

3 PROFESSIONALISIERUNG DER MODELLIERUNG

Rückblickend auf die letzten zwanzig Jahre lässt sich in der Klimamodellierung eine interessante Beobachtung machen. Aufgrund des enormen sozio-politischen Drucks, möglichst treffende Prognosen zukünftiger Klimaentwicklungen zu geben, und dem Umstand, dass diese Prognosen nur auf Basis der Computermodelle zu erstellen sind, findet seit gut zwei Jahrzehnten eine zunehmende Professionalisierung der Modellierung statt. Diese Professionalisierung zeigt sich an der internationalen Koordinierung und Synchronisierung der Modellierung, an der Ausdifferenzierung von Sub-Communities und an neuen Evaluationsstrategien. Vor allem die Synchronisierung ist einzigartig in der Wissenschaft und wird seit 1990 maßgeblich durch die Assessment Reports des IPCC Intergovernmental Panel on Climate Change vorangetrieben. Die Synchronisierung der IPCC Referenzmodelle umfasst dabei die Verbesserung, Evaluierung und Standardisierung der Modelle in einem international konzertierten Rhythmus. Vor allem die Möglichkeit, Modelle miteinander vergleichen zu können, basiert auf dieser Synchronisierungsleistung. Daher entwickelt sich die Klimamodellierung zur paradigmatischen Simulationswissenschaft, die neue Standards der Modellierung, des Experimentierens mit diesen Modellen und der Evaluation generiert, welche auch für andere Wissenschaftsbereiche handlungsleitend werden könnten. Ähnlich der Astronomie im 17. Jahrhundert, die im Kontext der neuzeitlichen Form der Erkenntnisproduktion auf Basis wissenschaftlicher Instrumente und mathematischer Methoden eine Leitfunktion innehatte, übernimmt die Klimaforschung eine ähnliche Leitfunktion für die Computational Sciences. Dabei durchläuft sie Restrukturierungsprozesse, die sie zu einer paradigmatischen E-Science machen.

Standardisierung

Voraussetzung für die Standardisierung der Modellierung ist die Etablierung einer internationalen Infrastruktur für den Austausch zwischen den einzelnen Forschergruppen und Institutionen. Die Meteorologie kann hier auf eine lange Tradition gemeinsamer Forschung und Messung zurückblicken. Bereits 1873 wurde die IMO International Meteorological Organisation gegründet und 1882/1883 bildete das 1. Internationale Polarjahr den Auftakt zu international koordinierten Messungen und Forschungen.¹ Heute vertritt die Nachfolgeorganisation WMO World Meteorological Organisation der UN United Nations 188 Nationen und führt mit dem WCP World Climate Programme seit 1980 zahlreiche Aktivitäten der internationalen Koordinierung der Klimaforschung durch (vgl. WMO 2009, WCP 2009). Dazu gehören seit 1979 das WCIRP World Climate Impact Assessment and Response Strategies Programme, seit 1980 das WCRP World Climate Research Programme, seit 1989 das IPCC Intergovernmental Panel on Climate Change und seit 1992 das GCOS Global Climate Observing System (vgl. WCRP 2009, IPCC 2009, GCOS 2009). Vor allem das WCRP Programm ist für die Standardisierung der Modellierung interessant, denn seine Mission ist es „to develop and evaluate climate system models for understanding, assessing and predicting Earth climate change and variations“ (WCRP 2009: Activities). Im Rahmen dieses Programms wurden 1980 die Working Group on Numerical Experimentation und 1997 die Working Group on Coupled Modelling gegründet. Dies sind nur einige Beispiele der international hochgradig vernetzten Meteorologie. Ein weiteres Beispiel auf EU Ebene ist das PRISM Programme for Integrated Earth System Modelling des ENES European Network for Earth System Modelling (vgl. PRISM 2009, ENES 2009), denn seit Anfang 2000 ist die Integration der bislang heterogenen Modelle und Modellkomponenten zu Erdsystemen das maßgebliche Ziel. So ist es Aufgabe von ENES „to develop an advanced software and hardware environment in Europe, under which the most advanced high resolution climate models can be developed, improved, and integrated“ (ENES 2009: Welcome). Voraussetzung für diese Integration ist die Standardisierung der Modellkopplungen und Softwareeinbettungen. Ein an diesen Prozessen beteiligter Klimaforscher formulierte die Situation 2005 folgendermaßen:

1 Gefolgt vom 2. Internationalen Polarjahr 1932/33, vom Internationalen Geophysikalischen Jahr 1957/58 und vom 3. Internationalen Polarjahr 2007/2008

„Die Idee ist, eine höhere Modularität zu erzielen, wenn man Standardkoppler für die Kommunikation der Modellteile benutzt. Dann kann man einen Modellteil durch einen anderen austauschen, der auch diesen Standard befolgt. Es gibt jetzt eine Reihe von Institute in Europa, die diese Koppler benutzen und damit ihre eigenen Komponenten mit anderen Komponenten koppeln. Also es gibt Institute, die benutzen unser Atmosphärenmodell, aber ein anderes Ozeanmodell oder umgekehrt. [...] Man kann mehrere Realisationen aus einem Pool von Modellen zusammenstellen und, vorausgesetzt es funktioniert auch wissenschaftlich gut, hat man die Möglichkeit, verschiedene Simulationen durchzuführen. Man bekommt eine größere Vielfalt und kann dann sehen: Ja, es kommt annäherungsweise die gleiche Antwort heraus, auch wenn ich technisch ein anderes System benutze. Oder es kommt etwas ganz anderes heraus“ (Interview 6, 2005).

Mittlerweile gehören gekoppelte Modelle zum Standard der Klimaforschung. In welche Richtung die aktuelle Entwicklung führt, dokumentiert das britische Projekt GENIE Grid Enabled Integrated Earth System Model:

„Our objectives are to build a model of the complete Earth system which is capable of numerous long-term (multi-millennial) simulations, using components which are traceable to state-of-the-art models, are scalable (so that high resolution versions can be compared with the best available, and there is no barrier to progressive increases in spatial resolution as computer power permits), and modular (so that existing models can be replaced by alternatives in future). Data archiving, sharing and visualisation will be integral to the system. The model will be used to quantitatively test hypotheses for the causes of past climate change and to explore the future long-term response of the Earth system to human activities“ (GENIE 2009: Scientific Research Challenge).

„The system will be accessed via a portal, which will allow a user to compose, execute and analyse the results from an Earth system simulation. After authenticating themselves with the portal, a user will have access to a library of components of the Earth system (for example, ocean, atmosphere) at different resolutions. The user constructs a composite application by selecting from these components; a selection informed by meta-data provided by a component's author and made available in the library. Together with these components the user selects appropriate mesh conversion tools, to enable data exchange at model boundaries, and defines an event-queue, which specifies the timing of the data exchanges and indicates what simulated data is to be archived for later inspection. The user also provides the data necessary to initialise the model. From this, an intelligent meta-scheduler determines the resource requirements and maps the processing required to a distributed Grid of compute resources using middleware such as Globus and Condor. At runtime each component produces distributed data, which can be monitored during execu-

tion and is also archived automatically as specified by the user. From the portal it is possible to browse this archive of results using post-processing visualization tools and reuse results from the archive to seed new calculations“ (GENIE 2009: Vision).

Die Komponenten für das Erdsystem liegen bereits vor. Woran gearbeitet wird, sind kompatible und effiziente Grid-fähige Koppler. Die Idee des Grid Computing ist es, auf einer internetbasierten Oberfläche Zugang zu Supercomputern, Simulationen und Messinstrumenten zu geben. In diesem Zusammenhang ist vom ‚virtual lab‘ und vom ‚web-portal based experiment management‘ die Rede. Ohne dass der Forscher sein Büro verlassen muss, kann er umfangreiche Experimentalsysteme je nach Fragestellung zusammenstellen und damit Experimente durchführen, da er auf Hard- und Software-Ressourcen weltweit Zugriff hat. Voraussetzung dafür sind nicht nur die als kompatible Module standardisierten Modelle und Koppler, sondern auch entsprechend vereinheitlichte Datenformate wie netCDF oder HDF5. netCDF Network Common Data Form ist das gängige Format wissenschaftlicher Daten und ein Standard-Interface für den Datenaustausch. HDF5 Hierarchical Data Format ist ein weitverbreitetes Satelliten-Datenformat. Ein weiteres, von der WMO standardisiertes Format ist GRIB Gridded Binary, das für meteorologische Reanalyse- und Vorhersagedaten verwendet wird. „In these [GRIB] codes the data are described in the message itself: it is the self-description feature. There will be sections at the beginning of the report, which define what data are transmitted in this message. These sections will in fact contain pointers towards elements in predefined and internationally agreed tables (stored in the official WMO Manual on Codes). Once these sections are read, the following part of the message containing the data (the Data Section) can be understood“ (WMO 2003: 2). Darüberhinaus gibt es Anstrengungen, die Metadaten wie Variablenamen, Koordinatensysteme oder Einheiten zu standardisieren.²

-
- 2 “The netCDF library is designed to read and write data that has been structured according to well-defined rules and is easily ported across various computer platforms. The netCDF interface enables but does not require the creation of self-describing datasets. The purpose of the CF conventions is to require conforming datasets to contain sufficient metadata that they are self-describing in the sense that each variable in the file has an associated description of what it represents, including physical units if appropriate, and that each value can be located in space (relative to earth-based coordinates) and time. An important benefit of a convention is that it enables software tools to display data and perform operations on specified subsets of the data with minimal user intervention. [...] This standard is intended for use with climate and forecast data, for atmosphere, surface and ocean,

Doch Computereperimente setzen eine weitere Standardisierung voraus, nämlich die der Verteilung der Messdaten. Bis heute liefert das verzweigte Netz aus Bodenstationen (ca. 260.000 Messdaten), Radiosondierungen bzw. Wetterballons (ca. 600 Messdaten) und Satelliten (ca. 250.000 Messdaten) keine ausreichende globale Abdeckung. Die Stationen sind unregelmäßig verteilt und die Messdaten fehlerhaft. Da Simulationen in einer gleichmäßigen Gitterstruktur gerechnet werden – typischerweise in einem rechtwinkligen Euklidischen Gitter und seit Neuestem in einem Ikosaeder Gitter – müssen die Messdaten assimiliert, d.h. durch Interpolation in gleichmäßig verteilte Datenstrukturen überführt werden. Dies ist keine triviale Aufgabe, wie die Beschreibung des DWD Deutschen Wetterdienstes illustriert:

„Die Grundaufgabe der vierdimensionalen (3 Raum plus 1 Zeitdimension) Datenassimilation lautet somit: Aus unvollständigen und fehlerhaften Beobachtungen zusammen mit einer näherungsweisen Beschreibung der Atmosphäre durch die prognostischen Gleichungen des Vorhersagemodells soll der wahrscheinliche augenblickliche Zustand der Atmosphäre analysiert sowie der Fehler dieser Analyse bestimmt werden. Die ‚operationell‘, also routinemäßig täglich angewandten Verfahren der vierdimensionalen Datenassimilation [...] beruhen auf zahlreichen Vereinfachungen, wie etwa Linearisierung einiger wichtiger Operationen und ad-hoc Spezifikation der erforderlichen Beschreibung (mathematisch gefasst in ‚Kovarianzmatrizen‘) der statistischen Fehler der verschiedenen Eingangsinformationen“ (DWD 2009: Datenassimilation).

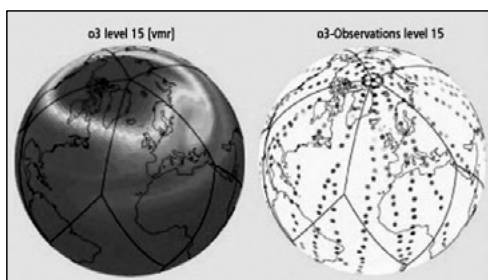


Abbildung 19: Datenassimilation (l.) und ENVISAT Messdatenverteilung (r.) (DEKLIM 2009)³

and was designed with model-generated data particularly in mind” (CF Metadata Convention 2009: 1).

- 3 „Ozonwerte aus rund 33 km Höhe, von einem Messinstrument des ENVISAT-Satelliten gewonnen, wurden mit Hilfe des neuen Verfahrens bearbeitet [SACADA Synoptic Analysis of Chemical Constituents by Advanced Satellite Data Assimilation]. Die rechte Graphik zeigt [das] Ergeb-

Primärdaten gibt es heute kaum noch. Messdatensätze werden mit Hilfe von Algorithmen automatisch in standardisierter Form aufbereitet und dafür nutzt man verschiedene mathematische Verfahren, aber auch Modelldaten. Noch deutlicher zeigt sich dies an den Satellitenbeobachtungen. Auch Satelliten liefern keine regelmäßig verteilten Messdaten und müssen entsprechend assimiliert werden. Wie umfangreich Datenassimilationen die Lücken ausgleichen, zeigt die Assimilation der ENVISAT Daten zur Ozonmessung, wie in Abbildung 19 dargestellt. Darüber hinaus sind Satellitendaten indirekte Daten, die erst durch Datenretrievals gewonnen werden. „Die Datensignale der Satelliten müssen umgedeutet werden,“ wie ein Klimamodellierer, spezialisiert auf Satellitendaten, erklärte. „Umdeuten heißt, interpretieren auf Basis von Modellen. Satelliten messen beispielsweise die Helligkeit im Blaubereich. Um daraus Informationen abzuleiten, müssen Annahmen über die Bewölkung getroffen werden, über die Dichte, Teilchengröße etc. Diese Unterscheidungen kann die Momentaufnahme nicht liefern. Man muss adäquate Annahmen [basierend auf Modellerkenntnissen] treffen. Praktisch sieht es so aus, dass es bessere oder schlechtere Datenretrievals gibt“ (Interview 9, 2005). Dennoch werden Satellitendaten als „Messdaten eingestuft, aber das sind sie eben nicht nur“ (Interview 9, 2005). Solche Datenretrievals werden automatisch auf Basis von Auswertungsalgorithmen durchgeführt.

Ein weiterer, wichtiger Bereich der Standardisierung sind Referenzdatensätze. Referenzdatensätze werden weltweit benutzt, um Modelle zu testen und zu evaluieren. Die Standardisierung liegt hier in der Auswahl bestimmter Daten und deren Aufbereitung für die gesamte Scientific Community. Zu diesen Referenzdatensätzen zählen auch Reanalysen historischer Daten. Historische Daten sind für die Evaluierung von Klimamodellen wichtig. Reproduzieren Modelle, initialisiert auf Basis historischer Messdaten, das aktuelle Klima, dann wird davon ausgegangen, dass sie auch zukünftige Entwicklungen prognostizieren können. Historische Messdaten liegen jedoch in einer sehr heterogenen Verteilung vor. Natürliche Datenarchive wie Eisbohrkerne, Baumstämme und Sedimente liefern nur lokale Informationen. Atmosphärendaten wurden bis in die 1940er Jahre in der Regel vom Boden her und Ozeandaten überhaupt nicht erhoben. Seit einiger Zeit werden daher historische Messdaten mit Assimilations-Techniken rekonstruiert bzw. neu berechnet, um räumlich

nis einer zehntägigen, fortschreitenden Assimilation“ (DEKLIM 2006: SACADA). Ozonmessungen mit Hilfe von Fernerkundung sind relativ genau, indem sie das reflektierte UV-Licht messen und damit auf die Ozonwerte rückschließen.

wie zeitlich gleichmäßig verteilte Datensätze zu erhalten. Ein Beispiel sind die NCEP/NCAR Reanalysedaten des US-amerikanischen NCEP National Centers for Environmental Prediction und des NCAR National Centre for Atmospheric Science. Mit diesen Daten werden verschiedene, rekonstruierte Messdatensätze der meteorologischen Variablen für den Zeitraum von 1948 bis 2002 im netCDF Format zur Verfügung gestellt (vgl. NCEP/NCAR 2009: Reanalysis). Mit diesem Referenzdatensatz werden Modelle weltweit evaluiert und für Prognosen initialisiert.⁴ Ein weiterer Reanalysedatensatz ist die ECMWF 40 Year Reanalysis (ERA-40) des ECMWF European Centre for Medium-Range Weather Forecasts. „The re-analysis project ERA-40 will cover the period from mid-1957 to mid-2002, overlapping the earlier ECMWF re-analysis, ERA-15, 1979 to 1993. [...] The data sets are based on quantities analysed or computed within the ERA-40 data assimilation scheme or from forecasts based on these analyses“ (ERA-40 2009). Referenzdatensätze liegen in der Regel im netCDF Format vor und können von internetbasierten Plattformen heruntergeladen werden. Eine Alternative dazu ist die Zusammenstellung eines individuellen Datensatzes aus einer Datenbank, die via Internet zugänglich ist. So bietet etwa das WDC World Data Center for Paleoclimatology mit der Ice Core Data Search einen solchen Service für Paleodaten an (vgl. NCDC/NOAA 2009: Ice Core Data Search). Insgesamt gibt es 52 dieser World Data Center.⁵

-
- 4 „The NCEP/NCAR Reanalysis 1 project is using a state-of-the-art analysis/forecast system to perform data assimilation using past data from 1948 to the present. A large subset of this data is available from PSD in its original 4 times daily format and as daily averages. However, the data from 1948-1957 is a little different, in the regular (non-Gaussian) gridded data. That data was done at 8 times daily in the model, because the inputs available in that era were available at 3Z, 9Z, 15Z, and 21Z, whereas the 4x daily data has been available at 0Z, 6Z, 12Z, and 18Z. These latter times were forecasted and the combined result for this early era is 8x daily. The local ingestion process took only the 0Z, 6Z, 12Z, and 18Z forecasted values, and thus only those were used to make the daily time series and monthly means here“ (NCEP/NCAR 2009: Reanalysis 1, Summary).
 - 5 “The World Data Center (WDC) system was created to archive and distribute data collected from the observational programs of the 1957-1958 International Geophysical Year. Originally established in the United States, Europe, Russia, and Japan, the WDC system has since expanded to other countries and to new scientific disciplines. The WDC system now includes 52 Centers in 12 countries. Its holdings include a wide range of solar, geophysical, environmental, and human dimensions data. These data cover timescales ranging from seconds to millennia and they provide baseline information for research in many ICSU [International Council for Science] disciplines, especially for monitoring changes in the geosphere and

Synchronisierung

Neben den zahlreichen Aktivitäten der Standardisierung von Daten, Modellen und Experimentalvoraussetzungen spielt die Synchronisierung der Klimamodellierung eine herausragende Rolle. Die seit 1990 publizierten IPCC Berichte – FAR First Assessment Report 1990, SAR Second Assessment Report 1995, TAR Third Assessment Report 2001, AR4 Fourth Assessment Report 2007 – führten zu einer Umstrukturierung der Klimaforschung, die einzigartig in der Wissenschaftslandschaft ist. Die sich beteiligenden Institutionen waren aufgrund des Publikationsrhythmus gezwungen, ihre Entwicklungsarbeit an den Modellen zeitlich anzupassen. Dies bedeutet, dass bereits während der Publikation eines IPCC Reports über Modellverbesserungen für den nächsten Report beraten wird. Institute, die IPCC Szenarien berechnen, durchlaufen dann einen international konzertierten Zyklus der Modellverbesserung, der Modellevaluation und schließlich der Szenarienberechnung und -auswertung von fünf bis sechs Jahren. Mit dieser Synchronisation werden die Vergleichbarkeit der Modelle, Simulationsbedingungen und Ergebnisse sowie die Güte der Forschung sichergestellt. Dieser Turnus ist bereits weltweit viermal durchlaufen worden und hat nun für den fünften Bericht, der 2014 publiziert werden soll, begonnen. Auf einer Expertentagung im Juli 2009 wurden die Ziele für den nächsten Bericht besprochen. Für die Klimamodellierung, die in WGI Working Group I organisiert ist, bedeutet dies: „The WGI assessment will address the full range of time scales (from decades to millennia) and space scales (from global to regional). This requires the consideration of the couplings between physical and biogeochemical processes and their influence on the sensitivity of the climate system, as well as the assessment of irreversible changes on the range of spatial scales in various components of the climate system. [...] Efforts targeted in these directions are already underway or planned within the climate modelling community (e.g. CMIP5 of WCRP)“ (IPCC 2009; 9, 10).⁶

biosphere - gradual or sudden, foreseen or unexpected, natural or man-made” (WDC 2009: About WDC).

- 6 „The IPCC Working Group I (WG I) assesses the physical scientific aspects of the climate system and climate change. The main topics assessed by WG I include: changes in greenhouse gases and aerosols in the atmosphere; observed changes in air, land and ocean temperatures, rainfall, glaciers and ice sheets, oceans and sea level; historical and paleoclimatic perspective on climate change; biogeochemistry, carbon cycle, gases and aerosols; satellite data and other data; climate models; climate projections, causes and attribution of climate change” (IPCC 2009: Working Groups/ Task Force).

Die Anforderungen an die Modellverbesserungen für AR5 sind aufgrund neuer Möglichkeiten wie Grid-Computing umfangreich. Bereits für AR4 waren die Modelle entscheidend verändert worden, denn es wurde erstmals mit gekoppelten Atmosphären-Ozeanmodellen gerechnet. Viele der Modelle waren zudem für Parallelrechner neu codiert und gleichzeitig umfangreich verbessert worden. Im vierten Bericht hieß es dazu: „Model improvements can, however, be grouped into three categories. First, the dynamical cores (advection, etc.) have been improved, and the horizontal and vertical resolutions of many models have been increased. Second, more processes have been incorporated into the models, in particular in the modelling of aerosols, and of land surface and sea ice processes. Third, the parametrizations of physical processes have been improved. [...] These various improvements, developed across the broader modelling community, are well represented in the climate models used in this report“ (IPCC AR4 WG1 2007: 596). Auf Basis der verbesserten Modelle, zu welchen auch das Atmosphärenmodell ECHAM5 gekoppelt mit dem MPI-OM Ozeanmodell gehörte, wurden ab 2005 verschiedene Szenarien berechnet. Beispielsweise simulierte das Max-Planck-Institut für Meteorologie mit seinem gekoppelten ECHAM5/MPI-OM Modell insgesamt achtzehn Szenarien. Dafür wurden allein im Jahr 2005 400.000 Rechnerstunden benötigt (vgl. Böttinger 2005).

Generell handelt es sich bei den IPCC Berichten um eine außergewöhnliche Form der Publikation. An den vierzehn Kapiteln des dritten Berichts, der Ende 2001 veröffentlicht wurde, arbeiteten ab Mitte 1998 122 verantwortlichen Autoren und 515 Co-Autoren. „One and a half year later, in January 2001, when the final versions of the chapters were accepted at the IPCC WGI plenary session in Shanghai, four revisions had been made of drafts of the chapters. The review rounds involved 420 experts and 100 governments. At the plenary session in Shanghai, the Summary for Policymaker (SPM) of the report was approved line by line by the governments in four days“ (Petersen 2006: 148). Jeder Kommentar im Rahmen des öffentlichen Schreibprozesses musste von den führenden Autoren berücksichtigt und geprüft werden. Einer der führenden Autoren des vierten Berichts erzählte im Gespräch, dass er über ein Jahr lang ausschließlich mit dem Schreiben und Korrigieren seines Kapitels beschäftigt war und auf hunderte Kommentare eingehen musste. Der Hintergrund für diesen Aufwand liegt in dem Wunsch, den Entstehungsprozess so offen und transparent wie möglich zu gestalten. Nahezu die gesamte Scientific Community der Klimaforschung ist in der einen oder anderen Weise an den Berichten beteiligt. Daher repräsentieren die IPCC Berichte den aktuellen Stand der Klimaforschung und die-

nen auch den Forschern selbst als Informationsquelle zu benachbarten Wissenschaftsgebieten.

Doch der Aufwand der IPCC Berichte hat einen weiteren Grund. Der soziopolitische Einfluss ist unübersehbar und die Berichte sind nicht unumstritten (vgl. beispielsweise Shackley et al. 1999; Skodvin 2000; Miller, Edwards 2001; Voss 2008). Ausgangspunkte der Kritik sind die enge Verbindung von Wissenschaft und Politik und die Unsicherheit der Prognosen. Während die beiden ersten Berichte relativ konkrete Angaben in Form von ‚best guess‘ Aussagen für ein ‚business-as-usual‘ Szenario machten – beispielsweise eine globale Erwärmung bis 2100 von 3°C (FAR) respektive 2°C (SAR) und eine Erhöhung der Meeresspiegel um 65 cm (FAR) respektive 15-95 cm (SAR) – formulierte man im dritten und vierten Bericht vorsichtiger.⁷ Zum einen wurden unterschiedliche sozioökonomische Szenarien als Ausgangsbasis der Projektionen entwickelt, zum anderen versuchte man, die Ergebnisse in Form von Wahrscheinlichkeitsaussagen zu relativieren. Bereits 1992 hatte man erstmals sechs Emissionsszenarien entwickelt, die für die Berechnung von Zukunftsprojektionen genutzt wurden (vgl. Leggett et al. 1992). Zwischen 1996 und 2000 wurden neue Szenarien entwickelt, welche den Projektionen des dritten und vierten Berichts zugrunde liegen und im SRES Special Report on Emission Scenarios veröffentlicht wurden (vgl. SRES 2000). In diesem Bericht sind vier Storylines (A1, A2, B1, B2) denkbarer sozioökonomischer Zukunftsentwicklungen auf Basis von Annahmen zum Bevölkerungszuwachs, zu technologischen und ökonomischen Entwicklungen, zum Energieverbrauch und der Landnutzung formuliert. Anhand dieser Szenarien berechnete man für insgesamt vierzig Szenarien die zu erwartenden Freisetzen von Treibhausgasen, um damit Vorhersagen zum Klimawandel zu ermöglichen, unter dem Motto: Was wäre wenn? Die Antworten finden sich in den beiden Berichten von 2001 und 2007. So wird im vierten Bericht für das moderateste der Szenarien (B1) eine globale Erwärmung bis 2099 von 1,1 bis 2,9°C und ein Anstieg des Meeresspiegels von 18 bis 38 cm prognostiziert. Für das

7 „In preparation of the TAR, a strong demand for a more systematic approach to uncertainties was identified and the subsequent discussions led to a so-called cross-cutting ‚Guidance Paper‘ on uncertainties (Moss and Schneider 2000). In that guidance paper, Moss and Schneider proposed that authors should use a probabilistic scale that expresses Bayesian confidence estimates about claims in five categories: very low confidence (0-5%), low confidence (5-33%), medium confidence (33-67%), high confidence (67-95%) and very high confidence (95-100%). [...] In WG1, however, the scale proposed by Moss and Schneider was changed into a likelihood scale which was not unequivocally defined by a Bayesian scale” (Petersen 2006: 160, 161).

dramatischste Szenario (A2FI) hingegen werden eine globale Erwärmung von 2,4 bis 6,4°C und ein Anstieg des Meeresspiegels von 26 bis 59 cm errechnet (vgl. IPCC AR4 WG1 Summary 2007: 13).⁸ Diese Projektionen werden im Kontext einer Wahrscheinlichkeitszuordnung als ‚likely‘ (66-90% Wahrscheinlichkeit) eingestuft.⁹

Die soziopolitische Bedeutung der IPCC Berichte sowie ihre Entstehung sind Thema sozial- und politikwissenschaftlicher Forschung. Ihre Bedeutung für die Modellierung und die Professionalisierung der Klimaforschung als Simulationswissenschaft ist bislang noch nicht im Detail untersucht worden. Doch bei näherer Betrachtung zeigt sich, dass die Synchronisierung der Modellierung und des Experimentierens im Rahmen der IPCC Berechnungen ein wichtiges Instrument im Umgang mit in-silico Experimentalsystemen vorangetrieben hat: den Modellvergleich.

„Das ist ja das schöne am IPCC Prozess, es gibt da Modellvergleiche. D.h. die Läufe werden ausgewertet. Derzeit ist es so, dass jede [an den IPCC Szenarien beteiligte] Gruppe einen Standardlauf macht. D.h. unter den gleichen Bedingungen führen alle Gruppen die gleichen Experimente durch. Und das liefert man Lawrence Livermore. Mit ihren riesigen Rechnern und einem großen Stab an Mitarbeitern schaufeln die das alles durch und werten alle Modellergebnisse aus und zwar so, dass sie wirklich vergleichbar sind. Und daher weiß man auch nach jedem IPCC Prozess, wie steht man denn selbst da. Wie verhält man sich zu anderen Modellen? Wo ist man besser, wo ist man schlechter? Verglichen zu Beobachtungen natürlich. D.h. man sucht sich dann die Punkte heraus, wo man schwach ist und versucht sie zu verbessern. Sei es, dass man Parame-

8 „The B1 storyline and scenario family describes a convergent world with the same global population, that peaks in mid-century and declines thereafter, as in the A1 storyline, but with rapid change in economic structures toward a service and information economy, with reductions in material intensity and the introduction of clean and resource-efficient technologies. The emphasis is on global solutions to economic, social and environmental sustainability, including improved equity, but without additional climate initiatives. [...] The A2 storyline and scenario family describes a very heterogeneous world. The underlying theme is self reliance and preservation of local identities. Fertility patterns across regions converge very slowly, which results in continuously increasing population. Economic development is primarily regionally oriented and per capital economic growth and technological change more fragmented and slower than other storylines” (IPCC AR4 WG1 2007: 18).

9 „Virtually certain > 99% probability of occurrence; very likely > 90% probability; likely > 66% probability; about as likely as not 33 to 66% probability; unlikely < 33% probability; very unlikely < 10% probability; exceptionally unlikely < 1% probability” (vgl. IPCC 2005: 4).

trisierung, die Andere schon verwenden auch verwendet, weil man das Gefühl hat, das könnte die Ursache sein dafür, dass sie besser sind. Sei es, dass man sich nach Neuem umsieht. Sei es, dass man selbst etwas entwickelt“ (Interview 19, 2006).

Bereits 1978 hatte Jule Charney drei Atmosphärenmodelle miteinander verglichen (vgl. Phillips 2000a). Doch für Klimamodelle begann die Ära des Modellvergleichs erst in den 1990er Jahren mit den großangelegten Projekten des WCRP World Climate Research Programme: AMIP und CMIP. Ab 1990 wurden umfangreiche Studien zum Modellvergleich im AMIP Atmospheric Model Intercomparison Project organisiert, das 1995 in das CMIP Coupled Model Intercomparison Project überführt wurde (vgl. AMIP 2009, CMIP 2009). Beide Projekte stehen im direkten Zusammenhang mit den IPCC Referenzmodellen und Szenarienberechnungen. Doch Modellvergleiche leisten mehr: „Overall, the vigorous, ongoing intercomparison activities have increased communication among modelling groups, allowed rapid identification and correction of modelling errors and encouraged the creation of standardised benchmark calculations, as well as a more complete and systematic record of modelling progress“ (IPCC AR4 WG1 2007: 594). Eine Folge des verstärkten Austausches zwischen den Forschergruppen ist die Einrichtung von internetbasierten Plattformen, von welchen Modelle, Ergebnisse und Testberichte heruntergeladen werden können. Die M&D Modelle & Daten Gruppe in Hamburg administriert neben dem World Data Center for Climate eine solche Plattform für den Modell- und Datenaustausch.

„While the development and operation of numerical climate models is becoming more complex, the model development is distributed among an increasing number of institutes. It is not just the pure number of program code-lines and the challenges for software engineers that are growing by the increasing complexity of models, model environments and computer architecture, but also the number of computed results and aspects of analysis. [...] The considered workflows are

- * Integration of new model components, like e.g. atmospheric chemistry, land surface processes and inland-ice, into a coupled numerical model system (Earth System model),
- * Porting of Earth System models to different computer architectures,
- * Operating complex Earth System models for cost-effective computations via chain experiments,
- * Postprocessing, analysis and visualisation of the computation results, size of data sets amounting to several Terrabytes,

- * Documentation, application-oriented long-term data storage and validation of the computation results for own and external projects,
- * International distribution of Earth System research data and their integration into common information systems.

The workflows are processed exemplarily by M&D and abstracted. From this knowledge tool sets are developed, which will facilitate the day-to-day work of scientists. The development of tool sets was highly supported by the cooperative work in the EU-project PRISM. In the framework of this project several Earth System models from different institutes equipped with a variety of computer architectures were made available for testing purposes. The aim of the activities of M&D is to develop and promote best practice methods to handle Earth System models and their results“ (M&D 2009: About us, Mission).

Modelle, die früher nur in den jeweiligen Instituten oder im informellen Austausch kursierten, werden nun auf solchen Plattformen als Community Modelle öffentlich zugänglich koordiniert. Dies bedeutet, dass weitere Standardisierungen von Nöten sind, da neue Nutzergruppen die Modelle und Daten für ihre Forschung verwenden. Während es in den Experimentalwissenschaften üblich ist, akribische Dokumentationen der Experimente und Experimentalanordnungen zu geben, „[is] this same level of methodological detail [...] only rarely provided for in numerical simulation or modeling studies. This is true in spite of the fact that numerical experimental systems are inherently reproducible; unlike a physical experiment, which may be subject to many sources of uncontrollable variability, a valid numerical result can always be replicated exactly“ (Thornton et al. 2005: 431). Mit der CERA Datenbank des World Data Center for Climate wird versucht, diese Standardisierung voranzutreiben. „Model describing MD helps data users to understand how data have been generated. This becomes an important issue in projects where multi-model data are analyzed by scientists not knowing the numerical models in detail (e.g. CMIP projects, IPCC). For these users it may be sufficient to know which schemes have been used and perhaps the values of characteristic parameters. While this will not enable them to reproduce model output, it may put them into the position to understand differences in model results from different models“ (M&D 2009: IMID, Model Metadata).

Diese neuen Nutzer führen in der Forschung mit den in-silico Experimentalsystemen zur Ausdifferenzierung von Sub-Communities: Neben den Modellierern wächst die Anzahl der Modellnutzer und auch der Modelldatennutzer. Diese Sub-Communities sind zumeist institutionell voneinander unterschieden. Während die Modellierung von gekoppelten

Atmosphären-Ozean Modellen, wie sie mittlerweile Standard für IPCC Szenarien sind, in einer überschaubaren Anzahl von Instituten wie dem GFDL Geophysical Fluid Dynamics Laboratory in Princeton (GFDL-CM), dem Max-Planck-Institut für Meteorologie in Hamburg (ECHAM/MPI-OM), dem IPSL Institut Pierre Simon Laplace in Paris (IPSL-CM), dem BCC Beijing Climate Centre in Beijing (BCC-CM) oder dem Hadley Centre for Climate Prediction in Exeter (UKMO-HadCM) geschieht, finden sich die Modell- und Modelldatennutzer vor allem im universitären Bereich oder in Instituten, die im Rahmen der zunehmenden Interdisziplinarität der Erdsysteme für ihre Forschungen auf spezifische Modell- und Datenauswertungen der Klimaforscher angewiesen sind. Durch den aktuellen Trend zur Modellierung von Erdsystemen mit Biosphäre, Ökosphäre und Anthroposphäre ist ein weiterer Zuwachs der Nutzergruppen zu erwarten. Während Modellnutzer bestehende Modelle modifizieren und dann für ihre spezifischen Experimente verwenden, greifen Datennutzer auf die in-silico und in-situ Daten der Community zurück und analysieren diese auf spezifische Fragestellungen hin. Diese Spezialisierung kann als ein Indiz der Professionalisierung im Umgang mit den Modellen gewertet werden. Der Wissensaustausch geht dabei in beide Richtungen. Die Nutzer von Modellen geben in der Regel den Modellierern ein Feedback mit Hinweisen zu Problemen und Modellverbesserungen, die dann in die Programmierung der nächsten Modellversion einfließen. Ebenso nutzen die Modellierer Analysen der Datennutzer zu ihren in-silico Experimentaldaten, um ggf. die Modelle zu verbessern. Meist werden diese Austauschprozesse in Workshops organisiert.

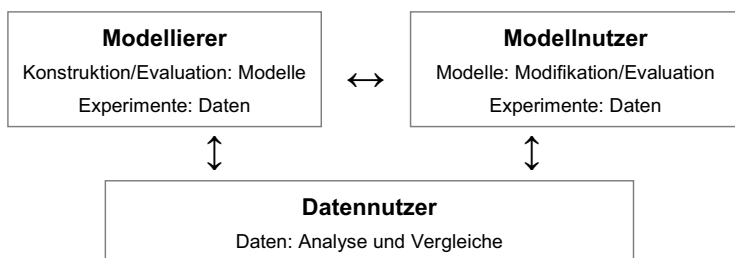


Abbildung 20: Sub-Communities der gekoppelten Atmosphären-Ozean-Modelle und Erdsysteme (Gramelsberger 2009)

Modellkomplexität

Die Ausdifferenzierung in Sub-Communities und die Zunahme der Modell- und Datennutzer weist auf eine weitere Entwicklung hin: die zunehmende Modellfülle und Modellkomplexität. Zirkulationsmodelle sind nur eine Art von Modellen, die in der Klimaforschung Verwendung finden. EMICs Earth System Models of Intermediate Complexity, Boxmodelle, konzeptuelle Modelle und Prozessmodelle sind weitere gängige Modellarten. EMICs werden seit gut zehn Jahren entwickelt und sie werden vor allem dafür genutzt, die Interaktion zwischen verschiedenen Modellkomponenten zu untersuchen. Da sie vereinfachte Versionen der gekoppelten Atmosphären-Ozean-Modelle sind und daher weniger Rechenkapazitäten benötigen, lassen sich mit ihnen wesentlich mehr Experimente und diese über längere Zeiträume durchführen. EMICs sind auch für Sensitivitätsstudien geeignet, um die Sensitivität des Modells auf Parameterjustierungen zu testen.

Die verschiedenen Modellarten werden oft in einer hierarchischen Einteilung von großen Modellen (Erdsysteme, Zirkulationsmodelle) zu kleinen Modellen (EMICs, konzeptuelle Modelle) gewertet (vgl. Henderson-Sellers, McGuffie 1987). Die Erfahrung der letzten Jahre bezüglich der Modellevaluierung zeigt jedoch, dass das gesamte Spektrum an Modellen notwendig ist, da sich die Modelle gegenseitig ergänzen und voneinander abhängig sind. So benötigen EMICs die getesteten Parametrisierungen der großen Modelle, während diese oft zu komplex sind, um detaillierte Studien zur Sensitivität der Modelle durchzuführen. Boxmodelle dienen, wie das Beispiel der Fehlersuche in der Vulkanparametrisierung zeigte, dem Test einzelner Modellkomponenten und konzeptuelle Modelle wiederum sind für die Veranschaulichung grundlegender Prozesse unabdingbar.

„Definition of climate in terms of the state and statistics of the climate system components is now commonly accepted, at least in the community of climate modelers. Because of the broad spectrum of typical time scales of the different components of the climate system, simulation of climate system dynamics requires different types of models. [...] Generally, we argue that there is a clear advantage in having available a spectrum of climate system models. Most EMICS are specifically designed for long-term simulations over many millennia, and some are designed to simulate the interaction of as many components of the climate system as possible in an efficient manner. Moreover, EMICs can explore the parameter space with some completeness. Thus, they are more suitable for assessing uncertainty, which CGCMs [coupled atmosphere and ocean general circulation models] can do to a significant lesser extent. On the other hand, it would not be sensible to apply EMICs to studies

which require high spatial resolution. EMICs can also be used to screen the phase space of climate or the history of climate to identify interesting time slices, thereby providing guidance for more detailed investigations to be undertaken by CGCMs. For the interpretation of model results, conceptual models appear to be very useful“ (Claussen et al. 2002: 584, 595).

Der Trend zur Modellierung von Erdsystemen zeigt, dass die Vielfalt der Modelle sowie ihre Kombinationsmöglichkeiten steigen. Die Vision einer ‚plug-and-play‘ Umgebung für in-silico Experimentalsysteme lässt jede Art von Modell zum integrierbaren Modul werden. Standardisierungen sowie die Entwicklungen im Software Engineering der Modelle – Kapselung, definierte Input/Output-Schnittstellen, etc. – unterstützen diese Tendenz. Doch neben der zunehmenden Fülle an Modellen steigt auch die Modellkomplexität rapide an.¹⁰ Während der dynamische Kern der Zirkulationsmodelle relativ alt ist und auf Vilhelm Bjerknes Konzept von 1904 zurückgeht, wächst die Anzahl der integrierten Prozesse in Form von Sub-Modellen und Parametrisierungen.

Parametrisierungen sind notwendig, da die Diskretisierung der dynamischen Grundgleichungen den meteorologischen Variablen nur einen Mittelwert je Volumenelement zuordnet, der subskalige Bewegungs- und Transportvorgänge nicht erfasst. Um diese subskaligen Prozesse in ihrer Wirkung auf die gemittelten Größen zu berücksichtigen, bedarf es expliziter Beschreibungen in Form von Parametrisierungen, wie das Beispiel des Schmelzvorgangs in stratiformen Wolken illustrierte. „Man löst dies Problem durch Hinzufügungen von Parametrisierungen, d.h. von Termen, die den Netto-Effekt der nichtaufgelösten Prozesse auf die durch Diskretisierung aufgelösten Prozesse darstellen. Diese Terme, oder Einflüsse von im wesentlichen Unbekanntem, werden als (deterministische) Funktionen der aufgelösten, mittleren Zustandsvariablen formuliert“ (von Storch, Güss, Heimann 1999: 109). Interessanter Weise gibt es weder in der Mathematik noch in der Meteorologie eine einheitliche Definition dessen, was unter dem Begriff des Parameters verstanden wird. „There are four possibilities to derive parameterizations: 1) derive an analytical expression of the process of interest from first principles. 2) Use results from laboratory studies and assume that the relationship amongst the variables is still valid in the real atmosphere. 3) Use data from field measurements: either long-term measurements or measurement campaigns. 4) Use detailed models which resolve the

10 Komplexität nimmt hier Bezug auf die Anzahl der integrierten Prozesse und nicht auf die mathematische Komplexität nicht-linearer Prozesse, wie sie in der Dynamik der Modelle Anwendung findet.

process of interest and derive a parameterization based on that“ (Lohmann et al. 2006: Submission).

Die Parametrisierungen in den Modellen wachsen enorm an und ermöglichen es, immer mehr Prozesse zu integrieren. So ‚schwimmen‘ in den simulierten Ozeanen mittlerweile Fische, denn auch sie haben Einfluss auf den Kohlendioxidhaushalt der Meere.¹¹ Doch Parametrisierungen gelten als problematisch, da sie nur selten auf physikalischen Grundprinzipien basieren (first principles). Meist beruhen sie auf funktionalen Zusammenhängen oder empirisch begründeten Annahmen, die aus Beobachtungen und Messkampagnen stammen. Die Problematik resultiert aus der Anpassung der empirisch gewonnenen Parametrisierungen an die großskaligen und gemittelten Modelle anhand von Messergebnissen, die sich auf Punktmessungen oder kleine Messareale beziehen und somit kleinräumige Effekte widerspiegeln. Parametrisierungen sind daher oft lokalen Beschränkungen unterworfen, während sie im Modell global gültig sind. Als Alternative zu diesen empirischen Parametrisierungen sind Prozessmodelle für kleinskalige Prozesse denkbar, beispielsweise hoch aufgelöste Wolkenmodelle, aus welchen dann Parametrisierungen gewonnen werden. Ein Großteil der Forschungs- wie der Modellierungsarbeit konzentriert sich daher auf die Integration adäquater Parametrisierungen in die Modelle. „Man benötigt viel Zeit, bis man neue Parametrisierungen modelliert und codiert hat. Diese werden erst im Einzelnen getestet und dann im Ensemble“ (Interview 19, 2006). Die

-
- 11 Ozeane nehmen CO₂ auf und speichern es. Daher ist es für die Klimaforscher von Interesse, diese Speicherprozesse in das Modell zu integrieren. CO₂ wird von dem, an der Ozeanoberfläche schwimmenden Phytoplankton aufgenommen. Phytoplankton steht am Beginn der Nahrungskette und wird von Zooplankton und Fischen gefressen. Dadurch gelangt der Kohlenstoff relativ schnell wieder in den Ozean und die Atmosphäre zurück. Sinkt das Plankton hingegen zu Boden, wird Kohlenstoff langfristig gespeichert. Dieser Prozess lässt sich modellieren und in die Ozeanmodelle integrieren. Aktuell geschieht das in Form eines Plankton-Fisch-Kohlenstoff Parameters. Dieser Parameter ist, wie all die anderen Parameter eines Klimamodells, ein mathematischer Term der charakteristischen Plankton-Fisch-Kohlenstoff Veränderungsrate gemittelt auf den gesamten Ozean. Veränderungsraten basieren auf mathematischen Termen, welche die Abnahmen (Senken) und Zunahmen (Quellen) eines Prozesses angeben. Im Beispiel des Plankton-Fisch-Kohlenstoff Parameters nimmt der freigesetzte Kohlenstoff zu, wenn die Anzahl der Fische zunimmt, da dann mehr Plankton konsumiert wird und dadurch der im Plankton gespeicherte Kohlenstoff über die Fische freigesetzt wird. Dieser komplexe Prozess wird als global gemitteltes, relativ einfaches Verhältnis abstrahiert und als mathematischer Term in das Modell eingefügt. Auf diese Weise werden zahlreiche subskalige Prozesse in die Modelle integriert.

Problematik der Parametrisierungen liegt darüber hinaus in jenen Parametern, die sich nicht messen lassen. Sie müssen justiert werden, doch es gibt kaum Anhaltspunkte, wie die Justierungen geeicht werden können. Dies ist einer der Hauptkritikpunkte an Klimamodellen.

„The parametrizations also involve numerical parameters that must be specified as input. Some of these parameters can be measured, at least in principle, while others cannot. It is therefore common to adjust parameter values (possibly chosen from some prior distribution) in order to optimise model simulation of particular variables or to improve global heat balance. This process is often known as ‚tuning‘. It is justifiable to the extent that two conditions are met: 1. Observationally based constraints on parameter ranges are not exceeded. Note that in some cases this may not provide a tight constraint on parameter values (e.g., Heymsfield and Donner, 1990). 2. The number of degrees of freedom in the tuneable parameters is less than the number of degrees of freedom in the observational constraints used in model evaluation. This is believed to be true for most GCMs“ (IPCC AR4 WG1 2007: 596).

Da die Zirkulationsmodelle jedoch alle auf derselben Dynamik beruhen, unterscheiden sie sich in erster Linie in ihren Parametrisierungen. „Es ist eigentlich so, dass wir neue Parametrisierungen einführen, die über das hinausgehen, was andere haben und wir von daher neue Fragen stellen und beantworten. Und es hat jedes Institut so seine Spezialitäten. Aber im Großen und Ganzen wird schon viel von der Community vorgegeben“ (Interview 19, 2006).

Evaluierung

Computerexperimente benötigen auf Grund ihres rein semiotischen Status neue Evaluierungsstrategien. Im Unterschied zu Laborexperimenten können sie nicht auf das Korrektivum der materialen Widerständigkeit hoffen und insofern auch nicht scheitern.¹² Als Erkenntnisinstrumente, die in der Tradition von quantitativen Prognosen und Präzisionsmessungen stehen, werden mit Computerexperimenten Fluten von Daten generiert. Die klassische Version der Hypothesenprüfung an „empirisch möglichst leicht nachprüfbarer bzw. anwendbarer singulärer Folgerungen

12 Allenfalls ein Computercrash lässt vermuten, dass innerhalb des Modells etwas konfligiert, wie das Beispiel der Fehlersuche in der Vulkanparametrisierung verdeutlichte. Dies muss allerdings nicht automatisch bedeuten, dass die neu hinzugefügte, codierte Theorie falsch ist. Ein Computerabsturz kann auch auf bereits bestehende Fehler oder Unzulänglichkeit im Modell hinweisen, die zuvor nicht zu Tage getreten sind.

(„Prognosen“)⁴ (Popper 1935/1989: 8) entspricht der Ja/Nein-Logik physikalischer Experimente. Die Evaluierungsstrategien, die in den letzten Jahrzehnten entwickelt wurden, tragen jedoch der Eigenart der Computerexperimente und ihren Datenfluten Rechnung. Einen guten Überblick über die Entwicklung liefert das achte Kapitel *Climate Models and Their Evaluation* des aktuellen IPCC Berichts (vgl. IPCC AR4 WG1 2007: 589-662). Dabei lassen sich zwei Arten von Evaluationsstrategien für die Resultate der Computerexperimente aufzeigen: die Evaluation an der Empirie (Messdaten) wie auch an Modellen. Zudem haben sich zwei unterschiedliche Perspektiven der Evaluation von in-silico Experimentalsystemen herausgebildet: die Evaluation auf Systemebene als auch auf Komponentenebene. „A climate model is a very complex system, with many components. The model must of course be tested at the system level, that is, by running the full model and comparing the results with observations. Such tests can reveal problems, but their source is often hidden by the model's complexity. For this reason, it is also important to test the model at the component level, that is, by isolating particular components and testing them independent of the complete model“ (IPCC AR4 WG1 2007: 594). Die Auflösung der Komponentenevaluierung reicht von Modellteilen und Modulen bis zu einzelnen Parametrisierungen.

Mathematische Tests zu numerischen Verfahren und Fallstudien zu Parametrisierungen werden in internationalen Messkampagnen, Projekten und Workshops organisiert. In solchen Workshops wird beispielsweise die Dynamik der teilnehmenden Modelle hinsichtlich ihrer Performance und Zuverlässigkeit unter standardisierten Vorgaben getestet und miteinander verglichen. Parametrisierungen werden in aufwendigen Kampagnen überprüft und verbessert. Das EUROCS European Cloud Systems Projekt beispielsweise untersuchte die Parametrisierung von Wolken in Klimamodellen. Ziel war es, durch Analyse von Mess- und Simulationsdaten die Wolkenschemata zu verbessern.

„Clouds probably remain the largest source of uncertainty affecting evaluations of climate change in response to anthropogenic change. That explains for a large part why the range of simulated temperature changes in response to a CO₂ doubling (1.5 to 4.5 C) is quite invariant for almost 20 years [...]. The recent interest to develop capability to predict regional climate changes stress the importance to better represent clouds in models. [...] The strategy used in EUROCS to address these issues is based on the use of a hierarchy of models and observations to integrate cloud studies across the full range of scales. Numerical models range from General Circulation Models (GCMs) through Single Column Models (SCMs) to Cloud Resolving Models (CRMs) and Large Eddy Simulations (LES). Observations which will be used will vary from

global satellite measurements to local observations of individual clouds through lidar and millimetric radars.“ (EUROCS 2003: Project Description).

Das Ziel der Evaluation einzelner Komponenten ist es, Benchmark Tests der Prozesse und Metriken zur Zuverlässigkeit der Resultate zu entwickeln.¹³ Doch um überhaupt Komponenten testen zu können, bedarf es zu allererst einer Evaluation des gesamten Modells auf Systemebene. Grundlegendster Test hierbei ist die Evaluation der Modellergebnisse im Vergleich zum aktuellen Klima. Dazu werden die Modelle mit historischen Daten gestartet, beispielsweise mit Klimadaten aus der Zeit vor der Industrialisierung oder noch früher. Repräsentieren die Ergebnisse das aktuelle Klima, dann wird davon ausgegangen, dass man das Modell auch für Zukunftsprojektionen nutzen kann. Die Evaluation eines Modells berücksichtigt auch bestimmte Muster, die im Wechsel der berechneten Jahreszeiten auftreten, wie die Konversion der Wetterlagen oder regelmäßig auftretende Wetterphänomene. Bildet ein Modell diese Muster nicht aus, ist es fehlerhaft. Eine weitere Möglichkeit der Evaluation bietet sich im Vergleich zwischen Klima- und Wettermodellen. Beide basieren auf derselben Dynamik und auf ähnlichen Prozessen. Doch Wettermodelle sind wesentlich höher aufgelöst, in der Regel auf einige Kilometer. Dadurch lassen sich bestimmte Prozesse zwischen beiden Modelltypen vergleichen und untersuchen.

Eine relative neue Art, Klimamodelle zu evaluieren sind Ensembleprognosen. Mit der steigenden Leistungsfähigkeit der Computer ist es möglich geworden, zunehmend umfangreichere aber auch zahlreichere Simulationsläufe zu berechnen. Da Klimamodelle komplexe Systeme sind, die sensitiv auf die gewählten Anfangsbedingungen reagieren, versucht man mit Ensemble-Prognosen den Grad der Sensitivität eines Modells zu bestimmen. Indem Hunderte von Modell-Läufen bei leicht variierenden Anfangsbedingungen durchgeführt werden, lässt sich erkennen, wie sensitiv sich das Modell verhält. Weichen die Resultate der einzelnen Modell-Läufe zu stark voneinander ab, da das Modell zu sensitiv reagiert, gilt es, die Prognosekraft des Modells kritisch zu hinterfragen. Doch Ensemble-Prognosen leisten noch mehr. Es wird davon ausgegangen, dass der aus den zahlreichen Läufen errechnete Mittelwert statis-

13 „A number of different observationally based metrics have been used to weight the reliability of contributing models when making probabilistic projections [...] The above studies show promise that quantitative metrics for the likelihood of model projections may be developed, but because the development of robust metrics is still at an early stage, the model evaluations presented in this chapter are based primarily on experience and physical reasoning, as has been the norm in the past” (IPCC AR4 WG1 2007: 594, 595).

tisch betrachtet eine höhere Prognosegüte hat als einzelne Experimente. Daher sind Projektionen basierend auf Ensembleprognosen zuverlässiger.

Ziel dieser Evaluationsbemühungen ist es, die Modelle wie auch Modellergebnisse verlässlicher zu machen. Wie die zu niedrigen Temperaturen in bestimmten Höhen basierend auf einer nicht adäquaten Wasserdampfverteilung im Modell (cold bias) zeigt, repräsentieren globale Klimamodelle regional nie exakt das aktuelle Klima. In der Summe gleichen sich diese Schwächen jedoch oft aus. Da Klima, dem Programm von Julius Hann folgend, der gemittelte Zustand der Atmosphäre basierend auf der Summe aller meteorologischen Phänomene ist (vgl. Hann 1908), ist die Regionalisierung in globalen Modellen eine Herausforderung für die Klimamodellierung. Da man sowohl für die Vorhersage zukünftiger Entwicklungen, als auch für Experimente mit dem Klima allein auf in-silico Experimentalsysteme angewiesen ist, lässt sich nachvollziehen, warum ein derartiger Aufwand für die Modellierung und Evaluierung betrieben wird. Dabei verändert sich sozusagen nebenbei die Forschungslandschaft als auch die Forschungspraxis und gestaltet die Meteorologie in eine paradigmatische e-Science um.

Reshaping Science

Die Umgestaltung der Meteorologie zur paradigmatischen e-Science mit ihrer Leitfunktion im Umgang mit in-silico Experimentalsystemen und Computerexperimenten, zeigt sich an den dargestellten Indikatoren: Der Standardisierung und Synchronisierung der in-silico basierten Forschungspraxis, der Ausdifferenzierung von Sub-Communities, der koordinierten und zunehmend standardisierten Modelldistribution innerhalb der Klima- und Erdsystemforschung, der Möglichkeit des Modellvergleichs, der Etablierung einer internationalen Infrastruktur der Modellevaluierung, der zunehmenden Standardisierung der Modelleinbettung für plug-and-play Erdsysteme und der, mit diesen Entwicklungen einhergehenden Professionalisierung der Modellierung. Eine wesentliche Voraussetzung hierfür ist, dass der Einfluss des Computers seit den 1950er Jahren zu einer Vereinheitlichung der Meteorologie führte. „As a result of the growth of science, the availability of computers, and the new style of research and practice, meteorology became a much more unified discipline“ (Nebeker 1995: 184). Wurde die Meteorologie bezüglich ihrer Messkampagnen im 20. Jahrhundert zur Großforschung, so ist sie es nun auch bezüglich ihrer in-silico Experimentalsysteme geworden. Aktuell werden die Vorhersagemodelle, wie die gekoppelten Atmo-

sphären-Ozean-Modelle oder die Erdsysteme, aus der Forschung herausgenommen und in einen operationellen Betrieb überführt.¹⁴ Die Administration dieser Modelle sowie ihre Zurverfügungstellung für immer mehr Nutzer, auch außerhalb der Meteorologie, erfordert diesen Schritt. Die Folge ist, dass sich die in-silico Experimentalsysteme in ihrem Status zunehmend zu Großforschungsinstrumenten ähnlich Satelliten wandeln, die Serviceleistungen erbringen – für die Wissenschaft oder, wie im Falle der Wettervorhersage und der Klimaprojektionen, für Industrie und Gesellschaft.

In Hinblick auf die Neuheit der Simulation als Instrument wissenschaftlicher Erkenntnisproduktion, lässt sich diese Entwicklung als zunehmende Etablierung der Simulation als gängiges Erkenntnisinstrument deuten. Dabei verändert sich jedoch nicht nur die Forschungspraxis erheblich, sondern auch die Logik und Epistemik der Forschung. „The large computer models, either for forecasting or for climate simulations, provided a way of tying together theoretical results. [...] Indeed, the research on the individual processes was often motivated by the desire to improve large models“ (Nebeker 1995: 184). Diese Entwicklung führt dazu, dass „a variety of physical processes could be synthesized mathematically within the computer“ (Charney 1972: 117). Als Synthesizer für Theorien verstanden, als synoptisches Theorieinstrument oder als Theoriebaukasten unterliegt der Computer einer anderen Logik, als der analytischen, auf einzelne Prozesse konzentrierten Erkenntnisinstrumente. Es geht um eine neue Art des Experimentierens mit Messdaten aufgeladenen Theorien. Doch wenn sich die Epistemik und Forschungslogik ändert, dann ändert sich auch das Bild der Welt, wenn sie aus dieser neuen Perspektive betrachtet wird. Der Vergleich mit der wissenschaftlichen Revolution in der Neuzeit scheint daher angebracht. Ebenso wie sich damals die Methoden und Blickweisen änderten und damit das, was als wissenschaftliche Erfahrung galt, vollzieht sich mit dem Computer eine ähnliche Veränderung. Analog zu den Transformationen des 17. Jahrhunderts verändert sich die Wissenschaft als Computational Science heute. „There is no one independently given class of practises that naturally correspond to the label „scientific experimentation“ [heute: scientific simulation]; there are many different practices, with their associated epistemological characterizations, that relate to experience and its place

14 Diese Entwicklung der Klimamodellierung hin zur Großforschungseinrichtung und zum operationellen Betrieb der Modelle lässt sich an den aktuellen Umstrukturierungen des NCAR National Center for Atmospheric Research in Boulder und dem dortigen CCM Climate Community Model erkennen.

in the creation of natural knowledge. In the seventeenth century [heute: twenty-first century] old practices changed and new ones appeared. Those changing practises represent shifts in the meaning of experience itself – shifts in what people saw [heute: see] when they looked at the events in the natural world“ (Dear 1995: 12, 13). Wie dieser „shift in the meaning of experience itself“ heute aussieht, wird im Rahmen der philosophischen Verortung der Computerexperimente zu klären sein.

