
7. VALIDACIÓ QUIMIOMÈTRICA. VISIÓ GLOBAL I UN ASPECTE ESPECÍFIC

F. Xavier Rius*

7.1. RESUM

En el present capítol hom proporciona una visió general del procés de validació de les hipòtesis i models més emprats en química analítica. Primerament, hom revisa de forma succinta, però intentant abraçar una àmplia perspectiva, els tests estadístics i les tècniques quimiomètriques més freqüents. En la segona part es descriu amb certa profunditat el procediment per a determinar la incertesa dels resultats obtinguts mitjançant anàlisi química.

7.2. INTRODUCCIÓ

Per validació quimiomètrica entenem l'acció de declarar vàlids, mitjançant tècniques, mètodes i procediments quimiomètrics, els resultats obtinguts en un procés químic de mesura. Per tant, la validació significa, de fet, una acció que incorpora l'ús de la quimiometria per a comprovar i documentar la validesa de qualsevol part del procediment de mesura químic.

La validesa és un terme qualitatiu que cal definir més específicament per a dotar-lo de ple sentit analític. Hom considera vàlid un mètode quan assoleix uns valors quantitius (i manté aquest nivell en el temps) suficientment elevats d'una sèrie de paràmetres de

* Departament de Química. Universitat Rovira i Virgili. Pl. Imperial Tàrraco, 1. 43005 Tarragona.

qualitat definits per l'usuari. En altres paraules més planeres, un mètode és vàlid si fa el que ha de fer (Kateman i Pijpers, 1981, Tranter, 1990). A la taula I hi consten els principals paràmetres de qualitat que acostumem a considerar. Sens dubte, aquí estem interessats en els que figuren a la columna de l'esquerra.

Podem veure, doncs, que som davant un camp molt ampli que pot incloure la determinació d'un conjunt de característiques molt variades per a validar, per exemple, des de mètodes analítics en la seva globalitat fins a un senzill programa d'ordinador que hom utilitzi per a calcular algun paràmetre estadístic d'interès.

En aquest treball ens hem fixat un doble objectiu: d'una banda, fer una revisió general de la validació de mètodes tot emprant conceptes i tècniques quimiomètriques; d'una altra, fixar-nos en un dels paràmetres de qualitat més importants de tot mètode analític, la precisió, i aprofundir en el procediment específic per a la seva avaluació.

7.3. LA VALIDACIÓ COM A ETAPA FINAL DEL DESENVOLUPAMENT DE MÈTODES

Existeixen nombroses tècniques quimiomètriques emprades en el procés de validació dels mètodes analítics. En el present capítol revisarem sobretot les emprades en la validació interna de mètodes nous. A més, mencionarem molt succintament les tècniques emprades en la validació al llarg del temps i les que hom empra per a la comparació i transferència de mètodes analítics.

TAULA I

Principals paràmetres de qualitat dels mètodes analítics

<i>Tipus quimiomètric</i>	<i>Tipus econòmic</i>
Exactitud	Inversió
Precisió	Manteniment
Sensibilitat	Despeses directes
Selectivitat	Despeses indirectes
Límits	Rapidesa
Robustesa	Facilitat de maneig

7.3.1. *Tècniques per a la validació interna de mètodes nous*

Els primers paràmetres de qualitat que es consideren en la validació inicial dels mètodes són l'exactitud i la precisió. Algunes normes, tals com l'EN-45001, parlen de *traçabilitat* com a concepte molt relacionat al d'*exactitud*. La determinació de la precisió (o incertesa) es tracta amb certa profunditat en la segona part d'aquesta ponència.

Existeixen nombroses tècniques estadístiques per a assegurar l'exactitud dels mètodes i dels procediments analítics (Cardone, 1983). Si hom disposa, per exemple, de mostres anomenades placebo, en les quals, per preparació, hom coneix exactament la concentració d'analític a determinar analíticament, o bé de materials de referència (Uriano i Cali, 1977), es pot calcular la mitjana aritmètica dels resultats obtinguts en aplicar repetidament el nostre procediment analític sobre el material de referència i comparar-la amb el valor considerat vertader. El test-*t* és l'indicat per a realitzar aquesta comparació estadística. Depenent del nombre de repeticions realitzades, hom utilitza els valors *z* de la distribució normal o els valors *t* de la distribució de Student (Miller i Miller, 1988; Massart *et al.*, 1988; Callao, Larrechi i Rius, 1992). També cal considerar si la comparació és la d'una mitjana amb un valor considerat vertader o la de dues mitjanes amb iguals o distintes variàncies (Miller i Miller, 1988; Massart *et al.*, 1988; Callao, Larrechi i Rius, 1992). En el procés estadístic de la validació de l'exactitud, hom ha de considerar també la possible presència de punts discrepants (Liteanu i Rica, 1980; Miller i Miller, 1988) i l'assumpció de distribucions estadístiques normals o gaussianes (Meier i Zünd, 1993).

Un aspecte important en la validació de l'exactitud és la consideració dels errors del primer tipus, és a dir, els derivats de considerar que un mètode és inexacte quan en realitat no ho és o és fals positiu, i dels errors de segona espècie, per exemple, els de considerar que un mètode és exacte quan en realitat no ho és o és fals negatiu. Una bona discussió d'aquests aspectes es pot trobar a Massart (1993) i a Flatman i Mullins (1985).

Un tipus de validació de l'exactitud del procediment analític indicat per a aquelles mostres en què no es disposa fàcilment de material ni de mètodes de referència i de les quals és fàcil trobar la matriu sense analític d'interès (per exemple, en l'anàlisi de drogues en fluids biològics) consisteix a afegir l'analític a analitzar, en diverses concentracions, a tantes altres mostres que hom està segur que no el contenen.

L'anàlisi posterior d'aquestes mostres (que en anglès reben el nom de *spiked samples*) dóna lloc a l'obtenció de dos conjunts de dades, les afegides i les trobades per anàlisi química, que es poden analitzar mitjançant els anomenats tests per parelles (*paired tests*) (Massart *et al.*, 1988; Massart, 1993) o mitjançant l'anàlisi de regressió. Aquest últim es revisa a l'apartat dedicat al calibratge lineal univariant.

Si hom vol comparar les mitjanes aritmètiques de diversos conjunts de dades, com per exemple quan són avaluats els resultats repetits produïts per diversos analistes, tècniques o laboratoris, ha d'emprar l'anàlisi de la variància o ANOVA (Massart *et al.*, 1988; Miller i Miller, 1988). Aquesta és una tècnica tradicional i elegant per a detectar la presència d'errors sistemàtics, o efectes que influeixen en la variació de les dades, en un o més conjunts de dades. Tot i això, l'ANOVA no és capaç de furnir informació sobre quin és el conjunt particular que difereix de forma estadísticament significativa de la resta. La tècnica és capaç de treballar amb dos, tres o més efectes; amb tot, té uns requisits estadístics tals com l'absència de punts discrepants o l'homogeneïtat de les distribucions, que cal conèixer. Les tècniques d'anàlisi multivariant com PCA, CFA o SMA (Mellinger, 1987; Thielemans, Lewi i Massart, 1988), tot i no tenir la capacitat que proporcionen els tests estadístics, poden ésser un bon complement a l'anàlisi de la variància. Colina *et al.* han publicat una sèrie d'indicacions valuoses per a l'ús de les diverses tècniques quimiomètriques i dels tests estadístics adequats al llarg del procediment analític (Colina *et al.*, 1988).

L'estadística més emprada per a elaborar models útils en l'anàlisi química és l'estadística paramètrica. Es basa en la suposició que les dades analitzades provenen de distribucions conegudes, comunament la distribució gaussiana, i a partir d'aquí, en la possibilitat de realitzar inferències futures per a nous conjunts de dades. Els valors *t* de les taules no sempre constitueixen els punts de referència més adequats, i en aquest sentit hom ha desenvolupat els tests d'aleatorietat, *randomization tests* (Vandenbosch, Therry i Massart, 1990; Edgington, 1987). Quan existeixen punts discrepants o la densitat de probabilitat de la distribució no és gaussiana, cal transformar les dades mitjançant operacions logarítmiques o d'un altre tipus (Box, Hunter i Hunter, 1978; Quenouille, 1966; Kateman i Pijpers, 1981). Si aquestes transformacions no donen bons resultats o bé hi ha massa pocs resultats per a considerar qualsevol tipus de distribució, cal aplicar tests no paramètrics (CETAMA, 1986;

Massart *et al.*, 1988; Rius, 1992). Cal, però, ésser conscients de les limitacions quant a la seva eficàcia i de la necessitat d'un nombre elevat de repeticions en aquestes tècniques.

La dificultat de decidir les tècniques de validació més adequades per a cada cas és gran a causa de la multiplicitat de les existents. Nogensmenys, els analistes poden trobar ajut en un sistema expert desenvolupat per a la validació quantitativa dels resultats analítics (Wolters, Den Broek i Kateman, 1990).

7.3.1.1. Calibratge lineal univariant

El model matemàtic $y_i = \beta_0 + \beta_1 x_i$, que correspon a una línia recta, és el més emprat per a descriure la relació entre respostes instrumentals i concentracions dels analits d'interès. Hom estima normalment els coeficients de regressió β_i mitjançant la tècnica de mínims quadrats, LS, d'on s'obtenen estimadors « b_i », que són els més precisos entre els obtinguts per les diferents tècniques anomenades no esbiaixades (Draper i Smith, 1981). Trobats aquests estimadors, hom obté per a cada punt experimental i que ha emprat en l'etapa de calibratge o de construcció del model l'equació $y_i = b_0 + b_1 x_i + \epsilon_i$.

En LS, per tal d'obtenir la predicció òptima, cal que sigui cert el model aplicat; però, a més, cal que les respostes tinguin la mateixa variància (siguin homoscedàstiques) (Davidian i Halland, 1990). Si existeixen respostes heteroscedàstiques, és a dir, amb excés de soroll, o amb presència de punts discrepants (*outliers*) i alta influència (*leverage*), aquestes respostes influeixen molt més que no pas d'altres en els estimadors i en les prediccions calculades. En termes estadístics, el teorema de Gauss-Markof (Rao, 1973) indica que els residus ϵ_i han d'ésser variables aleatòries de mitjana 0 i variància σ^2 i que $\text{cov}(\epsilon_i, \epsilon_j) = 0$. Si, a més, volem fer inferències respecte a la precisió dels coeficients de regressió, cal que tots els ϵ_i estiguin distribuïts amb una distribució normal $N(0, \sigma^2)$ als diferents nivells de concentració dels analits. Aquest últim requisit implica una independència dels valors de ϵ_i entre ells (Draper i Smith, 1981).

Hom hauria de validar mitjançant tests estadístics rigorosos la distribució gaussiana dels residus (Olucha i Rius, 1990), l'existència de diferències entre variàncies (heteroscedasticitat) (Massart *et al.*, 1988) o la manca d'acoblament del model lineal a les dades experimentals (Massart, 1988). Nogensmenys, aquest tipus de validació és molt tediosa i rarament es porta a terme. Una tècnica alternativa,

no tan rigorosa com la dels tests, és lo que fa ús dels diagrames de residus (Boqué, Rius i Massart, 1994).

Wolters i Kateman (1989) han estudiat el comportament dels errors en el calibratge lineal segons la tècnica LS i segons tècniques robustes en presència de distribucions no normals. Usualment, si no es compleix la condició de normalitat, hom ha d'intentar una transformació de les dades (Johnson i Wicharn, 1988). Si la situació és d'heteroscedasticitat, hom pot emprar la tècnica de mínims quadrats ponderats, WLS (Massart *et al.*, 1988; Levi, 1986), o aplicar mètodes no paramètrics (Garner i Robertson, 1988; Feinberg, 1988). En aquest últim cas, cal tenir present que, si bé s'obtenen valors molt més acurats de les variables a predir segons el model, hom no pot establir intervals d'incertesa ni avaluar estimadors de confiança que ens indiquin la precisió de la predicció.

La detecció de punts discrepants és un aspecte important de la regressió (Yuzhu, Smeyers-Verbeke i Massart, 1990). La seva presència no sols afecta una sèrie d'hipòtesis de la tècnica LS, sinó que, a més, pot modificar substancialment la validesa del model establert. Atès que la tècnica LS és poc robusta (una modificació petita dels punts de calibratge ja afecta els estimadors obtinguts dels coeficients de regressió i, per tant, qualsevol predicció futura), els diagrames de residus no sempre són capaços de detectar la presència de punts discrepants i aleshores cal aplicar tècniques de regressió robusta (Massart *et al.*, 1986; Chatterjee i Hadi, 1986; Rousseeuw i Leroy, 1987; Rius, Smeyers-Verbeke i Massart, 1989).

Sovint hom sobrevalora la capacitat de la tècnica d'addicions estàndard per a detectar errors sistemàtics en el calibratge (Bader, 1980). Aquesta tècnica només pot detectar els errors que són proporcionals a la concentració d'analit. Aquesta detecció es porta a terme mitjançant un test-*t* que compara el valor dels pendents de les rectes de calibratge obtingudes amb patrons aquosos i amb addicions estàndards. Depenent de la dispersió dels residus de calibratge, hom ha introduït modificacions de forma semblant a les del test-*t* que compara les mitjanes de dos conjunts de dades (Massart, Smeyers-Verbeke i Rius, 1989; Boqué, Rius i Massart, 1994).

En alguns casos, per exemple quan hom comprova la presència d'un analit en mostres naturals mitjançant la representació de les quantitats trobades analíticament enfront de les quantitats afegides per a diferents valors de la concentració, és útil aplicar un test individual del pendent per a comprovar si la seva diferència respecte a

la unitat és estadísticament significativa (Boqué, Rius i Massart, 1994; Massart *et al.*, 1988). A causa de la presència dels errors aleatoris, sempre existeix un interval de confiança al voltant del valor vertader del pendent, β_1 : $b_1 \pm t_{\omega, n-2} \cdot s.d. (b_1)$, i cal comprovar si el valor 1 es troba dins aquest interval calculat.

En altres casos, per exemple en comparar els resultats de dues metodologies diferents aplicades al mateix conjunt de mostres que contenen l'analit d'interès en diverses concentracions, ha estat apuntat que hi ha una correlació important entre el pendent i l'ordenada a l'origen. Per aquest fet, si hom aplica tests individuals del pendent i de l'ordenada pot cometre errors apreciables (Mandel i Linnig, 1957). Tot i així, ha de tenir-se present que en el procés de regressió lineal hom considera l'eix d'abscisses lliure d'error (o amb un error negligible respecte a l'eix d'ordenades). Quan es tracta de comparar dos mètodes, aquesta última hipòtesi no es compleix i fins recentment no han estat desenvolupats tests nous per a superar aquesta mancança (Hartmann, Smeyers-Verbeke i Massart, 1993; Riu i Rius, 1995; Riu i Rius, 1996).

Un aspecte important, amb repercussions pràctiques en els laboratoris de control de qualitat, és la validació de rectes de calibratge ja construïdes i emmagatzemades que es volen emprar per a la predicció de valors nous de la concentració de l'analit d'interès seguint una metodologia determinada. Aquí intervé l'interval de confiança per a una resposta futura, $y_0 = b_0 + b_1 x_0$, observada al punt x_0 , que correspon a un patró de calibratge de concentració coneguda. Cal comprovar si la resposta del patró mesurada en les condicions actuals es troba dins de l'interval de confiança calculat en les condicions de construcció de la recta de calibratge.

7.3.1.2. Límits de detecció

El límit de detecció (l.d.) és un paràmetre de qualitat molt relacionat, d'una banda, amb el procés de calibratge i, d'una altra, amb el concepte de soroll. A la bibliografia es poden trobar més de cinquants treballs científics relacionats directament amb límits de detecció. De tot aquest arsenal, sembla que hom pot extreure'n uns quants conceptes:

a) Currie és un dels primers científics que estudien en profunditat els l.d. La seva definició (Currie, 1968 i 1988) està formulada en funció de les respostes i no pas de les concentracions d'analit.

b) Kaiser (Kaiser, 1970) introdueix la importància estadística de considerar k vegades la desviació estàndard del senyal del blanc.

c) La translació de l'eix de concentracions dels límits descrits mitjançant la recta de regressió no considera els errors del procés de calibratge. Hi ha diverses tècniques, com la cromatografia o altres tècniques de flux, en les quals és difícil de quantificar el senyal del blanc i no es preveuen els errors de segona espècie.

d) La definició de la IUPAC (IUPAC, 1978) especifica que els l.d. s'expressen com a concentracions o quantitats. Nogensmenys, les bases que estableix són tan generals que deixen el camí obert a múltiples interpretacions i formulacions.

e) Hubaux i Vos (Hubaux i Vos, 1970), basant-se en els límits de confiança de la recta de regressió, defineixen un nou concepte del límit de detecció que no necessita mesures del blanc. Hom obté una estimació del soroll a concentració zero d'analit a partir del model de calibratge. Si bé es consideren els errors de l'etapa de calibratge, no queda ben establert el concepte d'error de segona espècie, β (Garner i Robertson, 1988).

f) Un dels passos definitius en el tema ha estat donat per Clayton *et al.* (Clayton, Hines i Elkins, 1987) en superar el diversos aspectes considerats prèviament. La metodologia de Clayton es basa a considerar el límit de detecció com un test d'hipòtesi en què intervenen l'hipòtesi nul·la, l'alternativa, l'error de primera espècie, α , i el de segona, β . Els valors dels l.d., segons aquest mètode, es poden calcular fàcilment emprant un programa d'ordinador desenvolupat recentment (Sarabia, 1994).

7.3.1.3. Regressió multivariant

Dins les tècniques per a la validació interna dels mètodes nous, actualment no es pot obviar la tendència creixent a la incorporació de models multivariants. Els sistemes instrumentals de detecció que permeten obtenir per a cada analit una resposta en forma de vector —tals com els detectors de sèrie de díodes en UV-Vis o tècniques acoblades tals com GC-MS, HPLC-FTIR, etc.— com l'esquematitzat a la figura 1, ens obliguen a estendre les tècniques de validació cap al calibratge multivariant (Martens i Naes, 1991; Lang i Kalivas, 1993), cap a la selectivitat i cap als límits de detecció multivariants (Boqué i Rius, 1996).

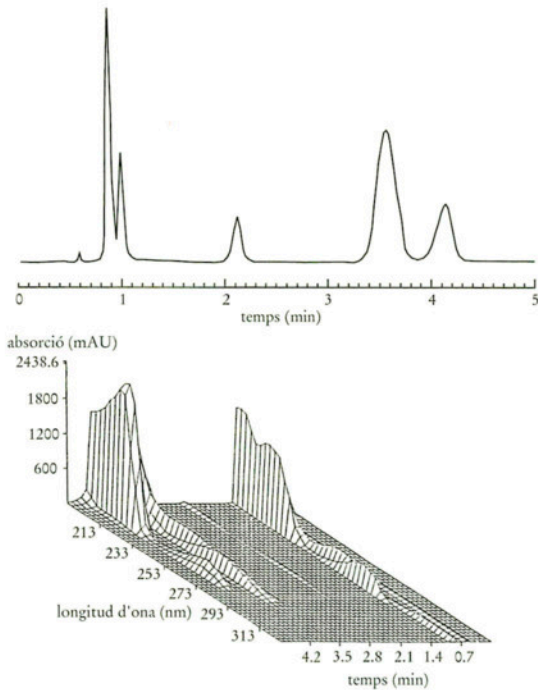


FIGURA 1. Representació del tipus de respostes multidimensional o multivariant.

Cal ésser conscient que en el calibratge univariant, sobretot el basat en la línia recta, els models emprats són relativament senzills i les respostes són considerades prou selectives. Aquestes característiques fan que la validació estigui fonamentada freqüentment en paràmetres estadístics del mateix conjunt de calibratge, tals com la desviació estàndard dels residus, el coeficient de correlació o altres estimadors (Vankeerberghen i Smeyers-Verbeke, 1992). Si el grau d'exigència és més elevat, cal efectuar un test estadístic del tipus ANOVA, en el qual és necessari realitzar repeticions (més o menys cares i tedioses) en el conjunt del calibratge. De fet, hom construeix un conjunt de predicció o prova ben controlat.

D'altra banda, en el calibratge multivariant, la validesa del model en l'etapa de calibratge, és a dir, l'acoblament del model a les dades experimentals obtingudes a partir dels patrons de calibratge, no és suficient, i cal utilitzar els errors de predicció per a validar els models proposats (Cruciani *et al.*, 1992).

Hi ha dues possibilitats per a calcular errors de predicció. Una és emprar un conjunt de predicció construït expressament per tal de verificar l'adequació del model proposat a les dades experimentals. Aquesta opció englobaria la tècnica de l'anàlisi de la variància multivariant (Draper i Smith, 1981), força cara a causa de l'experimentació necessària.

L'altra possibilitat consisteix a fer servir les dades del mateix conjunt de calibratge sense que calgui utilitzar nous patrons. Val a dir que en aquest cas el conjunt de calibratge sol ésser prou nombros, però, donada la complicació del model multivariant, totes les dades són emprades, a la vegada, per a la construcció del model i per a la seva validació. Les tècniques de la validació encreuada (*cross-validation*) i la del «tirant» (*bootstrapping*) permeten aquest estalvi (Geisser, 1975; Lepage i Billard, 1992).

L'error de predicció del model multivariant varia amb el nombre de factors o variables latents que hi intervenen. Així, doncs, cal trobar el model adequat. Malinowski fou un dels primers a proposar paràmetres estadístics indicadors del nombre òptim de factors (Malinowski, 1991).

Més recentment, han estat proposats diversos estadístics per als errors de predicció, tals com el Predictive Error Sum of Squares, PRESS (Draper i Smith, 1981; Baroni *et al.*, 1992), el PRESS robust (Osten, 1988), el Mean Square Error, MSE (Martens i Naes, 1991), o la desviació estàndard dels errors de predicció, SDEP (Cruciani *et al.*, 1992; Baroni *et al.*, 1992). El tema no és tancat, i s'hi produeixen contínuament aportacions noves com l'estudi de procediments totals de validació, *full validation procedures*, en problemes de regressió i calibratge multivariants (Lanteri, 1992), o com la Double Cross Full Validation (Forina *et al.*, 1993).

Reprement la validació quimiomètrica com un concepte global, Gemperline ha proposat un protocol interessant per a la validació dels mètodes basats en la detecció UV de sistemes multicomponents mitjançant PCR (Gemperline i Salt, 1989).

7.3.1.4. Selectivitat. Resolució

La selectivitat és un dels paràmetres de qualitat importants en tant que afecta almenys l'exactitud dels resultats. Tots els mètodes

analítics basats en tècniques de detecció univariants han d'ésser específics per evitar errors en els resultats finals. Això és la causa de la gran profusió de tècniques de separació i emmascarament que ha desenvolupat la química analítica en el passat. La tècnica d'addicions estàndard permet de detectar els efectes de la matriu quan aquests produeixen errors sistemàtics proporcionals (Bader, 1980; Boqué, Rius i Massart, 1994), però és ineficaç en presència d'errors sistemàtics constants. Cal ésser conscient que en qualsevol cas l'ús d'aquesta tècnica no valida el resultat final. La validació de l'exactitud ha de realitzar-se sempre amb les tècniques internes i externes comentades en aquest treball. En el camp multivariant, Kowalski i col·laboradors han desenvolupat la tècnica d'addicions estàndard generalitzada, GSAM, que en molts casos permet de superar els efectes de la matriu (Saxberg i Kowalski, 1979; Campins, 1988).

La selectivitat adquireix importància en les tècniques de flux, i molt especialment en cromatografia. El coneixement del nombre d'espècies diferents que són eluïts a un mateix temps de retenció és el primer pas indispensable per a una quantificació adequada d'aquests analits. En aquest sentit, la detecció multivariant ha estat l'avenç instrumental bàsic que ha permès el desenvolupament de nombroses tècniques quimiomètriques basades en l'anàlisi dels factors, *factor analysis* (Malinowski, 1991). Els treballs recents de Kvalheim han constituït un pas important envers la validació quimiomètrica de la cromatografia amb detecció multidimensional (Toft i Kvalheim, 1993). El mètode anomenat Heuristic Evolving Latent Projections, HELP, ha permès de resoldre mesclures d'isòmers, molt difícils de separar cromatogràficament, amb concentracions baixes dels components minoritaris (Liang i Kvalheim, 1993; Liang *et al.*, 1992; Kvalheim i Liang, 1992). El camp de l'*evolving factor analysis* és molt actiu actualment (Schostack i Malinowski, 1993) i les seves limitacions es defineixen per a la detecció de la puresa dels pics cromatogràfics (Vanslyke i Wentzell, 1993) i amb extensió als conjunts de dades de tres dimensions (Mitchell i Burdick, 1993).

Tauler i col·laboradors formen un dels pocs grups que apliquen amb èxit aquestes tècniques als equilibris en solució. No ens estenem en la descripció del seu treball, ja que, en ésser de casa nostra, és ben a l'abast de tothom (Gargallo *et al.*, 1996; Tauler *et al.*, 1995).

7.3.2. *Tècniques emprades en la validació al llarg del temps*

A la definició de validació, hom expressa que el nivell assolit en els valors dels paràmetres de qualitat desitjats ha de mantenir-se en el temps. Per tal de demostrar i documentar aquest fet, han estat desenvolupades tota una sèrie de tècniques estadístiques anomenades de control de qualitat, més recentment designades també amb el mot monitorització (Vernimont, 1977; Taylor, 1987).

Aquestes tècniques estan englobades dins l'anomenat sistema d'assegurament de la qualitat o de garanties de qualitat (Garfield, 1993; Dux, 1990; Valcàrcel i Ríos, 1992). Hartley ha publicat recentment una sèrie de programes d'ordinador que permeten posar fàcilment a punt diverses d'aquestes tècniques en el mateix laboratori (Hartley, 1987).

7.3.3. *Tècniques per a la comparació i la transferència de mètodes*

Una de les propietats importants dels resultats analítics és que aquests han de ser comparables. És a dir, el resultat obtingut del contingut d'un analit en una mostra amb una certa metodologia ha de ser comparable amb un altre resultat obtingut per al mateix analit en la mateixa mostra en un altre laboratori amb el mateix o amb un altre mètode analític.

Aquest fet implica que les metodologies analítiques han de validar-se no sols dins el laboratori on són utilitzades, sinó que també hi han d'intervenir agents exteriors. Entre els procediments de validació que tenen en compte agents exteriors fonamentals es troben els exercicis d'intercomparació (Youden i Steiner, 1975; Cardone, 1983; Vernimont, 1985; Horwitz, 1988; Rauret, 1992). Hi ha una extensa bibliografia estadística sobre aquest tema, que, per altra part, no interessa globalment a l'usuari final. Aquest sí que ha de conèixer, però, els principis bàsics i la interpretació dels resultats obtinguts (Danzer, Wank i Wienke, 1991). Mandel, en un treball recent, investiga el problema dels exercicis interlaboratori amb atenció als seus principals aspectes: plantejament dels experiments, anàlisi de les dades numèriques, tractament de les dades del qüestionari i possibles models (Mandel, 1991).

Un aspecte important en relació amb l'aplicació dels mètodes analítics a diversos laboratoris és la capacitat de transferència d'a-

quests mètodes. Se sap que un mètode ben establert en unes condicions determinades en un laboratori pot comportar-se de forma diferent en un altre laboratori, no sols respecte a la seva incertesa, sinó també respecte a la seva exactitud o altres paràmetres de qualitat. L'estandardització dels instruments, molt recent en el seu vessant multivariant, pot ésser una eina molt útil per a transferir models de calibratge, recalibrar al llarg del temps o recalibrar models per a diversos conjunts de mostres diferents (Wang, Veltkamp i Kowalski, 1991; Wang, Lysaght i Kowalski, 1992).

7.4. VALIDACIÓ DE LA INCERTESA DELS RESULTATS ANALÍTICS

Quant la traçabilitat dels procediments d'anàlisi ha estat validada mitjançant l'ús de materials o de mètodes de referència, d'exercicis interlaboratori o d'altres tècniques comentades a l'apartat de 7.3.1, hom pot arribar a les dues realitats següents:

a) La no acceptació de la hipòtesi nul·la del test- t significa que es refusa la validesa del procediment d'assaig. Això porta associada una revisió de totes les etapes, la qual pot incloure el recalibratge d'aparells i instruments de mesura, la revisió del procediment específic d'assaig, etc.

b) L'acceptació de la hipòtesi nul·la del test- t indica que el procediment emprat és vàlid pel que fa a la producció de resultats considerats exactes.

En aquest últim cas, pel que fa a la incertesa dels resultats, s'han generat durant el procés descrit dues magnituds que han de retenir-se: d'una banda, la variància que s'obté en analitzar repetidament la mostra de referència en el nostre laboratori, que anomenem ara s_L^2 (variància de laboratori); d'altra, la variància assignada al material de referència per la mateixa institució que el proporciona o la generada pel mètode de referència, que anomenem s_r^2 (variància de la referència). En certs casos, en lloc de variàncies es donen les incerteses associades al material o al mètode de referència, u_r . Aquestes incerteses es relacionen amb la variància mitjançant l'expressió $u_r = k s_r$, on k és el factor d'incertesa que es descriu més endavant. Fins en aquest moment, per a mantenir una certa senzillesa en els temes exposats, hom no ha tingut en compte altres factors que afecten la variància del laboratori, s_L^2 . En el següent apartat desenvoluparem de forma més completa el seu càlcul.

Quan no existeixen materials de referència adequats i, per diverses raons, no és factible aplicar mètodes de referència, l'única forma de validació del nostre procediment externa al laboratori és la participació en exercicis d'intercalibratge. En aquest últim cas, també es pot aconseguir una s_L^2 mitjançant la repetició de l'assaig de la mostra durant l'exercici en el mateix laboratori (si la quantitat de mostra ho permet i amb un valor de s_r^2 que correspondria a la variància dels n valors mitjans corresponents als n laboratoris participants amb resultats considerats acceptables. La norma EN 45001 indica en aquest últim cas, respecte a la traçabilitat, que «quan no sigui aplicable la traçabilitat en relació amb patrons nacionals o internacionals, el laboratori d'assaig haurà de posar de manifest satisfactòriament la correlació o l'exactitud dels resultats dels assaigs (per exemple, mitjançant la seva participació en una comparació d'assaigs interlaboratoris)».

Reprent els passos de la validació del mètode d'assaig que ens portarà a la determinació de la incertesa del nostre procediment, correspon examinar ara més detingudament una de les etapes del procés analític que pot influir considerablement en la incertesa global de l'assaig.

1. En tot assaig instrumental existeix l'anomenada etapa de calibratge del mètode. Hem de reconèixer aquí l'ús d'un mateix terme, calibratge, per a dos conceptes diferents, el calibratge dels instruments i el calibratge dels mètodes, però la nomenclatura in-

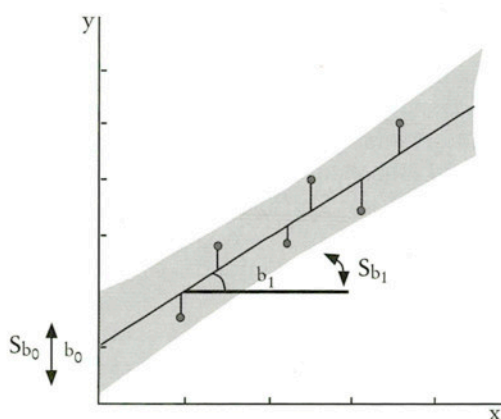


FIGURA 2. Corba de calibratge amb incerteses.

ternacionalment reconeguda en el camp de l'anàlisi química l'accepta i també anomena calibratge al procés de relacionar matemàticament la resposta d'un instrument amb la concentració d'analit responsable d'aquesta resposta. Encara que existeixen similituds conceptuals entre els dos significats del terme, crec que no ha d'existir cap dubte pel que fa a les diferències entre el procés de calibratge d'un instrument o aparell de mesura i el calibratge d'un mètode d'assaig o d'anàlisi.

El calibratge d'un mètode instrumental és un procés que ha estat molt estudiat i té les seves pròpies subtils. A banda dels comentaris realitzats a l'apartat 7.3.1.1, el lector pot aprofundir alguns conceptes en altres fonts d'informació tals com llibres especialitzats (Box, Hunter i Hunter, 1978; Massart *et al.*, 1989; Draper i Smith, 1981) o com alguns articles de revisió (Hunter, 1981; Rusling 1989; Miller 1991). Sabem que com a conseqüència del procés de calibratge s'obté generalment una relació del tipus $y = b_0 + b_1x$, on y representa la resposta instrumental mitjana i x és la concentració d'analit responsable de la dita resposta. Aquesta equació es representa en gràfics del tipus que es mostra a la figura 2, en el qual interessa posar en relleu la *presència d'una incertesa específica d'aquesta etapa del procés de mesura*. Val la pena de fer ressaltar que durant l'etapa de la construcció del model matemàtic ($y = b_0 + b_1x$) s'utilitzen patrons de calibratge. Aquests són generalment substàncies molt pures, que tenen una concentració d'analit perfectament coneguda. La incertesa generada en aquesta etapa és deguda a la incertesa associada al mesurament de la resposta de cadascun d'aquests patrons de calibratge. Per tant, durant aquest procés es té en compte la incertesa generada per l'instrument, però no la incertesa que pugui provenir de les etapes prèvies de l'anàlisi, tals com la presa de mostres, la pesada, la dessecació, etc. Perquè la incertesa de tot el procés de mesura química estigui continguda dins l'interval on es troba la recta de calibratge, els patrons de calibratge haurien d'haver sofert un tractament químic idèntic al que es dona a la mostra real. Aquesta pràctica, en realitat, rarament es fa perquè resultaria extremament llarga.

El coneixement de la presència d'aquests errors a l'etapa de calibratge i el de llurs efectes són importants perquè en la majoria de laboratoris dedicats a l'anàlisi química es porten a terme diàriament nombrosos calibratges i recalibratges de metodologies instrumentals.

Quan es pretén comprovar la validesa d'una recta de calibratge construïda anteriorment, és útil de saber que el valor mitjà de la resposta, y_0 , per a un cert patró de calibratge, x_0 , s'ha de trobar dins d'un interval de respostes donat per la següent expressió:

$$\hat{y}_i \pm t_{n-2, \alpha} \sqrt{\left(\frac{1}{N} + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right) \cdot s^2}$$

on N és el nombre de mesures repetides que es realitzen de x_0 , n és el nombre de parells utilitzats, (x_i, y_i) , corresponents als n patrons de calibratge, i s^2 representa la variància estimada dels errors de les respostes (o residus). Alguns llibres representen aquesta s^2 com a $s_{y/x}^2$:

$$s^2 = s_{y/x}^2 = \frac{\sum (y_i - \hat{y}_i)^2}{n - 2}$$

on \hat{y}_i , és el valor de la resposta calculada segons el model construït $\hat{y}_i = b_0 + b_1 x_i$, corresponent a un cert valor de la concentració d'analít x_i del patró de calibratge,

D'altra banda, l'analista està interessat a determinar la concentració d'analít desconeguda a partir de la mitjana de les respostes instrumentals. També és interessant conèixer la incertesa associada a aquest procés invers d'interpolació amb la finalitat de poder minimitzar-la. La variància associada al valor predit de la concentració de l'analít, x_{desc} , calculat a partir d'una resposta $y_{mesurada}$ d'una mostra que s'analitza, és donada per l'expressió

$$s_{calib}^2 = \left(\frac{1}{N} + \frac{1}{n} + \frac{(y_{mesurada} - \bar{y})^2}{b_1^2 \sum (x_i - \bar{x})^2} \right) \cdot \frac{s^2}{b_1^2}$$

on N és el nombre de mesures repetides de la resposta $y_{mesurada}$, la qual correspon a la mitjana de N alíquotes diferents de la mateixa mostra de concentració desconeguda x_{desc} . De fet, aquesta $y_{mesurada}$ és en realitat un valor mitjà de les diverses mesures respectives.

Aquesta variància, s_{calib}^2 , la corresponent al procés de calibratge, encara que de forma subtil, exerceix també la seva influència en el

valor final de la incertesa corresponent al resultat analític. Per tot això, és important minimitzar-la.

Queda encara un problema sense resoldre. La variància associada a totes les etapes prèvies al calibratge, encara que sabem que existeix i que en alguns casos pot arribar a ésser molt important, no s'ha tingut en compte fins ara. Aquesta variància es posa de manifest per l'evidència experimental que s'obtenen respostes instrumentals molt diferents per a un nombre de mostres iguals que han sofert idèntic tractament durant l'assaig químic. La solució al problema no és immediata; pel que fa al model matemàtic establert, $y = b_0 + b_1x$, no és vàlid per a transformar aquesta dispersió present en les respostes en una dispersió a l'eix d'abscisses o de concentracions. El model no és vàlid perquè, tal com s'ha esmentat anteriorment, no s'han tingut en compte aquests efectes (presa de mostra, pesada...) durant la construcció del model mitjançant l'ús de patrons de calibratge.

Per això és convenient, en aquelles mostres que no són homogènies, de calcular la contribució a la incertesa global de l'etapa corresponent a la presa de mostres. Si a partir de diverses alíquotes d'una mateixa mostra heterogènia realitzem el procés analític en estudi, obtenim un conjunt de senyals instrumentals (respostes) que es distribueixen al llarg de l'eix d'ordenades seguint un model gaussià. Des del punt de vista matemàtic és clar que una distribució

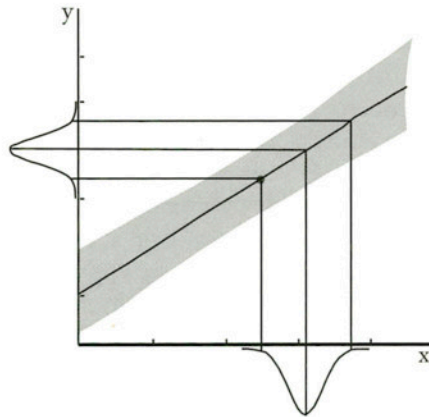


FIGURA 3. Corba de calibratge en la qual la incertesa de la resposta es transforma en incertesa a l'eix de concentracions.

normal de les respostes es converteix, per mitjà de la recta de calibratge, en una distribució normal de les concentracions. Així:

$$\begin{aligned} \text{probabilitat } \{y_2 < y < y_1\} &= 0,95 \\ \text{probabilitat } \{y_2 < b_0 + b_1 x < y_1\} &= 0,95 \\ \text{probabilitat } \{y_2 - b_0 < b_1 x < y_1 - b_0\} &= 0,95 \\ \text{probabilitat } \{(y_2 - b_0)/b_1 < x < (y_1 - b_0)/b_1\} &= 0,95 \end{aligned}$$

Per tant, a partir de la variància de les respostes instrumentals mesurades, s_{resposta}^2 , es calcula la variància de les concentracions associada a les etapes prèvies, s_{etapes}^2 , mitjançant l'expressió:

$$s_{\text{etapes}}^2 = s_{\text{resposta}}^2 / b_1^2$$

2. Anteriorment s'ha esmentat que per a validar l'exactitud del mètode s'ha de controlar prèviament el seu comportament en relació amb la precisió (assegurar les condicions de control estadístic). Per això, cal que el laboratori tingui una certa experiència de com es comporta la dispersió de resultats mitjançant el registre dels resultats d'analitzar el mateix material de referència en nombroses ocasions (amb motiu de diferents validacions del procediment al llarg del temps). Ens referim als gràfics o cartes de control de Shewhart (Taylor, 1987). En aquest cas, es tenen mesures (generalment no repetides) realitzades en diferents dies, per diferents analistes, sense utilitzar forçosament el mateix instrument de mesura, etc. Per tant, es tenen diferents valors de x_i , la concentració de l'analit en el material de referència. Aquesta informació és important perquè permet de trobar una incertesa que es traslladarà al resultat d'analitzar una mostra problema única. Aquest costum d'obtenir un sol resultat analític, que segueixen la majoria de laboratoris per a aconseguir una rendibilitat del seu treball habitual, no permet en si mateixa d'obtenir valors associats d'incertesa. És per això que només és acceptable quan coneixem suficientment bé el procés analític i tenim resultats quantitius del seu comportament.

A partir dels valors últims de x_i és immediat el càlcul de la variància obtinguda en condicions de reproductibilitat, s_R^2 :

$$s_R^2 = \frac{\sum (x_i - \bar{x})^2}{(n - 1)}$$

Com que les mesures de x_i solen ésser úniques (aquí no es tenen s_{etapes}^2 que han hagut de trobar-se separatament com s'ha indicat en paràgrafs anteriors), la variància del laboratori és la suma de dos termes,

$$s_L^2 = s_R^2 + s_{\text{etapes}}^2$$

és a dir,

$$s_L^2 = \frac{\sum (x_i - \bar{x})^2}{(n-1)} + \frac{s_{\text{resposta}}^2}{b_1^2}$$

on x_i correspon al valor de la concentració d'analit calculat mitjançant la corba de calibratge realitzat en les condicions i (dia, analista, instrument...), i \bar{x} correspon al valor mitjà de tots els x_i . Cadascun dels valors x_i trobats mitjançant la seva corresponent recta de calibratge durant el procés de coneixement del mètode, té una imprecisió que es quantificaria mitjançant la variància de l'etapa de calibratge, s_{calib}^2 . Si el nombre de valors i és suficientment elevat, la dispersió de cada valor x_i , introduïda per cada corba de calibratge individual, està englobada en la dispersió de tot el conjunt de valors x_i al voltant del valor mitjà \bar{x} .

3. Hem descrit fins ara els diferents tipus de variàncies parcials que contribueixen a la incertesa total del resultat analític.

La variància final, s_{final}^2 , que s'utilitzarà per al càlcul de la incertesa, està composta pels diferents termes estudiats. Per a deduir l'expressió que els combina tots, cal reprendre els conceptes de càlcul o estimació de les incerteses.

Recordem que una de les variàncies involucrades, la que hem anomenat variància del laboratori, s_L^2 , s'origina per causa dels errors aleatoris independents els uns dels altres perquè provenen d'etapes independents. Si aquests errors estiguessin distribuïts segons una funció de densitat de probabilitat normal o gaussiana, la incertesa associada al valor mitjà estimat dels resultats, deguda exclusivament a aquest tipus d'errors i expressada en les mateixes unitats que una desviació típica o estàndard, seria donada per l'expressió:

$$u_L = [(t_{nR} \cdot s_R / \sqrt{n_R})^2 + (t_{n \text{ etapes}} \cdot s_{\text{etapes}} / \sqrt{n_{\text{etapes}}})^2]^{1/2}$$

Amb aquesta expressió, segons els valors de t buscats a les taules per a cadascun dels graus de llibertat, s'aconsegueix d'establir els

límits d'un interval al voltant del valor mitjà per al qual hi ha una probabilitat coneguda que un resultat nou obtingut en les mateixes condicions caigui fora de l'interval $\bar{x} - u_L$, $\bar{x} + u_L$.

La Western European Calibration Cooperation (WECC) estima que no ha d'acceptar-se la certesa de posseir distribucions gaussianes i recomana substituir el terme t per un altre coeficient, w , que per a una probabilitat del 95 % pren els valors:

núm. de resultats individuals, n	w (factor pel qual cal multiplicar $s_i / \sqrt{n_i}$)
2	7,0
3	2,3
4	1,7
5	1,4
6	1,3
7	1,3
8	1,2
9	1,2
10 o més	1,0

amb el qual l'expressió anterior prendria la forma:

$$u_L = [(\omega_{nR} \cdot s_R / \sqrt{n_R})^2 + (\omega_{n_{etapes}} \cdot s_{etapes} / \sqrt{n_{etapes}})^2]^{1/2}$$

S'estima que emprant els valors d'aquesta taula es cobreix aproximadament el 95 % de probabilitat que el valor veritable es trobi dins de l'interval de semiamplitud $2 \cdot w \cdot s_L / \sqrt{n}$, centrat en el valor mitjà (adaptat del document WECC-19-1990 *Guidelines for the expression of the uncertainty of measurements in calibrations*). Com pot deduir-se, els valors resultants de la incertesa associada als errors aleatoris no difereixen de forma apreciable en aplicar una o altra expressió.

En el valor de la incertesa final del nostre resultat també intervé aquella part que prové de la incertesa associada al material o mètode de referència. És una incertesa que s'«importa» (fonamentalment es distingeix perquè incorpora les causes d'error controlades) i s'expressa com a u_r , essent $u_r = k \cdot s_r$. D'acord amb el que ja coneixem, les incerteses no poden sumar-se com a tals (se sumen les variàncies), per la qual cosa la incertesa composta, u_c , expressada novament com una desviació típica, és donada per l'expressió:

$$u_c = (s_r^2 + u_L^2)^{1/2} = [s_r^2 + (\omega_{nR} \cdot s_R / \sqrt{n_R})^2 + (\omega_{n_{etapes}} \cdot s_{etapes} / \sqrt{n_{etapes}})^2]^{1/2}$$

La incertesa global, o simplement incertesa, u , és donada per l'expressió

$$u = \pm k \cdot u_c$$

$$u = \pm k \cdot [s_r^2 + (\tau_{nR} \cdot s_R / \sqrt{n_R})^2 + (\tau_{n_{etapes}} \cdot s_{etapes} / \sqrt{n_{etapes}})^2]^{1/2}$$

on k és el «factor d'incertesa» o «factor d'inclusió» recomanat per la Western European Calibration Cooperation, el qual defineix l'interval de valors que pot atribuir-se raonablement al resultat. En definitiva, és un factor de seguretat que fa el paper de les t o w esmentades amb anterioritat sense que ara pugui afirmar-se que aquest valor proporciona nivells de confiança del 95 %, 99 % o d'un altre tant per cent. Normalment, k pren valors compresos entre 2 i 3. El valor recomanat és $k = 2$, encara que no té més importància ja que és necessari especificar-lo sempre.

Finalment, si durant el procés analític s'ha portat a terme una operació que comporta algun altre tipus d'incerteses no contemplades, aquestes es transmeten al resultat final de les anàlisis d'acord amb el principi de transmissió d'errors.

REFERÈNCIES BIBLIOGRÀFIQUES

- M. BADER (1980). *J. Chem. Educ.*, **57**, 703.
- M. BARONI, S. CLEMENTI, G. CRUCIANI, G. COSTANTINO, D. RIGANELLI i E. OBERRAUCH (1992). *J. Chemometrics*, **6**, 347-356.
- R. BOQUÉ, F. X. RIUS i D. L. MASSART (1994). *J. Chem. Educ.* [En premsa]
- R. BOQUÉ i F. X. RIUS (1996). *Chemom. and Intel. Lab. Systems*, **32**, 11.
- G. E. P. BOX, W. G. HUNTER i J. S. HUNTER (1978). *Statistics for Experimenters. An Introduction to Design, Data Analysis and Model Building*. Nova York: Wiley.
- M. P. CALLAO, M. S. LARRECHI i F. X. RIUS (1992). A: VALCÁRCCEL i RÍOS [eds.]. *La Calidad en los Laboratorios Analíticos*. Barcelona: Reverté.
- P. CAMPINS (1988). A: M. BLANCO i V. CERDÀ [Ed.]. *Quimiometría*. Bellaterra: Acta.

- M. J. CARDONE (1983). *J. Assoc. Off. Anal. Chem.*, **66**, 1257-1282.
- CETAMA (1986). *Statistique Appliquée à l'Exploitation des Mesures*. 2a ed. París: Masson.
- S. CHATTERJEE i A. S. HADI (1986). *Statist. Sci.*, **1**, 379.
- C. A. CLAYTON, J. W. HINES i P. D. ELKINS (1987). *Anal. Chem.*, **59**, 2506-2514.
- A. COLINA, M. C. GARNICA, J. LÓPEZ PALACIOS, M. C. ORTIZ, C. PIZARRO i L. A. SARABIA (1988). *J. Chemometrics*, **3**, 323-327.
- G. CRUCIANI, M. BARONI, S. CLEMENTI, G. COSTANTINO, D. RIGANELLI i B. SKAGERBERG (1992). *J. Chemometrics*, **6**, 335-346.
- L. A. CURRIE (1968). *Anal. Chem.*, **40**, 568.
- L. A. CURRIE [ed.] (1988). *Detection in Analytical Chemistry. Importance, Theory, and Practice*. Washington: American Chemical Society.
- K. DANZER, U. WANK i D. WIENKE (1988). *Chemom. and Intell. Lab. Systems*, **12**, 69-79.
- M. DAVIDIAN i P. D. HAALAND (1990). *Chemom. and Intell. Lab. Systems*, **9**, 231-248.
- N. DRAPER i H. SMITH (1981). *Applied Regression Analysis*. 2a ed. Nova York: Wiley.
- J. P. DUX (1990). *Handbook of Quality Assurance for the Analytical Chemistry Laboratory*. Nova York: Van Nostrand Reinhold.
- E. S. EDGINGTON (1987). *Randomization Tests*. 2a ed. Boca Raton: Marcel Dekker.
- M. H. FEINBERG (1988). *J. Chemometrics*, **3**, 103-114.
- G. T. FLATMAN i J. W. MULLINS (1985). «The Alpha and Beta of Chemometrics». A: J. J. BREEN, i P. E. ROBINSON, [eds.]. *Environmental Applications of Chemometrics*. Washington: American Chemical Society.
- M. FORINA, S. LANTERI, R. BOGGIA i E. BELTRÁN (1993). *Quím. Anal.*, **12**, 128-135.
- F. M. GARFIELD (1993). *Principios de Garantía de Calidad para Laboratorios Analíticos*. 2a ed. Arlington, VA: AOAC International.
- R. GARGALLO, R. TAULER, F. CUESTA-SÁNCHEZ i D.L. MASSART (1996). *Trends Anal. Chem.*, **15**, 279.
- F. C. GARNER i G. L. ROBERTSON (1988). *Chemom. and Intell. Lab. Systems*, **3**, 53-59.

- S. GEISSER (1975). *J. Am. Stat. Assoc.*, **70**, 328.
- P. J. GEMPERLINE i A. SALT (1989). *J. Chemometrics*, **3**, 343-357.
- T. F. HARTLEY (1987). *Computerized Quality Control: Programs for the Analytical Laboratory*. Nova York: Ellis Horwood.
- C. HARTMANN, J. SMEYERS-VERBEKE i D. L. MASSART (1993). *Analusis*, **21**, 125-132.
- W. HORWITZ (1988). *Pure and Appl. Chem.*, **60**, 855-864.
- A. HUBAUX i G. VOS (1970). *Anal. Chem.*, **42**, 849.
- J. S. HUNTER (1981). *J. Assoc. Off. Anal. Chem.*, **64**, 574-583.
- IUPAC (1978). *Spectrochim. Acta*, **33B**, 242-247.
- R. A. JOHNSON i D. W. WICHERN (1988). *Applied Multivariate Statistical Analysis*. 2a ed. Englewood Cliffs: Prentice Hall.
- H. KAISER (1970). *Anal. Chem.*, **42**, 26A.
- G. KATEMAN i F. W. PIJERS (1981). *Quality Control in Analytical Chemistry*. Nova York: Wiley.
- O. KVALHEIM i Y. LIANG (1992). *Anal. Chem.*, **64**, 936-946.
- P. M. LANG i J. KALIVAS (1993). *J. Chemometrics*, **7**, 153-164.
- S. LANTERI (1992). *Chemom. and Intell. Lab. Systems*, **15**, 159-169.
- R. LEPAGE i L. BILLARD (1992). *Exploring the Limits of Bootstrap*. Nova York: Wiley.
- R. DE LEVIE (1986). *J. Chem. Educ.*, **63**, 10-15.
- Y. LIANG, O. KVALHEIM, H. R. KELLER, D. L. MASSART, P. KIECHLE i F. ERNI (1992). *Anal. Chem.*, **64**, 946-953.
- Y. Z. LIANG i O. M. KVALHEIM (1993). *Chemom. and Intell. Lab. Systems*, **20**, 115-125.
- C. LITEANU i I. RICA (1980). *Statistical Theory and Methodology of Trace Analysis*. Nova York: Wiley.
- E. R. MALINOWSKI (1991). *Factor Analysis in Chemistry*. 2a ed. Nova York: Wiley. [1a ed., 1980]
- J. MANDEL (1991). *Chemom. and Intell. Lab. Systems*, **11**, 109-119.
- J. MANDEL i F. J. LINNIG (1957). *Anal. Chem.*, **29**, 743-749.
- H. MARTENS i T. NAES (1991). *Multivariate Calibration*. Chichester: Wiley.
- D. L. MASSART (1988). *Trends Anal. Chem.*, **7**, 157-158.
- D. L. MASSART (1993). «Method Validation. An Introduction». A: *European School on Chemometrics. Course Notes*. Santiago de Compostella.
- D. L. MASSART, L. KAUFMAN, P. J. ROUSSEEUW i A. LEROY (1986). *Anal. Chim. Acta*, **187**, 171.

- D. L. MASSART, B. G. M. VANDEGINSTE, S. N. DEMING, Y. MICHOTE i L. KAUFMAN (1989). *Chemometrics: a Textbook*. Amsterdam: Elsevier.
- D. L. MASSART, J. SMEYERS-VERBEKE i F. X. RIUS (1989). *Trends Anal. Chem.*, **8**, 49-51.
- P. C. MEIER i R. EL. ZÜND (1993). *Statistical Methods in Analytical Chemistry*. Nova York: Wiley.
- M. MELLINGER (1987). *Chemom. and Intell. Lab. Systems*, **2**, 61-77.
- J. C. MILLER i J. N. MILLER (1988). *Statistics for Analytical Chemistry*. 2a ed. Nova York: Wiley.
- J. N. MILLER (1991). *Analyst*, **116**, 3-14.
- B. C. MITCHELL i D. S. BURDICK (1993). *Chemom. and Intell. Lab. Systems*, **20**, 149-162.
- J. C. OLUCHA i F. X. RIUS (1990). *Trends Anal. Chem.*, **9**, 77-79.
- D. W. OSTEN (1988). *J. Chemometrics*, **2**, 39-48.
- M. H. QUENOUILLE (1966). *Introductory Statistics*. Oxford: Pergamon Press.
- C. R. RAO (1973). *Linear Statistical Inference and Its Applications*. 2a ed. Nova York: Wiley.
- G. RAURET (1992). «Ejercicios de intercomparación». A: M. VALCÁRCCEL, i A. RÍOS, [ed.]. *La Calidad en los Laboratorios Analíticos*. Barcelona: Reverté.
- J. RIU i F. X. RIUS (1995). *J. Chemometrics*, **9**, 343.
- J. RIU i F. X. RIUS (1996). *Anal. Chem.*, **68**, 1851.
- F. X. RIUS, J. SMEYERS-VERBEKE i D. L. MASSART (1989). *Trends Anal. Chem.*, **8**, 8-11.
- F. X. RIUS (1992). *La Calidad en los Laboratorios Analíticos*. Barcelona: Reverté.
- P. J. ROUSSEEUW i A. M. LEROY (1987). *Robust Regression and Outlier Detection*. Nova York: Wiley.
- J. F. RUSLING (1989). *Crit. Rev. Anal. Chem.*, **21**, 49-81.
- L. SARABIA i C. ORTIZ (1994). *Trends Anal. Chem.*, **13**, 1.
- B. E. H. SAXBERG i B. R. KOWALSKI (1979). *Anal. Chem.*, **51**, 1.031.
- K. J. SCHOSTACK i E. R. MALINOWSKI (1993). *Chemom. and Intell. Lab. Systems*, **20**, 173-182.
- R. TAULER, A. IZQUIERDO-RIDORSA, R. GARGALLO i E. CASASSAS (1995). *Chemom. and Intell. Lab. Systems*, **27**, 163.
- J. K. TAYLOR (1987). *Quality Assurance of Chemical Measurements*. Chelsea, Michigan: Lewis Publ.

- A. THIELEMANS, P. J. LEWI i D. L. MASSART (1988). *Chemom. Intell. Lab. Systems*, **3**, 277.
- J. TOFT i O. KVALHEIM (1993). *Anal. Chem.*, **65**, 2270-2275.
- R. L. TRANTER (1990). *Anal. Proceedings*, **27**, 299-300.
- G. A. URIANO i J. P. CALI (1977). «Role of reference materials and reference methods in the measurement process». A: J. R. DE VOL, [ed.]. *Validation of the Measurement Process*. Washington: American Chemical Society.
- M. VALCÁRCEL i A. RÍOS, [ED] (1992). *La Calidad en los Laboratorios Analíticos*. Barcelona: Reverté.
- C. VANDENBOSCH, V. THERRY i D. L. MASSART (1990). *Trends Anal. Chem.*, **9**, 283-284.
- P. VANKEERBERGHEN i J. SMEYERS-VERBEKE (1992). *Chemom. and Intell. Lab. Systems*, **15**, 195-202.
- S. J. VANSLYKE i P. D. WENTZELL (1993). *Chemom. and Intell. Lab. Systems*, **20**, 183-195.
- G. VERNIMONT (1977). A: J. R. DE VOL, [ed.]. *Validation of the Measurement Process*. Washington: American Chemical Society, 1-29. [ACS Symposium Series núm. 63]
- G. VERNIMONT (1985). *Use of Statistics to Develop and Evaluate Analytical Methods*. 2a ed. Arlington, VA: AOAC.
- Y. WANG, D. J. VELTKAMP i B. R. KOWALSKI (1991). *Anal. Chem.*, **63**, 2750-2756.
- Y. WANG, M. J. LYSAGHT i B. R. KOWALSKI (1992). *Anal. Chem.*, **64**, 562-564.
- R. WOLTERS i G. KATEMAN (1989). *J. Chemometrics*, **3**, 329-342.
- R. WOLTERS, A. C. M. DEN BROEK i G. KATERMAN (1990). *J. Chemometrics*, **9**, 143-175.
- W. J. YOUDEN i E. H. STEINER (1975). *Statistical Manual of the Association of Official Analytical Chemists*. Arlington, VA: AOAC.
- H. YUZHU, J. SMEYERS-VERBEKE i D. L. MASSART (1990). *Chemom. and Intell. Lab. Systems*, **9**, 31-44.