Data reconciliation in OpenModelica: state estimation of industrial systems

OpenModelica Annual Workshop February 5, 2024
Linköping University

Baptiste MAZURIÉ - baptiste.mazurie@edf.fr
Audrey JARDIN - audrey.jardin@edf.fr
Pascal BAILLY - pascal.bailly@edf.fr
Adrian LUCA - adrian.luca@edf.fr
Didier BOLDO - didier.boldo@edf.fr
Thomas PRADIER – thomas.pradier@edf.fr

"Data Validation and Reconciliation (DVR) offers the nuclear power industry plants a method of improving the reliability of Core Thermal Power (CTP) calculations by reducing single point measurement vulnerabilities.

DVR methodology uses analytical thermodynamic principles and measurement uncertainty analyses"

EPRI, Use of Data Validation and Reconciliation Methods for Measurement Uncertainty Recapture, Topical Report, 2020

Summary

1. What is Data Reconciliation ?

2. Application to a thermal-hydraulic testing laboratory

3. Future perspectives for monitoring the performance of power plants

1

What is Data Reconciliation ?

Algorithm, assumptions and implementation in OpenModelica

1. What is Data Reconciliation ?

Data Reconciliation (DR):

- **Correct measurements** to make them **physically consistent** by using **an optimization problem under constraints**

Goals:

- Improving the **reliability of system state estimation**
    - **Reducing** the effect of **random errors**

Usages:

- **Detection of failures** (instrumentation or process)
- Reduction of **measurements uncertainties**

Assumptions:

- **Redundant** measurements
- Estimation of their **initial uncertainties***
- **Behavioral model** considered as **perfect** which describes how measured quantities are physically related to each other

* VDI2048 norm’s additional assumptions:

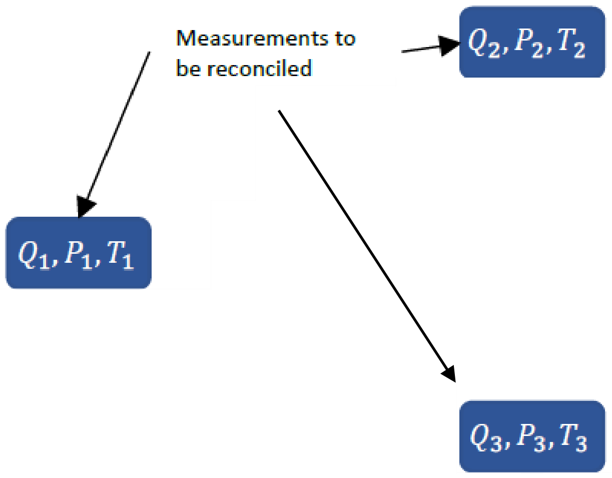
- **Uncertainties** follow a **Gaussian distribution**
- The observed process is in a **steady-state**

1. What is Data Reconciliation ?

$$X = [Q_i, P_i, T_i]$$
$$\sigma_X = [\sigma_{Q_i}, \sigma_{P_i}, \sigma_{T_i}]$$

Inputs:

- **Redundant measurement** data $\rightarrow X$
- Measurement **uncertainties** $\rightarrow \sigma_X$



1. What is Data Reconciliation ?

$X = [Q_i, P_i, T_i]$

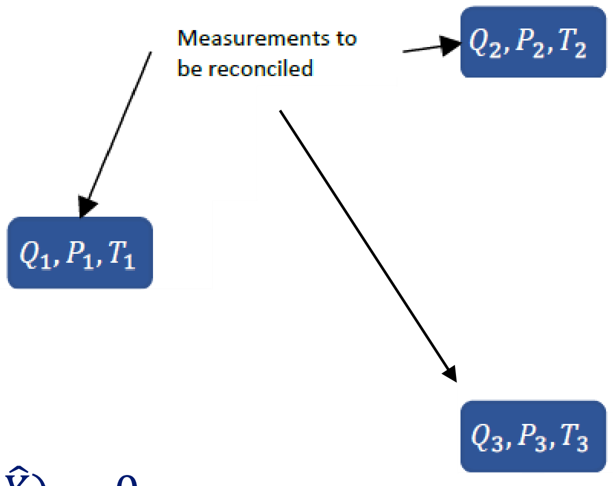
$\sigma_X = [\sigma_{Q_i}, \sigma_{P_i}, \sigma_{T_i}]$

+

Physical model

Inputs:

- **Redundant measurement** data $\rightarrow X$
- Measurement **uncertainties** $\rightarrow \sigma_X$
- **Representative behavioral model** of the system (assumed to be perfect $\rightarrow F(X) = 0$ and here steady-state)



$F(\hat{X}) = 0$

Fluid Splitter, *New Method to Perform Data Reconciliation with OpenModelica and ThermoSysPro, Bouskela et al. (2021)*

1. What is Data Reconciliation ?

$$X = [Q_i, P_i, T_i]$$
$$\sigma_X = [\sigma_{Q_i}, \sigma_{P_i}, \sigma_{T_i}]$$

+

Physical model

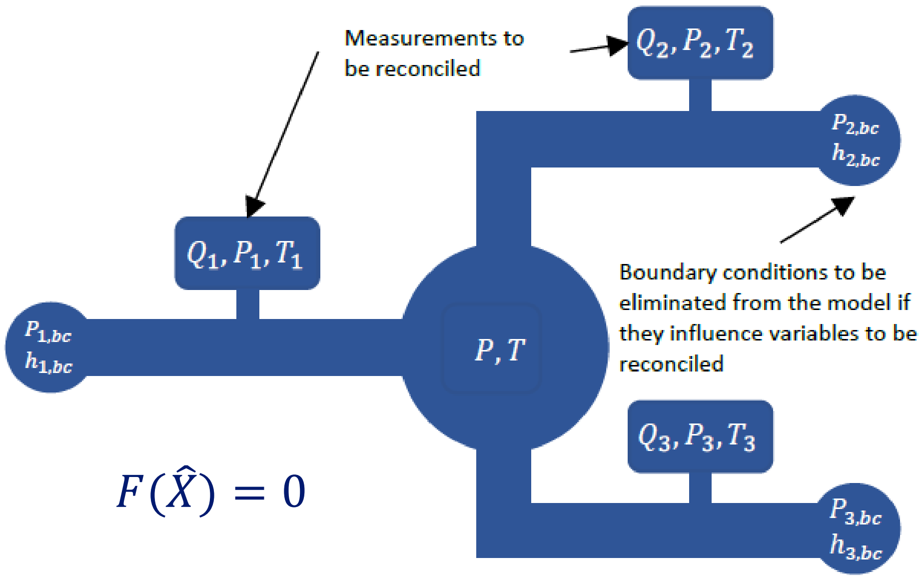
Inputs:

- **Redundant measurement** data $\rightarrow X$
- Measurement **uncertainties** $\rightarrow \sigma_X$
- **Representative behavioral model** of the system (assumed to be perfect $\rightarrow F(X) = 0$ and here steady-state)

Outputs:

- **Reconciled measurements:** $\rightarrow \hat{X}$
  - **Physically consistent** $\rightarrow F(\hat{X}) = 0$
  - **Improved**
    - **Closer** to the **"true" state** of the system $\rightarrow |\hat{X} - X^{True}| \leq |X - X^{True}|$
    - And associated to **reduced uncertainties** $\rightarrow \sigma_{\hat{X}} \leq \sigma_X$

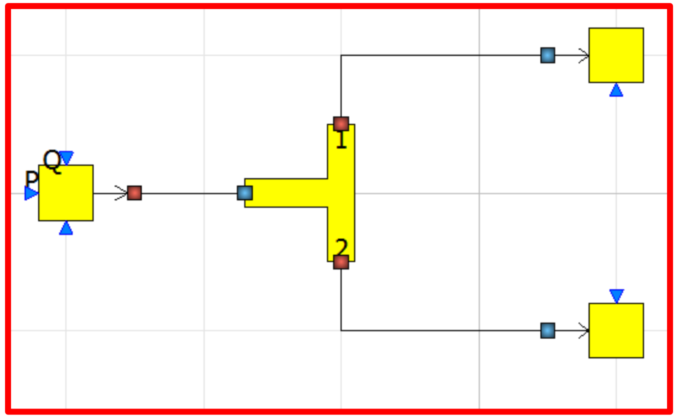
Better state estimation



Fluid Splitter, *New Method to Perform Data Reconciliation with OpenModelica and ThermoSysPro, Bouskela et al. (2021)*

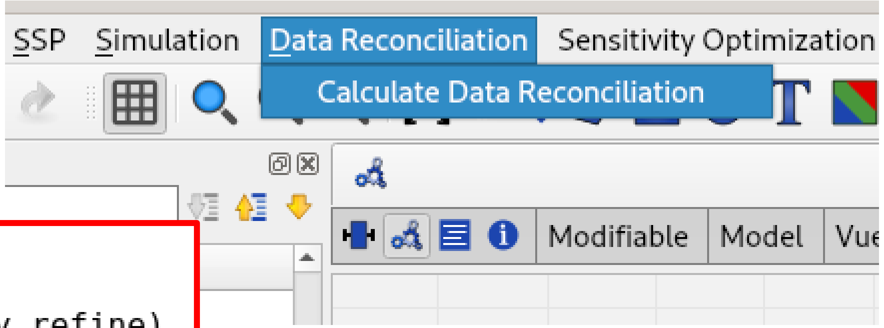
$$\hat{X} = [\widehat{Q_i}, \widehat{P_i}, \widehat{T_i}]$$
$$\sigma_{\hat{X}} = [\sigma_{\hat{Q}_i}, \sigma_{\hat{P}_i}, \sigma_{\hat{T}_i}]$$

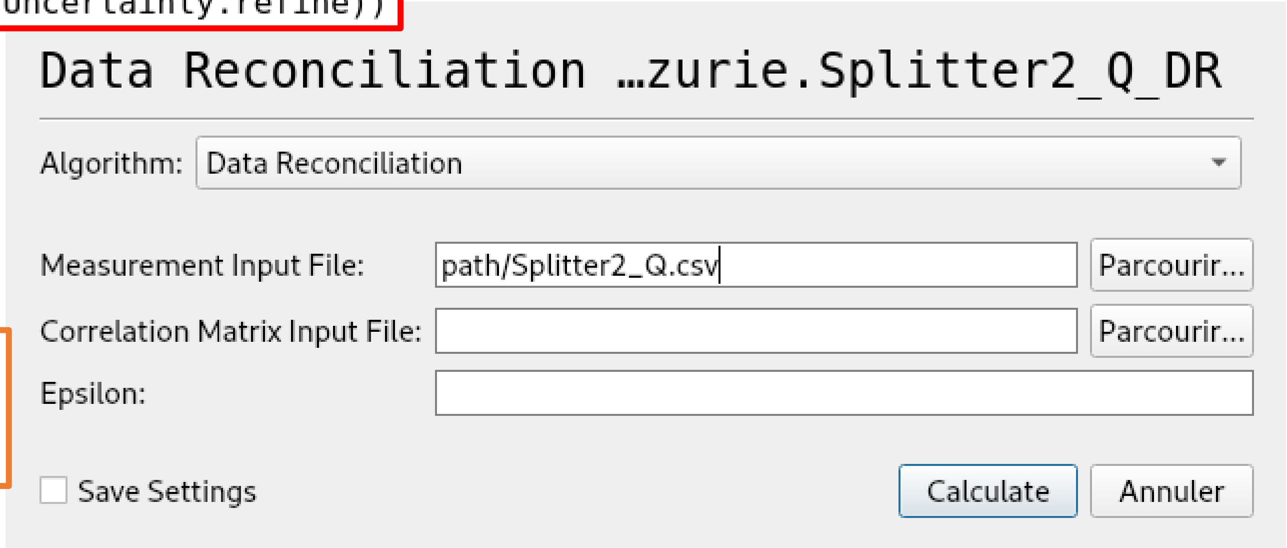
1. What is Data Reconciliation ? How to use it in OpenModelica ?



Annotations used for DR

```
model Splitter2_Q_DR
  Splitter2_Q splitter2_Q(
          Q1(uncertain = Uncertainty.refine),
          Q2(uncertain = Uncertainty.refine),
          Q3(uncertain = Uncertainty.refine))
```





OpenModelica Data Reconciliation interface

Inputs :

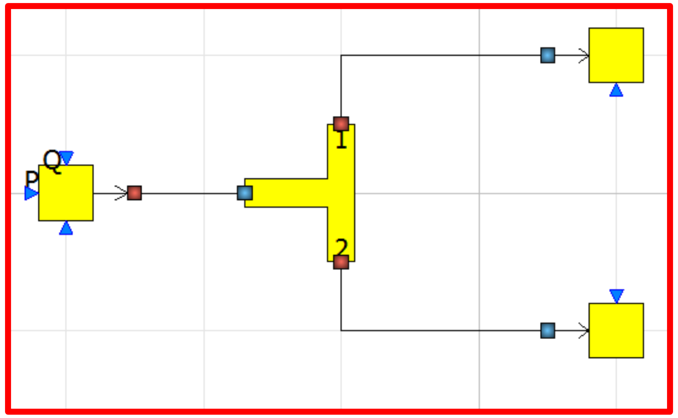
$$Q_1 = 12 \pm 2$$
$$Q_2 = 5 \pm 1$$
$$Q_3 = 5 \pm 1$$

```
Variable name;Measured value;Weight
splitter2_Q.Q1;12;2
splitter2_Q.Q2;5;1
splitter2_Q.Q3;5;1
```

CSV input file

- Functionality embedded in OpenModelica standard release
- For more details on how to use such reconciliation functionality → OpenModelica Users Guide/dataReconciliation

1. What is Data Reconciliation ?
   How to use it in OpenModelica ?



Inputs :

$$Q_1 = 12 \pm 2$$
$$Q_2 = 5 \pm 1$$
$$Q_3 = 5 \pm 1$$

Analysis:

Number of auxiliary conditions: 2
Number of variables to be reconciled: 3
Number of related boundary conditions: 1
Number of iterations to convergence: 2
Final value of (J*/r) : 0
Epsilon : 1e-10
Final value of the objective function (J) : 2.56107
Chi-square value : 5.99146
Result of global test : TRUE
Quality value (J/Chi-square) : 0.427453

Variables to be Reconciled	Initial Measured Values	Reconciled Values	Initial Half-width Confidence Intervals	Reconciled Half-width Confidence Intervals	Results of Local Tests	Values of Local Tests	Margin to Correctness(distance from 1.96)
exemple.Q1	12	10.6667	2	1.1547	TRUE	1.60033	0.359667
exemple.Q2	5	5.33333	1	0.57735	TRUE	0.800167	1.15983
exemple.Q3	5	5.33333	1	0.57735	TRUE	0.800167	1.15983

Outputs :

$$Q_1 = 10{,}67 \pm 1{,}15$$
$$Q_2 = 5{,}33 \pm 0{,}57$$
$$Q_3 = 5{,}33 \pm 0{,}57$$

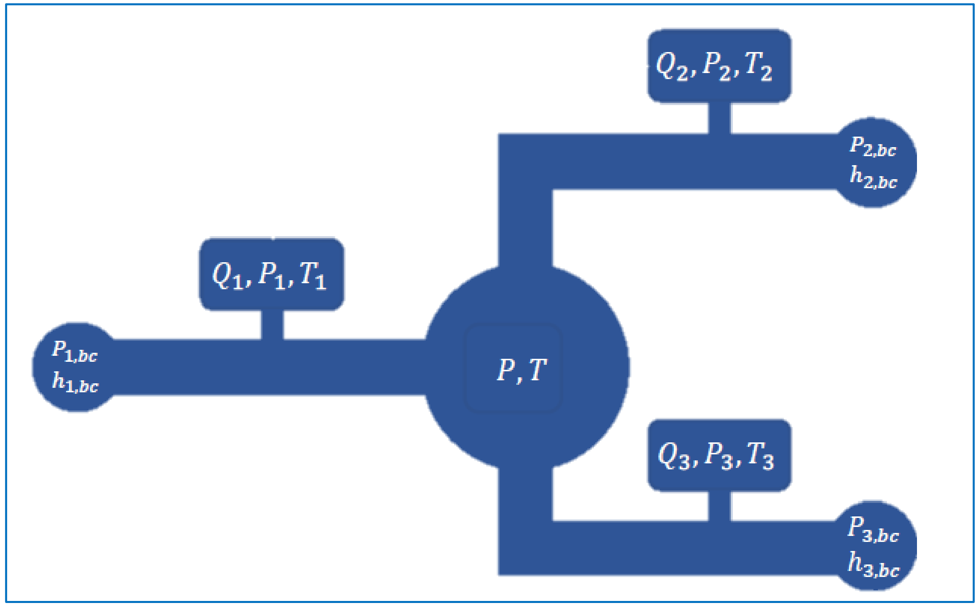
1. What is Data Reconciliation ?
How to diagnose system state from reconciled outputs

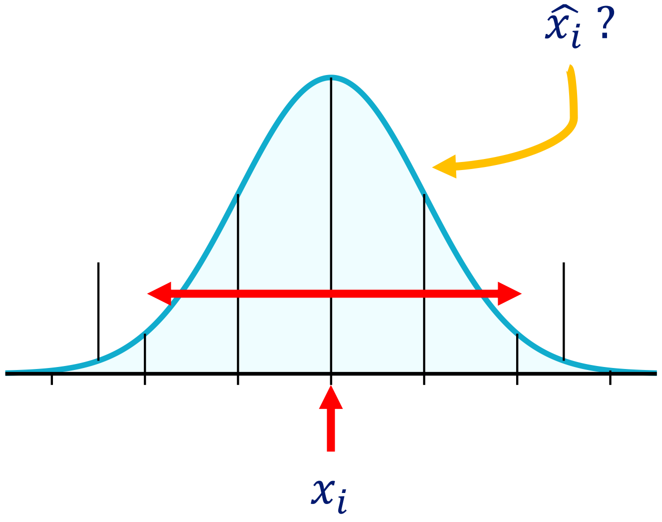
Different statistical criteria:

- **A global test C_1**
  - Are the measurements **consistent with the model**?
  - Consistent with initial assumptions on measurements uncertainties?
- **A set of local tests $C_{2,i}$ (one for each measurement i)**
  - Is the correction of i^{th} measured value within its confidence interval?

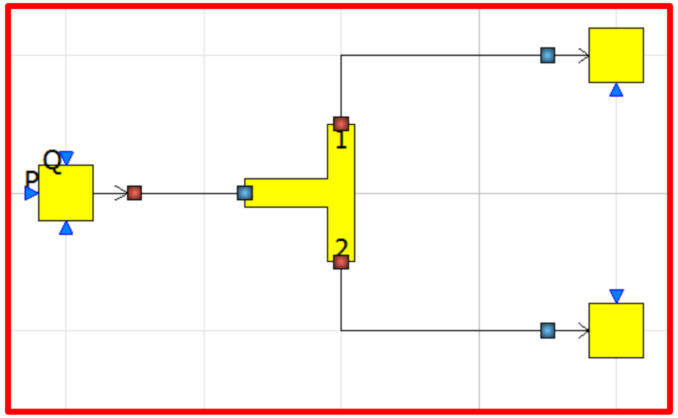
$$C_{2,i} = \frac{|\hat{x}_i - x_i|}{\sqrt{S_{v,i,i}}}$$

  - If $C_{2,i} > \lambda$: **Local failure is detected.** Root causes should be investigated to determine which assumption is not valid **(due to a faulty** sensor i or **local process default** not represented in the current model which would be hence no so perfect)





$S_{v,i,i}$ is the i^{th} diagonal element of the covariance matrix of the improvements $S_v = S_x - S_{\hat{x}}$



Inputs :

$$Q_1 = 12 \pm 2$$
$$Q_2 = 5 \pm 1$$
$$Q_3 = 5 \pm 1$$

Analysis:

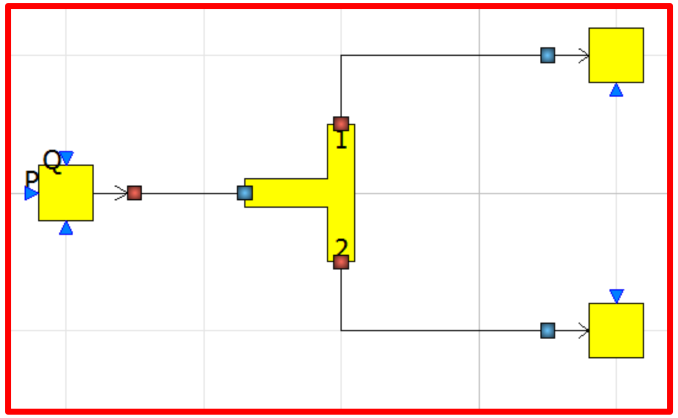
Number of auxiliary conditions: 2
Number of variables to be reconciled: 3
Number of related boundary conditions: 1
Number of iterations to convergence: 2
Final value of (J*/r) : 0
Epsilon : 1e-10
Final value of the objective function (J) : 2.56107
Chi-square value : 5.99146
Result of global test : TRUE
Quality value (J/Chi-square) : 0.427453

Variables to be Reconciled	Initial Measured Values	Reconciled Values	Initial Half-width Confidence Intervals	Reconciled Half-width Confidence Intervals	Results of Local Tests	Values of Local Tests	Margin to Correctness(distance from 1.96)
exemple.Q1	12	10.6667	2	1.1547	TRUE	1.60033	0.359667
exemple.Q2	5	5.33333	1	0.57735	TRUE	0.800167	1.15983
exemple.Q3	5	5.33333	1	0.57735	TRUE	0.800167	1.15983

- Conditions C1 & C2 are verified
  - Reconciled values are corrected within their confidence intervals
  - Uncertainties are reduced

Improved knowledge on the system state (with an estimation closer to the true state)

1. What is Data Reconciliation ? How to use it in OpenModelica ?



Analysis:

Number of auxiliary conditions: 2
Number of variables to be reconciled: 3
Number of related boundary conditions: 1
Number of iterations to convergence: 2
Final value of (J*/r) : 0
Epsilon : 1e-10
Final value of the objective function (J) : 16.0067
Chi-square value : 5.99146
Result of global test : FALSE
Quality value (J/Chi-square) : 2.67158

Variables to be Reconciled	Initial Measured Values	Reconciled Values	Initial Half-width Confidence Intervals	Reconciled Half-width Confidence Intervals	Results of Local Tests	Values of Local Tests	Margin to Correctness(distance from 1.96)
exemple.Q1	15	11.6667	2	1.1547	FALSE	4.00083	-2.04083
exemple.Q2	5	5.83333	1	0.57735	FALSE	2.00042	-0.0404166
exemple.Q3	5	5.83333	1	0.57735	FALSE	2.00042	-0.0404166

Inputs :

$$\cancel{Q_1 = 12 \pm 2}$$
$$Q_1 = 15 \pm 2$$
$$Q_2 = 5 \pm 1$$
$$Q_3 = 5 \pm 1$$

- Conditions C1 & C2 are not verified
  - Either the model is FALSE (ex: a leak is not represented)
  - Or the measurements are FALSE (ex: faulty sensor)

Detection of an inconsistency

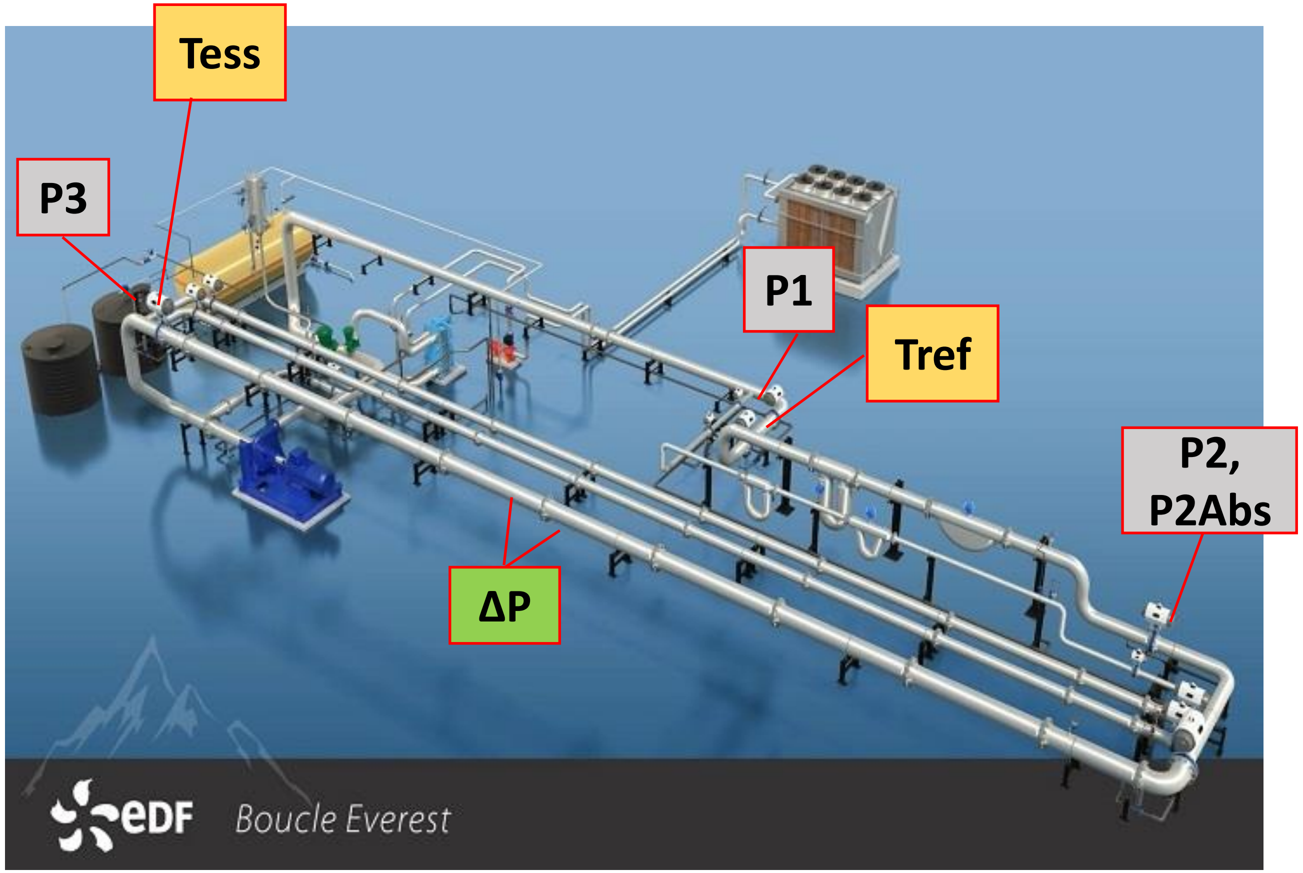
2

Application to a thermal-hydraulic testing laboratory

2. Application to a thermal-hydraulic testing laboratory
Study case : EVEREST testing laboratory

EVEREST at EDF R&D Chatou:

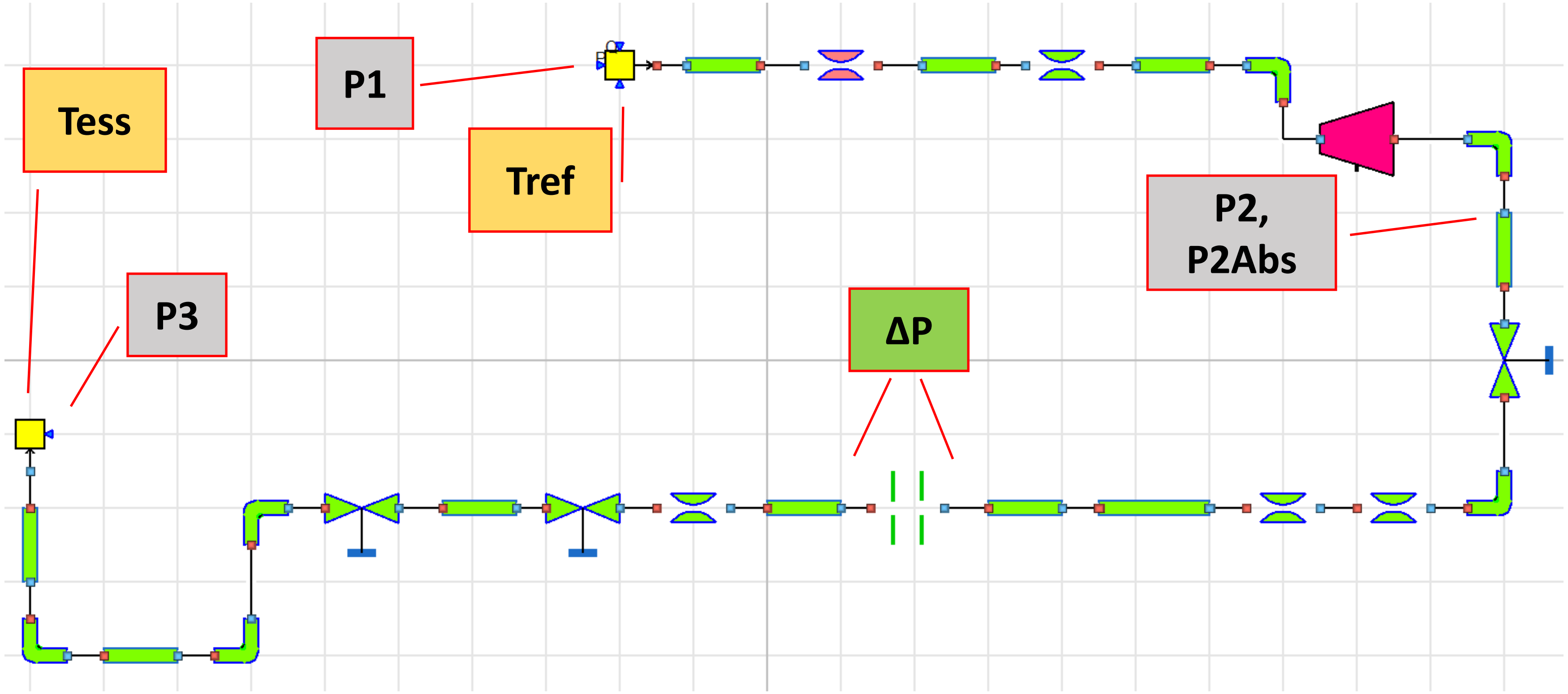
- An **experimental test facility** for analyzing the behavior of **measuring instruments** in high flow rates **water flows**
  - **Industrial scale** representative of nuclear power plant installations
  - **Reliable**: the loop is equipped with reference flowmeters (flow rate uncertainty less than 0.2%)
  - **Controlled**: temperature-, pressure- and flow-regulated to ensure on-demand thermodynamic conditions up to $1200\ m^3/h$
  - **Modular**: the circuit can be modified at will



Representation of the EVEREST test loop

Study case : Implementation with ThermoSysPro and OpenModelica

Implementation of the EVEREST test loop with ThermoSysPro library and OpenModelica:



ThermoSysPro model of the EVEREST test loop

2. Application to a thermal-hydraulic testing laboratory
Study case : EVEREST testing laboratory

EVEREST at EDF R&D Chatou:

- **Modular**: the circuit can be modified at will

Test campaign are run with **various defects** (deliberately machined on the test bench and precisely measured):

- **No defect** test
- **Trapezoidal** and **rectangular** bead **defects**
  - Made to represent weld beads inside pipes
- **Step defect**
  - Made to represent a slight change of diameter

Can it be detected with Data Reconciliation ?

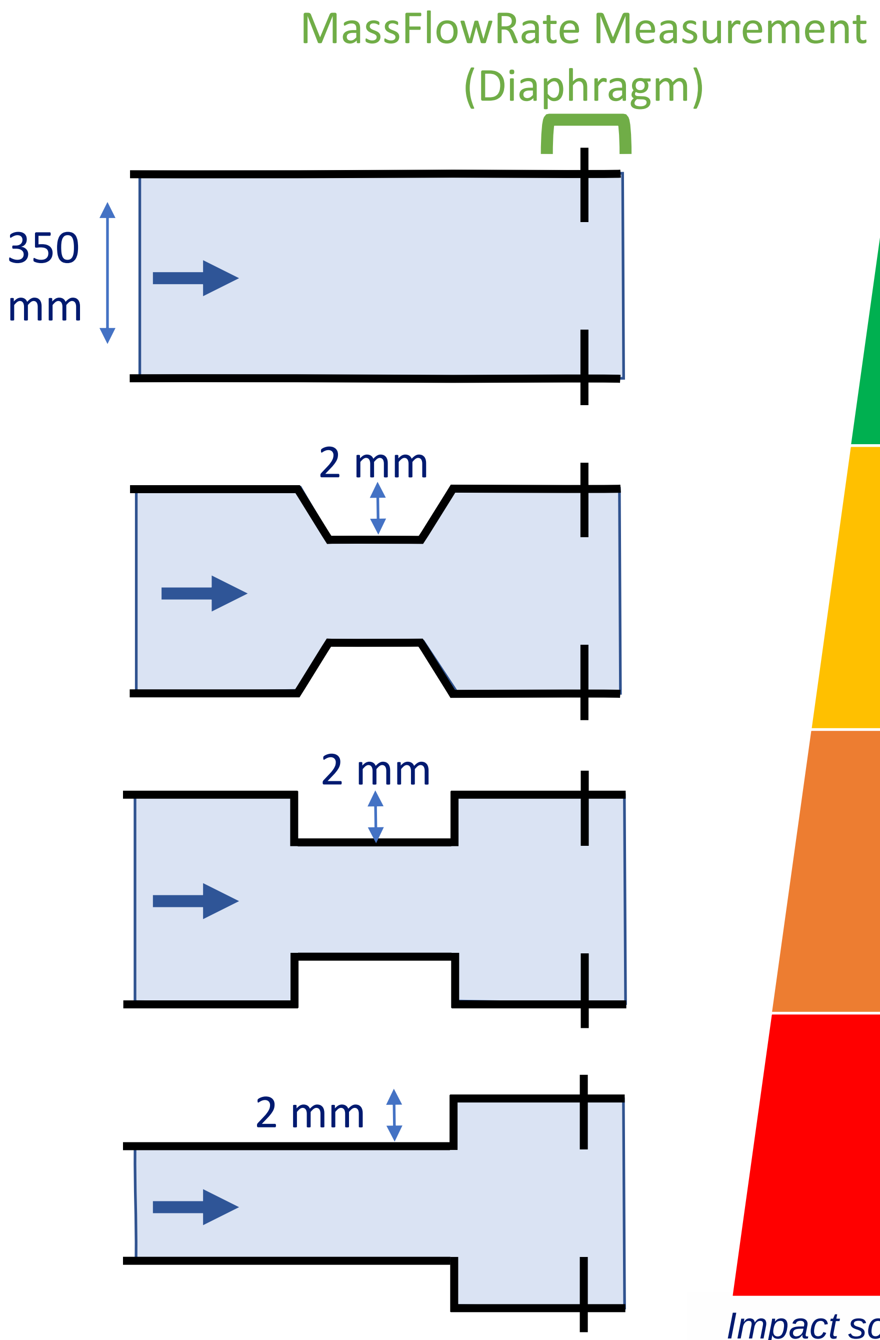
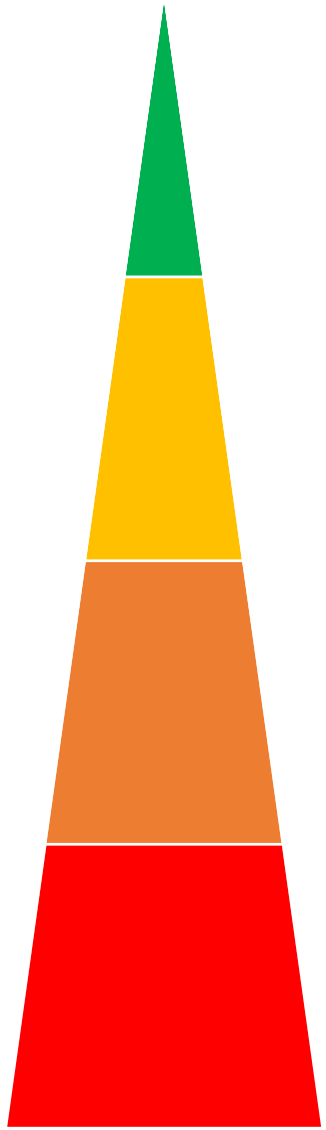


Diagram of the considered defects



Impact scale of the defects

2. Application to a thermal-hydraulic testing laboratory
Proof of defects detection

- **Step defect**
  - Made to represent a slight change of diameter

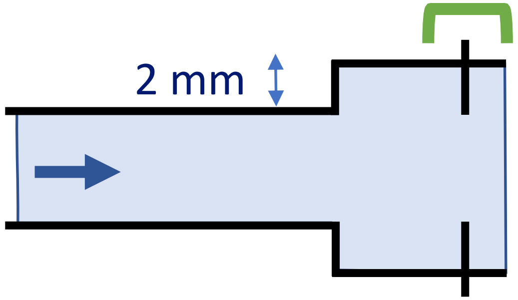


Diagram of the considered defect

Variable to be Estimated	Unit	Description	Initial Measured Value	Estimated Value	Initial Uncertainty	Estimated Uncertainty	Result of Local Test	Local Quality	Comment
everest.P2csv	bar		608443	607815	2900.8	1413.35	TRUE	0.247872	
everest.P2Abscsv	bar		709963	709140	3351.6	1413.35	TRUE	0.270552	
everest.Pacsv	bar		688197	691368	3351.6	1413.16	FALSE	1.04316	
everest.P3csv	bar		574430	573771	2214.8	1413.41	TRUE	0.386208	
everest.DeltaP_D_1	bar		29787.4	30302.5	377.47	41.6804	FALSE	1.373	
everest.DeltaP_D_2	bar		29835.3	30302.5	313.488	41.6804	FALSE	1.503	
everest.DeltaP_bride_1	bar		30799.6	31011.4	357.452	42.6581	TRUE	0.596	
everest.DeltaP_bride_2	bar		30800.1	31011.4	394.179	42.6581	TRUE	0.539	
everest.Trefcsv	degC		312.871	312.902	0.1421	0.0878243	TRUE	0.274	
everest.Tesscsv	degC		312.923	312.902	0.11172	0.0878243	TRUE	0.304	
everest.Qrefcsv			248.378	248.291	0.175255	0.170275	FALSE	1.557	

Analysis:

Number of auxiliary conditions: 8
Number of measured variables: 11
Number of unmeasured variables: 0
Number of related boundary conditions: 3
Number of iterations to convergence: 2
Final value of (J*/r) : 2.25184e-10
Epsilon : 1e-06
Final value of the objective function (J) : 23.5759
Chi-square value : 15.5073
Result of global test : FALSE
Quality (J/Chi-square) : 1.52031

Data reconciliation enables to detect such defect → OK but how effectively?

2. Application to a thermal-hydraulic testing laboratory
Test campaigns to evaluate the effectiveness of the reconciliation

- **Local tests $C_{2,i}$ (one for each measurement i)**
  - Is the correction of i^{th} measured value within its confidence interval?

$$C_{2,i} = \frac{|\widehat{x_i} - x_i|}{\sqrt{S_{v,i,i}}}$$

  - If $C_{2,i} > \lambda$: **Local failure** for sensor i
    - Run considered with a flaw
  - If $\forall i : C_{2,i} < \lambda$: Run considered with no flaw

- **Test campaigns** are run with **one defect** at a time:
  - **For each test campaign:**
    - Various **thermohydraulic conditions** (Mass Flow Rate variations)
      - 5 runs each
    - Comparison with **characterization and verification runs** (no defects)

105 measurement sets per test campaign (with and without the defect)

2. Application to a thermal-hydraulic testing laboratory
Test campaigns to evaluate the effectiveness of the method - indicators

- If $C_{2,i} > \lambda$: **Local failure** for sensor i
  - Run considered with a flaw
- If $\forall i : C_{2,i} < \lambda$: Run considered with no flaw

105 measurement sets per test campaign (with and without the defect)

Confusion matrix	Test with a defect	Test with no defect
Detected with a defect	True Positive (TP)	False Positive (FP) *False alarm*
Detected with no defect	False Negative (FN) *No detection*	True Negative (TN)

FP as few as possible

$$Sensitivity = \frac{TP}{TP + FN}$$

True positive rate: probability of detecting a defect for a test containing a defect

$$Specificity = \frac{TN}{TN + FP}$$

True negative rate: probability of not detecting any defect for a test that does not contain any

2. Application to a thermal-hydraulic testing laboratory

Test campaigns to evaluate the effectiveness of the method - indicators

$$Sensitivity = \frac{TP}{TP + FN}$$

True positive rate: probability of detecting a defect for a test containing a defect

$$Specificity = \frac{TN}{TN + FP}$$

True negative rate: probability of not detecting any defect for a test that does not contain any

Ideally Sensitivity and Specificity **should be equal to 1**

- All defects are detected
- No false alarms

Test campaigns to evaluate the effectiveness of the method - Step defect

$$Sensitivity = \frac{TP}{TP + FN}$$

True positive rate: probability of detecting a defect for a test containing a defect

$$Specificity = \frac{TN}{TN + FP}$$

True negative rate: probability of not detecting any defect for a test that does not contain any

Ideally Sensitivity and Specificity **should be equal to 1**

- All defects are detected
- No false alarms

Test campaign with **step defect**

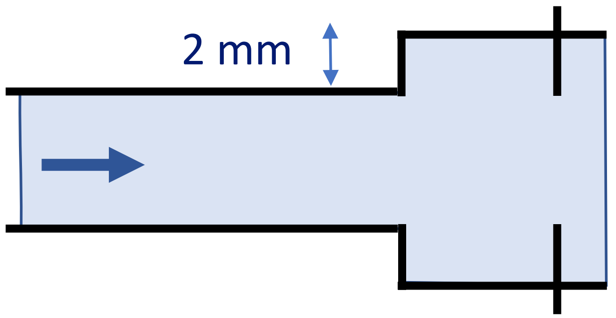


Diagram of the step defect considered

2. Application to a thermal-hydraulic testing laboratory
Test campaigns to evaluate the effectiveness of the method - Step defect

$$Sensitivity = \frac{TP}{TP + FN}$$

True positive rate: probability of detecting a defect for a test containing a defect

$$Specificity = \frac{TN}{TN + FP}$$

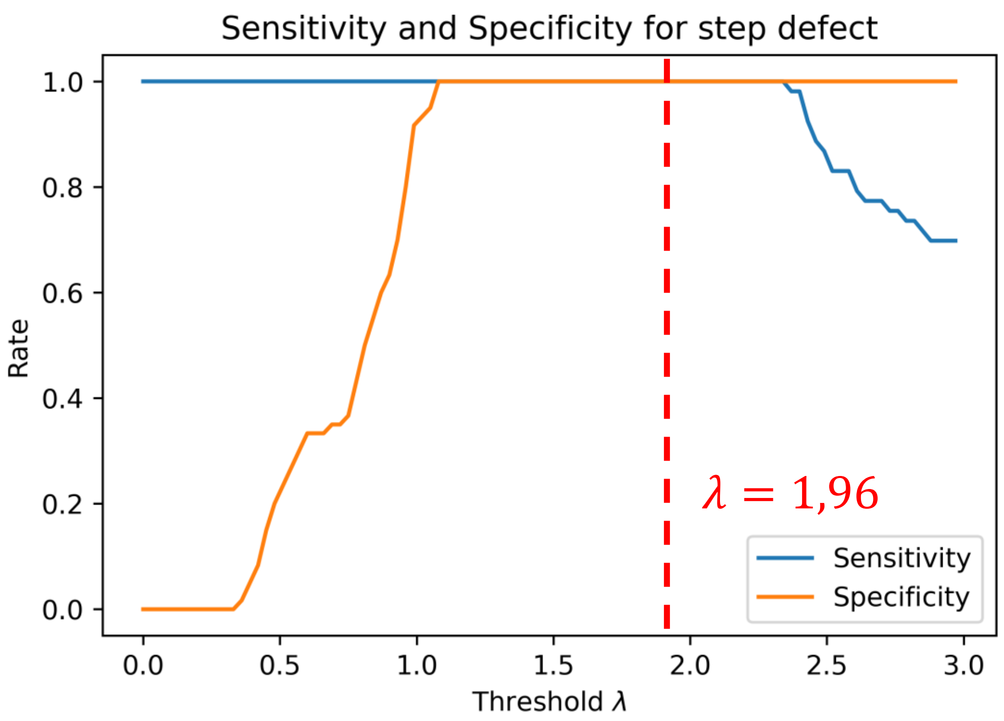
True negative rate: probability of not detecting any defect for a test that does not contain any

Ideally Sensitivity and Specificity **should be equal to 1**

- All defects are detected
- No false alarms

Test campaign with **step defect**

- Defect can be detected **perfectly** thanks to DR
- Specificity and Sensitivity **equal to 1** with $\boldsymbol{\lambda = 1,96}$



Always a perfect diagnosis?

Test campaigns to evaluate the effectiveness of the method - Rectangular bead defect

$$Sensitivity = \frac{TP}{TP + FN}$$

True positive rate: probability of detecting a defect for a test containing a defect

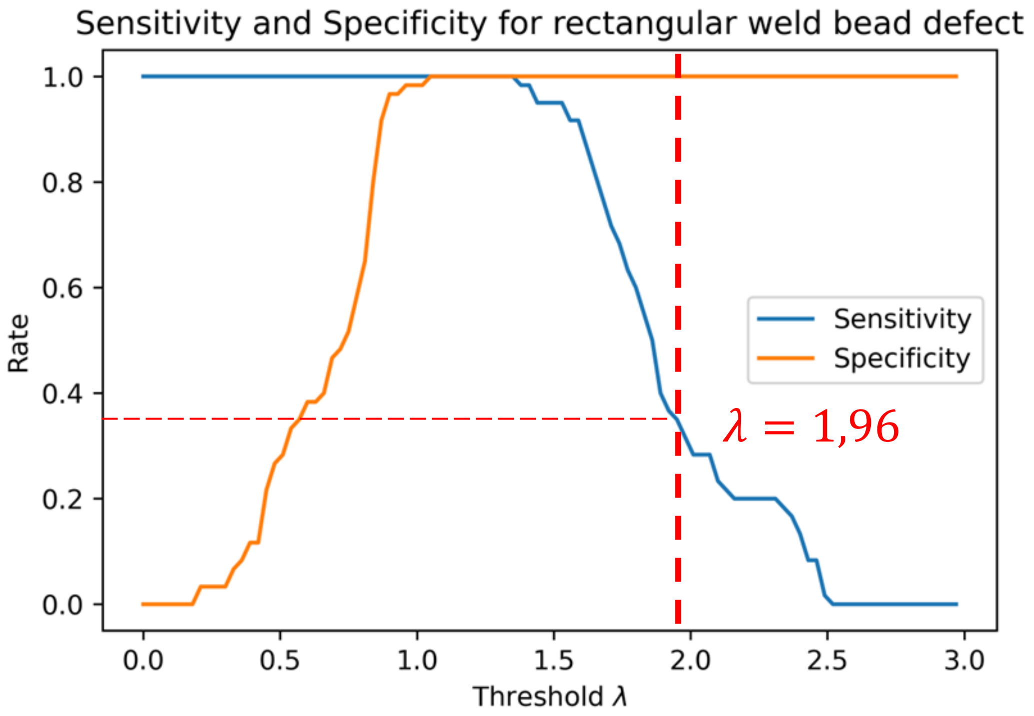
$$Specificity = \frac{TN}{TN + FP}$$

True negative rate: probability of not detecting any defect for a test that does not contain any

Test campaign for **rectangular weld bead defect**:

- Defect can be detected **perfectly** thanks to DR
    - But **not for** $\boldsymbol{\lambda = 1,96}$ where
        - $Sensitivity = 0{,}35$
- Specificity and Sensitivity equal to 1 with $\lambda = [1{,}05; 1{,}35]$

Impact of λ on the diagnosis



2. Application to a thermal-hydraulic testing laboratory
Test campaigns to evaluate the effectiveness of the method - Trapezoidal bead defect

$$Sensitivity = \frac{TP}{TP + FN}$$

True positive rate: probability of detecting a defect for a test containing a defect

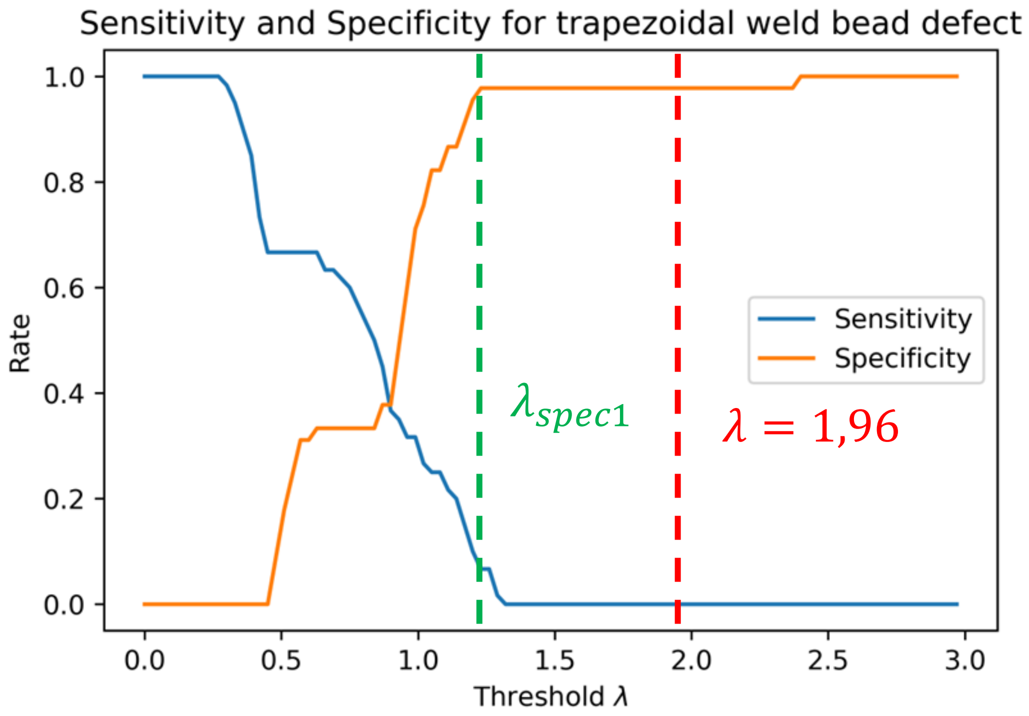
$$Specificity = \frac{TN}{TN + FP}$$

True negative rate: probability of not detecting any defect for a test that does not contain any

Test campaign for **trapezoidal weld bead** defect:
- Defect **cannot be perfectly detected**

- $\boldsymbol{\lambda = 1,96}$ not appropriate for every defect
  - $\boldsymbol{\lambda_{optimal}}$ depends on the defect

λ_{spec1} such as $Specificity = 1$
- **No false alarm**
- **Detection** becomes **less effective**



3

Future perspectives for monitoring the performance of power plants

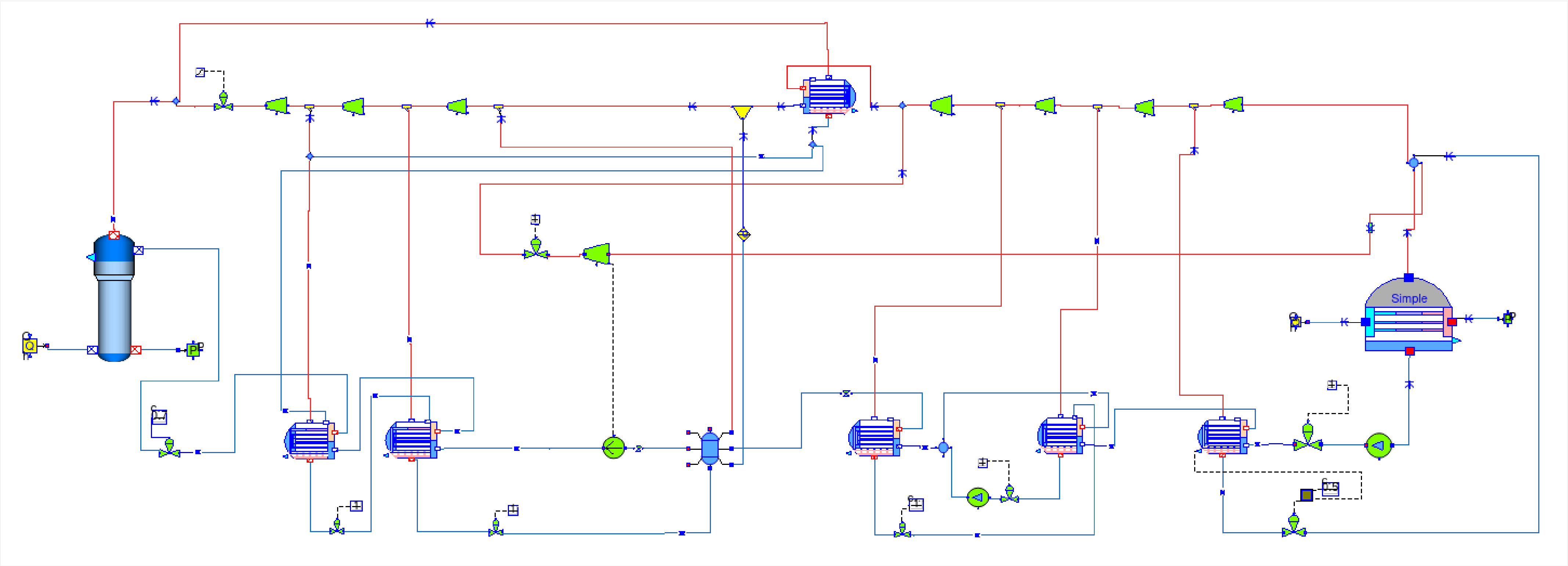
3. Future perspectives for monitoring the performance of power plants
Goals

"Data Validation and Reconciliation (DVR) offers the nuclear power industry plants a method of improving the reliability of CTP (Core Thermal Power) calculations by reducing single point measurement vulnerabilities."

EPRI, 2020

- Apply DR to more complex systems
  - Defect diagnosis on nuclear power plants
  - Consolidate measurements used in indicators such as CTP

3. Future perspectives for monitoring the performance of power plants
Use on larger ThermoSysPro models



ThermoSysPro model of the secondary loop of a 1300MW PWR

3. Future perspectives for monitoring the performance of power plants
Use of external databases collecting power plants on-site measurements

Behavioral models / Physical models

thermoSYSPRO

Measurements from updated databases

CSV

NEW

Scaling-up?

Data Reconciliation

Optimal estimation of the state of the system

Semi-automation using Bindings algorithm? (MODPROD'2018)

σ_X

Automatable?

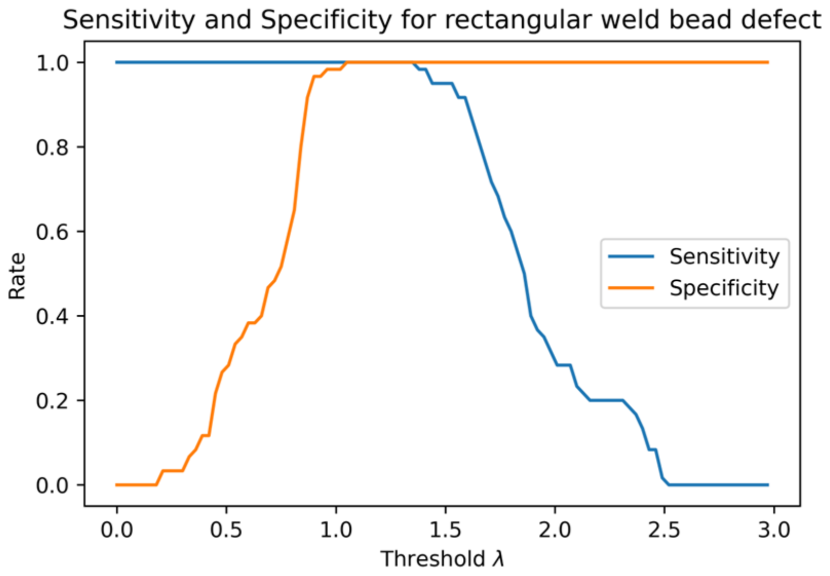
Measured quantity	Observation (or matching) function	Modelica variable
ARE101MD.Flowrate	Convert_T/H_to_Kg/s	ThermoSysPro. sensorAREQ1.signal

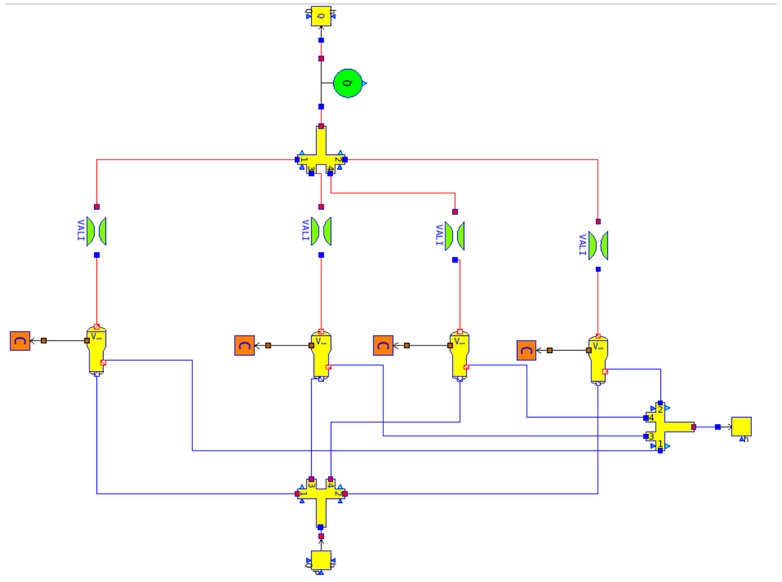
Conclusion

- **Data reconciliation detects defects intentionally reproduced on the test laboratory case**
- **Important choice of detection threshold λ**

Perspectives:

- **Use of DR on the EVEREST test laboratory as a monitoring tool to study sensors drift**
- **Use of DR on more complex models, such as the secondary loop, and automate as far as possible**
- **Use of DR (or similar algorithms) to help initialize Modelica models → see Luis Corona Mesa-Moles' talk at MODPROD'24**





ThermoSysPro model of the steam generators of a 1300MW PWR

Thank you