

Beispielhafte Problemlagen: Unsicherheiten in der Klima- und Umweltforschung

Andreas Oschlies (Kiel)

Bewertung von Modellqualität und Unsicherheiten in der Klimamodellierung

Abstract: The chapter discusses sources of uncertainties in climate models and their possible impacts on the model results. The three criteria “adequacy”, “consistency” and “representativeness” are suggested for a comprehensive assessment of the quality of climate models. The fit to data determines the model’s representativeness. For many climate variables, such as precipitation, cloudiness and the climate sensitivity, this has not significantly improved from the second-to-last to the last assessment report of the Intergovernmental Panel on Climate Change (IPCC). However, the level of detailed mechanistic descriptions has increased for a number of processes included in the models, yielding an improved adequacy of these models. Still, with current climate models being still unable to consistently reproduce glacial cycles driven only by orbital parameters, and with the amplitude of climate change expected until the end of the century being of similar amplitude as glacial-interglacial changes, there is still considerable uncertainty regarding how reliable current models’ projections of 21st century climate change can be. However, uncertainty must not hinder society to make informed decisions, and it is the responsibility of climate research to provide relevant information regarding the uncertainty of climate model projections.

Keywords: Erdklima – Erdsystemmodelle – Klimamodelle – CO₂-Emissionen – parametrische Unsicherheiten – Adäquatheit – Konsistenz – Repräsentativität

1 Einleitung

Ein Ziel der Klimawissenschaften ist es, ausgehend von bisher beobachteten Klimazuständen und ihren räumlichen und zeitlichen Veränderungen, ein Verständnis derjenigen Prozesse zu gewinnen, die für die Entwicklung des Erdklimas relevant sind. Von besonderer gesellschaftlicher Bedeutung ist dabei die Frage, wie das Klima auf anthropogene Eingriffe reagiert, vor allem auf die Emissionen von Kohlendioxid (CO₂). Diese steigen trotz aller politischen Absichtserklärungen, den Klimawandel begrenzen zu wollen, bisher praktisch ungebremst weiter an und erhöhen damit täglich das Risiko für erhebliche Klimaveränderungen mit vermutlich weitreichenden Auswirkungen auf Natur und Gesellschaft.

Eine in der gesellschaftlichen und klimapolitischen Diskussion etablierte Kenngröße zur Beschreibung des Klimawandels ist die *globale Mitteltemperatur*, die in der Regel die jährlich und räumlich gemittelte Lufttemperatur am Boden

über Land und im Wasser an der Meeresoberfläche bezeichnet. Während sie im Alltag in der Regel keine unmittelbare Relevanz hat, sind viele klimatisch und gesellschaftlich bedeutsamere Klimafaktoren wie Niederschlag, die Höhe des Meeresspiegels oder die Intensität von Extremwetterereignissen mit ihr korreliert. So nimmt die globale Mitteltemperatur als leicht verständliche, messbare und damit nachprüfbar und geographisch neutrale Kenngröße eine sinnvolle Rolle in der Klimadebatte ein.

Die globale Mitteltemperatur des Planeten kann seit ungefähr 150 Jahren aus direkten Messungen der Lufttemperatur an der Erdoberfläche recht zuverlässig bestimmt werden. Über diesen Zeitraum zeigt sich bis heute eine Erwärmung von knapp einem Grad Celsius. Projektionen in die Zukunft lassen bei unverändertem CO₂-Emissionsverhalten eine weitere Erwärmung von einigen Grad bis zum Ende des Jahrhunderts erwarten. Diese Erwärmung kann man zu den Temperaturunterschieden zwischen Eis- und Warmzeiten in Beziehung setzen, die etwa 4 bis 5 Grad Celsius betragen und überall auf der Erde gravierende Veränderungen der Umweltbedingungen zur Folge hatten. Der heutige Erwärmungstrend wird überlagert von starken zwischenjährlichen Schwankungen von mehreren Zehntel Grad, die eine Folge der natürlichen Klimavariabilität sind und vor allem durch Schwankungen in der Wärmeaufnahme des Ozeans verursacht werden (z. B. El Niño). Der absolute Wert der globalen Mitteltemperatur hängt stark davon ab, wie genau diese Größe im Einzelfall definiert wird (z. B. dem Höhenrelief folgend an der Erd- oder Eisschildoberfläche oder auf Meeresspiegelniveau; über den Weltmeeren als Luft- oder Oberflächenwassertemperatur). Werden jedoch die Veränderungen der globalen Mitteltemperaturen im zeitlichen Verlauf der letzten Jahrzehnte betrachtet, stimmen diese zwischen verschiedenen, voneinander unabhängigen Analysen unterschiedlicher Forschergruppen sehr genau überein. Die aus dem Grad der Übereinstimmung abgeschätzten Unsicherheiten in den bisherigen Änderungen der jährlichen Mitteltemperaturen betragen typischerweise weniger als ein Zehntel Grad und ergeben sich vor allem aus der räumlich und zeitlich variierenden Datenabdeckung mit nach wie vor großen Beobachtungslücken vor allem über dem Ozean und auf der Südhemisphäre (vgl. Morice et al. 2012).

Die wesentliche Ursache der bisher im 20. und 21. Jahrhundert beobachteten globalen Erwärmung sind anthropogene Emissionen von CO₂ (vgl. aktueller Sachstandsbericht des Weltklimarats, IPCC 2013). Emissionen anderer Treibhausgase (insbesondere Methan), natürliche Klimaschwankungen und Variationen in der Sonnenaktivität spielen für die beobachtete Erwärmung nur eine kleinere Rolle. Weitere CO₂-Emissionen werden nach allem Wissen zu einer weiteren Erderwär-

mung führen. Geht man von einem steigenden globalen Wirtschaftswachstum aus und dem damit im historischen Kontext stets bedingten Anstieg fossiler Energieerzeugung, dann ist sogar eine Beschleunigung der aktuellen Erwärmungstendenz zu erwarten.

Um konkretere Aussagen darüber zu erhalten, wie sich bei gegebenem Emissionsverhalten das Klima in der Zukunft verändern wird, werden *numerische Klimamodelle* verwendet, in denen, ausgehend von physikalischen Grundgesetzen, das Verhalten von Ozean, Atmosphäre sowie Meer- und Landeis beschrieben wird. Die grundlegenden Gleichungen sind dabei sehr gut bekannt (vgl. Heavens et al. 2013). Sie beruhen auf der klassischen Mechanik und Thermodynamik und sind in einer Vielzahl von Experimenten und Beobachtungen empirisch hervorragend bestätigt. Nicht bekannt ist jedoch, wie diese klassischen Gleichungen exakt gelöst werden können, sobald verschiedene Komponenten auf unterschiedlichen Raum- und Zeitskalen miteinander in Wechselwirkung treten. Aufgrund der fehlenden analytischen Lösung der vollständigen Gleichungen müssen in Klimamodellen numerische Lösungsverfahren angewandt werden, die alleine schon wegen der endlichen Rechenkapazität immer nur eine endliche Zahl von Punkten in Raum und Zeit erfassen können und damit nie das komplette Spektrum aller in der Natur auftretenden Raum- und Zeitskalen abbilden können. Die Effekte der nicht aufgelösten Prozesse (z. B. Wolkenbildung, Vermischung durch kleinskalige turbulente Wirbel in Luft und Wasser) werden durch sogenannte Schließungsansätze näherungsweise beschrieben. Bei der praktischen Anwendung der grundlegenden Gleichungen in Klimamodellen werden außerdem Prozesse herausgefiltert, die in der Natur zwar auftreten, für das betrachtete Problem aber als unwichtig erachtet werden (für Klimavorhersagen sind das z. B. Schallwellen).

Modelle, die neben rein physikalischen Prozessen auch den Kohlenstoffkreislauf und gegebenenfalls weitere Stoffkreisläufe (insbesondere die von Nährstoffen wie Stickstoff, Phosphor, Eisen) berücksichtigen und damit auch die Landvegetation sowie die Chemie und Biologie des Meeres behandeln, werden *Erdsystemmodelle* genannt. Ein konzeptuell wesentlicher Unterschied zu den zuvor skizzierten physikalischen Klimamodellen besteht darin, dass vor allem für die Beschreibung biologischer und ökologischer Prozesse keine etablierten oder aus einfachen Annahmen ableitbaren Gesetze oder Gleichungen bekannt sind. Für die Beschreibung von Ökosystemen ist nicht einmal die Wahl der beschreibenden Kategorien eindeutig und reicht von klassischen Aufteilungen in Pflanzen und Tiere (und ggf. Bakterien, Viren usw.) über Artengruppen mit gleichen Stoffwechselfunktionen hin zu genetischen Merkmalen. Es ist derzeit

unklar, welche Kategorisierung am besten geeignet ist, um den Austausch klimarelevanter Stoffe auch unter sich ändernden Umweltbedingungen und für sich ständig anpassende und evolutionär weiterentwickelnde Arten und Ökosysteme zuverlässig zu beschreiben. Zusätzlich zu den Komponenten des Ökosystems werden Gleichungen benötigt, die die Stoffflüsse zwischen den Komponenten beschreiben (u. a. Photosynthese, Nahrungsaufnahme, Verluste durch Fraß und Tod). Abgesehen von Masseerhaltung gibt es dabei keine allgemeingültigen ökologischen Grundgleichungen, so dass verschiedene Modelliergruppen auf unterschiedliche empirisch-pragmatische Ansätze zurückgreifen. Ein systematischer Vergleich der unterschiedlichen Modelle und ihrer Unsicherheiten ist schon aufgrund der unterschiedlichen Modellstruktur und unterschiedlichen Anzahl justierbarer Modellparameter schwierig. Auch wenn unterschiedliche Modelle die vorhandenen Beobachtungsdaten ähnlich gut bzw. schlecht reproduzieren können (vgl. Kriest et al. 2012; Kriest 2017), können sich dieselben Modelle sehr unterschiedlich verhalten, wenn sie auf bisher nicht beobachtete Klimazustände angewandt werden (vgl. Löptien/Dietze 2017).

Eine weitere Ursache für Fehler und Unsicherheiten in den Ergebnissen von Erdsystem- und Klimamodellen ergibt sich aus der *numerischen Implementierung* der Gleichungen. Numerische Verfahren liefern im Allgemeinen lediglich Näherungslösungen der analytischen Gleichungen. Deren Güte hängt neben den angewandten Schließungsansätzen (siehe oben) auch von der räumlichen und zeitlichen Gitterweite des Modells und den verwendeten numerischen Diskretisierungsverfahren ab. Zur Erklärung: Da das Modell Variablen und ihre zeitlichen Änderungen nur für jeden Gitterpunkt des diskreten Modellgitters ausrechnet (d. h., ein „Pixelbild“ der unbekanntem richtigen Lösung der Gleichungen produziert), müssen die kontinuierlichen Gleichungen so umgeformt werden, dass auch sie nur Werte an diesen Gitterpunkten benötigen und die Variablen zwischen den Gitterpunkten interpolieren (was zum Beispiel für die Berechnung von raum-/zeitlichen Ableitungen oder der Varianz von Variablen ein Problem ist). Diese Projektion der kontinuierlichen Gleichungen auf das Gitterraster des Modells nennt man Diskretisierung. Diskretisierungsverfahren variieren für verschiedene Klimamodelle (und in der Regel sogar für verschiedene Komponenten eines Modells), unter anderem weil sich Modellentwickler auf verschiedene Prozesse oder Diskretisierungsansätze fokussieren, aber auch aufgrund der jeweils zur Verfügung stehenden Rechnerleistung, die entweder in eine möglichst feine aber rechenintensive Auflösung des Gitterrasters oder aber in möglichst genaue und ebenfalls rechenintensive Diskretisierungsverfahren investiert werden kann.

Selbst bei identischen Annahmen über die zukünftigen CO₂-Emissionen liefern verschiedene Klima- und Erdsystemmodelle daher unterschiedliche Ergebnisse der simulierten Entwicklung des Erdklimas. Um der Verschiedenheit der Modellergebnisse Rechnung zu tragen, werden z. B. im Sachstandsbericht des Weltklimarats (IPCC 2013) die unterschiedlichen Modelltrajektorien (z. B. der zeitliche Verlauf der globalen Mitteltemperatur) oftmals gemittelt und die Modellstreuung retrospektiv als Unsicherheitsbereich angegeben (vgl. Knutti/Sedlacek 2013). Es ist bisher aber keineswegs klar, inwieweit die Modellstreuung tatsächlich ein Maß für die Unsicherheit der individuellen Modelle oder der Modellgesamtheit darstellt. Die Verbesserung der Abschätzung von Modellunsicherheiten ist daher derzeit ein aktives Forschungsfeld in der Klimamodellierung.

2 Qualitätsbewertung

Ein Modell ist immer ein vereinfachtes Abbild der Natur. Dies ist eine Prämisse für wissenschaftliche Anwendungen von Modellen, bei denen Hypothesen getestet werden und die komplexe Realität auf die dafür als wesentlich erachteten Prozesse reduziert wird. Dieser reduktionistische Ansatz gilt eingeschränkt ebenso für reine Vorhersageanwendungen, z. B. in der Wettervorhersage oder bei Crash-Tests von Autos. Auch hier muss immer vereinfacht werden, da eben nicht die Lage jedes Atoms zu jedem Zeitpunkt genau beschrieben werden kann und soll, in einer endlichen Zeit die für den Modellbetreiber relevanten Größen aber trotzdem möglichst gut beschrieben werden sollen. Eine perfekte Übereinstimmung eines Modells mit der Natur kann daher schon aus Prinzip niemals erreicht werden. Tatsächlich sind für individuelle Ergebnisse von Klimamodellen die Abweichungen zwischen Wirklichkeit und Modelllösung (z. B. der Temperatur an einem Ort und Zeitpunkt) in der Regel deutlich größer als die Genauigkeit, mit der wir die Realität beobachten können. Dennoch können dieselben Modelle durchaus in der Lage sein, zeitlich und räumlich ausreichend gemittelte Werte (z. B. die globale Mitteltemperatur) sogar innerhalb der Unsicherheiten der beobachtungsgestützten Abschätzungen zu reproduzieren. Welche Kriterien können wir also für ein „gutes“ Modell anlegen? Im Folgenden schlage ich die drei Kriterien Adäquatheit, Konsistenz und Repräsentativität vor, die für eine Bewertung von Modellen zum Beispiel der Klimaforschung angewandt werden sollten:

1. *Adäquatheit*: Ein adäquates Modell beinhaltet die für die zugrundeliegende Fragestellung relevanten Prozesse (z. B. Strahlung, Wärmetransport, CO₂-Emissionen) und vernachlässigt unwichtige Prozesse (z. B. Schallwellen). Es

beschreibt relevante Zustandsgrößen, wie z. B. Temperatur, Windgeschwindigkeit oder Biomasse, aber eben nicht die Lokalisation jedes Luft- oder DNA-Moleküls. Der Grad der Relevanz einzelner Prozesse für eine Modellaussage kann z. B. über eine Sensitivitätsanalyse bestimmt werden, bei der die Stärke eines Prozesses in verschiedenen Modellsimulationen unterschiedlich stark eingestellt wird. Wie stark sich die Modellergebnisse unterscheiden, zeigt dann, wie sensitiv das Ergebnis gegenüber einer expliziten Darstellung des Prozesses ist. Eine hohe Sensitivität impliziert eine hohe Relevanz. Während die für die Beschreibung der physikalischen Komponenten des Klimasystems verwendeten Zustandsgrößen recht gut etabliert sind, gibt es bisher keinen Konsens zu den Zustandsvariablen oder gar Zustandsgleichungen, mit denen Ökosysteme gut beschrieben werden könnten. Häufig werden historisch gewachsene Einteilungen in ‚Tiere‘, ‚Pflanzen‘ und ‚Bakterien‘ verwendet, wobei möglicherweise wichtige Gruppen wie ‚Viren‘ oder das zumindest im Ozean weit verbreitete Prinzip der Mixotrophie (Fähigkeit einiger Organismen, sowohl autotroph Photosynthese zu betreiben als auch heterotroph vom Abbau organischer Substanz zu leben) ignoriert werden. Wir können daher nicht ausschließen, dass relevante Prozesse in den heutigen Modellen noch nicht enthalten sind und dass insbesondere die bisher in Erdsystemmodellen verwendeten Ökosystemkomponenten möglicherweise nicht adäquat sind. Adäquat sein muss auch die numerische Implementierung der Gleichungen. Erdsystemmodelle sind häufig über Jahrzehnte gewachsen und bestehen aus vielen Programmpaketen mit insgesamt einigen hunderttausend Zeilen Code, an denen eine Vielzahl von Wissenschaftlern und Wissenschaftlerinnen bzw. Programmierern und Programmiererinnen geschrieben haben. Auch abgesehen von den trotz aller Sorgfalt immer noch unentdeckten Programmierfehlern, sind numerische Verfahren nicht immer optimal und z. B. häufig mitverantwortlich für die Verletzung der Energieerhaltung in heutigen Klimamodellen (siehe unten).

2. *Konsistenz*: Ein konsistentes Modell erfüllt die für eine Beschreibung des zugrundeliegenden Systems als wichtig erachteten Grundprinzipien wie z. B. Massenerhaltung. Interessanterweise sind heutige Klimamodelle generell zwar massen-, aber nicht energieerhaltend, was nicht wirklich befriedigend ist (z. B. Lucarini/Ragone 2011; Eden 2016). Dies liegt zum Teil (aber nicht nur!) an der Parametrisierung¹ turbulenter Vermischung in der

1 Parametrisierungen sind vereinfachte Beschreibungen von Prozessen, die in Modellen nicht vollständig beschrieben werden können.

Atmosphäre und auch im Ozean. Energieerhaltung ist eines der wenigen und erfolgreichen Grundprinzipien der Naturwissenschaften und sollte nicht ohne guten Grund aufgegeben werden. Um ein Modell auf Konsistenz zu prüfen, können neben der Massenerhaltung aber zusätzlich einige idealisierte Testfälle gerechnet werden, für die es analytische Lösungen gibt, z. B. für die Ausbreitung von Wellen. Ebenso geprüft werden kann die Konsistenz mit Vorgängerversionen desselben oder verwandter Modelle sowie zwischen Implementierungen desselben Modells auf unterschiedlichen Rechnern. Diese Konsistenzprüfung ist wichtig, um die Reproduzierbarkeit von Ergebnissen zu gewährleisten, ein wichtiger Pfeiler wissenschaftlicher Praxis. Potenzielle Probleme ergeben sich hier aus aktuellen Beobachtungen, dass Simulationsergebnisse einiger strukturell sehr komplexer, aber durchweg deterministischer Modelle auf parallelen Rechnerarchitekturen nicht immer exakt dieselben Ergebnisse produzieren und sich stärker als durch Rundungsfehler erklärbar unterscheiden, auch wenn das identische, bereits in Maschinsprache übersetzte Programm auf demselben Rechner mehrmals gerechnet wird. Mögliche Ursachen hierfür sind neben bisher unentdeckten Programmierfehlern im Klimamodell, unterschiedliche Zuweisungen von möglicherweise fehlerhaften Prozessoren bei verschiedenen Aufrufen desselben Programms, aber auch Fehler in den Programmen, die die Programmiersprache in Maschinsprache übersetzen.

3. *Repräsentativität*: Dieses Kriterium beinhaltet die Übereinstimmung der Modellergebnisse mit Beobachtungen. Beobachtungen beziehen sich naturgemäß nur auf die Vergangenheit. Da aus der Vergangenheit bisher keine guten Analogien für die zu erwartenden zukünftigen Klimaänderungen bekannt sind, ist eine gute Übereinstimmung mit Beobachtungen nur ein notwendiges, aber kein hinreichendes Kriterium für die Repräsentativität bezüglich der erwarteten zukünftigen Klimazustände. Hier ist zunächst zu klären, gegen welche Beobachtungsgrößen die Modellergebnisse verglichen werden sollen: Jahresmitteltemperatur, Monats-, Tages-, Stundenmittel? Räumliche Mittel? In welcher Höhe/Tiefe? Wie sollen Abweichungen zwischen Modell und Daten gewertet werden? Ist die Standardabweichung ein gutes Maß? Wie sollen Abweichungen von verschiedenen Größen (z. B. Bodentemperatur und Primärproduktion) miteinander verglichen werden? Wie zuverlässig sind die gemessenen Daten? Die Antworten auf diese Fragen bestimmen die Form der Metrik, mit der die Distanz zwischen Modellergebnissen und Beobachtungen gemessen und in eine für die Modellbewertung handliche Anzahl von einem oder wenigen Werten zusammengefasst wird.

Damit ist die Metrik von der jeweiligen wissenschaftlichen Fragestellung abhängig und beinhaltet immer subjektive Elemente, die sich letztlich auch auf die Modellbewertung auswirken.

A priori ist nicht klar, inwieweit aus einer Übereinstimmung von Modellergebnissen mit Beobachtungen aus dem relativ gut beobachteten, klimatisch aber recht stabilen Zeitraum der letzten Jahrzehnte auch auf die Modellqualität bezüglich der erwarteten größeren Klimaänderungen in der Zukunft geschlossen werden kann. Eine Strategie für die Beantwortung dieser Frage besteht in der Verwendung von Daten über große Klimaänderungen in der Erdgeschichte, z. B. den Übergängen zwischen Eis- und Warmzeiten. Solche Daten beruhen in der Regel auf Information aus Klimaarchiven (z. B. Sedimentkerne, Eisbohrkerne, Baumringe, Mineralablagerungen in Höhlen) und stellen häufig nur indirekte Informationen über vergangene Klimazustände dar, was weitere Unsicherheiten in die Modellbewertung induziert.

Alle drei Kriterien werden im Allgemeinen von den einzelnen Modelliergruppen bei der Modellentwicklung, Validation und Verifikation beachtet. Mit Validation wird geprüft, inwieweit ein Modell für den Einsatzzweck geeignet ist, also ob es adäquat und konsistent ist. Bei der Verifikation wird schließlich überprüft, wie genau das Modell die relevanten Aspekte der Realität wiedergeben kann, also wie repräsentativ es ist. Weder für Validation noch für Verifikation gibt es bisher einheitliche Regeln zur Bewertung von Modellqualität. Besonders bei hochgradig komplexen Erdsystemmodellen, die aus einer Vielzahl von gekoppelten Komponenten (z. B. Ozean-, Atmosphären-, terrestrisches Modell) bestehen, ist die Verifikation extrem schwierig. Oft können zwar einzelne Komponenten auf ihre Repräsentativität hin geprüft werden, die Auswirkungen der im nächsten Abschnitt behandelten Unsicherheiten sind für gekoppelte Systeme aber nur sehr schwierig zu kontrollieren und zu bewerten.

3 Unsicherheiten

Einzelne Simulationsergebnisse und damit die Beantwortung der oben diskutierten Frage, in welchem Maß ein Modell repräsentativ ist, hängen von einer Reihe von unsicheren Faktoren ab. Wie bereits erwähnt, gehören dazu die vereinfachte Darstellung von nicht explizit aufgelösten Prozessen (z. B. Turbulenz) sowie die nicht immer genau bekannten Fehler numerischer Diskretisierungsverfahren. Weitere Unsicherheitsquellen resultieren aus der Darstellung von prinzipiell bekannten Faktoren im Erdsystem, wie z. B. der Farbe von Schnee,

Wasser, Wolken, Wüsten oder Vegetation. Diese sind von wesentlicher Bedeutung für den Strahlungshaushalt des Planeten und werden in Klimamodellen häufig mit fest voreingestellten, empirisch gewählten „plausiblen“ Werten beschrieben, obwohl die Farbe jeder dieser Oberflächen in der Realität stark variieren kann. Die Darstellung von im Detail unbekanntem oder nur sehr schwer messbaren Prozessen, wie z. B. der Sterblichkeit von Algen oder des Fressverhaltens von Zooplankton, generiert ebenfalls Unsicherheiten, die sich auf die Modellergebnisse auswirken können. Effekte solcher *parametrischer Unsicherheiten* können durch Sensitivitätsexperimente abgeschätzt werden, in denen einzelne der unsicheren Parameterwerte variiert werden und der Effekt auf die Modellergebnisse untersucht wird. Dies ist prinzipiell möglich, erfordert jedoch eine Vielzahl von oftmals aufwendigen Modellevaluationen und wird z. B. in der Modellierung von marinen Ökosystemen nur selten durchgeführt (vgl. Arhonditsis/Brett 2004).

Unsicherheiten entstehen außerdem durch ungenau bekannte *Anfangsbedingungen*, von denen aus die Modelle gestartet werden. Ein kompletter Klimazustand ist zu keinem Zeitpunkt (z. B. vorindustriell) genau bekannt. In der Regel werden Modelle daher für viele hundert bis tausend Modelljahre gerechnet, um über diese sogenannte Einschwingzeit einen mit der Modelldynamik konsistenten Klimazustand zu entwickeln, der weitgehend unabhängig von den nur ungenau bekannten Anfangsbedingungen ist. Dennoch wirken sich einige unsichere Faktoren wie beispielsweise der anfängliche Nährstoff- oder Kohlenstoffgehalt des Ozeans auch nach vielen tausend Jahren noch wesentlich auf das Erdsystem aus. Unsicherheiten ergeben sich weiterhin aus den angewandten *Randbedingungen*, d. h. aus Auswirkungen von Prozessen, die nicht mehr Teil des Modellsystems sind und daher als externe Einflussfaktoren vorgeschrieben werden müssen. Dies sind in der Regel die Sonnenaktivität, Vulkanismus und natürlich der Einfluss des Menschen, insbesondere in Form von anthropogenen Treibhausgasemissionen und Landnutzungsänderungen.

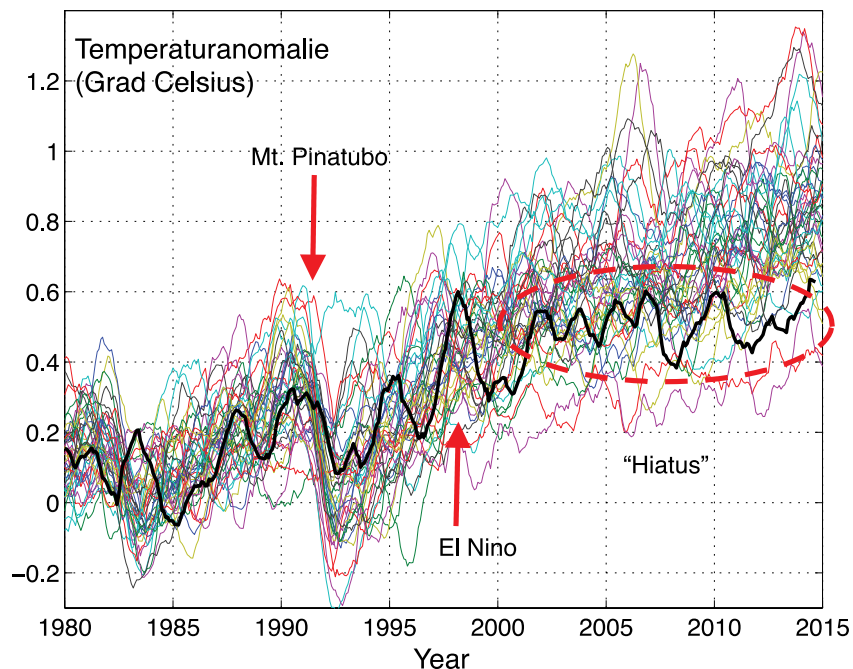
Klima- oder Erdsystemmodelle sind im Allgemeinen deterministisch formuliert, d. h., die Modelle beinhalten keine stochastischen Elemente (obwohl ein Einsatz solcher Elemente in einigen Bereichen vielversprechend erscheint, vgl. z. B. Palmer 2014). Unsicherheiten in den Modellergebnissen sind damit für heutige Modelle immer auf Unsicherheiten in den Modellparametern, Modellgleichungen und ihrer numerischen Lösung sowie in den Annahmen über Randbedingungen zurückzuführen. Unsicherheiten in den Randbedingungen können mit Ensemblesimulationen unter variierten Randbedingungen untersucht

werden, Unsicherheiten in den Modellparametern durch Kalibration gegen ebenfalls mit Unsicherheiten behaftete Beobachtungsdaten (vgl. Schartau et al. 2016). Unsicherheiten in den Modellgleichungen und ihrer numerischen Umsetzung werden selten diskutiert und in der Regel wider besseren Wissens als klein im Vergleich zu den parametrischen Unsicherheiten angenommen.

4 Auswirkung von Unsicherheiten

Ein häufig verwendeter Indikator für Modellunsicherheiten ist die Qualität von modellierten ‚Vorhersagen‘ des heutigen Klimazustands oder bereits vergangener Klimazustände. Dies ist in Abbildung 1 für 42 Simulationen verschiedener Klimamodelle für die Abweichungen der globalen Monatsmitteltemperatur der letzten 35 Jahre relativ zum Zeitraum 1961 bis 1990 abgebildet. Alle Modelle wurden gemäß des *Climate Model Intercomparison Projects (CMIP5)* mit denselben beobachteten Treibhausgaskonzentrationen, Landnutzungsänderungen, Aerosolkonzentrationen und astronomischen Parametern angetrieben (vgl. Taylor et al. 2012). Es zeigt sich, dass die Abweichungen der Modellergebnisse von den tatsächlichen Beobachtungen tendenziell mit dem Abstand zum Normierungszeitraum zunehmen (hier der Zeitraum 1961–1990, dessen Mittelwert in dieser Darstellung für jedes einzelne Modell und für die Beobachtungen abgezogen wurde). Abweichungen zwischen Modellen und Daten lassen sich also nicht alleine durch einen konstanten Temperatur-Offset erklären, sondern werden wesentlich durch unterschiedliche Erwärmungsraten als Antwort auf steigende Treibhausgaskonzentrationen verursacht. Diese Erwärmungsrate ist eine zentrale Größe, für die von Klimamodellen präzise Vorhersagen erwartet werden. Die Abbildung zeigt, dass diese Vorhersagen bereits auf dekadischen Zeitskalen (auf denen viele unsichere Prozesse, wie z. B. Eismassenverluste von polaren Eisschilden, noch gar nicht relevant sind) beträchtliche Unterschiede zwischen den einzelnen Modellen aufweisen. So ist die Streuung der simulierten Temperaturanomalien am Ende des Zeitraums mit ca. 0,8 Grad größer als das beobachtete Erwärmungssignal. Es ist weiter sichtbar, dass die interne zwischenjährliche und dekadische Variabilität in Beobachtungen und Modellrealisierungen einige Zehntel Grad beträgt und damit von derselben Größenordnung ist wie die über den betrachteten 35-Jahreszeitraum gemessene Erwärmung.

Abb. 1: Simulierte globale Mitteltemperatur von 42 Klimamodellen des Climate Model Intercomparison Project 5 (CMIP5, Taylor et al. 2012, für jede Simulation relativ zum Mittelwert des Zeitraums 1961–1990 und mit einem gleitenden Mittelwert über 12 Monate geglättet. Farbige Kurven stellen die verschiedenen Modelllösungen dar, die fettgedruckte schwarze Kurve die Beobachtungsdaten (HadCRUT4, Morice et al. 2012).



Einige Variationen, insbesondere die Abkühlung nach dem Ausbruch des philippinischen Vulkans Pinatubo im Jahr 1991, werden von den meisten Modellen gut wiedergegeben, was im Wesentlichen auf den externen Modellantrieb mit Aerosoleinträgen durch Vulkanismus zurückzuführen ist und Vertrauen in die Darstellung des kurzweiligen Strahlungsantriebs in Klimamodellen schafft. Andere Variationen, wie die beobachtete starke Erwärmung in den Jahren 1996/1997, die auf die interne Klimafuktuation El Niño zurückzuführen ist, werden von den meisten Modellen nicht zeitgleich korrekt wiedergegeben. Sie sind als interne Schwankung des gekoppelten Ozean-Atmosphäre-Systems nicht in den Antriebsdaten enthalten und finden in den Modellen daher nicht unbedingt zu denselben Zeiten statt wie in der Realität. Eine solche Phasenverschiebung kann trotz dynamisch richtiger Beschreibung der Prozesse zu großen lokalen Abwei-

chungen zwischen Modellergebnissen und Beobachtungsdaten führen. Es bleibt eine große Herausforderung an die Modellbewertung, diejenigen Anteile an den Diskrepanzen zwischen Modellergebnissen und Beobachtungen zu identifizieren, die auf Phasenverschiebungen von internen Schwankungen im Klimasystem zurückgeführt werden können und damit nicht notwendigerweise ein Hinweis auf eine fehlerhafte Beschreibung der Klimadynamik sind.

Längerperiodische interne Schwankungen im Klimasystem wurden u. a. auch als Ursache für den sogenannten „Hiatus“, d. h. die Verlangsamung der beobachteten Erwärmung zwischen Mitte der 1990er-Jahre bis 2014, vorgeschlagen (vgl. Schurer et al. 2015; Medhaug et al. 2017). Dass die meisten aktuellen Klimamodelle im Gegensatz zu den Beobachtungen eine ungebremste Fortsetzung des Erwärmungstrends zeigen (Abb. 1; Meehl et al. 2014), könnte demnach auf eine Phasenverschiebung von langperiodischen internen Schwankungen des Klimasystems hinweisen. Solche Schwankungen sind aufgrund des relativ kurzen Beobachtungszeitraums noch nicht komplett verstanden und können damit in Modellbewertungen vermutlich nicht adäquat berücksichtigt werden. Neuere Analysen von Beobachtungsdatensätzen deuten auf eine verstärkte Wärmeaufnahme des tiefen Ozeans als wichtiges Element der verlangsamten Erwärmung an der Erdoberfläche hin (vgl. Drijfhout et al. 2014; Gleckler et al. 2016). Diese verstärkte Wärmeaufnahme wird offensichtlich von den meisten heutigen Modellen zumindest nicht zur richtigen Zeit richtig wiedergegeben. Die Frage, ob es sich hierbei um systematische Modellfehler oder um eine Phasenverschiebung einer ansonsten richtig simulierten internen Variabilität im Klimasystem handelt, ist noch nicht abschließend beantwortet. Ein besseres Verständnis dieser Modell-Daten-Diskrepanz ist für die Bewertung von Modellqualitäten und typischen Unsicherheiten, besonders im Hinblick auf die weitere Entwicklung der globalen Erwärmung, von zentraler Bedeutung.

Während die Modellfehler für Simulationen des heutigen Klimas durch den Vergleich mit vorhandenen Beobachtungen identifiziert und sogar quantifiziert werden können, ist eine Abschätzung der Unsicherheiten der Projektionen in die Zukunft sehr viel schwieriger. In vielen Studien, u. a. in den Sachstandsberichten des Weltklimarats (IPCC 2013), wird die Streuung der Zukunftsprojektionen verschiedener Modelle als Maß der Unsicherheit verwendet. Eine Konvergenz unterschiedlicher Modelle ist jedoch nicht notwendigerweise gleichzusetzen mit einer Reduktion der Unsicherheiten. Konvergenz von Simulationsergebnissen kann z. B. auch durch pragmatisches Publikationsverhalten (die Veröffentlichung von Ausreißern ist oft schwieriger als die Veröffentlichung von Arbeiten, die frühere Ergebnisse bestätigen) oder durch Monopolisierung der Modelllandschaft

entstehen. Robustere Kriterien für die Güte von Modellen unter möglichen zukünftigen Klimabedingungen sollten sich aus dem Abgleich von Simulationen vergangener Klimazustände mit der Information aus geologischen Klimaarchiven entwickeln lassen. Dies erfordert allerdings eine sorgfältige Analyse der Beobachtungsdaten und ihrer Aussagekraft hinsichtlich vergangener Klimazustände (z. B. unter Aspekten wie Jahresmittel oder saisonale Mittel? Extremwerte? Wie lokal bzw. regional?), was selbst für die vergangene Eiszeit vor wenigen zehntausend Jahren schwierig ist und für weiter zurückliegende Klimaereignisse mit zunehmender zeitlicher Distanz immer problematischer wird. Dennoch bieten die großen Klimaschwankungen der Vergangenheit eine einzigartige Möglichkeit, die heutigen Modelle unter Bedingungen zu testen, die den erwarteten zukünftigen Änderungen besser entsprechen als der direkt beobachtete Zeitraum der letzten Jahrzehnte. Eine Simulation des letzten glazialen Maximums vor 21 000 Jahren durch heutige Modelle zeigt z. B. eine systematische Überschätzung der simulierten Temperaturabnahme in den Tropen und eine systematische Unterschätzung der Temperaturabnahme im Nordatlantik und für den europäischen Kontinent (vgl. IPCC 2013). Dies lässt auf systematische Modellfehler schließen, die möglicherweise mit dafür verantwortlich sind, dass bis heute kein dynamisch konsistentes und alleine durch Schwankungen der Erdbahnparameter angetriebenes Erdsystemmodell den Wechsel von Eis- und Warmzeiten in einer realistischen Amplitude beschreiben kann.

5 Schlussfolgerungen

Aus dem Vergleich des vierten und des fünften Sachstandsberichts des Weltklimarats (IPCC 2007, 2013) wird deutlich, dass sich die Abweichungen zwischen Modellergebnissen und Beobachtungen im Mittel für viele Größen wie die Oberflächentemperatur verringert haben. Die Übereinstimmung für andere wichtige Klimavariablen, insbesondere Niederschlag, Bewölkung und die Sensitivität des Klimas gegenüber Änderungen im atmosphärischen CO₂-Gehalt, hat sich dagegen nicht wesentlich verbessert (vgl. IPCC 2013). Ebenso hat sich die Streuung der simulierten Vorhersagen der globalen regionalen Mitteltemperatur und Niederschlagsänderungen für das 21. Jahrhundert nicht verringert (vgl. Knutti/Sedlacek 2013). Während sich aus den Modellergebnissen alleine damit keine Verbesserung der *Repräsentativität* der Modelle und der Genauigkeit von Prognosen ableiten lässt und aus Anwendersicht damit keine Verbesserung in der Qualität der Prognosen sichtbar ist, sind viele Prozesse in moderneren Modellen viel detailgetreuer aufgelöst. Diese Modelle haben daher mehr Freiheitsgrade, sind schwieriger zu kalibrieren und werden in der Regel auch eine stärkere in-

terne Variabilität aufweisen, die für viele Modelle im Vergleich zu den Beobachtungen immer noch zu gering ist. Eine detailliertere Prozessbeschreibung bei gleichbleibender Streuung der Modellvorhersagen kann jedoch als eine bessere Darstellung des Prozessverständnisses und damit eine Verbesserung bezüglich *Adäquatheit* und ggf. *Konsistenz* gedeutet werden. Trotz unveränderter Übereinstimmung mit Beobachtungen und der Modelle untereinander hat damit die Qualität der Modelle aus Modellierersicht zugenommen. Umgekehrt muss aber ein besseres Prozessverständnis nicht automatisch verringerte Unsicherheiten zur Folge haben.

Da gesellschaftliche Entscheidungen über die weiteren Wege der Klimapolitik wesentlich von den Simulationsergebnissen von Erdsystemmodellen beeinflusst werden, ist eine Bewertung der Modelle und ihrer Unsicherheiten dringend erforderlich. Bisher gibt es dafür keine einheitlichen Bewertungsmetriken, und die Modellbewertung läuft in der Regel nach bestem Wissen und Gewissen der jeweiligen Modellentwickler und Anwender ab. Die Sachstandsberichte des Weltklimarats stellen eine hervorragende Übersicht über die Qualität der heutigen Modelle bezüglich der Simulation vergangener und aktueller Klimazustände dar. Für viele, aber nicht für alle Klimavariablen werden die Abweichungen der Simulationsergebnisse von den Beobachtungen dabei mit der Zeit kleiner. Modelle simulieren jedoch immer noch eine im Allgemeinen zu geringe interne Klimavariabilität, und es gibt bisher kein dynamisch konsistentes und alleine durch Schwankungen der Erdbahnparameter angetriebenes Modell, das Übergänge zwischen Eis- und Warmzeiten simulieren könnte. Dies wirft Zweifel bezüglich der Zuverlässigkeit auf, mit der heutige Modelle die bereits für dieses Jahrhundert erwarteten Klimaänderungen simulieren können, die in ihrer Amplitude vermutlich ähnlich groß sein werden wie die Unterschiede zur letzten Eiszeit. Letztlich sind diese Modelle jedoch das beste Werkzeug, das wir für Projektionen in die Zukunft haben. Sie enthalten den besten Stand des Wissens und sind im wissenschaftlichen Wettbewerb einer ständigen Überprüfung ausgesetzt. Auch wenn weitere Modellverbesserungen, einhergehend mit verbesserten Methoden zur Modellbewertung, das Vertrauen in Modelle weiter stärken sollten, können und müssen politische und gesellschaftliche Entscheidungen natürlich auch unter den aktuellen Unsicherheiten getroffen werden. Entscheiden unter Unsicherheiten ist auch in anderen Lebensbereichen alltäglich, es sollte daher auch in der Klimapolitik gelingen können. Eine wesentliche Herausforderung an die Klimaforschung sehe ich darin, diese Unsicherheiten besser zu beschreiben, abzuschätzen und für die Allgemeinheit zugänglich darzustellen.

Literatur

- Arhonditsis, George B./Brett, Michael T. (2004): Evaluation of the current state of mechanistic aquatic biogeochemical modeling. *Marine Ecology Progress Series* 271, 13–26.
- Drijfhout, Sybren S./Blaker, Adam T./Josey, Simon A./Nurser, A. J. George/Sinha, Bablu/Balmaseda, Magdalena A. (2014): Surface warming hiatus caused by increased heat uptake across multiple ocean basins. In: *Geophysical Research Letters* 41.22, 7868–7874.
- Eden, Carsten (2016): Closing the energy cycle in an ocean model. In: *Ocean Modelling* 101, 30–42.
- Gleckler, Peter J./Durack, Paul J./Stouffer, Ronald J./Johnson, Gregory C./Forest, Chris E. (2016): Industrial-era global ocean heat uptake doubles in recent decades. In: *Nature Climate Change* 6.4, 394–398.
- Heavens, Nicholas G./Ward, Daniel S./Mahowald, Natalie M. (2013): Studying and Projecting Climate Change with Earth System Models. In: *Nature Education Knowledge* 4.5, 4.
- IPCC (2007): Solomon, Susan/Qin, Dahe/Manning, Martin/Chen, Zhenlin/Marquis, M./Averyt, K. B./Tignor, Melinda M. B./Miller, H. L. (Hrsg.): *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge.
- IPCC (2013): Summary for Policymakers. In: Stocker, Thomas F./Qin, Dahe/Plattner, Gian-Kaspe/Tignor, Melinda M. B./Allen, Simon K./Boschung, Judith/Nauels, Alexander/Xia, Yu/Bex, Vincent Pauline M. (Hrsg.): *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge/New York.
- Kriest, Iris (2017): Calibration of a simple and a complex model of global marine biogeochemistry. In: *Biogeosciences* 14.21, 4965–4984.
- Kriest, Iris/Oschlies, Andreas/Khatriwala, Samar (2012): Sensitivity analysis of simple global marine biogeochemical models. In: *Global Biogeochemistry Cycles* 26, GB2029, DOI: 10.1029/2011GB004072.
- Knutti, Reto/Sedlacek, Jan (2013): Robustness and uncertainties in the new CMIP5 climate model projections. In: *Nature Climate Change* 3.4, 369–373.
- Löptien Ulrike/Dietze, Heiner (2017): Effects of parameter indeterminacy in pelagic biogeochemical modules of earth system models on projections into a warming future: The scale of the problem. In: *Global Biogeochemistry Cycles* 31.7, 1155–1172.

- Lucarini, Valerio/Ragone, Francesco (2011): Energetics of climate models: Net energy balance and meridional enthalpy transport. In: *Reviews of Geophysics* 49.1, DOI: 10.1029/2009RG000323.
- Medhaug, Iselin/Stolpe, Martin B./Fischer, Erich M./Knutti, Reto (2017): Reconciling controversies about the 'global warming hiatus'. In: *Nature* 545, 41–47.
- Meehl, Gerald A./Teng, Hayian/Arblaster, Julie M. (2014): Climate model simulations of the observed early-2000s hiatus of global warming. In: *Nature Climate Change* 4.10, 898–902.
- Morice, Colin P./Kennedy, John J./Rayner, Nick A./Jones, Phil D. (2012): Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 dataset. In: *Journal of Geophysical Research* 117, DOI: 10.1029/2011JD017187.
- Palmer, Tim N. (2014): More reliable forecasts with less precise computations: a fast-track route to cloud-resolved weather and climate simulators? In: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 372 (2018).
- Schartau, Markus/Wallhead, Philip/Hemmings, John/Löptien, Ulrike/Kriest, Iris/Krishna, Shubham/Ward, Ben A./Slawig, Thomas/Oschlies, Andreas (2016): Reviews and syntheses: parameter identification in marine planktonic ecosystem modelling. In: *Biogeosciences Discussions*, DOI: 10.5194/bg-2016-242.
- Schurer, Andrew P./Hegerl, Gabriele C./Obrochta, Stephen P. (2015): Determining the likelihood of pauses and surges in global warming. In: *Geophysical Research Letters* 42.14, 5974–5982.
- Taylor, Karl E./Stouffer, Ronald J./Meehl, Gerald A. (2012): An overview of CMIP5 and the experiment design. In: *Bulletin of the American Meteorological Society* 93.4, 485–498.

Hermann Held (Hamburg)

Der ökonomische Wert von Klimainformation: Zur Neuinterpretation von Klimazielen unter antizipiertem Lernen

Abstract: How much would a rational decision maker be willing to invest in the reduction of uncertainty of climate projections by an order of magnitude? This seemingly technical question requires shedding some light on the foundations of the two leading schools of thought within climate economics: cost benefit and cost effectiveness analysis. While the former takes off from the most solid axiomatic basis, its results are currently not robust regarding some hard to determine input parameters. The latter operationalizes politically decided environmental targets that can be interpreted as an expression of strong sustainability. The 2°C target is of that sort. However under anticipated future learning fundamental conceptual difficulties appear within an interpretation as a hard limit. We offer a new, softer interpretation of the 2°C target that avoids these difficulties. Among other advantages of this new interpretation the question about the expected economic value of the reduction of climate response uncertainty regarding greenhouse gas emissions becomes a well-posed one. Perfect learning in that regard could on average save up to hundreds of billions of Euros per year if a stringent 2°C policy were pursued. It remains to be shown whether our softer interpretation is the only interpretation of a target that would be consistent with learning, or whether a third way between the traditional “hard” and our “soft” interpretation were possible.

Keywords: Unsicherheit – Risiko – Standardansatz – Vorsorgeansatz – Kosten-Nutzen-Analyse – Kosten-Effektivitäts-Analyse – Kosten-Risiko-Analyse – Klimapolitik

1 Einführung

Welchen ökonomischen Wert hätte es, die Unsicherheit von Klimaprojektionen zu verringern? Mit dieser Frage verbindet sich zunächst eine praktische Bedeutung, denn mit ihrer Beantwortung hängt eine zweite Frage eng zusammen, in welchem Umfang Klimawissenschaften vielleicht beschleunigt zu fördern wären, falls die Gesellschaft sich ernst gemeinte Klimaziele setzt. Dass der Eliminierung von Unsicherheit überhaupt ein ökonomischer Wert zuzuweisen ist, liegt letztlich daran, dass bessere Information über die Zukunft der Gesellschaft eine bessere Planbarkeit ermöglicht und sie sich dann nur noch gegenüber einem engeren Spektrum von Zukünften abzusichern braucht.

Hinter dieser technokratisch wirkenden Eingangsfrage verbergen sich jedoch bislang ungeklärte Grundsatzfragen zur Bewertung von Klimapolitik-Optionen unter heterogener Unsicherheit. Unter *Unsicherheit* verstehen wir hier die Wissensdifferenz zu perfektem Wissen (insbesondere über die Konsequenzen unserer Entscheidungen) und folgen hierin der Definition des Weltklimarates IPCC (*International Panel for Climate Change*) (vgl. Mastrandrea et al. 2010): In Vorbereitung auf seinen fünften Sachstandsbericht (2013/2014) hatte sich der IPCC erstmalig dem Anspruch unterzogen, sich auf einen für alle seine drei Arbeitsgruppen verbindlichen Umgang mit Unsicherheit zu verständigen. In Bezug auf den Begriff der Unsicherheit selbst musste er sich im Wesentlichen zwischen zwei Begriffstraditionen entscheiden – derjenigen der Klimawissenschaft und derjenigen der Ökonomie. Die Klimawissenschaft versteht unter *Unsicherheit* das Komplement des Wissens – also alles dasjenige, was im jeweiligen Kontext an wünschbarem Wissen fehlt. Probabilistisch formulierbares Wissen stellt hierbei nur eine von mehreren Unsicherheitskategorien dar. Hingegen grenzt ein Teil der ökonomischen Community seit Knight (1921) *Unsicherheit* von *Risiko* ab: *Risiko* bezeichnet hier die Möglichkeit, Handlungsfolgen-Kategorien vollständig und probabilistisch gewichtet prognostizieren zu können. *Unsicherheit* bezeichnet demgegenüber die Situation, keine solchen probabilistischen Gewichte oder noch nicht einmal alle Handlungsfolgen-Kategorien angeben zu können, markiert gewissermaßen „weicherer“ Wissen. Der IPCC ist der Tradition der Klimawissenschaft gefolgt (vgl. Mastrandrea et al. 2010) und wir schließen uns dem hier an: *Unsicherheit* stellt daher einen Sammelbegriff dar für alles, was wir in einem bestimmten Kontext nicht wissen, aber gern wissen würden. *Heterogene Unsicherheit* möge hier den Umstand bezeichnen, dass wir derzeit gewisse Subsysteme des gekoppelten Systems aus Klima und Gesellschaft mit qualitativ größerer Unsicherheit als andere Subsysteme zu beschreiben vermögen.

Im Folgenden wird ausgeführt, dass die beiden derzeit prominentesten und seit zwei Jahrzehnten betriebenen Schulen der Klimaökonomik grundlegende konzeptionelle Schwierigkeiten damit haben, unter unserer derzeitigen Wissens- und zugehörigen Unsicherheitsstruktur des Klimaproblems robuste Empfehlungen zu liefern. Aus diesem Grunde schlugen wir vor einigen Jahren einen Hybridansatz beider Schulen vor (vgl. Schmidt et al. 2011). Dieser erlaubt es u. a. erstmalig, den ökonomischen Wert der Verringerung von Unsicherheit unter einer 2°C-Politik als gut gestellte Frage zu formulieren und damit auch zu beantworten. Daher sollen nun zunächst die beiden Standardansätze der Klimaökonomik vorgestellt und in ihren Konsequenzen für Klimapolitik-Empfehlungen ausgeleuchtet werden.

Hierbei wird besonderes Augenmerk auf ihren Umgang mit Unsicherheit gerichtet. Schließlich werden unser Hybridansatz und dessen Konsequenzen erläutert.

Entlang dieser Sequenz wird auf Qualitätskriterien eingegangen werden, die uns als notwendig oder zumindest wünschenswert erscheinen, um beim heutigen Stand des Wissens (natürlich stilisierte) klimapolitische Empfehlungen unter Unsicherheit abzuleiten. Die Liste derartiger Qualitätskriterien liest sich wie folgt:

1. Dass Unsicherheit im betrachteten Problem einen Effekt erster Ordnung darstellt, legt nahe, Unsicherheit auch explizit bei klimapolitischen Entscheidungen zu berücksichtigen.
2. Da unstrittig ist, dass derzeit keine Obergrenze für die Sensitivität des Klimasystems gegenüber anthropogenen Treibhausgas-Antrieben angegeben werden kann, darf ein Entscheidungskriterium auch nicht rein numerisch-pragmatisch eine derartige Obergrenze als Hilfsgröße annehmen.
3. Angesichts andauernder Beobachtung des Klimasystems sowie fortgesetzt verfeinerter Beschreibung naturräumlicher Prozesse, die diese Sensitivität festlegen, ist davon auszugehen, dass wir kontinuierlich über die Sensitivität des Klimasystems dazulernen. Die Meinungen gehen darüber auseinander, wie zügig dieser Lernprozess voranschreiten kann. Wir halten es jedoch für wünschenswert, denjenigen, die eine bestimmte Erwartungshaltung hinsichtlich signifikanten Dazulernens einnehmen, ein Angebot zu machen – ein Angebot in Gestalt eines Entscheidungskalküls, das mit der Vorstellung eines Dazulernens über die Sensitivität des Klimasystems nicht strukturell überfordert ist.
4. Die Meinungen gehen weit darüber auseinander, bis zu welchem Grade die Klimawandelfolgen in ihrer Gesamtheit derzeit überhaupt darstellbar und dann auch noch monetarisierbar sind. Uns erscheint es als wünschenswert, denjenigen, die eine derartige Monetarisierung beim derzeitigen Kenntnisstand als so schwierig erachten, dass jedweder Unsicherheits-Formalismus überfordert wäre, ein Entscheidungskalkül anbieten zu können, das ohne eine derartige Monetarisierung auskommt.
5. Ferner ist als „Meta-Kriterium“ zu fordern, dass sämtliche Eingangsgrößen, die ein Entscheidungskalkül benötigt, auch tatsächlich und robust angegeben werden können.
6. Schließlich wird man dem „Meta-Kriterium“ zustimmen müssen, dass ein Kalkül eine Lösung zu liefern hat, andernfalls ist es schlicht nutzlos für eine Gesellschaft.

Diese Liste wünschbarer Eigenschaften von Entscheidungskalkülen ist selbstverständlich nicht vollständig. Etwa fehlt der Wunsch, dass sämtliche „Win-win-

Optionen“ abgeschöpft sein mögen, dass also z. B. nicht nutzlos „Geld verbrannt“ werde. Jedoch möchten wir hier auf diejenigen Eigenschaften fokussieren, hinsichtlich derer sich die prominenten klimaökonomischen Schulen in Bezug auf ihren Umgang mit heterogener Unsicherheit unterscheiden.

2 Unsichere Folgen des Klimawandels: Zwei Begründungsstränge für Klimaschutz

Die beiden oben erwähnten klimaökonomischen Schulen wurzeln in je unterschiedlichen Auffassungen darüber, welche Gründe für klimapolitische Eingriffe akzeptiert werden. Im Folgenden wird dieses an der Klimapolitik-Option ‚Vermeidung‘ (im Gegensatz zu den Optionen ‚Anpassung‘ und ‚Climate Engineering‘) illustriert, weil sich hierzu der klimaökonomische Diskurs als am ausgereiftesten darstellt und wir fragen können, was sich hiervon für die Beurteilung von Anpassungs- und *Climate-Engineering*-Optionen nutzen ließe.

Der Standardansatz der Umweltökonomie zielt darauf, das Kabinett sämtlicher Klimawandelfolgen auszuleuchten, diese zu listen und zu monetarisieren. Sollte sich herausstellen, dass ein sich selbst überlassener Klimawandel in der Summe schädlicher wäre als die Dekarbonisierung des Energiesystems, könnte dieses den „rationalen“ Grund für eine Vermeidungspolitik liefern.

Liest man hingegen den gegenwärtigen Wissensatlas der Klimafolgenforschung so, dass jenseits gewisser Erwärmungswerte die Folgen noch nicht näherungsweise absehbar, angebbar oder gar monetarisierbar sind, könnte dieses ein Argument dafür liefern, eine entsprechende Erwärmung zumindest solange nicht zuzulassen, wie sich die Wissensbasis entsprechend mager darstellt. Dieses würde eine Umsetzung des Vorsorgeprinzips bedeuten. Gemäß diesem zweiten und entgegen dem ersten Ansatz würde sich eine Empfehlung für eine Vermeidungspolitik gerade aus einem *Fehlen an Wissen* ableiten. (Hierbei sollen keineswegs die Verdienste der Klimafolgenforschung geschmälert werden. Ein großer Anteil der klimawissenschaftlichen Zunft hegt jedoch Zweifel daran, dass die Totalität der Klimawandelfolgen für beliebige Erwärmungen zurzeit näherungsweise monetarisierbar sei, aller bislang erfolgten Anstrengungen zum Trotz.) Jedem dieser beiden Ansätze entspricht nun eine „ökonomische Schule“.

2.1 Der umweltökonomische Standardansatz

Aus dem Standardansatz folgt, im Verein mit weiteren Annahmen, die Kosten-Nutzen-Analyse der Umweltökonomik, das axiomatisch am besten begründbare Entscheidungs-Kalkül (vgl. Kunreuther et al. 2014; Savage 1954; von Neumann/

Morgenstern 1944). Das bedeutet, dass dieses Entscheidungskalkül auf wenige, scheinbar unmittelbar einleuchtende Annahmen zurückgeführt werden kann. Aufwendungen zur Umrüstung des Energiesystems, die bereits heute beginnen könnte, werden mit dadurch künftig vermiedenen Schäden verrechnet. Hierbei bezieht sich „künftig“ auf die dem Klimasystem inhärente Antwort-Zeitskala von 50 bis 1000 Jahren. Hieraus folgt sofort, dass die Frage, wieviel in den kommenden Jahrzehnten in Vermeidungspolitik zu investieren ist, wesentlich von der ökonomischen Diskontierung der Zukunft abhängt (d. h. dem Ausmaß, in dem die Zukunft gegenüber der Gegenwart abgewertet wird; auf eine derartige Abwertung läuft auch der Zins hinaus). Ein seit den 1990er-Jahren häufig in Varianten gefundenes Ergebnis der Kosten-Nutzen-Analyse (siehe z. B. Nordhaus 2008) ist, dass eine Erwärmung um 3,5°C gegenüber dem vorindustriellen Niveau wohlfahrtsoptimal ist und in diesem Zuge in den kommenden drei Jahrzehnten im Wesentlichen die globalen Emissionen einem *Business-as-usual*-Szenario folgen könnten. Insofern wären gegenüber den bereits erfolgten Minderungsverpflichtungen der letzten *Conference of the Parties* (2015) keinerlei weitere Anstrengungen erforderlich, die Opferung des 2°C-Ziels wäre volkswirtschaftlich rational.

Kritiker haben auf die hohe Sensitivität dieser Empfehlung gegenüber Zusatzannahmen unter Unsicherheit hingewiesen (vgl. Kolstad et al. 2014; Nelson 2013; Stern 2013). Am wohl stärksten hat jedoch Weitzman (2009) die aktuelle Verwendbarkeit der Kosten-Nutzen-Analyse des Klimaproblems für Politikberatung in Frage gestellt: Er legte dar, dass eine konsequente Übernahme der Klimasensitivitäts-Wahrscheinlichkeitsverteilung (d. h. des Übertragungsfaktors von Treibhausgaskonzentration auf Temperatur) in eine Kosten-Nutzen-Analyse unter möglichen extremen Annahmen bezüglich Klimawandelfolgen zu einem „unendlich großen“ Klimawandelschaden führen könne. Auf der Linie dieses Ansatzes liegt dann, dass wir eigentlich sofort eine maximal mögliche Vermeidungspolitik umsetzen müssten. (Zwar verwiesen Horowitz und Lange [2014] Weitzmans Ansatz auf einen engeren Gültigkeitsbereich, doch bleibt die Validität des Arguments als solche bestehen.) Man kann nun argumentieren, dieses sei das Beste, was die ökonomische Zunft derzeit liefern könne – die getroffenen Annahmen müssten eben Stakeholdern gegenüber transparent und die je folgende Empfehlung, die von nahezu „keiner“ bis zu „totaler“ sofortiger Vermeidungsanstrengung reichen kann, klar kommuniziert werden. Dieses ist eine Haltung, die durchaus nachvollziehbar eingenommen werden kann.

Man könnte jedoch auch dafür plädieren, dass eine Kosten-Nutzen-Analyse derzeit die klimawissenschaftliche Community in Bezug auf ihre Eingangs-

datenseite überfordert. Die Unsicherheit der Klimawandelfolgen und ihrer Bewertung ist derzeit zu groß, um robuste Ergebnisse zu ermöglichen. Vermehrte Forschungsanstrengungen in entsprechender Richtung könnten hier schrittweise Abhilfe schaffen, jedoch besteht bereits heute gesellschaftlicher Bedarf an wissenschaftlicher Entscheidungs-Unterstützung, die Annahmen in greifbarer Form als im Fall der Kosten-Nutzen-Analyse von Stakeholdern verlangt. Wegen des Budget-Effekts von Treibhausgasen (d. h. der Erkenntnis, dass einem Temperaturziel approximativ ein bestimmtes Budget an Kohlendioxid oder Kohlendioxidäquivalenten, über die Zeit summiert, entspricht) schließt sich das Fenster für das 2°C-Ziel in den kommenden zehn bis zwanzig Jahren (vgl. Meinshausen et al. 2009). Zudem wird durch die notwendige Renovierung des OECD-Kraftwerksparks und den Aufbau eines Energiesystems in den Schwellenländern eine Klimapolitik dieser Akteure in den kommenden zwei Jahrzehnten wesentlich festgelegt. Sollte die Kosten-Nutzen-Analyse eines Tages adäquat operationalisierbar sein, könnte es für die Gestaltung des realweltlichen Klimaproblems daher bereits zu spät sein.

In Bezug auf unseren eingangs angegebenen Forderungskatalog können wir zunächst festhalten, dass dieses Standardverfahren die Kriterien 1–3 erfüllt, was wir hier nicht im Einzelnen begründen können (siehe z. B. Gollier 2001). Kriterium 4 kann und will qua Konstruktion nicht erfüllt werden. Das Meta-Kriterium 5 ist jedoch aus unserer Sicht bis auf weiteres verletzt. Das Standardverfahren fordert Informationen, die derzeit nur unzureichend zur Verfügung stehen, und seine Ergebnisse reagieren nahezu beliebig auf schwer bestimmbare Eingangsgrößen. Schließlich wird das Meta-Kriterium 6 qua Konstruktion erfüllt, weil dem Standardverfahren eine unbeschränkte Optimierung zugrunde liegt, zu der stets eine formale Lösung gefunden werden kann.

Aus diesem Grunde vertritt ein Teil der klimaökonomischen Community eine Temperaturziel-gebundene *Kosten-Effektivitäts-Analyse* als notwendigerweise weniger objektives – minimal weniger objektives! –, jedoch kurzfristig verfügbares Entscheidungsinstrument (vgl. Kunreuther et al. 2014).

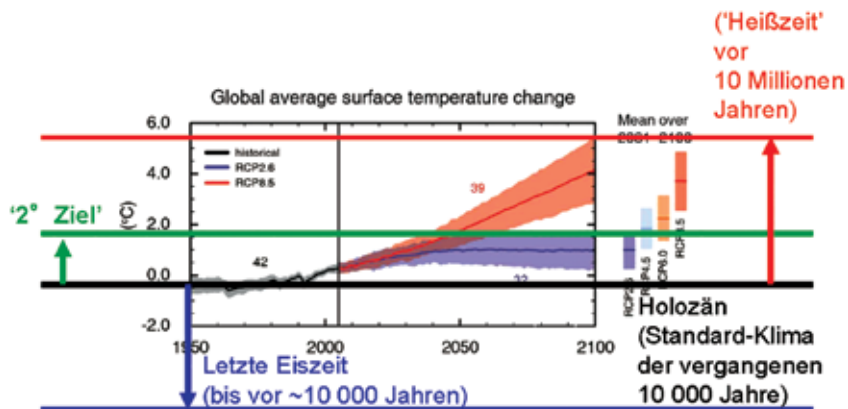
2.2 Eine mögliche Umsetzung des Vorsorgeprinzips: Die Kosten-Effektivitäts-Analyse des 2°C-Ziels

Wie ließe sich nun das Vorsorgeprinzip am Klimaproblem operationalisieren? Man kann fragen, welche natürliche Schwankungsbreite die globale Mitteltemperatur während der Entwicklung der Menschheit bereits erfahren hat. In der Tat wurden Erwärmungen gegenüber unserem vorindustriellen Standardklima von 1,5°C verzeichnet (vgl. Schellnhuber 2015: 453). Gesteht man Menschheit

und Natur ein gewisses Anpassungspotenzial zu und fordert zugleich eine für den politischen Prozess einprägsame Zahl, gelangt man so zu 2°C maximal zu erlaubender Erwärmung.

Umgekehrt könnte man auch fragen, welcher Temperaturbereich vielleicht zu meiden sei, weil er qualitativ von dem unseren derart verschieden ist, dass wir ihn unter Vorsorgegesichtspunkten meiden sollten. Fragt man etwa, wie weit man in der Erdgeschichte zurückgehen muss, um denjenigen Temperaturhub wiederzufinden, den wir bis 2100 bei ungebremster Erwärmung erwarten dürfen (bis zu 5°C; vgl. IPCC 2013: Abb. 7), so wären dieses mindestens zehn Millionen Jahre (vgl. Zachos et al. 2001). Außerdem bewegte man sich so in einer Temperaturänderungs-Skala, die derjenigen des Eiszeit-Warmzeit-Übergangs gleicht (vgl. Abb. 1 und Schneider von Deimling et al. 2006a), der einen dramatischen Bruch in Gestalt der natürlichen Rahmenbedingungen, unter denen sich Kultur entwickeln musste, markiert.

Abb. 1: Temperaturverlauf mit und ohne Klimapolitik (adaptiert aus IPCC 2013: Abb. 7), ins Verhältnis gesetzt zum Eiszeit-Holozän-Übergang



In der Eiszeit reichten die Eispanzer bis in die mittleren Breiten und rahmten daher einen qualitativ anderen Lebensraum als heute. Eine Differenz von 5°C bedeutet daher eine „große“ Änderung. Ohne Klimapolitik würde die Menschheit nochmals eine Differenz derselben Größe induzieren, das Klimasystem gewissermaßen in eine „Heißezeit“ katapultieren (vgl. Abb. 1). Dieses wäre ein Temperaturbereich, wie er seit einer Skala von 10 Millionen Jahren nicht aufgetreten ist (vgl. Zachos et al. 2001). Die Einhaltung des 2°C-Ziels würde bedeuten, näher am „Standardklima“ zu verbleiben (dem Klima des Holozäns, das seit 10.000 Jahren

vorherrscht) als am Klima der „Heißzeit“, wengleich 2°C auf dieser Skala auch keine „kleine“ Änderung bedeuten. Das 2°C -Ziel lässt sich daher auf doppelte Weise aus einem Vorsorgegedanken ableiten.

Schließlich liefert eine Festlegung auf 2°C eine für den politischen Prozess hilfreiche Vergrößerung in Gestalt einer einprägsamen, glatten Zahl, analog einer Geschwindigkeitsbeschränkung von 100 km/h. Das 2°C -Ziel will hingegen nicht sagen, dass bei dieser Temperatur ein objektiver „kritischer Schwellwert“ (etwa eine Bifurkation oder ein Phasenübergang) in der Natur vorliegt. Es handelt sich vielmehr um eine letztlich politisch gesetzte Orientierungsmarke, wengleich eine akademisch gut informierte, die vermutlich das europäische Wertesystem hervorragend abbildet. Auf diese Analogie werden wir noch zurückkommen, wenn wir später unser neues Entscheidungskalkül motivieren werden.

Das ökonomische Entscheidungsinstrument der Kosten-Effektivitäts-Analyse ersetzt nun in diesem Kontext den Versuch, Klimawandelfolgen zu projizieren, durch die Vorschrift, das vordefinierte umweltpolitische Ziel einzuhalten, d. h. im Sinne der COPs seit 2009, eine Erhöhung der globalen Mitteltemperatur gegenüber dem vorindustriellen Niveau auf 2°C zu begrenzen. Diese wichtigste Präferenz wird dann ergänzt durch die nachgeschaltete Vorschrift, unter den so erlaubten Klimapolitikpfaden nach dem Optimum der ökonomischen Wohlfahrt zu fragen. Vereinfacht gesagt, liefert die Kosten-Effektivitäts-Analyse den kostengünstigsten Energiemix, um das 2°C -Ziel gerade noch einzuhalten.

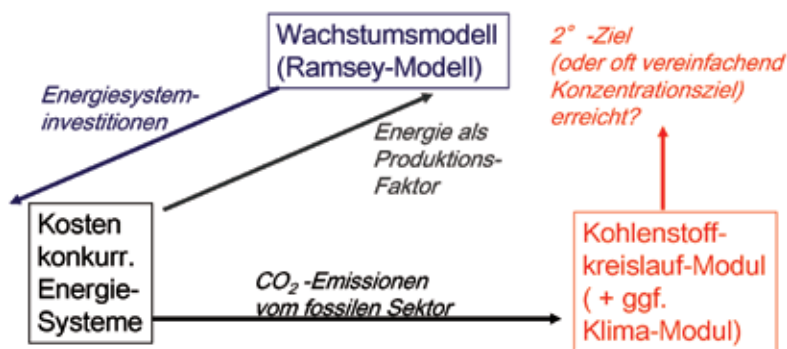
In Bezug auf unseren eingangs angegebenen Forderungskatalog ist festzuhalten, dass das hier zuletzt vorgestellte Verfahren Kategorie 4 qua Konstruktion erfüllt. Genau durch diesen Umstand umgeht es die Instabilitäten des Standardverfahrens in Bezug auf Kategorie 5. Leider werden wir weiter unten finden, dass es in seiner traditionellen Form Schwächen in Bezug auf Kategorien 1–3 aufweist.

2.3 Vergleichende Betrachtung von Standard- und Vorsorgeansatz

Vergleicht man die hinter Kosten-Effektivitäts- und Kosten-Nutzen-Analyse liegenden Ansätze, so könnte man sagen, dass hinter beiden eine je andere Arbeitsteilung zwischen ökonomischem Formalismus und menschlicher Intuition steckt (vgl. Neubersch et al. 2014): Die Kosten-Nutzen-Analyse überantwortet alles dem Formalismus, während in der Kosten-Effektivitäts-Analyse ein Teil des Systems durch die vorgeschaltete Setzung des Klimaziels für die Analyse als nicht handhabbar „weggeschnitten“ wird. Was dort vor sich geht, wird schemenhaft durch erste Indikationen der Klimafolgenforschung wahrgenommen und es wird dann der Intuition „Dahin gehen wir nicht!“ gefolgt. Insofern stellt das 2°C -Ziel auch eine Ausprägung starker Nachhaltigkeit dar (vgl. Hediger 1999).

Je mehr die Unsicherheit über Klimawandelfolgen reduziert wird, desto mehr bietet sich ein Wechsel von instinktbasierendem zu formalisiertem Entscheiden an. Diese sukzessive Verschiebung der Arbeitsteilung entfällt bei der Modellierung der Kosten der Umrüstung des Energiesystems: Dieses menschengemachte System ist qualitativ besser projizierbar als die Klimawandelfolgenkette (vgl. Stern 2007) und wird daher in beiden Entscheidungsinstrumenten formalisiert behandelt.

Abb. 2: Archetypischer Aufbau eines integrierten Modells zur temperaturzielbasierten Kosten-Effektivitäts-Analyse des Klimaproblems.



3 Das numerische Resultat einer Kosten-Effektivitäts-Analyse des 2°C-Ziels

Welchen volkswirtschaftlichen Aufwand bedeutet es nun, das 2°C-Ziel einzuhalten? Um diese Frage zu beantworten, werden Analysen mittels gekoppelter Klima-Energie-Ökonomie-Modelle benötigt, die eine Kosten-Effektivitäts-Analyse als Entscheidungskalkül in sich tragen (vgl. Abb. 2). Klimawandelfolgen werden nicht explizit modelliert, sondern es wird angenommen, dass diese stilisiert durch eine Temperaturobergrenze aufgefangen sind. Ein ökonomischer Kern erklärt globales Wachstum und projiziert das Ausmaß von Wohlfahrts-einbußen, das sich durch ein Aufzwingen eines Temperaturziels ergäbe. Die Ökonomie liefert Investitionen an den Energiesektor, dieser Energie als einen produktionssteigernden Faktor an die Volkswirtschaft zurück. Der Energiesektor ist in diverse Energietechnologien aufgespalten (bei hochauflösenden Modellen werden diese in einer Größenordnung von einhundert Technologien repräsentiert), worunter sich traditionelle fossile, aber auch konkurrierende, zunächst kostspieligere Niedrigemissions-Technologien befinden. Ein Klima-

modul, das die Projektion der Maximaltemperatur komplexer Klimamodelle auf Zehntelgrad genau emuliert, prüft, ob die Maximaltemperatur im Einklang mit dem Klimaziel steht. Ist dieses nicht der Fall, werden Investitionen in den fossilen Sektor zugunsten des Niedrigemissions-Sektors zurückgenommen. So wird intertemporal der kostenminimale Energiemix gefunden, der gerade noch mit dem Temperaturziel verträglich ist.

Im Zuge des letzten IPCC-Berichts wurden ca. 1000 Klima-Energie-Ökonomie-Szenarien, im Modus der Kosten-Effektivitäts-Analyse generiert, ausgewertet und in eine Hand voll „Klimaschutz-Klassen“ eingeteilt (IPCC 2014: 12). Betrachtet man diejenige Klasse, die in etwa dem 2°C-Ziel entspricht, so ergibt sich, dass die Implementierung einer 2°C-Politik die globale Wachstumsrate im Mittel um 0,06 %/Jahr-Punkte absenken würde (vgl. IPCC 2014). Sowohl gegenüber einem erwarteten globalen Wachstum von 1,6 %–3,0 %/Jahr als auch gegenüber der Unsicherheit in dieser Spanne bedeutet dies eine „kleine“ Zahl.

Insofern liegt es für viele Betrachter nahe, dem 2°C-Ziel „geringe“ Kosten zu attestieren. Folgt man dieser Diagnose, stünde nunmehr die Tür offen zur Umsetzung des 2°C-Ziels; die Gesellschaft könnte diese geringe „Versicherungsprämie“ leicht zahlen, um sich gegen schwer abschätzbare Klimawandelfolgen abzusichern (die Analogie zur Versicherungsprämie gilt nicht streng, kann aber als nützliche Heuristik dienen). Insofern könnte man sagen, dass eine Verschiebung der Arbeitsteilung zwischen Formalismus und Intuition das eingangs beschriebene „Kosten-Nutzen-Patt“ aufgelöst hat. Die Tatsache, Kosten des 2°C-Ziels angeben zu können, bedeutet auch, überhaupt eine Lösung dieses Kosten-Effektivitäts-Problems gefunden zu haben. Kriterium 6 ist also numerisch erfüllt. Sollte sich jedoch eine Umsetzung des 2°C-Ziels um Jahrzehnte verzögern, könnte der Fall eintreten, dass es keine Lösung mehr gibt (vgl. IPCC 2014: 15).

Doch ist es wirklich „fair“, allein auf Grund der bislang aufgeführten Argumente den „Punkt“ an die Kosten-Effektivitäts-Analyse, zu Lasten der Kosten-Nutzen-Analyse, zu vergeben? Wir erinnern daran, dass die Schwäche der Kosten-Nutzen-Analyse am eklatantesten durch Weitzmans konsequente Einbeziehung der Unsicherheit über die Klimasystem-Antwort auf Treibhausgas-Antrieb vorgeführt wurde. Sie findet ihren formalen Ausdruck u. a. in der konsequenten Einbeziehung der Wahrscheinlichkeitsverteilung der Klimasensitivität. Wurde diese formalisierbare Unsicherheit in den oben erwähnten 0,06 %/Jahr-Punkten Wachstumseinbuße, abgeleitet aus Kosten-Effektivitäts-Analysen, adäquat berücksichtigt?

3.1 Die probabilistische Verallgemeinerung der Kosten-Effektivitäts-Analyse

Was bedeutet es, das 2°C-Ziel einzuhalten, wenn die Temperaturantwort des Klimasystems unsicher ist? Zunächst hatte die Unsicherheit der Klimasensitivität bei der Formulierung des 2°C-Ziels wenig Beachtung gefunden. Es ist jedoch unmittelbar klar, dass die Klimasensitivität einen großen Einfluss auf das erlaubte Emissionsverhalten haben muss: Je größer die Sensitivität ist, desto stärker reagiert die globale Mitteltemperatur auf steigende Treibhausgas-Konzentrationen und desto weniger darf folglich pro vorgeschriebenem Temperaturziel emittiert werden.

Petschel-Held et al. (1999) und dann weiter ausführend Kriegler/Bruckner (2004) gaben einen analytischen Zusammenhang zwischen einem formulierten Temperaturziel T^* , dem dann erlaubten Emissions-Budget E und der Klimasensitivität γ an: Für E lässt sich eine strenge obere Schranke E^0 angeben, also $E < E^0$, mit $E^0 \propto (2^{T^*/\gamma} - 1)$. Derzeit kann die Klimawissenschaft keine Obergrenze für γ angeben. Wollte sich die Gesellschaft in ihrem Emissionsverhalten auf das gesamte Unsicherheits-Regime von γ vorbereiten, müsste sie sich gegen beliebig große γ wappnen und hätte bereits in der Vergangenheit kein Emissionsbudget aufhäufen dürfen (denn $\lim_{\gamma \rightarrow \infty} E^0 = 0$). Daraus folgt unmittelbar, dass sich jedes noch so laxe Temperaturziel T^* nur mit einer gewissen Wahrscheinlichkeit, aber nicht perfekt, einhalten lässt. Auch genügt es nicht, eine schlichte „Sensitivitätsstudie“ in Bezug auf γ in Kosten-Effektivitäts-Studien durchzuführen, weil nahezu jedes beliebige Emissionsbudget als „optimal“ eingeschätzt werden kann, je nachdem, welches γ verwendet wurde (vgl. Bürgermeier et al. 2006; Held et al. 2009). Insofern ist es erforderlich, die ökonomische Analyse so zu verallgemeinern, dass die Unsicherheit in γ zu einem integralen Bestandteil jedes Entscheidungskalküls erhoben wird.

In Kleinen (2005) wurde das Konzept eines probabilistischen Temperaturziels formuliert. Neben die normative Setzung eines T^* tritt dann noch die geforderte Wahrscheinlichkeit P^* , dieses Ziel auch einzuhalten. Dieses bedeutet zugleich, dass eine Gesellschaft davon ausgeht, mit der Wahrscheinlichkeit $1-P^*$ das Ziel zu verfehlen. Diese ethische Figur ist in unserer Gesellschaft tief verankert, etwa im Vorfeld der Genehmigung großtechnischer Anlagen wie z. B. Kernkraftwerken, bei denen eine ernste Störung nie ganz auszuschließen ist, deren Wahrscheinlichkeit aber als abschätzbar gilt. Diese „Weichspülung“ des 2°C-Ziels in ein nur noch probabilistisches Ziel wurde denn auch von der Öffentlichkeit ohne besondere Aufregung zur Kenntnis genommen. Im Vor-

feld der COP 2009 in Kopenhagen wurde eine „Copenhagen Diagnosis“¹ erstellt, in der das 2°C-Ziel stets gepaart mit einem P^* von 66 % bis 75 % auftrat. Zwar erscheint dieses P^* als ungewöhnlich niedrig, jedoch ist so das Ziel noch erreichbar (siehe Kriterium 6!) und kann daher weiterhin als Leitmotiv für die Klimaverhandlungen dienen.

Die entsprechende Verallgemeinerung einer Kosten-Effektivitäts-Analyse auf probabilistische Klimaziele wird „Chance Constraint Programming“ (CCP; vgl. Charnes/Cooper 1959) genannt. Im IPCC (2014) wurden ca. 1000 energieökonomische Szenarien ausgewertet, die im Modus der Kosten-Effektivitäts-Analyse bestimmt worden waren, also ohne explizite Berücksichtigung von Unsicherheit. Allerdings konnten Held et al. (2009) zeigen, dass sich in derartige Rechnungen näherungsweise eine CCP-Interpretation hineindeuten lässt. In der Tat wurden die nach Konzentrations-Kennziffern klassifizierten Einhaltungen von (T^*, P^*) -Zielen zugeordnet (vgl. IPCC 2014: 10). Jedoch bleibt offen, bis zu welchem Grade die so gefundenen Lösungen dann wohlfahrtsoptimal sind. Zusammenfassend kann daher festgehalten werden, dass im zentralen „0,06 %/Jahr-Punkte“-Ergebnis des letzten Sachstandsberichts des IPCC die Klimaantwort-Unsicherheit bereits gespiegelt und daher das Ergebnis als robust in dieser Hinsicht zu bezeichnen ist. In diesem Sinne sind näherungsweise Kriterien 1–2 erfüllt. Wie verhält es sich jedoch in Bezug auf Kriterium 3?

Schmidt et al. (2011) zeigen, dass in der Formulierung von CCP eine (folgeschwere) stillschweigende Annahme getroffen wurde, zu deren Rechtfertigung divergierende Meinungen existieren, nämlich dass künftiges klimawissenschaftliches Lernen über die Größe der Klimaantwort, insbesondere über die Klimasensitivität γ , vernachlässigbar sei. In der Tat haben sich seit dem Ende der 1970er-Jahre keine einschneidenden Änderungen der Unsicherheits-Einschätzung zu γ ergeben. Andererseits ist nicht auszuschließen, dass durch Assimilierung des Kohlendioxid-Temperatur-Zusammenhangs im letzten Glazial-Interglazial-Übergang qualitativ bedeutsames Lernen der Klimawissenschaft über γ und damit eine langfristige Reduktion der zugehörigen Unsicherheit möglich werden wird (vgl. Schneider von Deimling et al. 2006b). Auch Optionen, subskalige Wolkenprozesse aufzulösen, erscheinen vielversprechend (vgl. Klocke 2011). Insofern erscheint es als naheliegend, vom ökonomischen Entscheidungskalkül zu fordern, dass es auch antizipiertes künftiges Lernen zu verarbeiten, also Kriterium 3 zu entsprechen vermag.

1 Copenhagen Diagnosis, <http://www.copenhagendiagnosis.org>, Fig. 22, Abrufdatum: 08.01.2018.

3.2 Der „Quasi-Weitzman-Effekt“ in der Kosten-Effektivitäts-Traditionslinie

Diese scheinbar harmlose formale Erweiterung von CCP auf antizipiertes Lernen stößt insbesondere am Klimaproblem auf fundamentale konzeptionelle Schwierigkeiten (vgl. Schmidt et al. 2011), deren Grundfigur bereits in der Entscheidungstheorie der 1970er-Jahre bekannt war (vgl. Blau 1974): Während das Standardverfahren jedem künftigen Lernen einen nicht-negativen ökonomischen Wert zuweist, könnte der erwartete ökonomische Wert von Lernen im Falle von CCP negativ sein. Erweitert man etwa das CCP des gekoppelten Klima-Energie-Ökonomie-Modells MIND (vgl. Edenhofer et al. 2005) um antizipiertes Lernen (z. B. in 2030), wird genau dieser Effekt gefunden (vgl. Schmidt et al. 2011). Die Gesellschaft würde demnach sogar noch Mittel dafür aufwenden, gewisse Forschungsfragen *nicht* beantwortet zu bekommen. Dieses ist mit der Werteordnung einer Gesellschaft, die Entscheidungen unter möglichst fundierten, wissenschaftlich gefundenen Rahmenbedingungen treffen möchte, nicht vereinbar. CCP verletzt Kriterium 3.

Noch gravierender ist folgender Effekt: Ein Entscheider sollte sich bereits heute auf die Möglichkeit vorbereiten, in 2030 einen relativ hohen Wert für γ zu erlernen. Um auch in diesem Fall P^* überhaupt noch einhalten zu können, müsste wegen des Budgeteffekts bereits heute ein Übermaß an Emissionsreduktion geleistet werden (vgl. Schmidt et al. 2011), darin Weitzmans Ergebnis für die Kosten-Nutzen-Analyse gleichend (dieser „Quasi-Weitzman-Effekt“ der Kosten-Effektivitäts-Tradition ist bis heute von den Verfechtern von Temperaturzielen nahezu nicht zur Kenntnis genommen worden). Es kann darüber hinaus noch nicht einmal ausgeschlossen werden, einen derart hohen Wert für γ zu lernen, dass das P^* -Ziel auf Grund bereits erfolgter Emissionen nicht mehr einzuhalten ist. Insofern ist CCP als dysfunktional unter antizipiertem Lernen zu bezeichnen, es verletzt sogar das Meta-Kriterium 6.

Ein Hauptproblem scheint darin zu liegen, dass ein starres (T^* , P^*)-Ziel mit einer nicht benennbaren Obergrenze zu γ nicht vereinbar ist, wenn sich die Unsicherheit zu γ verringern könnte. Dieses sind Schwierigkeiten, die in einer Kosten-Nutzen-Analyse qua Konstruktion nicht auftreten können, weil letztere nicht über Ziele verfügt. Auch kann gezeigt werden, dass der erwartete Nutzen von antizipiertem Lernen nie negativ werden kann (siehe etwa Gollier 2001). Daher stellen Schmidt et al. (2011) die Frage, ob sich ein Entscheidungskalkül definieren ließe, das die Vorteile beider Entscheidungsschulen (Kosten-Nutzen vs. Kosten-Effektivität) erntet.

4 Das Hybridmodell „Kosten-Risiko-Analyse“ (KRA)

Schmidt et al. (2011) schlagen vor, das Konstrukt eines probabilistischen Klimaziels beizubehalten, jedoch eine weitere Aufweichung in der Interpretation des 2°C-Ziels vorzunehmen: Das Ziel darf überschossen werden, jedoch wird die Überschießungswahrscheinlichkeit in Kombination mit dem Grade des Übertritts eingepreist (siehe Gleichung 1):

$$\max W = \int_0^T \int_0^\infty p(\gamma) \{U(t) - \beta R(T(t, \gamma))\} e^{-\delta t} d\gamma dt$$

Gleichung (1)

Hierbei indizieren t die Zeit und s die möglichen Zustände, die γ annehmen kann. p_s bezeichnet die Wahrscheinlichkeit eines Zustands für γ , U die konsumgetriebene Nutzenfunktion, R das noch zu spezifizierende „Risiko“, das durch die Überschreitung der Maximaltemperatur empfunden wird, und $\exp(-\delta t)$ die bei intertemporaler Optimierung übliche Diskontierung. U repräsentiert den rein ökonomischen erfassten Anteil, in dem sich auch die Kosten der Transformation des Energiesystems spiegeln. R stellt das „empfundene Risiko“ der Überschreitung der 2°C-Obergrenze dar.

Formal handelt es sich bei (1) um ein Kosten-Nutzen-Funktional. Daher können die bei CCP für antizipiertes Lernen gefundenen konzeptionellen Schwierigkeiten nicht auftreten.

Wurde nun „durch die Hintertür“ wieder eine Schadensfunktion (in Gestalt von R) eingeführt, die lediglich anders bezeichnet wird? Dieses ist nicht der Fall. R bezeichnet eine Zahlungsbereitschaft, eine Temperaturüberschreitung zu vermeiden. Der tatsächlich eintretende Schaden braucht hierfür nicht gewusst zu werden. Wüsste man ihn, würde man eine Standard-Kosten-Nutzen-Analyse ausführen. Im Prinzip könnten nun R und sein Verrechnungsparameter β durch Interviews mit Entscheidern bestimmt werden. In Neubersch et al. (2014) haben wir jedoch der Community eine Interpretation der COP-2010-Formulierung des 2°C-Ziels (wiederum eines probabilistischen Zieles) in Gestalt einer Kosten-Risiko-Analyse (KRA) vorgeschlagen. Nach unserem Eindruck hat in den Klimaverhandlungen die Vorstellung, über γ dazuzulernen, kaum eine Rolle gespielt. Es war lediglich von „likely compliance“ mit dem 2°C-Ziel die Rede. Im „calibrated language“-Verständnis des IPCC (Mastrandrea et al. 2010) entspricht dies einem P^* von 66 %. Neubersch et al. (2014) kalibrierten nun den Verrechnungsparameter β derart, dass im Grenzfall ohne Lernen mit 66 % Wahrscheinlichkeit dem 2°C-Ziel entsprochen wurde. Alle übrigen Variablen wurden dann aus dem so

kalibrierten Modell diagnostiziert, insbesondere die Investitionsströme in die konkurrierenden Energietechnologien.

Wir betonen, dass dieses Entscheidungskalkül qua Konstruktion sämtlichen eingangs geforderten Qualitätskriterien genügt!

Doch mit welcher Funktion R wurde gearbeitet? Um auch hierfür einen stilisierten Vorschlag zu unterbreiten, wurde das Axiom der „Nicht-Opferung eines holozänahen Klimazustands“ (Neubersch et al. 2014) proklamiert. Es wurde gefordert, dass bei imaginiertem perfekten Lernen keinesfalls der Fall eintreten dürfe, wenn sich (etwa wegen zu großen γ 's) das 2°C-Ziel nicht mehr einhalten ließe, dann auf einen Standard-Emissionspfad ohne Klimapolitik einzuschwenken, „weil ja ohnehin alles egal ist“. Wir behaupten, dass dieses Axiom in der Tat die Werteordnung der Proponenten des 2°C-Ziels wiedergibt. Entscheidend ist hierbei, dass, wie vorne ausgeführt, dieses Ziel keinen scharfen Übergang zum „Weltuntergang“ kennzeichnet, sondern eine (akademisch informierte) politisch gesetzte Orientierungsmarke. Neubersch et al. (2014) zeigen dann, dass die Funktion $R(T)$ konvex und daher jenseits des Ziels mindestens linear ansteigen muss, um dem Axiom zu genügen. Eine lineare Risikofunktion stellt daher die konservativstmögliche Darstellung (aus Sicht einer an Emissionen gewöhnten Gesellschaft, d. h. im Mittel ihr Minimum) einer Zahlungsbereitschaft dar.

Mit dieser Funktion wurde in den wenigen bisher existierenden Anwendungen dieses neuen Entscheidungs-Kalküls gearbeitet. Die entsprechenden numerischen Ergebnisse seien hier kurz zusammengestellt:

- Die Investitionspfade aus KRA und CCP unterscheiden sich nicht signifikant für die kommenden Jahrzehnte, auch dann nicht, wenn antizipiertes Lernen hinzugenommen wird (vgl. Neubersch et al. 2014). Der „Quasi-Weitzman-Effekt“ von CCP ist entschärft. Dieses ist eine sehr gute Nachricht für all diejenigen, die zu den 1000 Szenarien des letzten IPCC-Berichts beigetragen haben. Vermutlich wird man zeigen können, dass auch bei der Verwendung komplexerer Modelle (wie dies in Neubersch et al. 2014 in Gestalt des MIND-Modells der Fall war) KRA und Kosten-Effektivitäts-Analyse sehr ähnliche Ergebnisse liefern. Insofern lassen sich die im letzten IPCC-Bericht präsentierten Rechnungen als Grenzfälle ohne Lernen interpretieren. Die zugehörigen Vermeidungskosten können als Obergrenzen der Kosten mit Lernen interpretiert werden, weil der erwartete Nutzen von Lernen in KRA qua Konstruktion nie negativ sein kann.
- Durch KRA ist es erstmalig möglich, den Erwartungsnutzen von klimawissenschaftlicher Information (hier stilisiert im Lernen zu γ vereinigt) konzeptionell sinnvoll zu formulieren. Es zeigt sich, dass sich durch heutiges perfektes

Lernen² ein Wohlfahrtsgewinn ergeben würde, der im zeitlichen Mittel bis zu 0,7 % Konsumgewinn entspräche. Zudem könnte im Mittel bis zu einem Drittel an Vermeidungskosten eingespart werden. Dieses bedeutet global hunderte von Milliarden Euro pro Jahr, falls man eine 2°C-Politik unterstützt. Es darf bezweifelt werden, dass der Forschung zur Reduktion von Klimaantwort-Unsicherheit derzeit Mittel in diesem Umfang zufließen.

- Die weichere Form des 2°C-Ziels in Gestalt der KRA erlaubt es, das Wertesystem, das hinter dem Ziel steht, in eine Zukunft zu extrapolieren, in der das Ziel vielleicht nicht mehr exakt einzuhalten sein wird, etwa durch eine weiterhin dem Ziel inadäquate Vermeidungspolitik. Roth et al. (2015) bestimmten die Vermeidungskosten unter verzögerter Klimapolitik. Sie sinken mit dem Grad der Verzögerung unter KRA, während sie unter einer Kosten-Effektivitäts-Analyse ansteigen (vgl. Luderer et al. 2013).

5 Zusammenfassung und Ausblick

Die Klimaökonomik muss mit Unsicherheiten sehr verschiedener Qualität umgehen. Hierbei verstehen wir, dem IPCC folgend, „Unsicherheit“ als die Gesamtheit des in einem spezifischen Kontext fehlenden Wissens, inklusive eines solchen, das sich durch Wahrscheinlichkeitsmaße ausdrücken lässt. Während die Unsicherheit der Klimaantwort auf Treibhausgasemissionen als gut formal handhabbar gilt, ist im Fall von Klimawandelfolgen noch nicht einmal klar, ob wir derzeit den relevanten Zustandsraum hinreichend erfassen, d. h., ob wir bereits alle wichtigen Größen kennen und ins Kalkül einbeziehen. In den vergangenen zwei Jahrzehnten haben sich unterschiedliche entscheidungstheoretische Ansätze dazu herausgebildet, welche klimapolitischen Handlungen in dieser Situation zu empfehlen seien. Die Kosten-Nutzen-Analyse und die Kosten-Effektivitäts-Analyse stellen hierbei die am häufigsten verwendeten Entscheidungskalküle dar. Dabei entspricht die Kosten-Nutzen-Analyse dem „klassischen“ Vorgehen der Umweltökonomik und sie kann auf die breiteste axiomatische Basis verweisen. Klimaziele sind darin Ergebnisse einer Abwägung aller zur Verfügung stehenden ökonomischen Informationen. Unsicherheiten werden stringent probabilistisch modelliert. Es hat sich gezeigt, dass das Ausmaß optimaler heutiger Vermeidungsanstrengungen so stark von schwer bestimmbareren Eingangsgrößen abhängt, dass nahezu das gesamte Spektrum möglicher Empfehlungen durch dieses Instrument in der Literatur gerechtfertigt wurde. Während der Schwerpunkt der Empfehlungen jenseits von

2 D. h. der Fiktion, heute „über Nacht“ perfektes Wissen über die Temperaturentwicklung in Reaktion auf unser Emissionsverhalten zu erlangen.

3°C Erwärmung zu finden war, eröffnete eine konsequentere Buchhaltung der Unsicherheit zur Klimasensitivität einen Pfad, sogar stärkere Vermeidungspolitiken als diejenige eines 2°C-Ziels zu rechtfertigen. Nach unserem Eindruck benötigt dieses formal fundierteste Entscheidungsinstrument im Lichte eines kombinierten Effekts aus Unsicherheit von Klimasensitivität und Klimawandelfolgen noch einen deutlich längeren Vorlauf an Grundlagenforschung, bevor mit seiner Hilfe robuste Politikempfehlungen abgeleitet werden können.

Hingegen benötigt die Kosten-Effektivitäts-Analyse ein Klimaziel – wie etwa das 2°C-Ziel – bereits auf der Eingabeseite des Kalküls. Es legt einen akademisch informierten, doch auch stark auf die Intuition zurückgreifenden Schnitt durch das System im Sinne dessen, was angesichts von „großer“ Klimawandelunsicherheit noch handhabbar bzw. möglicherweise nicht mehr handhabbar sei – sowohl akademisch als auch gesellschaftlich-praktisch. Dieses Ziel ersetzt dann eine explizit darzustellende Klimawandel-Folgen-Kostenfunktion der Kosten-Nutzen-Analyse. Die akademische Analyse kann sich dann auf diejenigen Teile des Systems beschränken, zu denen sie valide Aussagen zu treffen vermag: zum Zusammenhang von Treibhausgas-Emission und globaler Mitteltemperatur sowie zu den Kosten von Vermeidungspolitik. Auch diese Aussagen sind mit Unsicherheiten behaftet, die in derselben Größenordnung liegen wie die projizierten Effekte selbst; diese Unsicherheiten stellen daher „Effekte erster Ordnung“ dar. Doch hier ist das Systemverständnis ausreichend entwickelt, um Unsicherheiten zu formalisieren und ins Entscheidungskalkül der Kosten-Effektivitäts-Analyse oder in dessen Nachfolge-Kalküle einzubeziehen.

Für diese Einbeziehung musste die Kosten-Effektivitäts-Analyse in einem Doppelschritt verallgemeinert werden: Zum einen war dem Umstand Rechnung zu tragen, dass bislang für die Klimasensitivität keine Obergrenze angegeben werden kann. Daher muss das 2°C-Ziel als ein probabilistisches Ziel interpretiert werden („mit 66 % Wahrscheinlichkeit einzuhalten“). Jedoch läuft man bei der Vorstellung, eines Tages könnten sich derart hohe Sensitivitäts-Werte konsolidieren, dass dann diese Grenze nicht mehr einzuhalten wäre, in Selbstwidersprüche. Es wäre rational, sich bereits heute darauf vorzubereiten. Dann muss jedoch darüber verhandelt werden, wie eine Überschreitung der 2°C-Grenze zu bewerten ist.

Nachdem nun beide erwähnten klimaökonomischen Schulen angesichts von Unsicherheit gravierende praktische bzw. konzeptionelle Schwierigkeiten aufweisen, wenn es darum geht, der *heutigen* Gesellschaft Empfehlungen zu liefern, haben wir aus beiden Ansätzen einen Hybrid konstruiert und diesen *Kosten-Risiko-Analyse (KRA)* genannt. Die Figur des Temperaturziels wird beibehalten,

eine Überschreitung wird jedoch verrechenbar gemacht. Der neue Verrechnungsparameter wird am 2°C-Ziel der COP 2010 geeicht. Alle anderen Systemgrößen lassen sich dann daraus ableiten, wie etwa die Investitionen in den Energiesektor, der Nutzen klimawissenschaftlicher Information oder das Verhalten bei verzögerter Klimapolitik.

Es zeigt sich, dass wesentliche Empfehlungen traditioneller Kosten-Effektivitäts-Analysen, wie sie etwa im letzten IPCC-Bericht kompiliert wurden, durch KRA reproduziert wurden. Dieses bedeutet, dass es möglich ist, bisherigen Kosten-Effektivitäts-Analysen eine neue Interpretation zuzuweisen, so dass deren Ergebnisse gegen „Lernen über Klimasensitivitäts-Unsicherheit“ „robustifiziert“ werden. Unterschiede treten hingegen auf, wenn Klimapolitik verzögert umgesetzt wird. Dann würde eine Kosten-Risiko-Analyse in ihrer konservativsten Ausprägung (mit linearer Risikofunktion) weniger Vermeidung empfehlen als eine Kosten-Effektivitäts-Analyse. Künftige Arbeiten haben auch zu zeigen, ob es Alternativen gibt, einen Hybrid aus Kosten-Nutzen- und Kosten-Effektivitäts-Analyse zu bilden, oder ob die Kosten-Risiko-Analyse „alternativlos“ ist. Letztere Frage mag sich Vertretern starker Nachhaltigkeit stellen, die eine Verrechnung von Kosten und Risiken grundsätzlich ablehnen.

Diese Diskussion hat unmittelbare Konsequenzen für eine Beurteilung darüber, ob die Climate Engineering-Option „direkte Beeinflussung des Strahlungshaushalts“ im Verbund mit Vermeidungsoptionen eingesetzt werden sollte, denn die Nebenwirkungen dieser Option fällt in dieselbe Unsicherheits-Klasse wie die Folgen des Klimawandels selbst. Die Konsequenzen eines entsprechenden Konzept-Transfers werden derzeit von uns erforscht.

Entscheidung unter Unsicherheit am Klimaproblem stellt eine jahrzehntelang unterschätzte Herausforderung an die ökonomische Grundlagenforschung im Verbund mit den betroffenen klimawissenschaftlichen Bereichen dar – mit potentiell gravierenden gesellschaftspolitischen Konsequenzen. Sie bietet nicht nur die Chance auf faszinierende Forschungsagenden, sondern regt uns ebenfalls an, neu über die Arbeitsteilung von Akademia und Gesellschaft nachzudenken.

Danksagung

Ich danke Christian Dieckhoff für zahlreiche, sehr hilfreiche Hinweise in Bezug auf eine Vorversion dieses Textes.

Literatur

- Blau, Roger A. (1974): Stochastic Programming and Decision Analysis: an Apparent Dilemma. In: *Management Science* 21.3, 271–276.
- Bürgenmeier, Beat/Baranzini, Andrea/Ferrier, Catherine/Germound-Duret, Céline/Ingold, Karin/Perret, Sylvain/Rafaj, Peter/Kypreos, Socrates/Wokaun, Alexander (2006): Economics of climate policy and collective decision making. In: *Climatic Change* 79, 143–162.
- Charnes, Abraham/Cooper, William W. (1959): Chance constrained programming. In: *Management Science* 6, 73–79.
- Edenhofer, Ottmar/Bauer, Nico/Kriegler, Elmar (2005): The impact of technological change on climate protection and welfare — insights from the model MIND. In: *Ecological Economics* 54, 277–292.
- Gollier, Christian (2001): *The economics of risk and time*. Cambridge (Mass.)/London.
- Hediger, Werner (1999): Reconciling “Weak” and “Strong” Sustainability. In: *International Journal of Social Economics* 26, 1120–1143.
- Held, Hermann/Kriegler, Elmar/Lessmann, Kai/Edenhofer, Ottmar (2009): Efficient Climate Policies under Technology and Climate Uncertainty. In: *Energy Economics* 31, 50–61.
- Horowitz, John/Lange, Andreas (2014): Cost-Benefit Analysis under Uncertainty – A Note on Weitzman’s Dismal Theorem. In: *Energy Economics* 42, 201–203.
- IPCC (2013): Summary for Policymakers. In: Stocker, Thomas F./Qin, Dahe/Plattner, Gian-Kaspe/Tignor, Melinda M. B./Allen, Simon K./Boschung, Judith/Nauels, Alexander/Xia, Yu/Bex, Vincent Pauline M. (Hrsg.): *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge/New York.
- IPCC (2014): Summary for Policymakers. In: Edenhofer, Ottmar/Pichs-Madruga, Ramón/Sokona, Youba/Minx, Jan C./Farahani, Ellie/Kadner, Susanne/Seyboth, Kristin/Adler, Anna/Baum, Ina/Brunner, Steffen/Eickemeier, Patrick/Kriemann, Benjamin/Savolainen, Jussi/Schlömer, Steffen/von Stechow, Christoph/Zwicker, Timm (Hrsg.): *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge/New York.
- Kleinen, Thomas Christopher (2005): *Stochastic Information in the Assessment of Climate Change*. Diss. Universität Potsdam.
- Klocke, Daniel (2011): Assessing the Uncertainty in Climate Sensitivity. In: *Reports on Earth System Science* 95.

- Knight, Frank H. (1921): *Uncertainty and Profit*. Boston.
- Kolstad, Charles/Urama, Kevin/Broome, John/Bruvoll, Annegrete/Olvera, Michelle C./Fullerton, Don/Gollier, Christian/Hanemann, William Michael/Hassan, Rashid/Jotzo, Frank/Khan, Mizan R./Meyer, Lukas/Mundaca, Luis (2014): *Social Economic and Ethical Concepts and Methods*. In: Edenhofer, Ottmar/Pichs-Madruga, Ramón/Sokona, Youba/Minx, Jan C./Farahani, Ellie/Kadner, Susanne/Seyboth, Kristin/Adler, Anna/Baum, Ina/Brunner, Steffen/Eickemeier, Patrick/Kriemann, Benjamin/Savolainen, Jussi/Schlömer, Steffen/von Stechow, Christoph/Zwicker, Timm (Hrsg.): *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge/New York.
- Kriegler, Elmar/Bruckner, Thomas (2004): *Sensitivity Analysis of Emission Corridors for the 21st Century*. In: *Climatic Change* 66, 345–387.
- Kunreuther, Howard/Gupta, Shreekanth/Bosetti, Valentina/Cooke, Roger/Dutt, Varun/Ha-Duong, Minh/Held, Hermann/Llanes-Regueiro, Juan/Patt, Anthony/Shittu, Ekundayo/Weber, Elke (2014): *Integrated Risk and Uncertainty Assessment of Climate Change Response Policies*. In: Edenhofer, Ottmar/Pichs-Madruga, Ramón/Sokona, Youba/Minx, Jan C./Farahani, Ellie/Kadner, Susanne/Seyboth, Kristin/Adler, Anna/Baum, Ina/Brunner, Steffen/Eickemeier, Patrick/Kriemann, Benjamin/Savolainen, Jussi/Schlömer, Steffen/von Stechow, Christoph/Zwicker, Timm (Hrsg.): *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge/New York, 151–206.
- Luderer, Gunnar/Pietzcker, Robert C./Bertram, Christoph/Kriegler, Elmar/Meinshausen, Malte/Edenhofer, Ottmar (2013): *Economic Mitigation Challenges: How Further Delay Closes the Door for Achieving Climate Targets*. In: *Environmental Research Letters* 8.3, 034033.
- Mastrandrea, Michael D./Field, Christopher B./Stocker, Thomas F./Edenhofer, Ottmar/Ebi, Kristie L./Frame, David J./Held, Hermann/Kriegler, Elmar/Mach, Katherine J./Matschoss, Patrick R./Plattner, Gian Kasper/Yohe, Gary W./Zwiers, Francis W. (2010): *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties*, Intergovernmental Panel on Climate Change IPCC.
- Meinshausen, Malte/Meinshausen, Nicolai/Hare, William/Raper, Sarah C. B./Frieler, Katja/Knutti, Reto/Frame, David J./Allen, Myles R. (2009): *Greenhouse-Gas Emission Targets for Limiting Global Warming to 2°C*. In: *Nature* 458, 1158–1163.
- Nelson, Julie A. (2013): *Ethics and the Economist: What Climate Change Demands of Us*. In: *Ecological Economics* 85, 145–154.

- Neubersch, Delf/Held, Hermann/Otto, Alexander (2014): Operationalizing Climate Targets under Learning: An Application of Cost-Risk Analysis. In: *Climatic Change* 126, 305–318.
- Nordhaus, William D. (2008): *A Question of Balance: Weighing the Options on Global Warming Policies*. New Haven/London.
- Petschel-Held, Gerhard/Schellnhuber, Hans-Joachim/Bruckner, Thomas/Tóth, Ferenc L./Hasselmann, Klaus (1999): The Tolerable Windows Approach: Theoretical and Methodological Foundations. In: *Climatic Change* 41, 303–331.
- Roth, Robert/Neubersch, Delf/Held, Hermann (2015): Evaluating Delayed Climate Policy by Cost-Risk Analysis, begutachteter EAERE2015-Artikel, Programmankündigung: <http://www.eaere2015.org/programme.html>, conference booklet S. 59 (aufgerufen 8.4.2018).
- Savage, Leonard J. (1954): *The Foundations of Statistics*. New York.
- Schellnhuber, Hans-Joachim (2010): Tragic Triumph. In: *Climatic Change* 100, 229–238.
- Schmidt, Matthias G. W./Lorenz, Alexander/Held, Hermann/Kriegler, Elmar (2011): Climate Targets under Uncertainty: Challenges and Remedies. In: *Climatic Change* 104, 783–791.
- Schneider von Deimling, Thomas/Ganopolski, Andrey/Held, Hermann/Rahmstorf, Stefan (2006a): How Cold was the Last Glacial Maximum? In: *Geophysical Research Letters* 33, L14709, DOI: 10.1029/2006GL026484.
- Schneider von Deimling, Thomas/Held, Hermann/Ganopolski, Andrey/Rahmstorf, Stefan (2006b): Climate Sensitivity Estimated from Ensemble Simulations of Glacial Climates. In: *Climate Dynamics* 27, 463–483.
- Stern, Nicholas (2007): *The Economics of Climate Change – The Stern Review*. Cambridge.
- Stern, Nicholas (2013): The Structure of Economic Modelling of the Potential Impacts of Climate Change: Grafting Gross Underestimation of Risk onto already Narrow Science Models. In: *Journal of Economic Literature* 51, 838–859.
- von Neumann, John/Morgenstern, Oskar (1944): *Theory of Games and Economic Behavior*. Princeton.
- Weitzman, Martin L. (2009): Additive Damages, Fat-Tailed Climate Dynamics, and Uncertain Discounting. In: *SSRN Electronic Journal* 3, 1–24, <http://www.economics-ejournal.org/economics/journalarticles/2009-39> (aufgerufen 8.4.2018).
- Zachos, James/Pagani, Mark/Sloan, Lisa/Thomas, Ellen/Billups, Katharina (2001): Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. In: *Science* 292, 686–693.

Martin Scheringer (Brünn/Zürich)

Unsicherheit als zentrales Problem in der Risikobewertung für Chemikalien

Abstract: The risk assessment of chemicals is a scientific procedure that aims to determine the risks to human health and the environment that are associated with the use of commercially relevant chemicals. There are many uncertainties associated with the procedure. These uncertainties derive from the high number of chemicals on the market (several tens of thousands), a lack of data on chemical properties, erroneous and inaccurate chemical property data, bias in chemical property measurement methods, the huge variety of uses of chemicals in many consumer products and technical applications, the wide range of chemical properties such as vapor pressure, water solubility, degradation half-lives, toxicity and many others, and the wide range of possible adverse effects in humans and wildlife. Here different sources and types of uncertainty along with methods for dealing with the different types of uncertainty are presented. Overall, the uncertainties associated with the different elements of the chemical risk assessment procedure are substantial. It is essential that these uncertainties are better characterized in the future in order to make chemical risk assessment more rational and reliable.

Keywords: Industriechemikalien – Emissionsdaten – Stoffeigenschaften – Umweltverhalten – Persistenz – Bioakkumulation – Toxizität – „Datenbank-Unsicherheit“

1 Risikobewertung für Chemikalien: Vorgehensweise

Die Risikobewertung für Chemikalien ist ein Verfahren, in dem chemische Produkte im Hinblick auf ihre schädlichen Effekte für Mensch und Umwelt untersucht werden. Die Ergebnisse dieser Untersuchung bilden die Grundlage für die Regulierung chemischer Substanzen, die als Arzneimittel, Pflanzenschutzmittel, Biozide, Industriechemikalien usw. verwendet werden und somit kommerziell relevant sind (vgl. van Leeuwen/Vermeire 2007).

Das Verfahren fußt auf einer Erfassung verschiedener Arten von Daten: Ein erstes Element bilden die physikalisch-chemischen Substanzeigenschaften wie die Brennbarkeit, der Dampfdruck, die Wasserlöslichkeit etc., welche entweder nach standardisierten Verfahren gemessen oder aus der chemischen Struktur abgeschätzt werden können. Das zweite Element sind die toxischen Wirkungen einer Substanz, die sich in Testorganismen (*in vivo*) oder *in-vitro*-Testsystemen zeigen, und das dritte Element sind die Verwendungsmuster und Verwendungsmengen; diese beeinflussen (neben den chemischen Eigenschaften der betrachteten Sub-

stanz), welche Exposition von Mensch und Umwelt zu erwarten ist. Aus allen diesen Elementen wird dann ermittelt, welche Konzentrationen der betrachteten Substanz als Folge einer bestimmten Verwendung, z. B. als Lösungsmittel, in verschiedenen Umweltkompartimenten (Luft, Wasser, Boden) zu erwarten sind, und ob diese Konzentrationen eine aus den toxikologischen Befunden abgeleitete Nichtwirkungs-Schwelle überschreiten. Wenn sich abzeichnet, dass die Nichtwirkungs-Schwelle überschritten werden könnte, müssen letztendlich Maßnahmen zur Risikominderung getroffen werden (vgl. van Leeuwen/Vermeire 2007).

Das Bewertungsverfahren existiert in dieser Grundform in vielen Ländern und für verschiedene Arten von Chemikalien; im Folgenden wird vor allem auf den Kontext der Europäischen Union (EU) Bezug genommen. Die Entwicklung von Testmethoden für die diversen physikalischen, chemischen und toxikologischen Eigenschaften, die im Verfahren benötigt werden, wird seit über 40 Jahren von der Organisation für Wirtschaftliche Zusammenarbeit und Entwicklung (OECD) koordiniert (vgl. OECD 2016). Das Verfahren befindet sich seit den 1980er-Jahren in einer kontinuierlichen Entwicklung und wurde mehrfach stark überarbeitet. Der Grund dafür liegt darin, dass die zu bewertenden Chemikalien eine große Vielfalt an unterschiedlichen Eigenschaften haben, welche sich nicht mit einem abschließend festgelegten Raster von Testverfahren erfassen lassen, und dass es sehr viele Substanzen sind, die bewertet werden müssen, nämlich mindestens einige zehntausend (ECHA 2018a, 2018b). Wichtige Meilensteine bei der Weiterentwicklung des Verfahrens waren die Richtlinie und die Verordnung für die Risikobewertung alter und neuer Stoffe, welche 1991 in Kraft traten (vgl. Richtlinie 93/67/EEC und Verordnung (EG) 1488/94); für Pflanzenschutzmittel die Richtlinie 91/414 EEC sowie die Verordnung (EG) 1107/2009, welche die Richtlinie 91/414 EEC ersetzt hat, sowie REACH, eine umfassende neue Verordnung der EU für Industriechemikalien, welche 2007 in Kraft trat (Verordnung (EG) 1907/2006). REACH steht für *Registration, Evaluation, Authorisation and Restriction of Chemicals* (Registrierung, Bewertung, Zulassung und Beschränkung von Chemikalien).

Das Bewertungsverfahren hat sich von Beginn an im Spannungsfeld zwischen Wissenschaft und echter Grundlagenforschung einerseits und Routine/Vollzug andererseits befunden, und damit wird das Thema der Unsicherheit relevant. Wie erwähnt, kann sich das Verfahren nämlich nicht auf eine abschließend definierte Vorgehensweise stützen, welche sich von Behörden und Auftragslaboren routinemäßig durchführen ließe, sondern es muss fortwährend weiterentwickelt werden. Der Grund dafür ist, dass die Anzahl der Substanzen, die zu prüfen sind, so groß

ist, dass eine umfassende Untersuchung aller Substanzen nicht möglich ist. Es handelt sich um mehrere 10 000 verschiedene Substanzen, und hinzu kommt, dass für jede einzelne Substanz mehrere oder viele Verwendungsweisen zu prüfen sind, dass die Verteilungsmuster in der Umwelt sehr komplex sein können und dass es eine im Prinzip unbegrenzte Anzahl von schädlichen Effekten gibt, die zu prüfen wären. Schließlich kommt hinzu, dass jede Substanz für sich untersucht wird, was dazu führt, dass das Zusammenwirken verschiedener Substanzen in der Umwelt oder im menschlichen Körper im Rahmen des Bewertungsverfahrens systematisch ignoriert wird und die resultierenden Risiken systematisch unterschätzt werden. Hierin liegt der zentrale Punkt dieses Beitrags: Unsicherheiten sind ganz grundlegend mit dem Bewertungsverfahren für Chemikalien verbunden, und die Erfassung und Charakterisierung von Unsicherheiten muss als zentraler Bestandteil des Verfahrens angesehen werden.¹

Illustrieren lässt sich die Problematik von Unsicherheiten im Bewertungsverfahren für Chemikalien mit der Situation gegen Ende der 1990er-Jahre: Zu dieser Zeit waren die sogenannten Industriechemikalien (dies sind im Wesentlichen alle Substanzen, die nicht als Arzneimittel, Pestizide oder Biozide verwendet werden, z. B. Lösungsmittel, Flammschutzmittel, Weichmacher für Kunststoffe, Imprägniermittel, Farbstoffe, u. v. a. m.) noch in „Altstoffe“ und „Neustoffe“ eingeteilt. Altstoffe waren Substanzen, die bereits vor 1981 auf dem Markt waren, Neustoffe solche, die nach 1981 auf den Markt kamen. Für die Risikobewertung der Altstoffe waren nicht die Hersteller, sondern die Behörden der EU-Mitgliedstaaten verantwortlich; in Deutschland unterstützte das Beratergremium für umweltrelevante Altstoffe (BUA) die Regierung bei der Altstoffbewertung, und darüber hinaus betrieb auch die OECD ein Altstoffprogramm. Ende der 1990er-Jahre waren von den ca. 100 000 Altstoffen (Neustoffe gab es nur knapp 6 000) auf EU-Ebene erst weniger als 20 offiziell bewertet (vgl. EEA 1998). Die Altstoffbewertung war somit festgefahren und die Situation war eigentlich vor allem durch *Nichtwissen* und

1 In diesem Beitrag liegt der Fokus auf den physikalisch-chemischen Stoffeigenschaften und dem Verteilungsverhalten in der Umwelt. Unsicherheiten, welche mit toxischen Wirkungen verbunden sind, gehen über den Rahmen dieses Beitrags hinaus, sind aber ebenfalls von großer Bedeutung und sind z. T. sehr hoch. Dies zeigt die aktuelle Diskussion über endokrine (d. h. hormonähnliche) Wirkungen von Chemikalien, die bereits bei sehr tiefen Konzentrationen auftreten können und in etablierten toxikologischen Testverfahren nicht erfasst werden (vgl. UNEP/WHO 2013). Die Anzahl möglicher sogenannter endokriner Disruptoren (*endocrine disrupting chemicals*, EDC) ist groß, aber Stärke und Art des Effekts sind sehr variabel; über die Definition und Regulierung von EDC wird seit längerem äußerst heftig gestritten (vgl. Bergman et al. 2013).

nicht durch Unsicherheit gekennzeichnet (vgl. Scheringer 2013). Daher wurde ein grundlegender Neuansatz für nötig gehalten, und dieser wurde im sogenannten Weißbuch für REACH präsentiert (vgl. EC 2001). REACH trat dann im Jahr 2007 in Kraft und kann als ein großer Fortschritt angesehen werden, da es von den Herstellern und Importeuren von Chemikalien verlangt, dass diese die Substanzen, welche sie auf dem Markt behalten wollen, selbst testen (oder testen lassen) und dann die erforderlichen Daten der Europäischen Chemikalienagentur (ECHA) übermitteln. Dies drückt sich im REACH-Motto, „No data, no market“, aus. Dies soll dazu führen, dass nur noch Substanzen auf dem Markt gehalten werden, deren technische und ökonomische Bedeutung groß genug ist, um die Testung zu rechtfertigen, während alle anderen Substanzen vom Markt verschwinden.

Eine Hauptaussage des vorliegenden Beitrags ist, dass jedoch auch REACH das Problem nicht wirklich löst, da die Anzahl der zu untersuchenden Substanzen immer noch zu hoch ist. Es werden nun sehr viele Daten eingereicht, welche nicht korrekt sind und die dadurch die Qualität der Stoffdatenbank der ECHA erheblich beeinträchtigen (vgl. Stieger et al. 2014; Scheringer 2013; Springer et al. 2015; siehe auch unten Abschnitt 3).

2 Unsicherheiten bei der Bewertung einer einzelnen Substanz

Die mit der Risikobewertung verbundenen Unsicherheiten lassen sich gliedern, indem man von den Emissionsdaten zu den Stoffeigenschaften und dann zum Umweltverhalten (sowie den toxischen Wirkungen, hier nicht behandelt) übergeht.

- *Unsicherheiten bei Emissionsdaten:* Historisch sind Insektizide wie Dichlordiphenyltrichlorethan (DDT) und Industriechemikalien wie polychlorierte Biphenyle (PCB) gute Fallbeispiele, da sie einerseits ausführlich untersucht worden sind und andererseits auch heute immer noch relevant sind (vgl. Scheringer 2012). DDT steht seit Rachel Carsons Buch „Silent Spring“ (1962) als paradigmatische Substanz für das Problem der Umweltverschmutzung durch Chemikalien. Nach massivem Einsatz von DDT in den 1950er- und 1960er-Jahren erkannte man, dass das DDT überall in der Umwelt zu finden war und nicht einfach wieder verschwand. Für eine Substanz von großer chemischer Stabilität („Persistenz“) ist das eigentlich kein erstaunlicher Befund, aber Carsons Darstellung des Problems löste dennoch großes Erstaunen (und heftige Diskussionen) aus, und bald kam es zu ersten Verboten der Verwendung von DDT in der Landwirtschaft (gegen krankheitsübertragende Insekten ist die Verwendung von DDT bis heute möglich).

Verwendungs- und Emissionsdaten zu DDT und ähnlichen Pestiziden zeigen, dass die Unsicherheit der Emissionsmengen typischerweise einen Faktor drei jeweils nach oben und nach unten beträgt. Wenn fallspezifische Emissionsdaten bekannt sind, ist die Unsicherheit kleiner. Bei Industriechemikalien mit äußerst zahlreichen und zugleich sehr unterschiedlichen Anwendungen wie den polychlorierten Biphenylen sind die Unsicherheiten deutlich größer; hier ist es ein Faktor von zehn, den man um den Schätzwert der Emission herumlegen muss, um die Unsicherheit abzubilden, wodurch ein Unsicherheitsband von 100 entsteht (vgl. Breivik et al. 2007). Für perfluorierte Carboxylsäuren (PFCA), die u. a. bei der Herstellung von Teflon und von Imprägniermitteln für Bekleidung, Teppiche, Nahrungsmittelverpackungen etc. verwendet werden, ist die Unsicherheit kleiner; Wang et al. (2014) haben einen Faktor von acht zwischen einem Szenario mit tiefen Emissionsabschätzungen und einem Szenario mit hohen Emissionsabschätzungen ermittelt.

- *Unsicherheiten bei Stoffeigenschaften*: Ein sehr illustratives Beispiel ist hier der Verteilungskoeffizient zwischen Oktanol und Wasser (K_{ow}). Der K_{ow} ist in der Stoffbewertung von großer Bedeutung, weil er die relative Affinität einer Substanz für Wasser und für organisches Material (Böden, Vegetation, Fettgewebe in Organismen, Muttermilch) beschreibt. Er ist eine Basisgröße, die für jede Substanz erhoben und bei der Registrierung unter REACH eingereicht werden muss. Bei Substanzen mit geringer Wasserlöslichkeit ist der K_{ow} schwer zu messen, weil dann die Konzentration in der Wasserphase sehr niedrig ist. Beispielsweise ist für das Insektizid DDT die Konzentration im Wasser um ca. einen Faktor von einer Million tiefer als im Oktanol. Pontolillo/Eganhouse (2001) haben für das Insektizid DDT alle überhaupt verfügbaren K_{ow} -Werte zusammengestellt. Dabei hat sich ergeben, dass die Werte erstaunlich stark streuen, nämlich über ca. vier Größenordnungen, also einen Faktor 10 000. Dies ist ein irritierender Befund, da der K_{ow} eine wohldefinierte Stoffeigenschaft ist, die unter standardisierten Bedingungen erhoben wird und in die keine biologische Variabilität (wie sie bei toxikologischen Untersuchungen auftritt) einfließt. Es handelt sich also um reine Messungenauigkeiten. Bei anderen Stoffeigenschaften wie den Halbwertszeiten des biologischen Abbaus können ähnlich große Unsicherheiten auftreten. Für die Stoffbewertung heißt dies, dass für alle Eigenschaften einer Substanz die Unsicherheit der vorhandenen Messwerte erhoben werden und in die Beurteilung einbezogen werden muss. Die Unsicherheiten sind dabei deutlich größer als z. B. Abweichungen von einigen Prozent; sie erreichen einen Faktor 10 und mehr in jede Richtung.

- *Unsicherheiten bezüglich des Umweltverhaltens*: Nach der Charakterisierung von Stoffeigenschaften und Emissionsdaten ist die Verteilung in der Umwelt der nächste Schritt. Um die Verteilung einer Chemikalie in der Umwelt zu verstehen und ihre Konzentrationen in den verschiedenen Umweltkompartimenten zu bestimmen, verwendet man Modelle, welche die Stoffflüsse von Chemikalien zwischen Boden, Wasser und Luft berechnen und so eine Gesamtbilanz für die Substanz in einem System aus Boden, Wasser und Luft liefern (vgl. Scheringer 2015). Beispielsweise wird in einem solchen Modell berechnet, wie viel Substanz vom Boden in die Luft verdunstet und wie viel Substanz mit dem Regen wieder aus der Luft ausgewaschen und auf dem Boden deponiert wird. Dafür müssen der Dampfdruck und die Wasserlöslichkeit der Substanz (sowie weitere Stoffeigenschaften) als Eingabedaten in das Modell eingespeist werden. Sofern diese Daten mit einer quantifizierbaren Unsicherheit behaftet sind (Unsicherheitsbänder, siehe oben), können auch diese Bandbreiten in das Modell eingespeist werden. Für die Unsicherheiten bedeutet dies, dass sich die Unsicherheiten aller Stoffeigenschaften durch das Modell hindurch fortpflanzen, also auf die vom Modell berechneten Konzentrationen durchschlagen. Dabei reagiert das Modell allerdings mit unterschiedlichen Sensitivitäten gegenüber verschiedenen Eingabeparametern: Eine Unsicherheit in den Emissionsmengen überträgt sich direkt auf die berechneten Konzentrationen, wohingegen Unsicherheiten z. B. in der Regenrate oder der Windgeschwindigkeit nur schwächer auf die berechnete Konzentration wirken. Im Einzelnen hängen die Sensitivitäten eines solchen Modells gegenüber verschiedenen Parametern von der Struktur des Modells und der betrachteten Substanz ab.

Somit lässt sich festhalten: Methoden für die Erfassung und quantitative Behandlung von Unsicherheiten von den Emissionen bis zu den in der Umwelt zu erwartenden Konzentrationen sind etabliert, aber damit diese Methoden greifen, muss die Unsicherheit der Eingangsdaten quantifiziert, also ihrerseits recht gut charakterisiert sein. Ein Beispiel ist das Insektizid Endosulfan, für welches Becker et al. (2011) den hier skizzierten Prozess von den Emissionsdaten über die Stoffeigenschaften bis hin zu den in der Umwelt zu erwartenden Konzentrationen einschließlich aller Unsicherheiten durchgearbeitet haben. Auf der Grundlage der von Becker et al. (2011) erhaltenen Ergebnisse konnte dann das POP Review Committee der Stockholm-Konvention beschließen, dass Endosulfan die Eigenschaften eines persistenten organischen Schadstoffes im Sinne der Konvention erfüllt, und damit war die Grundlage für das weltweite Verbot von Endosulfan im Jahr 2011 gelegt. Trotz der mit allen Elementen der Bewertung verbundenen Unsicherheiten ergeben sich in solch einem recht gut dokumentierten Fall am

Ende des Bewertungsverfahrens belastbare Aussagen, die auch als politische Entscheidungsgrundlagen geeignet sind (vgl. Scheringer 2015). Als Fazit ergibt sich: Selbst für Chemikalien, die seit Längerem untersucht werden, bestehen einerseits noch erhebliche Unsicherheiten und Wissenslücken auf allen Stufen des Verfahrens, andererseits lassen sich diese Unsicherheiten in Form von Bandbreiten quantitativ abschätzen und damit für politische Entscheidungsprozesse „bändigen“ und handhaben (Beispiel Endosulfan).

3 Ein neues Problem: Datenbank-Unsicherheit

Unter REACH entsteht zurzeit ein neues, so nicht erwartetes Problem, welches eine neue Qualität von Unsicherheit hervorbringt, die hier als „Datenbank-Unsicherheit“ bezeichnet werden soll. Die mit den Registrierungs dossiers bei der ECHA eingereichten Stoffdaten werden in einer von der ECHA verwalteten Datenbank abgelegt. Wie sich nun, während kontinuierlich mehr und mehr Daten in die Datenbank eingespeist werden, zeigt, ist die Datenbank mit einer unbekannt Menge falscher Daten durchsetzt. Das Problem geht über die übliche Problematik eines (sehr) kleinen Anteils an falschen Datenpunkten hinaus, der bei großen Datenbeständen immer anfällt. Vielmehr handelt es sich um eine Kombination aus einerseits erheblichen Datenlücken und andererseits systematischen Messfehlern, die bei vielen Substanzen in ähnlicher Weise auftreten. Die Datenlücken werden dokumentiert in einem Bericht von Springer et al. (2015); untersucht wurden 1814 Dossiers von Substanzen mit einem Produktionsvolumen von über 1000 t/a (sog. *high-production-volume chemicals*). Von diesen 1814 Dossiers entsprach nur ein einziges (!) den Vorgaben, 58 % waren mangelhaft, und bei 42 % war die Datenlage unklar und es ließ sich im Rahmen einer standardisierten web-basierten Überprüfung nicht entscheiden, ob ein Mangel vorliegt oder nicht.

Das Problem der Messfehler haben Stieger et al. (2014) anhand einer Gruppe bromierter Flammschutzmittel illustriert. Zwei der zahlreichen Datenpunkte, die mit einem Dossier eingereicht werden müssen, sind der Oktanol-Wasser-Verteilungskoeffizient (K_{ow}) und die aquatische Toxizität, ausgedrückt als LC_{50} (Konzentration im Wasser, bei der 50 % der Testorganismen sterben). Die Analyse von Stieger et al. (2014) zeigt, dass der K_{ow} häufig zu tief ist, oft um mehrere Größenordnungen², und dass für die LC_{50} häufig zu hohe Werte angegeben werden, also eine zu geringe Toxizität ausgewiesen wird. Ursachen für diese Messfehler sind beim K_{ow} , dass insbesondere hohe K_{ow} -Werte schwer zu messen sind, weil

2 Eine Größenordnung ist ein Faktor 10.

sie eine analytische Bestimmung der Substanzkonzentration in Wasser erfordern und weil bei hohem K_{ow} die Konzentration im Wasser tief ist (vgl. Pontolillo/Eganhouse 2001). In vielen Fällen scheint die Nachweisgrenze des Messverfahrens nicht tief genug zu sein, um solche niedrigen Konzentrationen bestimmen zu können. Dadurch werden zu hohe Konzentrationen im Wasser und in der Folge zu tiefe K_{ow} -Werte angegeben. Diese Beobachtung wird dadurch ergänzt, dass Stieger et al. (2014) keine K_{ow} -Werte gefunden haben, die zu hoch liegen. Bei der LC_{50} für Fische oder Flusskrebse treten mehrere Probleme gemeinsam auf. Ein Faktor ist auch hier eine geringe Wasserlöslichkeit, weil bei tiefer Wasserlöslichkeit die Konzentration im Testsystem schwer einzustellen und zu kontrollieren ist. Immer wieder werden sogar nur nominale Konzentrationen angegeben, d. h. die Menge der Substanz, die dem Testsystem zugefügt wurde, wobei dies oft mehr ist, als im Wasser gelöst werden kann. Nominale Konzentrationen sind somit physikalisch und biologisch sinnlos. Ein weiteres Problem ist, dass bei tiefer Wasserlöslichkeit die Aufnahme in den Organismus langsam abläuft, so dass bei den primär geforderten Tests auf akute Toxizität die Testdauer zu kurz sein kann, um toxische Effekte sichtbar werden zu lassen.

Das Problem der Datenbank-Unsicherheit ergibt sich nun, weil die ECHA zum einen keine genügende Kapazität hat, um alle Daten auf ihre inhaltliche Korrektheit zu prüfen, und weil sie keine Handhabe hat, als fehlerhaft erkannte Daten zurückzuweisen bzw. bei Substanzen mit fehlerhaften Dossiers die Registrierung zu verweigern (vgl. UBA 2012). Dadurch akkumulieren sich fehlerhafte Daten in einem zunehmenden, insgesamt jedoch nicht genau bekannten Umfang.

Unter REACH besteht nun also eine Ambivalenz, welche zu „fragiler Evidenz“ (Scheringer 2013) führt: Einerseits sind viel mehr Daten zu Industriechemikalien verfügbar als je zu vor, andererseits ist der Datenbestand als ganzer nicht verlässlich, weil fehlerhafte Daten nicht markiert sind und daher nicht ohne weiteres erkennbar ist, ob ein Datenpunkt korrekt ist oder nicht. Wenn Anwender Daten aus der Datenbank abrufen, ist nicht klar, *ob* ein Datum falsch oder korrekt ist, und zweitens ist ebenfalls unklar, *wie stark* ein Datum allenfalls vom korrekten Wert abweicht. Der Inhalt der Datenbank ist sozusagen „kontaminiert“ mit einer systemischen Unsicherheit. Angesichts der Größe des Datenbestandes (im Februar 2018 gab es 17 885 registrierte Substanzen, wobei für jede Substanz zahlreiche Datenpunkte vorhanden sind) ließe sich diese Datenbank-Unsicherheit nur mit sehr großem Aufwand identifizieren und vermindern. Zurzeit gibt es seitens der ECHA keine Hinweise, wie das Problem der Datenbank-Unsicherheit angegangen werden soll. Immerhin wurde im März 2017, also zehn Jahre nach

dem Inkrafttreten von REACH, eine Möglichkeit geschaffen, eingereichte Stoffdaten von ca. 15 000 unter REACH registrierten Substanzen herunterzuladen (vgl. ECHA 2018c). Dieser Schritt war lange überfällig, denn größere Datensätze wie z. B. die K_{ow} -Werte oder der Biokonzentrationsfaktor sämtlicher bisher unter REACH registrierten Substanzen sind erforderlich, um einzelne Datenpunkte in den Kontext der Werte für andere, strukturell ähnliche Substanzen stellen und auf diese Weise fehlerhafte Werte identifizieren zu können (Stieger et al. 2014 haben dies für den K_{ow} von bromierten Flammschutzmitteln exemplarisch durchgeführt).

In diesem Zusammenhang ergeben sich interessante Folgefragen hinsichtlich der Qualität von (sehr) großen Datenbeständen, die (sehr) schnell aufgebaut werden (*big data*).

4 Sind die Unsicherheiten zu groß?

Auf naturwissenschaftlicher Seite gibt es einen weiteren Effekt, der bei der Risikobewertung von Chemikalien erschwerend wirkt: Ab einem gewissen Ausmaß von Unsicherheit wird unter Naturwissenschaftlerinnen und -wissenschaftlern häufig die Ansicht vertreten, dass man ein Problem nicht mehr sinnvoll bearbeiten könne. Gerade bei der Risikobewertung von Chemikalien, bei der die empirische Grundlage aufgrund der fehlenden Daten tatsächlich erstaunlich schwach ist, wird immer wieder eingewendet, dass die Unsicherheiten zu groß seien, um „sinnvolle“ oder „belastbare“ Aussagen zu erhalten. Diese Sichtweise führt dazu, dass Probleme von Datenbedarf und Datenqualität in der Stoffbewertung, die erhebliche praktische und politische Bedeutung haben, der wissenschaftlichen Bearbeitung verschlossen bleiben.³

Auch bei einem recht gut dokumentierten Fall wie der globalen Verteilung von Endosulfan (siehe Abschnitt 2) treten schnell Bandbreiten in der Größe von einem Faktor 10–30 auf, allerdings stammt bereits ein Faktor von knapp 10 allein aus der Unsicherheit der Emissionen (vgl. Becker et al. 2011). Wenn die Datelage schlechter ist (was oft der Fall ist), werden die Unsicherheiten noch größer und es kann durchaus sein, dass man auf die Bestimmung von Bandbreiten um

3 Wenn diesen Fragen wissenschaftsintern die Bearbeitbarkeit abgesprochen wird, ist dies u. a. ein Ausdruck des Streites über die Deutungshoheit in einem wissenschaftlichen Gebiet, denn die Definition, was als relevantes und bearbeitbares Problem gelten kann, ist ein zentraler Schritt beim „Abstecken“ von Ansprüchen und Rechten in einem Arbeitsgebiet und bei der Zuweisung eines Gebietes sowie der für seine Bearbeitung erforderlichen Ressourcen an Wissenschaftler/-innen und Institute.

einen mittleren Wert herum verzichten muss und z. B. nur noch Minimum- und Maximum-Szenarien abschätzen kann, um das gesuchte Resultat grob einzugrenzen (vgl. Morgan 2001). Dieses Beispiel zeigt jedoch auch, dass natürlich auch Fragestellungen mit großen Unsicherheiten *nicht* unbearbeitbar sind. Vielmehr müssen bei so großen Unsicherheiten die *Methoden* und die *Zielsetzungen* angepasst werden: Die Zielsetzung kann nicht mehr die möglichst genaue quantitative Bestimmung eines Wertes (wie z. B. der Konzentration einer Substanz an einem gewissen Ort zu einer gewissen Zeit) sein. Zielsetzung ist dann vielmehr zunächst nur die Bestimmung von plausiblen Minimal- und Maximalwerten der gesuchten Größe sowie die Identifizierung von Faktoren, die die gesuchte Größe beeinflussen; weiterhin die Identifizierung von „Zwischenproblemen“, deren Beantwortung auf dem Weg zur gesuchten Größe weiterhilft, und schließlich die Bestimmung des methodischen Bedarfs, also die Benennung dessen, was neu zu entwickelnde Methoden leisten müssen, wenn man der gesuchten Größe näherkommen will. Auf der Methodenseite kann dann statt einer quantitativen Fehlerfortpflanzung wie z. B. durch sogenannte „Monte-Carlo-Methoden“ (vgl. MacLeod et al. 2002) eine „Bounding Analysis“ (Morgan 2001) durchgeführt werden.

Ein weiterer Punkt, bei dem im Gebiet der Stoffbewertung Umdenken gefordert ist, betrifft das Gewicht, welches gemessenen Werten zugesprochen wird. Aus diversen Gründen (historisch, methodisch) werden in den Naturwissenschaften gemessene Werte, also unmittelbar empirische Befunde, als besonders aussagekräftig angesehen. Bei der Stoffbewertung gibt es jedoch immer wieder Zielgrößen, für welche die bestehenden Messverfahren nicht genügend genau sind oder nicht den tatsächlich relevanten Aspekt messen. Beispiele sind der K_{ow} (vgl. Pontolillo/Eganhouse 2001), die LC_{50} für aquatische Toxizität (vgl. Mayer/Reichenberg 2006), der Biokonzentrationsfaktor (BCF) (vgl. Jonker/van der Heijden 2007) und auch die Halbwertszeit für biologischen Abbau (vgl. Ng et al. 2011). Für diese Größen gibt es jedoch neben der Messung auch die Möglichkeit, sie aus der chemischen Struktur einer Substanz abzuschätzen (vgl. van Leeuwen/Vermeire 2007: Kap. 9 und 10; US EPA 2016).

Es ist ein zentrales Erkenntnisziel der Chemie, Zusammenhänge zwischen der Struktur und den Eigenschaften chemischer Substanzen zu etablieren; diese Zusammenhänge gehören zum Kern des chemischen Wissens überhaupt und fließen in die genannten Abschätzmethoden ein. Für Abschätzmethoden bestehen natürlich immer Grenzen ihrer Anwendbarkeit (die sogenannte Anwendungsdomäne, *application domain*); außerhalb dieser Domäne sind mit einer Abschätzmethode gewonnene Resultate inkorrekt oder zumindest unsicher (wobei das Ausmaß der Unsicherheit nicht bekannt ist), aber innerhalb der Anwendungsdomäne liefern

solche Struktur-Eigenschafts- oder Struktur-Aktivitäts-Beziehungen durchaus valide Resultate, deren Unsicherheit quantifiziert werden kann. Somit bestehen für einige in der Stoffbewertung zentrale Größen durchaus Alternativen in Form von Abschätzmethoden. Diese sind der direkten Messung vorzuziehen, wenn die derzeitig verfügbare Messmethode unzureichend ist und Artefakte erzeugt, hingegen das theoretische Verständnis gut genug ist, um brauchbare Schätzwerte zu liefern. Dies war die Vorgehensweise von Stempel et al. (2012) bei der Untersuchung der sogenannten PBT-Eigenschaften (Persistenz, Bioakkumulation, Toxizität) von 95 000 Chemikalien. Gemessene Werte für die hierbei benötigten Daten (Halbwertszeiten des biologischen Abbaus, Biokonzentrationsfaktor, aquatische Toxizität) waren nur für weniger als 3 000 von den 95 000 untersuchten Substanzen vorhanden. Die Vorgehensweise von Stempel et al. (2012) hat kontroverse Diskussionen über die Frage ausgelöst, ob eine zu großen Teilen auf abgeschätzten Daten beruhende Untersuchung dieser Größenordnung überhaupt als valide und als „wissenschaftlich“ gelten kann.

Wenn in solchen Diskussionen auf einer Priorität gemessener Daten insistiert und die Relevanz und Validität abgeschätzter Daten bestritten wird, hat dies zwei Aspekte: Von einem methodischen Standpunkt aus bedeutet diese Haltung, dass zentrale Elemente des etablierten chemischen Wissens ignoriert werden (siehe oben). Es geht in der Stoffbewertung jedoch bei weitem nicht nur um methodische Fragen, sondern auch um Machtfragen und ökonomische Aspekte. Daher führen Vertreterinnen und Vertreter der chemischen Industrie auch in Fällen, in denen Messungen größere Unsicherheiten und systematische Abweichungen erzeugen als Abschätzverfahren, immer wieder die angeblich höhere Gültigkeit gemessener Daten ins Feld. Ein Beispiel sind die viel zu tiefen K_{ow} -Werte, die von Stieger et al. (2014) gefunden wurden und die von der chemischen Industrie trotz offensichtlicher Abweichungen um viele Größenordnungen (konkret um bis zu einen Faktor von einer Milliarde; Stieger et al. 2014) als „valide, weil gemessen“ verteidigt wurden und werden.

5 Ausblick

Die diversen Quellen von Unsicherheit und die Größe und Komplexität des Problems erfordern eine fortwährende wissenschaftliche Weiterentwicklung und Fundierung des Verfahrens zur Chemikalienbewertung. Dass ein Teil des Verfahrens im Rahmen einer gewissen Routine abgewickelt werden kann (was durchaus sinnvoll und im Interesse der Industrie ist, weil damit Transparenz und Verlässlichkeit gewährleistet sind), darf nicht darüber hinwegtäuschen, dass genuine Forschungsleistungen notwendig sind, um das Verfahren weiterzuent-

wickeln und inhaltlich und methodisch abzusichern. Hier besteht zurzeit leider ein erhebliches Missverständnis dahingehend, dass es häufig als Signal für das Ende des Forschungsbedarfs gesehen wird, wenn eine Regulierung in Kraft getreten ist, wie z. B. REACH.

Gerade unter REACH, der Wasserrahmenrichtlinie sowie internationalen Abkommen wie der Stockholm-Konvention ist der Bedarf an umweltchemischer und ökotoxikologischer Forschung jedoch größer als je zuvor. Wenn eine Regulierung in Kraft getreten ist, ist damit verbindlich geworden, dass eine Problematik untersucht, Daten erhoben, Bewertungen vorgenommen werden müssen, und oft bedeutet dies nicht nur unmittelbaren „praktischen“ Arbeitsaufwand, sondern auch längerfristigen Bedarf an grundlegender Forschung: Es müssen neue Methoden zur Erhebung von Daten sowie Methoden zur Interpretation solcher Daten entwickelt werden, dann müssen diese Methoden in die praktische Anwendung gebracht werden, und erst damit werden die Grundlagen für die wirksame Umsetzung der Regulierung geschaffen (vgl. Scheringer 2018).

Die Stoffbewertung muss immer eine Brücke zwischen Wissenschaft und regulatorischer und politischer Umsetzung schlagen, und hier spielen die vorhandenen Unsicherheiten, und wie mit ihnen umgegangen wird, eine entscheidende Rolle. Beim Transfer des wissenschaftlichen Wissens in die behördliche und regulatorische Umsetzung können Unsicherheiten so gehandhabt werden, dass vorhandenes Wissen und seine Grenzen in Konsens-Erklärungen von Seiten der Wissenschaft gebündelt und in den Diskurs eingespeist werden. Beispiele für solche Konsens-Erklärungen, die zum Teil auch ausführlich in den Medien behandelt wurden, sind die Erklärungen zu bromierten Flammschutzmitteln („San Antonio-Statement“, Di Gangi et al. 2010) und zu fluorierten Substanzen für Imprägniermittel, Feuerlöschschäume etc. („Helsingør-Statement“, Scheringer et al. 2014; „Madrid-Statement“, Blum et al. 2015).

Literatur

- Becker, Linus/Scheringer, Martin/Schenker, Urs/Hungerbühler, Konrad (2011): Assessment of the environmental persistence and long-range transport of endosulfan. In: *Environmental Pollution* 159, 1737–1743.
- Bergman, Åke/Andersson, Anna-Maria/Becher, Georg/van den Berg, Martin/Blumberg, Bruce/Bjerregard, Poul et al. (2013): Science and policy on endocrine disruptors must not be mixed. A reply to a “common sense” intervention by toxicology journal editors. In: *Environmental Health* 12, 69. <http://www.ehjournal.net/content/12/1/69> (abgerufen 27.2.2018).

- Blum, Arlene/Balan, Simona/Scheringer, Martin/Trier, Xenia/Goldenman, Gretta/ Cousins, Ian/Diamond, Miriam/Fletcher, Tony/Higgins, Chris/Lindeman, Avery/ Peaslee, Graham/de Voogt, Pim/Wang, Zhanyun/Weber, Roland (2015): The Madrid Statement on Poly- and Perfluoroalkyl Substances (PFASs). In: *Environmental Health Perspectives* 123, A107–A111.
- Brevik, Knut/Sweetman, Andy/Pacyna, Jozef/Jones, Kevin (2007): Towards a global historical emission inventory for selected PCB congeners. A mass balance approach 3. An update. In: *Science of the Total Environment* 377, 296–307.
- Carson, Rachel (1962): *Silent Spring*. Boston.
- Di Gangi, Joe/Blum, Arlene/Bergman, Åke/de Wit, Cynthia/Lucas, Donald/Mortimer, David/Schechter, Arnold/Scheringer, Martin/Shaw, Susan/Webster, Thomas (2010): San Antonio Statement on Brominated and Chlorinated Flame Retardants. In: *Environmental Health Perspectives* 118, A516–A518.
- EC (European Commission) (2001): White Paper on the Strategy for a future Chemicals Policy. http://ec.europa.eu/environment/archives/chemicals/reach/background/white_paper.htm?uri=CELEX:52001DC0088 (abgerufen 27.2.2018).
- ECHA (European Chemicals Agency) (2018a): Registrierte Stoffe. <https://echa.europa.eu/de/information-on-chemicals/registered-substances> (abgerufen 27.2.2018).
- ECHA (European Chemicals Agency) (2018b): Vorregistrierte Stoffe. <https://echa.europa.eu/de/information-on-chemicals/pre-registered-substances> (abgerufen 27.2.2018).
- ECHA (European Chemicals Agency) (2018c): REACH study results are published. https://iuclid6.echa.europa.eu/view-article/-/journal_content/title/iuclid-6-version-1-2-0-is-availab-2 (abgerufen 27.2.2018).
- EEA (European Environment Agency) (1998): *Chemicals in the European Environment: Low Doses, High Stakes?* Copenhagen. <http://www.eea.europa.eu/publications/NYM2> (abgerufen 27.2.2018).
- Jonker, Michiel T. O./van der Heijden, Stephan A. (2007): Bioconcentration factor hydrophobicity cutoff: an artificial phenomenon reconstructed. In: *Environmental Science and Technology* 41, 7363–7369.
- MacLeod, Matthew/Fraser, Alison/Mackay, Donald (2002): Evaluating and expressing the propagation of uncertainty in chemical fate and bioaccumulation models. In: *Environmental Toxicology and Chemistry* 21, 700–709.
- Mayer, Philipp/Reichenberg, Fredrik (2006): Can highly hydrophobic organic substances cause aquatic baseline toxicity and can they contribute to mixture toxicity? In: *Environmental Toxicology and Chemistry* 25, 2639–2644.

- Morgan, M. Granger (2001): The neglected art of bounding analysis. In: *Environmental Science and Technology* 35, 162A–164A.
- Ng, Carla/Scheringer, Martin/Fenner, Kathrin/Hungerbühler, Konrad (2011): A framework for evaluating the contribution of transformation products to chemical persistence in the environment. In: *Environmental Science and Technology* 45, 111–117.
- OECD (Organisation for Economic Co-Operation and Development) (2016): Chemical safety and biosafety. Paris. <http://www.oecd.org/chemicalsafety/> (abgerufen 27.2.2018).
- Pontolillo, James/Eganhouse, Robert (2001): The search for reliable aqueous solubility (S_w) and octanol-water partition coefficient (K_{ow}) data for hydrophobic organic compounds: DDT and DDE as a case study. Reston. <http://pubs.water.usgs.gov/wri01-4201/> (abgerufen 27.2.2018).
- Scheringer, Martin (2012): Umweltchemikalien 50 Jahre nach *Silent Spring*: ein ungelöstes Problem. In: *GAIA* 21, 210–216.
- Scheringer, Martin (2013): Fragile Evidenz: Datenprobleme in der Risikobewertung für Chemikalien. In: *Technikfolgenabschätzung – Theorie und Praxis* 22.3, 25–30.
- Scheringer, Martin (2015): Multimedia mass- balance models for chemicals in the environment. Reliable tools or bold oversimplifications? In: *Integrated Environmental Assessment and Management* 11, 177–178.
- Scheringer, Martin (2017): Environmental chemistry and ecotoxicology: in greater demand than ever. In: *Environmental Sciences Europe* 29.3.
- Scheringer, Martin/Trier, Xenia/Cousins, Ian/de Voogt, Pim/Fletcher, Tony/Wang, Zhanyun/Webster, Thomas (2014): Helsingør Statement on poly- and perfluorinated alkyl substances (PFASs). In: *Chemosphere* 114, 337–339 [open access].
- Springer, Andrea/Herrmann, Henning/Sittner, Dana/Herbst, Uta/Schulte, Agnes (2015): REACH compliance. Data availability of REACH registrations part 1: screening of chemicals > 1000 tpa. UBA Texte 43/2015. Dessau-Roßlau.
- Stieger, Greta/Scheringer, Martin/Ng, Carla/Hungerbühler, Konrad (2014): Assessing the persistence, bioaccumulation potential and toxicity of brominated flame retardants. Data availability and quality for 36 alternative brominated flame retardants. In: *Chemosphere* 115, 118–123.
- Stempel, Sebastian/Scheringer, Martin/Ng, Carla/Hungerbühler, Konrad (2012): Screening for PBT chemicals among the “existing” and “new” chemicals of the EU. In: *Environmental Science and Technology* 46, 5680–5687.

- UBA (Umweltbundesamt) (2012): Pressemitteilung 20/2012. Dessau-Roßlau. <https://www.umweltbundesamt.de/presse/presseinformationen/informationslage-zu-chemikalien-verbessert> (abgerufen 27.2.2018).
- UNEP/WHO (United Nations Environment Programme/World Health Organization) (2013): State of the science of endocrine disrupting chemicals – 2012. Genf.
- US EPA (United States Environmental Protection Agency) (2016): EPI Suite – Estimation Program Interface. Washington D.C. <https://www.epa.gov/tsca-screening-tools/epi-suite-estimation-program-interface> (abgerufen 27.2.2018).
- van Leeuwen, Cornelis J./Vermeire, Theo G. (2007): Risk Assessment of Chemicals. An Introduction. Dordrecht.
- Wang, Zhanyun/Cousins, Ian/Scheringer, Martin/Buck, Robert/Hungerbühler, Konrad (2014): Global emission inventories for C₄–C₁₄ perfluoroalkyl carboxylic acid (PFCA) homologues from 1951 to 2030, Part I: production and emissions from quantifiable sources. In: Environment International 70, 62–75.

