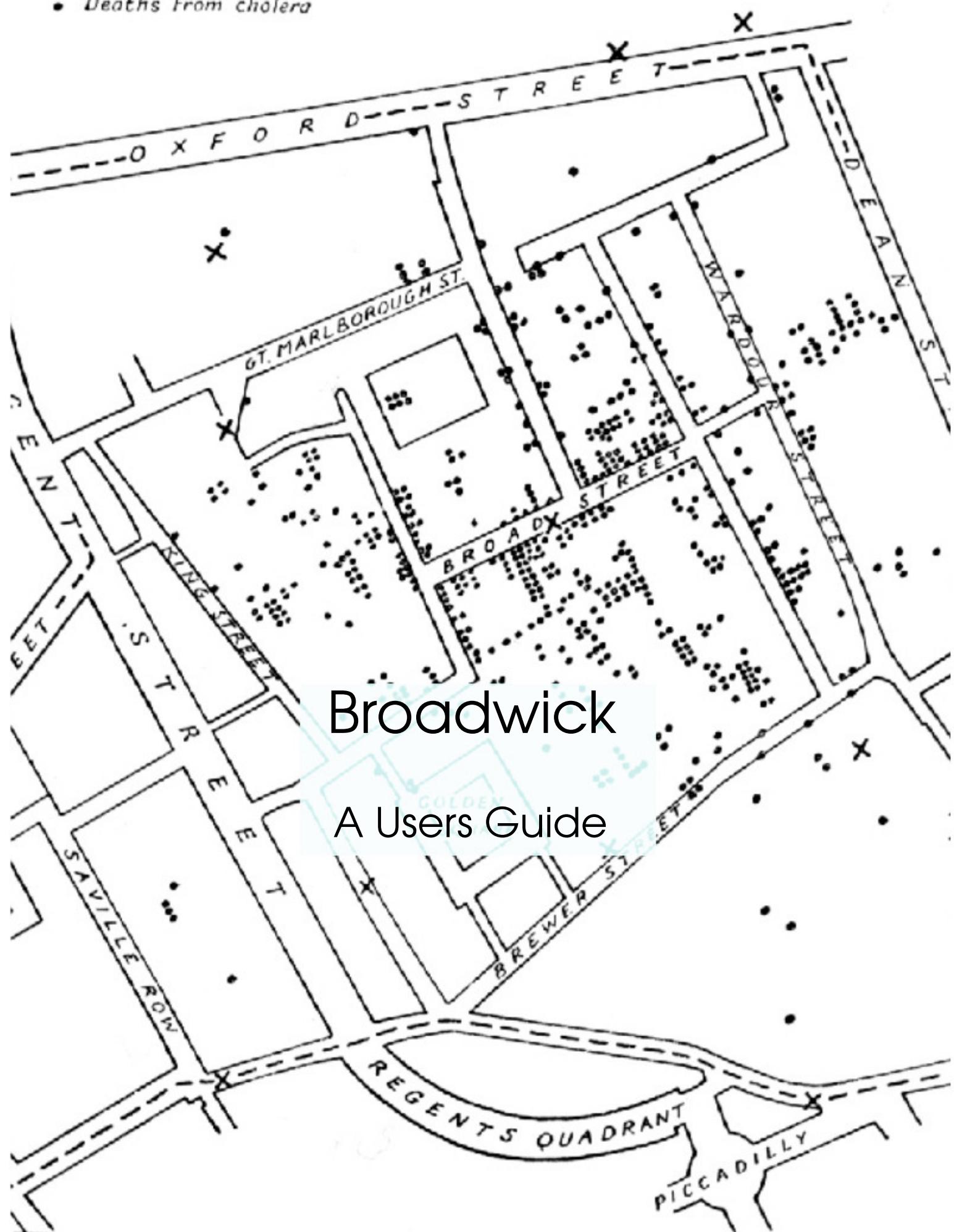
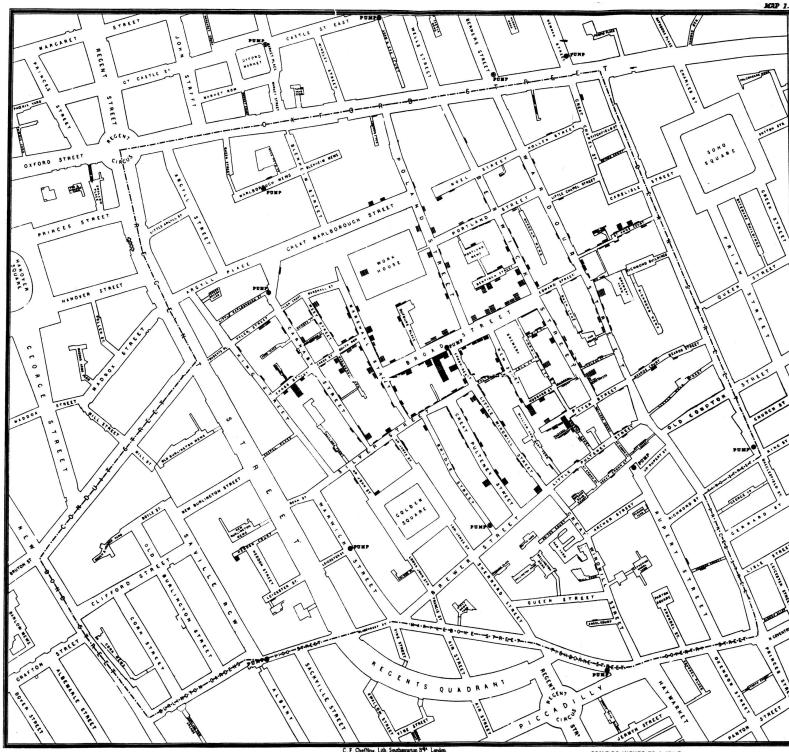


50 Yards  
100  
150  
200

- Deaths from cholera





Published by C.F. Cheffins, Lith, Southhampton Buildings, London, England, 1854 in Snow, John. On the Mode of Communication of Cholera, 2nd Ed, John Churchill, New Burlington Street, London, England, 1855. This image was originally from <http://en.wikipedia.org/wiki/File:Snow-cholera-map-1.jpg>

#### Cover Picture

A variant of the original map drawn by Dr. John Snow (1813-1858), a British physician who is one of the founders of medical epidemiology, showing cases of cholera in the London epidemics of 1854, clustered around the locations of water pumps.

This image is in the public domain because its copyright has expired. This applies to Australia, the European Union and those countries with a copyright term of life of the author plus 70 years.

The figures used in the chapter headings are cropped images from unsplash.com and are licensed under Creative Commons Zero license.

Copyright © 2013 University of Glasgow

<HTTP://EPICSCOTLAND.GITHUB.IO/BROADWICK/>

Licensed under the Apache License, Version 2.0 (the "License"); you may not use this file except in compliance with the License. You may obtain a copy of the License at <http://www.apache.org/licenses/LICENSE-2.0>. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

*First printing, March 2013*



# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	License	5
<b>2</b>	<b>Installation</b>	<b>7</b>
2.1	Installing the Distribution Jar	7
2.2	Installing From Source	7
2.3	Ubuntu	8
<b>3</b>	<b>Using Broadwick</b>	<b>11</b>
3.1	Creating a New Project	11
3.1.1	Using The Command Line	11
3.1.2	Using Netbeans	11
3.2	Configuration Files	14
3.3	Extending the Model	20
<b>4</b>	<b>Calculation Packages</b>	<b>23</b>
4.1	Markov Chains	23
4.2	Monte Carlo Simulation	24
4.2.1	Markov Chain Monte Carlo	25
4.3	Approximate Bayesian Computation (ABC)	26
4.4	Ordinary Differential Equations (ODEs)	28
	<b>Bibliography</b>	<b>31</b>

<b>Index</b>	.....	<b>33</b>
--------------	-------	-----------

# 1. Introduction

In the event of an outbreak it is important to have modelling tools in place to estimate the likely origin, speed of spread, and size, and to be able to predict the impact of intervention measures. However, different diseases in different animal populations require somewhat different approaches due to the detection, transmission and recovery (or not) characteristics of hosts infected with the pathogen, and the contact patterns between susceptible hosts (whether of the same species or not). Consequently specific models developed for one disease system are unlikely to be entirely suitable for another.

Broadwick is a framework for developing sophisticated epidemiological based mathematical models, and consists of several Java libraries and bespoke packages. The components of Broadwick are written in such a way that a scientist may combine them in order to rapidly prototype a model for a new specific scenario.

- Supports single (e.g. within herd) or structured populations (e.g. multi-species or locations)
- Inclusion of movement over network data (e.g. Cattle movement Tracing System)
- Stochastic Individual Based simulations (including fast approximate options)
- Approximate Bayesian Computation inference for estimating model parameters from data via simulations
- Markov Chain Monte Carlo inference for estimating model parameters from data

## 1.1 License

Broadwick is released under the Apache 2 license.





## 2. Installation

A compiled Broadwick jar is available from the EPIC Scotland website that can be added to your project. Broadwick uses maven as its build tool and so this manual is rather maven focused. You can use Broadwick with other tools or import it into your IDE settings though this is outside the scope of this manual.

### 2.1 Installing the Distribution Jar

If you plan on downloading the distribution jar file you should place it under

```
$HOME/.m2/repository/broadwick/broadwick/1.2
```

### 2.2 Installing From Source

To download the Broadwick sources

1. Create a directory for broad wick and ‘cd’ into that directory
2. Copy the broadwick sources.

```
git clone https://github.com/EPICScotland/Broadwick .
```

(this may take a while as there are release snapshots that are also on the site)

3. Build the jar library.

```
mvn package install; cd archetype; mvn install
```

You should see output similar to

```
INFO — maven-compiler-plugin:3.1:compile (default-compile) @ broadwick — INFO
Changes detected - recompiling the module! INFO Compiling 157 source files to
/XXXX/XXXX/XXXX/EPICScotland/Broadwick/target/classes WARNING bootstrap
class path not set in conjunction with -source 1.7 WARNING No processor claimed any
of these annotations: javax.xml.bind.annotation.XmlAccessorType,javax.xml.bind.annotation.XmlAttribute,lo
```

This will install the broadwick jar and the archetype under .m2/repository/broadwick/broadwick/1.2

Now you can create a new project using this archetype:

```
mvn3 archetype:generate -DarchetypeGroupId=broadwick -DarchetypeArtifactId=broadwick-archetype -DarchetypeVersion=1.2 -DgroupId=<unique id for your project> -DartifactId=<your project name> -Dversion=0.1 -Dpackage=<your package>
```

## 2.3 Ubuntu

The version of Java on Ubuntu 15.04 (other versions are also possibly affected) causes the following error when compiling Broadwick due to the jaxb plugin which is used to generate java sources.

```
WARNING: Error injecting: org.jvnet.mjiip.v_2.XJC2Mojo java.lang.NoClassDefFoundError:  
com/sun/xml/bind/api/ErrorListener at java.lang.ClassLoader.defineClass1(Native Method)  
at java.lang.ClassLoader.defineClass(ClassLoader.java:800) at java.security.SecureClassLoader.defineClass(SecureC
```

We recommend using the version of Java from Oracle. The following steps show how to download and install Java on Ubuntu (it is possible to revert back to Ubuntu's version afterwards).

1. Download the 32-bit or 64-bit Linux "compressed binary file" - it has a ".tar.gz" file extension.
2. Uncompress it

```
tar -xvf jdk-8u71-linux-x64.tar.gz (32-bit)
```

The JDK 8 package is extracted into `./jdk1.8.0_71` directory. N.B.: Check carefully this folder name since Oracle seem to change this occasionally with each update.

3. Move the JDK 8 directory to `/usr/lib`

```
sudo mkdir -p /usr/lib/jvm sudo mv ./jdk1.8.0_71 /usr/lib/jvm/
```

4. Now run

```
sudo update-alternatives --install "/usr/bin/java" "java" "/usr/lib/jvm/jdk1.8.0_71/bin/java" 1  
1 sudo update-alternatives --install "/usr/bin/javac" "javac" "/usr/lib/jvm/jdk1.8.0_71/bin/javac" 1  
1 sudo update-alternatives --install "/usr/bin/javaws" "javaws" "/usr/lib/jvm/jdk1.8.0_71/bin/javaws" 1
```

This will assign Oracle JDK a priority of 1, which means that installing other JDKs will replace it as the default. Be sure to use a higher priority if you want Oracle JDK to remain the default.

5. Correct the file ownership and the permissions of the executables:

```
sudo chmod a+x /usr/bin/java sudo chmod a+x /usr/bin/javac sudo chmod a+x /usr/bin-/javaws sudo chown -R root:root /usr/lib/jvm/jdk1.8.0_71
```

N.B.: Remember - Java JDK has many more executables that you can similarly install as above. `java`, `javac`, `javaws` are probably the most frequently required. This answer lists the other executables available.

6. Run

```
sudo update-alternatives --config java
```

You will see output similar to the one below - choose the number of `jdk1.8.0_71`

```
sudo update-alternatives --config java
```

There are 3 choices for the alternative java (providing `/usr/bin/java`).

```
Selection Path Priority Status ----- * 0
/usr/lib/jvm/java-7-openjdk-amd64/jre/bin/java 1071 auto mode 1 /usr/lib/jvm/java-7-
openjdk-amd64/jre/bin/java 1071 manual mode 2 /usr/lib/jvm/jdk1.8.0_71/bin/java 1
manual mode
```



## 3. Using Broadwick

### 3.1 Creating a New Project

Broadwick contains a set of packages that can be used as required. The framework is designed to be flexible and does not place any requirement on the user on how to use the framework. It is possible to use the classes and packages of Broadwick without using the powerful framework, creating your own main() method and taking responsibility for reading data files and configuration items though this is not the recommended way of using Broadwick.

#### 3.1.1 Using The Command Line

The Broadwick distribution contains a maven archetype for generating a skeleton project that contains all the configuration files, source code etc that is required to start a project based upon Broadwick. It uses apache maven as it's build tool. To generate a skeleton using this archetype on the command line (assuming that the broadwick-archetype jar is in your local repository)

```
mvn3 archetype:generate -DarchetypeGroupId=broadwick -DarchetypeArtifactId=broadwick-archetype -DarchetypeVersion=1.1 -DgroupId=broadwick.proj -DartifactId=StochasticSir -Dversion=0.1 -Dpackage=broadwick.stochasticsir
```

The groupId (maven uses the group id to uniquely identify your project), artifactId (is the name of your generated jar file without a version), version (the version number for your generated project) and the package to which the generated source will be created can be changed by modifying the -DgroupId, -DartifactId, -Dversion and -Dpackage arguments above.

#### 3.1.2 Using Netbeans

It is possibly easier to create a project using Netbeans (a free IDE available form Oracle, the ‘owners’ of Java). Open the Netbeans IDE and select File->New Project and choose a Maven project and “Project from Archetype” from the list of projects (see figproj1).

Click Next and choose the latest version of the broadwick-archetype from the “Known Archetypes” (see fig 3.2). The version of the broadwick archetype corresponds to the version of Broadwick.

The projects details can be specified on the next screen (fig 3.3).

Clicking “Finish” will create the project (fig 3.4).

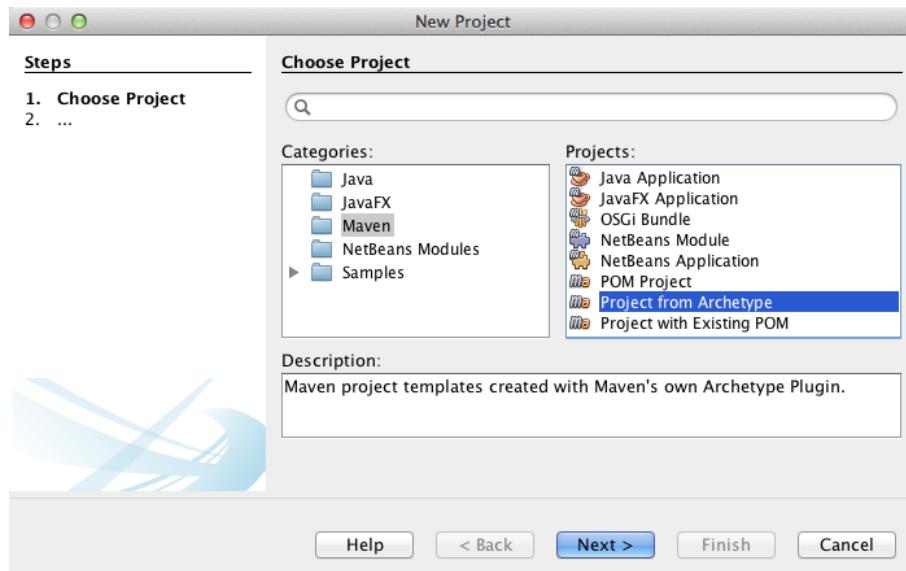


Figure 3.1: Creating a maven based project

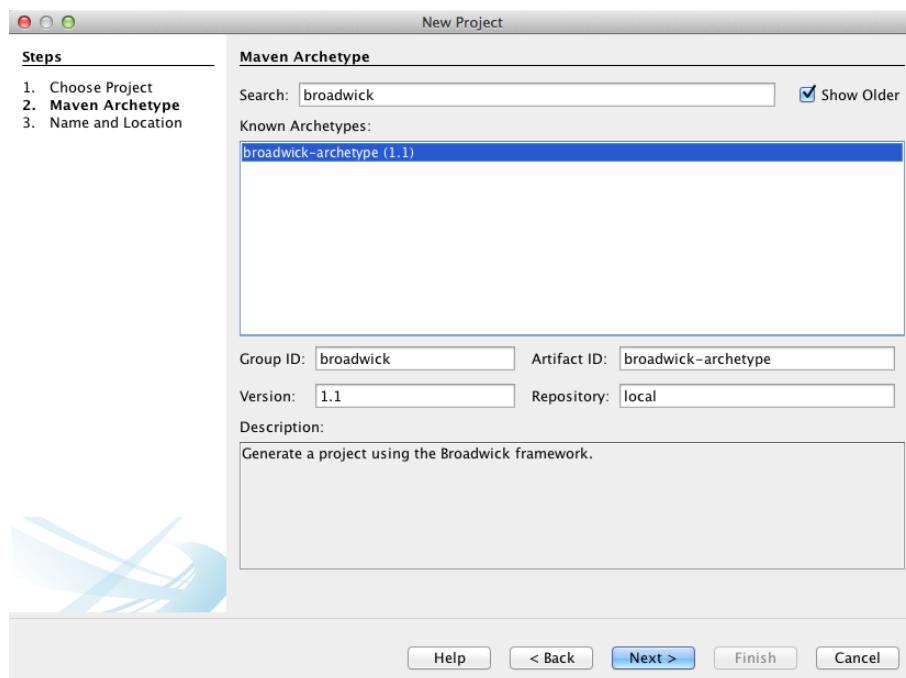


Figure 3.2: Using the Broadwick archetype to create a skeleton project.

A number of minor changes are needed in the generated project.

Select the name of the class (“App”) and right-click and select Refactor->Rename; Change the name of the class to StochasticSIR.

In the Broadwick.sh file change the \${artifactId} and \${version} to the artifactId and version specified when the project was created (StochasticSIR and 1.0 respectively). This is a shell script for running your project on Unix based systems, you will need to make it executable.

In Broadwick.xml (the configuration file for the generated project) change the name of the <classname> element to reflect the package and class (broadwick.stochasticsir.StochasticSIR)

We can build the generated project by selecting Run->Build Project from the menu bar or

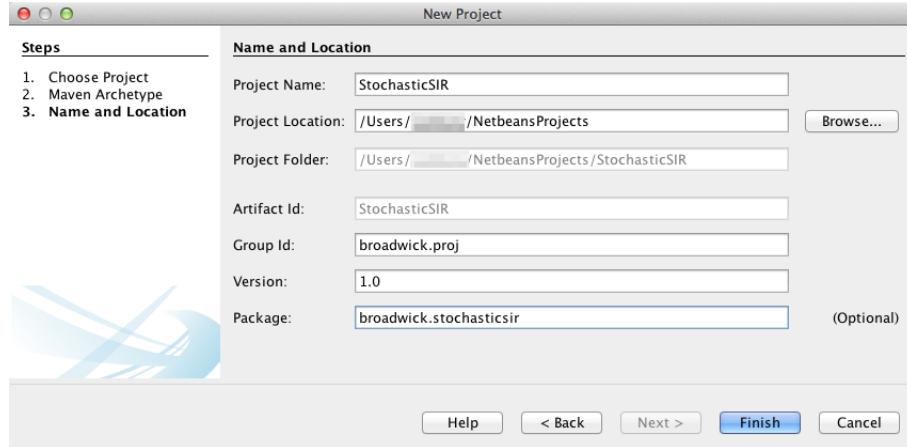


Figure 3.3: Setting project details for a Broadwick based project.

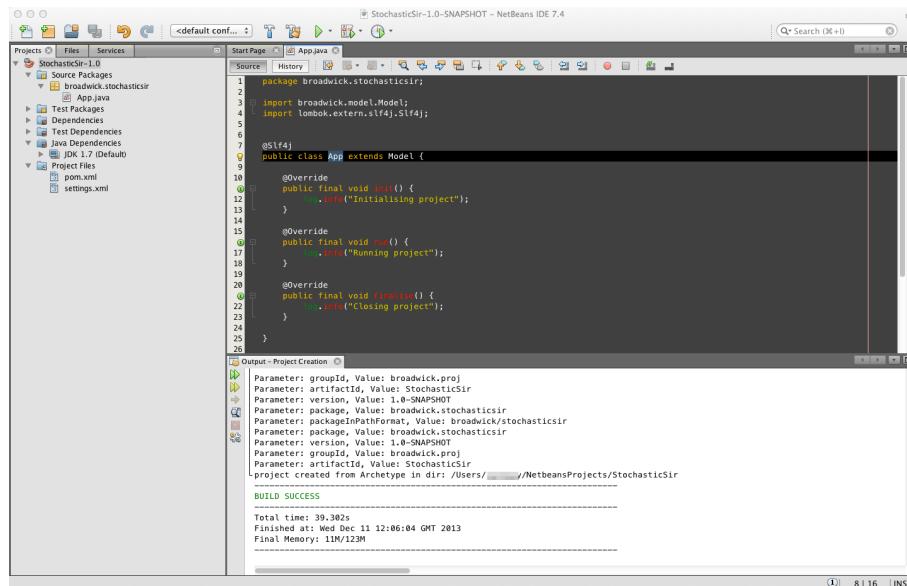


Figure 3.4: The generated project

clicking the icon in the toolbar (see fig 3.5 for some example output from this process). This will create a directory called target that contains, among other items, a jar file containing the compiled code and an executable jar file ending in .one-jar.jar. The Broadwick.sh file is a shell script that will run the executable jar file on \*NIX systems.

When Broadwick starts it looks for all the models specified in the <model> elements in the projects configuration file. It creates objects for each <model> found using the default (empty) constructor for the class given in the <classname> element of the model (this is why no constructor is generated for the project).

For each project object created Broadwick will call the init(), run() and finalise() methods in turn. In our skeleton project we simply logged the fact that these methods were called. A simplified outline of how Broadwick initialises itself is shown in fig 3.6.

A description of the configuration file is outlined in the next section.

```

26 Terminal - /NetbeansProjects/StochasticSir
bash-3.2$ cd /Users/.../NetbeansProjects/StochasticSir/
bash-3.2$ ./Broadwick.sh -c Broadwick.xml
[main] INFO Running broadwick Version 1.1 Build (d177045.stir.ac.uk - 80e92a8c63e7+ : 2013-12-06 12:19)
[main] INFO Running broadwick with the following models [Broadwick Project]
[pool-1-thread-1] INFO Running Broadwick Project [Broadwick,stochasticSIR]
[pool-1-thread-1] INFO Initialising project
[pool-1-thread-1] INFO Running project
[pool-1-thread-1] INFO Closing project
[main] INFO Simulation complete. 0:00:00.075
bash-3.2$

```

Figure 3.5: Figure caption

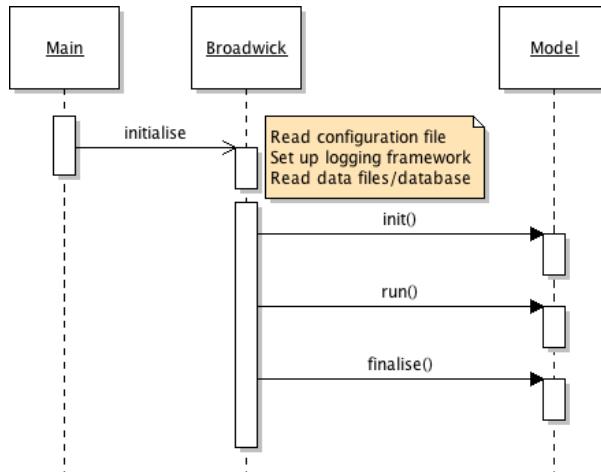


Figure 3.6: Schematic outline of the steps Broadwick performs on startup

### 3.2 Configuration Files

The configuration file MUST conform to the Broadwick.xsd specification that is supplied with the Broadwick source code. The configuration is contained within the `<project>` `</project>` tags and contains the following tags

---

<code>&lt;logs&gt;</code>	<code>&lt;console&gt;</code>	Contains <code>&lt;level&gt;</code> and <code>&lt;pattern&gt;</code> tags to define the level and structure of logging messages displayed on the console.
	<code>&lt;files&gt;</code>	Contains <code>&lt;level&gt;</code> and <code>&lt;pattern&gt;</code> tags like the <code>&lt;console&gt;</code> log but also a <code>&lt;name&gt;</code> for the name of the file to contain the log messages and a boolean <code>&lt;overwrite&gt;</code> tag that lets Broadwick know if any existing file with that name should be overwritten.
<code>&lt;data&gt;</code>	<code>&lt;databases&gt;</code>	Contains the <code>&lt;name&gt;</code> tag to specify the location of the database.
	<code>&lt;datafiles&gt;</code>	Contains the <code>&lt;DirectedMovementFile&gt;</code> <code>&lt;FullMovementFile&gt;</code> <code>&lt;BatchMovementFile&gt;</code> <code>&lt;PopulationFile&gt;</code> <code>&lt;LocationsFile&gt;</code> and <code>&lt;TestsFile&gt;</code> tags for the various file structures recognised by Broadwick. See the following tables in this section for a description of each.
<code>&lt;models&gt;</code>	<code>&lt;model&gt;</code>	Broadwick can run several models concurrently (though they will share the same logging and data configurations).

---

Each model element contains the following:

---

<code>&lt;classname&gt;</code>	Exactly one classname element that is the fully qualified name of the java class (that implements the <code>broadwick.model.Model</code> interface). This class <b>MUST</b> have a no argument constructor that Broadwick will use to create an instance through relection.
<code>&lt;priors&gt;</code>	The <code>&lt;priors&gt;</code> tag can contain as many elements as necessary, each prior contains an <code>id</code> attribute and optional <code>&lt;hint&gt;</code> and <code>&lt;initialVal&gt;</code> elements. The following priors are recognised by Broadwick <code>&lt;uniformPrior&gt;</code> additionally contains <code>&lt;min&gt;</code> and <code>&lt;max&gt;</code> elements for uniformly distributed priors. <code>&lt;gaussianprior&gt;</code> additionally contains <code>&lt;mean&gt;</code> and <code>&lt;deviation&gt;</code> tags for normally distributed priors.
<code>&lt;parameter&gt;</code>	The parameters for the model can be encoded in this tag using the <code>id</code> and <code>value</code> attributes to name the parameters and give the parameter value. These can be accessed by the <code>broadwick.model.Model getParameterAs[TYPE](id)</code> method that retrieves the parameter with the given id and converts it the the required type.

---

Each datafile contains specific information on it's layout, i.e. the columns in the file where the required data can be found. The structure of each recognised data file is outlined below.

`DirectedMovementFile`:

<name>	The name (including path from the configuration file) where the file can be found.
<alias>	An alias for the file.
<separator>	The character separating the columns in the datafile, e.g. ',' '<tab>'.
<idColumn>	
<movementDateColumn>	
<movementDirectionColumn>	
<locationColumn>	
<speciesColumn>	
<dateFormat>	The format the date is given in the data file, see below for details.
<customTags>	This optional field allows for optional information to be stored in the database.

---

**FullMovementFile:**

<name>	The name (including path from the configuration file) where the file can be found.
<alias>	An alias for the file.
<separator>	The character separating the columns in the datafile, e.g. ',' '<tab>'.
<idColumn>	
<departureDateColumn>	
<departureLocationIdColumn>	
<destinationDateColumn>	
<destinationLocationIdColumn>	
<marketIdColumn>	
<marketDateColumn>	
<speciesColumn>	
<dateFormat>	The format the date is given in the data file, see below for details.
<customTags>	This optional field allows for optional information to be stored in the database.

---

**BatchMovementFile:**

---

<name>	The name (including path from the configuration file) where the file can be found.
<alias>	An alias for the file.
<separator>	The character separating the columns in the datafile, e.g. ‘,’ ‘<tab>’.
<batchSizeColumn>	
<departureDateColumn>	
<departureLocationIdColumn>	
<destinationDateColumn>	
<destinationLocationIdColumn>	
<marketIdColumn>	
<marketDateColumn>	
<speciesColumn>	
<dateFormat>	The format the date is given in the data file, see below for details.
<customTags>	This optional field allows for optional information to be stored in the database.

---

## PopulationFile:

---

<name>	The name (including path from the configuration file) where the file can be found.
<alias>	An alias for the file.
<separator>	The character separating the columns in the datafile, e.g. ‘,’ ‘<tab>’.
<lifehistory>	
<population>	
<speciesColumn>	
<dateFormat>	The format the date is given in the data file, see below for details.
<customTags>	This optional field allows for optional information to be stored in the database.

---

## Lifehistory:

---

<name>	The name (including path from the configuration file) where the file can be found.
<alias>	An alias for the file.
<separator>	The character separating the columns in the datafile, e.g. ‘,’ ‘<tab>’.

---

## Population:

---

<name>	The name (including path from the configuration file) where the file can be found.
<alias>	An alias for the file.
<separator>	The character separating the columns in the datafile, e.g. ‘,’ ‘<tab>’.

---

## LocationFile:

---

<name>	The name (including path from the configuration file) where the file can be found.
<alias>	An alias for the file.
<separator>	The character separating the columns in the datafile, e.g. ',' '<tab>'.
<locationIdColumn>	The column in the file containing the id of the location.
<eastingColumn>	The column in the file containing the easting coordinate. Coordinates aren't strictly adhered to in Broadwick so a simple y-coordinate is sufficient.
<northingColumn>	The column in the file containing the northing coordinate. Coordinates aren't strictly adhered to in Broadwick so a simple x-coordinate is sufficient.
<dateFormat>	The format the date is given in the data file, see below for details.
<customTags>	This optional field allows for optional information to be stored in the database.

---

TestsFile:

---

<name>	The name (including path from the configuration file) where the file can be found.
<alias>	An alias for the file.
<separator>	The character separating the columns in the datafile, e.g. ',' '<tab>'.
<idColumn>	One of these is required, specifying whether the test is performed on an individual, group (e.g. herd) or location (must match the id in the Location file).
<groupIdColumn>	
<locationIdColumn>	
<testDateColumn>	
<positiveResultColumn>	
<negativeResultColumn>	
<dateFormat>	The format the date is given in the data file, see below for details.
<customTags>	This optional field allows for optional information to be stored in the database.

---

The date format specification follows the standard formatting for dates and times:

---

Symbol	Meaning
G	era
C	century of era ( $>=0$ )
Y	year of era ( $>=0$ )
x	year
w	week of weekyear
e	day of week
E	day of week
y	year
D	day of year
M	month of year
d	day of month
a	halfday of day
K	hour of halfday (0 11)
h	clockhour of halfday (1 12)
H	hour of day (0 23)
k	clockhour of day (1 24)
m	minute of hour
s	second of minute
S	fraction of second
z	time zone
Z	time zone offset/id
'	escape for text
''	single quote

---

A simplified configuration file is generated in the skeleton project. It contains configuration items for logging to console and to file for different logging levels (info, warning, error, debug, trace) and we can specify the pattern to apply to the log message.

```

<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<project>
    <logs>
        <console>
            <level>info</level>
            <pattern>[%thread] %-5level %msg %n</pattern>
        </console>
        <file>
            <name>broadwick.stochasticsir.log</name>
            <level>info</level>
            <pattern>[%thread] %-5level %msg %n</pattern>
            <overwrite>true</overwrite>
        </file>
    </logs>

    <models>
        <model id="Broadwick Project">
            <classname>broadwick.stochsir.StochasticSIR</classname>
        </model>
    </models>
</project>
```

Common logging patterns are:

---

%C	Outputs the fully-qualified class name of the caller issuing the logging request.
%M %method	Outputs the method name where the logging request was issued.
%L %line	Outputs the line number from where the logging request was issued.
%F %file	Outputs the file name of the Java source file where the logging request was issued. This is not very fast and should be avoided.
%d	Used to output the date of the logging event e.g. %dHH:mm:ss,SSS
%m (%msg)	Outputs the application-supplied message associated with the logging event.
%t (%thread)	Outputs the name of the thread that generated the logging event.
%n	Outputs the platform dependent line separator character or characters
%r	Outputs the number of milliseconds elapsed since the start of the application until the creation of the logging event.
%p %level	Outputs the level of the logging event.

---

More details on logging patterns can be found at <http://logback.qos.ch/manual/layouts.html>.

The model section requires a <classname> giving the fully qualified class name and optional <priors> and <parameter> sections.

### 3.3 Extending the Model

Our stochastic SIR model that we have created is a valid Broadwick model but does not perform any useful calculations. We will add some parameters to the configuration file and read (and log them) in the init() method.

Firstly, let us define beta and rho parameters for the susceptible->infectious rate and for the infectious->recovered rates respectively and parameters for the maximum time for which we will run the simulation and the name of a file in which we will save the time series data. To do this modify the configured model section by:

```
<model id="Broadwick Project">
    <classname>broadwick.stochasticsir.StochasticSIR</classname>

    <parameter id="beta" value="0.2" />
    <parameter id="rho" value="0.3" />
    <parameter id="tMax" value="100" />
    <parameter id="outputFile" value="broadwick.stochasticSIR.dat" />
</model>
```

Now edit the init() method of the StochasticSir class as shown in fig 3.7.

The ‘Model’ class contains getParameterValue(String), getParameterValueAsDouble(String), getParameterValueAsInteger(String), getParameterValueAsBoolean(String) methods to extract parameters from the configuration file as strings (default), doubles, integers and booleans (if the parameter is written as “true” or “false”).

The screenshot shows the NetBeans IDE interface. The top window displays the Java code for `StochasticSIR.java`. The code is annotated with several `@Override` and `@Parameter` annotations. It includes methods for initialization, running the model, and finalization, along with static variables for parameters like `beta`, `rho`, and `maxT`. The bottom window is a terminal window titled "Output - StochasticSIR-1.0" showing the execution of `broadwick.sh` for the `Broadwick` project. The terminal output shows the project is running, initializing, and saving data to `broadwick.stochasticsir.StochasticSIR.dat`. It also prints the values of `beta` (0.2), `rho` (0.3), and `Max.t` (100.0). The process concludes with a simulation complete message.

```
Start Page StochasticSIR.java
Source History ...
1 package broadwick.stochasticsir;
2
3 import ...2 lines
4
5 @Slf4j
6 public class StochasticSIR extends Model {
7
8     @Override
9     public final void init() {
10         log.info("Initialising project");
11
12         outputFileName = getParameterValue("outputFile");
13         beta = getParameterValueAsDouble("beta");
14         rho = getParameterValueAsDouble("rho");
15         maxT = getParameterValueAsDouble("tMax");
16
17         log.info("Saving data to {}", outputFileName);
18         log.info("beta = {}, rho = {}, maxT = {}", beta, rho, maxT);
19         log.info("Max t = {}", maxT);
20     }
21
22     @Override
23     public final void run() {...3 lines}
24
25     @Override
26     public final void finalize() {...3 lines}
27
28     private static String outputFileName;
29     private static double beta;
30     private static double rho;
31     private static double maxT;
32 }
33
34 Output - StochasticSIR-1.0 Terminal - ...
35 bash-3.2$ ./Broadwick.sh -- Broadwick.xml
36 [main] INFO Running broadwick Version 1.1 Build (d177045.stir.ac.uk - 80e92a8c63e7+ : 2013-12-06 12:19)
37 [main] INFO Running broadwick for the following models [Broadwick Project]
38 [pool-1-thread-1] INFO Running Broadwick Project [broadwick.stochasticsir.StochasticSIR]
39 [pool-1-thread-1] INFO Initialising project
40 [pool-1-thread-1] INFO Saving data to broadwick.stochasticsir.StochasticSIR.dat
41 [pool-1-thread-1] INFO beta = 0.2, rho = 0.3
42 [pool-1-thread-1] INFO Max.t = 100.0
43 [pool-1-thread-1] INFO Running project
44 [pool-1-thread-1] INFO Closing project
45 [main] INFO Simulation complete. 0:00:00.001
46 bash-3.2$
```

Figure 3.7: Reading some parameters to our model





## 4. Calculation Packages

Broadwick supports several simulation and fitting models. In this chapter we will give an outline of these methods and how they can be simulated (and combined) using the Broadwick framework. We will, for the most part, dispense with the theory and concentrate on how the methods are implemented in Broadwick. The code snippets in this chapter are taken from the examples that are distributed with the Broadwick source code.

### 4.1 Markov Chains

A Markov chain is a random sequence of states where the current state depends solely on the previous state. In this sense, it is a “memoryless” process as the transition from one state to the next does not depend on the sequence of states that preceded it.

A Markov Chain can be implemented in Broadwick using the MonteCarloStep and MarkovChain classes. The MonteCarloStep encapsulates the functionality of a state by maintaining a collection of the coordinates defining the state as a `java.util.Map<String,Double>` (i.e. the name and value of the state). A MarkovChain object is constructed using a MonteCarloStep object as an initial point and, optionally, a MarkovProposalFunction for generating the next step. The `generateNextStep` method uses the proposal function to generate the next step in the chain as the following code snippet demonstrates,

```
final Map<String, Double> coordinates = new LinkedHashMap<>();
{
    coordinates.put("x", 0.0);
    coordinates.put("y", 0.0);
}
final MonteCarloStep initialStep = new MonteCarloStep(coordinates);

final MarkovChain mc = new MarkovChain(initialStep);
for (int i = 0; i < chainLength; i++) {
    final MonteCarloStep nextStep = mc.generateNextStep(mc.getCurrentStep());
    mc.setCurrentStep(nextStep);

    log.trace("{}", nextStep.toString());
}
```

---

By default, a `MarkovNormalProposal` object is used to generate the next step by sampling from a `Normal` distribution centered on the current coordinate and with a standard deviation of 1. New proposal classes implement the `MarkovProposalFunction` interface and using this object when creating the `MarkovChain`.

---

```
final MarkovProposalFunction myProposer = new MarkovProposalFunction();
final MarkovChain mc = new MarkovChain(initialStep, myProposer);
```

---

## 4.2 Monte Carlo Simulation

Monte Carlo simulation is a broad class of methods that reply on repeated simulation of [random] processes to derive numerical results. Broadwick uses the `MonteCarloScenario` abstract class to encapsulate the process to be simulated, which is used by the `MonteCarlo` class to implement the Monte Carlo process.

Internally, the `MonteCarlo` class uses a producer-consumer pattern by creating a `ThreadFactory` to spawn several simulation processes which are in turn consumed by a `MonteCarloResults` object. This `MonteCarloResults` object takes the results of each simulation (the producer threads) and calculates the required statistics.

These classes (`MonteCarloScenario`, `MonteCarloResults`) are extended for each implementation. By way of example, we will calculate  $\pi$  by throwing darts randomly at a square dartboard and calculating the fraction that fall within the unit circle encompasses by the square.

First, we implement the `run` method of our class that extends the `MonteCarloScenario` class (this class has an internal random number generator, `rng`, to generate random number).

---

```
@Override
public MonteCarloResults run() {
    final MyResultsConsumer results = new MyResultsConsumer();

    final double x = rng.getDouble(-1, 1);
    final double y = rng.getDouble(-1, 1);

    final double r = Math.sqrt(x * x + y * y);
    if (r < 1) {
        results.addHit();
    } else {
        results.addMiss();
    }
    return results;
}
```

---

The `MonteCarloResults` class is responsible for both storing the results of each simulation and for acting as a consumer object that maintains a collection of the results returned by the producers.

---

```
class MyResultsConsumer implements MonteCarloResults {

    @Override
    public double getExpectedValue() {
        return hits.getSum() / (hits.getSum() + misses.getSum());
    }
}
```

```

@Override
public Samples getSamples() {
    return hits;
}

@Override
public String toCsv() {
    return String.format("\%d ; \%d", hits.getSize(), misses.getSize());
}

@Override
public MonteCarloResults join(final MonteCarloResults results) {
    // This is where the results of the producers are dealt with.
    final MyResultsConsumer r = (MyResultsConsumer) results;
    this.hits.add(r.hits);
    this.misses.add(r.misses);

    return this;
}

public void addHit() {
    hits.add(1);
}

public void addMiss() {
    misses.add(1);
}

@Override
public void reset() {
}

@Getter
private final Samples hits = new Samples();
@Getter
private final Samples misses = new Samples();
}

```

These two classes are utilised thus

---

```

MonteCarlo mc = new MonteCarlo(new Simulation(), 1000);
mc.setResultsConsumer(new MyResultsConsumer());
mc.run();

final MyResultsConsumer results = (MyResultsConsumer) mc.getResults();
log.info("Hits : Misses = {}", results.toCsv());
log.info("Estimation of Pi = {}", 4 * results.getExpectedValue());

```

---

#### 4.2.1 Markov Chain Monte Carlo

Markov chains can be combined with Monte Carlo simulation to explore a parameter space by using the Markov chain to ‘walk’ through the parameter space while Monte Carlo simulation is to determine the state of the system and each step in the walk. Thus, Markov chain Monte Carlo (MCMC) methods can be used to sample from a probability distribution by using the Markov chain (coupled with a rejection function to accept or reject proposed steps) to find the desired

distribution.

A `MarkovChainMonteCarlo` object will create a Markov chain and run the Monte Carlo simulation at each step using the Metropolis-Hastings algorithm (the `MetropolisHastings` class) to accept successive steps based on the results of the Monte Carlo simulation at the given step.

A `MarkovChainObserver` object can be added to the `MarkovChainMonteCarlo` object which will be informed when a step has been completed. This observer can be used to save the results of the simulation.

By default a Metropolis-Hastings, but it is easy to implement an alternative (in the following example we will use the Metropolis algorithm with a log-likelihood).

---

```

final MonteCarloStep step = new MonteCarloStep(<initial step>);
final MarkovChainMonteCarlo mcmc = new MarkovChainMonteCarlo(
    myModel, <numScenarios>,
    myMonteCarloScenarioResults,
    myMarkovStepGenerator);

mcmc.setAcceptor(new MonteCarloAcceptor() {
    @Override
    public boolean accept(final MonteCarloResults oldResult, final
        MonteCarloResults newResult) {
        final double ratio = newResult.getExpectedValue() -
            oldResult.getExpectedValue();
        return Math.log(generator.getDouble()) < ratio /
            smoothingRatio;
    }
});

MyMarkovChainObserver myMcmcObserver = new MyMarkovChainObserver(mcmc);
mcmc.addObserver(myMcmcObserver);

mcmc.run();

```

---

We can attach a `MarkovChainObserver` to the `MarkovChainMonteCarlo` object that observes the state of the chain (at each step the chain object informs the observer of its state which can be used, e.g., to save it to file).

---

```

final MarkovChainObserver observer = new MyMCObserver();
mcmc.addObserver(observer);

```

---

### 4.3 Approximate Bayesian Computation (ABC)

Approximate Bayesian computation (ABC) is a class of computational methods based on Bayesian Statistics [1–4]. It bypasses the evaluation of a likelihood function, which is often computationally expensive or even impossible.

At the heart of the ABC method is the ‘distance function’ which is a measure of close the proposed sample is to the desired (posterior). The distance function in Broadwick is specified by overriding the `AbcDistance` class and adding it to an `ApproxBayesianComp` object, by default a simple absolute value of the difference between the proposed and observed value is used (as can be seen in the following example).

An `AbcController` object is used to control (i.e. determine when the calculation should end) the ABC process.

The `ApproxBayesianComp` class runs the bayesian computation. It is constructed using

observed data (in the form of a `AbcNamedQuantity` object, which is a simple name-value map), an `AbcModel` object, an `AbcPriorsSampler` object (which specifies how samples are drawn from a prior distribution) and a sensitivity.

As a very simple example, we will sample from a posterior that is normally distributed around  $\pi$  (assuming a standard deviation of 1.0). Using a uniform prior (for the mean of the posterior distribution) in [3,4] we will sample from this using a simple `abs()` function as the distance function. We should observe a posterior distribution [normally] distributed around  $\pi$ .

---

```
// first set up the observed data (the mean value of my unknown
// distribution).
final Map<String, Double> observed = new LinkedHashMap<>();
observed.put("value", 3.142);
final AbcNamedQuantity observedData = new AbcNamedQuantity(observed);

// Next create a simple controller (we will sample 20000) points from
// the prior distribution.
final AbcController controller = new AbcController() {
    @Override
    public boolean goOn(final ApproxBayesianComp abc) {
        return abc.getNumSamplesTaken() <= 20000;
    }
};

// Create a dummy model to run
final AbcModel myModel = new AbcModel() {
    @Override
    public AbcNamedQuantity run(final AbcNamedQuantity parameters) {
        // this is trivially simple model. Instead of doing any
        // calculations we just return the parameters.
        return parameters;
    }
};

// Create a method for sampling our priors.
final AbcPriorsSampler priors = new AbcPriorsSampler() {
    @Override
    public AbcNamedQuantity sample() {
        // Another dummy method here, we uniformly sample 'value' in
        // the range [3,4]
        final LinkedHashMap<String, Double> sample = new
            LinkedHashMap<>();
        sample.put("value", generator.getDouble(3.0, 4.0));
        return new AbcNamedQuantity(sample);
    }
    private final RNG generator = new RNG(RNG.Generator.Well19937c);
};

final ApproxBayesianComp abc = new ApproxBayesianComp(observedData,
    myModel, priors, 0.05);
abc.setController(controller);
abc.run();
```

---

#### 4.4 Ordinary Differential Equations (ODEs)

Ordinary differential equations can be solved in Broadwick using the 4<sup>th</sup> order Runge-Kutta method. The RungeKutta4 object is constructed using an Ode objectlindexClass!Ode (which specifies the ODEs by implementing methods to specify the initial values and derivatives of each variable), start and end times and the step size.

An observer can be attached to keep track of the results generated by the solver and a controller can be used to ensure that the calculation stops (if, for example, you specify a negative step size resulting in infinite loops or obtain a negative value for a population size).

---

```

final Ode myOde = new SirModel(beta, rho, s0, i0, r0);
final OdeSolver solver = new RungeKutta4(myOde, tStart, tEnd, stepSize);

// we will need an observer to 'observe' our simulation and record the
// simulation states.
solver.getObservers().clear();
final MyObserver observer = new MyObserver(solver, outputFile);
solver.addObserver(observer);

// Create a simple controller object to tell the simulator when to stop.
final OdeController controller = new MyOdeController(tEnd);
solver.setController(controller);

solver.run();

```

---

We can also add triggered events to the solver. These events occur at predetermined times and can be used to model, e.g. immigration events, vaccination or culling strategies in a population.

---

```

// Register theta events, these are fixed events that will be triggered at set
// times.
solver.registerNewTheta(observer, 20.0, new MyThetaEvent(solver));

```

---



## Appendix

### Maven as a Build Tool

There is no requirement to use maven as a build tool but as Broadwick and it's examples are built using maven this section will give a brief outline of how maven is used to create a simple model.

Maven uses an xml file to describe the classes to be built as well as the dependencies, dynamically downloading required libraries as needed. It uses the 'convention over configuration' paradigm imposing the directory structure given in the table below.

Directory	Purpose
Project home	Contains the pom and all the subdirectories.
src/main/java	Contains the java source code for the project.
src/main/resources	Contains the xsd file for configuring the project.
src/test/java	Contains any [Junit or TestNG] test cases for the project.
src/test/resources	Contains resources necessary for testing.

Maven's directory structure

Maven's equivalent to a makefile or Ant's build.xml is a 'project object model' which is stored in a pom.xml file. A good reference for the maven pom is <http://maven.apache.org/pom.html>.





## Bibliography

- [1] T Toni et al. “Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems”. In: *J R Soc Interface* 6 (2007), pages 187–202 (cited on page 26).
- [2] J-M Marin et al. “Relevant statistics for Bayesian model choice”. In: *ArXiv:11104700v1 [mathST]* (2011) (cited on page 26).
- [3] P Marjoram et al. “Markov chain Monte Carlo without likelihoods”. In: *Proc Natl Acad Sci USA* 100 (2003), pages 15324–15328 (cited on page 26).
- [4] Excoffier L (2009) Wegmann D Leuenberger C. “Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood.” In: *Genetics* 182 (2009), pages 1207–1218 (cited on page 26).





# Index

## Symbols

$\pi$  ..... 22

## A

algorithm  
Metropolis-Hastings ..... 24  
Approximate Bayesian Computation (ABC)  
24

## B

Bayesian Statistics ..... 24

## C

Class  
AbcController ..... 24  
AbcDistance ..... 24  
AbcModel ..... 25  
AbcNamedQuantity ..... 25  
AbcPriorsSampler ..... 25  
ApproxBayesianComp ..... 24  
MarkovChain ..... 21  
MarkovChainMonteCarlo ..... 24  
MarkovChainObserver ..... 24  
MarkovNormalProposal ..... 22

MarkovProposalFunction ..... 21, 22  
MetropolisHastings ..... 24  
MonteCarlo ..... 22  
MonteCarloResults ..... 22  
MonteCarloScenario ..... 22  
MonteCarloStep ..... 21  
RungeKutta4 ..... 26  
ThreadFactory ..... 22  
Configuration Files ..... 14  
Creating a New Project ..... 11

## D

distance function ..... 24

## E

Extending the Model ..... 19

## L

License ..... 5  
Likelihood function ..... 24

## M

Markov Chain Monte Carlo ..... 23  
Markov Chains ..... 21  
Maven, Using ..... 27

Memoryless process ..... 21  
method  
    generateNextStep ..... 21  
    run ..... 22  
Monte Carlo Simulation ..... 22  
Monte carlo Simulation ..... 23

**N**

Normal distribution ..... 22

**O**

Ordinary Differential Equations (ODEs) .. 26

**P**

pattern  
    observer ..... 24  
    producer-consumer ..... 22

**R**

Runge-Kutta method ..... 26

**T**

triggered events ..... 26

**U**

Using Netbeans ..... 11  
Using The Command Line ..... 11