

Fakultät für Umwelt und Natürliche Ressourcen
der Albert-Ludwigs-Universität Freiburg
Professur für Umweltmeteorologie

Abhängigkeit der Qualität weltweiter Niederschlagsvorhersagen von Modellart und naturräumlicher Variabilität

Bachelorarbeit zur Erlangung des akademischen Grades
Bachelor of Science in Umweltnaturwissenschaften
mit Nebenfach Meteorologie und Klimatologie

Vorgelegt von Ursula Huonder
Matrikelnummer: 4521082

Freiburg, 20.07.2020

Betreut von:

Prof. Dr. Andreas Christen
Dr. Jochen Fründ

Technisch unterstützt von:

Dr. Sebastian Schlögl, Meteoblue AG

Zusammenfassung

In der vorliegenden Arbeit wurden die 24-Stunden-Niederschlagsvorhersagen der globalen numerischen Wettervorhersagemodelle GFS, ICON, MF und NEMS und dem Reanalysedatensatz ERA5 mit den Niederschlagsmessungen der Weltorganisation für Meteorologie verglichen. Dabei wurden die Jahres-, Monats- und Tagessummen der Modellvorhersagen den Messungen gegenübergestellt, die Vorhersagemodelle untereinander verglichen und die Genauigkeit der Vorhersagen auf räumliche Muster untersucht.

Es wurde nachgewiesen, dass die Vorhersagen in äquatornahen Gebieten und in Gebieten mit zunehmender Geländehöhe größere Fehler aufweisen. Weiterhin wurde die Genauigkeit der Modellvorhersagen auf die Distanz zum Meer untersucht. Dabei zeigte sich, dass die Modelle mit größerer Distanz zur Küste genauer sind. Die Modellvorhersagen wurden außerdem in Bezug auf die mittlere monatliche Temperatur eines Standorts mit den Messungen verglichen. Dabei zeigte sich, dass die Vorhersagen bei gemäßigten Temperaturen von 0 - 15 °C am besten mit den Messungen übereinstimmen, und die Modellabweichungen von Messwerten in kälteren oder wärmeren Bereichen stark zunehmen.

Für jedes Land der Welt wurde das Modell mit der genauesten Modellvorhersage bestimmt. Die Auswahlkriterien waren die zwei Fehlermaße mittlere absolute Differenz zwischen Vorhersage und Messung und der Heidke-Skill-Score.

Vorwort und Danksagung

Schon früh war für mich klar, dass ich meine Bachelorarbeit im Gebiet meines Nebenfaches Meteorologie und Klimatologie schreiben möchte. Das Thema dieser Bachelorarbeit wurde von der Professur für Umweltmeteorologie in Zusammenarbeit mit der Meteoblue AG, Basel, ausgeschrieben und hat sofort mein Interesse geweckt. Das Bearbeiten des Themas hat mir viel Freude bereitet und ich konnte vor allem im Umgang mit großen Datenmengen viel lernen. Das wäre ohne Unterstützung nicht möglich gewesen.

An erster Stelle möchte ich Herrn Dr. Sebastian Schlögl für die Begleitung während des ganzen Arbeitsprozesses, die investierte Zeit und die vielen Inputs herzlich danken. Die Einbindung in das Team hat stark zu einer guten Struktur und motivierenden Arbeitsatmosphäre während der vom Coronavirus geprägten Zeiten beigetragen. Weiter möchte ich der Meteoblue AG danken, dass sie mir die aufbereiteten Daten und einen Arbeitsplatz zur Verfügung gestellt hat.

Besonders möchte ich mich bei meiner Begleitperson Herrn Prof. Andreas Christen für die Unterstützung bedanken.

Schließlich möchte ich meiner Familie und meinen Freunden danken. Sie haben sich die Zeit genommen, mir den Rücken zu stärken und mich mit dem Gegenlesen von Texten unterstützt.

Inhaltsverzeichnis

ZUSAMMENFASSUNG	II
VORWORT UND DANKSAGUNG	III
ABBILDUNGSVERZEICHNIS	V
TABELLENVERZEICHNIS	VI
ABKÜRZUNGSVERZEICHNIS.....	VII
1 EINLEITUNG.....	1
2 STAND DES WISSENS	2
2.1 Numerische Niederschlagsvorhersagen	2
2.2 Niederschlagsmessungen	3
2.3 Verifikation von Wettervorhersagen	4
2.4 Geografische Abhängigkeiten	5
3 FRAGESTELLUNGEN UND ZIELSETZUNG	6
4 UNTERSUCHUNGSMETHODIK	7
4.1 Datengrundlage	7
4.2 Modelle.....	8
4.2.1 GFS	8
4.2.2 ICON	8
4.2.3 MF	9
4.2.4 NEMS.....	9
4.2.5 ERA5	9
4.3 Fehlermaße.....	9
4.3.1 Analyse der Jahressummen.....	9
4.3.2 Analyse der Monatssummen	10
4.3.3 Analyse der Tagessummen	11
4.4 Variabilität zwischen den Vorhersagemodellen	12
4.5 Statistische Analyse geografischer und zeitlicher Abhängigkeiten.....	12
4.5.1 Geländehöhe und geografische Breite.....	12
4.5.2 Distanz zum Meer	12
4.5.4 Temperatur	12
4.6 Globale Darstellung der Messstationen	13
5 RESULTATE	14
5.1 Vergleich zwischen Modellvorhersage und Messung.....	14
5.1.1 Analyse der Jahressummen.....	14
5.1.2 Analyse der Monatssummen	19
5.1.3 Analyse der Tagessummen	20
5.2 Variabilität zwischen den Vorhersagemodellen	24
5.3 Statistische Analyse geografischer und zeitlicher Abhängigkeiten.....	26
5.3.1 Vorhersagegenauigkeit in Abhängigkeit von Geländehöhe und geografischer Breite	26
5.3.2 Vorhersagegenauigkeit in Abhängigkeit von der Distanz zum offenen Gewässer.....	28
5.3.4 Temperatur	29
6. DISKUSSION	31
6.1 Vergleich zwischen Modellvorhersage und Messung.....	31
6.2 Variabilität zwischen den Vorhersagemodellen	32
6.3 Geografische und zeitliche Abhängigkeiten.....	32
7. SCHLUSSFOLGERUNG	33
LITERATURVERZEICHNIS.....	34
ERKLÄRUNG.....	36

Abbildungsverzeichnis

Abb. 1: Datenverfügbarkeit der 11'627 WMO-Niederschlagsmessstationen [%] räumlich als Karte dargestellt (links) und als Histogramm mit vertikal markierten Klassengrenzen (rechts).	7
Abb. 2: Stationen, welche eine Jahressniederschlagssumme von 0 mm aufweisen und für die weitere Auswertung nicht verwendet werden.	7
Abb. 3: Darstellungsproblem des Überdeckens einzelner Stationen bei hoher Stationsdichte, verdeutlicht anhand der MAD von GFS. Links wurden zuerst die kleinen absoluten Differenzen geplottet, in der rechten Abbildung zuerst die großen.	13
Abb. 4: MAD zwischen den Modellvorhersagen (GFS, ICON, MF, NEMS) sowie des Reanalysedatensatzes (ERA5) und den gemessenen Jahresniederschlagssummen.	15
Abb. 5: MD zwischen den Modellvorhersagen (GFS, ICON, MF, NEMS) sowie des Reanalysedatensatzes (ERA5) und den gemessenen Jahresniederschlagssumme.	16
Abb. 6: Die MAD der Jahressumme des Modells GFS für die Beispielländer Schweden, China, Argentinien, Indonesien und Schweiz.	17
Abb. 7: MAD der Jahressumme der vier Vorhersagemodelle (GFS, ICON, MF, NEMS) und dem Reanalysedatensatz (ERA5) am Beispiel von Indien.	18
Abb. 8: MD der Jahressummen für die vier Vorhersagemodelle (GFS, ICON, MF, NEMS) und den Reanalysedatensatz (ERA5) in Südafrika.	18
Abb. 9: Darstellung des Modells mit dem tiefsten MAD für jedes Land unter Berücksichtigung der vier Vorhersagemodelle (GFS, ICON, MF, NEMS) (oben) und dem zusätzlichen Hinzuziehen des Reanalysedatensatzes ERA5 (unten).	19
Abb. 10: Der gemittelte HSS für globale Niederschlagssummen von mehr als 1 mm für ICON im 2° Gitternetz.	21
Abb. 11: Der größte HSS für jedes Land, gemittelt unter Berücksichtigung der vier Vorhersagemodelle (GFS, ICON, MF, NEMS) (oben) und dem zusätzlichen Hinzuziehen vom Reanalysedatensatz ERA5 (unten). Die grau markierten Länder besitzen keine Messstation.	22
Abb. 12: Darstellung des HSS der monatlichen Niederschlagsereignisse am Beispiel von GFS.	23
Abb. 13: Die maximale Differenz der Jahressummen der Vorhersagemodelle und ERA5. Die Differenzen sind in mm angegeben.	24
Abb. 14: Maximale Differenz der Monatssummen der vier Vorhersagemodelle. Links der Monat Februar, rechts August.	24
Abb. 15: Mittlere Vorhersagemodelldifferenzen der Anzahl Regentage der Vorhersagemodelle (GFS, ICON, MF und NEMS).	25
Abb. 16: MAD in Abhängigkeit von der Geländehöhe für alle Vorhersagemodelle (GFS, ICON, MF, NEMS; als Kreise) und den Reanalysedatensatz (ERA5; als Quadrat).	26
Abb. 17: MAD in Abhängigkeit von der Geländehöhe, hier am Beispiel von GFS mit Trendgerade (rot) und dem 95 % Konfidenzintervall (grau). Rechts die Formeln der Trendgeraden aller Modelle.	26
Abb. 18: MAD in Abhängigkeit von der geografischen Breite für alle Vorhersagemodelle (GFS, ICON, MF, NEMS; als Kreise) und den Reanalysedatensatz (ERA5; als Quadrat).	27
Abb. 19: MAD in Abhängigkeit von der geografischen Breite auf der Nordhalbkugel, hier am Beispiel von GFS mit Trendgerade (rot) und dem 95 %-Konfidenzintervall (grau). Rechts die Formeln der Trendgeraden aller Modelle.	27
Abb. 20: MAD in Abhängigkeit von der Distanz zum Meer für alle Vorhersagemodelle (GFS, ICON, MF, NEMS; als Kreise) und den Reanalysedatensatz (ERA5; als Quadrat) mit logarithmierter x-Achse.	28
Abb. 21: Der HSS für die Vorhersagemodelle (GFS, ICON, MF, NEMS) und den Reanalysedatensatz (ERA5) für 5 °C – Temperaturklassen und jeden Monat.	29
Abb. 22: Der HSS für die verschiedenen Temperaturklassen der Nordhalbkugel für alle Monate. Am Beispiel des Vorhersagemodells NEMS.	30

Tabellenverzeichnis

Tabelle 1: Fehlerquellen der Niederschlagsmessung (WMO, 2018, S. 221).	4
Tabelle 2: Verwendete Wettervorhersagemodelle mit Reanalysedatensatz.	8
Tabelle 3: Vier-Felder-Tafel.	11
Tabelle 4: MAD, MD, PMD, RMSE und SD der Jahressummen für GFS, ICON, MF, NEMS und ERA5 für Stationen mit mindestens 70% verfügbaren Daten (oben), mindestens 50 % verfügbaren Daten (mittig) und mindestens 30% verfügbaren Daten (unten). In grün jeweils der beste Wert.	14
Tabelle 5: Unterschiede des MAD je nach Darstellungsform. Linke Spalte jede Station einzeln gewertet, mittlere Spalte im 2° Grid und rechte Spalte im 5° Grid.	17
Tabelle 6: PMD der vorhergesagten Monatssummen zu den gemessenen Monatssummen in Prozent.	20
Tabelle 7: MAE, MBE und MPE für die Vorhersagemodelle GFS, ICON, MF, NEMS und den Reanalysedatensatz ERA5.	20
Tabelle 8: Die mittlere Anzahl RT für Kontinente mit der jeweiligen Standardabweichung der gemessenen Werte (WMO), der vier Modellvorhersagen (GFS, ICON, MF NEMS) und des Reanalysedatensatzs (ERA5).	20
Tabelle 9: HSS, FAR und POD für die Klassengrenzen 1 mm, 10 mm und 50 mm für alle Modelle (GFS, ICON, MF, NEMS) und den Reanalysedatensatz ERA5. In grün jeweils der beste Wert.	21
Tabelle 10: HSS für Tage mit > 1 mm Niederschlag für die Monate Januar bis Dezember. In grün jeweils der beste Wert.	22

Abkürzungsverzeichnis

Abkürzung	Langbezeichnung / Beschreibung	Einheit
E	„Random luck“-Wert Wertebereich: zwischen 0 und 1	
FAR	Falschalarmquote (engl. False-alarm ratio) Wertebereich: zwischen 0 und 1; bester Wert bei 0	
HSS	Heidke Skill Score Wertebereich: zwischen $-\infty$ und 1; bester Wert bei 1	
MAD	Mittlere absolute Differenz. Wertebereich: zwischen 0 und ∞ ; bester Wert bei 0	mm
MAE	Mittlerer absoluter Fehler (engl. Mean absolute error) Wertebereich: zwischen 0 und ∞ ; bester Wert bei 0	mm
MD	Mittlere Differenz Wertebereich: zwischen $-\infty$ und ∞ ; bester Wert bei 0	mm
MBE	Mittlerer Fehler (engl. Mean bias error) Wertebereich: zwischen $-\infty$ und ∞ ; bester Wert bei 0	mm
MPE	Mittlerer prozentualer Fehler (engl. Mean percentage error) Wertebereich: zwischen $-\infty$ und ∞ ; bester Wert bei 0	%
PC	Anteil richtig vorhergesagte aller eingetroffenen Werte (engl. Portion correct) Wertebereich: zwischen 0 und 1; bester Wert bei 1	
PMD	Prozentuale mittlere Differenz Wertebereich: zwischen $-\infty$ und ∞ ; bester Wert bei 0	mm
POD	Erkennungswahrscheinlichkeit (engl. Probability of detection) Wertebereich: zwischen 0 und 1; bester Wert bei 1	
RMSE	Wurzel des mittleren Quadratfehlers (engl. Root mean square error) Wertebereich: zwischen 0 und ∞ ; bester Wert bei 0	mm
SD	Standardabweichung Wertebereich: zwischen 0 und ∞ ; bester Wert bei 0	mm
DWD	Deutscher Wetter Dienst	
ECMWF	Europäisches Zentrum für Mittelfristige Wettervorhersage	
ERA5	Reanalysedatensatz des ECMWF	
GFS	Wettervorhersagemodell des U.S. Wetterdienstes (Global Forecast System)	
ICON	Wettervorhersagemodell des Deutschen Wetterdienstes (Icosahedral Nonhydrostatic Modell)	
MF	Wettervorhersagemodell des Französischen Wetterdienstes (Météo France), dort unter dem Namen ARPEGE betrieben (Action de Recherche Petite Echelle Grand Echelle)	
NEMS	Wettervorhersagemodell der Meteoblue AG	
NetCDF	Network Common Data Form, ein Format um wissenschaftliche Daten zu speichern	
NWV	Numerische Wettervorhersage	
NWVM	Numerisches Wettervorhersagemodell	
WMO	Weltorganisation für Meteorologie	

1 Einleitung

Während die Wettervorhersage seit der Einführung der Telegraphie auf der synoptischen Betrachtung des Wettergeschehens basiert (Wengenmayr et al., 2017, S. 3), haben numerische Wettervorhersagemodelle (NWVM) seit der ersten Verwendung im Jahr 1950 (Platzman, 1979) ständig an Bedeutung hinzugewonnen. Durch die ständig steigenden Rechenkapazitäten wurde die Wettervorhersage immer komplexer. Es können mehr Variablen in die Berechnungen einbezogen und viele Modellläufe durchgeführt werden. Dadurch werden global viele Wettersimulationsdaten erzeugt.

Wegen des großen finanziellen und personellen Aufwands bei der Entwicklung und dem Unterhalt betreiben nur 14 nationale Wetterdienste und wenige private Anbieter globale numerische Wettervorhersagemodelle (DWD, 2020a). Qualitätskontrollen und Optimierungen der Vorhersagen werden durch Verifikationen erreicht. Bei Verifikationen werden die Vorhersagedaten dieser Modelle mit Wettermessungen abgeglichen, um so systematische Fehler zu entdecken und die Modelle zu verbessern. Dieser Schritt ist nicht nur für die Urheber des Modells, sondern auch für alle anderen Wetterprognosedienste relevant, da Wettervorhersagen neben einzelnen Modellen oft auch auf Kombinationen mehrerer Modelle beruhen.

Um zuverlässig das beste Modell für eine bestimmte Region der Welt wählen zu können, braucht es eine genaue Analyse der Daten der Vorhersagemodelle. Dabei werden die vorhergesagten Niederschlagsdaten nachträglich mit gemessenen Daten abgeglichen. Die Wetterdienste stellen ihre Modelldaten meist öffentlich zugänglich bereit und die Meteoblue AG legt durch die Aufbereitung dieser Daten den Grundstein für weitere Analysen.

Die Vorhersage von Niederschlägen ist beispielsweise im Vergleich zur Temperaturvorhersage noch sehr ungenau. Durch die hohe ortsabhängige Variabilität des Niederschlags sind oben genannte Kontrollprozesse bei der Niederschlagsvorhersage besonders wichtig.

2 Stand des Wissens

2.1 Numerische Niederschlagsvorhersagen

Die numerische Wettervorhersage hat zum Ziel, eine Prognose des Atmosphärenzustands zu einem bestimmten Zeitpunkt und Ort zu machen. Dabei wird ein System mathematischer Gleichungen, welche nicht exakt lösbar sind, mittels Supercomputern und numerischer Mathematik angenähert.

Bei globalen numerischen Wettervorhersagemodellen (NWVM) wird die Erde mit einem dreidimensionalen Gitter umspannt. Die Skala des Modells wird durch die Maschenweite des Gitternetztes beschrieben. Die vertikale Auflösung ist in bodennahen Schichten meist größer, um bodennahe Prozesse besser beschreiben zu können (DWD, 2017). Auch ändert sich die horizontale Auflösung der globalen Modelle teilweise erheblich, je nach Standort auf dem Globus und Interesse des Modellbetreibers an diesem Standort. So besitzt beispielsweise das Vorhersagemodell des französischen Wetterdienstes eine Auflösung von 7.5 km über Frankreich, auf der diametral gegenüberliegenden Seite des Globus nur noch von 37 km (CNRM, 2014).

Für jeden Gitternetzpunkt wird ein System von mathematische Näherungsformeln für die Wettergrößen aufgestellt. Um das Modell zu berechnen, also das Gleichungssystem anzunähern, werden passende Anfangs- und Randbedingungen als Eingangsgrößen benötigt. Als Anfangsbedingungen können aktuell beobachtete Wetterdaten verwendet werden. Diese werden im Prozess der Datenassimilation mit vorangegangenen Modellläufen und klimatologischen Informationen kombiniert und dann in das Modell eingebunden. Durch kleinste Änderungen der Anfangsbedingungen können große Abweichungen bei der Vorhersage entstehen. Deshalb ist bereits die Datenassimilation ein sehr aufwändiger Prozess. Um die Genauigkeit der numerischen Wettervorhersage zu erhöhen, wird weiterhin intensiv an der Optimierung der Datenassimilation geforscht und diese weiterentwickelt.

Neben der räumlichen Auflösung hat jedes numerische Vorhersagemodell auch eine zeitliche Auflösung. Um die Modellgleichungen eines numerischen Wettervorhersagemodells zu lösen, muss deshalb neben dem Raum auch die Zeit diskretisiert werden. Das heißt, dass die Gleichungen nur für die Gitternetzpunkte und diskrete Zeitschritte berechnet werden können. Alle Prozesse, deren räumliche oder zeitliche Ausdehnung kleiner sind als die vom Modell erfassbare Ausdehnung, wie beispielsweise ein starker lokaler Schauer, werden als subskalig bezeichnet und können nicht exakt vorhergesagt werden. Da subskalige Prozesse, vor allem in Bezug auf eine Niederschlagsvorhersage, nicht vernachlässigt werden dürfen, müssen diese kleinräumigen Vorgänge parametrisiert werden.

Je nach Modell werden verschiedene Prozesse miteinbezogen und unterschiedlich gewichtet. Der Deutsche Wetterdienst nennt beispielsweise für sein NWVM Strahlung und subskaligen Niederschlag als wichtigste parametrisierte Prozesse. Weitere für die Niederschlagsvorhersage wichtige parametrisierte Vorgänge sind nachfolgend aufgezählt:

- Schauer- und Gewitterbildung
- Entstehung und Auflösung von Bewölkung
- Turbulenter Austausch von Impuls und fühlbarer und latenter Wärme zwischen Erdoberfläche und Atmosphäre
- Kleinskalige orografische Effekte
- Prozesse in den obersten Schichten des Erdbodens
- Veränderungen von Meereis
- Änderung der Temperatur von Binnenseen (DWD, 2017)

Diese Vorgänge sind zu kleinskalig, um direkt mit den Gleichungen an den Gitternetzpunkten simuliert werden zu können. Bei der Niederschlagsvorhersage müssen beispielsweise zusätzlich die Übergänge zwischen Wasserdampf, Wolkeneis, Schnee, Wolkentröpfchen, Regentropfen, Graupel und Hagel und Niederschlag wegen Übersättigung betrachtet werden (Stull, 2017, S. 751).

Mit einer Parametrisierung wird die Auswirkung der parametrisierten Prozesse auf die skaligen Zustandsvariablen des Modells abgeschätzt und durch eine Anpassung ebendieser in das Modell miteinbezogen.

Diese Abschätzung erfolgt meist über eine komplexe Modellvorstellung des Vorganges, die weitere Variablen miteinbezieht. Besonders bei der Parametrisierung von konvektivem Niederschlag zeigen sich Probleme. Bei globalen Modellen sind konvektive Ereignisse zeitlich meist nicht subskalig, während sie räumlich meist subskalig sind. (DWD, 2017).

Mit Anfangsbedingungen unter Einbezug der parametrisierten Prozesse und Randbedingungen berechnet das Modell für zukünftige Zeitschritte das Wetter für jeden Gitternetzpunkt. Die Berechnung solcher Vorhersagemodelle ist sehr rechenintensiv und setzt daher Supercomputer voraus.

Durch NWVM können besonders großräumige Wetterentwicklungen erheblich zuverlässiger vorhergesagt werden als durch die früher angewandte manuelle Vorhersage erfahrener Meteorologen. So können beispielsweise Hoch- und Tiefdruckgebiete durch numerische Wettervorhersagemodelle wesentlich genauer vorausberechnet werden. Obwohl sich auch die Vorhersagegenauigkeit kleinerer Prozesse laufend verbessert (DWD, 2020b), ist dort das Optimierungspotential noch sehr groß.

Zusammenfassend gibt es drei stark limitierende Faktoren in der numerischen Wettervorhersage:

1. Rechenleistung: Die Vorhersage wird umso genauer, je feinmaschiger das Gitternetz ist und je mehr und detailliertere Näherungsformeln aus wetterrelevanten Variablen berechnet werden. Die Leistungsfähigkeit eines Modells nimmt somit mit der steigenden Rechenleistung der Supercomputer zu (Wengenmayr et al., 2017).
2. Anfangsbedingungen: Trotz der großen Anzahl an Wettermessstationen ist der aktuelle Zustand der Atmosphäre nie vollständig bekannt. Dadurch entsteht eine Unsicherheit bei der Datenassimilation.
3. Parametrisierung: Effektive physikalische Prozesse in der Atmosphäre müssen in den Vorhersagemodellen vereinfacht abgebildet werden. Durch die Parametrisierung kommt es zu ungenau beschriebenen Prozessen, welche durch Rückkopplungen in weiteren Rechenschritten verstärkt werden können.

Die Art des Niederschlags wirkt sich ebenfalls auf die Genauigkeit der Vorhersage aus. Es wird unterschieden zwischen stratiformem, orographischem und konvektivem Niederschlag. Stratiformer Niederschlag fällt über größere Zeiträume und Gebiete. Da stratiformer Niederschlag hauptsächlich durch Fronten verursacht wird, welche durch numerische Wettervorhersage mit hoher Richtigkeit vorhergesagt werden können, ist auch die entsprechende Niederschlagsvorhersage mit höherer Wahrscheinlichkeit richtig. Konvektiver und orographischer Niederschlag entstehen durch das Heben oder Senken von Luftmassen. Diese Prozesse sind meist kleinräumig und bei globalen Modellen oft subskalig. Bei konvektivem Niederschlag wird die vertikale Luftbewegung durch Temperaturänderungen ausgelöst, bei orographischem Niederschlag durch das Relief. Durch die geografische Kleinräumigkeit und schwankende Intensität fehlt eine gute, genug hochaufgelöste Datengrundlage, was zu fehlerhaften Annahmen bei der Parametrisierung und somit höheren Fehlern bei der Vorhersage führen kann. (Lin et al., 2001; Papadopoulos et al., 2005)

2.2 Niederschlagsmessungen

Niederschlag wird in der Einheit mm gemessen. Die Einheit meint damit mm Wassersäule, wobei 1 mm auch 1 l/m² entspricht. Da Niederschlag große lokale Unterschiede aufweist, ist ein dichtes einheitliches Messnetz wichtig. Als Kriterien für einen optimalen Messstandort empfiehlt die Weltorganisation für Meteorologie (WMO) für ein Niederschlagsmessgerät, im Gegensatz zu den meisten anderen meteorologischen Messgeräten, ein gewisses Maß an Windschutz. Die Messhöhe sollte möglichst ebenerdig sein, aber geschützt vor möglichem Herein- oder Herausspritzen von Regentropfen aus der Umgebung (WMO, 2018, S. 234).

Die Standards für die Niederschlagsmessgeräte werden von den nationalen Wetterdiensten verfügt und sind nicht einheitlich von der WMO festgelegt, deshalb gibt es mehrere verwendete Messmethoden. Das konventionelle Standardmessgerät ist der Niederschlagsmesser nach Hellmann (Häckel, 2012, S. 384). Dabei handelt es sich um ein zylinderförmiges Gefäß mit einem Trichter, welcher den Niederschlag in den im unteren Teil des Gefäßes angebrachten Sammler leitet. Das Abführen des Wassers in den Sammler vermindert das

Herausspritzen von Tropfen und reduziert die Verdunstung. Es gibt Niederschlagsmesser, bei welchen die Niederschlagshöhe vom Volumen abgeleitet oder abgelesen wird, und solche, die mit dem Gewicht arbeiten. Eine häufig verwendete Methode ist die Messung mit Kippwaagen. Dabei sind unter dem Trichter zwei Schalen angebracht, die beim Erreichen eines bestimmten Gewichtes von der einen Seite zu der anderen umschlagen. Durch das Zählen der Kippvorgänge kann die Niederschlagshöhe berechnet werden.

Niederschlagsmessungen weisen große Fehlermöglichkeiten auf. In der folgenden Tabelle sind die häufigsten Messfehler aufgeführt.

Tabelle 1: Fehlerquellen der Niederschlagsmessung (WMO, 2018, S. 221).

Fehlerquelle	Fehlerausmaß
Systematischer Windfehler	2 - 10% bei Regen, 10 - 20 % bei Schnee
Benetzungsfehler	Sowohl an den Wänden der Kollektoren als auch beim Leeren der Kollektoren: 2 - 15 % im Sommer, 1 - 8 % im Winter
Verdunstungsfehler	0 - 4 %
Technische Fehler	5 - 15 %, bei Starkniederschlägen noch höhere Werte
Spritzfehler	1 - 2 %
Spurenniederschläge	Niederschlagsmenge liegt unter der Messauflösung des Messgeräts (Koblinger, 2013)

Der Vorteil der konventionellen Niederschlagsmessung ist, dass der effektiv am Boden angekommene Niederschlag nur mit Abweichungen resultierend aus den Fehlerquellen in Tabelle 1 gemessen wird. Der Nachteil ist, dass damit nur punktförmige Messdaten vorliegen, womit nur schlecht auf ein ganzes Gebiet interpoliert werden kann und somit keine lokalen Unterschiede erfasst werden.

Eine weitere Möglichkeit der Niederschlagsmessung ist der Wetterradar. Es sendet Impulse mit einer bestimmten Wellenlänge aus und misst die von Niederschlag reflektierten Teil der Strahlung. Die Entfernung des Niederschlags wird anhand der Laufzeit des Impulses berechnet. Anhand der Rückstreuung der ausgesandten Radarstrahlen können flächendeckende und dreidimensionale Niederschlagsdaten aufgenommen werden (Wengenmayr et al., 2017, S. 8). Auch die Radarmessung weist mögliche Fehlerquellen auf. So entsteht bei wachsender Entfernung zum Radargerät eine Ungenauigkeit durch die Ausbreitung des Strahles. Durch variable Temperaturschichten kann es zu anormaler Ausbreitung der ausgesendeten Strahlung kommen. Bei Krümmung von der Erde weg wird beispielsweise bodennaher Niederschlag nicht wahrgenommen und bei starker Krümmung zur Erde hin, wird Strahlung von der Erdoberfläche zurückgestreut und somit zu viel Niederschlag erfasst. Ein weiteres Problem kann die Überschätzung des Niederschlags durch Verdunstung sein, wenn Niederschlag aus hoher Bewölkung in darunterliegende trockenere Luftschichten fällt. Der Vorteil der Niederschlagsmessung mittels Radar ist, dass hochreichende Gewitterzellen erkannt werden können und somit Unwetterwarnungen und kurzfristige Vorhersageanpassungen gegeben werden können.

Neben Radardaten sind auch Satellitendaten ein wichtiger Teil der numerischen Wettervorhersage. Geostationäre Wettersatelliten, welche in 36'000 km Höhe fliegen, sind dabei besonders wichtig. Sie liefern bis auf die Polregionen überwiegend flächendeckende Daten. Eine höhere Auflösung erreichen polarumlaufende Wettersatelliten, welche nur zwischen 800 und 900 km Höhe fliegen. (Wengenmayr et al., 2017, S. 8).

2.3 Verifikation von Wettervorhersagen

Um die Qualität der Vorhersagen hoch zu halten und stetig zu verbessern werden ständig Verifikationen der Modelle durchgeführt. Durch eine objektive Betrachtung der Vorhersagen können systematische Fehler erkannt werden. Um eine umfassende Verifikation durchzuführen braucht es verschiedene Fehlermaße und statistische Methoden, da NWVM zu komplex sind, um sie anhand eines Wertes zu beurteilen. Je nachdem ob

kontinuierliche, kategorische Variablen oder Wahrscheinlichkeitsaussagen betrachtet werden, müssen verschiedene Qualitätsmaße verwendet werden.

Für kontinuierliche Variablen, wie beispielsweise die Temperatur, werden kontinuierliche Maßzahlen wie der mittlere Fehler oder die Standardabweichung verwendet. Bei kategorischen Variablen, wie dem Auftreten von Gewittern, werden andere Fehlermaße verwendet, welche häufig als Scores bezeichnet werden. So werden etwa der Brier Score oder der Heidke-Skill-Score häufig verwendet. Da auch alle kontinuierlichen Variablen durch das Einführen von Schwellenwerten kategorisiert werden können, ist es möglich, dass auch die Niederschlagsvorhersage als kategorische Variable mit Scores beschrieben werden. (Damrath, 2002).

2.4 Geografische Abhängigkeiten

Die Abhängigkeit der Vorhersagegenauigkeit von den geografischen Gegebenheiten eines Standortes wurde in einigen Studien aufgearbeitet. So beschreibt das Paper von Lin et al. (2001) Voraussetzungen für starken orographischen Niederschlag in einem bestimmten Gebiet. Es wird versucht, durch das Verstehen orographischer Niederschläge das Auftreten ebendieser besser vorherzusagen, da orographisch induzierte Niederschlagsereignisse bisher eines der größten Probleme der numerischen Niederschlagsvorhersage sind. In Gebirgsregionen wird der Niederschlag bei großen Niederschlagsintensitäten systematisch unterschätzt, wohingegen bei kleinen Niederschlagsintensitäten eine systematische Überschätzung vorliegt (Isotta et al., 2014).

In einer Studie von Zhang et al. (2016) wird die Niederschlagsverteilung abhängig von der Geländehöhe der Messstationen und der Distanz zur Küste beschrieben. Es wird festgehalten, dass dieses Themengebiet von der Wissenschaft bisher wenig Beachtung fand, da allgemein akzeptiert wird, dass sich durch steigende Luftmassen an Gebirgen die relative Luftfeuchtigkeit erhöht. Dies kann zu Wolken und Niederschlag führen. Ausserdem erfahren Küstengebiete durch den starken maritimen Einfluss mehr Niederschlag als Gebiete im Landesinnern. In einer lokalen Studie für die Niederlande wurden die küstennahen Niederschläge untersucht. Es zeigte sich, dass die betrachteten Modelle einen zu starken Unterschied zwischen Land- und Meer berechneten und dadurch die küstennahen Modellsimulationen fehlerhaft waren (Attema & Lenderink, 2014).

3 Fragestellungen und Zielsetzung

In der vorliegenden Bachelorarbeit werden folgenden Fragestellungen bearbeitet:

- (1) Wie groß sind die absoluten und relativen Fehler der einzelnen Vorhersagemodelle für jährliche und monatliche Niederschlagssummen?
- (2) In welchen Gebieten wird die jährliche Niederschlagssumme am besten vorhergesagt?
- (3) In welchen Gebieten ist die Variabilität der einzelnen Modelle am größten? Sind in diesen Gebieten auch die Modellfehler im Vergleich zur Messung am größten?
- (4) Gibt es räumliche Muster, die für jede Region der Welt ein eindeutig bestes Vorhersagemodell empfehlen?

Das Ziel dieser Arbeit ist die Beantwortung der Fragestellungen durch das Vergleichen und Herausarbeiten der Unterschiede globaler NWVM. Dazu werden die absoluten und relativen Fehler der einzelnen Vorhersagemodelle für jährliche, monatliche und tägliche Niederschlagssummen bestimmt. Anhand der Geländehöhe, der geografischen Breite und der Distanz zum Meer werden Gebiete gesucht, in welchen die Variabilität der Modelle am größten und am kleinsten ist, sowie für jedes Land der Welt ein bestes Vorhersagemodell empfohlen.

Die vorliegende Arbeit befasst sich zudem mit dem Vergleich von Niederschlagsvorhersagen und Niederschlagsmessungen. Mögliche Ursachen für die Ungenauigkeiten der Niederschlagsprognosen werden nur in Bezug auf die geografischen Abhängigkeiten besprochen.

4 Untersuchungsmethodik

4.1 Datengrundlage

Als Grundlage dieser Arbeit dienen globale Niederschlagsmessungen der Weltorganisation für Meteorologie (WMO) des Jahres 2019. Die Tagesmesswerte der Niederschläge liegen in einer NetCDF-Datei vor. Die Originaldatei verzeichnet die Tagesmesswerte von 12'361 Niederschlagsmessstationen der WMO. Hierbei kamen die Koordinaten von 734 Stationen doppelt vor und mussten eliminiert werden. Somit sind für das Jahr 2019 im Messnetz der WMO Daten von 11'627 Stationen verfügbar. Die Verteilung der Stationen und die zeitliche Verfügbarkeit der Niederschlagsmessungen in Prozent ist in der folgenden Karte dargestellt (Abb. 1). Im Histogramm erkennt man, dass die meisten Stationen nahezu vollständige Zeitreihen vorweisen können.

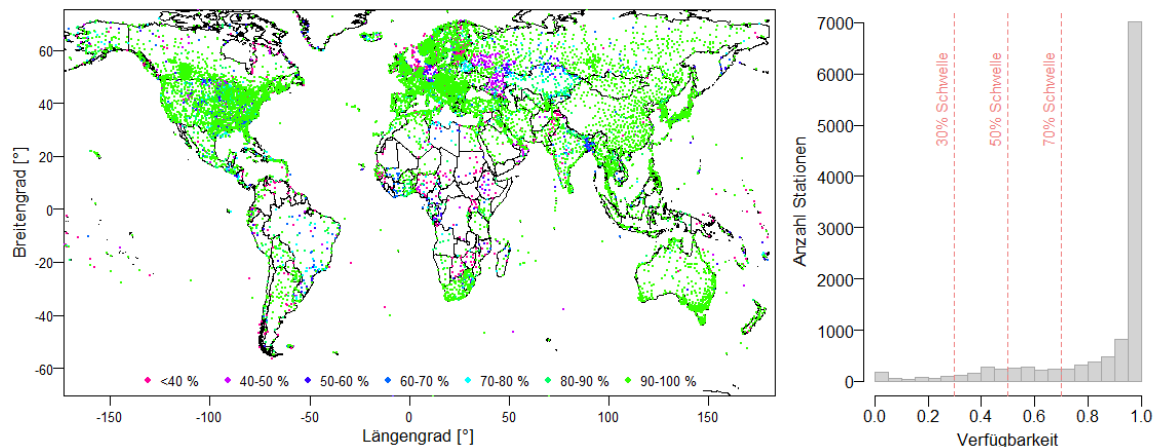


Abb. 1: Datenverfügbarkeit der 11'627 WMO-Niederschlagsmessstationen [%] räumlich als Karte dargestellt (links) und als Histogramm mit vertikal markierten Klassengrenzen (rechts).

Die Stationen wurden auf ihre Datenverfügbarkeit überprüft und die Plausibilität der Daten analysiert. Alle Stationen mit einer Datenverfügbarkeit unter 30 % wurden für die Analyse nicht berücksichtigt. Auffällig viele Messstationen haben ohne plausible Erklärung 0 mm Jahresniederschlag gemessen (Abb. 2). Da alle Vorhersagemodelle Niederschlag vorhergesagt haben und oft auch nahe gelegene Stationen Niederschlag gemessen haben, wurden diese Stationen ebenfalls nicht berücksichtigt.

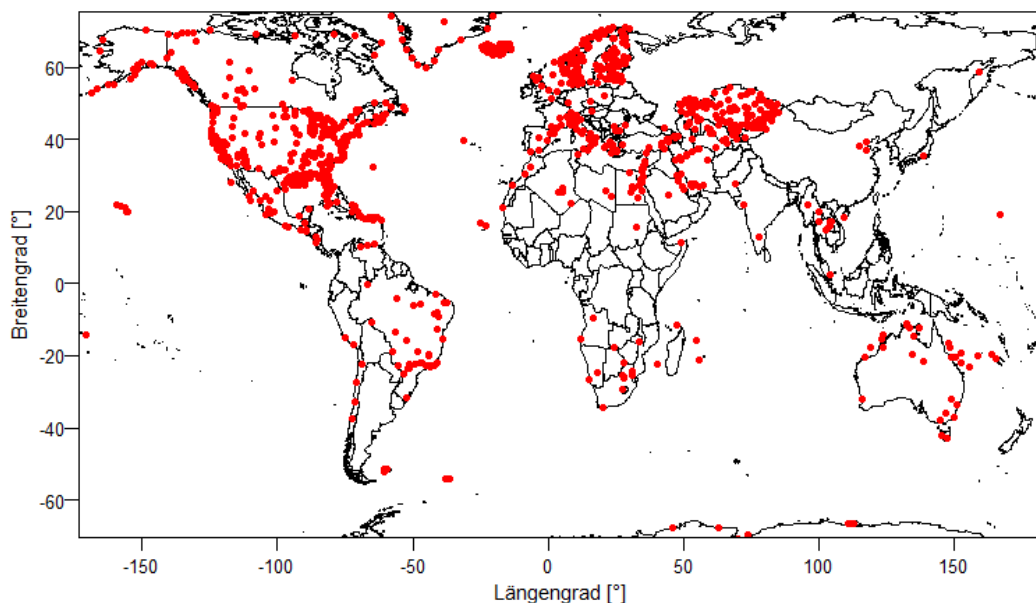


Abb. 2: Stationen, welche eine Jahresniederschlagssumme von 0 mm aufweisen und für die weitere Auswertung nicht verwendet werden.

Die restlichen Daten wurden in den Klassen mit Datenverfügbarkeit von mindestens 30 % (8741 Stationen), mindestens 50 % (8480 Stationen) und mindestens 70 % (8112 Stationen) analysiert.

4.2 Modelle

In dieser Arbeit wurden die Vorhersagedatensätze für Niederschlag der in Tabelle 2 aufgelisteten Modelle verwendet.

Tabelle 2: Verwendete Wettervorhersagemodelle mit Reanalysedatensatz.

Kurzname	Modellname	Horizontale Auflösung	Urheber
GFS	GFS40	40 km	Modell des U.S. Wetterdienstes NOAA
ICON	ICON13	13 km	Modell des Deutschen Wetterdienstes
MF	MFGLOBAL	40 km	Modell des französischen Wetterdienstes Météo-France (ARPEGE40 = MFGLOBAL)
NEMS	NEMSGLOBAL	30 km	Modell von Meteoblue AG
ERA5	ERA5	30 km	Reanalysedatensatz vom "European Centre for Medium-Range Weather Forecasts" (ECMWF)

Bei den Modelldaten handelt es sich um die 24-Stunden Niederschlagsvorhersage der Modelle. Die Daten dieser Modelle wurden von der Meteoblue AG vorbereitet und zur Verfügung gestellt. Für jedes Modell wurden die täglichen Niederschlagssummen für das Jahr 2019 an den Koordinaten der WMO-Messstationen heruntergeladen. Neben dem Längen- und Breitengrad und den 365 Tagessummen beinhalten die Datensätze auch die Höhe der Station über dem Meeresspiegel. In den folgenden Unterkapiteln werden die einzelnen Modelle genauer beschrieben.

4.2.1 GFS

Das Global Forecast System (GFS) ist ein globales numerisches Wettervorhersagemodell des Wetterdienstes der Vereinigten Staaten. Die horizontale Auflösung des Modells beträgt 28 km, die vertikale Auflösung besteht aus 64 Schichten. Das Modell wird viermal täglich berechnet (NCEP, 2020b, S. 3). Es werden dabei Vorhersagen von bis zu 16 Tagen gemacht, wobei die horizontale Auflösung dabei auf bis zu 70 km zurück geht (NCEP, 2020a). Der für diese Arbeit verwendete Datensatz hat eine Auflösung von 40 km. Das Modell ist aus vier einzelnen Modellen gekoppelt: Einem Atmosphärenmodell, einem Ozeanmodell, einem Land- und Bodenmodell sowie einem Meereissystem.

4.2.2 ICON

ICON ist ein globales numerisches Wettervorhersagemodell (NWP) des Deutschen Wetterdienstes in Zusammenarbeit mit dem Max-Planck-Institut für Meteorologie. Die Besonderheit an ICON ist der Grundaufbau als Ikosaeder, also einem aus gleichseitigen Dreiecken bestehenden Gitternetz. Um die gewünschte Auflösung zu erhalten, werden die Dreiecke in kleinere Dreiecke unterteilt. Bei der hier verwendeten Version hat jedes Dreieck eine Fläche von 173 km², was einer ungefähren horizontalen Auflösung von 13 km entspricht. Vertikal ist das Modell in 90 Schichten unterteilt. Die wichtigsten prognostischen Variablen sind die Luftdichte und die virtuelle potenzielle Temperatur, die horizontale und vertikale Windgeschwindigkeit, die Luftfeuchtigkeit, Wolkenwasser und -eis, Regen und Schnee. Da diabatische Prozesse wie Strahlung, Turbulenz und Niederschlag für gute Vorhersagen eine größere Auflösung brauchen, werden diese parametrisiert. (DWD, 2020a).

ICON hat die Möglichkeit, durch weitere Unterteilung des Dreieckgitters eine verfeinerte Domäne in der globalen Simulation zu berechnen. So ist die Unterteilung über Europa höher aufgelöst. (Reinert et al., 2020).

4.2.3 MF

MF global, auch ARPEGE40, ist ein globales Wettervorhersagemodell des französischen Wetterdienstes Météo France. Die horizontale Auflösung beträgt zwischen 7.5 km über Frankreich bis zu 37 km in der Region um Neuseeland. Die Anfangsbedingungen werden aus verschiedenen Quellen gewonnen, darunter Flugzeug- und Schiffsmessungen, Bodenstationen, Radar- und Fernerkundungsdaten. Die prognostischen Variablen, welche in das Modell einfließen, sind Wind, Temperatur, Luftfeuchtigkeit, Hydrometeore und die turbulente kinetische Energie. Parametrisiert werden Strahlung und Oberflächenprozesse. Das Modell wird viermal täglich berechnet. (CNRM, 2014).

4.2.4 NEMS

NEMS (NOAA Environment Monitoring System) ist ein globales Wettervorhersagemodell der Meteoblue AG. Es basiert ursprünglich auf dem NMM-System des Amerikanischen Wetterdienstes und wurde dort entwickelt. Es wird seit 2014 von der Meteoblue AG betrieben. Die horizontale Auflösung beträgt 40 km. Vertikal unterteilt das Modell in 60 atmosphärische Schichten. Es werden zwei Modellläufe pro Tag gerechnet. (Meteoblue AG, 2020a).

4.2.5 ERA5

ERA5 ist im Gegensatz zu den bereits beschriebenen Modellen kein Vorhersagemodell, sondern ein Reanalysemodell, von welchem erst nachträglich Daten vorliegen. Bei einem Reanalysemodell werden Simulationsdaten mit konkreten Messungen und Beobachtungen abgeglichen. Damit wird eine genaue Beschreibung des vergangenen Wetters möglich (Meteoblue AG, 2020b). So können flächendeckende Datensätze mit homogener räumlicher und zeitlicher Auflösung erstellt werden (Sun et al., 2018). ERA5 basiert auf dem IFS (Integrated Forecast System) CY41R des ECMWF. Es besitzt vertikal 137 Schichten und eine horizontale Auflösung von 31 km. Die zeitliche Auflösung beträgt 1 Stunde. (Hersbach et al., 2020).

4.3 Fehlermaße

Die Modellgenauigkeit der Wettervorhersagemodelle wurde mit verschiedenen Fehlermaßen bestimmt. Zum einen wird mit der absoluten Vorhersagegenauigkeit die Differenz zwischen den Niederschlagssummen betrachtet. Zum anderen wurde die relative Vorhersagegenauigkeit untersucht, welche die Niederschlagssummen in Relation zur mittleren Jahresniederschlagssumme setzt. Im Folgenden werden die Niederschlagsvorhersagen in verschiedenen Zeitskalen betrachtet.

4.3.1 Analyse der Jahressummen

In der ersten Analyse wurden die Vorhersagen der Jahressummen betrachtet. Dabei wurden die Messwerte der WMO-Stationen als Referenzwerte angesehen. Um die Jahressummen der Modellvorhersagen zu bilden, wurden jeweils nur die Tage verwendet, bei welchen auch ein Messwert der WMO-Messstation vorliegt. Liegt bei einer WMO-Station kein Messwert für einen Tag vor, wird die Tagessumme nicht mit aufsummiert.

Die Jahressummen der Modelle (x_{Modell}) jeder Station wurden mit den gemessenen Jahressummen (x_{WMO}) verglichen. Die Unterschiede wurden mit den folgenden Fehlermaßen bestimmt:

Das Mittel der absoluten Differenz (MAD), bei welchem sowohl die Unterschätzung wie auch die Überschätzung des Niederschlags positiv in den Fehler einfließen und sich so nicht aufheben können. Dabei wird über alle Stationen (n) summiert. Hierbei handelt es sich um eine räumliche Variante des für zeitliche Schritte definierten Mittleren Absoluten Fehlers, welcher in Kapitel 4.3.2 eingeführt wird.

$$MAD = \frac{1}{n} \sum_{i=1}^n |x_{Modell_i} - x_{WMO_i}| \quad (1)$$

Die einfachste Möglichkeit, den Fehler einer Vorhersage zu beziffern ist das Mittel der Differenz zwischen Vorhersage und Messwert (MD). Die MD zeigt an, ob ein Modell die Jahresniederschlagssumme unter- oder überschätzt.

$$MD = \frac{1}{n} \sum_{i=1}^n (x_{Modell_i} - x_{WMO_i}) \quad (2)$$

Das prozentuale Mittel der Differenz (PMD) gibt den relativen Fehler zur mittleren global gemessenen Jahressumme ($\bar{x}_{WMO}$) in Prozent an.

$$PMD = \frac{100\%}{n} \sum_{i=1}^n \frac{x_{Modell_i} - x_{WMO_i}}{\bar{x}_{WMO}} \quad (3)$$

Die Wurzel des mittleren Quadratfehlers (RMSE, engl. Root mean square error) ist ein weiteres oft benutztes Fehlermaß. Durch die Quadrierung reagiert es stärker auf wenige sehr große Abweichungen zwischen Messung und Modell als viele kleinere Abweichungen.

$$RMSE = \sqrt{(x_{Modell} - x_{WMO})^2} \quad (4)$$

Die Standardabweichung wurde für die Differenz zwischen Messung und Modell gebildet.

$$SD = \sqrt{\frac{\sum_{i=1}^n \left((x_{Modell_i} - x_{WMO_i}) - (\bar{x}_{Modell} - \bar{x}_{WMO}) \right)^2}{n - 1}} \quad (5)$$

Um die Robustheit der Vorhersagen zu testen, wurden diese Fehlermaße für 30 %, 50 % und 70 % verfügbare Daten bestimmt und anhand der Unterschiede diejenigen Stationen gewählt, welche für die weitere Analyse betrachtet werden.

Die MAD, die MD und PMD wurden für jedes der vier Vorhersagemodelle und den Reanalysedatensatz sowohl global als auch für jedes Land berechnet und als Karten dargestellt.

4.3.2 Analyse der Monatssummen

Für das Nachvollziehen vieler hydrologischer Prozesse und beispielsweise auch für die Agrarwirtschaft ist eine zeitlich höher aufgelöste Analyse notwendig. So wurden die Monatssummen für Modell und Messung berechnet, um die zeitlichen Abhängigkeiten im Verlaufe eines Jahres zu untersuchen. Als Fehlermaße wurden erneut MAD, MD, PMD, RMSE und SD verwendet.

Im Gegensatz zur räumlichen Mittelung über die Stationen, wie in Kapitel 4.3.1, wird für die Monatssummen auch zeitlich über die zwölf Monate gemittelt. Die folgende Formel zeigt den mittleren absoluten Fehler (MAE, engl. Mean absolute error), bei welchem sowohl die Unterschätzung als auch die Überschätzung des Niederschlags positiv in den Fehler einfließen und sich so nicht aufheben können. Dabei wird über die Monate summiert.

$$MAE = \frac{1}{12} \sum_{i=1}^{12} |x_{Modell,i} - \bar{x}_{Modell,i}| \quad (6)$$

Die einfachste Methode, den Fehler einer Vorhersage zu beziffern, ist die Berechnung des mittleren Fehlers (MBE, engl. Mean bias error). Dieser zeigt an, ob ein Modell die Jahresniederschlagssumme unter- oder überschätzt.

$$MBE = \frac{1}{12} \sum_{i=1}^{12} (x_{Modell,i} - x_{WMO,i}) \quad (7)$$

Der mittlere prozentuale Fehler (MPE) gibt den relativen Fehler zur mittleren global gemessenen Jahressumme ($\overline{x_{WMO}}$) in Prozent an.

$$MPE = \frac{100\%}{12} \sum_{i=1}^{12} \frac{x_{Modell,i} - x_{WMO,i}}{\overline{x_{WMO,i}}} \quad (8)$$

4.3.3 Analyse der Tagessummen

Die Methodik und Fehlermaße des Kapitels 4.3.3 *Analyse der Tagessummen* wurden in wesentlichen Teilen von Stull (2017, S. 780 ff) übernommen. Neben der Bestimmung der Exaktheit des Modells in Bezug auf Jahres- und Monatssummen sind bei 24-Stundenvorhersagen vor allem die täglichen Niederschlagssummen von großer Bedeutung für die Verifikation von Modellen. Als Regentage wurden in dieser Arbeit alle Tage mit einer Niederschlagssumme grösser als 1 mm definiert. So wurden die Tagessummen als kategorische Variablen analysiert, im Gegensatz zu den Monats- und Jahressummen, welche als kontinuierliche Variablen analysiert wurden. Die Vorhersagen der Tagessummen werden nebst den Regentagen in den zwei weiteren Intensitätsklassen betrachtet, welche als Events von mindestens 10 mm oder 50 mm Niederschlag pro Tag definiert sind. Dabei wird für jede Klasse ausgewertet, ob eine entsprechende Menge an Niederschlag vorhergesagt war und ob diese gemessen wurde. Dies lässt sich als Vier-Felder-Tafel darstellen:

Tabelle 3: Vier-Felder-Tafel.

		Messung	
		Ja	Nein
Vorhersage	Ja	a „Hit“	b „False Alarm“
	Nein	c „Miss“	d „Correct Rejection“

Um die Genauigkeit der Vorhersage der Events zu bestimmen, werden die Bestimmtheitsmaße der Erkennungswahrscheinlichkeit (Probability of Detection, POD), der Falschalarmquote (False-alarm ratio, FAR) und der Heidke-Skill-Score (HSS) berechnet.

Die POD gibt die Wahrscheinlichkeit an, dass ein beobachtetes Niederschlagsereignis auch vorhergesagt wurde. Der Wertebereich liegt zwischen 0 und 1, mit einem besten Wert bei 1.

$$POD = \frac{a}{a + c} \quad (9)$$

Weiterhin wird die FAR berechnet, welche das Verhältnis des Nicht-Eintreffens von allen vorhergesagten Events beschreibt. Der Wertebereich liegt zwischen 0 und 1, mit einem besten Wert bei 0.

$$FAR = \frac{b}{a + b} \quad (10)$$

Diese Bestimmtheitsmaße weisen keinen Skill vor, was sich besonders zeigt, wenn es an einem Standort immer oder nie regnen würde. Ein Skill bezieht die Überlegenheit einer Vorhersage gegenüber alternativen Prognosen. So wird im Folgenden der HSS berechnet, welcher die Überlegenheit zu zufällig richtigen Abschätzungen aufzeigt. Der HSS berechnet sich aus dem Anteil der richtig vorhergesagten von allen Werten (engl. Portion correct, PC) und dem „random luck“-Wert (E), welche sich auch aus den Werten der Vier-Felder-Tafel berechnen lassen.

$$PC = \frac{a + d}{n} \quad (11)$$

$$E = \left(\frac{a+b}{n}\right) \cdot \left(\frac{a+c}{n}\right) + \left(\frac{d+b}{n}\right) \cdot \left(\frac{d+c}{n}\right) \quad (12)$$

Wobei $n = a + b + c + d$, also der Anzahl der Messstationen entspricht.

Der HSS wird wie folgt berechnet:

$$HSS = \frac{PC - E}{1 - E} \quad (13)$$

Liegt der HSS nahe bei 1, ist die Vorhersage gut. Liegt der Wert hingegen bei 0, liegt kein Vorhersageskill mehr vor.

Der HSS, die POD und die FAR wurden für die drei Intensitätsklassen 1 mm, 10 mm und 50 mm berechnet und der HSS als Karte dargestellt.

4.4 Variabilität zwischen den Vorhersagemodellen

Um die Variabilität der vier Wettervorhersagemodelle zu bestimmen, wurden die Jahressummen, die Monatssummen und die Tagessummen der Niederschlagsvorhersagen betrachtet. Bei den Jahres- und Monatssummen wurde jeweils die Differenz zwischen dem Modell mit der höchsten und dem der niedrigsten Niederschlagsvorhersage bestimmt, bei den täglichen Events die Spannweite der Anzahl Regentage.

Die Differenzen zwischen den Modellen wurden als Karten dargestellt und die einzelnen Monate untereinander verglichen.

4.5 Statistische Analyse geografischer und zeitlicher Abhängigkeiten

Die naturräumliche Variabilität wurde in dieser Arbeit anhand der Geländehöhe, der geografischen Breite und der Distanz zum Meer betrachtet. Weiter wurden die Vorhersagemodelle auf Unterschiede und Gemeinsamkeiten von Gebieten mit gleicher durchschnittlicher Lufttemperatur analysiert.

4.5.1 Geländehöhe und geografische Breite

Um die räumliche Gegebenheit eines Standortes zu beschreiben, wurden die Geländehöhe und die geografische Breite analysiert. Dazu wurden die Stationen nach ihrer Höhe in Klassen von 0-200, 200-400, 400-600, 600-800, 800-1000, 1000-1500, 1500-2000, 2000-3000 m über Meer eingeteilt und die MAD für die Stationen in den entsprechenden Höhenstufen gemittelt. Die vier Vorhersagemodelle (GFS, ICON, MF, NEMS) und der Reanalysedatensatz (ERA5) wurden analysiert und auf eine mögliche Korrelation untersucht. Dabei wurde eine lineare Regression mit 95 % Konfidenzintervall durchgeführt, mit Programmteilen aus Dormann (2013).

Das gleiche Vorgehen wurde auch bei der geografischen Breite angewendet. Dabei sind die Breitengrade auf der Südhalbkugel als negative Werte angegeben und jene auf der Nordhalbkugel als positive. Die geografische Breite wurde dabei in Klassen zwischen -90, -60, -40, 23.5, 0, 23.5, 40, 60 und 90 ° eingeteilt

4.5.2 Distanz zum Meer

Um die Maritimität der Stationen zu quantifizieren, wurde die Distanz zu den Meeresgrenzen betrachtet. Die Standorte wurden in Klassen von 0-20, 20-50, 50-100, 100-500 und 500-1000 km Entfernung zum Meer eingeteilt. Die Daten zu der Lage der Küsten wurden aus dem R Paket `rworldmap` (South, 2011) im Datensatz «`coastsCoarse`» genommen. Anschließend wurden die mittleren absoluten, relativen und prozentualen Differenzen für jede Klasse berechnet.

4.5.4 Temperatur

Die Modellgenauigkeit wurde auf die Abhängigkeit zur Lufttemperatur untersucht. Dazu wurden die Messstationen nach der mittleren Monatstemperatur von ERA5 in Klassen eingeteilt. Die Klassen gehen in Fünfer-Schritten von -20 °C bis +30 °C. Der HSS für die Temperaturklassen wurden für jeden Monat dargestellt.

Um das sich gegenseitige Aufheben der Einflüsse von Sommer und Winter von der Nord- und der Südhalbkugel zu verhindern, wurden die Hemisphären getrennt betrachtet. Eine Temperaturklasse wurde dann verwendet, wenn mindestens fünf Stationen im gleichen Monat in diese Temperaturklasse fallen.

Dieselbe Analyse wurde weiter aufgeteilt und nach Anzahl Regentage pro Monat betrachtet. Dazu wurde der HSS für Stationen mit 0-5, 5-10, 10-20, >20 Regentage pro Monat in den verschiedenen Temperaturklassen betrachtet.

4.6 Globale Darstellung der Messstationen

Durch die hohe Dichte an Niederschlagsmessstationen ergeben sich durch das Überlappen einzelner Stationen Probleme bei der grafischen Darstellung auf den Karten. Die Fehlerklasse, die als letztes geplottet wird, erhält mehr Gewicht, da sie andere Fehlerklassen verdeckt. In der folgenden Abbildung ist das Problem veranschaulicht. Vor allem in den Bereichen Nordamerikas und Europas wird das Problem aufgrund der hohen Stationsdichte gut ersichtlich.

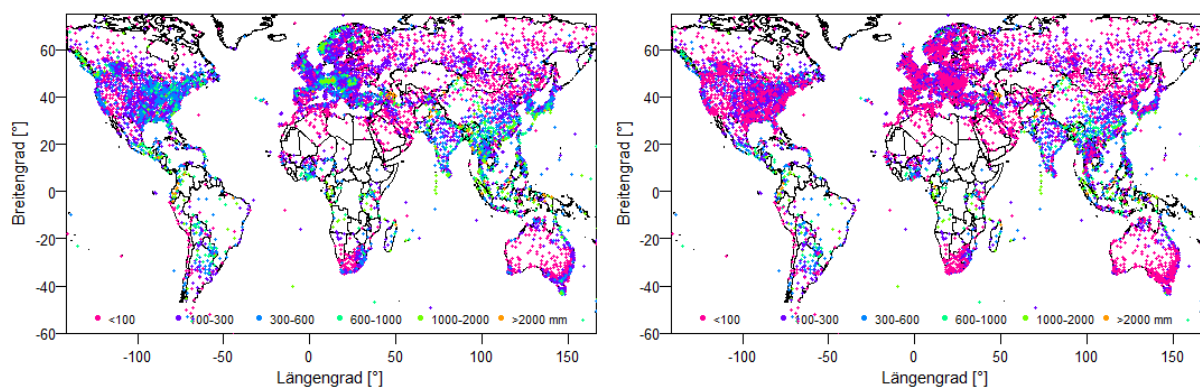


Abb. 3: Darstellungsproblem des Überdeckens einzelner Stationen bei hoher Stationsdichte, verdeutlicht anhand der MAD von GFS. Links wurden zuerst die kleinen absoluten Differenzen geplottet, in der rechten Abbildung zuerst die großen.

Um das Überlappen der Stationen zu verringern, wurden zwei Lösungsmethoden angewendet. Die Stationen wurden länderweise geplottet, so wurde das Überlappen durch die höhere Auflösung der Karte reduziert oder verhindert. Als zweite Lösung wurden die Stationen in einem Gitternetz dargestellt, wobei jeweils alle sich in einer Gitterzelle befindlichen Stationen gemittelt und zu einem Punkt zusammengefasst wurden.

5 Resultate

5.1 Vergleich zwischen Modellvorhersage und Messung

5.1.1 Analyse der Jahressummen

In der folgenden Tabelle 4 sind die Fehlermaße MAD, MD, PMD, RMSE und die Standardabweichung SD für die Vorhersagemodelle GFS, ICON, MF und NEMS und den Reanalysedatensatz ERA5 für Stationen mit einer Datenverfügbarkeit von jeweils mindestens 70 %, 50 % und 30 % der Messwerte dargestellt.

Tabelle 4: MAD, MD, PMD, RMSE und SD der Jahressummen für GFS, ICON, MF, NEMS und ERA5 für Stationen mit mindestens 70% verfügbaren Daten (oben), mindestens 50 % verfügbaren Daten (mittig) und mindestens 30% verfügbaren Daten (unten). In grün jeweils der beste Wert.

70 % Verfügbare Daten (8112 Stationen)					
	MAD [mm]	MD [mm]	PMD [%]	RMSE	SD
GFS	270	84	10.1	479	472
ICON	253	-51	-6.2	440	437
MF	281	131	15.8	469	450
NEMS	296	-144	-17.5	513	492
ERA5	234	-6	-0.8	516	516

50 % Verfügbare Daten (8480 Stationen)					
	MAD [mm]	MD [mm]	PMD [%]	RMSE	SD
GFS	271	89	10.9	479	470
ICON	251	-47	-5.8	435	433
MF	280	132	16.2	466	447
NEMS	294	-142	-17.5	510	490
ERA5	233	-4	-0.4	511	511

30 % Verfügbare Daten (8741 Stationen)					
	MAD [mm]	MD [mm]	PMD [%]	RMSE	SD
GFS	267	88	11.0	476	467
ICON	247	-47	-5.8	432	429
MF	276	129	16.2	462	444
NEMS	289	-140	-17.5	505	486
ERA5	230	-4	-0.5	505	505

Es ist zu erkennen, dass die MAD, MD und PMD der Jahressumme der Vorhersagen beim Reanalysedatensatz ERA5 deutlich kleiner sind als die Differenzen der Vorhersagemodelle. Generell ist die Spannweite zwischen den Vorhersagemodellen für die Jahresniederschlagssummen mit Werten zwischen 253 und 296 mm bei 70 % verfügbaren Daten gering.

Bei den Vorhersagemodellen schneidet das Modell des Deutschen Wetterdienstes ICON am besten ab. GFS und MF überschätzen die jährliche Niederschlagssumme tendenziell, wohingegen ICON, NEMS und ERA5 die jährliche Niederschlagssumme unterschätzen. Es fällt auf, dass sowohl die Standardabweichung als auch der RMSE trotz der relativ geringen mittleren Differenzen sehr hoch sind. Das liegt daran, dass bei diesen Fehlermaßen quadriert wird und somit Extremwerte stark ins Gewicht fallen. Stationen in Kolumbien, Venezuela und Indonesien haben absolute Differenzen von mehr als 10'000 mm bis zu einem maximalen Wert von 23'440 mm.

Anhand der Auswertung der Jahressummen wurden die Stationen ausgewählt, welche für die weitere Analyse verwendet werden. Die Fehler verändern sich durch die Änderung der Verfügbarkeit von 30 %, 50 % und 70 %

nur wenig. Die absolute Differenz wird höher, je mehr Messwerte vorhanden sind. Wie in Abb. 1 gut ersichtlich, ist der Unterschied der Anzahl Stationen, die hinzukommen, durch die allgemein hohe Verfügbarkeit sehr gering. Deshalb wurden für die weitere Datenanalyse nur noch Stationen mit mehr als 70% verfügbaren Daten verwendet.

In der folgenden Grafik ist die MAD für die vier Vorhersagemodelle GFS, ICON, MF und NEMS und den Reanalysedatensatz ERA5 dargestellt.

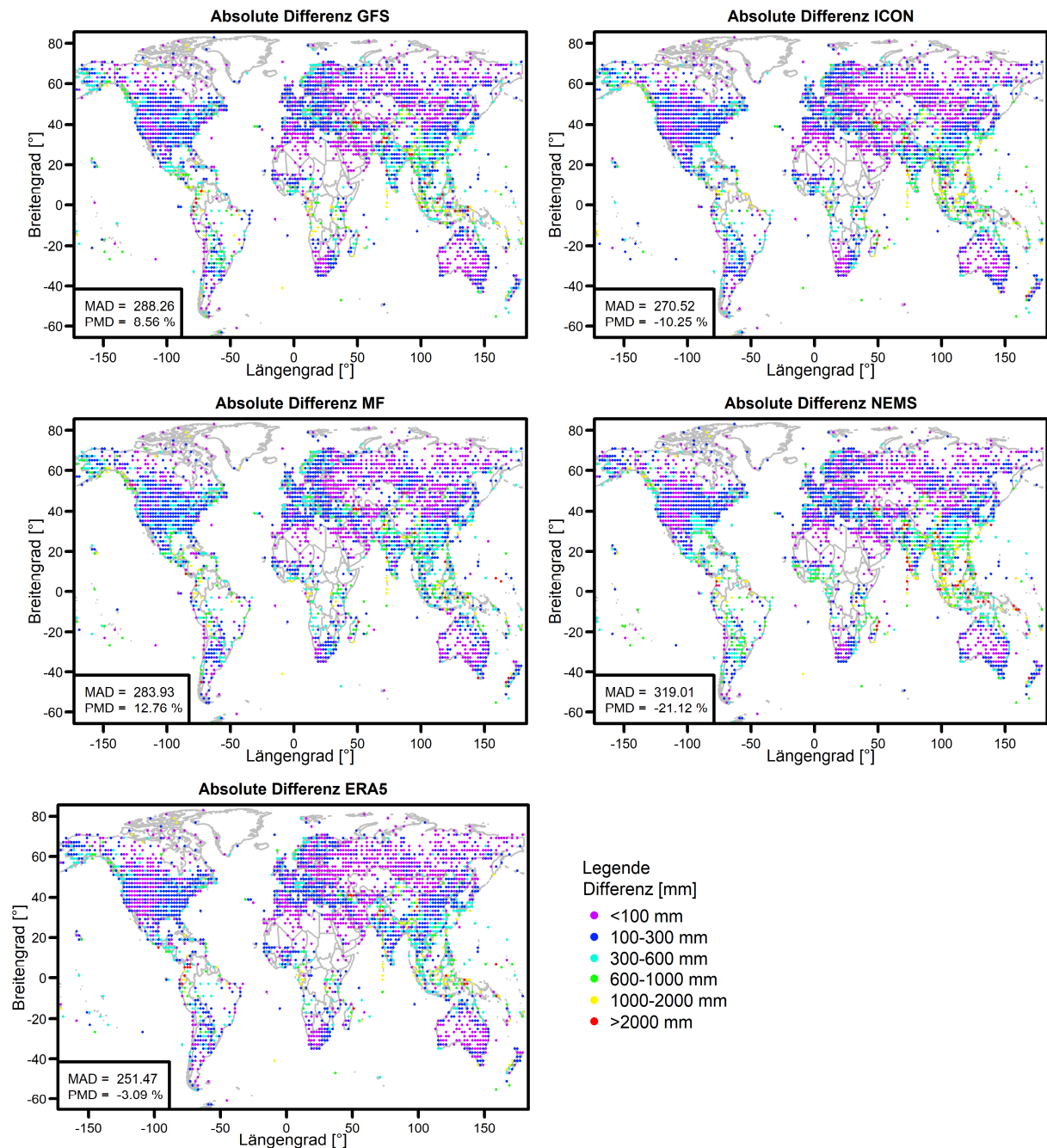


Abb. 4: MAD zwischen den Modellvorhersagen (GFS, ICON, MF, NEMS) sowie des Reanalysedatensatzes (ERA5) und den gemessenen Jahresniederschlagssummen.

Es ist zu erkennen, dass die Modelle im Süden Afrikas, in Nordafrika und in großen Teilen Russlands Vorhersagen mit weniger als 100 mm Abweichung zur gemessenen Jahressumme ergeben. In Europa und Nordamerika liegt der mittlere Fehler etwas höher. Hohe Abweichungen zeigen sich an der Ostküste Asiens und Australiens, mit Extremen in tropischen Regionen.

Bei der Betrachtung der MD in Abb. 5 ist wie schon in Tabelle 4 ersichtlich, dass GFS und MF die Niederschlagssummen eher überschätzen, während ICON, NEMS und ERA5 diese grundsätzlich eher unterschätzen. Bei den vier Vorhersagemodellen lassen sich die größten Differenzen in Südostasien feststellen. Während GFS und MF dort besonders überschätzen, unterschätzen ICON und NEMS in diesem Gebiet stark. Beim Reanalysedatensatz ERA5 wird in Südostasien sowohl über- als auch unterschätzt. Global betrachtet lassen sich, wie auch bei der MAD, in äquatornahen Gebieten größere Abweichungen erkennen. Die besten Vorhersagegenauigkeiten lassen sich von Europa bis Russland, im Westen der USA, in Südafrika und im Landesinneren Australiens feststellen. Bei allen Modellen wird eine Überschätzung der jährlichen Niederschlagssumme an der Westküste Kanadas und in Alaska festgestellt.

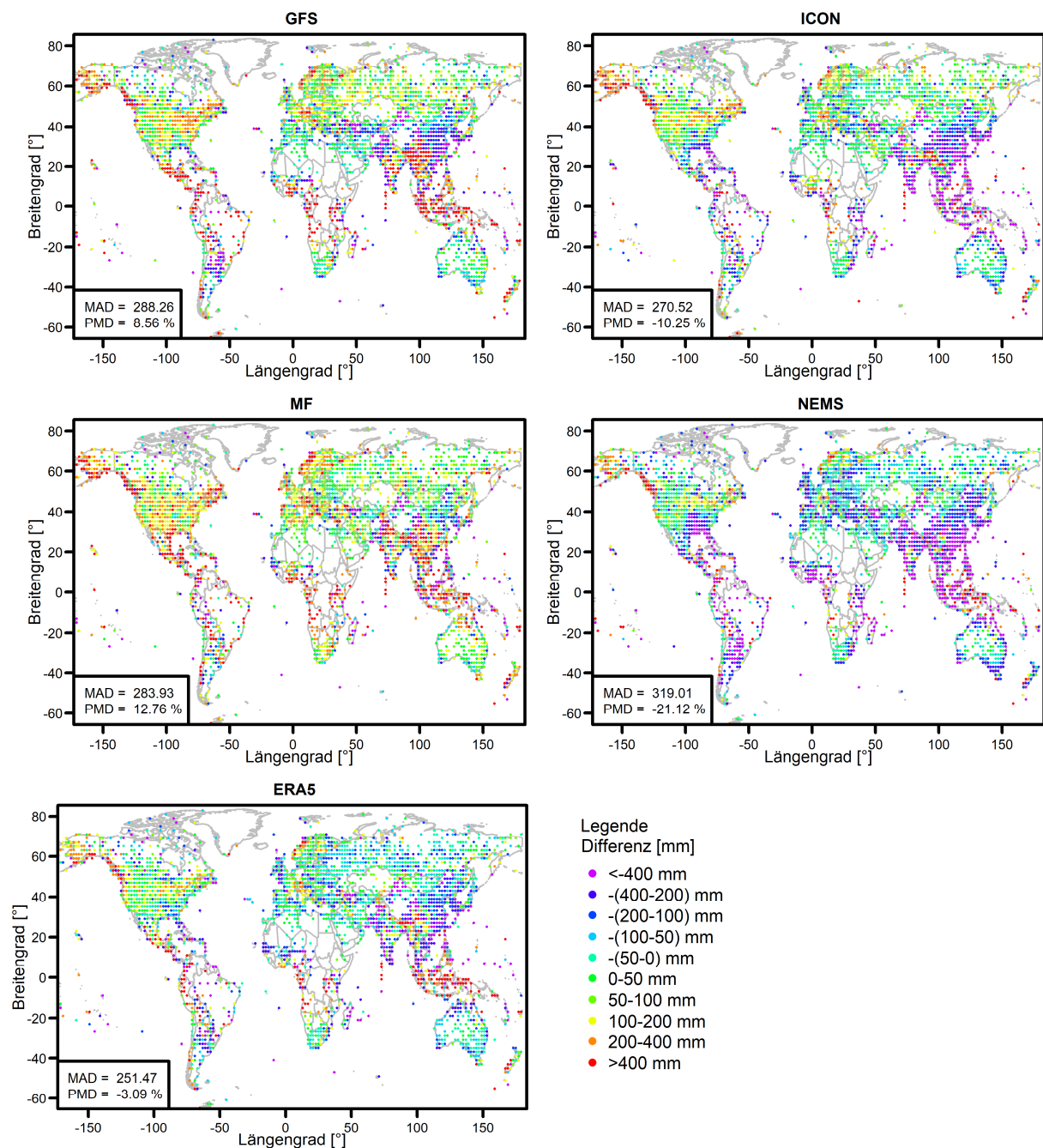


Abb. 5: MD zwischen den Modellvorhersagen (GFS, ICON, MF, NEMS) sowie des Reanalysedatensatzes (ERA5) und den gemessenen Jahresniederschlagssumme.

Mit dem Zusammenfassen mehrerer Stationen im Gitternetz, um das Überlappen der Stationen zu verhindern, geht eine Veränderung der Gewichtung einzelner Stationen einher. Die globale MAD verändert sich dadurch, da Regionen anhand ihrer Fläche und nicht nach der Anzahl Stationen gewichtet werden. In Tabelle 5 ist dies ersichtlich.

Tabelle 5: Unterschiede des MAD je nach Darstellungsform. Linke Spalte jede Station einzeln gewertet, mittlere Spalte im 2° Grid und rechte Spalte im 5° Grid.

	MAD Stationen [mm]	MAD 2° Grid [mm]	MAD 5° Grid [mm]
GFS	271	288	304
ICON	253	271	293
MF	281	284	304
NEMS	296	319	337
ERA5	234	251	273

Um die Modelle visuell ohne das Problem des Überlappens und ohne den Informationsverlust des Gitternetzes zu vergleichen, wurden die absoluten und relativen Differenzen für jedes Land einzeln als Karte dargestellt. In der folgenden Grafik sind die Karten von 5 Beispielländern für das Vorhersagemodell GFS dargestellt.

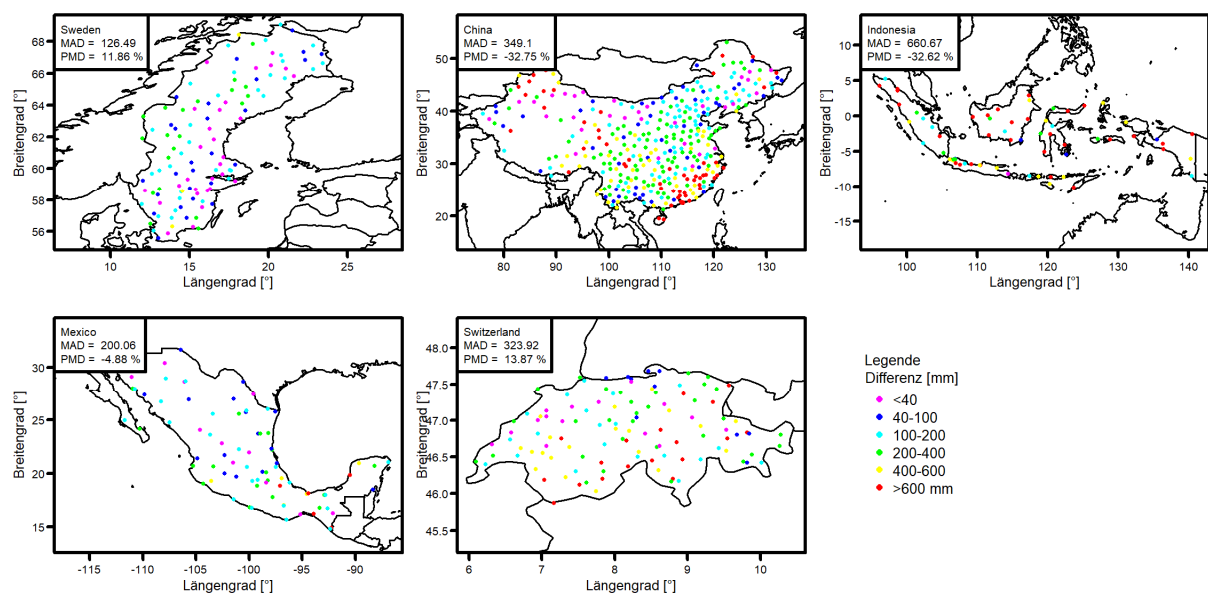


Abb. 6: Die MAD der Jahressumme des Modells GFS für die Beispielländer Schweden, China, Argentinien, Indonesien und Schweiz.

Die Karten in Abb. 6 zeigen, dass die absoluten Differenzen zwischen Vorhersage und Messung an der Ostküste Schwedens mit Niederschlagswerten unter 40 mm tendenziell kleiner sind als an der Westgrenze mit Werten im Bereich von 100 bis 400 mm. In China zeigt sich, dass die Differenzen in Küstennähe sehr viel größer sind als jene im Landesinneren. Indonesien zeigt mit einem gemittelten MAD von 661 mm wesentlich größere Werte als die anderen Länder. In der Schweiz lassen sich höhere Differenzen in den Gebirgsregionen erkennen. So sind die Werte an den Stationen in den Alpen und im Jura höher als jene in flachen Gebieten.

Die folgenden Karten zeigen die absoluten Differenzen der Modelle am Beispiel von Indien.

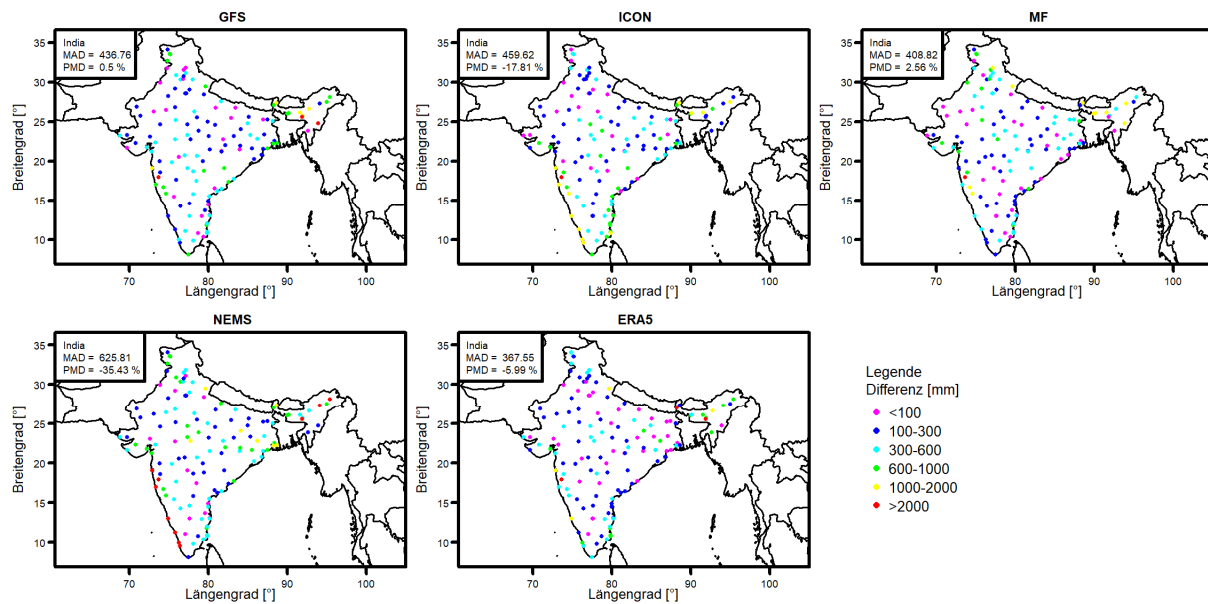


Abb. 7: MAD der Jahressumme der vier Vorhersagemodelle (GFS, ICON, MF, NEMS) und dem Reanalysedatensatz (ERA5) am Beispiel von Indien.

Es zeigt sich, dass beispielsweise GFS und NEMS im Landesinnern Indiens mit Werten zwischen 300 und 600 mm größere Differenzen aufweisen als die anderen Vorhersagemodelle. An der Westküste haben hingegen ICON und NEMS größere Differenzen. Den tiefsten gemittelten MAD der Vorhersagemodelle erreicht MF mit 409 mm, welcher jedoch vom Reanalysedatensatz ERA5 mit 378 mm unterboten wird.

An den Karten der MD und den eingedruckten Mittelwerten erkennt man in Abb. 8 die Unterschiede zwischen den Modellen am Beispiel von Südafrika.

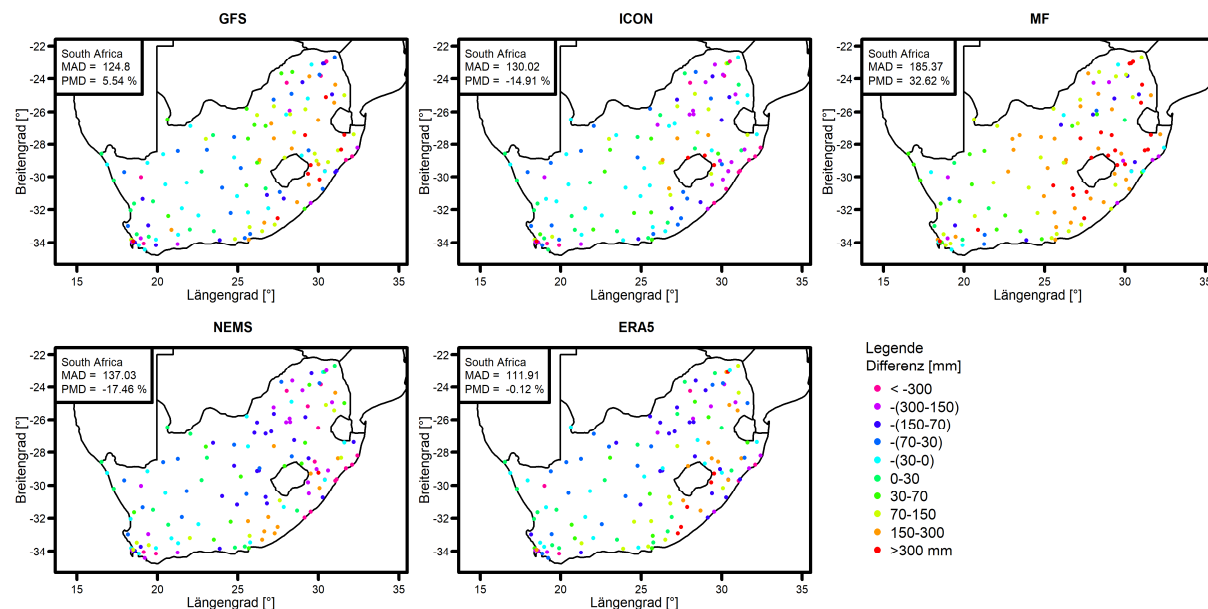


Abb. 8: MD der Jahressummen für die vier Vorhersagemodelle (GFS, ICON, MF, NEMS) und den Reanalysedatensatz (ERA5) in Südafrika.

Es zeigt sich, dass die Modelle die Niederschlagssummen im Osten tendenziell eher überschätzen als im Westen des Landes. Sehr deutlich fällt dies bei MF auf, welches die Niederschlagssummen auch im restlichen Teil des Landes überschätzt.

Anhand der MAD der Jahressummenvorhersage wurde in Abb. 9 eine Karte erstellt, welche jeweils das beste Niederschlagsvorhersagemodell für ein Land anzeigt. Die grau markierten Länder besitzen keine WMO-Messstationen.

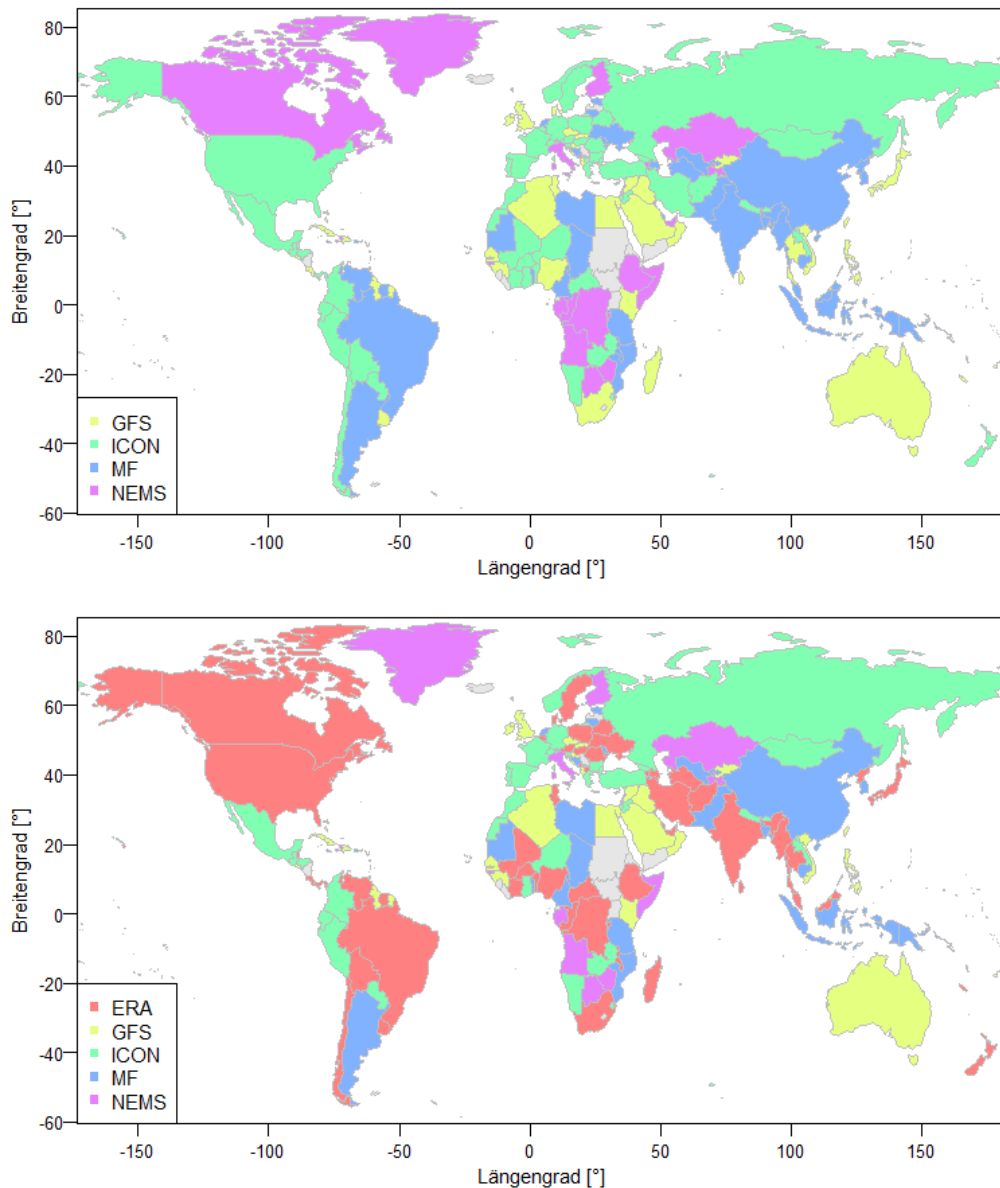


Abb. 9: Darstellung des Modells mit dem tiefsten MAD für jedes Land unter Berücksichtigung der vier Vorhersagemodelle (GFS, ICON, MF, NEMS) (oben) und dem zusätzlichen Hinzuziehen des Reanalysedatensatzes ERA5 (unten).

5.1.2 Analyse der Monatssummen

Bei der Analyse der Monatssummen des Niederschlags zeigte sich, dass die global gemittelte MAD in den Monaten Juli und August am größten ausfällt. Bei den Vorhersagemodellen liegt der Schnitt bei 40,6, bzw. 40,5 mm. Die kleinste gemittelte MAD wurde im Februar mit 27,7 mm erreicht. Die größte Abweichung zwischen den Vorhersagemodellen wird im April erreicht, hier beträgt die MAD von ICON 18 und die MAD von GFS 26,5.

Wie in Tabelle 6 ersichtlich, zeigt die PMD der vier Vorhersagemodelle (GFS, ICON, MF, NEMS) im Juli die größte Abweichung von 0. Im Januar liegt die PMD am nächsten bei 0 und ist somit am genauesten. Im Januar und von Juli bis Oktober wurde die über alle Modelle gemittelte monatliche Niederschlagssumme unterschätzt, in den restlichen Monaten hingegen überschätzt. Die geringste Abweichung erreichte ICON im März mit 0,3 %

Überschätzung der monatlichen Niederschlagssumme und einer über das Jahr gemittelten Unterschätzung von 5.9 %.

Beim Hinzuziehen des Reanalysedatensatzes ERA5 sind kleinere Unter- beziehungsweise Überschätzungen im Vergleich zu den Vorhersagemodellen zu erkennen. Nur im Januar (GFS, ICON) und März und April (ICON) sind die Vorhersagemodelle um bis zu 2.3 % besser als der Reanalysedatensatz.

Tabelle 6: PMD der vorhergesagten Monatssummen zu den gemessenen Monatssummen in Prozent.

	Jan	Feb	Mär	Apr	Mai	Jun	Jul	Aug	Sep	Okt	Nov	Dez
GFS	5.0	6.0	10.1	12.5	15.0	13.6	13.3	8.7	9.1	8.6	10.0	7.3
ICON	-9.0	-5.8	-4.4	0.3	1.0	-6.7	-12.2	-13.1	-8.5	-5.5	-2.3	-4.6
MF	5.7	12.8	18.6	23.1	21.8	16.3	18.6	11.7	13.6	12.9	18.7	16.4
NEMS	-11.6	-12.6	-12.4	-11.9	-13.9	-23.0	-25.3	-28.6	-20.4	-17.0	-13.6	-11.2
ERA5	-6.5	-2.5	0.2	2.6	2.5	-0.4	-2.5	-3.7	-0.6	3.1	-0.7	-0.9

Der MAE, der MBE und der MPE wurden für das Jahr 2019 anhand der Monatssummen berechnet. Sie sind in der Tabelle 7 dargestellt. Der Reanalysedatensatz ERA5 erzielt bei allen drei Fehlermaßen die niedrigsten Fehler. Von den Vorhersagemodellen erreicht ICON die besten Werte.

Tabelle 7: MAE, MBE und MPE für die Vorhersagemodelle GFS, ICON, MF, NEMS und den Reanalysedatensatz ERA5.

	MAE [mm]	MBE [mm]	MPE [%]
GFS	35.04	6.97	10.13
ICON	31.09	-4.26	-6.20
MF	36.48	10.90	15.85
NEMS	34.90	-12.04	-17.49
ERA5	30.30	-0.53	-0.78

5.1.3 Analyse der Tagessummen

Die mittlere Anzahl der Regentage (RT) für das Jahr 2019, definiert als Tage mit einer Niederschlagssumme grösser als 1 mm, sind mit der jeweiligen Standardabweichung in Tabelle 8 angegeben. Die meisten Regentage wurden in Europa gemessen. Die Vorhersagen der Modelle fielen jedoch sehr unterschiedlich aus. So sagten GFS, MF und ERA5 für Südamerika am meisten Regentage vorher. NEMS sagte die Anzahl Regentage auf allen Kontinenten bis auf Australien am besten vorher.

Tabelle 8: Die mittlere Anzahl RT für Kontinente mit der jeweiligen Standardabweichung der gemessenen Werte (WMO), der vier Modellvorhersagen (GFS, ICON, MF NEMS) und des Reanalysedatensatzs (ERA5).

	WMO		GFS		ICON		MF		NEMS		ERA5	
	RT	SD	RT	SD	RT	SD	RT	SD	RT	SD	RT	SD
Afrika	48	33	75	58	66	51	102	72	61	44	70	56
Asien	78	43	105	66	87	49	126	70	81	49	100	63
Australien	63	40	76	53	62	47	82	49	52	40	66	48
Europa	101	37	124	40	109	35	125	36	101	37	109	38
Nordamerika	85	34	116	37	100	32	124	36	92	32	101	33
Südamerika	74	46	129	75	97	62	139	72	88	59	112	67

Wie Tabelle 9 entnommen werden kann, ist die Vorhersage für kleine Niederschlagssummen viel genauer als für größere Summen. So werden tägliche Niederschlagssummen mit mehr als 1 mm von den Modellen wesentlich besser vorhergesagt als Niederschlagssummen mit mehr als 10 oder 50 mm. Bei kleinen täglichen Niederschlagssummen von mehr als 1 mm schneidet ICON mit einem HSS von 0.46 und der POD von 0.57 am besten ab.

Bei Niederschlagsereignissen von mehr als 50 mm ist der HSS mit Werten von weniger als 0.1 sehr tief und nahe bei «no skill». Je höher die Niederschlagssumme, desto besser wird der Reanalysedatensatz im Vergleich zu den Modellen.

Tabelle 9: HSS, FAR und POD für die Klassengrenzen 1 mm, 10 mm und 50 mm für alle Modelle (GFS, ICON, MF, NEMS) und den Reanalysedatensatz ERA5. In grün jeweils der beste Wert.

	täglicher Niederschlag (Stationen gemittelt)								
	> 1 mm			> 10 mm			>50 mm		
	HSS	FAR	POD	HSS	FAR	POD	HSS	FAR	POD
GFS	0.43	0.30	0.52	0.30	0.61	0.35	0.08	0.91	0.20
ICON	0.46	0.34	0.57	0.31	0.62	0.38	0.09	0.90	0.24
MF	0.42	0.27	0.50	0.30	0.60	0.38	0.08	0.91	0.20
NEMS	0.41	0.41	0.56	0.27	0.69	0.37	0.07	0.93	0.19
ERA5	0.45	0.33	0.56	0.32	0.61	0.39	0.10	0.90	0.25

Bei der globalen Betrachtung des HSS (Abb. 10) ist zu erkennen, dass tiefe Werte vor allem in Äquatornähe vorkommen. Weiterhin lassen sich tiefe Werte auch im Himalaya, in Alaska und in Neuseeland feststellen.

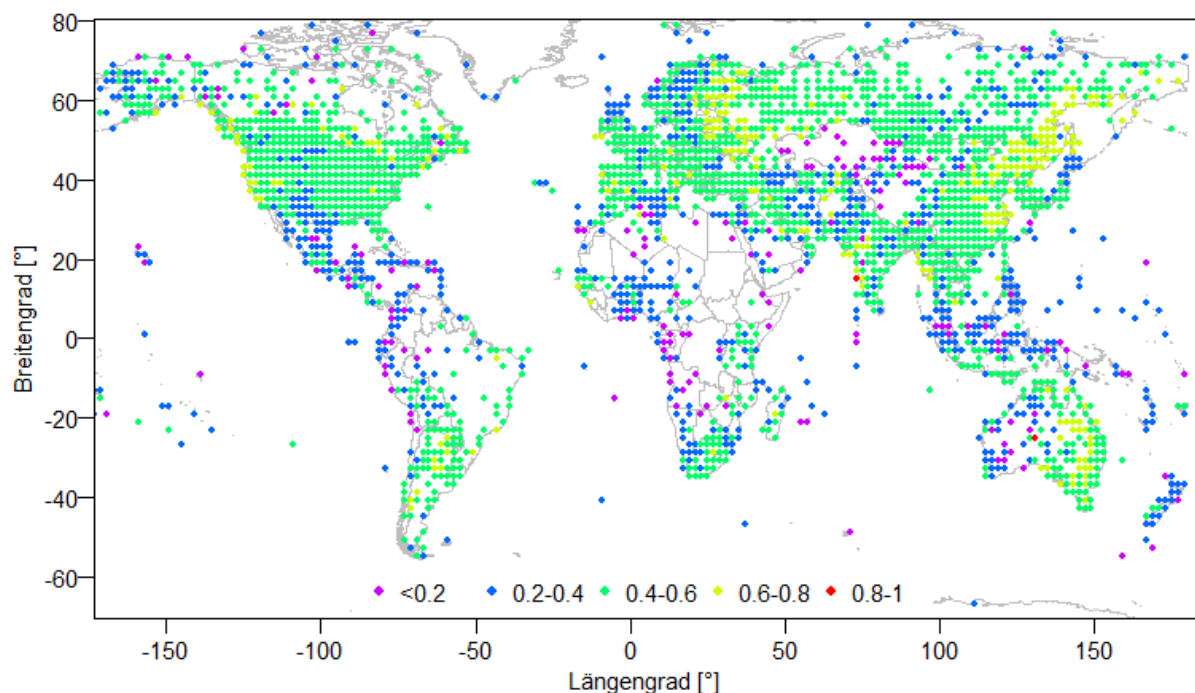


Abb. 10: Der gemittelte HSS für globale Niederschlagssummen von mehr als 1 mm für ICON im 2° Gitternetz.

In den folgenden Karten in Abb. 11 wird für jedes Land dasjenige Modell dargestellt, welches den höchsten HSS erreicht. Es lässt sich deutlich erkennen, dass ICON in weiten Teilen am besten abschneidet. In den Ländern mit generell tiefem HSS in Mittel und Nordafrika schneiden auch NEMS und GFS als beste Modelle ab. In einzelnen Ländern Europas bietet MF die beste Eventvorhersage.

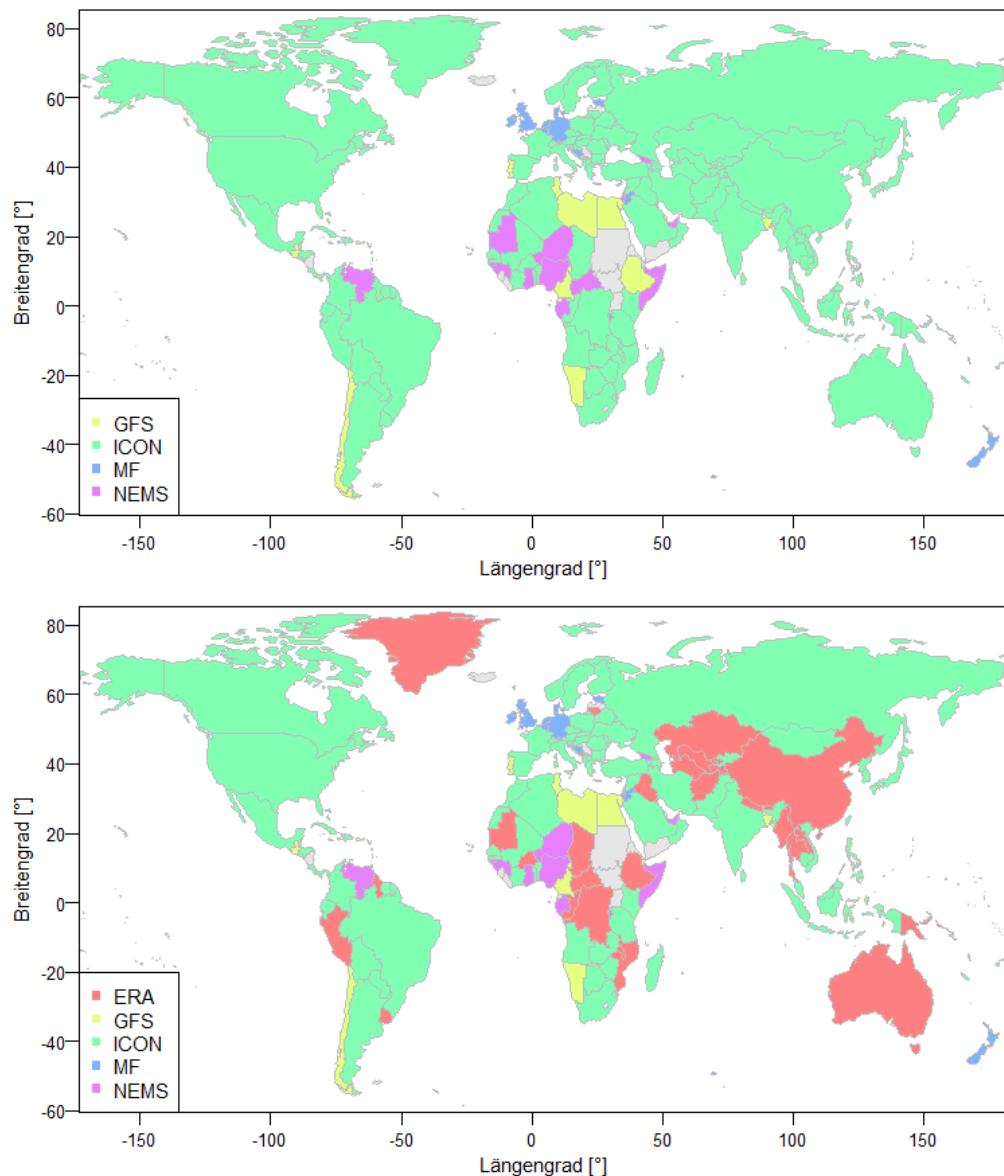


Abb. 11: Der größte HSS für jedes Land, gemittelt unter Berücksichtigung der vier Vorhersagemodelle (GFS, ICON, MF, NEMS) (oben) und dem zusätzlichen Hinzuziehen vom Reanalysedatensatz ERA5 (unten). Die grau markierten Länder besitzen keine Messstation.

Die folgende Tabelle 10 zeigt den mittleren HSS für Regentage mit mehr als 1 mm Niederschlag. Es zeigt sich, dass ICON für alle Monate, abgesehen von November, den höchsten HSS der Vorhersagemodelle aufweist und auch in jedem Monat höher liegt als der Reanalysedatensatz ERA5. Der größte HSS wird bei allen Modellen im Dezember erreicht. Dort liegen alle Modelle außer NEMS über 0.4. Mit 0.433 erreicht ICON den höchsten Wert aller Modelle. Die tiefsten Werte aller Modelle werden in den Monaten August und Juli erreicht. Dabei erreicht NEMS mit 0.297 den tiefsten Wert.

Tabelle 10: HSS für Tage mit > 1 mm Niederschlag für die Monate Januar bis Dezember. In grün jeweils der beste Wert.

	Jan	Feb	Mär	Apr	Mai	Jun	Jul	Aug	Sep	Okt	Nov	Dez
GFS	0.352	0.380	0.368	0.360	0.333	0.316	0.305	0.306	0.368	0.394	0.407	0.425
ICON	0.381	0.413	0.406	0.403	0.385	0.360	0.354	0.345	0.393	0.404	0.404	0.433
MF	0.360	0.386	0.372	0.368	0.339	0.304	0.279	0.289	0.350	0.381	0.384	0.406
NEMS	0.332	0.360	0.353	0.344	0.332	0.308	0.302	0.297	0.352	0.375	0.363	0.383
ERA5	0.364	0.391	0.389	0.395	0.374	0.345	0.335	0.328	0.383	0.391	0.385	0.414

Um die räumlichen Muster aufzuzeigen, wurden die in der vorhergehenden Tabelle 10 gemittelten HSS-Werte des Modells GFS als Karte (Abb. 12) für jeden Monat dargestellt.

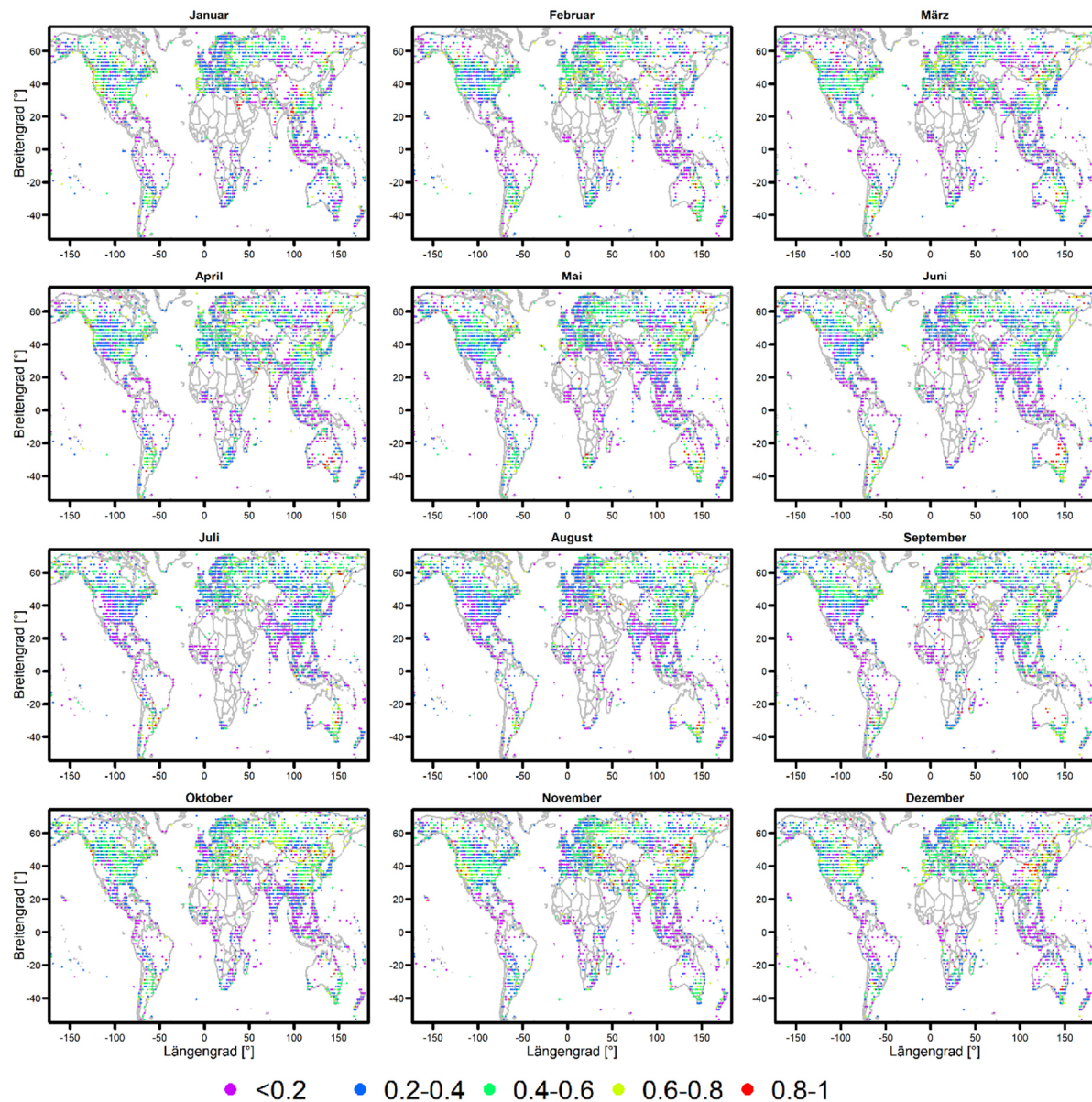


Abb. 12: Darstellung des HSS der monatlichen Niederschlagsereignisse am Beispiel von GFS.

In den Monaten Juli und August ist in Indien und im Golf von Thailand während des indischen Sommermonsuns ein tiefer HSS erkennbar. In Nordamerika ist der HSS ähnlich wie in Europa während der Wintermonate deutlich höher und erreicht im Mai seinen Tiefpunkt. Global gesehen ergeben sich die größten Werte im Dezember, wie in den Gebieten mit hoher Stationsdichte in Europa und Nordamerika.

5.2 Variabilität zwischen den Vorhersagemodellen

Die Unterschiede zwischen den Vorhersagemodellen werden durch die Spannweite der Vorhersagen beziffert. So wird die Differenz zwischen dem maximalen Vorhersagewert und dem minimalen Vorhersagewert berechnet. In Abb. 13 lässt sich erkennen, dass die Variabilität in den Tropen generell höher ist. Besonders hohe Abweichungen ergeben sich an der Ostküste Indiens in Indonesien, der nördlichen Ostküste Südamerikas und der Region um Nepal, Bhutan und Bangladesch.

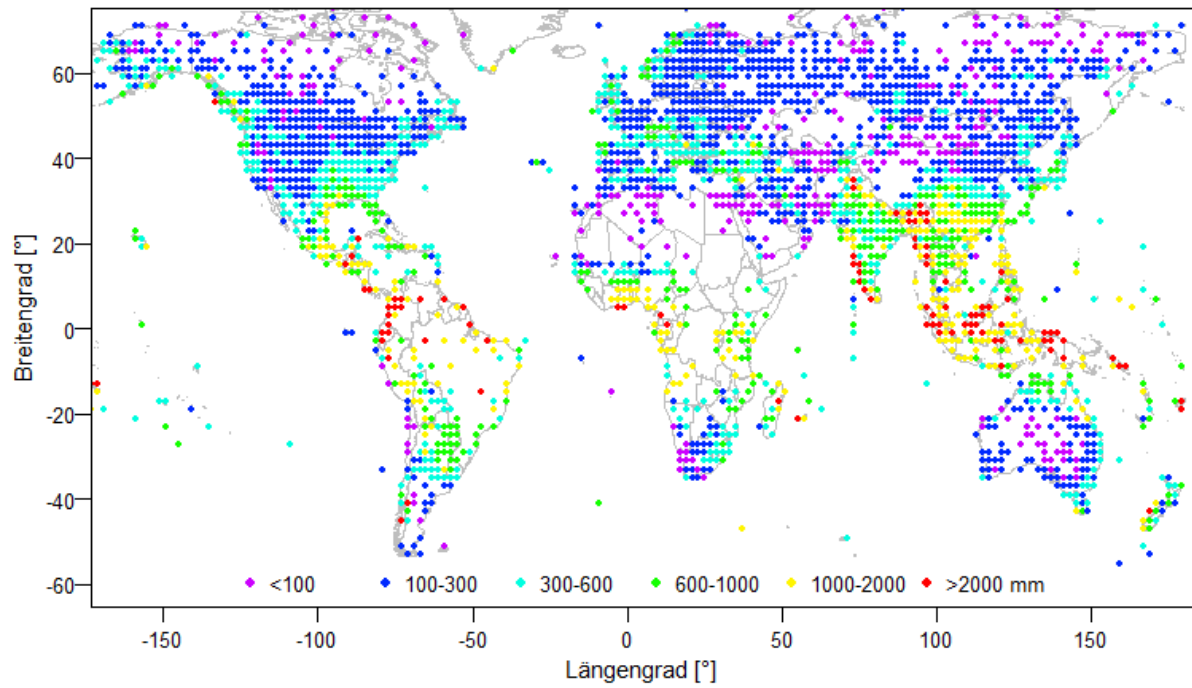


Abb. 13: Die maximale Differenz der Jahressummen der Vorhersagemodelle und ERA5. Die Differenzen sind in mm angegeben.

Die Modelldifferenz weist eine sehr ähnliche globale Verteilung wie die Differenzen zwischen den Modellen und den Messungen auf. So treten in den gleichen Regionen in Äquatornähe die größten Ungenauigkeiten auf.

Beim Vergleichen der Monate lässt sich in Abb. 14 erkennen, dass sich die Modelle zu verschiedenen Jahreszeiten unterschiedlich stark unterscheiden. Wie hier am Beispiel von Februar zu August zeigt sich, dass die höchsten Modelldifferenzen im Nordsommer tendenziell weiter nördlich liegen als im Nordwinter. Im August fallen auch die hohen Abweichungen im Bereich Indiens und Südasiens auf.

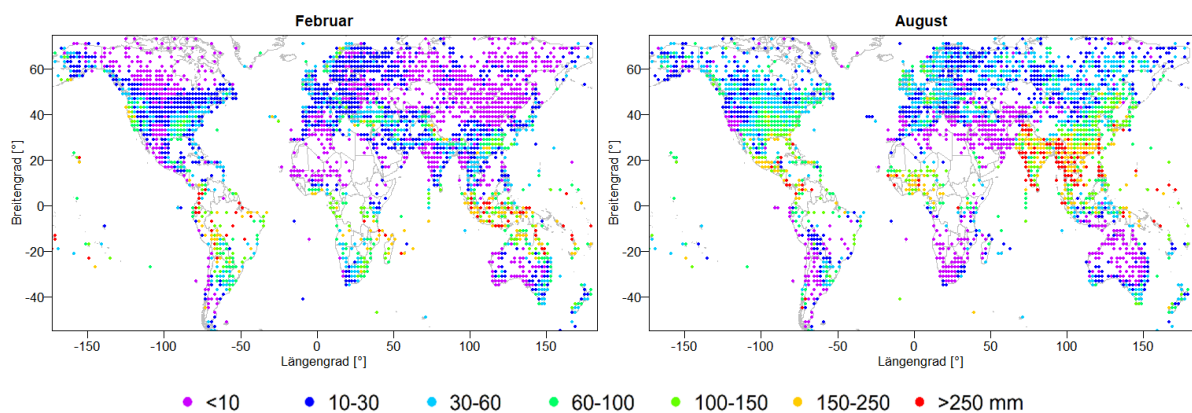


Abb. 14: Maximale Differenz der Monatssummen der vier Vorhersagemodelle. Links der Monat Februar, rechts August.

Die Unterschiede zwischen den Modellen wurden in Abb. 15 auch anhand der Tagessummen als kategorische Variablen herausgearbeitet. Es wird die Spannweite zwischen der aufsummierten Anzahl Regentage der vier Vorhersagemodelle dargestellt.

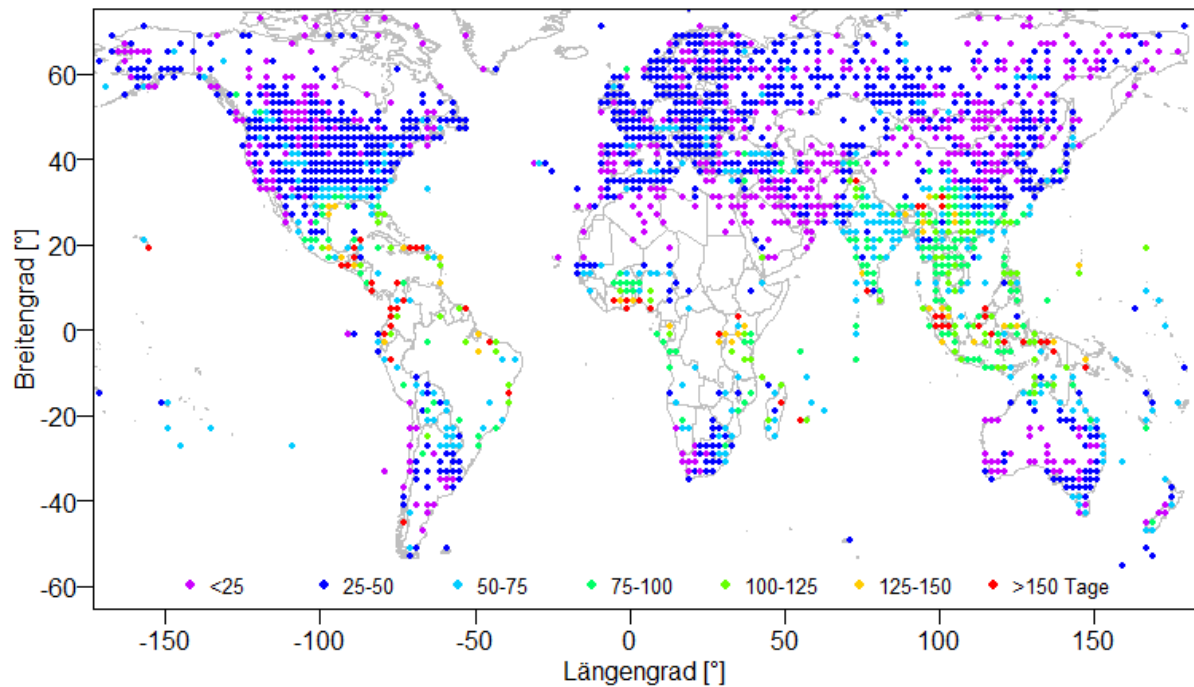


Abb. 15: Mittlere Vorhersagemodelldifferenzen der Anzahl Regentage der Vorhersagemodelle (GFS, ICON, MF und NEMS).

Die größten Unterschiede zwischen den vorhergesagten Regentagen finden sich in Südostasien, Mittelamerika und dem Golf von Guinea und somit erneut in Äquatornähe. Bei 154 Stationen wurden Abweichungen von über 150 Tagen gemessen, wobei die mittlere absolute geografische Breite dieser Stationen 11.0 ° beträgt. Beim Miteinbeziehen des Reanalysedatensatzes (ERA5) wird die Spannweite größer. So haben 169 Stationen Abweichungen von über 150 Tagen.

5.3 Statistische Analyse geografischer und zeitlicher Abhängigkeiten

5.3.1 Vorhersagegenauigkeit in Abhängigkeit von Geländehöhe und geografischer Breite

Die Analyse der MAD in Abhängigkeit der Geländehöhe in Abb. 16 lässt erkennen, dass die MAD mit der Geländehöhe zunimmt. Dieser Trend zeigt sich für alle 4 Modelle und den Reanalysedatensatz.

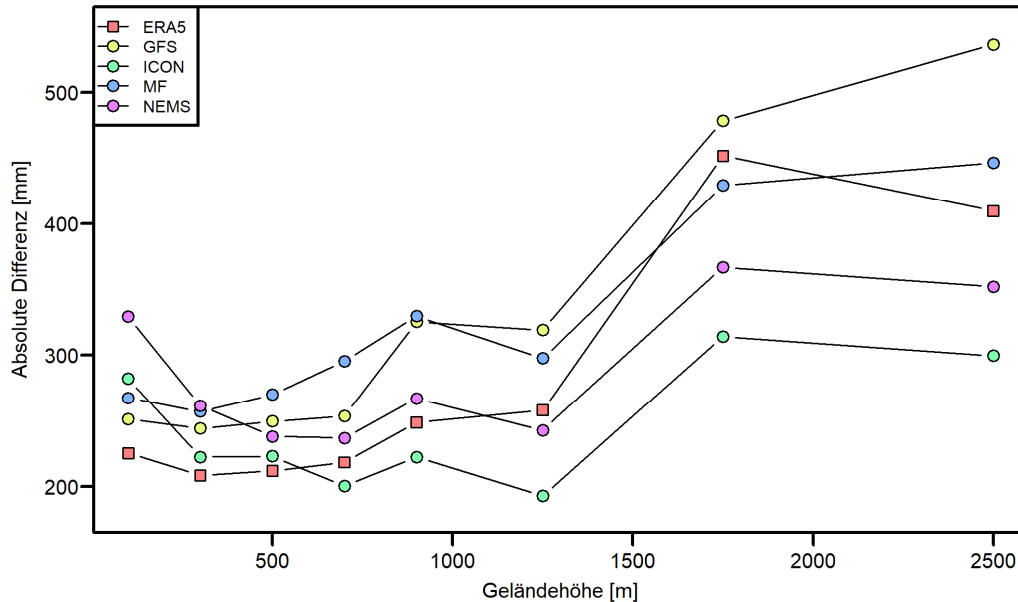
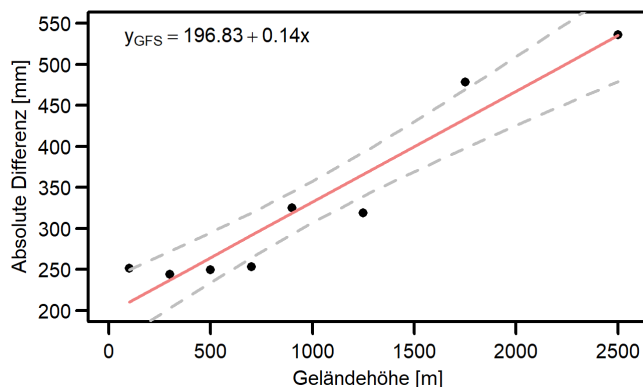


Abb. 16: MAD in Abhängigkeit von der Geländehöhe für alle Vorhersagemodelle (GFS, ICON, MF, NEMS; als Kreise) und den Reanalysedatensatz (ERA5; als Quadrat).

Die Abhängigkeit der MAD zur Geländehöhe wird in Abb. 17 am Beispiel des Modells GFS gezeigt. Es lässt sich eine Zunahme von 14 mm Differenz pro 100 Meter Geländehöhe feststellen. Dieser steigende Trend ist mit einem p-Wert von $1.95 \cdot 10^{-4}$ statistisch signifikant.



$$MAD_{Model} = \dots [mm] + \dots \left[\frac{mm}{m} \right] [m]$$

$$y_{GFS} = 196.83 + 0.14 \cdot x$$

$$y_{ICON} = 215.97 + 0.03 \cdot x$$

$$y_{MF} = 238.11 + 0.09 \cdot x$$

$$y_{NEMS} = 249.77 + 0.04 \cdot x$$

$$y_{ERA5} = 174.70 + 0.10 \cdot x$$

Abb. 17: MAD in Abhängigkeit von der Geländehöhe, hier am Beispiel von GFS mit Trendgerade (rot) und dem 95 % Konfidenzintervall (grau). Rechts die Formeln der Trendgeraden aller Modelle.

Anhand der Trendgeraden der Modelle erweisen sich ICON und NEMS mit Zunahmen der MAD von nur 3 mm, respektive 4 mm pro 100 Meter Geländehöhe im Vergleich mit den anderen Modellen als robuster gegenüber Höhenänderungen.

Bei der Analyse der MD fällt auf, dass die Überschätzung der Niederschlagssummen bei Geländehöhen in der Stufe zwischen 2000 und 3000 m wesentlich größer ausfällt als in den anderen Stufen. So liegen einzig in dieser Stufe alle Modellvorhersagen über den gemessenen Werten, auch jene der sonst unterschätzenden Modelle ICON und NEMS.

Bei der Analyse der MAD in Bezug auf die geografische Breite zeigte sich bei größerer Entfernung zum Äquator eine genauere Vorhersage der Jahressummen (Abb. 18). Die größten Differenzen wurden zwischen -10 ° und 20 ° geografischer Breite festgestellt. Es zeigt sich, dass die Genauigkeit der Vorhersagen auf der Nordhemisphäre bei allen Modellen gleichmäßig zunimmt. Auf der Südhalbkugel nimmt die Genauigkeit bei Breiten südlicher als -30 ° nochmals ab. Die Modelle MF und GFS haben in einer Kategorie von -55 bis -45 ° beziehungsweise von -75 bis -65 ° wesentlich größere MAD Werte als die anderen Modelle. Bei MF kommen diese hohen MAD in Neuseeland, Tasmanien und Chile zustande. Dabei unter- sowie überschätzte MF in Neuseeland und Tasmanien die Niederschlagssummen um jeweils mehr als 800 mm. Bei GFS entstehen die großen Differenzen an Stationen auf der Antarktis südlich Südamerikas.

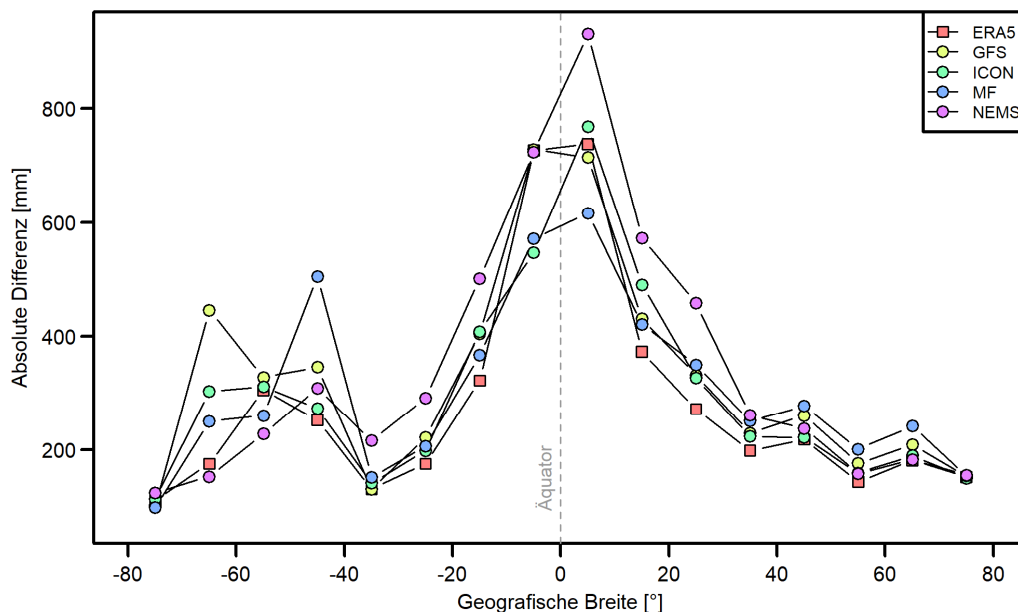
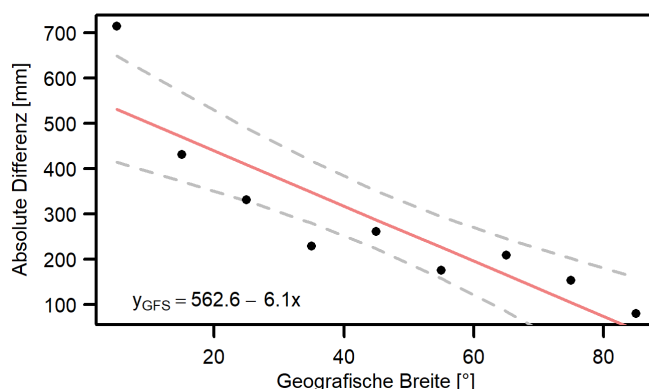


Abb. 18: MAD in Abhängigkeit von der geografischen Breite für alle Vorhersagemodelle (GFS, ICON, MF, NEMS; als Kreise) und den Reanalysedatensatz (ERA5; als Quadrat).

Es wurde eine lineare Regression für die absolute geografische Breite, die Nord- und die Südhemisphäre gerechnet. Die absoluten Differenzen sind in Äquatornähe wesentlich höher als in äquatorfernen Gebieten. Auf der Nordhalbkugel (Abb. 19) zeigt sich der abnehmende Trend mit der geografischen Breite besonders deutlich. So nimmt bei GFS die MAD mit 6.1 mm pro Breitengrad ab. Der abnehmende Trend ist mit einem p-Wert von 0.00166 statistisch signifikant.



$$MAD_{Model} = ...[mm] - ... \left[\frac{mm}{^\circ} \right] [^\circ]$$

$$y_{GFS} = 562.6 - 6.1 \cdot x$$

$$y_{ICON} = 601.9 - 7.0 \cdot x$$

$$y_{MF} = 533.5 - 5.5 \cdot x$$

$$y_{NEMS} = 733.1 - 8.8 \cdot x$$

$$y_{ERA5} = 529.9 - 6.0 \cdot x$$

Abb. 19: MAD in Abhängigkeit von der geografischen Breite auf der Nordhalbkugel, hier am Beispiel von GFS mit Trendgerade (rot) und dem 95 %-Konfidenzintervall (grau). Rechts die Formeln der Trendgeraden aller Modelle.

Anhand der Trendgeraden kann erkannt werden, dass bei allen Modellen ein stark negativer Trend vorliegt. MF wird mit einer Abnahme von 5.5 mm pro Breitengrad am robustesten gegenüber der geografischen Breite, wohingegen NEMS sich mit einer Abnahme von 8.8 mm am stärksten verändert.

5.3.2 Vorhersagegenauigkeit in Abhängigkeit von der Distanz zum offenen Gewässer

Die mittlere absolute Differenz nimmt mit der Distanz zur Küste ab (Abb. 20). Weiter ist ersichtlich, dass die Abweichungen zwischen den Modellen kleiner werden, wenn die Distanz zum Meer größer wird.

Die Variabilität zwischen den Modellvorhersagen ist in Küstennähe größer als im Landesinneren. Es lässt sich auch erkennen, dass die Niederschlagssumme in Küstennähe geringfügig tiefer eingeschätzt wird, was somit bei NEMS, ICON und ERA5 zu kleineren Abweichungen weiter im Landesinneren führt, bei GFS und MF hingegen zu einer Überschätzung des Niederschlags führt.

Bei der Betrachtung der MD zeigte sich ein leicht steigender Trend bis zu 500 km Distanz zum Meer. So sind die Vorhersagen der global meist unterschätzenden Modelle ICON und NEMS besser. Die Vorhersagen der global meist überschätzenden Modelle, hier vor allem GFS, werden durch die steigende Überschätzung schlechter und erreichen bei der Klasse 200-500 km Distanz zum Meer die am weitesten von 0 entfernten MD.

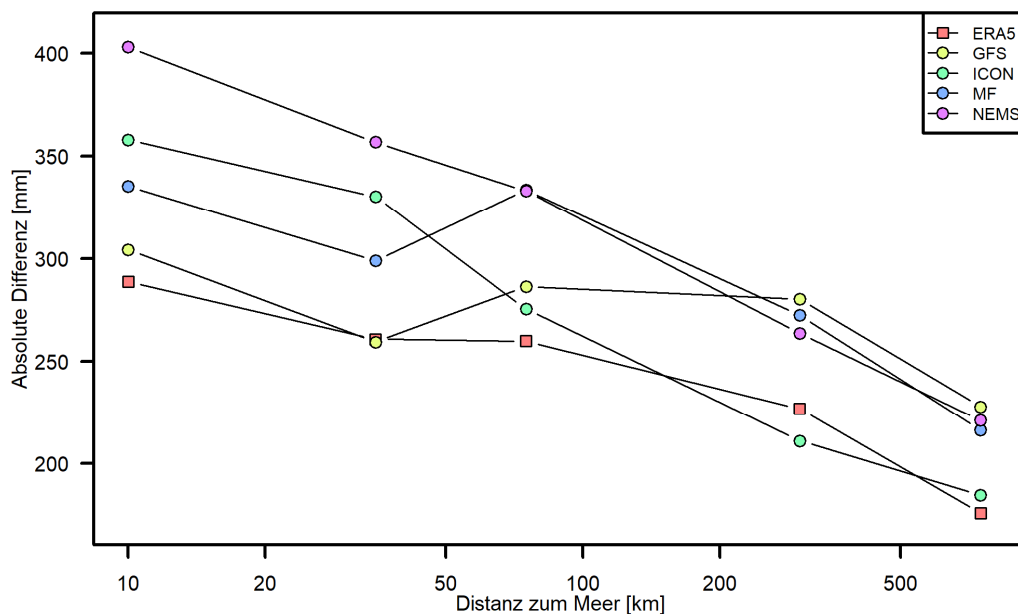


Abb. 20: MAD in Abhängigkeit von der Distanz zum Meer für alle Vorhersagemodelle (GFS, ICON, MF, NEMS; als Kreise) und den Reanalysedatensatz (ERA5; als Quadrat) mit logarithmierter x-Achse.

5.3.4 Temperatur

Die Temperaturabhängigkeit der Modellperformance wurde anhand des HSS für die einzelnen Monate betrachtet. Die folgende Abb. 21 zeigt die Modellperformance der vier Modelle und des Reanalysedatensatzes in zehn Temperaturklassen von -20 °C bis 30 °C in Schritten von 5 °C.

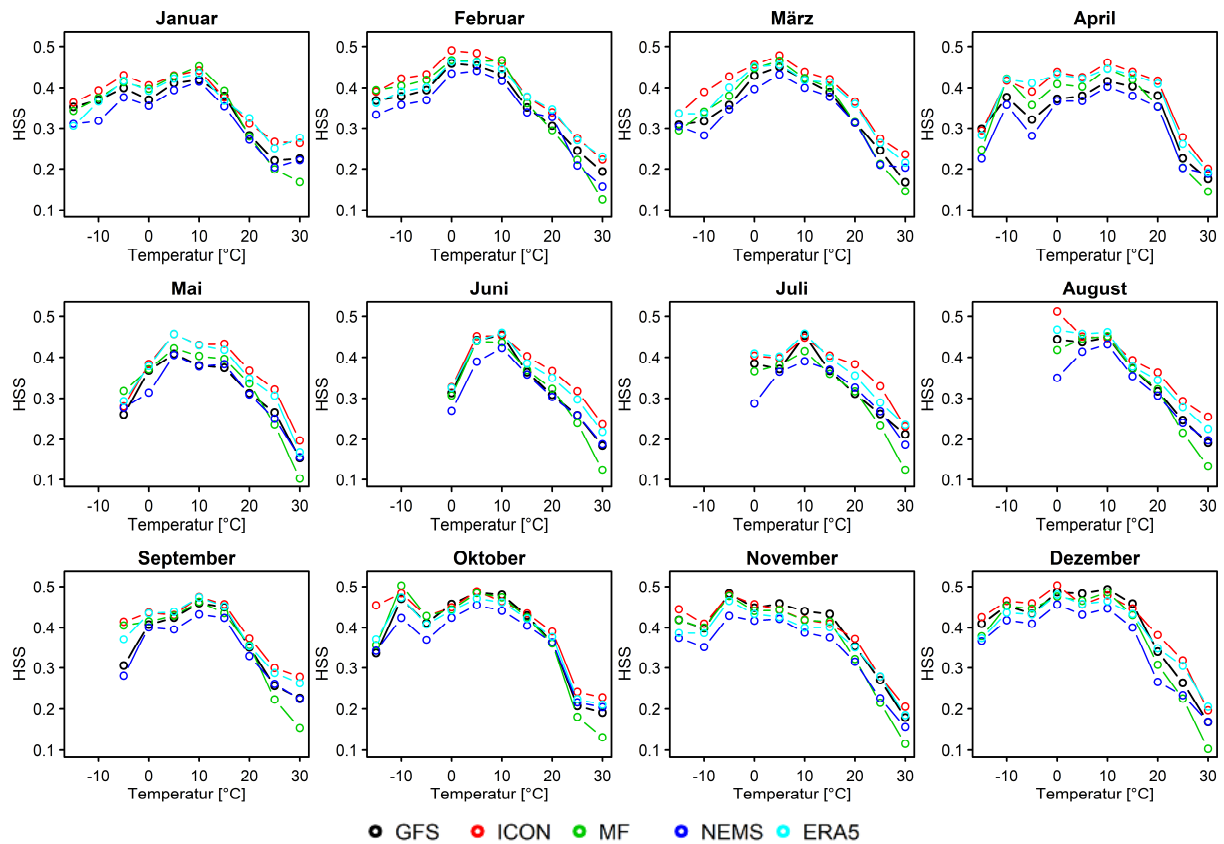


Abb. 21: Der HSS für die Vorhersagemodelle (GFS, ICON, MF, NEMS) und den Reanalysedatensatz (ERA5) für 5 °C – Temperaturklassen und jeden Monat.

Der HSS erreicht meist bei Temperaturen zwischen 0 und 10 °C den Maximalwert. Es lässt sich erkennen, dass der HSS bei hohen und tiefen Temperaturen deutlich tiefer liegt. Eine Klasse wurde nur dargestellt, wenn mindestens 5 Stationen darin vorkommen. Da in den Nordsummermonaten nur wenige Stationen unter 0°C aufweisen, wurden diese Temperaturklassen nicht abgebildet.

In einem weiteren Schritt wurden die Stationen getrennt nach Hemisphäre betrachtet, um ein Aufheben der Effekte der Jahreszeiten zu vermeiden. So sieht man in Abb. 22 den HSS für die Temperaturklassen auf der Nordhemisphäre.

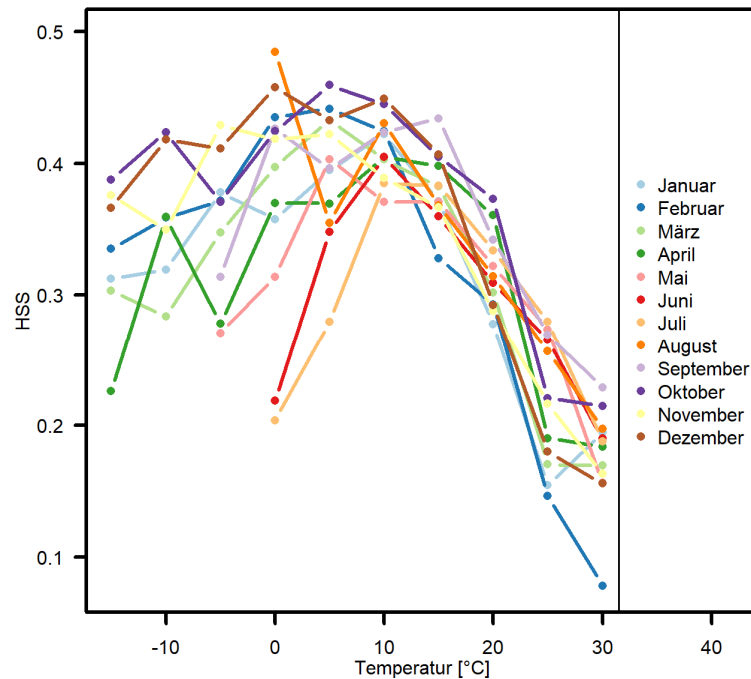


Abb. 22: Der HSS für die verschiedenen Temperaturklassen der Nordhalbkugel für alle Monate. Am Beispiel des Vorhersagemodells NEMS.

Es lässt sich erkennen, dass die höchsten HSS in den Sommermonaten bei Temperaturen zwischen 10 °C und 20 °C vorkommen, während in den Wintermonaten das Maximum meist schon bei 0 °C bis 10 °C erreicht wird. Bei Temperaturen über 20 °C nimmt der HSS stark ab.

6. Diskussion

6.1 Vergleich zwischen Modellvorhersage und Messung

Die Resultate haben gezeigt, dass die Modellvorhersagen von GFS, ICON, MF und NEMS sehr ähnliche Dynamiken aufweisen. So ist die Spannweite der MAD für die Jahresniederschlagssummen der einzelnen Vorhersagemodelle gering. Die deutlicheren Unterschiede zeigen sich in der globalen Verteilung. Die höheren Differenzen in den tropischen Regionen lassen sich durch die höheren Niederschlagssummen und die höhere Luftfeuchtigkeit erklären. Sehr hohe Differenzen treten in Südostasien auf, besonders an Standorten in Indonesien. In dieser Region kommt neben den tropischen Bedingungen durch die Inselstruktur der ausgeprägte maritime Einfluss dazu.

Neben diesen in allen Modellen auf gleiche Weise auftretenden Unterschieden der Vorhersagegenauigkeit lassen sich auch Unterschiede zwischen den einzelnen Modellvorhersagen bestimmen. In den nach Ländern aufgeschlüsselten Karten lassen sich große Unterschiede erkennen. Während MF beispielsweise in Indien auf Grund der genaueren Vorhersagen entlang der Küste als bestes der vier Vorhersagemodelle erscheint, schneidet MF in Südafrika im Vergleich zu den anderen Modellen am schlechtesten ab. Dies liegt insbesondere an den großen Differenzen entlang der Drakensberge im Südosten des Landes.

Die Darstellungsform, also ob jede Station einzeln gewertet wird oder Stationen in einem 2° oder 5° Gitternetz zusammengefasst werden, kann zu großen Unterschieden führen. Die MAD nimmt beispielsweise bei GFS von 271 mm (Stationen), 288 mm (2° Grid) bis 304 mm (5° Grid) stark zu (Tabelle 5). Ein ersichtlicher Grund dafür ist, dass, global gesehen, die Gebiete mit einer sehr hohen Dichte an Stationen, wie Nordamerika und Europa, relativ geringe MAD aufweisen. Da die Niederschlagsmessungen, die als Referenz für die Berechnung der Fehler genutzt werden, auf keiner einheitlichen Methode beruhen, kann auch dies eine Quelle von größeren Differenzen außerhalb der dicht beprobten Gebiete sein. Es ist somit möglich, dass die unterschiedlichen Messmethoden einen Einfluss auf die Genauigkeit der Niederschlagsdaten haben und dies wiederum zu einem scheinbaren Einfluss auf die Modellgenauigkeit führen kann. Ob die Vorhersagen aufgrund der hohen Messstationsdichte besser sind, lässt sich nicht daraus ableiten. Mit einem dichteren und einheitlicheren globalen Messnetz könnte diese Unsicherheit behoben werden.

Somit lässt sich kein eindeutiges Muster erkennen, welches jedem Land, bezogen auf die Jahresniederschlagssummen, ein bestes Vorhersagemodell zuordnet. Interessant ist, dass nur ICON in seinem eigenen Ursprungsland die beste Vorhersage macht. Auffällig ist auch, dass der Reanalysedatensatz ERA5 trotz der Mitbenutzung der erhobenen Messungen in vielen Ländern nicht als bestes Modell abschneidet.

Bei dem Vergleich zwischen den gemessenen und vorhergesagten Monatssummen zeigt sich, dass ERA5 überwiegend besser abschneidet als die Vorhersagemodelle. ERA5 weicht sowohl bei der MAD (11 von 12 Monaten), als auch bei der MD (9 von 12 Monaten) weniger von den Messwerten ab. So verbessert sich der Reanalysedatensatz, im Gegensatz zu den Vorhersagemodellen, bei höherer zeitlicher Auflösung. Die beobachtete höhere Vorhersagegenauigkeit in den Wintermonaten in Nordamerika deckt sich mit Ergebnissen einer Verifikation von GFS von Brown et al. (2014) für vier Bundesstaaten.

Bei den täglichen Niederschlagssummen nimmt die Qualität der Vorhersagen mit der Höhe der Niederschlagssummenkategorie ab. So erreicht beispielsweise GFS bei Niederschlagssummen von mehr als 1 einen HSS von 0.43, bei mehr als 10 mm einen HSS von 0.30 und bei mehr als 50 mm einen HSS von nur noch 0.08. Dies deckt sich mit einer Analyse für das Jahre 2017 (Meteoblue AG, 2018), welche HSS-Werte für GFS von 0.42, 0.30 und 0.12 gezeigt hat. Da das Erreichen der höheren beiden Niederschlagsklassen durch konvektive oder orographische Starkniederschläge begünstigt wird, lassen sich daraus die tiefen HSS und POD-Werte ableiten.

Bei der globalen Betrachtung der HSS Werte lässt sich feststellen, dass sich diese von der globalen Betrachtung der MAD unterscheiden. So zeigt sich beispielsweise bei Betrachtung des HSS ein starker Ost-West-Gradient in Australien, welcher bei der MAD nicht feststellbar ist. Durch die unterschiedliche Berechnung schneiden beim HSS trockene Gebiete nicht generell besser ab, als es Gebiete mit höheren Niederschlagssummen tun. Bei der

monatlichen Beurteilung der Vorhersagen anhand des HSS zeigte sich, dass der global gemittelte HSS in jedem Monat vergleichbar war, sich jedoch die globale Verteilung verändert.

Bei einem Großteil der Länder macht ICON, bezogen auf den HSS, die beste Niederschlagsvorhersage von den Vorhersagemodellen und in vielen Fällen auch beim Miteinbeziehen von ERA5. Eine mögliche Erklärung ist der Gitterzellenabstand von nur 13 km, welcher weniger als die Hälfte des Gitterzellenabstands der anderen betrachteten Modelle beträgt.

6.2 Variabilität zwischen den Vorhersagemodellen

Beim Vergleich der Jahressummen zeigt sich, dass sich die größten Unterschiede zwischen den vier Vorhersagemodellen und dem Reanalysedatensatz hauptsächlich in Äquatornähe befinden. Die monatliche Betrachtung macht deutlich, dass die Variabilität zwischen den Vorhersagemodellen stark jahreszeitenabhängig ist. Während die großen Unterschiede im Sommer nördlich des Äquators bis zum nördlichen Wendekreis auftreten, kommen sie im Winter vermehrt südlich des Äquators vor. Da in diesen Gebieten zu den genannten Jahreszeiten größere Niederschlagssummen zu erwarten sind, lässt sich auch die größere Spannweite der Vorhersagen darauf zurückführen.

Es zeigt sich, dass die Variabilität der Vorhersagemodelle in denjenigen Gebieten am größten ist, in welchen auch die Modellfehler am größten sind. Somit können die gewonnen Erkenntnisse die Fragestellung (3) beantworten.

6.3 Geografische und zeitliche Abhängigkeiten

Die Differenzen zwischen Modellvorhersage und Messungen nehmen bei allen Modellen mit der Geländehöhe zu. Die Dynamiken der Vorhersagemodelle und auch des Reanalysedatensatzes sind alle sehr ähnlich. Trotzdem lassen sich Unterschiede aufzeigen. Beispielsweise nehmen die Fehler bei GFS mit ansteigender Geländehöhe viel stärker zu, als dies beispielsweise bei ICON oder NEMS der Fall ist. Betrachtet wurde bei dieser Arbeit nur die absolute Geländehöhe. So wurde nicht zwischen weitläufigen Hochebenen und Gebirgen mit verschiedensten Höhenstufen auf kleinräumigem Gebiet unterschieden, obwohl die Niederschlagsentstehung und -intensität stark variiert. So wird die Niederschlagsmenge in weitläufigen Hochebenen vor allem durch den Feuchtigkeitsgehalt der Luft bestimmt und nimmt mit der Geländehöhe tendenziell ab, während die Niederschlagsmenge in Gebirgen durch orographische Prozesse mit der Geländehöhe zunimmt (Roe et al., 2003).

Mit zunehmender geografischer Breite nehmen die Differenzen zwischen Modellvorhersagen und Messungen ab. Bei der Untersuchung der geografischen Breite schneiden die Modelle vor allem auf der Nordhemisphäre sehr ähnlich ab, während sich auf der Südhemisphäre größere Unterschiede zeigen. Auffallend ist, dass ICON und NEMS, die bei steigender Geländehöhe am besten abschneiden, in äquatornäheren Gegenden grössere Differenzen aufweisen als die anderen Modelle. ICON und NEMS eignen sich für höhergelegene Gegenden besser, sind jedoch in Äquatornähe weniger geeignet als GFS und MF.

Die MAD nimmt mit steigender Distanz zum Meer ab. Die besten Vorhersagen wurden weit im Landesinneren gemacht. Eine Studie zu der Niederschlagsvorhersage des Modells SEAS5 in Südamerika, das wie ERA5 vom ECMWF stammt, kam zu einem gegenteiligen Schluss. Dort wurde die höchste Modellgenauigkeit unter anderem an der Küste von Kolumbien und Ecuador beobachtet (Gubler et al., 2020).

Bei der Analyse der Temperaturabhängigkeit zu der Qualität von Niederschlagsvorhersagen zeigte sich, dass die besten Vorhersagen bei 0 bis 15 °C erreicht werden. Die Abnahme des HSS bei höheren Temperaturen lässt sich zum einen durch den höheren Wassergehalt in der Atmosphäre bei höheren Temperaturen erklären, zum anderen durch Starkniederschläge infolge konvektiver Prozesse. So dehnen sich Luftmassen beim Erwärmen aus und werden durch konvektiven Transport angehoben. In den höheren, kälteren Schichten mit niedrigerem Luftdruck kommt es zu Kondensation und Niederschlägen (Stevens & Bony, 2013).

7. Schlussfolgerung

In dieser Bachelorarbeit wurden die Niederschlagsvorhersagen der NWVM GFS, ICON, MF, NEMS und der Reanalysedatensatz ERA5 mit den WMO-Niederschlagsmessungen des Jahres 2019 verglichen. Die Ergebnisse zeigen, dass bei der Niederschlagsvorhersage sowohl die absoluten als auch die relativen Fehler in Äquatornähe und in Gebieten mit hoher jährlicher Niederschlagssumme am höchsten sind. Anhand der Ergebnisse kann darauf geschlossen werden, dass die jährliche Niederschlagssumme in trockenen und eher kontinental geprägten Gebieten am besten vorhergesagt wird. Die Qualität der Vorhersagen aller Modelle nahm mit steigender Höhe ab und mit Entfernung zum Äquator zu. Küstennahe Gebiete weisen Modellvorhersagen mit größeren Differenzen zu den Messungen auf als küstenferne Gebiete.

Die zeitliche Begrenzung der Arbeit und die begrenzten öffentlich zugänglichen Informationen zur Parametrisierung des Niederschlags in den einzelnen Modellen erlaubten es nicht, eine noch genauere Kenntnis der Unterschiede zwischen den Modellen zu gewinnen. Eine präzise Aufarbeitung der Parametrisierung des Niederschlags der einzelnen Modelle könnte in einer künftigen Studie zu einer detaillierteren Unterscheidung der Vorhersagemodelle führen.

Da Niederschlagsmessungen immer mit großen Fehlerquellen verbunden sind, empfiehlt sich für eine weitere Arbeit die Miteinbeziehung von Satellitendaten oder Radarmessungen. So könnten die NWVM auch über die Fläche verglichen und verifiziert werden.

Der direkte Vergleich von globalen NWVM ist wichtig für die Wahl des besten Modells für eine bestimmte Region. Mit dem Betrachten der Modelle verschiedener Urheber konnten die aktuelle Leistungsfähigkeit, aber auch die Grenzen der numerischen Wettervorhersage aufgezeigt werden.

Literaturverzeichnis

- Attema, J. J., & Lenderink, G. (2014). The influence of the North Sea on coastal precipitation in the Netherlands in the present-day and future climate. *Climate Dynamics*, 42(1–2), 505–519. <https://doi.org/10.1007/s00382-013-1665-4>.
- Brown, J. D., Wu, L., He, M., Regonda, S., Lee, H., & Seo, D.-J. (2014). Verification of temperature, precipitation, and streamflow forecasts from the NOAA/NWS Hydrologic Ensemble Forecast Service (HEFS): 1. Experimental design and forcing verification. *Journal of Hydrology*, 519, 2869–2889. <https://doi.org/10.1016/j.jhydrol.2014.05.028>.
- CNRM, (Centre National de Recherches Météorologiques). (2014). ARPEGE—Centre National de Recherches Météorologiques. <https://www.umr-cnrm.fr/spip.php?article121&lang=fr>. Abgerufen am 19.06.2020.
- Damrath, U. (2002). Verifikation von Wettervorhersagen. *Deutscher Wetterdienst*, 28(2), 8–16.
- Dormann, C. F. (2013). *Parametrische Statistik*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-34786-3>.
- DWD, (Deutscher Wetter Dienst). (2017). Wettermodelle. Deutscher Wetterdienst. https://www.dwd.de/DE/fachnutzer/luftfahrt/download/produkte/wettermodelle/wettermodelle_download.pdf?__blob=publicationFile. Abgerufen am 02.07.2020.
- DWD, (Deutscher Wetter Dienst). (2020a). ICON (Icosahedral Nonhydrostatic) Model. Wetter und Klima - Deutscher Wetterdienst - Numerical weather prediction models -. https://www.dwd.de/EN/research/weatherforecasting/num_modelling/01_num_weather_prediction_modells/icon_description.html?nn=484268. Abgerufen am 10.06.2020.
- DWD, (Deutscher Wetter Dienst). (2020b). Qualität Vorhersage. Wetter und Klima - Deutscher Wetterdienst. https://www.dwd.de/DE/wetter/schon_gewusst/qualitaetvorhersage/qualitaetvorhersage_node.html. Abgerufen am 04.07.2020.
- Gubler, S., Sedlmeier, K., Bhend, J., Avalos, G., Coelho, C. a. S., Escajadillo, Y., Jacques-Coper, M., Martinez, R., Schwierz, C., de Skansi, M., & Spirig, C. (2020). Assessment of ECMWF SEAS5 Seasonal Forecast Performance over South America. *Weather and Forecasting*, 35(2), 561–584. <https://doi.org/10.1175/WAF-D-19-0106.1>.
- Häckel, H. (2012). *Meteorologie* (7.). utb.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., ... Thépaut, J. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, qj.3803. <https://doi.org/10.1002/qj.3803>.
- Isotta, F. A., Frei, C., Weilguni, V., Tadić, M. P., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., & Vertačnik, G. (2014). The climate of daily precipitation in the Alps: Development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *International Journal of Climatology*, 34(5), 1657–1675. <https://doi.org/10.1002/joc.3794>.
- Koblinger, S. (2013). *Korrektur von operationellen Niederschlagsdaten*. Universität Wien, Diplomarbeit. <https://doi.org/10.25365/thesis.27530>.
- Lin, Y.-L., Chiao, S., Wang, T.-A., Kaplan, M. L., & Weglarz, R. P. (2001). Some Common Ingredients for Heavy Orographic Rainfall. *WEATHER AND FORECASTING*, 16, 28.
- Meteoblue AG. (2018). Verification report for forecast and historical weather data. Meteoblue AG. <https://content.meteoblue.com/nl/research-development/processes/verification>. Abgerufen am 11.07.2020.
- Meteoblue AG. (2020a). Meteoblue Modelle. Meteoblue. <https://content.meteoblue.com/de/spezifikationen/datenquellen/wettersimulationsdaten/meteoblue-modelle>. Abgerufen am 19.06.2020.

- Meteoblue AG. (2020b). Reanalyse-Datensätze. Meteoblue. <https://content.meteoblue.com/de/spezifikationen/datenquellen/wettersimulationsdaten/reanalyse-datensaetze?/ru/switchlanguage/to/de/specifications/data-sources/weather-simulation-data/reanalysis-datasets&>. Abgerufen am 20.06.2020.
- NCEP, (National Center for Environmental Prediction). (2020a). Global Forecast System (GFS) | National Centers for Environmental Information (NCEI) formerly known as National Climatic Data Center (NCDC). <https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forecast-system-gfs>. Abgerufen am 10.06.2020.
- NCEP, (National Center for Environmental Prediction). (2020b). Product Description Document: NCEP Model Analyses & Guidance. https://mag.ncep.noaa.gov/docs/NCEP_PDD_MAG.pdf. Abgerufen am 19.06.2020.
- Papadopoulos, A., Chronis, T. G., & Anagnostou, E. N. (2005). Improving Convective Precipitation Forecasting through Assimilation of Regional Lightning Measurements in a Mesoscale Model. *Monthly Weather Review*, 133(7), 1961–1977. <https://doi.org/10.1175/MWR2957.1>.
- Platzman, G. W. (1979). The ENIAC Computations of 1950—Gateway to Numerical Weather Prediction. *American Meteorological Society*, 60(4), 302–312.
- Reinert, D., Prill, F., Frank, H., Denhard, M., & Zängl, G. (2020). Database Reference Manual for ICON and ICON-EPS. *Deutscher Wetterdienst*, 1.1.12. https://doi.org/10.5676/DWD_pub/nwv/icon_1.2.12.
- Roe, G. H., Montgomery, D. R., & Hallet, B. (2003). Orographic precipitation and the relief of mountain ranges: OROGRAPHIC PRECIPITATION AND RELIEF. *Journal of Geophysical Research: Solid Earth*, 108(B6). <https://doi.org/10.1029/2001JB001521>.
- South, A. (2011). rworldmap: A New R package for Mapping Global Data. *The R Journal*, 3(1), 35–43.
- Stevens, B., & Bony, S. (2013). Water in the atmosphere. *Physics Today*, 66(6), 29–34. <https://doi.org/10.1063/PT.3.2009>.
- Stull, R. (2017). *Practical Meteorology—An Algebra-based Survey of Atmospheric Science*. University of British Columbia.
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K.-L. (2018). A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. *Reviews of Geophysics*, 56(1), 79–107. <https://doi.org/10.1002/2017RG000574>.
- Wengenmayr, R., Friedrich, A., & Jaworski, V. (2017). Wie entsteht eine Wettervorhersage? (Nr. 6; Messen-Berechnen-Interpretieren). DWD, (Deutscher Wetterdienst). https://www.dwd.de/SharedDocs/broschueren/DE/presse/wettervorhersage_pdf.pdf?__blob=publicationFile&v=8. Abgerufen am 27.06.2020.
- WMO, (World Meteorological Organization). (2018). *Guide to meteorological instruments and methods of observation*. (8. Aufl.). World Meteorological Organization.
- Zhang, F., Yang, Z., Zhong, S., & Huang, Q. (2016). Exploring Mean Annual Precipitation Values (2003–2012) in a Specific Area (36°N–43°N, 113°E–120°E) Using Meteorological, Elevational, and the Nearest Distance to Coastline Variables. *Advances in Meteorology*, 2016, 1–13. <https://doi.org/10.1155/2016/2107908>.

Erklärung

Hiermit versichere ich, Ursula Huonder, dass ich die eingereichte Bachelorarbeit selbstständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommenen Inhalte als solche kenntlich gemacht habe.

Die eingereichte Bachelorarbeit ist oder war weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens. Sie wurde bisher nicht veröffentlicht.

Ort, Datum

Unterschrift