



Computation of fish larvae self-recruitment in using forward- and backward-in-time particle tracking in a Lagrangian model (SWIM-v2.0) of the simulated circulation of Lake Erie (AEM3D-v1.1.2)

Wei Shi^{1,2}, Leon Boegman², Josef D. Ackerman³, Shiliang Shan^{2,4}, and Yingming Zhao^{2,5}

¹Deep Sea and Polar Fisheries Research Center, and Key Laboratory of Mariculture, Ministry of Education, Ocean University of China, Qingdao, 266100, China

²Environmental Fluid Dynamics Laboratory, Department of Civil Engineering, Queen's University, Kingston, ON K7L 3N6, Canada

³Physical Ecology Laboratory, Department of Integrative Biology, University of Guelph, Guelph, ON N1G 2W1, Canada

⁴Department of Physics and Space Science, Royal Military College of Canada, Kingston, ON K7K 7B4, Canada

⁵Aquatic Research and Monitoring Section, Ontario Ministry of Natural Resources and Forestry Lake Erie Fishery Station, Wheatley, ON N0P 2P0, Canada

Correspondence: Wei Shi (weishi@ouc.edu.cn)

Received: 20 November 2024 – Discussion started: 25 March 2025

Revised: 25 December 2025 – Accepted: 15 January 2026 – Published: 10 February 2026

Abstract. Accurately estimating self-recruitment (SR), which is the fraction of recruits at a location that originated locally, is fundamental to understanding population connectivity. Biophysical models typically compute SR by releasing larval particles from source locations and tracking them forward in time. However, forward-tracking studies employ a variety of particle-release strategies (random, constant, area-scaled or production-scaled), which often leads to ambiguous SR estimates. Using theoretical analysis supported by numerical simulations of Lake Whitefish (*Coregonus clupeaformis*) larvae in Lake Erie, we show that SR is inherently dependent on larval production at all source locations. As a result, SR cannot be computed unambiguously in forward-tracking models unless the true larval production is known and released from every source location. In contrast, empirical parentage-analysis studies estimate SR directly as the fraction of locally produced juveniles among those sampled at a settlement location, without quantifying larval production at all sources. Motivated by this, we computed SR using backward-in-time particle tracking from the settlement location. We demonstrate that SR estimated using backtracking is independent of the number of particles released, which provides considerable benefit; namely the use of any magnitude

of particles released and eliminating the need to identify all contributing sources or their larval output. This reduces the effort required to estimate SR and provides a theoretically consistent basis for its calculation. Our findings provide a theoretical framework for estimating SR and highlight the advantages of applying backtracking models for studying larval dispersal and recruitment.

1 Introduction

Self-recruitment (SR), local retention (LR), and relative local retention (RLR) are commonly used metrics to assess the tendency for self-maintenance of local populations (Botsford et al., 2009; Burgess et al., 2014; Cowen and Sponaugle, 2009; Lett et al., 2015; Hogan et al., 2012). Among these, SR [(locally produced settlement)/(settlement from all origins); Botsford et al., 2009] has been the most widely used descriptor of larval dispersal and is generally easier to measure experimentally (Burgess et al., 2014). However, LR [(locally produced settlement)/(total local larval release); Botsford et al., 2009] is strongly related to self-persistence or the

ability of a population to sustain itself (Burgess et al., 2014; Lett et al., 2015; D’aloia et al., 2017); and RLR [(locally produced settlement)/(settlement originating from that location)]; Hogan et al., 2012; Lett et al., 2015] has gained recent attention because its denominator includes only successful dispersal events (Almany et al., 2017; Shi et al., 2024). Although these metrics share a common numerator (the number of locally produced settlers), their different denominators yield differences in their interpretation and calculation (see Sect. 2).

Typically, it is challenging to compute LR and RLR because it is difficult to measure the total number of eggs, larvae, or tracer particles produced at a source location for LR, or the total number of larvae that survive and successfully settle for RLR (Botsford et al., 2009; Lett et al., 2015; Burgess et al., 2014; Hogan et al., 2012). Consequently, biophysical models are usually coupled with forward-in-time Lagrangian particle-tracking models (hereafter “forward tracking models”) to estimate LR and RLR (Chaput et al., 2022; Saint-Amand et al., 2023; Sato et al., 2023; Gurdek-Bas et al., 2022; Karnauskas et al., 2022; Karnauskas et al., 2013). LR is then computed as the ratio of larvae that settle at their release location to the number initially released and RLR is similarly computed by excluding larvae that settle in unsuitable nursery habitats (Gurdek-Bas et al., 2022). Both LR and RLR are independent of the true larval production at the source location, which allows the release of a random number of simulated particles for their calculation, as demonstrated in this research.

In contrast, SR is generally easier to determine and has been studied more extensively (Lett et al., 2015; Burgess et al., 2014; D’aloia et al., 2013). For example, parentage analysis assigns sampled juveniles to their parents according to genetic markers, thereby identifying source locations and quantifying the number of locally produced settlers. SR has also been computed using biophysical models (Hiddink et al., 2013; Dubois et al., 2016; Klein et al., 2016; Faillettaz et al., 2018; Lequeux et al., 2018; Meerhoff et al., 2018; Hidalgo et al., 2019; Wolanski et al., 2021; Saint-Amand et al., 2023; Michie et al., 2024; Nadal et al., 2024; Corrochano-Fraile et al., 2022; Sato et al., 2023; Paris et al., 2005). In these models, larval particles are released from each source location and tracked forward in time, and SR is calculated as the number of particles that both originate from and settle at a given location divided by the total number of particles that settled there. A variety of particle-release strategies have been used, including: (1) releasing a random number of larvae (Sato et al., 2023); (2) releasing an equal number per location (Dubois et al., 2016; Faillettaz et al., 2018); (3) releasing larvae in proportion to habitat area (D’agostini et al., 2015); and (4) releasing larvae in proportion to estimated larval production (Nolasco et al., 2022). Notably, while the denominator (the total number of settlers at a location) reflects larval production from all potential sources, the numerator (the number of larvae both released from and set-

ling at that location) depends on only the local larval production at the source. As a result, changes in larval production at any external source can alter the denominator while they cannot affect the numerator, and thereby influence SR estimates. This presents two challenges for forward-tracking approaches: identifying all possible source locations and determining the appropriate number of particles to release from each source. For these reasons, SR cannot be unambiguously computed using forward tracking models.

The objective of this research is to theoretically determine, in forward tracking simulations, whether LR and RLR and SR are independent of larval production at the source location. We examine the results of this endeavour and show that computing SR with a forward-tracking model produces ambiguous estimates regardless of the particle-release strategies. In contrast, we find that SR computed with a back-tracking model informed by observations of Lake Whitefish (*Coregonus clupeaformis*) larvae collected in Lake Erie provide unambiguous results, i.e., it is independent of the number of recruits at the settlement location. We believe that these results are broadly applicable to both freshwater and marine species with pelagic larval phases.

2 Theoretical Development

2.1 Self-recruitment from forward tracking models

We introduce only the minimum mathematical notation required to compare forward- and backward-tracking formulations of SR for clarity. Suppose that there are n locations associated with larval hatching and subsequent settlement or recruitment (locations $1, 2, \dots, i, j, \dots, n$). Let N'_i denote the number of eggs or newly hatched larvae produced at location i . After hatching, larvae enter a pelagic phase during which they are transported by water currents before settling. The dispersal rate from patch i to patch j , D_{ij} , is defined as the proportion of larvae released from location i that settle at location j . Because D_{ij} is defined as a proportion (i.e., a fraction of the total larvae released), it is inherently independent of N'_i .

Local retention (LR) at location i , commonly computed in forward-tracking simulations (Saint-Amand et al., 2023; Sato et al., 2023), is defined following Shi et al. (2024):

$$\text{LR}_i = \frac{D_{ii} N'_i}{N'_i} = D_{ii} . \quad (1)$$

This result follows directly from the proportional nature of D_{ij} : the number of locally produced settlers scales linearly with N'_i , and the scaling cancels in the ratio.

At the end of dispersal, some larvae settle at suitable nursery habitats, whereas others settle in unsuitable habitats and are considered “unsuccessful settlers” (Almany et al., 2017). Excluding the unsuccessful settlers from the denominator of LR yields relative local retention (RLR) (Hogan et al., 2012;

Lett et al., 2015):

$$\text{RLR}_i = \frac{D_{ii}N'_i}{\sum_{j=1}^n D_{ij}N'_i} = \frac{D_{ii}}{\sum_{j=1}^n D_{ij}}. \quad (2)$$

Here, $\sum_{j=1}^n D_{ij}$ may be less than 1 because settlement in unsuitable habitats is excluded. As with LR, the proportional form of D_{ij} ensures that RLR is theoretically independent of N'_i .

Self-recruitment (SR = (locally produced settlement)/(settlement from all origins)) is given by (Botsford et al., 2009; Lett et al., 2015; Almany et al., 2017):

$$\text{SR}_i = \frac{D_{ii}N'_i}{\sum_{j=1}^n D_{ji}N'_j}. \quad (3)$$

In this expression, the denominator depends on larval production at all source locations ($N'_1, N'_2, \dots, N'_n$), not just location i . Therefore, SR is inherently sensitive to spatial variation in larval production. Any change in larval output at any other source location modified the total number of settlers arriving at i , and thus alters SR, even when local retention at i remains unchanged.

Although SR is often described informally as “the probability that a source location equals a settlement location”, Eq. (3) provides a formal expression for that probability. A central finding of this study is that this probability becomes dependent on larval production when computed with forward tracking, whereas backtracking yields a production-independent estimate.

If larval production is artificially assumed to be identical at all source locations (i.e., $N'_1 = \dots = N'_n$), then SR simplifies to:

$$\text{SR}_i = \frac{D_{ii}}{\sum_{j=1}^n D_{ji}}. \quad (4)$$

This expression resembles both LR and RLR in being independent of N'_i , but it is only valid under the restrictive and often unrealistic assumption of uniform larval production. When larval production varies across locations, Eq. (3) must be used, and substantial discrepancies between Eqs. (3) and (4) can arise – especially when many source locations contribute recruits. Because LR and RLR depend only on dispersal proportions D_{ij} , their values can be computed unambiguously in forward-tracking models using any sufficiently large number of simulated particles released from each source. Releasing an arbitrary (random) number of particles N_i is acceptable because the proportional structure of D_{ij} ensures convergence as particle number increases. As noted by Shi et al. (2024), LR provides a minimum bound for RLR because LR includes both successful and unsuccessful settlers in the denominator.

In contrast, SR cannot be uniquely determined from a forward-tracking simulation unless the true larval production values N'_j at all source locations are known and used.

Without this information, any particle-release strategy (equal numbers, area-weighted numbers, proportional to habitat, random, etc.) yields SR values that reflect the assumed production pattern rather than the true biological quantity. Thus, forward-tracking models cannot generally compute SR unambiguously unless realistic production values at all source locations are available.

2.2 Self-recruitment from backward tracking models

An alternative to forward tracking is to release larval particles from settlement locations (i.e., where larvae are recruited into the population) and track them backward-in-time to identify their origins. Suppose there are m settlement and hatching locations ($1, 2, \dots, i, j, \dots, m$). Let M'_i be the true number of larvae recruited at settlement location i . The recruitment rate R'_{ij} is defined as the proportion of recruits at location i that originated from source location j . Self-recruitment (SR) at location i is then:

$$\text{SR}_i = \frac{R'_{ii}M'_i}{M'_i} = R'_{ii}. \quad (5)$$

Thus, SR is theoretically identical to the recruitment rate from the local source, and is independent of the number of recruits at that location. This property is biologically intuitive: in parentage-analysis studies, it is rarely possible to sample all recruits in the field, yet SR is consistently estimated as the fraction of sampled juveniles assigned to local parents (D’aloia et al., 2013; Almany et al., 2017). Even when only a subset of recruits is sampled, this ratio provides an unbiased estimate of SR, reinforcing that SR depends on the proportion of local recruits, not the absolute number.

In backtracking simulations, a random number of larval particles M_i is typically released from settlement location i . By the end of the simulation, some particles will trace back to suitable source locations, while others will terminate in unsuitable ones; the latter constitute “unreal” recruits (Shi et al., 2024). The settlement rate R_{ij} is defined as the ratio of the number of larval particles that settle at location j and originate from location i , divided by the total number of larval particles released from location i . The difference between R_{ij} and R'_{ij} lies in the denominator: the former normalizes by M_i whereas the latter normalizes by the true number of recruits M'_i , respectively.

If R_{ij} and M_i are substituted into Eq. (5), the result is not SR, but rather the imaginary self-recruitment (ISR), which is (Shi et al., 2024):

$$\text{ISR}_i = \frac{R_{ii}M_i}{M_i} = R_{ii}. \quad (6)$$

ISR represents a minimum bound on SR, as its denominator includes unreal recruits in addition to real ones. To obtain SR from a backward-tracking simulation, unreal recruits must be

excluded from the denominator of ISR. This yields:

$$SR_i = \frac{R_{ii} M_i}{\sum_{j=1}^n R_{ij} M_j} = \frac{R_{ii}}{\sum_{j=1}^n R_{ij}}. \quad (7)$$

As Eq. (7) shows, SR computed by backward tracking is theoretically independent of the number of larval particles released from location i . This independence arises because increasing or decreasing M_i rescales the numerator and denominator equally; the settlement rates R_{ij} remain unchanged. This is confirmed in the numerical examples presented later. Consequently, a random number of particles can be released from the settling location without impacting SR.

The relative local retention (RLR) at location i in backward-tracking simulations is:

$$RLR_i = \frac{R_{ii} M_i}{\sum_{j=1}^n R_{ji} M_j}. \quad (8)$$

Unlike forward tracking (Eq. 2), RLR derived from backward tracking depends on the number of particles released from each location (M_i). This is expected because backward tracking does not redistribute particles across all source locations in proportion to their true production. Finally, backward tracking cannot compute LR because unsuccessful dispersal does not occur in backward time, just as forward tracking cannot compute ISR because unreal recruitment cannot occur in forward time.

2.3 Number of larvae produced and recruited at each location

The number of larvae produced at each source location N'_j may be computed for use in forward-tracking simulations to obtain unbiased estimates of SR. This can be inferred from the number of recruits observed at settlement locations, M'_i , together with the recruitment matrix estimated from backward tracking simulations.

Consider a backward-tracking simulation in which M_i particles are released from settlement location i , of these, $M_i \cdot \sum_{k=1}^n R_{ik}$ particles trace back to real (suitable) hatching locations, and $M_i \cdot (1 - \sum_{k=1}^n R_{ik})$ trace back to unreal (unsuitable) locations. To ensure that the number of real recruits, in the simulation, matches the observed number or real recruits M'_i at settlement location i , the following must hold $M_i \cdot \sum_{k=1}^n R_{ik} = M'_i$. Thus, the number of particles released from location i in a backward-tracking simulation is $M_i = M'_i / \sum_{k=1}^n R_{ik}$. The number of recruits at location i that originated from source location j is therefore:

$$N'_{ji} = R_{ij} \cdot M_i = M'_i \frac{R_{ij}}{\sum_{k=1}^n R_{ik}} \quad (9)$$

From the dispersal rate D_{ji} (the proportion of larvae released at j that settle at i), the total larval production at source location j is:

$$N'_j = \frac{N'_{ji}}{D_{ji}} = M'_i \cdot \frac{R_{ij}}{D_{ji}} \cdot \frac{1}{\sum_{k=1}^n R_{ik}}. \quad (10)$$

Interestingly, if we instead use the number of recruits observed at a different settlement location a (denoted is M'_a), an equivalent expression for N'_j is obtained:

$$N'_j = M'_a \cdot \frac{R_{aj}}{D_{ja}} \cdot \frac{1}{\sum_{k=1}^n R_{ak}}. \quad (11)$$

Equations (10) and (11) must agree for all recruitment locations i and a where both R_{ij} and R_{aj} are non-zero. Equating Eqs. (10) and (11) we obtain the number of recruits at location a as:

$$M'_a = M'_i \cdot \frac{R_{ij}}{D_{ji}} \cdot \frac{D_{ja}}{R_{aj}} \cdot \sum_{k=1}^n \frac{R_{ak}}{R_{ik}}. \quad (12)$$

This expression provides a constraint linking the number of recruits at different settlement locations. Both N'_j and M'_a are undefined when the dispersal, recruitment or settlement rates become zero. This reflects the fact that a location's production cannot be inferred from recruitment locations to which it contributes zero larvae, and vice versa.

3 Putting Theory into Practice: Application to Lake Erie

3.1 Study area

Shi et al. (2024) identified the principal Lake Whitefish larval hatching locations in Lake Erie using backward-tracking simulations. The locations were concentrated along the western and southern flanks of the western basin. Considering that Lake Whitefish spawn on hard substrates (Amidon et al., 2021), we selected four regions containing suitable substrate along these flanks as potential hatching locations (Fig. 1). These regions served as larval-release areas in our forward-tracking simulations. We refer to them as: Detroit River Mouth (region A), Western Shoreline (region B), Midlake Reefs (region C), and Bass Islands (region D) (Fig. 1c). The Midlake Reef and Bass Islands regions were also selected as settlement locations, from which larval particles were released in backward-tracking simulations. In this study, "settlement locations" refer specifically to regions C and D, identified in Shi et al. (2024) as primary juvenile collection zones based on empirical larval surveys.

3.2 The hydrodynamic model

Larval particle transport was driven by output from a hydrostatic three-dimensional Reynolds-averaged Navier–Stokes hydrodynamic model, the Aquatic Ecosystem Model (AEM3D) (HydroNumerics; <https://www.hydronumerics.com.au>, last access: 22 January 2026). AEM3D was applied in a continuous hindcast for 2017–2019 resolving water temperature and currents on a 500 m × 500 m horizontal grid with 45 vertical layers (Shi et al., 2024; Lin et al., 2022). Vertical

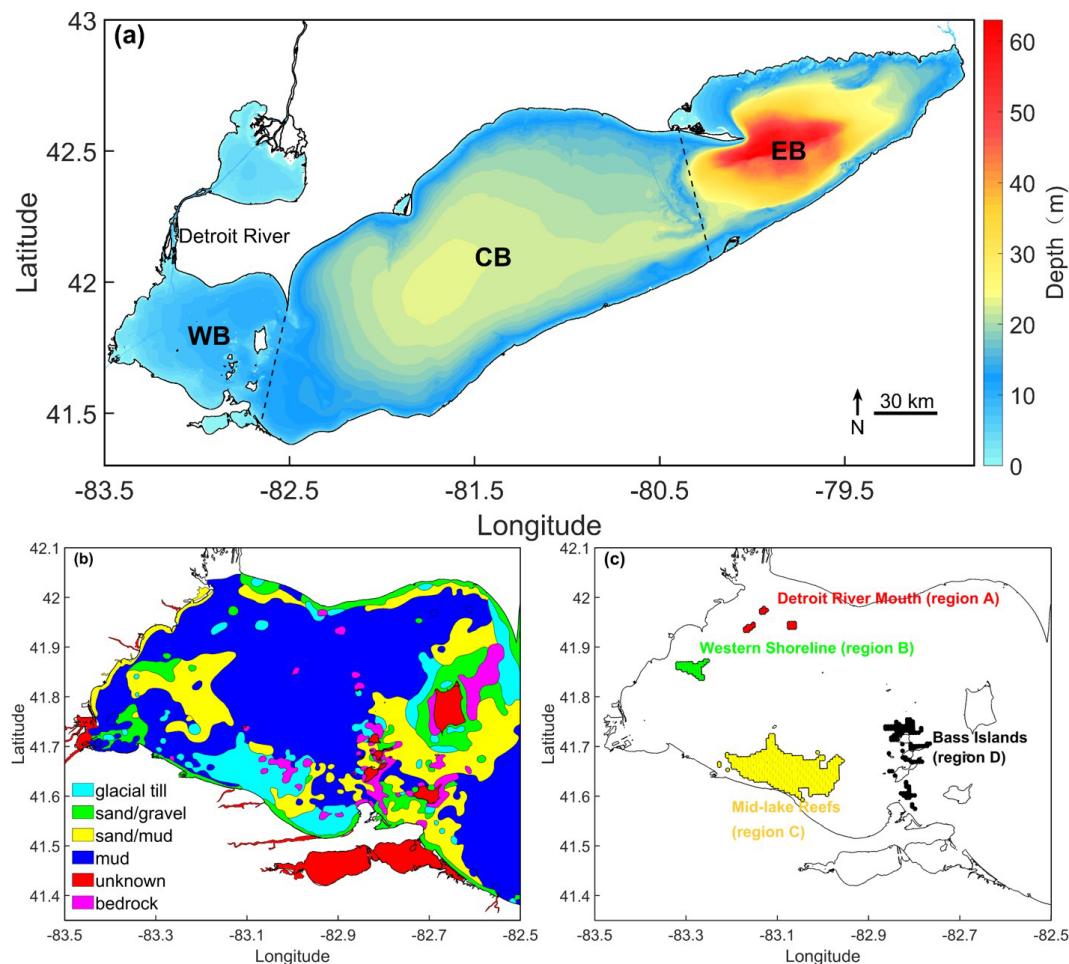


Figure 1. (a) Map showing the three basins of Lake Erie: western basin (WB), central basin (CB), and eastern basin (EB). The lake bathymetry was obtained from <https://www.ngdc.noaa.gov/mgg/greatlakes/erie.html> (last access: 22 January 2026). (b) Substrate distributions in the western basin from side-scan sonar transects (Haltuch et al., 2000). (c) Color coded Lake Whitefish (*Coregonus clupeaformis*) hatching locations (Detroit River Mouth in red, region A; Western Shoreline in green, region B; Mid-Lake Reefs in yellow, region C; Bass Islands in black, region D), where larval particles were released at the centre of each 500 m × 500 m AEM3D grid (black cross-hatching).

resolution was finest (0.5 m) through the surface layer, metalimnion, and near the bottom of the central basin, and coarser (5 m) in the deep hypolimnion of the eastern basin.

The model was forced with surface meteorological inputs (wind speed and direction, air temperature, relative humidity and long- and short-wave solar radiation) from four regional weather stations. Hydrological boundary conditions consisted of five major inflows (Detroit, Maumee, Sandusky, Cuyahoga and Grand rivers) and the Niagara River outflow.

AEM3D and its non-parallel predecessor, the Estuary and Lake Computer Model (ELCOM), have been extensively applied to Lake Erie. Previous applications include: backtracking Lake Whitefish larvae to determine hatching locations (Shi et al., 2024); hindcasting thermal structure (León et al., 2005); simulating internal wave dynamics (Valipour et al., 2015); surface wave and sediment transport modelling (Lin et al., 2021); simulating nutrient and chlorophyll-a distribu-

tions (Leon et al., 2011); modelling seasonal phytoplankton succession (Wang et al., 2024); forecasting storm surge and upwelling/downwelling events (Lin et al., 2022). These prior validation studies support the suitability of AEM3D for simulating larval dispersal processes in Lake Erie. Our model calibration and validation are published in the supplemental material of Shi et al. (2024). Briefly, in the western basin the mean modeled root-mean-square error (RMSE) was 0.99 °C during calibration (2018) and < 2 °C during validation (2017–2019). These are consistent with a previous ELCOM calibration RSME of 0.9 °C in the western basin for 2008 (Liu et al., 2014). The model reproduced the timing and magnitude of the current fluctuations ($\sim 10\text{--}20\text{ cm s}^{-1}$) driven by the 14 h surface seiche (Boegman et al., 2001) with modeled RMSE of 2.9–3.2 cm s^{-1} . These are comparable to published applications of ELCOM, giving RMSE of 4.3–4.4 cm s^{-1} in the central basin of Lake Erie (Lin et al.,

2021). Model error in predicting mean current fluctuations was < 10 % of measured values.

3.3 The Lagrangian particle tracking model

We used a Matlab®-based Lagrangian particle tracking model (SWIM-v2.0) to study SR. Earlier versions of SWIM have been applied both forward-in-time to simulate silver eel (*Anguilla rostrata* and *A. anguilla*) migration (Béguer-Pon et al., 2016) and backward-in-time to identify Lake Whitefish hatching locations (Shi et al., 2024). Biological behaviours, such as diel vertical migration and active swimming, were not included consistent with previous Lake Whitefish studies (Di Stefano et al., 2022; Rowe et al., 2022; Suca et al., 2022).

A horizontal turbulent diffusivity $K_h = 0.1 \text{ m}^2 \text{ s}^{-1}$ (Okubo, 1971) and timestep $dt_p = 600 \text{ s}$ were used for both forward- and backward-tracking simulations. Larval particles were released at 3 m depth and removed upon intersecting the lake boundary (Shi et al., 2024). A uniform release depth of 3 m was used to maintain consistency between forward and backward tracking. This depth approximates the mixed-layer habitat typically occupied by early-stage Lake Whitefish larvae.

Forward-tracking particle releases occurred daily at 12:00 LT (noon) between 21 March and 8 May 2018, across four regions (A–D; Table A1). Releases were repeated at 7 intervals (every 8 d) and were tracked for 12 d, corresponding to the pelagic larval duration of Lake Whitefish (Shi et al., 2024). Backward-tracking releases occurred daily at 12:00 LT (noon) between 2 April and 20 May 2018, and particles were tracked for 12 d. Preliminary back-tracking tests from regions A and B showed negligible settlement in any of the four regions (A, B, C and D). To avoid unnecessary computational expense, we thus restricted backward releases to Regions C and D, which represent the principal settlement regions in Shi et al. (2024).

To examine how the number of released particles influences SR, we conducted simulations with three release magnitudes per region (Tables 1 and 2). Release regions were subdivided into $500 \text{ m} \times 500 \text{ m}$ AEM3D grid cells, and particles were released at cell centers. Since the number of cells varies by region (e.g., region A: 24 cells; region B: 50 cells, Table A1), we scaled particles per cell to keep the total releases consistent within each scenario. In region A, we released 100, 500, 900 particles per cell, giving 16 800, 84 000 and 151 200 particles per simulation (e.g., $100 \text{ particles cell}^{-1} \text{ time}^{-1} \times 24 \text{ cells} \times 7 \text{ times} = 16 800 \text{ particles}$). In region B, we released 50, 200, 400 particles per cell, giving 17 500, 70 000 and 140 000 particles per simulation (e.g., $50 \text{ particles cell}^{-1} \text{ time}^{-1} \times 50 \text{ cells} \times 7 \text{ times} = 17 500 \text{ particles}$). Regions C and D were also partitioned in the same manner for consistency.

3.4 Nomenclature and data analysis

To clarify the variables used in our analysis, here we define the notation, and summarize all variable definitions in Table A2. In the forward-tracking simulations, N_i particles were released from each location i , with F_{ij} being the number that arrived at each destination j . For example, if N_A particles were released from region A, then F_{AC} particles settled in region C. In backward-tracking simulations, M_i particles were released from each location i , with B_{ij} being the number that traced back to location j . For example, if M_C particles were released from region C, then B_{CA} particles backtracked to region A, corresponding to recruits at C that originated from A.

The corresponding dispersal and recruitment (settlement) rates are:

$$D_{ij} = \frac{F_{ij}}{N_i}, \quad (13)$$

$$R_{ij} = \frac{B_{ij}}{M_i}. \quad (14)$$

From Sect. 2.3, estimating the true larval production at location j requires information from both forward-tracking dispersal rates and backward-tracking recruitment rates, along with the number of real recruits. In general, this is given by:

$$\begin{aligned} N'_j &= \frac{N'_{ji}}{D_{ji}} = M'_i \cdot \frac{R_{ij}}{D_{ji}} \cdot \frac{1}{\sum_{k=1}^n R_{ik}} \\ &= N_j \frac{M'_i}{M_i} \cdot \frac{B_{ij}}{F_{ji}} \cdot \frac{1}{\sum_{k=1}^n R_{ik}}, \end{aligned} \quad (15)$$

where forward- and backward-tracking quantities link larval release, settlement and recruitment.

Because only regions C and D were used as backward-tracking release locations in this study, larval production at location j is expressed as N'_{j-C} or N'_{j-D} :

$$N'_{j-C} = N_j \frac{M'_C}{M_C} \cdot \frac{B_{Cj}}{F_{jC}} \cdot \frac{1}{\sum_{k=1}^n R_{Ck}}, \quad (16)$$

$$N'_{j-D} = N_j \frac{M'_D}{M_D} \cdot \frac{B_{Dj}}{F_{jD}} \cdot \frac{1}{\sum_{k=1}^n R_{Dk}}. \quad (17)$$

For example, N'_A is the number of larvae produced from region A. To estimate N'_A , requires both backward and forward tracking simulations, and the number of recruits to a given location. Here, we only backtracked particles from region C and D; therefore, N'_A can be computed using the recruits at region C or D (M'_C or M'_D) and is denoted as N'_{A-C} or N'_{A-D} , respectively. Dividing N'_{A-C} by the number of AEM3D grid cells in region A (or the total area of the $500 \text{ m} \times 500 \text{ m}$ grid cells) gives the density d'_{A-C} .

4 Results

4.1 Larval distribution patterns

Most particles backtracked to regions west of their release locations, with relatively few tracing eastward (Fig. 2), consistent with the predominant west-to-east circulation in western Lake Erie (Beletsky et al., 2013). The exception was region A, where local circulation patterns were more variable. When particles were released from settlement region C and tracked backward for 12 d, they predominantly traced to the southern and western franks of the western basin (golden circles in Fig. 2c). These backtracking results agree with the modeled Lake Whitefish hatching locations reported by Shi et al. (2024), where larvae were sampled close to region C.

Forward-tracking simulations showed the opposite pattern: particles released from each of the four regions were generally transported eastward over the 12 d tracking period (dark green circles in Fig. 2), again matching the prevailing flow direction in the lake. An exception occurs Fig. 2b, where some particles move southeast or northeast of the release region. Although counter-intuitive, relative to the dominant flow, such deviations are reasonable given the complex bathymetry and variable wind-field in the western basin.

It is known that dispersal estimates are sensitive to the number of particles released (Béguer-Pon et al., 2016); therefore, we conducted a sensitivity analysis. When too few particles were released, dispersal patterns fluctuated markedly (Fig. 3). However, the results show that the estimated dispersal rate, local retention (Eq. 2) and self-recruitment all become invariant when more than 40 000 particles were released in both forward- and backward-tracking simulations (Fig. 3). This threshold was used to ensure statistically robust results.

4.2 Self-recruitment, local retention and relative local retention

In the forward-tracking simulations, the number of settled particles (F_{ij}) scaled directly with the number of particles released (N_i), whereas the dispersal rate (D_{ij}) remained independent of N_i (Table 1 and Fig. 3). Accordingly, both LR and RLR also showed negligible sensitivity to N_i . For example, LR at region C equalled $D_{CC} \approx 0.13$ and LR at region D equalled $D_{DD} \approx 0.024$. From Eq. (3), RLR was 0 at regions A and B, ~ 0.76 at region C, and ~ 0.99 at region D.

In contrast, SR estimates derived from forward-tracking varied markedly with N_i , reflecting the dependence of Eq. (3) on larval production at all source locations. Across the 81 forward-tracking configurations, SR at region C ranged from 0.22–0.95. For example, when releases were heavily weighted toward regions A and B (151 200 particles from A, 140 000 from B, 19 460 from C and 148 400 from D), $SR_C = 2537/(4668 + 4195 + 2537 + 37) = 0.22$ whereas when releases were weighted toward region C (16 800 particles from

A, 17 500 from B, 155 680 from C, and 18 550 from D), $SR_C = 20758/(527 + 501 + 20758 + 2) = 0.95$. These outcomes demonstrate that SR from forward tracking is not uniquely defined unless the actual larval production N'_i is known and applied. Due to D_{CC} being much larger than D_{AC} , D_{BC} and D_{DC} , increasing N_C disproportionately increases the numerator of SR_C , causing SR to approach 1. Conversely, reducing N_C or increasing releases from other regions forces SR toward 0. Therefore, SR computed from forward-tracking can span the range $[0, 1]$ solely as a function of the assumed release numbers and cannot be interpreted biologically unless the actual number of larvae produced at each location is known and is released (i.e., $N_i = N'_i$).

If all regions were assumed to release the same number of particles (a common but biologically unfounded modelling assumption), SR becomes only a function of D_{ij} . Under this scenario, $SR_C = D_{CC}/D_{AC} + D_{BC} + D_{CC} + D_{DC} = 0.68$ and $SR_D = D_{DD}/D_{AD} + D_{BD} + D_{CD} + D_{DD} = 0.32$. These values differ substantially from those obtained using backtracking.

In the backtracking simulations, the number of particles traced to each location B_{ij} varied with the number released M_i , but the settlement rate R_{ij} and SR were consistent with M_i (Table 2 and Fig. 3). This stability shows that SR calculated from backtracking is independent of M_i , as indicated in Eq. (7).

Comparing backtracking and forward-tracking SR with equal release assumptions (i.e., N_i was the same for all sources) reveals the bias introduced by unknown or mis-specified larval production. For example, $SR_C = 0.97$ and $SR_D = 0.18$ from backtracking, but only 0.68 and 0.32, respectively, from forward tracking. The discrepancies arise because assuming equal N_i artificially inflates the contribution of weak sources (e.g., A and B) and suppresses the contribution of strong sources (i.e., C and D), thereby distorting the SR denominator.

4.3 Larval production

Using forward- and backward-tracking results, together with real recruits at regions C and D (M'_C and M'_D), we estimated larval production at regions A, B, C and D based on Eqs. (10) and (16) (Table 3). When recruits at region C were used as the reference, estimated larval production at each region was largely insensitive to the number of particles released in the forward and backward tracking simulations, with the exception of region D, for which the dispersal rate from D to C (D_{DC}) was close to zero. When recruits at region D were used as the reference, estimated larval production at regions C and D also remained consistent, whereas estimates for regions A and B were zero due to the absence of recruitment to region D. These suggest that Eq. (10) provides a robust approach for estimating larval production.

The larval production density at region C computed from M'_C (d'_{C_C}) was approximately $0.013M'_C$, whereas that computed from M'_D (d'_{C_D}) was approximately $0.034M'_D$. Be-

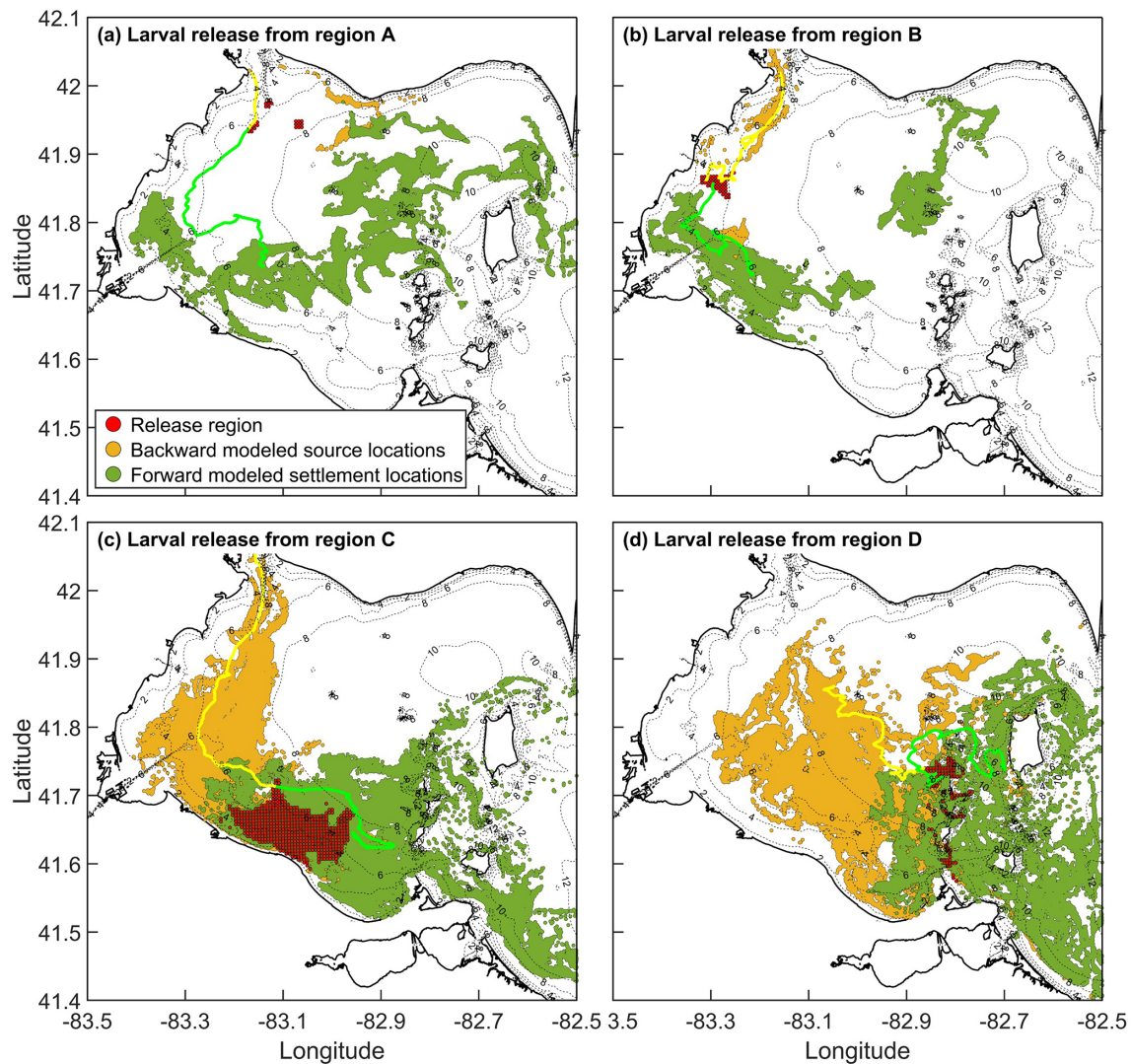


Figure 2. Modelled source location distributions (gold circles; backtracking) and settlement locations (dark green circle; forward tracking) for particles released from (a) region A, (b) region B, (c) region C, and (d) region D. Particles were tracked for 12 d in both forward and backward simulations. In backward tracking, the release region represents settlement location; in forward tracking, it represents the source location. The fraction of particles returning to their release region indicates imaginary self-recruitment (backtracking) or local retention (forward tracking). The ratio of particles returning to their release region, to the total settling in all four regions, indicates self-recruitment (backtracking). Yellow/green line shows example backward/forward trajectories.

cause d'_{C-C} should equal d'_{C-D} , this yields that $0.013M'_C = 0.034M'_D$, indicating that $M'_C \approx 2.6M'_D$. Similar, the larval production density at Region D computed from M'_D (d'_{D-C}) was around $0.031M'_C$, while that computed from M'_D (d'_{D-D}) was around $0.071M'_D$. Equating d'_{D-C} and d'_{D-D} yields $M'_C \approx 2.3M'_D$. The close agreement between these two independent estimates further supports the validity of Eq. (10) for estimating larval production. Notably, estimated larval production density followed a descending order from region D to region B to region C and finally to region A.

5 Discussion

5.1 Estimates of self-recruitment

We have shown, using both theoretical arguments and numerical simulations, that self-recruitment (SR) cannot be unambiguously computed using forward Lagrangian particle tracking models. Forward-tracking studies have implemented a wide range of particle release strategies, including: (1) releasing a random number of larvae (Sato et al., 2023), (2) releasing an equal number of larvae (Dubois et al., 2016; Faillettaz et al., 2018), (3) scaling releases to habitat area (D'agostini et al., 2015), and (4) scaling releases to esti-

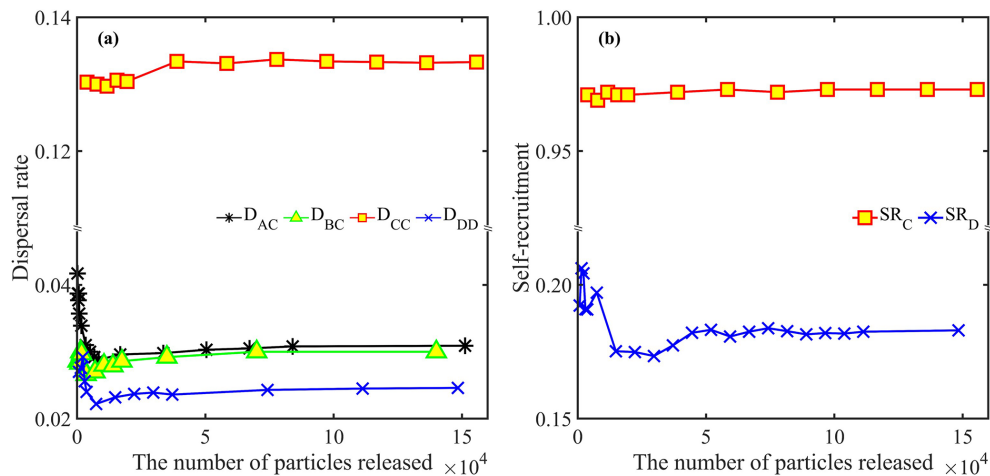


Figure 3. A sensitivity of results to the number of particles released. **(a)** In forward simulations, dispersal rates D_{ij} converge and become statistically stable when more than 40 000 particles are released. **(b)** In backward simulations, self-recruitment SR_i becomes stable above the same threshold.

Table 1. Number of particles released from each region (N_i), the number settling in region (F_{ij}), the resulting dispersal rates (D_{ij}), and the corresponding local retention (LR_i), and relative local retention (RLR_i) from forward-tracking simulations. The four regions are the Detroit River Mouth (A), Western Shoreline (B), Midlake Reefs (C), and Bass Islands (D).

N_A	F_{AA}	F_{AB}	F_{AC}	F_{AD}	D_{AA}	D_{AB}	D_{AC}	D_{AD}	LR_A	RLR_A
16 800	0	0	527	140	0	0	0.0313	0.0083	0	0
84 000	0	0	2587	665	0	0	0.0308	0.0079	0	0
151 200	0	0	4668	1189	0	0	0.0309	0.0079	0	0
N_B	F_{BA}	F_{BB}	F_{BC}	F_{BD}	D_{BA}	D_{BB}	D_{BC}	D_{BD}	LR_B	RLR_B
17 500	0	0	501	0	0	0	0.029	0	0	0
70 000	0	0	2097	0	0	0	0.030	0	0	0
140 000	0	0	4195	0	0	0	0.030	0	0	0
N_C	F_{CA}	F_{CB}	F_{CC}	F_{CD}	D_{CA}	D_{CB}	D_{CC}	D_{CD}	LR_C	RLR_C
19 460	0	0	2537	843	0	0	0.130	0.0433	0.130	0.751
77 840	0	0	10 440	3327	0	0	0.134	0.0427	0.134	0.758
155 680	0	0	20 758	6739	0	0	0.133	0.0433	0.133	0.755
N_D	F_{DA}	F_{DB}	F_{DC}	F_{DD}	D_{DA}	D_{DB}	D_{DC}	D_{DD}	LR_D	RLR_D
18 550	0	0	2	435	0	0	0.0001	0.0235	0.0235	0.995
74 200	0	0	19	1807	0	0	0.0002	0.0242	0.0242	0.990
148 400	0	0	37	3648	0	0	0.0002	0.0246	0.0246	0.990

mated larval production (Nolasco et al., 2022). Each strategy implicitly embeds an assumption about the spatial distribution of larval production, which is rarely known. As a result, SR estimates from forward tracking are sensitive to these assumptions.

For example, Sato et al. (2023) assumed equal larval production across 84 source locations in the Verde Island Passage and released 900 particles per location. After aggregating locations into three regions, they released 4500, 40 500 and 30 600 particles from Puerto Galera, the eastern region and the western region, respectively, giving $SR =$

$17.9/(17.9 + 4.4 + 71.9) = 0.19$ at Puerto Galera. However, if Puerto Galera was sub-divided into more spatial units, more particles would be released there, and the number of local settlers (17.9) would increase proportionally, altering SR. This example illustrates how SR can be influenced by arbitrary spatial partitioning.

Other studies apply different assumptions: D’agostini et al. (2015) proposed that larval production is proportional to habitat area; Saint-Amand et al. (2023) assumed a fixed larval density ($500 \text{ larvae km}^{-2}$); and Nolasco et al. (2022) scaled larval production to adult abundance and spawning in-

Table 2. Number of particles released from each region (M_i) in the backtracking simulations, the number traced back to each region (B_{ij}), the resulting settlement rates (R_{ij}), and corresponding self-recruitment values (SR_i). The regions are the Detroit River Mouth (region A), Western Shoreline (B), Midlake Reefs (C), and Bass Islands (D).

M_C	B_{CA}	B_{CB}	B_{CC}	B_{CD}	R_{CA}	R_{CB}	R_{CC}	R_{CD}	SR_C
19 460	17	94	3778	2	8.7×10^{-4}	0.0048	0.194	0.0001	0.971
77 840	75	337	15 300	8	9.6×10^{-4}	0.0043	0.196	0.0001	0.973
155 680	137	685	30 659	22	8.8×10^{-4}	0.0044	0.197	0.0001	0.973
M_D	B_{DA}	B_{DB}	B_{DC}	B_{DD}	R_{DA}	R_{DB}	R_{DC}	R_{DD}	SR_D
18 550	0	0	1911	407	0	0	0.103	0.022	0.176
74 200	0	0	7628	1718	0	0	0.102	0.023	0.184
148 400	1	0	15 302	3430	6.7×10^{-6}	0	0.103	0.023	0.183

Table 3. Estimated larval production for each region (N'_{i_C} and N'_{i_D}) based on the backtracking simulations initiated from regions C and D, respectively, together with the corresponding larval densities (number per AEM3D grid cell; d'_{i_C} and d'_{i_D}). The regions are the Detroit River Mouth (A), Western Shoreline (B), Midlake Reefs (C), and Bass Islands (D).

M_C	N_A	N'_{A_C}	d'_{A_C}	N_B	N'_{B_C}	d'_{B_C}	N_C	N'_{C_C}	d'_{C_C}	N_D	N'_{D_C}	d'_{D_C}
19 460	16 800	$0.139M'_C$	$0.0058M'_C$	17 500	$0.845M'_C$	$0.017M'_C$	19 460	$7.454M'_C$	$0.013M'_C$	18 550	$4.772M'_C$	$0.045M'_C$
	84 000	$0.142M'_C$	$0.0059M'_C$	70 000	$0.807M'_C$	$0.016M'_C$	77 840	$7.246M'_C$	$0.013M'_C$	74 200	$2.009M'_C$	$0.019M'_C$
	151 200	$0.142M'_C$	$0.0059M'_C$	140 000	$0.807M'_C$	$0.016M'_C$	155 680	$7.289M'_C$	$0.013M'_C$	148 400	$2.063M'_C$	$0.020M'_C$
77 840	16 800	$0.153M'_C$	$0.0064M'_C$	17 500	$0.751M'_C$	$0.015M'_C$	19 460	$7.488M'_C$	$0.014M'_C$	18 550	$4.734M'_C$	$0.045M'_C$
	84 000	$0.155M'_C$	$0.0065M'_C$	70 000	$0.718M'_C$	$0.014M'_C$	77 840	$7.278M'_C$	$0.013M'_C$	74 200	$1.993M'_C$	$0.019M'_C$
	151 200	$0.155M'_C$	$0.0065M'_C$	140 000	$0.718M'_C$	$0.014M'_C$	155 680	$7.321M'_C$	$0.013M'_C$	148 400	$2.047M'_C$	$0.019M'_C$
155 680	16 800	$0.139M'_C$	$0.0058M'_C$	17 500	$0.759M'_C$	$0.015M'_C$	19 460	$7.464M'_C$	$0.013M'_C$	18 550	$6.476M'_C$	$0.061M'_C$
	84 000	$0.141M'_C$	$0.0059M'_C$	70 000	$0.726M'_C$	$0.015M'_C$	77 840	$7.255M'_C$	$0.013M'_C$	74 200	$2.727M'_C$	$0.026M'_C$
	151 200	$0.141M'_C$	$0.0059M'_C$	140 000	$0.726M'_C$	$0.015M'_C$	155 680	$7.298M'_C$	$0.013M'_C$	148 400	$2.801M'_C$	$0.026M'_C$
M_D	N_A	N'_{A_D}	d'_{A_D}	N_B	N'_{B_D}	d'_{B_D}	N_C	N'_{C_D}	d'_{C_D}	N_D	N'_{D_D}	d'_{D_D}
18 550	16 800	0	0	17 500	0	0	19 460	$19.025M'_D$	$0.034M'_D$	18 550	$7.485M'_D$	$0.071M'_D$
	84 000	0	0	70 000	0	0	77 840	$19.282M'_D$	$0.035M'_D$	74 200	$7.236M'_D$	$0.068M'_D$
	151 200	0	0	140 000	0	0	155 680	$19.039M'_D$	$0.034M'_D$	148 400	$7.140M'_D$	$0.067M'_D$
74 200	16 800	0	0	17 500	0	0	19 460	$18.985M'_D$	$0.034M'_D$	18 550	$7.899M'_D$	$0.075M'_D$
	84 000	0	0	70 000	0	0	77 840	$19.242M'_D$	$0.035M'_D$	74 200	$7.606M'_D$	$0.072M'_D$
	151 200	0	0	140 000	0	0	155 680	$18.999M'_D$	$0.034M'_D$	148 400	$7.535M'_D$	$0.071M'_D$
148 400	16 800	$0.0064M'_D$	$2.7 \times 10^{-4}M'_C$	17 500	0	0	19 460	$18.890M'_D$	$0.034M'_D$	18 550	$7.822M'_D$	$0.074M'_D$
	84 000	$0.0068M'_D$	$2.8 \times 10^{-4}M'_C$	70 000	0	0	77 840	$19.146M'_D$	$0.034M'_D$	74 200	$7.532M'_D$	$0.071M'_D$
	151 200	$0.0068M'_D$	$2.7 \times 10^{-4}M'_C$	140 000	0	0	155 680	$18.904M'_D$	$0.034M'_D$	148 400	$7.462M'_D$	$0.070M'_D$

tensity. In these examples, SR varied with the assumed production pattern and not solely with dispersal dynamics. Since it is often unknown whether larval production is uniform, area-proportional, or scales with adult biomass, forward-tracking estimates of SR inherit this uncertainty. Moreover, unmodelled or unknown source locations may contribute additional settlers, further biasing SR estimates.

In contrast, backward-in-time Lagrangian particle-tracking provides a more direct and robust framework for estimating SR. Backtracking releases particles from locations where juveniles are sampled and traces them backward to their origins. This method is widely applied to identify spawning or hatching sites (Christensen et al., 2007;

Thygesen, 2011; Bauer et al., 2014; Gargano et al., 2022; Rowe et al., 2022; Chaput et al., 2023), but surprisingly few studies have used it to explicitly estimate SR. One exception is Torrado et al. (2021) who found high SR because most settlers originated from the same oceanographic region in which they settled.

A key advantage in backtracking is that the denominator of SR (the total number of settlers at location j) is simply the number of particles released from j and is allowed to be any convenient value (M_j). The key insight is that the resulting value of SR is independent of the absolute number of particles released (i.e., M_j in Eq. 7); our numerical results confirm this. Consequently, backtracking avoids the need to

specify larval production at all source locations or to include every source explicitly. Rather, SR is estimated directly as the proportion of sampled recruits whose trajectories lead back to local origins, consistent with parentage-analysis studies (D'Aloia et al., 2013; Almany et al., 2017).

Although SR estimated from backtracking does not depend on M_j , it does depend on the physical environment. Variability in hydrodynamic conditions, spawning time and environmental forcing all impact SR. In our simulations, SR varied substantially between region C (0.97) and D (0.18), despite similar larval durations. Comparable variability has been documented in marine systems; for example, Paris et al. (2005) found that mutton snapper (*Lutjanus analis*) spawning aggregations in Cuba exported larvae regionally in patterns that produced distinct SR values. Even species spawning at the same location but at different times, may experience contrasting SR, as shifts in currents between lunar phases can alter dispersal pathways Paris et al. (2005).

Connectivity probability matrices ($\mathbf{P}_T$), are widely used to summarise larval exchange among regions, and their diagonal elements are often interpreted as SR. However, because $P_{T,ij} = S_{T,ij} / \sum_i S_{T,ij}$, the denominator depends on larval production from all contributing source locations. Any change in assumed production therefore alters SR and all matrix elements. In contrast, normalising the transition matrix $\mathbf{S}_T$ by the total number of larvae released or surviving ($\sum_j S_{T,ij}$) yields production-independent connectivity matrices whose diagonal elements reflect LR or RLR. This distinction reinforces our central result: forward-tracking estimates of SR inherently inherit assumptions about larval production, whereas SR estimated from backtracking does not.

5.2 Estimates of local retention and relative local retention

As indicated earlier, relative local retention (RLR) is typically more challenging to evaluate empirically than self-recruitment (SR) (Lett et al., 2015). SR can be estimated by sampling juveniles at a settlement site and identifying the proportion that originated locally using genetic parentage or assignment methods. In contrast, estimating RLR requires quantifying all larvae produced at a source location that ultimately survive and settle anywhere in the system. This includes individuals dispersed to distant or unsampled habitats, meaning that incomplete sampling directly biases the denominator of RLR.

Local retention (LR) is even more challenging to measure because it requires the total number of larvae produced at a location (N'_i), including those lost to mortality or transported to unobserved locations. For example, Almany et al. (2017) assigned juveniles of *Amphiprion percula* and *Chaetodon vagabundus* to parents collected at eight sites Papua New Guinea, enabling estimation of the number of juveniles produced at each source location. However, it remained unknown whether larvae were transported exclusively among

these eight sites, or whether additional sinks existed outside the sampling domain. Moreover, the total number produced at each site, including the fraction that died before settlement could not be determined.

These empirical limitations explain why forward-tracking studies commonly adopt simplified particle release strategies. Such approaches circumvent the need to estimate N'_i , but they also mean that LR and RLR computed from forward models depend on assumptions about larval production.

5.3 Estimates of larval production and recruitment

Knowing the number of recruits at one settlement location (M'_i) or the larval production at one source location (N'_i) enables estimation of recruitment and production at all other locations from Eqs. (11) and (12). Equation (10) provides an approach to infer larval production at a source location (N'_i) based on the number of recruits at a settlement site (M'_j), together with the recruitment rate (R_{ji}) and the dispersal rate (D_{ij}). In principle, estimates of N'_i derived from different settlement locations should be consistent. For example, larval production at region C inferred using recruits from region C ($N'_{C,C}$) should equal that inferred using recruits from region D ($N'_{C,D}$). However, because the values of M'_C and M'_D are unavailable, this consistency cannot be directly verified. Nevertheless, it constrains their relative magnitudes, indicating that $M'_C \approx 2.6M'_D$. Similarly, larval production at region D inferred using recruits from regions C and D yields a comparable relationship, $M'_C \approx 2.3M'_D$. The close agreement between these two independent estimates provides partial support for the validity of Eq. (10) in estimating larval production.

Importantly, the framework also enables inference on relative larval production among locations. This information is of directly relevant to fisheries management, as it helps identify spawning and hatching grounds that contribute disproportionately to regional recruitment. By quantifying the relative contributions of different source locations, the approach provides a scientific basis for prioritizing habitat protection, restoration efforts, and spatially targeted management measures aimed at sustaining fish populations.

5.4 Implications and limitations of this research

Using forward and backward tracking models in combination provides a powerful framework for estimating not only SR and LR, but also larval production at each source location and the number of recruits at each settlement site. When used independently, backtracking models can only estimate SR (or imaginary SR) and settlement rates, while forward tracking models can only estimate LR (or relative LR) and dispersal rates. By integrating these two approaches, dispersal and settlement rates become linked, enabling the derivation of larval production at source locations and the number of recruits reaching all connected settlement sites once recruit-

ment at any one location is known. This joint framework, therefore, supports the estimation of total larval production and total recruitment within a region, offering improved insight into population connectivity and direct utility for population assessment and fisheries management. Importantly, it enables quantification of the relative contributions of different spawning and hatching grounds to regional recruitment, helping to identify key source areas that disproportionately sustain population persistence. Such information is directly applicable to fisheries stock assessment, spatial management, and marine spatial planning, including the prioritization of critical habitats for protection or restoration.

Despite these advantages, several limitations remain in applying a backtracking model approach to evaluate SR. For example, diffusion cannot be reversed in time; backward diffusion is ill-posed and must be approximated using convergent schemes (Thygesen, 2011). Our simulations also used passive particles moving two dimensions and did not incorporate larval behaviours such as vertical migration or active swimming. Numerous studies have shown that incorporating larval behaviour in forward-tracking models can substantially alter dispersal outcomes (Berenshtein et al., 2018; Burgess et al., 2021; Faillettaz et al., 2018; James et al., 2019; Leis, 2020; Paris et al., 2005). Finally, our boundary condition, of removing particles when they intersect the model boundary, may introduce deviations between simulated and true dispersal pathways, which impacts SR estimates. Comparisons with empirical datasets are required to validate the backtracking-derived values of SR.

6 Conclusions

Our results demonstrate that self-recruitment (SR) fundamentally depends on the true larval production at all source locations capable of supplying recruits to the settlement site. Consequently, SR cannot be computed unambiguously in forward-tracking models unless all potential source locations and their respective larval production are known and used. This limitation is compounded by the different particle-release strategies employed in previous studies, which produce different SR estimates solely because they assume different production patterns.

In contrast, empirical parentage analyses do not require quantifying larval production at every source location. SR is obtained directly as the proportion of sampled juveniles that originate locally. Backward-tracking models offer the same advantage: by releasing particles from a settlement site and tracing them backwards to their origins, SR is estimated as the fraction of locally sourced recruits. Our results confirm theoretically and numerically that SR computed from backtracking is independent of the number of particles released at the settlement locations, enabling the use of any release number and bypassing the need to identify all contributing source locations.

We further propose an approach to estimate larval production at each source location by combining information from forward- and backward-tracking models. When used separately, backtracking models provide only SR (or imaginary SR) and settlement rates, while forward-tracking models provide only LR (or relative LR) and dispersal rates. When used jointly they allow estimation of SR, LR, dispersal, settlement, larval production at source sites, and recruitment at all settlement sites, thereby offering a comprehensive framework for quantifying population connectivity.

Although demonstrated here for Lake Whitefish in Lake Erie, the theoretical and modelling framework apply broadly to both freshwater and marine species with pelagic larval phases. The ability to more accurately compute these metrics will improve understanding of population connectivity and support more informed management of spatially structured fish populations.

Appendix A

Table A1. Release regions, times and durations for forward and backward tracking simulations.

	Release regions	Release times	Tracking periods (day)	Number of cells
Forward	Detroit River Mouth (Region A)	21 March–8 May	12	24
	Western Shoreline (Region B)	21 March–8 May	12	50
	Midlake Reefs (Region C)	21 March–8 May	12	556
	Bass Islands (Region D)	21 March–8 May	12	106
Backward	Midlake Reefs (Region C)	2 April–20 May	12	556
	Bass Islands (Region D)	2 April–20 May	12	106

Table A2. The description of the notations in this research.

Notation	Description
B_{ij}	Number of particles recruited to site j that were released from site i in backtracking
d'_i	Number of larvae per cell ($500\text{ m} \times 500\text{ m}$) produced from site i
d'_{i-j}	Number of larvae per cell ($500\text{ m} \times 500\text{ m}$) produced from site i estimated based on the recruits to site j
D_{ij}	Proportion of larvae released from site i that recruit into the juvenile population at site j
F_{ij}	Number of particles recruited to site j that were released from site i in forward tracking
RLR_i	Ratio of local larval recruitment at site i to the number of larvae released locally, that settled in suitable nursery sites
M_i	Number of particles released from site i in the backtracking simulations
M'_i	Number of larval recruits to site i
N_i	Number of particles released from site i in forward tracking simulations
N'_i	Number of larvae produced at site i
N'_{i-j}	Number of larvae produced from site i estimated from the recruits to site j
N'_{ji}	Number of recruits at site i that were released from site j
R_{ij}	Proportion of particles released from patch i that settled in patch j in the backtracking simulations
SR_i	Ratio of local larval recruitment at site i to all recruitment at site i
LR_i	Ratio of local larval recruitment to site i to local larvae released
ISR_i	Ratio of particles that settled at site i to all the particles released from site i in backtracking simulations

Code and data availability. The AEM3D executable was used as a black-box hydrodynamic transport code. The AEM3D source code was not modified in this application but is available with permission from HydroNumerics. The model setup for AEM3D are available at <https://doi.org/10.5281/zenodo.14749408> (Boegman, 2025). The forward and backward particle tracking models were performed in Matlab. Their code and simulated date, and the velocity output from AEM3D are all available at <https://doi.org/10.5281/zenodo.14789098> (Shi, 2025). The SWIM-v2.0 model was implemented in MATLAB R2024a using standard functions and without reliance on specialized toolboxes. While we have not explicitly tested the code in GNU Octave, we expect that it should run in Octave without or with only minor modifications. Users are welcome to adapt the code for use with Octave or other compatible environments.

Author contributions. WS conceived the main study design, developed the theory, performed the simulations and analyses. WS wrote the first draft of manuscript and LB and JDA revised the draft significantly. LB and SLS co-supervised the research and provided resources. JDA, LB, SLS and YMZ acquired research funding. All authors contributed to the project conceptualization, and editing and revising the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. The authors bear the ultimate responsibility for providing appropriate place names. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Financial support. This work was supported by the Ontario Ministry of Natural Resources and Forestry, the Canada Ontario Agreement program and an NSERC Alliance Grant program to Josef D. Ackerman, Leon Boegman, and Shiliang Shan.

Review statement. This paper was edited by Andrew Yool and reviewed by Romain Chaput, Claire Paris, and three anonymous referees.

References

- Almany, G. R., Planes, S., Thorrold, S. R., Berumen, M. L., Bode, M., Saenz-Agudelo, P., Bonin, M. C., Frisch, A. J., Harrison, H. B., Messmer, V., Nanninga, G. B., Priest, M. A., Srinivasan, M., Sinclair-Taylor, T., Williamson, D. H., and Jones, G. P.: Larval fish dispersal in a coral-reef seascape, *Nat. Ecol. Evol.*, 1, 148, <https://doi.org/10.1038/s41559-017-0148-2017>.
- Amidon, Z. J., DeBruyne, R. L., Roseman, E. F., and Mayer, C. M.: Contemporary and Historic Dynamics of Lake Whitefish (*Coregonus clupeaformis*) Eggs, Larvae, and Juveniles Suggest Recruitment Bottleneck during First Growing Season, *Annales Zoologici Fennici*, 58, <https://doi.org/10.5735/086.058.0405>, 2021.
- Bauer, R. K., Gräwe, U., Stepputtis, D., Zimmermann, C., and Hammer, C.: Identifying the location and importance of spawning sites of Western Baltic herring using a particle backtracking model, *ICES Journal of Marine Science*, 71, 499–509, <https://doi.org/10.1093/icesjms/fst163>, 2014.
- Béguer-Pon, M., Shan, S., Thompson, K. R., Castonguay, M., Sheng, J., and Dodson, J. J.: Exploring the role of the physical marine environment in silver eel migrations using a biophysical particle tracking model, *ICES Journal of Marine Science*, 73, 57–74, <https://doi.org/10.1093/icesjms/fsv169>, 2016.
- Beletsky, D., Hawley, N., and Rao, Y. R.: Modeling summer circulation and thermal structure of Lake Erie, *J. Geophys. Res.-Oceans*, 118, 6238–6252, <https://doi.org/10.1002/2013jc008854>, 2013.
- Berenshtein, I., Paris, C. B., Gildor, H., Fredj, E., Amitai, Y., and Kiflawi, M.: Biophysical Simulations Support Schooling Behavior of Fish Larvae Throughout Ontogeny, *Frontiers in Marine Science*, 5, <https://doi.org/10.3389/fmars.2018.00254>, 2018.
- Boegman, L.: Lake Erie AEM3D-iWQ model setup and data, Zenodo [code and data set], <https://doi.org/10.5281/zenodo.14749408>, 2025.
- Boegman, L., Loewen, M. R., Hamblin, P. F., and Culver, D. A.: Application of a two-dimensional hydrodynamic reservoir model to Lake Erie, *Canadian Journal of Fisheries and Aquatic Sciences*, 58, 858–869, <https://doi.org/10.1139/cjfas-58-5-858>, 2001.
- Botsford, L. W., White, J. W., Coffroth, M. A., Paris, C. B., Planes, S., Shearer, T. L., Thorrold, S. R., and Jones, G. P.: Connectivity and resilience of coral reef metapopulations in marine protected areas: matching empirical efforts to predictive needs, *Coral Reefs*, 28, 327–337, <https://doi.org/10.1007/s00338-009-0466-z>, 2009.
- Burgess, S. C., Nickols, K. J., Griesemer, C. D., Barnett, L. A., Dedrick, A. G., Satterthwaite, E. V., Yamane, L., Morgan, S. G., White, J. W., and Botsford, L. W.: Beyond connectivity: how empirical methods can quantify population persistence to improve marine protected-area design, *Ecol. Appl.*, 24, 257–270, <https://doi.org/10.1890/13-0710.1>, 2014.
- Burgess, S. C., Bode, M., Leis, J. M., and Mason, L. B.: Individual variation in marine larval-fish swimming speed and the emergence of dispersal kernels, *Oikos*, 2022, <https://doi.org/10.1111/oik.08896>, 2021.
- Chaput, R., Sochala, P., Miron, P., Kourafalou, V. H., Iskandarani, M., and Kaplan, D. M.: Quantitative uncertainty estimation in biophysical models of fish larval connectivity in the Florida Keys, *ICES Journal of Marine Science*, 79, 609–632, <https://doi.org/10.1093/icesjms/fsac021>, 2022.
- Chaput, R., Quigley, C. N., Weppe, S. B., Jeffs, A. G., de Souza, J., and Gardner, J. P. A.: Identifying the source populations supplying a vital economic marine species for the New Zealand aquaculture industry, *Sci. Rep.-UK*, 13, 9344, <https://doi.org/10.1038/s41598-023-36224-y>, 2023.
- Christensen, A., Daewel, U., Jensen, H., Mosegaard, H., St. John, M., and Schrum, C.: Hydrodynamic backtracking of fish larvae by individual-based modelling, *Mar. Ecol. Prog. Ser.*, 347, 221–232, <https://doi.org/10.3354/meps06980>, 2007.
- Corrochano-Fraile, A., Adams, T. P., Aleynik, D., Bekaert, M., and Carboni, S.: Predictive biophysical models of bivalve larvae dispersal in Scotland, *Frontiers in Marine Science*, 9, <https://doi.org/10.3389/fmars.2022.985748>, 2022.
- Cowen, R. K. and Sponaugle, S.: Larval dispersal and marine population connectivity, *Ann. Rev. Mar. Sci.*, 1, 443–466, <https://doi.org/10.1146/annurev.marine.010908.163757>, 2009.
- D'Agostini, A., Gherardi, D. F., and Pezzi, L. P.: Connectivity of Marine Protected Areas and Its Relation with Total Kinetic Energy, *PLoS One*, 10, e0139601, <https://doi.org/10.1371/journal.pone.0139601>, 2015.
- D'Aloia, C. C., Bogdanowicz, S. M., Majoris, J. E., Harrison, R. G., and Buston, P. M.: Self-recruitment in a Caribbean reef fish: a method for approximating dispersal kernels accounting for seascape, *Mol. Ecol.*, 22, 2563–2572, <https://doi.org/10.1111/mec.12274>, 2013.
- D'Aloia, C. C., Daigle, R. M., Côté, I. M., Curtis, J. M. R., Guichard, F., and Fortin, M.-J.: A multiple-species framework for integrating movement processes across life stages into the design of marine protected areas, *Biol. Conserv.*, 216, 93–100, <https://doi.org/10.1016/j.biocon.2017.10.012>, 2017.
- D'Agostini, A., Gherardi, D. F. M., and Pezzi, L. P.: Connectivity of Marine Protected Areas and Its Relation with Total Kinetic Energy, *PLoS One*, 10, <https://doi.org/10.1371/journal.pone.0139601>, 2015.
- Di Stefano, M., Legrand, T., Di Franco, A., Nerini, D., and Rossi, V.: Insights into the spatio-temporal variability of spawning in a territorial coastal fish by combining observations, mod-

- elling and literature review, *Fisheries Oceanography*, 32, 70–90, <https://doi.org/10.1111/fog.12609>, 2022.
- Dubois, M., Rossi, V., Ser-Giacomi, E., Arnaud-Haond, S., López, C., and Hernández-García, E.: Linking basin-scale connectivity, oceanography and population dynamics for the conservation and management of marine ecosystems, *Global Ecol. Biogeogr.*, 25, 503–515, <https://doi.org/10.1111/geb.12431>, 2016.
- Faillietaz, R., Paris, C. B., and Irissou, J.-O.: Larval Fish Swimming Behavior Alters Dispersal Patterns From Marine Protected Areas in the North-Western Mediterranean Sea, *Frontiers in Marine Science*, 5, <https://doi.org/10.3389/fmars.2018.00097>, 2018.
- Gargano, F., Garofalo, G., Quattrocchi, F., and Fiorentino, F.: Where do recruits come from? Backward Lagrangian simulation for the deep water rose shrimps in the Central Mediterranean Sea, *Fisheries Oceanography*, 31, 369–383, <https://doi.org/10.1111/fog.12582>, 2022.
- Gurdek-Bas, R., Benthuyssen, J. A., Harrison, H. B., Zenger, K. R., and van Herwerden, L.: The El Nino Southern Oscillation drives multidirectional inter-reef larval connectivity in the Great Barrier Reef, *Sci. Rep.-UK*, 12, 21290, <https://doi.org/10.1038/s41598-022-25629-w>, 2022.
- Haltuch, M. A., Berkman, P. A., and Garton, D. W.: Geographic information system (GIS) analysis of ecosystem invasion: Exotic mussels in Lake Erie, *Limnol. Oceanogr.*, 45, 1778–1787, <https://doi.org/10.4319/lo.2000.45.8.1778>, 2000.
- Hidalgo, M., Rossi, V., Monroy, P., Ser-Giacomi, E., Hernandez-Garcia, E., Guijarro, B., Massuti, E., Alemany, F., Jadaud, A., Perez, J. L., and Reglero, P.: Accounting for ocean connectivity and hydroclimate in fish recruitment fluctuations within transboundary metapopulations, *Ecol. Appl.*, 29, e01913, <https://doi.org/10.1002/eap.1913>, 2019.
- Hiddink, J. G., Andreollo, M., Mouillot, D., Beuvier, J., Albouy, C., Thuiller, W., and Manel, S.: Low Connectivity between Mediterranean Marine Protected Areas: A Biophysical Modeling Approach for the Dusky Grouper *Epinephelus marginatus*, *PLoS One*, 8, <https://doi.org/10.1371/journal.pone.0068564>, 2013.
- Hogan, J. D., Thiessen, R. J., Sale, P. F., and Heath, D. D.: Local retention, dispersal and fluctuating connectivity among populations of a coral reef fish, *Oecologia*, 168, 61–71, <https://doi.org/10.1007/s00442-011-2058-1>, 2012.
- James, M. K., Polton, J. A., Brereton, A. R., Howell, K. L., Nimmo-Smith, W. A. M., and Knights, A. M.: Reverse engineering field-derived vertical distribution profiles to infer larval swimming behaviors, *P. Natl. Acad. Sci. USA*, 116, 11818–11823, <https://doi.org/10.1073/pnas.1900238116>, 2019.
- Karnauskas, M., Paris, C. B., Zapfe, G., Gruss, A., Walter, J. F., and Schirripa, M. J.: Use of the Connectivity Modeling System to estimate movements of gag grouper (*Mycteroperca microlepis*) recruits in the northern Gulf of Mexico, *SEDAR33-DW18*, SEDAR, North Charleston, SC, 12 pp., 2013.
- Karnauskas, M., Shertzer, K. W., Paris, C. B., Farmer, N. A., Switzer, T. S., Lowerre-Barbieri, S. K., Kellison, G. T., He, R., and Vaz, A. C.: Source–sink recruitment of red snapper: Connectivity between the Gulf of Mexico and Atlantic Ocean, *Fisheries Oceanography*, 31, 571–586, <https://doi.org/10.1111/fog.12607>, 2022.
- Klein, M., Teixeira, S., Assis, J., Serrão, E. A., Gonçalves, E. J., and Borges, R.: High Interannual Variability in Connectivity and Genetic Pool of a Temperate Clingfish Matches Oceanographic Transport Predictions, *PLoS One*, 11, <https://doi.org/10.1371/journal.pone.0165881>, 2016.
- Leis, J. M.: Perspectives on Larval Behaviour in Biophysical Modelling of Larval Dispersal in Marine, Demersal Fishes, *Oceans*, 2, 1–25, <https://doi.org/10.3390/oceans2010001>, 2020.
- Leon, L. F., Smith, R. E. H., Hipsey, M. R., Bocaniov, S. A., Higgins, S. N., Hecky, R. E., Antenucci, J. P., Imberger, J. A., and Guildford, S. J.: Application of a 3D hydrodynamic–biological model for seasonal and spatial dynamics of water quality and phytoplankton in Lake Erie, *J. Great Lakes Res.*, 37, 41–53, <https://doi.org/10.1016/j.jglr.2010.12.007>, 2011.
- León, L. F., Imberger, J., Smith, R. E. H., Hecky, R. E., Lam, D. C. L., and Schertzer, W. M.: Modeling as a tool for nutrient management in Lake Erie: a hydrodynamics study, *J. Great Lakes Res.*, 31, 309–318 2005.
- Lequeux, B. D., Ahumada-Sempoal, M. A., Lopez-Perez, A., and Reyes-Hernandez, C.: Coral connectivity between equatorial eastern Pacific marine protected areas: A biophysical modeling approach, *PLoS One*, 13, e0202995, <https://doi.org/10.1371/journal.pone.0202995>, 2018.
- Lett, C., Nguyen-Huu, T., Cuif, M., Saenz-Agudelo, P., and Kaplan, D. M.: Linking local retention, self-recruitment, and persistence in marine metapopulations, *Ecology*, 96, 2236–2244, <https://doi.org/10.1890/14-1305.1>, 2015.
- Lin, S., Boegman, L., Shan, S., and Mulligan, R.: An automatic lake-model application using near-real-time data forcing: development of an operational forecast workflow (COAST-LINES) for Lake Erie, *Geosci. Model Dev.*, 15, 1331–1353, <https://doi.org/10.5194/gmd-15-1331-2022>, 2022.
- Lin, S. Q., Boegman, L., Valipour, R., Bouffard, D., Ackerman, J. D., and Zhao, Y.: Three-dimensional modeling of sediment resuspension in a large shallow lake, *J. Great Lakes Res.*, 47, 970–984, <https://doi.org/10.1016/j.jglr.2021.04.014>, 2021.
- Liu, W., Bocaniov, S. A., Lamb, K. G., and Smith, R. E. H.: Three dimensional modeling of the effects of changes in meteorological forcing on the thermal structure of Lake Erie, *Journal of Great Lakes Research*, 40, 827–840, <https://doi.org/10.1016/j.jglr.2014.08.002>, 2014.
- Meerhoff, E., Yannicelli, B., Dewitte, B., Díaz-Cabrera, E., Vega-Retter, C., Ramos, M., Bravo, L., Concha, E., Hernández-Vaca, F., and Véliz, D.: Asymmetric connectivity of the lobster *Panulirus pascuensis* in remote islands of the southern Pacific: importance for its management and conservation, *Bulletin of Marine Science*, 94, 753–774, <https://doi.org/10.5343/bms.2017.1114>, 2018.
- Michie, C., Lundquist, C. J., Lavery, S. D., and Della Penna, A.: Spatial and temporal variation in the predicted dispersal of marine larvae around coastal Aotearoa New Zealand, *Frontiers in Marine Science*, 10, <https://doi.org/10.3389/fmars.2023.1292081>, 2024.
- Nadal, I., Picciulin, M., Falcieri, F. M., García-Lafuente, J., Sammartino, S., and Ghezzi, M.: Spatio-temporal connectivity and dispersal seasonal patterns in the Adriatic Sea using a retention clock approach, *Frontiers in Marine Science*, 11, <https://doi.org/10.3389/fmars.2024.1360077>, 2024.
- Nolasco, R., Dubert, J., Acuña, J. L., Aguión, A., Cruz, T., Fernandes, J. N., Geiger, K. J., Jacinto, D., Macho, G., Mateus, D., Rivera, A., Román, S., Thiébaud, E., Vazquez, E., and Queiroga, H.: Biophysical modelling of larval dispersal

- and population connectivity of a stalked barnacle: implications for fishery governance, *Mar. Ecol. Prog. Ser.*, 694, 105–123, <https://doi.org/10.3354/meps14097>, 2022.
- Okubo, A.: Oceanic diffusion diagrams, *Deep-Sea Res.*, 18, 789–802, 1971.
- Paris, C., Cowen, R. K., Claro, R., and Lindeman, K. C.: Larval transport pathways from Cuban snapper (*Lutjanidae*) spawning aggregations based on biophysical modeling, *Mar. Ecol. Prog. Ser.*, 296, 93–106, 2005.
- Rowe, M. D., Prendergast, S. E., Alofs, K. M., Bunnell, D. B., Rutherford, E. S., and Anderson, E. J.: Predicting larval alewife transport in Lake Michigan using hydrodynamic and Lagrangian particle dispersion models, *Limnol. Oceanogr.*, 67, 2042–2058, <https://doi.org/10.1002/lno.12186>, 2022.
- Saint-Amand, A., Lambrechts, J., and Hanert, E.: Biophysical models resolution affects coral connectivity estimates, *Sci. Rep.-UK*, 13, 9414, <https://doi.org/10.1038/s41598-023-36158-5>, 2023.
- Sato, M., Honda, K., Nakamura, Y., Bernardo, L. P. C., Bolisay, K. O., Yamamoto, T., Herrera, E. C., Nakajima, Y., Lian, C., Uy, W. H., Fortes, M. D., Nadaoka, K., and Nakaoka, M.: Hydrodynamics rather than type of coastline shapes self-recruitment in anemonefishes, *Limnol. Oceanogr.*, <https://doi.org/10.1002/lno.12399>, 2023.
- Shi, W.: Data for for manuscript: Computation of Self-Recruitment in Fish Larvae using Forward- and Backward-in-Time Particle Tracking in a Lagrangian Model (SWIM-v2.0) of the Simulated Circulation of Lake Erie (AEM3D-v1.1.2), Zenodo [code and data set], <https://doi.org/10.5281/zenodo.14789098>, 2025.
- Shi, W., Boegman, L., Shan, S., Zhao, Y., Ackerman, J. D., Amidon, Z., Jabbari, A., and Roseman, E.: A Larval “Recruitment Kernel” to Predict Hatching Locations and Quantify Recruitment Patterns, *Water Resour. Res.*, 60, <https://doi.org/10.1029/2023wr036099>, 2024.
- Suca, J. J., Ji, R., Baumann, H., Pham, K., Silva, T. L., Wiley, D. N., Feng, Z., and Llopiz, J. K.: Larval transport pathways from three prominent sand lance habitats in the Gulf of Maine, *Fisheries Oceanography*, 31, 333–352, <https://doi.org/10.1111/fog.12580>, 2022.
- Thygesen, U. H.: How to reverse time in stochastic particle tracking models, *J. Marine Syst.*, 88, 159–168, <https://doi.org/10.1016/j.jmarsys.2011.03.009>, 2011.
- Torrado, H., Mourre, B., Raventos, N., Carreras, C., Tintoré, J., Pascual, M., and Macpherson, E.: Impact of individual early life traits in larval dispersal: A multispecies approach using backtracking models, *Prog. Oceanogr.*, 192, <https://doi.org/10.1016/j.pocean.2021.102518>, 2021.
- Valipour, R., Bouffard, D., Boegman, L., and Rao, Y. R.: Near-inertial waves in Lake Erie, *Limnol. Oceanogr.*, 60, 1522–1535, <https://doi.org/10.1002/lno.10114>, 2015.
- Wang, Q., Boegman, L., Nakhaei, N., and Ackerman, J. D.: Multi-year three-dimensional simulation of seasonal variation in phytoplankton species composition in a large shallow lake, *Ocean Model.*, 189, <https://doi.org/10.1016/j.ocemod.2024.102374>, 2024.
- Wolanski, E., Richmond, R. H., and Golbuu, Y.: Oceanographic chaos and its role in larval self-recruitment and connectivity among fish populations in Micronesia, *Estuar. Coast. Shelf S.*, 259, <https://doi.org/10.1016/j.ecss.2021.107461>, 2021.