

Reflections on the Simulation of Complex Systems for Science

Fiona A. C. Polack¹

Paul S. Andrews¹ Adam T. Sampson²

Teodor Ghetiu¹ Mark Read¹ Susan Stepney¹

Jon Timmis³

¹Department of Computer Science, University of York, UK

²Computing Laboratory, University of Kent, UK

³Departments of Computer Science and of Electronics, University of York, UK

ICECCS March 2010



Complex systems simulation for science

Simulation: A model executed on a computer to reveal time-dependent behaviour

- e.g. Agent-based models to support scientific research
 - ▶ Modelling at individual level
 - ▶ Inherent modelling in time and space
- Problem: “opaque thought experiments” [Bullock et al]
 - ▶ Throw-away simulations that only address one question
 - ▶ Naive simulations that scientists cannot use effectively
 - ▶ Simulations that are not validated w.r.t. science



Complex systems simulation for science

Simulation: A model executed on a computer to reveal time-dependent behaviour

- e.g. Agent-based models to support scientific research
 - ▶ Modelling at individual level
 - ▶ Inherent modelling in time and space
- Problem: “opaque thought experiments” [Bullock et al]
 - ▶ Throw-away simulations that only address one question
 - ▶ Naive simulations that scientists cannot use effectively
 - ▶ Simulations that are not validated w.r.t. science



Complex systems simulation for science

Simulation: A model executed on a computer to reveal time-dependent behaviour

- e.g. Agent-based models to support scientific research
 - ▶ Modelling at individual level
 - ▶ Inherent modelling in time and space
- Problem: “opaque thought experiments” [Bullock et al]
 - ▶ Throw-away simulations that only address one question
 - ▶ Naive simulations that scientists cannot use effectively
 - ▶ Simulations that are not validated w.r.t. science



“Scientists don’t understand computer simulations”

But... scientists use other poorly understood tools:

- ▶ heuristic solutions
 - ▶ problems re invalid statistics and ODE solutions
- ▶ critical simplifications:
 - ▶ eg. “stable” communities used to model dynamic equilibria

And... scientists do construct their own simulations:

- ▶ one-off models to explore specific issues
- ▶ not engineered or validated
 - ▶ limited scalability and modifiability

Need simulations that scientists trust and understand.



“Scientists don’t understand computer simulations”

But... scientists use other poorly understood tools:

- ▶ heuristic solutions
 - ▶ problems re invalid statistics and ODE solutions
- ▶ critical simplifications:
 - ▶ eg. “stable” communities used to model dynamic equilibria

And... scientists do construct their own simulations:

- ▶ one-off models to explore specific issues
- ▶ not engineered or validated
 - ▶ limited scalability and modifiability

Need simulations that scientists trust and understand.



“Scientists don’t understand computer simulations”

But... scientists use other poorly understood tools:

- ▶ heuristic solutions
 - ▶ problems re invalid statistics and ODE solutions
- ▶ critical simplifications:
 - ▶ eg. “stable” communities used to model dynamic equilibria

And... scientists do construct their own simulations:

- ▶ one-off models to explore specific issues
- ▶ not engineered or validated
 - ▶ limited scalability and modifiability

Need simulations that scientists trust and understand.



“Scientists don’t understand computer simulations”

But... scientists use other poorly understood tools:

- ▶ heuristic solutions
 - ▶ problems re invalid statistics and ODE solutions
- ▶ critical simplifications:
 - ▶ eg. “stable” communities used to model dynamic equilibria

And... scientists do construct their own simulations:

- ▶ one-off models to explore specific issues
- ▶ not engineered or validated
 - ▶ limited scalability and modifiability

Need simulations that scientists trust and understand.



“Scientists don’t understand computer simulations”

But... scientists use other poorly understood tools:

- ▶ heuristic solutions
 - ▶ problems re invalid statistics and ODE solutions
- ▶ critical simplifications:
 - ▶ eg. “stable” communities used to model dynamic equilibria

And... scientists do construct their own simulations:

- ▶ one-off models to explore specific issues
- ▶ not engineered or validated
 - ▶ limited scalability and modifiability

Need simulations that scientists trust and understand.



“Scientists don’t understand computer simulations”

But... scientists use other poorly understood tools:

- ▶ heuristic solutions
 - ▶ problems re invalid statistics and ODE solutions
- ▶ critical simplifications:
 - ▶ eg. “stable” communities used to model dynamic equilibria

And... scientists do construct their own simulations:

- ▶ one-off models to explore specific issues
- ▶ not engineered or validated
 - ▶ limited scalability and modifiability

Need simulations that scientists trust and understand.



Can we engineer trustworthy simulation for science?

Scientific motivations:

- ▶ problems that cannot be addressed using current techniques

Trustworthiness:

- ▶ apply best (appropriate) engineering practice
 - ▶ demonstrably reliable procedures and results
 - ▶ eg. continuity in modelling and implementation
 - ▶ eg. challenging assumptions
 - ▶ explicit decisions re abstraction and design
- ▶ tailor practices to working in collaboration

Good engineering is implicit in trusted simulations: eg. Calder et al., Cohen and Harel et al.



Can we engineer trustworthy simulation for science?

Scientific motivations:

- ▶ problems that cannot be addressed using current techniques

Trustworthiness:

- ▶ apply best (appropriate) engineering practice
 - ▶ demonstrably reliable procedures and results
 - ▶ eg. continuity in modelling and implementation
 - ▶ eg. challenging assumptions
 - ▶ explicit decisions re abstraction and design
- ▶ tailor practices to working in collaboration

Good engineering is implicit in trusted simulations: eg. Calder et al., Cohen and Harel et al.



Can we engineer trustworthy simulation for science?

Scientific motivations:

- ▶ problems that cannot be addressed using current techniques

Trustworthiness:

- ▶ apply best (appropriate) engineering practice
 - ▶ demonstrably reliable procedures and results
 - ▶ eg. continuity in modelling and implementation
 - ▶ eg. challenging assumptions
 - ▶ explicit decisions re abstraction and design
- ▶ tailor practices to working in collaboration

Good engineering is implicit in trusted simulations: eg. Calder et al., Cohen and Harel et al.



BIG Question: When is a simulation valid?

- **When is a simulation valid?**
 - ▶ When it looks ok?
 - ▶ When it produces the right results?
 - ▶ When its internals are like the real system?
- Need acceptable results from adequately equivalent processes
- Requires collaboration between scientists and simulation engineers
 - ▶ CoSMoS-related examples used to explore collaborative working
 - ▶ Ongoing CoSMoS work on validation of simulations



BIG Question: When is a simulation valid?

- When is a simulation valid?
 - ▶ When it looks ok?
 - ▶ When it produces the right results?
 - ▶ When its internals are like the real system?
- Need acceptable results from adequately equivalent processes
- Requires collaboration between scientists and simulation engineers
 - ▶ CoSMoS-related examples used to explore collaborative working
 - ▶ Ongoing CoSMoS work on validation of simulations



BIG Question: When is a simulation valid?

- When is a simulation valid?
 - ▶ When it looks ok?
 - ▶ When it produces the right results?
 - ▶ When its internals are like the real system?
- Need **acceptable results** from **adequately equivalent processes**
- Requires collaboration between scientists and simulation engineers
 - ▶ CoSMoS-related examples used to explore collaborative working
 - ▶ Ongoing CoSMoS work on validation of simulations



CoSMoS-related examples of simulation

Title	Scientists	Engineer	Format
Lymphocyte migration	Coles, Scott: York CII	Andrews, Sampson	CS developers with CII collaborator
Granuloma formation	Moore, Kaye: York CII	Flügge	CS MSc with CII collaborator
Autoimmune EAE	Kumar: Torrey Pines	Read	CS PhD working with TP collaborator
Species-abundance Ecology	Bown: Abertay	Ghetiu	CS PhD co-supervised
Auxin transport	Leyser: Biology York	Garnett	CS PhD co-supervised by scientist

Observations for Successful Simulation

- ▶ Clarity of roles and objectives
 - ▶ Scientist responsible for science
 - ▶ Engineer responsible for simulation
- ▶ Strong commitment to collaboration
 - ▶ Time and effort on both sides
 - ▶ Willingness to engage with the other side's problems

Here, focus on the roles in relation to the Domain and Domain Modelling



Domain and Domain Model

Establishing roles and commitments of collaboration:

- ▶ Scientists own and represent the **domain**: the science
 - ▶ sort out contentious or problematic scientific issues
 - ▶ provide and interpret information about the domain
 - ▶ work out how simulation results contribute to their research
- ▶ Jointly-owned **domain model** is a representation of the scientists' view of part of the domain
 - ▶ Collaborate to find appropriate levels of abstraction
 - ▶ Engineer needs to understand the view of science
 - ▶ Engineer is responsible for abstracting to an agreed model that can map to agents and the agents' environment
- ▶ Joint interpretation of results



Domain to Domain Model: examples 1

Lymphocyte migration domain is leishmaniasis research at York Centre for Immunology and Infection

- ▶ Domain model derived and developed (over time) through:
 - ▶ guided access to and discussion of relevant immunology literature
 - ▶ informal modelling and discussion of potential abstractions

Autoimmune EAE and Auxin Transport examples are similar

- ▶ EAE domain: Kumar, Torrey Pines Inst. for Molecular Studies, CA
- ▶ Auxin domain: Leyser, Biology, York

Domain modelling involves a significant investment in mutual understanding

- ▶ the trust-based relationship grows out of mutual respect and this learning process



Domain to Domain Model: examples 1

Lymphocyte migration domain is leishmaniasis research at York Centre for Immunology and Infection

- ▶ Domain model derived and developed (over time) through:
 - ▶ guided access to and discussion of relevant immunology literature
 - ▶ informal modelling and discussion of potential abstractions

Autoimmune EAE and Auxin Transport examples are similar

- ▶ EAE domain: Kumar, Torrey Pines Inst. for Molecular Studies, CA
- ▶ Auxin domain: Leyser, Biology, York

Domain modelling involves a significant investment in mutual understanding

- ▶ the trust-based relationship grows out of mutual respect and this learning process



Domain to Domain Model: examples 2

Domain modelling without close active collaboration ...

- ▶ Species-abundance ecology example works from existing simulations and published papers
 - ▶ Initial collaboration identified issues and objectives
 - ▶ Some opportunity to follow up discrepancies in existing code and papers
- ▶ The granuloma formation example tests an abstract concept from the biological research
 - ▶ Well-defined problem identified by domain experts
 - ▶ Appropriate sources identified
 - ▶ Results interpreted in limited context only

... more ad hoc decisions, and less reliable results

- ▶ harder to establish scientific credibility



Domain to Domain Model: examples 2

Domain modelling without close active collaboration ...

- ▶ Species-abundance ecology example works from existing simulations and published papers
 - ▶ Initial collaboration identified issues and objectives
 - ▶ Some opportunity to follow up discrepancies in existing code and papers
- ▶ The granuloma formation example tests an abstract concept from the biological research
 - ▶ Well-defined problem identified by domain experts
 - ▶ Appropriate sources identified
 - ▶ Results interpreted in limited context only

... more ad hoc decisions, and less reliable results

- ▶ harder to establish scientific credibility



Domain to Domain Model: examples 2

Domain modelling without close active collaboration ...

- ▶ Species-abundance ecology example works from existing simulations and published papers
 - ▶ Initial collaboration identified issues and objectives
 - ▶ Some opportunity to follow up discrepancies in existing code and papers
- ▶ The granuloma formation example tests an abstract concept from the biological research
 - ▶ Well-defined problem identified by domain experts
 - ▶ Appropriate sources identified
 - ▶ Results interpreted in limited context only

... more ad hoc decisions, and less reliable results

- ▶ harder to establish scientific credibility



“Simulation development is agile: it does not need models”

”Agile”:

- ▶ Collaboration is more important than modelling
- ▶ Iterative development and intrinsic verification are key
- ▶ Changing understanding is inherent

But ... modelling is used here:

- ▶ to develop and maintain mutual understanding
- ▶ to support interpretation of results
- ▶ to support the next phase of simulation
 - ▶ Simulations raise new scientific questions
 - ▶ Initial simulations become the basis for continuing scientific collaboration
 - ▶ e.g. Autoimmunity EAE example ...



“Simulation development is agile: it does not need models”

”Agile”:

- ▶ Collaboration is more important than modelling
- ▶ Iterative development and intrinsic verification are key
- ▶ Changing understanding is inherent

But ... modelling is used here:

- ▶ to develop and maintain mutual understanding
- ▶ to support interpretation of results
- ▶ to support the next phase of simulation
 - ▶ Simulations raise new scientific questions
 - ▶ Initial simulations become the basis for continuing scientific collaboration
 - ▶ e.g. Autoimmunity EAE example ...



Motivation for Autoimmune EAE domain modelling

courtesy of M Read, example

- ▶ Necessary pre-requisite to creation of a simulation.
- ▶ ... for Domain Expert to validate understanding
 - ▶ *before* jumping into code
 - ▶ ... but iterative learning process as well
- ▶ Diagrams used as a tool for guiding discussion
- ▶ Raise a lot of questions:
 - ▶ Detail needed for agent model: “How could we know that?”
 - ▶ Things you thought you understood, but actually don't
 - ▶ Identifying and addressing omissions
- ▶ Contract between modeller & Domain Expert - scope of work

More than anything else, a tool with which to explore understanding of the domain



Summary

- ▶ Agent-based simulation can be, and is being, used to support scientific research
 - ▶ Simulation can fill gaps in capabilities of scientific techniques
 - ▶ Successful simulation leads to new questions for science and simulation
 - ▶ There is a steep learning curve, the first time, for both stakeholders
- ▶ Collaboration needs clear responsibilities and mutual respect
- ▶ Existing engineering approaches help but often need adaptation
 - ▶ support for communication and collaboration is more important than engineering dogma

CoSMoS is looking at the process of simulation, modelling, and validation of simulations against science



Summary

- ▶ Agent-based simulation can be, and is being, used to support scientific research
 - ▶ Simulation can fill gaps in capabilities of scientific techniques
 - ▶ Successful simulation leads to new questions for science and simulation
 - ▶ There is a steep learning curve, the first time, for both stakeholders
- ▶ Collaboration needs clear responsibilities and mutual respect
- ▶ Existing engineering approaches help but often need adaptation
 - ▶ support for communication and collaboration is more important than engineering dogma

CoSMoS is looking at the process of simulation, modelling, and validation of simulations against science

