

2. Philosophie der Klimawissenschaften

*Richard Bradley, Roman Frigg, Katie Steele, Erica Thompson
und Charlotte Werndl¹*

1 Einführung

Klimawissenschaften ist ein Sammelbegriff, der sich auf wissenschaftliche Disziplinen, die Aspekte des Erdklimas studieren, bezieht. Darunter fallen unter anderem Teile der Atmosphärenwissenschaften, Ozeanographie und Glaziologie. Als Folge der öffentlichen Diskussion über angemessene Reaktionen auf den Klimawandel sind Teile der Entscheidungstheorie und der Ökonomie für Klimafragen herangezogen worden, weswegen Beiträge aus diesen Gebieten auch zu den Klimawissenschaften im weiteren Sinne gezählt werden. Im Zentrum der Philosophie der Klimawissenschaften steht die Betrachtung der Methodologie, die zum Erreichen von Schlussfolgerungen bezgl. des Klimas verwendet wird. Die Philosophie der Klimawissenschaften ist ein Teilgebiet der Wissenschaftstheorie, das sich zu Beginn des 21. Jahrhunderts (als Wissenschaftstheoretiker begonnen haben, sich genauer mit den Methoden der Klimawissenschaften auseinanderzusetzen) herauskristallisiert hat. Mittlerweile beinhaltet sie Betrachtungen über beinahe alle Aspekte der Klimawissenschaften, inklusive der Beobachtung von Daten, Methoden des Nachweises und der Zuschreibung, Modell-Ensembles und Entscheidungsfindung bei Unsicherheit. Da der Teufel bekanntlich im Detail steckt, operiert die Philosophie der Klimawissenschaften in engem Kontakt mit den Wissenschaften selbst und richtet viel Aufmerksamkeit auf wissenschaftliche Details. Aus diesem Grund gibt es keine klare Trennung zwischen den Klimawissenschaften und der Philosophie derselben und Konferenzen werden oft sowohl von Wissenschaftler/innen wie auch von Philosophen/innen besucht.

Ziel dieses Artikels² ist es, eine aktuelle Zusammenfassung der wichtigsten Probleme und Fragen in den Grundlagen der Klimawissenschaften zu liefern.

¹ Die Autoren sind in alphabetischer Ordnung aufgelistet. Frigg, Thompson und Werndl sind für die Sektionen 2 bis 6 verantwortlich, Bradley und Steele für die Sektionen 7 und 8.

² Wir danken Reto Knutti, Wendy Parker, Lenny Smith, Dave Stainforth und zwei anonymen Gutachtern für wertvolle Diskussionen und Vorschläge. Außerdem danken wir Markus Hierl, welcher bei der Übersetzung und Formatierung des Artikels mitwirkte. Diese Forschung wurde vom britischen Arts and Humanities Research Council (grant number AH/J006033/1) und dem ESRC Centre for Climate Change Economics and Policy, welches durch das Economic and Social Research Council (Grant Nummer ES/

Im 2. Kapitel sprechen wir über das Problem, Klima zu definieren. Im 3. Kapitel stellen wir Klimamodelle vor. Im 4. Kapitel diskutieren wir Probleme des Nachweises und der Zuschreibung von Klimawandel. Im 5. Kapitel untersuchen wir die Bestätigung von Klimamodellen und die Grenzen der Vorhersagbarkeit. Im 6. Kapitel besprechen wir Klassifikationen von Unsicherheit und die Verwendung von Modell-Ensembles. Im 7. Kapitel wenden wir uns der Entscheidungstheorie zu und diskutieren, in welchem theoretischen Rahmen Klima-Entscheidungsprobleme diskutiert werden sollen. Im 8. Kapitel stellen wir unterschiedliche Entscheidungsregeln vor. Im 9. Kapitel bieten wir einen Ausblick an.

Zwei Klarstellungen seien vorangestellt. Erstens, wir besprechen Probleme und Fragen von einer wissenschaftstheoretischen Perspektive aus und mit speziellem Fokus auf erkenntnistheoretische und entscheidungstheoretische Probleme. Natürlich ist dies nicht die einzige Perspektive. Man kann die Klimawissenschaften auch aus dem Blickwinkel der Wissenschaftsforschung, der Soziologie der Wissenschaften, der politischen Theorie und Ethik betrachten. Aus Platzgründen können wir auf diese Gebiete nicht eingehen.

Zweitens sollte, um möglichen Missverständnissen vorzubeugen, darauf hingewiesen werden, dass eine kritische philosophische Reflexion über die Ziele und Methoden der Klimawissenschaften keinesfalls gleichbedeutend mit dem Einnehmen einer Position ist, die als Klimaskeptizismus bekannt ist. Klimaskeptiker/innen sind eine heterogene Gruppe von Leuten, welche die Ergebnisse der »Mainstream«-Klimawissenschaften nicht akzeptieren. Die Gruppe umfasst ein breites Spektrum, angefangen von denjenigen, welche schlichtweg die grundlegende Physik des Treibhauseffekts (und den Einfluss von menschlichen Aktivitäten auf das Klima der Erde) leugnen, bis zu einer kleinen Minderheit, die sich aktiv an der wissenschaftlichen Forschung und Diskussion beteiligt und zu Konklusionen gelangt, in denen der Klimawandel als sehr gering eingeschätzt wird. Kritische Wissenschaftstheorie ist keine Handlangerin des Klimaskeptizismus. Darum sollte betont werden, dass wir dem Klimaskeptizismus nicht zustimmen. Unser Ziel ist es, zu verstehen, wie die Klimawissenschaften funktionieren, und über ihre Methoden zu reflektieren.

2 Definitionen von Klima und Klimawandel

Diskussionen über das Klima sind sowohl in den populären Medien als auch im akademischen Bereich allgegenwärtig. Dies verschleiert die Tatsache, dass *Klima* ein komplexer Begriff ist, und die Definition von Klima und Klimawandel Gegenstand wissenschaftlicher Kontroversen ist. Um ein Verständnis des Begriffs

K006576/1) finanziert wird, unterstützt. Professor Friggs Forschung wurde des Weiteren auch durch das Spanische Ministerium für Wissenschaft und Innovation (Grant Nummer FFI2012-37354) gefördert.

Klima zu erlangen, ist es wichtig, ihn vom Begriff Wetter zu unterscheiden. Intuitiv gesprochen ist das Wetter an einem bestimmten Ort zu einer bestimmten Zeit der Zustand der Atmosphäre an diesem Ort zu dieser Zeit. Beispielsweise ist das Wetter in Berlin-Mitte am 1. Januar 2015 um 14:00 dadurch charakterisiert, dass die Temperatur zwölf Grad beträgt, die Luftfeuchtigkeit bei 65 Prozent liegt, etc. Im Kontrast dazu liefert Klima eine summarische Aussage über Wetterbedingungen: Es ist eine Verteilung von bestimmten Variablen (die sogenannten Klimavariablen), welche sich aus einer bestimmten Konfiguration des Klimasystems ergibt.

Die Frage ist, wie man diese Idee präzisieren kann. In der aktuellen Diskussion findet man zwei generelle Ansätze. Beim ersten wird Klima als *Verteilung über eine Zeitspanne* definiert, und beim zweiten wird Klima als *Ensemble-Verteilung* definiert. Zu den Klimavariablen zählen bei beiden Ansätzen jene, die den Zustand der Atmosphäre und des Ozeans beschreiben, und manchmal auch Variablen wie jene, welche den Zustand von Gletschern und Eisdecken beschreiben.

Verteilung über eine Zeitspanne. Der Wert der Klimavariablen hängt von externen Bedingungen ab, wie z. B. der von der Sonne empfangenen Energie und Vulkanaktivitäten. Wir nehmen nun an, dass in einer gewissen Zeitspanne diese externen Bedingungen als konstant behandelt werden können, weil diese nur kleine Schwankungen um einen Mittelwert c aufweisen. Man kann Klima dann für diese Zeitspanne als Verteilung der Klimavariablen während dieser Zeitspanne unter konstanten Bedingungen c definieren (z. B. Lorenz 1995). Unter Klimawandel versteht man dann die Tatsache, dass aufeinanderfolgende Zeitspannen durch verschiedene Verteilungen charakterisiert werden. Allerdings sind die externen Bedingungen in der Realität nicht konstant, und selbst wenn es nur leichte Schwankungen um c gibt, kann es sein, dass die dadurch entstehenden Verteilungen sehr unterschiedlich sind. Aus diesem Grund ist diese Definition unbefriedigend.

Dieses Problem kann vermieden werden, wenn man Klima als empirisch beobachtbare Verteilung über eine gewisse Zeitspanne definiert, wobei die externen Bedingungen variieren können. Unter Klimawandel versteht man dann auch wieder, dass aufeinanderfolgende Zeitspannen durch verschiedene Verteilungen charakterisiert werden. Diese Definition ist populär, weil das Klima aufgrund von Beobachtungen berechnet werden kann, wie z. B. von den Zeitreihen der letzten 30 Jahre (Hulme et al. 2009). Ein Problem dieser Definition kann an einem Beispiel illustriert werden: In der Mitte einer Zeitspanne wird die Erde von einem Meteoriten getroffen, wodurch sie zu einem viel kälteren Ort wird. Offensichtlich unterscheidet sich das Klima vor dem Einschlag von dem danach. Dieser Klimawandel kann aber durch diese Definition nicht eingefangen werden, da sie Klima als Verteilung der tatsächlichen Werte der Klimavariablen über die Zeitspanne definiert und deshalb gegenüber plötzlichen radikalen Veränderungen blind ist. Um dieses Problem zu umgehen, führt Werndl (2015) die Idee von Systemen mit variierenden externen Bedingungen ein und schlägt vor, Klima als die Verteilung

der Klimavariablen über die Zeit, die unter bestimmten Regimen von variierenden externen Bedingungen auftreten, zu definieren.

Ensemble-Verteilung. Ein Ensemble von Klimasystemen (nicht zu verwechseln mit einem Ensemble von Modellen, dem wir uns weiter unten zuwenden) ist eine Sammlung von virtuellen Kopien des Klimasystems. Nun betrachten wir das Sub-Ensemble mit denjenigen Kopien, in denen die Werte der Klimavariablen in einem bestimmten Intervall um die im aktuellen Klimasystem *gemessenen Werte* liegen (das sind die Werte, welche mit der Messgenauigkeit kompatibel sind). Zudem betrachten wir wieder eine Zeitspanne, während der die externen Bedingungen relativ stabil sind und nur kleine Schwankungen um einen Mittelwert c aufweisen. Das Klima eines zukünftigen Zeitpunktes t wird dann häufig als die Verteilung der Werte der Klimavariablen definiert, welche auftritt, wenn alle Systeme im Sub-Ensemble sich von jetzt bis t unter konstanten externen Bedingungen c entwickeln (z. B. Lorenz 1995). Anders ausgedrückt: Das Klima in der Zukunft ist die Verteilung der möglichen Werte der Klimavariablen, welche mit den *aktuellen Beobachtungen unter der Annahme von konstanten externen Bedingungen* c konsistent sind.

Allerdings sind die externen Bedingungen in der Realität nicht konstant. Selbst wenn es nur kleine Schwankungen um den Mittelwert gibt, kann dies zu verschiedenen Verteilungen führen (Werndl 2015). Deshalb liegt die Modifikation nahe, dass man die Entwicklung des Ensembles unter den *aktuellen* externen Bedingungen studiert. Das Klima zu einem zukünftigen Zeitpunktes t ist dann die Verteilung der Klimavariablen, welche auftritt, wenn das Ensemble sich vorwärts zum Zeitpunkt t entwickelt (unter den tatsächlichen externen Bedingungen).

Diese Definition hat mehrere begriffliche Probleme. Erstens macht sie das Klima von unserem Wissen abhängig (durch Messgenauigkeit), was aber unseren Intuitionen widerspricht, denn wir verstehen Klima als etwas Objektives und unabhängig von unserem Wissen Existierendes. Zweitens ist die obige Definition eine Definition von *zukünftigem* Klima und es ist schwer zu erkennen, wie das gegenwärtige oder vergangene Klima definiert werden soll. Ohne einen Begriff von gegenwärtigem und vergangenem Klima lässt sich jedoch Klimawandel nicht definieren. Ein drittes Problem ist, dass Ensemble-Verteilungen nicht mit vergangenen Beobachtungsreihen in Verbindung stehen, was impliziert, dass das Klima nicht mit ihnen berechnet werden kann.

3 Klimamodelle

Ein Klimamodell ist eine Repräsentation von bestimmten Aspekten des Klimasystems. Eines der einfachsten Klimamodelle ist das Energiebilanz-Modell, welches die Erde als Linie mit eindimensionaler Atmosphäre betrachtet. Es basiert auf dem einfachen Prinzip, dass im Gleichgewichtszustand die einfallende und die ausfallende Energie gleich sein müssen (für eine Diskussion solcher Modelle

siehe Dessler 2011: Kap. 3–6). Dieses Modell kann durch eine Einteilung der Erde in Zonen (wobei der Energietransfer zwischen den Zonen erlaubt ist) oder durch eine Einführung eines Vertikalprofils der atmosphärischen Charakteristika verfeinert werden. Trotz ihrer Einfachheit liefern diese Modelle ein gutes qualitatives Verständnis des Treibhauseffektes.

Die modernen Klimawissenschaften zielen darauf ab, Modelle zu konstruieren, welche möglichst viel vom aktuellen Wissensstand integrieren (für eine Einführung in Klimamodellierung siehe McGuffie/Henderson-Sellers 2005). Normalerweise geschieht dies durch Einteilung der Erde (Atmosphäre und Ozean) in Gitterzellen. Aktuelle Klimamodelle arbeiten mit einer horizontalen Gitterskala von etwa 150 km. Klimaprozesse können dann durch Flüsse zwischen diesen Zellen beschrieben werden, wie z. B. Hitze oder Dampf von einer Zelle zur anderen. Diese Flüsse werden mathematisch durch Gleichungen beschrieben. Diese Gleichungen bilden den »dynamischen Kern« eines sogenannten »Global Circulation Model« (GCM). Diese Gleichungen können normalerweise nicht analytisch gelöst werden, weshalb leistungsstarke Supercomputer verwendet werden, um sie numerisch zu integrieren. Aus diesem Grund werden sie oft als Simulationsmodelle bezeichnet. Um die Gleichungen numerisch lösen zu können, wird die Zeit diskretisiert. Gegenwärtige Simulationen verwenden diskrete Zeitschritte von ca. 30 Minuten. Auf Supercomputern dauert es Wochen oder Monate, um ein Jahrhundert Klimaentwicklung zu simulieren.

Um eine hypothetische Entwicklung des Klimasystems zu berechnen, benötigen wir auch *Anfangs-* und *Randbedingungen*. Die ersteren sind mathematische Beschreibungen vom Zustand des Klimasystems zu Beginn der simulierten Periode. Die letzteren sind Werte für alle Variablen, die das System betreffen, aber nicht direkt berechnet werden. Diese beinhalten z. B. die Konzentration von Treibhausgasen in der Atmosphäre zu einem bestimmten Zeitpunkt, die Konzentration von Aerosolen und die Menge der Sonneneinstrahlung auf der Erde. Weil dies Schlüsselfaktoren des Klimawandels sind, werden sie oft als *externe Faktoren* oder *externe Bedingungen* bezeichnet.

Wenn Prozesse auf einer kleineren Skala als dem Gitter auftreten, dann können sie im Modell durch sogenannte Parametrisierungen berücksichtigt werden, wobei das Nettoergebnis des Prozesses gesondert als eine Funktion der Gitter-Variablen berechnet wird. Wolkenbildung beispielsweise ist ein physikalischer Prozess, der nicht direkt simuliert werden kann, weil Wolken typischerweise viel kleiner als eine Gitterzelle sind und das Wissen über Wolkenbildung sehr begrenzt ist. Der Nettoeffekt von Wolken wird deswegen normalerweise in jeder Gitterzelle parametrisiert (als Funktion von Temperatur, Luftfeuchtigkeit etc.) und für die Berechnungen verwendet. Prozesse, die kleiner als eine Gitterzelle sind, sind eine der Hauptquellen für die Unsicherheit in Klimamodellen.

Es gibt zurzeit ca. 20 Klimamodelle, welche kontinuierlich von nationalen Modellierungszentren wie der NASA, dem britischen Met Office und dem Beijing Climate Center weiterentwickelt werden. Um die Ergebnisse dieser unterschied-

lichen Modelle zu vergleichen, hat das sogenannte »Coupled Model Intercomparison Project« (CMIP) eine Reihe von Standardexperimenten definiert, die von jedem Klimamodell durchlaufen werden. Eines der Standardexperimente ist, jedes Modell unter Verwendung der Faktoren, die während des 20. Jahrhunderts vorherrschend waren, laufen zu lassen. Dies erlaubt einen direkten Vergleich der Simulation mit den wirklichen Klimadaten.

Klimamodelle werden an vielen Stellen in den Klimawissenschaften verwendet und ihre Verwendung wirft wichtige Fragen auf. Diese Fragen besprechen wir in den nächsten drei Kapiteln.

4 Nachweis und Zuschreibung von Klimawandel

Jede empirische Studie des Klimas beginnt mit der Beobachtung des Klimas. Meteorologische Beobachtungen messen allerlei Variablen wie z.B. die Lufttemperatur nahe der Erdoberfläche unter Verwendung von Thermometern. Jedoch wurden mehr oder weniger systematische Beobachtungen erst seit ca. 1750 gemacht. Um das Klima davor zu rekonstruieren, müssen Wissenschaftler/innen sich auf *Klimaproxys* verlassen. Dies sind Daten für Klimavariablen, welche aus anderen natürlichen Phänomenen wie z.B. Baumringen, Eisbohrkernen und Ozeansedimenten abgeleitet werden.

Die Verwendung von Klimaproxys ruft eine Reihe von methodologischen Problemen hervor. Insbesondere gaben sogenannte *Hockeyschläger-Diagramme*, welche eine auf Klimaproxys basierende Rekonstruktion des Temperaturverlaufs der nördlichen Hemisphäre darstellen, Anlass zu hitzigen Debatten (Mann, Bradley/Hughes 1998). Klimaskeptiker/innen verfolgen zwei Argumentationslinien. Sie bezweifeln die Verlässlichkeit der verfügbaren Daten und behaupten, dass die zur Gewinnung der Daten verwendeten Methoden eine Kurve in Hockeyschlägerform produziert hätten, ganz egal welche Daten man verwendet hätte. Die von den Skeptikern/innen veröffentlichten Artikel haben wichtige Probleme aufgeworfen und weitere Untersuchungen angeregt, es wurde aber festgestellt, dass ihren Schlussfolgerungen gravierende Fehler zugrunde liegen. Mittlerweile gibt es mehr als zwei Dutzend Rekonstruktionen dieses Temperaturverlaufs, die unterschiedliche statistische Methoden und Klimaproxy-Quellen verwenden. Obwohl es aufgrund der begrenzten Daten und Methoden tatsächlich eine Vielfalt an möglichen Temperaturkurven gibt, unterstützen diese Studien den Konsens, dass die Temperaturen während des späten 20. Jahrhunderts sehr wahrscheinlich die wärmsten der letzten 1400 Jahre sind (Frank et al. 2010).

Zeigen steigende Temperaturen an, dass es einen Klimawandel gibt, und falls ja, kann der Klimawandel menschlichen Handlungen zugeschrieben werden? Diese beiden Probleme sind als *Nachweis* und *Zuschreibung* bekannt. Das »Intergovernmental Panel on Climate Change« (IPCC) definiert diese folgendermaßen:

»Ein Nachweis des Klimawandels ist definiert als eine Demonstration, dass sich das Klima, oder ein anderes vom Klima beeinflusstes System, in einem statistisch beschreibbaren Sinn verändert hat, ohne einen Grund für diese Veränderung anzugeben. Eine bestimmte Veränderung wird als in Beobachtungen nachgewiesen betrachtet, wenn die Wahrscheinlichkeit, dass die Veränderung durch Zufall oder aufgrund der internen Variabilität aufgetreten ist, als klein eingestuft werden kann.« [...] Zuschreibung ist definiert als »der Prozess der Evaluation der relativen Beiträge verschiedener kausaler Faktoren zu einer bestimmten Veränderung oder einem Ereignis, wobei eine statistische Konfidenz anzugeben ist« (IPCC 2013; Übersetzung der Autoren/innen).

Diese Definitionen werfen mehrere Probleme auf. Die Wurzel der Probleme ist die Aussage, dass Klimawandel nur dann nachgewiesen worden ist, wenn eine beobachtete Veränderung im Klima sehr wahrscheinlich nicht aufgrund von interner Variabilität erfolgt ist. *Interne Variabilität* ist das Phänomen, dass Klimavariablen wie Temperatur und Niederschlag sich aufgrund von internen Dynamiken des Klimasystems über die Zeit verändern: Es würde wärmere und kältere Jahrzehnte und Jahrhunderte geben, auch wenn es überhaupt keine Menschen gäbe.

Diese Definition von Nachweis hat die Konsequenz, dass es keinen internen Klimawandel geben kann. Die Eiszeiten beispielsweise würden nicht als Klimawandel gelten, wenn sie aufgrund von interner Variabilität entstanden wären. Dies widerspricht nicht nur grundsätzlichen Intuitionen über das Klima und den gängigsten Definitionen von Klima als eine endliche Verteilung über eine relativ kurze Zeitspanne (wo interner Klimawandel möglich ist); es führt auch zu Problemen in Hinblick auf Zuschreibung: Wenn nachgewiesener Klimawandel *ipso facto* nicht von interner Variabilität herrührt, dann ist es von Beginn an ausgeschlossen, dass bestimmte Faktoren (nämlich interne Klimadynamiken) zur Veränderung des Klimas führen können.

Im Fall der Eiszeiten würden wohl viele betonen, dass interne Variabilität zu unterscheiden ist von natürlicher Variabilität. Weil die Orbitalbewegung die Eiszeiten erklärt und die Orbitalbewegung etwas Natürliches, aber Externes ist, ist dies ein Fall von externem Klimawandel. Dadurch werden zwar einige Probleme gelöst, aber das Problem, dass es keinen allgemein akzeptierten Weg gibt, wie man interne von externen Faktoren unterscheidet, bleibt bestehen. Derselbe Faktor wird manchmal als intern und manchmal als extern klassifiziert. Vergletscherung beispielsweise wird manchmal als interner Faktor und manchmal als externer Faktor behandelt. Ebenso wird die Biosphäre manchmal als externer Faktor behandelt, manchmal jedoch auch intern modelliert. Man könnte sogar so weit gehen, zu fragen, ob menschliche Handlungen ein externer Faktor für das Klimasystem sind. Studien behandeln menschliche Handlungen zwar normalerweise als externe Faktoren, aber es könnte konsistent dafür argumentiert werden, dass sie ein interner dynamischer Prozess sind. Die angemessene Definition hängt schlicht und einfach davon ab, für welche Fragen man sich interessiert.

Die Effekte der internen Variabilität sind auf allen Zeitskalen gegenwärtig: von Schwankungen, die kürzer als einen Tag dauern und als Wetter diskutiert werden, bis zu langfristigen Veränderungen aufgrund der Zyklen der Vergletscherung. Da interne Variabilität das Ergebnis eines hochkomplexen nichtlinearen Systems ist, ist es auch unwahrscheinlich, dass die statistischen Eigenschaften von interner Variabilität über die Zeit hinweg konstant sind. Aktuelle Klimamodelle zeigen sowohl in der Größe der internen Variabilität als auch was die Zeitskala von Veränderungen betrifft signifikante Unterschiede. Schätzungen der internen Variabilität im Klimasystem werden von Klimamodellen selbst produziert (Hegerl et al. 2010), was zu einer möglichen Zirkularität führt. Dies unterstreicht die Schwierigkeiten von Aussagen bzgl. Zuschreibungen, die sich auf die obige Definition stützen, welche einen beobachteten Klimawandel nur dann anerkennt, wenn er sehr wahrscheinlich nicht auf interner Variabilität basiert.

Weil die Definitionen des IPCC von vielen Klimawissenschaftlern/innen verwendet werden, wird die weitere Diskussion auf diesen Definitionen basieren. Nachweis des Klimawandels stützt sich auf statistische Tests. Studien zum Nachweis betreffen oft die Wahrscheinlichkeit eines bestimmten Ereignisses oder einer Abfolge von Ereignissen, welche man in Abwesenheit von Klimawandel vorfinden würde. In der Praxis ist es eine Herausforderung, eine angemessene Nullhypothese (das erwartete Verhalten des Systems in Abwesenheit von sich verändernden externen Faktoren) zu definieren, mit welchen dann die Beobachtungen verglichen werden können. Dies stellt keine triviale Aufgabe dar, da das Klimasystem ein dynamisches System mit nichtlinearen Prozessen und Feedbacks ist, welche auf allen Skalen aktiv sind. Die Wichtigkeit der Nullhypothese wird durch die Studie von Cohn/Lins (2005) illustriert, welche dieselben Daten mit verschiedenen Nullhypothesen vergleichen. Die Ergebnisse unterscheiden sich in ihrer Signifikanz in 25 Größenordnungen. Dies zeigt nicht, dass nur gewisse Nullhypothesen angemessen sind, aber es veranschaulicht die Sensibilität des Resultats in Bezug auf die gewählte Nullhypothese. Dies wiederum unterstreicht, wie wichtig die Wahl der Nullhypothese ist und wie schwierig es ist, eine derartige Wahl zu treffen, wenn die zugrundeliegenden Prozesse nur schlecht verstanden werden.

In der Praxis wird die beste verfügbare Nullhypothese oft vom besten verfügbaren Modell abgeleitet. Dieses Modell wird verwendet, um *Kontrolldurchläufe* durchzuführen: Unter konstanten externen Bedingungen wird dabei die interne Variabilität des Modells abgeschätzt. Klimawandel wird dann als nachgewiesen betrachtet, wenn die beobachteten Werte außerhalb eines im Vorhinein festgelegten Bereichs der internen Variabilität des Modells fallen. Die Schwierigkeit bei dieser Methode ist, dass es nicht ein einziges »bestes« Modell zur Auswahl gibt: Viele solche Modelle existieren, die in vielerlei Hinsicht ähnlich gut sind, aber sie können, wie oben schon erwähnt, zu sehr unterschiedlicher interner Variabilität führen.

Die Unterschiede zwischen verschiedenen Modellen sind für die eindeutigsten Resultate, wie etwa die aktuelle Steigerung der globalen Durchschnittstemperatur, relativ unwichtig. Hier ist, wie von Parker (2010) betont wurde, der Nachweis über verschiedene Modelle hinweg robust (in Kapitel 6 wird Robustheit näher behandelt werden) und es gibt eine Vielfalt an Evidenz dafür, dass sich die globale Durchschnittstemperatur erhöht hat. Die Frage, welche Nullhypothese man verwendet und wie interne Variabilität quantifiziert wird, kann dennoch wichtig für den Nachweis von kleineren und lokalen Klimaveränderungen sein.

Wenn Klimawandel nachgewiesen worden ist, dann stellt sich die Frage der Zuschreibung. Dies kann eine Zuschreibung entweder einer direkten Klimaveränderung (wie etwa einer erhöhten globalen Durchschnittstemperatur) oder einer Auswirkung einer Klimaveränderung (wie z. B. der von Waldbränden vernichteten Fläche) zu einer identifizierten Ursache (wie z. B. einem erhöhten CO₂-Gehalt in der Atmosphäre) sein. Wo eine indirekte Einwirkung betrachtet wird, ist ein zwei- oder mehrstufiger Ansatz möglicherweise angemessen, welcher zuerst eine direkte Klimaveränderung ausfindig macht und dann im zweiten Schritt die indirekte Klimaveränderung als deren Konsequenz zuschreibt. Ein Beispiel dafür, welches dem IPCC Good Practise Guidance Artikel (Hegerl et al. 2010) entnommen ist, ist die Zuschreibung der Auswirkung von Verkalkungen der Korallenriffe auf den CO₂-Gehalt. In dieser Zuschreibung wird eine Zwischenstufe verwendet: Zuerst werden die Veränderungen der Karbon-Eisen-Konzentration der Erhöhung des CO₂-Gehalts und dann die Verkalkungen den Veränderungen in der Karbon-Eisen-Konzentration zugeschrieben. Dies zeigt die Notwendigkeit eines klaren Verständnisses der beteiligten physikalischen Mechanismen auf, um eine verlässliche Zuschreibung durchführen zu können.

Statistische Analysen messen die Stärke der Beziehung unter den vereinfachenden Annahmen, welche dem Verfahren der Zuschreibung zu Grunde liegen, aber der Konfidenzgrad in diese vereinfachenden Annahmen muss außerhalb des Rahmens bewertet werden. Dieser Konfidenzgrad ist von dem IPCC in diskrete (aber subjektive) Kategorien standardisiert worden: fast sicher (> 99%), extrem wahrscheinlich (> 95%), sehr wahrscheinlich (> 90%), etc. Das Ziel ist, die begrenzte Datenverfügbarkeit, die Adäquatheit des verwendeten Modells und die Präsenz von verzerrenden Faktoren zu berücksichtigen. Die Schlussfolgerungen werden dann entsprechend qualifiziert. Ein Beispiel ist die folgende Aussage:

»Es ist extrem wahrscheinlich, dass mehr als die Hälfte der beobachteten Erhöhung der globalen Oberflächendurchschnittstemperatur von 1951 bis 2010 durch den anthropogenen Zuwachs an der Treibhausgaskonzentration und andere anthropogene Faktoren verursacht ist.« (IPCC 2013: Summary for Policymakers: section D.3; Übersetzung der Autoren/innen).

Der sogenannte *Optimale Fingerabdruck* ist eine Methode, um ein solches Resultat zu erreichen. Diese Methode assoziiert mit jedem potentiellen Schlüsselfaktor (z. B. Treibhausgasen oder der Veränderung der Sonneneinstrahlung) ein

raumzeitliches Muster von Veränderung (Fingerabdruck), normalisiert relativ zur internen Variabilität. Dieses Muster wird mit einem Klimamodell berechnet. Danach wird eine statistische Regression der beobachteten Daten im Hinblick auf lineare Kombination dieser Muster durchgeführt. Die restliche Variabilität, nachdem die Beobachtungen dem jeweiligen Faktor zugeschrieben wurden, sollte mit der internen Variabilität konsistent sein; wenn nicht, dann deutet dies darauf hin, dass eine wichtige Ursache für Variabilität nicht berücksichtigt wurde.

Wie von Parker (2010) betont wurde, beruhen Fingerabdruck-Studien auf mehreren Annahmen. Die erste ist Linearität, d. h., dass die Reaktion des Klimasystems, wenn mehrere Faktoren gegenwärtig sind, gleich der linearen Kombination dieser Faktoren ist. Da das Klimasystem nichtlinear ist, ist dies offensichtlich eine Quelle von methodologischen Schwierigkeiten, obwohl für Prozesse auf globaler Ebene (im Kontrast zu Prozessen auf regionaler Ebene) gezeigt wurde, dass die Additivität eine gute Annäherung ist. Eine andere Annahme ist, dass Klimamodelle die kausalen Prozesse, welche im Klimasystem der Erde auftreten, genau genug simulieren. Der Erfolg von Fingerabdruck-Studien (welcher nicht trivial ist) kann dabei helfen, Bedenken über diese Annahmen zu entkräften, wie Parker auch argumentiert. Ein weiteres Problem ist einmal mehr die Notwendigkeit, die Charakteristiken von interner Variabilität zu definieren. Wie schon besprochen, wird in der Praxis interne Variabilität mit Hilfe der besten Klimamodelle abgeschätzt, welches die Problematik der Zirkularität aufwirft.

In Aussagen über Zuschreibung ist der Konfidenzgrad in erster Linie vom physikalischen Verständnis der involvierten Prozesse abhängig. Wo es einen klaren, einfachen und wohlverstandenen Mechanismus gibt, sollte es größeres Vertrauen in die statistischen Ergebnisse geben; wo die Mechanismen unzusammenhängend oder mehrstufig sind oder wo ein komplexes Modell als Bindungsglied verwendet wird, ist das Vertrauen dementsprechend niedriger. Das Guidance Paper warnt:

»Wo Modelle in der Zuschreibung verwendet werden, da sollte die Fähigkeit des Modells, kausale Zusammenhänge realistisch zu repräsentieren, beurteilt werden. Das sollte eine Beurteilung verzerrender Tendenzen sowie der Fähigkeit des Modells, die relevanten Prozesse auf den relevanten Skalen darzustellen, beinhalten.«
(Hegerl 2010, 5; Übersetzung der Autoren/innen)

Ebenso gibt es ein größeres Vertrauen in die Resultate von Zuschreibungen, wenn die Resultate robust sind und es eine Vielfalt an Evidenz gibt, wie Parker (2010) argumentiert. Beispielsweise hat sich die Erkenntnis, dass der Temperaturanstieg gegen Ende des 20. Jahrhunderts hauptsächlich von Treibhausgasen verursacht worden ist, als robust bei einer großen Auswahl an Modellen, verschiedenen Analysetechniken und verschiedenen Faktoren herausgestellt, und es gibt eine Vielfalt an Evidenz, welche diese Behauptung stützt. Dadurch ist unser Vertrauen hoch, dass Treibhausgase die globale Erwärmung erklären.

Es gibt eine interessante Frage bzgl. des Status von Methoden der Zuschreibung wie der Fingerabdruck-Methode. Der Treibhauseffekt ist gut dokumen-

tiert und in Laborexperimenten einfach zu beobachten. Manche argumentieren, dass dies ein gutes qualitatives Verständnis des Klimasystems liefert, was ausreicht, um mit Zuversicht zu behaupten, dass die globale Erwärmung real ist und dass menschlich bedingtes CO₂ als Ursache für den Anstieg der globalen Durchschnittstemperatur identifiziert worden ist (z. B. Betz 2013). Dann würden statistische Debatten über Methoden der Zuschreibung, welche zeigen, dass vom Menschen verursachte CO₂-Emissionen die globale Erwärmung verursachen, nicht mehr unbedingt benötigt. Diese Argumentationslinie wird aber nicht von allen akzeptiert. Winsberg/Goodwin (2015) etwa argumentieren, dass Fingerabdrücke entscheidend für die Zuschreibung der Erhöhung der globalen Durchschnittstemperatur durch menschlich bedingte CO₂-Emissionen sind.

5 Bestätigung und Vorhersage

Zwei Fragen stehen in Verbindung mit Klimamodellen im Vordergrund: Wie werden Modelle bestätigt und was ist ihre Vorhersagekraft? *Bestätigung* betrifft die Frage, ob und in welchem Ausmaß ein bestimmtes Modell von den Daten bestätigt wird. Lloyd (2009) argumentiert, dass viele Klimamodelle von Daten über die Vergangenheit bestätigt werden. Parker widerspricht dieser Behauptung. Sie argumentiert, dass die Annahme, dass Klimamodelle an sich bestätigt werden können, nicht ernsthaft in Erwägung gezogen werden kann, weil alle Klimamodelle falsch und empirisch nicht adäquat sind. Parker (2009) spricht sich für ein Umdenken von Bestätigung zu *Adäquatheit für einen Zweck* aus: Hypothesen von Klimamodellen können nur für bestimmte Zwecke bestätigt werden. Beispielsweise könnte man behaupten, dass ein gewisses Klimamodell die globale Temperatur bis 2100 bis maximal auf einen Fehler von 0,5 Grad adäquat vorhersagt. Zur gleichen Zeit könnte man jedoch meinen, dass den Vorhersagen des globalen Niederschlagsmittelwerts bis 2100 durch das gleiche Modell nicht getraut werden kann.

Katzav (2014) warnt, dass der Begriff der Adäquatheit für einen Zweck nur begrenzt nützlich ist. Er argumentiert, dass normalerweise nicht bewertet werden kann, ob ein Modell adäquat für einen Zweck ist, weil es höchst unklar ist, welche der beobachtbaren Folgerungen des Modells zeigen könnten, dass das Modell angemessen für einen Zweck ist. Stattdessen behauptet er, dass Klimamodelle bestenfalls über mögliche Zukunftsszenarien Auskunft geben können. Katzav betont zu Recht, dass Abschätzungen, ob Modelle adäquat für einen Zweck sind, schwieriger sind, als sie auf den ersten Blick erscheinen. Aber die Methode der Adäquatheit für einen Zweck sollte nicht als Ganzes verworfen werden; tatsächlich wird sie erfolgreich in verschiedenen Wissenschaften verwendet (z. B. wenn bestätigt wird, dass ideale Gasmodelle nützlich für bestimmte Zwecke sind). Ob Bewertungen der Adäquatheit für einen Zweck möglich sind oder nicht, hängt vom vorliegenden Fall ab.

Wenn man herausfindet, dass ein Modell richtige Vorhersagen macht und ein anderes Modell nicht, will man die Gründe wissen, warum das erste erfolgreich ist und das zweite nicht. Lenhard/Winsberg (2010) argumentieren, dass dies hoffnungslos ist: Für komplexe Klimamodelle ist es aufgrund einer *starken Version des Bestätigungsholismus* unmöglich herauszufinden, wo die Fehler und Erfolge liegen. Insbesondere behaupten sie, dass es unmöglich ist, die Vorzüge und Probleme der Sub-Modelle und Teile der Modelle zu bewerten. Sie belegen ihren Standpunkt mit Fallstudien (z. B. über das Coupled Model Intercomparison Project) und behaupten, dass diese starke Version des Bestätigungsholismus auch in der Zukunft bestehen bleiben wird. Diese Behauptung, und ob dieser Bestätigungsholismus alle Modelle betrifft, ist aber fraglich. Komplexe Modelle haben verschiedene Module für die Atmosphäre, den Ozean und das Eis. Diese Module können individuell und auch zusammen gebraucht werden und es scheint plausibel, dass man so zumindest in einigen Fällen herausfinden kann, welche Vorzüge und Probleme die separaten Module haben.

Ein weiteres Problem betrifft die Verwendung von Daten bei der Konstruktion von Modellen. Die Werte der Parameter von Modellen werden oft aufgrund von Beobachtungen bestimmt, ein Prozess der als *Schätzung* bekannt ist. Beispielsweise wird das Ausmaß der Aerosoleffekte oft mithilfe von Daten geschätzt. Wenn die Daten zur Schätzung verwendet worden sind, stellt sich die Frage, *ob die gleichen Daten nochmal für die Bestätigung des Modells verwendet werden können*. Wenn Daten für die Bestätigung verwendet werden, welche nicht bereits zur Schätzung verwendet worden sind, sind sie *neuwertig*. Wenn Daten sowohl für die Bestätigung als auch die Schätzung verwendet werden, wird dies als *double-counting* bezeichnet.

Wissenschaftler/innen und Philosophen/innen haben gleichermaßen argumentiert, dass double-counting nicht legitim ist und dass die Daten neuwertig sein müssen, um bestätigen zu können (Lloyd 2010; Shakley et al. 1998; Worall 2010; Frisch [i. E.]). Steele/Werndl (2013) widersprechen dieser Schlussfolgerung und argumentieren dafür, dass zumindest im Bayesianischen Rahmenwerk double-counting legitim ist. Des Weiteren behaupten Steele/Werndl (2015), dass die klassische *Model Selection Theory* ein nuancierteres Bild der Verwendung von Daten anbietet als die allgemein diskutierten Positionen. Hier gibt es zwei Möglichkeiten. Erstens gibt es Methoden wie das Kreuzvalidierungsverfahren, bei welchen die Daten neuwertig sein müssen. Für das Kreuzvalidierungsverfahren werden die Daten in zwei Gruppen aufgeteilt: Die erste Gruppe wird für die Schätzung verwendet, die zweite zur Bestätigung. Zweitens gibt es Methoden wie das Akaike-Informationskriterium, für welche die Daten nicht neuartig sein müssen.

Dies bringt uns zu dem zweiten Punkt: *Vorhersage*. Im Klima-Kontext wird typischerweise von Prognosen gesprochen. »Prognose« ist in der Klimamodell-Literatur ein technischer Ausdruck und bezeichnet eine Vorhersage, welche von einem bestimmten Strahlungsantrieb-Szenario abhängt (das Szenario ergibt sich aus den zukünftigen sozioökonomischen und technologischen Entwicklungen).

Das Strahlungsantrieb-Szenario wird entweder durch die Menge an Treibhausgas-Emissionen und Aerosolen bestimmt, welche der Atmosphäre zugeführt werden, oder direkt durch ihre Konzentration in der Atmosphäre.

Modelle können unterschiedlich verwendet werden. Insbesondere können aus Simulationsresultaten entweder physikalische Einsichten in das System gewonnen oder aus diesen Prognosen zum zukünftigen Klima abgeleitet werden (siehe Held (2005) für eine Erörterung dieses Vergleichs). Ein großer Teil der Forschung heutzutage konzentriert sich auf Prognosen über die tatsächliche zukünftige Evolution des Klimasystems, auf Basis derer dann politische Entscheidungen getroffen werden sollen. In diesen Fällen ist es notwendig zu verstehen und zu quantifizieren, wie gut diese Prognosen voraussichtlich sein werden.

Es ist zweifelhaft, ob dies mit herkömmlichen Methoden bewerkstelligt werden kann. Eine solche Methode wäre etwa, sich auf die Bestätigung eines Modells gegenüber historischen Daten zu beziehen (Kapitel 9 des IPCC (2013) befasst sich mit im Detail mit Modellevaluation) und zu argumentieren, dass die Fähigkeit eines Modells, historische Daten erfolgreich zu reproduzieren, uns Zuversicht geben sollte, dass auch die Vorhersagen in der Zukunft richtig sein werden. Das Problem ist, dass Klimaprognosen einen Bereich außerhalb eines bereits erfahrenen Strahlungsantrieb-Szenarios betreffen, und zumindest *prima facie* gibt es keinen Grund anzunehmen, dass Erfolg in niedrigen Strahlungsantrieb-Kontexten zu Erfolg in hohen Strahlungsantrieb-Kontexten führen wird; beispielsweise könnte ein Modell, dessen Parameter durch Daten einer Welt, in der die Arktische See von Eis bedeckt ist, geschätzt wurden, nicht mehr so gut funktionieren, wenn das Eis komplett geschmolzen ist und die relevanten dynamischen Prozesse anders sind. Aus diesem Grund haben Schätzungen mit vergangenen Daten bestenfalls begrenzte Relevanz für den Erfolg der Vorhersagen eines Modells (Oreskes et al. 1994; Stainforth et al. 2007a; 2007b; Steele/Werndl 2013).

Diese Beobachtung unterstreicht, dass es keinen Konsens über die Frage der Vertrauenswürdigkeit von Modell-Ergebnissen gibt. Es herrscht jedoch Einigkeit darüber, dass Vorhersagen für Durchschnitte von großen Zeitintervallen, für größere räumliche Durchschnitte, für niedrige Spezifität und bei Vorhandensein von gutem physikalischem Verständnis besser sind; bei sonst gleichen Bedingungen sind Ereignisse in näherer Zukunft leichter vorherzusagen als Ereignisse in der fernen Zukunft. Weiter werden die Trends der globalen Durchschnittstemperatur als vertrauenswürdig erachtet, und es wird allgemein akzeptiert, dass die Durchschnittstemperatur weiter steigen wird (Oreskes 2007). Der aktuellste IPCC-Report (IPCC 2013: Summary for Policemakers, Section D.1) postuliert etwa, dass modellierte Muster und Trends der Oberflächentemperatur, auf einen globalen und kontinentalen Maßstab bezogen, vertrauenswürdig sind.

Ein schwierigeres Problem betrifft die Verwendung von Modellen um detaillierte Informationen über das *lokale* zukünftige Klima zu erlangen. Beispielsweise hat das UKCP09-Projekt das Ziel, hochauflösende, wahrscheinlichkeitstheoretische Vorhersagen des Klimas bis zum Jahr 2100 zu machen. Diese basieren

auf dem globalen Klimamodell HadCM3, welches am UK Met Office Hadley Centre entwickelt wurde. Die Wahrscheinlichkeiten werden für Ereignisse auf einem 25-km-Gitter angegeben. Spezifische Ereignisse wie z. B. die Temperaturveränderungen des wärmsten Tages im Sommer, Änderungen des Niederschlags am nassesten Tag im Winter oder die Veränderung der durchschnittlichen Wolkenmenge im Sommer werden mit Prognosen angegeben, welche in sich überschneidende Segmente von 30 Jahren eingeteilt sind, die bis zum Jahr 2100 reichen. Es ist beispielsweise prognostiziert, dass die Wahrscheinlichkeit (für ein mittleres Emissionsszenario) für eine 20–30%ige Reduktion des durchschnittlichen Niederschlags im Zentrum von London im Jahr 2080 bei 0.5 liegt.

Die Vertrauenswürdigkeit von solchen Prognosen und deren Relevanz für die Politik wurde angefochten. Ein Modell hat einen strukturellen Modell-Fehler, wenn die Dynamik des Modells sich von der Dynamik des Zielsystems unterscheidet. Frigg et al. (2014a) weisen darauf hin, dass jeder strukturelle Modell-Fehler in chaotischen Modellen möglicherweise dazu führt, dass entscheidungsrelevanten Vorhersagen nicht getraut werden kann. Ferner gibt es wenig Grund anzunehmen, dass die Nachbearbeitung von Modell-Ergebnissen die Konsequenzen solcher Fehler korrigieren kann (Frigg et al. 2014b; 2015). Dies wirft Zweifel bzgl. der Fähigkeit auf, zum jetzigen Zeitpunkt vertrauenswürdige hochauflösende Prognosen bis ans Ende des Jahrhunderts zu machen. Die entscheidende Frage für die Forschung wäre, die Zeitskalen, bei der solche Prognosen wahrscheinlich verlässlich sind, zu bestimmen und außerhalb dieser Zeitskalen den Effekt der Inadäquatheit des Modells abzuschätzen. Wo liegt die Grenze zwischen vertrauenswürdigen und nicht-vertrauenswürdigen Prognosen? Das heißt, wo zwischen globalen Temperatur-Trends und präzisen Prognosen auf einem 25-Kilometer-Gitter endet die Vertrauenswürdigkeit? Das ist eine entscheidende – und für die Politik enorm wichtige – Frage in der Philosophie der Klimawissenschaften, und sie ist bisher nicht gelöst. Die Schlussfolgerungen von Frigg et al. (2014a) werden von Winsberg/Goodwin (2016) bestritten. Sie argumentieren, dass hier die Beschränkungen durch Modell-Fehler überbewertet werden.

6 Verstehen und Quantifizieren von Unsicherheit

Unsicherheit – ein prominentes Thema in vielen Diskussionen über die Klimawissenschaften – ist ein schlecht verstandener Begriff, der viele komplizierte Fragen aufwirft. Im Allgemeinen wird unter Unsicherheit ein Mangel an Wissen verstanden. Die erste Herausforderung ist, genauer zu beschreiben, was mit »Unsicherheit« gemeint ist und was die Quellen der Unsicherheit sind. Es wurden dahingehend bereits einige Vorschläge gemacht, die Diskussion ist aber immer noch in einer »vor-paradigmatischen« Phase. Smith/Stern (2011, 4821–4824) identifizieren vier relevante Arten von Unsicherheit: Ungenauigkeit, Mehrdeutigkeit, Nicht-Analysierbarkeit und Unbestimmtheit. Spiegelhalter/Riesch (2011)

betrachten eine fünfstufige Struktur mit drei Binnen-Modell-Stufen – Ereignis-, Parameter- und Modell-Unsicherheit – und zwei Extra-Modell-Stufen, welche bekannte und unbekannte Unzulänglichkeiten im Modellierungsprozess betreffen. Wilby und Dessai diskutieren das Thema mit Bezug auf sogenannten Kaskaden der Unsicherheit: Sie studieren, wie sich Unsicherheiten vergrößern, wenn man von Annahmen über zukünftige Treibhausgasemissionen Konsequenzen für lokale Klimaanpassungen ableitet. Petersen (2012: Kap. 3 und 6) stellt eine Unsicherheits-Matrix vor, in welcher er die Quellen von Unsicherheit in vertikaler und die Arten von Unsicherheit in horizontaler Richtung auflistet. Lahsen (2005) betrachtet das Thema aus dem Blickwinkel der Wissenschaftsforschung und diskutiert die Verteilung von Unsicherheit als eine Funktion der Distanzen vom Ort der Wissensproduktion. Dies sind nur einige von vielen zurzeit verfügbaren Vorschlägen.

Das nächste Problem ist die Messung und Quantifizierung von Unsicherheit bezüglich Klimavorhersagen. Bei den für diese Herausforderung entwickelten Ansätzen spielen Modell-Ensembles eine zentrale Rolle. Aktuelle Berechnungen der Klimasensitivität und des Anstiegs der globalen Durchschnittstemperatur unter verschiedenen Emissionsszenarien beispielsweise werden von aus Klimamodellen bestehenden Ensembles abgeleitet. Der Grund, Ensembles zu verwenden, ist die anerkannte Unsicherheit in individuellen Modellen, welche sowohl die Anfangsbedingungen, die Parameter des Modells, als auch die Modellstruktur betrifft. Es ist eine allgemein übliche Annahme, dass Ensembles dabei helfen, die Effekte dieser Unsicherheiten zu mindern, entweder durch »robuste« Vorhersagen oder durch das Bereitstellen von numerischen Abschätzungen der Unsicherheit. (Für eine Diskussion von Modell-Ensembles siehe Parker (2013).)

Bevor wir die epistemische Funktion von Ensembles diskutieren, müssen wir zwischen zwei Typen von Ensembles unterscheiden. Wie oben bereits erwähnt, enthalten Klimamodelle eine Vielzahl an Parametern. Manche repräsentieren physikalische Größen wie z. B. die Viskosität des Wassers, während andere »effektive Zusammenfassungen« von Prozessen sind, welche kleiner als das Gitter sind (wie z. B. die Wolkenbedeckung). Ein »perturbed parameter ensemble« (PPE) untersucht durch mehrmaligen Durchlauf des Modells, wie sensibel die Ergebnisse eines Modells von den Parametern abhängen. Dabei werden bei jedem Durchlauf andere Werte für die Parameter verwendet. Dadurch erforscht das Ensemble den Einfluss parametrischer Unsicherheit auf Vorhersagen; d. h. es bietet eine Sensibilitätsanalyse in Hinblick auf die gewählten Parameter. Im Gegensatz dazu besteht ein Multi-Modell-Ensemble aus mehreren verschiedenen Modellen – d. h. Modelle, welche sich nicht nur in den Werten der Parameter, sondern auch in der mathematischen Struktur und dem physikalischen Gehalt unterscheiden. Solche Ensembles werden verwendet, um zu untersuchen, wie Vorhersagen von relevanten Klimavariablen von Unsicherheit bzgl. der Modellstruktur beeinflusst werden.

Ein Resultat ist *robust*, wenn alle (oder die meisten) Modelle im Ensemble das gleiche Ergebnis aufweisen (für eine allgemeine Diskussion der Analyse von Ro-

bustheit siehe Weisberg (2006)). Wenn beispielsweise alle Modelle in einem Ensemble bei einem bestimmten Emissionsszenario mehr als 4 °C Erwärmung (der globalen Durchschnittstemperatur) bis zum Ende des Jahrhunderts vorhersagen, dann ist dieses Resultat bezüglich des Ensembles robust. Rechtfertigt Robustheit größere Zuversicht? Lloyd (2010; 2015) argumentiert, dass Robustheits-Argumente in Verbindung mit Klimamodellen sehr stark sind und zumindest Kernaussagen, wie jener, dass die globale Erwärmung im 20. Jahrhunderts real ist, Glaubwürdigkeit verleihen. Im Kontrast dazu ist die Schlussfolgerung von Parker (2011) nüchtern: »Wenn die heutigen Klimamodelle darin übereinstimmen, dass eine interessante Hypothese über zukünftige Klimaveränderungen wahr ist, kann daraus nicht gefolgert werden [...], dass die Hypothese wahrscheinlich wahr ist, dass die Zuversicht der Wissenschaftler in die Hypothese zunehmen sollte oder dass die Behauptung, dass es Evidenz für die Hypothese gibt, sicherer ist« (2011, 579; Übersetzung der Autoren/innen). Eines der Hauptprobleme ist, dass die heutigen Modelle sich die gleichen Annahmen bezüglich des Klimasystems teilen und die gleichen, in der Rechnerarchitektur begründeten, technologischen Begrenzungen haben. Und aus diesem Grund teilen sie sich zwangsläufig auch einige gemeinsame Fehler. Tatsächlich wurden solche gemeinsamen Fehler umfassend diskutiert (siehe z. B. Knutti et al. 2010) und mehrere Studien haben einen Mangel an Unabhängigkeit von Modellen aufgezeigt (Bishop und Abramowitz 2013; Jun et al. 2008). Wenn Modelle aber nicht unabhängig voneinander sind, dann hat ist unklar, ob Übereinstimmung zwischen ihnen epistemisches Gewicht hat.

Wenn Ensembles nicht zu robusten Vorhersagen führen, dann wird die Bandbreite an Resultaten innerhalb des Ensembles oft zur quantitativen Abschätzung der Unsicherheit des Ergebnisses verwendet. Es gibt diesbezüglich hauptsächlich zwei Ansätze. Der erste Ansatz beruht darauf, das Histogramm der Modell-Resultate direkt in eine Wahrscheinlichkeitsverteilung zu übersetzen. Das Leitprinzip hier ist, dass die Wahrscheinlichkeit eines Ergebnisses proportional zu der Anzahl an Modellen im Ensemble ist, die zu diesem Ereignis führen. Der Gedanke, welcher dieser Methode zu Grunde liegt, scheint zu sein, dass man Modelle als austauschbare Informationsquellen behandeln kann (und deshalb auch, dass es keinen Grund gibt, einem Ensemblemitglied mehr zu trauen als irgendeinem anderen) und dies deshalb zu einer Art von Häufigkeits-Interpretation bezüglich der Wahrscheinlichkeiten führt. Wie aber oben angesprochen wurde, ist die Annahme, dass Modelle unabhängig sind, aus mehreren Gründen problematisch. Darüber hinaus gibt es ein weiteres Problem: Auch die Modelle der besten zurzeit verfügbaren Ensembles wie z. B. CMIP5 sind nicht dazu entworfen worden, um systematisch alle Möglichkeiten zu untersuchen, und es ist daher vorstellbar, dass es große Klassen von Modellen gibt, welche völlig andersartige Resultate produzieren würden. Solange kein Grund gefunden werden kann, dies auszuschließen, ist es nicht klar, warum die Anzahl an Ensemble-Resultaten als Richtwert für Wahrscheinlichkeiten gelten soll. Das IPCC anerkennt diese Ein-

schränkungen (siehe Diskussion im 12. Kapitel von IPCC (2013)) und vermindert die festgesetzte Wahrscheinlichkeit von Konfidenzintervallen, welche durch Ensembles hergeleitet werden (Thomson/Frigg/Helgeson 2016).

Ein gemäßigerer Ansatz betrachtet Ergebnisse von Ensembles als einen Indikator von Möglichkeiten anstatt von Wahrscheinlichkeiten. Gemäß dieser Auffassung zeigt die Bandbreite des Ensembles den Umfang der Ereignisse, welche nicht ausgeschlossen werden können. Dieser Umfang liefert eine untere Grenze der Unsicherheit (Stainforth et al. 2007b). In diesem Sinn argumentiert Katzav (2014), dass die Konzentration auf Vorhersagen fehlgeleitet ist und Modelle dazu verwendet werden sollten zu zeigen, dass gewisse Szenarien echte Möglichkeiten darstellen.

Obwohl zweifelsfrei weniger stark als der Zugang über Wahrscheinlichkeiten, wirft auch dieser Ansatz viele Fragen auf. Die erste ist das Verhältnis zwischen Möglichkeit und Resultaten, die nicht ausgeschlossen werden können. Weiter: Zeigen Resultate, welche nach aktuellem Wissensstand nicht ausgeschlossen werden können, wirklich Möglichkeiten an? Wenn nicht, was ist dann ihre Relevanz für die Berechnung von unteren Grenzen der Unsicherheit? Es ist ferner wichtig, nicht zu vergessen, dass Modell-Ensembles sicherlich nicht die *gesamte* Breite an Möglichkeiten aufzeigen, wodurch bestimmte Typen von formalisierten Entscheidungsfindungsprozessen unmöglich werden. Für eine weiterführende Diskussion dieser Probleme siehe Betz (2009; 2010).

Letztendlich betonen eine Vielzahl von Autoren/innen die Beschränkungen der modellbasierten Methoden (wie etwa Ensemble-Modellen) und argumentieren, dass jede realistische Beurteilung von Unsicherheiten sich auch auf andere Gesichtspunkte verlassen müssen wird, insbesondere auf Urteile von Experten. Petersen (2012: Kap. 4) diskutiert die Methode der Netherlands Environmental Assessment Agency (PBL), welche Expertinnen- und Expertenurteile als eine wichtige Komponente der Einschätzung von Unsicherheit betrachtet. Aspinall (2010) schlägt die Verwendung von strukturierten Experten/innen-Umfragen vor.

Im Lichte der oben aufgeworfenen Probleme stellt sich die Frage, wie Unsicherheit in den Klimawissenschaften den Entscheidungsträgern/innen gegenüber kommuniziert werden soll. Der prominenteste Rahmen für die Kommunikation von Unsicherheit ist jener des IPCC. Dessen neueste Version, welcher im fünften Sachbestandsbericht (AR5) verwendet worden ist, wird in der »Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties« eingeführt und in Mastrandrea et al. (2011) weiter erläutert. Der Rahmen basiert auf zwei Maßeinheiten für die Kommunikation von Unsicherheit. Die erste, ein qualitatives »Konfidenz-Maß«, hängt sowohl von der Art der Evidenz als auch vom Grad der Übereinstimmung unter den Experten/innen ab. Das zweite Maß ist eine quantitative Skala zur Repräsentation von Wahrscheinlichkeiten (oder etwas genauer: Intervallen von Wahrscheinlichkeiten) für relevante Klima- und Ökonomievariablen. Das folgende Zitat veranschaulicht die Verwendung dieser beiden Maße für die Kommunikation von

Unsicherheit in AR5: »Die Veränderung der durchschnittlichen globalen Oberflächentemperatur ist für die Zeitspanne von 2016–2035 relativ zu 1986–2005 für die vier RCPs ähnlich und wird wahrscheinlich im Bereich zwischen 0,3°C und 0,7°C (mittlere Konfidenz) liegen« (IPCC 2013; Übersetzung der Autoren/innen). »RCP« ist die Abkürzung für »representative concentration pathway«, also die zukünftige Konzentration von Treibhausgasen in der Atmosphäre. Eine Erörterung dieses Rahmens kann in Budescu et al. (2014) und Adler/Hadorn (2014) gefunden werden.

An dieser Stelle sollte auch darauf hingewiesen werden, dass die Rolle von ethischen und sozialen Werten in Verbindung mit Unsicherheiten in den Klimawissenschaften kontrovers diskutiert wird. Winsberg (2012) beruft sich auf komplexe Simulationsmodellierung, um zu argumentieren, dass es für Klimawissenschaftler/innen nicht möglich ist, Resultate zu produzieren, welche nicht von ethischen und sozialen Werten beeinflusst sind. Noch konkreter: Er argumentiert, dass die Zuordnung von Wahrscheinlichkeiten zu Hypothesen über zukünftige Klimaveränderungen von ethischen und sozialen Werten abhängt, weil diese Werte bei der Konstruktion und der Bewertung von Klimamodellen ins Spiel kommen. Parker (2014) entgegnet, dass oft pragmatische Faktoren und nicht soziale und ethische Werte während der Bewertung und Konstruktion von Klimamodellen eine Rolle spielen. Sie argumentiert weiter, dass der Fokus nicht auf präzise, wahrscheinlichkeitstheoretische Berechnungen von Unsicherheit gelegt werden soll. Unsicherheit über zukünftige Klimaveränderungen wird durch gröbere Berechnungen angemessener dargestellt und diese gröberen Berechnungen sind weniger von Werten beeinflusst. Sie argumentiert, dass diese Einwände zeigen, dass Winsbergs Argumente zwar nicht falsch sind, der Einfluss von ethischen und sozialen Werten aber überbetont wird. Parker argumentiert weiter, dass traditionelle Probleme bezüglich des Ideals von wertfreier Wissenschaft auch in den Klimawissenschaften zum Tragen kommen. Man könnte nämlich argumentieren, dass Berechnungen von Unsicherheit selbst immer einigermaßen unsicher sind. Dadurch würde die Entscheidung für eine bestimmte Berechnung von Unsicherheit wiederum Werturteile involvieren.

7 Entscheidungen im Licht von Unsicherheit

Klima-Entscheidungen haben Konsequenzen, welche sowohl Individuen als auch Gruppen an verschiedenen Orten zu verschiedenen Zeiten betreffen. Außerdem involvieren die Umstände von vielen dieser Entscheidungen Unsicherheit und Meinungsverschiedenheiten. Diese sind manchmal schwerwiegend und weitreichend; sie betreffen nicht nur den Zustand des Klimas und die weiteren sozialen Konsequenzen von Handlungen, sondern auch die Bandbreite der verfügbaren Handlungen und die Signifikanz, die wir ihren möglichen Konsequenzen beimessen. Diese Überlegungen machen Klima-Entscheidungsfindung sowohl wichtig als

auch schwierig. Es steht viel auf dem Spiel und die Standard-Entscheidungstheorie sieht sich mit großen Schwierigkeiten konfrontiert. Deshalb ist eine philosophische Beschäftigung mit der Entscheidungstheorie besonders relevant.

Beginnen wir mit einer Betrachtung der Akteure/innen, die im Bereich des Klimas eine Rolle spielen, und der Art von Entscheidungsproblemen, mit denen sie konfrontiert sind. Innerhalb der Entscheidungstheorie ist es üblich, zwischen drei Hauptgebieten zu unterscheiden: Individuelle Entscheidungstheorie (welche sich mit Entscheidungsproblemen eines/r einzelnen Agenten/in beschäftigt), Spieltheorie (welche sich auf Fälle von strategischer Interaktion zwischen rationalen Agenten/innen konzentriert) und Sozialwahltheorie (welche Vorgänge betrifft, in denen eine Vielzahl von Agenten/innen kollektiv handeln). Alle drei Bereiche sind für das Dilemma des Klimawandels relevant, ob das Anliegen nun Anpassung an den Klimawandel, Linderung des Klimawandels oder beides ist.

Die angemessenen Perspektiven und Interaktionen zwischen den Agenten/innen zu bestimmen, ist wichtig, weil sonst die Bemühungen einer Entscheidungsmodellierung vergeblich sein können. Es ist z. B. möglicherweise nutzlos, sich auf die Not einzelner zu konzentrieren, wenn die Macht, eine Veränderung herbeizuführen, in Wirklichkeit bei den Staaten liegt. Es ist möglicherweise ebenso irreführend, die Erwartungen an eine gemeinsame Handlung in der Klimapolitik zu analysieren, wenn die vermeintlichen Mitglieder der Gruppe sich selbst nicht als an einer gemeinsamen Entscheidung beteiligt sehen, welche gut für die gesamte Gruppe ist. Es wäre auch irreführend, von dem Entscheidungsmodell eines/r individuellen Agenten/in den Einfluss anderer Agenten, welche in ihrem Umfeld strategisch handeln, auszuschließen. Dies soll aber nicht nahelegen, dass Entscheidungsmodelle die vorliegenden Entscheidungsprobleme immer realistisch repräsentieren müssen; der Punkt ist vielmehr, dass wir Entscheidungsmodelle nicht in einer naiven Art und Weise verwenden sollten.

Die Perspektiven und Interaktionen der Agenten/innen richtig zu bestimmen, ist nur der erste Schritt in der Formulierung eines Entscheidungsproblems. Die Aufgabe, die Details des Entscheidungsproblems von der angemessenen epistemischen und evaluativen Perspektive aus zu repräsentieren, bleibt bestehen. Unser Fokus liegt hier auf der individuellen Entscheidungstheorie, einerseits aus Platzgründen und andererseits, weil die meisten Entscheidungssituationen die Entscheidung eines Individuums betreffen, wobei ein Individuum eine Einzelperson oder eine Gruppe sein kann.

Das Standardmodell der (individuellen) Entscheidungsfindung unter Unsicherheit leitet sich von den klassischen Arbeiten von Neumanns/Morgensterns (1944) und Leonard Savages (1954) ab. Es behandelt Handlungen als Funktionen, die möglichen Zuständen der Welt Konsequenzen zuordnet, welche die Ergebnisse der fraglichen Handlungen in diesem Zustand der Welt sind. Die einzige Art von Unsicherheit ist die Unsicherheit über den Zustand der Welt, und diese wird mittels einer Wahrscheinlichkeitsfunktion über mögliche Zustände der Welt quantitativ beschrieben. Die Wahrscheinlichkeiten reflektieren entweder objek-

tives Risiko oder den Glaubensgrad des/r Entscheidungsträgers/in (oder eine Kombination von beidem). Der relative Wert der Konsequenzen wird von einer intervallskalierten Nutzenfunktion über diese Konsequenzen repräsentiert. Entscheidungsträgern/innen wird empfohlen, die Handlung mit dem größten *erwarteten Nutzen* (EN) zu wählen; der EN für eine Handlung ist die Summe des wahrscheinlichkeitsgewichteten Nutzens der möglichen Konsequenzen der Handlung.

Das Standardmodell ist für viele Klima-Entscheidungen inadäquat, weil es den multidimensionalen Charakter der Unsicherheit, welcher Entscheidungsträger/innen gegenüberstehen, nicht angemessen repräsentiert. Unsicherheit ist nicht nur empirische Unsicherheit bezüglich des Zustands der Welt (Zustands-Unsicherheit). Weitere empirische Unsicherheiten betreffen die vorhandenen Optionen und die Konsequenzen einer jeden Option für einen bestimmten Zustand (Options-Unsicherheit). Im Folgenden verwenden wir »empirische Unsicherheit« als Sammelbegriff für Zustands- und Options-Unsicherheit. Außerdem stehen Entscheidungsträger/innen einer nicht-empirischen Form von Unsicherheit gegenüber – ethischer Unsicherheit: der Frage, welche Werte möglichen Konsequenzen zugeschrieben werden sollen.

Wie schon erwähnt, geht die Entscheidungstheorie davon aus, dass alle empirische Unsicherheit als eine Wahrscheinlichkeitsverteilung über die Zustände der Welt dargestellt werden kann. Dies wirft zwei Probleme auf. Das erste ist, dass es im Fall von komplexen Entscheidungen unnatürlich wirkt, alle Unsicherheit im Zustandsraum zu verorten. Tatsächlich sind klimabezogene Entscheidungsmodelle natürlicher und weniger kompliziert, wenn man erlaubt, dass die Unsicherheit über Zustände von den möglichen Handlungen abhängig ist (vgl. Richard Jeffreys (1965) Erwartung-mal-Wert-Modell), und wenn weiter Unsicherheit bezüglich der Konsequenzen zugelassen wird (dies ist ein Aspekt der Option-Unsicherheit). Es kann zum Beispiel nützlich sein, das Entscheidungsproblem mit einer Zustandsraum-Aufteilung hinsichtlich des möglichen Anstiegs der durchschnittlichen globalen Temperatur über eine gegebene Zeitperiode darzustellen. In diesem Fall würden unsere Glaubensgrade in die Zustände von der gewählten Schadensminderung-Option abhängen.

Das zweite Problem ist, dass die Verwendung einer präzisen Wahrscheinlichkeitsfunktion das Ausmaß der Unsicherheit falsch darstellen kann. Selbst wenn man beispielsweise annimmt, dass die Position der wissenschaftlichen Gemeinschaft bezüglich des Klimawandels unter einer Schadensminderungsoption gut durch eine präzise Wahrscheinlichkeitsfunktion über dem Zustandsraum repräsentiert werden kann, dann ist es trotzdem wenig plausibel, dass präzise Wahrscheinlichkeiten für die mögliche Nahrungsmittelproduktion und andere ökonomische Konsequenzen angegeben werden können. Globale Entscheidungssituationen werden oft mit Hilfe eines sogenannten Integrated Assessment Modells (IAM) analysiert, welches Abhängigkeiten zwischen Schadensminderungsstrategien und Klima- sowie Ökonomievariablen einschließt. Was die empirische Unsicherheit betrifft, kann Nordhaus' (2008) Vertrauen auf »beste Schätzungen«

für Parameter wie Klimasensitivität mit Sterns (2007) »Konfidenzintervallen« verglichen werden. Frisch (2013) argumentiert nichtsdestotrotz dafür, dass *alle bestehenden* IAMs die Unsicherheit bezüglich der Modellparameter nicht adäquat repräsentieren (vgl. Weitzman 2009) und empfiehlt die Verwendung von Mengen von plausiblen Wahrscheinlichkeiten.

Der Gebrauch von Mengen von Wahrscheinlichkeitsfunktionen, um schwerwiegende Unsicherheit zu repräsentieren, ist eine minimale Generalisierung des Standard-Entscheidungsmodells. Unsicherheit wird noch immer mit Wahrscheinlichkeiten beschrieben, aber die Größe der Menge reflektiert, wie schwerwiegend die Unsicherheit ist (siehe z. B. Walley 1991). In manchen Kontexten scheint es angemessen, die möglichen Wahrscheinlichkeitsverteilungen hinsichtlich ihrer Plausibilität zu gewichten (Halpern (2003) diskutiert verschiedene Ansätze). Dieser letztere Zugang passt möglicherweise besser zu den IPCC-Repräsentationen von Unsicherheit bei entscheidungsrelevanten Klima- und Ökonomievariablen. Es ist in der Tat eine wichtige Frage, ob und wie die Repräsentationen von Unsicherheit des IPCC in dem *Rahmen der unpräzisen Wahrscheinlichkeiten* übersetzt werden können. Eine Alternative dazu ist, die Konfidenzmaße und Wahrscheinlichkeiten des IPCC zu einer ungewichteten, unpräzisen Menge von Wahrscheinlichkeitsverteilungen zu kombinieren, welche für eine Eingabe in ein angemessenes Entscheidungsmodell geeignet sind.

Entscheidungsträger/innen begegnen nicht nur empirischer, sondern auch ethischer Unsicherheit. Letztere kann eine Vielzahl unterschiedlicher Ursachen haben. Die wichtigsten treten in Zusammenhang mit Urteilen darüber auf, wie Kosten und Nutzen der Schadensminderung und Anpassung auf verschiedene Regionen und Länder verteilt werden sollen, wie Präferenzen von Personen, deren Lebensumstände von den gewählten Handlungen bezüglich des Klimawandels abhängen, und wie Menschen in der Zukunft berücksichtigt werden sollen. (Für eine Diskussion dieser ethischen Fragen verweisen wir den/die Leser/in auf Bear (2010).) Von diesen ist das letztere das am häufigsten diskutierte Thema, weil die Entscheidungen, wie drastisch Kohlenstoff-Emissionen gekürzt werden sollen, stark vom Diskontsatz abhängen (wie von Broome (2008) hervorgehoben wurde).

Der Diskontsatz wird zur Gewichtung des Wohlergehens zu verschiedenen Zeitpunkten angewendet, wobei ein positiver Satz impliziert, dass zukünftiges Wohlergehen weniger Gewicht hat als gegenwärtiges Wohlergehen. Der gesamte »soziale Diskontsatz« in ökonomischen Modellen involviert neben der unverfälschten Zeitpräferenz auch andere Faktoren, die Güter und Konsum und nicht Wohlergehen *per se* betreffen (siehe Broome (1992) und Parfit (1984) für eine hilfreiche Diskussion). Viele Philosophen/innen betrachten jegliche unverfälschte Diskontierung von zukünftigem Wohlergehen als völlig ungerechtfertigt. Damit soll nicht bestritten werden, dass die zeitliche Dimension möglicherweise mit Eigenschaften der Verteilung von Wohlergehen, die tatsächlich ethisch bedeutsam sind, korreliert. Wenn Leute beispielsweise in der Zukunft finanziell bessergestellt sind, ist es vernünftig, betreffend ihren Interessen weniger besorgt zu sein

als gegenüber jenen der gegenwärtigen Generation. Aber die bloße Tatsache, dass ein Vorteil zu einer bestimmten Zeit eintritt, kann für seinen Wert nicht relevant sein.

Nichtsdestotrotz diskontieren Ökonomen/innen Wohlergehen oft in ihren Modellen. Zudem herrscht erhebliche Uneinigkeit darüber, welcher Diskontsatz verwendet werden soll. Die Stern Review repräsentiert die oben beschriebene unparteiische Perspektive und argumentiert, dass nur ein sehr kleiner Diskontsatz (in der Größenordnung von 0,5%) gerechtfertigt ist, und dies auf Grund der kleinen Wahrscheinlichkeit des Aussterbens der menschlichen Population. Andere Ökonomen/innen hingegen betrachten eine partiische Sichtweise als angemessener. Nach einer solchen Sichtweise, beispielsweise vertreten von Nordhaus (2007) und Arrow (1995a), sollte der Diskontsatz durch die Präferenzen der heute lebenden Menschen bestimmt sein. Aber typische Schätzungen des durchschnittlichen unverfälschten Zeit-Diskontsatzes von beobachtetem Marktverhalten sind viel höher als die von Stern verwendeten Werte (etwa 3%). Obwohl bezweifelt wurde, dass die Verwendung dieser Daten ein adäquates Maß für durchdachte Präferenzen liefert (siehe z. B. Sen 1982; Dreze/Stern 1990; Broome 1992), bleibt das Problem bestehen, dass jede plausible Methode zur Bestimmung der Einstellung der gegenwärtigen Generation gegenüber dem Wohlergehen von zukünftigen Generationen wahrscheinlich eine höhere Rate liefert als die vom Stern Review vorgeschlagene. In dem Maß, wie die Debatte über die ethische Basis für Diskontierung ungelöst bleibt, wird es ethische Unsicherheit über den Diskontsatz in klimapolitischen Entscheidungen geben. Die ethische Unsicherheit kann möglicherweise analog zur empirischen Unsicherheit repräsentiert werden – durch Ersetzung der standardmäßigen, präzisen Nutzenfunktion durch eine Menge von möglichen Nutzenfunktionen.

8 Umgang mit Unsicherheit

Wie sollte ein/e Entscheidungsträger/in zwischen verschiedenen Handlungsalternativen wählen, wenn er/sie Entscheidungen unter großer Unsicherheit treffen muss? Das Problem, welchem Entscheidungsträger/innen in Klimafragen gegenüberstehen, ist, dass die genauen Nutzen- und Wahrscheinlichkeitswerte, welche von der Standard-Entscheidungstheorie gefordert werden, möglicherweise nicht ohne Weiteres zur Verfügung stehen. Es gibt im Wesentlichen drei mögliche Antworten auf dieses Problem.

Erstens kann der/die Entscheidungsträger/in die bittere Pille schlucken und versuchen, genaue Nutzen- und Wahrscheinlichkeitsurteile für die relevanten Eventualitäten anzugeben. Orthodoxe Entscheidungstheoretiker/innen argumentieren, dass die Rationalität gebietet, Entscheidungen so zu treffen, dass sie den erwarteten Nutzen der entscheidenden Person maximieren. Broome (2012, 129; Übersetzung der Autoren/innen) verteidigt diesen Ansatz entschlossen: »Der

Mangel an stabilen Wahrscheinlichkeiten ist kein Grund, das Erwartungswert-Modell aufzugeben ... Bleiben Sie bei diesem Modell, denn es ist gut begründet, und tun Sie Ihr Bestes, um Wahrscheinlichkeiten und Werte zu finden.« Insbesondere wenn es um umweltpolitische Entscheidungen geht, ist es in Bezug auf die Repräsentation von Unsicherheit möglicherweise bereits ein großer Schritt, eine wohlfundierte, präzise Wahrscheinlichkeits- und Nutzenfunktion einzuführen (vgl. Steele 2006). Weitzmann (2009) beispielsweise argumentiert, dass echte Anerkennung von nicht vernachlässigbaren katastrophalen Klimakonsequenzen die Bewertung von Schadensminderung-Optionen radikal verändert. In vielen Fällen bleibt also die Frage offen, wie man Broomes Empfehlung folgen soll: Wie soll der/die Entscheidungsträger/in sich in einer nicht willkürlichen Weise für eine klare Meinung bzgl. entscheidungsrelevanten Themen entscheiden? Es gibt zwei zusammenhängende Strategien: Er/Sie kann weiter reflektieren und/oder widersprüchliche Ansichten verbinden. Die erste zielt auf die Konvergenz von Meinung, während die zweite darauf abzielt, angesichts eines bestehenden Konflikts einen Kompromiss zu finden. (Für eine Diskussion von Beratung siehe Fishkin/Luskin (2005); mehr zur Aggregation findet man z.B. in Genest/Zidek (1986), Mongin (1995), Sen (1970), List/Puppe (2009). Es gibt vergleichsweise wenig formale Literatur über Beratung, ein fruchtbarer Beitrag ist Lehrers/Wagners (1981) Modell für die Aktualisierung von Wahrscheinlichkeitsgraden.)

Zweitens kann der/die Entscheidungsträger/in versuchen, die Entscheidung, oder zumindest Aspekte davon, hinauszuzögern, in der Hoffnung, dass seine/ihre Unsicherheit kontrollierbar wird, wenn mehr Informationen verfügbar sind oder wenn gewisse Widersprüchlichkeiten sich durch eine Einstellungsänderung von selbst auflösen. Das Hauptmotiv für das Hinauszögern einer Entscheidung ist, kostenfrei *Flexibilität* aufrecht zu erhalten (siehe Koopmans 1962; Kreps und Porteus 1978; Arrow 1995b). Nehmen wir an, dass wir uns zwischen dem Bau einer billigen, aber niedrigen Ufermauer und einer hohen, aber teuren entscheiden müssen, wobei der Nutzen dieser beiden Handlungsalternativen von unbekanntem Faktoren wie dem Ausmaß des Anstiegs des Meeresspiegels abhängt. In diesem Fall wäre es vernünftig, die Mauer zuerst niedrig zu bauen, aber die Möglichkeit, sie später zu erhöhen, offen zu lassen. Wenn dies ohne zusätzliche Kosten gemacht werden kann, ist es klar die beste Option. In vielen Anpassungsszenarien ist das, was einer »niedrigen Ufermauer« entspricht, möglicherweise eine soziale Maßnahme, welche eine verzögerte Reaktion auf Klimaveränderungen ermöglicht, was auch immer die Details dieser Veränderung sein mögen. In vielen Fällen ist es jedoch schlicht eine Illusion, einen kostenlosen Aufschub einer Entscheidung zu erwarten, weil eine Verzögerung die Möglichkeiten aufgrund von Veränderungen eher senkt als erhöht. Dies trifft etwa bei Anpassungsentscheidungen aufgrund von Klimaveränderungen zu, ganz zu schweigen von Entscheidungen über eine Schadensminderung.

Zu guter Letzt kann der/die Entscheidungsträger/in eine Entscheidungsregel verwenden, welche nicht von dem Erwartungswert-Modell vorgegeben wird und

viel geringere Ansprüche hinsichtlich der benötigten Informationen mit sich bringt. Es existieren in der Literatur viele verschiedene Vorschläge für solche Regeln. Sie stellen mehr oder weniger starke Abweichungen von der orthodoxen Theorie dar und variieren in den Anforderungen, welche sie an die benötigten Informationen stellen. Es sollte zu Beginn angemerkt werden, dass es bei diesen nicht-standardmäßigen Entscheidungsregeln eine allgemein akzeptierte Auflage an Rationalität gibt: Wenn eine Option gemäß allen zulässigen Paaren von Wahrscheinlichkeits- und Nutzenfunktionen einen geringeren erwarteten Nutzen als eine andere hat, dann ist diese Option keine zulässige Wahl. Dies ist eine relativ minimale Bedingung, aber sie kann in manchen Entscheidungsszenarien zu einer eindeutigen Wahl einer Handlung führen. In solchen Fällen ist die schwerwiegende Unsicherheit nicht wirklich entscheidungsrelevant. Es kann beispielsweise der Fall sein, dass von der Perspektive eines/r globalen Entscheidungsträgers/in aus eine bestimmte Schadensminderungs-Option besser ist, als die Dinge wie gewohnt weiter laufen zu lassen, was auch immer die Details der Unsicherheit des Klimasystems sind. Das ist sogar noch plausibler, je nachdem inwieweit die Schadensminderungsoption als »Win-win«-Strategie gesehen werden kann (Maslin/Austin 2012), d. h. inwieweit sie andere positive Effekte z. B. auf die Luftqualität oder die Energiesicherheit ohne Bezug zur Schadensminderung nach sich zieht. Jedoch ist es in vielen schwierigeren Entscheidungskontexten so, dass diese Auflage an die Rationalität nur einige wenige Optionen als Wahlmöglichkeiten ausschließt.

Ein Faktor, der oft ins Feld geführt wird, um weiter zwischen Optionen zu unterscheiden, ist *Vorsicht*. In der Tat ist dies eine wichtige Facette des populären, aber unklar definierten Vorsorgeprinzips. Vorsichtige Entscheidungsregeln legen mehr Gewicht auf »down-side«-Risiken: die möglichen negativen Folgen der Wahl einer Handlung. Die Maxmin-EN-Regel beispielsweise empfiehlt, die Handlung mit dem größten minimal erwarteten Nutzen zu wählen (siehe Gilboa/Schmeidler 1989; Wallex 1991). Die Regel ist einfach anzuwenden, aber wohl viel zu vorsichtig, da sie die Bandbreite von möglichem erwarteten Nutzen nicht beachtet. Im Kontrast dazu empfiehlt die α -Maxmin-Regel die Handlung mit der größten α -gewichteten Summe des minimalen und maximalen erwarteten Nutzens, welcher mit ihr assoziiert wird. Die relative Gewichtung für den minimal und maximal erwarteten Nutzen kann entweder als Ausdruck von Pessimismus des/der Entscheidungsträgers/in angesichts der Unsicherheit oder als Grad von Vorsicht betrachtet werden (siehe Binmore 2009; für einen umfassenden Überblick über Nicht-Standard-Entscheidungstheorie siehe Gilboa/Marinacci (2012)).

Eine Art von Regeln, für welche mehr Informationen benötigt werden, sind jene, welche Überlegungen bezüglich der *Zuversicht* und/oder *Verlässlichkeit* in Betracht ziehen. Der Gedanke ist, dass ein/e Agent/in mehr oder weniger zuversichtlich ist hinsichtlich der verschiedenen Wahrscheinlichkeits- und Nutzenfunktionen, welche seine/ihre Unsicherheit beschreiben. Zum Beispiel ist dies der Fall, wenn die Berechnungen von verschiedenen Modellen oder Experten/

innen stammen und der/die Entscheidungsträger/in einige Modelle als durch die verfügbare Evidenz besser bestätigt betrachtet als andere oder aber die Urteile eines/r Experten/in als verlässlicher als die eines/r anderen betrachtet. In diesen Fällen ist es *ceteris paribus* vernünftiger, Handlungen zu bevorzugen, von denen man denkt, dass sie vorteilhafte Konsequenzen haben werden. Man könnte dies etwa umsetzen, indem jeder erwartete Nutzen, der mit einer Handlung verbunden ist, in Übereinstimmung damit, wie zuversichtlich man hinsichtlich der Urteile ist, die sie unterstützen, gewichtet wird und dann die Handlung mit dem maximalen, nach Zuversicht gewichteten, erwarteten Nutzen gewählt wird (siehe Klibanoff et al. 2005). Diese Regel ist aber nicht sehr verschieden vom Maximieren des erwarteten Nutzens, und in der Tat könnte man Konfidenzgewichtung als eine Aggregationstechnik statt als alternative Entscheidungsregel betrachten. Erwägungen zur Konfidenz können allerdings nützlich sein, auch wenn präzise Konfidenzgewichtungen nicht zur Verfügung stehen. Gärdenfors/Sahlin (1982/1988) beispielsweise schlagen vor, einfach von den Erwägungen jene auszuschließen, welche unter einen Verlässlichkeits-Grenzwert fallen und dann vorsichtig aus den verbleibenden auszuwählen. Hill (2013) verwendet ein Ordinal-Maß der Konfidenz, das Grenzwerte von Verlässlichkeit zulässt, die vom Vorhaben abhängen und welche dann mit verschiedenen Vorsichtsniveaus kombiniert werden können.

Schließlich gibt es verschiedene Entscheidungsregeln, welche in einem etwas anderen Sinne vorsichtig sind – sie vergleichen Optionen in Bezug auf ihre Robustheit gegenüber Unsicherheit unter der Annahme, dass ein problemspezifisches zufriedenstellendes Niveau des erwarteten Nutzens gegeben ist. Bessere Alternativen sind solche, die einen erwarteten Nutzen aufweisen, der angesichts von Unsicherheit immer noch gut genug oder zumindest nicht zu schlecht ist. Die von Ben-Haim (2001) entwickelte »Information-Gap-Theory« stellt eine Formalisierung dieser Grundidee dar, welche sich in der Umwelt-Management-Theorie etabliert hat. Ein anderer prominenter Ansatz zur robusten Entscheidungsfindung ist jener von Lempert/Popper/Bankes (2003). Für einen Vergleich dieser beiden theoretischen Ansätze siehe Hall et al. (2012). Die betreffende Unsicherheit kann mehrere Facetten in Bezug auf die Wahrscheinlichkeiten der Zustände/Ausgänge oder die Werte der Endausgänge haben. Die meisten Entscheidungsregeln, die auf Robustheit abzielen, setzen voraus, dass eine beste Schätzung der relevanten Variablen verfügbar ist und betrachten dann Abweichungen von dieser Schätzung. Eine robuste Option hat einen befriedigenden erwarteten Nutzen relativ zu einer Klasse von Berechnungen, welche von der besten in gewissem Maße abweichen; je größer die fragliche Klasse, desto robuster die Option. Viel hängt davon ab, welche Höhe an erwartetem Nutzen als befriedigend erachtet wird. Für Schadensminderungs-Entscheidungen ist eine anerkannte ausreichende Stufe von erwartetem Nutzen eine 50 %-Chance, dass die durchschnittliche globale Temperatur sich um 2°C oder weniger erhöht. Zu beachten ist, dass man das Ziel einer solchen Schadensminderung möglicherweise anders interpretiert, nämlich als

Einschränkung dessen, was als durchführbare Option gilt. Mit anderen Worten, Schadensminderungen, die ihr Ziel nicht erreichen, sind einfach unzulässige Optionen, die nicht berücksichtigt werden sollen. Für Anpassungsentscheidungen würde das zufriedenstellende Niveau vom lokalen Kontext abhängen, aber, grob gesprochen, sind die robusten Optionen jene, welche für alle ungünstigen Klimaszenarien mit nicht vernachlässigbarer Wahrscheinlichkeit vernünftige Ergebnisse liefern. Die Optionen, die dieses Kriterium erfüllen, sind meistens jene, die in allen der zuvor erwähnten Klimaszenarios einen zufriedenstellenden Effekt hätten. (Robuste Entscheidungsfindung werden beispielsweise von Dessai et al. (2009) und Wilby/Dessai (2010) befürwortet, welche diese Art von Entscheidungsregeln mit Resilienz-Strategien verbinden. Siehe auch Linkov et al. (2014) für eine Diskussion von Resilienz-Strategien.)

9 Ausblick

In diesem Artikel haben wir Themen und Fragen im Zusammenhang mit den Klimawissenschaften aus wissenschaftstheoretischer Perspektive beleuchtet. Die meisten dieser Themen sind Gegenstand aktueller Forschung und verdienen weitere Aufmerksamkeit. Anstatt diese Punkte hier zu wiederholen, möchten wir ein Thema erwähnen, dem nicht die Aufmerksamkeit zugekommen ist, die es verdient: die epistemische Signifikanz von Konsens bei der Akzeptanz von Ergebnissen. Wie die gegenwärtige Kontroverse über die Arbeit von Cook et al. (2013) zeigt, scheinen doch viele der Meinung zu sein, dass der Grad des Expertenkonsenses ein wichtiger Grund dafür ist, an den Klimawandel zu glauben. Umgekehrt ist es eine klassische Taktik der anderen Seite, den Expertenkonsens anzugreifen und Zweifel zu säen. Die Rolle von Konsens im Kontext des Klimawandels verdient größere Aufmerksamkeit, als ihr bisher zuteilgeworden ist.³

Literatur

- Adler, Carolina E., und Gertrude H. Hadorn (2014). The IPCC and treatment of uncertainties: topics and sources of dissensus. *Wiley Interdisciplinary Reviews: Climate Change* 5.5, 663–676.
- Arrow, Kenneth J. (1995a). Discounting Climate Change: Planning for an Uncertain Future. *Lecture given at Institut d'Économie Industrielle, Université des Sciences Sociales, Toulouse*. Online verfügbar unter <http://idei.fr/doc/conf/annual/paper_1995.pdf> (aufgerufen 22.08.2014).

³ Wir verzichten hier auf ein- bzw. weiterführende Literaturempfehlungen zur Philosophie der Klimawissenschaften, da die Disziplin noch sehr jung ist und es bislang keine zweckmäßige Literatur gibt.

- (1995b). A Note on Freedom and Flexibility. In: Kaushik Basu, Prasanta Pattanaik und Kōtarō Suzumura (Hg.). *Choice, Welfare and Development*. Oxford: Oxford University Press, 7–15.
- Aspinall, Willy (2010). A route to more tractable expert advice. *Nature* 463, 294–295.
- Bear, Paul (2010). Adaption to Climate Change: Who Pays Whom?. In: Stephen Gardiner, Simon Caney, Dale Jamieson und Henry Shue (Hg.). *Climate Ethics: Essential Readings*. Oxford: Oxford University Press, 247–262.
- Ben-Haim, Yakov (2001). *Information-Gap Theory: Decisions Under Severe Uncertainty*. London: Academic Press.
- Betz, Gregor (2009). What range of future scenarios should climate policy be based on? Modal falsificationism and its limitations. *Philosophia naturalis* 46, 133–158.
- (2010). What's the worst case? *Analyse und Kritik* 32, 87–106.
- (2013). Chaos, plurality and model metrics in climate science. In: Ulrich v. Gähde et al. (Hg.). *Models, Simulation, and the Reduction of Complexity*. Berlin: de Gruyter, 255–263.
- Binmore, Ken (2009). *Rational Decisions*. Princeton, NJ: Princeton University Press.
- Bishop, Craig H., und Gab Abramowitz (2013). Climate model dependence and the replicate Earth paradigm. *Climate Dynamics* 41, 885–900.
- Broome, John (1992). *Counting the Cost of Global Warming*. Cambridge: The White Horse Press.
- (2008). The Ethics of Climate Change. *Scientific American* 298, 96–102.
- (2012). *Climate Matters: Ethics in a Warming World*. New York: Norton.
- Budescu, David V., Han-Hui Por, Stephen B. Broomell und Michael Smithson (2014). The interpretation of IPCC probabilistic statements around the world. *Nature Climate Change* 4, 508–512.
- Cohn, Timothy A., und Harry F. Lins (2005). Nature's style: naturally trendy. *Geophysical Research Letters* 32, L23402.
- Cook, John, et al. (2013). Quantifying the consensus on the anthropogenic global warming in the scientific literature. *Environmental Research Letters* 8, 1–7.
- Dessai, Suraje, et al. (2009). Do We Need Better Predictions to Adapt to a Changing Climate?. *Eos* 90.13, 111–112.
- Dessler, Andrew (2011). *Introduction to Modern Climate Change*. Cambridge: Cambridge University Press.
- Drèze, Jean und Nicholas Stern (1990). Policy reform, shadow prices, and market prices. *Journal of Public Economics* 42.1, 1–45.
- Fishkin, James S., und Robert C. Luskin (2005). Experimenting with a Democratic Ideal: Deliberative Polling and Public Opinion. *Acta Politica* 40, 284–298.
- Frank, David, Jan Esper, Eduardo Zorita und Rob Wilson (2010). A noodle, hockey stick, and spaghetti plate: A perspective on high-resolution paleoclimatology. *Wiley Interdisciplinary Reviews: Climate Change* 1.4, 507–516.
- Frigg, Roman, Seamus Bradley, Hailiang Du und Leonard A. Smith (2014a). Laplace's Demon and the Adventures of His Apprentices. *Philosophy of Science* 81.1, 31–59.

- Frigg, Roman, David A. Stainforth und Leonard A. Smith (2014b). The Myopia of Imperfect Climate Models: The Case of UKCP09. *Philosophy of Science* 80.5, 886–897.
- (2015). An Assessment of the Foundational Assumptions in High-Resolution Climate Projections: The Case of UKCP09 2015. *Synthese* 192(12), 2015, 3979–4008.
- Frisch, Mathias (2013). Modeling Climate Policies: A Critical Look at Integrated Assessment Models. *Philosophy and Technology* 26, 117–137.
- (i. E.). Predictivism and old evidence: a critical look at climate model tuning. Forthcoming in: *European Journal for Philosophy of Science*.
- Gärdenfors, Peter, und Nils-Eric Sahlin [1982] (1988). Unreliable probabilities, risk taking, and decision making. In: Peter Gärdenfors und Nils-Eric Sahlin (Hg.). *Decision, Probability and Utility*. Cambridge: Cambridge University Press, 313–334.
- Genest, Christian, und James V. Zidek (1986). Combining Probability Distributions: A Critique and Annotated Bibliography. *Statistical Science* 1.1, 113–135.
- Gilboa, Itzhak, und Massimo Marinacci (2012). Ambiguity and the Bayesian Paradigm. In: Daron Acemoglu, Manuel Arellano und Eddie Dekel (Hg.). *Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress of the Econometric Society*. Cambridge: Cambridge University Press, 179–242.
- Gilboa, Itzhak, und David Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18, 141–153.
- Goodwin, William, und Eric Winsberg (2016). Missing the Forest for the Fish: How Much Does the ›Hawkmoth Effect‹ Threaten the Viability of Climate Projections? *Philosophy of Science* 83 (5), 1122–1132.
- Hall, Jim W., Robert Lempert, Klaus Keller, Andrew Hackbarth, Christophe Mijere und David J. McInerney (2012). Robust Climate Policies Under Uncertainty: A Comparison of Robust Decision-Making and Info-Gap Methods. *Risk Analysis* 32.10, 1657–1672.
- Halpern, Joseph Y. (2003). *Reasoning About Uncertainty*. Cambridge, MA: MIT Press.
- Hegerl, Gabriele C., Ove Hoegh-Guldberg, Gino Casassa, Martin P. Hoerling, Sari Kovats, Camille Parmesan, David W. Pierce und Peter A. Stott (2010). Good Practice Guidance Paper on Detection and Attribution Related to Anthropogenic Climate Change. In: Thomas F. Stocker, Christopher B. Field, Qin Dahe, Vincente Barros, Gian-Kasper Plattner, Melinda Tignor, Pauline M. Midgley und Kristie L. Ebi (Hg.). *Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Detection and Attribution of Anthropogenic Climate Change*. Switzerland: IPCC Working Group I Technical Support Unit, University of Bern.
- Held, Isaac M. (2005). The Gap between Simulation and Understanding in Climate Modeling. *Bulletin of the American Meteorological Society* 80, 1609–1614.
- Hill, Brian (2013). Confidence and Decision. *Games and Economic Behavior* 82, 675–692.
- Hulme, Mike, Suraje Dessai, Irene Lorenzoni und Donald R. Nelson (2009). Unstable Climates: exploring the statistical and social constructions of climate. *Geoforum* 40, 197–206.

- IPCC (2013). *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 and New York: Cambridge University Press.
- Jeffrey, Richard C. (1965). *The Logic of Decision*. Chicago: University of Chicago Press.
- Jun, Mikiyoung, Reto Knutti und Douglas W. Nychka (2008). Local eigenvalue analysis of CMIP3 climate model errors. *Tellus A* 60.5, 992–1000.
- Katzav, Joel (2014). The epistemology of climate models and some of its implications for climate science and the philosophy of science. *Studies in History and Philosophy of Modern Physics* 46, 228–238.
- Klibanoff, Peter, Massimo Marinacci und Sujoy Mukerji (2005). A smooth model of decision making under ambiguity. *Econometrica* 73, 1849–1892.
- Knutti, Reto, Reinhard Furrer, Claudia Tebaldi, Jan Cermak, und Gerald A. Meehl (2010). Challenges in Combining Projections from Multiple Climate Models. *Journal of Climate* 23.10, 2739–2758.
- Koopmans, Tjalling C. (1962). On flexibility of future preference. *Cowles Foundation for Research in Economics, Yale University, Cowles Foundation Discussion Papers* 150.
- Kreps, David M. und Evan L. Porteus (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* 46.1, 185–200.
- Lahsen, Myanna (2005). Seductive Simulations? Uncertainty Distribution Around Climate Models. *Social Studies of Science* 35.6, 895–922.
- Lehrer, Keith, und Carl Wagner (1981). *Rational Consensus in Science and Society*. Dordrecht: Reidel.
- Lempert, Robert J., Steven W. Popper und Steven C. Bankes (2003). *Shaping the Next One Hundred Years: New Methods for Quantitative Long-Term Policy Analysis*. Santa Monica, CA: RAND Corporation, MR-1626-RPC.
- Lenhard, Johannes, und Eric Winsberg (2010). Holism, entrenchment, and the future of climate model pluralism. *Studies in History and Philosophy of Modern Physics* 41, 253–262.
- Linkov, Igor, et al. (2014). Changing the resilience program. *Nature Climate Change* 4, 407–409.
- List, Christian, und Clemens Puppe (2009). Judgment aggregation: a survey. Paul Anand, Prasanta Pattanaik und Clemens Puppe (Hg.). *Oxford Handbook of Rational and Social Choice*. Oxford: Oxford University Press, Kapitel 19.
- Lorenz, Edward N. (1995). Climate is what you expect. Prepared for publication by NCAR. Unpublished, 1–33.
- Lloyd, Elisabeth A. (2009). Varieties of Support and Confirmation of Climate Models. *Proceedings of the Aristotelian Society Supplementary Volume LXXXIII*, 217–236.
- (2010). Confirmation and robustness of climate models. *Philosophy of Science* 77, 971–984.
 - (2015). Model robustness as a confirmatory virtue: The case of climate science. *Studies in History and Philosophy of Science* 49, 58–68.

- Mann, Michael E., Raymond S. Bradley und Malcom K. Hughes (1998). Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* 392, 779–787.
- Maslin, Mark, und Patrick Austin (2012). Climate models at their limit?. *Nature* 486, 183–184.
- Mastrandrea, Michael D., Katherine J. Mach, Gian-Kasper Plattner, Ottmar Edenhofer, Thomas F. Stocker, Christopher B. Field, Kristie L. Ebi, und Patrick R. Matschoss (2011). The IPCC AR5 guidance note on consistent treatment of uncertainties: a common approach across the working groups. *Climatic Change* 108, 675–691.
- McGuffie, Kendal und Ann Henderson-Sellers (2005). *A Climate Modelling Primer*. New Jersey: Wiley.
- Mongin, Philippe (1995). Consistent Bayesian Aggregation. *Journal of Economic Theory* 66.2, 313–351.
- von Neumann, John, und Oskar Morgenstern (1944). *Theory of Games and Economic Behaviour*. Princeton: Princeton University Press.
- Nordhaus, William D. (2007). A Review of the Stern Review on the Economics of Climate Change. *Journal of Economic Literature* 45.3, 686–702.
- (2008). *A Question of Balance*. New Haven, CT: Yale University Press.
- Oreskes, Naomi (2007). The Scientific Consensus on Climate Change: How Do We Know We're Not Wrong?. In: Joseph F. C. DiMento und Pamela Doughman (Hg.). *Climate Change: What It Means for Us, Our Children, and Our Grandchildren*. Boston: MIT Press, 65–99.
- , Kristin Shrader-Frechette und Kenneth Belitz (1994). Verification, validation, and confirmation of numerical models in the Earth Science. *Science New Series* 263.5147, 641–646.
- Parfit, Derek (1984). *Reasons and Persons*. Oxford: Clarendon Press.
- Parker, Wendy S. (2009). ›Confirmation and Adequacy for Purpose in Climate Modelling.« *Aristotelian Society Supplementary Volume* 83.1, 233–249.
- (2010). Comparative Process Tracing and Climate Change Fingerprints. *Philosophy of Science* 77, 1083–1095.
- (2011). When Climate Models Agree: The Significance of Robust Model Predictions. *Philosophy of Science* 78.4, 579–600.
- (2013). Ensemble modeling, uncertainty and robust predictions. *Wiley Interdisciplinary Reviews: Climate Change* 4.3, 213–223.
- (2014) ›Values and Uncertainties in Climate Prediction, Revisited.« *Studies in History and Philosophy of Science Part A* 46 (2014), 24–30.
- Petersen, Arthur C. (2012). *Simulating Nature: A Philosophical Study of Computer-Simulation Uncertainties and Their Role in Climate Science and Policy Advice*. Boca Raton, Florida: CRC Press.
- Savage, Leonard J. (1954). *The Foundations of Statistics*. New York: John Wiley & Sons.
- Sen, Amartya K. (1970). *Collective Choice and Social Welfare*. San Francisco: Holden-Day Inc.

- (1982). Approaches to the choice of discount rate for social benefit–cost analysis. *Discounting for Time and Risk in Energy Policy* (ed. R. C. Lind), 325–353. Washington, DC: Resources for the Future.
- Smith, Leonard A., und Nicholas Stern (2011). Uncertainty in science and its role in climate policy. *Phil. Trans. R. Soc. A* 369, 4818–4841.
- Spiegelhalter, David J., und Hauke Riesch (2011). Don't know, can't know: embracing deeper uncertainties when analysing risks. *Phil. Trans. R. Soc. A* 369, 4730–4750.
- Stainforth, David A., Myles Allen, Edward R. Tredger und Leonard A. Smith (2007a). Confidence, Uncertainty and Decision-support Relevance in Climate Predictions. *Philosophical Transactions of the Royal Society A* 365, 2145–2161.
- Stainforth, David A., Thomas E. Downing, Richard Washington, Ana Lopez und Mark New (2007b). Issues in the Interpretation of Climate Model Ensembles to Inform Decisions. *Philosophical Transactions of the Royal Society A* 365, 2163–2177.
- Steele, Katie (2006). The precautionary principle: a new approach to public decision-making?. *Law Probability and Risk* 5, 19–31.
- und Charlotte Werndl (2013). Climate Models, Confirmation and Calibration. *The British Journal for the Philosophy of Science* 64, 609–635.
- und Charlotte Werndl (2015). Model Selection Theory: The Need for a More Nuanced Picture on Use-Novelty and Double-Counting. *The British Journal for the Philosophy of Science*, doi: 10.1093/bjps/axw024.
- Stern, Nicholas (2007). *The Economics of Climate Change: The Stern Review*. Cambridge: Cambridge University Press.
- Erica Thompson, Roman Frigg, and Casey Helgeson (2016): Expert Judgment for Climate Change Adaptation. *Philosophy of Science* 83(5), 1110–1121.
- Walley, Peter (1991). *Statistical Reasoning with Imprecise Probabilities*. New York: Chapman and Hall.
- Weitzman, Martin L. (2009). On Modeling and Interpreting the Economics of Catastrophic Climate Change. *The Review of Economics and Statistics* 91.1, 1–19.
- Werndl, Charlotte (2015). On defining climate and climate change. *The British Journal for the Philosophy of Science* 67 (2), 337–364. doi, 10.1093/bjps/axu48.
- Wilby, Robert L., und Suraje Dessai (2010). Robust adaptation to climate change. *Weather* 65.7, 180–185.
- Winsberg, Eric (2012). Values and Uncertainties in the Predictions of Global Climate Models. *Kennedy Institute of Ethics Journal* 22, 111–127.

