UCONN HEALTH

Running COPASI biochemical simulations with HTCondor

Pedro Mendes

Center for Cell Analysis & Modeling
University of Connecticut School of Medicine

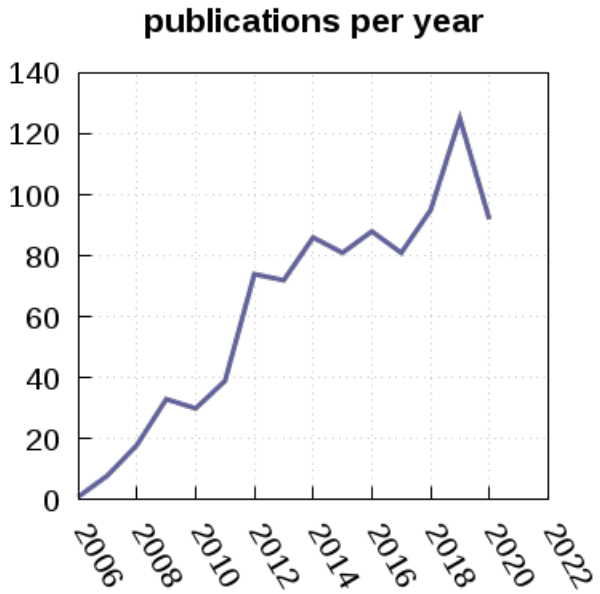
pmendes@uchc.edu

twitter: @gepasi

http://www.comp-sys-bio.org
http://copasi.org

COPASI

- Biochemical system simulator, directed at biologists/chemists (ie little computing or maths expertise)
- GUI-based standalone software (QT)
- Has now a well-established worlwide user-base
- Around 100 papers published yearly that use COPASI
- Over 10,000 downloads per year

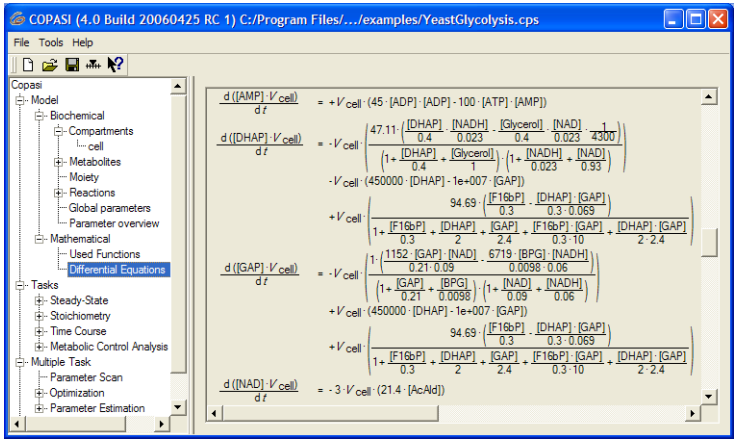


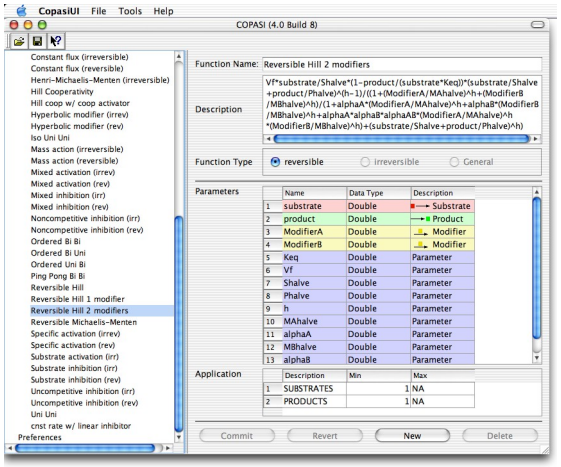
http://copasi.org **@COPASI_software**

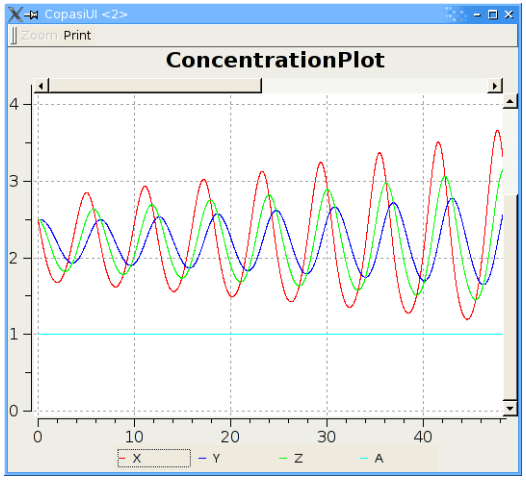
COPASI

- ODE, SDE, or SSA
- Discrete events
- Steady states
- Stability analysis
- Parameter scans
- Sensitivity analysis
- Metabolic Control Analysis
- Time scale separation analysis

- Optimisation,
- Parameter estimation
- Cross-sections
- Lyapunov exponents
- Network diagrams
- 2D and surface plots
- GUI and command line versions
- Reads/writes SBML







http://copasi.org **@COPASI_software**

University of Manchester FEPS HTCondor pool

- Setup by Ian Cottam around 2009
- 3000 CPU cores mostly from student terminal room workstations
- During daytime about 100 machines available
- Overnight, weekends and holidays, full power of 3000 CPUs

Using COPASI with HTCondor

1) Edit model in COPASI
2) Set appropriate report format
3) Mark job for execution
4) Prepare HTCondor job specification file
5) Prepare output directories
6) Submit to HTCondor
7) Edit output to match required order
8) Create summary plots

File: `stochastic_test_1.job`

```
executable = /home/ed/software/copasi/bin/CopasiSE
universe       = vanilla
arguments = --nologo --home . stochastic_test_batch_mode.cps
transfer_input_files = stochastic_test_batch_mode.cps
log =  stochastic_test.log
error = stochastic_test.err
output = stochastic_test.out
Requirements = (OpSys == "LINUX" && Arch == "X86_64" && Memory > 0 )
should_transfer_files = YES
when_to_transfer_output = ON_EXIT
queue
```

User-friendly use of COPASI with HTCondor ?

- Yes!
- Condor-COPASI is a package that will do that
- Needs to be installed in a Condor pool
- Registered users can then use it easily
- Condor-COPASI written in python, uses a web server
- Manages all Condor issues and emails user when job is complete
- Makes summary reports and plots automatically

`github.com/copasi/condor-copasi`

Kent E, Hoops S, Mendes P (2012) Condor-COPASI: High-Throughput Computing for Biochemical Networks. *BMC Systems Biology* **6**:91

Ed Kent

Kent *et al. BMC Systems Biology* 2012, **6**:91
http://www.biomedcentral.com/1752-0509/6/91

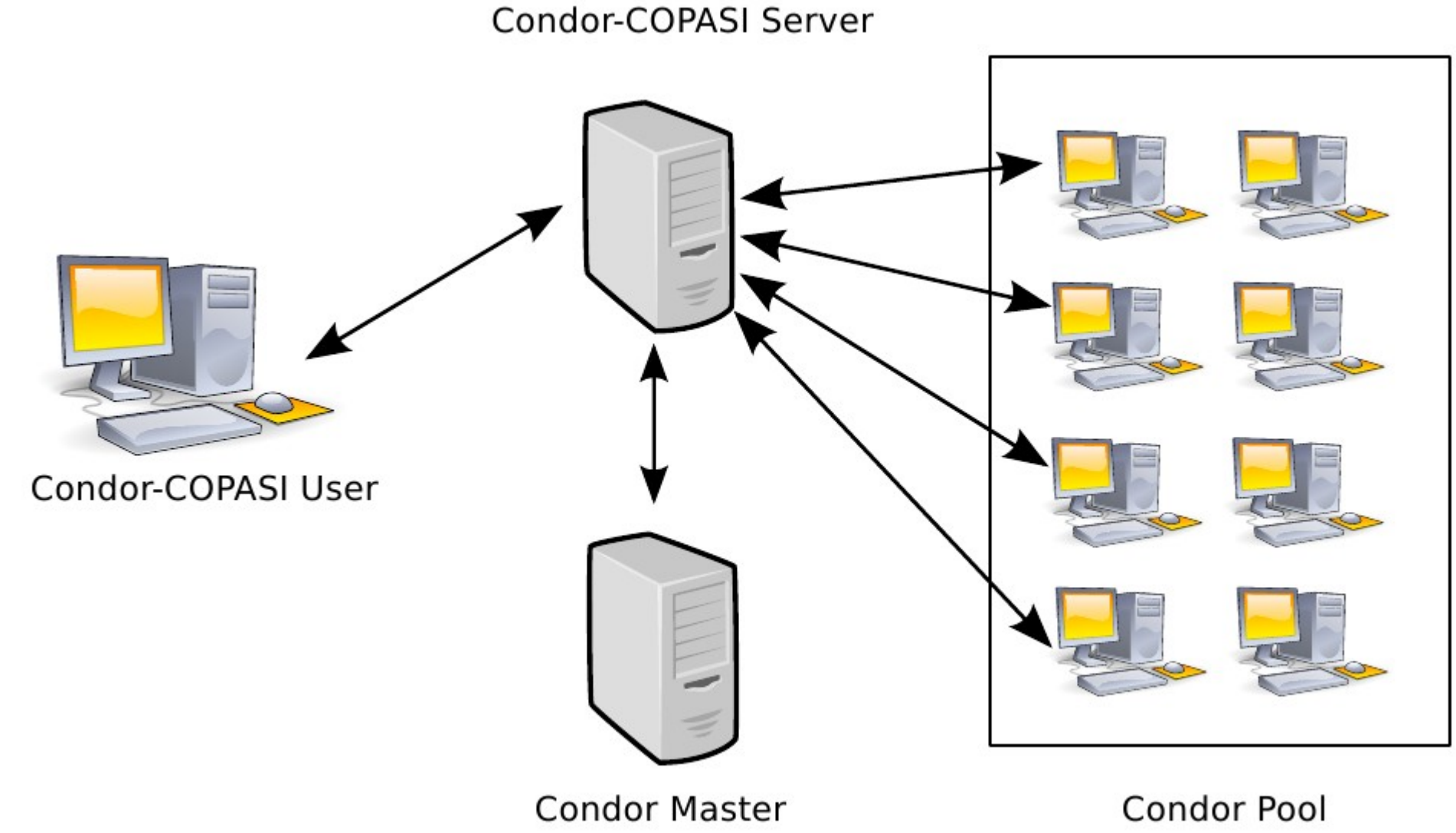
BMC
Systems Biology

SOFTWARE Open Access

Condor-COPASI: high-throughput computing for biochemical networks

Edward Kent[1], Stefan Hoops[2] and Pedro Mendes[2,3*]

Condor-COPASI high-throughput computing

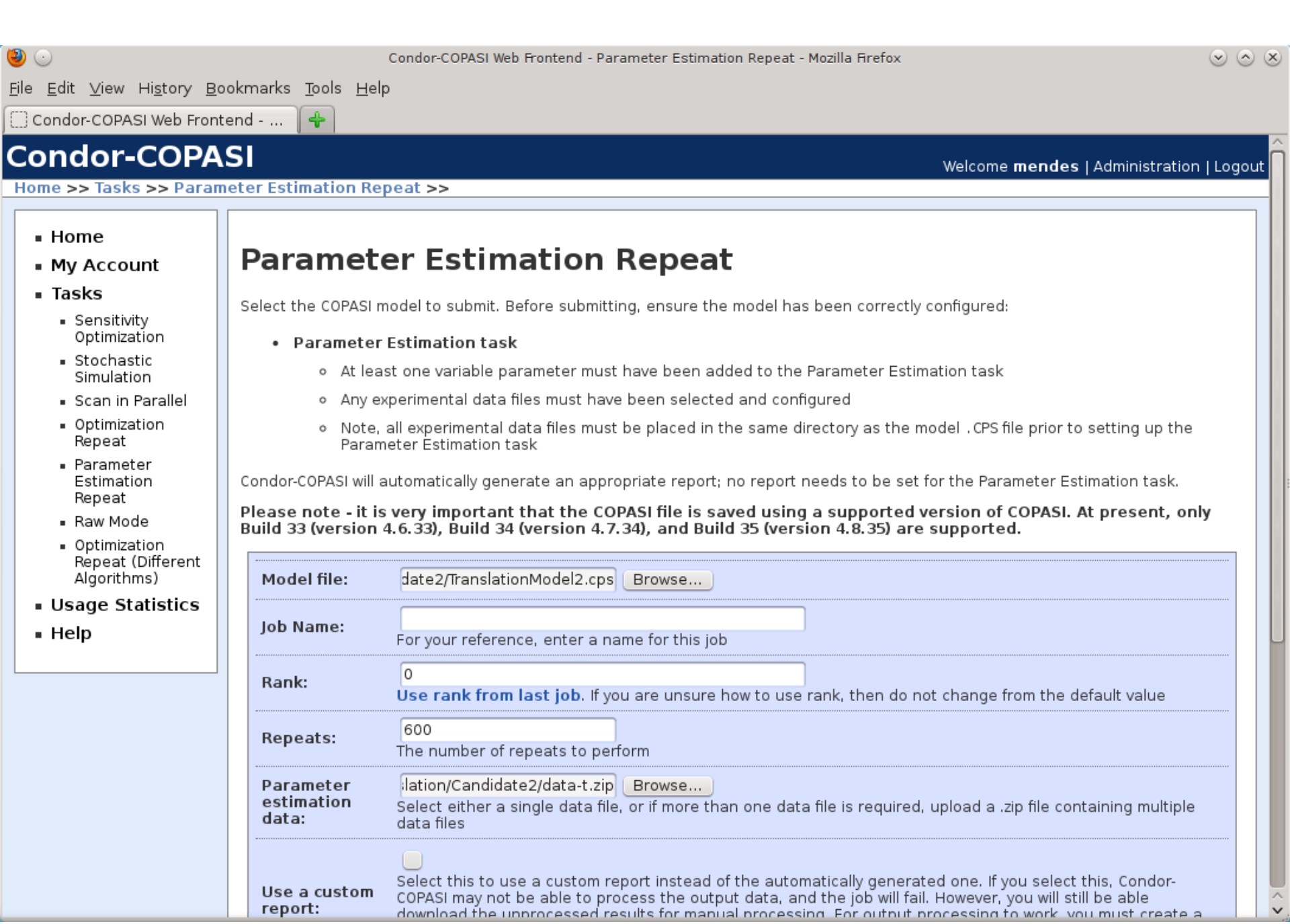


Condor-COPASI

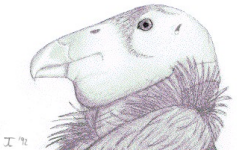
- Registered users (email needed)
- COPASI files with required tasks are uploaded to web server
- (Optionally) Condor-COPASI measures speed of each subtask
- Condor-COPASI breaks tasks in several subtasks, minimizing overhead costs
- Jobs are submitted to HTCondor
- Return results are collated in correct order
- Further processing of some results
- All results zipped
- User alerted by email that results are ready

Condor-COPASI Tasks

- **Stochastic simulation repeats**: same simulation repeated to draw time-dependent distributions
- **Parameter scans**: parameters varies and simulated repeated
- **Sensitivity optimization**: finding maximum and minimum values of local sensitivities in wide parameter space
- **Parameter estimation repeat**: repeats same (stochastic) optimization algorithm
- **Optimization repeat**: repeats same (stochastic) optimization algorithm
- **Parameter estimation with all algorithms**: runs the same problem with all algorithms available


Condor-COPASI Web Frontend - Parameter Estimation Repeat - Mozilla Firefox
File Edit View History Bookmarks Tools Help
Condor-COPASI Web Frontend - ...
Condor-COPASI
Welcome mendes | Administration | Logout
Home >> Tasks >> Parameter Estimation Repeat >>
Home
My Account
Tasks
Sensitivity Optimization
Stochastic Simulation
Scan in Parallel
Optimization Repeat
Parameter Estimation Repeat
Raw Mode
Optimization Repeat (Different Algorithms)
Usage Statistics
Help
Parameter Estimation Repeat
Select the COPASI model to submit. Before submitting, ensure the model has been correctly configured:
Parameter Estimation task
At least one variable parameter must have been added to the Parameter Estimation task
Any experimental data files must have been selected and configured
Note, all experimental data files must be placed in the same directory as the model .CPS file prior to setting up the Parameter Estimation task
Condor-COPASI will automatically generate an appropriate report; no report needs to be set for the Parameter Estimation task.
Please note - it is very important that the COPASI file is saved using a supported version of COPASI. At present, only Build 33 (version 4.6.33), Build 34 (version 4.7.34), and Build 35 (version 4.8.35) are supported.
Model file:
date2/TranslationModel2.cps
Browse...
Job Name:
For your reference, enter a name for this job
Rank:
0
Use rank from last job. If you are unsure how to use rank, then do not change from the default value
Repeats:
600
The number of repeats to perform
Parameter estimation data:
lation/Candidate2/data-t.zip
Browse...
Select either a single data file, or if more than one data file is required, upload a .zip file containing multiple data files
Use a custom report:
Select this to use a custom report instead of the automatically generated one. If you select this, Condor-COPASI may not be able to process the output data, and the job will fail. However, you will still be able download the unprocessed results for manual processing. For output processing to work, you must create a

Condor-COPASI high-throughput computing



Condor-COPASI Web Frontend - Stochastic Simulation - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Condor-COPASI Web Frontend - Stochas...

Condor-COPASI

Welcome **mendes** | Administration | Logout

Home >> Tasks >> Stochastic Simulation >>

- Home
- My Account
- Tasks
  - Sensitivity Optimization
  - Stochastic Simulation
  - Scan in Parallel
  - Optimization Repeat
  - Parameter Estimation Repeat
  - Optimization Repeat (Different Algorithms)
- Help

Stochastic Simulation

Select the COPASI model to submit. Before submitting, ensure the model has been correctly configured:

- **Time Course task:**
  - The Time Course should be set up as if a single run were to take place on the local machine
  - An appropriate stochastic method must be selected

Condor-COPASI will automatically generate an appropriate report; no report needs to be set for the Time Course task.

Please note - it is very important that the COPASI file is saved using a supported version of COPASI. At present, only Build 33 (version 4.6.33) and Build 34 (version 4.6.34) are supported.

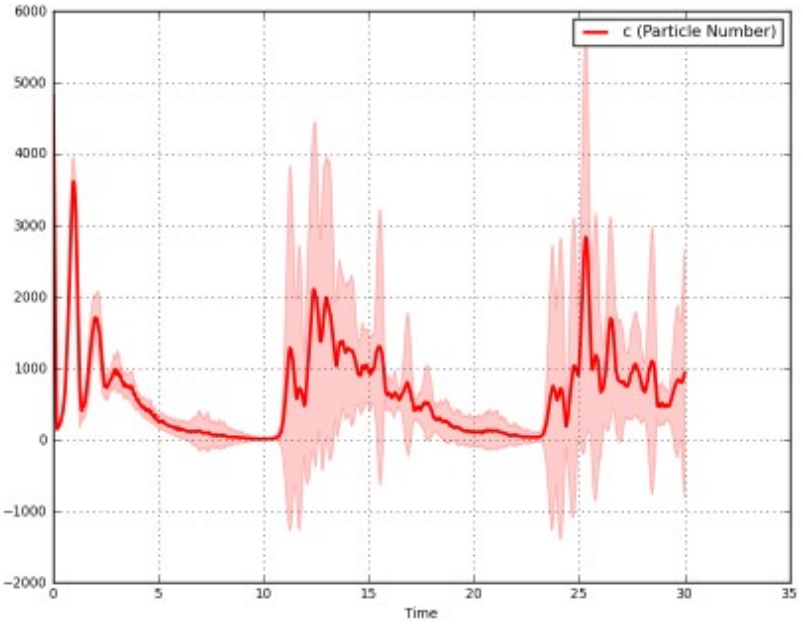
Model file: Browse...

Job Name:
For your reference, enter a name for this job

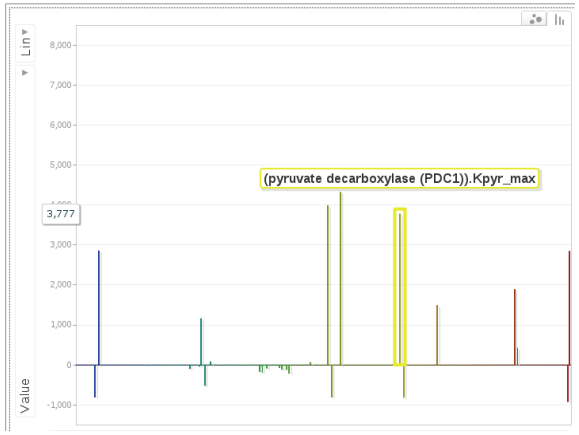
Skip load balancing step:
Select this to skip the automatic load balancing step, and make the run time of each parallel job as short as possible. **Use with caution! This has the potential to overload the Condor system with huge numbers of parallel jobs.** Not applicable for some job types - see documentation for further details.

Repeats:
The number of repeats to perform

Submit



Compare Global Sensitivity Job Output



Load-balancing

- How many subtasks should be submitted to each node?
- If too few, lots of jobs with high overhead
- If too many, very few jobs and little benefit from parallelization
- Ideally jobs should execute for ~20min

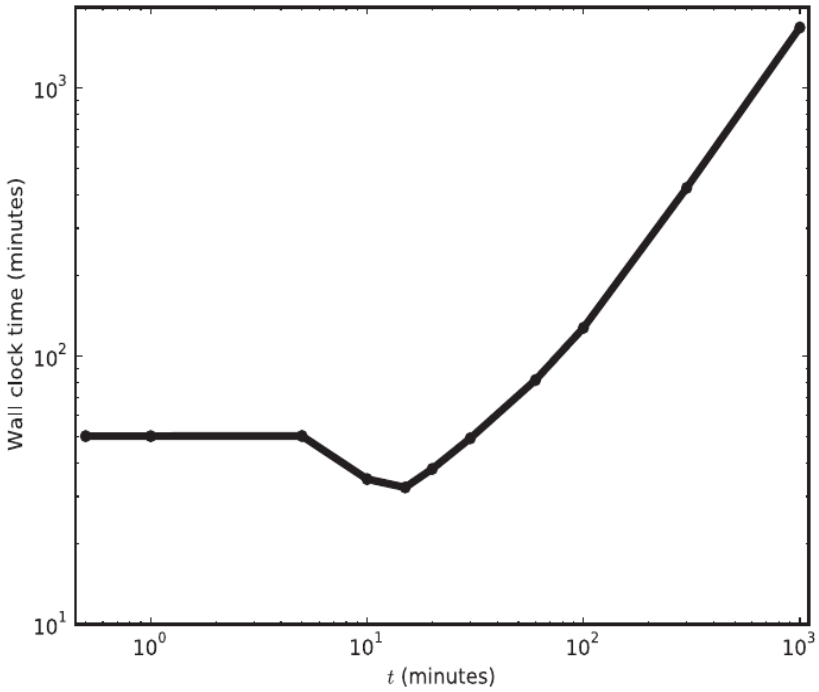
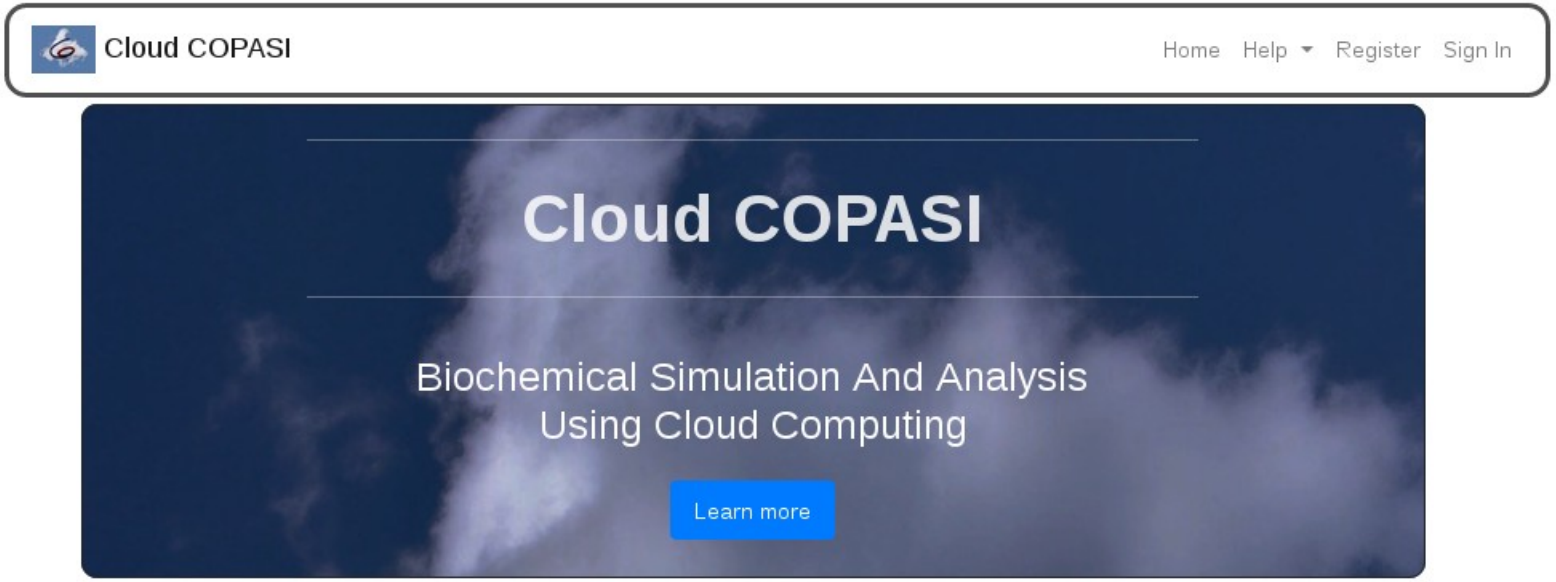


Figure 4 Tuning the load balancing algorithm.

Work in progress: Cloud-COPASI



- Updated web front-end (updated Django and Python 3)
- Uses Bosco to support different task managers
- Allow execution on AWS
- Add new use cases

Hasan Baig

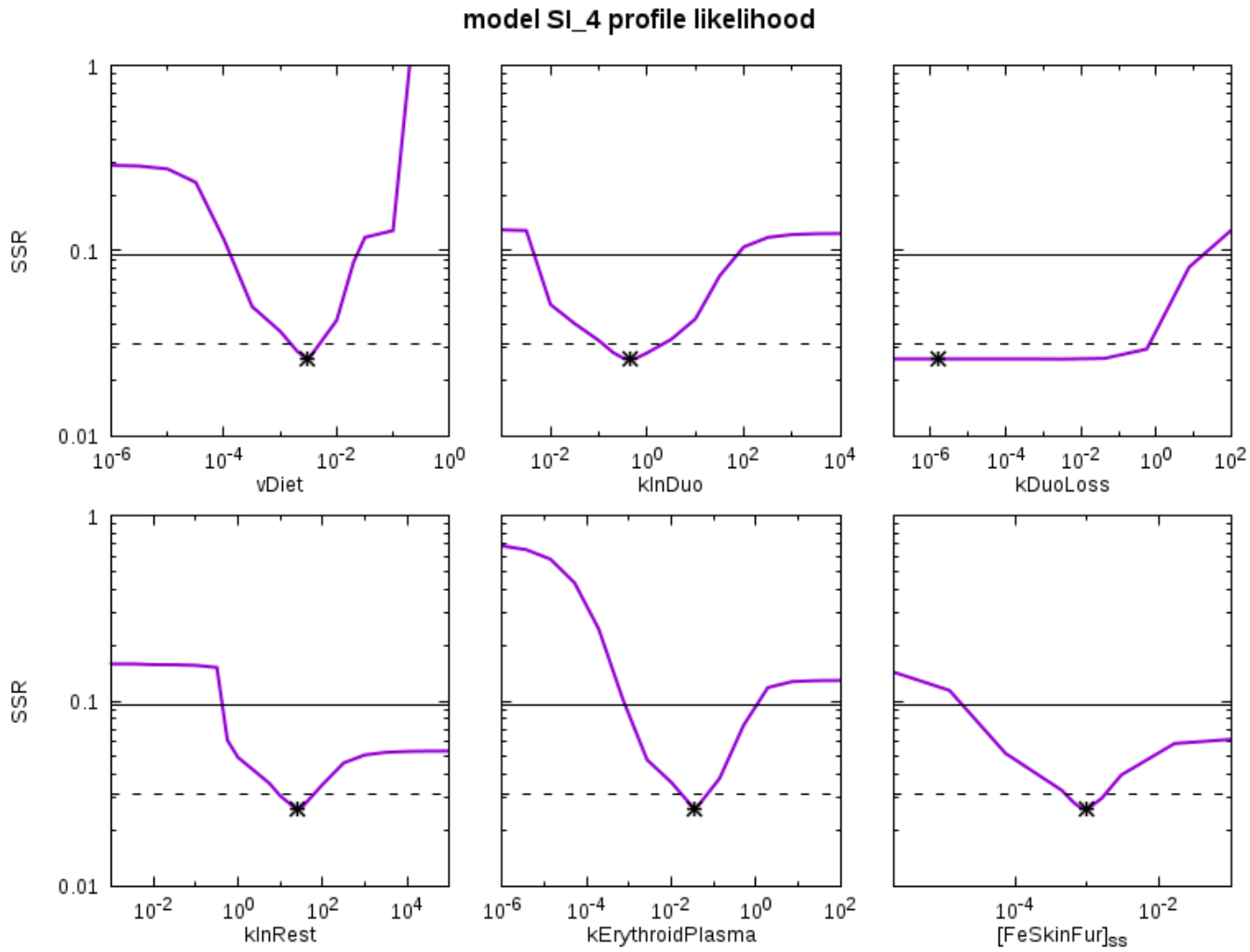
1D likelihood profiles for parameter identifiability analysis

$$LP(p_i) = min_{p_{j \neq i}}(SSR(p_j))$$

- Re-optimizing all parameters $p_{j \neq i}$ for fixed values of p_i around the solution
- Confidence intervals for each parameter can be estimated based on likelihood contours (C_{LC})or on likelihood ratios (C_{LR}), *e.g.*:

$$C_{LR} = \{ p : SSR(p) \leq SSR(\hat{p}) e^{\chi^2_\alpha / n} \}$$

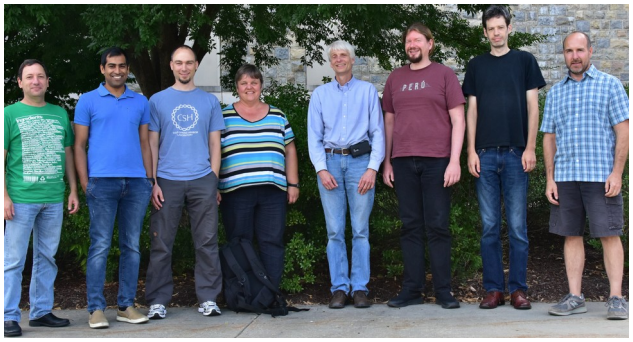
Example: 1D likelihood profiles to assess parameter identifiability



Acknowledgments

Ed Kent (Uni. Manchester)
Hasan Baig (Uconn Health)
Stefan Hoops (Uni. Virginia)
Brian Klahn (Uni. Virginia)

COPASI Team



Funding:

BBSRC and EPSRC DTC program

NIH: R01 GM080219

NIH: R24 GM137787, Mechanistic Modeling of Cellular Systems

(together with the VCell team: Les Loew, Ann Cowan, Ion Moraru, Michael Blinov, Boris Slepchenko)