UNIVERSITY OF Southampton

MLSB09

Matching Models to Data in Modelling Morphogen Diffusion

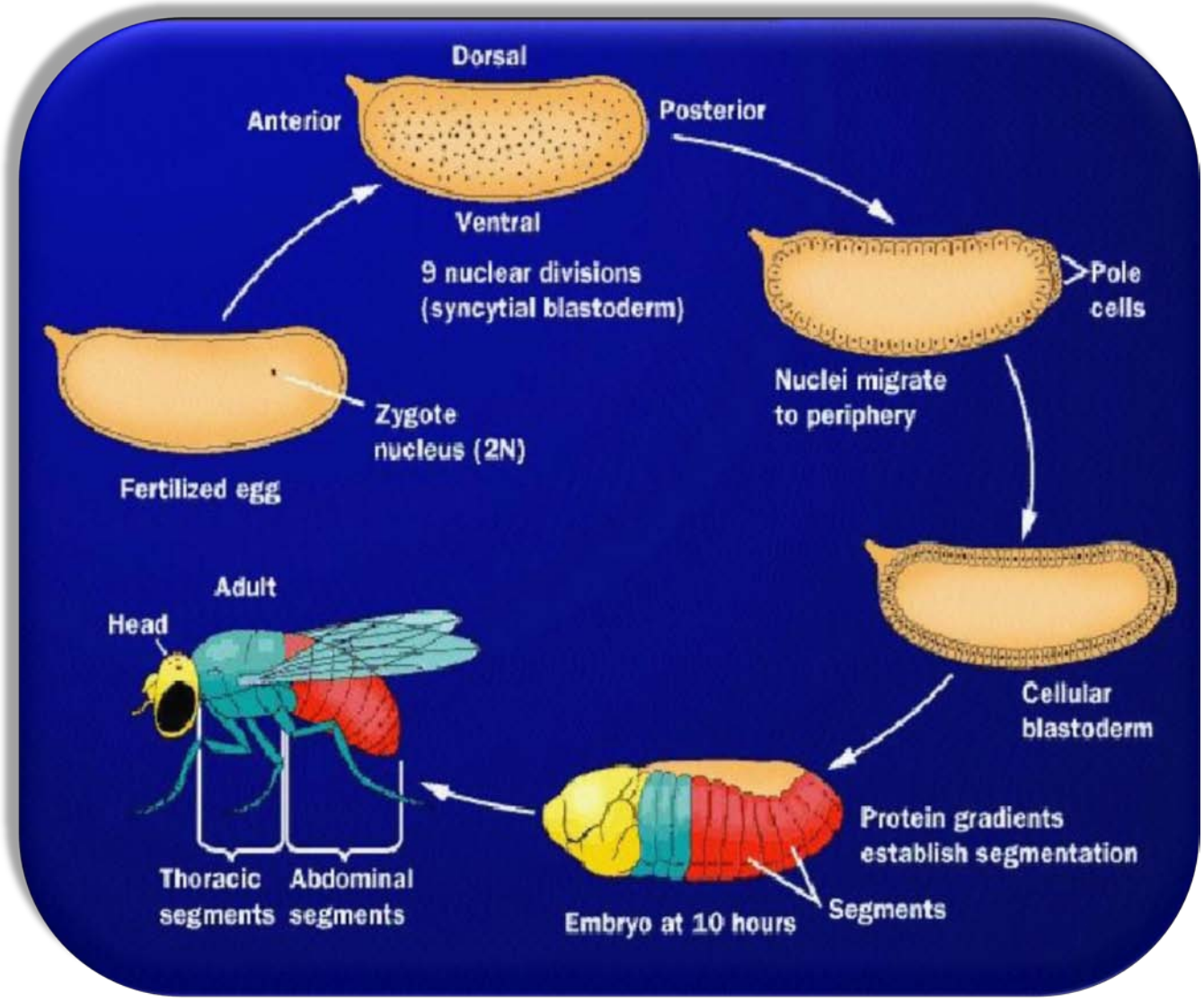
Wei Liu and Mahesan Niranjan

School of Electronics and Computer Science
University of Southampton
United Kingdom
{wl08r, mn}@ecs.soton.ac.uk

Drosophila

- Small flies
- Length of embryo: 500 μm
- Short generation time
- Genetics
- Common model

Drosophila Development



- After Fertilization
- Nuclear division
- No cytoplasm cleavage
- Syncytium
- Pattern formation
- Cellular blastoderm
- Body axes segment boundaries

from LIFE: The Science of Biology, Purves et al, 1998

Outline

- Passive diffusion models for spatial patterns establishment
- Constant supply
- Bicoid morphogens
- Constant supply followed by exponential decay
- Models vs Measured data
- Parameter estimation

What is the morphogen?

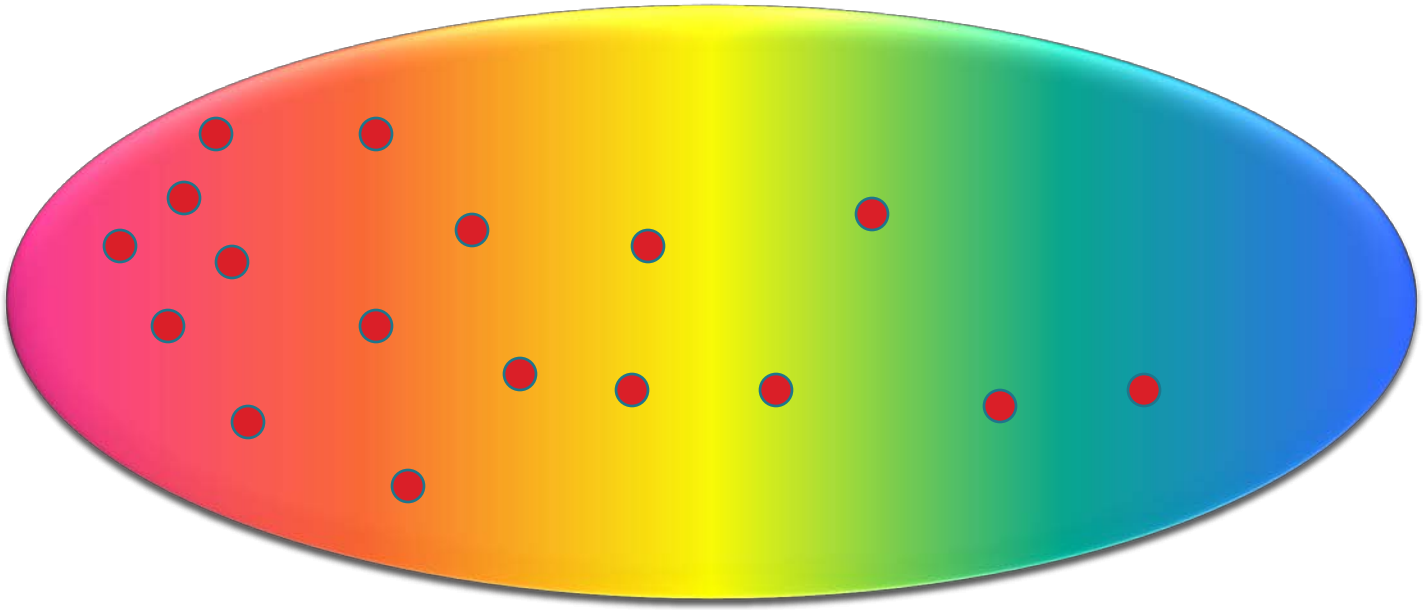
- Molecules in multi-cellular organism.
- Establishing spatial patterns of gene expression.
- Cells far from source: low level
- Cells close to source: high level
- Subdivided into different types.
- Different cells organise to form different organs

Turing, A.*: The chemical basis of morphogenesis.* Philosophical Transactions of the Royal Society B 237(641) (1952) 37–72

Bicoid Morphogen

- Drosophila body plan and position information.
- Contributing to set up the anterior posterior axis.
- Controlling cells fate along 70% of this axis.

Bicoid Morphogen Concentration



Anterior part

Posterior part

Reaction-diffusion Equation

- The reaction-diffusion equation of single-morphogen concentration system is below:

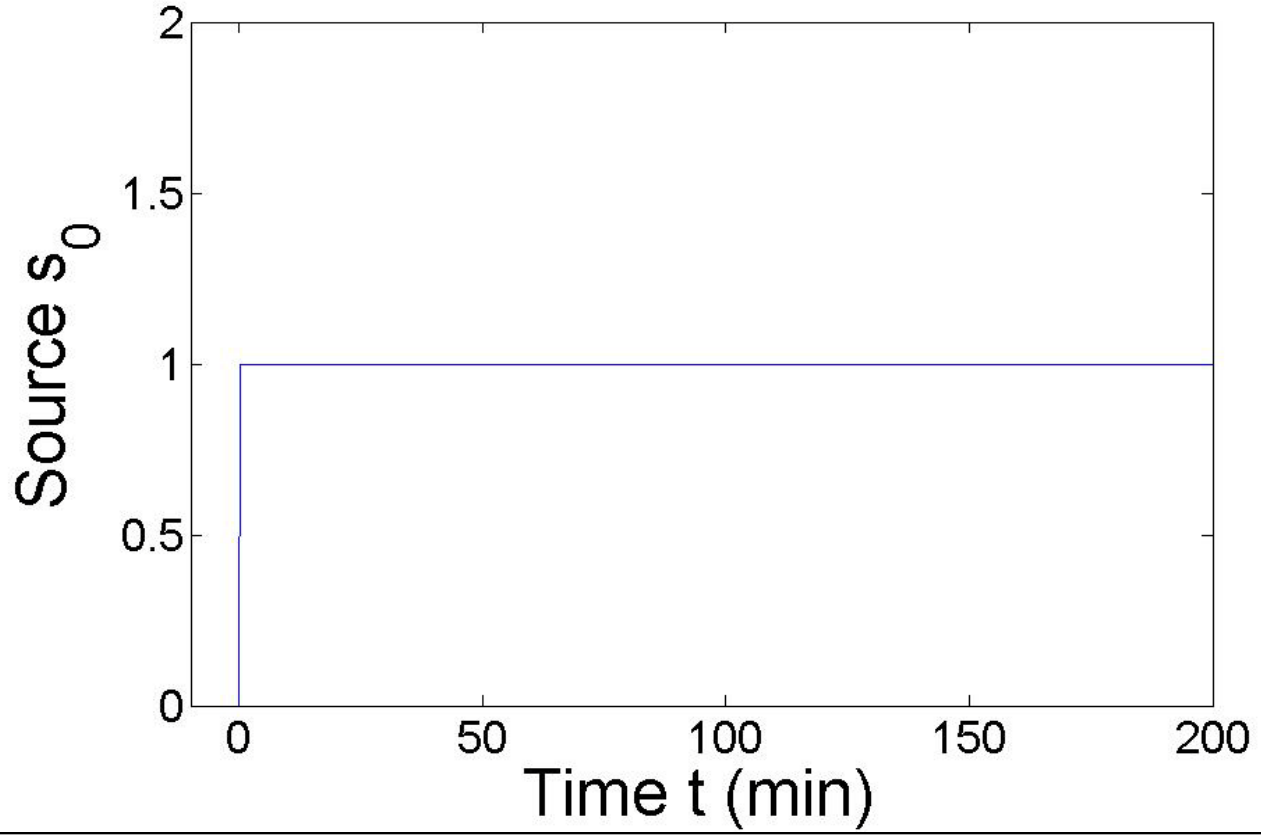
$$\frac{\partial}{\partial t} M(x,t) = D \frac{\partial^2}{\partial x^2} M(x,t) - \tau_p^{-1} M(x,t) + S(x,t)$$

- *M(x,t)* is morphogen concentration
- *S(x,t)* is a general source term at the anterior pole
- *D* is diffusion constant
- τ_p is half-life of the morphogen protein

Constant Source

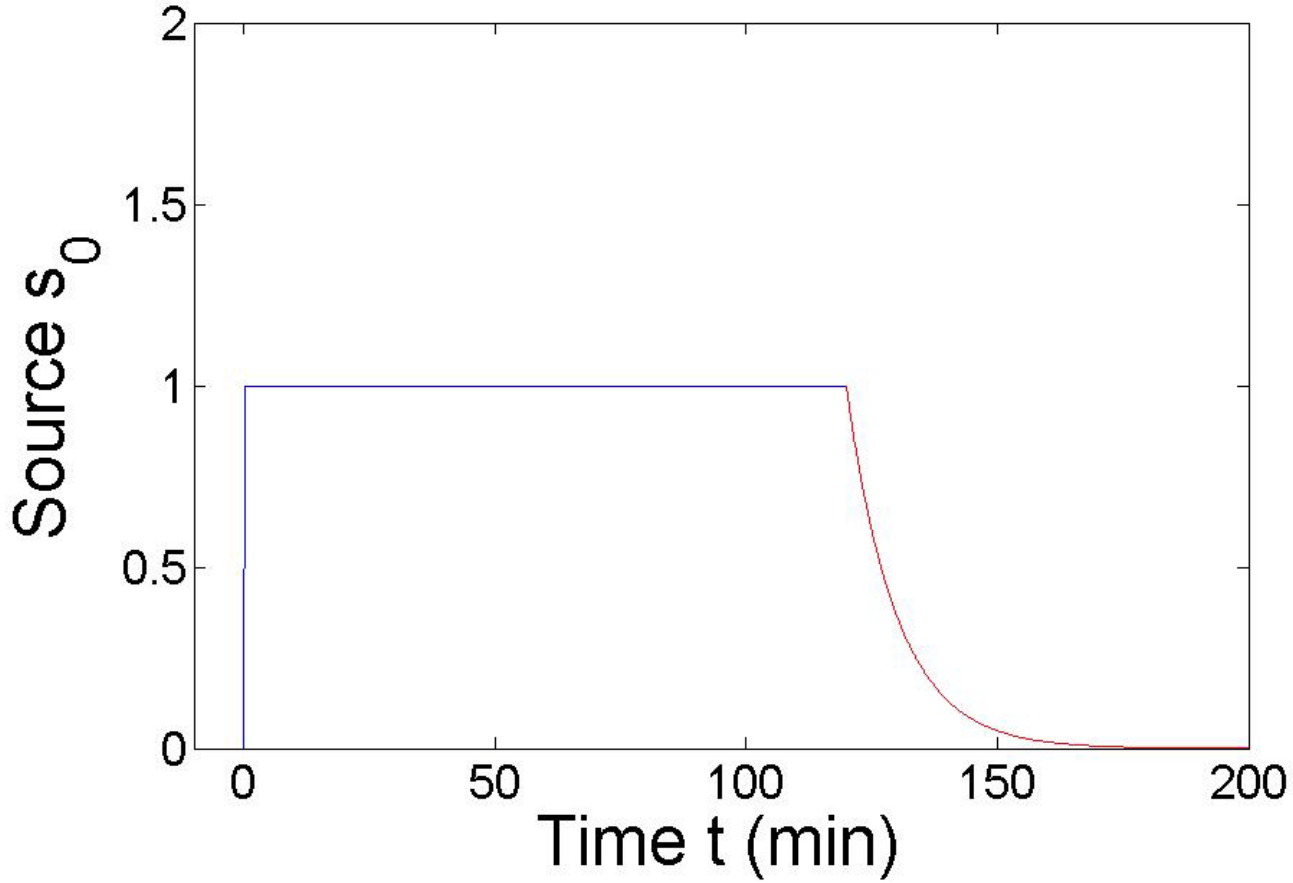
- Usual assumption

$$S_{con}(x,t) = S_0 \delta(x) \Theta(t)$$



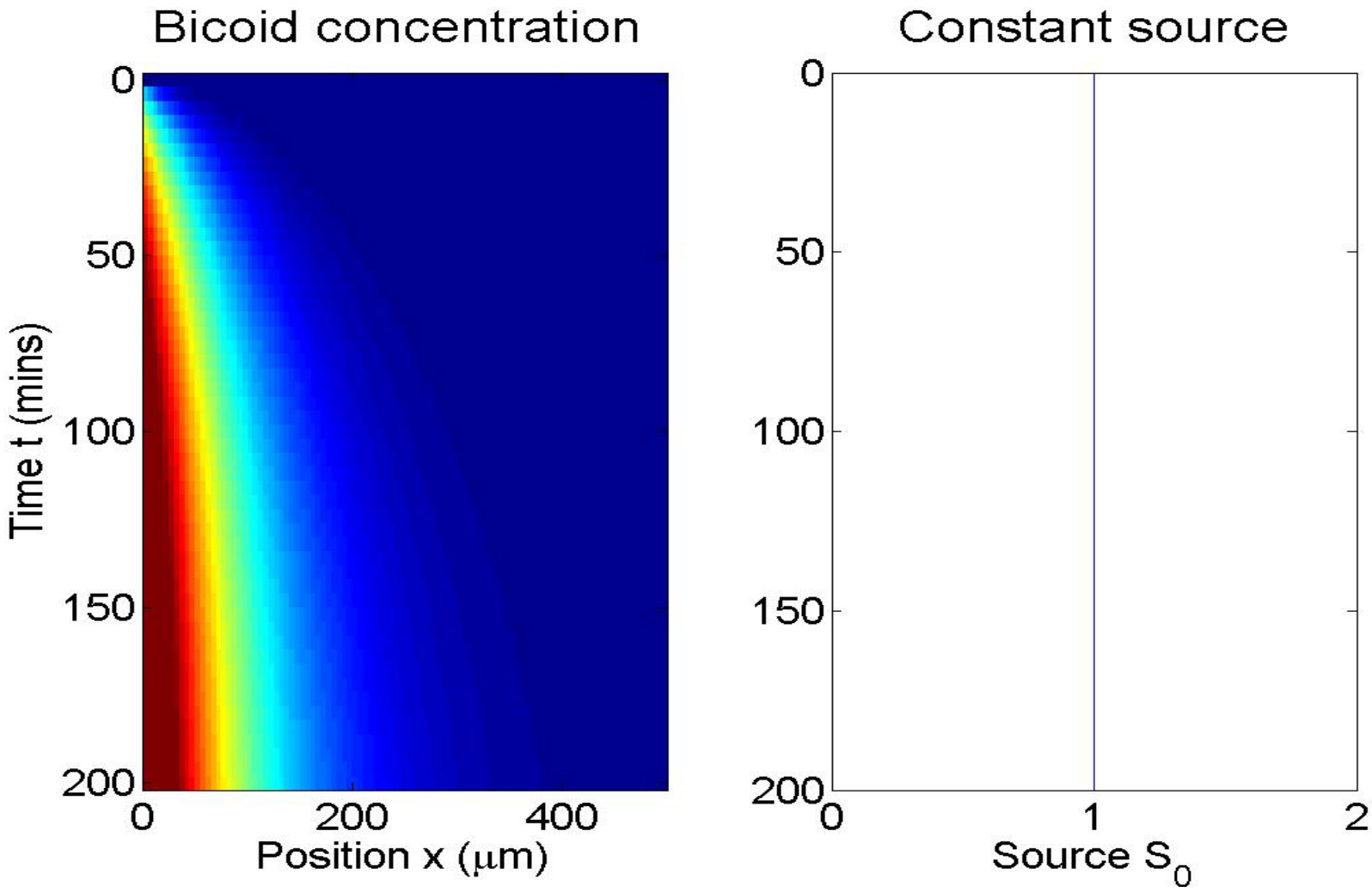
New Source Model

$$S_{com} = S_0\delta(x)(\Theta(t) - \Theta(t - t_0)) + S_0\delta(x)\Theta(t - t_0)\exp\left\{-\frac{t - t_0}{\tau_m}\right\}$$



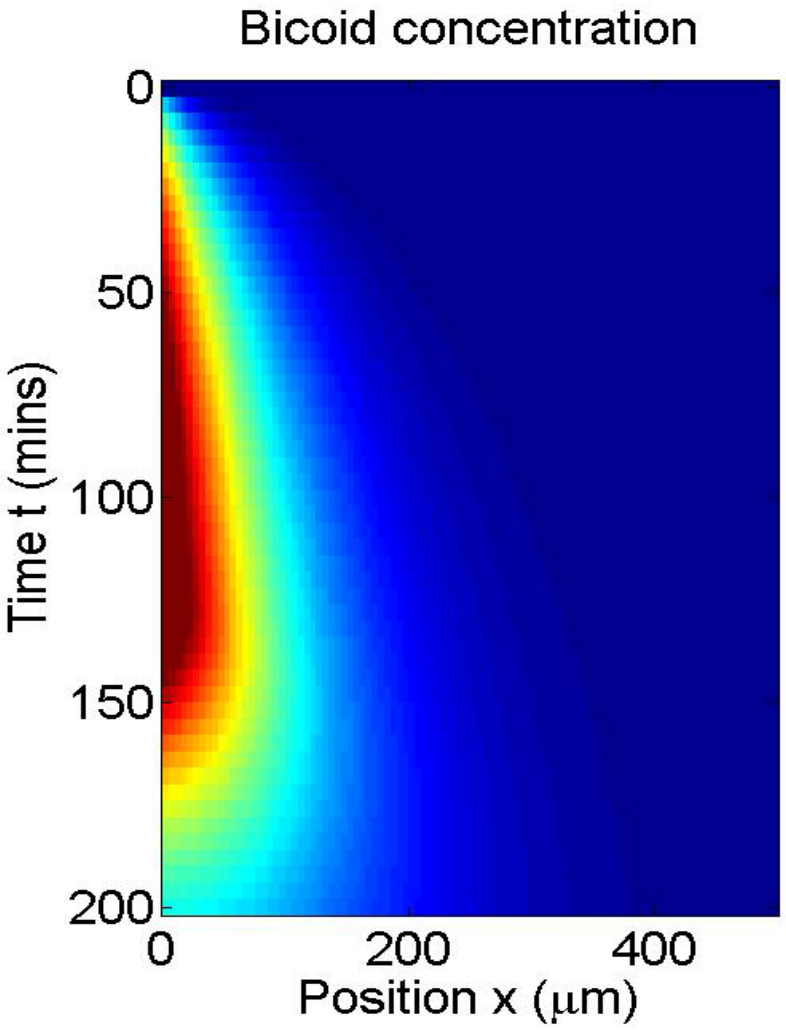
Surdej, P., Jacobs-Lorena, M.: ***Developmental regulation of bicoid mrna stability is mediated by the first 43 nucleotides of the 3' untranslated region***. Molecular and Cellular Biology 18(5) (1998) 2892–2900

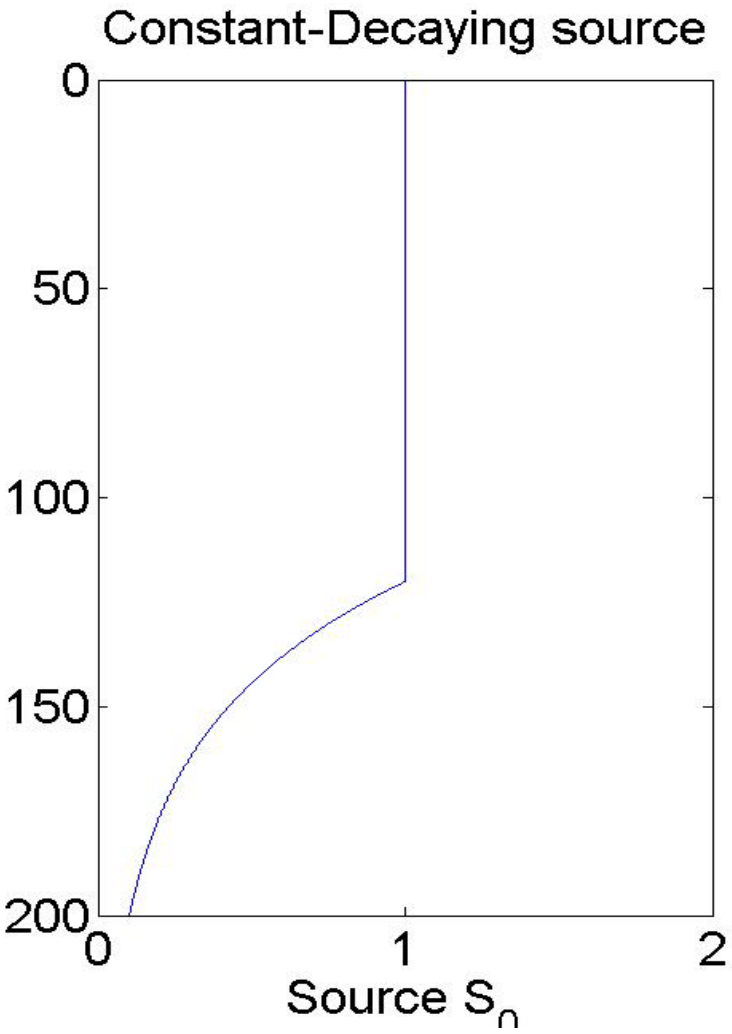
Widely used Model with constant source



The more realistic model

- More realistic





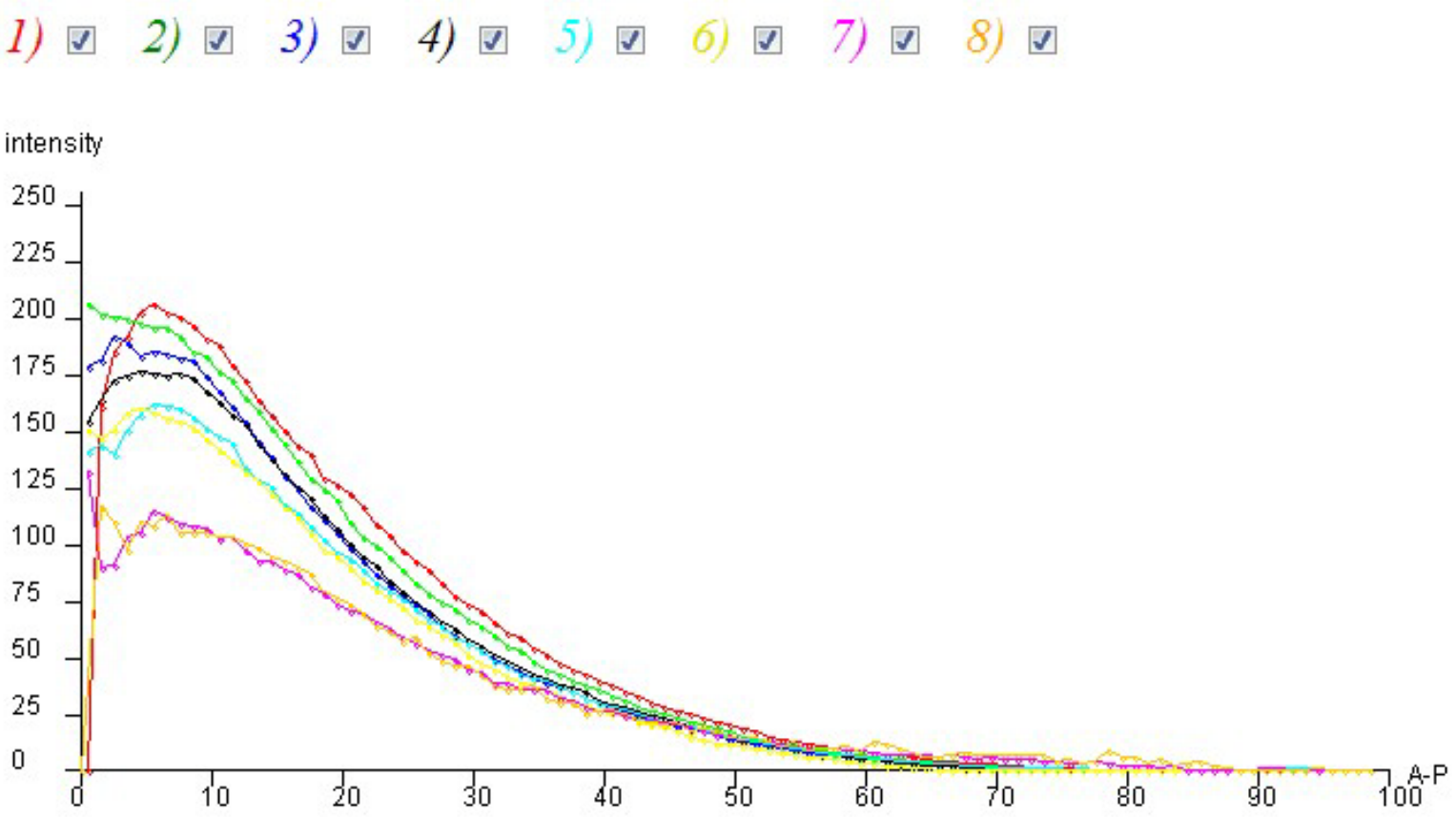
Measured Data I

- Flyex Database to estimate parameters of the model.
- Measured data: one dimension Bicoid integrated data in nuclear cleavage cycle 14A.
- Cycle 14A : 50 mins in duration; 8 equal temporal classes; 6.5 mins each class.

Andrei Pisarev, Ekaterina Poustelnikova, Maria Samsonova, John Reinitz (2009) **FlyEx, the quantitative atlas on segmentation gene expression at cellular resolution**. Nucl. Acids Res.; 37: D560 - D566.
Ekaterina Poustelnikova, Andrei Pisarev, Maxim Blagov, Maria Samsonova, and John Reinitz (2004). **A database for management of gene expression data *in situ***

Measured Data Ⅱ

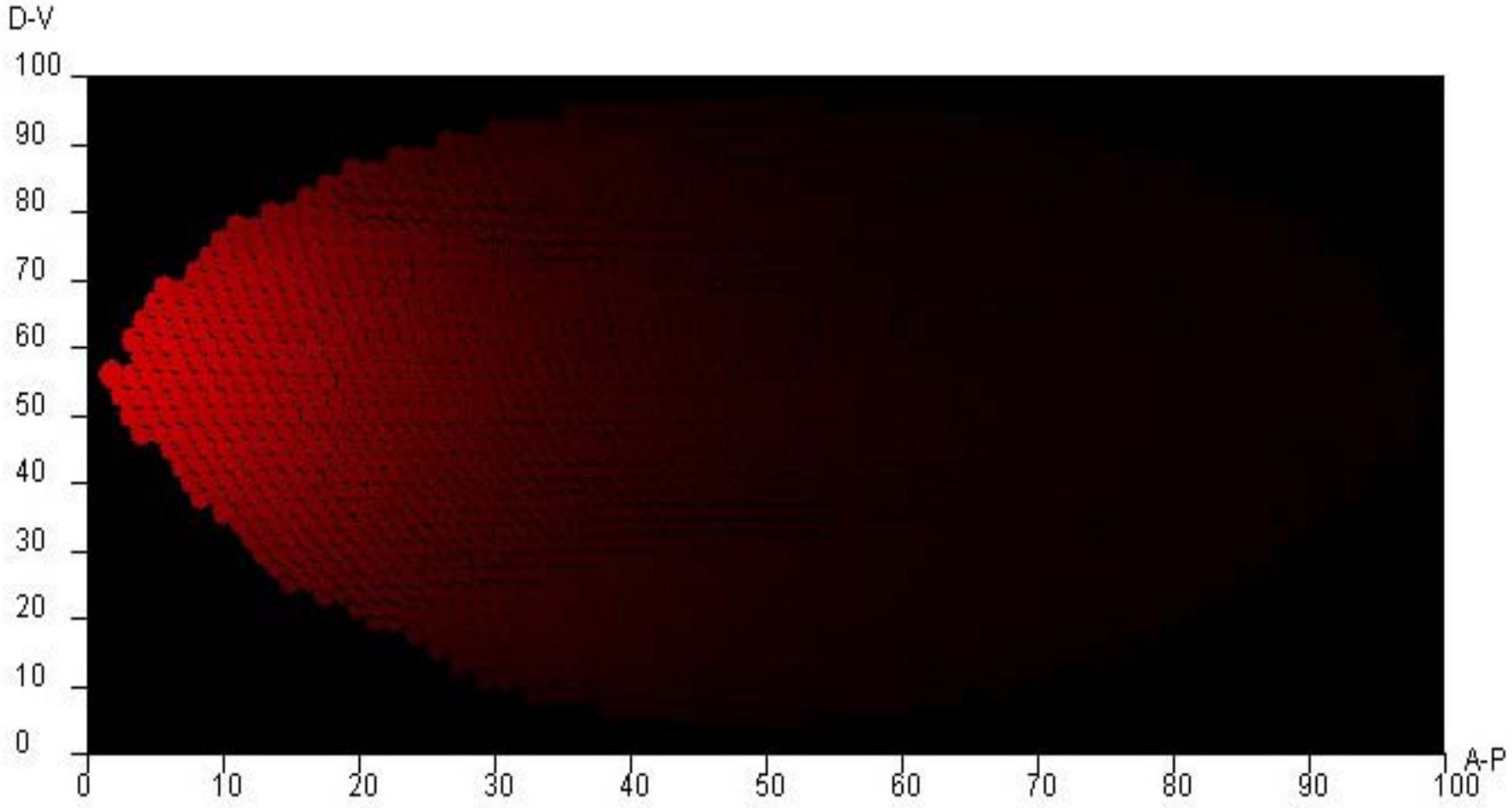
- 1D integrated data – cycle11- cycle 14A (1-8 classes)



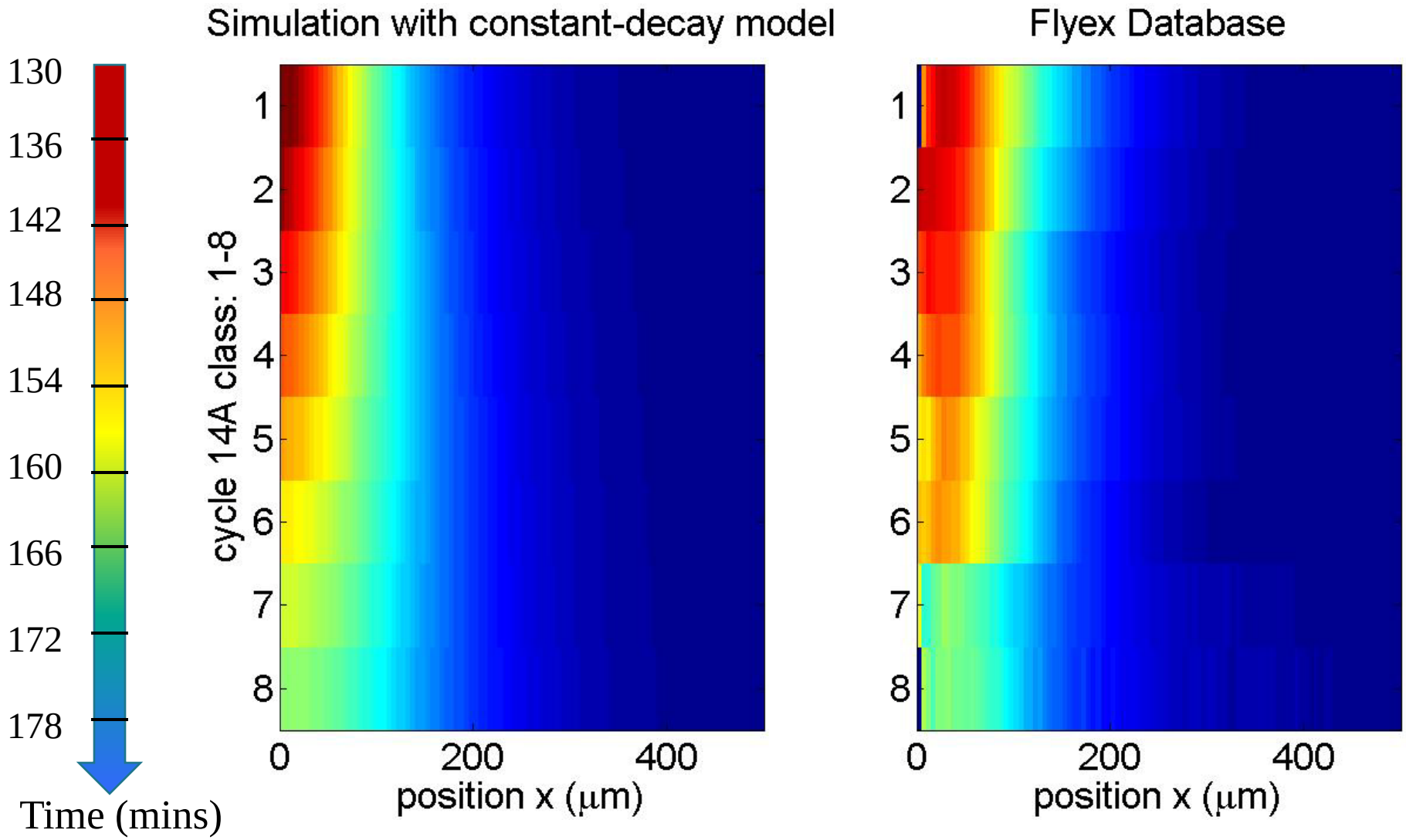
From FlyEx Database:

Measured Data Ⅲ

- Integrated 2D patterns – reconstructed image 14A-2



Comparison of model based and measured data of bicoid intensities



Matching Parameter Values to Data I

- Squared error between model output and measured intensities to evaluate error.

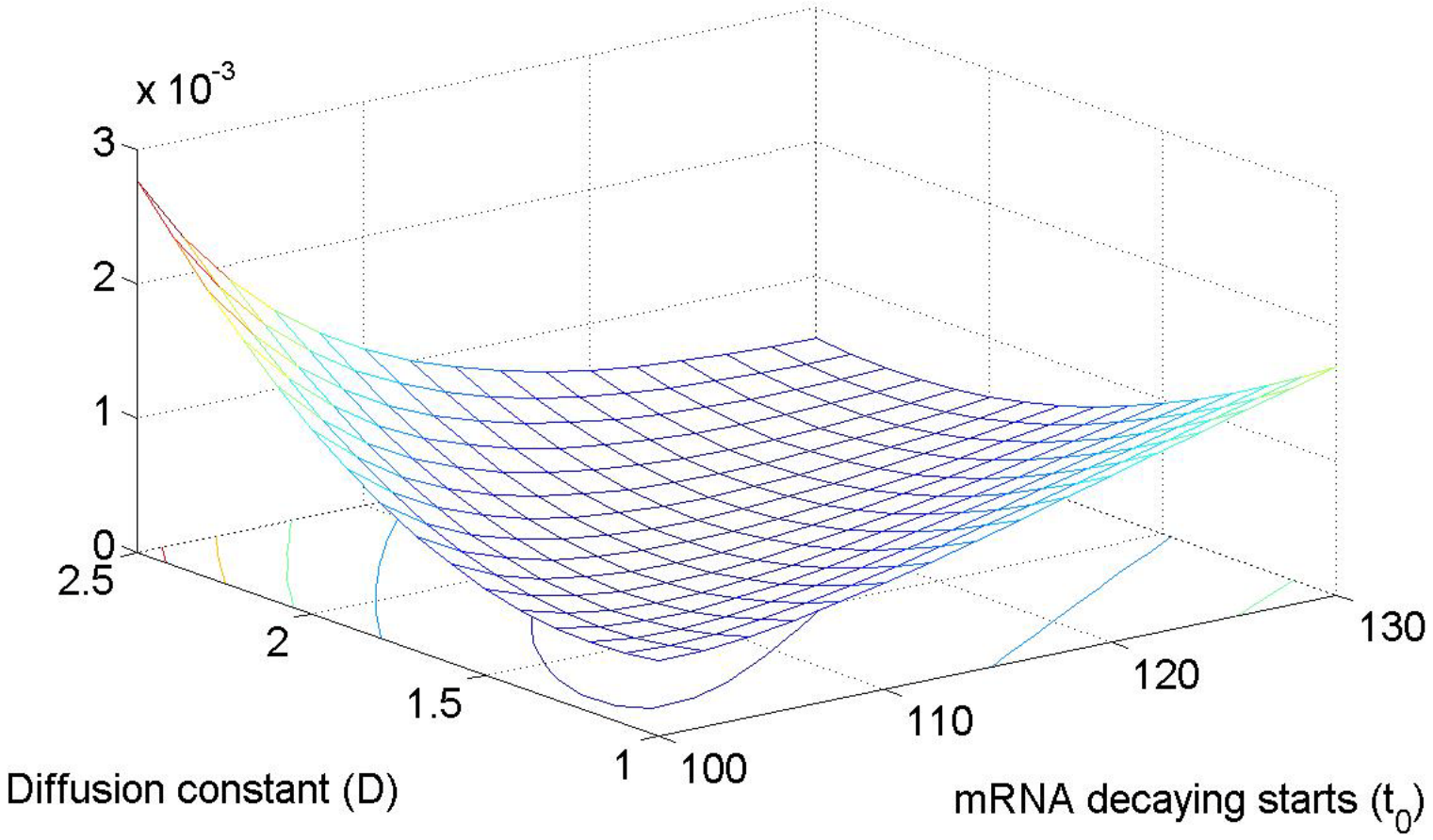
$$E = \sum_{t=T_1}^{T_2} \sum_{x=1}^{L} \{M(x,t) - M_d(x,t)\}^2$$

- Parameters estimation:

•Diffusion constant $D = 1.8 \mu m^2 / s$,
•The time mRNA starts to decay t_0 = 120mins,
• mRNA half-life τ_m = 29mins
•Bicoid protein half-life τ_p = 111mins.

Matching Parameter Values to Data Ⅱ

- The errors in the joint space of diffusion constant and maternal mRNA decay onset time.



Finding Optimal Values for all the Parameters Ⅲ

- Best combination of parameter values simultaneously.
- Diffusion constant D = 1.83 $\mu m^2 / s$,
- The time mRNA starts to decay t_0 = 118mins,
- mRNA half-life τ_m= 28.4mins
- Bicoid protein half-lifeτ_p = 120mins.

Conclusion and Future Work

- Widely used model with a constant source is unrealistic.
- By matching models output to data.
- The single measurements without uncertainties.

- Developing a stochastic model for a population of embryos i.e. Master equation model
- Developing data driven model for embryo spatio-temporal data i.e. Kriged Kalman Filter

UNIVERSITY OF Southampton

MLSB09

Thanks !

Kriged Kalman Filter

- KF – linear Gaussian state space model

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{w}_t$$

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{v}_t$$

  - KKF –modelling spatio-temporal data (Mardia *et. al 1998*)

$$M(x,t) = \mathbf{h}(x)^T \mathbf{a}(t) + \varepsilon(x,t)$$

$$\mathbf{a}(t) = P\mathbf{a}(t-1) + \eta(x,t)$$

FlyEx Database

- **Step 1 : Data Acquisition**

  Acquisition of quantitative data on gene expression in individual embryos
- **Step 2 : Data Registration**

  Excision of 10 % stripe of quantitative data

  Feature extraction

  Registration
- **Step 3: Data Normalization**

  Rescaling of data to bring data to unified standard form with a zero background
- **Step 4: Data Averaging**

  Construction of the integrated pattern of each gene expression