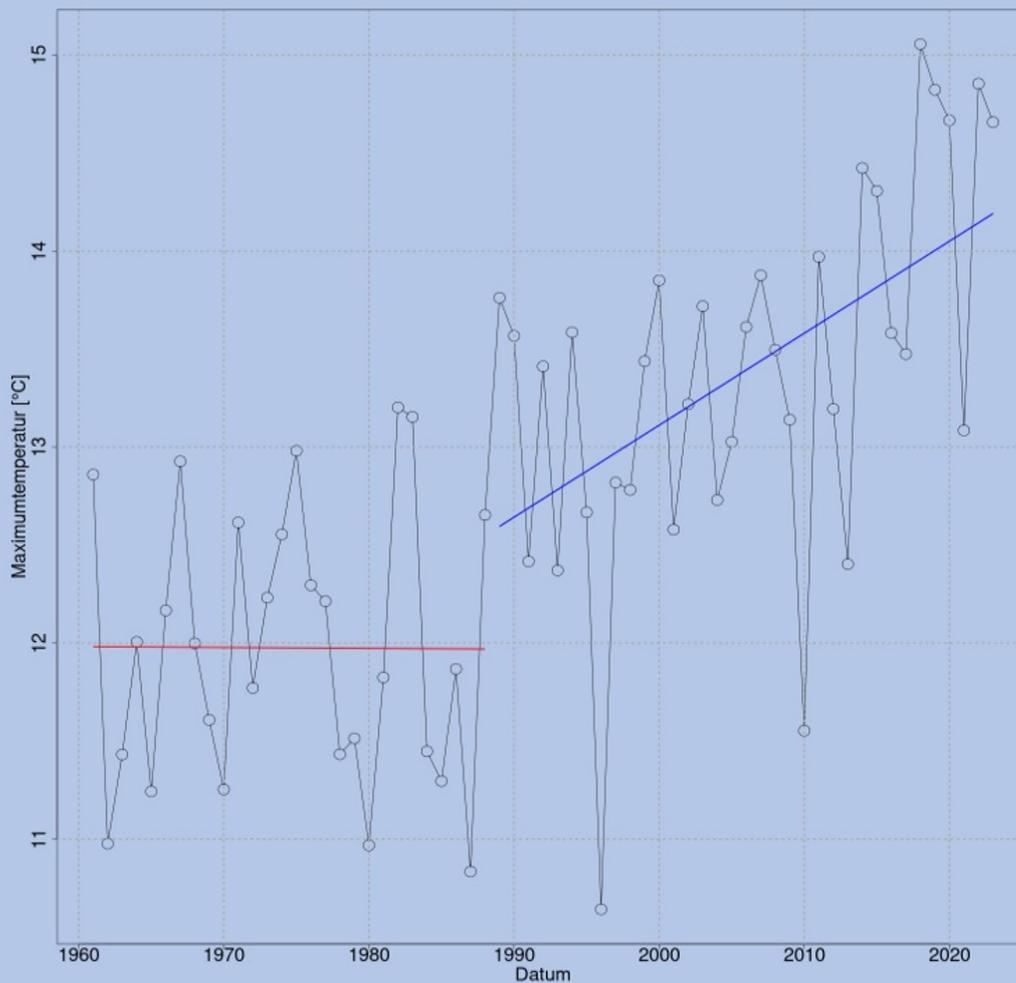


# Klima-Referenzdatensatz

## Version 3.0 (1961-2023)

Schriftenreihe, Heft 5/2025



Ableitung von Produkten aus COPERNICUS-Da-  
ten für kommunale Bedarfe und das Klimafol-  
gen-Monitoring: Fortschreibung des Klima-Refere-  
nzdatensatzes zur Version 3.0 zur Verwendung  
im Klimafolgenbereich

Dr. Philipp Körner; Dr. Uta Moderow, beide iamk GmbH

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b> .....	<b>9</b>
<b>2</b>	<b>Untersuchungsgebiet</b> .....	<b>10</b>
<b>3</b>	<b>Erstellung Stations- und Rasterdaten</b> .....	<b>11</b>
3.1	Datengrundlage .....	11
3.2	Datenaufbereitung .....	12
3.2.1	Metadaten .....	12
3.2.2	Stationsdaten .....	13
3.2.3	Aufbereitung von Rohdaten .....	13
3.3	Lücken füllen und Zeitreihen verlängern .....	16
3.3.1	Methodik zur Füllung der Lücken.....	17
3.3.2	Kreuzvalidierung Lückenfüllen .....	18
3.4	Berechnen von abgeleiteten Größen .....	19
3.5	Prüfung auf Homogenität.....	23
3.6	Rasterdatensatz erstellen.....	25
3.6.1	Methodik .....	25
3.6.2	Kreuzvalidierung.....	26
3.7	Monats- und Jahresdaten .....	27
3.7.1	Stationsdaten .....	27
3.7.2	Rasterdaten.....	28
3.8	Datenbereitstellung.....	28
<b>4</b>	<b>Vergleich mit Referenzdatensatz 2.0 und der Reanalyse ERA5-land</b> .....	<b>29</b>
4.1	Vergleich mit ERA5-land .....	29
4.1.1	Potentielle Verdunstung (ET) .....	30
4.1.2	Mittlere Temperatur .....	31
4.1.3	Mittlere Windgeschwindigkeit in 10 m Höhe .....	33
4.1.4	Globalstrahlung .....	35
4.1.5	Korrigierter Niederschlag .....	37
4.2	Vergleich mit Referenzdatensatz 2.0 .....	39
4.2.1	Potentielle Verdunstung.....	40
4.2.2	Grasreferenzverdunstung.....	41
4.2.3	Mittlere Temperatur .....	43
4.2.4	Minimum Temperatur .....	45
4.2.5	Maximum Temperatur .....	47
4.2.6	Mittlere Windgeschwindigkeit.....	49

4.2.7	Globalstrahlung .....	51
4.2.8	Korrigierter Niederschlag .....	54
4.2.9	Unkorrigierter Niederschlag.....	56
4.2.10	Dampfdruck .....	58
4.2.11	Relative Feuchte .....	60
4.3	Übersicht des Vergleiches .....	61
<b>5</b>	<b>Zusammenfassung .....</b>	<b>63</b>
	<b>Literaturverzeichnis.....</b>	<b>64</b>

## Abbildungsverzeichnis

Abbildung 1: Übersicht über das Untersuchungsgebiet im Referenzdatensatz 3.0.....	10
Abbildung 2: Höhenabhängigkeit der mittleren Temperatur der gemessenen Daten und des lückengefüllten Datensatzes .....	19
Abbildung 3: Zeitliche Verteilung der Bruchpunkte für die Maximum Temperatur.....	24
Abbildung 4: Mittlerer zeitlicher Verlauf aller Stationen für die Maximum Temperatur. Rote Linie: Mittlerer Trend vor dem häufigsten Bruchpunkt. Blaue Linie: Mittlerer Trend nach dem häufigsten Bruchpunkt. ....	25
Abbildung 5: Potentielle Verdunstung des Referenzdatensatzes 3.0 für 1961-1990.....	30
Abbildung 6: Potentielle Verdunstung aus ERA5-land für 1961-1990 .....	30
Abbildung 7: Differenz der potentiellen Verdunstung aus ERA5-land und Referenzdatensatz 3.0 1961-1990.....	31
Abbildung 8: Mittlere Temperatur des Referenzdatensatzes 3.0 für 1961-1990 .....	32
Abbildung 9: Mittlere Temperatur aus ERA5-land für 1961-1990 .....	32
Abbildung 10: Differenz der mittleren Temperatur aus ERA5-land und Referenzdatensatz 3.0 1961-1990.....	33
Abbildung 11: Mittlere Windgeschwindigkeit des Referenzdatensatzes 3.0 für 1961-1990.....	34
Abbildung 12: Mittlere Windgeschwindigkeit aus ERA5-land für 1961-1990 .....	34
Abbildung 13: Differenz der mittleren Windgeschwindigkeit aus ERA5-land und Referenzdaten- satz 3.0 1961-1990 .....	35
Abbildung 14: Globalstrahlung des Referenzdatensatzes 3.0 für 1961-1990 .....	36
Abbildung 15: Globalstrahlung aus ERA5-land für 1961-1990.....	36
Abbildung 16: Differenz der Globalstrahlung aus ERA5-land und Referenzdatensatz 3.0 1961-1990.....	37
Abbildung 17: Korrigierter Niederschlag des Referenzdatensatzes 3.0 für 1961-1990 .....	38
Abbildung 18: Niederschlag aus ERA5-land für 1961-1990.....	38
Abbildung 19: Differenz des Niederschlages aus ERA5-land und Referenzdatensatz 3.0 1961-1990.....	39
Abbildung 20: Potentielle Verdunstung des Referenzdatensatzes 3.0 für 1961-1990.....	40
Abbildung 21: Potentielle Verdunstung des Referenzdatensatzes 2.0 für 1961-1990.....	40
Abbildung 22: Differenz der potentiellen Verdunstung aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961-1990.....	41
Abbildung 23: Grasreferenzverdunstung des Referenzdatensatzes 3.0 für 1961-1990.....	42
Abbildung 24: Grasreferenzverdunstung des Referenzdatensatzes 2.0 für 1961-1990.....	42
Abbildung 25: Differenz der Grasreferenzverdunstung aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961-1990.....	43
Abbildung 26: Mittlere Temperatur des Referenzdatensatzes 3.0 für 1961-1990 .....	44

Abbildung 27: Mittlere Temperatur des Referenzdatensatzes 2.0 für 1961-1990 .....	44
Abbildung 28: Differenz der mittleren Temperatur aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961-1990 .....	45
Abbildung 29: Minimum Temperatur des Referenzdatensatzes 3.0 für 1961-1990 .....	46
Abbildung 30: Minimum Temperatur des Referenzdatensatzes 2.0 für 1961-1990 .....	46
Abbildung 31: Differenz der Minimum Temperatur aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961-1990 .....	47
Abbildung 32: Maximum Temperatur des Referenzdatensatzes 3.0 für 1961-1990 .....	48
Abbildung 33: Maximum Temperatur des Referenzdatensatzes 2.0 für 1961-1990 .....	48
Abbildung 34: Differenz der Maximum Temperatur aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961-1990 .....	49
Abbildung 35: Mittlere Windgeschwindigkeit des Referenzdatensatzes 3.0 für 1961-1990 .....	50
Abbildung 36: Mittlere Windgeschwindigkeit des Referenzdatensatzes 2.0 für 1961-1990 .....	50
Abbildung 37: Differenz der mittleren Windgeschwindigkeit aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961-1990 .....	51
Abbildung 38: Globalstrahlung des Referenzdatensatzes 3.0 für 1961-1990 .....	52
Abbildung 39: Globalstrahlung des Referenzdatensatzes 2.0 für 1961-1990 .....	52
Abbildung 40: Differenz der Globalstrahlung aus Referenzdatensatz 2.0 und Referenzdaten- satz 3.0 1961-1990 .....	53
Abbildung 41: Korrigierter Niederschlag des Referenzdatensatzes 3.0 für 1961-1990 .....	54
Abbildung 42: Korrigierter Niederschlag des Referenzdatensatzes 2.0 für 1961-1990 .....	55
Abbildung 43: Differenz des korrigierten Niederschlags aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961-1990 .....	55
Abbildung 44: Unkorrigierter Niederschlag des Referenzdatensatzes 3.0 für 1961-1990 .....	56
Abbildung 45: Unkorrigierter Niederschlag des Referenzdatensatzes 2.0 für 1961-1990 .....	57
Abbildung 46: Differenz des unkorrigierten Niederschlags aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961-1990 .....	57
Abbildung 47: Dampfdruck des Referenzdatensatzes 3.0 für 1961-1990 .....	58
Abbildung 48: Dampfdruck des Referenzdatensatzes 2.0 für 1961-1990 .....	58
Abbildung 49: Differenz der Dampfdruck aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961-1990 .....	59
Abbildung 50: Relative Feuchte des Referenzdatensatzes 3.0 für 1961-1990 .....	60
Abbildung 51: Relative Feuchte des Referenzdatensatzes 2.0 für 1961-1990 .....	60
Abbildung 52: Differenz der relativen Feuchte aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961-1990 .....	61

## Tabellenverzeichnis

Tabelle 1: Übersicht über die verwendeten Zeitreihen je Messgröße .....	12
Tabelle 2: Kennwerte der Kreuzvalidierung zur Abschätzung der Qualität der Lückenfüllung .....	18
Tabelle 3: Koeffizienten der Niederschlags-Korrekturfunktion in Abhängigkeit von der Niederschlagsart und der Geschützttheit der Messstation (nach RICHTER, 1995) .....	20
Tabelle 4: Jahresgang der Koeffizienten zur Berechnung der Globalstrahlung .....	21
Tabelle 5: Berechnete Rasterdatensätze für meteorologische und abgeleitete Größen .....	26
Tabelle 6: Kennwerte der Kreuzvalidierung zur Abschätzung der Qualität der räumlichen Interpolation .....	27
Tabelle 7: Prozentuale Korrektur des Niederschlages in Abhängigkeit der Stationslage und des Zeitraumes .....	28
Tabelle 8: Vergleich der Datensätze RefDat3 mit RefDat2 und RefDat3 mit ERA5-land.....	62

## **Abkürzungsverzeichnis**

LfULG	Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie
DWD	Deutscher Wetterdienst
CHMI	Tschechisches Hydrometeorologisches Institut
IMGW	polnisches Institut für Meteorologie und Wasserwirtschaft
TX	Maximum Temperatur
TM	Mittlere Temperatur
RF	Relative Feuchte
RR	Unkorrigierter Niederschlag
BG	Bewölkungsgrad
FM	Mittlere Windgeschwindigkeit
FX	Maximale Windgeschwindigkeit
DD	Dampfdruck
SN	Schneehöhe
SD	Sonnenscheindauer
ES	Sättigungsdampfdruck
SRTM	Shuttle Radar Topography Mission
WMO	World Meteorological Organisation

## 1 Einleitung

Langfristige, qualitativ gesicherte Datengrundlagen sind essenziell für ein fundiertes Klima- und Klimafolgen-Monitoring. Eine einheitliche Datengrundlage führt dazu, dass die Ergebnisse aus der Impactmodellierung besser interpretiert werden können. Werden beispielsweise für die Analyse von Temperaturentrends in einer Region unterschiedliche Datenquellen herangezogen, können schon Abweichungen in den Messmethoden oder der Datenaufbereitung zu widersprüchlichen Ergebnissen und Unsicherheiten führen. Ziel dieses Projektes ist es, den bestehenden klimatologischen Referenzdatensatz 2.0 des LfULG (KÖRNER et al., 2022), zu erweitern und bis einschließlich 2023 fortzuschreiben.

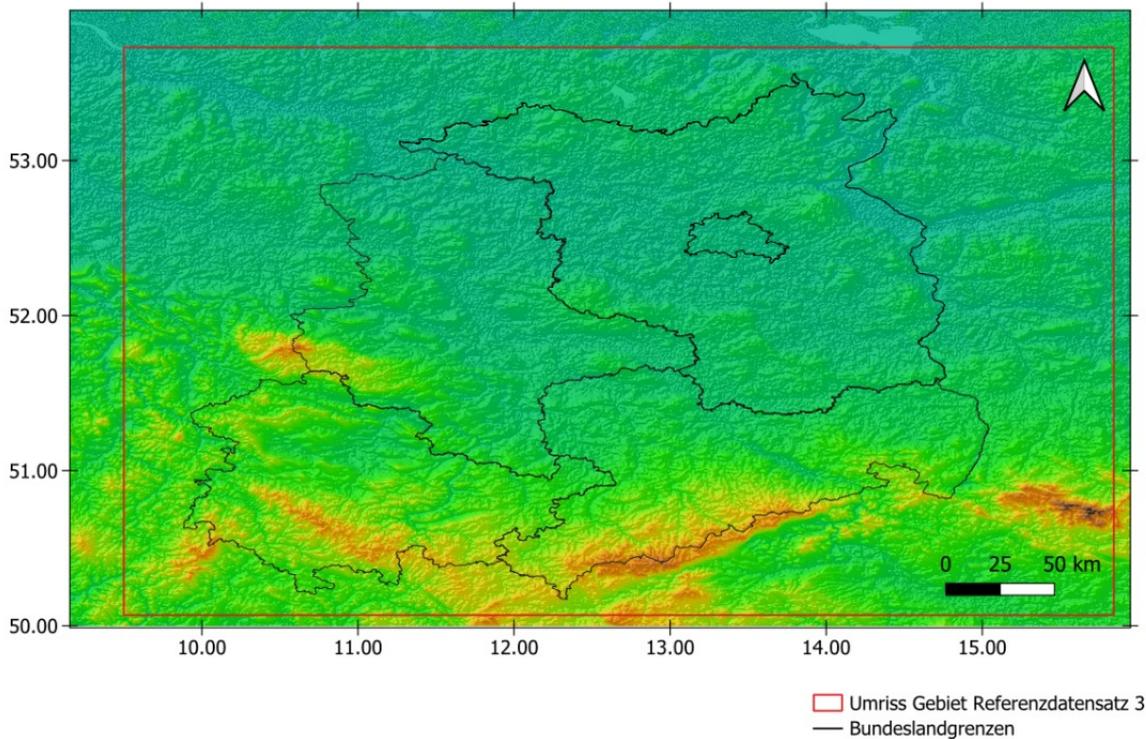
Der erweiterte Referenzdatensatz 3.0 umfasst stations- und rasterbasierte Teildatensätze (1x1 km Auflösung) für den Zeitraum vom 1. Januar 1961 bis zum 31. Dezember 2023. Die Datensätze liegen separat für Tages-, Monats- und Jahreswerte vor und decken die Bundesländer Sachsen, Sachsen-Anhalt, Thüringen sowie Brandenburg ab. Mit dieser Fortschreibung wird eine wesentliche Grundlage geschaffen, um die Auswirkungen des Klimawandels in der Region langfristig analysieren und geeignete Anpassungsstrategien entwickeln zu können.

Neben einem erweiterten Untersuchungsgebiet im Vergleich zum bisherigen Referenzdatensatz 2.0 wird für diesen Bericht die Datengrundlage erweitert: Einzelne neue Stationen im Untersuchungsgebiet kommen hinzu, bestehende Stationen wurden vom DWD nachdigitalisiert. Zusätzlich stehen erstmals aktuelle polnische Stationen frei zur Verfügung.

Die Methodik zur Verlängerung der Zeitreihen wurde verfeinert, so dass nun vor allem die Extremereignisse der jeweiligen Größe besser abgebildet sind. Eine detailliertere Beschreibung hierzu findet sich in den jeweiligen Kapiteln.

## 2 Untersuchungsgebiet

Das Untersuchungsgebiet umfasst die Länder Brandenburg, Sachsen-Anhalt, Sachsen und Thüringen vollständig. Ebenso ist Berlin vollständig enthalten. Aufgrund des dort dominierenden und hier nicht behandelten Stadtklimas ist das hier erstellte Produkt für Berlin weniger geeignet. Zusätzlich umfasst das Untersuchungsgebiet Teile der Bundesländer Niedersachsen, Mecklenburg-Vorpommern, Schleswig-Holstein, Hessen und Bayern. Ebenso sind Gebiete Tschechiens und Polens enthalten.



**Abbildung 1: Übersicht über das Untersuchungsgebiet im Referenzdatensatz 3.0**

Stationsdaten werden aus dem gesamten in Abbildung 1 rot umrandeten Gebiet verwendet, das Rasterprodukt umfasst ebenso das gesamte rot umrandete Gebiet.

### **3 Erstellung Stations- und Rasterdaten**

Dieses Kapitel beschreibt den zeitlichen und räumlichen Umfang und die Datenquellen der meteorologischen Größen. Weiterhin wird die Methodik beschrieben, wie die meteorologischen Daten aufbereitet werden, Lücken in den Daten gefüllt und die Zeitreihen auf den gesamten Untersuchungszeitraum verlängert werden. Aus diesen vollständigen Stationsdaten werden abschließend Rasterprodukte erstellt. Für Lückenfüllung und räumliche Interpolation wird außerdem eine umfassende Kreuzvalidierung durchgeführt, um Aussagen zur Qualität der Zeitreihen und Rasterprodukte treffen zu können.

#### **3.1 Datengrundlage**

Die Datengrundlage bilden meteorologische Daten in täglicher Auflösung für den Zeitraum 01.01.1961 bis 31.12.2023, soweit diese bis zum Stichtag 31.03.2024 verfügbar waren. Diese Daten umfassen Zeitreihen verschiedener meteorologischer Größen von insgesamt 2479 Stationen. Hierin enthalten sind 2116 Stationen des Deutschen Wetterdienstes (DWD), 227 Stationen des Tschechischen Hydrometeorologischen Institutes (CHMI) sowie 136 Stationen des polnischen Instituts für Meteorologie und Wasserwirtschaft (IMGW). Zusätzlich werden Zeitreihen unterschiedlicher meteorologischer Größen von insgesamt 192 weiteren Stationen als Stützstellen verwendet. Diese Stationen stammen aus verschiedenen Landesmessnetzen im Untersuchungsgebiet. Die Daten der nationalen Wetterdienste sind öffentlich zugänglich, die Daten der Landesmessnetze wurden von den Landesämtern für dieses Projekt zur Verfügung gestellt. Tabelle 1 gibt einen Überblick über die verwendeten Stationsdaten. Dargestellt ist jeweils die Anzahl der direkt in den Datensätzen vorhandenen Messreihen.

**Tabelle 1: Übersicht über die verwendeten Zeitreihen je Messgröße**

Größe	Abkürzung	Einheit	Anzahl Stationen	Bemerkungen
Minimum Temperatur	TN	°C	416	
Maximum Temperatur	TX	°C	418	
Mittlere Temperatur	TM	°C	422	
Relative Feuchte	RF	%	414	
Unkorrigierter Niederschlag	RR	mm	2.462	
Bewölkungsgrad	BG	Achtel	19	In Polen mangels Daten für Sonnenscheindauer und Strahlung als Ausgangsdaten für Sonnenscheindauer verwendet
Mittlere Windgeschwindigkeit	FM	m/s	221	
Maximale Windgeschwindigkeit	FX	m/s	159	Teilweise aus FM abgeleitet
Dampfdruck	DD	hPa	308	Für Polen und Tschechien aus TM und RF abgeleitet
Schneehöhe	SN	cm	2.295	Nur Stationsdaten
Sonnenscheindauer	SD	h	238	Nur Stationsdaten, Ausgangsgröße für Berechnung der Strahlung

## 3.2 Datenaufbereitung

Die folgenden Abschnitte umfassen die Darstellung der Datenaufbereitung der Tagesdaten. Die Erzeugung der Monats- und Jahresdaten ist unter Abschnitt 3.7 ausgeführt.

### 3.2.1 Metadaten

Der DWD und das CHMI bieten mit den Zeitreihen eine Reihe von Metadaten an. Unter anderem wird der Verlauf der Lage des Standortes mit geografischen Koordinaten sowie die Geländehöhe und die Höhe der Messung der Windgeschwindigkeit angegeben. Die zeitliche Änderung der Stationslage betrifft jede Station unterschiedlich stark. Einige Stationen haben keinen Standortwechsel zu verzeichnen, bei anderen Stationen beträgt dieser bis zu fünf Kilometer sowie bis zu 100 Meter für die Geländehöhe. Die Lage der Stationen spielt für die räumliche Interpolation sowie für die Berechnung abgeleiteter Größen wie der Strahlung oder der Verdunstung eine entscheidende Rolle. Aus Konsistenzgründen werden immer die aktuellsten Metadaten der jeweiligen Stationen verwendet.

Für die polnischen Stationen des IMGW lagen zum Stichtag 31.03.2024 lediglich geografische Koordinaten ohne Höhenangabe vor. Es gab weder Daten zum Verlauf noch Angaben zur Geländehöhe oder zur Messhöhe der Windgeschwindigkeit. Die geografischen Koordinaten haben eine Genauigkeit von einem

sechzigstel Grad und somit etwas weniger als 2 km. Die Geländehöhe wurde für die polnischen Stationsdaten mittels SRTM (Shuttle Radar Topography Mission)-Daten abgeleitet.

### 3.2.2 Stationsdaten

Zeitreihen, die zur Erstellung des Referenzdatensatzes verwendet wurden, mussten grundsätzlich zwei Kriterien erfüllen: Zum einen müssen Sie eine Mindestlänge von wenigstens 365 Tagen und somit einem Jahr aufweisen. Kürzere Zeitreihen können mangels Messwerten über ein vollständiges Jahr nur unzureichend verlängert werden. Das zweite Kriterium ist die Korrelation der jeweiligen Zeitreihe mit Nachbarstationen. Zeitreihen, die über unplausibel geringe Korrelationskoeffizienten mit Nachbarstationen verfügen, werden für die weiteren Berechnungen nicht verwendet.

Die verbliebenen Datensätze (Tabelle 1) werden weiter bereinigt. Daten, die außerhalb des plausiblen Wertebereiches der jeweiligen Messgröße liegen, werden entfernt. Ebenso werden Temperaturdaten an einzelnen Stationen an den Tagen nicht verwendet, an denen das Kriterium

$$TX \geq TM \geq TN$$

(1)

nicht erfüllt ist. Insgesamt ist die Plausibilität und Konsistenz der Daten der nationalen Wetterdienste sehr hoch, so dass nur einige wenige Zeitpunkte einzelner Stationen die Kriterien nicht erfüllen und entsprechend für die weitere Verarbeitung nicht verwendet werden.

### 3.2.3 Aufbereitung von Rohdaten

Für den Großteil der Zeitreihen ist die Aufbereitung an dieser Stelle abgeschlossen. Für die Windgeschwindigkeiten, Dampfdruck und Sonnenscheindauer ist ein weiterer Aufbereitungsschritt nötig.

#### ■ Windgeschwindigkeit

Sowohl die mittlere als auch die maximale Windgeschwindigkeit werden in der Standardhöhe von 10 m gemessen. Allerdings weicht ein Teil der Messungen von dieser Höhe ab. Diese Abweichung kann sich je nach Station über die Zeit ändern. Für die Stationen des DWD und des CHMI liegen diese zeitlichen Verläufe der Messhöhen vor. Für die Daten des IMGW gibt es keine derartigen Informationen, so dass von der Standardmesshöhe von 10 m über den gesamten Zeitraum ausgegangen werden muss. Diese Annahme ist mit einem unbekanntem Fehler behaftet, so dass sowohl die Zeitreihen als auch die Rasterprodukte für den polnischen Teil des Gebietes mit einer höheren Unsicherheit versehen sind.

Mithilfe des logarithmischen Windprofils kann die Windgeschwindigkeit auf die Messhöhe von 10 m umgerechnet werden. Es werden Parameter für den WMO-Standard kurzes Gras verwendet.

$$u_2 = u_1 \cdot \frac{\log \frac{z_2 - d_2}{z_{02}}}{\log \frac{z_1 - d_1}{z_{01}}} \quad (2)$$

$u_2$  *Windgeschwindigkeit in Zielhöhe*

$u_1$  *Windgeschwindigkeit in Messhöhe*

$z_2$  *Zielhöhe, also 10 Meter.*

$z_1$  *Messhöhe aus den Metadaten.*

$z_{01}$  *dynamische Rauigkeitslänge für Gras, in dem Fall 0,02 Meter.*

$z_{02}$  *dynamische Rauigkeitslänge für Gras, in dem Fall 0,02 Meter.*

$d_1$  *Verdrängungshöhe und wird mit 0,1 Metern angesetzt.*

$d_2$  *Verdrängungshöhe und wird mit 0,1 Metern angesetzt.*

Durch die Berechnung des logarithmischen Windprofils werden sowohl die mittlere (FM) als auch die maximale (FX) Windgeschwindigkeit einheitlich auf 10 m Höhe reduziert.

Der DWD und das CHMI stellen neben der mittleren Windgeschwindigkeit auch Daten der maximalen Windgeschwindigkeit zur Verfügung. Diese Daten fehlen beim IMGW. Um Datenlücken zu vermeiden, werden die maximalen Windgeschwindigkeiten für alle verwendeten polnischen Zeitreihen aus der mittleren Windgeschwindigkeit rekonstruiert.

Diese Rekonstruktion erfolgt mithilfe von Gradient Boosting Modellen. Es werden alle Stationen zum Training verwendet, an denen sowohl die mittlere als auch die maximale Windgeschwindigkeit bekannt ist. Zielgröße des Modells ist die maximale Windgeschwindigkeit. Unabhängige Größen (Prädiktoren) sind die anderen meteorologischen Größen sowie die geografische Lage der Stationen. Die wichtigste Größe hierbei ist die mittlere Windgeschwindigkeit.

Insgesamt stehen etwas weniger als 1,5 Millionen Tagesdaten zum Anlernen des Modells zur Verfügung. Eine unabhängige Kreuzvalidierung zeigt hierbei sehr gute Ergebnisse. Hierbei wurde das Modell für 64 Stationen angelernt. Für weitere 63 Stationen wurden es angewendet und die Modellergebnisse mit den Messungen überprüft: Der mittlere Fehler weist einen Wert von 1,5 m/s auf. Die systematische Abweichung (Bias) beträgt weniger als 0,1 m/s. Diese Werte sind besser als die der räumlichen Interpolation der maximalen Windgeschwindigkeit. Um die maximale Qualität zu erreichen, werden für die polnischen Stationen die auf diese Weise rekonstruierten Zeitreihen der maximalen Windgeschwindigkeit verwendet.

## ■ Dampfdruck

Werte des Dampfdruckes werden vom DWD angeboten, nicht aber vom CHMI und vom IMGW. Der Dampfdruck lässt sich jedoch aus Temperatur (TM) und relativer Feuchte (RF) ableiten. Hierfür wird der Sättigungsdampfdruck (ES) mithilfe der Magnus-Formel bestimmt.

$$ES = c_1 \cdot l \frac{c_2 \cdot TM}{c_3 + TM} \quad (3)$$

Für Temperaturen größer gleich 0 °C wird der folgende Parametersatz verwendet:

$$c_1 \quad 6,112 \text{ hPa}$$

$$c_2 \quad 17,62$$

$$c_3 \quad 243,12 \text{ °C}$$

Für Temperaturen kleiner 0 °C wird der folgende Parametersatz verwendet:

$$c_1 \quad 6,112 \text{ hPa}$$

$$c_2 \quad 22,46$$

$$c_3 \quad 272,62 \text{ °C}$$

Der Dampfdruck (DD) berechnet sich anschließend aus der folgenden Formel:

$$DD = \frac{RF}{100} * ES \quad (4)$$

## ■ Sonnenscheindauer

Die Stationsdaten des DWD und des CHMI verfügen über Daten zur Sonnenscheindauer. Für die polnischen Stationen liegen solche Daten zum Zeitpunkt der Datenbeschaffung nicht vor. Allerdings stehen Daten zum Bedeckungsgrad an 19 Stationen zur Verfügung. Aus dem Bedeckungsgrad lässt sich die Sonnenscheindauer ableiten.

Der Ableitung liegt die folgende Annahme zu Grunde: Der Bedeckungsgrad gibt die Reduzierung der maximal möglichen Sonnenscheindauer an. Wenn die maximal mögliche Sonnenscheindauer an einem Tag beispielsweise acht Stunden beträgt und der Bedeckungsgrad zwei Achtel beträgt, dann wird eine Sonnenscheindauer von sechs Stunden angenommen. Die Sonnenscheindauer berechnet sich über die folgenden Schritte:

$$d = 23.45 \cdot \sin(360/365 \cdot (284 + n)) \cdot 180/\pi \quad (5)$$

*d* Deklinationswinkel, der Winkel, unter dem die Sonne zur Erdachse steht (abhängig von der Jahreszeit).

*n* Der Tag im Jahr (1 bis 365; für Schaltjahre bis 366).

$$\omega = \arccos(-\tan(\phi) \cdot \tan(d)) \quad (6)$$

$\omega$  *Stundenwinkel, der mit dem Breitengrad und Deklinationswinkel in Beziehung steht. Er beschreibt die Zeitspanne zwischen dem Sonnenhöchststand und Sonnenaufgang/-untergang.*

$\phi$  *geografischer Breitengrad des Standortes*

$$S_0 = 24/\pi \cdot \omega \quad (7)$$

$S_0$  *Maximal mögliche Sonnenscheindauer*

$$SD = S_0 \cdot BG/8 \quad (8)$$

Anschließend berechnet sich die tatsächliche Sonnenscheindauer aus dem Bedeckungsgrad und der maximalen Sonnenscheindauer.

Die für die polnischen Stationen berechneten Sonnenscheindauern zeigen im Vergleich zu den tschechischen und deutschen Stationen eine systematische Abweichung von im Mittel mehr als 30 %. Das heißt die Sonnenscheindauer und alle abgeleiteten Größen wie die Strahlung und Verdunstung wären für die polnischen Gebiete sehr deutlich erhöht. Da diese Unterschiede meteorologisch nicht erklärt werden können, werden die Daten der Sonnenscheindauer und deren abgeleitete Größen für die polnischen Stationen im weiteren Datenverarbeitungsprozess nicht berücksichtigt.

Nach Durchführung aller in diesem Abschnitt 3.2 angewendeten Qualitätskriterien und Umformungen steht ein Datensatz zur Verfügung, der für die Lückenfüllung und Verlängerung der Zeitreihen bereit ist.

### 3.3 Lücken füllen und Zeitreihen verlängern

Die bis zum hier beschriebenen Stand aufbereiteten Zeitreihen weisen Datenverfügbarkeiten zwischen einem und 63 Jahren auf. In diesem Aufbereitungsschritt werden die Zeitreihen aller Stationen und aller meteorologischen Größen vollständig gefüllt, so dass anschließend alle Zeitreihen eine Datenverfügbarkeit von 63 Jahren im Zeitraum 1961 bis 2023 aufweisen.

Zur Anwendung kommen hierfür Gradient-Boosting-Modelle. Hierfür wird für jede unvollständige Zeitreihe ein Modell erstellt. Dieses Modell hat die jeweilige unvollständige Zeitreihe als Zielgröße. Unab-

hängige Größen (Prädiktoren) sind jeweils die gleiche Messgröße anderer Stationen. Zusätzlich fungieren auch die anderen meteorologischen Größen als Prädiktoren. Das jeweilige Modell wird für alle Zeitpunkte, an denen Daten der Zielgröße vorhanden sind, trainiert (angelernt). Anschließend wird das Modell für alle Zeitpunkte, an denen für die Zielgröße keine Daten vorhanden sind, angewendet. Die so erzeugten Daten vervollständigen die jeweilige Zeitreihe. Ausführlicher ist das Verfahren in KÖRNER ET AL. 2018 beschrieben. In ähnlicher Form ist diese Methodik beim Referenzdatensatz 2.0 (LfULG, 2021) angewandt wurden.

Zur Unterscheidung zwischen originalen Messdaten und rekonstruierten Daten wird neben den vollständigen Zeitreihen ein Produkt angeboten, welches per flagging über die Datenherkunft aufklärt. Unterschieden wird dabei zwischen gemessen, lückengefüllt und verlängert.

### **3.3.1 Methodik zur Füllung der Lücken**

Für jede Zeitreihe jeder meteorologischen Größe wird ein eigenes Modell aufgebaut. Dieses Modell wird für den Zeitraum, in dem die jeweilige Zeitreihe Messdaten aufweist, trainiert. Das Vorgehen soll hier einmal anhand einer Beispielzeitreihe des Niederschlages dargestellt werden. Diese Niederschlagszeitreihe weist von 1970 bis 1980 Messdaten auf. Das Modell wird nun für genau diesen Zeitraum trainiert. Die wichtigsten Prädiktoren sind Niederschlagszeitreihen der umliegenden Stationen, wobei lange Zeitreihen bevorzugt werden, die eine hohe Korrelation mit der zu füllenden Zeitreihe aufweisen. Wichtig ist, dass für die Prädiktoren sowohl in der Trainingsphase, hier 1970 bis 1980, als auch in der Anwendungsphase, hier 1961–1969 und 1981–2023, Daten vorliegen. Weitere Prädiktoren sind mittels räumlicher Interpolation bestimmte Niederschlagswerte der jeweiligen Stationen (Methode siehe Abschnitt 3.6.1). Zusätzlich kommen noch die weiteren meteorologischen Größen der Nachbarstationen hinzu, wie Temperatur, Windgeschwindigkeit und relative Luftfeuchte, Wieder werden diejenigen Prädiktoren ausgewählt, die die höchsten Korrelationen mit der Zielgröße aufweisen. Ist das Modell angelernt, dann wird es auf die Zeitpunkte angewendet, in der die Zielgröße keine Messdaten aufweist. In diesem Beispiel also 1961–1969 und 1981–2023. Ergebnis sind die modellierten Werte der Zeitreihe für diese Zeitpunkte.

Je nach Datenverfügbarkeit der Zeitreihe wird dieses Vorgehen mehrmals durchgeführt: Wenn im o.g. Beispiel einige Nachbarstationen im Zeitraum 1961–1980 Daten aufweisen und andere von 1970–2023, dann werden zwei Modelle angelernt: Das erste Modell mit den Stationen 1961–1980 wird von 1970–1980 angelernt und von 1961–1969 für die Lückenfüllung angewendet. Das zweite Modell wird mit anderen Stationen als Prädiktoren für 1970–1980 angelernt und dann für den Zeitraum 1981–2024 angewendet. Insgesamt werden so zur Lückenfüllung aller Zeitreihen aller meteorologischen Größen einige 10.000 Modelle trainiert und angewendet.

Die auf die beschriebene Weise erreichte Qualität der Lückenfüllung ist im Mittel sehr hoch, weist aber im Bereich der Extreme wie beispielsweise der Extremniederschläge größere Abweichungen auf. Die Abweichungen treten in der Form auf, dass die modellierten Extreme weniger stark sind als die gemessenen Extreme. Um diese Abweichungen zu minimieren, wurde zusätzlich ein Quantil-Quantil-Mapping (qq-Mapping) vorgenommen. Mithilfe der im folgenden Abschnitt 3.3.2 dargestellten Kreuzvalidierung ist es möglich, gemessene und modellierte Werte direkt miteinander zu vergleichen. Diese Wertepaare aus gemessenen und modellierten Werten werden genutzt, um mittels Splines ein qq-Mapping durchzuführen. Dieses qq-Mapping reduziert die Qualitätskriterien  $R^2$  und RMSE minimal, verbessert aber die Abschätzung der Extremwerte erheblich. Auch wird der mittlere Fehler (MAE) und der BIAS dadurch reduziert.

### 3.3.2 Kreuzvalidierung Lückenfüllen

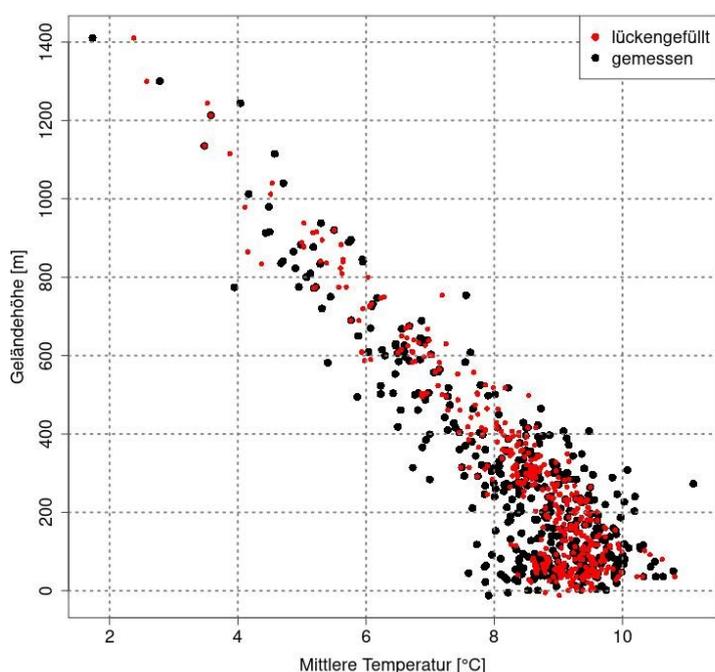
Zur Bewertung der Qualität der Lückenfüllung und der Verlängerung der Zeitreihen wurde folgende Validierungsstrategie verwendet: Jede Zeitreihe einer meteorologischen Größe wurde in der Mitte in zwei gleich große Teile aufgeteilt. Ein Gradient-Boosting-Modell wurde zunächst mit der ersten Hälfte der Daten trainiert und anschließend genutzt, um die zweite Hälfte vorherzusagen. Dieser Prozess wurde umgekehrt wiederholt, sodass am Ende für jeden Messwert ein vom Modell berechneter Wert vorlag. Die Ergebnisse sind in Tabelle 2 dargestellt. Diese zeigen die Qualitätskriterien nach dem in Abschnitt 3.3.1 beschriebenen qq-Mapping.

**Tabelle 2: Kennwerte der Kreuzvalidierung zur Abschätzung der Qualität der Lückenfüllung**

Größe	RMSE	MAE	BIAS	$R^2$
TN	0,91 °C	0,67 °C	-0,005 °C	0,982
TX	0,71 °C	0,51 °C	-0,003 °C	0,994
TM	0,59 °C	0,43 °C	-0,004 °C	0,994
RF	4,16 %	3,19 %	0,02 %	0,87
RR	1,61 mm/d	0,58 mm/d	-0,002 mm	0,85
FM	0,58 m/s	0,43 m/s	-0,005 m/s	0,86
FX	1,47 m/s	1,08 m/s	-0,08 m/s	0,86
DD	0,48 hPa	0,34 hPa	-0,005 hPa	0,986
SN	1,13 cm	0,26 cm	-0,02 cm	0,90
SD	1,03 h	0,67 h	0,01 h	0,942

Die statistischen Eigenschaften (Minimum, Maximum, Varianz etc.) der Zeitreihen verändern sich beim Verlängern derselben: So haben Temperaturzeitreihen, die beispielsweise nur für den Zeitraum 1961–1970 Messdaten aufweisen, in diesem Zeitraum niedrigere Minimum- und Maximum- und Mittelwerte als in der verlängerten Zeitreihe 1971–2023. Grundlegende physikalische Zusammenhänge bleiben aber erhalten. Dazu gehört die Höhenabhängigkeit der Temperatur. Abbildung 2 zeigt die Abhängigkeit

der Temperatur mit der Geländehöhe des gemessenen Datensatzes (schwarze Punkte) und des vollständig gefüllten Datensatzes (rote Punkte). Zu sehen sind insbesondere zwei Dinge: Zum einen bleibt der grundsätzliche Zusammenhang erhalten: Je höher eine Station gelegen ist, desto kälter ist es dort im Mittel. Dieser Temperaturgradient beträgt etwa 0,6 Kelvin pro 100 Höhenmeter. Zusätzlich ist zu sehen, dass die schwarzen Punkte, also die Messdaten, stärker streuen als der gefüllte Datensatz. Ursache hierfür ist, dass die Zeiträume der Messungen teilweise mehrere Jahrzehnte auseinanderliegen. Beispielsweise misst eine Station 1961–1970 und eine andere 2021–2023. Jede Station spiegelt damit die mittleren Verhältnisse der Messdauer wider. Durch den Anstieg der Durchschnittstemperatur im Untersuchungsgebiet um etwa 2 Kelvin im Untersuchungszeitraum ist die Streuung entsprechend groß.



**Abbildung 2: Höhenabhängigkeit der mittleren Temperatur der gemessenen Daten und des lückengefüllten Datensatzes**

### 3.4 Berechnen von abgeleiteten Größen

In diesem Abschnitt wird die Methodik vorgestellt, mit deren Hilfe abgeleitete Größen berechnet werden. Abgeleitete Größen sind jene Größen, die die Wetterdienste selbst nicht zur Verfügung stellen, die aber aus den zur Verfügung gestellten Größen abgeleitet werden können.

#### ■ Sättigungsdampfdruck ES

Der Sättigungsdampfdruck ES ist von der Temperatur (TM) abhängig und wird somit für alle TM-Zeitreihen berechnet. Es wird wieder die Magnus-Formel angewendet, Formel (3), wie sie bei der Berechnung des Dampfdruckes in Abschnitt 3.2.3 beschrieben ist.

## ■ Korrigierter Niederschlag RK

Die Korrektur des Niederschlages erfolgt nach RICHTER (1995). Die Korrekturverfahren unterscheiden sich hierbei nach zeitlicher Auflösung der Daten. In diesem Abschnitt wird zunächst auf die Korrektur der Tagesdaten eingegangen.

Der korrigierte Niederschlag berechnet sich dabei wie folgt:

$$RK = RR + b \cdot RR^\varepsilon \quad (9)$$

$RR$  unkorrigierter Niederschlag

$b, \varepsilon$  Koeffizienten der Korrekturfunktion

Die Koeffizienten  $b$  und  $\varepsilon$  sind von der Geschüttheit der Station sowie der Art des Niederschlages abhängig und lassen sich aus der Tabelle 3 entnehmen.

**Tabelle 3: Koeffizienten der Niederschlags-Korrekturfunktion in Abhängigkeit von der Niederschlagsart und der Geschüttheit der Messstation (nach RICHTER, 1995)**

Koeffizient	b				ε
	frei	leicht geschützt	mäßig geschützt	stark geschützt	
Regen (Sommerhalbjahr)	0,345	0,31	0,28	0,245	0,38
Regen (Winterhalbjahr)	0,34	0,28	0,24	0,19	0,46
Mischniederschlag	0,535	0,39	0,305	0,185	0,55
Schnee	0,72	0,51	0,33	0,21	0,82

Bei Tagesmitteltemperaturen über 3,0 °C wird von Regen, bei unter -0,7 °C von Schnee ausgegangen. Liegt die Temperatur dazwischen, so ist von Mischniederschlag auszugehen. Die Temperaturen (TM) an den Stationen zur Bestimmung der Niederschlagsart wurden mittels „Thin Plate Splines“-Verfahren interpoliert. Näheres dazu ist in Abschnitt 3.6 zu finden.

## ■ Globalstrahlung RG

Die Berechnung der Globalstrahlung richtet sich nach dem Merkblatt DWA-M 504-1. Das Merkblatt schlägt zur Berechnung der extraterrestrischen Strahlung die folgenden Gleichungen vor:

$$R_0 = 245 \cdot (9,9 + 7,08 \cdot \sin \zeta + 0,18 \cdot (\phi - 51) \cdot (\sin \zeta - 1)) \quad (10)$$

mit

$$\zeta = 0,0172 \cdot n - 1,39 \quad (11)$$

$n$  Der Tag im Jahr (1 bis 365; für Schaltjahre bis 366)

$\phi$  geografische Breitengrad des Standorts

Als Vereinfachung zur Berechnung der maximalen Sonnenscheindauer (Formeln (5) bis (7)) schlägt das Merkblatt die folgende Gleichung vor:

$$S_0 = 12,3 + \sin \zeta \cdot \left(4,3 + \frac{\phi - 51}{6}\right) \quad (12)$$

Die Globalstrahlung berechnet sich nun aus

$$RG = R_0 \cdot \left(a + b \cdot \frac{SD}{S_0}\right) \quad (13)$$

mit den Koeffizienten a und b nach Tabelle 4.

**Tabelle 4: Jahresgang der Koeffizienten zur Berechnung der Globalstrahlung**

	JAN	FEB	MRZ	APR	MAI	JUN	JUL	AUG	SEP	OKT	NOV	DEZ
a	0,19	0,19	0,19	0,19	0,19	0,19	0,19	0,19	0,19	0,19	0,19	0,19
b	0,54	0,57	0,57	0,57	0,57	0,55	0,54	0,53	0,54	0,54	0,54	0,53

Die Globalstrahlung RG wird in den Einheiten  $\frac{W}{m^2}$  sowie  $\frac{J}{cm^2}$  berechnet und zur Verfügung gestellt.

#### ■ Potentielle Verdunstung ET

Die potentielle Verdunstung wird nach Turc-Wendling im Merkblatt DWA-M 504-2 berechnet. Sie ist von den meteorologischen Größen Temperatur (TM) und Globalstrahlung (RG) abhängig und wird wie folgt berechnet:

$$ET = \frac{s}{s + \gamma} \cdot \left(0,71 \cdot \frac{RG}{L} + 0,27 \cdot KF\right) \quad (14)$$

$s$  Steigung der Sättigungsdampfdruckkurve

$\gamma$  Psychrometerkonstante 0,65 hPa/K

$L$  Verdampfungswärme

$KF$  Küstenfaktor

mit

$$s = ES \cdot \frac{c_4}{(c_5 + TM)^2} \quad (15)$$

$c_4$  4284 bei Temperaturen größer gleich 0 °C, 6123 bei kleiner 0 °C

$c_5$  243.12 °C bei Temperaturen größer gleich 272.62 °C bei kleiner 0 °C

und

$$L = 249,8 - 0,242 \cdot TM \quad (16)$$

und

$$KF = 1 - 0,8 \cdot \left( \phi - 53 - \frac{\lambda - 8}{6} \right) \quad (17)$$

$\lambda$  geografische Länge

mit  $KF=1$  bei  $KF>1$  und  $KF=0,5$  bei  $KF<0,5$ .

Die potentielle Verdunstung wird in den Einheiten  $W/m^2$  sowie mm zur Verfügung gestellt.

Aufgrund der meteorologischen Eingangsgrößen steht die potentielle Verdunstung für alle Stationen zur Verfügung, die sowohl Zeitreihen für TM als auch SD enthalten. Da die Stationen mit SD eine Teilmenge der Stationen mit TM sind, entspricht die Anzahl der Stationen mit ET-Zeitreihen den Stationen mit SD-Zeitreihen.

### ■ Grasreferenzverdunstung GR

Die Grasreferenzverdunstung wird nach Penman-Monteith abgeleitet und nach Merkblatt DWA-M 504-2 wie folgt berechnet:

$$GR = \frac{s \cdot RN^*}{s + \gamma^*} + \frac{90 \cdot \gamma}{s + \gamma^*} \cdot FM_2 \cdot \frac{ES}{TM + 273} \cdot \left( 1 - \frac{RF}{100\%} \right) \quad (18)$$

$s$  Steigung der Sättigungsdampfdruckkurve

$\gamma$  Psychrometerkonstante 0,65 hPa/K

$\gamma^*$  modifizierte Psychrometerkonstante 0,65 hPa/K

$FM_2$  Mittlere Windgeschwindigkeit in 2 Metern Höhe

$RN^*$  Verdunstungsäquivalent der Nettostrahlung

Die dafür nötigen Eingangsgrößen werden wie folgt berechnet:

$$\gamma^* = \gamma \cdot (1 + 0,34 \cdot FM_2) \quad (19)$$

$$RN^* = 0,004 \cdot RN \quad (20)$$

*RN* Nettostrahlung

$$RN = (1 - \alpha) \cdot RG - \sigma \cdot (TM + 273,15)^4 \cdot (0,1 + 0,9 \cdot \frac{SD}{S_0} \cdot (0,34 - 0,044 \cdot \sqrt{DD})) \quad (21)$$

$\alpha$  Albedo von 0,23

$\sigma$  Stefan-Boltzmann-Konstante

Aufgrund der Abhängigkeit der Grasreferenzverdunstung (GR) von den meteorologischen Größen FM, DD, SD und TM lässt sich diese nur für die Schnittmenge dieser Zeitreihen berechnen. Das entspricht insgesamt 161 Stationen. Die Größe GR wird in den Einheiten W/m<sup>2</sup> und mm zur Verfügung gestellt.

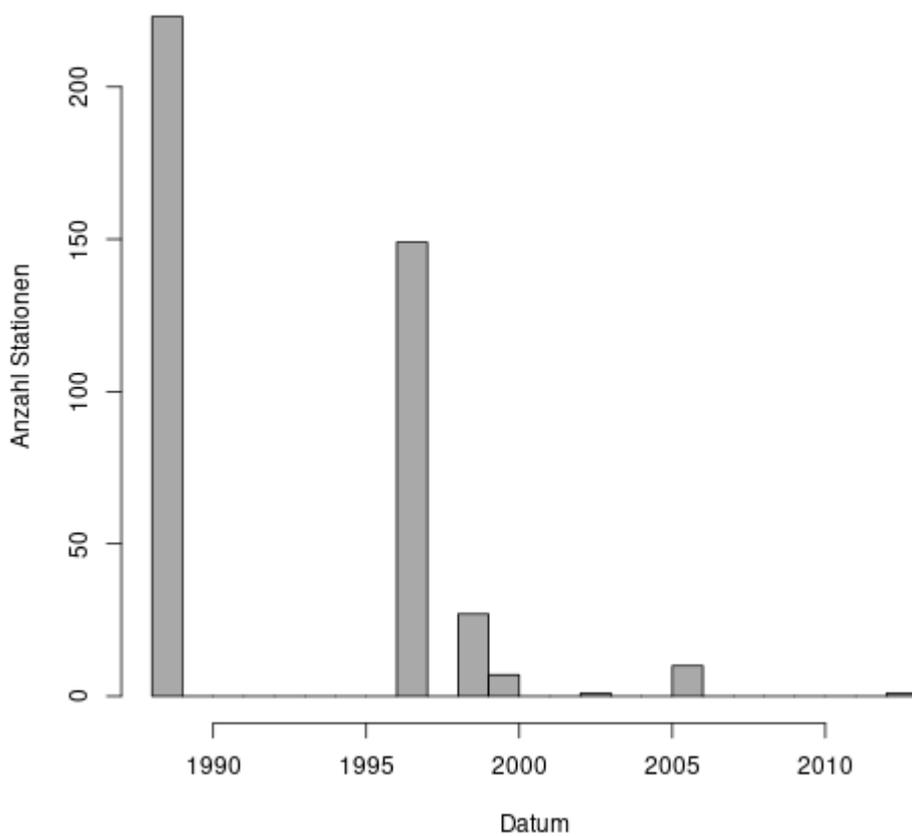
### 3.5 Prüfung auf Homogenität

Homogenitätstests für Zeitreihen dienen dazu, Bruchpunkte oder Trends in den untersuchten Daten zu identifizieren. Dabei wird zwischen absoluten und relativen Homogenitätstests unterschieden. Absolute Homogenitätstests betrachten ausschließlich die zu analysierende Zeitreihe. Relative Homogenitätstests setzen die untersuchte Zeitreihe in Bezug zu einer zweiten, bereits als homogen getesteten Referenzzeitreihe.

Da es jedoch keine zweifelsfrei homogenen Referenzzeitreihen für verschiedene Messgrößen gibt, wird in diesem Bericht ein absoluter Homogenitätstest eingesetzt: Der Pettitt-Test (PETTITT, 1979) gilt als sehr robust und erkennt Bruchpunkte in Zeitreihen zuverlässig.

Inhomogenitäten in Zeitreihen können sowohl interne als auch externe Ursachen haben. Interne Ursachen sind beispielsweise Standortwechsel, Gerätewechsel, Beobachterwechsel, Neukalibrierung von Geräten oder Veränderungen der Umgebung (z. B. Baumfällungen oder Gebäudebau).

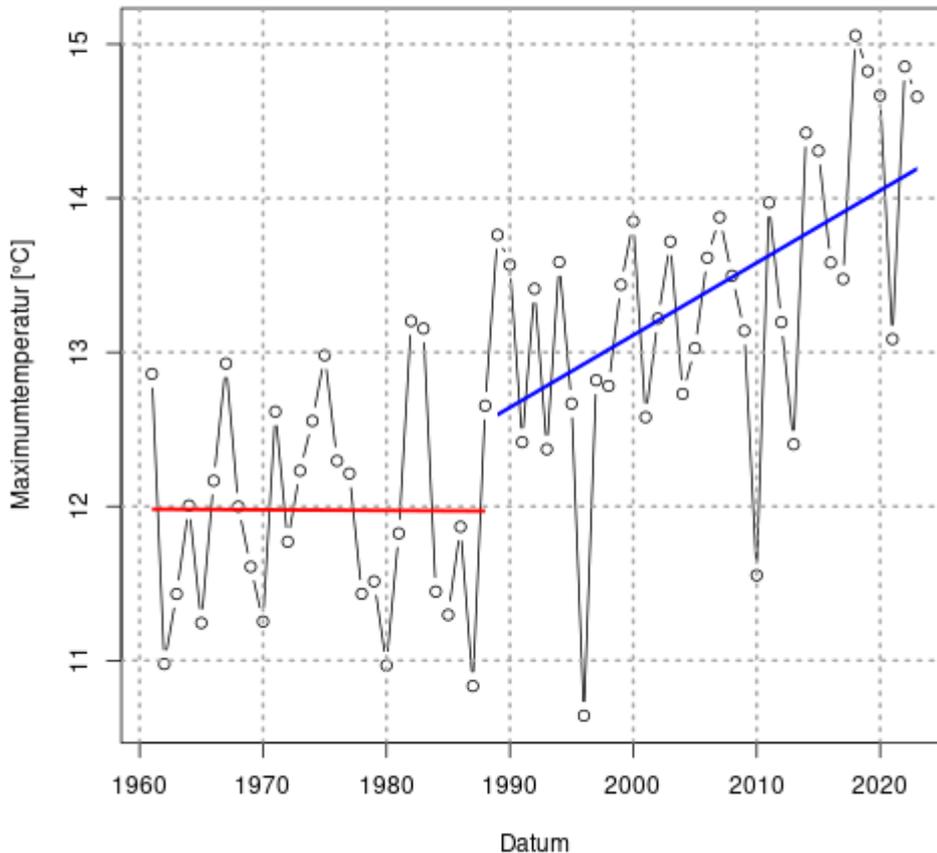
Externe Ursachen sind großräumige Einflussfaktoren, die gleichzeitig an mehreren Messstationen Veränderungen hervorrufen können. Darunter zählen auch Klimaänderungen. Im Referenzdatensatz 2.0 (KÖRNER ET AL, 2022) wurden diese Homogenitätstests bereits durchgeführt. Die mit diesem erweiterten Datensatz ermittelten Inhomogenitäten entsprechen weitgehend denjenigen des Vorgängerberichtes. Unterschiede gibt es durch die um drei Jahre längeren Zeitreihen: Hierdurch verschieben sich Bruchpunkte auf spätere Zeitpunkte. Als Beispiel sei die Maximum Temperatur genannt:



**Abbildung 3: Zeitliche Verteilung der Bruchpunkte für die Maximum Temperatur**

Im Vergleich zum Referenzdatensatz 2.0 (1961–2020) treten weniger Bruchpunkte im Jahr 1988, dafür mehr im Jahr 1997 auf. An der grundsätzlichen Aussage gibt es hingegen keine Änderung. So steigt die

Temperatur an den Bruchpunkten sprunghaft an und der Temperaturanstieg pro Jahr nimmt ebenfalls zu (Abbildung 4).



**Abbildung 4: Mittlerer zeitlicher Verlauf aller Stationen für die Maximum Temperatur. Rote Linie: Mittlerer Trend vor dem häufigsten Bruchpunkt. Blaue Linie: Mittlerer Trend nach dem häufigsten Bruchpunkt.**

### 3.6 Rasterdatensatz erstellen

Aus dem stationsbasierten Referenzdatensatz werden rasterbasierte Datensätze berechnet. Die Berechnung erfolgt mit den räumlichen Interpolationsverfahren „Thin Plate Spline Smoothing“ (Tps) und „Kriging“ (Krig). Die Ergebnisse werden als NetCDF-Datei in verschiedenen Projektionen zur Verfügung gestellt. Dieser Abschnitt beinhaltet ebenso wie die Lückenfüllung eine Einschätzung der Qualität der räumlichen Interpolation mittels Kreuzvalidierung.

#### 3.6.1 Methodik

Die Erstellung der Raster erfolgt mit zwei Methoden zur räumlichen Interpolation mittels der R-Bibliothek "fields". Für den Niederschlag (RR und RK) hat sich eine Kombination aus „Kriging“ und „Thin Plate Spline Smoothing“ als qualitativ am hochwertigsten herausgestellt. Zusätzlich erfolgt eine

Indikatorinterpolation. Bei dieser Methode werden pro Station 0 (kein Niederschlag) und 1 (Niederschlag) interpoliert. Bei Werten ab 0,5 wird der interpolierte Niederschlag verwendet. Werte unter 0,5 werden auf null mm gesetzt. Für alle anderen Größen werden die Raster mittels „Thin Plate Spline Smoothing“ interpoliert. Eine Ausnahme bilden hier die Verdunstungsgrößen ET und GR. Diese werden nicht aus den Stationswerten interpoliert. Vielmehr werden die meteorologischen Ausgangsgrößen, also FM, TM, RG, RN und DD interpoliert. Aus diesen Rastern werden anschließend nach den Formeln (14) bis (21) die Verdunstungswerte abgeleitet.

**Tabelle 5: Berechnete Rasterdatensätze für meteorologische und abgeleitete Größen**

Größe	Abkürzung	Einheit	Anzahl Stützstellen	Bemerkungen
Minimum Temperatur	TN	°C	416	
Maximum Temperatur	TX	°C	418	
Mittlere Temperatur	TM	°C	422	
Relative Feuchte	RF	%	414	
Unkorrigierter Niederschlag	RR	mm	2.462	Interpolation mit Tps, Kriging und Indikatorberechnung
Korrigierter Niederschlag	RK	mm	2.462	
Mittlere Windgeschwindigkeit	FM	m/s	221	
Maximale Windgeschwindigkeit	FX	m/s	178	
Dampfdruck	DD	hPa	413	
Sättigungsdampfdruck	ES	hPa	422	
Globalstrahlung	RG	W/m <sup>2</sup> und J/cm <sup>2</sup>	257	
Potentielle Verdunstung	ET	mm und W/m <sup>2</sup>	-	Verdunstung wird aus Rastern der jeweiligen Größen mit Formeln (14)-(21) berechnet
Grasreferenzverdunstung	GR	mm und W/m <sup>2</sup>	-	

### 3.6.2 Kreuzvalidierung

Zur Abschätzung der Qualität der räumlichen Interpolation wurde die folgende Validierungsstrategie verwendet: Für alle meteorologischen Messgrößen, welche auch interpoliert werden, wurde eine 10-fache Kreuzvalidierung angewendet. Das bedeutet, dass für jede Größe für jeden Zeitpunkt 10 % der Stationsdaten zufällig entfernt werden. Die Funktion zur räumlichen Interpolation wird anhand der Daten

der 90 % verbleibenden Stationen berechnet. Diese Funktion wird auf die 10 % entfernten Daten angewendet. Für jeden Zeitschritt wird dieses Verfahren 10-mal wiederholt, so dass letztendlich für jeden Messwert ein interpolierter Schätzwert zur Verfügung steht. Zur Beurteilung der Qualität werden statistische Kennwerte der gemessenen und interpolierten Zeitreihen bestimmt. In der folgenden Tabelle 6 wurde der Median je Größe von allen kreuzvalidierten Stationen berechnet und dargestellt.

**Tabelle 6: Kennwerte der Kreuzvalidierung zur Abschätzung der Qualität der räumlichen Interpolation**

Größe	RMSE	MAE	BIAS	R <sup>2</sup>
TN	1,16 °C	0,86 °C	0,01 °C	0,976
TX	0,79 °C	0,57 °C	0,01 °C	0,993
TM	0,73 °C	0,54 °C	0,01 °C	0,992
RF	4,50 %	3,53 %	-0,23 %	0,87
RR	1,76 mm/d	0,69 mm	0,045 mm	0,82
FM	1,21 m/s	0,92 m/s	0,05 m/s	0,67
FX	1,76 m/s	1,32 m/s	0,11 m/s	0,82
DD	0,48 hPa	0,34 hPa	-0,01 hPa	0,987

### 3.7 Monats- und Jahresdaten

Aus den Stations- und Rasterdaten der Tageswerte werden jeweils Daten auf Monats- und Jahresbasis aggregiert. Die jeweilige Methodik ist in diesem Abschnitt beschrieben.

#### 3.7.1 Stationsdaten

Der DWD stellt für die meteorologischen Größen unkorrigierter Niederschlag, Sonnenscheindauer sowie mittlere, minimale und maximale Temperatur neben Tagesdaten auch Zeitreihen in Monats- und Jahresauflösung bereit. Zur Erstellung dieser Zeitreihen werden die lückenlosen Tagesdaten zunächst auf Monats- und Jahreswerte aggregiert. Anschließend werden die aggregierten Daten mit den vom DWD bereitgestellten Monats- und Jahreswerten zusammengeführt. Dabei werden bei Abweichungen stets die originalen DWD-Monats- oder Jahreswerte bevorzugt. Sind keine DWD-Werte verfügbar, werden die aggregierten Tageswerte verwendet. Alle anderen Zeitreihen werden auf Jahres- und Monatsbasis aggregiert, ohne dass es zu einer zusätzlichen Datenzusammenführung kommt.

Eine Ausnahme bildet der korrigierte Niederschlag. Nach Richter (1995) werden die Monats- und Jahreswerte des unkorrigierten Niederschlages für beide zeitliche Auflösungen korrigiert.

**Tabelle 7: Prozentuale Korrektur des Niederschlages in Abhängigkeit der Stationslage und des Zeitraumes**

Zeitraum\Stationslage	frei	leicht geschützt	mäßig geschützt	stark geschützt
Januar	31,6	23,3	17,3	11,5
Februar	33,5	24,5	17,9	11,8
März	26,9	20,3	15,5	10,7
April	18,3	15,1	12,7	10
Mai	12,5	11,1	10,1	8,6
Juni	10,4	9,8	8,8	7,7
Juli	10,8	10	9,1	8
August	10,5	9,5	8,5	7,5
September	12,6	11,5	10,2	8,7
Oktober	15,5	12,7	11	8,8
November	21,8	16,8	13,3	9,5
Dezember	26,5	19,8	15	10,3
Jahr	18,2	14,6	12	9,3

Die Korrektur erfolgt prozentual anhand der in Tabelle 7 gezeigten Faktoren. Eine ungeschützte Station erhält demnach im Januar einen Aufschlag von 31,6 % auf den unkorrigierten Niederschlag, im gesamten Jahr einen Aufschlag von 18,2 %.

Insgesamt wird die Datenbasis durch die Verwendung der Monats- und Jahreswerte des DWD nur minimal erweitert: Für den Niederschlag kommen 15, die Sonnenscheindauer 19 und bei den Temperaturen jeweils eine Station dazu.

Bei auftretenden Lücken in den zusätzlichen Zeitreihen wird wie in Abschnitt 3.3 beschrieben verfahren, so dass auch für diese Zeitreihen je ein lückenloser Datensatz zur Verfügung steht.

### 3.7.2 Rasterdaten

Die Berechnung der Rasterdaten erfolgt analog der in Abschnitt 3.6 beschriebenen Methodik. Die Stationsdaten werden mittels „Thin Plate Splines“ und „Kriging“ räumlich interpoliert. Ausnahme bilden wieder die potentielle und die Grasreferenzverdunstung: Diese Größen werden durch die Aggregation der Tagesraster zu entsprechenden Monats- und Jahresrastern erstellt.

## 3.8 Datenbereitstellung

Die Stations- und Rasterdaten für alle zeitlichen Auflösungen und räumlichen Projektionen werden über das ReKIS Data Center (RDC) frei zur Verfügung gestellt.

## **4 Vergleich mit Referenzdatensatz 2.0 und der Reanalyse ERA5-land**

In diesem Abschnitt sollen die im Rahmen dieses Berichtes erstellten Daten mit dem Vorgängerprodukt, dem Referenzdatensatz 2.0, verglichen werden. Der Vergleich erfolgt für den gemeinsam abgedeckten Zeitraum 1961–2020 und das gemeinsam abgedeckte Gebiet des Mitteldeutschen Kernensembles in einer Auflösung von 1x1 km. Unterschiede beider Produkte sind in gewissem Rahmen zu erwarten, da sowohl bei der Datenverfügbarkeit als auch der Methodik Verbesserungen vorgenommen wurden. Durch die ständige Nachdigitalisierung älterer DWD-Stationen wurde die Datengrundlage im Untersuchungsgebiet erweitert. Durch die Erweiterung des Gebietes auf ganz Brandenburg wurde die Datengrundlage zusätzlich erweitert. Außerdem wurde durch die Verfügbarmachung der polnischen Daten ein bisher nicht ausreichend abgedecktes Gebiet in die Datenverarbeitung aufgenommen. Durch die Verbesserung der im Abschnitt 3.3 beschriebenen Lückenfüllung ergeben sich weitere potentielle Änderungen im Vergleich zum vorhergehenden Produkt. Während bei konservativen Größen wie Temperatur oder Feuchte insbesondere bei zeitlich gemittelten Werten keine größeren Abweichungen zu erwarten sind, können beim Niederschlag aufgrund der großen räumlichen Heterogenität oder bei den Windgeschwindigkeitsgrößen aufgrund geringer Stationsdichte einige zusätzliche Stationen zu einer erheblichen lokalen Abweichung führen.

Außerdem soll der hier erstellte Datensatz mit dem Reanalyseprodukt ERA5-land verglichen werden. Dieses weist eine hohe zeitliche (1950–2024) und räumliche (weltweit) Abdeckung auf, bei geringerer räumlicher Auflösung und größeren Abweichungen von den Messwerten.

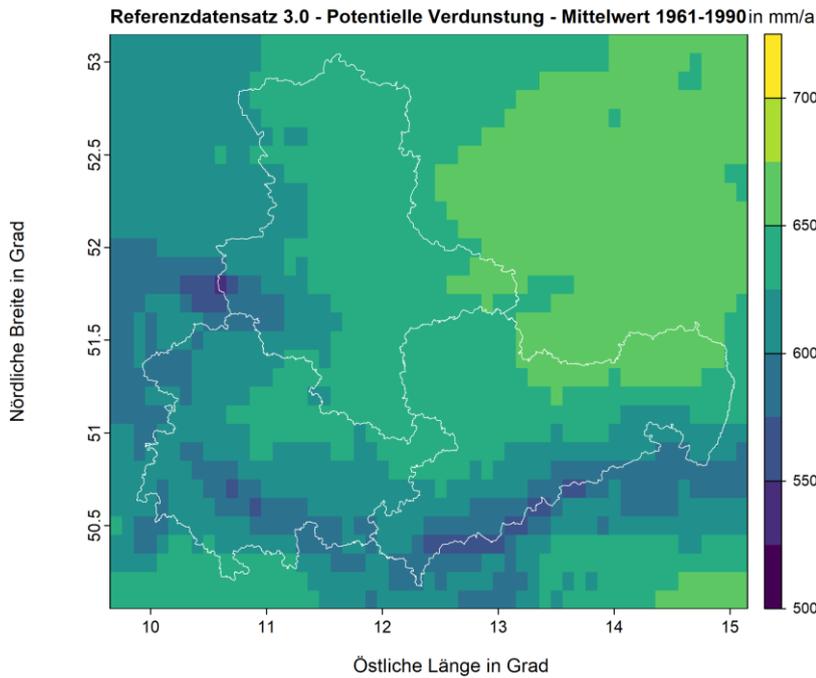
Hierfür werden aggregierte Tageswerte der jeweiligen Produkte von 1961–1990 miteinander verglichen. In Abschnitt 4.3 folgt eine Zusammenfassung der Vergleiche mittels statistischer Kennwerte für die beiden Klimanormalperioden 1961–1990 und 1991–2020.

### **4.1 Vergleich mit ERA5-land**

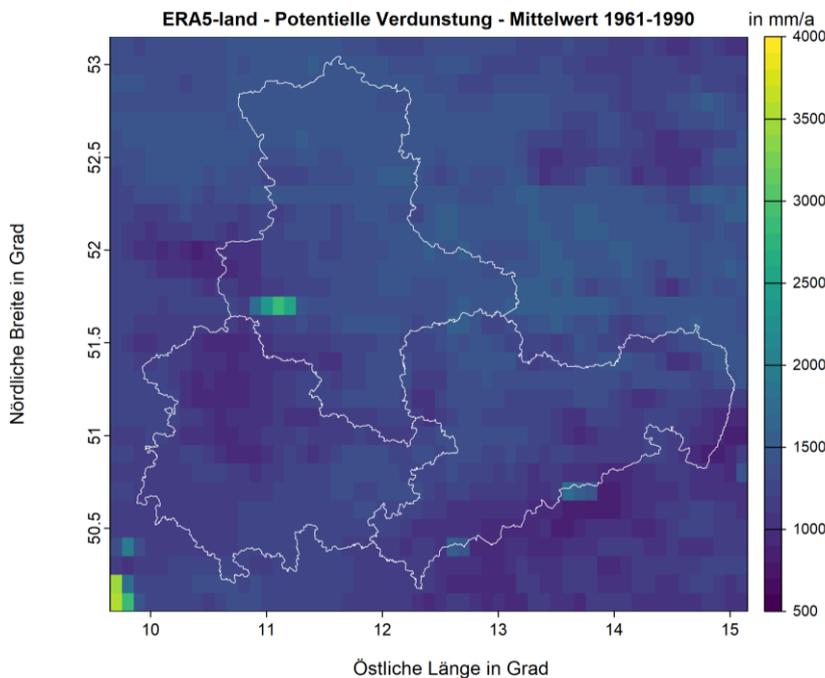
Im Referenzdatensatz 3.0 und dem Datensatz ERA5-land überschneiden sich fünf meteorologische Größen, die miteinander verglichen werden können. Das betrifft die potentielle Verdunstung (ET in mm), die mittlere Temperatur (TM in °C), die mittlere Windgeschwindigkeit (FM in m/s), die Globalstrahlung (RG in W/m<sup>2</sup>) und den Niederschlag (RK in mm). Die Auflösung beträgt etwa 0,1°, weswegen für den Vergleich der Referenzdatensatz 3.0 auf diese Auflösung aggregiert wurde. Der hier gewählte Vergleichszeitraum 1961–1990 ist repräsentativ für den Gesamtzeitraum. Eine Betrachtung des Zeitraumes 1991–2020 ändert an den jeweiligen Aussagen nichts.

### 4.1.1 Potentielle Verdunstung (ET)

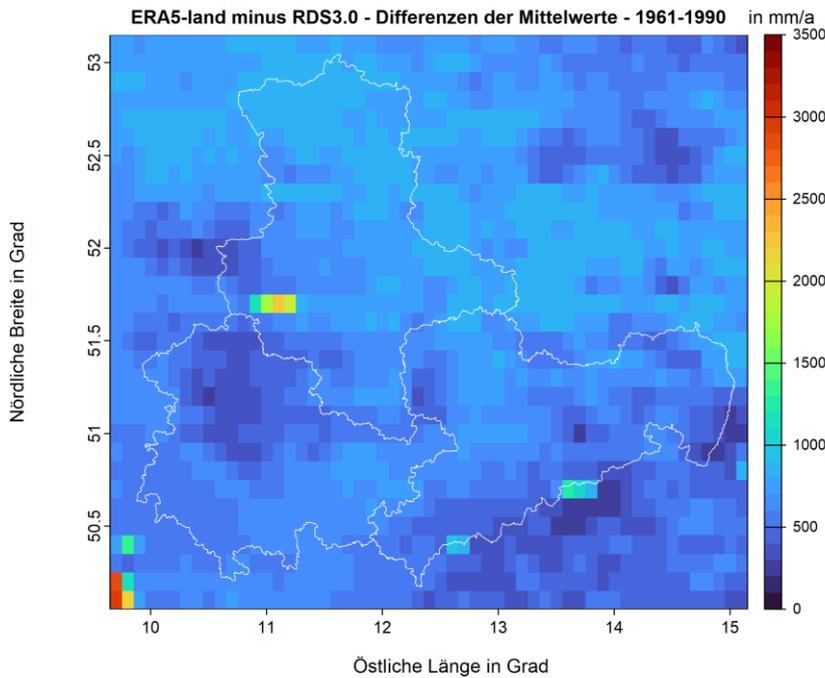
Der Vergleich zeigt die mittlere jährliche potentielle Verdunstung für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.



**Abbildung 5: Potentielle Verdunstung des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 6: Potentielle Verdunstung aus ERA5-land für 1961–1990**

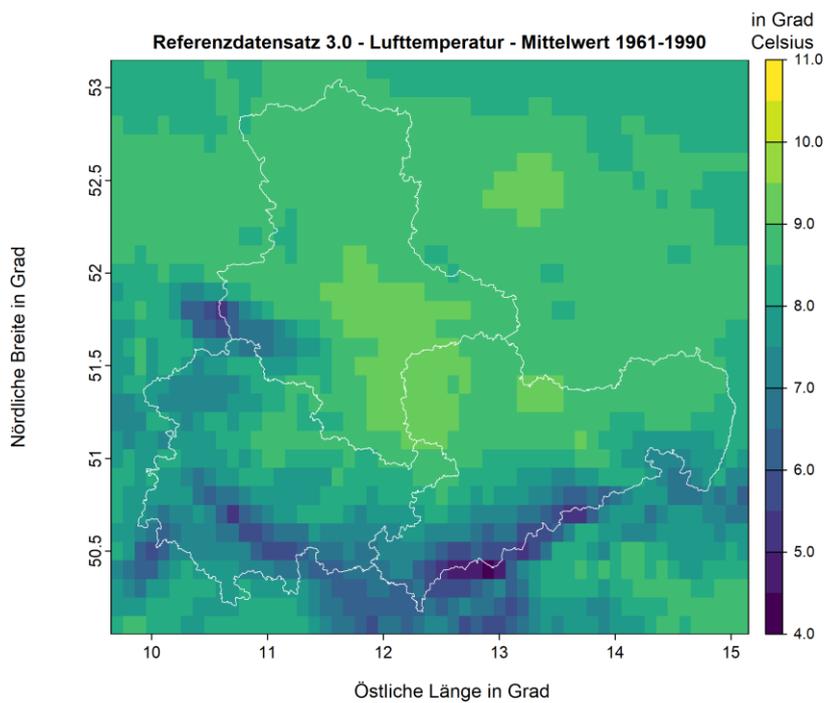


**Abbildung 7: Differenz der potentiellen Verdunstung aus ERA5-land und Referenzdatensatz 3.0 1961–1990**

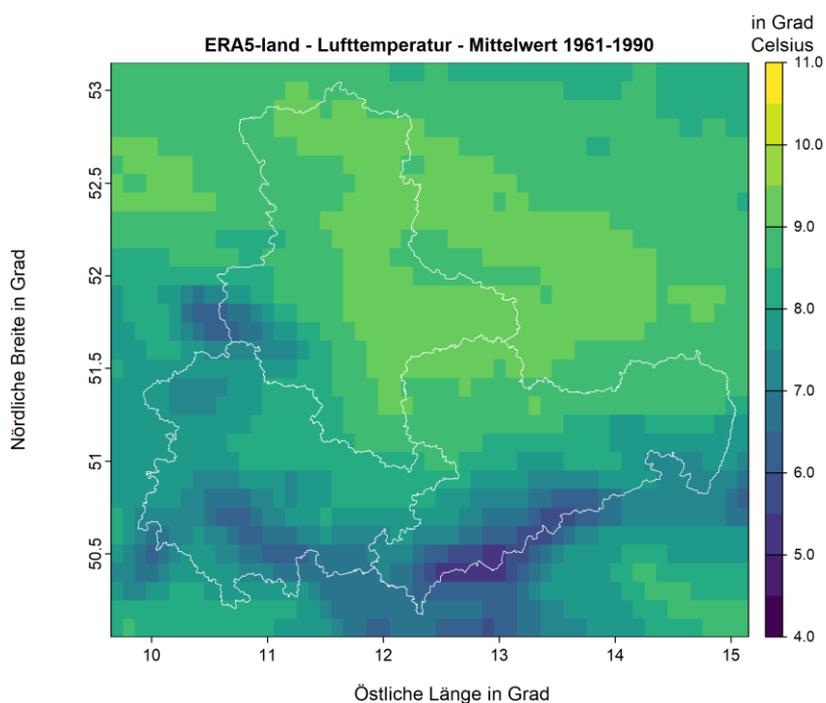
In Abbildung 5 ist die potentielle Verdunstung des Referenzdatensatzes dargestellt. Demgegenüber steht die potentielle Verdunstung aus dem ERA5-land-Produkt in Abbildung 6. Die Differenzendarstellung in Abbildung 7 zeigt, dass das ERA5-Land-Produkt weder in der Größenordnung noch in der räumlichen Verteilung Gemeinsamkeiten mit dem Referenzdatensatz aufweist. Die Verdunstungswerte sind unrealistisch hoch und übertreffen den Niederschlag um bis zu 300 %. Die fehlerhaften Werte des Produktes der potentiellen Verdunstung für ERA5-land treten weltweit auf (XU ET AL. 2024). Von einer Verwendung dieses Produktes ist daher abzuraten.

#### 4.1.2 Mittlere Temperatur

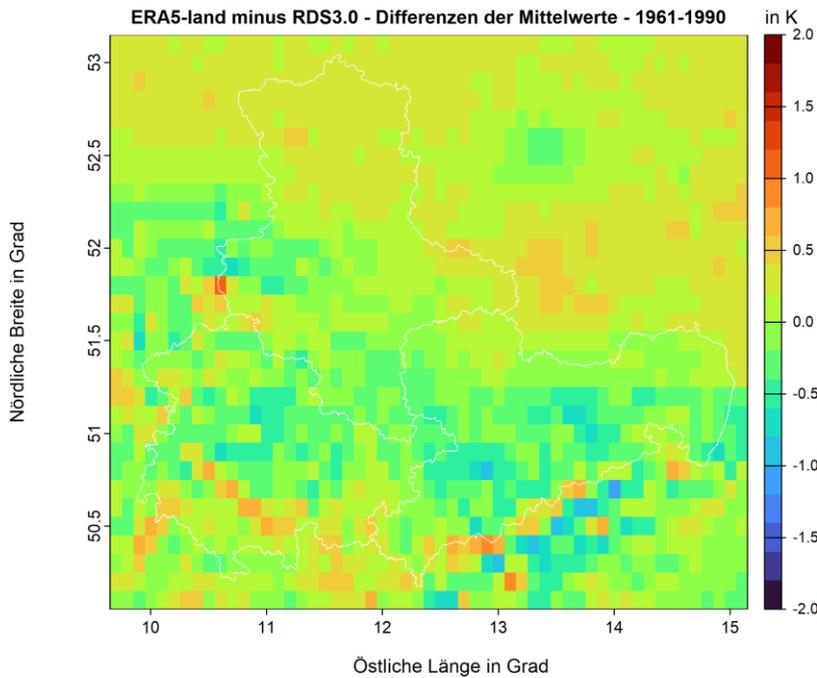
Der Vergleich zeigt die mittlere jährliche Temperatur für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.



**Abbildung 8: Mittlere Temperatur des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 9: Mittlere Temperatur aus ERA5-land für 1961–1990**

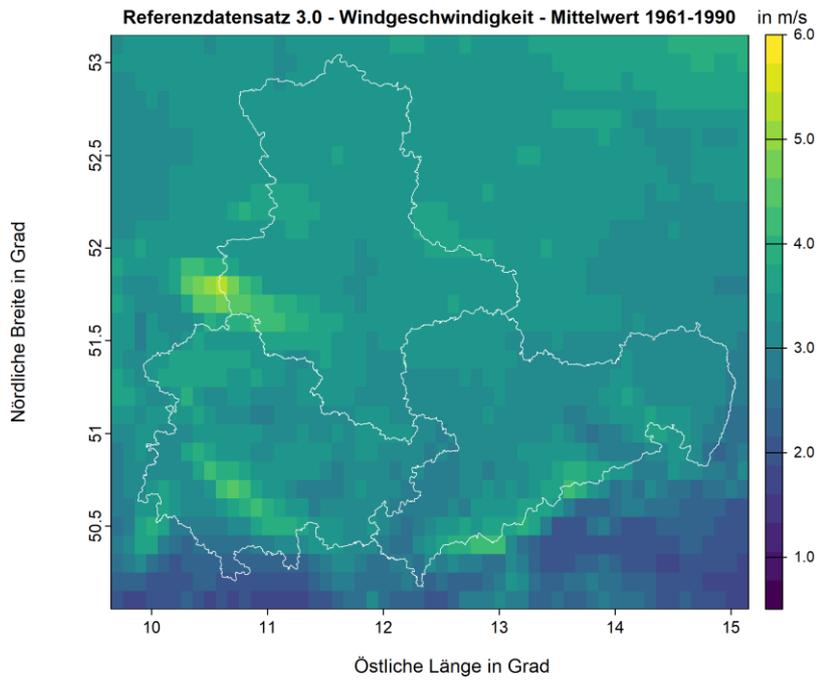


**Abbildung 10: Differenz der mittleren Temperatur aus ERA5-land und Referenzdatensatz 3.0 1961–1990**

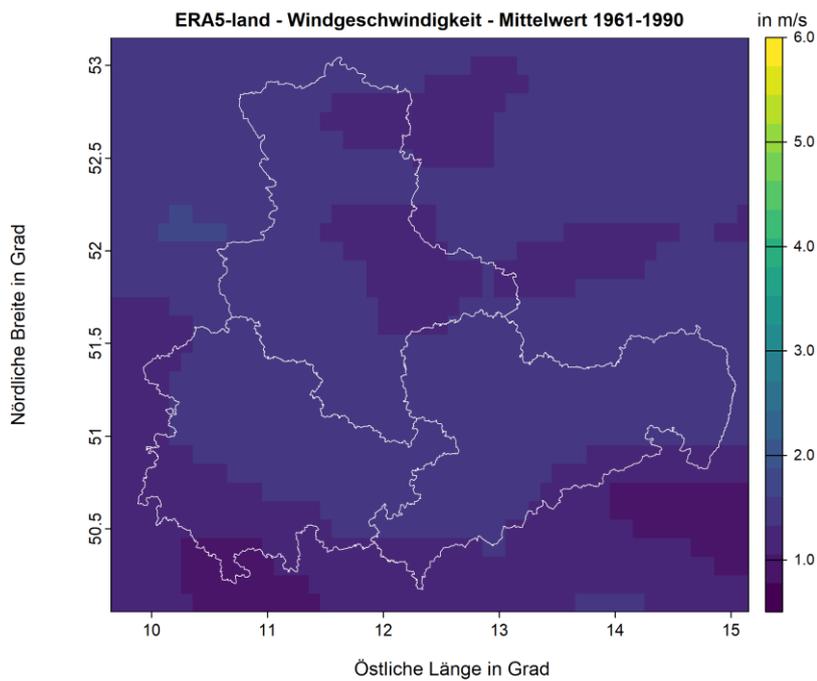
In Abbildung 8 ist die mittlere Temperatur des Referenzdatensatzes dargestellt. Demgegenüber steht die mittlere Temperatur aus dem ERA5-land-Produkt in Abbildung 9. Die Differenzendarstellung in Abbildung 10 zeigt, dass Abweichungen verhältnismäßig gering sind. Allerdings ist die Übereinstimmung der Temperatur in den Mittelgebirgen geringer, das ERA5-land-Produkt weist hier eine höhere Temperatur aus als der Referenzdatensatz 3.0.

#### 4.1.3 Mittlere Windgeschwindigkeit in 10 m Höhe

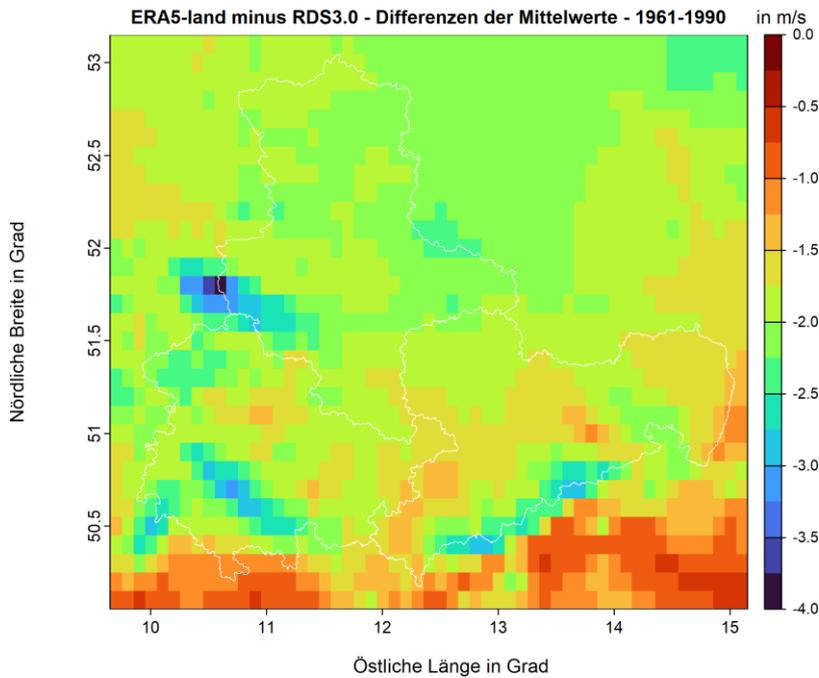
Der Vergleich zeigt die mittlere jährliche Windgeschwindigkeit für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.



**Abbildung 11: Mittlere Windgeschwindigkeit des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 12: Mittlere Windgeschwindigkeit aus ERA5-land für 1961–1990**

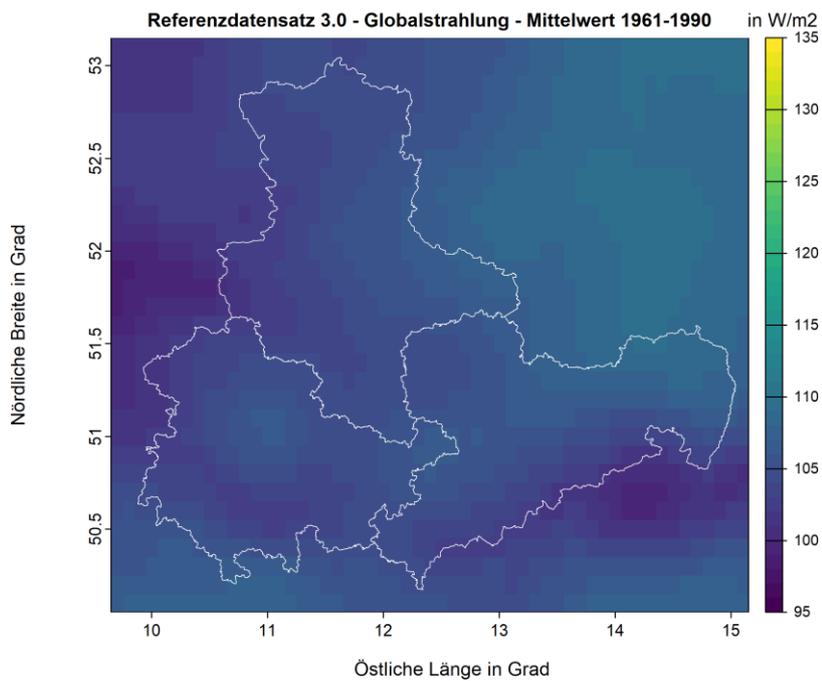


**Abbildung 13: Differenz der mittleren Windgeschwindigkeit aus ERA5-land und Referenzdatensatz 3.0 1961–1990**

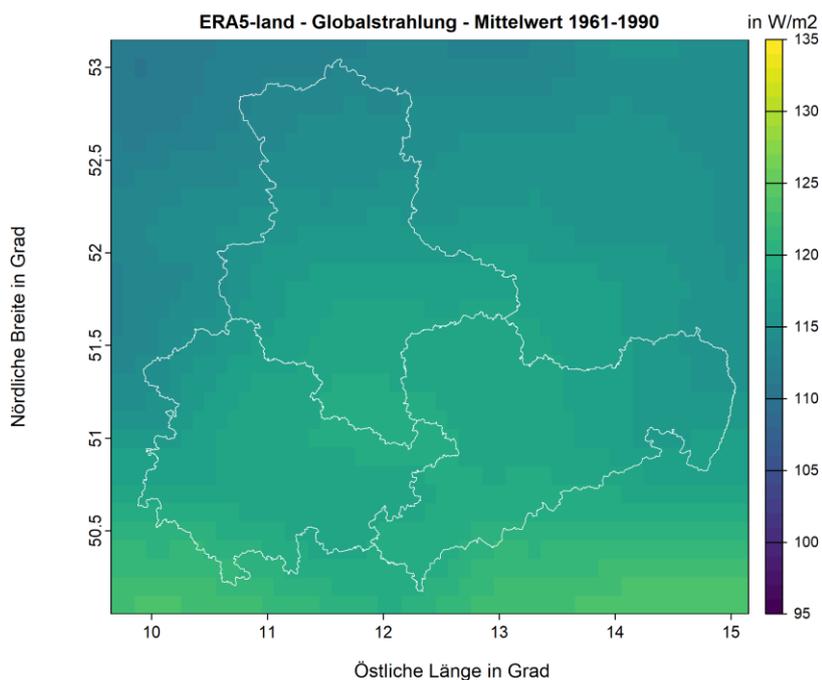
In Abbildung 11 ist die mittlere Windgeschwindigkeit des Referenzdatensatzes dargestellt. Demgegenüber steht die mittlere Windgeschwindigkeit aus dem ERA5-land-Produkt in Abbildung 12. Die Differenzdarstellung in Abbildung 13 zeigt, dass ähnlich wie bei der potentiellen Verdunstung zwischen beiden Produkten kein Zusammenhang besteht. Gleiches stellt auch KHADKA ET AL. (2022) fest. Die Windgeschwindigkeit des ERA5-land Produktes ist durchgehend zu homogen und deutlich zu niedrig.

#### 4.1.4 Globalstrahlung

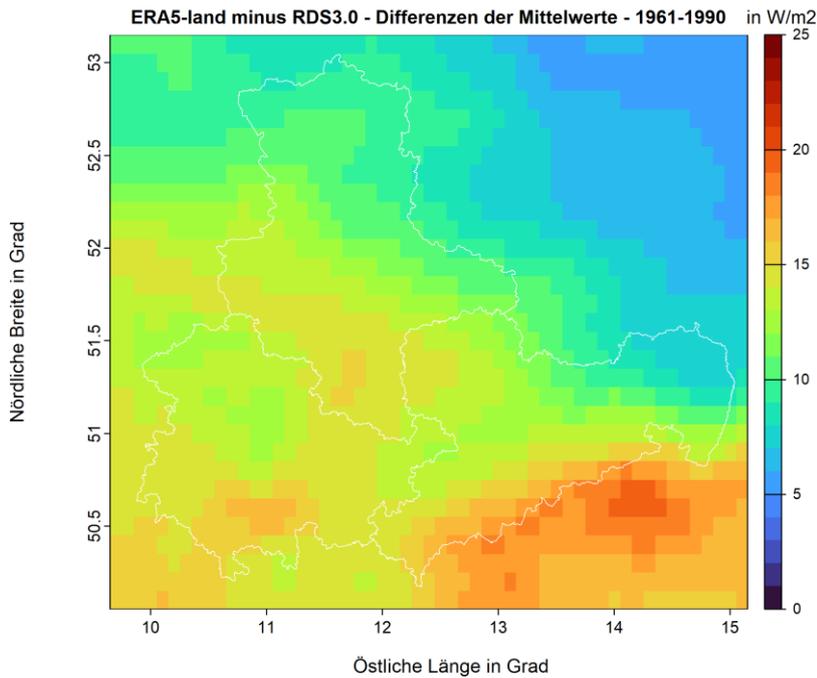
Der Vergleich zeigt die mittlere jährliche Globalstrahlung für den Bereich des Mitteldeutschen Kernens in der Klimareferenzperiode 1961–1990.



**Abbildung 14: Globalstrahlung des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 15: Globalstrahlung aus ERA5-land für 1961–1990**

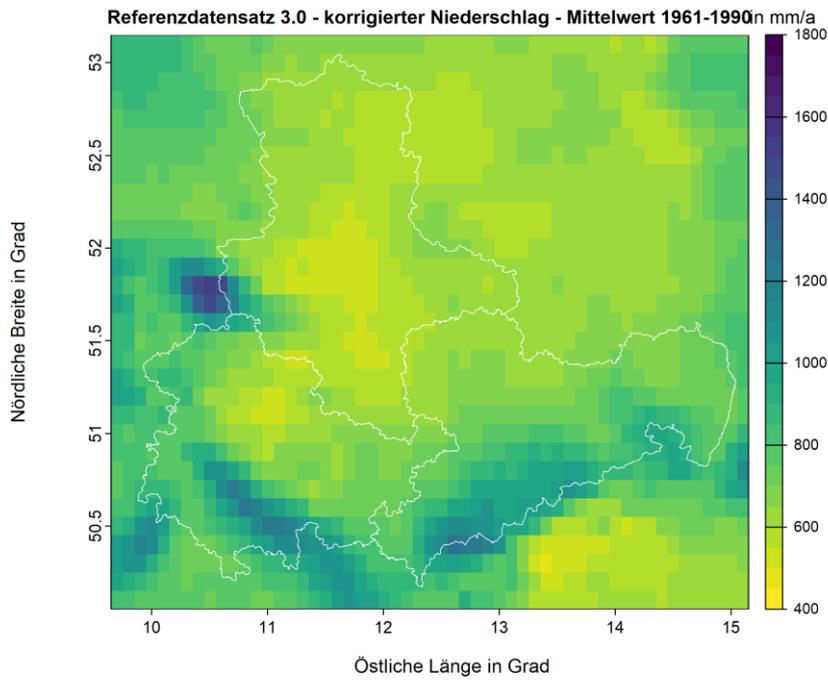


**Abbildung 16: Differenz der Globalstrahlung aus ERA5-land und Referenzdatensatz 3.0 1961–1990**

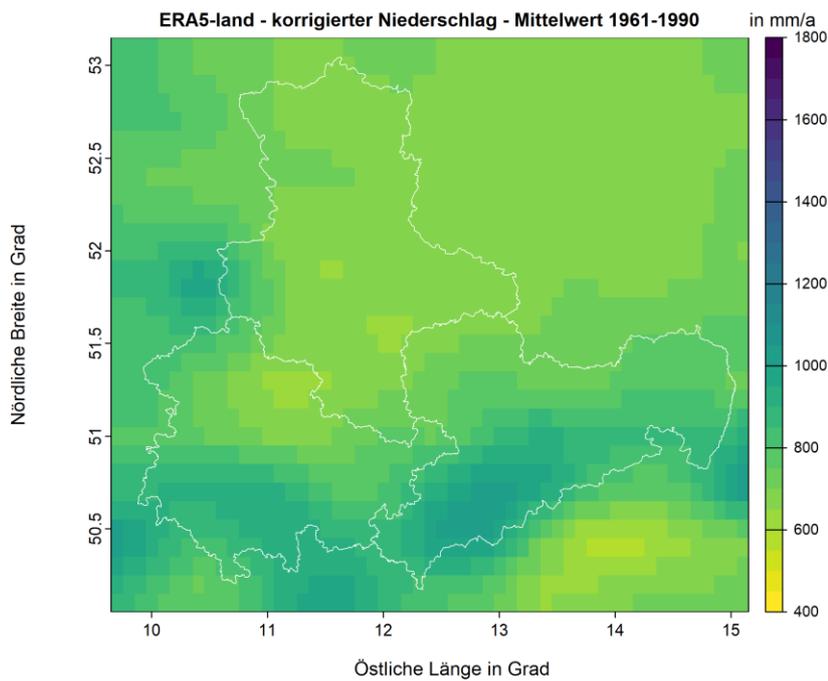
In Abbildung 14 ist die Globalstrahlung des Referenzdatensatzes dargestellt. Demgegenüber steht die Globalstrahlung aus dem ERA5-land-Produkt in Abbildung 15. Die Differenzendarstellung in Abbildung 16 zeigt, dass die Globalstrahlung im gesamten Vergleichsgebiet beim ERA5-land Produkt höher ist, als beim Referenzdatensatz 3.0. Die Literatur macht hier unterschiedliche Angaben zum Vergleich der Globalstrahlung mit gemessenen Werten (PELOSI ET AL., 2020, TIAN ET AL., 2024, GAO ET AL., 2023). Es gibt Gebiete mit höherer, ähnlicher und niedrigerer Globalstrahlung aus dem ERA5-land-Produkt im Vergleich zu Messungen. Eine endgültige Einordnung ist darum schwierig. Festzustellen ist jedoch, dass eine räumliche Korrelation zwischen beiden Produkten der Globalstrahlung praktisch nicht vorhanden ist.

#### 4.1.5 Korrigierter Niederschlag

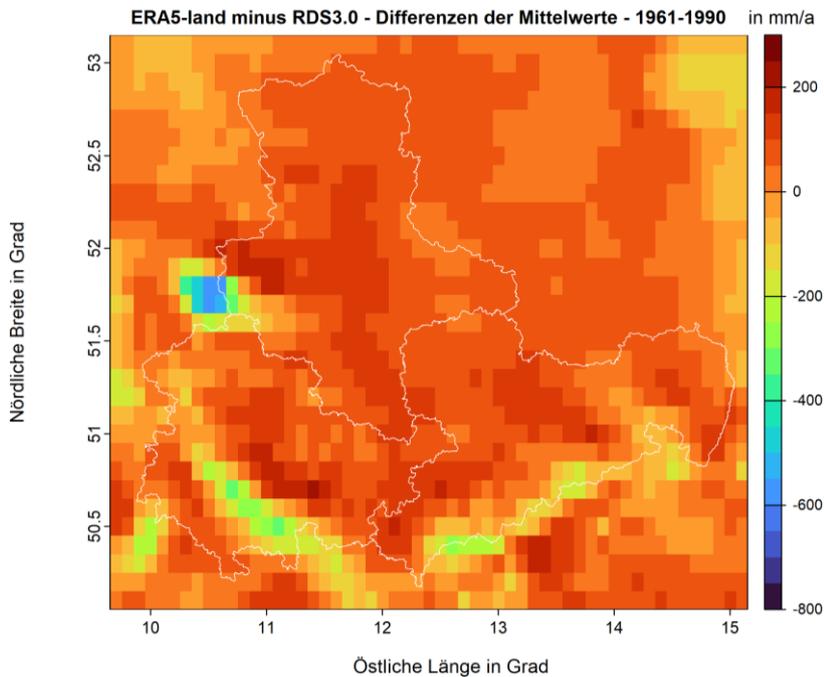
Der Vergleich zeigt die Summe des jährlichen Niederschlages für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990. Es wurde das Niederschlagsprodukt aus ERA5-land mit dem korrigierten Niederschlag des Referenzdatensatzes 3.0 verglichen, da davon ausgegangen wird, dass modellierter Niederschlag aus einem dynamischen Modell keinen Messfehlern unterliegt.



**Abbildung 17: Korrigierter Niederschlag des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 18: Niederschlag aus ERA5-land für 1961–1990**



**Abbildung 19: Differenz des Niederschlages aus ERA5-land und Referenzdatensatz 3.0 1961–1990**

In Abbildung 17 ist der korrigierte Niederschlag des Referenzdatensatzes dargestellt. Demgegenüber steht der Niederschlag aus dem ERA5-land-Produkt in Abbildung 18. Die Differenzendarstellung in Abbildung 19 zeigt, dass der Niederschlag räumlich insgesamt leicht höher ist als der korrigierte Niederschlag im Referenzdatensatz 3.0. In den Mittelgebirgen, insbesondere im Harz wird der Niederschlag vom ERA-land-Produkt jedoch stark unterschätzt.

Zusammenfassend ist festzustellen, dass die Produkte aus ERA5-land teilweise extreme, physikalisch nicht zu erklärende Abweichungen gegenüber Messungen und den Daten des Referenzdatensatzes 3.0 aufweisen. Am geringsten sind die Abweichungen bei der Temperatur und beim Niederschlag, wobei jeweils die Daten in den Gebieten der Mittelgebirge am größten voneinander abweichen. Bei der Strahlung sind systematisch größere Werte festzustellen, bei der Windgeschwindigkeit und der Verdunstung liegt offenbar ein Fehler im Produkt vor, so dass diese Daten aus ERA5-land nicht verwendet werden sollten.

## 4.2 Vergleich mit Referenzdatensatz 2.0

Im Referenzdatensatz 3.0 und dem Referenzdatensatz 2.0 überschneiden sich elf meteorologische Größen, die miteinander verglichen werden können. Das betrifft die Verdunstung (ET und GR in mm), die Temperatur (TM, TN und TX in °C), die mittlere Windgeschwindigkeit (FM in m/s), die Globalstrahlung (RG in W/m<sup>2</sup>), den Niederschlag (RK und RR in mm), den Dampfdruck (DD in hPa) und die relative Feuchte (RF in %). Die Auflösung beträgt jeweils 1 km, der Raumausschnitt umfasst das Gebiet des Mitteldeutschen Kernensembles. Der hier gewählte Vergleichszeitraum 1961–1990 ist repräsentativ für

den Gesamtzeitraum. Eine Betrachtung des Zeitraumes 1991–2020 ändert an den jeweiligen Aussagen nichts.

#### 4.2.1 Potentielle Verdunstung

Der Vergleich zeigt die Summe der jährlichen potentiellen Verdunstung (ET in mm) für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.

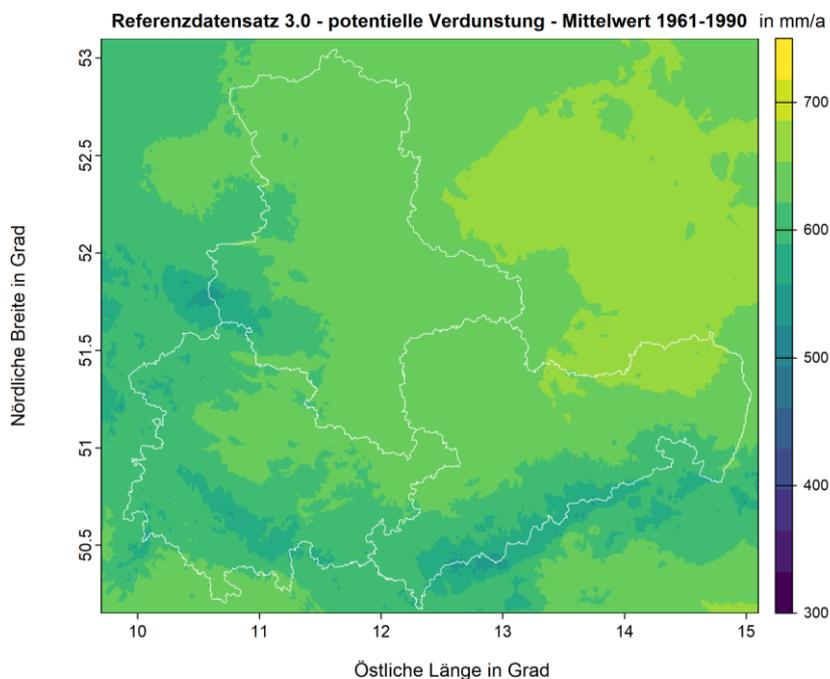


Abbildung 20: Potentielle Verdunstung des Referenzdatensatzes 3.0 für 1961–1990

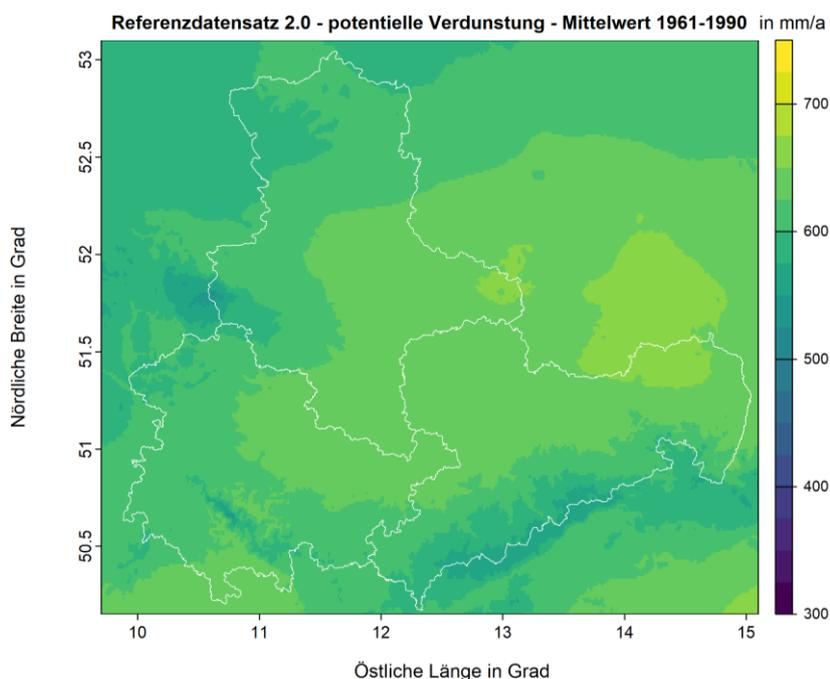
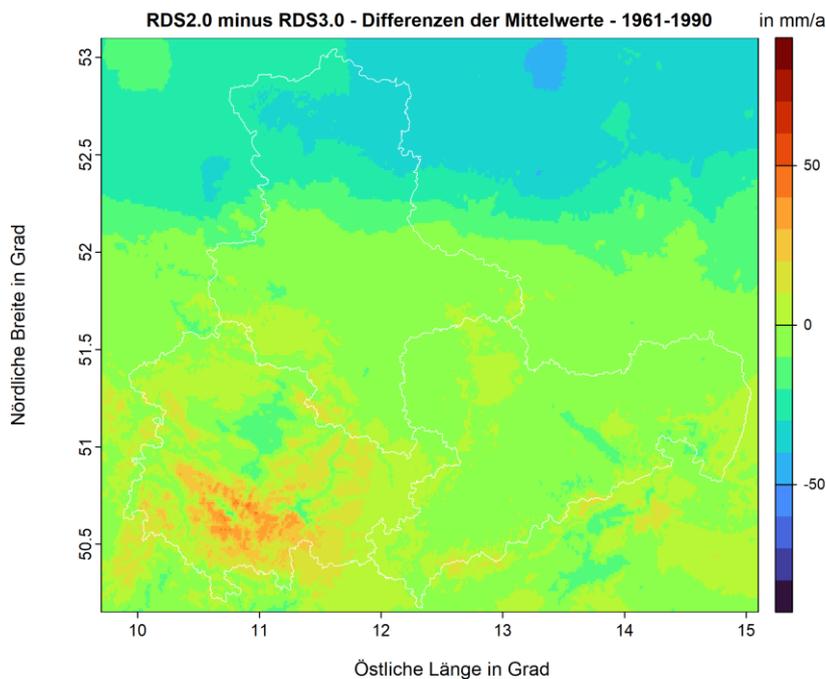


Abbildung 21: Potentielle Verdunstung des Referenzdatensatzes 2.0 für 1961–1990

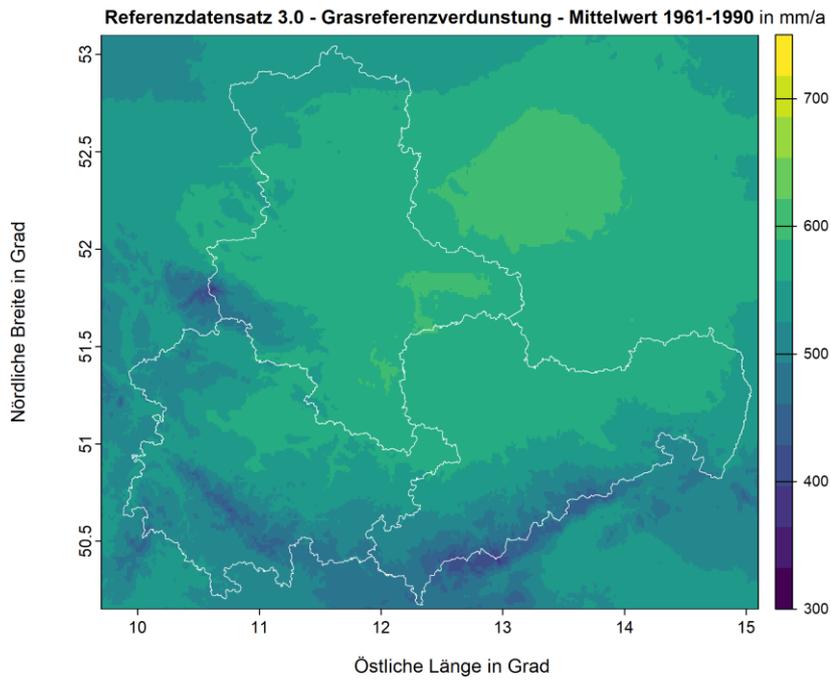


**Abbildung 22: Differenz der potentiellen Verdunstung aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961–1990**

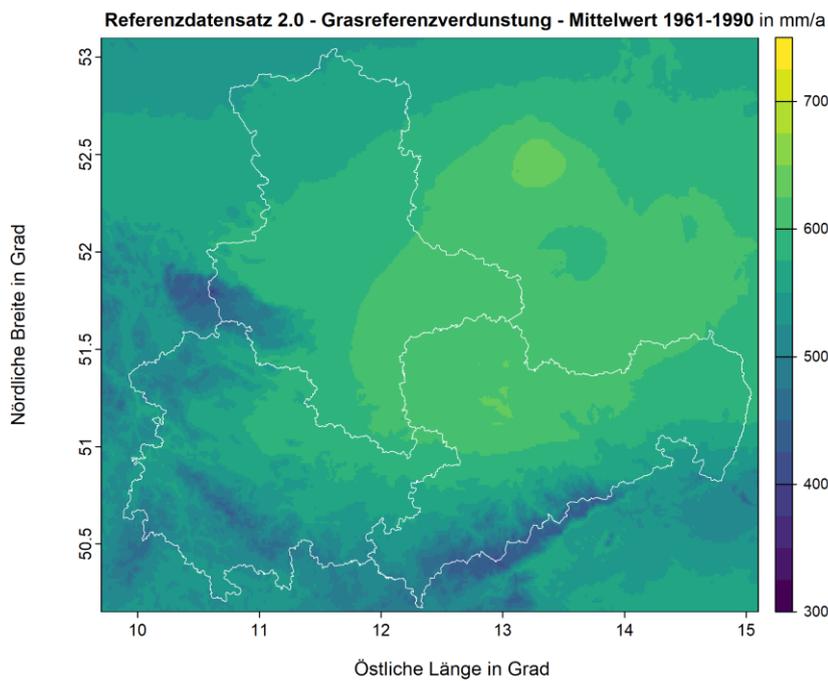
In Abbildung 20 ist die Summe der jährlichen potentiellen Verdunstung des Referenzdatensatzes 3.0 dargestellt. Demgegenüber steht die Summe der jährlichen potentiellen Verdunstung aus dem Referenzdatensatz 2.0 in Abbildung 21. Die Differenzendarstellung in Abbildung 22 zeigt, dass die Abweichungen insgesamt sehr gering sind. Bei den Gebieten mit den höchsten Abweichungen um Berlin, sowie im Thüringer Wald, ist die Stationsdichte gestiegen: Im Gebiet des MDK sind beispielsweise neun zusätzliche Stationen mit Messungen der Sonnenscheindauer verarbeitet worden. Sechs allein davon befinden sich im Gebiet um den Thüringer Wald. Etwa 20 zusätzliche Stationen sind im Norden des Gebietes durch die Erweiterung des Untersuchungsgebietes hinzugekommen.

#### 4.2.2 Grasreferenzverdunstung

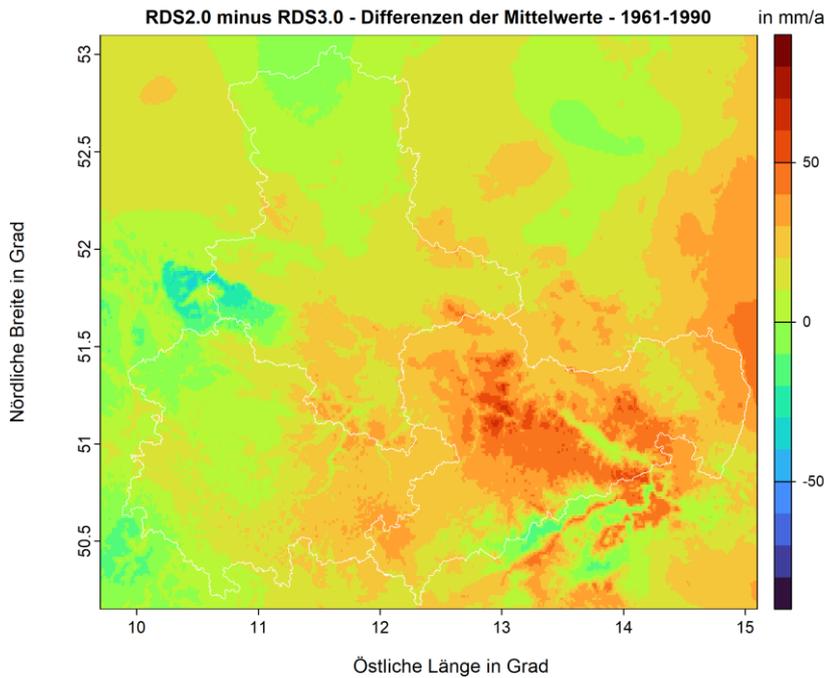
Der Vergleich zeigt die Summe der jährlichen Grasreferenzverdunstung (GR in mm) für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.



**Abbildung 23: Grasreferenzverdunstung des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 24: Grasreferenzverdunstung des Referenzdatensatzes 2.0 für 1961–1990**

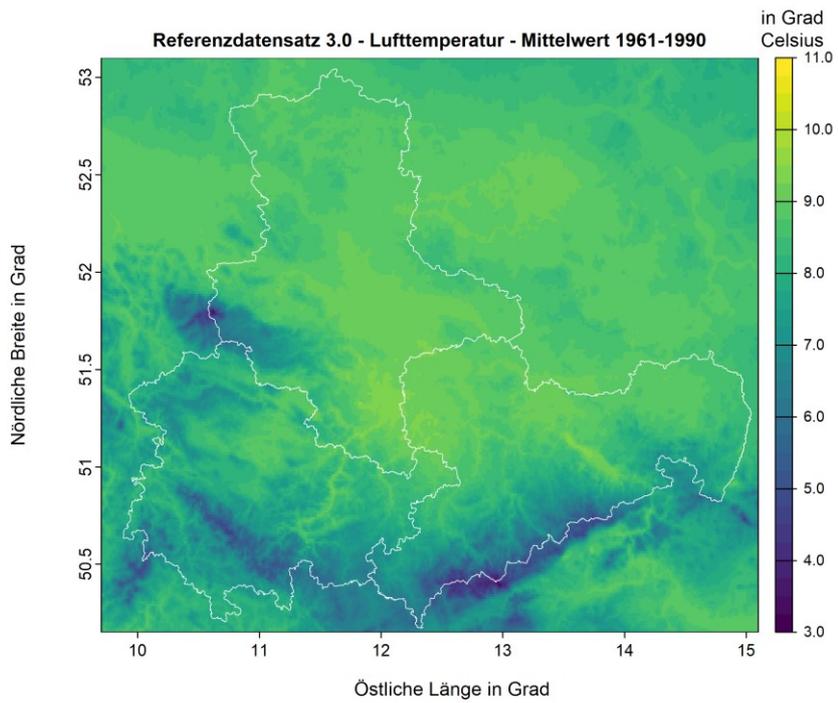


**Abbildung 25: Differenz der Grasreferenzverdunstung aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961–1990**

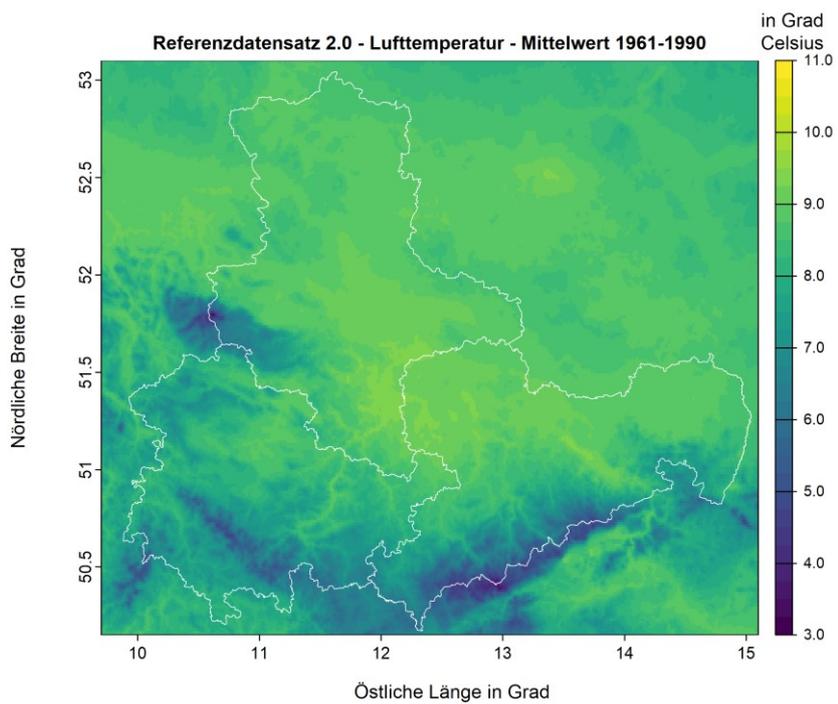
In Abbildung 23 ist die jährliche Summe der Grasreferenzverdunstung des Referenzdatensatzes 3.0 dargestellt. Demgegenüber steht die jährliche Summe der Grasreferenzverdunstung aus dem Referenzdatensatz 2.0 in Abbildung 24. Die Differenzendarstellung in Abbildung 25 zeigt, dass die mittlere Abweichung relativ gering ist, lokal, vor allem im Osten Sachsens aber Größenordnungen von bis zu 10 % erreichen können. Die Ursachen hierfür sind multikausal: Zunächst wird eine im Vergleich zum Referenzdatensatz 2.0 angepasste Gleichung nach DVWK-M-504-2 verwendet (18). Die Grasreferenzverdunstung hängt somit neben der Strahlung auch von der Feuchte, der Temperatur, dem Sättigungsdampfdruck und der Windgeschwindigkeit ab. Bei jeder dieser Größen gibt es kleinere bis mittlere Abweichungen des Referenzdatensatzes 3.0 zum Referenzdatensatz 2.0, wie die folgenden Abschnitte zeigen werden.

#### 4.2.3 Mittlere Temperatur

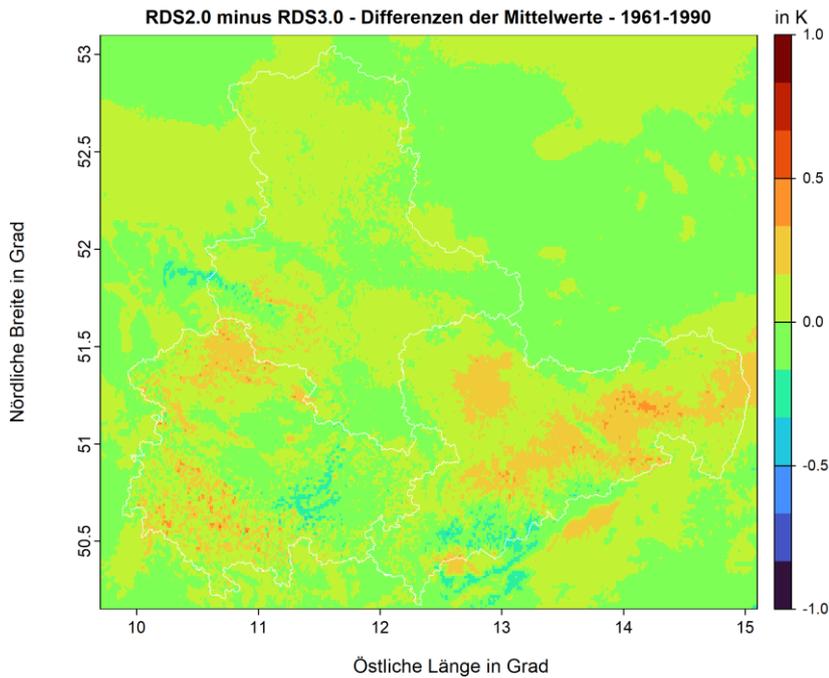
Der Vergleich zeigt die mittlere Temperatur (TM in °C) für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.



**Abbildung 26: Mittlere Temperatur des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 27: Mittlere Temperatur des Referenzdatensatzes 2.0 für 1961–1990**

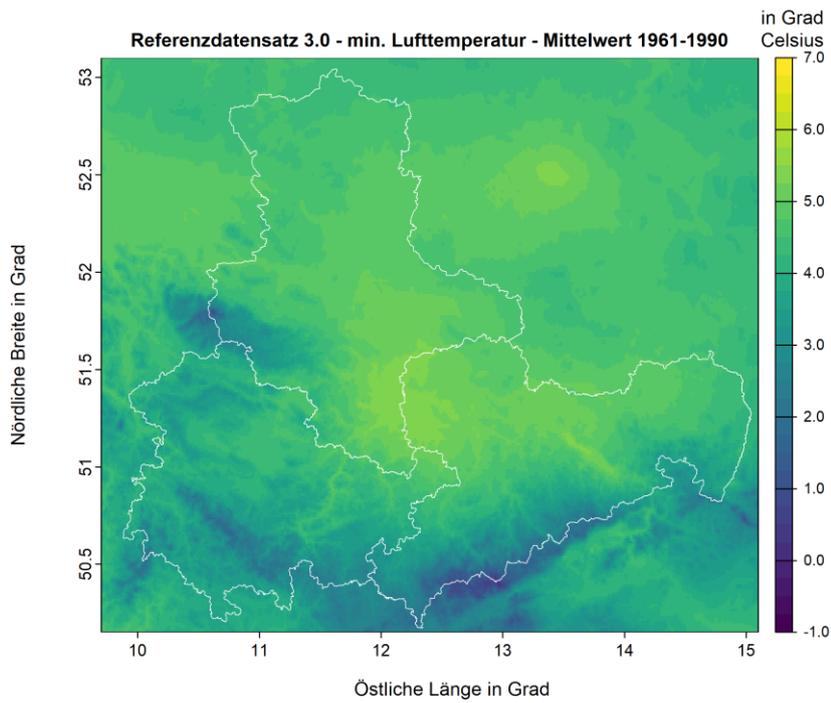


**Abbildung 28: Differenz der mittleren Temperatur aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961–1990**

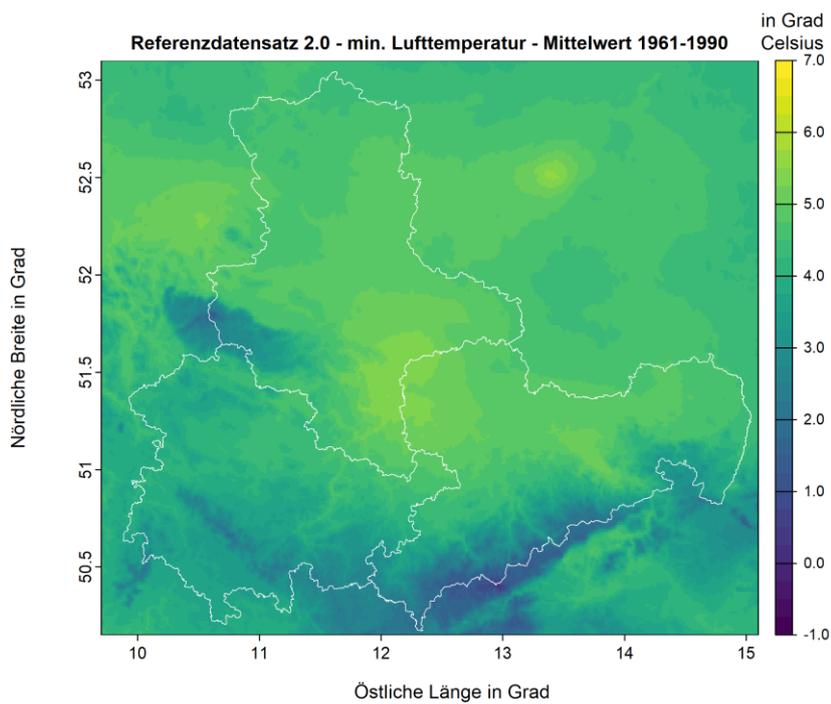
In Abbildung 26 ist die mittlere Temperatur des Referenzdatensatzes 3.0 dargestellt. Demgegenüber steht die mittlere Temperatur aus dem Referenzdatensatz 2.0 in Abbildung 27. Die Differenzendarstellung in Abbildung 28 zeigt, dass die Differenzen aus beiden Produkten sehr gering sind. Maximale Unterschiede von bis zu 0.4 Kelvin beschränken sich auf einzelne Pixel. Ursache hierfür sind voneinander unabhängige Dinge: Zum einen gibt es räumlich verteilt neue Stationen, die zusätzlich berücksichtigt wurden. Zum anderen weist die Temperatur eine starke Höhenabhängigkeit auf und nimmt mit zunehmender Höhe im Mittel ab. Für den Referenzdatensatz 2.0 wurde ein digitales Geländemodell (DGM) verwendet, welches die Geländehöhe bundesweit einheitlich erfasst. Dieses DGM wird auch im Interpolationstool RaKliDa verwendet. Für den hier erstellten Referenzdatensatz 3.0 wurde ein von den Landesämtern angepasstes DGM verwendet. Dieses beinhaltet zusätzlich die Informationen höher aufgelöster Landes-DGM für Sachsen, Sachsen-Anhalt und Thüringen. Im Mittel unterscheiden sich beide DGM um etwa einen Meter. Für einzelne Rasterzellen können die Unterschiede jedoch bis zu 100 m betragen. Aufgrund dieses Unterschiedes kann es bei höhenabhängigen Größen wie der Temperatur zu größeren lokalen Abweichungen kommen.

#### 4.2.4 Minimum Temperatur

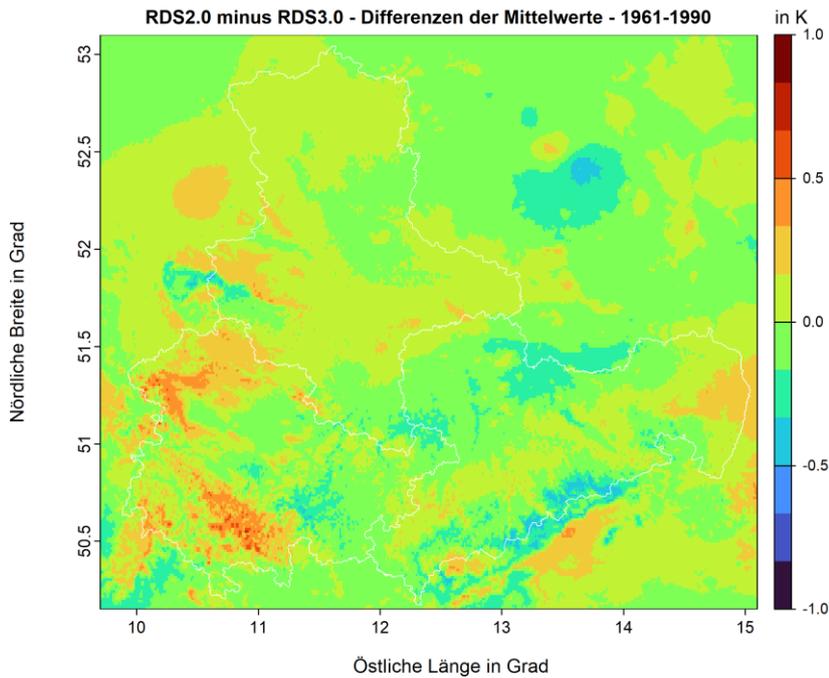
Der Vergleich zeigt die mittlere tägliche Minimum Temperatur (TN in °C) für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.



**Abbildung 29: Minimum Temperatur des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 30: Minimum Temperatur des Referenzdatensatzes 2.0 für 1961–1990**

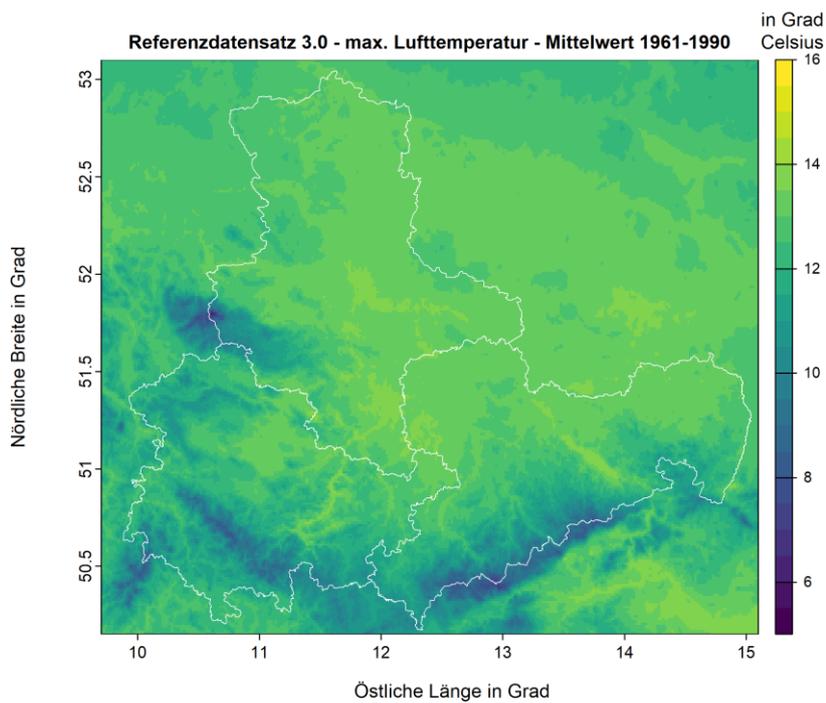


**Abbildung 31: Differenz der Minimum Temperatur aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961–1990**

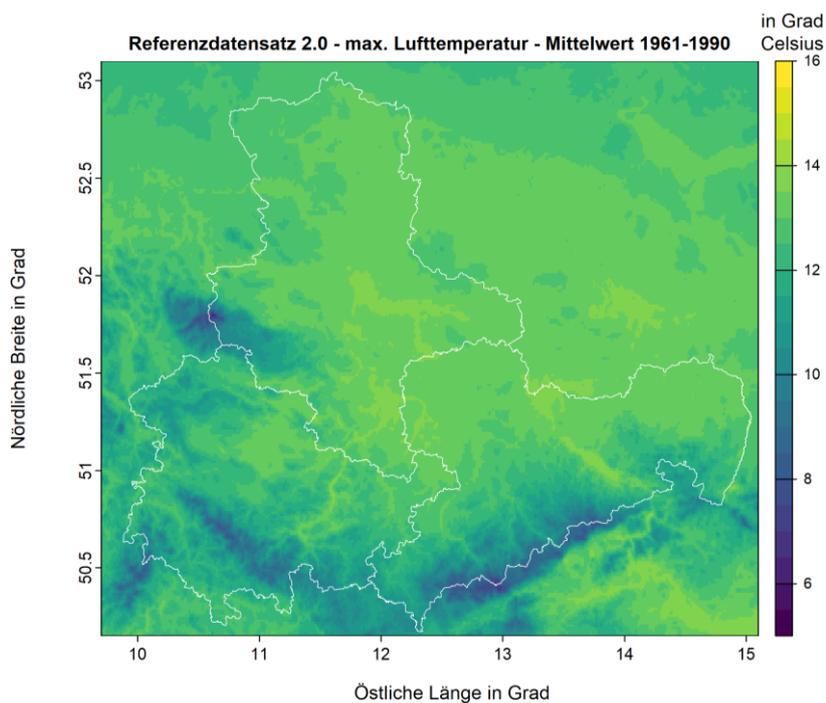
In Abbildung 29 ist das Mittel der täglichen Minimum Temperaturen des Referenzdatensatzes 3.0 dargestellt. Demgegenüber steht das Mittel der täglichen Minimum Temperaturen aus dem Referenzdatensatz 2.0 in Abbildung 30. Die Differenzendarstellung in Abbildung 31 zeigt, dass die Abweichungen im Mittel wieder gering sind, jedoch maximale Werte bis zu 0,5 Kelvin erreichen können. Grund hierfür ist wieder das aktuellere DGM und zusätzlich verwendete Stationen.

#### 4.2.5 Maximum Temperatur

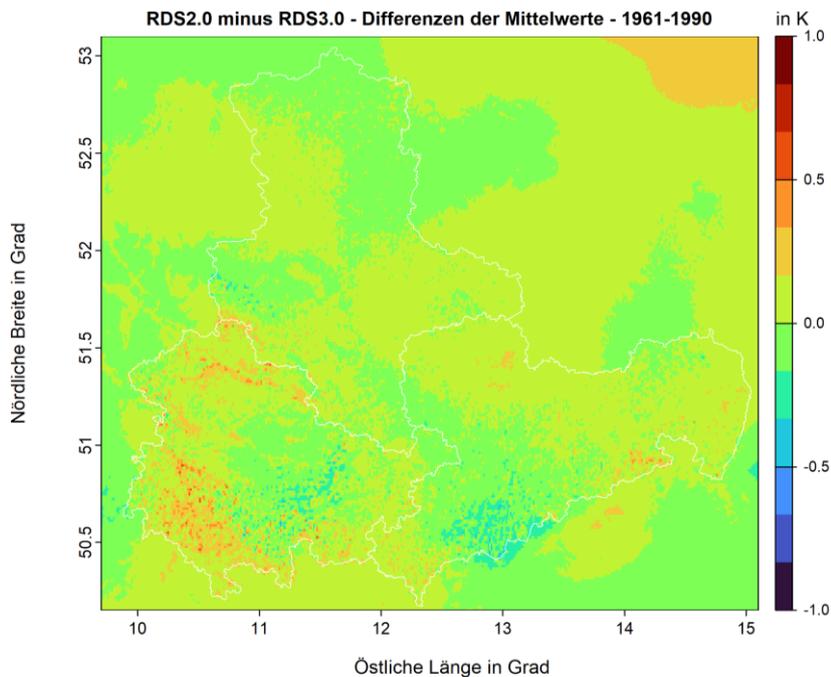
Der Vergleich zeigt die mittlere tägliche Maximum Temperatur (TX in °C) für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.



**Abbildung 32: Maximum Temperatur des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 33: Maximum Temperatur des Referenzdatensatzes 2.0 für 1961–1990**

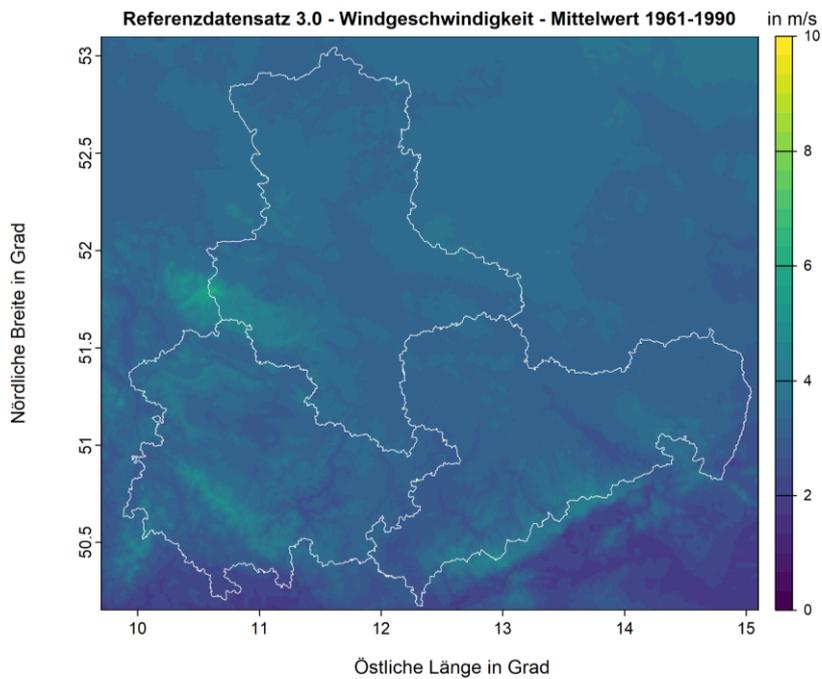


**Abbildung 34: Differenz der Maximum Temperatur aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961–1990**

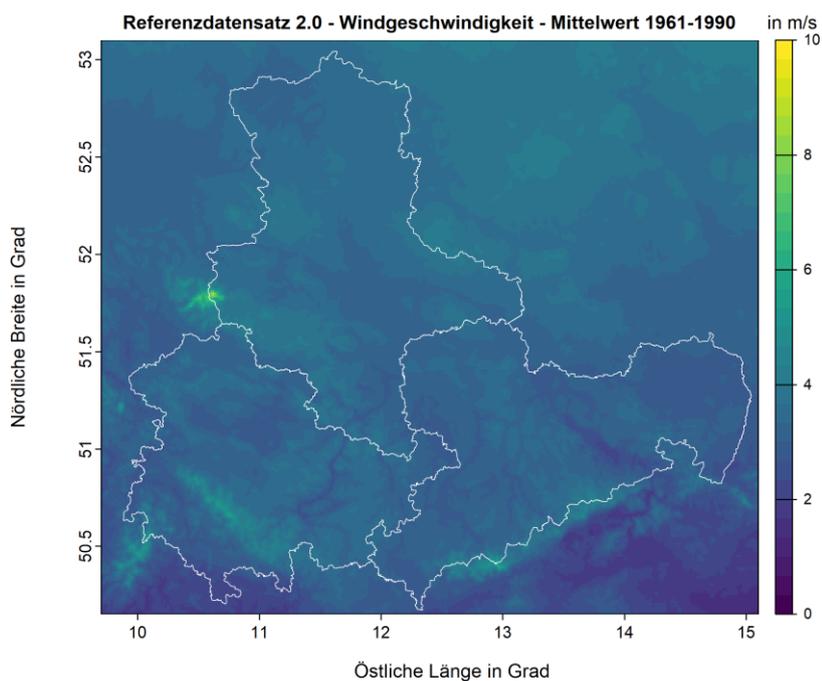
In Abbildung 32 ist das Mittel der täglichen Maximum Temperatur des Referenzdatensatzes 3.0 dargestellt. Demgegenüber steht das Mittel der täglichen Maximum Temperaturen aus dem Referenzdatensatz 2.0 in Abbildung 33. Die Differenzendarstellung in Abbildung 34 zeigt, dass die Abweichungen im Mittel wieder gering sind, jedoch maximale Werte bis zu 0,5 Kelvin erreichen können. Grund hierfür ist wieder das aktuellere DGM und zusätzlich verwendete Stationen.

#### 4.2.6 Mittlere Windgeschwindigkeit

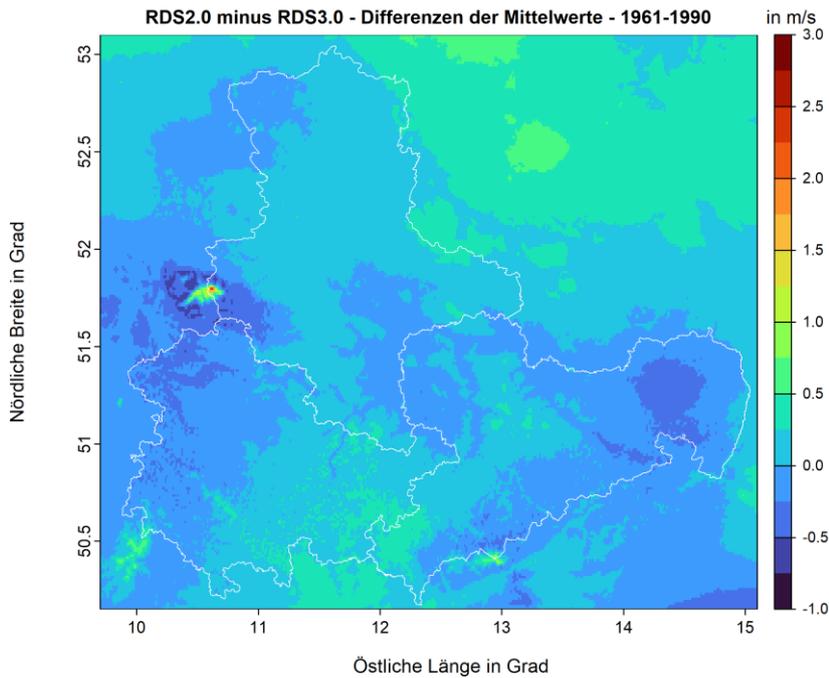
Der Vergleich zeigt die mittlere Windgeschwindigkeit (FM in m/s) für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.



**Abbildung 35: Mittlere Windgeschwindigkeit des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 36: Mittlere Windgeschwindigkeit des Referenzdatensatzes 2.0 für 1961–1990**

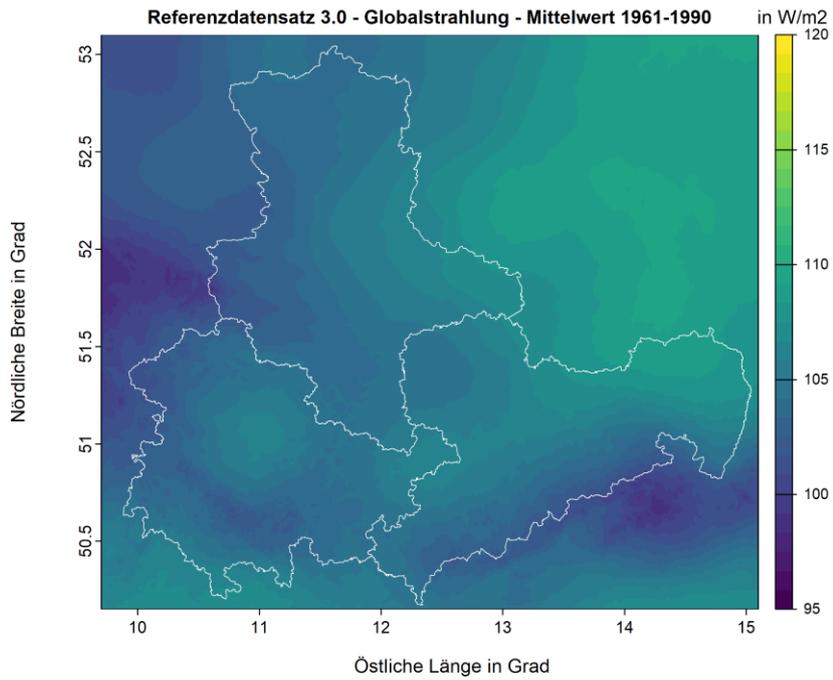


**Abbildung 37: Differenz der mittleren Windgeschwindigkeit aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961–1990**

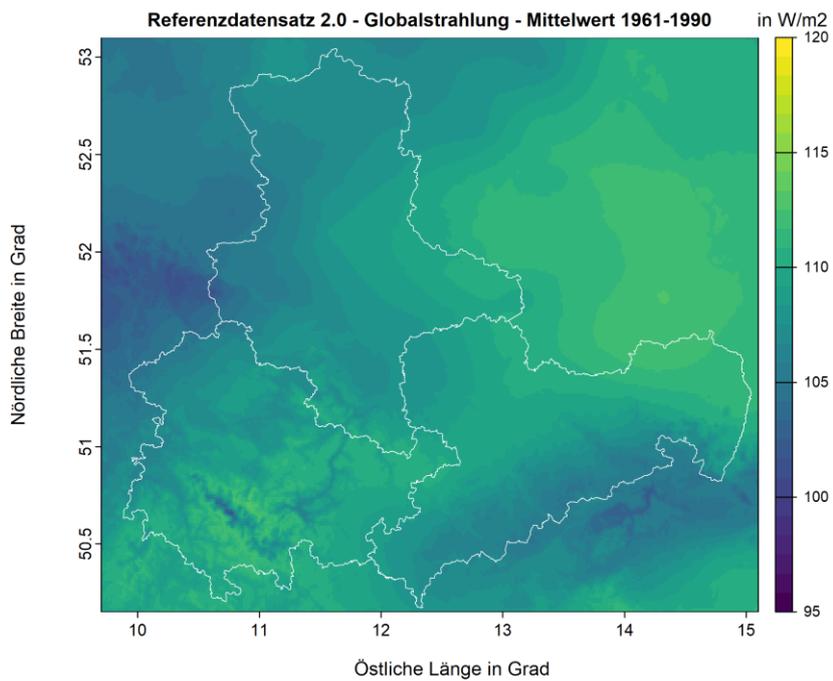
In Abbildung 35 ist die mittlere Windgeschwindigkeit des Referenzdatensatzes 3.0 dargestellt. Demgegenüber steht die mittlere Windgeschwindigkeit aus dem Referenzdatensatz 2.0 in Abbildung 36. Die Differenzendarstellung in Abbildung 37 zeigt, dass die Abweichungen im Mittel gering, lokal aber sehr groß sein können. Besonders ins Auge fällt der Bereich des Brockens rund um den Harz. Die Windgeschwindigkeit wird meistens, aber nicht immer in einer Standardmesshöhe von 10 m gemessen. Im Referenzdatensatz 2.0 war eine räumliche Interpolation der Windgeschwindigkeit zunächst nicht vorgesehen, so dass lediglich lückenlose Tageswerte der Stationen generiert wurden. Im Nachgang des Projektes wurde die Möglichkeit jedoch im Interpolationstool RaKliDa freigeschaltet. Die Zeitreihen im Referenzdatensatz 2.0 sind jedoch nicht auf 10 m korrigiert, was zur Folge hat, dass diese teilweise höhere, teilweise niedrigere Werte aufweisen als in der Standardhöhe von 10 m. Im Referenzdatensatz 3.0 sind wie oben beschrieben, alle Winddaten auf 10 m Messhöhe korrigiert. Für den Bereich um den Brocken bedeutet das konkret, dass statt im Mittel bis zu 9 m/s Windgeschwindigkeit durch die Messhöhe von 28 m, nun maximal noch um 7 m/s mittlere Windgeschwindigkeiten für das Jahr interpoliert werden. Für die anderen Stationen trifft das in ähnlicher Form auch zu, jedoch werden an diesen Stationen nicht so hohe absolute Windgeschwindigkeiten wie auf dem Brocken erreicht.

#### 4.2.7 Globalstrahlung

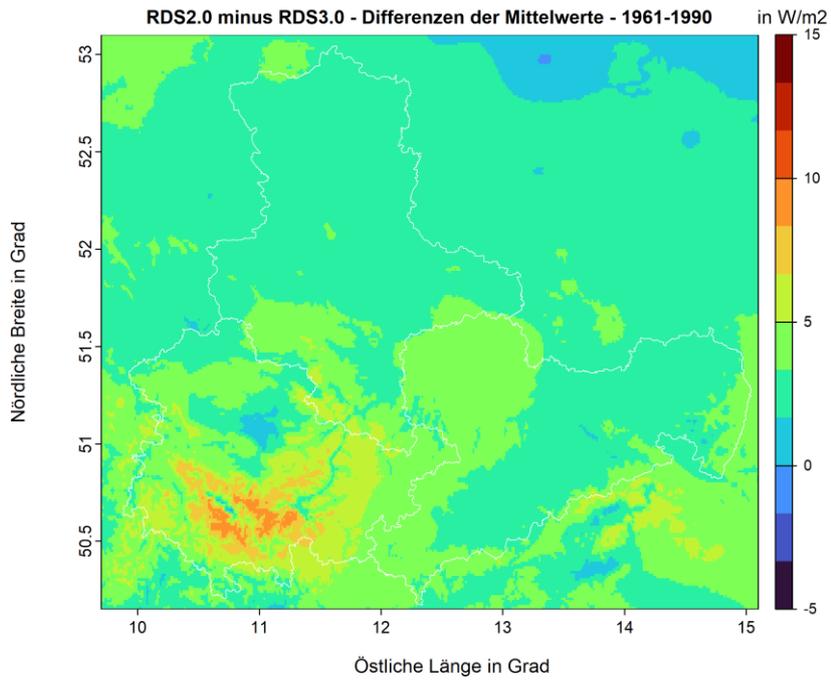
Der Vergleich zeigt die mittlere Globalstrahlung (RG in  $W/m^2$ ) für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.



**Abbildung 38: Globalstrahlung des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 39: Globalstrahlung des Referenzdatensatzes 2.0 für 1961–1990**

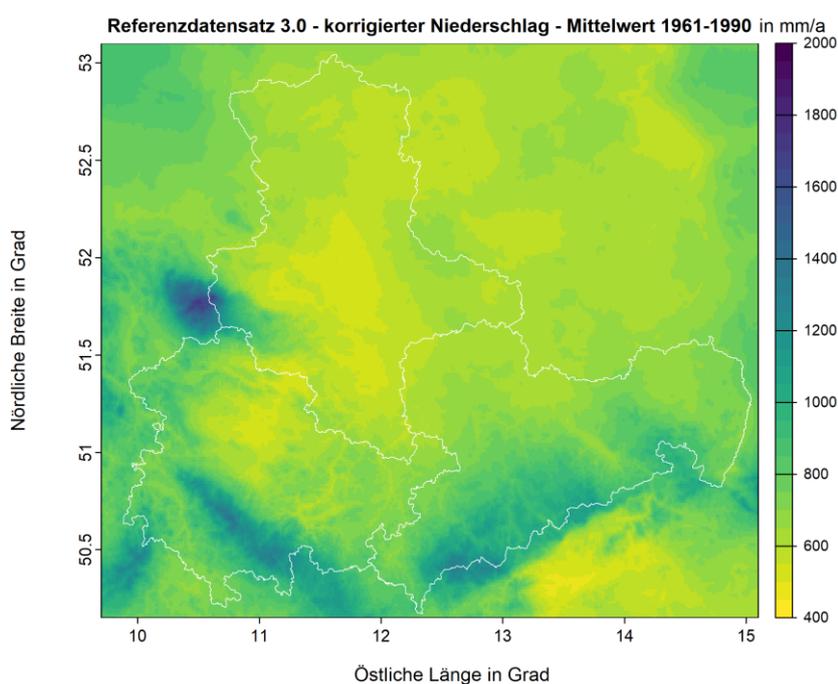


**Abbildung 40: Differenz der Globalstrahlung aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961–1990**

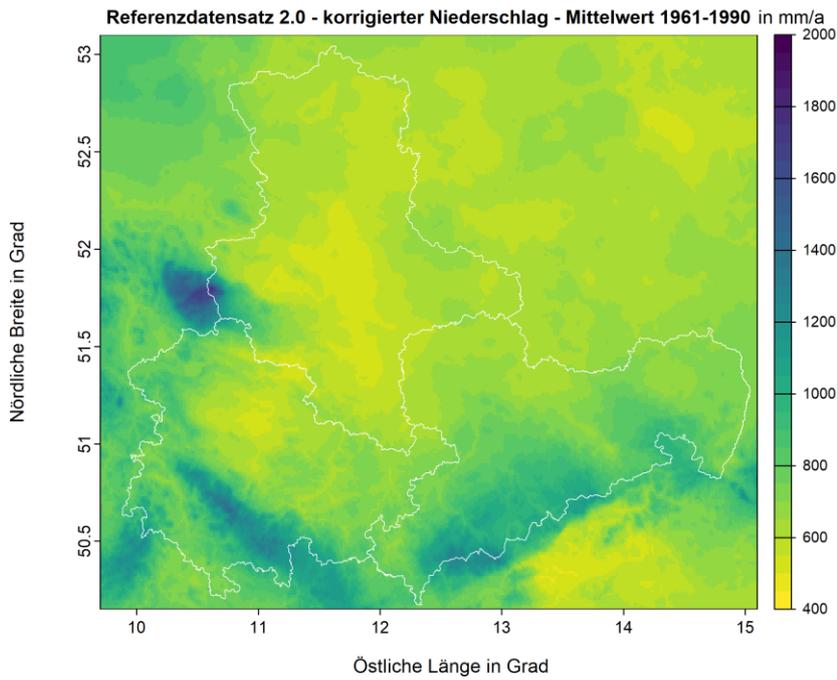
In Abbildung 38 ist die Globalstrahlung des Referenzdatensatzes 3.0 dargestellt. Demgegenüber steht die Globalstrahlung aus dem Referenzdatensatz 2.0 in Abbildung 39. Die Differenzendarstellung in Abbildung 40 zeigt, dass insgesamt sehr geringe Abweichungen von unter 5 % zu verzeichnen sind. Eine Ausnahme bildet das Gebiet um den Thüringer Wald. Ursächlich hierfür sind insgesamt sechs im Vergleich zum Referenzdatensatz 2.0 neue Stationen in diesem Gebiet. Diese sorgen für eine höhere Stationsdichte und somit für eine bessere Qualität des aktuellen Produktes.

#### 4.2.8 Korrigierter Niederschlag

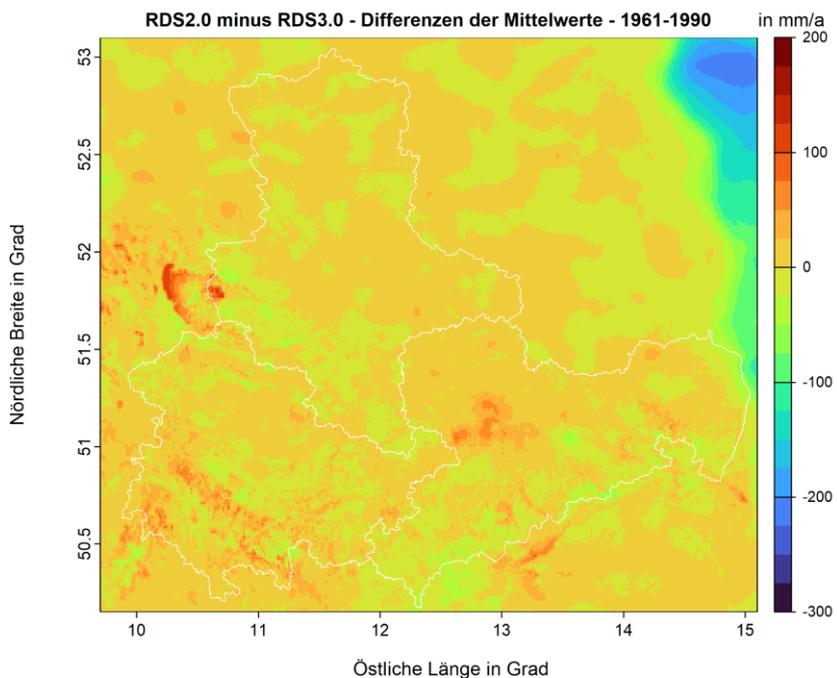
Der Vergleich zeigt die Summe des jährlichen korrigierten Niederschlags (RK in mm) für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.



**Abbildung 41: Korrigierter Niederschlag des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 42: Korrigierter Niederschlag des Referenzdatensatzes 2.0 für 1961–1990**



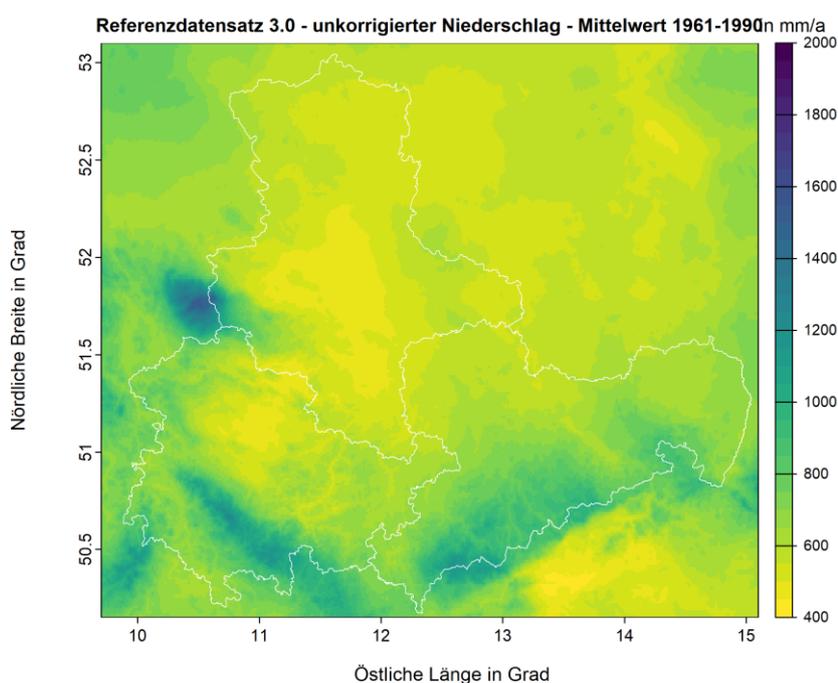
**Abbildung 43: Differenz des korrigierten Niederschlags aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961–1990**

In Abbildung 41 ist der korrigierte Niederschlag des Referenzdatensatzes 3.0 dargestellt. Demgegenüber steht der korrigierte Niederschlag aus dem Referenzdatensatz 2.0 in Abbildung 42. Die Differenz-

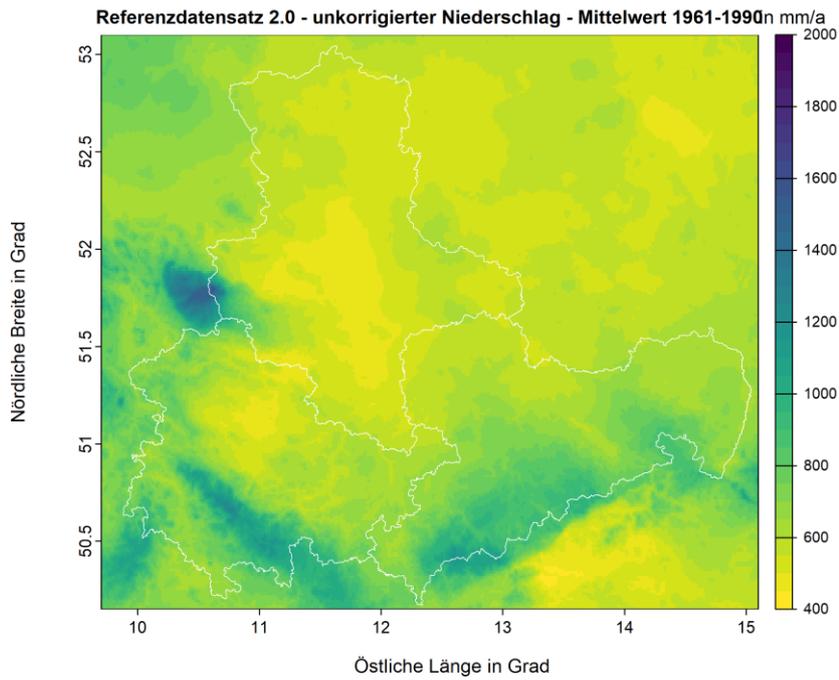
endarstellung in Abbildung 43 zeigt, dass die Unterschiede im Mittel praktisch 0 mm betragen, in einigen Gebieten aber durchaus größere Werte annehmen können. Neben der Höhenabhängigkeit durch das aktuelle DGM ist hier die insgesamt hohe Stationsdichte Ursache für die lokal sehr heterogenen Unterschiede: Zusätzliche Stationen haben nur einen sehr geringen räumlichen Einfluss auf den Niederschlag, da bereits wenige Kilometer weiter eine andere Station den Niederschlag im Gebiet beeinflusst. Ausnahme hiervon ist das polnische Gebiet im Osten des Untersuchungsgebietes: Hier sind zahlreiche neue Stationen hinzugekommen, die das Niederschlagsprodukt entsprechend stark beeinflussen.

#### 4.2.9 Unkorrigierter Niederschlag

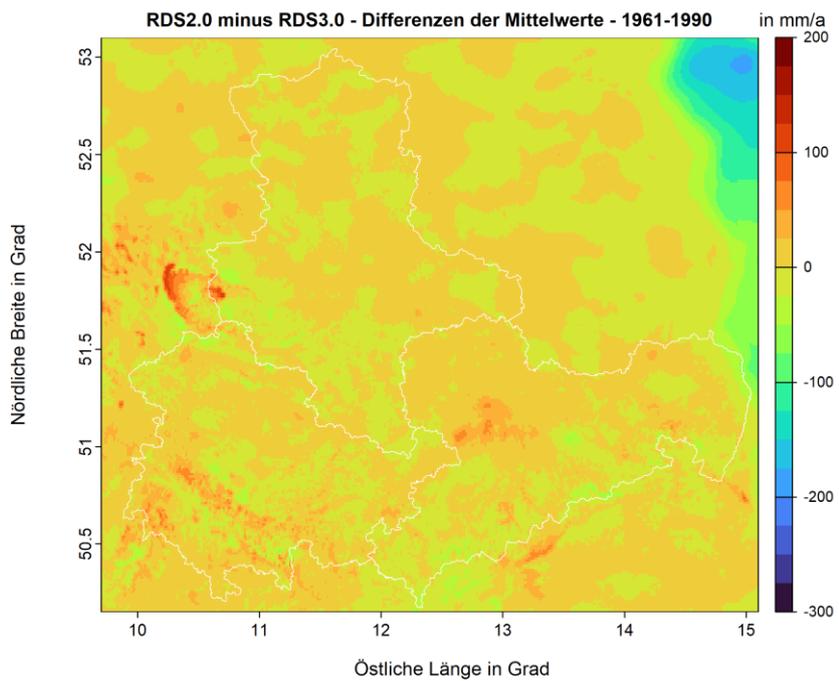
Der Vergleich zeigt die Summe des jährlichen unkorrigierten Niederschlags (RR in mm) für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.



**Abbildung 44: Unkorrigierter Niederschlag des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 45: Unkorrigierter Niederschlag des Referenzdatensatzes 2.0 für 1961–1990**

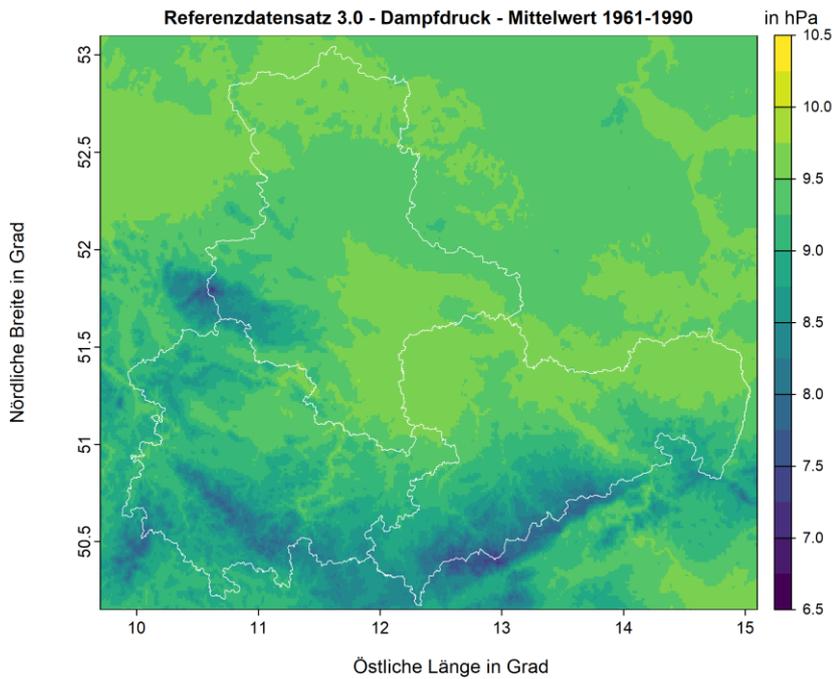


**Abbildung 46: Differenz des unkorrigierten Niederschlags aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961–1990**

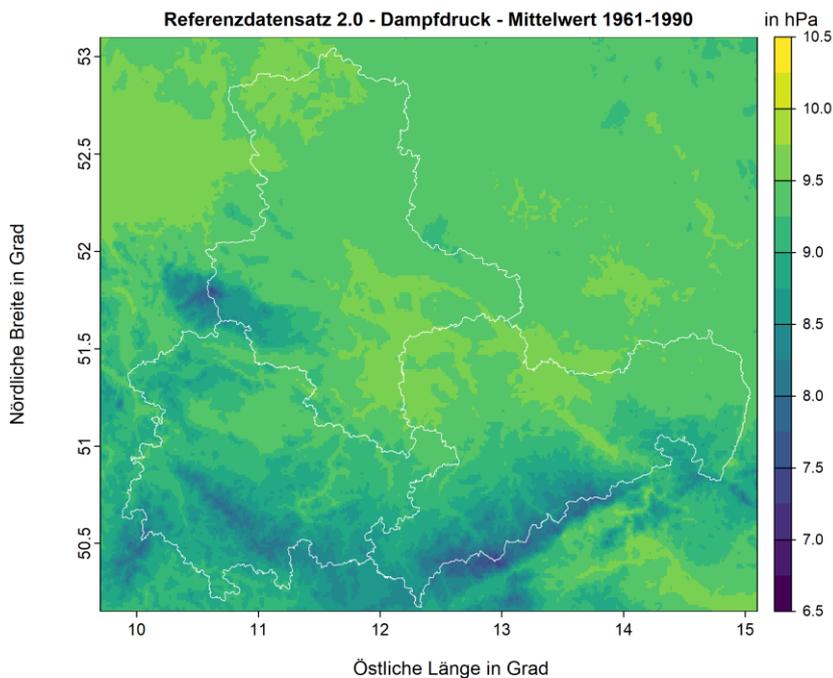
Die Ergebnisse des unkorrigierten Niederschlags sind äquivalent zum korrigierten Niederschlag.

#### 4.2.10 Dampfdruck

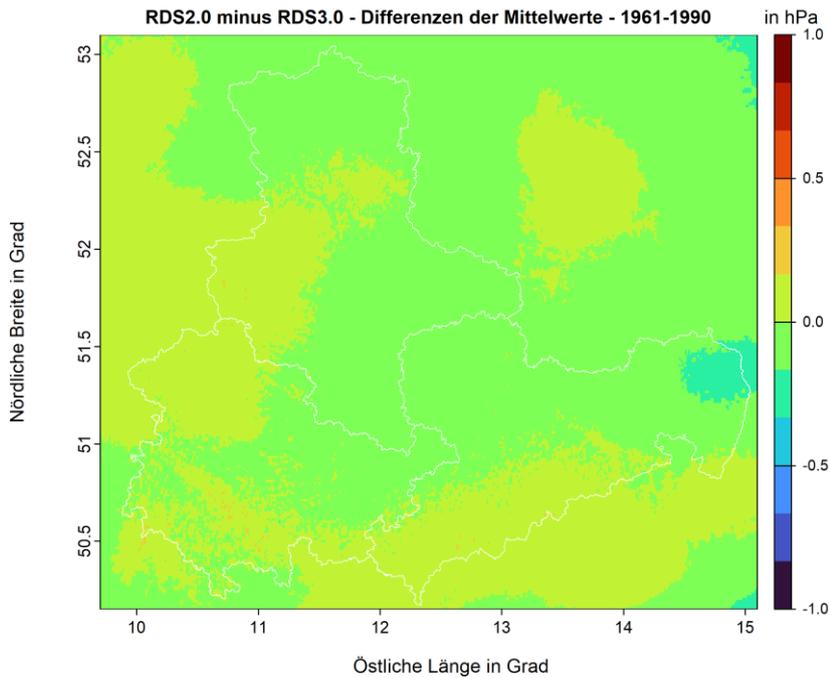
Der Vergleich zeigt den mittleren Dampfdruck (DD in hPa) für den Bereich des Mitteldeutschen Kernens in der Klimareferenzperiode 1961–1990.



**Abbildung 47: Dampfdruck des Referenzdatensatzes 3.0 für 1961–1990**



**Abbildung 48: Dampfdruck des Referenzdatensatzes 2.0 für 1961–1990**



**Abbildung 49: Differenz der Dampfdruck aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961–1990**

In Abbildung 47 ist der Dampfdruck des Referenzdatensatz 3.0 dargestellt. Demgegenüber steht der Dampfdruck aus dem Referenzdatensatz 2.0 in Abbildung 48. Die Differenzendarstellung in Abbildung 49 zeigt, dass die Abweichungen im räumlichen Mittel sehr gering sind. Es sind einige wenige Gebiete sichtbar, wo sich die Differenz auf das neue DGM zurückführen lässt. Zusätzliche Stationen haben kaum Einfluss auf die räumliche Verteilung des Dampfdrucks.

### 4.2.11 Relative Feuchte

Der Vergleich zeigt die mittlere relative Feuchte (RF in %) für den Bereich des Mitteldeutschen Kernensembles in der Klimareferenzperiode 1961–1990.

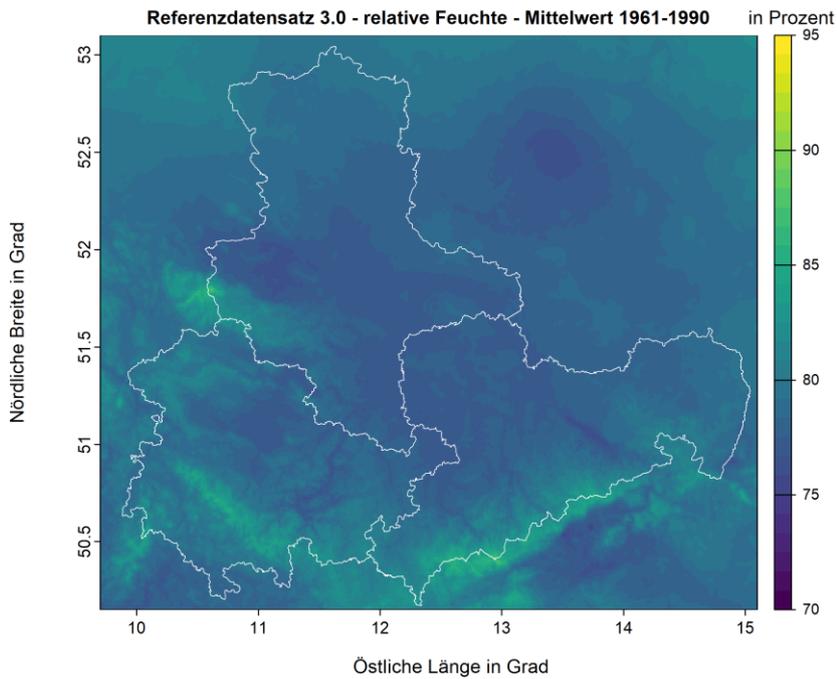


Abbildung 50: Relative Feuchte des Referenzdatensatzes 3.0 für 1961–1990

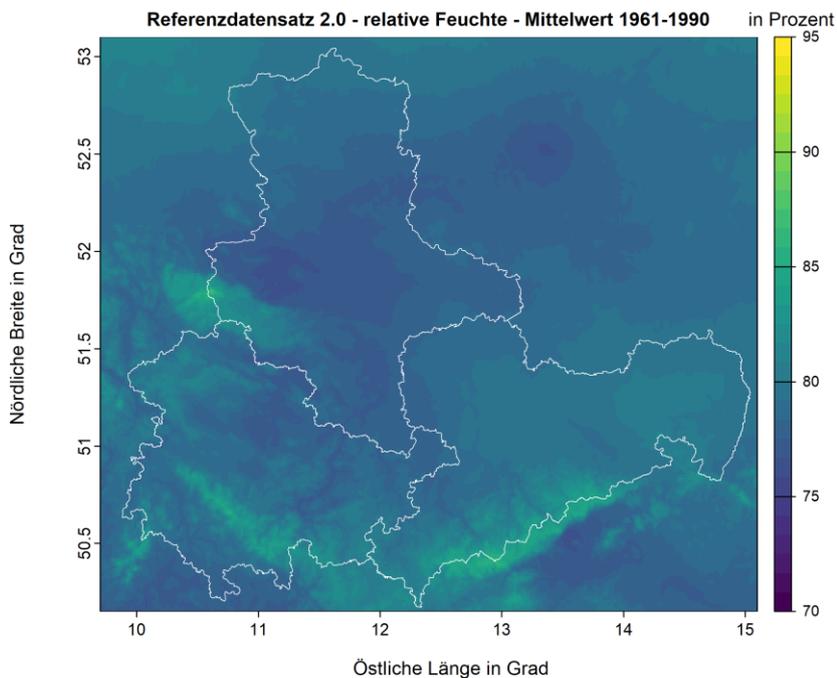
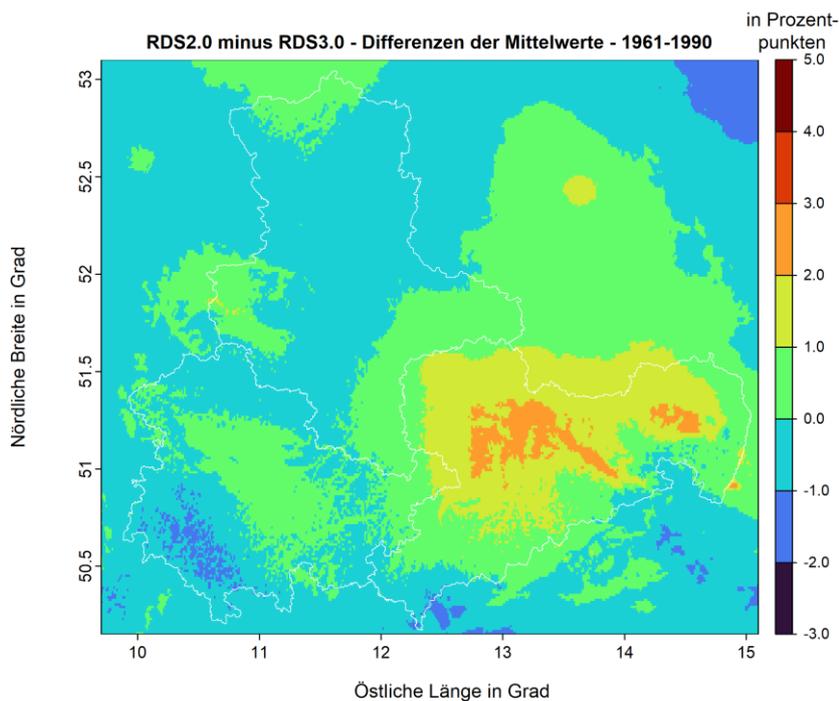


Abbildung 51: Relative Feuchte des Referenzdatensatzes 2.0 für 1961–1990



**Abbildung 52: Differenz der relativen Feuchte aus Referenzdatensatz 2.0 und Referenzdatensatz 3.0 1961–1990**

In Abbildung 50 ist die relative Feuchte des Referenzdatensatzes 3.0 dargestellt. Demgegenüber steht die relative Feuchte aus dem Referenzdatensatz 2.0 in Abbildung 51. Die Differenzendarstellung in Abbildung 52 zeigt, dass, ähnlich wie bei der Temperatur, die relative Feuchte im Mittel nur sehr geringe Abweichungen aufweist. Aufgrund der negativen Korrelation mit der Geländehöhe, wirkt sich das aktuelle DGM zusammen mit neuen Stationen für manche Gebiete stärker aus, so dass teilweise bis über 2 % Abweichungen im Mittel zu verzeichnen sind.

### 4.3 Übersicht des Vergleiches

Die Übersicht in Tabelle 8 fasst die räumlichen Abweichungen der verschiedenen Größen zusammen und erweitert diese Zusammenfassung auf die Klimanormalperiode 1991–2020. Grundsätzlich gibt es Abweichungen zwischen den beiden Klimanormalperioden 1961–1990 und 1991–2020.

**Tabelle 8: Vergleich der Datensätze RefDat3 mit RefDat2 und RefDat3 mit ERA5-land**

Periode	1961 bis 1990						1991 bis 2020					
	RefDat3 vs. RefDat2			RefDat3 vs. ERA5-l			RefDat3 vs. RefDat2			RefDat3 vs. ERA5-l		
Vergleich	RMSE	BIAS	R <sup>2</sup>	RMSE	BIAS	R <sup>2</sup>	RMSE	BIAS	R <sup>2</sup>	RMSE	BIAS	R <sup>2</sup>
DD (hPa)	0,06	-0,02	0,98	-	-	-	0,09	-0,05	0,96	-	-	-
ET (mm)	17,81	-10	0,62	668,18	643,95	0,14	18,07	-13,26	0,76	777,41	752,78	0,08
FM (m/s)	0,22	0,08	0,85	1,89	-1,85	0,37	0,23	0,02	0,79	1,78	-1,74	0,24
GR (mm)	20,43	16,1	0,89	-	-	-	23,45	17,7	0,85	-	-	-
RF (%)	0,67	0,06	0,80	-	-	-	0,67	0,08	0,88	-	-	-
RG (W/m <sup>2</sup> )	3,42	3,23	0,81	12,31	11,79	0,01	0,78	-0,2	0,92	11,3	10,81	0,11
RK (mm)	36,8	-4,64	0,94	85,88	33,35	0,72	31,07	-12,44	0,96	82,42	-5,55	0,72
RR (mm)	30,1	-4,32	0,95	-	-	-	25,44	-10,84	0,97	-	-	-
TM (°C)	0,09	0,01	0,99	0,28	0,04	0,89	0,11	0,02	0,98	0,26	0,07	0,91
TN (°C)	0,13	0	0,97	-	-	-	0,14	-0,01	0,96	-	-	-
TX (°C)	0,1	0,02	0,99	-	-	-	0,09	0,01	0,99	-	-	-

An der grundsätzlichen Aussage gibt es jedoch keine Änderungen. Die Abweichungen des Referenzdatensatzes 3.0 gegenüber Referenzdatensatz 2.0 sind insgesamt sehr klein, die größten Abweichungen sind bei der Windgeschwindigkeit und der potentiellen Verdunstung zu verzeichnen. Die Abweichungen zwischen Referenzdatensatz 3.0 und ERA5-land sind insgesamt sehr viel höher. Bei Strahlung, Windgeschwindigkeit und Verdunstung ist das räumliche Bestimmtheitsmaß zwischen beiden Datensätzen nicht vorhanden bis sehr gering. Bei Niederschlag und Temperatur ist die Übereinstimmung höher, dennoch sehr viel geringer als zwischen beiden Referenzdatensätzen.

## 5 Zusammenfassung

Dieser Bericht legt die Fortschreibung des Klima-Referenzdatensatzes v3.0 für den Zeitraum 1961 bis 2023 dar. Ziel ist es, langfristige, qualitativ hochwertige Daten für Sachsen, Sachsen-Anhalt, Thüringen und Brandenburg bereitzustellen, um Klima- und Klimafolgen-Monitoring zu unterstützen. Grundlage bilden meteorologische Daten von über 2400 Stationen, ergänzt gegenüber dem Vorgängerprodukt durch nationale und internationale Messnetze sowie Nachdigitalisierungen.

Die Aufbereitung umfasst Qualitätskontrollen, Lückenfüllung und Zeitreihenverlängerung. Für Lücken in den Daten werden Gradient-Boosting-Modelle verwendet, die mithilfe korrelierter Stationsdaten präzise Vorhersagen treffen. Zusätzliche Validierungen wie Kreuzvalidierungen und Quantil-Quantil-Mapping (qq-Mapping) verbessern die Genauigkeit, insbesondere bei Extremereignissen.

Die Daten werden von Tages- zu Monats- und Jahreszeitreihen aggregiert, wobei bei Abweichungen die Originaldaten des DWD priorisiert werden. Einige Messgrößen, wie korrigierter Niederschlag oder Grasreferenzverdunstung, werden auf Basis auflösungsabhängiger Berechnungen angepasst. Die Berechnung der Globalstrahlung erfolgt nach DWA-Standards und zeigt teils deutliche Unterschiede zu ERA5-Daten, die physikalisch fragwürdige Werte liefern.

Die Fortschreibung verbessert die räumliche Abdeckung und Datenqualität, etwa durch höhere Stationsdichte und ein präziseres digitales Geländemodell. Der Vergleich mit dem Vorgänger-Referenzdatensatz 2.0 zeigt geringfügige, lokal jedoch signifikante Unterschiede, die aus veränderten Methoden, neuen Stationen und aktualisierten Datengrundlagen resultieren.

Abschließend stellt dieser Bericht eine detaillierte Fortschreibung in der regionalen Klimadatenverarbeitung dar. Die bereitgestellten Daten bilden eine wesentliche Grundlage für künftige Klimaanalysen.

## Literaturverzeichnis

- GAO, LIMING, YAONAN ZHANG, UND LELE ZHANG. „Validation and Spatiotemporal Analysis of Surface Net Radiation from CRA/Land and ERA5-Land over the Tibetan Plateau“. *Atmosphere* 14, Nr. 10 (Oktober 2023): 1542. <https://doi.org/10.3390/atmos14101542>.
- KHADKA, ARBINDRA, PATRICK WAGNON, FANNY BRUN, DIBAS SHRESTHA, YVES LEJEUNE, UND YVES ARNAUD. „Evaluation of ERA5-Land and HARv2 Reanalysis Data at High Elevation in the Upper Dudh Koshi Basin (Everest Region, Nepal)“, 1. August 2022. <https://doi.org/10.1175/JAMC-D-21-0091.1>.
- KÖRNER, PHILIPP, IVAN VOROBEVSKII, RICO KRONENBERG, UND AHMED HOMOUDI. „Klima- Referenzdatensatz Sachsen 1961 - 2020“. Schriftenreihe des LfULG. Dresden, 2022.
- PELOSI, ANNA, FABIO TERRIBILE, GUIDO D'URSO, UND GIOVANNI BATTISTA CHIRICO. „Comparison of ERA5-Land and UERRA MESCAN-SURFEX Reanalysis Data with Spatially Interpolated Weather Observations for the Regional Assessment of Reference Evapotranspiration“. *Water* 12, Nr. 6 (Juni 2020): 1669. <https://doi.org/10.3390/w12061669>.
- PETTITT, A. N. „A Non-Parametric Approach to the Change-Point Problem“. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 28, Nr. 2 (1979): 126-35. <https://doi.org/10.2307/2346729>.
- RICHTER, DIETER. Ergebnisse methodischer Untersuchungen zur Korrektur des systematischen Meßfehlers des Hellmann-Niederschlagsmessers. Berichte des Deutschen Wetterdienstes 194. Offenbach am Main: Selbstverlag des Deutschen Wetterdienstes, 1995.
- TIAN, QIANQIAN, SHUHUA ZHANG, WEILI DUAN, UND GUANGHUI MING. „Evaluation of Downward Shortwave Radiation Products Over the Loess Plateau“. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 17 (2024): 3167–80. <https://doi.org/10.1109/JSTARS.2023.3346032>.
- XU, CHAO, WEN WANG, YANJUN HU, UND YONGWEI LIU. „Evaluation of ERA5, ERA5-Land, GLDAS-2.1, and GLEAM potential evapotranspiration data over mainland China“. *Journal of Hydrology: Regional Studies* 51 (1. Februar 2024): 101651. <https://doi.org/10.1016/j.ejrh.2023.101651>.

**Herausgeber**

Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie (LfULG)  
Pillnitzer Platz 3, 01326 Dresden  
Telefon: + 49 351 2612-0; Telefax: + 49 351 2612-1099  
E- Mail: Poststelle.LfULG@smekul.sachsen.de  
www.lfulg.sachsen.de

**Autoren**

Dr. Philipp Körner, Dr. Uta Moderow  
iamk GmbH  
Institut für Angewandte Meteorologie und Klimatologie  
Dr.-Wilhelm-Külz-Straße 70, 01896 Pulsnitz  
Telefon: + 49 176 - 45747729; Telefax: + 49 351 XXXX-X  
E-Mail: philipp.koerner@iamk-gmbh.de

**Redaktion**

Dr. Johannes Franke, Dr. Daniel Hertel  
Abteilung 5 / Referat 55  
Söbrigener Str. 3a, 01326 Dresden-Pillnitz  
Telefon: + 49 351 2612-5500; Telefax: + 49 351 2612-5099  
E-Mail: johannes.franke@smekul.sachsen.de

**Bildnachweis**

Dr. Philipp Körner, iamk GmbH (Titelbild)

**Redaktionsschluss**

21.02.2025

**ISSN**

1867-2868

**Bestellservice**

Die Broschüre steht nicht als Printmedium zur Verfügung, kann aber als PDF-Datei heruntergeladen werden aus der Publikationsdatenbank des Freistaates Sachsen (<https://publikationen.sachsen.de>).

**Hinweis**

Diese Publikation wird im Rahmen der Öffentlichkeitsarbeit vom LfULG (Geschäftsbereich des SMUL) kostenlos herausgegeben. Sie ist nicht zum Verkauf bestimmt und darf nicht zur Wahlwerbung politischer Parteien oder Gruppen eingesetzt werden.

Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des vom Sächsischen Landtag beschlossenen Haushaltes.

*Täglich für ein gutes Leben.*

www.lfulg.sachsen.de