

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.

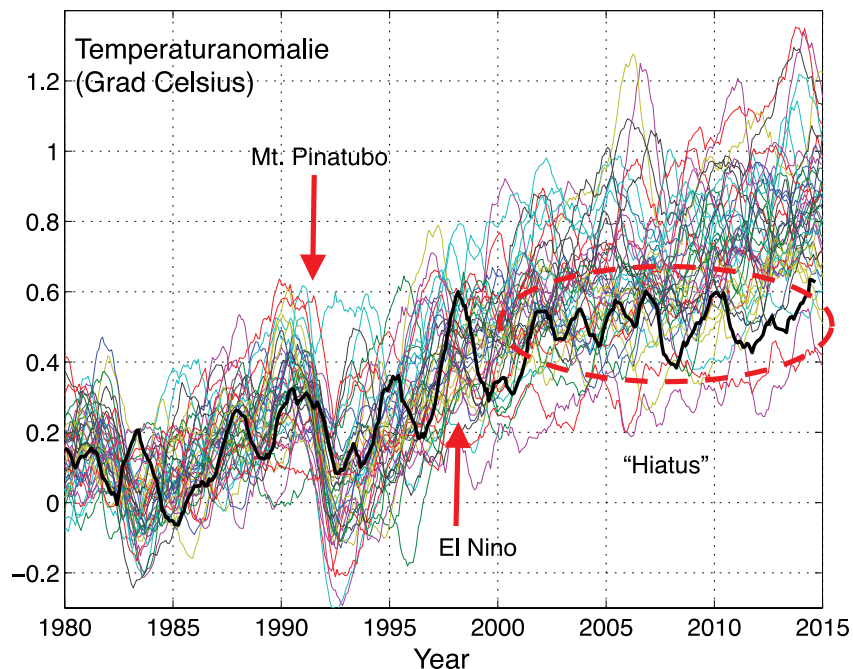
Klima- oder Erdsystemmodelle sind im Allgemeinen deterministisch formuliert, d. h., die Modelle beinhalten keine stochastischen Elemente (obwohl ein Einsatz solcher Elemente in einigen Bereichen vielversprechend erscheint, vgl. z. B. Palmer 2014). Unsicherheiten in den Modellergebnissen sind damit für 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.