

Ein Ansatz zur Prozessverbesserung

Multiple lineare Regression

Dr.-Ing. Thomas Pfeilsticker, TEQ® Training & Consulting GmbH

Im vorliegenden Aufsatz wird das Anlegen und Auswerten einer Regressionsstudie anhand von Prozessdaten gezeigt. Gesucht wird dabei ein empirisches Modell $y=f(x_1, x_2, x_3...)$ für die Prozessdaten in Bezug ihrer Wirkung auf die Zielgröße.

Der Spritzgießprozess eines Thermoplasts produzierte Bauteile, die bei der späteren Montage aufgrund ihrer zu großen Schrumpfung zu Problemen führten. Zur Untersuchung einer möglichen Verringerung der prozentualen Schrumpfung wurden gezielt Daten von **Einspritztemperatur, Einspritzgeschwindigkeit und Halteback** erhoben. Wird die allgemeine Modellgleichung angewandt, ergibt sich: **Prozentuale Schrumpfung = f (Einspritztemperatur, Einspritzgeschwindigkeit, Halteback)**.

Welche der drei Einflussgrößen haben eine besonders große Wirkung auf die prozentuale Schrumpfung des Spritzgießprozesses? Und wie müssen die Einflussgrößen eingestellt sein, damit sich eine möglichst kleine prozentuale Schrumpfung der Spritzgussteile ergibt.

Folgende zuverlässige Prozessbeobachtungsdaten liegen vor:

	Einspritztemperatur	Einspritzgeschwindigkeit	Halteback	prozent. Schrumpfung
1	296,00	206,20	37,31	0,87
2	296,80	206,20	36,48	0,83
3	294,70	204,30	36,01	0,87
4	296,40	206,50	37,51	0,82
5	293,20	207,40	37,48	0,82
6	296,30	207,40	37,21	0,86
7	294,70	207,40	37,21	0,88
8	296,30	206,90	37,21	0,86
9	294,20	204,20	37,21	0,83
10	296,10	204,30	36,87	0,86
11	296,00	206,80	34,52	0,89
12	294,30	204,90	36,66	0,91
13	293,80	203,50	37,23	0,84
14	296,30	206,30	37,36	0,86
15	294,40	206,70	36,40	0,86

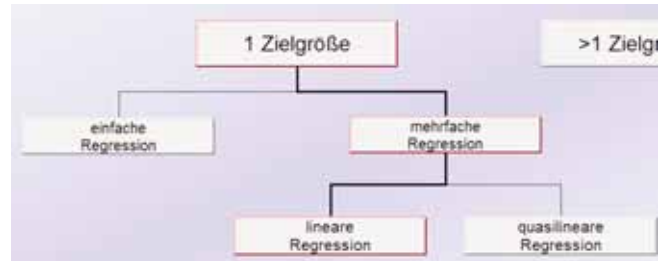
Prozessdaten

Der Mittelwert der prozentualen Schrumpfung beträgt 0,886%.

1 Wahl Regressionsmethode -mehrfache lineare Regression- und Modellstruktur

Die Software destra® unterstützt den Anwender beim Anlegen und Auswerten von Prozess-/Felddaten sowie bei den durch systematische Planung gewonnenen Versuchsdaten. destra® hilft, Fragen nach den stark wirkenden Faktoren (deren Betrag) und deren Wirkrichtung (deren Vorzeichen), nach Wechselwirkungen und optimalen Faktoreinstellungen zu klären.

Die erhobenen Prozessdaten werden mit einer mehrfachen linearen Regressionsanalyse untersucht:



Zur Wahl und zum Vergleich stehen die mehrfache lineare oder die mehrfache quasilineare Regression zur Verfügung.



Mit dem Navigations-Icon „Weiter“  gelangen wir zur Auswertung.

2 Auswertung einer mehrfachen linearen Regression

Korrelationskoeffizient	r	1	0,8814
Bestimmtheitsmaß	B	0,7768	77,68%
korrigiertes Bestimmtheitsmaß	B*	0,7158	71,58%
Restvarianz	s ²	0,00421	0,00421
Reststandardabweichung	s	0,0205	0,0205
$y = -4,56489 + 0,0239612 \cdot X1 + 0,00498306 \cdot X2 - 0,0173386 \cdot X3$			

Hier sind wichtige Kenngrößen zur Modellgüte des Regressionsmodells zusammengefasst.

Das **Bestimmtheitsmaß (B)** = 77,68% ist recht hoch. (B) gibt an, wie gut die Varianz der Zielgröße, der prozentualen Schrumpfung, durch die drei Einflussgrößen erklärt werden kann oder gleichbedeutend damit, wie genau ist die Streuung der Schrumpfungswerte alleine durch die Berücksichtigung der drei ausgesuchten Einflussgrößen vorherbestimmt.

Das **korrigierte Bestimmtheitsmaß (B*)** ist ein faires Vergleichsmaß zwischen alternativen Modellen, beispielsweise dem linearen und quadratischen, da die Anzahl der Regressionsparameter und die Anzahl der Messwerte berücksichtigt wird.

Die nicht erklärte Streuung der prozentualen Schrumpfung wird durch die **Restvarianz (s²)** bzw. die **Reststandardabweichung (s)** ausgedrückt.

3 Schätzung der Koeffizienten (Parameterschätzung)

Zur rechnerischen Ermittlung des Regressionsmodells wird die Summe der quadrierten Abweichungen (Residuen²) als minimal geschätzt. Es gibt kein Regressions-

modell, das eine kleinere **Summe der quadrierten Abweichungen** aufweist.

R = 77,652%		R² = 71,537%	
Merkmale	Merkmale	xi	bi
4	Schrumpfung	f(x1, x2)	
	Konst.	-8,585	+13,635 -2,444
1	Einspritztemperatur	X1	0,0240 0,0083 0,0396 0,00759 3,376**
2	Einspritzgeschwindigkeit	X2	0,0000 -0,0170 -0,0272 0,0101 0,483
3	Haltedruck	X3	-0,0172 -0,0316 -0,0028 0,00619 2,829*

$y = 6,56489 + 0,0239512 \cdot X1 + 0,00498306 \cdot X2 - 0,0173386 \cdot X3$

Nach den Bezeichnungen für die Ziel- und Einflussgrößen werden die **Regressionskoeffizienten (bi)** sowie deren **Vertrauensbereiche** (95%, Vertrauensniveau) **bi (...)** angegeben. Zur Bewertung der Regressionskoeffizienten werden die jeweiligen **Standardabweichungen s (bi)** und die, durch Division des Regressionskoeffizienten durch dessen Standardabweichung errechnete, **t-Testgröße** ausgegeben. Zur Bewertung der Signifikanz stehen eine Markierung der t-Testgröße mit hochgestellten Sternchen, eine Balkengrafik mit den drei üblichen durch rote Linien markierte Signifikanzniveaus und alternativ der **P-Wert** zur Verfügung.

Wird die $\alpha = 5\%$ Hürde überschritten, ist die Einflussgröße signifikant, der entsprechende Balken ist gelb gefärbt; wird auch noch die $\alpha = 1\%$ Hürde übertroffen, ist der Balken rot. Der P-Wert muss kleiner 5% (0,05) sein, er ist das Restrisiko, eine Signifikanz irrtümlicherweise zu erkennen. Einspritztemperatur und Haltedruck erweisen sich als signifikant, die Einspritzgeschwindigkeit aber nicht.

Die **VIF Werte** (Varianzinflationsindex) erhellen, ob eine gegenseitige Abhängigkeit der einzelnen Einflussgrößen vorliegt.

Regressionsmodelle werden oftmals abgebaut, das heißt, um nicht signifikante Einflussgrößen und deren Wechselwirkungen bereinigt, anschließend wird das Modell erneut berechnet und bewertet.

Einspritztemperatur und **Haltedruck** haben eine besonders **starke Wirkung** auf die prozentuale Schrumpfung. Dies wird im reduzierten Modell noch deutlicher.

R = 77,190%		R² = 73,389%	
Merkmale	Merkmale	xi	bi
4	Schrumpfung	f(x1, x2)	
	Konst.	-8,178	-9,755 -2,596
1	Einspritztemperatur	X1	0,0262 0,0145 0,0378 0,00533 4,902***
3	Haltedruck	X3	-0,0178 -0,0316 -0,0040 0,00632 2,818*

$y = 6,17574 + 0,0261521 \cdot X1 - 0,0178006 \cdot X3$

Eine Zunahme der Einspritztemperatur um 1°C bewirkt eine Vergrößerung der prozentualen Schrumpfung von ca. 0,026% und eine Vergrößerung des Haltedrucks um 1MPa eine Verringerung um ca. 0,018%.

4 Modellbeurteilung

4.1 Modellsignifikanz

Dieser F-Test prüft, ob die Einflussgrößen (Koeffizienten) sich gemeinsam auf die Zielgröße signifikant auswirken und nicht einzeln wie unter 3. beim t-Test der Koeffizienten. Wird die Nullhypothese abgelehnt (rot), so ist mindestens ein Einfluss signifikant.



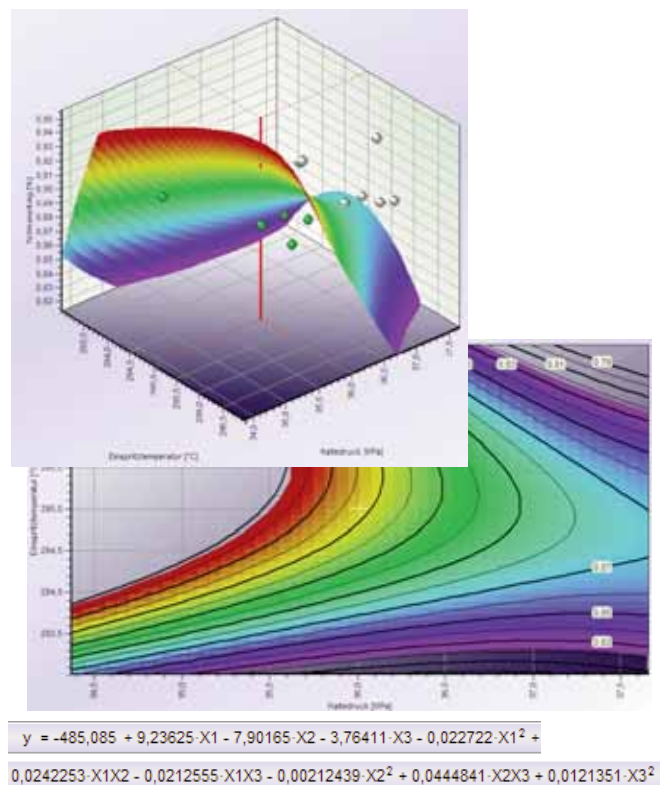
4.1 Modellanpassung - Linearitäts-F-Test (LoF)

Zur Modellbeurteilung werden neben dem Bestimmtheitsmaß und dem Signifikanz-t-Test der Koeffizienten noch ein varianzanalytischer Linearitäts-F-Test, auch lack of fit-Test (LoF) genannt, genutzt, um die Anpassung eines Modells zu testen. Der Test setzt aber mehrere Beobachtungen zu identischen Prozesseinstellungen voraus und ist deshalb vor allem für geplante Versuchsdaten wichtig.

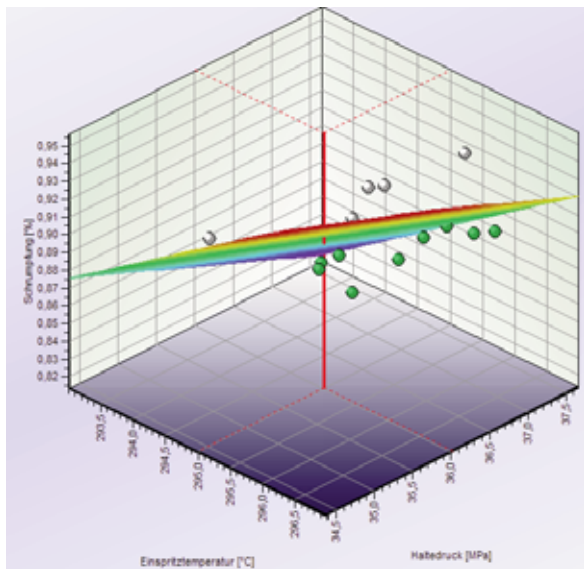
5 Übersicht Funktion 3D und Höhenlinien:

Die Grafiken zeigen den Einfluss zweier Einflussgrößen auf die Zielgröße unter konstanten Bedingungen der dritten Einflussgröße.

Die 3-dimensionale Darstellung der Variablen ist vor dem Abbau des Modells für den quasilinearen Ansatz (inklusive Wechselwirkungen) gezeigt; ebenso ist die Projektion der prozentualen Schrumpfung in die durch die Einflussgrößen aufgespannte Ebene abgebildet:



Nach dem Abbau des linearen Modells ergibt sich folgendes Bild:

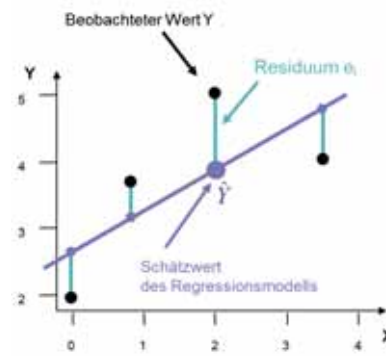


$$y = -6,17574 + 0,0261521 \cdot X1 - 0,0178006 \cdot X3$$

6 Einflussgrößenübersicht:

Die Einflussgrößen-Übersicht zeigt für jede gewählte Einflussgröße, wie diese wirkt und mit welchem Prognosewert bei vorgegebener Einstellung der Einflussgröße zu rechnen ist. 0,85% prozentuale Schrumpfung ist bei den durch die roten Linien markierten Einstellwerten vorhergesagt.

7 Analyse der Residuen

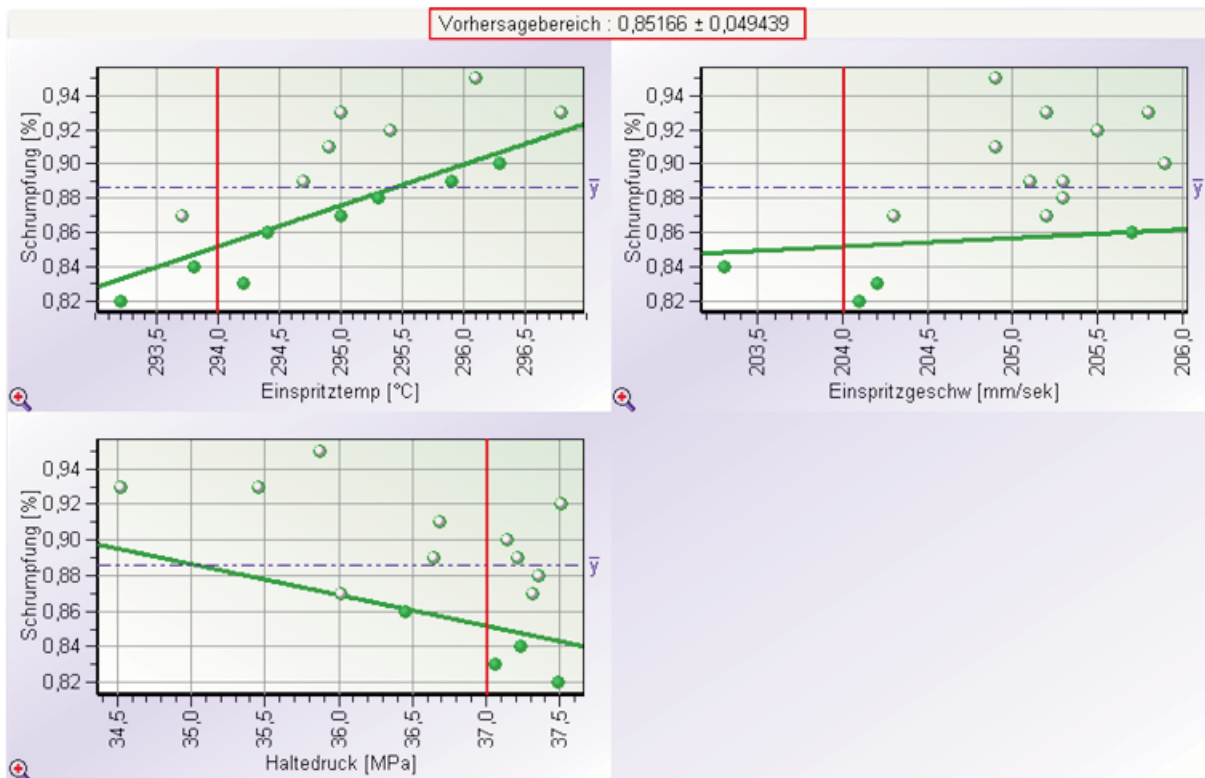


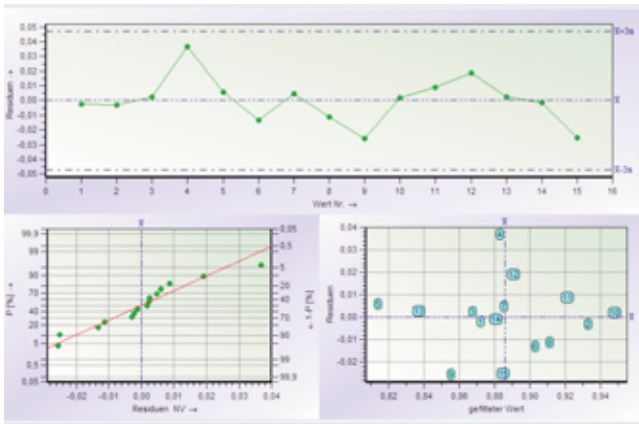
Residuen (lat. Rest, Singular Residuum) sind die Differenzen zwischen den beobachteten (y_i) und den durch das Modell geschätzten Werten (\hat{y}) der Zielgröße. Folglich sind die Residuen die durch das Modell nicht erklärten Abweichungen.

Ihre grafische Analyse gibt Aufschluss darüber, ob die Modellannahmen erfüllt sind.

In der oberen Grafik werden die Residuen über der Wertenummer dargestellt. Im Idealfall sollten diese zufällig verlaufen, also keine Struktur aufweisen.

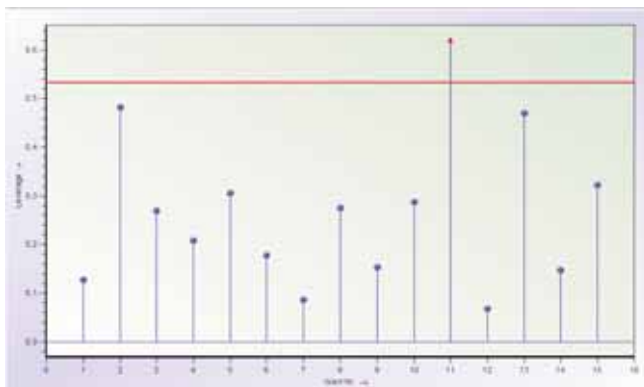
Links unten werden die Residuen im Wahrscheinlichkeitsnetz auf Normalverteilung geprüft. Die Annahme der Normalverteilung ist Voraussetzung für einige Tests im Rahmen der Regressionsanalyse.





Unten rechts werden die geschätzten Werte (gefittete Werte) den Residuen gegenübergestellt. Die Werte sollten in dieser Grafik zumindest näherungsweise von links nach rechts zufällig um die Nulllinie auftreten.

8 Leverage, Cook:



Diese Diagnostika dienen der Prüfung, ob einzelne Werte (Datensätze) einen übergebührlichen Einfluss auf das gesamte Regressionsmodell ausüben.

Der Leverage (Hebel), dient der Überprüfung, ob einzelne Wertesätze der Einflussgrößen als Ausreißer in x-Richtung interpretiert werden können.

Die Cook-Distanz erkennt einen einflussreichen Wert, der sich ebenfalls stark auf das gesamte Modell (seinen eigenen Schätzwert im Modell) auswirkt.

9 VIF und Red%

Der Varianzinflationsindex (VIF) zeigt die gegenseitige Abhängigkeit der einzelnen Einflussgrößen an, der VIF sollte < 10 sein. Ohne gegenseitige Beeinflussung ist der VIF = 1. Wird der VIF größer, werden die Schätzungen der Koeffizienten ungenauer, dies kann sich auf den Signifikanztest der Koeffizienten auswirken.

Red% ist ein Indikator für die prozentuale Verringerung des Bestimmtheitsmaßes, falls eine bestimmte Einflussgröße aus dem Modell entfernt (reduziert) wird.

Red% ist ein Streuungs-Komplement zu der Koeffizienten-t-Statistik, die den Erwartungswert der Zielgröße, d. h. deren Steigerungsmöglichkeit untersucht.

Red% dagegen untersucht die Wirkung einzelner Einflussgrößen auf die Streuung der Zielgrößen.

23,13% bedeuten, dass das Bestimmtheitsmaß ohne die Einspritztemperatur nur noch 77,68% - 23,13% = 54,55% ist.

	VIF	Red%
	---	---
Einspritztemperatur	1,752	23,13
Einspritzgeschwindigkeit	1,774	0,492
Haltdruck	1,079	14,03

Die Einspritztemperatur ist auch bei der Streuungsverringerung wichtiger als der Haltdruck.

10 Optimierung

Vor der Prozessoptimierung betrug der Mittelwert der prozentualen Schrumpfung 0,886%.

Nach der Ermittlung der optimalen Parametereinstellungen sagen diese eine Verringerung der prozentualen Schrumpfung mit 0,825%. Im Kontrollversuch wurde die Prozessverbesserung nachgewiesen.

Suche Optima

Maximieren

Minimieren

Startpunkte

Anzahl zufälliger Startpunkte: 5

Auflösung der Gitterpunkte: 2

Index	Index	1
Faktoren	Einspritztemperatur	293,2
	Haltdruck	37,51
Schrumpfung	Vorhersage	0,82437
	Standardabw.	0,022717

messbar voraus

WANZEL

Rundum präzise...

- △ Bohrungsmessung
- △ Form- und Konturmessung
- △ Rauheitsmessung
- △ mit qs-STAT Anbindung



www.wanzel.com

△ Messen △ Spannen △ Kalibrieren △ Software △ Seminare

Wanzel Projektmanagement Ges.m.b.H. • 1220 Wien • Wagramer Str.173/D
 +43 1 259 3616 • +43 1 259 3617 • wanzel@wanzel.com