



vodní hospodářství®

www.vodnihospodarstvi.cz

ročník 57

1
2007

10 let

výroby
kompozitních
konstrukcí v ČR

 KOMPOZITY



Moravská Třebová 3. - 4. 4. 2007
"Nové metody a postupy při provozování ČOV"
tradiční odborný seminář XII. ročník

pořádá VHOS, a.s. ve spolupráci s AČE ČR

program, přihlášky: www.vhos.cz

informace: J. Kotoučková, telefon: 461 357 103, e-mail: tr.sek@vhos.cz

PŘÍLOHA
• ČL •

Statistické zpracování vodohospodářských dat

3. Intervalové odhady a míry přesnosti v kalibraci

Milan Meloun

Klíčová slova

Kalibrace - přímý odhad - Naszodiho korekce - Kručkovův inverzní odhad - Schwarzův nelineární odhad - kritická úroveň - limita detekce - limita stanovení - uzlový bod - spline - interval spolehlivosti

Souhrn

Lineární a nelineární kalibrace patří k základním úlohám v každém laboratoři. Kalibrace se skládá ze dvou fází, z konstrukce kalibračního modelu $y = f(b, x)$ a z použití kalibračního modelu, když naměřeného signálu y^* (např. absorbance) je určován bodový odhad neznámé koncentrace či obsahu látky x^* . Vedle bodového odhadu neznámé koncentrace se vyčísluje také intervalový odhad, který je u lineárního kalibračního modelu symetrický zatímco u nelineárního asymetrický. V případě lineárních modelů se obě fáze provádějí s využitím kalibrační přímky. Vedle přímého odhadu x^* se užívá i Naszodiho korekce na vychýlení, dále i Kručkovův inverzní odhad a Schwartzův nelineární odhad. V případě nelineárních kalibračních modelů se používá splinová regrese s pevnými uzly, které uživatel volí. U kalibrace se přesnost kalibrace vyjadřuje limitními hodnotami, které souvisejí s úrovní koncentrace, pro kterou je signál ještě statisticky významný a odlišitelný od šumu: *kritická úroveň* (slepý pokus) představuje horní mez 100(1 - α)-%ního intervalu spolehlivosti predikce signálu z kalibračního modelu pro koncentraci rovnou nule. *Limita detekce* odpovídá hodnotě koncentrace, pro kterou je dolní mez 100(1 - α)-%ního intervalu spolehlivosti predikce signálu kalibračního modelu rovna kritické úrovni. *Limita stanovení* je nejmenší hodnota signálu, pro kterou je relativní směrodatná odchylnka predikce kalibračního modelu dostatečně malá a obvykle rovna 10%.

1 Úvod

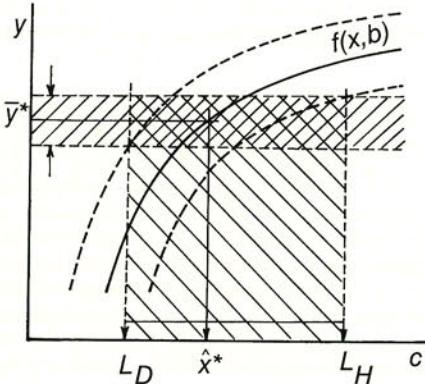
Kalibrace patří k základním úlohám, které se řeší s využitím regresních metod. Používá se při konstrukci snímačů fyzikálních veličin, adjustaci přístrojů a při vývoji nových instrumentálních metod. Skládá se ze dvou fází: (a) sestavení kalibračního modelu, (b) použití kalibračního modelu. Sestavení kalibračního modelu je totožné s úlohou hledání regresního modelu, tj. pro závislost signálu y na cílovou veličinu x . Ve druhé fázi se řeší úloha inverzní, tj. pro naměřenou odezvu y^* se hledá odpovídající hodnota x^* a její statistické charakteristiky. Z numerického hlediska jde o úlohu hledání kořene nelineární funkce. Ze statistického hlediska jde o problém nelineárního vychýleného odhadu.

2 Druhy kalibrace

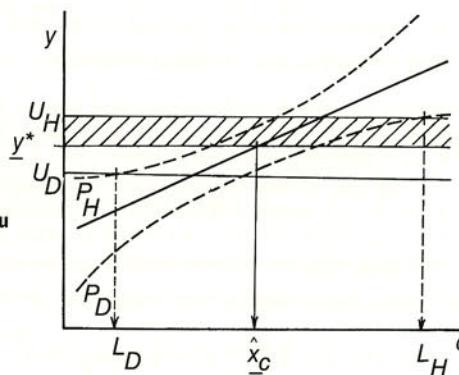
Rozdělení kalibračních úloh podle různých hledisek, určujících způsob statistického zpracování, je uvedeno v práci Rosenblatta a Spiegelmana [1]:

(1) *Absolutní kalibrace* je nejběžnější v technické praxi. Při sestavení kalibračního modelu se hledá vztah mezi měřitelnou veličinou η , nazývanou signál (potenciál, napětí článku, proud, elektrický odpor, pH, absorbance, atd.), a cílovou veličinou ξ , která určuje stav nebo vlastnosti systému (obsah, koncentrace, objem, teplota, čas, atd.). Ta může být obtížnější měřitelná a bývá proto určována jako výsledek druhé fáze kalibrace. Příkladem může být kalibrace absorbance roztoku (η) na jeho koncentraci (ξ). Při kalibračním experimentu se u n vzorků se znázými nebo přesně měřitelnými hodnotami proměnné ξ změří odpovídající veličiny η . Vzhledem k tomu, že jsou ve většině případů obě veličiny naměřené, bude pro n bodů $\{\eta_i, \xi_i\}$, $i = 1, \dots, n$, platit $\eta_i = \eta_0 + \epsilon_i$, $\xi_i = \xi_0 + \delta_i$, kde ϵ_i a δ_i jsou experimentální chyby. Pokud je proměnná ξ měřena přesně, nebo je užito definovaných standardů, je $\delta_i = 0$, $i = 1, \dots, n$. Veličina η_0 je approximována kalibračním modelem $f(x, b)$ a úloha zpracování dat vede na odhad parametrů b . Ve fázi použití kalibračního modelu se pro naměřené hodnoty při M opakování měření signálu

Obr. 1. Zadání absolutní kalibrace a postup při určení koncentrace pro průměrnou hodnotu signálu z kalibrační křivky. L_D , L_H jsou dolní a hornímez konfidenčního intervalu. Šrafován je vyznačen konfidenční pás signálu a konfidenční oblast koncentrace.



Obr. 2. Určení intervalu spolehlivosti $[L_D, L_H]$ koncentrace c z kalibrační přímky. Šrafování značí poloviční šíři intervalu spolehlivosti signálu



$\{y_j^*\}, j = 1, \dots, M$ určuje průměrná hodnota vlastnosti \bar{x}^* a její interval spolehlivosti. Případ kalibrace absorbance na koncentraci je znázorněn na obr. 1. u kalibrační křivky a obr. 2. u kalibrační přímky, kde symboly L_D a L_H značí dolní a hornímez asymetrického nebo symetrického konfidenčního intervalu koncentrace. V tomto článku o kalibraci budeme věnovat pozornost především absolutní kalibraci.

(2) *Komparativní kalibrace* je postup, při kterém se jeden přístroj kalibruje vůči druhému a je libovolné, který se užije jako standard. Příkladem může být měření koncentrace z absorbance dle Beerova zákona (1. metoda) a titračně (2. metoda). Porovnávají se hodnoty absorbance vůči objemu titračního činidla. Chybou δ , nejsou zanedbatelné a je třeba při sestavení kalibračního regresního modelu užít ortogonální regrese, tj. regrese pro případ obou proměnných zatížených náhodnými chybami, šumem.

S ohledem na vlastní práci s kalibračním modelem je možné uvažovat: (a) jedno použití, kdy se z kalibračního modelu, vzniklého z n bodů $\{x_i, y_i\}$, $i = 1, \dots, n$, počítá odhad cílové veličiny \hat{x}^* se svým intervalem spolehlivosti pro jednu hodnotu y^* ; (b) vícenásobné použití, kdy se pro různé hodnoty signálů určují odhad \hat{x}^* na základě jednoho kalibračního modelu; (c) jedno nebo více použití v kombinaci s dalšími měřeními, kdy se výsledek druhé fáze kalibrace užívá spolu s dalšími údaji k určení veličiny, která je funkcí více proměnných. Zde se velikost vychýlení odhadu \hat{x}^* projeví v systematické chybě výsledku.

3 Rozptyl predikce cílové veličiny x^*

Složitost řešení úlohy kalibrace souvisí zejména s užitým kalibračním modelem. Pro lineární regresní modely lze vyjádřit konfidenční pásy kolem modelu pomocí známých kvadratických rovnic. Složky vektoru \mathbf{x} jsou funkemi cílové veličiny, obvykle koncentrace. Obyčejně se uvažují polynomické modely, u kterých jednotlivé složky odpovídají mocninám měřené vlastnosti. Při hledání hodnoty cílové veličiny \hat{x}^* se řeší úloha hledání kořene polynomu. Pro nelineární regresní modely se hledá řešení ve tvaru $\hat{x}^* = f^{-1}(y^*)$. Na základě Taylorova rozvoje této funkce lze nalézt přibližnou formuli pro rozptyl $D(\hat{x}^*)$ ve tvaru uvedeném v citaci

[2], tj. $D(\hat{x}^*) \approx \left[\frac{\partial f(x, b)}{\partial x} \right]^{-2} \left[\frac{D(y^*)}{M} + D(f(x, b)) \right]$, kde $D(y^*)$ je rozptyl y^* -ových hodnot, který je obyčejně roven σ^2 , a $D(f(x, b)) = D(\bar{y})$ je rozptyl predikce, který se určuje také z Taylorova rozvoje funkce $f(x, b)$. Pro lineární regresní modely je rozptyl predikce roven

$$D(\bar{y}) = \sigma^2 \left[\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] = \sigma^2 \left[\frac{1}{n} + \frac{(y^* - \bar{y})^2}{b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2} \right]$$

kde b_1 představuje odhad směrnice regresní přímky. Po dosazení dostaneme pro rozptyl odhadované koncentrace

$$D(\hat{x}^*) \approx \frac{\sigma^2}{b_1^2} \left[\frac{1}{M} + \frac{1}{n} + \frac{(y^* - \bar{y})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right]. \text{ Problémem je, že rozdelení veličiny } \hat{x}^* \text{ je obecně nesymetrické. Jedině pro případ kalibrační přímky a malý reziduální rozptyl lze rozdelení veličiny } \hat{x}^* \text{ považovat za přibližně symetrické a normální [3]. Za předpokladu, že jak } y\text{-ové, tak i } y^*\text{-ové hodnoty jsou náhodné proměnné s normálním rozdelením platí, že rozdíl } \Delta = y^* - f(x^*, b) \text{ bude mít také normální rozdelení. Standardizovaná náhodná veličina } \Delta / \sqrt{D(\Delta)} \text{ má pak Studentovo rozdelení se stupni volnosti, které byly užity při určení } D(\Delta). \text{ Pro } 100(1-\alpha)\% \text{ konfidenční interval hodnoty odhadované koncentrace } \hat{x}^* \text{ platí, že}$$

$$D(\Delta) = D(\bar{y}^*) + \sum_{j=1}^m \left[\frac{\delta f(\hat{x}_j^*, b)}{\delta b_j} \right]^2 D(b_j) + 2 \sum_{i=1}^{m-1} \sum_{j>i}^m \frac{\delta f(\hat{x}_i^*, b)}{\delta b_i} \frac{\delta f(\hat{x}_j^*, b)}{\delta b_j} \text{ cov}(b_i, b_j)$$

kde m je počet regresních parametrů. Speciálně pro model kalibrační přímky $y = b_1(x - \bar{x}) + \bar{y}$ vyjde dosazením

$$D(\Delta) = \sigma^2 \left[\frac{1}{M} + \frac{1}{n} + \frac{(\hat{x}^* - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right].$$

Při znalosti $D(\Delta)$ lze již nalézt krajní meze konfidenčního intervalu pro \hat{x}^* . Jde o úlohu hledání kořenů kvadratické rovnice

$$[(\bar{y}^* - \bar{y} - b_1(\hat{x}^* - \bar{x}))^2] = F_{1-\alpha}(I, n-2) D(\Delta)$$

vzhledem k proměnné \hat{x}^* .

4 Kalibrační přímka

Model kalibrační přímky patří v laboratořích k nejpoužívanějším. Předpokládá se obvykle, že tento model vyhovuje v celém sledovaném rozsahu proměnných x a y . Příkladem může být Lambertův-Beerův zákon, $A = \epsilon c d$, vyjadřující lineární vztah mezi absorbancí A a koncentrací c , kde molární absorpcní koeficient ϵ a délka kyvety d jsou konstanty. V některých případech však model kalibrační přímky platí pouze v omezeném intervalu a nad hraničním bodem $\{x_A, y_A\}$ dochází k výrazným odchylkám od linearity. Příkladem může být Kubelkův-Munkův vztah mezi funkcí remise $(1 - R^2)/2R$ a koncentrací c , který platí jen pro nízké koncentrace. Konečně i výše citovaný Lambertův-Beerův zákon platí pro některé roztoky jen do určité prahové koncentrace, od které pak dochází k výrazným nelineárním odchylkám. Z hlediska statistického zpracování lze užit kalibračního modelu $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i=1, \dots, n$, pro určení velikosti signálu neznámé koncentrace $y_j = b_0 + b_1 x_j$, $j=1, \dots, M$. Úlohou kalibrace je potom nalezení odhadu \hat{x}^* parametru κ jako primárního a odhadu parametrů β_1 , β_2 jako doplňkových. Vychází se opět z představy normality chyb ε_i a ε_j . Odhad \hat{x}^* a jeho odpovídající konfidenční interval je možné určit několika způsoby:

(1) *Přímý odhad parametru κ ve tvaru $\hat{x}^* = \bar{x} + y^* - \bar{y} / b_1$, kde y^* je měřená hodnota signálu (průměr \bar{y}^* pro $M > 1$ opakování měření) a b_1 je odhad směrnice kalibrační přímky. Tento odhad je obecně vychýlený.*

(2) *Korekci na vychýlení lze provést pomocí Naszodiho modifikovaného odhadu $\hat{x}_B^* = \bar{x} + \frac{(y^* - \bar{y}) b_1}{b_1^2 + \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$.*

(3) *Kručkov [6] navrhl inverzní odhad*

$$\hat{x}_I^* = \bar{x} + (y^* - \bar{y}) \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2}, \text{ který vychází z inverzního regresního modelu } E(x/y) = \alpha_1(y - \bar{y}) + \alpha_0. \text{ Na základě rozboru odhadu } \hat{x}_I^* \text{ bylo zjištěno, že jde také o vychýlený odhad, který není lepší než přímý odhad } \hat{x}^*. \text{ Navíc se při odhadování parametrů } \alpha_0 \text{ a } \alpha_1 \text{ chyběně předpokládá, že } y\text{-ové hodnoty jsou měřeny se zanedbatelnými chybami vůči } x\text{-ovým hodnotám.}$$

(4) V práci Schwartzte [4] je navržen nelineární odhad

$$\hat{x}_N^* = \left(\sum_{i=1}^n x_i \exp \left[\frac{-(y^* - b_0 - b_1 x_i)^2}{2\sigma^2} \right] \right) / \left(\sum_{i=1}^n \exp \left[\frac{-(y^* - b_0 - b_1 x_i)^2}{2\sigma^2} \right] \right),$$

který je však založen na předpokladu normality reziduů.

Příklad 1. Bodové odhady koncentrace z kalibrace AAS

Metodou atomové absorpční spektrofotometrie AAS byla změřena absorbance roztoků o různé koncentraci lithia, **tabulka 1**. Je třeba určit kalibrační přímku a z ní pak odhadnout koncentraci lithia pro naměřené absorbance vzorků $A_1 = 0.0002$, $A_2 = 0.5$, $A_3 = 1.0$ o neznámých koncentracích.

Řešení: Klasickou metodou nejmenších čtverců byla určena rovnice kalibračního modelu $A = 0.02525 (\pm 1.138 \cdot 10^{-4})c + 0.0002 (\pm 2.753 \cdot 10^{-3})$ s korelačním koeficientem, $R = 0.9999$. V **tabulce 2**, jsou uvedeny bodové odhady \hat{x}^* , \hat{x}_B^* , \hat{x}_I^* , \hat{x}_N^* pro absorbanční signál od tří vzorků, a to a $A = 0.0002$, 0.5 a 1.0 . Vzhledem k dostatečné přesnosti dat jsou všechny odhady, až na nelineární odhad \hat{x}_N^* shodné.

Závěr: Úloha ukázala, že pro přesná data s malým rozptylem kolem regresní přímky zcela postačuje klasický přímý odhad \hat{x}^* .

Tabulka 1. Absorbance roztoků pro různé koncentraci lithia c [g Li/25 cm⁻³] poskytuje kalibrační přímku.

A	0.063	0.120	0.189	0.251	0.316	0.393	0.442	0.502
c	2.5	5.0	7.5	10.0	12.5	15.0	17.5	20.0
A	0.568	0.639	0.694	0.749	0.821	0.884	0.947	1.010
c	22.5	25.0	27.5	30.0	32.5	35.0	37.5	40.0

Tabulka 2. Odhad koncentrace, vyčíslené rozličnými metodami

Absorbance	Absorbance			
	\hat{x}^* (Přímý odhad)	\hat{x}_B^* (Naszodiho korekce)	\hat{x}_I^* (Kručkovův inverzní odhad)	\hat{x}_N^* (Schwarzův nelineární odhad)
0.0002	0	4.32 · 10 ⁻⁴	6.0499 · 10 ⁻³	2.5
0.5	19.795	19.795	19.795	20
1.0	39.597	39.597	39.592	40

5 Nelineární model kalibrační křivky

Kromě nelineárních modelů, vycházejících z fyzikálně chemických vztahů mezi signálem a odevzou se v praxi používají také nelineární empirické modely. Značného zjednodušení lze dosáhnout využitím speciálních modelů, které vedou k odhadům, získaným lineární metodu nejmenších čtverců. Mezi takové dostatečně flexibilní modely patří například, polynomické spline funkce: u modelu kalibrační křivky se k approximaci často volí lokálně definované funkce, které budou v místech vzájemného styku, tj. v uzlech, spojité ve funkčních hodnotách a hodnotách zadaných derivací, **obr. 1**. Vhodné interpolační funkce tohoto typu jsou složeny z polynomických úseků a platí pro ně, že jsou ze třídy $C^m[a, b]$. Obecně jsou funkce třídy $C^m[a, b]$ na intervalu $[a, b]$ spojité v prvních m derivacích a funkčních hodnotách. Hladké jsou všechny funkce od třídy C^0 . Pro funkce třídy C^m platí, že m -tá derivace je lineární lomená závislost, $(m+1)$ -derivace je po částech konstantní a $(m+2)$ -derivace je po částech nulová, tj. není definovaná v uzlových bodech ξ_i .

6 Intervalové odhady cílové veličiny x

Při konstrukci intervalů spolehlivosti cílové veličiny x pro odhady \hat{x}^* nebo \hat{x}_B^* u silněji rozptýlených dat je nejjednodušší užít k určení $D(\hat{x}^*)$ předpokladu asymptotické normality. Meze 95%ho intervalu spolehlivosti se pak vypočtou dle *approximativního vztahu*, a to pro dolní mez $L_D = \hat{x}^* - 1.96 \sqrt{D(\hat{x}^*)}$ a pro horní mez $L_H = \hat{x}^* + 1.96 \sqrt{D(\hat{x}^*)}$.

Grafický způsob určení intervalu spolehlivosti pro cílovou veličinu x je na **obr. 2**. Z obrázku je patrné, že v případě opakování měření signálu y a určení střední hodnoty y^* je třeba stanovit konfidenční přímky U_D a U_H , a řešit úlohu hledání průsečíku U_H s dolní konfidenční parabolou P_D kalibrační přímky, kdy výsledkem je bod L_H , respективě průsečíku přímky U_D s horní konfidenční parabolou P_H kalibrační přímky, kdy výsledkem je bod L_D . Při znalosti rozptylu měření σ^2 lze snadno definovat $100(1-\alpha)\%$ interval spolehlivosti pro signál y^* ve tvaru $U_{D,H} = y^* \pm u_{1-\alpha/2} \sigma$, kde $u_{1-\alpha/2}$ je kvantil normovaného normálního rozdělení. Pokud σ^2 není známo, lze využít nerovnosti $\sigma^2 < [(n-2)\sigma^2]/[\chi^2_{\alpha/2}(n-2)M]$, kde $\chi^2_{\alpha/2}$ je

dolní kvantil χ^2 -rozdělení. Interval spolehlivosti signálu $U_{D,H}$ se potom vypočte ze vztahu $U_{D,H} = \bar{y} \pm u_{1-\alpha/2} \frac{\sigma}{\sqrt{M}} \sqrt{\frac{n-2}{\chi^2_{\alpha/2}(n-2)}}$.

Místo kvantitu $u_{1-\alpha/2}$ se v této rovnici pro $M=1$ užívá kvantil Studentova rozdělení $t_{1-\alpha/2}(n-2)$ a rozptyl σ^2 je nahrazen jeho odhadem σ^2 . Pro celou regresní přímku budou hraniční $100(1-\alpha)\%$ ní paraboloidy

$$P_{D,H} = b_1 x + b_0 \pm \sigma \sqrt{2F_{1-\alpha}(2, n-2) \left[\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right]}^{1/2}$$

Hraniční hodnota L_H je potom řešením rovnice $U_H = P_D$, vzhledem k proměnné x . Hraniční hodnota L_D je řešením rovnice opět vzhledem k proměnné x , tj. $U_D = P_H$. Obě rovnice jsou vzhledem k proměnné x kvadratické. Z obr. 2. je také patrné, že v některých případech nemusí průsečík přímky s parabolou existovat nebo v jiném případě protne konfidenční přímka signálu parabolu kalibrační přímky ve dvou bodech. To ukazuje na příliš velký rozptyl v datech, kdy směrnice kalibrační přímky je málo významná. Taková kalibrační přímka pak není vhodná.

Kvalita intervalu spolehlivosti cílové veličiny je příznivě ovlivněna: (1) Opakováním měření signálu y^* , čili růstem M . (2) Zvýšení konfidenčních parabol lze dosáhnout eliminací vlivných bodů, což bude ukázáno v příštím sdělení této publikaci řady. (3) Zmenšením reziduálního rozptylu σ^2 , a tedy buď zpřesněním měření, nebo užitím správného kalibračního modelu.

Příklad 2. Intervaly spolehlivosti koncentrace z kalibrace AAS

Pro kalibrační data Lambertova-Beerova zákona z Příkladu 1 je třeba vypočítat 95%ní interval spolehlivosti koncentrace pro hodnoty absorbance $A_1^* = 0.0002$, $A_2^* = 0.5$, $A_3^* = 1.0$, $A_4^* = 0.51$ (vzniklé jako průměr hodnot 0.5 a 0.52, zde $M=2$) a $A^* = 0.977$ (vzniklé jako průměr hodnot 0.95, 0.98, 1.00, kde $M=3$) s využitím všech výše uvedených vztahů.

Řešení: Vyčíslené hranice $L_{D,H}$ pro 95%ní intervaly spolehlivosti odhadované koncentrace jsou uvedeny v tabulce 3. Intervaly spolehlivosti jednoduché approximace neberou v úvahu korekci na větší přesnost stanovení y^* . Při grafickém určenímezí intervalu spolehlivosti byl užit zjednodušený výraz pro konfidenční přímky signálu $U_{D,H} = \bar{y} \pm u_{1-\alpha/2} \frac{\sigma}{\sqrt{M}}$ při volbě $\alpha=0.05$. Z tabulky 3. je patrné, že vzhledem k dostatečné přesnosti dat vede approximativní vztah i grafický způsob ke stejným výsledkům.

Závěr: Pro data s malým rozptylem kolem regresní přímky je výhodné použít spíše jednodušší approximace.

Tabulka 3. Interval spolehlivosti stanovené koncentrace, vyčíslený dvěma rozličnými vztahy: L_D značí dolní a L_H horní mez intervalu spolehlivosti.

M	A^*	Užitý vztah pro interval spolehlivosti		
		Aproximativní vztah	Grafický způsob	
1	0.0002	L_D	L_D	
		-0.46	-0.398	
1	0.50	L_H	L_H	
		0.46	0.417	
2	(0.50; 0.52)	19.37	19.39	
		20.22	20.20	
1	1.00	19.89	19.90	
		20.50	20.48	
3	(0.95; 0.98; 1.00)	39.16	39.19	
		40.05	40.01	
		38.37	38.43	
		38.97	38.91	

7 Přesnost kalibrace

K vyjádření přesnosti kalibrace se definují limitní hodnoty, které souvisejí s takovou úrovňí koncentrace, pro kterou je signál ještě statisticky významně odlišný od šumu. V souvislosti s vyjádřením přesnosti a citlivosti kalibračních metod se definují tři specifické úrovny signálu:

1. Kritická úroveň y_c představuje horní mez $100(1-\alpha)\%$ ního intervalu spolehlivosti predikce signálu z kalibračního modelu pro koncentraci rovnou nule, tzv. *slepý pokus*. Pro kritickou úroveň y_c platí vztah

$$y_c = \bar{y} - b_1 \bar{x} + \sigma \sqrt{t_{1-\alpha/2}(n-2)} \sqrt{1 + \frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}. \text{ Potom platí, že}$$

nad hodnotou y_c lze signál odlišit od šumu. Koncentrace x_c , odpovídající hodnotě kritické úrovni, se určí z kalibračního modelu pomocí vztahu

$$x_c = \frac{y_c - \bar{y}}{b_1} + \bar{x} \quad (\text{viz obr. 3}).$$

2. Limita detekce y_D odpovídá hodnotě koncentrace, pro kterou je dolní mez $100(1-\alpha)\%$ ního intervalu spolehlivosti predikce signálu z kalibračního modelu rovna y_c , obr. 3. Pro lineární kalibrační model lze

$$\text{psát } y_D = y_c + \sigma \sqrt{t_{1-\alpha/2}(n-2)} \sqrt{1 + \frac{1}{n} + \frac{(x_D - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}.$$

Odpovídající koncentrace x_D se vypočte podle vztahu

$$x_D = \frac{y_D - \bar{y}}{b_1} + \bar{x}. \text{ Limita detekce udává skutečnou úroveň signálu,}$$

která umožnuje ještě detekci koncentrace. Číselná velikost x_D udává minimální koncentraci, kterou lze ještě s pravděpodobností $(1-\alpha)$ odlišit od nulové hodnoty.

3. Limita stanovení y_s je nejmenší hodnota signálu, pro kterou je relativní směrodatná odchylka predikce z kalibračního modelu dostatečně malá a rovna číslu C. Pro číslo C se volí obyčejně hodnota $C=0.1$. Označme predikci v místě x_s výrazem $y(x_s) = \bar{y} + b_1(x_s - \bar{x})$. Podmínka pro určení y_s je pak rovna $\sqrt{D(y(x_s))} / \bar{y}(x_s) = C$.

$$\text{Dosazením bude } y_s = \frac{\sigma}{C} \sqrt{1 + \frac{1}{n} + \frac{(x_s - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}. \text{ K praktickým}$$

výpočtům v laboratoři se však často užívá approximace

$$y_s \approx \frac{\sigma}{C} \sqrt{1 + \frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}. \text{ Odpovídající koncentrace } x_s \text{ je pak}$$

$$x_s = \frac{y_s - \bar{y}}{b_1} + \bar{x}. \text{ Z uvedených charakteristik lze snadno konstruovat}$$

limitu detekce y_D a limitu stanovení y_s i pro nelineární kalibrační modely a pro případy dat, kdy rozptyly měření nejsou konstantní. Obecně platí porovnání limit, že $y_c \leq y_D \leq y_s$.

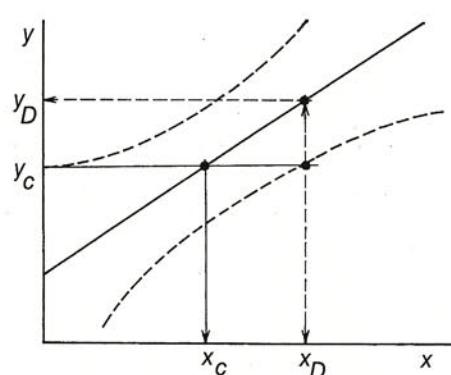
Příklad 3. Limity přesnosti u kalibrace AAS

Pro kalibrační data Lambertova-Beerova zákona z Příkladu 1 je třeba vypočítat limity přesnosti kalibrace, tj. kritickou úroveň y_c , limitu detekce y_D a limitu stanovení y_s pro relativní směrodatnou odchylku $C=0.1$, a k těmto limitám i odpovídající koncentrace.

Řešení: Pouhým dosazením do příslušných vztahů lze vypočítat všechny tři úrovně absorbance a jim odpovídající koncentrace: $y_c = 0.0061$, $x_c = 0.234$, $y_D = 0.0119$, $x_D = 0.464$, $y_s = 0.0593$, $x_s = 2.340$. Jelikož naměřená hodnota prvního neznámého vzorku $A^* = 0.0002$ leží pod hodnotou limity detekce $y_D = 0.0119$, nelze tuto koncentraci stanovit a na její místo se uvede slovo "nedetectovatelná koncentrace".

Závěr: Pro lineární kalibrační modely lze jednoduše stanovit všechny limity. K určení citlivosti metody kalibrace postačuje určení limity detekce. Limita stanovení ukazuje na spodnímez běžné použitelnosti kalibrační

Obr. 3. Definice kritické úrovni y_c , limity detekce y_D a jim odpovídající koncentrace x_c a x_D



Tabulka 4. Data kalibrační přímky obsahují hodnoty signálu ze tří rozličně přesných přístrojů: A nejméně přesný, B přesnější, C nejpřesnější přístroj s nejmenší hladinou šumu.

c [mmol dm ⁻³]	Hodnoty signálu		
	A	B	C
1.0	4.823	5.108	5.101
2.0	5.197	5.207	5.199
3.0	5.937	5.311	5.300
4.0	5.424	5.399	5.400
5.0	5.255	5.497	5.500
6.0	5.702	5.606	5.599
7.0	5.790	5.700	5.699
8.0	5.962	5.769	5.800
9.0	5.734	5.889	5.899
10.0	5.786	6.008	5.999
11.0	6.117	6.099	6.101
12.0	6.555	6.203	6.200
13.0	6.570	6.289	6.299
14.0	6.815	6.399	6.400
15.0	6.187	6.494	6.500
16.0	6.552	6.606	6.598
17.0	6.947	6.697	6.702
18.0	7.090	6.807	6.801
19.0	7.159	6.896	6.900
20.0	7.291	6.994	7.001

závislosti. Je třeba si uvědomit, že hodnoty signálu, ležící pod hodnotou limity detekce y_D se nezapisují ale uvádí se jako "nedetektovatelné", t.zn. ležící v oblasti nerozlišitelného šumu.

Příklad 4. Vliv přesnosti přístroje na limity přesnosti

Užitím tří rozličně přesných přístrojů A, B a C byly změřeny signály y pro 20 úrovní koncentrace x, a data jsou v tabulce 4. Pro navržený regresní model kalibrační přímky $E(y/x) = \beta_1 x + \beta_0$ je třeba vypočítat limitu detekce a limitu stanovení regresní analýzou. Pro úroveň signálu $y^* = 6$ z jednoho měření ($M = 1$) je třeba odhadnout koncentraci a její 95% interval spolehlivosti. K této úloze byl užit vzorek o koncentraci 0.10 mol dm⁻³, vykazující signál y = 6.

Řešení: Vzhledem k tomu, že jde o ukázku vlivu přesnosti přístroje především na odhad parametrů b_0 a b_1 , nebyla zde provedena celá regresní diagnostika. V tabulce 5. jsou uvedeny odhad parametrů kalibrační přímky b_0 a b_1 , dále korelační koeficient R a reziduální směrodatná odchylka $s(\hat{e})$, určené metodou nejmenších čtverců.

V tabulce 6. jsou uvedeny hodnoty kritické úrovni (slepého pokusu) $\{y_c, x_c\}$, limity detekce $\{y_D, x_D\}$ a limity stanovení $\{y_s, x_s\}$ po data ze tří užitých přístrojů A, B, C. Hodnoty limit jsou ovlivněny poměrně velkou hodnotou úseku kalibrační přímky, která ukazuje na vysokou hodnotu signálu slepého pokusu $\{y_c, x_c\}$. Lze proto doporučit, aby experimentální měření bylo v budoucnu provedeno s automatickou korekcí slepého pokusu.

V tabulce 7. jsou různé bodové odhady neznámé koncentrace \bar{x} pro úroveň signálu $y^* = 6$ při jediném měření $M = 1$.

V tabulce 8. je uveden 95% interval spolehlivosti pro koncentraci \bar{x} při úrovni signálu $y^* = 6$.

Závěr: Rozptýlení dat kolem kalibrační přímky souvisí s přesností přístroje a ovlivňuje výrazně především limitní údaje a intervaly spolehlivosti koncentrace. Při vyhodnocování kalibračních experimentů z toho pak vyplývá velká důležitost analýzy kvality modelu a kvality dat metodami regresní diagnostiky.

Tabulka 5. Odhad parametrů b_0 a b_1 kalibračního modelu z dat ze tří rozličně přesných přístrojů

Přístroj	b_1	b_0	R	$s(\hat{e})$
A	0.1120	4.9685	0.9373	0.253
B	0.0997	5.002	0.9999	0.096
C	0.1000	4.9995	1	0.001

Tabulka 6. Limity přesnosti kalibrace tří rozličně přesných přístrojů.

	y_c	x_c	y_D	x_D	y_s	x_s
A	5.309	3.037	5.649	6.075	6.103	10.125
B	5.015	0.129	5.028	0.258	5.045	0.430
C	5.000	0.013	5.002	0.026	5.004	0.043

Tabulka 7. Odhady neznámé koncentrace pro naměřenou úroveň signálu neznámého vzorku $y^* = 6$.

	\bar{x}_A^* (Prímý odhad)	\bar{x}_B^* (Naszodiho korekce)	\bar{x}_C^* (Kručkovův inverzní odhad)	\bar{x}_N^* (Schwarzův nelineární odhad)
A	9.209	9.219	9.366	9.209
B	10.009	10.009	10.010	10.00
C	10.001	10.001	10.001	10.00

Tabulka 8. Interval spolehlivosti koncentrace $L_D \leq \bar{x} \leq L_H$ pro úroveň signálu $y^* = 6$ s využitím dvou rozličných metod: dolní hranice L_D (horní řádek) a dolní hranice L_H (dolní řádek)

	Aproximativní vztah	Grafický způsob
A	4.662	4.680
	13.760	13.900
B	9.817	9.820
	10.200	10.200
C	9.980	9.980
	10.020	10.020

8 Závěry

Bylo ukázáno, že kalibrační úloha se rozpadá do několika částí:

- U kalibračních úloh se pro kalibrační data naleze kalibrační model, lineární či nelineární. Za kritérium vhodnosti navrženého modelu se bere těsnost proložení experimentálních bodů vypočtenou kalibrační křivkou.
- Pro kalibrační model se pro daný signál y^* neznámého vzorku vypočte bodový a intervalový odhad cílové hodnoty x^* .
- Před vlastním užitím kalibračního modelu (lineárního i nelineárního) je vhodné výčíslit limitu detekce a limitu stanovení eventuelně i hodnotu slepého pokusu čili kritickou úroveň, které určují dolní hranici ještě použitelného kalibračního modelu.
- Často je vhodné aplikovat v rámci vyhodnocení kalibrace i regresní diagnostiku regresního tripletu (data, model, metoda), jež bude předmětem příštěho sdělení a především eliminovat lživé body, tj. odlehle hodnoty v kalibrační závislosti a současně ověřit splnění všech předpokladů nejmenších čtverců.

Poděkování

Článek vznikl za finanční podpory vědeckého záměru Ministerstva školství mládeže a tělovýchovy č. MSM0021627505.

9 Literatura

- Rosenblatt J.R. a Spiegelman C. H.: *Technometrics* **23**, 329 (1981).
- Ebel S. a Becht U.: *Fresenius Z. Anal. Chem.*, **158** (1987).
- Schwartz L. M.: *Anal. Chem.* **48**, 2287 (1976).
- Schwartz L. M.: *Anal. Chem.* **49**, 2062 (1977).
- Naszodi L. J.: *Technometrics* **20**, 201 (1978).
- Krutchkoff R. G.: *Technometrics* **9**, 425 (1967).
- Meloun M., Miličký J.: *Statistické zpracování experimentálních dat*, Academia Praha 2004 (4. vydání).
- Meloun M., Miličký J.: *Kompendium statistického zpracování experimentálních dat*, Academia Praha 2002 (1. vydání), 2006 (2. rozšířené vydání).
- ADSTAT, TriloByte Statistical Software s. r. o., Pardubice 1990.



Prof. RNDr. Milan Meloun, DrSc.
Katedra analytické chemie
Chemickotechnologická fakulta
Univerzita Pardubice,
nám. Čs. Legií 565, 532 10 Pardubice,
<http://meloun.upce.cz>
e-mail: milan.meloun@upce.cz,
telefon: 466 037 026

Computer-Assisted Statistical Data Analysis: 3. Interval Estimates And Precision Measures In Calibration (Meloun, M.)

Key Words

Calibration - point straight estimate - Naszodi correction - Krutckhoff inversion estimate - critical level - detection limit - determination limit - interval estimate of concentration - spline function.

Calibration consists of two steps, building a calibration model and application of the calibration model to determine an unknown

Milan Meloun, Jiří Militký:

Kompendium statistického zpracování dat (2., přepracované a rozšířené vydání)

Doporučená prodejní cena 750, Kč – včetně doprovodného CD.

Nakladatelství Academia připravuje nové vydání publikace, jež je zaměřena na statistickou analýzu experimentálních dat. V každé kapitole je uveden postup analýzy dat, vzorový příklad s komentovaným výstupem z počítače a vzorem protokolu. Úlohy jsou podle zaměření dále tříděny na úlohy z laboratoří technické kontroly kvality (chemická data), technické kontroly strojírenských závodů (fyzikální data), hygienických stanic, zdravotních ústavů a laboratoří klinické chemie (biochemická data), výrobci léčiv (farmakologická data), potravinářské, zemědělské a vodohospodářské inspekce (potravinářská a zemědělská data), kontroly životního prostředí a energetiky (environmentální data), hutní analýzy a zpracování rud (hutní a mineralogická data) a konečně managementu a ekonomie (ekonomická data). Vzorový software a data k úlohám jsou na doprovodném kompaktním disku, a to ve formátu Excel, ASCII atd. Publikace je výtanou pomůckou pro studenty vysokých škol technického, přírodovědného, ale také ekonomického směru. Je dále určena pracovníkům výzkumných ústavů, podnikových a kontrolních laboratoří a zkušeben ke kontrole kvality.

Milan Meloun, Jiří Militký: Statistická analýza experimentálních dat

Doporučená prodejní cena 690

Čtenáři se dostává do rukou ojedinělá publikace, v níž se odráží současná doba charakteristická tím, že vznikají, archivují se a zpracovávají se velké objemy dat. Ukazuje, že k získávání informací z takovýchto databází je nezbytné volit statistické metody a postupy založené

concentration x^* . For a calibration model proposed for the given signal value y^* , the quality of the independent variable x^* together with its confidence interval is estimated. A straight line is the usual calibration model in a laboratory even that in some cases the straight-line model is valid only in a limited interval and a calibration curve is preferred. For sufficiently precise data with a small spread around the regression straight line, the classical straight estimate is satisfactory even that the Naszodi correction or the Krutckhoff inversion estimate of the unknown concentration is preferably enumerated. The quality of the confidence interval around the parameter x is improved by repeating the signal measurement y^* or decreasing the residual variance and so increasing the precision of measurement, or by the use of a correct calibration model. To express the precision of a calibration, the critical level, the detection limit and the determination limit of the concentration for which the measurement signal is still statistically significantly different from the noise are usually calculated and applied. These precision limits determine the allowable lower limit of the calibration model. For non-linear calibration graph the quadratic spline is preferred.

Recenze

na zkoumání statistických zvláštností dat a adaptivní tvorbě datově orientovaných modelů. Pozornost autorů je nejprve zaměřena na teorii chyb instrumentálních měření i s analýzou nejistoty výsledků měření, a dále na vyšetřování statistických zvláštností dat průzkumovou analýzou, vícerozměrné statistické metody, lineární a nelineární regresní modely, analýzu rozptylu,

korelace a korelační modely, interpolace a approximace. Jádrem výkladu látky je vždy počítačem podporovaný rozbor příkladu.

Milan Meloun, Jiří Militký, Martin Hill: Počítačová analýza vícerozměrných dat v příkladech

Doporučená prodejní cena 450,

Počítačové orientovaná statistická analýza vícerozměrných dat je vysvětlena na 50 příkladech s mnoha grafy a diagramy. Kniha se soustředí na prakticky Nejpoužívanější metody, jako je průzkumová analýza dat, metoda hlavních komponent, faktorová analýza, diskriminační analýza, kanonická korelace, logistická regrese, analýza shluků, vícerozměrné škálování a korespondenční analýza. Příklady jsou voleny především z biochemické a klinické praxe, dále z analytické chemie, ale také ze sociologie a psychologie. Kniha poslouží jako učebnice studentům přírodních, technických i společenských věd, kteří začínají s vícerozměrnou analýzou pomocí počítače. Součástí knihy je CD s programem STATISTIKA a s databází vstupních dat všech uvedených příkladů včetně jejich počítačových výstupů řešení.

Knihy je možné koupit ve speciálních knihkupectvích nebo i s pomocí následujících kontaktů:

tel.: +420 224 943 164, +420 774 135 605

fax.: +420 224 941 982

e-mail: tomanova@academia.cz

www: www.academia.cz

Diskuse měsíce ke 40 letům vzniku podniku Povodí, díl třetí

Pokračujeme v diskusních příspěvcích spojených s loňským čtyřicetiletým výročím existence podniku Povodí. Nyní odpovídá technický ředitel Povodí Odry, pan Petr Březina. Dúraz je v rozhovoru kladen především na ty oblasti oboru, které jsou pro toto povodí specifické.

Stránský: Vaše povodí je nejmenším v ČR, do kterého je ale koncentrováno velice mnoho průmyslu a je hodně ovlivněno těžařskými aktivitami. Jak to ovlivňuje správu povodí Odry?

Březina: Je známo, že do našeho prostoru, hydrologickým potenciálem velmi skromného, je již z minula soustředěno značné množství průmyslu, vázaného často tak či onak na hlubinnou těžbu černého uhlí v Ostravsko – karvinské pánvi. Potřeby průmyslu bylo potřeba postupně pokrýt zdroji vody, v období po druhé světové válce zdroji především v podobě budovaných přehrádek. Mimo to si koncentrace průmyslu a obyvatel vyžádala i výstavbu vodárenských údolních nádrží. Vznikla tak poměrně složitá a vzájemně provázaná vodohospodářská soustava, která si kládla - a klade i nadále - značné nároky na správu povodí, kdy je neustále nutno dbát na všechny vazby a souvislosti jak na úseku kvantity vody, kvality

vody, ale i na poli zabezpečnosti před povodněmi. Nadále říkám i přesto, že určitým útlumem průmyslu a jeho restrukturalizací odběry vody proti začátku devadesátých let minulého století klesly až o 45 % a projevuje se to někdy i ve změně priorit při využívání území a pozemků.

Specifikem našeho území jsou i vlivy poddolování v nejexponovanější části povodí, v ostravské aglomeraci, kam se všechny naše větší toky stékají do jednoho prostoru. Tyto vlivy mimo některé plošné dopady zasahují především do úrovně povodňové ochrany. Poklesy terénu způsobené hlubinnou těžbou mění sklonové poměry na tocích a tím někde způsobují ve vztahu k okolí relativní zvýšení úrovně povodňových hladin, jinde zase erozní účinky v důsledku zvětšení sklonu nivelety dna toku. To si samozřejmě vyžaduje cílené a průběžně se vyvíjející zásahy do koryt těchto toků a dosí náročnou koordinaci s ostatními „uživateli“ okolo. Mezi nimi mám na mysli třeba správce řady mostních a toky křížujících objektů, jejichž výškové uspořádání je mnohdy pro zabezpečnost před povodněmi rozhodující, souběžný s toky a hrázemi, u nichž výšková úroveň se musí během času často zvyšovat, atd.

Ukončením téžby uhlí v západní, ostravské části revíru po r. 1995 nyní účinky dolování na zemský povrch ustaly, těží se ale dál v části východní, karvinské. Na první pohled by se celkově zdálo, že tyto vnější změny mimo vodní hospodářství v posledních 10 či 15 letech by mohly vést k určitému zjednodušení správy povodí, avšak není to pravda. Akcenty na jiné, zvláště ekologické aspekty, úroveň legislativního „ošetření“ každé