

Leitlinien zur Interpretation regionaler Klimamodelldaten

Erstellt durch das

Bund- Länder- Fachgespräch

"Interpretation regionaler Klimamodelldaten"

(Stand: Nov. 2020)

Die vorliegenden Leitlinien sind das Ergebnis eines mehrjährigen Arbeitsprozesses und einer permanenten Fortschreibung. Das Bund-Länder-Fachgespräch „Interpretation regionaler Klimamodelldaten“ besteht seit 2005 als ein informeller Erfahrungsaustausch. Darin sind nachstehende Einrichtungen regelmäßig vertreten (Stand Nov. 2020):

Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg (LUBW),
Bayerisches Landesamt für Umwelt (LfU),
Senatsverwaltung für Stadtentwicklung und Umweltschutz Berlin (SenStadtUm),
Landesamt für Umwelt, Brandenburg (LfU),
Der Senatorin für Klimaschutz, Umwelt, Mobilität, Stadtentwicklung und Wohnungsbau
Bremen (SKUMS),
Behörde für Stadtentwicklung und Umwelt Hamburg (BSU),
Hessisches Landesamt für Naturschutz, Umwelt und Geologie (HLNUG),
Ministerium f. Energie, Infrastruktur, Landesentwicklung Mecklenburg-Vorpommern (EMMV),
Niedersächsisches Ministerium für Umwelt, Energie, Bauen und Klimaschutz (NMU),
Niedersächsischer Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (NLWKN),
Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen (LANUV),
Landesamt für Umwelt, Rheinland-Pfalz (LfU),
Rheinland-Pfalz, Kompetenzzentrum für Klimawandelfolgen (RLPKK),
Ministerium für Wirtschaft, Arbeit, Energie und Verkehr Saarland (MWAEV),
Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie (LfULG),
Landesamt für Umweltschutz Sachsen-Anhalt (LAU),
Landesamt für Landwirtschaft, Umwelt und ländliche Räume Schleswig-Holstein (LLUR),
Thüringer Landesanstalt für Umwelt und Geologie (TLUG),
Umweltbundesamt (UBA; KomPass),
Deutscher Wetterdienst (DWD),
Bundesanstalt für Gewässerkunde (bfg),
Climate Service Center Germany (GERICS),

Kontakt Daten für das Fachgespräch „Interpretation regionaler Klimamodelldaten“:

Fachgesprächsleitung:

Dr. Matthias Zimmer

Landesamt für Umwelt Rheinland-Pfalz

Rheinallee 97-101

55118 Mainz

Telefon 0 61 31 60 33-1644

Telefax 0 61 31 60 33-1680

Matthias.Zimmer@lfu.rlp.de

<http://www.klimawandel-rlp.de/>

Redaktionsleitung:

Carsten Linke

Landesamt für Umwelt Brandenburg

Seeburger Chaussee 2

14476 Potsdam

Telefon: 0 331201 442-322

carsten.linke@lfu.brandenburg.de

<http://www.lfu.brandenburg.de>

Zitierhinweis für die vorliegenden Leitlinien:

Linke, C. et al, 2020: Leitlinien zur Interpretation regionaler Klimamodelldaten des Bund-Länder-Fachgespräches „Interpretation regionaler Klimamodelldaten“, Potsdam, Nov.2020

Inhalt

1.	Motivation zur Erstellung der Leitlinien	5
2.	Allgemeines zur Klimamodellierung.....	6
2.1	Beobachtungsdaten	6
2.2	Szenarien.....	6
2.3	Globale und regionale Klimamodelle.....	7
2.4	Wirkmodelle in einer Modellkette.....	8
2.5	Grenzen der Modellierbarkeit	9
3.	Regionale Klimaprojektionen – eine Übersicht	10
3.1	Verfügbare Simulationen.....	10
3.2	Zeitliche und räumliche Betrachtung	10
3.3	Bezugsgrößen der Klimaänderungssignale.....	11
3.4	Prüfung von Klimamodelldaten	11
4.	Auswertung und Darstellung	13
4.1	Rahmenbedingungen.....	13
4.2	Ensemblebetrachtung	13
4.3	Bias-Minimierung	14
4.4	Statistische Auswertung und Extremwerte	15
4.5	Robustheit der Klimaänderungssignale	16
4.6	Ergebnisdarstellung	17
5.	Interpretation und Vermittlung der Ergebnisse.....	19
6.	Hinweise zur Wirk- bzw. Impaktmodellierung	21
7.	Konsolidierte Wortwahl.....	23
8.	Glossar.....	25
9.	Literaturverzeichnis	30
10.	Anlagen	34
10.1	Anlage 1: Beispiel Parameterdarstellung.....	34
10.2	Anlage 2 Ensembledarstellung.....	35
10.3	Anlage 3: Beispiel Gesundheit / Hitzestress.....	37
10.4	Anlage 4: Beispiel Klimatische Leitplanken	38
10.5	Anlage 5: Beispiel Obstbau (Pilzkrankheit: Apfelschorf)	43
10.6	Anlage 6: Beispiel Tourismus (Schwimmbadbesucher).....	44
10.7	Anlage 7: Prüfung von Modellergebnissen auf Plausibilität	46
10.8	Anlage 8: Beispiel für die Berechnung einer Bemessungsschwelle.....	50
10.9	Anlage 9: Beispiel zur Identifizierung von robusten Klimaänderungen.....	57
10.10.	Anlage 10: Ein Kernensemble von regionalen Klimaprojektionen	60

1. Motivation zur Erstellung der Leitlinien

Im Zusammenhang mit dem Klimawandel sind für die Entwicklung von Anpassungsstrategien und Anpassungsmaßnahmen zu den verschiedenen Handlungs- und Politikfeldern Angaben zu möglichen Klimaentwicklungen und deren Bandbreiten notwendig.

Für Deutschland liegen die Ergebnisse von Projektionsläufen mehrerer regionaler Klimamodelle und die Ergebnisse des ENSEMBLES-Projektes (<http://www.ensembles-eu.org/>) sowie der CORDEX-Initiative (<http://www.euro-cordex.net/>) vor. Für die Nutzung, Auswertung und Interpretation solcher Klimaprojektionen sind spezielle Kenntnisse und eine sachgerechte Herangehensweise notwendig, da sonst erhebliche Fehlinformationen vermittelt werden können. Der fachgerechte Umgang mit den umfangreichen und komplexen Daten erfordert ein Instrumentarium nach dem Stand der Wissenschaft. Die hier vorgelegten Leitlinien sind ein erster Schritt auf dem Weg zu einem 'Good-Practice-Standard'.

Die hier vorgelegten Leitlinien dienen:

- als Grundlage für eine sachgerechte Datennutzung, der Qualitätssicherung von Auswertungen und einer zweckmäßigen Darstellung der Ergebnisse,
- der Vermeidung von Fehlinterpretation von Sachverhalten und der besseren Entscheidungsfindung sowie Politikberatung,
- als Instrument der Homogenisierung der zahlreichen Arbeiten im Bereich der regionalen Klimadatenanalyse und der Vergleichbarkeit von Auswertungen durch Bundes-, Landeseinrichtungen und Institute für die Bundesländer, Regionen und sachlichen Teilräume,
- als Grundlage zur Findung und Nutzung einer gemeinsamen Fachsprache und des besseren Problemverständnisses,
- als Voraussetzung für die sachgerechte und zielgerichtete Nutzung von Klimadaten als Antrieb für die Wirkmodelle in den zahlreichen Handlungsfeldern im Bereich der Anpassungsforschung.

Die Leitlinien erläutern wesentliche Elemente der Klimamodellierung und geben Hinweise über die Nutzung von Klimaprojektionsdaten. Es werden Empfehlungen zur Auswertung von Klimamodelldaten, zu deren Darstellung und zu deren Interpretation gegeben.

Die Teilnehmer am Bund-Länder-Fachgespräch „Interpretation regionaler Klimamodelldaten“ haben das Selbstverständnis, dass diese Leitlinien bei der Auswertung, Interpretation und Nutzung von Klimadaten eingehalten werden. Eine Anwendung dieser Leitlinien auch durch Dritte (z.B. Bundes-, Landesbehörden, Institute, Ingenieurbüros) wird empfohlen.

Die Leitlinien stellen den Stand der Praxis und der Wissenschaft dar und werden nach Bedarf und Erkenntnisstand fortgeschrieben.

2. Allgemeines zur Klimamodellierung

Die Erkenntnisse zum gegenwärtigen Klima und dem beobachteten Klimawandel beruhen auf der Auswertung von langen Messreihen (Beobachtungsdaten). Die Abschätzung des Klimawandels in der Zukunft, der vom Menschen wesentlich verursacht ist, ist nur durch Klimamodellierung möglich.

2.1 Beobachtungsdaten

Globaler, regionaler und bedingt auch lokaler Klimawandel können durch die statistische Analyse der Zeitreihen von Beobachtungsdaten nachgewiesen werden. Die Analyse erfordert eine Qualitätsprüfung der Beobachtungsdaten bezüglich der Homogenität des Standortes, der Messverfahren und der Messinstrumente sowie von Lücken, Ausreißern, etc. Die Trendermittlung sollte auf robusten statistischen Verfahren beruhen und durch einen Signifikanztest gesichert sein (s. z.B. Hennemuth et al., 2013).

Untersuchungen können eine Übertragung der stationsbasierten Beobachtungsdaten in die Fläche oder auf ein Gitter erforderlich machen. Die dabei verwendeten Interpolationsverfahren müssen speziell für die jeweilige Aufgabe geeignet sein.

Die Übertragung von weltweit routinemäßig durch die Wetterdienste gewonnenen Beobachtungsdaten auf ein dreidimensionales Gitter erfolgt durch die sogenannte Reanalyse. Hierbei werden die beobachteten Daten in ein Wettermodell eingelesen und durch die im Modell enthaltenen physikalisch-mathematischen Gleichungen werden für jeden Ort (auch dort, wo keine Beobachtungsdaten vorliegen) entsprechende Daten berechnet. Diese sog. Datenassimilation gewährleistet, dass die recht inhomogen verteilten Datenpunkte im resultierenden Gitter (oder auch Raster) physikalischen Zusammenhängen folgen und konsistent zueinander sind. Reanalysedaten werden durch verschiedene Organisationen bereitgestellt und sind unter <http://icdc.zmaw.de/reanalysis.html> und <http://reanalysis.org> beschrieben.

Gegitterte (oder gerasterte) bodennahe Werte sind beim Deutschen Wetterdienst u.a. in Form des HYRAS-Datensatzes mit einer räumlichen Auflösung von 5 km für den Zeitraum 1951-2006 verfügbar. Dieser Datensatz enthält aktuell (Stand Juli 2015) die 2m-Tagesmitteltemperatur (Temperatur in 2 m Höhe über Grund), die relative Feuchte (Frick et al., 2014) und die Tagessumme des Niederschlags (Rauthe et al., 2013). Der Datensatz eignet sich besonders gut für die Validierung von Modellen. Es gibt aber auch andere Beobachtungsdatensätze wie den E-OBS Datensatz für Europa (<http://eca.knmi.nl/download/ensembles/ensembles.php>) mit einer räumlichen Auflösung von 0.25 ° für den Zeitraum 1950 – 2013 und verschiedene globale Datensätze von CRU (Climatic Research Unit; <http://www.cru.uea.ac.uk/cru/data/>) mit einer räumlichen Auflösung von 0.5 ° für Zeiträume 1901 – 2006, teilweise bis 2009, auch mit einer Auflösung von 0.1 ° für 1901 – 2000.

2.2 Szenarien

Zukunftsforscher aus den Bereichen Sozial-, Wirtschafts- und Gesellschaftswissenschaften treffen Annahmen über mögliche zukünftige Entwicklungen, die durch den Menschen ausgelöst werden und auf das weltweite Klima einen Einfluss ausüben. In Zusammenarbeit von verschiedenen Experten mit dem Weltklimarat (Intergovernmental Panel on Climate Change; IPCC) wurden Szenarien entwickelt.

Die Annahmen des 4. Sachstandsberichtes (2007) stellten Szenarien über die weltweite demographische, ökonomische, politische und technologische Entwicklung in der Zukunft dar und ermittelten daraus den Ausstoß an Treibhausgasemissionen und die resultierende Treibhausgaskonzentration in der Atmosphäre. Es handelt sich dabei um Special Reports on Emission Scenarios (SRES Emissionsszenarien, Nakicenovic et al. 2000). Die Treibhausgaskonzentration verändert den Strahlungshaushalt der Atmosphäre. Die Klimamodelle berücksichtigen die zusätzliche Absorption langwelliger Wärmestrahlung der Erdoberfläche durch die Treibhausgase sowie die dadurch erhöhte atmosphärische Gegenstrahlung (Strahlungsantrieb).

Die neuen Szenarien für den 5. Sachstandsbericht beschreiben repräsentative Pfade des Strahlungsantriebs der Treibhausgaskonzentrationen von 2005 bis 2100 (engl.: Representative Concentration Pathways; RCP), die mit gekoppelten Energie-Ökonomie-Klima-Landnutzungs-Ozean-Modellen unter Vorgabe des Strahlungsantriebs am Ende des Jahrhunderts ermittelt wurden.

(Moss et al., 2010, <http://www.klimanavigator.de/dossier/artikel/012038/index.php> und IPCC AR5 WG1, <http://www.climatechange2013.org>)

2.3 Globale und regionale Klimamodelle

Die Simulation des weltweiten Klimas in Vergangenheit und Zukunft erfolgt durch globale Klimamodelle (Zirkulationsmodelle engl. General Circulation Models, GCM oder Erdsystemmodelle engl. Earth System Models, ESM).

Ein Klimamodell ist ein Computermodell, in dem die für das Klima und dessen Veränderung wichtigen Komponenten und Prozesse im Klimasystem beschrieben sind. Es basiert auf den physikalischen Grundgleichungen der Massen-, Impuls- und Energieerhaltung. Das Modell umfasst die für das Klima wichtigen Prozesse in der Atmosphäre, im Ozean, an der Erdoberfläche, im Boden, in der Biosphäre, in den Eisschilden und Gletschern mit ihren vielfältigen Wechselwirkungen. Für die Modellierung des Klimas müssen bspw. die Sonneneinstrahlung, die chemische Zusammensetzung der Atmosphäre und die Beschaffenheit der Erdoberfläche bekannt sein.

Das Modellgitter der globalen Klimamodelle hat etwa eine horizontale Maschenweite von 100 - 200 km. Im Modul Atmosphäre wird auf 20 – 30 vertikalen Schichten bis in die Stratosphäre gerechnet. (http://www.climate-service-center.de/011606/index_0011606.html.de)
(weitere Angaben in <http://www.klimanavigator.de/dossier/dossiers/011976/index.php>)

Zur Simulation des vergangenen Klimas wird ausgehend von einem vorgegebenen Anfangszustand der Modellkomponenten Ozean, Eismassen, Vegetation und Atmosphäre mit der Treibhausgaskonzentration vor der Industrialisierung (etwa 1850) begonnen und unverändert über einige Jahrhunderte gerechnet (Kontroll-Lauf). Der Antrieb des Klimasystems sind die Sonneneinstrahlung und die festgehaltene Treibhausgaskonzentration. Für den Zeitraum ab 1850 bis 2000 oder 2005 wird mit beobachteten Treibhausgaskonzentrationen gerechnet (20C/historical-Läufe). Von der Gegenwart in die Zukunft werden Vorgaben aus den verschiedenen Szenarien genommen (siehe Abschnitt 2.2).

In einem Regionalmodell erfolgt eine differenzierte Darstellung für einen räumlichen Ausschnitt aus einem Globalmodell unter dessen Vorgaben der Klimaparameter an den Rändern bzw. zur Beschreibung des Anfangs- und Ozeanzustands in der Fläche des Regionalmodells. Durch die geringere räumliche Ausdehnung kann die Gitterweite deutlich verringert werden - bis auf horizontale Gitterabstände von 10 - 20 km. Dadurch wird eine bessere Auflösung von

Prozessen erreicht, die beispielsweise durch die Orographie, die Landnutzung oder durch Küstenlinien beeinflusst werden.

Alternativ oder zusätzlich kann ein Downscaling auch durch statistische Verfahren (Empirisch-statistisches Downscaling (engl. Empirical Statistical Downscaling ESD)) erfolgen. Diese statistischen Verfahren basieren auf Stationsbeobachtungen und geben deren räumliche Auflösung wieder. Konstruktionsbedingt weichen sie nur geringfügig von den beobachteten Datensätzen ab. Auf Grund des geringeren Rechenzeitaufwands lassen sich große Ensembles leicht realisieren. Der Nachteil des statistischen Downscaling liegt in der fehlenden physikalischen Kopplung an das antreibende Modell (globales oder regionales Klimamodell; GCM oder RCM). Neue Zustände der Atmosphäre sind daher nur bedingt abbildbar.

2.4 Wirkmodelle in einer Modellkette

Die Auswirkungen des Klimawandels auf nicht explizit durch Klimamodelle simulierte Größen können nur durch nachgeschaltete spezielle Wirkmodelle (auch Impaktmodelle genannt) ermittelt werden. In diese Modelle gehen als Eingangsdaten die Ausgabedaten (Ergebnisse) der Klimamodelle ein. Die Modellkette stellt sich dann folgendermaßen dar.

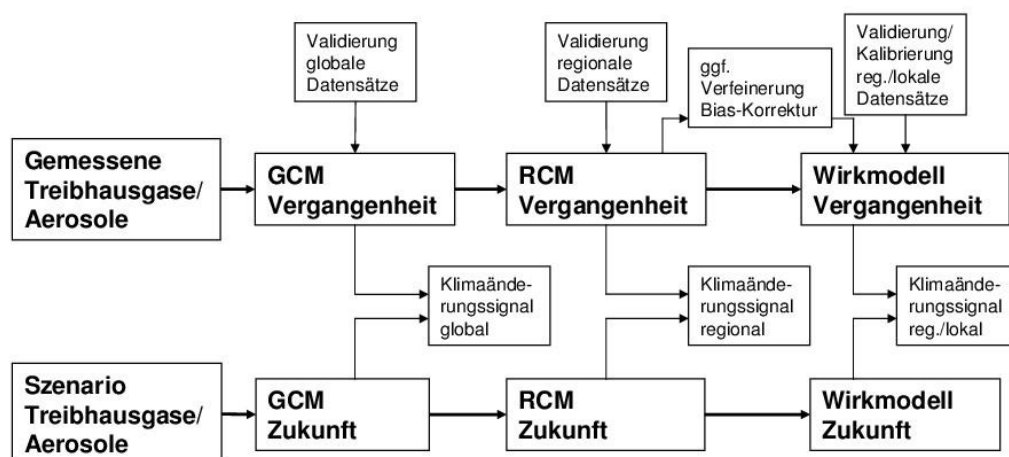


Abbildung 1: Modellkette mit Wirkmodell

Dazu müssen unter Umständen räumliche und zeitliche Verfeinerungen der Daten vorgenommen werden (Regionalisierung). Ferner kann eine Bias-Minimierung einer Größe notwendig sein (siehe Abschnitt 4.3). Für diese Korrekturen gibt es zurzeit keine Standardverfahren. Es ist darauf zu achten, dass Extremwerte durch Verfeinerungs- und Interpolationsverfahren (Regionalisierung) oder durch eine Bias-Minimierung so wenig wie möglich verfälscht werden, dass bei Interpolation von Stationsdaten in die Fläche Luv- und Leeeffekte berücksichtigt werden müssen und dass die Konsistenz physikalischer Zusammenhänge nicht durch die Minimierung vernachlässigt wird.

Zahlreiche, vor allem in Deutschland publizierte Ergebnisse basieren auf dem GCM ECHAM. ECHAM (Akronym aus ECMWF und Hamburg) ist ein globales atmosphärisches Zirkulationsmodell, das am Hamburger Max-Planck-Institut (MPI) für Meteorologie entwickelt wurde.

2.5 Grenzen der Modellierbarkeit

Jedes Klimamodell stellt nur ein begrenztes Abbild der Realität dar, da nicht alle räumlichen und zeitlichen Skalen und auch nicht alle Prozesse des Klimasystems simuliert werden können. In einer Modellkette zur Simulation zukünftiger Klimaänderungen werden zusätzliche Beschränkungen hinzugefügt. Grundsätzlich kann man folgende Modellierungs-grenzen identifizieren:

SZENARIENANNAHMEN

Jedes Szenario basiert auf Annahmen über zukünftige Entwicklungen,

MODELLBESCHRÄNKUNG

Beschränkung der Modellgenauigkeit durch Auflösung, Parameterisierung, Rechen-schema, ggf. mangelndes Prozessverständnis.

KLIMAVARIABILITÄT

z.B. große Variabilität vieler Klimaparameter auf dekadischen Skalen,

ZUSÄTZLICH IN DER MODELLKETTE

Unterschiede in den Regionalisierungsverfahren und in den Wirkmodelle.

Die meisten dieser Beschränkungen sind systeminhärent und nicht zu vermeiden (insbes. Szenarienannahmen und Klimavariabilität). Sie bewirken eine Bandbreite der Ergebnisse verschiedener Klimamodelle. Eine bewährte Strategie, mit dieser Bandbreite umzugehen, ist die Nutzung möglichst aller verfügbaren Modelle und die Darstellung der Simulationen als statistische Größen wie Mittelwert und Varianz oder Minimum und Maximum oder Median und anderer Quantile wie 15 % und 85 % (s.a. Klimaatlas des Deutschen Wetterdienstes, www.dwd.de)

Inhaltlich unterscheidet sich die Szenarienunsicherheit von den anderen Beschränkungen, da sie in einer Klimasimulation durch die Wahl des verwendeten Szenarios festgelegt ist. Daher wird empfohlen, die Auswertung der Bandbreiten der Simulationsergebnisse getrennt für die verschiedenen Szenarien vorzunehmen.

Aufgrund der verschiedenen Ursachen der genannten Modellierungsgrenzen existieren jeweils leicht unterschiedliche Ansätze zu deren Quantifizierung. Die von der hohen Komplexität und Nichtlinearität des Klimasystems herrührende (aleatorische) Unsicherheit durch die natürliche Klimavariabilität wird durch Variation der Anfangsbedingungen erfasst. Dazu werden mehrere Simulationen mit verschiedenen Startterminen, jedoch mit demselben Klimamodell und Emissionsszenario durchgeführt (Anfangsbedingungen-Ensemble). Die Abweichungen in den Ergebnissen resultieren dann allein aus den Unterschieden im anfänglichen Systemzustand, d.h. sie sind mögliche und plausible Realisierungen des Klimas im Rahmen der natürlichen Klimavariabilität.

Neben der dem System selbst innewohnenden und damit auch theoretisch nicht gänzlich reduzierbaren aleatorischen Komponente kommen weitere Unsicherheiten durch die unvollständige Abbildung des Klimas in den Modellen hinzu (Modellbeschränkung). Dieser Teil wird auch epistemische Unsicherheit genannt und ergibt sich aus der notwendigen Vereinfachung bestimmter physikalischer Vorgänge sowie aus der Vernachlässigung einzelner, bislang nicht oder nicht vollständig verstandener Prozesse. Diese Form der Unsicherheit lässt sich durch die Kombination mehrerer unterschiedlicher Globaler und Regionaler Klimamodelle (so genanntes Multi-Modell-Ensemble) abschätzen (siehe Abschnitt

4.2). Eine weitere Möglichkeit besteht in der Verwendung einer jeweils leicht unterschiedlichen Beschreibung der physikalischen Prozesse durch ein und dasselbe Klimamodell (Modellphysik-Ensemble (engl.: perturbed physics ensemble)). Diese Vorgehensweise erlaubt die Abschätzung des Fehlers, der durch die bewusst eingeführten Vereinfachungen im Klimamodell entsteht. Zur Auswertung mit dem Ziel der Entwicklung von Anpassungsmaßnahmen ist die Betrachtung des Gesamtensembles notwendig.

3. Regionale Klimaprojektionen – eine Übersicht

3.1 Verfügbare Simulationen

In Deutschland wurden in der Vergangenheit vorwiegend die Projektionen vier verschiedener regionaler Klimamodelle zur Abschätzung der zukünftigen Klimaentwicklung genutzt. Die Modelle CCLM (COSMO Climate Local Model) und REMO (Regional-Modell) gehören dabei zu der Gruppe der dynamischen regionalen Klimamodelle (RCM), d.h., ebenso wie die globalen Klimamodelle simulieren sie das Klima auf Basis physikalischer Gleichungen (vgl. Abschnitt 2.3). Bei den Modellen WETTREG (wetterlagenbasiertes Regionalisierungsverfahren) und STARS (statistisches Regionalmodell) handelt es sich dagegen um statistische Ansätze (ESD). Ihre Projektionen fußen auf der Neuordnung beobachteter Klimadaten unter Berücksichtigung der Vorgaben der globalen Klimamodelle. Weiterhin sind durch die Projekte ENSEMBLES (van der Linden 2009) und PRUDENCE (Christensen 2007) weitere Projektionsdatensätze verfügbar.

Für Europa steht seit 2014 innerhalb der CORDEX (Coordinated Downscaling Experiment) Initiative durch den europäischen Zweig EURO-CORDEX ein Ensemble von Projektionen mehrerer Regionalmodelle, angetrieben durch mehrere Globalmodelle zur Verfügung. Die Simulationen basieren auf den RCP-Szenarien (Abschnitt 2.2). Die Region umfasst Europa und Nordafrika mit einer horizontalen Auflösung von 0.11° (ca. 12 km). Die nahe Vergangenheit wird durch die Periode 1951 – 2005 beschrieben. Gerechnet wurden 14 20C-Läufe mit unterschiedlichem Anfangszeitraum (1950/1969 – 2005) und 29 Szenarienrechnungen (2006 bis 2100), durchgeführt mit neun verschiedenen Regionalmodellen auf der Basis von sechs Globalmodellen für die Szenarien RCP2.6, RCP4.5 und RCP8.5. Weitere Informationen sind unter <http://www.euro-cordex.net/EURO-CORDEX-Simulations.1868.0.html> zu finden.

3.2 Zeitliche und räumliche Betrachtung

Klimaprojektionen sind grundsätzlich für einen längeren Zeitraum auszuwerten. Der Grund ist die große Klimavariabilität der Klimaparameter auf dekadischer Skala. Die Betrachtungszeiträume für mögliche Klimaänderungen sollten im Regelfall mindestens 30 Jahre umfassen, also die Länge der von der Weltorganisation für Meteorologie (WMO) definierten Klimanormalperiode (Trewin, 2007). Je nach Fragestellung kann abweichend auch eine Auswertung von Dekaden sinnvoll sein. Es sei hier aber ausdrücklich darauf hingewiesen, dass eine Nutzung der Ergebnisse für die kommenden Dekaden nicht sinnvoll ist, obwohl sie vielfach nachgefragt wird. Die Dekade 2020 bis 2030 wird auch geprägt durch die vorangegangene Dekade, und die kann in einer Klimaprojektion nicht zeitgenau wiedergegeben werden (siehe Projekt MIKLIP – Mittelfristige Klimaprognosen; <http://fona-miklip.de>). In manchen Fällen kann auch die Verwendung eines längeren, über 30 Jahre hinausgehenden Zeitabschnitts sinnvoll oder sogar notwendig sein. Abhängig ist dies von der allgemeinen Variabilität des betrachteten Parameters sowie vom Typ des gewählten

Untersuchungsmerkmals. So erfordert z.B. die Analyse von Extremereignissen einen deutlich längeren Zeitraum als die Betrachtung von Mittelwerten einzelner Parameter.

Neben der zeitlichen ist bei der Auswertung von Klimaprojektionsdaten im Allgemeinen auch eine räumliche Aggregation vorzunehmen. Grund hierfür ist die Tatsache, dass die effektive Auflösung dynamischer Simulationsmodelle, d.h. die Größenordnung vollständig vom Modell erfasster Prozesse, ein Vielfaches des eigentlichen Gitterpunktabstandes beträgt. Dementsprechend sollten basierend auf den Empfehlungen der Modellentwickler besser mehrere Gitterpunkte zusammengefasst werden. Als kleinste Einheit wird von vielen Klimamodellierern ein Raum von 3 x 3 Gitterboxen empfohlen (bei Gitterboxen von 10 km Kantenlänge entspricht das ca. 900 km²). Ähnliches gilt auch für stationsbasierte Auswertungen. Um eine Übertragbarkeit der Daten auf eine Region zu ermöglichen, ist eine Auswertung über mehrere, in ihren klimatischen Eigenschaften vergleichbare Stationen anzustreben. In orographisch stark gegliedertem Gelände oder an Küsten muss darauf geachtet werden, dass die Gitterboxen / Stationen ähnliche Eigenschaften aufweisen. Verwendet man Mittelwerte über Gitterboxen, die Land- und Wasserflächen oder Flachland- und Gebirgsflächen repräsentieren, so ist ein solcher Mittelwert problematisch. In solchen Fällen sollte man eher Mittelwerte über gleichartige, aber u.U. unregelmäßig verteilte Gitterboxen bilden. Ist die räumliche Mittelung der originalen Klimaprojektionszeitreihen mit zu großen Nachteilen verbunden, z.B. beim Einsatz von komplexen Multiparameter-Wirkmodellen (siehe Abschnitt 2.4), so kann die Aggregation auch nachträglich vorgenommen werden.

3.3 Bezugsgrößen der Klimaänderungssignale

Speziell die Gegenwartssimulationen dynamischer Klimamodelle weisen häufig einen systematischen Fehler, auch als Bias bezeichnet, auf. Die direkte Verwendung der von diesen Modellen für die Zukunft berechneten Datenzeitreihen ist daher ungünstig. Das bedeutet, dass das Ergebnis verfälscht wird. Die einfachste Möglichkeit, dieses Problem zu umgehen, ist die Angabe des Klimaänderungssignals. Dabei wird das zukünftige Klima durch Bildung einer Differenz oder eines Verhältnisses auf das gegenwärtige Klima bezogen. Es wird z.B. der Anstieg der Temperatur oder die relative Niederschlagsänderung im Vergleich zur Referenzperiode (siehe Abschnitt 4.1) berechnet. Entsprechend Abschnitt 3.2 sollten Projektions- und Bezugsperiode dabei einen längeren und möglichst auch gleichlangen Zeitraum umfassen.

Dieser Vorgehensweise liegt die Annahme zugrunde, dass Bias aus Projektionszeitraum und Bezugsperiode weitestgehend übereinstimmen und sich daher aufheben. Allerdings ist nicht bekannt, ob diese Voraussetzung erfüllt ist.

3.4 Prüfung von Klimamodelldaten

Zur Abschätzung der Belastbarkeit von Klima- und Klimafolgenstudien auf der Basis von Projektionsdaten ist es sinnvoll, die problemspezifische Qualität der verwendeten Simulationen zu prüfen.

Die Prüfung von Klimamodelldaten kann grundsätzlich auf zweierlei Art und Weise erfolgen: zum einen anhand des Vergleichs der Gegenwartsläufe mit Beobachtungsdaten (innerhalb eines Referenzzeitraumes), im Folgenden Evaluierung genannt, zum anderen durch Überprüfung der zukünftigen Änderungen einer Größe (Signale) bzw. der absoluten Modellwerte in der Zukunft auf Plausibilität.

Bei beiden genannten Verfahren kann eine zeitlich differenzierte Betrachtung der Daten stattfinden, z.B. für gemittelte Gebietszeitreihen oder einzelne Stationszeitreihen (also ohne räumliche Differenzierung). Hierbei ist zu beachten, dass Klimamodelle Entwicklungstendenzen von Klimagrößen nur über längere Zeiträume sinnvoll wiedergeben, nicht aber in Form einzelner Monats- oder Jahreswerte. Dementsprechend sollten bei einer Evaluierung bzw. Plausibilitätskontrolle auch nur über einen längeren Zeitraum (z.B. 30 Jahre) gemittelte Werte verglichen werden bzw. Werte ohne chronologischen Bezug. Einzeljahre können beispielsweise nicht verglichen werden.

Andererseits kann auch eine räumlich differenzierte Betrachtung von Klimamodelldaten erfolgen. In diesem Zusammenhang sollte der Betrachtungsraum nicht zu klein gewählt werden, da die Prozesse innerhalb der Klimamodelle nur über größere Räume (mehrere Gitterzellen, s. Abschnitt 3.2) gemittelt berechnet bzw. interpretiert werden sollten. Durch die Zusammenfassung vieler Punktdaten (Stationsdaten) innerhalb eines solchen Betrachtungsraums verringert sich der Fehler im Vergleich zu den Flächendaten (Gitterboxen) aus den dynamischen Modellen.

Bei einem Vergleich von Modellergebnissen der nahen Vergangenheit (20C) und real beobachteten Referenzdaten geht man davon aus, dass die Aussagekraft der Klimamodelldaten umso belastbarer ist, je genauer die Übereinstimmung ausfällt. Die Abweichungen zwischen Beobachtungs- und Modelldaten können z. B. über ein Gütekriterium, etwa den mittleren quadratischen Fehler oder die prozentuale bzw. absolute Differenz bewertet und objektiv verglichen werden (s.a. Hennemuth et al., 2013, Kap. 8.5.1). In diesem Zusammenhang kann die Modellgüte je nach Art der Beobachtungsdaten (z.B. Messdaten, Reanalysedaten) unterschiedlich ausfallen. Das Klimaänderungssignal kann dagegen nur für sich fachlich beurteilt werden, z. B. ob sich die Klimagröße zukünftig in physikalisch „realistischen“ Größenordnungen bewegt oder ob sich ggf. zwei oder mehr Größen „logisch“ (physikalisch sinnvoll) zueinander verhalten. Weiterführende ausführliche und anschauliche Erklärungen sind in dem Beispiel „Plausibilität“ im Anhang zu finden.

Allerdings muss darauf hingewiesen werden, dass es problematisch sein kann, einzelne Modellergebnisse auf Grund einer hohen Abweichung von 20C-Ergebnissen zu den beobachteten Daten für weitere Betrachtungen auszuschließen. Auch das begründete Aussortieren einer oder mehrerer Klimaprojektionen aus dem zur Verfügung stehenden Gesamtensemble allein auf Basis einer Evaluierung muss gründlich abgewogen werden. Es besteht immer die Gefahr, wertvolle Informationen über das zukünftige Klima zu vernachlässigen. Zudem ist keineswegs sichergestellt, dass sich bei der Untersuchung verschiedener Variablen oder mehrerer Parametereigenschaften stets dieselben Projektionsläufe als besonders unzuverlässig herausstellen. Das Ensemble von Ergebnissen müsste also sinnvollerweise jeweils um andere Projektionen reduziert werden, was den späteren Vergleich erschwert (siehe Abschnitt 4.2).

4. Auswertung und Darstellung

4.1 Rahmenbedingungen

Für die Entwicklung von Anpassungsstrategien und Anpassungsmaßnahmen zu den verschiedenen Handlungsfeldern der Deutschen Anpassungsstrategie (DAS) sowie der Anpassungsstrategien der Bundesländer sind plausible Angaben zur Bandbreite möglicher Klimaentwicklungen notwendig. Insbesondere werden auch abgeleitete Klimakenngrößen, wie z.B. Heiztage, Kühlgradtage, Vegetationsbeginn, heiße Tage etc. und deren Variabilität benötigt, um strategische Entscheidungen zur Gestaltung von Anpassungsmaßnahmen in den nächsten Dekaden zu treffen. Voraussetzung dafür ist ein fachlich korrekter Umgang mit den Projektionen regionaler Klimamodelle.

Klimaprojektionen sind grundsätzlich für einen längeren Zeitraum auszuwerten. (vgl. Abschnitt 3.2). Bei der Auswertung (und Darstellung) der zukünftigen Entwicklung von Klimagrößen sind ausgewählte, mindestens 30jährige Perioden der Klimaszenarien immer mit den jeweiligen Referenzperioden der Modelle (20C-Zeitraum) zu vergleichen und nicht mit Beobachtungsdaten. Als Referenzzeitraum wird von der WMO dauerhaft die Periode 1961 – 1990 für die Bewertung des langfristigen Klimawandels empfohlen. Darüber hinaus gibt es parallel dazu die 30-jährigen „klimatologischen Standardnormalwerte“, für die als aktuelle Periode 1981-2010 verwendet wird. Diese 30-jährigen klimatologischen Standardnormalwerte werden alle 10 Jahre fortgeschrieben, d.h., am Anfang der nächsten Dekade wird dann auf den Zeitraum 1991-2020 umgestellt. In manchen Fällen lassen sich die Empfehlungen der WMO z.B. auf Grund der Datenverfügbarkeit oder zum Zwecke der besseren Vergleichbarkeit mit anderen Untersuchungen jedoch nicht eingehalten. Eine derartige Vorgehensweise sollte aber stets gut begründet werden.

Als Vergleichszeiträume für die Zukunft sollten die Zeitspannen 2021 - 2050 und 2071 - 2100 abgebildet werden, um auch hier die Vergleichbarkeit zu anderen Studien zu gewährleisten. Bei saisonalen Betrachtungen werden die meteorologischen Jahreszeiten Winter (Dezember, Januar, Februar - DJF), Frühjahr (März, April, Mai - MAM), Sommer (Juni, Juli, August - JJA) und Herbst (September, Oktober, November - SON) empfohlen, um die Variabilität der Klimagrößen im Jahresverlauf zu erfassen.

4.2 Ensemblebetrachtung

Zur quantitativen Abschätzung der bereits in Abschnitt 2.5 angesprochenen Unsicherheiten von Klimaprojektionen werden Ensembles verwendet, d.h. mehrere Einzelsimulationen werden kollektiv ausgewertet. Hierdurch ergibt sich eine Ergebnisspanne, mit der gearbeitet werden muss.

Das Klima ist im mathematischen Sinne ein „deterministisch chaotisches System“. Das heißt, dass das System physikalischen Gesetzmäßigkeiten unterliegt, aber dennoch extrem sensitiv auf veränderte Start- und Randbedingungen reagiert. Rechnet man zwei Klimasimulationen mit nur minimal anderen Anfangsbedingungen (z.B. Änderungen in der Nachkommastelle eines meteorologischen Parameters), so können die Modellergebnisse langfristig gesehen, innerhalb der klimatischen Variabilität unterschiedlich ausfallen. Beide Ergebnisse sind jedoch physikalisch als gleichwertig „richtig“ und möglich zu betrachten. Deshalb ist für die Auswertung und Interpretation von Klimamodelldaten der statistische Umgang mit Ensembles unabdingbar erforderlich. Die Ensemblebetrachtung sollte als Kombination mehrerer Modellketten (global und regional) mit jeweils verschiedenen Läufen (veränderten Anfangsbedingungen) durchgeführt werden. Dieser Ansatz wird Multimodellansatz genannt.

Für Simulationen der Vergangenheit hat sich gezeigt, dass der Ensemblemittelwert den real eingetroffenen Zustand am besten abbildet. Daher wird auch für die Zukunft der Ensemblemittelwert oftmals als der am glaubhaftesten eintreffende Zustand angesehen. Er sollte aber nur dann verwendet werden, wenn alle Ensemblemitglieder eine Normalverteilung bilden. Anderenfalls ist der Median die bessere Wahl.

Für die Wiedergabe der Ergebnisse ist grundsätzlich die gesamte Bandbreite des Ensembles aufzuzeigen. Dies kann durch Angabe der Spannweite oder aber mittels Darstellung von Minimum und Maximum erfolgen. Unter bestimmten Voraussetzungen lassen sich durch die Ermittlung von Perzentilen auch robuste Aussagen zu konkreten Änderungskorridoren ableiten (bspw. das 15. und 85. Perzentil wie im Deutschen Klimaatlas des DWD).

Es ist jedoch wichtig hier anzumerken, dass jedes genutzte Ensemble nur ein Gelegenheitsensemble (engl.: 'ensemble of opportunity'; vgl. Parker, 2011; Meehl et al., 2007; Carter, 2010 und Knutti et al., 2010) darstellt. Das bedeutet, dass die Ensemble-Mitglieder nicht notwendiger Weise den vollständigen Umfang der Bandbreite darstellen müssen. Auch sind die Modelle nicht alle voneinander unabhängig. Einige Modellkomponenten werden z.B. von mehreren Modellgruppen genutzt. (Knutti et al., 2013; Pennel und Reichler, 2011 und Yuan und Wood, 2012). Dementsprechend kann es bei der Bestimmung von Bandbreiten aus Klimaprojektionsensembles sinnvoll sein, eine Gewichtung der einzelnen Simulationen vorzunehmen. Darüber hinaus gilt es abzuwägen, ob als unrealistisch einzuordnende Projektionen aus dem Ensemble ausgeschlossen werden sollen oder nicht. Bei dieser Entscheidung konkurrieren eine möglicherweise zu große Auffächerung der Ergebnisse und die zu starke Reduktion der Bandbreite miteinander (Plagemann et al., 2014).

4.3 Bias-Minimierung

Viele Wirkmodelle eignen sich nicht für den Einsatz mit Klimaänderungssignalen, da sie für den Antrieb mit meteorologischen Beobachtungswerten entwickelt wurden und sich somit ausschließlich mit Absolutwerten betreiben lassen. Sollen diese Wirkmodelle mit den Ergebnissen von Klimaprojektionen betrieben werden, so müssen statt Änderungssignalen die Klimaprojektionsdaten selbst als Input verwendet werden. Hierfür müssen aber zunächst etwaige modellinterne Fehler der Klimamodelle quantifiziert und ggf. so weit wie möglich korrigiert werden. Dafür stehen, je nach Klimaparameter, verschiedene Verfahren zur Bias-Minimierung (auch Bias-Korrektur genannt), also zur Verringerung des systematischen Fehlers, zur Verfügung. Diese basieren auf einem Vergleich zwischen beobachteten Daten und den 20C-Simulationen der Klimamodelle innerhalb der Referenzperiode. Hierfür müssen die Beobachtungswerte auf demselben Gitter vorliegen wie die Klimaprojektionsdaten, d.h., im Allgemeinen ist eine Übertragung der in Form von Stationsdaten vorliegenden Beobachtungen in die Fläche notwendig.

Die einfachste Art der Bias-Minimierung bei einer annähernd linearen Beziehung zwischen Modell- und Beobachtungsdaten ist eine Bearbeitung durch einen konstanten Versatz oder eine konstante prozentuale Minimierung. Das trifft meist auf Jahresmittelwerte zu. Weitere Bias-Minimierungsverfahren ziehen die fehlerbehafteten Häufigkeitsverteilungen der Projektionsdaten über die Bestimmung mittlerer Abbildungsvorschriften an die realen Beobachtungen heran (s. Hennemuth et al., 2013).

Diese Art der Postprozessierung weist jedoch auch gewisse Nachteile auf. Insbesondere bei dem mit der individuellen Bearbeitung der verschiedenen Parameter verbundenen Verlust der meteorologischen Konsistenz (Ehret et al. 2012) handelt es sich beim Einsatz komplexer Multiparameter-Wirkmodelle um einen störenden Nebeneffekt. Zudem lassen sich die

Ausprägungen einiger typischer Modelldefizite auch mit Bias-Minimierungsverfahren nicht oder nur teilweise beheben.

Darüber hinaus kann die Anwendung der Bias-Minimierung auch zu geänderten Klimasignalen führen. Dieser Effekt tritt insbesondere dann auf, wenn zum Beispiel beim Quantilemapping die einzelnen Bereiche mit deutlich unterschiedlichen Faktoren adjustiert werden müssen. Ebenso können die vorhandenen Beobachtungsfehler sowie die durch die Rasterung der Stationsdaten entstehenden Fehler durchaus einen nachteiligen Effekt auf die angepassten Zeitreihen ausüben.

Die Anwendung von Bias-Minimierungsverfahren erfordert daher immer eine umfangreiche Kontrolle, sowohl der Qualität der resultierenden Datenzeitreihen als auch der Sensitivität des Wirkmodells gegenüber Inkonsistenzen zwischen den Eingangsvariablen.

4.4 Statistische Auswertung und Extremwerte

Klimamodelldaten und abgeleitete Impact- oder Wirkmodelldaten für die Gegenwart und für Zukunftsprojektionen sind sehr komplex. Zur Bearbeitung der Daten mit dem Ziel der Qualitätsanalyse und Interpretation der Ergebnisse müssen daher statistische Methoden und Auswerteverfahren eingesetzt werden. Diese Verfahren können für verschiedene Fragestellungen und Sektoren sehr unterschiedlich sein. Daher können keine für alle Anwendungen gleichermaßen gültigen Verfahren angegeben werden.

Einige häufig angewendete Methoden sollen hier genannt und kurz erläutert werden.

Einen guten Überblick über die Struktur der simulierten oder berechneten Daten geben Häufigkeitsverteilungen. Bei geeigneter Auflösung (siehe z.B. Abbildung 7 und Anlage 6: Beispiel Tourismus (Schwimmbadbesucher)) lässt eine Häufigkeitsverteilung Aussagen über die am häufigsten vorkommenden Werte, die seltensten (extremen) Werte und auch Verschiebungen von einer Zeitscheibe zur anderen zu. Falls es zur weiteren Analyse nötig ist, kann auch die Art der Verteilung und deren charakterisierende Parameter ermittelt werden (bspw. eine Gauß-Verteilung mit den Parametern Mittelwert und Standardabweichung).

Weitergehende und gezielte Aussagen werden durch Zeitreihenanalysen möglich, die die zeitliche Abfolge der Werte berücksichtigen. Beispielsweise kann eine Analyse von zusammenhängenden Ereignissen wie Trockentagen, Starkregentagen u.a. durchgeführt werden. Die Zeitreihenanalyse wird auch zur Ermittlung von Trends in den Datensätzen genutzt. Dabei gibt es einfache Verfahren, die einen linearen Trend annehmen und berechnen (z.B. lineare Regressionsanalyse), aber auch Verfahren für instationäre und nichtlineare Trends.

Ein durch eine Zeitreihenanalyse ermittelter Trend sollte durch einen Signifikanztest auf statistische Signifikanz überprüft werden. Auch hier ist die Bandbreite der möglichen Verfahren groß und eine Entscheidung zwischen parametrischen und nicht-parametrischen Tests ist notwendig. Es sei bereits hier gesagt, dass auch ein statistisch signifikanter Trend noch keine Aussage über die Eintrittswahrscheinlichkeit enthält (s.u.).

In vielen Fragestellungen ist eine zeitliche und/oder räumliche Verfeinerung der Klimadaten gewünscht (etwa bei Stadtplanungs- oder Stadtentwässerungsfragen). Hier muss ein Downscaling- oder Interpolationsverfahren angewendet werden. Diese Verfahren sind mit besonderer Vorsicht einzusetzen, da die Klimamodelle die Prozesse, bis zu denen aufgelöst werden soll, nicht simulieren können. So kommen etwa die Strömungen zwischen kühlen Parkflächen und überhitzten Straßenkomplexen in Klimamodellen mit einer horizontalen Auflösung von 25 km oder 10 km nicht vor, sind aber gerade für Stadtplaner wichtig. Generell ist es besser, größere Gebiete, die durch Klimamodelle aufgelöst werden, zu betrachten. Ist

jedoch eine Verfeinerung notwendig, muss sichergestellt sein, dass alle Einflussgrößen in dem Interpolationsschema berücksichtigt sind. Eine einfache räumliche Interpolation kann u.U. nur über relativ homogenen Oberflächen – bspw. der Meeresoberfläche – eingesetzt werden.

Ebenfalls häufig gewünscht, da von hoher gesellschaftlicher und ökonomischer Relevanz, ist eine Aussage über die zeitliche Entwicklung von Extremwerten, was eine Extremwertstatistik voraussetzt. Diese Verfahren sind besonders komplex und fordern eine gute Kenntnis der zu analysierenden Daten. Im Allgemeinen ist die Verteilung der Extremwerte nicht aus der allgemeinen Häufigkeitsverteilung abzuleiten und muss gesondert ermittelt werden. Die Analyse umfasst folgende Schritte: Auswahlverfahren der Daten, Parameterschätzung der Verteilung, Analyse durch empirische Methoden oder Extremwertanalyseverfahren.

Ein weiteres statistisches Verfahren, das oft angewendet wird, ist die Ermittlung von Kennzahlen zur Charakterisierung der Daten. Hierbei kann es sich um Kennzahlen zur Angabe der Modellgüte (Modellevaluierungskennzahlen, siehe Abschnitt 3.4) oder um die Bestimmung komplexer klimatischer Indizes oder Indikatoren handeln.

Besonders wichtig für alle Analysen von Klimamodell- und Impaktmodelldaten ist die Verwendung eines Ensembles von Modellen, um die Aussagekraft der Ergebnisse besser bewerten zu können. Hierfür müssen Methoden der Ensembleanalyse (vgl. Abschnitt 4.2) eingesetzt werden. Das InterAcademy Council (InterAcademy Council, 2010) benennt Verfahren zur Bewertung von Modellergebnissen, die auf Ensembles beruhen, wie etwa ein Maß für das Verständnisniveau (level-of-understanding scale), ein Konfidenzmaß und ein Wahrscheinlichkeitsmaß. Nur durch die Beschreibung der Ergebnisse auf der Basis solcher Maße kann die Bandbreite der Ergebnisse abgebildet werden und bspw. eine Aussage über die Robustheit der Ergebnisse ermöglicht werden.

Bei der Anwendung aller statistischen Verfahren ist es notwendig, die Eignung für die jeweiligen Datensätze und Fragestellungen zu prüfen. Das betrifft nicht nur die Anzahl und die Art der analysierten Parameter (unbeschränkte oder beschränkte Wertebereiche), sondern auch die Voraussetzungen, die durch die Verfahren vorgegeben sind (bspw. Annahme einer Gaußverteilung). Die Nutzer können also häufig nicht auf Standardverfahren zurückgreifen.

Viele gängige Verfahren sind in der statistischen Literatur beschrieben, etwa in Schönwiese (2013), Mudelsee (2010), von Storch H, Zwiers FW (1999). Für Nutzer von Klimadaten ist es jedoch eher hilfreich, Verfahren zu verwenden, die schon erfolgreich in ähnlichen Fragestellungen eingesetzt wurde. Daher wurden vom Climate Service Center bei vielen Nutzern von Klimamodell- und Anpassungsprojekten Beschreibungen ihrer ausgewählten statistischen Verfahren zusammen getragen und bereitgestellt (Hennemuth et al., 2013). Die Verfahren sind nach übergeordneten Kategorien (bspw. ‚Bias-Minimierung‘, ‚Zeitreihenanalyse‘) sortiert und in einer standardisierten Tabelle vorgestellt. So gehören zur Verfahrensbeschreibung auch Angaben über die Voraussetzungen für die Anwendung und zur Interpretation der Ergebnisse. Großer Wert bei der Zusammenstellung der statistischen Verfahren wurde auf die Bewertung des jeweiligen Verfahrens und auf eine Beispielanwendung gelegt.

4.5 Robustheit der Klimaänderungssignale

Wegen der hohen Klimavariabilität, welche sich auch in den Klimamodellsimulationen widerspiegelt, ist die Erkennung von schwachen Klimaänderungssignalen erschwert. Ein Signal sollte entsprechend erst dann als Klimaänderungssignal interpretiert werden, wenn die Signalamplitude das Klimarauschen übersteigt. In der internationalen Literatur wird als Bemessungsschwelle oft ein prozentualer Anteil der Standardabweichung auf der Basis eines 30 jährigen Referenzzeitraums genutzt. Erst wenn ein Änderungssignal (auch wieder als Mittel

über einen Zeitraum von 30 Jahren) oberhalb der Bemessungsschwelle liegt kann die Änderung als ein Klimaänderungssignal betrachtet werden. Darüber hinaus ist auch die Anwendung komplexer Signifikanztests möglich. Ein Beispiel für einen solchen Test ist in Anlage 8 dieses Textes beschrieben.

Bei der Ensemble-Modellierung geht es nicht nur um die Erkennbarkeit (Signifikanz) eines Klimasignals, sondern auch um die Frage, ob alle Ensemble-Mitglieder dieses Signal hinsichtlich Richtung und Ausprägung gleichermaßen aufweisen und das Signal auch weitgehend unabhängig von Unsicherheiten ist. In diesem Zusammenhang spricht man auch vom Model Agreement (IPCC, 2013). Dieses kann über verschiedene Kriterien geprüft werden, wie etwa den Prozentsatz der Übereinstimmung bzgl. der Richtung des Änderungssignals aus allen Modellen, die Signifikanz der Veränderung und die Sensitivität der Ergebnisse gegenüber einer zeitlichen Verschiebung der Vergleichsperioden. Beispiele für eine Robustheitsprüfung finden sich in der Anlage 9 und in den sogenannten Klimasignalkarten (http://www.climate-service-center.de/031443/index_0031443.html.de). Hier werden die Übereinstimmung der Modelle in der Richtung der Klimaänderung (Zunahme oder Abnahme), die Signifikanz der Klimaänderung und eine Unabhängigkeit von geringen zeitlichen Verschiebungen der 30-Jahres Zeitscheiben geprüft und bewertet. Eine Robustheitsprüfung bietet eine verlässliche Basis für Aussagen zu Klimaänderungssignalen. Umgekehrt können im Fall von nicht robusten Signalen keine Aussagen gemacht werden.

4.6 Ergebnisdarstellung

Die Darstellung von Analyseergebnissen ist ein sensibler Arbeitsschritt bei der Auswertung von regionalen Klimamodelldaten. Er stellt die Schnittstelle zwischen dem Erzeuger und dem Nutzer von Klimamodelldaten dar. Daher sind einheitliche und klar verständliche Strukturen für die Darstellung wichtig, um die Ergebnisse richtig zu interpretieren und mit anderen Untersuchungen vergleichen zu können. Nachfolgende grundsätzliche Aspekte sollten bei der Auswertung berücksichtigt und bei der visuellen Darstellung der Ergebnisse benannt werden. Daneben sind die bereits zuvor genannten zeitlichen und räumlichen „Rahmenbedingungen“ zu beachten (siehe Abschnitt 4.1)

In allen Darstellungen sind die Modellkette, auf der das dargestellte Ergebnis beruht, sowie die Vergleichszeiträume fest zu verankern. In jeder Darstellung (Flächendarstellungen oder Diagramme) muss die Angabe des oder der Global-, Regional- und ggf. des Wirkmodells erkennbar sein, da sonst die Ergebnisse zu einer unzulässigen Verallgemeinerung und Missverständnissen führen oder diesen zumindest Vorschub leisten können.

Bei der Interpolation von punktuell vorliegenden Werten (z.B. zur Erstellung eines Rasters bzw. einer flächenhaften Darstellung in einer Karte) muss dafür Sorge getragen werden, dass die interpolierten Werte die Realität nicht verfälschen. Dies kann vor allem bei Extremwerten der Fall sein und zu Problemen bei der Darstellung und Interpretation führen. So besteht beispielsweise die Gefahr, bei der Interpolation des Niederschlages zwischen „nassen“ (> 0 mm) und „trockenen“ ($= 0$ mm) Werten bzw. Zellen ein Raster zu produzieren, das (je nach Interpolationsverfahren) am Ende immer einen Niederschlagswert größer Null aufweist und somit keine trockenen Bereiche mehr beinhaltet (die in den Punktwerten aber noch vorhanden waren). Ähnlich verhält es sich bei Maximalwerten und Verfahren, die zusätzliche Informationen berücksichtigen, wie etwa die Höhenlage eines Stationsortes oder Gitterpunktes. Liegt zwischen zwei Punktwerten z.B. ein Höhenrücken, kann dies durch Interpolation zu (Extrem)Werten in diesem Bereich führen, die in der Realität vielleicht gar nicht vorkommen.

Bei der Visualisierung der Ergebnisse (hier speziell mit Karten) sollte ein besonderes Augenmerk auf die Wahl der Farb- und Werteskala gelegt werden (vgl. auch Spekat et al.,

2007). In der Regel sollten nicht mehr als 9 - 10 Klassen in der Farbskala dargestellt werden (je größer die mit den Werten verbundene Unsicherheit ist, desto weniger Klassen sollten benutzt werden). Das Klassenintervall sollte größer sein als die modellspezifische Variabilität auf der entsprechenden Zeitskala, welche z.B. über statistische Signifikanztests ermittelt werden kann. Bei positiven und negativen Werten (z.B. bei der Darstellung von Änderungssignalen) sollte die mittlere Klasse den Nullwert, mit einem ausreichend großen Bereich größer und kleiner Null, beinhalten, um zu verdeutlichen, dass hier kein (signifikantes) Änderungssignal vorliegt. Die verbleibenden Klassen sollten sich symmetrisch um den Nullwert anordnen. Beim Vergleich von Werten eines Parameters, z.B. über verschiedene Jahreszeiten oder Zeitabschnitte, sollte immer eine einheitliche Werteskala benutzt werden, um die Vergleichbarkeit der Einzelergebnisse zu gewährleisten.

Besitzen die dargestellten Werte lediglich ein Vorzeichen, sollte die Farbskala auf einer Grundfarbe mit verschiedenen Abstufungen basieren. Bei unterschiedlichen Vorzeichen sollte zum einen auf farbpsychologische Effekte geachtet werden (z.B. Rot = warm / trocken, Blau = kalt / nass), zum anderen sollte der Nullwertbereich neutral (z.B. weiß) dargestellt werden (keinesfalls mit einer Signalfarbe). Auch wird empfohlen auf eine barrierefreie Darstellung zu achten (z.B. bei Rot-Grün-Blindheit). Ein hilfreiches Tool zur Farbgestaltung von Legenden ist beispielsweise unter www.colorbrewer2.org zu finden.

Auf Farb- und Helligkeitssprünge zwischen benachbarten Klassen sollte geachtet werden, um keine nicht vorhandenen Gradienten vorzutäuschen. Aus selbigem Grund sollte ebenso eine gleichmäßige Helligkeit der Kartendarstellung gewährleistet sein.

Bei der Ergebnisdarstellung von Modell-Ensembles (Ensemble-Mitglieder namentlich benennen) sollte stets deren Bandbreite visualisiert werden. Als mögliche obere / untere Grenze der Bandbreite können beispielsweise ausgewählte Perzentile genutzt werden (wie z.B. im Deutschen Klimaatlas des DWD). Werte außerhalb dieser definierten Bandbreite sollten ggf. (mit z. B. möglichen Gründen für deren Vorkommen) kommentiert werden (siehe v. a. auch Beispiele „Darstellung“ und „Darstellung_Ensembles“). Weitere Darstellungsbeispiele finden sich in der Statistikbroschüre des CS2.0 (s. Hennemuth et al., 2013 Kap. 4.4) unter der Rubrik „Ensembles“.

5. Interpretation und Vermittlung der Ergebnisse

Bei der Interpretation der Ergebnisse der Klimamodellrechnungen muss immer wieder darauf hingewiesen werden, dass es sich nicht um Prognosen über einen künftigen Verlauf lokaler, regionaler oder globaler Klimaverhältnisse handelt, sondern lediglich um Szenarien, die sich aus Annahmen über künftige Entwicklungen (z. B. Emissionen und Landnutzung) ergeben. Aus diesem Grund erfordert die Bewertung von Modellergebnissen einen zurückhaltenden Umgang mit den möglichen Zukunftspfaden.

Die Interpretation der Ergebnisse ist nur unter Beachtung der Grenzen der Modellierbarkeit, der Plausibilitätsprüfung der Ergebnisse und der Robustheit der Änderungssignale möglich. Es sind nur Aussagen über längere zukünftige Zeiträume möglich (siehe Abschnitt 3.2). Die von vielen Nutzern gewünschten Aussagen über die nächste oder übernächste Dekade können u.a. wegen der großen Klimavariabilität nicht geleistet werden. Bei der Interpretation der Ergebnisse geht es nicht vordringlich um kleinräumige und detaillierte Einzelaussagen, sondern um eine möglichst quantifizierte Trendaussage. Hierzu hat der IPCC ein standardisiertes Maß der Wahrscheinlichkeit der Ergebnisse eingeführt (Solomon et al., 2007). Danach können Aussagen nach dem Grad der Übereinstimmung eines Ergebnisses in einem Ensemble auf einer Skala von ‚ausgesprochen unwahrscheinlich‘ über ‚unwahrscheinlich‘, ‚wahrscheinlich‘ bis ‚ausgesprochen wahrscheinlich‘ bewertet werden.

Anhand der modellierten globalen Temperaturentwicklung bis 2100 wird deutlich, dass alle Szenarien einen weiteren Temperaturanstieg anzeigen. Dies ist ein robustes und sehr wahrscheinliches Signal (siehe Abschnitt 4.5). Die Höhe des Anstiegs ist jedoch variant. Sicher ist auch, dass die Höhe des Temperaturanstiegs umso größer ist, je mehr Treibhausgase in die Atmosphäre gelangen. Mit Hilfe dieser Herangehensweise können aus den Unsicherheiten gesicherte Aussagen selektiert werden.

Bei der Interpretation der Ergebnisse ist immer darauf zu achten, dass es sich bei Szenarienergebnissen stets um WENN-DANN-Aussagen handelt. Die Eingangsbedingungen sind daher immer zu nennen. Dazu gehören insbesondere das genutzte Emissions-/Konzentrations-Szenario sowie die verwendete Modellkette (GCM-RCM). Gegebenenfalls sind weitere Erläuterungen dazu notwendig. Dies trifft besonders bei der Interpretation der Ergebnisse auf Basis der RCP-Szenarien zu, da diese z. T. bereits konkrete Annahmen/Maßnahmen zum Schutze des Klimas enthalten.

So muss bei der Ergebnisinterpretation einer RCP2.6-Projektion immer darauf hingewiesen werden, dass das RCP2.6 aktiven Klimaschutz der Weltgemeinschaft voraussetzt, um das globale Ziel der maximalen Erwärmung um 2 Grad gegenüber der vorindustriellen Zeit einzuhalten. Es wird empfohlen, die Ensembleergebnisse nach Szenarien getrennt darzustellen, da die Voraussetzungen der Szenarien sich grundsätzlich unterscheiden.

Die Bewertung der Ergebnisse bzw. deren Qualität ist nur unter Beachtung der in den vorangegangenen Kapiteln genannten Aspekte möglich, zugleich aber auch notwendig. An dieser Stelle werden noch einmal einige wesentliche Punkte bzw. Schritte aufgezeigt:

- Die Grenzen der Modellierbarkeit müssen berücksichtigt werden (Abschnitt 2.5)
- Die zugrunde liegenden Modelldaten sind grundsätzlich geeignet für die Fragestellungen (der Wirkmodellierung) hinsichtlich:
 - räumlich-zeitlicher Auflösung (Abschnitt 3.2),
 - Plausibilität der relevanten Klimagrößen (Abschnitt 3.4)
- Die Bandbreiten der Ergebnisse können mit Wahrscheinlichkeitsaussagen hinterlegt werden (Abschnitt 4.2 und 4.4).

Bei der Interpretation und Vermittlung der Projektionsergebnisse müssen zusätzlich nachstehende Hinweise beachtet werden, die teilweise durch Beispiele illustriert sind.

Es gibt keine sichere Aussage über das Ausmaß an Klimaänderungen, daher sind Handlungsoptionen / Anpassungsmaßnahmen nur auf der Grundlage der Bandbreite möglichst aller verfügbaren Modellsimulationen zu treffen (Abschnitt 4.2, Bsp. „Darstellung Ensembles“).

Überschreitungswahrscheinlichkeiten sind aus der Bandbreite (Bsp. „Leitplanken“) ableitbar. Interpretationen erfolgen nicht auf der Basis einer Kurve sondern auf der Basis von Wahrscheinlichkeiten (Bsp. „Schwimmbadbesucher“, „Hitzestress“ und „Obstbau“).

Trendaussagen müssen mit statistischen Tests hinterlegt sein, die so zu wählen sind, dass sie der Charakteristik (z.B. Variabilität) der jeweiligen Untersuchungsgröße angepasst sind. Bestenfalls sind mehrere Trendanalyseverfahren anzuwenden und zu vergleichen.

Die Ergebnisse müssen objektiv und nach dem Stand der Wissenschaft dargestellt werden, gerade bei geringen oder uneinheitlichen Veränderungen, um eine Fehl- bzw. Überinterpretation durch Dritte zu vermeiden (Abschnitt 4.6, Bsp. „Parameterdarstellung“).

Wenn Klimaänderungssignale für einzelne Parameter nicht eindeutig sind (Änderungssignale mit unterschiedlichen Vorzeichen oder keine robusten Signale), dann können nur Handlungsempfehlungen gegeben werden, die auch unter derzeitigen Bedingungen eine Verbesserung darstellen (sog. „no-regret-Maßnahmen“).

6. Hinweise zur Wirk- bzw. Impaktmodellierung

Regionale Klimasimulationen liefern als Ergebnisse Zeitreihen von Klimavariablen, die nach entsprechender Aufbereitung als Eingangsgröße für die modellgestützte Abschätzung von Klimawirkungen (Impaktmodellierung) dienen können. Informationen und Daten über künftige Klimaänderungen sind daher eine notwendige, aber nicht hinreichende Basis für die Ermittlung von Klimawirkungen. Darüber hinaus sind detaillierte Kenntnisse (Expertenwissen) des Untersuchungsgegenstands (zum Beispiel Flusseinzugsgebiete, städtische Regionen, Ökosysteme) erforderlich, für die Klimawirkungen ermittelt werden sollen. Dies gilt sowohl für die Beschreibung von Klimawirkungen in bestimmten Sektoren als auch in sektorübergreifender Perspektive. Ferner können weitere Datensätze (bspw. ökologische, demographische, sozio-ökonomische) angekoppelt werden. Eine der größten Herausforderungen für die Verarbeitung von Klimamodelldaten in Wirkmodellen liegt in der Vielzahl der zu verarbeiteten Zeitreihen.

Für die Ausgestaltung der Schnittstelle zwischen Klimaprojektionsdaten und Wirkmodellen sind folgende Punkte zu beachten:

Für die Bewertung von Klimasimulationen hat sich mittlerweile der Ensemble-Ansatz durchgesetzt. Dies gilt im übertragenen Sinne auch für die Ermittlung von Klimafolgen. Die Aussagekraft eines Simulationslaufs eines Wirkmodells, angetrieben durch einen Lauf eines einzelnen Klimamodells ist eingeschränkt, da die Bandbreite möglicher Klimafolgen hierdurch nicht angemessen abgebildet werden kann. Daher muss ein geeignetes Ensemble aus Emissionsszenarien, Global- und Regionalmodellen und Wirkmodellen ausgewählt werden, um möglichst eine plausible Bandbreite abzubilden. Zur Orientierung können dabei die aktuelle Entwicklung und/oder in Bezug auf den Untersuchungsgegenstand „worst case“/„best case“ Projektionen verschiedener Klimaparameter dienen.

Die Klimasensitivität des Untersuchungsgegenstandes muss durch Sensitivitätsstudien ermittelbar sein. Dadurch wird gezeigt, wie stark verschiedene Klimaprojektionen und Beobachtungsdaten die Ausprägung von Klimawirkungen beeinflussen. Ein weiteres Ziel von Sensitivitätsstudien kann die Identifikation von Schwellenwerten sein. Schwellenwerte geben an, dass eine relevante Änderung im untersuchten System ausgelöst wird.

Es gibt unterschiedlich komplexe Wirkmodelle mit unterschiedlichen Anforderungen an die Eingangsdaten aus Klimamodellen. Wirkmodelle, die mit Beobachtungsdaten kalibriert und validiert wurden, können nicht ohne Weiteres mit Klimamodelldaten angetrieben werden. Die Validierbarkeit der Ergebnisse von Wirkmodellen muss – analog wie bei Klimamodellen – möglich sein und durchgeführt werden und sich auf einen Referenzzeitraum beziehen.

Ist eine Diskrepanz zwischen Klimamodelldaten und Beobachtungsdaten festgestellt worden, so muss ggf. eine Bias-Minimierung vorgenommen werden.

Die räumliche und zeitliche Auflösung von regionalem Klima- und Wirkmodell müssen aufeinander abgestimmt sein. Die Interpretation von Klimawirkmodellergebnissen muss auf die unterschiedlichen Auflösungen der verwendeten Daten Bezug nehmen. Die Daten mit der größten Auflösung begrenzen die Aussagekraft in Raum und Zeit. Anderenfalls muss eine räumliche und zeitliche Verfeinerung des grob aufgelösten Datensatzes vorgenommen werden. Dabei reicht eine einfache Interpolation nicht aus, da durch ein solches Verfahren keine neue Information hinzukommt.

Die Anwendung von Wirkmodellen kann zu weiteren Unsicherheiten führen und daher die Bandbreiten der Ergebnisse vergrößern. Daher wird auch bei Wirkmodellen die Nutzung verschiedener Modelle empfohlen (Modell- und Methodenvielfalt; s.o. erster Punkt), auch wenn dies einen erheblichen Mehraufwand bedeuten kann.

Als Beispiele für eine gelungene Kopplung von Klima- und Klimafolgenmodellen wird exemplarisch auf zwei Studien hingewiesen:

- Die Vulnerabilitätsstudie des UBA (Zebisch et al. 2005): Für den Bereich Wasserhaushalt / Wasserwirtschaft konnten die Ergebnisse verschiedener Klimamodelle und Klimaszenarien verwendet werden, um Aussagen u. a. über den künftigen Abfluss zu generieren.
- Die Ermittlung von Auswirkungen des Klimawandels auf die Elbe (Horsten et al. 2011) sowie auf Rhein und Donau (www.kliwas.de) im Kontext des KLIWAS-Projekts: Hier wurden die Änderungen von Wasserhaushaltskomponenten im Elbegebiet durch eine Impactmodellierung auf der Basis von 24 Klimaprojektionen berechnet.

Wenn für den zu modellierenden Sachverhalt die entscheidenden Schwellenwerte, die klimatischen Eckpunkte bzw. eine Kombination verschiedener Klimaparameter (zum Beispiel in Form eines „Ökogramms“, Spellmann et al. 2011) fachlich geklärt sind, sollte das Fachwissen und die Erfahrung der Klimamodellierer genutzt werden, um zu entscheiden, wie ein geeignetes Ensemble für dieses Wirkmodell ausgewählt werden kann. Dazu stehen der Deutsche Wetterdienst, die Klimafachzentren und Landesämter der Bundesländer sowie des Umweltbundesamtes (KOMPASS), das Climate Service Centers 2.0 und die Klimabüros der Helmholtz-Gemeinschaft zur Verfügung.

7. Konsolidierte Wortwahl

Aus Sicht des Fachgespräches ist es notwendig mittelfristig zu einer neuen Qualität der Auswertung von Klimaprojektionen zu gelangen. Es wird empfohlen, in den Texten zum Klimawandel eine einheitliche, definierte Sprache zu entwickeln und zu nutzen. Sie soll helfen um:

- Das Vertrauen in die Stichhaltigkeit der Erkenntnisse, basierend auf der Art, der Menge, der Qualität, der Konsistenz der Belege und den Grad der Übereinstimmung zu beschreiben. Das Vertrauen wird qualitativ beschrieben.
- Ein quantifiziertes Maß der Unschärfe der Erkenntnisse ausgedrückt über Wahrscheinlichkeitsaussagen bereitzustellen.

Basis der folgenden Ausführungen sind Texte die im Zusammenhang mit der Erstellung des 5. Sachstandberichtes des IPCC entstanden sind. Das Kapitel stellt eine komprimierte Zusammenführung und Übersetzung der Texte Mastrandrea et al. 2010, IPCC 2013, IPCC-DE 2014 und APCC 2014 dar.

In Texten und Zusammenfassungen für politische Entscheidungsträger sollten für die Beschreibung der verfügbaren **Belege der Änderungen** (der Art, der Menge, der Qualität und der Konsistenz der Belege) im Rahmen des Klimawandels die folgenden summarischen Ausdrücke verwendet werden: *schwache*¹, *mittel* oder *robust*; und für den **Grad der Übereinstimmung**: *gering*, *mittel* und *hoch*.

Tabelle 1: Bestimmung der Vertrauensbereiche (Grad des Vertrauens)

Grad der Übereinstimmung (zu einzelnen Aussagen)	Hohe Übereinstimmung Schwache Beweislage	Hohe Übereinstimmung Mittlere Beweislage	Hohe Übereinstimmung Robuste Beweislage
	Mittlere Übereinstimmung Schwache Beweislage	Mittlere Übereinstimmung Mittlere Beweislage	Mittlere Übereinstimmung Robuste Beweislage
	Geringe Übereinstimmung Schwache Beweislage	Geringe Übereinstimmung Mittlere Beweislage	Geringe Übereinstimmung Robuste Beweislage
Beweislage (der Art, der Menge, der Qualität und der Konsistenz der Belege)			

Der **Vertrauensgrad** wird durch die fünf Ausdrücke *sehr gering*, *gering*, *mittel*, *hoch* und *sehr hoch* beschrieben und kursiv gesetzt, z.B. *mittleres Vertrauen*. Für eine bestimmte Belegs-

¹ Die deutsche Übersetzung der Zusammenfassung für politische Entscheidungsträger (IPCC-DE 2014) nutzt hier das Wort ‚beschränkt‘.

und Übereinstimmungsangabe können unterschiedliche Vertrauensgrade angegeben werden, aber ein steigender Umfang an Belegen und Grad der Übereinstimmung sind mit zunehmendem Vertrauen verbunden. In Tabelle 2 wird jedem Vertrauensgrad eine Maßzahl an Belegen zugeordnet.

Tabelle 2: Beschreibung des Vertrauensgrades

Begriff	Vertrauensgrad
<i>sehr hohes Vertrauen</i>	In mindestens 9 von 10 Fällen korrekt
<i>hohes Vertrauen</i>	In etwa 8 von 10 Fällen korrekt
<i>mittleres Vertrauen</i>	In etwa 5 von 10 Fällen korrekt
<i>geringes Vertrauen</i>	In etwa 2 von 10 Fällen korrekt
<i>sehr geringes Vertrauen</i>	In weniger als 1 von 10 Fällen korrekt

Alternativ / ergänzend dazu ist es auch möglich eine Angabe der abgeschätzten **Wahrscheinlichkeit eines Befundes oder Resultats** durch die in Tabelle 3 benannten Ausdrücke anzugeben.

Tabelle 3: Verwendete Begriffe im Bereich Wahrscheinlichkeiten²

Begriff	Wahrscheinlichkeit
<i>praktisch sicher</i>	≥ 99 % Wahrscheinlichkeit
<i>sehr wahrscheinlich</i>	≥ 90 % Wahrscheinlichkeit
<i>wahrscheinlich</i>	≥ 66 % Wahrscheinlichkeit
<i>ebenso wahrscheinlich wie nicht</i>	33–66% Wahrscheinlichkeit
<i>unwahrscheinlich</i>	0–33% Wahrscheinlichkeit
<i>sehr unwahrscheinlich</i>	0–10% Wahrscheinlichkeit
<i>besonders unwahrscheinlich</i>	0–1% Wahrscheinlichkeit

Zusätzliche Ausdrücke (*äußerst wahrscheinlich* ≥ 95%, *eher wahrscheinlich als nicht* >50% und *äußerst unwahrscheinlich* 0–5%) können falls angebracht auch verwendet werden. Die abgeschätzte Wahrscheinlichkeit ist kursiv gesetzt, z.B. *sehr wahrscheinlich*.

² Der Wahrscheinlichkeitsbegriff ist in diesem Zusammenhang NICHT im mathematischen oder Wahrscheinlichkeitstheoretischen Sinne zu verstehen, sondern eher als subjektive Expertise, die häufig angewendet wird um eine Einschätzung zu den Ergebnissen aus Klimamodellensembles vorzunehmen. Beispiel dafür ist die "Guidance Note for Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties".

8. Glossar

Einige wesentliche und häufig genutzte Begriffe werden im Folgenden definiert.

(Die Auflistung erfolgt in alphabetischer Reihenfolge)

20C-SIMULATION / HISTORISCHE SIMULATION / KONTROLL-LAUF

Mit einem globalem Klimamodell (GCM) werden Klimaverhältnisse auf Grund des Strahlungsantriebes durch sich ändernde Treibhausgas-Konzentrationen (THG-Konzentrationen) errechnet. Für die Zukunft bilden Szenarien die Grundlage für die Veränderung dieses Antriebes.

GCMs werden zudem aber auch eingesetzt, um das gegenwärtige Klima – auf der Basis der gegenwärtigen THG-Verhältnisse – zu simulieren. In der Regel handelt es sich um den Zeitraum des 20. Jahrhunderts, weswegen dafür auch der Begriff 20C-Simulation (engl.: 20th Century) benutzt wird. Häufig ist dieser Zeitraum auf das rezente Klima, also nach 1950, eingegrenzt. Wird der Zeitraum noch weiter in die Vergangenheit ausgedehnt, beispielsweise um den durch gleichbleibenden THG-Antrieb charakterisierten Atmosphärenzustand der vorindustriellen Zeit zu beschreiben, so heißen diese Läufe auch historische Simulationen. Ein wichtiger Grund, 20C-Simulationen durchzuführen ist die Notwendigkeit, zu testen, inwiefern das GCM das durch Messungen bekannte Gegenwarts-klima beschreibt. Bisweilen wird in diesem Zusammenhang fälschlich auch die Bezeichnung Kontroll-Läufe verwendet.

(Anmerkung: In der Klimamodellierung wird der Begriff Kontroll-Lauf auch für die Simulation der vorindustriellen Bedingungen genutzt. Diese Läufe simulieren konstante THG-Konzentrationen, enthalten jedoch die vom Modell simulierte natürliche Klimavariabilität. Da der exakte Zustand der Atmosphäre zur vorindustriellen Zeit nicht genau bekannt ist, werden häufig mehrere Simulationen des 20. Jahrhunderts zu verschiedenen Zuständen der natürlichen Klimavariabilität im Modell gestartet. Dies bildet dann ein Anfangsbedingungen-Ensemble).

BIAS-MINIMIERUNG (ENGL.: BIAS-ADJUSTMENT)

Als Bias (systematischer Fehler) werden die Abweichungen der einzelnen Klimamodelle zu Beobachtungsdaten des rezenten Klimas bezeichnet. Wegen dieser Abweichungen ist die direkte Verwendung der von diesen Modellen für die Zukunft berechneten Datenzeitreihen für viele Anwendungen nicht empfehlenswert, da das Vorhandensein eines Bias einen verfälschenden Einfluss auf das Klimaänderungssignal haben kann.

Einen Ausweg bietet die Anpassung des simulierten Datensatzes an die Beobachtungsdaten zur Verringerung der Abweichung. Der oft genannte Begriff Bias-Korrektur ist dabei irreführend, da der systematische Fehler nicht eliminiert, sondern durch die Nachbearbeitung lediglich gemindert wird. Zudem wird der systematische Fehler nur für das rezente Klima tatsächlich gemindert, jedoch nicht unbedingt auch für die Zukunft, da man im Allgemeinen nicht von einem stationären Bias ausgehen kann.

ENSEMBLE

Hierbei handelt es sich um eine Gruppe von Modell-Läufen, die zusammengefasst betrachtet und ausgewertet werden. Es gibt verschiedenen Arten, Ensembles aufzubauen.

Anfangsbedingungs-Ensemble (engl.: initial conditions ensemble): In diesem Ensemble wird dasselbe Modell, dasselbe antreibende Szenario verwendet, aber eine Anzahl von Simulationen, die zu verschiedenen Simulations-Zeitpunkten in der vorindustriellen Epoche (s. Eintrag zu den *20C-Simulationen* -> *Kontroll-Lauf*) aufgesetzt wurden. Verbreitet wird dieses Ensemble auch als *multiple Läufe desselben Modells* bezeichnet.

Modellphysik-Ensembles (engl.: perturbed physics ensemble): Hierbei handelt es sich um dasselbe Modell und dasselbe antreibende Szenario. Es werden mehrere Läufe verwendet, die jeweils mit verschiedenen Annahmen, Approximationen und Parametrisierungen in den mathematischen Gleichungen des Modells erfolgten (Murphy et al., 2009).

Multimodell-Ensemble: Es werden verschiedene (globale oder regionale) Modelle jedoch dasselbe antreibende Szenario, z.B. SRES A1B zur Bildung des Ensembles verwendet.

Multimodell-Multiszenario-Ensemble: Es werden verschiedene Modelle und verschiedene Szenarien zur Bildung des Ensembles verwendet.

ESD

Empirisch-statistisches Downscaling (engl. Empirical Statistical Downscaling, z.T. auch statistisches Downscaling genannt) ist ein Oberbegriff für statistische Methoden (z.T. auch statistische Modelle genannt), die eine Verbindung zwischen großräumiger Klimainformation (z.B. aus einem GCM) und dem lokalen Klima (z.B. an einer Beobachtungsstation) herstellen (s. a. Benestad et al., 2005). Im Unterschied zum dynamischen Downscaling verwenden die statistischen Verfahren üblicherweise weniger komplexe Informationen aus dem GCM als Antrieb und lassen sich sehr viel schneller berechnen.

EXTREMEREIGNISSE / EXTREMWERTE

Extremereignisse sind per Definition (sehr) seltene Ereignisse. Diese Definition ist bezogen auf den jeweiligen Ort und dem Zeitpunkt im Jahr. Extremereignisse können auch durch eine Akkumulation von mittleren Wetterereignissen entstehen (z.B.: Trockenheit oder Hochwasser). Ein Extremwert beschreibt ein Extremereignis mit einem oder mehreren Kennwerten.

GCM / ESM

Globale Klimamodelle, auch allgemeine Zirkulationsmodelle (engl. Global Climate Models oder General Circulation Models, beides mit dem Kürzel GCM bezeichnet), Klimasystemmodelle oder Erdsystemmodelle (engl. Earth System Models, ESM) genannt, sind numerische (Computer-)Modelle, welche die zeitliche Entwicklung der globalen Atmosphäre, des Ozeans, des Bodens, der Biosphäre und der Kryosphäre simulieren. In ihnen werden die physikalischen Prozesse des Klimasystems durch mathematische Algorithmen repräsentiert. ESM enthalten zusätzlich biogeochemische Prozesse und Wechselwirkungen, inklusive des Kohlenstoffkreislaufes. Übliche Gitterboxdimensionen sind in der Größenordnung von 150 bis 200 km. (Eine gute Darstellung der globalen Klimamodellierung findet sich in McGuffie und Handerson.-Sellers, 2005.

KLIMAÄNDERUNGSSIGNAL

Beim Klimaänderungssignal handelt es sich um den Unterschied zwischen dem Wert einer Klimagröße in einer Periode (z.B. in der Vergangenheit) und in einer späteren Periode (z.B. in der Zukunft). Diese Betrachtung hat den Vorteil, dass damit systematische Modellfehler (siehe Eintrag *Bias-Minimierung*) teilweise eliminiert werden können und Ergebnisse zwischen verschiedenen Modellen vergleichbar werden. Diese Änderung als *Signal* zu bezeichnen, erfordert, dass bestimmte Schwellen der Veränderung überschritten werden. Diese Schwellen werden beispielsweise durch die Analyse der Variabilität (siehe Eintrag *Klimavariabilität*) dieser Größe in der ersten Periode bestimmt.

KLIMAPROJEKTION

Eine Klimaprojektion ist keine Klimavorhersage. Eine Klimaprojektion ist die Beschreibung eines möglichen und plausiblen Zukunfts-Zustands des Klimasystems nebst der zeitlichen Entwicklungslinie, die dorthin führt. Sie wird üblicherweise mit einem Klimamodell (GCM, RCM oder ESD) erstellt.

KLIMAVARIABILITÄT

Diese Größe beschreibt zeitliche Schwankungen im Klimasystem um einen Mittelzustand (z.B. der globalen oder regionalen Jahres- oder Monatswerte der Lufttemperatur, der Niederschlagsmengen, der Extremwerte o.ä.). Betrachtungen der Klimavariabilität erstrecken sich häufig auf Zeitbereiche von Monaten oder Jahren bis hin zu mehreren Jahrzehnten. Der Begriff natürliche Klimavariabilität wird für denjenigen Anteil verwendet, der nicht der menschlichen Aktivität zugeschrieben werden kann. Die Klimavariabilität darf nicht fälschlich als Klimasignal (siehe Eintrag *Klimasignal*) interpretiert werden. Daher werden bei Klimauntersuchungen üblicherweise lange Zeiträume (z.B. 30 Jahre) betrachtet.

MODELLKETTE

Im Kontext der Regionalisierung/Downscaling erscheint dieser Begriff in zwei verschiedenen Zusammenhängen (i) Wie in den Terminologie-Erläuterungen zu RCM, ESD und Regionalisierung erläutert, gibt es verschiedene Strategien, um eine Regionalisierung zu erreichen. Sie erfordern jeweils eine Kette oder Kaskade von Modellen verschiedener Auflösung (angefangen mit einem GCM kann entweder ein dynamisches oder ein statistisches Modell oder auch eine Abfolge von Downscalingmethoden genutzt werden). (ii) In der Terminologie des IPCC handelt es sich bei einer Kette oder Kaskade um eine Reihe von Modellen, die beispielsweise zum Entwickeln eines Emissionsszenarien zum Einsatz kommen: Aufbauend auf Modell-basierten Annahmen zur Wirtschaftsentwicklung transformiert ein Modell diese Information in Treibhausgas (THG)-Emissionen (sozio-ökonomisches Modell). Diese wiederum werden von einem Kohlenstoffkreislauf-Modell in atmosphärische THG-Konzentrationen umgerechnet und daraus wird mit einem weiteren Modell der Strahlungsantrieb (s.u.) bestimmt. Letzterer ist Eingangsgröße für ein GCM, das dann im Sinne der Definition (i) zum Antrieb eines RCM oder eines ESD genutzt werden kann. In der Folge können die Ergebnisse der RCM/ESD-Läufe als Antriebsdaten, z.B. für Klimafolgen- oder Wirkmodelle genutzt werden (s.a. Abschnitt 2.4).

RCM

Ein (dynamisches) regionales Klimamodell (Regional Climate Model) ist ein Ausschnittmodell, welches ein begrenztes Gebiet der Erde mit einer höheren räumlichen

Auflösung im Vergleich zum GCM simuliert. Derzeit liegt die typische Gitterweite bei etwa 12 km. Die benötigten Antriebsinformationen aus dem Bereich außerhalb dieses Ausschnitts werden aus einem GCM verwendet. Das dynamische Downscaling liefert für den simulierten Bereich alle relevanten physikalischen Parameter (z.B.: Temperatur, Feuchte, Wind, Wolken, Druck, etc.), nicht nur am Boden, sondern auch in der Höhe, mit einer direkten Kopplung an das antreibende globale Modell.

REGIONALISIERUNG/DOWNSCALING

In der Meteorologie und Klimaforschung beschreibt der Begriff Regionalisierung den Prozess, mit dem eine grob aufgelöste Information (z.B. aus einem GCM) auf eine höhere räumliche Auflösung gebracht wird (engl. Downscaling). In Anlehnung an die Definition in Benestad et al. (2008) handelt es sich dabei um den Vorgang, eine Verbindung zwischen einer Variablen, die die großräumige Skala repräsentiert und einer Variablen, die eine erheblich kleinere Skala repräsentiert, herzustellen. Eine Herangehensweise der Regionalisierung ist die Anwendung eines RCM (siehe Eintrag *RCM*) in einer Zielregion (dynamische Downscaling); ein alternativer Ansatz nutzt statistische Methoden (statistische Downscaling, siehe Eintrag *ESD*).

STRAHLUNGSANTRIEB

Der Strahlungsantrieb ist die Änderung der Netto-Strahlungsflussdichte (engl.: radiative forcing; physikalische Einheit: W/m^2) am Oberrand der Atmosphäre verursacht durch die Änderung des externen Antriebs (z.B. CO_2 Konzentration). Er wirkt sich sowohl auf das Strahlungs- als auch das Temperatur-Gleichgewicht der Atmosphäre aus. Veränderungen in der chemischen Zusammensetzung der Atmosphäre, beispielsweise beim Anteil der Treibhausgase können die Eigenschaften des Strahlungsflusses durch die Atmosphäre ändern. In einem GCM werden die sich ändernden Atmosphären-Bedingungen (in Vergangenheit, Gegenwart oder einer möglichen Zukunftsentwicklung) als Strahlungsantrieb ausgedrückt.

(KLIMA)SZENARIEN

Ein Szenario ist eine Beschreibung einer möglichen Zukunft auf Grund von Annahmen. Das kann durch eine in sich schlüssige Kette von Annahmen bezüglich der politischen, wirtschaftlichen und ökologischen Bedingungen in der Zukunft und daraus abgeleiteten Veränderungen der THG-Emissionen geschehen. Es handelt sich dabei um Special Reports on Emission Scenarios (so genannte SRES-Szenarien, Nakicenovic et al., 2000) aus dem 4. Assessment-Report (AR). Die neuen Szenarien des 5. AR geben nur verschiedene vorherrschende Strahlungsantriebe (s. oben) zum Ende des 21. Jahrhunderts an, mit denen Klimamodelle angetrieben werden und aus denen mögliche Pfade von THG-Konzentrationen entwickelt werden. Diese werden als Representative Concentration Pathways bezeichnet (kurz RCP-Szenarien, Moss et al., 2008 und 2010).

VORHERSAGE

Vorhersagen leiten aus dem vergangenen und aktuellen Zustand der Atmosphäre Aussagen über dessen zukünftigen Zustand ab. Traditionell beinhaltet eine Vorhersage die Wetterentwicklung der nächsten ein bis zehn Tage.

Ein aktueller Forschungsgegenstand ist die Abschätzung der Entwicklung über diesen Zeithorizont hinaus für Zeitskalen von mehreren Monaten bis wenigen Dekaden (Keenlyside und Ba, 2010).

VULNERABILITÄT

Der Grad, zu dem ein System (z.B. Erdsystem, Ökosystem, wirtschaftliches System) nachteilig auf Störungen oder Gefährdungen reagiert, also die Anfälligkeit eines Systems für nachteilige Einwirkungen (Impacts).

9. Literaturverzeichnis

- APCC (2014): Österreichischer Sachstandsbericht Klimawandel 2014 (AAR14). Austrian Panel on Climate Change (APCC), Verlag der Österreichischen Akademie der Wissenschaften, Wien, Österreich, 1096 Seiten. Verfügbar unter: <http://www.apcc.ac.at/4%20-%20Report.html>
- Benestad, R.E., I. Hanssen-Baue und D. Cheng, 2008: Empirical-statistical Downscaling. World Scientific Publishing Co. Pte. Ltd., Singapore, 215 pp.
- Bosshard, T., Kotlarski, S., Ewen, T. und Schär, C., 2011: Klimaszenarien für Klimaimpaktstudien in der Schweiz. – IAC, Eidgenössische Technische Hochschule Zürich (interner Bericht).
- Carter, T (2010) Representing multi-model climate projection uncertainties in modelling impact risks and adaptation options: Recent advances in europe. In: IPCC Expert Meeting on Assessing and Combining Multi Model Climate Projections. Boulder, Colorado, USA, 25-27 January 2010
- CH2011, 2011: Swiss Climate Change Scenarios CH2011. - published by C2SM, MeteoSwiss, ETH, NCCR Climate and OcCC, Zurich, Switzerland, 88 pp.
- Christensen, J.H. and O.B. Christensen, 2007: A summary of the PRUDENCE model projections of changes in European climate by the end of this century. Climatic Change, 81(S1), 7-30
- Deutschländer, T., Imbery, F., Plagemann, S., Namyslo, S., Früh, B. und Becker, P., 2009: Zum Umgang mit Klimaprojektionen in der Klima- und Klimafolgenforschung. – Homepage des DWD.
- Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, and J. Liebert (2012): Should we apply bias correction to global and regional climate model data? Hydrol. Earth Syst. Sci., 16, 3391-3404, www.hydrol-earth-syst-sci.net/16/3391/2012/, doi:10.5194/hess-16-3391-2012
- Fabig, I., 2007: Die Niederschlags- und Starkregenentwicklung der letzten 100 Jahre im Mitteldeutschen Trockengebiet als Indikator möglicher Klimaänderungen. Dissertation an der Martin-Luther-Universität Halle-Wittenberg. E-Dokument, ULB Sachsen-Anhalt.
- Frick et al., 2014: Central European high-resolution gridded daily datasets (HYRAS): Mean temperature and relative humidity Meteorologische Zeitschrift, Vol. 23, 15-32. DOI: 10.1127/0941-2948/2014/0560
- Hennemuth, B., S. Bender, K. Bülow, N. Dreier, E. Keup-Thiel, O. Krüger, C. Mundersbach, C. Radermacher, R. Schoetter (Hrg.), 2013: Statistische Verfahren zur Auswertung von Klimadaten aus Modell und Beobachtung. CSC Report 13, Climate Service Center, Germany, 135pp.
- Horsten, T.; Krahe, P.; Nilson, E.; Belz, J.U.; Ebner von Eschenbach, A.-D.; Larina, M. 2011: Änderungen von Wasserhaushaltskomponenten im Elbegebiet – Herausforderungen und Lösungsansätze. In: BMVBS (Hrsg.) Tagungsband KLIWAS - Auswirkungen des Klimawandels auf Wasserstraßen und Schifffahrt in Deutschland. 2. Statuskonferenz Oktober 2011, Berlin. S. 93-97.
- Imbery, F., Plagemann, S., and Namyslo, J.: Processing and analysing an ensemble of climate projections for the joint research project KLIWAS, Adv. Sci. Res., 10, 91-98, doi:10.5194/asr-10-91-2013, 2013.

- InterAcademy Council, Committee to Review the Intergovernmental Panel on Climate Change, 2010: Climate change assessments - Review of the processes and procedures of the IPCC, 2010. ISBN 978 90 6984 617 0, <http://reviewipcc.interacademycouncil.net/report.html>
- IPCC, 2007: Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Pachauri, R.K and Reisinger, A. (eds.)]. IPCC, Geneva, Switzerland, 104 pp.
- IPCC, 2013: Working Group I, Contribution to the IPCC Fifth Assessment Report (AR5), Climate Change 2013: The Physical Science Basis
- IPCC-DE, 2014: Zusammenfassung für politische Entscheidungsträger. In: Klimaänderung 2013: Wissenschaftliche Grundlagen. Beitrag der Arbeitsgruppe I zum Fünften Sachstandsbericht des Zwischenstaatlichen Ausschusses für Klimaänderung (IPCC), Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (Hrsg.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 Seiten. Deutsche Übersetzung durch ProClim, Deutsche IPCC-Koordinierungsstelle, Österreichisches Umweltbundesamt, Bern/Wien/Bonn, 2014. Verfügbar unter: http://climatechange2013.org/images/report/WG1AR5_SPM_brochure_de.pdf
- Keenlyside, N.S. und J. Ba, 2010: Prospects for decadal climate prediction. WIREs Climate Change, 1, 627-635.
- Knutti R, Abramowitz G, Collins M, Eyring V, Gleckler P, Hewitson B, Mearns L (2010) Good practice guidance paper on assessing and combining multi model climate projections. In: Stocker T, Qin D, Plattner G-K, Tignor M, Midgley P (eds) Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Assessing and Combining Multi Model Climate Projections, IPCC Working Group I Technical Support Unit, University of Bern, Bern, Switzerland
- Knutti, R., Masson, D., Gettelmann, A. 2013: Climate model genealogy: Generation CMIP5 and how we got there. GEOPHYSICAL RESEARCH LETTERS, VOL. 40, 1–6, doi:10.1002/grl.50256
- LUBW, 2012: Vergleich regionaler Klimaprojektionen für Baden-Württemberg. (2012).
- Mc Guffie, K. und A. Henderson-Sellers, 2005: A Climate Modelling Primer, Third Edition. Wiley & Sons, Chichester, UK., 280 pp.
- MacHardy, W. E., Gadoury, D. M., 1989: A Revision of Mills's Criteria for Predict Apple Scab Infection Periods. Phytopathology, 79, pp. 304-310.
- Mastrandrea, M.D., C.B. Field, T.F. Stocker, O. Edenhofer, K.L. Ebi, D.J. Frame, H. Held, E. Kriegler, K.J. Mach, P.R. Matschoss, G.-K. Plattner, G.W. Yohe, and F.W. Zwiers (2010): Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties. Intergovernmental Panel on Climate Change (IPCC). Verfügbar unter: <https://www.ipcc.ch/pdf/supporting-material/uncertainty-guidance-note.pdf>
- Meehl G, Stocker T, Collins W, Friedlingstein P, Gaye A, Gregory J, Kitoh A, Knutti R, Murphy J, Noda A, Raper S, Watterson I, Weaver A, Zhao Z-C (2007) Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL, eds.). Cambridge University Press, Cambridge, UK and New York, USA. chapter 10, pp 747–846
- Moss, R., M. Babiker, S. Brinkman, E. Calvo, T. Carter, J. Edmonds, I. Elgizouli, S. Emori, L. Erda, K. Hibbard, R. Jones, M. Kainuma, J. Kelleher, J.F. Lamarque, M. Manning, B. Matthews,

- J. Meehl, L. Meyer, J. Mitchell, N. Nakicenovic, B. O'Neill, R. Pichs, K. Riahi, S. Rose, P. Runci, R. Stouffer, D. van Vuuren, J. Weyant, T. Wilbanks, J.P. van Ypersele und M. Zurek, 2008: Report: Towards New Scenarios for Analysis of Emissions, Climate Change, Impacts, and Response Strategies. 19--21 September 2007, Noordwijkerhout, The Netherlands, IPCC, 132 pp.
- Moss, R.H., J. A. Edmonds, K. A. Hibbard, M. R. Manning, S. K. Rose, D. P. van Vuuren, T. R. Carter, S. Emori, M. Kainuma, T. Kram, G. A. Meehl, J. F. B. Mitchell, N. Nakicenovic, K. Riahi, S. J. Smith, R. J. Stouffer, A. M. Thomson¹, J. P. Weyant und T. J. Wilbanks, 2010 The next generation of scenarios for climate change research and assessment. *Nature* 463, 747-756, doi:10.1038/nature08823
- Mudelsee M., 2010: *Climate Time Series Analysis: Classical Statistical and Bootstrap Methods*. Springer, Dordrecht, 474 pp.
- Murphy, J.M., S.M.H. Sexton, G.J. Jenkins, P.M. Boorman, B.B.B. Booth, C.C. Brown, R.T. Clark, M. Collins, G.R. Harris, E.J. Kendon, R.A. Betts, S.J. Brown, T.P. Howard, K.A. Humphrey, M.P. McCarthy, R.E. McDonald, A. Stephens, C. Wallace, R. Warren, R. Wilby and R.A. Wood, 2009: UK Climate Projections Science Report: Climate change projections. Met Office Hadley Centre, Exeter, 194 pp.
- Nakićenović N, Alcamo J, de Vries B, Fenhann J, Gaffin S, Gregory K, Grübler A, Jung T, Kram T, Rovere E L, Michaelis L, Mori S, Morita T, Pepper W, Pitcher H, Price L, Raihi K, Roehrl A, Rogner H-H, Sandovski A, Schlesinger M, Shukla P, Smith S, Swart R, van Rooijen S, Victor N, Dadi Z (2000) Emissions Scenarios; A Special Reports of IPCC Working Group III. Cambridge University Press, Cambridge, UK
- NLWKN (2017): Globaler Klimawandel – Wasserwirtschaftliche Folgenabschätzung für das Binnenland. Abschlussbericht KliBiW Phase 4 (Hochwasser): In: NLWKN-Schriftenreihe Oberirdische Gewässer, Band 41, Norden.
- Parker, W (2011) When climate models agree: The significance of robust model predictions. *Phil Sci* 78: 579–600
- Pennell, Ch, Reichler, T., 2011: On the Effective Number of Climate Models. *Journal of Climate*, Vol. 24, 2358-2367, DOI: 10.1175/2010JCLI3814.1
- Plagemann, S., Imbery, F., Namyslo, J., 2014: Validierung und Bewertung von Klimaprojektionen – Bereitstellung von Klimaszenarien für den Binnenbereich. Schlussbericht KLIWAS-Projekt 1.02. KLIWAS-29/2014. DWD / BfG, Koblenz. DOI: 10.5675/Kliwas_29/2014_1.02
- Pfeifer, S., Bülow, K., Gobiet, A., Hänsler, A., Mudelsee, M., Otto, J. Rechid, D. Teichmann, C. and Jacob, D. (2015): Robustness of Ensemble Climate Projections Analyzed with Climate Signal Maps: Seasonal and Extreme Precipitation for Germany. *Atmosphere* 6, 677-698. doi: 10.3390/atmos6050677
- Rauthe et al., 2013: A Central European precipitation climatology – Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS). *Meteorologische Zeitschrift*, Vol. 22, 235-256. DOI: 10.1127/0941-2948/2013/0436
- Schönwiese, C.-D., 2008: Extremereignisse aus meteorologisch-statistischer Sicht. *promet*, Jahrg. 34, Nr. 1/2, 61-65, http://www.geo.uni-frankfurt.de/iau/klima/PDF_Dateien/Sw_Promet_2008.pdf (26.04.2012).
- Schönwiese, C.-D., 2013: *Praktische Statistik für Meteorologen und Geowissenschaftler*. Borntraeger, 319 pp.

- Spekat, A. and Kreienkamp, F. (2007): Somewhere over the rainbow – advantages and pitfalls of colourful visualizations in geosciences, *Adv. Sci. Res.*, 1, 15-21, doi:10.5194/asr-1-15-2007
- Spellmann, H, M. Albert, M. Schmidt, J. Suttmöller und M. Overbeck (2011): Waldbauliche Anpassungsstrategien für veränderte Klimaverhältnisse. *AFZ-DerWald*
- von Storch H, Zwiers FW., 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge, 484 pp.
- Tebaldi, C., Arblaster, J.M., Knutti, R. (2011): Mapping model agreement on future climate projections. *Geophys. Res. Lett.*, 38, L23701.
- Trewin, B., 2007: The role of climatological normals in a changing climate. Report. WMO - WCDMP No. 61, WMO-TD No. 1377.
- van der Linden, P. and J.F.B. Mitchell (eds.), 2009: *ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project*. Met Office Hadley Centre, Exeter, UK, 160 pp.
- Yuan X, Wood, E. F., 2012: On the clustering of climate models in ensemble seasonal forecasting. *GEOPHYSICAL RESEARCH LETTERS*, VOL. 39, L18701, doi:10.1029/2012GL052735.
- Zebisch, M.; Grothmann, T.; Schröter, D.; Haße, C.; Fritsch, U.; Cramer, W. (2005): *Klimawandel in Deutschland – Vulnerabilität und Anpassungsstrategien klimasensitiver Systeme* Potsdam-Institut für Klimafolgenforschung.

10. Anlagen

10.1 Anlage 1: Beispiel Parameterdarstellung

Am Beispiel der Änderung der Niederschlagsmengen im Herbst für Deutschland (Vergleich 2071/2100 zu 1971/2000) anhand des Modells WETTREG2010 (auf Basis ECHAM5; A1B) wird eine sinnvolle bzw. nicht sinnvolle Darstellung von Ergebnissen aufgezeigt (siehe Abbildung 2). Während das linke Bild den Sachverhalt entsprechend den zuvor genannten Empfehlungen visualisiert, weist das Bild rechts mehrere ungünstige Eigenschaften auf. Zunächst wird mit einer farbpsychologisch verkehrten Farbskala gearbeitet (Rot = nasser, Grün = trockener). Außerdem befindet sich der Nullwert der Veränderungen nicht in der Mitte einer symmetrischen Skala (-15 bis +30%), und er besitzt eine Signalfarbe (Orange). Zudem beinhaltet die Skala die Farben Rot und Grün (nicht barrierefrei) und die Intervallgröße der Klassen liegt unterhalb der Signifikanzschwelle der Veränderungen (vgl. weißer Bereich im Bild links von 10%), so dass Regionen wie Niedersachsen oder Sachsen nun differenzierter wirken als sie aufgrund der realen Daten interpretiert werden sollten.

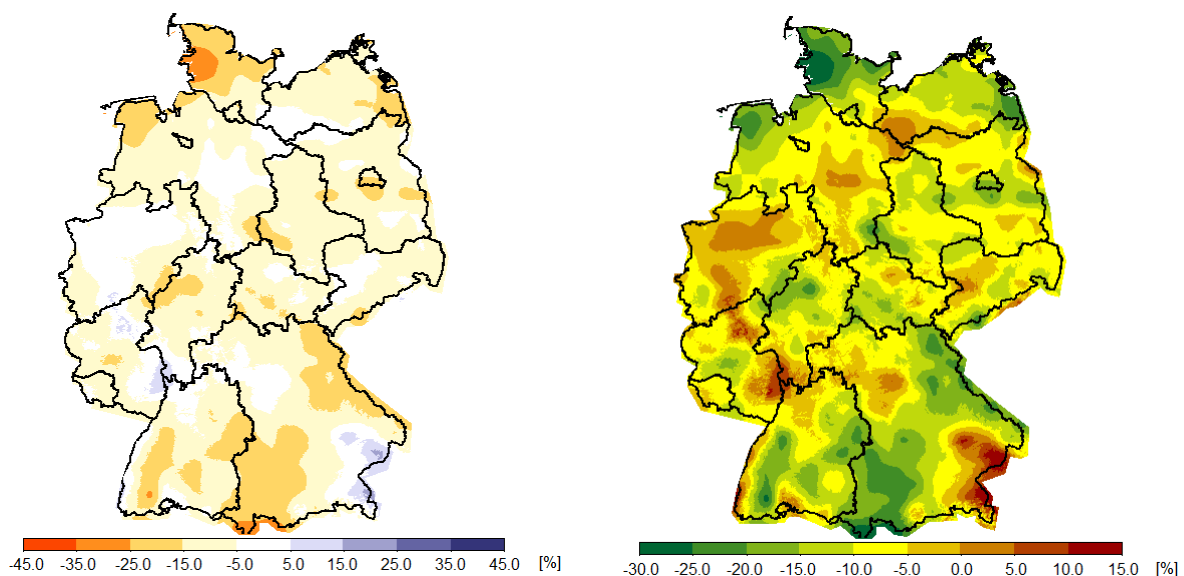


Abbildung 2: Anomalien der Niederschläge in Deutschland zwischen 2071/2100 und 1971/2000 für den Herbst (SON) anhand von WETTREG2010 (ECHAM5, A1B); links: zweckmäßige Darstellung, rechts: irreführende Darstellung

10.2 Anlage 2 Ensembledarstellung

Folgende Auswertungen beziehen ausschließlich auf den Standort Jena. Die verwendeten Beobachtungsdaten entstammen der Station des Deutschen Wetterdienstes (Jena – Sternwarte) im Schillergässchen der Innenstadt. Außerhalb des Stadtgebietes Jena können daher durchaus abweichende Absolutwerte auftreten. Die Änderungssignale bleiben davon jedoch unbeeinflusst.

Die erstellten Diagramme basieren auf der Auswertung mehrerer regionaler Klimaprojektionsmodelle (angetrieben durch ECHAM5/A1B) mit jeweils mehreren Modellläufen. Das daraus resultierende Ensemble besteht aus 16 Ensemblemitgliedern (10 Läufe WETTREG2010, 2 Läufe CCLM, 3 Läufe STAR, ein REMO-Lauf). Daraus wird eine Häufigkeitsverteilung gebildet und anhand der am häufigsten auftretenden Werte der wahrscheinlichste Zustand inklusive der vorhandenen Schwankungsbreite bestimmt.

In den Diagrammen sind jeweils die Messwerte (farbige Balken), sowie Median (schwarze strichlierte Linie) und die Bandbreite (grauer Bereich) der Modellprojektionen als gleitender Mittelwert bzw. Summe über zehn Jahre dargestellt. Die Obergrenze der Bandbreite ist als 80% Perzentil der Häufigkeitsverteilung aus allen Ensemblemitgliedern definiert und die Untergrenze wird durch das 20% Perzentil bestimmt (Vertrauensbereich). Diese empirische und in Anlehnung an andere Studien vorgenommene Eingrenzung soll „Modell-Ausreißer“ von der Betrachtung ausschließen.

Anzahl von Hitze-/Frostperioden

Eine Periode ist definiert als mindestens drei aufeinanderfolgende Tage des entsprechenden Temperaturgrenzwertes.

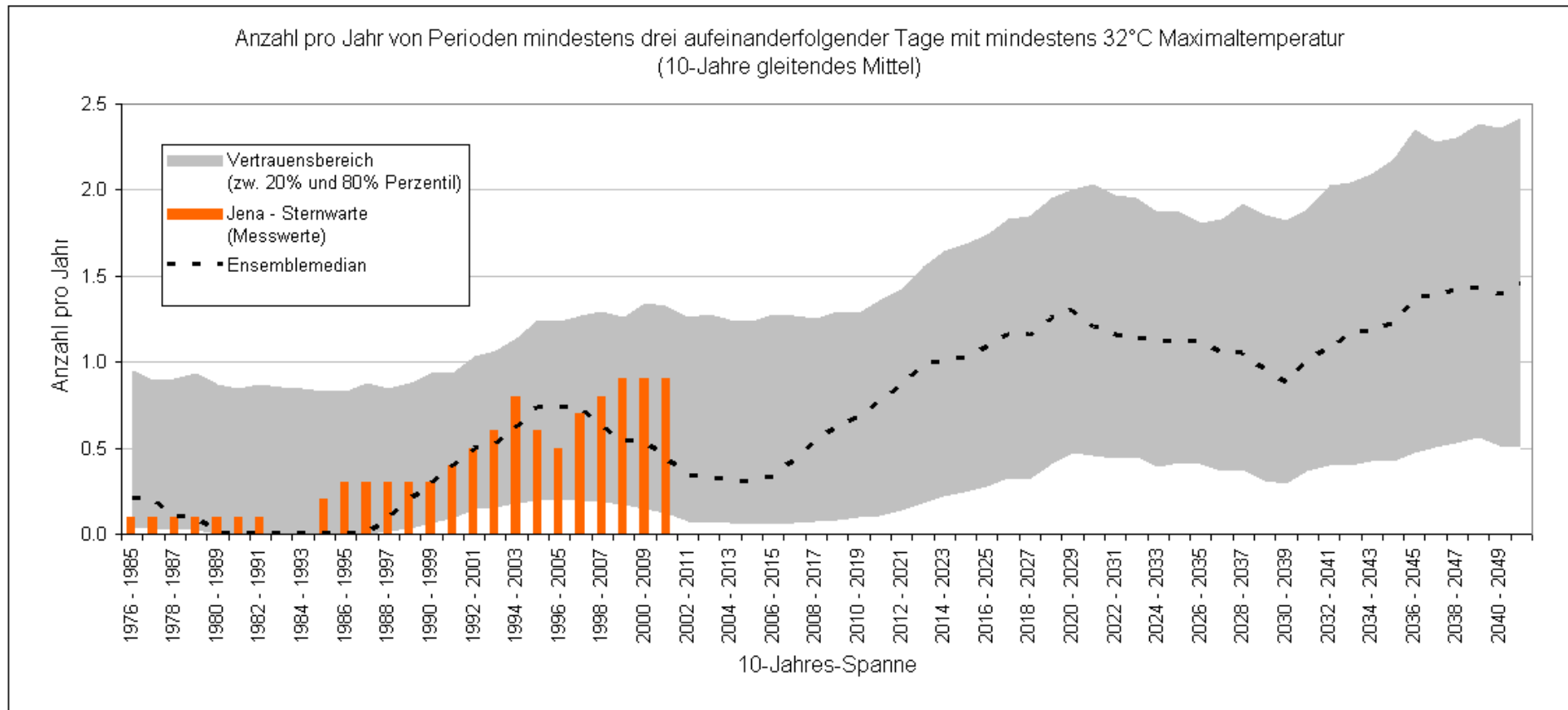


Abbildung 3: Jährliche Anzahl von Perioden von mindestens drei aufeinanderfolgenden Tagen mit mindestens 32°C Maximaltemperatur an der Wetterstation Jena-Sternwarte (Messwerte: DWD; Projektionen: Szenario A1B, ECHAM5, Modelle: WETTREG2010, REMO, CCLM, STAR)

10.3 Anlage 3: Beispiel Gesundheit / Hitzestress

Datengrundlage und Diagrammbeschreibung: Für die Tagesmitteltemperatur (schwarze durchgehende Kurve) wurde der Juli der Jahre 2021 bis 2050 (Standort Artern) ausgewählt. Die Werte entstammen dem Modell WETTREG2010 (Basis ECHAM5). Bezüglich der Ensemblebildung wurden für das Szenario A1B, alle zehn WETTREG2010-Realisierungen benutzt. Die gestrichelten Geraden sind die gefühlte Temperatur (Windchill) für 10, 20 und 30 km/h Windgeschwindigkeit. Die Einteilung der Wärmebelastungsbereiche erfolgte nach VDI 3787 Blatt 2. Es ist zu beachten, dass es sich hier um eine mittlere tägliche Wärmebelastung (d. h. durchschnittliche 24-stündige Wärmebelastung) handelt (Tag + Nacht), weil die Windchill-Temperatur mit der Tagesmitteltemperatur errechnet wurde.

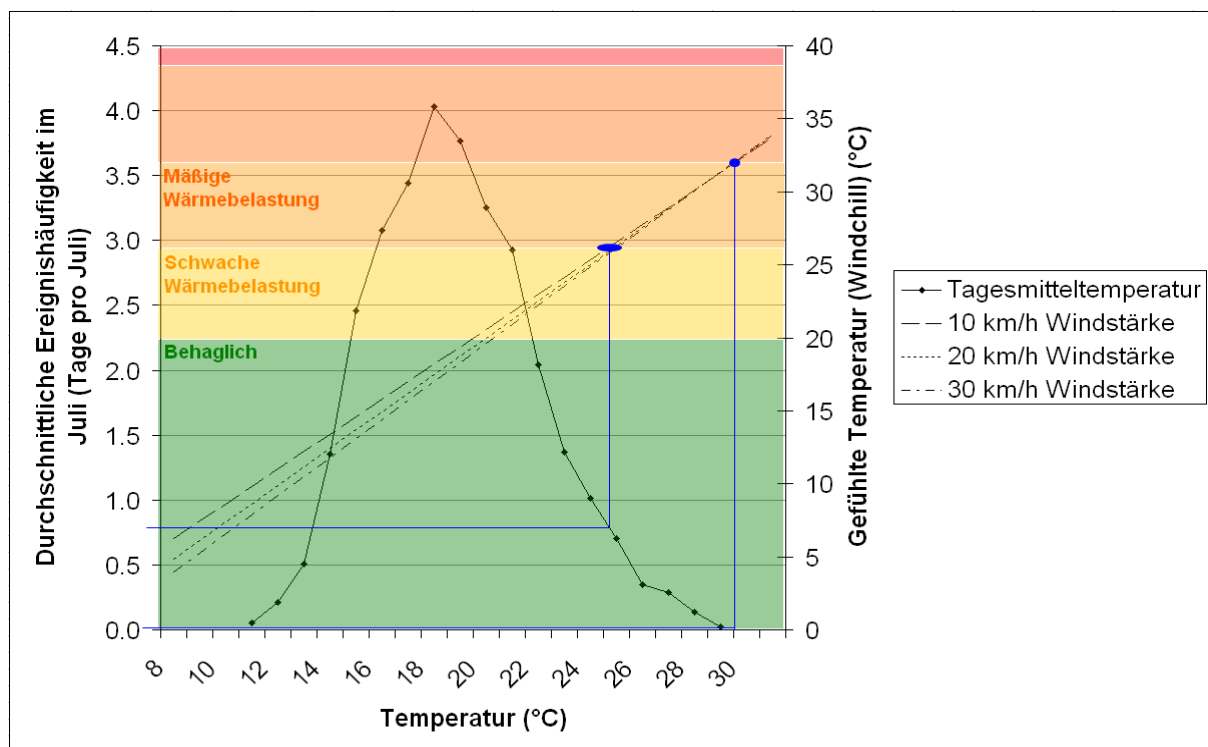


Abbildung 4: Windchilltemperatur für drei verschiedene Windgeschwindigkeiten und Häufigkeitsverteilungen der Tagesmitteltemperatur im Juli für den Zeitraum 2021 – 2050 am Beispiel des Standorts Artern (SRES A1B, ECHAM5, WETTREG2010)

Aus dem Diagramm kann die Ereignishäufigkeit der einzelnen Wärmebelastungsklassen abgeleitet werden. Mittlere mäßige Wärmebelastung (= Windchill-Temperatur von 26 – 32°C) tritt umgerechnet bei einer Tagesmitteltemperatur von 25 – 29°C auf. Die Ereignishäufigkeit des Bereiches von 25°C bis 29°C ergibt sich aus der Summe der Einzelhäufigkeiten. In diesem Beispiel liefert das einen Wert von 1.49. Das heißt, dass eine mittlere mäßige Wärmebelastung im Juli (2021 – 2050) durchschnittlich ein- bis zweimal pro Juli auftritt. Innerhalb des Zeitraumes von 2021 bis 2050 ist demnach mit ca. 45 Ereignissen dieser Art zu rechnen.

10.4 Anlage 4: Beispiel Klimatische Leitplanken

Die Ermittlung „klimatischer Leitplanken“ für die Entwicklung von Anpassungsstrategien und –maßnahmen soll am Beispiel einer Auswertung regionaler Klimaprojektionen für Baden-Württemberg demonstriert werden (LUBW, 2012). 29 regionale Klimaprojektionen mit verschiedenen Modellen lagen dieser Untersuchung zu Grunde. Nach einer Prüfung und Plausibilisierung wurden für die nahe Zukunft (2021-2050) 24 Klimaprojektionen und für die ferne Zukunft (2071-2100) 15 Klimaprojektionen verwendet.

Für die Auswertung der Klimaprojektionen wurde zum einen der Median der Klimaprojektionen betrachtet. Der Median ist der Mittelpunkt einer Bandbreite von Ergebnissen der Klimaprojektionen, die alle als gleich wahrscheinlich zu betrachten sind. Zum anderen wurde aus der Bandbreite der Klimaprojektionen ein Korridor bestimmt, in dem sich nach dem A1B Szenario der Klimawandel in Baden-Württemberg entwickeln kann.

Die Werte dieses Korridors werden als zukunftsbezogene „klimatische Leitplanken“ verstanden, da die unteren und oberen Bereiche der Auswirkungen des Klimawandels angegeben werden. Durch die dargestellten Bandbreiten wird zum einen die Streuung der Ergebnisse bei den derzeitigen Klimaprojektionen ersichtlich, zum anderen kann aber auch die Stärke und Richtungssicherheit der zu erwartenden Änderungen eingeschätzt werden. Die Stärke kennzeichnet das Ausmaß und die Höhe des Klimasignals. Richtungssicherheit liegt vor, wenn die Klimaprojektionen weitgehend die gleiche Tendenz (zunehmend, abnehmend oder unverändert) zeigen.

Trotz der Verwendung von einer großen Zahl von Klimamodellen in dieser Untersuchung können die wirklich eintretenden Werte in der Zukunft auch außerhalb dieser Leitplanken liegen. Diese Unsicherheit hängt zum einen mit der tatsächlichen zukünftigen Entwicklung zusammen. In diesem Zusammenhang sind beispielsweise die durch das gewählte Szenario bestimmten Treibhausgasemissionen zu nennen. Zum anderen können auch die aktuellen Klimamodelle nicht alle Elemente des Klimasystems bzw. des Wasserhaushalts berücksichtigen. So fehlt z.B. den globalen Klimamodellen noch ein gekoppelter Kohlenstoffkreislauf. Mit dem 5. IPCC-Sachstandsbericht (Basis CMIP5) sind weitere Modellverbesserungen zu erwarten. Aufgrund der hier durchgeführten Berücksichtigung einer großen Anzahl von Klimaprojektionen steigt jedoch die Wahrscheinlichkeit, dass die angegebenen Leitplanken vertrauenswürdig sind.

Um eine Entscheidung im Hinblick auf eine Anpassungsstrategie zu finden, müssen aus dem Ensemble der Klimaprojektionen handhabbare und fachlich korrekte Werte abgeleitet werden können. Wird aus dem Ensemble der Klimaprojektionen der Median gewählt, so wird der Mittelpunkt einer Spanne von Simulationen verwendet, die alle als gleich wahrscheinlich zu betrachten sind. Die alleinige Verwendung des Medians der Ergebnisse der Klimaprojektionen ist jedoch nicht objektiv zu begründen. Daher werden neben dem Median auch die Grenzen des Korridors angegeben, die sich aus der Analyse der Klimaprojektionen ergeben. In wie weit sich eine Anpassungsmaßnahme am oberen oder unteren Rands des Ensemblekorridors oder am Median des Ensembles orientiert, hängt von der jeweiligen Fragestellung im Handlungsfeld ab und ist von den Fachexperten zu entscheiden.

Als Korridor bzw. „klimatische Leitplanken“ wird der Bereich des 15. Perzentils und des 85. Perzentils vorgeschlagen. Die Betrachtung der Perzentile bietet den Vorteil, eine gewisse Bandbreite von Modellergebnissen zu erlauben und gleichzeitig Ausreißerwerte unberücksichtigt zu lassen. In Abbildung 4 (links) ist die Ermittlung der Perzentile anhand der Ergebnisse für die mittlere jährliche Lufttemperatur der Klimaprojektionen schematisch dargestellt:

Die mittleren Werte aus den regionalen Klimaprojektionen werden der Größe nach sortiert.

Das 50. Perzentil ist der Wert, für den jeweils die Hälfte der Modellsimulationen höhere bzw. niedrigere Änderungen zeigt. Er repräsentiert somit den Median und innerhalb des Ensembles ein mittleres Klimaänderungssignal.

Neben dem 50. Perzentil werden das 15. und 85. Perzentil angegeben, um die Bandbreite der Modellergebnisse zu verdeutlichen, ohne dass einzelne, besonders extreme Ergebnisse diesen Eindruck verfälschen.

Das 15. Perzentil ist der oberste Grenzwert der untersten 15% aller Modellsimulationen. Das heißt, 15% aller Modellergebnisse liegen unterhalb dieses Werts oder erreichen diesen genau. Die übrigen 85% der Modellsimulationen zeigen größere Änderungen. Entsprechend umgekehrt verhält es sich mit dem 85. Perzentil.

Zwischen das 15. und 85. Perzentil fallen somit insgesamt 70% aller Modellergebnisse.

Die Perzentile der verschiedenen Kennzahlen wurden für jedes Rasterelement ermittelt und flächendifferenziert für die Landesfläche von Baden-Württemberg dargestellt (Abbildung 4 rechts oben). Um die Vergleichbarkeit der Klimaprojektionen zu wahren, wurden für die flächenhaften Perzentildarstellungen die vier Klimaprojektionen in der 7x7km-Auflösung weggelassen und nur die Klimaprojektionen in 25x25km-Auflösung betrachtet. Da für die ferne Zukunft nicht alle Klimaprojektionen verfügbar sind, ergeben sich für Ist-Zustand und nahe Zukunft 20, für die ferne Zukunft 15 Klimaprojektionen.

Zudem wurden für die Analyse der Kennzahlen Box-Whisker-Diagramme mit den Flächenmittelwerten über Baden-Württemberg erzeugt (Abbildung 4 rechts Mitte, Legende unten):

Die farbigen Marker neben den Box-Whisker-Diagrammen beschreiben die jeweiligen Simulationsergebnisse der vorliegenden Klimaprojektionen.

Der rote Balken in den Box-Whisker-Diagrammen markiert den Median (50. Perzentil) der Klimaprojektionen.

Die Ausdehnung der Box gibt den Bereich des 1. Quartils (25. Perzentil) bis zum 3. Quartil (75. Perzentil) an und umfasst damit 50 % der Klimaprojektionen.

Der grün gestrichelte Bereich bis zu den zwei schwarzen Balken (Whisker) beschreibt den Bereich, in dem alle Klimaprojektionen liegen, die nicht als Ausreißer angesehen werden (1,5-facher Interquartilsabstand).

Werte von Klimaprojektionen außerhalb der Whisker gelten als Ausreißer.

Dabei wurden bei den Box-Whisker-Diagrammen für den Ist-Zustand und die nahe Zukunft 24 Klimaprojektionen und für die ferne Zukunft 15 Klimaprojektionen ausgewertet. Durch die unterschiedliche Anzahl von Klimaprojektionen bei den Perzentilen und bei den Box-Whisker-Diagrammen ist der Median einer Kennzahl nicht identisch (vgl. Abbildung 4: Median bei den Perzentilen 9,7 °C, bei den Box-Whisker-Diagrammen 9,6 °C).

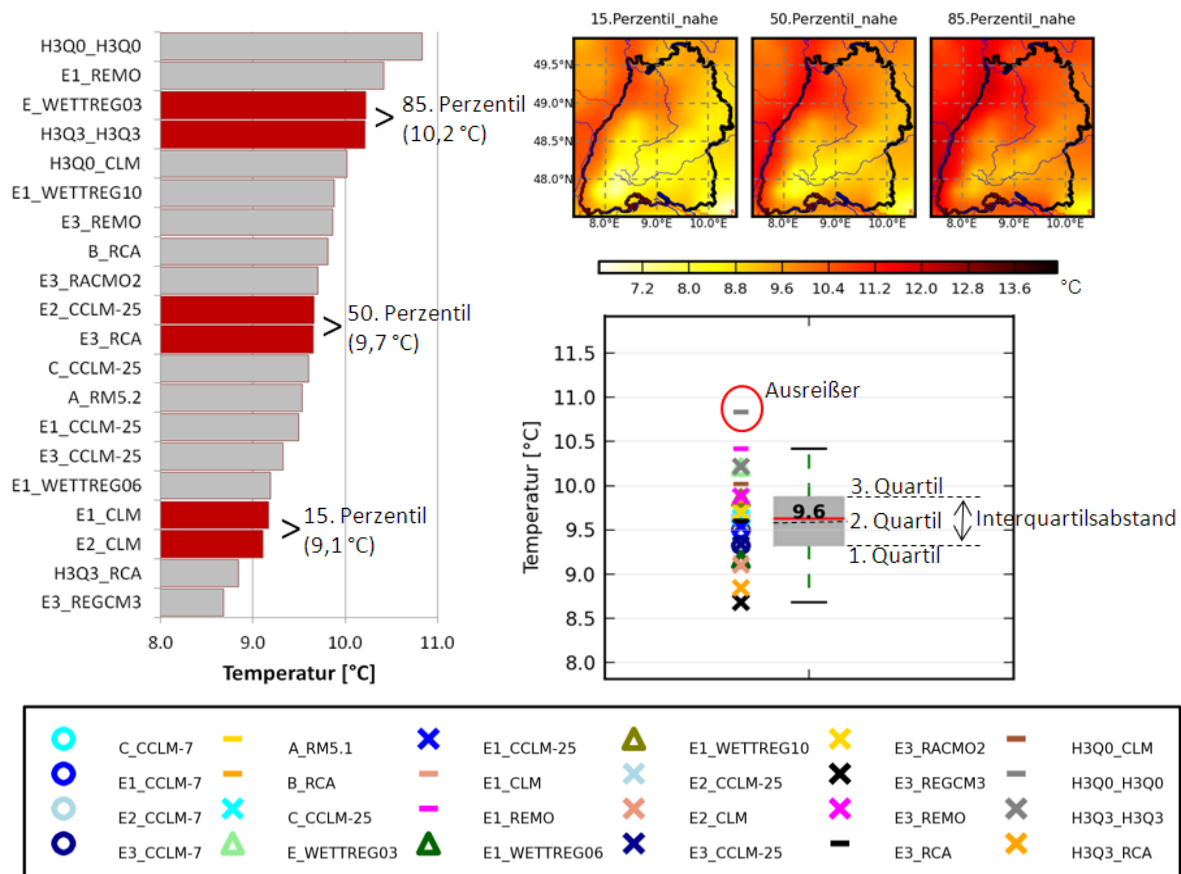


Abbildung 4: Exemplarische Erläuterung der Perzentile und der Box-Whisker-Diagramme am Beispiel der Jahreslufttemperatur der nahen Zukunft:

Perzentile: ermittelt aus den 20 Klimaprojektionen in der Auflösung 25x25km² (links: Perzentile am Beispiel der über die Fläche Baden-Württemberg gemittelten Temperatur; rechts oben: flächenhafte Darstellung der 15., 50. und 85. Perzentile):

Box-Whisker-Diagramme (Mitte rechts und unten): über die Fläche von Baden-Württemberg gemittelte Temperatur, ermittelt aus den 24 Klimaprojektionen in der Auflösung 25x25km² und 7x7km².

Eine direkte Übernahme der Werte für die zukünftigen „klimatischen Leitplanken“ aus den Simulationsergebnissen der Klimaprojektionen ist nicht empfehlenswert. Viele Modellprojektionen zeigen eine systematische Über- oder Unterschätzung bezüglich bestimmter Klimaparameter.

Stattdessen wird für eine bestmögliche Abschätzung einer zukünftigen Entwicklung einer Kennzahl das Delta-Change-Verfahren gewählt (Hay et al. 2000, Andréasson et al. 2004). Dabei werden die relativen Veränderungen der Werte zwischen der Simulation des Ist-Zustands und der Zukunftsszenarien auf die beobachteten Werte übertragen (vgl. Abbildung 5).

Ergibt sich z.B. als Median für die nahe Zukunft die Lufttemperatur von 9,6 °C im Vergleich zum Ist-Zustand mit 8,5 °C und somit eine Temperaturerhöhung um +1,1 °C, so wird dieser Wert als Klimaänderungssignal verstanden und auf den Beobachtungswert (hier: 8,4 °C)

aufaddiert, um die Temperatur der nahen Zukunft zu erhalten (somit 9,5 °C). Für die ferne Zukunft ergeben sich mit diesem Verfahren dann 11,4 °C.

Diese Kombination aus Beobachtungswert und Klimasignal stellt nur eine Näherung dar.

Voraussetzung für die Anwendung des Delta-Change-Verfahrens ist, dass die grundsätzlichen klimatischen Bedingungen des Ist-Zustands und der Zukunft vergleichbar bleiben.

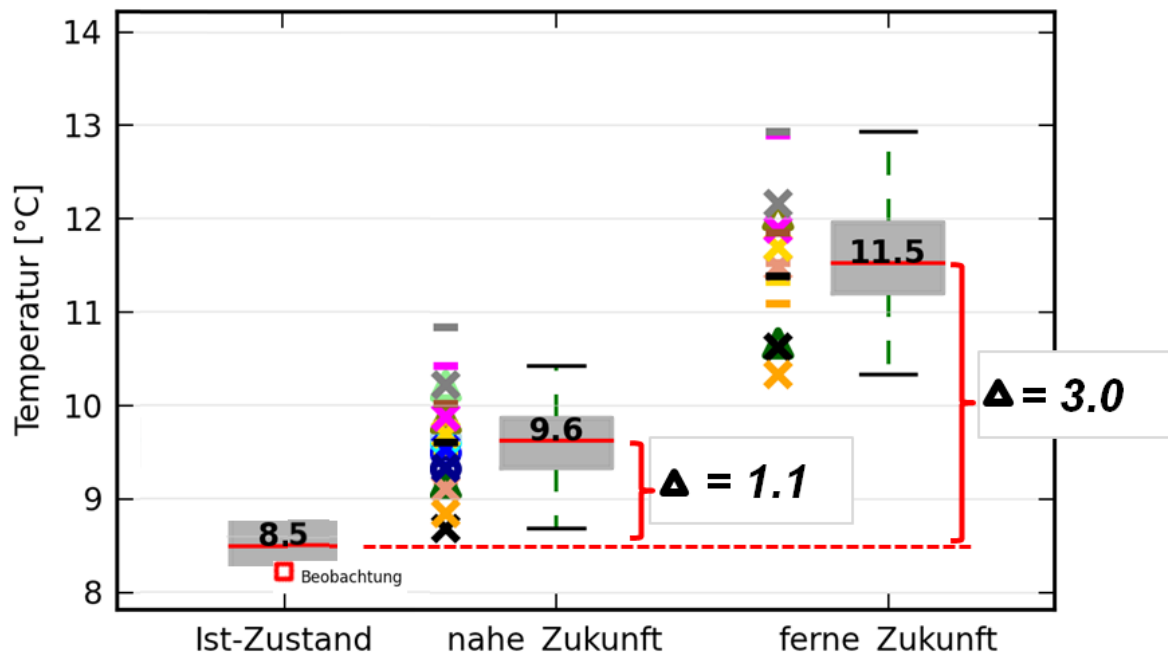


Abbildung 5: Schematische Darstellung der Ableitung des Klimasignals aus den Ergebnissen der Klimaprojektionen zur Übertragung auf die Beobachtungsdaten (Delta-Change-Verfahren).

Eine Übertragbarkeit des Klimasignals auf Beobachtungsdaten ist vor allem dann zu hinterfragen, wenn der Median der Simulationsergebnisse des Ist-Zustands weit vom Beobachtungswert abweicht oder die Unsicherheit der Modellergebnisse generell sehr hoch ist. Weitere Vor- und Nachteile des Verfahrens finden sich z.B. in Bosshard et al. (2011).

Da das Delta-Change-Verfahren robust und nachvollziehbar ist, wird es hier für die Übertragung des Klimasignals auf die Beobachtungsdaten umgesetzt. Auch die aktuellen Szenarien zur Klimaänderung in der Schweiz (CH2011) verwenden ein modifiziertes Delta-Change-Verfahren.

Eine detailliertere Beschreibung der umfangreichen Auswertung würde den Rahmen an dieser Stelle sprengen. Hier sei verwiesen auf den Bericht „Vergleich regionaler Klimaprojektionen für Baden-Württemberg“ (2012). Hinweise zum Umgang mit Klimaprojektionen und zur Auswertung finden sich auch bei Imbery F. et al. (2013) und Deutschländer et al. (2009, Homepage des DWD).

Für die hier dargestellte Studie wurden 28 Klimakennwerte ausgewertet (siehe Tabelle 4).

Tabelle 4: Ausgewertete Kennzahlen.

Temperatur	Niederschlag
Jahresmitteltemperatur	Niederschlagssumme im Jahr
Heiztag	Niederschlagssumme im hydrologischen Winterhalbjahr

Kühlgradtage	Niederschlagssumme im hydrologischen Sommerhalbjahr
Vegetationsbeginn	Anzahl der Tage ohne Niederschlag
Tropennacht	Anzahl der Tage mit Starkniederschlag
Frosttag	Starkniederschlags-Höhe
Datum letzter Frost	Niederschlagssumme im meteorologischen Frühjahr
Heißer Tag (Tropentag)	Niederschlagssumme in der Vegetationsperiode
Sommertag	Trockenperioden in der Vegetationsperiode
Eistag	Längste Dauer von Trockenperioden
Windgeschwindigkeit	Globalstrahlung
Mittlere Windgeschwindigkeit	Summe der Globalstrahlung/Jahr
Standardabweichung der Windgeschwindigkeit	Standardabweichung der Globalstrahlung
Maximale mittlere Windgeschwindigkeit	Minimale mittlere Globalstrahlung
Anzahl der Tage > 8 m/s Windgeschwindigkeit	Maximale mittlere Globalstrahlung

10.5 Anlage 5: Beispiel Obstbau (Pilzkrankheit: Apfelschorf)

Datengrundlage und Diagrammbeschreibung: Für die Tagesmitteltemperatur (schwarze durchgehende Kurve) wurde der Juli der Jahre 2021 bis 2050 (Standort Artern) ausgewählt. Die Werte entstammen dem Modell WETTREG2010. Bezüglich der Ensemblebildung wurden für das Szenario A1B alle zehn WETTREG2010-Realisierungen benutzt. Der Apfelschorf ist eine Pilzkrankheit, die stark von der Temperatur und der Blattnässedauer abhängig ist. Die Infektion mit Apfelschorf setzt nicht sofort ein, sondern benötigt eine gewisse Zeit mit vorhandenem Wasser auf den Blättern. Die gestrichelte Kurve zeigt die bis zur Infektion notwendige Blattnässedauer in Stunden in Abhängigkeit von der Temperatur. Die Berechnung erfolgte nach einer Formel von MacHardy und Gadoury (1989). Dabei ist zu beachten, dass für die Berechnung der Blattnässedauer die Tagesmitteltemperatur benutzt wurde, sodass im Einzelfall Abweichungen auftreten könnten.

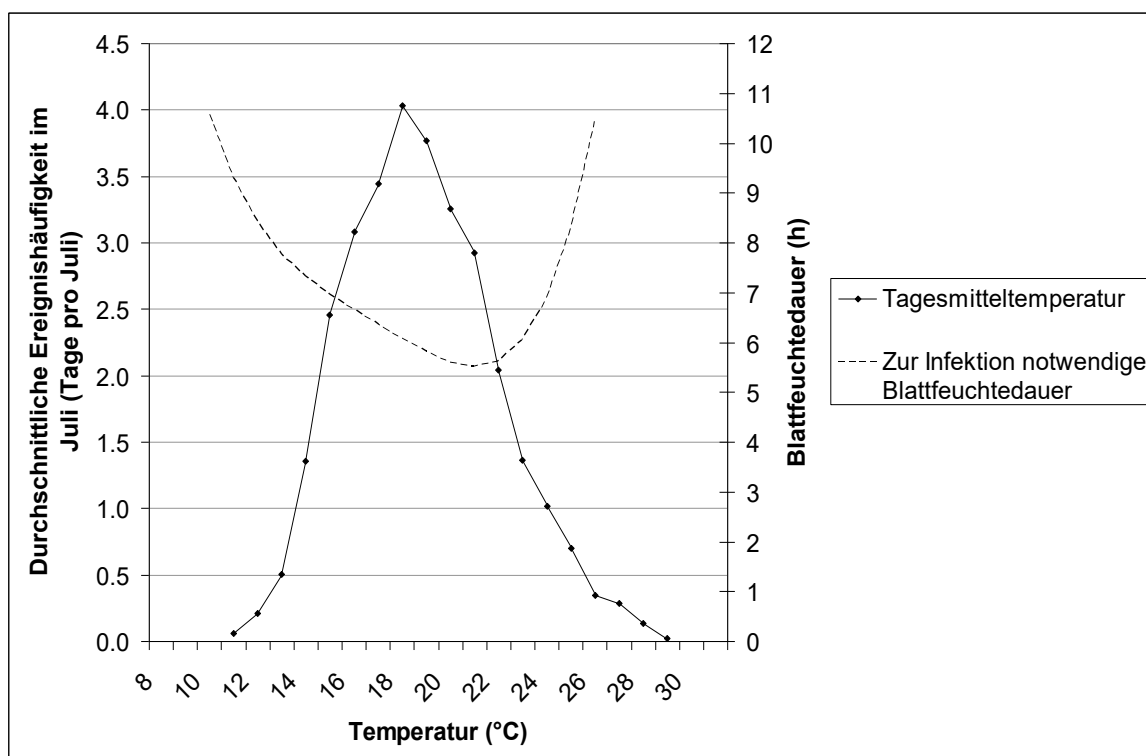


Abbildung 6: Zur Apfelschorfinfektion notwendige Blattfeuchtedauer in Abhängigkeit von der Temperatur und Häufigkeitsverteilungen der Tagesmitteltemperatur im Juli für den Zeitraum 2021 – 2050 am Beispiel des Standorts Artern (SRES A1B, ECHAM5, WETTREG2010).

Bei einer Temperatur von 21°C besteht das höchste Infektionsrisiko. Eine Blattnässedauer von ca. 5h 30min ist dabei bereits für das Auslösen der Infektion ausreichend. Im Gegensatz dazu wird bei über 26°C und unter 11°C eine Blattnässedauer von über 10h benötigt. Würde sich die Wahrscheinlichkeitsdichtefunktion der Temperatur in weiterer Zukunft abflachen (d.h. extreme Temperaturereignisse nehmen zu) und das Niederschlagsverhalten gleichzeitig unverändert bleiben, so ergäbe sich ein insgesamt geringeres Risiko einer Apfelschorfinfektion. Eine Verengung der Wahrscheinlichkeitsverteilung oder ein Versatz um plus 2 - 3°C führt zu einer Erhöhung des Infektionsrisikos.

10.6 Anlage 6: Beispiel Tourismus (Schwimmbadbesucher)

Datengrundlage und Diagrammbeschreibung: Für die Tagesmitteltemperatur (dunkelblaue Kurve) wurde der Juli der Jahre 2021 bis 2050 (Standort Artern) ausgewählt. Die Werte entstammen dem Modell WETTREG2010. Bezüglich der Ensemblebildung wurden für das Szenario A1B alle zehn WETTREG2010-Realisierungen benutzt. Bei der hellblauen Kurve handelt es sich um die identische Häufigkeitsverteilung wie für die dunkelblaue Kurve, nur ist diese um 2°C nach rechts verschoben. Dies steht hier symbolhaft für die Temperaturerhöhung bis 2100 bei sonst gleichbleibender Verteilung. Bei der Abhängigkeit der Schwimmbadbesucher von der Temperatur (orange Kurve) handelt es sich um ein fiktives Beispiel.

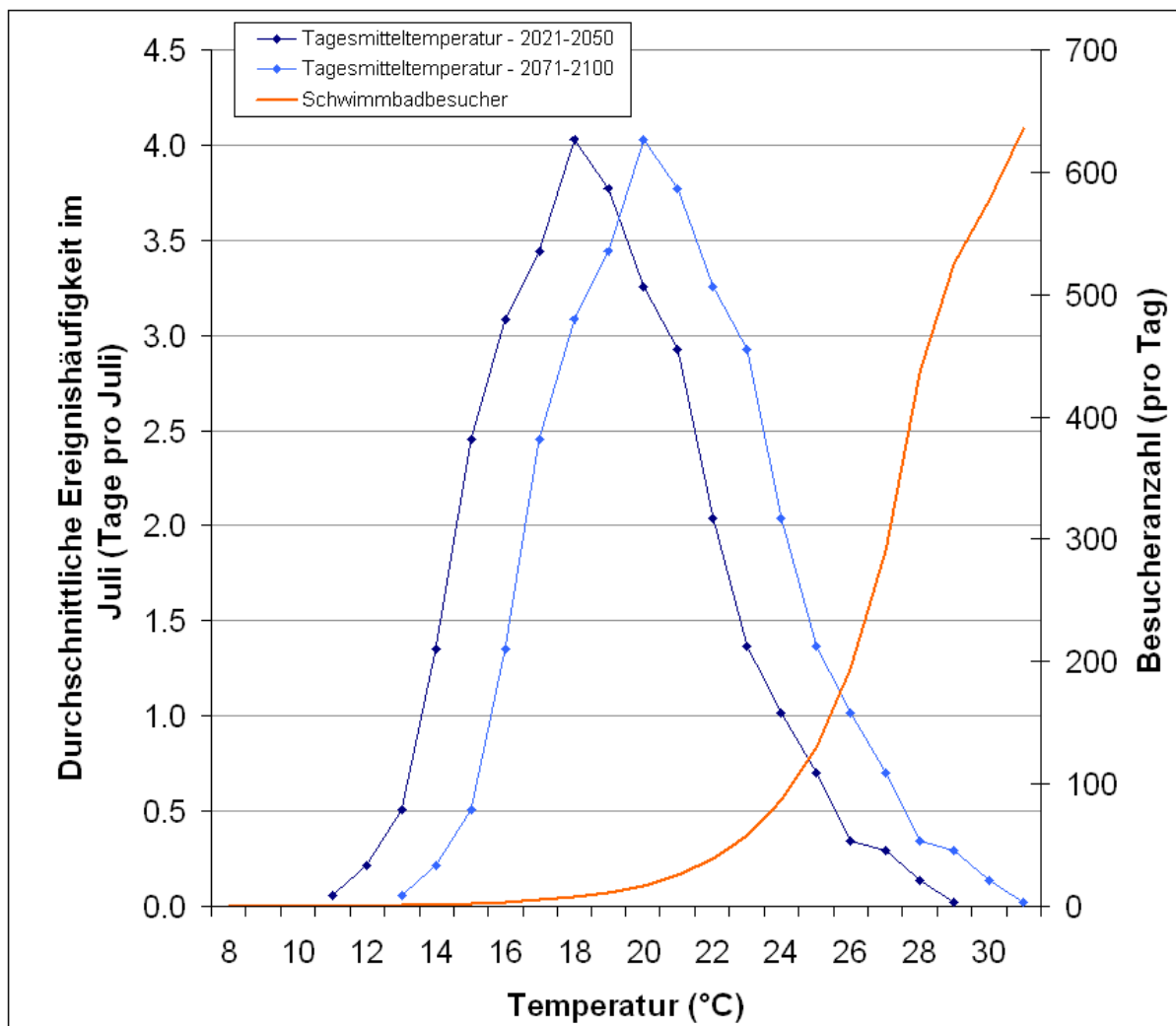


Abbildung 7: Durchschnittliche Schwimmbadbesucheranzahl je Tagesmitteltemperatur und Häufigkeitsverteilungen der Tagesmitteltemperatur im Juli für die Zeiträume 2021 – 2050 und 2071 – 2100 am Beispiel des Standorts Artern (SRES A1B, ECHAM5, WETTREG2010).

Der Schwimmbadbesitzer benötigt die in Zukunft zu erwartende Schwimmbadbesucheranzahl, um finanziell planen zu können. Dazu muss die Ereignishäufigkeit der

jeweiligen Temperatur mit der dazugehörigen Besucherzahl multipliziert werden. Die Aufsummierung aller Produkte mit anschließender Division mit 31 Julitagen ergibt die durchschnittlich zu erwartende Besucheranzahl pro Julitag. Für die dunkelblaue Kurve ergibt sich dabei ein Wert von 26 Besuchern pro Julitag. Dieselbe Rechnung mit der hellblauen Kurve (2°C höhere Durchschnittstemperatur) liefert einen Wert von 52 Besuchern pro Julitag. Der Schwimmbadbetreiber könnte demnach (durchschnittlich) mit einer Verdopplung seiner Einnahmen rechnen.

10.7 Anlage 7: Prüfung von Modellergebnissen auf Plausibilität

a) Prüfung anhand von Beobachtungsdaten (Evaluierung)

In diesem Beispiel soll eine Aussage darüber gemacht werden, mit welcher Güte zwei Klimamodelle (hier: REMO BfG und WETTREG Version2006; Szenario A1B; ECHAM5) innerhalb eines Referenzzeitraumes (hier: 1971 - 2000) die Niederschläge für ein ausgewähltes Untersuchungsgebiet für die hydrologische Modellierung (Einzugsgebiet von Aller und Leine in Niedersachsen, Größe rund 15.000 km²) wiedergeben. Als Vergleichsdatensatz dienen Stationsmessungen des Deutschen Wetterdienstes.

Die zeitlich differenzierte Prüfung der Klimamodelldaten wird zunächst anhand der Jahresniederschlagssummen durchgeführt. Hierzu werden als erstes gemittelte Gebietsreihen des Niederschlags aus den Zeitreihen der Gitterboxen (REMO) bzw. der Stationen (WETTREG) für das Einzugsgebiet gebildet. Aus diesen werden die Jahressummen berechnet und diese der Größe nach sortiert dargestellt (Abbildung 8). Außerdem werden die gemittelten Gebietsreihen der Modelle in Form von Monatsmitteln ins Verhältnis zu den Beobachtungsdaten gesetzt (Abbildung 9).

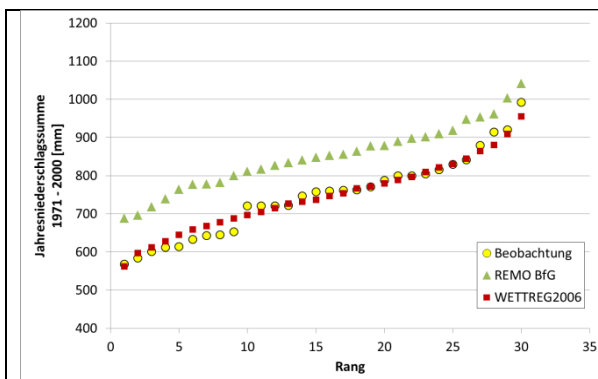


Abbildung 8: Vergleich der nach Größe sortierten Jahressummen des Niederschlags im Aller-Leine-Gebiet (1971 – 2000)

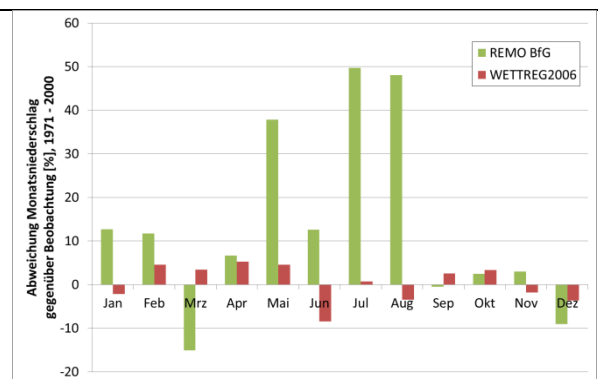
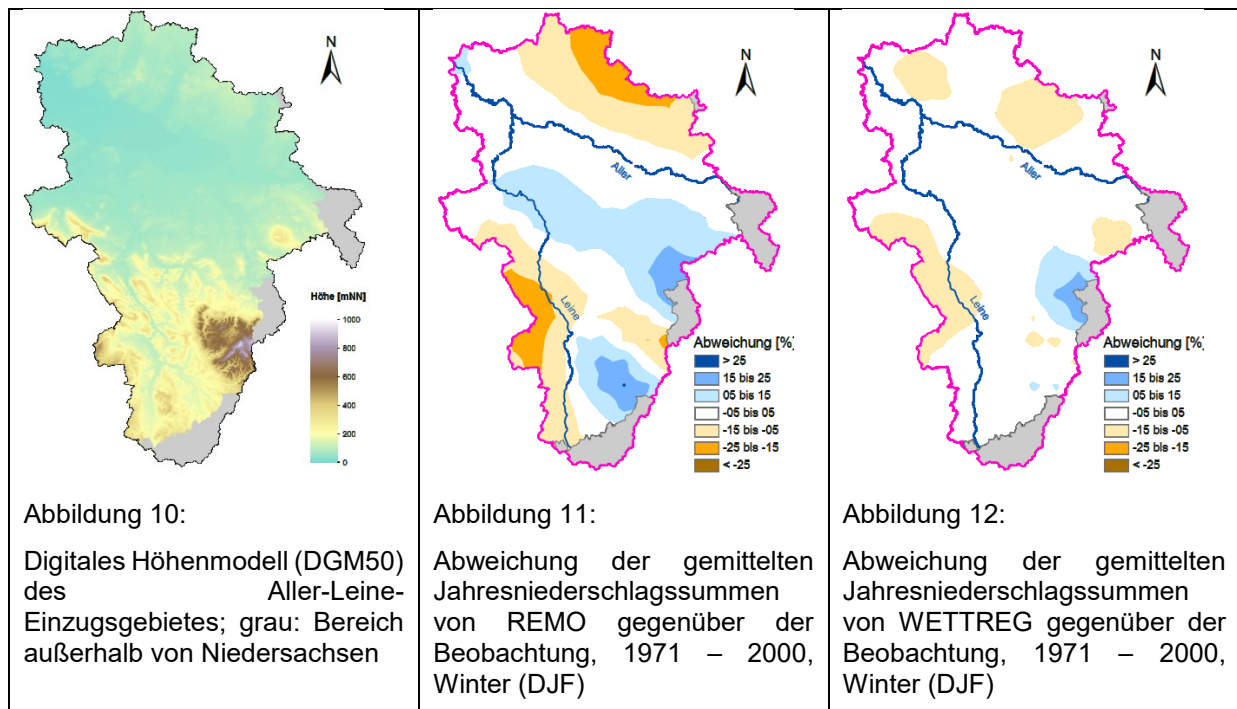


Abbildung 9: Vergleich der Abweichungen der mittleren Monatsniederschläge von der Beobachtung im Aller-Leine-Gebiet (1971 – 2000)

Es zeigt sich, dass REMO die Jahresniederschlagssummen grundsätzlich überschätzt, wobei „nasse“ Jahre noch am Nächsten an der Beobachtung liegen (rechte Seite von Abbildung 8). WETTREG gibt die Jahressummen insgesamt relativ gut wieder. Die differenzierte Betrachtung der Monatsmittel macht deutlich, dass die gesamtheitliche Überschätzung von REMO (Abbildung 8) vor allem auf den Abweichungen in den Sommermonaten beruht (Abbildung 9, zwischen Juni und August von +10 bis +50 %). Daher sollten Absolutwerte von Wirkmodellen (z. B. hydrologische Modelle), die diese Daten als Eingabe nutzen, für den genannten Zeitraum mit Vorsicht interpretiert werden bzw. nochmals gesondert geprüft werden. Die Abweichungen von WETTREG liegen in allen Monaten unter 10 %, was grundsätzlich ein guter Wert ist, jedoch keine Garantie darstellt, dass die Werte der Szenariensimulation (bis 2100) mit diesem Modell näher an der Wirklichkeit der Zukunft liegen müssen.

Die räumlich differenzierte Prüfung der Daten erfolgt anhand der auf ein 10 x 10 km Raster interpolierten Beobachtungs- bzw. Klimamodelldaten, gemittelt über den Referenzzeitraum

1971 - 2000. Die Daten von REMO bzw. WETTREG werden wiederum in das Verhältnis zu den Beobachtungswerten gesetzt (vgl. Abbildung 11 und Abbildung 12). Als zusätzliche Information wird hierbei noch das digitale Geländemodell herangezogen (Abbildung 10).



Der Vergleich zeigt, dass die Abweichungen von REMO in den Wintermonaten (DJF), die sich auch in Abbildung 9 gezeigt haben (etwa -10 bis +10 % gegenüber der Beobachtung), auch räumlich noch differenziert bewertet werden können. Während der Bereich südlich der Erhebungen des Harzes sowie im südlichen Aller-Flachland hinsichtlich des Niederschlags überschätzt wird, erscheinen das Leinebergland und die Lüneburger Heide zu trocken. Die entsprechenden Abweichungen von WETTREG tauchen nur lokal auf und sind aufgrund ihrer geringen Ausprägung (zumeist < 10 %) kaum signifikant, da sie vom Betrag her kleiner sind als die modellinterne Variabilität, d. h. den Abweichungen innerhalb dieses Zeitraums, die sich allein zwischen den Realisationen des Modells ergeben.

Insgesamt betrachtet erscheinen die Daten des mittleren Niederschlags von WETTREG für den Zeitraum 1971 - 2000 plausibel, während REMO BfG hier sowohl bzgl. der Werte als auch der räumlichen Strukturen leichte bis zum Teil erheblichen Schwächen aufweist. Ein Rückschluss auf die Qualität der Daten in den Zukunftsszenarien kann hieraus jedoch nicht zwingend erfolgen.

b) Prüfung der zukünftigen Signale (Plausibilitätskontrolle)

Es soll überprüft werden, ob eine modifizierte Variante des REMO-Modells plausible Daten für das Zukunftsszenario A1B innerhalb eines ausgewählten Einzugsgebietes (Aller-Leine Gebiet in Niedersachsen) produziert.

Zunächst werden dazu die Daten zeitlich differenziert anhand einiger ausgewählter Klimagrößen stichprobenhaft analysiert. In diesem Fall sind es der mittlere Jahresniederschlag sowie die mittlere Sonnenscheindauer und der Bedeckungsgrad. Alle Größen werden chronologisch (von 1960 bis 2100) in Diagrammen aufgetragen, der Niederschlag für sich alleine (Abbildung 13), die Sonnenscheindauer und der Bedeckungsgrad zusammen (Abbildung 14), da diese beiden Größen grundsätzlich miteinander zusammenhängen.

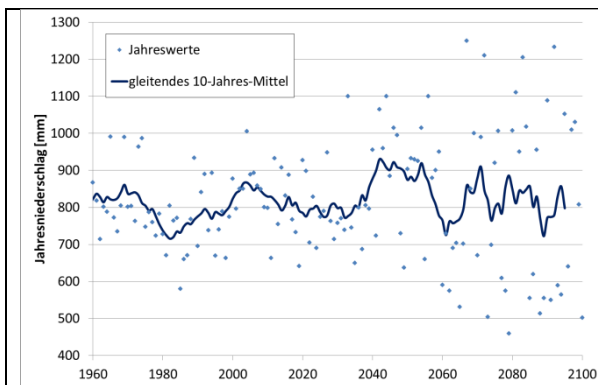


Abbildung 13: Jahresniederschlag REMO (modifiziert), Szenario A1B, Aller-Leine Gebiet

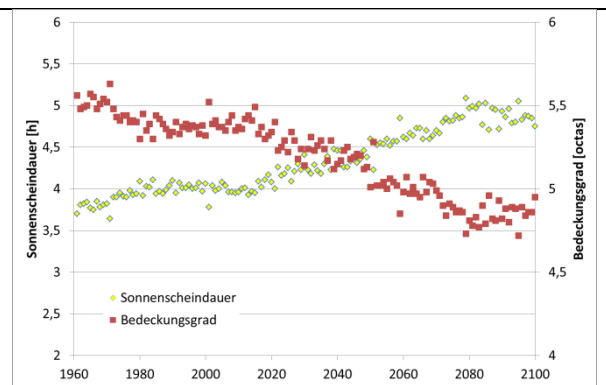
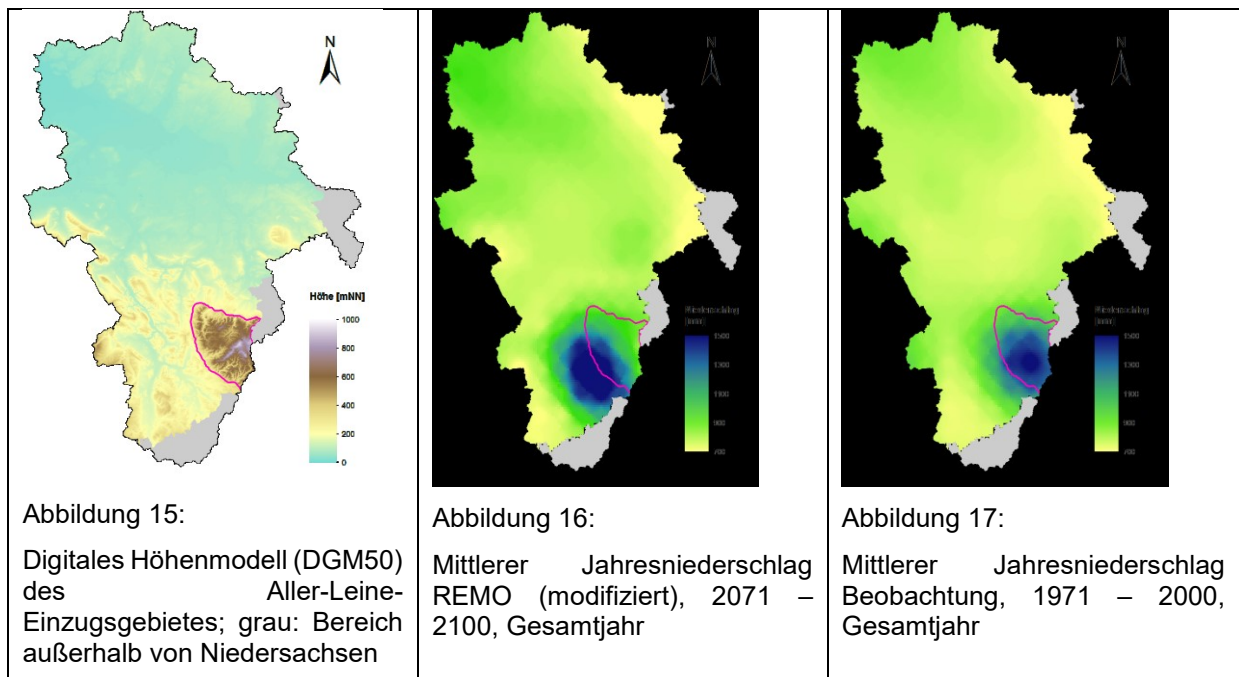


Abbildung 14: Sonnenscheindauer und Bedeckungsgrad REMO (modifiziert), Szenario A1B, Aller-Leine Gebiet

Das gleitende 10jährige Mittel der Niederschläge (Abbildung 13) zeigt einen schwankenden Verlauf der Niederschlagsmengen bis 2100 ohne eindeutigen Trend. Die Werte variieren, bewegen sich aber im Mittel grundsätzlich zwischen 700 und 900 mm pro Jahr, einem für diesen Raum durchaus nachvollziehbaren Wertebereich. Betrachtet man jedoch die einzelnen Jahreswerte, so ist zu erkennen, dass die Bandbreite der Werte zunimmt (bis zum Jahr 2100 auf ca. 500 bis 1250 mm). Dieses sind nach wie vor mögliche Werte für nasse bzw. trockene Jahre. Es stellt sich hier jedoch die Frage, in wie weit es „realistisch“ erscheint, dass es zukünftig (ab ca. 2040) in diesem Gebiet scheinbar nur noch sehr nasse bzw. trockene Jahre gibt, aber praktisch kaum noch durchschnittliche Verhältnisse. Zudem weist der ständige Wechsel der Extreme von einem Jahr auf das andere auf klimatische Zustände hin, die für ein Gebiet in mittleren Breiten mit bisher eher gemäßigttem Klima ungewöhnlich erscheinen, da sie bisher so nicht beobachtet wurden.

Die Gegenüberstellung der Klimagrößen Sonnenscheindauer und Bedeckungsgrad zeigt dagegen ein durchaus plausibles Bild. Während die Sonnenscheindauer zukünftig zunimmt, nimmt der Bedeckungsgrad in etwa in gleichem Maße ab, was logisch erscheint. Auch die Absolutwerte bewegen sich in „realistischen“ Bahnen. Ergänzend könnte z. B. noch der Temperaturverlauf herangezogen werden, da dieser bei zunehmender Sonnenscheindauer i. d. R. auch zunehmen müsste.

Die räumlich differenzierte Überprüfung erfolgt anhand der mittleren Jahresniederschlagsmengen über den Zeitraum 2071 – 2100. Diese werden anhand eines 10 x 10 km Rasters für das Gebiet flächenhaft dargestellt (Abbildung 16). Ein Niederschlagsschwerpunkt scheint im Bereich der Erhebungen des Harzes zu liegen, was durchaus plausibel wäre. Bei genauerer Betrachtung fällt jedoch auf, dass dieser Schwerpunkt bei dem REMO-Modell im südwestlichen Harzvorland liegt. Dies könnte auf den ersten Blick noch damit begründet werden, dass zukünftig die Hauptanströmrichtung der Niederschlagsgebiete von Südwesten erfolgt und die Luftmassen an den Ausläufern des Harzes aufsteigen, abkühlen und entsprechend bereits im Vorland abregnen. Diese Vermutung erscheint jedoch aufgrund der nicht ausreichenden Höhe des Harzes als unplausibel. Ein Vergleich mit heutigen Verhältnissen (Abbildung 17) zeigt, dass sich ein solches Bild bisher nicht ergeben hat bzw. der Niederschlagsschwerpunkt leicht verschoben in den Hochlagen des Harzes liegt. Als weitere Entscheidungsunterstützung könnten noch die Ergebnisse anderer Modelle herangezogen werden oder es sollte geprüft werden, ob die zukünftigen Großwetterlagen eine solche Hauptanströmrichtung mit entsprechenden Niederschlagsverhältnissen vertretbar erscheinen lassen.



Insgesamt erscheint die Beurteilung bzw. Kontrolle auf Plausibilität von zukünftigen Verhältnissen aufgrund der fehlenden Vergleichsdaten i. d. R. schwieriger als die Evaluierung von Klimamodelldaten anhand von Beobachtungsdaten in der Vergangenheit. Eine Bewertung von Klimamodelldaten im Zukunftszeitraum erfolgt in erster Linie über die Einschätzung der physikalischen Plausibilität der Modelldaten. Hilfsweise können hierbei ggf. auch Erfahrungswerte mit Beobachtungsdaten in der Vergangenheit herangezogen werden, wobei eine Abweichung der Modelldaten aufgrund der veränderten Randbedingungen in der Zukunft innerhalb gewisser Grenzen durchaus möglich sein kann.

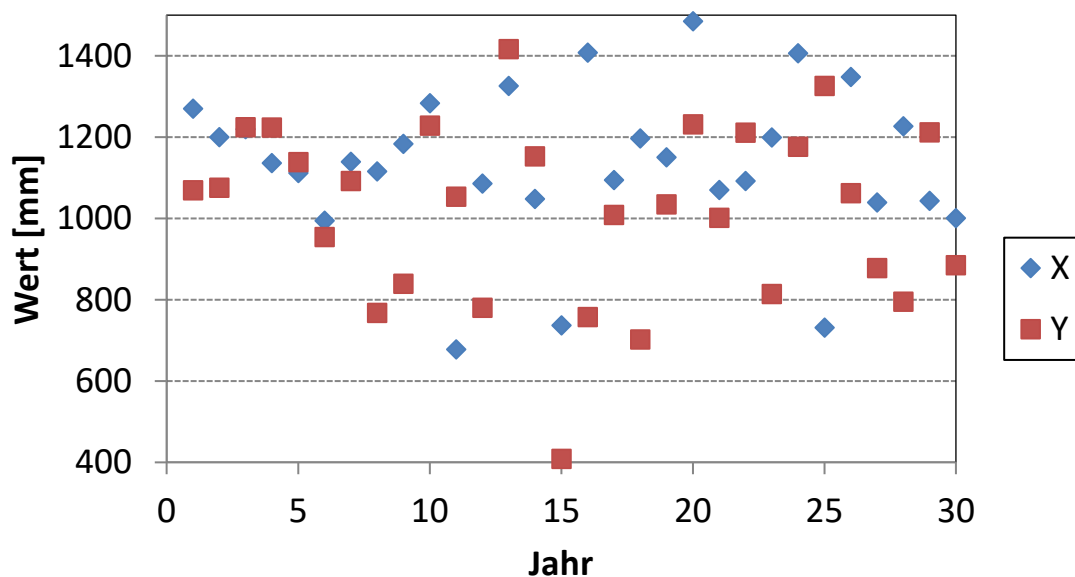
10.8 Anlage 8: Berechnung einer Bemessungsschwelle für signifikante Änderungen

Die Prüfung, ob ein berechneter Änderungswert eine Signifikante Änderung darstellt, erfolgt im Rahmen eines mehrstufigen Prozesses. Voraussetzung dafür ist die **Annahme der Normalverteilung**. Diese kann im Fall der Temperatur als erfüllt angesehen werden. Für andere meteorologische Parameter ist gegebenenfalls eine Transformation durchzuführen, z.B. Monatswerte des Niederschlags oder die Anzahl feuchter Tage.

Beispiel:

Die Beschreibung der Tests erfolgt auf der Basis der folgenden Daten (Einheit mm):

Jahr	1	2	3	4	5	6	7	8	9	10
X	1270,1	1200,0	1219,3	1136,5	1111,7	994,6	1139,1	1115,5	1183,4	1283,3
Y	1069,3	1075,3	1224,5	1223,4	1138,6	954,3	1092,0	767,9	839,4	1228,2
Jahr	11	12	13	14	15	16	17	18	19	20
X	678,2	1085,8	1326,3	1048,2	737,0	1407,6	1094,1	1196,7	1150,2	1485,3
Y	1053,4	779,9	1416,2	1152,7	408,7	757,5	1008,9	702,4	1034,7	1231,8
Jahr	21	22	23	24	25	26	27	28	29	30
X	1070,2	1092,0	1198,8	1406,3	731,4	1347,7	1039,7	1226,8	1043,1	1001,0
Y	1001,3	1211,2	814,0	1176,4	1326,2	1062,0	878,0	795,3	1212,0	885,0



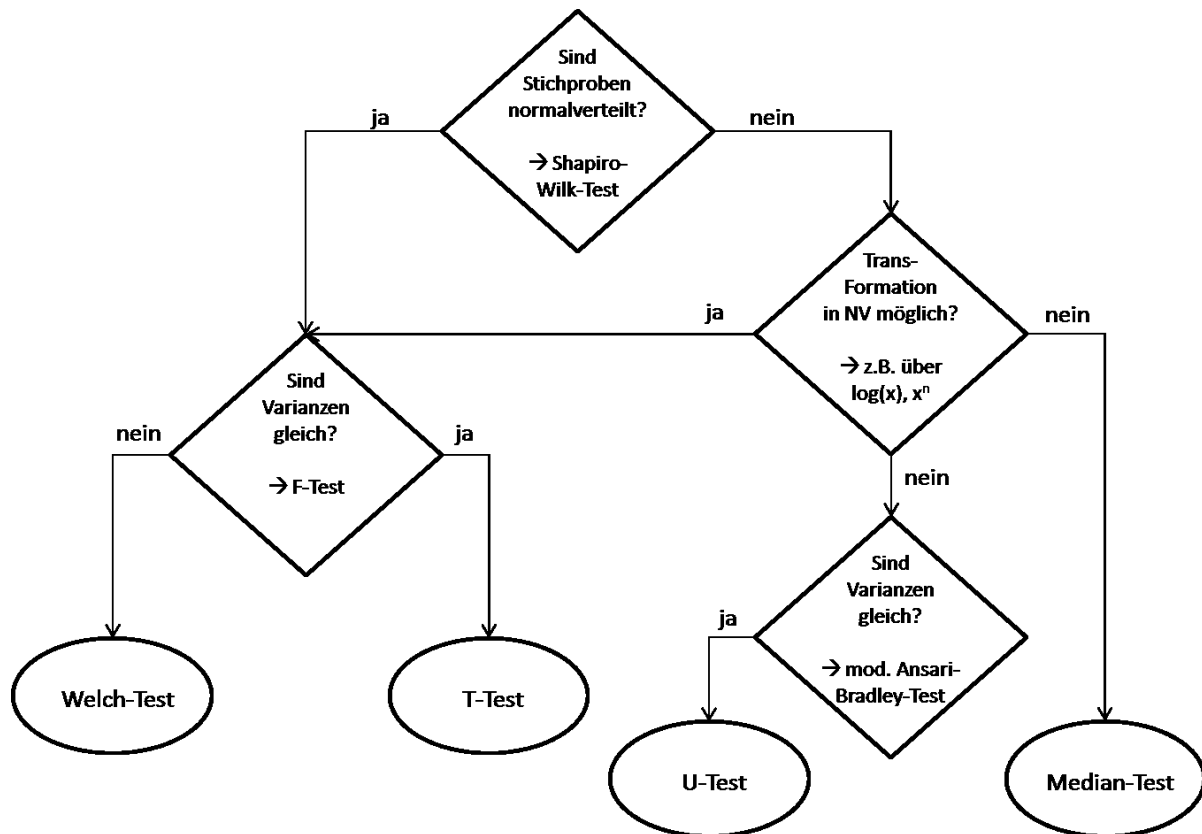
Statistische Werte:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 1134,0 \text{ mm}; \quad s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = 35402$$

$$\bar{y} = \frac{1}{m} \sum_{i=1}^m y_i = 1017,4 \text{ mm}; \quad s_y^2 = \frac{1}{m-1} \sum_{i=1}^m (y_i - \bar{y})^2 = 48229$$

$$\bar{y} - \bar{x} = -116,6 \text{ mm } (-10,29 \%)$$

In Abhängigkeit des Ergebnisses sind weitere Tests notwendig. Das Flussdiagramm gibt eine Übersicht über die weitere Verfahrensweise.



Test auf signifikanten Unterschied der Varianzen: F-Test

Vor der Anwendung des eigentlichen Tests auf Signifikanz des Änderungssignals ist zu prüfen, ob die Varianzen des Referenz- und des Projektionszeitraumes sich unterscheiden. Die Prüfung erfolgt auf Basis des **F-Testes** zur Hypothese $H_0: s_x^2 = s_y^2$ gegen $H_1: s_x^2 \neq s_y^2$.

$$F = \frac{s_x^2}{s_y^2}$$

wobei s_x^2 und s_y^2 die Varianzen im Referenz- bzw. Projektionszeitraum sind ($X_1 \dots X_n$ und $Y_1 \dots Y_m$). Für ein α von 5% und einen Stichprobenumfang von jeweils 30 (Jahres- oder Saisonwerten), liegt der kritische Bereich des Testes bei $\leq 0,482$ bzw. $\geq 2,074$ (die Quantile der F-Verteilung $F_{(n-1, m-1; 1-\alpha/2)}$ sind tabelliert und werden außerdem von den gängigen statistischen Programm-Paketen berechnet). Fällt der Quotient in den Annahmebereich:

$$0,482 \leq \frac{s_x^2}{s_y^2} \leq 2,074$$

dann kann der t-Test genutzt werden. Ansonsten muss der Welch-Test, der sich nur wenig vom t-Test unterscheidet, eingesetzt werden.

Beispiel:

Im Beispiel beträgt der Wert:

$$\frac{s_x^2}{s_y^2} = \frac{35402}{48229} = 0,734$$

Die beiden Varianzen unterscheiden sich also **nicht signifikant**. Damit kann für das Beispiel der t-Test eingesetzt werden.

Der F-Test hat jedoch abgesehen von der Entscheidung t-Test/Welch-Test auch noch weitergehende Bedeutung bei der Auswertung von Klimadaten. Unabhängig davon, ob sich der Mittelwert eines gewählten meteorologischen Parameters ändert, kann eine geänderte Variabilität ebenso wichtig sein. Läge der Wert der Statistik F im Beispiel über 2,074, so wäre die Variabilität von $X_1 \dots X_n$ signifikant größer als die Variabilität von $Y_1 \dots Y_n$.

Test auf signifikanten Unterschied der Mittelwerte bei gleicher Varianz: t-Test

Unterscheiden sich beide Varianzen nicht wesentlich, dann wird zur Prüfung des Änderungssignals $\bar{x} - \bar{y}$ (Mittelwerte des Referenz- und des Projektionszeitraumes) der **t-Test für unabhängige Stichproben** durchgeführt mit $H_0: \bar{x} = \bar{y}$ und $H_1: \bar{x} \neq \bar{y}$.

$$t = \frac{\sqrt{\frac{n \cdot m}{n + m}} \cdot \frac{\bar{x} - \bar{y}}{s}}$$

mit

$$s^2 = \frac{(n - 1)s_x^2 + (m - 1)s_y^2}{n + m - 2}$$

Wobei n und m die Stichprobenumfänge des Referenz- und des Projektionszeitraumes sind. Der Annahmebereich (Bereich nicht-signifikanter Änderung) für den Mittelwert des Projektionszeitraumes berechnet sich aus.

$$\bar{x} \pm t_{(1-\alpha/2; n+m-2)} \cdot \frac{s}{\sqrt{\frac{n \cdot m}{n + m}}}$$

Für den hier vorliegenden Fall von jeweils 30 Werten und bei einer Wahl von $\alpha=0,05$ (zweiseitig) hat das entsprechende Quantil $t_{1-\alpha/2; n+m-2}$ den Wert $t_{0,975; 58} \approx 2,00$. (Diese Werte sind ebenfalls tabelliert bzw. werden von statistischen programm-Paketen berechnet.)

$$\bar{x} \pm 2 \cdot \frac{s}{\sqrt{\frac{30 \cdot 30}{30 + 30}}} = \bar{x} \pm 0,516 \cdot s$$

Liegt der Änderungswert des Mittelwertes innerhalb der Schranken:

$$-0,516 \cdot s \leq \Delta \bar{x} \leq 0,516 \cdot s$$

so ist die Änderung **nicht signifikant**.

Beispiel:

Im Beispiel beträgt der s^2 Wert:

$$s^2 = \frac{(n-1) \cdot s_x^2 + (m-1) \cdot s_y^2}{n+m-2} = \frac{(30-1) \cdot 35402 + (30-1) \cdot 48229}{n+m-2} = 41815$$

$$s = 204,49$$

$$0,516 \cdot s = 105,52$$

Das bedeutet, dass bei dem Beispiel ab einer Änderung 105,52 mm die Änderung signifikant ist. Die vorhandene Änderung beträgt 116,6 mm und ist damit **signifikant**.

Test auf signifikanten Unterschied der Mittelwerte bei unterschiedlicher Varianz: Welch-Test

Unterscheiden sich die Varianzen des Referenz- und des Projektionszeitraumes wesentlich, so muss an Stelle des t-Testes der **Welch-Test** genutzt werden. Die Teststatistik t wird identisch zum t-Test berechnet.

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}}$$

Der kritische Bereich des Welch-Testes wird ebenfalls durch die t-Verteilung approximiert, jedoch mit einer modifizierten Anzahl von Freiheitsgraden v :

$$v = \frac{(n-1)(m-1) \left(\frac{s_x^2}{n} + \frac{s_y^2}{m} \right)^2}{(n-1) \left(\frac{s_x^2}{n} \right)^2 + (m-1) \left(\frac{s_y^2}{m} \right)^2}$$

Entsprechend Taubenheim 1969 S. 94 lassen sich die 5% Schrankenwerte des zweiseitigen Vertrauensbereiches mit folgender Näherungsformel darstellen:

$$t_{(0,025;v)} \approx 1,96 + \frac{2,44}{v-1} \quad (\text{für } v \geq 3)$$

Liegt der berechnete Testwert t innerhalb des Bereiches

$$-t_{(0,025;v)} \leq t \leq t_{(0,025;v)}$$

ist die Änderung **nicht signifikant**.

Beispiel:

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} = \frac{1134,0 - 1017,4}{\sqrt{\frac{35402}{30} + \frac{48229}{30}}} = 2,2093$$

$$v = \frac{(n-1)(m-1) \left(\frac{s_x^2}{n} + \frac{s_y^2}{m} \right)^2}{(n-1) \left(\frac{s_x^2}{n} \right)^2 + (m-1) \left(\frac{s_y^2}{m} \right)^2} = \frac{(30-1)(30-1) \left(\frac{35402}{30} + \frac{48229}{30} \right)^2}{(30-1) \left(\frac{35402}{30} \right)^2 + (30-1) \left(\frac{48229}{30} \right)^2}$$

$$v = 56,667$$

$$t_{(0,025;v)} \approx 1,96 + \frac{2,44}{v-1} = 1,96 + \frac{2,44}{56,667-1} = 2,0038$$

Für das gewählte Beispiel wird die Testbedingung

$$-t_{(0,025;v)} \leq t \leq t_{(0,025;v)}$$

nicht erfüllt. Aus der Perspektive des Welch-Testes ist damit die Änderung **signifikant**.

Alle bisher genannten Tests sind nur unter Voraussetzung der Normalverteilung gültig. Dabei ist der F-Test besonders unrobust gegen Abweichungen von der Normalverteilung, was sich in einer hohen Rate von falsch-positiven Testergebnissen niederschlägt (Test nimmt fälschlich die Hypothese der identischen Variabilität an). Der t- und der Welch-Test sind dagegen deutlich robuster. Solange die Abweichung von der Normalverteilung nicht dramatisch ist, sind deren Testergebnisse durchaus zuverlässig. Falls also Zweifel an der Gültigkeit des F-Testes bestehen, weil die Verteilung der zu betrachtenden meteorologischen Variable nur ungefähr der Normalverteilung entspricht, so könnte immer noch der weniger empfindliche Welch-Test angewendet werden, der für beliebige Variabilität gilt.

Nichtparametrischer Test auf signifikanten Unterschied der Mittelwerte: Median-Test

Sollte es nicht möglich sein, eine geeignete Transformation zu finden, die die Verteilung eines gewählten meteorologischen Parameters in eine (annähernde) Normalverteilung umformt, so sind die bisher genannten Tests ungültig. Alternativ käme der **Median-Test** in Frage. Da dieser nichtparametrische Test aber eine geringe Teststärke besitzt, kann sich ein erhöhter Aufwand bei der Transformation durchaus lohnen.

Zur Berechnung des Tests wird zunächst der gemeinsame Median der kombinierten Stichproben $m_{(X,Y)}$ ermittelt. Die Werte werden dann nach Stichprobenzugehörigkeit und danach, ob sie größer oder kleiner als der gemeinsame Median sind, in 4 Gruppen geordnet, die eine 4-Felder-Kontingenztafel bilden. Diese Kontingenztafel kann dann mit einem Chi-Quadrat-Test ausgewertet werden.

	Stichprobe X	Stichprobe Y	Σ
Element $< m_{(X,Y)}$	a	b	a+b
Element $\geq m_{(X,Y)}$	c	d	c+d
Σ	a+c=n	b+d=m	n+m

Die Teststatistik:

$$\chi^2 = \frac{n(a \cdot d - b \cdot c)^2}{(a+b)(c+d)(a+c)(b+d)}$$

hat eine annähernde Chi-Quadrat-Verteilung mit Freiheitsgrad 1, sodass bei $\alpha=0,05$ die Änderung nicht signifikant ist, wenn

$$\chi^2 < 3,84.$$

Beispiel:

Der Median beider Wertegruppen beträgt 1093,1 mm.

	Stichprobe X	Stichprobe Y	Σ
Element $< m_{(X,Y)}$	a = 11	b = 19	a+b = 30
Element $\geq m_{(X,Y)}$	c = 19	d = 11	c+d = 30
Σ	a+c=n = 30	b+d=m = 30	n+m = 60

$$\chi^2 = \frac{(n+m)(a \cdot d - b \cdot c)^2}{(a+b)(c+d)(a+c)(b+d)} = \frac{(30+30)(11 \cdot 11 - 19 \cdot 19)^2}{(11+19)(19+11)(11+19)(19+11)} = 4,27$$

Somit zeigt auch der Median-Test, dass die Unterschiede zwischen X und Y **signifikant** sind.

Soll der Mann-Whitney-Wilcoxon-Test zum Nachweis signifikanter Änderungssignale verwendet werden, ist zu beachten, dass sich die Verteilungen der beiden Stichproben gemäß der für diesen Test geltenden Voraussetzungen nur durch eine Verschiebung des Mittelwertes voneinander unterscheiden dürfen. Unterscheidet sich zusätzlich auch deren Variabilität, so kann es zur fälschlichen Ablehnung der Hypothese $H_0: \bar{x} = \bar{y}$ kommen, das heißt der Trend wird als signifikant bewertet, obwohl in Wirklichkeit nur die Varianz verschieden ist.

Nichtparametrischer Test auf signifikanten Unterschied der Varianzen: Modifizierter Ansari-Bradley-Test nach Fligner und Killeen (1976)

Der **Modifizierte Ansari-Bradley-Test** ist eine nichtparametrische Alternative zum F-Test. Wie auch der Median-Test hat er als verteilungsfreier Test eine geringere Teststärke als die parametrischen Tests. Er ist deshalb nicht zum Test der Voraussetzungen des t-Tests geeignet, da der t-Test selbst nur unter Normalverteilung gilt. Jedoch kann ein Test der Hypothese $H_0: s_x^2 = s_y^2$ auch unabhängig von einer Verschiebung des Mittelwertes von Interesse sein. In diesem Fall, bei Nichterfüllung der Normalitätsannahme, sollte der Modifizierte Ansari-Bradley-Test zum Einsatz kommen.

Zur Berechnung wird zunächst in beiden Stichproben (Referenz- und Projektionszeitraum) der Median gebildet und die Stichprobe damit zentriert.

$$\tilde{X}_i = X_i - \text{med } X, \quad \tilde{Y}_j = Y_j - \text{med } Y$$

Die Absolutbeträge der $\tilde{X}_i$ und $\tilde{Y}_j$ werden gemeinsam der Größe nach geordnet, wobei R_i die Ränge der $|\tilde{X}_i|$ bezeichnet. Nur die Ränge der $|\tilde{X}_i|$ werden nun aufsummiert.

$$T = \sum_{i=1}^n R_i$$

Unter der Hypothese besitzt diese Teststatistik die Wilcoxon-Verteilung (Quantile des Wilcoxon-Rangsummen-Tests). Für große Stichprobenumfänge ($n, m \geq 20$) kann die Verteilung durch eine Normalverteilung $\mathcal{N}(n(n+m+1)/2, n \cdot m(n+m+1)/12)$ approximiert werden. Für $\alpha=0,05$ und $n=m=30$ ist die Abweichung der Variabilität also nicht signifikant, falls

$$\frac{30 \cdot 61}{2} - 1,96 \cdot \sqrt{\frac{30 \cdot 30 \cdot 61}{12}} \leq T \leq \frac{30 \cdot 61}{2} + 1,96 \cdot \sqrt{\frac{30 \cdot 30 \cdot 61}{12}}$$

bzw.

$$915 - 1,96 \cdot 67,64 \leq T \leq 915 + 1,96 \cdot 67,64$$

Beispiel:

Die Summe der $|\tilde{X}_i|$ Ränge beträgt 786.

$$915 - 1,96 \cdot 67,64 \leq T \leq 915 + 1,96 \cdot 67,64$$

$$782,43 \leq 786 \leq 1047,57$$

Somit zeigt der Test, dass die Abweichung der Variabilität zwischen X und Y **nicht signifikant** sind.

Test auf signifikanten Unterschied zwischen zwei Wahrscheinlichkeiten: Binomial-Test

Speziell für das Testen von dichotomen Variablen (z.B. Niederschlag ja/nein bzw. die Anzahl trockener/nasser Tage) wird der **Binomial-Test** verwendet. Dabei wird der Anteil von Elementen mit/ohne Merkmal in beiden Stichproben ermittelt (p_X, p_Y) und mit dem Anteil in der kombinierten Stichprobe $p = (p_X \cdot n + p_Y \cdot m)/(n + m)$ verglichen:

$$b = \frac{p_X - p_Y}{\sqrt{p(1-p)(1/n + 1/m)}}$$

Für große n und m wird das Quantil durch eine Standard-Normalverteilung approximiert, für $\alpha=0,05$: $z_{1-\alpha/2} = 1,96$. Falls also

$$-1,96 < b < 1,96$$

so ist die Änderung nicht signifikant.

10.9. Anlage 9: Beispiel zur Identifikation von robusten Klimaänderungen

Im Zuge des Projektes „Globaler Klimawandel – Wasserwirtschaftliche Folgenabschätzung für das Binnenland“ (KliBiW; NLWKN, 2017) erfolgte die Abschätzung der Robustheit von Klimaänderungssignalen auf Basis der Ansätze von Pfeifer et al. (2015) und Tebaldi et al. (2011). Ziel der Analysen in diesem Zusammenhang war die Beantwortung der Frage, wie unsicher bzw. verlässlich das mittlere Signal eines Modell-Ensembles in einer Rasterkartendarstellung ist, da hier eine Darstellung von Bandbreiten zur Verdeutlichung der Unsicherheiten nicht sinnvoll erschien. Nach dem Schema in Abbildung 18 wurden dabei drei Einstufungen der Robustheit³ unterschieden. Die Einschätzung der Signifikanz erfolgte dabei anhand der Methode in Anlage 8. Die Schwellwerte sind projektbezogen, können für andere Anwendungen auch neu definiert werden und dienen in den Leitlinien der Veranschaulichung der Methode.

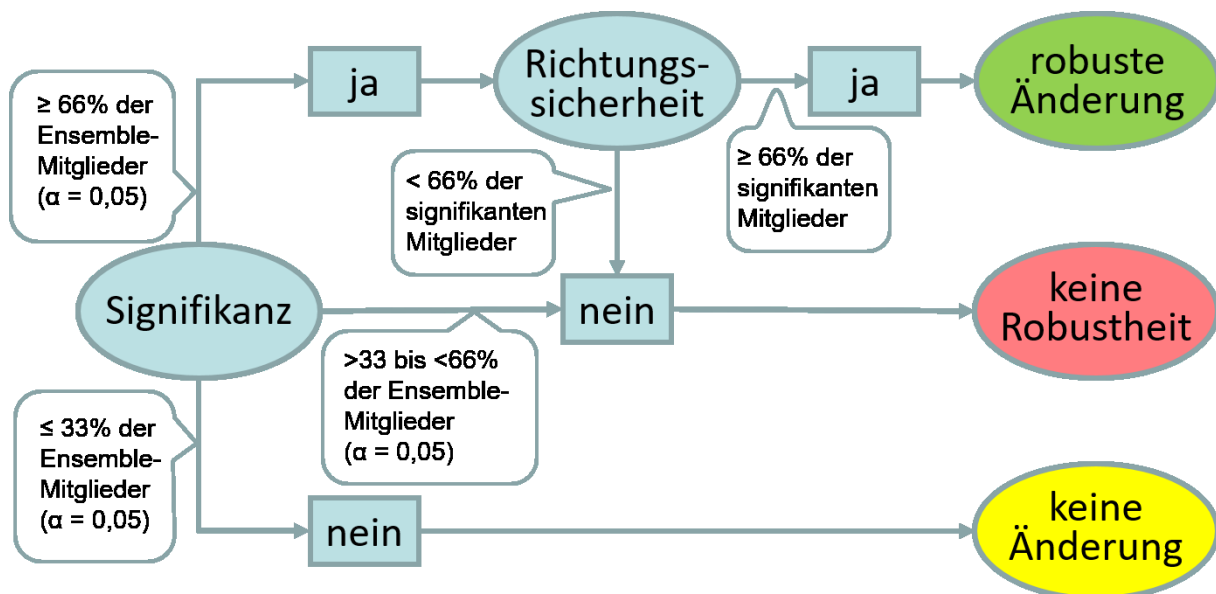


Abbildung 18: Ablaufschema zur Einschätzung der Robustheit von Klimaänderungssignalen im Projekt KliBiW

Fall 1 Robuste Zu- oder Abnahme einer Klimagröße (robuste Änderung): Mindestens 66 % der Mitglieder des Klimamodellensembles zeigen eine statistisch signifikante Änderung und von allen statistisch signifikanten Signalen haben mindestens 66 % dasselbe Vorzeichen.

³ Der Begriff Robustheit wird in diesem Zusammenhang benutzt, um einen gewissen (ausreichenden) Grad der Übereinstimmung in Bezug auf Änderungssignal eines Ensembles aus beliebig vielen (ausgewählten) Mitgliedern auszudrücken; siehe auch Abschnitt 4.5

- Fall 2 Keine Robustheit:
- Mindestens 66 % der Mitglieder des Klimamodellensembles zeigen eine statistisch signifikante Änderung, von allen statistisch signifikanten Signalen haben aber weniger als 66 % dasselbe Vorzeichen.
 - Von allen Mitgliedern des Klimamodellensembles zeigen weniger als 66 %, aber mehr als 33 % eine statistisch signifikante Änderung.
- Fall 3 Keine Änderung:
- Mindestens 66 % der Mitglieder des Klimamodellensembles zeigen eine statistisch *nicht* signifikante Änderung.

In einem ersten Schritt wurde die Signifikanz der Einzelsignale der jeweiligen Ensemble-Mitglieder analysiert. Für jede Rasterzelle der Modelldaten wurde dabei die Stichprobe der 30 Jahreswerte für den Zukunftszeitraum (z.B. 2071-2100) den 30 Jahreswerten des Referenzzeitraumes (z.B. 1971-2000) anhand der Methode in Anlage 8 gegenübergestellt. Wenn mindestens 66% der Ensemble-Mitglieder demnach eine signifikante Abweichung gegenüber dem Referenzzeitraum aufwiesen (bei einer Irrtumswahrscheinlichkeit von $\alpha = 0,05$), wurde zusätzlich noch die Richtungssicherheit des Ensembles betrachtet. Dieses Vorgehen wurde für jede Rasterzelle der einzelnen Ensemble-Mitglieder durchgeführt. Das Ergebnis ist eine Kartendarstellung des mittleren Ensemble-Signals in Kombination mit der Aussage zu dessen Robustheit (vgl. Abbildung 19).

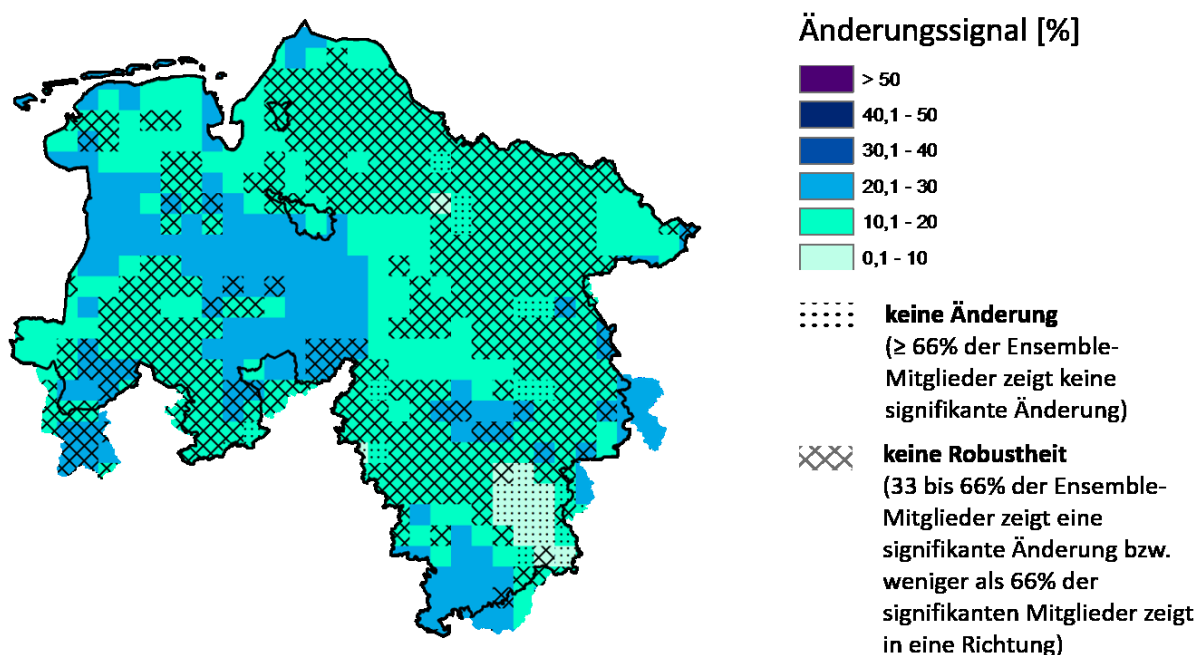


Abbildung 19: Darstellung des Änderungssignals und seiner Robustheit am Beispiel des 99%-Quantils der Tagesniederschläge in den Wintermonaten (Dezember, Januar, Februar) der fernen Zukunft (2071-2100) in Niedersachsen gegenüber dem Referenzzeitraum (1971-2000)

Die Karte zeigt, dass die Regenmengen bei großen Niederschlagsereignissen (99%-Quantil der Tagesniederschläge) in allen Teilen Niedersachsens rein rechnerisch in den Wintermonaten im Mittel ansteigen werden. In einigen (gepunkteten) Bereichen sind diese Veränderungen jedoch bei einer Mehrzahl der Einzelsignale des Ensembles nicht signifikant, d.h. die meisten Modelle sagen, die Veränderungen bewegen sich im Bereich der natürlichen Klimavariabilität und sind somit keine Änderungssignale im engeren Sinn. In größeren (gekreuzten) Bereichen zeigt ein gewisser Anteil der Ensemble-Mitglieder ein signifikantes Signal. Dieser Anteil ist zu groß, um von keiner Änderung zu sprechen. Er ist aber auch gleichzeitig zu klein, um von einer robusten Änderung zu sprechen. Daher kann für diese Bereiche keine verlässliche Aussage bzgl. der zukünftigen Veränderung gemacht werden. Nur in den Bereichen ohne Schraffur sind sich ausreichend viele Modelle des Ensembles bzgl. der Signifikanz und der Richtung des Signals einig.

Für eine schärfere Aussagekraft können die Schwellenwerte für die jeweilige Einstufung (anhand einer definierten Mehrheit, hier 66 %) angehoben werden. Dies würde bedeuten, dass die Sicherheit einer Veränderung in den Bereichen ohne Schraffur größer würde. Gleichzeitig kann es jedoch auch dazu führen, dass weniger Bereiche als robust identifiziert werden.

10.10. Anlage10: Ein Kernensemble von regionalen Klimaprojektionen

Mit steigender Anzahl der zur Verfügung stehenden regionalen Klimaprojektionen sieht sich ein Großteil der Wirkmodellierer und Nutzer immer weniger in der Lage, den Empfehlungen der Klimaforschung folgend als Basis für zukunftswirksame Entscheidungen Simulationsrechnungen für das gesamte Ensemble durchzuführen. Vor dem Hintergrund dieser Problematik wird zunehmend die Frage gestellt, ob und ggf. wie ein großes Ensemble auf eine deutlich kleinere und damit handhabbarere Anzahl reduziert werden kann, ohne dass wichtige Informationen (wie z.B. die Bandbreite der Ergebnisse) über das zukünftige Klima verloren gehen.

Das Bund-Länder-Fachgespräch hat sich 2018 deshalb auf folgende „Anforderungen an ein Kernensemble verständigt:

- sollte alle Anforderungen der Qualitätskontrolle und Qualitätssicherung erfüllen,
- möglichst breit nutzbar (viele Anwendungsmöglichkeiten),
- sollte die Bandbreite des plausibilisierten Referenzensembles abdecken,
- basiert auf einem Referenzensemble eines festen Stichtags und
- **das Kernensemble ist eine Empfehlung** (d.h. Abweichungen für einzelne Anwendungen werden nicht ausgeschlossen).

Anforderung an die Qualitätssicherung:

- Simulationsergebnisse müssen auf regionaler Skala (z.B. Naturräume) für den Zeitraum (1971-2000) plausibel/realitätsnah sein. Für das nächste Kernensemble soll diese Plausibilisierung methodisch diskutiert und für ganz Deutschland ausgewertet werden.
- Jedes Ensemblemitglied sollte das gesamte ReKliEs-De Gebiet abdecken (Deutschland und die nach Deutschland entwässernden Flusseinzugsgebiete, mit Ausnahme der Oder),
- Alle erforderlichen Parameter (pr, tas, tasmin, tasmax, hurs, huss, rsds, sund, sfcWind, ps, psl) sollen vorliegen.
- Alle Parameter sollten qualitätskontrolliert sein (mind. QC mit DKRZ-Tool) um unrealistische Werte (z.B. negativer Sonnenschein) auszuschließen.
- Simulationen sollten auf grundlegende Ungereimtheiten (z.B. konstante THG-Konzentrationen) geprüft werden. (Info sollte in Metadaten enthalten sein)
- Transparente und offene Dokumentation von „Unzulänglichkeiten“ oder „Fehlern“ in Modellläufen

Das Kernensemble muss nachvollziehbar bestimmt werden. Darüber hinaus kann empfohlen werden, dass noch weitere Modellläufe hinzugenommen werden können. Mit der vom DWD entwickelten Methode soll aus den verbleibenden Klimamodellläufen ein Kernensemble bestimmt werden.“ [Protokollauszug Fachgespräch]

Ausgangspunkt ReKliEs-De:

Im Bundesforschungsvorhaben ReKliEs-De wurden als Ergänzung des EORO-CORDEX-Projekts die Klimaänderungssignale zur nahen Zukunft (2021-2050) und zum Ende des 21. Jahrhunderts (2071-2100) im Vergleich zur Referenzperiode 1971-2000 für das ReKliEs-De-Gebiet untersucht.

Das EURO-CORDEX/ReKliEs-De Ensemble stellt durch seine Größe, die hohe räumliche Auflösung und die Kombination aus statistischen und dynamischen Regionalisierungsverfahren eine weltweit einzigartige Datenbasis für die Erforschung des Klimawandels sowie der Abschätzung von dessen Folgen dar. Die Bandbreite der Klimaänderungssignale für das RCP8.5 und das RCP2.6 wird durch die in ReKliEs-De zusätzlich erzeugten Simulationen für viele Größen erweitert. Die zentralen Aussagen vorheriger Studien zur Temperaturänderung sowie zur Niederschlagsänderung im Sommer und Winter bleiben erhalten. Mehr dazu unter: <http://reklies.hlnug.de>

Das RCP8.5-Ensemble besteht aus 37 Klimaprojektionen. Diese Ensemblegröße stellt in einigen Fällen eine erhebliche, wenn nicht gar überfordernde Vielfalt für die Impactmodellierung dar. Es ist davon auszugehen, dass jede Projektion, sofern nicht physikalisch widerlegt, eine plausible Klimazukunft darstellt, und die Einschränkung der Spannbreite des Ensembles einen Informationsverlust bewirkt. Aus pragmatischen, nicht aus wissenschaftlichen Gründen, erscheint deshalb eine Ensemblereduzierung notwendig. Diese sollte nach klar definierten Parametern erfolgen. Im Ergebnis sollte ein Kernensemble entstehen, welches mindestens für weiterführende Modellierungen und Untersuchungen anzuwenden ist.

Reduktion in zwei Schritten:

Die DWD-Reduktionsmethode wurde nicht direkt auf das ReKliEs-De-Ensemble angewandt. Diesem Prozessschritt vorgeschaltet waren Qualitätsprüfungen seitens des DWD und das bayerische Klimaaudit.⁴

Auf der Homepage des DWD wird eine Liste mit allen verfügbaren Modellläufen aus dem EURO-CORDEX Projekt und dem ReKliEs-DE Projekts geführt.⁵ Alle Modellläufe, die vom DWD aus diesem Referenzensemble ausgeschlossen wurden, sind gekennzeichnet und es findet sich eine Begründung für den Ausschluss. Weitere Ausschlüsse erfolgten in Bezug auf die o.g. Anforderungen an ein Kernensemble bzw. die Qualitätssicherung. So fehlen beispielsweise beim Regionalmodell WETTREG die Flussgebiete die nach Deutschland entwässern. Somit wird das ReKliEs-De-Gebiet (Deutschland und die nach Deutschland entwässernden Flusseinzugsgebiete) nicht vollständig abgedeckt.

⁴ „Bayerisches Landesamt für Umwelt [Hrsg.] (2020): Das Bayerische Klimaprojektionsensemble - Audit und Ensemblebildung. UmweltSpezial. Augsburg: 55S.

⁵ https://www.dwd.de/DE/klimaumwelt/klimaforschung/klimaprojektionen/fuer_deutschland/liste-klimaprojektionen.html

Das bayerische Klimaaudit bezog sich daraus resultierend auf die verbleibenden 21 Ensemblemitglieder als aktuelles Referenzensemble. Die Qualitätsprüfung erfolgte anhand von vier Indikatoren (absolute Abweichung der mittleren Jahreswerte und der mittleren Monatswerte, die Kling-Gupta-Effizienz der Monatswerte sowie dem Bestimmtheitsmaß der Einzelrastrer) und einem darauf aufbauenden Punktesystem. Diese Analyse wurde für vier Regionen (Alpen, Alpenvorland, Mittelgebirge, Sichtstufenland) vorgenommen. Im Ergebnis blieben 10 Simulationen als Referenzempfehlung des BLFG übrig.⁶

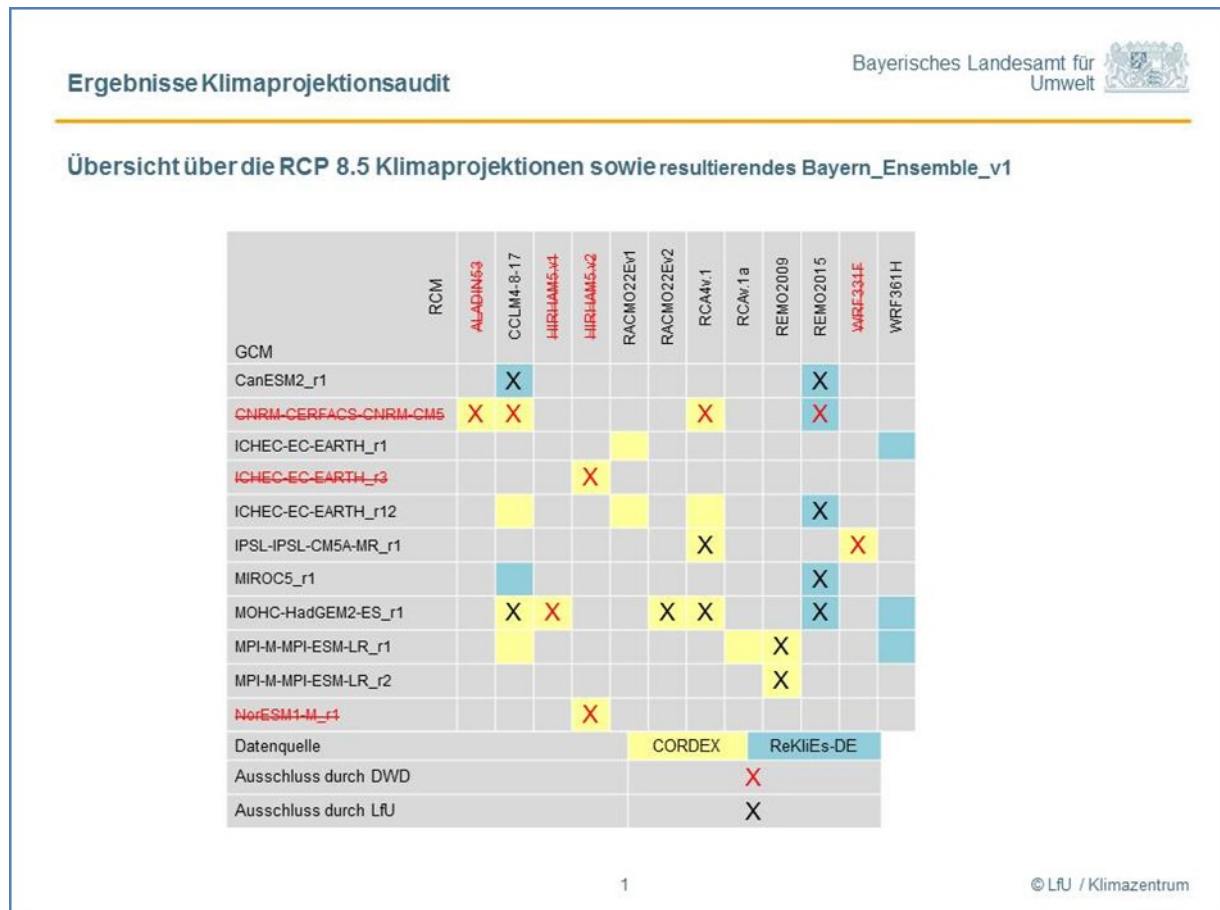


Abb.10.1 Ergebnis Klimaprojektionsaudit; LfU Bayern

"Das Ziel der Methode ist, diejenigen Simulationen auszuwählen, die die stärksten Änderungen (in Temperatur, Niederschlag, Wind) zeigen. Alle Simulationen, die Änderungssignale zwischen diesen Extremen liegen, werden nicht weiter berücksichtigt. Insbesondere wenn zwei Modelle ähnliche Ergebnisse zeigen wird nur eines davon für das Kernensemble zugelassen." [DWD, Dr. C. Dalelane]

⁶ „BLFG-Referenzensemble“, „Bayerisches Referenzensemble“ oder „10er Ensemble Bay“ sind identisch.

Das Verfahren arbeitet multivariat und berücksichtigt die wichtigsten klimatologischen Parameter einschließlich jahreszeitlicher und regionaler Schwankungen in mehreren Zukunftsperioden, wobei die Dimensionalität des Problems zunächst durch eine Hauptkomponentenanalyse reduziert wird. Der Abstand der Modelle untereinander wird aus den Werten der führenden Hauptkomponenten ermittelt, woraus wiederum unter Verwendung einer nichtlinearen Transformation die paarweise Ähnlichkeit der Projektionen berechnet wird. Als Selektionsgewicht dient die Inverse der effektiven Wiederholung einer Projektion im Ensemble (Summe der paarweisen Ähnlichkeiten mit allen anderen Projektionen=Redundanz), wobei das Ensemble schrittweise durch Ausschluss derjenigen Projektion verkleinert wird, die die höchste Redundanz innerhalb der verbliebenen Projektionen aufweist.

Die Anwendung des Verfahrens auf das aktuelle CORDEX EUR-11 Ensemble zum Treibhausgasszenario RCP8.5 wurde im Journal of Applied Meteorology and Climatology veröffentlicht (<http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-17-0141.1>).

Zur Erstellung des deutschen Kernensembles zum Szenario RCP8.5, wurden die jeweiligen Projektionen aus CORDEX EUR11 und ReKliEs-De, die bis zum Stichtag 31.05. 2017 vorlagen, und den Qualitätskriterien des bayerischen Klimaaudits entsprachen, verwendet. Die Parameter Mittlere Temperatur, Niederschlagssumme, Anteil des Niederschlags an sehr nassen Tagen, Autokorrelationskoeffizient des Tagesniederschlags, und Mittlere Windgeschwindigkeit in monatlicher und regionaler (acht Klima-Regionen innerhalb Deutschlands) Auflösung bzgl. naher und ferner Zukunft berücksichtigt. Die Qualitätsgewichtung wurde nicht angewandt, da alle Projektionen einer multivariaten Biasadjustierung unterzogen werden, die den Bias bzgl. der hier verwendeten Maße weitgehend eliminiert. Vorgabe war, dass die Variabilität des Ursprungsensembles zu 90 % erhalten bleibt (siehe Tabelle 10-1 Abdeckung des Variabilitätserhalts).

Abdeckung des Variabilitätserhalts im Kernensemble

	Nahe Zukunft 2021-2050	Ferne Zukunft 2071-2100
Monatliche Mitteltemperatur	0,97	0,98
Monatliche Niederschlagssumme	0,95	0,89
Monatlicher Index für Extremniederschlag	0,74	0,77
Monatlicher Persistenz-Index für Niederschlag	0,88	0,88
Monatliche mittlere Windgeschwindigkeit	0,95	0,93
gesamt	0,90	0,89

Tab.10-1 Ergebnis zum Variabilitätserhalt (DWD-Reduktionsmethode)

Ergebnis Ensemblereduzierung

Nachdem also im ersten Schritt alle fehlerhaften oder unplausiblen Simulationen entfernt wurden (Qualitätskontrolle Bayern), wurden im zweiten Schritt aus den verbleibenden Simulationen diejenigen ausgewählt, deren Änderungssignale in den Parametern Temperatur, Niederschlag und Wind die Bandbreite des Referenzensembles bestmöglich widerspiegeln. Diese ausgewählten Simulationen ergeben folgendes Kernensemble:

Liste der Klimaprojektionen nach der Reduktion: Kernensemble2020 zu RCP8.5:

Globalmodell	Regionalmodell
ICHEC-EC-EARTH_r1	RACMO22Ev1
ICHEC-EC-EARTH_r12	RACMO22Ev1
MIROC5_r1	CLM4-8-17
MOHC-HadGEM2-ES_r1	WRF361H
MPI-M-MPI-ESM-LR_r1	CLM4-8-17
MPI-M-MPI-ESM-LR_r1	WRF361H

Tab.10.2 Kernensemble2020 zu RCP8.5

Für die Regionen, in denen die Flusseinzugsgebiete außerhalb Deutschlands eine untergeordnete Rolle spielen, kann dieses Ensemble durch zwei statistische Regionalisierungen ergänzt werden:

Erweiterung des Kernensembles2020 um zwei Wettreg-Läufe: 8er-Ensemble zu RCP8.5:

Globalmodell	Regionalmodell
ICHEC-EC-EARTH_r1	RACMO22Ev1
ICHEC-EC-EARTH_r12	RACMO22Ev1
MIROC5_r1	CLM4-8-17
MOHC-HadGEM2-ES_r1	WRF361H
MPI-M-MPI-ESM-LR_r1	CLM4-8-17
MPI-M-MPI-ESM-LR_r1	WRF361H
MOHC-HadGEM2-ES_r1	WETTREG2018
MPI-M-MPI-ESM-LR_r1	WETTREG2018

Tab.10.3 Erweitertes Kernensemble2020 (8er-Ensemble)

Beispiel:

Nachstehend einige Parametervergleiche für die Region Brandenburg-Berlin in Abhängigkeit der Ensemblesnutzung (1971-2000 zu 2071-2100 für RCP8.5). Deutlich wird, dass die Bandbreite der Ergebnisse vom 10er-Ensemble des bayerischen Klimaaudits im Wesentlichen im Kernensemble erhalten bleibt. Abweichungen entstehen durch die Hinzunahme der Wettreg-Läufe zu dem 8er-Ensemble. Deren Ergebnisse liegen anteilig außerhalb des Referenzensembles (10er-Ensemble).

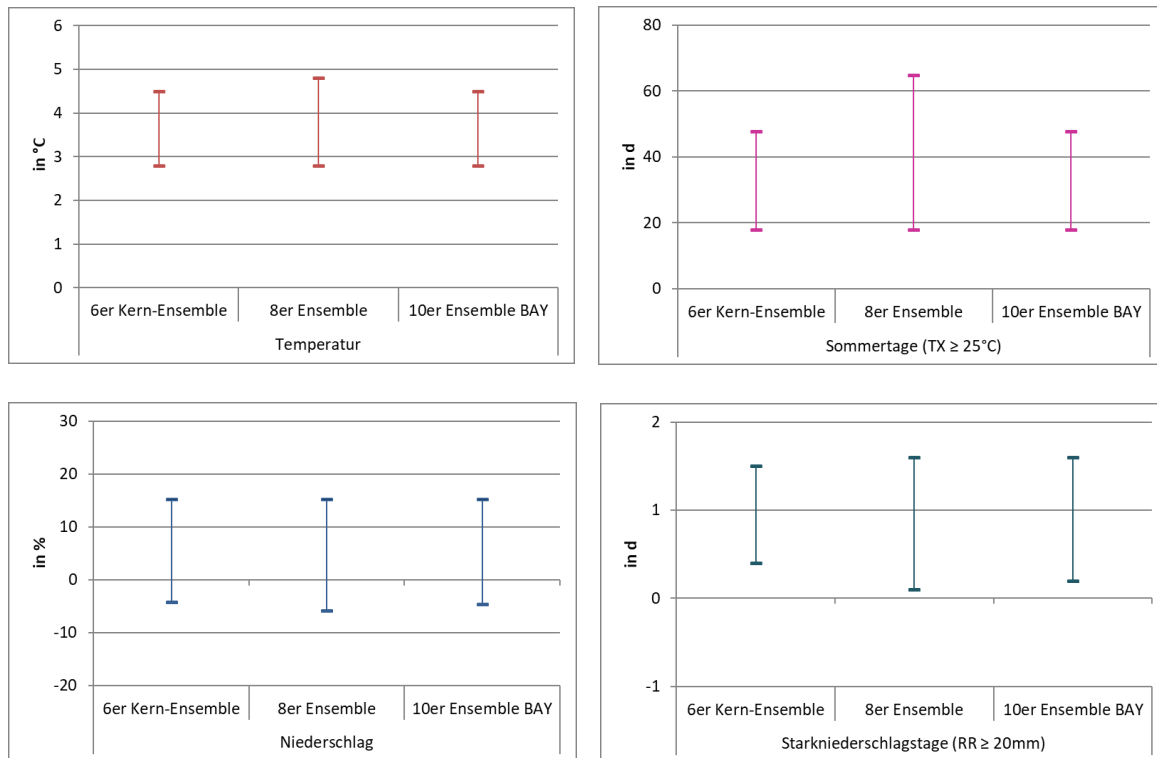


Abb. 10.2 Ergebnisübersicht Ensemblevergleich (Min-Max-Werte)
LfU Brandenburg

...