

Statistische Prozeßmodellierung in der Holzwerkstoffindustrie

Dipl.-Holzwirt Kai Kruse

Universität Hamburg

Ord. für Holztechnologie

Abstrakt

Im Bereich der Holzwerkstoffproduktion werden in zunehmendem Maße Systeme zur Prozeßdatenerfassung und Prozeßleitung eingesetzt. Hierdurch wird eine wesentliche Erleichterung bei der Steuerung des Prozesses vor allem bei Produktionsumstellungen erreicht. Des weiteren können die erfaßten Prozeßdaten zur Entwicklung und Fortschreibung von Prozeßmodellen verwendet werden. Zur Bestimmung funktionaler Zusammenhänge in industriellen Prozessen wie der Produktion von Holzwerkstoffen kann dann das Verfahren der experimentell-statistischen Prozeßmodellierung eingesetzt werden. Verwendungsmöglichkeiten von Prozeßmodellen im Bereich der Holzwerkstoffproduktion sind die Vorhersage der Produkteigenschaften, die Prozeßsimulation sowie verfahrenstechnische und kostenbezogene Optimierungen des Herstellungsprozesses. Ziel der hier vorgestellten Arbeit war die Bildung eines Modells für den Herstellungsprozeß von Spanplatten durch Anwendung der Regressionsanalyse mit dem SAS-Programmpaket. Zur Abschätzung der Vorhersagegüte der gebildeten Prozeßmodelle wurde ein Verfahren entwickelt, das die Überprüfung von Modellen an einer neuen Stichprobe simuliert. Es wurde festgestellt, daß eine Modellbildung unter alleiniger Berücksichtigung statistischer Kennzahlen zu einer zu optimistischen Einschätzung der Modelle führen kann. Aus diesem Grunde ist es bei der Modellbildung erforderlich, vor Anwendung eines statistischen Variablenauswahlverfahrens eine Reduzierung der Variablenzahl anhand technologischer und anlagenspezifischer Überlegungen vorzunehmen. Es konnten Modelle entwickelt werden, mit denen sich wichtige Produkteigenschaften (Qualitätsmerkmale) vorhersagen lassen. Dabei konnten aus der großen Fülle vorliegender Prozeßparameter die bedeutendsten identifiziert und in die Prozeßmodellierung integriert werden.

Produktion von Holzwerkstoffen

Holzwerkstoff ist ein Sammelbegriff für verschiedene Produkte, die durch Zerlegen des Holzes und anschließendes Zusammenfügen entstehen. Plattenförmige Produkte wie Span- und Faserplatten stellen den größten Anteil der Holzwerkstoffproduktion dar. 1995 wurden in Westeuropa allein annähernd 29 Mio m³ Spanplatten erzeugt (Pena-Möller 1996). Hauptverwendungsbereiche von Holzwerkstoffen sind der Möbelbau und das Bauwesen. Für die Produktion von Holzwerkstoffen werden als Rohstoffe Holz und Leim/Harze eingesetzt. Beim Holz entfällt ein Anteil von derzeit ca. 30% auf Stammholzsortimente. Die anderen 70% setzen sich aus Industrierestholz (Reststoffe der Holzverarbeitenden Industrie wie Paletten, Brettabschnitte, Furnierreste) und Hackschnitzeln zusammen. Als Leime/Harze werden zu etwa 90% Polykondensationsharze und zu etwa 10% MDI-Systeme (Methyldiisocyanat) eingesetzt. Das Fertigprodukt, die Holzwerkstoffplatte selbst, besteht aus ca. 85% Holz, ca. 8% Leim/Harz sowie ca. 7% Wasser und sonstigen Zuschlägen.

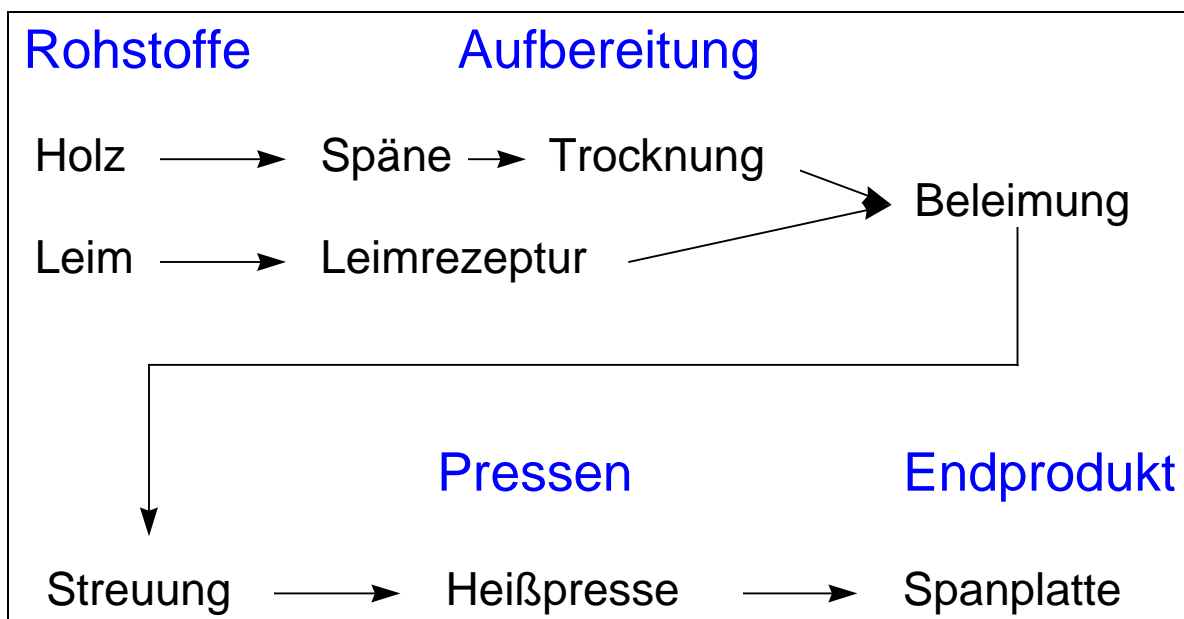


Abb. 1: Flow-chart der Herstellung von Spanplatten

Der Herstellungsprozeß von Spanplatten gliedert sich in verschiedene Teilprozesse der Rohstofflagerung, der Rohstoffaufbereitung, des Pressens und der Weiterbearbeitung zum Endprodukt.

Nach der Lagerung wird das Holz zunächst mittels verschiedener Hacker und Zerspaner zu Spänen verarbeitet. Den in der Regel noch feuchten Spänen werden in

großen Trocknern das Wasser entzogen. Die nunmehr in der Feuchte und den Abmessungen homogenisierten Späne werden in Beleimungsmischern mit dem mit Zusätzen aufbereiteten Leim beaufschlagt. Das Span-Leimgemisch wird in Bunkern zwischengelagert und anschließend auf ein Formband gestreut. Der gestreute „Spankuchen“ wird in einer Heißpresse unter Hitze und Druck zur fertigen Spanplatte verpreßt, die wiederum definierte Qualitätsmerkmale aufweisen muß.

Der Anlagenbetreiber befindet sich in der Situation, daß er zum einen ein ganzes Rohstoffmix mit unterschiedlichen Eigenschaften verarbeitet. Des weiteren ist der Gesamtprozeß in vielfältige Teilprozesse untergliedert die alle separat geregelt werden. Darüber hinaus werden an die fertigen Platten diverse Qualitätsanforderungen gestellt. Neben den normativen Forderungen nach DIN und EN, die vorrangig Eigenschaften der Mechanik und des Verhaltens gegenüber Wasser/feuchter Luft betreffen, werden von der weiterverarbeitenden Industrie darüber hinausgehende Anforderungen hauptsächlich hinsichtlich der Oberflächenqualität gestellt (Frühwald et al. 1995). Bislang erfolgt die Prozeßsteuerung dergestalt, daß in größeren Zeitabständen (oft mehrere Stunden) Proben aus dem Prozeß entnommen werden und an diesen die Produkteigenschaften ermittelt werden. Zwischen den Probennahmen versucht der Anlagenbetreiber die Prozeßparameter konstant zu halten, ohne aber konkrete Informationen über die Produktqualität zu haben.

Unter den Restriktionen dieser Gegebenheiten muß der Anlagenbetreiber Strategien entwickeln die dennoch eine sichere und ökonomische Produktion erlauben. Dabei ist für eine stabile und zuverlässige Produktion die Kenntnis der vielschichtigen Wechselwirkungen zwischen den Eigenschaften der Rohstoffe, den Prozeßparametern und der Platteneigenschaften notwendig. Diese Wechselwirkungen sind immer anlagenspezifisch und dem Anlagenbetreiber in der Regel nicht vollständig bekannt.

In diesem Zusammenhang bieten statistische Prozeßmodelle unter anderem Möglichkeiten einer stärkeren Transparenz des Produktionsprozesses und der (Echtzeit-) Vorhersage von Produkteigenschaften. Nach Entwicklung robuster Prozeßmodelle eröffnen sich Möglichkeiten einer kostenminimalen Prozeßoptimierung und der Simulation (Lobenhoffer et al. 1995, Bernardy 1995).

Statistische Prozeßmodellierung in der Holzwerkstoffindustrie

Für die statistische Prozeßmodellierung des Herstellungsprozesses von Spanplatten hat sich die multiple, lineare Regressionsanalyse als geeignet erwiesen (Lobenhoffer 1990). Die Prozeßmodelle werden dabei in bekannter Form dargestellt. Allgemein lautet die Vorhersagegleichung der linearen Regression:

$$\hat{y} = \hat{b}_0 + \hat{b}_1 x_1 + \hat{b}_2 x_2 + \dots + \hat{b}_k x_k \quad (1)$$

Die Gleichung besagt, daß der Vorhersagewert \hat{y} für die Zielgröße Y aus den Werten für die Einflußgrößen $X_1 = x_1$ bis $X_k = x_k$ bestimmt werden kann. Die Regressionskoeffizienten \hat{b}_1 bis \hat{b}_k sowie das Absolutglied \hat{b}_0 sind zunächst unbekannt und werden aus den Datensätzen (Beobachtungen) geschätzt.

Die einzelnen Datensätze lassen sich durch folgende Modellgleichung darstellen:

$$y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_k x_{ki} + u_i \quad (2)$$

Der Index $i = 1 \dots N$ bezeichnet dabei die jeweilige Datensatznummer. y_i und x_{1i} bis x_{ki} stellen die Meßwerte des i -ten Datensatzes dar. u_i ist der Stör-Term bei der i -ten Beobachtung. Der Stör-Term ist auf unkontrollierbare Einflüsse wie natürliche Schwankungen, Meßungenaugkeiten, Meßfehler etc. zurückzuführen.

Als Zielgrößen Y werden die nach DIN/EN festgelegten Produkteigenschaften (Mechanik (Festigkeiten), Verhalten gegenüber Feuchtigkeit) betrachtet. Als Einflußgrößen X gelten die physikalischen (wie Dimensionen, Feuchte), chemischen (wie pH-Wert, Pufferkapazität) und biologischen (wie Holzarten) Eigenschaften der Rohstoffe. Auch die Prozeßparameter wie Stoffmengen, Temperatur, Druck, Verweilzeit stellen Einflußgrößen dar. Im vorliegenden Fall sind die Zielgrößen und Einflußgrößen weiter zu differenzieren. Bei den Zielgrößen handelt es sich zum einen um die Zielgrößen Y_z , die nicht on-line meßbaren sondern im Labor festzustellenden Qualitätsmerkmale der Platten. Zum anderen existieren auch on-line meßbare Platteneigenschaften Y_a . Auch die Einflußgrößen sind zu differenzieren; in nicht einstellbare Einflußgrößen X_s und einstellbare Einflußgrößen X_c .

Die Voraussetzungen für die Anwendbarkeit der linearen Regressionsanalyse sind für den Prozeß der Holzwerkstoffherstellung nach Lobenhoffer (1990) erfüllt. Die unabhängige Verteilung der nicht einstellbaren Einflußgrößen x_{sj} und der Stör-Terme u_i kann als gesichert angesehen werden. Ebenso gilt für den Erwartungswert von u_i , daß $E(u_i) = 0$ und für die Varianz $\sigma_i^2 = const$. Auch liegen keine Autokorrelationen

zwischen den Werten der Zielgrößen vor. Die Vermeidung von Multikollinearitäten kann durch eine geeignete Vorauswahl der Einflußgrößen weitgehend ausgeschlossen werden. Dabei werden u.a. offensichtliche lineare Abhängigkeiten berücksichtigt. Liegen darüber hinaus nach der Modellbildung statistisch gesicherte multiple Bestimmtheitsmaße vor, so kann dies nach Jahn et al. (1972) als hinreichendes Kriterium angesehen werden, daß keine Multikollinearitäten vorliegen. Auch die Voraussetzung, daß annähernd lineare Zusammenhänge zwischen den Einflußgrößen und der Zielgröße bestehen, kann mit guter Näherung akzeptiert werden. Zunächst ist von vielen Einflußgrößen ein linearer Zusammenhang mit den Zielgrößen bekannt. Darüber hinaus schwanken zahlreiche Variablen bei Betriebsversuchen in einem vergleichsweise engen Rahmen, weswegen sich eventuelle Krümmungen kaum auswirken.

Ergebnisse praktischer Untersuchungen

Die im folgenden dargestellten Ergebnisse basieren auf Untersuchungen die über mehrere Wochen in einem Spanplattenwerk (Jahreskapazität 300.000 m³) durchgeführt wurden (León-Méndez und Thömen 1996). Untersucht wurde die Produktion eines Plattentypes (Dicke 16 mm, KUF-Verleimung). Erfasst wurden 132 Variable. 90 von diesen waren on-line meßbare Prozeßparameter. Zur umfassenden Charakterisierung der Rohstoffe wurden 32 weitere Parameter zusätzlich manuell aufgenommen. 8 im Labor festgestellte Produkteigenschaften wurden als Zielgrößen betrachtet. Insgesamt wurden 44 Datensätze (N) erhoben.

Aufgrund dieser Datenkonstellation ist es erforderlich die Anzahl der in den Modellansätzen zu berücksichtigenden Einflußgrößen (k) zu reduzieren, da mit einer Zunahme des Verhältnisses N/k die statistische Sicherheit des Modells nach Überschreitung eines Optimums abnimmt. Auch in der Literatur wird wiederholt auf die Gefahr hingewiesen, Modelle die bei einer solchen Datenkonstellation gebildet wurden zu optimistisch zu interpretieren (Farnum und Stanton 1989). Für das Verhältnis von Stichprobenumfang zur Anzahl der Einflußgrößen gibt Mager (1982) eine Faustformel von $N = 4k$ an, die hier als grobe Orientierung verstanden wird.

Die Reduzierung von Einflußgrößen kann sowohl nach statistischen¹ als auch nach technologischen Gesichtspunkten vorgenommen werden. Hierbei werden solche Einflußgrößen entfernt, die entweder bekanntermaßen unbedeutend sind, ausschließlich der Regelung eines Teilprozesses dienen, im Beobachtungsraum

¹ Statistische Auswertungen wurden mit dem Programmpaket SAS/STAT (Version 6.10) durchgeführt. Die verwendeten Prozeduren sind nach in folgenden Textstellen kursiv angegeben.

keine Variation aufweisen oder deren Informationsgehalt vollständig in anderen Einflußgrößen enthalten ist. Des weiteren können Einflußgrößen anhand partieller Korrelationskoeffizienten (*SAS proc korr*) und technologischer Kriterien in Gruppen zusammengefaßt werden. Unter Verwendung der beschriebenen Reduzierungsmöglichkeiten verblieben von den ehemals 124 Einflußgrößen 37 Prozeßvariablen für die Modellbildung.

Für die Bildung des Prozeßmodells für den Herstellungsprozeß von Spanplatten wurden die 37 Prozeßvariablen nochmals anhand unterschiedlicher Variablenauswahlverfahren reduziert:

1. Auswahl zwecks Maximierung des multiplen Bestimmtheitsmaßes R^2 (*SAS proc reg maxR*)
2. Auswahl anhand der F-Werte für die einzelnen Einflußgrößen (*SAS proc reg stepwise*)
3. Auswahl anhand technologischer Aspekte

Bei dem Variablenauswahlverfahren 1 wurde die Anzahl (5, 8, 11) der zu berücksichtigenden Variablen vorgegeben. Bei dem Variablenauswahlverfahren 2 wurde einerseits ein Abbruchkriterium vorgegeben ($p \leq 0.15$) oder andererseits die Anzahl (5, 8, 11) der zu berücksichtigenden Variablen in Abhängigkeit ihrer p-Werte festgelegt.

Bei der Variablenauswahl nach technologischen Gesichtspunkten wurden Variablen in das Modell aufgenommen, die nach wissenschaftlichem Kenntnisstand (Lobenhoffer 1990, Schweitzer 1992, Roll 1994) sowie nach dem Erfahrungswissen der Anlagenbetreiber für die Ausbildung der Produkteigenschaften besonders bedeutend sind.

Um zu einer Einschätzung der tatsächlichen Güte eines Modells zu gelangen, ist üblicherweise eine neue Stichprobe erforderlich. Aufgrund der in diesem Fall besonders zeitintensiven Datenerhebung, wurde hier eine abgewandelte Form des *cross-validation* Verfahrens eingesetzt. Üblicherweise wird beim *cross-validation* Verfahren der i -te Datensatz bei der Schätzung der Parameter nicht berücksichtigt. Dann kann die Zielgröße \hat{y}_i des ausgesonderten Datensatzes durch die berechnete Vorhersagegleichung geschätzt und mit dem im Labor gemessenen Wert der Produkteigenschaft y_i verglichen werden (Aitkin et al. 1989). Die Auswahl der Einflußgrößen wurde dabei allerdings auf Basis aller Datensätze vorgenommen. Der Wert \hat{y}_i wird also durch ein Modell geschätzt, in das die Informationen des unberücksichtigten Datensatzes bereits eingeflossen sind. Dem Anlagenbetreiber stehen aber die Informationen des vorherzusagenden Datensatzes nicht im vorwiegenden zur Verfügung. Er muß die Modellbildung ohne Berücksichtigung des zu schätzenden

Datensatzes vornehmen. Im vorliegenden Fall wurde daher bereits bei der Auswahl der Einflußgrößen der nicht zu berücksichtigende Datensatz ignoriert. Somit erfolgte sowohl die Auswahl der Einflußgrößen als auch die Bestimmung der Regressionskoeffizienten n -mal. Als Kriterium der Vorhersagegüte der Produkteigenschaften wurde das Abweichungsmaß a verwendet (Kruse et al. 1997):

$$a = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-1}} \quad (3)$$

In Abbildung 1 ist das Ergebnis der Modellbildung der Querzugfestigkeit (einer für die Gütecharakterisierung der Verleimung besonders bedeutende Größe) nach dem Verfahren der Maximierung des Bestimmtheitsmaßes R^2 dargestellt. Es zeigt, daß sich das mathematische Modell mit einem Bestimmtheitsmaß von $R^2=0.87$ und einem Abweichungsmaß von $a = 0.026 \text{ N/mm}^2$ gut an die beobachteten Meßwerte mit einer Standardabweichung von 0.071 N/mm^2 anpaßt.

In Abbildung 2 wurden auf Basis aller Datensätze die Einflußgrößen ausgewählt. Nur die Regressionsparameter wurden $n-1$ neu bestimmt. Die Bestimmtheitsmaße der einzelnen Regressionsansätze liegen zwischen 0.85 und 0.88. Das Abweichungsmaß hat einen Wert von $a = 0.037$.

In Abbildung 3 wurde die abgewandelte Form des cross-validation Verfahrens angewandt und auch die Einflußgrößen n -mal ausgewählt. Die Bestimmtheitsmaße liegen wiederum sehr hoch zwischen 0.81 und 0.88. Das Abweichungsmaß liegt allerdings nur noch geringfügig unter der Standardabweichung s der Beobachtung. Die Vorhersage der Produkteigenschaft war mit diesen Modellen somit im Mittel nicht besser als die mit den bisherigen Prozeßsteuerungsmethoden erreichte Schwankung.

Die Modell deren Bildung anhand technologischer Aspekte erfolgte zeigten des öfteren keine befriedigenden Ergebnisse. Auch hier lagen die Abweichungsmaße häufig nur geringfügig unter der Standardabweichung der Beobachtung.

Zur Interpretation von Modellgleichungen im Hinblick auf die funktionalen Zusammenhänge zwischen Prozeßvariablen und Produkteigenschaften ist es sinnvoll sich auf die Betrachtung eines zuverlässigen Modells zu beschränken. Deshalb wurde in einem weiteren Schritt **ein** Prozeßmodell für die Zielgröße Querzugfestigkeit entwickelt, bei dem die Auswahl der Einflußgrößen durch eine Kombination von statistischen und technologischen Aspekten erfolgte. Es wurde von einem Startmodell mit 18 Einflußgrößen ausgegangen, das die nach technologischen

Gesichtspunkten ausgewählten sowie die durch die statistischen Auswahlverfahren häufig ermittelten Einflußgrößen enthielt. Durch schrittweises Eliminieren (*SAS proc reg stepwise*) der nach dem F-Test am wenigsten gesicherten Variable, ergab sich ein Prozeßmodell mit 13 Einflußgrößen. Von 10 Einflußgrößen ist der Einfluß auf die Zielgröße mindestens auf dem 95%-Niveau gesichert, für die übrigen 3 auf dem 90%-Niveau. Das Bestimmtheitsmaß für diesen Ansatz beträgt $R^2=0.85$, das Abweichungsmaß beträgt $\alpha=0.042$ N/mm². Der Shapiro-Wilk-Test ergab für die Störterme, daß auf dem Signifikanzniveau von $\alpha = 0.05$ Normalverteilung angenommen werden kann.

Auf Grundlage dieses Prozeßmodells war es möglich besonders wichtige Prozeßparameter dieser Produktionsanlage zu identifizieren. Auch konnten Entwicklungsdefizite derzeit noch fehlender on-line Meßtechniken definiert werden. Darüber hinaus wurde das Prozeßmodell zur Eingrenzung und Analyse von Schwachstellen des Produktionsprozesses und damit zu dessen Optimierung genutzt.

Schlußfolgerungen

Aufgrund der vorgestellten Ergebnisse kann abschließend festgehalten werden, daß die Vorhersage der Produktqualität von Holzwerkstoffen (hier Spanplatten) mit Hilfe der statistischen Prozeßmodellierung möglich ist. Dabei ist es notwendig die Prozeßvariablen durch eine Kombination von statistischen und technologischen Aspekten auszuwählen. Die erforderlichen statistischen Werkzeuge stellt das SAS-Programmpaket zur Verfügung. Die Überprüfung der Modellgüte ist für die Prozeßsteuerung unerlässlich.

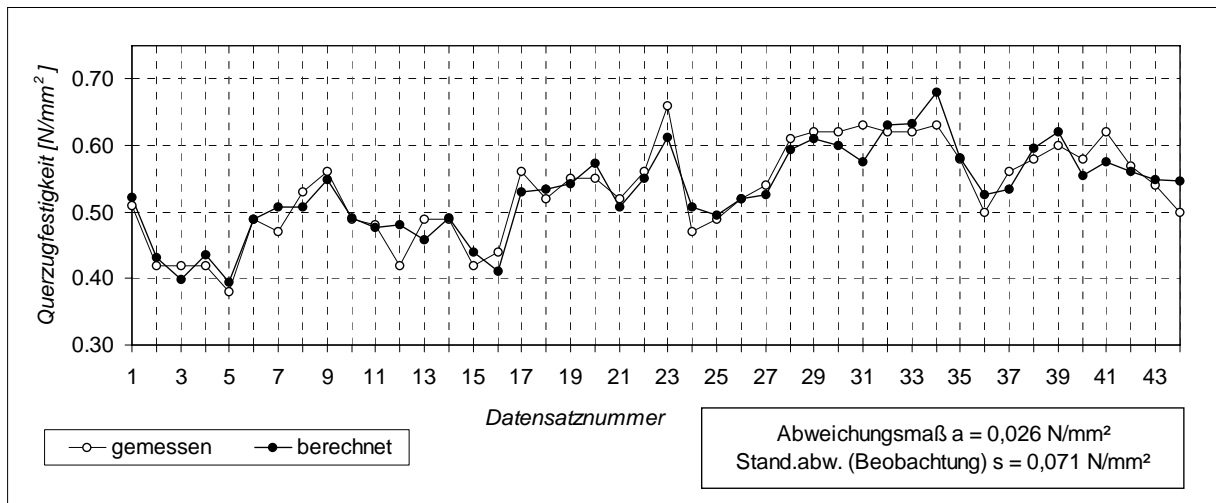


Abb. F-1: Gemessene und berechnete Werte für die Querzugfestigkeit (Variablenauswahlverfahren MAXR; 11-Variablen-Modell; einfache Anpassung)

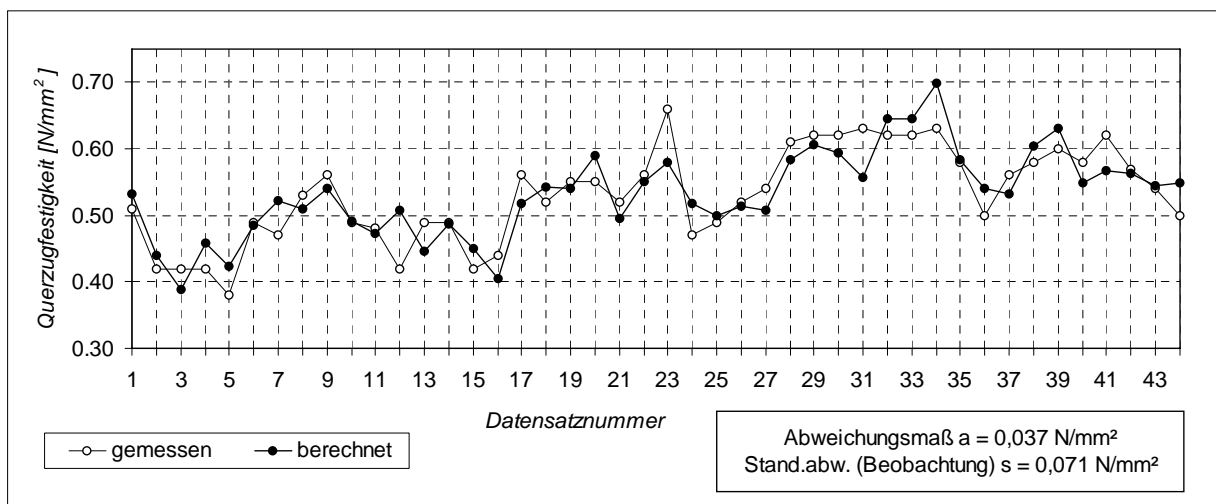


Abb. 2: Gemessene und berechnete Werte für die Querzugfestigkeit (Variablenauswahlverfahren MAXR; 11-Variablen-Modell; *cross-validation* Verfahren)

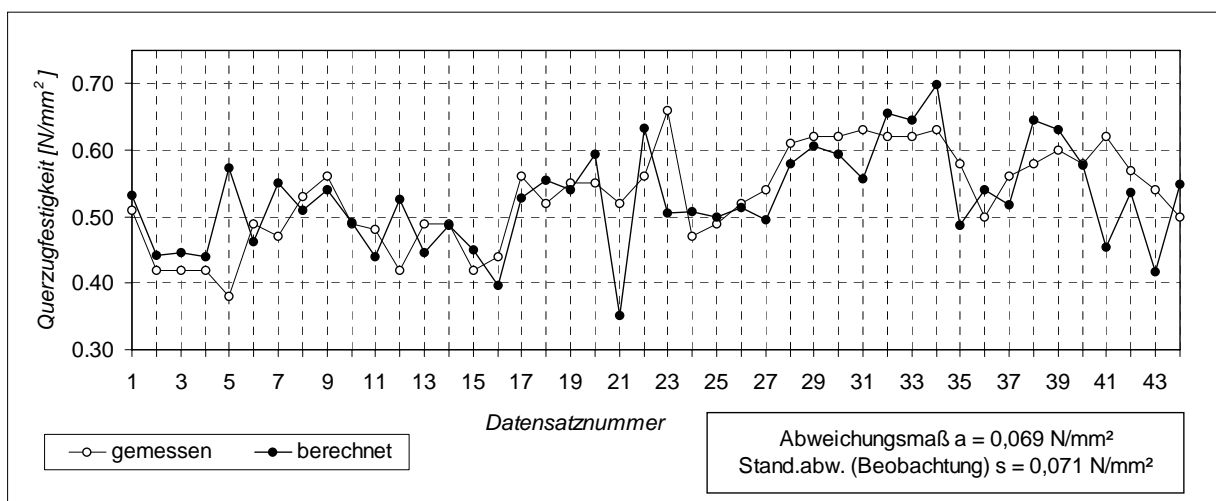


Abb. 3: Gemessene und berechnete Werte für die Querzugfestigkeit (Variablenauswahlverfahren MAXR; 11-Variablen-Modell; abgewandeltes *cross-validation* Verfahren)

Literatur

AITKIN, M. et al.: Statistical modelling in GLIM. Oxford 1989.

BERNARDY, G.: SPOC: Echtzeit-Qualitätskontrolle, Kostenoptimierung, Simulation - Erschließung wesentlicher Einsparpotentiale. Vortrag beim ATR Symposium Prozeßleittechnik, Prozeßoptimierung am 26.5.1995 in Hannover

FARNUM, N.; STANTON, L.: Quantitative forecasting methods. Boston: PWS-KENT Publ. Co. 1989

FRÜHWALD, A.; RUF, A.; STEFFEN, A. UND KRUSE, K.: Perspektiven der Qualitätssicherung in der Spanplattenindustrie. Holz-Kunststoffverarb. 1995 (30) Nr. 11, S. 1375-1379.

JAHN, W.; KÜHNE, P.; VAHLE, H.: Analyse und statistische Prozeßmodellierung für die Prozeßsteuerung. VEB Verlag Technik Berlin 1972, 67 S.

KRUSE, K.; THÖMEN, H.; MAURER, H.; STEFFEN, A. UND LEÓN-MÉNDEZ, R.: Optimierung des Spanplattenproduktionsprozesses mit Hilfe der Prozeßmodellierung. Holz Roh- Werkst. 1997 (55) Nr. 1, S. 17-21.

LEÓN-MÉNDEZ, R. UND THÖMEN, H.: Untersuchungen zur Optimierung eines Prozeßleitsystems in einem Spanplattenwerk. Diplomarbeit, Universität Hamburg, Fachbereich Biologie (1996).

LOBENHOFFER, H.; JENSEN, L.; MAURER, H.; BECKER, R.: Das Prozeßleitsystem (PLS) als Mittel der kundennahen Entwicklung und Produktion. Holz-Zentralblatt (1995) Nr. 6, S. 97

LOBENHOFFER, H.: Qualitätskontrolle in Echtzeit und Prozeßoptimierung bei der Spanplattenfertigung. Holz-Zentralblatt 1988 Nr. 8 (114), S. 101-108

LOBENHOFFER, H.: Qualitätsbedingte Regelung eines Spanplattenformstrangs. Dissertation Georg-August-Universität Göttingen, Göttingen 1990

MAGER, H.: Moderne Regressionsanalyse. Otto Salle Verlag Frankfurt/Main, 1992

- PENA-MÖLLER, J.: The European Wood-Based Panel Industry. Proceedings: New Challenges for the Wood-Based Panles Industry, WKI/FESYP, Braunschweig 1996.
- ROLL, H.: Prozeßüberwachung und -lenkung in der Plattenherstellung unter Berücksichtigung der DIN ISO 9000 ff. Vortrag beim Holzwerkstoffsymposium der Mobil Oil AG am 9.9.1994 in Bad Wildungen
- SCHWEITZER, F.: Modellierung des Heißpreßvorgangs zur Herstellung von Spanplatten in diskontinuierlich arbeitenden Pressen. Dissertation TU Dresden, Dresden 1992