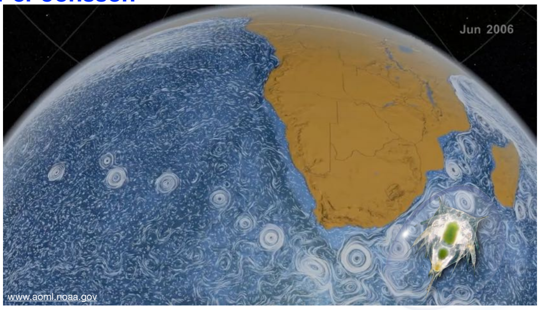
1

Biophysical models of dispersal

Per Jonsson



PhD course autumn 2023: *Land, River, and Seascape Genomics*

2

Time table

Lecture 1: 14.30 - 15.00

Coffee break: 15.00 - 15.20

Lecture 2: 15.20 - 16.15

Short break: 16.15-16.25

Discussion: 16.25-17.00

3

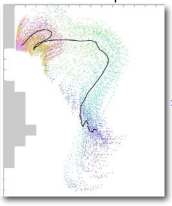
One main task in scape genomics is to identify dispersal barriers and potential for genetic differentiation in relation to environmental factors

- Main method: analysis of genetic markers
- Complementary method: identification of subpopulations and barriers through dispersal modelling, e.g. by using **Biophysical models**

4

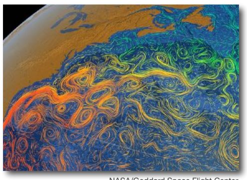
What is a biophysical model?

Biophysical model of dispersal in the land/seascape



=

Hydrodynamic (aerodynamic) model of water & air transport



NASA/Goddard Space Flight Center

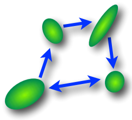
+

Biological characteristics of propagules (gametes, spores, seeds, larvae)

Goals for today:
1. Orientation about how biophysical models are constructed
2. Understand descriptions of biophysical models in articles about scape-genetics/genomics

5

Dispersal and Migration



Dispersal is movement away from the birth site to potential sites of reproduction. This implies potential for gene flow. Dispersal is often partly ***passive***, e.g. through air or water transport of gametes, seeds, spores or larvae. But dispersal may also include ***active*** movements, e.g. juvenile dispersal in birds and mammals.

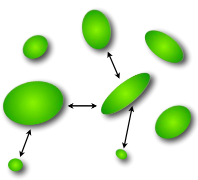
Migration in ecology is ***active***, round-trip movement between different environments, e.g. breeding and feeding areas. This kind of migration does not primarily lead to gene flow.

Migration in population genetics are all movements, ***passive or active***, that lead to gene flow.

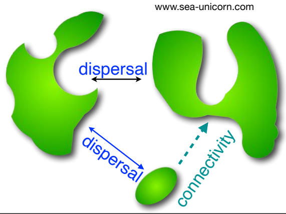
6

Connectivity

- **Structural Connectivity**, a notion purely related to the physical characteristics of the landscape, measuring its heterogeneity and structuring
- **Functional Connectivity**, which represents all the movements of organisms that result in the exchange of genes, biomass or energy between heterogeneous habitat patches. These are either caused, facilitated or hampered by Structural Connectivity patterns

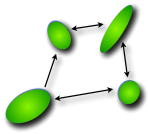


www.sea-unicorn.com



7

Importance of dispersal



ecological time

- Fluctuations in population size
- Local extinction
- Invasion of non-native species

evolutionary time

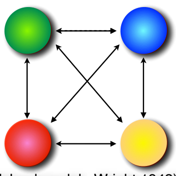
- Genetic diversity & differentiation
- Local adaptation & speciation

- Management of harvested species
- Design of reserves & protected areas

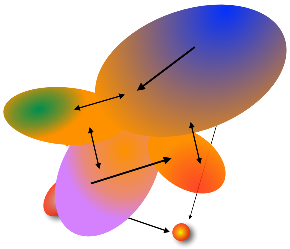
8

How to measure and estimate dispersal

ideal case



(Island model, Wright 1943)



real-scape cases

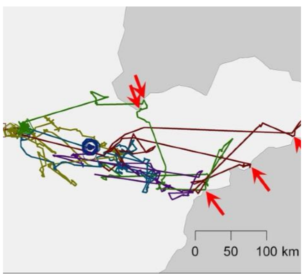
subpopulation

9

How to measure and estimate dispersal

Direct methods

Real-time tracking

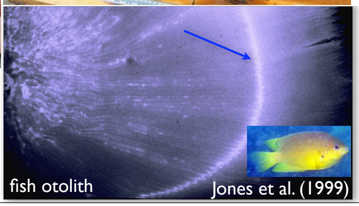


10

How to measure and estimate dispersal

Direct methods

Mark-recapture

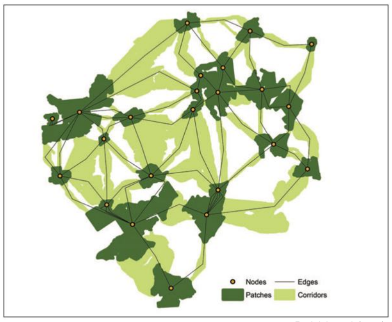


11

How to measure and estimate dispersal

Indirect methods

Landscape and seascape mapping (GIS models)



Rudnick *et al.* (2012)

12

How to measure and estimate dispersal

Indirect methods

Chemical markers: geographic signatures from water chemistry, e.g. using fish otoliths and mussel shells

Chemical markers: isotope signatures related to area-specific food sources

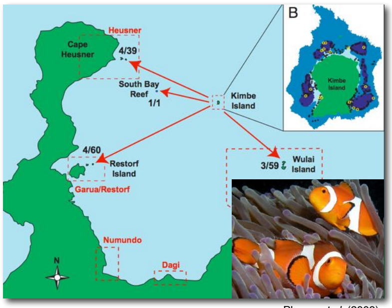
Genetic markers: assignment to sources showing different genetic signatures

Genetic markers: model gene flow from genetic differentiation

Estimate dispersal with genetic markers

Indirect methods

Assignment tests, e.g. using parentage analysis



Current dispersal

Planes *et al.* (2009)

13

Estimate dispersal/gene flow with models

Estimate (model) dispersal from population differentiation

Allelic differences:

1. Mutations
2. Genetic drift
3. Selection
4. Gene flow
5. (Non-random mating)

historic dispersal

$$F_{ST} \approx \frac{1}{(4mN + 1)}$$

14

Problems with genetic differentiation to model dispersal/gene flow

- **Markers may be under selection**
- **Reflects historic connectivity (and demography)**
- **Even very low dispersal (Nm) will erode neutral differentiation**
- **Populations with high N_e show little differentiation**

15

16

Problems with non-genetic methods to estimate dispersal

- **Propagules often small, even microscopic**
- **Large number of propagules are produced**
- **Dispersal may reach great distances**
- **Dispersal into areas that are inaccessible**

Generally difficult to observe and follow dispersal of most propagules

17

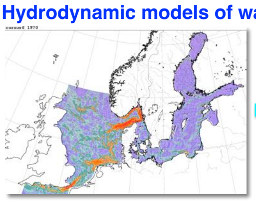
Dispersal in the sea is difficult to study

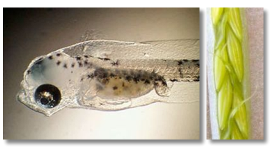
- 70% of invertebrates and fish have a planktonic larval stage
- Most marine propagules (spores & larvae) are numerous, sub-mm, and ***drift with ocean circulation***
- Duration of planktonic dispersal: often many weeks

18

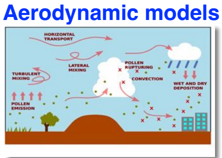
Biophysical modelling of dispersal

Hydrodynamic models of water transport

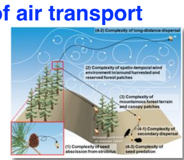




Aerodynamic models of air transport



Nickovic *et al.* (2023), Science of the Total Environment 864,160879

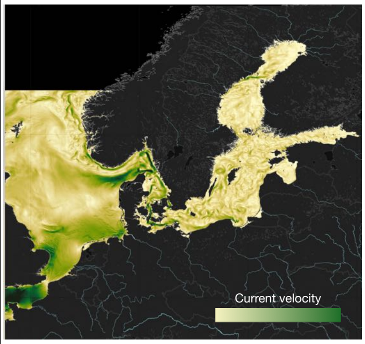


Kim *et al.* (2022), Forest 13, 659

19

Biophysical modelling in the sea

Ocean Circulation Model

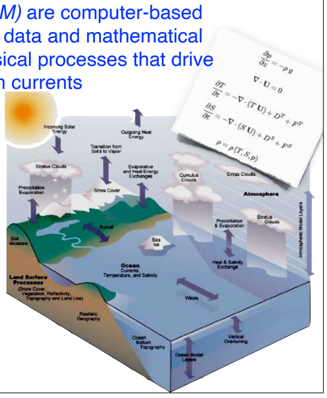


20

What is an ocean circulation model?

Ocean Circulation Models (OCM) are computer-based simulations that use measured data and mathematical equations to re-create the physical processes that drive temperature, salinity and ocean currents

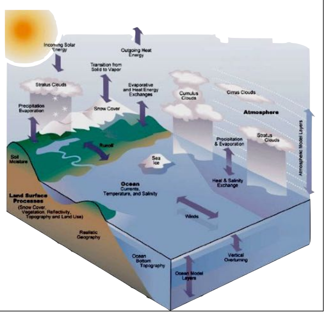
- atmosphere
- oceans
- land surface
- ice



21

Environmental processes driving an OCM

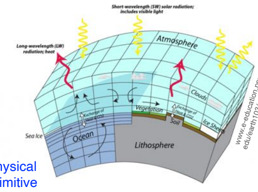
- wind
- atmospheric pressure
- tide
- heat exchange
- precipitation-evaporation
- freshwater outflows
- bathymetry - land contour



How does an OCM work?

Ocean models divide the continuous ocean and atmosphere into a discrete ocean with:

- discrete spatial grid net
- discrete time steps



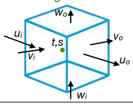
www.e-education.psu.edu/earth103/node/524

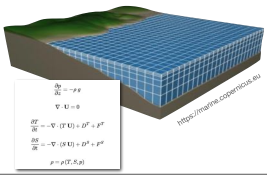
Ocean models describe the ocean properties using physical (and often biogeochemical) relationships called the primitive equations, mainly:

- conservation of mass
- conservation of momentum
- a thermal energy balance

Simplified, the ocean circulation model predicts the hydrodynamic flow of the water in each grid cell:

6 velocities on the sides
1 temperature in the middle
1 salinity in the middle





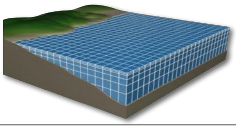
https://marine.copernicus.eu

OCMs simulate several properties from the surface to the sea floor

OCMs are run on super computers (clusters), solving millions of equations in each grid cell

Physical parameters

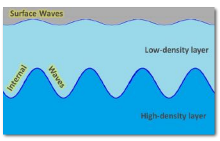
- temperature
- salinity
- currents in 3 dimensions
- sea surface height
- sea ice
- (wave height)



OCMs differ in their complexity

Some mechanisms of water motion is not included in OCMs. Often missing are:

- Surface wave motion
- Internal waves within the water column



Processes that may be important for larvae reaching near-coast waters

25

Extent, resolution and boundaries of OCMs

Extent:

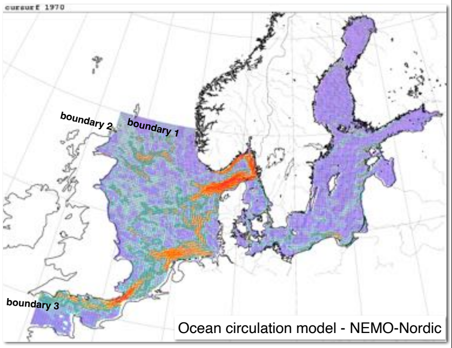
North Sea, Skagerrak, Kattegat and Baltic Sea

Resolution:

Horizontal: 3.7 km
Vertical: 3-22 m (56 layers)
Temporal: 6 min

Model boundaries:

Boundary 1: N 59.5°
Boundary 2: Shetland
Boundary 3: English Channel

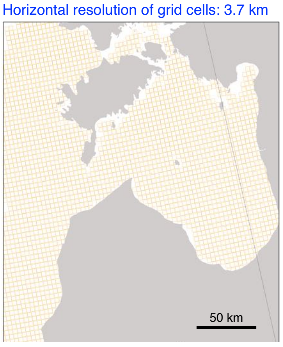


Ocean circulation model - NEMO-Nordic

26

Resolution of OCMs

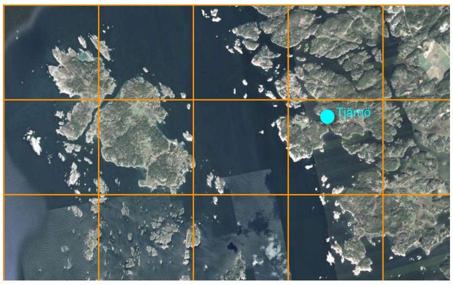
Horizontal resolution of grid cells: 3.7 km



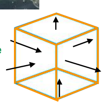
27

Resolution of OCMs

Horizontal resolution: 3.7 km

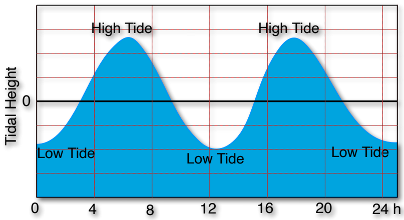


6 velocities on the sides
1 temperature in the middle
1 salinity in the middle
1 water depth



Resolution of OCM

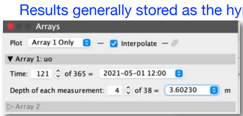
Temporal resolution
To resolve tides, at least a resolution of 3 hours



Existing OCM data, e.g. on data portals, rarely saved with temporal resolution greater than 1 day

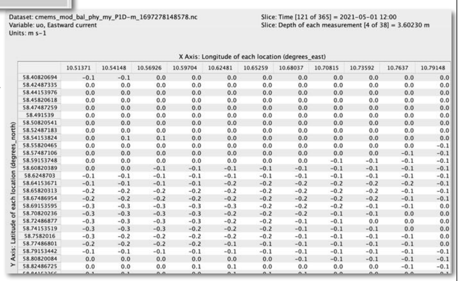
Results from an OCM

Results generally stored as the hyper-file format: *Network Common Data Form (NetCDF)*



Data matrix of east current; one "slice" for each depth layer and time point

NetCDF readers, e.g. Panoply and scripts in:
R, Matlab, and Python



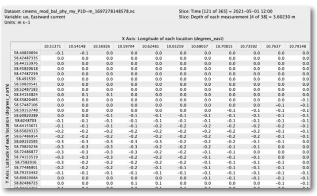
Dataset: cmems_mod_bal_phy_my_P1D-m_1697278148578.nc
Variable: uo, Eastward current
Units: m s-1

Slice: Time [121 of 365] = 2021-05-01 12:00
Slice: Depth of each measurement [4 of 38] = 3.60230 m

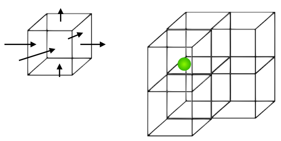
X Axis: Longitude of each location (degrees_east)
Y Axis: Latitude of each location (degrees_north)

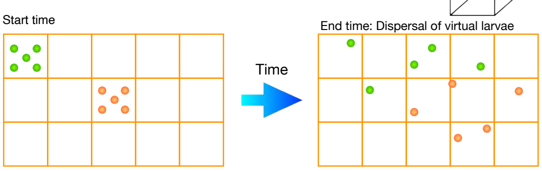
	10.51371	10.54148	10.56926	10.59704	10.62481	10.65259	10.68037	10.70815	10.73592	10.7637	10.79148
58.40820694	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
58.42487335	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
58.44153976	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
58.45820618	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
58.47487259	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
58.491539	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
58.50820541	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
58.52487183	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
58.54153824	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
58.55820465	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.1
58.57487106	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.1	-0.1
58.59153748	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.1
58.60820389	0.0	0.0	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1
58.6248703	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1
58.64153671	-0.1	-0.1	-0.1	-0.1	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.1
58.65820313	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.1	-0.1
58.67486954	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.1	-0.1
58.69153595	-0.3	-0.3	-0.3	-0.3	-0.3	-0.2	-0.2	-0.2	-0.1	-0.1	0.0
58.70820236	-0.3	-0.3	-0.3	-0.3	-0.2	-0.2	-0.2	-0.1	-0.1	0.0	0.0
58.72486877	-0.3	-0.3	-0.3	-0.3	-0.2	-0.2	-0.1	-0.1	0.0	0.0	0.0
58.74153519	-0.3	-0.3	-0.3	-0.2	-0.2	-0.2	-0.1	-0.1	-0.1	0.0	0.0
58.7582016	-0.3	-0.2	-0.2	-0.2	-0.2	-0.2	-0.1	-0.1	-0.1	-0.1	0.0
58.77486801	-0.2	-0.2	-0.2	-0.2	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	0.0
58.79153442	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1
58.80820084	0.0	0.0	0.0	0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.1	-0.1
58.82486725	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	-0.1	-0.1

The biological part of biophysical models



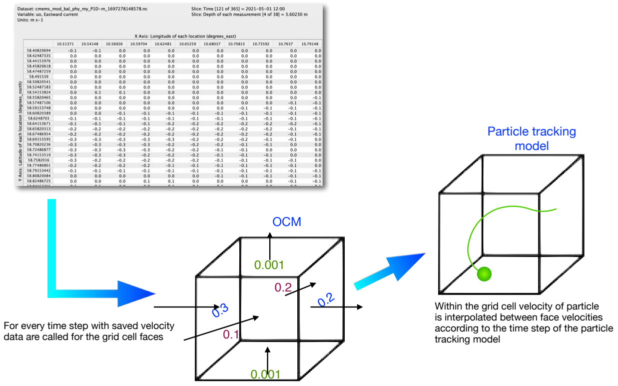
The OCM drives a **Lagrangian particle tracking** model that moves "particles" according to the velocity in the OCM





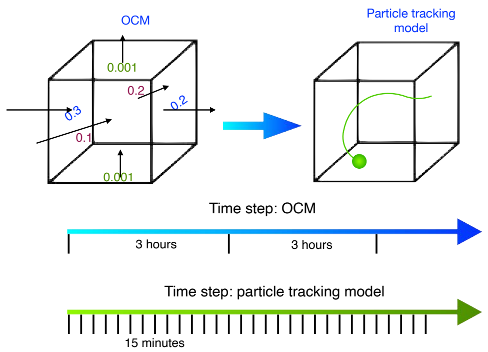
31

The biological part of biophysical models



32

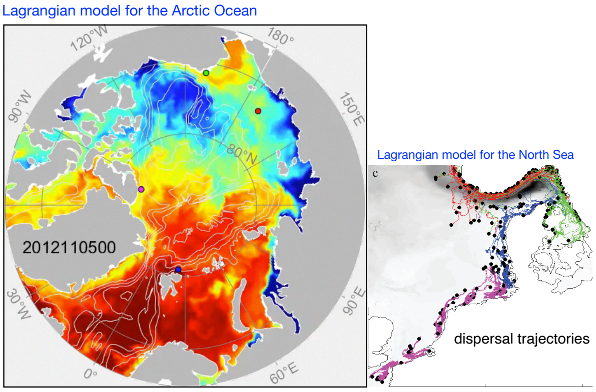
The biological part of biophysical models



33

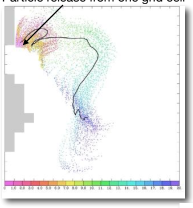
Example of Lagrangian particle model

Lagrangian model for the Arctic Ocean

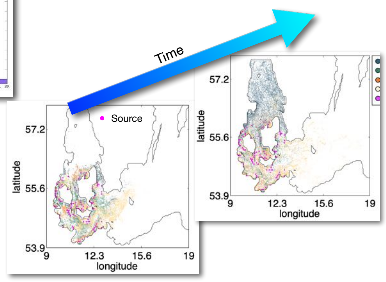


Coverage in space and time

Particle release from one grid cell



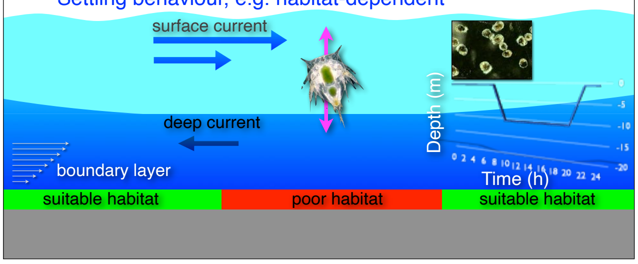
Usually a very large number of particles are released from multiple sources, at several release times



34

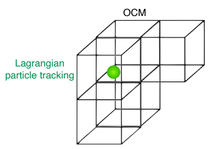
But larvae are not neutral, passive particles and not always

- Spawning time
- Drift depth - vertical behaviour
- Pelagic Larval Duration (PLD)
- Settling behaviour, e.g. habitat-dependent



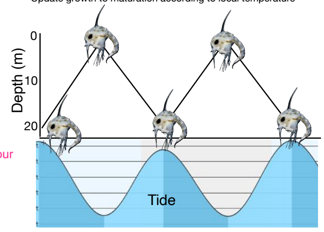
35

Individual Based Models (IBMs)



Every time step in the particle tracking model:

- **Update velocity in 3 dimensions and move particle**
- Keep a specific depth, change depth with time of day or with maturation
- Keep track of PLD - how long time to drift (time or size goal)
- Search for suitable settling substrate in an included habitat map
- Update growth to maturation according to local temperature

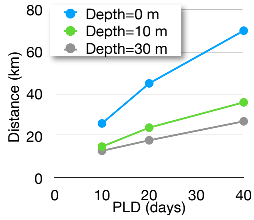


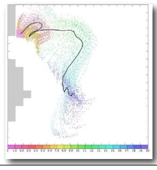
Usually poor knowledge about larval behaviour

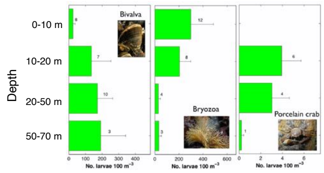
36

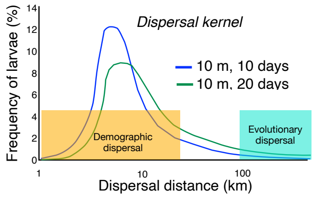
37

Effect of PLD and drift depth





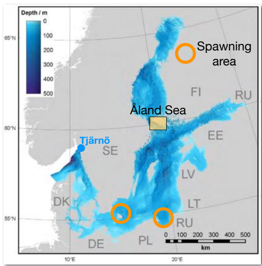




38

Biophysical modelling - an example

Dispersal of cod larvae



Bergström *et al.* (in manus)

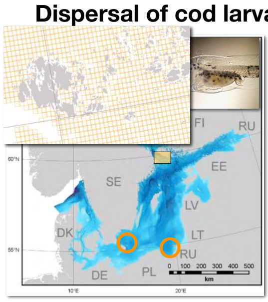
- Is there dispersal from present spawning areas?
- Is there a signature of isolation with local spawning?
- Are there local adaptations, e.g. in egg buoyancy?

Task for biophysical model: Model larval dispersal from spawning areas and test for connectivity with Åland Sea

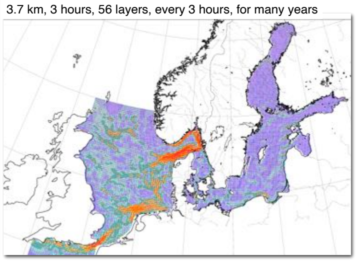
39

Biophysical modelling - the OCM

Dispersal of cod larvae



OCM: NEMO-Nordic
3.7 km, 3 hours, 56 layers, every 3 hours, for many years



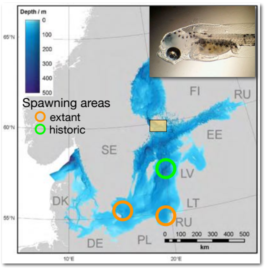
- Is the resolution relevant?
- Are the boundaries sufficiently distant?
- Are relevant oceanographic processes included?
- How many years are available, and with what temporal resolution?

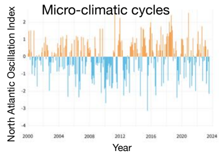
40

Biophysical modelling - the particle tracking model

Model design

- Where to release particles from source areas?
- Time of release?
- Multiple release times - to include daily variation, annual variation?
- What PLD is relevant?
- Are particles neutral or should they have some control of their vertical position, which requires an IBM?
- How many particles to release?
- How to analyse the results?

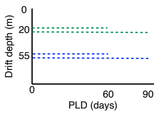




41

Biophysical modelling - the particle tracking model

Vertical position and PLD



Spawning time
May, June, July, August

Years included
1995, 1996, 1998, 2000, 2001, 2002
spanning the range of NAO index

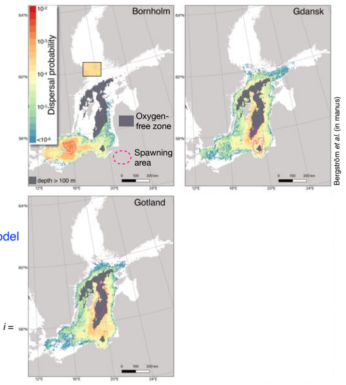
Selection of Lagrangian particle tracking model
TRACMASS (de Vries & Döös 2001)

Number of particles
23 million particles

Results

1. Dispersal probability from spawning area to grid cell i = number of particles ending up in grid cell i / total released particles from spawning area

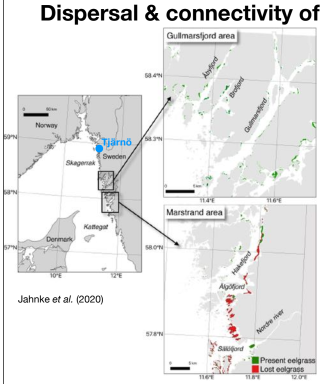
2. Proportion of particles from spawning area ending in the target area (yellow)



42

Biophysical modelling - an example

Dispersal & connectivity of eelgrass (*Zostera marina*)

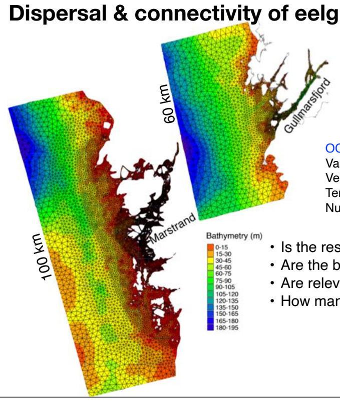


Jahnke *et al.* (2020)

- Which meadows are most important to protect for metapopulation connectivity from a network perspective
- Which are the most valuable extinct meadows to restore?

Biophysical modelling - the OCM

Dispersal & connectivity of eelgrass (*Zostera marina*)



OCM: DHI MIKE 3
Variable grid cell size: 100 m - 1 km
Vertical resolution: 1 m
Temporal resolution: 1 hour
Number of years: 5

- Is the resolution relevant?
- Are the boundaries sufficiently distant?
- Are relevant oceanographic processes included?
- How many years are available?

43

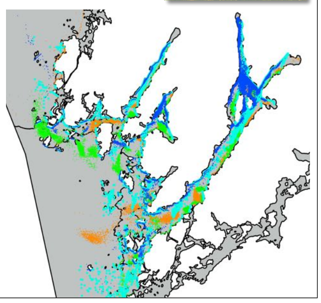
Biophysical modelling - the particle tracking model

Task:

- Model seed dispersal between 140 and 237 eelgrass meadows for the two areas, respectively.
- From dispersal probability construct connectivity matrices and analysing the network in search for essential meadows

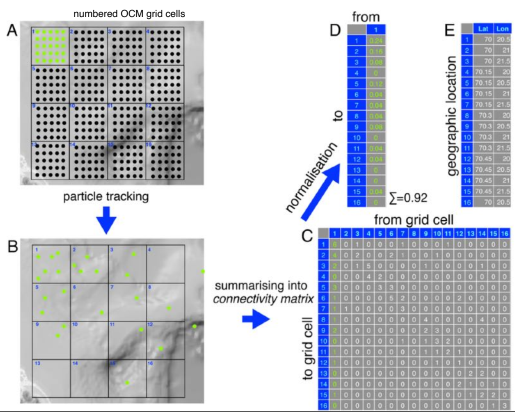
Model design

- Release particles from all meadows
- Release on 7 occasions during August during 5 years
- Drift time (PLD): 1-30 days
- 15 million particles released in total
- Lagrangian particle tracking model tool: DHI MIKE ECO-lab



44

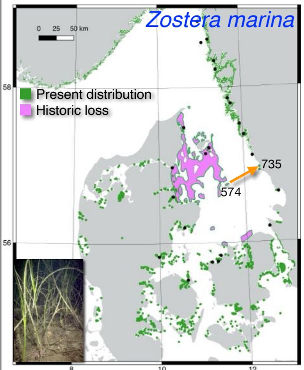
Biophysical modelling - the connectivity matrix (CM)



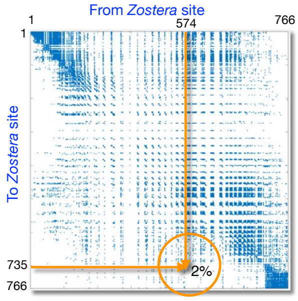
45

46

Connectivity matrix and habitat

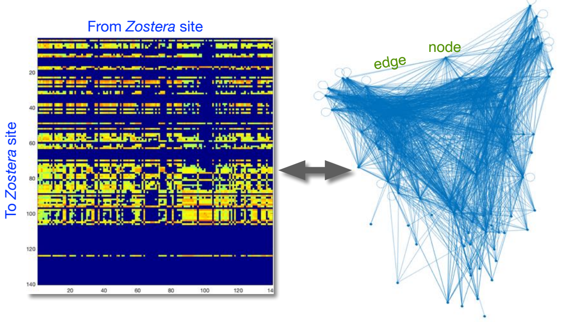


Jahnke *et al.* (2018)



47

Analysis of the connectivity matrix



A node in the seascape may represent a habitat patch (meadow), a local population, a spawning area, an MPA, or simply a model grid cell

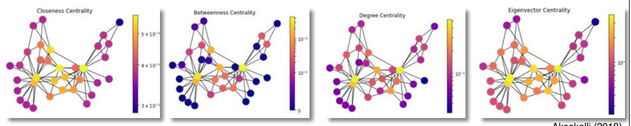
48

Analysis of the connectivity matrix

Network analysis (Graph theory)

- What nodes are most important for the metapopulation connectivity?
- What nodes are most important as sources of seeds or larvae?

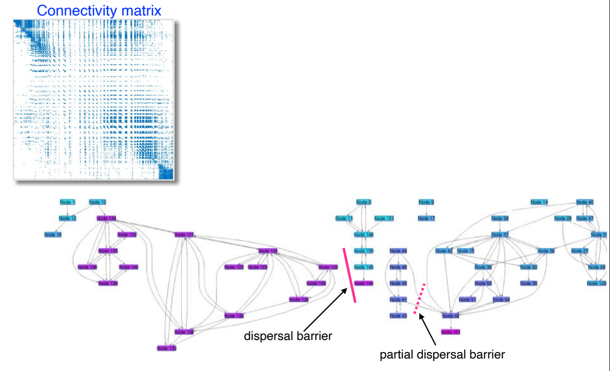
Some network metrics



Aksakalli (2018)

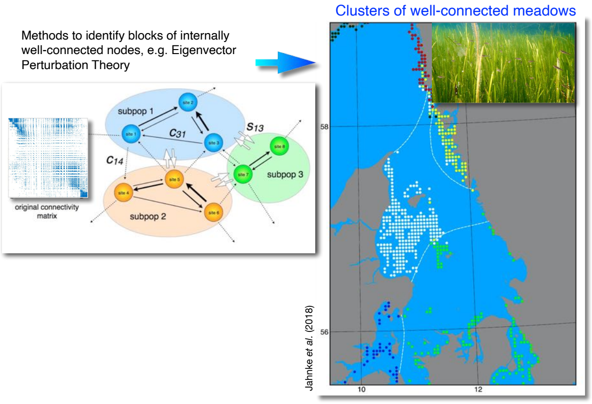
Analysis of dispersal barriers

49



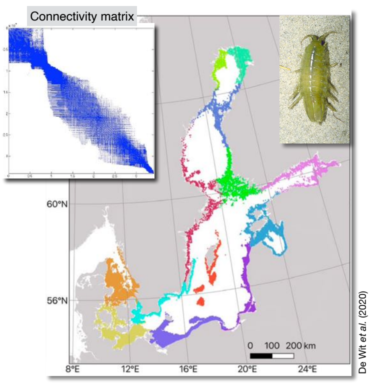
Analysis of dispersal barriers

50



Analysis of dispersal barriers

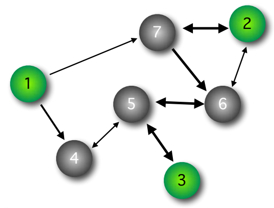
51



52

Biophysical model to estimate multi-generation connectivity - stepping-stone dispersal

Biophysical models predict **single** generation dispersal, which may not correlate well with gene flow that is the result of dispersal over many generations



Sampled site

Not sampled site

CM (from / to)

	1	2	3
1	0	0	0
2	0	0	0
3	0	0	0

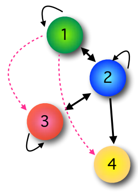
	1	2	3	4	5	6	7
1	0	0	0	0	0	0	0
2	0	0	0	0	0	0.2	0.5
3	0	0	0	0	0.4	0	0
4	0.3	0	0	0	0.2	0	0
5	0	0	0.5	0.3	0	0.4	0
6	0	0.2	0	0	0.4	0	0.4
7	0.2	0.5	0	0	0	0	0

53

The multi-generation connectivity matrix

Gene flow occurs over stepping-stone dispersal across generations

Demographic dispersal



Connectivity matrix for 1-generation dispersal (from / to)

0.3	0.3	0	0
0.7	0	1	0
0	0.3	0	0
0	0.4	0	0

By multiplying the connectivity matrix with itself each generation, the probability for all possible steps are summed

0.3	0.3	0	0
0.7	0	1	0
0	0.3	0	0
0	0.4	0	0

*

0.3	0.3	0	0
0.7	0	1	0
0	0.3	0	0
0	0.4	0	0

=

0.3	0.1	0.3	0
0.2	0.5	0	0
0.2	0	0.3	0
0.3	0.4	0	0

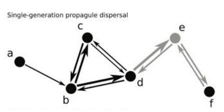
$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

c_{41} = 0*0.3 + 0.4*0.7 + 0*0 + 0*0 ≈ 0.3

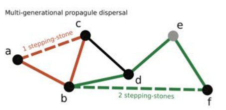
Multiple demographic dispersal events with reproduction will lead to gene flow on a larger spatial scale

54

The multi-generation connectivity matrix



from Boulanger *et al.* (2020)



CM (from / to)

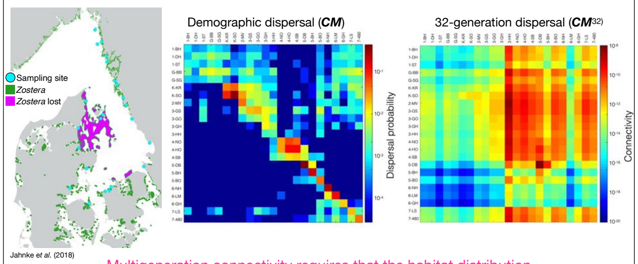
	a	b	c	d	e	f
a	0	0	0	0	0	0
b	0.1	0	0.3	0.3	0	0
c	0	0.5	0	0.1	0	0
d	0	0.5	0.1	0	0.1	0
e	0	0	0	0.5	0	0.3
f	0	0	0	0	0.1	0

CM*CM*CM=

	aa	bb	cc	dd	ee	ff
aa	0	0	0	0	0	0
bb	0.03/0	0.3/0.03	0.09/0.03	0.03/0.3	0.003/0.03	0.009/0
cc	0.005/0.05	0.16/0.05	0.16/0.18	0.036/0.15	0.015/0.01	0.003/0.03
dd	0.005/0.05	0.18/0.05	0.036/0.15	0.23/0.21	0.024/0.02	0.03
ee	0.025/0	0.025/0.25	0.07/0.05	0.12/0	0.08	0.02/0
ff	0	0.025/0	0.005/0	0.05	0.008/0	0.03

55

The multi-generation connectivity matrix



Multigeneration connectivity requires that the habitat distribution (with all stepping stones) is included

56

Strength and weakness of biophysical modelling

Strengths

- High coverage in space and time
- Potentially less expensive
- No requirement for samples, sometimes from inaccessible areas
- Can be adapted to many species
- Suitable for modelling management actions, e.g. protected areas
- Can suggest areas for genetic investigations, e.g. putative dispersal barriers
- May allow for multi-generation projections
- Can be used to project future dispersal

Weaknesses

- Only applicable to species showing passive dispersal
- Still low spatial resolution of OCMs, especially along complex coasts
- Not all oceanographic mechanism included in models
- Poor knowledge about vertical behaviour of larvae
- Commonly gives only potential dispersal rather than realised dispersal

57

Biophysical model predictions and genetic/genomic scape patterns

So, do biophysical models seem to explain population genetic/ genomic patterns in the seascape?

Marlene Jahnke will this evening talk about if and when biophysical models can explain patterns of genetic/genomic data sampled in the seascape

58