



## Fast and accurate numerical method for predicting gas chromatography retention time



Carlos Alberto Claumann<sup>a</sup>, André Wüst Zibetti<sup>a,\*</sup>, Ariovaldo Bolzan<sup>a</sup>,  
Ricardo A.F. Machado<sup>a</sup>, Leonel Teixeira Pinto<sup>b</sup>

<sup>a</sup> Laboratório de Controle de Processos, Departamento de Engenharia Química e de Alimentos, Centro Tecnológico, Universidade Federal de Santa Catarina (UFSC), P.O. Box: 476, Zipcode: 88010-970, Florianópolis, SC, Brazil

<sup>b</sup> Laboratório de neuroengenharia computacional, Departamento de Engenharia Química e de Alimentos, Centro Tecnológico, Universidade Federal de Santa Catarina (UFSC), P.O. Box: 476, Zipcode: 88010-970, Florianópolis, SC Brazil

### ARTICLE INFO

#### Article history:

Received 11 March 2015

Received in revised form 3 June 2015

Accepted 3 June 2015

Available online 9 June 2015

#### Keywords:

Gas chromatography

Prediction

Retention time

Programmed temperature gas

chromatography (PTGC)

Thermodynamic properties

Numerical method

### ABSTRACT

Predictive modeling for gas chromatography compound retention depends on the retention factor ( $k_i$ ) and on the flow of the mobile phase. Thus, different approaches for determining an analyte  $k_i$  in column chromatography have been developed. The main one is based on the thermodynamic properties of the component and on the characteristics of the stationary phase. These models can be used to estimate the parameters and to optimize the programming of temperatures, in gas chromatography, for the separation of compounds. Different authors have proposed the use of numerical methods for solving these models, but these methods demand greater computational time. Hence, a new method for solving the predictive modeling of analyte retention time is presented. This algorithm is an alternative to traditional methods because it transforms its attainments into root determination problems within defined intervals. The proposed approach allows for  $t_r$  calculation, with accuracy determined by the user of the methods, and significant reductions in computational time; it can also be used to evaluate the performance of other prediction methods.

© 2015 Elsevier B.V. All rights reserved.

### 1. Introduction

The methods for predicting compound retention time ( $t_r$ ) in gas chromatography (GC) have garnered the interest of many a research group. Some software packages have been developed for this end, combining the possibility of identifying compounds with mass spectrometry.

Different approaches are used to calculate the retention time of a compound on a capillary column, given a stationary phase and its main dimensions. The main approach tackles the determination of the retention factor of a compound departing from its thermodynamic properties ( $\Delta H$ ,  $\Delta S$  and  $\Delta C_p$ ). The estimation of these parameters can be obtained under the isothermal conditions of the chromatographic column, and such estimate presents an analytical solution. However the same does not occur if temperature ramps

are applied, which would require solving the problem with the use of numerical methods [1].

According to [1–3], the use of temperature ramps when estimating parameters produces better results in terms of  $t_r$  prediction than in case of parameters obtained in constant conditions. However, the first problem results in a non-linear regression, and an elevated number of evaluations of  $t_r$  can be required until its convergence.

Usage of this model enables other applications, such as determining the programming of the temperature in optimizing the analysis of components of a mixture. In this case, the variables to be optimized would be the slopes and duration of temperature ramps in relation to time. This type of application would also require a large number of model evaluations. Additionally, the objective function to be optimized must be specified. Depending on the problem and the objective function chosen, it can be possible to use the classic method (based on differential calculus) [4–7]; however, these are not robust and demonstrate a high dependence on initial conditions. Thus, it is usually necessary to resolve the model several times to ensure that the optimal, or at least a good, solution has been obtained. Another possibility is to use methods from Computational Intelligence; for example: *genetic algorithms* [8]. Such methods are more robust than those based on calculation (*differential calculus*),

\* Corresponding author. Tel.: +55 48 3721 4073.

E-mail addresses: [carlos.claumann@prosgad.ufsc.br](mailto:carlos.claumann@prosgad.ufsc.br) (C.A. Claumann), [azibetti@gmail.com](mailto:azibetti@gmail.com), [andre.zibetti@ufsc.br](mailto:andre.zibetti@ufsc.br) (A. Wüst Zibetti), [ariovaldo.bolzan@ufsc.br](mailto:ariovaldo.bolzan@ufsc.br) (A. Bolzan), [ricardo.machado@ufsc.br](mailto:ricardo.machado@ufsc.br) (R.A.F. Machado), [leonel.p.teixeira@ufsc.br](mailto:leonel.p.teixeira@ufsc.br) (L.T. Pinto).

but demand longer computational time [9]. This way, it is difficult to define, a priori, what the best optimization methodology for the described problems would be.

The optimization of temperature programming by way of computer simulations has been dealt with by some authors, who proposed algorithms for retention time calculations and temperature programming optimization strategies for the separation of compounds. Different optimization strategies have been proposed [4,6,7,10–15], but it is notably difficult to develop a heuristic method that aims at reducing the computational time of the model solution, keeping the numerical accuracy. Consequently, it is necessary to develop efficient methods for solving the retention time prediction model.

This paper introduces an efficient and accurate way with which to solve the model for the prediction of retention time of an analyte, with low computational time and a reduced number of iterations. The method is relatively simple and can be applied to solve problems such as parameter estimation and the optimization of multiple ramps in gas chromatography.

The proposed methodology combines the use of integration techniques based on Gaussian Quadrature [16], which results in less integrand evaluations. Additionally, the determination of retention time is converted into a root-finding problem in a known interval [17]. This way, retention time can be obtained with an accuracy that is defined by the user of the method, where it would be enough to simply select the stopping criteria of the root determining problem and the number of points of quadrature in the integration. Thus, the proposed method can be used to verify the accuracy obtained via other methods of prediction of GC retention time.

In this work, optimization is not described; however, it is important to stress that the proposed method was developed so as to be used for multi-ramps optimization in PTGC, for in these cases millions of evaluations of the retention time model are needed.

## 2. Theory

A model for predicting the retention time of an analyte requires, basically, information about: (1) the mobile flow phase, which equally influences all compounds; and (2) its retention, which is specific to each component.

Providing there is good approximation, the displacement of the mobile phase consists of the flow of a compressible fluid in a tube. If the diameter of the capillary columns employed in analysis is diminutive, the flow regime will preferably be laminar, which implies a parabolic profile for the mobile phase velocity due to the diameter of the column. Considering the velocity profile known, a rigorous model for hold-up time ( $t_M$ ) can be derived; however, due to uncertainties relating to the size of the columns, it is recommended that estimation be done experimentally  $t_M$ . The hold-up time will depend directly on the mobile phase viscosity, which in turn will depend on the operating temperature [18].

The thermodynamic approach for predicting the retention time of a solute in a temperature-programmed GC (TPGC) is well applied [19–23]. In GC the solvation of an analyte ( $i$ ) in the carrier gas inside a capillary column with a specific stationary phase is expressed by the relationship between  $\Delta G_i(T)$  and the distribution coefficient  $K_i(T)$ . The temperature dependence of  $\Delta G_i$  can be established by the basic thermodynamic relationship in terms of  $\Delta H(T_0)$ ,  $\Delta S(T_0)$  and  $\Delta C_p$ . The  $\Delta H(T_0)$ ,  $\Delta S(T_0)$  and  $\Delta C_p$ , represent the changes in enthalpy, and entropy associated with the transfer of solute from mobile phase to stationary phase at a given temperature  $T$ ,  $T_0$  is a reference temperature, and  $\Delta C_p$  is the change in its isobaric

heat capacity heat capacity for the transfer. The thermodynamic relationship is expressed in  $K_i(T)$  as [19],

$$\Delta H(T(t)) = \Delta H(T_0) + \Delta C_p(T(t) - T_0) \quad (1)$$

$$\Delta S(T(t)) = \Delta S(T_0) + \Delta C_p(\ln T(t) - \ln T_0) \quad (2)$$

$$\ln K_i(T(t)) = -\frac{\Delta H}{RT(t)} + \frac{\Delta S}{R} \quad (3)$$

where  $R$  is the gas constant.

The distribution coefficient  $K_i$  is related to retention factor  $k_i$  according to Eq. (4). In this case, Eq. (3) can be expressed directly in terms of  $k_i$  as shown in Eq. (5).

$$k_i(t) = \frac{K_i(t)}{\beta} \quad (4)$$

where  $\beta$  is the phase ratio of the column.

$$\ln k_i(T(t)) = -\frac{\Delta H}{RT(t)} + \frac{\Delta S}{R} - \ln \beta \quad (5)$$

If hold-up time can be described as a function only of time (but not of the position in the column), it is possible to apply separation of variables in the differential equation [3,24,25], describing the motion of an analyte, resulting in the well-known equation GC (Eq. (6)),

$$1 = \int_0^{t_{r,i}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} \quad (i = 1 \dots \text{NC}) \quad (6)$$

where  $t_{r,i}$  is retention time of the compound  $i$  and NC indicates the number of compounds (analytes).

Analytical solution of Eq. (6) in terms of  $t_{r,i}$  is possible when the temperature is kept constant [2,19,22], otherwise numerical methods need to be applied to obtain its value [1,23,26].

The proposed approach is compared with the classic method of integration used with numerous authors to predict the retention time as described in section 3.

## 3. Classic methodology for determining retention time

A classic procedure used to determine the retention time involves discretizing Eq. (6) in accordance with Eq. (7),

$$1 = \sum_{i=1}^{N_{p,i}} \frac{\Delta t}{t_M(t)[1 + k_i(t)]} \quad (i = 1, \dots \text{NC}) \quad (7)$$

where  $\Delta t$  is the discretization step.

The summation contained on the right side of Eq. (7) is done until it exceeds the unit value. In this case,  $N_{p,i}$  corresponds to the smallest whole value that makes the right side of Eq. (7) surpasses the unit. Once  $N_{p,i}$  is determined, retention time can be easily calculated according to Eq. (8).

$$t_{r,i} = N_{p,i} \cdot \Delta t \quad (8)$$

For a rectangular discretization, as shown in Eq. (7), a small time step must be utilized. Applications described in literature use  $\Delta t$  between 0.05 and 0.1 min [4,11,13,14].

It is worth stressing that the number of integrand evaluations is numerically equal to the value of  $N_{p,i}$ . The question is that  $N_{p,i}$  can turn out to be of a few hundreds. As an example, for a  $t_{r,i} = 10$  min and  $\Delta t = 0.01$  min it will imply in 1000 evaluations of the integrand of Eq. (6).

One possibility for increasing the accuracy in estimating  $t_{r,i}$  with the classic method is to interpolate the values of the summation from the right side of Eq. (7) for  $N_{p,i}$  and  $N_{p-1,i}$ , which was admitted in this work.

The number of evaluations of the integrand may be reduced by using a high-order integrator, such as a Simpson or trapezoid,

**Table 1**  
Definition and discretization of the programmed temperature.

Time ( $t$ )	Temperature ( $T$ )	Ramps ( $j$ )
$t_1 = 0$	$T_1$	1
$t_2$	$T_2$	2
$\vdots$	$\vdots$	$\vdots$
$t_{NR}$	$T_{NR}$	NR
$t_{(NR+1)}$	$T_{(NR+1)}$	

which allows for the utilization of a larger time step; however, a heightened  $\Delta t$  will incur a higher interpolation error.

#### 4. Algorithm

The description of the new method begins with the discretization of the temperature programming on at time intervals, with well-defined linear ramps.

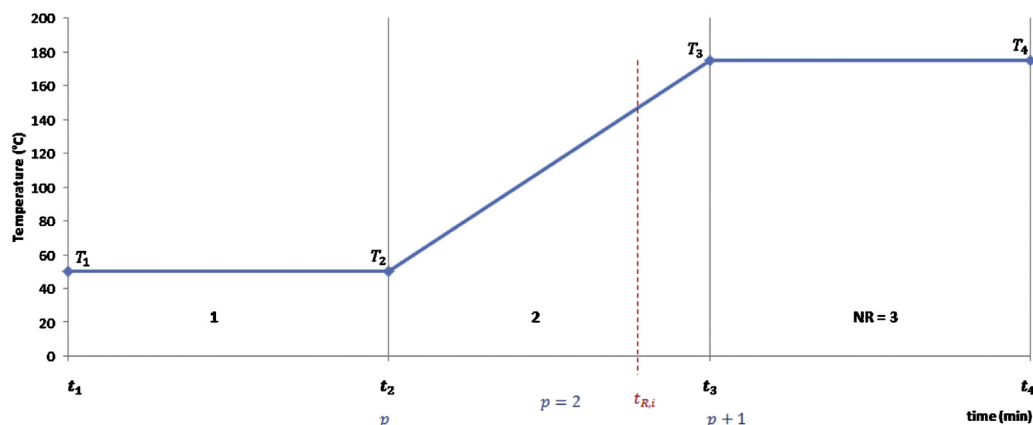
##### 4.1. Programming of multiple temperature ramps

Any given linear ramp temperature or isotherms program ( $j = 1 \dots NR$ ) can be described as a set of  $(NR + 1)$  points in time (independent variable) and temperature (dependent variable), where NR is the number of ramps and isotherms.

The temperature at any time  $t$  can be obtained by interpolating the values from Table 1. Because linear behavior of temperature with time is assumed, only the tabulated values are necessary, prior and post the desired  $t$ .

In temperature programming (see Fig. 1) it is assumed that this always increases with time so as to decrease retention of the compounds. Some GC equipments do not allow decrease in temperature programming.

It is desirable that all analytes be detected at the end of the established analysis time  $t_{(NR+1)}$ . In any case, it is possible that a component remains in the column for a longer time than  $t_{(NR+1)}$ . In this case, to obtain an estimate of retention time when  $t_r$  is longer than  $t_r$  it is assumed that the temperature remains constant and equal to  $T_{(NR+1)}$  for lengthier time.



**Fig. 1.** Hypothetical gas chromatography programmed temperature, given any number of ramps (NR), with discrete temperature  $T_x$ , time  $t_x$  and retention time  $t_{r,i}$  of a compound  $i$ , and  $p$  being its ramp index.

##### 4.2. Proposed methodology for determining retention time

The integral of the right side of Eq. (6) wherein retention time was replaced by an arbitrary time  $t$  (Eq. (9)) to further clarify the proposal,

$$\int_0^t \frac{dt}{t_M(t)[1 + k_i(t)]} \quad (10)$$

Eq. (9) can be calculated differently, dividing the integration domain in intervals corresponding to the times of each ramp (Eq. (10)), see Fig. 1,

$$\int_0^t \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} = \sum_{j=1}^{p-1} \int_{t_j}^{t_{j+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} + \int_{t_p}^t \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} \quad (i = 1 \dots NC) \quad (10)$$

where  $p$  is defined as the highest index between the ( $j = 1, \dots NR$ ) ramps, so that  $t \geq t_p$ , that is,  $t$  satisfied  $t_p \leq t < t_{p+1}$ .

The advantage of the representation described by Eq. (10) is that in each of the intervals the integrand is a smooth function of the integration time. The smooth behavior is related to the continuity of the function and its derivatives. In the case of attempting to run an integration in time directly between 0 and, one would be dealing with a discontinuous function departing from its first-order derivative. Thus, discontinuities were eliminated by appropriately choosing integration limits.

Seeing as one of the objectives of this paper is to reduce computational time, the representation shown in Eq. (10) is important, for it makes possible the use of an integrator of the Gaussian quadrature type which is very precise, even for a small number of quadrature points, but which demands elevated smoothness (existence of high-order derivatives) of the integrand.

At the application of the representation of Eq. (10) for solving Eq. (9) the upper limit of integration  $t = t_r$  is not known a priori. In fact, this is the unknown that is to be determined. Thus, the proposed approach for solving Eq. (6) consisted of two main steps for each analyte of interest: (a) initially one must determine in which interval retention time is located, that is, it is desired to know the value of the index  $p$  such that  $t_p \leq t_r < t_{p+1}$ ; and (b) subsequently, retention time proper is calculated. For this research, said step was turned into a problem of determining the root in the  $[t_p; t_{p+1}]$  interval, as described in section 4.3.

4.3. Determining interval  $[t_p; t_{p+1}]$  where the root (the retention time,  $t_r$ ) is located

The root search interval can be determined by verifying in which ramp the derived restrictions in this section will be satisfied.

Obtaining inequalities begins by evaluating Eq. (10) for  $t = t_{r,i}$ . Such equation is equivalent to the original problem to be solved (Eq. (6)) and thus can be equalized to unit

$$1 = \sum_{j=1}^{p-1} \int_{t_j}^{t_{j+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} + \int_{t_p}^{t_{r,i}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} \quad (i = 1 \dots NC) \tag{11}$$

4.3.1. Obtaining inequality upper bound for retention time,  $t_r$

Adding and subtracting  $\int_{t_{r,i}}^{t_{p+1}} dt/(t_M(t) \cdot [1 + k_i(t)])$  from the right side of Eq. (11), the following is obtained:

$$1 = \sum_{j=1}^{p-1} \int_{t_j}^{t_{j+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} + \int_{t_p}^{t_{r,i}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} + \left( \int_{t_{r,i}}^{t_{p+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} - \int_{t_{r,i}}^{t_{p+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} \right) \quad (i = 1 \dots NC) \tag{12}$$

---


$$\text{Res}_i(\varepsilon) = 1 - \left( \underbrace{\sum_{j=1}^{p-1} \int_{t_j}^{t_{j+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]}}_{I_{p,i}} + \int_{t_p}^{\varepsilon} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} \right) \quad (t_p \leq \varepsilon \leq t_{p+1}) \text{ and } (i = 1 \dots NC) \tag{18}$$

Bringing together the integral between  $t_p$  and  $t_{r,i}$  with integral between  $t_{r,i}$  and  $t_{p+1}$ :

$$1 = \sum_{j=1}^{p-1} \int_{t_j}^{t_{j+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} + \int_{t_p}^{t_{p+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} - \underbrace{\int_{t_{r,i}}^{t_{p+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]}}_{>0} \quad (i = 1 \dots NC) \tag{13}$$

Eliminating the last integral in Eq. (13) this is converted into an inequality:

$$1 \leq \sum_{j=1}^{p-1} \int_{t_j}^{t_{j+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} + \int_{t_p}^{t_{p+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} \quad (i = 1 \dots NC) \tag{14}$$

$$1 \leq \sum_{j=1}^p \int_{t_j}^{t_{j+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} \quad (i = 1 \dots NC) \tag{15}$$

4.3.2. Obtaining inequality lower bound for retention time,  $t_r$

Inequality inferior to unit can be obtained by directly eliminating the last term from the right side of Eq. (11).

$$1 = \sum_{j=1}^{p-1} \int_{t_j}^{t_{j+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} + \underbrace{\int_{t_p}^{t_{r,i}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]}}_{>0} \quad (i = 1, \dots NC) \tag{16}$$

$$1 \geq \sum_{j=1}^{p-1} \int_{t_j}^{t_{j+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} \quad (i = 1, \dots NC) \tag{17}$$

Interval  $p$ , where retention time is located, can be determined by verifying which  $p$  index value that satisfies inequalities, given Eq. (15) and Eq. (17).

If there is not a value  $p$  between  $(1 \dots NR)$  which satisfies inequalities Eq. (15) and Eq. (17), then retention time is located in interval  $[t_{(NR+1)}, \infty)$ .

4.4. Determining the action of retention time in a known interval (root determination problem)

In the case of a problem of root calculation, one must initially define the residue equation and warrant that the root exists. Additionally, it is necessary to stipulate the stopping criteria and choose the numerical method for solving the residue equation.

4.4.1. The residual equation

Residual (or error) equation was defined from Eq. (11) and it corresponds to Eq. (18):

where  $I_{p,i}$  is the accumulated value of the integral obtained in previous intervals for each compound.

It is important to note that  $I_{p,i}$  is a constant value which remains fixed during the root determination, irrespective of the  $\varepsilon$  value indicated by the numerical value in the search interval.

The lower search limit of the integration of the residual equation  $t_p$  is not necessarily the same for all compounds, thus, formally,  $t_p$  is a shortcut for  $t_{p(i)}$ .

For the calculation of the residual function, the defined integral must be reassessed by way of the Gaussian quadrature method for each new ( $\varepsilon$ ) indicated by the root search method.

4.4.2. Existence of one root

The integrand of Eq. (6) is always positive, which implies that the integral on the right side of the Eq. (6) will be monotonically increasing alongside time, thus Eq. (18) will be a monotonically decreasing function in relation to  $\varepsilon$ . Hence, there is solely one root for the type of problem in question in  $[0, \infty)$ . Different root determination methods must be applied if  $t_{r,i} < t_{(NR+1)}$  or  $t_{r,i} \geq t_{(NR+1)}$ .

4.4.3. Numerical solution method

There are numerous methods for determining the root of a function of one variable, such as: bisection, Newton-Raphson, false position, etc. [27]. For this research the Newton-Raphson method was applied, and this is because it presented the fastest convergence among the described.

Another important reason for choosing the Newton-Raphson method was that the calculation of the residual function derived implies in a small increase of computational time. The derived from Eq. (18) is described by Eq. (19):

$$\frac{d\text{Res}_i(\epsilon)}{d\epsilon} = -\frac{1}{t_M(\epsilon) \cdot [1 + k_i(\epsilon)]} \quad (i = 1, \dots, \text{NC}) \quad (19)$$

It can be observed that in Eq. (19) it is only necessary to evaluate the integrand of Eq. (6) for time  $t = \epsilon$ .

Thus, for each iteration of the Newton-Raphson method ( $NQ + 1$ ) evaluations of the integrand from Eq. (6) are done, in which  $NQ$  evaluations result from calculating residue function, plus an additional one, coming from the calculation of its derived.

The initial value  $\epsilon_{0,i}$  applied for the Newton-Raphson method was obtained in all cases by linear interpolation of the calculated residues  $\text{Res}_i(t_p)$  and  $\text{Res}_i(t_{p+1})$ , considering respective times  $t_p$  and  $t_{p+1}$ , that is,  $\epsilon_{0,i}$  is the value of time on interval  $[t_p; t_{p+1}]$ , for which  $\text{Res}_i(t) = 0$ .

Immediately after, the Newton-Raphson method is applied until the established convergence criterion is reached. From the converged solution, it is assumed that the retention time of the component be numerically equal to the obtained  $\epsilon$  value ( $t_{r,i} \leftarrow \epsilon$ ).

The methodology described can be used for calculating retention time, considering any temperature program based on linear heating ramps, as well as any class of compounds, where it suffices to know the parameters of the compounds and of the migration time of the mobile phase.

Definitions of the residue function, stopping criterion, and solution method are necessary only if retention time is inferior to  $t_{(NR+1)}$ . As previously described, this is the commonest case, however, if  $t_{r,i}$  is superior to  $t_{(NR+1)}$  the former can be explicitly obtained, therefore there is no need to apply the numerical method. In such case it is known that  $p = NR + 1$  on Eq. (11) which entails:

$$1 = \sum_{j=1}^{NR} \int_{t_j}^{t_{j+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} + \int_{t_{(NR+1)}}^{t_{r,i}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]} \quad (20)$$

$$1 = \underbrace{\sum_{j=1}^{NR} \int_{t_j}^{t_{j+1}} \frac{dt}{t_M(t) \cdot [1 + k_i(t)]}}_{I_{NR+1,i} = \text{constant}} + \int_{t_{(NR+1)}}^{t_{r,i}} \underbrace{\frac{dt}{t_M(t_{(NR+1)}) \cdot [1 + k_i(t_{(NR+1)})]}}_{\text{constant integrand}} \quad (21)$$

$$1 = \frac{(t_{r,i} - t_{(NR+1)})}{t_M(t_{(NR+1)}) \cdot [1 + k_i(t_{(NR+1)})]} + I_{NR+1,i} \quad (22)$$

On Eq. (22), except for  $t_{r,i}$ , the numerical value of all other quantities involved is known, so  $t_{r,i}$  can be obtained explicitly from this equation.

#### 4.4.4. Stopping criterion

According to the established aim, two stopping criteria have been used for the Newton-Raphson method, both based on  $\epsilon$  variations between two consecutive iterations; and so the error criteria on the variable were:

a)  $|\Delta \epsilon| \leq \left(\frac{\text{EP}}{10}\right) \text{ min}$ : The precision of chromatography equipment's (EP)<sup>1</sup> is  $10^{-3}$  min. As a result, an adequate and

practical value for the stopping criteria is to consider it equal to the 10% error margin from the equipment, that is,  $10^{-4}$  min.

b)  $|\Delta \epsilon| \leq 10^{-12}$  min: This criterion was utilized to calculate the value of  $t_r$  with arbitrary accuracy. Seeing as the analytic solution of Eq. (6) is not available for temperature ramps, the most accurate possible solution was defined in each problem as that calculated with the larger number of available quadrature points available (8 points) and  $|\Delta \epsilon| \leq 10^{-12}$  min. This criterion was utilized for the obtainment of the exact solution.

Calculation of  $t_r$  taking criterion (a) into account suffices for practical applications, that is, the error made on resolving Eq. (18) can be deemed irrelevant in relation to the experimental error. Criterion (b) was utilized for determining retention times that serve as reference—that is, to evaluate the accuracy of solutions obtained from criterion (a).

## 5. Numerical results and discussion

The works of Dose (1987) [20] and Gonzalez (1997) [3] were utilized to test the new approach proposed in this paper. With the former, analytes correspond to six types of nitrobenzenes, employed in the production of pesticides, and with the latter, analytes are hydrocarbons.

With the aim of comparing the performance of the methods used (classic and the one proposed in this research), two error levels for predicting retention time were defined: (1) good quality: the error between the “exact” solution and the calculated one is smaller than  $1 \text{ s} = 0.017 \text{ min}$ ; (2) excellent quality: said error is smaller than  $0.1 \text{ s} = 0.0017 \text{ min}$ . The first case suffices for the majority of applications, and the second offers high levels of accuracy.

### 5.1. Identifying retention time of nitrobenzene

In the case of Dose's work (1987), a comparison was made between the retention time of nitrobenzenes for temperature programming, called  $D$ : [PT:  $3^\circ\text{C}/\text{min}$  from  $140^\circ\text{C}$  to  $220^\circ\text{C}$ ]  $\rightarrow$  [I: hold], where I—Isothermal step; PT—linear programmed temperature step.

In relation to hold-up time, Dose's work suggests usage of liquid marker (methylene chloride) with a lower boiling point [20] in order to determine  $t_M$ . The hold-up time ( $t_M$ ) was obtained from the linear temperature dependence of methylene chloride. Under isotherms, this compound is essentially unretained above  $150^\circ\text{C}$  and the retention time follows the increase in the viscosity with good agreement (from  $150^\circ\text{C}$  until  $220^\circ\text{C}$ ). This could be explained by the increasing difference of the boiling point ( $\sim 39.6^\circ\text{C}$ ) with the column temperature, which leads to an exponential increase in vapor pressure, reducing interaction time (retention) with stationary phase.

The values of the experimental points contained in Dose's work were digitalized. From the data obtained with high temperature, a linear regression was performed, and coefficients can be seen on Eq. (23):

$$t_M[\text{min}] = 8.3267 \times 10^{-4} \cdot T[^\circ\text{C}] + 0.583 \quad (23)$$

The parameters used in the simulations are shown in Table 2:

For this comparison the correction of enthalpy and entropy in relation to temperature was not considered, which results in  $\Delta C_p = 0$ . Applying the methodology described in section 3 and 4, retention time for nitrobenzenes was determined.

Tables 3 and 4 show a comparison between the performances achieved by the classic method and the one proposed on this paper, in terms of error and the number of evaluations of the integrand from Eq. (6), respectively. The sampling times ( $\Delta t$ ) when using the

<sup>1</sup> Precision of retention time given by chromatography software.

**Table 2**  
Thermodynamic parameters of each compound used for simulation, from ref. [18].

Compound	$-\Delta H$ (kcal/mol)	$-\Delta S$ (cal/mol K)
Nitrobenzene	9.553	12.13
1,3-Dinitrobenzene	12.801	15.71
2,6-Dinitrotoluene	12.801	15.33
2,4-Dinitrotoluene	13.249	15.76
1,3,5-Trinitrobenzene	15.216	18.35
2,4,6-Trinitrotoluene	15.243	18.12

classic method were de 0.2 s and 2 s and, on the latter, the number of quadrature points varied between 2 and 5.

As described, in order to evaluate the prediction errors of retention times, what was considered as the “exact” solution was that which was obtained by using the method proposed in this work, according to the rigorous criterion described in section 4. The “exact” values for each analyte can be observed on the respective column in Table 3. The columns titled  $NQ=2-5$  correspond to the number of points of quadrature used in each simulation of the proposed method. The errors in relation to the exact solution of each analyte are shown in these columns. In the simulations, the stopping criterion (a) defined in section 4.4.3 was considered.

When the classic method is used, as it is possible to observe in Table 3, a reduction from 2 s to 0.2 s in  $t$  implies approximately in the reduction of one order of magnitude of the solution error.

In Table 4 it can be observed that the reduction mentioned causes an increase in computational time, due to an increase in the number of evaluations of the integrand (approximately a magnitude order). It can be thus concluded that the decrease in prediction error of  $t_r$  is proportional to sampling time.

The classic method presents a direct correlation between retention time and the number of necessary evaluations of the integrand. This can be easily observed in Eq. (8), given fixed  $t_{r,i}$  and  $t$ . Thus, by comparing Tables 3 and 4 it can be verified that the number of evaluations increases with the respective retention time value.

When applying the method proposed in this work, low levels of error were obtained in all cases (excellent performance) with only two quadrature points, which implies in a small number of

evaluations of the integrand from Eq. (6). Once the interval where the root is located was determined, it was noticed that the number of iterations needed for the convergence of the Newton method was between 4 and 7. Thus, the increase in retention time does not imply a significant increase in the number of evaluations. The only change is the number of evaluations associated with determination of the interval where the root is located.

Upon comparing the performances of the methods, the one proposed in this work reaches the same level of error, but with a far smaller number of evaluations of the integrand from Eq. (6). The observed difference between the numbers of evaluations is of one or two orders of magnitude, according to the desired accuracy.

## 5.2. Identifying retention time of hydrocarbons

In the work of Gonzales, a comparison was made between retention times for “Program 5”, which corresponds to temperature programming: [I: 2.5 min] → [PT: 5 °C/min until 80 °C] → [I: 0.5 min] → [PT: 10 °C/min until 250 °C], where I—Isothermal step, PT—programmed temperature linear step.

Program 5 [25] corresponds to a case in which inlet pressure is kept constant for the duration of the tests. Inlet and outlet pressures remained, respectively, in 1546 and 770 Torr.

The hold-up time was described as a function of temperature, according to Eq. (24). The exponent of that is directly related to the viscosity of the carrier gas used ( $N_2$ ).

$$t_M (\text{min}) = t_{m,0} \left( \frac{T(\text{K})}{T_0} \right)^{0.725} \quad (24)$$

where:  $t_{m,0}$  = retention time for the initial temperature in the program (50 °C) = 1.489 min;

$T_0$ : initial temperature in the program 323.15 [K].

The parameters used are shown in Table 5:

According to reference [25], the ratio  $\alpha/\beta$  in Table 5 relates to the retention of the analyte, according to Eq. (25).

$$k = \left( \frac{\alpha}{\beta} \right) \exp \left( \frac{-\Delta H_c}{RT} \right) \quad (25)$$

**Table 3**  
Comparison of errors in calculated retention times between the proposed method and the classic one.

Compound	Error (min) <sup>a</sup>	Proposed method					
		Proposed method				Classic method	
		Exact solution <sup>b</sup>	NQ. 2	NQ. 3	NQ. 4	NQ. 5	$\Delta t=2$ s
Nitrobenzene	1.946512	-2.988E-06	-2.999E-06	-2.999E-06	-2.999E-06	-1.533E-03	-1.538E-04
1,3-Dinitrobenzene	8.110278	1.821E-04	2.136E-08	-1.572E-09	-4.613E-10	-8.895E-03	-8.900E-04
2,6-Dinitrotoluene	9.154818	2.627E-04	2.455E-07	-5.636E-09	-2.656E-09	-9.655E-03	-9.651E-04
2,4-Dinitrotoluene	11.017025	4.520E-04	1.937E-06	-7.937E-08	-6.350E-08	-1.101E-02	-1.101E-03
1,3,5-Trinitrobenzene	17.813573	-1.263E-03	9.430E-05	-1.414E-05	-1.344E-05	-1.436E-02	-1.436E-03
2,4,6-Trinitrotoluene	18.959965	-2.299E-03	1.426E-04	-6.962E-06	-6.043E-06	-1.462E-02	-1.462E-03

<sup>a</sup> Error = exact solution—calculated retention time from methods.

<sup>b</sup> Exact solution—described in section (4.4.4-b).

**Table 4**  
Comparison of the number of evaluations of the integrand, between the proposed method and the classic method, for solutes from [18].

Compound	Number of evaluations of the integrand					
	Proposed method			Classic method		
	NQ. 2	NQ. 3	NQ. 4	NQ. 5	$\Delta t=2$ s	$\Delta t=0.2$ s
Nitrobenzene	11	15	19	23	59	584
1,3-Dinitrobenzene	17	23	29	35	244	2434
2,6-Dinitrotoluene	17	23	29	35	275	2747
2,4-Dinitrotoluene	17	23	29	35	331	3306
1,3,5-Trinitrobenzene	17	23	29	35	535	5345
2,4,6-Trinitrotoluene	17	23	29	35	570	5689
Mean	16	22	27	33	336	3351

**Table 5**  
Thermodynamic parameter of each solute used in simulation, from Ref [3].

Compound	Parameters from Gonzalez [3]		Model parameters	
	$-\Delta H/R$ (K)	$(\alpha/\beta) \times 10^6$	$-\Delta H$ (kJ/mol)	$-\Delta S$ (J/mol K)
n-Octane	4175	5.533	34.713	54.762
p-Xilene	4262	6.924	35.436	52.897
1,3,5-Trimethylbenzene	4661	4.417	38.754	56.635
1-Undecane	5313	1.729	44.175	64.433
Naphtalene	5107	4.623	42.462	56.256
n-Dodecane	5559	1.695	46.220	64.598
n-Tetradecane	5990	1.545	49.804	65.369
n-Hexadecane	6760	0.635	56.206	72.761

**Table 6**  
Calculated retention times by different methods.

Compound	Retention time [min]	Error (min) <sup>a</sup>						
		Exact solution <sup>b</sup>	Proposed method				Classic method	
			NQ_2	NQ_3	NQ_4	NQ_5	$\Delta t=2$ s	$\Delta t=0.2$ s
n-Octane	4.575023	-2.734E-06	-2.768E-06	-2.769E-06	-2.769E-06	-3.518E-03	-3.525E-04	
p-Xilene	6.013931	1.573E-05	-1.083E-05	-1.075E-05	-1.075E-05	-6.069E-03	-6.075E-04	
1,3,5-Trimethylbenzene	8.840355	7.013E-04	-5.088E-06	3.153E-08	-1.534E-10	-1.007E-02	-1.007E-03	
1-Undecane	12.536080	6.897E-04	-7.161E-05	-6.601E-05	-6.590E-05	-1.474E-02	-1.475E-03	
Naphtalene	13.944925	1.940E-03	-3.212E-05	4.948E-07	-3.236E-08	-1.526E-02	-1.527E-03	
n-Dodecane	14.690550	4.616E-03	-1.166E-04	2.246E-06	-4.468E-07	-1.577E-02	-1.577E-03	
n-Tetradecane	17.860037	2.939E-02	-1.212E-03	-2.247E-05	-6.095E-05	-1.640E-02	-1.640E-03	
n-Hexadecane	20.7574795	7.998E-02	-3.045E-03	-1.712E-04	2.227E-05	-1.660E-02	-1.640E-03	

<sup>a</sup> Error = exact solution—calculated retention time from methods.

<sup>b</sup> Exact solution—described in section (4.4.4-b).

**Table 7**  
Comparison of the number of evaluations of the integrand, between proposed method and the classic one, for solutes from [3].

Compound	Number of evaluations of the integrand					
	Proposed method			Classic method		
	NQ_2	NQ_3	NQ_4	NQ_5	$\Delta t=2$ s	$\Delta t=0.2$ s
n-Octane	13	18	23	28	138	1373
p-Xilene	13	18	23	28	181	1805
1,3,5-Trimethylbenzene	9	13	17	21	266	2653
1-Undecane	20	28	36	44	377	3762
Naphtalene	23	32	41	50	419	4184
n-Dodecane	23	32	41	50	442	4408
n-Tetradecane	26	36	41	50	537	5359
n-Hexadecane	26	36	41	56	624	6228
Mean	19	27	33	41	373	3722

Once rati $\alpha/\beta$  is known, the  $\Delta S_c$  of the retention model can be defined.

Similarly to nitrobenzene, the methodologies described in Sections 3 and 4 were applied to determine the retention time of the hydrocarbons.

Performance comparison between the traditional method and the approach proposed are shown in Tables 6 and 7. In the first case, sampling times ( $\Delta t$ ) of 0.2 and 2 s, and, in the latter, the number of quadrature points was varied between 2 and 5.

When using the method proposed in this work, low error levels were found in practically all cases (excellent performance) with only three quadrature points, which entails a small number of integrand evaluations from Eq. (6).

In the same way as it was noted with for nitrobenzenes, the same performance differences were found between the classic method and the one proposed in this paper, when referring to hydrocarbons.

Upon comparing the performance between methods, in Tables 6 and 7, it can be noted that the one proposed in this work reaches the same levels of error, with a far lower need for evaluations of the integrand from Eq. (6). The difference observed between

the numbers of evaluations is of one or two orders of magnitude, according to the desired accuracy.

Considering that for optimization problems in PTGC the number of evaluations of the fundamental equation could easily be in the house of millions, these results are important. For this type of optimization problem, the use of robust methods [28] (like genetic algorithms) is usually recommended for a good solution, which can be highly demanding, computationally. Thus, in such scenarios, a faster way of calculating retention time, without loss in accuracy, becomes very important.

## 6. Conclusions

This paper presents an efficient, robust and precise manner with which to solve the model for predicting the retention time of an analyte. The method combines the use of integration techniques based on Gaussian quadrature, which results in low computational time. Additionally, determination of the retention time is converted into a root search problem in a known range. The value of the retention time can be obtained with the desired accuracy, simply by setting the stopping criteria for determining the root problem and

the number of points in the quadrature integration. Thus, the proposed method can be used to verify the accuracy obtained with other prediction methods of GC retention time. Considering the prediction of retention time of a series of nitrobenzenes and hydrocarbons, the method proposed in this paper had a much higher performance, in terms of number of evaluations of integrand, than the traditional method, while achieving the same level of accuracy; in addition, differences of one and two orders of magnitude in time resolution were observed. The proposed method will be used in the future to solve problems that require a large number of evaluations of the prediction model GC retention time: the estimate of the thermodynamic parameters for that model and the optimization of temperature ramps programming.

### Conflict of interest

The authors confirm that there are no conflicts of interest regarding this paper.

### Acknowledgments

The authors thank CAPES, CNPq for their financial support. A.W. Zibetti thanks to CAPES-PNPD (2014/2015) for their assistantship.

### References

- [1] G. Castello, P. Moretti, S. Vezzani, Retention models for programmed gas chromatography, *J. Chromatogr. A* 1216 (2009) 1607–1623.
- [2] Z. Wu, Retention simulation in gas chromatography, *J. Chromatogr. A* 840 (1999) 137–143.
- [3] F.R. Gonzalez, A.M. Nardillo, Retention in multistep programmed-temperature gas chromatography and flow control Linear head pressure programs, *J. Chromatogr. A* 757 (1997) 109–118.
- [4] J. Sanz, I. Martínez-Castro, G. Reglero, M.D. Cabezudo, Prediction of the separation in gas chromatography, *Anal. Chim. Acta* 194 (1987) 91–98.
- [5] L.R. Snyder, D.E. Bautz, J.W. Dolan, Computer simulation as an aid in method development for gas chromatography, *J. Chromatogr.* 541 (1991) 35–58.
- [6] D.E. Bautz, J.W. Dolan, L.R. Snyder, Computer simulation as an aid in method development for gas chromatography, *J. Chromatogr.* 541 (1991) 1–19.
- [7] E.V. Dose, Off-line optimization of gas chromatographic temperature programs, *Anal. Chem.* 59 (1987) 2420–2423.
- [8] D.E. Goldberg, *Genetic Algorithms*, Pearson Education, 2006.
- [9] M.S. Bazaraa, H.D. Sherali, C.M. Shetty, *Nonlinear Programming: Theory and Algorithms*, John Wiley & Sons, 2013.
- [10] V. Bártů, S. Wicar, G.-J. Scherpenzeel, P.A. Leclercq, Optimization of temperature programming in gas chromatography with respect to separation time, *J. Chromatogr.* 370 (1986) 235–244.
- [11] V. Bártů, S. Wičar, G.-J. Scherpenzeel, P.A. Leclercq, Optimization of temperature programming in gas chromatography with respect to separation time, *J. Chromatogr.* 370 (1986) 219–234.
- [12] V. Bártů, Calculation of the retention time and peak width for the purpose of optimized gas chromatographic analysis, *J. Chromatogr.* 260 (1983) 255–264.
- [13] V. Bártů, S. Wičar, Optimization of temperature programs in gas chromatography, *Anal. Chim. Acta* 150 (1983) 245–252.
- [14] D.E. Bautz, J.W. Dolan, W.D. Raddatz, L.R. Snyder, Computer simulation (based on a linear-elution-strength approximation) as an aid for optimizing separations by programmed-temperature gas chromatography, *Anal. Chem.* 62 (1990) 1560–1567.
- [15] J.W. Dolan, L.R. Snyder, D.E. Bautz, Computer simulation as an aid in method development for gas chromatography, *J. Chromatogr.* 541 (1991) 21–35.
- [16] S. Chapra, R. Canale, *Numerical Methods for Engineers*, sixth ed., McGraw-Hill Education, Boston, 2010.
- [17] J. Epperson, *An Introduction to Numerical Methods and Analysis*, second ed., Wiley, New Jersey, 2013.
- [18] R.B. Bird, W. Stewart, E. Lightfoot, *Transport Phenomena*, second ed., John Wiley & Sons, New York, 2007.
- [19] F.R. Gonzalez, A.M. Nardillo, Retention index in temperature-programmed gas chromatography, *J. Chromatogr. A* 842 (1999) 29–49.
- [20] E.V. Dose, Simulation of gas chromatographic retention and peak width using thermodynamic retention indexes, *Anal. Chem.* 59 (1987) 2414–2419.
- [21] T.M. McGinitie, H. Ebrahimi-Najafabadi, J.J. Harynuk, A standardized method for the calibration of thermodynamic data for the prediction of gas chromatographic retention times, *J. Chromatogr. A* 1330 (2014) 69–73.
- [22] Y. Thewalim, F. Aldaeus, A. Colmsjö, Retention time prediction of compounds in Grob standard mixture for apolar capillary columns in temperature-programmed gas chromatography, *Anal. Bioanal. Chem.* 393 (2009) 327–334.
- [23] Y. Thewalim, I. Sadiqsis, A. Colmsjö, Comparing columns for gas chromatography with the two-parameter model for retention prediction, *J. Chromatogr. A* 1218 (2011) 5305–5310.
- [24] D.M. Hernandez-Baez, A. Reid, A. Chapoy, B. Tohidi, R. Bounaceur, Establishing the maximum carbon number for reliable quantitative gas chromatographic analysis of heavy ends hydrocarbons. Part 2. Migration and separation gas chromatography modeling, *Energy Fuels* 27 (2013) 2336–2350.
- [25] F. Gonzalez, A. Nardillo, Integration of the equation of peak motion in programmed-pressure and -temperature gas chromatography, *J. Chromatogr. A* 766 (1997) 147–152.
- [26] F. Aldaeus, Y. Thewalim, A. Colmsjö, Prediction of retention times and peak widths in temperature-programmed gas chromatography using the finite element method, *J. Chromatogr. A* 1216 (2009) 134–139.
- [27] S.C. Chapra, *Applied Numerical Methods with MATLAB for Engineers and Scientists*, 3rd ed., McGraw-Hill, 2012.
- [28] P. Nikitas, a. Pappa-Louisi, P. Agrafiotou, Multilinear gradient elution optimization in reversed-phase liquid chromatography using genetic algorithms, *J. Chromatogr. A* 1120 (2006) 299–307.