# Использование программной реализации модели гемиклональных популяционных систем

Для настройки серии экспериментов используются файл hps.bat и файлы, содержащиеся в папке inputs и её подпапках.

## Моделирование серии экспериментов на одном начальном состоянии

### hps.bat

Это текстовый файл, который содержит команды для запуска Java-приложения. Здесь же указаны аргументы, которые будут переданы в программу.

Содержание этого файла можно разделить на то, что не касается настройки моделирования (неинтересные) и то, что его касается (**интересные**):

"c:\Program Files\Java\jre8\bin\java.exe" -Xms256M -Xmx768M -jar hps-2.jar  
**-y 100 -e 1..5 -p 1..40 --name “Modeling name” --statistic “only\_genotypes”**

**-y 100** задает число моделируемых лет.

**-e 1..5** задает диапазон моделируемых экспериментов. Используется именно диапазон, а не одно число, ввиду небольших проблем с именованием файлов при запуске на кластере. При запуске на одной машине, имеет смысл первым числом всегда указывать 1.

**-p 1..40** задает диапазон моделируемых начальных состояний. Каждое из начальных состояний имеет номер. Этот аргумент помогает ограничить диапазон моделируемых начальных состояний, что позволяет легко распределить моделирование серии экспериментов по нескольким машинам.

**--name “..”** позволяет определить имя папки с выходными данными для серии экспериментов. Все выходные данные сохраняются в подпапку папки outputs/. Например, если указать **--name “Test-RR-population”**, то все выходные данные будут сохранены в папки outputs/**Test-RR-population/settings** (вычисленные настройки для моделируемого начального состояния[[1]](#footnote-1)) и   
outputs/**Test-RR-population/statistic** (краткая сводка результатов моделирования для каждого начального состояния и развернутая статистика по каждому эксперименту).

**--statistic “..”** параметры вывода статистики. Здесь, перечисляя ключевые слова, можно указать этапы, после которых должна сохраняться численность популяции, можно запросить выводить лишь краткую статистику или статистику лишь по генотипам (без учета возрастов).  
only\_genotypes не выводить численность особей каждого конкретного возраста.  
only\_matures не учитывать неполовозрелых особей.  
only\_short не выводить подробную статистику по каждому эксперименту.  
after\_each сохранять численность популяции после каждого этапа моделирования.  
after\_movement сохраняться численность популяции после этапа перемещения.  
after\_growing сохраняться численность популяции после этапа взросления.  
after\_reproduction сохраняться численность популяции после этапа размножения.  
after\_competition сохраняться численность популяции после этапа конкуренции.  
after\_dieing сохраняться численность популяции после этапа вымирания.  
По умолчания после каждого этапа моделирования сохраняется подробная численность популяции.

### inputs/Habitat-X/

В папке inputs/ содержаться папки с описанием зон. Каждая inputs/Habitat-X/ хранит набор файлов, описывающих одну зону обитания.

### inputs/Habitat-X/Initial-composition.csv

Это табличный файл, в котором описан начальный состав особей в зоне Habitat-X.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| xRxR-0 | xRxR-1 | xRxR-2 | **...** | xRyR-8 |
| 1740 | 478 | 190 | **...** | 2 |

### inputs/Habitat-X/Neighbors.csv

Это табличный файл, в котором описаны относительные вероятности перехода особей из зоны Habitat-X в другие зоны. В шапке таблицы записываются имена соседних зон (имя зоны совпадает с именем папки, в которой храниться её описание).

|  |  |  |
| --- | --- | --- |
| Habitat-4 | Habitat-7 | Habitat-2 |
| 0.3 | 0.3 | 0.4 |

### inputs/Habitat-X/Posterity.csv

Здесь, вроде, вопросов не должно быть.

### inputs/Habitat-X/Resources.csv

Это табличный файл с одной заполненной ячейкой. В ней записано число ресурсов в зоне Habitat-X.

### inputs/Habitat-X/Scenario.csv

Это табличный файл, который описывает хронологию событий в эксперименте. Может быть указана иммиграция особей и параметры изменения числа ресурсов.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Year# | Resource + | Resource \* | Resource = | xRyR-5 | yL(xR)-5 |
| 10 | 20000 |  |  |  | 20 |
| 15 |  | 1.75 |  | 10 |  |
| 20 |  |  | 1500 |  |  |
| 30 | 8500 |  |  |  | 20 |
| 50 |  | 0.5 |  |  |  |

Последние две колонки, в данном примере, описывают иммиграцию. Resource + указывает число ресурсов, на сколько увеличиваются ресурсы зоны. Resource \* умножает ресурсы. Resource = устанавливает новое число ресурсов в зоне.

### inputs/Habitat-X/Viability.csv

Здесь, вроде, вопросов не должно быть.

### inputs/Dimensions.csv

При запуске моделирования серии экспериментов на одном начальном состоянии, этот файл должен быть пуст.

## Моделирование серии экспериментов на множестве начальных состояний

Чтобы разобраться в том, как описать множество начальных состояний, необходимо морально подготовиться к неестественным для пользователя рассуждениям – все не так тривиально, как хотелось бы. Для начала стоит понимать, что программа различает два принципиально разных типа запуска: одиночный (одно начальное состояние) и множественный (на множестве начальных состояний).

В первом случае *inputs/Dimensions.csv* будет пуст, а остальные входные файлы будут содержать лишь абсолютные, неизменяемые значения, которые и описывают одно начальное состояние. Во втором случае *inputs/Dimensions.csv* будет содержать информацию об изменяемых группах параметров, а остальные файлы будут содержать частично абсолютные значения, а частично изменяемые (переменные) значения. На основании заполненного *inputs/Dimensions.csv* и других входных файлах с изменяемыми параметрами будет генерироваться

Множество начальных состояний обычно состоит из таких точек конфигурационного пространства, которые отличаются друг от друга значениями одного или нескольких из сотен параметров. При проведении серии экспериментов мы исследуем влияние значения некоторого параметра на исход жизни популяции, – мы не пытаемся сравнить возможные конечные состояния для двух совершенно разных начальных состояний.

Программа реализована так, что ей на вход подается не список полных описаний каждого начального состояния, а шаблон состояния, в котором указаны изменяемые параметры.

### Простой пример

Здесь, сначала будут описаны общие рассуждения о том, как программе описать множество начальных состояний. После этого будут даны конкретные инструкции по заполнению входных файлов.

К примеру, мы хотим найти наиболее правдоподобные значения параметра «продолжительность жизни» для особей с генотипами xRxR и xRyR. Обозначим продолжительность жизни самок символом ***А***, а продолжительность жизни самцов – ***Б***. В таком случае, первые строки файла с параметрами выживаемости будут соответствовать следующему шаблону:

|  |  |  |
| --- | --- | --- |
|  | xRxR | xRyR |
| Продолжительность жизни | ***А*** | ***Б*** |
| Возраст 1-го нереста | 5 | 4 |
| Выживаемость | 0.5 | 0.45 |
| Возраст достижения выживаемости S | 5 | 3 |
| … | | |

Но программе необходимо знать в каком диапазоне, и с каким шагом следует перебирать значения этих параметров. Возьмем одинаковый диапазон для А и Б: [6, 12]. Шаг будет равен единице. Возможные значения параметров А и Б можно записать в таблицу или изобразить на диаграмме:

|  |  |
| --- | --- |
| ***А*** | ***Б*** |
| 6 | 6 |
| 6 | 7 |
| 6 | 8 |
| 6 | 9 |
| 6 | 10 |
| 6 | 11 |
| 6 | 12 |
| 7 | 6 |
| 7 | 7 |
| 7 | 8 |
| 7 | 9 |
| 7 | 10 |
| 7 | 11 |
| 7 | 12 |
| 8 | 6 |
| 8 | 7 |
| 8 | 8 |
| 8 | 9 |
| 8 | 10 |
| 8 | 11 |
| 8 | 12 |
| 9 | 6 |
| 9 | 7 |
| 9 | 8 |
| 9 | 9 |
| 9 | 10 |
| 9 | 11 |
| 9 | 12 |
| 10 | 6 |
| 10 | 7 |
| 10 | 8 |
| 10 | 9 |
| 10 | 10 |
| 10 | 11 |
| 10 | 12 |
| 11 | 6 |
| 11 | 7 |
| 11 | 8 |
| 11 | 9 |
| 11 | 10 |
| 11 | 11 |
| 11 | 12 |
| 12 | 6 |
| 12 | 7 |
| 12 | 8 |
| 12 | 9 |
| 12 | 10 |
| 12 | 11 |
| 12 | 12 |

Программа запустит серии экспериментов на 49-ти начальных состояниях.

Остается понять, как по средству заполнения входных файлов описать вышеуказанный шаблон.

Всё представляется как текст.

1. Если серия экспериментов проводиться над множеством начальных состояний, то входные файлы буду всё равно одни, только они будут содержать специальные конструкции, описывающие изменяемые параметры, которые требуют перевычисления для каждого конкретного начального состояния. [↑](#footnote-ref-1)