15TH NATIONAL CONFERENCE ON
Transportation Asset Management
MONDAY–THURSDAY
August 25–28, 2025
CHICAGO, ILLINOIS

Hybrid Data Processing Approach for Overcoming Data Challenges in AI-powered Pavement Preservation:

An Iowa DOT Case Study

Alireza Sassani
Sara Arezoumand
Omar Smadi
Ashley Buss

IOWA STATE UNIVERSITY
Institute for Transportation

Roads

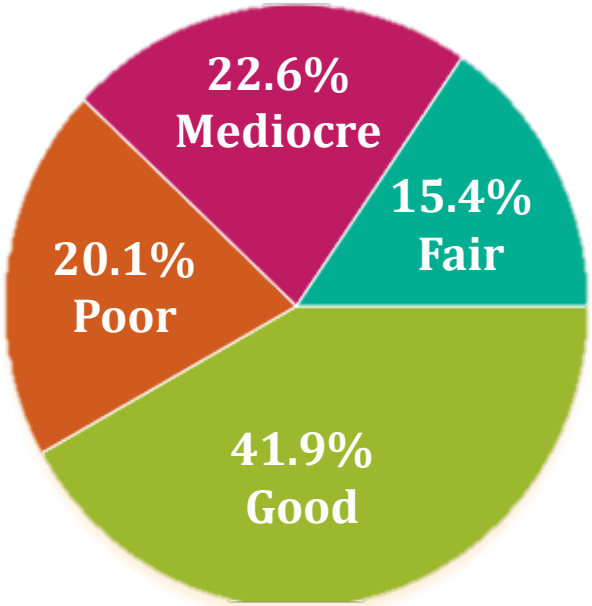
D+

Pavements in the U.S.

Over **4 million** miles of public roadways in the United States

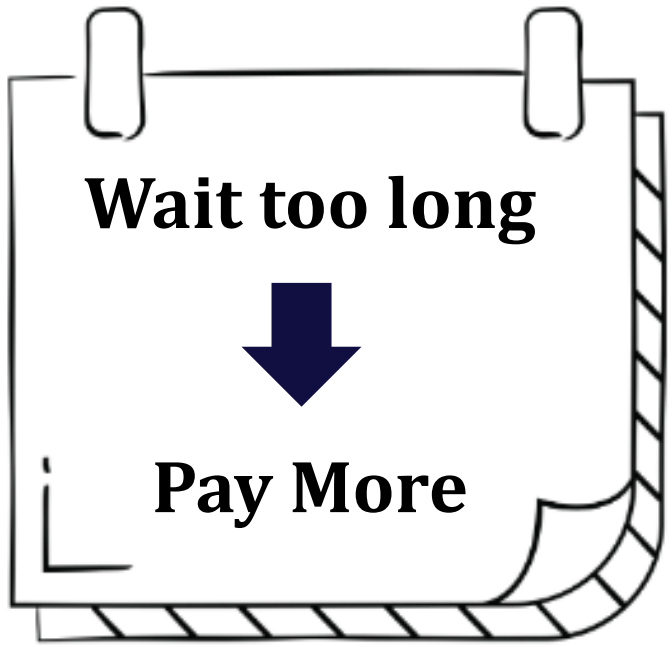
Deteriorated roads add nearly **$130 billion** each year to vehicle repairs and operating costs.

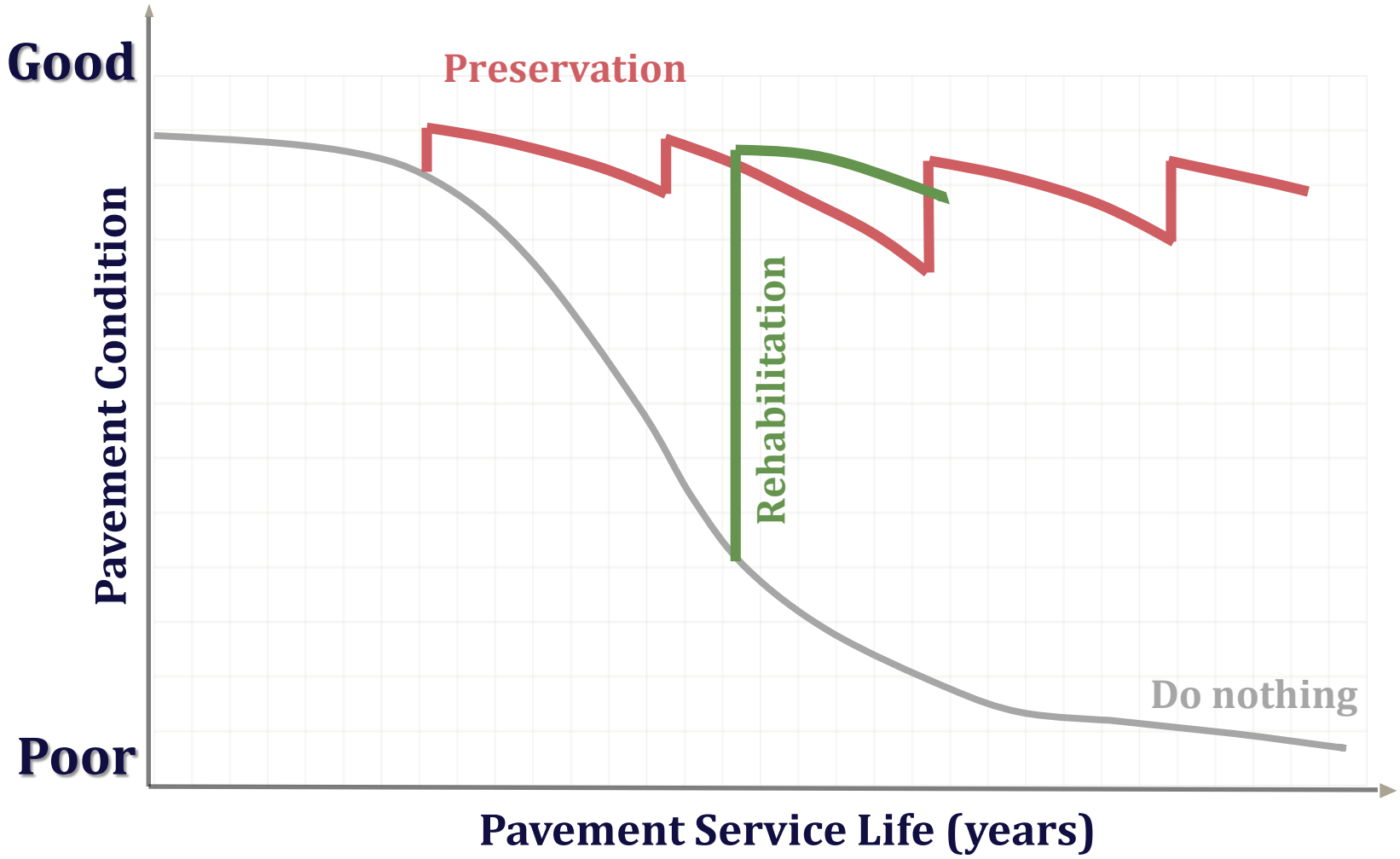
Only ***42%*** of the pavements in good condition



2021 ASCE Infrastructure Report Card

Maintenance and Rehabilitation





Introduction

Treatments

- Right pavement
- Right time
- Right treatment

https://www.pritchettcartoons.com/dot2.htm

Methodology - Phase I

Data Source & Scope

- Pavement Management Information System (PMIS)-Iowa DOT (1998-2020)
- Bid Express software -Cost data(2015-2022)

Analysis Approach:

- Robust regression (RANSAC) to handle outliers.
- Benefit–Cost Analysis (BCR, EUAC) for cost-effectiveness.
- Developed decision matrices by pavement type, pavement condition, and truck traffic

Indicators of Effectiveness:

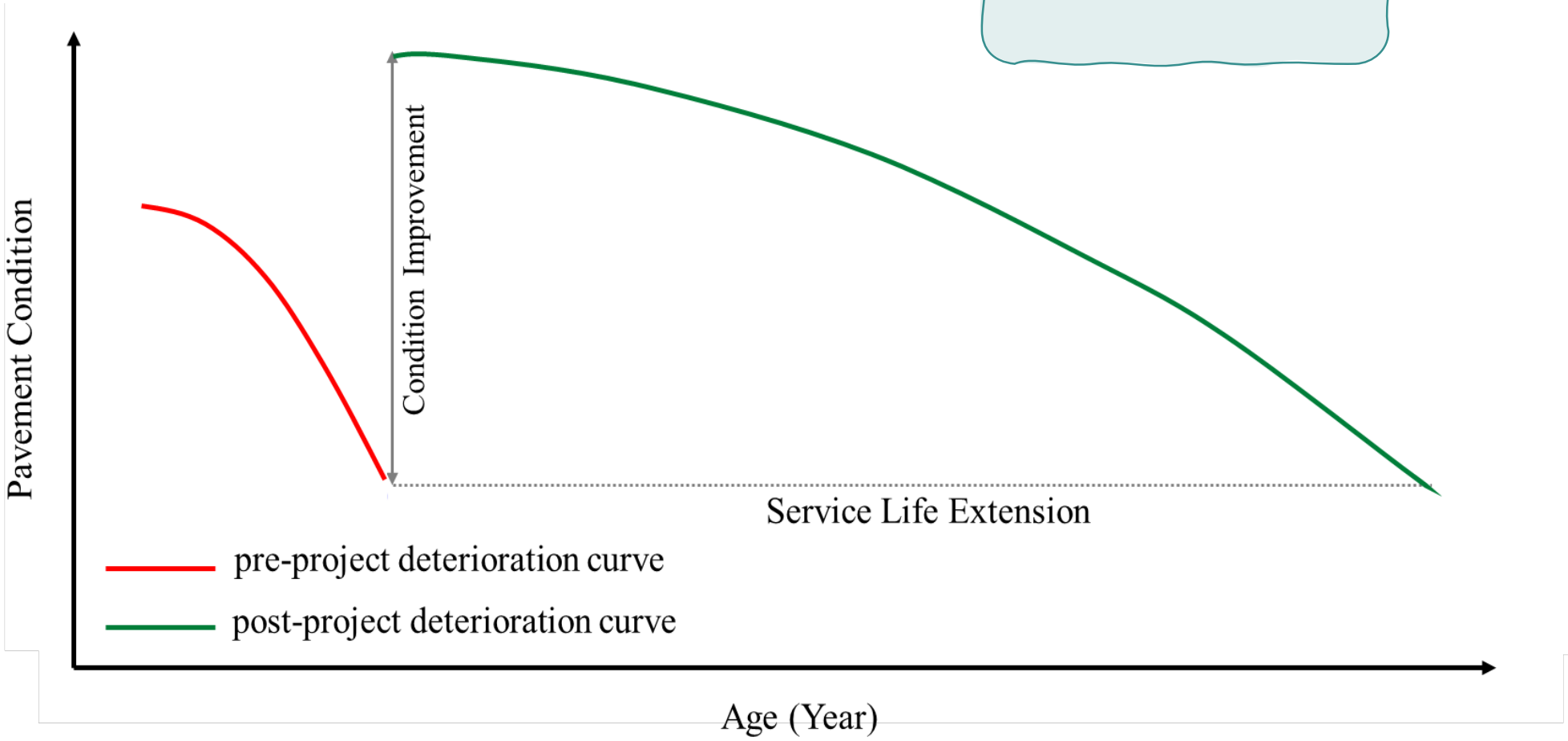
- Life Extension (LE): years of service life gained.
- Index Benefit (IB): area between pre- and post-treatment deterioration curves.

Methodology

Pavement Performance Modeling

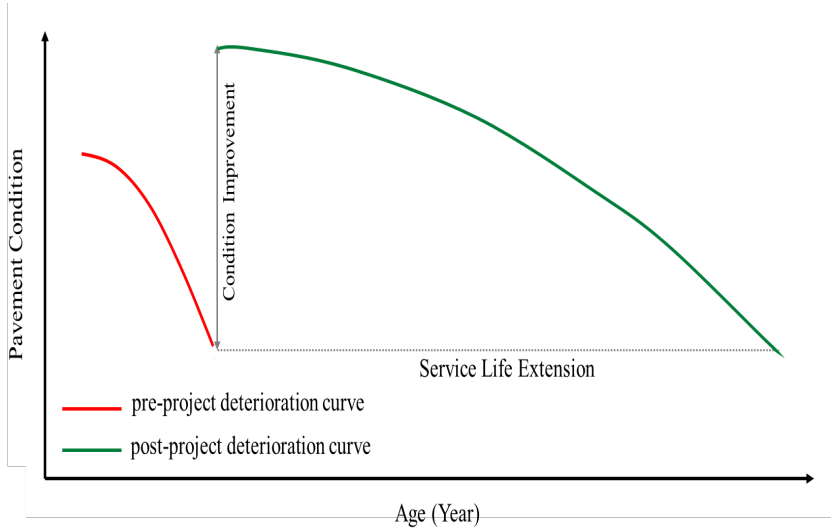
On the new PMIS

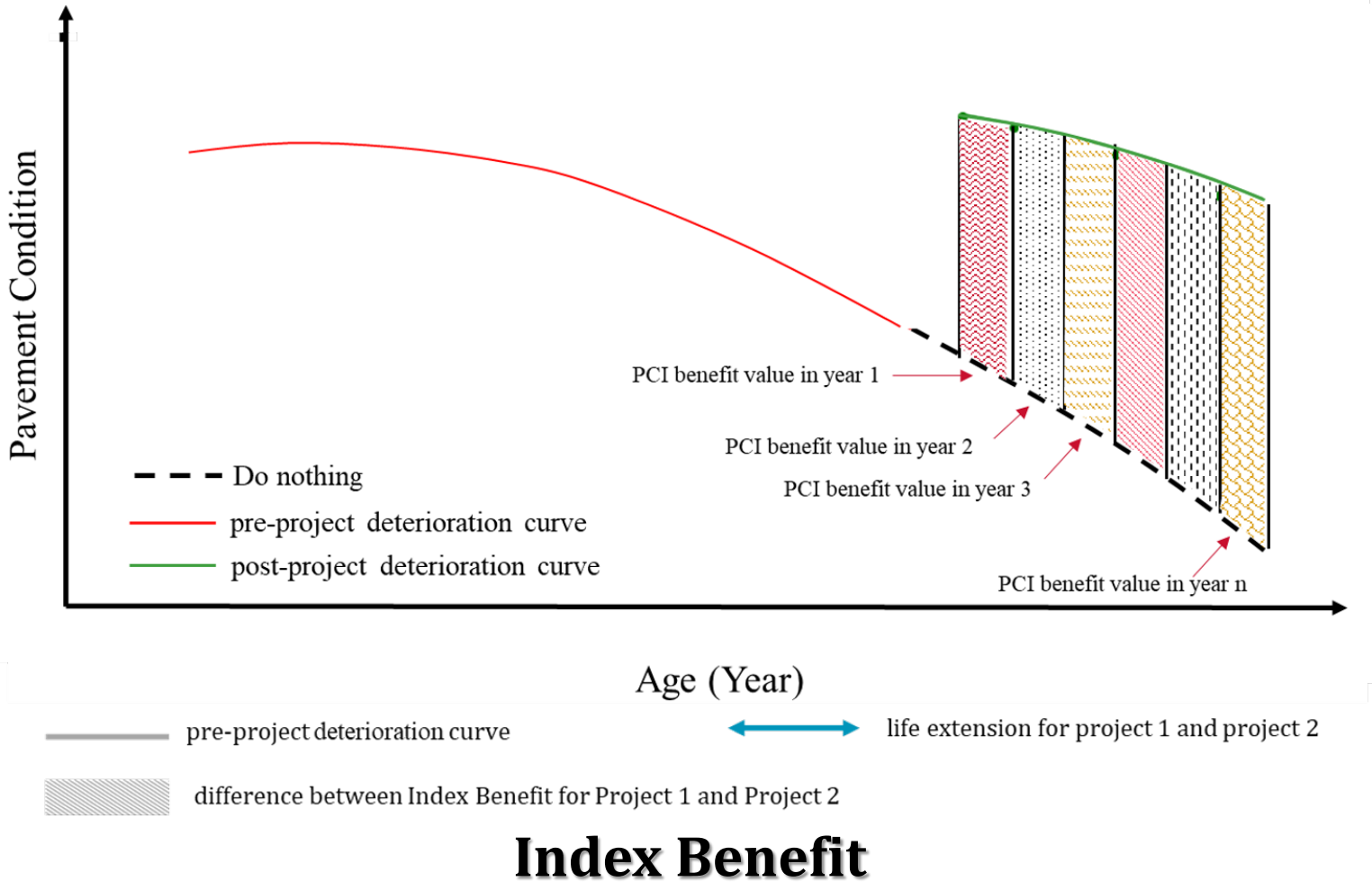
Life Extension



Methodology

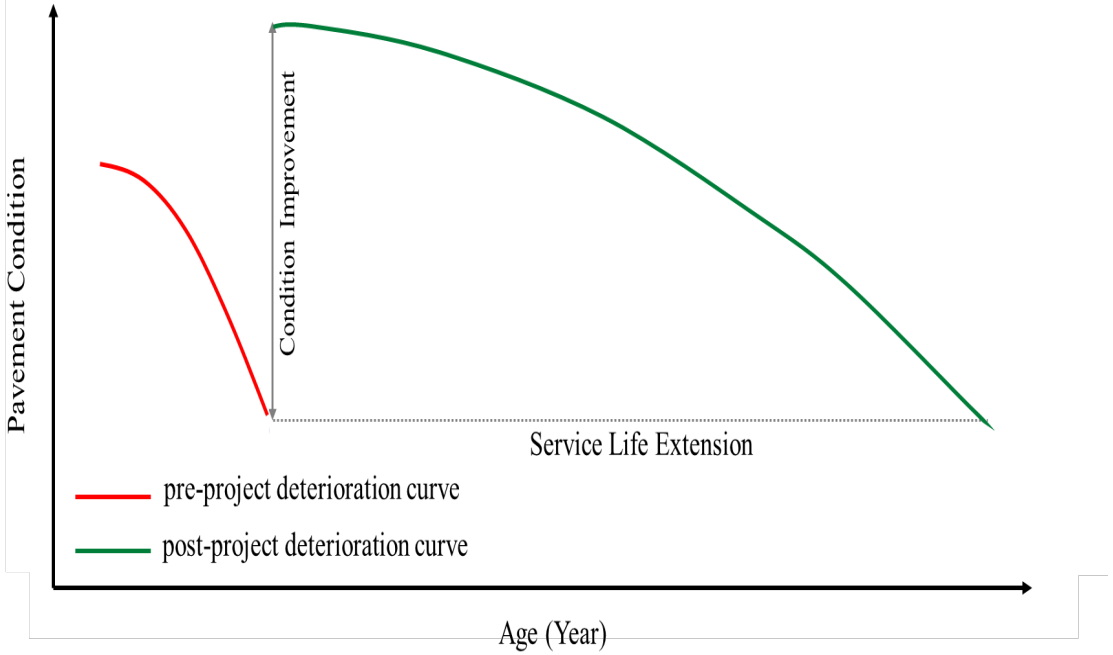
Performance Modeling



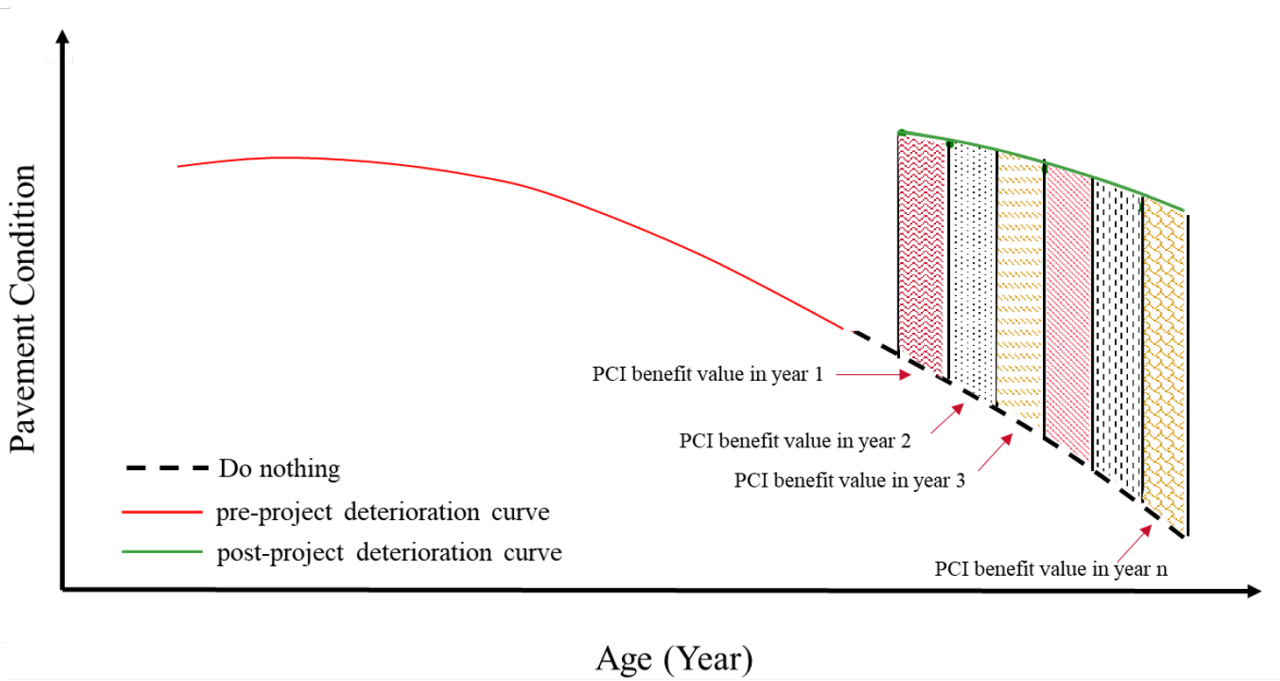


Index Benefit

Methodology



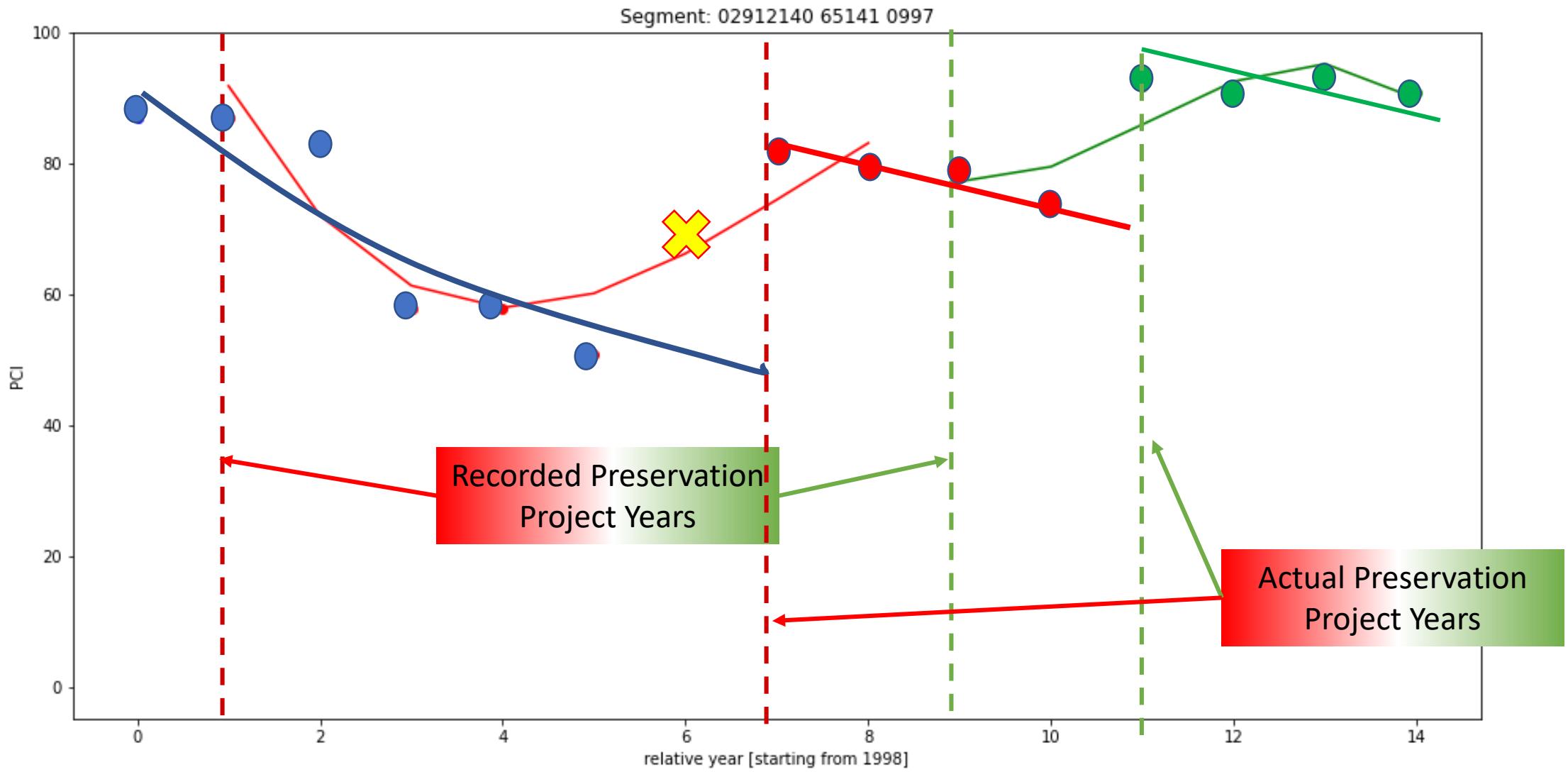
Life Extension



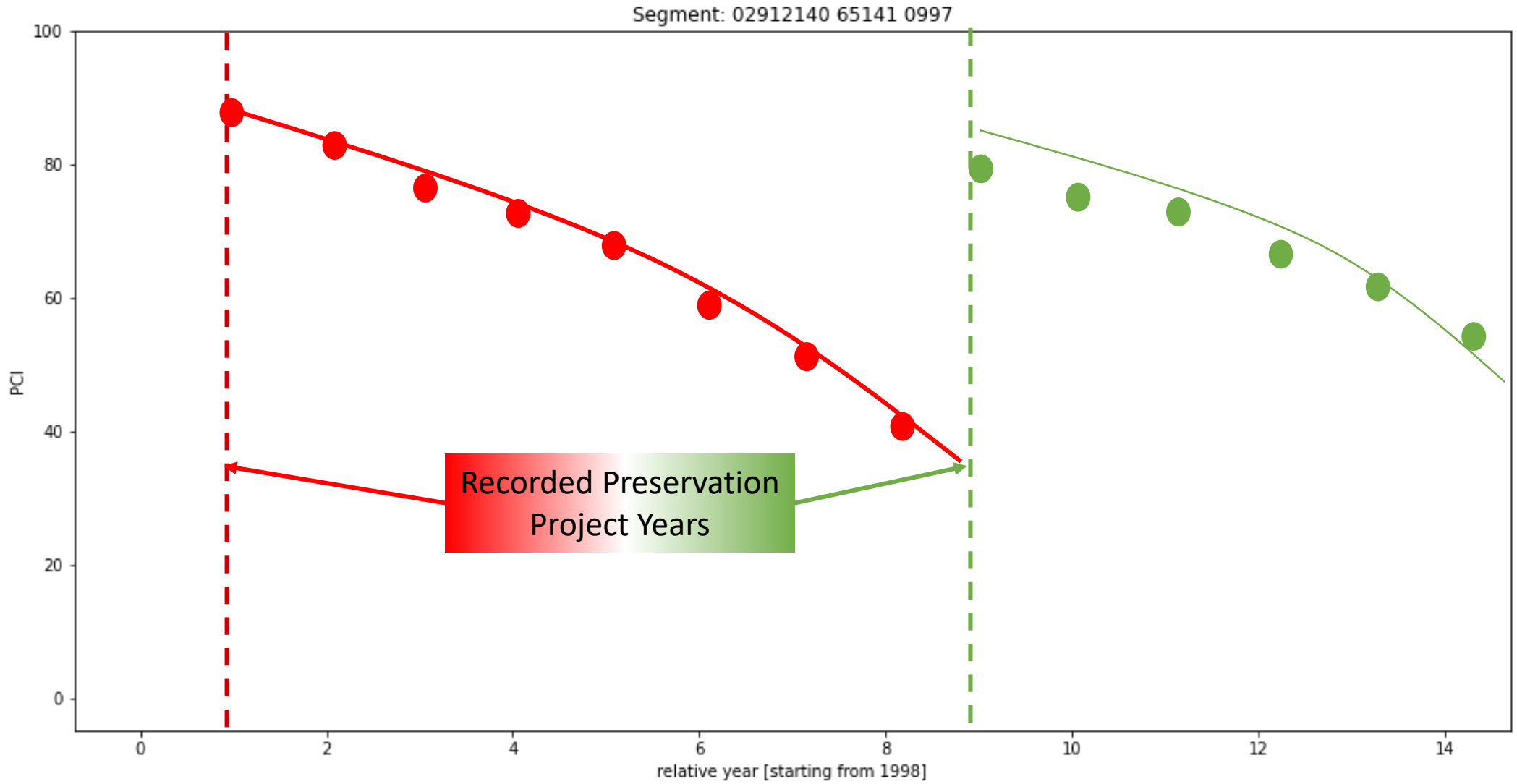
Index Benefit

- ☑ **Treatment's Effectiveness**
- ☑ **Right Time**

Methodology



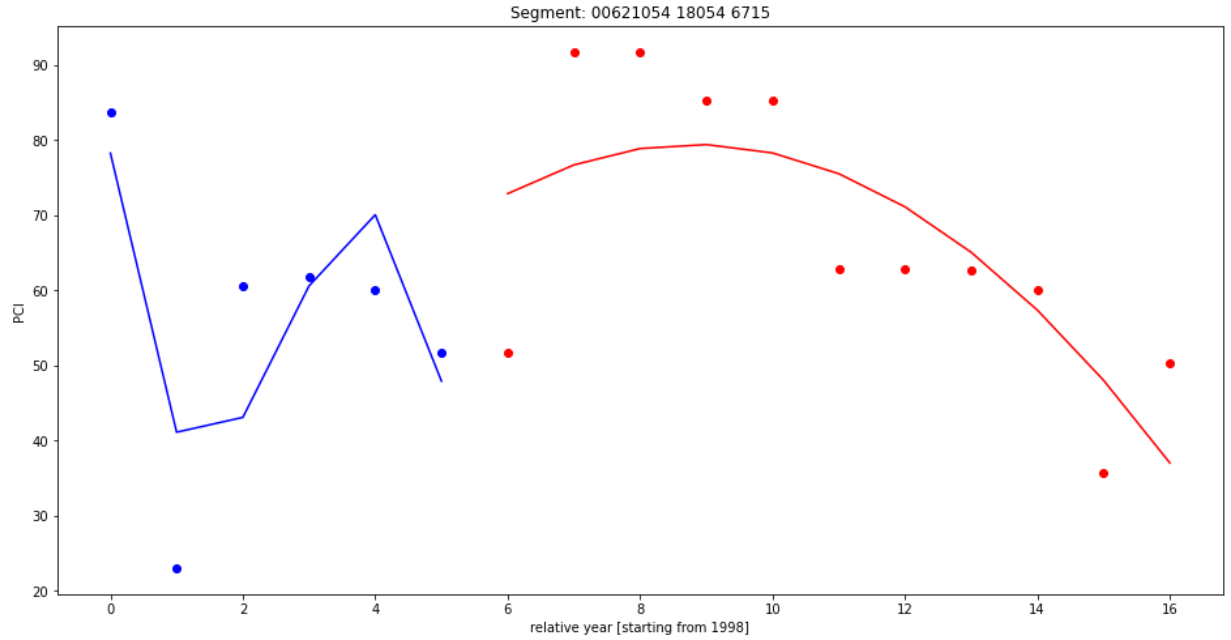
Methodology



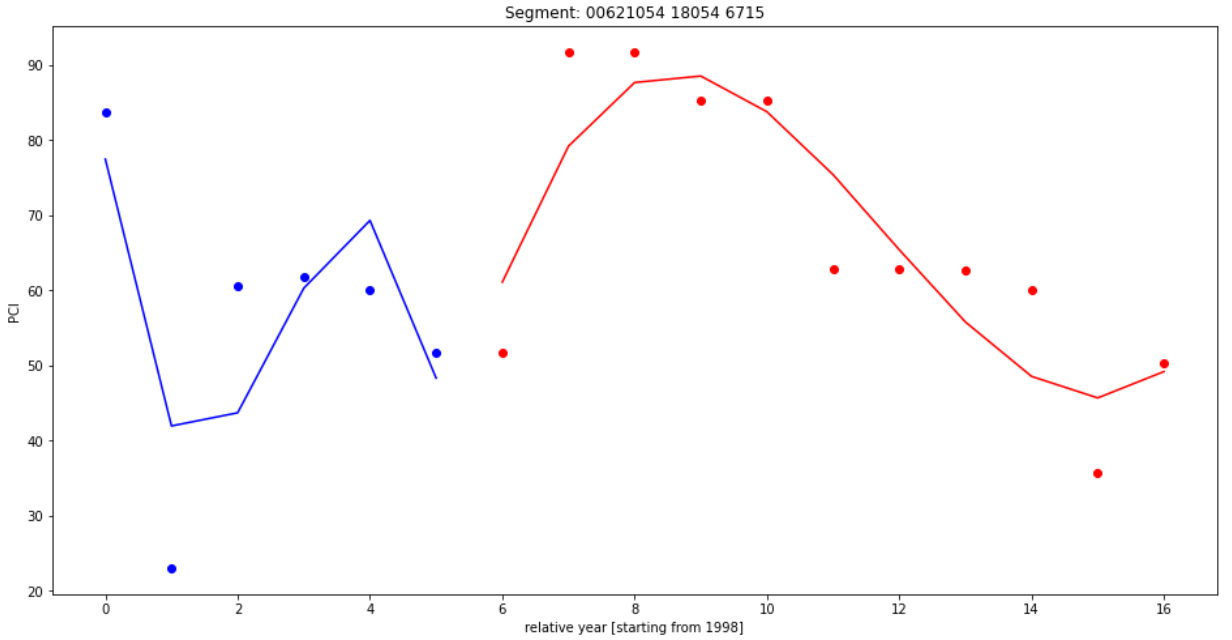
Methodology

Pavement Performance Modeling

Raw Data from PMIS



Conventional Regularization



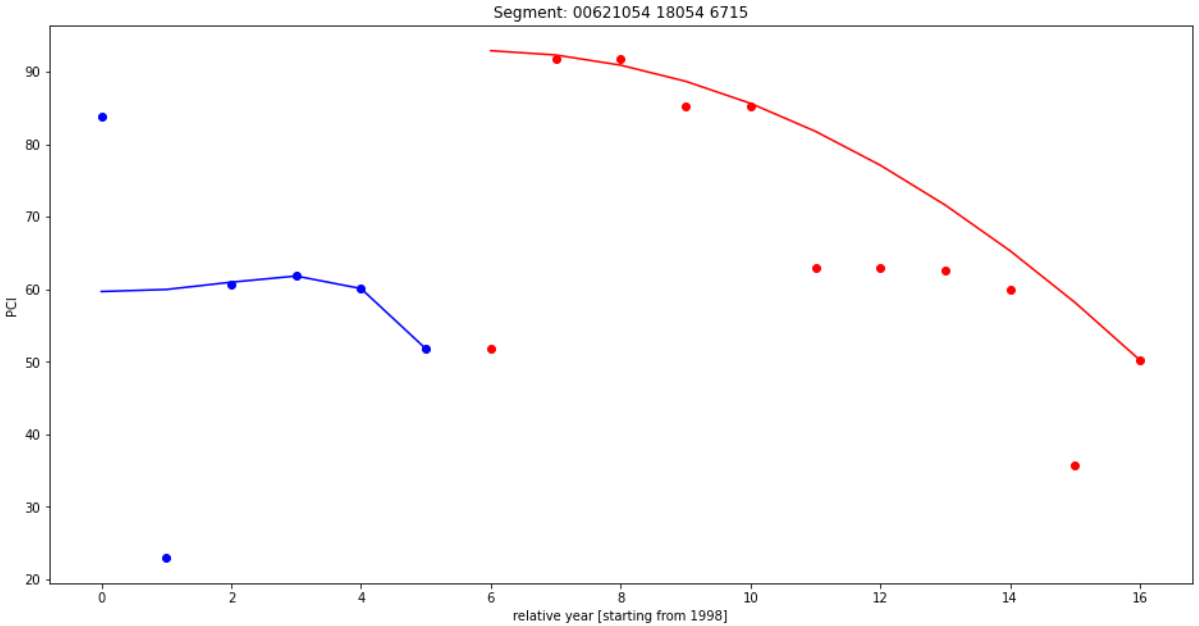
Regression+Regularization

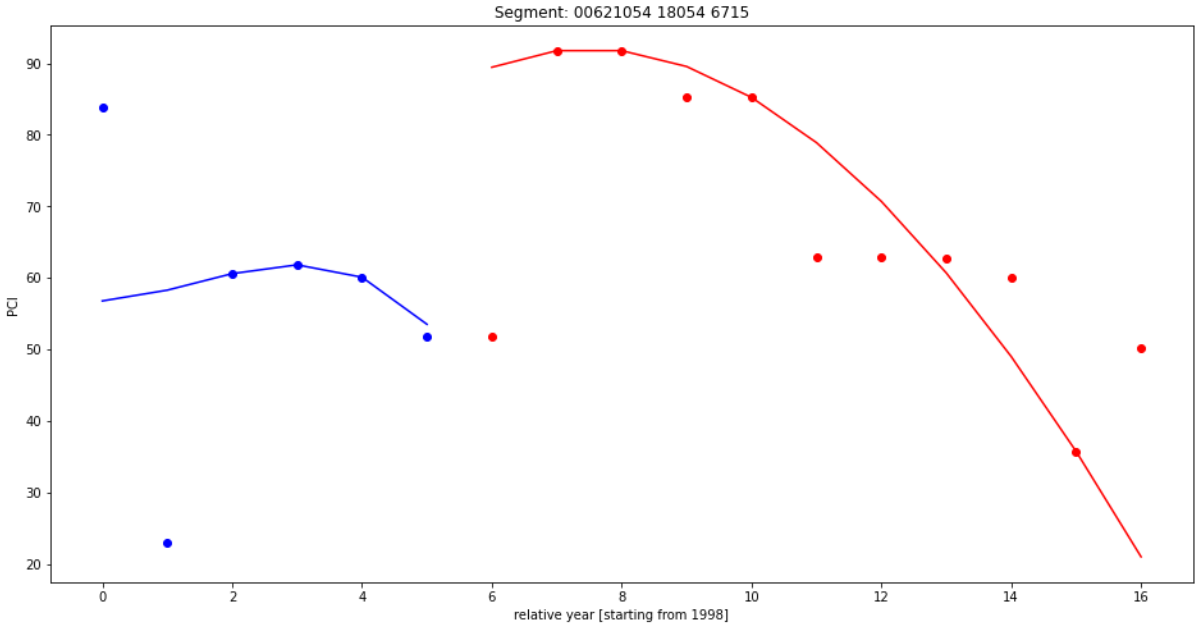
Methodology

Pavement Performance Modeling

Robust Regression

Raw Data from PMIS





Methodology - Phase I

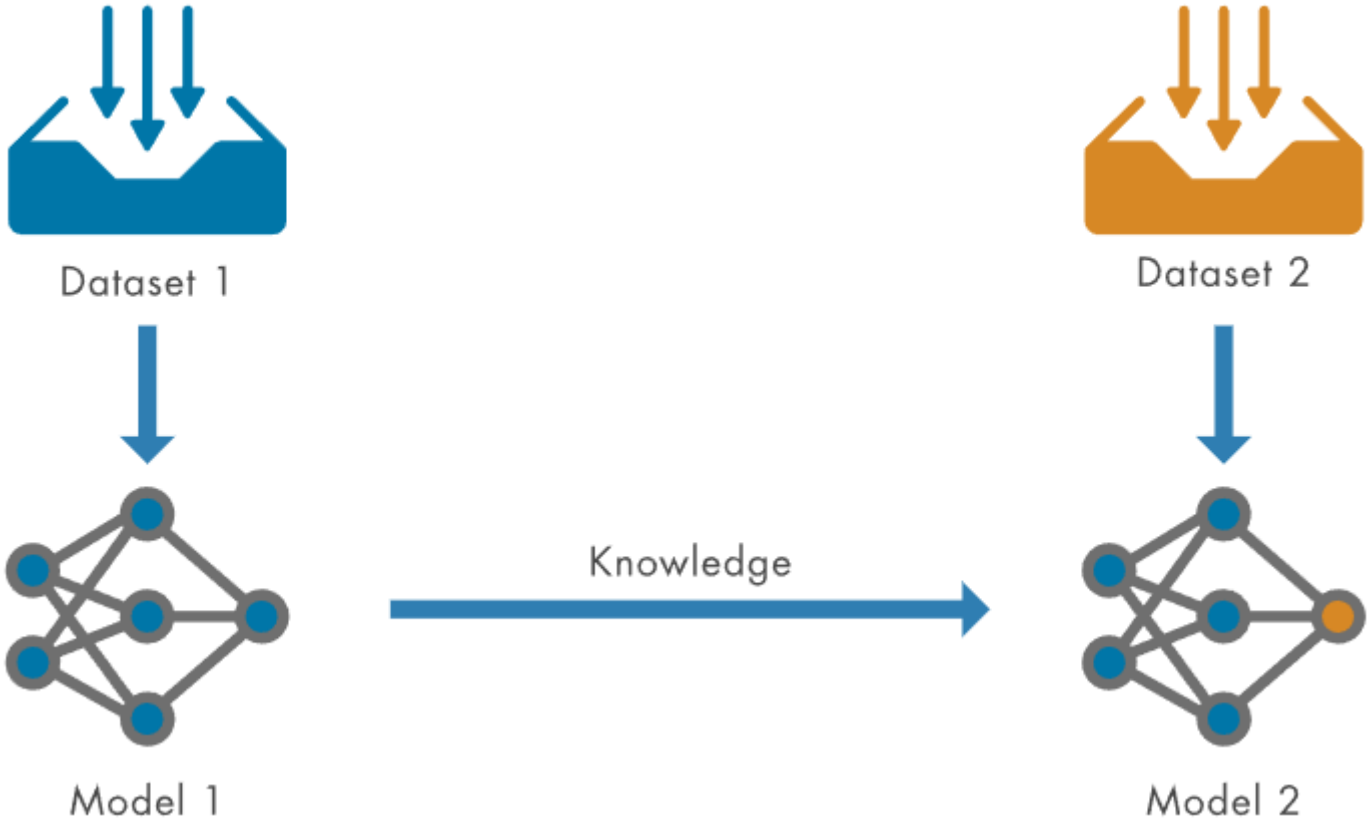
Challenges & Limitations

Data Issues:

- **Inconsistencies** from different vendors and unrecorded maintenance.
- **Outliers kept** → handled via robust regression, but still bias possible.
- **Insufficient reliable records in some categories** → "NA" in decision matrices.
- **Heavy skew toward HMA overlays** (potential analysis bias).

Methodology - Phase II

Transfer Learning



https://www.mathworks.com/discovery/transfer-learning.html

Methodology - Phase II

Data

LTPP data set
(1989-2023)
Pavement Data
Climate Data

Pavement Management Information System (PMIS)
Iowa DOT (1998-2023)

Bid Express software
(2015-2023)
Cost Data

Analysis

Transfer learning
From LTPP to PMIS

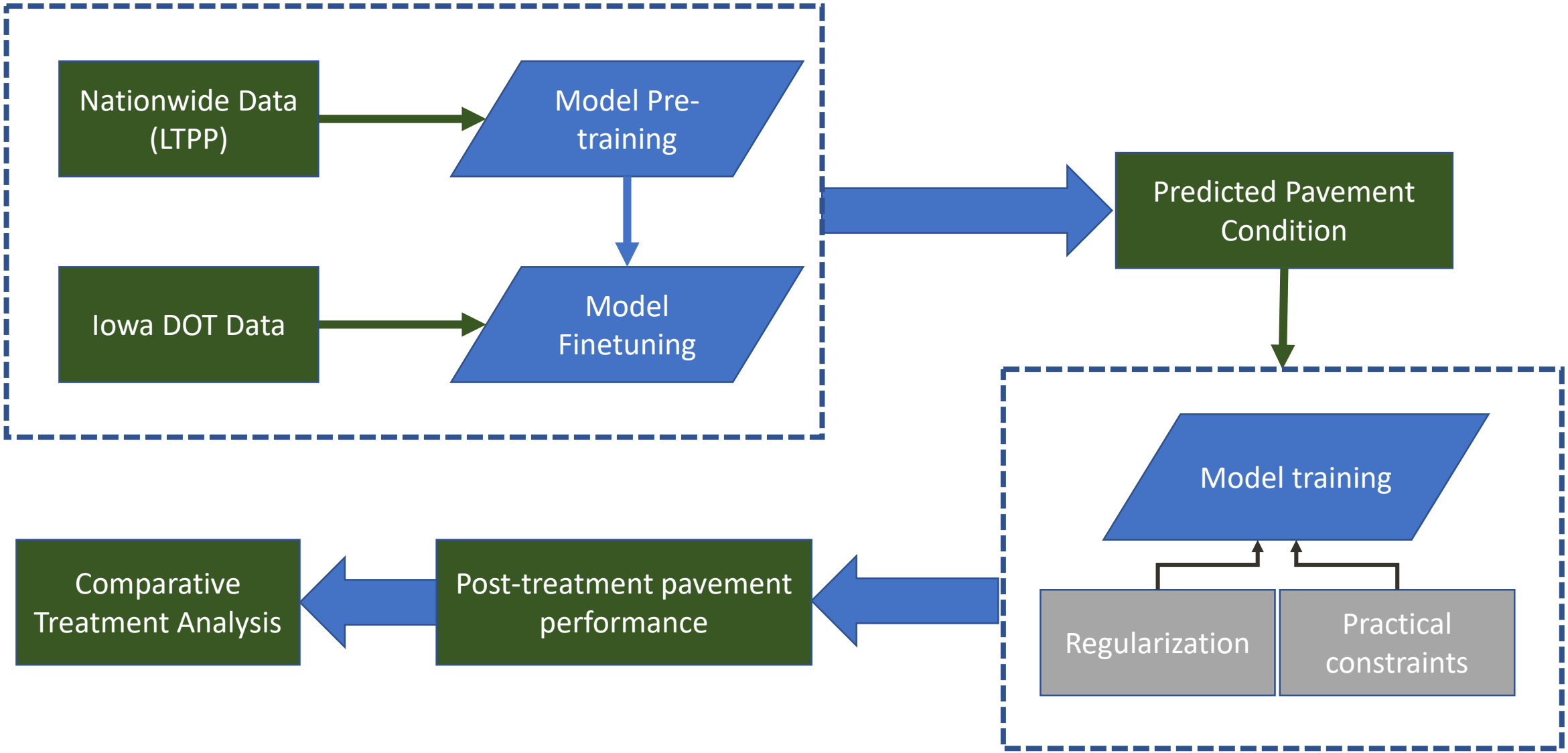
Pavement Performance Modeling

Benefit-Cost Analysis (BCA)

Results

Determining the Most Cost-Effective Pavement Preservation Treatment ✓

Methodology - Phase II



Methodology

Transfer Learning

Objective:

- Uses **LTPP dataset** as the knowledge source
- Adapts pretrained models to **PMIS dataset**
- Improves prediction of **Pavement Condition Index (PCI)**
- Handles noisy local data

Features as Input:

- Previous Cracking Index
- Previous Rutting Index
- Previous Riding Index
- Previous PCI
- Section's Age
- Traffic
- Truck traffic
- Precipitation
- Average temperature
- Freezing index
- Freeze and thaw

Methodology

Transfer Learning Models

Neural Network

- Pretrained on **LTPP sequences**, then fine-tuned on **PMIS**
- selective layer freezing preserves transferable knowledge while adapting to local data

RMSE : **6.98**
R-squared: **0.68**

Hybrid Model (LSTM + XGB)

- LSTM encoder captures long-term deterioration trends from historical sequences
- Embeddings are combined with exogenous/tabular features and passed to XG-Boost for final transfer learning

RMSE: **6.22**
R-squared: **0.75**

Limitation

- Learning from a three-year sequence causes significant data loss

Methodology

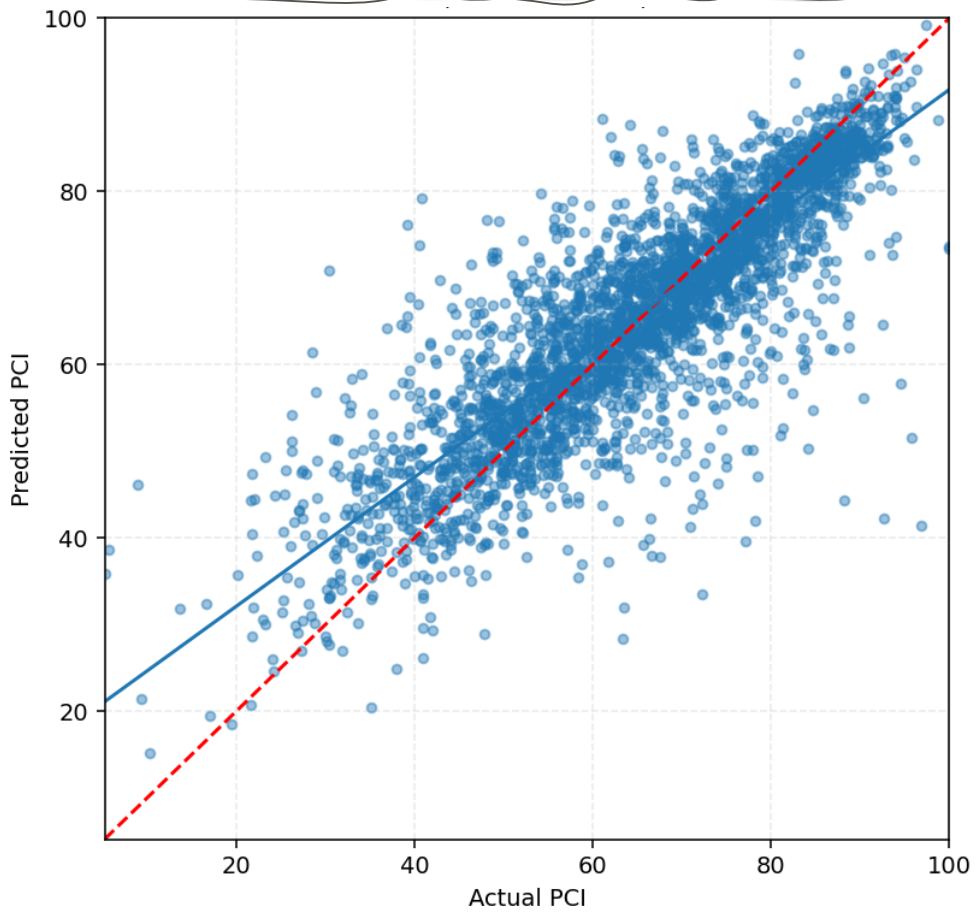
Chosen Model

XG-Boost Model

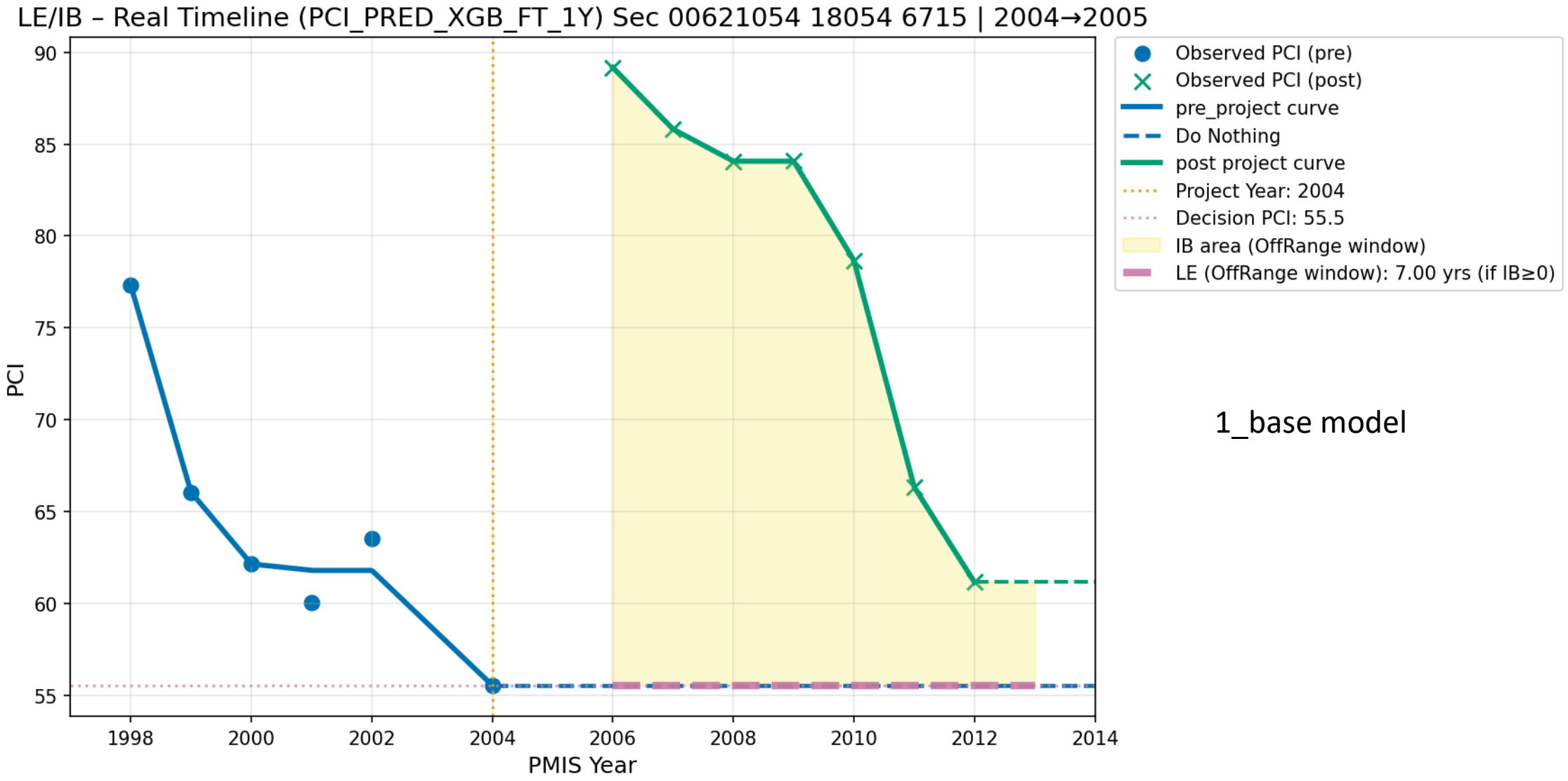
- Gradient boosting with monotone constraints to enforce engineering logic (e.g., older pavements shouldn't improve PCI)
- Fine-tuned on PMIS;
- Tested **joint fine-tuning**, blending weighted LTPP + PMIS samples to reduce domain shift

Outcome:
Cleaner, less noisy PCI in PMIS

RMSE : **6.45**
R-squared: **0.73**

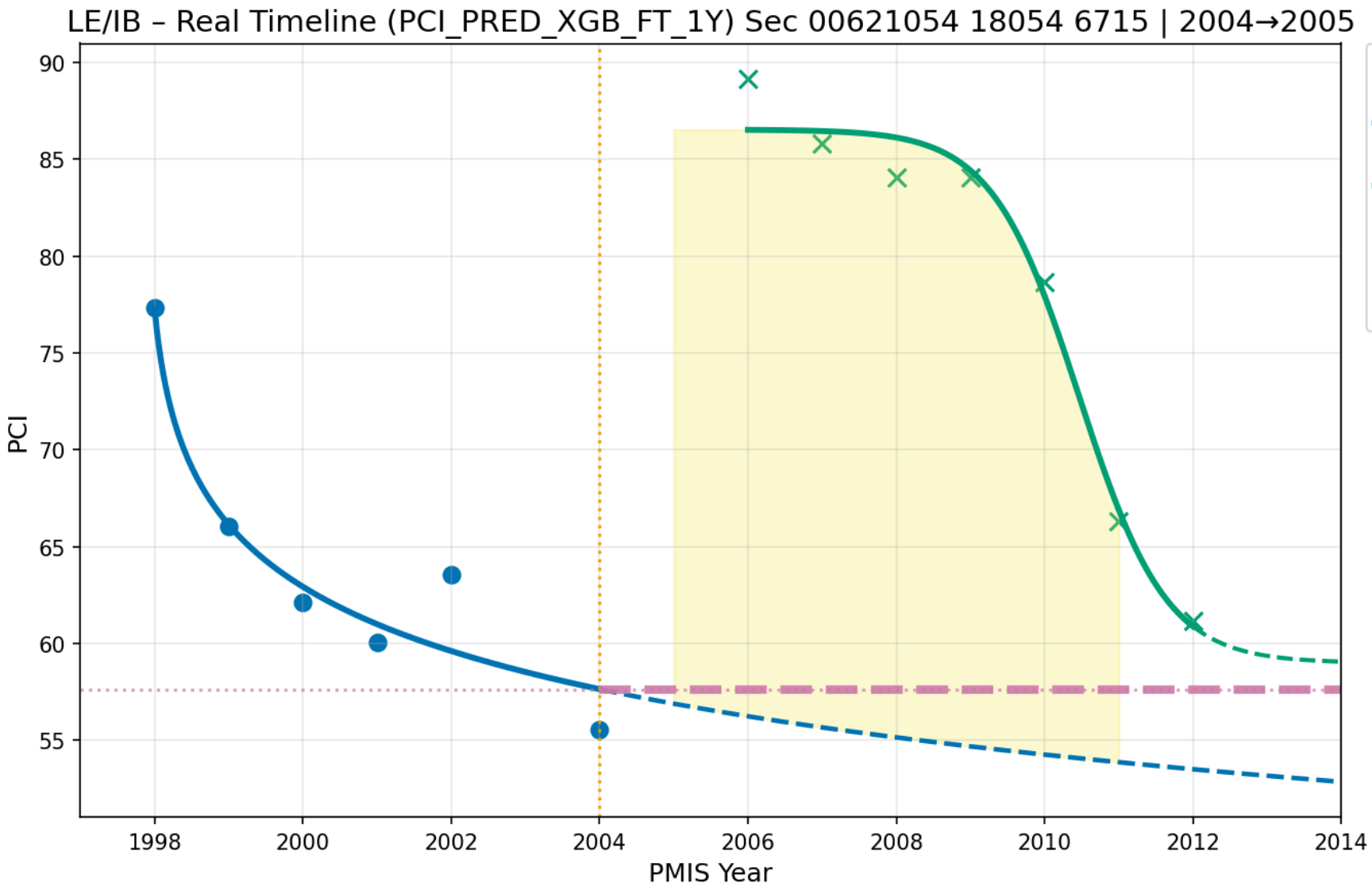


Methodology



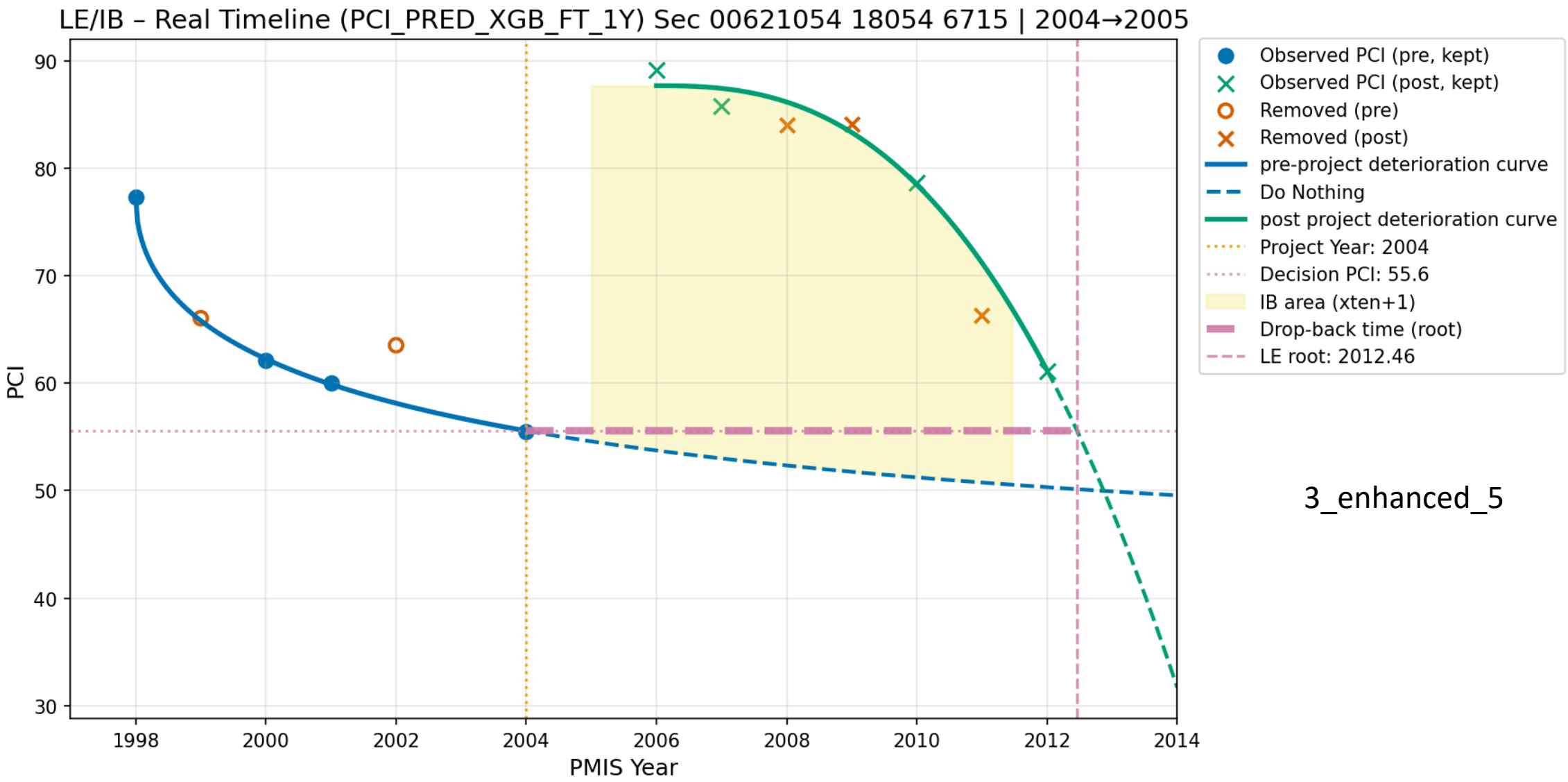
1_base model

Methodology



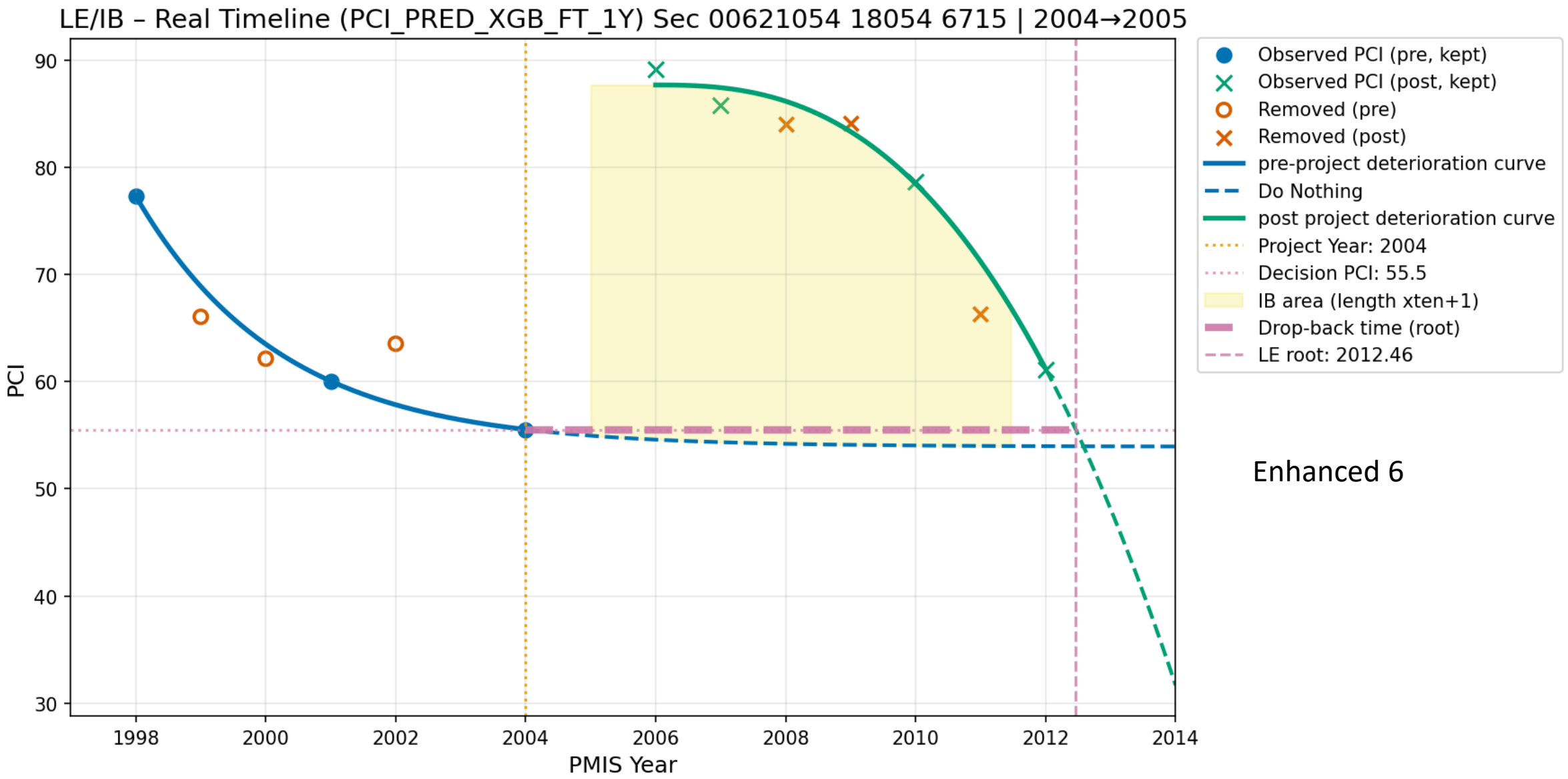
2_enhanced

Methodology



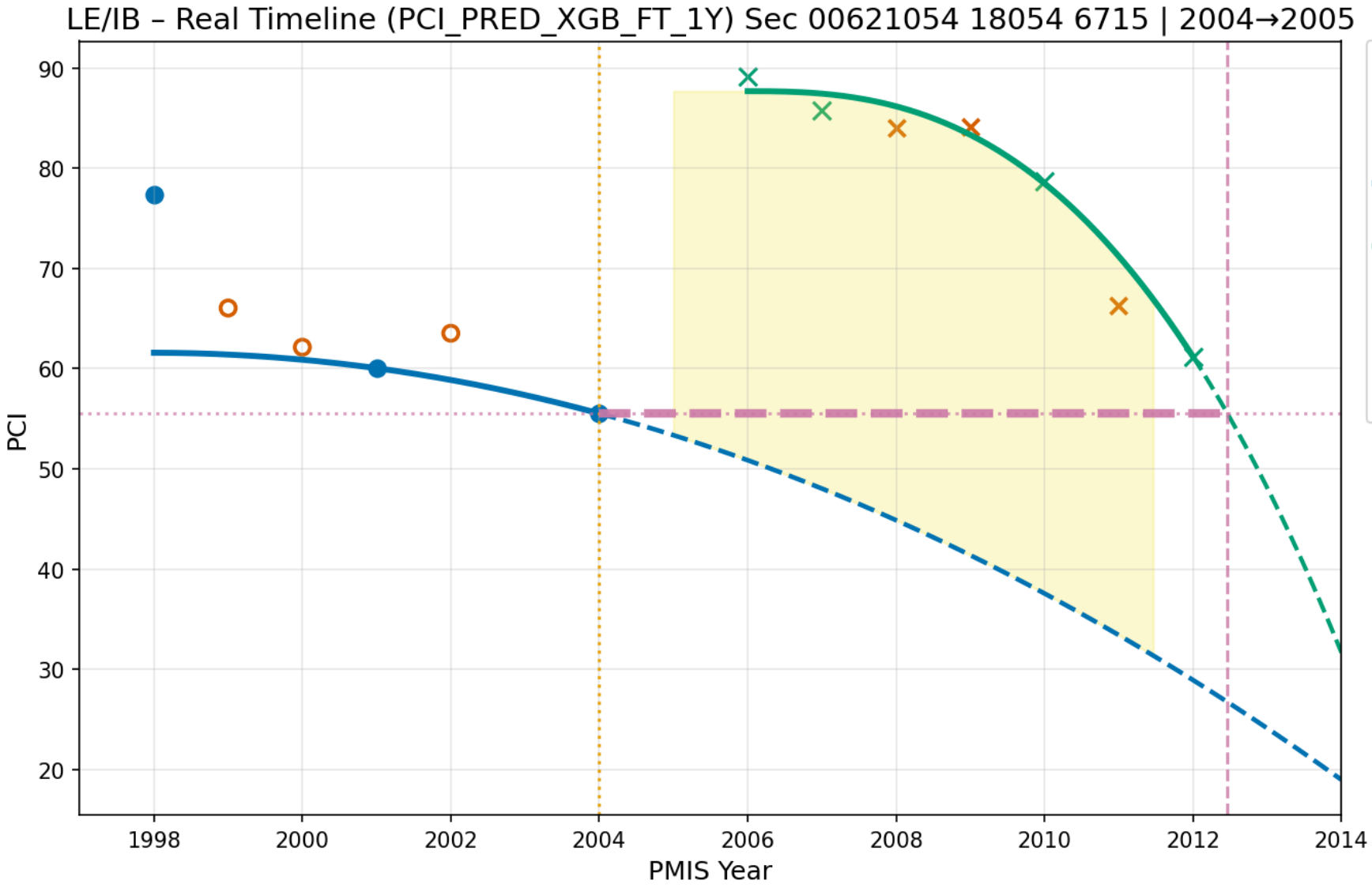
3_enhanced_5

Methodology



Enhanced 6

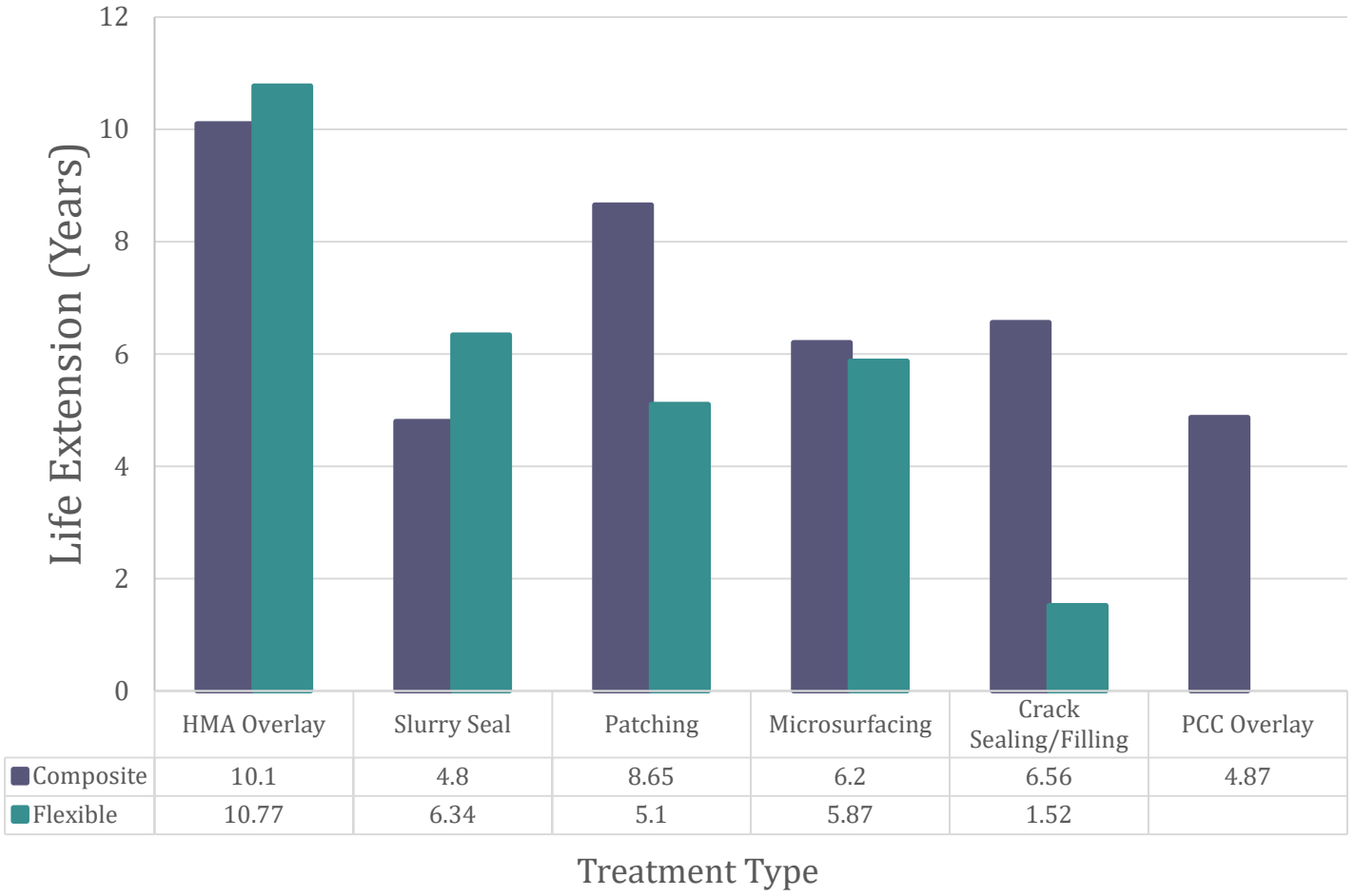
Methodology

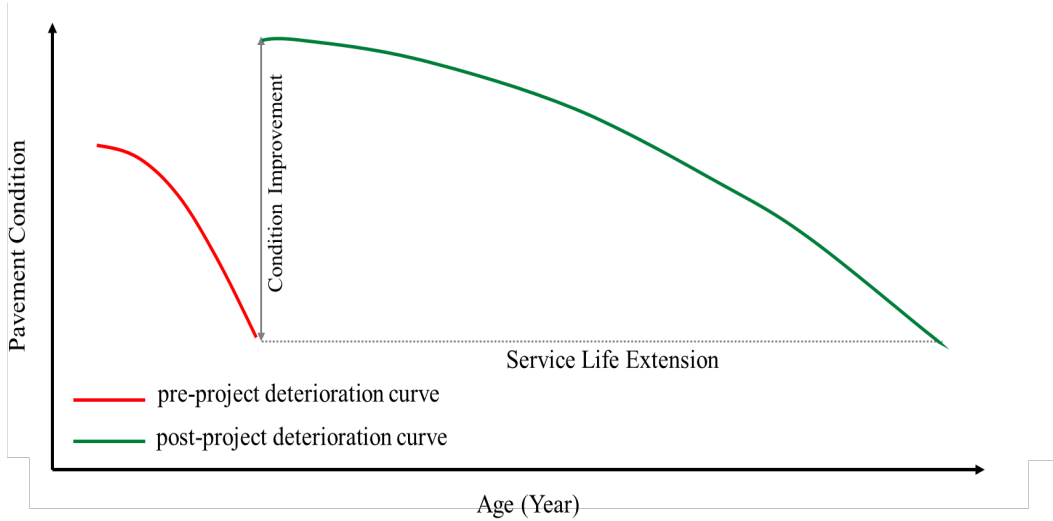


XGB_Bootstrap_Curvature Filter

Methodology

Performance Modeling





Methodology

https://www.pritchettcartoons.com/dot2.htm

☑ **Right Pavement**

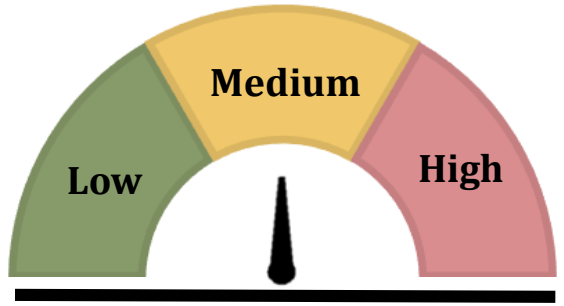
Iowa DOT PMIS Data

- **3,942 pavement section**
- **Pavement type** (Composite, Flexible)
- **Truck traffic**
- **Pavement condition**

Methodology

Truck Traffic

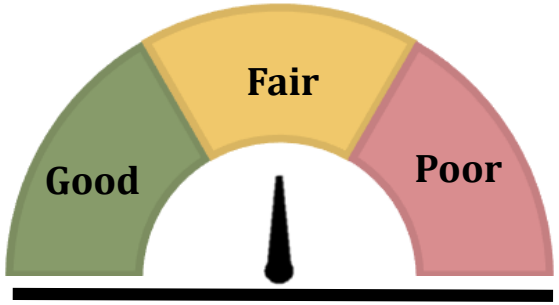
Low	Medium	High
AADTT < 500	500 =< AADTT < 1500	AADTT >= 1500



Traffic & Truck Traffic level

Pavement Condition

Good	Fair	Poor
PCI >75	50 < PCI <= 75	PCI =< 50



Pavement Condition

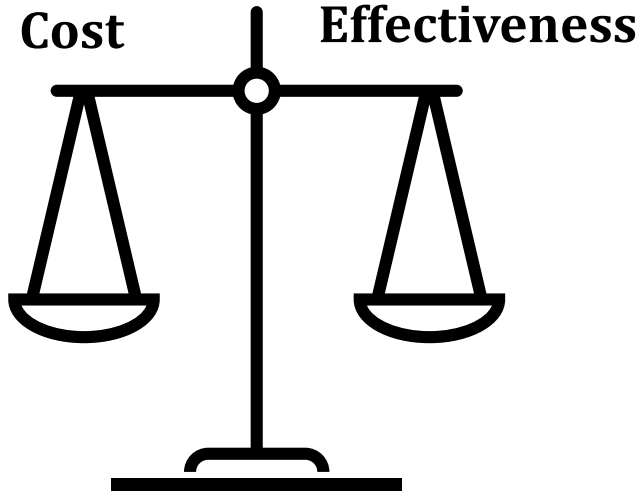
Benefit-Cost Analysis

$$NPV = \sum_{0}^{n_k} Preservation\ Cost \left[\frac{1}{(1+i)^{n_k}}\right] + \sum_{0}^{n_k} Maintanance\ Cost \left[\frac{1}{(1+i)^{n_k}}\right] - Salvage\ value \times \left[\frac{1}{(1+i)^{n_k}}\right]$$

Life-Cycle Cost of Asset

- ✓ Preservation Cost
- ✓ Maintenance Cost
- ✓ Salvage Cost
- ✓ Discount Rates(i)
- ✓ Year of expenditure(n_k)
- ✓ Analysis period

$$EUAC = NPV \times \frac{discount\ rate(1+discount\ rate)^{analysis\ period}}{(1+discount\ rate)^{analysis\ period} - 1}$$



Methodology

Right Treatment

$$Benefit/Cost\ Ratio = \frac{Index\ Benefit}{EUAC}$$

Greater ratio
=
Greater cost-effectiveness

Example: Composite Pavement

Treatments	Traffic level	Truck traffic	Pavement condition	EUAC	BCR
Crack Sealing/Filling	Medium	Low	Fair	1973	0.020892702
Slurry Seal	Medium	Low	Fair	3517	0.012825688
Microsurfacing	Medium	Low	Fair	8043	0.009556438
HMA Overlay	Medium	Low	Fair	17914	0.004813586
Patching	Medium	Low	Fair	55429	0.003861155

Cost Benefit Analysis Dashboard

Cost–Benefit Dashboard

LE_IB_Results_Strata_ML\summary_ newcat_no zero in LE

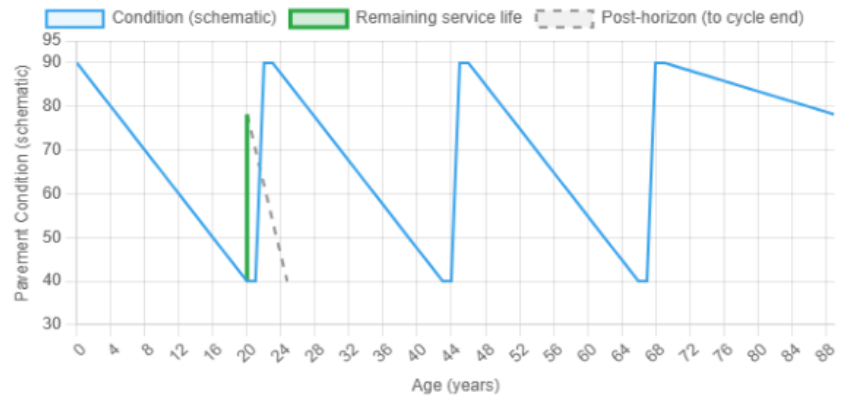
Combination	NewTruck	PavementFamily	Treatment	Cost basis
MR_NewTruck_Type_Compos	High	Composite	Crack Sealing/Filling	Lane-mile

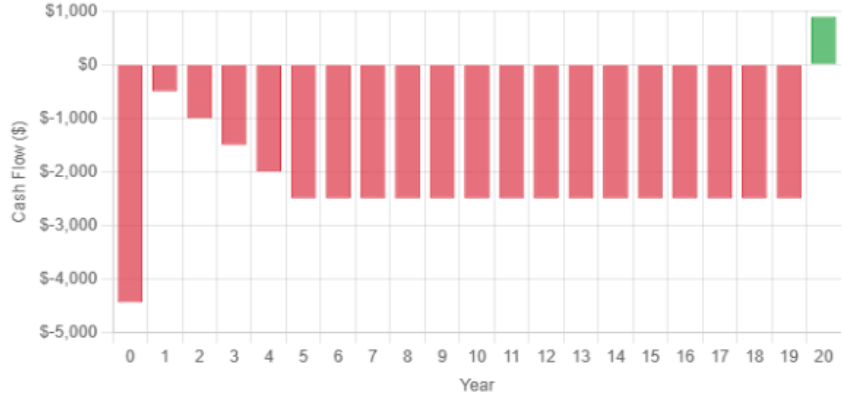
Life Extension (yrs)	Index Benefit (IB)	Age (info)	Analysis Period (yrs)	Discount Rate
6.18	845.00	13.00	20	2.2% (20y)

Maintenance Cap ($/yr): 2500

$2,980	0.284	$47,806	$3,393
EUAC	BCR (IB/EUAC)	NPV	Salvage

LE < N → salvage = 0; benefits stop at LE; maintenance continues to N.





Results

Pavement type	PCI-Condition	Dominant Distress Indicator		
		Cracking	IRI	Rutting
Flexible	Good	Slurry Sealing	Slurry Sealing	Slurry Sealing
	Fair	Microsurfacing	Slurry Sealing	Microsurfacing
	Poor	Slurry Sealing	Slurry Sealing	HMA Overlay

Pavement type	PCI-Condition	Dominant Distress Indicator		
		Cracking	IRI	Rutting
Composite	Good	Crack Sealing/Filling	Crack Sealing/Filling	Crack Sealing/Filling
	Fair	Crack Sealing/Filling	Microsurfacing	HMA Overlay
	Poor	Crack Sealing/Filling	Microsurfacing	Microsurfacing

Thank you!