

Gewichtete lineare Regression zur Kalibrierung

Hans-Joachim Kuss, Putzbrunn

In der Chromatographie erwartet man im Regelfall einen linearen Zusammenhang zwischen dem Messwert und der Konzentration. Dies muss natürlich verifiziert werden. Die beiden wichtigsten Detektoren in der Chromatographie, UV-Detektor in der HPLC und FID in der GC haben einen großen linearen Bereich. Massenspektrometer haben einen ausreichend großen linearen Bereich.

Sinnvollerweise wird man sich bei einer Chromatogrammserie immer mal wieder überzeugen, ob eine Leerprobe tatsächlich keine Veränderung der Basislinie an den Stellen zeigt, an denen bei Analysenproben Peaks erwartet werden. Das heißt, dass bei einer Konzentration von Null auch der Messwert Null sein und die lineare Kalibrierfunktion durch den Nullpunkt gehen muss.

Im Gegensatz dazu hat man bei der Messung der UV-Absorption in einer Küvette einen Messwert für die Untergrundabsorption. In der HPLC mit UV-Detektion wird diese Untergrundabsorption durch das Nullen (Autozero) des Detektors für alle Messwerte abgezogen. Ein chromatographisches Messsignal in einer Leerprobe an der Stelle des erwarteten Peaks kann nur als unzureichende Methode gewertet werden, was bedeutet, dass die Methode verbessert werden muss. Dies ist nicht zu verwechseln mit der stetigen Gefahr, eine zusätzliche interferierende Substanz unter dem Peak einer einzelnen Analysenprobe zu finden. Dagegen ist keine Methode vollständig gefeit.

Trotzdem wird in der chromatographischen Auswertung generell eine lineare Regression mit Achsenabschnitt durchgeführt. Dabei ist der Achsenabschnitt fast immer nicht signifikant. Auch aus statistischen Gründen muss dann der Achsenabschnitt auf Null gesetzt werden, weil sein Wert nur durch den Zufall bedingt ist.

Bei der Auswertung spielt ein nichtsignifikanter Achsenabschnitt nur dann eine Rolle, wenn der Arbeitsbereich groß ist. In diesem Fall kann ein (zufallsbedingter, d.h. nicht reproduzierbarer) Achsenabschnitt die kleinen Konzentrationen um mehr als 100% verfälschen. Ist der Achsenabschnitt

signifikant, sollte ebenfalls an eine unzureichende Methode gedacht werden.

Die Regression wird auch Methode der kleinsten quadratischen Abweichungen genannt. Zur Kalibrierung wird eine lineare Funktion mit n ($K=1$ bis n ; $n=5$ bis 10) Kalibrierpunkten verwendet, die durch eine sehr exakt bekannte Konzentration x_K und ein mit einem Rauschen versehenen Messsignal y_K beschrieben werden.

$$y = bx + a$$

In der allgemeingültigen Form wird das Minimum für das Quadrat aus der Differenz des Messwerts y_K und dem aus der Geradengleichung errechneten Messwert $y(x_K)$ errechnet, normiert auf die entsprechenden Varianzen, d.h. die x-Abweichungen werden als vernachlässigbar angesehen.

$$\sum_{K=1}^n \frac{(y_K - y(x_K))^2}{S y_K^2} = \text{Minimum}$$

Um so rechnen zu können, muss die Standardabweichung Sy an jedem Kalibrierpunkt bekannt sein. Aus statistischer Sicht wäre es wünschenswert 10 Kalibrierpunkte 10-mal zu messen. Aus der Sicht des analytischen Chemikers wäre das ein unakzeptabel hoher Aufwand.

Zur Vereinfachung nimmt man an, dass die Standardabweichungen im Arbeitsbereich y_{K1} bis y_{Kn} gleich sind. Zur Absicherung macht man nur am untersten und obersten Kalibrierpunkt eine Mehrfachmessung und berechnet den F-Wert nach:

$$F = \frac{S y_{Kn}^2}{S y_{K1}^2}$$

Findet man keinen signifikanten Unterschied, nimmt man an, dass obige Annahme gerechtfertigt ist und es können die Residuen R minimiert werden:

$$R = \sum (y_K - y(x_K))^2 = \text{Minimum}$$

Die Residuen, d.h. der Abstand der gemessenen und berechneten Signalwerte sollten grundsätzlich in einem Residuenplot graphisch aufgetragen werden. Man kann daraus leicht die Homogenität der Abweichungen über den Konzentrationsbereich oder einen Trend erkennen, der evtl. auf eine quadratische Regressionsfunktion hinweist.

$$\bar{x} = \text{Mittelwert } x \quad \bar{y} = \text{Mittelwert } y$$

$$a = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2} = \frac{SAW_{xy}}{SAW_{xx}} \quad \text{SAW=Summe der Abweichungen}$$

Die Bestgerade geht immer durch den Zentroid \bar{x} und \bar{y} . Ist a bekannt, lässt sich b berechnen:

$$b = \bar{y} - a\bar{x}$$

Aus der Varianz der y-Werte Sy^2 und der RSD lässt sich der Anteil der Varianz berechnen, der durch den Zusammenhang zwischen x- und y-Werten erklärt wird:

$$\text{Bestimmtheitsmaß} \quad r^2 = \frac{Sy^2 - RSD^2}{Sy^2}$$

Besteht überhaupt keine Korrelation zwischen x und y, dann ist $RSD=Sy$ und $r^2=0$. Je stärker der Zusammenhang wird, desto kleiner wird RSD gegen Sy . Für $RSD=0$ ergibt sich ein vollständiger Zusammenhang mit $r^2=1$ und dem Korrelationskoeffizienten $r=1$.

Sind a und b bekannt, lässt sich aus gemessenen Signalwerten die Konzentration einer Analysenprobe berechnen:

$$x = \frac{y}{b} - \frac{a}{b}$$

Die RSD ist eine Standardabweichung in y-Richtung. Sie kann analog zur Umrechnung von y- in x-Werte durch Division durch die Steigung in x-Richtung dargestellt werden. Damit erhält man die in den Validierungsrichtlinien genannte Verfahrensstandardabweichung. Diese

dividiert man durch \bar{x} , um den Verfahrensvariationskoeffizienten VVK zu erhalten:

$$VSD = \frac{RSD}{a} \qquad VVK = \frac{RSD}{x \cdot a} = \frac{VSD}{x}$$

Diese Kenngrößen sind nicht in den üblichen Statistikprogrammen enthalten, lassen sich aber in Excel einfach berechnen.

Die RSD ist die zentrale Größe zur Berechnung der Standardabweichung der Steigung und des Achsenabschnitts, der Nachweis- und der Bestimmungsgrenze sowie der Ergebnisunsicherheit U der errechneten Analysenwerte.

Leider findet sich in der Chromatographie bereits bei einem Arbeitsbereich von 10 oder größer häufig ein signifikanter F-Test. Das liegt daran, dass in der Chromatographie eher der Relativfehler konstant ist, als der Absolutfehler, wie bei der vereinfachten Regression angenommen, d.h.:

$$VK = const. = \frac{Sy}{y}$$

Bei einem konstanten Verhältnis zwischen Standardabweichung des Messwerts und dem Messwert selbst kann der Messwert als Schätzung der Standardabweichung benutzt werden. Damit wird obige Gleichung zu:

$$\sum \frac{(y_k - y(x_k))^2}{y_k^2} = Minimum$$

Dies ist nichts Anderes als die bekannte $1/y^2$ Gewichtung der Varianzen, die bei den Integrationsprogrammen angewählt werden kann. Übliche weitere Gewichtungen sind $1/y$, auch $1/x$ und $1/x^2$ und auch $1/y^{1/2}$. Da ein Schätzwert für SDy gesucht wird, ist generell eine y-Gewichtung sinnvoller, als eine x-Gewichtung.

Den Fall gleicher Varianzen (Varianzenhomogenität; keine Gewichtung) könnte man „Apparatevarianz“ nennen. Dies wird angenähert der Fall sein, wenn wir nur eine einfache physikalische Messung durchführen. Die $1/y^2$ Gewichtung, die auf einen konstanten VK im Messbereich zurückzuführen ist, könnte man „Volumenabhängige Varianz“ nennen. Dies beschreibt den

Fall, wenn die Probenvorbereitung die Abweichungen dominiert. Die sehr wahrscheinlichen Volumenfehler müssen per se konzentrationsabhängig sein.

Nun sind alle Fälle zwischen diesen beiden Extremen einer Gewichtung w mit 1 (=keine Gewichtung) und einer $1/y^2$ Gewichtung vorstellbar, die allgemein durch $1/y^{WE}$ beschrieben werden können. Alle oben genannten y -Gewichtungen sind enthalten, aber zusätzlich ungeradzahlige Gewichtungen erlaubt. Der ungewichtete Fall ist durch den Wichtungsexponenten $WE=0$ beschrieben. Damit wird:

$$\sum w(y_k - y(x_k))^2 = \min \quad \text{mit} \quad g = \frac{1}{y^{WE}} \quad \text{und} \quad w = \frac{n^* g}{\sum g}$$

Das Problem besteht darin, WE aus den vorhandenen Daten möglichst sinnvoll abzuschätzen. Dies ist aus dem Varianzquotienten (F-Wert) unter Beachtung des y -Arbeitsbereichs nach folgender Gleichung möglich:

$$WE = \frac{\log F}{\log AB_y} = \frac{\log S_{y_{Kn}}^2 - \log S_{y_{K1}}^2}{\log y_{Kn} - \log y_{K1}}$$

Dieses Verfahren könnte Varianzquotientengewichtung genannt werden.

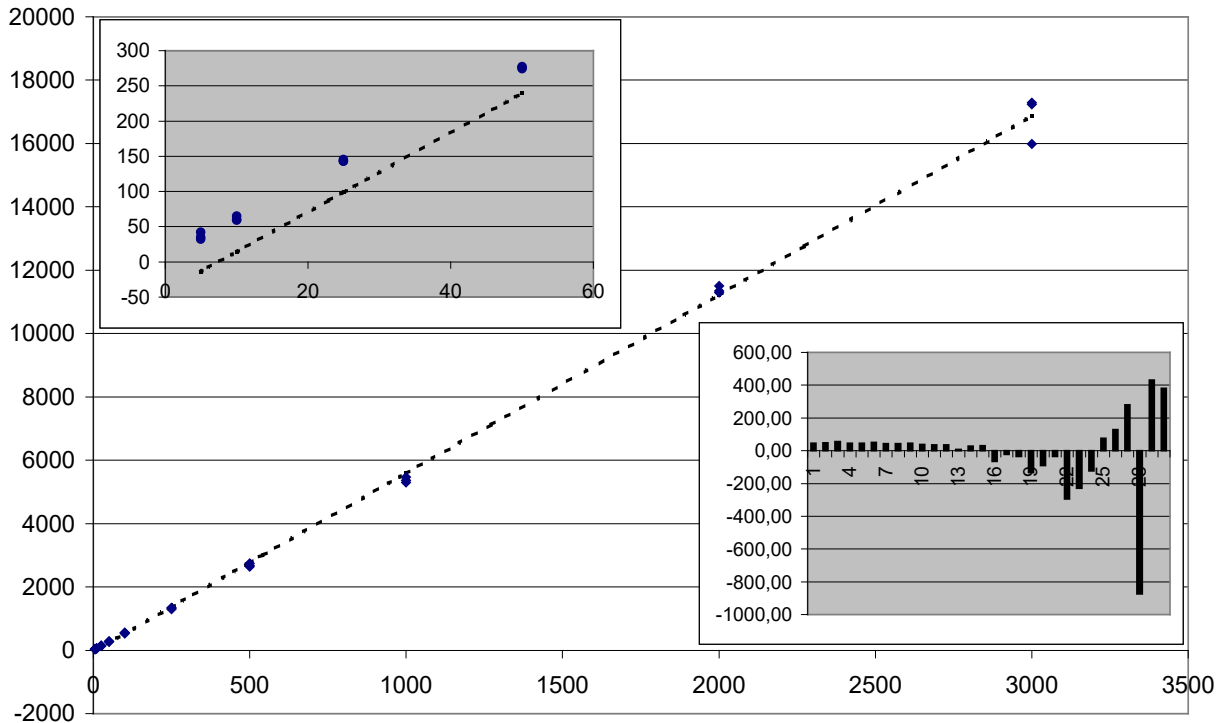
Johnson, 1988, zeigt eines der wenigen Beispiele, bei denen Originalmessungen veröffentlicht wurden und bei dem Mehrfachmessungen für jeden Kalibrierpunkt vorliegen, so dass die jeweiligen S_y bekannt sind.

x	y	y(x)	Resid	x(y)	x(y)Gewicht	x(y)Null
5	32	-13,79	45,79	13,1	4,5	5,7
5	35	-13,79	48,79	13,7	5,1	6,2
5	42	-13,79	55,79	14,9	6,3	7,5
10	60	14,37	45,63	18,1	9,6	10,7
10	59	14,37	44,63	17,9	9,4	10,5
10	65	14,37	50,63	19,0	10,5	11,6
25	143	98,85	44,15	32,8	24,8	25,5
25	143	98,85	44,15	32,8	24,8	25,5
25	145	98,85	46,15	33,2	25,1	25,8
50	277	239,65	37,35	56,6	49,2	49,4
50	274	239,65	34,35	56,1	48,7	48,8
50	276	239,65	36,35	56,5	49,0	49,2
100	530	521,25	8,75	102	95	94

100	550	521,25	28,75	105	99	98
100	552	521,25	30,75	105	99	98
250	1302	1366,05	-64,05	239	236	232
250	1343	1366,05	-23,05	246	244	239
250	1331	1366,05	-35,05	244	242	237
500	2642	2774,05	-132,05	477	481	471
500	2683	2774,05	-91,05	484	488	478
500	2740	2774,05	-34,05	494	499	488
1000	5295	5590,06	-295,06	948	965	944
1000	5360	5590,06	-230,06	959	977	955
1000	5468	5590,06	-122,06	978	997	974
2000	11297	11222,08	74,92	2013	2060	2013
2000	11352	11222,08	129,92	2023	2070	2023
2000	11502	11222,08	279,92	2050	2098	2050
3000	15981	16854,10	-873,10	2845	2915	2848
3000	17286	16854,10	431,90	3077	3153	3080
3000	17235	16854,10	380,90	3068	3144	3071

Die Originalwerte (peak height ratio response) in Spalte 2 wurden mit 1000 multipliziert, um sie leserlicher zu machen. Das beeinflusst die Rückrechnung in Konzentrationen nicht.

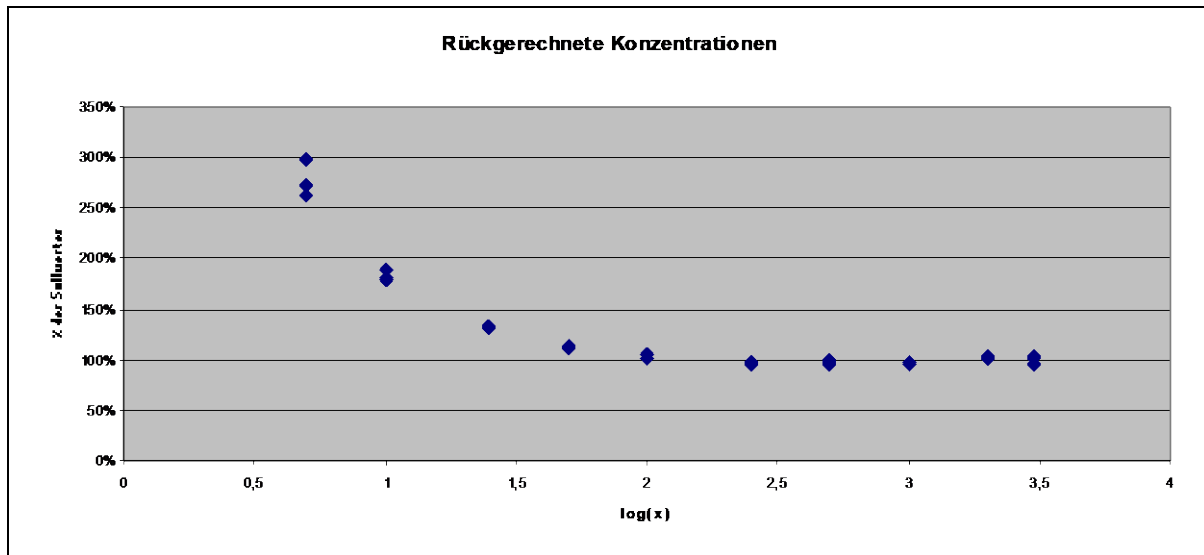
Die lineare Regression ohne Gewichtung ergibt die Geradengleichung $y=5,63x-41,96$, die graphisch plausibel aussieht – es fällt nur die große Variabilität des obersten Punktes auf. Anders sieht es aus, wenn man nur die ersten vier Kalibrierpunkte betrachtet, die man in der großen Graphik nicht mehr erkennen kann, die aber in der kleinen Graphik links oben deutlich zu sehen sind. Es ist offensichtlich, dass die unteren Kalibrierpunkte praktisch überhaupt nicht berücksichtigt werden. Dies zeigt auch die Abbildung der Residuen $R=y_K-y(x_K)$ in der kleinen Graphik rechts unten. Die $x(y)$ -Werte in Spalte 5 sind schon in der Arbeit von Johnsson enthalten, der eine $1/y$ -Gewichtung als Lösung des Problems vorschlägt.



Jede Abweichung von der Geraden, d.h. jedes Residuum zieht die Gerade zu sich hin, sogar sehr stark, weil ja die Abweichungsquadrate aufsummiert werden. Wenn im unteren Teil der Gerade nur kleine Abweichungen auftreten, im oberen Teil nur große Abweichungen (inhomogene Varianzen, Heteroskedastizität), dann bleiben die unteren Punkte unberücksichtigt, weil sie kein Gewicht haben und die Gerade kann an ihnen zu weit vorbeigehen. Die Abweichungen bei den kleineren Konzentrationen haben die höhere Gewichtung „verdient“, weil sie (in Absoluteinheiten) erheblich genauer gemessen werden können.

Nehmen wir an, dass wir Messwerte hätten, die gerade den Kalibrierwerten entsprechen. Zur Zurückrechnung in Konzentrationen würde man graphisch einen Strahl in Höhe des Signals parallel zur x-Achse, d.h. durch die Messpunkte bis zur Geraden ziehen und von dort zur x-Achse, um die Konzentration abzulesen. Dadurch, dass die Gerade deutlich unterhalb der untersten Kalibrierpunkte verläuft und einen negativen, aber nicht signifikanten Achsenabschnitt hat, ergeben kleine Signalwerte viel zu hohe Konzentrationen.

Noch deutlicher zeigen dies die aus den Messwerten der Kalibrierkonzentrationen nach der Geradengleichung zurückgerechneten Konzentrationen in % des Soll-Konzentrationswertes.



Die leichten negativen Abweichungen im mittleren oberen Konzentrationsbereich wiegen die großen relativen Abweichungen im unteren Konzentrationsbereich auf, weil sie als Absolutwert gleich sind. Es ist offensichtlich, dass die Abweichungen der niedrigsten Kalibrierkonzentrationen völlig unakzeptabel sind. Bei einem Messwert von Null errechnen sich bereits 7,5 ng/ml (=b/a). Diese Linearverschiebung beeinflusst die unteren Punkte viel mehr, als die oberen Punkte.

Der Wichtungsexponent errechnet sich nach obiger Gleichung zu 1,62. Dieser Wert wurde in eine entsprechende Excel-Tabelle eingesetzt:

20748	F-Wert
4,3170	log(F)
463	AB y
2,6659	log(AB y)
1,62	WE

Die unteren 4 Kalibrierpunkte liegen mit Gewichtung fast exakt auf der Geraden. Es ergibt sich aus der Geradengleichung $y=5,48x+7,25$ ein positiver Achsenabschnitt, der 1,3 ng/ml entspricht. Die rückgerechneten Ergebnisse sind in der Spalte x(y)Gewicht dargestellt. Der Vergleich von x(y) und x(y)Gewicht zeigt, welche ungenügenden Ergebnisse die normale

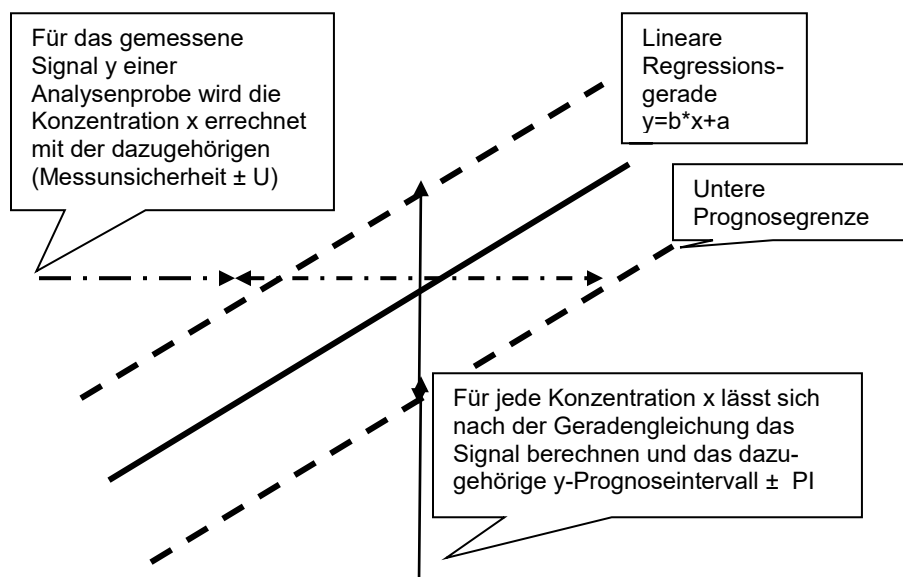
lineare Regression bringt. Mit der Nullpunktsgerechten weichen die Werte $x(y)$ nur etwas mehr von den wahren Werten im unteren Konzentrationsbereich ab, als mit der Gewichtung.

Ein weiterer wesentlicher Grund, eine Gewichtung vorzunehmen, liegt in der adäquaten mathematischen Abbildung realistischer Messunsicherheiten U . Zu diesem Thema schrieb Doerffel 1990: „Deshalb muss die gewichtete Regression immer dann angewandt werden, wenn aus den Messungen Aussagen zur Präzision der Ergebnisse gefordert sind.“

Die bekannten Gleichungen zum Prognoseintervall PI (y -Abweichungen) und zur Messunsicherheit U (x -Abweichungen) verkomplizieren sich durch die Gewichtung nur geringfügig:

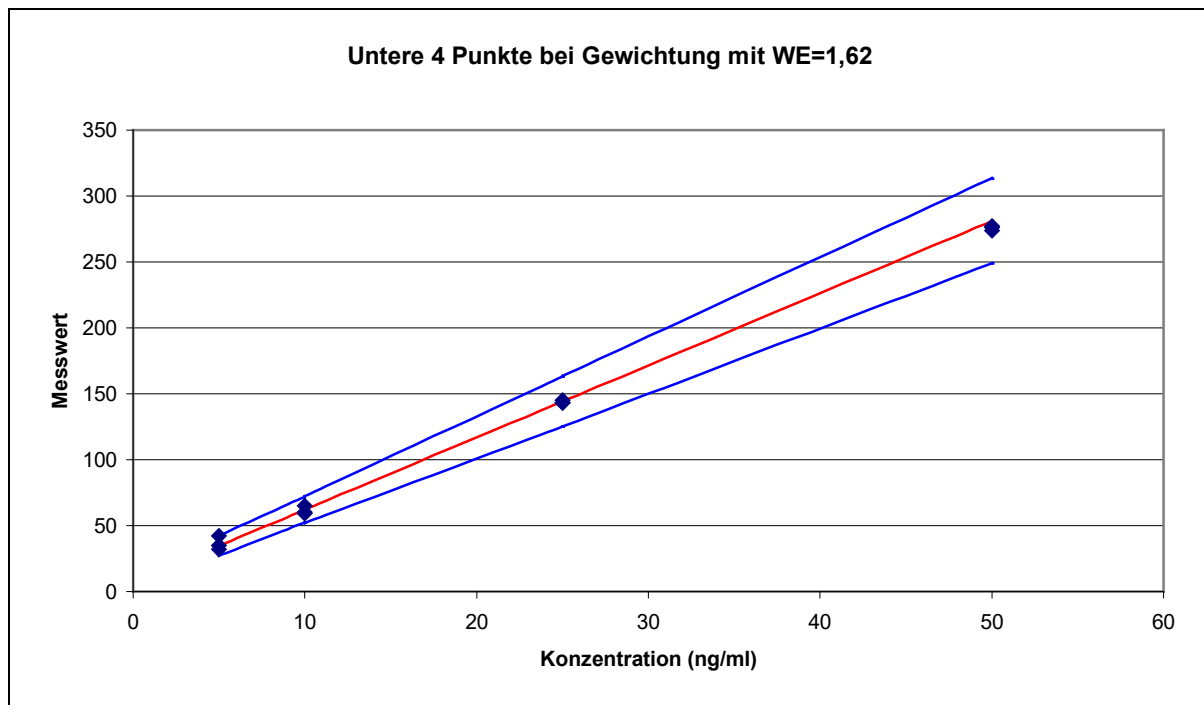
$$PI = t * RSD * \sqrt{\frac{1}{w_k} + \frac{1}{\sum w} + \frac{(x_k - \bar{x}_w)^2}{\sum (x_k - \bar{x}_w)^2}}$$

$$U = \frac{t * RSD}{b} \sqrt{\frac{1}{w_a} + \frac{1}{\sum w} + \frac{(y_a - \bar{y}_w)^2}{b^2 \sum (x_a - \bar{x}_w)^2}}$$



Beide Gleichungen ähneln sich sehr, weil U auf PI zurückgeführt wird. Die Unsicherheit der errechneten Konzentrationen resultiert aus der Unsicherheit der Messwerte. Graphisch wird dieselbe Grenzlinie einmal in y und einmal in x Richtung abgelesen.

Mit Gewichtung sind auch die unteren Konzentrationen im zulässigen Bereich der Messunsicherheit, wie die Abbildung demonstriert.



Insbesondere bei einem großen Arbeitsbereich macht man sich mit normaler linearer Regression durch eine letztlich falsche Rechnung (falls der F-Test signifikant ausfällt) die eigene meist mühsam optimierte Messung im unteren Konzentrationsbereich kaputt. Das schmerzt.

Literatur:

Doerffel, K, Statistik in der analytischen Chemie, Deutscher Verlag für Grundstoffindustrie GmbH 1990.

Johnson, EL, Reynolds, DL, Wright, DS, Pachla, LA, J Chromtogr Sci 26, 372 (1988).

Kuss, HJ, Weighted Least-Squares Regression in Practice: Selection of the Weighting Exponent; LCGC Europe, December 2003.

Kuss, HJ, Quantifizierung in der Chromatographie; in: [HPLC-Tipps 2](#) (Hrsg.: Stavros Kromidas), Hoppenstedt Bonnier Zeitschriften GmbH 2003.

Miller, JN, Miller, JC, Statistics and Chemometrics for Analytical Chemistry, Pearson Education Limited 2000.