

# Big Data meistert die Barriere

## Durch Datenanalyse Barriereigenschaften prognostizieren und steuern

Mithilfe eines strukturierten Vorgehens können viele Materialeigenschaften von Bahnkunststoffen und Folien gut modelliert, prognostiziert und mit einer modellgestützten Regelung auch exakt gesteuert werden. Der Ausschuss wird so verringert, und die Produktionskosten sinken erheblich.

Schlagworte wie „Deep Learning“ und „Big Data“ verführen zu der Annahme, man könne mit freier Software und Datenanalysten auf einfache Weise Lösungen für technische Probleme finden. Die Analyse gescheiterter Industrieprojekte zeigt aber, dass Erfolge eben nur dann möglich sind, wenn erprobte Konzepte und Prozesswissen mit neuen Methoden sinnvoll zu Lösungen kombiniert werden. Dabei beträgt der Anteil der Big-Data-Methoden an solchen Lösungen maximal 20%. Zum Erfolg führt die richtige Kombination aus systematischem Vorgehen, klassischen Modellen und stabilen neuronalen Netzen oder anderen „Black-Box“-Methoden, also Methoden, die ohne spezifisches Prozess-Know-how arbeiten.

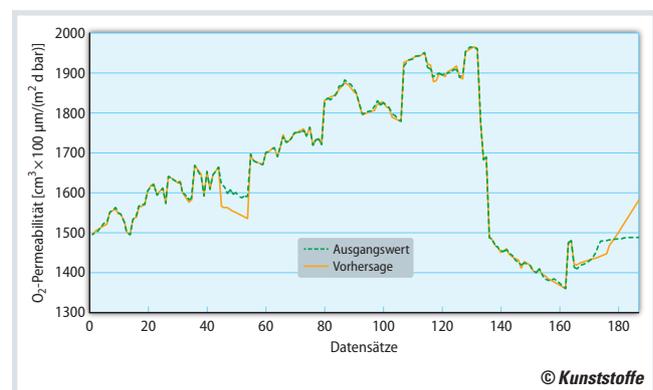
Der VDI hat die neue Richtlinie VDI3714 erarbeitet, die voraussichtlich im Juni 2019 als Gründruck erscheinen wird. Mit dem Titel „Implementierung und Betrieb von Big-Data-Anwendungen in der produzierenden Industrie“ [1] wurde sie in den letzten zwei Jahren von dem fachlich sehr breit aufgestellten Richtlinienausschuss RA 7.24 erarbeitet und beschreibt ein Vorgehen gemäß dem aktuellen Stand der Technik. Das Dokument befasst sich in sieben Dokumenten (Blättern) mit allen Inhalten, die für dieses Thema relevant sind und geht auf mehr als 180 Seiten auf Datenqualität, Datenimport (ETL-Prozess, Extract – Transport – Load), Analyseverfahrensklassen, Modelle, Modellvalidierung und auf die Online-Anwendung dieser Systeme ein. Wird diese Richtlinie konsequent angewendet und mit tiefem Prozesswissen, professioneller Software und dem Wissen über Prozessführung kombiniert, können damit Systeme entwickelt werden, die jeden technischen Produktionsprozess beherrschbar machen.

### SIPOC – Das vollständige Bild der Produktionsdaten

Immer wieder gibt es Versuche, aus reinen „Maschinendaten“ eine Aussage über Produktqualitäten zu erhalten. Da die Rohstoffe meist entscheidenden Einfluss haben, sind solche Versuche meistens zum Scheitern verurteilt. Zur erfolgreichen Prozessmodellierung und damit zur Prozessoptimierung ist ein vollständiges Gesamtbild aller Vorgänge um den Produktionsprozess herum zwingend erforderlich – es sein denn, der Rohstoff ist absolut konstant.

Eines der nach VDI3714 möglichen organisatorischen Prozessmodelle für Big-Data-Projekte ist die Nutzung von Six-Sigma-Konzepten. Six Sigma ist beschrieben in der Norm DIN-

ISO13053–2 [2]. Das SIPOC-Konzept („Supplier, Input, Process, Output, Customer“) wird im „Factsheet 9“ als Methode genauer ausgeführt. Im Wesentlichen geht es dabei darum, methodisch sicherzustellen, dass zur Modellierung eines Prozesses die Vollständigkeit von Daten gewährleistet ist.



**Bild 1.** Hohe Präzision des Modells: Vergleich der Modellprognose (orangefarbene Linie) mit Testdaten (grüne Linie) (Quelle: atlan-tec)

- „Supplier“ sowie „Input“ verweisen darauf, dass man sehr genaue Kenntnis über seine Quellen, Lieferanten und das Rohmaterial haben muss. Bei Kunststoffen bedeutet dies, man muss ggf. Farbwerte, mittlere Kettenlänge, Dichte, MFI/MFR (Melt Flow Index, Melt Flow Rate) und andere Werte der Rohstoffe kennen, die in den Prozess eingebracht werden. Im Idealfall müssen auch Rückführungen und Mischungsverhältnis der Rohstoffe zu jedem Zeitpunkt möglichst genau bekannt sein, die dem Prozess zugeführt werden. Ist die Rohstoffzufuhr unbestimmt, wird auch das Produkt unbestimmt sein.
  - „Process“ bedeutet, dass alle relevanten Prozessgrößen aus dem Prozess, also dem Extruder und den Walzen, gemessen werden müssen, die auf die Produktqualität wirken können.
  - „Output“ sowie „Customer“ bedeuten, dass es zumindest regelmäßige Laboranalysen oder Messungen geben muss, um die Produktqualität zu erfassen, und im Idealfall als KPI (Key Performance Indicator) auch eine Bewertung des Endkunden.
- Um Qualitätsgrößen dem Rohstoff und den Maschinendaten sinnvoll zuordnen zu können, müssen die Probenahmezeit- ➤

punkte und die Laufzeiten des Materials in der Maschine exakt bekannt sein. Je genauer und umfassender alle diese Daten und die Erfassungszeitpunkte bekannt sind, desto zuverlässiger lässt sich ein Prozess mithilfe der neuen Methoden modellieren. Die Qualität vorhandener Daten ist in der Praxis allerdings eines der größten Hindernisse bei der Nutzung intelligenter Technologien.

Übertragen auf Extrusion und Blasformen bedeutet dies, dass die wichtigsten physikalischen Eigenschaften der Rohstoffe bekannt sein müssen. Zudem sollten Rückführungen von Recyclingmaterial mengenmäßig ebenfalls genau erfassbar sein und einer Materialverfolgung unterliegen. Außerdem müssen zu dokumentierten Zeitpunkten Qualitätsanalysen oder Probenentnahmen erfolgen. Kurzum: Es ist eine „Good Manufacturing Practice“ erforderlich, wie sie sich eigentlich für jeden qualitätsbewussten Hersteller gehört. Ob diese Anforderungen erfüllt sind, lässt sich mit einem Readiness Check prüfen, der einen Aufwand von zwei Manntagen erfordert und aus ca. 150 Fragen besteht, die mit einem professionellen Berater bearbeitet werden müssen.

### Datenbereinigung und Modellierung

Stehen die vollständigen Daten zur Verfügung, müssen diese erst einer Aufbereitung und Datenbereinigung unterworfen werden. Hierfür stehen verschiedene Methoden zur Verfügung, von denen einige in professionellen Softwarewerkzeugen auch vollständig automatisiert angeboten werden. Ziel ist es, das Produkt durch den Prozess zu verfolgen, Ausreißer zu erkennen (z.B. fehlerhafte Laboranalysen), Inkonsistenzen und Redundanzen

## Die Autoren

**Dipl.-Ing. Thomas Froese** ist Vorsitzender des VDI/VDE/GMA-Richtlinienausschusses RA 7.24 Big Data und Geschäftsführer der atlan-tec Systems GmbH, Mönchengladbach (ab August 2019); T.Froese@atlan-tec.com

**Dipl.-Ing. Christoph Kugler** ist Mitarbeiter der Arbeitsgruppe Prozessmesstechnik am SKZ Würzburg und spezialisiert auf maschinelles Lernen und Industrie 4.0; c.kugler@skz.de

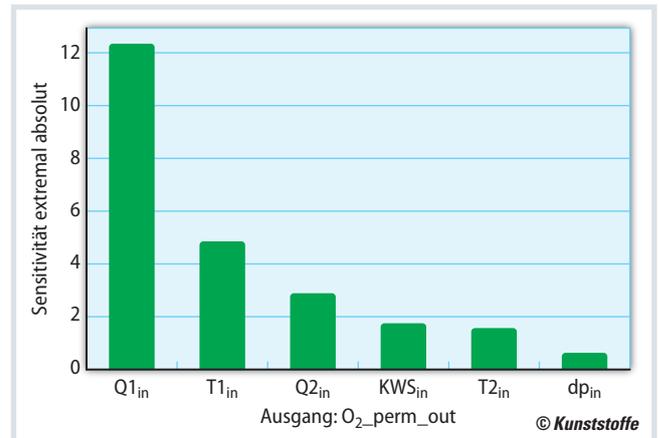
## Entropieanalyse

In der Informationstheorie ist Entropie der mittlere Informationsgehalt einer Nachricht oder eines zeitlichen Abschnitts einer Zeitreihe. Gleicht man den Informationsgehalt zweier Zeitreihen miteinander ab, spricht man von einer Verbundentropie. Aus der Verbundentropie lässt sich eine Art Korrelationswert berechnen, der erheblich genauer ist als die bekannten statistischen Korrelationsverfahren, da der Grad der Nichtlinearität keine Rolle spielt. Somit eignet sich die Methode, um wichtige Einflussgrößen, z.B. auf eine Qualität, genau zu bestimmen.

## Service

### Digitalversion

Ein PDF des Artikels finden Sie unter  
[www.kunststoffe.de/2019-06](http://www.kunststoffe.de/2019-06)



**Bild 2.** Einfluss der Modelleingänge auf die Permeabilität: Die Sensitivitätsanalyse zeigt die maximale Veränderung der Permeabilität bei Veränderung eines Eingangs um 1% seines Wertebereichs (Quelle: atlan-tec)

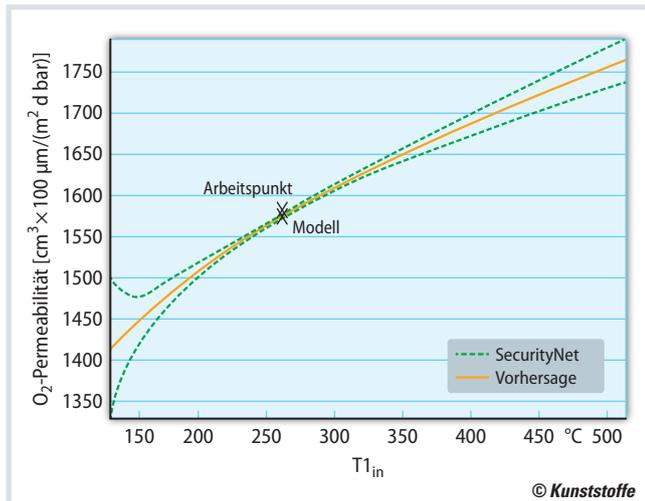
aus den Datensätzen zu entfernen und damit eine saubere und widerspruchsfreie Datenbasis zu schaffen, die eine genaue Analyse erst erlaubt.

Stand der Technik ist dann die Anwendung von Korrelationsverfahren, wobei lineare Korrelationen und PCA (Principal Component Analysis, Hauptkomponentenanalyse) sich kaum eignen, da sie nur lineare Korrelationen finden. Am besten sind neuere Methoden geeignet, wie Mutual Information (ein informationstheoretischer Begriff, der die Stärke des Zusammenhangs zwischen zwei Größen angibt) oder Entropieanalyse (siehe **Kasten**), da sie eine erste Orientierung geben, welche Rohstoffeigenschaften und Maschinenmessdaten einen Einfluss auf die Produktqualitätsparameter haben. Einige Erkenntnisse sind trivial, da zum Beispiel die zugeführte Materialmenge und die Ziehgeschwindigkeit ganz sicher einen Einfluss auf die Dicke einer Folie haben werden. Solche Sachverhalte kann man nur mit Prozesswissen erkennen und z.B. nicht dazu passende Daten aussortieren oder berichtigen.

Besondere Vorsicht ist hier geboten, da die Verfahren nur Korrelationen und keine Kausalitäten aufdecken. Dennoch lohnt sich der kritische Vergleich des Erfahrungswissens der Maschinenbediener mit den Ergebnissen dieser Analysen. Zunächst lassen sich so Erklärungen oder bisher unbekannte Einflussgrößen finden, und Ergebnisse der statistischen Analysen können mit Erfahrungswissen überprüft werden. Sind die Haupteinflussgrößen bekannt, lassen sich die Daten im Anschluss verwenden, um daraus Modelle zu erstellen. Auch hier gibt es moderne Softwarelösungen, die solche Modelle vollautomatisch aus den Daten erstellen. Diese Modelle werden dann mit Datensätzen validiert, die nicht zum Training der Modelle verwendet wurden.

### Modelle zur Vorhersage der Sauerstoffpermeabilität von Folien

In diversen Tests an einem gleichsinnig drehenden Doppelschneckenextruder, einer Breitschlitzdüse und einem Walzwerk wurde am SKZ, Würzburg, ein Atline-Permeabilitätsmessung installiert und die Heliumpermeabilität gemessen, woraus sich die Sauerstoffpermeabilität mit einer linearen Gleichung berechnen lässt. Im Technikum wurden die Materialien HDPE, PA46, PET, PP



**Bild 3.** Schnitt durch das 7-dimensionale Modell (alle bis auf einen der sechs Modelleingänge verbleiben auf ihrem Mittelwert): Das Beispiel zeigt, dass sich bei Erhöhung einer bestimmten Prozesstemperatur auch die Permeabilität erhöht. Orange: Modellausgang; grün: Modellunschärfe wegen fehlender Datensätze (Quelle: atlan-tec)

und PS getestet; in der späteren Installation bei einem industriellen Auftraggeber nur PP und PE. Mithilfe einer Entropieanalyse ließen sich zwei Rohstoffeigenschaften, ein Differenzdruck, die Walzendrehzahl und zwei Temperaturen als wichtigste Einflussgrößen auf die Sauerstoffpermeabilität identifizieren und für die Modellierung verwenden. Dann wurde das Modell an 186 Testdaten getestet, die nicht in der Modellierung genutzt worden waren.

Das Modell zeigt an einem HDPE-Material mit Pigmenten die tatsächlich gemessenen Permeabilitäten der produzierten Folie im Vergleich mit den vorhergesagten Permeabilitäten des neuronalen Modells und sagt so die Permeabilität der Proben sehr genau vorher (**Bild 1**); kleinere Abweichungen sieht man nur zwischen den Datensätzen 45 bis 55 und gegen Ende. Diese Abweichungen können aber auch Laborfehler bei der Analyse der Probe sein, zumal das Modell an den anderen Stellen eine sehr hohe Genauigkeit zeigt.

An den Modellen lassen sich dann Sensitivitätsanalysen durchführen und einzelne kausale Zusammenhänge zwischen Eingangsgrößen und Qualitätsgrößen untersuchen. Dazu werden die Eingangsdaten des Modells variiert, und die Reaktion der Permeabilität wird beobachtet und statistisch ausgewertet.

**Bild 2** zeigt eine daraus resultierende Gewichtung der Modellvariablen nach ihrer Bedeutung. Hier zeigt sich, dass die Qualität eines bestimmten Zuschlagsstoffes ( $Q_{1,in}$ ) die Permeabilität sehr stark beeinflusst. **Bild 3** zeigt exemplarisch den Einfluss einer bestimmten geregelten Prozesstemperatur auf die Permeabilität.

### Softsensoren und intelligente Prozessregelung

Damit eignet sich das Modell als sogenannter „Softsensor“; es kann also eine Laboranalyse ersetzen oder mindestens ergänzen und deren Ergebnis prognostizieren. Statt einzelner Stichproben und einer damit verbundenen destruktiven Probenahme durch Ausstanzen von Probekörpern kann ein solcher Softsensor sekundlich eine genaue Prognose der Permeabilität vorwegneh-

men und im Fall einer Qualitätsabweichung sofort alarmieren, sodass Gegenmaßnahmen eingeleitet werden können. Schon durch diese Maßnahme steigt die Prozessfähigkeit deutlich, da jede Qualitätsabweichung umgehend angezeigt werden kann. Dazu wird der Softsensor an die Maschinensteuerung angebunden und prognostiziert dort als virtuelle Messgröße sekundlich die bei dem momentanen Betriebszustand zu erwartende Sauerstoffpermeabilität.

Das Modell lässt sich auch nutzen, um den Prozess so zu steuern, dass es gar nicht erst zu einer Abweichung der Qualität kommen kann. Hier können zum Beispiel Extruderdrehzahl, Walzendrehzahl und Temperaturen so gesteuert werden, dass Änderungen der Rohstoffeigenschaften auf diese Weise ausgeglichen werden. Hierzu werden spezielle Optimierungsalgorithmen auf der Basis genetischer Algorithmen verwendet, die auf die Maschinensteuerung einwirken und Sollwerte verändern.

### Fazit

Mithilfe einer standardisierten Lösung, die Gegenstand mehrerer Normen ist und bereits in anderen Industriebereichen, wie Chemieindustrie, Raffinerien und Wasseraufbereitung im Einsatz ist, ließ sich in kurzer Zeit eine Lösung erarbeiten, die Laboranalysen einspart und eine deutliche Verbesserung der Prozessfähigkeit und Qualität ermöglicht. Neben den Barriereigenschaften von Folien lassen sich damit auch andere Größen zuverlässig prognostizieren und steuern. ■