

# SB pipe documentation

Release 2.7.0

# Piero Dalle Pezze and Nicolas Le Novère

# CONTENTS

1	User	manual	1			
	1.1	Introduction	1			
		1.1.1 Requirements	1			
		1.1.2 Installation	4			
	1.2	How to use SBpipe	4			
		1.2.1 Preliminary configuration steps	4			
		1.2.2 Running SBpipe	5			
		1.2.3 Pipeline configuration files	6			
	1.3	Reporting bugs or requesting new features	9			
2	Deve	loper manual	11			
	2.1	Introduction	11			
	2.2	Development model	11			
		2.2.1 Conventions	11			
		2.2.2 Work flow	11			
		2.2.3 New releases	12			
	2.3	Package structure	12			
		2.3.1 docs	13			
		2.3.2 sbpipe	13			
		2.3.3 scripts	14			
		2.3.4 tests	15			
	2.4	Miscellaneous of useful commands	15			
		2.4.1 Git	15			
3	Sour	ce code	17			
	3.1	Python modules	17			
		3.1.1 sbpipe package	17			
4	Meta	a information	37			
	4.1	Copyright	37			
5	Indic	ees	39			
Py	Python Module Index					
Inc	Index					

# **USER MANUAL**

Copyright © 2015-2018, Piero Dalle Pezze and Nicolas Le Novère.

SBpipe and its documentation are released under the GNU Lesser General Public License v3 (LGPLv3). A copy of this license is provided with the package and can also be found here: https://www.gnu.org/licenses/lgpl-3.0.txt.

Contacts: Dr Piero Dalle Pezze (piero.dallepezze AT babraham.ac.uk) and Dr Nicolas Le Novère (lenov AT babraham.ac.uk)

Affiliation: The Babraham Institute, Cambridge, CB22 3AT, UK

Mailing list: sbpipe AT googlegroups.com

Forum: https://groups.google.com/forum/#!forum/sbpipe

# Introduction

This package contains a collection of pipelines for dynamic modelling of biological systems. It aims to automate common processes and speed up productivity for tasks such as model simulation, single/double parameter scan, and parameter estimation.

# Requirements

In order to use SBpipe, the following software must be installed:

- Python 2.7+ https://www.python.org/
- R 3.3.0+ https://cran.r-project.org/

SBpipe can work with the following simulators (at least one must be installed):

- Copasi 4.19+ http://copasi.org/ (for model simulation, parameter scan, and parameter estimation)
- Any R / Python / Octave / Java simulator (for model simulation. Users must install the dependencies)

If LaTeX/PDF reports are also desired, the following software must also be installed:

• LaTeX 2013

Depending on your operating system, LaTeX can be downloaded at these websites:

• GNU/Linux: https://latex-project.org/ftp.html

• Windows: https://miktex.org/

#### **GNU/Linux**

It is advised that users install Python, R and (optionally) LaTeX packages using the package manager of their GNU/Linux distribution. Users need to make sure that the packages python-pip and

texlive-latex-base (only for reports). In most cases, the installation via the package manager will automatically configure the correct environment variables.

If a local installation of Python, R, or LaTeX is needed, users need to add the following environment variables to \$PATH in their \$HOME/.bashrc file as follows:

```
# Path to R
export PATH=$PATH:/path/to/R/binaries/

# Path to Python. Scripts is the folder (if any) containing the Python
# script `pip`. pip must be available via command line.
export PATH=$PATH:/path/to/Python/:/path/to/Python/Scripts/

# Path to LaTeX
export PATH=$PATH:/path/to/LaTeX/binaries/
```

The correct installation of Python, R, and LaTeX can be tested by running the commands:

```
# If variables were manually exported, reload the .bashrc file
$ source $HOME/.bashrc

$ python -V
Python 2.7.12
$ pip -V
pip 8.1.2 from /home/ariel/.local/lib/python2.7/site-packages (python 2.7)

$ R --version
R version 3.2.3 (2015-12-10) -- "Wooden Christmas-Tree"
Copyright (C) 2015 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

$ pdflatex -v
pdfTeX 3.14159265-2.6-1.40.16 (TeX Live 2015/Debian)
kpathsea version 6.2.1
Copyright 2015 Peter Breitenlohner (eTeX)/Han The Thanh (pdfTeX).
```

As of 2016, Copasi is not available as a package in GNU/Linux distributions. Users must add the path to Copasi binary files manually editing their GNU/Linux \$HOME/.bashrc file as follows:

```
# Path to CopasiSE
export PATH=$PATH:/path/to/CopasiSE/
```

The correct installation of CopasiSE can be tested by running the command:

```
# Reload the .bashrc file
$ source $HOME/.bashrc

$ CopasiSE -h
COPASI 4.16 (Build 104)
```

At this stage, Python, R, Copasi, and (optionally) LaTeX should be installed correctly. SBpipe requires the configuration of the environment variable \$SBPIPE which must also be added in the \$HOME/.bashrc file. The package also needs to be added to \$PATH. To do so, users need to add the following lines to their \$HOME/.bashrc file:

```
# SBPIPE
export SBPIPE=/path/to/sbpipe
export PATH=$PATH:$SBPIPE/scripts
```

Now you should reload the .bashrc file to make the previous change effective:

```
# Reload the .bashrc file
$ source $HOME/.bashrc
```

Before testing the correct installation of SBpipe, users need to install Python and R dependency packages used by SBpipe. Two scripts are provided to perform these tasks automatically.

To install SBpipe Python dependencies on GNU/Linux, run:

```
$ cd $SBPIPE/
$ ./install_pydeps.py
```

To install SBpipe R dependencies on GNU/Linux, run:

```
$ cd $SBPIPE/
$ R
# Inside R environment, answer 'y' to install packages locally
> source('install_rdeps.r')
```

If R package dependencies must be compiled, it is worth checking that the following additional packages are installed in your machine: build-essential, liblapack-dev, libbas-dev, libcairo-dev, libsal-dev, libcurl4-openssl-dev, and gfortran. After installing these packages, install\_rdeps.r must be executed again.

The correct installation of SBpipe can be tested by running the command:

```
$ sbpipe.py -v
2.1.0
```

#### **Windows**

Windows users are also strongly advised to install the package:

• Cygwin 2.6.0 https://www.cygwin.com/

Cygwin offers a GNU/Linux-like shell. This makes the installation of dependencies easier as this follows the configuration for GNU/Linux users.

Windows users may need to edit the PATH environment variable so that the binary files for the previous packages (Copasi, Python, R, and (optionally) LaTeX) are correctly found. Specifically for Python, the python scripts pip.py and easy\_install.py are located inside the folder Scripts within the Python root directory. The path to this folder must also be added to PATH.

Therefore, the following environment variables must also be added:

```
SBPIPE=\path\to\sbpipe
PATH=[previous paths];%SBPIPE%\scripts
```

**NOTE for Cygwin:** Environment variables can also be configured directly within the .bashrc file in cygwin/home/USERNAME/. In the beginning of this file, users should place:

```
# Path to R
export PATH=$PATH:/path/to/R/binaries/

# Path to Python
export PATH=$PATH:/path/to/Python/:/path/to/Python/Scripts/

# Path to LaTeX
export PATH=$PATH:/path/to/LaTeX/binaries/

# Path to CopasiSE
export PATH=$PATH:/path/to/CopasiSE/binaries/

# SBPIPE
export SBPIPE=/path/to/sbpipe
export PATH=$PATH:$SBPIPE/scripts
```

After configuring the environment variables directly or internally in Cygwin, the next step is to install Python and R packages used by SBpipe. Two scripts are provided to perform these tasks automatically.

1.1. Introduction 3

To install SBpipe Python dependencies using Cygwin on Windows, run:

```
$ cd /cygdrive/PATH/TO/SBPIPE/
$ python.exe install_pydeps.py
```

To install SBpipe R dependencies using Cygwin on Windows, run:

```
$ cd /cygdrive/PATH/TO/SBPIPE/
$ R.exe
# Inside R environment, answer 'y' to install packages locally
> source('install_rdeps.r')
```

# Installation

If desired, SBpipe can be installed in your system. To do so, run the command inside the sbpipe folder:

```
$ cd $SBPIPE
$ python setup.py install
```

The correct installation of SBpipe and its dependencies can be checked by running the following commands inside the SBpipe folder:

```
$ cd $SBPIPE/tests
$ ./test_suite.py
```

# How to use SBpipe

# **Preliminary configuration steps**

#### **Pipelines using Copasi**

Before using these pipelines, a Copasi model must be configured as follow using CopasiUI:

#### pipeline: simulation

- Tick the flag *executable* in the Time Course Task.
- Select a report template for the Time Course Task.
- Save the report in the same folder with the same name as the model but replacing the extension .cps with .csv.

#### pipelines: single or double parameter scan

- Tick the flag executable in the Parameter Scan Task.
- Select a report template for the Parameter Scan Task.
- Save the report in the same folder with the same name as the model but replacing the extension .cps with .csv.

# pipeline: parameter estimation

- Tick the flag *executable* in the Parameter Estimation Task.
- Select the report template for the Parameter Estimation Task.
- Save the report in the same folder with the same name as the model but replacing the extension .cps with .csv.

For tasks such as parameter estimation using Copasi, it is recommended to move the data set into the folder Models/ so that the Copasi model file and its associated experimental data files are stored in the same folder.

#### Pipelines using R, Python, Octave, or Java

### pipeline: simulation

- The program must be a functional and invokable via Rscript, python, octave, or java -jar, respectively.
- The Jar file for Java models must include a manifest.mf specifying the main class.
- The program must receive the report file name as input argument (see examples in \$SBPIPE/tests/).
- The program must save the report to file including the *Time* column. Report fields must be separated by TAB, and row names must be discarded.

#### pipeline: parameter estimation

- The program must be a functional and invokable via Rscript, python, octave, or java -jar, respectively.
- The Jar file for Java models must include a manifest.mf specifying the main class.
- The program must receive the report file name as input argument (see examples in \$SBPIPE/tests/).
- The program must save the report to file. This includes the objective value as first column column, and the estimated parameters as following columns. Rows are the evaluated functions. Report fields must be separated by TAB, and row names must be discarded.

# **Running SBpipe**

SBpipe is executed via the command *sbpipe.py*. The syntax for this command and its complete list of options can be retrieved by running *sbpipe.py* -h.

As of Sep 2016 the output is as follows:

```
$ sbpipe.py -h
Usage: sbpipe.py [OPTION] [FILE]
Pipelines for systems modelling of biological networks.
List of mandatory options:
        -h, --help
                Show this help.
        -c, --create-project
                Create a project structure using the argument as name.
        -s, --simulate
                Simulate a model.
        -p, --single-param-scan
                Simulate a single parameter scan.
        -d, --double-param-scan
                Simulate a double parameter scan.
        -e, --param-estim
                Generate a parameter fit sequence.
        -1, --license
                Show the license.
        -v, --version
                Show the version.
Exit status:
0 if OK,
1 if minor problems (e.g., a pipeline did not execute correctly),
 2 if serious trouble (e.g., cannot access command-line argument).
Report bugs to sbpipe@googlegroups.com
SBpipe home page: <a href="https://pdp10.github.io/sbpipe">https://pdp10.github.io/sbpipe</a>
For complete documentation, see README.md .
```

The first step is to create a new project. This can be done with the command:

```
$ sbpipe.py --create-project project_name
```

This generates the following structure:

Models must be stored in the Models/ folder. The folder Data/ is meant for collecting experimental data files and analyses in one place. Regarding Copasi, once the data files (e.g. for parameter estimation) are generated, **it is advised** to move them into the Models/ folder so that the Copasi (.cps) file and its associated experimental data files are stored in the same folder. To run SBpipe, users need to create a configuration file for each pipeline they intend to run (see next section). These configuration files should be placed in the Working\_Folder/. This folder will eventually contain all the results generated by SBpipe.

For instance, the pipeline for parameter estimation configured with a certain configuration file can be executed by typing:

```
$ cd project_name/Working_Folder/
$ sbpipe.py -e my_config_file.conf
```

# Pipeline configuration files

Pipelines are configured using files (here called configuration files). These files are INI files and are therefore structured as follows:

```
[pipeline_name]
option1=value1
option2=value2
...
```

where pipeline\_name can be: - simulate, for deterministic or stochastic model simulation; - single\_param\_scan, for scanning one model parameter; - double\_param\_scan, for scanning two model parameters; - param\_estim, for parameter estimation.

In SBpipe each pipeline executes three tasks: data generation, data analysis, and report generation. These tasks can be activated in each configuration files using the options:

- generate\_data=True
- analyse\_data=True
- generate\_report=True

The <code>generate\_data</code> task runs a simulator accordingly to the options in the configuration file. Hence, this task collects and organises the reports generated from the simulator. The <code>analyse\_data</code> task processes the reports to generate plots and compute statistics. Finally, the <code>generate\_report</code> task generates a LaTeX report containing the computed plots and invokes the utility <code>pdflatex</code> to produce a PDF file. This modularisation allows users to analyse the same data without having to re-generate it, or to skip the report generation if not wanted.

Pipelines for parameter estimation or stochastic model simulation can be computationally intensive. SBpipe allows users to generate simulated data in parallel using the following options in the pipeline configuration file:

- cluster=pp
- pp\_cpus=7
- runs=250

The cluster option defines whether the simulator should be executed locally (pp: Parallel Python), or in a computer cluster (sge: Sun Grid Engine (SGE), lsf: Load Sharing Facility (LSF)). If pp is selected, the pp\_cpus option determines the maximum number of CPUs to be allocated for local simulations. The runs option specifies the number of simulations (or parameter estimations for the pipeline param\_estim) to be run.

Assuming that the configuration files are placed in the Working\_Folder of a certain project, examples are given as follow:

#### **Example 1:** configuration file for the pipeline *simulation*

```
[simulate]
# True if data should be generated, False otherwise
generate_data=True
# True if data should be analysed, False otherwise
analyse_data=True
# True if a report should be generated, False otherwise
generate_report=True
# The relative path to the project directory (from Working_Folder)
project_dir=..
# The name of the configurator (e.g. Copasi, Rscript, Python, Octave, Java)
simulator=Copasi
# The model name
model=insulin_receptor_stoch.cps
# The cluster type. pp if the model is run locally,
# sqe/lsf if run on cluster.
cluster=pp
# The number of CPU if pp is used, ignored otherwise
pp_cpus=7
# The number of simulations to perform.
\# n>=1 for stochastic simulations.
runs=40
# An experimental data set (or blank) to add to the
# simulated plots as additional layer
exp_dataset=insulin_receptor_dataset.csv
# True if the experimental data set should be plotted.
plot_exp_dataset=True
# The label for the x axis.
xaxis_label=Time [min]
# The label for the y axis.
yaxis_label=Level [a.u.]
```

#### **Example 2:** configuration file for the pipeline *single parameter scan*

```
[single_param_scan]
# True if data should be generated, False otherwise
generate_data=True
# True if data should be analysed, False otherwise
analyse_data=True
# True if a report should be generated, False otherwise
generate_report=True
# The relative path to the project directory (from Working_Folder)
project_dir=..
# The name of the configurator (e.g. Copasi)
simulator=Copasi
# The model name
model=insulin_receptor_inhib_scan_IR_beta.cps
# The variable to scan (as set in Copasi Parameter Scan Task)
scanned_par=IR_beta
# The cluster type. pp if the model is run locally,
# sge/lsf if run on cluster.
cluster=pp
# The number of CPU if pp is used, ignored otherwise
pp_cpus=7
# The number of simulations to perform per run.
\# n>=1 for stochastic simulations.
runs=1
# The number of intervals in the simulation
simulate intervals=100
# True if the variable is only reduced (knock down), False otherwise.
```

```
single_param_scan_knock_down_only=True
# True if the scanning represents percent levels.
single_param_scan_percent_levels=True
# The minimum level (as set in Copasi Parameter Scan Task)
min level=0
# The maximum level (as set in Copasi Parameter Scan Task)
max_level=100
# The number of scans (as set in Copasi Parameter Scan Task)
levels_number=10
# True if plot lines are the same between scans
# (e.g. full lines, same colour)
homogeneous_lines=False
# The label for the x axis.
xaxis_label=Time [min]
# The label for the y axis.
yaxis_label=Level [a.u.]
```

#### **Example 3:** configuration file for the pipeline *double parameter scan*

```
[double_param_scan]
# True if data should be generated, False otherwise
generate_data=True
# True if data should be analysed, False otherwise
analyse data=True
# True if a report should be generated, False otherwise
generate_report=True
# The relative path to the project directory (from Working_Folder)
project_dir=..
# The name of the configurator (e.g. Copasi)
simulator=Copasi
# The model name
model=insulin_receptor_inhib_dbl_scan_InsulinPercent__IRbetaPercent.cps
# The 1st variable to scan (as set in Copasi Parameter Scan Task)
scanned_par1=InsulinPercent
# The 2nd variable to scan (as set in Copasi Parameter Scan Task)
scanned_par2=IRbetaPercent
# The cluster type. pp if the model is run locally,
# sge/lsf if run on cluster.
cluster=pp
# The number of CPU if pp is used, ignored otherwise
pp_cpus=7
# The number of simulations to perform.
\# n>=1 for stochastic simulations.
runs=1
# The simulation length (as set in Copasi Time Course Task)
sim_length=10
```

## **Example 4:** configuration file for the pipeline *parameter estimation*

```
[param_estim]
# True if data should be generated, False otherwise
generate_data=True
# True if data should be analysed, False otherwise
analyse_data=True
# True if a report should be generated, False otherwise
generate_report=True
# True if a zipped tarball should be generated, False otherwise
generate_tarball=True
# The relative path to the project directory (from Working_Folder)
project_dir=..
# The name of the configurator (e.g. Copasi, Rscript, Python, Octave, Java)
simulator=Copasi
# The model name
```

```
model=insulin_receptor_param_estim.cps
# The cluster type. pp if the model is run locally,
# sge/lsf if run on cluster.
cluster=pp
# The number of CPU if pp is used, ignored otherwise
pp_cpus=7
# The parameter estimation round which is used to distinguish
# phases of parameter estimations when parameters cannot be
# estimated at the same time
round=1
# The number of parameter estimations
# (the length of the fit sequence)
runs=250
# The threshold percentage of the best fits to consider
best_fits_percent=75
# The number of available data points
data_point_num=33
# True if 2D all fits plots for 66% confidence levels
# should be plotted. This can be computationally expensive.
plot_2d_66cl_corr=True
\mbox{\#} True if 2D all fits plots for 95% confidence levels
# should be plotted. This can be computationally expensive.
plot_2d_95cl_corr=True
# True if 2D all fits plots for 99% confidence levels
# should be plotted. This can be computationally expensive.
plot_2d_99cl_corr=True
# True if parameter values should be plotted in log space.
logspace=True
# True if plot axis labels should be plotted in scientific notation.
scientific_notation=True
```

Additional examples of configuration files can be found in:

```
$SBPIPE/tests/insulin_receptor/Working_Folder/
```

# Reporting bugs or requesting new features

SBpipe is a relatively young project and there is a chance that some error occurs. The following mailing list should be used for general questions:

```
sbpipe AT googlegroups.com
```

All the topics discussed in this mailing list are also available at the website:

https://groups.google.com/forum/#!forum/sbpipe

To help us better identify and reproduce your problem, some technical information is needed. This detail data can be found in SBpipe log files which are stored in \${HOME}/.sbpipe/logs/. When using the mailing list above, it would be worth providing this extra information.

Issues and feature requests can also be notified using the github issue tracking system for SBpipe at the web page:

https://github.com/pdp10/sbpipe/issues.

**CHAPTER** 

**TWO** 

# **DEVELOPER MANUAL**

Mailing list: sbpipe AT googlegroups.com

Forum: https://groups.google.com/forum/#!forum/sbpipe

# Introduction

This guide is meant for developers and contains guidelines for developing this project.

# **Development model**

This project follows the Feature-Branching model. Briefly, there are two main branches: master and develop. The former contains the history of stable releases, the latter contains the history of development. The master branch contains checkout points for production hotfixes or merge points for release-x.x.x branches. The develop branch is used for feature-bugfix integration and checkout point in development. Nobody should directly develop in here. The develop branch is versionless (just call it -dev).

#### **Conventions**

To manage the project in a more consistent way, here is a list of conventions to follow:

- Each new feature is developed in a separate branch forked from *develop*. This new branch is called *featureNUMBER*, where *NUMBER* is the number of the GitHub Issue discussing that feature. The first line of each commit message for this branch should contain the string *Issue #NUMBER* at the beginning. Doing so, the commit is automatically recorded by the Issue Tracking System for that specific Issue. Note that the sharp (#) symbol is required.
- The same for each new bugfix, but in this case the branch name is called bugfixNUMBER.
- The same for each new hotfix, but in this case the branch name is called hotfixNUMBER and is forked from *master*.

# Work flow

The procedure for checking out a new feature from the develop branch is:

```
$ git checkout -b feature10 develop
```

This creates the feature10 branch off develop. This feature10 is discussed in *Issue #10* in GitHub. When you are ready to commit your work, run:

```
$ git commit -am "Issue #10, summary of the changes. Detailed description of the changes, if any."
$ git push origin feature10  # sometimes and at the end.
```

As of June 2016, the branches master and develop are protected and a status check using Travis-CI must be performed before merging or pushing into these branches. This automatically forces a merge without fast-forward. In order to merge any new feature, bugfix or simple edits into master or develop, a developer must checkout a new branch and, once committed and pushed, merge it to master or develop using a pull request. To merge feature10 to develop, the pull request output will look like this in GitHub Pull Requests:

```
base:develop compare:feature10 Able to merge. These branches can be automatically merged.
```

A small discussion about feature 10 should also be included to allow other users to understand the feature.

Finally delete the branch:

```
$ git branch -d feature10  # delete the branch feature10 (locally)
```

#### **New releases**

When the develop branch includes all the desired feature for a release, it is time to checkout this branch in a new one called release-x.x.x. It is at this stage that a version is established. Only bugfixes or hotfixes are applied to this branch. When this testing/correction phase is completed, the master branch will merge with the release-x.x.x branch, using the commands above. To record the release add a tag:

```
git tag -a v1.3 -m "PROGRAM_NAME v1.3"
```

To transfer the tag to the remote server:

```
git push origin v1.3 # Note: it goes in a separate 'branch'
```

To see all the releases:

```
git show
```

# Package structure

This section presents the structure of the SBpipe package. The root of the project contains general management scripts for installing Python and R dependencies (install\_pydeps.py and install\_rdeps.r), and installing SBpipe (setup.py). Additionally, the logging configuration file (logging\_config.ini) is also at this level.

In order to automatically compile and run the test suite, Travis-CI is used and configured accordingly (.travis.yml).

The project is structured as follows:

These folders will be discussed in the next sections. In SBpipe, Python is the project main language. Instead, R is essentially used for computing statistics (see section configuration file in the user manual) and for generating plots. This choice allows users to run these scripts independently of SBpipe if needed using an R environment like Rstudio. This can be convenient if further data analysis are needed or plots need to be annotated or edited.

#### docs

The folder docs/contains the documentation for this project. The user and developer manuals in markdown format are contained in docs/source. In order to generate the complete documentation for SBpipe, the following packages must be installed:

- · python-sphinx
- pandoc
- · texlive-fonts-recommended
- texlive-latex-extra

By default the documentation is generated in html and LaTeX/PDF. Instruction for generating or cleaning SBpipe documentation are provided below.

To generate the source code documentation:

```
$ cd $SBPIPE/docs
$ ./gen_doc.sh
```

#### To clean the documentation:

```
$ cd $SBPIPE/docs
$ ./cleanup_doc.sh
```

The complete source code documentation for this project is stored in docs/build/html (html format) and docs/build/latex (LaTeX/PDF format). A shortcut to the documentation in html format is available at the page docs/index.html.

UML class diagrams for the packages sbpipe.pl and sbpipe.simul are stored in docs/source/uml\_diagrams.

### sbpipe

This folder contains the source code of the project SBpipe. At this level a file called \_\_main\_\_.py enables users to run SBpipe programmatically as a Python module via the command:

```
$ python sbpipe
```

Alternatively sbpipe can programmatically be imported within a Python environment as shown below:

```
$ cd $SBPIPE
$ python
# Python environment
>>> import sbpipe.main as sb
>>> sb.version()
'2.0.0'
```

The following subsections describe sbpipe subpackages.

# рl

The subpackage sbpipe.pl contains the class Pipeline in the file pipeline.py. This class represents a generic pipeline which is extended by SBpipe pipelines. These are organised in the following subpackages:

- create: creates a new project
- ps1: scan a model parameter, generate plots and report;
- ps2: scan two model parameters, generate plots and report;
- pe: generate a parameter fit sequence, tables of statistics, plots and report;
- sim: generate deterministic or stochastic model simulations, plots and report.

All these pipelines can be invoked directly via the script \$SBPIPE/scripts/sbpipe.py. Each SBpipe pipeline extends the class Pipeline and therefore must implement the following methods:

```
# executes a pipeline
def run(self, config_file)

# process the configuration file lines
# imported by Pipeline.config_parser()
def read_config(self, lines)
```

The former contains the procedure to parse an INI configuration file. The latter parses the pipeline options. The Pipeline class implements the INI parser which is therefore available to each pipeline. The INI parser returns the configuration file as a list of lines.

#### R

This folder contains a collection of R utility methods for plotting and generating statistics. These utilities are used by the pipelines during data analysis.

## report

The subpackage sbpipe.report contains Python modules for generating LaTeX/PDF reports.

#### simul

The subpackage sbpipe.simul contains the class Simul in the file simul.py. This is a generic simulator interface used by the pipelines in SBpipe. This mechanism uncouples pipelines from specific simulators which can therefore be configured in each pipeline configuration file. As of 2016, the following simulators are available in SBpipe:

- Copasi, package sbpipe.simul.copasi, which implements all the methods of the class Simul;
- Rscript, package sbpipe.simul.rscript;
- Python, package sbpipe.simul.python;
- Octave, package sbpipe.simul.octave;
- Java, package sbpipe.simul.java.

Rscript, Python, Octave, and Java only implement the methods sim() and pe() of Simul (time course simulation and parameter estimation). Pipelines can dynamically load a simulator via the class method Pipeline.get\_simul\_obj(simulator). This method instantiates an object of subtype Simul by refractoring the simulator name as parameter. A simulator class (e.g. Copasi) must have the same name of their package (e.g. copasi) but start with an upper case letter. A simulator class must be contained in a file with the same name of their package (e.g. copasi). Therefore, for each simulator package, exactly one simulator class can be instantiated. Simulators can be configured in the configuration file using the field simulator.

## utils

The subpackage sbpipe.utils contains a collection of Python utility modules which are used by sbpipe. Here are also contained the functions for running commands in parallel.

### scripts

The folder scripts contains the scripts: cleanup\_sbpipe.py and sbpipe.py. sbpipe.py is the main script and is used to run the pipelines. cleanup\_sbpipe.py is used for cleaning the package including the test results.

#### tests

The package tests contains the script test\_suite.py which executes all sbpipe tests. It should be used for testing the correct installation of SBpipe dependencies as well as reference for configuring a project before running any pipeline. Projects inside the folder \$SBPIPE/tests/ have the SBpipe project structure:

- Data: (e.g. training / testing data sets for the model);
- Model: (e.g. models, Copasi models, R-Python-Octave-Java models, data sets directly used by Copasi models);
- Working\_Folder: (e.g. pipelines configurations and parameter estimation results, time course, parameter scan, etc).

Examples of configuration files (\*.conf) using Copasi can be found in \$SBPIPE/tests/insulin\_receptor/Working\_Folder/.

To run tests for R models, the R packages deSolve, sde, rshape2, and minpack.lm must be installed. To run tests for Python models, the Python packages numpy, scipy, and pandas must be installed. These additional dependencies should not be included to SBpipe main dependencies as they are user-specific. In principle, users may define their R, Python, Octave or Java models using arbitrary packages.

As of 2016, the repository for SBpipe source code is github.com. This is configured to run Travis-CI every time a git push into the repository is performed. The exact details of execution of Travis-CI can be found in Travis-CI configuration file \$SBPIPE/.travis.yml. Importantly, Travis-CI runs all SBpipe tests using nosetests.

# Miscellaneous of useful commands

#### Git

#### Startup

```
# clone master
$ git clone https://github.com/pdp10/sbpipe.git
# get develop branch
$ git checkout -b develop origin/develop
# to get all the other branches
$ for b in `git branch -r | grep -v -- '->'`; do git branch
--track ${b##origin/} $b; done
# to update all the branches with remote
$ git fetch --all
```

#### **Update**

```
# ONLY use --rebase for private branches. Never use it for shared # branches otherwise it breaks the history. --rebase moves your # commits ahead. For shared branches, you should use # `git fetch && git merge --no-ff` $ git pull [--rebase] origin BRANCH
```

# File system

```
$ git rm [--cache] filename
$ git add filename
```

#### **Information**

```
$ git status
$ git log [--stat]
$ git branch # list the branches
```

#### Maintenance

```
$ git fsck  # check errors
$ git gc  # clean up
```

# Rename a branch locally and remotely

```
git branch -m old_branch new_branch  # Rename branch locally
git push origin :old_branch  # Delete the old branch
git push --set-upstream origin new_branch  # Push the new branch, set
local branch to track the new remote
```

#### Reset

```
git reset --hard HEAD  # to undo all the local uncommitted changes
```

#### Syncing a fork (assuming upstreams are set)

```
git fetch upstream
git checkout develop
git merge upstream/develop
```

# **THREE**

# **SOURCE CODE**

# **Python modules**

# sbpipe package

**Subpackages** 

sbpipe.pl package

**Subpackages** 

sbpipe.pl.create package

**Submodules** 

### sbpipe.pl.create.newproj module

This module initialises the folder tree for a new project.

#### **Parameters**

- data\_folder the folder containing the data
- working\_folder the folder to store the results

run (project\_name)

Create a project directory tree.

Parameters project\_name - the name of the project
Returns 0

**Module contents** 

sbpipe.pl.pe package

**Submodules** 

#### sbpipe.pl.pe.parest module

This module provides the user with a complete pipeline of scripts for running model parameter estimations

```
classmethod analyse_data (simulator, model, inputdir, outputdir, fileout_final_estims, fileout_all_estims, fileout_param_estim_details, fileout_param_estim_summary, sim_plots_dir, best_fits_percent, data_point_num, plot_2d_66cl_corr=False, plot_2d_95cl_corr=False, logspace=True, scientific_notation=True)
```

The second pipeline step: data analysis.

#### **Parameters**

- **simulator** the name of the simulator (e.g. Copasi)
- model the model name
- inputdir the directory containing the simulation data
- outputdir the directory to store the results
- **fileout\_final\_estims** the name of the file containing final parameter sets with Chi^2
- **fileout\_all\_estims** the name of the file containing all the parameter sets with Chi^2
- **fileout\_param\_estim\_details** the name of the file containing the detailed statistics for the estimated parameters
- **fileout\_param\_estim\_summary** the name of the file containing the summary for the parameter estimation
- sim\_plots\_dir the directory of the simulation plots
- best\_fits\_percent the percent to consider for the best fits
- data\_point\_num the number of data points
- plot\_2d\_66cl\_corr True if 2 dim plots for the parameter sets within 66% should be plotted
- plot\_2d\_95cl\_corr True if 2 dim plots for the parameter sets within 95% should be plotted
- plot\_2d\_99cl\_corr True if 2 dim plots for the parameter sets within 99% should be plotted
- logspace True if parameters should be plotted in log space
- scientific\_notation True if axis labels should be plotted in scientific notation

**Returns** True if the task was completed successfully, False otherwise.

- **simulator** the name of the simulator (e.g. Copasi)
- model the model to process

- inputdir the directory containing the model
- cluster\_type pp for parallel python, lsf for load sharing facility, sge for sun grid engine
- pp\_cpus the number of cpu for parallel python
- nfits the number of fits to perform
- outputdir the directory to store the results
- sim\_data\_dir the directory containing the simulation data sets
- updated\_models\_dir the directory containing the models with updated parameters for each estimation

**Returns** True if the task was completed successfully, False otherwise.

classmethod generate\_report (model, outputdir, sim\_plots\_folder)

The third pipeline step: report generation.

#### **Parameters**

- model the model name
- outputdir the directory to store the report
- **sim\_plots\_folder** the folder containing the plots

**Returns** True if the task was completed successfully, False otherwise.

```
read_config (lines)
run (config_file)
```

#### **Module contents**

#### sbpipe.pl.ps1 package

#### **Submodules**

#### sbpipe.pl.ps1.parscan1 module

This module provides the user with a complete pipeline of scripts for computing single parameter scans.

```
classmethod analyse_data (model, scanned_par, knock_down_only, outputdir, sim_data_folder, sim_plots_folder, runs, percent_levels, min_level, max_level, levels_number, homogeneous_lines, xaxis_label, yaxis_label)
```

The second pipeline step: data analysis.

- model the model name
- scanned\_par the scanned parameter
- knock\_down\_only True for knock down simulation, false if also scanning over expression.
- outputdir the directory containing the results
- sim\_data\_folder the folder containing the simulated data sets

- sim\_plots\_folder the folder containing the generated plots
- runs the number of simulations
- **percent\_levels** True if the levels are percents.
- min level the minimum level
- max level the maximum level
- levels number the number of levels
- homogeneous\_lines True if generated line style should be homogeneous
- **xaxis** label the name of the x axis (e.g. Time [min])
- yaxis\_label the name of the y axis (e.g. Level [a.u.])

**Returns** True if the task was completed successfully, False otherwise.

#### **Parameters**

- **simulator** the name of the simulator (e.g. Copasi)
- model the model to process
- scanned\_par the scanned parameter
- cluster\_type pp for local Parallel Python, lsf for Load Sharing Facility, sge for Sun Grid Engine.
- **pp\_cpus** the number of CPU used by Parallel Python.
- runs the number of model simulation
- simulate intervals the time step of each simulation
- single\_param\_scan\_intervals the number of scans to perform
- inputdir the directory containing the model
- **outputdir** the directory to store the results

**Returns** True if the task was completed successfully, False otherwise.

**classmethod generate\_report** (*model*, *scanned\_par*, *outputdir*, *sim\_plots\_folder*)

The third pipeline step: report generation.

## **Parameters**

- model the model name
- **scanned\_par** the scanned parameter
- outputdir the directory containing the report
- **sim\_plots\_folder** the folder containing the plots

**Returns** True if the task was completed successfully, False otherwise.

```
read_config (lines)
run (config_file)
```

#### **Module contents**

#### sbpipe.pl.ps2 package

#### **Submodules**

#### sbpipe.pl.ps2.parscan2 module

```
class sbpipe.pl.ps2.parscan2.ParScan2 (data_folder='Data',
                                                                     models_folder='Models',
                                             working_folder='Working_Folder',
                                             sim_data_folder='double_param_scan_data',
                                             sim_plots_folder='double_param_scan_plots')
     Bases: sbpipe.pl.pipeline.Pipeline (page 23)
```

This module provides the user with a complete pipeline of scripts for computing double parameter scans.

classmethod analyse\_data (model, scanned\_par1, scanned\_par2, inputdir, outputdir, runs) The second pipeline step: data analysis.

#### **Parameters**

- model the model name
- scanned par1 the first scanned parameter
- **scanned\_par2** the second scanned parameter
- inputdir the directory containing the simulated data sets to process
- **outputdir** the directory to store the performed analysis
- runs the number of model simulation

**Returns** True if the task was completed successfully, False otherwise.

classmethod generate\_data (simulator, model, sim\_length, inputdir, outputdir, cluster\_type, pp\_cpus, runs)

The first pipeline step: data generation.

#### **Parameters**

- **simulator** the name of the simulator (e.g. Copasi)
- model the model to process
- **sim\_length** the length of the simulation
- inputdir the directory containing the model
- outputdir the directory to store the results
- cluster\_type pp for local Parallel Python, lsf for Load Sharing Facility, sge for Sun Grid Engine.
- pp\_cpus the number of CPU used by Parallel Python.
- runs the number of model simulation

**Returns** True if the task was completed successfully, False otherwise.

classmethod generate report (model, scanned parl, scanned par2, outputdir, sim plots folder)

The third pipeline step: report generation.

#### **Parameters**

- model the model name
- scanned\_par1 the first scanned parameter
- **scanned\_par2** the second scanned parameter
- outputdir the directory containing the report
- **sim\_plots\_folder** the folder containing the plots.

**Returns** True if the task was completed successfully, False otherwise.

```
read_config (lines)
run (config_file)
```

#### Module contents

sbpipe.pl.sim package

#### **Submodules**

#### sbpipe.pl.sim.sim module

This module provides the user with a complete pipeline of scripts for running model simulations

#### **Parameters**

- model the model name
- **inputdir** the directory containing the data to analyse
- outputdir the output directory containing the results
- **sim\_plots\_dir** the directory to save the plots
- exp\_dataset the full path of the experimental data set
- plot\_exp\_dataset True if the experimental data set should also be plotted
- **xaxis\_label** the label for the x axis (e.g. Time [min])
- yaxis\_label the label for the y axis (e.g. Level [a.u.])

**Returns** True if the task was completed successfully, False otherwise.

```
classmethod generate_data (simulator, model, inputdir, outputdir, cluster_type='pp', pp\_cpus=2, runs=1)

The first pipeline step: data generation.
```

#### **Parameters**

- **simulator** the name of the simulator (e.g. Copasi)
- model the model to process
- inputdir the directory containing the model
- **outputdir** the directory containing the output files
- cluster\_type pp for local Parallel Python, lsf for Load Sharing Facility, sge for Sun Grid Engine.
- pp\_cpus the number of CPU used by Parallel Python.
- runs the number of model simulation

**Returns** True if the task was completed successfully, False otherwise.

```
{\bf classmethod\ generate\_report\ }({\it model}, {\it outputdir}, {\it sim\_plots\_folder})
```

The third pipeline step: report generation.

#### **Parameters**

- model the model name
- outputdir the output directory to store the report
- sim\_plots\_folder the folder containing the plots

Returns True if the task was completed successfully, False otherwise.

```
read_config (lines)
run (config_file)
```

#### **Module contents**

#### **Submodules**

#### sbpipe.pl.pipeline module

Generic pipeline.

#### **Parameters**

- data\_folder the folder containing the experimental (wet) data sets
- models\_folder the folder containing the models
- working folder the folder to store the results
- **sim\_data\_folder** the folder to store the simulation data
- **sim\_plots\_folder** the folder to store the graphic results

```
config_parser (config_file, section)
```

Return the configuration for the parsed section in the config\_file

#### **Parameters**

- config\_file the configuration file to parse
- **section** the section in the configuration file to parse

**Returns** the configuration for the parsed section in the config\_file

```
get data folder()
```

Return the folder containing the experimental (wet) data sets.

**Returns** the experimental data sets folder.

```
get_models_folder()
```

Return the folder containing the models.

**Returns** the models folder.

```
get_sim_data_folder()
```

Return the folder containing the in-silico generated data sets.

**Returns** the folder of the simulated data sets.

```
get_sim_plots_folder()
```

Return the folder containing the in-silico generated plots.

**Returns** the folder of the simulated plots.

#### classmethod get\_simul\_obj (simulator)

Return the simulator object if this exists. Otherwise throws an exception. The simulator name starts with an upper case letter. Each simulator is in a package within *sbpipe.simulator*.

**Parameters** simulator – the simulator name

Returns the simulator object.

#### get\_working\_folder()

Return the folder containing the results.

**Returns** the working folder.

#### classmethod read common config(lines)

Parse the common parameters from the configuration file

**Parameters** lines – the lines to parse.

**Returns** return a tuple containing the common parameters

#### read\_config(lines)

Read the section lines from the configuration file. This method is abstract.

**Returns** a tuple containing the configuration

run (config\_file)

Run the pipeline.

**Parameters** config\_file – a configuration file for this pipeline.

Returns True if the pipeline was executed correctly, False otherwise.

#### **Module contents**

#### sbpipe.report package

#### Submodules

#### sbpipe.report.latex reports module

Initialize a Latex header with a title and an abstract.

## **Parameters**

- pdftitle the pdftitle for the LaTeX header
- title the title for the LaTeX header
- abstract the abstract for the LaTeX header

#### **Returns** the LaTeX header

Generate a generic report.

- **outputdir** the output directory
- **sim\_plots\_folder** the folder containing the simulated plots
- model noext the model name
- **filename\_prefix** the prefix for the LaTeX file
- caption True if figure captions (=figure file name) should be added

```
sbpipe.report.latex_reports.latex_report_dps (outputdir, sim_plots_folder, filename_prefix, model_noext, scanned_par1, scanned_par2)
```

Generate a report for a double parameter scan task.

#### **Parameters**

- **outputdir** the output directory
- **sim\_plots\_folder** the folder containing the simulated plots
- **filename\_prefix** the prefix for the LaTeX file
- model noext the model name
- scanned\_par1 the 1st scanned parameter
- scanned\_par2 the 2nd scanned parameter

Generate a report for a parameter estimation task.

#### **Parameters**

- outputdir the output directory
- sim\_plots\_folder the folder containing the simulated plots
- model\_noext the model name
- filename prefix the prefix for the LaTeX file

```
sbpipe.report.latex_reports.latex_report_sim(outputdir, sim_plots_folder, model_noext, filename_prefix)
```

Generate a report for a time course task.

#### **Parameters**

- **outputdir** the output directory
- **sim\_plots\_folder** the folder containing the simulated plots
- model\_noext the model name
- filename\_prefix the prefix for the LaTeX file

Generate a report for a single parameter scan task.

#### **Parameters**

- **outputdir** the output directory
- **sim\_plots\_folder** the folder containing the simulated plots
- filename prefix the prefix for the LaTeX file
- model\_noext the model name
- scanned\_par the scanned parameter

```
sbpipe.report.latex_reports.pdf_report (outputdir, filename)
Generate a PDF report from LaTeX report using pdflatex.
```

- **outputdir** the output directory
- filename the LaTeX file name

#### **Module contents**

sbpipe.simul package

**Subpackages** 

sbpipe.simul.copasi package

#### **Submodules**

#### sbpipe.simul.copasi.copasi module

```
class sbpipe.simul.copasi.copasi.Copasi
    Bases: sbpipe.simul.simul.Simul (page 31)
```

Copasi simulator.

```
collect_pe_results (inputdir, outputdir, fileout_all_fits, file_out_best_fits)
```

**pe** (model, inputdir, cluster\_type, pp\_cpus, runs, outputdir, sim\_data\_dir, updated\_models\_dir)

ps2 (model, sim\_length, inputdir, outputdir, cluster\_type='pp', pp\_cpus=2, runs=1)

sim (model, inputdir, outputdir, cluster\_type='pp', pp\_cpus=2, runs=1)

#### sbpipe.simul.copasi\_parser module

```
class sbpipe.simul.copasi.copasi_parser.CopasiParser
```

Retrieve information from a Copasi file.

```
classmethod get_param_estim_val (file_in)
```

Parse a Copasi file and retrieve information on the parameters to estimate.

Parameters file\_in - the Copasi file including absolute path to parse

**Returns** a tuple containing the report file name, the parameter lower bounds, names, starting values, and upper bounds

### sbpipe.simul.copasi\_utils module

Collect all the parameter estimates from the Copasi parameter estimation report. Results are stored in filename\_out.

#### **Parameters**

- path\_in the path to the input files
- path\_out the path to the output files
- **filename** out the filename to store the final estimates

Collect the final parameter estimates from the Copasi parameter estimation report. Results are stored in filename out.

- path\_in the path to the input files
- path\_out the path to the output files

• **filename\_out** – the filename to store the final estimates

sbpipe.simul.copasi.copasi\_utils.get\_params\_list (filein)

Return the list of parameter names from filein

Parameters filein – a Copasi parameter estimation report file

**Returns** the list of parameter names

sbpipe.simul.copasi.copasi\_utils.replace\_str\_copasi\_sim\_report (report)

Replace a group of annotation strings from a generated copasi report file

Parameters report - The report file with absolute path

sbpipe.simul.copasi.copasi\_utils.write\_all\_fits (files, path\_out, filename\_out)
Write all the estimates to filename out

#### **Parameters**

- **files** the list of Copasi parameter estimation reports
- path\_out the path to store the file combining all the estimates
- filename\_out the file containing all the estimates

sbpipe.simul.copasi.copasi\_utils.write\_best\_fits (files, path\_out, filename\_out)
Write the final estimates to filename out

#### **Parameters**

- **files** the list of Copasi parameter estimation reports
- path\_out the path to store the file combining the final (best) estimates (file-name\_out)
- **filename\_out** the file containing the final (best) estimates

#### sbpipe.simul.copasi.randomise module

class sbpipe.simul.copasi.randomise.Randomise(path, filename\_in)

This class generates multiple copies of a Copasi file configured for parameter estimation task, and randomises the starting values of the parameters to estimate.

#### **Parameters**

- path the path to filename in
- **filename\_in** the Copasi file to process.

```
get_copasi_obj()
```

Return the Copasi parser object

Returns the Copasi parser object

```
get_lower_bounds_list()
```

Return the list of parameter lower bounds

**Returns** the list of parameter lower bounds

```
get_param_names_list()
```

Return the list of parameter names

**Returns** the list of parameter names

# get\_path()

Return the path containing the template Copasi file

**Returns** the path to the Copasi file

## get\_report\_filename()

Return the name of the template parameter estimation report

```
Returns the name of the report file name for parameter estimation
```

```
get_start_values_list()
```

Return the list of parameter starting values

**Returns** the list of parameter starting values

```
get_template_copasi_file()
```

Return the name of the template Copasi file

Returns the name of the Copasi file

```
get_upper_bounds_list()
```

Return the list of parameter upper bounds

**Returns** the list of parameter upper bounds

```
print_params_2_estim()
```

Print the parameter names, lower/upper bounds, and starting value, as extracted from COPASI template file

```
randomise (num_files, idstr)
```

Randomise the starting values for the parameter to estimate.

#### **Parameters**

- num\_files the number of files (instances) to generate
- idstr an ID string to label the generated files (e.g. a timestamp)

```
replicate (num_files, idstr)
```

Generate num\_files files and add an ID string to Copasi file/report names

#### **Parameters**

- num\_files the number of files (instances) to generate
- idstr an ID string to label the generated files (e.g. a timestamp)

# **Module contents**

### sbpipe.simul.java package

#### Submodules

# sbpipe.simul.java.java module

Java Simulator.

```
class sbpipe.simul.java.java.Java
    Bases: sbpipe.simul.pl_simul.PLSimul(page 29)
```

## **Module contents**

#### sbpipe.simul.octave package

#### Submodules

#### sbpipe.simul.octave.octave module

```
class sbpipe.simul.octave.octave.Octave

Bases: sbpipe.simul.pl_simul.PLSimul (page 29)
```

Octave Simulator.

#### **Module contents**

#### sbpipe.simul.python package

#### **Submodules**

```
sbpipe.simul.python.python module
```

```
class sbpipe.simul.python.python.Python
    Bases: sbpipe.simul.pl_simul.PLSimul(page 29)
    Python Simulator.
```

#### **Module contents**

#### sbpipe.simul.rscript package

#### **Submodules**

### sbpipe.simul.rscript.rscript module

```
class sbpipe.simul.rscript.rscript.Rscript
    Bases: sbpipe.simul.pl_simul.PLSimul(page 29)
    Rscript Simulator.
```

#### **Module contents**

#### **Submodules**

#### sbpipe.simul.pl\_simul module

```
class sbpipe.simul.pl_simul.PLSimul (lang, lang_err_msg, options)
    Bases: sbpipe.simul.simul.Simul (page 31)

A generic simulator for models coded in a programming language.

collect_pe_results (inputdir, outputdir, fileout_all_fits, file_out_best_fits)

get_lang()
    Return the programming language name :return: the name

get_lang_err_msg()
    Return the error if the programming language is not found :return: the error message

get_lang_options()
    Return the options for the programming language command :return: the options. Return None, if no options are used.

pe (model, inputdir, cluster_type, pp_cpus, nfits, outputdir, sim_data_dir, updated_models_dir)

sim (model, inputdir, outputdir, cluster_type='pp', pp_cpus=2, runs=1)
```

#### sbpipe.simul.pl\_simul\_utils module

Collect all the parameter estimates. Results are stored in filename\_out.

#### **Parameters**

- path\_in the path to the input files
- path\_out the path to the output files
- **filename\_out** the filename to store the final estimates

```
sbpipe.simul.pl_simul_utils.get_best_fits(path_in='.', path_out='.', file-name out='final estimates.csv')
```

Collect the final parameter estimates. Results are stored in filename\_out.

#### **Parameters**

- path\_in the path to the input files
- path\_out the path to the output files
- **filename\_out** the filename to store the final estimates

```
sbpipe.simul.pl_simul_utils.get_input_files (path)
    Retrieve the input files in a path.
```

Parameters path – the path containing the input files to retrieve

Returns the list of input files

```
sbpipe.simul.pl_simul_utils.get_params_list (filein)
```

Return the list of parameter names from filein

Parameters filein – a report file

**Returns** the list of parameter names

```
sbpipe.simul_pl_simul_utils.move_report_files (outputdir, group_model, groupid)

Move the report files :param outputdir: the output directory :param group_model: the model file name :param groupid: the group id of the reports
```

```
sbpipe.simul.pl_simul_utils.replace_str_pl_report (report)
Replace a group of annotation strings from report generated with a programming language
```

Parameters report – The report file with absolute path

```
sbpipe.simul.pl_simul_utils.write_all_fits (files, path_out, filename_out)
Write all the estimates to filename_out
```

#### **Parameters**

- **files** the list of parameter estimation reports
- path\_out the path to store the file combining all the estimates
- filename\_out the file containing all the estimates

```
sbpipe.simul.pl_simul_utils.write_best_fits (files, path_out, filename_out)
Write the final estimates to filename out
```

- **files** the list of parameter estimation reports
- path\_out the path to store the file combining the final (best) estimates (file-name out)
- **filename\_out** the file containing the final (best) estimates

```
sbpipe.simul.pl_simul_utils.write_params (col_names, path_out, filename_out)

Write the list of parameter names to filename_out
```

#### **Parameters**

- col\_names the list of parameter names
- path\_out the path to store filename\_out
- **filename\_out** the output file to store the parameter names

#### sbpipe.simul.simul module

```
class sbpipe.simul.simul.Simul
```

Generic simulator.

 $\verb|collect_pe_results| (input dir, output dir, file out\_all\_fits, file\_out\_best\_fits)|$ 

Collect the parameter estimation results in two files.

#### **Parameters**

- inputdir input folder containing the independent parameter estimation files
- outputdir the output folder to store the new files to create
- **fileout\_all\_fits** a global file containing all fits from the independent parameter estimations.
- **file\_out\_best\_fits** a global file containing the best fits from the independent parameter estimations.

**pe** (model, inputdir, cluster\_type, pp\_cpus, nfits, outputdir, sim\_data\_dir, updated\_models\_dir) parameter estimation.

#### **Parameters**

- model the model to process
- inputdir the directory containing the model
- **cluster\_type** pp for parallel python, lsf for load sharing facility, sge for sun grid engine
- pp\_cpus the number of cpu for parallel python
- nfits the number of fits to perform
- outputdir the directory to store the results
- sim\_data\_dir the directory containing the simulation data sets
- updated\_models\_dir the directory containing the models with updated parameters for each estimation

- model the model to process
- scanned\_par the scanned parameter
- **simulate\_intervals** the time step of each simulation
- single\_param\_scan\_intervals the number of scans to perform
- inputdir the directory containing the model
- **outputdir** the directory to store the results
- **cluster\_type** pp for local Parallel Python, lsf for Load Sharing Facility, sge for Sun Grid Engine.
- pp\_cpus the number of CPU used by Parallel Python.

• runs – the number of model simulation

**ps2** (*model*, *sim\_length*, *inputdir*, *outputdir*, *cluster\_type='pp'*, *pp\_cpus=2*, *runs=1*) Double paramter scan.

#### **Parameters**

- model the model to process
- sim length the length of the simulation
- inputdir the directory containing the model
- **outputdir** the directory to store the results
- cluster\_type pp for local Parallel Python, lsf for Load Sharing Facility, sge for Sun Grid Engine.
- pp\_cpus the number of CPU used by Parallel Python.
- runs the number of model simulation

**sim** (*model*, *inputdir*, *outputdir*, *cluster\_type='pp'*, *pp\_cpus=2*, *runs=1*)
Time course simulator.

#### **Parameters**

- model the model to process
- inputdir the directory containing the model
- **outputdir** the directory containing the output files
- cluster\_type pp for local Parallel Python, lsf for Load Sharing Facility, sge for Sun Grid Engine.
- pp\_cpus the number of CPU used by Parallel Python.
- runs the number of model simulation

#### Module contents

#### sbpipe.utils package

#### Submodules

#### sbpipe.utils.io module

```
sbpipe.utils.io.files_with_pattern_recur (folder, pattern)
Return all files with a certain pattern in folder+subdirectories
```

#### **Parameters**

- **folder** the folder to search for
- pattern the string to search for

**Returns** the files containing the pattern.

```
sbpipe.utils.io.get_pattern_pos(pattern, filename)
```

Return the line number (as string) of the first occurrence of a pattern in filename

#### **Parameters**

- pattern the pattern of the string to find
- **filename** the file name containing the pattern to search

Returns the line number containing the pattern or "-1" if the pattern was not found

```
sbpipe.utils.io.refresh(path, file_pattern)
```

Clean and create the folder if this does not exist.

#### **Parameters**

- path the path containing the files to remove
- **file\_pattern** the string pattern of the files to remove

```
sbpipe.utils.io.replace_str_in_file (filename_out, old_string, new_string)
```

Replace a string with another in filename\_out

#### **Parameters**

- filename\_out the output file
- old\_string the old string that should be replaced
- new\_string the new string replacing old\_string

```
sbpipe.utils.io.write_mat_on_file(path, filename_out, data)
```

Write the matrix results stored in data to filename\_out

#### **Parameters**

- path the path to filename\_out
- filename\_out the output file
- data the data to store in a file

### sbpipe.utils.monitor module

```
class sbpipe.utils.monitor.Monitor
```

This is a monitor. It is a callback class for collecting information about finished processes. It is used by Parallel Python (pp).

```
add (pid, value)
```

The callback function

#### **Parameters**

- pid this is callbackargs passed to parallel python *submit()* method
- **value** the return value of the parallelised function. It is the callback value.

```
get_count()
```

Return the counter

**Returns** the number of running processes.

```
get_value()
```

Return the internal status.

**Returns** True if the counter is empty.

## sbpipe.utils.parcomp module

```
sbpipe.utils.parcomp.parcomp(cmd, cmd_iter_substr, cluster_type, runs, output_dir, pp_cpus=1)
```

Generic funcion to run a command in parallel

### **Parameters**

- cmd the command string to run in parallel
- **cmd\_iter\_substr** the substring of the iteration number. This will be replaced in a number automatically
- cluster\_type the cluster type among pp (multithreading), sge, or lsf
- runs the number of runs

- output\_dir the output directory
- pp\_cpus the number of cpus that pp should use at most

sbpipe.utils.parcomp.run\_cmd\_instance(cmd)

Run a command using Python subprocess.

**Parameters** cmd – the string of the command to run

sbpipe.utils.parcomp.run\_command\_pp(cmd, cmd\_iter\_substr, runs, server, monitor=<sbpipe.utils.monitor.Monitor instance>)

Run instances of a command in multithreading using parallel python (pp).

#### **Parameters**

- cmd the command string to run in parallel
- **cmd\_iter\_substr** the substring of the iteration number. This will be replaced in a number automatically
- runs the number of runs
- **server** the server that pp should use
- monitor the mutex object to count the jobs

sbpipe.utils.parcomp.run\_jobs\_lsf(cmd, cmd\_iter\_substr, out\_dir, err\_dir, runs)
Run jobs using a Load Sharing Facility (LSF) cluster.

#### **Parameters**

- cmd the full command to run as a job
- cmd\_iter\_substr the substring in command to be replaced with a number
- out\_dir the directory containing the standard output from bsub
- err\_dir the directory containing the standard error from bsub
- runs the number of runs to execute

sbpipe.utils.parcomp.run\_jobs\_pp (cmd, cmd\_iter\_substr, runs, pp\_cpus=1)
Run jobs using parallel python (pp) locally.

## Parameters

- cmd the full command to run as a job
- cmd\_iter\_substr the substring in command to be replaced with a number
- runs the number of runs to execute
- **pp\_cpus** The number of available cpus. If pp\_cpus <=0, all the available cores will be used.

sbpipe.utils.parcomp.run\_jobs\_sge(cmd, cmd\_iter\_substr, out\_dir, err\_dir, runs)
Run jobs using a Sun Grid Engine (SGE) cluster.

#### **Parameters**

- cmd the full command to run as a job
- cmd\_iter\_substr the substring in command to be replaced with a number
- out\_dir the directory containing the standard output from qsub
- err\_dir the directory containing the standard error from qsub
- runs the number of runs to execute

### sbpipe.utils.rand module

sbpipe.utils.rand.get\_rand\_alphanum\_str(length)

Return a random alphanumeric string

**Parameters** length – the length of the string

**Returns** the generated string

sbpipe.utils.rand.get\_rand\_num\_str(length)

Return a random numeric string

Parameters length – the length of the string

**Returns** the generated string

#### sbpipe.utils.re\_utils module

```
sbpipe.utils.re_utils.nat_sort_key(str)
```

The key to sort a list of strings alphanumerically (e.g. "file10" is correctly placed after "file2")

**Parameters** str – the string to sort alphanumerically in a list of strings

**Returns** the key to sort strings alphanumerically

#### **Module contents**

#### **Submodules**

sbpipe.\_\_main\_\_ module

### sbpipe.main module

```
exception sbpipe.main.Usage (msg)
```

Bases: exceptions. Exception

This class is used for printing a generic exception

sbpipe.main.check\_args(args, msg)

Check that at least one argument is passed.

### **Parameters**

- args the list of arguments
- msg the message to print

Raise Usage exception if less than one argument is passed

Returns no output

sbpipe.main.help()

Return help message.

Returns the help message

sbpipe.main.license()

Return the license

Returns the license

sbpipe.main.logo()

Return sbpipe logo.

Returns sbpipe logo

sbpipe.main.main(argv=None)

SB pipe main function.

**Parameters** argv – options for sbpipe. Type python -m sbpipe -h for a full list of options.

**Returns** 0 if OK, 1 if minor problems, or 2 if serious trouble.

```
\verb|sbpipe.main.version|()|
```

Return the version

**Returns** the version

## sbpipe.sb\_config module

```
\verb|sbpipe.sb_config.isPyPackageInstalled|(package)|
```

Utility checking whether a Python package is installed.

Parameters package - a Python package name

**Returns** True if it is installed, false otherwise.

sbpipe.sb\_config.which(cmd\_name)

Utility equivalent to which in GNU/Linux OS.

Parameters cmd\_name - a command name

Returns return the command name with absolute path if this exists, or None

#### **Module contents**

**CHAPTER** 

**FOUR** 

# **META INFORMATION**

# Copyright

Copyright © 2015-2018, Piero Dalle Pezze and Nicolas Le Novère.

SB pipe and its documentation are released under the GNU Lesser General Public License v3 (LGPLv3). A copy of this license is provided with the package and can also be found here: https://www.gnu.org/licenses/lgpl-3.0.txt.

Contacts: Dr Piero Dalle Pezze (piero.dallepezze AT babraham.ac.uk) and Dr Nicolas Le Novère (lenov AT babraham.ac.uk)

Affiliation: The Babraham Institute, Cambridge, CB22 3AT, UK

# **CHAPTER**

# **FIVE**

# **INDICES**

- genindex
- modindex
- search

40 Chapter 5. Indices

## S

```
sbpipe, 36
sbpipe.___main___,35
sbpipe.main, 35
sbpipe.pl, 24
sbpipe.pl.create, 17
sbpipe.pl.create.newproj, 17
sbpipe.pl.pe, 19
sbpipe.pl.pe.parest, 18
sbpipe.pl.pipeline, 23
sbpipe.pl.ps1, 20
sbpipe.pl.psl.parscan1, 19
sbpipe.pl.ps2, 22
sbpipe.pl.ps2.parscan2,21
sbpipe.pl.sim, 23
sbpipe.pl.sim.sim, 22
sbpipe.report, 26
sbpipe.report.latex_reports, 24
sbpipe.sb_config, 36
sbpipe.simul, 32
sbpipe.simul.copasi, 28
sbpipe.simul.copasi.copasi,26
sbpipe.simul.copasi.copasi_parser,26
sbpipe.simul.copasi.copasi_utils,26
sbpipe.simul.copasi.randomise, 27
sbpipe.simul.java, 28
sbpipe.simul.java.java,28
sbpipe.simul.octave, 29
sbpipe.simul.octave.octave, 28
sbpipe.simul.pl_simul, 29
sbpipe.simul.pl_simul_utils,30
sbpipe.simul.python, 29
sbpipe.simul.python.python, 29
sbpipe.simul.rscript,29
sbpipe.simul.rscript.rscript, 29
sbpipe.simul.simul, 31
sbpipe.utils, 35
sbpipe.utils.io, 32
sbpipe.utils.monitor, 33
sbpipe.utils.parcomp, 33
sbpipe.utils.rand, 35
sbpipe.utils.re_utils,35
```

A	get_all_fits() (in module
add() (sbpipe.utils.monitor.Monitor method), 33	sbpipe.simul.copasi.copasi_utils), 26
analyse_data() (sbpipe.pl.pe.parest.ParEst class method), 18	get_all_fits() (in module sbpipe.simul.pl_simul_utils), 30
analyse_data() (sbpipe.pl.ps1.parscan1.ParScan1 class	get_best_fits() (in module
method), 19	sbpipe.simul.copasi.copasi_utils), 26
analyse_data() (sbpipe.pl.ps2.parscan2.ParScan2 class method), 21	get_best_fits() (in module sbpipe.simul.pl_simul_utils), 30
analyse_data() (sbpipe.pl.sim.sim.Sim class method),	get_copasi_obj() (sbpipe.simul.copasi.randomise.Randomise method), 27
	get_count() (sbpipe.utils.monitor.Monitor method), 33
C	get_data_folder() (sbpipe.pl.pipeline.Pipeline method),
check_args() (in module sbpipe.main), 35	23
collect_pe_results() (sbpipe.simul.copasi.copasi.Copasi	get_input_files() (in module sbpipe.simul.pl_simul_utils), 30
method), 26 collect_pe_results() (sbpipe.simul.pl_simul.PLSimul	get_lang() (sbpipe.simul.pl_simul.PLSimul method), 29
method), 29 collect_pe_results() (sbpipe.simul.simul.Simul	get_lang_err_msg() (sbpipe.simul.pl_simul.PLSimul method), 29
method), 31 config_parser() (sbpipe.pl.pipeline.Pipeline method),	get_lang_options() (sbpipe.simul.pl_simul.PLSimul method), 29
23 Copasi (class in sbpipe.simul.copasi.copasi), 26	get_latex_header() (in module
CopasiParser (class in sopipe.simur.copasi.copasi), 20	sbpipe.report.latex_reports), 24
sbpipe.simul.copasi_copasi_parser), 26	get_lower_bounds_list()
	(sbpipe.simul.copasi.randomise.Randomise
F	method), 27
files_with_pattern_recur() (in module sbpipe.utils.io), 32	get_models_folder() (sbpipe.pl.pipeline.Pipeline method), 23
0	get_param_estim_val() (sbpipe.simul.copasi.copasi_parser.CopasiParser
G	class method), 26
generate_data() (sbpipe.pl.pe.parest.ParEst class	get_param_names_list()
method), 18	(sbpipe.simul.copasi.randomise.Randomise
generate_data() (sbpipe.pl.ps1.parscan1.ParScan1 class	method), 27
method), 20 generate_data() (sbpipe.pl.ps2.parscan2.ParScan2 class	get_params_list() (in module
method), 21	sbpipe.simul.copasi.copasi_utils), 27
generate_data() (sbpipe.pl.sim.sim.Sim class method),	get_params_list() (in module
22	sbpipe.simul.pl_simul_utils), 30
generate_report() (sbpipe.pl.pe.parest.ParEst class method), 19	get_path() (sbpipe.simul.copasi.randomise.Randomise method), 27
generate_report() (sbpipe.pl.ps1.parscan1.ParScan1	get_pattern_pos() (in module sbpipe.utils.io), 32
class method), 20	get_rand_alphanum_str() (in module sbpipe.utils.rand), 35
generate_report() (sbpipe.pl.ps2.parscan2.ParScan2	get_rand_num_str() (in module sbpipe.utils.rand), 35
class method), 21	get_report_filename() (sbpipe.simul.copasi.randomise.Randomise
generate_report() (sbpipe.pl.sim.sim.Sim class	method), 27
method), 22	get_sim_data_folder() (sbpipe.pl.pipeline.Pipeline

method), 23 get_sim_plots_folder() (sbpipe.pl.pipeline.Pipeline method), 23	ParScan1 (class in sbpipe.pl.ps1.parscan1), 19 ParScan2 (class in sbpipe.pl.ps2.parscan2), 21 pdf_report() (in module sbpipe.report.latex_reports), 25
get_simul_obj() (sbpipe.pl.pipeline.Pipeline class method), 23 get_start_values_list() (sbpipe.simul.copasi.randomise.R	pe() (sbpipe.simul.copasi.copasi.Copasi method), 26 pe() (sbpipe.simul.pl_simul.PLSimul method), 29
method), 28	Pipeline (class in sbpipe.pl.pipeline), 23
get_template_copasi_file()	PLSimul (class in sbpipe.simul.pl_simul), 29 print_params_2_estim() (sbpipe.simul.copasi.randomise.Randomise
get_upper_bounds_list()	method), 28
(sbpipe.simul.copasi.randomise.Randomise method), 28	ps1() (sbpipe.simul.copasi.copasi.Copasi method), 26 ps1() (sbpipe.simul.simul.Simul method), 31
get_value() (sbpipe.utils.monitor.Monitor method), 33 get_working_folder() (sbpipe.pl.pipeline.Pipeline method), 24	ps2() (sbpipe.simul.copasi.copasi.Copasi method), 26 ps2() (sbpipe.simul.simul.Simul method), 32 Python (class in sbpipe.simul.python.python), 29
Н	R
help() (in module sbpipe.main), 35	Randomise (class in sbpipe.simul.copasi.randomise),
isPyPackageInstalled() (in module sbpipe.sb_config),	randomise() (sbpipe.simul.copasi.randomise.Randomise method), 28
36	read_common_config() (sbpipe.pl.pipeline.Pipeline class method), 24
J	read_config() (sbpipe.pl.pe.parest.ParEst method), 19 read_config() (sbpipe.pl.pipeline.Pipeline method), 24
Java (class in sbpipe.simul.java.java), 28	read_config() (sbpipe.pl.ps1.parscan1.ParScan1 method), 20
latex_report() (in module sbpipe.report.latex_reports),	read_config() (sbpipe.pl.ps2.parscan2.ParScan2
24	method), 21 read_config() (sbpipe.pl.sim.sim.Sim method), 23
latex_report_dps() (in module sbpipe.report.latex_reports), 24	refresh() (in module sbpipe.utils.io), 32
latex_report_pe() (in module	replace_str_copasi_sim_report() (in module sbpipe.simul.copasi.copasi_utils), 27
sbpipe.report.latex_reports), 25 latex_report_sim() (in module	replace_str_in_file() (in module sbpipe.utils.io), 33 replace_str_pl_report() (in module
sbpipe.report.latex_reports), 25 latex_report_sps() (in module	sbpipe.simul.pl_simul_utils), 30
sbpipe.report.latex_reports), 25	replicate() (sbpipe.simul.copasi.randomise.Randomise method), 28
license() (in module sbpipe.main), 35 logo() (in module sbpipe.main), 35	Rscript (class in sbpipe.simul.rscript.rscript), 29
M	run() (sbpipe.pl.create.newproj.NewProj method), 17 run() (sbpipe.pl.pe.parest.ParEst method), 19
main() (in module sbpipe.main), 35	run() (sbpipe.pl.pipeline.Pipeline method), 24
Monitor (class in sbpipe.utils.monitor), 33	run() (sbpipe.pl.ps1.parscan1.ParScan1 method), 20 run() (sbpipe.pl.ps2.parscan2.ParScan2 method), 22
move_report_files() (in module sbpipe.simul.pl_simul_utils), 30	run() (sbpipe.pl.sim.sim.Sim method), 23
N	run_cmd_instance() (in module sbpipe.utils.parcomp), 34
nat_sort_key() (in module sbpipe.utils.re_utils), 35	run_command_pp() (in module sbpipe.utils.parcomp), 34
NewProj (class in sbpipe.pl.create.newproj), 17	run_jobs_lsf() (in module sbpipe.utils.parcomp), 34
0	run_jobs_pp() (in module sbpipe.utils.parcomp), 34 run_jobs_sge() (in module sbpipe.utils.parcomp), 34
Octave (class in sbpipe.simul.octave.octave), 28	
Р	S shrine (module) 36
parcomp() (in module sbpipe.utils.parcomp), 33	sbpipe (module), 36 sbpipemain (module), 35
ParEst (class in sbpipe.pl.pe.parest), 18	sbpipe.main (module), 35

44 Index

```
sbpipe.pl (module), 24
                                                        write_best_fits()
                                                                                       (in
                                                                                                        module
sbpipe.pl.create (module), 17
                                                                  sbpipe.simul.pl_simul_utils), 30
sbpipe.pl.create.newproj (module), 17
                                                         write_mat_on_file() (in module sbpipe.utils.io), 33
sbpipe.pl.pe (module), 19
                                                         write_params()
                                                                                      (in
                                                                                                        module
sbpipe.pl.pe.parest (module), 18
                                                                  sbpipe.simul.pl_simul_utils), 30
sbpipe.pl.pipeline (module), 23
sbpipe.pl.ps1 (module), 20
sbpipe.pl.ps1.parscan1 (module), 19
sbpipe.pl.ps2 (module), 22
sbpipe.pl.ps2.parscan2 (module), 21
sbpipe.pl.sim (module), 23
sbpipe.pl.sim.sim (module), 22
sbpipe.report (module), 26
sbpipe.report.latex_reports (module), 24
sbpipe.sb_config (module), 36
sbpipe.simul (module), 32
sbpipe.simul.copasi (module), 28
sbpipe.simul.copasi.copasi (module), 26
sbpipe.simul.copasi.copasi_parser (module), 26
sbpipe.simul.copasi.copasi utils (module), 26
sbpipe.simul.copasi.randomise (module), 27
sbpipe.simul.java (module), 28
sbpipe.simul.java.java (module), 28
sbpipe.simul.octave (module), 29
sbpipe.simul.octave.octave (module), 28
sbpipe.simul.pl_simul (module), 29
sbpipe.simul.pl simul utils (module), 30
sbpipe.simul.python (module), 29
sbpipe.simul.python.python (module), 29
sbpipe.simul.rscript (module), 29
sbpipe.simul.rscript.rscript (module), 29
sbpipe.simul.simul (module), 31
sbpipe.utils (module), 35
sbpipe.utils.io (module), 32
sbpipe.utils.monitor (module), 33
sbpipe.utils.parcomp (module), 33
sbpipe.utils.rand (module), 35
sbpipe.utils.re_utils (module), 35
Sim (class in sbpipe.pl.sim.sim), 22
sim() (sbpipe.simul.copasi.copasi.Copasi method), 26
sim() (sbpipe.simul.pl_simul.PLSimul method), 29
sim() (sbpipe.simul.simul.Simul method), 32
Simul (class in sbpipe.simul.simul), 31
U
Usage, 35
version() (in module sbpipe.main), 36
W
which() (in module sbpipe.sb_config), 36
write all fits()
                             (in
                                               module
         sbpipe.simul.copasi.copasi utils), 27
write_all_fits()
                             (in
                                               module
         sbpipe.simul.pl_simul_utils), 30
write best fits()
                                               module
```

Index 45

sbpipe.simul.copasi.copasi\_utils), 27