Cross-Cell Battery Health Transfer Learning
(SOM → CNN Embedding + Domain Projection)

Dataset: NASA PCoE Li-ion Aging (B0005, B0006, B0007)

Author: Xinxin Sun

Motivation

- Battery SOH estimation usually requires full labels and cell-specific supervised training.
- Real-world systems lack ground-truth health labels for new cells.
- Goal: transfer knowledge across cells (B0005 → B0006, B0007) without target labels.
- A two-step SOM + CNN embedding + domain alignment approach.

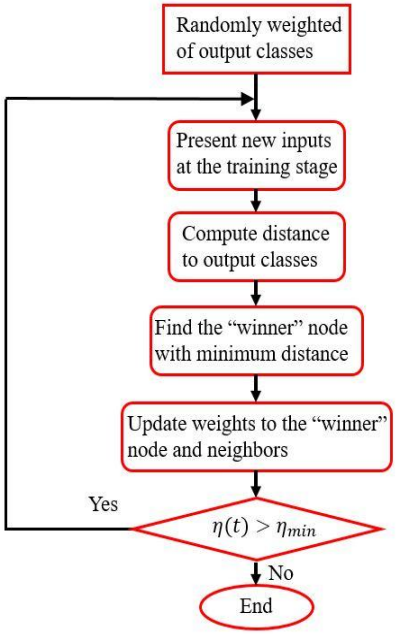
Dataset Overview

- Dataset: NASA PCoE Li-ion Aging
- Source: NASA Ames Prognostics Center of Excellence
- Cells: B0005 (reference labeled), B0006 & B0007 (unlabeled target)
- Input Signal: per-cycle voltage curve
- Task: Zero-label domain transfer health stage inference

Method Part-1: 9-D Structural Descriptor + Unsupervised SOM Alignment

Goal: Learn interpretable degradation morphology without labels.

Feature Group	Feature Name	Symbol	Type	Sensitivity	Why Important for Battery Aging
Intensity	Normalized Intensity	f1	Global	Overall trend level	Captures baseline SoH shift
Gradients	Gradient Magnitude	f2	Local diff	Local slope changes	Detects aging onset
	Laplacian	f3	Curvature	Inflection points	Identifies instability patterns
Local Variance	Local Std	f4	Texture	Fluctuation roughness	Reflects micro-cycle noise
Frequency	FFT (Low-Band)	f5	Spectral	Slow drift	Tracks long-term wear
	FFT Residual	f6	Spectral	Fast anomalies	Captures transient failure
Robustness	MAD	f7	Statistic	Noise-immune variance	Stable under sensor distortion
Phase-Shape	Phase Curve Index	f8	Geometry	Curve morphology shift	Models hysteresis evolution
Dynamics	Entropy	f9	Complexity	Irregularity level	Identifies late-stage chaos



Workflow Summary

1.Extract **9-D pixel-level structural descriptors**.
2.Apply **PCA → 6-D** to preserve dominant variance.
3.Train **SOM (unsupervised)** to form stable interpretable clusters.
4.Use **Elbow** to auto-select cluster K.
5.Assign new cells via **prototype matching + embedding visualization**.

Key Advantages

- **No labels required** — fully self-supervised.
- **Physically meaningful** — each dimension retains signal morphology.
- **Transfer-ready** — centroids reusable across new cells.
- **Explainable** — SOM clusters correspond to interpretable degradation modes.
- **Lightweight deployment** — CPU-only, no fine-tuning required.
- **Better than raw-sequence ML** — no overfitting to battery-specific biases.

Unlike RUL-based regression or cycle-index ML, our method **models intra-cycle structural evolution**, enabling **pixel-level cross-cell alignment** and **transferable degradation prototypes**.

Directly supports SOH diagnosis via pixel-level degradation morphology.

Method Part-2: SOM-Supervised CNN Latent Encoder & Cross-Cell Transfer

Goal: Distill SOM knowledge into a transferable latent encoder.

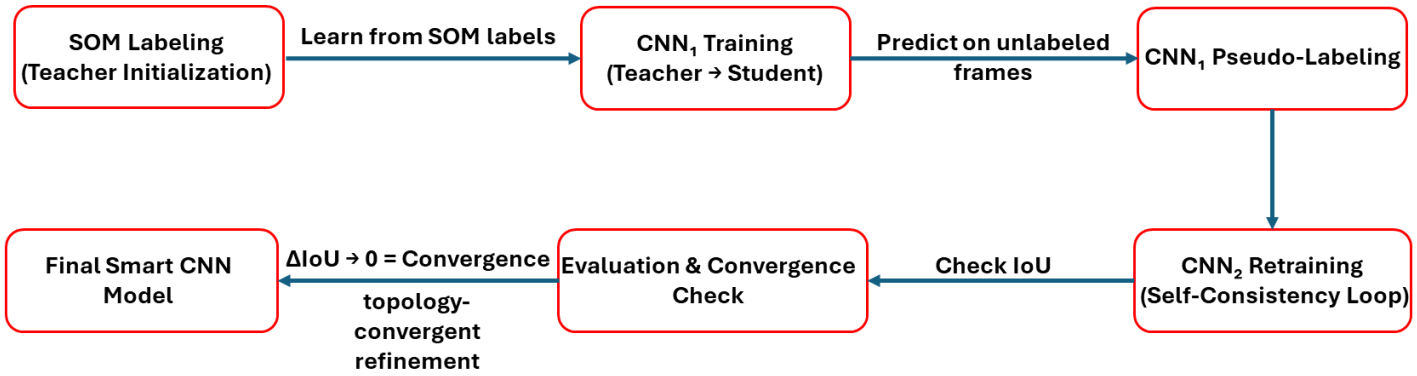
Objective: Build a **reusable latent representation** that generalizes across battery cells.
Training data: Only **B0005**, supervision = **SOM-derived pseudo labels**
Network
1D CNN encoder: Conv → ReLU → Conv → ReLU → AvgPool → FC → **64-D embedding**
Training rule

$$Loss = CrossEntropy(f(x), SOM\ pseudo\ label)$$

This converts SOM clusters into **learnable geometric prototypes** in latent space.
Transfer to B0006/B0007 (no fine-tuning)

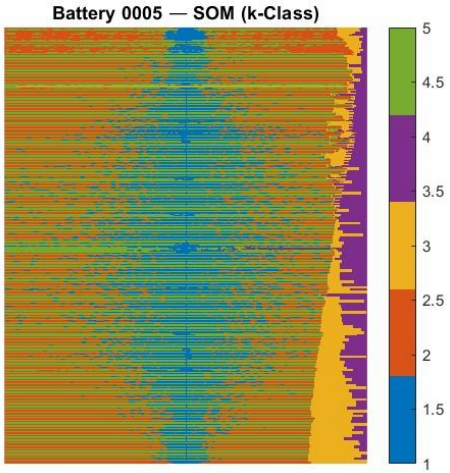
$$\hat{z} = f(x_{test}), \check{y} = arg^{\min}_{c\in\{1,2,3\}} \|\check{z} - \mu_c\|_2$$

Property	Meaning
Hybrid supervised	SOM distills interpretable labels; CNN compresses & generalizes
Domain-transferable	Same encoder reused across battery cells
Explainable topology	Predictions tied to SOM prototypes, not black-box regression
Lightweight & low-cost	Only voltage signal; no degradation metadata required
Knowledge-distillation topology	Student CNN preserves SOM manifold geometry

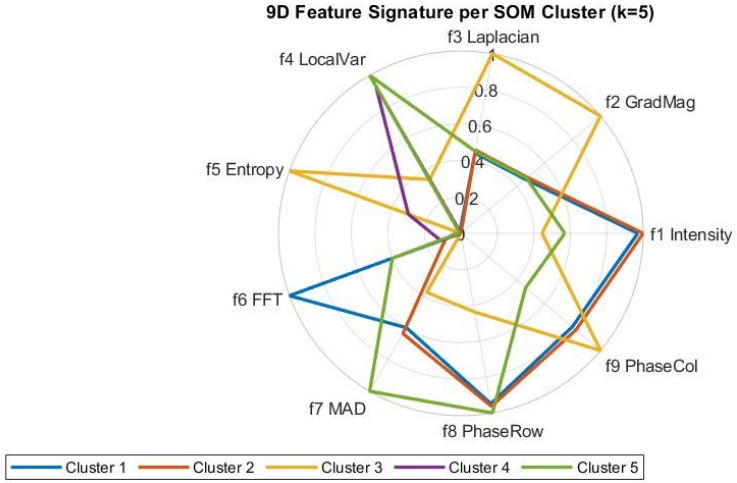


SOM provides **topology-preserving supervision**, while CNN **compresses the manifold into a reusable latent space**, enabling **zero-retraining cross-cell transfer**. *Unlike RUL-based deep models, our encoder operates **label-free, cell-free, and cycle-index free**, requiring **only voltage signal** and **no retraining** when transferring across cells.*

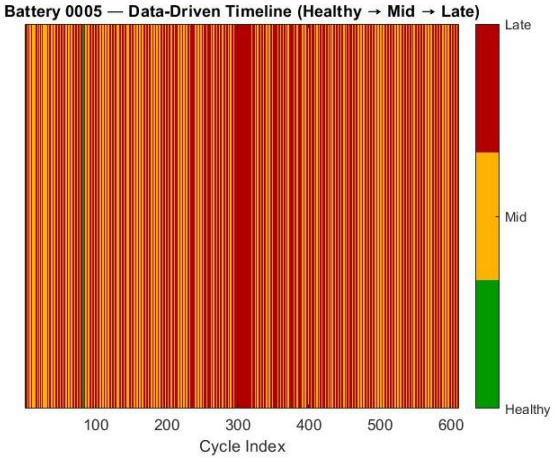
SOM-Based Degradation Staging – NASA Cell B0005



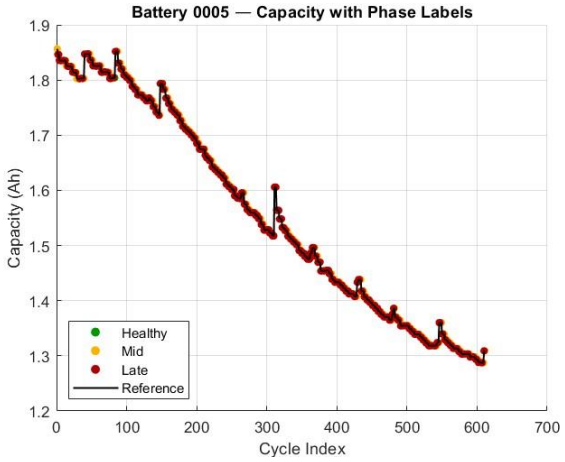
B0005 – SOM k=5 Pixel-Level Clusters



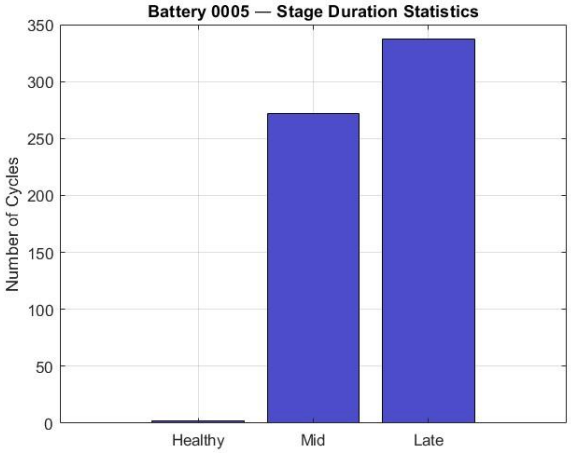
9-D Structural Signature of SOM Clusters



Unsupervised 3-Stage Mapping (Healthy → Mid → Late)



Capacity vs. Cycle, Colored by SOM Stages

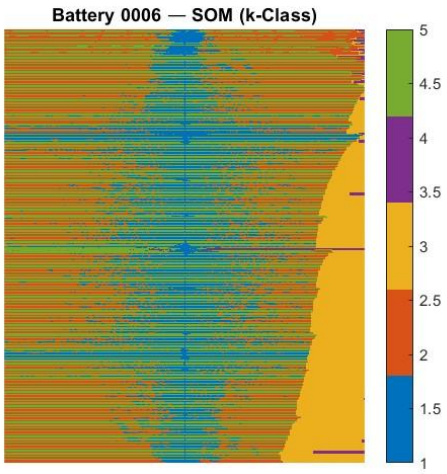


Stage Duration (Number of Cycles)

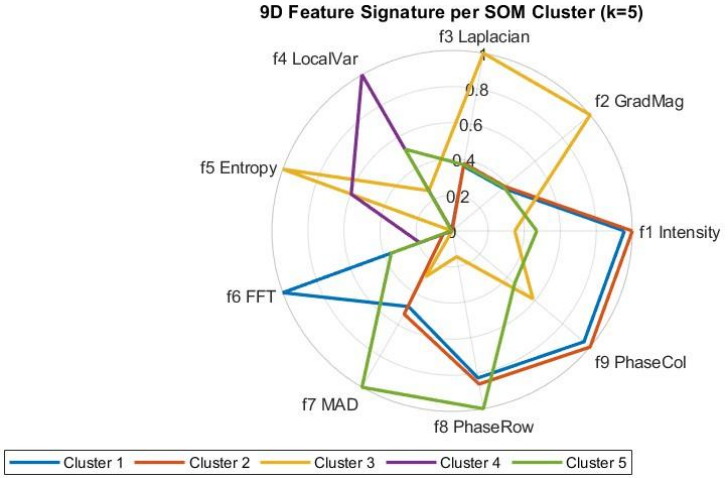
Note: Pure SOM-based staging on B0005 (no CNN yet).

Cross-Cell Transfer: B0006 via Reused SOM-CNN (No Retraining)

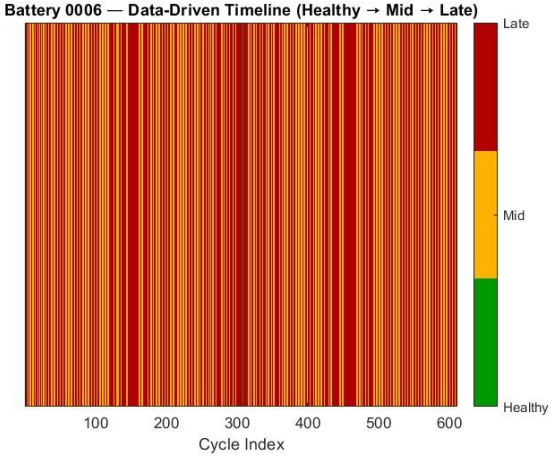
Using B0005-trained SOM prototypes + CNN encoder directly on B0006. No labels. No fine-tuning. Fully transferable degradation interpretation.



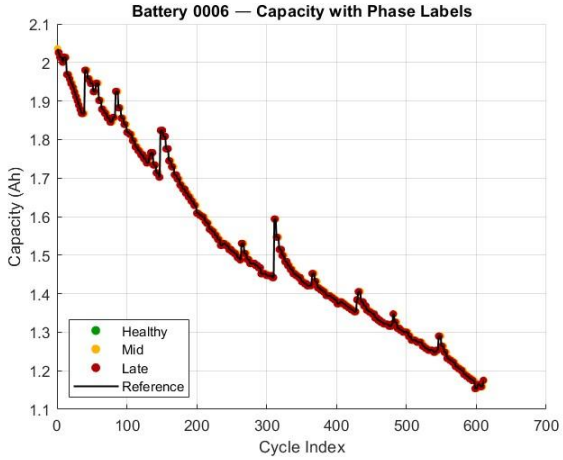
B0006 – SOM k=5 Pixel-Level Clusters



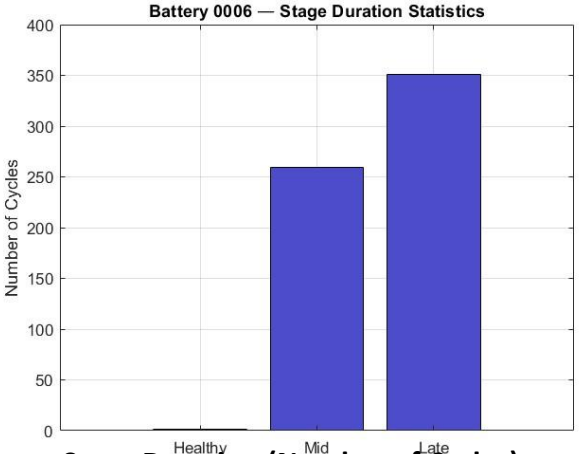
9-D Structural Signature of SOM Clusters



Unsupervised 3-Stage Mapping (Healthy → Mid → Late)



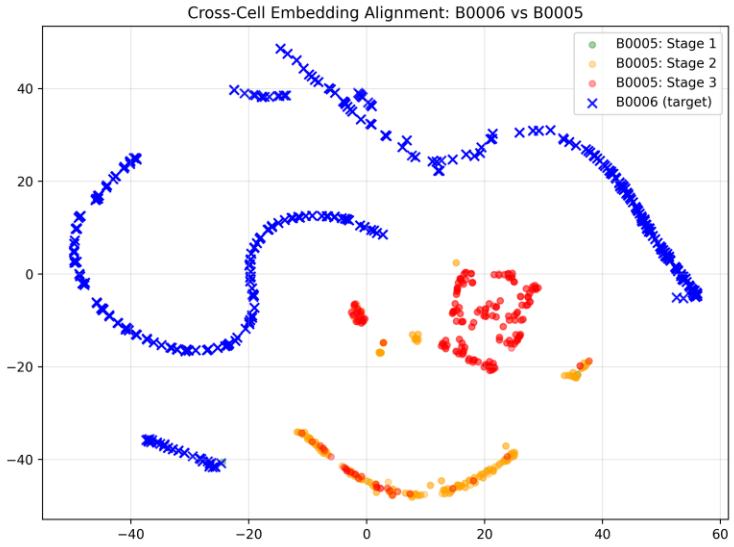
Capacity vs. Cycle, Colored by SOM Stages

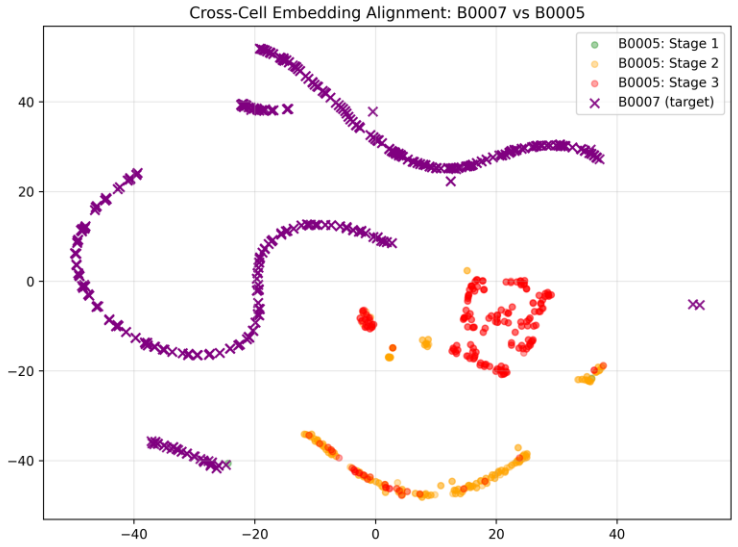


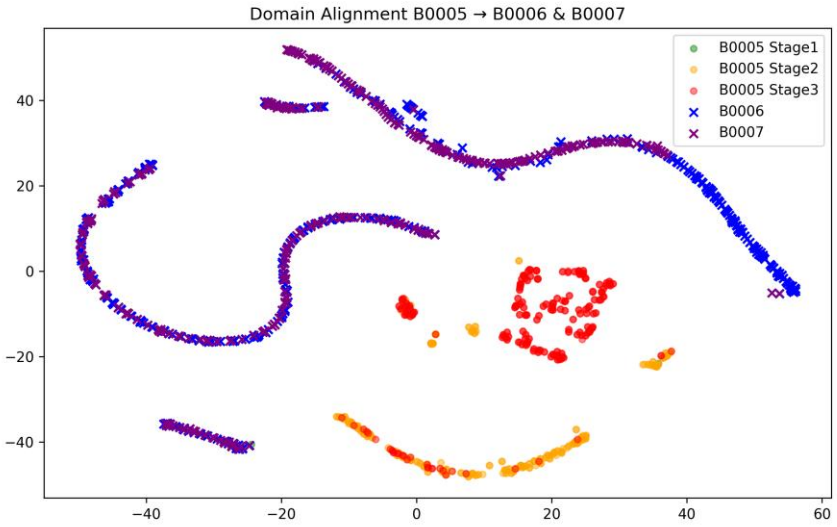
Stage Duration (Number of Cycles)

SOM weight-vector topology reused across cells → stable 3-stage prediction under domain shift.
Cross-cell reuse of SOM centroids enables zero-shot transfer, crucial for large battery fleets.

Cross-cell Embedding Projection (t-SNE)





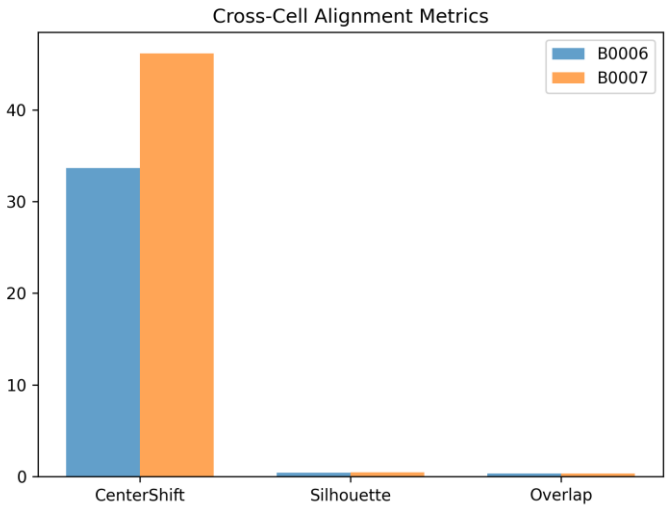


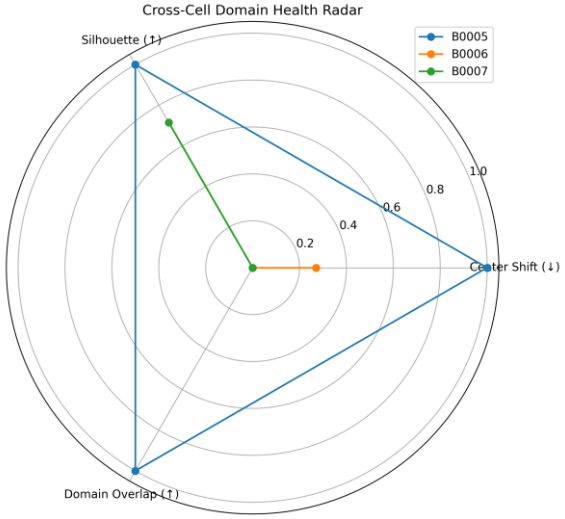
These projections use the CNN latent encoder distilled from B0005's SOM labels, showing that B0006/B0007 cycles occupy consistent, transferable geometry without any retraining.

These embeddings preserve degradation manifold structure, enabling early-cycle prognosis.

No labels, no retraining — only voltage signal is used.

Domain Gap Metrics & Radar Plot

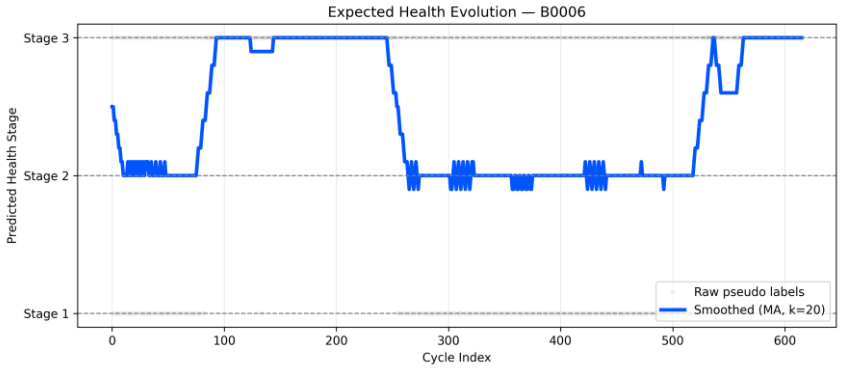


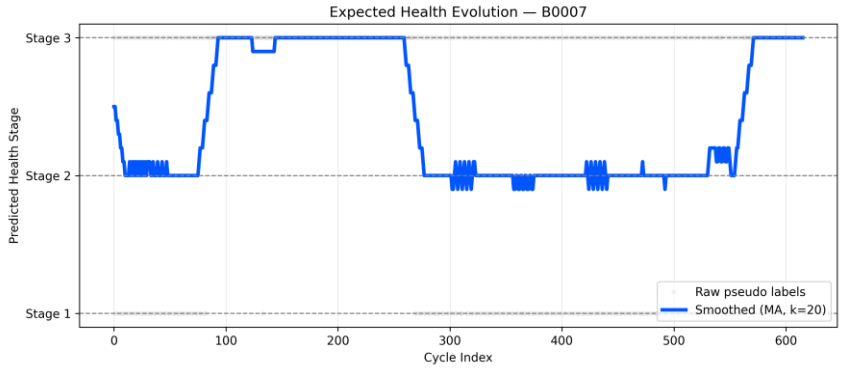


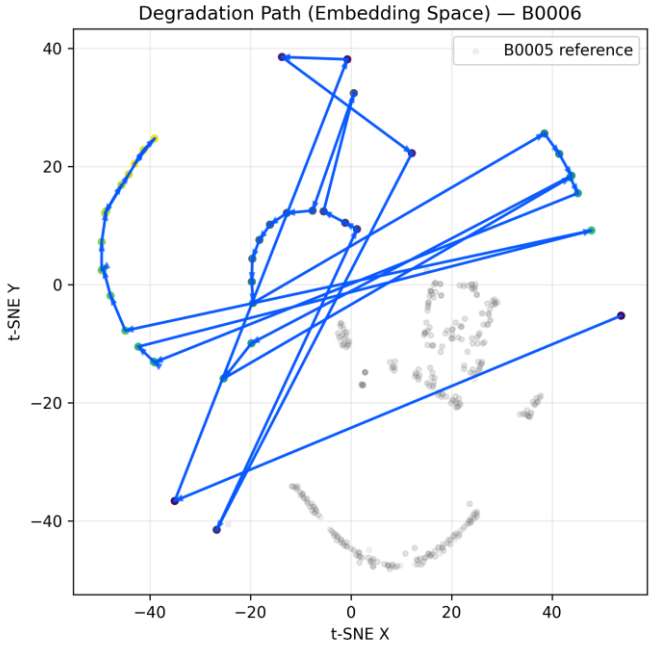
Cross-cell alignment metrics show that the B0005-trained encoder preserves degradation topology when transferred to B0006/B0007 (no retraining).

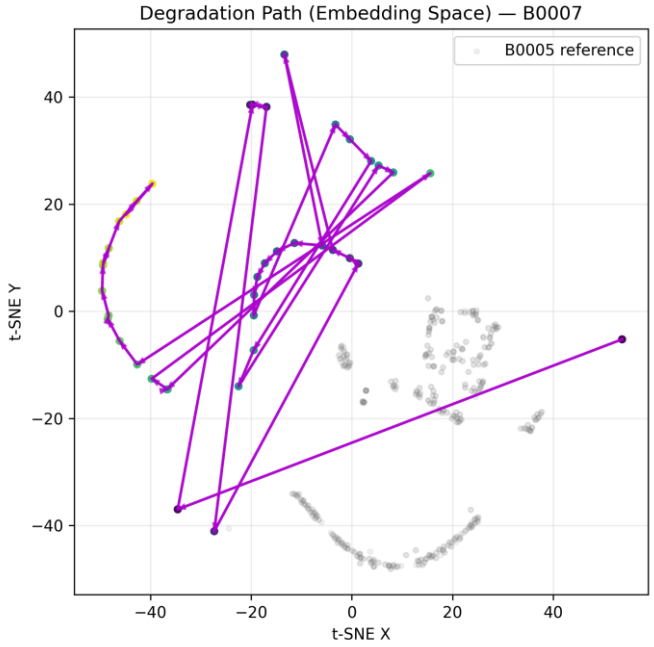
Provides interpretable quantification of cross-cell domain gap, enabling scalable deployment.

Health Stage Evolution (Smoothed)









Cross-cell trajectories show that B0006 / B0007 preserve the same 3-stage progression and manifold path learned from B0005 (no retraining). The encoder reproduces consistent degradation paths across cells, confirming stability of the learned structural manifold.

Conclusion

- Successfully generated target health stage estimation without labels.
- Embedding space alignment shows partial consistency.
- Domain metrics suggest feasible knowledge transfer.
- Future: physics-informed loss + temporal consistency training.

Why this pipeline is a strong match for Argonne's Battery Diagnosis & Prognosis Program

- Fully unsupervised → solves the "no labels / complex datasets" problem in the JD.
- Physics-aware 9D descriptors → compatible with electrochemical knowledge.
- SOM topology encodes interpretable degradation stages → supports diagnosis.
- CNN latent encoder distills SOM geometry → enables transferable prognosis embeddings.
- Zero-shot transfer to B0006/B0007 → ideal for large-scale battery fleets.
- Domain gap metrics quantify cross-cell generalization → important for deployment.
- Lightweight, fast, and HPC-friendly → fits Argonne's big-team HPC workflow.