Systems Biology

An introduction to dynamic pathway modelling

Olaf Wolkenhauer

www.sbi.uni-rostock.de

Systems Biology and Bioinformatics in Rostock

Current Research Projects:

- Systems Biology of the AMP-activiated kinase pathway (AMPKIN), '06-'09. EU FP6 STREP.
- Computational Systems Biology of Cell Signalling (COSBICS), '05-'08. EU FP6 STREP.
- Modelling and Simulation in Cell Signalling (MOSAICS), '05-'08. BMBF NGFN II.
- Microarray Research, Innovation, and Exploitation (MARIE), '02-'07. UK BBSRC.
- Exploring Coordination and Control in Cell Signalling (ECCCS). '04-'07. EU & M-V.
- Transcriptomics and Network Modelling for *M.bovis.* '03-'06. UK DEFRA.
- Systems Biology for Medical Applications (SysBioMed). '06-'08. EU FP6 SSA.

Motivation:

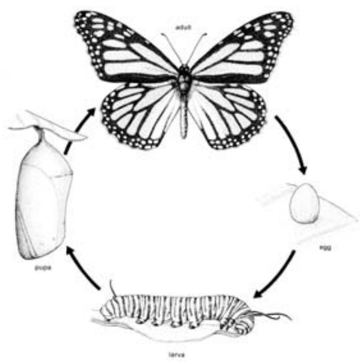
How does the cell realise its functions?

Regulation, control and coordination of cell function.
The cell is alive – why is that?
How do we deal best with biological complexity?

Personal Motivation

Total Metamorphosis: one genome, two proteomes...

... the genome provides only the notes: the music is played elsewhere!



Picture from P.J.Gullan and P.S.Cranston: *The Insects*, Blackwells Science, 2000

The Origins

Mike Mesarović in *System Theory and Biology*, 1968:

"In spite of the considerable interest and efforts, the application of systems theory in biology has not quite lived up to expectations."

"One of the main reasons for the existing lag is that systems theory has not been directly concerned with some of the problems of vital importance in biology."

"The real advance in the application of systems theory to biology will come about only when the biologists start asking questions which are based on the system-theoretic concepts rather than using these concepts to represent in still another way the phenomena which are already explained in terms of biophysical or biochemical principles."

"Then we will not have the 'application of engineering principles to biological problems' but rather a field of systems biology with its own identity and in its own right."

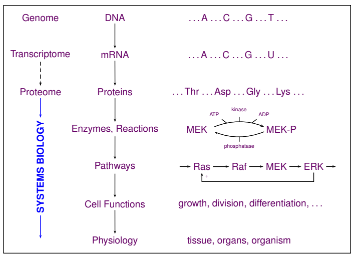
Outline

Systems biology integrates experimental and modelling approaches to explain the functional organisation of cell-biological systems
It aims at quantitative experimental results and building predictive models.

Why Systems Biology:
- Mining vs. Systems Approach
- Structural vs. Funct. Organization

Why Mathematical Modelling:
- Managing uncertainty
- Modelling causal entailment
- *Example: Mutual-activation switch*

Dynamic Principles and Motifs:
- The role of feedback; delays.
- *Example: Ras/Raf/MEK/ERK pathway.*



What is Systems Biology?

... *not* Genomics and Bioinformatics:

- ✘ Identification & molecular characterisation of the components that make up a cell.
- ✘ Storage of data, information, and mining of those resources.

Systems biology is a *merger* of cell biology with (dynamic) systems theory.

Key Questions:

- ✘ *How do the components within a cell interact, so as to bring about the cell's **structure** and realise its **functioning**?* (intra-cellular dynamics)
- ✘ *How do cells interact, so as to develop and maintain higher levels of **structural** and **functional organisation**?* (inter-cellular dynamics)

Key Concept:

Pathways as *networks* of biochemical reactions

Key Dogma:

(Feedback) dynamics determine cell function

Statistical Significance Tests

Testing for Differences

- ▷ ... between sample and expected value: one-sample t-test.
- ▷ ... between two samples from the same population: paired t-test.
- ▷ ... between two samples from two populations: two-sample t-test.

Two-Sample t-Test:

Step 1: Null-hypothesis: Mean of the differences *is not* different from zero.

Step 2: The test statistic: $t = \frac{\text{mean difference}}{\text{standard error of difference}} = \frac{\bar{X}-\bar{Y}}{\overline{\text{SE}}_d}$.
Assuming variance of both populations is the same

$$\overline{\text{SE}}_d = \sqrt{\left(\overline{\text{SE}}_X\right)^2 + \left(\overline{\text{SE}}_Y\right)^2}$$

Step 3: Calculate the significance probability P

Step 4: Decision:

- ✘ If $P < 0.05$, reject the null hypothesis.
- ✔ If $P \geq 0.05$, there is no evidence to reject the null hypothesis.

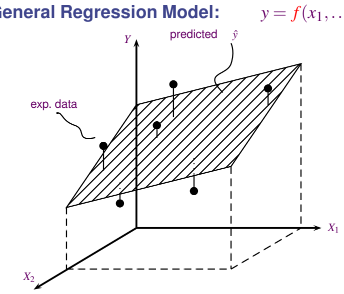
Step 5: 95% CI(difference) $= \bar{X} - \bar{Y} \pm \left(t_{(n_X+n_Y-2)}(5\%) \times \overline{\text{SE}}_d\right)$

Correlation and Regression

Co-Variation: (association, coincidence) $r_{x,y} = \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2 \sum_i y_i^2}}$

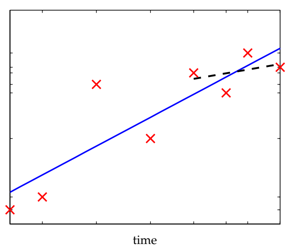
Correlations do not imply *causal entailment*: we cannot predict y from x!

General Regression Model: $y = f(x_1, \ldots, x_m)$ "y is a function of..."



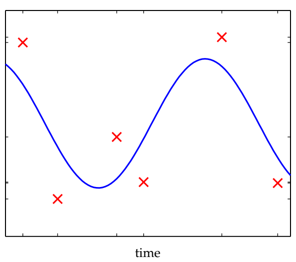
Linear model:

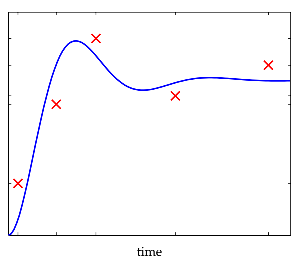
$\hat{y} = \theta_0 + \theta_{y_1} x_1 + \cdots + \theta_{ym} x_m$



Trend Model: $\hat{y}(t) = \theta_0 + \theta_1 t$

Time Series Analysis





General Linear Regression Model:

$$\hat{y} = \theta_0 + \theta_{y_1} x_1 + \theta_{y2} x_2 + \cdots + \theta_{ym} x_m + w$$

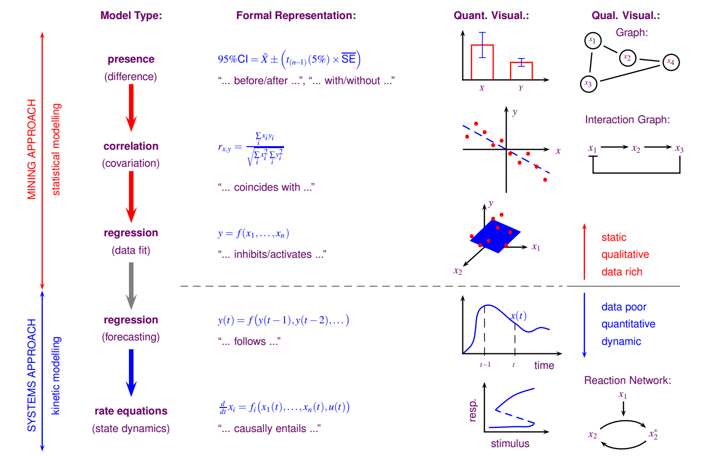
Trend Model:

$$\hat{y}(t) = \theta_0 + \theta_1 t$$

Autoregressive Forecasting Model:

$$\hat{y}(t+1) = \theta_0 + \theta_1 y(t) + \theta_2 y(t-1) + \cdots + \theta_k y(t-k+1)$$

Preview



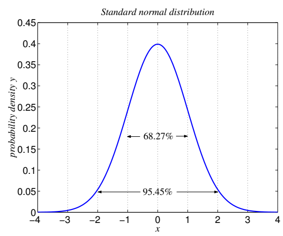
Why Mathematical Modelling?

Statistical Modelling: Testing Differences

- ▷ Wild-type, mutant, knock-out/-in, sRNAi data
- ▷ Comparing conditions, expression levels etc.
- ▷ Testing the presence of a component
- ▷ Using 2D gels, immunoblots, microarrays etc.

Uncertainty: Nonbiological Variability

- ▷ n repeated experiments: sample X
- ▷ Abstraction: *random variable*
- ▷ Probability density: mean μ, variance σ^2
- ▷ **(Abstract) stochastic model:**



$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

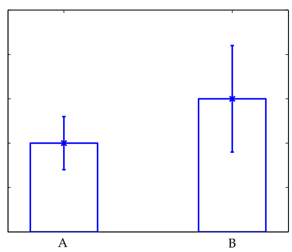
Statistical Modelling

Testing Differences in two samples A and B:

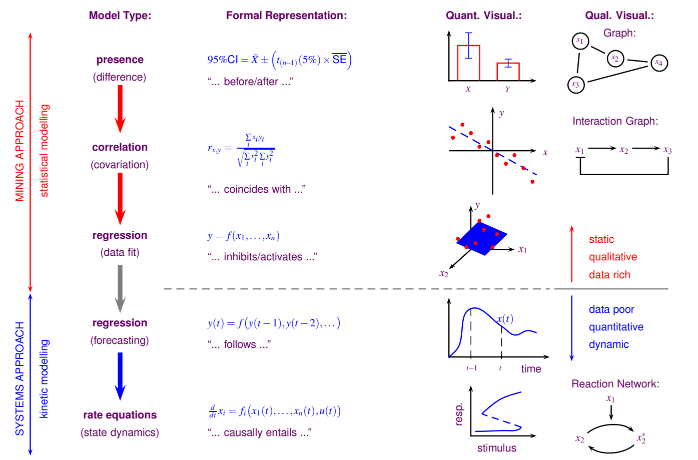
- ▷ Sample Mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$ Sample Variance $\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})^2$

How do we know the difference is significant?

- ▷ The sample mean is an estimate: *random variable* following a t-distribution
- ▷ Variability of sample means: *standard error:* $\overline{\text{SE}} = \frac{s}{\sqrt{n}}$, where $s = \sqrt{\frac{n}{n-1}\sigma_n^2}$
- ▷ *Confidence Interval:* 95% CI$(mean) = \bar{X} \pm \left(t_{(n-1)}(5\%) \times \overline{\text{SE}}\right)$



Classification of Models



Mathematical Modelling

Interpretations of the Regression Model:

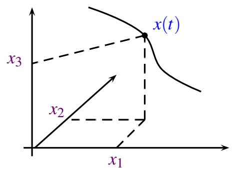
(i) A representation or summary of a dataset (data fitting).
(ii) To predict the value of y from $x_1, \ldots, x_m$ (data fitting).
(iii) To predict the value of y as a *causal* consequence of *changes* in x_j (dynamic modelling).

... which 'principle' (mechanism) generated the data?!

State-Space Model: ... rate equations, nonlinear ordinary differential equations

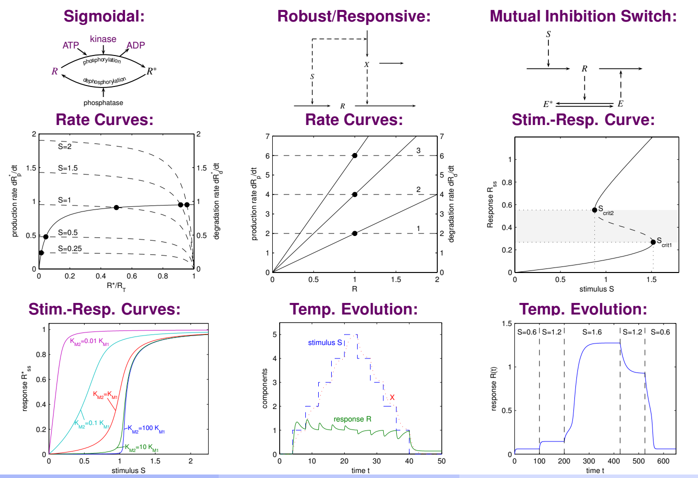
$$\underbrace{\frac{d}{dt}x_i(t)}_{\text{rate of change}} = f_i\big(\underbrace{x_1(t), \ldots, x_n(t)}_{\text{state } x(t)};\ \underbrace{u(t)}_{\text{stimulus}}\big)$$

$$\underbrace{y(t)}_{\text{response}} = h\big(x_1(t), \ldots, x_n(t)\big)$$



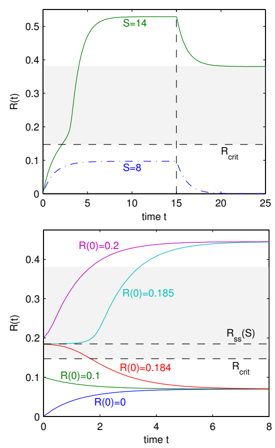
Causation is the principle of explanation of change!
... it is a relationship, not between things, but between changes of states of things.

Dynamic Motifs

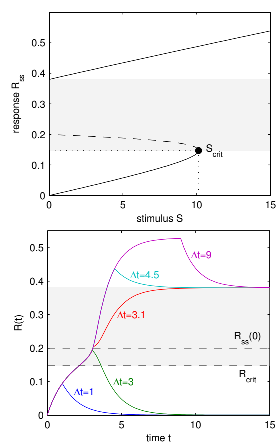


Dynamic Pathway Modelling

Step Response:

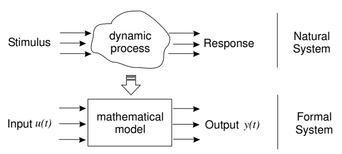


Stimulus Response Curve:

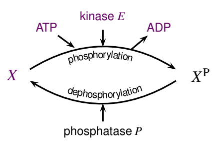


Mathematical Models as Mediators

- ✘ *A model is an abstraction.*
- ✘ *Abstraction implies a reduction of complexity.*
- ✘ *To abstract is to model.*
- ✘ *The art is to make appropriate assumptions.*



(De-)Activation Cycle:



$$X + E_1 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C_1 \overset{k_2}{\longrightarrow} E_1 + X^P$$

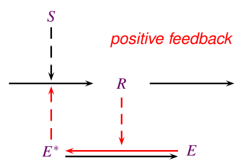
$$X^P + E_2 \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} C_2 \overset{k_4}{\longrightarrow} E_2 + X$$

... assuming ATP, temperature, volume and water-balance are constant.
... assume conservation $X^T = X(t) + X^P(t) + C_1(t) + C_2(t)$ to hold.
... assume quasi-steady state for C_1 and C_2.
... assume $X + X^P \gg C_1 + C_2$
... implying a relatively small kinase concentration...

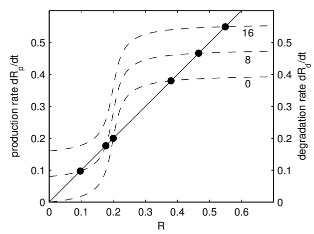
Dynamic Pathway Modelling

A cell is built up of molecules, as a house is with stones.
But a soup of molecules is no more a cell
than a heap of stones is a house.

Pathway Map:



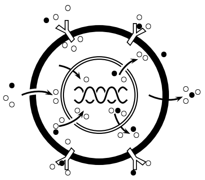
Rate Curves:

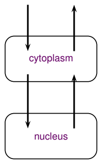


Rate Equation:

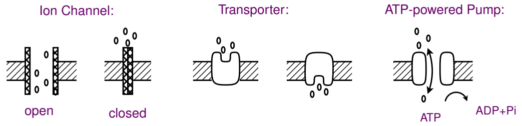
$$\frac{d}{dt}R(t) = \underbrace{k_0 E^*(R) + k_1 S(t)}_{\text{production}} \underbrace{- k_2 R(t)}_{\text{degradation}} \qquad \text{where} \qquad E^*(R) = \text{GK-function}$$

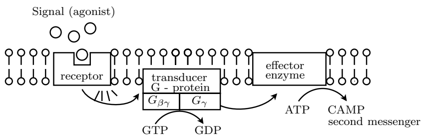
Structural Organisation of the Cell





Membranes, Organelles, ExtraCellular Matrix
DNA, RNA, Proteins, Enzymes, Metabolites
Pores, Channels, Transporter, Pumps
Ligands, Receptors, Messengers





Functional Organisation of the Cell

Signals encoding *extra-cellular* balances and information:

- ▷ physical ECM-to-cell and chemical cell-cell contacts
  ... related to cell adhesion, tissue formation, and development.
- ▷ transmission of information to regulate gene expression.

Signals encoding *intra-cellular* balances and information:

- ▷ transmission of information to regulate cell growth and cell cycle.
- ▷ homeostasis of pH level, temperature and water imbalances.

Dynamic Principles:

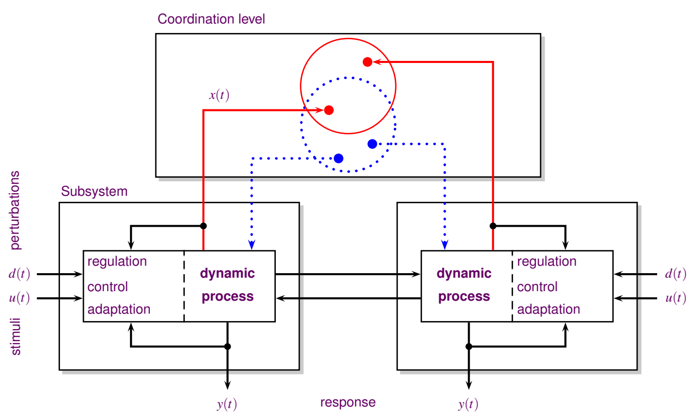
Regulation: the system maintains the level of a variable against external influences;
... the system is **robust** against perturbations.

Control: the system tracks/follows an input or reference signal;
... the system is **responsive** to stimuli.

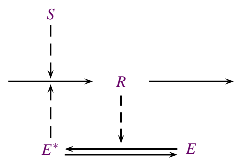
Adaptation: the system modifies itself in consequence of changes in environmental conditions.

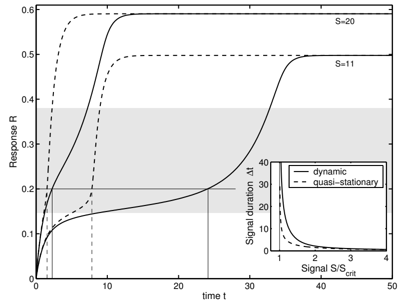
Coordination: harmonization of the system's behavior at different (structural and functional) **levels**.

Coordination: Cross-Talk among Pathways



Mathematical Models as Mediators





Quasi-Stationary Activation Model:

$$\frac{d}{dt}R(t) = k_0 E_P(R) + k_1 S(t) - k_2 R(t)$$ where $E_P(R)$ is the Goldbeter-Koshland function.

Dynamic Activation Model:

$$\frac{d}{dt}E_p(t) = \frac{k_3\,(1-E_p)R(t)}{K_{M3}+1-E_p(t)} - \frac{k_4\,E_p(t)}{K_{M4}+E_p(t)}$$

Mathematical Models as Mediators

Goldbeter-Koshland Ultrasensitive Switch:

$$E_p(R) = G(k_3R, k_4, K_{M3}, K_{M4}) = G(V_1, V_2, K_1, K_2)$$
$$= \frac{2V_1K_2}{V_2 - V_1 + V_2K_1 + V_1K_2 + \sqrt{(V_1 - V_2 + V_2K_1 + V_1K_1)^2 - 4(V_2 - V_1)V_1K_2}}$$

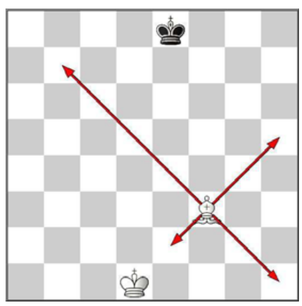
Choosing an Appropriate Modelling Framework:

Systems
Static — **Dynamic**
Time-varying — **Time-invariant**
Linear — **Nonlinear**
Discrete state — **Continuous state**
Event driven — **Time driven**
Stochastic — **Deterministic**
Discrete-time — **Continuous-time**

... idealizations force us into *reduced models*.
... assumptions force us into *phenomenological models*.

Functional vs. Structural Organization

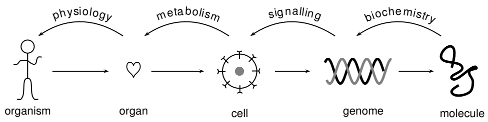
The shape determines only direct *interactions*...



... but not *function*.

Systems Biology: Will it work?

What we get funding for...



... uumpf!

- **Summary and Conclusions**
  - Modelling is the art of making appropriate assumptions.
  - Hypothesis are nets: only he who casts, will catch.
  - There is nothing more practical than a good theory.
- **Cellular Weather Forecasting:**
  - High-dimensional, nonlinear spatio-temporal dynamics.
  - Feedback mechanisms and delays.
  - Quantitative, sufficiently rich time series.

If something seems impossible, you increase your chances of success by trying it.

Scientific Pornography?

Ronald Plasterk, *Current Biology*, Vol.15, No.21, 2005:

"The fundamental misconception of systems biology advocates is that one could create a 'virtual cell', and use big computers to model life and make discoveries."

... I agree.

"None of these modellers ever predicted that small microRNAs would play a role. One makes discoveries by watching, working, checking. They want to be Darwin, but do not want to waste years on the Beagle. They want sex but not love, icing but no cake. Scientific pornography."

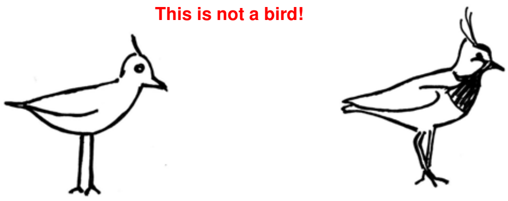
... *characterisation* of components is not the purpose of kinetic modelling!!

... but you need systems biology to understand the *role* of microRNA.

Plasterk's System Biologist

Mathematical Models as Mediators

This is not a bird!



Mathematics is an art that makes us realise reality.
Art is a lie that makes us realise truth (P.Picasso)

Why Modelling?

- *Managing uncertainty and complexity.*
- *Modelling implies abstraction and hence a reduction of complexity.*
- *It provides a language to generate and validate hypotheses.*