

Hledání chemického modelu regresní analýzou VIS spekter

Milan Javůrek¹ a Milan Meloun²

¹ Katedra řízení procesů, Fakulta elektrotechniky a informatiky, Univerzita Pardubice,
Milan.Javurek@upce.cz

² Katedra analytické chemie, Fakulta chemicko-technologická, Univerzita Pardubice,
Milan.Meloun@upce.cz

Abstrakt: *Určení chemického modelu, tj. stechiometrie částic, jejich koncentrace a disociační konstanty (nebo konstanty stability), je velmi důležité pro posouzení vlastností a acidobazického chování látek v roztocích. V článku je řešena věrohodnost stanovení při použití nelineární regrese absorbančních responzních ploch.*

1. Úvod

Záporný dekadický logaritmus disociační konstanty při porovnání se stupnicí pH určuje, zda se daná látka chová jako kyselina příp. zásada. Jsou-li jednotlivé komponenty barevné, nejčastěji se používá pro svoji jednoduchost, spolehlivost i přesnost metoda měření a analýzy spekter. Pro různé koncentrace jednotlivých komponent a proměňované hodnoty pH v potřebném rozsahu jsou pak naměřena celá spektra ve viditelné oblasti. Nevýhodou této metody je omezená platnost Lambertova – Beerova zákona, kdy můžeme měřit pouze v určitém rozsahu koncentrací a pak některé z částic nemusí být identifikovatelné pro svou malou koncentraci. V těchto případech musíme použít jinou metodu – např. potenciometrickou titraci, kdy rozsah koncentrací může být o několik řádů větší. Text této přednášky vychází z článku [1], kde lze najít další podrobnosti vč. kompletního přehledu literatury.

2. Data a model

Acidobazické chování látek v roztocích popisuje Guldbergův – Waageův zákon, který charakterizuje rovnovážný stav popsany rovnicí:



pomocí protonační (disociační resp. deprotonační pro opačný průběh) konstanty:

$$K_H = \frac{[HL^z]}{[L^{z-}][H^+]} \quad (2)$$

Hranaté závorky vyjadřují rovnovážný (ustálený) stav. Tak se v roztoku mohou vyskytovat různě protonizované částice látky: L, HL, H₂L, H₃L, tj. obecně H_rL, takže lze definovat konstanty stability libovolně protonovaných částic:

$$\beta_{qr} = \frac{[L_q H_r]}{[L]^q [H]^r} \quad (3)$$

Každá z těchto částic může absorbovat světlo, tj. má své vlastní spektrum. Absorpci světla jednotlivými částicemi popisuje Lambertův – Beerův zákon:

$$A_{ij} = \sum_{n=1}^{n_c} \varepsilon_{j,n} c_n \quad (4)$$

kde $\varepsilon_{j,n}$ je tzv. molární absorpční koeficient charakteristický pro látku n a vlnovou délku j a c_n je koncentrace látky n . Veličinu $A_{j,n}$ nazýváme absorbance, vyjadřuje nám míru barevnosti roztoku, resp. míru absorpce světla roztokem.

Při měření spekter pak získáváme celkovou absorbanci, tj. vlastně součet příspěvků absorbancí světla jednotlivými částicemi.

3. Metoda

Předpokladem k určení chemického modelu je znalost skutečného počtu barevných částic v roztoku. Absorbanční matice pro n koncentrací a j vlnových délek je podrobena faktorové analýze podle Wernimontovy – Kankareho metody [2]. Zde zlom na Cattelově grafu směrodatné odchylky reziduí absorbance určuje počet barevných částic.

Při vlastním regresním výpočtu použijeme metodu nejmenších čtverců, tj. počítáme sumu odchylek absorbancí experimentálních a vypočtených. Vypočtená absorbance se vyčísľuje jako součet příspěvků jednotlivých látek, přičemž koncentrace odpovídajících částic se počítají z rovnice (3) pro daný odhad konstanty stability. Doplnkovou podmínkou je znalost celkových koncentrací jednotlivých komponent. Rovněž iteračním způsobem se dopočítávají molární absorpční koeficienty těchto částic.

Východiskem pro každý iterační krok jsou tedy odhadnuté hodnoty disociačních konstant (tj. pro zvolený chemický model). Zlepšující kroky hodnot těchto konstant se počítají z celé matice absorbancí, což v případě derivační metody představuje součty parciálních derivací ve všech bodech (derivujeme podle parametrů, tj. hodnot disociačních konstant).

4. Verifikace výsledků

Výše popsaným způsobem lze provést více výpočtů pro různé varianty chemického modelu. Zde potřebujeme jednoznačná kritéria pro porovnání dosažených výsledků, resp. dosažené těsnosti proložení regresního výpočtu.

Prvotní je fyzikální významnost parametrů (hodnot disociačních konstant), jejich odhadnuté chyby z nelineární regrese a spočtené volné koncentrace jednotlivých částic. Koncentrace nižší než 5% se zpravidla zanedbávají, tj. příslušná částice nemá fyzikálně chemický význam.

Pro hodnocení kvality proložení máme v případě nelineární regrese k dispozici pouze statistickou analýzu reziduí. Jelikož však data jsou tvořena celou maticí, je možné analyzovat rezidua buď po řádcích (pro určitou koncentraci přes všechny vlnové délky) nebo po sloupcích (přes všechny koncentrace po jednotlivých vlnových délkách). Užívanější je postup první. Vypočítávají se statistické charakteristiky, tj. všechny čtyři centrální momenty, medián a Hamiltonův R-faktor, které nám dostatečně vypovídají o parametrech polohy a tvaru daného souboru reziduí. Zajímavé je porovnání aritmetického průměru a mediánu, které by se pro normální rozdělení měly shodovat.

Kromě klasických reziduí je také možno užít i jiné varianty – zde připadají v úvahu standardizovaná a Jackknifé rezidua, která se používají k identifikaci odlehlých bodů. Zde ovšem musíme ještě dopočítávat projekční matici $\mathbf{P} = \mathbf{J} (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T$ (kde \mathbf{J} je Jakobián, matice prvních parciálních derivací, viz [3]), kdy se potřebné derivace opět sčítají po řádcích absorbanční matice.

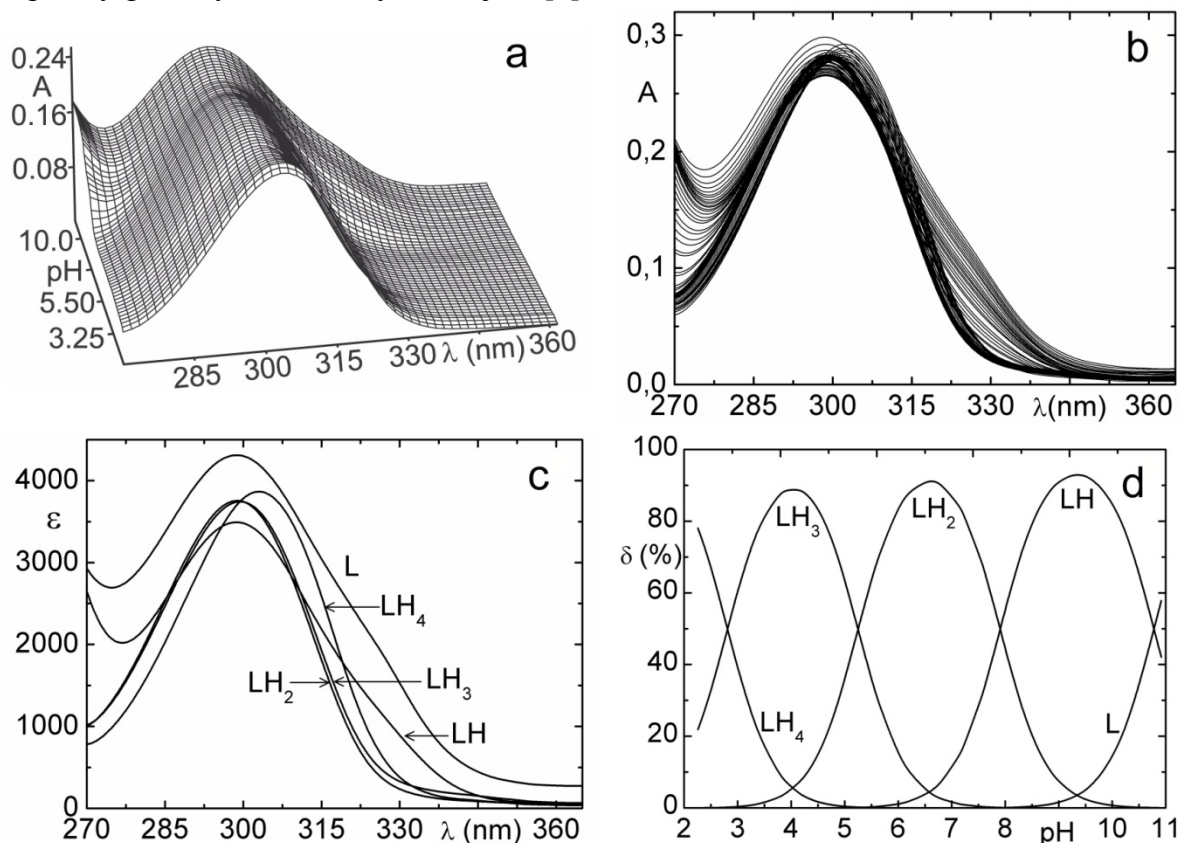
5. Simulace dat

Velmi významným nástrojem pro ověření kvality výpočtu je zařazení možnosti vygenerování syntetických dat pro předvolené hodnoty koncentrací a parametrů. Tato matice absorbancí je zatížena náhodnými chybami s normálním rozdělením a následně analyzována stejně jako data experimentální. Shoda předvolených a nalezených hodnot parametrů včetně

analýzy reziduí je jednoznačným kritériem při posouzení především správnosti výpočtu. Tím je také vytvořena platforma pro hodnocení různých algoritmů.

6. Testování algoritmů

Pro analyzování spektrálních dat byla publikována řada algoritmů. Na našem pracovišti je po řadu let s úspěchem používán program SQUAD(84), viz [4]. Dále se podařilo získat program SPECFIT/32, viz [5], který je podstatně novější. Bylo provedeno jejich testování jak na autentických experimentálních datech, tak na sérii dat simulovaných s různou hodnotou směrodatné odchylky souboru chyb s normálním rozdělením, kterými jsou zatěžována vygenerovaná data. Celý průběh testování, shrnutý do přehledné tabulky a doplněný grafickými schématy, lze najít v [1].



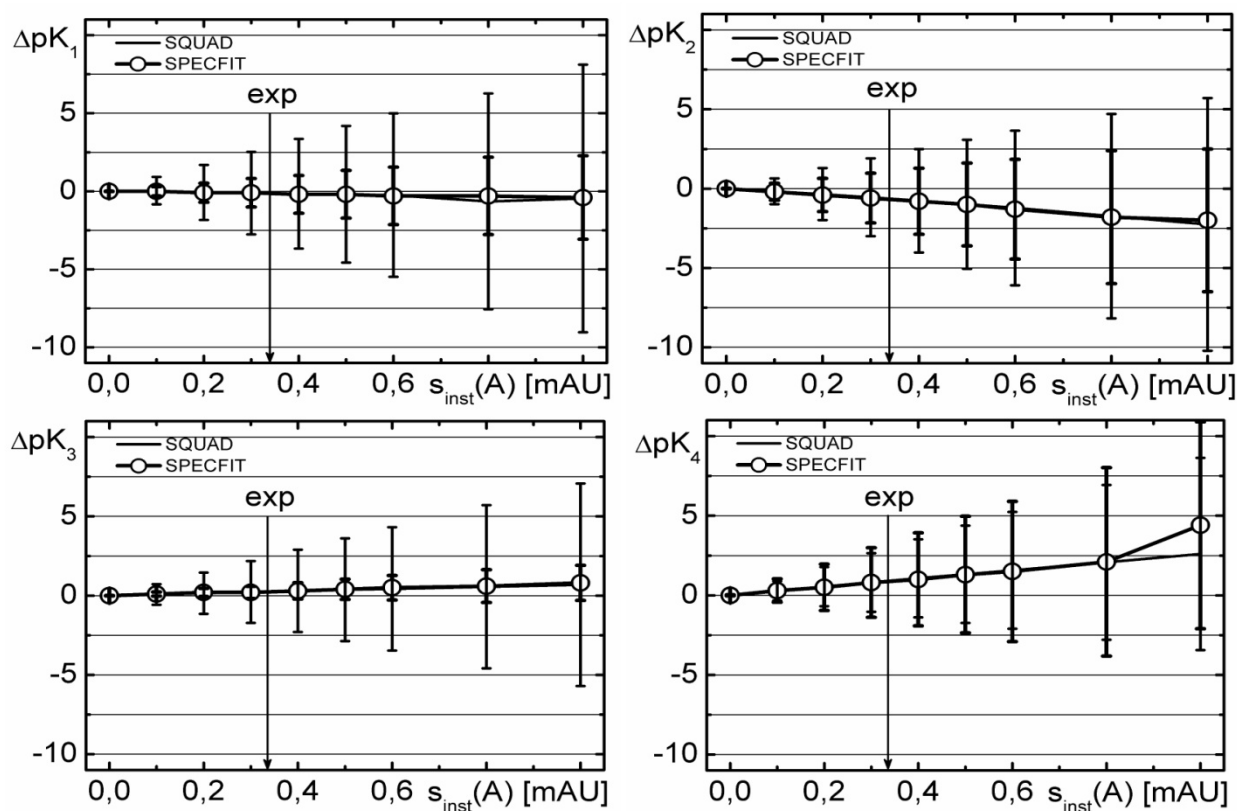
Obr. 1 (a) Absorbanční responzní plocha čili (b) vėjř spekter tvořř data, (c) křřvky molárnřch absorpčních koeficientů a distribuční diagram částic protonačního modelu tvořř vřsledek.

Závěr

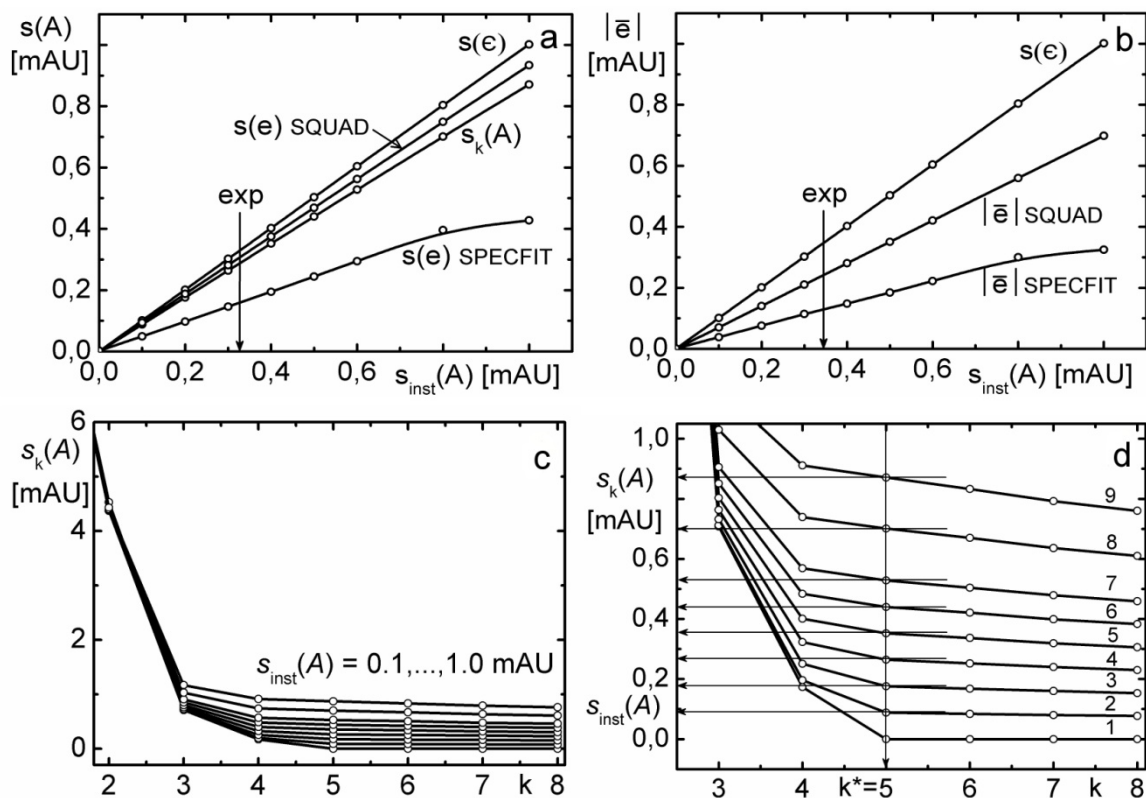
Bylo provedeno testování dvou rozličných algoritmů pro analýzu spektrálních dat za účelem stanovení chemického modelu. Program SPECFIT/32 je novější, je z numerického hlediska podstatně vyspělejší – je zde užita Levenbergova – Marquardtova minimalizační metoda, proto rychleji a přesněji konverguje. Program SQUAD(84) užívá metodu Newtonovu – Raphsonovu. Obsahuje však řadu diagnostických nástrojů, především kompletní analýzu reziduí, což umožňuje objektivněji posuzovat průběh výpočtu zejména v případě, analyzujeme-li neznámý vzorek.

Po provedeném testování lze říci, že oba programy dosahovaly shodných výsledků co do správnosti nalezených parametrů, ale program SPECFIT/32 dosahoval vyšší těsnosti proložení a tudíž i vyšší přesnosti stanovených parametrů.

Pro spolehlivé určení chemického modelu je klíčová přesnost a správnost experimentálních dat. To je umožněno na dnes již běžných měřicích přístrojích, které umožňují měřit absorbanci s přesností na šest desetinných míst.



Obr. 2 Správnost a přesnost nalezených odhadů čtyř disociačních konstant v závislosti na vloženém šumu $s_{inst}(A)$ přístroje. Svislice značí šum použitého spektrofotometru firmy GBC.



Obr. 3 Věrohodnost určení počtu světlo-absorbujících částic Cattelovým grafem v závislosti na velikosti šumu přístroje $s_{inst}(A)$.

Literatura

- [1] MELOUN, M., FERENČÍKOVÁ, Z., JAVŮREK, M. Reliability of dissociation constants and resolution capability of SQUAD(84) and SPECFIT/32 in the regression of multiwavelength spectrophotometric pH-titration data, *Spectrochim. Acta A: Mol. Biomol. Spectrosc.* (2011), doi: 10.1016/j.saa.2011.10.041.
- [2] KANKARE, J.J. Computation of Equilibrium Constants for Multicomponent Systems from Spectrophotometric Data, *Anal. Chem.* 42 (1970) 1322.
- [3] MELOUN, M., MILITKÝ, J. *Statistická analýza experimentálních dat*, Praha: Academia, 2004, 953 s., ISBN 80-200-1254-0.
- [4] MELOUN, M., JAVŮREK, M., HAVEL, J. Multiparametric curve fitting. 10. A structural classification of programs for analyzing multicomponent spectra and their use in equilibrium-model determination, *Talanta* 33 (1986), 513-524
- [5] SPECFIT/32, Spectrum Software Associates, 197M Boston Post Road West, Marlborough, MA, 01752 USA, 2004.