Firstly, I think this is a great manuscript and I enjoyed reading it – a novel idea with potentially very useful results for those involved in marine debris management efforts on Galapagos (and other remote islands if this methodology were implemented elsewhere). The effectiveness of the clean-up strategies proposed in this manuscript will depend on the veracity of the assumptions a significant proportion of debris undergoes resuspension, but this is clearly stated in the conclusion. So overall, I think this manuscript will be valuable for island managers, and researchers working in this field. I do have some technical questions/requests for clarification though:

- I wonder whether you carried out a sensitivity analysis to test whether you released sufficient particles? With such high resolution hydrodynamical data forcing your particle-tracking, significant dispersion can occur over a 60-day integration time with an original separation of <4km (e.g. the off-diagonal cells in the transition matrix are ‘noisy’, and this is probably why). This in itself is not an issue since practitioners will probably not be using your transition matrix, but it would be nice to know how robust, say, Figure 7 is to particle number. I’m guessing that you were limited to 700k particles due to storage (since you were saving particle positions with a very high output frequency) but, if it is tractable, a quick sensitivity test might provide some assurance.
- I’m struggling to completely follow section 2.5. If I understand correctly, the equation on line 162 models the distribution of debris on coastlines in the limit of 100% resuspension. You simulate clean-up as removing the outgoing nodes of a clean-up target cell. You then say that “particles will accumulate at the target cleanup node”, but how can particles accumulate if you’re constantly removing them?
- It’s also not clear to me why you based your definition of steady state on the number of particles in the ocean – does the vector $\mathbf{v}$ not reach steady state?
- I’m also finding the “benefit” metric quite difficult to follow. “Zero connectivity between different nodes”, to me, implies that 100% of resuspended debris enters the ocean, but I don’t think this is what you meant?
- I’m not 100% convinced by your sensitivity tests to the initial macroplastic distribution (3.3). You’ve tested how robust your method is to uncertainty in the initial distribution by using a completely random initial distribution, i.e. assuming that the mass of plastic on beaches is completely decorrelated across length scales $L > 4$km. Is this realistic?
Your result that the efficacy of node rankings remains the same with a random initial distribution is not surprising to me, since mesoscale ocean structures are much larger than 4km so will on average still see a ‘uniform’ distribution of resuspended debris. But given that van Sebille et al. (2019) showed that most debris incident to Galapagos arrives from the East, and that there could be large-scale effects from wind shadows, wake eddies, etc., I’d have expected that there would be some large-scale structure in the distribution of debris. I wonder if a more realistic way to model an uncertain initial macroplastic distribution might be by generating perlin noise with a wavelength larger than 4km (e.g. maybe the length scale of an island).

- This is not a criticism, but I wonder to what extent using a time-mean transition matrix affects your results. For instance, if we have sites A, B and C each releasing 1 unit of debris per time-step with probability $P(A \rightarrow B)=1/6$ and $P(B \rightarrow C)=1/6$, the probability of transition $P(A \rightarrow C)$ is 1/36 (over two timesteps). After 12 time-steps, 1/3 units would have been transported from A→C. But if these transition probabilities were time-varying and turned out to be $P(A \rightarrow B)= P(B \rightarrow C)=1$ during time-steps 1-2 and 0 otherwise, 1 unit would have been transported $A \rightarrow C$ – 3x more than the time-mean case, even though $P(A \rightarrow C)= P(B \rightarrow C)=1/6$ in both cases when averaged over 12 time-steps. This is obviously an artificially bad case and I completely accept that this goes beyond the scope of your study, but I didn’t see a mention of this in your discussion of limitations so I was wondering whether you thought this was a limitation (or if there is evidence that this time variability in the transition matrix is not important).