

# Schlussbericht

## Einleitung

Ziel des Vorhabens war die Bereitstellung regionalisierter Klimainformationen auf der dekadischen Zeitskala für Deutschland und Europa. Aus wissenschaftlichen Studien (z.B. Smith et al., 2007) ist bekannt, dass der Nord-Atlantik eine Region ist die Potential für eine Vorhersage auf der dekadischen Zeitskala zeigt. Diese und andere Studien dienten als Grundlage für das gesamte MiKlip Programm. Die Regionalisierung von Klimainformationen auf der dekadischen Zeitskala wurde in MiKlip erstmals durchgeführt und lebt von der Vorhersagegüte des globalen Systems. Aus den in REDCLIP vorhandenen Erfahrungen der saisonalen Vorhersage wurde ein umfangreiches Ensemble auf der regionalen Skala geplant. Darüber hinaus waren verschiedene Ensembletechniken sowie die statistische Aufbereitung zur Verbesserung der Vorhersagegüte der Simulationsdaten Gegenstand der Planungen. Des weiteren wurde die Anwendbarkeit einer statistischen Aufbereitung der Simulationsergebnisse zur Verbesserung der Vorhersagegüte nach Weigel et al. (2009) im Laufe des Vorhabens untersucht. Hierzu wurde eine Kooperation mit Andreas Weigel (damals am Bundesamt für Meteorologie und Klimatologie der Schweiz) vereinbart, um die Methode der Re-Kalibrierung von Simulationsergebnissen (engl. „climate conserving recalibration“ CCR) auf die regionalisierten Simulationen anzuwenden. Außerdem wurde die Anwendung der Ensembletechniken innerhalb MiKlips mit dem Modul-C Projekt LACEPS abgestimmt.

Während der Projektlaufzeit stellte sich heraus, dass der hierfür vorgesehene Supercomputer am Deutschen Klimarechenzentrum mehrfach überbucht war. Daher musste das Vorhaben mit massiven Rechenzeitkürzungen durchgeführt werden, die besonders zu Lasten der Untersuchungen zu den Ensembletechniken und der Modellauflösung ging. Die Untersuchung unterschiedlicher Ensembletechniken wurde weitestgehend von LACEPS übernommen und in REDCLIP durch eine Untersuchung der internen Variabilität von dekadischen, regionalisierten Klimainformationen ersetzt. In der zweiten Phase von MiKlip wurde dann die Auflösung der Simulationen vergrößert, um die nötige Rechenzeit für ein Ensemble mit vielen Startterminen zu erhalten.

## Validierung des regionalen Klimamodells REMO

Bevor REMO zum Regionalisieren von dekadischen Vorhersagen benutzt werden konnte, mussten einige Vorarbeiten erbracht werden. Zum einen musste überprüft werden, ob REMO zufriedenstellende Ergebnisse bei der Simulation des Klimas im Zielgebiet liefert. Zum anderen war es wichtig für die jeweiligen Starttermine der zu simulierenden Hindcasts passende Startwerte für Bodenparameter wie Feuchte und Temperatur zu erzeugen.

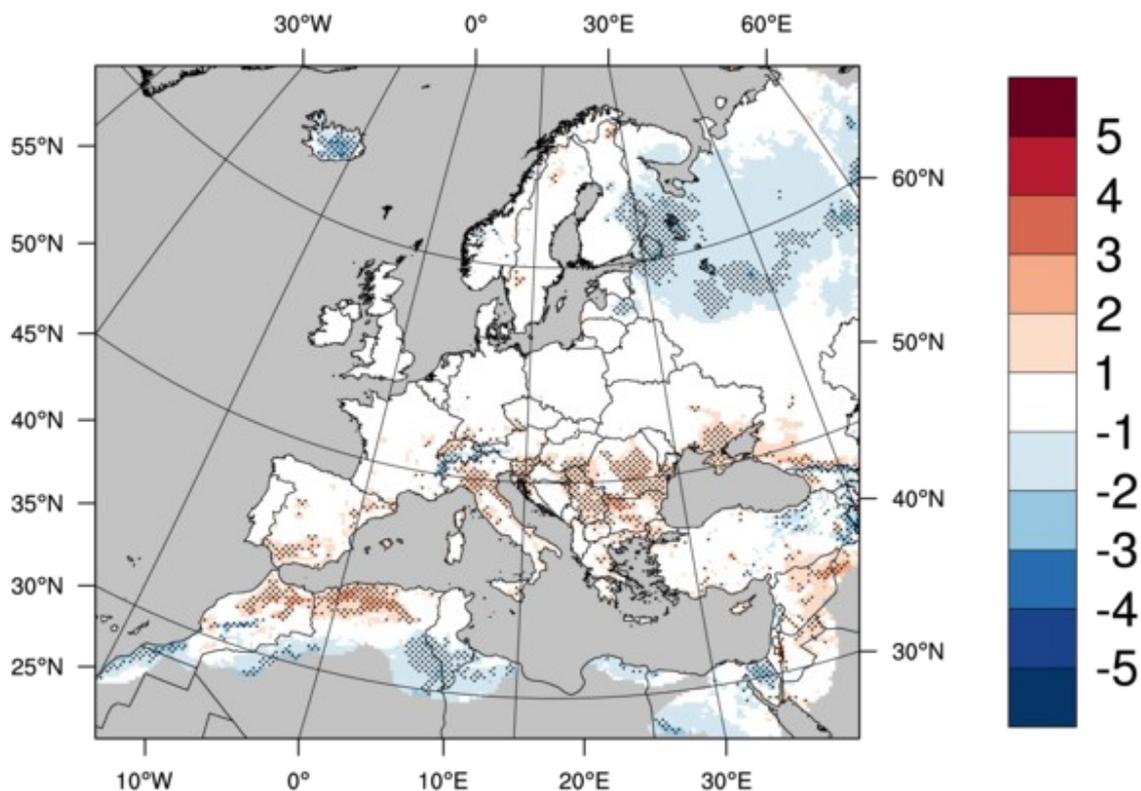


Abbildung 1: Differenz zwischen der REMO ERA-Interim Simulation und E-OBS Beobachtungen in der Jahresmitteltemperatur von 1979-2008 in [K]. Abweichungen, die größer als der Standardfehler aus E-OBS sind, sind gepunktet markiert. Für grau eingefärbte Flächen liegen keine Beobachtungsdaten vor.

### Vergleich mit Beobachtungsdaten

Um die Güte von REMO auf dem neu erstellten Simulationsgebiet abzuschätzen, wurde eine sogenannte Evaluierungssimulation mit ERA-Interim Reanalysen vorgenommen. Diese Simulation liegt für die Jahre 1979-2011 vor. Die Parameter Temperatur und Niederschlag aus der ERA-Interim Simulation wurden mit gegitterten E-OBS Beobachtungsdaten (Version 7.0; Haylock et al., 2008) verglichen. Um auch Unsicherheiten in den Beobachtungen mit einfließen zu lassen, wurde der in E-OBS

angegebene Standardfehler (Hofstra et al., 2009) mit in die Analyse einbezogen. Dieser Fehler bezieht sich nur auf die Unsicherheit der Interpolationsmethode von Stations- auf Gitterdaten und beinhaltet keine weiteren Unsicherheitsfaktoren wie z. B. Messfehler, und ist daher als eine untere Schranke der Unsicherheit anzusehen (Hofstra et al., 2009).

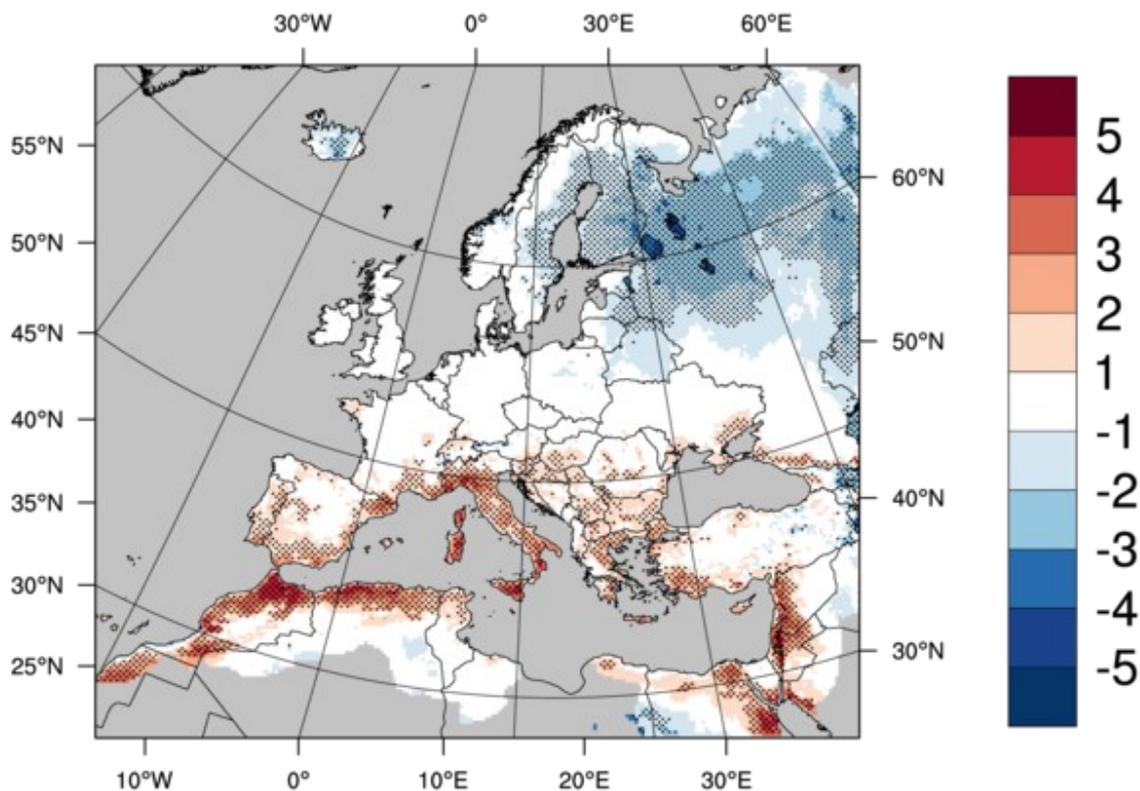


Abbildung 2: Wie in Abbildung 1 nur für das Jahresmittel der Tagesmaximaltemperatur in [K].

Im Vergleich zu E-OBS Beobachtungen zeigen sich im Jahresmittel der Temperatur nur vereinzelt Abweichungen (siehe Abbildung 1), die größer als der Standardfehler in E-OBS selbst sind (dargestellt durch die gepunkteten Flächen). Die Temperatur im Nordosten Europas wird um ca. 2 K zu kühl simuliert, während der Mittelmeerraum und der Südosten Europas teilweise um 2 K zu warm berechnet wird. Die Simulation der Tagesmaxima ergibt ein ähnliches Bild (Abbildung 2), aber mit stärkerer Ausprägung. Es zeigen sich Überschätzungen im Mittelmeerraum von bis zu 4 K. Auffällig ist, dass sich diese Überschätzungen besonders auf die Küstenregionen konzentrieren. Im Nordosten Europas werden die Maxima systematisch um 2-3 K unterschätzt.

Bei den Tagesminima zeigt sich ein uneinheitliches Bild (Abbildung 3). In weiten Teilen Mitteleuropas werden die Werte um 1-3 K überschätzt (ausgenommen der Alpenraum), während im Nordosten Europas (2-3 K), sowie in Teilen Nordafrikas (bis zu 5 K) die Werte unterschätzt werden.

Für den Jahresniederschlag (Abbildung 4) gibt es eine Tendenz zu einer Überschätzung. Diese liegt allerdings in den meisten Regionen niedriger als die Unsicherheit in der Interpolationsmethode von E-OBS. Darüber hinaus werden die Niederschlagsmengen in E-OBS generell unterschätzt (Hofstra et al., 2009). Die starken Über- und Unterschätzungen in den trockenen Regionen im östlichen Mittelmeerraum lassen sich durch die Darstellung als relative Abweichungen erklären. In fast allen diesen Regionen ist die Unsicherheit der Beobachtung größer als die Abweichung in der Simulation.

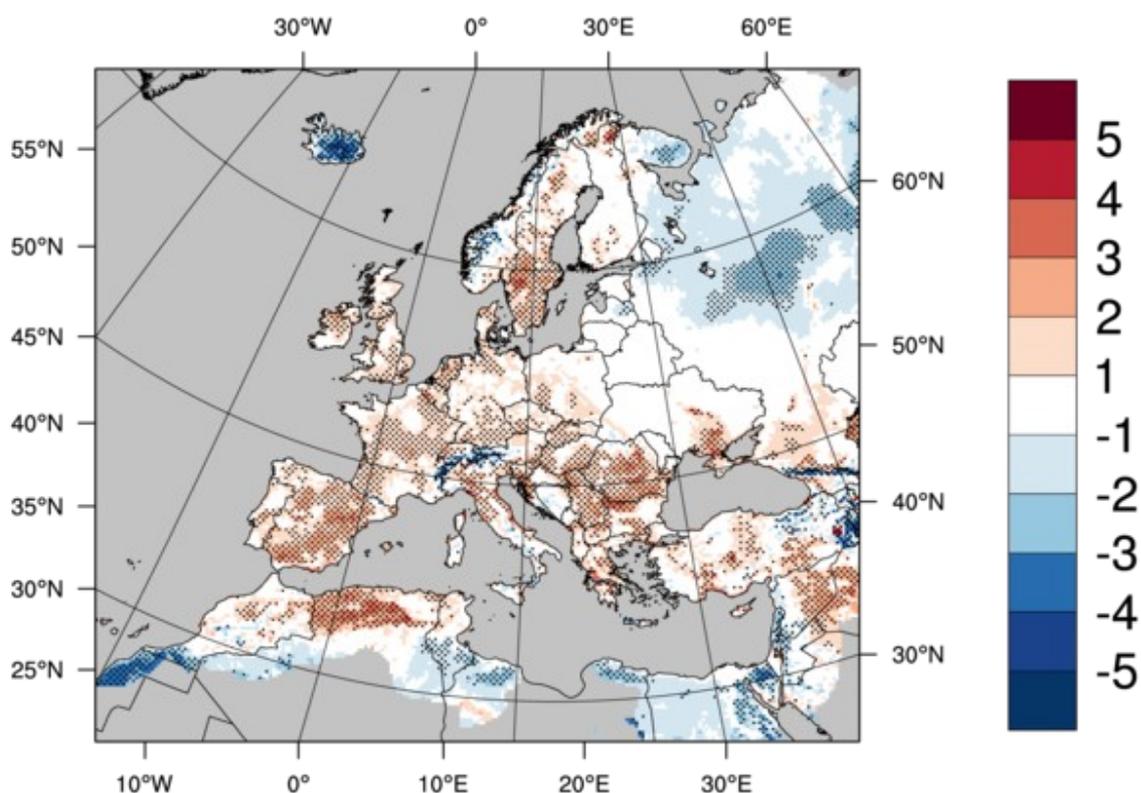


Abbildung 3: Wie in Abbildung 1 nur für das Jahresmittel der Tagesminimaltemperatur in [K].

In den Jahresmittelwerten zeigt REMO eine zufriedenstellende Güte. Dies gilt besonders dann, wenn die Unsicherheit der Beobachtung mit in Betracht gezogen wird. Bei den Temperaturminima und -maxima zeigen sich größere Abweichungen die es genauer zu untersuchen gilt. Zum Teil zeigen sich „künstliche“ Strukturen in dem E-OBS Beobachtungsdatensatz (siehe z. B. Abbildung 3 östlich des Schwarzen Meeres). Einige dieser Strukturen lassen sich durch Fehler in E-OBS erklären, die in aktuelleren Versionen des Datensatzes korrigiert wurden. Darüber hinaus zeigt sich, dass der Vergleich mit anderen Beobachtungsdaten systematische Fehler aufdecken kann. Diese Analyse ist besonders für die Anwendung der „Climate Conserving Re-calibration (CCR)“ Methode wichtig und wird weiter unten dargestellt.

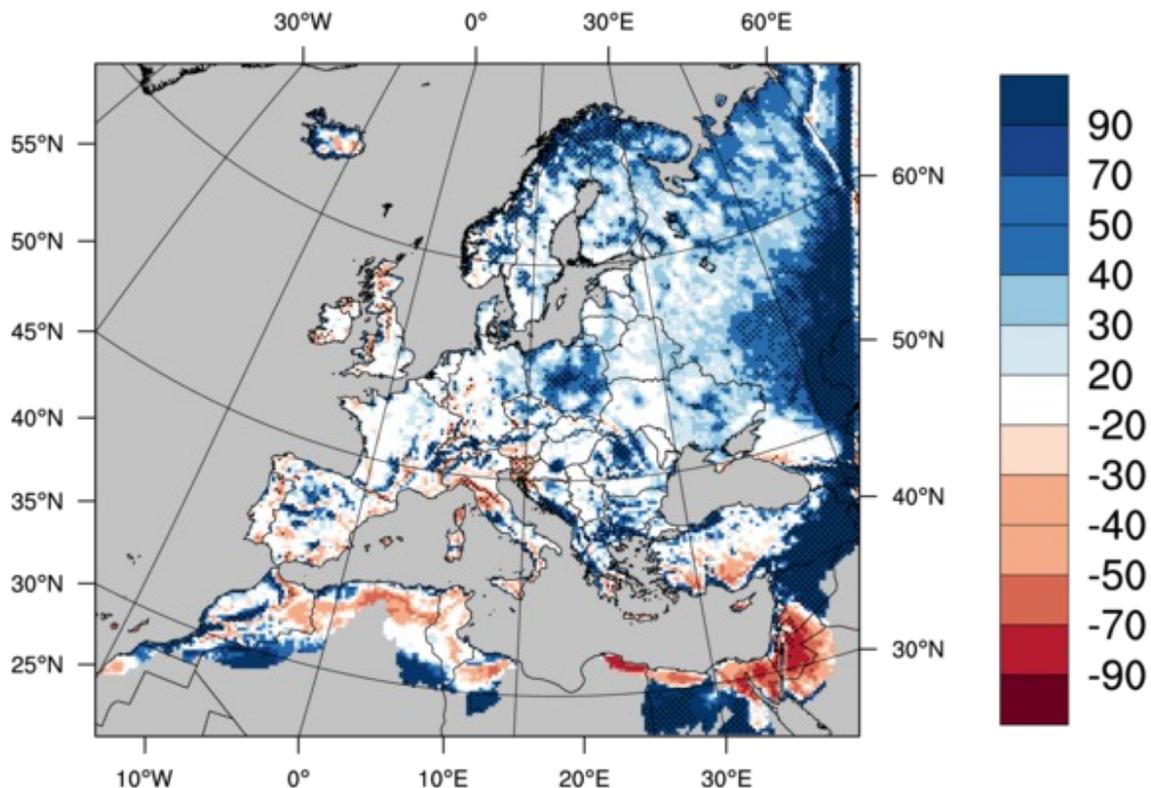


Abbildung 4: Wie in Abbildung 1 nur für den Jahresniederschlag in [%].

## Vergleich mit bekannten Korrelationsmustern

In Müller et al. (2012) wird die Vorhersagegüte des MiKlip Baseline0 Vorhersagesystems auf saisonaler Ebene untersucht. Dabei zeigt die Analyse von Beobachtungsdaten, dass es für den Sommer (Juni-August) signifikante Korrelationen zwischen vier Jahre gemittelten Meeresoberflächentemperaturen im Nordatlantik und bodennahen Temperaturen in Südost Europa gibt. Eine wichtige Frage für die Regionalisierung ist, ob solche Korrelationen auch vom Regionalmodell wiedergegeben werden können. Da es sich hierbei um großräumige Fernwirkungen handelt, muss diese Information von außen durch die Randbedingungen in das Regionalmodell hineingebracht werden. Da das bisherige Vorhersagesystem diese Korrelation nur ungenügend wiedergibt, wird zunächst die Evaluierungssimulation dahingehend untersucht.

Für die Meeresoberflächentemperatur (englisch „Sea-surface-temperature“, SST) ergibt sich dabei kein Unterschied zu den Reanalysen, da die Werte direkt in REMO übernommen werden. Der SST-Index aus Müller et al. (2012) berechnet sich als Mittel der SST in einem Gebiet von 40°W-15°W und 50°N-60°N im Nordatlantik. Eine Zeitreihe von gleitenden vier Jahresmitteln der saisonal gemittelten Sommertemperatur (Juni-August)

dieses Index wird dann mit allen Gitterpunkten der Oberflächentemperatur (zeitlich auf die gleiche Art gemittelt) korreliert.

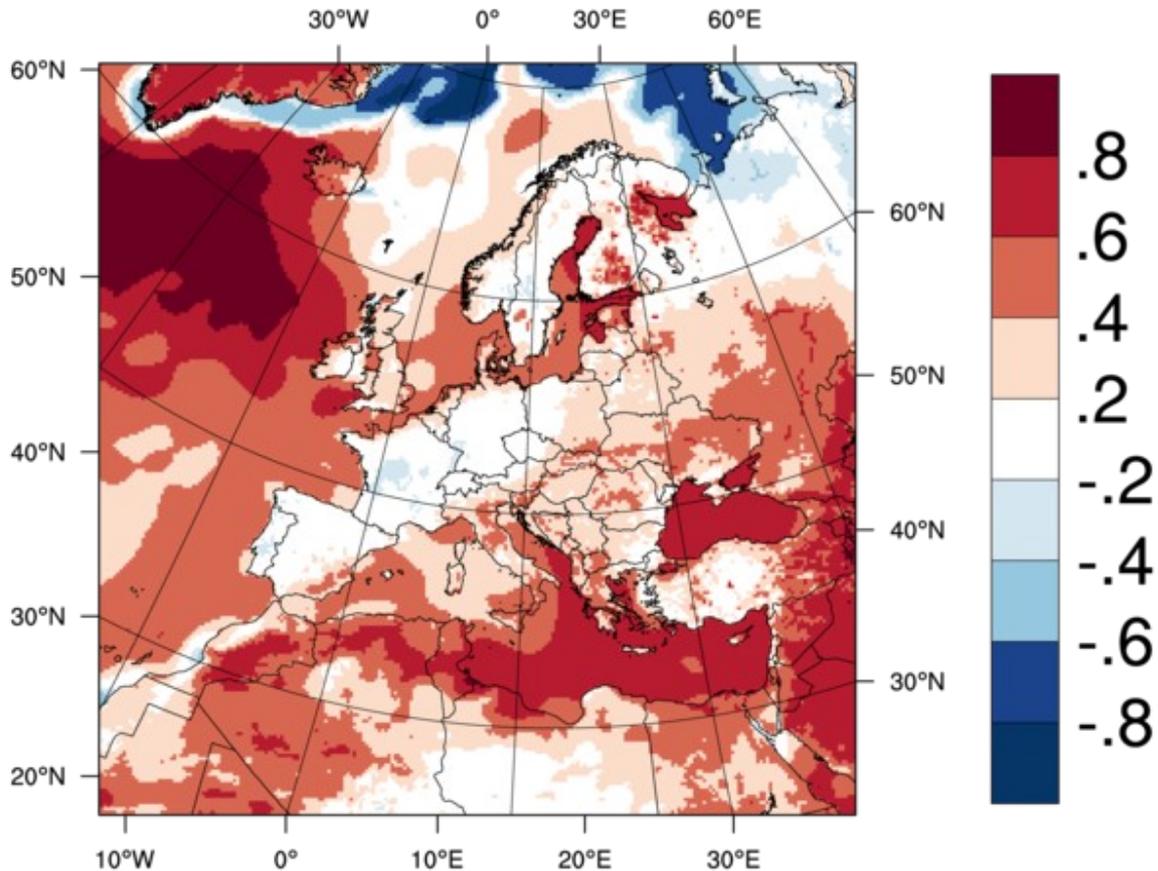


Abbildung 5: Korrelationskoeffizient zwischen Meeresoberflächentemperatur im Nordatlantik ( $40^{\circ}W-15^{\circ}W$ ,  $50^{\circ}N-60^{\circ}N$ ) und der Oberflächentemperatur der Evaluationssimulation als gleitende vier Jahresmittel der Sommertemperaturen.

Abbildung 5 zeigt den Korrelationskoeffizienten für jeden Gitterpunkt des Modellgebietes. Das großräumige Muster entspricht in etwa den Erwartungen aus den Beobachtungen. Die SST über dem Nordatlantik zeigt eine deutlich positive Korrelation mit den bodennahen Temperaturen im östlichen Mittelmeerraum und nördlich des Schwarzen Meeres. Damit im Zusammenhang steht ein charakteristisches Korrelationsmuster (vgl. u. a. Müller et al., 2012) zwischen SST und reduziertem Bodendruck (Abbildung 6). Ein wichtiger Unterschied zu den Ergebnissen aus Müller et al. (2012) ist die Verschiebung der positiven Korrelation östlich der Ostsee nach Osten. Sie erklärt auch den leicht nach Osten verschobenen Bereich der positiven Korrelationen in der Temperatur.

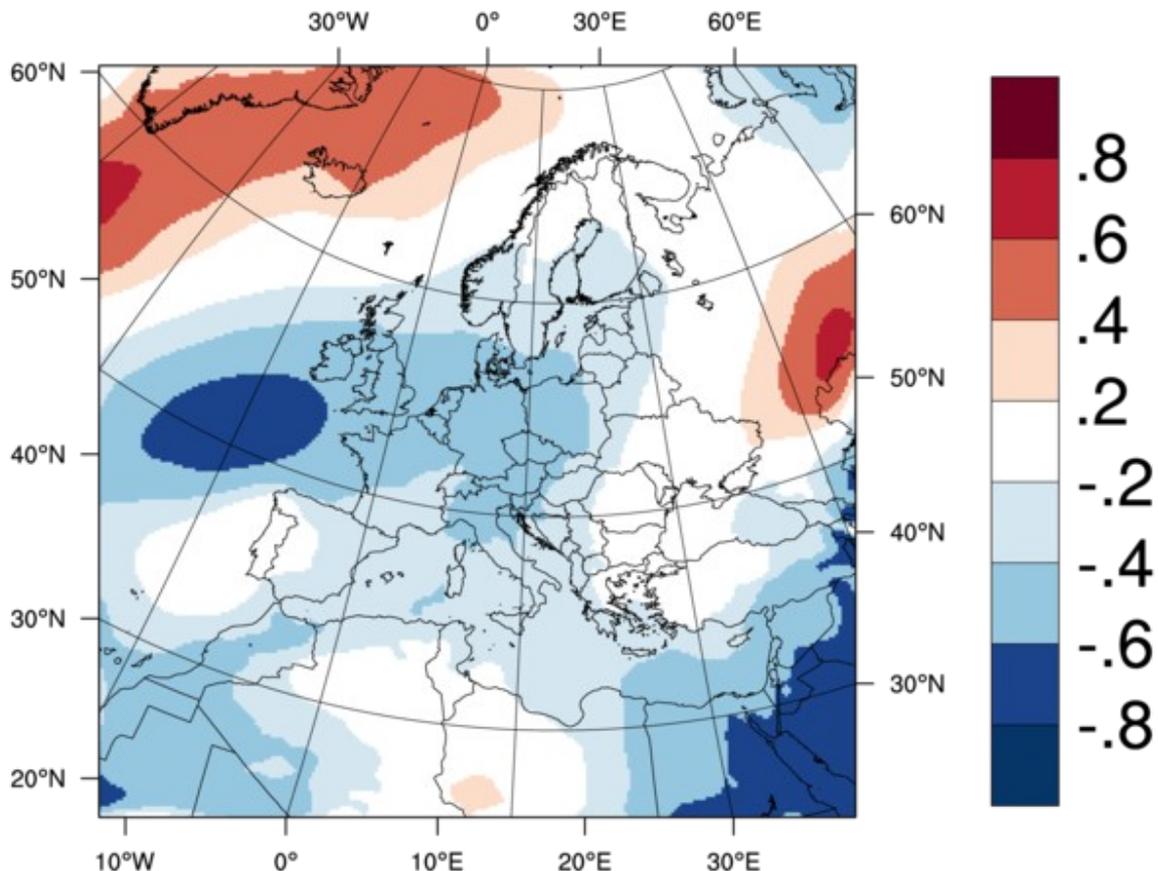


Abbildung 6: Wie Abbildung 5 nur zwischen Meeresoberflächentemperatur und reduziertem Bodendruck.

## Modellinitialisierung

In einer Sensitivitätsstudie zur Initialisierung dekadischer Vorhersagen, wurden jeweils eine Simulation mit und ohne „eingeschwungenen Boden“ initialisiert. Ein eingeschwungener Boden oder der sogenannte Warmstart bedeutet in diesem Fall, dass die Bodenvariablen wie Bodentemperatur und -feuchte aus einer Modellsimulation stammen, bei der diese Werte mit der Modellklimatologie im Gleichgewicht sind. Wird das Modell dagegen direkt mit Antriebsdaten initialisiert (Kaltstart), kommt es häufig zu einer mehrjährigen Drift der Bodenvariablen. Dieser Effekt wirkt sich besonders in den ersten fünf Jahren einer Simulation auf die Ergebnisse aus.

Abbildung 7 zeigt die Differenz der über die Jahre 2-5 gemittelten Sommertemperaturen (Juni-August) zwischen einem Warm- und einem Kaltstart mit jeweils gleichen Antriebsdaten. Besonders westlich des Schwarzen Meeres gibt es zum Teil deutliche Unterschiede von bis zu 1.5 K in der bodennahen Temperatur zwischen den beiden Simulationen. Ein Blick auf die Unterschiede in der Bodenfeuchte zeigt, dass die trockeneren Bereiche wärmer und nassere kühler simuliert werden (Abbildung 8). Dieser

Effekt nimmt mit zunehmender Simulationsdauer ab. Es kann also festgehalten werden, dass die Initialisierung der Bodenfeuchte im Regionallmodell einen großen Einfluss auf Ergebnisse in den ersten Jahren haben kann. Nähere Untersuchungen zum Einfluss der Bodeninitialisierung auf die Vorhersagbarkeit werden in anderen Projekten innerhalb von MiKlip unternommen.

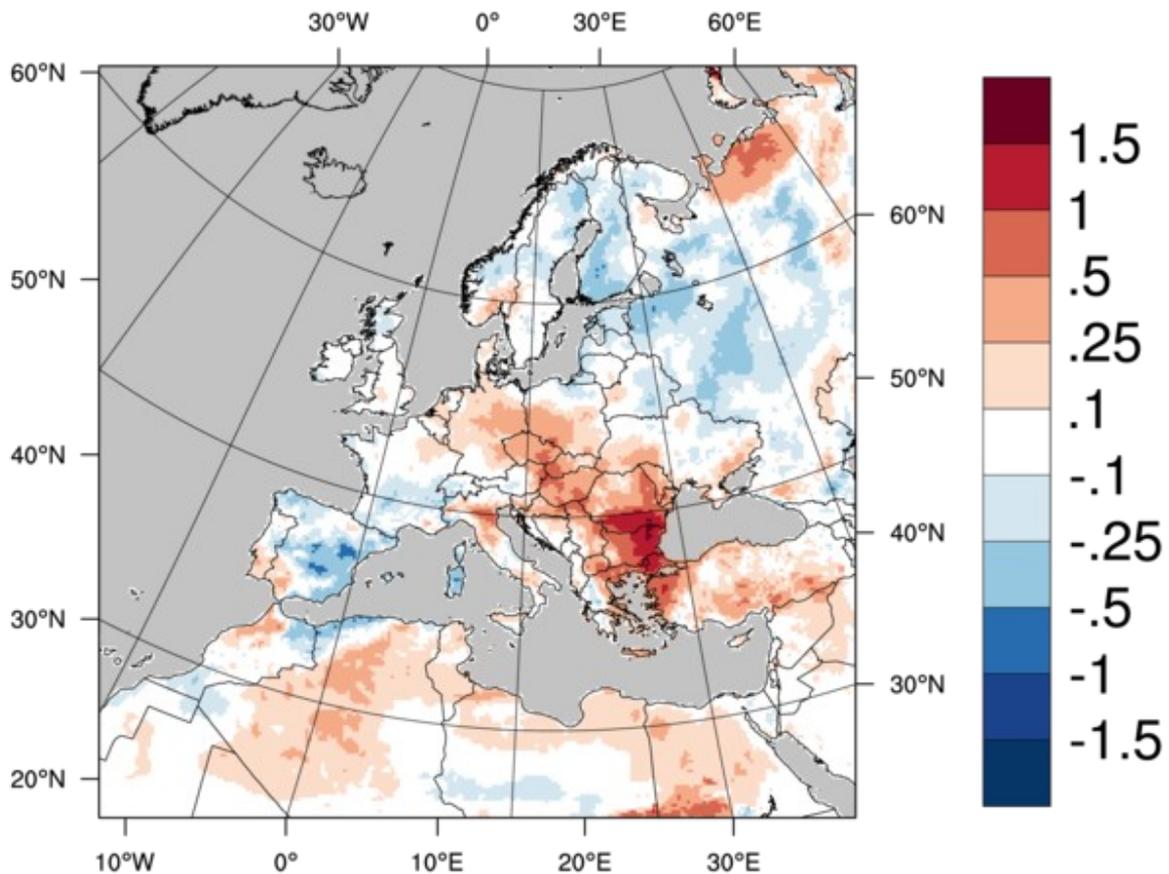


Abbildung 7: Differenz zwischen Kaltstart und Warmstart der gemittelten Sommertemperatur über die Jahre 2-5 nach Modellstart in [K].

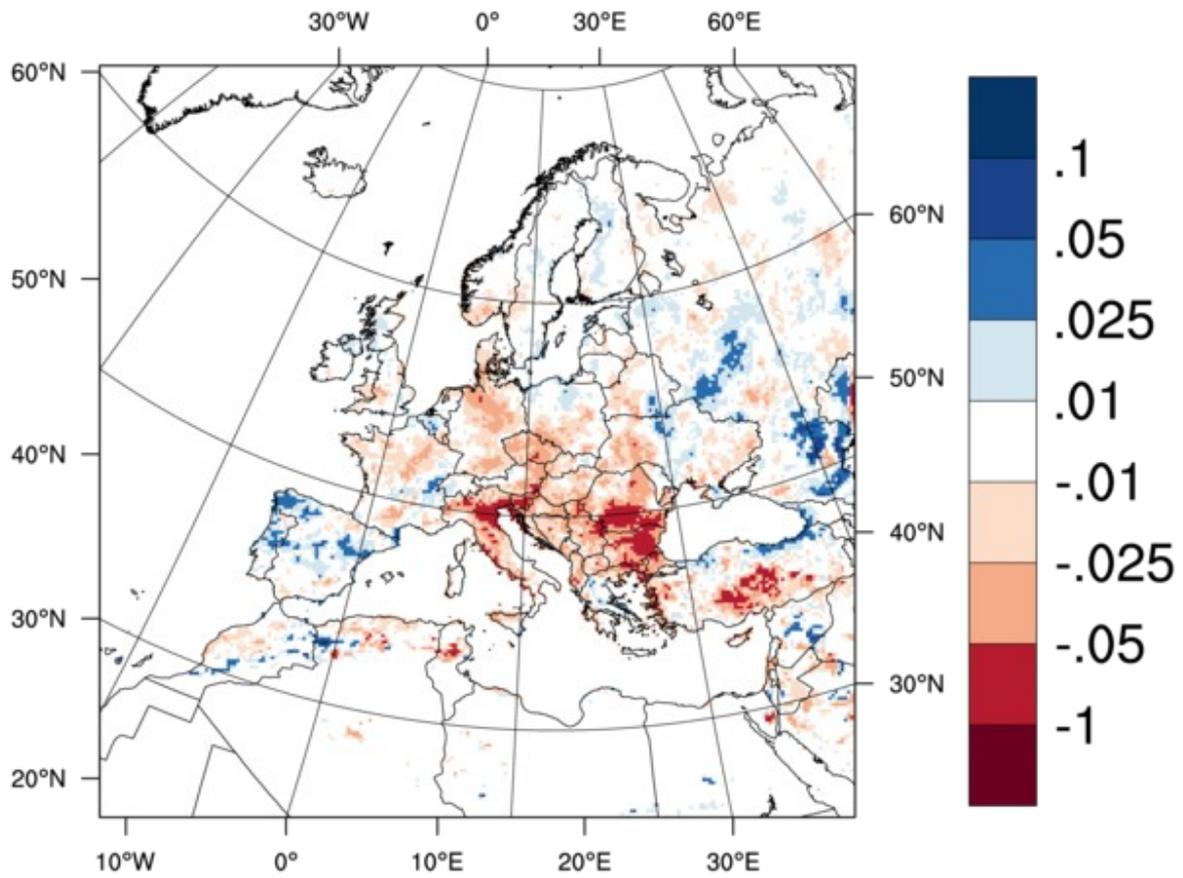


Abbildung 8: Wie Abbildung 7 nur für die Bodenfeuchte in [m].

## Ergebnisse

In den unterschiedlichen Phasen des Projektes, den so genannten „Development Stages (DS)“, wurden verschiedenste Ergebnisse erzielt. Hauptziel in DS1 und DS2 war es, die Simulationen für das regionale MiKlip Ensemble zu erstellen. Hierbei wurden in DS1 Simulationen auf einer Auflösung von  $0.22^\circ$  in der Horizontalen durchgeführt. In DS2 wurde eine doppelt so grobe ( $0.44^\circ$ ) Auflösung für REMO verwendet, damit ein großes Ensemble von vielen Simulationen ( $\sim 300$  für Modul-C) in DS3 zur Verfügung stand. In DS3 wurde dann die Re-Kalibrierungsmethode „Climate Conserving Re-calibration (CCR)“ auf die Ergebnisse aus DS2 angewendet.

### 0.22° Simulationen aus DS1

Zum Ende von DS1 wurden alle vereinbarten Simulationen für die Auswertung innerhalb von MiKlip zur Verfügung gestellt. Tabelle 1 fasst die Simulationen für DS1 nochmals zusammen.

*Tabelle 1: Für DS1 vereinbarte Simulationen ( $0.22^\circ$  horizontale Auflösung) mit dem MiKlip baseline0 System*

	Decadal1960	Decadal1970	Decadal1980	Decadal1990	Decadal2000
Simulationen	2	2	2	2	10

Ein wichtiges wissenschaftliches Ergebnis aus REDCLIP stellt die Untersuchung der internen Variabilität auf der dekadischen Zeitskala dar. Es zeigt sich, dass REMO selbst für mehrjährige Sommermitteltemperaturen andere Signale aufweist als das antreibende Globalmodellsystem. Dies bedeutet, dass REMO sich auf dieser Zeitskala vom antreibenden System entkoppeln kann. Abbildung 9 zeigt dies für den reduzierten Bodendruck und Abbildung 10 für den Niederschlag.

Für die Regionalisierung können daraus wichtige Konsequenzen folgen. Sollte die Vorhersagegüte nur durch die Randbedingungen in das Modellgebiete gebracht werden, müssten dann auch für die Regionalisierung Ensembles von jedem einzelnen Ensemblemitglied des globalen Systems gerechnet werden, um ein robustes regionales Signal zu erhalten. Auf der anderen Seite könnte die Entkopplung zur einer Verbesserung der Vorhersagegüte führen, wenn es lokale Quellen, wie z. B. Landoberflächenprozesse, für die Vorhersagbarkeit gibt. In Mieruch et al., 2014 wird zumindest für den Niederschlag so ein Mehrwert durch die Regionalisierung in MiKlip gefunden. Weitergehende Untersuchungen konnten auf Grund der bekannten Probleme bei der Rechenzeitvergabe nicht durchgeführt werden.

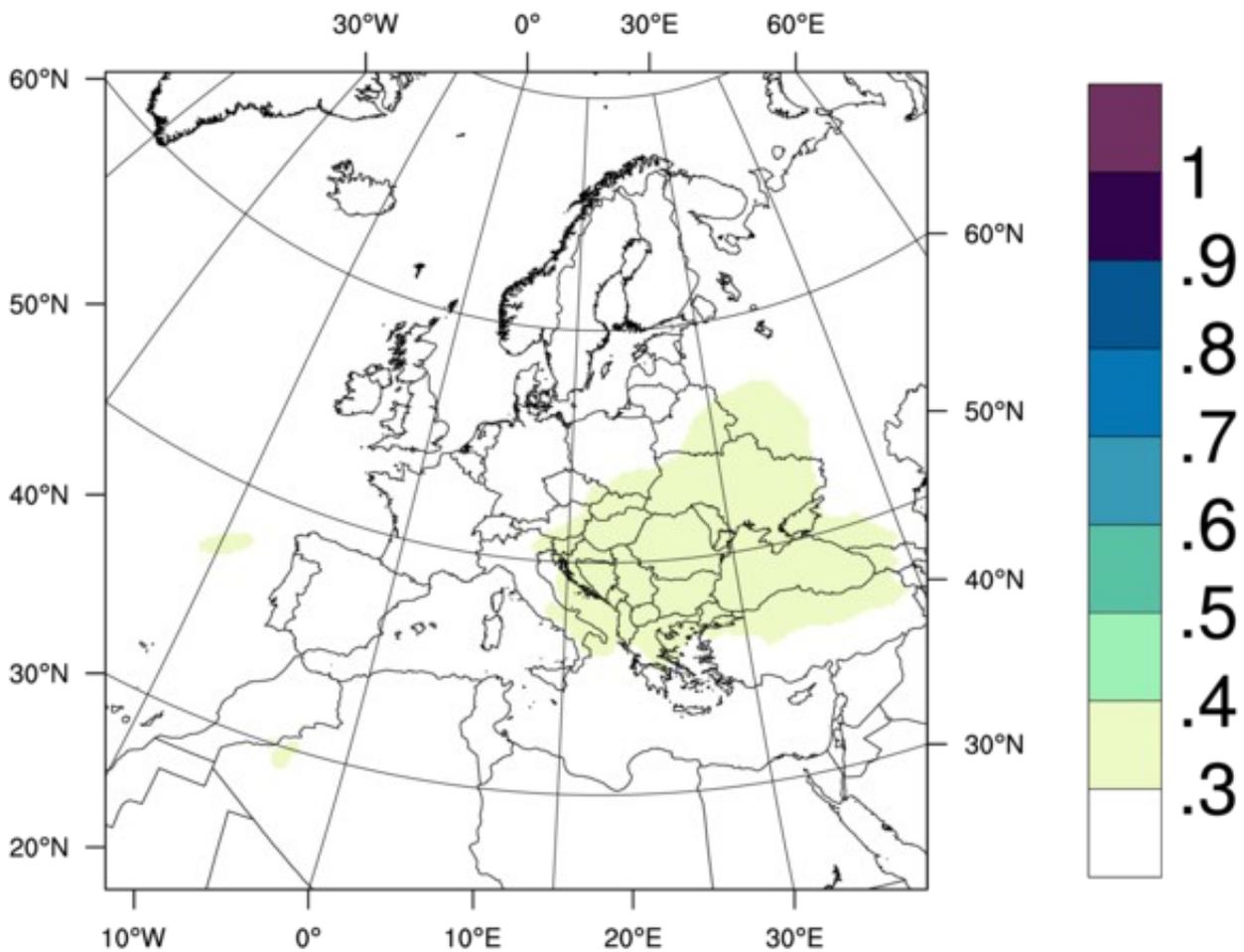


Abbildung 9: Entkopplung des reduzierten Bodendrucks (4-jähriges Sommermittel) der REMO Simulationen vom globalen Vorhersagesystem. Ein Wert von 1 bedeutet vollständige Entkopplung und 0 vollständige Kopplung zwischen den Modellen.

Obwohl eine Auflösung von  $0.22^\circ$  oder besser wünschenswert für die weitere Verarbeitung der Ergebnisse wäre, zeigte sich, dass der Mangel an Rechenzeit nur eine unzureichende Anzahl an Simulationen für tiefer gehende statistische Auswertungen lieferte. Diesem Umstand wurde in DS2 Rechnung getragen.

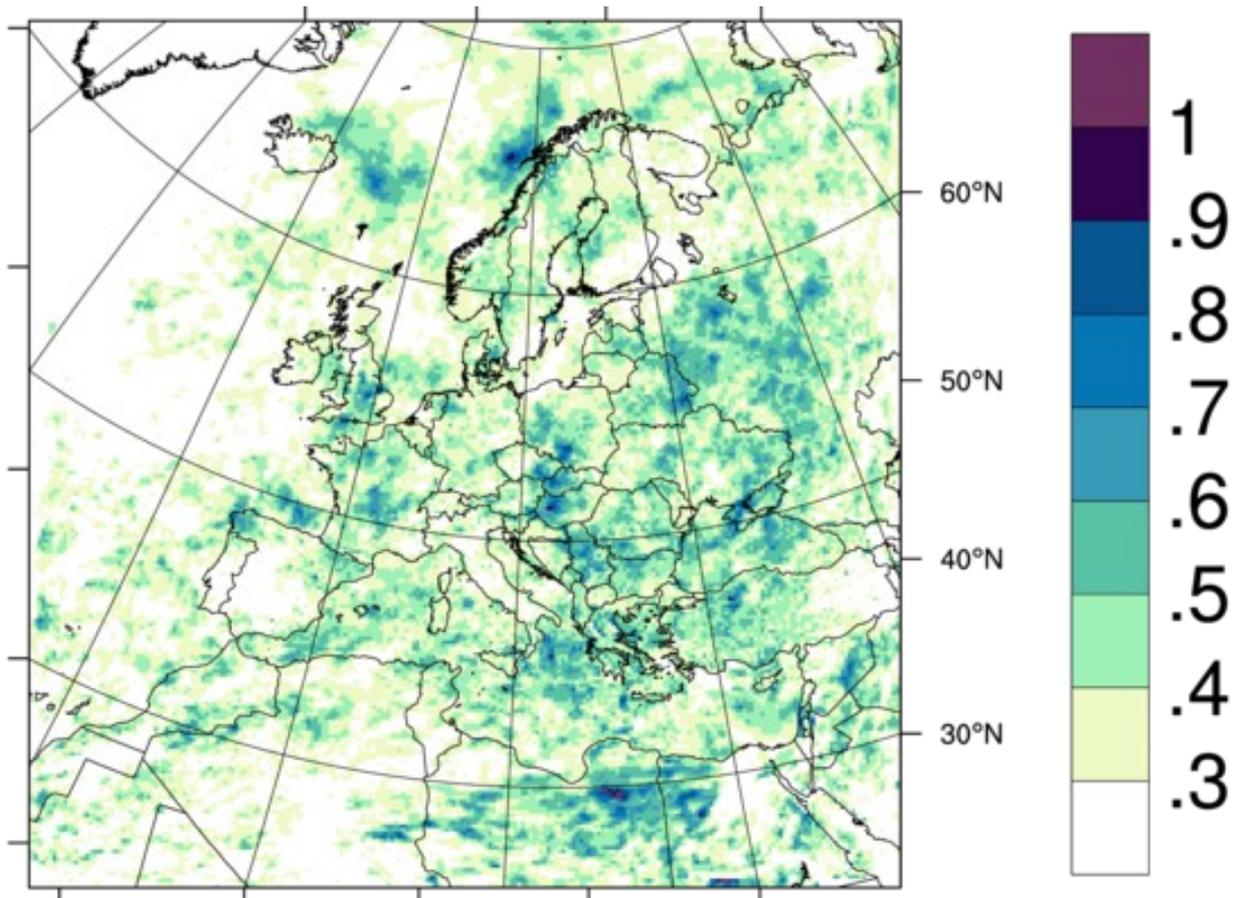


Abbildung 10: Wie Abbildung 9 nur für den Niederschlag.

## 0.44° Simulationen aus DS2

Mit dem Übergang zu DS2 wurde nicht nur das globale Vorhersagesystem von baseline0 auf baseline1 geändert. Auch die Auflösung für die Simulationen wurde von 0.22° auf 0.44° vergrößert. Innerhalb von Modul-C wurde diese Maßnahme getroffen, da der signifikant geringere Rechenzeitbedarf (ca. Faktor 10) ein wesentlich größeres Ensemble erlaubt. Für robustere statistische Aussagen (u.a. eine Forderung aus Modul-D) und für die Vorhaben in REDCLIP im Speziellen war diese Änderung notwendig. Ergebnisse aus ENSEMBLES (van der Linden and Mitchell, 2009) haben zu dem gezeigt, dass sich an den großräumigen Mustern im simulierten Klima keine wesentlichen Änderungen zwischen beiden Auflösungen ergeben. Durch die Änderung der Auflösung und die Einbeziehung der Starttermine ab 1960 wurde zusätzlich zu einer vorhandenen ERA-Interim Simulation (1979-2008) noch eine Simulation mit ERA40 Daten ab 1959 durchgeführt, um Initialisierungsdaten für die Bodenparameter über den gesamten Hindcastzeitraum zu haben.

Mit REMO wurden von 1960-2003 jährlich 2 von 10 globalen Vorhersagen regionalisiert. Für den Vergleich mit DS1 wurden für die Dekade 2000 alle 10 globalen Simulationen regionalisiert (siehe Tabelle 2). Ausgewählte Parameter (nach Absprache mit anderen Modulen) aus den Ergebnissen aller Simulationen stehen seit Januar 2015 auf dem MiKlip

Server zur Auswertung bereit. Für das regionale Ensemble (REMO + CLM) sind in Zusammenarbeit mit DroughtClip und INTEGRATION zusätzlich noch sog. „bias-korrigierte“ Daten berechnet und in das MiKlip Evaluierungssystem integriert worden.

*Tabelle 2: In DS2 durchgeführte Simulationen (0.44° horizontale Auflösung) mit dem MiKlip baseline1 System*

	Decadal1960-2003 (jährlich)	Decadal1990	Decadal2000
Simulationen	2	7	10

Die Ergebnisse bezüglich der Re-Kalibrierung der Simulationen zeigen, dass die Bestimmung der „Verlässlichkeit“ (Reliability) des regionalen Ensembles auf Grund des kleinen Ensembles (2 REMO + 4 CLM Simulationen) mit einer statistischen Unschärfe behaftet ist. Die Bestimmung der Reliability ist eine Voraussetzung zur Entscheidung über die Methode der Re-Kalibrierung. Im Prinzip sagt die Reliability aus, dass der mittlere quadratische Fehler des Ensemblemittels gleich der zeitlich gemittelten Varianz der Ensemblemitglieder entsprechen muss (siehe u.a. Weigel et al., 2009 oder Palmer et al., 2006). Exemplarisch ist in Abbildung 11 die Reliability des regionalen Ensembles für den Vorhersagezeitraum 2-5 Jahre nach der Initialisierung nach Weigel et al. (2009) dargestellt. Als Basis wurden die oben genannten bias-korrigierten Daten verwendet. Es zeigt sich, dass das regionale Ensemble in weiten Teilen Europas „overconfident“ also zu scharf ist (positive Werte). Es gibt allerdings auch Regionen in Süd-Ost-Europa, für die negative Werte angezeigt werden. Hier ist das Ensemble zu weit („underconfident“). Es gibt eine Reihe an Re-Kalibrierungsmethoden für zu scharfe Ensembles. Für zu weite Ensembles sind noch nicht viele Methoden entwickelt worden (Eade et al., 2014).

Bei der Bestimmung der Reliability hat sich heraus gestellt, dass sich die in MiKlip zur Verfügung stehenden Beobachtungsdaten in einigen Punkten unterscheiden. Sowohl in der räumlichen Homogenität als auch im Trend konnten hier Unterschiede in drei miteinander verglichenen Beobachtungsdatensätzen gefunden werden. Im einzelnen wurden CRU 3.21 (CRU, 2013), Willmott-Matsuura 3.01 (Willmott and Matsuura, 2001) und E-OBS 11.0 (Haylock, 2008) miteinander verglichen.

Abbildung 12 zeigt die Standardabweichung der drei Beobachtungsdatensätze für zwei 4-jährige Mittel der Temperatur. Für das Mittel um das Jahr 1975 zeigen sich starke, örtlich begrenzte Abweichungen in Skandinavien. Zu einem späteren Zeitpunkt (1991) treten solche Abweichungen nicht auf. Insgesamt kommen noch andere Zeiträume vor, in denen an unterschiedlichen Orten, verursacht durch verschiedene Datensätze, Abweichungen in vergleichbarer Größenordnung auftreten. Es ist daher nicht zu entscheiden, welcher Datensatz „der Beste“ ist. Vielmehr wurde in REDCLIP für die Anwendung der CCR Methode ein Ensemble von Beobachtungsdaten benutzt um den Unsicherheiten in den Beobachtungen Rechnung zu tragen. Ein weiteres Problem sind unterschiedlich starke Trends in den Datensätzen (nicht gezeigt), die regional auch verschieden sein können. Diese Unsicherheit wurde nur durch die Ensemblebildung berücksichtigt.

Unter Verwendung des Beobachtungsensembles gelingt es mit Hilfe von CCR das regionale Ensemble „reliable“ zu machen (siehe Abbildung 13). Wird nur ein Beobachtungsdatensatz verwendet gibt es immer noch Regionen, in denen die Spannweite des Ensembles zu weit oder scharf ist. Die Abhängigkeit der Ergebnisse vom

Beobachtungsdatensatz hat sich auch in anderen Projekten von Modul-C gezeigt (z. B. Regio\_Predict).

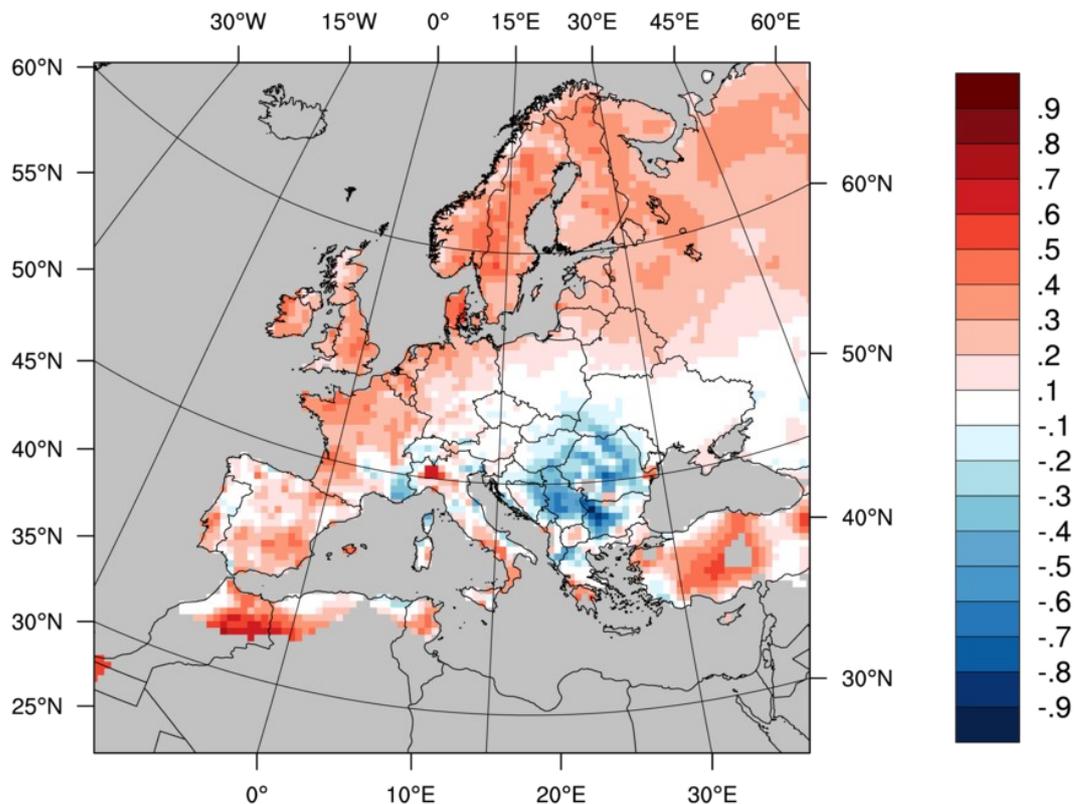
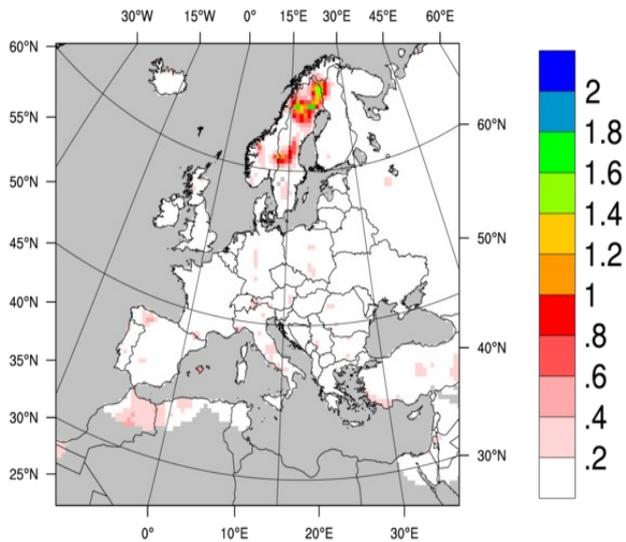


Abbildung 11: Reliability des regionalen MiKlip Ensembles nach Weigel et al. (2009). Positive Werte zeigen Regionen mit einem zu scharfen Ensemble an („overconfidence“). Negative Werte zeigen Regionen mit zu weitem Ensemble („underconfidence“).

Observation std. dev. (4-year running mean) 1975



Observation std. dev. (4-year running mean) 1991

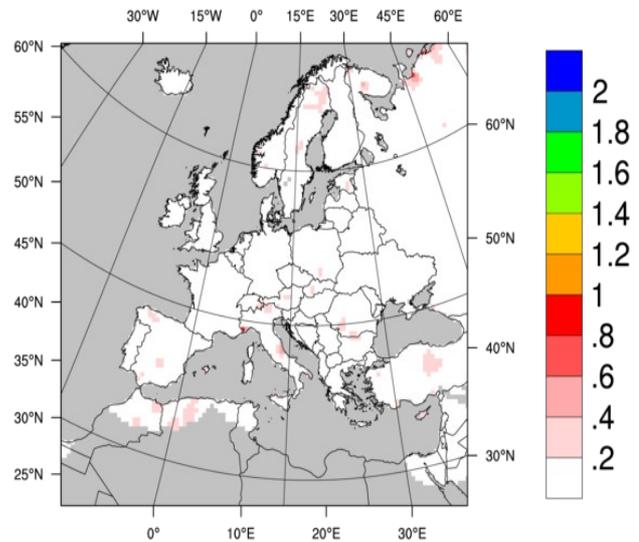


Abbildung 12: Standardabweichung in der Temperatur in [K] zwischen verschiedenen Beobachtungsdatensätzen (CRU v3.21, Willmott-Matsuura v3.01 und E-OBS v11.0). Links für das gleitende Mittel um das Jahr 1975 und rechts für 1991.

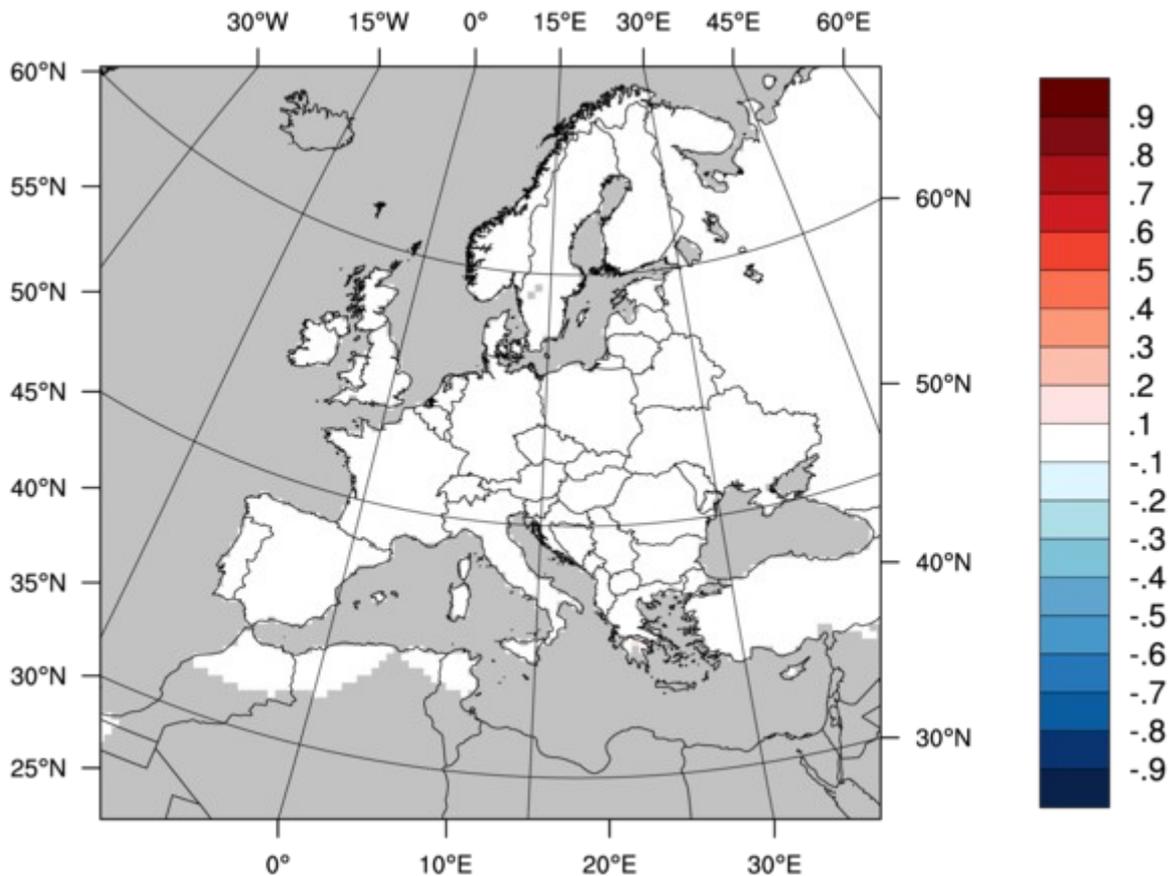


Abbildung 13: Reliability wie in Abbildung 11 nach Anwendung der CCR Methode unter Verwendung eines Ensembles an Beobachtungsdaten.

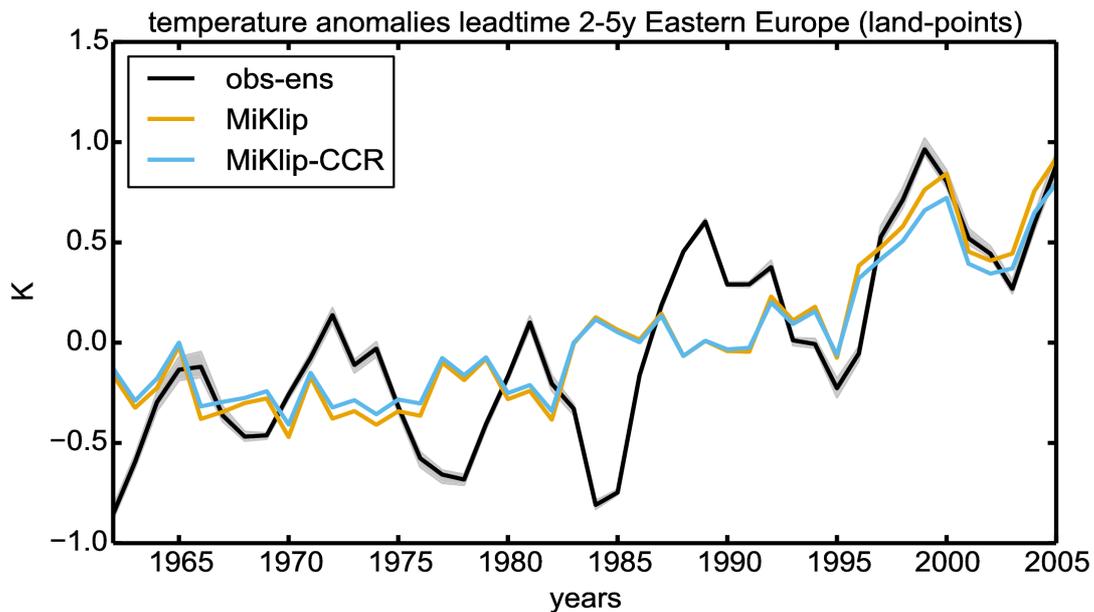


Abbildung 14: Anomalien der 4-jährigen Mitteltemperatur gegenüber der jeweiligen Referenzperiode (1970-2003) für das Ensemble der Beobachtungen (schwarze Linie), das regionale MiKlip Ensemble (gelbe Linie) und das re-kalibrierte MiKlip Ensemble (blaue Linie) in der Region „Eastern Europe“ ( $16^{\circ}\text{E}$ - $30^{\circ}\text{E}$ ,  $44^{\circ}\text{N}$ - $55^{\circ}\text{N}$ ).

Die Größe des regionalen Ensembles aus DS2 hat erstmals die Möglichkeit ergeben, regionale Temperaturtrends der initialisierten Simulationen mit Beobachtungen zu vergleichen. In Abbildung 14 sind die Anomalien der 4-jährigen Mitteltemperatur (gemittelt über die Jahre 2-5) gegenüber der Referenzperiode 1970-2003 für das regionale MiKlip Ensemble, das rekalibrierte regionale MiKlip Ensemble und dem oben genannten Ensemble an Beobachtungen für die Region „Eastern Europe“ ( $16^{\circ}\text{E}$ - $30^{\circ}\text{E}$ ,  $44^{\circ}\text{N}$ - $55^{\circ}\text{N}$ ) dargestellt. Bis in die 1990er Jahre gibt es keinen Zusammenhang zwischen Beobachtungen und simulierten Daten. Erst danach korreliert das regionale MiKlip Ensemble mit den Beobachtungen. Hier besteht ein möglicher Zusammenhang zu dem Zustand der Atlantischen Multidekaden-Oszillation (AMO). In den 1990er Jahren hat sich diese von einem negativen in einen positiven Zustand verschoben (siehe z. B. Sutton, 2005). Es ist spekulativ aber durch die oben beschriebene Korrelation von Meeresoberflächentemperatur und Lufttemperatur in Ost-Europa nicht ganz unplausibel, dass die unterschiedlichen Phasen der AMO unterschiedliche Auswirkungen auf die Vorhersagbarkeit der Temperatur haben. Weitere Untersuchungen im Nachgang zu MiKlip oder in MiKlip II können hier mehr Klarheit schaffen.

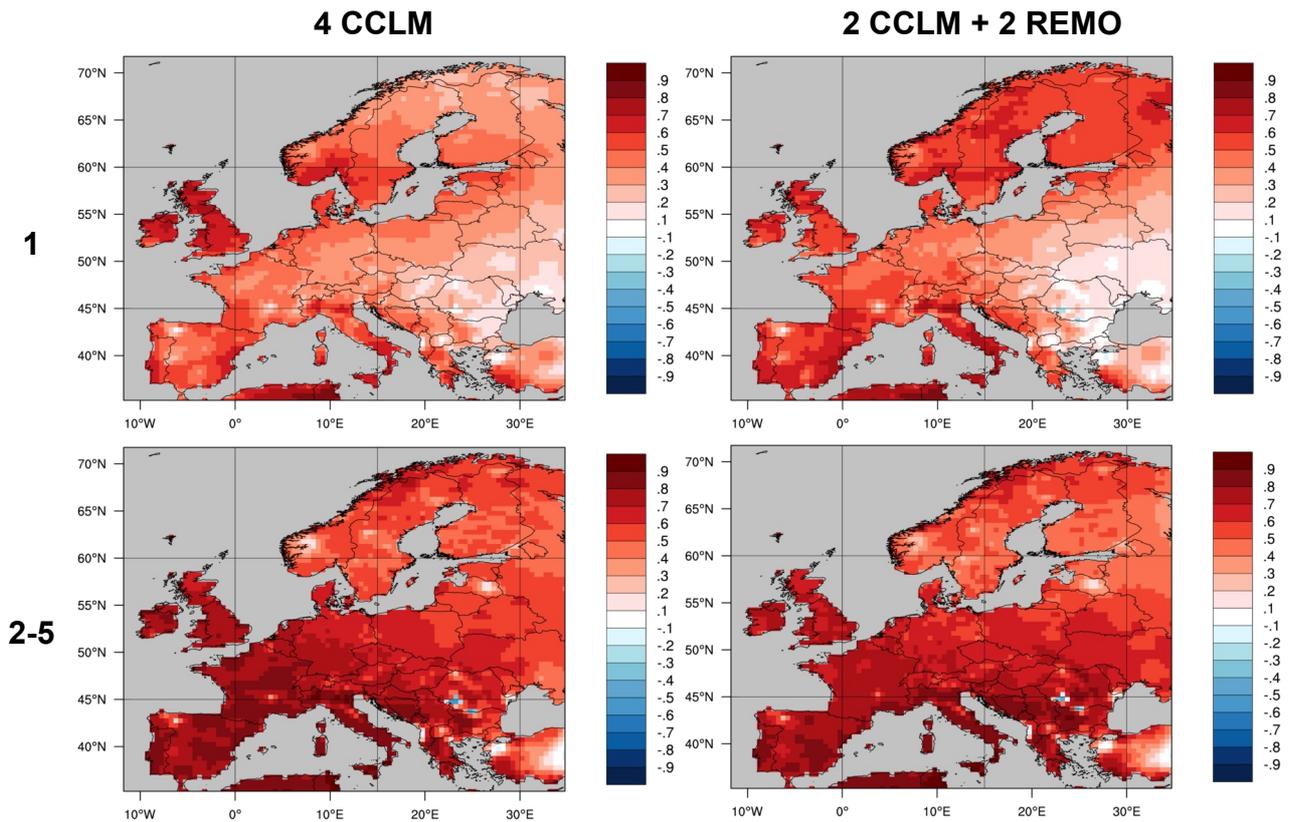


Abbildung 15: Korrelation der Temperatur zwischen Ensemblemittel und Beobachtungen (E-OBS 9.0) für das erste Jahr (obere Reihe) und gemittelt über die Jahre 2-5 (untere Reihe) für zwei Teilensembles des regionalen MiKlip Ensembles. Ein Ensemble wurde dabei aus 4 Simulationen mit CCLM (linke Spalte) gebildet und das andere aus 2 CCLM und 2 REMO Simulationen zusammengesetzt (rechte Spalte). Rote Farben stehen für positive, blaue Farben für negative Korrelationen. Für die grauen Bereiche liegen keine Daten in dem Beobachtungsdatensatz vor.

Eine weitere Untersuchung beschäftigte sich mit dem Mehrwert von zwei Regionalmodellen gegenüber einem einzelnen Modell. Abbildung 15 zeigt die Korrelation zwischen einem Ensemble mit jeweils vier Läufen und Beobachtungen. Auf der linken Seite zusammengestellt aus vier CCLM Simulationen und auf der rechten Seite aus zwei CCLM und zwei REMO Simulationen. Besonders im ersten Jahr bietet die Verwendung zweier Modelle eine Verbesserung der Korrelation mit Beobachtungen in den meisten Teilen Europas außer dem Ostteil. Für die folgenden Jahre gibt es nur wenig Unterschiede zwischen beiden Ensembles.

## Fazit

Im Vergleich zu Beobachtungen zeigt REMO, dass es zufriedenstellende Ergebnisse für die Parameter Temperatur und Niederschlag liefert. Insbesondere wenn die Unsicherheiten der Beobachtungen mit in Betracht gezogen werden. Die kritische Betrachtung von Referenzdatensätzen ist auch ein wichtiges Ergebnis aus REDCLIP. Immer wieder hat sich im Laufe des Projektes gezeigt, dass sich unter Einbeziehung mehrerer Beobachtungsdatensätze vermeintliche Fehler im Modell als nicht gravierend oder gar nicht vorhanden herausstellten. Vielmehr stellten sich die Beobachtungsdaten hier als fehlerhaft heraus. Auf der dekadischen oder noch kürzeren Zeiträumen können sich zufallsbedingte Fehler in den Referenzdaten besonders auswirken.

In REDCLIP hat sich gezeigt, dass bei der Regionalisierung auf der dekadischen Zeitskala bekannte Korrelationsmuster im Regionalmodell erhalten bleiben. Dies betrifft auch die Vorhersagegüte (siehe z.B. Marotzke et al., 2016). Die Verwendung mehrerer Regionalmodelle bei nur einem Globalmodellsystem bietet zumindest bei der Vorhersage der Temperatur kaum Vorteile. Nur im ersten Jahr zeigt sich eine Verbesserung der Vorhersagegüte (hier Korrelation). Dies liegt in erster Linie an dem Signal, das schon vom Globalmodellsystem vorgegeben wird. Beim Niederschlag gibt es im Prinzip einen Mehrwert (Mieruch et al., 2014). Allerdings ist die Vorhersagegüte bei dem Niederschlag eher gering.

Die Re-Kalibrierung hat gezeigt, dass das Ensemble zumindest verlässlich gemacht werden kann. Verlässlich heißt hier, dass die Spannweite des Ensembles im Prinzip so korrigiert werden kann, dass sie dem zeitlich gemitteltem Fehler des Ensembledittels entspricht. Diese bisher nur in der saisonalen Vorhersage benutzte Methode kann also auch im Prinzip auch auf dekadische Vorhersagen angewandt werden. Um statistisch stabile Parameter für die Re-Kalibrierung zu erhalten, sind allerdings qualitativ hochwertige Beobachtungsdatensätze und viele Jahre an Hindcasts mit einem großen Ensemble notwendig. Die in REDCLIP gezeigte Anwendung sollte also nur als Prototyp gesehen werden.

Für die zweite Phase von MiKlip sollten aus Sicht von REDCLIP zwei Dinge beachtet werden. Erstens: Es muss sicher gestellt werden, dass genügend Rechenzeit zur Ensemblebildung zur Verfügung steht. Es sollte besonders auch Wert auf die Hincasts gelegt werden, da aus diesen Daten die Korrekturparameter für die Verbesserung der Vorhersagegüte abgeleitet werden müssen. Zweitens: Es sollten für die Erstellung nutzerrrelevanter Produkte alle verfügbaren Daten verwendet werden. Dies bedeutet, dass auch die Ergebnisse anderer dekadischer Vorhersagesysteme mit einbezogen werden sollten. Sowohl für die saisonale Vorhersage als auch für Klimaprojektionen hat die Verwendung eines Multi-Modellansatzes in der Regel immer zu robusteren und verlässlicheren Klimainformationen geführt.

## Referenzen

- University of East Anglia Climatic Research Unit (CRU)** [Phil Jones, Ian Harris] (2013): "CRU TS3.21: Climatic Research Unit (CRU) Time-Series (TS) Version 3.21 of High Resolution Gridded Data of Month-by-month Variation in Climate (Jan. 1901 - Dec. 2012)", [Internet]. NCAS British Atmospheric Data Centre, 11.3.2016. Available from [http://badc.nerc.ac.uk/view/badc.nerc.ac.uk\\_\\_ATOM\\_\\_ACTIVITY\\_0c08abfc-f2d5-11e2-a948-00163e251233](http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__ACTIVITY_0c08abfc-f2d5-11e2-a948-00163e251233) ; doi: 10.5285/D0E1585D-3417-485F-87AE-4FCECF10A992
- Eade** et al. (2014): "Do seasonal-to-decadal climate predictions underestimate the predictability of the real world?" *Geophys. Res. Lett.*, 41, 2014GL061146
- Haylock**, M.R., N. Hofstra, A.M.G. Klein Tank, E.J. Klok, P.D. Jones, M. New. (2008): "A European daily high-resolution gridded dataset of surface temperature and precipitation", *J. Geophys. Res (Atmospheres)*, 113, D20119, doi:10.1029/2008JD10201
- Hofstra**, N.; Haylock, M.; New, M. & Jones, P. D. (2009): "Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature" *J. Geophys. Res.*, 114
- Marotzke**, et al. (2016): "MiKlip – a National Research Project on Decadal Climate Prediction", *BAMS*, submitted.
- Mieruch**, S., Feldmann, H., Schädler, G., Lenz, C.-J., Kothe, S., and Kottmeier, C. (2014): "The regional MiKlip decadal forecast ensemble for Europe: the added value of downscaling", *Geosci. Model Dev.*, 7, 2983-2999, doi:10.5194/gmd-7-2983-2014.
- Müller**, W. A.; Baehr, J.; Haak, H.; Jungclaus, J. H.; Kröger, J.; Matei, D.; Notz, D.; Pohlmann, H.; von Storch, J. S. & Marotzke, J. (2012): "Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology" *Geophys. Res. Lett.*, 39, L22707
- Palmer**, T. N., R. Buizza, R. Hagedorn, A. Lawrence, M. Leutbecher, and L. Smith, (2006): "Ensemble prediction: A pedagogical perspective". *ECMWF Newsletter*, No. 106, ECMWF, Reading, United Kingdom, 10–17.
- Smith**, D. M., Cusack, S., Colman, A.W., Folland, C.K., Harris, G.R. & Murphy, J. M. (2007): "Improved surface temperature prediction for the coming decade from a global climate model", *Science* 317, 796-799.
- Sutton**, R. T.; Hodson, L. R. (2005). "Atlantic forcing of North American and European summer climate". *Science* 309 (5731): 115–118
- 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, FitzRoy Road, Exeter EX1 3PB, UK. 160pp
- Weigel** et al. (2009): "Seasonal Ensemble Forecasts: Are Recalibrated Single Models Better than Multimodels?" *Monthly Weather Review*, 137, 1460-1479
- Willmott, C.J. and K. Matsuura** (2001): "Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1900-2010)", Internet: [http://climate.geog.udel.edu/~climate/html\\_pages/README\\_ghcn\\_clim.html](http://climate.geog.udel.edu/~climate/html_pages/README_ghcn_clim.html)