Data Reconciliation for Industrial Experiments

Baptiste Mazurié[1] Audrey Jardin[1] Pascal Borel[1] Frans Davelaar[1] Didier Boldo[1] Luis Corona Mesa-Moles[1]

[1]R&D, Electricité de France, France, `firstname.lastname@edf.fr`

Abstract

This paper presents the use of data reconciliation to improve the characterization of the operating state of experimental installations in an industrial context. It focuses on the development of a data reconciliation approach using the OpenModelica prototype to study: the detection of the defects in an experimental hydraulic test loop; and to validate measurement data of an HVAC testing facility. Data reconciliation is shown to be effective for the most pronounced defects intentionally introduced in the system. Regarding the characterization of HVAC measurements, data reconciliation identified two initially unnoticed invalid experiments.

Keywords: data validation, data reconciliation, default detection, uncertainties, measurements, Modelica, model reuse, industrial experiments

1 Introduction

Control of measurement equipment and associated acquisition chains is a key element for monitoring and diagnosis of industrial installations, in particular in the domain of electricity production, where constraints related to safety and resilience for securing the supply of energy networks require a sound and accurate knowledge of the true operating state of the system. To that aim, several experimental installations exist in premises of EDF Lab to improve the characterization of uncertainties associated with the examination of the main circuits of interests by testing new methods, using either direct measurements (via new sensor technologies) or indirect measurements (via the complementary use of knowledge models).

This article falls into this second category showing how a physical simulation model combined with statistical assessments can be used to:

- detect deviations from nominal operation and alert operators of the occurrence of possible faults
- validate measurement campaigns.

In practice, the work focuses on the use of a technique known as Data Reconciliation (DR). The aim is to obtain an estimate of the most likely state of the system by correcting the measured values on the actual installation so that the reconciled values are consistent with the equations governing the behaviour of the system.

While data reconciliation exists for several decades and its application to energy systems is ruled by a German standard (VDI 2048 2017), it is finding a renewed interest in the nuclear field for Measurement Uncertainty Recapture (MUR) and potential associated power uprates (EPRI 2020).

Dedicated tools (VALI, PROCESSPLUS, Thermal System Monitoring Enterprise, BILCO ...) exist to perform data reconciliation, but they require the physical models to be specifically developed for that purpose. They use their own proprietary language and are sometimes restricted to a specific domain (e.g. material processing industry). This makes data reconciliation costly and difficult to use. A possible answer to that problem is to apply data reconciliation on existing Modelica general purpose models, developed and validated for system operation.

This is why DR has been implemented in OpenModelica enabling the reuse of simulation models by automatically extracting the equations that should be considered for binding the measured variables together.

The principle and use on a simple example has been documented in (Bouskela et al. 2021). This paper presents first Data Reconciliation algorithm (Section 2) and then shows its applicability on two industrial experiments: one to control liquid flow rate measurements (Section 3) and a second to measure air flows (Section 4).

2 Data Reconciliation Theory

Industrial processes are usually monitored with a number of measurement equipments which can be more or less isolated on different sections of the circuit and of different technologies leading to differences in terms of representativeness, accuracy and reliability. The problem is to determine whether all these measurements, which are necessarily subject to error, are physically and statistically consistent with each other and whether they are a reliable mirror of the installation.

Data reconciliation is an optimization process that allows numerical correction of measurement values so that they satisfy the physical equations of the system (the model). If the corrections are too significant, the measurements are invalidated: in that case reconciliation alerts to a possible instrumentation defect (e.g., poor calibration or a faulty sensor), a process defect (e.g., an unmodeled leak) or a poor knowledge of sources of measurements uncertainties (e.g., lack of correlation effects). If the measurements are validated, the reconciliation exploits the redundancy of information to reduce uncertainties and obtain an estimate as close as possible to the true state of the system.

2.1 Mathematical Foundations

Data reconciliation aims at improving the accuracy of measurements by reducing the effect of random errors in the data. The main difference between data reconciliation and others techniques for propagating uncertainties such as Quadratic Combination or Monte Carlo simulation (Baudin et al. 2017; EPRI 2020) is that data reconciliation uses a model to express the physical constraints on the variables of interest. The measured values of these variables are then adjusted such that the estimates satisfy the constraints: the variables are thus reconciled.

Let $\{X_0^t, \ldots, X_d^t\}$ be a vector of d physical quantities constrained by a set C of r equations or auxiliary conditions derived from physical laws. These quantities must satisfy Equation 1 to ensure consistency.

$$\begin{aligned} C : \mathbb{R}^d &\rightarrow \mathbb{R}^r \\ 0 &= C(X^t) \end{aligned} \tag{1}$$

Measurements $\hat{X}$ of X^t inevitably have uncertainties, which are commonly modeled by a random variable X with a multivariate Gaussian distribution characterized by its mean vector μ and its variance-covariance matrix Σ.

$$\pi(X) = \mathcal{N}(\mu, \Sigma) \tag{2}$$

The contradiction vector U of a given state $\{X_0, \ldots, X_d\}$ is given by:

$$U = C(X) \tag{3}$$

The objective of data reconciliation is to define a conditional probabilistic distribution of measurements, called the reconciled distribution, such that Equation 1 is satisfied by the random variable $X|C(X) = 0$. This reconciled distribution is a multivariate Gaussian characterized by a mean vector μ' and a variance-covariance matrix Σ'.

$$\pi(X|C(X) = 0) = \mathcal{N}(\mu', \Sigma') \tag{4}$$

(Bouskela et al. 2021) present a detailed methodology for solving this optimization problem using Lagrange multipliers as stated in (VDI 2048 2017). The optimization problem consists in minimizing the following objective function in order to determine the set of reconciled values μ':

$$J(\mu') = (\mu' - \mu).\Sigma.(\mu' - \mu)^\top \tag{5}$$

The variance-covariance matrix of reconcilied values derive from the general formula of uncertainty propagation:

$$\Sigma' = \partial_\mu \mu'.\Sigma.\partial_\mu \mu'^\top \tag{6}$$

Relevancy of the reconciled distribution can be assessed using local and global test : χ^2, reflecting the relevancy of the estimated reconciled distribution based on prior measurements uncertainty knowledge. The global test involves verifying the following statistical hypothesis, where r denotes the number of degrees of freedom and p represents the significance level associated with the χ^2 distribution:

$$J(\mu') \leq \chi^2_{r,p} \tag{7}$$

The improvement vector v is defined as:

$$v + \mu = \mu' \tag{8}$$

Local tests, measurement by measurement, can help to identify measurements with a too large improvement pushing the most probable reconciled value to be out of its confidence range defined by the prior measurement knowledge as presented in Equation 9.

$$|\mu_i' - \mu_i| < \Sigma_{ii}^v \lambda_p \tag{9}$$

where μ_i is the i^{th} measurement, μ_i' is the i^{th} reconciled measurement which is also the mean of the i^{th} marginal of the reconciled distribution and λ_p is the confidence interval width, usually taken as 1.96 corresponding to 95% of the possible values. Σ^v is the variance-covariance matrix of the improvement v defined as:

$$\Sigma^v = \partial_\mu v \Sigma \partial_\mu v^\top \tag{10}$$

2.2 Measurement Uncertainty: GUM

The measurement of physical quantity inevitably has an uncertainty which has to be estimated. Let X_i^t the true value of physical quantity to measure, $\hat{X}_i$ the corresponding measurement with an uncertainty ε_i

$$X_i = \hat{X}_i + \varepsilon_i \tag{11}$$

ε_i is a random variable with a null mathematical expectation which can by decomposed in elementary uncertainty sources as follows:

$$\varepsilon_i = \varepsilon_i^0 + \ldots + \varepsilon_i^n \tag{12}$$

One can follow the guide (ISO/IEC GUIDE 98-3 2008) to establish a model of the random variable ε_i by proposing a multivariate statistical distribution $\pi(\varepsilon_i)$. Typically, there are five sources of uncertainty: sensor calibration, sensor environment impact, sensor acquisition chain, measurement methodology, and repeatability uncertainties. The first four are epistemic uncertainties, which must be determined a priori, usually based on sensor provider documentation or calibration reports. The last source, if it exists, is a stochastic uncertainty that must be evaluated by empirical samples collected during measurement.

When the uncertainty sources are assumed to be Gaussian, it is necessary to specify the variance-covariance matrix of the uncertainty sources, denoted as Σ. Estimating the correlation is a challenging aspect that can significantly impact uncertainty propagation. In practice, without empirical samples, the correlation is often assumed to be either null or total.

DOI 10.3384/ecp21845

Complex measurements of quantities Y can be based on several measurements of elementary quantities that are combined with a function g.

$$Y = g(X) \tag{13}$$

Uncertainty propagation with methods such as first-order Taylor series expansion or Monte Carlo simulation aims to estimate the statistical distribution of Y. In the context of data reconciliation, estimating the multivariate distribution of Y, which is assumed to be a multivariate Gaussian distribution, is a crucial step that must be performed meticulously. This includes accurately estimating the correlation between the uncertainties of different measurements. The construction of this multivariate distribution can be achieved using OpenTURNS (Baudin et al. 2017), as illustrated in Figure 1.

2.3 Data Reconciliation in OpenModelica

The main added-value of performing DR in OpenModelica lies in the reuse of models initially developed for other purposes than data validation or diagnosis.

In particular the equational nature of the Modelica language is used to automatically extract from a complete simulation model (i.e. a square system of equations with the same numbers of equations than unknowns) a minimal subset of auxiliary equations that will bind the measured variables together (i.e. here a non-square system of equations since the redundancy of measurements is exploited for the reconciliation).

The extraction algorithm has already been fully described in (Bouskela et al. 2021) and is now available in the standard release of OpenModelica.

For the sake of clarity, Figure 1 summarizes the main steps required to perform a complete data reconciliation process in OpenModelica. For more details and illustration on the simple example of a splitter, the user is invited to read the documentation accessible online (*OpenModelica Documentation: Data Reconciliation* 2021).

On Figure 1 light blue refers to steps implemented in OpenModelica while dark blue highlights steps performed in a third-party tool dedicated to uncertainty quantification (in this case OpenTURNS (Baudin et al. 2017) but other tools can be chosen). Basically, the user :

1. states the reconciliation problem by specifying which variables should be reconciled and which equations should be voluntary not be considered as an auxiliary condition (typically the boundary conditions). This is done by decorating the Modelica model with dedicated annotations;

2. defines the set of measurements and quantifies their corresponding uncertainties (potentially with the help of a dedicated tool). The numerical information to be reconciled can be provided in the form of a csv file;

3. launches the DR functionality which will extract the auxiliary conditions and computes the reconciled data. Based on the results, a report is generated in the form of an html file with reconciled outputs and statistical tests.

4. analyses the validity of reconciled data with respect to the initial assumptions.

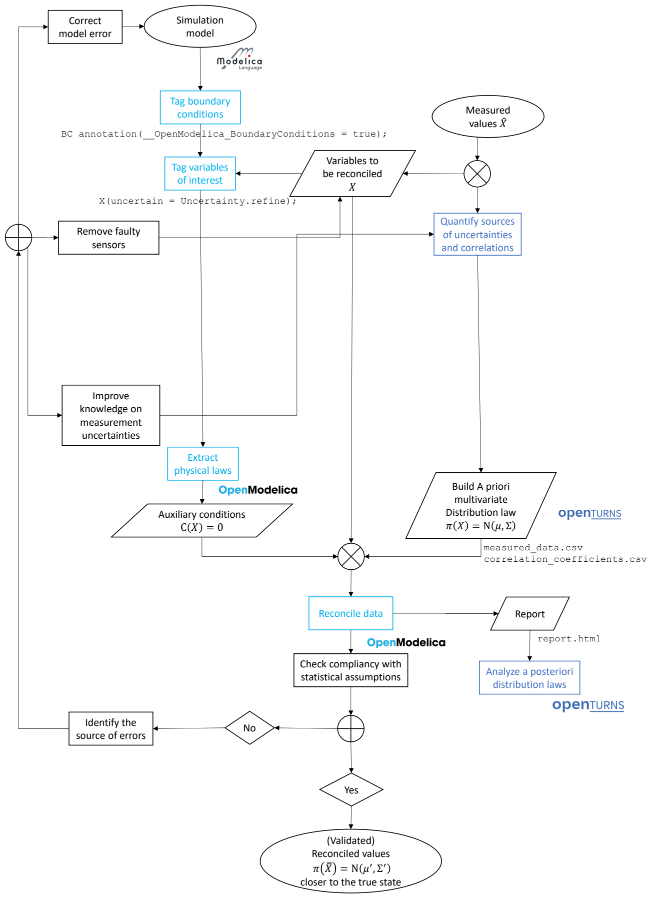


Figure 1. Data reconciliation workflow in OpenModelica

3 Flowrate Measurement with EVEREST Facility

3.1 Description of the test loop

3.1.1 Global Description

EVEREST (Thibert et al. 2019) is a closed loop facility with a liquid flow rate regulation (Table 1). The fluid is clean demineralized water. EVEREST consists of three parts: a flow metering reference section, an operation section and a modular flow metering test section (Figure 2). Table 1 shows the EVEREST design specifications.

The test bench provides a steady water stream with Reynolds numbers as low as 5.10^4 up to 2.10^6. Flow circulation is provided by a variable centrifugal pump (315

Prime features	Details
Calibration test type	Closed regulated loop
Flow meter reference	Master meters method
Test section pipe size	From DN40 up to DN350
Pipe compositions	Stainless steel
Fluid	Demineralised water
Pressure range	From 1 up to 10 bar
Fluid temperature range	From 20°C up to 40°C
Environment Temperature	From 10°C up to 35°C
Flow rate range	From 5 $m^3.h^{-1}$ up to 1200 $m^3.h^{-1}$
Flow metering uncertainty	0.2% (volume and mass flow rate)

Table 1. Overview of the design specification of EVEREST

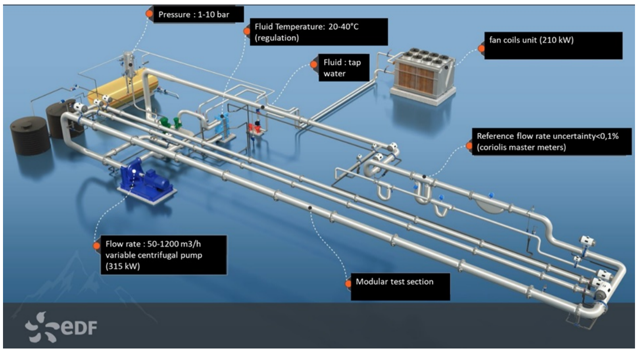


Figure 2. EVEREST test facility layout design

kW). Depending on the geometrical configuration of the test section, the flow can range from 25 $m^3.h^{-1}$ up to 1100 $m^3.h^{-1}$. Thanks to regulation systems, a steady flow rate is maintained in the reference and the test sections. The heat produced by the pump is absorbed by the pumped liquid flowing in the pipes. Due to the EVEREST loop configuration, this phenomenon consequently increases the fluid temperature. To compensate this phenomenon, the water temperature is regulated thanks to a plate heat exchanger (215 kW). The cooling system including the fan coils unit is located at the outside of the experimental hall. The fluid temperature can thus be maintained at a specific value according to the desired test conditions, generally between 20°C and 40°C. The fluid pressure is regulated at a set-point between 1 bar up to 10 bar thanks to a pressuriser.

The EVEREST loop is designed to minimise its environmental impact by managing efficiently its water consumption. Three water storage tanks are used to fill and empty the circuits between two test campaigns requiring different pipe configurations.

3.1.2 Reference Flow Rate

The reference flow rate is calculated from four different master meters based on the Coriolis technology. Two meters are used for low flow rate calibration tests, from $50m3.h^{-1}$ up to $350m3.h^{-1}$. They are installed on a DN100 stainless steel pipe. The other two devices operate under high flow rate calibration tests conditions, from $150m3.h^{-1}$ up to $1200m3.h^{-1}$. They are installed on a DN250 stainless steel pipe. This configuration aims at decreasing uncertainty over the whole loop flow rate rangeability, so each meter is used in its nominal operative flow rate.

For each reference line, one Coriolis meter is used as the reference flow rate while the other is used as a "drift detection" meter. This choice allows monitoring and detecting any reference sensor drift between two consecutive calibrations.

The EVEREST reference flow rate can be calculated either as a volume or as a mass flow rate. The EVEREST test loop has been ISO17025 accredited with a volume (resp. mass) flow rate uncertainty of 0.2% from 25 $m^3.h^{-1}$ (resp. $t.h^{-1}$) up to 1100 $m^3.h^{-1}$ (resp. $t.h^{-1}$).

3.1.3 Modular Test Section

The flow metering test section is located at the end of EVEREST (see Figure 3) and resides completely inside the building. Devices to be tested are mounted on stainless steel pipes, they can be invasive as well as non-invasive according to the sensor technology. This section constitutes an available space with a length of 25 m, a width of 5 m and with a high ceiling of 5 m.

Figure 3. EVEREST test facility test section

This test section is designed to perform calibration tests not only under ideal thermo-hydraulic conditions (fully developed and swirl-free velocity profile thanks to a long straight pipes at the upstream of the tested meter) but also calibration tests under non-ideal thermo-hydraulic conditions due to the presence of various fittings located upstream or downstream of the tested meter.

Summerizing, the test section of the loop allows the recreation of numerous process pipe geometries or to install a reduced scale mock-up of typical industrial components for metrological purposes. This flexibility is a key feature of the loop as it allows EDF R&D to investigate the impact on the accuracy of flowmeters installed in real industrial conditions.

3.2 Detection of Defects

As described in §3.1, the EVEREST test loop is a representative and reliable experimental test facility. Therefore, it is an interesting use case for Data Reconciliation

DOI 10.3384/ecp21845

in order to improve the reliability of the facility's state estimation. Indeed, the test loop being modular, deliberately machined defects can be introduced into the system to impact its behaviour. The defects are introduced upstream the mass flow rate measurement and are shown in Figure 4. They are representative of what could be found in thermal hydraulic installations, such as weld bead defects (trapezoidal or rectangular) or changes in diameter at the junction of two pipes (referred to as step defects). Those defects have more or less impact on the flow and are sorted by order of impact in Figure 4, with the one having no defect showing no impact, and the change in diameter having the greatest impact on the flow. Thus, the goal is to determine whether data reconciliation will allow the detection of these defects.

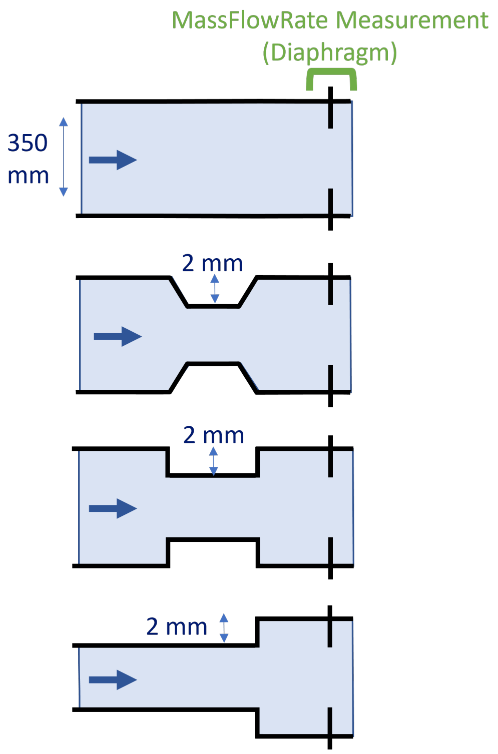


Figure 4. Diagram of the different defects introduced in the EVEREST test loop: no defect, trapezoidal weld bead defect, rectangular weld bead defect, change of diameter (also refered as step defect) sorted by order of impact from the one having the less impact (no defect added), to the one having the more impact on the flow (step defect).

A data reconciliation approach using the OpenModelica prototype is thus developed to study the detection of these defects thanks to an experimental test campaign on the loop. This workflow is inspired from (EPRI 2020). First a Modelica model of the facility is developed using manufacturer's data of the different equipment and calibrated by inversion of n parameters based on n historical measurements from the facility. Initial tests of data reconciliation are run to make sure that the data reconciliation problem is well-posed. Those first steps ensure a good numerical representativity of the installation and that the data reconciliation problem is well-posed. Once the initial phase completed, test campaigns are conducted, both with and without defects, to allow comparison with a control test. The data reconciliation algorithm is run and challenges the prior uncertainties with the posterior results. If the results are significantly different (in the sense of a χ^2 test as described in §2.1), it can be concluded that an inconsistency has been detected: either the model is incorrect or the measurements are false.

A readily available ThermoSysPro (El Hefni and Bouskela 2019) model of EVEREST test loop shown Figure 5 was adopted and calibrated. Test campaigns are then

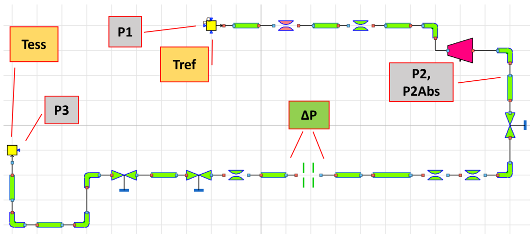


Figure 5. ThermoSysPro model of the EVEREST test loop with the different measurements used (Pressure, Temperature and MassFlowRate).

conducted, starting with a characterisation test, followed by one defect at a time. For each test campaign (each defect), different thermal hydraulic conditions are tested (variation of mass flow rates) both with and without the defect. A test campaign for one defect thus consists of 105 measurement sets (with and without the defect). Effectiveness of defect detection with data reconciliation can be evaluated in many ways. For instance, the global and local tests for data reconciliation can be used in order to identify if a test is detected as faulty by comparing it to a threshold λ as described by Equation 9. If so, the confusion matrix for such a case can be written as shown in Table 2. For industrial use case defect detection, it is desirable to have as few *False Positive (FP)* as possible.

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)

Table 2. Confusion matrix for defect detection

With numerous measurements to evaluate the effectiveness of defect detection using data reconciliation, new indicators can be computed. Sensitivity and Specificity are defined. Sensitivity (or true positive rate) is the probability of detecting a defect for a test containing a defect (Equation 14). Specificity (or true negative rate) is the probability of not detecting any defect for a test that does not contain any (Equation 15). In a perfect scenario, Sensitivity and Specificity should be equal to 1 as all defects would be detected and no false alarms would appear.

$$Sensitivity = \frac{TP}{TP+FN} \quad (14)$$

$$Specificity = \frac{TN}{TN+FP} \quad (15)$$

For each test campaign, the data reconciliation algorithm is applied to each measurement set. As the result of the Sensitivity and Specificity for each set depends on the threshold λ, it is plotted with respect to λ. Starting with the defect having the greatest impact on the flow, Figure 6 shows the Sensitivity and Specificity against the threshold for the step defect. It can be seen that with the default value of $\lambda = 1.96$, this defect can be detected perfectly as Sensitivity and Specificity are equal to 1.

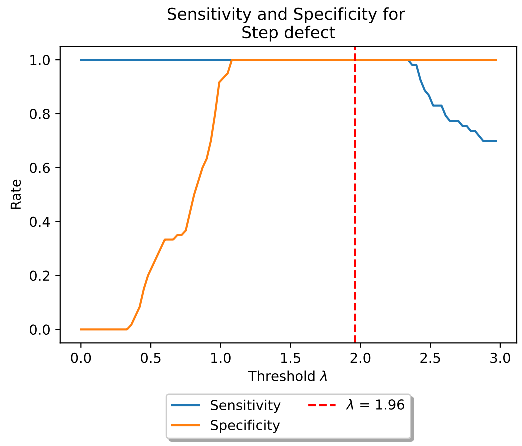


Figure 6. Sensitivity and Specificity against threshold λ for the Step defect. The defect can be detected perfectly with $\lambda_{default} = 1.96$ as Sensitivity and Specificity are equal to 1.

However, the default value of λ does not always lead to a perfect diagnosis. Indeed as shown on Figure 7, the rectangular weld bead defect can be detected perfectly, but not with $\lambda_{default} = 1.96$. Indeed with this default value, the Sensitivity is equal to 0.35. λ should be in the interval $[1.05; 1.35]$ so that the diagnosis is perfect. This difference has an impact on the diagnosis as more *False Negatives* would have been found.

Finally, for the trapezoidal weld bead defect, which has the smallest impact on the flow, it is seen on Figure 8 that the defect can not be perfectly detected as no range of threshold λ gives a Sensitivity and Specificity equal to 1. However the threshold could be reduced so that the Specificity is still high, but that some *True Positives* have been found.

Thus, data reconciliation can detect defects that were intentionally introduced in the system. On the most pronounced defects as the step defect, the impact is such that global and local tests fail with the default threshold value

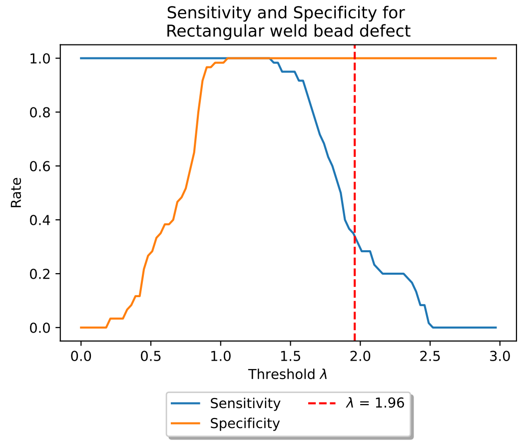


Figure 7. Sensitivity and Specificity against threshold λ for the Rectangular Weld Bead defect. The defect can be detected perfectly, but not with $\lambda_{default} = 1.96$, but with $\lambda \in [1.05; 1.35]$.

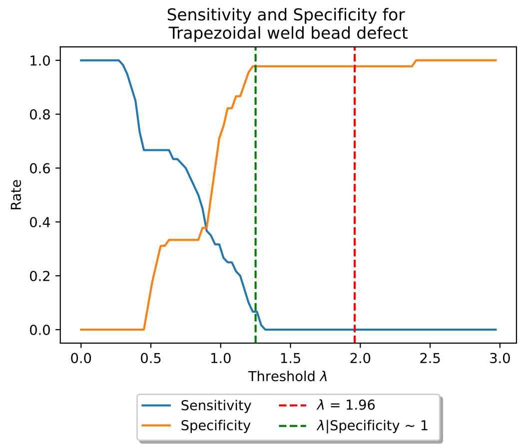


Figure 8. Sensitivity and Specificity against threshold λ for the Trapezoidal Weld Bead defect. The defect can not be detected perfectly.

of $\lambda_{default} = 1.96$. However with less serious defects, $\lambda_{default}$ is not always appropriate. The optimal threshold $\lambda_{optimal}$, depends on the defect that is studied and should be fixed to a value minimising the number of false alarms.

4 ZEPHYR Facility for Testing HVAC Cooling Coil with Latent Exchanges

4.1 In search for good measurement data

The modular test facility ZEPHYR provides air at a precise flow and temperature. On hot days, the air is cooled by a heat exchanger and under humid conditions, condensation can occur. To predict the performance of this heat exchanger at very high temperature and humidity, a validated simulation model is needed. While several theoretical models are available, all of them need "good" input data to calibrate their parameters. Therefore, an experimental test campaign on the cooling coil of the main Air Handling Unit of ZEPHYR and illustrated on Figure 9,

DOI 10.3384/ecp21845

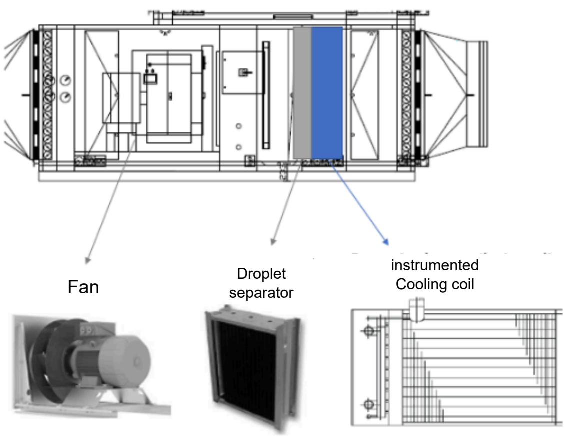


Figure 9. Instrumented cooling coil of Air Handling Unit of ZEPHYR laboratory

was developped to create a set of validated measurement data.

4.2 Experimental campaign

Sensors were installed on the (cool) water side and the (hot) air side of the heat exchanger. On the water side, input and output temperatures as well as the flow are measured. Air measurements at the inlet include temperature and humidity, and at the outlet, temperature and flow. For latent heat, the quantity of condensate is measured. Additionally, the uncertainty of each calibrated sensor has been determined, primarily from data sheets and measured data points. The uncertainty distribution of each sensor can be considered Gaussian. The outlet temperature at one point can be considered representative, because the uniformity of the outlet air temperature was confirmed using an optical fiber installation with approximately 60 measurement points evenly distributed across the entire cross-section.

Over fifteen steady-state experiments were conducted over two days, with varying flows and inlet temperatures (coolant and air). One of these experimental data points is shown in Table 3.

Label	Description	Value	σ	Units
Q_w	Coolant flow	7.91	0.09	kg/s
ΔT_w	Coolant temperature difference	1.62	0.03	°C
Q_a	Air flow	5.82	0.09	kg/s
T_{ai}	Inlet air temperature	29.0	0.30	°C
T_{ao}	Outlet air temperature	20.0	0.47	°C
RH_{ai}	Inlet air relative humidity	42.9	1.5	%
Qc	Condensate flow	1.62	0.03	kg/h

Table 3. Example of sensor data : values and uncertainty (k=1)

It can be observed that the uncertainty on the air side is higher than on the coolant side. One reason for this is that the inlet air was taken from outside the building, causing the temperature and humidity to change gradually throughout the day.

While the amount of condensate at the exit can be measured precisely, the phenomena related to condensation on the plates inside the heat exchanger are complex. Discrepancies, such as the transport time and the thickness of the condensate layer, may lead to invalid experimental data. The idea was to apply data reconciliation to eliminate these invalid experiments and simultaneously improve the accuracy of the measurement set. Since this set contained some level of redundancy, data reconciliation based on basic conservation laws was possible.

4.3 Driving equations

The purpose of this paragraph is to model the thermal exchanges inside an HVAC exchanger that cools humid air coming from outside, as illustrated in Figure 10. The air temperature inside the exchanger can drop below the dew point, caused by the relatively cold surface temperature of the coils, and produce condensation. The main energy balance equations are :

$$\begin{cases} P_w = Q_w(h_{wi} - h_{wo}) \\ P_a = Q_{ai}h_{ai} - Q_c h_c - Q_{ao}h_{ao} \\ 0 = P_w + P_a \\ Q_{ai} = Q_{ao} + Q_c \end{cases} \tag{16}$$

where h denotes the fluid enthalpy, Q the massflowrate and P the thermal power exchanged between the two fluids within the exchanger. The subscript $_a$ denotes air, $_w$ denotes water, $_c$ denotes condensate, $_i$ denotes inlet, and $_o$ denotes outlet. Psychrometric charts of humid air are typically used to visualize the thermodynamic processes occurring within the HVAC cooling coil, as shown in Figure 11.

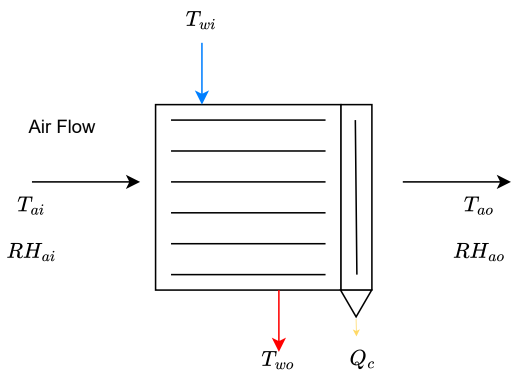


Figure 10. Scheme of the studied HVAC cooling coil with latent exchanges

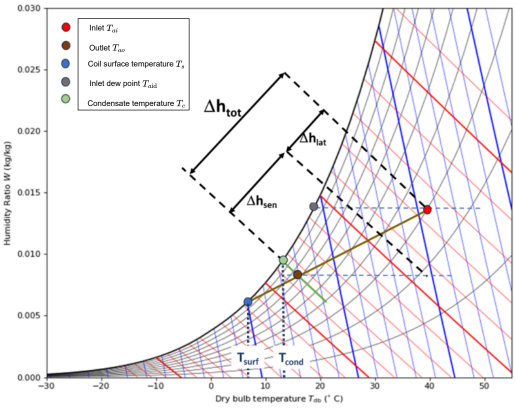


Figure 11. Typical thermodynamic process within an HVAC cooling coil

4.4 Data Reconciliation

Using the conservation equations Equation 16 and the experimental measurements listed in Table 3, it becomes evident that one measurement is redundant and can be deduced from the others. The enthalpy of the humid air and water, assuming the total pressure is equal to atmospheric pressure, is computed using T and RH for air, and T for water, utilizing the media package of the `Modelica Standard Library` as declared below:

Listing 1. HVAC Modelica media

```
replaceable package MediumAir = Modelica.
    Media.Air.MoistAir;
replaceable package MediumWater = Modelica.
    Media.Water.StandardWater;
```

Condensate rate is computed from the other measurements, solving a non linear system to respect mass and energy balances involving condensate rate. Measurements are declared as presented in Listing 2

Listing 2. Declare measurment for data reconciliation

```
/*Measurement*/
Modelica.SIunits.MassFlowRate Qa_mes(
    uncertain=Uncertainty.refine, start=1)
    "Air flowrate [kg/s]";
Modelica.SIunits.MassFlowRate Qe_mes(
    uncertain=Uncertainty.refine) "water
    flowrate [kg/s]";
Modelica.SIunits.Temperature Tai_mes(
    uncertain=Uncertainty.refine) "air
    inlet temperature [K]";
Modelica.SIunits.Temperature Tao_mes(
    uncertain=Uncertainty.refine) "air
    outlet temperature [K]";
Modelica.SIunits.Temperature Twi_mes(
    uncertain=Uncertainty.refine) "water
    inlet temperature [K]";
Modelica.SIunits.TemperatureDifference
    DeltaTw_mes(uncertain=Uncertainty.
    refine) "water temperature rise [K]";
Real RH_ai_mes(uncertain=Uncertainty.refine
    ) "air inlet relative humidity [-]";
Modelica.SIunits.MassFlowRate Q_c(uncertain
    =Uncertainty.refine) "condensate rate [
    kg/s]";
```

In this case, the number of obtained auxiliary conditions r is equal to one. The results of the data reconciliation for two typical experiments are presented on Figure 12 produced by interfacing `OpenModelica` data reconciliation with an OpenTURNS (Baudin et al. 2017) `https://openturns.github.io/` module as illustrated in Listing 3.

Listing 3. Perform data reconciliation with OpenTURNS and OpenModelica

```
covarianceMatrix = priorDistribution
    .getCovariance()
priorMeasurements =
    priorDistribution.getMean()
dataR = otDataR.DataReconciliationOM
    (modelName, VarNames,
    priorMeasurements,
    covarianceMatrix, casePath,
    simuMatPath)
dataR.reconcileData()
reconciledDistribution = dataR.
    getPosteriorDistribution()
```

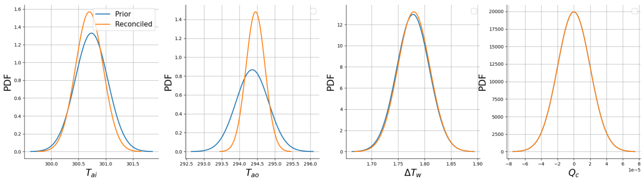
For the first experiment, the data is considered consistent, the result of the local being positive, which is qualitatively in agreement with the prior and posterior Gaussian distribution of the measurements. For the second experiment depicted on Figure 12, the prior and posterior distribution are clearly inconsistent. One can notice that the results of the data reconciliation process indicate that the air temperature measurement may be potentially incorrect, while the condensate rate measurement shows little variation between the prior and posterior distributions. This result is strongly driven by the specified prior distribution. The condensate rate measurement, relying on the mass of the condensate collected over a period, is specified as a measurement with a very low prior uncertainty while some uncertainty sources can have been misidentified due to the complex condensation phenomena occurring within the exchanger. It underlines that output of data reconciliation process has to be carefully analysed to identify correctly defectuous sensors in case of measurements set declared as non consistent.

Measurement with uncertainty before and after reconciliation are illustrated in Figure 13 and Figure 14. It can be observed that experiment number 8 leads to inconsistent measurement after data reconciliation, as the cooling coil outlet state is above the saturation curve. Therefore, this experiment has been discarded.

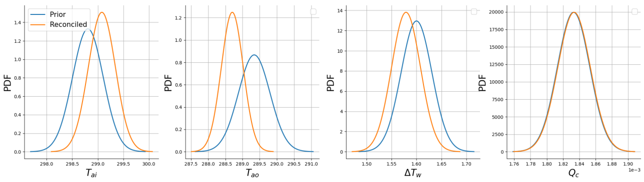
5 Conclusions and future work

The implementation of data reconciliation with OpenModelica has been successfully tested on two industrial experimental installations. For the EVEREST test loop

 DOI 10.3384/ecp21845



(a) Positive local test and improve measurements



(b) False local test for some measurements

Figure 12. Example of reconciled data leading to: improve outlet air temperature measurement (a) - inconsistent measurements (b).

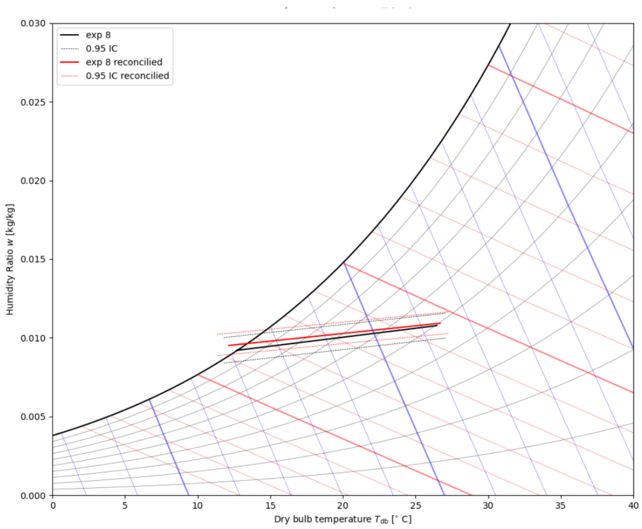


Figure 13. Measurements with uncertainty before and after data reconciliation for experiment number 8

aiming at controlling measurements of liquid flow rates, some defects were intentionally introduced and data reconciliation detected efficiently the most pronounced defects. The detection threshold λ should however be carefully chosen as it has a great impact on the detection results: for less serious defects such as the trapezoidal weld bead defect, the detection is not perfect and *False alarms/No detection signals* will be returned. Further work could be to integrate data reconciliation as a monitoring tool to study sensor drift of the facility in addition to other defect detection methods like comparing a measurement set to a validated reference characterisation test campaign to observe the relative difference through time.

Data reconciliation on the HVAC application for measuring air flows not only improved the accuracy and uncertainty of the measurements but it also identified two initially unnoticed invalid experiments. This high confidence reconciled data set, contains thus "good" measurements that can be used as inputs for parameter calibration of simulation models.

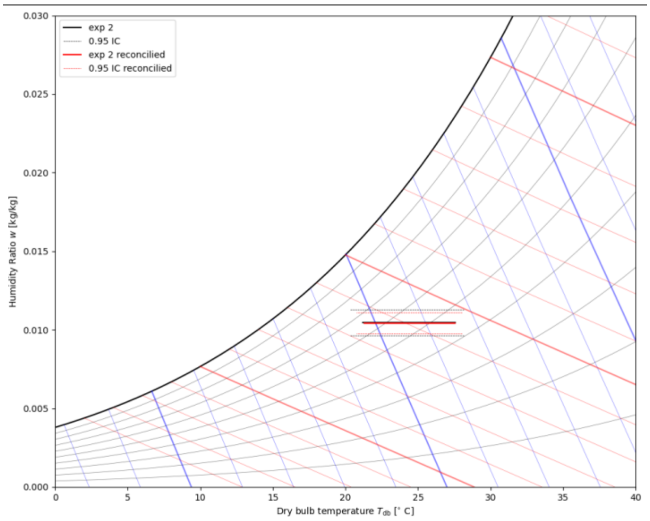


Figure 14. Measurements with uncertainty before and after data reconciliation for experiment number 2

Data reconciliation tools exist and have already proven their worth in the industrial sector. However, they are limited to a restricted modelling domain and often use a dedicated, proprietary modelling environment. In practice, this restricts the scope of reconciliation: apart from use cases where reconciliation is used as a real-time monitoring tool to gain operating margins, few use cases can afford this costly environment. Other applications can nevertheless be imagined: support for experimental installations as presented in this article, support for the design of systems to assist in the positioning and selection of sensors and provision of rigorous proofs to justify the control of uncertainties of sensitive quantities, etc. The integration of data reconciliation into OpenModelica opens up new possibilities: users no longer need to create a dedicated reconciliation model in a dedicated environment; they can derive it from pre-existing models, such as those used for classical simulation studies. The Modelica language is used there as a sole and standardized format for modelling the physical phenomena of the system hence bridging the gap between design tools and operation tools.

From the software standpoint, although the current implementation of data reconciliation in OpenModelica is still in the prototype stage, the article demonstrates its ability to handle industrially realistic models. The prototype documentation will be updated based on the application cases described here, particularly with regard to the methodology to be followed when using third-party tools to calculate uncertainties for pre- and post-processing of data reconciliation. Further improvements to the prototype are also expected to enable a better scaling for larger or more complex models. This will include improvements to the user interface for providing inputs to be reconciled - especially for linking model variables to uncertainty data - and for optimising the visualisation of reconciled outputs.

It will also include adjustments of the extraction algorithm to handle models embedding algebraic parts which cannot be easily inlined and symbolically processed (e.g. if-then-else clauses, function with external codes...).

Acknowledgements

The authors would like to thank Arunkumar Palanisamy, Lennart Ochel and Adrian Pop for their implementation of data reconciliation in OpenModelica. This functionality has been initially supported by DGE (France) and Vinnova (Sweden) in the scope of the ITEA3 OPENCPS project and then improved within the Open Source Modelica Consortium thanks to EDF funding.

References

Baudin, Michaël et al. (2017). "Openturns: An industrial software for uncertainty quantification in simulation". In: *Handbook of uncertainty quantification*. Springer, pp. 2001–2038.

Bouskela, Daniel et al. (2021). "New Method to Perform Data Reconciliation with OpenModelica and ThermoSysPro". In: *Modelica Conferences*, pp. 453–462. DOI: https://doi.org/10.3384/ecp21181453.

El Hefni, Baligh and Daniel Bouskela (2019). *Modeling and Simulation of Thermal Power Plants*. Springer.

EPRI (2020). *Use of Data Validation and Reconciliation Methods for Measurement Uncertainty Recapture*. Last accessed 24 March 2025. URL: https://www.nrc.gov/docs/ML2105/ML21053A030.pdf.

ISO/IEC GUIDE 98-3 (2008). *Uncertainty of measurement - Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)*.

OpenModelica Documentation: Data Reconciliation (2021). URL: https://openmodelica.org/doc/OpenModelicaUsersGuide/latest/dataReconciliation.html (visited on 2025-04-29).

Thibert, Emmanuel et al. (2019). "EDF R&D new test bench for liquid industrial flow meters calibration". In: *FLOMEKO 2019. The 18th International Flow Measurement Conference*.

VDI 2048 (2017). *Blatt 1 "Control and quality improvement of process data and their uncertainties by means of correction calculation for operation and acceptance tests*.