Nonlinear model predictive control of *Escherichia coli* culture

Merouane ABADLI [1,2], Sihem TEBBANI [1], Didier DUMUR[1], Alain VANDE WOUWER[2], Laurent DEWASME[2]

[1]: Laboratoire des Signaux et Systèmes, CentraleSupélec, CNRS, Univ. Paris Sud, Univ. Paris-Saclay, Control Department, 3 rue Joliot-Curie, 91192 Gif sur Yvette, France

2: Automatic Control Laboratory, University of Mons, 31, Boulevard Dolez , 7000 Mons, Belgium

Contact : merouane.abadli@centralesupelec.fr

Context

- *Escherichia coli* is a popular microorganism in biotechnology applications, and the most commonly used host cell for the production of recombinant proteins and many other biopharmaceutical products.
- Computer control of the biochemical state variables can help to increase performance significantly.
- To maximize the biomass production and reach high cell densities, a substrate feeding strategy must be considered.
- Overfeeding the culture can lead into acetate production, a cell growth inhibiting byproduct.
- To maintain the culture in optimal operating conditions, an optimal closed-loop control algorithm coupled with a state estimator is developped.

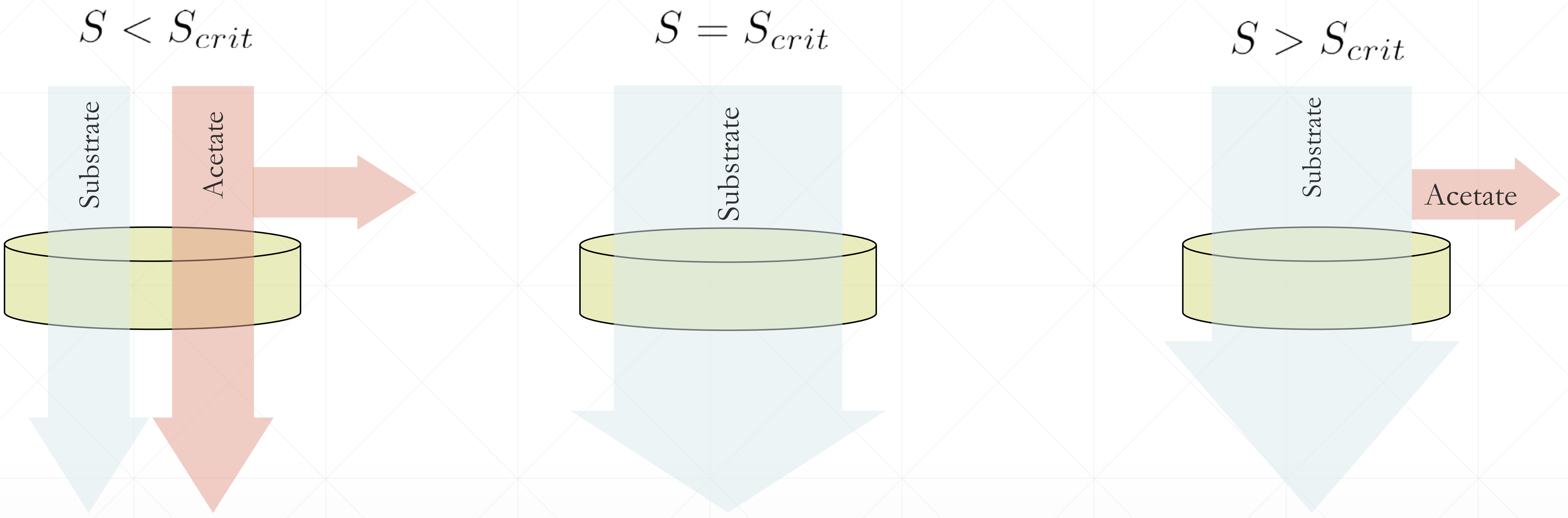
OUTLINE

1. Process model
2. Model predictive control
3. Unscented Kalman Filter
4. Simulation results
5. Conclusion and perspectives

Progress

1. **Process model**
2. Model predictive control
3. Unscented Kalman Filter
4. Simulation results
5. Conclusion and perspectives

Bottleneck assumption



Respirative regime

Optimal operating conditions

Respiro-fermentative regime

$$Substrate + Oxygen \rightarrow Biomass + Carbon\ dioxide$$

$$Substrate + Oxygen \rightarrow Biomass + Acetate + Carbon\ dioxide$$

$$Acetate + Oxygen \rightarrow Biomass + Carbon\ dioxide$$

Model

The macroscopic model of E. coli follows the reaction scheme :

$$
\begin{aligned}
Substrate\ oxidation &: \quad k_{s1}S + k_{o1}O \xrightarrow{\mu_1 X} k_{x1}X + k_{c1}C \\
Substrate\ fermentation &: \quad k_{s2}S + k_{o2}O \xrightarrow{\mu_2 X} k_{x1}X + k_{A2}A + k_{c2}C \\
Acetate\ oxydation &: \quad k_{A3}A + k_{o3}O \xrightarrow{\mu_3 X} k_{x3}X + k_{c3}C
\end{aligned}
$$

- X, S, A, O, et C are respectively, the biomass, glucose, acetate, dissolved oxygen and carbon dioxyde.
- k_{ξ_i} $(i = 1, 2, 3)$ are the yield coefficients.
- μ_i $(i = 1, 2, 3)$ are the specific growth rates.

The specific growth rates:

$$\mu_1 = \frac{min(q_s, q_{s_{crit}})}{k_{s1}}$$
$$\mu_2 = \frac{max(0, q_s - q_{s_{crit}})}{k_{s2}}$$
$$\mu_3 = \frac{min(0, q_{AC})}{k_{A3}}$$

The consumption rates:

$$q_s = q_{s_{max}} \frac{S}{K_s + S}$$
$$q_{s_{crit}} = \frac{q_{O_{max}}}{k_{OS}} \frac{K_{iA}}{K_{iA} + A}$$
$$q_{AC} = \frac{k_{OS}(q_{s_{crit}} - q_s)}{k_{OA}} \frac{A}{K_A + A}$$

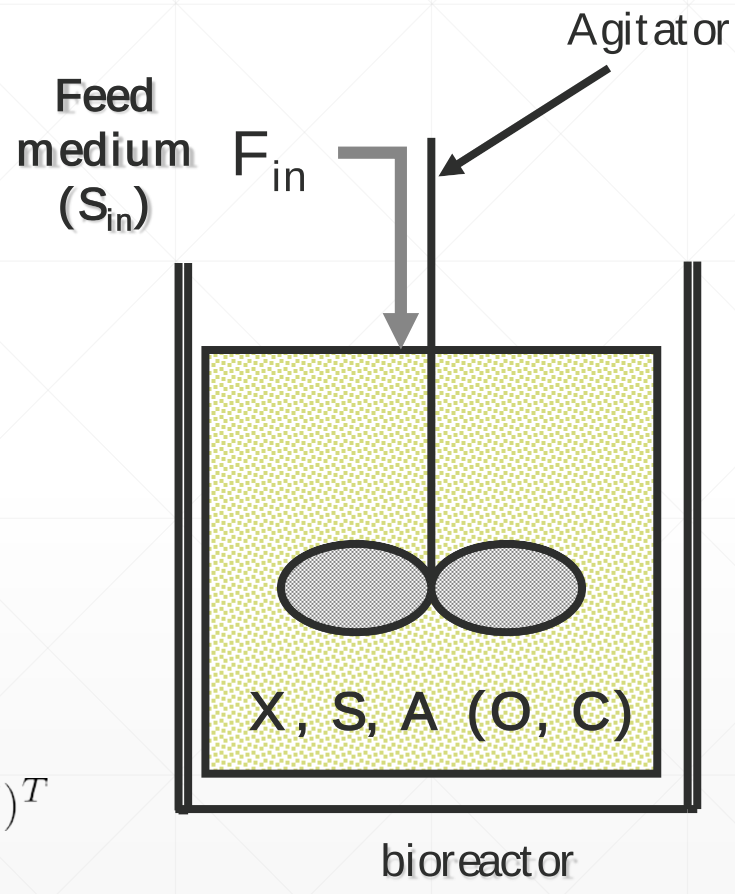
- q_s et $q_{s_{max}}$: Glucose consumption rate and its maximal value.
- q_{AC} Acetate consumption rate.

State space model

A mass balance modeling considering homogeneous well-stirred **fed-batch** reactor leads to:

$$
\begin{aligned}
\dot{X} &= (k_{x1}\mu_1 + k_{x2}\mu_2 + k_{x3}\mu_3)X - \frac{F_{in}}{V} X \\
\dot{S} &= -(\mu_1 + \mu_2)X - \frac{F_{in}}{V} (S - S_{in}) \\
\dot{A} &= (k_{A2}\mu_2 - \mu_3)X - \frac{F_{in}}{V} A \\
\dot{V} &= F_{in}
\end{aligned}
$$

- X, S, A, and V are respectively, the biomass, glucose, acetate concentrations, and the culture volume.
- F_{in} is the medium inlet feed-rate.
- The dynamics of O & C are not considered in this model.

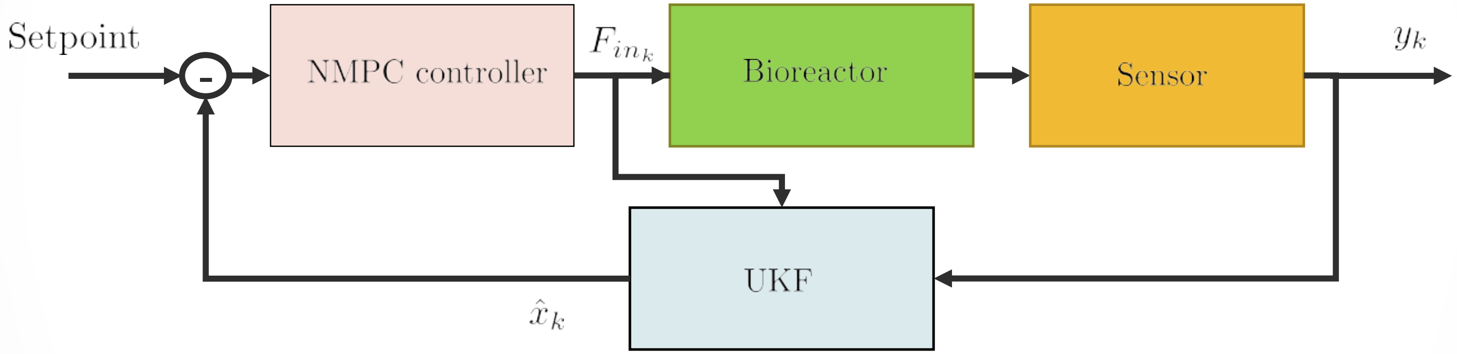


To sum up:

$$\dot{x} = f_0(x, F_{in}) \qquad x = (X \ S \ A \ V)^T$$

Or in a discrete form:

$$x_{k+1} = f(x_k, F_{in})$$



Objective:

Maximize the **biomass** growth

Our strategy

Develop a control law that **tracks a reference profile** in a fed-batch *E. coli* culture process using a nonlinear predictive control **(NMPC)** strategy coupled to an Unscented Kalman Filter **(UKF)** estimator.

Progress

1. Process model
2. **Model predictive control**
3. Unscented Kalman Filter
4. Simulation results
5. Conclusion and perspectives

Model predictive control

- Regulation of Biomass concentration X to a reference profile X^r while the feed-rate F_{in} to track a specified feed-rate profile $F_{in\ ref}$.
- The optimization problem considers minimizing the following quadratic cost function over a finite horizon N, applying only the first control value according to the receding horizon strategy:

$$\min_{\hat{x}_k \ldots \hat{x}_{k+N-1}, F_{in_k} \ldots F_{in_{k+N-1}}} \quad \sum_{i=1}^{N} \|\hat{X}_{k+i} - X^r_{k+i}\|^2 + \lambda \sum_{i=1}^{N} \|F_{in_{k+i}} - F_{in_{ref_{k+i}}}\|^2$$

$$\text{s.t} \begin{cases} \hat{X}_{k+1} = Hf(\hat{x}_k, F_{in_k}) \\ \vdots \\ \hat{X}_{k+N} = Hf(\hat{x}_{k+N-1}, F_{in_{k+N-1}}) \\ \hat{x}_k \geq 0 \quad \forall k \in \mathbb{N} \\ F_{max} \geq F_{in_k} \geq 0 \quad \forall k \in \mathbb{N} \end{cases}$$

$H = [1 \quad 0 \quad 0 \quad 0]$

$\hat{X}$ is the predicted output

$\hat{x}$ is the predicted state vector

N is the prediction horizon

λ is the control weighting factor

Nonlinear model predictive control applied to E. Coli culture

- Difficulties when solving this problem:

  - Such control requires discretization of the model with a small sampling time, so that the discretized model, remains significant compared to the continuous one.

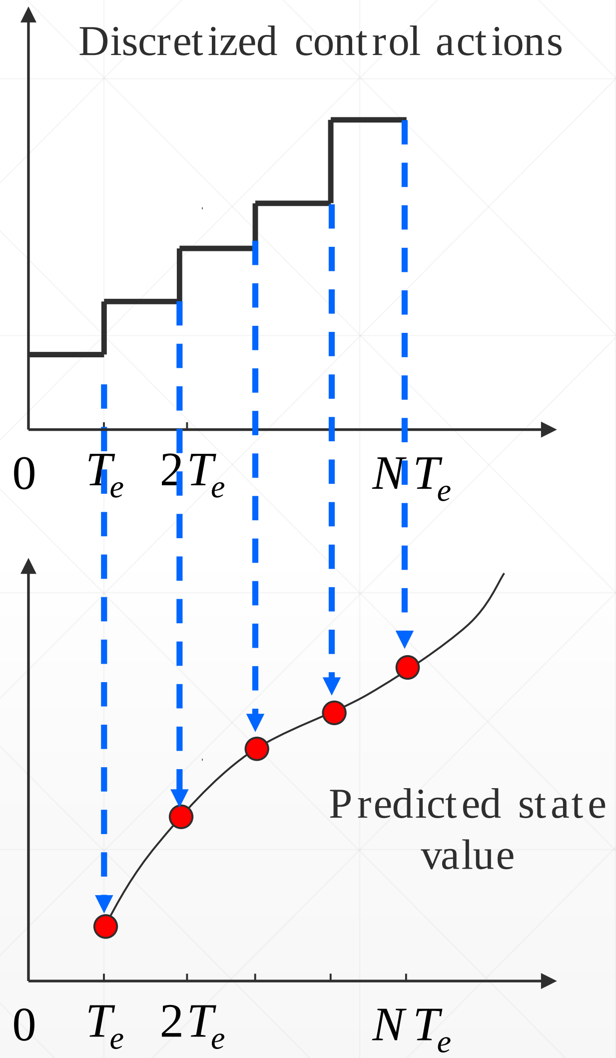
    ➔ This leads to a sampling time much too short compared to the time response of the system.

  - The presence of nonlinear constraints increases the on-line computation time when solving the optimization problem.

- How can we avoid these difficulties?

Nonlinear model predictive control applied to E. Coli culture

- The idea is to move the classical NMPC formulation into a nonlinear programming (NLP) problem.
- The resulting strategy is based on the Control Vector Parametrization (CVP) technique:
  - Only the control actions are discretized with respect to time. The sampling time can thus be chosen much larger than in the case of classical discretization.
  - A piecewise constant approximation of such control actions is considered for the continuous-time computation of the predicted state vector, without discretizing the state variables.



Nonlinear model predictive control applied to E. Coli cultures

- The formulation of the NMPC problem becomes:

$$\min_{\hat{x}_k \ldots \hat{x}_{k+N-1}, F_{in_k} \ldots F_{in_{k+N-1}}} \sum_{i=1}^{N} \|\hat{X}_{k+i} - X^r_{k+i}\|^2 + \lambda \sum_{i=1}^{N} \|F_{in_{k+i}} - F_{in_{ref_{k+i}}}\|^2$$

$$\text{s.t} \quad F_{max} > F_{in_k} > 0 \quad \forall k \in \mathbb{N}$$

- The number of constraints is drastically decreased.

- The tuning parameters are N and λ
  - N chosen as a trade-off between computation burden and performance (anticipation effect)
  - λ chosen as a trade-off between the control smoothness and performance (accuracy)

Progress

1. Process model
2. Model predictive control
3. **Unscented Kalman Filter**
4. Simulation results
5. Conclusion and perspectives

UKF estimation

- Goal : on-line estimation of the acetate and the glucose concentration based on the biomass measurement
- Nonlinear dynamics: Kalman filtering.
- Unscented Kalman filter:
  - Derivative-free
  - Propagation of the nonlinear dynamics through *Sigma* points
  - Estimated state is given by linear regression of these points.
  - 3 steps:
    - Calculate the Sigma points
    - Prediction
    - Update

Unscented Kalman Filtering Algorithm (1/2)

- Consider the nonlinear discrete system:

$$x_{k+1} = f(x_k, u_k) + v_k$$
$$y_k = h(x_k) + w_k$$

$$v_k \sim \mathrm{N}(0, Q), \quad w_k \sim \mathrm{N}(0, R)$$
$$\hat{x}_0 = E[x_0], \; P_0 = E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]$$
$$n = \dim(x)$$

- Step 1: Selection of the Sigma points:

$$(\mathcal{X}_{k-1})_0 = \hat{x}_{k-1}$$
$$(\mathcal{X}_{k-1})_i = \hat{x}_{k-1} + \gamma.\left(\sqrt{P_{k-1}}\right)_i, \; i = 1, \ldots, n,$$
$$(\mathcal{X}_{k-1})_i = \hat{x}_{k-1} - \gamma.\left(\sqrt{P_{k-1}}\right)_{i-n}, \; i = n+1, \ldots, 2n$$

$$\gamma = \sqrt{n + \lambda_u}$$
$$\lambda_u = \alpha^2(n + \kappa) - n$$

- Step 2: Prediction

$$\mathcal{X}_{k|k-1} = f[\mathcal{X}_{k-1}, u_{k-1}]$$
$$\mathcal{Y}_{k|k-1} = h[\mathcal{X}_{k|k-1}]$$

$$\Rightarrow$$

$$\hat{x}_k^- = \sum_{i=0}^{2n} W_i^{(m)} \mathcal{X}_{i,k|k-1}$$
$$\hat{y}_k^- = \sum_{i=0}^{2n} W_i^{(m)} \mathcal{Y}_{i,k|k-1}$$

$$W_0^{(m)} = \frac{\lambda_u}{n+\lambda_u}, \; W_0^{(c)} = \frac{\lambda_u}{n+\lambda_u} + 1 - \alpha^2 + \beta$$
$$W_i^{(m)} = W_i^{(c)} = \frac{1}{2(n+\lambda_u)}$$

Unscented Kalman Filtering Algorithm (2/2)

- Step 2: Prediction

$$P_k^- = \sum_{i=0}^{2n} W_i^{(c)} [\mathcal{X}_{i,k|k-1} - \hat{x}_k^-][\mathcal{X}_{i,k|k-1} - \hat{x}_k^-]^T + Q$$

$$P_{\tilde{y}_k\tilde{y}_k} = \sum_{i=0}^{2n} W_i^{(c)} [\mathcal{Y}_{i,k|k-1} - \hat{y}_k^-][\mathcal{Y}_{i,k|k-1} - \hat{y}_k^-]^T + R$$

$$P_{y_k x_k} = \sum_{i=0}^{2n} W_i^{(c)} [\mathcal{X}_{i,k|k-1} - \hat{x}_k^-][\mathcal{X}_{i,k|k-1} - \hat{x}_k^-]^T$$

- Step 3: Update

$$\mathcal{K}_k = P_{y_k x_k} P_{\tilde{y}_k\tilde{y}_k}^{-1}$$

$$\hat{x}_k = \hat{x}_k^- + \mathcal{K}_k (y_k - \hat{y}_k^-)$$

$$P_k = P_k^- - \mathcal{K}_k P_{\tilde{y}_k\tilde{y}_k} \mathcal{K}_k^T$$

UKF tuning parameters : **covariance matrices $\boldsymbol{Q}$, $\boldsymbol{R}$, $\boldsymbol{P_0}$**, UKF parameters α, β, κ

Progress

1. Process model
2. Model predictive control
3. Unscented Kalman Filter
4. **Simulation results**
5. Conclusion and perspectives

Reference trajectory

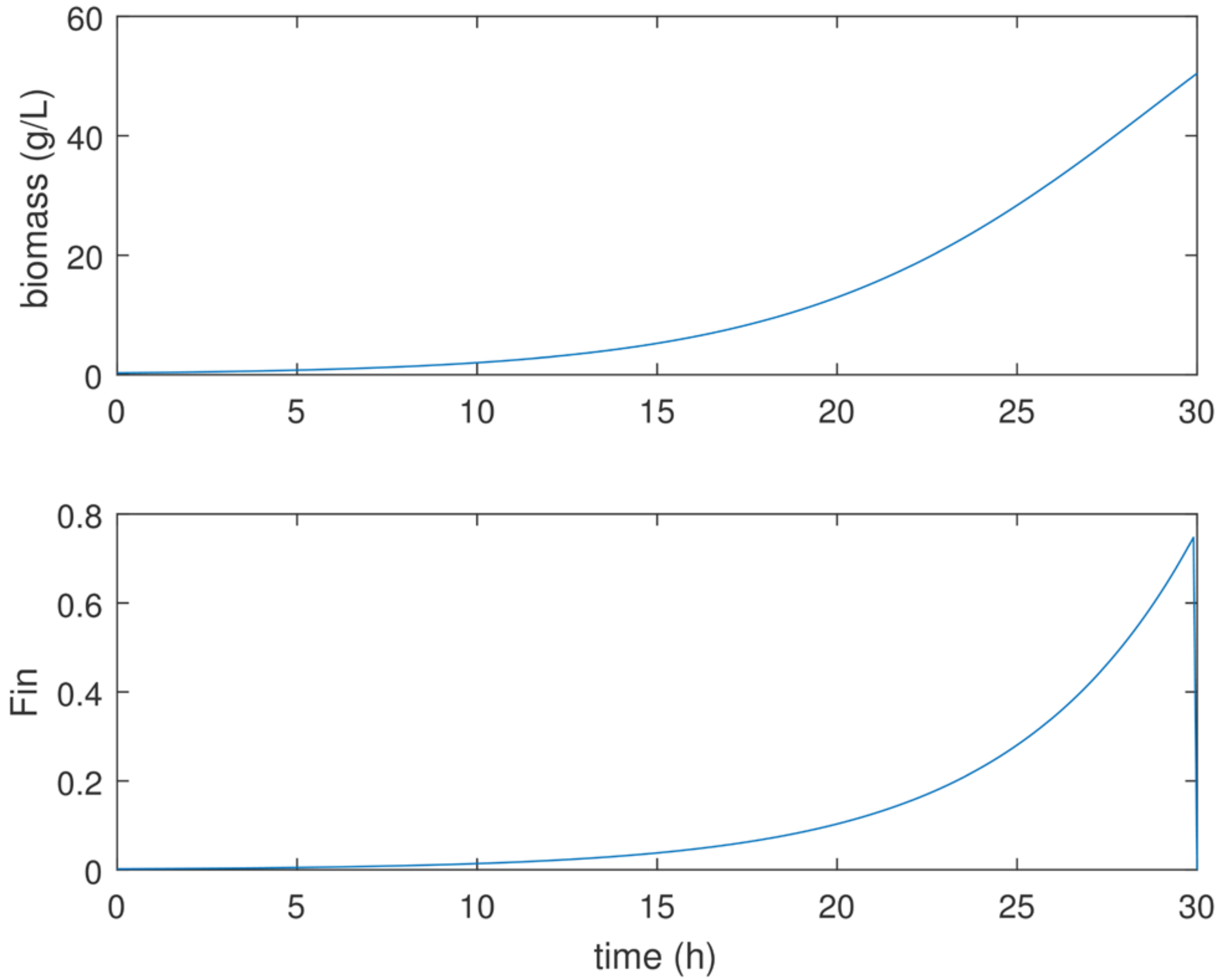
Initial conditions for reference profile

Variable	Value	Unit
X_0	0.3	g/L
S_0	0	g/L
A_0	0.1	g/L
V_0	3.15	L

Exponential feeding profile: [1]

$$F_{in} = F_0 e^{\eta(t-t_0)}$$

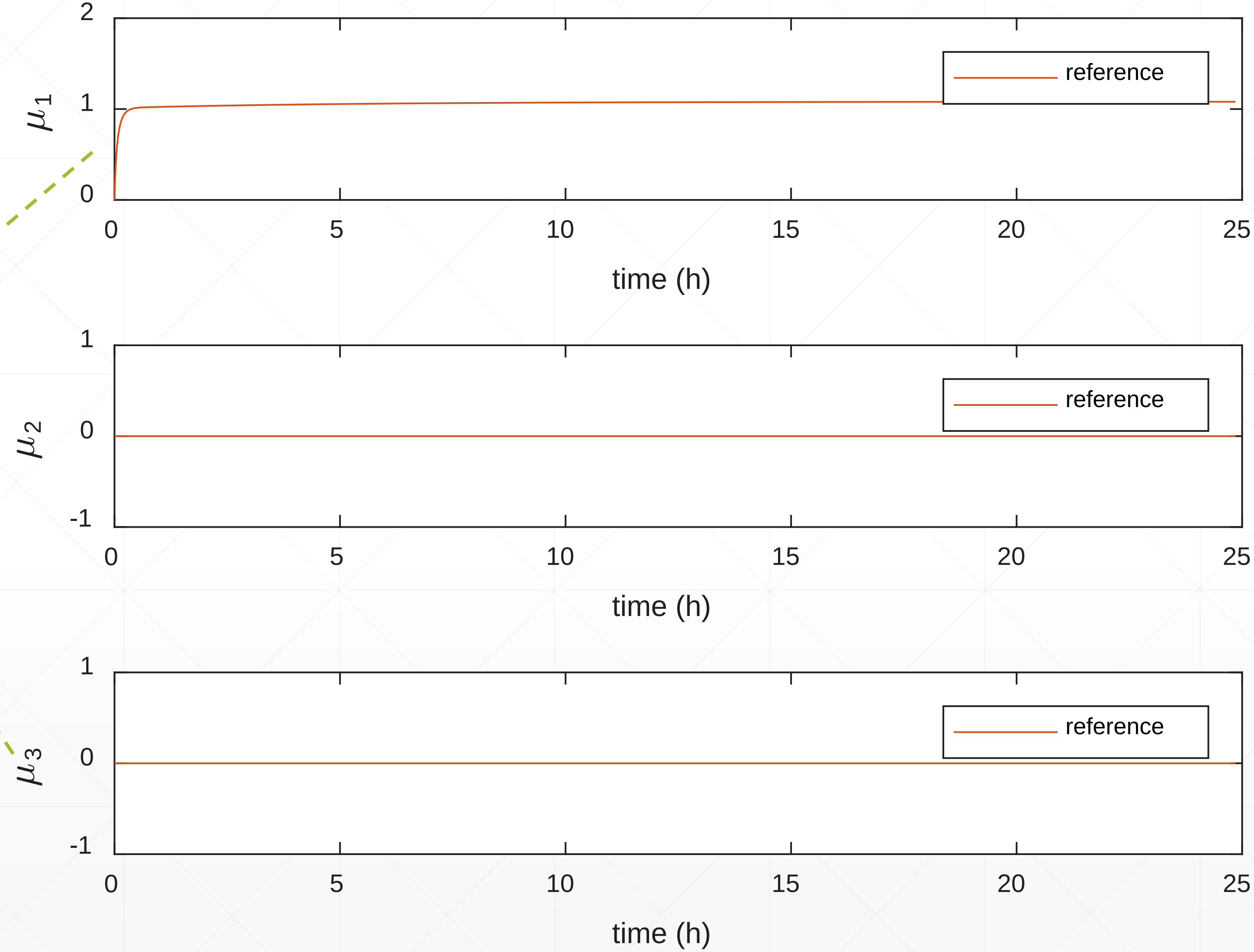
$F_0 = 4.8e^{-04} L/h$ $\eta = 0.2$



[1] C. Retamal, L. Dewasme, A.-L. Hantson, A. Vande Wouwer *Parameter estimation of a dynamic model of Escherichia coli fed-batch cultures*. Biochemical Engineering Journal

Reference trajectory

$$k_{s1}S + k_{o1}O \xrightarrow{\mu_1 X} k_{x1}X + k_{c1}C$$
$$k_{s2}S + k_{o2}O \xrightarrow{\mu_2 X} k_{x1}X + k_{A2}A + k_{c2}C$$
$$k_{A3}A + k_{o3}O \xrightarrow{\mu_3 X} k_{x3}X + k_{c3}C$$



UKF & NMPC : performance test

Estimator initial conditions:

Variable	Value	Unit
$\hat{X}_0$	$X_0 + 50\%$	g/L
$\hat{S}_0$	$S_0 + 10\%$	g/L
$\hat{A}_0$	$A_0 + 10\%$	g/L
$\hat{V}_0$	$V_0 + 10\%$	L

Estimator parameters:

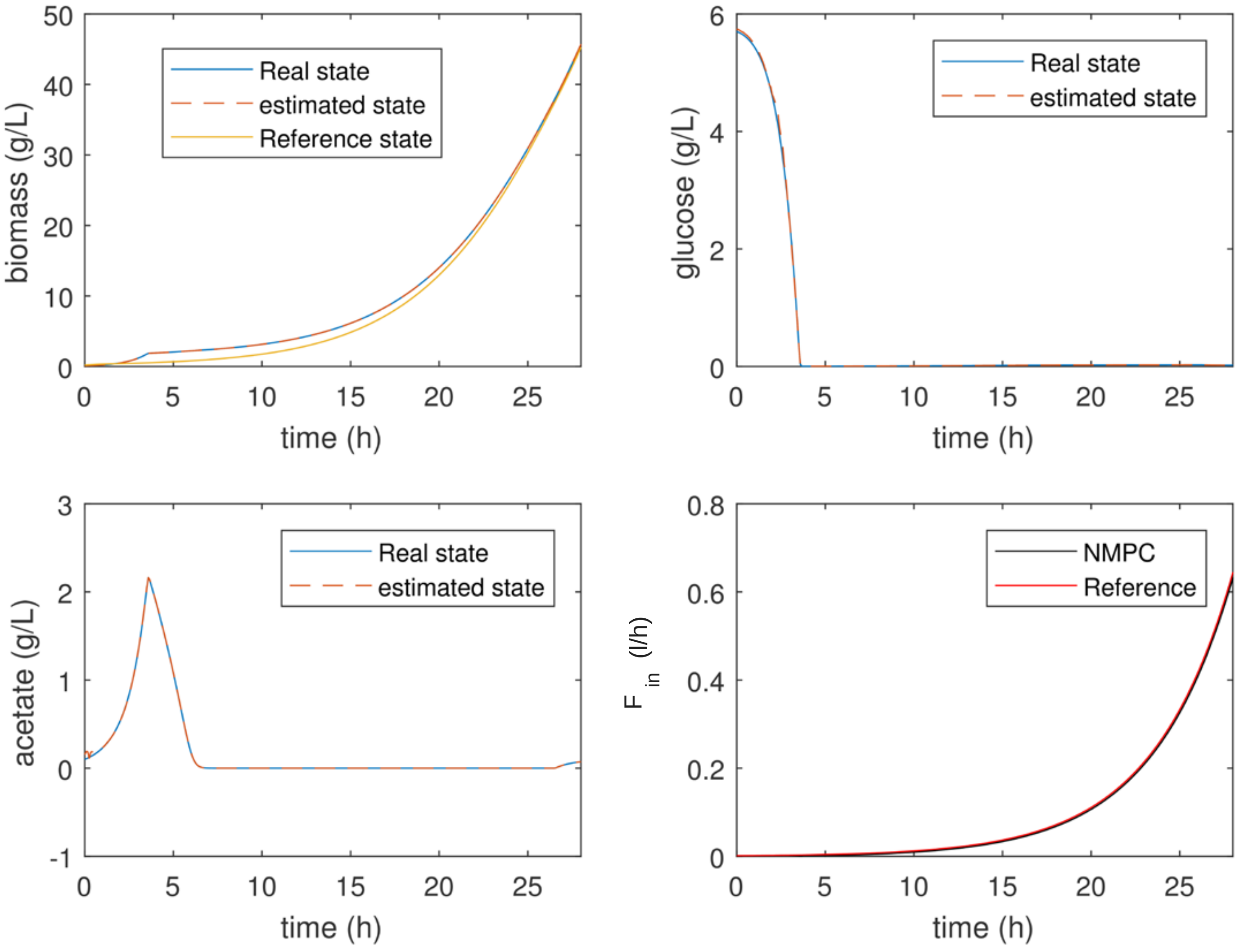
$$Q = \begin{pmatrix} 10^{-2} & 0 & 0 & 0 \\ 0 & 10^{-2} & 0 & 0 \\ 0 & 0 & 10^{-2} & 0 \\ 0 & 0 & 0 & 10^{-2} \end{pmatrix} \qquad R = 10^{-2}(g/L)^2$$

$$P_0 = \begin{pmatrix} 5\ 10^{-3} & 0 & 0 & 0 \\ 0 & 5\ 10^{-2} & 0 & 0 \\ 0 & 0 & 1\ 10^{-2} & 0 \\ 0 & 0 & 0 & 10^{-1} \end{pmatrix}$$

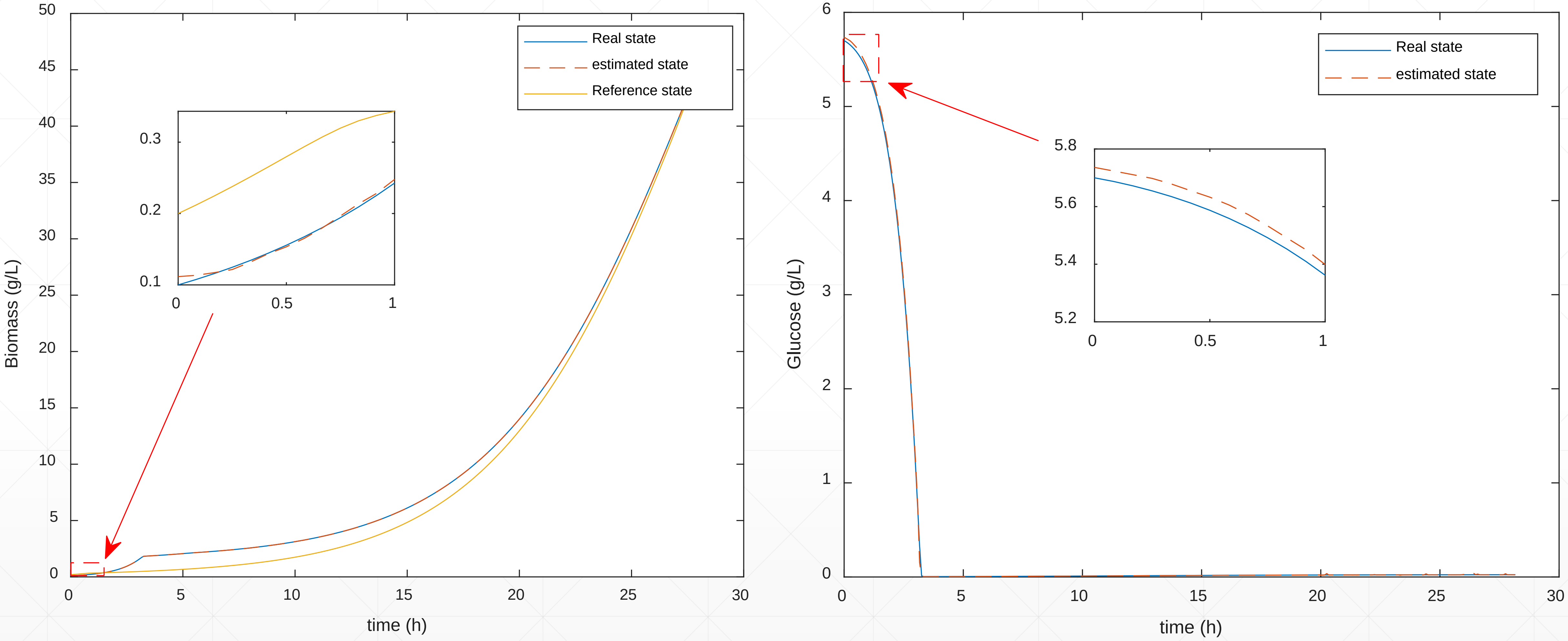
α	10^{-2}
β	2
κ	0

NMPC parameters :

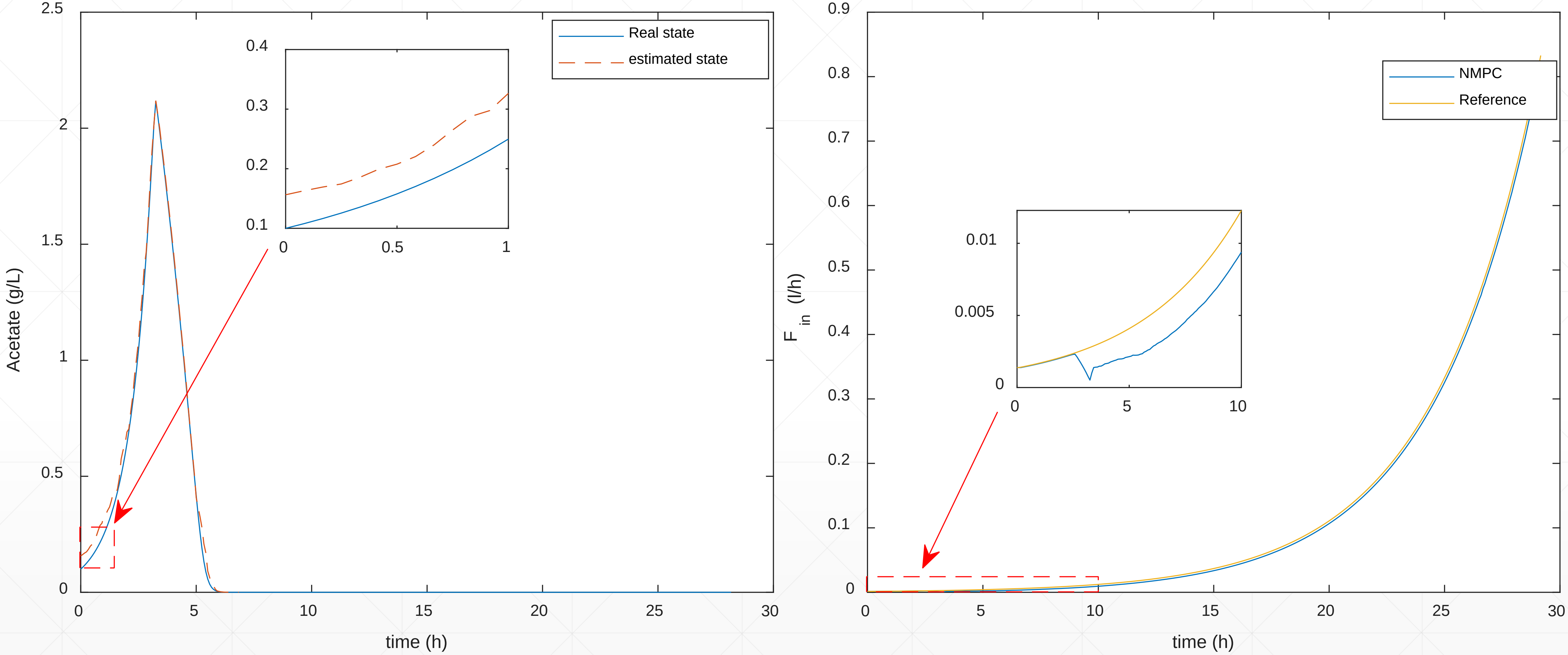
Variable	Value	Unit
N	10	–
T_e	5	min
λ	0.8	–



Good performance of the NMPC and UKF algorithms



Biomass (g/L)
time (h)
Real state
estimated state
Reference state
Glucose (g/L)
time (h)
Real state
estimated state



After a transitory phase, the estimator and the controller give a good performance

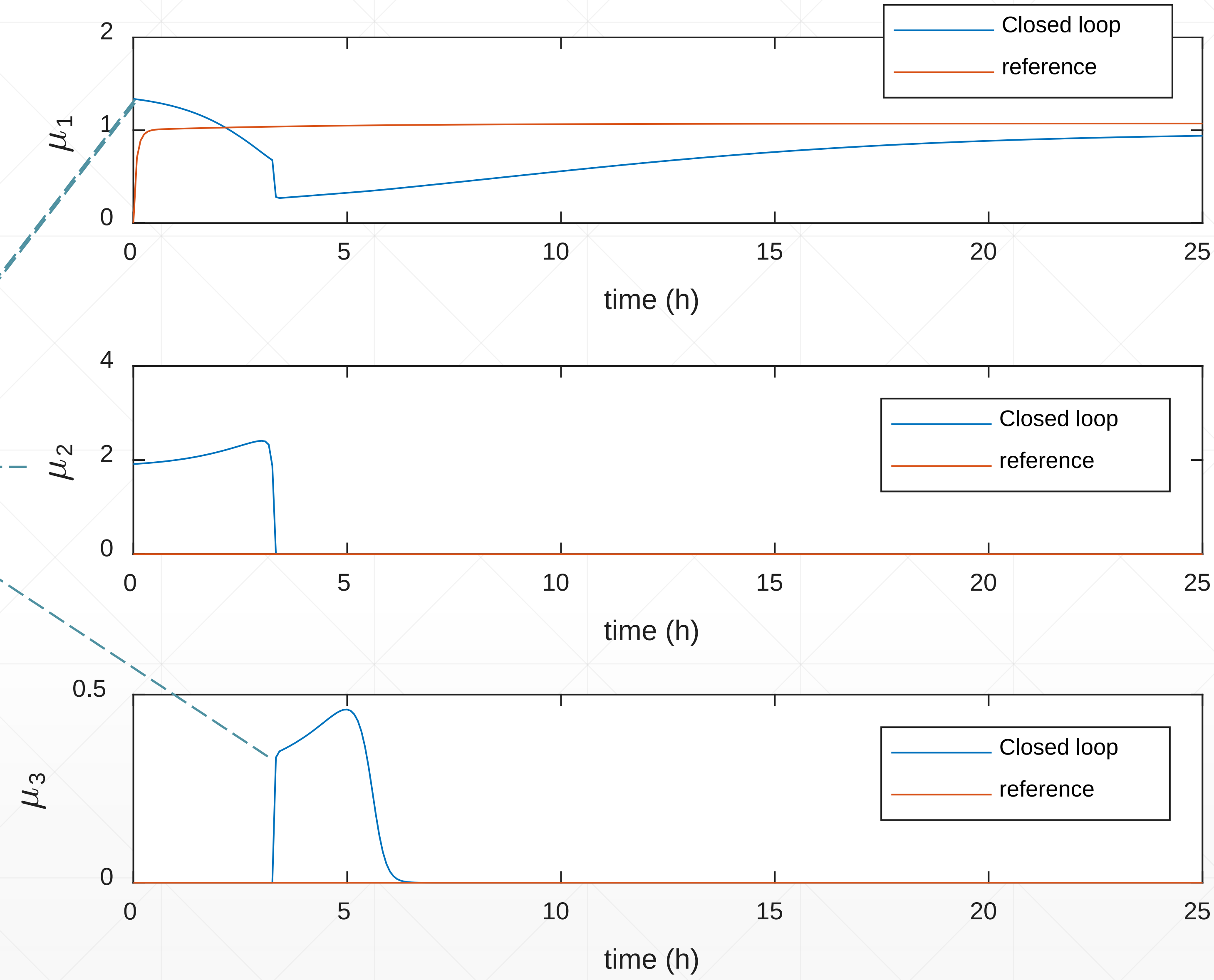
$$k_{s1}S + k_{o1}O \xrightarrow{\mu_1 X} k_{x1}X + k_{c1}C$$

$$k_{s2}S + k_{o2}O \xrightarrow{\mu_2 X} k_{x1}X + k_{A2}A + k_{c2}C$$

$$k_{A3}A + k_{o3}O \xrightarrow{\mu_3 X} k_{x3}X + k_{c3}C$$

Respiro-fermentative regime

Respirative regime



Preliminary robustness test: model mismatch

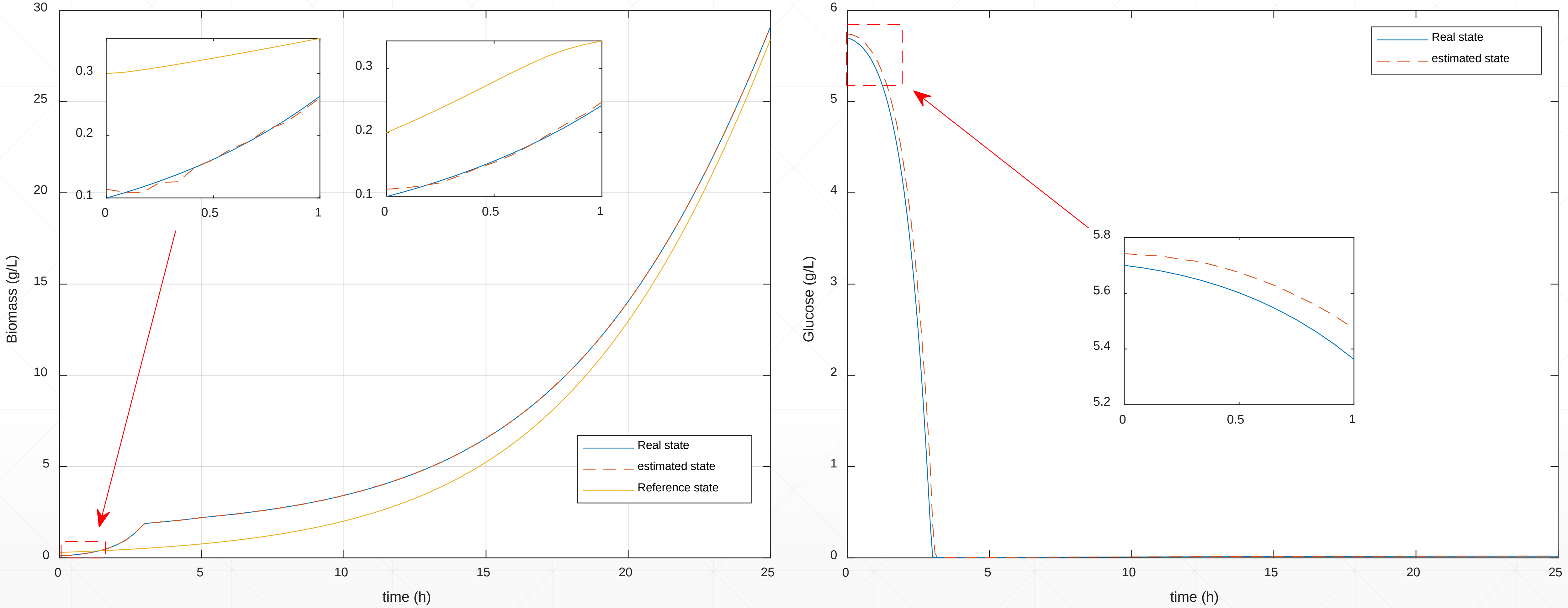
- Three parameters were altered to test the robustness of the control and estimation algorithms
- These parameters appear in the following equations which represent the specific growth rate expressions in the respiro-fermentative regime:

$q_{s_{max}}$ $q_{O_{max}}$ K_{iA}

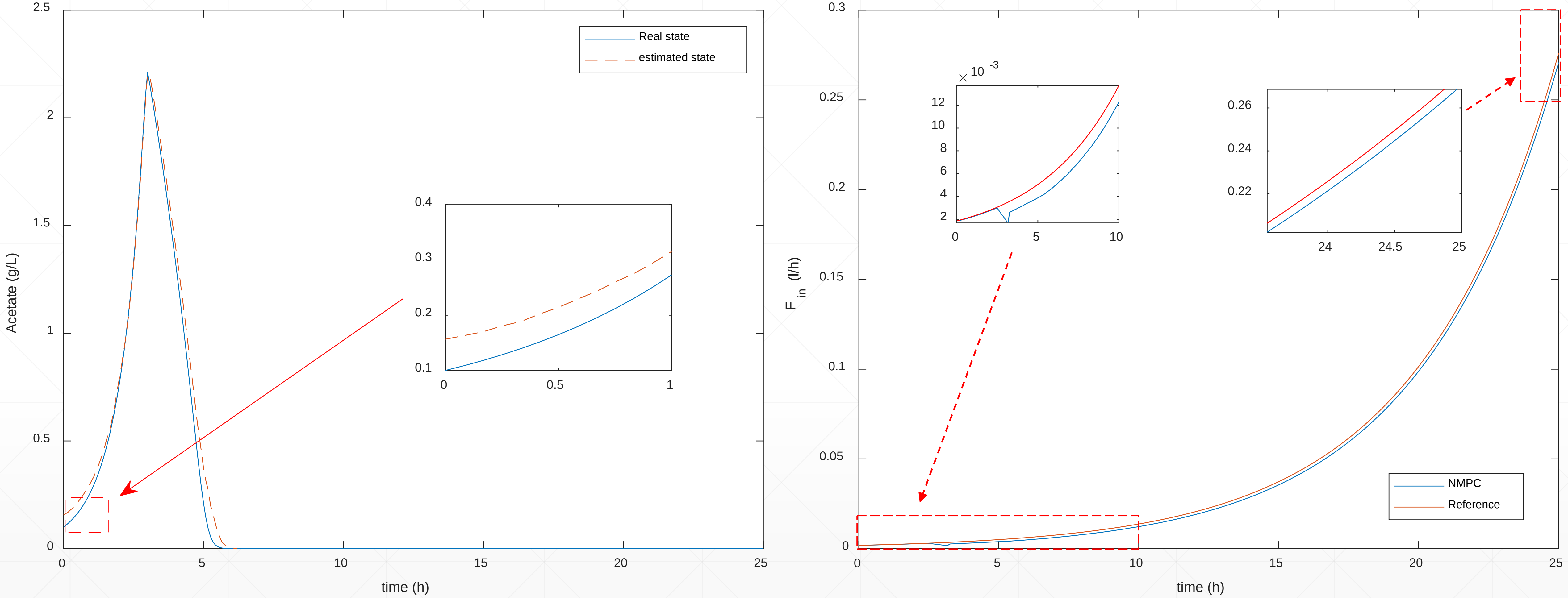
$$q_s = q_{s_{max}} \frac{S}{K_s + S} \qquad q_{s_{crit}} = \frac{q_{O_{max}}}{k_{OS}} \frac{K_{iA}}{K_{iA} + A}$$

$$\mu_1 = \frac{min(q_s, q_{s_{crit}})}{k_{s1}} \qquad \mu_2 = \frac{max(0, q_s - q_{s_{crit}})}{k_{s2}}$$

- The parameters q_s *and* q_{smax} *were* altered by 20%
- The parameter K_{iA} was altered by 10%



Biomass (g/L)
time (h)
Real state
estimated state
Reference state
Glucose (g/L)
time (h)
Real state
estimated state



Good performance of the NMPC and UKF algorithms

Progress

1. Process model
2. Model predictive control
3. Unscented Kalman Filter
4. Simulation results
5. **Conclusion and perspectives**

Conclusion

- A macroscopic model describing the metabolism of *E. coli* was presented
- Development of the NMPC controller to track a biomass and a feeding reference trajectory.
- State estimation using an Unscented Kalman Filter presents advantages due to the nonlinearity of the system
- Simulation results show the efficiency of the proposed strategy (NMPC controller coupled to an unscented kalman filter).

Perspectives

- Determination of an optimal trajectory according to the bottleneck theory
- Analyze further the performance and robustness of the proposed strategy
- Experimental validation of the proposed control strategy on an *E. coli* culture.
- Online optimization of the biomass growth : determination of the appropriate criterion (growth rate, biomass concentration, ..)
- Robustification of the control and estimation strategies w.r.t. model mismatch and measurement errors