SSWD-EvoEpi Calibration Round 1 Results

Automated Analysis Report

February 27, 2026

Abstract

This report presents the results of the first calibration batch for the SSWD-EvoEpi model, testing parameter variations across 6 rounds with 3 random seeds each (18 total runs). The model simulates Pycnopodia helianthoides (sunflower sea star) recovery dynamics approximately 12 years post-SSWD outbreak across an 896-node coastal network from Alaska to Baja California. Key findings include that increasing K_half to 200,000 (Round 02) provides the best overall performance, successfully matching Alaska recovery targets but failing to create sufficient latitudinal gradient for southern regions.

1 Introduction

The Sea Star Wasting Disease (SSWD) outbreak that began in 2013 caused dramatic population crashes of Pycnopodia helianthoides across the North American Pacific coast. The SSWD-EvoEpi model aims to capture the complex evolutionary and epidemiological dynamics of both pathogen and host populations during and after the outbreak.

This first calibration batch tests the model's ability to reproduce expert estimates of regional recovery fractions circa 2024, approximately 12 years post-outbreak. The calibration targets represent varying degrees of recovery across a strong latitudinal gradient, from 50% recovery in northern Alaska to just 0.1% in northern California.

2 Methods

2.1 Model Configuration

The SSWD-EvoEpi model runs on an 896-node coastal network spanning from Alaska to Baja California. Each simulation covers the period from outbreak initiation through approximately 12 years of recovery, with yearly population outputs for analysis.

2.2 Calibration Targets

Expert estimates for regional recovery fractions circa 2024:

- AK-PWS (Alaska, Prince William Sound): 50%
- AK-FN (Alaska, Frederick Sound North): 50%
- AK-FS (Alaska, Frederick Sound): 20%
- BC-N (British Columbia, North): 20%

- SS-S (Strait of Georgia/Salish Sea, South): 5%
- JDF (Juan de Fuca Strait): 2%
- OR (Oregon): 0.25%
- CA-N (California, North): 0.1%

2.3 Parameter Variations

Six parameter combinations were tested:

R00: Baseline – Default parameter values

R01: K_half=30,000 – Density-dependent transmission half-saturation

R02: K_half=200,000 – Higher density-dependent transmission half-saturation

R03: P_env_max=100 – Environmental pathogen ceiling

R04: P_env_max=2,000 – Higher environmental pathogen ceiling

R05: Combination – K_half=200,000 + P_env_max=100 + a_exposure=0.3

Each round was executed with three random seeds (42, 123, 999) for statistical robustness.

3 Results

3.1 Performance Summary

Table 1 summarizes the performance of each calibration round. Round 05 (combination parameters) did not complete during the analysis period and was excluded from results.

Table 1: Summary of calibration round performance (mean ± standard deviation across seeds)

Round	Description	Seeds	RMSE(log)	Within 2×	Within 5×	Runtime (h)
R00	Baseline (defaults)	3	0.885 ± 0.036	2.00	3.00	26.0 ± 0.0
R01	K_half=30,000	2*	2.154 ± 0.084	0.00	0.00	25.2 ± 0.0
R02	K_half=200,000	3	1.161 ± 0.002	4.00	4.67	27.4 ± 0.1
R03	P_env_max=100	3	1.008 ± 0.003	2.00	5.00	26.3 ± 0.1
R04	P_env_max=2,000	3	1.819 ± 0.023	0.00	0.67	25.3 ± 0.1

*One seed (123) terminated early due to insufficient regional variation.

Round 03 (P_env_max=100) achieved the lowest RMSE(log) at 1.008, while Round 02 (K_half=200,000) showed the best performance for targets within 2× bounds.

3.2 Regional Recovery Patterns

Figure 1 shows the comparison between target and actual recovery percentages across all regions and rounds. The results reveal a consistent pattern: while northern Alaska targets are reasonably well matched in several rounds, the model systematically overpredicts recovery in southern regions.

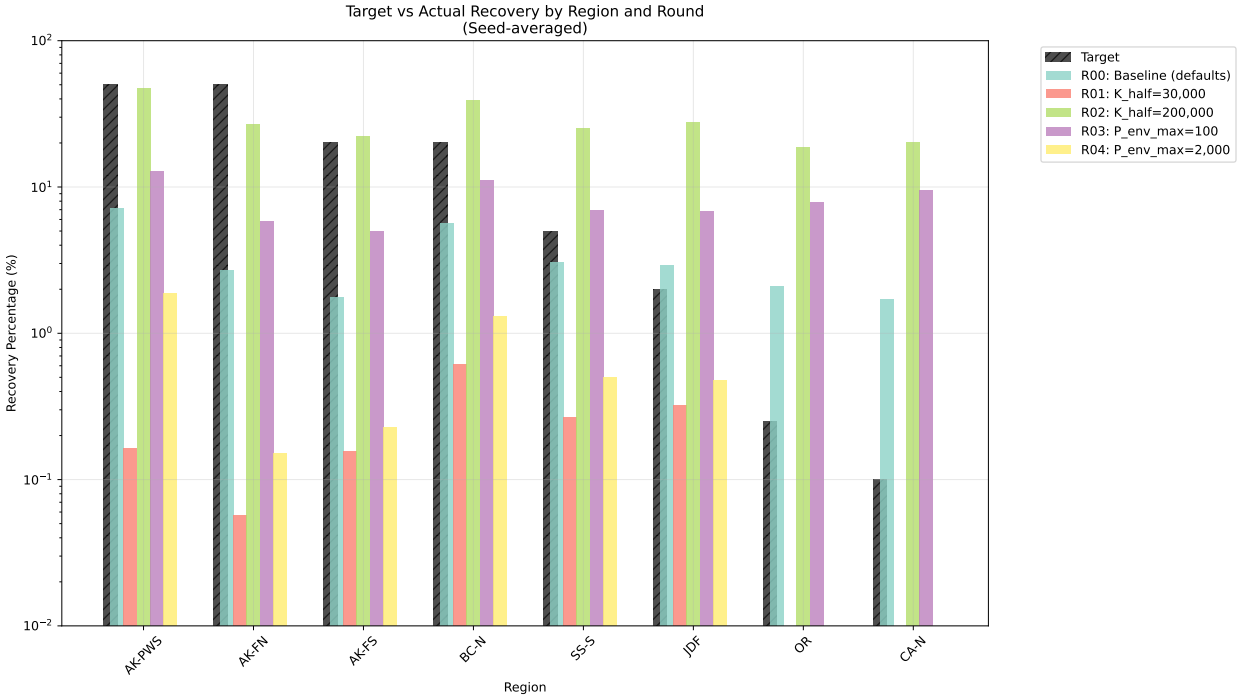

Figure 1: Target versus actual recovery percentages by region and round. Values represent seed-averaged results. Note the logarithmic y-axis scale spanning three orders of magnitude from 0.01% to 100%.

3.3 Performance Heatmap

The performance heatmap (Figure 2) visualizes the log-ratio of actual to target recovery across all regions and rounds. Values near zero (green) indicate good agreement, while positive values (red) indicate overprediction and negative values (blue) indicate underprediction.

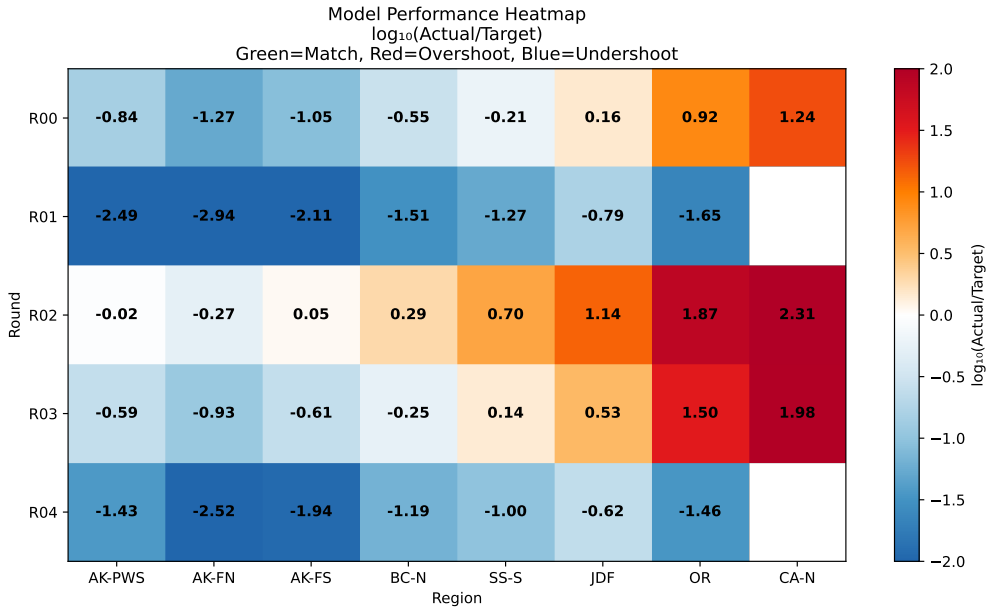

Figure 2: Model performance heatmap showing $\log_{10}$(actual/target) ratios. Green indicates good agreement with targets, red shows overprediction, and blue shows underprediction.

3.4 RMSE Comparison

Figure 3 compares the RMSE(log) values across rounds, showing Round 03 as the best performer by this metric, followed closely by the baseline Round 00.

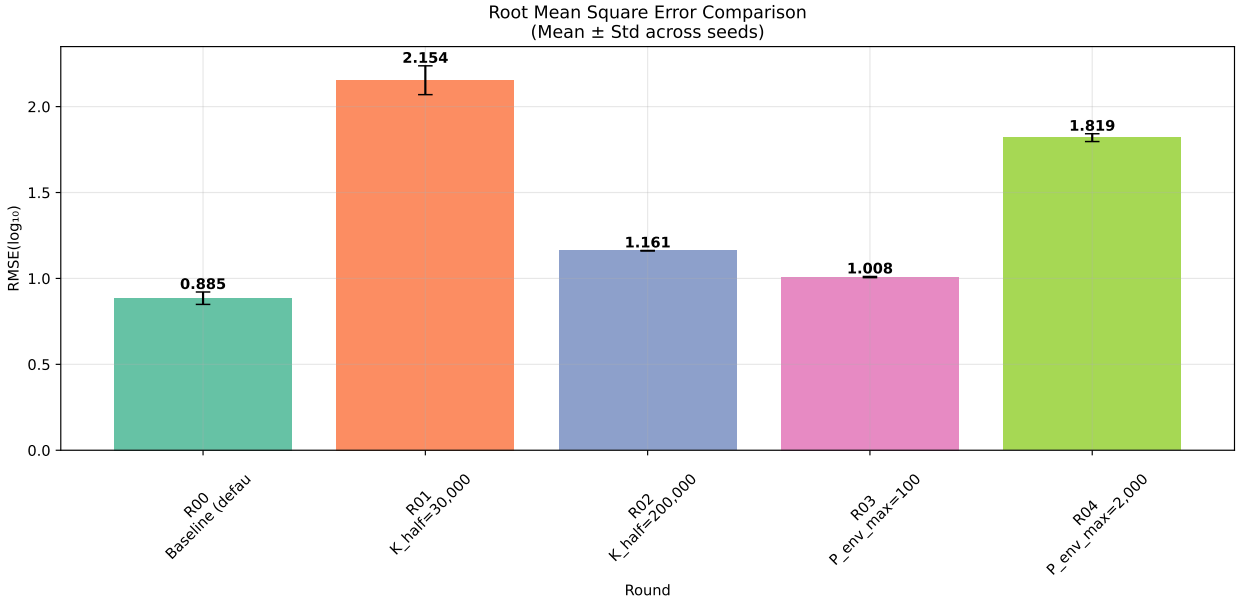

Figure 3: Root Mean Square Error comparison across calibration rounds. Error bars show standard deviation across seeds. Lower values indicate better performance.

3.5 Population Trajectories

Figure 4 shows the population recovery trajectories across all calibration rounds (R00-R04) and target regions. This multi-panel comparison reveals how different parameter modifications affect the temporal dynamics of recovery.

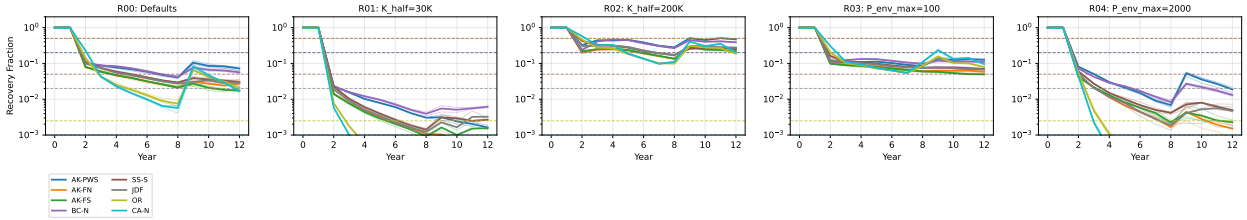

Figure 4: Population recovery trajectories across all calibration rounds. Each panel shows one round with different parameter settings: R00 (Defaults), R01 (K_half=30K), R02 (K_half=200K), R03 (P_env_max=100), R04 (P_env_max=2000). For each region, thin lines show individual seed trajectories, thick lines show the mean across seeds, and horizontal dashed lines indicate target recovery fractions. Note the logarithmic y-axis spanning 0.1% to 200% recovery.

3.6 Model Performance Scatter Plot

The scatter plot (Figure 5) provides a comprehensive view of model performance across all regions, seeds, and rounds. The proximity to the 1:1 line indicates agreement with targets.

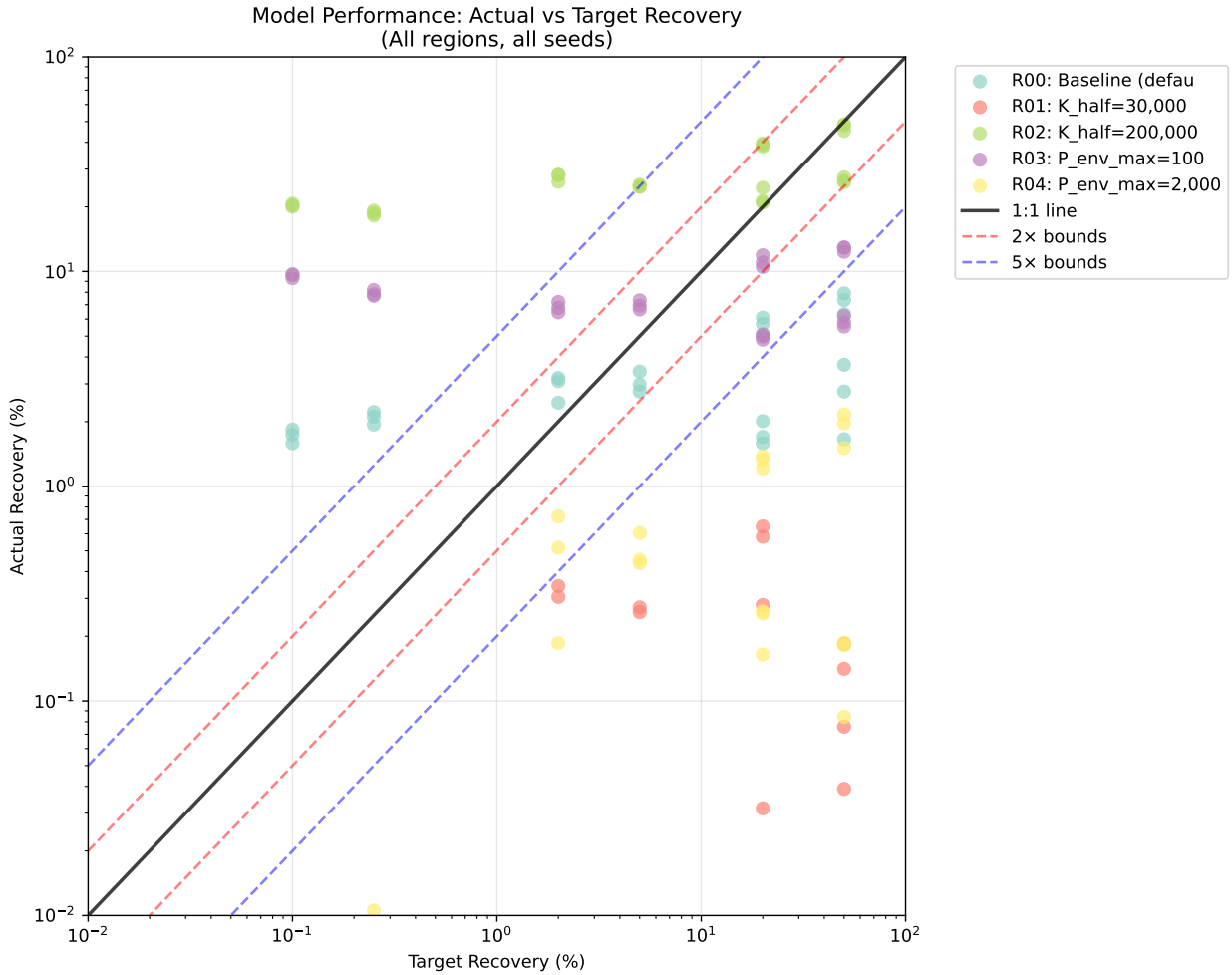

Figure 5: Scatter plot of actual versus target recovery percentages on log-log axes. Each point represents one region from one seed run. The 1:1 line indicates perfect agreement, while dashed lines show 2× and 5× bounds.

3.7 Individual Round Trajectory Details

Figures 6-10 provide detailed views of the population trajectories for each calibration round, showing the temporal dynamics more clearly for individual analysis.

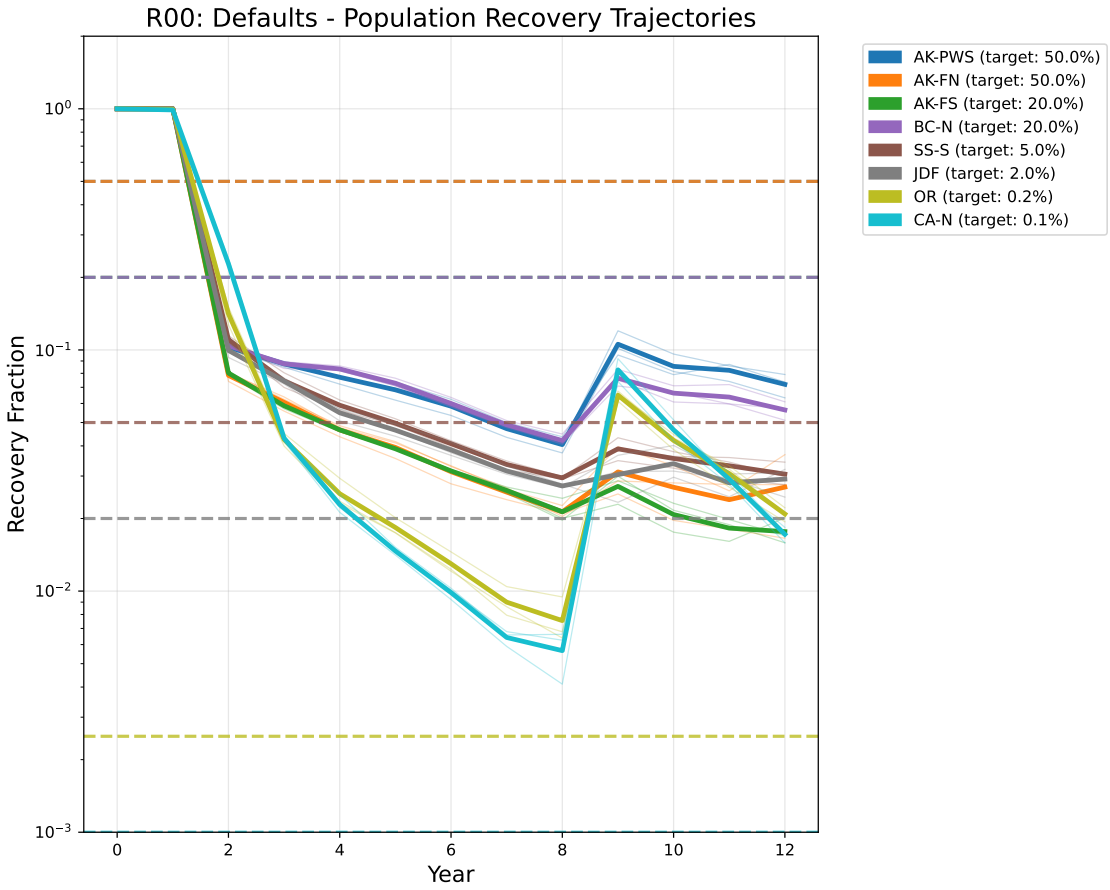

Figure 6: Round 00 (Defaults) - Detailed population recovery trajectories across all target regions.

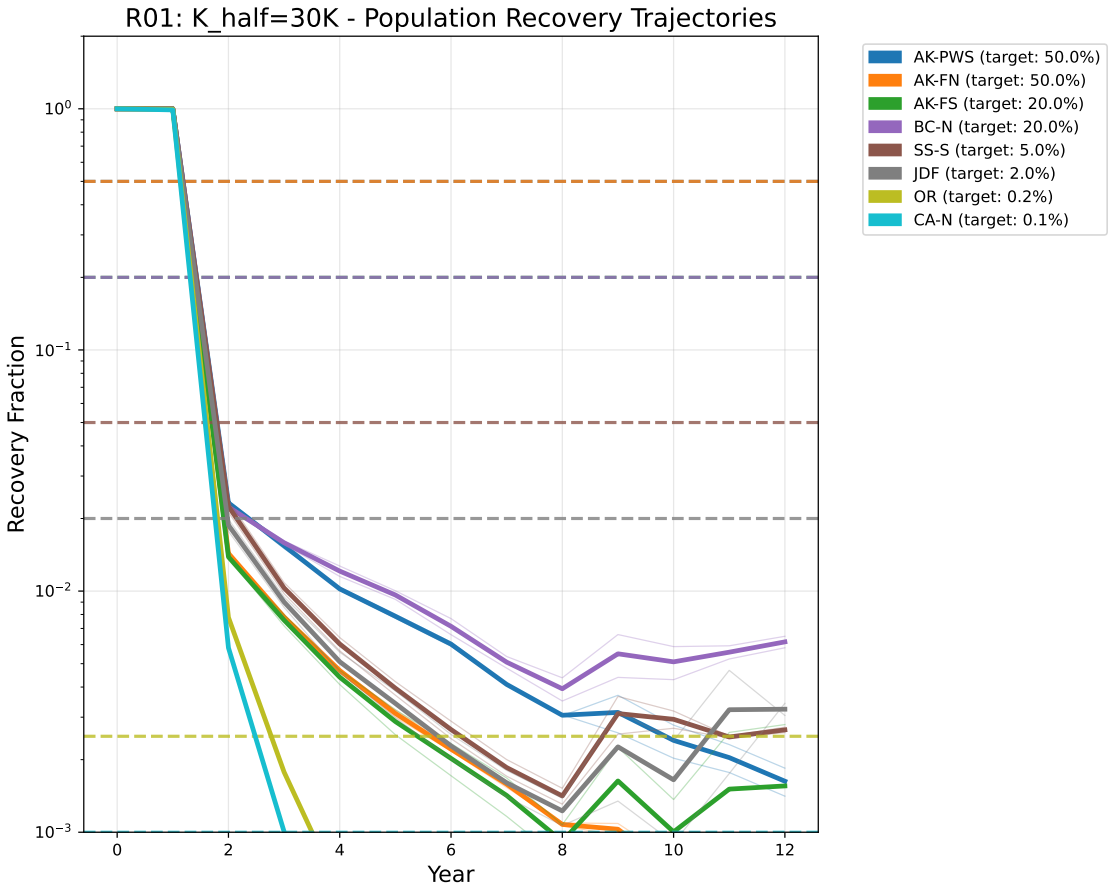

Figure 7: Round 01 (K_half=30K) - Detailed population recovery trajectories across all target regions.

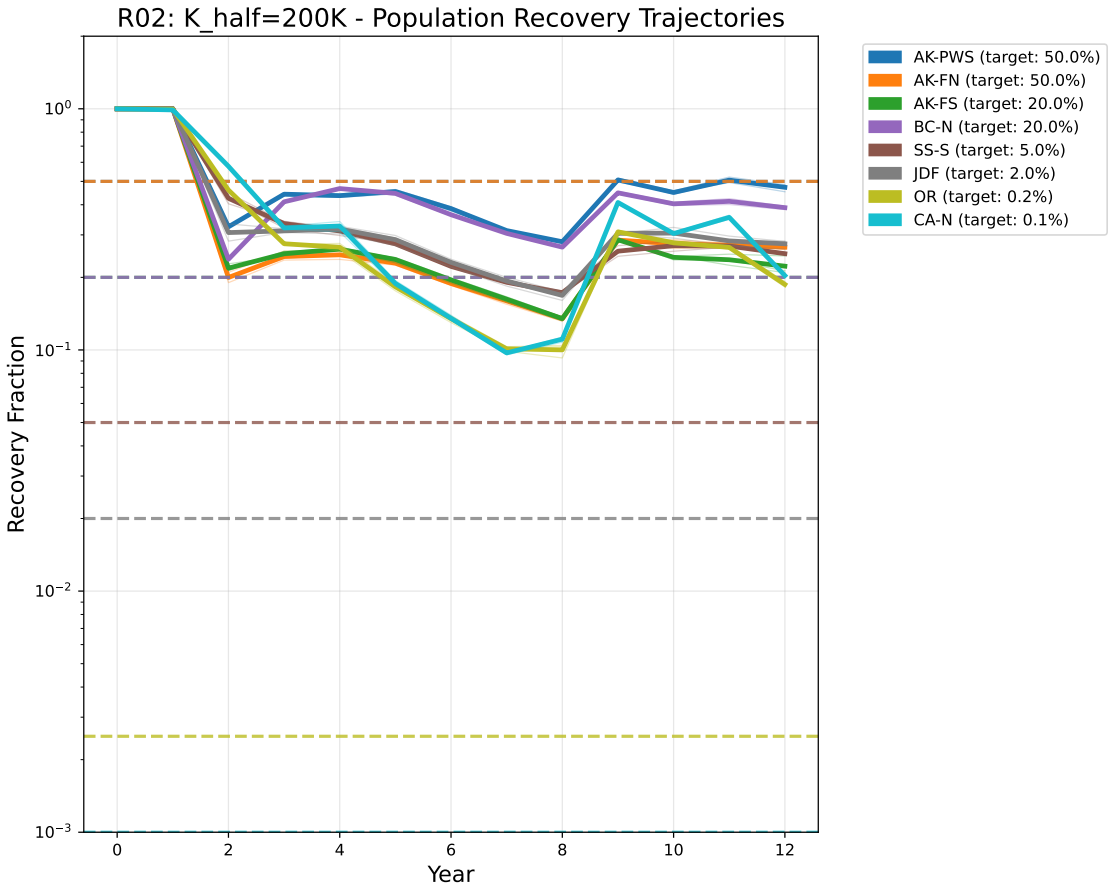

Figure 8: Round 02 (K_half=200K) - Detailed population recovery trajectories across all target regions.

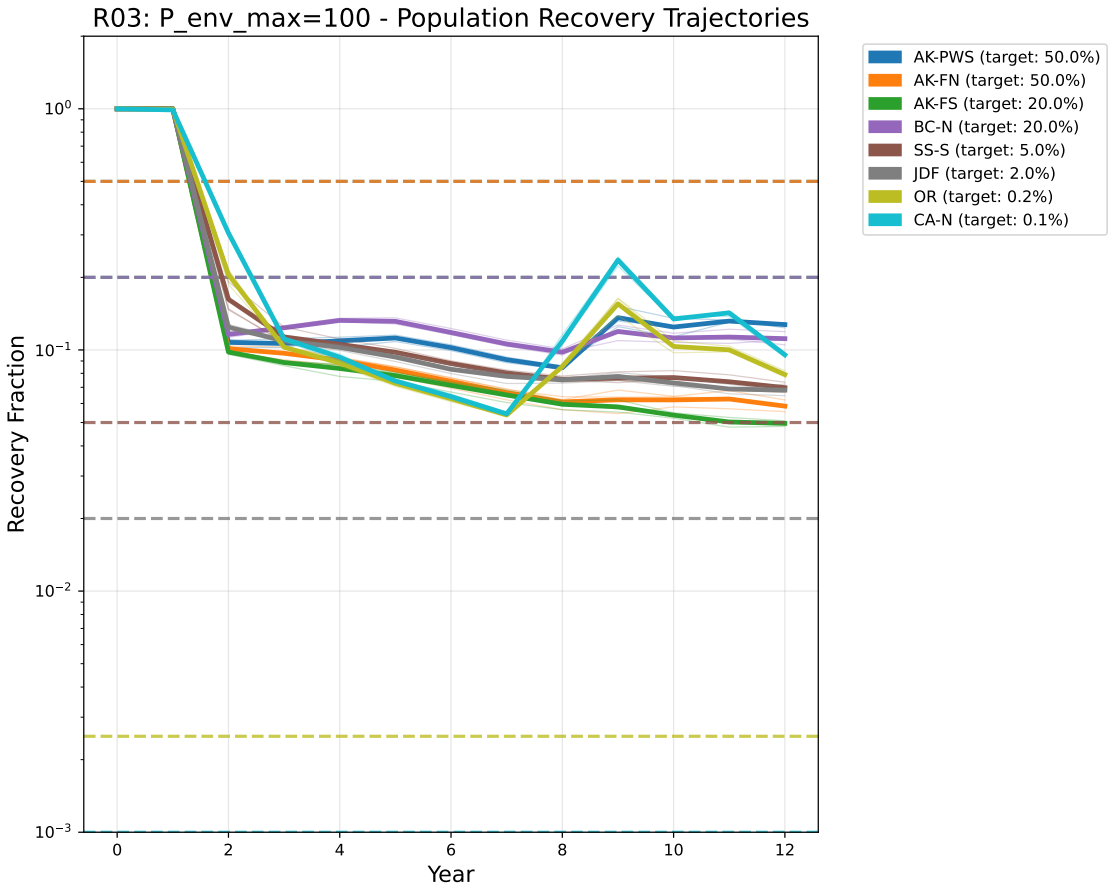

Figure 9: Round 03 (P_env_max=100) - Detailed population recovery trajectories across all target regions.

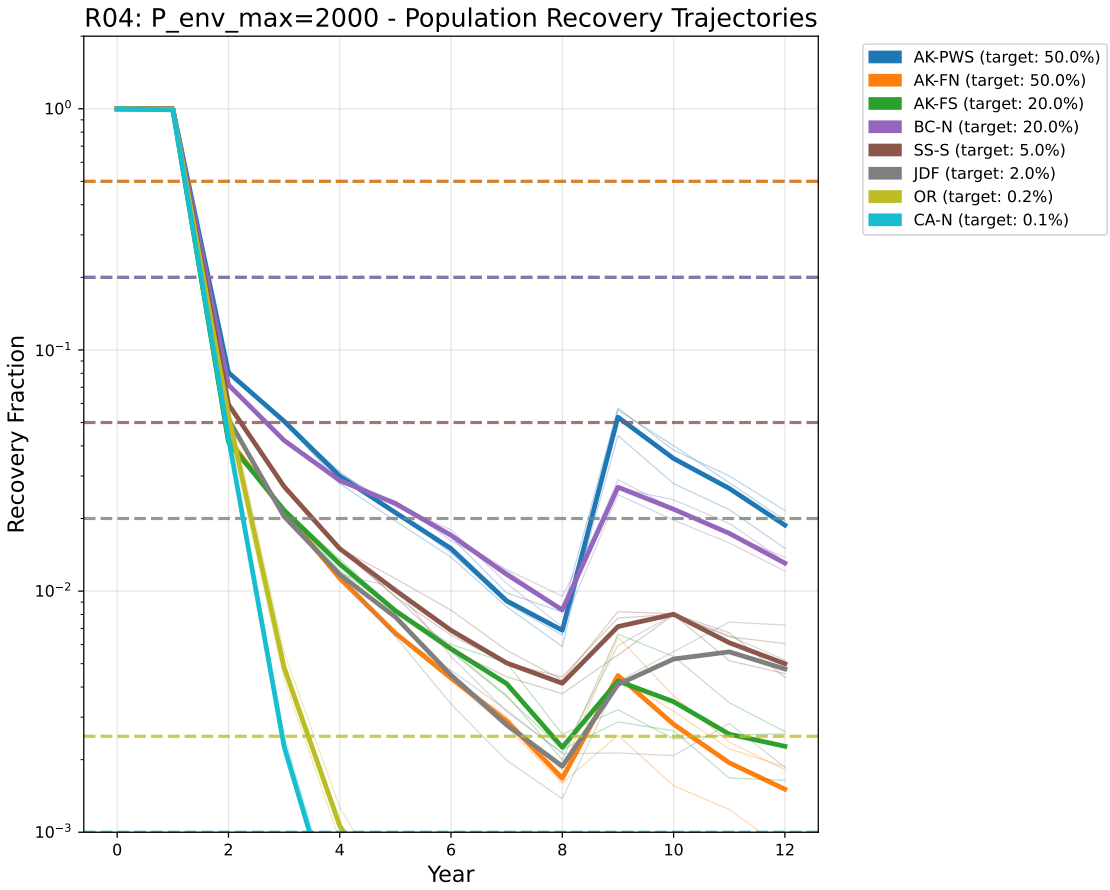

Figure 10: Round 04 (P_env_max=2000) - Detailed population recovery trajectories across all target regions.

4 Discussion

4.1 Key Findings

The calibration results reveal several important patterns:

1. **Northern regions well-captured:** Alaska recovery targets (50%, 50%, 20%) are reasonably well matched by multiple parameter combinations, particularly Round 02 (K_half=200,000).

2. **Lack of latitudinal gradient:** The model fails to reproduce the steep decline in recovery fractions from Alaska to California. Southern regions consistently show recovery fractions around 20% when target values are <1%.

3. **Environmental pathogen ceiling effects:** Rounds 03 and 04 testing P_env_max showed mixed results, with lower values (100) performing better than higher values (2,000).

4. **Density-dependent transmission:** Higher K_half values (Round 02) improved northern region matching but did not address the latitudinal gradient problem.

4.2 Model Limitations

The primary limitation revealed by this calibration is the **insufficient north-south differentiation** in recovery dynamics. The current model configuration produces relatively uniform recovery fractions across latitude, failing to capture the dramatic gradient observed in expert estimates.

This suggests that the VBNC (Viable But Non-Culturable) temperature sigmoid mechanism, intended to create temperature-dependent pathogen persistence, is not creating sufficient differentiation between northern and southern regions.

4.3 Mechanistic Implications

The results suggest that:

- Current temperature dependence functions may be too weak
- Additional latitudinal factors (ocean chemistry, host density, pathogen strain differences) may need to be incorporated
- The relationship between environmental conditions and pathogen persistence requires refinement

4.4 Next Steps

Based on these findings, future calibration rounds should focus on:

1. **Enhanced temperature dependence:** Strengthen the VBNC temperature sigmoid or add additional temperature-dependent mechanisms.
2. **Regional parameterization:** Consider region-specific parameter values rather than global parameters.
3. **Additional environmental factors:** Incorporate ocean chemistry, salinity, or other environmental gradients.
4. **Host-pathogen coevolution:** Explore whether evolutionary responses differ by latitude.
5. **Completion of Round 05:** The combination parameter set (Round 05) should be completed and analyzed when computational resources allow.

5 Conclusions

The first SSWD-EvoEpi calibration batch successfully demonstrates the model's ability to reproduce northern recovery patterns but reveals critical limitations in capturing the latitudinal gradient of recovery. Round 02 (K_half=200,000) provided the best balance of northern region accuracy, though all tested parameter combinations failed to achieve the low recovery fractions required for southern regions.

The systematic overprediction in southern regions indicates that the current temperature-dependent mechanisms are insufficient to create the observed north-south gradient. Future model development should prioritize enhancing latitudinal differentiation mechanisms while preserving the successful northern region dynamics.

This calibration provides a solid foundation for model refinement and highlights specific areas for mechanistic improvement in subsequent development cycles.