**Title:**

Data presented in Zhao et al. 2024, ‘Boosting efficiency of mussel spat collection for ecological sustainability: identifying critical drivers and informing management’ in the *Journal of Applied Ecology*.

**Creators:**

Zhiyuan Zhao, Jacob J. Capelle, Jaco C. de Smit, Theo Gerkema, Johan van de Koppel, Lin Yuan, Tjeerd J. Bouma

**License:**

CC-0

**These files include the raw data used to create figures in the manuscript, organized as follows:**

1. The 11-year dataset

2. Field experiments

a) Larval abundance

b) Spat settlement

**The 11-year dataset**

* *Data*

“Seed mussel harvest data.xlsx”

* *Figures*

“Fig 2.png”

“Fig 4.png”

“Fig 5.png”

“Fig 6.png”

* *Description*

From 2006, the local management agency of the Dutch Wadden Sea launched investigations targeting mussel farmers who deployed SMCs on government-leased lots. To keep their permit, farmers were obliged to complete separate questionnaires during the deployment and harvesting periods each year, reporting on *i*) the location, date, type, size, etc. of deployed SMCs, and *ii*) the location, date, batch, size, harvest, etc. of harvested SMCs. The efficiency of SMCs was calculated using collected information, expressed as biomass per unit of substrates (i.e., rope and net). These data can be found in the data frame within the following file:

* “Seed mussel harvest data.xlsx”.

This file includes one sheet, which shows the data of SMCs in 63 plots across eight locations from 2011 to 2021. The columns include “Year”, indicating the year of data collection; “Location”, showing the installation location of the SMCs; “Plot”, indicating the commercial plot where the SMCs were installed; “InstallationDate”, showing the date when the SMCs were installed; “HarvestDate”, showing the date when the SMCs were removed; “InstallationPeriod\_Days”, indicating the period the SMCs were installed, measured in days; “Substrate”, showing the type of installed SMCs, including rope-based and net-based; “RopeLength\_NetArea”, indicating the size of the installed SMCs, with rope-based SMCs shown as rope length (m) and net-based SMCs shown as net area (m²); “Harvest\_mt”, indicating the final harvest amount of the SMCs in metric tons (mt); “Harvest\_kg”, indicating the final harvest amount of the SMCs in kilograms (kg); “BiomassHarvest\_rope.net\_Kg.m”, showing the harvest efficiency of the installed SMCs, measured in kg m-1; “Starfish”, indicating the starfish abundance at the SMC installation site, divided into 6 levels, from 0 to 5. The above data can be used to create “Fig 2.png”.

This file also records environmental conditions during the SMCs installation period. These data were obtained from public data platforms. The columns include “Average\_temperature\_D”, indicating the average annual temperature in degrees Celsius; “Max\_temperature\_D”, indicating the highest daily average temperature of the year in degrees Celsius; “Min\_temperature\_D”, indicating the lowest daily average temperature of the year in degrees Celsius; “Average\_windspeed\_m.s”, indicating the average annual wind speed in meters per second (m s-1); “Max\_windgust\_m.s”, indicating the maximum wind gust speed in meters per second (m s-1); “MaxWaveHeight\_m”, indicating the maximum wave height in meters (m); “MeanWaveHeight\_m”, indicating the mean wave height in meters; “Salinity”, indicating the average annual salinity; “Velocity\_Vector”, indicating the average current velocity in meters per second (m s-1); “Chlorophyll”, indicating the average annual chlorophyll-a concentration. The above data can be used to create “Fig 4.png”.

The above two sets of data were integrated and applied to machine learning, resulting in “Fig 5.png” and “Fig 6.png”.

**Field experiments - Larval abundance**

* *Data*

“Larvae.csv”

* *Figure*

“Fig 3a.png”

* *Description*

To identify the effect of larval abundance on SMCs efficiency, continuous monitoring was conducted at four locations in the Dutch Wadden Sea throughout the mussel reproduction season (March to June) from 2017 to 2020. The related data can be found in the data frame “Larvae.csv”. This file includes one sheet. The columns include “Year”, which indicates the year the experiment was conducted; “Week”, which describes the number of weeks since the start of the experiment; “Date”, which records the date of each monitoring; “Location”, indicating the location of the experiment; and “mussel larvae”, which records the abundance of larvae detected, with the unit being ind/100L. This data can be used to produce the left subplot (a) in “Fig 3.png”.

**Field experiments - Spat settlement**

* *Data*

“settlement.csv”

* *Figure*

“Fig 3b.png”

* *Description*

To identify the effect of spat settlement rate on SMCs efficiency, continuous monitoring was conducted at four locations in the Dutch Wadden Sea throughout the mussel reproduction season (March to June) from 2017 to 2020. The related data can be found in the data frame “settlement.csv”. This file includes one sheet. The columns include “Year”, which indicates the year the experiment was conducted; “Period” describes the round of the experiment; “Date” records the date of each experiment (start or end); “in.uit” marks the beginning or end of each round, with “in” indicating the start and “uit” indicating the end; “Location” specifies the experiment’s location; “settlement” records the detected abundance of spat, measured in individuals per 0.15 square meters (ind / 0.15m²). This data can be used to produce the right subplot (b) in “Fig 3.png”.