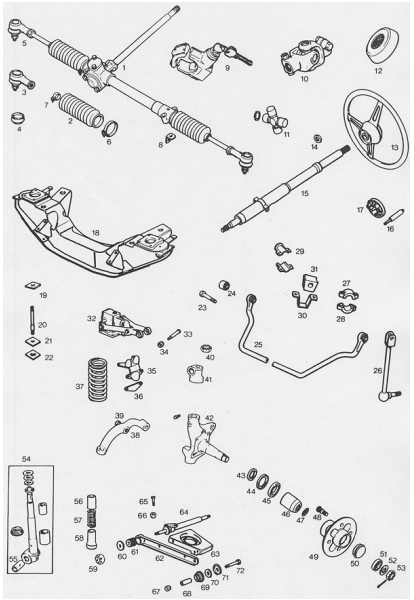
BioModel Engineering

David Gilbert

School of Information Systems, Computing & Mathematics
Centre for Systems & Synthetic Biology
Brunel University, London UK

Brunel UNIVERSITY LONDON

Bioinformatics, Systems Biology, Synthetic Biology

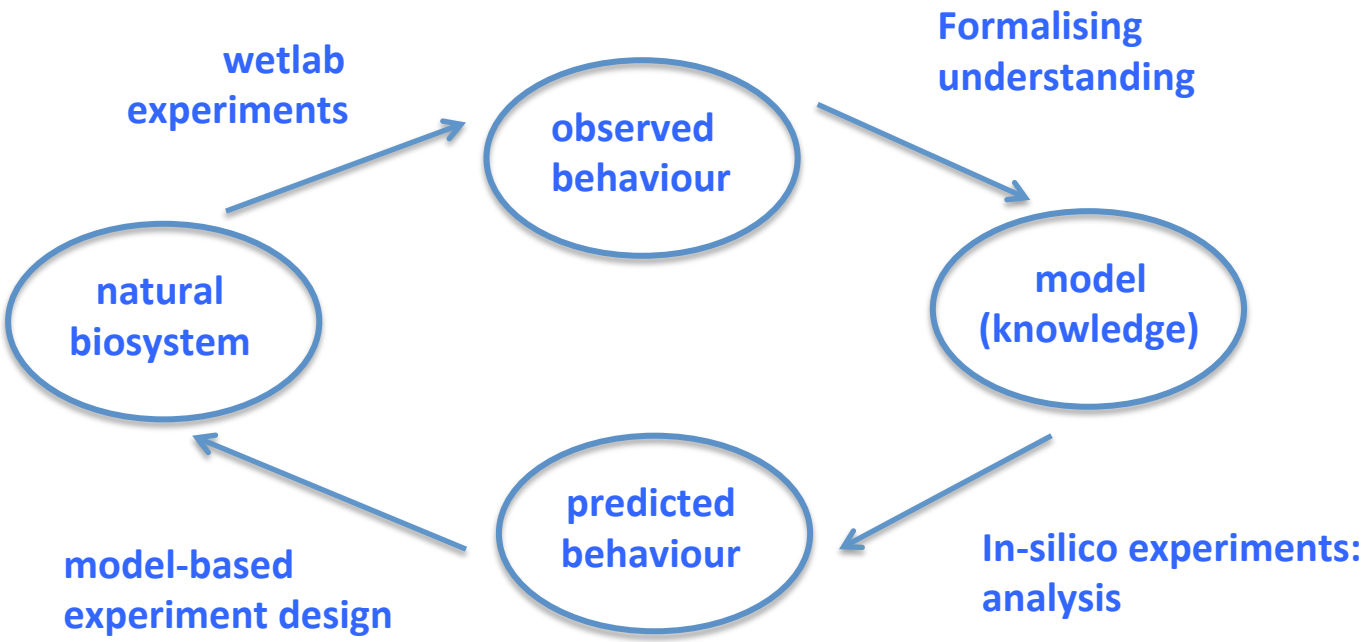


But how do these work together?

Build me a better one!

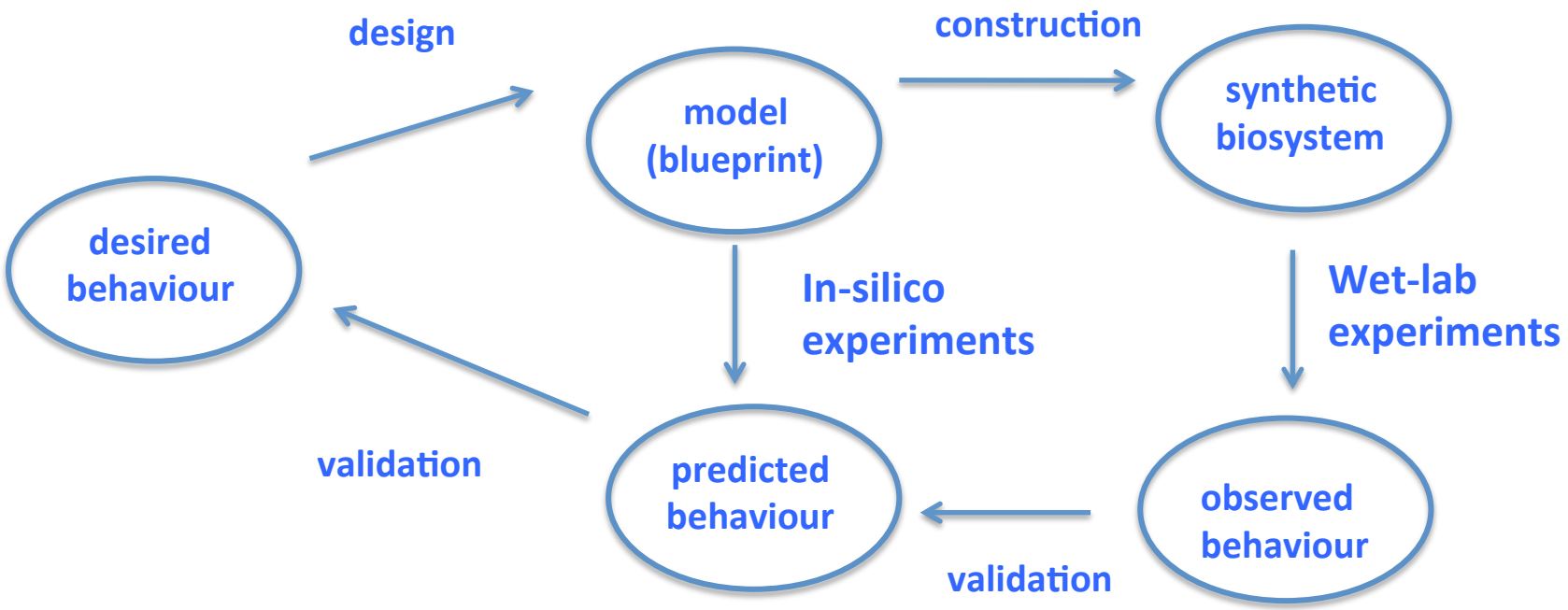
Brunel UNIVERSITY LONDON

Systems biology



Brunel UNIVERSITY LONDON

Synthetic Biology



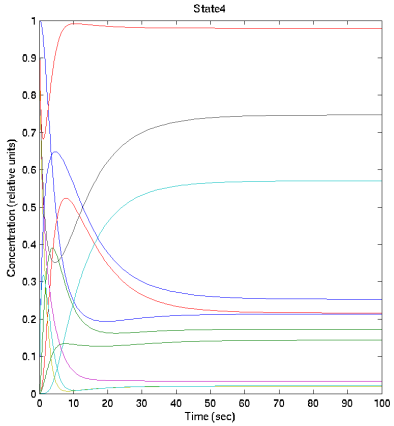
Brunel UNIVERSITY LONDON

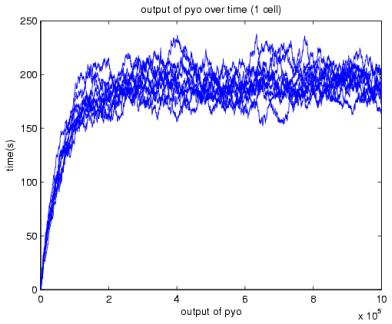
Validation & verification

- **Validation** – '***You built the right product?***'. (Quality control)
  - Product / system accomplishes its intended requirements.
  - But you didn't tell me you wanted a **red** bus!

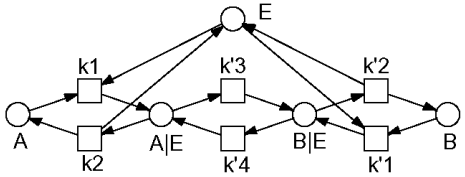
- **Verification** - '***You built the product right?***'. (Quality assurance)
  - System complies with its specification
- Possible for the product to produce the required outcome, but not in accord with its specification. (?)

Where to start?





$$\frac{d[A]}{dt} = -k_1 \times [A]$$

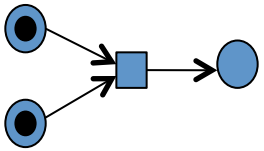


Brunel UNIVERSITY LONDON

We need a Framework!

Molecules/Levels
CTL, LTL

Qualitative PN



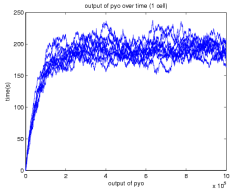
Time-free

Timed, Quantitative

Abstraction

Abstraction

Markov chain
Molecules/Levels
Stochastic rates
CSL



Stochastic PN

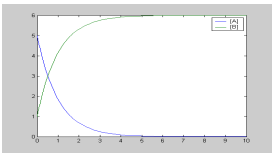
Approximation

Approximation

Continuous PN

$$\frac{d[A]}{dt} = -k_1 \times [A]$$

Map to ODES:
Concentrations
Deterministic rates
LTLc



DiscreteState Space

Continuous State Space

Gilbert, Heiner and Lehrack. ``A Unifying Framework for Modelling and Analysing Biochemical Pathways Using Petri Nets.'' Proc CMSB 2007

Brunel UNIVERSITY LONDON

BioModel Engineering

- Takes place at the interface of computing science, mathematics, engineering & biology.
- A systematic approach for **designing**, **constructing** and **analyzing** computational models of biological systems.
- Inspiration from efficient software engineering strategies.
- Not engineering biological systems *per se,* but
  - describes their structure and behaviour,
  - in particular at the level of intracellular molecular processes,
  - using computational tools and techniques in a principled way.

Biomodel engineering

1. Problem identification
2. Construction
3. Simulation
4. Analysis & interpretation
5. Management & development

Building a computational model

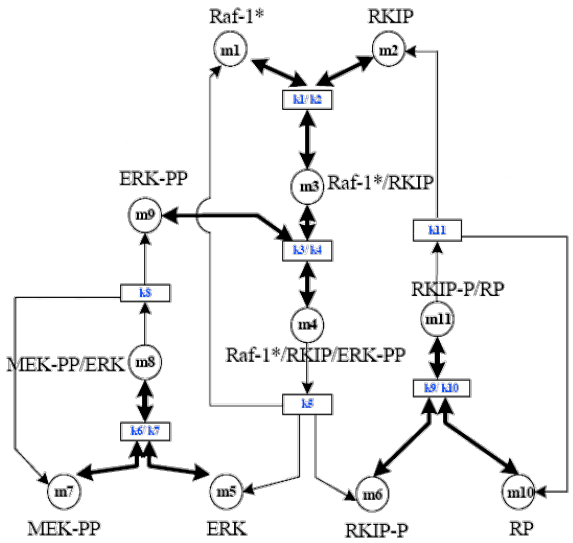
1: Identification

- **Identify the biological pathway** to model (what)
  - RKIP
  - EGF and NGF activated MAPK
- Or, more importantly, **identify the biological question** to answer (why)
  - What influence does the Raf Kinase Inhibitor Protein (RKIP) have on the Extracellular signal Regulated Kinase (ERK) signalling pathway?
  - How do EGF and NGF cause differing responses in ERK activation, transient and sustained, respectively?

(2. Construction)

What is a biochemical network model?

1. Structure



bipartite graph

QUALITATIVE Petri net

2. Kinetics

$$d[Raf1^*]/dt = k1*m1*m2 + k2*m3 + k5*m4$$

$$k1 = 0.53;\ k2 = 0.0072;\ k5 = 0.0315$$

reaction rates

QUANTITATIVE (PN)

3. **Initial conditions**

$$[Raf1^*]_{t=0} = 2\ \mu Molar$$

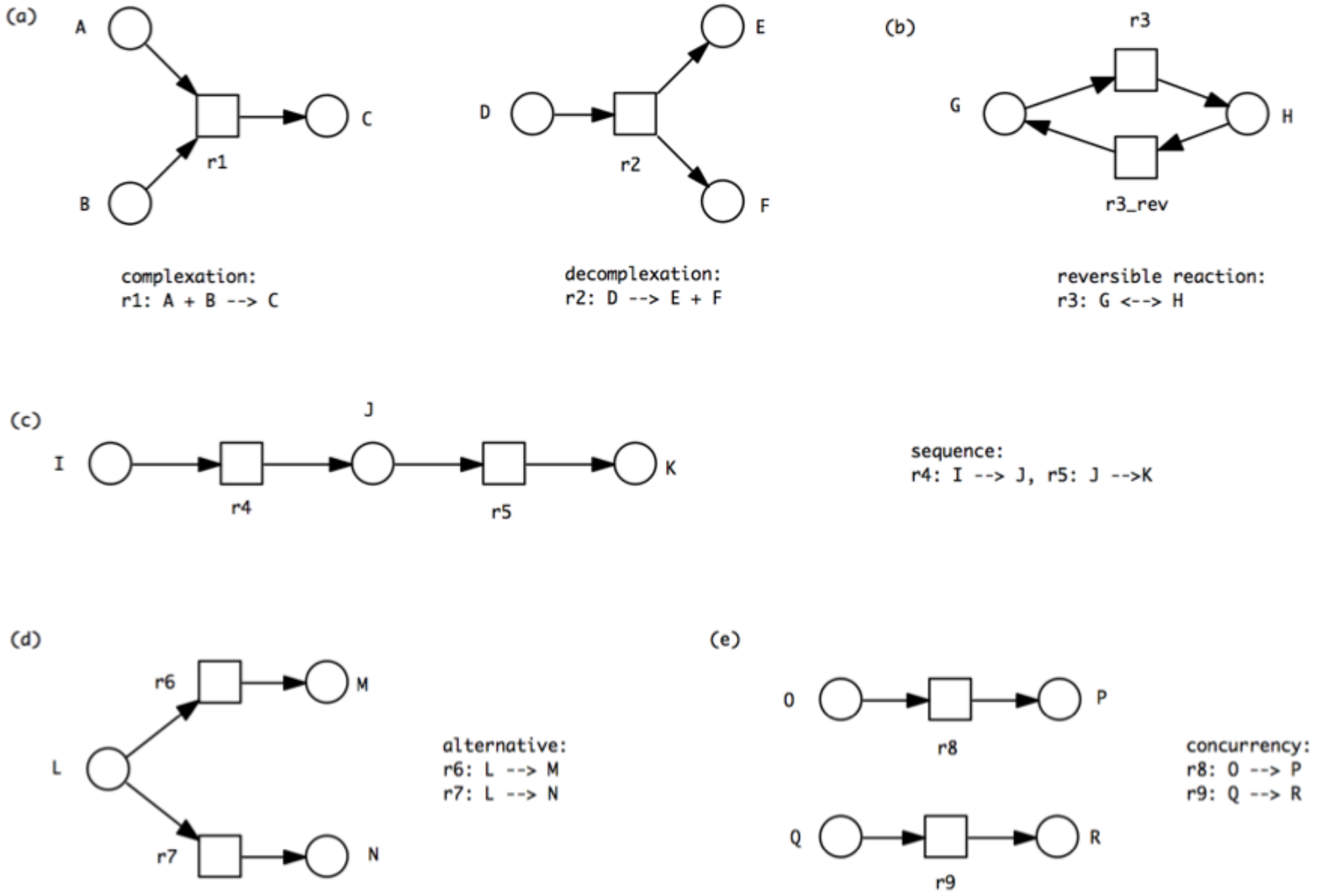
marking , concentrations

QUANTITATIVE (PN)

Brunel UNIVERSITY LONDON

Model Design Patterns

Some biochemical reaction patterns



Brunel UNIVERSITY LONDON

MA1: mass-action enzymatic reaction

A: substrate

B: product

E: enzyme

E|A complex

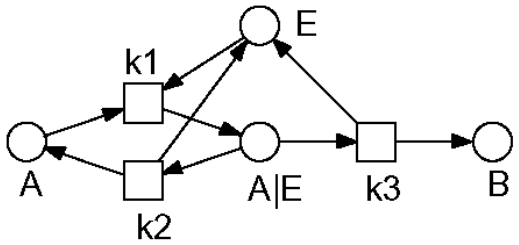
Biochemistry

$$A \xrightarrow{E} B$$

Biochemistry
Mass Action

$$A + E \underset{k_2}{\overset{k_1}{\rightleftarrows}} A \mid E \xrightarrow{k_3} B + E$$

Petri net



Ordinary Differential Equations

$$\frac{d[A]}{dt} = -k_1 \times [A] \times [E] + k_2 \times [A \mid E]$$

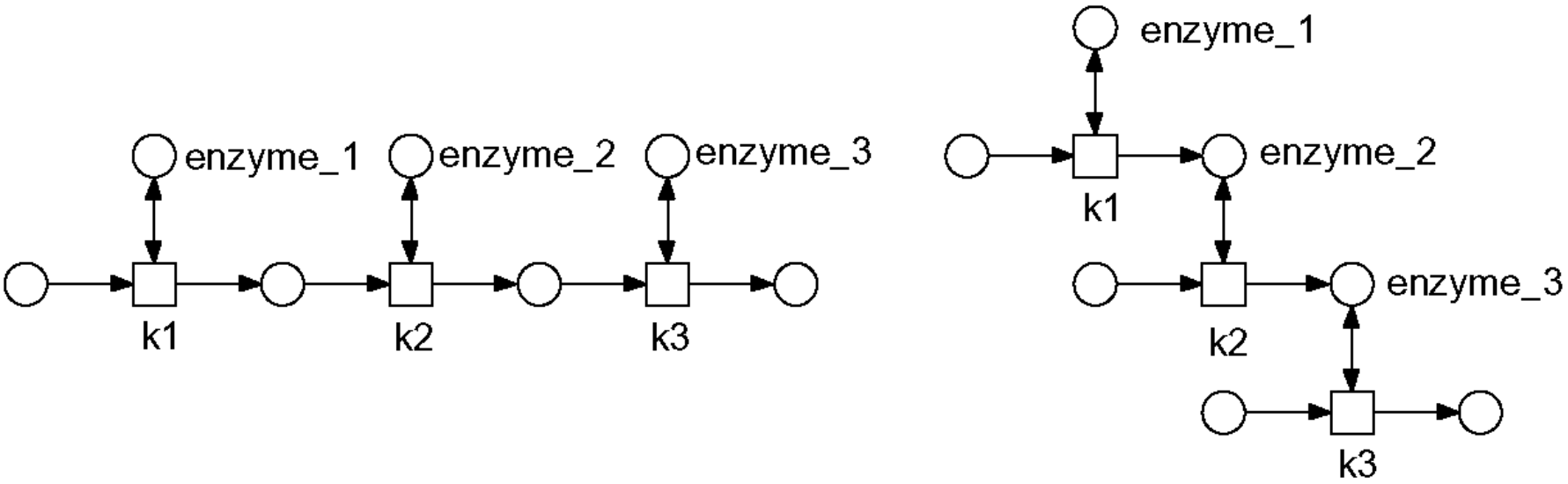
$$\frac{d[A \mid E]}{dt} = +k_1 \times [A] \times [E] - k_2 \times [A \mid E] - k_3 \times [A \mid E]$$

$$\frac{d[B]}{dt} = + k_3 \times [A \mid E]$$

$$\frac{d[E]}{dt} = -k_1 \times [A] \times [E] + k_2 \times [A \mid E] + k_3 \times [A \mid E]$$

Brunel UNIVERSITY LONDON

Metabolic pathways vs Signalling Pathways



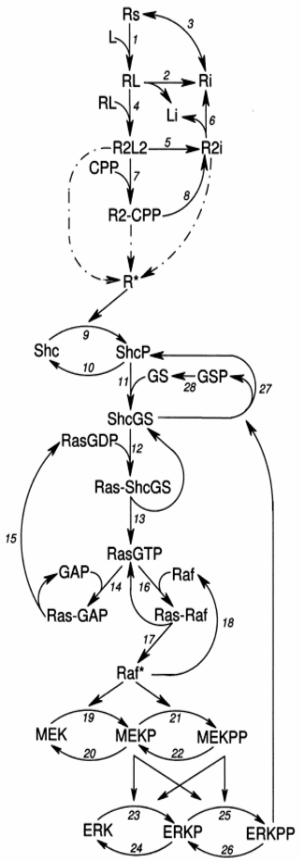
Automated construction…

- Automated
  - **Fitting** from target behaviour (topology, rates)
  - **Formal derivation** time series data (topology, afternoon session)

- Modular modelling (afternoon session)

A future vision

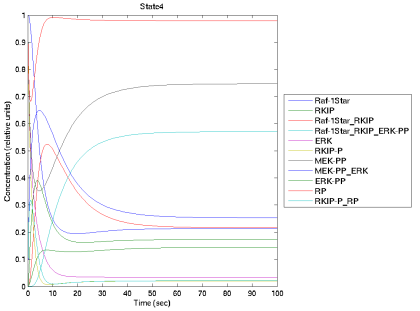
Model construction



observations

validate

Desired behaviour



construct

verify

Biological system

Component library

Standard biological parts

Biochemical Models
Construction Based on Reuse of Components

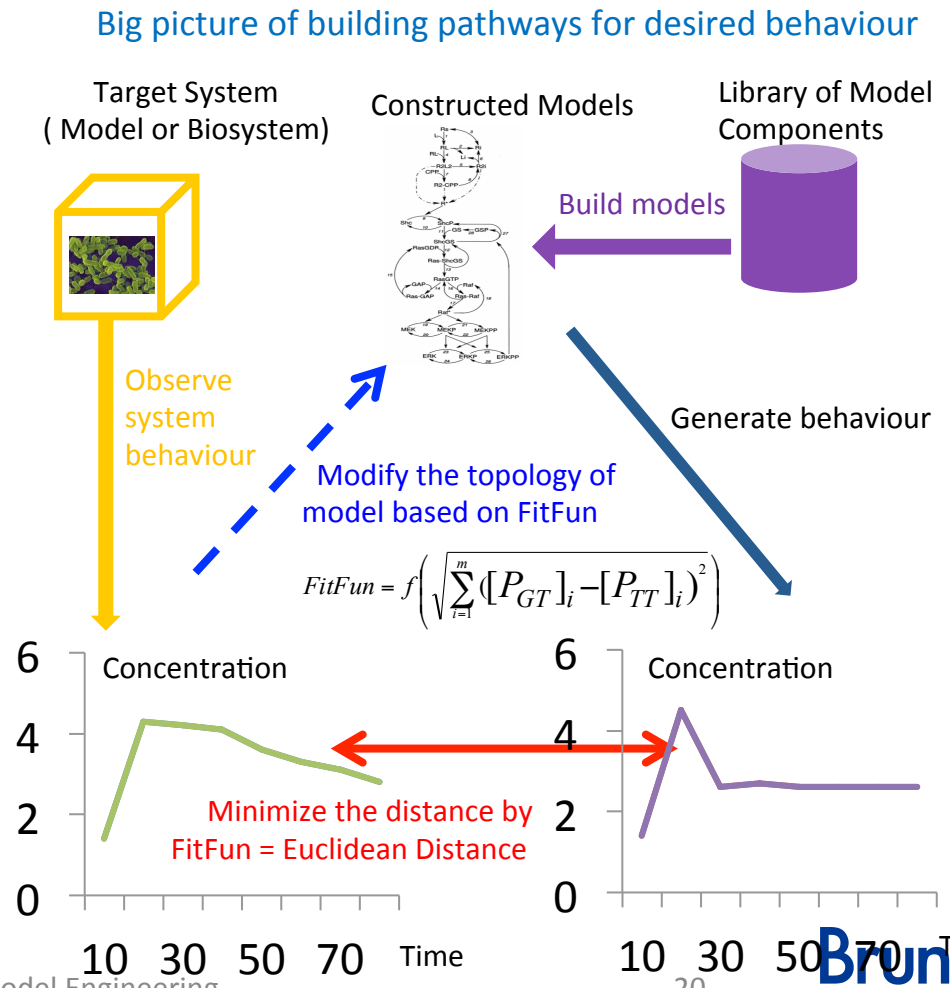
Zujian Wu

Aims:

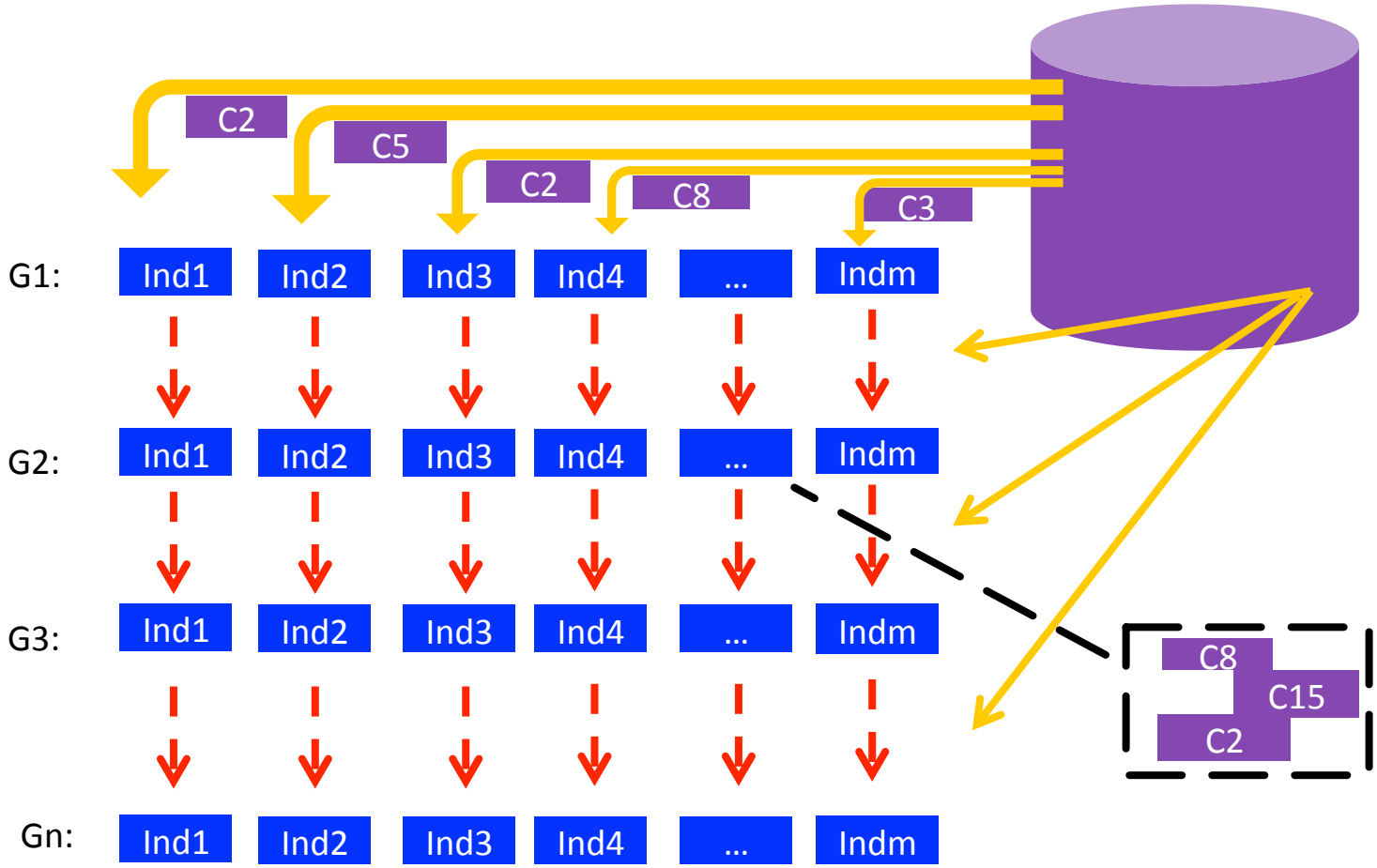
- To achieve the target-driven construction of biochemical models by reference to their desired behaviours
- To address the construction problem by
  - building a library for storing reliable biochemical functional submodels (as components)
  - intelligently selecting, combining and mutating these submodels to generate complex systems

Methods:

- Petri Nets: Components & Model
- Simulated Annealing: Kinetic rates
- Evolutionary Algorithms: Topology



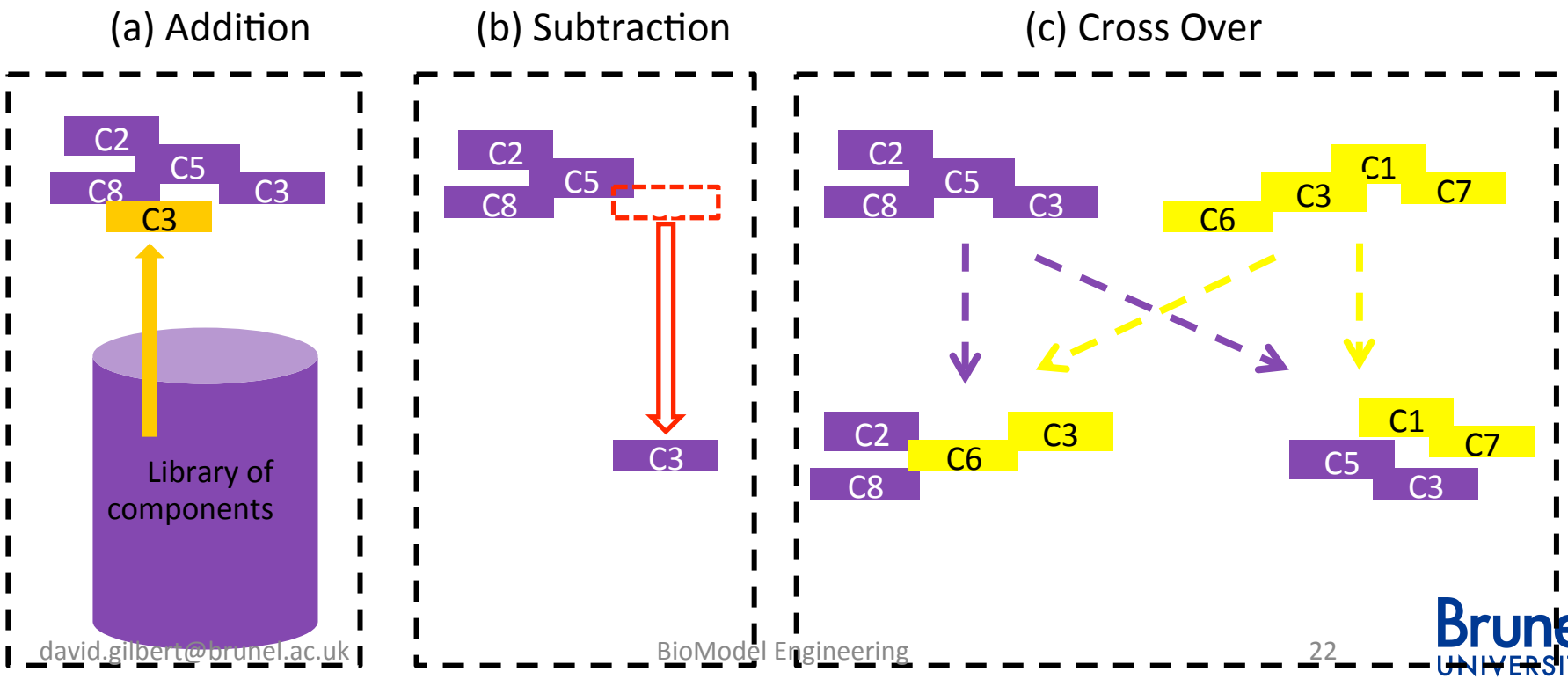
Evolution Strategy – Topology Optimization



Brunel UNIVERSITY LONDON

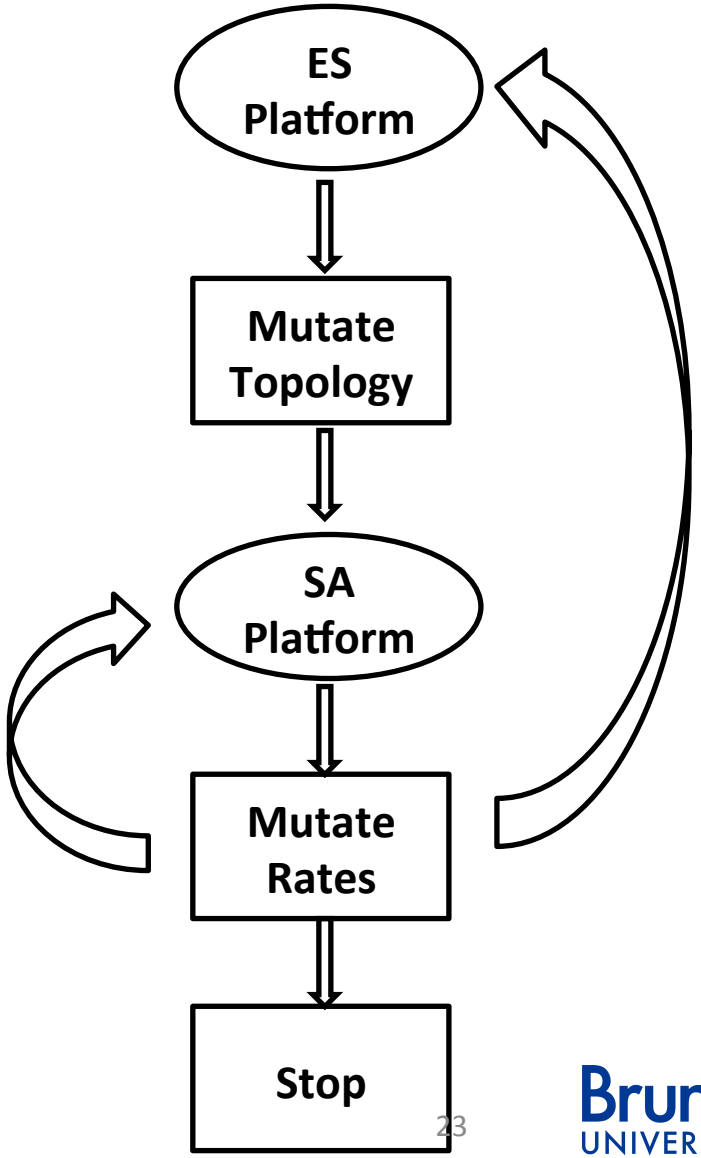
Genetic operators

- Addition – composing component to an existing network
- Subtraction – removing component from an existing network
- Cross Over – recombining two networks by 'cut and splice'



Hybrid Optimization of Topology and Kinetic Rates based on Evolution Strategy(ES) and Simulated Annealing(SA)

Algorithm: A hybrid piecewise modelling framework
Require: CompLib, Composition Rules
Ensure: BioNbest
1: Initiate the population;
2: **while** Not reached maximum generation (ES layer) **do**
3: **for** Each individual in the population **do**
4: Mutate the topology of individual by Addition or Subtraction;
5: Check the mutated topology of the individual;
6: Evaluate the mutated individual;
7: **if** The kinetic rates are required to be optimized **then**
8: **while** Not reach minimum temperature (SA layer) **do**
9: Optimize the kinetic rates by Gaussian distribution;
10: Evaluate the mutated kinetic rates;
11: **end while**
12: **end if**
13: **end for**
14: Crossover the individuals;
15: Select offspring for next generation;
16: **end while**
17: Return BioNbest

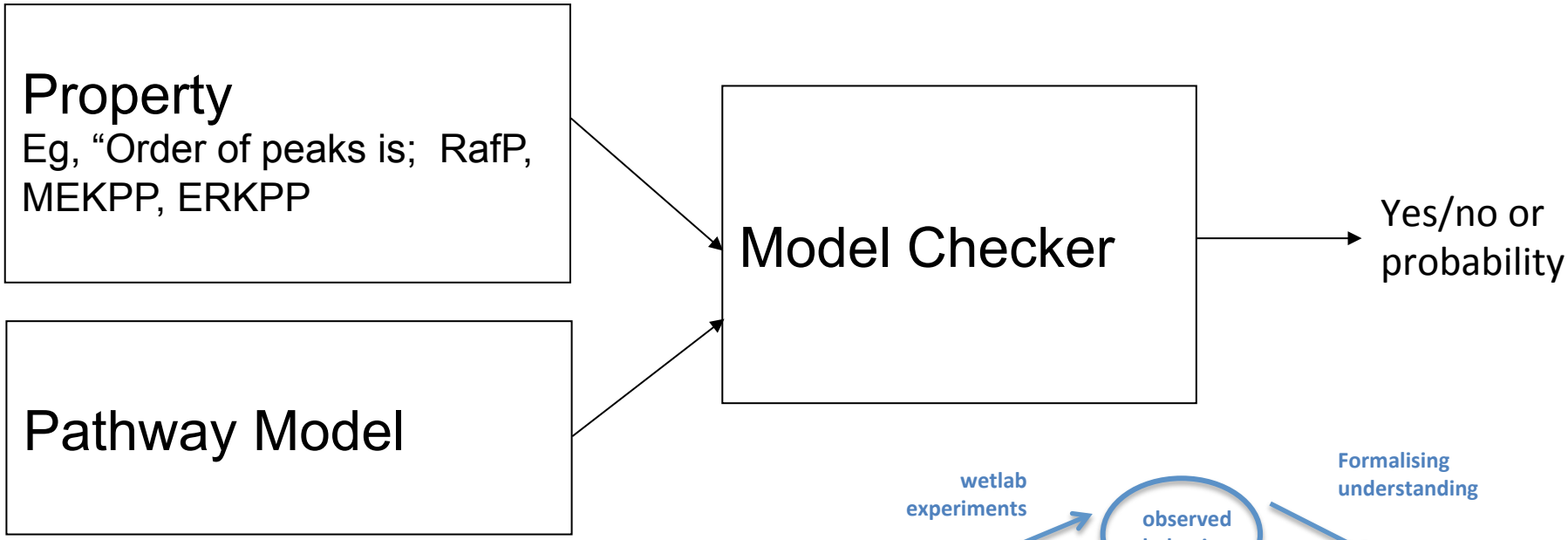


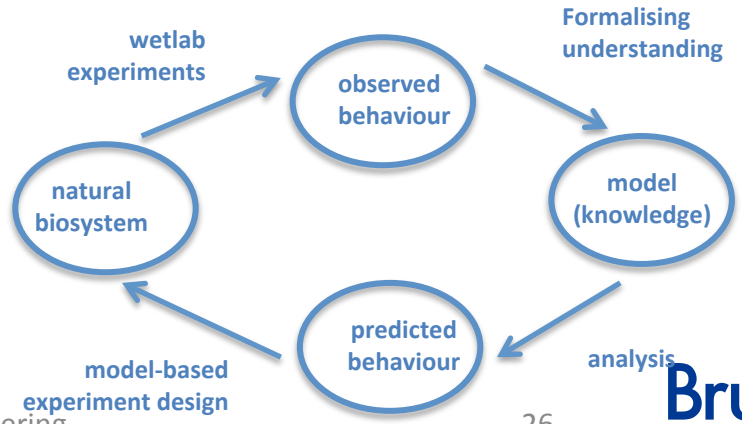
Construction & simulation

Snoopy!

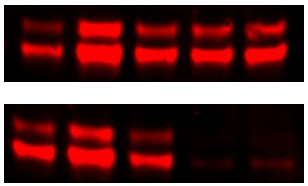
Analysis & interpretation

Model Checking Biochemical Pathways





Simulation-based Model Checking Biochemical Pathways



Property

Eg, “Order of peaks is RafP, MEKPP, ERKPP”

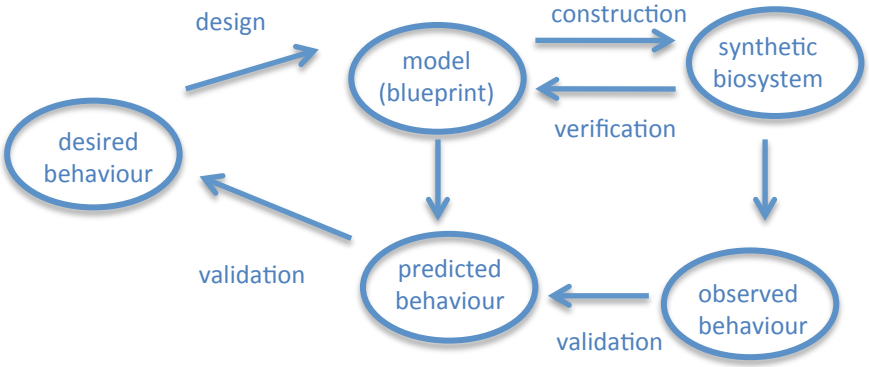
Time series data

Model

Lab

Behaviour Checker

Yes/no or probability



Brunel UNIVERSITY LONDON

Managing models

- Version control
  - **Improving the reuse of computational models through version control,** Dagmar Waltemath et al, J. Bioinformatics 2013
- Database of models
  - Storage, searching, retrieval
  - Biomodels.org

Model searching

Peaks at least once

(rises then falls below 50% max concentration)

$P_{>=1}[\ ErkPP <= 0.50*max(ErkPP) \wedge d(ErkPP) > 0\ U\ (\ ErkPP = max(ErkPP) \wedge F(\ ErkPP <= 0.50*max(ErkPP)\ )\ )\ ]$

- ***Brown***
- ***Kholodenko***
- ***Schoeberl***

Rises and remains constant

(99% max concentration)

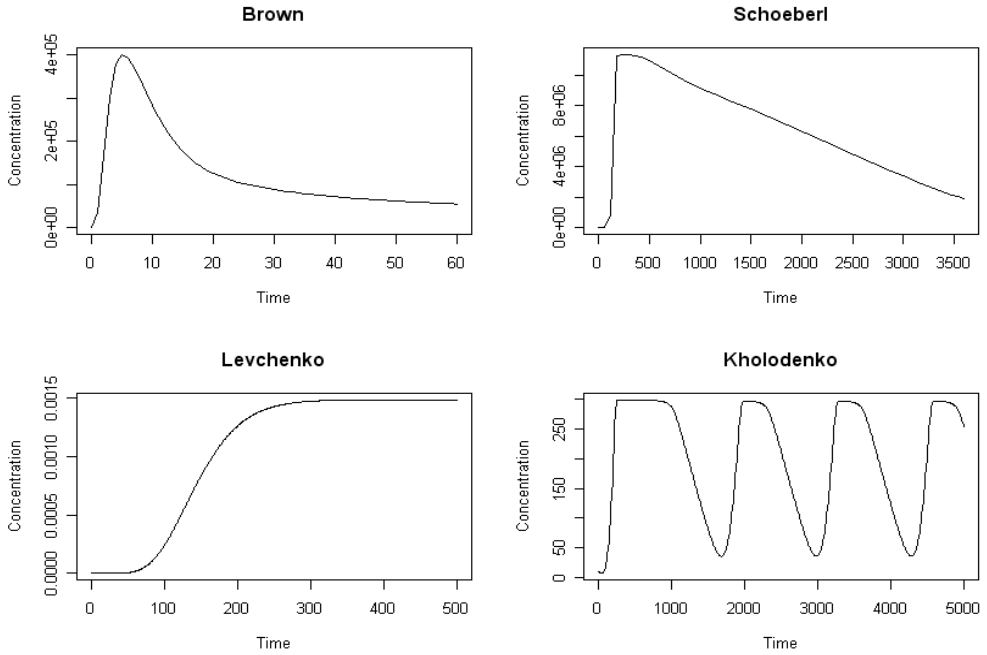
$P_{>=1}[ErkPP <= 0.50*max(ErkPP) \wedge (\ d(ErkPP) > 0\ )\ U\ (\ G(ErkPP >= 0.99*max(ErkPP))\ )\ ]$

- ***Levchenko***

Oscillates at least 4 times

$P_{>=1}[\ F(\ d(ErkPP) > 0 \wedge F(\ d(ErkPP) < 0 \wedge \ldots\ )\ )\ ]$

- **Kholodenko**



Brunel UNIVERSITY LONDON