#Comments from Romain Chaput

I have carefully reviewed the manuscript by Shi et al., which presents a valuable theoretical and modelling framework for estimating self-recruitment (SR) in larvae by using both forward and backward particle tracking in a biophysical model. The approach is based on a Lagrangian model. This methodology is especially used in the context of population connectivity studies, where the accurate estimation of SR is important for ecological research, conservation planning, and ecosystem management.

The manuscript is clearly written, well-structured, and provides a strong demonstration of the methodology through a case study in Lake Erie. Notably, the authors show how this approach can be linked with genetic studies to validate or complement connectivity estimates, which significantly broadens the applicability of their method.

The study is timely and offers a potentially general tool for assessing self-recruitment, larval dispersal and connectivity in aquatic systems. Overall, I support the publication of this manuscript after minor revisions. My main concern relates to the number of particles used in the simulations and how this might influence the robustness of SR estimates.

Response: We sincerely thank Dr Chaput for the carefully reading, revision and recognition. The comments have been addressed in our point-by-point responses below.

Minor Comments:

1. The manuscript would benefit from a more detailed discussion on whether the number of released particles is sufficient to saturate the system, especially given the stochasticity introduced in the simulations. A paragraph in the *Discussion* section addressing how particle number affects SR estimates, and the potential biases introduced by under-sampling, would strengthen the paper. Ideally, some justification or sensitivity analysis could be added or referenced.

Response: We thank the reviewer for the insightful comment. In response, we have conducted a sensitivity analysis and added a paragraph in the Results Section. Our results on local retention and self-recruitment are reliable when more than 40000 particles are released in both forward and backward tracking simulations. However, releasing >40,000 particles in forward-tracking simulations does not inherently eliminate the self-recruitment (SR) biases caused by under-sampling. As shown in Eq. (3), such biases can only be avoided if realistic larval production is released from all potential source locations.

It is known that the tracking results could be sensitive to the number of particles released (Béguer-Pon et al., 2016). When too few particles are released, the

results fluctuate (Figure 3). To avoid these errors and achieve statistically stable results, a sensitivity analysis was conducted on the number of particles released. The results show that the estimated dispersal rate, local retention (Eq. 2) and self-recruitment are largely invariant when the number of particles released exceeds 40000 in both forward and backward tracking simulations (Figure 3).

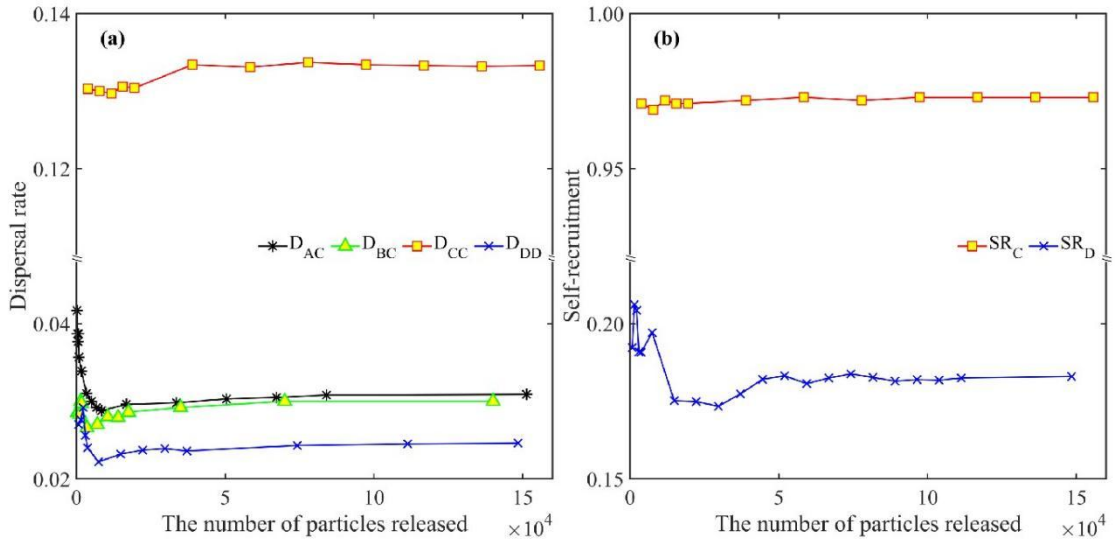

Figure 3. A sensitivity analysis on the number of particles released. (a) In forward simulations, the dispersal rate D_{ij} becomes statistically stable when more than 40000 particles are released. (b) In backtracking simulations, the self-recruitment SR_i becomes statistically stable when more than 40000 particles are released.

2. It would be valuable to include a short discussion on how larval mortality during dispersal and settlement could influence SR estimates. Furthermore, if the model outputs or the analytical framework could allow for the inference of mortality (e.g., when both larval production and SR are known), this should be briefly discussed.

Response: We thank the reviewer for this comment, it's really a good point. We have written a description in the Discussion Section as the follows:

Local retention (LR) is typically more challenging to evaluate empirically, compared to SR (Lett et al., 2015). SR can be estimated by sampling recruits/juveniles at a location and determining the proportion of recruits that originate locally using DNA relationships. However, estimating LR requires sampling all of the eggs/larvae produced at a location that successfully grow into juveniles and disperse to many other locations, yet some juveniles may be transported to unknown locations and be missed. Theoretical local retention (TLR) is thus more challenging to evaluate, as it requires knowledge of the total number of larvae produced ($N_i^{'}$), including those lost to mortality. For example, Almany et al. (2017) sampled adult and juvenile *Amphiprion percula* and *Chaetodon vagabundus* from eight different locations in Papua New Guinea and assigned juveniles to their parents according to DNA

relationships. The location of the parents served as the source location of the juveniles, allowing the researchers to determine the number of juveniles produced from each source location. However, it remained unknown whether all larvae were transported exclusively to these eight locations, as well as the total number of larvae produced. The difficulty in sampling newly hatched larvae, i.e., measuring $N_i^{'}$, is likely why it is common to apply different larval particle release strategies (e.g., releasing a random, constant, or number of particles proportional to the area) in forward tracking simulations. As such, larval mortality (Mor) can affect the estimation of LR (Eq. 16). Specifically, increasing mortality reduces $\sum_{j=1}^{n} D_{ij}$ (the denominator of LR), while its impact on the numerator D_{ii} remains unknown, causing the exact mechanisms by which mortality influences LR to be poorly understood. SR depends on either the recruitment rate (Eq. 5) or settlement rate (Eq. 7), and whether larval mortality affects these rates requires further research.

$$\text{Mor} = \frac{N_i^{'} - \sum_{j=1}^{n} D_{ij} N_i^{'}}{N_i^{'}} = 1 - \sum_{j=1}^{n} D_{ij} = 1 - \frac{TLR_i}{LR_i} \ , \quad (16)$$

Specific Comments:

1. Lines 15–17: The sentence beginning by "However, various strategies have been employed…" is somewhat confusing. Consider rephrasing for clarity.

Response: We have revised the sentence as:
However, various strategies have been employed for releasing larval particles, including releasing a random number, a constant number, or a number particles proportional to the location area or the larval production. The lack of a consistent approach leads to ambiguous results.

2. Line 231: The number of larvae released should be explicitly stated in the main text, along with a brief rationale for varying numbers across different release locations.

Response: We have added the sentences in Section 3.3 as:
To study the impacts, on SR, of varying the number of particles released from a source location, three different numbers of particles were released from each region (Tables 1 and 2). For example, in forward tracking, 16800, 84000, 151200 particles were released in region A, and 17500, 70000, 140000 particles were released in region B. Each release region was divided into 500 m × 500 m AEM3D grid cells, and particles were released at the centers of the cells. Since the cell counts vary across regions (Table A1), for example there are 24 cells in region A while 50 cells in region B, we adjusted the number of particles released in each cell to standardize the total number of particles released per region. For example, we released 100, 500, 900 particles in each cell in region A (100 particles/cell/time × 24 cells × 7 times = 16800

particles), and 50, 200, 400 particles in each cell in region B (50 particles/cell/time × 50 cells × 7 times = 17500 particles).

3. Line 232: Mention that the tracking duration corresponds to the pelagic larval duration (PLD) of the target species.

Response: We have added a sentence as:
this tracking duration corresponds to the pelagic larval duration of Lake Whitefish.

4. Line 235: Please clarify why forward tracking is conducted from four regions while backtracking is performed from only two. Does this reflect known settlement areas or observed recruitment patterns in Lake Erie?

Response: We have clarified this in Section 3.3 as:
In our preliminary tests, backtracking simulations from regions A and B showed negligible settlement in any of the four regions (A, B, C and D), we thus restricted particle release to regions C and D in subsequent backtracking simulations for the sake of computational efficiency.

5. Table 1: Define LR_FA clearly in the caption. Also, include a brief explanation of whether the particle release numbers are sufficient for system saturation.

Response: We have added the description of $\mathrm{LR\_F}_i$ in the caption. We have conducted a sensitivity analysis and added a paragraph in the Results Section as presented above.

6. Table 2: Add the term SR_{ij} to the caption.

Response: We have added the description of $\mathrm{SR\_B}_i$ in the caption.

7. Table 3: The methodology used to derive the larval numbers in this table is not immediately clear. Are these values inferred from combined forward and backward simulations? Please clarify this in the caption with a brief methodological summary.

Response: We have rewritten the caption of Table 3 as: The number of larvae produced in the four regions was computed based on the backtracking simulations from region C ($\boldsymbol{N'_{i\_C}}$) and from region D ($\boldsymbol{N'_{i\_D}}$) and the corresponding larval density (number per cell) $\boldsymbol{d'_{i\_C}}$ and $\boldsymbol{d'_{i\_D}}$.

A detailed description of N'_{i_C} and d'_{i_C} are given in the last paragraph of Section 3.4.

8. Line 339: The phrase "different larval particle release strategies" needs clarification. Does this refer to spatial distribution, timing, number of particles, or something else?

Response: We have clarified it as:

The difficulty in sampling newly hatched larvae, i.e., measuring $N_i^{'}$, is likely why it is common to apply different larval particle release strategies (e.g., releasing a random, constant, or number of particles proportional to the area) in forward tracking simulations.

This is a strong and valuable contribution, and with the suggested clarifications and additions, I believe the manuscript will be of high interest to the modelling and marine connectivity communities.

Response: We sincerely thank the reviewer for the time and effort in reviewing our submission and these kind words.

Reference

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, 10.1093/icesjms/fsv169, 2016.

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, 10.1890/14-1305.1, 2015.