

## „Kalibrierung, Linearität und Korrelation – Teil 1“

Ein Artikel unseres Referenten Dipl.-Ing. Wolfgang Gottwald

Jeder Analytiker macht sich zwangsläufig Gedanken bei einer indirekten analytischen Methode über den Zusammenhang zwischen den erhaltenen Signalen (z. B. Peakfläche, Extinktion etc.) und den Konzentrationen des Analyten in einer zu analysierenden Probe. Dieser Vorgang wird Kalibrierung genannt. Der Validierungsparameter, der diesen Zusammenhang beschreibt, wird als „Linearität“ bezeichnet – leider. Der Begriff suggeriert, dass der Zusammenhang zwischen dem Signal und der Konzentration durch eine „Gerade“ beschrieben werden soll. Das ist auch bei vielen Methoden der Fall, z. B. wenn die physikalischen Gesetzmäßigkeiten ein „Geradengesetz“ aufweisen. Zum Beispiel ist die Extinktion  $E$  durch das Lambert-Beer-Gesetz in gewissen Konzentrationsbereichen linear abhängig von der Konzentration  $c$  des Analyten und ergibt im Idealfall eine Gerade. Auch chromatografische Methoden weisen üblicherweise dieses Verhalten auf. Es gibt jedoch Methoden, bei denen eine Kalibriergerade durch das zugrunde liegende Verfahren ausgeschlossen ist, z. B. Immunoassays-Methoden. Und vor allem Störgrößen, wie z. B. Einflüsse bei der Probenvorbereitung, können auch an sich lineare Methoden so verändern, dass es zu keiner Geraden kommt. Bei einer Untersuchung über die Linearität wird daher geprüft, wie sich das Signal in Abhängigkeit von der Konzentration beschreiben lässt und welchen mathematischen Zusammenhang es zwischen den beiden Größen  $x$  und  $y$  gibt. Durch den Einsatz diverser Softwarepakete (z. B. MVA, Validat, SQS etc.) oder durch die Verwendung von EXCEL-Funktionen werden durch eine „lineare“ oder „quadratische“ Regression die Größen: Steigung (Empfindlichkeit)  $b$ , Ordinatenabschnitt  $a$ , Korrelationskoeffizient  $R$  berechnet. Dazu kommt zusätzlich im Falle eines quadratischen Zusammenhangs der quadratische Term  $c$  in  $y = c \cdot x^2 + b \cdot x + a$ .

Anschließend muss der Analytiker entscheiden, welches mathematische Modell - ob „linear“ oder „quadratisch“ - die besten Umrechnungsergebnisse zwischen Signal und Konzentration liefert und ob die analytische Methode in bestimmten Grenzen akzeptabel ist. Dabei ist zuvor zu prüfen, ob es ggf. bei dem Kalibrierexperiment zu groben Fehlern gekommen ist, die zu Ausreißern geführt haben.

Im Rahmen dieser kleinen Abhandlung soll der Schwerpunkt der Betrachtung auf die Korrelation gelegt werden. Prinzipiell stellt eine Korrelation den Zusammenhang zwischen zwei Größen fest. Aber Vorsicht! Eine Korrelation beschreibt keine Ursache-Wirkungs-Beziehung. Aus einem starken Zusammenhang folgt nicht zwangsweise, dass es auch eine eindeutige Ursache-Wirkungs-Beziehung gibt, das gilt vor allem dann, wenn mehrere Größen zusammenwirken.

Beispiel: In einem heißen Sommer wird zur Abkühlung viel Mineralwasser getrunken und gleichzeitig tritt eine erhöhte Zahl von Sonnenbränden auf, eine klassische steigende Korrelation. Trotzdem kann nicht daraus geschlossen werden, dass Mineralwassertrinken zu Sonnenbrand führt. Die gemeinsame Größe ist die Sonneneinstrahlung. Im analytischen Bereich erwartet man jedoch eine eindeutige Korrelation zwischen Signal und Konzentration, die aber nicht immer gegeben ist.

Bei der Abschätzung der Korrelation stehen drei Fragen im Vordergrund:

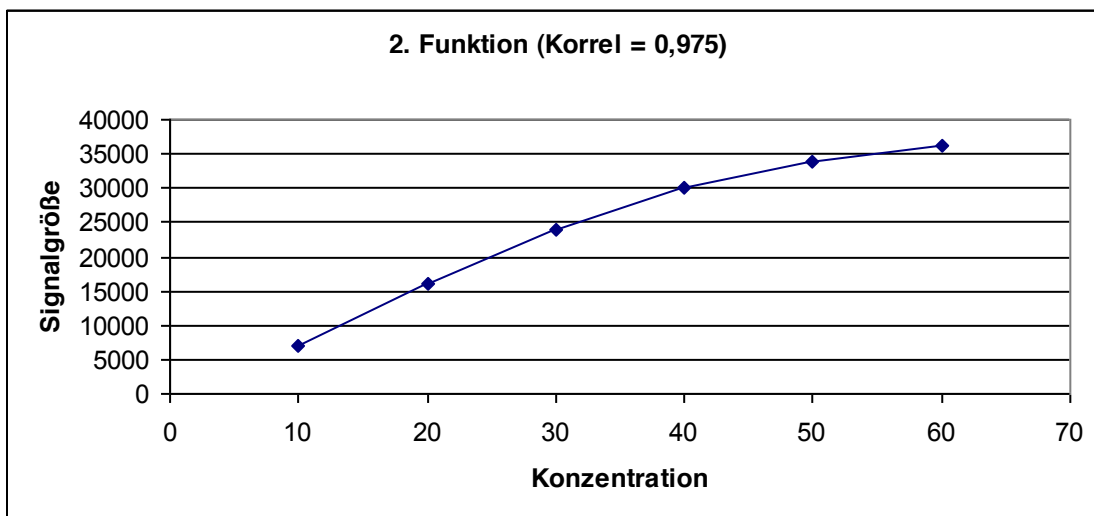
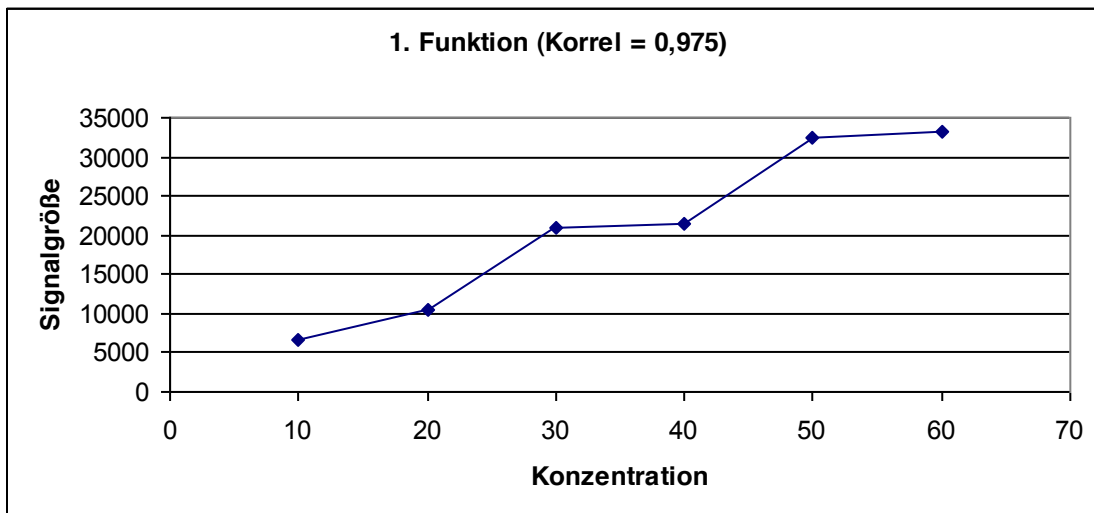
1. Wie stark ist die Korrelation, also der Zusammenhang zwischen Signal und Konzentration?
2. Welche Richtung hat die Korrelation, also absteigend oder aufsteigend?
3. Ist das mathematische Modell geeignet oder gibt es ein anderes, besseres Modell?

Zur Beurteilung der Korrelation wird meistens mit den einschlägigen Statistikpaketen eine Indexzahl berechnet, die Korrelationskoeffizient  $R$  genannt wird. Er wurde vorgeschlagen

von dem britischen Mathematiker Karl Pearson (\* 1857 in London; † 1936 in London) nach Vorarbeiten des französischen Physikers Auguste Bravais (\* 1811 in Annonay; † 1863 in Le Chesnay).

Dieser Indexwert bewegt sich zwischen  $R = -1,000$  (strenge absteigende Korrelation),  $R = 0$  (keine Korrelation zu erkennen) und  $R = +1,000$  (strenge aufsteigende Korrelation). Aber der Indexwert Korrelationskoeffizient  $R$  ist in der Analytik nur mit großer Vorsicht zu verwenden.

Zum Beispiel haben beide Funktionen in Abb. 1 denselben hohen Korrelationskoeffizienten von  $R = 0,975$ . Während die 1. Funktion zwar einen prinzipiellen linearen Verlauf mit erheblichen Streuungen aufweist (zufälliger Fehler), unterliegt die 2. Funktion eindeutig einem nichtlinearen mathematischen Modell (systematischer Fehler). Die oft gehörte Aussage, „eine Funktion mit  $R > 0,980$  ist ausreichend linear“, ist daher so nicht haltbar. Eine Unterscheidung, ob eine systematische oder zufällige Abweichung vorliegt, ist mit dem Korrelationskoeffizienten nicht zu beurteilen – eine entscheidende Schwäche bei der Verwendung der Indexzahl in der Analytik.

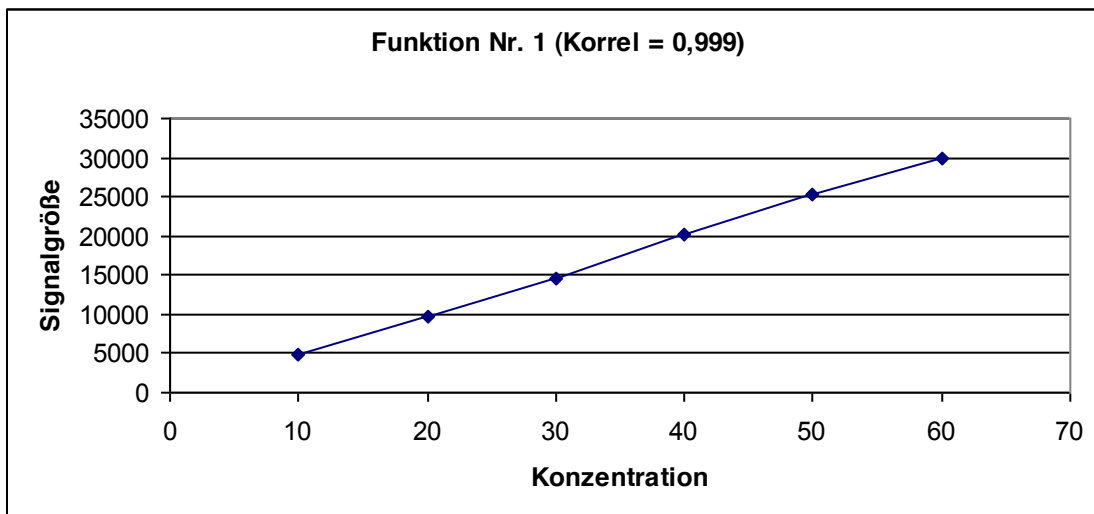


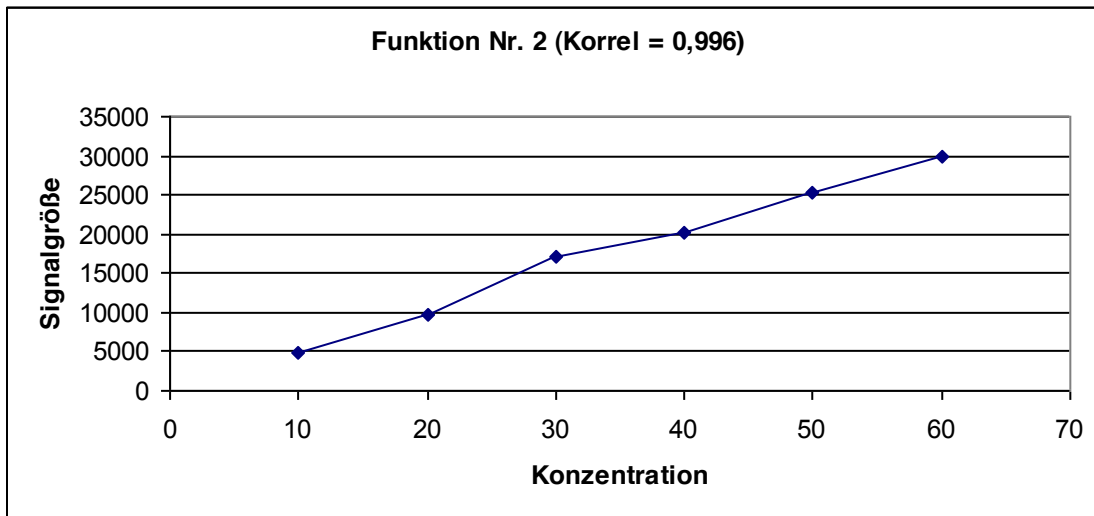
**Abb. 1.** Zwei Funktionen mit dem Korrelationskoeffizient  $R = 0,975$

Eine oft praktizierte Vorgehensweise ist die Bearbeitung des x/y-Datenkollektives nach der Methode der linearen Regression und dann mit der Methode der Polynomregression 2. Ordnung. („Quadratische Regression“). Man erhält u. a. den jeweiligen Korrelationskoeffizienten. Wenn der Korrelationskoeffizient nach der linearen Regression größer ist als der nach der Polynomregression, gibt man dem linearen Verfahren den Vorzug. Diese Vorgehensweise ist aus den bereits genannten Gründen ebenfalls nicht empfehlenswert.

Eine weitere Einschränkung zur Verwendung der Korrelationskoeffizienten bei einer Strategiebewertung besteht darin, dass der Korrelationskoeffizient die Empfindlichkeit  $b$  (Steigung) der analytischen Methode nicht oder nur indirekt berücksichtigt. Ein analytisches Verfahren mit hoher Empfindlichkeit weist naturgemäß eine höhere Streuung (zufälliger Fehler) auf als ein Verfahren mit geringer Empfindlichkeit, damit verschlechtert (verkleinert) sich der Korrelationskoeffizient, ob wohl das per se nicht nachteilig sein muss.

Manchmal wird der Korrelationskoeffizient auch dazu benutzt, um bei einer bereits akzeptierten linearen Strategie ein neu gewonnenes Datenkollektiv auf Ausreißer zu überprüfen (Motto: „Sobald der Korrelationskoeffizient  $R > 0,996$  ist, sind die neuen Kalibrierdaten o.k.“). Auch diese Vorgehensweise ist mit Vorsicht anzuwenden. In Abb. 2 unterscheiden sich beide Funktionen nur durch einen Ausreißer in der Funktion Nr. 2 (bei  $x = 30$ ). Der Korrelationskoeffizient verändert sich aber nur von  $R = 0,999$  zu  $R = 0,996$ , er ist „schwerfällig“.





**Abb. 2.** Zwei lineare Funktionen, mit und ohne Ausreißer

Aus einem hohen Wert des Korrelationskoeffizienten  $R$  kann man daher nur schwer auf Ausreißerfreiheit schließen.

Fazit: Der Korrelationskoeffizient  $R$  ist zwar ein unverzichtbarer und wichtiger Parameter zur Beurteilung einer Kalibrierung, aber er ist mit Vorsicht anzuwenden und es gilt, den Indexwert behutsam zu interpretieren! Mit geeigneten Methoden muss der Analytiker unterscheiden, ob systematische oder eher zufällige Einflüsse die Kalibrierung beeinflussen.

***Es müssen daher noch weitere - und bessere - Kriterien zur Beurteilung der Kalibriergüte zur Verfügung stehen. Davon soll ein weiterer Artikel in dieser Zeitschrift handeln.***

Beschäftigen Sie sich mit Kalibrierungen und Validierungen? Dann besuchen Sie das NOVIA Anwenderforum „Qualitätssicherung im analytischen Labor“ am 20. November 2014 mit dem Schwerpunktthema „Validierung in der Analytik“ oder buchen Sie den neuen NOVIA Fernlehrgang „Validierungsbeauftragte/r für analytische Methoden“ (Start: 03. Februar 2014)

Ihr Kontakt:

**NOVIA**

Chromatographie- und Messverfahren GmbH  
Industriepark Höchst - Gebäude B 845  
65926 Frankfurt am Main

Telefon: +49 69 3 05 - 43843

Telefax: +49 69 30 91 59

E-Mail: [info@novia.de](mailto:info@novia.de)

Internet: [www.provadis-novia.de](http://www.provadis-novia.de)