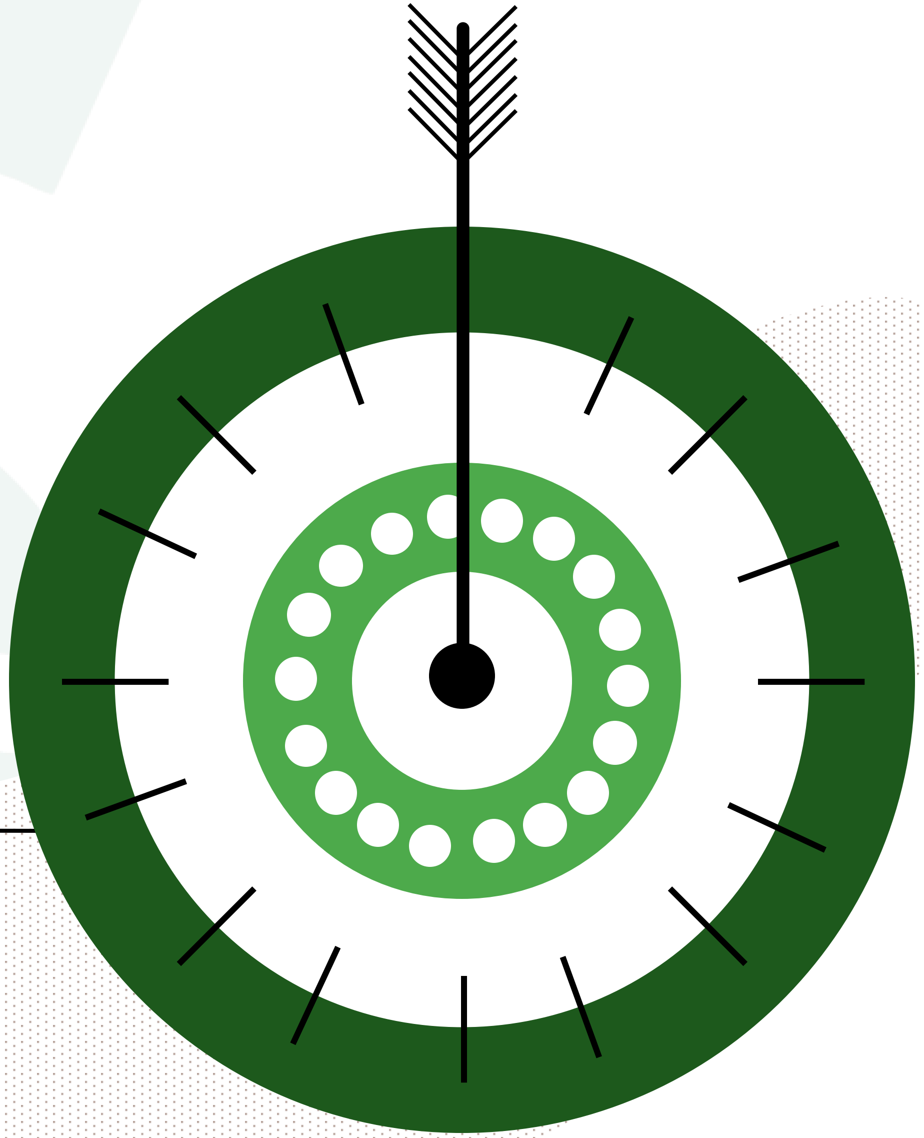
3.3 “Innovization” approach for improving scalability

3.4 Distributed computing approach for improving scalability

Timeline of the Project

Research Goal

To reduce search space for optimization by creating new knowledge based on physical characteristics of a study area.

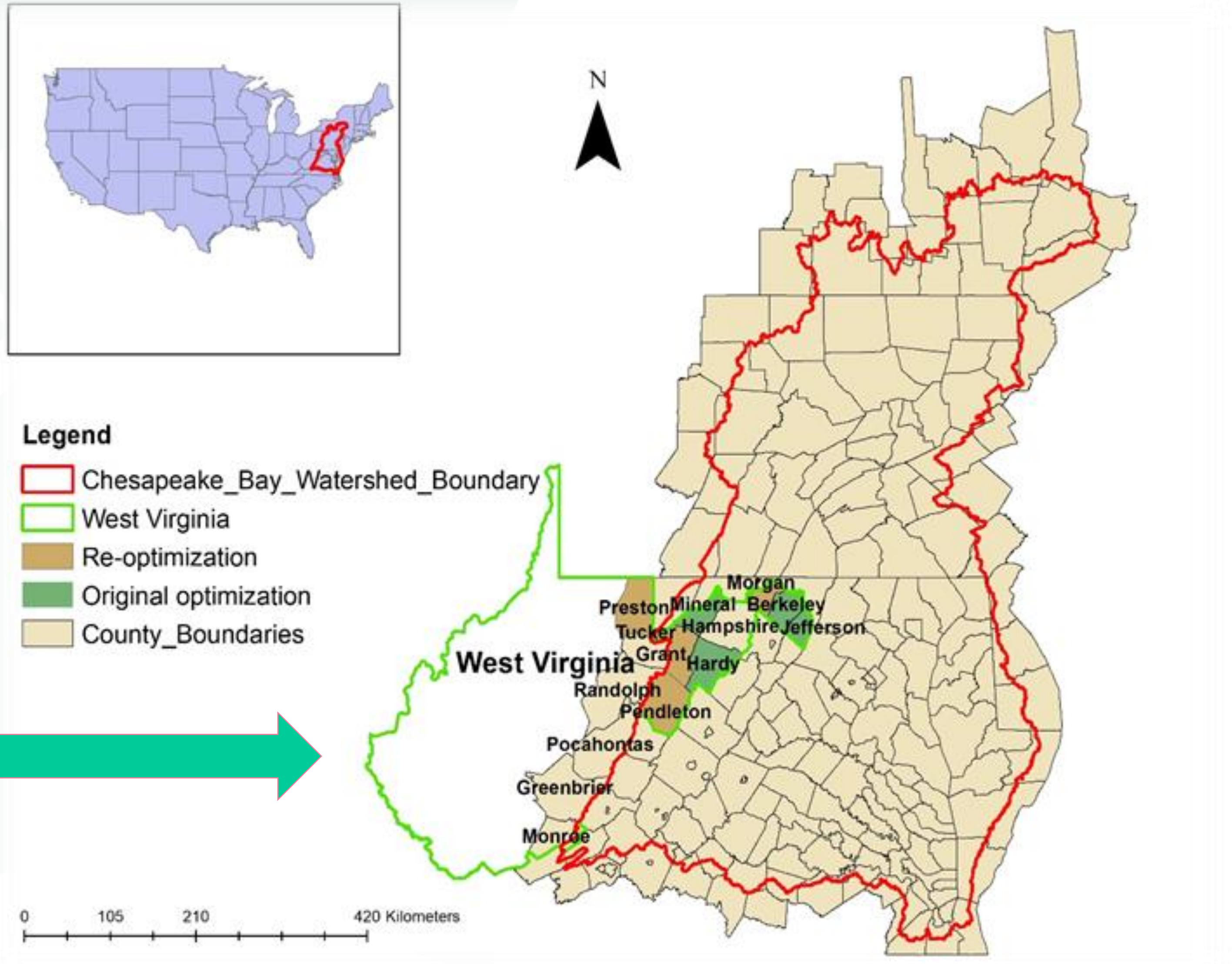


MICHIGAN STATE
UNIVERSITY

Methodology

Chesapeake Bay Watershed

- The **largest estuary** in the United States and the **third-largest** in the world.
- Over **300 counties**
- **WEST VIRGINIA**



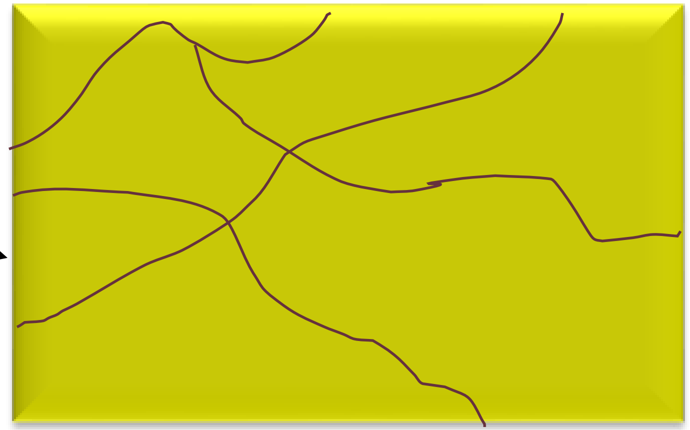
Problem Significance (Search Space Complexity)

- **Let's talk about one county out of 300**

County

Problem Significance (Search Space Complexity)

- **<u>Each county</u>** on average has **10 land-river segments**,



MICHIGAN STATE UNIVERSITY

Problem Significance (Search Space Complexity)

- **Each county** on average has **10 land-river segments,**
- **Each landriver segment can accept up to 200 BMPs options**

Problem Significance (Search Space Complexity)

- Each county on average has **10 land-river segments**,
- Each landriver segment can accept up to 200 BMPs options (i, e, variables)
- **Each landriver segment has 50 land use options**
- **Making the overall search space for a county to have about $50^{2,000}$ different management options.**
- **Highly constrained:** ~more than **100,000 constraints.**

Overview of the Methodology

Watershed Modeling (CoreCAST)
Evaluation of BMP impacts on water quality

Multi objectives optimization framework (Pareto Front)
- Cost of BMP implementation
- BMP Nitrogen reduction (Environmental efficiency)

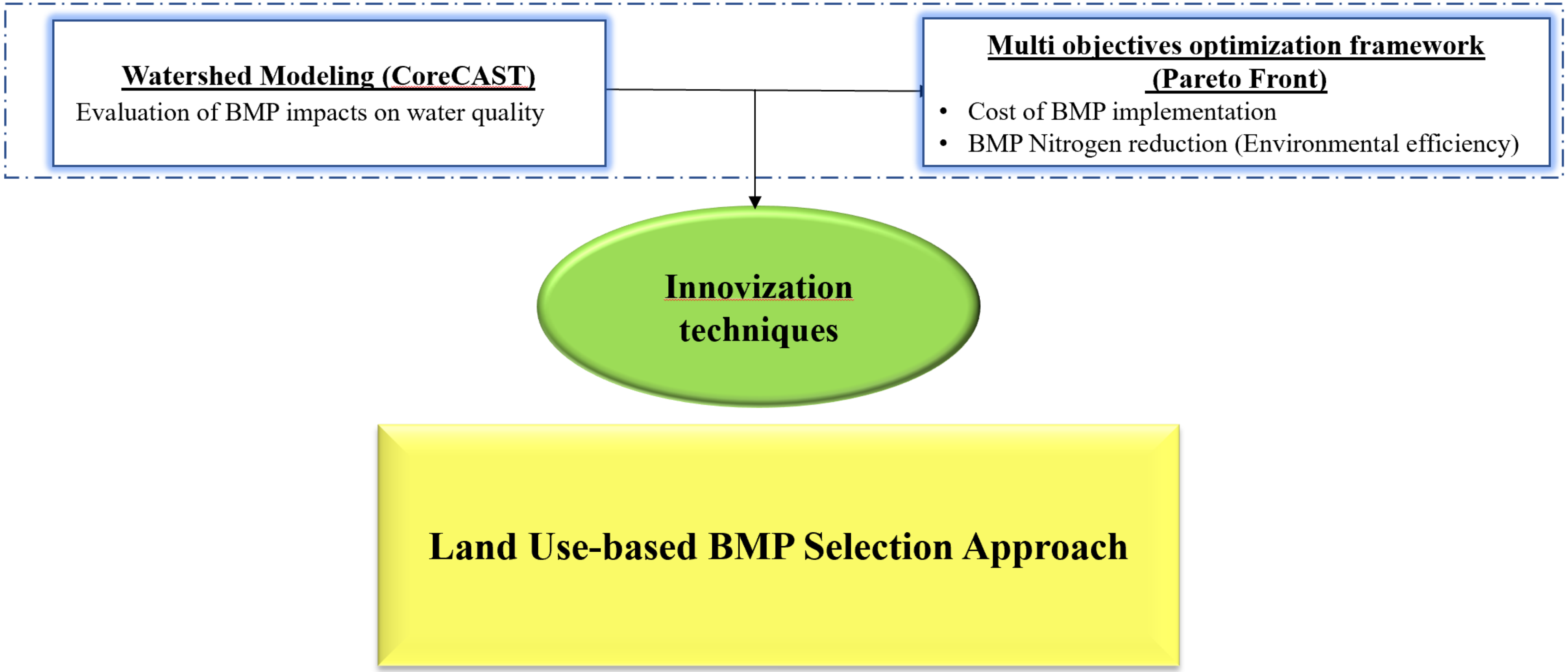
Innovization techniques

Innovation through Optimization

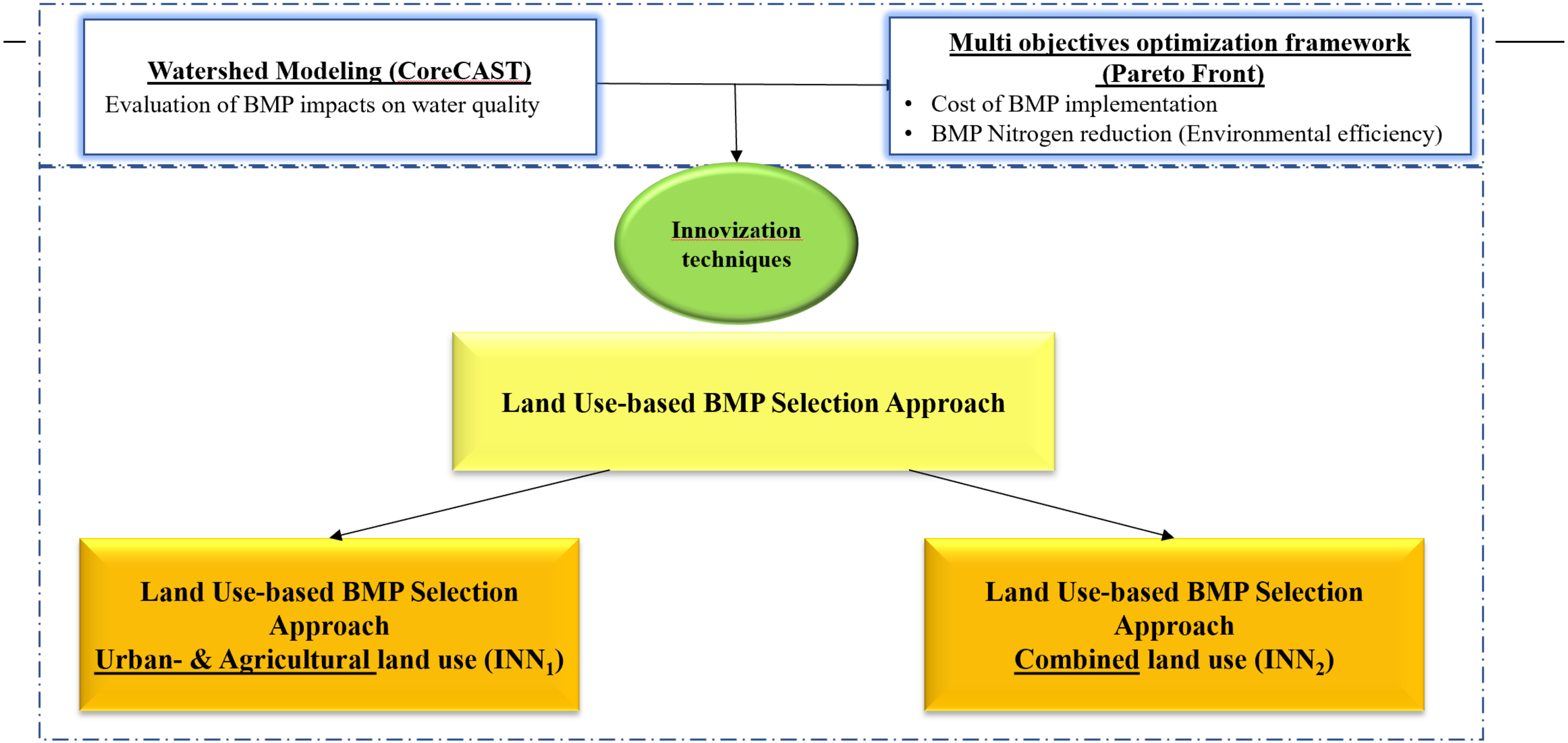
- **What is innovization?**

Create new knowledge from optimization solutions

Overview of the Methodology



Overview of the Methodology



Methodology

The Land Use-based BMP Selection Approach (Innovization based on the Urban- and Agricultural-dominated Land Use Characteristics) (INN1)

We hypothesize that since urban and agricultural-dominated areas are basically different in terms of pollution generation, fate, and transport. Therefore, different types of BMPs should be selected to address these challenges.

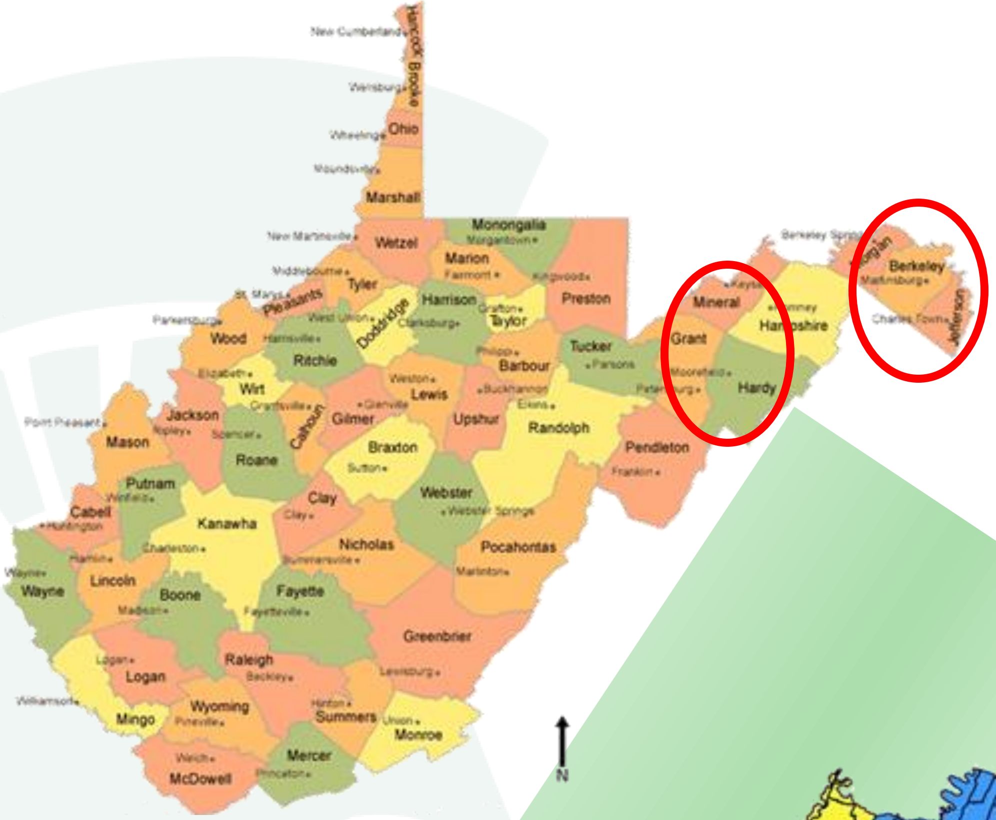
Methodology

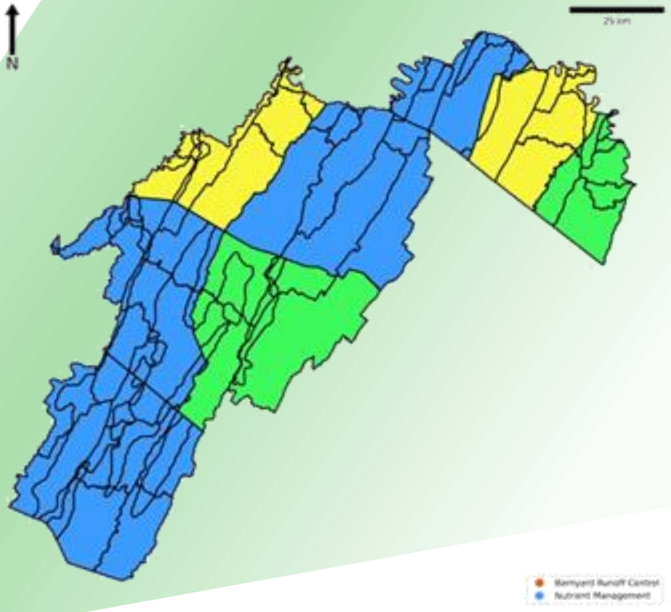
BMP Selection ranking strategy based on Land use:

In **West Virginia**, we identified the top two counties with the highest areas of urban and agricultural land uses.

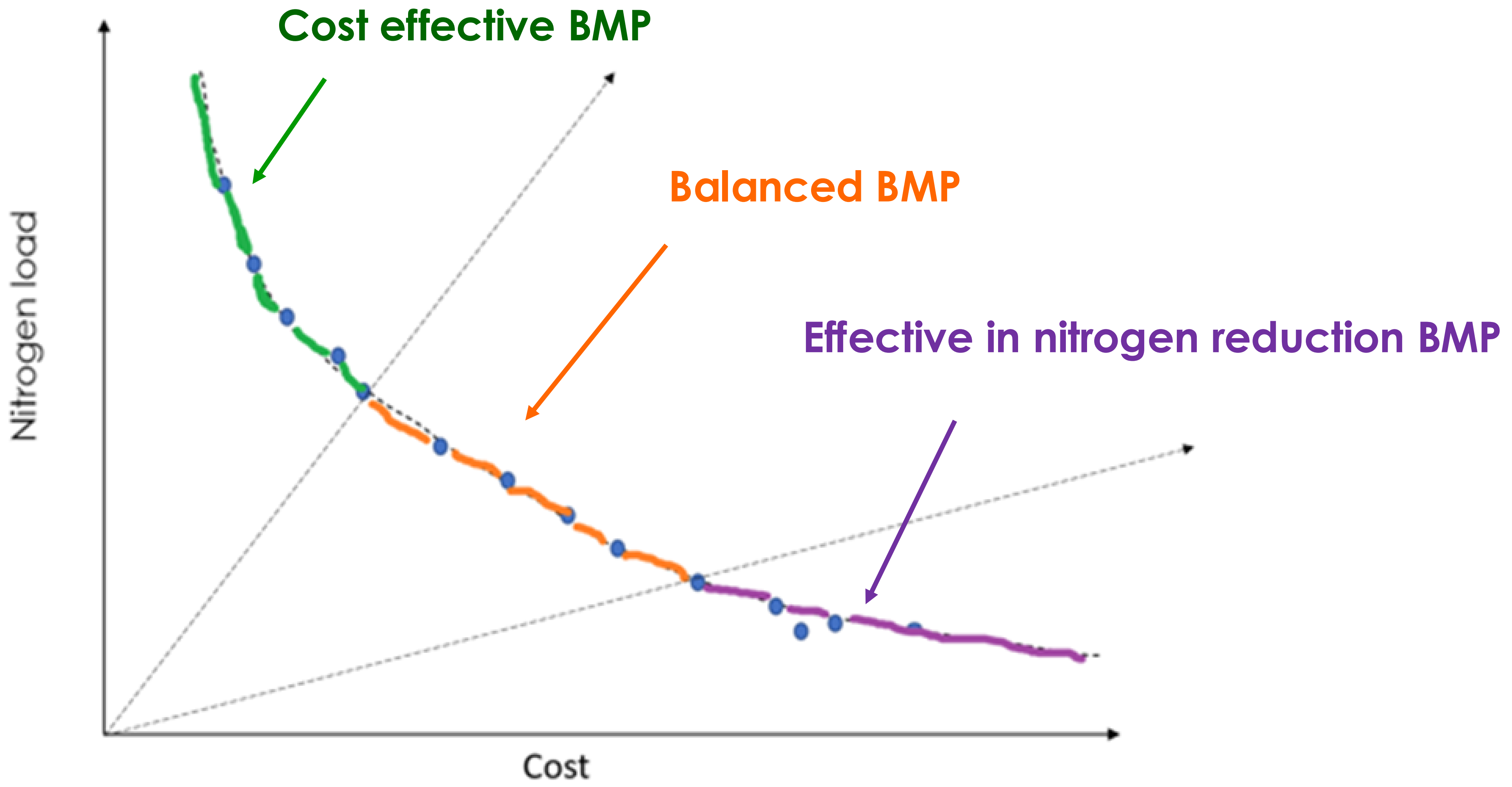
(Berkeley and Mineral): Urban dominated

(**Jefferson and Hardy**): Agricultural dominated





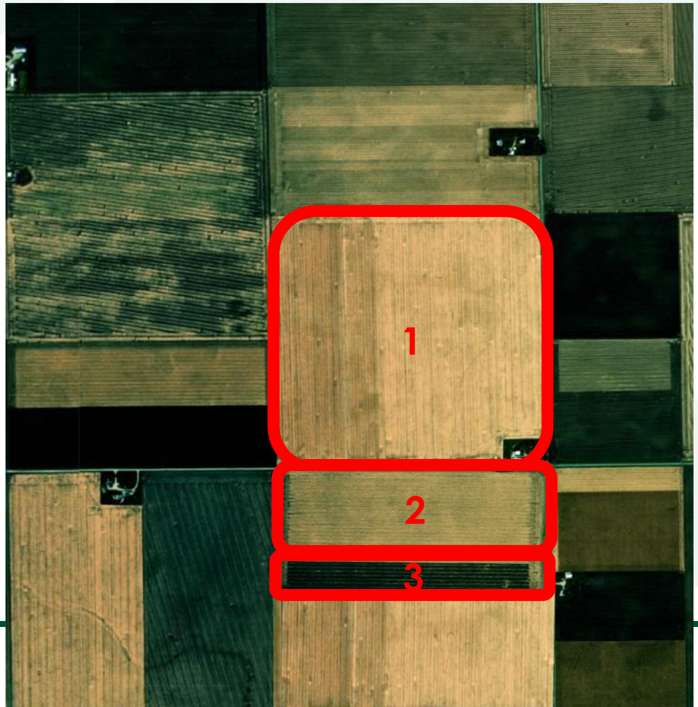
Pareto Front



Methodology

1 Examine different ranking methodologies to **identify the top BMPs**

- **Ranking strategy 1)** rank the top BMPs based on the **implementation acreages**;
- **Ranking strategy 2)** rank the top BMPs based on the percentage of **maximum allowable** acreages;
- **Ranking strategy 3)** rank the top BMPs based on the amount of **nitrogen reduction per dollar spent.**



MICHIGAN STATE UNIVERSITY

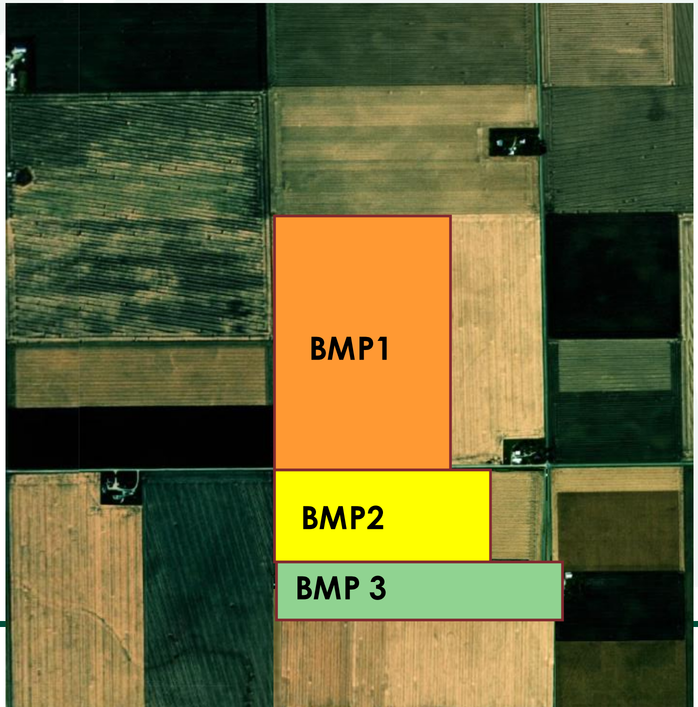
Methodology

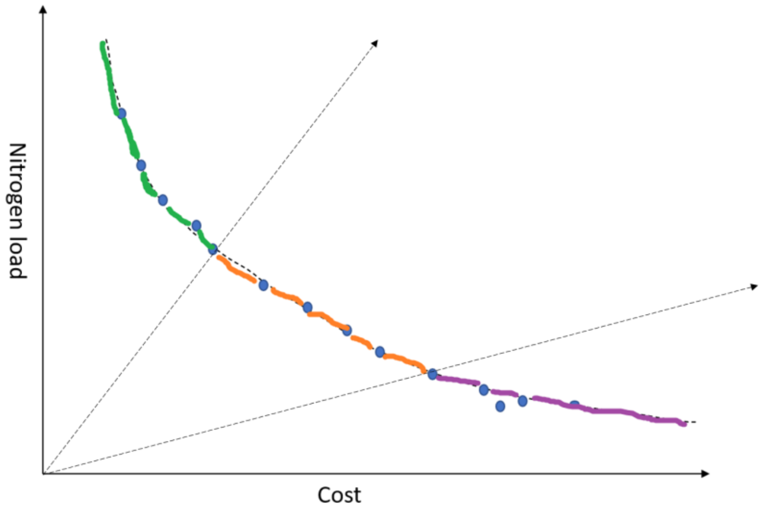
1 Examine different ranking methodologies to **identify the top BMPs**,

- **Ranking strategy 1) rank the top BMPs based on the implementation acreages;**
- **Ranking strategy 2)** rank the top BMPs based on the percentage of maximum allowable acreage
- **Ranking strategy 3)** rank the top BMPs based on the amount of nitrogen reduction per dollar spent.

Ranking strategy):

BMP1
BMP2
BMP3





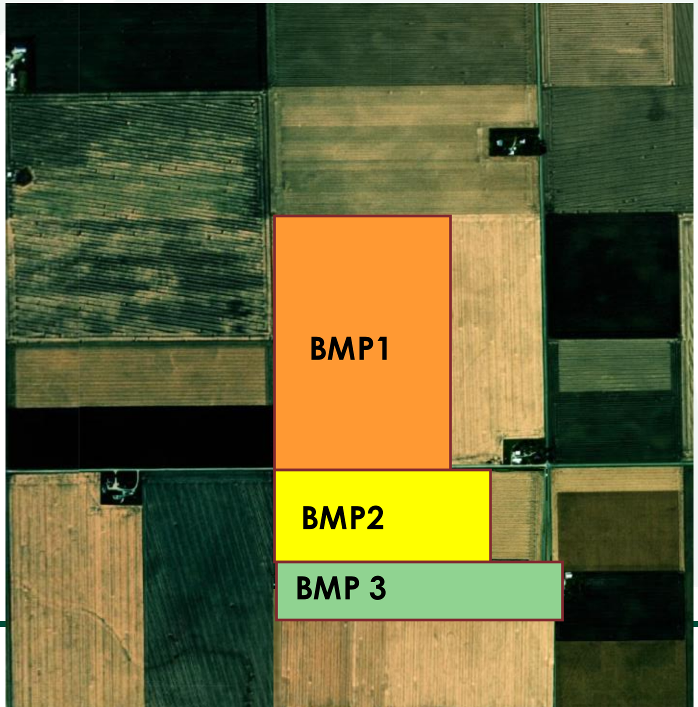
Methodology

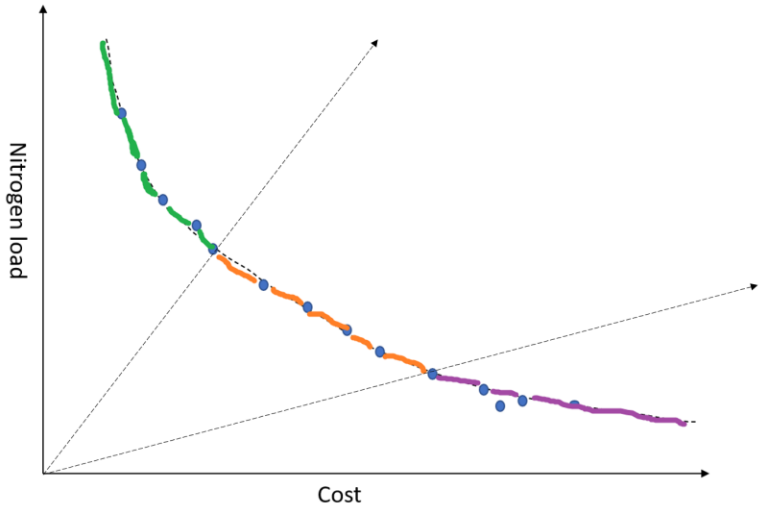
1 Examine different ranking methodologies to **identify the top BMPs**,

- **Ranking strategy 1) rank the top BMPs based on the implementation acreages;**
- **Ranking strategy 2) rank the top BMPs based on the percentage of maximum allowable acreages;**
- **Ranking strategy 3)** rank the top BMPs based on the amount of nitrogen reduction per dollar spent.

Ranking strategy 2):

BMP3
BMP2
BMP1





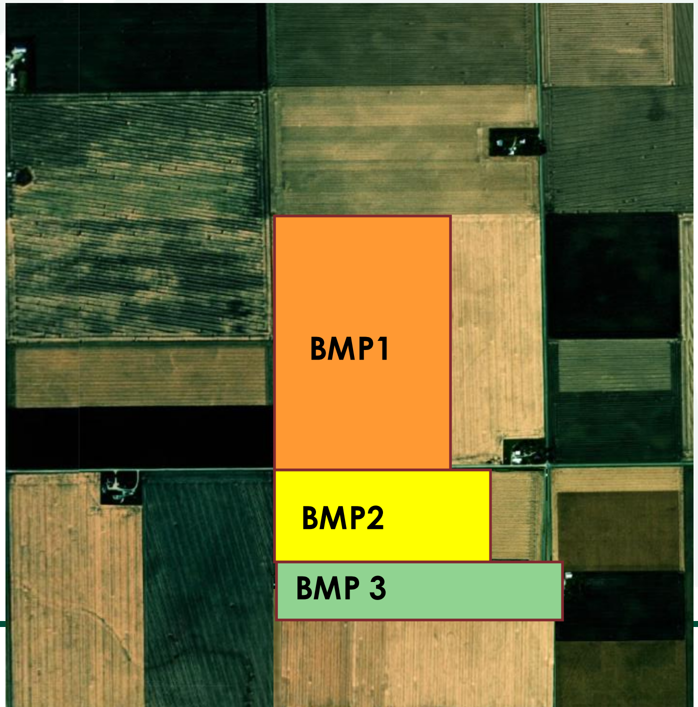
Methodology

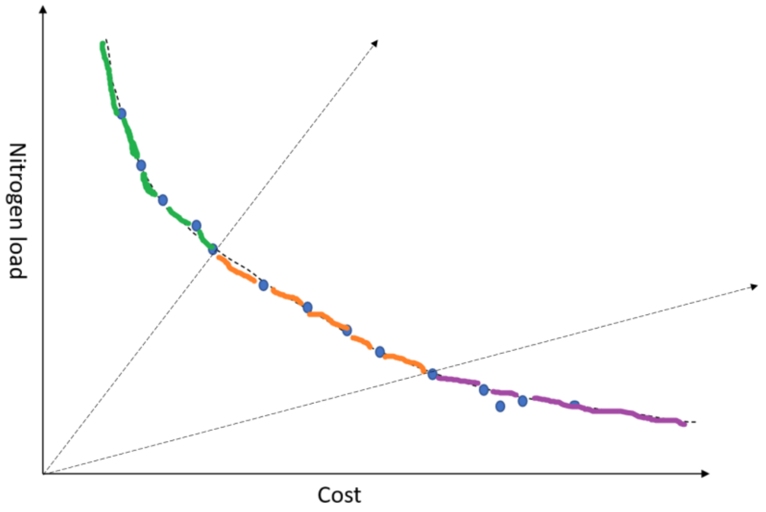
1 Examine different ranking methodologies to **identify the top BMPs**,

- **Ranking Strategy 1) rank the top BMPs based on the implementation acreages;**
- **Ranking Strategy 2) rank the top BMPs based on the percentage of maximum allowable acreages;**
- **Ranking Strategy 3) rank the top BMPs based on the amount of nitrogen reduction per dollar spent.**

Ranking methodology3):

BMP2 ($12/lb N)
BMP3 ($15/lb N)
BMP1($24/lb N)





MICHIGAN STATE UNIVERSITY

Methodology

Land Use-based BMP Selection Approach (Innovization based on the Combined Urban- and Agricultural-dominated Land Use Characteristics) (INN2)

Development of Strategies:

1. INN2-STG1: Combines BMPs with the **highest implementation acreages** from **both** land use types.
2. INN2-STG2 & INN2-STG3: Use information from strategies focusing on **maximum area of implementation** and **cost-effectiveness in nitrogen reduction**, respectively.
3. INN2-STG4: A hybrid strategy that combines BMPs from INN1-STG1 & INN1-STG3, considering both **high implementation acreages and cost-effectiveness.**

Methodology

Statistical methods

- The **base/control scenario** considering **all potential BMPs (up to 200)** and the **innovization method** (**INN1 & INN2**) consider up to **10 BMPs**.
- Compare the new INN1 and INN2 approaches with the **regular optimization (Control)**

The Wilcoxon-Mann-Whitney test method

Results (INN1) Urban vs. Ag

Results

1

Examine different ranking methodologies to **identify the top BMPs,**

Table: Top choice BMPs based on the all three-ranking methodology

Urban Top-Ranked Counties			Agricultural Top-Ranked Counties		
Strategy 1	Strategy 2	Strategy 3	Strategy 1	Strategy 2	Strategy 3
Nutrient Management Plan High-Risk Lawn	Barnyard Runoff Control	Nutrient Management Plan High-Risk Lawn	Nutrient Management Plan High-Risk Lawn	Barnyard Runoff Control	Nutrient Management Plan High-Risk Lawn
Off-Stream Watering Without Fencing	Agricultural Stormwater Management	Off-Stream Watering Without Fencing	Off-Stream Watering Without Fencing	Nutrient Management Plan High-Risk Lawn	Off-Stream Watering Without Fencing
Nutrient Management N Timing	Nutrient Management Plan High-Risk Lawn	Nutrient Management N Rate	Nutrient Management N Timing	Nutrient Management N Rate	Nutrient Management N Rate
Precision Intensive Rotational/Prescribed Grazing	Soil Conservation and Water Quality Plans	Forest Harvesting Practices	Nutrient Management N Rate	Nutrient Management N Timing	Nutrient Management N Timing
Cover Crop Traditional Rye Early Drilled	Off-Stream Watering Without Fencing	Nutrient Management N Timing	Cover Crop Traditional Rye Early Drilled	Agricultural Stormwater Management	Cover Crop Traditional Rye Early Drilled
Nutrient Management N Rate	Nutrient Management N Placement	Cover Crop Traditional Rye Early Drilled	Nutrient Management N Placement	Nutrient Management N Placement	Nutrient Management N Placement
Nutrient Management N Placement	Nutrient Management N Rate	Nutrient Management N Placement	Barnyard Runoff Control	Off-Stream Watering Without Fencing	Cover Crop Traditional Rye Early Other

All three-ranking methodologies:

MICHIGAN STATE UNIVERSITY

Results

1 Examine different ranking methodologies to **identify the top BMPs,**

All three-ranking methodologies:

Table: Top choice BMPs based on the all three-ranking methodology

Urban Top-Ranked Counties			Agricultural Top-Ranked Counties		
Strategy 1	Strategy 2	Strategy 3	Strategy 1	Strategy 2	Strategy 3
Nutrient Management Plan High-Risk Lawn	Barnyard Runoff Control	Nutrient Management Plan High-Risk Lawn	Nutrient Management Plan High-Risk Lawn	Barnyard Runoff Control	Nutrient Management Plan High-Risk Lawn
Off Stream Watering Without Fencing	Agricultural Stormwater Management	Off Stream Watering Without Fencing	Off Stream Watering Without Fencing	Nutrient Management Plan High-Risk Lawn	Off Stream Watering Without Fencing
Nutrient Management N Timing	Nutrient Management Plan High-Risk Lawn	Nutrient Management N Rate	Nutrient Management N Timing	Nutrient Management N Rate	Nutrient Management N Rate
Precision Intensive Rotational/Prescribed Grazing	Soil Conservation and Water Quality Plans	Forest Harvesting Practices	Nutrient Management N Rate	Nutrient Management N Timing	Nutrient Management N Timing
Cover Crop Traditional Rye Early Drilled	Off Stream Watering Without Fencing	Nutrient Management N Timing	Cover Crop Traditional Rye Early Drilled	Agricultural Stormwater Management	Cover Crop Traditional Rye Early Drilled
Nutrient Management N Rate	Nutrient Management N Placement	Cover Crop Traditional Rye Early Drilled	Nutrient Management N Placement	Nutrient Management N Placement	Nutrient Management N Placement
Nutrient Management N Placement	Nutrient Management N Rate	Nutrient Management N Placement	Barnyard Runoff Control	Off Stream Watering Without Fencing	Cover Crop Traditional Rye Early Other

MICHIGAN STATE UNIVERSITY

Results:

1 Examine different ranking methodologies to **identify the top BMPs,**

Table: Top choice BMPs based on the all three-ranking methodology

All three-ranking methodologies:

More than once BMPs

Once repeated BMPs

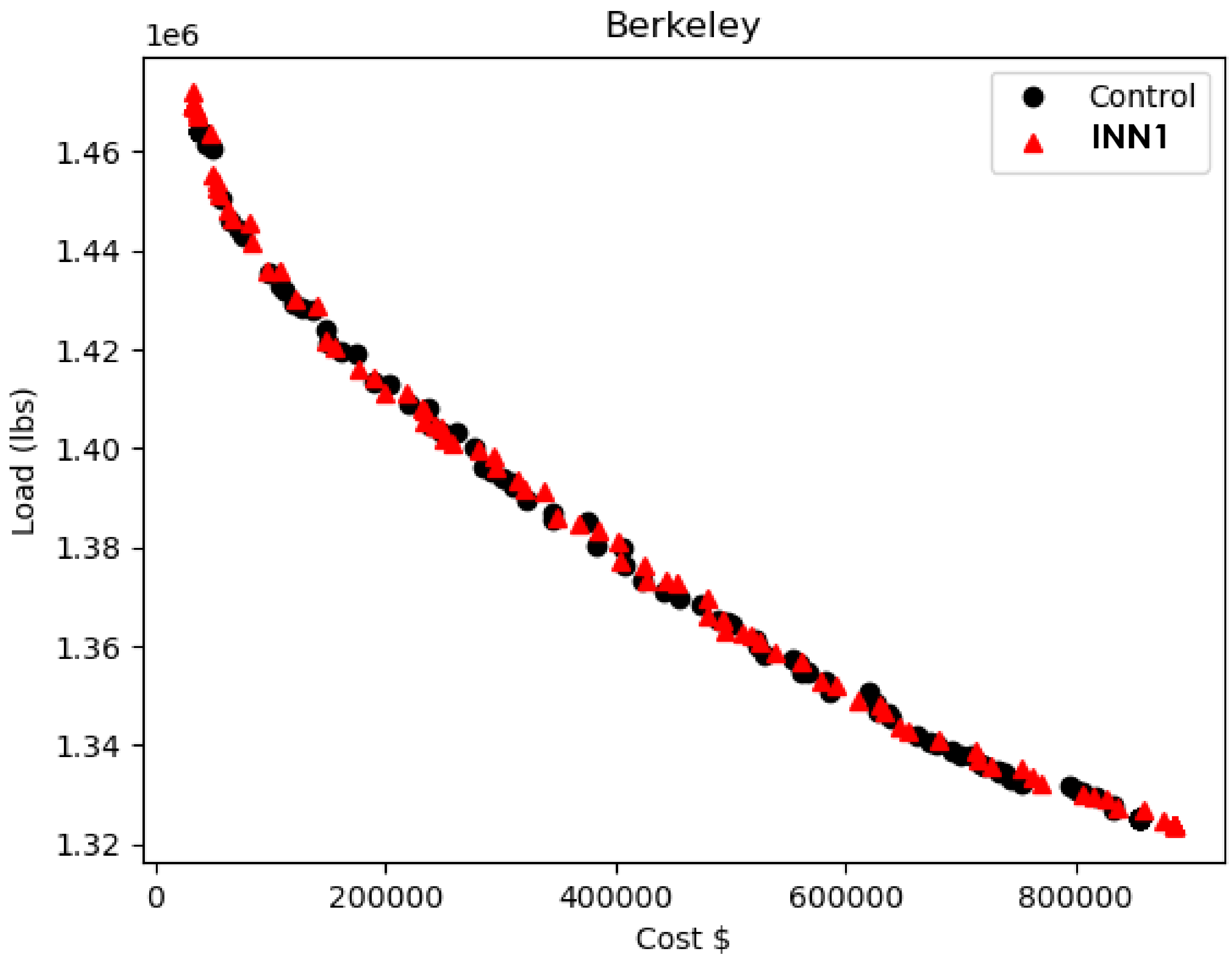
Urban Top-Ranked Counties			Agricultural Top-Ranked Counties		
Strategy 1	Strategy 2	Strategy 3	Strategy 1	Strategy 2	Strategy 3
Nutrient Management Plan High-Risk Lawn	Barnyard Runoff Control	Nutrient Management Plan High-Risk Lawn	Nutrient Management Plan High-Risk Lawn	Barnyard Runoff Control	Nutrient Management Plan High-Risk Lawn
Off-Stream Watering Without Fencing	Agricultural Stormwater Management	Off-Stream Watering Without Fencing	Off-Stream Watering Without Fencing	Nutrient Management Plan High-Risk Lawn	Off-Stream Watering Without Fencing
Nutrient Management N Timing	Nutrient Management Plan High-Risk Lawn	Nutrient Management N Rate	Nutrient Management N Timing	Nutrient Management N Rate	Nutrient Management N Rate
Precision Intensive Rotational/Prescribed Grazing	Soil Conservation and Water Quality Plans	Forest Harvesting Practices	Nutrient Management N Rate	Nutrient Management N Timing	Nutrient Management N Timing
Cover Crop Traditional Rye Early Drilled	Off-Stream Watering Without Fencing	Nutrient Management N Timing	Cover Crop Traditional Rye Early Drilled	Agricultural Stormwater Management	Cover Crop Traditional Rye Early Drilled
Nutrient Management N Rate	Nutrient Management N Placement	Cover Crop Traditional Rye Early Drilled	Nutrient Management N Placement	Nutrient Management N Placement	Nutrient Management N Placement
Nutrient Management N Placement	Nutrient Management N Rate	Nutrient Management N Placement	Barnyard Runoff Control	Off-Stream Watering Without Fencing	Cover Crop Traditional Rye Early Other

Results:

Reduce the optimization variables & search space

	Control	INN1
Berkeley	14,090	510 (3%)
Hardy	18,607	725 (3%)
Jefferson	12,303	456 (3%)
Mineral	20,260	765 (3%)

Compare control against INN 1 (Berkeley County)



Statistical analysis

Compare the new INN1 approach with the regular optimization (Control/Base)

The Wilcoxon-Mann-Whitney test method (p-value)

Innovization	Urban Dominated counties		Agricultural Dominated counties		All counties
	Berkeley	Mineral	Jefferson	Hardy	
INN1	0.065	4.08E-05	0.00354	4.08E-05	0.338

MICHIGAN STATE UNIVERSITY

Results (INN2) Combined

Results:

Statistical methods:

The Wilcoxon-Mann-Whitney test- Statistical analysis for INN2 section:

- **P value of (Hyper volume-HV) for all four strategies for 5 and 10 top BMP (INN2) compared to the Control case**

	Control	*Control*	*Control*	*Control*	*Control*
Strategies	**Berkeley**	**Hardy**	**Jefferson**	**Mineral**	**all counties**
Strategy-1-10	0.032986	0.146712	0.000319	0.050334	0.2654347
Strategy-1-5	0.178967	0.000041	0.007558	0.000041	0.0060929
Strategy-2-10	0.000041	0.001564	0.000041	0.083953	0.3380517
Strategy-2-5	0.000041	0.000053	0.000041	0.000041	0.0060929
Strategy-3-10	0.000041	0.000091	0.000041	0.000811	0.0060929
Strategy-3-5	0.000646	0.000646	0.000041	0.000041	0.0060929
Strategy-4-10	0.000319	0.000053	0.000041	0.118609	0.0060929
Strategy-4-5	0.000041	0.000646	0.000117	0.000041	0.0060929

Conclusion

Conclusion Innovization

1.First Innovization Technique:

This method greatly **simplified the BMP selection process** by **reducing the number** from approximately 200 to just seven. This simplification aids in:

1. **Simplicity and Manageability**
2. **Reduce Optimization Constrained**

2. Second Innovization Technique:

This approach was **validated** across various regions, showing improved performance in adapting the innovization method beyond the initial study area.

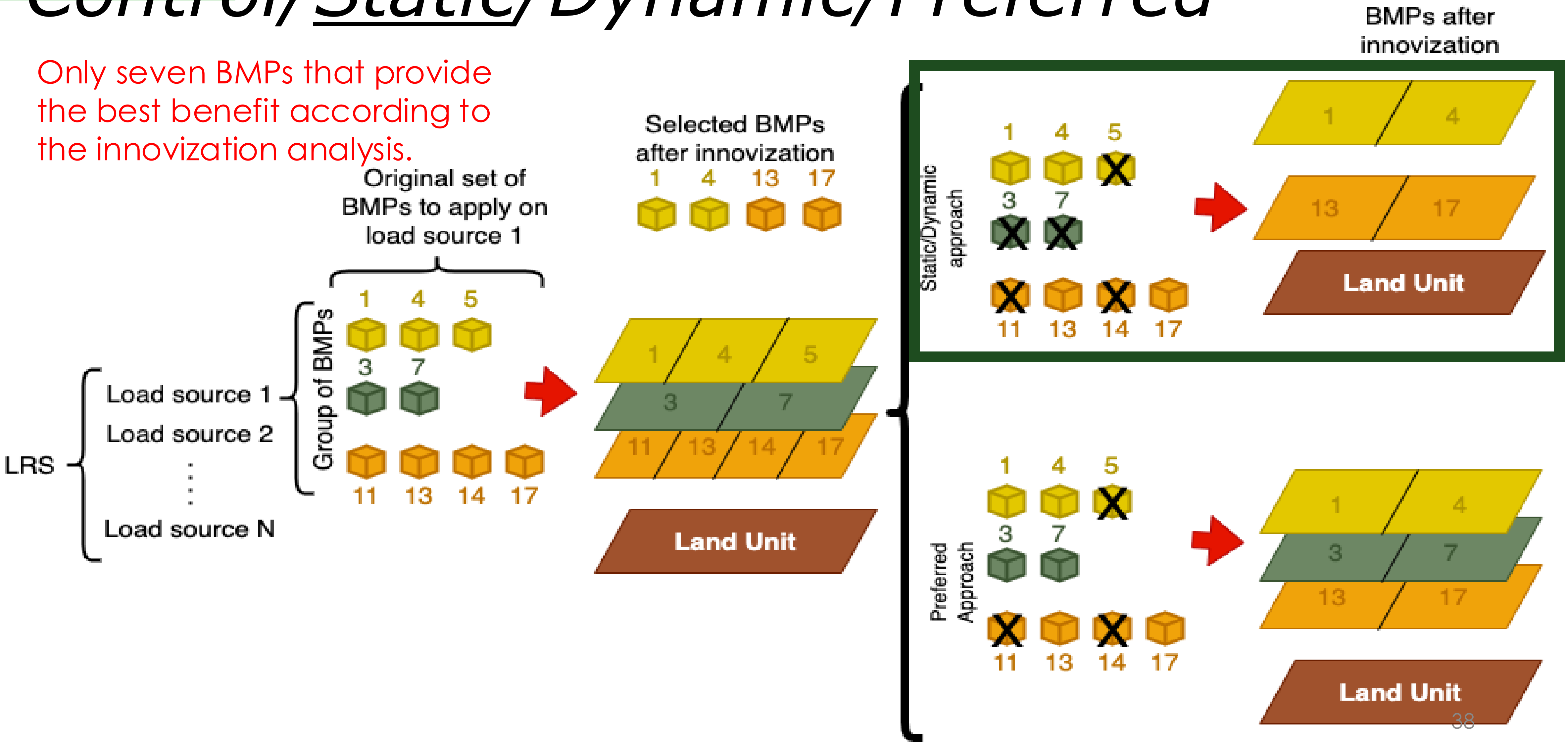
MICHIGAN STATE UNIVERSITY

Re-Optimization with Reduced BMPs

- **Purpose**:
  - Verify that a reduced set of BMPs can still produce optimal results.
  - Ensure that the simplified BMPs are still effective and efficient.
- **Validation**:
  - Confirm that the selected BMPs from innovization are capable of achieving desired outcomes.
  - Re-optimization ensures robustness and reliability of the reduced BMP set.
- **Efficiency Gains:**
  - Faster run times due to fewer variables.
  - Reduced computational effort.
- **Quality Results:**
  - Maintain or improve solution quality with fewer BMPs.
  - Demonstrate that the reduced set is effective in achieving goals.

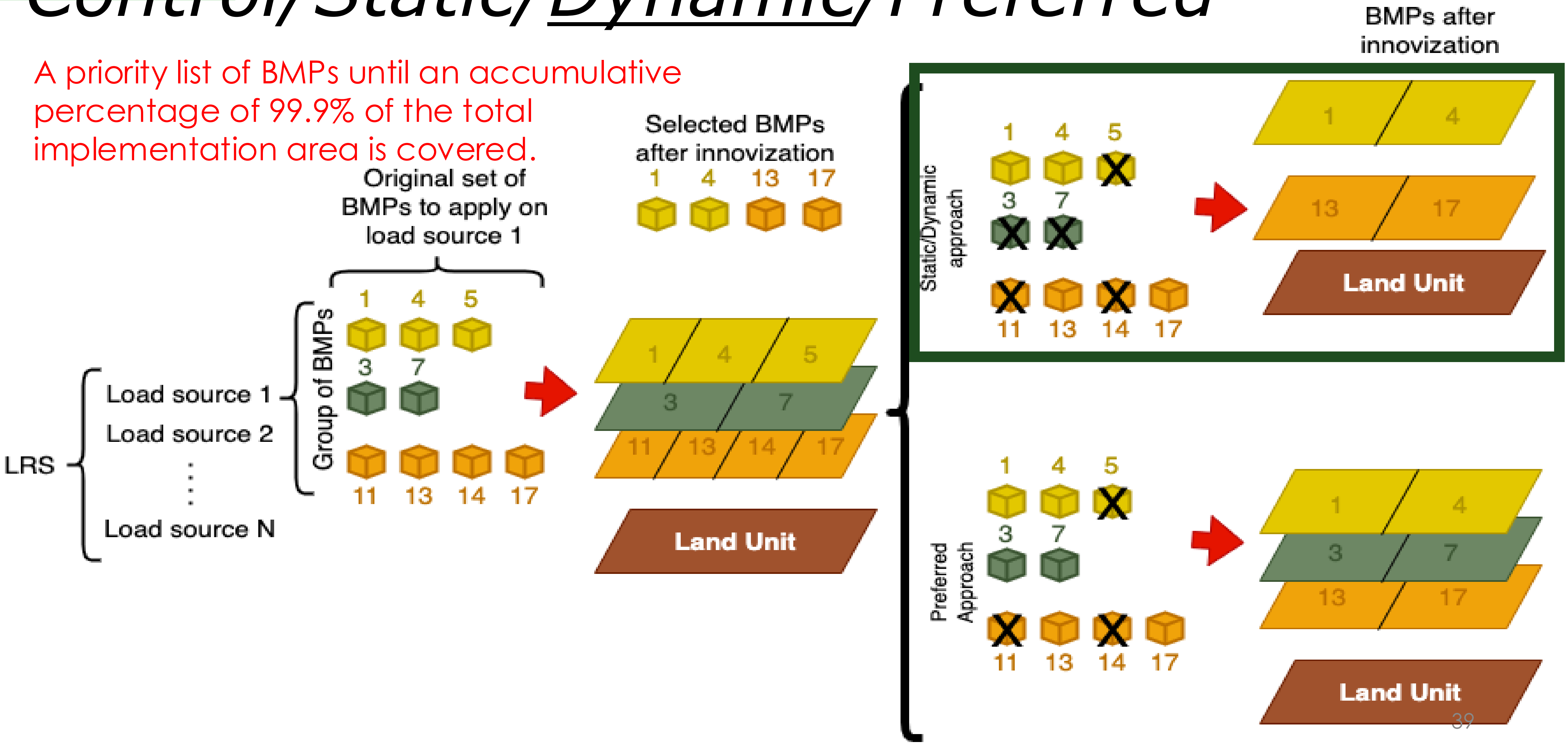
Re-optimization Strategies Control/Static/Dynamic/Preferred

Only seven BMPs that provide the best benefit according to the innovization analysis.



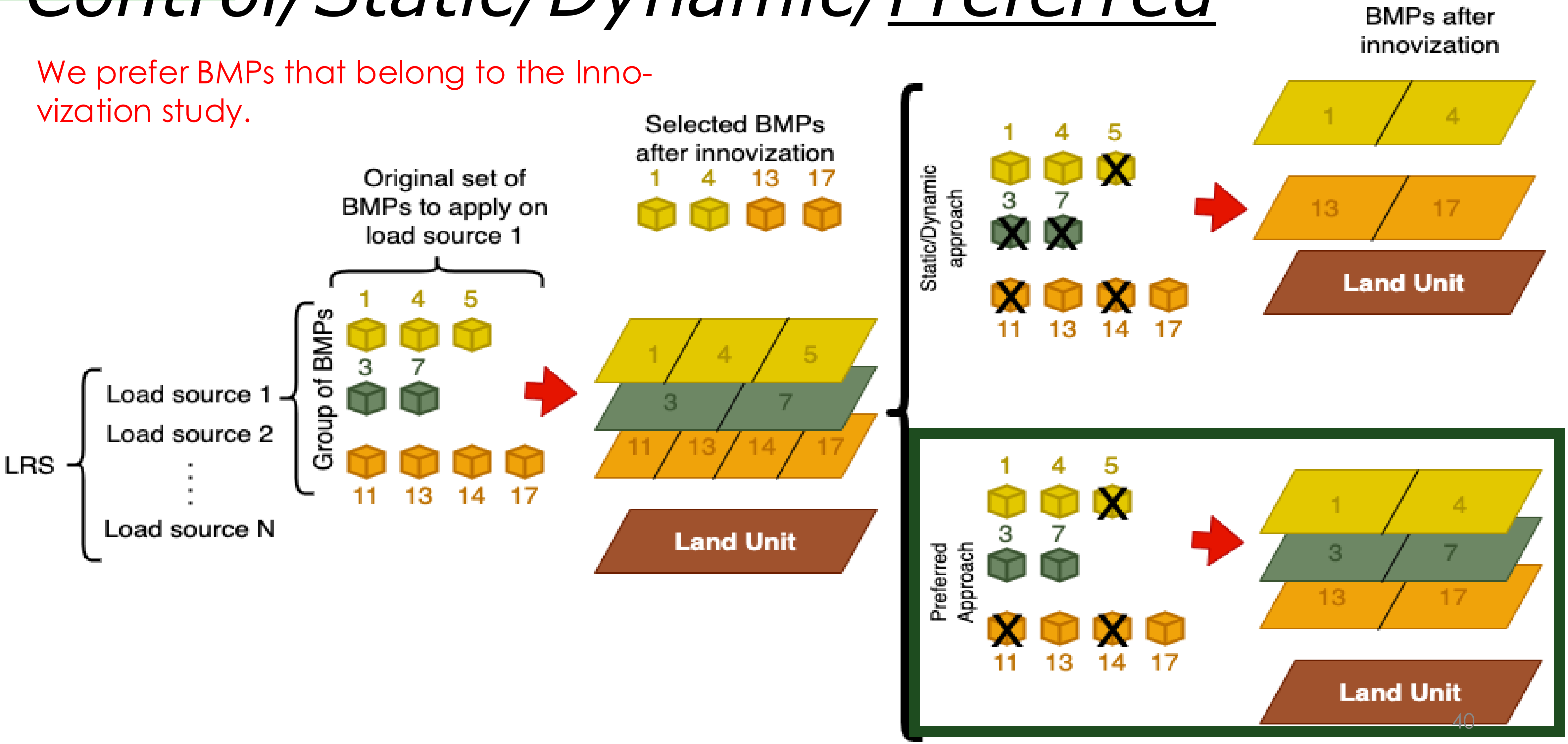
Re-optimization Strategies Control/Static/Dynamic/Preferred

A priority list of BMPs until an accumulative percentage of 99.9% of the total implementation area is covered.

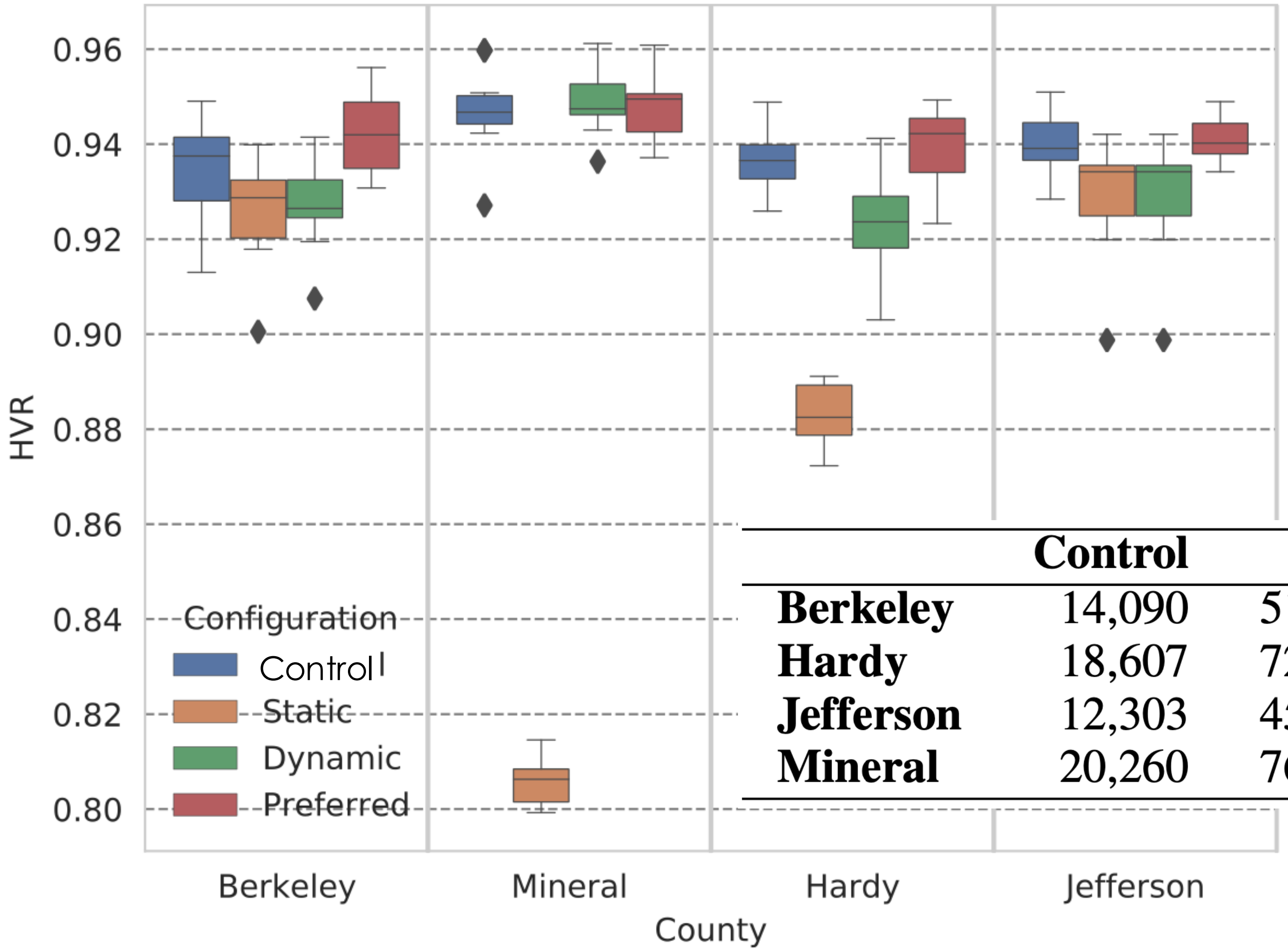


Re-optimization Strategies Control/Static/Dynamic/Preferred

We prefer BMPs that belong to the Innovization study.



Validation with Hypervolume Ratio

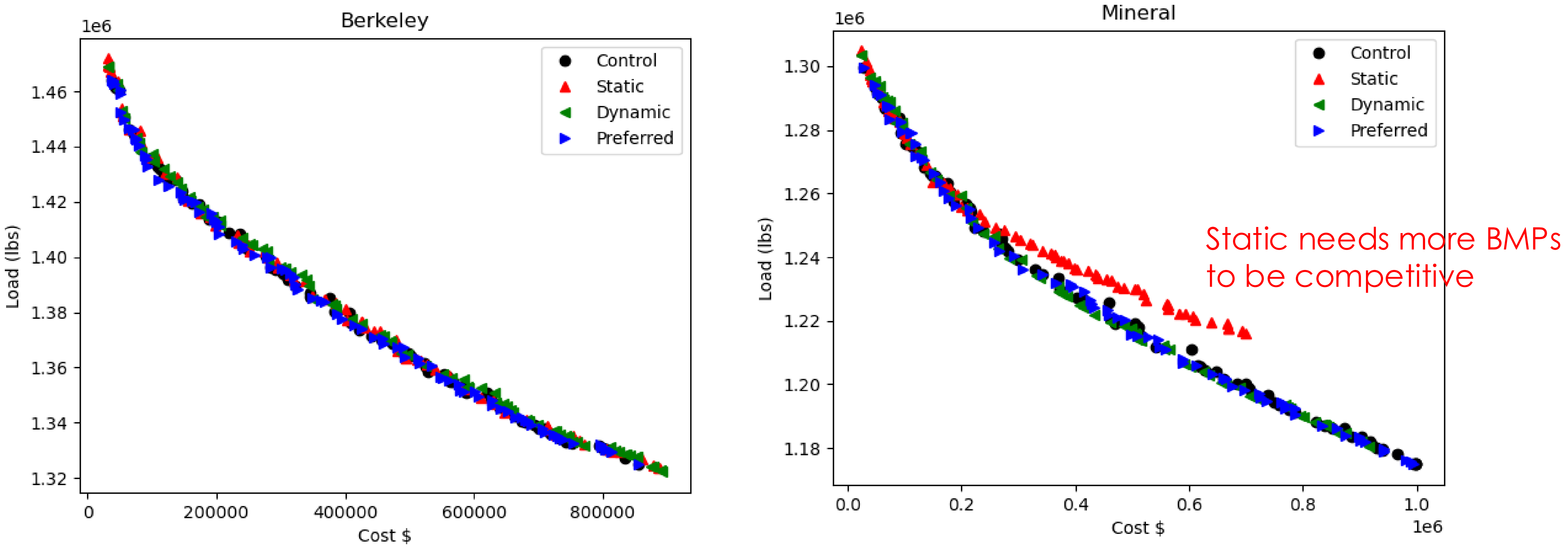


- Competitive Results
- Fraction of the variables.
- Static failed in some counties (it couldn't find a competitive solution with such a small BMP list).
- Preferred outperformed the other approaches (better exploitation of the area).

An order of magnitude less regarding the number of variables

	Control	**Static**	**Dynamic**	**Preferred**
Berkeley	14,090	510 (3%)	1,023 (7%)	4,823 (34%)
Hardy	18,607	725 (3%)	751 (4%)	5,307 (28%)
Jefferson	12,303	456 (3%)	456 (3%)	4,079 (33%)
Mineral	20,260	765 (3%)	1,650 (8%)	6,415 (31%)

Pareto Fronts *(Representative)*



Re-optimizing with fewer variables can produce similar performance on the large-scale problem: Our Suggestion: Preferred strategy

Achievements with Re-optimization

- With a fraction of variables, we achieve similar results
- Validated on single and multiple counties

Transfer Learning

- **Definition:**
  - Transfer learning is the application of knowledge gained from solving one problem to a different but related problem.
- **Simplified Explanation:**
  - Imagine learning to drive a car. Once you know how to drive one car, you can easily drive any other car without having to learn from scratch again.

Why Use Transfer Learning?

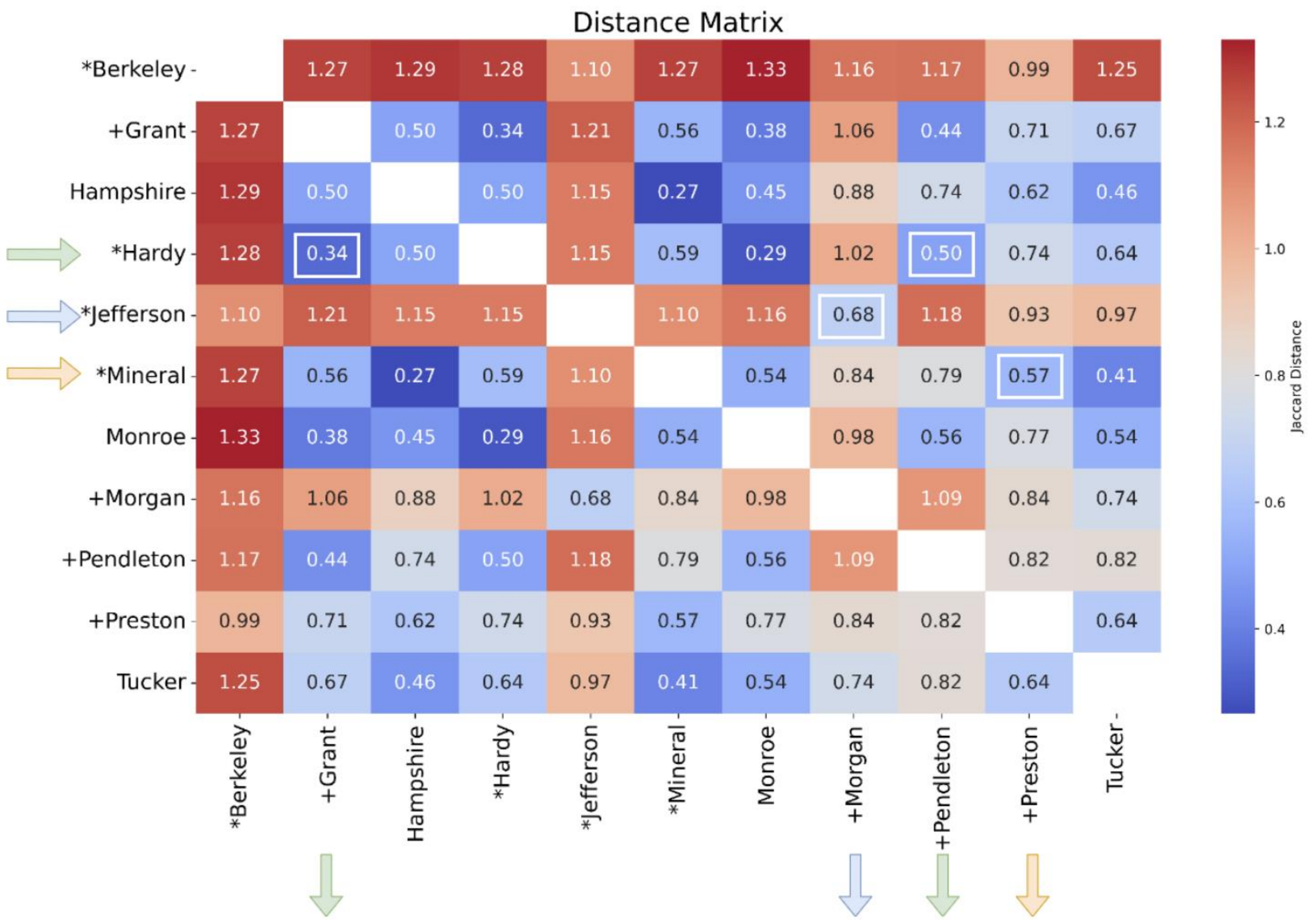
- **Efficiency**: Reduces the need to start from scratch for every new scenario.
- **Consistency**: Maintains the quality of solutions across different scenarios.
- **Scalability**: Easily applied to multiple new scenarios or regions.

Applying Transfer Learning to BMPs

- Select New Counties (Scenarios):
  - Identify new areas where BMPs need to be implemented.
- Similarity Measure:
  - Use a measure (e.g., geographic, demographic, or environmental data) to determine how similar these new counties are to previously analyzed ones (Jaccard distance over the load-sources of all LRS's of a county).
- Transfer BMPs:
  - Apply the selected BMP list from similar, previously analyzed counties to the new ones.

Transfer Learning to Other Counties

Innovative Transfer Learning for Water Quality Optimization

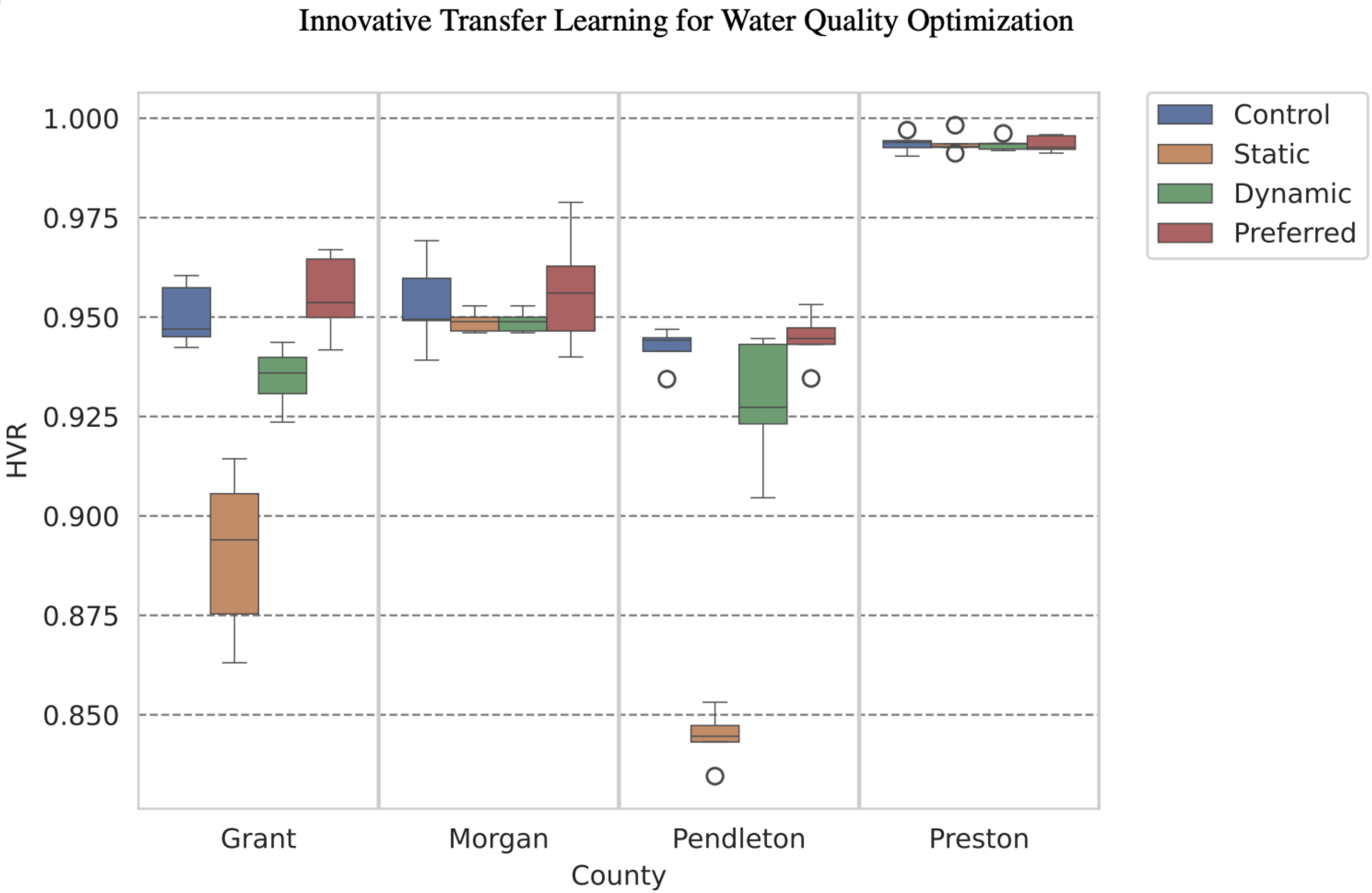


Transfer Learning to Other Counties

	Control	Static (% reduction)	Dynamic (% reduction)	Preferred (% reduction)
Grant	25,228	962 (3.8%)	1217 (4.8%)	8,036 (31%)
Morgan	11,880	466 (3.9%)	466 (3.9%)	3,947 (33%)
Pendleton	33,083	1,207 (3.6%)	1,851 (5.6%)	10,483 (31%)
Preston	1,470	53 (3.5%)	118 (8%)	447 (30%)

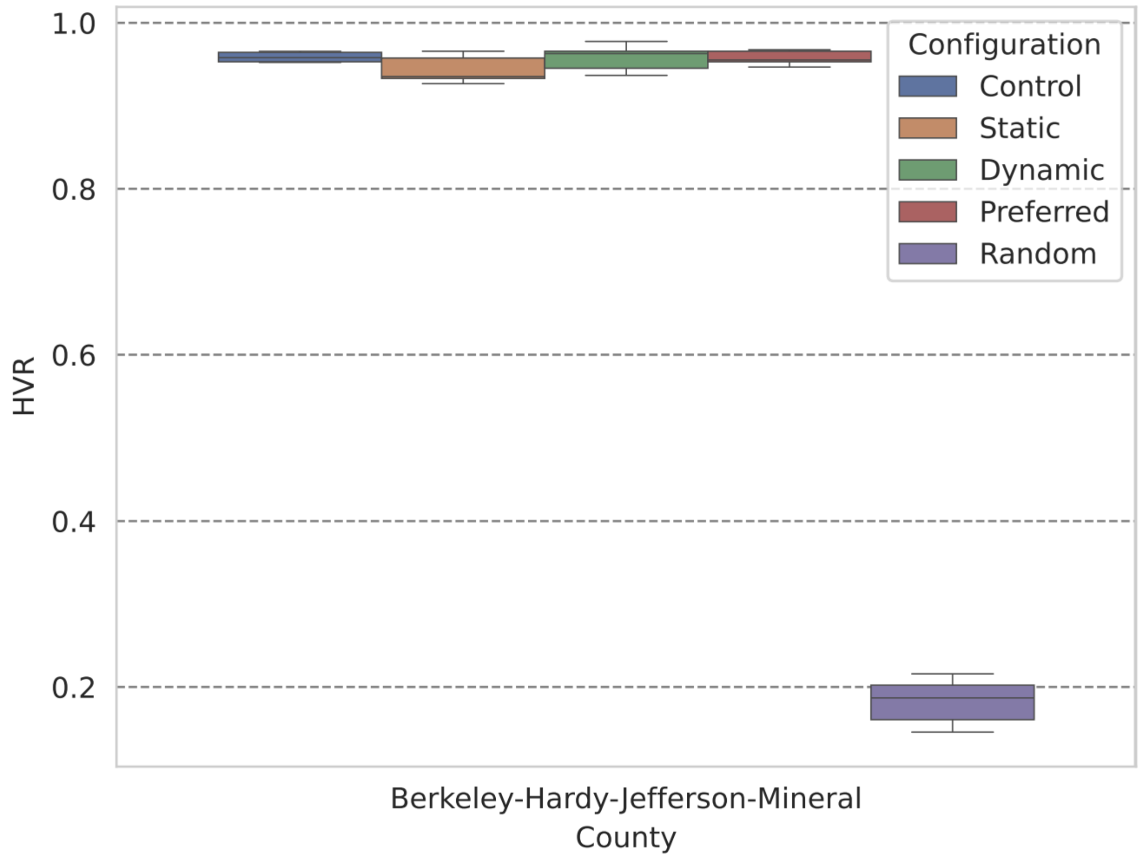
Above counties were not seen by the innovization study, but the right identification of BMPs could be successfully transferred.

Transfer Learning to Other Counties



Transfer Learning (Multi-county)

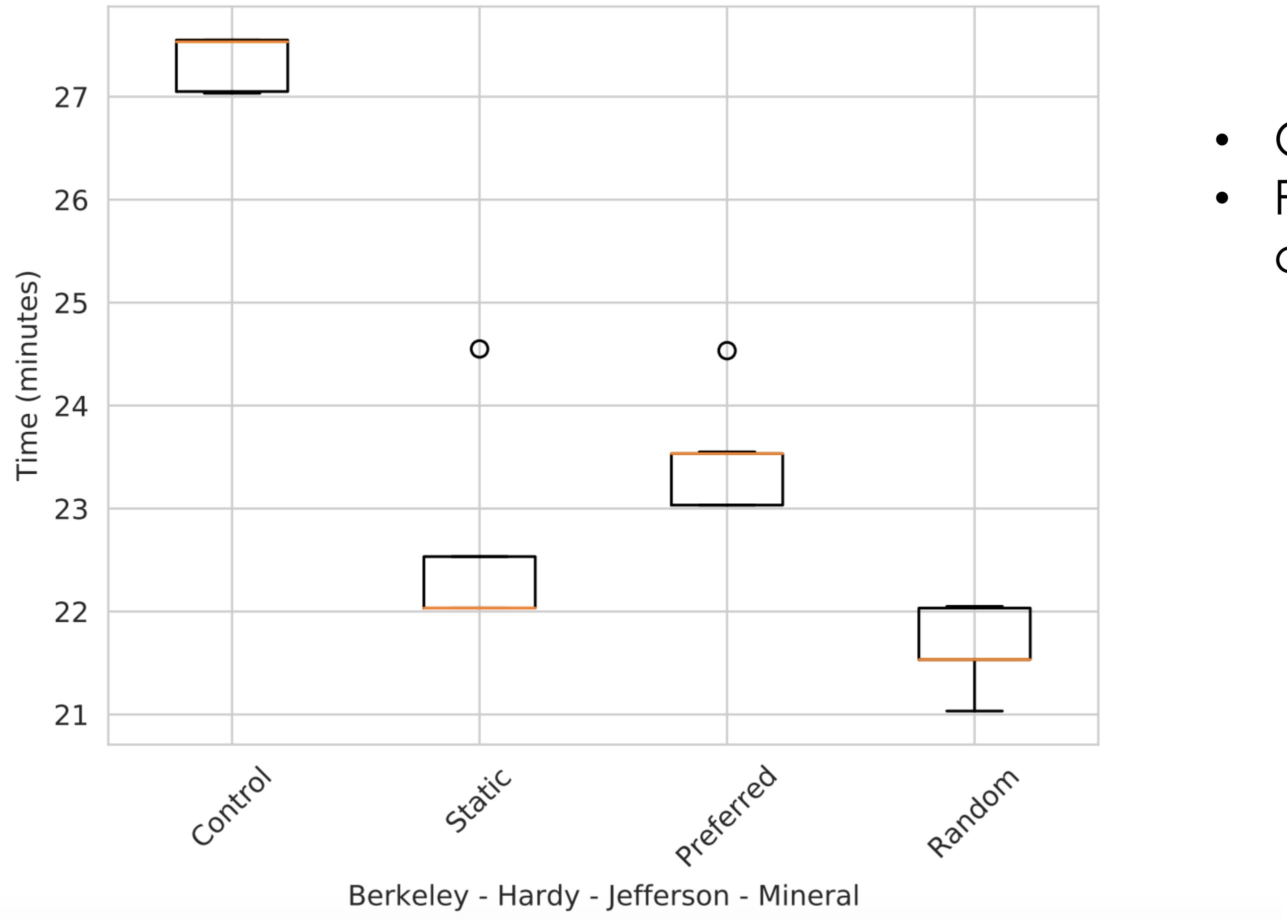
Re-optimization	# Variables
Control	65,260
Static	2,456
Dynamic	5,882
Preferred	2,456



The random approach could not converge in a multi-county scenario when BMPs were selected at random.

Even when the number of variables increased, re-optimizing with the selected BMPs produced a competitive result.

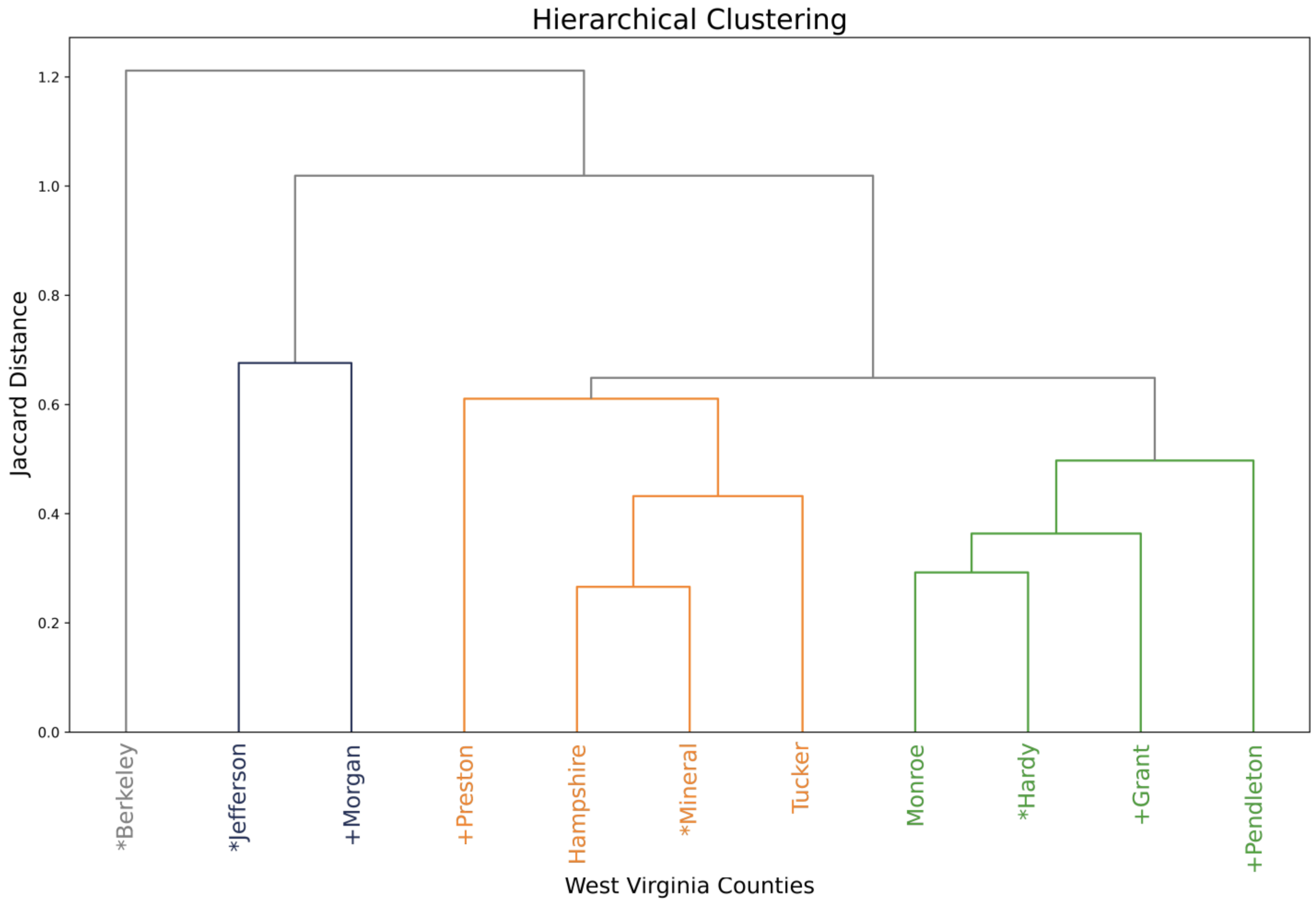
Transfer Learning (Multi-county)



- Our study suggests:
- Preferred strategy is a good compromise

Transfer Learning to Other Counties

Innovative Transfer Learning for Water Quality Optimization



Conclusions / Future work

- **Real-world** large-scale application with dozens of potential users (**reduce pollutants**).
- "Innovization" helps to perform a **comput. efficient search**.
- Re-optimization validated the use of reduced variables
- Transfer learning allowed a reduced BMP list obtained from a few counties to apply to a few more similar counties.
- Transfer learning reduces time, and complexity and shows promise to solve large-scale problems.

iEMSs 2024 Biennial Conference: Abstracts

1. Toscano, G. A. P. Nejadhashemi, K. Deb, H. Razavi. L. Linker, 2024. Advancing Watershed Management: A Multiobjective Optimization and Multicriteria Decision-Making Platform.
2. Deb, K., Lu, Z.,Kropp, I., Hernandez-Suarez, S., Hussein, R., Miller, S. and Nejadhashemi, A. P., 2024, Reliable Decision-making Under Uncertainty in Hierarchical Multi-Criterion Problems.
3. Deb, K., Goodman, E., and Chikumbo, O. 2024, Multi-objective Land Use Management and Decision-making.
4. H. Razavi. A. P. Nejadhashemi, K. Deb, Toscano, G. L. Linker, 2024. Water Resources Management: A Comprehensive Analysis of Elements, Interconnections, and Emerging Synergies.
5. H. Razavi. , G. L. Toscano, A. P. Nejadhashemi, K. Deb, Linker, 2024. Innovative Ranking Methods for Parameter Size Reduction in Large Scale Multi-Objective Optimization Problem.

Journal Papers

- A. P. Nejadhashemi, K. Deb, G. Toscano, H. Razavi, L. Linker. Leveraging Innovization and Transfer Learning to Optimize Best Management Practices in Large-Scale Watershed Management. Journal of Environmental Modelling & Software. Accepted with minor revisions.

NEXT STEPs

- Scale up to watershed level optimization

- Parallel computing platform for even faster execution

- Workshops with CBP users for feedback and improvement of our approaches

MICHIGAN STATE
UNIVERSITY

Computational Optimization and Innovation

Thank you

C-ECO

Computational Ecohydrology

Thank you!