Dynamic Neural Networks for Actuator Fault Diagnosis: Application to the DAMADICS Benchmark Problem

Krzysztof PATAN[†] and Thomas PARISINI[‡]

[†]Institute of Control and Computation Engineering
University of Zielona Góra Poland
e-mail: k.patan@issi.uz.zgora.pl

[‡]Dept. of Electrical, Electronic and Computer Engineering
DEEI-University of Trieste, Italy
e-mail: parisini@univ.trieste.it

OUTLINE

1. Introduction
2. Dynamic neural network
3. Stochastic approximation
4. Sugar actuator
5. Experiments
6. Concluding remarks

INTRODUCTION

Motivations

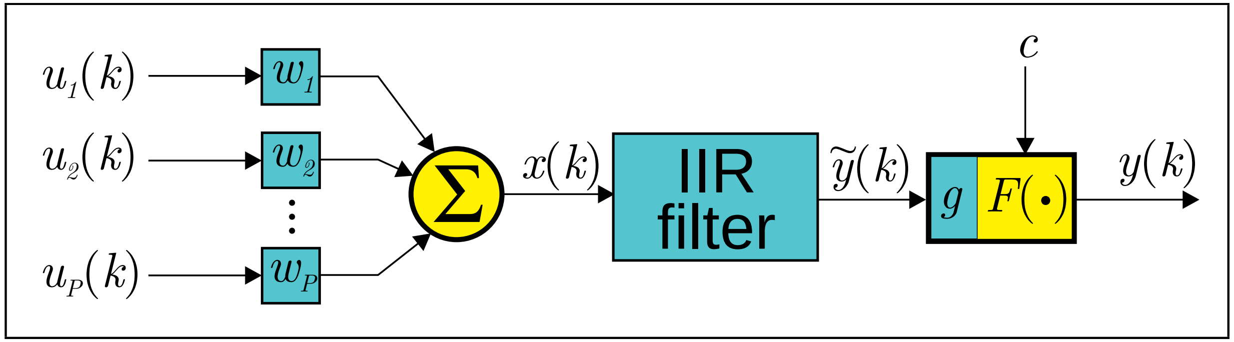
- Artificial neural networks – a useful and efficient tool for modelling of nonlinear dynamic processes
- Fault detection applications – a high model quality required
- Back-propagation based methods suffer from entrapment in local minima of an error function

Objectives

- Designing of an actuator FDI system using dynamic neural networks
- Application of a stochastic algorithm for dynamic neural network training
- Evaluation of the proposed FDI system using real process data (sugar evaporation process)

DYNAMIC NEURAL NETWORKS

➢ Dynamic model with IIR filter



- weighted adder

$$x(k) = \sum_{p=1}^{P} w_p u_p(k)$$

- filter module

$$\widetilde{y}(k) = \sum_{i=0}^{n} b_i x(k-i) - \sum_{i=1}^{n} a_i \widetilde{y}(k-i)$$

- activation function

$$y(k) = F(g \cdot \widetilde{y}(k) + c)$$

LEARNING PROCESS

- A vector of all unknown network parameters $\boldsymbol{\theta} = [\mathbf{w}, \mathbf{a}, \mathbf{b}, \mathbf{g}, \mathbf{c}]$
- Optimization task

$$\boldsymbol{\theta}^{\star} = \min_{\theta \in C} J(\boldsymbol{\theta}); \quad \boldsymbol{\theta} \in \mathbb{R}^p$$

C – a set of constraint, and cost function

$$J(k; \boldsymbol{\theta}) = \frac{1}{N} \sum_{i=1}^{N} \left(\mathbf{y}(k) - \hat{\mathbf{y}}(k; \boldsymbol{\theta})\right)^2$$

N – number of learning patterns

STOCHASTIC APPROXIMATION

- Recursive form of Stochastic Approximation (SA)

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k - a_k \hat{\boldsymbol{g}}_k(\hat{\boldsymbol{\theta}}_k)$$

where $\hat{\boldsymbol{g}}_k(\hat{\boldsymbol{\theta}}_k)$ – gradient estimate $\partial J/\partial \hat{\boldsymbol{\theta}}$, a_k – small positive number

Simultaneous Perturbation Stochastic Approximation (SPSA)

- Gradient estimate

$$\hat{\boldsymbol{g}}_{ki}(\hat{\boldsymbol{\theta}}_k) = \frac{L(\hat{\boldsymbol{\theta}}_k + c_k \Delta_k) - L(\hat{\boldsymbol{\theta}}_k - c_k \Delta_k)}{2 c_k \Delta_{ki}}$$

where $L(\cdot)$ – cost function measurement, Δ_k – random perturbation vector,
c_k – small positive number

- Random distribution

  1. Symmetrically distributed about zero
  2. Finite inverse moments $E(|\Delta_{ki}|^{-1})$

  ✓ Bernoulli ± 1 distribution satisfies these conditions

  ✗ popular normal and uniform distributions do not fulfill condition 2

- Gain sequences

  (i) $a_k, c_k > 0 \quad \forall k; \quad \lim_{k\to\infty} a_k \to 0; \quad \lim_{k\to\infty} c_k \to 0$

  (ii) $\sum_{k=0}^{\infty} a_k = \infty; \quad \sum_{k=0}^{\infty} \left(\frac{a_k}{c_k}\right)^2 < \infty$

  Gains calculation

$$a_k = \frac{a}{(A+k)^\alpha}, \quad c_k = \frac{c}{k^\gamma}$$

Advantages

✓ To calculate gradient estimate 2 measurements needed

✓ Low numerical complexity in contrast to gradient based methods and other stochastic algorithms

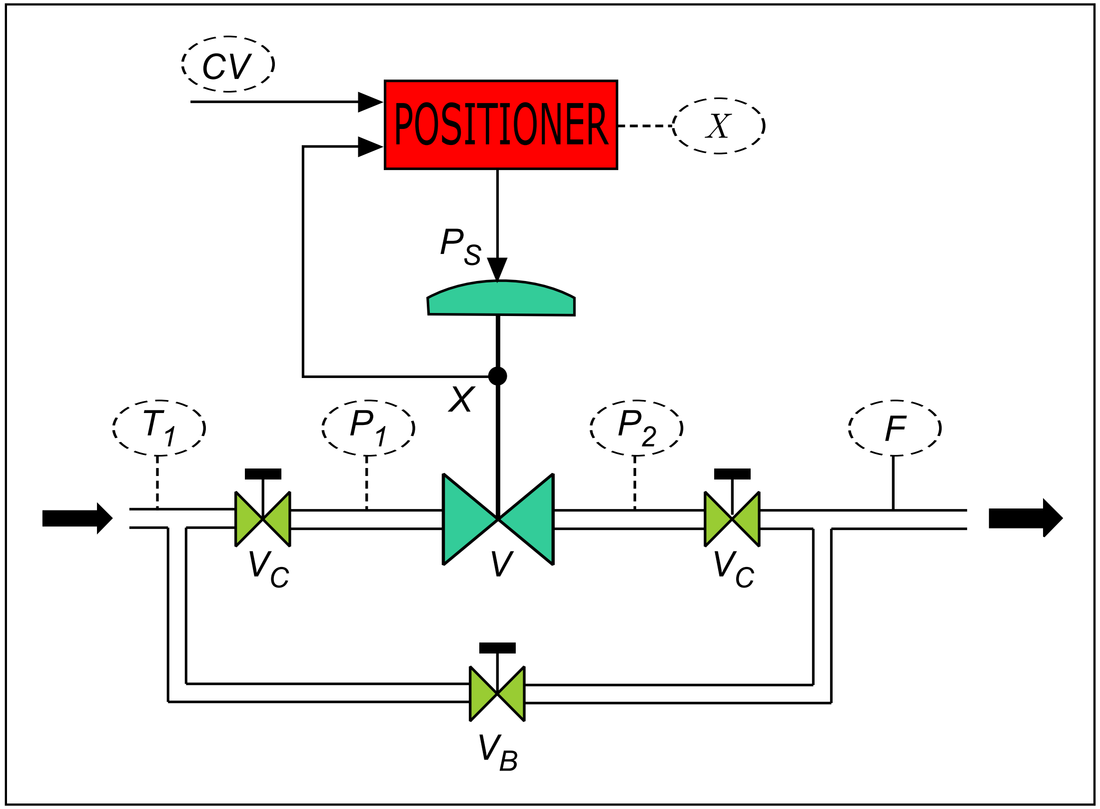
✓ Useful in the case of noisy data

✓ Global optimization property

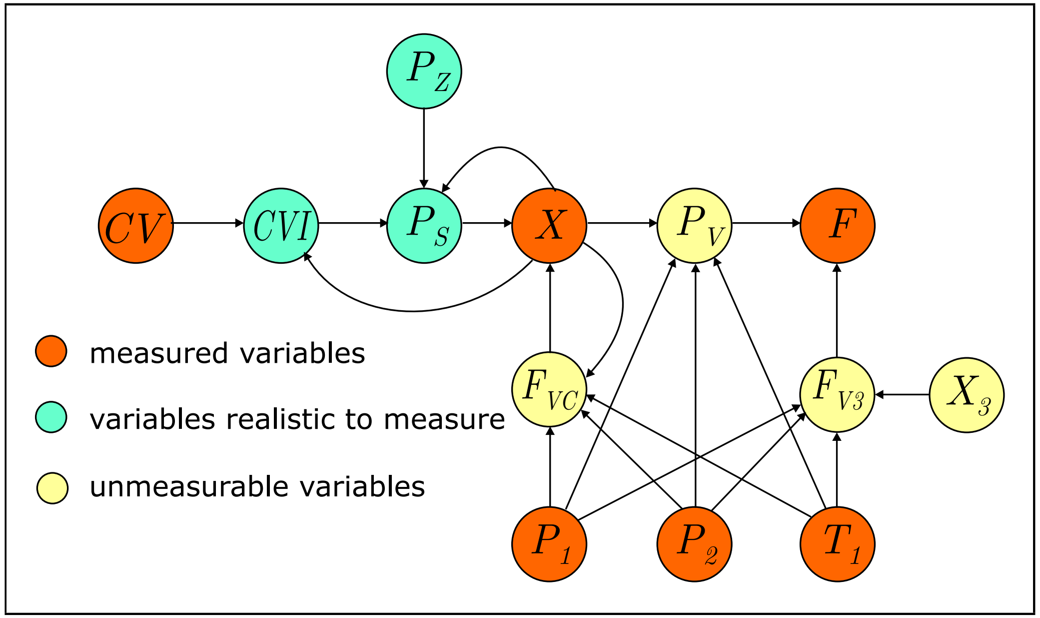
1. Chin D.C.: A more efficient global optimization algorithm based on Styblinski and Tang, *Neural Networks*, 7, pp. 573–574, 1994
2. Maryak J.L. and Chin D.C.: Global random optimization by Simultaneous Perturbation Stochastic Approximation, Proc *American Control Conference*, 25-27 June 2001, Arlington, VA, pp. 756-762

SUGAR ACTUATOR

- positioner
- servo-motor
- control valve



Causal graph



- Servo-motor rod displacement

$$X = r_1(CV, P_1, P_2, T_1, X)$$

- Flow through the valve

$$F = r_2(X, P_1, P_2, T_1)$$

Examined faulty scenarios

f_1 – positioner supply pressure drop

interpretation: oversized system air consumption, air leading pipes breaks, etc.

fault nature: rapidly developing

f_2 – unexpected pressure change across the valve

interpretation: media pump station failure, increased pipes resistance,
external media leakage

fault nature: rapidly developing

f_3 – fully opened by-pass valve

interpretation: valve corrosion, seat sealing wear, etc.

fault nature: abrupt

EXPERIMENT

Neural models

$$\begin{bmatrix} X \\ F \end{bmatrix} = NN(P_1, P_2, T_1, CV)$$

F - flow through the valve,

X - servo-motor rod displacement

P_1 - pressure on the valve inlet

P_2 - pressure on the valve outlet

T_1 - juice temperature on the valve inlet

CV - control signal

Data preprocessing

1. Inputs normalized to zero mean and standard deviation of one
2. Outputs transformed to fall in the range [-1,1]

Modelling

- Neural models – chosen using information criteria: AIC and FPE

Fault	**Structure**	**Filter order**	**Activation function**
f_0	$N^2_{4,5,2}$	2	hyperbolic tangent
f_1	$N^2_{4,7,2}$	1	hyperbolic tangent
f_2	$N^2_{4,7,2}$	1	hyperbolic tangent
f_3	$N^2_{4,5,2}$	1	hyperbolic tangent

- Data - the sugar campaign 2001
- Assumed accuracy - 0.01
- Parameters of the training procedure:
  - ◇ SPSA: $A = 100$, $a = 0.01$, $c = 0.01$, $\alpha = 0.602$, $\gamma = 0.101$

Decision making

- Let us assume that residual $r(k)$ is $\mathcal{N}(m, v)$
- A significance level β corresponds to probability that a residual exceeds a random value t_β with $\mathcal{N}(0, 1)$

$$\beta = prob(|r(k)| > t_\beta)$$

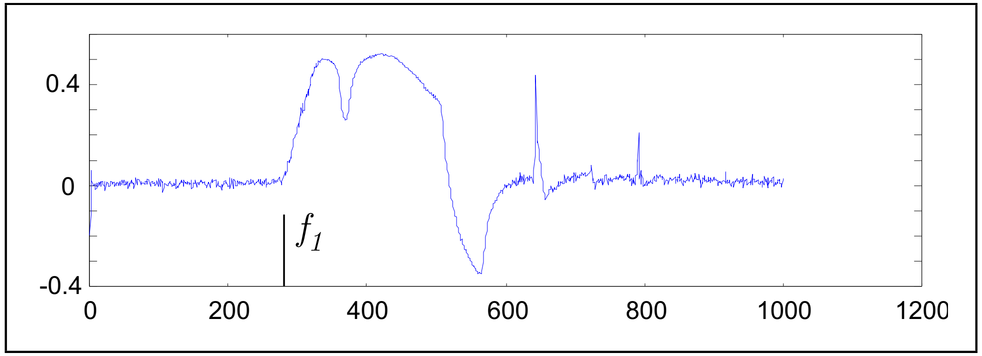
- Fixed threshold

$$T = t_\beta v + m$$

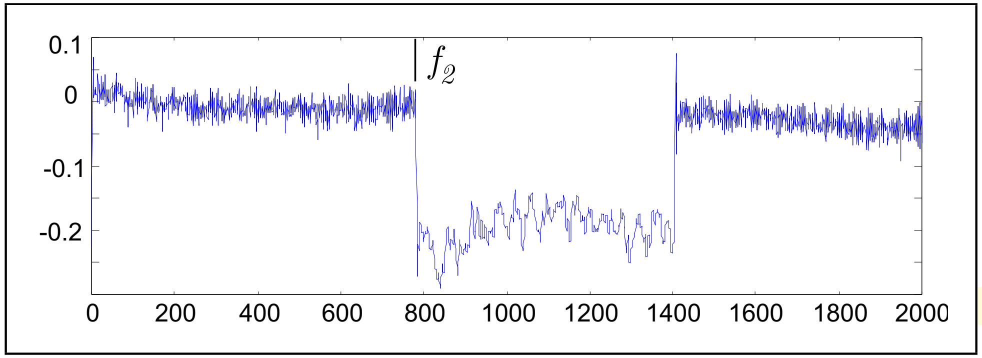
- Significance level used – $\beta = 0.05$

Fault detection – nominal model

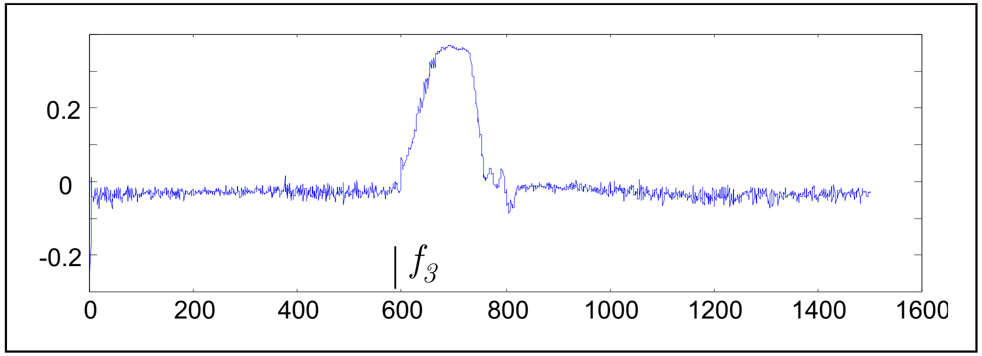
(a) fault f_1



(b) fault f_2

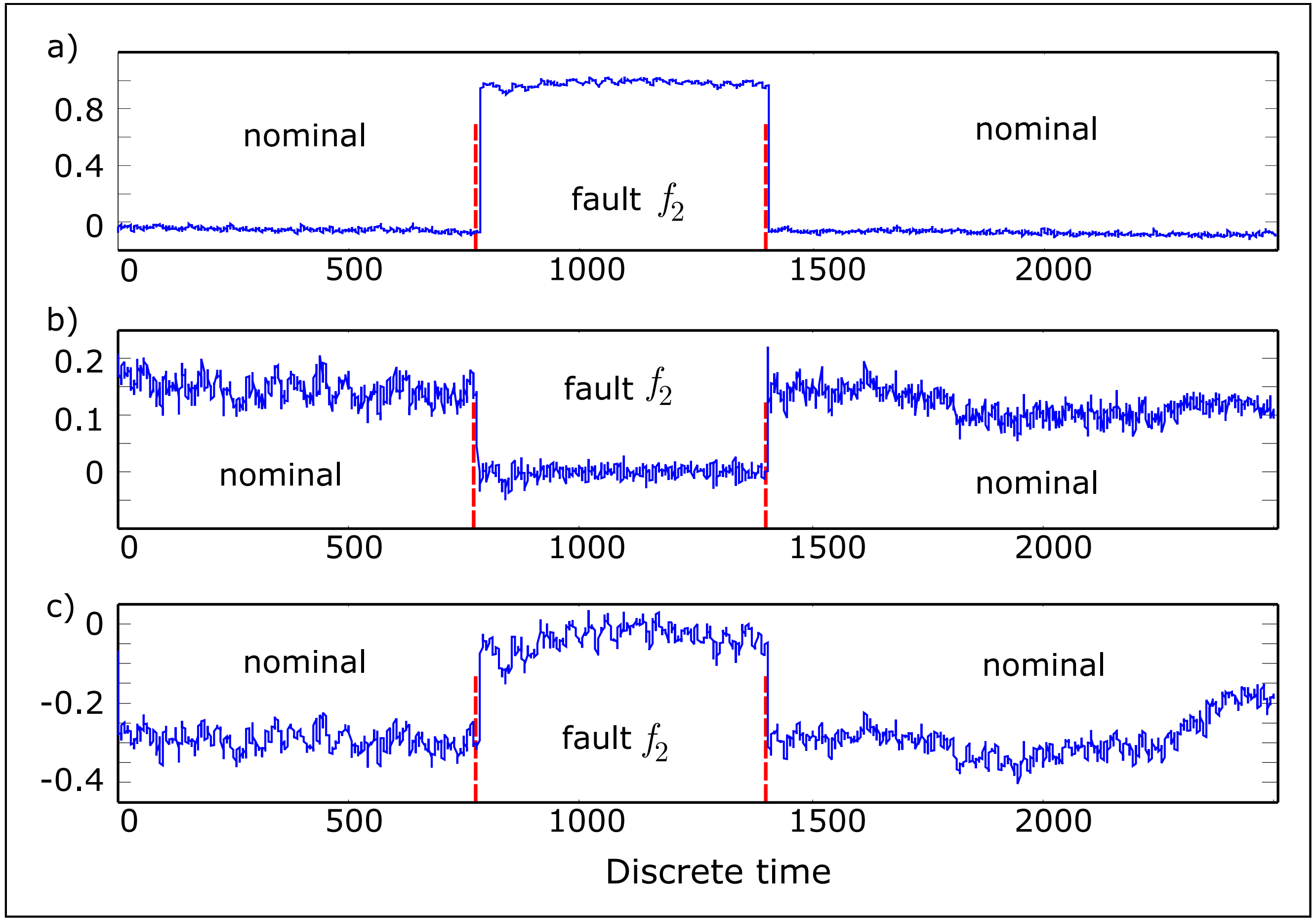


(c) fault f_3



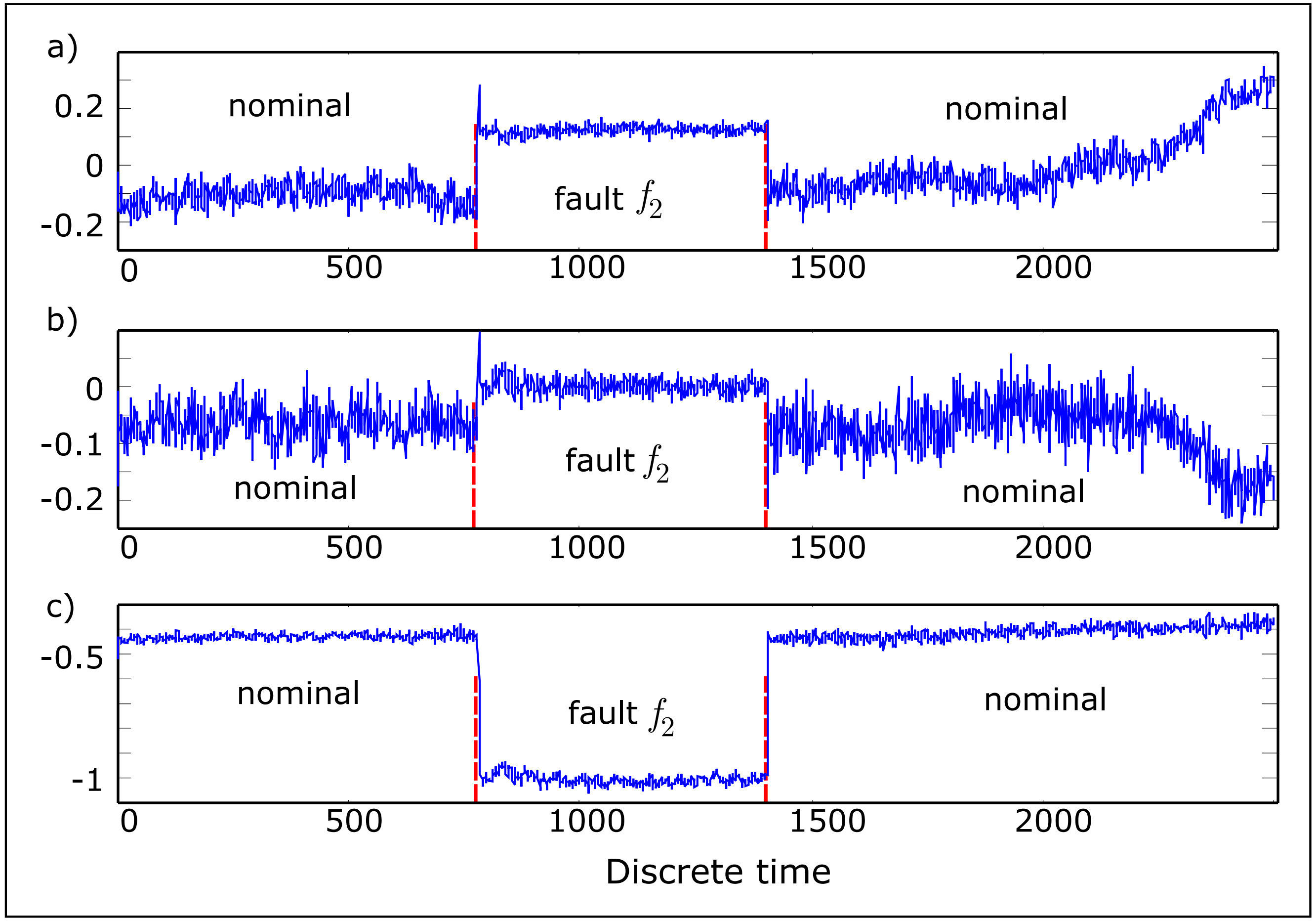
Fault isolation – fault f_2, output F

Fault model f_1 (a), fault model f_2 (b), fault model f_3 (c)



Fault isolation – fault f_2, output X

Fault model f_1 (a), fault model f_2 (b), fault model f_3 (c)



Fault diagnosis results

- Fault detection results (X – detectable, N – non detectable)

Fault	f_1	f_2	f_3
flow output	X	X	X
rod displacement output	X	X	X

- Fault isolation results (X – isolable, N – non isolable)

Fault	f_1	f_2	f_3
flow output	N	N	N
rod displacement output	N	X	X

Comparative study

- Sum of squared errors SSE
- Detection time t_{dt}
- False detection rate r_{fd}
- Isolation time t_{it}
- False isolation rate r_{fi}

- Modelling quality – SSE

  DN - Dynamic Network

  ARX - Auto-Regressive with eXogenous input

  NNARX - Neural Networks ARX

Method	f_0		f_1		f_2		f_3	
	F	X	F	X	F	X	F	X
DN	0.73	$\boxed{0.46}$	$\boxed{0.02}$	0.91	$\boxed{0.098}$	$\boxed{0.139}$	2.32	$\boxed{12.27}$
ARX	2.52	5.38	4.93	14.39	11.92	16.96	19.9	4.91
NNARX	$\boxed{0.43}$	0.71	0.089	$\boxed{0.1551}$	0.6	$2, 17$	$\boxed{0.277}$	22.5

➢ FDI properies

T_f – threshold used for the flow output

T_x – threshold used for the rod displacement output

Index	DN			NNARX		
	f_1	f_2	f_3	f_1	f_2	f_3
t_d	4	5	81	10	3	37
t_i	1	7	92	1	5	90
r_{fd}	0.34	0.26	0.186	0.357	0.42	0.45
r_{id}	0.08	0.098	0.091	0.145	0.0065	0.097
T_f	0.0164	0.0191	0.0468	0.0245	0.0541	0.0215
T_x	0.0936	0.0261	0.12	0.0422	0.0851	0.2766

CONCLUDING REMARKS

- ❐ Dynamic neural networks can be easily and effectively applied to design model-based fault detection and isolation systems
- ❐ Impossibility to model all potential system faults
- ❐ Data for faulty scenarios can be simulated
- ❐ Simultaneous Perturbation Stochastic Approximation algorithm
  - strong alternative to gradient based methods
  - useful when the search direction can not be determined accurately
  - property of the global optimization