StochSS: Stochastic Simulation Service

User Guide and Tutorial

Contents

1

Installation

1.1 Try StochSS

It is possible to try StochSS online at try.stochss.org by signing up with your e-mail address. Nothing has to be installed locally, and all modeling capabilities of StochSS are available. It is *not* possible to take advantage of advanced cloud capabilities (see Section 6) through the try.stochss.org-server, so for the user with the need to run large-scale computations, we recommend to consider a local install after trying out StochSS online.

1.2 Install StochSS

StochSS supports Mac OS X, Linux, and Windows through the Docker platform. To install StochSS, you will need to install Docker on your computer, and then use the appropriate StochSS launch application to initialize and run your StochSS Docker container. Docker is free software that automates the deployment of Linux applications inside software containers. To download Docker, visit https://www.docker.com/.

For instructions on how to install the latest version of StochSS for your platform, please see the StochSS website http://www.stochss.org/. The source code, together with installation instructions, is also available on GitHub: https://github.com/StochSS/stochss/tree/master/stochss-launcher.

2

Basic Introduction

This tutorial will guide you through the basic features of StochSS. You will become familiar with the **Model Editor** and the **Simulation Manager**. You will learn how to create your own model, which can be population or concentration-based, and how to simulate it using either an ordinary differential equation (ODE) solver or the stochastic simulation algorithm (SSA).

2.1 Creating Administrator and Standard User Accounts

At the end of a successful installation process, your default browser will launch and you will be asked to create an admin account as in Figure 2.1 (there is only one admin account for the entire system). Once the admin account is created you will be forwarded to a regular login page where you can log in to StochSS.

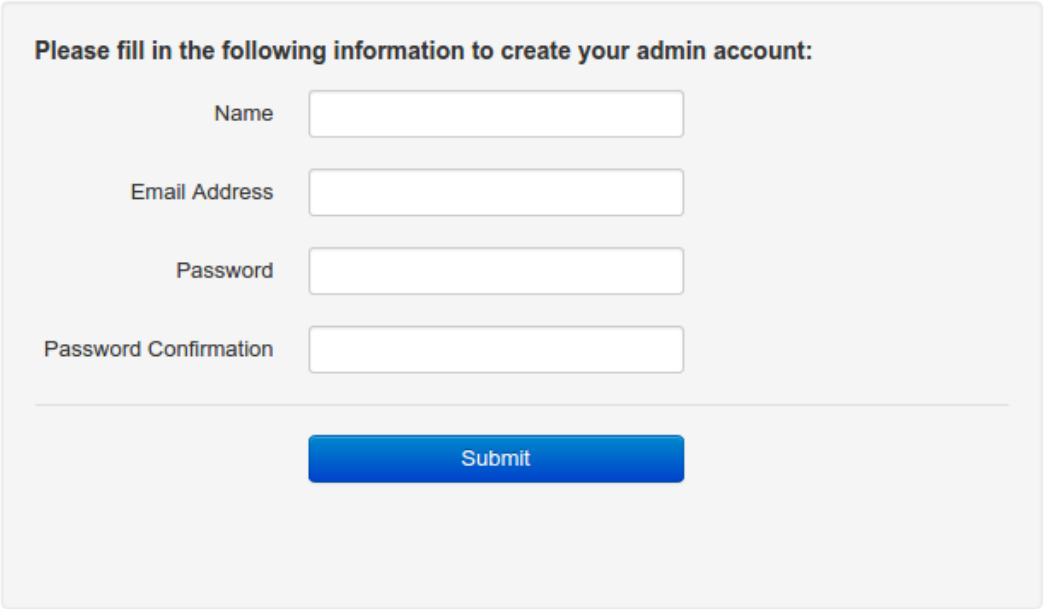

Figure 2.1: Administrator login page

Users can click **Create Account** to request an account. The account will not be accessible until the admin approves it in the **Admin Panel** (the admin can also delete active users as well as reset their passwords there).

2.2 Importing a Model

The **Model Editor** will let you create new or modify existing well-mixed stochastic biochemical models as well as deterministic models based on ODEs. The best way to get started with the model editor is to import an example model and look at the different sections.

Importing an existing model

Option 1: StochSS Public Library

1. Navigate to the **Model Editor** page.
2. Click **Import from Public Library** in the right-hand toolbar.
3. Select a model from the Public Library and click **Copy Model to Library**.

Option 2: Stochkit2 XML

1. Navigate to the **Model Editor** page.
2. Click **Import from .XML** in the right-hand toolbar.
3. Select an XML file. A collection of example models can be found in the *examples* directory within the StochSS install folder.
4. Click **Import**.

Option 3: SBML

1. Navigate to the **Model Editor** page.
2. Click **SBML** in the right-hand toolbar.
3. Select an SBML file and click submit.

Note that due to limitations with the SBML standard, StochSS only supports non-spatial models in this format.

After importing the model, StochSS should display the imported model in the model editor. Look through the page to see how the different **Species, Parameters**, and **Reactions** are defined.

By clicking **Export to .zip**, **Export to Public Library**, or **Export to SBML**, the model can be shared across computers or shared amongst users of the same StochSS.

2.3 Creating a New Model

An example population model is defined by the following two reactions:

$$\begin{aligned} S0 + S0 &\xrightarrow{k1} S1 \\ S1 &\xrightarrow{k0} \emptyset. \end{aligned} \tag{2.1}$$

To create this model:

1. Navigate to the **Model Editor**.
2. Click **Add Model** and select *Population, Well-mixed*.
3. Rename the model to *example*.
4. Click **Create Species** twice to create two species.
5. By default the species are named $S0$ and $S1$. Set the initial condition for $S0$ to 1000 and the initial condition for $S1$ to 0.
6. Similarly to above, click **Add Parameter** twice to add two parameters.
7. By default they will be named $k0$ and $k1$. Set $k0$ to 0.0001 and $k1$ to 0.05.
8. Click **Add Reaction** to add two reactions. Select the reactants, products, rates and reaction types corresponding to (2.1). Compare to Figures 2.2 and 2.3 to verify the settings.

Figure 2.2: Dimerization reaction

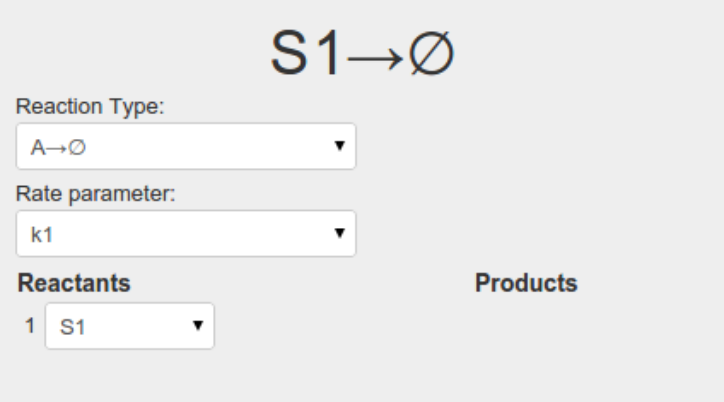

Figure 2.3: Decay reaction

2.4 Running a Simulation and Visualizing results

For this section, create or import a model using the directions above.

1. Navigate to the **Simulation Manager** page.
2. Select the model you wish to simulate and click **Next**.
3. Setup your simulation parameters: name, time, data storage frequency, realizations and solver type.
    a) If you are simulating a population-based model you can choose between the deterministic and the stochastic solvers.
    b) Concentration-based models can only be simulated using the deterministic solver.
4. Click **Run Locally**. You will be automatically forwarded to the **Job Status** page.
5. Click **View** to open the **Job summary** page, where you can visualize the simulation's trajectories.
6. Click **Access Local Data** to download a raw copy of the data. This can be used to share data between StochSS installations or perform manual data analysis.

Converting a concentration model to population

Create a concentration model or use the directions above to import one. Both Lotka-Volterra examples are concentration based and are available both as XML files and Public Library models.

1. Select the newly minted concentration model.
2. Click **Convert to Population** on the right-hand toolbar to start the conversion process. The model conversion page will open.
3. To convert a concentration model to a population, a system volume must be specified.

4. Given a volume, StochSS converts initial conditions through

$$\text{initial\_condition\_population} = \text{initial\_condition\_concentration} \times \text{volume} \tag{2.2}$$

   and attempts to convert the reaction rates. It is not always possible to convert reaction rates automatically. During the conversion process StochSS lists all the reactions, and notifies the user of which reaction rates were successfully converted and which were not. If automatic conversion fails the conversion still proceeds, but the user now has to correct the reaction rates that were not automatically converted.

5. Click **Finish conversion** at the bottom left of the page to create a population-based model. This newly created population model can be simulated using both deterministic and stochastic solvers.

6. Click **Cancel conversion** to cancel the conversion.

The conversion process operates correctly only if the model to be converted is entirely based on the mass action kinetics allowed in Gillespie Stochastic Simulation Algorithm [**?**]. If the model to be converted is not entirely based on mass action kinetics, the conversion tool only converts what it can.

2.5 Backup and Transfer your Data

You can backup or share your saved models and simulations with the StochSS ZIP format. There are three ways to create StochSS ZIP files:

- Navigate to the **Backup** page in the left-hand toolbar and click **Export**. This exports a ZIP containing all models and simulation results for the current user. There is an option to export all data for all users if this page is accessed with the admin account.
- Select a model on the **Model Editor** page and click **Export to .zip**.
- Click **View** on the **Job Status** page, and then click **Access local data**.

These ZIP files can all be imported into StochSS on the **Backup** page. To import the contents of a ZIP file into StochSS:

1. Navigate to the **Backup** page.
2. Click **Import**.
3. Select the ZIP file to upload. The file should automatically begin uploading, and then appear in a table of ZIP archives below.
4. Select the ZIP file in the table.
5. Define the behavior of the import by either limiting what files get imported or specifying how overlapping names are handled.
6. Click **Import** at the bottom of the page.

Exporting data from an old version of StochSS (1.2 or previous)

To create a backup archive from an older version of StochSS, execute the following command from a terminal window in the directory of your new StochSS installation:

```
./exportserver.py path_to_your_old_StochSS_installation
```

You can import the backup archive you created as described above.

3

Sensitivity Analysis

StochSS implements forward sensitivity analysis based on the CVODES ODE solver [**?**].

For instance, if we consider the population of P and the parameter $Vmax$ in the Michaelis-Menten model in the **Public Library**. The sensitivity of the rate of the increase in population P to the parameter $Vmax$ is defined as the incremental rate of change in P due to incremental changes in $Vmax$:

$$\frac{\partial P(t)}{\partial V_{max}}. \tag{3.1}$$

These are the sensitivities that StochSS produces. There is no rescaling to make these sensitivities unitless.

3.1 Example: Michaelis-Menten

In this example we will study the Michaelis-Menten model from the StochSS Public Library. Details on the model can be found in [**?**].

1. Import the model from the public library. By default, it will be called *michaelis_menten*.
2. Navigate to the **Simulation Manager**, select the model, and click **Next**.
3. Under **Simulation type**, select *Deterministic + Sensitivity*.
4. You can now select the parameters for which you wish to compute sensitivities. For instance you could select Km, $k3$, mu, and $Vmax$, to perform sensitivity analysis with only these parameters.
5. Click **Run Locally**.
6. Navigate to the **Job Summary** page to analyze the output. You can plot species trajectories as well as sensitivities.

4

StochOptim: Parameter Estimation

StochSS implements parameter estimation for stochastic well-mixed biochemical systems, StochOptim, via the Monte Carlo expectation-maximization with Modified Cross-Entropy method (MCEM^2) [**?**]. MCEM^2 computes maximum likelihood parameter estimates (MLEs) and associated uncertainties in three consecutive phases: cross-entropy, Monte Carlo expectation-maximization (MCEM), and uncertainty quantification [**?**].

4.1 Example: Estimate parameters of a birth-death model

We consider the Birth-Death model that can be imported from the StochSS Public Library (see Section 2 for instructions on how to import models in StochSS). After the model has been imported:

1. Select **Parameter estimation** from the menu on the left of the screen.
2. Select the *birthdeath* model and click **Next**. This will open the **Simulation page**.
3. Example files (StochOptim input data) for the initial conditions and the trajectories are provided in the *examples* folder (included in your StochSS package) and are named *birthDeathInitial.txt* and *birthDeathTrajectories.txt*, respectively.
4. Upload and select these example files.
5. Configure the checkboxes to only perform parameter estimation on $k1$. For the example model, the value of $k2$ is already optimized.
6. Click **Run Locally** to start the parameter estimation.
7. Select **Job Status** from the menu on the left of the screen to monitor the status of your submitted job. For the example trajectory data, $k1 = 1.0$ and $k2 = 0.03$ are the correct parameter values so the simulation should converge to somewhere close to these.
8. Click **View Progress** to access the **Job Summary** page and to view more details about status of your job. Parameter estimation calculations using MCEM^2 are typically time consuming computations. This simple example takes a little less than 20 minutes on a desktop Intel i7 computer. A more realistic job will take much longer to run.
9. When the job has finished, you can generate a new model using the final estimates of the MCEM^2 calculation by scrolling to the bottom of the **Job Summary** page and click **Create Model from Current Estimates**.

10. If the job doesn't complete in a reasonable amount of time the job can be stopped manually on the **Job Status** page and the parameters can be extracted by clicking **Create Model from Current Estimates**.

StochOptim input file format

The StochOptim input data format consists of two tab-separated text files. The first file represents the initial conditions of the system, and the second file represents trajectories of it. Each file has three base columns. *Time*, *Rep*, and *Weight*. Additional columns are added for every species in the model that parameter estimation is to be run on.

- The *Time* column contains the time values of the various trajectories (and should be set to zero for the initial conditions file).
- The *Rep* column is used to include multiple trajectories in one file for fitting. Basically each trajectory should have a different *Rep* number.
- The *Weight* column is currently unused and should be set to 1.
- The rest of the columns should be named after the species (case-sensitive) in the model the data will be fit against, and the columns themselves should contain integers representing the population counts at the various time points.

An example of this can be seen by comparing the files *birthDeathTrajectories.txt* and *birthDeathTrajectoriesMulti.txt* in the *examples* folder included in the StochSS package.

5

Spatial Stochastic Simulations

The spatial stochastic simulation capabilities in StochSS are based on PyURDME [**?**]. PyURDME is a general software framework for modeling and simulation of stochastic reaction-diffusion processes on unstructured, tetrahedral (3D) and triangular (2D) meshes. The current core simulation algorithm is based on the mesoscopic reaction-diffusion master equation (RDME) model. The default solver is an efficient implementation of the next subvolume method (NSM) [**?**].

5.1 Example: Annihilation of two species in a cylinder

We will build a simple annihilation model based on an cylinder geometry. At each end of the cylinder, different chemicals will be produced. When they diffuse and meet at the center, they will annihilate each other.

1. Navigate to the main **Model editor**.
2. Add a new model. Select *Population, spatial* in the dropdown menu.
3. Click **Mesh** and select *Cylinder*. The cylindrical mesh is divided into three subdomains which can be visualized with the controls below the wireframe view.
4. Add two species, A and B, both with diffusion constant 1.
5. Click **Initial Condition**, select *scatter*, and add 500 molecules of species A in subdomain 1 and 500 molecules of species B in subdomain 3.
6. Add two parameters, $k0$ and $k1$, and set their values to 1 and 100, respectively.
7. Add three reactions:

$$\text{R1}: \quad \emptyset \xrightarrow{k_1} A$$

$$\text{R2}: \quad \emptyset \xrightarrow{k_1} B$$

$$\text{R3}: \quad A + B \xrightarrow{k_0} \emptyset$$

8. Reaction $R1$ should be restricted to subdomain 1 and reaction $R2$ to subdomain 3. Reaction $R3$ should be allowed throughout the whole domain.
9. The model is now complete and ready to be simulated.

10. Navigate to the **Simulation manager** page.
11. Select the spatial model you just created and click **Next**.
12. Setup the simulation parameters: name, time, data storage frequency, and number of realizations.
13. You can specify a random seed for the random number generator under **Advanced Settings**.
14. Click **Run locally**.
15. In a few seconds you will be directed to the **Job Status** page where you can check the status of your simulation.
16. Once your simulation is complete, click **View results** to open the **Job summary** page, where you can visualize the diffusion of the two species over time within the cylindrical container and download the output files of the simulation.

6

Cloud Computing

6.1 Credentials

StochSS provides the options to run jobs using the Amazon cloud infrastructures. In order to use Amazon Elastic Computing Cloud (EC2), Simple Storage Service (S3) and DynamoDB database, which are all required for running jobs in the cloud, you need an Amazon Web Services (AWS) account and a credential pair (consisting of a secret key and an access key) in hand.

More information regarding how to create an AWS account and get credentials can be found here: `http://aws.amazon.com`

Setting Credentials in StochSS

You must set your credentials in StochSS manually (Figure 7.1). Once StochSS validates these, you will be able to launch compute nodes and run jobs in the Amazon cloud.

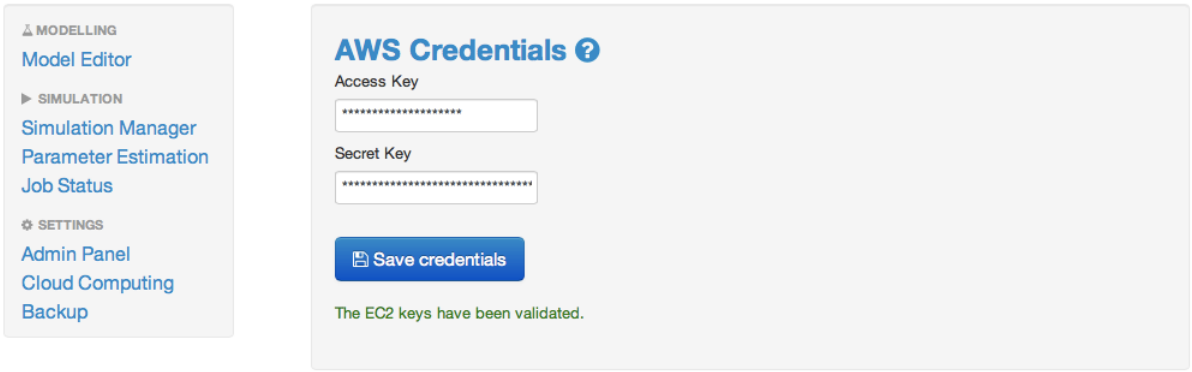

Figure 6.1: Cloud Computing page - Credentials section

1. Navigate to the main **Cloud Computing** page.
2. Copy your access key to the **Access Key** text box.
3. Copy your secret key to the **Secret Key** text box.

4. Click **Save credentials** to validate and save your credentials.

Launching and Shutting Down Nodes

Once valid credentials are entered, clicking **Launch nodes** (Figure 7.2) by default will launch one c3.large Amazon instance. This first node is designated the **head node**. Any other nodes launched are called **worker nodes**. There can be zero or more worker nodes, and they are c3.larges instance types by default. For information on Amazon instance types, look here: `http://aws.amazon.com/ec2/instance-types/`. The head node must be a c3.large or c3.xlarge. The worker nodes can be any combination of t1.micro, m1.small, m3.medium, m3.large, c3.large, and c3.xlarge nodes (chosen in the **Advanced settings** menu).

Only one head node is needed to run jobs in the cloud. It is possible to access cloud data even when no head nodes are launched. Compute resources and storage resources are billed separately by Amazon. More details can be found at: `http://aws.amazon.com/ec2/pricing/`.

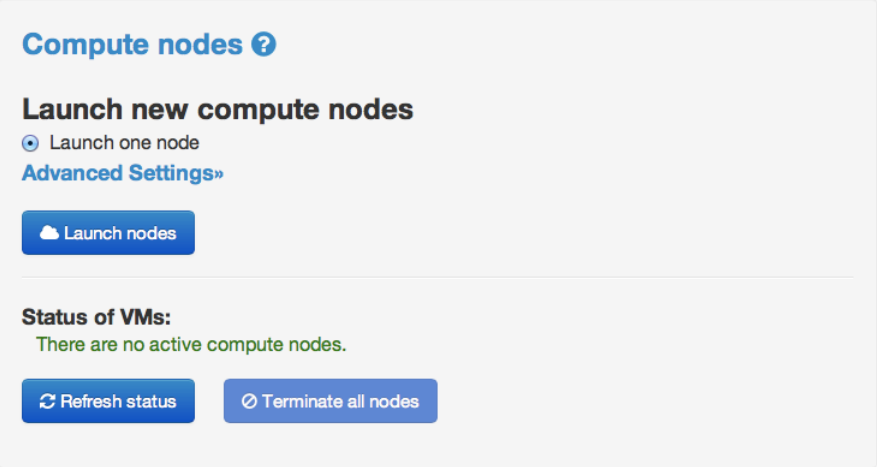

Figure 6.2: Cloud Computing page - Compute nodes default setting section

Launching a node takes time. The **Refresh status** button can be used to check the launch progress. The **Terminate all nodes** button terminates all the nodes that StochSS started.

6.2 Job Reproduction

StochSS provides the flexibility to store simulation output in the cloud or delete it and regenerate it later. If simulations are fast but produce large amounts of data, reproducing data only when it is needed can save money.

An Example on Job Reproduction

Reproducing a cloud job is simple:

1. Launch a compute node.

2. Run a well-mixed or spatial model (everything except parameter estimation jobs can be reproduced).

3. Navigate to the **Job Status** page.

4. Click **view** beside the job you would like to reproduce.

5. Click **Delete Output** to delete output in the cloud. No reproduction action is available until you delete the output.

6. Once the output is deleted, the option to reproduce the job will be appear as shown in Figure 7.3.

7. Choose a node type for reproduction. If there is no such instance type running, a warning will show up to guide you to the **Cloud Computing** page to launch one.

8. Click **Reproduce Results** to submit the reproduction request. This will automatically redirect to the **Job Status** page where the new job's status can be monitored.

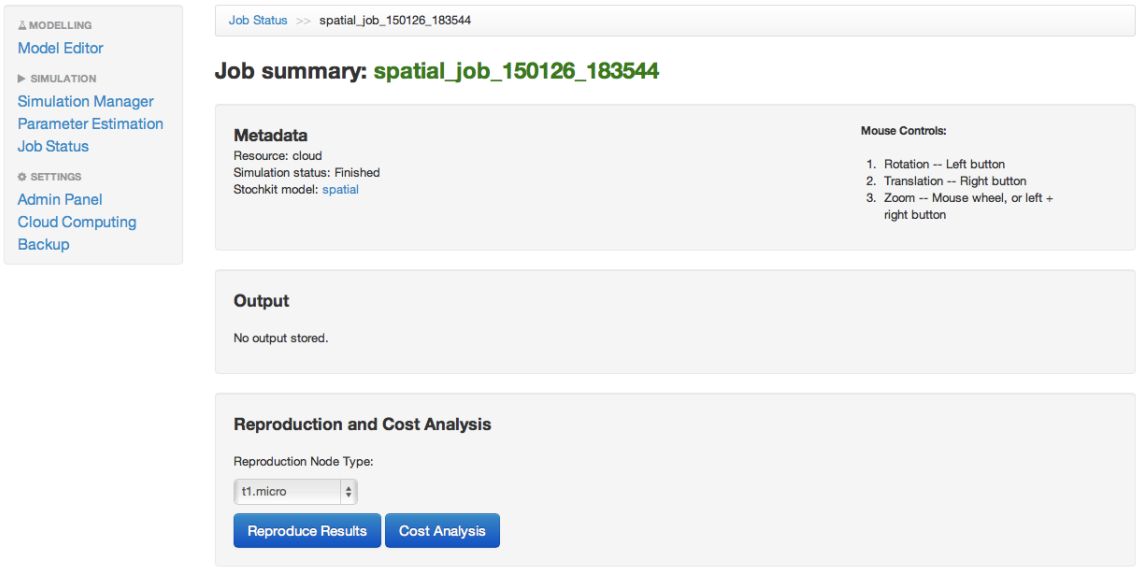

Figure 6.3: Job summary page - reproduction available

6.3 Cost Analysis

Because different instance types cost different amounts of money, it is not obvious which nodes are the cheapest for any given job type. StochSS allows manual measurement of job cost with the cost-analysis tool.

Figure 6.4: Cost Analysis page

An Example on Job Reproduction

1. Run a well-mixed or spatial model.
2. Navigate to the **Job Status** page.
3. Click **view** beside the job you would like to analyze.
4. Click **Cost Analysis** in the **Reproduction and Cost Analysis** section.
5. By default, cost analysis should be available for whatever instance type the job was run on.
6. Click **Analyze** with any other node type you would like to run and analyze the job on. At least one node of this instance type must already be running.
7. The run times and costs of simulating the jobs are plotted on the screen as in Figure 6.4.

7

Cluster Computing

StochSS provides an option to run jobs using traditional high performance computing infrastructure, such as university clusters. In order to use a cluster, you will need to upload your SSH secret key to the cluster you wish to use.

7.1 Cluster Prerequisites

The cluster you wish to use with StochSS must have the following:

1. **qsub** - To create a job is to submit an executable script to a batch server. StochSS uses qsub to submit jobs for execution on a cluster. For more information, see `http://docs.adaptivecomputing.com/torque/4-0-2/Content/topics/commands/qsub.htm`.
2. **Docker** - StochSS executes jobs on a cluster inside Docker containers. Docker containers provide a means to package up the complex software of StochSS jobs into a portable format. For more information, see `https://www.docker.com/`. In order to run a Docker container, operating system privileges are required. For any user, these are obtained by adding the user to the **docker group** on the cluster. *Please contact your cluster administer to obtain this permission and add your username to the docker group.*

7.2 Setting Credentials in StochSS

You must provide your credentials in StochSS.

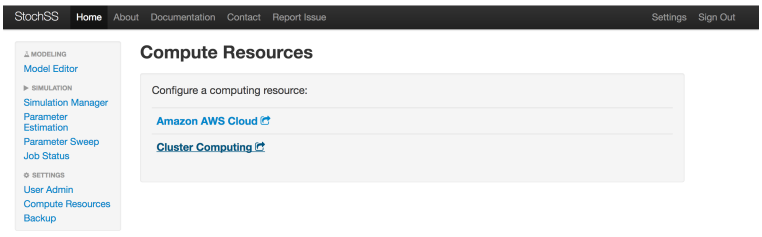

Figure 7.1: Compute Resources Page

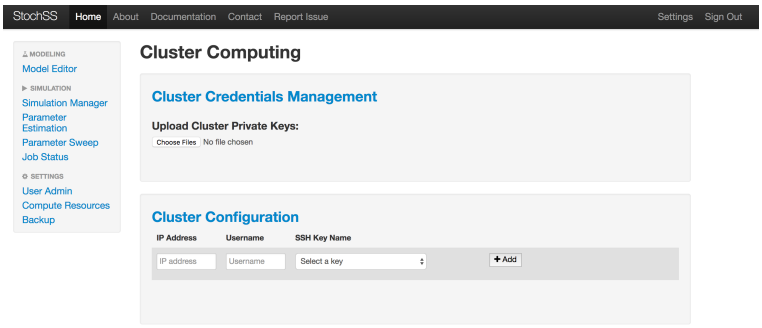

Figure 7.2: Configure Cluster Credentials

1. Navigate to the main **Compute Resources** page. (Figure 7.1)
2. Click on **Cluster Computing**.
3. On the **Configure Cluster Credentials** page (Figure 7.2):
   a) Under *Cluster Credentials Management*, click on **Choose Files** and upload your SSH key to the cluster.
   b) Under *Cluster Configurations*, enter your username associated with the previously uploaded SSH key and the ip address of the cluster submit node. Click **Add**.

7.3 Running a job on a cluster

After setting your credentials in StochSS, it is possible to use a cluster for the following types of experiments:

1. Parameter Sweeps
2. Simulations

An Example of performing an experiment on a cluster

The following steps illustrate running a parameter sweep on a cluster. It is assumed that cluster credentials have been uploaded to StochSS, as described in section 7.2.

1. From the menu displayed in the left hand column of the screen, click on *Parameter Sweep*.
2. Select the model you wish to use and click next. (Figure 7.3)
3. On the *Run Parameter Sweep* page, configure different parameters of the experiment, if required and then at the bottom of the page, under *Choose a compute resource*, click the drop-down menu and select the cluster you previously configured in StochSS. (Figure 7.4)
4. Click run to start the experiment on the cluster.
5. You will be automatically redirected to the results page when they are available.

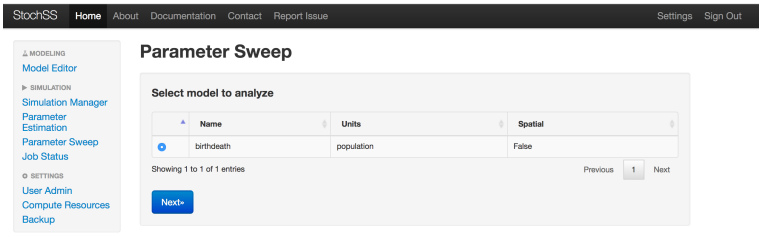

Figure 7.3: Select the model you wish to use and click next

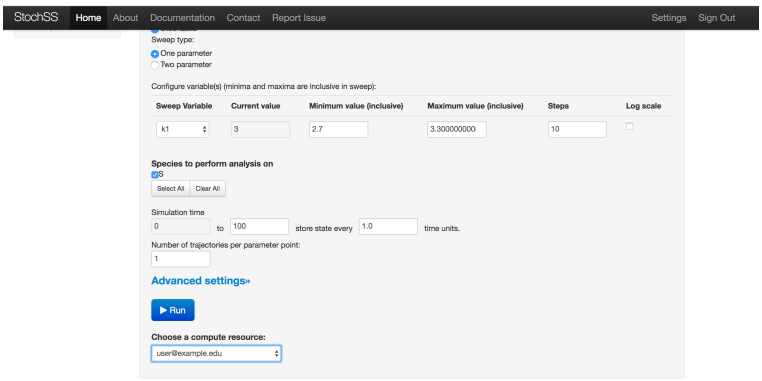

Figure 7.4: Select the cluster you wish to use