Dugga 2025-03-20

TBMT19 / TBMT37

Write your Dugga-ID on all pages and your answers in Swedish or English. There are in total 5 questions, each worth 3 points. You need at least 10 points with 2 points per question, or 12 points in total to pass. Good luck! /William

1 Model formulation and model parts

You have the following model:

$$\frac{dA}{dt} = -v1 + v2 \qquad v1 = k1 \cdot A \cdot u \qquad v2 = k2 \cdot C$$

$$\frac{dB}{dt} = v1 - v3 \qquad v3 = k3 \cdot B \qquad v4 = k4 \cdot B$$

$$\frac{dC}{dt} = v4 - v2 - v5 \qquad v5 = k5 \cdot C$$

$$[k1, k2, k3, k4, k5] = [1, 1, 2, 2, 3]$$

$$[A(0), B(0), C(0)] = [1, 0, 0]$$

$$u = 2$$

(a) Give an interaction graph corresponding to the model (1 point).

(b) List the model states (1 point).

(c) Make the reaction $\emptyset \rightarrow C$ saturated with respect to B (1 point).

2 Model simulation

(a) Use the Euler forward method to simulate the following model with step-length $\Delta t = 0.1$. What are the values of X and Y in t = 0.1, and t = 0.2? (1 point)

$$d/dt(X) = v1 - v2, \qquad v1 = k1 \cdot Y$$

$$d/dt(Y) = -v1 + v2, \qquad v2 = k2 \cdot X$$

$$k1 = 2, \quad k2 = 2, \quad X(0) = 100, \quad Y(0) = 0$$

(b) What do you think happens with X and Y, if we simulate the model for a long time? (1 point)

(c) Why do we typically numerically simulate the ODEs instead of solving them analytically? (1 point)

3 Parameter estimation

(a) You have two model simulations ($\hat{y}_1$ and $\hat{y}_2$) with three different time-points each for a given parameter set (θ):

$$\hat{y}_1(\theta) = [2,2,2]$$

$$\hat{y}_2(\theta) = [3,3,3]$$

You also have corresponding experimental data (y_1 and y_2) with measurement uncertainties (SEM_1 and SEM_2):

$$y_1 = [4,4,4], \qquad SEM_1 = [2,2,2]$$

$$y_2 = [2,2,2], \qquad SEM_2 = [1,1,1]$$

What is the cost for the parameter set θ? (1 points)

(b) Why do we square the residuals? (1 point)

(c) What is the difference between local and global optimization solvers? (1 point)

4 Statistical tests

Assume that we again have the same model simulations and data as in question 3 above.

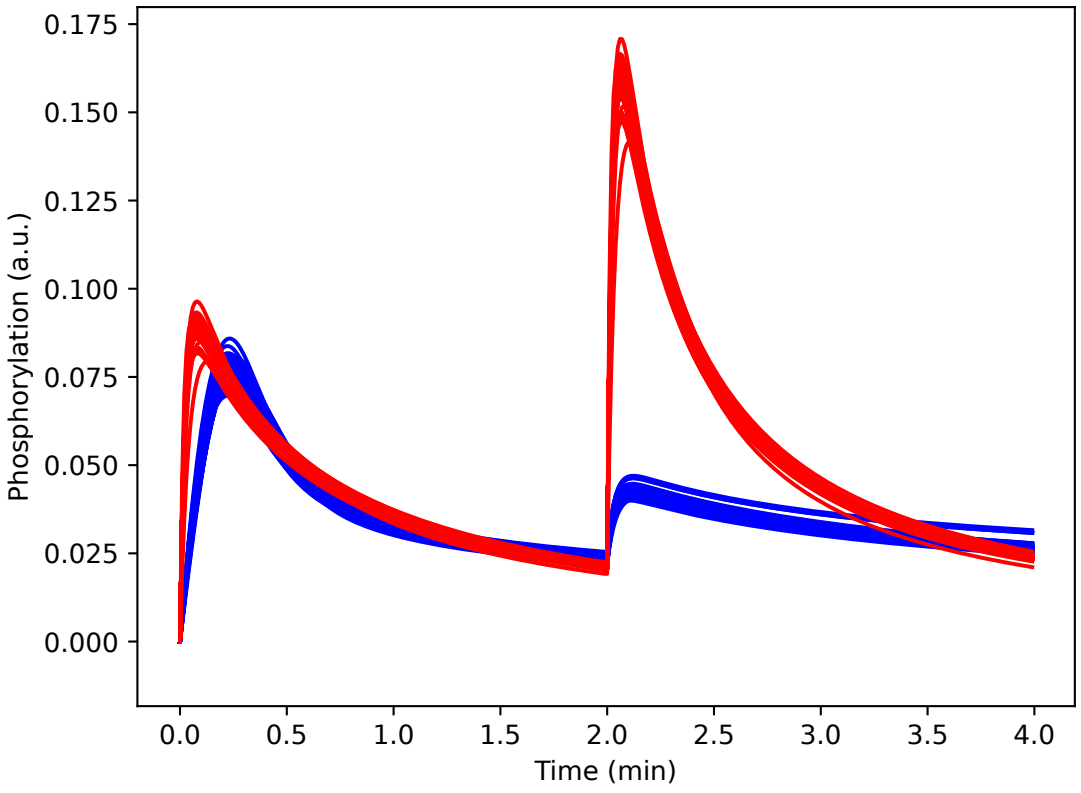
(a) If the χ^2-limit is 12.6, should you reject the model simulation given the parameter set θ? What is the conclusion of the test? (1 point)
If you did not calculate the cost in 3(a), assume that the cost is 11.8.

(b) Would you reject the model simulation with a whiteness test? Why/why not? (1 point)

(c) Based on your conclusion from 4(a), should you reject the *model* if the parameter set θ is *not* the optimal parameter set? (1 point)

5 Predictions and experimental design

Assume that you have the following model predictions from two different models: a *red* model and a *blue* model. The predictions correspond to the phosphorylation of a protein over time.



(a) Why would it be more useful to (experimentally) measure the phosphorylation between 2 and 4 minutes rather than between 0 and 2 minutes? (1 point)

(b) Sketch the figure (or use this paper and remember to hand it in). Draw one experimental data point (with a reasonable uncertainty). Explain which conclusions you can draw based on your data point. (1 point)

(c) What is the risk if we underestimate the model uncertainty (i.e. that the model uncertainty is narrower than it should be)? (1 point)

Answers: Dugga 2025-03-20

1 Model formulation and model parts

You have the following model:

$$\frac{dA}{dt} = -v1 + v2 \qquad v1 = k1 \cdot A \cdot u \qquad v2 = k2 \cdot C$$

$$\frac{dB}{dt} = v1 - v3 \qquad v3 = k3 \cdot B \qquad v4 = k4 \cdot B$$

$$\frac{dC}{dt} = v4 - v2 - v5 \qquad v5 = k5 \cdot C$$

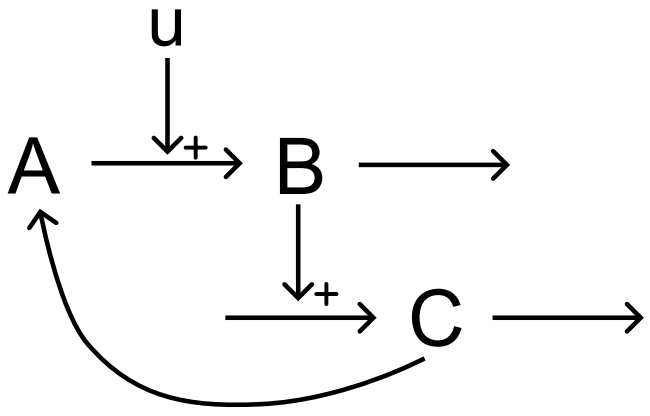
$$[k1, k2, k3, k4, k5] = [1, 1, 2, 2, 3]$$

$$[A(0), B(0), C(0)] = [1, 0, 0]$$

$$u = 2$$

(a) Give an interaction graph corresponding to the model (1 point).

Answer:
The interaction graph can look like this:



(b) List the model states (1 point).

Answer:
The model states are A, B, and C.

(c) Make the reaction $\emptyset \rightarrow C$ saturated with respect to B (1 point).

Answer:
To make $v4$ be saturated with respect to B update the reaction rate to:

$$v4 = Vmax \cdot \frac{B}{kM + B}$$

Value for the new introduced parameter (made up): $Vmax = 2$, $kM = 1$.

2 Model simulation

(a) Use the Euler forward method to simulate the following model with step-length $\Delta t = 0.1$. What are the values of X and Y in t = 0.1, and t = 0.2? (1 point)

$$d/dt(X) = v1 - v2, \qquad v1 = k1 \cdot Y$$

$$d/dt(Y) = -v1 + v2, \qquad v2 = k2 \cdot X$$

$$k1 = 2, \quad k2 = 2, \quad X(0) = 100, \quad Y(0) = 0$$

Answer:
In the Euler forward method, we start with the initial values of the states and calculate the values of the ODEs at the initial time-point. We then take a small step in the direction of the ODEs, and use the new values to calculate the ODEs at the next time-point. This is repeated until the time-frame asked for by the user has been simulated.

The Euler forward method can be written as:

$$x(t + \Delta t) = x(t) + d/dt(x(t)) \cdot \Delta t$$

For this model, to calculate the values of X and Y in t = 0.1, and t = 0.2 we start by calculating the values for t = 0.1 using the following steps:

$$v1(0) = k1 \cdot Y(0) = 2 \cdot 0$$

$$v2(0) = k2 \cdot X(0) = 2 \cdot 100 = 200$$

$$d/dt(X(0)) = v1(0) - v2(0) = 0 - 200 = -200$$

$$X(0.1) = X(0) + d/dt(X(0)) \cdot \Delta t = 100 - 200 \cdot 0.1 = 80$$

$$d/dt(Y(0)) = -v1(0) + v2(0) = 0 + 200 = 200$$

$$Y(0.1) = Y(0) + d/dt(Y(0)) \cdot \Delta t = 0 + 200 \cdot 0.1 = 20$$

Then we calculate the values for t = 0.2 using the following steps:

$$v1(0.1) = k1 \cdot Y(0.1) = 2 \cdot 20 = 40$$

$$v2(0.1) = k2 \cdot X(0.1) = 2 \cdot 80 = 160$$

$$d/dt(X(0.1)) = v1(0.1) - v2(0.1) = 40 - 160 = -120$$

$$X(0.2) = X(0.1) + d/dt(X(0.1)) \cdot \Delta t = 80 - 120 \cdot 0.1 = 68$$

$$d/dt(Y(0.1)) = -v1(0.1) + v2(0.1) = -40 + 160 = 120$$

$$Y(0.2) = Y(0.1) + d/dt(Y(0.1)) \cdot \Delta t = 20 + 120 \cdot 0.1 = 32$$

So, the values of X and Y in t = 0.1, and t = 0.2 are:

$$X(0.1) = 80, \quad X(0.2) = 68$$

$$Y(0.1) = 20, \quad Y(0.2) = 32$$

(b) What do you think happens with X and Y, if we simulate the model for a long time? (1 point)

Answer:
If we simulate the model for a long time, X and Y will reach a steady-state. Since both X and Y have both positive and negative reactions, the steady state will be some positive, non-zero, non-infinity value.

Extra answer, not needed for point:
To find the steady-state, we can set the ODEs to zero and solve for the states. In this case we get

$$0 = v1 - v2$$

$$0 = -v1 + v2 \iff 0 = v1 - v2$$

In other words:

$$0 = v1 - v2 \iff v1 = v2 \iff k1 \cdot Y = k2 \cdot X \iff 2 \cdot Y = 2 \cdot X \iff Y = X$$

Since we started with $X = 100$ and $Y = 0$, and X can only be converted into Y and vice versa, the sum of X and Y will be 100. Together we get that:

$$\left.\begin{array}{l} X + Y = 100 \\ \\ X = Y \end{array}\right\} \implies 2X = 100 \iff X = Y = 50$$

So, the model will reach a steady state when $X = Y = 50$.

(c) Why do we typically numerically simulate the ODEs instead of solving them analytically? (1 point)

Answer:
In systems biology, we often have complex models with many states and reactions, and it is often not feasible to solve these ODEs analytically. Therefore, we need to use numerical simulation to solve the ODEs and obtain the model simulations.

3 Parameter estimation

(a) You have two model simulations ($\hat{y}_1$ and $\hat{y}_2$) with three different time-points each for a given parameter set (θ):

$$\hat{y}_1(\theta) = [2,2,2]$$

$$\hat{y}_2(\theta) = [3,3,3]$$

You also have corresponding experimental data (y_1 and y_2) with measurement uncertainties (SEM_1 and SEM_2):

$$y_1 = [4,4,4], \qquad SEM_1 = [2,2,2]$$

$$y_2 = [2,2,2], \qquad SEM_2 = [1,1,1]$$

What is the cost for the parameter set θ? (1 points)

Answer:
The cost for the parameter set θ is calculated as follows:

$$residuals = y - \hat{y}$$

$$residuals_1 = [4,4,4] - [2,2,2] = [2,2,2]$$

$$residuals_2 = [2,2,2] - [3,3,3] = [-1,-1,-1]$$

$$v(\theta) = \sum_{t=1}^{N} \frac{(y(t) - \hat{y}(t,\theta))^2}{SEM(t)^2} = \sum_{i=1}^{M}\sum_{t=1}^{N_i} \frac{(y_i(t) - \hat{y}_i(t,\theta))^2}{SEM_i(t)^2} =$$

$$= \sum_{t=1}^{N_1} \frac{(y_1(t) - \hat{y}_1(t,\theta))^2}{SEM_1(t)^2} + \sum_{t=1}^{N_2} \frac{(y_2(t) - \hat{y}_2(t,\theta))^2}{SEM_2(t)^2} =$$

$$= \left(\frac{2^2}{2^2} + \frac{2^2}{2^2} + \frac{2^2}{2^2}\right) + \left(\frac{(-1)^2}{1^2} + \frac{(-1)^2}{1^2} + \frac{(-1)^2}{1^2}\right) =$$

$$= \left(\frac{4}{4} + \frac{4}{4} + \frac{4}{4}\right) + \left(\frac{1}{1} + \frac{1}{1} + \frac{1}{1}\right) = 3 + 3 = 6$$

(b) Why do we square the residuals? (1 point)

Answer:
So that the sign of the residuals should not be important. Extra (not needed) answer: Also, squaring the residuals gives more weight to larger deviations between the model and the data. Furthermore, it makes the cost (when normalized with the squared residuals) follow a chi-squared distribution.

(c) What is the difference between local and global optimization solvers? (1 point)

Answer:
Local optimization solvers strictly search in directions which improve cost. This means that they might get stuck in local minima. Global optimization solvers, on the other hand, can accept solutions that increase the cost at some occasions, which allows them to escape local minima.

4 Statistical tests

Assume that we again have the same model simulations and data as in question 3 above.

(a) If the χ^2-limit is 12.6, should you reject the model simulation given the parameter set θ? What is the conclusion of the test? (1 point)
If you did not calculate the cost in 3(a), assume that the cost is 11.8.

Answer:
The χ^2-limit is 12.6, and the cost is 6 (or 11.8). Since the cost is smaller than the χ^2-limit, we should not reject the model simulation given the parameter set θ. The conclusion of the test is that we do not draw any conclusions.

(b) Would you reject the model simulation with a whiteness test? Why/why not? (1 point)

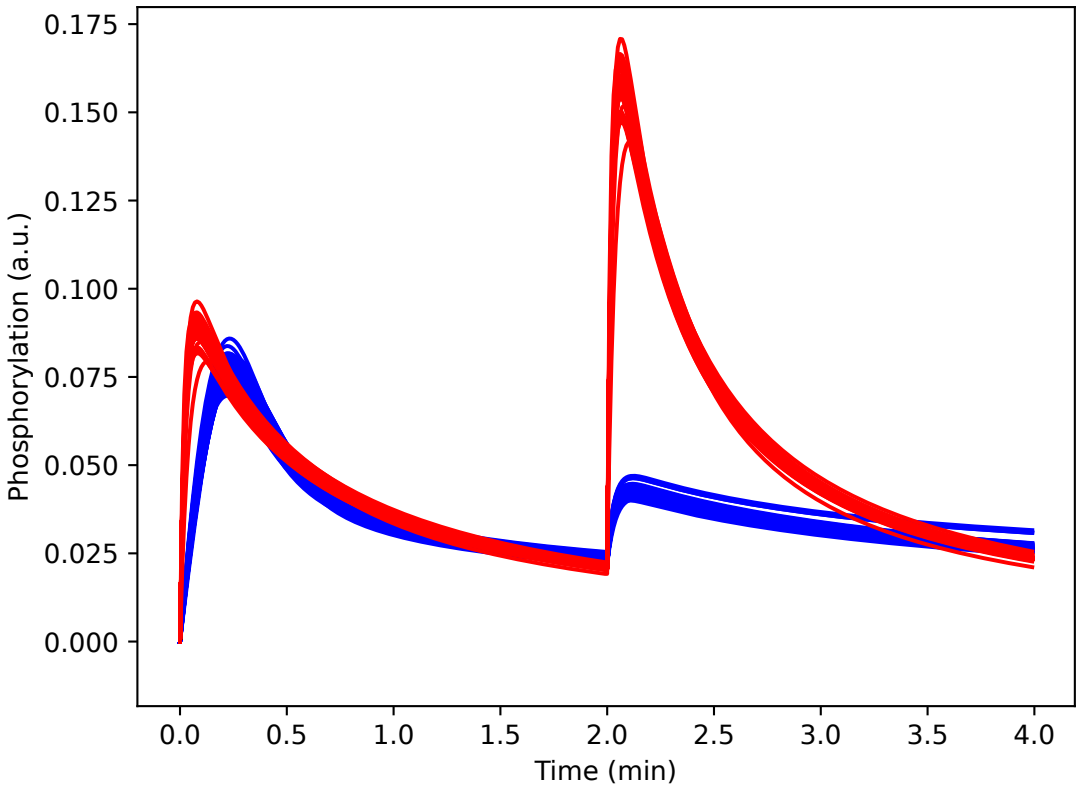
Answer:
Yes, the residuals for both simulations are correlated ($y_1 > \hat{y}_1$, and $y_2 < \hat{y}_2$ for all time-points). Therefore, we would reject the model simulation with a whiteness test (the residuals are too correlated).

(c) Based on your conclusion from 4(a), should you reject the *model* if the parameter set θ is *not* the optimal parameter set? (1 point)

Answer:
No, the parameter set θ gives simulations which cannot be rejected with a χ^2-test. Thus, we know that the model can give simulations which are not rejected, and we cannot reject the model based on the results from 4(a).

5 Predictions and experimental design

Assume that you have the following model predictions from two different models: a *red* model and a *blue* model. The predictions correspond to the phosphorylation of a protein over time.

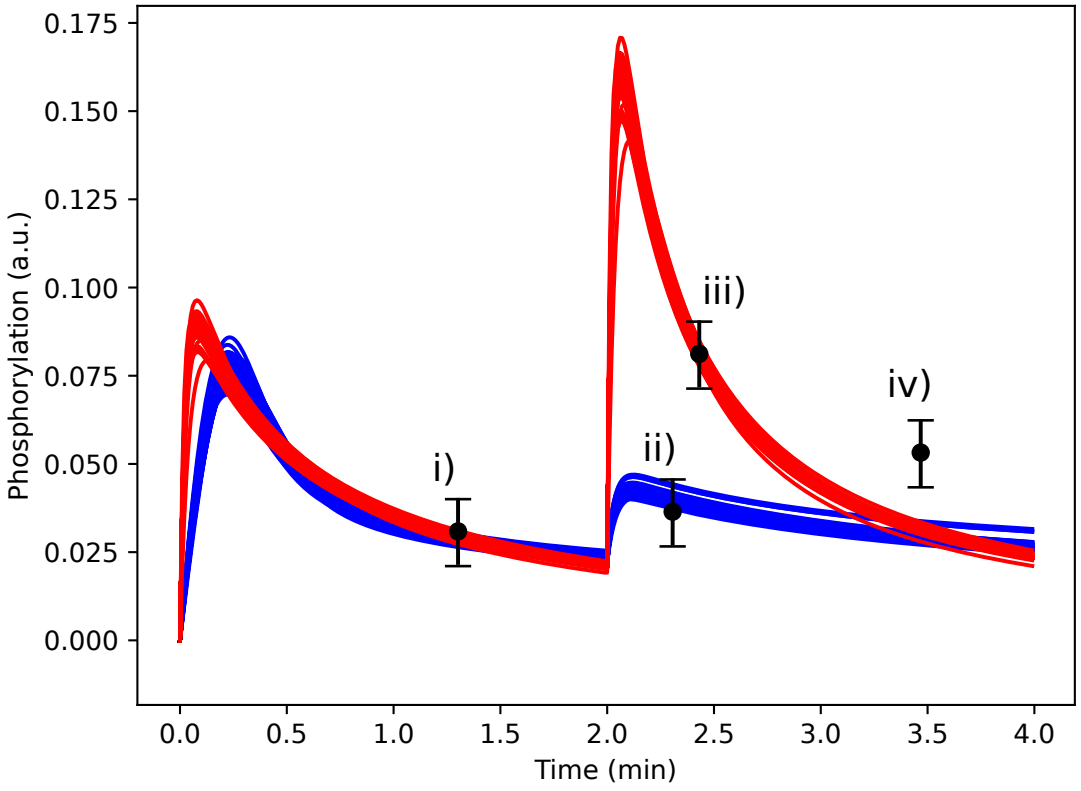


(a) Why would it be more useful to (experimentally) measure the phosphorylation between 2 and 4 minutes rather than between 0 and 2 minutes? (1 point)

Answer:
If the experimental data has a reasonable uncertainty, and the two models are behaving differently, we will be able to reject at least one of the models.

(b) Sketch the figure (or use this paper and remember to hand it in). Draw one experimental data point (with a reasonable uncertainty). Explain which conclusions you can draw based on your data point. (1 point)

Answer:
One data point is enough, but I here provide four different (equally correct) answers:



i) Reject neither of the the models
ii) Reject the red model
iii) Reject the blue model
iv) Reject both models

(c) What is the risk if we underestimate the model uncertainty (i.e. that the model uncertainty is narrower than it should be)? (1 point)

Answer:
If we did not quantify the uncertainty correctly, we could draw the wrong conclusions from the data. If the uncertainty is underestimated, we might reject a model that is actually correct.