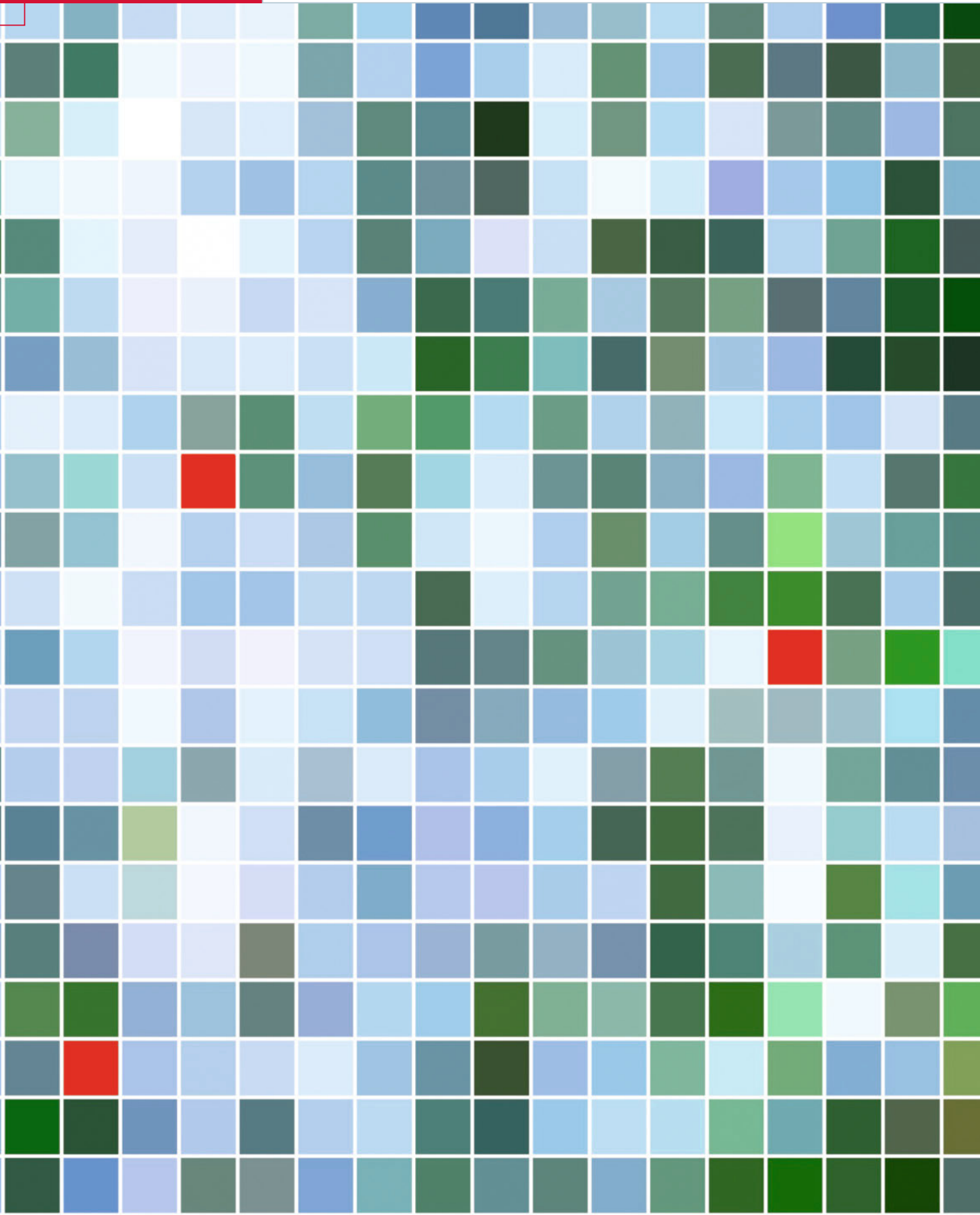


Verteilern.
elektronischen
Angeboten sowie
Internet- und
Intranet- und
Nicht zur
Verwendung in
QZ-Archiv
www.qm-
München
Verlag
© 2011 Carl Hanser Verlag, München



Mit einer zufälligen Stichprobe aus der Stichprobe zieht man sich gleichsam am eigenen Schopfe aus dem Datensumpf – und entwickelt die DoE-Methodik weiter.

FRÜHE ABSCHÄTZUNG DER PROZESSFÄHIGKEIT AUS EINEM DOE

Am eigenen Schopfe ...

Kombiniert mit modernen Simulationstechniken vermag die Methodik des Design of Experiments (DoE), auch die Robustheit der ermittelten optimalen Produktrezeptur oder Prozessfahrweise abzuschätzen. Eine neue methodische Variante erlaubt es jetzt sogar, Faktoreinstellfehler zu berücksichtigen. Ein diffiziler Fertigungsprozess der Schaeffler Gruppe profitierte davon und ermöglicht einen Methodenvergleich.

Andreas Orth, Eschborn, und
Mathias Probst, Herzogenaurach

Auf dem Weg von der Idee bis zur Marktreife hat sich das Design of Experiments (DoE) als wichtiger Systematisierungsschritt in der Produkt- und Prozessentwicklung etabliert. Auch als „statistische Versuchsplanung“ bezeichnet, umfasst die Methodik über die Planung hinaus auch Aus- und Bewertungsmethoden und dient dem Auffinden von Optima und der Absicherung erzielter Verbesserungen. Wurden etwa in einem Brainstorming Maßnahmen oder Faktoren entwickelt, die potenziell zur gesuchten Verbesserung führen könnten, so hilft DoE zu entscheiden, welche von ihnen zum gesuchten Optimum eines Produktionsprozesses führten, und die Entscheidung abzuschern. Dazu kann in einem DoE simultan ein fester Satz von Einflussparametern systematisch gegeneinander variiert werden und unter Konstanzhaltung von Störeinflüssen anschließend mit der multiplen Regressionsanalyse als gängiger Modellierungsmethode ausgewertet werden.

Bei der Bewertung von DoE-Ergebnissen lassen sich zufallsbasierte Simulationstechniken verwenden, die die enorme Rechenkapazität moderner Computer nutzen:

■ **Monte-Carlo-Methoden** werden zur Optimierung und Validierung insbesondere dann sehr effektiv eingesetzt, wenn rigorose, analytische Methoden zur Fehlerabschätzung scheitern. Der Trick besteht darin, Zufallszahlen aus einer bekannten Verteilung zu ziehen (typischerweise der Normalverteilung,

deren Lage und Breite aus der Versuchsplanauswertung geschätzt werden) zu ziehen.

■ **Die Bootstrap-Methode** basiert auf dem Prinzip: „Sampeln aus dem eigenen Sample“ [1]. Sie zieht also eine Probe aus der Probe und kommt damit der legendären Leistung des Barons von Münchhausen nahe, der sich rettete, indem er sich an seinem eigenen Haarschopf aus dem Sumpf zog [5].

Zufällige Fehler in den Faktoreinstellungen

Solche Simulationstechniken ermöglichen, was die klassische Methode nicht erlaubt: die Berücksichtigung von Faktoreinstellfehlern. Im Kontext der Prozess-

modellierung und -optimierung sind neben systematischen Fehlern diverser Quellen zufällige, unkontrollierbare Fehler und deren Folgen schwer quantifizierbar und schlecht vorab schätzbar. Betrachtet man einen Produktionsprozess beziehungsweise einen Teilprozess oder Prozessschritt als Black Box mit einstellbaren Eingangsfaktoren und messbaren Antworten, so sieht man sich (meist kleinen) zufälligen Fehlern in den Faktoreinstellungen und (häufig größeren) Versuchsfehlern bei den Antworten gegenüber. In Kombination führen diese Fehler zu Schätz- und Prognosefehlern bei der Optimumsbestimmung (Bild 2). Die zufallsbasierten Simulationstechniken erlauben es dann, zusammen mit Modell- und Optimumsbestimmung auch ▷

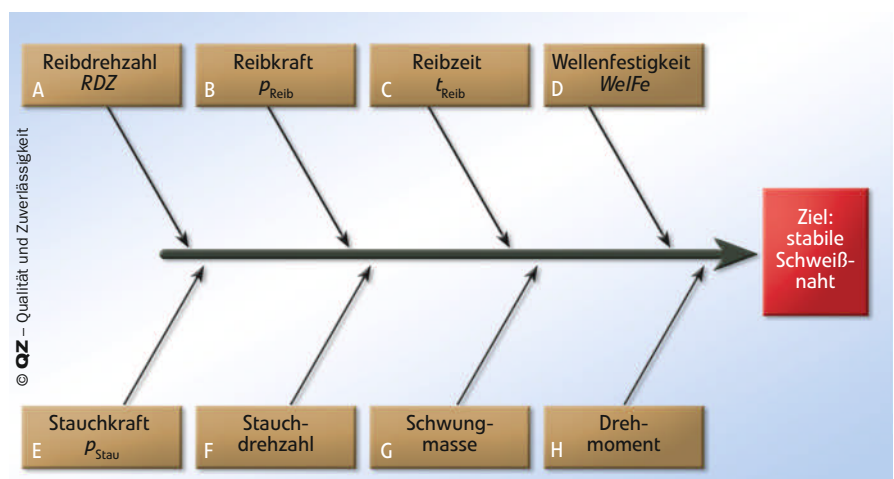


Bild 1. Ergebnis eines Brainstormings: Die Parameter A bis E sollen nun in einem DoE, einem 2(5-1)-Design, untersucht werden, die Parameter F bis H werden zunächst konstant gehalten.

Kongress

Moderne Weiterentwicklungen und Anwendungen des Design of Experiments sind auch Thema des 10. Kongresses „DoE – Get the best of it“ am 6.10.2011 in Frankfurt am Main. Programm und Anmelde-möglichkeit finden Sie unter:

www.ihk-kassel.de

Literatur

- 1 Efron, B; Tibshirani, R.J.: An Introduction to the Bootstrap. Chapman & Hall / CRC, London / New York 1993
- 2 Davison, A.C.; Hinkley, D.V.: Bootstrap Methods and their Applications. Cambridge University Press, 1997
- 3 Orth, A.; Oberacker, H.: Logarithmus oder nicht? Effizientere Entwicklungsversuche durch DoE. QZ 49 (2004) 7, S. 25–30
- 4 Soravia, S.; Orth, A.: Design of Experiments (DoE). Ullman's Encyclopedia of Industrial Chemistry. Wiley-VCH, Weinheim 2001
- 5 http://en.wikipedia.org/wiki/Baron_Munchhausen
Stand: 1.8.2011

Autoren

Prof. Dr. Andreas Orth, geb. 1958, hat sich als mathematischer Physiker auf Design of Experiments und statistische Datenanalyse spezialisiert. Er ist Professor an der FH Frankfurt/Main und Gesellschafter der Umesoft GmbH in Eschborn.
Dipl.-Ing. (FH) Mathias Probst, geb. 1962, ist Referent des „Kompetenzzentrums Statistik“ in der Abteilung „Wälzlager Grundlagen“ bei der Schaeffler Technologie GmbH & Co. KG, Herzogenaurach.

Kontakt

Andreas Orth
T 06173 608780
orth@umesoft.de

www.qm-infocenter.de

Diesen Beitrag finden Sie online unter der Dokumentennummer: **QZ110441**

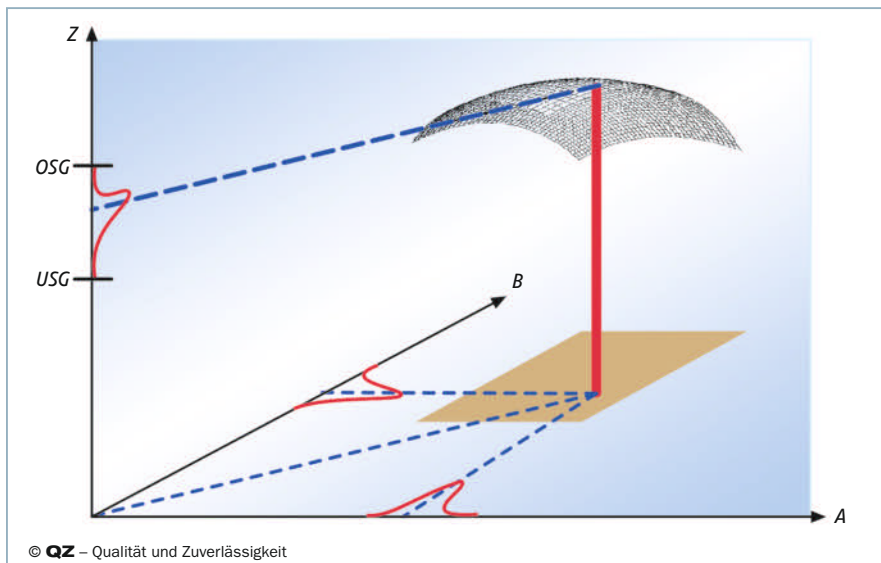


Bild 2. Exemplarische 3D-Modellfläche für die Abhängigkeit der Zielgröße Z von den Faktoren A und B; alle Modellprognosen – auch Optima – sind mit Fehlern behaftet, einer Summe aus Modellfehlern, Versuchsfehlern und Einstellfehlern.

gleich den Prozessfähigkeitsindex C_p oder C_{pk} abzuschätzen.

Diese Anforderung ergab sich jüngst auch bei einem Fertigungsprozess der Schaeffler Gruppe in Herzogenaurach. Mit rund 70 000 Mitarbeitern ist das Familienunternehmen weltweit führender

Anbieter von Wälzlagern und Linearprodukten und beliefert die Automobilindustrie mit Präzisionskomponenten und Systemen in Motor, Getriebe und Fahrwerk. Bei der Auslegung eines Reibschweißprozesses zur Verbindung zweier Wellen aus unterschiedlich legierten Stäh-

BAUANLEITUNG

Zweifacher Bootstrap zur Auswertung eines DoE

In folgenden Schritten lässt sich der zweifache Bootstrap zur Abschätzung von Prognoseintervallen für Messwertsimulationen aus DoE-Modellen (Modellen des Linear Modelling) durchführen.

Voraussetzung: Es wurde ein Versuchsplan mit n Einzelversuchen durchgeführt und die Versuchsergebnisse stehen zur Verfügung.

- Das Modell wird angepasst, Koeffizienten, Prognosen und normale sowie studentisierte (also nach „Weglassen“ berechnete) Residuen werden ermittelt.
- Ein „Würfel“ wird gebaut, auf dessen Flächen die studentisierten Residuen platziert werden. Für jeden Einzelversuch wird ein Residuum gewürfelt und auf den Prognosewert addiert. Aus diesem Satz gewürfelter Versuchswerte wird ein Satz sogenannter Bootstrap-Koeffizienten gewonnen. Dieser Ablauf wird $R = 1000$ Mal wiederholt.

- Nun wird eine neue Einstellkombination der Faktoren gewählt (zum Beispiel ein Optimum). Hieraus berechnet man mithilfe der ursprünglichen Koeffizienten die eine echte Prognose sowie mithilfe der 1000 Bootstrap-Koeffizientensätze 1000 Bootstrap-Prognosen.
- Parallel dazu würfelt man $K = 1000$ Mal aus den studentisierten Residuen, gleichzeitig aus den Faktorfehlerverteilungen, und addiert dann die gewürfelten Residuen auf die echten (also mit den originären Koeffizienten berechneten) Prognosewerte der gewürfelten Faktoreinstellungen.
- Man bildet schließlich die R mal K Differenzen und enthält somit eine Verteilung, die die Fehler der Messwertsimulation beschreibt. Aus ihr lassen sich Prognoseintervalle, Histogramme und Prozessfähigkeitsindizes ermitteln.

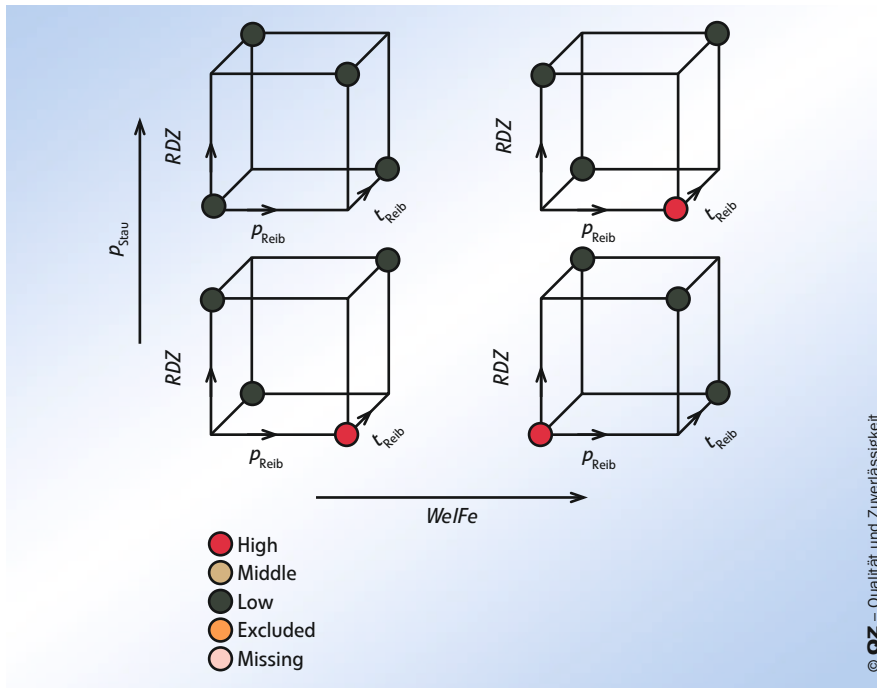


Bild 3. $2^{(5-1)}$ -Design für fünf Parameter ohne Zentrumspunkt, aber zur Abschätzung der Versuchsstreuung einmal komplett repliziert

len wurde DoE erfolgreich eingesetzt. Nach Voruntersuchungen zur Bestimmung des ungefähren Arbeitsbereichs

wurden die Wirkungen der prozessrelevanten Schweißparameter wie Reibdrehzahl, Reibkraft, Reibzeit, Stauchdrehzahl

und Stauchkraft sowie der Einfluss der Festigkeit des Wellenwerkstoffs untersucht. Zielgrößen waren neben der Schweißnahtfestigkeit diverse andere Spezifikationsgrößen. Verwendet wurde ein teilfaktorierter $2^{(5-1)}$ -Versuchsplan mit der Auflösung V zur Bestimmung der Haupteffekte und der Zweifachwechselwirkungen. Durch Wiederholversuche wurde die Prozessstreuung (Variance) ermittelt (Bild 3).

Abweichungen der Einstellfaktoren werden berücksichtigt

Neben der Bestimmung und Quantifizierung der Einflussfaktoren (Bild 4) war es von besonderem Interesse, den Prozess zu optimieren und die Auswirkung möglicher Abweichungen der Einstellfaktoren auf den Prozessfähigkeitsindex C_{pk} abzuschätzen. Hierzu wurden für die optimalen Einstellgrößen die zu erwartenden Schwankungen ermittelt beziehungsweise die maximal zulässigen Abweichungen spezifiziert. Unter der Annahme, dass die so ermittelten Schwankungen zufällig sind und somit einer Gleichverteilung \triangleright

PROGRAMMCODE

Zweifacher Bootstrap für Faktoreinstellfehler

Der nachstehende Algorithmus dient der Berechnung des zweifachen Bootstraps für das Prognoseintervall einer Modellvorhersage unter Berücksichtigung von Faktoreinstellfehlern (in der Variablen *blur*)

für den Plan in Tabelle 1. Er ist verfasst in der Programmierungssprache R. Dessen Interpreter kann kostenlos heruntergeladen werden unter:

<http://cran.r-project.org>

```
#Datenimport vom Clipboard (zuvor muss der kleine Versuchsplan in
#Tabelle 1 in die Zwischenablage kopiert werden - Spaltennamen bitte
#genau einhalten und Kleinbuchstaben verwenden)
boot_ex <- read.delim("clipboard",dec=","")
attach(boot_ex)      #... um einfacher an die Spalten zu kommen
head(boot_ex)       #müsste die ersten 6 Zeilen der Tabelle anzeigen
#Durchführung der Regressionsrechnung
X <- matrix(c(1,1,1,1,1,1,1,1,1,1,zt,temp,pre,zt*temp),ncol=5)
I <- solve(t(X)%*%X) #Informationsmatrix  $X^T X^{-1}$ 
g <- diag(I)         #Diagonale der I-Matrix
b <- I%*%t(X)%*%z    #Modellkoeffizienten b
H <- X%*%I%*%t(X)   #Hat-Matrix zur Berechnung der Prognosen
h <- diag(H)         #Diagonale der H-Matrix (leverages)
p <- H%*%z           #Prognosen p
e <- z - p           #Residuen e
r <- e/sqrt(1-h)     #Studentisierte Residuen r
#Vorbereitung des Bootstrap
r <- r-mean(r)       #Zentrieren der studentisierten Residuen
R <- 1000            #Probengröße für den äußeren Bootstrap
K <- 1000            #Probengröße für den inneren Bootstrap
delta <- array(0,dim=c(R,K)) #Matrix der Bootstrapergebnisse
x_plus_blur <- array(0,dim=c(K,5))
x_plus <- matrix(c(0,0,0)) #neuer Prognosepunkt
x_plus_ext <- c(1,x_plus,x_plus[1]*x_plus[2])
#erweiterter Prognosepunkt
blur <- matrix(c(0.5,0.5,0.5)) #Einstellungsunsicherheit
x_plus_blur[,1] <- array(1,dim=K) #Verteilung um Prognosepunkt
x_plus_blur[,2] <- rnorm(m,x_plus[1],blur[1])
x_plus_blur[,3] <- rnorm(m,x_plus[2],blur[2])
x_plus_blur[,4] <- rnorm(m,x_plus[3],blur[3])
x_plus_blur[,5] <- x_plus_blur[,2]*x_plus_blur[,3]
p_plus <- x_plus_ext%*%b #Prognose am Prognosepunkt
p_plus_blur <- x_plus_blur%*%b #Prognosen um den Prognosepunkt
#Bootstrap auf zwei Ebenen, er zieht zufällig aus den stud. Residuen
for(i in 1:R){ #äußere Ebene für Koeffizientenfehler
  z_star <- p + sample(r,replace=T) #Residuen würfeln
  b_star <- I%*%t(X)%*%z_star #Bootstrap-Koeffizienten
  p_star <- t(x_plus_ext)%*%b_star #Prognosen mit Bootstrap-Modell
  epsil <- sample(r,size=m,replace=T) #innere Ebene f. Prognosefehler
  delta[i,] <- array(p_star,dim=c(K,1))-(p_plus_blur+epsil)
} #Matrix der simulierten Fehler
#Quantile, Quartile, Histogramm
p_plus-sort(delta[,]) [c(m*975,m*500,m*25)]
p_plus-sort(delta[,]) [c(m*750,m*500,m*250)]
simulated_new_measurements<-p_plus-sort(delta[,])
hist(simulated_new_measurements)
#Aufräumen
detach(boot_ex)
```

unterliegen, wurde der Prozess mit seinen Einstellfehlern mittels Design Space Validation simuliert (Implementierung der Monte-Carlo-Methode mit 1 Million Zufallszahlen zur Simulation des Prozesses unter Berücksichtigung der Modell- und Einstellfehler) und die ermittelten Prognosen den Spezifikationsgrenzen gegenübergestellt (Bild 5). So ergab sich beispielsweise für die geforderte Schweißnahtfestigkeit ein Prozessfähigkeitsindex $C_{pk} > 3$, was auf einen hinreichend robusten Prozess hindeutet.

Der Reibschweißprozess konnte also mithilfe von Design of Experiments analysiert und optimiert und seine Prozessfähigkeit mittels der Monte-Carlo-Methode beurteilt werden.

Abschätzung von Konfidenz- und Prognosebereichen

Bei der Abschätzung der Fehler von Modellprognosen stellt sich zunächst die Frage, ob einen Modellschätzfehler oder Einzelprognosefehler interessieren. Der Schätzfehler ist die zu erwartende Standardabweichung der Modellschätzung an einem Prognosepunkt. Würde man einen Versuchsplan bis in alle Ewigkeit immer wieder wiederholen, so wäre die entsprechende Modellschätzung 100%ig präzise. Die entsprechenden Konfidenzintervalle würden also zu einem Punkt zusammenschrumpfen. Nicht zu einem Punkt schrumpft dabei der Einzelprognosefehler zusammen, denn jeder neue Versuchswert hat gleichsam sein eigenes Recht zu streuen – was auch durch wiederholtes Prüfen nicht zu ändern ist. Es gibt folglich zwei zu unterscheidende Fragen:

- Wie streuen die Mittelwerte (also Schätzer)?
- Wie streuen die Einzelprognosen?

Die **Versuchsplanung** hat zwar etwas komplexere Fehlerfortpflanzungsregeln als die Mittelwertbestimmung, dennoch kann man – unter der idealen Voraussetzung eines Modells ohne Schwächen (also ohne Bias) und unter der Annahme unabhängiger und ungefähr normal verteilter Versuchsfehler – Standardfehler und Konfidenzintervalle von Modellschätzer und Einzelprognose unmittelbar ausrechnen. Faktoreinstellfehler vermag die klassische Methode nicht mit zu berücksichtigen, zumal wenn diese nicht als normal verteilt, sondern – wie im Fall des Reibschweißens – als gleich verteilt anzu-

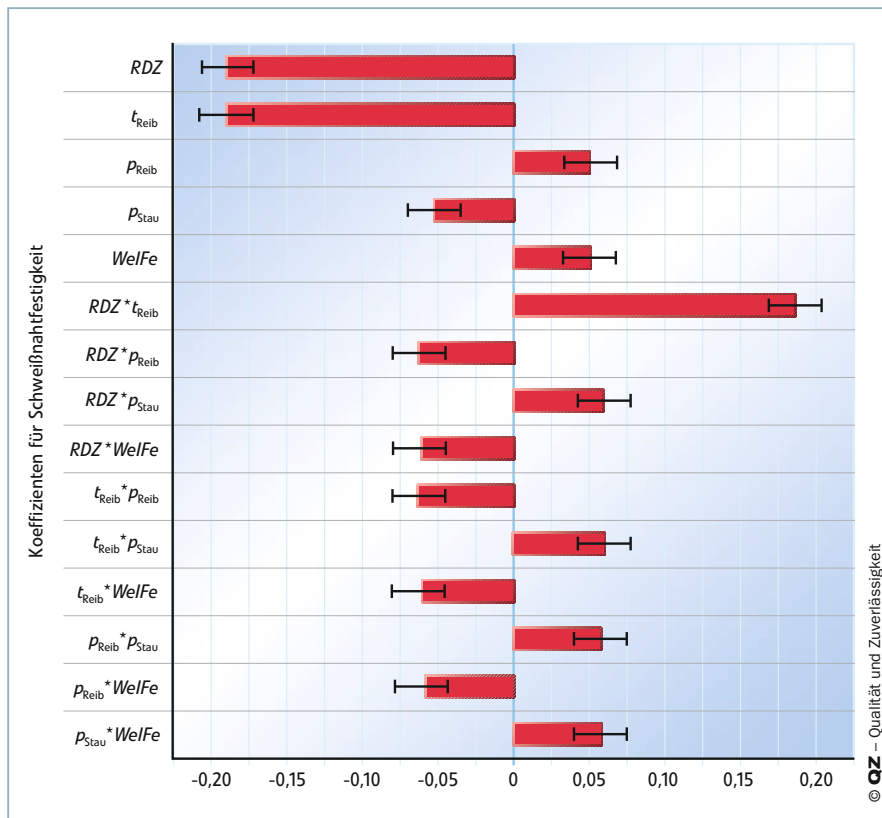


Bild 4. Quantifizierung der Einflussfaktoren als Koeffizientengrafik: Die 16 Restfreiheitsgrade bilden eine solide Basis für die Berechnung der Residualstandardabweichung und die spätere Monte-Carlo-Simulation.

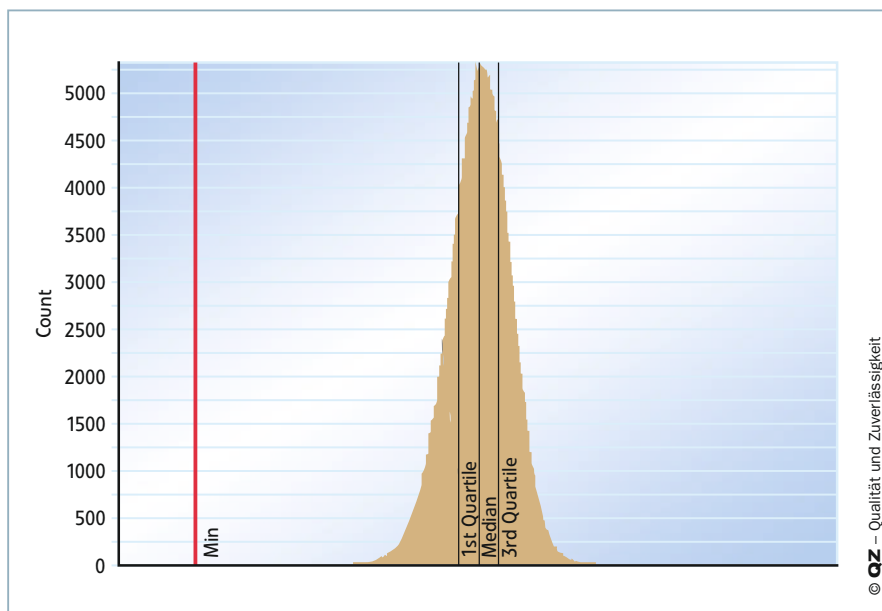


Bild 5. Monte-Carlo-Simulation der Prognosen für den Reibschweißprozess unter Berücksichtigung von Versuchsfehlern und Faktoreinstellfehlern (ohne Transformation, symmetrische Verteilung)

sehen sind. Abhilfe schaffen hier die zufallsbasierten Simulationen.

Die **Monte-Carlo-Methode** beruht darauf, gemäß einer vermuteten Fehlerverteilung Zufallszahlen zu ziehen oder zu

würfeln. Wie bei der klassischen Berechnung der Konfidenz- und Prognoseintervalle werden der Fehlerfortpflanzung die Annahmen der Modellvalidität und der unabhängigen und normal verteil-

zt	temp	pre	z
-1	-1	-1	68,31
1	-1	-1	86,96
-1	1	-1	94,30
1	1	-1	98,72
-1	-1	1	94,07
1	-1	1	97,41
-1	1	1	98,84
1	1	1	99,77
0	0	0	96,52
0	0	0	96,40

Tabelle 1. Der kleine Versuchsplan, der für die Bootstrap-Simulationen mit der Programmiersprache R verwendet werden soll (Infokasten „Programmcode“)

ten Versuchsfehler zugrunde gelegt. Von dieser Verteilung wird dann gesampelt, es werden also beispielsweise $R = 1000$ Zufallszahlen generiert. Auch Faktoreinstellfehler einer beliebigen Verteilung können so berücksichtigt werden – indem man parallel auch von ihr sampelt (Bild 5).

Die Bootstrap-Methode verzichtet auf die Fehlerfortpflanzungsmethode und somit teilweise auf die Annahmen der Modellvalidität und der unabhängigen und normal verteilten Versuchsfehler [1]. Gesampelt wird nicht von einer Verteilung, sondern direkt von den Residuen, also den Abweichungen von Modellwert und beob-

achtetem Wert an den Versuchspunkten. Für den Einsatz bei der Auswertung eines DoE muss der zweifache Bootstrap eingesetzt werden (Infokasten „Bauanleitung“). Er sampelt einmal für den Modellfehler und ein zweites Mal für den Prognosefehler (Infokasten „Programmcode“, Bild 6).

Von den drei Methoden ist die klassische Fortpflanzungsformel der Versuchsplanung die schnellste, allerdings kann sie keine Faktoreinstellfehler berücksichtigen. Die Monte-Carlo-Methode hingegen erlaubt es (gegebenenfalls nach unterschiedlichen Verteilungen) Faktoreinstellfehler mit zu simulieren. Sie ist schnell genug, und die notwendigen Iterationen können leicht auf dem PC durchgeführt werden. Die Bootstrap-Methode hängt am wenigsten von Modell- und Fehlerverteilungsannahmen ab, liefert also auch bei leicht asymmetrischen Verteilungen vernünftige Ergebnisse (Bild 6). Wegen der R -fachen Koeffizientenbestimmung ist die Methode langsamer als die Monte-Carlo-Methode (wobei dank der internen Parallelisierungsfunktionalitäten der verwendeten Software Werte bis $R = 1000$ und $K = 1000$ durchaus praktikabel sind).

Alle drei Methoden setzen voraus, dass der Versuchsplan für das gewählte Modell adäquat ausgesucht wurde, dass hinreichend viele Restfreiheitsgrade (also Redundanz in den Versuchen) vorliegen und dass bei der Modellbildung die Residuen-

analyse und die Modellverfeinerung durchgeführt wurden [4]. Diese Funktionalität wird heute von den meisten gängigen Softwareprodukten zu Modelling and Design of Experiments angeboten.

Frühschätzung kann strategisch relevant sein

Alle drei Methoden können in Kombination mit varianzstabilisierenden Transformationen der beteiligten Zielgrößen, etwa der Logarithmustransformation [3], kombiniert werden und auch im Fall sehr asymmetrischer Fehlerverteilungen zuverlässige Ergebnisse liefern.

Für die Berechnung der Defects per Million Opportunities (DPMO) und von C_p - und C_{pk} -Werten kann man im Prinzip alle drei Verfahren verwenden. Auf jeden Fall ist darauf zu achten, dass die so ermittelten Indexschätzungen lediglich als Richtwerte verwendet werden. Denn erstens reichen die Restfreiheitsgrade (also die Versuchsredundanz) zumeist nicht für eine zuverlässige Bestimmung aus, und zweitens erfolgt die Schätzung in der Regel nicht im Kontext der tatsächlichen Produktion, sondern im Entwicklungsumfeld. Dennoch ist eine Frühschätzung solcher Indizes sehr nützlich und kann sogar strategische Entscheidungen statistisch absichern und sie dadurch unterstützen. □

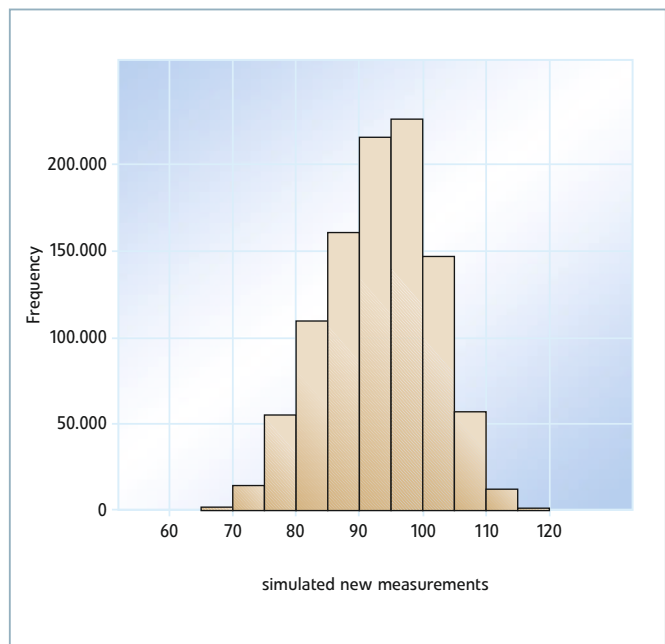
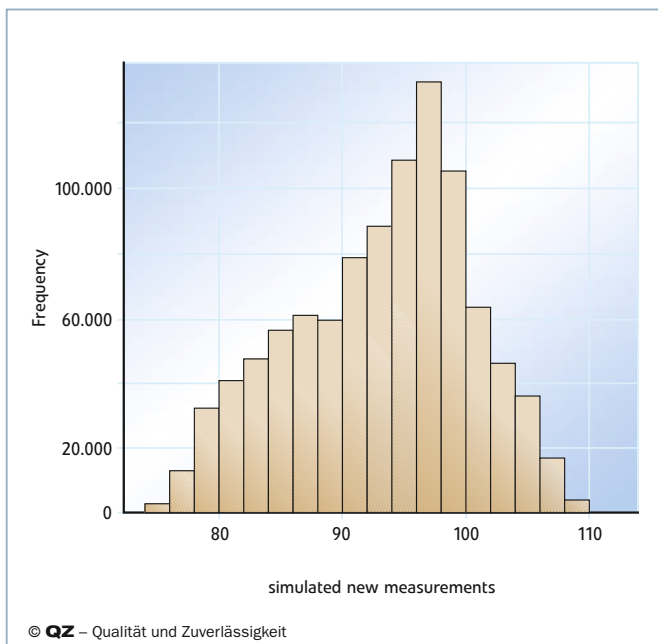


Bild 6. Ergebnis der Bootstrap-Simulation in Infokasten „Programmcode“ für den Versuchsplan aus Tabelle 1; links das Ergebnis nach der Methode von Davison und Hinkley [2], rechts das Ergebnis nach der neuen Methode in Infokasten „Programmcode“ unter Berücksichtigung der Faktorvariationen. Es wurden $R = 1000$ Iterationen für den Bootstrap und $K = M = 1000$ zufällige Prognosepunkte von der Normalverteilung gesampelt.