Loughborough University

This item was submitted to Loughborough's Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/

Non-linear Dynamic Data Reconciliation For Industrial Processes

Xuemin Tian, Bokai Xia, Zuojun Yu, Shuang-Hua Yang (SMIEEE)

Abstract—This paper investigates and improves a technique known as Nonlinear Dynamic Data Reconciliation (NDDR) for a real industrial process. NDDRS is a technique for data reconciliation that requires an objective function to be minimised subject to both algebraic and differential, equality and inequality constraints. These constraints are obtained from the mathematical description of the process and ensure that the measurement data can be optimised to conform as closely as possible to the true behaviour of the process. One of the difficulties of using the original NDDR is that a rigorous process dynamic model is required as a constraint. Unfortunately it is very hard to establish a rigorous dynamic model for a complex industrial process, particularly for data reconciliation purpose. A transfer function matrix model has been introduced in this new NDDR method. Therefore the rigorous dynamic model is avoided. The real industrial data from FCCU is used to illustrate the efficiency of the new NDDR method. Copyright © 2006 IEEE.

Keywords: Fluid Catalytic Cracking Unit (FCCU); transfer function; non-linear dynamic data reconciliation.

I. INTRODUCTION

Data processing and reconciliation allows the quality of measurement data from a process to be improved upon, which involves procedures for the reduction of the errors present in the measurements of the process, and for the estimation of those aspects of the process that are unmeasured. The reasons for conducting data processing and reconciliation are that it allows for a more detailed knowledge of the state of the observed process, enabling decisions to be made with the best possible information. An improved knowledge of the state of the process allows for maintaining the process at its optimal level, increasing product quantity or quality and providing better safety and environmental levels.

1.1 Random errors and gross errors

There are two different types of error that can be present in a data set. The first and most common are random errors and the second and less frequent are gross errors. Random errors can come from a number of sources; most commonly, they are the result of measurement error and fluctuation of the process. Random errors are expected to have a normal distribution and an expected value of zero. Errors of this type are the easiest to deal with, and there exist numerous different methods by which they can be minimised. Gross

errors (Ozyurt and Pike, 2004; Bagajewicz, 2000; Amand et al., 2001), sometimes known as systematic errors, result from occurrences such as instrument malfunction or mis-calibration, process leaks or malfunctions, poor sampling and inaccurate transcription of data. Gross errors are usually larger and less frequent than random errors and their expected values are not zero. The majority of data reconciliation procedures are based upon the assumption that errors are random with a zero mean; the presence of gross errors invalidates their statistical basis. For this reason it is necessary for gross errors to be identified and either corrected or removed from the data set for data reconciliation to be used.

1.2 Dynamic processes

Dynamic processes are those in which the process variables change with time. Data reconciliation for these systems is more complex than for steady state systems, primarily due to the fact that the optimisation problem that is being solved contains more complex constraints. Dynamic processes will have more information available than a similarly sized steady state process because of the presence of temporal redundancy in the data between the different sampling times, in addition to the spatial redundancy observed in steady state systems.

Wavelets are an extension of Fourier analysis and allow data to be expressed in terms of averages and differences between the data. Wavelet techniques can be used to reduce the errors in data by taking into account the temporal redundancy in the data. They do not however make use of the spatial redundancy of the data. Additionally, wavelets can be used to represent the differential equation present in the optimisation problem, and therefore the problem can be discretised and converted to a more standard algebraic problem. The application of this technique is discussed in Kong et al. (2002). Another approach taken to conducting dynamic data reconciliation is to use commercial simulation software to produce the model for the process, removing the need for a complex equation based model. Methods can then be applied to this model. This method is discussed by Alici & Edgar (2002). Both Kong et al. (2002) and Alici & Edgar (2002) applied their techniques to a continuous-flow stirred tank reactor system. Both obtained better results than the techniques they were compared with. Unfortunately, these two methods cannot be directly compared as they use different metrics to determine the performance of the techniques and different inputs to the system.

1.3 Objectives of the paper

This paper focuses on the data reconciliation for random

X Tian, B Xia, and Z Yu are with College of Information and Engineering, University of Petroleum, East China, Shangdong, 257061, PR China; e-mail: tianxm@hdpu.edu.cn).
S H Yang is with Computer Science Department, Loughborough University, Loughborough, LE11 3TU, UK; e-mail: s.h.yang@lboro.ac.uk.

5291

errors of dynamic processes, and in particular investigates the application of dynamic data reconciliation techniques for large scale complex industrial processes. The investigation is based on the non-linear dynamic data reconciliation (NDDR) technique. Since NDDR involves a rigorous dynamic model for a complex industrial process the improved NDDR proposed in this work aims to use a simple transfer function model to replace the rigorous dynamic model. The rest paper is organized as follows. Section 2 gives the description of the original NDDR technique. Section 3 presents the application of the original NDDR on a well-known example. Section 4 extends the NDDR technique to real measurement data taken from an industrial Fluid Catalytic Cracking Unit (FCCU) where a rigorous dynamic model is not available. Section 5 is the conclusions.

II. DESCRIPTION OF NDDR

NDDR is a technique for data reconciliation that requires an objective function to be minimised subject to both algebraic and differential, equality and inequality constraints. These constraints are obtained from the mathematical description of the process and ensure that the measurement data can be optimised to conform as closely as possible to the true behaviour of the process. NDDR is applicable to dynamic processes and where data is absent from gross errors. The NDDR approach can be described mathematically as shown in Equation 1 (Liebman et al., 1992).

$$\underset{\hat{y}(t)}{\text{Min}} \ \Phi\left[y, \hat{y}(t), \sigma\right]$$ (1)

s.t.

$$f\left[\frac{d\hat{y}(t)}{dt}, \hat{y}(t)\right] = 0$$

$$h\left[\hat{y}(t)\right] = 0$$

$$g\left[\hat{y}(t)\right] \geq 0$$

Where:
Φ: objective function;
$\hat{y}(t)$: estimate functions
y: discrete measurements;
σ: measurement noise standard deviations;
f: differential equation constraints;
h: algebraic equality constraints;
g: inequality constraints

The nature of the problem for which NDDR is to be applied will determine the nature of the objective function to be used. For this study only a weighted least squares objective function will be considered. The weighted least squares objective function can be described mathematically as shown in Equation 2 (Romagnoli & Sanchez, 2000).

$$\Phi\left[y, \hat{y}(t_k), \sigma\right] =$$
$$\sum_{k=0}^{c} \frac{1}{2} \left[\hat{y}(t_k) - y(t_k)\right]^T \Psi^{-1} \left[\hat{y}(t_k) - y(t_k)\right]$$ (2)

Where:
c: current time
Ψ: variance-covariance matrix where $\Psi_{ii} = \sigma_{ii}^2$
t_k: discrete time instance

Thus, the aim of the NDDR technique is to select the optimum values of $\hat{y}$ for each instance in time so that the value of the objective function is minimised. The values of $\hat{y}$ that can be chosen are constrained by the need for them to satisfy the constraints, which are themselves derived from the description of the process on which the NDDR is being performed.

III. APPLICATION OF NDDR TO A SIMULATED CSTR

This example is selected because it appears relatively frequently in the literature on data reconciliation, and has become to some extent the first call example in the demonstration on the utility of any new technique in dynamic data reconciliation. The example was first used in Liebman et al. (1992), and also appears in Romagnoli & Sanchez (2000), as well as in the works of Alici & Edgar (2002) and Kong et al. (2002).

Fig. 1 shows a CSTR (continuous stirred tank reactor) with a first order, exothermic reaction and defined by a mathematical model shown in Equation 3 (Liebman et al., 1992) with the parameters shown in Table 1.
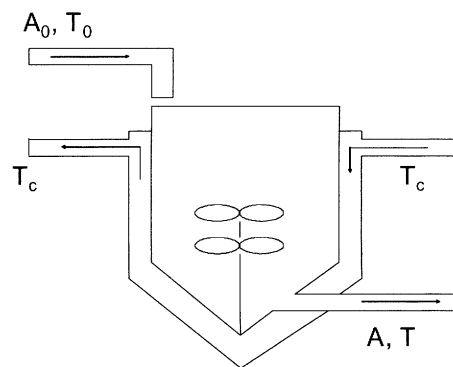

FIG. 1 CSTR

$$\frac{dA}{dt} = \frac{q}{V}(A_0 - A) - KA$$
$$\frac{dT}{dt} = \frac{q}{V}(T_0 - T) + \frac{-\Delta H}{\rho Cp} KA - \frac{UA_R}{\rho CpV}(T - T_c)$$ (3)

Where:

5292

$$K = K_0 e^{\frac{-E_A}{T}}$$

A_0: Feed concentration, gmol cm^{-3}

T_0: Feed temperature, K

A: Tank concentration, gmol cm^{-3}

T: Tank temperature, K

CSTR is modelled over a period of 250 seconds, with samples taken every 2.5 second. The process was initialised at a steady state condition with the following values. A0 = 6.5, T0 = 3.5, A = 0.1531, and T = 4.6091. At a time of 75s, the feed concentration is stepped from 6.5 to 7.5. The simulated values are treated as the process measurements by adding noise to them. The noise added to the simulated values was normally distributed with a mean of zero and a standard deviation of 0.05. The simulated values are denoted as a solid line, the measurements as a dotted line in Fig. 2.

Table 1 Parameters of CSTR

Parameter	Value	Units
q	10.0	$cm^3 s^{-1}$
V	1000.0	cm^3
ΔHr	-27,000.0	cal $gmol^{-1}$
ρ	0.001	g cm^{-3}
C_p	1.0	cal $(g\ K)^{-1}$
U	5.0×10^{-4}	cal $(cm^2\ s\ K)^{-1}$
A_R	10.0	cm^2
T_c	340.0	K
K_0	7.86×10^{12}	s^{-1}
E_A	14,090.0	K
α_d	1.0	-

Using Equation 2 as the optimal objective in NDDR for the above CSTR, NDDR can be straightforwardly applied to CSTR. We implement NDDR in MATLAB by calling an inbuilt MATLAB function fmincon. The fmincon function requires two input functions, one describing the objective function to be minimised, and the other describing the constraints on the minimisation variables.

The results of the reconciliation are shown in Fig. 2. As can be seen, the estimates produced by the reconciliation contain considerably less noise than the measurements on which they were performed.

Table 2 shows numerically the reduction in the standard deviation of the error between the measurements and the estimates, along with a comparison with the reconciliation carried out by Liebman et al. (1992) on the same process under the same conditions. The results are similar to those of Liebman et al. (1992), providing confirmation that the NDDR technique has been correctly implemented in MATLAB. These results were obtained using a moving window of 10 time steps, i.e. 25 seconds.

As can be seen from Table 2, the estimates for the feed concentration (A0) have a large increase in standard deviation. This is caused by the lag, which can be seen from the graph, where the estimates fail to respond quickly to the step change in the feed concentration. This lag comes about due to the assumption that the inputs to the CSTR are constant. The size of the lag is directly related to the length of the moving window. The estimates for the feed concentration outside of the area affected by the lag have a reduced standard deviation, but overall this reduction is masked by the effect of the lag.

IV. EXTENTION OF NDDR TO A REAL INDUSTRIAL PROCESS - FCCU

It will be very difficult to build a rigorous mathematical model for any real industrial processes. By introducing an acceptable step change into the processes it is possible to identify a transfer function model for them. In this section we investigate the efficiency of NDDR by replacing the rigorous mathematical models with transfer function models. An industrial fluid catalyst reactor unit (FCCU) is used as a case study in this investigation.

4.1 FCCU transfer function models

A typical FCC process is shown in Fig. 3. It includes a reactor-regenerator section, main fractionators and gas processing facilities and converts heavy oils into gasoline, light cycle oil (LCO), butanes and gas. The combined feed is preheated against the heavy cycle gas oil pump around from the main fractionators, then combined with regenerated catalyst at the bottom of the riser and then reacts endothermically in the riser to form lighter hydrocarbons. A series of step changes have been introduced into the operation of the FCC process. The responses of step change in the temperature (TIC2201) and the flow rate (FIC2201) of the combined feed are described in the following transfer function model shown in Equation 4. The output variables are the temperatures of the 1st and 2nd regenerators (Treg1, Treg2), the carbon dioxide concentration of the 1st regenerator (C1), and the oxygen concentration of the 2nd regenerator (C2). The input variables are the temperature and the flow rate of the feed (Tf, Ff).

$$G = \begin{bmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \end{bmatrix} \quad (4)$$

where

$$g_{11} = \frac{0.0085(1+s)e^{-12s}}{59.2s^2 + 7.7s + 1} \; ; \; g_{12} = \frac{0.085e^{-22s}}{13.7s^2 7.4s + 1} \; ;$$

$$g_{13} = \frac{0.32e^{-11}}{59.2s^2 + 8.3s + 1} \; ; \; g_{14} = \frac{0.175(1+2s)e^{-13s}}{15s^2 + 5s + 1} \; ;$$

$$g_{21} = \frac{0.024e^{-27s}}{30.9s^2 + 7.8s + 1} \; ; \; g_{22} = \frac{0.085e^{-22s}}{8.16s^2 + 4s + 1} \; ;$$

$$g_{23} = \frac{-0.47(1-15s)e^{-20s}}{100s^2 + 9s + 1} \; ; \; g_{24} = \frac{-0.6e^{-24s}}{30.9s^2 + 7.8s + 1} \; ;$$
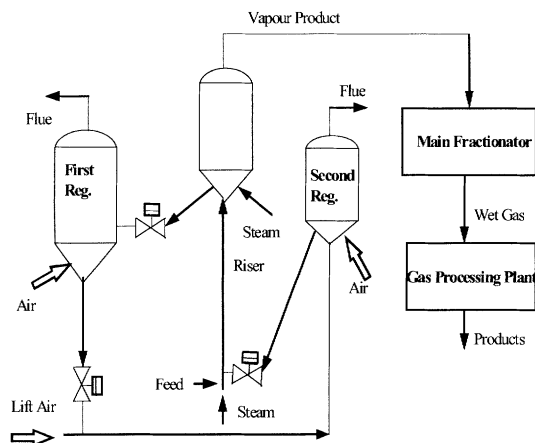
5293

Fig. 3. FCCU reactor-regenerator system (Yang et al., 1996)

Using the above transfer functions to replace the equality constraints shown in Equation 1in NDDR the model constraint is implemented in terms of the MatLab inbuilt function such as ilaplace, which allows the inverse Laplace transform, therefore the data reconciliation is carried out. The results of the reconciliation are shown in Fig. 4 and Table 3. It can be seen that the estimates achieve a good reduction in error during the dynamic portions of the responses. However, outside of these portions the estimates are little more than a smoothing of the measurements. The reason that the reduction in standard deviation is so low for the responses to the transfer functions g23 and g24 is that the measurements of these responses deviate significantly from the response defined by the transfer function, particularly in the later sample times. This causes the values of standard deviation to be dominated by these measurements and results in the masking of any noticeable reduction in the errors. This points out that replacing a rigorous mathematical model with a transfer function model in NDDR is only suitable for the range of the transfer function model.

V. CONCLUSIONS

The NDDR technique described in Section 2 has been successfully implemented in two case studies in Section 3 and 4 and demonstrated to be a capable method of performing data reconciliation on process measurements. Large reductions in the random errors present in the measurements of the process have been made even when the process is in a highly dynamic state.

Although the NDDR technique performed well on a process that was described by a system of differential equations, one of the main weaknesses of the technique was using a rigorous mathematical model as an equality constraint. The way of overcoming this weakness is to use a transfer

function model to replace the rigorous mathematical model in the data reconciliation. It has been shown that the NDDR technique has some utility in a real industrial application by employing a transfer function model, however its application is very much limited within the range of the transfer function models, and has not been ideal outside the range.

VI. REFERENCES

[1] Alici, S. and Edgar, T. F. (2002). Nonlinear dynamic data reconciliation via process simulation software and model identification tools. Industrial Engineering Chemical Research, 41(16), pp. 3984-3992.

[2] Amand, T., Heyen, G. and Kalitventzeff, B. (2001). Plant monitoring and fault detection - Synergy between data reconciliation and principal component analysis. Computers and Chemical Engineering, 25(4-6), pp. 501-507.

[3] Bagajewicz, M. J. (2000). A brief review of recent developments in data reconciliation and gross error detection/estimation. Latin American Applied Research, 30(4), pp. 335-342.

[4] Kong, T. F., Chen, B. Z. and He, X. R. (2002). Wavelet-based regularization of dynamic data reconciliation. Industrial Engineering Chemical Research, 41(14), pp. 3405-3412.

[5] Liebman, M. J., Edgar, T. F. and Lasdon, L. S. (1992). Efficient data reconciliation and estimation for dynamic processes using non-linear programming techniques. Computers and Chemical Engineering, 16(10-11), pp. 963-986.

[6] Özyurt, D. B. and Pike, R. W. (2004). Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes. Computers and Chemical Engineering, 28, pp. 381-402.

[7] Romagnoli, J. A. and Sanchez, M. C. (2000). Data processing and reconciliation for chemical process operations, Academic, San Diego, California.

[8] Yang, S.H., Wang, X.Z. and McGreavy, C. (1996). Multi-variable Coordinated Control System Based on Predictive Control Strategy for FCC Reactor-Regenerator System. Chemical Engineering Science, 51(11), 2977-2982.

5294

Table 2 Results for step change in feed concentration

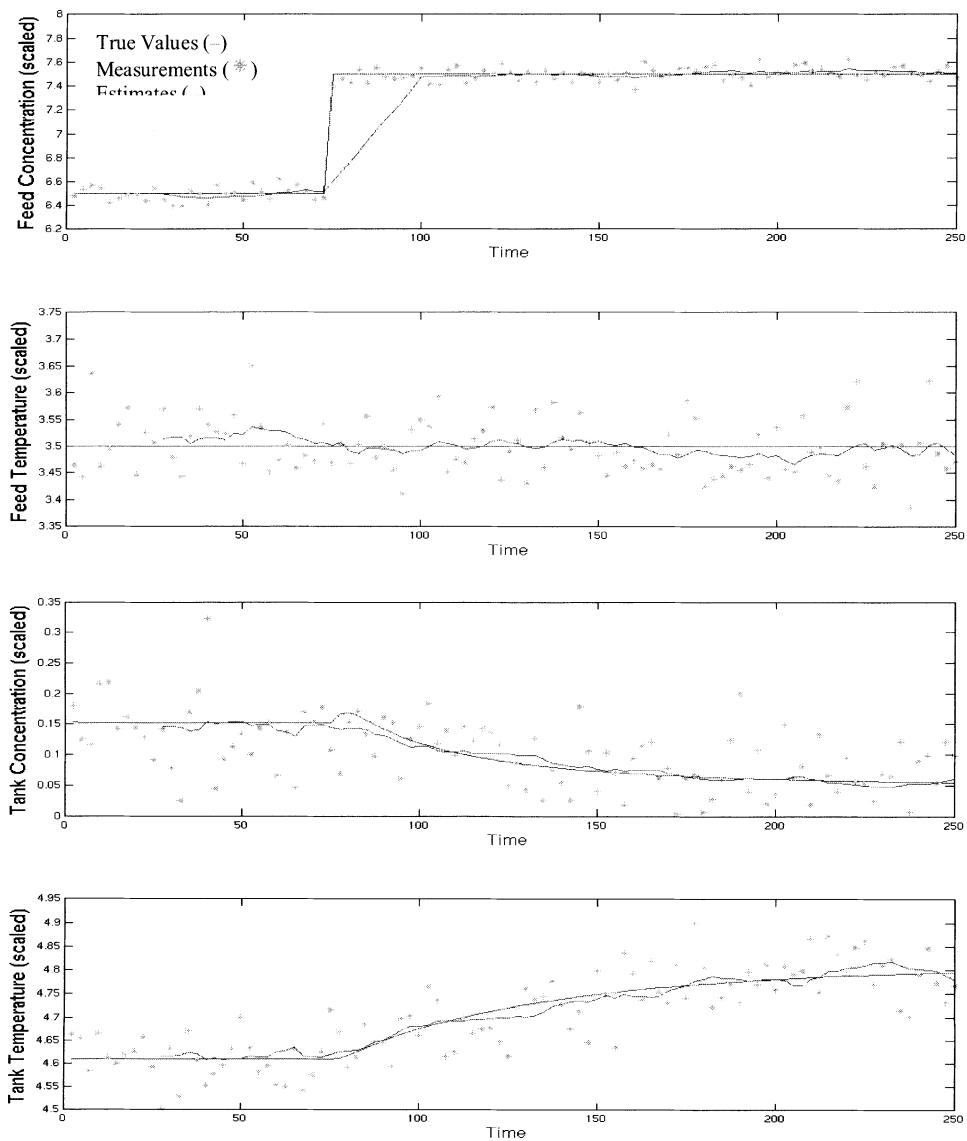| Variable	Measurement Error		Estimate Error		%SD Reduction	%SD Reduction (Liebman, 1992)
	Mean	SD	Mean	SD		
A_0	-3.77×10^{-4}	0.0578	5.63×10^{-2}	0.1833	-217.3	-
T_0	1.84×10^{-3}	0.0509	7.58×10^{-4}	0.0142	72.11	65.7
A	-6.49×10^{-4}	0.0512	1.86×10^{-3}	0.0084	83.58	87.8
T	2.58×10^{-3}	0.0521	-6.96×10^{-4}	0.0121	76.67	77.1

Fig. 2. CSTR response with measurements and estimates

5295

Table 3 Results for reconciliation based on transfer function models

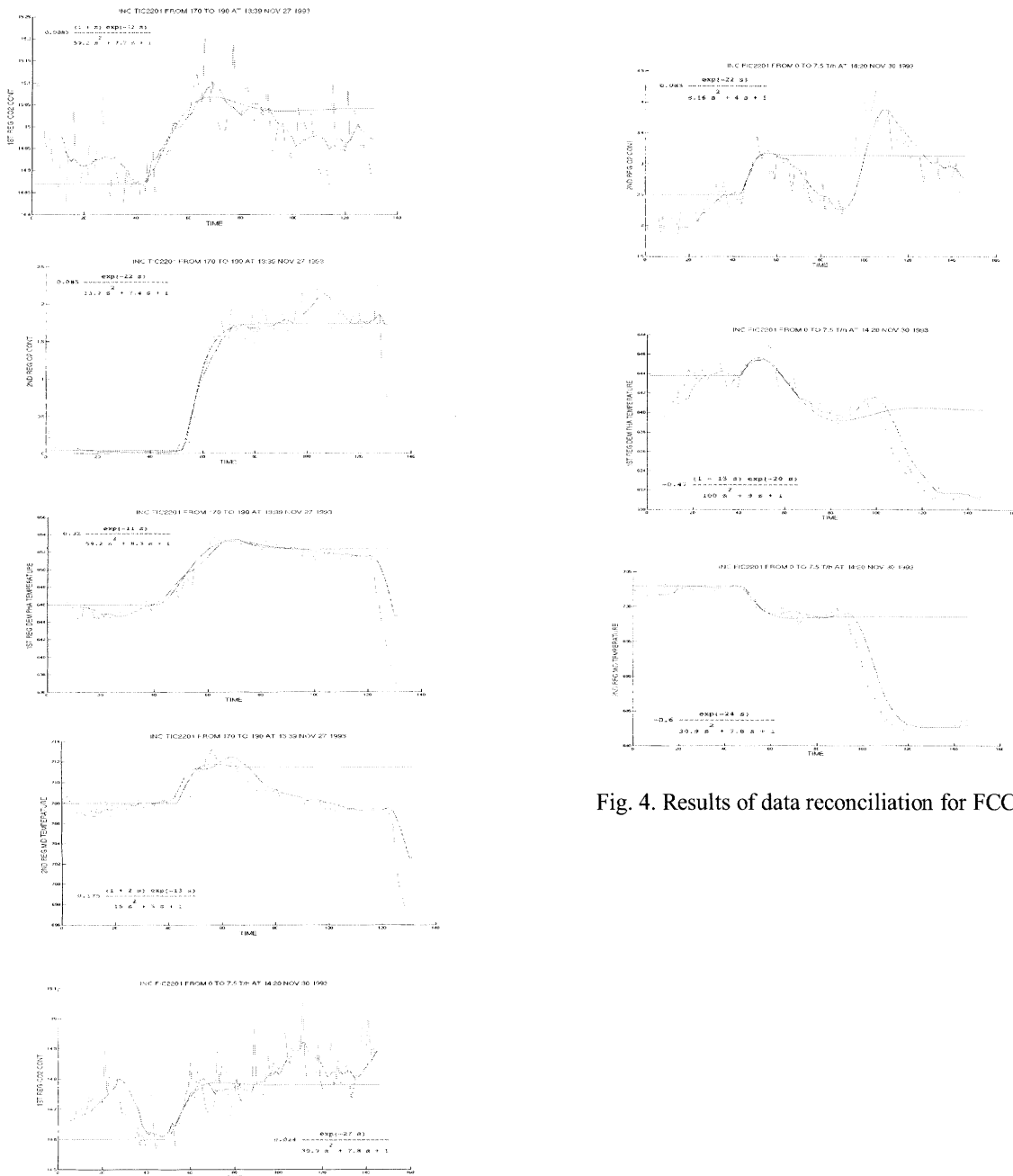| Transfer function	Measurement Error		Estimate Error		%SD Reduction
	Mean	SD	Mean	SD	
g_{11}	-0.0047	0.0797	-0.0077	0.0467	41.41
g_{12}	0.0154	0.2025	0.0256	0.1099	45.73
g_{13}	-1.0757	2.8875	-0.6619	1.3307	53.92
g_{14}	-2.1550	3.1702	-1.7882	2.2082	30.35
g_{21}	0.0563	0.0825	0.0533	0.5860	28.97
g_{22}	-0.1593	0.4470	-0.1448	0.3894	12.89
g_{23}	-2.2802	3.7999	-2.0233	3.5954	5.38
$g24$	-4.8702	6.9560	-4.6273	6.8008	2.23

Fig. 4. Results of data reconciliation for FCCU

5296