Mechanisms of retention in HPLC

María Celia García-Álvarez-Coque

Department of Analytical Chemistry
University of Valencia
Valencia, Spain

https://sites.google.com/site/fuschrom/

Part 5

Index

5.1. Estimation of dead time

5.1.1. Definition and problems

The **void volume** in liquid chromatographic systems has been defined as the volume of mobile phase that fills the **space between the injector and the detector cell,** which includes the accessible interstitial or **interparticle volume** (the volume between column particles) and the **intraparticle volume** (the pores of the column packing), as well as the volume of tubing and any other component in the system (the **extra-column volume**).

mobile phase

interparticle volume

intraparticle volume
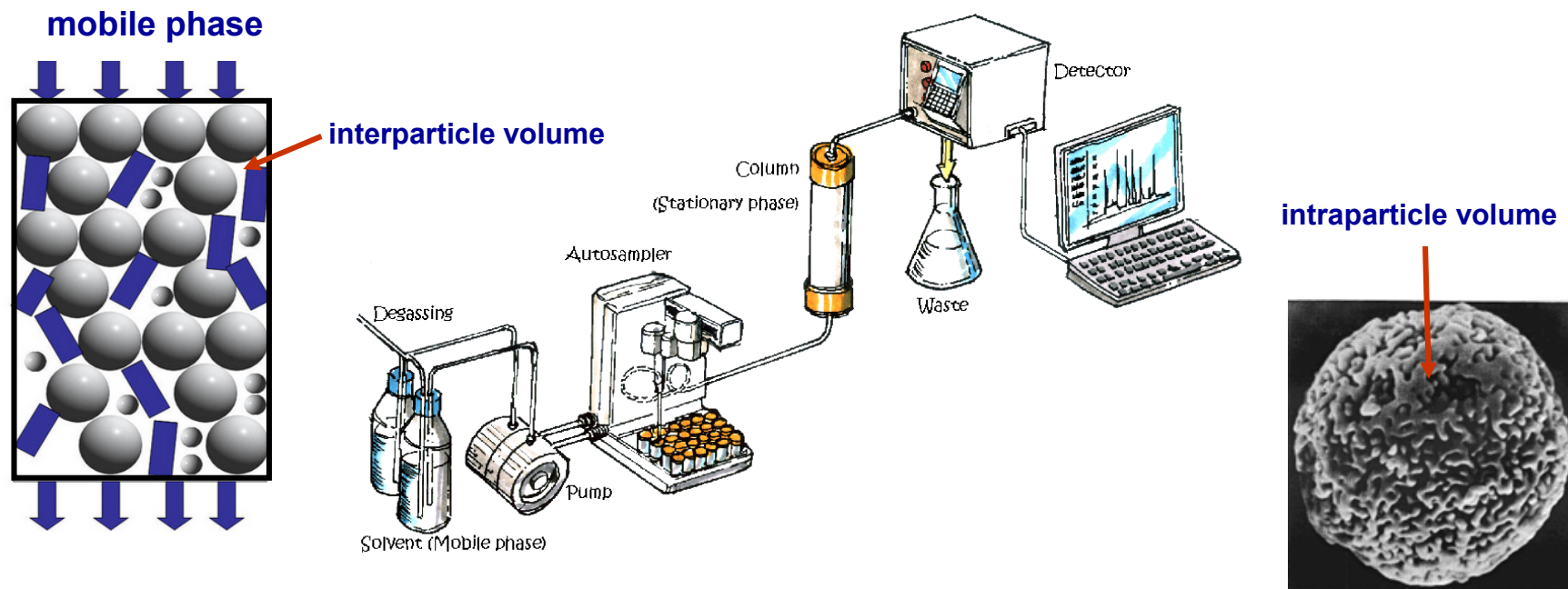
Related with the concept of void volume is the dead time, which is the time that an ideal unretained compound (a compound that does not interact with the stationary phase) needs to cross the distance between the injector and the detector, when eluted at constant flow rate.
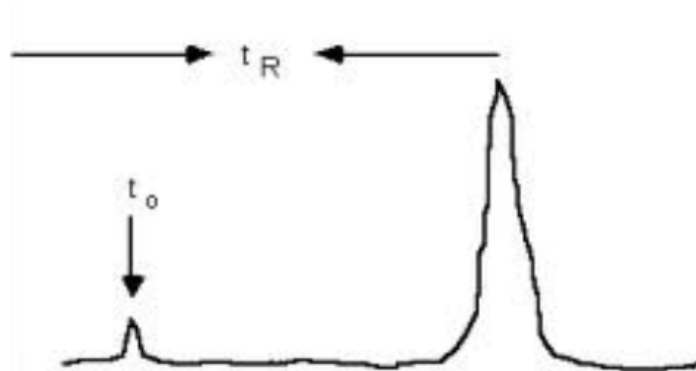
Retention factor

● **Dead time estimation is the basis for the calculation of retention factors.**

● **This calculation implies moving the origin of retention times to the time of an unretained compound, and correction of the column length or flow effects through division by the dead time.**

$$k = \frac{t_R - t_0}{t_0}$$ (5.1)

t_R : solute retention time

t_0 : dead time

The **retention factor** normalises the retention, and allows comparison among **different column lengths,** or among **different flow rates for the same column.**

Need of retention factors

Accurate knowledge of retention factors is needed in several fields:

● **Prediction of retention for optimisation of chromatographic separations**

$$\log k = c_0 + c_1\, \varphi + c_2\, \varphi^2$$

● **Estimation of partition coefficients and other thermodynamic parameters**

$$\log k = c_0 + c_1 \log P_{o/w}$$

● **Establishment of correlations with several chemical-physical properties**
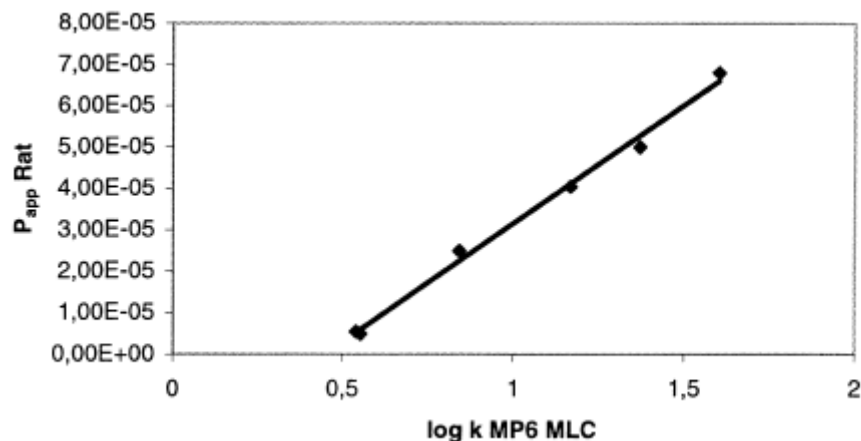
P_{app} : apparent permeability coefficients of rat intestinal segments

QRAR: Quantitative retention-activity relationships

7

Problems in the estimation of the dead time

● **The stationary phase adsorbs a certain amount of organic solvent,** forming a layer and thus reducing the accessible volume inside the column (especially in RPLC).

● Solute **molecules may be partially or completely excluded** from the stationary phase pores. This means that **different solutes will have different associated dead times:** small solutes and portions of some larger solutes can **access the average pores,** while other solutes may not be able to enter the pores.

● **Electrostatic exclusion** may also happen.
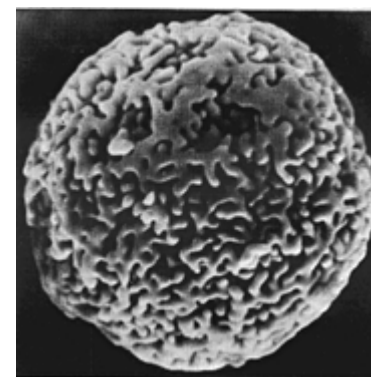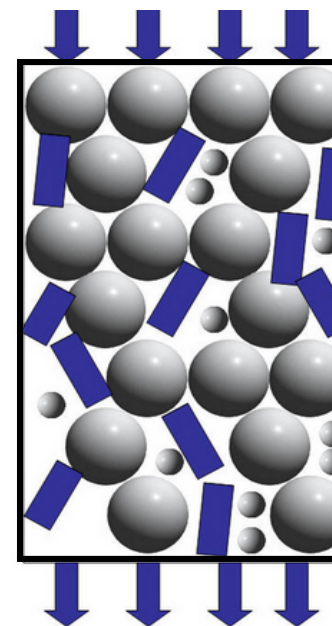
Methods for dead time estimation

Static methods

no flow and column kept at atmospheric pressure

Dynamic methods

mobile phase is flowing and there is a linear pressure gradient along the column

The results yielded by different methods can be expected to differ:

different properties are measured

particular experimental uncertainties
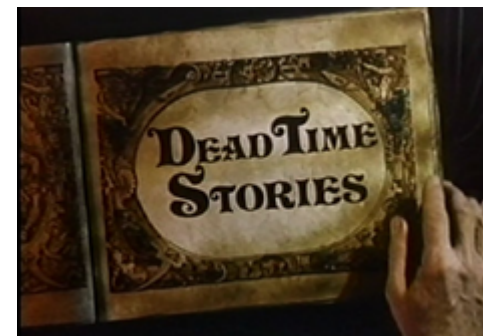
The discussion and controversy on the dead time estimation has kept alive during decades. The general opinion is that there is no acceptable method for the accurate evaluation of dead time yet.
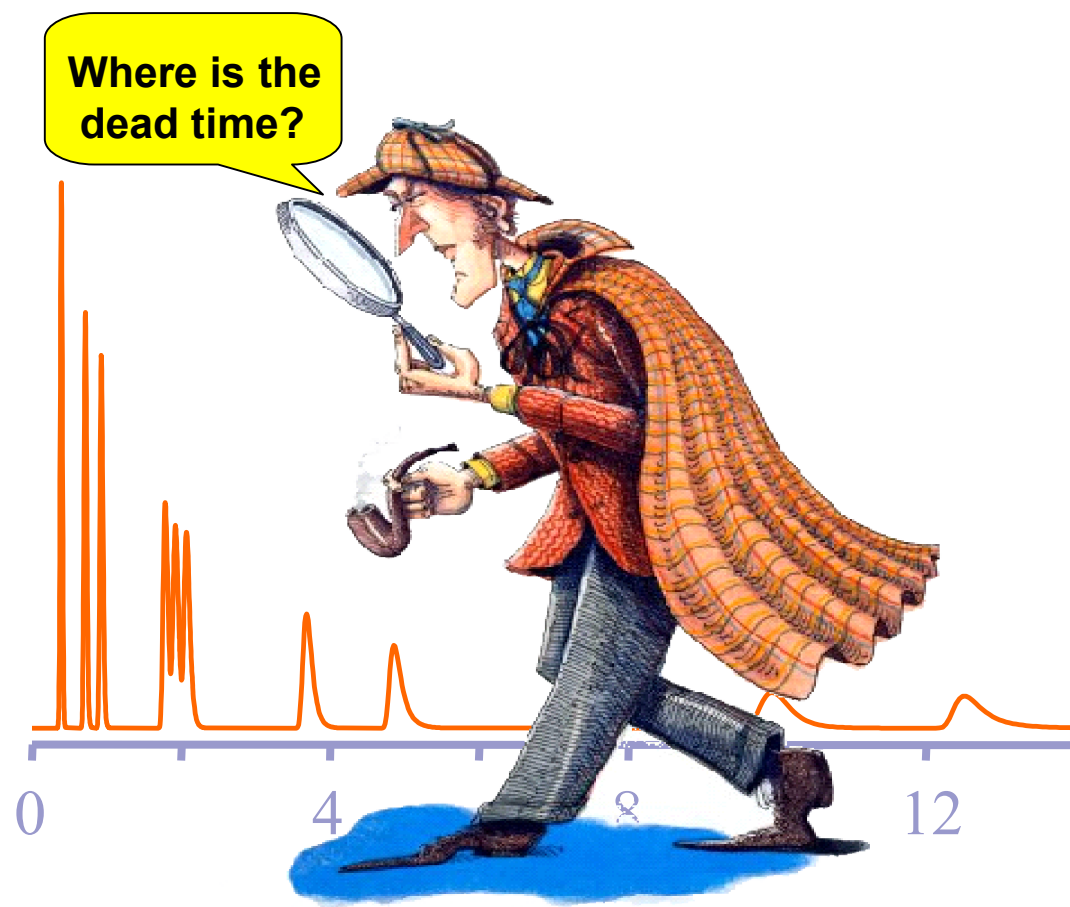
Eye-catching image in:

S. Pous-Torres, J.R. Torres-Lapasió, J.J. Baeza-Baeza, M.C. García-Álvarez-Coque,
Alternating iterative regression method for dead time estimation from experimental designs,
Anal. Bioanal. Chem. 394 (2009) 625–636.

5.1.2. Static methods

● **Pycnometry or weight difference method:** the packed **column is filled successively with two solvents** of sufficiently different density (carbon tetrachloride and methanol), followed by **weighting**; the column volume is then obtained from the differences in density and weight.

　　carbon tetrachloride: 1.5867 g cm^{-3}

　　methanol: 0.7918 g cm^{-3}

⟹ **Solvation** of the stationary phase by the mobile phase components is **ignored,** giving rise to an error in the **estimation of column dead volume.**

⟹ **Error derived from the estimation of a small magnitude from the subtraction of two larger magnitudes:** weights of column filled with each solvent.

● **One solvent method:** the column is **flushed with water and dried** with a nitrogen stream, followed each by weighting.

⟹   **Solvation is ignored** giving rise to **overestimation of column dead volume.**

⟹   Error derived from the **estimation of a small magnitude from the subtraction of two larger magnitudes:** weights of column filled with each solvent.

⟹   Once the column is dried, it may be irreversibly **damaged.**

● Measurement of the **difference between the volumes of the empty column tube and the packed material.**

⟹   Implementation requires the **use of special instruments**

⟹   It needs a long time to be applied

Static methods are not routinely used !!!

5.1.3. Dynamic methods : mobile phase is flowing

Direct methods: injection of unretained compounds
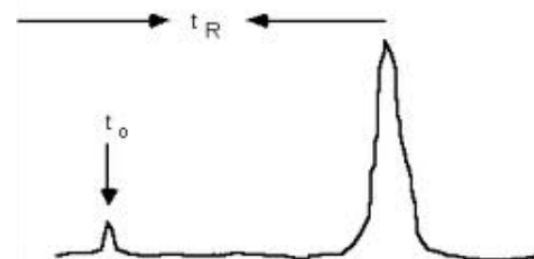
● **mobile phase components (minor disturbance method):**

⟹ water, organic solvent, solution with composition slightly different from mobile phase

● **Refractive index detector are not necessary** to measure the solvent disturbance peak.

● **UV detectors** evidence changes in the refractive index at low wavelengths

Direct methods :

injection of unretained compounds

● **UV-absorbing markers:**
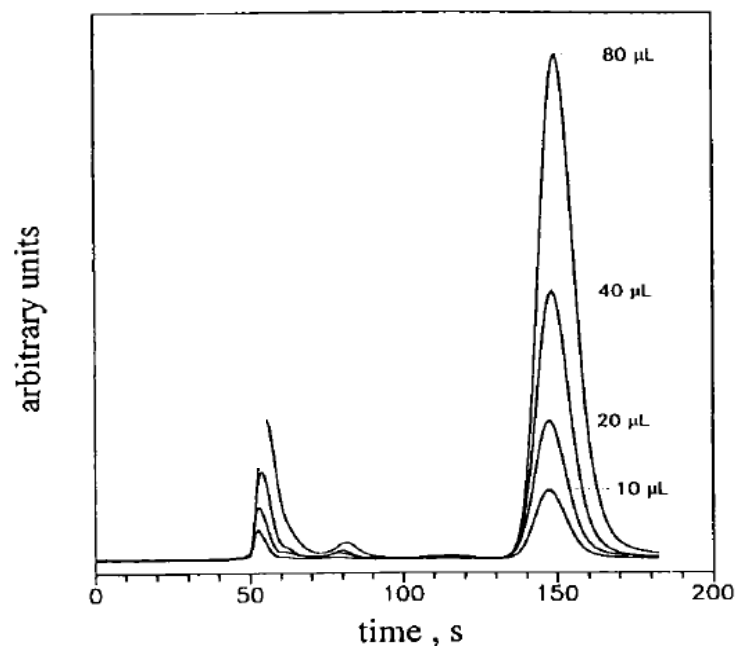
⟹ **organic compounds (acetone, benzoic acid, nitrobenzene, picric acid, uracil)**

⟹ **inorganic compounds (KI, KBr, KNO$_3$, NaCl, NaNO$_3$ and NaNO$_2$)**

RPLC

✓ Uracil   ✓ Pyridine   ✗ Toluene   ✗ Hexane

-1   0   1   2   3   Analyte LogP

The **ideal unretained compound** should be small enough to access the whole accesible stationary phase volume, and hydrophilic enough to stay out of the stationary phase. The main problem of direct methods is that all probe compounds are either slightly retained or excluded from the stationary phase. Also, the measurements depend on the experimental conditions.

Uracil and KBr seem to be the best markers !!!

Indirect methods: mathematical methods

● **Elution of successive members in a homologous series**

based on an assumed linear relationship

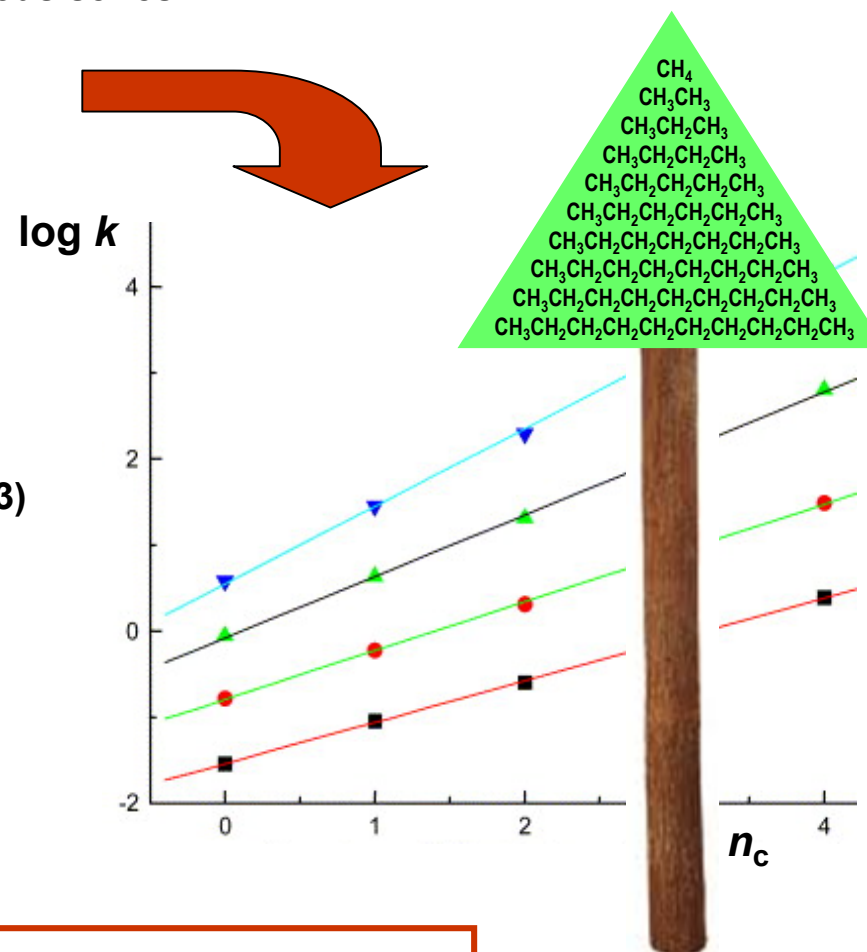$$\log k = c_0 + c_1\, n_c \qquad (5.2)$$

$$k = \frac{t_R - t_0}{t_0} \qquad (5.1)$$

$$t_R = t_0\,(1 + k) = t_0\,(1 + k_0\, e^{c_1\, n_c}) \qquad (5.3)$$

n_C : homologue carbon number

k_0 : residual retention factor for $n_C = 0$

c_1 : linear relationship slope

t_0 , k_0 and c_1 are calculated by non-linear regression

Drawbacks of the homologous series method

● **The assumed linear relationship is not always valid** along the whole series: Deviations from linearity have been observed for the smaller homologues and for the homologues exceeding the alkyl chain of the bonded phase.

● The choice of the homologous series is based on their availability, solubility in the mobile phase, retention and detection.

● A single series of homologues may not cover the entire range of mobile phase compositions, since the solubility and retention of the highest homologues become rapidly inappropriate with decreasing organic solvent content.
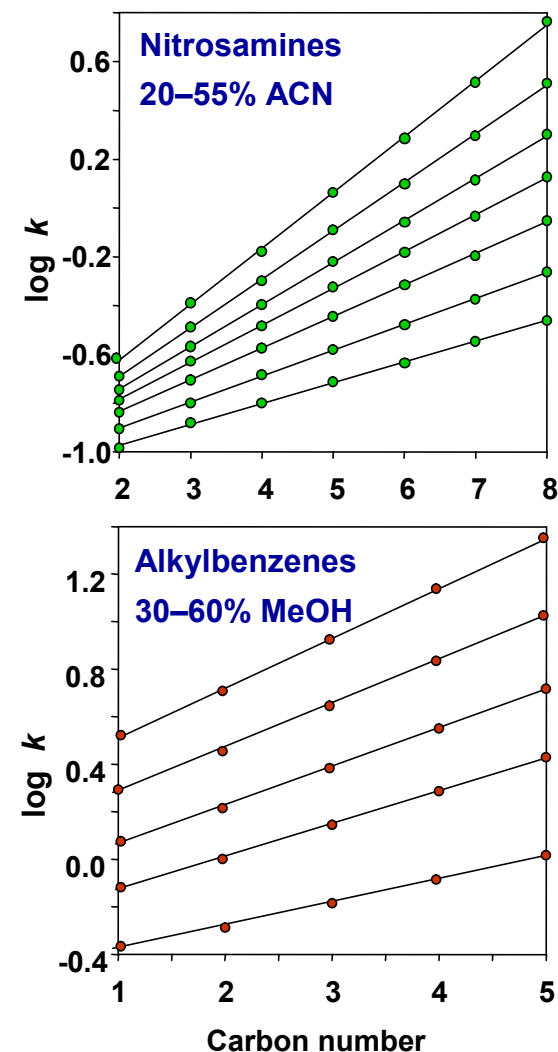
Drawbacks of the homologous series method

● **The method is more time-consuming than other dynamic methods.**

● **The obtained dead time is an extrapolated value and, therefore, it requires highly precise and accurate data.**

$$t_R = t_0\,(1 + k) = t_0\,(1 + k_0\,e^{c_1\,n_C})$$

● **The estimated dead time depends significantly on:**

choice of homologues

retention times of homologues

the mathematical approach

data quality

Solute retention versus composition relationships: a new mathematical method

> The approach processes the retention times of two or more compounds at several mobile phase compositions. It can make use of the data for the same compounds for which the dead time knowledge is needed, without requirement of external compounds.

- Any retention model can be used, but the **linear relationship** between $\log k$ and φ is the most convenient, fitted for a **narrow solvent concentration range,** where the linear $\log k$ relationship is valid and the column dead time is negligibly affected by changes in composition.
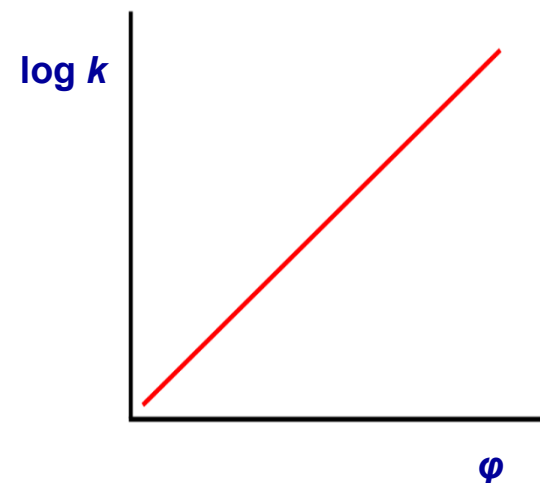
$$\log k = c_0 + c_1\,\varphi \qquad\qquad k = \frac{t_R - t_0}{t_0}$$

$$t_R = t_0\,(1 + k) = t_0\,(1 + k_0\,e^{c_1\varphi}) \qquad (5.4)$$

better convergence behaviour $\longrightarrow$

$$t_R = t_0 + e^{a + b\varphi} \qquad (5.5)$$

$\log k$

φ

- Model parameters can be obtained through **non-linear regression,** by fitting the **retention times at several mobile phase compositions.**

The data from a single compound can be used to estimate the dead time. However, these estimations are strongly affected by the magnitude of the retention times being processed. Reliable estimations are obtained using the retention data of several compounds eluted under the same conditions, which are treated simultaneously in an iterative algorithm.

● The simultaneous use of retention data from several solutes compensates the lack of accuracy or inadequacy of some data.

● Compounds of any kind can be used, without any special requirement except that they should not be excessively retained in the experimental conditions.

● The approach is implicitly accepting that:

⇒ measured void volume is the space accessible to solutes

⇒ accessibility to the stationary phase pores of the solutes treated simultaneously is similar

⇒ void volume is not significantly affected by changes in mobile phase composition
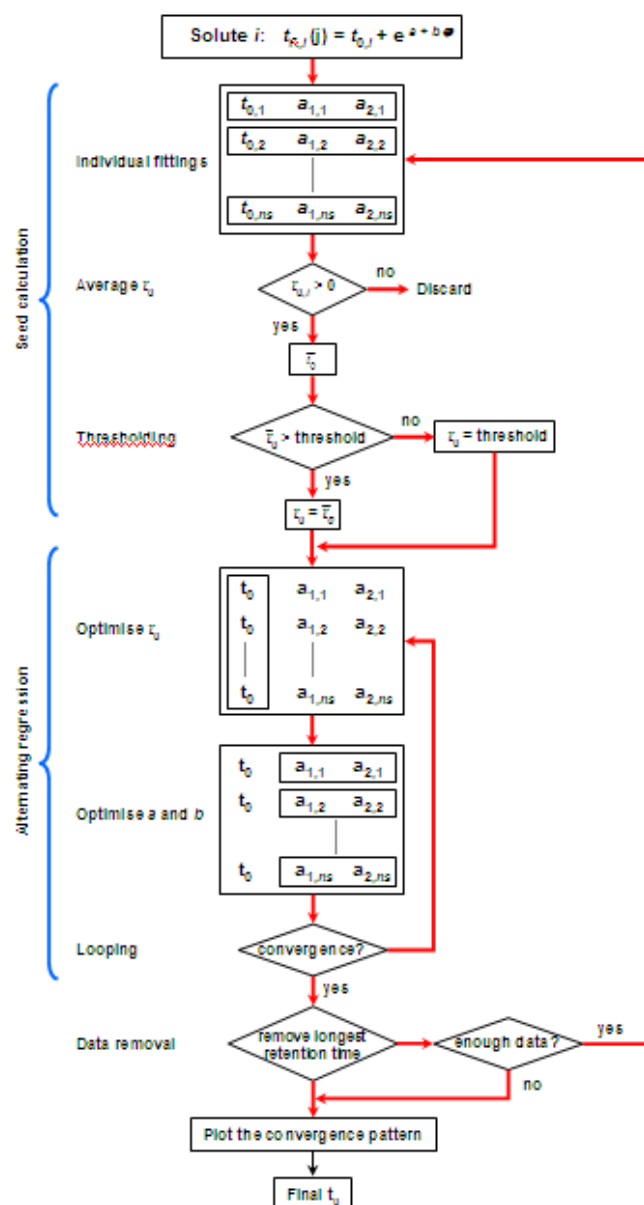
Alternating iterative regression method for dead time estimation from experimental designs

S. Pous-Torres, J.R. Torres-Lapasió, J.J. Baeza-Baeza, M.C. García-Álvarez-Coque, *Alternating iterative regression method for dead time estimation from experimental designs,* Anal. Bioanal. Chem. 394 (2009) 625–636.

5.1. Estimation of dead time

　　5.1.1. Definition and problems

　　5.1.2. Static methods

　　5.1.3. Dynamic methods

　　5.1.4. Use of solute retention versus composition relationships

5.2. Effect of pressure

　　5.2.1. Origin of the deviations of retention factors

　　5.2.2. Effect of pressure on Chromolith columns

5.3. Effect of temperature

　　5.3.1. Van't Hoff equation

　　5.3.2. Combined effect of solvent content, pH and temperature

5.4. Enhancing the predictions of retention
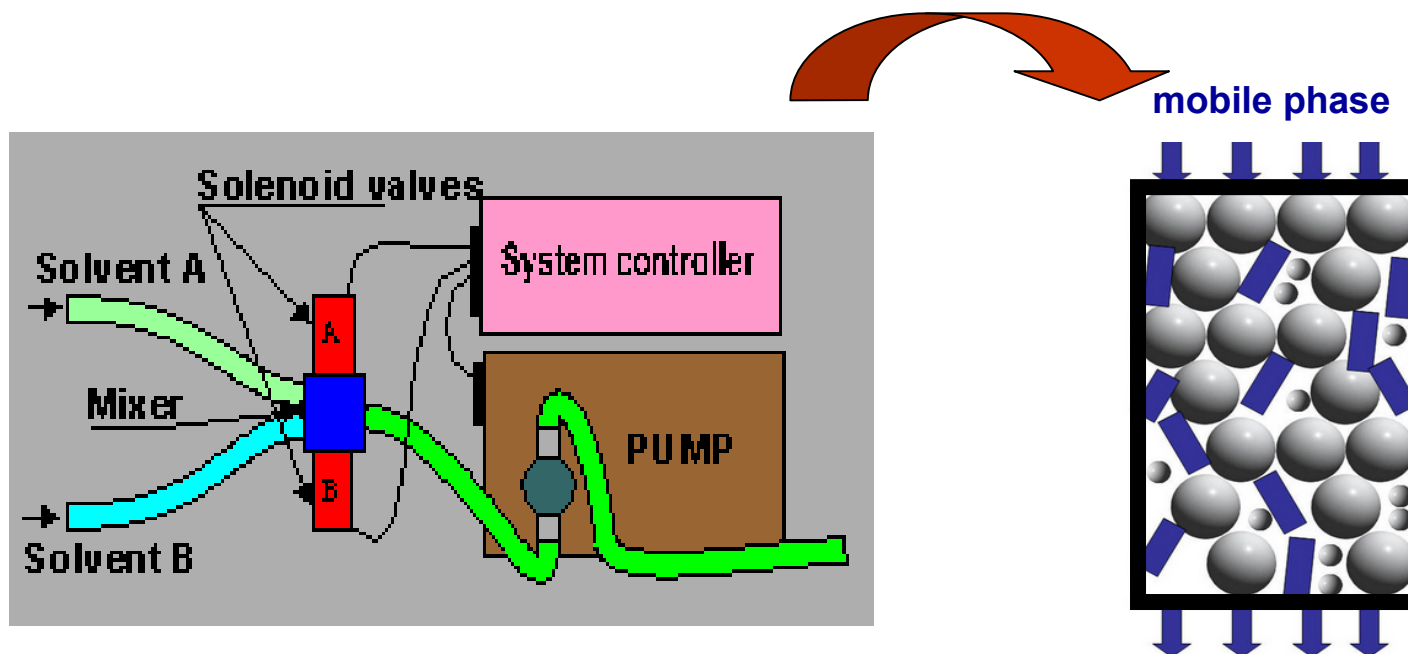
5.5. Recommended literature

5.2. Effect of pressure

5.2.1. Deviations of retention factors

In an **ideal chromatographic system,** the mobile phase, packing and column cartridge are **not affected by pressure.** When the system departs from the ideal behaviour, the **linear velocity profile** of the mobile phase is altered by pressure, and deviations in retention are expected.

mobile phase

Deviations in retention are due to …

- Changes in the mobile phase volume and viscosity

- Deviations in Darcy's law (describes the flow of a fluid through a porous medium)

- Column expansion

- Packing compression

- Deviations as a result of the perturbation of the equilibria inside the column at high pressure, when the solute partial molar volume varies during the equilibrium process. This is especially significant for large molecular-weight compounds.

- Forcing a stream of liquid at high velocity through a low permeability bed is not possible without generating a certain amount of heat (frictional heating). This will affect the partition equilibria, decreasing the retention.

Discriminating the contribution of the different sources to the observed deviations is not easy.

5.2.2. Effect of pressure on Chromoliths

> The **deviations** in retention time **are only evidenced** when working **at varying flow rate,** or at **varying pressure and constant flow rate** (which is unusual in practice).


The **interest on these effects has increased in recent years due to the development of fast chromatography.**

A **particular behaviour is observed with silica-based monolithic columns.**

Silica-based monolithic columns are made of a silica skeleton rod encapsulated within a PEEK tube, resulting in a **network of macropores interconnected by channels** through which the mobile phase percolates.
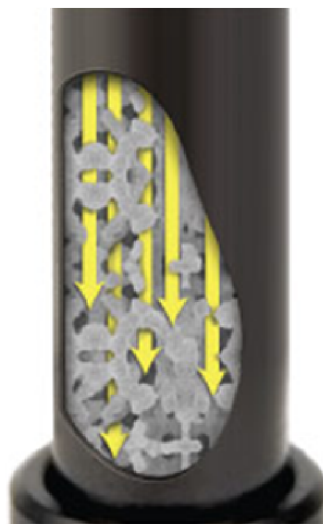
The **volume of the channels** in a monolithic column is **larger than the volume between the particles in microparticle columns,** and their structure is less tortuous. This allows **large flow rates** beyond those feasible for conventional packed columns at relatively low pressure.
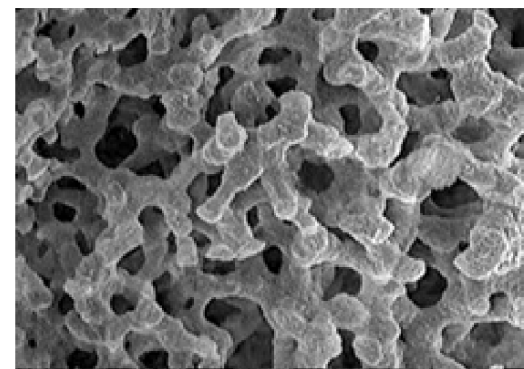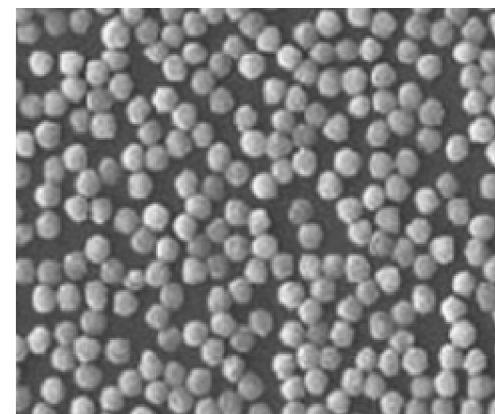

Macropores: 2µm

With Chromoliths, flow rate becomes an important factor, in addition to the mobile phase composition, in order to achieve good resolution at sufficiently low analysis time.

● However, increasing flow rates means increasing pressure.

● As a result of the bed and tube elasticity, mobile phase linear velocity may vary with pressure to a certain extent.

● The initially cylindrical column at atmospheric pressure may be deformed under the influence of the pressure gradient. The column becomes approximately a truncated cone.
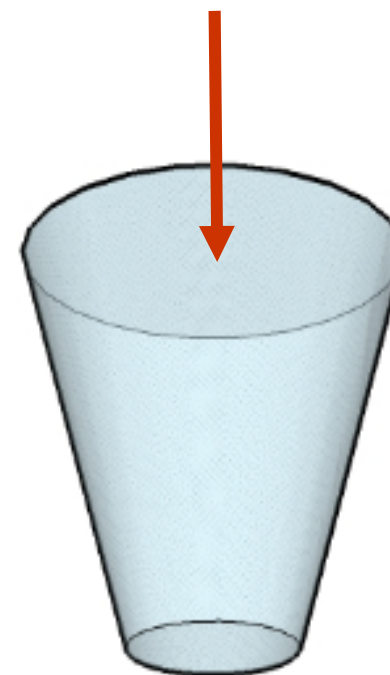
● The **deformation is small for stainless steel** tubes and pressures up to 1000 bar (only some tenths percent)

Young modulus (stainless steel) = 210 GPa

● **Chromolith should suffer a larger stress** with pressure tending to inflate the column and increase its volume.

Young modulus (PEEK) = 3.6 GPa

(>50 times more elastic than stainless steel).

● This **decreases the linear velocity** inside Chromoliths, and **increases the retention** at relatively low pressure (< 200 bar, maximal pressure for the PEEK material).

● In contrast, **frictional heating,** which is an issue for microparticle columns, seems to be **less significant for Chromoliths,** owing to the smaller resistance to the flow.
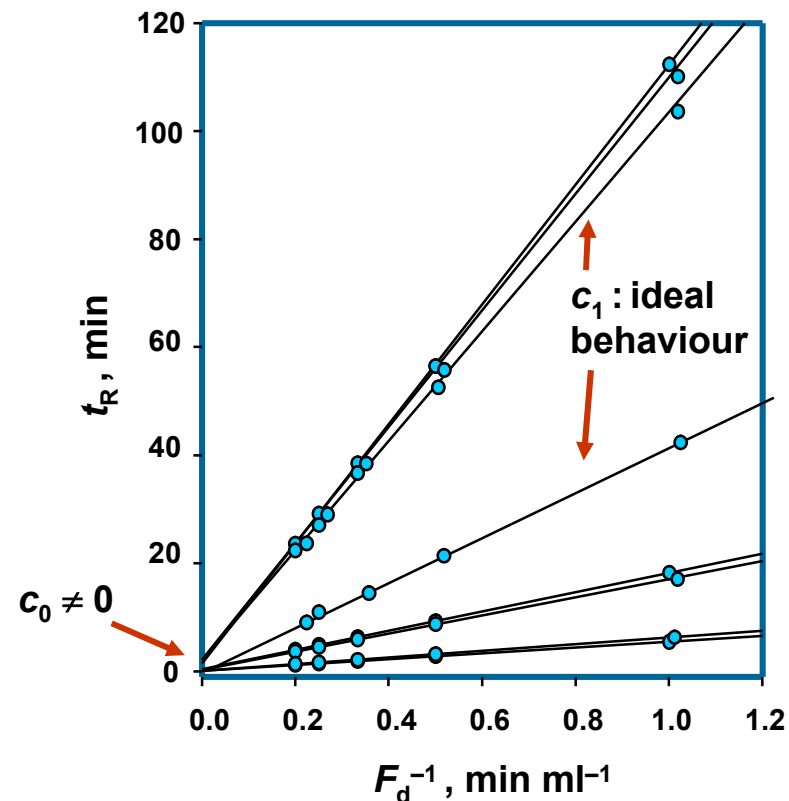
$$k = \frac{t_R - t_0}{t_0}$$

The retention factor should not depend on the flow rate. However, under the effect of pressure, significant changes are observed with Chromoliths.
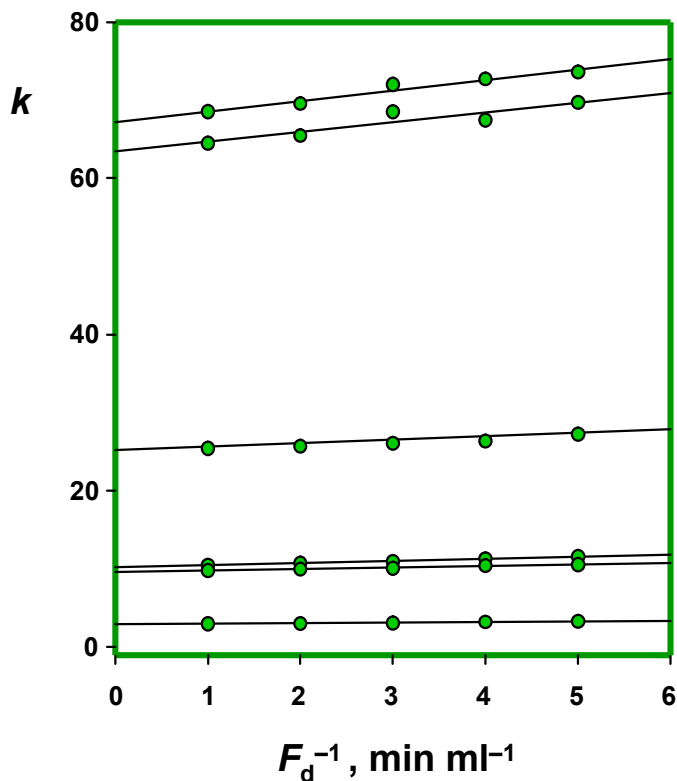
The plots of the retention time versus the inverse of delivered flow rate (F_d) should exhibit intercept of zero ($c_0 = 0$), but positive deviations are observed: more significant for the most retained compounds and higher flow-rate.

k

F_d^{-1}, min ml^{-1}

t_R, min

c_1 : ideal behaviour

$c_0 \neq 0$

F_d^{-1}, min ml^{-1}

Due to the **correlation of log c_0 with solute polarity,** the **deviation in the retention time for the dead time marker is smaller.** This makes the retention factors depend on the flow rate.
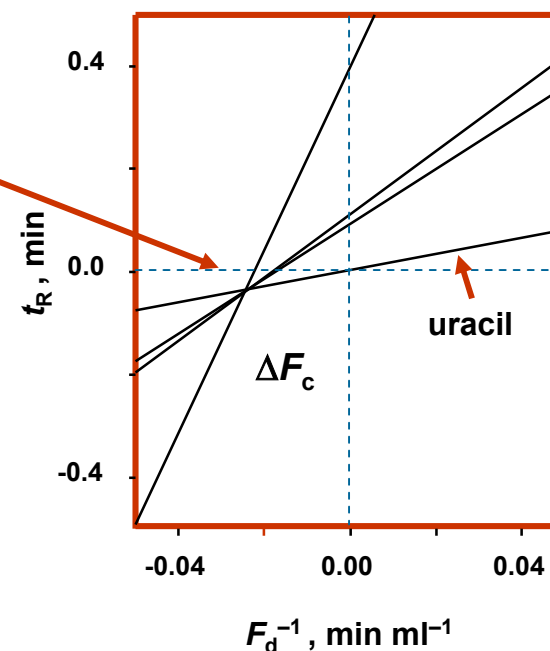
$$k = \frac{t_R - t_0}{t_0}$$

$$t_R = \frac{c_1}{F_d} + c_0 \qquad (5.6)$$

$$\frac{1}{F_{app}} = \frac{1}{F_d} + \frac{1}{\Delta F_c} \qquad (5.7)$$

converge approximately in a common point that **allows the estimation of an apparent flow rate** (F_{app})

F_d : **experimental flow rate**

ΔF_c : **flow rate deviation**

29

Several authors have reported **models** with RPLC microparticle packed columns **that include both organic solvent and flow rate as factors,** such as:

$$y = c_0 + c_1\,\varphi + c_2\,F + c_{12}\,\varphi\,F + c_{11}\,\varphi^2 + c_{22}\,F^2 \qquad (5.8)$$

y is the t_R or log *k*

Correction of retention factors to the ideal behaviour

(independently of the origin of the deviations)

$$k_i^c = \frac{t_{R,i}^c - t_0^c}{t_0^c} = \frac{(t_{R,i} - c_{0,i}) - (t_0 - c_{0,0})}{(t_0 - c_{0,0})} \qquad (5.9)$$

t_0 : experimental dead time at each flow rate

t_0^c : corrected dead time

$c_{0,i}$: deviation in retention time for a retained compound

$c_{0,0}$: deviation in retention time for the marker

$$\log k = c_0 + c_1\,\varphi + c_2\,\varphi^2$$

Once retention factors are corrected, the models do not need to include the flow rate as factor.

5.1. Estimation of dead time

    5.1.1. Definition and problems

    5.1.2. Static methods

    5.1.3. Dynamic methods

    5.1.4. Use of solute retention versus composition relationships

5.2. Effect of pressure

    5.2.1. Origin of the deviations of retention factors

    5.2.2. Effect of pressure on Chromolith columns

5.3. Effect of temperature

    5.3.1. Van't Hoff equation

    5.3.2. Combined effect of solvent content, pH and temperature

5.4. Enhancing the predictions of retention

5.5. Recommended literature

5.3. Effect of temperature

5.3.1. Van't Hoff equation

> The **influence of temperature on retention** can be described from thermodynamic considerations, using the **Vant'Hoff equation.**

$$\ln k = \ln K + \ln \frac{V_S}{V_m} = \frac{\Delta S}{R} - \frac{\Delta H}{RT} + \ln \phi \qquad \textbf{(5.10)}$$

K : partitioning constant Φ : system phase ratio

ΔH : standard enthalpy change ΔS : standard entropy change

R : universal gas constant T : absolute temperature

For **sufficiently narrow temperature** ranges (about 90ºC in RPLC), where ΔH and ΔS are constant, Eq. (5.10) can be transformed into a simple **two-parameter equation:**

$$\ln k = c_0 + \frac{c_1}{T} \qquad \textbf{(5.11)}$$

For **wider ranges,** due to the dependence of ΔH and ΔS with temperature:

$$\ln k = c_0 + \frac{c_1}{T} + \frac{c_2}{T^2} \qquad \textbf{(5.12)}$$

The usefulness of temperature to improve separations is rather controversial.

Retention times are decreased with temperature, but solute diffusion is collaterally increased, which deteriorates the resolution to a greater or lesser extent.

However, improvements in peak efficiency with temperature (due to enhanced kinetics in adsorption processes) have been reported, which increase the resolution.

Something more to know …

- The **slope of the Van't Hoff equation** for different solutes may **differ** enough to significantly **affect the selectivity,** occasionally giving rise to peak reversals.

- Changes in selectivity with temperature are **larger for ionisable or polar compounds.**

- The selectivity changes are **particularly intense for large molecules,** such as proteins, that can exhibit diverse conformations depending on temperature.
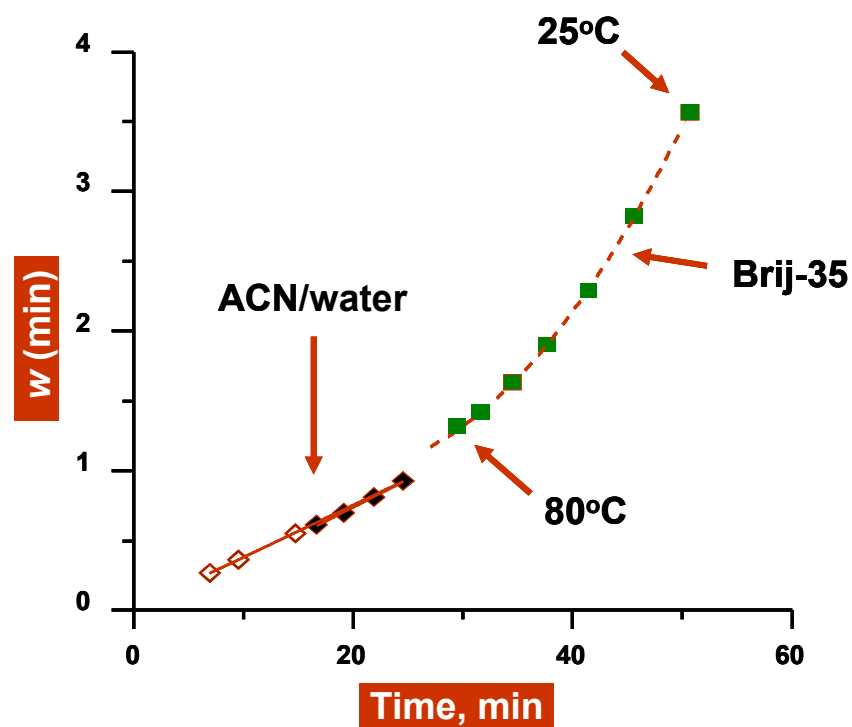
Flavonoids

ACN / water

34
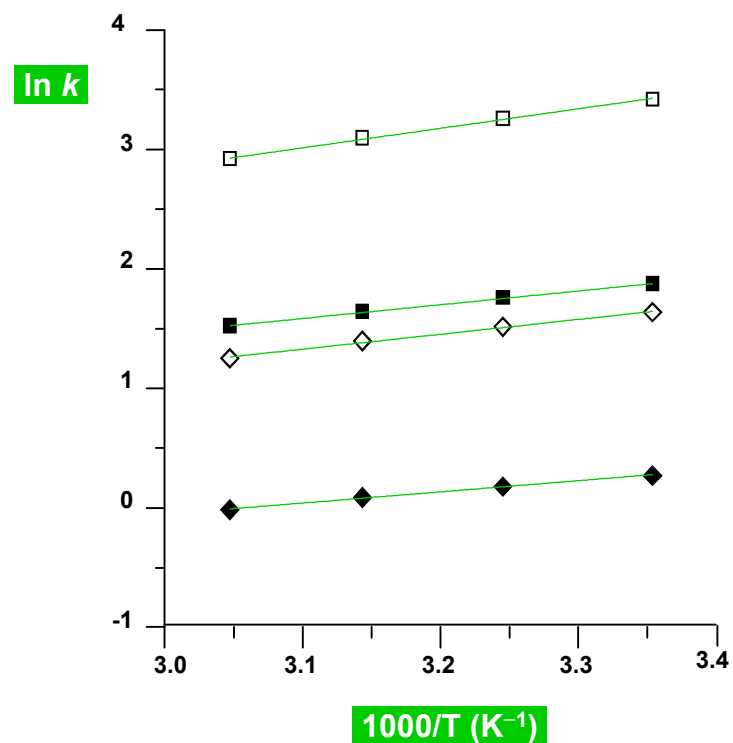
Enhancements in peak shape with temperature

**Flavonoids**

Brij-35 / water

The **effects of temperature and elution strength** (organic solvent content) on selectivity are approximately **orthogonal** to each other. Therefore, a peak pair that is difficult to separate by optimising the elution strength, may be separated by optimising the temperature. Although generally **not providing as much influence** on the selectivity as the organic modifier content, gradient steepness, solvent type and pH, temperature can be **still a worth factor** in method development, **due to the reduction in analysis time.**
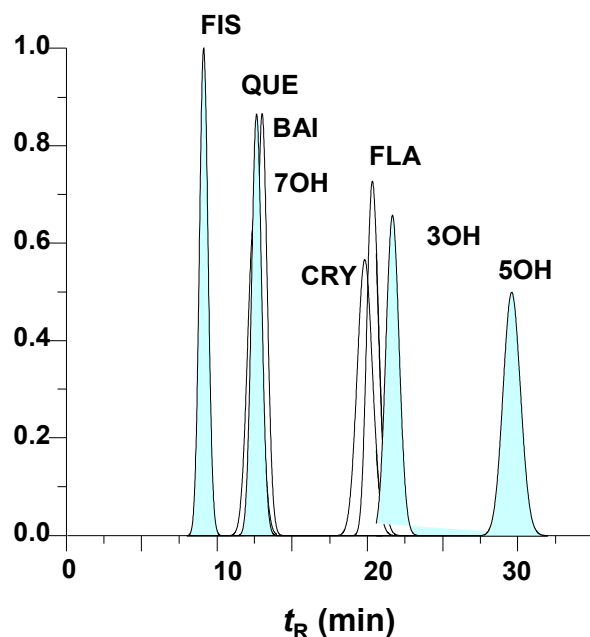
**50 mM Brij-35 at 25°C**   **40 mM Brij-35 at 55°C**
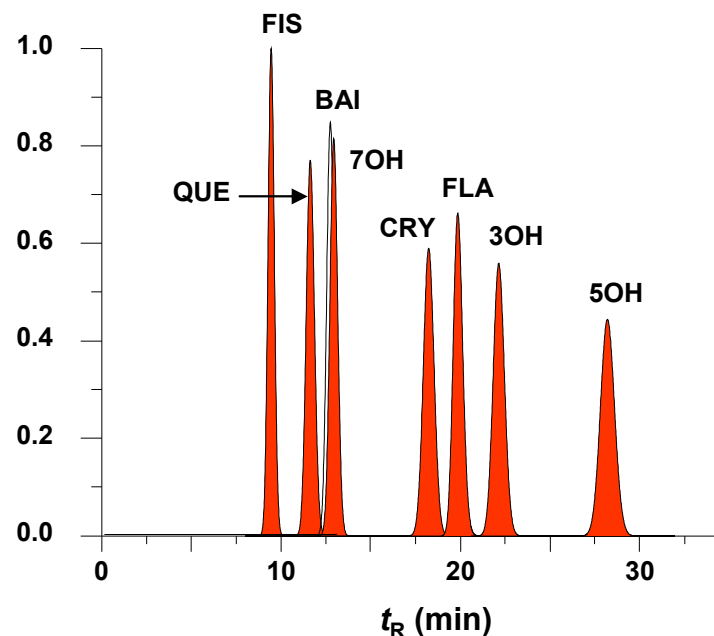
5.3.2. Combined effect of solvent content, pH and temperature

For **ionisable compounds, interactions between solvent content and pH, and temperature and pH** can be expected. The intensity of these interactions makes modelling of retention a **challenging problem.** In addition, a **unique experimental design** should satisfy the information requirements of several compounds simultaneously, which may make the location of experiments in the design, critical.

Diuretics and β-blockers

acidic　　　　　　basic

29.0% acetonitrile/50.0°C/pH 4.12

39.4% acetonitrile/20.6°C/pH 3.00

Time, min

37
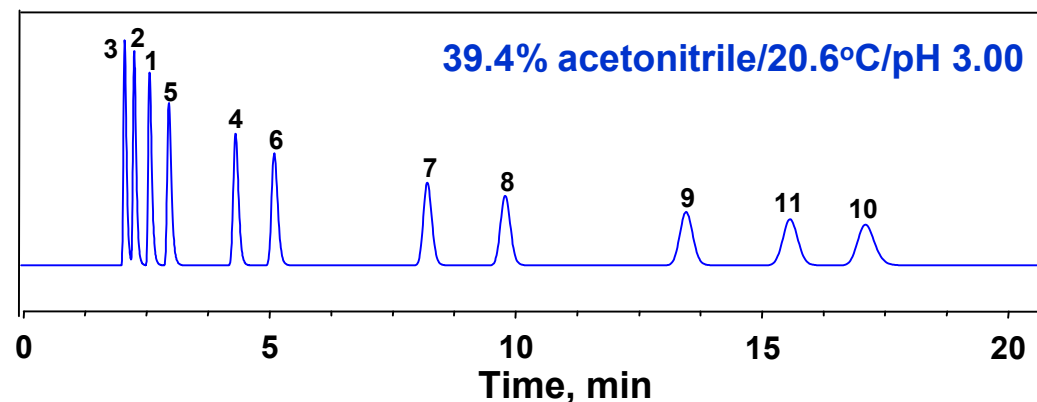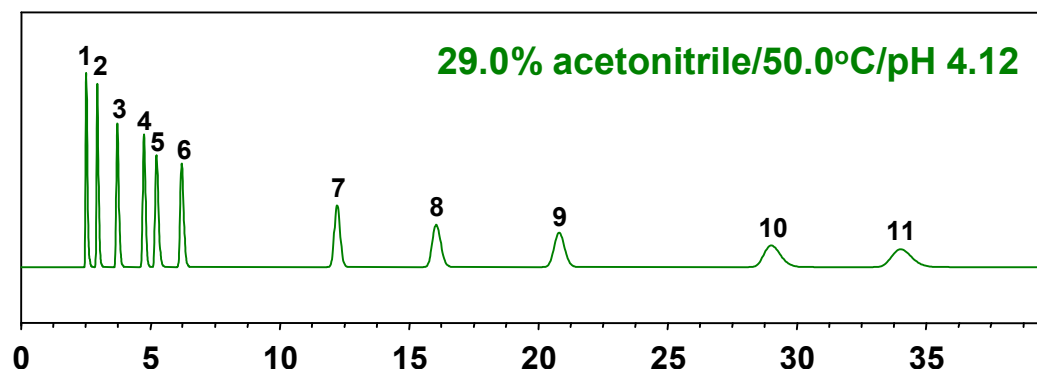
Simultaneous effects of temperature and pH

$$\log k = \log k_{HA} + \log\left[f + \frac{K h}{1 + K h}(1 - f)\right] = \log k_{HA} + \log\left(f + \frac{10^{\log K} h}{1 + 10^{\log K} h}(1 - f)\right) \qquad (4.23)$$

$$\log k = c_0 + \frac{c_1}{T} \qquad (5.11)$$

$$\log K = c_2 + \frac{c_5}{T} \qquad (5.13)$$

$$\log k = c_1 + \frac{c_2}{T} + \log\left(c_3 + \frac{10^{c_4 + \frac{c_5}{T}} h}{1 + 10^{c_4 + \frac{c_5}{T}} h}(1 - c_3)\right) \qquad (5.14)$$

The **model parameters** are characteristic for a given solute and separation system.

Simultaneous effect of temperature, pH and solvent content: polarity model

$$\log k = c_0 + c_1 P_M^N + \log\left(f + \frac{10^{(\log K_w + m\varphi)} h}{1 + 10^{(\log K_w + m\varphi)} h}(1 - f) \right) \qquad (4.29)$$

$$\log k = c_1 + \frac{c_2}{T} + \log\left(c_3 + \frac{10^{c_4 + \frac{c_5}{T}} h}{1 + 10^{c_4 + \frac{c_5}{T}} h}(1 - c_3) \right) \qquad (5.14)$$

$$\log k = c_1 + \frac{c_2}{T} + c_3 P_M^N + \log\left(c_4 + \frac{10^{c_5 + c_6\varphi + \frac{c_7}{T} + c_8\frac{\varphi}{T}} h}{1 + 10^{c_5 + c_6\varphi + \frac{c_7}{T} + c_8\frac{\varphi}{T}} h}(1 - c_4) \right) \qquad (5.15)$$

Eqs. (4.29) and (5.14) are particular cases of the general description given by Eq. (5.15).

5.4. Enhancing the predictions of retention

5.4.1. Considerations on the fittings

Reasons of fluctuation in peak parameters

Replicated measurements of retention times at a given mobile phase composition always fluctuate inside a certain range, due to several reasons.

- pump flow rate fluctuations

- irregularities in sample injection

- changes in temperature

- changes in power supply

- stationary phase degradation

- mobile phase mispreparation

(deviations from the nominal composition or evaporation of organic solvent)

t_R
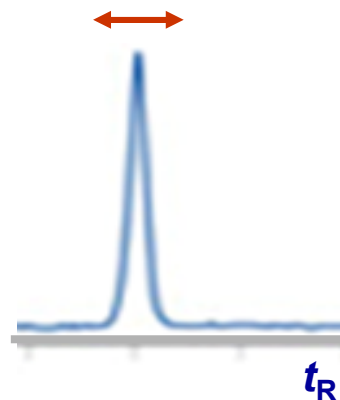
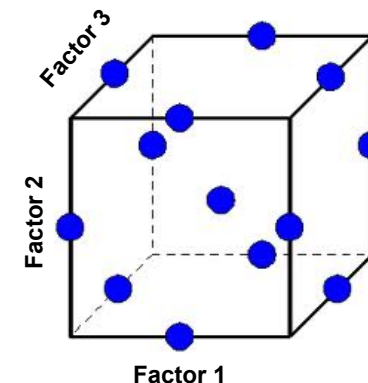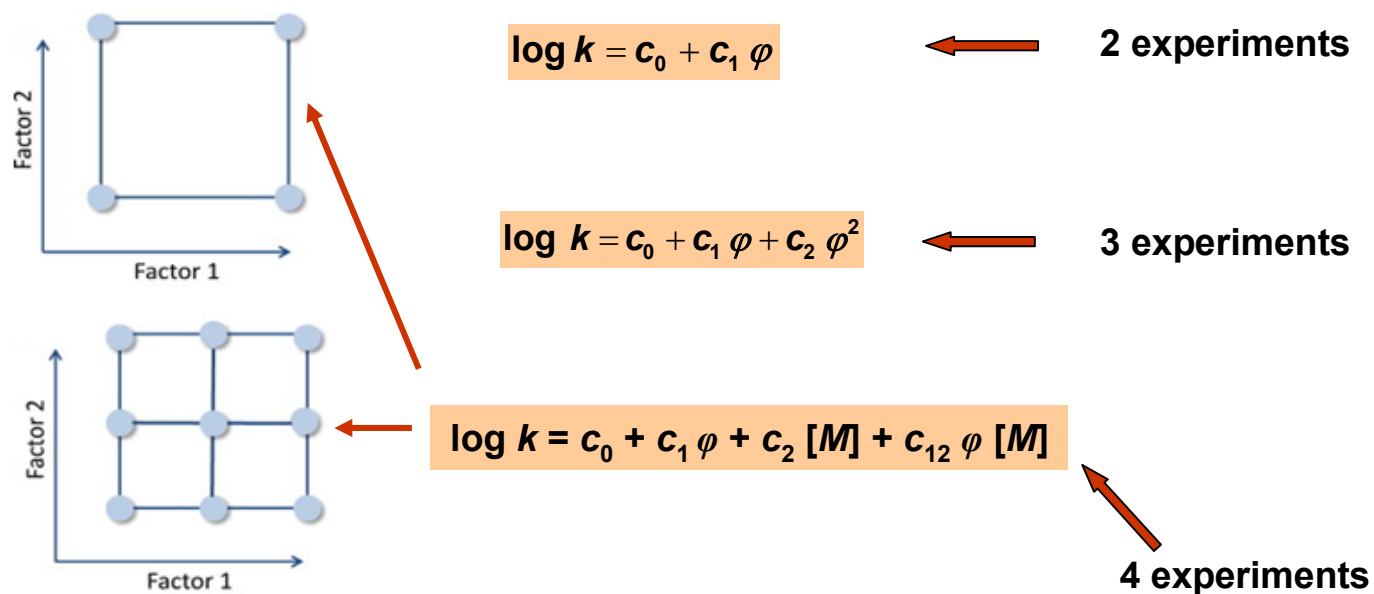These sources of error can deteriorate the accuracy of the predictions !!!

41

Peak tracking

● **When performing modelling experiments, the assignation of each peak to each solute**

   can be facilitated by the injection of standards of:

   ➡ **individual compounds**

   ➡ **mixtures of two or more compounds with separate retention** ⬅ **more practical**

● **If needed, peak tracking can be carried out:**

   ➡ **by varying the concentration of the injected standards**

   ➡ **with the aid of a selective detection technique**

   ➡ **by the assistance of Chemometrics**

Accuracy

● As a rule, with only the modifier and additive contents or temperature as factors, predictions are accurate, even using minimal experimental designs with only one degree of freedom.

● Often, no degree of freedom is left, since the reliability of the models has been extensively demonstrated.

$$\log k = c_0 + c_1\,\varphi$$ ⟵   **2 experiments**

$$\log k = c_0 + c_1\,\varphi + c_2\,\varphi^2$$ ⟵   **3 experiments**

$$\log k = c_0 + c_1\,\varphi + c_2\,[M] + c_{12}\,\varphi\,[M]$$

4 experiments

● When significant deviations are found in a further prediction, more experiments should be added to the design.

pH as a factor for ionisable compounds …

The complexity of modelling ionisable compounds (sudden and particular logarithmic drops) always recommends that the experimental design assures a good description of the factor space, independently of the number of parameters in the retention model.
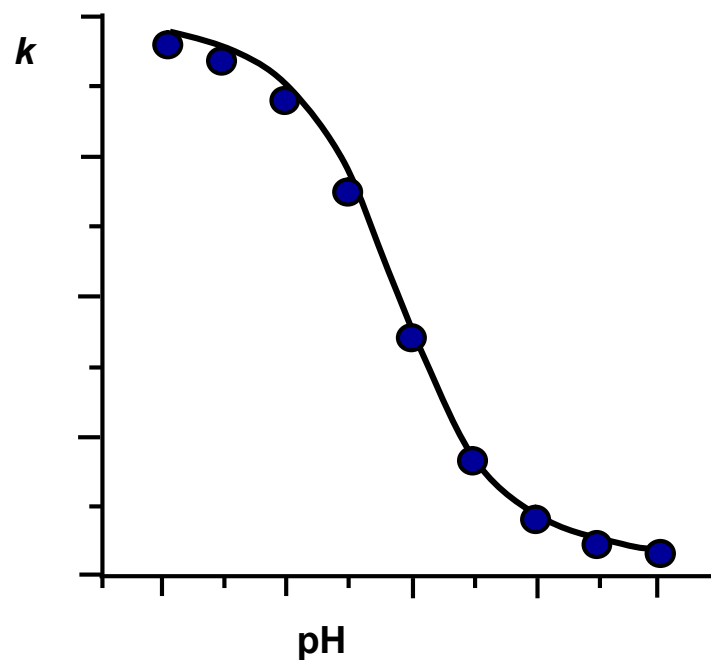
pH as a factor for ionisable compounds …

● **Non-significant terms are detrimental** for the fitting quality, and consequently, for predictions, especially when the retention models are empirical and the predictions are carried out far away from the experimental domain.

$$\frac{1}{k} = a_0 + a_1[M] + a_2\,\varphi + a_3\,\text{pH} + a_{12}[M]\,\varphi + a_{13}[M]\,\text{pH} + a_{23}\,\varphi\,\text{pH} + a_{123}[M]\,\varphi\,\text{pH} + a_{33}\,\text{pH}^2$$

● **The same experimental design must provide satisfactory information for several solutes.** Different solutes may demand different pH regions for good predictions.

● **The complexity** of finding common regions of high resolution grows with the number of solutes in the mixture.

● The fact that the retention factor of some solutes remain unchanged, while others suffer sudden drops, gives rise to multiple peak crossing and is translated in unrobust separations.

More experiments are needed with respect to those attending exclusively to the number of parameters in the fitted equations !!!
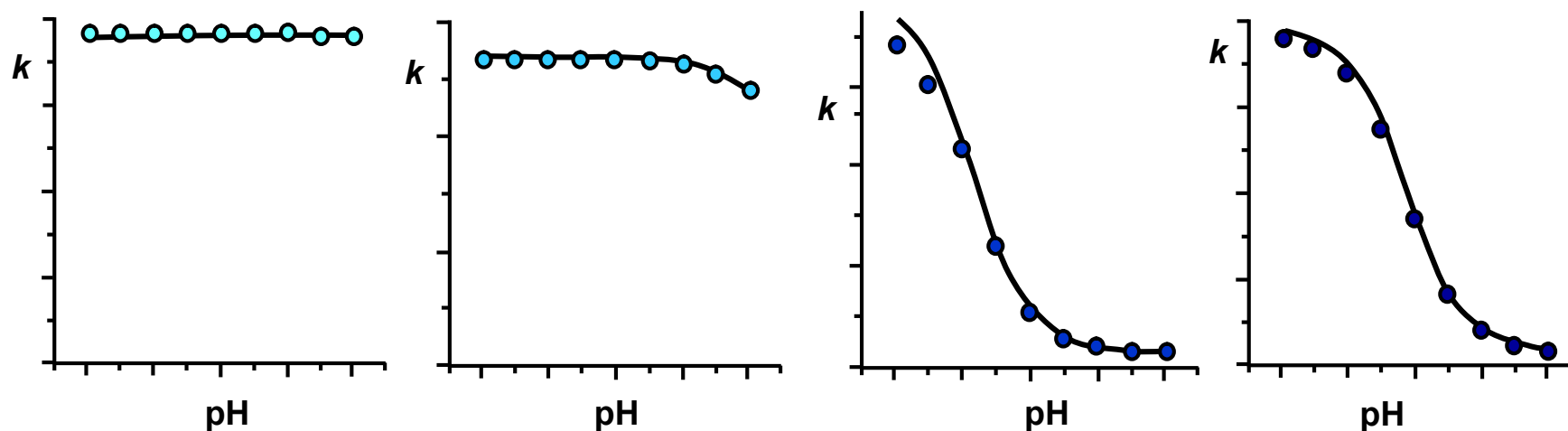
Reproducing a predicted optimal separation in the chromatograph often requires small corrections in composition in the pH direction.

5.4.2. Regression process

> The regression process builds the **best possible relationship between response and predictors.** This is often done by **minimising the sum of the squared residuals** (the squared difference between actual and predicted responses, extended to the whole set of experimental data).

● When **the response is transformed** to achieve a more convenient linear relationship (e.g. log k), the set of optimal (regressed) parameters will **minimise the residuals for the transformed response** (log k), but not for the original response (k or t_R), which is the actual interest.
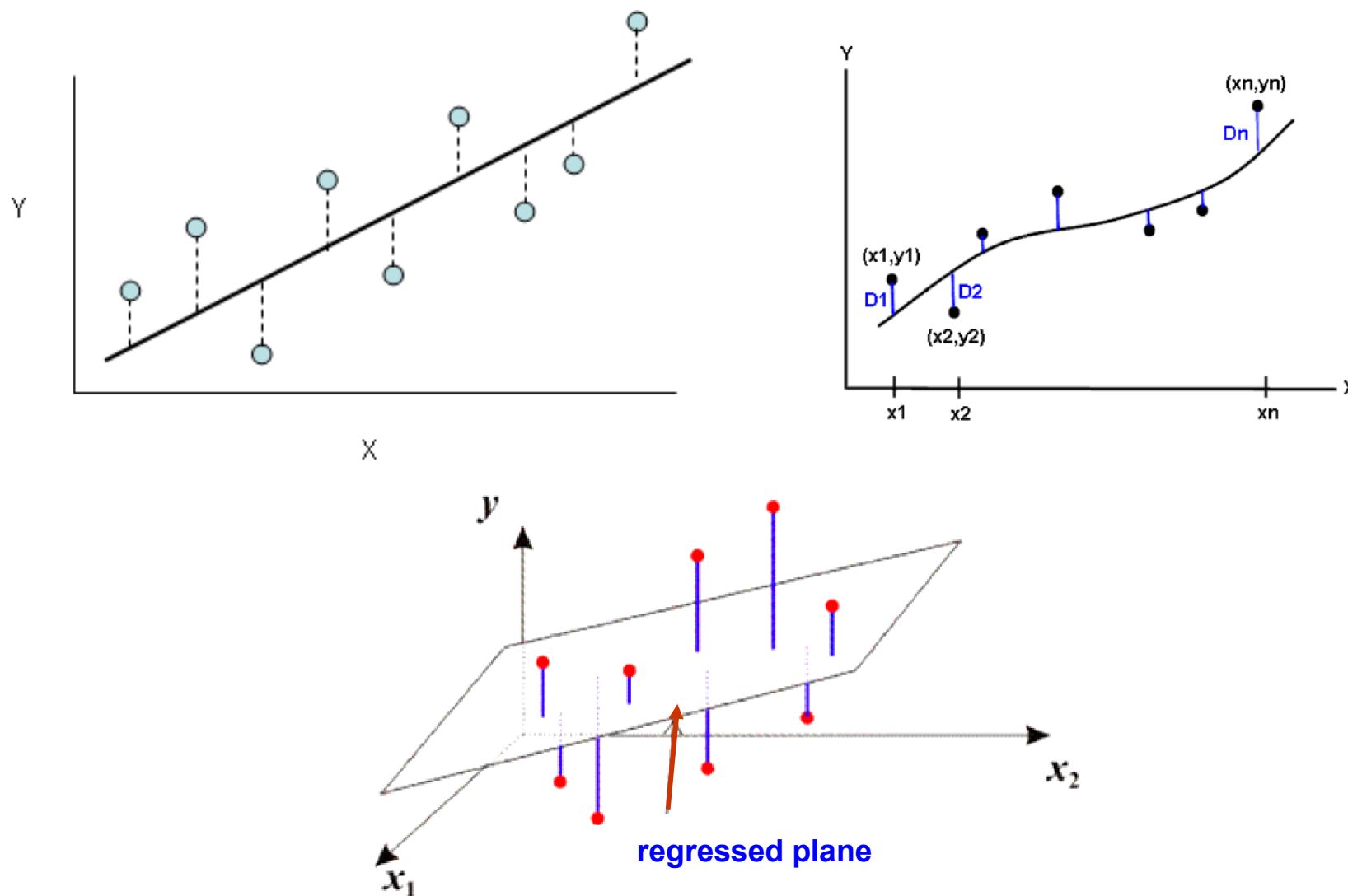
$$\log k = c_0 + c_1\,\varphi \qquad\qquad k = \frac{t_R - t_0}{t_0}$$

$$t_R = t_0\,(1+k) = t_0\,(1+k_0\,e^{c_1\varphi})$$

Least-squares regression

regressed plane

Unweighted regression

Ordinary unweighted linear least-squares regression gives rise a homoscedastic error distribution: the responses are predicted with a uniform absolute error.

- When a transformation is carried out, the unweighted linear regression will yield a homoscedastic error distribution for the transformed response (e.g. log k). Consequently, the error distribution for the original response will not be uniform.

- Since the relative influence in the fitting of data of large magnitude is decreased, predictions are deteriorated. When the retention covers almost 2 orders of magnitude, deviations between experimental and predicted retention factors may exceed 10 units for highly retained compounds.
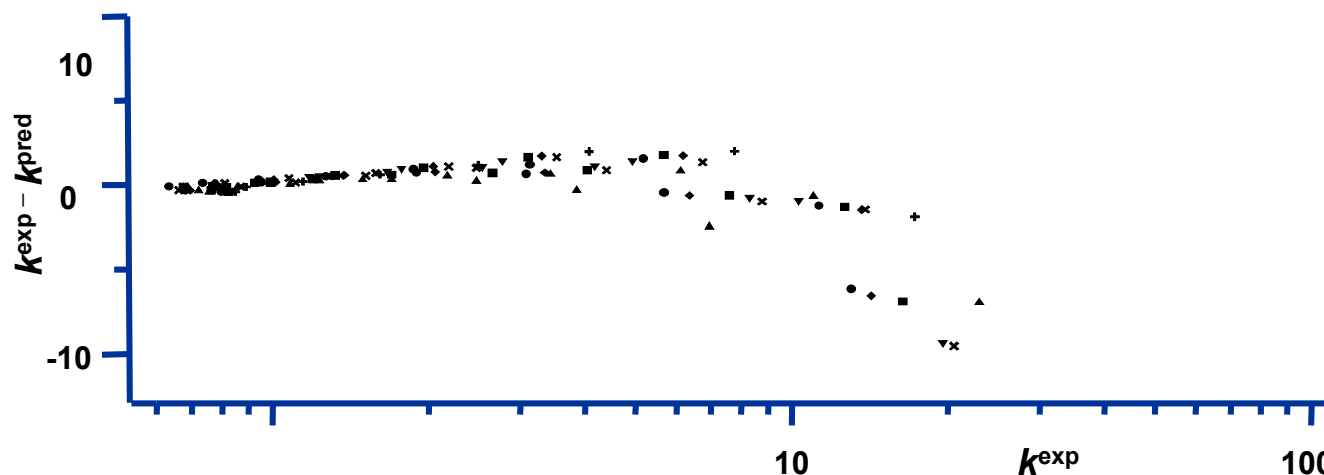
Weighted regression

The fitting bias can be compensated through the **application of weights,** which can be obtained by **applying the error theory.**

$$w = \left(\frac{\partial F}{\partial f}\right)^2 = \frac{1}{\left(\dfrac{\partial \log k}{\partial k}\right)^2} = (2.303\ k)^2 \qquad (5.16)$$

F : non-linearised function f : linearised function

- The use of weights is recommended as a general rule.

- Weighted regression, however, deteriorates the accuracy for faster eluents. Thus, if incidentally an optimal condition is predicted with a fast eluent, predictions should then be rebuilt using unweighted regression.
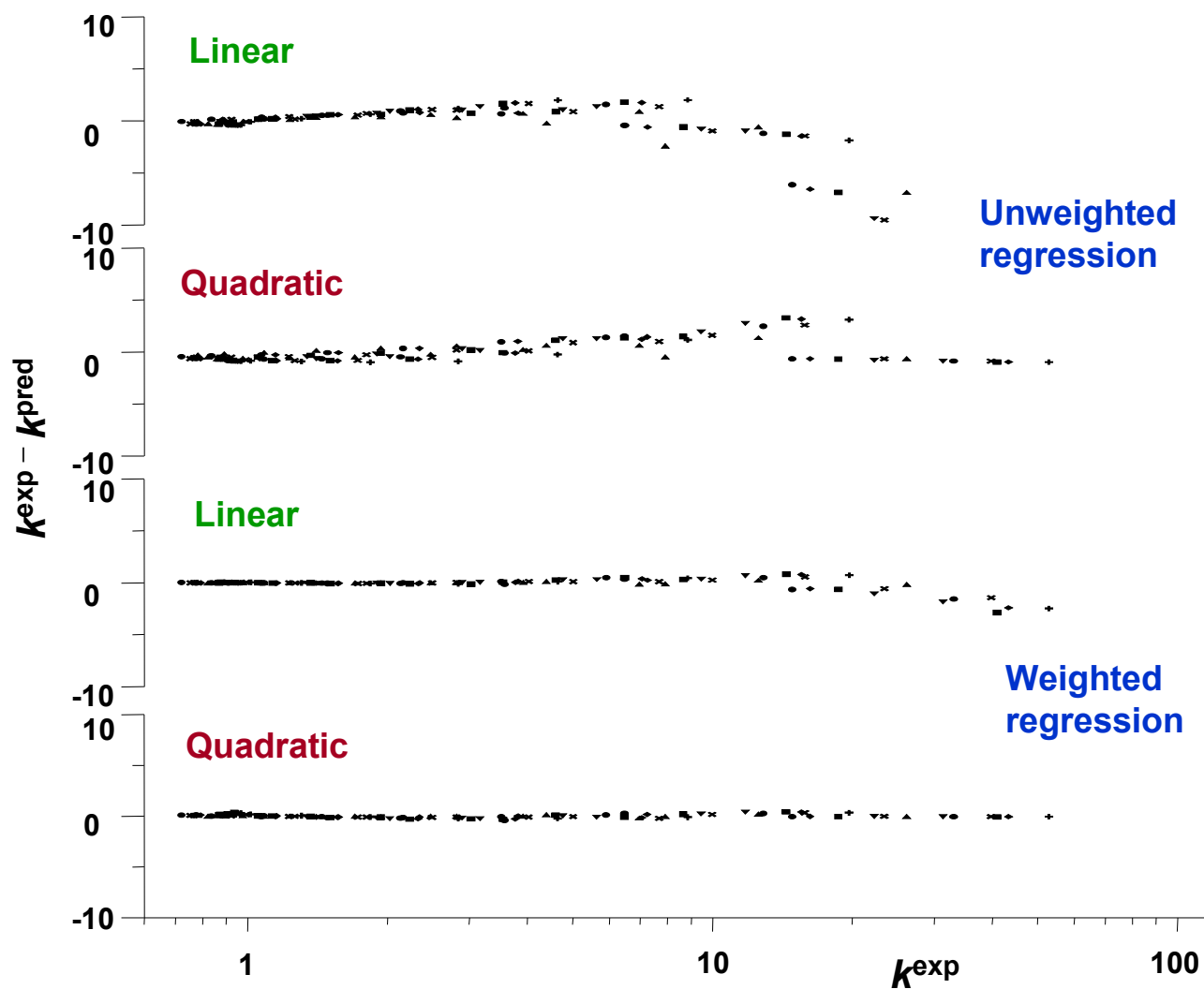
13 phenols eluted with 9 mobile phases in the range 20–100% ACN

5.4.3. Other approaches to enhance the predictions

● **Addition of new experiments to the training set:** For a *posteriori* improvement of the retention model. The new experiment should be carried out in the region of the predicted optimal conditions, so that the local accuracy is enhanced.

● **Use of reference compounds:** To compensate errors associated with mobile phase mispreparation. These compounds are a sort of internal standards aimed to correct not the solute concentration but the mobile phase composition. The references should be present in each sample injection and can be:

⇒ compounds not included in the sample,
    added to the injected solution

⇒ reagent needed in an incidental pre-column derivatisation reaction

⇒ by-products