

Evolutionary Algorithm Optimisation

Fomekong-Nanfack *et al.* (2007). *Bioinformatics* **23**, 3356–3363.

An alternative optimisation strategy is to use an **evolutionary algorithm**:

- Treat a parameter set as the “genome” of an individual.
- Each individual has a “fitness” determined by a combination of the cost function and a penalty (to account for the ‘feasibility’ of the parameters).
- At each generation, rank individuals on fitness and select the fittest to seed the next generation (selection).
- ‘Mutation’ and ‘recombination’ (parameter changes) allow parameter space to be searched.

7

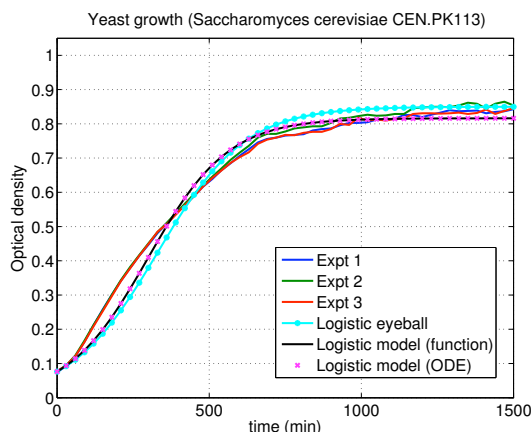
Local Optimisation

1. Optimise locally to speed up convergence.
2. One option is to move using steepest descent of the cost function, but requires evaluation of the derivative. No analytical expression and costly to approximate.
3. Use downhill simplex (Nelder-Mead). Evaluate the cost function at $n+1$ points (for an n -dimensional parameter space). Treat each point as a vertex of a simplex. Move the worst point to search for local minima (with progressively smaller moves).
4. Improves goodness of fit and speed of convergence.

8

Logistic growth? $\frac{dx}{dt} = rx\left(1 - \frac{x}{K}\right)$

- Consider experimental data on the growth of yeast.
- Dynamics look a bit like logistic growth ...



- But logistic growth is too slow at first, and too fast later.

9

A harder problem

- **Challenge:** Infer values of a and b for which the following model (a Bliss-Painter-Marr negative feedback model) best accounts for the data
- x_1 represents the mRNA in question, x_2 and x_3 are proteins (unmeasured in the experiment)
- Know from other data that $k = 1$ and $100 \leq a \leq 300$ and $0.05 \leq b \leq 0.3$

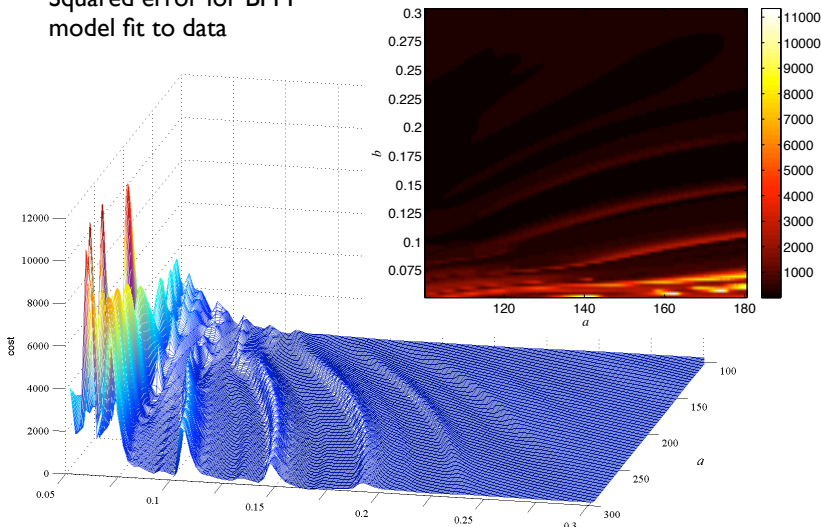
$$\begin{aligned} \frac{dx_1}{dt} &= \frac{a}{1+x_3} - bx_1 \\ \frac{dx_2}{dt} &= b(x_1 - x_2) \\ \frac{dx_3}{dt} &= bx_2 - c \frac{x_3}{1+kx_3} \end{aligned} \quad c = \begin{cases} 5, & t < 0 \\ 5 + 0.2t, & 0 \leq t < 50 \\ 15, & t \geq 50 \end{cases}$$

- **Method:** Solve the equations for different values of a and b and evaluate the squared error cost function specified earlier
- Use a search algorithm to search the allowed parameter values to identify the optimal values (lowest cost function value)

13

Fitness Landscape for BPM Model

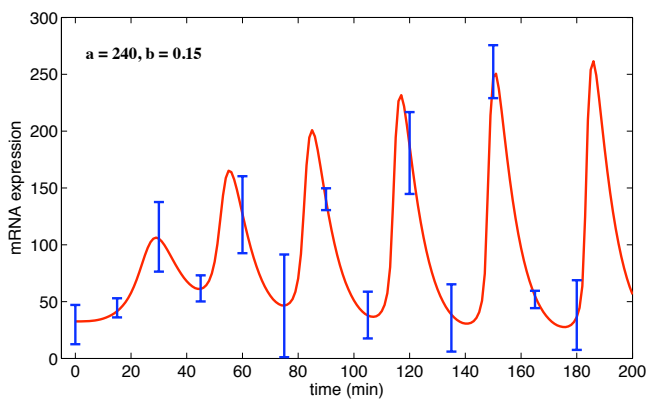
Squared error for BPM model fit to data



14

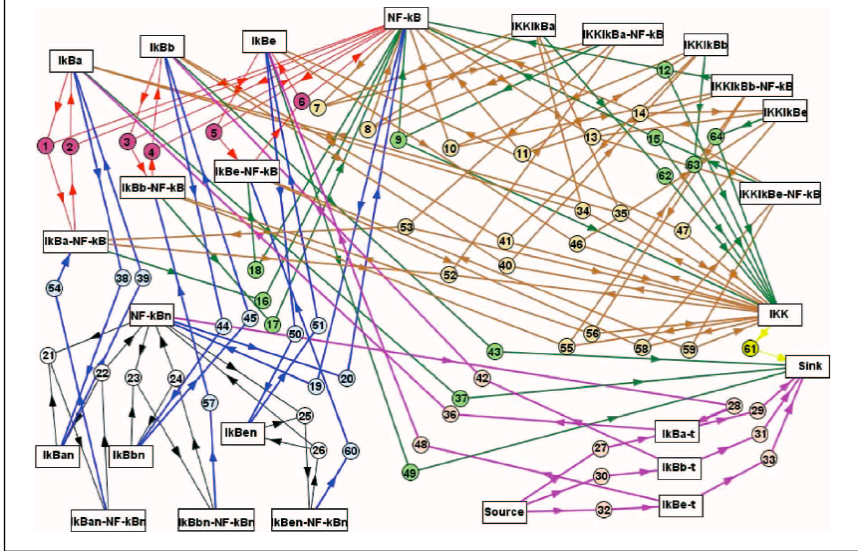
Best fit solution from Nelder-Mead search

cost = 0.2898



15

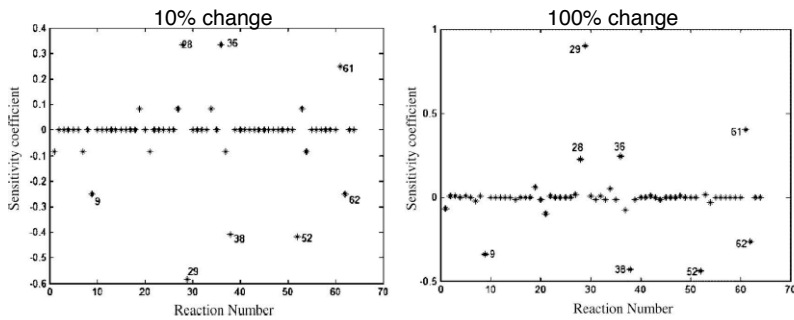
NF-κB Model Network Diagram



19

Parameter Sensitivity (T3)

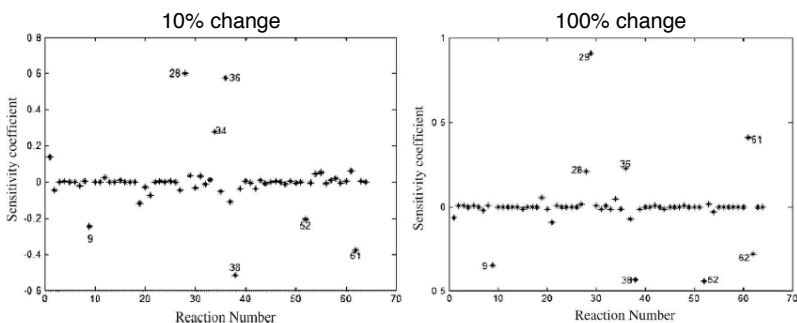
- Construct ODE model representing the interaction network (26 variables, 64 parameters)
- Find parameters that reproduce observed oscillations
- Assess parameter sensitivity of features: $S_p^\phi = \frac{\delta\phi/\phi}{\delta p/p}$



20

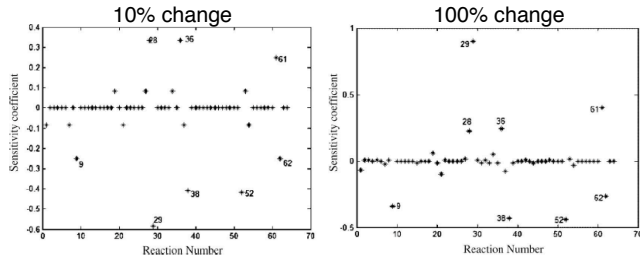
Parameter Sensitivity (A3)

- Construct ODE model representing the interaction network (26 variables, 64 parameters)
- Find parameters that reproduce observed oscillations
- Assess parameter sensitivity of features: $S_p^\phi = \frac{\delta\phi/\phi}{\delta p/p}$



21

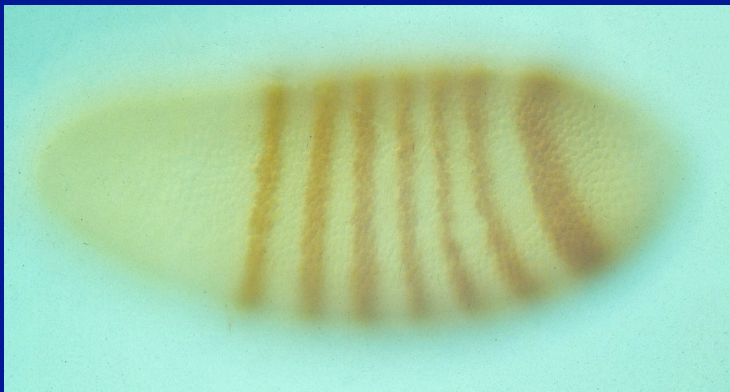
Parameter Sensitivity



- Only 9 out of 64 parameters have a significant impact when altered by 10% (ISI > 0.2). The same parameters were significant for other features.
- The most significant parameters are different for larger parameter changes (100%), due to model nonlinearity.
- All 9 parameters refer to reactions involving only free IKK and $I\kappa B\alpha$ — suggesting that model reduction might be possible.

22

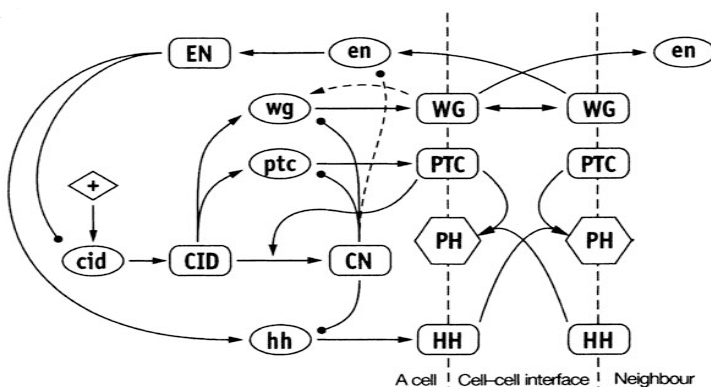
Segmental gene expression in the *Drosophila* embryo: pair-rule stripes



23

The segment polarity network

- A spatially distributed network, involving signalling between neighbouring cells.
- Ovals = mRNAs, rectangles = proteins, hexagons = protein complexes.



von Dassow, G. et al. Nature 406, 188 (2000)

24

