FAST AND AUTOMATED EXPLORATION OF THE STATISTICAL PROPERTIES OF AGENT-BASED MODELS

Andrea Vandin

Sant'Anna School of Advanced Studies – Pisa

Institute of Economics

Department of Excellence 2018 - 2022

EMbeDS Economics and Management in the era of Data Science

DTU Danmarks Tekniske Universitet

Joint work with

Daniele Giachini, Francesco Lamperti, Francesca Chiaromonte

LEM Working Paper available at:
www.lem.sssup.it/WPLem/2020-31.html

Tool and models available at:
https://bit.ly/MultiVeStATool

Tools for Stochastic Modelling and Evaluation (performance, dependability, security and verification) 12/11/2021

FAST AND AUTOMATED EXPLORATION OF THE STATISTICAL PROPERTIES OF AGENT-BASED MODELS

Andrea Vandin

Sant'Anna School of Advanced Studies – Pisa

Institute of Economics

Department of Excellence 2018 - 2022

EMbeDS Economics and Management in the era of Data Science

DTU Danmarks Tekniske Universitet

But also

Stefano Sebastio, INRIA

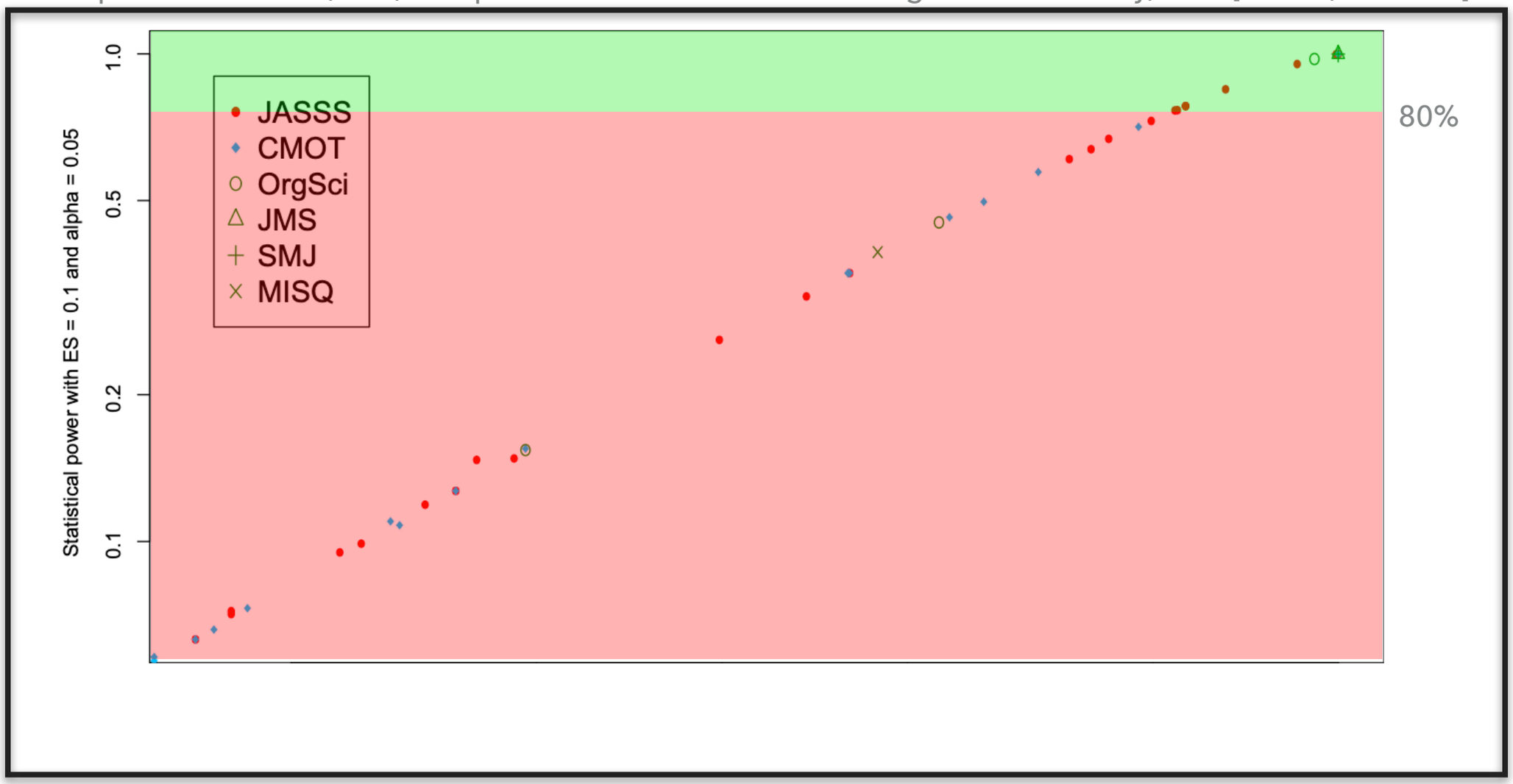
Stephen Gilmore, University of Edinburgh

Daniël Reijsbergen, Singapore University of Technology and Design

Tools for Stochastic Modelling and Evaluation (performance, dependability, security and verification) 12/11/2021

‘Quality’ of Statistical Analysis on 55 ABM from Management & Organisational Research

Adapted from Secchi, Seri, Computational and Mathematical Organization Theory, 2017[Secchi,Seri2017]



- The importance of designing well simulation-based analysis.
  - Power analysis on ‘are the expected outcomes of different configurations of parameters the same’?
- Power is 1 - P(Type II error)
  - Roughly, P(test=‘outcomes are different’ | outcomes are different)
  - “The value that seems to be more commonly accepted is 80%”
- “We need to encourage researchers to be more precise in the determination of the number of runs”

Similar studies can be found also in other communities

A systematic review of statistical power in software engineering experiments

Tore Dybå [a,b,*], Vigdis By Kampenes [a], Dag I.K. Sjøberg [a]

[a] *Simula Research Laboratory, P.O. Box 134, NO-1325 Lysaker, Norway*
[b] *SINTEF ICT, NO-7465 Trondheim, Norway*

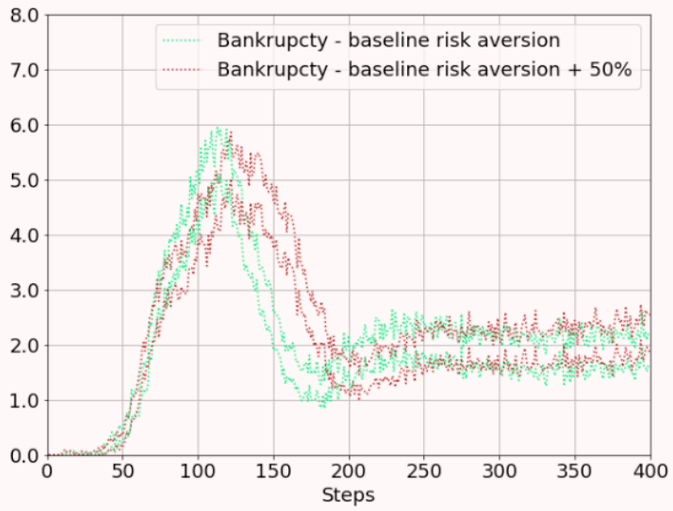
Received 11 May 2005; revised 24 August 2005; accepted 31 August 2005
Available online 3 November 2005

Abstract

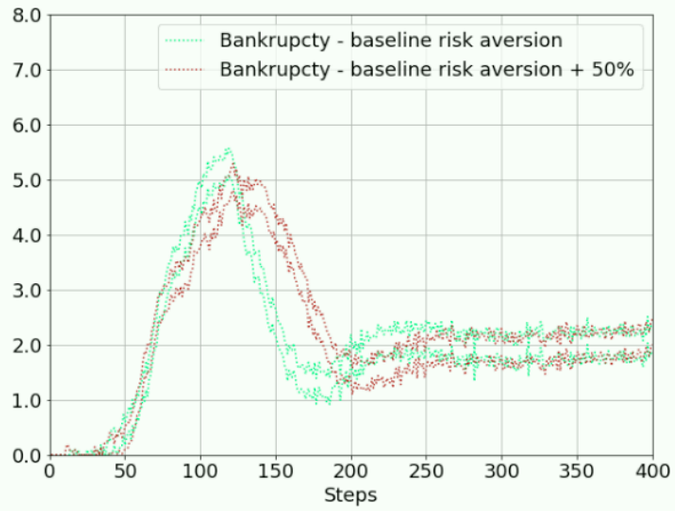
Statistical power is an inherent part of empirical studies that employ significance testing and is essential for the planning of studies, for the interpretation of study results, and for the validity of study conclusions. This paper reports a quantitative assessment of the statistical power of empirical software engineering research based on the 103 papers on controlled experiments (of a total of 5,453 papers) published in nine major software engineering journals and three conference proceedings in the decade 1993–2002. The results show that the statistical power of software engineering experiments falls substantially below accepted norms as well as the levels found in the related discipline of information systems research. Given this study's findings, additional attention must be directed to the adequacy of sample sizes and research designs to ensure acceptable levels of statistical power. Furthermore, the current reporting of significance tests should be enhanced by also reporting effect sizes and confidence intervals.

Statistically Meaningful Counterfactual Analysis

97.5% CI
100 Simulations



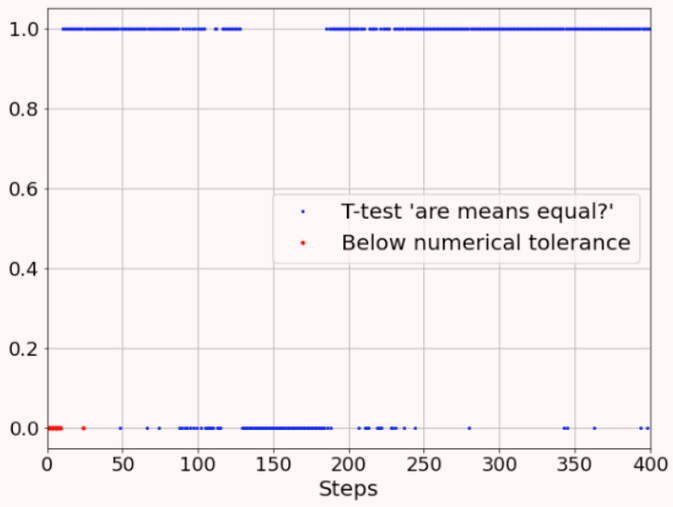
(a) CIs width for $\alpha = 0.025$ and $N = 100$ simulations



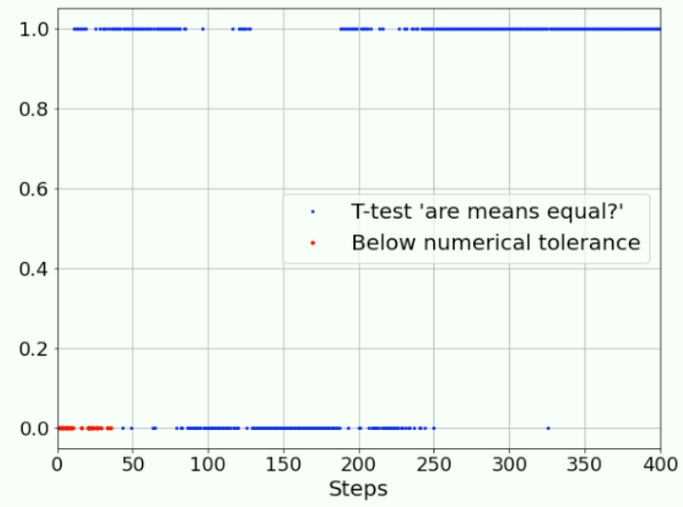
(b) CIs width for $\alpha = 0.025$ and $\delta = 0.5$

97.5% CI
MultiVeStA
'**Right**' number of simulations

Welch's t-test



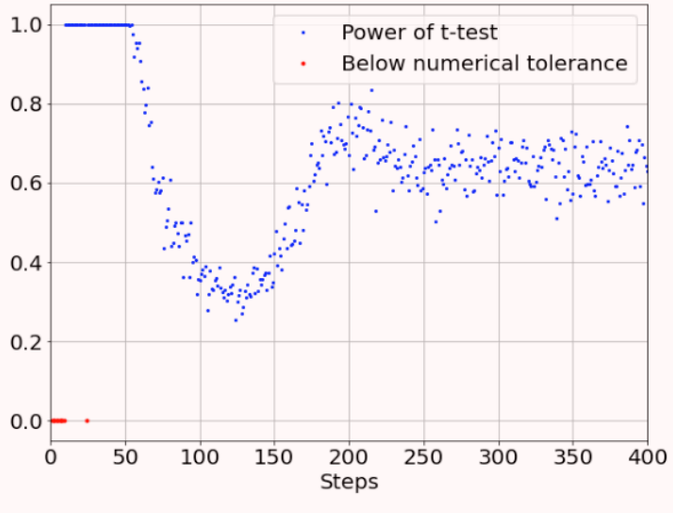
(c) T-test *are means in (a) point-wise equal* for significance $\alpha = 0.025$



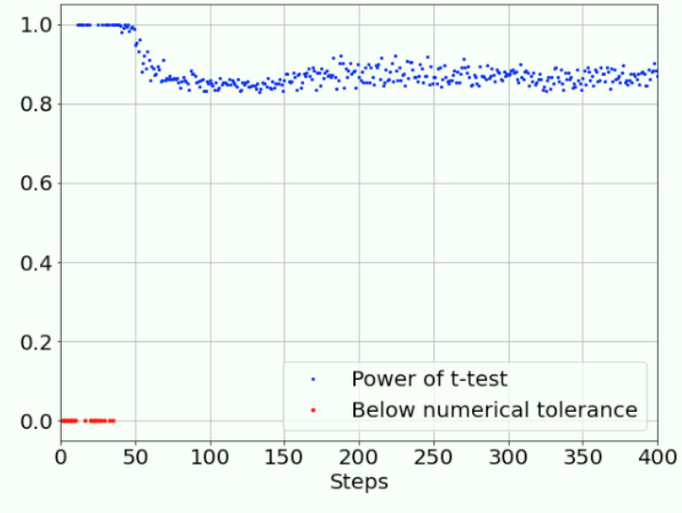
(d) T-test *are means in (b) point-wise equal* for significance $\alpha = 0.025$

Welch's t-test

Power of the test
P(Test=0 | Real=0)



(e) Power of t-test in (c) for difference $\varepsilon = 0.5$



(f) Power of t-test in (d) for difference $\varepsilon = 0.5$

Power of the test
P(Test=0 | Real=0)

Steady-State Analysis: Market Selection

Arbitrary choice of
- Number of sims
- Warmup period
- Time horizon

from [Kets et al2014]

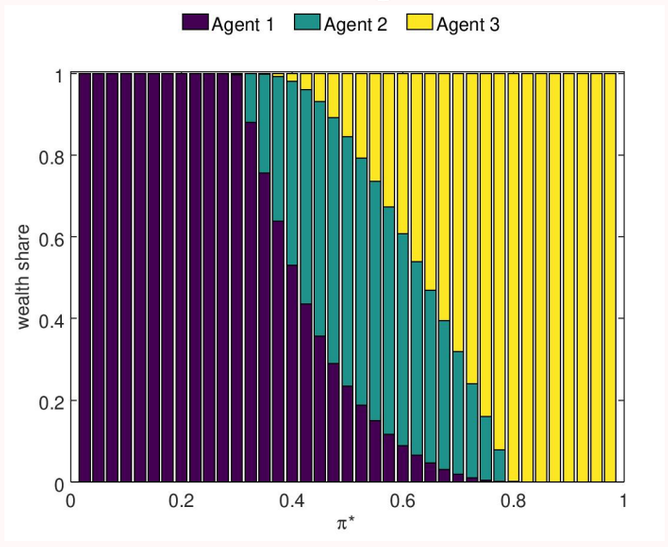
Automated choice of
- Number of sims
- Warmup period
- Time horizon

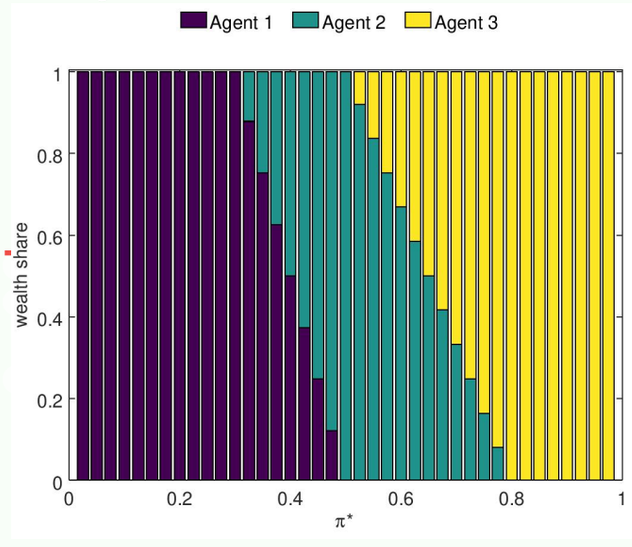
MultiVeStA

Same as analytical solution

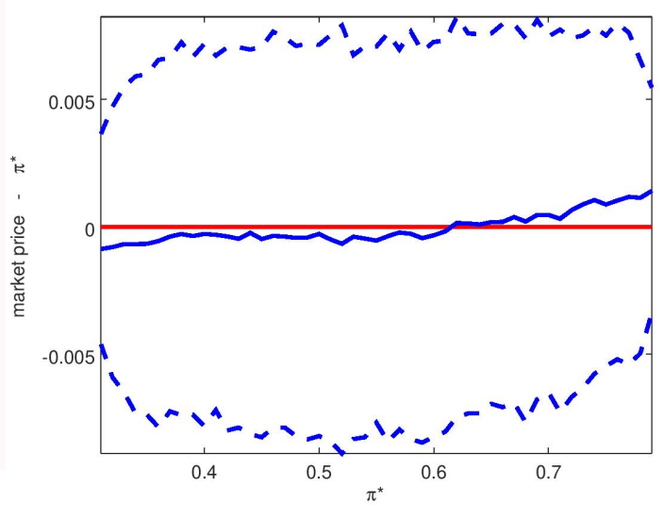
from [Bottazzi,Giachini2019]

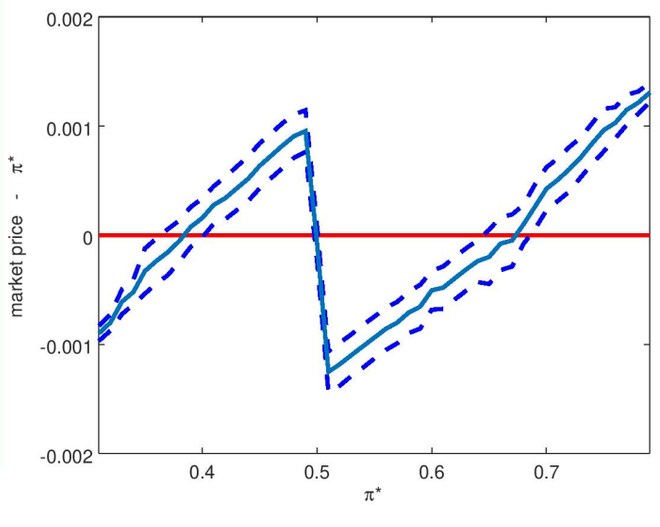
Agents wealth at steady state



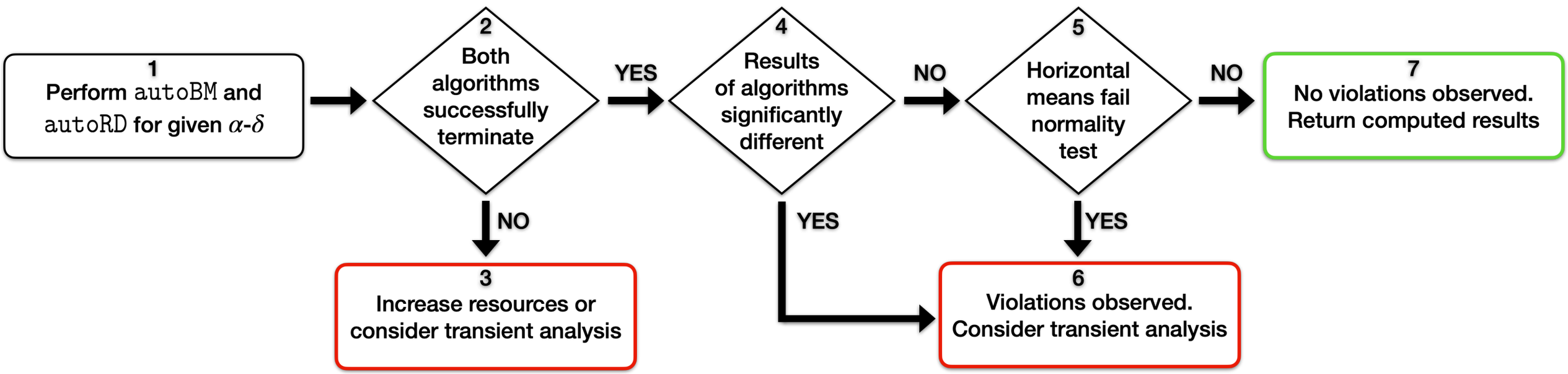


Does the market price match π^*?





A Methodology for Ergodicity Diagnostics



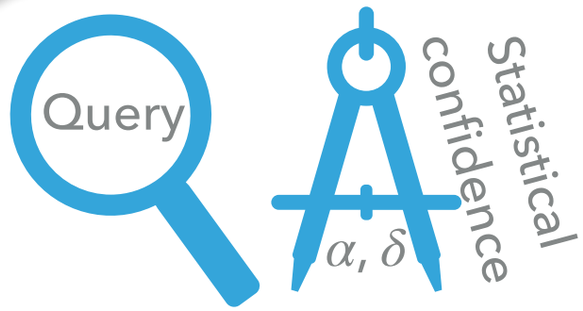
Our Proposed Approach to Simulation-Based Analysis

newstalkzb.co.nz/news/education/modern-lego-sets-more-complex-less-inspiring/

https://www.alamy.com/

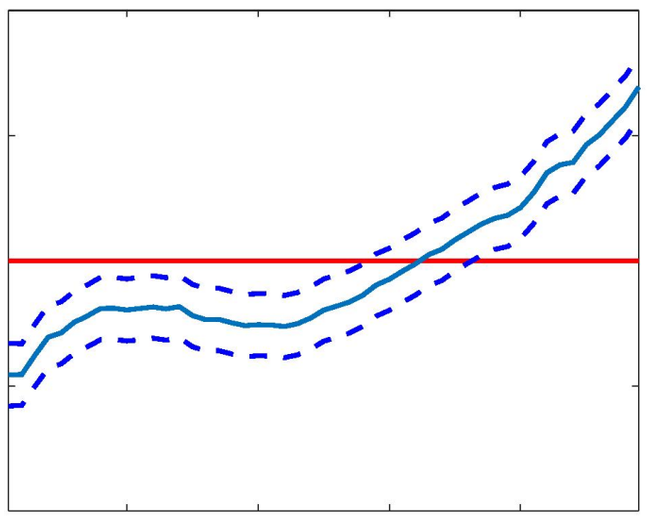
Handcrafted

- Mainly manual process
  - Time-consuming
  - Problems with replicability
  - Error-prone
    - Modify model, interpret CSV
- Ad-hoc implementations
  - Reliability? Efficiency?



Statistical Model Checking

- Automatic
  - Time-saving and Reproducible
  - Promotes use of *standard* analysis
- Reference implementation
  - Reliable and Efficient



MultiVeStA: SMC For Discrete-Event Simulators

Decides number of steps and of simulations, and performs analysis by interacting with the simulator

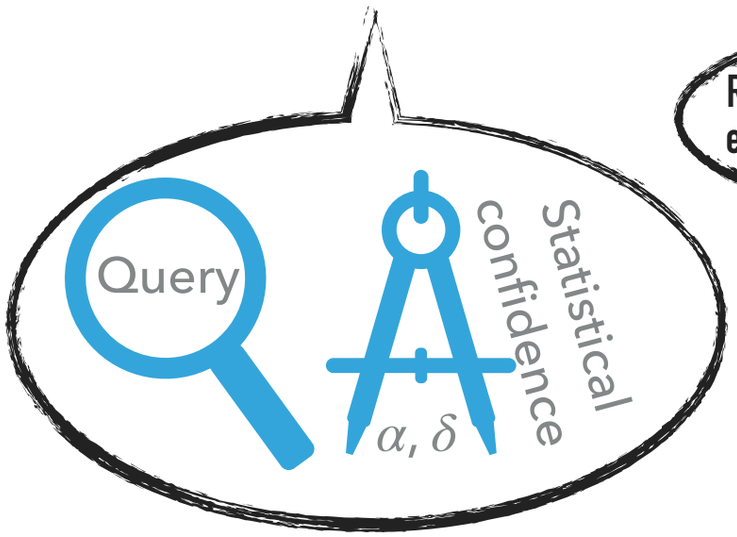
Performs basic tasks as required by MultiVeStA: reinitialize, perform one step, evaluate observation

MultiVeStA

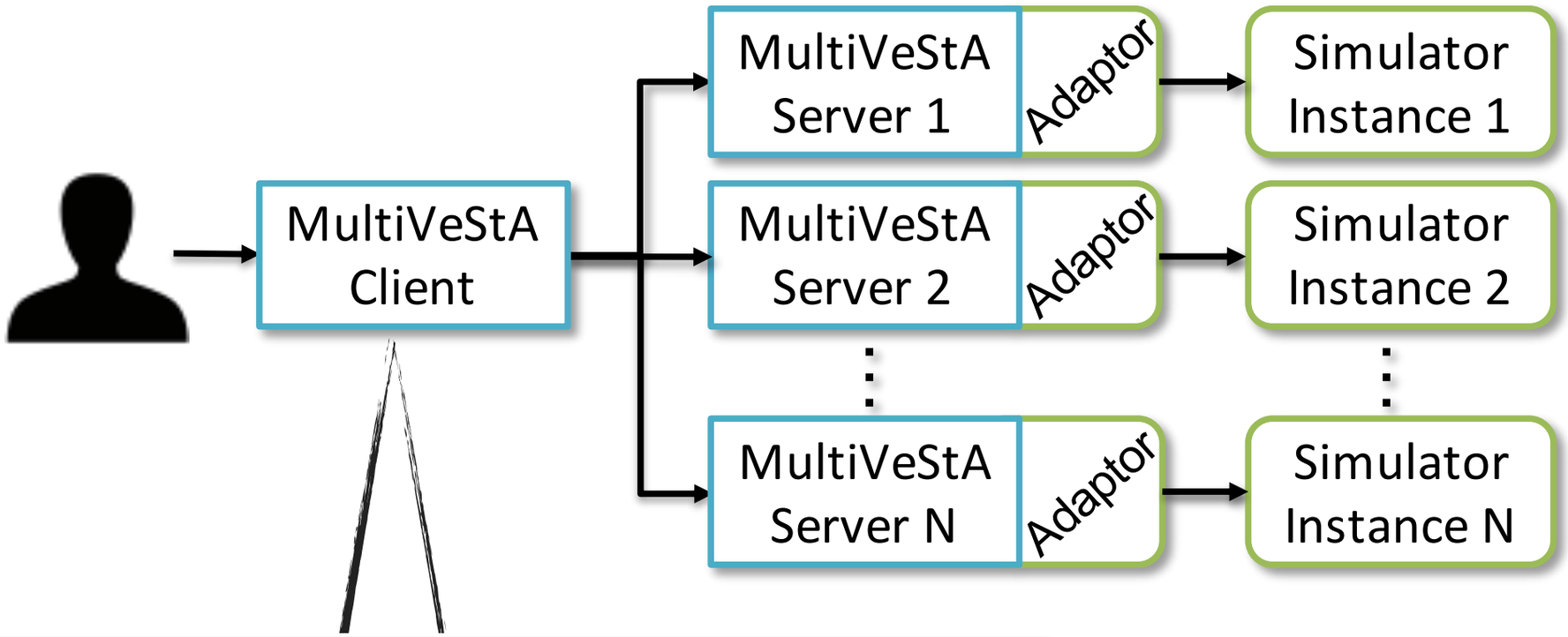
Simulator

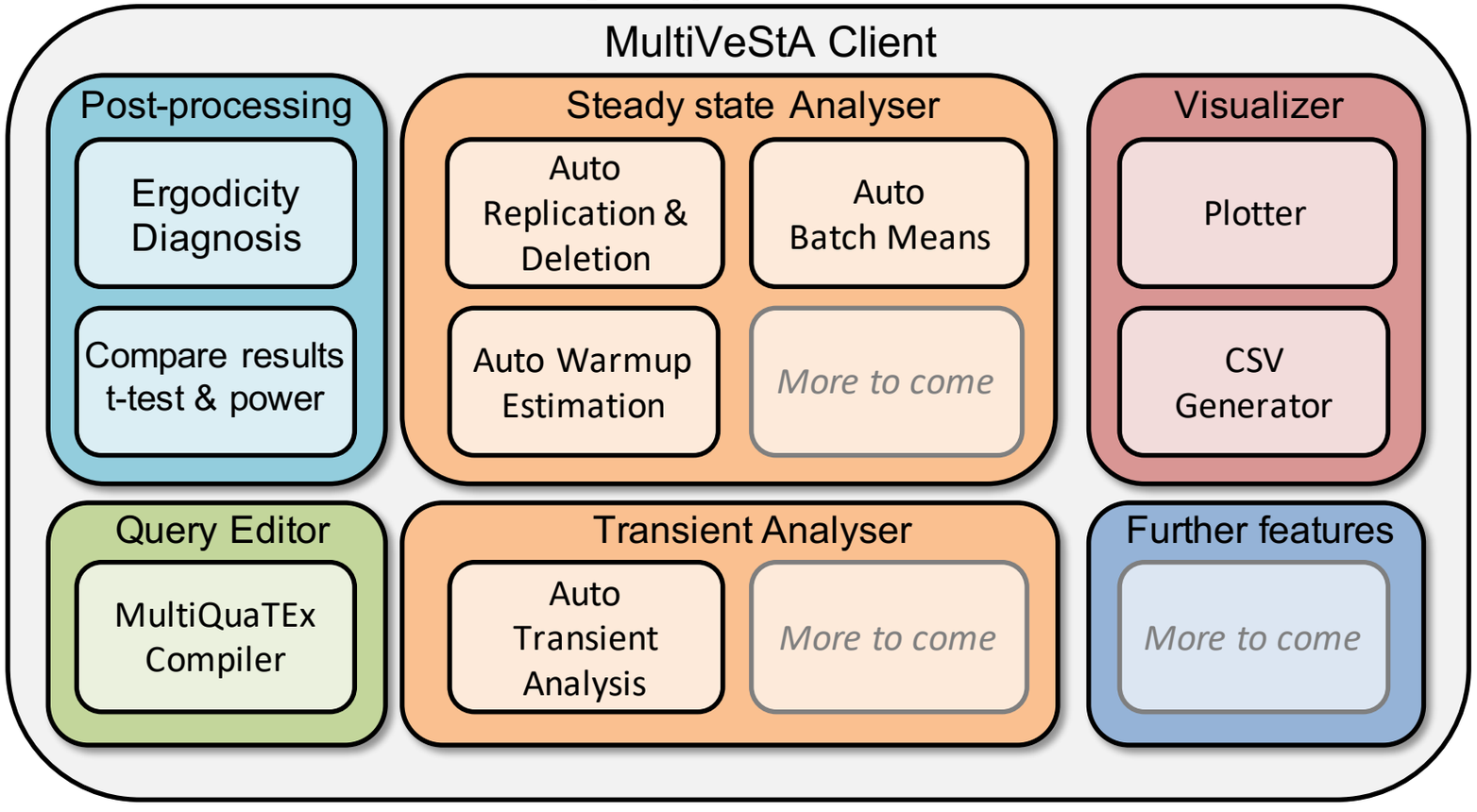
Returns to the user plots and CSV files with estimated means, confidence intervals, . . .

Can be kept to a 'minimum complexity' No need to implement analysis techniques

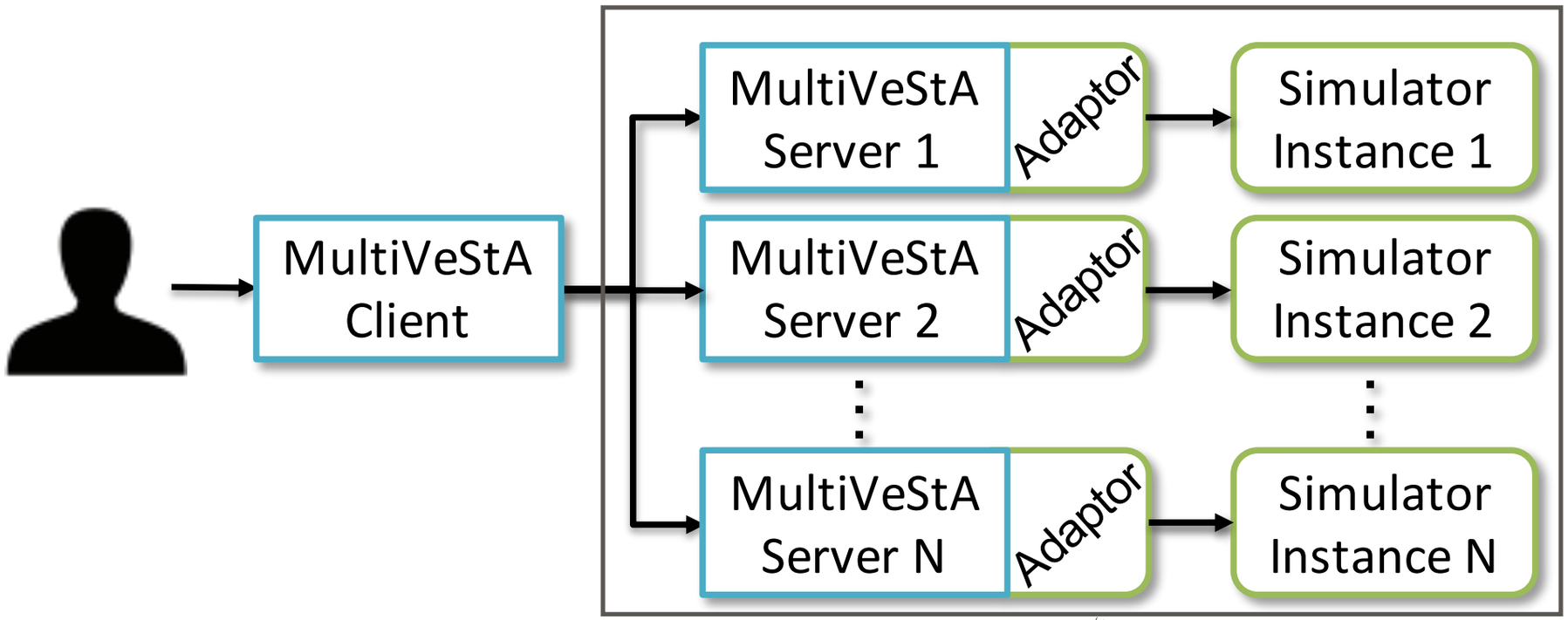


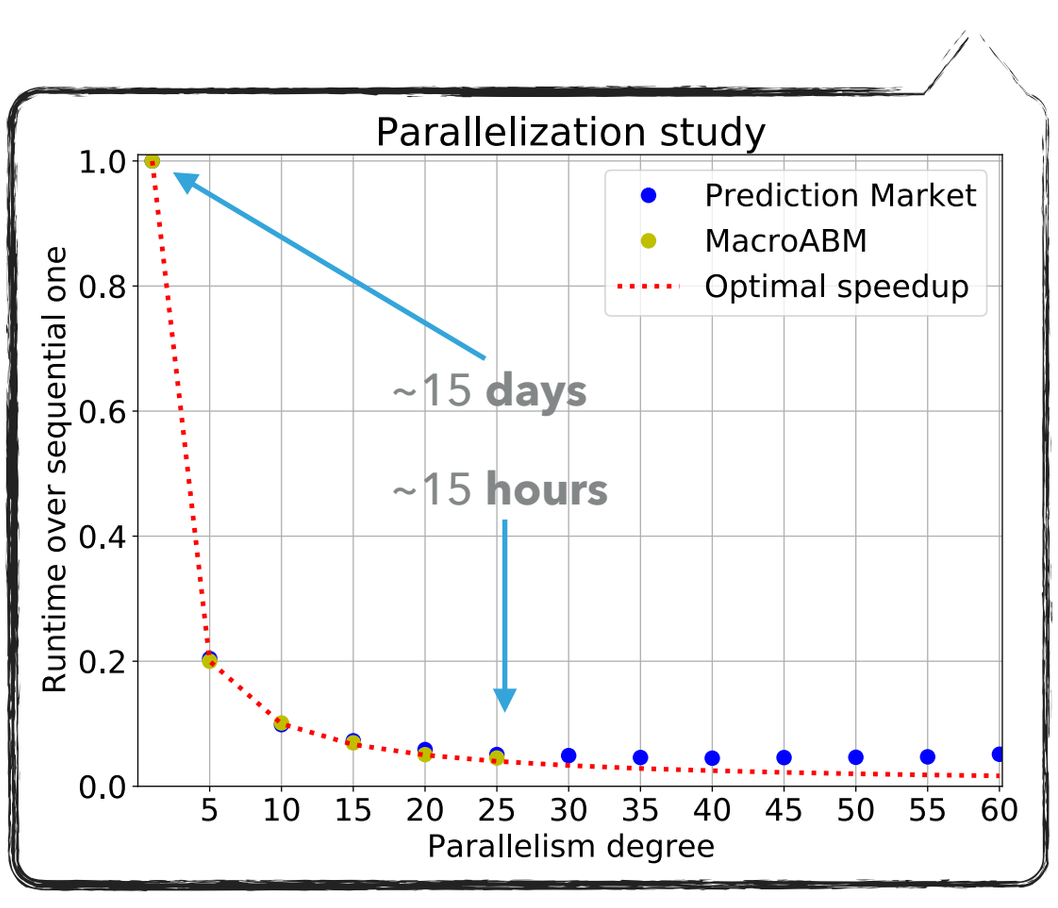
MultiVeStA: SMC For Discrete-Event Simulators

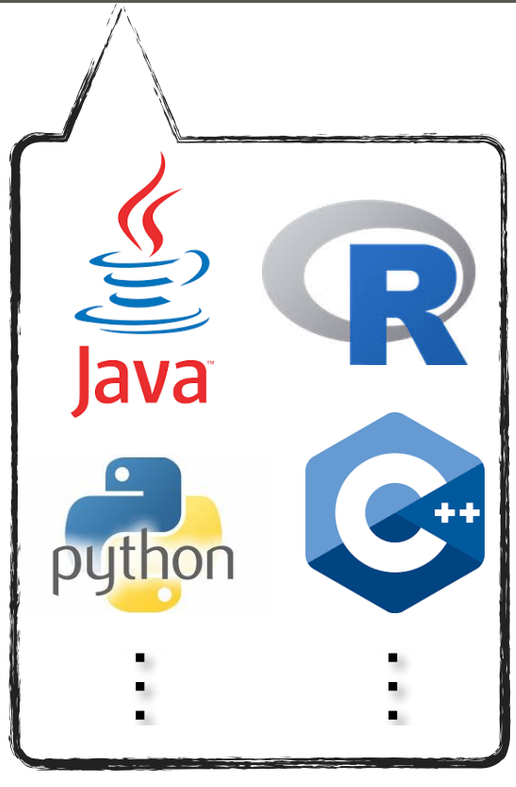




MultiVeStA: SMC For Discrete-Event Simulators







REFERENCES

Papers on MultiVeStA

- Automated and Distributed Statistical Analysis of Economic Agent-Based Models. A Vandin, D Giachini, F Lamperti, F Chiaromonte. Under revision at Journal of Economic Dynamics and Control
- Transient and steady-state statistical analysis for discrete event simulators. S Gilmore, D Reijsbergen, A Vandin IFM2017
- MultiVeStA: Statistical model checking for discrete event simulators. S Sebastio, A Vandin. ValueTools2013

Papers using MultiVeStA

- A formal approach for the analysis of BPMN collaboration models. F Corradini, F Fornari, A Polini, B Re, F Tiezzi, A Vandin. JSS (2021)
- Quantitative Security Risk Modeling and Analysis with RisQFLan. MH ter Beek, A Legay, A Lluch Lafuente, A Vandin. COSE (2021)
- Variability meets Security. MH ter Beek, A Legay, A Lluch Lafuente, A Vandin. VAMOS2021
- Statistical analysis of CARMA models: an advanced tutorial. V Galpin, A Georgoulas, M Loreti, A Vandin. WSC2021
- Qflan: A tool for the quantitative analysis of highly reconfigurable systems. A Vandin, MH Ter Beek, A Legay, A Lluch Lafuente. FM2018
- A framework for quantitative modeling and analysis of highly (re) configurable systems. MH ter Beek, A Legay, A Lluch Lafuente, A Vandin. TSE (2018)
- Statistical model checking for product lines. MH Ter Beek, A Legay, A Lluch Lafuente, A Vandin. ISOLA2016
- A tool-chain for statistical spatio-temporal model checking of bike sharing systems. V Ciancia, D Latella, M Massink, R Paškauskas, A Vandin. ISOLA2016
- Statistical analysis of probabilistic models of software product lines with quantitative constraints. MH ter Beek, A Legay, A Lluch Lafuente, A Vandin. SPLC2015
- Quantitative analysis of probabilistic models of software product lines with statistical model checking. MH ter Beek, A Legay, A Lluch Lafuente, A Vandin. FMSPLE2015
- Modelling and analyzing adaptive self-assembly strategies with Maude. R Bruni, A Corradini, F Gadducci, A Lluch Lafuente, A Vandin. SCP (2015)
- The SCEL language: design, implementation, verification. R De Nicola et al. Software Engineering for Collective Autonomic Systems (2015)
- Tools for Ensemble Design and Runtime. Software Engineering for Collective Autonomic Systems 2015
- An analysis pathway for the quantitative evaluation of public transport systems. S Gilmore, M Tribastone, A Vandin. IFM2014
- Distributed statistical analysis of complex systems modeled through a chemical metaphor. D Pianini, S Sebastio, A Vandin. HPCS2014
- Reasoning (on) service component ensembles in rewriting logic. L Belzner, R De Nicola, A Vandin, M Wirsing. SAS2014
- ....

THANK YOU FOR YOUR ATTENTION!

QUESTIONS?

FEEDBACK?

LEM Working Paper available at:
www.lem.sssup.it/WPLem/2020-31.html

Tool and models available at:
https://bit.ly/MultiVeStATool