Optimization of Separation Efficiency in Temperature Programmed Gas Chromatography: Application of Response Surface Methodology

Habtewold Deti Waktola

Acknowledgments

My deepest gratitude goes to my advisors Professor Svein A. Mjøs and Professor Bjørn Grung for their invaluable advice and continuous support from the beginning of this work till the end by guiding, commenting and correcting the thesis. It is my pleasure to mention also the tireless effort made by professor Svein A. Mjøs in teaching me MATLAB computer program by preparing different tutorials and giving me clarifications whenever I needed.

I am thankful also to the University of Bergen department of chemistry for accepting me as a student to carry out my research project using its laboratory facilities.

My gratitude goes to the University of Cadiz which first accepted me as a student and where I took the whole course work for the master program. I enjoyed all the laboratory facilities and the support I received from every professor I encountered there. I am indebted to Professor Miguel Palma the program director of EMQAL who has first welcomed me to Spain and whom kind support and help was always there throughout my stay in the University of Cadiz.

I am very grateful to the founders and coordinators of Erasmus Mundus Quality in Analytical Laboratories (EMQAL) program. It is such a wonderful program which brought together students from different corners of the world and made us share our academic and social experiences in addition to the program trainings. I am indebted also to Professor Isabel Cavaco the program coordinator who has welcomed us to the program and taught us also a couple of courses. It is my pleasure also to thank all professors involved in this program in teaching courses as well as guiding us in laboratories.

Last but not least I am very grateful to my parents, sisters, brothers and friends who have contributed to my academic success in one way or the other.

Index

Abstract

Chromatographic efficiency can be defined as the maximal number of chromatographic peaks that can be separated within a section of a chromatogram. In isothermal gas chromatography (GC) it is usually measured by number of theoretical plates and modeled by the van Deemter equation. Number of theoretical plates is not a valid measure of efficiency in temperature programmed GC due to its variability with temperature. The van Deemter equation is also not useful since it does not account for the temperature factor. Thus the purpose of this work has been to investigate how the efficiency depends on carrier gas velocity and temperature rate in temperature programmed GC by expanding the van Deemter equation to a function that accounts for temperature rate in addition to carrier gas velocity, and to develop methodology for optimizing the separation efficiency.

Fatty acid methyl esters of saturated homologous and unsaturated fatty acids were used as analytes. The study involved nine different columns of varying polarity using helium, hydrogen and nitrogen as a carrier gases.

Efficiency was defined as the inverse of peak width in retention index units and thus peak width was used to measure efficiency where its minimum corresponds to maximum efficiency. Peak width was then used to model van Deemter equation where carrier gas velocity and temperature rate were varied. The data obtained from measurements at different levels of combinations of the varied factors resulted in good fit of van Deemter models at all levels of temperature rate and for all carrier gases used (overall mean $R^2 = 0.9885$). Then by adding the temperature rate factor the ordinary van Deemter curve was expanded to a response surface function which explains peak width in retention index units as a function of carrier gas velocity and temperature rate. The response surface function contained main, interaction and quadratic terms. But since the interacting factors may vary and there is no theoretical basis that explains which of the terms in the response surface function are significant, a backward elimination procedure was employed to detect significant terms. The result revealed that five terms are always required to define the models obtained using each column and every carrier gas. The terms that appear significant depend on the type of carrier gas used.

Experiments were carried out in general at 9 levels of carrier gas velocities and 3 levels of temperature rates. Experimental designs of full 9 x 3, 5 x 3, 3 x 3, skewed 3 x 3 and Doehlert designs of carrier gas velocity and temperature rate were compared. From comparison of

predicted optimum velocity and peak width measurements the 5 x 3 design showed similar result as the full 9 x 3 design. The 3 x 3, skewed 3 x 3 and Doehlert design had lower performance than the 5 x 3 design. Nevertheless they may be good enough for practical purposes. The maximum difference in predicted optimum velocity from results of full design was 0.98 cm/s for velocities measured in the range of 10-15, 20-30 and 30-40 cm/s for nitrogen, helium and hydrogen respectively whereas the difference in peak width measurements were less than 0.0005 equivalent chain length units.

In chromatography there is normally a trade-off between analysis time and separation efficiency. The time of analysis which was best represented with the retention time of the last eluting compound is related to both carrier gas velocity and temperature rate by power functions. This was taken as a basis for deriving response surface function of the retention time of the last eluting compound, which was combined with the models of efficiency. An optimum line of best time efficiency trade-off was then found by an iterative procedure from the overlapped plots and was found to lay 2-5%, 3-6% and 3-8% above the velocity that gives the minimum in van Deemter curve in optimum velocity for helium, hydrogen and nitrogen respectively. Comparison of the carrier gases indicated that at similar efficiency with helium hydrogen as carrier gas reduces the time of analysis by 24-31% whereas there is 80-92% increase in retention time with nitrogen as a carrier gas. On the other hand at similar retention time with helium hydrogen gives 6-8% more efficiency than helium while 12-18% decreases in efficiency with nitrogen is observed.

The performance of different polar columns with the right selectivity for FAME was evaluated and compared. The efficiency of the columns was found to depend on polarity of the columns. The less polar DB-23 and BP-20 columns are the most efficient columns whereas the more polar BPX70 and IL100 columns have moderate and less efficiency than the others.

Lists of abbreviations

CCD	Central Composite Design
DoE	Design of Experiment
ECL	Equivalent Chain Length
FAMEs	Fatty Acid Methyl Esters
FID	Flame Ionization Detector
GC	Gas Chromatography
HETP (H)	Height Equivalent to Theoretical Plate
PPC	Peaks per Carbon
RMSE	Root Mean Square Error
RMSEP	Root Mean Square Error of Prediction
SN	Separation Number
VD	Van Deemter equation / model
VD+Int	Expanded van Deemter model (with interaction)

1. Introduction

1.1. Fatty acids

Fatty acids constitute an important component of food that is used to generate energy as well as form biological membranes and thereby influencing membrane properties like integrity, fluidity, permeability and activities of membrane bound enzymes. The health of human beings is affected by the type and amount of fatty acids present in the regular diet. The two classes of essential polyunsaturated fatty acids: omega-3 and omega-6 are associated with health and disease conditions since they play important physiologic roles in the body. They are essential because our body cannot synthesise them and solely depends on their supply by diet. It has been reported that omega-3 fatty acids are important in preventing many diseases like coronary heart disease, hypertension, type 2 diabetes, rheumatoid arthritis, chronic obstructive pulmonary disease and have many other effects [1, 2]. They are known to affect also favourably inflammatory diseases and behavioural disorders [3]. A good balance between omega-3 and omega-6 fatty acids is important for good health since the two groups of fatty acids have opposing metabolic effects. High intake of omega-6 fatty acids is associated with increase in blood viscosity, vasospasm and decrease in bleeding time [1].

Fatty acids of plant, animal and microbial origin are characterized in general with a carbon chain (usually unbranched and with an even number of carbons), a carboxylic functional group at one end and typically zero to six double bonds. Common fatty acids in animal tissue typically vary in chain length from 12 to 24, but occasionally vary from 2 to 36 or even more. Double bonds usually have *cis* geometry and are separated by a single methylene group if there are more than one. Fatty acids from animal tissue may have one to six double bonds whereas those from higher plants and microbes rarely have more than three and one, respectively. Saturated fatty acids from animal and plant tissues are found in nature in esterified form. They are named systematically as saturated hydrocarbons with the same number of carbons, but with the final –e replaced with –oic. In shorthand they can be notated as A:B, with A representing the number of carbon atoms in the fatty acid chain and B representing the number of double bonds [4]. The double bond positions in unsaturated fatty acids can be specified by the notation A:B n-C, where C is the position of the first double bond counted from the methyl end of the fatty acid chain [5], A and B has the same meaning as above and it is assumed that double bonds in polyunsaturated fatty acid named by this system are methylene-interrupted.

1.2. Gas chromatography in fatty acid analysis

Quantitative analyses of fatty acid composition are usually performed by gas chromatography. The fatty acids that are esterified in various lipid classes are converted to fatty acid methyl esters (FAME) before the gas chromatographic analysis [6].

Martin and Synge in 1941 for the first time described the theory for a new type of chromatography based on partition of solute between two liquid phases [7]. After ten years James and Martin then described the first application of gas-liquid partition chromatography to the separation of free fatty acids [8]. Since then gas chromatography has been widely employed for the analysis of volatile and thermally stable organic compounds. The introduction of open tubular capillary columns led to great improvements in separation power and allowed analyses of complex samples with hundreds of analytes.

Given the high capability of modern capillary columns for separation of complex mixtures, GC analyses of FAMEs can be carried out routinely in laboratories. There are many commercially available capillary columns made especially for the analysis of FAMEs over a wide range of chain lengths and number of double bonds. Capillary columns have advantages of high resolution capacity over a packed column. But they are easily overloaded with sample, which may decrease their resolution and quantitation capabilities [9].

1.3. The theory of Gas Chromatography

A typical gas chromatograph contains an injection system where a sample is introduced; the column in which the separation of components takes place and the detector, which gives the signal of the components eluted from the column.

When a volatile sample is injected to gas chromatograph solutes in the sample immediately partition between a stationary phase and a gaseous mobile phase (often called carrier gas). Helium, hydrogen, nitrogen and argon are typically used as carrier gas. The sample is carried through the column where separation of the sample components takes place based on their ability to distribute between stationary phase and mobile phase [10, 11, 12]. The distribution of solutes between stationary phase and mobile phase is expressed by the **retention factor**, *k*, of the solute and given by the equation:

Equation 1: $$k = \frac{amount\ of\ analyte\ in\ stationary\ phase}{amount\ of\ analyte\ in\ mobile\ phase}$$

All molecules spend the same amount of time in the mobile phase and separation is dependent on the difference in retention of the solutes by the stationary phase. The retention

factor can be affected by factors like temperature, type of stationary phase, stationary phase thickness, and column diameter [11]. In isothermal gas chromatography where temperature is the same throughout the run retention factor can also be given in terms of retention times.

Equation 2: $$k = \frac{t_R - t_M}{t_M} = \frac{t'_R}{t_M}$$

Where t_R is the **retention time** of a compound, which is the time taken by the compound from introduction of the sample to GC to the appearance of its peak maximum. The **hold up time**, t_M, also called 'dead time', is the time it takes for a non-retained compound to pass through the column. The difference between the retention time of certain compound and the holdup time is called **adjusted retention time**, t'_R, and this is the time the compound spend in the stationary phase (Figure 1).

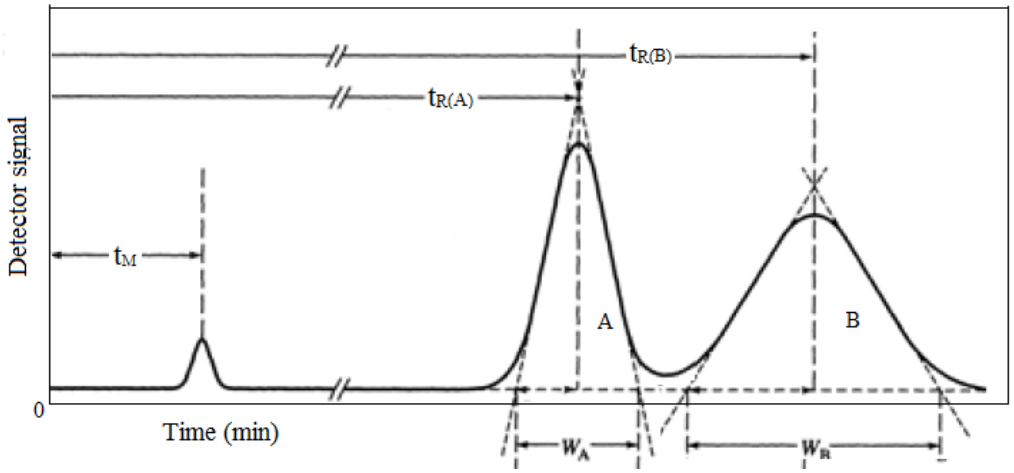


Figure 1: Two closely eluted peaks of compounds A and B, with slight overlapping of the peaks.

The goal of chromatography is to separate sample components into a series of chromatographic peaks, each representing a single component in the sample mixture. The degree of separation between two chromatographic peaks A and B can be measured by **resolution** (R_s), which is given by [12]:

Equation 3: $$R_S = \frac{t_{R(A)} - t_{R(B)}}{\frac{1}{2}(w_{b(A)} + w_{b(B)})} = \frac{\Delta t_R}{\bar{w}_b}$$

Where $t_{R(A)}$ and $t_{R(B)}$ are the retention times of compounds A and B respectively, and $w_{b(A)}$ and $w_{b(B)}$ are the corresponding peak width at baseline of the peaks of the compounds (Figure 1).

The separation between two peaks increases if R_s is increased. This can be achieved either by increasing distance between the two peaks, Δt_R, or decreasing the peak width, w_b. The Δt_R is increased by increasing the difference in retention between the solutes, which means increasing the **selectivity**. Chromatographic selectivity (α) or relative retention between two peaks can be given by the retention factor as:

Equation 4: $$\alpha = \frac{k_B}{k_A}$$

Where k_A and k_B are retention factors of solute A and B, respectively.

The other alternative for increasing resolution is by narrowing the peak width. Initially when chromatographic separation starts to take place narrow bands of finite width of solutes appear. But as the separation proceeds through the column a phenomenon called band broadening will take place which increases the width of the solutes band. Quantitatively this column broadening is measured by column **efficiency**. Martin and Synge described a chromatographic column as different discrete sections where a partitioning of the solutes occurs between the stationary phase and mobile phase [7]. These discrete sections are called theoretical plates. Chromatographic efficiency is therefore traditionally reported as the number of theoretical plates (N):

Equation 5: $$\mathrm{N} = 16\left(\frac{t_R}{w_b}\right)^2$$

Where t_R and w_b are the retention time and base peak width respectively of the peak under consideration. According to Equation 5 at a given retention time for a solute increases in N, thus in efficiency leads to a decrease in the peak width. The number of theoretical plate is also dependent on column length, L, and thus related to each other as:

Equation 6: $$\mathrm{H} = \frac{\mathrm{L}}{\mathrm{N}}$$

Where H is called height equivalent to theoretical plate (HETP) and the smaller its value the higher the efficiency per meter column. When optimizing the efficiency one therefore seek to minimize H.

The three factors leading to chromatographic separation efficiency, selectivity and retention are summarized in one equation called Purnell equation.

Equation 7: $$R_s = \left[\frac{\sqrt{\mathrm{N}}}{4}\right]\left[\frac{\alpha-1}{\alpha}\right]\left[\frac{k_B}{1+k_B}\right]$$

Where the terms in the first, second and third brackets accounts for efficiency, selectivity and retention, respectively. The equation tells us where to put effort if we need improved resolution. To double resolution through N keeping other factors constant requires a column four times as long as the original. Improving resolution through k_B is important only when k_B is low , while if poor resolution can be improved through improving α it is usually the best choice.

1.3.1. Models of band broadening

Different factors are proposed as causes of chromatographic band broadening. These are the multiple path effect, longitudinal diffusion and resistance to mass transfer.

Multiple paths: In packed columns different paths exist for solute to pass through as it moves through the column. Thus the time it takes the solute to elute out of the column depends on the length of the path followed. This elution time difference for solutes in the same band causes band broadening. Non homogeneous packing and large particles will increase this factor.

Longitudinal diffusion: Solutes are constantly in motion and thus diffuse through the mobile phase. Given the higher concentration of the solute in the centre of the chromatographic band more solute diffuse towards the band's forward and rear edges than to the centre resulting in band broadening. To minimize longitudinal diffusion one should minimize the time the analyte spend in the mobile phase, which in practice means that the mobile phase velocity should be high.

Mass transfer: There is a continuous exchange of molecules between the mobile and the stationary phase. However, the exchange between the two phases takes time and before the molecules can move from one phase to another they must diffuse first to the interface between the two phases. While some molecules of an analyte are trapped in the stationary phase the molecules in the mobile phase will move further down the column and the distance they have moved depend on the carrier gas velocity. This is therefore a band broadening effect that increases with mobile phase velocity.

In his equation van Deemter put all the three terms together expressing H as a function of carrier gas velocity (u) [12].

Equation 8: $$\mathrm{H} = \mathrm{A} + \frac{\mathrm{B}}{u} + \mathrm{C}u$$

Where the A term describes the multiple path effect, the B term describes the molecular diffusion of the solute in the mobile phase and the C term describes the resistance to mass transfer of the solute. The effects of these three terms are illustrated in Figure 2. The contribution from the A term is independent of mobile phase velocity. The effect of B term is higher at low mobile phase velocity and then rapidly decreases with mobile phase velocity whereas the contribution from the C term increases with mobile phase velocity. The mobile phase velocity where the sum of the three terms is at the minimum is referred to an **optimal velocity**. Since H has a minimum at optimal velocity the partial derivative of Equation 8 with respect to u is equal to zero. Then solving for the optimal velocity, u_{opt} gives us:

Equation 9: $$u_{opt} = \sqrt{\frac{B}{C}}$$

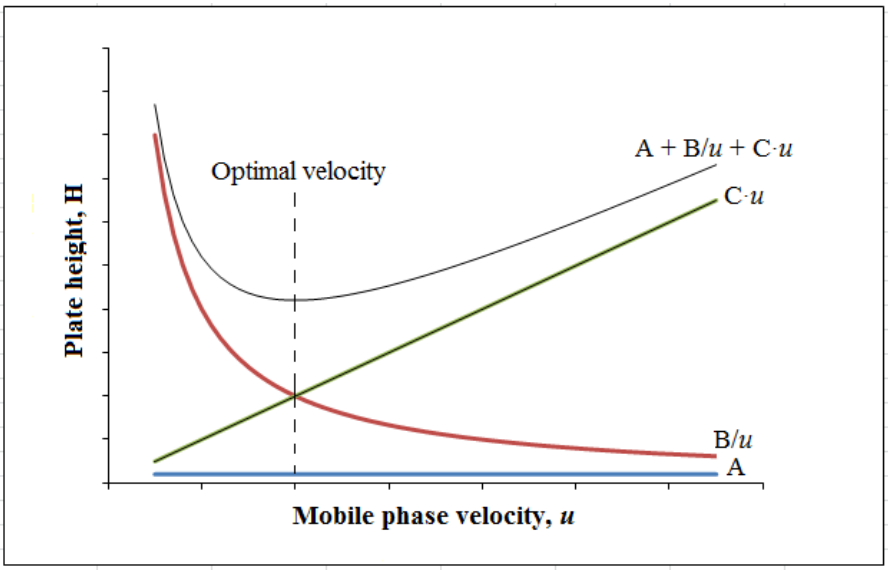


Figure 2: The van Deemter curve.

In capillary columns the effect of band broadening due to multiple pathways should be absent since there is no column packing. Thus removing the A-term from the van Deemter equation gives the Golay equation:

Equation 10: $$H = \frac{B}{u} + Cu$$

The van Deemter equation (Equation 8) and the Golay equation (Equation 10) are the most common equations that explain band broadening in chromatography but there are several alternatives and variants of these that fit observations better under certain circumstances [13, 14, 15].

The carrier gas velocity,u, is not constant throughout the column. Since it is made to pass through the column by applying pressure and gas is compressible it has higher velocity at the end of the column than in the beginning. Therefore it always refers to average carrier gas velocity. L.M. Blumberg [13] in his discussion published on Journal of chromatography A explained the incorrectness of the form of equations that assumes dependence of H on a carrier gas time-averaged linear velocity, $u = {}^{L}/_{t_M}$(average velocity). The reason is that the equations have not been proven and it is in sufficient disagreement with experimental data. For a thin film capillary column with high pressure drop the correct formula converges to:

Equation 11: $$\mathrm{H} = \frac{\mathrm{B}}{u^2} + \mathrm{C}u^2$$

1.4. Carrier gases

Even though basically any gas can be used as carrier gas in gas chromatography commonly used gases are helium, hydrogen and nitrogen. In capillary gas chromatography the choice of carrier gas is particularly important since the optimal conditions are very dependent on the size and diffusion of the gas molecule. Figure 3 shows van Deemter curves for the three carrier gases where lowest H is achieved with nitrogen in an optimum velocity of approximately 10 cm/s. When its velocity is increased above optimum velocity H decreases rapidly and results in rapid loss of efficiency. Helium on the other hand has somewhat higher minimum H but is flatter than nitrogen, making it possible to run at a velocity higher than its optimum without remarkable loss in efficiency. Hydrogen with only slightly higher minimum H than nitrogen has even higher optimum velocity and flatter curve than helium. Even by doubling the velocity hydrogen results in only small loss of efficiency [10, 11].

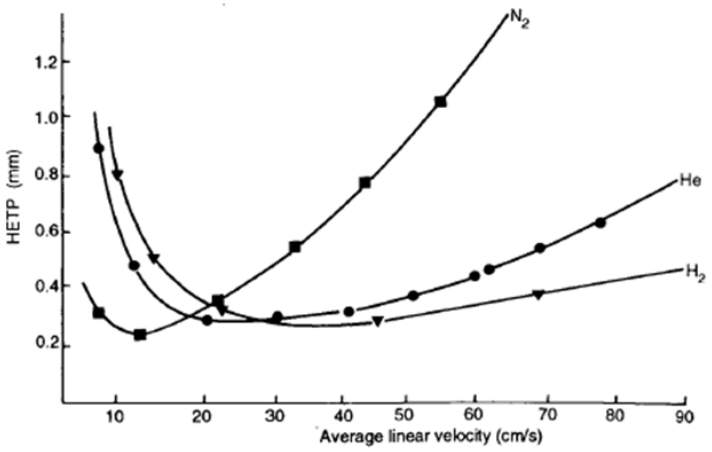


Figure 3: van Deemter curve for helium, hydrogen and nitrogen as carrier gases (non-polar column: 25 m, 0.25 internal diameter) [11].

Hydrogen has a 20 and 40% speed advantage over helium at low and high column pressure drops, respectively, whereas helium in turn has about 55 and 25% speed advantages over nitrogen at low and high pressure drops, respectively [16]. Thus hydrogen is the preferred carrier gas when speed of analysis is the target. But there is a safety concerns in using hydrogen as a carrier gases in some laboratories. This concern can be addressed by using accurate safety informations, safety interlocks, hydrogen generators with limited capacity and inherently safe instrument designs [16].

1.5. Temperature programmed gas chromatography

In temperature-programmed gas chromatography the oven temperature is increased with time, which allows analysis of analytes with a broader range of volatility than isothermal GC. This leads to different solute-stationary phase and solute-mobile phase interactions over the time of analysis. Under such conditions the retention factor (*k*) varies, so equations based on *k* are no longer valid. In addition Equation 5 is also not valid since there is a relationship between N and *k*. Thus selectivity and efficiency for temperature programmed GC must be redefined.

1.5.1. Selectivity in temperature-programmed GC

Based on different proposals how to use retention data from different published results the expression of retention relative to a single standard substance became widely used. This is usually referred to as relative retention and the principle is similar to that described by Equation 4. However, in temperature programmed GC these numbers are largely system dependent, especially when the chemical properties of the analytes differ largely from the reference. The difficulty of having a single standard always close to the substances of interest and the temperature dependency of relative retention lead Kovats to the proposal of the so called retention index system [17]. Retention index systems express the retention behaviour of the compounds of interest relative to a series of homologous standard substances. Thus the following equation is derived to determine the retention index (I_x) of any given substance *x*:

Equation 12: $$I_X = 100\left[\frac{log\, t'_{R(x)} - log\, t'_{R(z)}}{log\, t'_{R(z+1)} - log\, t'_{R(z)}} + z\right]$$

Where *x* is the compound of interest, *z* is the n-alkane with *z* carbon atoms eluting before the compound of interest and *z+1* is the n-alkane with *z+1* carbon atoms eluting after the compound of interest [18]. The above equation applies to isothermal conditions. In a linear

temperature program a similar formula that takes retention times straight without logarithmic term and adjustments is applied [19, 20].

Equation 13: $$I_X = 100\left[\mathrm{n}\frac{t_{R(x)}-t_{R(z)}}{t_{R(z+n)}-t_{R(z)}}+z\right]$$

Where n is the difference in carbon number of the two n-alkanes used as a reference while the other terms are the same as Equation 12 above.

F. P. Woodford and C. M. Van Gent, 1960 has also revealed the linear relationship between the logarithm of the retention time of the fatty acid methyl esters (FAMEs) and their chain length (‘‘carbon number’’). They demonstrated that also the saturated esters have integral carbon-numbers, whereas esters with branched chains and unsaturated esters have non-integral carbon numbers [21]. With what seems to be derived from this ‘‘carbon number’’ concept, Equivalent chain length (ECL) is used to describe retention of fatty acid derivatives. The FAMEs of straight chain saturated fatty acids are used as the reference compounds. For a programmed temperature GC a modification of the van den Dool and Kratz equation can be used [22, 23, 24].

Equation 14: $$ECL_x = n\frac{t_{R(x)}-t_{R(z)}}{t_{R(z+n)}-t_{R(z)}}+z$$

Where $t_{R(x)}$ is the retention time of compound *x*, $t_{R(z)}$ is the retention time of a saturated straight chain FAME eluting before *x* and *z* is the number of carbons in the fatty acid chain of this molecule without carbon of the methyl group, $t_{R(z+n)}$ is the retention time of a saturated straight chain FAME eluting after *x* and *n* is the difference in carbon between the two reference FAMEs.

1.5.2. Efficiency in temperature-programmed GC

In temperature programmed gas chromatography N is not a valid measurement of efficiency. Thus an alternative measure must be applied. L. S. Ettre, 1975 first introduced the concept of separation number to express column performance [25]. Separation number is defined as the number of peaks that are separated between two consecutive members of a homologous series. It is expressed with the following equation:

Equation 15: $$\mathrm{SN} = \frac{t_{R(z+1)}-t_{R(z)}}{w_{h(z+1)}+w_{h(z)}} - 1$$

Where $t_{R(z)}$ and $t_{R(z+1)}$ are the retention time of the two members of the homologous series with *z* and *z+1* carbon numbers respectively, and $w_{h(z)}$ and $w_{h(z+1)}$ are the respective peak widths at half peak heights [26].

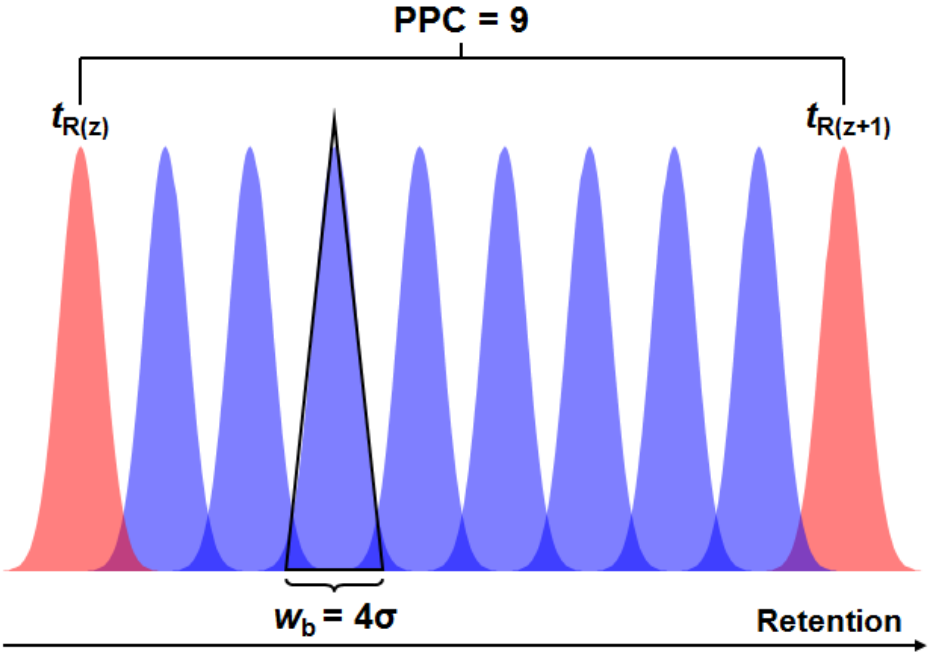


Figure 4: Peaks separated between two homologous series with chromatographic resolution of one.

Since SN is a rough approximation of the number of peaks that can be eluted between two members of a homologous series its inverse is not a good replacement of H. This is because when SN is zero still the homologous are separated, which means there is some separation efficiency. Thus as an alternative to SN the peak per carbon (PPC) is used in this study, this is the number of peaks that can be separated with chromatographic resolution equal to one per compound in a homologous series. Mathematically PPC can be expressed as the difference in retention time between the two homologous compounds divided by the average peak width at baseline.

Equation 16: $$\text{PPC} = \frac{t_{R(j+1)} - t_{R(j)}}{0.5(w_{b(j+1)} + w_{b(j)})}$$

The above equation can be simplified if the measurement of retention and peak width is on retention index scale where the retention difference between the homologous is given by definition (equal to 1 for equivalent chain lengths, ECL, and equal to 100 for Kovats indices).

Equation 17: $$PPC = \frac{1}{w_{b,RI}}$$

There is also a similar equation that relates SN to peak width in ECL units [26]. In a temperature programmed GC, peak width in retention index units should therefore be minimized to attain maximum efficiency which is similar to minimizing H in isothermal GC. Finally resolution (R_S), peak per carbon (PPC) and equivalent chain length (ECL) are related by the simple relationship:

Equation 18: $$R_s = \Delta ECL \cdot PPC$$

1.6. Response surface methodology

Response surface methodology is a methodology in which response function(s) are obtained from experiments conducted in accordance to predetermined plan by varying the values of predictor variables. The predetermined plan is worked out by Design of Experiments (DoE). The response functions are typically polynomial models that link the response to the experimental settings, and are obtained by regression [26]. For two variables system the model typically looks like:

Equation 19: $$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2$$

Where x_1 and x_2 represents the main effects, x_1x_2 represents the interaction and; x_1^2 and x_2^2 represents the squared terms of the variables 1 and 2 respectively. The above model includes quadratic terms. Depending on the number of the variables and their effect on the response the model may assume a higher order polynomial or a first order function where only the main effects and the interaction terms are included.

In a temperature programed GC it is possible to assume that peak widths in retention index units (the inverse of efficiency) follows a response function of the two independent variables carrier gas velocity and temperature gradient. However, the van Deemter equation, which explains the inverse of the efficiency as a function of the carrier gas velocity, is not a quadratic function. Assuming that peak width in retention index units (w) follows the van Deemter equation with carrier gas velocity (u) and a quadratic function of temperature gradient (i), a response function that combines the two functions can be generated.

Equation 20: $$w = a + \frac{b}{u} + c \cdot u + d \cdot i + e\frac{i}{u} + f \cdot i \cdot u + g \cdot i^2$$

Where a, b and c-terms are the terms in the original van Deemter equation, d explains the linear effect of i on w, e explains the effect of i on the b-term in the van Deemter equation, f explains the effect of i on the c-term in the van Deemter equation and g explains the quadratic effect of i on w. This expansion of the van Deemter equation to account for an interaction will

be referred to as a VD+Int model throughout the thesis. From the VD+Int model by inserting values for *i* it is possible to calculate ordinary VD models at any temperature gradient if the coefficients a-f are known.

At an optimum velocity the partial derivative of Equation 20 with respect to *u* is equal to zero, since *w* has a local minimum at an optimum velocity value. One can therefore estimate u_{opt} at any temperature gradient if the parameters b, c, e and f are known:

Equation 21: $$u_{opt} = \sqrt{\frac{b+e.i}{c+f.i}}$$

1.7. Experimental design (DoE)

The use of experimental design in chromatographic science has grown rapidly in the past two decades. D. Brynn Hibbert has reported that chromatographic research in which DoE has been applied were increased from less than 20 papers published in the year 1991 to more than 200 papers in the year 2010. This growing need for DoE is necessitated to increase the efficiency of scientific discovery and decrease the cost of experimentation [27].

There are different types of experimental designs used in the field of analytical chemistry and their use depends on the purpose and the kind of experiments that the experimenter needs to conduct [26, 28]. In general when the need is to investigate the most important factor(s) experimental designs such as factorial or Plackett-Burman can be applied whereas to optimize a process or to obtain a response function experimental designs such as central composite design, Box-Behnken design, Doehlert design or others can be applied [28, 29]. The experimental designs applied in this study are briefly discussed below. There are two factors to be investigated in this study: the velocity of the carrier gas (helium, hydrogen and nitrogen) and the temperature gradient.

1.7.1. Factorial design

A full factorial design investigates experiments at every combination of factor levels in the experiments. For k number of factors at L levels it requires L^k number of experiments to be conducted. At two levels only four experiments are needed for two factors while eight experiments should be carried out for three factors. The number of experiments increases drastically as the number of levels increase. For three levels and three factors it requires twenty seven experiments to be run. As the number of factors increase a fractional factorial design may be used to decrease the number of experiments to a reasonable number and thereby minimizing the cost of experimentation, but it puts limitations on what information

that can be gained and the reliability of the results may also be reduced. In fractional factorial design only a fraction of the total experiments covered by full factorial design should be run. Factorial designs are widely used for screening purpose to investigate main and interaction effects [26]. With two level factorial or fractional factorial the models that fit to the design are first order models. In order to represent quadratic or higher order models three or more level factorial designs are mandatory [29]. It is also possible with factorial design to investigate different factors at different number of levels [28]. The 3 × 3, 5 × 3 and 9 × 3 (carrier gas velocity x temperature gradient) factorial design experiments carried out in this study are examples of this scenario.

1.7.2. Central composite design

Central composite designs are created by combining two level factorial designs and additional star and centre points. This enables the central composite design to determine linear and quadratic models. The factorial points are important for determining the interaction terms while the star points are used to determine the quadratic terms. CCD has the properties of orthogonality or rotatability and the design approximates a spherical surface [26, 29]. Depending on the distance of the star points from the central point there are three different forms of the CCD. One in which the star points are located equidistant to the factorial points called circumscribed design, a second when the star points are in the space of the factorial design called inscribed design and third when they lie on the faces of the factorial design called face centred design. The total number of experiments or design points required in CCD is equal to $L^k + Lk + n_c$, where L, k and n_c are number of levels, factors and centre point respectively [27]. The three level full factorial design (3 × 3) run in this study can be considered also as face centred central composite design.

1.7.3. Doehlert design

Doehlert design was introduced first by Doehlert in 1970. Characterized by its uniform shell design nature Doehlert design maximises the coverage of a spherical experimental domain. Doehlert designs are not rotatable in the sense that they have different number of levels on different parameters. This may be beneficial when a factor appears to be more important or has stronger effect than other factors. A Doehlert design needs a fewer experiment than the central composite design. For two factors one central point and six points of a regular hexagon forms the Doehlert design. One of the two factors assumes five levels while the other takes three levels in the design. The number of experiments in Doehlert design is equal to $k^2 + k + n_c$ points [26, 27, 28, 29]. Doehlert design has received many applications in the

area of chromatographic, spectrometric and electroanalytical fields in optimizing chemical or instrumental variables. It is often employed for optimization of extraction steps and/or optimization of instrumental conditions for determination or separation of analytes [28, 29].

2. Experimental

2.1. Gas chromatograph

All analyses were performed using two Agilent 7890A gas chromatographs. One of the chromatographs has a possibility of alternating between helium and hydrogen as carrier gases whereas the other has the possibility of alternating between helium and nitrogen. Both chromatographs were equipped with split/splitless injector, electronic pressure control, autosampler and FID detector. The GC systems were controlled by Agilent Chemstation B.04.03. A 5 µL syringe size was used to inject 1 µL of FAMEs sample to the injection port. The injection mode used was splitless injection at 250°C. A pre and post wash of the injection needle were performed using methanol and isooctane. The FID detector was heated at a temperature of 260 ^{0}C and the flows of the carrier gas, air and make up gas were at 40, 400 and 40 mL/min respectively. The purity of helium and nitrogen gases used was 99.999% while hydrogen was generated by a Parker Balston FID1000 gas generator.

All experiments were performed in constant flow mode, which means that the mass flow of carrier gas from the column was constant throughout the chromatographic run. Because of gas expansion as the oven temperature increases, in reality the carrier gas velocity continuously increases. Thus the term "nominal carrier gas velocity" refers to the estimated average velocity at injection temperature (60 °C), assuming that actual column dimensions were identical to nominal dimensions. All velocities in the results part refer to the nominal average velocities, and these were estimated by the built-in algorithm in the chromatographs.

2.2. Columns

The capillary columns that are employed in this study have low to high polarity. The general purpose DB-5 columns with 95% methylpolysiloxane stationary phase are non-polar and characterized by low bleeding and high temperature limit. The medium polar DB-23 column with 50% cyanopropyl substituted polysiloxane is designed for the analysis of FAMEs and resolution of cis- and trans-isomers.

The BPX70 columns are highly polar columns with 70% cyanopropyl polysilphenylene-siloxane stationary phase. They are also mainly designed to analyse FAMEs and characterized by their high temperature limit and low column bleeding. The BP-20 is a polar column with polyethylene glycol polymer stationary phase. Polyethylene glycol based columns vary only based on the cross linking and deactivation process involved during manufacturing of the columns [30].

The IL columns are columns with ionic liquid phase which are different physically and chemically from the polysiloxane and polyethylene glycol columns. While all IL columns are polar columns, the IL82 and IL100 are highly polar columns with IL100 the most polar one, and IL61 has a polarity that is comparable to DB23 and BP20.

Table 1: Description of columns used in the study.

	Column type	**Stationary phase**	**L (m)**	**I.D (mm)**	**D_f (µm)**	**Temperature limit (^{0}C)**
1	DB-5[a]	5% phenyl, 95% methylpolysiloxane	30	0.25	0.1	350
2	DB-5[a]	5% phenyl, 95% methylpolysiloxane	30	0.25	0.25	350
3	DB-23[a]	(50% Cyanopropyl)-methylpolysiloxane	30	0.25	0.25	260
4	BP-20[b]	Polyethylene Glycol (PEG)	30	0.25	0.25	260
5	BPX70[b]	70% Cyanopropyl Polysilphenylene-siloxane	30	0.22	0.25	260
6	BPX70[b]	70% Cyanopropyl Polysilphenylene-siloxane	60	0.25	0.25	260
7	SLB-IL6[c]	1,12-Di(tripropylphosphonium)dodecane bis(trifluoromethylsulfonyl)imide trifluoromethylsulfonate	30	0.25	0.2	290
8	SLB-IL82[c]	1,12-Di(2,3-dimethylimidazolium)dodecane bis(trifluoromethylsulfonyl)imide	30	0.25	0.2	270
9	SLB-IL100[c]	1,9-Di(3-vinylimidazolium)nonane bis(trifluoromethylsulfonyl)imide	30	0.25	0.2	230

L - Length of column; I.D - Internal diameter; D_f - Stationary phase film thickness

a- Agilent, Santa-Clara, CA, USA, b- SGE, Ringwood, Victoria, Australia and c- Supelco, Bellefonte, PA, USA

2.3. Samples

Fatty acid methyl esters of saturated homologous fatty acids and unsaturated fatty acids were used for the study (Table 2). The FAMEs were distributed in two samples to avoid overlaps of some FAMEs except for BPX70 where there is no overlaps of peaks of the applied FAMEs. The samples were injected in 1-2 ng amounts of each compound, which gave symmetric peaks.

Table 2: Saturated, monoenoic and polyunsaturated fatty acid methyl esters (FAME) with their systematic name, trivial name and shorthand designation which are used in the study.

Systematic Name	Trivial Name	Shorthand Designation
Dodecanoic	Lauric	12:0
Tridecanoic		13:0
Tetradecanoic	Myristic	14:0
Pentadecanoic		15:0
Hexadecanoic	Palmitic	16:0
9-hexadecenoic	Palmitoleic	16:1 n-7
Heptadecanoic	Margaric	17:0
Octadecanoic	Stearic	18:0
9-octadecenoic	Oleic	18:1 n-9
Di-trans-9,12-octadecadienoic	Linolelaidic	18:2 n-6 tt
6,9,12-octadecatrienoic	γ-Linolenic	18:3 n-6
Nonadecanoic		19:0
Eicosanoic	Arachidic	20:0
8,11,14-eicosatrienoic	homo-γ -linolenic	20:3 n-6
5,8,11,14,17-eicosapentaenoic	EPA	20:5 n-3
Heneicosanoic		21:0
Docosanoic	Behenic	22:0
4,7,10,13,16,19-docosahexaenoic	DHA	22:6 n-3
Tetracosanoic	Lignoceric	24:0
Hexacosanoic		26:0

The systemic name given is followed by 'acid methyl ester' and all the double bonds have *cis* geometry unless specified.

2.4. Experimental designs

The FAME samples were run in nine or more levels of carrier gas velocities and at three levels of temperature gradients on all columns with the three carrier gases used. The carrier gas velocities and temperature levels applied are listed in Table 4. Higher temperature gradients were employed for non-polar columns while lower gradients were used for polar columns with the exception of IL61 that was run at higher gradient.

The general oven heating conditions for all columns under study was as follows: Samples were injected at a column temperature of 60 ^{0}C that was held for 3 minutes for proper analyte focusing. Then the oven was heated with 60 ^{0}C/min to start temperature of A ^{0}C, which was followed by the main temperature rate of B ^{0}C/min until the last component had eluted. The runs were carried out at three levels of gradient temperatures, except on BPX70 with hydrogen that was run at four levels that was used to develop the retention time model (Table 3). The start temperature was set in such a way that FAME of saturated fatty acid of 19:0 should elute at approximately 27.6 min with helium at 2 ^{0}C/min and 30 cm/s carrier gas velocity.

Table 3: Start and gradient temperatures used for different column types.

	Column type	**Start temperature (^{0}C)**	**Gradient temperature (^{0}C)**
1	DB-5 (0.1 µm)	131.72	2, 4, 6
2	DB-5 (0.25 µm)	164.14	2, 4, 6
3	DB-23	131.09	1, 2, 3
4	BP-20	148.15	1, 2, 3
5	BPX70 (30 m)	125.02	1, 2, 3
6	BPX70 (60 m)	125.02	1, 2, 3, 4*
7	SLB-IL61	134.75	2, 4, 6
8	SLB-IL82	119.00	1, 2, 3
9	SLB-IL100	108.47	1, 2, 3

* - fourth level temperature rate only with hydrogen

The experimental runs were performed in a randomised sequence. First the experiments were listed on excel with each temperature level and carrier gas velocity combination in increasing order of both factors and then randomized on excel using a random number generated in a next column to the list of the experiments. Then the GC sequence analysis is set up for run with the randomized sequence. The lists of experiments conducted are given in Table 4.

Table 4: List of experiments carried out on different columns studied.

Experiment number	Temperature rate (^{0}C/min)	Carrier gas velocity (cm/s)			Experiment number	Temperature rate (^{0}C/min)	Carrier gas velocity (cm/s)		
		He	H_2	N_2			He	H_2	N_2
1	1	14	14	8	1	2	14	18	8
2	1	18	22	10	2	2	18	26	10
3	1	22	30	12	3	2	22	34	12
4	1	26	38	14	4	2	26	42	14
5	1	30	46	16	5	2	30	50	16
6	1	34	54	18	6	2	34	58	18
7	1	38	62	20	7	2	38	66	20
8	1	42	70	22	8	2	42	74	22
9	1	46	78	24	9	2	46	82	24
10	2	14	14	8	10	4	14	18	8
11	2	18	22	10	11	4	18	26	10
12	2	22	30	12	12	4	22	34	12
13	2	26	38	14	13	4	26	42	14
14	2	30	46	16	14	4	30	50	16
15	2	34	54	18	15	4	34	58	18
16	2	38	62	20	16	4	38	66	20
17	2	42	70	22	17	4	42	74	22
18	2	46	78	24	18	4	46	82	24
19	3	14	14	8	19	6	14	18	8
20	3	18	22	10	20	6	18	26	10
21	3	22	30	12	21	6	22	34	12
22	3	26	38	14	22	6	26	42	14
23	3	30	46	16	23	6	30	50	16
24	3	34	54	18	24	6	34	58	18
25	3	38	62	20	25	6	38	66	20
26	3	42	70	22	26	6	42	74	22
27	3	46	78	24	27	6	46	82	24

2.5. Calculations and software

2.5.1. Software

Chrombox C and Chrombox O (Optimizer) were used to handle and analyse the GC data. Chrombox C, which reads Agilent Chemstation raw data, was used to integrate, identify based on retention indices and measure peak width of chromatographic peaks of the FAMEs. It can convert chromatograms on a retention time scale to retention indices so that peaks are identified based on templates of retention indices and peak widths are measured on the retention index scale (ECL). Chrombox O which reads data from Chrombox C was used for setting up experimental designs, developing models of van Deemter (VD), van Deemter plus interaction (VD+Int) and retention time models. It was also used to calculate model fits and errors.

The general procedure followed using both Chrombox C and Chrombox O is expressed as follows:

- First the data obtained in Chrombox C from the Agilent Chemstation was imported from a raw data folder to the import box window of Chrombox C.
- The peaks in the imported chromatogram were integrated.
- The retention scale was converted to retention index scale and the peaks for the saturated FAMEs were calibrated by typing their corresponding retention index which is equal to the number of carbons in the fatty acid chain. Then the peaks were identified and the chromatogram was saved.
- In the design window of Chrombox O the design was defined by importing the design from a csv file (comma separated values) created in advance.
- Result files from Chrombox C were imported to the experiments window and assigned to different conditions in the design. Then the experiment was saved.
- Finally by opening the windows for models experiments were loaded to the window which contains functions for van Deemter equation and modifications of it. Here the van Deemter models for individual FAMEs as well as average models were obtained. The regression coefficients were found by least square regression. Optimal carrier gas velocities at different temperature levels and average peak widths at optimal velocity as well at other values were also obtained. Response surface models for retention time and inverse of efficiency (peak width) were solved and their plots were obtained. Additionally, model fits and errors were also calculated.

Equation 22: $$\mathrm{RMSE} = \sqrt{\sum_1^n \frac{(y_{\mathrm{meas}} - y_{\mathrm{pred}})^2}{n-p}}$$

Where y_{meas} is the measured peak width in retention index units or retention times, y_{pred} is the predicted peak width in retention index units or retention times; p is the number of coefficients in the applied model and n is number of samples predicted for the FAME. In general the models shown are mean models of the individual models of different FAMEs. In these cases the reported RMSE are average RMSE for individual models. However, reported R^2 are calculated from predicted versus measured values of all observations. For example if models are based on nine experiments and 17 FAMEs in the chromatograms, 17 RMSE values are calculated (one per FAME) and the reported value is the average RMSE. The R^2 value is calculated directly from predicted versus measured values of the total 153 (9x17) measured responses.

2.5.2. Retention time and retention index

Retention indices were related by a stepwise procedure using local second order regression as explained in [31]. Unbranched saturated fatty acids from 12:0 to 26:0 (13 and 23:0 are not included) were used as a references. To calculate the relationship for any interval between two references (n and n + 1) a polynomial regression with, $f_1(t_x)$, was fitted to three reference compounds n−1, n, and n + 1 and a second polynomial regression $f_2(t_x)$, was fitted to three reference compounds n, n + 1, and n + 2. Since the range between n and n + 1 is covered by both polynomial functions then any retention time in the interval is converted to the corresponding retention index by weighting the two functions.

Equation 23: $$RI = (1 - w) \times f_1(t_x) + w \times f_2(t_x), \quad w = \frac{t_x - t_n}{t_{n+1} - t_n}$$

The procedure is only used to calculate the ECL values for retentions between the second and the second last reference compound. Retention indices for the next interval are calculated by increasing n by one. Figure 5 shows section of chromatogram of FAMEs in retention time scale converted to ECL scale using Equation 23.

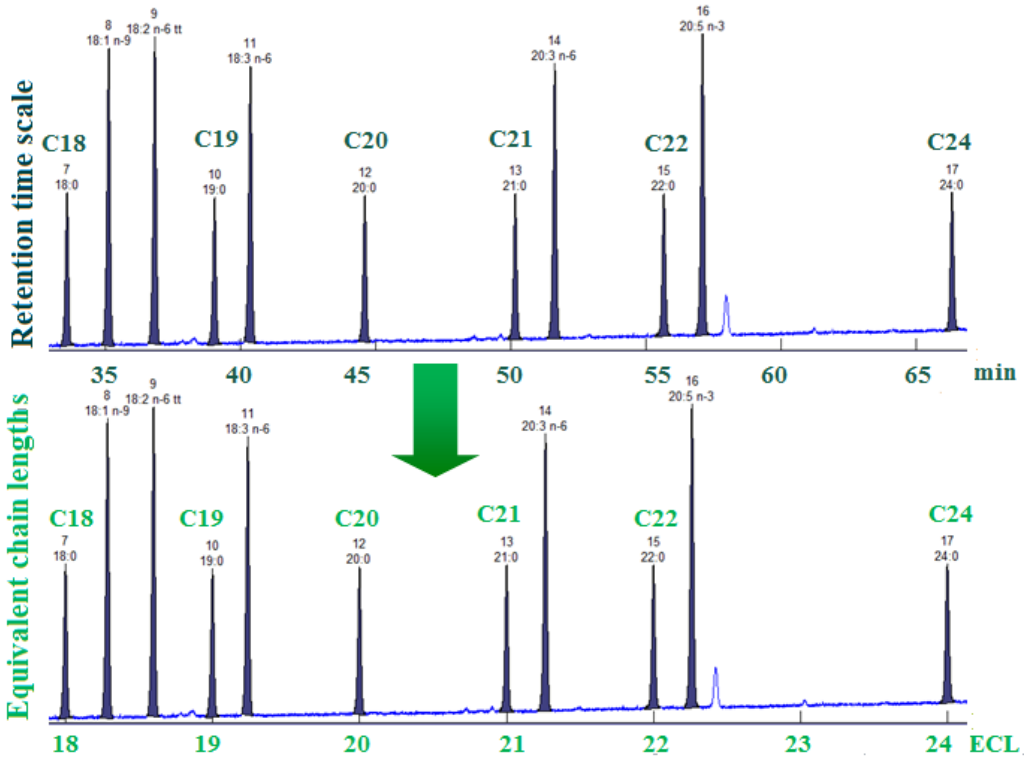


Figure 5: Chromatogram of FAMEs on retention time and equivalent chain length scale (only section of chromatogram from C-18 to C-24 is shown).

2.5.3. Fatty acid methyl esters (FAMEs) included in models

All the FAMEs of the fatty acids listed in Table 2 were used during the peak width model fitting and the response surface equation evaluation except the first eluting FAME of 12:0 and the last eluting FAME of 26:0. The two FAMEs are excluded since the peak width of the first eluting FAME may be affected by the start temperature of the gradient and 26:0 has low solubility on some of the stationary phases, which may lead to asymmetric peaks. In the work described in Section 3.6 the FAMEs of the saturated fatty acids from 19:0 to 24:0 were also excluded. This is because long chain saturated fatty acids will typically give asymmetric peaks on the most polar phases. These fatty acids are not abundant in nature and they are therefore of limited interest for the evaluation of column performance.

3. Results and discussion

3.1. The van Deemter's curve fitting

The ordinary van Deemter equation (Equation 8 and Figure 2) is traditionally used to model H as a function of carrier gas velocity (u) in isothermal chromatography. Since N is not a valid measure of efficiency in temperature programmed GC it can be substituted by peak per carbon (PPC) whereas peak width (w_b) is used instead of H (Equation 17). Thus w_b can be used to model van Deemter equation in temperature programmed GC, but it is necessary to evaluate how well the model fits the data.

First individual models for each FAME compound were obtained from peak width measurements made at the experimental conditions. As shown in Figure 6a the peak widths of each compound fitted fairly well to the van Deemter equation. The A terms are in this case close to zero, which is in accordance with the theory of capillary columns. With capillary columns positive A-terms may explain extra-column effects, such as poor focusing of the analytes at injection or band broadening in the detector. The A term may also explain lack of fit for the B and C terms since there are non-ideal conditions (the carrier gas velocity is nominal and not real velocity). Since usually interests are on the average performance of GC program and not the performance for individual compounds average models were calculated from individual models and any further discussions on different models are based on average models (Figure 6b). In general plots show average values and average response surface. The errors are calculated on individual FAME.

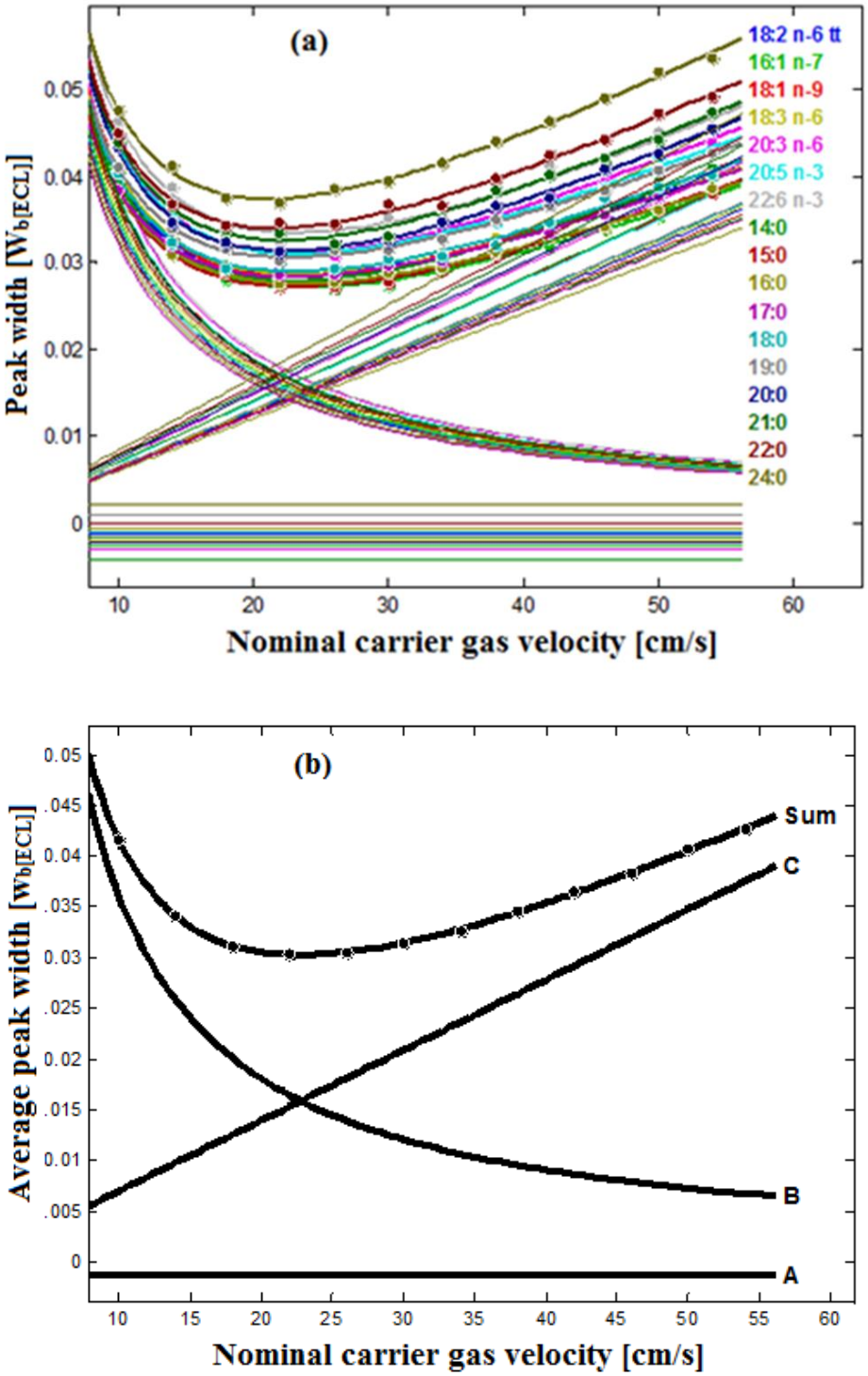


Figure 6: van Deemter models (a) individual FAMEs model (b) average model (column: DB5 30 m, 0.25 mm, 0.1 µm; 2 ^{0}C; helium as a carrier gas).

In isothermal chromatography H used to model chromatographic systems where resolution is proportional to the square root of N (Equation 7). Unlike in the isothermal GC resolution is directly proportional to PPC in temperature programmed GC (Equation 18). In a similar fashion with the isothermal GC then w_b^2 which is proportional to squared PPC can be used to

model van Deemter equation in temperature programed GC. Thus in addition to peak width (w_b**)** squared peak width (w_b^2) was also used to measure efficiency and to model the function in the van Deemter equation.

Experiments conducted on DB-5 and BPX70 columns using helium, hydrogen and nitrogen as carrier gases were used to develop and evaluate the models. Correlation coefficient (R^2) and root mean square error (RMSE) were used to compare the results from models with w_b and w_b^2. All peak width measurements were made on w_b scale and w_b^2 is used only during modeling and after modeling the square root of the predicted w_b^2 was taken to calculate the RMSE.

Both approaches resulted in good models of the van Deemter equation as shown in Figure 7. The plots show that regardless of whether w_b or w_b^2 is used as response the models are fitted well to the experimental points. However, it has to be noted that there is higher difference between the lowest point and the highest point in the plot, about 4 times, where peak width is squared than when normal peak width is used, where the difference is approximately two.

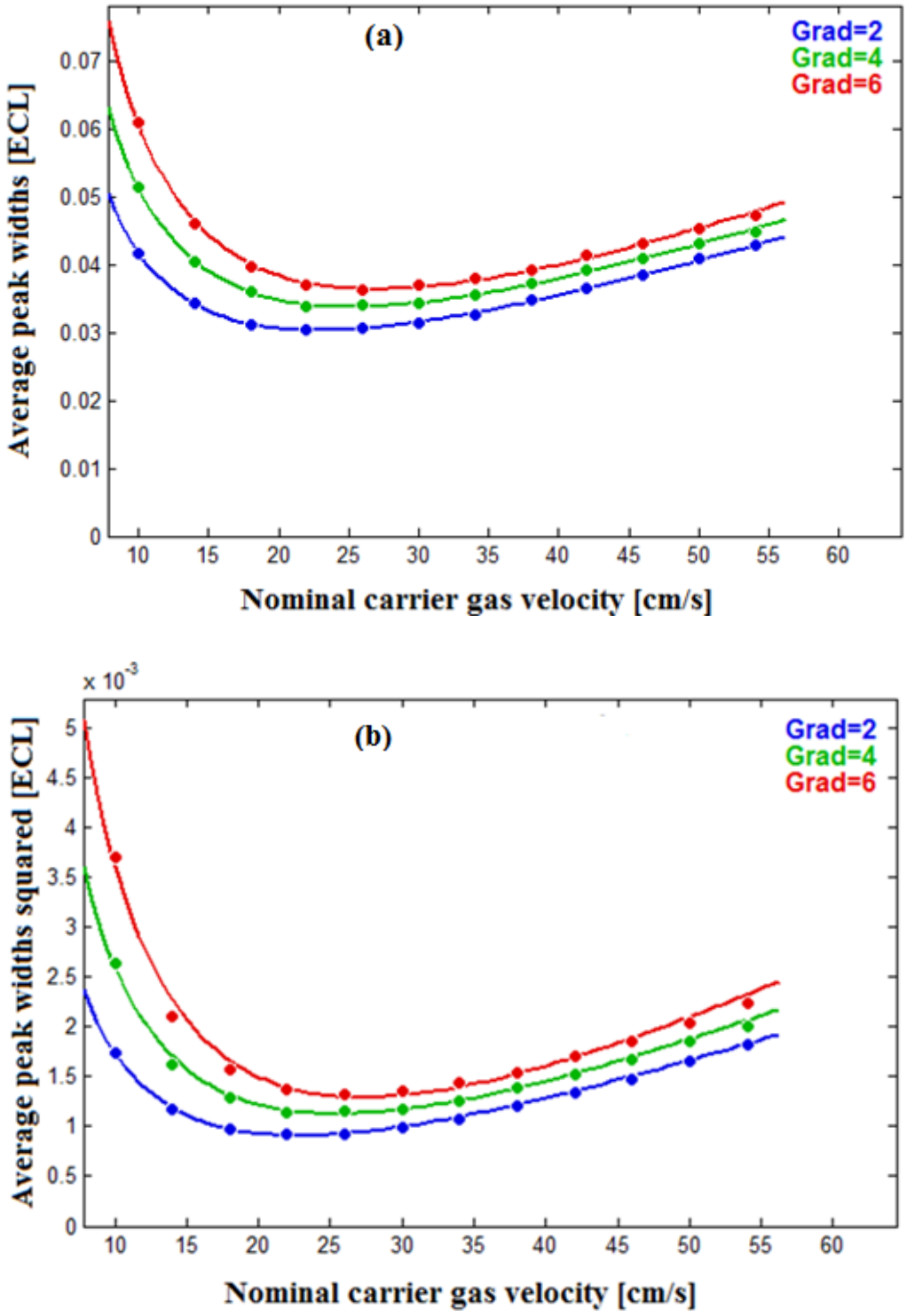


Figure 7: van Deemter curve calculated by inserting the gradient temperatures and using (a) normal peak width and (b) squared peak width (column: DB5 30 m, 0.25 mm, 0.1 µm, helium as a carrier gas).

Correlation coefficients (R^2) and adjusted root mean square errors (RMSE) of the resulting models from both approaches were evaluated and presented as shown in Table 5 and Table **6** respectively. Both R^2 and RMSE are obtained at different constraints on the A-term in the

van Deemter equation, i.e, when A-term is included, kept at its mean value and excluded from the model. When A-term is kept at its mean value all models for individual fatty acids assume the same value for the A-term and it is independent of the temperature rate used. This is important when extra column effects like effects of the injector volume and band spreading in detector is significant in band broadening which are assumed to be equal for all compounds and independent of temperature rate. The exclusion of the A-term is when its column and extra column effects are considered insignificant.

Table 5: Correlation coefficients (R^2) for models obtained with normal peak width (w_b) and squared peak width (w_b^2) for different column types.

	Rate	Overall R^2, A incl			Overall R^2, A = mean			Overall R^2, A = 0		
	(^{0}C/min)	w_b		w_b^2	w_b		w_b^2	w_b		w_b^2
DB-5 30 m, 0.25 mm, 0.1 µm										
Helium	2	0.9928	>	0.9921	0.9839	>	0.9464	0.9914	>	0.8735
	4	0.9913	>	0.9896	0.9882	>	0.9714	0.9883	>	0.8401
	6	0.9936	>	0.9862	0.9929	>	0.9806	0.9841	>	0.7985
Hydrogen	2	0.9947	>	0.9936	0.9931	>	0.9826	0.9893	>	0.9386
	4	0.9923	<	0.9925	0.9908	>	0.9874	0.9876	>	0.9342
	6	0.9887	>	0.9886	0.9862	>	0.9858	0.9845	>	0.9201
Nitrogen	2	0.9950	<	0.9952	0.9877	>	0.9655	0.9862	>	0.9826
	4	0.9923	>	0.9922	0.9908	>	0.9866	0.9905	>	0.9606
	6	0.9856	>	0.9850	0.9827	>	0.9823	0.9826	>	0.9439
DB-5 30 m, 0.25 mm, 0.25 µm										
Helium	2	0.9903	>	0.9898	0.9806	>	0.9452	0.9888	>	0.8680
	4	0.9876	>	0.9820	0.9838	>	0.9625	0.9838	>	0.8194
	6	0.9879	>	0.9793	0.9863	>	0.9720	0.9772	>	0.7794
Hydrogen	2	0.9866	<	0.9879	0.9810	>	0.9702	0.9773	>	0.9250
	4	0.9791	<	0.9798	0.9760	>	0.9709	0.9733	>	0.9116
	6	0.9817	<	0.9829	0.9790	>	0.9770	0.9783	>	0.8932
Nitrogen	2	0.9923	>	0.9921	0.9812	>	0.9565	0.9798	>	0.9725
	4	0.9820	=	0.9820	0.9797	>	0.9732	0.9793	>	0.9453
	6	0.9851	>	0.9850	0.9829	>	0.9816	0.9828	>	0.9396
BPX70 30 m, 0.25 mm, 0.22 µm										
Helium	1	0.9869	>	0.9868	0.9794	>	0.9627	0.9778	>	0.9117
	2	0.9852	>	0.9845	0.9807	>	0.9735	0.9800	>	0.8965
	3	0.9777	>	0.9772	0.9722	>	0.9672	0.9719	>	0.8765
Hydrogen	1	0.9905	>	0.9902	0.9842	>	0.9757	0.9687	>	0.9409
	2	0.9892	<	0.9898	0.9839	>	0.9796	0.9739	>	0.9344
	3	0.9865	>	0.9861	0.9830	>	0.9768	0.9726	>	0.9281
Nitrogen	1	0.9915	<	0.9918	0.9891	>	0.9826	0.9824	>	0.9702
	2	0.9922	<	0.9928	0.9904	>	0.9877	0.9836	>	0.9707
	3	0.9910	>	0.9909	0.9896	>	0.9892	0.9862	>	0.9577
Overall mean		0.9885	>	0.9876	0.9844	>	0.9738	0.9816	>	0.9123
Mean with He		0.9881	>	0.9853	0.9831	>	0.9646	0.9826	>	0.8515
Mean with H_2		0.9877	<	0.9879	0.9841	>	0.9784	0.9784	>	0.9251
Mean with N_2		0.9897	=	0.9897	0.9860	>	0.9784	0.9837	>	0.9603

In general a good fit of the models (R^2) is observed at all levels of temperature gradients applied with all carrier gases used. However there is a decrease in R^2 when the A-term is at

its mean value or removed from the model. This is observed in both cases when peak width and squared peak width are used to model the function in the van Deemter equation (Table 5 and Figure 9 to Figure 11). In most cases when the A-term is included and in all cases when the A-term is either set at its mean value or removed from the model the R^2 is greater when normal peak width is used than when squared peak width is applied.

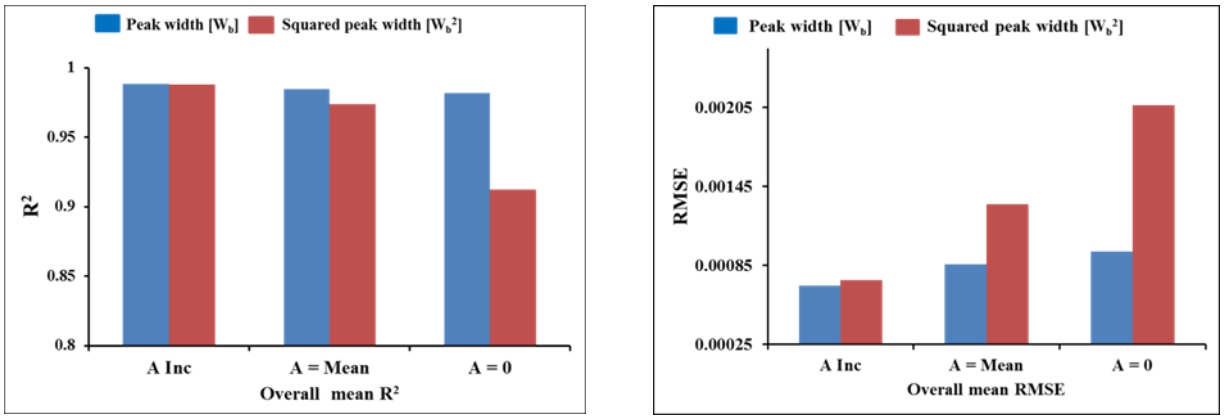


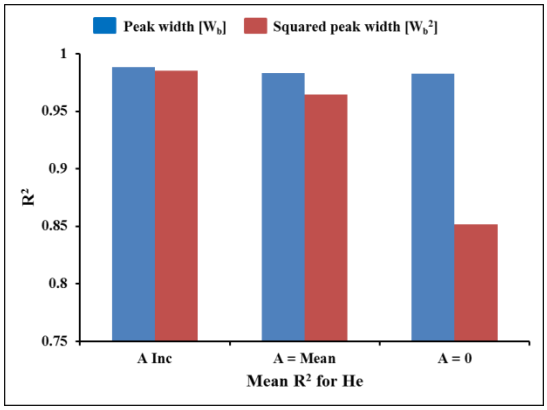
Figure 8: Overall mean R^2 and RMSE of models obtained with normal peak width and squared peak width.

The corresponding RMSE for the two ways of modeling the function was also evaluated. Similarly the RMSE value is lower when the A-term is included than when A is set to its mean value or removed from the model for both peak width and squared peak width at all temperature gradients and with all carrier gases used (Table 6 and Figure 9 to Figure 11). In all cases where A-term is at its mean value or excluded from the models the value of RMSE are lower for models obtained with peak width than for models obtained with squared peak width. And in most cases it is also lower in value for normal peak width than squared peak width models when the A-term is included. The overall mean values of R^2 and RMSE are compared in Figure 8.

Table 6: Adjusted root mean square error ($RMSE_{adj}$) of models obtained with normal peak width (w_b) and squared peak width (w_b^2) for different column types

	Rate	$RMSE_{Adj}$, A incl			$RMSE_{Adj}$, A = mean			$RMSE_{Adj}$, A = 0		
	(^{0}C/min)	w_b		w_b^2	w_b		w_b^2	w_b		w_b^2
DB-5 30 m, 0.25 mm, 0.1 µm										
Helium	2	0.00051	<	0.00055	0.00114	<	0.00311	0.00055	<	0.00249
	4	0.00063	<	0.00071	0.00092	<	0.00188	0.00082	<	0.00317
	6	0.00069	<	0.00107	0.00073	<	0.00127	0.00132	<	0.00416
Hydrogen	2	0.00050	<	0.00055	0.00058	<	0.00108	0.00074	<	0.00174
	4	0.00061	=	0.00061	0.00067	<	0.00088	0.00079	<	0.00188
	6	0.00074	=	0.00074	0.00082	<	0.00087	0.00084	<	0.00217
Nitrogen	2	0.00042	>	0.00041	0.00069	<	0.00127	0.00068	<	0.00076
	4	0.00056	=	0.00056	0.00061	<	0.00079	0.00057	<	0.00124
	6	0.00081	<	0.00083	0.00088	<	0.00091	0.00082	<	0.00157
DB-5 30 m, 0.25 mm, 0.25 µm										
Helium	2	0.00063	<	0.00065	0.00141	<	0.00361	0.00065	<	0.00277
	4	0.00086	<	0.00106	0.00115	<	0.00228	0.00107	<	0.00369
	6	0.00109	<	0.00146	0.00116	<	0.00172	0.00166	<	0.00485
Hydrogen	2	0.00074	>	0.00072	0.00090	<	0.00144	0.00101	<	0.00194
	4	0.00102	=	0.00102	0.00111	<	0.00135	0.00115	<	0.00224
	6	0.00102	>	0.00099	0.00109	<	0.00114	0.00105	<	0.00264
Nitrogen	2	0.00051	=	0.00051	0.00079	<	0.00141	0.00077	<	0.00092
	4	0.00087	=	0.00087	0.00093	<	0.00111	0.00088	<	0.00147
	6	0.00085	<	0.00086	0.00092	<	0.00094	0.00085	<	0.00175
BPX70 30 m, 0.25 mm, 0.22 µm										
Helium	1	0.00056	=	0.00056	0.00071	<	0.00109	0.00069	<	0.00138
	2	0.00059	<	0.00061	0.00068	<	0.00085	0.00065	<	0.00160
	3	0.00075	<	0.00076	0.00086	<	0.00093	0.00080	<	0.00185
Hydrogen	1	0.00060	<	0.00061	0.00075	<	0.00125	0.00167	<	0.00174
	2	0.00068	>	0.00067	0.00084	<	0.00099	0.00163	<	0.00207
	3	0.00082	>	0.00081	0.00092	<	0.00108	0.00176	<	0.00227
Nitrogen	1	0.00054	=	0.00054	0.00061	<	0.00088	0.00078	<	0.00106
	2	0.00058	>	0.00055	0.00064	<	0.00076	0.00083	<	0.00108
	3	0.00063	=	0.00063	0.00069	<	0.00070	0.00076	<	0.00135
Overall mean		0.00070	<	0.00074	0.00086	<	0.00132	0.00096	<	0.00207
Mean with He		0.00070	<	0.00083	0.00097	<	0.00186	0.00091	<	0.00288
Mean with H_2		0.00075	=	0.00075	0.00085	<	0.00112	0.00118	<	0.00208
Mean with N_2		0.00064	=	0.00064	0.00075	<	0.00097	0.00077	<	0.00124

The decrease in R^2 and the increase in RMSE for models of squared peak width is significant. This indicates that the A-term is important for these models. There is only a minor increase in errors by using the mean A-term or excluding the A-term for models with normal peak width. This is in agreement also to the theory of open tubular column where A-term should not be present in the van Deemter equation. Moreover simpler models are always preferred.



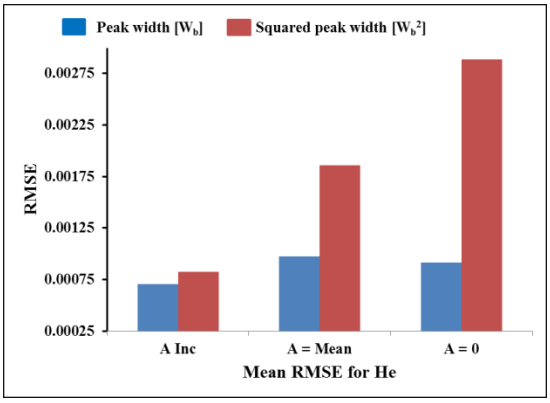
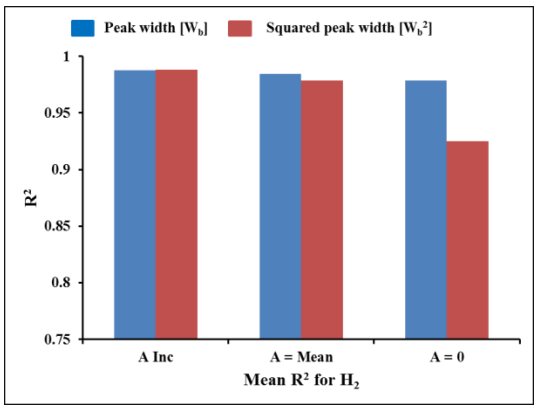


Figure 9: Mean R^2 and RMSE of models obtained with normal peak width and squared peak width using helium as a carrier gas.



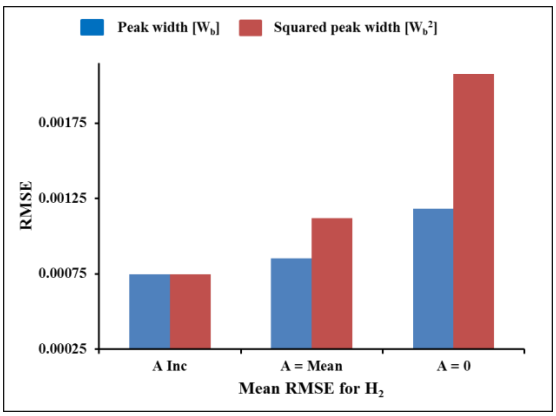
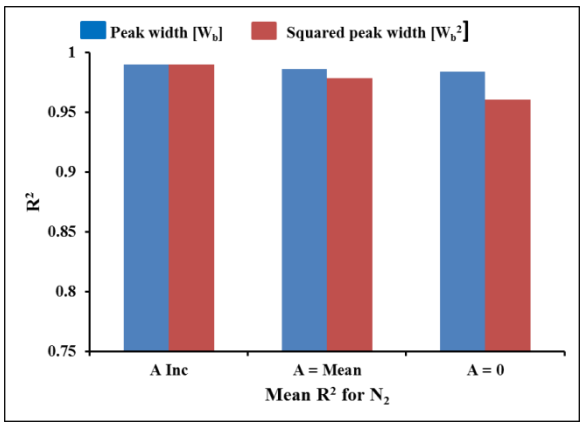


Figure 10: Mean R^2 and RMSE of models obtained with normal peak width and squared peak width using hydrogen as a carrier gas.



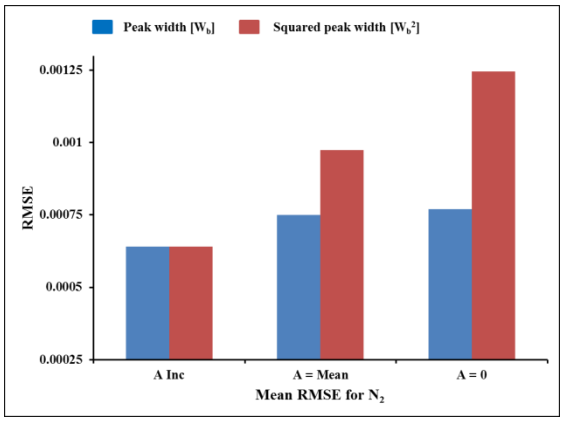


Figure 11: Mean R^2 and RMSE of models obtained with normal peak width and squared peak width using nitrogen as a carrier gas.

From this experiment it is concluded that normal peak width should be used to model the functions in van Deemter since in most cases it gives a better model than when squared peak

width is used. In addition when normal peak width is used excluding the A-term does not significantly affect the model and this agrees with the theory of open tubular column.

3.2. Response surface equations for peak width (w_b)

As explained in the preceding section peak widths in retention index units can be used to model the van Deemter equation in temperature programed GC. The van Deemter equation expresses the dependence of inverse of efficiency on carrier gas velocity. But efficiency is also affected by the rate of temperature employed in temperature programmed GC. Thus it is important to account for the temperature effect, which can be done by expanding the ordinary van Deemter equation to the response surface equation for peak width of the type given in Equation 20.

Since the interacting factors may vary and there is no theoretical basis that explains which of the terms that are significant in Equation 20 it is an important step to evaluate which term is more significant than the others by adding and removing them. Thus evaluation of the significant terms is carried out for models obtained from the chromatographic experiments conducted on DB-5, BPX70 and IL61 columns using helium, hydrogen and nitrogen as mobile phases. The RMSE after excluding different terms one at a time following backward elimination procedure was used to decide on the significance of each term. A model with low RMSE and low number of terms is preferred. A backward elimination procedure was followed since evaluating all possible combinations of the terms is practically difficult because of many possible combinations and many experiments to be evaluated. The backward elimination was performed by starting with a full model with all seven terms in Equation 20 and calculating the RMSE. Thereafter all possible models with 6 terms was evaluated (a,d,e,f, or g kept out) and the model with the lowest RMSE was used as new basis for models with five terms. The process was continued down to three terms. The b and c terms were never deleted since these are needed to describe the dependence of the model on carrier gas velocity. Finally the insignificant terms were those terms that were removed before the RMSE starts to increase sharply and the rest of the terms were considered significant for the function and used to define the model.

In all of the cases the g-term disappeared in the first or the second round while the other terms disappeared earlier or later depending on the column type and/or the mobile phase used (Table 10). For example, the a-term disappeared in the second round for DB5 columns and in the first round for BPX70 60 m column when the mobile phase was helium. In all of the cases

the terms b, c and e are the terms that remained at the end of the backward elimination processes resulting in the simplest model that contains both variables of velocity and temperature gradients.

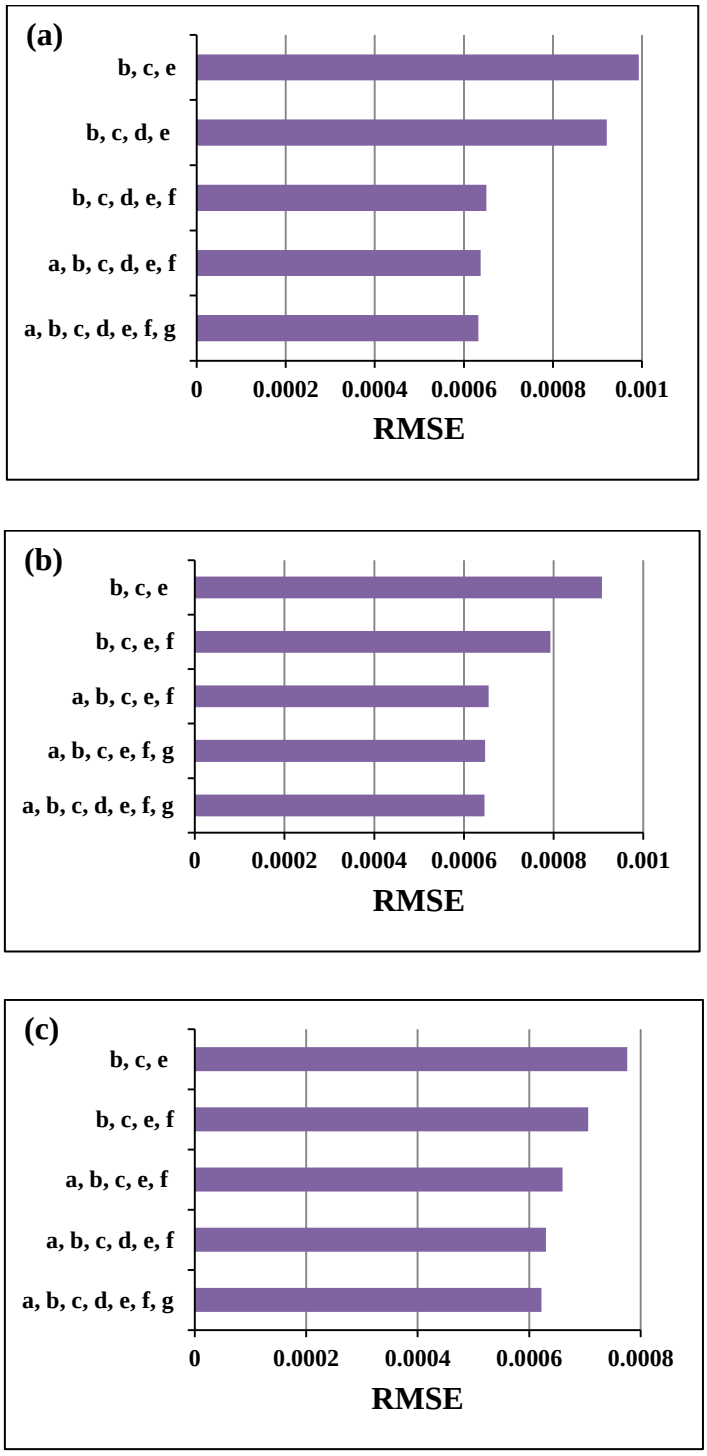


Figure 12: Mean RMSE of best models with 3 to 7 terms (a) helium (b) hydrogen and (c) nitrogen as carrier gases (column: DB-5 30 m, 0.25 mm, 0.1 µm).

Elimination of the first two terms does not show a significant increase in RMSE of the resulting models while at the time the third term is removed the RMSE shows significant increase in all the cases where helium or hydrogen is used as a mobile phase. When nitrogen is used though there is a higher increase in RMSE when the third term is removed the increase is not so significant as compared to when helium or hydrogen is used (Figure 12). Accordingly the a and g-terms when helium is used as a mobile phase; the g and d-terms when hydrogen or nitrogen is used as a mobile phase appears to be insignificant terms that can be excluded from the VD+Int. models. Therefore, the three response surface equations that should be used can be expressed as follows:

Helium: $w = \frac{\mathrm{b}}{u} + \mathrm{c} \cdot u + \mathrm{d} \cdot i + \mathrm{e}\frac{i}{u} + \mathrm{f} \cdot i \cdot u$

Hydrogen: $w = \mathrm{a} + \frac{\mathrm{b}}{u} + \mathrm{c} \cdot u + \mathrm{e}\frac{i}{u} + \mathrm{f} \cdot i \cdot u$

Nitrogen: $= \mathrm{a} + \frac{\mathrm{b}}{u} + \mathrm{c} \cdot u + \mathrm{e}\frac{i}{u} + \mathrm{f} \cdot i \cdot u$

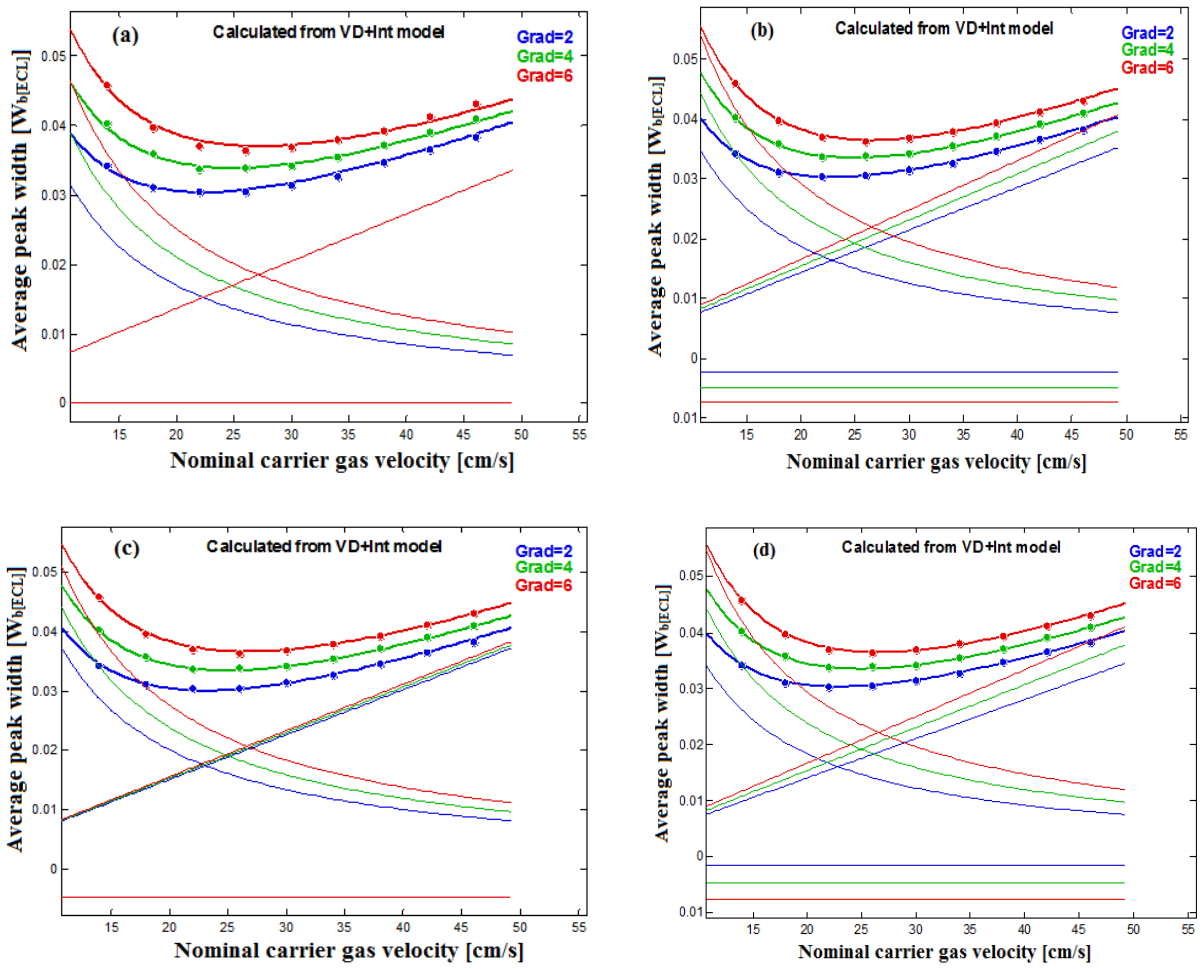


Figure 13: The VD models calculated from VD+Int models with terms included (a) b, c and e (b) b, c, d, e and f (c) a, b, c, e and f (d) a, b, c, d, e and f (column: DB-5 30 m, 0.25 mm, 0.1 µm, helium as a carrier gas).

Consequently five terms are always required for the best accuracy of the models as shown in the three equations above for helium, hydrogen and nitrogen. Four of the five terms required are b, c, e and f; which are common to all equations. The fifth term alternates between a-term and d-term. But these two terms are much related in such a way that the d-term in VD+Int model (the case of helium) indicates the presence of a significant A-term in VD models, which changes linearly with temperature gradient (Figure 13b). Alternately the a-term in VD+Int model (the case of hydrogen and nitrogen) indicates the presence of a significant A-term in VD models, which is identical for all models (Figure 13c). Therefore a model with a, b, c, e and f-terms was compared to models with b, c, d, e and f-terms in all the experiments conducted with the objective of considering the choice of one model over the other. But the difference in RMSE was high and was not encouraging to replace one alternative model by the other. The possibility of including both a-term and d-term to build models of six terms

was also evaluated. However, the gain from reduction in RMSE with six term models was minor compared to the complexity associated to the model built with one more term included. Thus, the five term equations stated above are preferable to represent the response function that best describe the effect of carrier gas velocity and the interacting factor temperature rate on peak width.

3.3. Evaluation of experimental designs

Once the response surface function is established then experiments at different levels of the involved factors can be conducted and evaluated for optimization of the chromatographic separation process. Basically in this study experiments were conducted at three levels of temperature rate and nine levels of carrier gas velocity for all columns and carrier gases studied. This is a 9 x 3 design that involves all points in the experiment conducted. It is a costly experiment both by time and resource used to conduct such number of experiments. So there is no question for the need of experimental designs that require less number of experiments and give comparable result with the full design. Thus in addition to the full 9 x 3 design experimental designs with less number of experiments such as 5 x 3, 3 x 3, a skewed 3 x 3 design and the Doehlert design were evaluated (Figure 14). Some experimental designs, like central composite design, was not utilized here because it requires more levels on the temperature rate.

The experimental designs that are applied to this study have some advantages and disadvantages. The 5 x 3 design with 15 experiments has nearly half the experiments as the full 9 x 3 design, which has 27 experiments (Figure 14 a and b). Thus it may give a similar result as the full design while decreasing the number of experiments by nearly half. But still there are experimental designs that use a less number of experiments while giving a good result. One of these designs is Doehlert design which uses only 7 experiments (Figure 14e). It also avoids the combination of extreme experiments like high and low level combination of the carrier gas velocity and temperature rate. However, with the experimental domain used in this work, the optimal velocity at the lowest temperature gradient is sometimes found slightly outside the design space of the Doehlert Design. The 3 x 3 design with 9 experiments is also one of the experimental designs with few experiments (Figure 14c). Nevertheless it includes low and high levels combination of the factors which may lead to the domination of the results from extreme combinations. It is also a problem when the column has low temperature limit to conduct experiments at higher temperature required for lower carrier gas velocity. This problem is avoided by the skewed 3 x 3 design, while keeping the same number of

experiments as the 3 x 3 design (Figure 14d). Moreover, apart from the full design it is the design that keeps all the 9 levels of the carrier gas velocity.

The comparison of the different designs in this study are demonstrated using experiments conducted on the DB-5 and IL61 columns that were run at temperature gradients of 2, 4 and 6 ^{0}C with the three carrier gases. An experiment using nitrogen as carrier gas on IL61 was not conducted because of column damage after the experiments with helium and hydrogen. The optimal velocity and the peak width at all the three temperature gradients used are determined using the full 9 x 3 design, and other designs with less number of experiments are compared against this. In addition, RMSE of possible validation points were applied for the evaluation (Table 7 and Table 8).

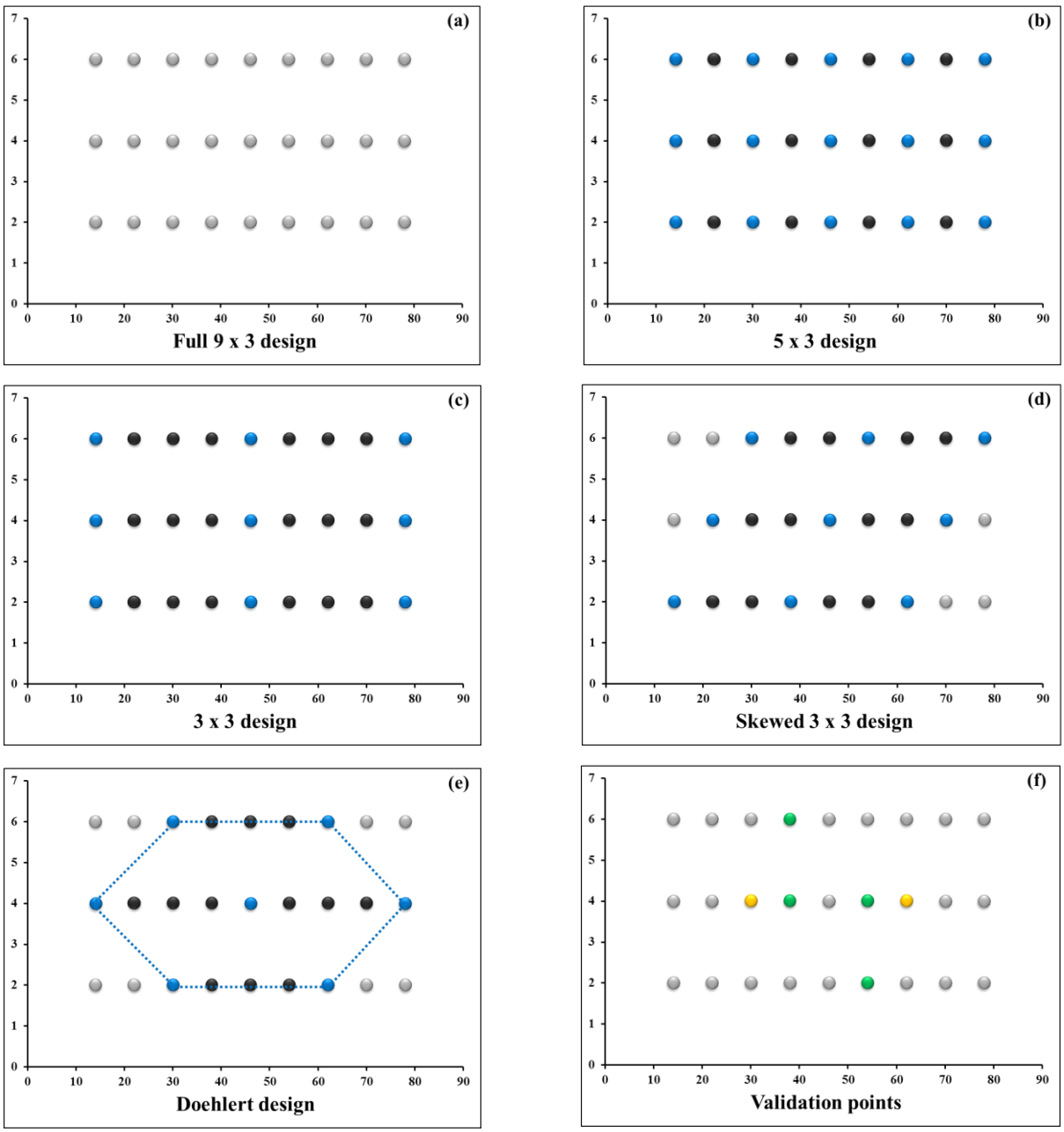


Figure 14: Types of experimental designs used for optimization of efficiency (a-e) and common validation points (f). The blue and the black points are calibration and validation points respectively. The green points are common validation points for all designs whereas yellow and green points together are common validation points for Doehlert, 3 x 3 and skewed 3 x 3 designs.

Response surface plots for average peak width (inverse of efficiency) and corresponding calculated VD models from VD+Int models for all temperature gradient levels are obtained with the different designs employed (Figure 15 to Figure 17).

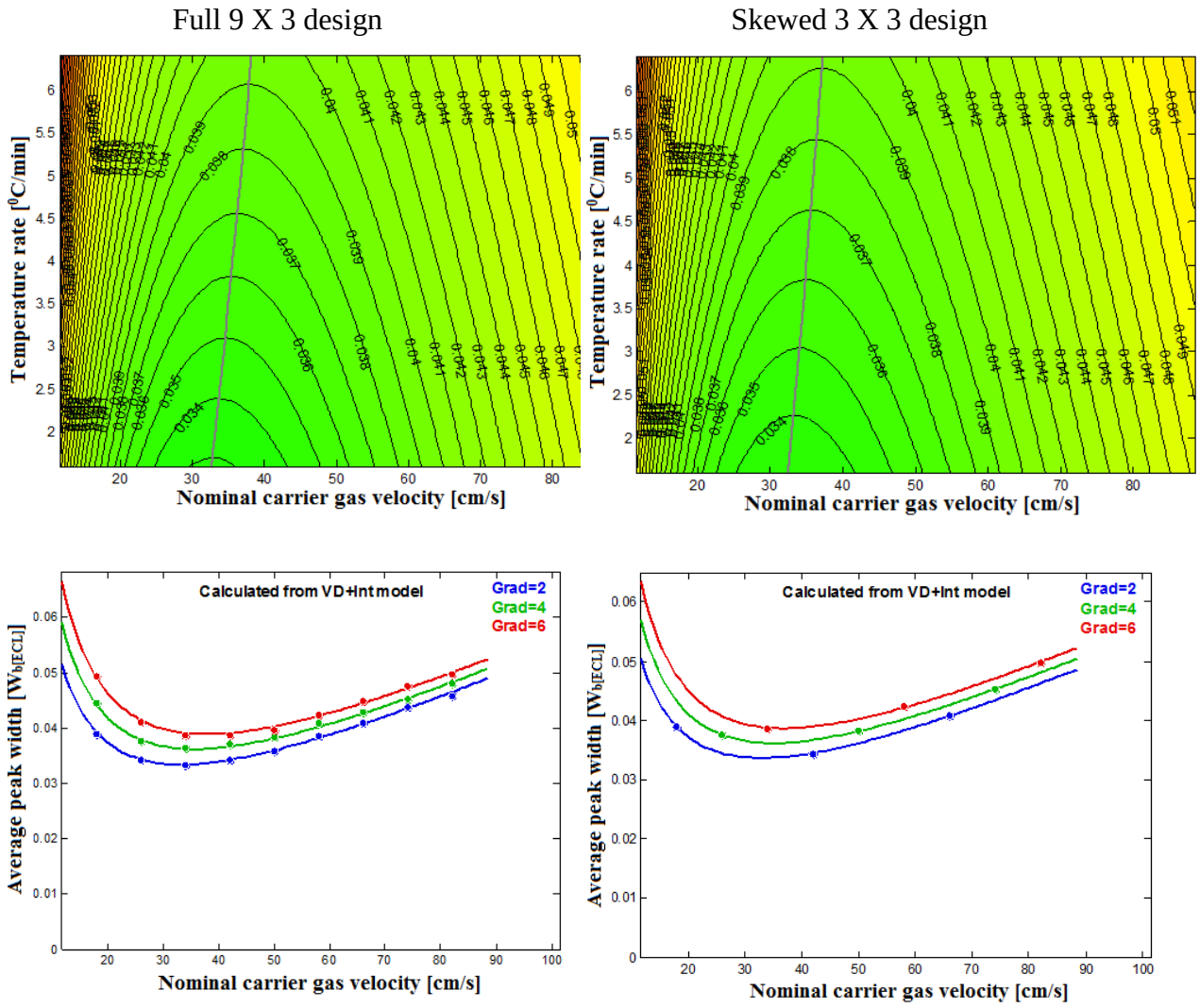


Figure 15: Surface plot for inverse of efficiency (average peak width) and corresponding VD model calculated from VD+Int model for all temperature gradient levels (model from full 9 x 3 design: R^2=0.9664 and RMSE = 0.00094; model from skewed 3 x 3 design R^2=0.9784 and RMSE = 0.00087) (column: DB-5 30 m, 0.25 mm, 0.25 µm; hydrogen as a carrier gas).

The grey line crossing down the plots is the optimal velocity line representing the minimum peak width or the maximum efficiency that can be achieved at a given temperature rate (Equation 21). The VD models are calculated by inserting values for *i* in the VD+Int models. It can be observed from the plots that the peak width increases as temperature gradient increases. This means efficiency decreases with temperature gradients. The optimum velocity is observed to increase with temperature gradient.

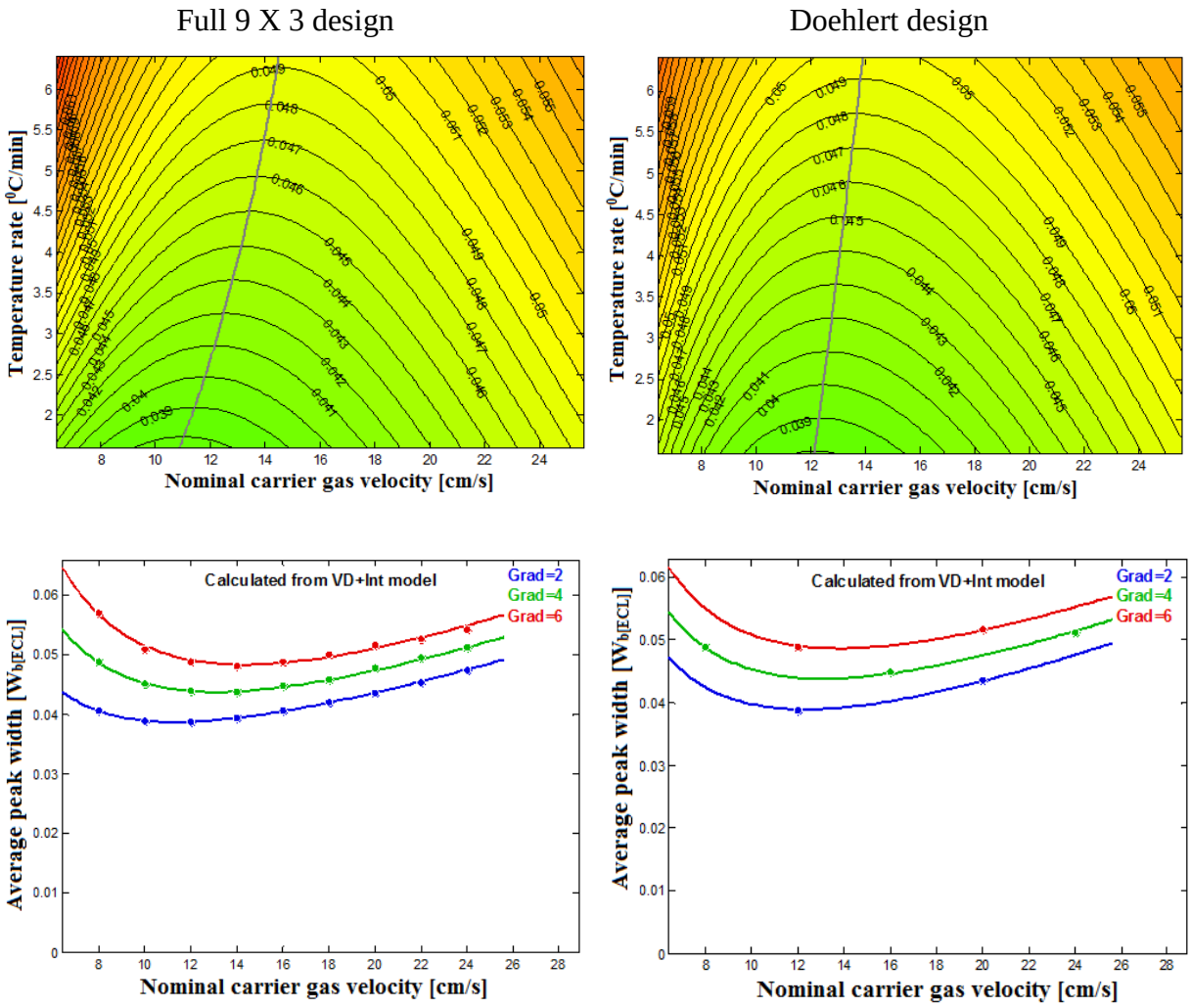


Figure 16: Surface plot for inverse of efficiency (average peak width) and corresponding VD model calculated from VD+Int model for all temperature gradient levels (model from full 9 x 3 design: R^2=0.9777 and RMSE = 0.00080; model from Doehlert design: R^2=0.9910 and RMSE = 0.00071) (column: DB-5 30 m, 0.25 mm, 0.25 µm; nitrogen as a carrier gas).

In general the plots for the full 9 x 3 design and the other designs with reduced number of experiments are similar. But some differences are worth noting, for instance in Figure 16 there is a curvature in the optimum velocity line for the 9 x 3 design but not for Doehlert design. This results in large difference in optimum velocity at lower gradients.

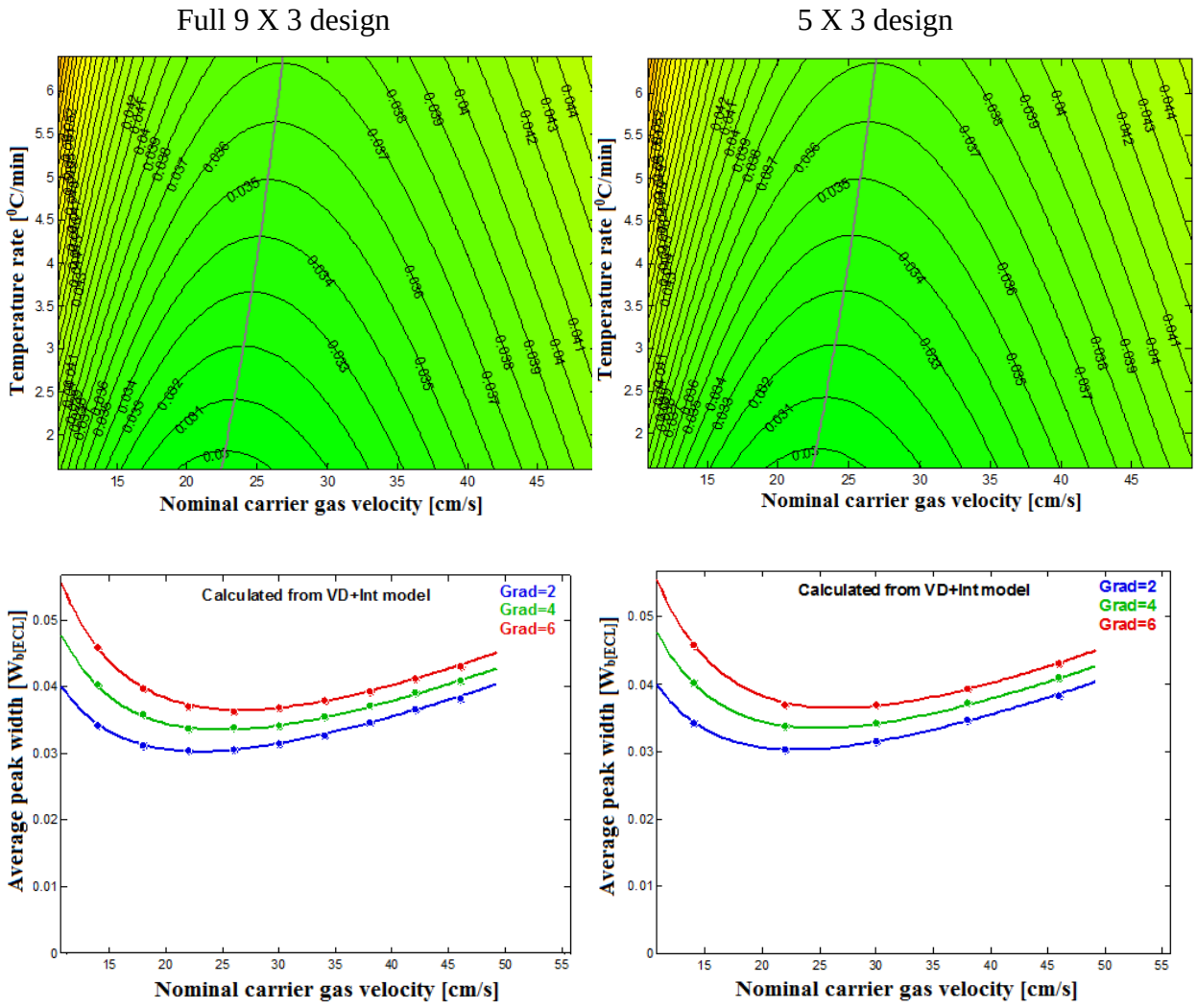


Figure 17: Surface plot for inverse of efficiency (average peak width) and corresponding VD model calculated from VD+Int model for all temperature gradient levels (model from full 9 x 3 design: R^2=0.9806 and RMSE = 0.00060; model from 5 x 3 design R^2=0.9857 and RMSE = 0.00061) (column: DB-5 30 m, 0.25 mm, 0.1 µm; helium as a carrier gas).

The comparison of the designs with reduced number of experiments to the full 9 x 3 design was made with respect to optimum velocity and peak width (Table 7). In general the values of the optimum velocities obtained by other designs have small differences from the value obtained from the full 9 x 3 design. The absolute values of the deviations in percent of the optimum velocities show that the maximum deviations of 1.05%, 1.96%, 4.84% and 8.65% for 5 x 3, 3 x 3, skewed 3 x 3 and Doehlert designs respectively. This indicates that 5 x 3 design has the best performance in the optimum velocity measurement compared to the other designs. The smaller difference obtained with the 5 x 3 design from the full design is related

to its higher number of experiments compared to the rest of the designs with reduced number of experiments. The 3 x 3 design has also better performance than the skewed 3 x 3 and Doehlert design. The reason why the skewed 3 x 3 design has lower performance than the 3 x 3 design might come from the fact that the skewed 3 x 3 design even though it includes all levels from the velocity factor it only includes one experiment from each level which might lead to inaccuracy of measurements. Doehlert design has less performance than all other designs. This is because in Doehlert design optimum velocities at lower temperature lays outside the design space in this experimental set up due to the nature of the Doehlert design, which is also evidenced by the fact that the maximum deviation is at lower temperature rate with the Doehlert design. Nevertheless the deviations from the full design by all the other designs are not very high. The mean absolute deviation in percent of the optimum velocities from the full design are 0.35%, 0.65%, 1.34% and 1.67% for 5 x 3, 3 x 3, Doehlert and skewed 3 x 3 designs. Carrier gas velocities are usually given without decimal points. Which means a difference in velocity of +/- 1 cm/s has limited effect whereas the maximum difference observed in this study was only 0.98 cm/s (Table 7). All the designs may therefore be suitable if the purpose is to find an approximate value for the optimal carrier gas velocity.

Table 7: Comparison of the experimental designs with fewer numbers of experiments to the full 9 x 3 design. Difference in values from the full design is given for the other types of designs with ‘‘-’’ indicting the value is less otherwise more by that amount than that of the full design.

Type of design	Gradient 2		Gradient 4		Gradient 6	
	Optimal velocity	Minimum w_b	Optimal velocity	Minimum w_b	Optimal velocity	Minimum w_b
DB-5 30 m, 0.25 mm, 0.1 µm (Helium)						
9X3 (Full design)	**22.92**	**0.03031**	**24.94**	**0.03352**	**26.56**	**0.03653**
5x3	0.00	-0.00001	0.03	-0.00002	0.07	-0.00005
Doehlert	0.16	0.00012	0.06	-0.00006	-0.05	-0.00023
Skewed 3x3	-0.20	0.00003	-0.05	0.00002	0.07	-0.00003
3x3	0.02	-0.00004	0.09	0.00005	0.16	0.00012
DB-5 30 m, 0.25 mm, 0.1 µm (Hydrogen)						
9X3 (Full design)	**31.80**	**0.02899**	**34.52**	**0.03179**	**36.93**	**0.03444**
5x3	0.10	0.00001	0.11	0.00002	0.13	0.00001
Doehlert	0.55	0.00028	0.26	0.00001	0.03	-0.00023
Skewed 3x3	-0.25	0.00015	-0.61	-0.00005	-0.92	-0.00022
3x3	0.03	0.00008	0.11	0.00017	0.15	0.00026
DB-5 30 m, 0.25 mm, 0.1 µm (Nitrogen)						
9X3 (Full design)	**11.29**	**0.03377**	**12.91**	**0.03885**	**14.27**	**0.04347**
5x3	0.00	-0.00006	0.07	0.00005	0.15	0.00012
Doehlert	0.44	0.00024	0.08	-0.00012	-0.20	-0.00033
Skewed 3x3	-0.48	0.00025	-0.57	0.00008	-0.69	0.00000
3x3	-0.14	0.00003	0.08	0.00031	0.26	0.00048
DB-5 30 m, 0.25 mm, 0.25 µm (Helium)						
9X3 (Full design)	**23.82**	**0.03474**	**25.81**	**0.03801**	**27.41**	**0.04106**
5x3	-0.02	-0.00004	-0.03	0.00004	-0.03	0.00012
Doehlert	0.21	0.00023	-0.13	0.00007	-0.40	-0.00001
Skewed 3x3	-0.13	-0.00002	0.03	-0.00009	0.16	-0.00019
3x3	-0.06	-0.00008	0.00	0.00014	0.08	0.00034
DB-5 30 m, 0.25 mm, 0.25 µm (Hydrogen)						
9X3 (Full design)	**33.20**	**0.03343**	**35.62**	**0.03624**	**37.81**	**0.03891**
5x3	0.16	-0.00009	0.20	-0.00005	0.22	-0.00001
Doehlert	0.72	0.0001	0.10	-0.00009	-0.46	-0.00022
Skewed 3x3	-0.27	0.00022	-0.63	-0.00003	-0.96	-0.00024
3x3	0.22	-0.00016	0.34	-0.00009	0.44	-0.00004
DB-5 30 m, 0.25 mm, 0.25 µm (Nitrogen)						
9X3 (Full design)	**11.33**	**0.03875**	**13.08**	**0.04382**	**14.31**	**0.04840**
5x3	0.01	0.00002	0.05	0.00011	0.10	0.00017
Doehlert	0.98	0.00019	-0.01	0.00003	-0.52	0.00024
Skewed 3x3	0.05	-0.00004	-0.42	0.00007	-0.46	0.00033
3x3	-0.04	0.00008	0.11	0.00026	0.28	0.00035
IL61 30 m, 0.25 mm, 0.2 µm (Helium)						
9X3 (Full design)	**23.62**	**0.04282**	**25.99**	**0.04841**	**27.92**	**0.05364**
5x3	-0.10	-0.00009	0.01	-0.00009	0.09	-0.00012
Doehlert	0.08	0.00003	-0.06	-0.0001	-0.22	-0.00019
Skewed 3x3	-0.24	0.00012	-0.21	0.00012	-0.10	0.00011
3x3	-0.09	-0.00015	0.08	-0.00012	0.21	-0.00014
IL61 30 m, 0.25 mm, 0.2 µm (Hydrogen)						
9X3 (Full design)	**32.63**	**0.04120**	**35.63**	**0.04578**	**38.21**	**0.05013**
5x3	0.01	0.00012	0.19	0.00015	0.36	0.00015
Doehlert	0.48	0.00022	0.14	-0.00006	-0.13	-0.0003
Skewed 3x3	-0.34	0.00033	-0.60	0.00004	-0.83	-0.00022
3x3	-0.03	0.00023	0.25	0.00036	0.49	0.00043

The deviations in peak width calculated using all other designs than the full 9 x 3 designs are small. The absolute values of the maximum deviations in percent of the peak widths are 0.35%, 0.80%, 0.97% and 1.10% whereas the mean absolute values of deviations of peak width are 0.17%, 0.32%, 0.40% and 0.46% both for 5 x 3, skewed 3 x 3, Doehlert and 3 x 3 designs respectively. Even though the results indicate that 5 x 3 design is the best design by performance, the deviations by other designs with reduced number of experiments are fairly similar. The peak width deviations are less than 0.0005 ECL (Table 7). With minimum peak width measured equal to 0.02899 ECL in this study where two significant digits can be considered acceptable the difference from the full design of the other types of designs in the third or fourth digit is not important also for practical purpose. The higher deviations are observed at higher gradient which is expected since peak widths are higher and also there is higher inaccuracy of peak width measurements at higher temperature gradients. With respect to the type of carrier gases used less variation in optimum carrier gas velocity as well as peak width is observed when helium is used than when either hydrogen or nitrogen is used (Table 7).

The RMSEP of the experimental designs with less number of experiments than the full 9 x 3 design were determined using validation points for the individual designs. The validation points are the black points in the designs in Figure 14b-e, which are inside the designs but not used for calibration purpose. The RMSEP is calculated using Equation 22 where p is zero since the validation points are independent of calibration points. It is calculated on all individual squared errors (17 FAMEs x number of validation points), contrary to reported RMSE for the response surface models that are average RMSE for each FAME model. The calculation was done manually in Excel from predicted and measured values.

The calculated errors of prediction are given in Table 8. The averages of the RMSEP were calculated and are found to be in the range of 7.22×10^{-4} to 11.11×10^{-4} for all types of columns and carrier gases evaluated. The design that gives a better RMSEP in most of the cases is the 5 x 3 design whereas higher RMSEP is shown by 3 x 3 design. This is due to the difference in the number of experiments in the two designs. However, the skewed 3 x 3 design with the same number of experimental points as the 3 x 3 design gives a better RMSEP. The reason for this may be that the skewed design is not dominated by the low and high value combinations of the temperature gradient and the carrier gas velocity. Doehlert design gives a higher RMSEP than skewed 3 x 3 design. This might be due to the lower number of experiments in Doehlert design. The difference between the best and the worst

RMSEP ranges from 0.7 x 10^{-4} to 1.79 x 10^{-4} which is 6.3 to 24.8%. Therefore a better prediction is dependent on the number of experiments that can be conducted and the nature of the design used. Thus skewed 3 x 3 design is found to be the best design when it is not possible to afford the number of experiments to be conducted in 5 x 3 design.

Table 8: Comparison of RMSE of predicted peak width of different experimental designs.

	Type of design	RMSE ($\cdot 10^{-4}$)	RMSE* ($\cdot 10^{-4}$)	RMSE**($\cdot 10^{-4}$)
DB-5 30 m, 0.25 mm, 0.1 µm				
Helium	5x3	6.69		6.71
	Doehlert	7.18	7.43	6.96
	Skewed 3x3	7.40	7.37	7.38
	3x3	7.59	7.98	7.78
	Mean RMSE	**7.22**	**7.59**	**7.21**
Hydrogen	5x3	7.29		7.88
	Doehlert	8.12	8.10	9.05
	Skewed 3x3	7.97	7.31	7.94
	3x3	8.03	7.54	8.16
	Mean RMSE	**7.85**	**7.65**	**8.26**
Nitrogen	5x3	7.46		7.76
	Doehlert	7.09	6.55	7.07
	Skewed 3x3	6.83	6.42	7.03
	3x3	8.62	8.08	9.24
	Mean RMSE	**7.50**	**7.02**	**7.78**
DB-5 30 m, 0.25 mm, 0.25 µm				
Helium	5x3	9.17		9.25
	Doehlert	10.85	10.64	10.18
	Skewed 3x3	10.53	10.46	10.35
	3x3	9.83	10.12	9.58
	Mean RMSE	**10.10**	**10.41**	**9.84**
Hydrogen	5x3	10.78		11.01
	Doehlert	10.88	11.28	10.97
	Skewed 3x3	10.98	10.77	10.42
	3x3	11.81	12.49	11.78
	Mean RMSE	**11.11**	**11.51**	**11.05**
Nitrogen	5x3	8.69		9.09
	Doehlert	9.60	10.00	10.01
	Skewed 3x3	8.81	9.71	9.89
	3x3	10.07	10.25	10.41
	Mean RMSE	**9.29**	**9.99**	**9.85**
IL61 30 m, 0.25 mm, 0.2 µm				
Helium	5x3	7.75		8.02
	Doehlert	8.45	8.56	8.97
	Skewed 3x3	8.42	8.57	9.03
	3x3	8.38	8.14	8.26
	Mean RMSE	**8.25**	**8.42**	**8.57**
Hydrogen	5x3	7.57		7.17
	Doehlert	7.28	7.64	7.63
	Skewed 3x3	7.56	7.09	6.97
	3x3	8.90	8.07	8.05
	Mean RMSE	**7.83**	**7.60**	**7.46**

RMSE*- common validation point except for 5 x 3 design (Figure 14f – Green plus yellow)

RMSE**- common validation point for all designs (Figure 14f – Green colour)

Common validation points (green points in Figure 14f) for all the designs with less number of experiments than the full design and common validation points for Doehlert, 3 x 3 and skewed 3 x 3 designs (yellow + green points in Figure 14f) were also used to calculate the RMSEP. In both cases the validation points are inside the designs but not used for calibration purpose. The purpose of using common validation points was to compare the different designs using the same validation experiments. For the 5 x 3 design fewer common validation points were available than for the other reduced designs. It is therefore one set of results including the 5x3 design and one without.

The results from common validation points are in a similar range as RMSEP calculated for the individual designs using individual validation points (Table 8). This leads to the same conclusion given above. The results are summarized in Figure 18. The different experiments have different degree of error. To compare the RMSEP the values for each design in a specific experiment was therefore normalized to the mean RMSEP for the experiment. Figure 18 show the average of these values for each design and validation set. Values below 100% mean that the design performs better than the average and values above 100% mean that the design performs worse.

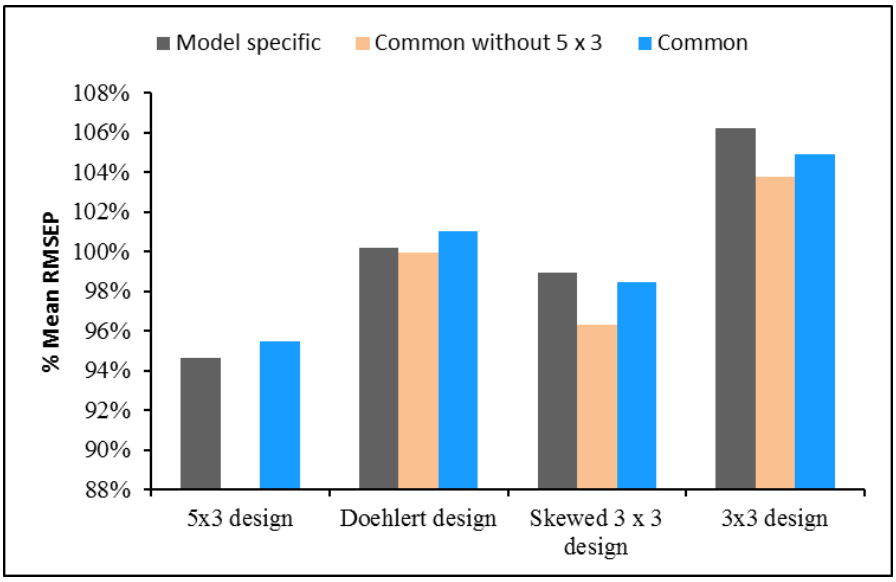


Figure 18: Average RMSEP relative to the mean value for the 8 experiments calculated using model specific, common without 5 x 3 design and common to all designs validation points for different designs employed.

In general irrespective of which set of validation points (model specific, common or common without 5 x 3 design) that is used to evaluate the results the order of performance from the

best to the poorest is: 5 x 3, skewed 3 x 3, Doehlert and 3 x 3 designs. Moreover the RMSEP calculated based on the three different sets of data are quite similar for the same design and experiment (Table 8).

3.4. Efficiency and time of analysis

In a chromatographic analysis temperature rate and carrier gas velocity affects not only efficiency but also the time of analysis, which can be best expressed with the retention time of the last eluting compound. Shortest time of analysis is usually preferred, especially in a routine analysis of samples in a laboratory. There are combinations of carrier gas velocity and temperature rate that minimize the time for a required efficiency, or that maximize the efficiency that can be achieved within a certain amount of time.

In this section a methodology for evaluating the trade-off between time and efficiency is demonstrated using a DB-5 column and three types of carrier gases. The optimization process is demonstrated by using a full 9 x 3 design.

3.4.1. Models of retention time

Before evaluation of the effect of temperature rate and carrier gas velocity on time of analysis as well as finding the time efficiency trade-off there is a need for a model that accurately explain the analysis time in terms of retention time of the last eluting compound. Thus a retention time model was developed using an experiment conducted on 60 m BPX70 column, which was the only experiment that was performed on four levels of temperature rate. Retention time is related to both temperature gradient and carrier gas velocity. Increasing either of the two or both factors decreases retention time. When retention time versus temperature rate and retention time versus carrier gas velocity graphs are plotted the data fits to curves with power function of $y = ax^b$type (Figure 19a and Figure 20a).

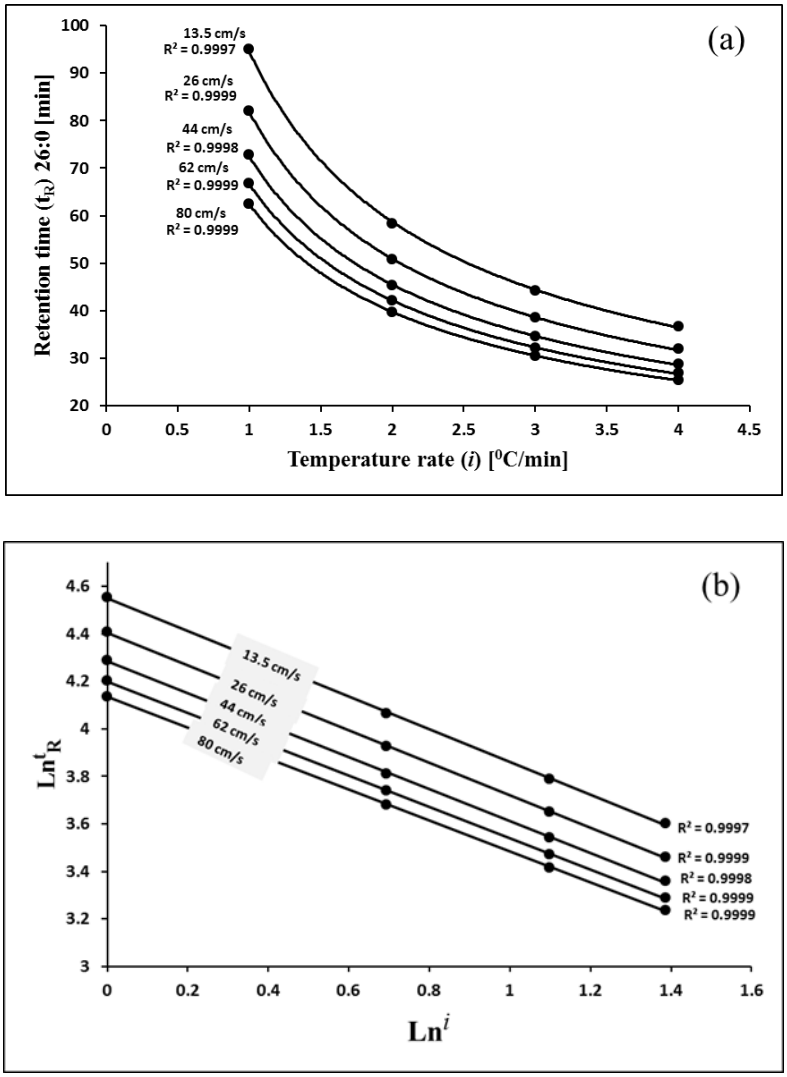
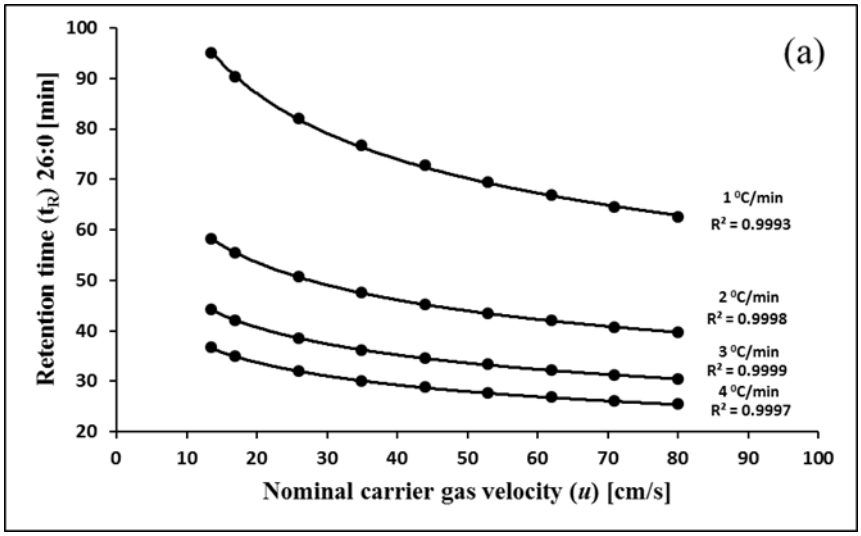


Figure 19: Retention time related to temperature rate in power function (a) and logarithmic function (b) at different carrier gas velocities (column: BPX70 60 m, 0.25 mm, 0.25 µm; hydrogen as a carrier gas).



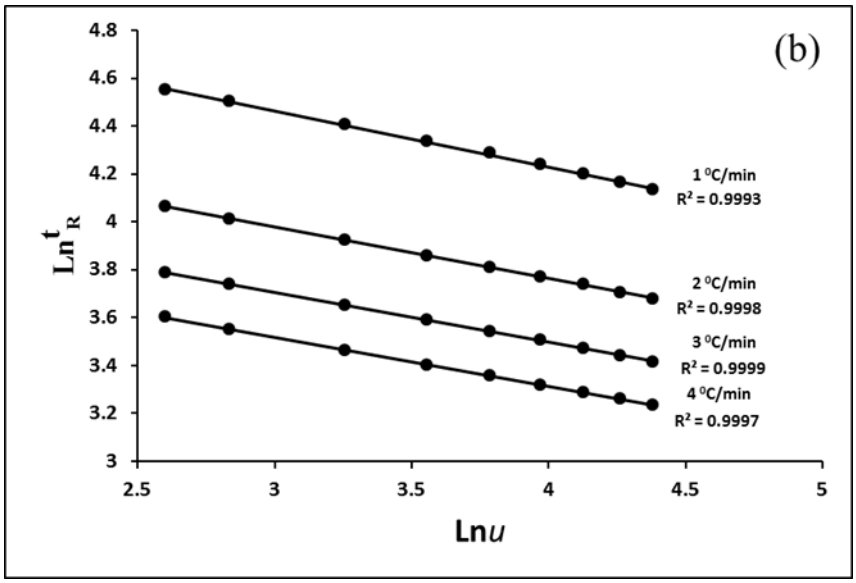


Figure 20: Retention time related to carrier gas velocity in power function (a) and logarithmic function (b) at different temperature rates (column: BPX70 60 m, 0.25 mm, 0.25 µm; hydrogen as a carrier gas).

Taking the logarithm of both sides of the power equation linearize the relationships (Figure 19b and Figure 20b). Thus including all the parameters involved the linear function was modelled by response surface without quadratic terms.

Equation 24: $$\ln t_R = \mathrm{A} + \mathrm{B}\ln u + \mathrm{C}\ln i + \mathrm{D}\ln u \ln i$$

Where t_R is the retention time of the last eluting compound, u is carrier gas velocity, i is temperature gradient and A, B, C and D are coefficients. After the model was created and the

parameters were found the exponent of the equation was taken to achieve a response surface that directly explains the retention time as a function of *u* and *i*. Thus a response surface of the retention time of the last eluting FAME which is the methyl ester of the saturated fatty acid of 26:0 was obtained (Figure 21) in the original form of measurements of the parameters, not in logarithmic form. From the resulting surface plots it can be easily noted that shorter retention time is achieved at higher temperature rate and at higher carrier gas velocity. It is also observed that increasing temperature rate has a larger effect than increasing carrier gas velocity.

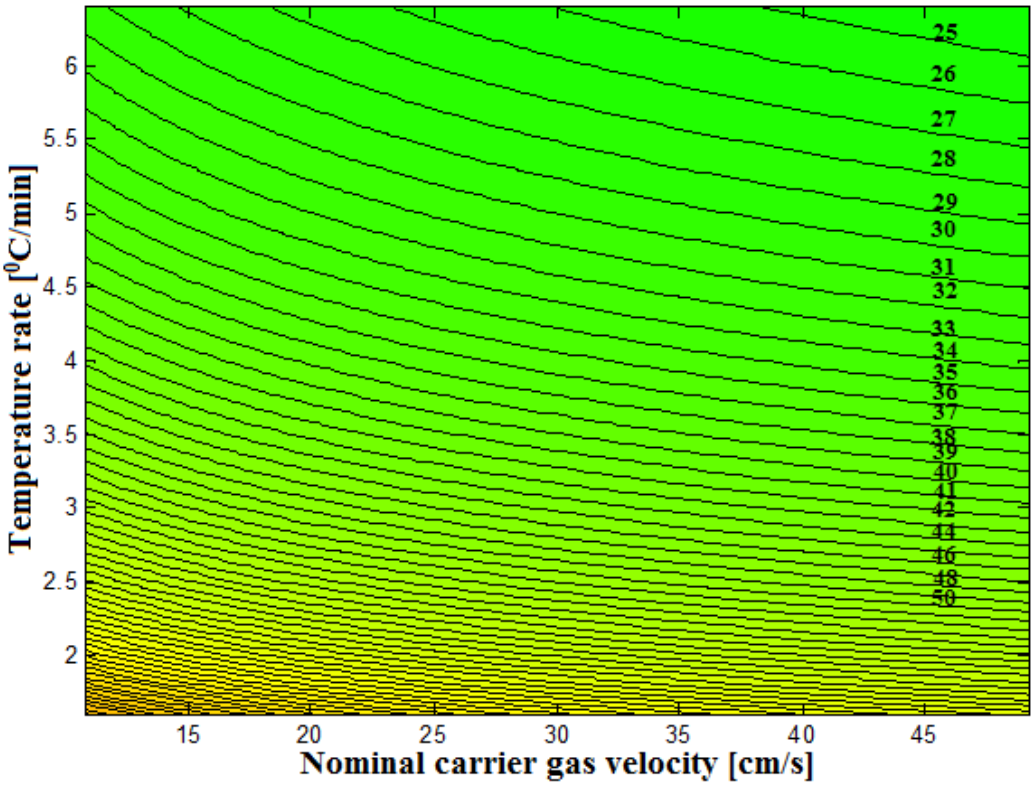


Figure 21: Surface response plot of the retention time in minutes of the last eluting FAME (26:0), R^2 = 0.9999 and RMSE = 0.183 (column: DB-5 30 m, 0.25 mm, 0.1 µm, helium as carrier gas). Numbers on iso-lines represents retention time in minutes

It is always the wish of the analyst to get analytical result in shortest time possible. But shorter retention time may not be obtained without loss in efficiency. So to indicate the shortest possible time that can be achieved we have to evaluate the efficiency at all combinations of the temperature rate and carrier gas velocities used to obtain the retention time model.

3.4.2. Time-efficiency trade-off

To explain the time efficiency trade of it is important first to bring together the response surface for both retention time and efficiency to understand the relation between the two responses. Figure 22 shows the overlap of the efficiency response plot over the retention time response plot. The white isolines are the lines that represent peak widths (inverse of efficiency) from efficiency response plot overlaid on the time response plot from Figure 21, while the white line crossing down the plots is the optimal velocity line representing the minimum peak width or the maximum efficiency that can be achieved at a given temperature rate (Equation 21). From the overlapped surface plots it is easy to notice that the best efficiency is located at lower temperature rate and optimal carrier gas velocity whereas shortest time of analysis is at higher temperature rate and higher carrier gas velocity. This means *if we are willing to wait for longer time of analysis it is possible to get higher efficiency,* or *if we are willing to accept reduced efficiency it is possible to analyse our sample in a shorter period of time*. But longer analysis time as well as lower efficiency is not desirable. Thus we need to make a compromise between time of analysis we afford and the level of efficiency we need.

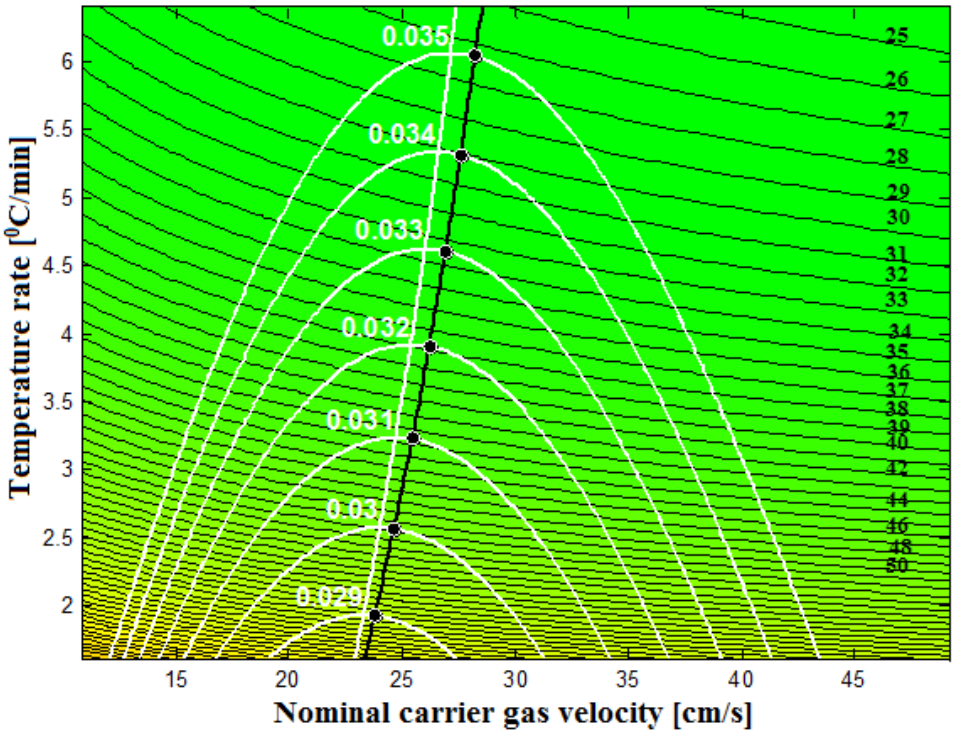


Figure 22: Surface plot for inverse of PPC (R^2 = 9820 and RMSE = 0.00055) overlapped on the surface plot of retention time for last eluting FAME (26:0) (R^2 = 0.9999 and RMSE = 0.183) (column: DB-5 30 m, 0.25 mm, 0.1 µm, helium as carrier gas). Numbers on white iso-lines are w_b in ECL units.

To find out the points on the plot where best time efficiency trade-off can be made an iterative procedure was applied. The procedure follows the iso-lines for the VD+Int model and from the retention time model it finds the conditions that minimize the retention time on the iso-lines. These are the black points in the response surface. A spline function is thereafter fitted to the points and represented by the black line. This line represents the optimal conditions for the time-efficiency trade-off. For any conditions that are not along the black line one can argue that the same efficiency can be achieved with shorter analysis time or that better efficiency can be achieved using the same time (Figure 22 and Figure 23). This line indicates the best conditions of temperature rate and carrier gas velocity for best trade-off. The point along the line to choose is dependent on how much efficiency is needed or how long analysis time one is willing to accept. The velocities along this line are referred to as time optimal velocities (u_{topt}).

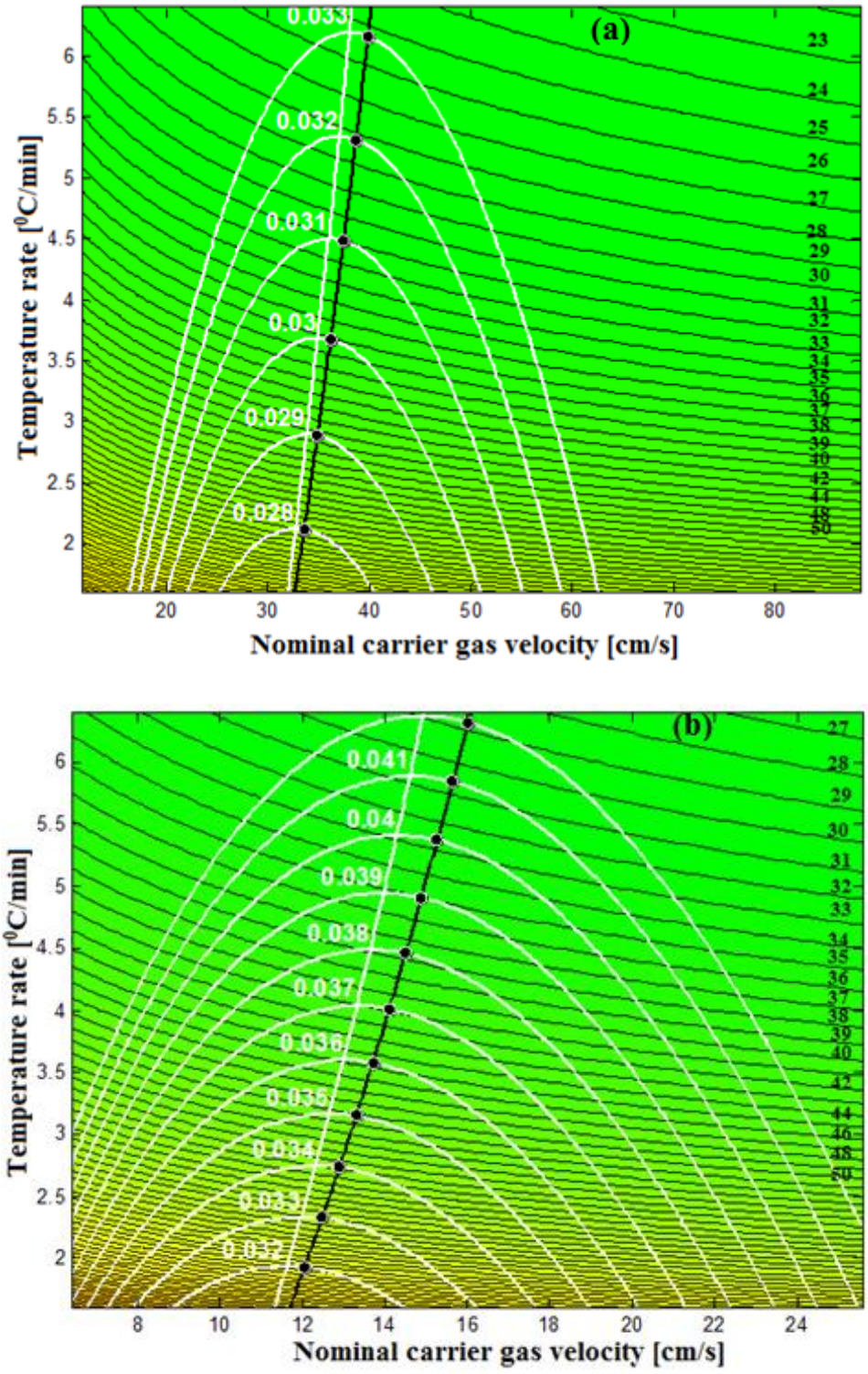


Figure 23: Surface plots showing best line of efficiency/time trade-off (the black line) in (a) hydrogen (for retention time model R^2 = 0.9999 and RMSE = 0.162; for peak width model R^2=0.9809 and RMSE = 0.00061) (b) nitrogen as carrier gases (for retention time model R^2 = 0.9998 and RMSE = 0.221; for peak width model R^2 = 0.9835 and RMSE = 0.00061) (column: DB-5 30 m, 0.25 mm, 0.1 µm).

This line of best time-efficiency trade-off as it can be observed from Figure 22 and Figure 23 is found at slightly higher velocities than the minimum in the van Deemter curves. The line of best efficiency time of analysis trade-off (u_{topt}) lays 2-5%, 3-6% and 3-8% above the minimum in van Deemter curve (u_{opt}) for helium, hydrogen and nitrogen respectively (Table 9 and Appendix Table 11).

Table 9: Optimum carrier gas velocity and efficiency at the minimum point on van Deemter curve and at best time efficiency trade-off point (column: DB-5 30 m, 0.25 mm, 0.1 μm).

Temperature rate (^{0}C/min)	u_{opt} (cm/s)	w_b (opt)	u_{topt} (cm/s)	w_b (topt)	% (u_{opt} - u_{topt})
Helium as a carrier gas					
2.0	23.4	0.0291	23.9	0.0291	2.1
2.5	23.9	0.0299	24.5	0.0299	2.5
3.0	24.5	0.0306	25.2	0.0307	2.9
3.5	25.0	0.0314	25.7	0.0314	2.8
4.0	25.4	0.0321	26.3	0.0321	3.5
4.5	25.9	0.0328	26.8	0.0329	3.5
5.0	26.3	0.0335	27.3	0.0336	3.8
5.5	26.7	0.0342	27.7	0.0343	3.7
6.0	27.1	0.0349	28.2	0.0349	4.1
Hydrogen as a carrier gas					
2.0	32.5	0.0278	33.4	0.0278	2.8
2.5	33.3	0.0285	34.2	0.0285	2.7
3.0	34.0	0.0291	35.1	0.0291	3.2
3.5	34.6	0.0298	35.9	0.0298	3.8
4.0	35.3	0.0304	36.7	0.0304	4.0
4.5	35.9	0.0310	37.5	0.0310	4.5
5.0	36.6	0.0316	38.2	0.0316	4.4
5.5	37.2	0.0322	38.9	0.0322	4.6
6.0	37.8	0.0328	39.6	0.0328	4.8
Nitrogen as a carrier gas					
2.0	11.7	0.0322	12.1	0.0322	3.4
2.5	12.1	0.0334	12.7	0.0334	5.0
3.0	12.5	0.0346	13.2	0.0346	5.6
3.5	12.9	0.0358	13.6	0.0358	5.4
4.0	13.3	0.0369	14.1	0.0370	6.0
4.5	13.7	0.0380	14.5	0.0381	5.8
5.0	14.0	0.0391	15.0	0.0392	7.1
5.5	14.4	0.0402	15.4	0.0403	6.9
6.0	14.7	0.0412	15.8	0.0413	7.5

- u_{opt} -Velocity that gives maximum efficiency at gradient
- w_b(opt) - Inverse efficiency at u_{opt}
- u_{topt} - Optimal velocity at gradient (efficiency/time trade-off)
- w_b(topt) - Inverse efficiency at u_{topt}

From the results obtained it can be concluded that it is important to know in which conditions to operate chromatographic separations so that one can benefit from either the maximum efficiency that can be achieved at a given time or the time that can be gained and still

achieving acceptable efficiency. This can be simply achieved by working on conditions on the optimal time efficiency trade-off line, where the exact point on the line to choose depends on the required efficiency or the analysis time one are willing to accept.

3.5. Evaluation of the carrier gases used

As shown in the surface plots (Figure 22 and Figure 23) using the three carrier gases helium, hydrogen and nitrogen we obtain different ranges of peak width or time of analysis along the optimal conditions. And it is simple to observe that hydrogen as a carrier gas is more efficient (smaller peak width) and faster while nitrogen is less efficient (wider peak width) and slower. Helium is in the middle of the two carrier gases in terms of efficiency as well as time of analysis. While this fact is generally true one may want to know how much efficiency and/or time of analysis will be gained or given off by switching from one carrier gas to the other. Thus comparing u_{topt} from the analysis of FAMEs on the same column at similar temperature gradient but using different carrier gases it is possible to show the loss or gain of the time and efficiency when changing the carrier gas. This is demonstrated by using experiments conducted on DB-5 with results from experiment on helium being at the centre taken as starting point. Thus by changing the carrier gas from helium to either hydrogen or nitrogen the gain or loss in efficiency and/or time are determined (Figure 24 and Figure 25).

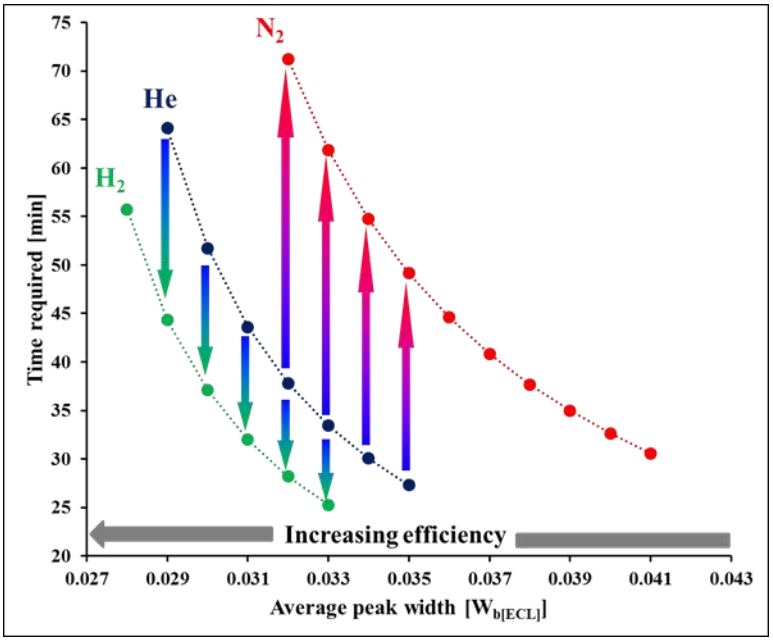


Figure 24: Gain and loss of time in switching carrier gas from helium to hydrogen and from helium to nitrogen respectively while keeping the same efficiency (column: DB-5 30 m, 0.25 mm, 0.1 µm).

At similar efficiency with helium hydrogen as carrier gas reduces the time of analysis by 24-31% whereas there is 80-92% increase in retention time with nitrogen as a carrier gas (Figure 24). This is mainly related to the optimal velocity of the carrier gas. The optimum velocity of hydrogen as a carrier gas is about 1.3 times that of helium while it is nearly 3 times as fast as nitrogen. On the other hand by switching the carrier gas from helium to hydrogen or nitrogen at similar retention time hydrogen gives 6-8% more efficiency than helium, while 12-18% decreases in efficiency with nitrogen is observed (Figure 25).

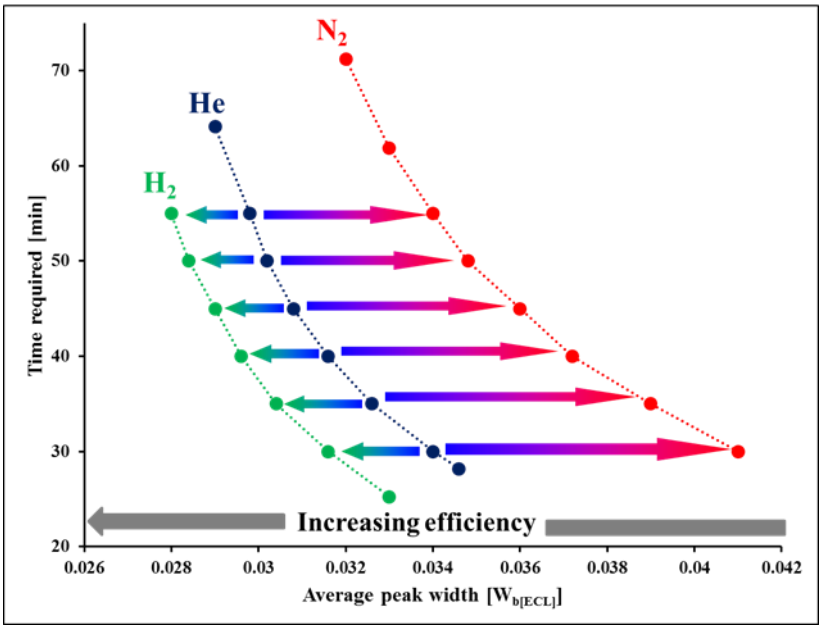


Figure 25: Gain and loss of efficiency in switching carrier gas from helium to hydrogen and from helium to nitrogen respectively while keeping the same retention time (column: DB-5 30 m, 0.25 mm, 0.1 µm).

Some basic differences can be noted from comparison of isothermal GC and temperature programmed GC. In temperature programmed GC there is around 25% reduction in retention time when the carrier gas is changed from helium to hydrogen (Figure 24) while increasing from 20 cm/s (helium) to 40 cm/s (hydrogen) in isothermal GC reduces retention time approximately by half without significant reduction in efficiency (Figure 3). This is because in isothermal GC the retention time is determined by only carrier gas velocity and they are inversely related whereas in temperature programmed GC retention time is determined by carrier gas velocity and temperature rate. At relatively higher velocities the temperature rate has more effect on retention time while at lower velocities increasing velocity has more effect

on reducing retention time. At medium to high velocities the iso-lines for the time in Figure 21 to Figure 23 are almost horizontal.

Van Deemter curves calculated from the equations from a typical temperature programmed GC experiment is shown in Figure 26. The minimum in the van Deemter curve is lower for nitrogen than the other two gases in isothermal GC whereas in temperature programmed GC it is higher for nitrogen. This can be explained from the Purnell equation (Equation 7). In temperature programmed GC increasing temperature gradually decreases the retention factor (*k*) and when the analyte reaches the end of the column *k* is very low. This leads the last factor in the Purnell equation to approach zero and the chromatography to be inefficient. The lower the carrier gas velocity is the more severe the problem is, which is the case with nitrogen.

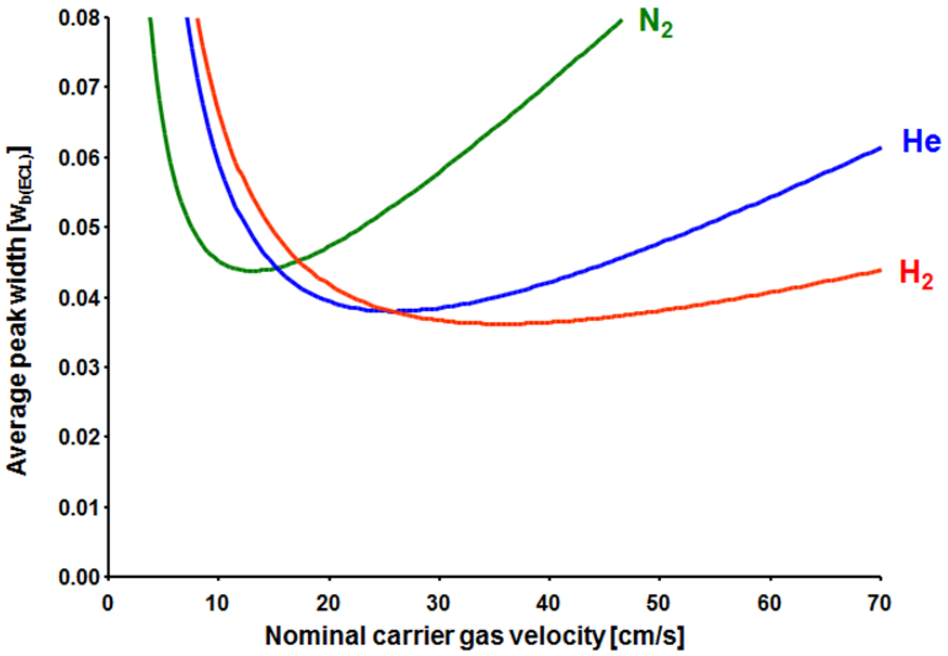


Figure 26: 'van Deemter curve' in temperature programed gas chromatography (column: DB-5 30 m, 0.25 mm, 0.25 µm; 4 ^{0}C).

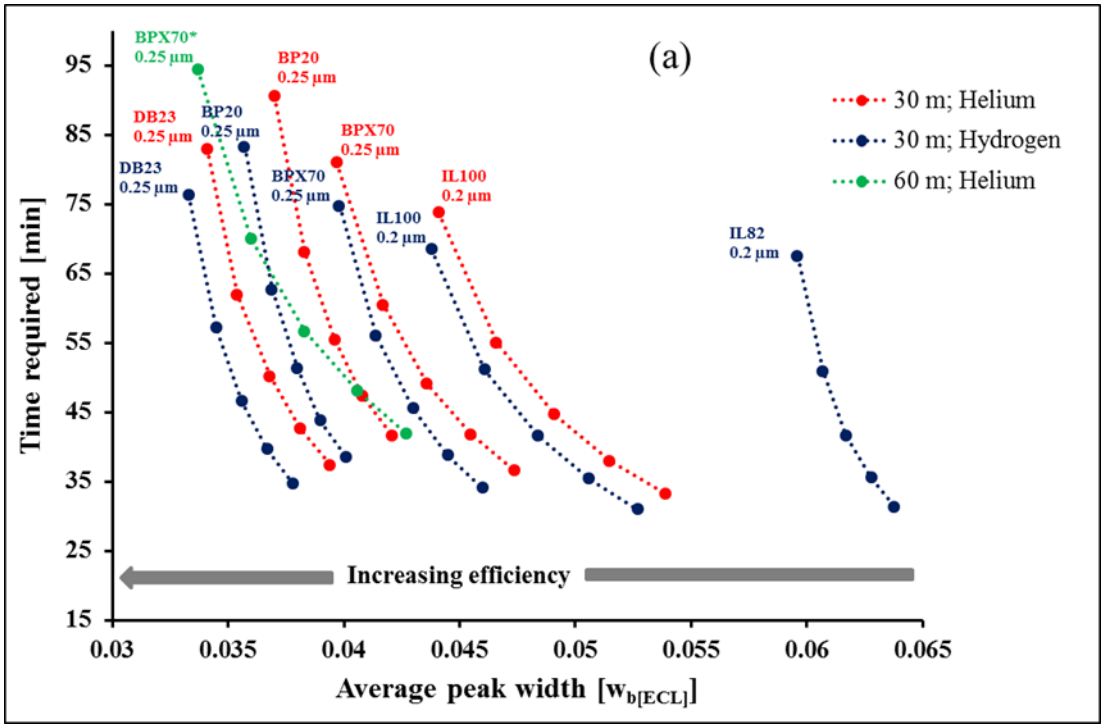
3.6. Comparison of polar columns for FAMEs analysis

Polar chromatographic columns have an important role in the analysis of FAMEs. Some columns are even designed and produced to analyse specifically FAMEs, and they have a selectivity that is optimized for important FAMEs. However, efficiency is equally important and there have been little focus on differences in efficiency between these phases. The selection of column for the analysis of FAMEs samples therefore should depend on the

performance of the columns, both with regard to efficiency and time of analysis that the analyst wants to attain.

The performance of different commonly employed polar columns for the analysis of FAMEs using the three carrier gases were evaluated and compared. Polar columns: DB-23, BP-20, BPX70 and IL100 are selected and compared because these columns have polar stationary phase that have the right selectivity for the FAMEs. In this experiment FAMEs of the long chain fatty acid from 19:0 to 26:0 were excluded due to the asymmetry of the peaks on the most polar phases. The experiments were conducted at temperature gradients of 1, 2 and 3 ^{0}C because most of these columns have too low temperature limit to run the analyses at higher gradients.

The efficiency at optimal velocity using the three carrier gases was evaluated (Figure 27). In general DB-23 and BP-20 are found to be the most efficient columns in the analysis of FAMEs whereas the IL100 column is less efficient. The 30 m BPX70 column has a moderate efficiency. The 60 m BPX70 has high efficiency at longer analysis time but approaches the 30 m BPX70 as the retention time goes down. This shows that to take advantage of the possible high efficiency of long columns one has to use low temperature gradients and accept longer analysis time. The IL82 column run with hydrogen appeared far behind IL100 in efficiency because the column used was an old column (Figure 27a). It is presented here just to demonstrate how efficiency is affected if an aged column is used for analysis.



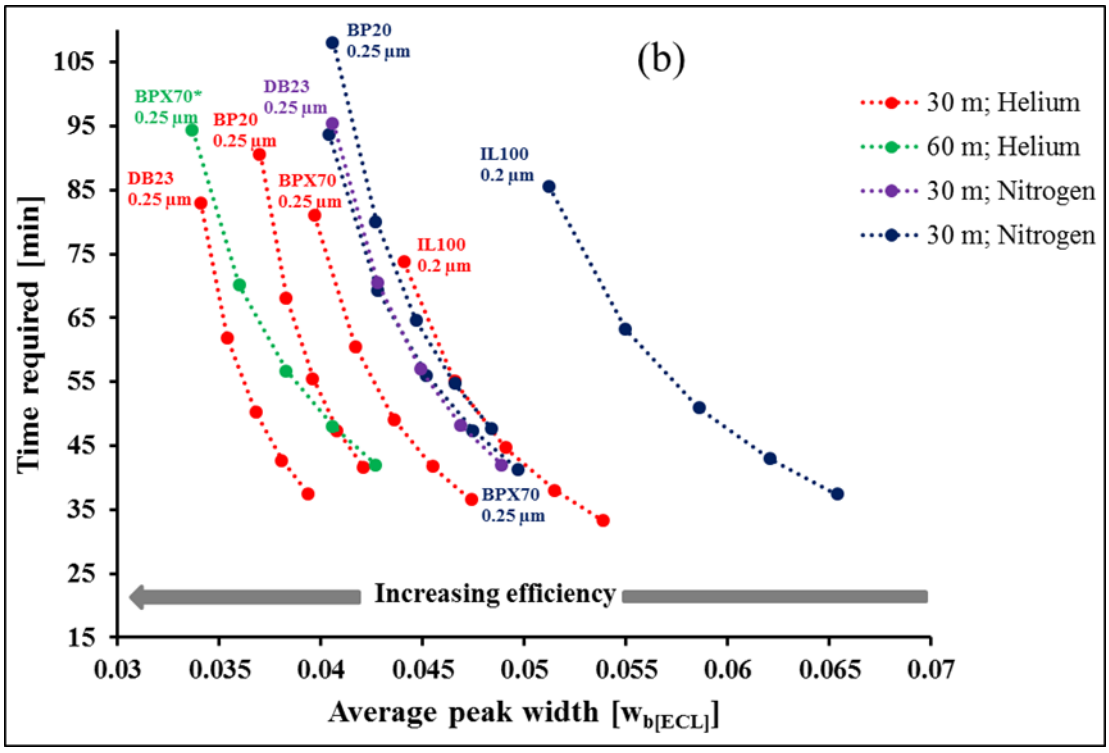


Figure 27: Efficiency-time relation in different columns (a) helium and hydrogen (b) helium and nitrogen as carrier gases (all at temperature levels of 1, 2 and 3 ^{0}C).

This difference in efficiency arises mainly from the difference in the polarity of the stationary phases in the columns. Column efficiency is observed to decrease with increasing polarity. The DB-23 column coated with (50% Cyanopropyl)-methylpolysiloxane and the BP-20

column coated with polyethylene glycol stationary phase have relatively less polarity and thus have higher efficiencies. The BPX70 coated with 70% cyanopropyl polysilphenylene-siloxane and the ionic IL100 columns are more polar columns with moderate and less efficiency respectively. The ECL value of the FAME 22:6 n-3 is used to compare the polarity of the columns. The higher the ECL values the more polar the column is. The ECL values using helium as a carrier gas (at 2 ^{0}C and 30 cm/s) were 23.98, 24.15, 24.82 and 25.40 for DB-23, BP-20, BPX70 and IL100 respectively.

The difference in column efficiency between the different types of columns is more easily observed when hydrogen is used as a carrier gas than when helium or nitrogen is used. Using nitrogen as a carrier gas there is no much difference in column efficiency between different column types except for IL100. A possible explanation is that the C term in the van Deemter equation can be split into a contribution from the stationary phase, C_s, and a contribution from the mobile phase, C_m [10]. Nitrogen has a high C term irrespective of the stationary phase, meaning that C_m is high. C_s will therefore be of minor importance for the sum of the terms, C. With hydrogen and helium C_m is much lower and the contribution from the stationary phase, C_s, is therefore more important. The IL100 when nitrogen is used appeared very different, possibly because of its high polarity the effect of the stationary phase dominated over the effect of carrier gas also with Nitrogen.

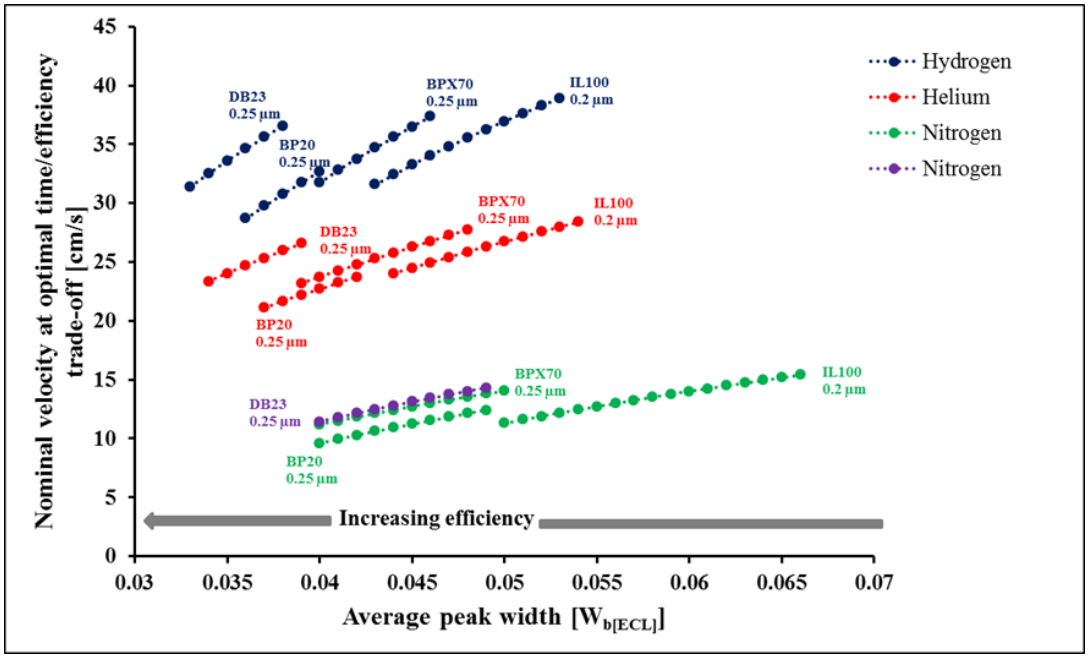


Figure 28: Column efficiency for different types of columns using hydrogen, helium and nitrogen as carrier gases (all at temperature level of 1, 2 and 3 ^{0}C).

The nominal velocity at optimal time efficiency trade-off is plotted versus inverse of efficiency for all the three carrier gasses (Figure 28). In general the optimum velocities of the three carrier gases used are in the range of 10-15, 20-30 and 30-40 cm/s for nitrogen, helium and hydrogen respectively. The range is wider for hydrogen and narrower for nitrogen. This is because the van Deemter curve for hydrogen is more flat than nitrogen around its optimum velocity (Figure 26). The optimum velocities required for different columns vary also when hydrogen is used than when helium or nitrogen is used, for instance when hydrogen is used the optimum velocity for BPX70 starts from where for BP20 ends.

4. Conclusions and further work

4.1. Conclusions

From the experiments carried out in this study it is possible to conclude the following in temperature programmed gas chromatography:

- Peak widths measured in retention index units can be explained by the van Deemter equation. Similar to the height equivalent to theoretical plate (H) in the isothermal van Deemter equation the minimum the peak width is the higher the efficiency.
- Five terms are always required in the expanded van Deemter equation to show the dependence of peak width on carrier gas velocity and on temperature gradient for all the three carrier gases employed (helium, hydrogen and nitrogen).
- Experimental designs like Doehlert design and factorial designs of 5 x 3, 3 x 3 or skewed 3 x 3 designs can be used to optimize the efficiency response function. The 5 x 3 design attributed to its higher number of experiments gives a model with lower RMSEP than the other designs. When a lower number of experiments has to be conducted skewed 3 x 3 design is better than Doehlert design and 3 x 3 design.
- The logarithm of the retention time of the last eluting FAME (26:0) is linearly related to the logarithm of carrier gas velocity and logarithm of temperature gradient and thus it can be modeled by a simple response surface function without a quadratic term.
- Higher efficiency is obtained at an optimal velocity and lower temperature gradient whereas minimum time of analysis is at higher carrier gas velocity and higher temperature gradient. Thus it always requires a trade-off between time and efficiency.
- The line of best time efficiency trade-off lays 2-5%, 3-6% and 3-8% above the velocities that represent the minima in the van Deemter curves for helium, hydrogen and nitrogen respectively.
- At similar retention time with helium hydrogen gives 6-8% more efficiency than helium while 12-18% decreases in efficiency with nitrogen is observed. At similar efficiency with helium hydrogen as carrier gas reduces the time of analysis by 24-31% whereas there is 80-92% increase in retention time with nitrogen as a carrier gas for a 30 m column.
- In FAMEs analysis column efficiency is found to decrease with increasing column polarity.

4.2. Limitations of the study and recommendations for further work

The carrier gas velocity used in this study is a nominal velocity, while the real velocity will increase throughout the chromatographic runs because of gas expansion and because the instruments were applied in constant mass flow mode. An alternative that has not been investigated and that may have some influence on the accuracy of the models is to replace the nominal velocity in the equations with the carrier gas flow. In constant flow mode the column head pressure is increased with the oven temperature to compensate for the increased viscosity of the carrier gas at higher temperatures. The instruments can also be operated in constant pressure mode, and it has not evaluated how well the methodology works with constant pressure. In constant pressure mode the column head pressure could be used instead of velocity in the equations.

Also in this study only temperature and velocity factors were changing with constant column dimensions. By including a third factor, column dimension, a better model may be developed which better explains the chromatographic separation process. Whether reducing the column length instead of or in addition to increasing the temperature rate when shorter analysis time is needed is also something that is poorly understood. This may be answered by starting with a long column and then performing experiments with gradually reduced column length.

References

1. Simopoulos A.P.; *Essential fatty acids in health and chronic disease*. Am J Clin Nutr, 1999. **70**: p. 560S–9S.
2. Simopoulos A.P.; *Omega-3 Fatty Acids and Cardiovascular Disease: The Epidemiological Evidence -review*. Environ. Health and Prev. Med., 2002. **6**: p. 203–209.
3. Connor W.E. *Importance of n-3 fatty acids in health and disease.* Am J Clin Nutr, 2000. **71**: p. 171S–5S.
4. William W. Christie.; *Gas chromatography and lipids (A Practical Guide)*. First edition 1989: The Oily Press Ltd, bridgwater, Somerset. p. 6-8.
5. Mjøs S.A., Grahl-Nielsen O.; *Prediction of gas chromatographic retention of polyunsaturated fatty acid methyl esters*. J. Chromatogr. A, 2006. **1110**: p. 171–180.
6. Seppänen-Laakso T., Laakso I., Hiltunen R.; *Analysis of fatty acids by gas chromatography, and its relevance to research on health and nutrition-review*. Anal. Chim. Acta, 2002. **465**: p. 39–62.
7. Martin A.J.P., Synge R.L.M. *A new form of chromatogram employing two liquid phases. A theory of chromatography and application to the micro-determination of the higher monoamino-acids in proteins.* Biochem. J., 1941. **35**: p. 1358-1368.
8. James A.T., Martin A.J.P.; *Gas-liquid Partition Chromatography: the Separation and Micro-estimation of Volatile Fatty Acids from Formic Acid to. Dodecanoic Acid*. Biochem. J., 1952. **50**: p. 679-690.
9. Eder K.; *Gas chromatographic analysis of fatty acid methyl esters -review*. J. Chromatogr. B, 1995. **671**: p. 113-131.
10. Wittkowski R., Matissek R.; *Capillary Gas Chromatography in Food Control and Research*. 1993: by Technomic publishing company Inc. Lancaster, Pennsylvania, USA. p. 19-33.
11. Fowlis I. A.; *Gas Chromatography*. 1995: by John Wiley & Sons Ltd. 2nd ed. University of Greenwich, Uk. p. 1- 100.
12. Harvey D.; *Modern analytical chemistry*. 2000: by The McGraw-Hill Companies, Inc., United States of America. p. 547-562.
13. Blumberg L.M.; *Plate height formula widely accepted in GC is not correct*. J. Chromatogr. A, 2011. **1218**: p. 8722-8723.
14. Usher K. M., Simmons C. R., Dorsey J. G. *Modeling chromatographic dispersion: A comparison of popular equations*. J. Chromatogr. A, 2008. **1200**: p. 122–128.

15. Blumberg L. M. *Theory of fast capillary gas chromatography- Part 3: Column performance versus gas flow rate*. J. High Resol. Chromatogr., 1999. **22** (7): p. 403-414.
16. Klee M.S., Blumberg L.M. *Theoretical and practical aspects of fast gas chromatography and method translation*. J. chromatogr. Sci., 2002. **40**: p. 234-246.
17. Kováts E.; *Gas-chromatographishe Charakterisierung organischer Verbindungen. Teil 1: Retentionsindices aliphatischer Halogenide, Alkohole, Aldehyde und Ketone.* Helv. Chim. Acta, 1958. **41**: p. 1915-1932.
18. Ettre L.S.; *The Retention Index System; Its Utilization for Substance Identification and Liquid Phase Characterization*.; Chromatographia, 1973. **6** (11): p. 489-495.
19. Ettre L.S.; *Retention Index Expressions*.; Chromatographia, 2003. **58** (7-8): p.491-494.
20. Castello G.; *Retention index systems: alternatives to the n-alkanes as calibration standards-review*. J. Chromatogr. A, 1999. **842**: p. 51–64.
21. Woodford F.P., Van Gent C.M. *Gas-liquid chromatography of fatty acid methyl esters: the "carbon-number" as a parameter for comparison of columns*. J. Lipid Res. January, 1960. **1**(2): p. 188-190.
22. Van Den Dool H., Kratz P.D.; *A generalization of the retention index system including linear temperature programmed gas-liquid partition chromatography*. J. Chvomalog., 1963. **11**: p. 463-471.
23. Wasta Z., Mjøs S.A.; *A database of chromatographic properties and mass spectra of fatty acid methyl esters from omega-3 products*. J. Chromatogr. A, 2013. **1299**: p. 94–102.
24. Mjøs S. A.; *Prediction of equivalent chain lengths from two-dimensional fatty acid retention indices*. J. Chromatogr. A, 2006. **1122:** p. 249–254.
25. Ettre L.S.; *Separation Values and Their Utilization in Column Characterization. Part II: Possible Improvement in the Use of the Separation Values to Express Relative Column Performance.* Chromatographia, 1975. **8** (7): p. 355-357.
26. Skartland L.K., Mjøs S.A., Grung B.; *Experimental designs for modeling retention patterns and separation efficiency in analysis of fatty acid methyl esters by gas chromatography–mass spectrometry*. J. Chromatogr. A, 2011. **1218**: p. 6823– 6831.
27. Hibbert D.B.; *Experimental design in chromatography: A tutorial review*. J. Chromatogr. B, 2012. **910**: p. 2– 13.

28. Ferreira S.L.C, Bruns R.E., da Silva E.G.P., et al.; *Statistical designs and response surface techniques for the optimization of chromatographic systems-review*. J. Chromatogr. A, 2007. **1158**: p. 2–14.
29. Ferreira S.L.C., dos Santos W.N.L, Quintella C.M., et al.; *Doehlert matrix: a chemometric tool for analytical chemistry—review.* Talanta, 2004. **63**: p. 1061–1067.
30. http://www.netascientific.com/NetaAdmin/Uploads/ContentDocument/Agilent%20GC%20Column%20Selection%20Guide%20(J&W).pdf. Agilent J&W GC Column Selection Guide.
31. Mjøs S. A. *Two-dimensional fatty acid retention indices*. J. Chromatogr A, 2004. **1061**: p. 201–209.

Appendix

Table 10: Table: Insignificant terms found by back ward elimination of the terms (RMSE given is the error of the resulting model after excluding the less significant term)

Column type	Carrier gas	1st round elimination		2nd round elimination		3rd round elimination		4th round elimination		Insignificant terms
		Less sig. term	RMSE	Less sig. term	RMSE	Less sig. term	RMSE	Less sig. term	RMSE	
DB-5 (0.1 µm)	He	g	0.000638	a	0.000651	f	0.000921	d	0.000993	a and g
	H_2	d	0.000647	g	0.000655	a	0.000793	f	0.000908	d and g
	N_2	g	0.000630	d	0.000660	a	0.000706	f	0.000776	d and g
DB-5 (0.25 µm)	He	g	0.000892	a	0.000914	f	0.00114	d	0.00121	a and g
	H_2	d	0.000949	g	0.000972	f	0.00108	a	0.00119	d and g
	N_2	g	0.000790	d	0.000813	a	0.000860	f	0.000949	d and g
BPX70 (30 m)	He	d	0.000633	g	0.000665	a	0.000743	f	0.000924	d and g
	H_2	d	0.000734	g	0.000733	f	0.00102	a	0.00179	d and g
	N_2	d	0.000609	g	0.000639	a	0.000838	f	0.00107	d and g
BPX70 (60 m)	He	a	0.000535	g	0.000561	d	0.000843	f	0.00109	a and g
	H_2	d	0.000649	g	0.000706	a	0.00101	f	0.00132	d and g
IL61	He	g	0.000908	a	0.000957	f	0.00140	d	0.00143	a and g
	H_2	g	0.00101	d	0.00118	f	0.00140	a	0.00162	d and g

Table 11: Optimum carrier gas velocity and efficiency at the minimum point on van Deemter curve and at best time efficiency trade-off point (column: DB-5 30 m, 0.25 mm, 0.25 µm).

Temperature rate (^{0}C/min)	u_{opt} (cm/s)	w_b (opt)	u_{topt} (cm/s)	w_b (topt)	% (u_{opt} - u_{topt})
Helium as a carrier gas					
2.0	24.4	0.0337	25.0	0.0338	2.5
2.5	24.9	0.0345	25.7	0.0345	3.2
3.0	25.4	0.0353	26.3	0.0353	3.5
3.5	25.9	0.0360	26.8	0.0361	3.5
4.0	26.3	0.0368	27.4	0.0368	4.2
4.5	26.8	0.0375	27.9	0.0375	4.1
5.0	27.2	0.0382	28.4	0.0382	4.4
5.5	27.6	0.0389	28.8	0.0390	4.3
6.0	28.0	0.0396	29.3	0.0397	4.6
Hydrogen as a carrier gas					
2.0	34.0	0.0324	35.0	0.0324	2.9
2.5	34.7	0.0331	35.9	0.0331	3.5
3.0	35.3	0.0337	36.7	0.0338	4.0
3.5	35.9	0.0344	37.4	0.0344	4.2
4.0	36.5	0.0350	38.2	0.0350	4.7
4.5	37.1	0.0356	38.9	0.0357	4.9
5.0	37.7	0.0362	39.6	0.0363	5.0
5.5	38.3	0.0369	40.3	0.0369	5.2
6.0	38.8	0.0375	41.0	0.0375	5.7
Nitrogen as a carrier gas					
2.0	11.7	0.0371	12.3	0.0371	5.1
2.5	12.1	0.0383	12.8	0.0384	5.8
3.0	12.5	0.0396	13.3	0.0396	6.4
3.5	12.8	0.0407	13.7	0.0408	7.0
4.0	13.2	0.0419	14.2	0.0420	7.6
4.5	13.5	0.0430	14.6	0.0431	8.1
5.0	13.8	0.0442	15.0	0.0443	8.7
5.5	14.2	0.0452	15.4	0.0454	8.5
6.0	14.5	0.0463	15.7	0.0465	8.3

- u_{opt} -Velocity that gives maximum efficiency at gradient
- w_b(opt) - Inverse efficiency at u_{opt}
- u_{topt} - Optimal velocity at gradient (efficiency/time trade-off)
- w_b(topt) - Inverse efficiency at u_{topt}