

Meeting with STAC Members  
June 13, 2023

# CAST Optimization: Intermediate Status

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

MICHIGAN STATE UNIVERSITY



# Agenda

1

Introduction to Michigan State University (MSU) Team

4

Current Status of the Project

2

STAC Workshop Report, 2016: Goals and Applications

5

Short Demo (video)

3

Objectives and Main Tasks of the Project

6

Future Plan

# Kalyanmoy Deb



## • **Title:**

- **University Distinguished Professor**
- **Koenig Endowed Chair Professor**
  - **Dept of Electrical and Computer Engineering**
  - **Dept of Computer Science and Engineering**
  - **Dept of Mechanical Engineering**

## • **Expertise and Achievements:**

- **Optimization, Multi-objective optimization, Machine Learning, Modeling**
- **36 years of experience in optimization and its applications**
- **Author of popular evolutionary optimization methods: NSGA-II, NSGA-III**
- **Author of two text-books on optimization, 610 research papers**
- **185,000 Google Scholar citations, h-index: 133**
- **Director, Computational Opt. and Innovation (COIN) Lab at MSU**

# Pouyan Nejadhashemi



- **Title:**
  - **University Foundation Professor**
    - **Department of Biosystems and Agricultural Engineering**
    - **Department of Plant, Soil and Microbial Sciences**
  - **Elected board member**
    - **International Environmental Modelling & Software Society**
- **Expertise and Achievements:**
  - **Soft computing applications in water resources management**
  - **Computational Ecohydrology**
  - **Evaluation and development of watershed and water quality models**
  - **\$41M in grant funding**
  - **130 peer-reviewed publications**
  - **180 scientific presentations**
  - **Director, Center for Intelligent Water Resources Engineering (CIWRE)**

# Gregorio Toscano



- **Title:**

- CBPO CAST optimization researcher
- Associate Professor - Center for Research and Advanced Studies, Mexico
- PhD in Evol. Multi-Criterion Optimization, 2005

- **Expertise and Achievements:**

- Multi-objective optimization, Computational Intelligence, and Machine Learning
- Multi-objective Micro-GA, Multi-objective PSO
- Full Stack
- Programming Languages
- 7,692 Google Scholar citations

# Hoda Razavi



- **Title:**

- **PhD Student, Biosystems and Agricultural Engineering, Michigan State University**
- **MS Water and Hydraulic Structures, Civil Engineering, Khajeh Nasir Toosi University of Technology**
- **BS Civil Engineering University of Tehran**

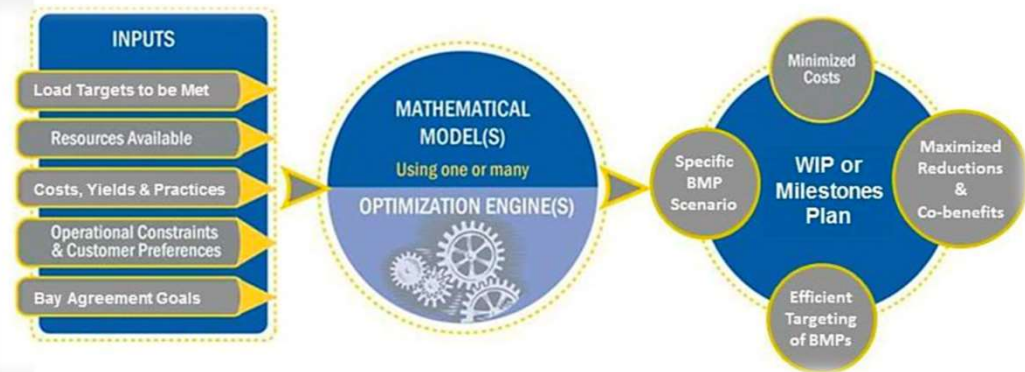
- **Expertise:**

- **Watershed/water quality modeling**
- **Environmental flow**
- **Multi-objective optimization**
- **Uncertainty quantification**
- **Water resources management**
- **Data-driven models**
- **Climate change impacts**

# “Cracking the WIP” - Designing an Optimization Engine to Guide Efficient Bay Implementation *STAC Workshop Report, 2016*

## Goals:

1. **Cost minimization** was a key goal for the partners
2. **Maximizing co-benefits**, particularly those supporting Chesapeake Bay Watershed Agreement goals
3. **Maximizing load reduction** reliability
4. Equitable distribution of effort among jurisdictions
5. Equitable distribution of effort among source sectors
6. Limits on retirement of agricultural land
7. Ability to use the tool at **various scales**



# Applications: STAC Workshop Report, 2016

## Under development

- Address optimization of multiple co-benefits
- Minimize costs of BMP implementation
- Optimal use of BMPs on land use by county
- Make Bay TMDL load targets achievable
- Make Phase III WIP scenario that achieves nitrogen, phosphorus, and sediment targets for the lowest cost with the ability to tweak to see the different scenario costs

## Future Application

- Rethink the allocation of responsibilities by sector, by geography, by funding
- Recognize the value/influence of ecosystem services in local decision making
- Assist with the development of grant applications
- Help local governments document co-benefits of WIP implementation
- Identify cost savings within a source sector
- Provide a basic resource for planners to understand advantages and disadvantages of implementation options
- Assist progress towards other management strategy objectives
- Help in development of state implementation plans (Bay Milestones and WIPs)
- Help in development of local implementation plans (local and Bay TMDLs)
- Develop sector implementation plans cost effectively

# Objective of the MSU-Optimization Project



Investigate, develop, program, verify, and implement an optimization system built around the CBP's CAST Model to:



- Improve the water quality
- At the lowest cost



# Timeline of the Project

Calendar Year

Calendar Quarter

Project Year

**Task 1: Development of an efficient single-objective optimization procedure for cost-effective BMP allocation**

- 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
- 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

**Task 2: Development of an efficient multi-objective (MO) optimization procedure for cost-loading trade-off BMP allocation**

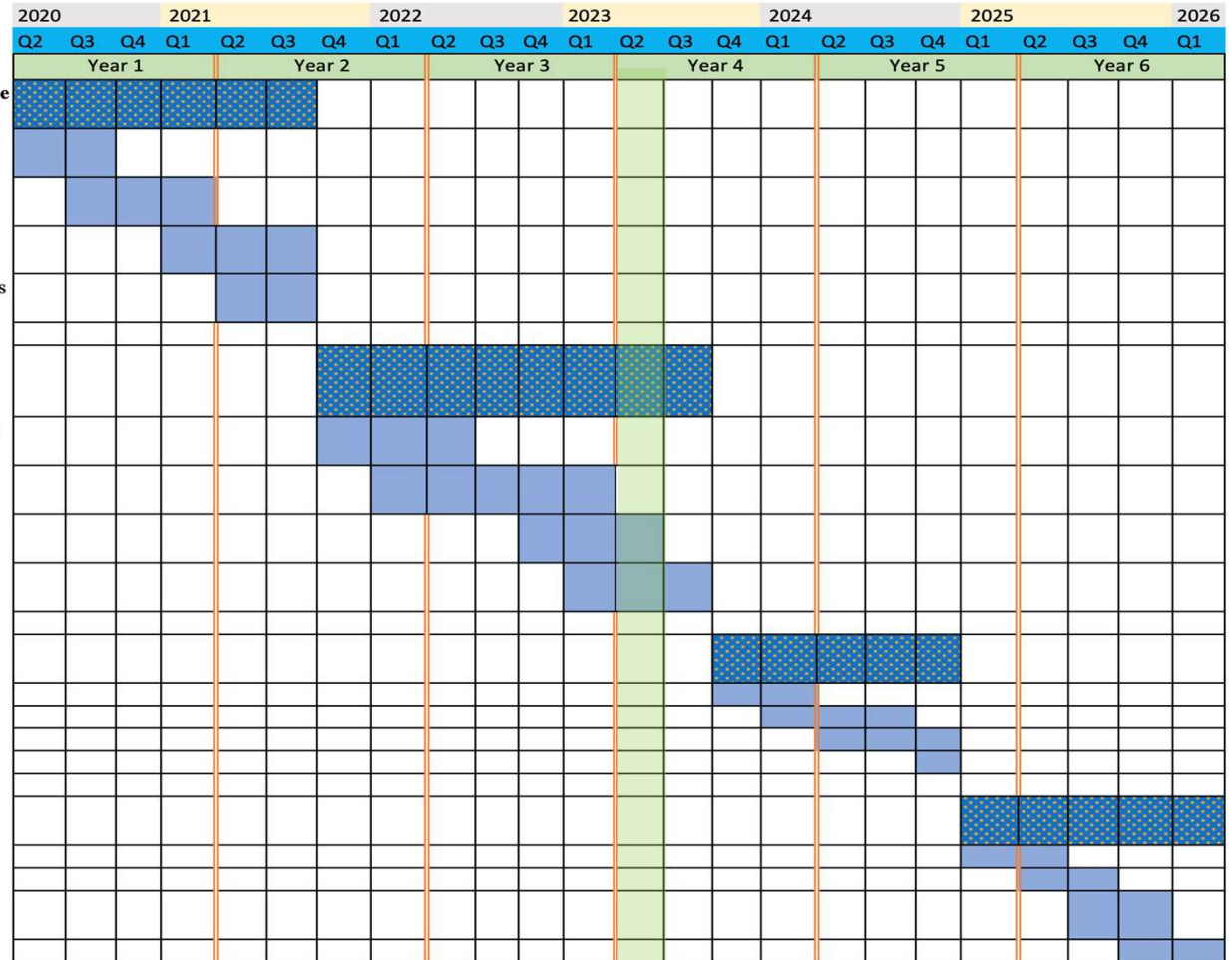
- 2.1: Develop generative MO optimization using hybrid optimization procedure developed at Task 1
- 2.2: Develop simultaneous MO customized optimization using population-based evolutionary algorithms
- 2.3: Comparison of generative & simultaneous procedures and validation with CBP users & decision-makers
- 2.4: Develop an interactive multi-criterion decision-making aid for choosing a single preferred solution

**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

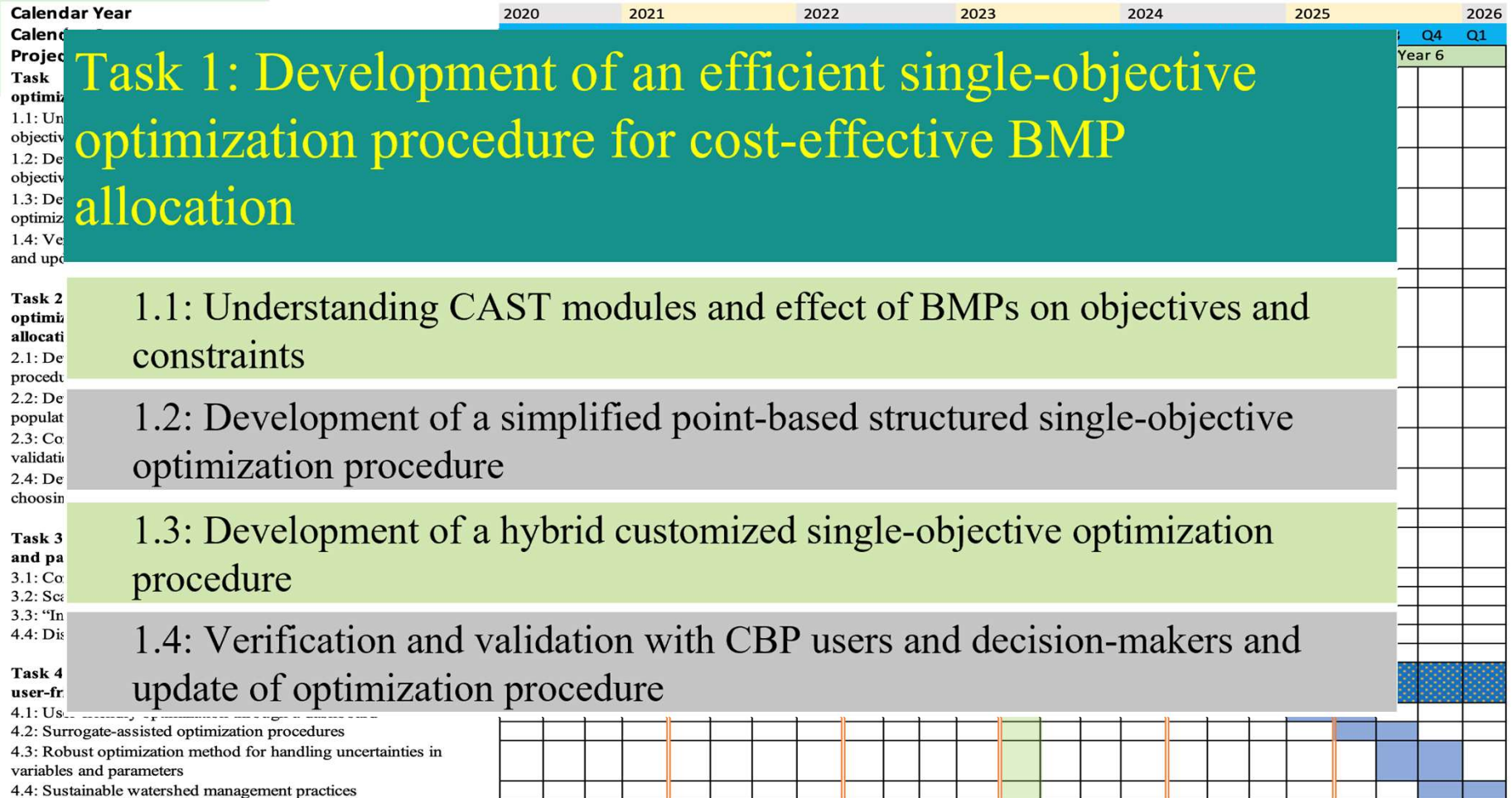
**Task 4: Interactive optimization and decision-making using user-friendly dashboard**

- 4.1: User-friendly optimization through a dashboard
- 4.2: Surrogate-assisted optimization procedures
- 4.3: Robust optimization method for handling uncertainties in variables and parameters
- 4.4: Sustainable watershed management practices



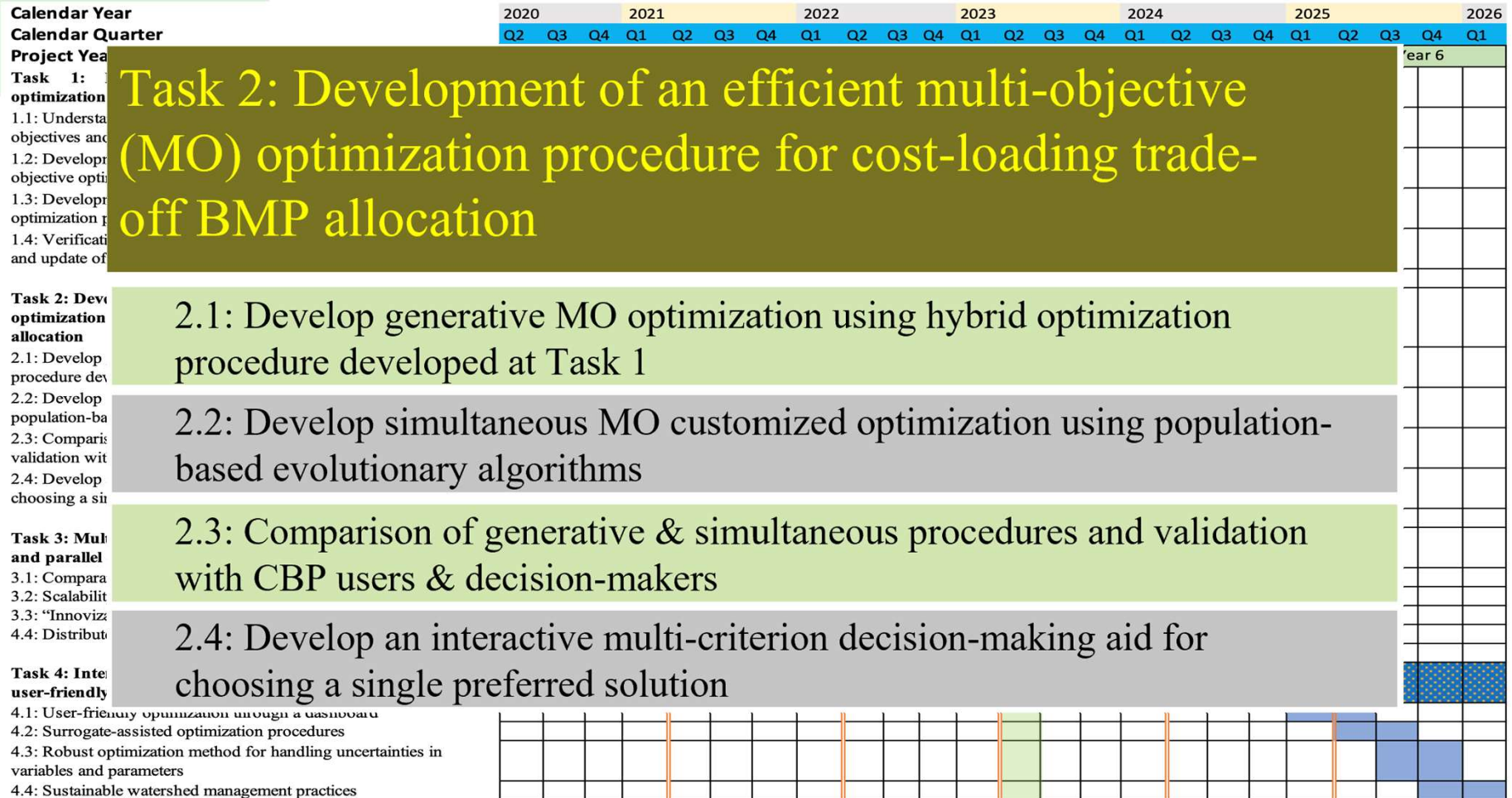
We are here

# Timeline of the Project



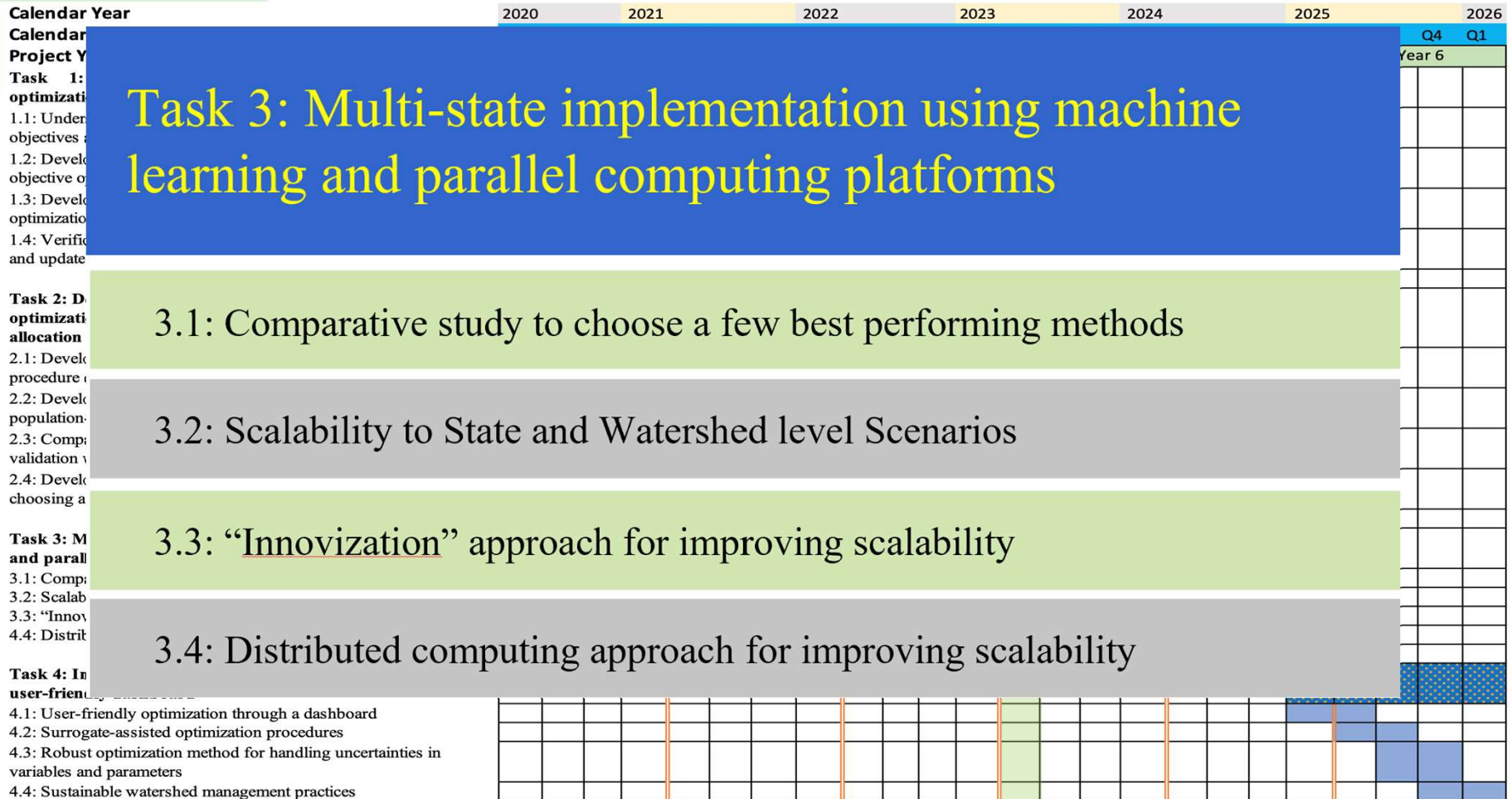
We are here

# Timeline of the Project



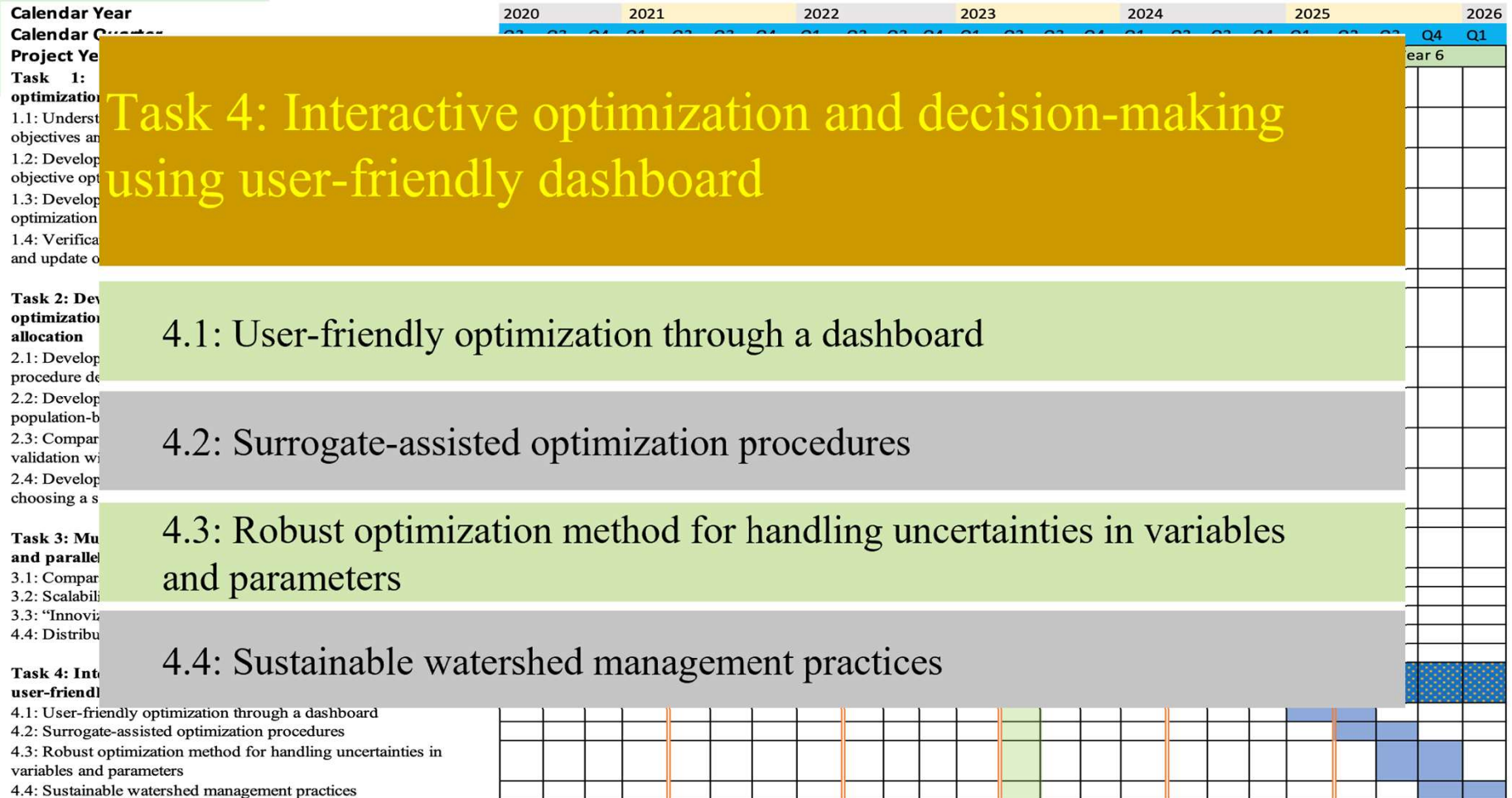
We are here

# Timeline of the Project



We are here

# Timeline of the Project



**Calendar Year**  
**Calendar Quarter**  
**Project Year**  
**Task 1: optimization**  
 1.1: Underst objectives an  
 1.2: Develop objective opt  
 1.3: Develop optimization  
 1.4: Verifica and update o  
**Task 2: Dev optimization allocation**  
 2.1: Develop procedure de  
 2.2: Develop population-b  
 2.3: Compar validation wi  
 2.4: Develop choosing a s  
**Task 3: Multi and parallel**  
 3.1: Compar  
 3.2: Scalabili  
 3.3: "Innoviz  
 4.4: Distribu  
**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  
 4.3: Robust optimization method for handling uncertainties in variables and parameters  
 4.4: Sustainable watershed management practices

**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

4.3: Robust optimization method for handling uncertainties in variables and parameters

4.4: Sustainable watershed management practices

We are here

# Related Optimization Research at MSU



## Point and population-based optimization and their combined use

A case study on a billion-variable resource allocation problem



## Multi-objective optimization (MO) and decision-making

- **Evolutionary multi-objective optimization (EMO)** algorithms to find multiple Pareto solutions
- **"Innovization"**: Knowledge extraction from learning from Pareto solutions
- **Multi-criterion decision-making (MCDM)** to choose a preferred Pareto solution

# Point and Population Based Optimization Methods: Adv and Disadv

Best for simplistic problems, difficult to modify for different problem classes

Ideal for local search

Single-obj Opt.:

Search for a point which minimizes an objective function satisfying constraints

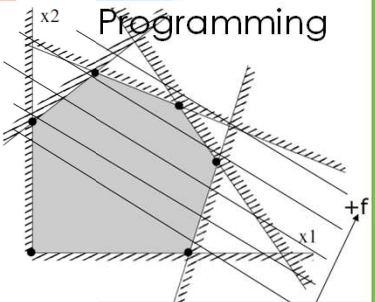
Point-Based

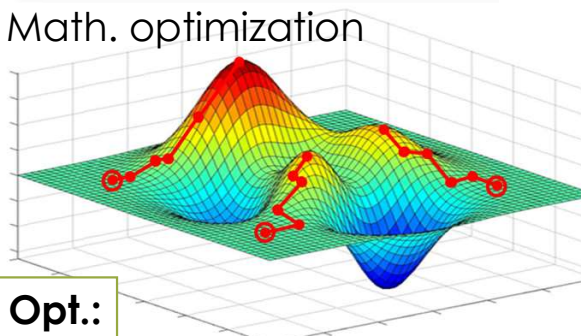
Nonlinear Programming

$$\nabla f(x) - \sum_{j=1} u_j \nabla g_j(x) - \sum_{k=1} v_k \nabla h_k(x) = 0$$

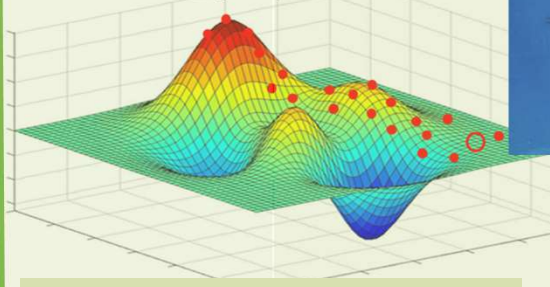
Linear Programming

$g_j(x) \geq 0 \quad j = 1$   
 $h_k(x) = 0 \quad k = 1$   
 $u_j g_j(x) = 0 \quad j = 1$   
 $u_j \geq 0 \quad j = 1$






Evolutionary optimization



Other Metaheuristics



## Population-Based



Best for complex problems, flexible to suit a problem class

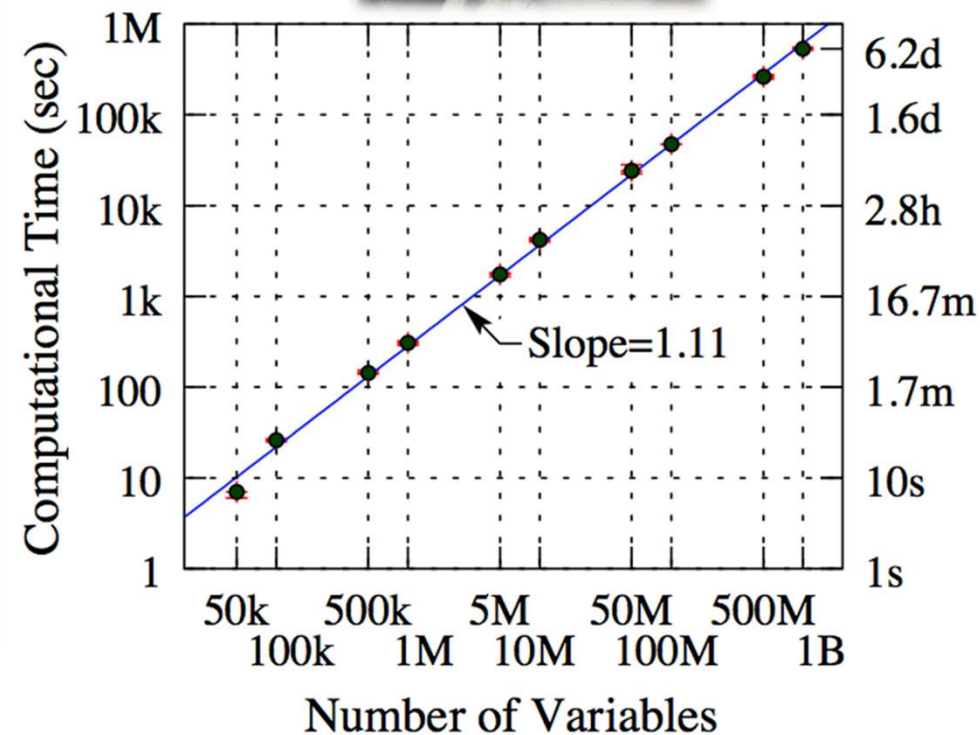
Ideal for global search

# Population-based Optimization:

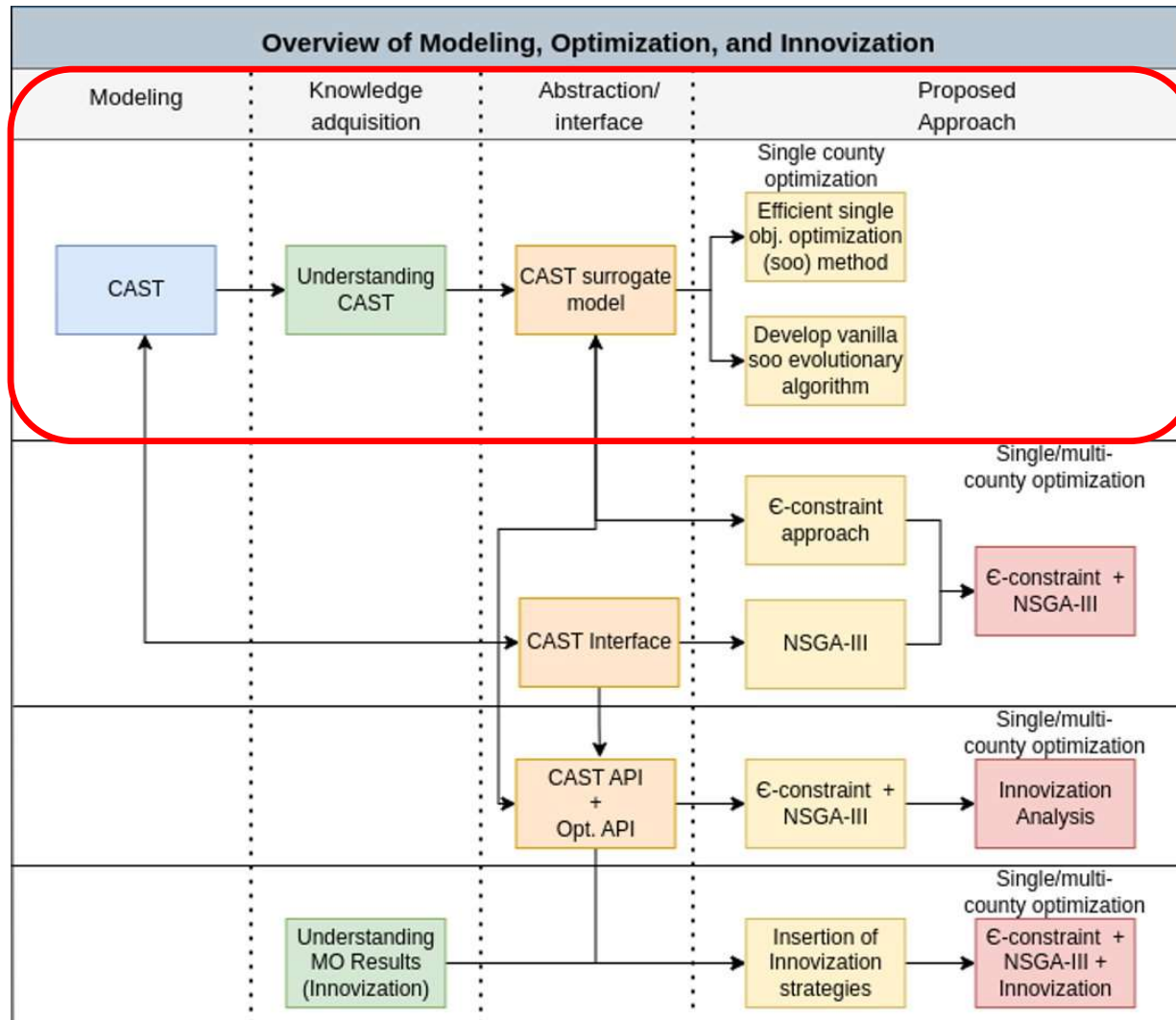
A Case Study on Resource Allocation Problem



- Casting scheduling problem requiring **50,000 variables**
- CPLEX software could not solve 2,000-variable version
- Customized Evol. Algorithm (population based) solves **Billion-variable** version
  - Polynomial time performance



Deb, K. and Myburgh, C. (2017). A Population-Based Fast Algorithm for a Billion-Dimensional Resource Allocation Problem with Integer Variables. *European Journal of Operational Research*, 261(2), 460–474.



# Chesapeake Bay Watershed Optimization Problem

## West Virginia Counties

County	#Variables	#Constraints	Base N <sub>2</sub> (f <sub>2</sub> <sup>base</sup> )
Berkeley	14,090	1,813	977,896
Grant	25,228	3,448	1,049,450
Hampshire	12,783	1,700	1,012,797
Hardy	18,607	2,491	1,344,295
Jefferson	12,303	1,606	1,018,012
Mineral	20,260	2,698	763,864
Monroe	3,102	399	48,655
Morgan	11,880	1,665	271,134
Pendleton	33,083	4,352	1,133,327
Preston	1,470	193	4,683
Tucker	1,012	144	1,702
<b>Total</b>	<b>153,818</b>	<b>20,509</b>	<b>7,625,818</b>

**Allocate a specific BMP to**

- Land River Segment (LRS)
- Agency
- Load source

**Minimize {Cost, Loadings}**

Large number of variables will require **large computational time**

**Surrogate model:**

**s: LRS**  
**h: Agency**  
**u: Load source**  
**b: BMP**

$$\text{Min. } f_1(\mathbf{x}) = \sum_{s \in S} \sum_{h \in H_s} \sum_{u \in U} \sum_{b \in B_u} \tau_b x_{s,h,u,b},$$

$$\text{Min. } f_2(\mathbf{x}) = \sum_{s \in S} \sum_{h \in H_s} \sum_{u \in U} \left[ \alpha_{s,h,u} \phi_{s,h,u} \prod_{G^B \in \mathcal{G}^B} \left( 1 - \sum_{b \in G^B} \eta_{s,h,b}^N \frac{x_{s,h,u,b}}{\alpha_{s,h,u}} \right) \right],$$

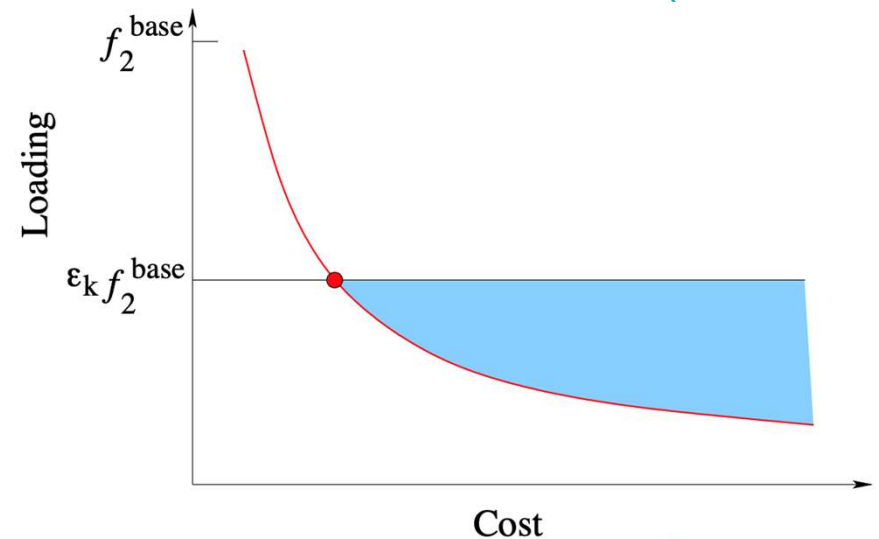
$$s.t. \quad \sum_{b \in G^B} x_{s,h,u,b} = \alpha_{s,h,u}, \quad \forall s \in S, h \in H_s, u \in U_s, G^B \in \mathcal{G}^B,$$

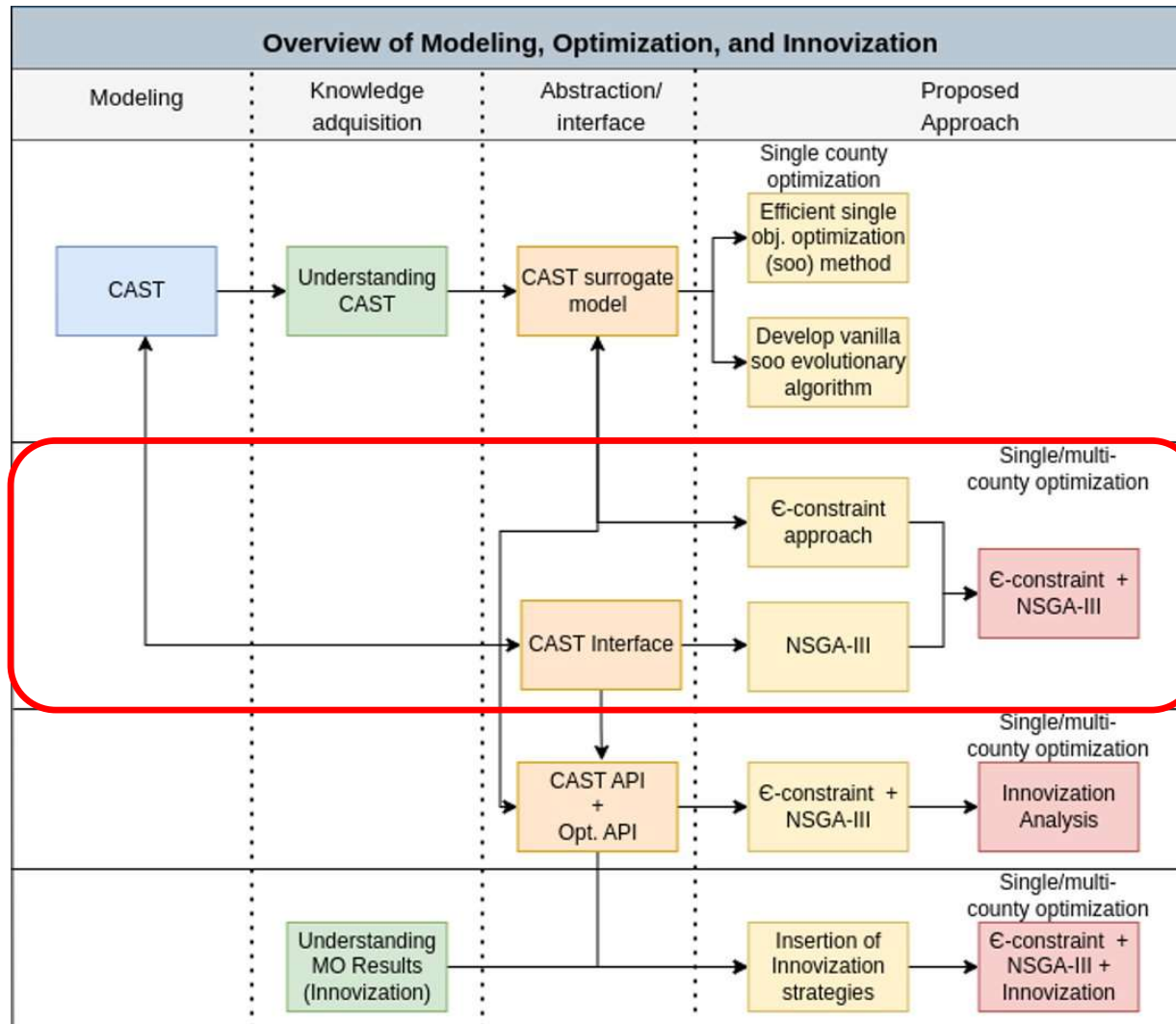
$$x_{s,h,u,b} \geq 0, \quad \forall s \in S, h \in H_s, u \in U_s, b \in B_u.$$

# Converting Multiple Objectives Into One

- Convert second objective into a constraint
- **Epsilon-Constraint method**
  - Vary  $\epsilon_k$  to generate a set of trade-off solutions

$$\begin{array}{ll} \text{Minimize} & f_1(\mathbf{x}), & (\text{Cost}) \\ \text{Subject to} & f_2(\mathbf{x}) \leq \epsilon_k f_2^{\text{base}}, & (\text{N2}) \\ & \mathbf{x} \in \mathbf{X}, & \end{array}$$



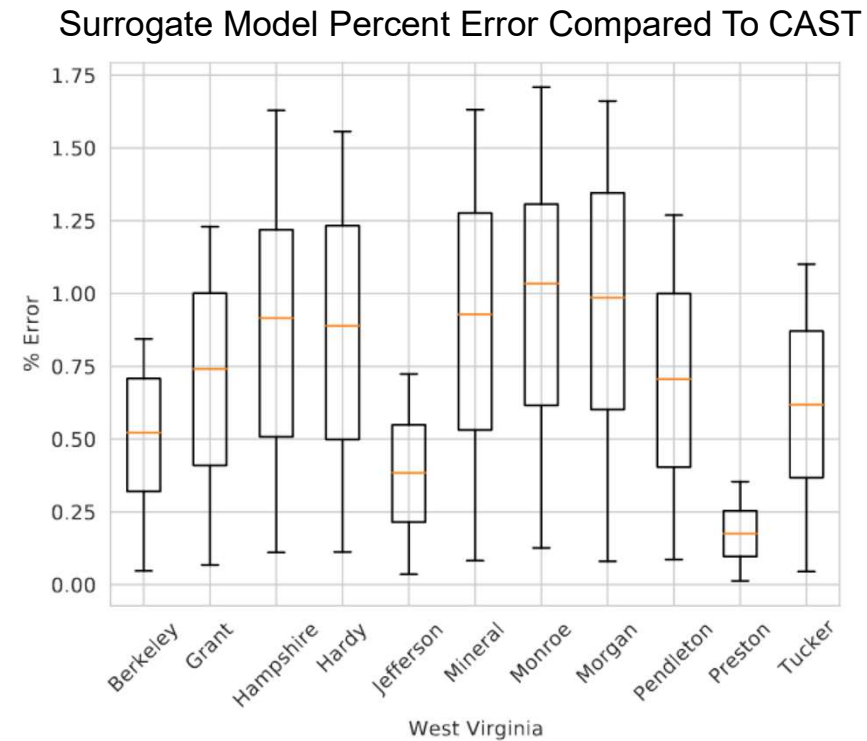


# Surrogate Model Error

- **10,000 BMP scenarios** are evaluated using **surrogate model and CAST** on West Virginia counties

- Observed **small error** in Nitrogen loading value

- **Supports the use of surrogate model based optimization, if needed**

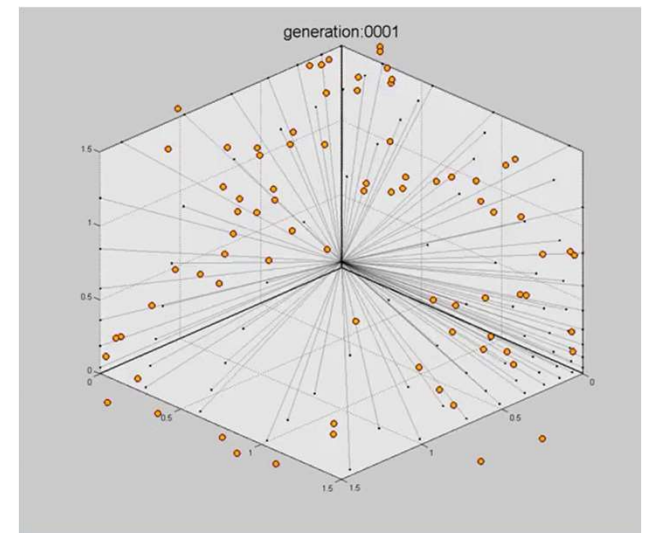
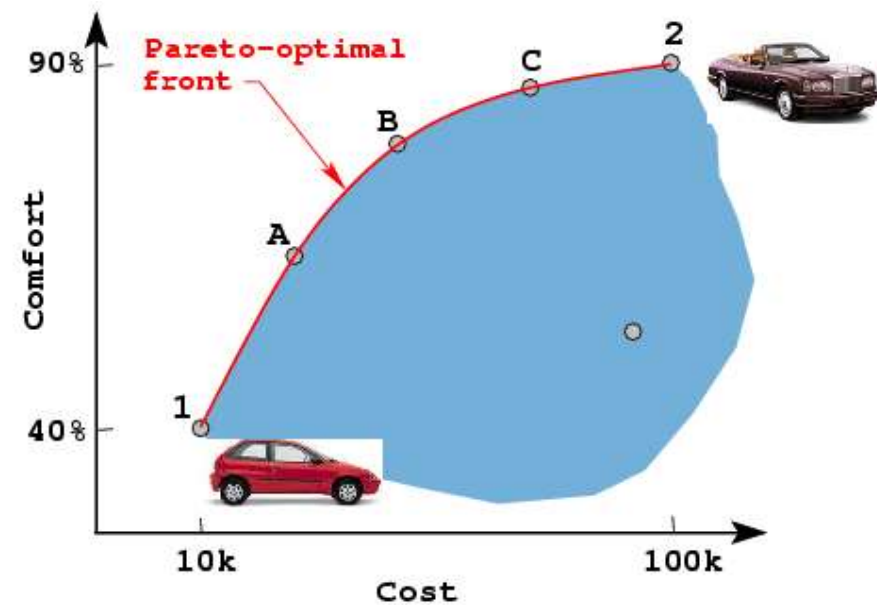


# Multi-Objective Optimization

Results in a set of Pareto-optimal solutions

- **Step 1:** Find multiple trade-off solutions
- **Step 2:** Choose a preferred solution

- **Evol. Multi-objective optimization (EMO)**
- NSGA-III can handle 2-15 objectives with constraints

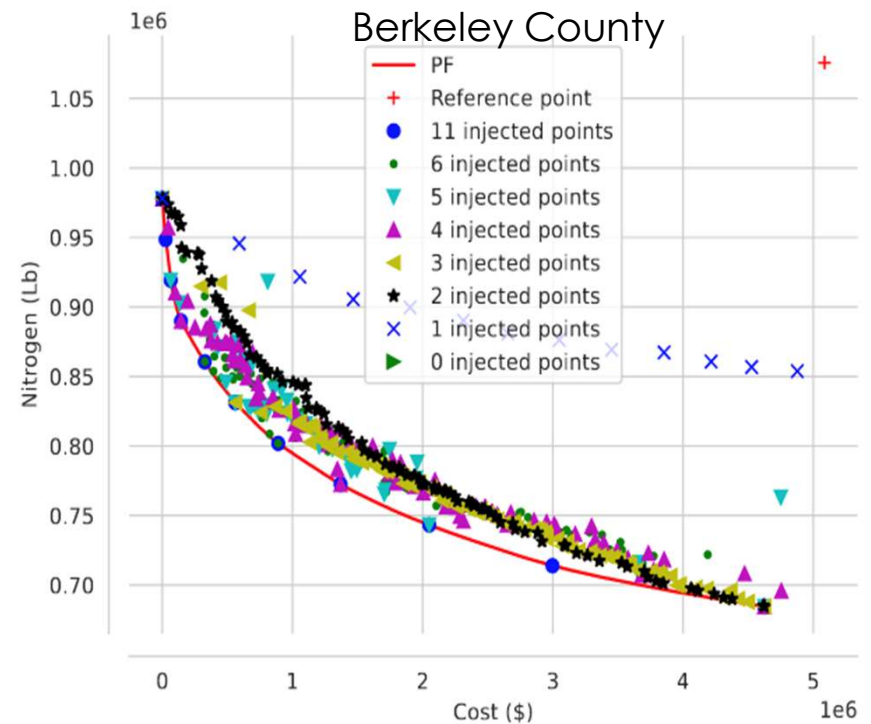


# Customized NSGA-III for CBW Problem

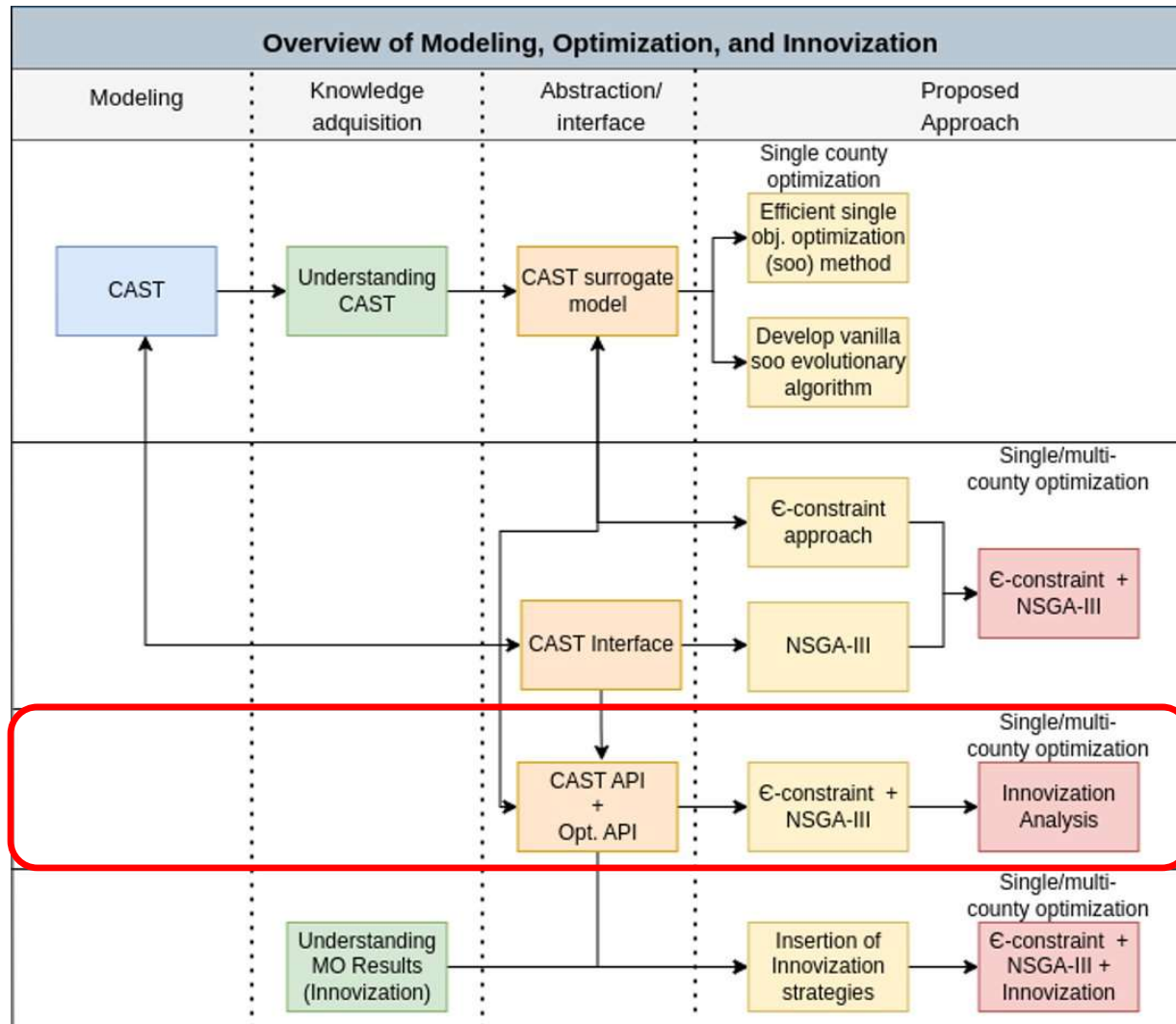
- NSGA-III initialized with **Eps-Constraint solutions**
- Repair operator to fix constraint violation
- Optimize surrogate model

## Major finding:

At least 3 injected solutions make NSGA-III efficient



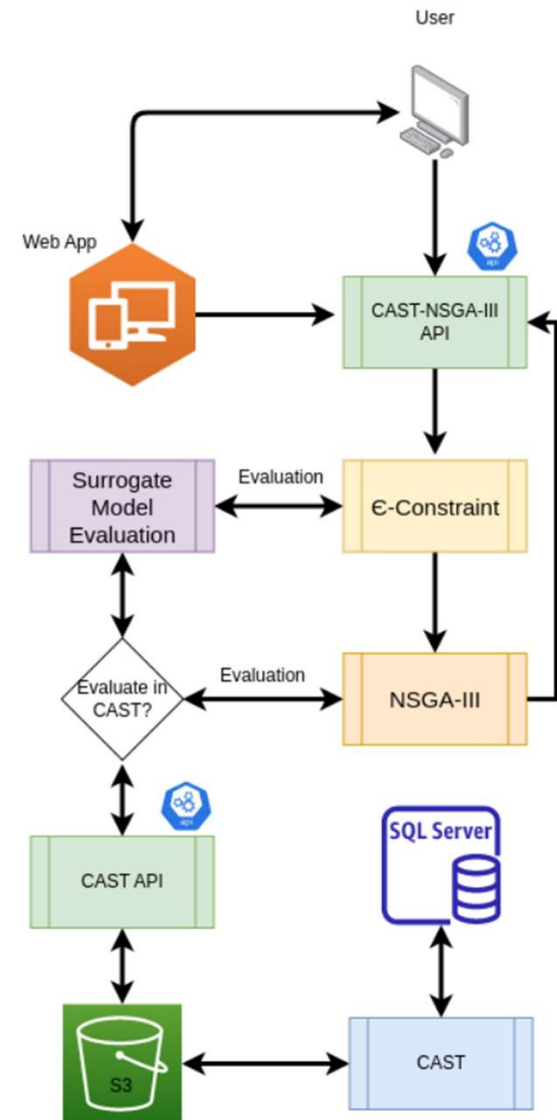
Toscano, G., Hernandez-Suarez, J. S., Blank, J., Nejadhashemi, P., Deb, K. and Linker, L. (2022). Large-scale Multi-objective Optimization for Water Quality in Chesapeake Bay Watershed. Proceedings of 2022 Congress on Evolutionary Computation (CEC-2022), IEEE Press. (pp. 1–8). **BEST PAPER AWARD**



# API-based Linking of CAST with NSGA-III

## Automatic Programming Interface (API)

- Allows multiple users with different programming environments to interact
- Makes application modular



# NSGA-III Linked with CAST

NSGA-III calls CAST to evaluate using Restful APIs

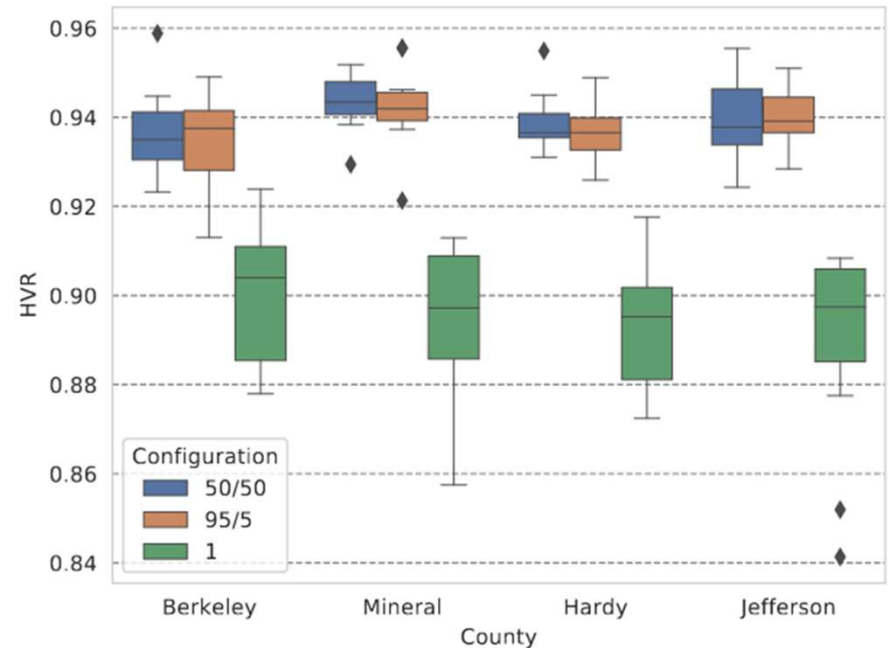
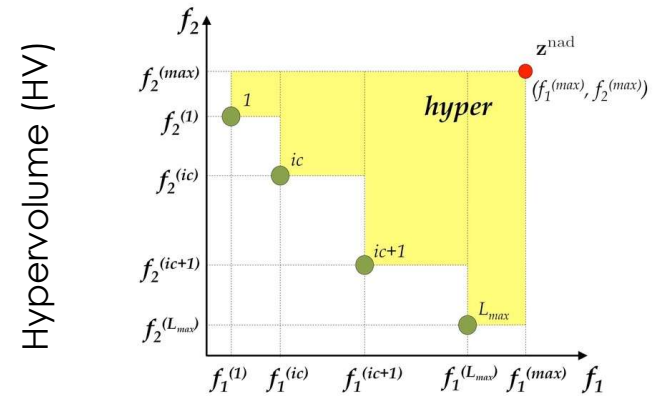
## Mixed Heterogeneity:

- **50/50**: 50% surrogate, 50% CAST
- **95/5**: 95% surrogate, 5% CAST
- **1**: All surrogates, evaluate final solutions by CAST

Time (min)	50/50	95/5	1
<b>Best</b>	138.90	17.53	2.70
<b>Worst</b>	152.96	20.55	3.80
<b>AVG</b>	143.79	18.20	3.18
<b>STD</b>	2.61	0.69	0.33

Major finding:

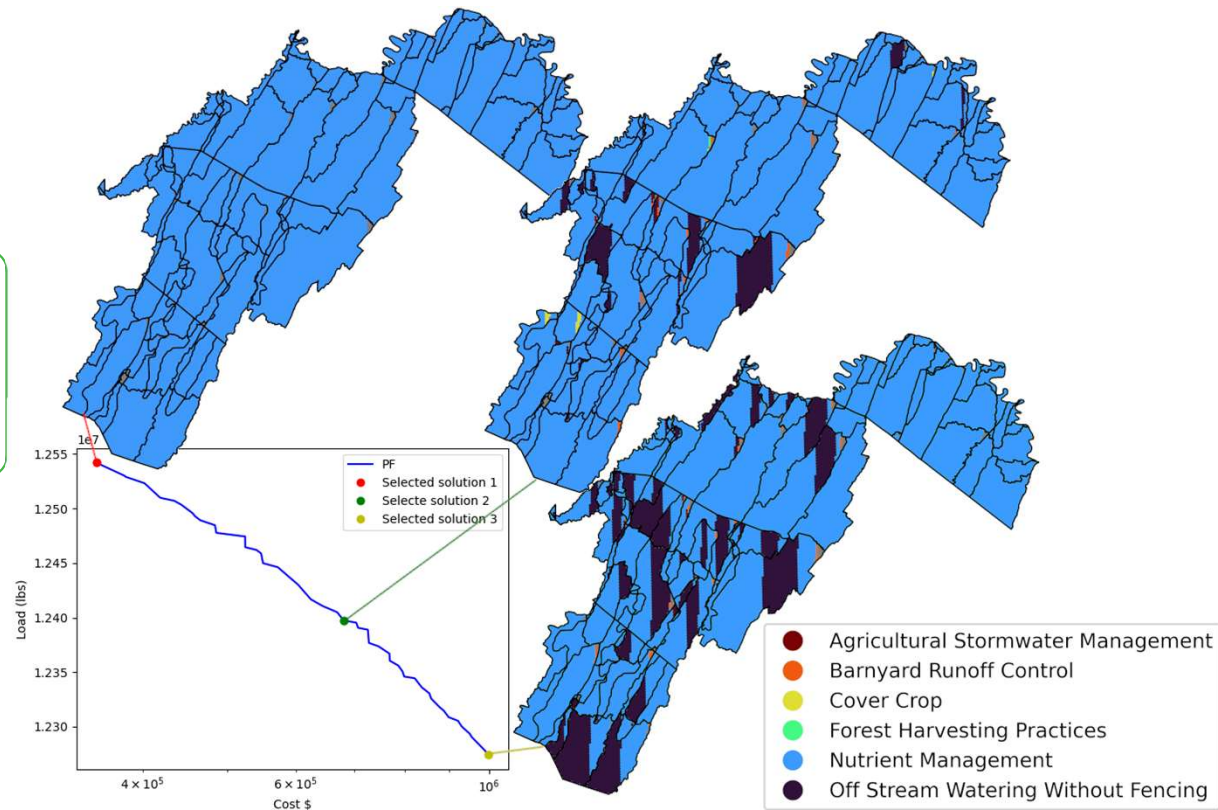
95/5 is almost as good but requires less time



# Alternate Solutions Using Multi-objective Optimization

Evaluate alternate solutions before picking a single preferred solution

Analyze solutions for extracting knowledge for future use



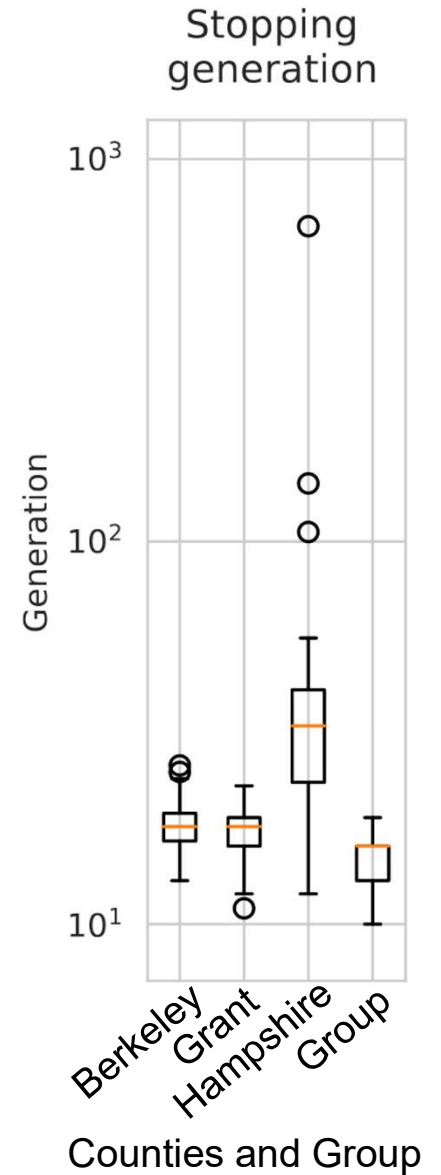
# Multi-County Optimization

## Observation:

- Some combinations of counties make the problem easier to solve compared to individual counties
  - Ex: Group gets optimized faster than Hampshire county

## Major finding:

This provides promise for extending the proposed optimization method to multi-state and to watershed level



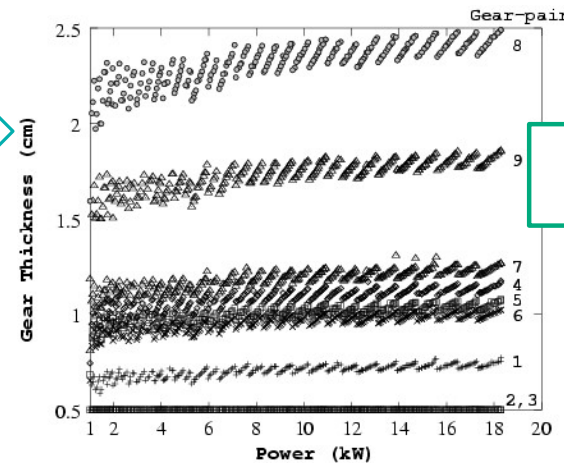
# “Innovization”

## Innovation through Optimization

Finding **patterns** in Pareto-optimal solutions for knowledge discovery using AI/ML methods

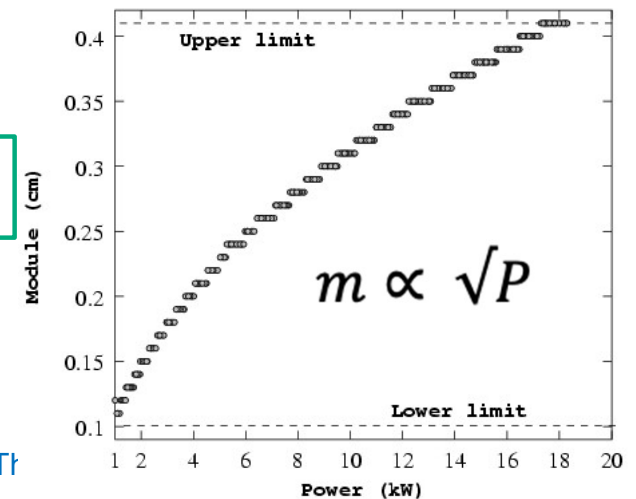
- Optimality conditions must manifest as common patterns
- Extract patterns and re-optimize

A Gear-box design with 28 variables

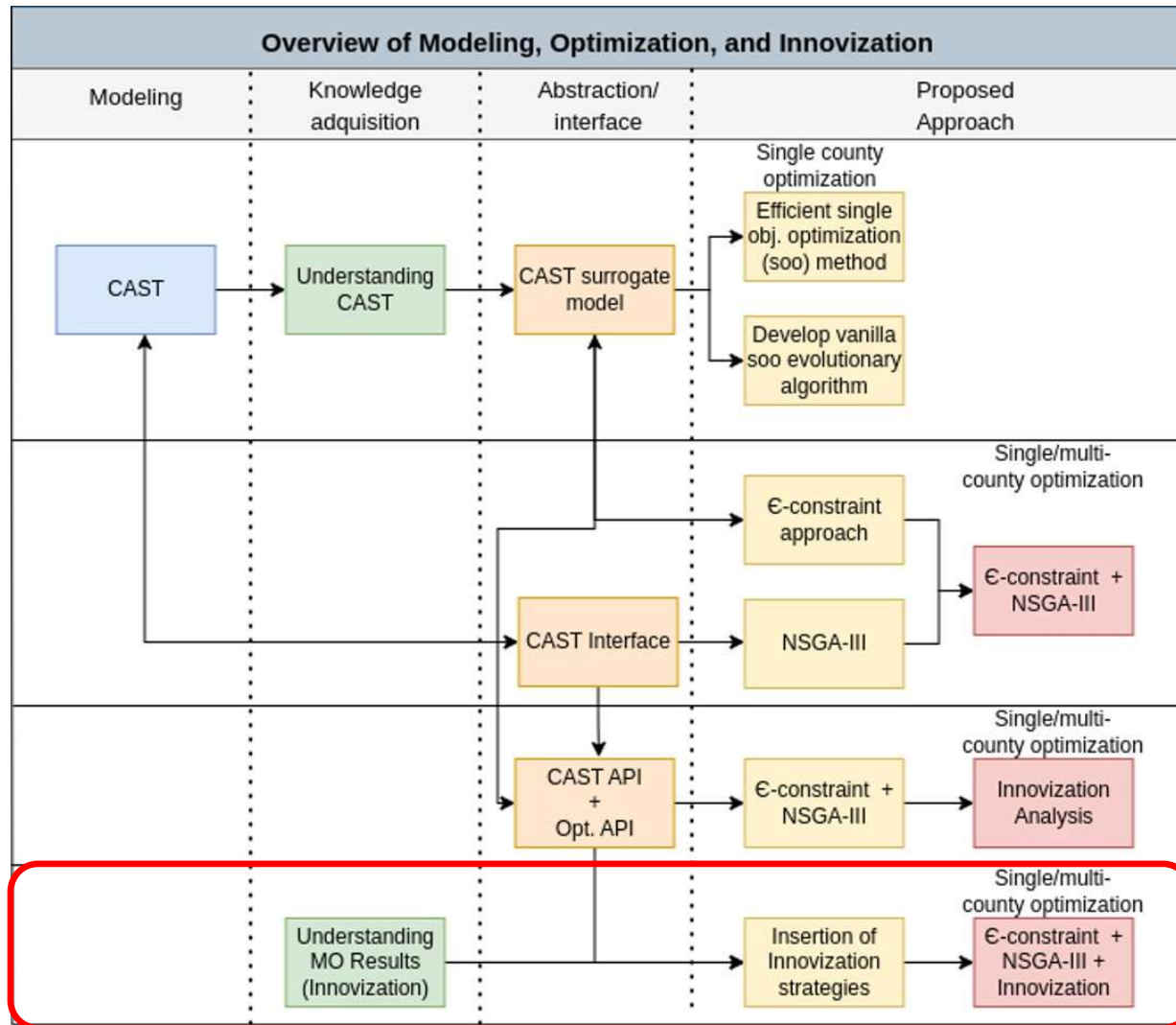


Not much change in 27 variables

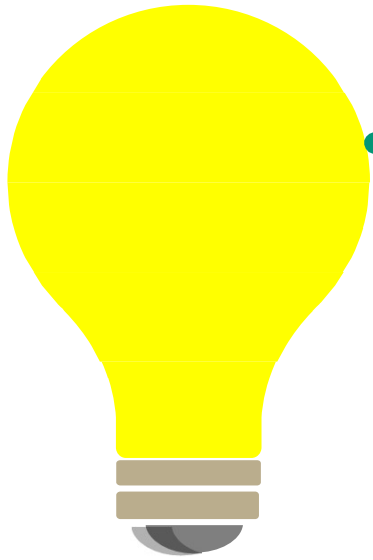
Large change in one variable



Deb, K. and Srinivasan, A. (2006). *Innovization: Innovating design principles through optimization. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO2006)*, New York: Tl Association of Computing Machinery (ACM), (pp. 1629–1636)



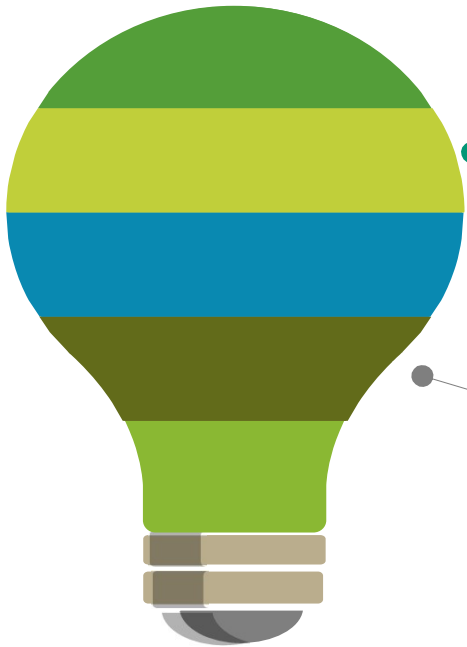
# Knowledge Discovery Using Optimization



What are the benefits of optimization?

- Identify the **best solutions** for the problem in hand.
- **Generating knowledge** to solve future problems.

# Innovization Analysis



## • What is innovization?

1

Learning from optimization results and introducing new ideas, products, and services different from the existing ones.

## What innovization can do to CBPO?

2

- Provide information for better decision-making for BMP selections (**farmers**)
- Identify the high priority areas for BMP implementation (**regulators**)
- Help with resources allocation (**policymakers**)

# Methodology

## BMP Selection ranking methodology based on Land use:

- *Overall goal:* learn from optimization results to

1

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

2

Identify the **similarities and differences** between top-ranked BMPs,

3

Provide recommendations to **reduce the optimization time**

# Methodology



BMP Selection ranking methodology based on Land use:



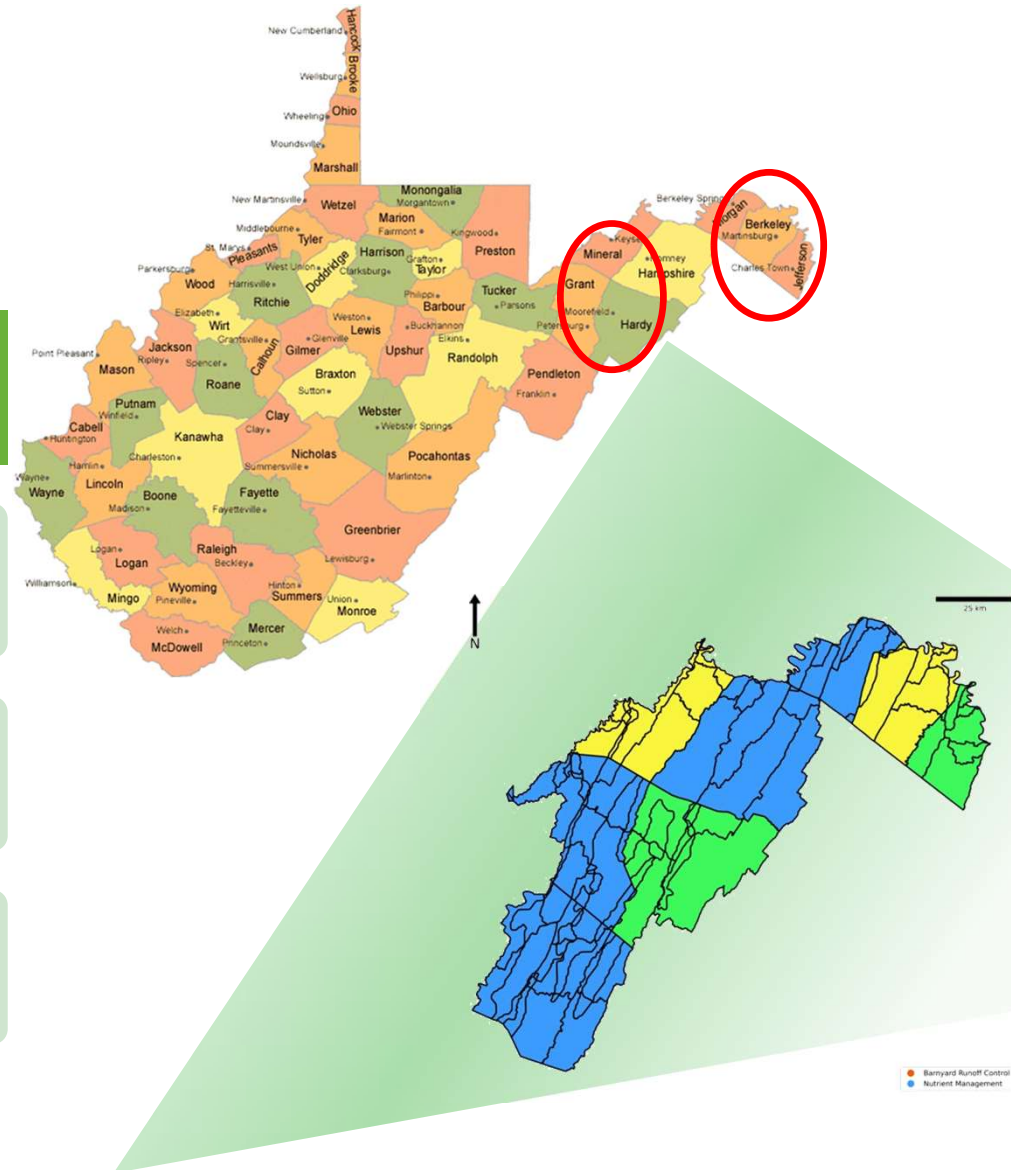
In **West Virginia**, we identified the top two counties with the highest areas of urban and agricultural land uses.



**(Berkeley and Mineral)**: Urban dominated



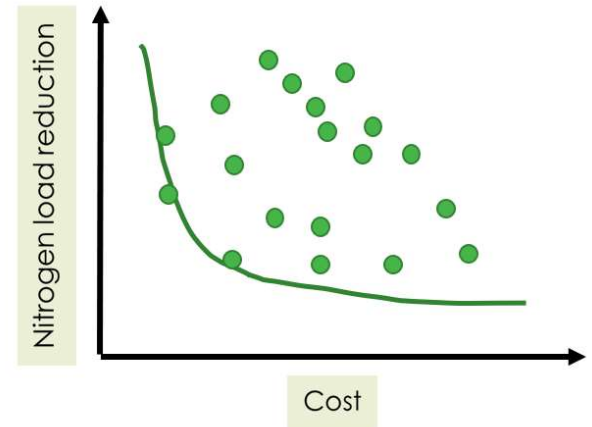
**(Jefferson and Hardy)**: Agricultural dominated



# Methodology



**BMP Selection ranking methodology based on Land use:**



1

Running CAST-optimization algorithm resulted in 220 solutions for **each county**

2420 solutions for 11 counties in about thousands land river segments.

2

3

Identified the **best solutions from optimization.**

# Methodology

1

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

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



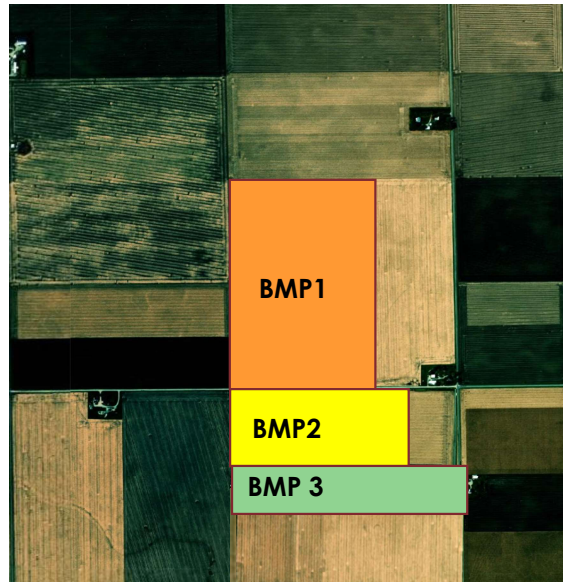
# Methodology

1

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

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

**BMP1**  
**BMP2**  
**BMP3**



# Methodology

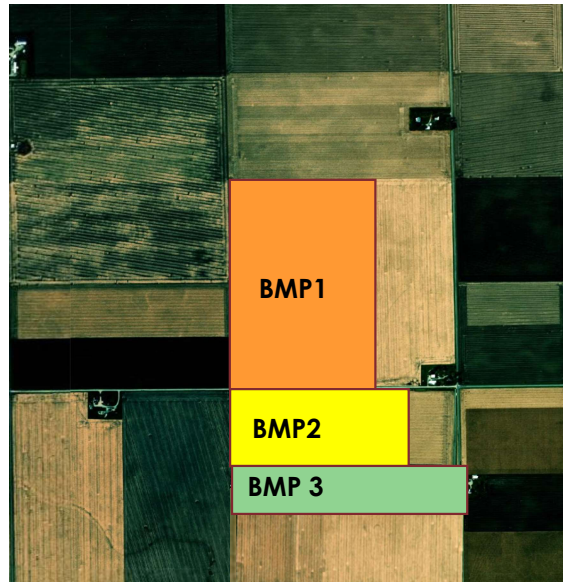
1

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

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

Ranking methodology 2):

**BMP3**  
**BMP2**  
**BMP1**



# Methodology

1

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

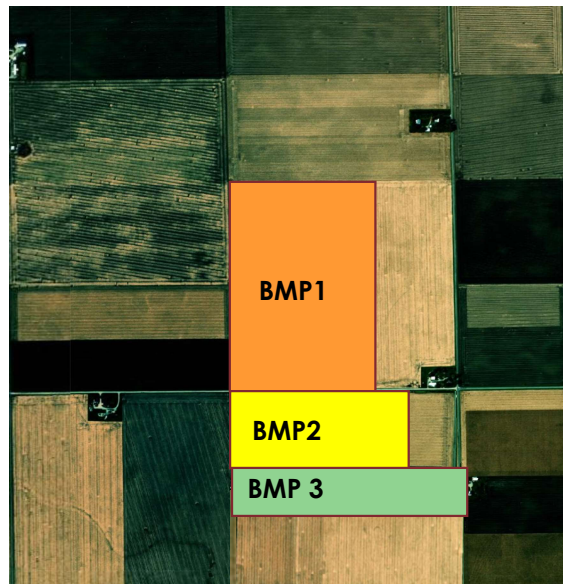
- **Ranking methodology 1)** rank the top BMPs based on the **implementation acreages;**
- **Ranking methodology 2)** rank the top BMPs based on the percentage of **maximum allowable acreages;**
- **Ranking methodology 3)** rank the top BMPs based on the amount of **nitrogen reduction per dollar spent.**

Ranking methodology3):

**BMP2 (\$12/lb N)**

**BMP3 (\$15/lb N)**

**BMP1(\$24/lb N)**



# Methodology

1

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

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

Ranking methodology1):

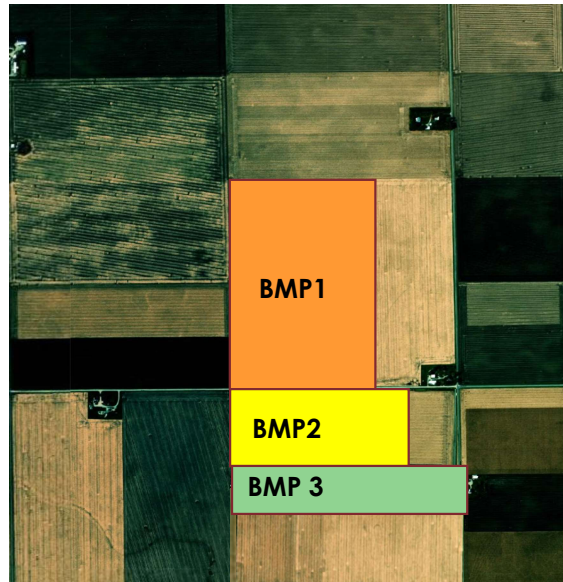
**BMP1**  
**BMP2**  
**BMP3**

Ranking methodology2):

**BMP3**  
**BMP2**  
**BMP1**

Ranking methodology3):

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



# Results:

2

Identify the **similarities and differences** between top-ranked 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:

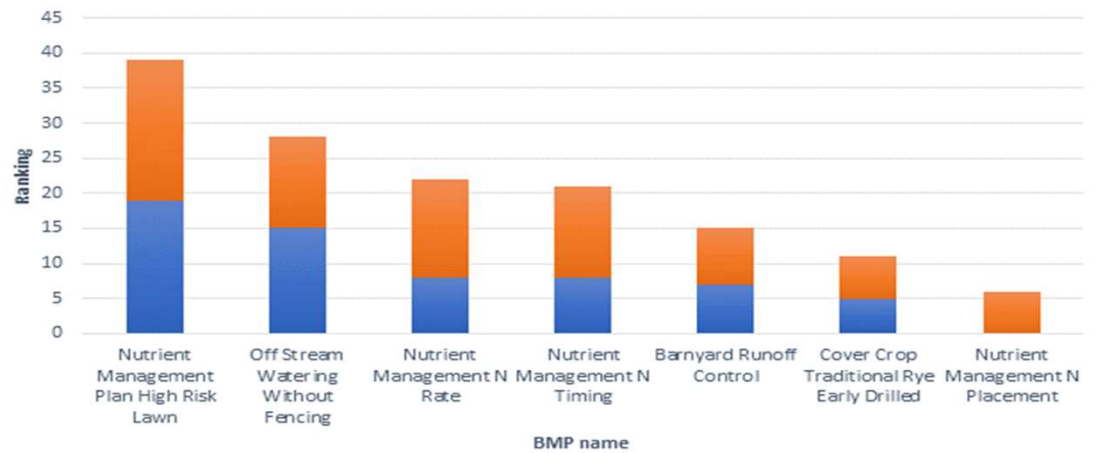
# Results:

3

Provide recommendations to **reduce the optimization time**



Obtaining the total ranking of each BMP by adding the associated ranking to individual BMPs.



# Results:

2

Identify the **similarities and differences** between top-ranked BMPs,

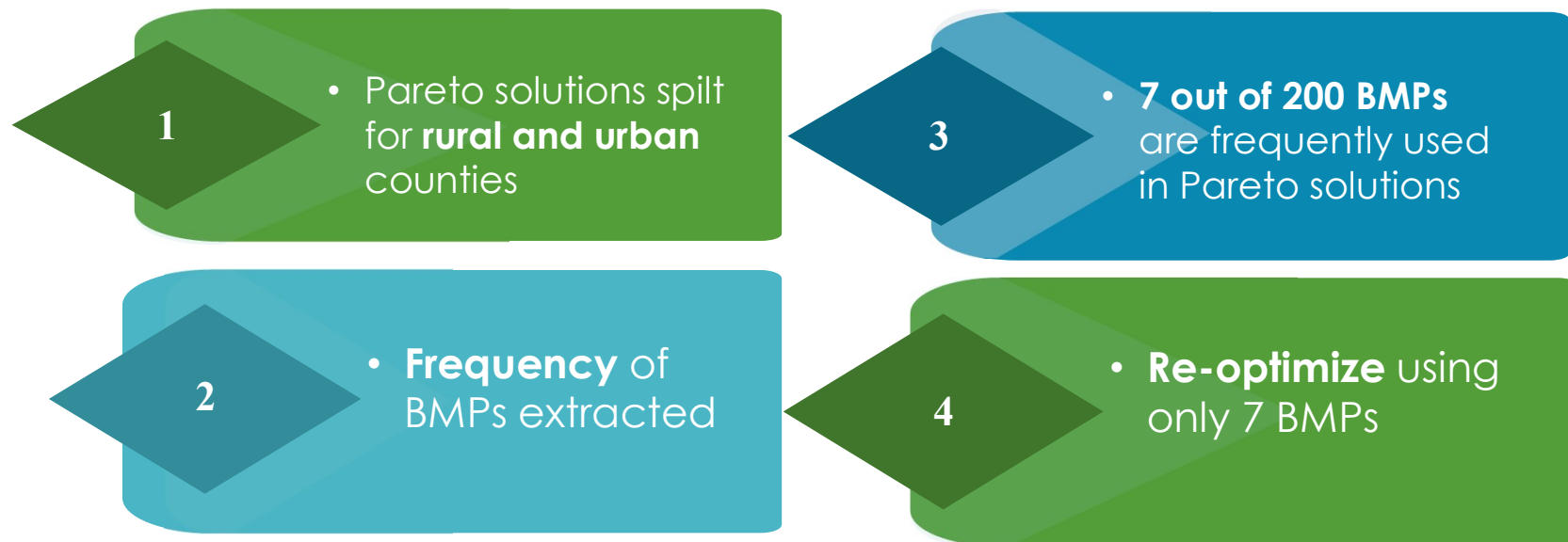
## Similarities:

**Top BMP choice:** Nutrient management lawn or farm  
**The pasturelands:** Off-stream watering facilities

## Differences:

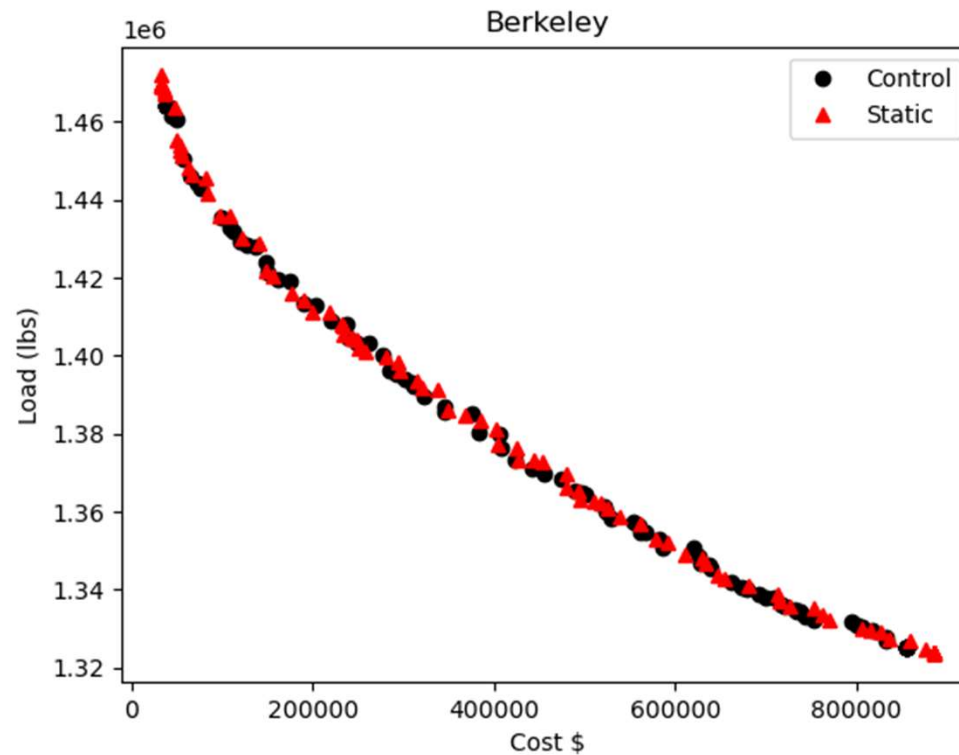
More diversity in BMP types was in agricultural settings compared to urban ones.

# “Innovization” Study on CBW Problem Variable Reduction

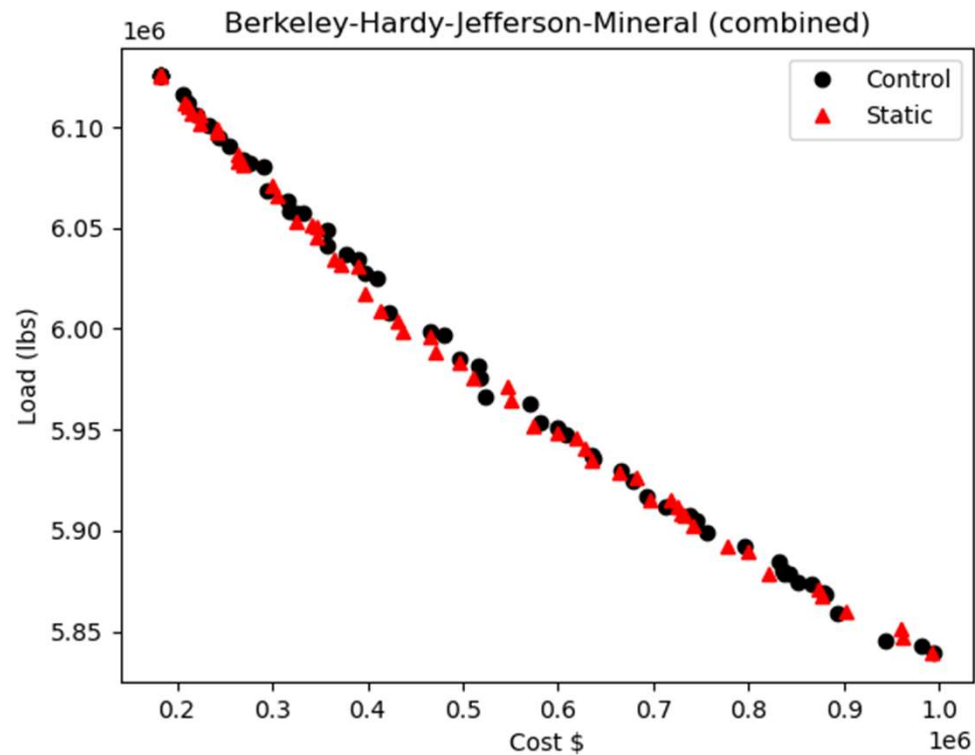


	Control	Static
<b>Berkeley</b>	14,090	510 (3%)
<b>Hardy</b>	18,607	725 (3%)
<b>Jefferson</b>	12,303	456 (3%)
<b>Mineral</b>	20,260	765 (3%)

# Reoptimization Using Innovization Results on Berkeley County



# Reoptimization Using Innovization Results on Berkeley, Hardy, Jefferson, and Mineral Counties (Combined)



# Results:

3

Provide recommendations to **reduce the optimization time**



Recommend the selection of **the top seven BMPs from the overall column** for optimization.

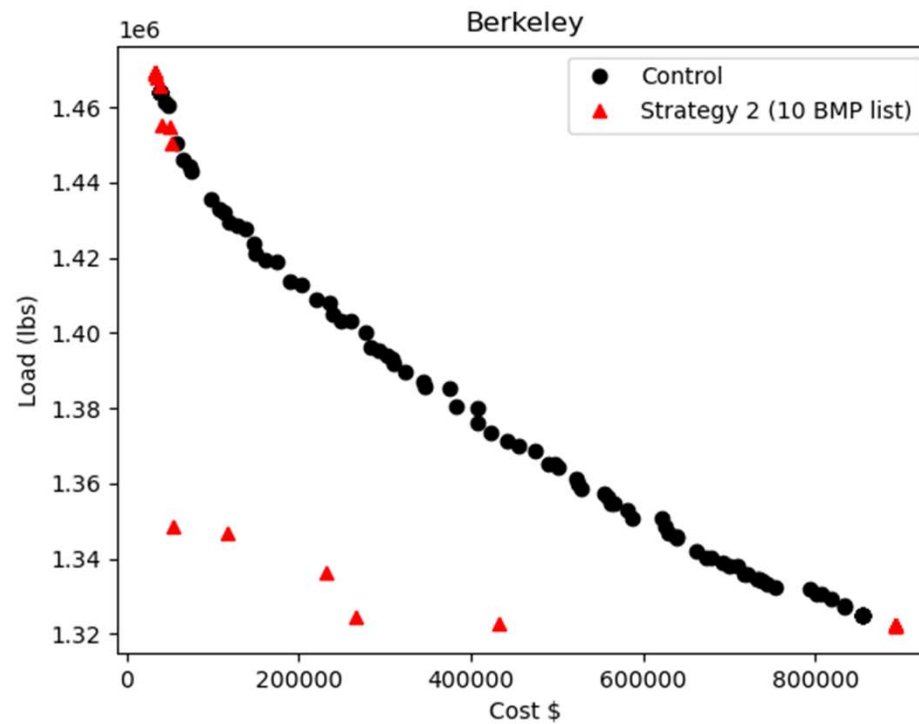


Can be used in developing the initial population in other counties within the state of West Virginia.



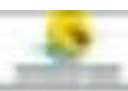
**Hypothesis:** this approach could significantly reduce the **optimization processing time** while producing more cost-effective BMP implementation plans.

# Reoptimization Using Innovization Results on Berkeley County



# Prototype Interactive Web Tool

- 1 • Input-output through **the web portal** (Done)
- 2 • Collects scenario for optimization (Done)
- 3 • NSGA-III is invoked and calls CAST for evaluation (Done)
- 4 • Calls Decision-making Dashboard for analysis of Pareto solutions (Remaining)
- 5 • Re-optimize using "**Innovized**" principles until satisfied (Partially Done)



Decision Making Tool for  
the Chesapeake Bay  
Program developed by  
Michigan State  
University



# Decision-Making Methods

## Decision-making:

A systematic approach to pick a preferred solution

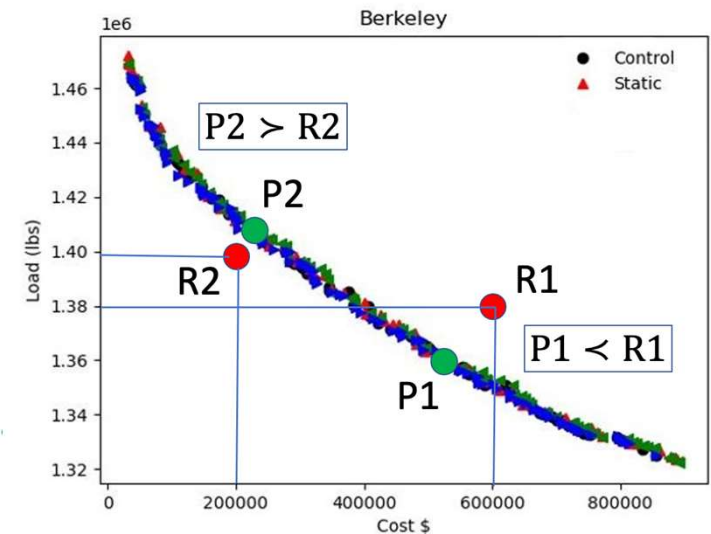
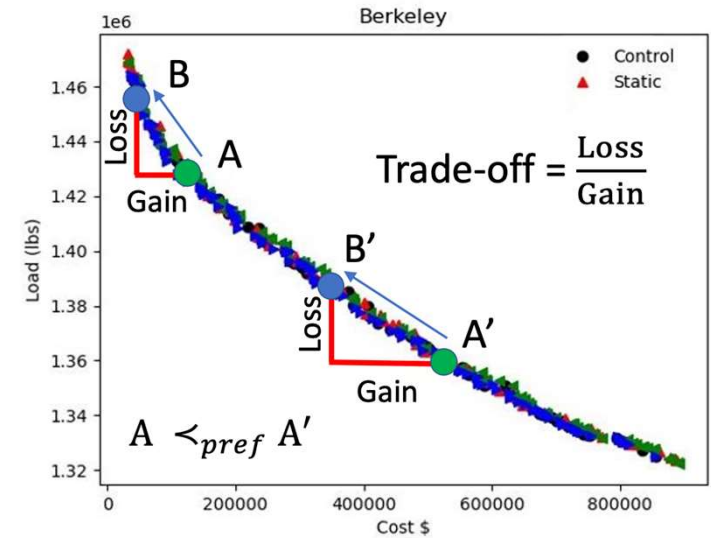
**A-posteriori** Trade-off analysis

**A-priori** Aspiration based approach

**Interactive** EMO-MCDM

DM provides preference information during optimization

Deb, K., Sundar, J., Reddy, Uday, B., and Chaudhuri, S. (2006). Reference point based multi-objective optimization using evolutionary algorithms. *International Journal of Computational Intelligence Research (IJCIR)*, 2(6), 273–286.



# Remaining Tasks

Completion of all BMP types:  
**Land conversion, Animal, Manure transport, etc.**

**Multi-criterion Decision-making (MCDM)** to choose a single solution

Converted Oxygen optimization to combine **multiple loadings**

Dashboard for interactive applications

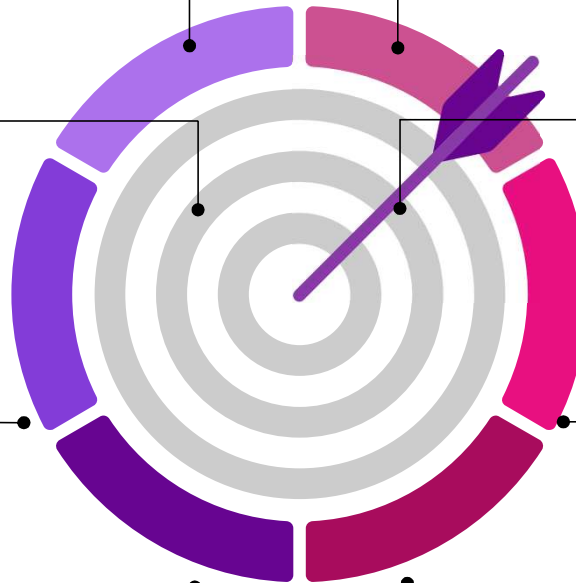
- A partial framework is completed

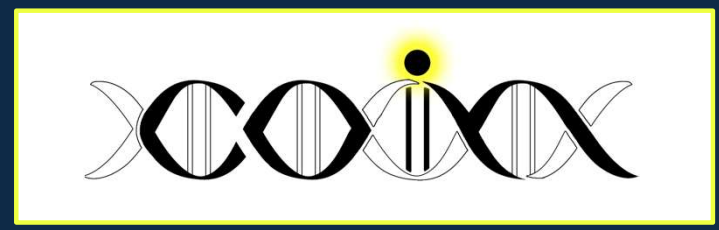
More than two-objective optimization

Harnessing hardware parallelism

**Scale-up** study to multi-state and watershed level optimization

Demonstration through workshops and tutorials





Computational Optimization and Innovation

**Thank you**

