# Summary on the pool data for the Daphnia metapopulation.

### Compiled by Dieter Ebert

Version, January 2018

## 1. People involved

The following people have been involved in the data collection and analysis to an extent that they deserve being co-authors on papers resulting from analysis of the data to which they contributed:

Maxime Dubart1,+, Patrice David1, Frida Ben-Ami2,3, Christoph R. Haag1,3, Peter D. Fields3,5, V. Ilmari Pajunen3,4 & Dieter Ebert3,5

**Addresses:**

1 CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

2 School of Zoology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 6997801, Israel

3 Tvärminne Zoological Station, University of Helsinki, J.A. Palménin tie 260, 10900 Hanko, Finland

4 Department of Biological and Environmental Sciences, University of Helsinki, P. O. Box 65, Viikinkaari 1, Helsinki 00014, Finland

5 University of Basel, Department of Environmental Sciences, Zoology, Vesalgasse 1, 4051 Basel, Switzerland

The following people have been involved in the field work for data collection:

Jürgen Hottinger, Andrea Cabalzar, Mikko Lehto, Jennifer Lohr, Mathew Hall, David Preiswerk, David Duneau, Katharina Ebert, Gleb Ebert, Anna Marcelino, Y. Haag, E. Haag, C. Liautard-Haag, C. Reisser, C. Molinier, E. Hürlimann, and C. Mills.

## 2. Data available on metapopulation ecology

#### 2.1. Presence absence data on 3 Daphnia species

These are the core data of this project. Presence/absence data for *D. magna, D. pulex* and *D. longispina* are available for nearly every spring and summer sample since summer 1982. Spring 1993 are missing (NA were included in the data file for this seasons).

Datafile: Daphnia\_dynamics\_1982\_2017\_2.csv

##### Variables:

"year" Year when data were recorded (integer)

"sample" First (1, May/early June)) or second (2; late July/ August) sample in the year (integer)

"island" Island code (string)

"pool" Pool number (string)

"magna" Presence/absence (1/0) for Daphnia magna

"longispina" Presence/absence (1/0) for D. longispina

"pulex" Presence/absence (1/0) for D. pulex

“author” Initials of person collecting the data (VIP: Ilmari, DE: Dieter, ...)

"poolname" island + "-" + pool

"water" Presence/absence (y/n) of water in pool during screening. For data before 2009, this variable is always "y".

#### 2.2. Presence absence data on fish in rock pools.

Occasionaly we see that pools are populated by sticklebacks. This means usually that the Daphnia cannot survive. We record fish in the "remack" section of the field protocol. Here they are converted to a variable called "fish". 1 means that fish where observed, no entry means no fish. Thus the file only contains data of fish observations.

Records about fish were only collected since 2009.

Current datafile: 2009\_2017-2\_fish\_ver2.csv

##### Variables:

“day” Date of sampling

“month” Month of sampling

"year" Year when data were recorded (integer)

"sample" First (1, May/early June)) or second (2; late July/ August) sample in the year (integer)

"poolname" island + "-" + pool

"island" Island code (string)

"pool" Pool number (string)

“remark” remark as taken from field protocol

“fish” 1 for presence of fish. Anything else means no fish. No record means also no fish.

##### Notes:

These are dynamic data and are only present since 2009. They cannot be used as dynamic variable for an analysis of the entire time period from 1982. They can be helpful for an analysis of pool characteristics (fish are more often seen in pools closer to the sea) to describe the impact of the sea on the pool environment.

#### 2.3. pH and water conductivity

pH and water conductivity data have been collected at irregular occasions since 1998. All data are compiled in one data file. Data from

1998 or 1999: Data are from the first screening of D. magna pools in the core area by the Ebert group. Data were only collected for pools with D. magna (about 20 % of all pools in core area).

2005 and 2006: VIP collected in 2005 and 2006 pH and conductivity data for most pools. In some cases VIP reported the average across 2 to 5 measures, by this taken out some of the individual variation.

2012 – 2017: Data collected since 2012 in irregular intervals. In some years only a subset of pools where samples, because we did not have enough time to finish all pools. Focus was placed in these cases on pools for which we had less data so far. The dataset has now mostly 5 or more pH and conductivity measures per pool.

Current datafile: MetapopData\_pH\_conductivity\_1998\_2017.csv

##### Variables:

“run” collecting sequence of pools (variable has no meaning here)

“day” Date of sampling

“month” Month of sampling

"year" Year when data were recorded (integer)

"sample" First (1, May/early June)) or second (2; late July/ August) sample in the year (integer)

"island" Island code (string)

"pool" Pool number (string)

“conduct\_uS” conductivity of water in micro Siemens

“pH” pH of water

“author” person collecting data

##### Notes:

Water chemistry (pH and conductivity) was check in irregular intervals. One may use the data to create a static variable (mean or median) of the repeated measures. Furthermore, one may include the variance (or Coeff. of var.) of the data across time for a pool, as a measure of pool stability.

#### 2.4. Plant coverage

The presence of vascular plants in rock pools has been first recorded by Florian. He found that it influences the time to desiccation of a pool. "plants suck pools dry!". Florian's data are in the file: "pools.txt". He describes his data as:

plants\_F Presence (1) or absence (0) of plants (moss, grass, bushes, trees) in the pool/around the pool with direct contact to pool that might increase evaporation. Recorded May/June 2006. (Altermatt) These are not used here, as they are not quantitative.

Since summer 2013 we screen all pools every summer for plant cover. Plant cover is an approximation of the pool surface area covered by plants. It is judged by eye during field sampling. Even dry pools can easily judged for plant cover. This works well as the plants remain viable for rather long periods even in dry conditions. The following categories are used: "plants gives the approximate coverage in %, "plants\_rank" is coverage expressed as an ordinal variable with ranks from 1 to 7.

|  |  |  |
| --- | --- | --- |
| **plants** | **plants-rank** | **comment** |
| 0 % | 1 | no plants present |
| 5 % | 2 | few plants (e.g. 1 bushel of gras in a corner of the pools) (1 to 5 % coverage) |
| 25 % | 3 | 6 to 25 % coverage |
| 50 % | 4 | 26 - 50% coverage |
| 75 % | 5 | 51 - 75 % coverage |
| 95 % | 6 | 76- 95 % coverage. A little bit of free water left. Daphnia could be present in principle, but is unlikely |
| 100 % | 7 | totally covered. No free water left. Daphnia cannot be present. |

Data are available only for the summer sampling, as in spring plants are hardly out yet. The dataset has for 2013 to 2017 no missing values.

Current datafile: plantcover\_2003\_2017.csv

##### Variables:

“day” Date of sampling

“month” Month of sampling

"year" Year when data were recorded (integer)

"sample" First (1, May/early June)) or second (2; late July/ August) sample in the year (integer)

"poolname" Name of island (redundant with the two following variables)

"island" Island code (string)

"plants" gives the approximate coverage in %

“plants\_rank” coverage expressed as an ordinal variable with ranks from 1 to 7

“water” whether pool had water at the time of sampling. These data are also in the file on Daphnia prersence/absence (see above).

##### Notes:

* Plant cover was check only since 2013. One my use the data to create a static variable (mean or median) of the repeated measures. One may also include the variance of the data across time for a pool, as a measure of pool stability.
* For 2013, water is NA in this datafile. These data are available, but have to be taken from the Daphnia presence/absence file. However, plant cover was estimated even in case the pool was dry, as the plants are still visible.

#### 2.5. Pool hard data

The file Pool\_hard\_data contains data on pool features that do not change over time (at least not in time scales considered in the current projects), such as pool dimensions, location relative to sea, etc. This file is update when new pools are included, new data are generated, and when inconsistencies show up.

Variables with \* can be ignored. These are variables used to calculate other variables. They are often incomplete. Other variables are used to keep track of missing or changing data (e.g. missing, miss\_comment).

Current datafile: Pools\_hard\_data\_vers12.csv

##### Variables:

|  |  |  |
| --- | --- | --- |
| run\* | variable indicating the sequence in which we visit the pools on the islands |  |
| poolname | island + '-' + pool (string) |  |
| island | island short name (string) |  |
| pool | pool number (string) |  |
| missing\* | here a 1 is placed when some data are missing, unreliable or otherwise odd for this pool. They need to be (re-)collected next time in the field. |  |
| miss\_comment\* | comment to previous variable |  |
| core | 1 if part of Ilmari Pajunen's core area, otherwise 0 |  |
| part\_of\_core\_since | most pools are part of the core metapop. since 1982. Pools included later have the year when we added into the dataset. |  |
| length\* | length of longest axis of pool (units: cm) |  |
| width\* | Width perpendicular to length at widest place (units: cm) |  |
| surface\_calculated\_DE\* | pool surface estimated from ((length \*width)/2) (units: square meter) |  |
| surface\_area\_VIP\* | pool surface calculated by Ilmari. Not clear how he estimated this exactly. (units: square meter) |  |
| surface\_area | pool surface combined from DE and VIP data. (units: square meter) |  |
| depth | pool depth at the deepest place in the pool (units: cm) |  |
| height\_aboveSea | meter above sea level. (Note, the sea level changes by about 1 m during the summer season (even more in winter). (units: cm) |  |
| sometimes\_submerged | pools very close to the sea maybe at times under sea water. Pools with observed submergence at least at one time have a 1. Otherwise zero. |  |
| distance\_toSea | Shortest distance from sea shore to pool (units: cm) |  |
| distance\_toTree\* | distance to the next tree. Was used formerly to estimate influence of plants on the pool. We use now other variable, e.g. plant coverage, pH and conductivity to estimate this. (units: meter). Problem: many islands have no trees. |  |
| watersheet | All pools with the same watersheet name are connected with each other. The following 5 variables describe how the water cascade was shaped. The watersheet is usually named after the pool on top of the cascade. (Note, some watersheet have 2 pools on top. Only one gives the name). |  |
| casc\_top | Top pool of a cascade |  |
| casc\_middle\_top | Intermediate level pool of a cascade |  |
| casc\_middle | Intermediate level pool of a cascade |  |
| casc\_middle\_bottom | Intermediate level pool of a cascade |  |
| casc\_bottom | Bottom pool of a cascade |  |
| catchment\_area\_FA\* | catchment area as estimated by GPS by Florian (units: square meter) |  |
| catchment\_area\_DE\* | catchment area as estimated by GPS and by hand by Dieter (units: square meter) |  |
| catchment\_area | compete catchment area list (compiled from the two previous variables). (units: square meter) |  |

##### Notes:

* These are all static (hard) data and do not change over the time periods considered here.
* pool\_volume: To estimate volume we assume an inverted pyramid: Volume = (surface \* depth)/3

#### 2.6. New pools and unsuitable pools

Unsuitable pools are considered being not available for colonisation at the time of sampling. This is a way to deal with the problem that when a new pool is defined, we do not know what was before we defined it. So we define the earlier sampling events as being unsuitable.

New pools come about in two ways:

1) when a pool splits into 2 ppols (e.g. because of a lowering of the water level (G-37) or because a land-bridge forms and splits the pool (N-28, M-69)) the main part of the pool keeps the original name and the newly formed part will have a new name, typically with an A added (eg N-28 split into N-28 and N-28A). Since the new pool will have no previous record it is defined as being "unsuitable" in all previous years.

2) when a pool is newly included in the dataset (eg FO-18A, FO-19A, FS-6A) it will be declared unsuitable for all previous years. This might be because pools suddenly showed up (e.g. after removal of plants, or a tree fall over).

As a consequence of this logic, a pools is defined and listed as unsuitable any time before it shows up for the first time in the dataset. Afterwards it is defined as being suitable.

Current datafile: 1982\_2017\_unsuitable\_pools\_only\_ver2

##### Variables:

"year" Year when data were recorded (integer)

"sample" First (May/June) or second (July/ August) sample per year (1/2)

"island" Island code (string)

"pool" Pool number (string)

"pool\_suitable" n for pools not being suitable. Only unsuitable pools are listed. All other pools are considered suitable in the entire sampling period from 1982 to today.

##### Notes:

This variable was necessary to be created because we need a way to declare what was the state of a pool before we newly included it into our regular sampling rounds in spring and summer. These are all pools within the core sampling area! The variable “pool\_suitability” does not take any other information into account, other than if the pool was part of the sampling scheme before it was included in the seasonal survey or not.

A way to deal with them is to simply use the variable "pool\_suitable" as it is. Alternatively, exclude all pools from the entire dataset that were added since 2009. Excluding these pools would not make much of a difference as they are few and they are only present since at most 8 years. These pools are (sorted by the year they were added with the youngest addition being first: M-69A was added in 2016, SK-20A was added in 2009):

* M-69A
* SK-26A
* FO-19A
* FS-6A
* M-30A
* N-28A
* SK-40A
* M-14B
* FO-18A
* G-111A
* G-12A
* G-122A
* G-37A
* G-37B
* G-50A
* N-15A
* N-24A
* N-85A
* LA-18A
* SK-20A

#### 2.7. Pool coordinates

Coordinate format is the standard GPS data format with latitude and longitude as one number with typically 6 or more digits after the dot.

Correction of coordinates was done after plotting GPS coordinates and comparing the plots with areal photographs. Unprecise GPS coordinates are a consequence of the low precision of the GPS device by the time of data collection.

Current datafile: Pools\_coordinates\_2017\_vers7.xls

##### Variables:

name name of pool

island name of island

pool number of pool

core part of core dataset (1 for yes)

latitude\_TZ coordinates as collated by Thomas Zumbrunn

longitude\_TZ coordinates as collated by Thomas Zumbrunn

latitude\_2012 coordinates as collected by D. Ebert starting in 2012

longitude\_2012 coordinates as collected by D. Ebert starting in 2012

latitude\_mean average between estimate of Th. Zumbrunn and D. Ebert

latitude\_correction correction of coordinate if necessary.

latitude\_corr corrected value

longitude\_mean average between estimate of Th. Zumbrunn and D. Ebert

longotide\_correction correction of coordinate if necessary.

longitude\_corr corrected value

##### Note: For any project on this metapopulation only the variables **"latitude\_corr"** and **"longitude\_corr"** are to be used. The other variables are only of historic value.

#### 2.8. Island\_measures

Pool are located on islands of the Tvaerminne archipelago. The following datafile gives estimates regarding island features, such as island size.

Current datafile: Island\_measures\_vers4.csv

##### Variables:

island\_name: not all islands have real names, therefore the code is used in some cases.

island: code of islands. This abbreviation (code) is used in all other datafiles.

number\_pools\_approx: This is the number of pools on the island.

island\_area\_sq\_meter: area of an island in square meters.

island\_perimeter\_meter: perimeter of an island in meters.

group\_100: groups of islands into sets with less than 100 meter distance between islands

group\_200: groups of islands into sets with less than 200 meters distance between islands

## 3. Other data related to metapopulation

### 3.1. Data to estimate sampling error

#### 3.1.1. Sampling error of Daphnia sampling

During the time period from 2009 to 2012 the entire metapopulation was sampled by VIP and the DE group within 2 weeks of each other. This allows to estimate the error if recording the presence and absence of Daphnia in the pools. To estimate the degree of miss sampling, the old data files where reconstructed with the aim to compare VIPs and our data side by side. Only pools should be included here that were sampled by VIP and the Ebert group in the same sampling time window.

Current datafile: VIP\_DEdata5.csv

##### Variables:

"poolname" island + "-" + pool

"year" Year when data were recorded (integer)

"sample" First (1, May/early June)) or second (2; late July/ August) sample in the year (integer)

"island" Island code (string)

"pool" Pool number (string)

"water\_VIP" Presence/absence (y/n) of water. For VIP this is always "y".

"magna\_VIP" Presence/absence (1/0) for Daphnia magna by VIP

"pulex\_VIP" Presence/absence (1/0) for D. pulex by VIP

"longispina\_VIP" Presence/absence (1/0) for D. longispina by VIP

"water\_DE" Presence/absence (y/n) of water, as recorded by DE.

"magna\_DE" Presence/absence (1/0) for Daphnia magna by DE + group

longispina\_DE Presence/absence (1/0) for Daphnia longispina by DE + group

pulex\_DE Presence/absence (1/0) for Daphnia pulex by DE + group

##### Notes:

The current datafile contains the data for 2009 to 2012.

## 5. Overview tables

### 5.1. Overview of ecological variables of pools.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Recording window** | **Dynamic or static** | **Data type** | **Expected correlation with marine impact** | **Expected corr. with habitat persistence** | **Author** | **Comment** |
| fish | 2009 – ongoing | dynamic | 0/1 | positive | negative | DE |  |
| holds\_no\_water | 2009 – ongoing | dynamic | 0/1 | - | - | DE |  |
| conduct\_uS | 1998, 2005, 2006, 2012 – ongoing | dynamic | real number | positive | - | DE, other |  |
| pH | 1998, 2005, 2006, 2012 – ongoing | dynamic | real number | positive | - | DE, other |  |
| Plants, plants\_rank | 2013 – ongoing | dynamic | integer | negative | negative | DE |  |
| meanhydro | 1982-2006 | static | real number | - | positive | FA | One may generate a dynamic version of this variable |
| meandesi | 1982-2006 | static | real number | - | negative | FA | One may generate a dynamic version of this variable |
| mediandesi | 1982-2006 | static | real number | - | negative | FA | One may generate a dynamic version of this variable |
| medianhydro | 1982-2006 | static | real number | - | positive | FA | One may generate a dynamic version of this variable |
| surface\_area | NA | static | real number | - | positive | DE/VIP |  |
| depth | NA | static | real number | - | positive | DE/VIP |  |
| height\_aboveSea | NA | static | real number | negative | - | DE |  |
| sometimes\_submerged | NA | static | real number | positive | negative | DE |  |
| distance\_toSea | NA | static | real number | negative | - | DE |  |
| catchment\_area | NA | static | real number | - | positive | FA/DE |  |

### 5.2. Overview of other variables of pools.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Recording window** | **Dynamic or static** | **Who recorded** | **Comment** |
| pool\_suitable | 1982 - ongoing | - | DE |  |
| watersheet | - | static | DE |  |
| casc\_top | - | static | DE |  |
| casc\_middle\_top | - | static | DE |  |
| casc\_middle | - | static | DE |  |
| casc\_middle\_bottom | - | static | DE |  |
| casc\_bottom | - | static | DE |  |
| latitude, longitude | - | static | DE |  |
| island\_area\_sq\_meter |  | static | DE |  |
| island\_perimeter\_meter |  | static | DE |  |
| group\_100 |  | static | DE |  |
| group\_200 |  | static | DE |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

### 5.3. Climate variables

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Time period** | **Data file** | **resolution** |
| Temperature | 1970 - 2017 | temp\_rain\_TZS\_1970\_2005.xls  temp\_rain\_TZS\_2006-2017.xlsx | Average, min and max per day |
| Percipitation and snow fall | 1970 - 2017 | temp\_rain\_TZS\_1970\_2005.xls  temp\_rain\_TZS\_2006-2017.xlsx | Average, min and max per day |
| WindSpeed | 2006 - 2018 | WindSpeed\_Russaro\_2006-2018.csv | One measure per hour (might be average per hour?) |
|  |  |  |  |