

SSWD-EvoEpi: Wavefront Calibration Report

Rounds W01–W04: Dispersal Kernel Sweep with Wavefront Disease Propagation

SSWD-EvoEpi Calibration Pipeline

February 28, 2026

Abstract

We report results from the first four wavefront-enabled calibration rounds (W01–W04) of the SSWD-EvoEpi agent-based model for sea star wasting disease in *Pycnopodia helianthoides*. These rounds sweep the pathogen dispersal kernel (D_P) from 50 km to 200 km while holding the wavefront activation threshold at 1.0. The wavefront mechanism successfully produces south-to-north disease progression, but propagation is approximately $2\times$ too fast: all regions are reached within 12–20 months regardless of D_P , whereas observed timing spans 6–42 months. The recovery gradient remains flat across regions (~ 21 – 46%), failing to reproduce the 3-order-of-magnitude gradient from California ($<0.1\%$) to Alaska ($\sim 50\%$) seen in field data. RMSE on log-scale recovery is nearly identical across rounds (~ 1.19), indicating that the dispersal kernel has minimal influence on outcomes when the activation threshold is too permissive. We recommend increasing the activation threshold to slow propagation and exploring T_{VBNC} to steepen the recovery gradient.

Contents

1	Introduction	2
2	Methods	2
2.1	Wavefront Mechanism	2
2.2	Parameter Configuration	2
2.3	Calibration Targets	3
2.3.1	Recovery Fractions	3
2.3.2	Arrival Timing	3
2.3.3	Scoring	3
3	Results	3
3.1	Recovery Fractions	3
3.2	Wavefront Arrival Timing	5
3.3	RMSE Comparison	7
3.4	Cross-Seed Consistency	8
4	Discussion	9
4.1	Why the Wavefront Is Too Fast	9
4.2	Why the Recovery Gradient Is Flat	9
4.3	What D_P Does and Does Not Control	9

5	Next Steps	10
5.1	Priority 1: Increase Activation Threshold	10
5.2	Priority 2: T_{VBNC} Tuning for Recovery Gradient	10
5.3	Priority 3: Combined Approach	10
5.4	Positive Findings to Preserve	10

1 Introduction

Sea star wasting disease (SSWD) caused catastrophic population declines in *Pycnopodia helianthoides* beginning in 2013 along the Pacific coast of North America. The disease spread as a waterborne wavefront from Southern California northward to Alaska over approximately 3.5 years [?]. Southern populations experienced near-total mortality with minimal recovery, while Alaska populations retained ~50% of pre-disease abundance.

The SSWD-EvoEpi model is a coupled eco-evolutionary epidemiological agent-based model that simulates disease transmission, host population dynamics, and trait evolution (resistance, tolerance, and recovery rate) across a spatial network of habitat nodes spanning the species' range. Previous calibration rounds used simultaneous disease introduction across all regions; rounds W01–W04 represent the first implementation of **wavefront disease propagation**, where disease seeds at origin nodes in Southern California and propagates northward as a waterborne wavefront.

These four rounds sweep the pathogen dispersal kernel distance (D_P) to evaluate its effect on wavefront speed and disease severity:

- **W01:** $D_P = 50$ km, $D_{P,\max} = 175$ km
- **W02:** $D_P = 100$ km, $D_{P,\max} = 350$ km
- **W03:** $D_P = 150$ km, $D_{P,\max} = 525$ km
- **W04:** $D_P = 200$ km, $D_{P,\max} = 700$ km

All rounds share: $K_{\text{half}} = 200,000$, wavefront enabled, activation threshold = 1.0, 78 origin nodes in the CA-S region, $K = 5,000$ individuals per node, 13-year simulation with disease introduction in year 1.

2 Methods

2.1 Wavefront Mechanism

In the wavefront model, disease is initially seeded only at designated origin nodes (78 nodes in the CA-S region, corresponding to Southern California). Disease then spreads to adjacent regions through pathogen dispersal, but a node can only become *activated* (disease-permissive) when the cumulative pathogen pressure from already-infected neighbors exceeds the **activation threshold**. Once activated, the node follows normal disease dynamics. This creates a traveling wavefront of disease that propagates outward from the origin.

2.2 Parameter Configuration

Table 1 summarizes the parameter settings for each round. The only parameters varied between rounds are D_P (characteristic dispersal distance) and $D_{P,\max}$ (maximum dispersal range, set at $3.5 \times D_P$).

Table 1: Parameter configurations for rounds W01–W04.

Parameter	W01	W02	W03	W04
D_P (km)	50	100	150	200
$D_{P,\max}$ (km)	175	350	525	700
K_{half}		200,000		
Wavefront enabled		true		
Activation threshold		1.0		
Origin nodes		78 (CA-S region)		
K (per node)		5,000		
Simulation years		13		
Seeds		42, 123, 999		

2.3 Calibration Targets

2.3.1 Recovery Fractions

Recovery fractions represent the proportion of pre-disease population remaining at the end of the 13-year simulation. Targets are derived from Gravem et al. (2021) field observations and expert estimates, spanning three orders of magnitude from CA-N (0.1%) to AK-PWS (50%).

2.3.2 Arrival Timing

Arrival timing targets represent the number of months after origin seeding when disease first reaches each region. Targets range from 0 months (CA-S, origin) to 42 months (AK-WG, AK-AL).

2.3.3 Scoring

Model fit is evaluated using RMSE on \log_{10} -transformed recovery fractions across 8 target regions. Arrival timing fit is assessed using mean absolute error (MAE) in months across 17 regions.

3 Results

3.1 Recovery Fractions

Figure 1 shows recovery fractions for all four rounds compared to targets. The most striking finding is the **flat recovery gradient**: all model regions show recovery fractions between 21–46%, regardless of their distance from the origin. The targets span 0.1% (CA-N) to 50% (AK-PWS), a 500-fold range.

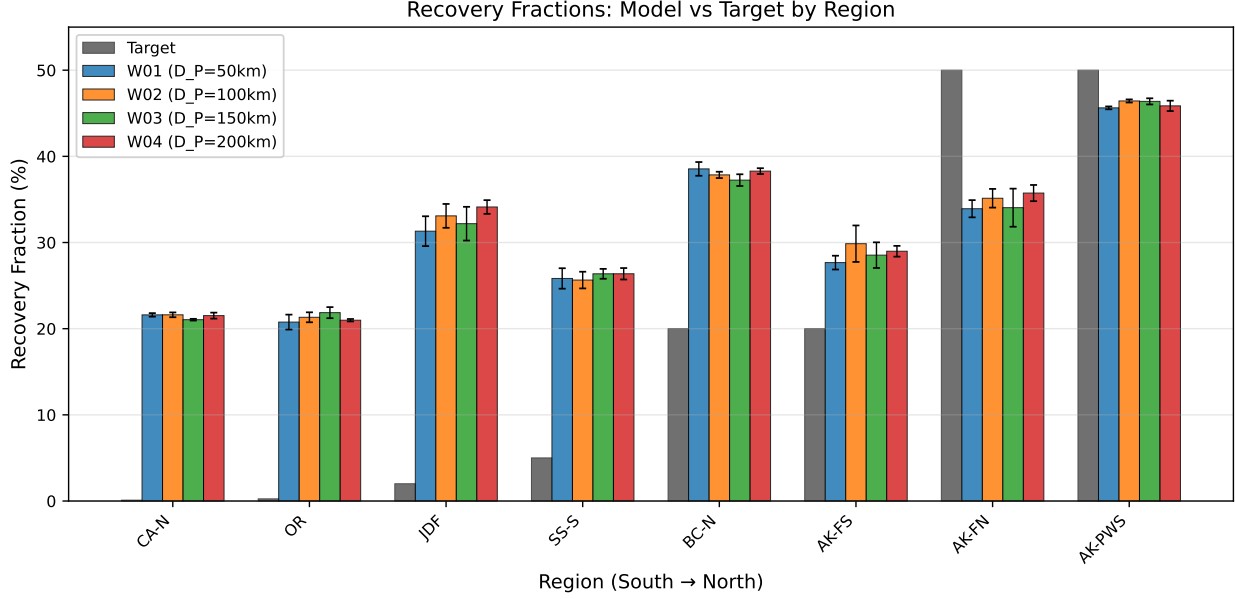


Figure 1: Recovery fractions by region for rounds W01–W04 compared to calibration targets. Regions are ordered south to north. Error bars show ± 1 standard deviation across 3 seeds. The model produces a nearly flat recovery gradient, failing to capture the steep south-to-north gradient observed in the field.

Table 2: Recovery fractions (%) by region and round (mean \pm std across 3 seeds).

Region	Target	W01	W02	W03	W04
CA-N	0.10	21.6 \pm 0.2	21.6 \pm 0.3	21.0 \pm 0.1	21.5 \pm 0.3
OR	0.25	20.8 \pm 0.9	21.3 \pm 0.6	21.9 \pm 0.6	21.0 \pm 0.1
JDF	2.00	31.3 \pm 1.7	33.1 \pm 1.4	32.2 \pm 2.0	34.1 \pm 0.8
SS-S	5.00	25.8 \pm 1.2	25.6 \pm 1.0	26.4 \pm 0.6	26.4 \pm 0.7
BC-N	20.00	38.5 \pm 0.8	37.8 \pm 0.4	37.2 \pm 0.7	38.3 \pm 0.3
AK-FS	20.00	27.7 \pm 0.8	29.9 \pm 2.1	28.5 \pm 1.5	29.0 \pm 0.6
AK-FN	50.00	33.9 \pm 1.0	35.1 \pm 1.1	34.0 \pm 2.2	35.7 \pm 0.9
AK-PWS	50.00	45.6 \pm 0.2	46.4 \pm 0.2	46.4 \pm 0.4	45.9 \pm 0.6

Key observations:

- **AK-PWS is close to target:** 45.6–46.4% modeled vs. 50% target, confirming the $K_{\text{half}} = 200,000$ setting from prior rounds.
- **Southern regions far too high:** CA-N at $\sim 21\%$ (target 0.1%), OR at $\sim 21\%$ (target 0.25%), representing $200\times$ overestimation.
- **D_P has negligible effect:** Recovery fractions are virtually identical across all four rounds, varying by less than 2 percentage points for any region.

Figure 2 presents the same data as a log-log scatter plot, clearly showing the divergence from the 1:1 line. Northern regions (AK-PWS, BC-N, AK-FS) fall near the diagonal, while southern regions are displaced by 1–2 orders of magnitude.

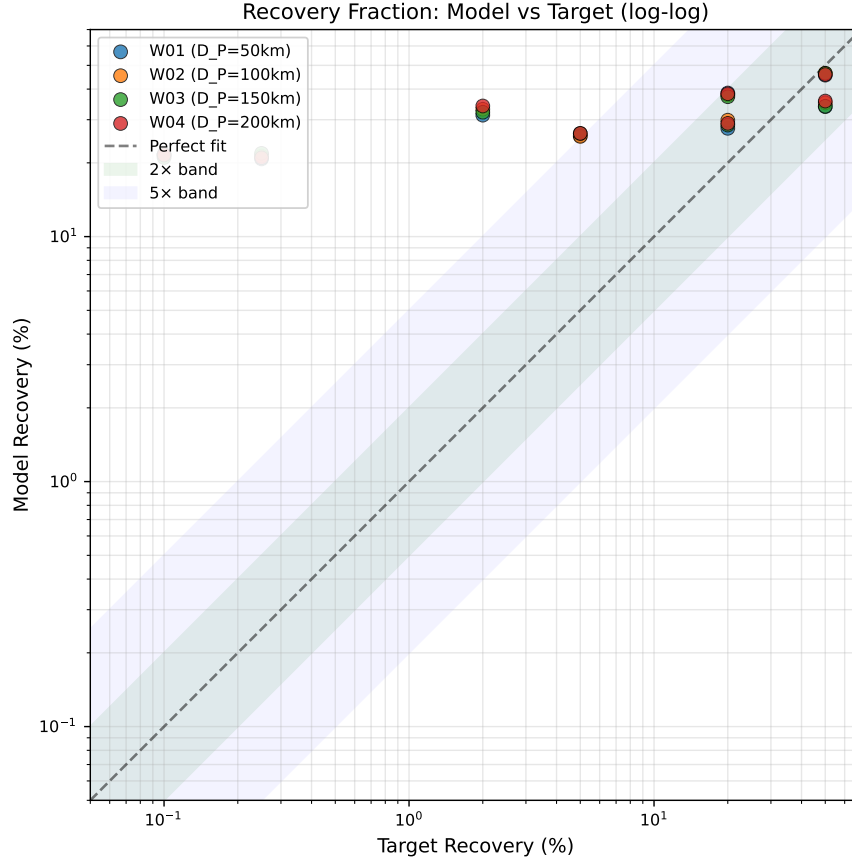


Figure 2: Model vs. target recovery fractions on log-log axes. Perfect agreement falls on the dashed diagonal. Northern regions (upper right) show reasonable fit; southern regions (lower left targets) show massive overestimation.

3.2 Wavefront Arrival Timing

Figure 3 shows arrival timing for all regions compared to targets. All four rounds produce substantially faster wavefront propagation than observed.

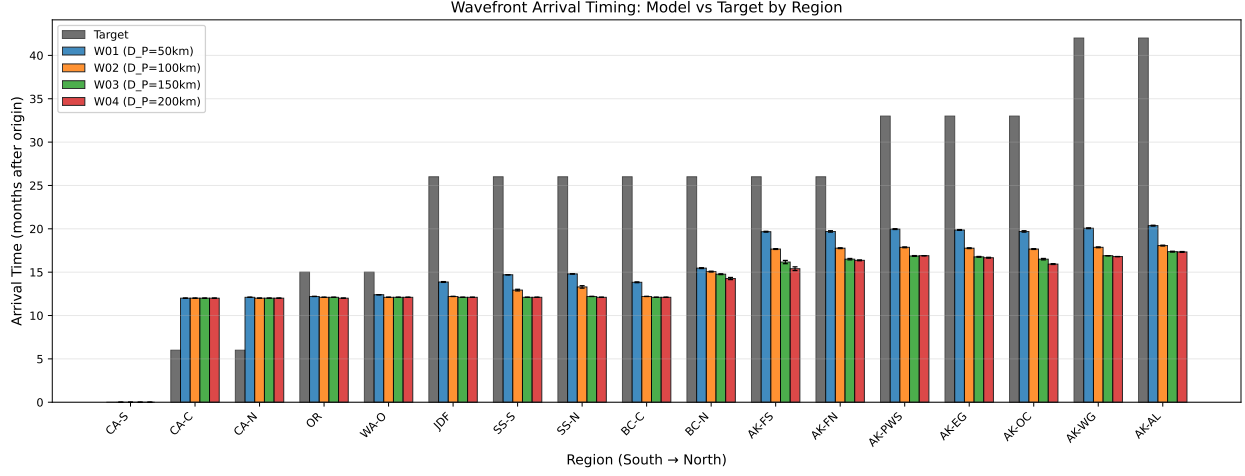


Figure 3: Wavefront arrival timing by region for rounds W01–W04 compared to targets. The model consistently reaches all regions faster than the observed timing, with the discrepancy growing from south to north.

The wavefront propagation is shown most clearly in Figure 4, the key diagnostic figure for these runs.

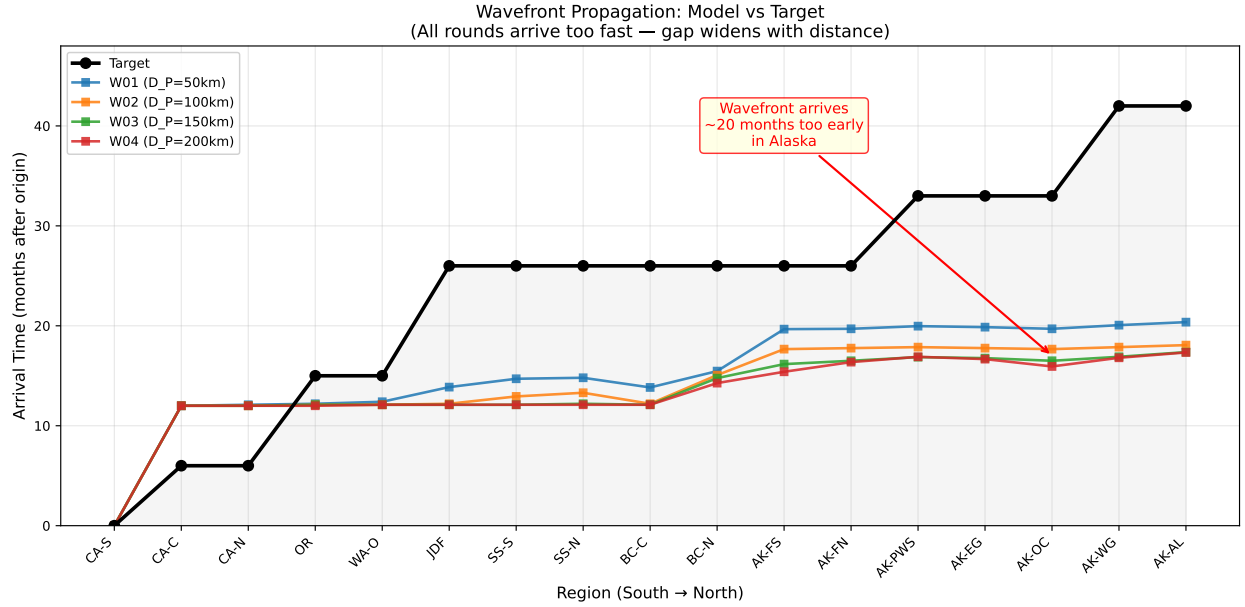


Figure 4: Wavefront propagation: arrival month vs. region for all rounds and targets. The target curve rises steeply from 0 to 42 months over the full range; all model curves plateau at 12–20 months. The gap widens dramatically for northern regions. Paradoxically, higher D_P produces *faster* propagation.

Table 3: Arrival timing (months) by region and round (mean across 3 seeds). Standard deviations are <0.2 months for all entries.

Region	Target	W01	W02	W03	W04
CA-S	0	0.0	0.0	0.0	0.0
CA-C	6	12.0	12.0	12.0	12.0
CA-N	6	12.1	12.0	12.0	12.0
OR	15	12.2	12.1	12.1	12.0
WA-O	15	12.4	12.1	12.1	12.1
JDF	26	13.9	12.2	12.1	12.1
SS-S	26	14.7	12.9	12.1	12.1
SS-N	26	14.8	13.3	12.2	12.1
BC-C	26	13.8	12.2	12.1	12.1
BC-N	26	15.5	15.1	14.8	14.3
AK-FS	26	19.7	17.7	16.2	15.4
AK-FN	26	19.7	17.8	16.5	16.4
AK-PWS	33	20.0	17.9	16.9	16.9
AK-EG	33	19.9	17.8	16.8	16.7
AK-OC	33	19.7	17.7	16.5	15.9
AK-WG	42	20.1	17.9	16.9	16.8
AK-AL	42	20.4	18.1	17.4	17.3

Key timing observations:

- **CA-C and CA-N arrive at exactly 12.0 months** in all rounds—double the 6-month target. Since disease is introduced in year 1, this suggests the wavefront reaches CA-C/CA-N at the start of the second simulation year.
- **Paradoxical D_P effect on speed:** Increasing D_P makes the wavefront *faster*, not slower. W01 ($D_P = 50$ km) reaches AK-AL at 20.4 months; W04 ($D_P = 200$ km) reaches it at 17.3 months. This occurs because longer-range dispersal leapfrogs intermediate nodes, allowing the wavefront to skip ahead.
- **Wavefront compresses in the north:** The target timing spans 0–42 months (42-month range), but all model rounds compress arrival into 0–20 months (20-month range at best). Alaska regions that should arrive 33–42 months post-origin arrive at 16–20 months.
- **Mean absolute error increases with D_P :** MAE is 10.0 months (W01), 11.3 months (W02), 11.9 months (W03), and 12.1 months (W04).

3.3 RMSE Comparison

Figure 5 shows the RMSE on log-scale recovery across rounds. All values are nearly identical at ~ 1.19 .

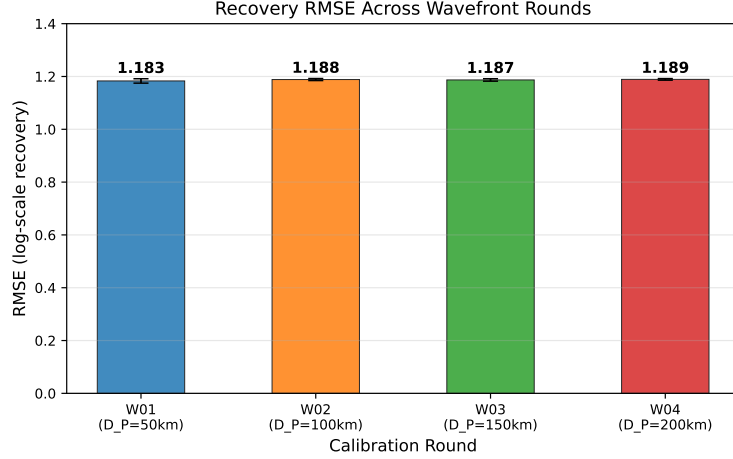


Figure 5: RMSE on \log_{10} -transformed recovery fractions across rounds. Error bars show ± 1 std across 3 seeds. All rounds are indistinguishable, indicating D_P does not materially affect disease severity outcomes when the activation threshold permits rapid spread.

3.4 Cross-Seed Consistency

Figure 6 shows per-seed variability for both recovery fractions and arrival timing.

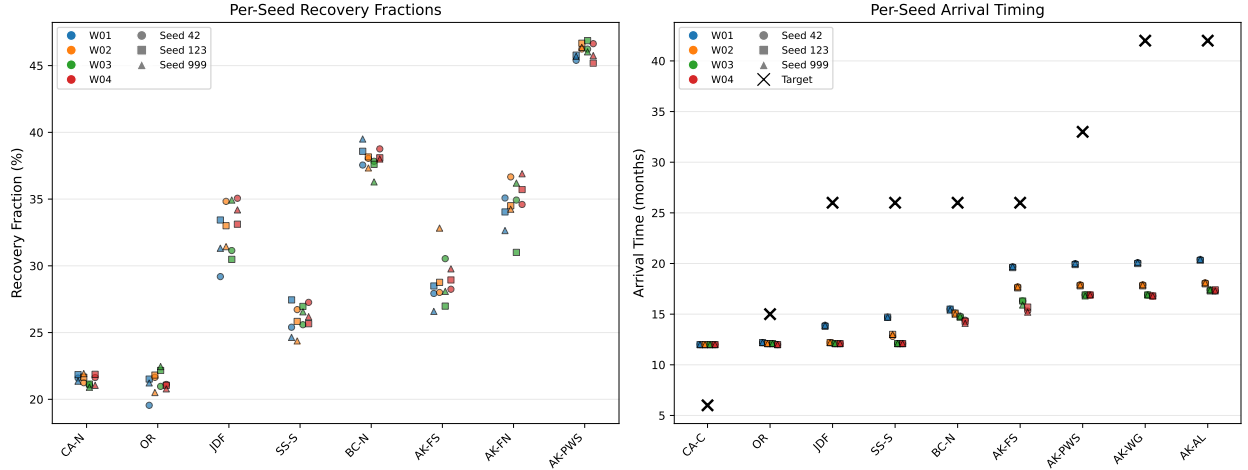


Figure 6: Per-seed variability in recovery fractions (left) and arrival timing (right) across all rounds. Variability is remarkably low: recovery fractions vary by < 2 percentage points and timing by < 0.5 months across seeds for any given round.

The low cross-seed variability indicates that results are robust and that the observed patterns are deterministic consequences of the parameter settings rather than stochastic artifacts. With 5,000 individuals per node and 78 origin nodes ($\sim 390,000$ origin-region agents), the system is large enough to suppress stochastic fluctuations.

4 Discussion

4.1 Why the Wavefront Is Too Fast

The activation threshold of 1.0 is too permissive. With this setting, even minimal pathogen pressure from a single infected neighbor is sufficient to activate a node. The wavefront therefore propagates at the speed of pathogen diffusion rather than being gated by a buildup threshold. Since pathogen dispersal occurs every time step, the wavefront advances rapidly regardless of D_P .

The 12.0-month arrival time for CA-C/CA-N (adjacent to the CA-S origin) suggests that the wavefront crosses region boundaries within 1 simulation year. Given that the simulation uses annual time steps for reporting, this represents the minimum detectable propagation time. The actual within-year dynamics likely produce even faster spreading.

The paradoxical result that higher D_P produces *faster* arrival is explained by the dispersal kernel’s maximum range ($D_{P,\max} = 3.5 \times D_P$). At $D_P = 200$ km, pathogen can disperse up to 700 km in a single step, effectively skipping over intermediate regions and creating a more uniform arrival pattern.

4.2 Why the Recovery Gradient Is Flat

The wavefront mechanism controls *when* disease arrives at each region, but not *how severely* it impacts populations. Once a region is activated, it experiences the same disease dynamics regardless of its distance from the origin. Key mechanisms that should create a gradient are:

1. **Time for evolution:** Southern populations have longer exposure to disease and therefore more time for resistance/tolerance evolution. However, the model’s evolutionary rates are insufficient to create measurable differences over the 12–20 month timing spread.
2. **Dose-dependent severity:** The wavefront delivers similar pathogen loads to all regions once activated, because the threshold is low enough that any exposure triggers full disease dynamics.
3. **Environmental gradients:** The current model does not implement latitude-dependent environmental parameters that might modulate disease severity (e.g., temperature effects on pathogen virulence or host immune response).

The flat gradient problem existed in pre-wavefront runs (simultaneous disease introduction) and persists here, confirming that it is fundamentally a disease severity issue rather than a timing issue.

4.3 What D_P Does and Does Not Control

These results definitively show that the pathogen dispersal kernel distance (D_P) is *not* the primary control on either wavefront speed or disease severity at the current activation threshold. Increasing D_P from 50 km to 200 km (4 \times):

- Changes Alaska arrival by only ~ 3 months (20.1 \rightarrow 17.3 for AK-WG)
- Changes recovery fractions by $< 2\%$ for any region
- Changes RMSE by < 0.006 ($< 0.5\%$ relative change)

This indicates that the wavefront speed is limited by other factors—most likely the activation threshold and the spatial structure of the habitat network itself.

5 Next Steps

Based on these findings, we recommend the following calibration strategies:

5.1 Priority 1: Increase Activation Threshold

The activation threshold should be increased substantially (e.g., 10, 50, 100, 500) to slow wavefront propagation. Higher thresholds require more pathogen buildup before a node becomes activated, which should:

- Slow the wavefront to better match 6–42 month observed timing
- Create a more gradual south-to-north progression
- Potentially create a dose-dependent severity gradient if regions at the wavefront edge receive sub-lethal pathogen loads before full activation

Recommended sweep: W05–W08 with activation thresholds of 10, 50, 100, 500 at a fixed $D_P = 100$ km (intermediate value).

5.2 Priority 2: T_{VBNC} Tuning for Recovery Gradient

The T_{VBNC} parameter (viable-but-not-culturable duration) controls how long pathogen persists in the environment. Reducing T_{VBNC} for northern regions could allow disease pressure to decay faster at the wavefront edge, creating the severity gradient needed for the observed recovery pattern. This could be modulated by latitude or water temperature.

5.3 Priority 3: Combined Approach

The ultimate calibration will likely require simultaneously tuning:

- Activation threshold (wavefront speed)
- D_P (local dispersal range)
- T_{VBNC} or an environmental severity modifier (recovery gradient)
- Possibly latitude-dependent disease mortality rates

5.4 Positive Findings to Preserve

Two aspects should be preserved in future rounds:

- $K_{\text{half}} = 200,000$: Produces AK-PWS recovery of $\sim 46\%$, close to the 50% target.
- **Wavefront mechanism itself**: The qualitative south-to-north propagation pattern is correct; only the speed and severity gradient need adjustment.

References

1. Gravem, S.A., et al. (2021). *Pycnopodia helianthoides*. *The IUCN Red List of Threatened Species*.