DETERMINING COMMINGLING VOLUME AND ITS UNCERTAINTY IN BATCH TRANSFERS FROM NOISY DATA

Lauro Carvalho, e-mail: carvalholauro@hotmail.com[1,2,3]
Maria Laura Martins-Costa, e-mail: laura@mec.uff.br[2,3,4]
Rogério Martins Saldanha da Gama, e-mail: rsggama@gmail.com[5]
Felipe Bastos de Freitas Rachid, e-mail: rachid@vm.uff.br[3,4]

[1]Graduate Student
[2]Laboratory of Theoretical and Applied Mechanics (LMTA)
[3]Mechanical Engineering Graduate Program (PGMEC)
[4]Department of Mechanical Engineering (TEM)
Universidade Federal Fluminense - Rua Passo da Pátria, 156 , Niterói, RJ, 24210-240 - Brazil
[5]Mechanical Engineering Graduate Program (FEN)
Universidade do Estado do Rio de Janeiro - Rua São Francisco Xavier, 524, Rio de Janeiro, RJ, 20550-013 - Brazil

Abstract: *This paper presents a statistical methodology to compute the commingling volume and its uncertainty that takes place in batch transfers in pipelines. The method is based on the mass density signature of the products which is acquired in typical SCADA systems. Its performance is tested against noisy and filtered data, generated through the well-known theoretical concentration profile by assuming constant dispersion coefficient and constant flow rates. The nosy data we have been filtered by employing the Savitzky-Golay smoothing technique. The obtained results for low and high noise amplitudes and frequencies have shown the if the data are smoothed, then the methodology has almost always been able to compute the commingling volume and its uncertainty with great accuracy.*
Keywords: *batch transfer, commingling volume, noisy data, digital filters, uncertainty analysis*

1. INTRODUCTION

The transfer of dissimilar products through a same pipeline is a common practice in the petroleum industry and is usually known as batch transfer (Aunicky, 1970; Austin and Palfrey, 1964; Baptista et al., 2000a; Freitas Rachid et al., 2002; Krantz and Wasan, 1974; Netchval et al., 1972; Sjenitzer, 1958; Smith and Schulze, 1948a,b; Taylor, 1954). When this task is carried out without using scrapers or pigs to separate the products, a commingling zone develops at the products' interface, and increases in extent as it travels along the line. By the time this zone reaches the receiving point, the commingling volume, as well as its uncertainty, must be properly identified and determined, so that it can be segregated and accommodated in a separate tank, to be shipped back to the refinery for later reprocessing. The feasibility of such operation depends heavily on the proper identification of the beginning and the end of commingling zone in real time during the transfer (Freitas Rachid et al., 2002; Melo and Freitas Rachid, 2010). It is usually done by continuously monitoring and acquiring a physical property of the products, capable to discern them, such as density and sonic velocity among others Baptista et al. (2000b). However, in real world applications such data are not free of noise, what poses difficulties and sometimes compromises this task. To get rid of such an inconvenience, it is proposed in this paper a modified statistical strategy capable to not only identify the ends of the commingling zone and determine its volume for any pair of admissible contamination previously established, but also to automatically compute its uncertainty, even in noisy data. To test and validate the performance and robustness of the proposed methodology, numerical examples are presented for data generated with different levels of noise, which are artificially introduced in the mass density of the products through the concentration profile predicted by a classical well-known theoretical model with constant dispersion coefficient and constant flow rates.

2. PROBLEM STATEMENT

Aiming at being better acquainted with the batch transfer process and its terminologies, it is convenient to describe how such a transfer is conducted in practice and also how the commingling volume is computed. Consider the schematic pipeline installation shown in Fig. 1 which is used to sequentially pump two distinct products designated as "A"and "B". The pipeline has a constant diameter D and a length L, which is measured from the junction of the pump discharge lines until the receiving point at the other pipeline end. By an appropriate valve switching, at time instant $t = 0$, the pumping of fluid "A"is interrupted and, at the same time, the displacement of fluid "B"is started, so that fluid "B"begins to push fluid "A"along the line. Let us designate by $C_\imath(x,t) \in [0,1]$, with $\imath \in \{A, B\}$, the time-averaged mean concentration in volume of fluid $\imath$ within the mixture at the cross-section of the pipeline as a function of the spatial position x along the pipe and the time t. Thus, the beginning of the sequential transfer can be described in terms of C_B by,

$$C_B(x = 0^-, t = 0) = 1 \text{ and } C_B(0^+ \leq x \leq L, t = 0) = 0 \tag{1}$$

or in terms of C_A by,

$$C_A(x = 0^-, t = 0) = 0 \text{ and } C_A(0^+ \leq x \leq L, t = 0) = 1 \tag{2}$$

since, for all (x, t), the following relationship must hold:

$$C_A + C_B = 1. \tag{3}$$

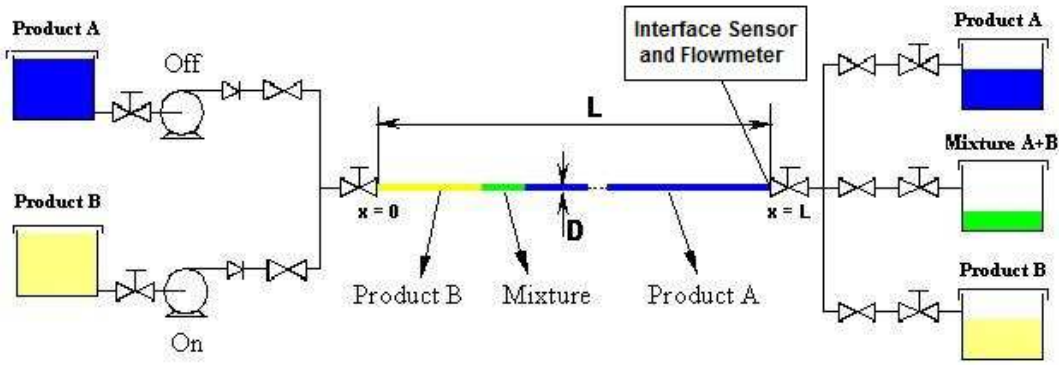


Figura 1: Schematic representation of a batch transfer installation in a pipeline.

During the passage of the products through the pipeline a mixture zone is formed at the boundary of the two adjacent products. Such a zone of mixture increases in length as it travels along the line and can be characterized by the interval $[x_f, x_i] \subset [0, L]$, with $x_f = x_f(t)$ and $x_i = x_i(t)$, such that $1 = C_B(x = x_f, t) > C_B(x, t) > C_B(x = x_i, t) = 0$ or $0 = C_A(x = x_f, t) < C_A(x, t) < C_A(x = x_f, t) = 1$, for $x \in (x_f, x_i)$. When the beginning of the mixture zone reaches the pipeline end, that is $x_i(t) = L$, the flow is directed towards the mixture tank when the products are incompatible, such as gasoline and diesel. Finally, when the end of the mixture zone reaches the position $x_f(t) = L$, the fluid flow is conducted to the tank containing product "B". In practice, small degrees of contamination are allowed to take place at the beginning and at the end of the mixture zone, which are expressed in terms of the maximum admissible concentration C_{BA} and C_{AB}, respectively. Once these values (which are not necessarily equal) have been chosen in such a way that the products technical specifications are not altered, the commingling volume, which should be stored in the tank for later re-processing, is defined as:

$$V_c := \int_{t_{BA}}^{t_{AB}} Q(t)\, dt, \tag{4}$$

in which $Q(t)$ denotes the volumetric flow rate at the receiving point (that is, the position $x = L$), t_{BA} is the time instant associated to the arrival of the beginning of the contaminated interface with concentration C_{BA} at $x = L$ and t_{AB} is the time instant associated to the arrival of the end of the contaminated interface with concentration C_{AB} at $x = L$. In other words, t_{BA} and t_{AB} are such that $C_B(x = L, t = t_{BA}) = C_{BA}$ and $C_A(x = L, t = t_{AB}) = C_{AB}$, respectively.

In order to evaluate the integral in Eq. (4) one must keep track of both the concentration of the products and the flow rate at the position $x = L$, as illustrated in Fig. 1. The passage of the mixture zone at $x = L$ is usually detected by continuously monitoring and acquiring a physical property of the products, capable to discern them, such as the mass density and sonic velocity among others. If mass density is chosen to identify the products, in-line densitometers are used with a constant scan rate denoted herein as Δt, which in general depends on the supervisory control and data acquisition system (SCADA) installed on the line.

3. ACQUISITION SYSTEM AND DATA REDUCTION PROCEDURE

The data reduction procedure used to convert the products' signature into concentration relies on the scan rate (whose reciprocate is the acquisition data frequency) as well as on a theoretical relationship between mass density and concentration. If the mass density is used as a property to discern the products, then this relationship reads as:

$$C_B = \frac{\rho - \rho_A}{\rho_B - \rho_A}, \tag{5}$$

in which ρ is the mass density of the mixture and ρ_A and ρ_B stand for the mass densities of the products "A"and "B"with 100% of purity, respectively. The filed data of mass density $\rho(t_j)$ and flow rate $Q(t_j)$ at $x = L$ during the passage of the batch are recorded in a data file for $j = 1, \cdots, N$; being N the total number of registers. Since the data acquisition was done with constant frequency $\Delta t = t_{j+1} - t_j = constant$.

To better characterize the process used to delimit the mixture zone, consider the following time instants $t_o < t_{BA} < t_{AB} < t_l$ of the database time chronology,

$$t_1, \ldots, t_r, \ldots, t_o, \ldots, t_{p-1}, t_p, t_{p+1}, \ldots, t_{k-1}, t_k, t_{k+1}, \ldots, t_l, \ldots, t_s, o \ldots, t_N, \tag{6}$$

in which t_o stands for the last time instant before the passage of the mixture zone (that is, for $t \leq t_o$ it is registered in the database the presence of fluid "A"with 100% purity) and t_l the first time instant after the passage of the mixture zone (that is, for $t \geq t_l$ it is registered in the database the presence of fluid "B"with 100% puriry).

The data reduction procedure aims at identifying the extreme limits of the time interval $[t_o, t_l]$ in which the mass density is changing due to the mixture process only. If the mass density signature were free of noise mainly due to the local temperature variations and pressure transients in the line, this task would be trivial. However, in real-world applications statistical analyses will be necessary to carry out the task with certain level of confidence. To do so, a step-by-step procedure originally proposed by Baptista et al. (2000b) is presented next:

1. Using the values of $\rho(t_j)$, $j = 1, \cdots, N$, it is determined the time instants t_o and t_l, which characterize the mixture interfaces, as well as the mean values of ρ_A and ρ_B.

2. Based on Eq. (5), $C_B(t_j)$ can be evaluated based on $\rho(t_j)$ for $j = o, o+1, \cdots, l-1, l$, by noting that $C_B(t_o) = 0$ and $C_B(t_l) = 1$.

3. With $C_B(t_j)$, for $j = o, o+1, \cdots, l-1, l$, in one hand the pre-set values of the maximum admissible concentrations C_{AB} and C_{BA} at the other it is estimated the time instants t_{BA} and t_{AB}. Finally, with t_{BA}, t_{AB} and $Q(t_j)$, for $j = o, o+1, \cdots, l-1, l$, a numerical approximation for V_c is evaluated.

If pressure and temperature conditions at the point the mass density data is acquired were kept constants and also there were no errors in the process of data acquisition, then the mass densities $\rho_\imath$, with $\imath \in \{A, B\}$, as well as the time instants t_o and t_l would be promptly determined. In such a case, t_o would be equal to the greatest time instant of the set $\{t_j\}$, $j = 1, 2, \cdots, N$, for which $\rho(t_o) = \rho_A$. Analogously, t_l would be equal to the smallest time instant of the set $\{t_j\}$, $j = o+1, 2, \cdots, N$, for which $\rho(t_l) = \rho_B$. Since the above conditions do not prevail in practice, it becomes necessary to establish a mathematical criterion in order to determine t_o, t_l and $\rho_\imath$, with $\imath \in \{A, B\}$. To achieve this goal, the following statistical procedure is employed.

Consider two arbitrary time instants t_r and t_s, with $t_r \leq t_s$ of the set $\{t_j\}$, $j = 1, 2, \cdots, N$. Let $\bar{\rho}_A$ and $\bar{\rho}_B$ represent, respectively, the arithmetic mean of the mass density samples $\{\rho(t_1), \cdots, \rho(t_r)\}$ and $\{\rho(t_s), \cdots, \rho(t_N)\}$, that is

$$\bar{\rho}_A = \frac{\sum_{j=1}^{r} \rho(t_j)}{r} \quad \text{and} \quad \bar{\rho}_B = \frac{\sum_{j=s}^{N} \rho(t_j)}{N-s+1}. \tag{7}$$

By assuming that the above sets represent observations of a multiple-sample experiment, the confidence interval $\bar{h}_\imath$ associated to the mean value $\bar{\rho}_\imath$, $\imath \in \{A, B\}$, can be expressed as Moffat (1988):

$$\bar{h}_A = Z_A \frac{S_A}{\sqrt{r}} \quad \text{and} \quad \bar{h}_B = Z_B \frac{S_B}{\sqrt{N-s+1}}, \tag{8}$$

in which $Z_\imath$, $\imath \in \{A, B\}$, is the Student's statistic appropriate for the number of samples and a confidence level desired (95% confidence in our case) and $S_\imath$, $\imath \in \{A, B\}$, is the standard deviation of these sets:

$$S_A = \left\{ \frac{\sum_{j=1}^{r} (\rho(t_j) - \bar{\rho}_A)^2}{r-1} \right\}^{1/2}, \quad \text{and} \quad S_B = \left\{ \frac{\sum_{j=s}^{N} (\rho(t_j) - \bar{\rho}_B)^2}{N-s} \right\}^{1/2}. \tag{9}$$

Based on the past definitions, we proposed that t_o and t_l be the time instants t_r and t_s, respectively, of the set $\{t_j\}$, $j = 1, 2, \cdots, N$ which verify the following properties:

1. the number of elements of the sequences $\{\rho(t_1), \cdots, \rho(t_r)\}$ and $\{\rho(t_s), \cdots, \rho(t_N)\}$ are maximized without excluding the intermediate ones;

2. the intervals of confidence $\bar{h}_\imath$, $\imath \in \{A, B\}$, defined in (8) are minimized.

Once the above algorithm is implemented **?**, it provides the time instants t_o and t_l as well as the mass densities $\rho_\imath$, $\imath \in \{A, B\}$, of the following and leading products for a 100 % purity.

At the end of this step we have $o := r$, $l := s$ so that

$$\rho_A := \bar{\rho}_A, \quad \rho_B := \bar{\rho}_B, \quad h_A := \bar{h}_A \text{ and } h_B := \bar{h}_B. \tag{10}$$

Equation (5) is systematically used to transform the mass density signature $\rho(t_j)$ into volumetric concentration $C_B(t_j)$ for $j = o, o+1, \cdots, l-1, l$ at the receiving point $x = L$.

4. CONCENTRATION PROFILE AND COMMINGLING VOLUME COMPUTATION

To compute the commingling volume it becomes necessary to numerically approximate the integral in Eq. (4). To do so, we have used the trapezoidal rule and have assumed that the time instants t_{BA} and t_{AB} (which define the beginning and the end of the mixture region for a pre-set couple of admissible concentrations C_{BA} and C_{AB}) are such that $t_{p-1} \leq t_{BA} \leq t_p$ and $t_k \leq t_{AB} \leq t_{k+1}$ with $\Delta t \ll t_{BA} - t_{AB}$. As a result, the commingling volume can be written as:

$$V_c = \int_{t_{BA}}^{t_p} Q(t)dt + \int_{t_p}^{t_k} Q(t)dt + \int_{t_k}^{t_{AB}} Q(t)dt =$$

$$\frac{\Delta t}{2}\Big[Q(t_p)\big(1 + 2\alpha_{BA} - \alpha_{BA}^2\big) + \alpha_{BA}^2 Q(t_{p-1}) + Q(t_k)\big(1 + 2\alpha_{AB} - \alpha_{AB}^2\big) +$$

$$\alpha_{AB}^2 Q(t_{k+1})\Big] + \Delta t \sum_{j=p+1}^{k-1} Q(t_j) + O(\Delta t^3), \tag{11}$$

in which α_{BA} and α_{AB} are determined from $C_B(t_j)$ by means of linear interpolations as:

$$\alpha_{BA} = \frac{C_B(t_p) - C_{BA}}{C_B(t_p) - C_B(t_{p-1})} \qquad ; \qquad \alpha_{AB} = \frac{1 - C_{AB} - C_B(t_k)}{C_B(t_{k+1}) - C_B(t_k)}. \tag{12}$$

5. UNCERTAINTY ANALYSIS

To properly report the commingling volume computed via Eq. (11) according to the statistical data reduction procedure presented in Section 3we next report the way uncertainties are evaluated.

Let $U(\rho_A)$, $U(\rho_B)$, $U(\rho)$, and $U(Q)$ represent the total uncertainties associated with the measurements of ρ_A, ρ_B, $\rho(t_j)$ and $Q(t_j)$, for $j = o, o+1, \cdots, N$, respectively. As usual, we consider the total uncertainty as being due to errors of fixed and random natures. Fixed errors are those inherited to the tolerances of the instruments and will be referred to as $U^f(Q)$ and $U^f(\rho)$. Since a same instrument is used to measure ρ, ρ_A, ρ_B we have $U^f(\rho) = U^f(\rho_A) = U^f(\rho_B)$. Any other non-fixed errors are classed as random. We refer to them as $U^r(\rho)$, $U^r(\rho_A)$, $U^r(\rho_B)$ and $U^r(Q)$. Since it seems reasonable to assume that random errors exist for the mass density only, as a result of pressure transients in the line as well as temperature variations, therefore we consider that $U^r(Q) = 0$.

On the other hand, based on the analysis reported on Section 3.we promptly see that

$$U^r(\rho_\imath) = |h_\imath| \qquad \text{for} \qquad \imath \in \{A, B\}. \tag{13}$$

In contrast to what happens to $\rho_\imath$, $\imath \in \{A, B\}$, for which multiple samples are available, only one measure is carried out for the mass density of the mixture $\rho(t_j)$, for $t_j \in [t_o, t_l]$. Thus, we consider $U^r(\rho) = \max\{U^r(\rho_A), U^r(\rho_B)\}$.

Finally, for each measurement the total uncertainty is computed by taking the square-root-sum of the fixed and random errors.

To compute the uncertainty associated to $C_B(t_j)$ we note through Eq. (5) that the correlation between C_B and ρ may be regarded as function of the following independent quantities $(\rho(t), \rho_A, \rho_B)$. Thus, the uncertainty $U(C_B(t_j))$ may be written as combination of the individual terms by a root-sum-square method Moffat (1988):

$$U(C_B(t_j)) = \sqrt{\left(\frac{\partial C_B}{\partial \rho} U(\rho)\right)^2 + \left(\frac{\partial C_B}{\partial \rho_A} U(\rho_A)\right)^2 + \left(\frac{\partial C_B}{\partial \rho_B} U(\rho_B)\right)^2}. \tag{14}$$

It should be pointed out that $U(C_B(t_j))$ depends on the time instant t_j through $\rho(t_j)$. As it will be seen next, we shall make use of $U(C_B(t_j))$ in evaluating the uncertainty of the commingling volume.

The commingling volume uncertainty is also evaluated by using the root-sum-square method described in Moffat (1988). To achieve this goal, we first note that the commingling volume expression given by Eq. (4) can be regarded as a functional relationship of the form $V_c = V_c\big(t_{BA}, t_{AB}, Q(t)\big)$ of independent arguments. By denoting the uncertainties associated to the time instants t_{BA} and t_{AB} as $U(t_{BA)}$ and $U(t_{AB})$, respectively, the uncertainty associated to the commingling volume can be computed as:

$$U(V_c) = \sqrt{\left(\frac{\partial V_c}{\partial t_{BA}} U(t_{BA})\right)^2 + \left(\frac{\partial V_c}{\partial t_{AB}} U(t_{AB})\right)^2 + \left(\frac{\partial V_c}{\partial Q(t)} U(Q)\right)^2}. \tag{15}$$

In the above equation, the partial derivatives represent the sensitivity coefficients of V_c with respect to the independent measurements and are given by,

$$\frac{\partial V_c}{\partial t_{BA}} = -Q(t_{BA}), \qquad \frac{\partial V_c}{\partial t_{AB}} = Q(t_{AB}), \qquad \frac{\partial V_c}{\partial Q(t)} = t_{BA} - t_{AB}. \tag{16}$$

To complete the evaluation of $U(V_c)$ it remains to determine $U(t_{BA})$ and $U(t_{AB})$. Since Eq. (5) associates one value of $C_B(t_j)$ to a single time instant $t_j \in [t_o, t_l]$, the uncertainty $U(t(t_j))$, for $j = BA$ and $j = AB$, can be computed from $C_B(t_j)$ through the following relationships

$$U(t_{BA}) = \left(\frac{\partial C_B}{\partial t} \Big]_{t_{BA}} \right)^{-1} U(C_B(t_{BA})), \quad \text{and} \quad U(t_{AB}) = \left(\frac{\partial C_B}{\partial t} \Big]_{t_{AB}} \right)^{-1} U(C_B(t_{AB})), \tag{17}$$

in which $U(C_B(t_{BA}))$ and $U(C_B(t_{AB}))$ are approximated by:

$$U(C_B(t_{BA})) \quad = U(C_B(t_p)) - \alpha_{BA}\big[U(C_B(t_p)) - U(C_B(t_{p-1}))\big], \tag{18}$$

$$U(C_B(t_{AB})) \quad = U(C_B(t_k)) + \alpha_{AB}\big[U(C_B(t_{k+1})) - U(C_B(t_k))\big]. \tag{19}$$

It should be noticed that we have used Eq. (4) instead of Eq. (11) in computing the uncertainty associated to the commingling volume. In other words, we have disregarded the numerical approximation errors of order $O(\Delta t^3)$ as well as those inherited from the linear interpolations.

6. DATA GENERATION WITH RANDOM NOISE

In the absence of field data to test the procedure to compute the commingling volume and its uncertainty we appeal to theoretical data with superimposed random noise. A straightforward simplification of the problem described in Section 2.comes out when the virtual dispersion coefficient is assumed to be constant. This is equivalent to assume that the batch transfer is carried out in a constant diameter pipe with constant flow rate and that the products being transferred have similar kinematic viscosities as well as similar mass densities. Under these assumptions, it can be shown that the space-time evolution of the concentration can be described by the following equation Taylor (1954); Botros (1984):

$$\frac{\partial C_i}{\partial t^*} + \frac{\partial C_i}{\partial x^*} = K^* \frac{\partial}{\partial x^*} \left(\frac{\partial C_i}{\partial x^*} \right), \quad \text{for} \quad i \in \{A, B\}, \tag{20}$$

in which K^* is the dimensionless virtual dispersion coefficient and x^* and t^* stand for the dimensionless independent variables, which are given by:

$$K^* = \frac{KA}{DQ} \quad ; \quad t^* = \frac{tA}{DQ}; \quad ; \quad x^* = \frac{x}{D}, \tag{21}$$

being $A = \pi D^2/4$ the cross-sectional area of the pipe.

Equation (20) along with the initial conditions given by Eq. (1) or (2) form an initial-value problem which is the basis for a number of models available in the literature used to estimate commingling volumes. The dispersion coefficient K (or its dimensionless counterpart K^*) given by Eq. (21) accounts for the relative motion of the fluid flow with respect to the bulk-average velocity as well as for molecular and eddy diffusion in the axial direction. As reported in some analytical studies (Krantz and Wasan, 1974; Netchval et al., 1972) this coefficient presents a strong dependence not only on the Reynolds number Re but also on the mean velocity and eddy diffusivity profiles in the turbulent core and pipe wall region. Also, a somewhat less important dependence of K^* on the Schmidt number Sc is shown to take place at low turbulent Reynolds numbers. Since Taylor (1954) first conceived a theoretical expression for $K^* = K^*(Re, Sc)$ for laminar and soon after extended it to turbulent flows, several other correlations have been proposed since then. Among them, the correlation proposed by (Krantz and Wasan, 1974) is the most precise since the mean velocity and eddy diffusivity distributions used to compute K^* satisfy the equations of motion and boundary conditions in the wall region, providing a smooth and continuous transition to the universal mean velocity profile valid in the turbulent core. Based on these considerations, the Krantz and Wasan's correlation for K^* will be adopted in this study.

The functional dependence of K^* on Re and Sc , which is disregarded in several other models e.g., Taylor (1954); Levenspiel (1958); Sjenitzer (1958); Aunicky (1970); Ovádi and Török (1977), appears when the kinematic viscosity and the mass density of the mixture is expressed in terms of the concentrations of the products in the mixture, such as (Gambill, 1959):

$$\rho_{50-50} = 0.5\rho_A + 0.5\rho_B, \tag{22}$$

$$\nu_{50-50}^{1/3} = 0.5\nu_A^{1/3} + 0.5\nu_B^{1/3}. \tag{23}$$

To take the assumptions made so far into account, the value of K^* computed according to correlation proposed by Krantz and Wasan is determined by using a Reynolds number calculated with the mixture density and the mixture viscosity on the basis of a 50%-50 % blend of "A"and "B"products through Eq. (22) and (23). One of the effects of considering K^* as being dependent on the concentration is the asymmetric shape of the concentration distribution profile, as reported by Netchval et al. (1972)and pointed out by Austin and Palfrey (1964) in field tests and by Botros (1984) in an analytical study.

As long as K^* is constant the initial-value problem described by Eq. (20) along with Eq. (1) or (2) admits a simple analytical solution given according to:

$$C_B(\zeta) = \frac{1}{2}\left[1 - erf(\zeta)\right], \quad \text{with} \quad \zeta = \frac{x^* - t^*}{2\sqrt{K^* t^*}}, \tag{24}$$

in which $erf(\zeta)$ stands for the error function of the variable ζ . Finally, if the definition given by Eq. (4) is employed at a position $x^* = L/D$ then it is possible to prove that the resulting theoretical or exact commingling volume can be evaluated by the following expression,

$$V_c = \frac{\pi D^2}{2}\sqrt{K^* DL}\left[erf^{-1}(1-2C_{BA}) + erf^{-1}(1-2C_{AB})\right], \tag{25}$$

where $erf^{-1}(\zeta)$ stands for the inverse of the error function of the variable ζ. Equation (25) is a well-known result which shows that the commingling volume is proportional to the square-root of the distance traveled by the conventional half-length of the mixture.

To emulate the data provided by the SCADA during the passage of a batch through the position $x = L$, Eq. (5) is used with Eq. (24) to generate the mass density registers $\rho(t_j)$ for t_j belonging to the database time chronology given by Eq. (6). Aiming at giving a field-data character to the theoretical data provided by $\rho(t_j)$ a random noise is introduced in this signal. Thus, instead of using $\rho(t_j)$ we have considered a noisy data $\widehat{\rho}(t_j)$ obtained as follows:

$$\widehat{\rho}(t_j) = \rho(t_j) + \rho^a \theta_\rho(t_j) sin\left(\theta_\omega(t_j)\omega t_j + \theta_\phi(t_j)\phi\right), \tag{26}$$

in which $\theta_\rho(t_j)$, $\theta_\omega(t_j)$ and $\theta_\phi(t_j)$ are random numbers within the interval $[0, 1]$ with uniform probability density generated at each time t_j. In Eq. (26), $\rho^{(a)}$, ω and ϕ are constants and represent the mass density amplitude, the frequency and phase of the pseudo-noise, respectively.

As it has already been pointed out, the application of the data reduction procedure described in Section 3to compute the commingling volume and its uncertainty may fail when applied to noisy data, especially the computation of the uncertainty. To circumvent such a problem, it is proposed in this paper the use of low-pass digital filters on the raw data given by Eq. (26), before applying the data reduction procedure. The central role behind low-pass digital filters is to smooth noisy data by replacing each data point by some kind of local average of surrounding data points. Since nearby points measure vary around the same underlying value, averaging can reduce the level of noise without biasing the value obtained. Among several existing filters, we have chosen the well-known Savitzky-Golay smoothing filter due to its simplicity and effectiveness (Savitzky and Golay, 1964; Savitzky, 1989). The Savitzky-Golay filter makes use of the convolution technique to successively fit sub-sets of adjacent data points with a low-degree polynomial by the method of linear least squares (Savitzky and Golay, 1964). The main parameters in using the filter are m, n_l and n_r, which represent the degree of the smoothing polynomial and the number of the leftward and rightward data points in the moving window average, respectively.

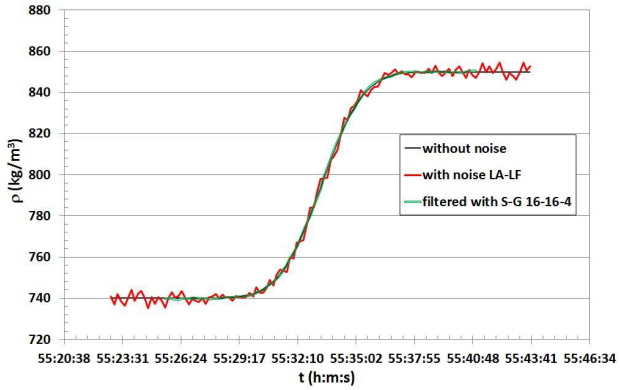
7. RESULTS AND DISCUSSION

In what follows, we analyze the performance of the data reduction procedure to compute the commingling volume and its uncertainty, as outlined in Sections 3.4and 5.for the noisy data generated according to Section 6with and without applying the Savitzky-Golay smoothing filter. As a batch transfer installation we consider a pipeline with $D = 10$ "of diameter and $L = 200$ km long. The transfer is carried out at a constant flow rate of $Q = 182$ m^3/h and the products being transported are gasoline and diesel, being the gasoline the leading fluid ($\gamma_A = 740$ kgf/m^3, $\nu_A = 0.9$ cSt) and the diesel the following fluid ($\gamma_B = 850$ kgf/m^3, $\nu_B = 7.6$ cSt). These data render a Reynolds number $Re = 6.4 \times 10^7$ and a Schmidt number $Sc \approx 1000$, which in turn are used to evaluate the dimensionless dispersion coefficient $K^* = 0.0943$ by using the correlation of Krantz and Wasan (1974). By admitting a pair of admissible concentrations $C_{AB} = C_{BA} = 1.5\%$, the theoretical or exact commingling volume computed at $x = L$ with Eq. (25) is 21.524 m^3.

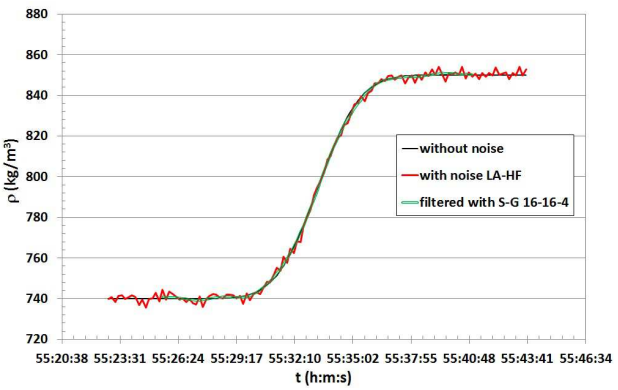
Tabela 1: **Physical parameters of the noise in the mass density signal.**

Case	ρ^a (kg/m^3)	ω (Hz)
LA-LF	5	0.001
LA-HF	5	10
HA-LF	10	0.001
HA-HF	10	10

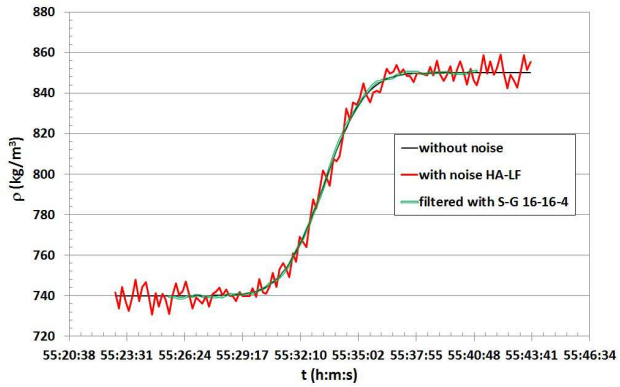
With this flow rate the conventional half-length mixture takes 55:33:20 hours to reach the end of the pipeline, where the acquisition of the flow rate and mass density is carried out with a scan-rate of $\Delta t = 10$ s. This acquisition rate is used in Eqs. (24) and (26) to generate a noisy data base with a total of $N = 125$ registers in time, contemplating the passage of the mixture zone a few moments before and after it has reached the position $x = L$. In order to highlight the effects of the amplitude ρ^a and the frequency ω associated with the noise on the mass density signal (see Eq. (26)) on the computation of the commingling volume and its uncertainty, two values of amplitude and two values of frequency were considered in the analysis. The low and high frequencies, herein designated LF and HF, represent $100f$ and $0.01f$, respectively, being $f := 1/\Delta t = 0.1$ Hz the data acquisition frequency. The low and high amplitudes, herein denoted LA and HA, correspond to 5 and 10 times the uncertainty $U(\rho)$ associated with the in-line interface sensor. Thus, by making all possible combinations of these parameters we get four cases LA-LF, LA-HF, HA-LF and HA-HF, which are summarized in Table 1.



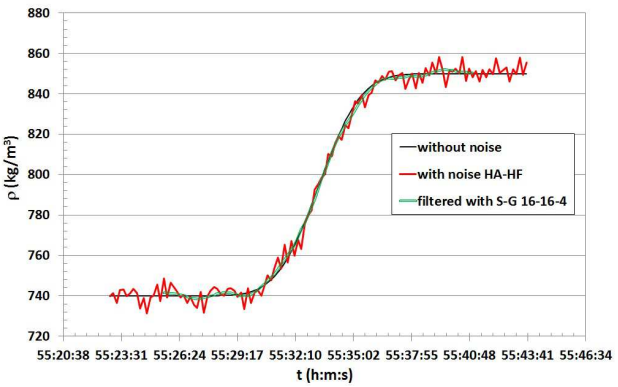
(a) Mass density vs time at $x = 200$km for LA-LF



(b) Mass density vs time at $x = 200$km for LA-HF



(c) Mass density vs time at $x = 200$km for HA-LF



(d) Mass density vs time at $x = 200$km for HA-HF

Figura 2: Mass density vs time at $x = 200$km for the four cases LA-LF, LA-HF, HA-LF and HA-HF. The signals without noise, with noise and filtered with the Savitzky-Golay(S-G) filter are displayed with dark, red and green lines, respectively.

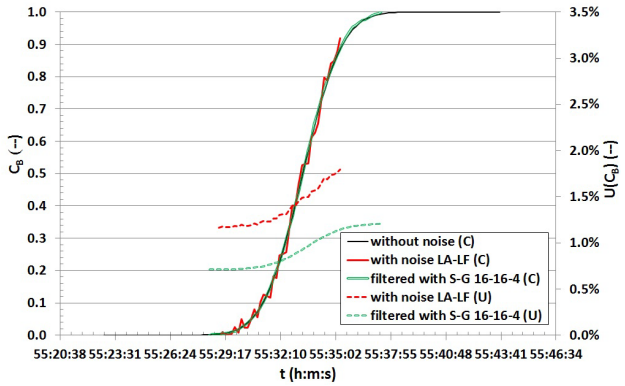
The generated mass density signals with noise versus time (in the format hh:mm:ss (hour:minute:second)) registering the passage of the mixture zone at $x = L$ are displayed in red in Figs. 2(a), 2(b), 2(c) and 2(d) for the cases LA-LF, LA-HF, HA-LF and HA-HF, respectively. To facilitate the establishment of proper comparisons, the signal without noise is shown in dark line and the corresponding filtered signals are displayed in green in these figures. The filtered signals were smoothed with the Savitzky-Golay(S-G) filter by using $n_l = n_r = 16$ and $m = 4$. As it can be seen in these figures, the filter presents a very good performance in almost all cases, being capable to severely eliminate the noise without distorting the original signal. The lower the amplitude is, the better the performance becomes. The worst performance is observed for the case HA-HF.

To evaluate the performance of the data reduction procedure used to compute the commingling volume and its uncertainty the mass density signals with noise and pos-processed with the S-G filter are considered. The obtained results are presented in Table 2 along with the original signal without noise. The second column of this table shows the time instants t_i & t_f (in the format h:m:s) associated with the passage of the beginning and the end of the commingling zone at $x = L$ with admissible concentrations of $C_{AB} = C_{BA} = 1.5\%$. The third column exhibits the computed commingling volume (C.C.V.) and its uncertainty (with a confidence interval of 95%) followed by the lower-upper confidence bounds (L.B.-U.B.) for easier comparison purpose. Finally, the last two columns show the exact commingling volume (E.C.V.) and the relative error, respectively.

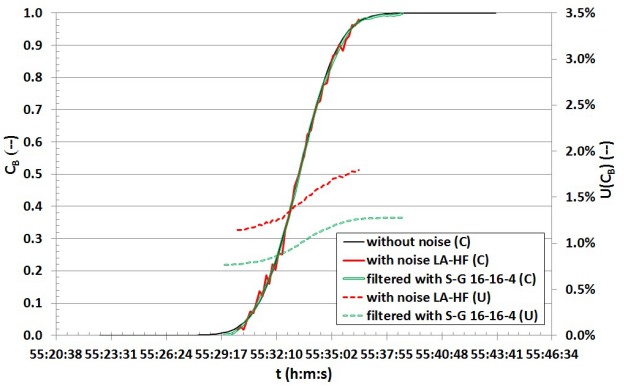
It can be inferred form Table 2 that the presence of noise in the data severely compromise the identification of the time instants t_i & t_f and/or the computation of the C.C.V. an its uncertainty, inasmuch as these parameter are either not identified (NI) or not computed (NC). With respect to the proper identification of these time instants, the discrepancy with relation to those observed for the signal without noise do not exceed three scan-rates, except for the HA-HF case for which it has reached eleven scan-rates or equivalently 110 seconds. A difference of up to three scan-rates in the identification of t_i & t_f (once shifted towards the same direction for both limits) is not so significant in terms of the computed commingling volume as it can be observed in Table 2. However, its impact on the uncertainty can be crucial, since its varies from 0.781 to 2.47 m^3 for the LA-LF and LA-HF filtered cases, respectively. Except in the LA-LF case, it should be noticed that the noise presence hampers the computation of the uncertainty. On the other hand, once these signals are filtered the computation of the the commingling volume and its uncertainty are allowed. Only for the case HA-HF filtered the exact commingling volume does not fall inside the lower-upper confidence bounds, by demonstrating the effectiveness of the S-G filter. The greatest relative errors of the computed commingling volumes are observed for the HA-HF case. In this case in particular, although the filter has allowed the computation of the commingling volume and its uncertainty the obtained result is not good.

Tabela 2: **Commingling volume (C.V.) results obtained by means of the data reduction procedure for the cases without noise, LA-LF, LA-LF filtered, LA-HF, LA-HF filtered, HA-LF, HA-LF filtered, HA-HF and HA-HF filtered. NI is not identified and NC stands for not computed.**

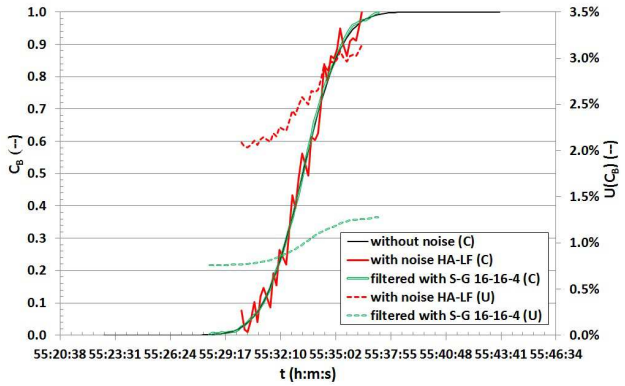
Case	t_i & t_f (h:m:s)	C.C.V. (m^3) (L.B.-U.B.) (m^3)	E.C.V. (m^3)	R.E. (%)
Without Noise	55:29:59 & 55:36:37	21.177±1.826 (19.351-23.003)	21.524	1.61
LA-LF	55:29:42 & 55:36:23	20.351±0.781 (19.570-21.132)	21.524	5.45
LA-LF Filtered	55:29:53 & 55:36:49	21.128±1.258 (19.870-22.386)	21.524	1.84
LA-HF	NI	NC	21.524	NC
LA-HF Filtered	55:30:03 & 55:37:07	21.439±2.470 (18.969-23.909)	21.524	0.39
HA-LF	NI	NC	21.524	NC
HA-LF Filtered	55:29:53 & 55:36:51	21.198±1.053 (20.145-22.251)	21.524	1.51
HA-HF	55:30:37 & 55:36:37	18.732±NC (NC-NC)	21.524	13.0
HA-HF Filtered	55:30:08 & 55:38:41	25.940±1.553 (24.387-27.493)	21.524	20.5



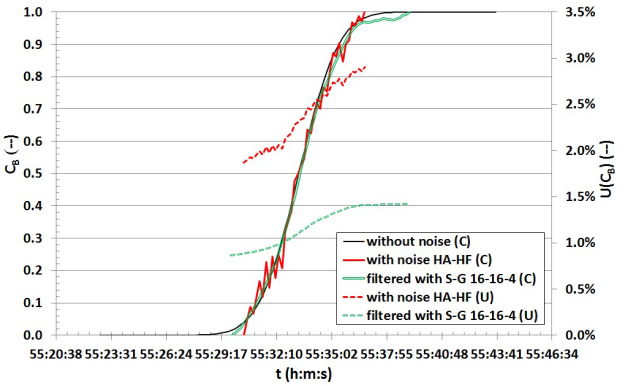
(a) Concentration and its uncertainty vs time at $x = 200$km for LA-LF



(b) Concentration and its uncertainty vs time at $x = 200$km for LA-HF



(c) Concentration and its uncertainty vs time at $x = 200$km for HA-LF



(d) Concentration and its uncertainty vs time at $x = 200$km for HA-HF

Figura 3: Concentration and its uncertainty vs time at $x = 200$km for the four cases LA-LF, LA-HF, HA-LF and HA-HF. The signals without noise, with noise and filtered with the Savitzky-Golay(S-G) filter are displayed with dark, red and green lines, respectively.

To better understand the values of the volume uncertainties presented in Table 2, it is convenient to report to the concentration curve and its uncertainty along the time for the eight cases, as illustrated in Figs. 3(a), 3(b), 3(c) and 3(d). Based on Eq. (15) and especially on Eq. (17), we can see that the commingling volume uncertainty depends mainly on the uncertainty of the concentration curve and the inverse of the time derivative of the concentration curve at both ends of the mixing zone for which $C_{AB} = C_{BA} = 1.5\%$. Since the time derivative of the concentration curve changes locally at $C_{AB} = C_{BA} = 1.5\%$ due to the noise and also due do the filter action, different values can result, by impacting directly on the commingling volume uncertainty. That is, the filtered curve will not necessarily provide best result. However, it can be seen in all curves in Fig. 3 that the uncertainties of the concentration filtered curves are allays lower than those of the noisy ones, what explains the general tendency of having small commingling volumes uncertainties for filtered signals. It also seems apparent in Fig. 3 that the higher the noise amplitude is, the greater the uncertainty concentration becomes. On the contrary, the high frequency noise seems not to interfere significantly on the uncertainty concentration.

8. FINAL REMARKS

It has been proposed in this paper a modified statistical strategy capable to not only identify the ends of the commingling zone and determine its volume for any pair of admissible contamination previously established, but also to automatically compute its uncertainty, even in noisy data. To do so, the Savitzky-Golay smoothing filter has been used. To test and validate the performance and robustness of the proposed methodology, numerical examples are presented for data generated with different levels of noise both in amplitude and frequency. The noise was artificially introduced in the mass density of the products through the concentration profile predicted by a classical well-known theoretical model for constant dispersion coefficient and constant flow rates. The nosy data we have been filtered by employing the Savitzky-Golay smoothing technique. The obtained results for low and high noise amplitudes and frequencies have shown the if the data are smoothed, then the methodology has almost always been able to compute the commingling volume and its uncertainty with great accuracy.

9. ACKNOWLEDGEMENTS

The authors kindly acknowledge the partial financial support provided by the Brazilian Council of Science and Technology (CNPq).

10. REFERENCES

Aunicky, Z., 1970, "The longitudinal mixing of liquids flowing succesively in pipelines," *The Canadian Journal of Chemical Engineering*, Vol. 48, pp. 12–16.

Austin, J. E., and Palfrey, J. R., 1964, "Mixing of miscible but dissimilar liquids in a serial flow in a pipeline," *Proc. Institution of Mechanical Engineers*, Vol. 178, Part1, No.15, pp. 377–395.

Baptista, R. M., Freitas Rachid, F. B., and Carneiro de Araujo, J. H., 2000a, "Estimating mixing volumes between batches in multiproduct pipelines," *Proc. of the ASME International Pipeline Conference 2000*, Vol. 2, pp. 1165–1170.

Baptista, R. M., Freitas Rachid, F. B., and Carneiro de Araujo, J. H., 2000b, "Mixing volume determination in batch transfers through sonic detectors," *Proc. of the 8th Brazilian Congress of Thermal Engineering and Sciences 2000*, in Cd-Rom.

Botros, K. K., 1984, "Estimating contamination between batches in products lines," *Oil and gas Journal*, Vol. 13, pp. 112–114.

Freitas Rachid, F. B., Carneiro Araujo, J. H. and Baptista, R. M., 2002, "Predicting Mixing Volumes in Serial Transport in Pipelines," *ASME Journal of Fluids Engineering*, Vol. 124, pp. 528–534.

Gambill, W. R., 1959, "How to estimate mixtures viscosities," *Chemical Engineering*, Vol. 66, pp. 151–152.

Krantz, W. B., and Wasan, D. T., 1974, "Axial dispersion in turbulent flow of power-law fluids in straight tubes," *Industrial Chemical, Fundamentals*, Vol. 13, No. 1, pp. 56–61.

Levenspiel, O., 1958, "How much mixing occurs between batches?," *Pipe Line Industry*, pp. 51–54.

Melo, S. R., and Freitas Rachid, F. B., 2010, "Computing Transmix in Complex Batch Transfers via the Mixing-Volume-Equivalent-Pipe Concept," *Proc. of the 8th Int. Pipeline Conf.*, Paper IPC2010-31607, Sep. 27th-Oct. 1st, Calgary, C.A..

Moffat, R. J., 1988, "Describing the Uncertainties in Experimental Results", *Experimental Thermal and Fluid Science*, Vol. 1, pp.3–17.

Netchval, M. V., Tugunov, P. I., and Slesáreva, V. G., 1972, "Mixture formation during the process of sequential pumping of petroleum products of varying viscosites," *Neftyanoy Khozyaistvo*, Vol. 50, No. 8, pp. 57–59. In Russian.

Ovádi, Z., and Török, E., 1977, "Examination of pipeline transport from the viewpoint of goods quality," *Köolaj és Földgáz*, Vol. 10, No. 4, pp. 121–123. In Hungarian.

Savitzky, A., and Golay, M.J.E., 1964. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures," *Analytical Chemistry*, Vol. 36, No. 8, pp. 1627–1639.

Savitzky, A.b,1989. "A Historic Collaboration," *Analytical Chemistry*,Vol. 61, No. 15, pp. 921A–923A.

Sjenitzer, F., 1958, "How much do products mix in a pipeline?," *The Pipeline Engineer*, pp. D31–D34.

Smith, S. S., and Schulze, R. K., 1948a, "Interfacial mixing characteristics of products in products pipe line - Part 1," *The Petroleum Enginner*, Vol. 20, pp. 94–104.

Smith, S. S., and Schulze, R. K., 1948b, "Interfacial mixing characteristics of products in products pipe line - Part 2," *The Petroleum Enginner*, Vol. 20, pp. 7–12.

Taylor, G. I., 1954, "The dispersion of matter in turbulent flow through a pipe," *Proceedings of Royal Society*, Series A, Vol. 223, pp. 446–468.

Wylie, E. B., and Streeler, V. L., 1993, "Fluid Transients in Sytems," Prentice Hall, New York.

11. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.