Print ISSN 2682-3993
Online ISSN 2682-4000

Data Reconciliation As A Key To Enable Digitalisation Of Agrifood Industrial Processes: A Preliminary Case Study

Andrea Galeazzi [1,2], Francesca Girotto [2,3*], Lorenzo Rizzoli [2], Kristiano Prifti [1,2], Flavio Manenti [1,2], and Laura Piazza [2,3]

[1]CMIC "Giulio Natta" Department, Politecnico di Milano, Piazza Leonardo da Vinci 32, Milano, 20133, Italy
[2]Department of Environmental Science and Policy, Universita` degli Studi di Milano, Via Celoria 2, Milano, 20133, Italy
[3]Centre for Sustainable Process Engineering Research (SuPER), Politecnico di Milano, Milano, Italy
Corresponding author: francesca.girotto@unimi.it

Abstract: Through the adoption of the industry 4.0 matrix technology, the Agrifood 4.0 era calls for more big data to be collected and analyzed to help the transition from the traditional industry towards a digital twin-based and energy-saving industry. However, proper handling of those diverse information remains challenging. The often-overlooked prerequisite to digitalization is to verify the reliability of such data. Measured material streams can be validated by performing data reconciliation, a technique that uses material and conservation equations to minimize the actual measurement error in-process data. An application example is here proposed in order to provide an understandable guideline of data reconciliation procedure at a food industry level. Data reconciliation was, indeed, applied to the measured production data (material flows) of an industrial orange juice production line processing fresh raw material into concentrated juice (up to the desired dissolved sugar content of 65 °Bx). Energetic data assessment was performed using Aspen HYSYS v11.0. Results proved the importance of getting reconciled data in order to close material balances clearing the measurement inaccuracies due to operators' errors or sensors bias. Besides, through Aspen HYSYS software, the surplus to requirements of the steam applied in the concentration unit operation was pointed out. The main perspective for future research will be the application of the present outputs to create the digital twin through which material and energy flows can be simulated to identify the best process optimization strategy.

Keywords: Agrifood 4.0; process optimization; digitalization; data reconciliation; orange juice production line

I. INTRODUCTION

One of the greatest challenges of our century is to seek and build the tools to enable the agrifood industry to cope with the growing demand for food underlying a world population that is expected to reach 9.7 billion people in 2050. As a matter of fact, the agrifood sector must rethink methods to improve productivity while reducing the consumption of resources, minimizing waste generation, and improving traceability. At each phase of the food value chain, a significant amount of data is generated. Such data provide important information to the agents involved in the processing and transport of food products from farm to fork [1]. During raw material processing, proper handling of data has a crucial role in providing safe, quality, and sustainable products.

Consumers seems to reward those food companies keen at improving their environmental sustainability through product and process innovations finalized to maintain or improve their performances while minimizing the consumption of resources, mainly water and energy. Indeed, the water-energy-food nexus describes the complex and interrelated nature of our global resources systems while defining a new approach in support of food security and sustainable agriculture [2]. Digitalization is taking place in the food sector due to the implementation of advanced technologies (e.g. sensors, robots, drones) [1] which help monitoring and managing huge amounts of data in view of the transition from traditional industry towards a digital twin-based and energy-saving industry [3,4]. The combined productivity-profitability-sustainability purposes for which digital tools (advanced robotics, cyber-physical production systems, Internet of Things, big data analysis, Artificial Intelligence, Virtual Realities, etc.) are deployed have been named Industry 4.0 [3,4]. Through the adoption of the industry 4.0 matrix technology, the new Agrifood 4.0 era calls for more big data to be collected and analyzed to provide improved insights while leading to better process control, supply chain management, traceability and ultimately decision making [3]. Most

Print ISSN 2682-3993
Online ISSN 2682-4000

importantly, the major pillars needing to be managed in terms of Agenda 2030 and Sustainable Development Goals are waste, energy, economy, and environment [5].
In this context of industrial digitalization, the concept of the digital twin has recently emerged as a means for a more versatile process operational management [3]. Digital twins have advanced fast in various industries, but they are just emerging in postharvest supply chains [6]. Indeed, at the processing companies' level, system boundaries are narrower and comprehensive numerical tools and software platforms for improving insights and optimizing designs and processes must still be verified for their reliability in the food sector [7]. This is mainly due to the unpredictable quality and characteristics of the input, the complex material properties, and the multiscale, multiphase, and multiphysics coupling in combination with mechanistic kinetics of biological and chemical processes involved [3]. This also means that food process models are highly diverse in the domain and one general implementation approach such as for computational fluid dynamics, although advocated [3,8], is not available today. Based on computational intelligence and machine-learning methods, data-driven modeling can be a solution as robust as data are reliable.
Following the Chemical Engineering approach, the first step for implementing a process optimization study in terms of resource consumption through the use of digital twins is the exact quantification of material and energy flows. To obtain accurate data, it is necessary to perform data reconciliation, a model-based technique that reduces measurement errors by making use of redundancies in process data [9]. In particular, data reconciliation is right at the core of data rectification, a collection of techniques for correcting data which comprises also the steps of variable classification, gross error detection, and parameter estimation [10]. Indeed, measurements of process variables, both on- and off-line, are subject to gross errors (outliers), and random errors (noise), and, consequently, the collected data generally do not even satisfy basic process constraints, such as mass and energy balances [9]. Inaccurate process data could result in misleading conclusions, thus leading to poor decisions that can adversely affect many plant functions. Thus, to reduce the impacts of measurement errors in plant measurements and to increase the value of data accessible through implemented data management systems, data rectification should be employed [9,11].
Compared to the use of raw data, reconciled values allow a more accurate evaluation of the overall process yield while helping to better identify energy inefficiencies. Only employing reconciled data, it will be possible to create a reliable numerical model of the plant, i.e., digital twin, to have an exhaustive representation of the system on which operating and simulating material and energy flows to identify inefficiencies and the best strategy for process techno-economic optimization [12,13]. Digital twins shall be operated on-line, directly connected to the stream of reconciled data. In order to do so, the digital twin solutions must be fast enough, and this cannot be obtained with complex process models, thus strategies like model reduction, e.g., surrogate modeling [14,15] or computation power increase, especially with cloud implementations [16], are necessary.
This paper relates to the data reconciliation methodology applied to a real case study, an industrial orange juice production line receiving fresh raw material (oranges) and processing it into concentrated orange juice 65 °Bx. The goal of this article is to guide researchers in the understanding and application of an effective methodology (not yet applied to the food sector but only to the chemical and pharma ones) in order to build a robust basis of verified data for the creation of accurate digital models (not part of the current paper) resembling industrial food processes. The article is structured as follows: Section (II) elaborates the mathematical background in the data reconciliation applied to the process material flows and introduces the software used for the energy assessment. Section (III) shows the results of both the data reconciliation and the energy analysis. Section (IV) concludes the paper and gives suggestions for further research.

II. METHODS AND TECHNIQUES

A. Industrial application

In this paper, an Italian manufacturer of fruit processing machines, the Bertuzzi Food Processing S.r.l., located in Busto Arsizio, Varese, was selected as a case study. In eighty years of activity, the Company designed and engineered more than 1000 fruit processing plants which were commercialized in over 100 countries. The attention was here focused on a concentrated orange juice production line. The complete process description is explained in section (III, **A**). Each material (oranges and juice) flux was measured, and an energetic assessment was performed over the three main operational units namely the juice evaporation, and the two-step thermal treatment units in accordance with their time-temperature conditions.

B. Steady-state data reconciliation

The key parameter in steady-state data reconciliation is the degree of redundancy (DOR) defined as the minimum amount of process variables that can be independently determined by the remaining variables [17]. DOR is measured (Eq. 1) as the sum between the number of material balances (equations) that can be performed at each

Print ISSN	2682-3993
Online ISSN	2682-4000

operational unit level and the number of measurable material fluxes (measures) minus the total number of fluxes that can be reconciled (reconciled).

$$DOR = equations + measures - reconciled \qquad (1)$$

Once this information is collected, and the DOR is verified as higher than zero, data reconciliation can be performed. At this point, there are two possible scenarios. In the first case, all process streams are measured, thus all the variables can be reconciled. In the second case, several unmeasured variables exist, and the problem must be split into two sub-problems. The first sub-problem consists in reconciling the measured variables, while the second consists in calculating the unmeasured ones. The technique applied to solve the unmeasured variables scenario is the QR factorization method [10, 18-20]. This work deals with the case of existing unmeasured variables.

Let us decouple the measured values from the ground truth-value and the normally distributed random error, with zero mean and known variance as described in Eq.2:

$$y = x + \varepsilon \qquad (2)$$

with y being a (n × 1) vector of measured variables, x being a (n × 1) vector of the measured variables truth-values, and ε being a (n × 1) vector of random errors. Considering x as the vector of measured variables and u as the vector of unmeasured variables, then the material balance under study can be represented by Eq. 3:

$$Axx + Auu = 0 \qquad (3)$$

where matrices Ax (m × n) and Au (m × p) have known constants, and u is a (p × 1) vector of unmeasured variables. In order to solve the data reconciliation in presence of measurement errors, it is necessary to solve the following least square problem (Eq. 4):

$$\min_x (y - x)^T \Sigma^{-1} (y - x) \qquad (4)$$

with Σ being the (n × n) covariance matrix of random error ε, subject to Eq. 3.

By introducing a projection matrix, P [21] defined according to Eq. 5 such that:

$$PAu = 0 \qquad (5)$$

it is possible to remove from Eq. 3 the terms Au and u, leaving only (Eq. 6):

$$PAxx = 0 \qquad (6)$$

The problem is now finding a matrix P such that Eq. 5 is respected. In order to do so, the QR factorization method is introduced [10,19] where matrix Au (Eq. 7) is factorized into the product of the orthogonal matrix or Q(m × m) and the upper triangular matrix R(m× p).

$$A_u = QR = [Q_1 \;\; Q_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \qquad (7)$$

In Eq. 7, R1 is defined as a (p × p) nonsingular upper triangular matrix, while Q is split into Q1 and Q2 based on the dimension of matrix R1. By multiplying both sides with $[Q_1^T, Q_2^T]$T of Eq. 7, the following Eq. 8 is obtained:

$$\begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} A_u = \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \qquad (8)$$

thus showing that Q_2^T is the sought-after projection matrix from Eq. 5 (Eqs. 9 and 10):

$$Q_2^T A_u = 0 \qquad (9)$$

$$Q_2^T = P \qquad (10)$$

The values of vector x can now be reconciled by combining Eqs. 10 and 4, thus giving Eq. 11:

$$\hat{x} = y - \Sigma(PA_x)^T [(PA_x)\Sigma(PA_x)^T]^{-1} (PA_x)y \qquad (11)$$

The reconciled solution vector $\hat{x}$ can now be exploited in Eq. 5 to find the solution vector $\hat{u}$ (Eq. 12):

$$\hat{u} = -R_1^{-1} Q_1^T A_x \hat{x} \qquad (12)$$

The implementation of steady-state data reconciliation method mentioned above was operated through MATLAB.

C. Energetic assessment

Energetic data assessment was performed using the Aspen HYSYS v11.0 software. Unitary operations namely first thermal treatment, concentration, and second thermal treatment, were mathematically modeled, and, in particular, the energy balance was performed. By inserting the reconciled steady-state conditions, the heat exchanges (process enthalpy), based on mass and energy balances, and thermodynamics, allow getting information about vapor-liquid equilibrium and heat transfer. Detailed knowledge about the specific heat capacity of food products is required for the design, performance evaluation, operation, and optimization of the heat transfer equipment [22]. In addition to water content, soluble solids, such as sugar, pectin, organic acids, water-solute, and solute-solute interactions play an essential role in the values of sugar-rich systems like juices [22]. The assumption used for this work was to consider the orange juice as a mixture of just water and sugars, i.e., glucose, fructose, and sucrose. The thermodynamic modeling of this mixture was performed with an NRTL thermodynamic equation of state [23] with specific binary interaction parameters. Steam properties were calculated using the NIST REFPROP [24] thermodynamic package implemented in Aspen HYSYS, with the assumption of considering it composed uniquely by water. The software contains various modules that can be combined to give a full

Print ISSN 2682-3993
Online ISSN 2682-4000

representation of the unitary operations to be described [24]. In particular, the plate heat exchanger module was used to simulate both thermal treatment operations, while the double-effect evaporator was applied for modeling the concentration step.

III. RESULTS AND DISCUSSION

A. *Industrial process description and data collection*

The processing line is equipped with an electric and control board with PLC (Programmable Logic Controller). The main operating functions of the plant, such as motors and the electro valves status, and important process variables as temperature, pressure, flow, levels, PID (Proportional Integral Derivative) process regulator status, alarms, and signalization of activated interlocks, are controlled by an operator. The automatic Cleaning-In-Place station to wash the main plant components by means of hot basic detergent (75 °C), hot acid detergent (60 °C), and rinse using water was not part of the current study.

A block flow diagram was drawn in order to have a clear representation of the system and its operational units. Through such industrial orange juice production line (Fig. 1), 5000 kg of fresh oranges can be processed into around 271 kg of concentrated orange juice per hour. The fresh raw material is washed with normal and then ionized water, brushed, and scrubbed to remove dirt and part of the essential oils from the outer part of the orange peels. The subsequent calibration ensures to keep in the process line only fruits with a diameter between 45 and 100 mm. Such fruits are then squeezed and a juice rich in fibers is collected. The latter enters a helicoidal extractor characterized by a slow rotation speed to remove fiber parts while avoiding the inclusion of air bubbles within the product. The following step is juice first thermal treatment operated in a plate heat exchanger where the juice temperature is raised up to 95 °C for 30 s. Such conditions are fundamental to inactivate the pectin-esterase enzyme which may favor cloud instability in citrus juices [25] and which is responsible for browning reactions [26]. After going through high-speed decanting for additional fibers removal, the juice enters a depressurized deaerator. Here, essential oils and remaining air bubbles are removed. Hence, the juice is ready for excess water removal within a dual-effect evaporator. The first effect consists in raising the juice temperature to about 75 °C–78 °C for a typical residence time of 60 s, while the second effect allows the juice to reach 45 °C–47 °C for additional 60 s. Subsequently, the concentration unit allows reaching the requested dissolved sugar concentration of 65 °Bx within the product (juice). The final processing step before the aseptic filling is the second thermal treatment. Such operation is performed in a plate heat exchanger where juice temperature reaches 95 °C for 1 min. Organic waste is generated at the following unitary operations level: manual sorting - fruits with physiopathies are discarded, calibration - fruits out of size are removed, squeezing - peels and seeds are withdrawn, and refining (both with helicoidal extractor and decanter) - pulp is separated from the juice and discarded. During juice concentration, water is evaporated. As represented in Fig. 1, 20 are the material fluxes both as input and output of the 13 operational units.

Material flowrates used for this study were retrieved from a two-hour operation time frame with a sampling of 1 minute for each time step, thus generating a database of 240 points.

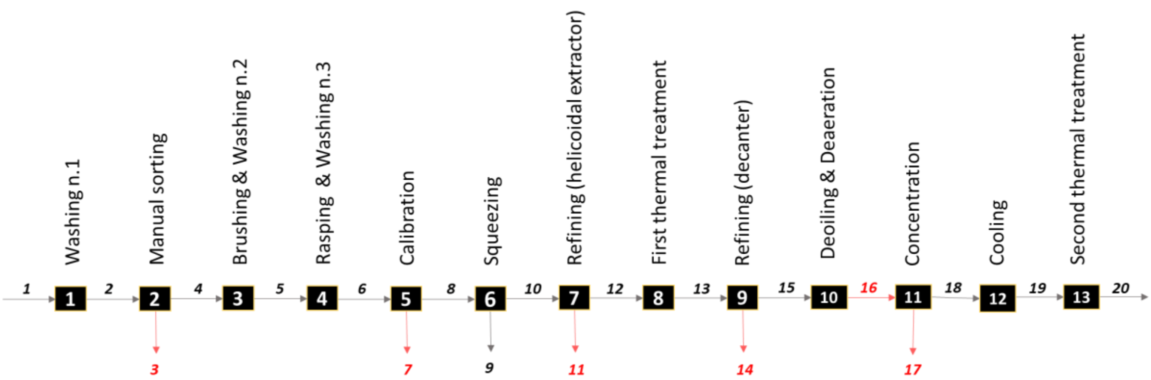

Fig. 1: Block flow diagram of the industrial orange juice production line. The 13 black boxes represent the sequential operational units, while the numbers reported in italics are referred to the 20 material flows. Unmeasured material fluxes are highlighted

Print ISSN 2682-3993
Online ISSN 2682-4000

B. Material flows data reconciliation

As shown in Fig. 1, 20 material fluxes and 13 material balance equations (one at each operational unit) could be counted. Recalling that the definition of redundancy (Eq. 1) is the difference between the amount of available information (constraints and measurements) and the total number of fluxes to be reconciled, the system had a highly positive DOR:

$DOR = 14 + 13 - 20 = 7$

Therefore, the problem could be solved and 14 out of 20 streams could be reconciled following the method described in Section (II, **B**). The six remaining unmeasured streams could only be calculated and not reconciled. The coefficients of the resulting equations were used to compose the 13 × 20 matrix A, as represented in Fig. 2. Mean and variance values of streams are shown in Fig. 3 and 4, together with reconciliation results.

Results of the data reconciliation procedure are shown in Fig. 3 and 4, for mean and variance analyses, respectively. Relative deltas, in Fig. 3, show that the process was indeed operating on a steady-state basis with presumably normally distributed measurement errors, since the mean values of reconciled data and raw data showed very little difference, even though variance was effectively reduced for all streams (as shown in Fig. 4), except for the unmeasured streams and stream 15 which had a too small local DOR (being between two unmeasured streams). Plots of resulting reconciled versus raw data are shown in Fig. 5 for streams 1, 11, 13, and 20, and in Fig. 6 and Fig. 7 for streams 10 and 7, respectively.

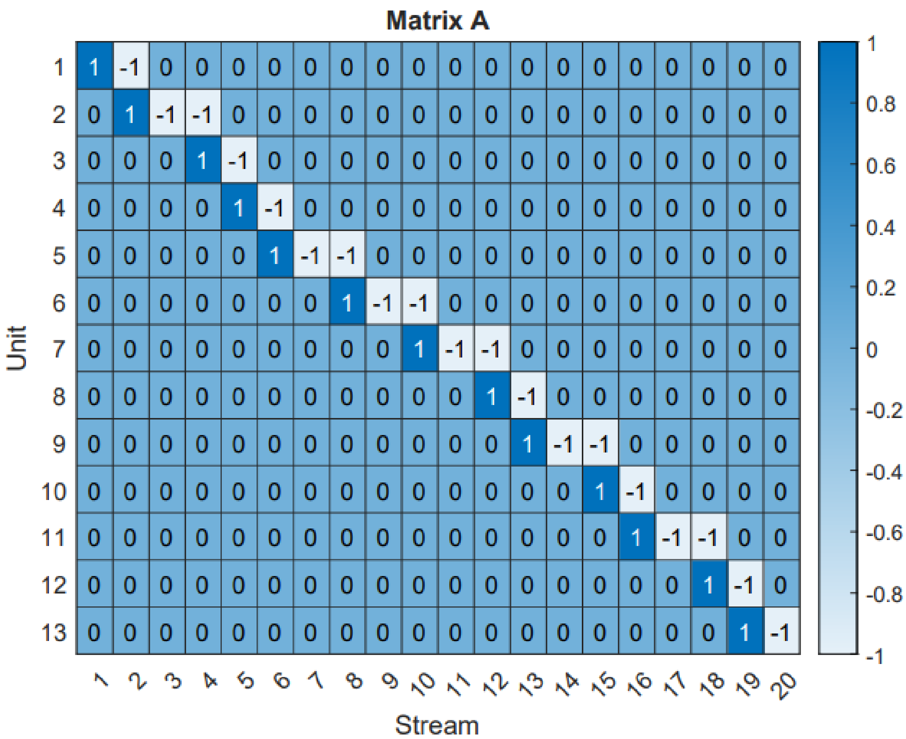

Figure 2: Matrix of the material balances coefficients

Print ISSN 2682-3993
Online ISSN 2682-4000

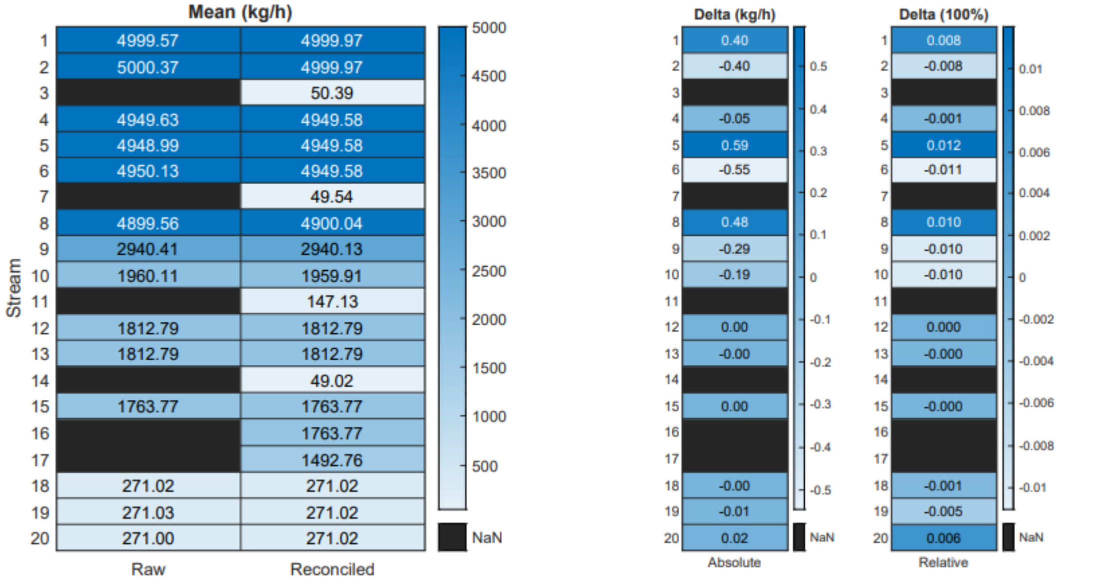

Figure 3: Mean values of raw versus reconciled data with absolute and relative differences, on the right. Values were calculated for the two-hour investigated time frame. Unmeasured (NaN) streams are represented by black cells

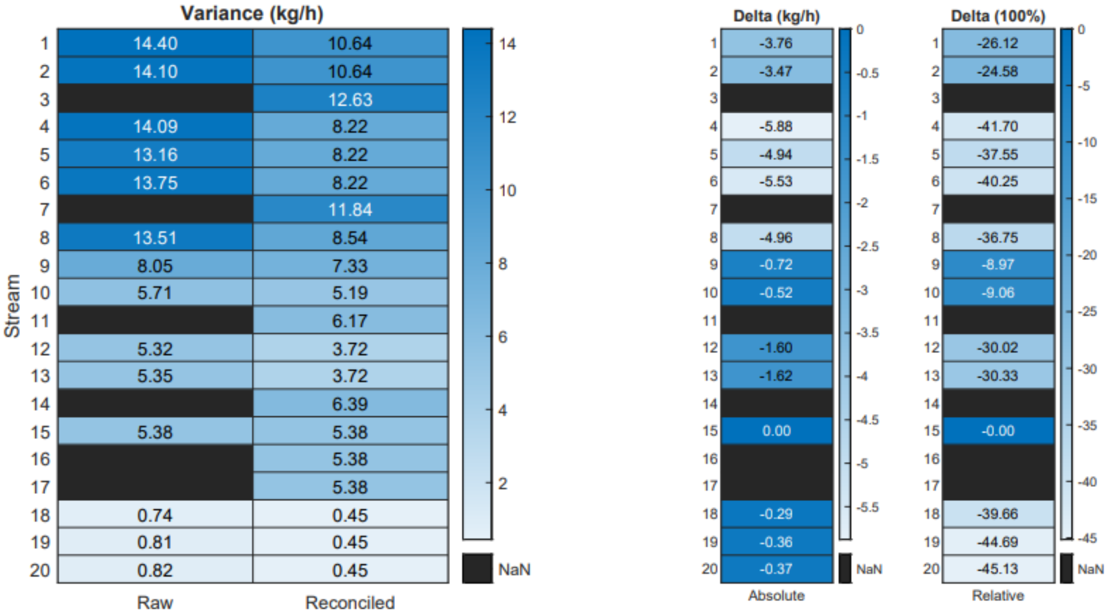

Fig. 4:Analysis of the variance calculated between measured and reconciled data for the two-hour investigated time frame. Variance reduction with data reconciliation can be seen in the two columns on the right. Unmeasured (NaN) streams are represented by black

Print ISSN 2682-3993
Online ISSN 2682-4000

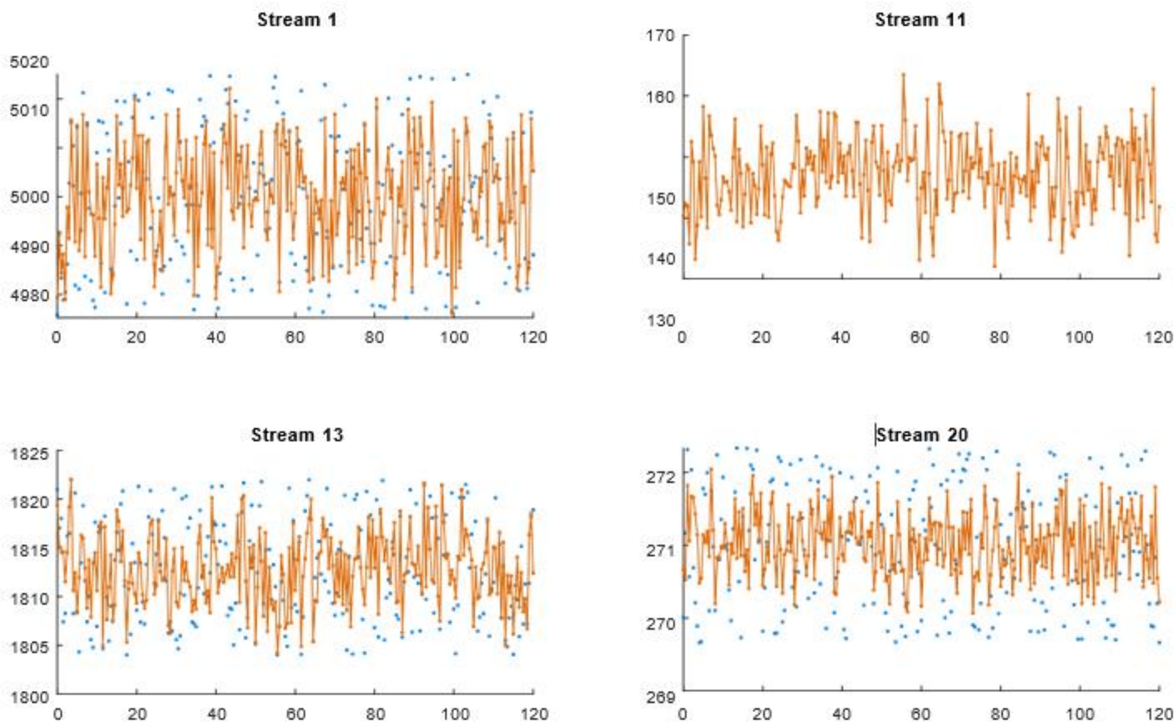

Fig. 5:Data reconciliation results for selected streams (1, 11, 13, 20): experimental data (blue dots), reconciled data (orange dots). Flowrate (kg/h) is displayed on the y-axis, while time (min) is on the x-axis.

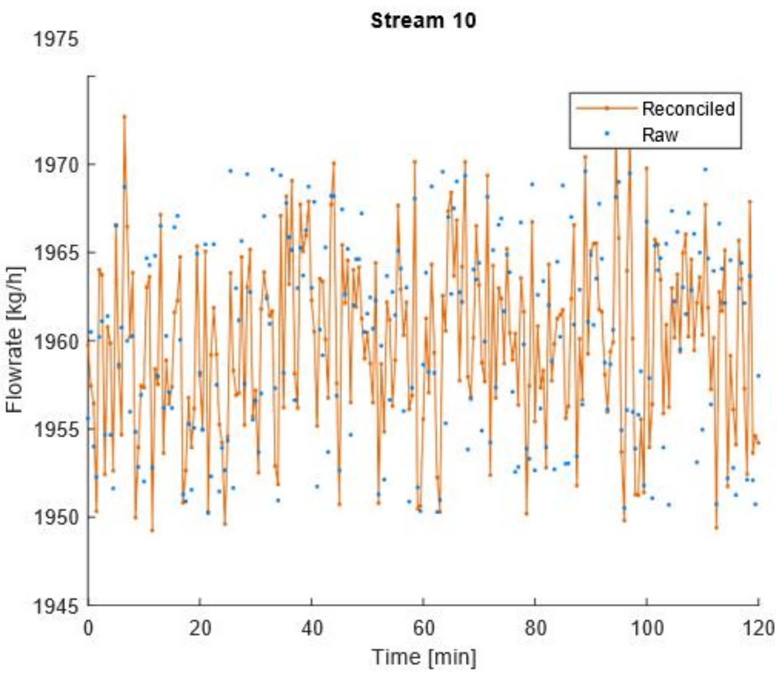

Fig. 6:Stream 10 (measured) data reconciliation results

Print ISSN 2682-3993
Online ISSN 2682-4000

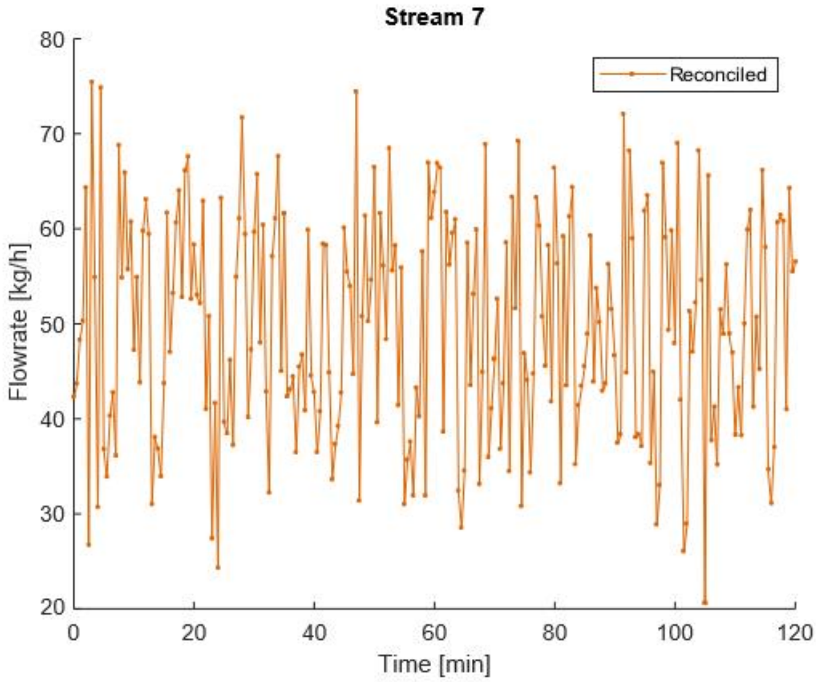

Fig. 7:Stream 7 (unmeasured) data reconciliation results

C. Energy data assessment

The energetic analysis performed with Aspen HYSYS allowed the evaluation of latent heat exchanges and the amount of energy expenditure or intake by steam and juice, respectively. The NIST REFPROP database v.10.0 [27], already implemented in Aspen HYSYS, was used to calculate the fluid thermodynamic and transport properties of pure water, i.e. steam. Raw orange juice is composed mainly of water, sugars (8 %), proteins (0.7 %), and fibers (0.2 %)[28]. For this study, the fiber content could be easily neglected due to the double decanting that is performed in the process. Proteins, on the other hand, are not so easily removed but the added complexity in terms of thermodynamic modeling should not make a noticeable impact in this study given their low concentration, and other approximations used. The juice mixture was then modeled as a mixture of water and sucrose, fructose, and glucose in a 2.1:1.1:1 ratio, respectively [29]. The thermodynamic model chosen for this mixture was the NRTL [23] with an ideal gas approximation for the vapor phase. The binary interaction parameters, for water and each sugar only, were taken from the NISTV100.NIST-IG data bank existing in Aspen Plus was created in collaboration with NIST.

In the case under analysis, the energetic assessment was performed on the three operational units that foreseen an active heat exchange, namely first thermal treatment, concentration, and second thermal treatment. As mentioned in Section (II ,**C**), the plate heat exchanger module was used to simulate both thermal treatment operations, while the double-effect evaporator was applied for modeling the concentration step (Fig. 8). Only water vapor was used as process duty fluid. Exergy efficiency (also known as the second-law efficiency) was assumed to be equal to 100 %. Through the process simulation software, it was possible to verify the factual thermal energy exchanges and, therefore, to identify the enthalpy jumps within the adiabatic system Table 1. To carry out the following energetic analysis, the chosen candidate operative point was the first reconciled time step, i.e., the first point in time in the two-hour set of reconciled data. The reason is that the use of real-time reconciled data would be coupled directly on-line with digital twins to give live insights into the process operation, time step by time step.

In a co-current dual effect evaporator, as shown in Fig. 8, the vapor stream exiting the first effect was used to heat the stream entering the second effect, after an isenthalpic expansion of 0.75 bar of the saturated liquid stream exiting the first effect through valve VLV-100 (see Fig. 8). In this configuration, the external heat was supplied only to the first effect evaporator, so the energetic assessment was performed only in this regard. The heat transfer medium of the three examined operational units was always steam generated within a boiler (not included in the present study). Steam was introduced as saturated (vapor fraction equal to 1) at the pressure of 8 barg and temperature of approximately 175 °C. The two-step thermal treatment was performed in plate heat exchangers, as represented in Fig. 8. Here, steam and juice entered in counter-current and, respectively, released and uptook thermal energy spreading out over the plates that separated them. During both operations, juice entered with a

Print ISSN 2682-3993
Online ISSN 2682-4000

temperature of around 25 °C at atmospheric pressure and exited with a temperature of 95 °C. The temperature of 95 °C was typically maintained for 30 s and 1 min of residence time in the first and second steps of thermal treatment, respectively. The calculated heat flows resulted to be 141.55 kW during the first thermal treatment and 12.64 kW during the second one. In the latter case, due to the lower flow rate of juice to be processed (1819.32 kg/h were thermally treated during the first step while 270.67 kg/h were treated during the second one), the steam mass flow required was far lesser than the former. On the other hand, the evaporator was a more complex system that worked under low-pressure conditions in such a way as to reduce the water boiling temperature and, therefore, to avoid juice thermal damage. The juice flow entering the dual effect evaporator was 1764.72 kg/h (11.8 °Bx). In normal process operative conditions, output stream 19 (with a reconciled flowrate of 1491.56 kg/h, in the first time step) was supposed to be concentrated to 65 °Bx. However, data reconciliation results coupled with the HYSYS simulation made it possible to show that during the operation time frame investigated the concentration unit operation was working over specification, by concentrating the juice up to 76 °Bx as shown in Table 2. The resulting enthalpy balance was equal to 558.11 kW, the highest among the three investigated operational units.

Table 1:Energetic assessment of two-step thermal treatment and concentration units performed with Aspen HYSYS. It was applied only to the duty currents (external energy)

			First thermal treatment	**Concentration (E-100)**	**Second thermal treatment**
Steam		**Flow (kg/h)**	251.00	989.64	22.41
	Inlet	**Vapor Fraction**	1	1	1
		Temperature (°C)	175.4	175.4	175.4
		Pressure (barg)	8	8	8
	Outlet	**Vapor Fraction**	0	0	0
		Temperature (°C)	175.4	175.4	175.4
		Pressure (barg)	8	8	8
Juice		**Flow (kg/h)**	1819.32	1764.72	270.67
	Inlet	**Vapor Fraction**	0	0	0
		Temperature (°C)	25	60	25
		Pressure (atm)	1	1	1
	Outlet	**Vapor Fraction**	0	0.482	0
		Temperature (°C)	95	101	95
		Pressure (atm)	1	1	1
		Heat Flow (kW)	141.55	558.11	12.64

Note: Comprehensive streams' data are available in the Supplementary Material file.

Print ISSN 2682-3993
Online ISSN 2682-4000

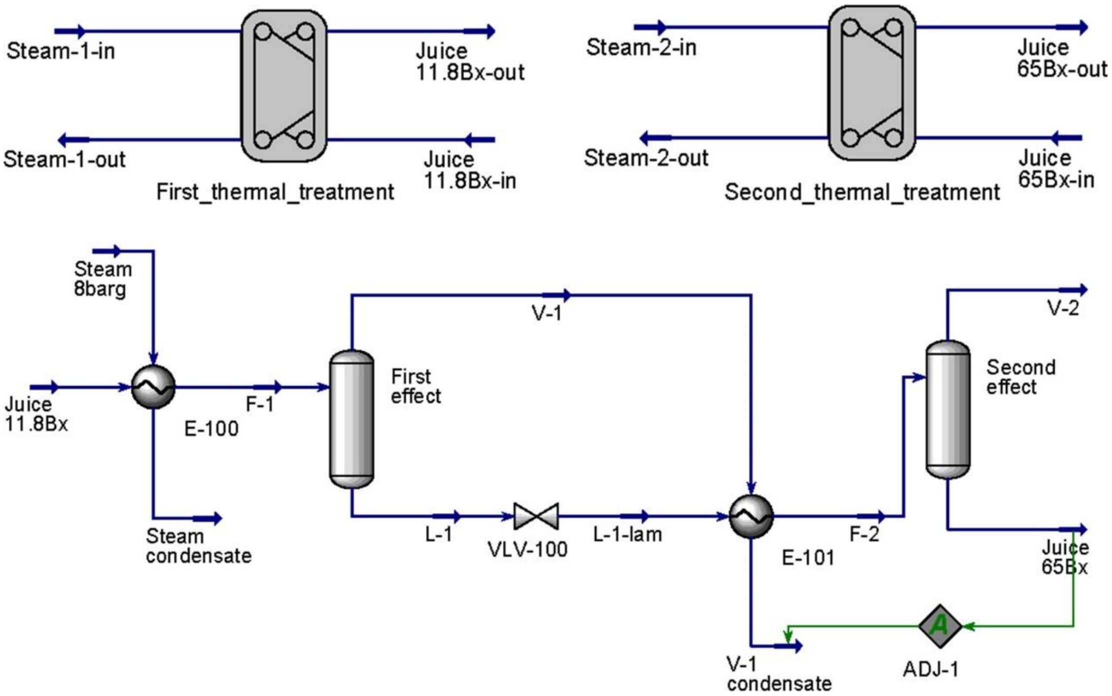

Fig. 8:Operational units represented using Aspen HYSYS: first thermal treatment (top left), second thermal treatment (top right), concentration unit (bottom)

Table 2:Simulation of concentration outlet properties (stream 18) using (a) the reconciled flow value and (b) the design specification of the desired concentration

		Inlet	Outlet^a	Outlet^b
Flow (kg/h)		1764.72	**270.67**	319.59
Composition (−)	**Fructose**	0.031	0.192	0.169
	Glucose	0.028	0.183	0.155
	Sucrose	0.059	0.385	0.326
	Water	0.882	0.240	0.350
Concentration (°Bx)		11.8	76.0	**65.0**

IV. CONCLUSIONS

Agrifood 4.0 era is characterized by cyber-physical systems and internet-oriented technologies. The possibility to operate several simulations of a real process in order to find out the optimal operative conditions without wasting edible raw materials and energy is being made possible by the spreading of digital twins in the food sector. However, data recovery and validation are often the bottlenecks in drawing an accurate virtual copy of the real industrial system. In the present paper, a well-known industrial food process, namely the production of concentrated orange juice, was taken as an example to be analyzed in terms of quantified sustainability and, more specifically, for the application of a data validation methodology, data reconciliation, as a necessary primary step to implement further digitalization approaches. Data reconciliation means applying a quantitative mathematical procedure to measure material and energy flow along the process to estimate their accuracy and, eventually, correct them.

For the specific case study, reconciled values of material flows were found as the factual descriptors of the process. The applied method allowed to find out the appropriate corrections to be made on measured data in order to close material balances highlighting the errors behind those values either measured by plant operators and generated by in-line sensors. The energetic assessment, carried out with Aspen HYSYS, enabled outlining enthalpy changes within the operational units characterized by thermal energy exchanges (two-step thermal treatments and

Print ISSN 2682-3993
Online ISSN 2682-4000

concentration). Thanks to reconciled flow streams, it was possible to detect that the process was working over specification, thus wasting steam, i.e. energy.
Based on reconciled data, the creation of a precise digital twin resembling the reality under study is expected as the future development of this research through which real industrial process data were measured, collected, and validated. The digital twin will mimic several process conditions geared towards finding the optimal one. Possibly coupling the digital twin with online reconciled data would offer live insights to decision makers for the industrial process.

REFERENCES

[1] Zeb A, Soininen JP, Sozer N. Data Harmonisation as a Key to Enable Digitalisation of the Food Sector: A Review. Food Bioprod Process 2021;127:360–370.
[2] FAO, Food and Agriculture Organization of the United Nations. The Water-Energy-Food Nexus: A New Approach in Support of Food Security and Sustainable Agriculture; 2014. https://www.fao.org/3/bl496e/bl496e.pdf
[3] Verboven P, Defraeye T, Datta AK, Nicolai B. Digital Twins of Food Process Operations: The next Step for Food Process Models? Curr Opin Food Sci 2020;35:79–87.
[4] Teng SY, Tous M, Leong WD, How BS, Lam HL, Masa V. Recent Advances on Industrial Data-Driven Energy Savings: Digital Twins and Infrastructures. Renewable Sustainable Energy Rev 2021;135:110208.
[5] Rezek Jambrak A, Nutrizio M, Djekic I, Pleslic S, Chemat F. Internet of Nonthermal Food Processing Technologies (IoNTP): Food Industry 4.0 and Sustainability. Appl Sci 2021;11:686.
[6] Defraeye T, Shrivastava C, Berry T, Verboven P, Onwude D, Schudel S, Buhlmann A, Cronje P, Rossi RM. Digital Twins Are Coming: Will We Need Them in Supply Chains of Fresh Horticultural Produce? Trends Food Sci Technol 2021;109:245–258.
[7] Girotto F, Galeazzi A, Manenti F, Gueguen S, Piazza L. Water–Food–Energy Nexus: Assessing Challenges in the Trend toward Digitalization: The Case Study of an Italian Winemaking Industry. Environ Prog Sustain Energy 2022;e13893.
[8] Datta A K. Toward Computer-Aided Food Engineering: Mechanistic Frameworks for Evolution of Product, Quality and Safety during Processing. J Food Eng 2016;176:9–27.
[9] Camara MM, Soares RM, Feital T, Anzai TK, Diehl FC, Thompson PH, Pinto JC. Numerical Aspects of Data Reconciliation in Industrial Applications. Process 2017;5:56.
[10] Narasimhan S, Jordache C. Data Reconciliation and Gross Error Detection: Elsevier; 1999.
[11] Johnston LPM, Kramer MA. Maximum Likelihood Data Rectification: SteadyState Systems. AIChE J 1995;41:2415–2426.
[12] Battisti R, Galeazzi A, Prifti K, Manenti F, Machado RAF, Marangoni C. Techno-Economic and Energetic Assessment of an Innovative Pilot-Scale Thermosyphon-Assisted Falling Film Distillation Unit for Sanitizer-Grade Ethanol Recovery. Appl Energy 2021;297:117185.
[13] Prifti K, Galeazzi A, Barbieri M, Manenti F. In Computer Aided Chemical Engineering; Montastruc L, Negny S, Eds; 32 European Symposium on Computer Aided Process Engineering; Elsevier, 2022; Vol. 51; pp 1321–1326.
[14] Ors E, Schmidt R, Mighani M, Shalaby M. A Conceptual Framework for AI-based Operational Digital Twin in Chemical Process Engineering. 2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC) 2020; pp 1–8.
[15] Galeazzi A, Prifti K, Gallo F, Manenti F. In Computer Aided Chemical Engineering; Montastruc L, Negny S, Eds; 32 European Symposium on Computer Aided Process Engineering; Elsevier, 2022; Vol. 51; pp 1543–1548.
[16] Galeazzi A, Nasti R, Bozzano GL, Verotta L, Marzorati S, Manenti F. In Computer Aided Chemical Engineering; Turkay M, Gani R, Eds; 31 European Symposium on Computer Aided Process Engineering; Elsevier, 2021; Vol. 50; pp 1035–1040.
[17] Manenti F, Grottoli MG, Pierucci S. Online Data Reconciliation with Poor Redundancy Systems. Ind Eng Chem 2011;50:14105–14114.
[18] Sanchez M, Romagnoli J. Use of Orthogonal Transformations in Data Classification-Reconciliation. Comput Chem Eng 1996;20:483–493.
[19] Romagnoli JA, Sanchez MC. Data Processing and Reconciliation for Chemical Process Operations: Elsevier Science & Technology; 1999.
[20] Bisotti F, Galeazzi A, Gallo F, Manenti F. In Simulation and Optimization. In Process Engineering; Bortz M, Asprion N, Eds; Elsevier, 2022; pp 161–199.
[21] Crowe CM, Campos YAG, Hrymak A. Reconciliation of Process Flow Rates by Matrix Projection. Part I: Linear Case. AIChE J 1983;29:881–888.

Print ISSN 2682-3993
Online ISSN 2682-4000

[22] Sanchez-Romero MA, Garcıa-Coronado P, Rivera-Bautista C, Gonzalez-Garcıa R, Grajales-Lagunes A, Abud-Archila M, Ruiz-Cabrera MA. Experimental Data and Predictive Equation of the Specific Heat Capacity of Fruit Juice Model Systems Measured with Differential Scanning Calorimetry. J Food Sci 2021;86:1946–1962.
[23] Renon H, Prausnitz JM. Local Compositions in Thermodynamic Excess Functions for Liquid Mixtures. AIChE J 1968;14:35–144.
[24] Damartzis T, Michailos S, Zabaniotou A. Energetic Assessment of a Combined Heat and Power Integrated Biomass Gasification–Internal Combustion Engine System by Using Aspen Plus®. Fuel Process Technol 2012;95:37–44.
[25] Bayındırlı A, Alpas H, Bozoglu F, Hızal M. Efficiency of High Pressure Treatment on Inactivation of Pathogenic Microorganisms and Enzymes in Apple, Orange, Apricot and Sour Cherry Juices. Food Control 2006;17:52–58.
[26] Quoc AL, Mondor M, Lamarche F, Ippersiel D, Bazinet L, Makhlouf J. Effect of a Combination of Electrodialysis with Bipolar Membranes and Mild Heat Treatment on the Browning and Opalescence Stability of Cloudy Apple Juice. Food Res Int 2006;39:755–760.
[27] Lemmon EW, Bell I, Huber ML, McLinden MO. NIST Standard Reference Database 23: Reference Fluid Thermodynamic and Transport Properties-REFPROP, Version 10.0; 2018.
[28] U.S. Department of Agriculture. Food Data Central. Orange juice, raw (Includes foods for USDA’s Food Distribution Program) SR LEGACY, 2018; 169098.
[29] Zhang JX, Ritenour MA. Sugar composition analysis of commercial citrus juice products. Proceedings of the Florida State Horticultural Society 2016;129:178–180. ISSN 0886-7283.