Prof. Davide Manca – Politecnico di Milano

Dynamics and Control of Chemical Processes

Solution to Lab #5

Model identification

POLITECNICO
MILANO 1863

Model identification

In defining a black-box model, the output data (y) of the system are calculated from the input data (u) and from the past history (y_{old}, u_{old}):

$$\mathbf{y} = \mathbf{f}\left(\mathbf{y}_{old}, \mathbf{u}_{old}\right)$$

In general, in order to have a model as close as possible to the real system some adaptive parameters (p) are introduced:

$$\mathbf{y} = \mathbf{f}\left(\mathbf{y}_{old}, \mathbf{u}_{old}, \mathbf{p}\right)$$

It is also possible to introduce in the model the error (e) that is defined as y_{real}-y:

$$\mathbf{y} = \mathbf{f}\left(\mathbf{y}_{old}, \mathbf{u}_{old}, \mathbf{e}_{old}, \mathbf{p}\right)$$

Model identification

The system to be identified has the following structure:

$$\mathbf{y}(t)=\mathbf{f}\Big[y_1(t-1),\dots,y_1\left(t-n_{y_1}\right),\dots,y_r(t-1),\dots,y_r\left(t-n_{y_r}\right),$$
$$u_1(t-1),\dots,u_1\left(t-n_{u_1}\right),\dots,u_m(t-1),\dots,u_m\left(t-n_{u_m}\right),$$
$$e_1(t-1),\dots,e_1\left(t-n_{e_1}\right),\dots,e_r(t-1),\dots,e_r\left(t-n_{e_r}\right)\Big]$$

The vector $\boldsymbol{\varphi}$ is the vector of the regressors:

$$\boldsymbol{\varphi}(t)=\Big[y_1(t-1),\dots,y_1\left(t-n_{y_1}\right),\dots,y_r(t-1),\dots,y_r\left(t-n_{y_r}\right),$$
$$u_1(t-1),\dots,u_1\left(t-n_{u_1}\right),\dots,u_m(t-1),\dots,u_m\left(t-n_{u_m}\right),$$
$$e_1(t-1),\dots,e_1\left(t-n_{e_1}\right),\dots,e_r(t-1),\dots,e_r\left(t-n_{e_r}\right)\Big]^T$$

Model identification

The function **f**, by means of the parameters **p**, maps the vector of the regressors in the output variables **y:**

$$\mathbf{y}(t) = \mathbf{f}\left[\boldsymbol{\varphi}(t), \mathbf{p}\right]$$

The simplest function **f** is:

$$\mathbf{y}(t) = \underbrace{\mathbf{p}}_{\text{Row vector}} \times \underbrace{\boldsymbol{\varphi}(t)}_{\text{Column vector}}$$

Mathematical models may have scalar, vector or mixed structure:

- **SISO**: Single Input – Single Output
- **MISO**: Multiple Input – Single Output
- **MIMO**: Multiple Input – Multiple Output

Identification procedure

1. Determination of the system limits and necessary variables
2. Design of experiments
3. Selection of the model structure
4. Parameters evaluation
5. Simulation and validation

Identification procedure

1. **Definition of the system limits and necessary variables**
    - The exact number of input (u) and output (y) variables is defined.
    - The variability range of the variables is identified to create a sutiable sampling domain for the next identification step.

2. **Design of experiments**
    - Once the variables are identified, the sampling frequency is assigned
    - All the input variables must be disturbed

Identification procedure

3. Selection of the model structure

- We have to define:
    - The length of the regressors vector (see the following point)
    - The order of the model respect to every variable
    - The linearity or non-linearity respect to the regressors and the parameters

4. Parameters evaluation

- We have to choose the numerical algorithm for the evaluation of the model parameters
- The models may be classified as:
    - Deterministic (error minimization)
    - Stochastic (maximum likelihood method)

Identification procedure

5. Simulation and validation

- Once the model is identified, it is required to test its predictive capability and its goodness by using a set of not formerly used data
- The validation procedure is based of a validation data set (cross-validation set), properly chosen *a priori* and kept separated from the learning set

Disturbance sequence generation

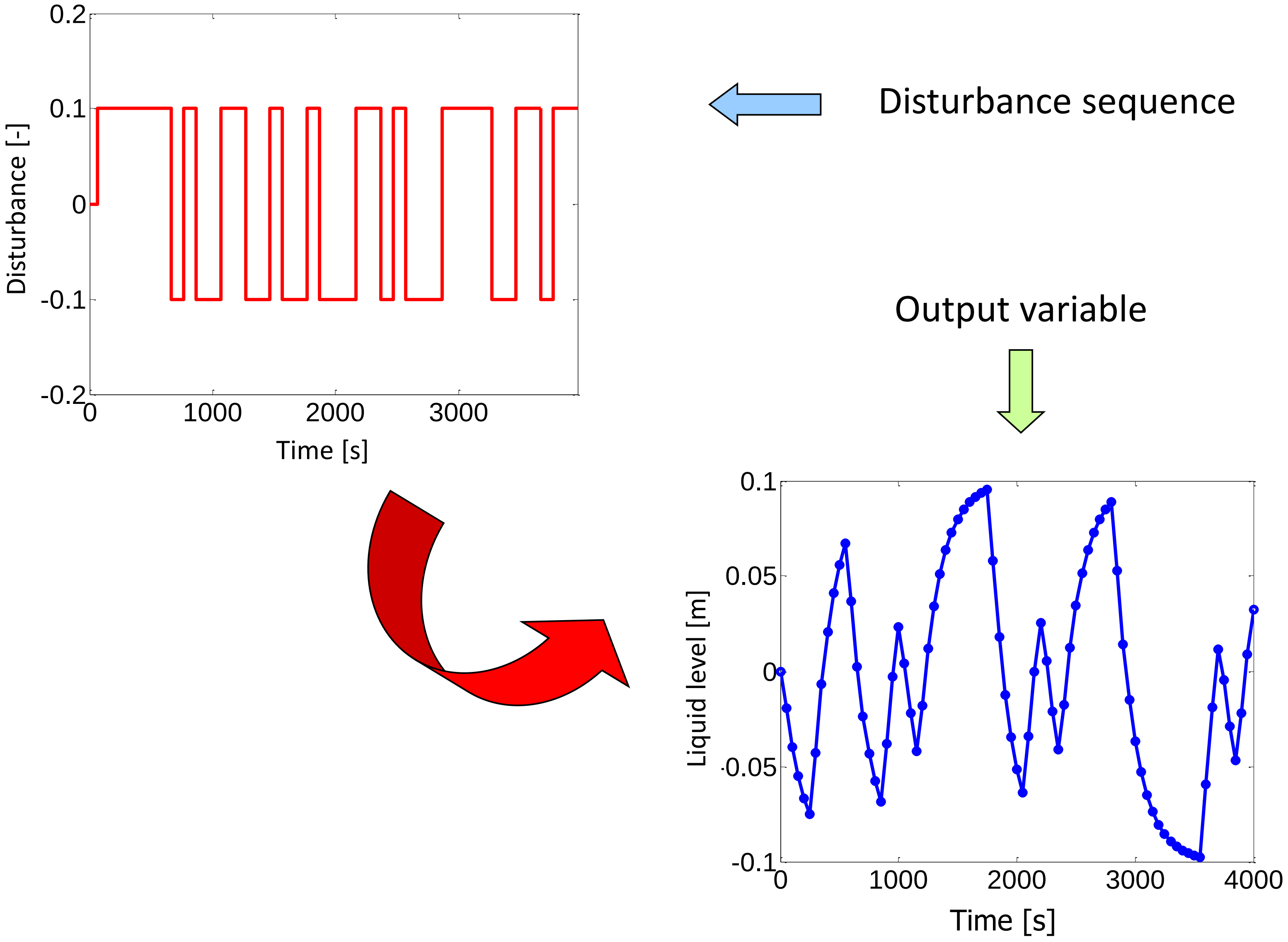
- The input-output data collection for the identification and validation procedures is obtained by disturbing the process input variables.

- The **PRBS** (Pseudo Random Binary Sequence) method is used:
  - Two bounds are chosen, uMIN, uMAX, for the variability range of the disturbed variable u
  - The variable quantity is varied randomly. The variable can assume only the bound values (*i.e.* uMIN and uMAX)
  - The corresponding output vector is measured

PRBS sequences

1. Input variables can assume only two values, equal in amplitude but with opposite sign, ±Δu, respect to the stationary conditions
2. The shift from the positive condition to the negative one, and vice-versa, is made randomly in order to give the sequence a kind of white noise behaviour (*i.e.* null average)
3. The disturbance on the input variables is made every n sampling times (t_s = sampling time)
4. Usually the range of $n \times t_s$ is equal to the 20% of the time needed by the system to end its transient
5. The amplitude of the disturbance Δu should be high enough to eliminate the measurement error due to the system noise

Disturbance sequences generation



Disturbance sequence

Output variable

Data pre-processing

- At the stage of field data collection it is possible to apply some appropriate **mathematical operators** able to damp the excessive oscillations (the moving average for example)
- It is possible to apply some high-cut, low-cut **filters** in order to remove sudden variations beyond the normal operating intervals
- It is possible to remove the so called **outliers** by means of appropriate techniques of statistical analysis
- **DETREND**: the average value is subtracted to the sampled data. By doing so, the sampled variables express the deviation from either the stationary conditions or the mean operating conditions. As a consequence, it is also possible to use the model (at the cost of lower quality results) even for other steady state conditions.

ARX models

Features

- The ARX model is linear both in the regressors and in the parameters
- As such, it is not able to describe different steady states
- By definition it cannot describe non-linear behaviours
- Its identification is quite simple
- The computational time for one prediction is extremely low

ARX – SISO models

SISO models:

$$y(t)+a_1\,y(t-1)+a_2\,y(t-2)+\ldots+a_{n_y}\,y\left(t-n_{n_y}\right)=$$
$$=b_1\,u(t-1)+b_2\,u(t-2)+\ldots+b_{n_u}\,u(t-n_u)$$

In order to make a prediction, n_y values of the dependent variable (*y*) and n_u values of the independent variable (*u*) are needed.

Example: evaluation based on 3 previous times for both the independent and dependent variables

$$y(t)+a_1\,y(t-1)+a_2\,y(t-2)+a_3\,y(t-3)=$$
$$=b_1\,u(t-1)+b_2\,u(t-2)+b_3\,u(t-3)$$

ARX – MIMO models

MIMO models:

Example: system characterized by:

- 3 independent variables (u)
- 2 dependent variables (y)
- 4 previous times

$$
\begin{aligned}
y(t) &+ a_1 y_1(t-1) + a_2 y_1(t-2) + a_3 y_1(t-3) + a_4 y_1(t-4) + \\
&+ a_5 y_2(t-1) + a_6 y_2(t-2) + a_7 y_2(t-3) + a_8 y_2(t-4) = \\
&= b_1 u_1(t-1) + b_2 u_1(t-2) + b_3 u_1(t-3) + b_4 u_1(t-4) + \\
&+ b_5 u_2(t-1) + b_6 u_2(t-2) + b_7 u_2(t-3) + b_8 u_2(t-4) + \\
&+ b_9 u_3(t-1) + b_{10} u_3(t-2) + b_{11} u_3(t-3) + b_{12} u_3(t-4)
\end{aligned}
$$

8 + 12 = 20 parameters

Parameters determination (SISO)

Given the following measurements on the real system

Different sets of parameters can be derived

$$\varphi = \left[y(t-1) \quad y(t-2) \quad \dots \quad y(t-n_y) \quad u(t-1) \quad u(t-2) \quad \dots \quad u(t-n_u) \right]^T$$

$$y(t)$$

$$y(t) = \mathbf{p} \times \varphi(t)$$

$$p_1 = \left[a_1 \quad a_2 \quad \dots \quad a_{n_y} \quad b_1 \quad b_2 \quad \dots \quad b_{n_u} \right]$$

First parameters set

$$\varphi = \left[y(t) \quad y(t-1) \quad \dots \quad y(t-n_y+1) \quad u(t) \quad u(t-1) \quad \dots \quad u(t-n_u+1) \right]^T$$

$$y(t+1)$$

$$y(t) = \mathbf{p} \times \varphi(t)$$

$$p_2 = \left[a_1 \quad a_2 \quad \dots \quad a_{n_y} \quad b_1 \quad b_2 \quad \dots \quad b_{n_u} \right]$$

Second parameters set

Parameter determination

Searching for the best parameters set, an objective function should be optimized

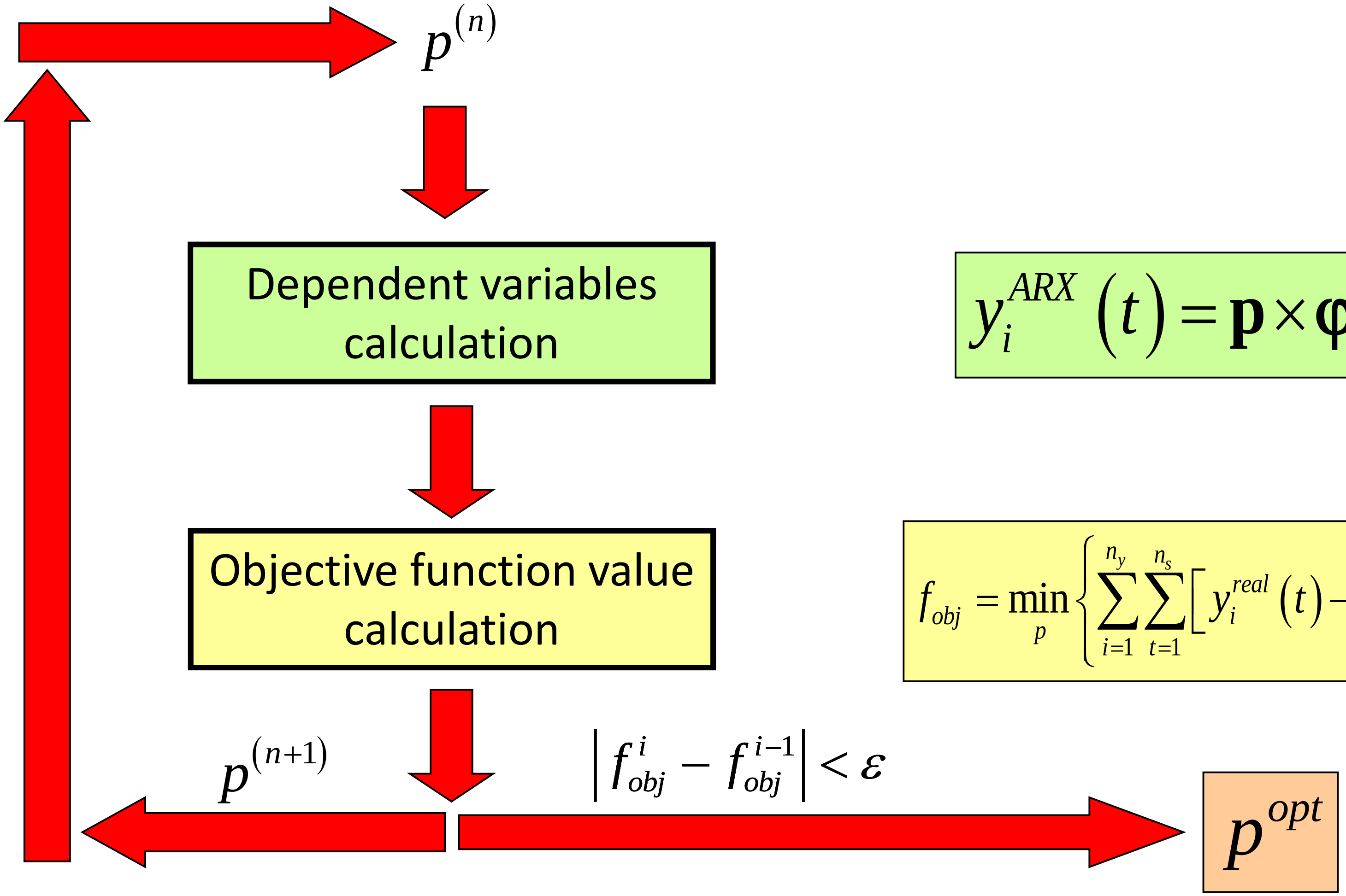
⇒ Least squares method

$$f_{obj} = \min_{p} \left\{ \sum_{i=1}^{n_y} \sum_{t=1}^{n_s} \left[y_i^{real}(t) - f_i\left(\boldsymbol{\varphi}(t), \mathbf{p}\right) \right]^2 \right\}$$

⇓

$$f_{obj} = \min_{p} \left\{ \sum_{i=1}^{n_y} \sum_{t=1}^{n_s} \left[y_i^{real}(t) - y_i^{ARX}(t) \right]^2 \right\}$$

Parameters determination



$p^{(n)}$

Dependent variables calculation

$$y_i^{ARX}(t) = \mathbf{p} \times \boldsymbol{\varphi}(t)$$

Objective function value calculation

$$f_{obj} = \min_{p} \left\{ \sum_{i=1}^{n_y} \sum_{t=1}^{n_s} \left[y_i^{real}(t) - y_i^{ARX}(t) \right]^2 \right\}$$

$p^{(n+1)}$

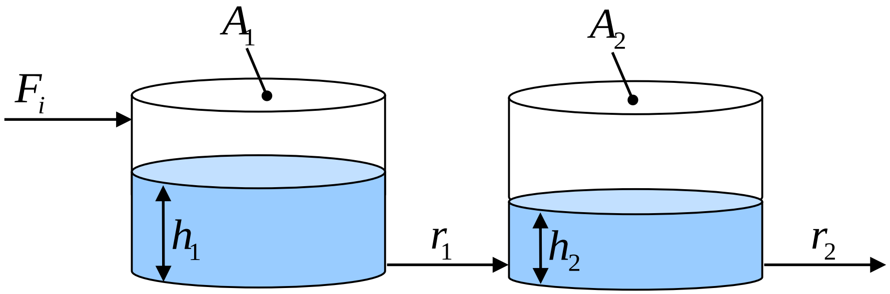
$\left| f_{obj}^{\,i} - f_{obj}^{\,i-1} \right| < \varepsilon$

p^{opt}

Practical

- Consider two interacting tanks
- Find the parameters of an ARX model considering that the independent variable is the inlet flowrate and the dependent variable is the level of the second tank
- In order to compute the dependent variable at time t, consider 2 old values for both the independent and the dependent variables
- Consider that the independent variable oscillates around the steady state value (which is assigned) of ±10%

System model



$$\begin{cases} A_1 \dfrac{dh_1}{dt} = F_i - \dfrac{h_1 - h_2}{r_1} \\ A_2 \dfrac{dh_2}{dt} = \dfrac{h_1 - h_2}{r_1} - \dfrac{h_2}{r_2} \end{cases}$$

Data: $F_i = 10 \ \mathrm{m^3/s}$

Tank 1:

$A_1 = 40 \ \mathrm{m^2}$

$r_1 = 0.9 \ \mathrm{s/m^2}$

Tank 2:

$A_2 = 30 \ \mathrm{m^2}$

$r_2 = 2.1 \ \mathrm{s/m^2}$

I.C.: $h_1(0) = h_1^{(s)}$

$h_2(0) = h_2^{(s)}$

Solution procedure

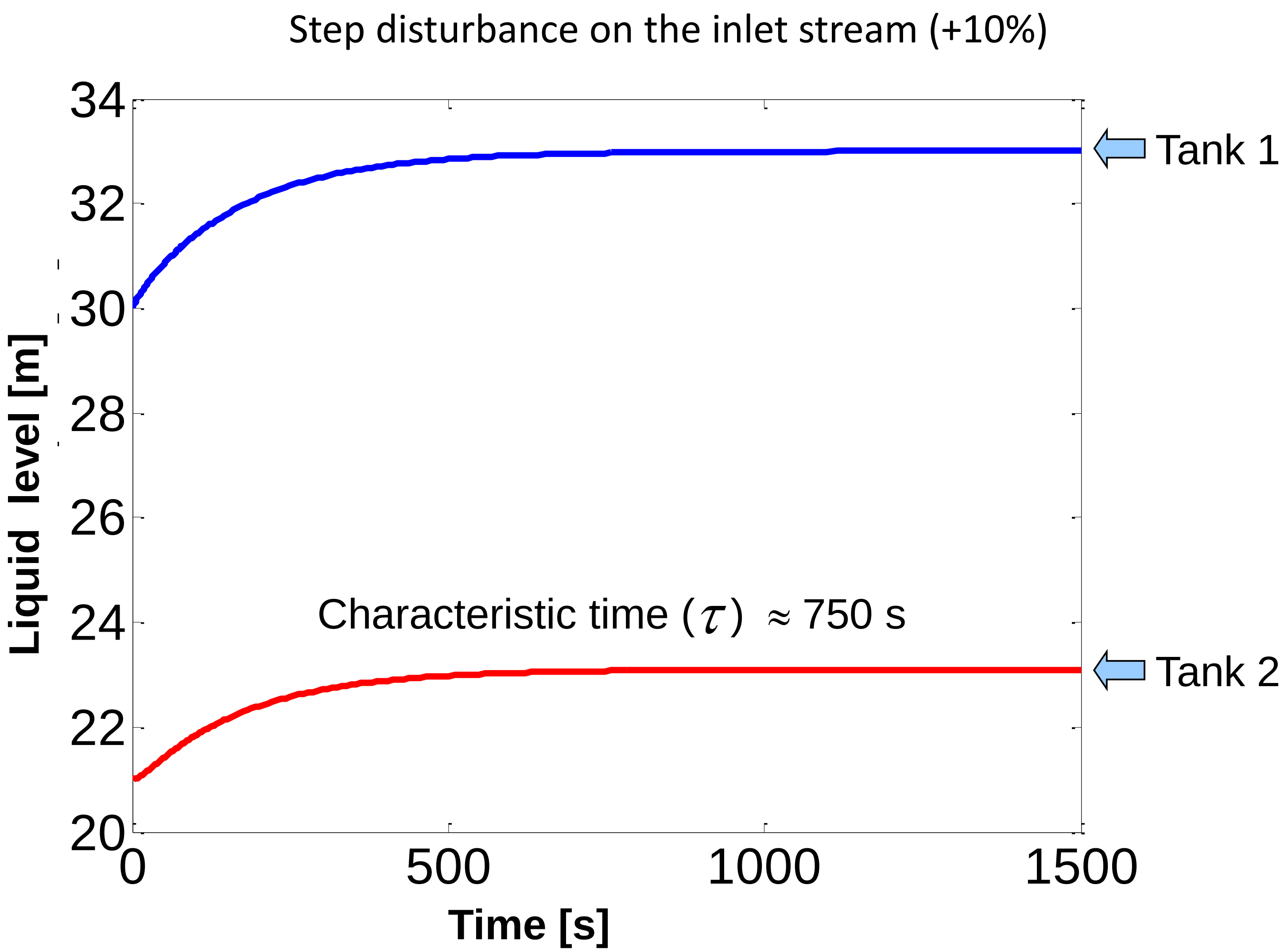
1. Determine the steady state conditions

$$\begin{cases} A_1 \dfrac{dh_1}{dt} = 0 = F_i - \dfrac{h_1 - h_2}{r_1} \\ A_2 \dfrac{dh_2}{dt} = 0 = \dfrac{h_1 - h_2}{r_1} - \dfrac{h_2}{r_2} \end{cases} \Rightarrow \begin{cases} h_1^{(s)} = \left(r_1 + r_2\right) F_i \\ h_2^{(s)} = r_2 F_i \end{cases}$$

2. Give a disturbance to the system in order to assess the characteristic time (for example +10% F_i)

3. Evaluate the system dynamics according to the PRBS method

Characteristic time assessment



Characteristic time (τ) $\approx$ 750 s

Determination of the times

- Interval between two disturbances(t_d):

$$t_d = 0.2\tau = 150 \text{ s}$$

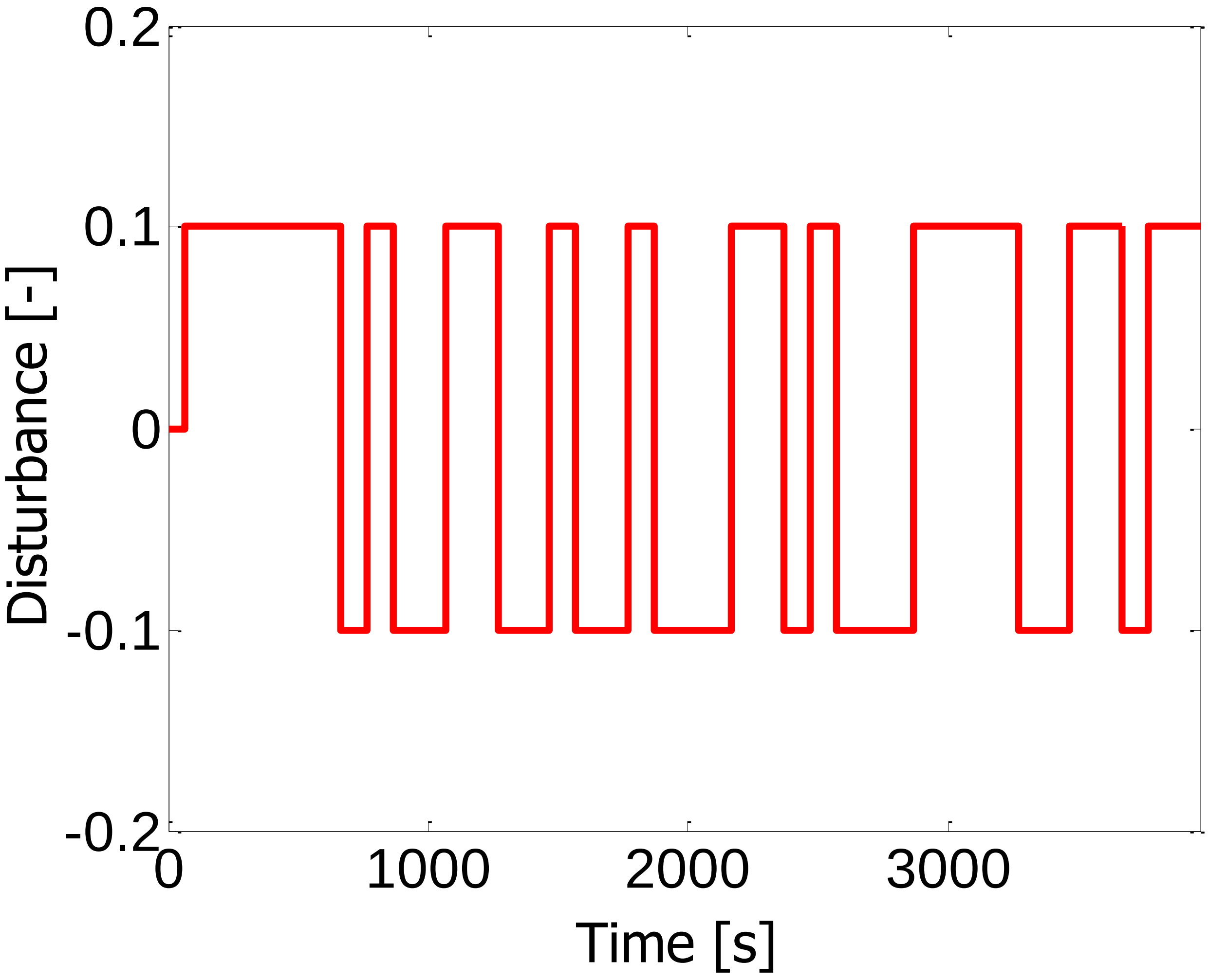
- Sampling time (t_s):

$$t_s = t_d / 3 = 50 \text{ s}$$

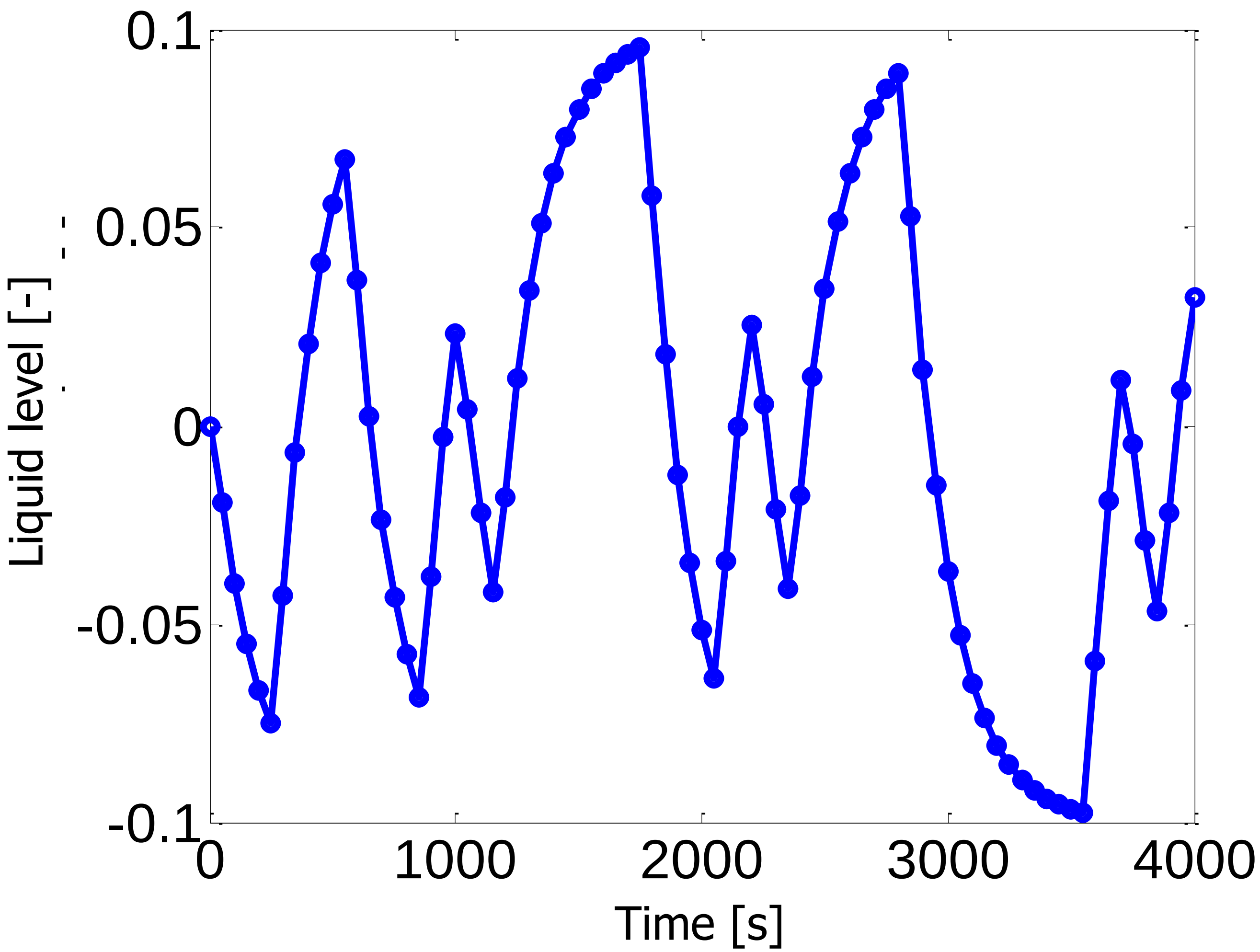
MatLab implementation

```
for i = 1:nSteps
   cont = cont + 1;
   if(cont == 3)
      randNum = rand();
      if(randNum <= 0.5)
           Fi = Fi0 * 1.1;
      else
           Fi = Fi0 * 0.9;
      end
    cont = 0;
    end
    … system dynamics evaluation
end
```

Disturbance sequence

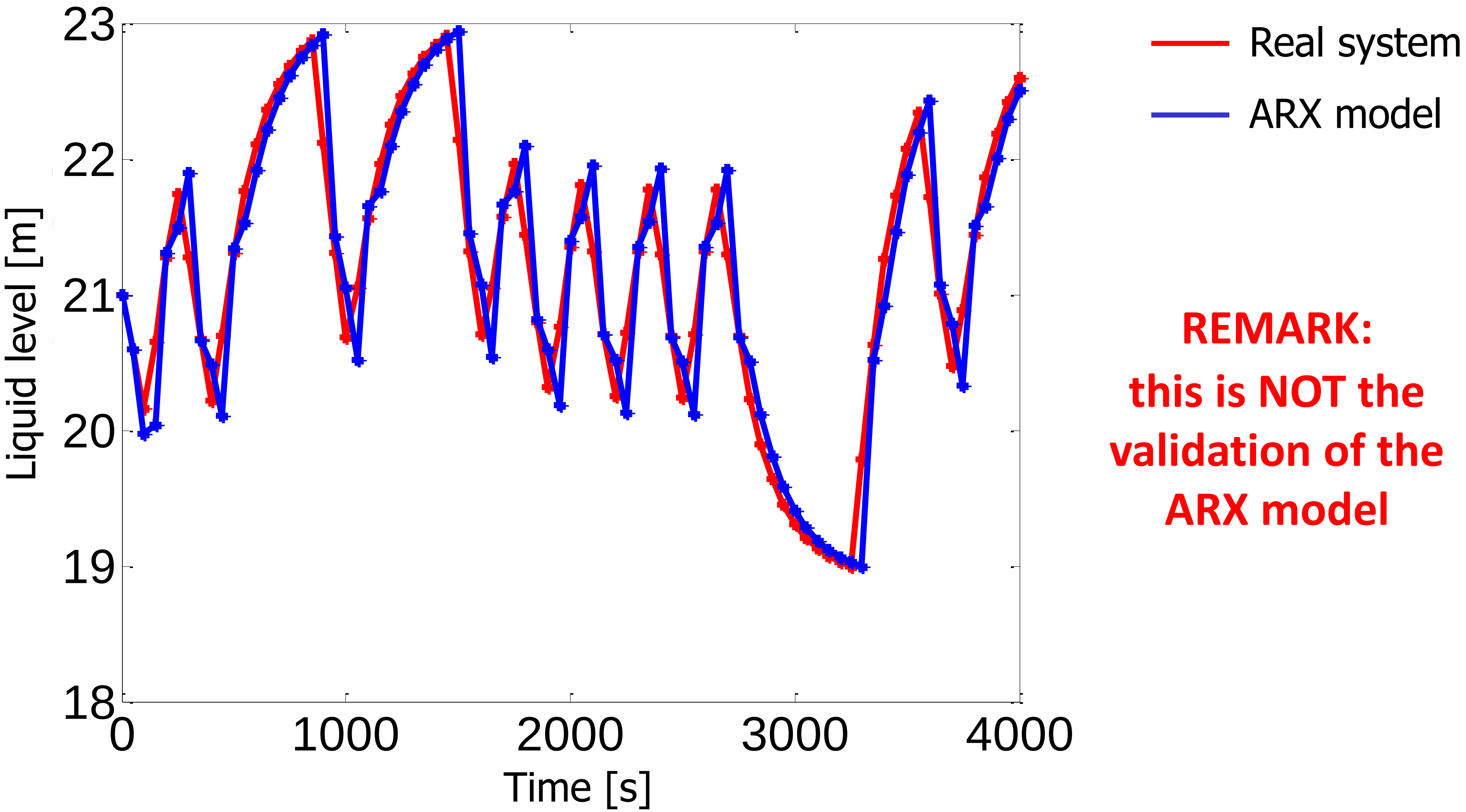


Real system response



ARX performance assessment

Comparison of the ARX model with the original identification data



REMARK:
this is NOT the validation of the ARX model

ARX validation