

# **Bemerkungen zur Entwicklung der Internationalität der Forschung – Bibliometrische Untersuchungen am SCI**

MICHAEL HEINZ

*In der Arbeit werden verschiedene Kennziffern zur Messung der Internationalität der Forschung untersucht. Die Grundlage bilden die Daten des Science Citation Index (SCI) von 1980 bis 2002 in der CD-ROM Version. Alle betrachteten Kennziffern weisen einen einheitlichen Gesamttrend in diesem Zeitraum auf, der die Hypothese der Zunahme der Internationalität in der Forschung bestätigt. Zwei Kennziffern, der mittlere Anteil eines Landes an einem Artikel und die Diversität, gemessen durch die Shannonsche Entropie des Vektors der Anteile der Länder am SCI, zeigen eine charakteristische Verstärkung der Trends ab 1987, was für eine erhöhte Zunahme des Internationalisierungsprozesses der Forschung ab Mitte der 80er Jahre des vergangenen Jahrhunderts spricht. Darüber hinaus werden Zusammenhänge zwischen der ökonomischen Leistung eines Landes, seinem Anteil am SCI und seiner internationalen Forschungskooperation aufgezeigt.*

## **Ziel**

Von der Zunahme der Internationalität in der Forschung ist in unserer Zeit häufig zu lesen. Viele Arbeiten beziehen sich auf die Forschungskooperation von Wissenschaftlergruppen, Institutionen bzw. Ländern, oder es steht eine einzelne Disziplin im Mittelpunkt der Betrachtungen. Andere Arbeiten versuchen den Anteil einzelner Länder an der Forschung der Welt zu bestimmen und Zusammenhänge zum Bruttoinlandsprodukt und zu den Aufwendungen der Länder für Forschung und Entwicklung herzustellen, so wie es in dem interessanten Artikel von King [6] geschieht. Bei diesen Untersuchungen wird nicht selten auf Daten des Science Citation Index (SCI) zurückgegriffen. Diese interdisziplinäre Datenbank spiegelt den Stand und die Entwicklungstendenzen der Natur- und Technikwissenschaften gut wider. Sie diente und dient daher Generationen von Scientometrikern, Infometrikern bzw. Bibliometrikern als Grundlage für ihre Forschung.

Die vorliegende Arbeit basiert auch auf dem SCI, wertet alle dort verzeichneten Artikel aus und versucht ebenfalls, die Entwicklung der internationalen Repräsentanz im SCI anhand verschiedener berechneter Kennziffern sichtbar zu machen. Dabei spielen Maße der internationalen Kooperation, der Diversität bzw. Vielfalt und der Konzentration bzw. Ungleichheit, die auf die Verteilung der Länderanteile am SCI aufsetzen, eine besondere Rolle. Bei der Analyse werden Kennziffern und Maße verwendet, die sich bei ähnlichen Fragestellungen auch innerhalb anderer Fachgebiete wie der Biologie oder Ökonomie in der Literatur bereits bewährt haben.

### **Daten**

Den Untersuchungen liegen die Daten der CD-ROM Ausgaben des SCI aus den Jahren 1980 bis 1999 und 2002 zugrunde. Diese CD-ROM umfassen zusammen ca. 12 Millionen Artikel. Für die Untersuchungen wurden von den Artikeln das Adressfeld und das Publikationsjahr ausgewertet. Die Adressangaben beim SCI haben eine bestimmte Struktur. Die Länderangaben stehen jeweils am Ende von ihnen. Insgesamt wurden in den 21 untersuchten Jahren bei den 12 Millionen Artikeln 292 verschiedene Länderbezeichnungen gefunden. Diese wurden geprüft, und Verschreibungen wurden korrigiert. Darüber hinaus wurden die Bezeichnungen für die beiden deutschen Staaten auch für die Jahre von 1980 bis 1990 zu GERMANY und die Teile Großbritanniens, England, Wales, Nordirland und Schottland, die im SCI einzeln auftreten, generell zu UNITED-KINGDOM zusammengefasst. Ausschlaggebend für die jahresbezogenen Auswertungen in dieser Arbeit ist das Publikationsjahr des Artikels und nicht das Jahr der CD-ROM-Ausgabe. Damit ist bei den Randjahren der Untersuchung 1999 und 2002 keine Vollständigkeit vorhanden, da in den CD-ROM-Ausgaben der Folgejahre 2000, 2001 bzw. 2003, 2004 noch Artikel der Publikationsjahre 1999 bzw. 2002 enthalten sein werden. Der fehlende Anteil für die Jahre 1999 und 2002 wird, entsprechend der Situation in den anderen Jahren, kleiner sein als 10 %. Da in der Regel relative Kennziffern (Anteile, Mittelwerte) berechnet werden, lässt sich dies vernachlässigen.

### **Methoden**

#### Fraktionale Zählweise

Bei der Bestimmung des Anteils eines Landes am SCI benutzen wir eine fraktionale Zählweise. Wenn mehrere Länder an einem Artikel beteiligt sind,

also mehrere Länder in den Adressen eines Artikels vorkommen, so bekommen sie auch jeweils nur einen Teil des Artikels angerechnet. Diese Zählweise wurde in [8] für die Anrechnung der Autorschaft bei Artikeln mit mehreren Autoren vorgeschlagen.

In [3] stellen die Autoren verschiedene Zählweisen für die Autorschaft bzw. den Länderanteil vor und diskutieren ihre Vor- und Nachteile. Wir berechnen den Anteil eines Landes an einem Artikel proportional zur Häufigkeit, mit der dieses Land bei den Adressen des Artikels auftritt. Wenn ein Artikel 10 Adressenangaben von beteiligten Institutionen hat und dabei sechsmal USA, dreimal Frankreich und einmal Deutschland auftritt, dann erhält die USA den Anteil 0,6, Frankreich den Anteil 0,3 und Deutschland den Anteil 0,1 am Artikel angerechnet. Die Anzahl der Artikel eines Landes im SCI in einem Jahr ergibt sich aus der Summe der Anteile dieses Landes an Artikeln des Jahres. Diese Zahl wird in der Regel eine rationale und keine natürliche Zahl sein. Der Anteil eines Landes am SCI in einem Jahr ergibt sich aus der Anzahl der Artikel dieses Landes dividiert durch die Gesamtzahl der Artikel im SCI in dem entsprechenden Jahr.

Die einzelnen Anteile der Länder sortieren wir der Größe nach absteigend und fassen sie in einem Vektor  $X = (x_1, x_2, \dots, x_n)$  zusammen. Für diesen Vektor gilt demzufolge:

$$\sum_{i=1}^n x_i = 1 \quad \text{und} \quad x_i \geq x_{i+1} \quad \text{für alle} \quad i = 1, \dots, n-1.$$

Er wird für die Berechnung von Kennziffern der Diversität bzw. der Konzentration verwendet. Die Zahl  $n$  gibt gleichzeitig die Anzahl der beteiligten Länder am SCI in dem Jahr an.

#### Mittelwerte

Für die Messung der Kooperation von Wissenschaftlern wurde in [1] der Kooperationskoeffizient (collaboration coefficient)  $cc$  eingeführt:

$$cc = 1 - \frac{1}{N} \sum_{j=1}^k \frac{1}{j} \cdot F_j$$

wobei  $\sum_{j=1}^k F_j = N$  gilt,  $N$  die Anzahl der Artikel,  $F_j$  die Anzahl der

Artikel mit  $j$  Autoren und  $k$  die höchste auftretende Autorenzahl bei einem Artikel ist.

Es geht pro Artikel der mittlere Autorenanteil also  $1/j$  bei  $j$  Autoren in die Mittelwertberechnung ein. Die Differenz zu 1 ergibt den Kooperationskoeffizienten  $cc$ . Dieser Kooperationskoeffizient wurde bereits in vielen wissenschaftlichen Publikationen, wie beispielsweise in [5] oder [7], benutzt.

Seien  $y_1, \dots, y_m$  die Anteile eines Landes an Artikeln in einem Jahr, dann berechnet sich der mittlere Artikelanteil  $p$  eines Landes als arithmetisches Mittel der Anteile:

$$p = \frac{1}{m} \sum_{i=1}^m y_i.$$

Für  $p$  gilt stets  $0 \leq p \leq 1$ . Der mittlere Anteil der internationalen Kooperation bei einem Artikel eines Landes  $q$  ergibt sich, wie bei dem Kooperationskoeffizienten  $cc$ , aus der Differenz von  $p$  zu 1,  $q = 1 - p$ .

Zur Berechnung eines mittleren Artikelanteils  $r$  von Ländern insgesamt im SCI eines Jahres berechnen wir zunächst für jeden Artikel das arithmetische Mittel der Länderanteile ( $1/j$  wenn  $j$  Länder an dem Artikel beteiligt sind) und dann von diesen Mittelwerten wiederum das arithmetische Mittel.

Bei der Berechnung des Kooperationskoeffizienten  $cc$  wird analog vorgegangen.

#### Diversitäts- und Konzentrationsmaße

Diversität ist ein in der Biologie seit langem benutzter Begriff zur Kennzeichnung der Artenvielfalt der Flora oder Fauna zum Beispiel einer Region. Dabei geht die Anzahl der verschiedenen Arten, aber auch die Anzahl der Individuen innerhalb der Arten, also die Häufigkeitsverteilung der Individuen auf die Arten, in die Bestimmungen der Diversität ein. Je mehr Arten es gibt und je gleichmäßiger die Verteilung der Individuen auf die Arten ist, desto größer wird die Diversität eingeschätzt. Maße der Diversität müssen diese Eigenschaft haben. Sie müssen auch unempfindlich sein in Bezug auf Permutationen der Häufigkeiten in der Verteilung, denn es kommt bei der Diversität nicht darauf an, welche Art die Häufigkeit hat. Es kommt nur auf die Häufigkeiten selbst an.

#### *Shannonsche Entropie*

Die Shannonsche Entropie

$$H(X) = - \sum_{i=1}^n x_i \cdot \text{ld}(x_i) \quad (\text{ld bedeutet Logarithmus zur Basis 2})$$

wurde eigentlich eingeführt, um Unbestimmtheit und, da nach Shannon Information beseitigte Unbestimmtheit ist, Information zu messen. Sie erfüllt aber auch die Bedingungen für ein Maß der Diversität. Je höher die Diversität ist, desto unbestimmter ist es, die Art eines zufällig gewählten Individuums vorherzusagen. Bei unseren Untersuchungen spielen die Länder die Rolle der Arten und die Artikel die der Individuen.

Der Begriff Konzentration, wie wir ihn hier benutzen, ist dem Begriff der Diversität entgegengesetzt. Er wird in der Ökonomie unter anderem benutzt, um die Verteilung der Einkommen oder des Kapitals auf Personen zu kennzeichnen. Eine hohe Konzentration des Einkommens bedeutet, wenige Personen bekommen viel und viele eben wenig. In [4] werden Konzentrationsmaße auch für die Messung der Internationalität von Unternehmen eingesetzt. Wir werden drei Maße für die Konzentration in der Arbeit verwenden. In [2] werden diese Maße ausführlich behandelt, und es werden Bedingungen für gute Konzentrationsmaße formuliert. Die hier benutzten Maße erfüllen diese Bedingungen.

#### *Lorenz-Kurve*

Die Lorenz-Kurve wurde von dem amerikanischen Mathematiker Max Otto Lorenz entwickelt, um grafisch anschaulich Ungleichheit bzw. Konzentration bei statistischen Verteilungen sichtbar zu machen. Bei der Lorenz-Kurve wird die Ordnung der Häufigkeiten im Vektor vorausgesetzt. Wir betrachten als Beispiel die Verteilung  $X = (0,5; 0,25; 0,125; 0,125)$ . Zur Darstellung ihrer Lorenz-Kurve geht man zu der kumulierten Verteilung  $C = (0,5; 0,75; 0,875; 1)$  über. Die Lorenz-Kurve verbindet dann die Punkte mit den Koordinaten:

$$(0; 0), (0,25; 0,5), (0,5; 0,75), (0,75; 0,875), (1; 1).$$

Diese Lorenz-Kurve ist in der Abbildung 1 dargestellt. Je weiter sich die Kurve von der Diagonalen entfernt, desto größer ist die Konzentration.

#### *Gini-Koeffizient*

Der Gini-Koeffizient  $G(X)$  einer Verteilung  $X = (x_1; x_2; \dots; x_n)$  baut unmittelbar auf die Lorenz-Kurve der Verteilung auf. Er ergibt sich aus dem doppelten Flächeninhalt zwischen der Lorenz-Kurve der Verteilung und der Diagonalen und berechnet sich auf folgende Weise:

$$G(X) = 1 - \frac{2}{n} \left( \sum_{i=1}^n i \cdot x_i - 0,5 \right).$$

Der Gini-Koeffizient wurde von dem italienischen Mathematiker Carrado Gini als Maß für die Konzentration einer Verteilung eingeführt und liegt im Bereich von 0 bis 1 ( $0 \leq G(X) \leq 1$ ). Je näher er an 1 liegt, desto größer ist die Konzentration bzw. Ungleichheit der Verteilung, je näher er bei 0 liegt, desto größer ist die Diversität bzw. Vielfalt der Verteilung.

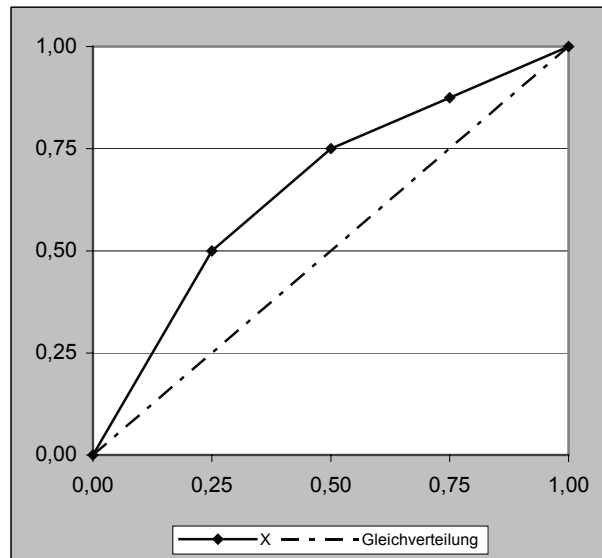


Abb. 1: Lorenz-Kurve für das Beispiel  $X$

#### Theil-Maß

Das Teil-Maß  $T(X)$  wurde in [9] eingeführt. Es stellt die Differenz der Entropie zur maximalen Entropie bei gegebener Dimension  $n$  des Häufigkeitsvektors  $X$  dar. Der Maximalwert der Entropie bei vorgegebener Dimension  $n$  der Verteilung ist  $ld(n)$ , und so ergibt sich:

$$T(X) = ld(n) - H(X) = ld(n) - \sum_{i=1}^n x_i \cdot ld(x_i)$$

Das Teil-Maß entspricht damit dem Ansatz, dass Diversität und Konzentration entgegengesetzte Begriffe darstellen.

## Ergebnisse

### SCI insgesamt

Die Abbildung 2 zeigt die Entwicklung der Zahl der Länder, die an den im SCI verzeichneten Artikeln im jeweiligen Jahr beteiligt waren. Der Anstieg der Länderzahl in den Jahren 1987 bis 1993 hängt insbesondere mit dem Zerfall der ehemaligen UdSSR, der ČSSR und Jugoslawiens zusammen. Der fallende Trend nach 1993 erklärt sich dadurch, dass nach dem Entstehen der neuen Länder über einige Jahre auch die alten Bezeichnungen noch in den Artikeln auftauchen. Das kann zum Beispiel an relativ langen Zeiträumen zwischen dem Einreichen und dem Erscheinen der Artikel bei Zeitschriften liegen. Allgemein kann festgehalten werden: Die Länder der Welt sind im SCI vertreten.

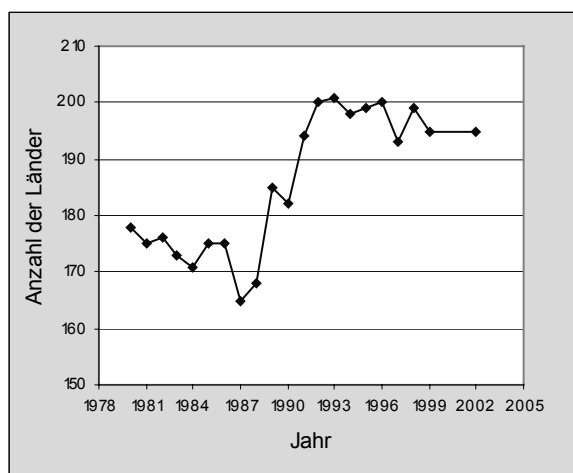


Abb. 2: Darstellung der Entwicklung der Zahl der beteiligten Länder an den im SCI verzeichneten Artikeln

Wenn wir die Anteile der einzelnen Länder am SCI den jeweiligen Kontinenten, zu denen sie gehören, zuordnen, ergeben sich für die Jahre 1992-2002 die in der Abbildung 3 dargestellten Tendenzen. Der Anteil Nordamerikas, der wesentlich durch den der USA bestimmt wird, ist in dieser Periode um ca. 6 % zurückgegangen. Um diesen Wert nimmt gerade der asiatische Anteil in dieser Zeit zu. Dafür sind insbesondere China, dessen Anteil in dem Zeitraum von 1,1 % auf 3,4 %, und Südkorea, dessen Anteil von 0,3 % auf 1,7 % gestiegen ist, verantwortlich.

Der Anteil Europas nimmt leicht ab, und der Anteil des Restes der Welt, der Afrika, Australien/Ozeanien, Süd- und Mittelamerika umfasst, nimmt im Beobachtungszeitraum um etwa 1 % zu. Ab dem Jahre 1996 ist der europäische Anteil am SCI größer als der nordamerikanische. Ein analoges Resultat findet man in [6]. Dort wird festgestellt, dass der Anteil der Europäischen Union am gesamten Web of Science, der neben dem SCI auch den Social Science Citation Index (SSCI) und den Art & Humanities Citation Index (A&HCI) umfasst, ab 1997 größer ist als der Anteil der USA.

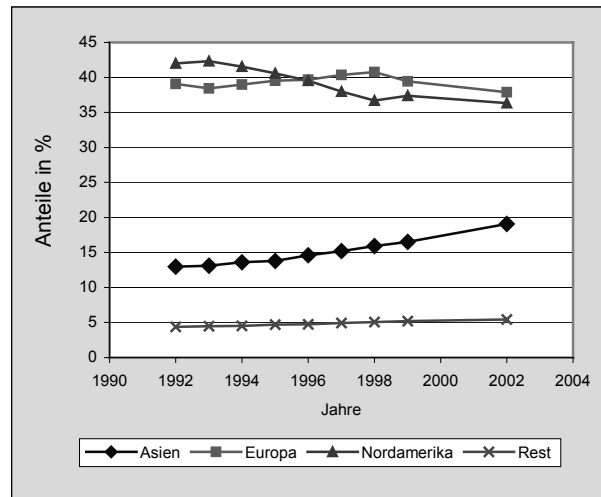


Abb 3: Darstellung der Entwicklung der Anteile der Kontinente am SCI 1992-2002

Die Abbildung 4 enthält die drei Lorenz-Kurven für die Vektoren der Länderanteile am SCI 1980 bzw. 2002 und den vom Bruttoinlandsprodukt (BIP) aus dem Jahre 2002. Die Daten der Bruttoinlandsprodukte wurden online vom Server des Internationalen Währungsfonds im April 2006 herunter geladen (<http://www.imf.org/external/pubs/ft/weo/2006/01/data/index.htm>).

Es zeigt sich, dass die Konzentration der Artikelverteilung auf die Länder im SCI von 1980 bis 2002 abgenommen hat. Die Konzentration der natur- und technikwissenschaftlichen Forschung ist aber auch 2002 noch größer als die der Bruttoinlandsprodukte der Länder. Die 10 % wirtschaftlich stärksten Länder der Welt erzielten im Jahre 2002 zusammen etwa 84 % des gesamten Bruttoinlandsproduktes der Welt, dagegen repräsentieren die 10 % forschungsmäßig stärksten Länder der Welt im Jahre 2002 sogar 90 % der Artikel im SCI.



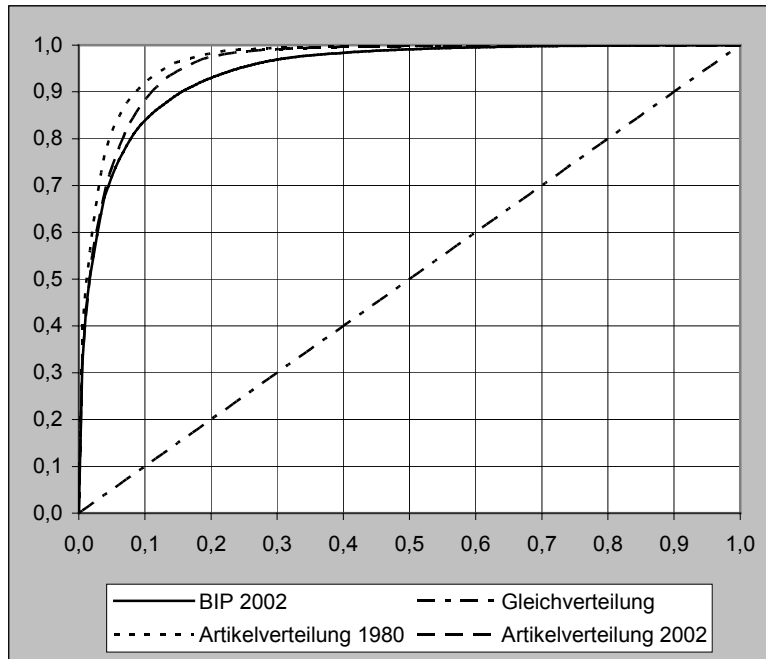


Abb. 4: Darstellung der Lorenz-Kurven für die Artikelverteilungen der Länder 1980 und 2002 sowie der Verteilung der Bruttoinlandsprodukte der Länder für das Jahr 2002

In der folgenden Tabelle 1 werden die 50 Länder mit den höchsten Artikelzahlen im SCI im Jahre 2002 aufgelistet. Die Spalte C enthält die Bruttoinlandsprodukte (BIP) dieser Länder und die Spalte D die erwarteten BIP entsprechend der Regressionsgeraden aus Abbildung 5. Die Spalte E gibt die Quotienten der Werte der Spalte C durch die Werte der Spalte D an.

Tabelle 1: Übersicht der Artikelzahlen und der BIP der 50 Länder mit den höchsten Artikelzahlen im Jahre 2002 im SCI

A	B	C	D	E
Länder	Artikel im SCI 2002	BIP 2002 in Millionen USD	Erwartetes BIP	C/D
USA	214256,7	10469600	10560601	0,99
JAPAN	55571,9	3911581	2726334	1,43
UNITED-KINGDOM	47542,9	1574003	2329940	0,68
GERMANY	41718,7	2025798	2042398	0,99

FRANCE	29015,2	1464204	1415226	1,03
ITALY	22854,5	1223272	1111071	1,10
CANADA	22605,2	735601	1098766	0,67
PEOPLES-R-CHINA	22123,2	1453837	1074970	1,35
SPAIN	15728,5	688501	759263	0,91
AUSTRALIA	13563,1	413570	652359	0,63
RUSSIA	13509,2	345071	649696	0,53
NETHERLANDS	12226,9	439357	586387	0,75
SOUTH-KOREA	11479,5	546935	549492	1,00
INDIA	11197,1	493335	535548	0,92
SWEDEN	9343,4	244314	444030	0,55
SWITZERLAND	8050,6	277113	380206	0,73
BRAZIL	7799,5	460612	367811	1,25
TAIWAN	7692,3	294876	362514	0,81
ISRAEL	6072,2	104266	282530	0,37
BELGIUM	5920,5	252721	275041	0,92
POLAND	5797,7	198039	268979	0,74
TURKEY	5117,2	182973	235383	0,78
DENMARK	4756,5	174420	217573	0,80
FINLAND	4586,1	133024	209161	0,64
AUSTRIA	4566,6	208422	208198	1,00
GREECE	3375,6	134456	149402	0,90
ARGENTINA	3099,4	97732	135762	0,72
MEXICO	3091,8	649078	135391	4,79
NORWAY	2805,5	191514	121253	1,58
NEW-ZEALAND	2492,2	59765	105788	0,56
SINGAPORE	2480,1	88468	105191	0,84
CZECH-REPUBLIC	2456,1	73756	104007	0,71
HUNGARY	2322,0	65562	97386	0,67
PORTUGAL	2181,6	127906	90452	1,41
SOUTH-AFRICA	2162,7	111138	89521	1,24
UKRAINE	1916,5	42393	77365	0,55
IRELAND	1725,3	123261	67927	1,81
EGYPT	1483,2	87506	55973	1,56
IRAN	1266,3	116412	45263	2,57
CHILE	1207,2	67266	42348	1,59
ROMANIA	827,2	45825	23587	1,94
SLOVAKIA	821,3	24239	23296	1,04
THAILAND	802,1	126877	22348	5,68
SLOVENIA	790,0	22292	21747	1,03

BULGARIA	751,8	15614	19865	0,79
SAUDI-ARABIA	569,1	188803	10845	17,41
CROATIA	559,6	22798	10374	2,20
YUGOSLAVIA	525,0	15528	8666	1,79
VENEZUELA	518,8	92889	8358	11,11
BELARUS	443,2	14654	4627	3,17

Wenn der Wert in der Spalte E kleiner als 1 ist, dann ist das betreffende Land gemessen an seinem BIP im SCI überdurchschnittlich und im anderen Fall unterdurchschnittlich vertreten. Überdurchschnittlich sind beispielsweise die englischsprachigen Länder Großbritannien, Kanada, Australien und Neuseeland vertreten. Die portugiesisch- bzw. spanischsprachigen Länder Portugal, Brasilien, Mexiko hingegen, wie auch die großen asiatischen Länder Japan und China sind unterdurchschnittlich vertreten.

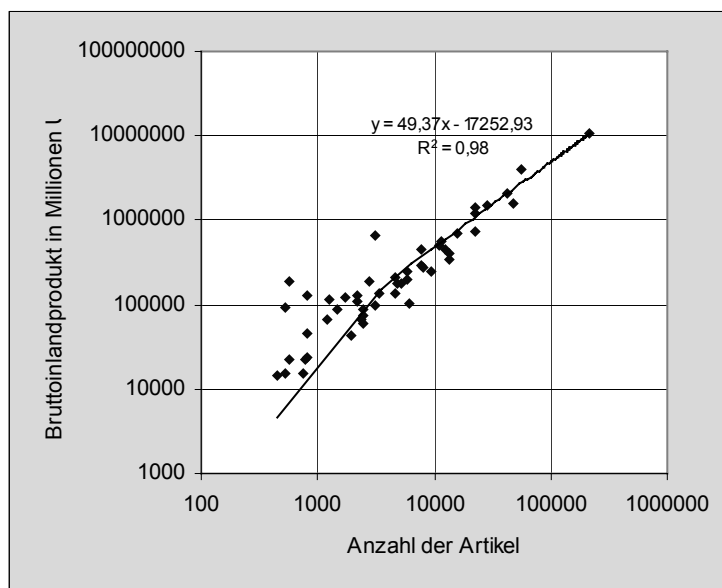


Abb.5: Darstellung der linearen Regression zwischen den Anzahlen der Artikel eines Landes 2002 im SCI und dem Bruttoinlandsprodukt des Landes im Jahre 2002 (hier wurden nur die 50 im Jahre 2002 am stärksten im SCI vertretenen Länder berücksichtigt)

Die Abbildung 5 stellt den linearen Zusammenhang zwischen den Spalten B und C der Tabelle 1 dar. Die Regressionsgerade ist wegen der doppelt-logarithmischen Darstellung etwas verzerrt. Das Bestimmtheitsmaß ist mit  $R^2=0,98$  hoch. Die Erhöhung der Anzahl der Artikel um eins korrespondiert mit einer Erhöhung des Bruttoinlandsproduktes um ca. 49 Millionen USD.

In der Abbildung 4 wurde mit Hilfe der Lorenz-Kurven die Konzentration der Artikelverteilungen auf die Länder im Jahre 1980 und 2002 sichtbar gemacht. In den folgenden drei Abbildungen 6-8 wird die Entwicklung der Konzentration bzw. Diversität dieser Verteilung im Untersuchungszeitraum 1980-2002 mit Hilfe des Gini-Koeffizienten, des Theil-Maßes und der Shannonschen Entropie dargestellt.

Die Verläufe der Kurven des Gini-Koeffizienten und des Theil-Maßes sind relativ analog. Wir erkennen jeweils eine tendenzielle Abnahme der Konzentration von 1980 bis 1987, eine tendenzielle Zunahme der Konzentration von 1987 bis 1993 und eine starke Abnahme nach 1993. Die Zunahme der Konzentrationswerte in den Jahren 1987-1993 basiert wesentlich auf der Erhöhung der Länderzahlen  $n$  in diesem Zeitraum.

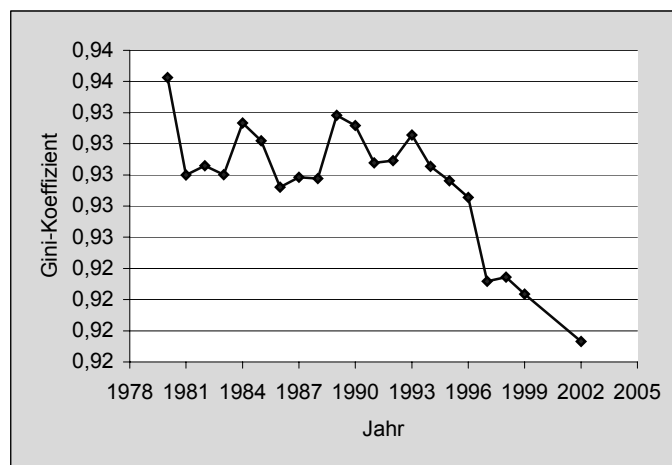


Abb. 6: Darstellung der Entwicklung des Gini-Koeffizienten der Verteilung der Artikel des SCI auf die Länder

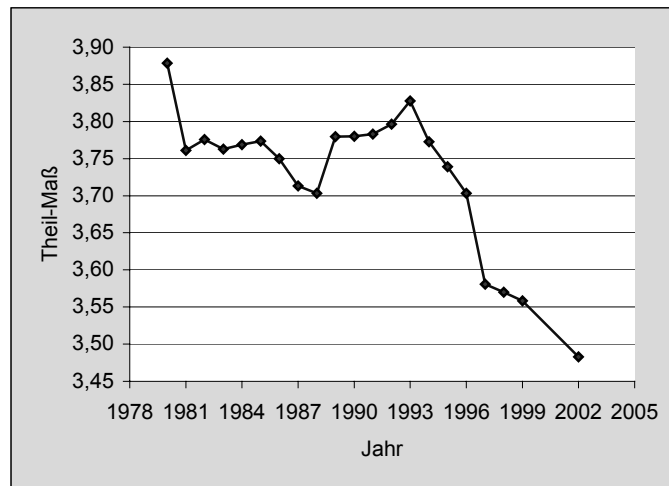


Abb. 7: Darstellung der Entwicklung des Theil-Maßes der Verteilung der Artikel des SCI auf die Länder

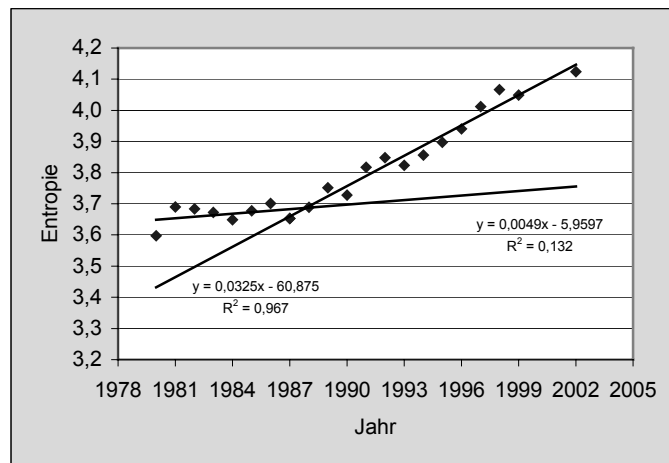


Abb. 8: Darstellung der Entwicklung der Entropie der Verteilung der Artikel des SCI auf die Länder mit eingezeichneten linearen Trendlinien für 1980-1987 und 1987-2002

Bei der Entropieberechnung geht die Erhöhung der Länderzahl  $n$  durch Aufteilung von Ländern (UdSSR, ČSSR, Jugoslawien) auch in die Berechnung ein und erhöht den Wert der Entropie. Dies geschieht allerdings hier nur mit dem Gewicht des Anteils der Länder, die aufgeteilt werden. Diese Anteile waren bei der UdSSR, der ČSSR und Jugoslawiens relativ gering. Dadurch zeigt die Kurve der Entwicklung der Entropie in Abbildung 8 nicht die drei Phasen wie in den Abbildungen 6 und 7. Wir erkennen nur zwei Phasen: Eine Phase 1980-1987 mit einem schwachen Anstieg und eine Phase 1987-2002 mit einem bedeutend stärkeren linearen Anstieg der Entropie und damit der Diversität der im SCI widergespiegelten internationalen Forschung.

Die Abbildung 9 zeigt die Entwicklung der mittleren Länderanteile  $r$  im gesamten SCI. Der mittlere Länderanteil  $r$  hat sich von 0,975 im Jahre 1980 auf 0,904 verringert. Damit hat sich der mittlere Anteil der internationalen Kooperation bei einem Artikel in diesem Zeitraum von 2,5 % auf 9,6 % erhöht. Wir erkennen, analog zur Entwicklung der Entropie, die beiden Phasen 1980-1987 und 1987-2002. Die linearen Trends für die beiden Phasen sind hier noch prägnanter, was die sehr hohen Werte der Bestimmtheitsmaße anzeigen.

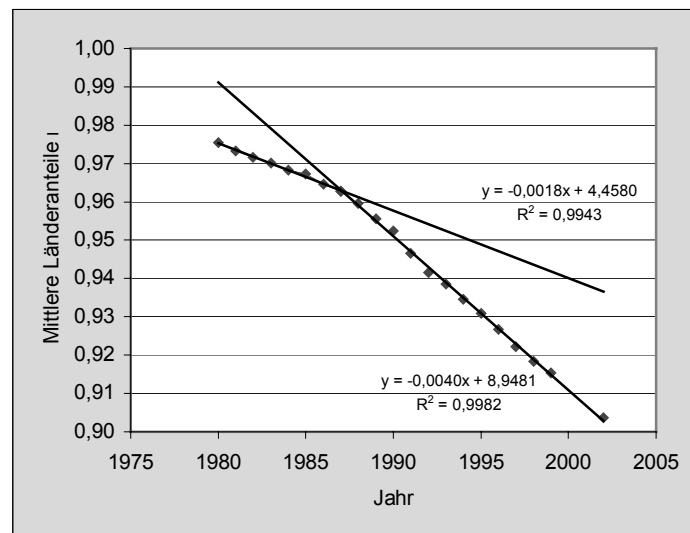


Abb. 9: Darstellung der Entwicklung der mittleren Artikelanteile  $r$  des gesamten SCI mit eingezeichneten linearen Trendlinien für die Zeiträume 1980-1987 und 1987-2002

Der Trendwandel 1987 hat nichts mit dem Zerfall der UdSSR, der ČSSR und Jugoslawiens und dem Entstehen von ca. 20 neuen Nachfolgestaaten nach 1987 zu tun. Die von uns durchgeführten Kontrollberechnungen führten zum quasi identischen Ergebnis – als hätte der Zerfall nicht stattgefunden. An der Änderung des Trends Mitte der 1980er Jahre hat sicher die Entwicklung der modernen Informations- und Kommunikationstechnologien und dabei die des Internets einen wichtigen Anteil. Zu Beginn der 1980er Jahre wurden der Personal Computer eingeführt, und ab der Mitte der 1980er Jahre waren wichtige Dienste des Internets wie ftp und email, zumindest zwischen wissenschaftlichen Einrichtungen, verfügbar. Diese Dienste haben die internationale Zusammenarbeit auf dem Gebiet der Forschung bedeutend bequemer, schneller und billiger werden lassen. Bemerkenswert ist, dass der Trendwandel nicht erst Anfang der 1990er Jahre mit der Entwicklung des World Wide Webs einsetzte.

Andere Faktoren, wie die Zunahme der Kosten der naturwissenschaftlichen und technikkwissenschaftlichen Forschung, die kleinere Länder bei größeren Projekten allein nicht mehr übernehmen können, tragen ebenfalls zu dem beobachteten Trendwandel bei.

#### Entwicklung der mittleren Artikelanteile $p$ ausgewählter Länder

Die nachfolgenden Abbildungen 10-14 dokumentieren anschaulich die Entwicklung der mittleren Artikelanteile  $p$  ausgewählter Länder im Untersuchungszeitraum.

Abbildung 10 zeigt, dass sich der Verlauf der Kurve des mittleren Artikelanteils  $r$  im gesamten SCI von dem Verlauf der Kurve des mittleren Artikelanteils  $p$  der USA kaum unterscheidet. 2002 liegt  $p$  bei den USA bei 0,88 bzw. 88 % und ist damit bedeutend höher als der der vier ökonomisch starken europäischen Länder Frankreich, Deutschland, Italien und Großbritannien, deren Entwicklungen von  $p$  in Abbildung 11 dargestellt sind. Die Kurven Frankreichs und Deutschlands verlaufen fast identisch. Sie erreichen im Jahre 2002 ein  $p$  von etwa 75 % und damit einen mittleren Anteil internationaler Kooperation  $q$  von etwa 25 %. Bei den skandinavischen und den ausgewählten osteuropäischen Ländern, deren Ergebnisse in den Abbildungen 12 und 13 dargestellt sind, werden im Jahre 2002 noch kleinere Werte für  $p$  erreicht. Dort spielt also die internationale Kooperation eine noch größere Rolle. Offensichtlich gibt es einen Zusammenhang zwischen der ökonomischen Stärke eines Landes und seinem mittleren Artikelanteil. Der Korrelationskoeffizient nach Pearson für den Zusammenhang zwischen den Bruttoinlandsprodukten und den mittleren Artikelanteilen  $p$  der 50 am stärksten vertretenden Länder im SCI im Jahre 2002 beträgt 0,42.

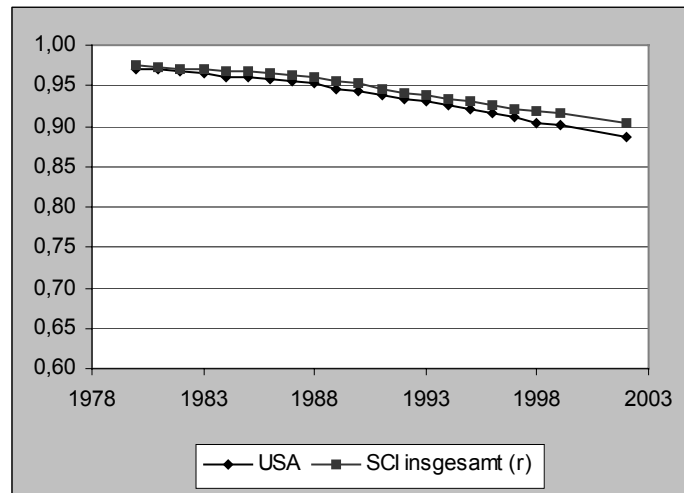


Abb. 10: Entwicklung des mittleren Artikelanteils  $p$  der USA und des mittleren Artikelanteils  $r$  im gesamten SCI

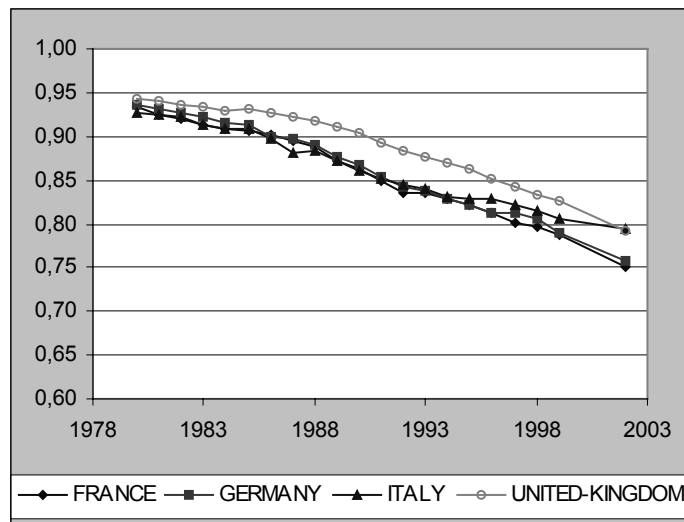


Abb. 11: Entwicklung der mittleren Artikelanteile  $p$  ausgewählter, ökonomisch starker europäischer Länder



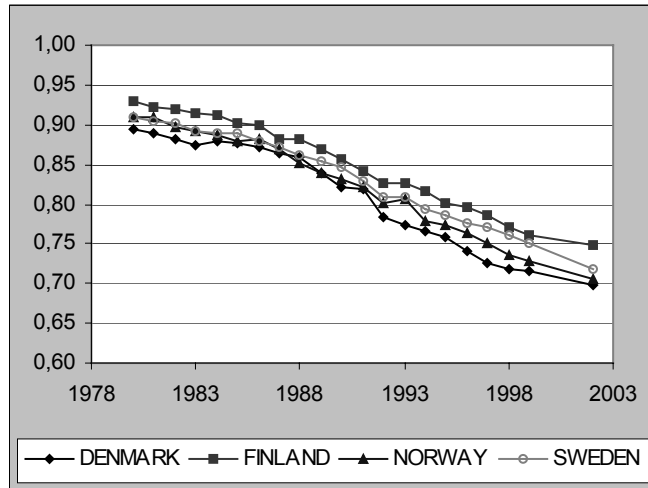


Abb. 12: Entwicklung der mittleren Artikelanteile  $p$  der skandinavischen Länder

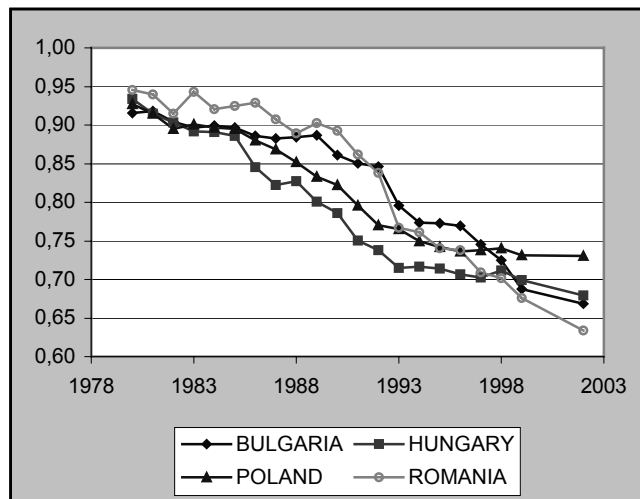


Abb. 13: Entwicklung der mittleren Artikelanteile  $p$  ausgewählter osteuropäischer Länder

Die folgende Abbildung 14 zeigt die Entwicklung von  $p$  bei bedeutenden asiatischen Ländern. Die Kurven für Japan und Indien verlaufen fast identisch und enden 2002 mit einem  $p$ -Wert, der dem der USA etwa entspricht. Die Kurvenverläufe Südkoreas und Chinas unterscheiden sich stark von denen der anderen Länder. Bei China haben wir eine teilweise starke Abnahme von  $p$  in den Jahren 1980-1990 zu verzeichnen. Von 1990 an ist  $p$  und damit auch  $q$  relativ gleich geblieben. Bei Südkorea haben wir ein Ansteigen von  $p$  bis etwa 1985 zu verzeichnen, und von diesem Jahr an liegt  $p$  etwa bei 86 %. Die  $p$ -Werte der vier asiatischen Länder haben sich im Verlauf der Untersuchungsperiode stark angenähert. Sie sind aber viel größer als die Werte, die für die Länder in Europa gefunden wurden. Dies spricht für eine geringere internationale Zusammenarbeit in der Forschung in den asiatischen Ländern gegenüber den europäischen.

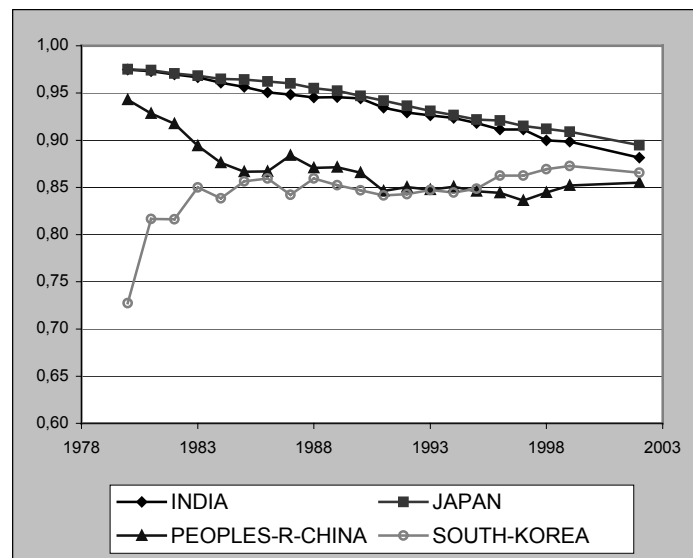


Abb. 14: Entwicklung der mittleren Artikelanteile  $p$  ausgewählter asiatischer Länder

### Danksagung

An dieser Stelle möchte ich mich für viele interessante wissenschaftliche Diskussionen mit Professor Dr. Walther Umstätter in den letzten anderthalb

Jahrzehnten herzlich bedanken. Häufig ging es um Fragen der Bibliometrie, Informetrie oder Scientometrie, und nicht selten spielten in diesen Diskussionen die Shannonsche Entropie und natürlich der Science Citation Index eine wesentliche Rolle. Sie fanden nicht unmittelbar in Vorbereitung dieser Arbeit statt, haben diese aber mittelbar sicher beeinflusst.

Für diverse Hinweise zum Gelingen der vorliegenden Arbeit bin ich Herrn Dr. Frank Havemann dankbar.

### Literatur und Internetquellen

- 1 AJIFERUKE, L., BURREL, Q. & TAGUE, J. (1988). Collaborative Coefficient - A single Measure of the Degree of Collaboration in Research. *Scientometrics*, 14; 421-433
- 2 EGGHE, L. (2005). *Power Laws in the Information Production Process: Lotkaian Informetrics*. Amsterdam: Elsevier
- 3 EGGHE, L., ROUSSEAU, R. & VAN HOOYDONK, G. (2000). Methods for accrediting publications to authors or countries: Consequences for evaluation studies. *Journal of the American Society for Information Science*, 51, 145-157.
- 4 FISCHER, L. D. (2006). *Internationalität der Unternehmung: aktueller Forschungsstand, Analyse und Konzeptualisierung*. Bamberg, Univ. Diss. 2006, urn:nbn:de:bvb:473-opus-803.
- 5 HAVEMANN, F. (2001). Bibliometrische Analyse von Kooperation und Produktivität biomedizinischer Forscher im Jahrfünft 1980-84 und vierzehn Jahre später. *Wissenschaftsforschung Jahrbuch 2000*. Hrsg. H. Parthey et al., Gesellschaft für Wissenschaftsforschung, 121-131.
- 6 KING, D. A. (2004). The scientific impact of nations. *Nature*, 439, 311-316.
- 7 OSAREH, F. (2006). Giant collaboration in astronomy knowledge production at international level. In *Proceedings of International Workshop on Webometrics, Informetrics and Scientometrics & Seventh COLLNET Meeting*. 10-12 05 2006 (pp. 273-280).
- 8 PRICE, D. DE SOLLA (1981). Letter to the editor. *Science*, 212, 987.
- 9 THEIL, H. (1967). *Economics and Information Theory*. Amsterdam: North-Holland.