

Algebraisches Erdöl

Prof. Dr. Martin Kreuzer
Fakultät für Informatik und Mathematik
Universität Passau
94030 Passau

Dr. Hennie Poulisse
Shell International Exploration & Production
NL-2288 GD Rijswijk, Niederlande

Martin.Kreuzer@uni-passau.de
Hennie.Poulisse@shell.com



Zusammenfassung

Mit Hilfe der Computeralgebra werden neuartige Modellgleichungen berechnet, die es gestatten, das Verhalten eines Ölfelds unter Produktionsbedingungen über längere Zeiträume korrekt vorherzusagen. Dabei werden symbolische Berechnungen mit numerischen Verfahren kombiniert, um ein sogenanntes *approximatives Verschwindungsideal* einer Menge von Datenpunkten zu bestimmen. Approximatives Verschwinden bedeutet hier, dass der Wert einer Modellgleichung an den Input-Datenpunkten nur ungefähr null sein muss, da die Daten Messfehler enthalten können. Das Verfahren wird an einem konkreten Beispiel unter Verwendung des Computeralgebrasystems ApCoCoA ausprobiert.

Die Erdölförderung und ihre Probleme

Bei der Förderung von Erdölvorkommen treten eine Reihe von Problemen auf, die bisher mit traditionellen Techniken der Geologie, der Physik und der angewandten Mathematik nicht zufriedenstellend gelöst werden konnten. Eine der Hauptaufgaben dabei ist, Modellgleichungen für die Produktion von Öl oder Gas aus Reservoiren zu finden, die es erlauben, das Verhalten des Systems über längere Zeiträume vorherzusagen. Der klassische Ansatz dafür ist, physikalische Gleichungen aus der "Flüssigkeitsdynamik in porösen Medien" zu verwenden und sie an die jeweilige geologische Situation anzupassen. Auf Grund seismischer und anderer Messungen hat man eine grobe Vorstellung von der Gestalt des Reservoirs.

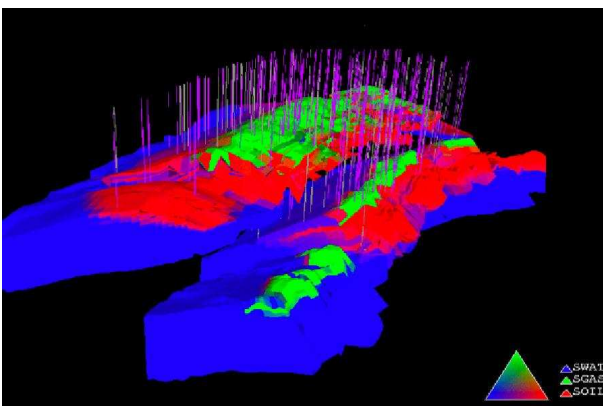


Abb. 1: Simulation eines Ölfelds und seiner Bohrungen

Allerdings sind die klassischen Modelle aus mehreren Gründen für Vorhersagen nicht geeignet. Zum einen kennt man die meisten der in den partiellen Differentialgleichungen der physikalischen Modelle auftretenden

Strukturkonstanten nicht bzw. nicht genau genug (z.B. die exakte Durchlässigkeit der Erdschichten, unterirdische Verwerfungen etc.), und andererseits besitzen die Differentialgleichung meist viele Lösungen, so dass sie so ziemlich an jede denkbare zukünftige Entwicklung angepasst werden können. Derartige Modelle spiegeln also mehr unsere Ansichten über das physikalische System als dessen wahre Struktur wieder.

Die Folgen dieser Unkenntnis sind dramatisch. In fast allen Fällen wird weniger als 30% des in einem Reservoir vorhandenen Erdöls oder Erdgases gefördert, bevor eine weitere Förderung unmöglich wird. Eine typische Ursache dafür ist die Bildung sogenannter *Dry Spots*, d.h. das Reservoir zerfällt in viele kleinere Reservoirs, aus denen nicht mehr wirtschaftlich gefördert werden kann.

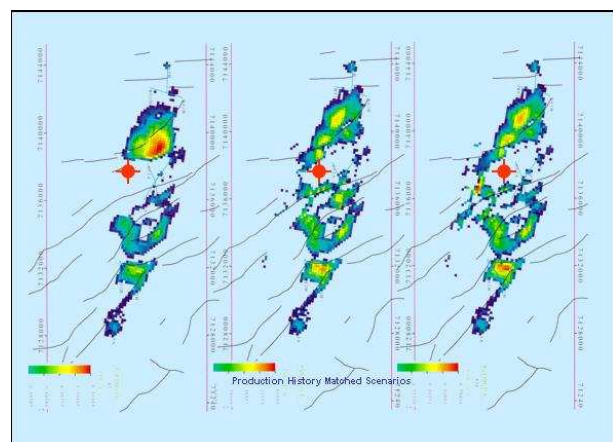


Abb. 2: Zeitliche Entwicklung eines Ölfelds

Ferner kann es zum sogenannten *Water Break-through* kommen, d.h. ein falsches Reservoirmanagement führt zu einem frühzeitigen Wassereinbruch, und

ab diesem Zeitpunkt wird nur noch Wasser gefördert. Oder aber zu Beginn wird zu viel Gas gefördert, so dass der Reservoir-Druck zu schnell abfällt und eine weitere Förderung wegen des Druckverlusts unmöglich wird. All diese Prozesse sind nach heutigem Kenntnisstand irreversibel, so dass in diesen Fällen eine weitere Bewirtung des Felds ausgeschlossen ist und große Mengen Öl im Boden verbleiben müssen. Die Folgen für die Umwelt sind beträchtlich: da seit Jahren kein großes Ölvorkommen mehr gefunden wurde, wird inzwischen in ökologisch sehr sensiblen Bereichen wie Alaska gebohrt und gefördert.

Modellgleichungen aus der Computeralgebra

Der Ansatz des *Algebraic Oil* Projekts, das die Firma Shell Int. Exploration & Production in Zusammenarbeit mit dem Lehrstuhl für Symbolic Computation der Universität Passau betreibt, ist radikal anders. Als Grundlage für die gesuchten Modellgleichungen sollen nur Messdaten herangezogen werden. Aus diesen Daten werden die Gleichungen dann mit Hilfe der Computeralgebra berechnet. Um dieses Verfahren zu erklären, beginnen wir zunächst mit der klassischen Situation bei approximativen Interpolationsaufgaben.

Gegeben sei eine Menge von (exakten) Punkten $X = \{p_1, \dots, p_s\}$ in \mathbb{R}^n . Die Gleichungen, die wir suchen, sind *Polynome*. Ein Polynom ist ein Ausdruck der Form $f = c_1 t_1 + c_2 t_2 + \dots + c_m t_m$, wobei die c_i reelle Zahlen und die t_i Terme $t_i = x_1^{a_1} \dots x_n^{a_n}$ (mit $a_j \geq 0$) in den Unbestimmten x_1, \dots, x_n sind. Wir suchen nun nach Polynomen, die an den Punkten von X *Nullstellen* besitzen, d.h. wenn $p_j = (p_{j1}, \dots, p_{jn})$ die Koordinatendarstellung des Punkts p_j ist, so liefert die Einsetzung $x_k \mapsto p_{jk}$ in f das Ergebnis $f(p_{j1}, \dots, p_{jn}) = 0$.

Die Menge aller Polynome ist der *Polynomring* $\mathbb{R}[x_1, \dots, x_n]$. Die Teilmenge $I(X)$ aller Polynome, die an allen Punkten von X Nullstellen besitzen, bildet das sogenannte *Verschwindungsideal* von X . Die Elemente des Verschwindungsideals können als polynomiale Modelle aufgefasst werden, d.h. als Gleichungen, die polynomiale Relationen unter den Input-Daten darstellen. Modelliert man eine Output-Datenreihe ebenfalls durch ein polynomiales Modell, so kann man die Äquivalenz zweier solcher Modelle bzgl. der Input-Daten überprüfen. Man braucht dazu nur zu testen, ob das Differenzpolynom der beiden Modelle im Verschwindungsideal liegt, was in der Computeralgebra mit Hilfe einer Gröbner-Basis möglich ist.

Liegt die Punktmenge X exakt vor, so kann man ihr Verschwindungsideal mittels des Buchberger-Möller Algorithmus [1] berechnen. Für Berechnungen mit echten Messdaten, die ja meist Messfehler beinhalten, ist dieser Algorithmus weniger geeignet. In diesem Fall sucht man ja auch nur Polynome, die auf den Datenpunkten *approximativ* verschwinden, d.h. deren Funktionswert betragsmäßig kleiner als eine vorgegebene Schranke $\epsilon > 0$ ist. Da alle Polynome $f = c_1 t_1 + \dots + c_m t_m$ mit extrem kleinen Koeffizienten c_j and den

Punkten von X approximativ verschwinden, sind wir in Wirklichkeit sogar nur an solchen Polynomen interessiert, deren Koeffizientenvektor normiert ist, d.h. für die $c_1^2 + \dots + c_m^2 = 1$ gilt, und die an den Punkten von X approximativ verschwinden. Genau diese Polynome berechnet der approximative Buchberger-Möller Algorithmus, der in [2] eingeführt wurde.

Ein Beispiel mit ApCoCoA

Zur Illustration betrachten wir das folgende Beispiel, wobei wir die Implementation des approximativen Buchberger-Möller Algorithmus im Computeralgebra-system ApCoCoA (vgl. [3]) verwenden. Mit Hilfe der Befehle

```
M:=Mat ([[0.9817,-0.191],[0.191,0.9817]]);
U:=[];
P:=Mat ([[1],[0]]);
For I:=1 To 500 Do
  P:=M*P;
  P[1,1]:=FloatApprox(P[1,1],10^(-6));
  P[2,1]:=FloatApprox(P[2,1],10^(-6));
  Append(U,[P[1,1],P[2,1]]);
EndFor;
```

erzeugen wir zunächst 500 Punkte, die in der Nähe des Einheitskreises liegen. (Aufgabe: Analysiere dieses Programm! Tipp: Der Punkt (1,0) wird jeweils um ca. 11° weitergedreht.) Dann betrachten wir die Kreise vom Radius $\sqrt{2}$ um den Punkt (2,2), die in den Ebenen $E_1: x - z = 0$ und $E_2: x + y - z = 6$ liegen. Mit

```
A:=[[2+P[2],2+P[1]*1.414,2+P[2]]|P In U];
B:=[[2+P[1]+P[2]*0.577, 2-2*P[2]*0.577,
     2-P[1]+P[2]*0.577] | P In U];
C:=Concat(A,B);
```

erzeugen wir dann 1000 Punkte, die wie zwei Gürtel stark gestört um die beiden Kreise liegen. (Aufgabe: Erkläre dieses Programm!)

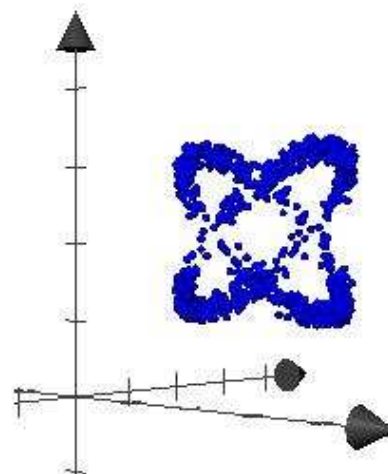


Abb. 3: 1000 Punkte in der Nähe zweier Kreise

Nun suchen wir nach polynomialen Relationen, die an diesen Punkten approximativ verschwinden. Wir machen über die Gestalt der zu berechnenden Gleichungen nur geringe Annahmen: die in den Relationen vorkommenden Variablen seien bekannt, und die gesuchten

Relationen seien polynomialer Natur. Der ApCoCoA-Befehl

```
L:=Numerical.GBasisOfPoints(C,0.08,False);
```

berechnet eine sogenannte Gröbner-Basis des approximativen Verschwindungsideals der 1000 Punkte, wobei $\epsilon = 0.08$ verwendet wird. Die resultierende Liste $L = [f_1, f_2, \dots]$ enthält Polynome f_1, f_2 mit

$$f_1 \approx -0.07x^2 - 0.07y^2 - 0.07z^2 + 0.29x + 0.29y + 0.29z - 0.71 \quad \text{und}$$

$$f_1 - 2.3 \cdot f_2 \approx -0.07x^2 - 0.07xy + 0.07yz + 0.07z^2 + 0.43x - 0.43z$$

Das erste ist ziemlich genau das -0.07-fache der definierende Gleichung der Sphäre

$$(x - 2)^2 + (y - 2)^2 + (z - 2)^2 = 2$$

und das zweite entspricht der Gleichung

$$(x - z)(x + y + z - 6) = 0$$

der beiden Ebenen, die wir ja zur Konstruktion der Punkte verwendet hatten.

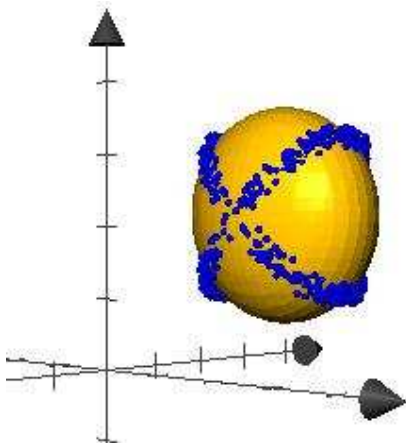


Abb. 4: Die Sphäre enthält die Punkte approximativ

Der klassische Buchberger-Möller Algorithmus hätte hier nicht funktioniert, da er jeden Punkt als exakte Nullstelle der Polynome realisiert hätte. Die approximative Variante hingegen findet in den (z.B. durch Messfehler) gestörten Daten die eigentliche Information und liefert auch bei 1000 Datenpunkten einfache und verständliche Ergebnisse.

An dieser Stelle können wir zwei weitere Vorteile des neuen Algorithmus vermerken. Zum einen ist es möglich, auch nicht-polynomiale Modelle zu finden, zum Beispiel Wurzeln oder exponentielle Funktionen der Variablen. Dazu genügt es, diese als zusätzliche Datenreihen der Berechnung hinzuzufügen. Bei der Anwendung in der Ölförderung etwa wurden Wurzeln aus Differenzen von Druckmessungen verwendet, da diese typische Bestandteile von physikalischen Gesetzen der Flüssigkeitsdynamik sind. Dass dieses Vorgehen funktioniert, liegt an der Stabilität des Verfahrens gegenüber

Erweiterungen des Modellhorizonts: fügt man dem Input unnötige neue Datenreihen hinzu, so werden diese automatisch ignoriert und die ursprüngliche Lösung wird gefunden. Wurden die neuen Datenreihen aus den vorhandenen berechnet, sind aber von diesen genügend unabhängig und für die Modellbildung nützlich, so werden sie mit einbezogen.

Der zweite und entscheidende Vorteil dieser Art der Modellbildung ist der, dass nicht nur eine optimale Approximation der Input-Daten bestimmt wird (wie es bei klassischen Methoden der approximativen Interpolation der Fall wäre), sondern polynomiale Gleichungen möglichst einfacher Form. Dies wird auch durch die Anwendung in der Ölförderung bestätigt, in der nicht nur bereits bekannte, sondern auch vollständig neue, durch weitere Messungen verifizierte Gesetzmäßigkeiten entdeckt wurden. Die so gewonnenen Ergebnisse erlauben Vorhersagen hoher Güte, wie die folgende Abbildung zeigt.

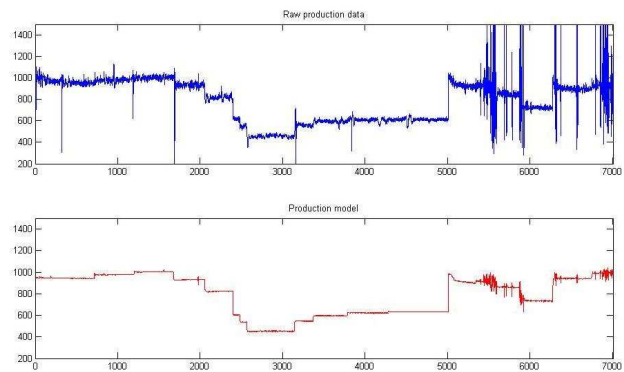


Abb. 5: Gemessene vs. vorhergesagte Gesamtproduktion

Interaktionen und ihre Interpretationen

Um das Entstehen von "Dry Spots" und anderen unerwünschten Effekten zu verhindern, ist es wichtig, die Ölförderung geeignet zu steuern. Dazu muss man die physikalischen Interaktionen innerhalb eines Reservoirs verstehen.

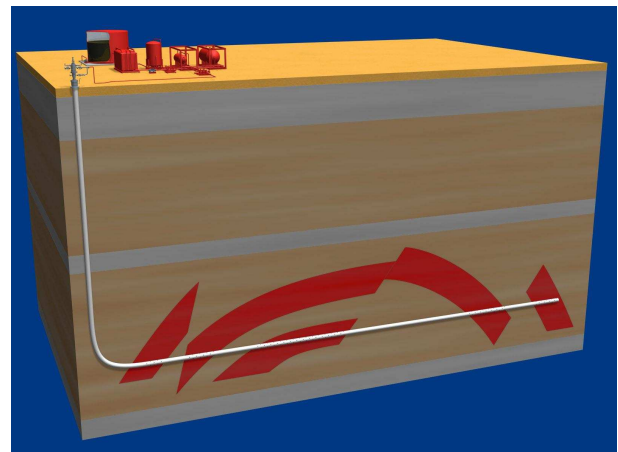


Abb. 6: Schematischer Aufbau einer Multi-Zonen-Quelle

Eine Multi-Zonen-Ölquelle besteht typischerweise aus verschiedenen *Taschen*, also Bereichen, die Öl oder Gas beinhalten. Zu den Taschen korrespondieren verschiedene Förderpunkte, die mit Sensoren für Drücke und Temperaturen ausgestattet sind. Man kann die Produktion einzelner Förderpunkte in einem Testbetrieb separat messen. Lässt man jedoch alle Förderpunkte gleichzeitig produzieren, so beeinflussen sie sich gegenseitig, weil das Öl in verbundenen Taschen zu verschiedenen Förderpunkten sickern kann. Die Kenntnis über die Verbundenheit bzw. Unabhängigkeit von Taschen spielt eine entscheidende Rolle bei der optimalen Bewirtschaftung eines Felds.

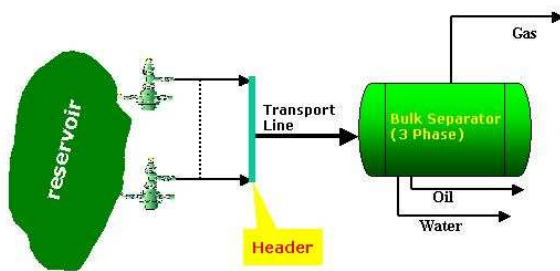


Abb. 7: Schematischer Aufbau der Messumgebung

Um die Interaktionen verschiedener Taschen zu berechnen, stellen wir die Modellgleichung f der Gesamtproduktion als Kombination der Modellgleichungen f_i der einzelnen Förderpunkte dar: $f = g_1 f_1 + \dots + g_m f_m$. Dabei sind die Koeffizienten g_i Polynome in *allen* Unbestimmten, also auch in den nicht zum i -ten Förderpunkt gehörigen. Die Berechnung einer derartigen Darstellung nennt man das explizite Idealzugehörigkeitsproblem. Sie ist mit den Ergebnissen des approximativen Buchberger-Möller Algorithmus problemlos möglich. Aus den Polynomen g_i kann man

Rückschlüsse über die tatsächliche Struktur des Felds ziehen. Dadurch kann man das Verhalten des Ölfelds vorhersagen und somit das Produktionssystem wesentlich gezielter steuern. Zum Beispiel kann eine hohe Gasproduktion eines Förderpunkts genutzt werden, um das Öl eines anderen Förderpunkts leichter zu machen und somit effizienter zu gewinnen, allerdings nur wenn das Mischverhältnis korrekt ist.

Mit der skizzierten Berechnung des approximativen Verschwindungsideals konnten feldspezifische Gesetzmäßigkeiten nachgewiesen werden, die durch physikalische und statistische Untersuchungen bestätigt wurden, aber durch ihre unerwartete Form höchstwahrscheinlich unentdeckt geblieben wären. Dabei ist zu bedenken, dass insbesondere die hohe Komplexität des Systems und die vielen Einzelgrößen eine klassische Deduktion dieser Gesetzmäßigkeiten verhindern. Die Geschwindigkeit der Berechnung ist hoch genug um auch komplexe Datensätze von mehreren 10000 Punkten bearbeiten zu können. Die vielversprechenden Ergebnisse machen Hoffnung, dass man die beschriebenen Methoden auch auf andere Problemstellungen übertragen kann.

Literaturverzeichnis

- [1] B. Buchberger und H.M. Möller, *The construction of multivariate polynomials with preassigned zeros*, LNCS **144**, Springer, Heidelberg 1982.
- [2] D. Heldt, M. Kreuzer, S. Pokutta und H. Poulisse, *Approximate computation of zero-dimensional polynomial ideals*, verfügbar unter staff.fim.uni-passau.de/~kreuzer
- [3] ApCoCoA – Applied Computations in Commutative Algebra, siehe www.apcocoa.org