Protonační rovnováhy léčiv faktorovou analýzou a nelineární regresí absorbanční responzní plochy

Ing. Sylva Bordovská,
Katedra analytické chemie, Univerzita Pardubice, 532 10 Pardubice,
sylva.bordovska@seznam.cz
a
Prof. RNDr. Milan Meloun, DrSc.,
Katedra analytické chemie, Univerzita Pardubice, 532 10 Pardubice,
milan.meloun@upce.cz

Souhrn: *Existuje široká škála technik pro analýzu spektrofotometrických dat jak pro kvalitativní, tak i pro kvantitativní účely. Jedná se především o techniky faktorové analýzy (FA) a metody hlavních komponent (PCA). Všechny tyto techniky jsou výpočtově náročné. Mají-li výše uvedené metody přinášet přesné a správné výsledky, je nutná vysoká kvalita vstupního experimentu a dobrá znalost fyzikálně-chemických vlastností systému.*

1) Určení počtu světlo-absorbujících částic

Metoda hlavních komponent (PCA): Pro určení počtu světlo-absorbujících částic lze použít metodu hlavních komponent[1,2,3]. Tato metoda je určena pro zjišťování počtu tzv. *latentních proměnných*, tj. proměnných, které popisují všechnu variabilitu datové matice A o n řádcích (v našem případě jednotlivá spektra) a m sloupcích (v našem případě vlnové délky) nezpůsobenou experimentální chybou. Počet světlo-absorbujících částic je právě roven počtu signifikantních latentních proměnných absorbanční matice.

2) Postup určení počtu latentních proměnných v PCA

2.1 Rozklad zdrojové matice: Prvním krokem metody hlavních komponent je rozklad zdrojové (v našem případě absorbanční) matice $\boldsymbol{A}$ rozměru n řádků (spekter) a m sloupců (vlnových délek) na součin matic $\boldsymbol{T}$ a $\boldsymbol{P}^{\mathrm{T}}$ a matici nepopsané variability E

$$\boldsymbol{A} = \boldsymbol{T}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{E} \qquad (1)$$

kde $\boldsymbol{T}$ je matice latentních proměnných (score matice) o rozměrech n x k, k jsou sloupcové vektory označované jako hlavní komponenty, matice $\boldsymbol{P}$ je matice zátěží (loadings matice) o rozměrech n x k, kde k jsou sloupcové vektory vyjadřující míru příspěvku dané latentní proměnné k variabilitě zdrojové matice.Výpočet matic $\boldsymbol{T}$ a $\boldsymbol{P}$ je založen na diagonalizaci kovarianční matice $\boldsymbol{Z}$

$$\boldsymbol{Z} = \boldsymbol{A}^{T}\boldsymbol{A}, \qquad (2)$$

čímž dostáváme vlastní čísla g_a matice $\boldsymbol{Z}$ a matici vlastních vektorů $\boldsymbol{Q}$, která je totožná s maticí $\boldsymbol{P}$. Matice $\boldsymbol{T}$ se pak dopočítá podle vztahu

$$\boldsymbol{T}^{\mathrm{T}} = (\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P})^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}. \qquad (3)$$

3) Regresní analýza A-pH spekter

Analýza spektrofotometrických dat za účelem určení počtu světlo-absorbujících částic v roztoku, protonačních konstant pK a jejich molárních absorpčních koeficientů $\varepsilon_{qr,\lambda}$ je předmětem nelineární regresní analýzy spekter pomocí programu SQUAD(84), SPECFIT/32. Absorbanční matice popisující k světlo-absorbujících částic v rovnovážné směsi je tvořena m roztoky pro n vlnových délek.

3.1 Hledané parametry: Nejdůležitějšími parametry, jež jsou předmětem odhadu programem SQUAD(84)[4] a SPECFIT/32[5], je protonační konstanta pK a molární absorpční koeficient $\varepsilon_{qr,\lambda}$ měřené vlnové délky. Oba tyto parametry jsou určovány pro všechny zúčastněné světlo-absorbující částice. Parametry pK a $\varepsilon_{qr,\lambda}$ jsou odhadovány na základě nelineární regresní analýzy absorbanční matice řádu n x m obdržené z experimentu.

3.2 Věrohodnost parametrů: Věrohodnost určených parametrů lze posoudit na základě dosažené těsnosti proložení vypočtené regresní spektrální křivky experimentálními body statistickou analýzou reziduí[6]. U nejtěsnějších proložení vykazují rezidua Gaussovské normální rozložení a jejich velikost není větší než $s_{inst}(A)$. Analýzu reziduí popisují následující statistiky:

a) Aritmetický průměr reziduí musí byt roven nule, $\bar{e} \approx 0$.

b) Průměrná reziduum musí byt menší nebo rovno než $s_{inst.}(A)$.

c) Koeficient šikmosti rozdělení reziduí musí být roven 0.

d) Koeficient špičatosti rozdělení reziduí musí být roven 3.

e) Hamiltonův R-faktor[6] menší než 1%.

I když rezidua nevykazují vždy Gaussovo rozložení, je jejich analýza užitečnou pomůckou k objektivnímu posouzení těsnosti proložení a výběru nejlepší testované hypotézy rovnovážného modelu.

4) Debyeův-Hückelův vztah

Protonační rovnováhu lze popsat koncentrační, smíšenou nebo termodynamickou rovnovážnou konstantou. Smíšená rovnovážná konstanta je pro q = 1 charakterizována následujícím vztahem

$$\beta_{1r} = \frac{[\mathrm{LH}_r]}{[\mathrm{L}]a_{H^+}^r} = K_{11} \cdot K_{12} \ldots\ldots K_{1r} . \qquad (4)$$

Vztah mezi protonační smíšenou a koncentrační konstantou pomocí následující rovnice

$$\log\ K_{1r,smíš} = \log\ K_{1r,c} + \log \gamma_{H^+} . \qquad (5)$$

Pro Debye-Hückelovu rozšířenou rovnici je možné z definice rovnovážné konstanty a na základě definice aktivitního koeficientu vyjádřit vztah závislosti smíšené protonační konstanty na iontové síle, kde pro nízké iontové síly $I \leq 0.1$ lze vztah zjednodušit na tzv. zkrácený Debyeův-Hückelův zákon

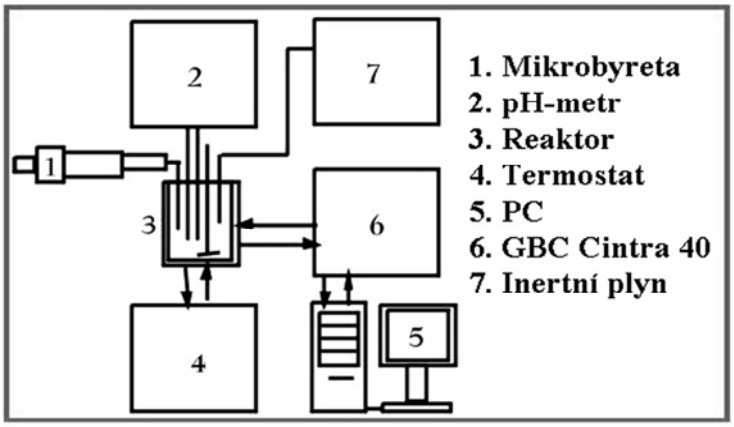
$$\log\ K_{1r,smíš} = \log\ K_{1r}^T - \frac{A|z_+ z_-|\sqrt{I}}{1 + Bå\sqrt{I}} . \qquad (6)$$

Technické zabezpečení experimentu

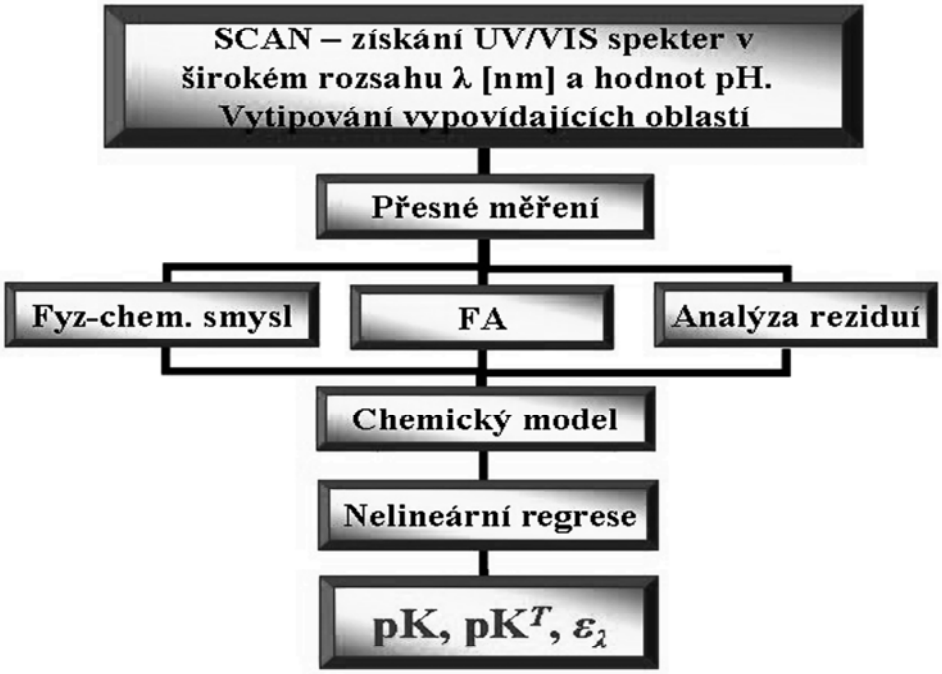
Přístroje
- UV/VIS spektrofotometr GBC Cintra 40
- pH metr Radelkis OP-271
- Termostat
- Mikrobyrety

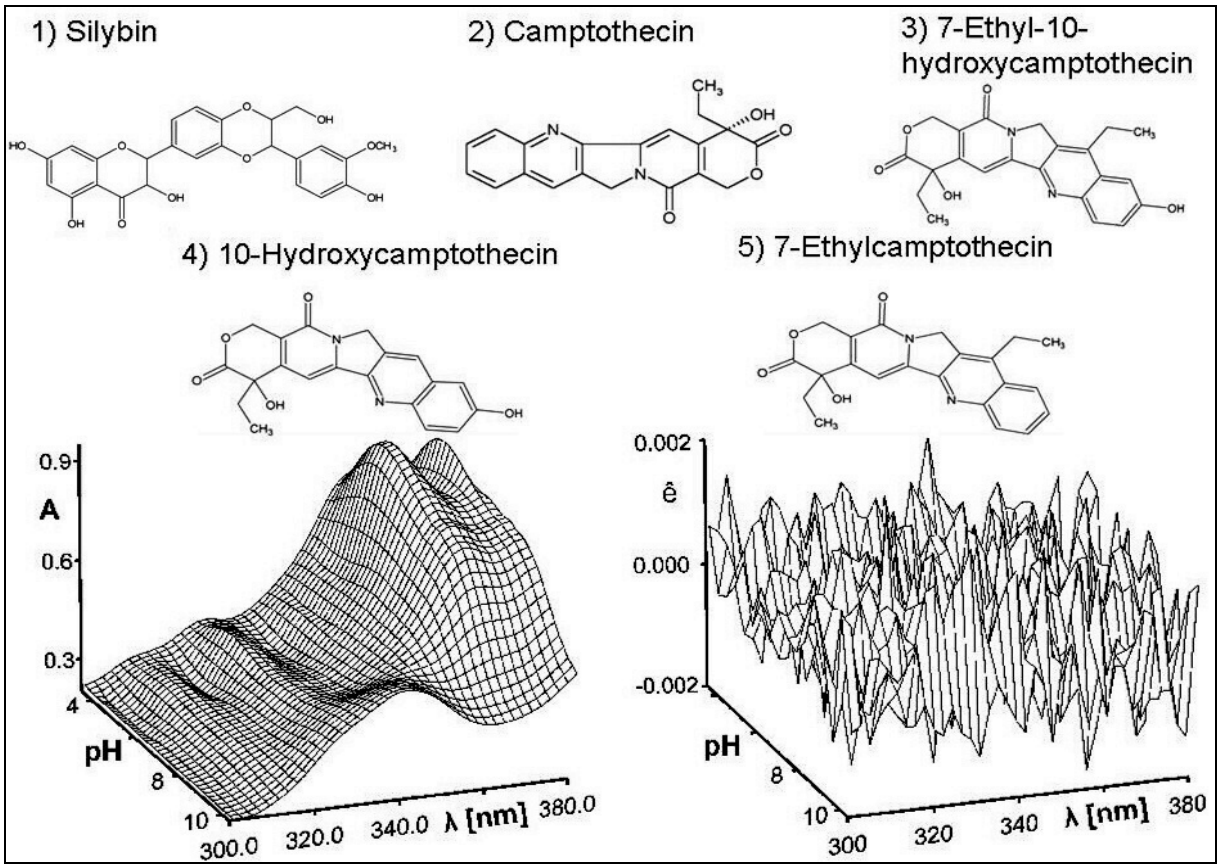
Chemikálie
- 1M HCl
- 1M KOH
- Standartní pH pufry
- Léčiva

Programy
- SQUAD(84)
- SPECFIT/32
- ADSTAT 1.25
- S-Plus
- MS EXCEL 2002
- Microcal Origin



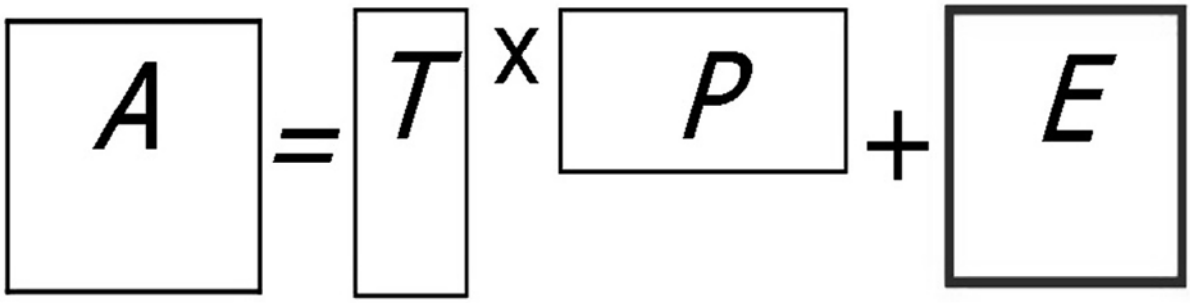
Strategie experimentu





Faktorová analýza - FA

Rozklad zdrojové matice
Matice A o rozměrech n x m, kde $A = TP^{T} + E$, kde T má rozměr n x k a P *rozměr k x m.*



Výpočet matic T a P se provádí z kovarianční matice Z, kde $Z = A^{T}A$.
Matice latentních proměnných lze vypočítat ze vztahu: $T^{T} = (P^{T} P)^{-1} P^{T} A^{T}$

Krátký cyklus
Rekonstrukce datové matice použitím prvních k nejvýznamějších latentních proměnných. $A^{pred} = T P^{T}$

Metody faktorové analýzy určující počet latentních proměnných

Kankareho metoda	**Metoda $\chi 2$- kriteria**	Reziduální variace	**Malinowského F-test**
Metoda reálné chyby	Metoda sm.odch. vlastních čísel	Metoda log. vlastních čísel	*Bartlettův test izotropie*
Metoda extrahované chyby	*Metoda průměru vlastních čísel*	Metoda zabudované chyby	
Metoda průměrné chyby	*Metoda Exnerovy funkce*	Faktorová indikátorová funkce	

Tučně – spolehlivé metody, *kurzíva* – selhávající metody

Kritéria pro určení počtu latentních proměnných:

1) **Bod zlomu křivky**
2) **Precizní metody – reziduálová směrodatná odchylka**
3) **Aproximativní metody – testační kritéria**
4) **Druhá derivace, třetí derivace, poměr derivací**

Kankareho metoda

Vychází z druhého momentu M analyzované absorpční matice *A*.

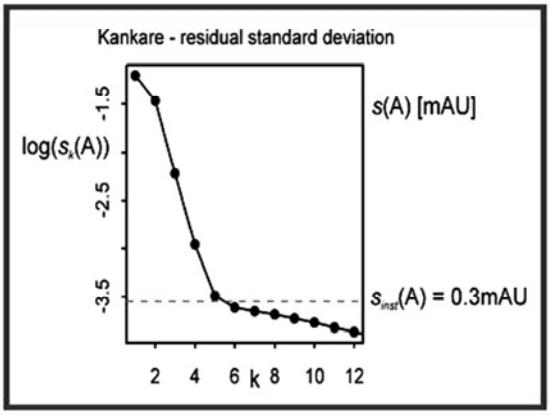
$$\mathbf{M} = \frac{1}{n_W} A^T A$$

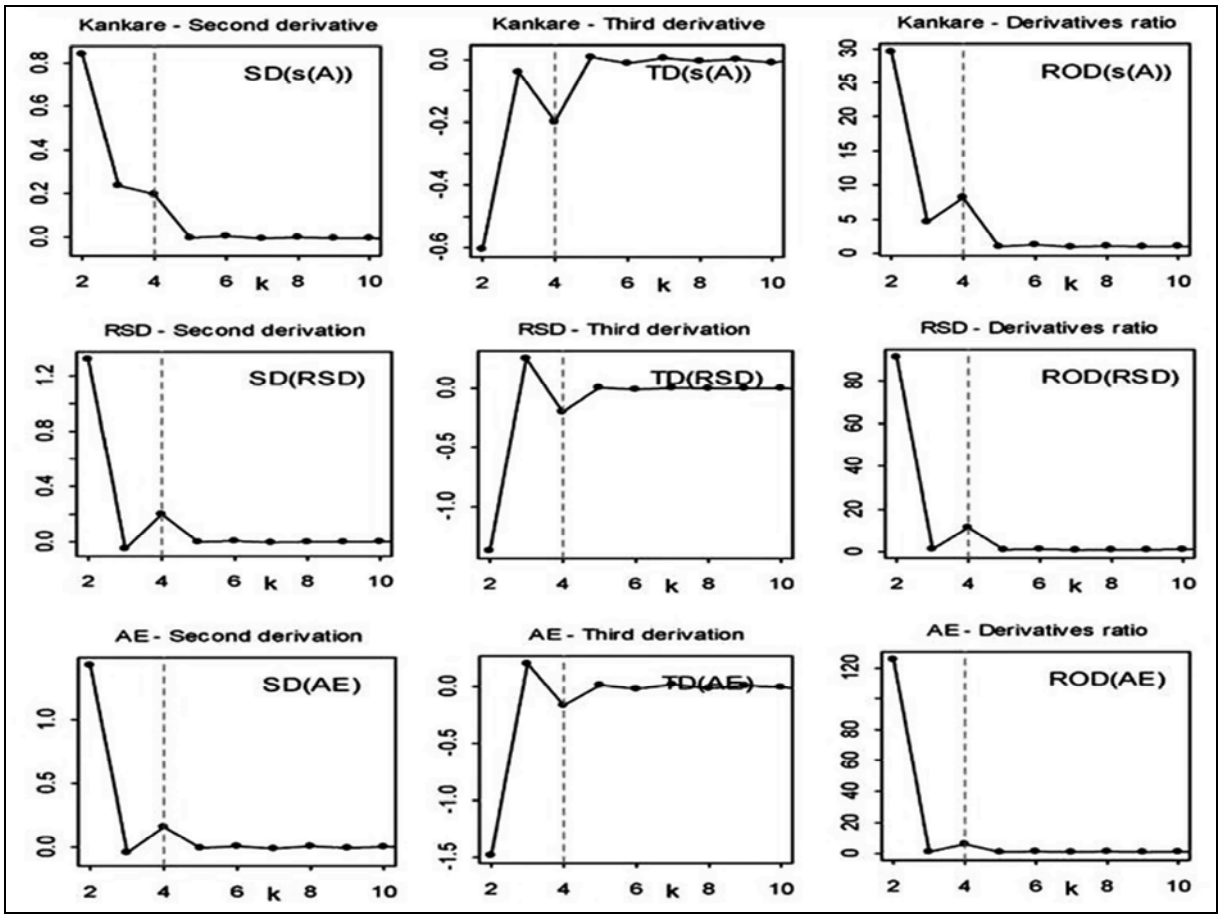
Předpoklad: Prvky matice jsou zatíženy experimentálním šumem jehož úroveň je dána $s_{inst}(A)$.

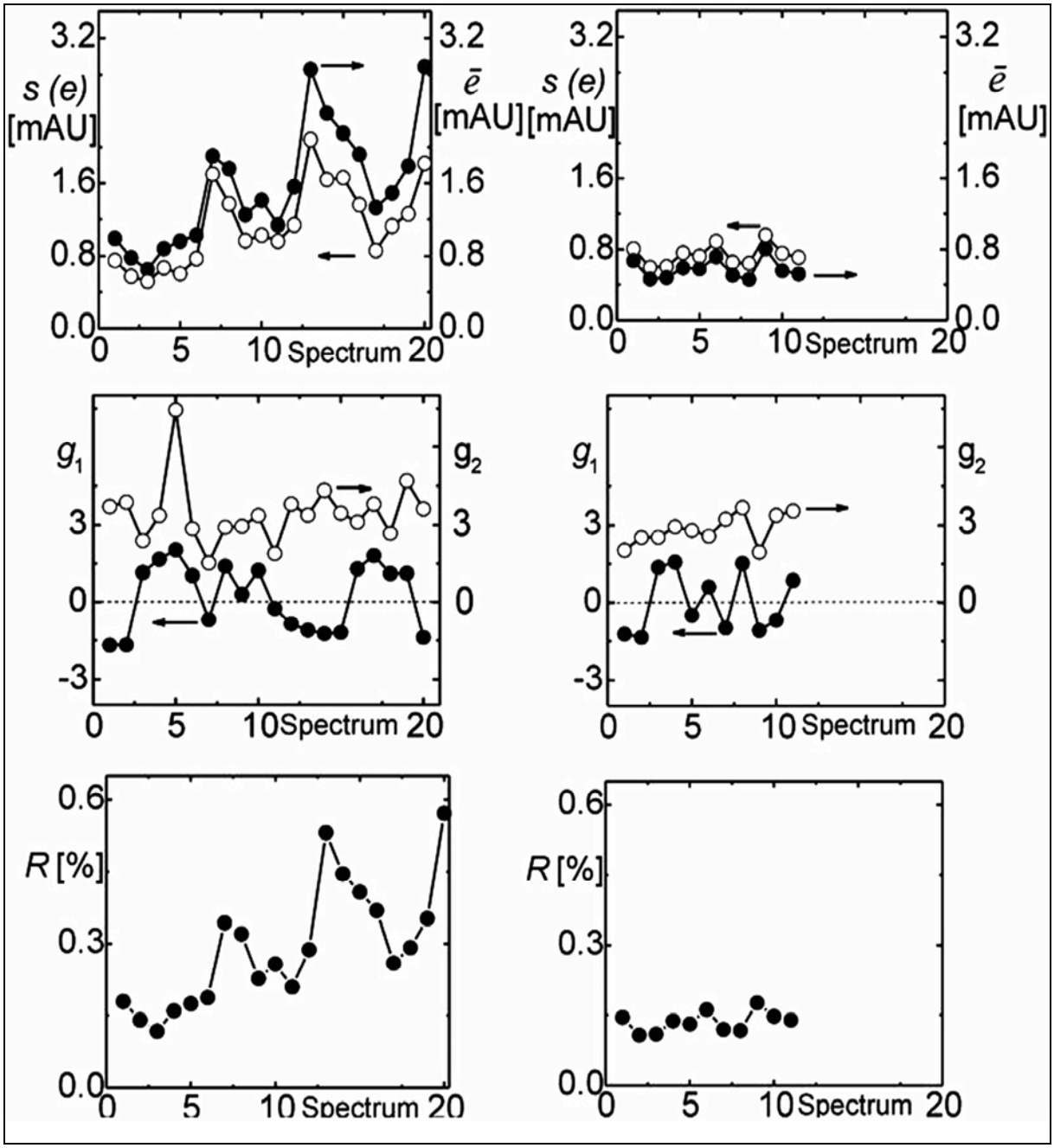
Zbytková reziduálová chyba $s_k(A)$

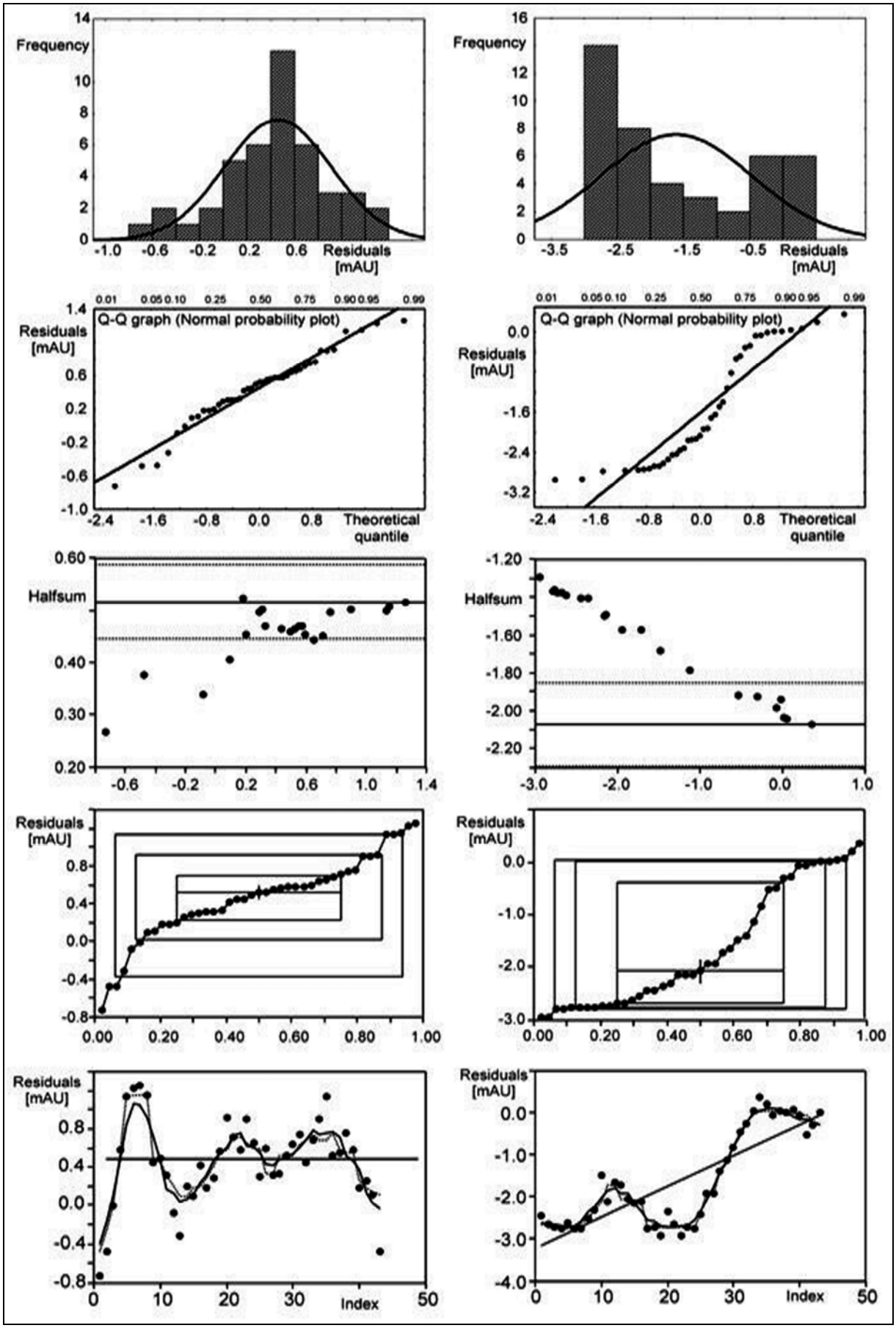
$$s_k(A) = \sqrt{\frac{tr(\mathbf{M}) - \sum_{a=1}^{k} r_a}{n - k}}$$

Za hodnost matice je pokládána ta hodnota, kdy platí $s_k(A) = s_{inst}(A)$, přičemž hodnosti matice je roven počet latentních proměnných.

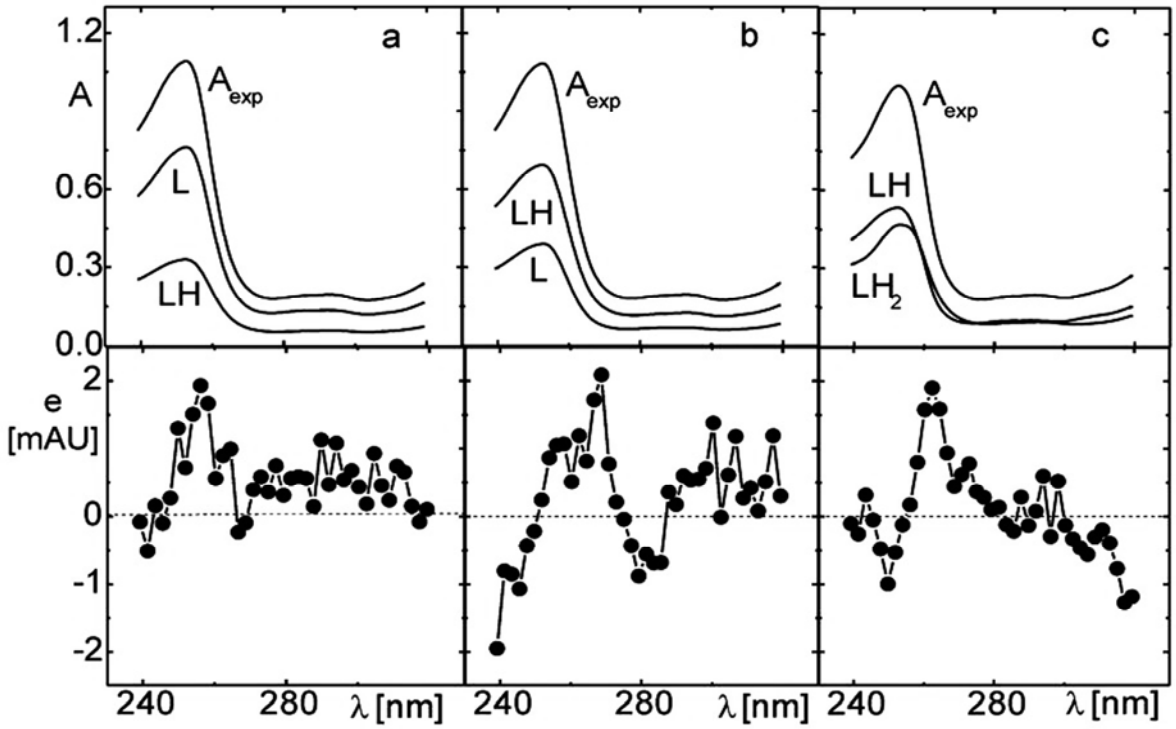




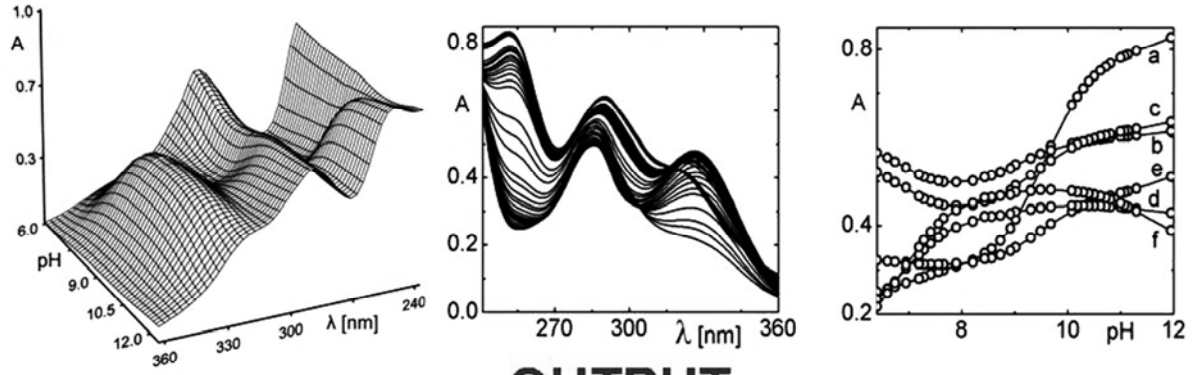

s (e)
[mAU]
ē
[mAU]
3.2
1.6
0.8
0.0
0
5
10
Spectrum
20
g1
g2
3
0
-3
R [%]
0.6
0.3


Frequency
Residuals [mAU]
Q-Q graph (Normal probability plot)
Theoretical quantile
Halfsum
Index

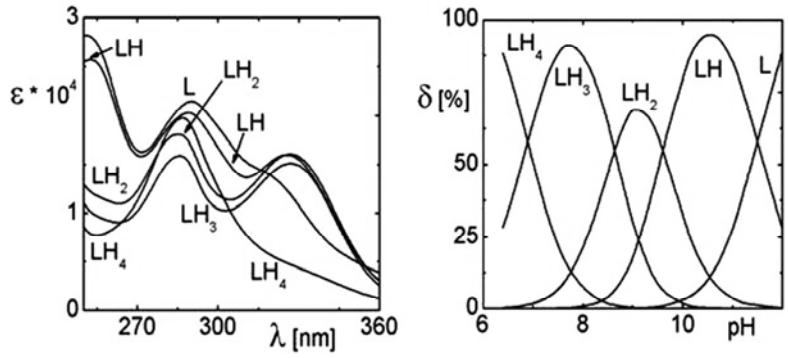
Dekonvoluce spekter



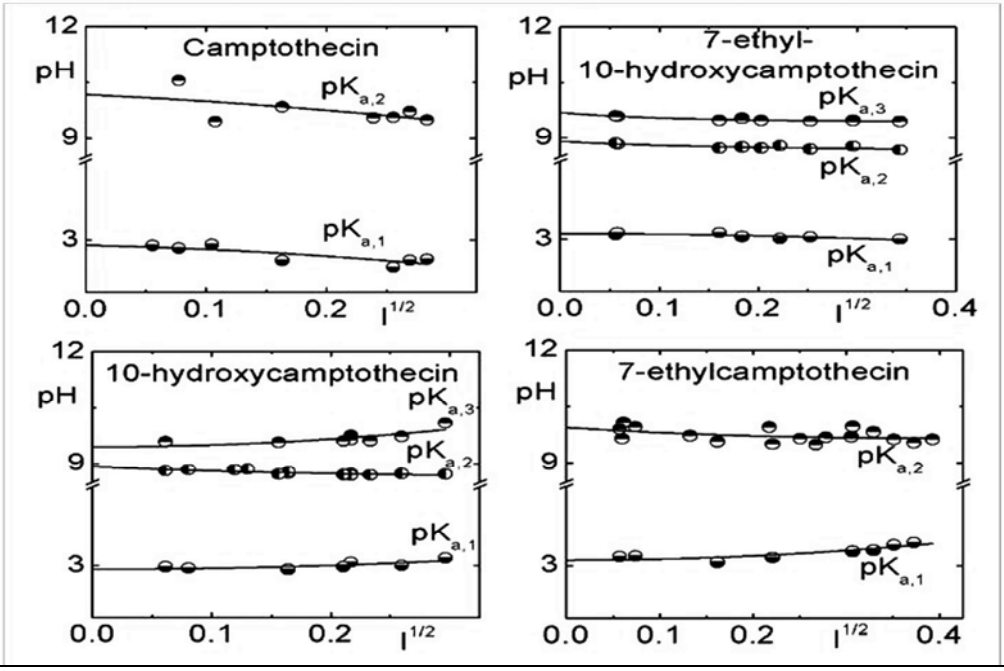
INPUT



OUTPUT



pK_a při různých iontových silách Debye-Hückelovým vztahem



Závěr

Přehled pK léčiv: **(v závorce směrodatná odchylka posledních platných číslic) .**

		SPECFIT/32		SQUAD(84)	
		Value at 25°C	Value at 37°C	Value at 25°C	Value at 37°C
Camptothecine	$pK^T_{a,1}$	2.90(7)	3.02(8)	2.83(9)	2.92(8)
	$pK^T_{a,2}$	10.18(30)	10.23(8)	10.11(36)	10.43(3)
7-Ethyl-10-hydroxycamptothecine	$pK^T_{a,1}$	3.11(2)	2.46(6)	3.04(5)	2.30(6)
	$pK^T_{a,2}$	8.91(4)	8.74(3)	8.90(3)	8.84(3)
	$pK^T_{a,3}$	9.70(3)	9.47(8)	9.71(5)	9.53(10)
10-Hydroxycamptothecine	$pK^T_{a,1}$	2.93(4)	2.84(5)	2.92(4)	2.77(5)
	$pK^T_{a,2}$	8.93(2)	8.92(2)	8.93(3)	8.90(2)
	$pK^T_{a,3}$	9.45(10)	9.98(4)	9.46(9)	10.02(7)
7-Ethylcamptothecine	$pK^T_{a,1}$	3.10(4)	3.30(16)	2.94(3)	3.26(22)
	$pK^T_{a,2}$	9.94(9)	10.98(18)	9.73(9)	10.96(18)
Silybin	$pK^T_{a,1}$	6.99(4)	6.84(4)	7.00(4)	6.86(5)
	$pK^T_{a,2}$	8.77(11)	8.76(3)	8.77(11)	8.77(3)
	$pK^T_{a,3}$	9.57(8)	9.62(1)	9.57(8)	9.62(1)
	$pK^T_{a,4}$	11.66(3)	11.37(1)	11.66(3)	11.38(1)

Poděkování: Práce vznikla za podpory grantu Ministerstva zdravotnictví NR9055-4/2004 a vědeckých záměrů Ministerstva školství, kultury a mládeže MSMT0021627502.

Doporučená literatura:

1. Čepan M., Pelikán P., Liška M.: *Chem. Listy* 1991, 85, 1057.
2. Pytela O.: *Chemometrie pro organické chemiky*, VŠCHT Pardubice 1993.
3. Malinowski E. R.: *Anal. Chem.* 1977, 49, 606.
4. Meloun M., Havel J.: *Computation of Solution Equilibria 1. Spectrophotometry*, UJEP Brno 1985.
5. SPECFIT/32: *Spectrum Software Associates*, 197M Boston Post Road West, Marlborough, MA, 01752 U.S.A. (2004), (http://www.bio-logic.info/rapid-kinetics/specfit.html).
6. Meloun M., Militký J., *Statistické zpracování experimentálních dat v chemometrii, biometrii, ekonometrii a v dalších oborech přírodních, technických a společenských věd*, East Publishing, a.s., Praha, 1998, s. 589-611 s. 680, Academia 2004.
7. Meloun M., Militký J. , Forina M.: *Chemometrics for Analytical Chemistry, Volume 1. PC-Aided Statistical Data Analysis,* Ellis Horwood, Chichester 1992.
8. Brereton R. G. *Multivariate Pattern Recognition in Chemometrics*, Illustrated by Case Studies, Elsevier 1992,
9. Krzanowski W. J.: *Principles of Multivariate Analysis, A User's Perspective*, Oxford Science Publications 1988,
10. Meloun M. , Militký J., *Statistické zpracování experimentálních dat*, Plus Praha 1994, Academia Praha 2004..
11. Meloun M. , Militký J., Hill M., *Počítačová analýza vícerozměrných dat v příkladech,* Academia Praha 2005.
12. Meloun M., Militký J., *Kompendium statistického zpracování experiment-tálních dat*, Academia Praha 2002, 2006.