# Title

Review of biophysical modelling in investigating connectivity patterns in marine species with a pelagic larval stage.

Assessing the impact of parameterization decisions in biophysical dispersal models on connectivity metrics

Making the right choices: a meta-analysis of connectivity metrics derived from biophysical dispersal models

# Introduction

Biophysical dispersal modelling (BDM) has been used to investigate marine dispersal patterns for close to 30 years (Bartsch and Backhaus 1988), and has become increasingly employed in the last 15 years to estimate patterns of connectivity. BDMs produce potential patterns of connectivity using a physical (i.e. oceanographic) model coupled with a model of the marine species of interest, using the knowledge about their biology [CITE]. These BDMs lend themselves to studies inferring distributional patterns (e.g. connectivity) of marine larval dispersal for species with two distinct life stages, a pelagic larval and contrasting sedentary adult phases (Cowen and Sponaugle 2009). BDMs are powerful research tools, allowing us to investigate potential connectivity patterns over wide spatial scales (Andrello *et al.* 2013; Treml *et al.* 2015) and long temporal scales (Richar *et al.* 2014; Berglund *et al.* 2012), that are experimentally difficult to capture using current empirical methods (Cowen and Sponaugle 2009). These models also provide us the ability to investigate how different early-life history traits (e.g. mortality, swimming ability, and orientation) affect the patterns of connectivity (Butler IV *et al.* 2011; Staaterman *et al.* 2012), questions that would typically be intractable *in situ* due to the inherent difficulties of studying the larval phase of marine species [CITE].

Complexity and realism in the BDMs used by studies has progressively increased; driven in part by advancements in parallel programming coupled with greater access to high performance computing, and open access high resolution oceanographic flow models [CITE]. Empirical research studies are also behind these increases, they are increasingly demonstrating the ability of pelagic larvae to affect their own fate via a variety of behaviours (Leis 2010). This trend was clear in Miller’s (2007) review of ichthyoplankton biophysical modelling (referred to by Miller as individual-level, coupled physical-biological models), where 43.3% of studies implemented behaviour,predominately age-dependent vertical migration behaviour. This increased focus on the importance of the biology of these dispersal models has led to important insights about the importance of larval behaviour to dispersal. Vertical movement of larval fish has been shown to increase retention near the natal site, and affect connectivity patterns (Paris and Cowen 2004; Brochier *et al.* 2008; Robins *et al.* 2013; Paris *et al.* 2007). Horizontal swimming has not been included in BDMs to the same extent as vertical migration, and it can likely only influence connectivity patterns if the swimming is orientated (Leis 2007). From empirical studies we know larval fish can achieve sizeable swimming speeds, potentially giving the ability to out-swim the surrounding current, and it improves substantially with growth (Leis 2006). Horizontal swimming with orientation increased settlement success, including settlement at or near the natal site (Staaterman *et al.* 2012). Spawning behaviour (along with the location of the natal site) has been shown to significantly affect dispersal patterns (Puckett *et al.* 2014).

This previous work has made it clear that the biological components of transport models can strongly affect connectivity patterns. And this emphasis has been recognised in the ICES publication, "Manual of recommended practices for modelling physical-biological interactions during fish early life", which recommends their inclusion and provides guidance for how to do this (Anon 2009). Larval behaviour has also been emphasised in published BDM tools, such as the Connectivity Modelling System (Paris *et al.* 2013) & Connie [CITE]. But despite this increasing attention paid to the inclusion of biology in BDM, there remains little consensus on the relative importance of different parameters. If BDMs are to be a robust tool for understanding the connectivity of marine populations, such a concensus is ultimately required. It may be that there is considerable context dependence to the relative importance of some biological parameters (the importance of horizontal swimming may depend on the strength of the flow regime or the length of the larval period for instance) while the inclusion of others may always make a large difference to connectivity (mortality for instance).

In this study we review the literature on BDMs to evaluate the ability of this collective body of work to provide insight into the ramifications of the choices we might make when attempting to include biological realism in transport models. We achieve this aim by firstly evaluating if the field is generally progressing in its inclusion of biological parameters. We look at the collection of parameters included in models, how this is changing over time, and if there are trends in how parameters are included (or excluded) based on the context of models. Secondly, we look at the ramifications of the inclusion of different parameters on model results. Here we evaluate if there are observable differences in key metrics of connectivity as a result of the inclusion of specific biological parameters. We also assess the degree to which the physical context of the models may regulate any observed effect of biological parameters. This information will serve to guide the application of BDMs into the future and by highlighting key biological parameters, allows for focusing of empirical studies where they are most needed.

# Methods

## Literature search

We searched for studies to review using the keywords; connectivity, model\* and marine, querying the online search engines ISI Web of Knowledge and SCOPUS. We found 76 studies matching these criteria, published in 34 different journals (see the full list of references and the data extracted in the Supplementary materials). We constrained the search to publications from 2009, as this was the publication date of a manual recommending the way forward for biophysical connectivity models (Anon 2009), and also close to the date of the last review in this area (Miller 2007). Whilst our primary research interest is the connectivity patterns of ichthyoplankton, we included connectivity studies modelling other marine species with a pelagic larval stage (e.g. molluscs, crustaceans and corals), owing to often negligible differences in parameterised behaviours.

## Data classification

Most studies performed connectivity simulations using variations of their BDM, involving different input values for one or more parameters. This was often undertaken to test the effect of different parameters (representing different species/taxa or variations in behaviours of the same species/taxa) on resulting connectivity patterns. In this review, we separately identify each of these scenarios within a study and refer to each as a model. Nearly all studies consisted of a small number of models, although one study contained 99 models.

We characterised each model within a study according to a variety of traits, which are broadly grouped into four themes. The first theme, Model context, was general information about the study, such as study motivation, location, model species, and whether the system was tropical or temperate. To see if there was an effect of the modelled geographical area on settlement, the regions were broadly categorised into open (coastal, island) or closed (bay or inland sea). We classified the motivation of a study based on its broad aims: climate change, marine park connectivity, dispersal and settlement, and theoretical. Climate change objectives identified temporal and spatial changes in larval connectivity patterns with increases in ocean temperatures. Marine park connectivity studies determined the extent of connectedness of either existing or potential marine parks for different species. Dispersal and settlement approaches concerned either the dispersal or settlement of a species using known spawning or settlement sites. Theoretical approaches to connectivity studies investigated how different larval parameters effected the patterns of connectivity.

The second model characteristic theme focused on the inputs used to parametrise the physical model. Assessing the physical model included details such as the oceanographic model used, the particle tracking software, diffusion methods and the model time-step. The third theme considered the input parameters of the biological model, recording larval traits such as pelagic larval duration, settlement competency, spawning times and behaviours, e.g. mortality, vertical migration, and growth. The length of the settlement competency window was determined as the time-period between the first age of settlement competency and the end of the pelagic larval duration.

The fourth model characteristic theme recorded the metrics used to measure the connectivity patterns of the model and their specific values. The metrics used varied considerably, with in-excess of 25 different metrics used to describe the patterns of connectivity. These metrics ranged from common metrics (i.e. dispersal kernels and connectivity matrices) to some which were unique to the study (defined in Table 1). These metrics can be broadly categorised into two related classes. The first class included those metrics which summarised the dispersal kernel (mean, variance directionality) while the second captured metrics which identified connectivity pathways more explicitly ( self-recruitment, local retention and settlement success). Specific values for every metric were not available from every study. Where values were included in figures we used the software package GraphClick (version 3.0.3; http://www.arizona-software.ch/graphclick/) to obtain them.

## Data analysis

To address our aim of evaluating the application of biological realism in BDMs, we summarized model characteristics from our first three themes and evaluated their relative occurrence and how this may have changed over time. We conducted all data analysis using the R programming language version 3.3.2 (R Core Team n.d.), specifically the package *tidyverse* to manipulate the data (Wickham 2016) and the package *ggplot2* to produce the figures (Wickham 2009).

To address our aim of identifying the importance of including various biological parameters on different metrics of connectivity as well as the context dependence of any such relationships, we compared each of the most commonly reported connectivity metrics (self-recruitment, local retention, settlement success and dispersal distance) amongst models grouped by the different biological parameters included. Due to variation in the nature of metrics reported (or which could be obtained) from study to study, samples sizes for these comparison varied considerably. We required there be a minimum of at least X studies for a given metric within a parameter group in order to conduct a comparison. We performed linear regressions and Student's t-tests (checking for equal variance using Bartlett's test) with the R base package. For each dispersal metric, we used means to describe the data in a single statistic, allowing for comparisons against the physical and biological model parameters.

Table 1: The definition of terms used when referring to the output of biophysical connectivity studies in this review, definitions were taken from (Burgess et al. 2014; Cowen and Sponaugle 2009)

|  |  |
| --- | --- |
| Term | Definition |
| Self-recruitment | The percentage of settlers to a site, that (regardless of their origin site) were natal to that site. |
| Local retention | The percentage of the larvae natal to a site that recruited back to that site. |
| Settlement success | The percentage of recruits natal to a site that recruit to any site. |
| Dispersal kernel | The probability density function of the dispersal distance for the larvae released in the system. |
| Connectivity matrix | A matrix with columns specifying *i* natal sites and rows specifying *j* settlement sites, showing the probability of a larva spawned at site*i* that settles at site*j* at position*ij*. |

# Current trends in biophysical connectivity modelling

For the period we reviewed, there was a rapid increase in the number of studies per year which peaked in 2012 (20 studies), with a steady decline following this (Figure 1). The average number of BDM studies published per year was 15.3 in our review. In the last 8 years of Miller's (2007) review an average of 6.6 studies were published, demonstrating BDM studies have become more prevalent as a research tool. In this section we explore the current trends in both input parameters and output metrics of connectivity studies, using the four model characteristic themes of model context, physical model parameters, biological model parameters, and the metrics of connectivity used.

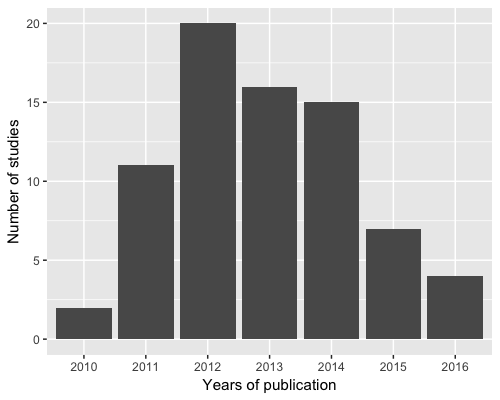


Figure 1: The number of published studies per year included in the review

## Model context

Ten specific and one generic marine taxonomic groups were identified in the studies we reviewed (Figure 2). The generic taxonomic group was defined as a suite of biological parameters that could potentially apply to multiple species, e.g. many species of reef fish most likely have a pelagic larval duration of 20 days. Generic species were used in 40.2% of the models, although one study used 99 models (28.4% of the total analysed models in the review) implementing a suite of generic parameters (Treml *et al.* 2015). The 10 known taxonomic groups were comprised from 47 species and another seven where the modelled biological group was specified at the family level, e.g. Labridae or Sparidae. The two most common taxa modelled were fish (39%) and bivalves (10%). The number of ichthyoplankton BDMs that were published each year was consistent with that reported by Miller (2007).

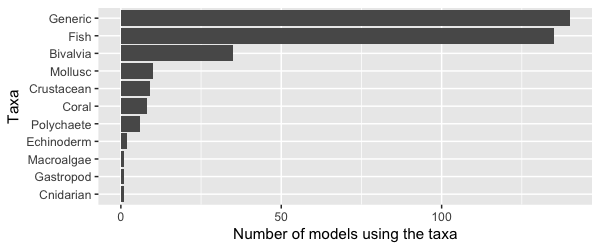


Figure 2: The different taxonomic groups that were used to parameterise the biological models in the reviewed connectivity studies

Most of the connectivity studies used an existing BDM (82.6%), the rest either build a BDM to answer the study question (16.0%) or information about the BDM was missing from the methods (1.4%). This high model reuse suggests an increasing maturity in the model output and the ease of adoption for end-users. The three most popular BDMs used amongst the reviewed studies were Ichthyop (18.7%) (Lett *et al.* 2008), Connectivity Modelling System (12.0%) (Paris *et al.* 2013), & MGET (8.0%) (Roberts *et al.* 2010). Twenty-one different BDMs (ignoring version numbers) were used amongst the studies, although we were unable to identify the exact particle tracking model used for 22.7% of the studies. The studies all used forecasting for their models, except for one hindcast solution (whereby the larvae are initially tracked from settlement sites to determine their natal sites) (Wren and Kobayashi 2016).

The overwhelming motivation for using a BDM was to answer questions about dispersal and settlement (68.0%). Dispersal models without settlement (13.3%) and studies motivated by investigating marine park connectivity (12.0%) occurred in roughly equal frequencies. Theoretical based studies (4.0%) and studies investigating how climate change effects connectivity (2.7%) were more sporadic motivations. A handful of studies used a purely theoretical approach to understanding the biology that influences connectivity, although within another applied context, e.g. using different species in a study that have varied behaviours or traits, studies indirectly accomplished this.

The studies in this review used hydrodynamic models covering 20 oceanographic regions across the world, two studies even investigated connectivity patterns at a global scale (Figure 3). Similarly to patterns observed by Miller (2007), modelling effort concentrated on highly industrialised fishing regions, such as the North-East Atlantic and the Mediterranean Sea. As in Miller (2007) most studies were of temperate regions (64.9%), though the number of studies in tropical oceanic regions increased from 4% to 32.4%. The early-life histories of tropical marine species are commonly researched, and thus this increase focus on tropical connectivity patterns is not unexpected.

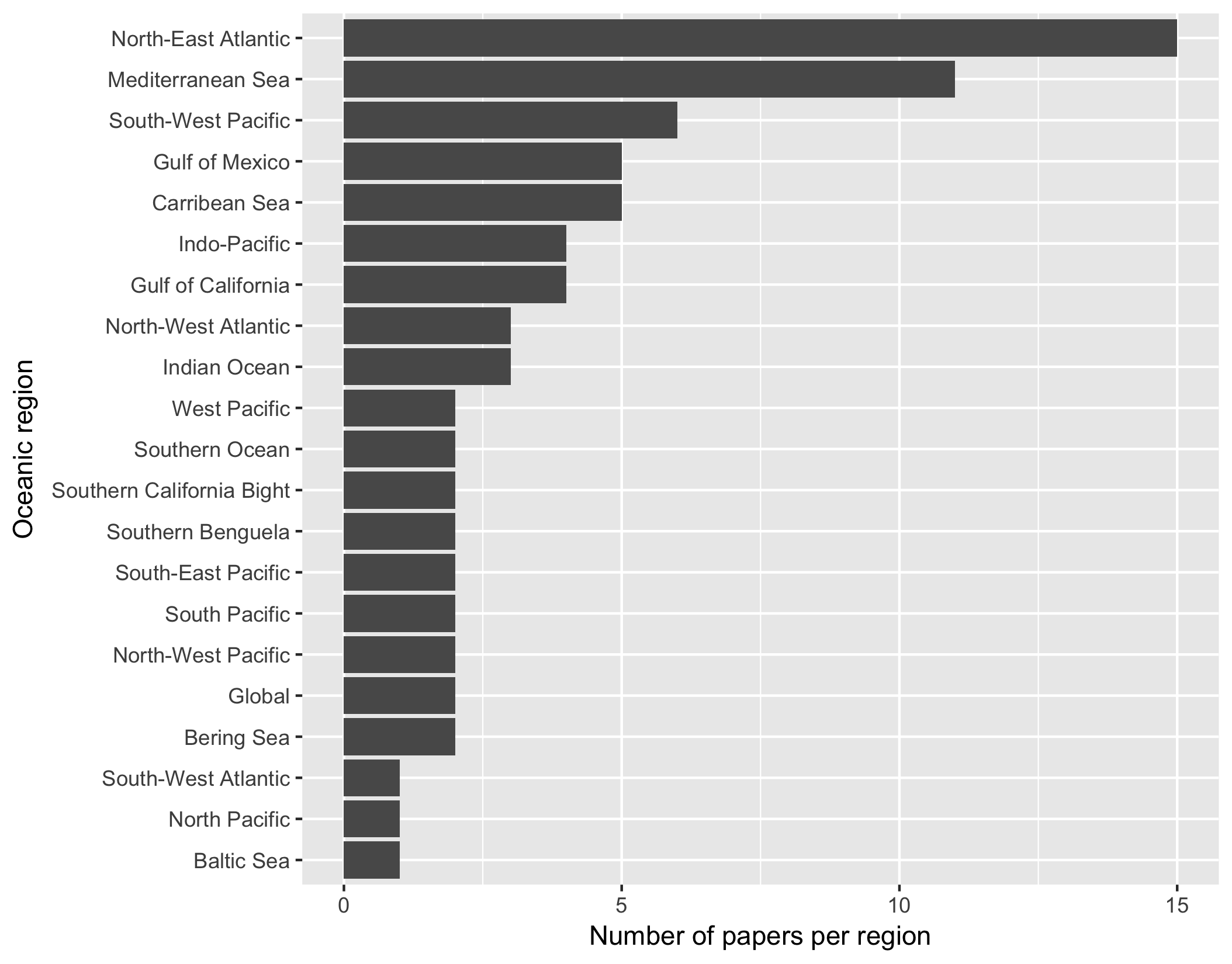


Figure 3: The oceanic regions used in the reviewed studies

The models used different time-scales of oceanographic data depending on the trends of interest (inter- or intra-annual). Studies trended towards running the BDM over a handful of years (median = 3 years), with the majority using a single year in their connectivity study (26.9%). One study used the average of multiple years input data to interpret mean connectivity patterns (Feutry *et al.* 2013). Other outliers included studies investigating climate change scenarios, which used periods of 60 years (Aiken *et al.* 2011) and 130 years (Andrello *et al.* 2014). In these studies, the earliest year modelled was 1950, and the latest year modelled was 2100 (using ocean current data based on past averages). The average of the dates modelled was a start date of 2001 with an end date of 2008. Although not all models used continuous years as input choosing rather to use specific years. As high powered computing becomes more readily available and with the known variability between consecutive years in hydrodynamic modelling, the trend for BDMs should hopefully be towards using longer periods of ocean data as input, though there is little evidence for this trend over the past 20 years based on Miller (2007) and this review. Its important to move in this direction as modelling longer periods of time allows for inter-annual comparisons of connectivity patterns.

## How connectivity studies are parameterising the physical model

The choice of oceanographic model depended upon the study region. The two most common ocean circulation models used were the Regional Ocean Modelling System (ROMS; 25.3%) and the Hybrid Coordinate Ocean Model (HYCOM; 17.3%). Nested circulation models were used in the more complex particle trackers (16.2%) to bypass issues of low resolution in modelling. Miller (2007) showed in their review the majority of studies used Finite Element Models (FEM; 20%) or Hamburg Shelf Ocean Model (HAMSOM; 12.9%) and ROMS was chosen by only 7.1% of studies. It shows a trend towards open-source models, which are readily available on the Internet and present day connections allow for expeditious download times. The mean resolution used in these connectivity models was 5.6km2, with the finest resolution 50m2 (Brennan *et al.* 2014) to the coarsest resolution of 33km2 (Crochelet *et al.* 2013). The mean resolution is half that of Miller's 2007 review (10.8km2), so while within studies there was no trend over time to higher resolution models, between reviews hydrodynamic precision has increased.

The tracking time-step for larvae in the system (environmental forcing) were either fixed time-steps or varied per other factors. Runge-kutta fourth-order integration allows for interpolating the larvae accurately at smaller time-steps than the time interval used in the physical model and was used in 48.1% of the studies (usage could be higher as 33.8% of the studies did not specify the integration scheme used). The median time-step was one hour, and ranged between 60 seconds (Moritz *et al.* 2013) and 24 hours (Mora *et al.* 2012). Miller (2007) reported the same median time-step, with similar ranges, except the maximum value found in that review was 48 hours. The time-step used was not reported in 37.3% of the studies. No clear rationale was given for the choice of time-step amongst the studies, likely determined by the choice of physical model and computational reasons.

## How connectivity studies are parametrising the biological model

The biological parameters implemented were not consistent across studies, apart from: pelagic larval duration (PLD; 100%) and spawning behaviour (98.3% - as mentioned previously there was a single hindcast study where spawning is not applicable; Table 2). This was expected, as these traits are fundamental to connectivity studies for marine larvae, as they need to be spawned somewhere and there must be a time specified for the pelagic phase before the larvae can settle. A reason for the lack of consistency in the other traits and behaviours captured by the biological models could be that researchers do not believe that these traits and behaviours have much influence on connectivity patterns. This reason seems unlikely given the recent publications acknowledging the influence of behaviour on connectivity patterns. The more likely explanation is that there is not enough knowledge about the early-life history of many marine species, making parametrisation of the model difficult for certain traits and behaviours.

Table 2: The behaviours that have been implemented in the modelled connectivity studies. Note that multiple behaviours can be implemented in the one model.

|  |  |
| --- | --- |
| Functionality | Implemented (%) |
| Pelagic larval duration | 100.0 |
| Settlement competency window | 48.3 |
| Spawning strategies | 98.3 |
| Mortality | 41.0 |
| Growth | 6.7 |
| Sensory ability | 58.3 |
| Orientation | 2.6 |
| Swimming behaviours | 25.7 |

### Pelagic larval duration

The studies we reviewed exhibited two different approaches when implementing the pelagic larval duration. The standard approach (92.4% of models) was to make the PLD value fixed for each larva. If the individual larva reaches this time, it either must settle (if possible) or it is considered dead. The other approach was to implement variability in the PLD (5.6% of models), either using a Gaussian distribution (based on knowledge of the population) or instead of using time, make the PLD dependent on ocean temperature. The fixed PLD values ranged from a brief 2 days to a prolonged 420 days, with a median of 30 days. The PLDs were taken from known ranges found in the literature where available, with the longest PLDs occurring for invertebrate species e.g. the long duration (~400 days) of the phyllosoma stage of lobster larvae (Butler IV *et al.* 2011; Kough *et al.* 2013).

### 

### Spawning strategies

A second fundamental element, included by necessity, in every study in our review was a spawning strategy, which determined how the larvae were released in the model. The spawning strategy was typically dependent on the biology of the modelled taxa and included pelagic eggs, demersal eggs, or brooding. To implement these different strategies, studies changed depth for the release of the larvae, e.g. demersal eggs were released on the bottom of the natal reef and pelagic eggs were released higher up the water column. The two strategies for releasing the particles at depth were either to use a fixed depth (37.5%), or to randomly release the larvae over a range of depth values (60.8%), i.e. release larvae in in the 0-20 m range of the water column. The timing of when larvae were released varied from hourly intervals (Jacobi *et al.* 2012), daily (Kough *et al.* 2013), monthly (Staaterman *et al.* 2012), on lunar cycles (Davies *et al.* 2014) or a one-off spawning event (Herbert *et al.* 2011). The spawning windows relied on the known biology and the times of year at which the species is known to produce larvae. Typically the spawning events were pre-determined, but several studies used temperature as a spawning initiation (Andrello *et al.* 2014; Rochette *et al.* 2012) or even wave height was used to release abalone larvae in one model (Miyake *et al.* 2011). The release sites and settlement sites were the same for most studies. However, several studies (15.6%) used a limited set of key release sites with more available settlement sites. The larvae spawned at each site ranged from 10 larvae up to 10 million, with a median of 6,800. The total larvae tracked in the system (spawning site multiplied by the number of sites multiplied by the number of releases) had a similarly large range, from 328 to 610 million larvae, with a median total of 3.2 million. To simplify tracking these large numbers of larvae in the system, five studies (6.5%) used a super-individual concept, whereby a single larva in the system represents a sizeable number of larvae with a normalised collection of attributes. There are large computational benefits to this approach, especially in models that use a continuous space. The caveat is that you must consider how much individual variation you are willing to lose by modelling these super-individuals.

### 

### Mortality and growth

Larval mortality was implemented in only 41% of the models (Table 2). However, the implementation of mortality varied across the models, and the most common mortality function used was linear (82.2%), whereby a fixed percentage of the larvae were killed each day. The daily mortality rate in these studies varied from 1.7% to 30.0% with an average of 19.4%. Other mortality implementations were to kill the larvae based on a known temperature or salinity threshold (7.9%), using a decay function that decreased mortality based on age (7.9%) or using a Weibull function (2.1%). Unlike in Miller's 2007 review, where growth was implemented in one-third of the studies, only 6% of the models reviewed had the ability for larvae to grow (Table 2). The rationale is that the models used in this review are solely investigating connectivity, where growth is not considered such an important factor affecting the strength of connectivity amongst sites, unless it is co-implemented with a sinking velocity dependent on mass.

### Behaviour

The degree to which larvae can behave, specifically sense, orient and swim directionally (vertically or horizontally) is a fundamental concern in broader field of larval biology which has potentially important ramifications for connectivity. As discussed previously, this has been recognised in the world of BDMs, with general support for the inclusion of such behaviours. In this review we classified models according to which of the four major larval behaviours were implemented; no movement (passive movement), movement (e.g. horizontal or vertical swimming), orientation (ability to orientate towards a reef using cues) and settlement (settle at a site using a sensory buffer or sticky zones).

Larvae were given the ability to move in 26% of the models, conversely in 74% of the models the larvae were assumed to be passive and moved solely using the physical model. Miller (2007) observed that 43% of the fish studies in the reviewed papers implemented a facet of larval behaviour, thus the percentage of BDM studies implementing larval behaviour is decreasing. Of the behaviours classified as movement, diel vertical migration was the most commonly implemented behaviour (45.4%; Table 3). Ontogenetic vertical migration, the idea that marine species change their vertical position in the water column as a factor of their ontogeny, was used in 27.2% of the models. Horizontal swimming was implemented in 12.5% of the models, although it must be mentioned that random movement to mimic turbulence does capture an extent of undirected horizontal swimming. A smaller number implemented vertical swimming in the larvae (4.5%). Other movements included in models were prompted by environmental conditions such as tides (circatidal migration 5.7%,), salinity (halocline migration; 1.1%), or migrating with the pycnocline (6.8%). Another movement implementation, as mentioned above, was using sinking velocity (3.4%) in conjunction with growth, e.g. the heavier the larvae the faster it sinks. The last strategy involved giving the pelagic eggs a buoyancy (4.5%) based on the salinity level in the water column, thus they can float at different vertical migrations.

Table 3: The percentage of ichthyoplankton models in this review which implemented different swimming behaviours. Note the percentages sum to greater than 1 because multiple swimming behaviours can be implemented in a single model.

|  |  |
| --- | --- |
| Swimming behaviour | Implemented (%) |
| Horizontal swimming | 12.5 |
| Vertical swimming | 4.5 |
| Ontogenetic vertical swimming | 27.2 |
| Diel vertical migration | 45.4 |
| Halocline migration | 1.1 |
| Circatidal migration | 5.7 |
| Pycnocline migration | 6.8 |
| Sinking velocity | 3.4 |
| Egg buoyancy | 4.5 |

The ability of the marine larvae to sense settlement habitat was implemented in 50.9% of the models. Of these, 95.4% opted for a buffer zone implementation, where each settlement site has a buffer added to it, increasing its size. The mean size of the buffer zone was 8.18 km (SD = 5.2 km). The other implementation, which essentially works on similar principles, was to give the larvae a detection distance relative to the centre of the settlement site (Staaterman *et al.* 2012). This settlement cue was used to allow the larvae to orientate towards the reef if the larvae was within the cue distance ranging between 1-10 km (median distance was 5 km), with a combination of early and late onset orientating abilities.

For the studies modelling fish, the implementation of only passive movement decreased over the reviewed period. In 2015, for the first time there were more studies using movement behaviours than other forms of behaviour (Figure 4). Implementation of settlement sensory strategies remain constant over the published time period. Interestingly, the ability of fish to orientate towards settlement sites was only found to be implemented in one study in 2012. For invertebrates, there is no real change over time in the pattern of implemented behaviours (Figure 5).

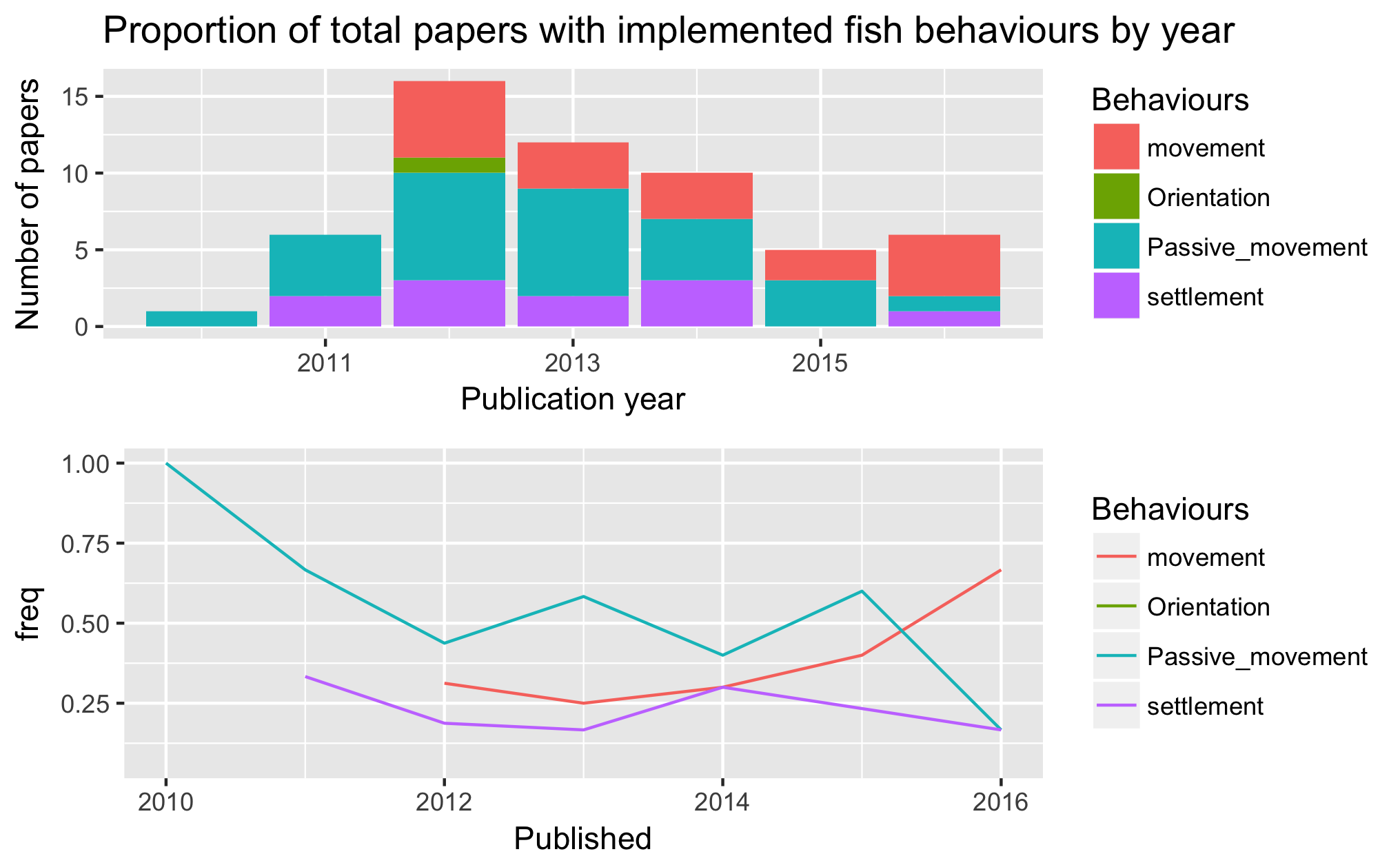


Figure 4: The total number (a) and proportion (b) of behaviours implemented in BDMs by year.

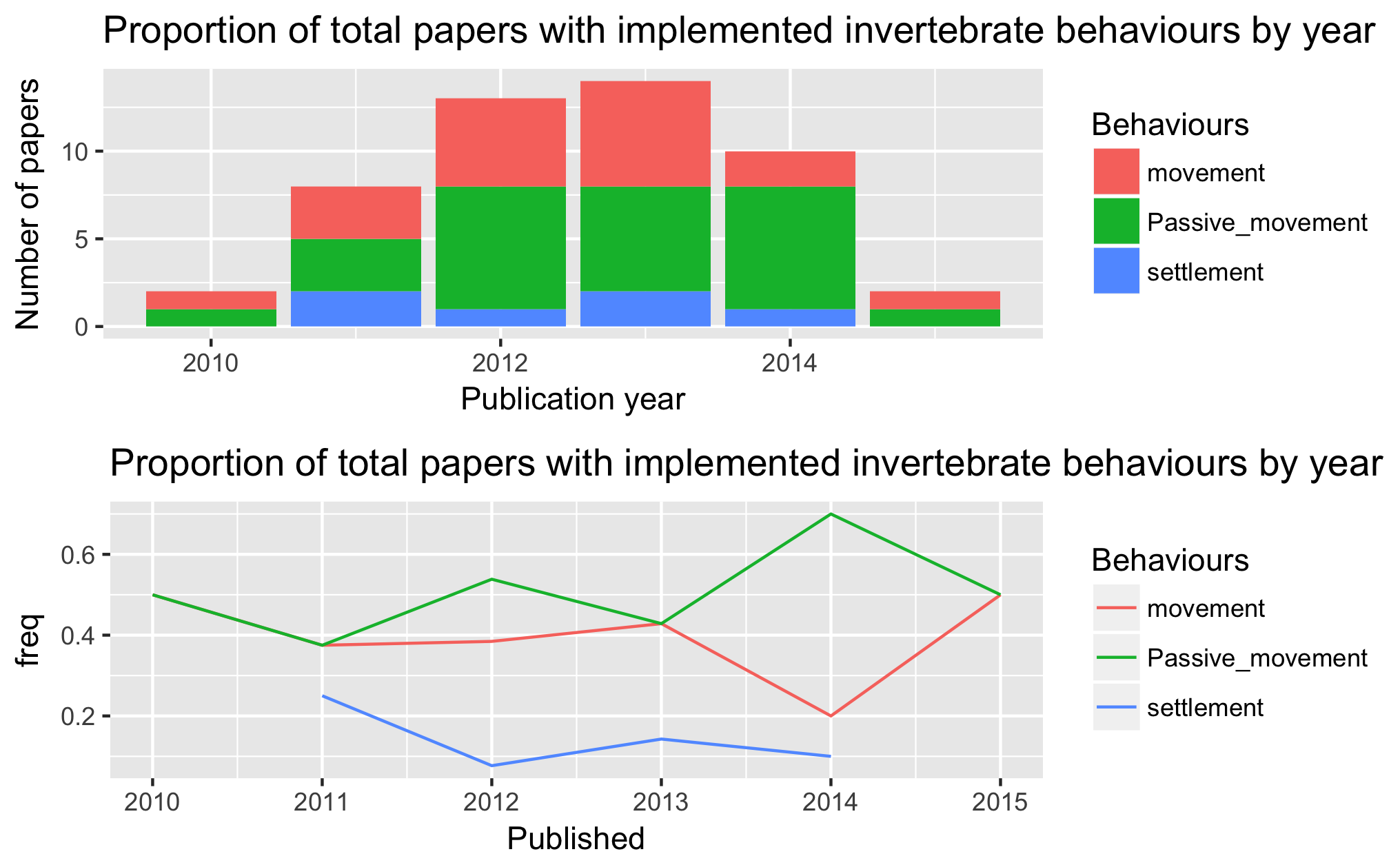


Figure 5: The proportion of behaviours in the published invertebrate studies by year

This finding is extremely interesting from a biological viewpoint. There is empirical knowledge of fish behaviour influencing the outcome of connectivity (Leis 2006; Cowen and Sponaugle 2009), BDM studies demonstrating that connectivity changes with behaviour, and studies suggesting the way forward is to include behaviour (Staaterman and Paris 2014; Anon 2009). However, there are still many studies including models with no implemented behaviour at all - passive larvae were used in 75% of model configurations. There is also a well established body of empirical literature which highlights the often extensive swimming abilities of ichthyoplankton (Leis 2006, 2010). Even where horizontal swimming is not important, it is well understood zooplankton exhibit diel vertical migration patterns to below the mixed layer, hypothesised due to predator avoidance [CITE]. Several studies in the review showed that vertical migration (either diel or ontogenetic) increased self-recruitment and limit the dispersal kernel (Young *et al.* 2012; Aiken *et al.* 2011). Orientation was only implemented in one study, yet there is an increasing body of research in the last decade regarding the ability of reef fish larvae to orientate towards natal reefs using olfactory mechanisms (Gerlach *et al.* 2007), celestial cues (Leis *et al.* 2014; Faillettaz *et al.* 2015) and a magnetic compass (Bottesch *et al.* 2016).

One possible reason for not implementing behaviour is that the complete (or even partial) early life history is often unknown for many species and researchers are trying to avoid using any behavioural assumptions in their models. Fish larval behaviour is known to be species specific, yet many models are limited to using known data from the family taxonomic level due to the lack of knowledge on a particular species. It was been argued that this approach is fraught with danger, as behaviour can vary widely amongst species within a family taxonomic grouping (Leis 2010). The lack of species specific knowledge poses the question, can we only model species where we fully understand the early life history, or is there a degree of assumption that is deemed to still be scientifically valid to answer the study question.

## The metrics used to describe connectivity

There were more than 25 different measures used to describe the patterns of dispersal and settlement. The common approaches were to use combinations of dispersal kernels (64.0%) and connectivity matrices (57.2%; Table 4). Less common measures of model output were analysis using graph theory (8.1%), larval centre of mass (average geographical position of the larvae; 2.6%) (Lacroix *et al.* 2013; Miyake *et al.* 2011), isotropy (describes the inertia around the centre of mass; 1.3%) (Ayata *et al.* 2010), positive area (number of areas that had larvae in them, 1.3%) (Puckett *et al.* 2014), and proportion of sites settled (1.3%) (Robins *et al.* 2013).

Table 4: Those most common metrics published in the connectivity studies evaluated in this review and the parameters reported per metric used.

|  |  |  |
| --- | --- | --- |
| Metric | Sub-metric | Reported (%) |
| Dispersal kernel |  | **64.0** |
|  | Mean distance | 57.1 |
|  | Maximum distance | 33.3 |
|  | Median distance | 11.9 |
|  | Upper quartile distance | 11.9 |
|  | Minimum distance | 4.8 |
| Connectivity matrix |  | **57.2** |
|  | Self-recruitment | 70.0 |
|  | Local retention | 62.0 |
|  | Settlement success | 42.0 |
| Graph theory |  | **8.1** |

The trend in biophysical modelling is to present data using visual techniques such as heat maps, which while being effective tools for visually communicating complex patterns qualitatively, they are not amenable to the extraction of exact metric values. The most common connectivity values reported in the reviewed studies–when a connectivity matrix was presented–were self-recruitment (39.9%), local retention (35.3%) and global settlement success of larvae (23.9%; Table 4). However, while these response metrics were explicitly described, we were only able to extract actual values from 70.0%, 62.0%, and 42.0% of the studies in which connectivity matrices were described for self-recruitment, local-retention and settlement success respectively.

Based on the extracted data, there was a large spread of the mean self-recruitment amongst the models, ranging from 0-98%, depending on the study, with a median self-recruitment percentage of 13.0%. Local retention, while mentioned for 31.7% of studies, was only available for 5.5% of the models. Mean local retention ranged from 0.3-46.0%, with a median value of 6.0%. Overall mean settlement success of the larvae (available from 18.0% of the studies), ranged from 0.1-80% with a median of 17.0%. The mean dispersal distance (available from 15.2% of the studies) had a median value of 78.2 km (range of 9.1-952 km).

The percentage of successfully settled larvae per model seems unusual high. It was expected that the settlement success would be extremely low, even if mortality is not included in the system. Perhaps because many studies have reported potential connectivity instead of realised connectivity (realised is a more accurate measure offset by biological processes such as egg production and mortality). The large ranges amongst the metrics for self-recruitment, local retention and settlement success, was also unexpected. However, the large range of input parameters must be responsible for such variations - in the next section we explore how these modelling decisions affect these common metrics.

# Influence of modelling decisions on connectivity values

As indicated above, much of the variation observed in connectivity values and patterns across the reviewed study will be driven by modelling environments and especially parameterisation decision. Here we attempt to use the body of reviewed literature to assess the impact of various model and parameter choices on resulting connectivity values. As indicated in the methods section above, we have restricted consideration to only those model characteristics/input parameters and output connectivity metrics that were shared amongst the majority of studies. Specifically we look at the impact of four modelling environment characteristics (choice of model, model resolution, model design and study system) and four biological parameterisation decisions (larval duration, competency window, larval mortality and movement behaviour) on four key metrics of connectivity (self-recruitment, local retention, settlement success & dispersal distance). Note there are situations where there were not enough data points to make valid comparison of specific connectivity metrics.

## Modelling environment

### Model choice

There appears to be no clear effect of either physical or BDM model choice on resulting connectivity metrics (Figure 6; Figure 7). The BDM model CMS had consistent reported self-recruitment values between studies, but highly variable settlement success. The PTM has higher values than the median for self-recruitment, settlement success and dispersal distance. Icthyop had low settlement success values, except for two studies. Studies using HYCOM and POM had the largest ranges amongst the metrics. ROMS and HANSOM had the least variation of metrics for the common hydrodynamic models.

*Note: Potentially reduce these graphs to the most popular models (put the cut-off at 5%?)*

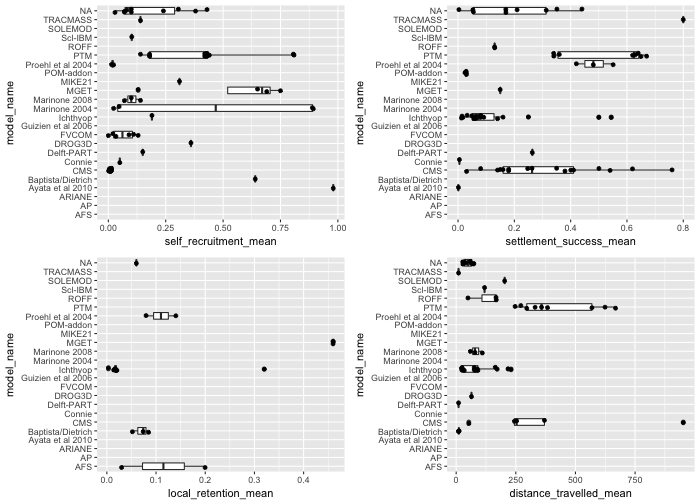


Figure 6: Self recruitment (a), local retention (b), settlement success (c) and distance travelled (d) for each BDM used by studies in the review. Box plot represents …

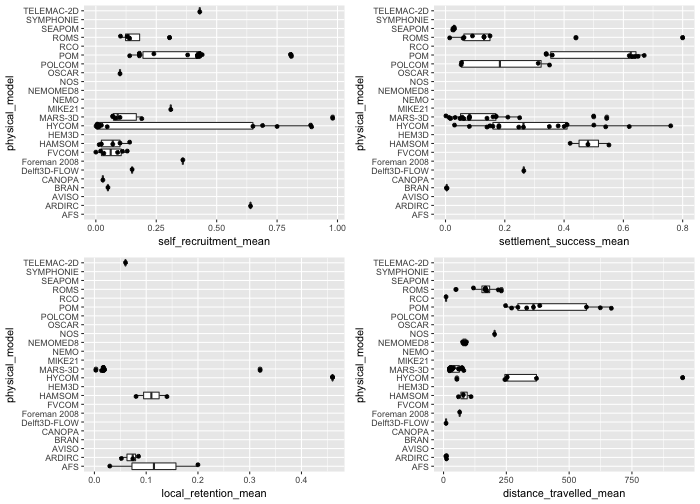


Figure 7: How four common connectivity metrics (self-recruitment, local retention, settlement success and dispersal distance) are influenced by choice of hydrodynamic model

### Model resolution

There was a significant positive relationship between model resolution and local retention (r2=0.82, F(1,17)=79.11, p<0.05), but no relationship between model resolution and self-recruitment (r2=0.05, F(1,57)=2.75, p>0.05), settlement success (r2=0.01, F(1,64)=0.77, p>0.05) or dispersal distance (r2=0.01, F(1,51)=0.69, p>0.05). This was unexpected given that sub-mesoscale processes are expected to be more accurately captured in the hydrodynamic model as the resolution increases (spatial scale decreases), as would the connectivity patterns of the modelled larvae, e.g. increasing self-recruitment as small eddies trap larvae close to natal reefs. Local retention increases as model resolution gets coarser, which is difficult to explain in absence of no relationship with self-recruitment or settlement success and model resolution.

#### Nested model design

The self-recruitment at natal sites was significantly reduced (28.8% versus 0.9%) when nested models (i.e. sub-models with higher resolution near settlement sites) were used (t=7.15, df=49.13, p<0.05), however, caution should be used when evaluating this result due to the large unbalance between the groups (only 10 models describing self-recruitment had nested designs versus 50 which did not). There was no significant difference in the settlement success or dispersal distance when using nested models, and there were no studies using nested models where we were able to get local retention values. Again, the results are not expected, as similar to above, increasing model resolution theoretically increases self-recruitment and local retention, while decreasing the dispersal distance.

### Open systems vs closed systems

There was no evidence that the geographic openness of the modelled system was consistently affecting the connectivity results derived from the models. While both self-recruitment and settlement success (38% and 22% higher respectively) were greater in closed system, the difference was not significant (self-recruitment: t=1.24, df=16.49, p>0.05; settlement success: t=0.97, df=64, p>0.05). This result is unexpected given the presumably elevated rates of larval loss in open systems, but it should be noted that the samples sizes here are small. Of course, this lack of difference could also be indicative of researchers appropriately choosing the boundaries of the modelled systems based on the biology of the modelled taxa. In this case, even seemingly open systems, may in fact be demographically closed ones. Dispersal distance was 3.4 times longer in open systems, compared to closed (t=-3.75, df=41.82, p<=0.05). This can be explained by the geographical nature of an open system, potentially allowing more larvae to be dispersed away from settlement sites and into large oceanic features. There were not enough data points on local retention for a meaningful comparison.

## Biological parameterisation

### Pelagic larval duration

Local retention to the natal site showed a significant positive relationship with the length of PLD (r2=0.57, F(1,16)=20.95, p<0.05), although this relationship is based on limited data (n= ?). There was no relationship of PLD with either self-recruitment (r2=0.02, F(1,54)=1.12, p>0.05) or settlement success (r2=0.05, F(1,59)=3.08, p>0.05). There was a significant positive relationship of dispersal distance with PLD (r2=0.23, F(1,45)=13.63, p<0.05). Other studies have shown a relationship between PLD and dispersal distance [CITE]. It is interesting that pelagic larval duration does not appear to influence self-recruitment or settlement success at all. Suggesting that the evolutionary pressures for longer pelagic larval duration is being able to disperse further, combined with other potential factors such as increased size before settlement to reduce recruitment mortality.

### Settlement competency window

The settlement competency window size, dependent on the pelagic larval duration, was strongly negatively correlated with local retention (r2=0.70, F(1,9)=20.83, p<0.05) and weakly positively correlated with self-recruitment (r2=0.27, F(1,18)=6.69, p<0.05). The was no significant relationship with settlement success (r2=0.09, F(1,24)=2.39, p>0.05) or dispersal distance (r2=0.04, F(1,19)=0.75, p>0.05). One would assume a greater competency window size would enhance opportunities for successful settlement and to do so over a greater distance. Thus the lack of effect here is surprising though its possibly any observable relationships are masked by the impact of parameterization decisions made by these models which may have stronger effects on settlement success (e.g. mortality, see below). The strong tendency for models with longer window sizes to have lower self-recruitment is seemingly contradictory. However, as this increased window size will affect individuals at later stages of the larval period only, and we know that PLD is positively correlated with dispersal distance (above), any larvae old enough to benefit from the extended window are likely to be far from the natal site, and thus unlikely to contribute to self-recruitment.

### Mortality

Models which included mortality had similar levels of self-recruitment (t=-1.54, df=58, p=>0.05) and settlement success (t=0.70, df=64, p>0.05) to models run without mortality (Figure 8). Including mortality also had no effect on the dispersal distance of the marine larvae (t=-1.16, df=9.13, p>0.05). There was a significant difference in local retention when mortality was applied compared to when it was not (t=-3.65, df=4.30, p<0.05). Mortality did reduce the variation of the settlement success and the dispersal distance (apart from two outliers) values, which was expected. The variation of self-recruitment values increased, which would occur if settlement success also decreased, as less larvae would be arriving from other sites. Local retention, while going up with mortality, which is an expected results, again there are not enough data points to draw strong conclusions.

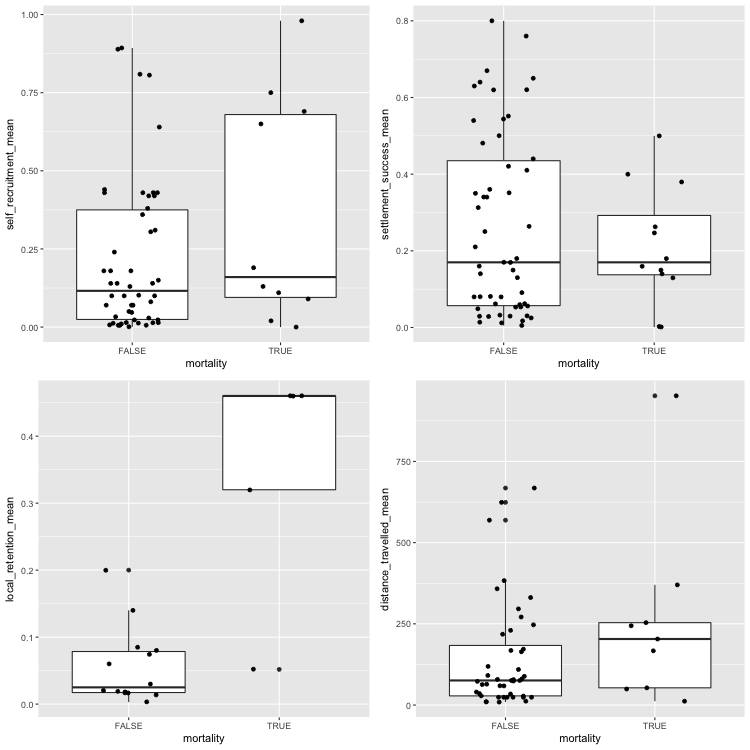


Figure 8: Comparing models that implemented mortality with self-recruitment, local retention, settlement success and dispersal distance values

### Movement behaviours

The inclusion of movement or settlement sensory abilities reduced the self-recruitment of a model (Figure 9). While the data indicates including both of them increases self-recruitment, this is likely due to a low number of data points (n=?). There were slight increases for including either movement or settlement sensory behaviour in the models, but including both movement and settlement greatly increased the settlement success seen the model (Figure 9). Local retention was not commonly reported for studies including movement behaviours, therefore not many conclusions can be drawn from the data. Dispersal distance did not change with movement added, the dispersal distance had increased variability when settlement sensory was included in the model. The ability to sense a settlement site favours particles that have travelled far from the natal site. Potentially sensory zones can overcome coarse near-shore hydrodynamic modelling issues, increasing settlement of larvae. Although when movement and sensory ability was included, this variation was greatly reduced. Unfortunately, there was not enough data on models with orientation to get a good understanding how it affects the four metrics, but the strong assumption is that it would also increase settlement success, as shown in individual studies (Staaterman *et al.* 2012).

*In this section I am still analysing how to statistically analyse the data, as per the email I sent on (15/3)*

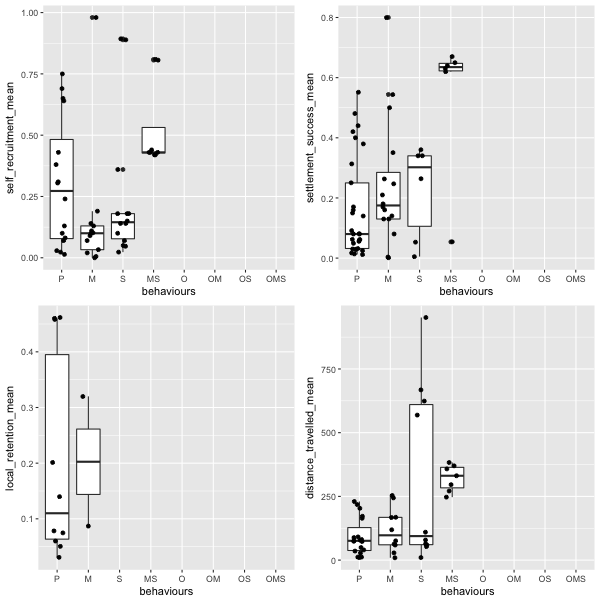


Figure 9: The effect of larval behaviour (P=passive, M=movement, O=orientation, S=settlement sensory, MO=movement and orientation, MS=movement and settlement sensory, OS=orientation and sensory, MOS=movement, orientation and settlement sensory) on the mean self-recruitment, settlement success, local retention and distance travelled in the reviewed connectivity studies

# Conclusion and recommendations

The popularity of BDMs as a connectivity research tool has been on a downward trend since 2012. However, the implementation of these models appears to be getting more complex, especially for ichthyoplankton, where for the first time in 2016 there were more models with movement than without. The proportion of studies using settlement sensory strategies was consistent across the review period for models of either fish or invertebrate species, not keeping up with the trend in implementing movement. Therefore it appears researchers are heeding the recommendations of publications such as the ICES manual and other research showing the importance of movement in influencing connectivity patterns.

There was a large range of values for all the complexity metrics analyzed in this study. This variability is likely due to the non-trivial number of parameter choices made in BDMs, many of which are optional or species specific. These are complex systems and as such it is not easy to standardise the output. Seemingly obvious explanations such as open/closed systems did not explain the variance. Another potential key driver is the inherent variability of the oceanographic currents, which we know to be huge influence on connectivity patterns (Cowen and Sponaugle 2009). Model resolution seemingly had no effect on the connectivity metrics, although using a nested design did reduce self-recruitment. Increasing the size of the settlement window, thus allowing larvae to settle at an earlier age, increased local retention and self-recruitment. Mortality did not seem to affect the connectivity metrics, although including both movement and a sensory ability to select a settlement site increased settlement, self-recruitment and the distance travelled by larval.

The current state of the field feels like is not moving forward as coherently it should be. Studies have been calling for behaviour to be implemented for nearly 10 years, however, many studies are still not including larval behavioural traits such as movement - although this is changing. This is despite the evidence that behaviours do change the described connectivity patterns. Obviously there is a problem with the lack of empirical evidence for many marine species, and this should be a priority research area. The second priority to is provide clarity in metrics so we can meaningfully compare the outcomes of connectivity studies and progress the field further. A consistent approach, including providing underlying data would assist in progressing the field for future comparisons, allowing us to understand why there is so much variation amongst reported measures of connectivity.

# References

Aiken, C. M., Navarrete, S. A., Pelegrí, J. L. (2011) Potential changes in larval dispersal and alongshore connectivity on the central Chilean coast due to an altered wind climate. *Journal of Geophysical Research*.

Andrello, M., Mouillot, D., Beuvier, J., Albouy, C., Thuiller, W., Manel, S. (2013) Low connectivity between mediterranean marine protected areas: a biophysical modeling approach for the dusky grouper *Epinephelus marginatus*. *PLoS ONE*. **8**, e68564.

Andrello, M., Mouillot, D., Somot, S., Thuiller, W., Manel, S. (2014) Additive effects of climate change on connectivity between marine protected areas and larval supply to fished areas. *Diversity and Distributions*. **21**, 139–150.

Anon (2009) *Manual of recommended practices for modelling physical-biological interactions during fish early life*. ICES Cooperative Research Report.

Ayata, S.-D., Lazure, P., Thiébaut, E. (2010) How does the connectivity between populations mediate range limits of marine invertebrates? A case study of larval dispersal between the Bay of Biscay and the English Channel (North-East Atlantic). *Progress in Oceanography*. **87**, 18–36.

Bartsch, J., Backhaus, J. (1988) Numerical simulation of the advection of vertically migrating herring larvae in the North Sea. *Meeresforschung-Reports On Marine Research*. **32**, 30–45.

Berglund, M., Jacobi, M. N., Jonsson, P. R. (2012) Optimal selection of marine protected areas based on connectivity and habitat quality. *Ecological Modelling*.

Bottesch, M., Gerlach, G., Halbach, M., Bally, A., Kingsford, M. J., Mouritsen, H. (2016) A magnetic compass that might help coral reef fish larvae return to their natal reef. *Current biology : CB*. **26**, R1266–R1267.

Brennan, G., Kregting, L., Beatty, G. E., Cole, C., Elsasser, B., Savidge, G., Provan, J. (2014) Understanding macroalgal dispersal in a complex hydrodynamic environment: a combined population genetic and physical modelling approach. *Journal of The Royal Society Interface*. **11**, 20140197.

Brochier, T., Lett, C., Tam, J., Freon, P., Colas, F., Ayon, P. (2008) An individual-based model study of anchovy early life history in the northern Humboldt Current system. *Progress in Oceanography*. **79**, 313–325.

Burgess, S. C., Nickols, K. J., Griesemer, C. D., Barnett, L. A. K., Dedrick, A. G., Satterthwaite, E. V., Yamane, L., Morgan, S. G., White, J. W., Botsford, L. W. (2014) Beyond connectivity: how empirical methods can quantify population persistence to improve marine protected-area design. *Ecological Applications*. **24**, 257–270.

Butler IV, M. J., Paris, C. B., Goldstein, J. S., Matsuda, H., Cowen, R. K. (2011) Behavior constrains the dispersal of long-lived spiny lobster larvae. *Marine Ecology Progress Series*.

Cowen, R. K., Sponaugle, S. (2009) Larval Dispersal and Marine Population Connectivity. *Annual Review of Marine Science*. **1**, 443–466.

Crochelet, E., Chabanet, P., Pothin, K., Lagabrielle, E., Roberts, J. J., Pennober, G., Lecomte-Finiger, R., Petit, M. (2013) Validation of a fish larvae dispersal model with otolith data in the Western Indian Ocean and implications for marine spatial planning in data-poor regions. *Ocean and Coastal Management*. **86**, 13–21.

Davies, S. W., Treml, E. A., Kenkel, C. D., Matz, M. V. (2014) Exploring the role of Micronesian islands in the maintenance of coral genetic diversity in the Pacific Ocean. *Molecular Ecology*. **24**, 70–82.

Faillettaz, R., Blandin, A., Paris, C. B., Koubbi, P., Irisson, J.-O. (2015) Sun-Compass Orientation in Mediterranean Fish Larvae. *PLoS ONE*. **10**, e0135213.

Feutry, P., Vergnes, A., Broderick, D., Lambourdière, J., Keith, P., Ovenden, J. R. (2013) Stretched to the limit; can a short pelagic larval duration connect adult populations of an Indo-Pacific diadromous fish (*Kuhlia rupestris*)? *Molecular Ecology*. **22**, 1518–1530.

Gerlach, G., Atema, J., Kingsford, M. J., Black, K. P., Miller-Sims, V. (2007) Smelling home can prevent dispersal of reef fish larvae. *Proceedings of the National Academy of Sciences*. **104**, 858–863.

Herbert, R. J. H., Willis, J., Jones, E., Ross, K., Hübner, R., Humphreys, J., Jensen, A., Baugh, J. (2011) Invasion in tidal zones on complex coastlines: modelling larvae of the non-native Manila clam, Ruditapes philippinarum, in the UK. *Journal of Biogeography*. **39**, 585–599.

Jacobi, M. N., André, C., Döös, K., Jonsson, P. R. (2012) Identification of subpopulations from connectivity matrices. *Ecography*. **35**, 1004–1016.

Kough, A. S., Paris, C. B., Butler IV, M. J. (2013) Larval connectivity and the international management of fisheries. *PLoS ONE*. **8**, e64970.

Lacroix, G., Maes, G. E., Bolle, L. J., Volckaert, F. A. M. (2013) Modelling dispersal dynamics of the early life stages of a marine flatfish (*Solea solea* L.). *Journal of Sea Research*. **84**, 13–25.

Leis, J. M. (2006) Are larvae of demersal fishes plankton or nekton? In: *Advances in marine biology, vol 51* (eds A. J. Southward & D. W. Sims) pp. 57–141 Australian Museum, Sydney, NSW 2000, Australia. Leis, JM (reprint author), Australian Museum, 6-8 Coll St, Sydney, NSW 2000, Australia. Academic Press Ltd-Elsevier Science Ltd, London.

Leis, J. M. (2007) Behaviour as input for modelling dispersal of fish larvae: behaviour, biogeography, hydrodynamics, ontogeny, physiology and phylogeny meet hydrography. *Marine Ecology Progress Series*.

Leis, J. M. (2010) Ontogeny of behaviour in larvae of marine demersal fishes. *Ichthyological Research*. **57**, 325–342.

Leis, J. M., Paris, C. B., Irisson, J.-O., Yerman, M. N., Siebeck, U. E. (2014) Orientation of fish larvae in situ is consistent among locations, years and methods, but varies with time of day. *Marine Ecology Progress Series*.

Lett, C., Nguyen-Huu, T., Cuif, M., Saenz-Agudelo, P., Kaplan, D. M. (2015) Linking local retention, self-recruitment, and persistence in marine metapopulations. *Ecology*. **96**, 2236–2244.

Lett, C., Verley, P., MULLON, C., Parada, C., Brochier, T., PENVEN, P., BLANKE, B. (2008) A Lagrangian tool for modelling ichthyoplankton dynamics. *Environmental Modelling & Software*. **23**, 1210–1214.

Miller, T. J. (2007) Contribution of individual-based coupled physical-biological models to understanding recruitment in marine fish populations. *Marine Ecology Progress Series*.

Miyake, Y., Kimura, S., Kawamura, T., Kitagawa, T., Takahashi, T., Takami, H. (2011) Population connectivity of Ezo abalone on the northern Pacific coast of Japan in relation to the establishment of harvest refugia. *Marine Ecology Progress Series*.

Mora, C., Treml, E. A., Roberts, J. J., Crosby, K., Roy, D., Tittensor, D. P. (2012) High connectivity among habitats precludes the relationship between dispersal and range size in tropical reef fishes. *Ecography*. **35**, 89–96.

Moritz, C., Meynard, C. N., Devictor, V., Guizien, K., Labrune, C., Guarini, J. M., Mouquet, N. (2013) Disentangling the role of connectivity, environmental filtering, and spatial structure on metacommunity dynamics. *Oikos*.

Paris, C. B., Cherubin, L. M., Cowen, R. K. (2007) Surfing, spinning, or diving from reef to reef: effects on population connectivity. *Marine Ecology Progress Series*.

Paris, C. B., Cowen, R. K. (2004) Direct evidence of a biophysical retention mechanism for coral reef fish larvae. *Limnology and Oceanography*. **49**, 1964–1979.

Paris, C. B., Helgers, J., Sebille, E. van, Srinivasan, A. (2013) Connectivity Modeling System: A probabilistic modeling tool for the multi-scale tracking of biotic and abiotic variability in the ocean. *Environmental Modelling & Software*. **42**, 47–54.

Puckett, B. J., Eggleston, D. B., Kerr, P. C., Luettich Jr, R. A. (2014) Larval dispersal and population connectivity among a network of marine reserves. *Fisheries Oceanography*. **23**, 342–361.

R Core Team R: A language and environment for statistical computing.

Richar, J. I., Kruse, G. H., Curchitser, E. N., Hermann, A. J. (2014) Patterns in connectivity and retention of simulated tanner crab (*Chionoecetes bairdi*) larvae in the eastern Bering Sea. *Progress in Oceanography*.

Roberts, J. J., Best, B. D., Dunn, D. C., Treml, E. A., Halpin, P. N. (2010) Marine Geospatial Ecology Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++. *Environmental Modelling & Software*. **25**, 1197–1207.

Robins, P. E., Neill, S. P., Giménez, L., Jenkins, S. R., Malham, S. K. (2013) Physical and biological controls on larval dispersal and connectivity in a highly energetic shelf sea. *Limnology and Oceanography*. **58**, 505–524.

Rochette, S., Huret, M., Rivot, E., Le Pape, O. (2012) Coupling hydrodynamic and individual-based models to simulate long-term larval supply to coastal nursery areas. *Fisheries Oceanography*. **21**, 229–242.

Staaterman, E., Paris, C. B. (2014) Modelling larval fish navigation: the way forward. *Ices Journal of Marine Science*. **71**, 918–924.

Staaterman, E., Paris, C. B., Helgers, J. (2012) Orientation behavior in fish larvae: A missing piece to Hjort’s critical period hypothesis. *Journal of Theoretical Biology*.

Treml, E. A., Roberts, J. J., Halpin, P. N., Possingham, H. P., Riginos, C. (2015) The emergent geography of biophysical dispersal barriers across the Indo-West Pacific. *Diversity and Distributions*. **21**, 465–476.

Wickham, H. (2009) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wickham, H. (2016) *tidyverse: Easily Install and Load ’Tidyverse’ Packages*. R package version 1.0.0.

Wren, J. L. K., Kobayashi, D. R. (2016) Exploration of the larval pool: development and ground-truthing of a larval transport model off leeward Hawaii. *PeerJ*.

Young, E. F., Rock, J., Meredith, M. P., Belchier, M., Murphy, E. J., Carvalho, G. R. (2012) Physical and behavioural influences on larval fish retention: contrasting patterns in two Antarctic fishes. *Marine Ecology Progress Series*.