

University of Groningen

Data validation and reconciliation for error correction and gross error detection in multiphase allocation systems

Badings, Thom S.; van Putten, Dennis S.

Published in:
Journal of Petroleum Science and Engineering

DOI:
[10.1016/j.petrol.2020.107567](https://doi.org/10.1016/j.petrol.2020.107567)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Badings, T. S., & van Putten, D. S. (2020). Data validation and reconciliation for error correction and gross error detection in multiphase allocation systems. *Journal of Petroleum Science and Engineering*, 195, Article 107567. <https://doi.org/10.1016/j.petrol.2020.107567>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

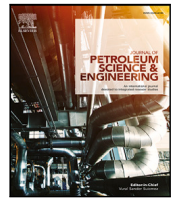
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

journal homepage: www.elsevier.com/locate/petrol

Data validation and reconciliation for error correction and gross error detection in multiphase allocation systems

Thom S. Badings^{a,b}, Dennis S. van Putten^{a,c,*}^a DNV GL, Energieweg 17, 9743 AN Groningen, The Netherlands^b University of Groningen, Faculty of Science and Engineering, Nijenborgh 4, NL 9747 AG Groningen, The Netherlands^c University of Twente, Engineering Fluid Dynamics, Drienerlolaan 5, 7522 NB Enschede, The Netherlands

ARTICLE INFO

Keywords:

Data validation and reconciliation
 Sales allocation
 Gross error detection
 Multiphase allocation systems

ABSTRACT

In upstream oil and gas production, flow measurement errors are common at multiple locations in the production system. These measurement errors lead to imbalances, which are reconciled by using allocation methods. The allocation method should redistribute the imbalance in a fair manner to all involved stakeholders. In this paper, data validation and reconciliation (DVR) is proposed as an alternative allocation method in multiphase production systems. DVR is a model-based optimization method that exploits redundancies in process data to minimize random measurement errors, and simultaneously provides a basis for the detection of gross errors. To study the applicability of DVR compared to conventional allocation methods, a model for generic multiphase production systems is developed, and the corresponding DVR problem is formulated. In order to deal with nonlinear flow effects (e.g. interphase mass transfer) the model is linearized around the operating conditions and the DVR problem is solved iteratively. From the simulation studies, it is concluded that DVR reduces measurement errors by up to 56% for gas allocation systems and 33% for multiphase systems. Gross errors in interior locations of the network can be located precisely, while allocating gross errors to exterior locations is generally not possible, due to limited network redundancy. Compared to conventional methods, DVR provides a better means for the detection of gross errors, and results in more accurate attribution of oil/gas ownership and more fair division of revenues between stakeholders.

1. Introduction

Over the past decades, oil and gas production systems have become increasingly complex in terms of geography, topology, involved stakeholders, and flow compositions (Wee, 2014). In order to reduce capital expenditure, oil and gas producers construct shared production or transport facilities and use joint flow metering for different wells (Johnsen and Dahl, 2018; Stewart and Skelton, 2004). Furthermore, fields can span over multiple concessions, and pipelines can be shared between stakeholders (Kanshio, 2020). Global production of natural gas is estimated at 4000 BCM (Alverà et al., 2018), while yearly growth rates are still around 1.5–2.0% (Bahadori, 2014). As a result, accurate sales allocation and flow metering are crucial to ensure fair division of revenues between the involved stakeholders.

The goal of the allocation process is to determine in the most fair manner the quantities of oil, gas, or other material produced and transported over a given period for each contributing producer (Amin, 2016; Kanshio, 2020). The resulting allocation should be auditable and defensible, and minimizes disputes between stakeholders in a production

agreement (Pobitzer et al., 2016). The problem of allocation is faced in many industries, including oil and gas production systems (Bjørk et al., 2016), but also in power systems (Cherukuri and Cortes, 2017; Gil et al., 2005) and process facilities (Abu-El-Zeet et al., 2002; Meyer et al., 1993). As allocation of gas ownership is based on flow measurements, performance of metering equipment has direct financial implications (Wee, 2014). However, flow meters in production systems often show large measurement uncertainty, which leads to significant measurement errors and unfair sales allocation (Chebiyyam, 2010; Stewart and Skelton, 2004). This discrepancy between the observed and true flow values caused by measurement errors is the fundamental issue of allocation (Pobitzer et al., 2016). As illustrated in Fig. 1, measurement errors cause the sum of the measured inlet mass flows to differ from the reading of the output (reference) meter. Hence, the observed measurements do not satisfy theoretical mass balances, and sales allocation cannot be based directly on flow meter measurements.

Improving the performance of flow meter equipment and obtaining measurements at every valuable position in a production system is

* Corresponding author.

E-mail addresses: thombadings@gmail.com (T.S. Badings), d.s.vanputten@gmail.com (D.S. van Putten).¹ For brevity, phase superscripts *c* are omitted.

Abbreviations

BDA	By-difference allocation
CT	Custody transfer
DVR	Data validation and reconciliation
PC	Principal component
PRA	Pro-rata allocation
UBA	Uncertainty-based allocation
VFMS	Virtual flow measurement system
WHP	Wellhead platform

Nomenclature¹

α	Confidence level (also allocation factor)
Δe	Error reduction
Δr	Fractional reconciliation factor
Δy	Measurement adjustment
δ	Gross error
γ	Global test statistic
\hat{y}	Reconciled variable
C	Set of phases in system
\mathcal{P}	Set of pipes
\mathcal{T}	Set of intersections
μ, η	Convergence limit parameter
ϕ, ψ	Convergence criterion
σ	Absolute standard deviation (uncertainty)
σ^*	Relative standard deviation (uncertainty)
ε	Random error
A	System constraint matrix
A^P	Pipe incidence matrix
A^T	Intersection incidence matrix
e	Measurement imbalance (residual)
f	Interphase mass transfer coefficient
H	Covariance matrix of residual
J	Objective function
n_e	Number of edges
n_t	Number of nodes
p	PC test statistic
Q	Variance-covariance matrix
r	Reconciliation factor
S^*	Relative sensitivity
W	Eigenvector column matrix
y	Measured variable
y^*	True/optimal value of variable
Z	Critical value of standard normal distribution
z_i	Penalty factor of variable i
$z_{e,i}$	Nodal test statistic of constraint i

highly expensive or can even be physically impossible and is, therefore, not a feasible solution to the allocation problem (Pobitzer et al., 2016; van Putten et al., 2019). As a result of the increasing complexity, the costs associated with keeping the total system imbalance within a certain specification (through meter calibrations or frequent in-field inspections) are increasing, and smarter, digital solutions are demanded.

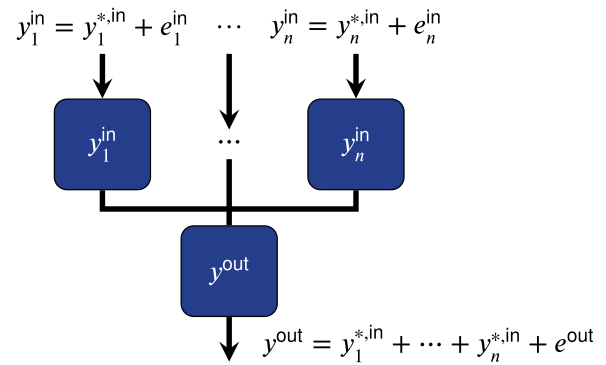


Fig. 1. Single-tier sales allocation problem, with n producing wells and one output. The figure shows the measurement errors (e) caused by discrepancies between the true value (y^*) and measured value (y) of individual flow meters.

1.1. Allocation methods

Allocation methods are such digital solutions which are applied to flow measurement data of production systems, to minimize measurement errors and optimize the resulting sales allocation (Kanshio, 2020; Stockton and Spence, 2008). Proportional allocation methods, such as by-difference allocation (BDA) and pro-rata allocation (PRA), are widely used to mitigate measurement imbalances in oil and gas production systems (Amin, 2016; Bjørk et al., 2016; Lunde et al., 2002; Stewart and Skelton, 2004). For a production system with n_e flow meters, define y , \hat{y} , and $y^* \in \mathbb{R}^{n_e}$ as the vectors of observed flow measurements, corrected (allocated) flow values, and true flow values, respectively. A functional allocation method reduces the measurement errors present in the data, and is formulated as

$$\hat{y} = f_{\text{allocation}}(y, \dots) \text{ such that } |\hat{y} - y^*| \leq |y - y^*|, \quad (1)$$

where the dots indicate that other input data can also be considered, such as pressure, temperature, and measurement uncertainty data (Couput et al., 2017; Webb et al., 2002). A perfect allocation method fully mitigates any measurement error present in the flow data, such that the allocated values approximate the true values as follows:

$$\hat{y} = f_{\text{allocation}}(y, \dots) \approx y^*, \quad (2)$$

In BDA, sales allocation is done by attributing the complete imbalance in a single measurement location, such that the production balance of all fields involved is always satisfied (Bjørk et al., 2016; Stewart and Skelton, 2004). Compared to the perfect allocation method in Eq. (2), BDA shows limited performance, as it assumes that only one variable contains a measurement error. Using PRA, the observed flow measurements are adjusted proportional to the absolute flow, such that the balance is satisfied (Lunde et al., 2002; Amin, 2016). Third, UBA adjusts the observed measurements proportional to the absolute uncertainty of metering equipment (Stockton and Spence, 2008; Webb et al., 2002).

For all conventional allocation methods, the resulting sales allocation satisfies the expected mass balances. However, the methods are limited in terms of error correction performance, and can only be applied to the single-tier system shown in Fig. 1. Furthermore, although UBA methods can reduce random measurement errors, the detection of gross (or systematic) errors remains difficult. Consequently, the accuracy and versatility of conventional allocation methods is limited, justification of the resulting allocation is difficult, and more sophisticated methods are demanded.

1.2. Research contributions

More recently, optimization-based methods, such as data validation and reconciliation (DVR) have been proposed to solve allocation

problems in oil and gas production systems (Couput et al., 2010; Narasimhan and Jordache, 2000; Oliveira and Aguiar, 2009). The method has a long history in process industry (Abu-El-Zeit et al., 2002; Meyer et al., 1993), where it is often simply referred to as data reconciliation. DVR is a model-based optimization method that exploits redundancies in process data to minimize measurement errors (Amin et al., 2016; Kanshio, 2020). The method takes into account the measurement uncertainty in the overall balance of the production system and leads to the detection of measurements that are performing out-of-spec (Bagajewicz and Jiang, 2000; van Putten et al., 2019).

In this work, DVR is proposed as a more versatile allocation method that can be used under much more general conditions than conventional methods and provides a means for the detection of gross errors. Although the use of DVR in the natural gas industry was motivated, e.g. by Oliveira and Aguiar (2009), specific application in oil, gas or multiphase allocation systems is still very limited. Consequently, there is little knowledge on the applicability and limitations of DVR for error detection and correction in multiphase allocation systems. Therefore, the main contributions of this paper are threefold:

1. A model for multiphase production systems is developed, which takes into account interphase mass transfer effects.
2. The applicability and limitations of DVR compared to conventional allocation methods are studied, for the correction of measurement errors in multiphase allocation systems.
3. It is demonstrated how gross error detection methods can be used in combination with DVR for the detection of gross errors in multiphase allocation systems.

In order to achieve these research contributions, the remainder of this paper is organized as follows:

- In Section 2, the model for multiphase production systems is developed, which is used as input model for the DVR method. Nonlinear transfer effects are quantified using the software Aspen HYSYS, by linearization of the production system around the operating conditions.
- In Section 3, random and gross measurement errors are discussed, and DVR is formally introduced and compared analytically to BDA, PRA and UBA, to study the similarity, applicability and limitations of the methods. In addition, a functional comparison of the different allocation methods is provided in tabular format.
- In Section 4, the concept of network redundancy is introduced, and statistical tests for the detection of gross errors are discussed.
- In Section 5, simulation studies are performed to demonstrate the performance of DVR and the error detection methods for the detection of gross errors in multiphase allocation systems.
- Finally, in Section 6, conclusions of the current work are provided, and future perspectives are discussed.

In this paper, calligraphic letters, e.g. \mathcal{A} , are reserved for sets, vectors are denoted by bold letters (e.g. \mathbf{y}), and capital letters are reserved for matrices (e.g. A). $\mathbb{1}_n \in \mathbb{R}^{n \times 1}$ denotes a column vector of ones, and I_n an identity matrix of size n . Furthermore, hat-notation (\hat{y}_i) denotes the reconciled flow value, normal notation (y_i) the observed measurement value, and star superscripts (y_i^*) the true value. A full list of symbols is provided in the Nomenclature section.

2. Modeling of multiphase production systems

A typical multiphase production system has a converging structure towards the custody transfer (CT) point (Harrison and Safar, 2004). The inputs of the system are comprised of individual wells, which are connected through wellhead platforms (WHPs) to one or a few CT points. In this section, a model for generic multiphase production systems is developed. First, a linear flow network model is developed based on graph theory. Thereafter, a method to take into account multiphase flow effects is proposed.

2.1. Linear multiphase flow network model

In this work, production systems are modeled as the directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, consisting of a set of $\mathcal{V} = \{1, \dots, n_v\}$ nodes and $\mathcal{E} = \{1, \dots, n_e\}$ directed edges (Mesbahi and Egerstedt, 2010). The set of nodes includes pipelines and intersections, while the edges represent mass flow streams. Pipes can only be connected to intersections and vice versa, and mass flow is only possible between objects that are physically connected. The side of an edge connected as inlet stream to a pipe or intersection is defined as the positive end, while the outlet is the negative end. Using this notation, any generic production system is characterized by a set of \mathcal{V} nodes and \mathcal{E} edges. The set of nodes is the union of pipelines and intersections, i.e. $\mathcal{V} = \mathcal{P} \cup \mathcal{T}$, where $\mathcal{P} = \{1, \dots, n_p\}$ is the set of pipes, and $\mathcal{T} = \{1, \dots, n_t\}$ the set of intersections. As a node is either a pipe or an intersection, the total number of nodes is the sum of the numbers of pipes and intersections, i.e. $n_v = n_p + n_t$.

Denote by C the set of phases associated with the multiphase production system. The production system is characterized by system matrix $A^c \in \mathbb{R}^{n_v \times n_e}$ for each phase $c \in C$. In particular, the phase-specific nodal mass balances are written as

$$A^c \mathbf{y}^{c*} = \begin{bmatrix} A^{P,c} \\ A^T \end{bmatrix} \mathbf{y}^{c*} = 0, \quad \forall c \in C, \quad (3)$$

where $\mathbf{y}^{c*} \in \mathbb{R}^{n_e}$ is the vector of true flow values of phase c , and $A^{P,c} \in \mathbb{R}^{n_p \times n_e}$ and $A^T \in \mathbb{R}^{n_t \times n_e}$ are the phase-specific system matrices associated with pipes and intersections, respectively. Note that $A^{P,c}$ is unique for each phase $c \in C$, while A^T is equal for all phases. As pipes have exactly one inlet (positive end) and one outlet (negative end), each row of $A^{P,c}$ has two non-zero elements. In particular, matrix $A^{P,c}$ is the directed incidence matrix associated with all pipes in graph \mathcal{G} (Bullo, 2018), with the entries $a_{i,j}^{P,c}$ for phase c between pipe i and edge j defined as follows:

$$a_{i,j}^{P,c} = \begin{cases} f_i^c & \text{if } i \text{ is connected to the positive end of } j \\ -1 & \text{if } i \text{ is connected to the negative end of } j \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where f_i^c is the interphase mass transfer coefficient associated with phase c in pipe i , which is defined explicitly in Section 2.2. For convenience, the edges connected to pipe i are always adjacent in measurement vector \mathbf{y} , thus simplifying Eq. (4) to the block-diagonal matrix defined by

$$A^{P,c} = \text{diag} \left([f_i^c, -1]_{i \in \mathcal{P}} \right). \quad (5)$$

A typical production system has a converging structure, with many inputs and only a few outputs. Although systems with two outputs can also be encountered, for this work it is assumed that all intersections have multiple inputs, but exactly one output.¹ Therefore, each row of the incidence matrix A^T has multiple positive entries and exactly one negative entry. In addition, the volume at intersections is assumed to be zero, and any transfer effects are negligible. Consequently, the entries $a_{k,j}^T$ for intersection k and edge j are given by

$$a_{k,j}^T = \begin{cases} 1 & \text{if } k \text{ is connected to the positive end of } j \\ -1 & \text{if } k \text{ is connected to the negative end of } j \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

Note that Eq. (6) is defined similarly to Eq. (4), but with all transfer coefficients $f_j^c = 1$, which implies that mass transfer effects are negligible at intersections. Hence, the term $A^T \mathbf{y}^{c*} = 0$ in Eq. (3) implies that, for each phase at any intersection, the outlet mass flow is equivalent to the sum of all inlet flows.

¹ Note that this assumption is not a restriction of the modeling framework, as it is easily modified to consider multiple outputs.

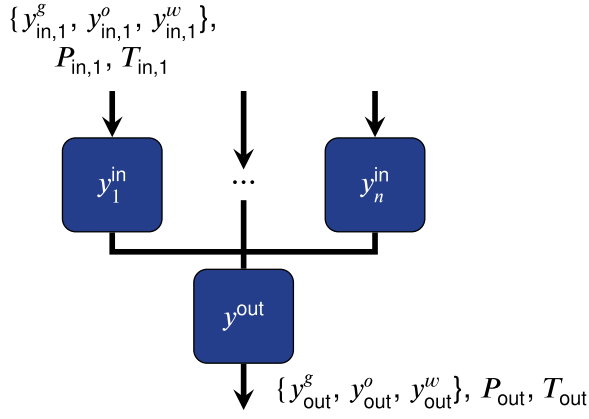


Fig. 2. Multiphase flow in a single-tier gas, oil (condensate), and water production system.

2.2. Nonlinear multiphase flow effects

The interphase mass transfer coefficients f_i^c in Eq. (5) characterize multiphase flow effects caused by temperature and pressure variations in the network. To illustrate this, consider the single-tier production system in Fig. 2 of n producing wells with phases $C = \{g, o, w\}$, which correspond to gas, oil, and water, respectively. Due to multiphase flow effects and mixing of streams, an increase in flow of one phase of a specific input does generally not cause the same increase in combined output flow (Bikmukhametov and Jäschke, 2020). Although this effect is nonlinear, it can be linearized around the operating point, which is characterized by the mass flow, composition, pressure, and temperature in each stream. In particular, the interphase mass transfer coefficient, f_i^c , in Eqs. (4) and (5) for phase c in inlet pipe i , associated with flow variable $y_{in,i}^c$, is calculated as

$$f_i^c = \frac{\partial y_{out}^c}{\partial y_{in,i}^c}, \quad (7)$$

where y_{out}^c is the outlet flow of the connected intersection. In this work, multiphase flow effects in production systems are simulated and quantified in the chemical process simulation software Aspen HYSYS. Based on the thermodynamic conditions, fluid compositions and flow data, HYSYS sequentially solves for the flow changes in Eq. (7) for each phase c in each pipe i , and the interphase mass transfer coefficients are calculated.

3. DVR for multiphase allocation problems

In this section, measurement errors are discussed, and the DVR method is formally introduced. Thereafter, DVR is compared analytically to conventional allocation methods, and a functional comparison of the methods is provided. Finally, an iterative DVR implementation for multiphase allocation systems is proposed.

3.1. Measurement errors and imbalance

All measurements in industrial processes are subject to random and nonrandom errors (Cámara et al., 2017). Random errors include natural disturbances, sampling errors, and unreliable sensor readings, and are typically considered independent with zero mean and thus irreproducible (Benqilou, 2004; Taylor and Del Pilar Moreno, 2013). In oil and gas production systems, random errors are induced by flow measurement equipment, which are assumed independent of each other. Therefore, random errors in this work are also considered independent. In contrast, nonrandom errors, also called systematic or gross errors,

include sensor bias or failure, process leaks, and modeling errors (Drysdale et al., 2011; Hodouin and Everell, 1980). Gross errors can be related to instrument performance (e.g. measurement bias or miscalibration), but can also be constraint model-related (e.g. unaccounted loss of material) (Narasimhan and Jordache, 2000).

The effects of random and gross errors on the set of observed flow measurements $\mathbf{y} \in \mathbb{R}^{n_e}$ as in Eq. (3) (for brevity, phase superscripts c are omitted here) are both modeled as additive contributions to the true value, denoted by \mathbf{y}^* :

$$\mathbf{y} = \mathbf{y}^* + \boldsymbol{\varepsilon} + \boldsymbol{\delta}, \quad (8)$$

where $\boldsymbol{\varepsilon}, \boldsymbol{\delta} \in \mathbb{R}^{n_e}$ reflect the random and gross errors, respectively. While the gross error has a specific value, the random error follows a normal distribution with zero mean, i.e. $E(\varepsilon_i) = 0$, and $\text{var}(\varepsilon_i) = E(\varepsilon_i^2) = \sigma_i^2$, with σ_i the true standard deviation of variable i . In this paper, errors are assumed to be uniform with Gaussian distribution, so the standard deviation also equals the standard uncertainty. The measurement imbalance (constraint residual) for the nodal mass balance constraint in Eq. (3) is given by

$$\mathbf{e} = \mathbf{A}\mathbf{y} = \mathbf{A}(\mathbf{y}^* + \boldsymbol{\varepsilon} + \boldsymbol{\delta}), \quad (9)$$

where $\mathbf{e} \in \mathbb{R}^{n_v}$ is the measurement imbalance vector. As the theoretical mass balance is always satisfied, i.e. $\mathbf{A}\mathbf{y}^* = 0$, and $E(\boldsymbol{\varepsilon}) = 0$, the expected value of the measurement balance defined in Eq. (9) is written as

$$E(\mathbf{e}) = \mathbf{A}\boldsymbol{\delta}. \quad (10)$$

It is important to note that only the measured variable \mathbf{y} is known, while the true value, random error, and gross error are all unknown. However, the distribution of the random error can in general be approximated, e.g. using historical data.

3.2. Data validation and reconciliation (DVR)

Steady-state DVR, or simply DVR, is a model-based data correction technique that makes use of redundancies in process data to reduce measurement errors (Bikmukhametov and Jäschke, 2020; Câmara et al., 2017). In general, it is formulated as the constrained weighted least-square optimization problem, which results from the maximization of the Gauss likelihood function (Amin et al., 2016; Benqilou, 2004; Szega, 2018). The DVR problem for phase c in the production system model is formulated as follows:

$$\text{minimize}_{\hat{\mathbf{y}}^c} \left[(\hat{\mathbf{y}} - \mathbf{y})^\top \mathbf{Q}^{-1} (\hat{\mathbf{y}} - \mathbf{y}) \right]^c \quad (11a)$$

$$\text{subject to } \mathbf{A}^c \hat{\mathbf{y}}^c = 0, \quad (11b)$$

where the notation $[\cdot]^c$ implies that all variables are with respect to phase c . Furthermore, $\mathbf{Q}^c \in \mathbb{R}^{n_e \times n_e}$ is the variance-covariance matrix, and $\mathbf{A}^c \in \mathbb{R}^{n_v \times n_e}$ is the system constraint matrix given by Eq. (3). Assuming that no systematic error is present in the system, the covariance between any two terms is zero (Oliveira and Aguiar, 2009), and

$$\mathbf{Q}^c = \left[\text{diag}(\sigma_1^c, \dots, \sigma_{n_e}^c) \right]^2, \quad (12)$$

where σ_i^c is the absolute standard deviation of measurement i in phase c (Jiang et al., 2014). The analytical solution to Eq. (11), as given by Narasimhan and Jordache (2000), is

$$\hat{\mathbf{y}}^c = \left[\mathbf{y} - \mathbf{Q}\mathbf{A}^\top \mathbf{H}^{-1} \mathbf{A}\mathbf{y} \right]^c, \quad (13)$$

where covariance matrix $\mathbf{H}^c \in \mathbb{R}^{n_v \times n_v}$ of the measurement imbalance in Eq. (9) is written as

$$\mathbf{H}^c = \text{cov}(\mathbf{e}^c) = \left[E\{(\mathbf{A}\mathbf{e})(\mathbf{A}\mathbf{e})^\top\} \right]^c = \left[\mathbf{A}\mathbf{Q}\mathbf{A}^\top \right]^c. \quad (14)$$

The cumulative error reduction is denoted by $\Delta e \in \mathbb{R}$, and is calculated as

$$\Delta e = \frac{\sum_{i \in \mathcal{E}, c \in C} |\hat{y}_i^c - y_i^{c*}|}{\sum_{i \in \mathcal{E}, c \in C} |y_i^c - y_i^{c*}|} - 1. \quad (15)$$

Cost matrix Q^c contains absolute uncertainty values of the corresponding variables. Alternatively, the uncertainty of stream i is expressed as the relative fraction of the absolute measurement, i.e. the relative uncertainty σ_i^{c*} :

$$\sigma_i^{c*} = \frac{\sigma_i^c}{y_i^c}. \quad (16)$$

The reconciliation factor of variable i is defined as the ratio between the optimal reconciled value to Eq. (11) and the original measurement:

$$r_i^c = \frac{\hat{y}_i^c}{y_i^c}. \quad (17)$$

Furthermore, denote by $\Delta r_i^c = r_i^c - 1$ the reconciliation adjustment factor from the original measurements. Finally, the relative sensitivity S_i^{c*} of the output flow to the inlet flow measurement i is defined as follows:

$$S_i^{c*} = \frac{\partial \ln y_{\text{out}}^c}{\partial \ln y_i^c} = \frac{y_i^c}{y_{\text{out}}^c} \frac{\partial y_{\text{out}}^c}{\partial y_i^c} = \frac{f_i^c y_i^c}{y_{\text{out}}^c}, \quad (18)$$

where the latter relation follows from Eq. (7).

3.3. Analogy between DVR and conventional allocation methods

Depending on the conditions and input data, DVR shows similarities with conventional allocation methods. As outlined in Section 1.2, a number of propositions and corollaries are presented, to compare DVR analytically with BDA, PRA, and UBA methods. The results provide valuable information about the characteristics, advantages and limitations of DVR as allocation method on multiphase production systems. For brevity, phase superscripts c are omitted in this section, as the analogy is equivalent for all phases.

Proposition 1. Consider the single-phase, single-tier allocation problem in Fig. 1 with input flows $\mathcal{N} = \{1, \dots, n\}$ and single output, where for a single input i , the relative uncertainty $\sigma_{in,i}^* \gg \sigma_{in,j}^* \forall j \in \mathcal{N}, j \neq i$, while $\sigma_{out}^* \approx 0$. Then, application of DVR reduces to the BDA method, and the complete error is allocated to stream i .

Proof. The general solution to the DVR problem for this single-tier system is given by Eq. (13), with $\mathbf{y} \in \mathbb{R}^{n+1}$, cost matrix $Q \in \mathbb{R}^{(n+1) \times (n+1)}$ and row vector $A = \mathbf{a} \in \mathbb{R}^{1 \times (n+1)}$ defined as

$$\mathbf{a} = [\mathbf{1}_n^T, -1]. \quad (19)$$

As cost matrix Q is diagonal, it follows from Eq. (19) that $H = AQA^T = \text{tr}(Q)$, and the solution in Eq. (13) is written in terms of the vector of measurement adjustments:

$$\Delta \mathbf{y} = \hat{\mathbf{y}} - \mathbf{y} = -QA^T H^{-1} \mathbf{a} \mathbf{y} \quad (20a)$$

$$= -Q[\mathbf{1}_n^T, -\mathbf{1}_m^T]^T \text{tr}(Q)^{-1} \mathbf{e} \quad (20b)$$

$$= -e \left[\sigma_{in,1}^2, \dots, \sigma_{in,n}^2, 0 \right]^T \left[\sum_{p \in \mathcal{N}} \sigma_p^2 \right]^{-1}, \quad (20c)$$

where the imbalance $e \in \mathbb{R}$ is a scalar, and the output uncertainty is set to zero, i.e. $\sigma_{out}^* = 0$. It follows from Eq. (20c) that $\sum_{p \in \mathcal{N}} (\Delta y_p) = -e$, i.e. the sum of variable adjustments exactly mitigates the initial measurement imbalance. Consider the case where the uncertainty of one variable i is infinitely higher than of all other variables. Then, for all inputs $k \in \mathcal{N}$, it must hold that

$$\Delta y_k = -e \sigma_k^2 \left[\sum_{p \in \mathcal{N}} \sigma_p^2 \right]^{-1} \approx \begin{cases} -e & \text{if } k = i \\ 0 & \text{otherwise,} \end{cases} \quad (21)$$

which implies that the complete imbalance is mitigated by the adjustment of variable i , while all other adjustments are zero. This is analogous with the BDA method, so Proposition 1 holds.

Proposition 2. Consider the single-phase, single-tier allocation problem in Fig. 1 with input flows $\mathcal{N} = \{1, \dots, n\}$ of equal relative uncertainty $\sigma_{in}^* = \bar{\alpha}$, and single output of relative uncertainty $\sigma_{out}^* = \bar{\beta}$. Then, for all inputs $i \in \mathcal{N}$, the reconciliation factor r_i is proportional to the absolute uncertainty, $\sigma_{in,i}$.

Proof. It is observed that the vector of reconciliation adjustment factors, $\Delta \mathbf{r} = [\text{diag}(\mathbf{y})]^{-1} \Delta \mathbf{y}$, which is used to rewrite Eq. (20c), yielding

$$\Delta \mathbf{r} = -e \left[\frac{\sigma_{in,1}^2}{y_{in,1}}, \dots, \frac{\sigma_{in,n}^2}{y_{in,n}}, 0 \right]^T \left[\sum_{p \in \mathcal{N}} \sigma_p^2 \right]^{-1} \quad (22a)$$

$$= -e \left[\sigma_{in,1}, \dots, \sigma_{in,n} \right] \bar{\alpha}, 0 \left[\sum_{p \in \mathcal{N}} \sigma_p^2 \right]^{-1}, \quad (22b)$$

where Eq. (22b) follows because $\frac{\sigma_i}{y_i} = \sigma_i^* = \bar{\alpha}$ for all input streams $i \in \mathcal{N}$. The only term in Eq. (22b) that can cause mutual differences in the reconciliation factors is the absolute input measurement uncertainty, $\sigma_{in,i}$. Hence, it is concluded that for all input streams, the reconciliation adjustment factor is proportional to the absolute measurement uncertainty, so Proposition 2 holds.

Corollary 1. Define the relative variance $\text{var}(y)^*$ of a flow variable as the ratio between absolute variance and measurement, i.e. $\text{var}(y)^* = \frac{\text{var}(y)}{y}$. Consider, again, the single-tier allocation problem in Proposition 2, but with equal relative variance $\text{var}(y_i)^* = \bar{\alpha}$ for all input flow streams $i \in \mathcal{N}$. Then, application of DVR reduces to the PRA method.

Proof. As all input streams have equal relative variance, i.e. $\text{var}(y_i)^* = \bar{\alpha} \forall i \in \mathcal{N}$, Eq. (22b) is rewritten as

$$\Delta \mathbf{r} = -e \left[\text{var}(y_{in,1})^*, \dots, \text{var}(y_{in,n})^*, 0 \right]^T \left[\sum_{p \in \mathcal{N}} \sigma_p^2 \right]^{-1} \\ = -e \left[\mathbf{1}_n^T \bar{\alpha}, 0 \right]^T \left[\sum_{p \in \mathcal{N}} \sigma_p^2 \right]^{-1}. \quad (23)$$

Finally, as the reconciliation adjustment factor is equal for all input streams, it is observed that Eq. (23) is equivalent to the PRA method, so Corollary 1 holds.

Proposition 3. Consider the single-tier allocation problem in Fig. 1 with input flow streams $\mathcal{N} = \{1, \dots, n\}$ and single output with $\sigma_{out}^* \approx 0$. Then, for any flow and uncertainty conditions, application of DVR reduces to the UBA method.

Proof. First of all, Eq. (20c) is rewritten in terms of the vector of allocated flow values, $\hat{\mathbf{y}}$:

$$\hat{\mathbf{y}} = \mathbf{y} - e \left[\sigma_{in,1}^2, \dots, \sigma_{in,n}^2, 0 \right]^T \left[\sum_{p \in \mathcal{N}} \sigma_p^2 \right]^{-1} \quad (24)$$

Equivalently, the allocated flow value for every variable is written as

$$\hat{y}_k = y_k - \frac{\sigma_i^2}{\sum_{p \in \mathcal{N}} \sigma_p^2} \cdot e, \quad \forall i \in \mathcal{N}, \quad (25)$$

which is analogous with the UBA method, so it is concluded that Proposition 3 holds.

3.4. Functional comparison of allocation methods

To provide a functional comparison of the different allocation methods, consider the single-phase, single-tier production system as illustrated in Fig. 1, with 4 producing wells and a single output. The flow measurements, uncertainty values, and allocation results obtained by

Table 1

Functional comparison of allocation methods for the single-tier production system with 4 wells and single output.

	Input wells				Output y^{out}	Imbalance e (kg s ⁻¹)
	y_1^{in}	y_2^{in}	y_3^{in}	y_4^{in}		
y (kg/s)	100.0	200.0	150.0	150.0	700.0	-100.0
σ^*	10%	10%	15%	5%	1%	
BDA						
\hat{y} (kg/s)	100.0	300.0	150.0	150.0	700.0	0.0
r	1.000	1.500	1.000	1.000	1.000	
PRA						
\hat{y} (kg/s)	116.7	233.3	175.0	175.0	700.0	0.0
r	1.167	1.167	1.167	1.167	1.000	
UBA						
\hat{y} (kg/s)	124.8	225.2	187.2	162.8	700.0	0.0
r	1.248	1.126	1.248	1.085	1.000	
DVR						
\hat{y} (kg/s)	124.4	224.4	186.6	162.2	697.6	0.0
r	1.244	1.122	1.244	1.081	0.997	

applying different methods to this problem are all presented in Table 1. It is observed that the initial measurement imbalance is $e = -100 \text{ kg s}^{-1}$, and that each allocation method mitigates this imbalance exactly. In the remainder of this section, the results of each allocation method are briefly discussed.

First of all, using BDA, the complete imbalance was attributed to well 2, with a corresponding reconciliation factor of 0.667.² Note that all other reconciliation factors are one, as the allocated values equal the initial measurements. Second, in the case of PRA, the measurement imbalance was distributed proportionally over all inputs, and all reconciliation factors are equal. Third, using UBA the imbalance is distributed proportional to both the flow measurements and the relative measurement uncertainty. Finally, DVR considers the problem formulation given in Eq. (11) to mitigate the measurement imbalance. Although UBA and DVR show similar results, it is important that UBA generally considers the output measurement as fixed, while DVR also considers the uncertainty of the output measurement to allocate the corresponding variable. Note that if the measurement uncertainty of the output meter is assumed zero, DVR reduces to UBA, as also reflected in Proposition 3.

3.5. Iterative DVR implementation for multiphase systems

Thus far, two interdependent optimization steps have been discussed. First of all, the interphase mass transfer coefficients in Eq. (7) are calculated through the linearized sensitivity analysis in HYSYS. Secondly, the DVR problem in Eq. (11) is applied to determine the reconciled flow variables. As the HYSYS and DVR problems are interdependent, both steps are performed iteratively, until the solution has converged. In particular, the convergence criterion, denoted by $\phi[k]$, is specified as follows:

$$\phi[k] = \left| \Delta r[k] - \Delta r[k-1] \right| < \mu, \quad (26)$$

where $\Delta r[k] \in \mathbb{R}^{n_c \cdot n_e}$ is the stacked vector of reconciliation adjustment factors for all phases $c \in C$ in iteration k , with $\Delta r[0] = \mathbf{0}$, and μ is the convergence limit parameter. A similar convergence criterion is formulated based on the changes in interphase mass transfer coefficients:

$$\psi[k] = \left| \Delta f[k] - \Delta f[k-1] \right| < \eta, \quad (27)$$

where $\Delta f[k]$ is the concatenated vector of coefficients in Eq. (7), with $\Delta f[0] = \mathbf{1}_{n_c \cdot n_e}$. Eq. (27) is not implemented as criterion in the iterative algorithm, but is used for evaluation of the simulation results. The complete iterative HYSYS-DVR problem is summarized in Algorithm 1. The HYSYS model and DVR problems for all individual phases are solved iteratively, until either convergence or the maximum number

Algorithm 1 Iterative HYSYS-DVR problem

```

1: Read initial flow data:  $\mathbf{y}^c[0] \forall c \in C$ 
2: Construct system matrices:  $A^c \forall c \in C$ 
3: Initialize counter:  $k = 1$ 
4: while  $|\Delta r^c[k] - \Delta r^c[k-1]| \geq \mu$  and  $k < k_{\max}$  do
5:   Build and solve HYSYS model using:  $\mathbf{y}^c[k-1]$ 
6:   Calculate coefficients:  $f_i^c[k] \forall c \in C, i \in \mathcal{V}$ 
7:   for  $c \in C$  do
8:     Solve DVR problem using:  $\mathbf{y}^c[k-1], f^c[k]$ 
9:     Save DVR solution:  $\mathbf{y}^c[k]$ 
10:  end for
11:   $k = k + 1$ 
12: end while
13: Gross error detection on solution:  $\mathbf{y}^c[k] \forall c \in C$ 

```

of iterations is reached. Thereafter, gross error detection is performed on the final solution, using methods that are discussed in more detail in Section 4.

4. Statistical tests for gross error detection

DVR can be applied to reduce random measurement errors in flow data (Bagajewicz and Jiang, 1998), but the presence of gross errors invalidates the statistical basis of the method (Benqlilou, 2004). Hence, gross errors must be detected and eliminated before valid reconciliation can be performed to reduce random errors (Bagajewicz and Jiang, 1998). In this section, statistical tests for gross error detection are discussed in detail. By means of outlier detection of the DVR results, these tests exploit redundant measurements to detect gross errors (Ragot et al., 1991; Tamhane and Mah, 1985). In all tests, the null hypothesis, H_0 , is that no gross error is present. The tests can be implemented iteratively with Algorithm 1, to achieve valid reconciliation and gross error detection at the same time.

4.1. Network redundancy

Network redundancy is an essential prerequisite for gross error detection in allocation systems (Leibman et al., 1992; Romagnoli and Sánchez, 2000). A measurement is redundant if the system is still observable when that measurement is removed (Stanley and Mah, 1981). Hence, redundancy can be increased, by installing additional meters, as they are coupled to the existing set of measurements via the model constraints. Higher redundancy yields more reference points to reconcile the measurements of interest and is, therefore, a beneficial property for DVR.

Practically nonredundant variables are variables that are theoretically redundant, but behave as nonredundant, such as parallel streams and variables that are only bounded by a single model constraint (Narasimhan and Jordache, 2000). In particular, variables in the outermost layer of any transport system, such as well measurements, are practically nonredundant. Practical nonredundancy makes gross error detection and the attribution of errors to a single input or output stream highly difficult.

4.2. Global imbalance test

Imbalances in an allocation system can be tested collectively by defining the global test statistic for phase c (Bagajewicz, 2000; Tamhane and Mah, 1985):

$$\gamma^c = \left[\mathbf{e}^T \mathbf{H}^{-1} \mathbf{e} \right]^c, \quad (28)$$

² Note that the choice for this specific well is arbitrary in this example.

where $\gamma^c \in \mathbb{R}$, and $e^c \in \mathbb{R}^{n_v}$ is defined in Eq. (9). Under H_0 , γ^c follows a $\chi^2(v)$ -distribution with $v = \text{rank}(A)$ degrees of freedom (Drysdale et al., 2011; Loyola-Fuentes and Smith, 2019). The criterion $\chi^2_{1-\alpha, v}$ is chosen as the critical χ^2 distribution value for a confidence level of $1-\alpha$ (Jiang et al., 2014). If $\gamma > \chi^2_{1-\alpha, v}$, a gross error is detected globally, and H_0 is rejected.

4.3. Nodal constraint test

The nodal measurement imbalances can be used to derive a nodal test for each constraint (Narasimhan and Jordache, 2000; Ragot et al., 1991), given in vector form by

$$z_e^c = \left[\text{diag}(H^c) \right]^{-1/2} e^c, \quad (29)$$

where $z_e^c \in \mathbb{R}^{n_v}$. The nodal constraint test gives rise to n_v univariate tests, i.e. one for each constraint $i \in \mathcal{V}$. Under H_0 , the statistics in z_e^c follow a standard normal distribution (Bagajewicz, 2018). The two-sided critical value of the standard normal distribution is denoted by $Z_{1-\beta/2}$. For proper control of Type I errors (i.e. rejection of a true null hypothesis), the modified significance level β with $k = n_v$ is used (Mah and Tamhane, 1982), which is defined using α as follows:

$$\beta = 1 - (1 - \alpha)^{\frac{1}{k}}. \quad (30)$$

If the test statistic $|z_{e,i}^c| > Z_{1-\beta/2}$ for any constraint $i \in \mathcal{V}$, H_0 is rejected, and a gross error is detected in that specific constraint.

4.4. Measurement (penalty factor) test

Gross errors can be detected at variable level, by defining penalty factors $z^c \in \mathbb{R}^{n_e}$ for each edge (Bagajewicz, 2018; Mah and Tamhane, 1982), given in vector form by

$$z^c = \left[\text{diag}(\sigma^c) \right]^{-1} (\hat{y}^c - y^c) = \left[\text{diag}(\sigma^{c*}) \right]^{-1} \Delta r^c, \quad (31)$$

where σ^{c*} is the vector of relative uncertainty in phase c , and Δr^c the vector of reconciliation adjustment factors. The statistics in z^c provide n_e univariate tests, and follow a standard normal distribution under H_0 (Narasimhan and Jordache, 2000; Tamhane and Mah, 1985). The modified significance level β with $k = n_e$ as defined in Eq. (30) is used. If $z_i > Z_{1-\beta/2}$ for any variable $i \in \mathcal{E}$, H_0 is rejected, and a gross error is detected in that specific variable.

Based on Eq. (31), the following useful proposition is derived:

Proposition 4. Consider the single-phase, single-tier allocation problem in Fig. 1 with input flow streams $\mathcal{N} = \{1, \dots, n\}$ and single output with $\sigma_{out}^* \approx 0$. Then, for all $i \in \mathcal{N}$, the penalty factors z_i are proportional to the absolute uncertainty, σ_i .

Proof. First of all, Eq. (31) is rewritten by inserting Eq. (22a), yielding

$$z = -e \left[\sigma_{in,1}, \dots, \sigma_{in,n}, 0 \right]^T \left[\sum_{p \in \mathcal{N}} \sigma_p^2 \right]^{-1} \quad (32)$$

As the measurement imbalance e and summation term are shared among all variables, it follows from Eq. (32) that the penalty factor is proportional to the absolute uncertainty for all input variables. Hence, it is concluded that Proposition 4 holds.

In gross error detection practice, the variable associated with the highest test statistic (i.e. largest violation with the expected distribution) is marked for inspection or calibration. Therefore, Proposition 4 implies that under the measurement test, any gross error in the system is always allocated to the flow meter with highest absolute uncertainty. In addition, if all flow meters show similar relative measurement uncertainty, the gross error is always allocated to the meter showing the highest absolute flow.

4.5. Principal component tests

The nodal and measurement tests discussed above only exploit the information in the diagonals of the covariance matrices, while all other information is lost. In the principal component (PC) test, the eigenvectors of the covariance matrix are exploited to construct a new set of uncorrelated variables, known as principal components (Drysdale et al., 2011; Narasimhan and Jordache, 2000). Based on the eigenvectors of the covariance matrix, a new coordinate system is constructed, in which the variables are uncorrelated and the covariance matrix is diagonal (Amand et al., 2000). As a result, PC tests help control Type I errors, and can increase the detectability of subtle gross errors (Tong and Crowe, 1995).

The vector of principal components, $p_a \in \mathbb{R}^n$ of component $a \in \mathbb{R}^n$ with covariance matrix $\Sigma_a = \text{cov}(a)$ is given as follows:

$$p_a = W_a^T a, \quad (33)$$

where the columns of W_a are the eigenvectors of Σ_a , satisfying $W_a = U_a \Lambda_a^{-1/2}$ (Narasimhan and Jordache, 2000). Matrix U_a contains the orthonormalized eigenvectors of Σ_a , such that $U_a U_a^T = I_n$, and matrix Λ_a contains the eigenvalues on the diagonal (Tong and Crowe, 1995), i.e.

$$\Lambda_a = U_a^T \Sigma_a U_a = \text{diag}(\lambda_{a_1}, \dots, \lambda_{a_n}). \quad (34)$$

Under H_0 , the principal components follow the normal distribution $p_a \sim N(0, I_n)$, while the original components follow $a \sim N(0, \Sigma_a)$. Principal component tests can be performed for constraint residuals by letting $a = e$ and $\Sigma_a = \text{cov}(e) = H$, or for measurement adjustments by choosing $a = y - \hat{y}$ and $\Sigma_a = \text{cov}(y - \hat{y}) = Q^T A^T H^{-1} A Q$, respectively (Narasimhan and Jordache, 2000).

5. Numerical simulation studies

To demonstrate the performance of DVR and gross error detection methods, two numerical simulation studies are performed, both on the 4-layer allocation system topology shown in Fig. 3. In the first numerical study, Monte-Carlo simulations are performed on single-phase allocation systems. In the second numerical study, the performance of the HYSYS-DVR algorithm to solve multiphase allocation problems is demonstrated. In both studies, the relative flow metering uncertainty as given in Eq. (16) is 1% in the CT point, 5% in the WHPs, and 10% in wells. By default, only single flow metering is performed at the inlet of pipes, which implies that the outgoing edges i of all pipes show infinitely high uncertainty, i.e. $\sigma_i^* \approx \infty$. All methods are implemented in Python, with PICOS as linear solver interface (Sagnol and Stahlberg, 2019).

5.1. Experimental set-up

In the first numerical study, Monte-Carlo simulations are performed to demonstrate the detectability of gross errors in gas flow measurements using the global, nodal, measurement and principal component tests. To this end, a simplified, single-phase gas allocation system of the topology in Fig. 3 is considered. The total mass flow at the CT point is fixed at $y_{CT} = 100 \text{ kg s}^{-1}$, and is divided randomly at each intersection over the individual input streams. To generate realistic flow data, random measurement errors are imposed on all gas flow variables. In addition, a single gross error caused by sensor bias of magnitude up to $\delta_i = 20\sigma_i^*$ is randomly imposed on exactly one gas flow variable i in either a well, WHP, or the CT point. For all cases, 5000 simulation iterations are performed, in which a different randomized flow distribution is determined, new random errors are imposed, and a new random location for the gross error is determined. The resulting allocation problem is solved using the DVR formulation in Eq. (11), and the gross error detection methods are applied. In addition, to demonstrate that gross error detection performance can be improved

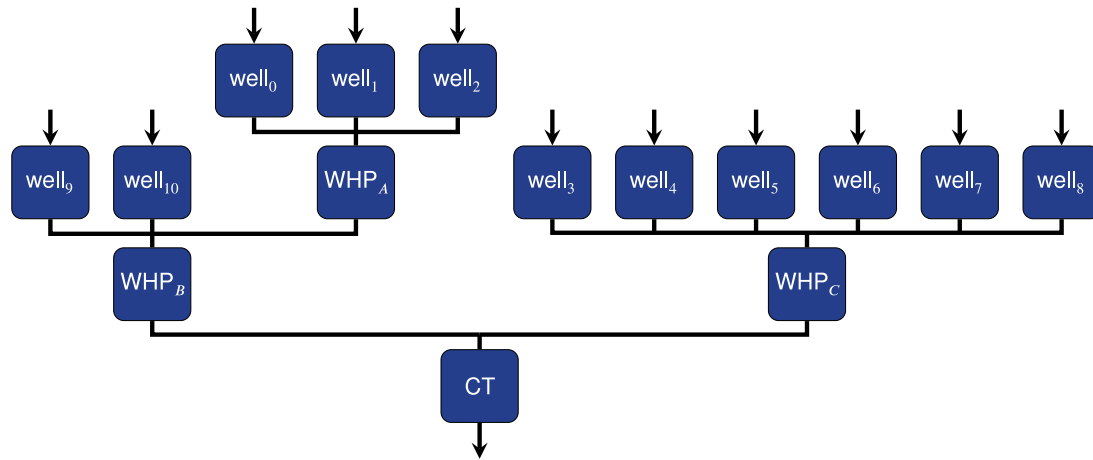


Fig. 3. Allocation system topology of the numerical simulation studies, with each block as pipe of the described type.

Table 2

Gross error detection performance of the measurement test and cumulative error reduction by applying DVR for different types of well metering systems in the presence of a single gross error of magnitude $\delta_i = 5\sigma_i^*$.

Metering type	σ_{II}^*/σ_I^*	Error in well		Error in WHP		Error in CT	
		Performance	Error reduction (Δe)	Performance	Error reduction (Δe)	Performance	Error reduction (Δe)
Single well metering	N/A	20.28%	-28.70%	96.58%	-50.04%	0.12%	-2.33%
Double well metering	1	73.66%	-45.41%	96.98%	-55.76%	0.96%	-16.09%
Double well metering	2	44.76%	-43.34%	90.44%	-54.10%	0.08%	-24.80%
Double well metering	3	29.06%	-45.91%	73.22%	-56.11%	0.02%	-32.58%

by increasing redundancy, experiments are performed on both single and double well metering systems. In case of single well metering, only one physical flow meter is present per well, which is located at the inlet of well pipes. The measurement uncertainty of the primary flow meters is denoted by σ_I^* , and is specified for each meter. In the case of double well metering, a secondary flow meter with measurement uncertainty denoted by σ_{II}^* is introduced at the outlet of well pipes. Double flow metering systems are characterized by the ratio between the secondary and primary measurement uncertainty: σ_{II}^*/σ_I^* . Experiments for double well metering are performed with uncertainty ratios of 1, 2, and 3.

In the second simulation study, the complete multiphase allocation system depicted in Fig. 3 is considered with realistic flow, thermodynamic conditions and fluid compositions. The multiphase allocation system is based on a real field situation and produces mainly gas, but also a significant amount of condensate and water, accumulating to ± 5 -20 wt% of the total production per well. Total mass flow at the CT point is around 84 000 kg h⁻¹, with individual wells contributing for up to 11 800 kg h⁻¹. Realistic gas compositions are obtained, and oil compositions are assumed symmetric around C₇ in the range of C₃-C₁₀. In three cases, different gross errors are imposed:

1. a single gross error of magnitude $\delta = 5\sigma^*$ in the gas flow meter of well 5;
2. a single miscalibration of the oil/water fraction in WHP A;
3. a combination of 6 gross errors throughout the network.

Algorithm 1 is implemented in Python, to iteratively solve the interconnected HYSYS-DVR problem with convergence limit parameter $\mu = 0.001$. The goal is to demonstrate the convergence behavior of the interconnected HYSYS-DVR problem, and to determine the error correction performance of DVR. In addition, the performance of gross error methods is also explored.

5.2. Experimental results

In Fig. 4, the gross error detection performance of the different methods in the Monte Carlo study is depicted. The figures show the

Table 3

DVR and gross error detection results on multiphase HYSYS-DVR simulations.

Simulation case	1	2	3	
Simulation time (s)	157.94	154.68	156.86	
Iterations	4	4	4	
Error reduction, Δe	-32.31%	-32.99%	-18.59%	
Mass transfer coefficients				
f^s	min	0.92373	0.92378	0.92951
	max	1.00000	1.00000	1.00000
f^o	min	0.94381	0.94390	0.94506
	max	1.12464	1.12712	1.14797
f^w	min	1.00000	1.00000	1.00000
	max	1.01090	1.01091	1.01098
Gross error detection tests				
Global detection	No	No	Yes	
Nodal detection	Yes	Yes	Partial	
Measurement det.	No	Partial	Partial	

detectability of gross errors in a random well (left) and random WHP (right), respectively, as a function of the gross error magnitude. It is observed that the global test can only detect gross errors of very high magnitude. The standard nodal test performs very well for isolating the approximate location of a gross error, regardless of the error location. The measurement test is highly suitable for exactly locating gross errors in WHPs, but shows poor detection performance for errors in wells. Although it was postulated that principal component tests can perform better than the standard tests (Narasimhan and Jordache, 2000), slightly lower detection performance is observed in all cases. As the considered system matrices A are primarily of block-diagonal structure and Q is fully diagonal, any associated covariance matrix is also diagonally dominant, and the extra data that is exploited using principal component tests is negligible. In addition, the presence of unknown flow variables in the outgoing edges of pipe nodes results in numerical difficulties in determining the eigenvectors and eigenvalues. Hence, principal component tests do not perform better than the standard tests, as reflected in Fig. 4.

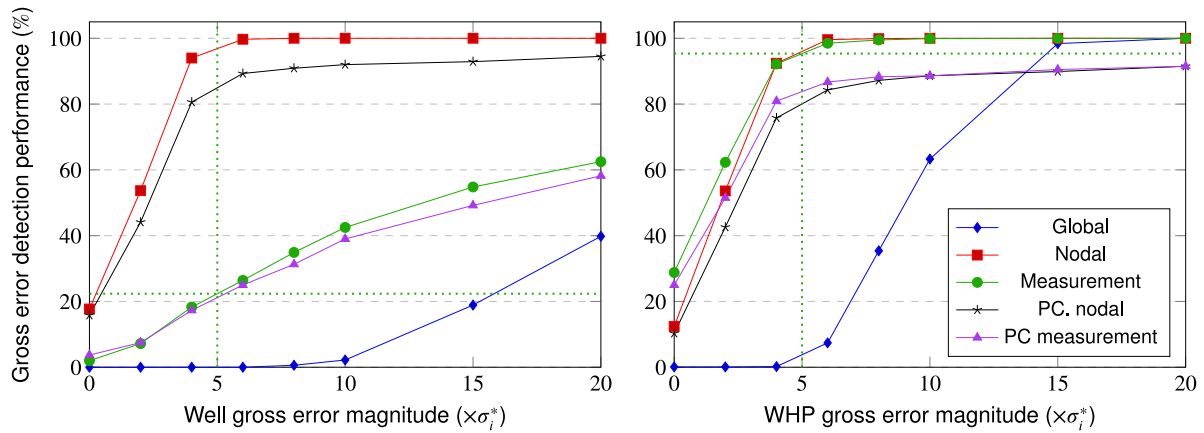


Fig. 4. Probability of detecting a gross error (measurement bias) located in a well (left) or WHP (right) of the single well metering gas allocation system, using different detection methods.

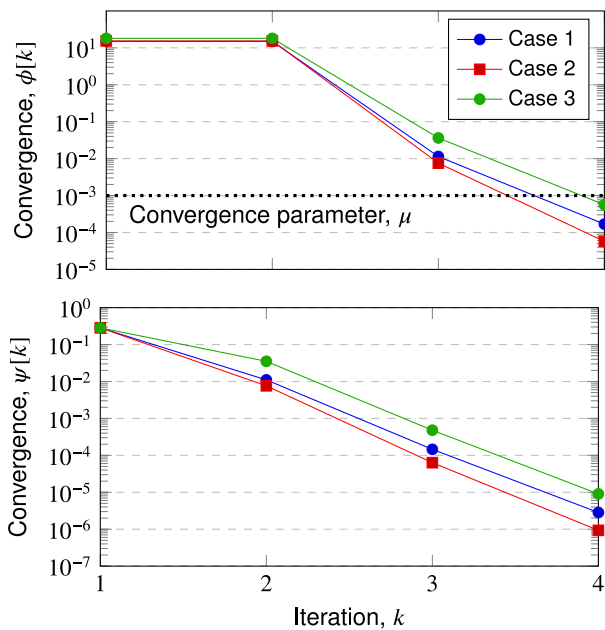


Fig. 5. Convergence criteria on reconciliation factors (top) and mass transfer coefficients (bottom) of iterative HYSYS-DVR simulations.

In Table 2, the detectability of gross errors (by means of the measurement test) and cumulative error reduction (by applying DVR) are shown for the simplified gas allocation system under different types of well metering. The detectability of gross errors is described as detection performance, and the error reduction, denoted by Δe , is calculated using Eq. (15). The performance value in Table 2 corresponds with the detection performance for the measurement test in Fig. 4 at a magnitude of $\delta_i = 5\sigma_i^*$ (indicated by the green dotted lines in both figures). As also reflected in Fig. 4, detectability of gross errors in WHPs of single metering systems is very high, while the detection of errors in wells and the CT point is much more difficult. However, the detectability of gross errors in wells improves significantly in double well metering systems with low σ_{11}^*/σ_1^* ratio. Cumulative error reduction is also improved, even under higher secondary uncertainty. Despite this, detectability of gross errors in WHPs is lower for double than single metering systems, because the number of meters that can contain the error (and are thus considered in the tests) is also increased.

The DVR and gross error detection results for the HYSYS-DVR simulations are presented in Table 3, and convergence behavior for the criteria in Eqs. (26) and (27) is shown in Fig. 5. All cases required 4

iterations to achieve convergence and showed similar simulation times. As shown in Fig. 5, the HYSYS-DVR algorithm approximately converges linearly with respect to mass transfer coefficients for all iterations, and linearly to reconciliation factors after the first iteration. Error reduction is highest for case 1–2 around -33%, although case 3 showed the highest absolute reduction, due to the presence of multiple gross errors. Only errors in interior locations of the system (i.e. WHPs) were located exactly, while errors in wells were only located approximately. For case 2, partial detection performance of the measurement test implies that the error was only detected in the water phase, but not in the oil phase. For case 3, partial detection indicates that only the first few errors (of the total of 6) were successfully detected. The global test was only triggered in the presence of multiple gross errors (case 3), because the cumulative imbalance was too low for case 1–2. Hence, the gross error detection results in Table 3 are in agreement with Fig. 4. In all simulations, the interphase mass transfer coefficients of water are approximately one, thus showing linear relative sensitivity. In contrast, the coefficients are in the range of 0.924–1.000 for gas, and 0.944–1.148 for oil. Interestingly, for some wells and WHPs, the mass transfer coefficients of both gas and oil are both below, or both above one. For example, the HYSYS-DVR allocation results for well 6 of the multiphase system under multiple gross errors (case 3) are as follows:

$$\begin{aligned} f_6^g &= 0.95822, & \hat{y}_6^g &= 10320.96 \text{ kg h}^{-1}, \\ f_6^o &= 0.95147, & \hat{y}_6^o &= 471.30 \text{ kg h}^{-1}, \\ f_6^w &= 1.00152, & \hat{y}_6^w &= 798.40 \text{ kg h}^{-1}. \end{aligned} \quad (35)$$

Hence, any extra production of gas in this well, denoted by ∂y_6^g , results in a lower increase of the gas WHP flow, i.e. $\partial y_{\text{WHP}}^g = f_6^g \partial y_6^g < \partial y_6^g$. Vice versa, any additional oil production results in a lower increase of the oil WHP flow. Instead, part of the additional gas production is condensed in the oil phase after the streams are merged at the WHP, while additionally produced oil evaporates into the gas phase. This interphase mass transfer effect is caused by pressure and temperature variations and co-mingling of streams from multiple contributing wells, and highlights the importance of compensating for interphase mass transfer effects in allocation systems (Bikmukhametov and Jäschke, 2020).

6. Conclusions and future perspectives

In this paper, DVR was proposed for error correction and gross error detection in multiphase allocation systems. A model for multiphase production systems was developed, which takes into account nonlinear interphase mass transfer effects, and HYSYS was used to linearize the system around the operating conditions. The developed model is highly versatile and can easily be applied to model larger or more complex

production systems than considered in the numerical studies. It was shown analytically that DVR can be very similar to UBA, BDA, and PRA, depending on the characteristics of the allocation system. The iterative HYSYS-DVR algorithm was successfully used in the numerical simulation studies to solve the corresponding multiphase allocation problem, and converged linearly for all cases. Based on the convergence behavior, it was shown that the iterative implementation yields a better solution than a single-step algorithm would provide. Furthermore, it was demonstrated that DVR can reduce cumulative measurement errors by up to 56% for gas allocation systems and 33% for multiphase systems, and that the results can be used for gross error detection at global, nodal, and individual variable levels.

Based on this work, it is concluded that the detectability of gross errors strongly depends on the error location and magnitude. The global test is only triggered in the presence of gross errors of very high magnitude. For errors in exterior locations (i.e. wells and the CT point) only the approximate location can be determined using nodal constraint tests, unless the error magnitude is extremely high. Allocating gross errors to individual well or CT flow measurements is in general not possible, as these variables are only involved in a single model constraint and are, therefore, practically nonredundant. On the contrary, flow variables in the interior of the system are associated with multiple model constraints. Consequently, gross errors in these variables can be located precisely using the measurement test. The detectability of gross errors in wells can be improved by increasing the network redundancy, e.g. by introducing secondary well flow meters. However, this requires that the secondary meter exhibits sufficiently low measurement uncertainty.

6.1. Recommendations for future work

In future research, the developed HYSYS-DVR algorithm can potentially be extended towards a Virtual Flow Measurement System, or VFMS (Coupot et al., 2010, 2017; Haouche et al., 2012). A VFMS adds additional layers of redundancy upon the standard DVR method, by providing a coherent, accurate set of process data and system constraints (Bikmukhametov and Jäschke, 2020; Szega, 2018). By improving the accuracy of process models and increasing network redundancy, both random and gross measurement errors can be detected and corrected more accurately (Wising et al., 2008). However, open challenges remain in developing such advanced models, and it is likely that a VFMS yields a bilinear or nonlinear problem, which requires more sophisticated procedures to solve (Bagajewicz and Jiang, 2000; Taylor and Del Pilar Moreno, 2013).

To conclude, when sufficient data and network redundancy can be exploited, DVR is a valuable method to reduce overall measurement errors and to detect gross error of significant magnitude in multiphase allocation systems. However, care must be taken before conclusions are made, and the system under consideration must be analyzed in advance to determine the applicability of DVR. In addition, nonlinear interphase mass transfer effects in allocation systems can be significant. It is, therefore, important to compensate for these effects and consider the complete production system topology, in order to derive the most fair sales allocation between all involved stakeholders.

CRedit authorship contribution statement

Thom S. Badings: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **Dennis S. van Putten:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Abu-El-Zeet, Z.H., Roberts, P.D., Becerra, V.M., 2002. Enhancing model predictive control using dynamic data reconciliation. *AIChE J.* 48, 324–333. <http://dx.doi.org/10.1002/aic.690480216>.
- Alverà, M., Carroll, D.C., Marten, I., 2018. Global gas report 2018. In: SNAM, World Gas Conference 2018. SNAM, URL https://www.snam.it/export/sites/snam-ri/repository/file/gas_naturale/global-gas-report/global_gas_report_2018.pdf.
- Amand, T., Heyen, G., Kalitventzeff, B., 2000. Plant monitoring and fault detection: Synergy between data reconciliation and principal component analysis. *Comput. Aided Chem. Eng.* 8, 757–762. [http://dx.doi.org/10.1016/S1570-7946\(00\)80128-5](http://dx.doi.org/10.1016/S1570-7946(00)80128-5).
- Amin, A., 2016. Using measurement uncertainty in production allocation. In: Upstream Production Measurement Forum. Houston, Texas. URL <https://upmforum.com/lectures-presentations-forum-documents/papers/UPM%2016010%20-%20Allocation%20Uncertainty%20Paper%204%20-%20Draft.pdf>.
- Amin, A., Webb, R.A., Kalitventzeff, B., Vos, G.D., 2016. Application of data validation and reconciliation to production allocation. In: 34th International North Sea Flow Measurement Workshop. URL <https://nfoqm.no/wp-content/uploads/2019/02/2016-10-Application-of-Data-Validation-and-Reconciliation-to-Production-Allocation-Amin-Letton-Hall-Group.pdf>.
- Bagajewicz, M.J., 2000. A brief review of recent developments in data reconciliation and gross error detection/estimation. *Lat. Am. Appl. Res.* 30, 335–342. URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.604.5250&rep=rep1&type=pdf>.
- Bagajewicz, M.J., 2018. Process Plant Instrumentation Design and Upgrade. CRC Press, <http://dx.doi.org/10.1201/9781482279108>.
- Bagajewicz, M.J., Jiang, Q., 1998. Gross error modeling and detection in plant linear dynamic reconciliation. *Comput. Chem. Eng.* 22, 1789–1809. [http://dx.doi.org/10.1016/S0098-1354\(98\)00248-8](http://dx.doi.org/10.1016/S0098-1354(98)00248-8).
- Bagajewicz, M.J., Jiang, Q., 2000. Comparison of steady state and integral dynamic data reconciliation. *Comput. Chem. Eng.* 24, 2367–2383. [http://dx.doi.org/10.1016/S0098-1354\(00\)00498-1](http://dx.doi.org/10.1016/S0098-1354(00)00498-1).
- Bahadori, A., 2014. Natural Gas Processing: Technology and Engineering Design. Gulf Professional Publishing, URL <https://www.sciencedirect.com/book/9780080999715/natural-gas-processing>.
- Benqilou, C., 2004. Data Reconciliation as a Framework for Chemical Processes Optimization and Control (ProQuest dissertations and theses). URL <https://upcommons.upc.edu/handle/2117/93748>.
- Bikmukhametov, T., Jäschke, J., 2020. First principles and machine learning virtual flow metering: A literature review. *J. Petrol. Sci. Eng.* 184, 106487. <http://dx.doi.org/10.1016/j.petrol.2019.106487>.
- Bjørk, R.N., Skålvik, A.M., Pobitzer, A., Mosland, A., Sætre, E., Frøysa, K., 2016. Analysis of field and ownership allocation uncertainty in complex multi-field configurations. In: North Sea Flow Measurement Workshop. URL <https://www.prototech.no/publications/10672/analysis-of-field-and-ownership-allocation-uncertainty-in-complex-multi-field-configurations/>.
- Bullo, F., 2018. Lectures on Network Systems, first ed. CreateSpace, URL <http://motion.me.ucsb.edu/book-1ns/>.
- Câmara, M., Soares, R., Feital, T., Anzai, T., Diehl, F., Thompson, P., Pinto, J., 2017. Numerical aspects of data reconciliation in industrial applications. *Processes* 5, 56. <http://dx.doi.org/10.3390/pr5040056>.
- Chebiyyam, C., 2010. Reconciliation - An inevitable challenge in upstream oil industry. In: SPE Oil and Gas India Conference and Exhibition. <http://dx.doi.org/10.2118/127486-ms>.
- Cherukuri, A., Cortes, J., 2017. Iterative bidding in electricity markets: rationality and robustness. *IEEE Trans. Netw. Sci. Eng.* 1–13. <http://dx.doi.org/10.1109/TNSE.2019.2921056>, arXiv:1702.06505.
- Coupot, J.-P., Caulier, R., Wising, U., 2010. Field and installation monitoring using on line data validation and reconciliation — Application to offshore fields in middle east and west Africa. In: SPE Intelligent Energy Conference and Exhibition. Utrecht, <http://dx.doi.org/10.2118/128717-MS>.
- Coupot, J., Laiani, N., Richon, V., 2017. Operational experience with virtual flow measurement technology. In: 35th International North Sea Flow Measurement Workshop. URL <https://nfoqm.no/wp-content/uploads/2019/02/2017-11-Virtual-flow-measurements-Coupot-Total.pdf>.
- Drysdale, E., Little, H., Stockton, P., Wilson, A., 2011. Gross meter error detection and elimination by data reconciliation. In: 29th North Sea Flow Measurement Workshop. https://www.researchgate.net/publication/311593251_Gross_Meter_Error_Detection_and_Elimination_by_Data_Reconciliation.
- Gil, H.A., Galiana, F.D., Conejo, A.J., 2005. Multiarea transmission network cost allocation. *IEEE Trans. Power Syst.* 20, 1293–1301. <http://dx.doi.org/10.1109/TPWRS.2005.851951>.

- Hauouche, M., Tessier, A., Deffous, Y., Authier, J.F., Couput, J.P., Caulier, R., Vrielynck, B., 2012. Smart metering: An online application of data validation and reconciliation approach. In: Society of Petroleum Engineers Intelligent Energy International 2012, Vol. 1. Utrecht, <http://dx.doi.org/10.2118/149908-MS>.
- Harrison, G., Safar, F., 2004. Modern E and P data management in Kuwait Oil Company. J. Petrol. Sci. Eng. 42, 79–93. <http://dx.doi.org/10.1016/j.petrol.2003.12.002>.
- Hodouin, D., Everell, M.D., 1980. A hierarchical procedure for adjustment and material balancing of mineral processes data. Int. J. Miner. Process. 7, 91–116. [http://dx.doi.org/10.1016/0301-7516\(80\)90002-2](http://dx.doi.org/10.1016/0301-7516(80)90002-2).
- Jiang, X., Liu, P., Li, Z., 2014. Data reconciliation and gross error detection for operational data in power plants. Energy 75, 14–23. <http://dx.doi.org/10.1016/j.energy.2014.03.024>.
- Johnsen, A.J., Dahl, A.M., 2018. Maria to Kristin allocation from project to operation, a real life experience. In: NFOGM Hydrocarbon Management Workshop. URL <https://nfogm.no/wp-content/uploads/2018/09/04.Johnsen-Anita-Dahl-Arne-M-Maria-to-Kristin-Allocation-from-Project-to-Operation-a-Real-Life-Experience.pdf>.
- Kansho, S., 2020. A review of hydrocarbon allocation methods in the upstream oil and gas industry. J. Petrol. Sci. Eng. 184, <http://dx.doi.org/10.1016/j.petrol.2019.106590>.
- Leibman, M.J., Edgar, T.F., Lasdon, L.S., 1992. Efficient data reconciliation and estimation for dynamic processes using nonlinear programming techniques. Comput. Chem. Eng. 16, 963–986. [http://dx.doi.org/10.1016/0098-1354\(92\)80030-D](http://dx.doi.org/10.1016/0098-1354(92)80030-D).
- Loyola-Fuentes, J., Smith, R., 2019. Data reconciliation and gross error detection in crude oil pre-heat trains undergoing shell-side and tube-side fouling deposition. Energy 183, 368–384. <http://dx.doi.org/10.1016/j.energy.2019.06.119>.
- Lunde, P., Frøysa, K.-E., Neumann, S., Halvorsen, E., 2002. Determination of measurement uncertainty for the purpose of wet gas hydrocarbon allocation. In: North Sea Flow Measurement Workshop. URL <https://nfogm.no/wp-content/uploads/2019/02/2002-03-Determination-of-Measurement-Uncertainty-for-the-Purpose-of-Wet-Gas-Hydrocarbon-Allocation-Webb-BP.pdf>.
- Mah, R., Tamhane, A., 1982. Detection of gross errors in process data. AIChE J. 28, 828–830. <http://dx.doi.org/10.1002/aic.690280519>.
- Mesbahi, M., Egerstedt, M., 2010. Graph Theoretic Methods in Multiagent Networks, first ed. Princeton University Press, New Jersey, p. 403. <http://dx.doi.org/10.1073/pnas.0703993104>.
- Meyer, M., Koehret, B., Enjalbert, M., 1993. Data reconciliation on multicomponent network process. Comput. Chem. Eng. 17, 807–817. [http://dx.doi.org/10.1016/0098-1354\(93\)80065-U](http://dx.doi.org/10.1016/0098-1354(93)80065-U).
- Narasimhan, S., Jordache, C., 2000. Data Reconciliation & Gross Error Detection. Gulf Publishing Company, Houston, <http://dx.doi.org/10.1016/B978-0-88415-255-2.X5000-9>.
- Oliveira, E.C., Aguiar, P.F., 2009. Data reconciliation in the natural gas industry: Analytical applications. Energy Fuels 23, 3658–3664. <http://dx.doi.org/10.1021/ef9001428>.
- Pobitzer, A., Skålvik, A.M., Bjørk, R.N., 2016. Allocation system setup optimization in a cost-benefit perspective. J. Petrol. Sci. Eng. 147, 707–717. <http://dx.doi.org/10.1016/j.petrol.2016.08.025>.
- van Putten, D.S., van Luijk, L., Wahab, I.S., Johari, S.F., 2019. Assessment of allocation systems: combining data validation & reconciliation scheme and PVT simulations. In: North Sea Flow Measurement Workshop. URL https://www.tekna.no/contentassets/eac487e9aabb4773a1337fda3c28b70e/06-assessment-of-allocation-systems_dnvgl_dennis-putten.pdf.
- Ragot, J., Aitouche, A., Kratz, F., Maquin, D., 1991. Detection and location of gross errors in instruments using parity space technique. Int. J. Miner. Process. 31, 281–299. [http://dx.doi.org/10.1016/0301-7516\(91\)90031-D](http://dx.doi.org/10.1016/0301-7516(91)90031-D).
- Romagnoli, J.A., Sánchez, M.C., 2000. Data Processing and Reconciliation for Chemical Process Operations. Academic Press, San Diego, California, URL <https://www.sciencedirect.com/bookseries/process-systems-engineering/vol/2>.
- Sagnol, G., Stahlberg, M., 2019. PICOS: A Python interface to conic optimization solvers. URL <https://picos-api.gitlab.io/picos/>.
- Stanley, G.M., Mah, R.S., 1981. Observability and redundancy in process data estimation. Chem. Eng. Sci. 36, 259–272. [http://dx.doi.org/10.1016/0009-2509\(81\)85004-X](http://dx.doi.org/10.1016/0009-2509(81)85004-X).
- Stewart, D., Skelton, M., 2004. Equity exposure in wet gas allocation metering. In: North Sea Flow Measurement Workshop. URL <https://nfogm.no/wp-content/uploads/2019/02/2004-04-Equity-Exposure-in-Wet-Gas-Allocation-Metering-Stewart-NEL.pdf>.
- Stockton, P., Spence, A., 2008. Experiences in the use of uncertainty based allocation in a North Sea offshore oil allocation. In: Production and Upstream Flow Measurement Workshop. https://www.researchgate.net/publication/286883319-Experiences_in_the_Use_of_Uncertainty_Based_Allocation_in_a_North_Sea_Offshore_Oil_Allocation_System.
- Szega, M., 2018. Extended applications of the advanced data validation and reconciliation method in studies of energy conversion processes. Energy 161, 156–171. <http://dx.doi.org/10.1016/j.energy.2018.07.094>.
- Tamhane, A.C., Mah, R.S., 1985. Data reconciliation and gross error detection in chemical process networks. Technometrics 27, 409–422. <http://dx.doi.org/10.1080/00401706.1985.10488080>.
- Taylor, J.H., Del Pilar Moreno, R., 2013. Nonlinear dynamic data reconciliation: In-depth case study. In: Proceedings of the IEEE International Conference on Control Applications. IEEE, pp. 746–753. <http://dx.doi.org/10.1109/CCA.2013.6662839>.
- Tong, H., Crowe, C., 1995. Detection of gross errors in data reconciliation by principle components analysis, 41. American Institute of Chemical Engineers, <http://dx.doi.org/10.1002/aic.690410711>.
- Webb, R.A., Letton, W., Basil, M., 2002. Determination of measurement uncertainty for the purpose of wet gas hydrocarbon allocation. In: North Sea Flow Measurement Workshop. URL <https://nfogm.no/wp-content/uploads/2019/02/2002-03-Determination-of-Measurement-Uncertainty-for-the-Purpose-of-Wet-Gas-Hydrocarbon-Allocation-Webb-BP.pdf>.
- Wee, A., 2014. Using multiphase and wetgas meters for revenue allocation. In: North Sea Flow Measurement Workshop. URL <https://nfogm.no/wp-content/uploads/2014/02/Using-Multiphase-and-Wetgas-meters-for-revenue-allocation.pdf>.
- Wising, U., Campan, J., Vrielynck, B., Dos Anjos, C., Kalitventzef, P.-b., 2008. Improving your real-time data infrastructure using advanced data validation and reconciliation. In: Rio Oil & Gas Expo and Conference. URL <https://www.osti.gov/etdweb/servlets/purl/21169042>.