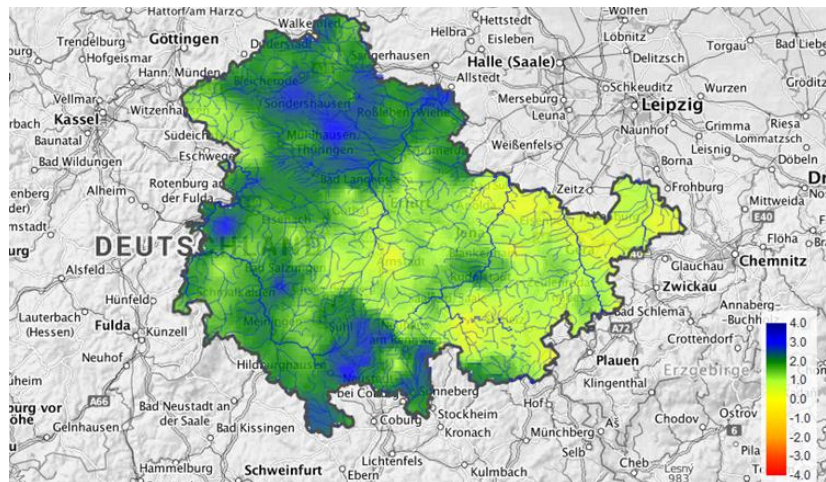


ZWISCHENBERICHT

Dürretrend Thüringen



Thüringer Fernwasserversorgung (TFW)

Projektnummer

WBW 001064-26

Projektname

Analyse Vorhersagequalität

Dokumentversion

v1.5

Datum

11.09.2026

Inhalt

1	VERANLASSUNG UND AUFGABENSTELLUNG	5
2	EINFÜHRUNG	6
2.1	Modelle	7
2.2	Dürre-Indizes.....	10
2.2.1	Standardized Precipitation Index (SPI)	10
2.2.2	Soil Moisture Anomaly (SMA)	10
2.2.3	Standardized Streamflow Index (SSI)	11
2.3	Vorhersagen.....	12
2.4	Datenaufbereitung.....	12
2.5	Dashboard.....	13
3	ANALYSE VORHERSAGEN.....	14
3.1	Methodik.....	15
3.2	Qualitative Analyse Vorhersagen.....	16
3.3	Quantitative Analyse.....	19
3.3.1	Genauigkeit / Accuracy	20
3.3.2	Verlässlichkeit (Reliability) und Trennschärfe (Discrimination)	28
3.3.3	Skill	31
3.3.4	Bias	36
4	ZUSAMMENFASSUNG.....	37
5	ANHANG	39

Abbildungsverzeichnis

Abbildung 1: Räumliche Ausdehnung der verwendeten hydrologischen Modelle J2000g (rot) und Talsim (grün) in Relation zur Landesgrenze von Thüringen (schwarz)	7
Abbildung 2: Niederschlagsprodukt HYRAS-DE-PR im Bereich der Modellgebiete (dargestellt sind Monatswerte für Januar 2025; je dunkelblauer desto höher der Niederschlag)	8
Abbildung 3: Lage der DWD-Klimastationen von denen gemessene Temperaturdaten als Antrieb für das J2000g-Modell verwendet werden	9
Abbildung 4: Lage der DWD-Klimastationen von denen gemessene Temperaturdaten als Antrieb für das Talsim-Modell verwendet werden	9
Abbildung 5: GCFS-EPISODES Vorhersagedaten für Niederschlag (Beispiel) in Relation zu den Modellgebieten (dargestellt sind Tageswerte eines Ensemble-Mitglieds der Januar-Vorhersage 2025 für 02.02.2025; je dunkelblauer desto höher der vorhergesagte Niederschlag)	12
Abbildung 6: Screenshot des Dashboards (Stand Mai 2025)	14
Abbildung 7: Karte adaptierte klimatische Naturräume Thüringen	15
Abbildung 8: Thüringer Wald / Schiefergebirge (Skala 6 Monate) - alle Vorhersagen SPI	17
Abbildung 9: Gebiet Ringgau / Harzvorland (Skala 6 Monate) - alle Vorhersagen SPI	17
Abbildung 10: Thüringer Wald / Schiefergebirge (Skala 6 Monate) - Vorhersagehorizont 2 Monate SPI	18
Abbildung 11: Unstrut / Thüringer Becken (Skala 6 Monate) - Vorhersagehorizont 2 Monate SPI	18
Abbildung 12: Thüringer Wald / Schiefergebirge (Skala 6 Monate) - Vorhersagehorizont 4 Monate SPI	19
Abbildung 13: Unstrut / Thüringer Becken (Skala 6 Monate) - Vorhersagehorizont 4 Monate SPI	19
Abbildung 14: Thüringen CRPS - SPI alle Skalen	20
Abbildung 16: Thüringen CRPS - Karten SPI.6M (Vorhersagehorizonte 1, 3, 5 und 6 Monate)	21
Abbildung 16: Thüringen CRPS Regionen - Heatmap nach Zielzeitraum SPI.6M (Vorhersagehorizonte 2, 4 und 6 Monate)	23
Abbildung 17: Thüringen CRPS - Karten SMA.30d (Vorhersagehorizonte 1, 3, 5 und 6 Monate)	23
Abbildung 18: Thüringen CRPS - Karten SSI.3M (Vorhersagehorizonte 1, 3, 5 und 6 Monate)	24
Abbildung 20: Anzahl Schwellwertunterschreitungen Januar 2023 - März 2026	25
Abbildung 20: Thüringen Brier Score - Vergleich Skalen (Schwellwert 0)	26
Abbildung 21: Thüringen Brier Score - Vergleich Skalen (Schwellwert -1)	27
Abbildung 22: Thüringen Brier Score - Karten SPI.6M (Schwellwert -1; Vorhersagehorizonte 1, 3, 5 und 6 Monate)	28
Abbildung 23: Thüringer Wald / Schiefergebirge Reliability-Diagramm - SPI.6M Schwellwert 0 (Vorhersagehorizonte 2, 4, 5 und 6 Monate)	29
Abbildung 24: Nordthüringen Discrimination-Diagramm - SPI.12M Schwellwert 0 (Vorhersagehorizont 1 Monat)	30
Abbildung 25: Altenburg / Ronneburg Discrimination-Diagramm - SPI.6M Schwellwert -1 (Vorhersagehorizonte 1, 3, 5 und 6 Monate)	31
Abbildung 26: Thüringen Brier-Skill-Score - Karten SPI.3M (Schwellwert -1; Vorhersagehorizonte 1, 3 und 5 Monate)	33
Abbildung 27: Thüringen Brier-Skill-Score - Karten SPI.6M (Schwellwert -1; Vorhersagehorizonte 1, 3 und 5 Monate)	34
Abbildung 28: Thüringen CRPS-Skill-Score - Karten SPI.6M (Vorhersagehorizonte 1, 3 und 6 Monate)	35
Abbildung 29: Thüringen Bias - Karten SPI.6M (Vorhersagehorizonte 1, 3, 5 und 6 Monate)	36
Abbildung 30: Thüringen Bias - Karten für einzelne Ereignisse	37
Abbildung 32: Thüringen CRPS - Karten SPI.3M (Vorhersagehorizonte 1, 3, 5 und 6 Monate)	39

Abbildung 32: Thüringen CRPS - Karten SPI.12M (Vorhersagehorizonte 1, 3, 5 und 6 Monate)	40
Abbildung 33: Thüringen CRPS Regionen - Heatmap nach Zielzeitraum SPI.3M (Vorhersagehorizonte 2, 4 und 6 Monate)	41
Abbildung 34: Thüringen CRPS Regionen - Heatmap nach Zielzeitraum SPI.12M (Vorhersagehorizonte 2, 4 und 6 Monate)	42
Abbildung 35: Thüringen Brier Score - Karten SPI.3M (Schwellwert -1; Vorhersagehorizonte 1, 3, 5 und 6 Monate)	43
Abbildung 36: Thüringen Brier Score - Karten SPI.12M (Schwellwert -1; Vorhersagehorizonte 1, 3, 5 und 6 Monate)	44
Abbildung 37: Thüringen Brier-Skill-Score - Karten SPI.12M (Schwellwert -1; Vorhersagehorizonte 1, 3 und 5 Monate)	45
Abbildung 38: Thüringen CRPS-Skill-Score - Karten SPI.3M (Vorhersagehorizonte 1, 3 und 6 Monate)	46
Abbildung 39: Thüringen CRPS-Skill-Score - Karten SPI.12M (Vorhersagehorizonte 1, 3 und 6 Monate)	47

Abkürzungsverzeichnis

DWD	Deutscher Wetterdienst
GCFS-EPIISODES	Saisonales Vorhersageprodukt des DWD
HYRAS-DE-PR	Niederschlagsprodukt des DWD
SMA	Soil Moisture Anomaly Index
SPI	Standardized Precipitation Index
SSI	Standardized Streamflow Index
TFW	Thüringer Fernwasserversorgung
TLUBN	Thüringer Landesamt für Umwelt, Bergbau und Naturschutz

Dürretrend Thüringen

1 VERANLASSUNG UND AUFGABENSTELLUNG

Im Forschungsvorhaben TASK2 „Operativer Einsatz saisonaler hydro-meteorologischer Vorhersagen mit Wasserbilanz- und Bodenwasserhaushaltsmodellierung zur Anpassung an Trockenheit durch den Klimawandel“¹, gefördert durch das Bundesministerium für Umwelt, Naturschutz, Bau und Reaktorsicherheit, bearbeitet von SYDRO Consult GmbH in Zusammenarbeit mit der Thüringer Fernwasserversorgung (TFW) und der Stadt Erfurt im Zeitraum 2020 bis 2023, wurde ein webbasiertes, interaktives Dashboard zur Visualisierung verschiedener Dürre-Indizes für Thüringen entwickelt². Im Dashboard werden Dürre-Indizes für den historischen Zeitraum seit dem Jahr 2000 bis heute, sowie auf Basis von saisonalen Vorhersagen berechnete Indizes für die Zukunft dargestellt.

Während der Projektlaufzeit wurde das Dashboard monatlich mit aktuellen Daten und Vorhersagen aktualisiert. Gemeinsam mit dem Thüringer Landesamt für Umwelt, Bergbau und Naturschutz (TLUBN) wurde das Dashboard im Jahr 2025 weitergeführt und weiterentwickelt, nun unter dem Namen „Dürretrend Thüringen“.

Die Thüringer Fernwasserversorgung (TFW) hat die Fortführung des Dashboards für das Jahr 2026 beauftragt.

¹ Projekthomepage: <https://task.sydro.de>

² Dashboardseite: <https://duerretrend.thueringen.de/>

2 EINFÜHRUNG

Die interaktive Webanwendung **Dürretrend Thüringen** (im Folgenden „Dashboard“ genannt) ermöglicht die Visualisierung von verschiedenen Dürre-Indizes für Thüringen. Das Dashboard ist unter der Adresse <https://duerretrend.thueringen.de> erreichbar.

Es ist ein Produkt entstanden aus dem vom Bundesministerium für Umwelt, Klimaschutz, Naturschutz und nukleare Sicherheit (BMU, aktuell BMUKN) geförderten Forschungsvorhaben "Operativer Einsatz saisonaler hydro-meteorologischer Vorhersagen mit Wasserbilanz- und Bodenwasserhaushaltsmodellierung zur Anpassung an Trockenheit durch den Klimawandel" (TASK2).

Mehr Informationen zum Forschungsvorhaben sind auf der Projektwebseite task.sydro.de verfügbar.

Seit Januar 2026 wird die fortlaufende Aktualisierung der Anwendung durch die freundliche Unterstützung der Thüringer Fernwasserversorgung (TFW) ermöglicht und vom Thüringer Landesamt für Umwelt, Bergbau und Naturschutz (TLUBN) begleitet.

In vorliegendem Bericht wird die Analyse der Qualität der bereits durchgerechneten Vorhersagen vorgestellt.

2.1 Modelle

Die Dürre-Indizes werden teilweise auf Basis von mit hydrologischen Modellen berechneten Simulationsergebnissen berechnet. Dabei kommen zwei Modelle bzw. Datensätze zum Einsatz:

- J2000g-Modell³: Dieser Datensatz deckt das gesamte Bundesland Thüringen sowie die stromauf gelegenen Einzugsgebiete in den Nachbarländern ab.
- Talsim-Modell⁴: Dieser Datensatz deckt nur die Einzugsgebiete von Gera und Scherkonde ab ("Pilotgebiet" im TASK2-Vorhaben), hat dafür aber einen höheren Detaillierungsgrad und bildet auch den wasserwirtschaftlichen Betrieb (z.B. Talsperren) im Gebiet ab.

Die räumliche Ausdehnung der beiden Modelle ist in folgender Abbildung 1 dargestellt:

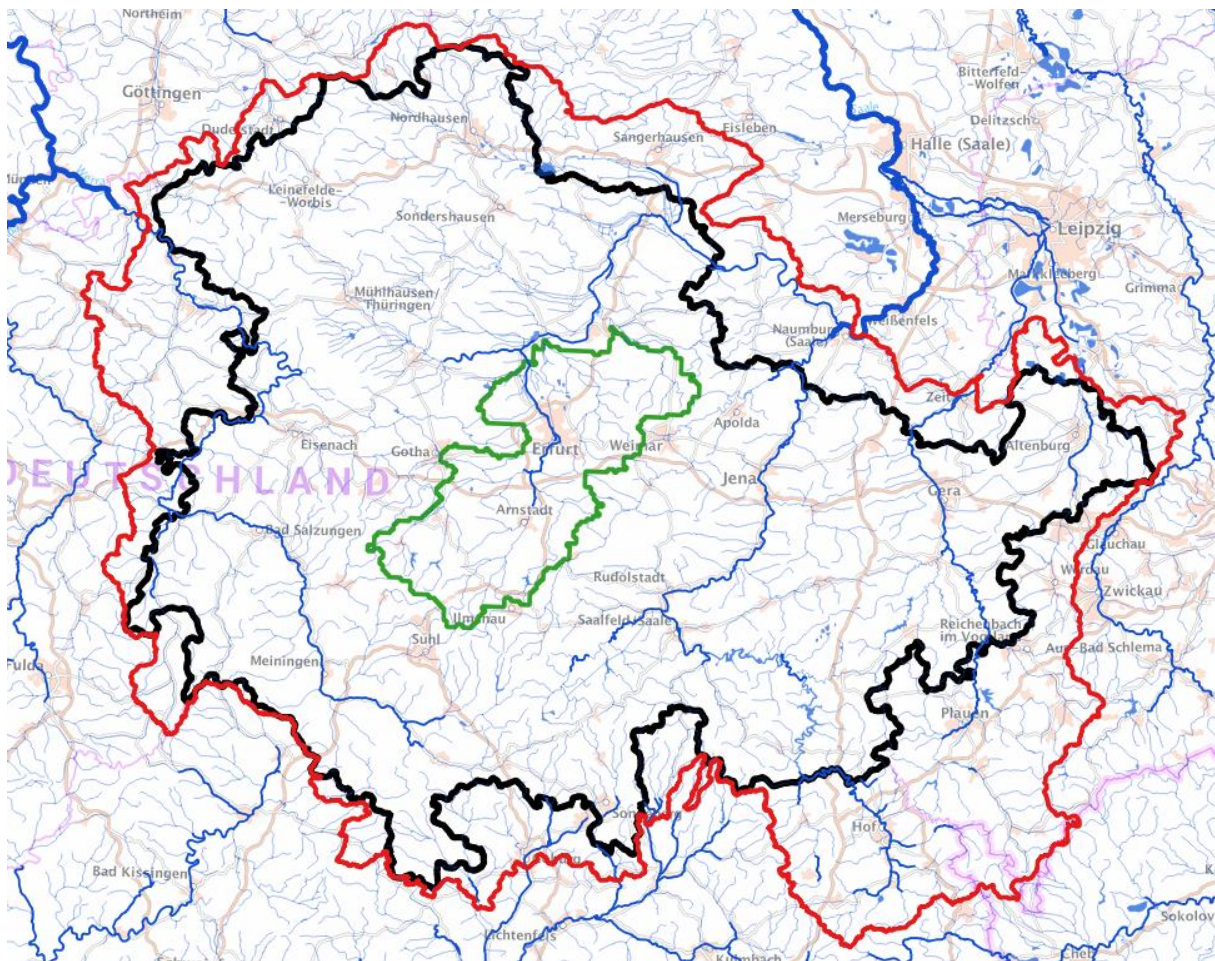


Abbildung 1: Räumliche Ausdehnung der verwendeten hydrologischen Modelle J2000g (rot) und Talsim (grün) in Relation zur Landesgrenze von Thüringen (schwarz)

³ https://jams.uni-jena.de/ilmswiki/index.php/Hydrological_Model_J2000g

⁴ <https://www.talsim.de>

Die beiden hydrologischen Modelle werden jeweils mit Niederschlags- und Temperaturdaten in täglicher Auflösung angetrieben und mit einem Simulationszeitschritt von einem Tag simuliert.

Die Datengrundlage für den gemessenen Niederschlag ist das DWD-Produkt HYRAS-DE-PR⁵ mit einer räumlichen Auflösung von 1x1km und einer zeitlichen Auflösung von 1 Tag.

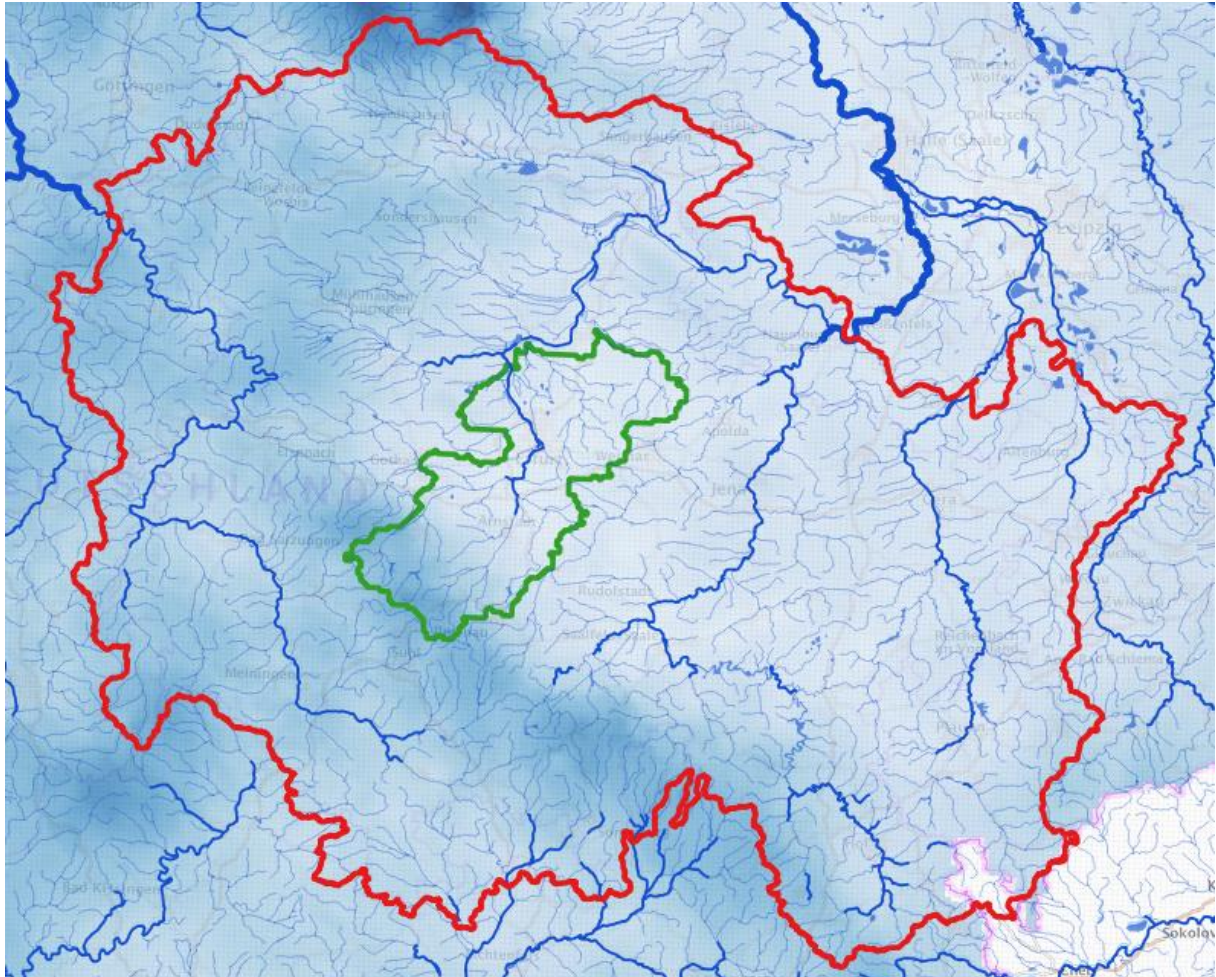


Abbildung 2: Niederschlagsprodukt HYRAS-DE-PR im Bereich der Modellgebiete (dargestellt sind Monatswerte für Januar 2025; je dunkelblauer desto höher der Niederschlag)

⁵ [Raster der Tagessumme des Niederschlags in mm für Deutschland - HYRAS-DE-PR, Version v6.1](#)

Die Datengrundlage für die gemessene Temperatur, welche modellintern u.a. dazu verwendet wird, die Verdunstung zu berechnen, bilden Temperaturdaten von DWD-Bodenstationen⁶ mit einer zeitlichen Auflösung von 1 Tag (Lage siehe folgende Abbildung 3 und Abbildung 4).

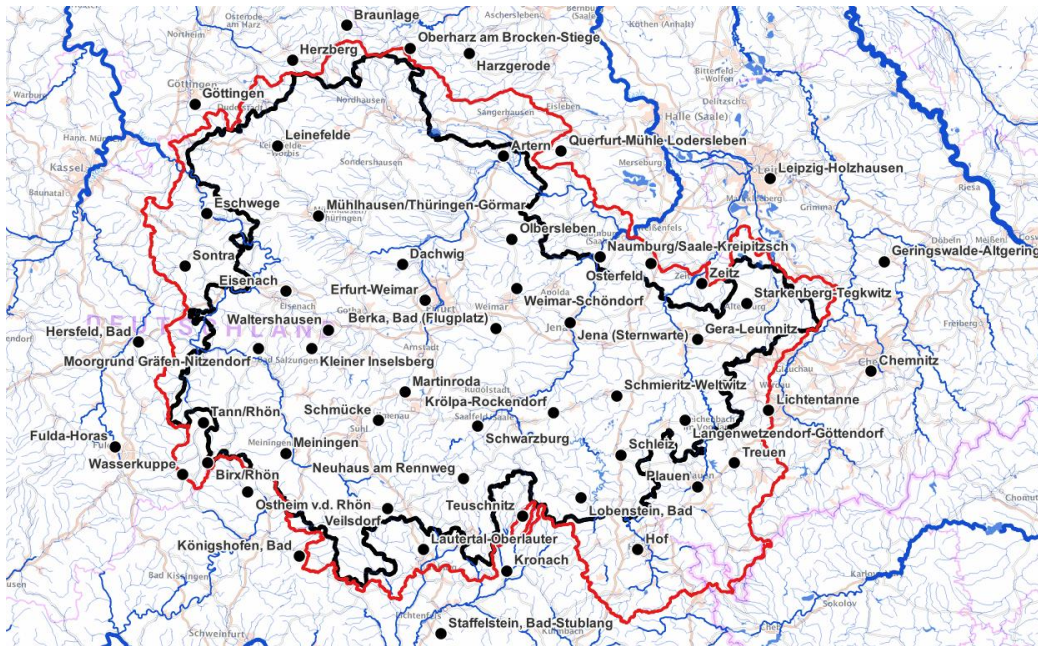


Abbildung 3: Lage der DWD-Klimastationen von denen gemessene Temperaturdaten als Antrieb für das J2000g-Modell verwendet werden

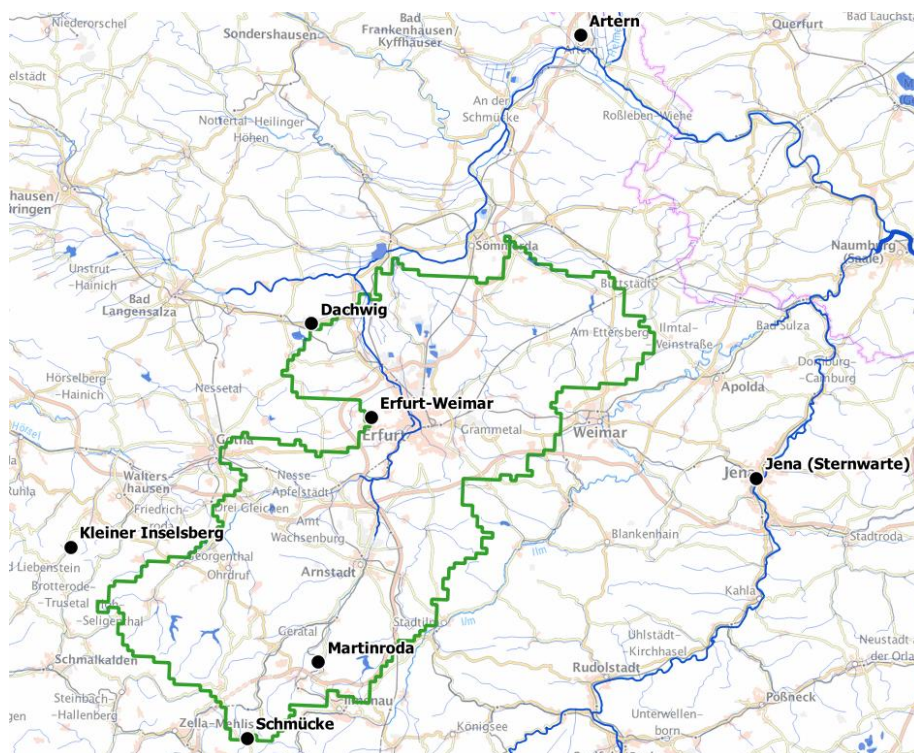


Abbildung 4: Lage der DWD-Klimastationen von denen gemessene Temperaturdaten als Antrieb für das Talsim-Modell verwendet werden

⁶ [Tägliche Stationsbeobachtungen \(Temperatur, Druck, Niederschlag, Sonnenscheindauer, etc.\) für Deutschland, Version v24.3](#)

Für die Berechnung von Index-Vorhersagen werden Niederschlags- und Temperaturdaten aus dem DWD-Produkt GCFS-EPISODES für den Antrieb der Modelle verwendet (siehe Abschnitt 2.3).

2.2 Dürre-Indizes

Im Dashboard werden verschiedene Dürre-Indizes angezeigt, die im Folgenden beschrieben werden. Die jeweils zur Verfügung gestellten Aggregationszeiträume (Skalen) wurden in Absprache mit den Projektbeteiligten und Anwendern des TASK2-Projekts festgelegt. Für die Berechnung aller Indizes wird als Referenzzeitraum der Zeitraum 1952-2011 verwendet.

2.2.1 Standardized Precipitation Index (SPI)

Der Standardized Precipitation Index (SPI) ist ein weit verbreiteter Niederschlagsindex zur Ermittlung von Niederschlagsüberschüssen und -defiziten. Die Berechnung erfolgt auf Basis der Niederschlagsdaten.

Der Wertebereich des Index reicht von etwa -4 bis 4, wobei negative Werte unterdurchschnittlichen und positive Werte überdurchschnittlichen Niederschlag anzeigen.

Es werden Aggregationszeiträume (Skalen) von 3, 6 und 12 Monaten berechnet und im Dashboard dargestellt.

Weitere Informationen zum Index SPI:

- [DWD: Erläuterungen zum Standardized Precipitation Index \(SPI\)](#)
- [WMO: Standardized Precipitation Index User Guide \(2012\)](#)

Die Datengrundlage für die Berechnung des SPI ist das DWD-Produkt HYRAS-DE-PR⁷ mit einer räumlichen Auflösung von 1x1km und zeitlicher Auflösung von Monatssummen.

2.2.2 Soil Moisture Anomaly (SMA)

Der Indikator für die Bodenfeuchteanomalie (SMA) ist ein weit verbreiteter Index. Er wird unter anderem auch von der Europäischen Dürrebeobachtungsstelle (EDO) verwendet. Der Index wird zur Bestimmung des Beginns und der Dauer von Dürreperioden in der Landwirtschaft verwendet, die auftreten, wenn die Verfügbarkeit von Bodenfeuchte für Pflanzen auf ein kritisches Niveau sinkt. Hierzu berücksichtigt der Index in seiner Berechnung neben der über die Zeit variablen Bodenfeuchte auch konstante Bodeneigenschaften wie Feldkapazität und Welkepunkt.

Der Wertebereich des Index reicht von etwa -4 bis 4, wobei negative Werte für Trockenheit stehen.

Die Berechnung des SMA erfolgt auf Basis der mit den beiden Modellen berechneten Bodenfeuchten.

Es werden Aggregationszeiträume (Skalen) von 10 und 30 Tagen berechnet und im Dashboard dargestellt.

Weitere Informationen zum Index SMA:

- [EDO - Soil Moisture Anomaly Indicator Factsheet \(archivierte Version\)](#)

⁷ [Raster der Tagessumme des Niederschlags in mm für Deutschland - HYRAS-DE-PR, Version v6.1](#)

2.2.3 Standardized Streamflow Index (SSI)

Der Standardized Streamflow Index (SSI) ist ein hydrologischer Index, der zur Bewertung von Abflussmengen genutzt wird. Der SSI ermöglicht eine standardisierte Analyse von Abweichungen des aktuellen Abflusses im Vergleich zu einem festgelegten Referenzzeitraum.

Der SSI-Index bietet einen Wertebereich von etwa -4 bis 4, wobei negative Werte auf niedrigen Abfluss und positive Werte auf hohe Abflussmengen im Vergleich zum Referenzzeitraum hinweisen.

Es werden Aggregationszeiträume (Skalen) von 1 und 3 Monaten berechnet und im Dashboard angezeigt.

Die Berechnung des SSI erfolgt auf Basis der mit den beiden Modellen berechneten Abflüsse in Gewässern (Talsim) bzw. an ausgewählten Pegelstationen (J2000g).

2.3 Vorhersagen

Datengrundlage für die Indexvorhersagen sind mit dem EPISODES Verfahren regionalisierte, saisonale Klimavorhersagen für Niederschlag und Temperatur des DWD aus dem German Climate Forecast System (GCFS), genannt GCFS-EPISODES.

Diese saisonalen Vorhersagen werden monatlich aktualisiert, reichen 7 Monate in die Zukunft und bestehen aus jeweils 50 Ensemble Mitgliedern. Sie besitzen eine räumliche Auflösung von 5x5 km und eine zeitliche Auflösung von 1 Tag. Abbildung 5 zeigt beispielhaft für einen Zeitschritt die räumliche Auflösung der Vorhersagedaten im Vergleich zu den Modellgebieten.

Weitere Informationen zum Vorhersageprodukt:

- [DWD - Deutsches Klimavorhersage System](#)
- Hoff, Amelie; Pasternack, Alexander; Wehring, Sabrina; Fröhlich, Kristina; Lorenz, Philip; Paxian, Andreas; Kreienkamp, Frank; Früh, Barbara: Seasonal Climate Predictions for Germany (EPISODES) version 2022.01 https://doi.org/10.5676/DWD/SEASONAL-EPI-DE_V2022.01

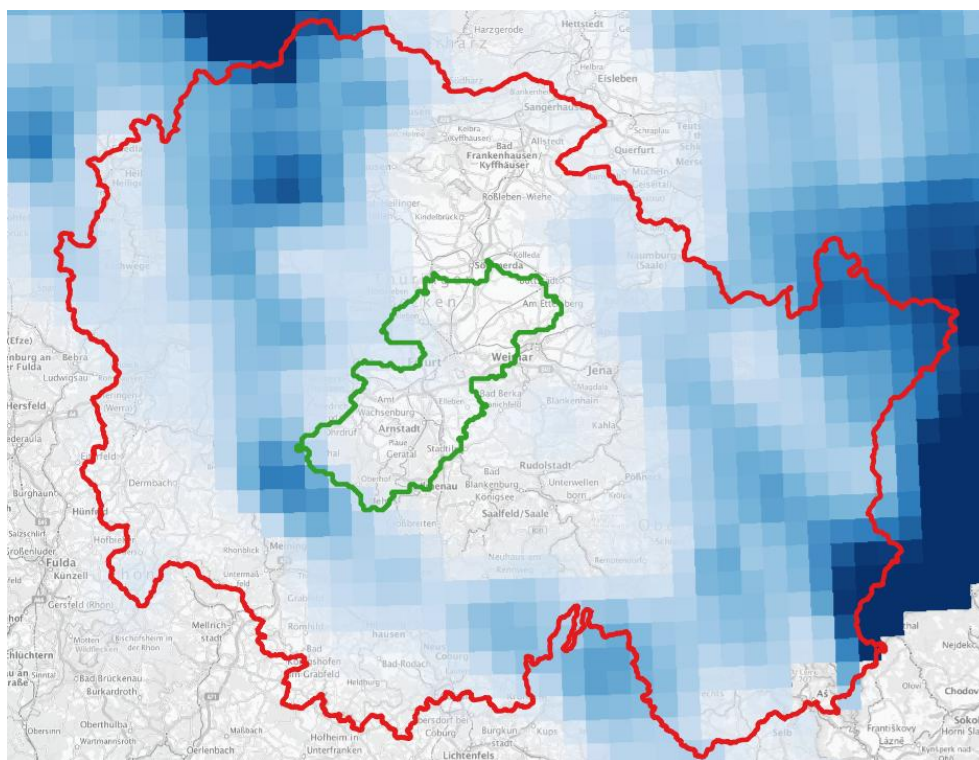


Abbildung 5: GCFS-EPISODES Vorhersagedaten für Niederschlag (Beispiel) in Relation zu den Modellgebieten (dargestellt sind Tageswerte eines Ensemble-Mitglieds der Januar-Vorhersage 2025 für 02.02.2025; je dunkelblauer desto höher der vorhergesagte Niederschlag)

2.4 Datenaufbereitung

Für die Datenaufbereitung und Berechnung der Indizes werden monatlich die folgenden Schritte ausgeführt:

- Aktualisierung der Messdaten von Niederschlag und Temperatur bis zum aktuellen Zeitpunkt
- Für ganz Thüringen:
 - Berechnung des meteorologischen Index SPI
 - Hydrologische Simulation mit dem thüringenweiten J2000g-Modell
 - Berechnung der Indizes SMA und SSI auf Basis der J2000g-Ergebnisse

- Aufbereitung der Ergebnisse für das Dashboard
- Für das Pilotgebiet:
 - Hydrologische Simulation mit dem Talsim-Modell des Pilotgebiets
 - Berechnung der Indizes SMA und SSI auf Basis der Talsim-Ergebnisse
 - Aufbereitung der Ergebnisse für das Dashboard

Die gleichen Bearbeitungsschritte werden dann nochmal in Kombination mit vorhergesagtem Niederschlag und Temperatur aus der saisonalen Vorhersage des aktuellen Monats durchgeführt. Das hierfür verwendete Vorhersageprodukt GCFS-EPISODES ist eine probabilistische Vorhersage bestehend aus 50 Ensemblemitgliedern. Die Simulationen und Index-Berechnungen werden zunächst für alle 50 Ensemblemitglieder einzeln durchgeführt und die Ergebnisse dann in Form von statistischen Kennwerten wie Quantilen des Ergebnis-Ensembles ausgewertet und für das Dashboard aufbereitet.

2.5 Dashboard

Das Dashboard ist eine interaktive Webanwendung aufrufbar unter <https://duerretrend.thueringen.de>, die im Wesentlichen aus drei Darstellungen besteht:

- Eine Karte zeigt die räumliche Verteilung eines Index für einen Zeitpunkt.
- Ein Radardiagramm zeigt mehrere Indizes für einen Ort und Zeitpunkt.
- Eine Zeitreihenanzeige zeigt den zeitlichen Verlauf der Indizes für einen Ort.

Mit einem Klick in die Karte wird für den ausgewählten Ort der zeitliche Verlauf der Indizes sowie das Radardiagramm angezeigt.

Durch Auswahl von Index und zugehöriger Skala wird die Kartenanzeige sowie die Zeitreihen- und Radardiagrammanzeige entsprechend aktualisiert.

Durch das Ändern des Anzeigedatums wird die Karte sowie das Radardiagramm aktualisiert.

Bei Auswahl eines Anzeigedatums, das in der Zukunft liegt, werden in der Karte sowie im Radardiagramm das 50%-Quantil (Median) der Vorhersage angezeigt. In der Zeitreihenanzeige wird die Spanne des Vorhersageensembles zwischen den 10% und 90%-Quantilen dargestellt.

Das Dashboard enthält zwei unterschiedliche Ansichten, eines für die Indizes für ganz Thüringen (auf Basis von J2000g) sowie eines für die Indizes im Pilotgebiet (auf Basis von Talsim). Weiterhin ist eine Info-Seite hinterlegt mit Beschreibungen und Nutzungshinweisen.

Die technische Umsetzung des Dashboards basiert auf dem Framework Dash⁸ und es wird derzeit auf der Plattform Heroku⁹ gehostet.

Folgende Abbildung 6 zeigt einen Screenshot des Dashboards.

⁸ <https://dash.plotly.com/>

⁹ <https://www.heroku.com/>

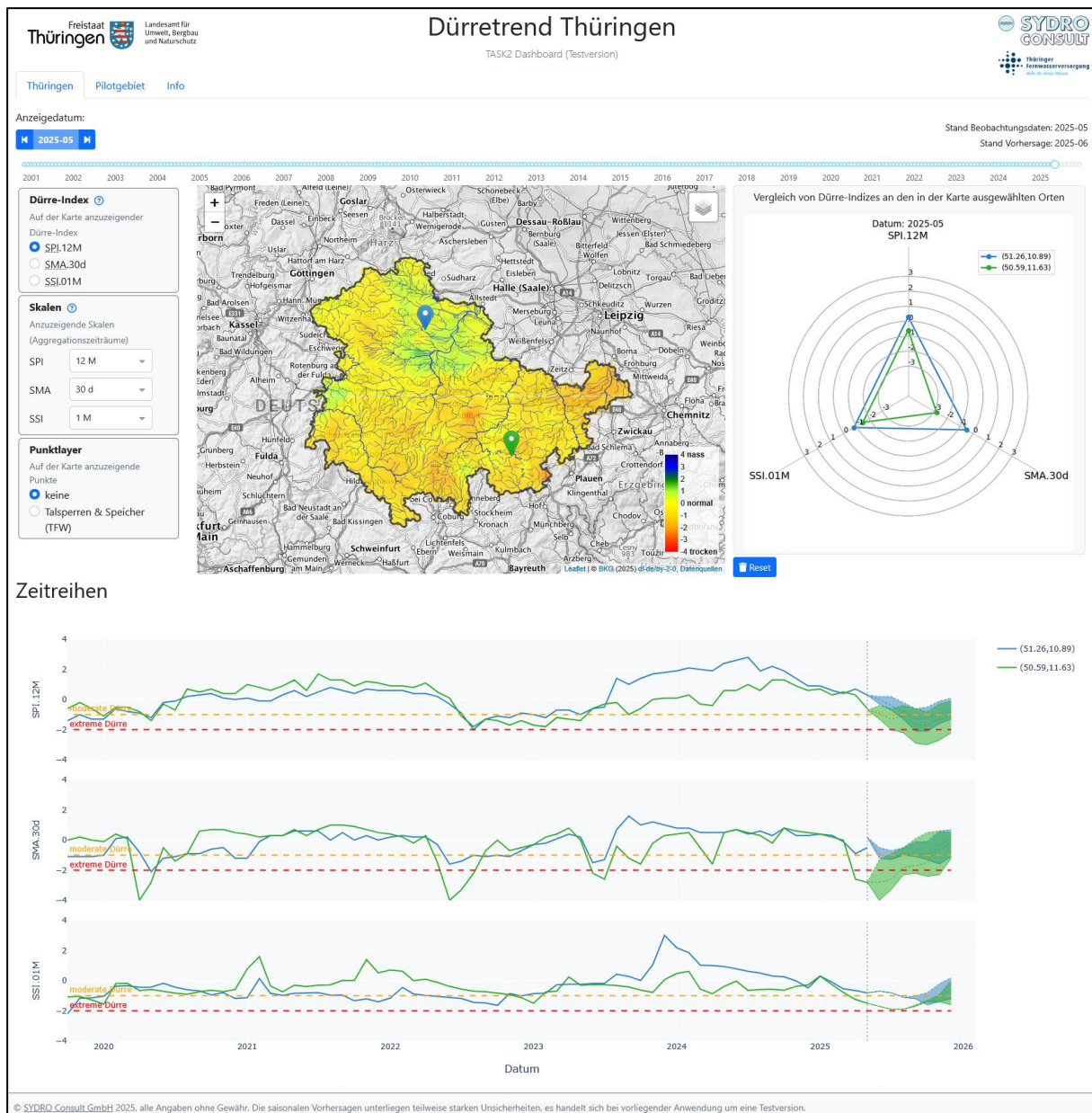


Abbildung 6: Screenshot des Dashboards (Stand Mai 2025)

3 ANALYSE VORHERSAGEN

In den folgenden Kapiteln wird eine detaillierte Analyse der Vorhersagen der Indizes vorgestellt. Dafür wurden insgesamt 39 Vorhersagen von Januar 2023 – März 2026 analysiert. Der Fokus liegt dabei auf der Analyse des Index SPI. Da die Umstellung des German Climate Forecast Systems (GCFS) auf Version 2.2, einschließlich der Verlängerung des Vorhersagezeitraums auf sieben Monate, während des Untersuchungszeitraums erfolgte, werden im Folgenden sowohl Version 2.1 als auch Version 2.2 berücksichtigt.

Die Erkenntnisse aus den Analysen der Vorhersagen sollen zudem als Grundlage für die Berechnung der Vorhersagegüte aktueller Prognosen dienen sowie für deren Darstellung im Dashboard genutzt werden.

Für einen besseren raum-zeitlichen Vergleich wurden Naturräume in Thüringen herangezogen. Dafür wurde der DWD-Datensatz „Klimatische Naturräume“¹⁰ für Thüringen adaptiert und zu neun Thüringer Naturräumen zusammengefasst, die in Abbildung 7 dargestellt sind.

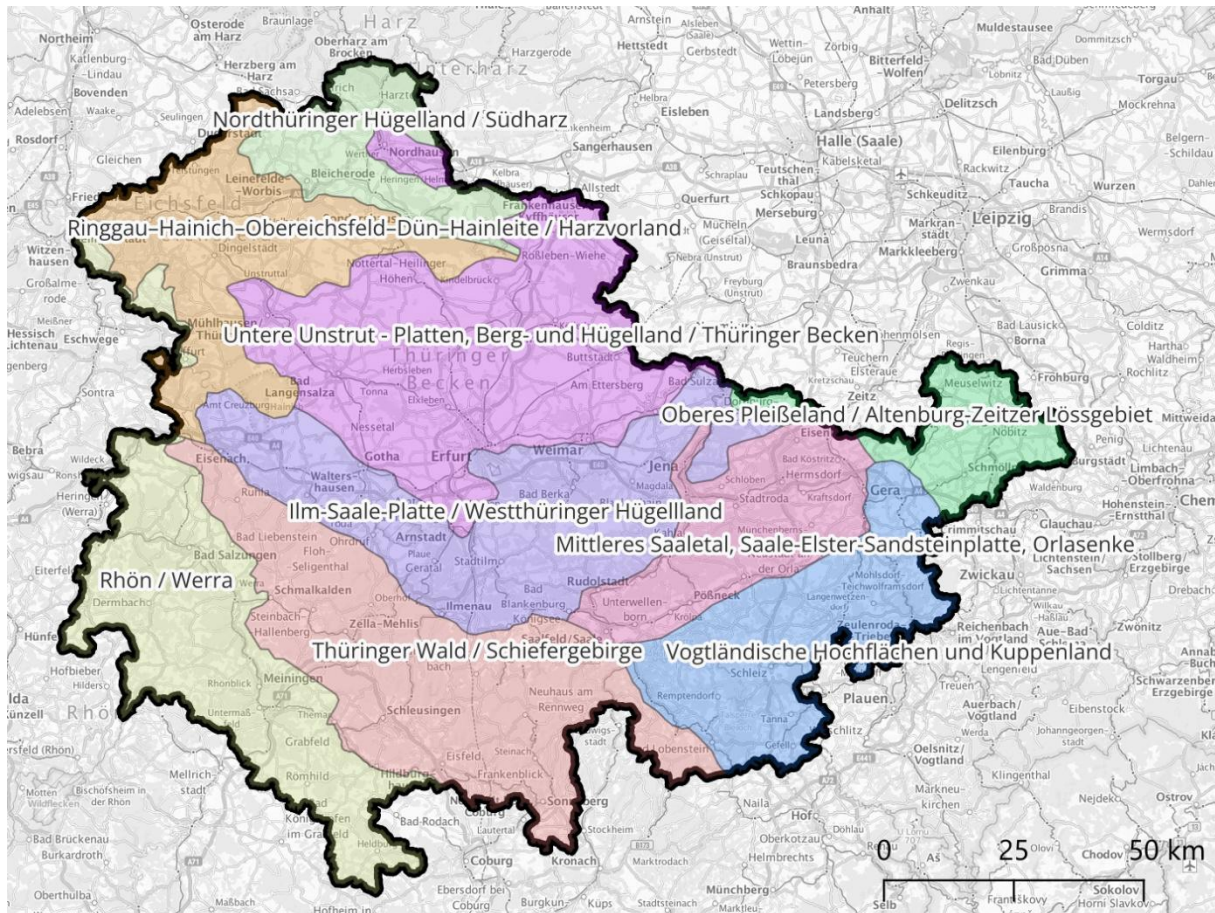


Abbildung 7: Karte adaptierte klimatische Naturräume Thüringen

3.1 Methodik

Für die Entwicklung der Methodik wurde primär die Publikation „Guidelines on the Verification of Hydrological Forecasts“¹¹ der WMO herangezogen. Im Folgenden werden Analysen hinsichtlich verschiedener Aspekte betrachtet:

- **Genauigkeit (Accuracy):** Die Genauigkeit (Accuracy) beschreibt, wie stark eine Vorhersage von den Beobachtungen abweicht.
- **Verlässlichkeit (Reliability):** Die Verlässlichkeit (Reliability) zeigt, ob die vorhergesagte Eintrittswahrscheinlichkeit eines Ereignisses mit dessen tatsächlich beobachteter Häufigkeit übereinstimmt. Sie beantwortet also die Frage, ob z.B. bei einer prognostizierten Wahrscheinlichkeit von 20% für die Unterschreitung eines Schwellwerts, dieser Schwellwert auch tatsächlich in 20% der Fälle unterschritten wurde. Die Vorhersage gilt als verlässlich, wenn die vorhergesagten Wahrscheinlichkeiten im Mittel mit der beobachteten Häufigkeit übereinstimmen.

¹⁰ Rybka et al. (2025): Klimatische Naturräume in Deutschland – Methodik und Dokumentation, Stand Oktober, 2025. Online unter: https://www.dwd.de/DE/leistungen/klimaprojektionsdaten/pdf/natur-raeume.pdf?__blob=publicationFile&v=1

¹¹ World Meteorological Organization (2025): Guidelines on the Verification of Hydrological Forecasts. WMO-No. 1364. Genf. Online verfügbar unter: https://library.wmo.int/viewer/69478/download?file=WMO-1364-2025_en.pdf&type=pdf&navigator=1

- **Trennschärfe (Discrimination):** Die Trennschärfe (Discrimination) gibt an, wie gut das Modell zwischen dem Eintreten eines Ereignisses und dem Nicht-Eintreten des Ereignisses unterscheiden kann.
- **Vorhersagegüte (Skill):** Der Skill misst, wie gut das Vorhersagemodell im Vergleich zu einer Referenz abschneidet.

Verwendete Begriffe:

- **Vorhersagehorizont:** Bezeichnet den Zeitraum, für den die Vorhersage gemacht wird. Wird die Vorhersage z.B. im Oktober 2023 erstellt, entspricht ein Vorhersagehorizont von einem Monat dem Oktober 2023, ein Vorhersagehorizont von zwei Monaten dem November 2023 usw.
- **Zieljahr:** Das Jahr für das die Vorhersage gilt. Wird eine Vorhersage z.B. im Oktober 2023 erstellt, ist das Zieljahr für die Vorhersagehorizonte 1, 2 und 3 Monate das Jahr 2023, und für die Horizonte 4, 5 und 6 Monate das Jahr 2024.

3.2 Qualitative Analyse Vorhersagen

Abbildung 8 zeigt sämtliche SPI-Vorhersagen für das Gebiet Thüringer Wald / Schiefergebirge für den Aggregationszeitraum (Skala) von sechs Monaten, während Abbildung 9 analog dazu jene Vorhersagen für das Gebiet Ringgau / Harzvorland darstellt. Hierbei werden teils Unterschiede bei der Entwicklung des SPI und bei der Treffersicherheit der Vorhersagen sichtbar.

Im Gebiet Thüringer Wald / Schiefergebirge fiel die Feuchtperiode im Vergleich zum Gebiet Ringgau / Harzvorland zu Beginn des Jahres 2024 weniger ausgeprägt aus; zugleich war die Treffersicherheit der Vorhersagen höher. Auf der anderen Seite waren die Vorhersagen im Frühjahr 2025 – insbesondere bei längeren Vorhersagehorizonten – deutlich zu feucht. Hier zeigte sich ebenfalls abweichend zur Beobachtung ab ca. 10/2024 mit jeder Vorhersage ein grundsätzlich ansteigender Index. Erst mit der Vorhersage 06/2025, mit Abstrichen 05/2025 zeigte sich wieder eine Zunahme des Index entsprechend der Beobachtung. Der ab 10/2025 neuerlich absinkende Index-Wert aus der Beobachtung wird von der Vorhersage nicht bzw. erst spät, ab 12/2025 hinsichtlich der Dynamik erkannt.

Ebenso war es im Frühjahr 2025 im Thüringer Wald deutlich trockener als im Gebiet Ringgau / Harzvorland. Für letzteres konnte im gleichen Zeitraum eine höhere Treffersicherheit festgestellt werden. Jedoch ist auch hier festzustellen, dass die Vorhersagen für Winter 2023/Frühjahr 2024 unterhalb der Beobachtung lagen und die Vorhersagen im Frühjahr 2025 und 2026 oberhalb der Beobachtung lagen.

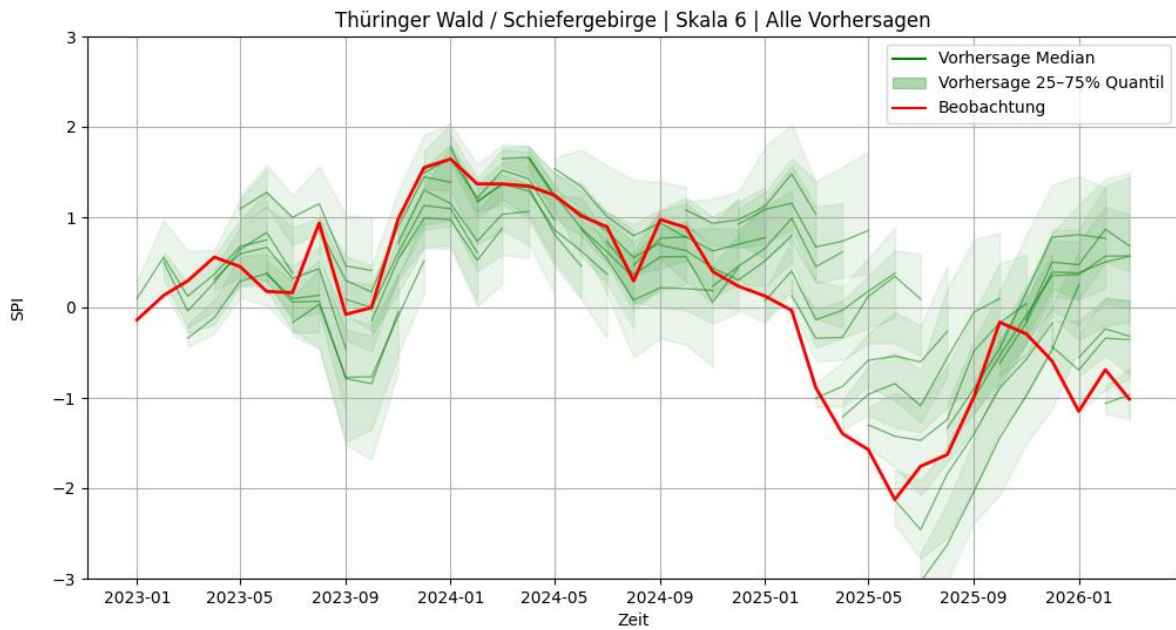


Abbildung 8: Thüringer Wald / Schiefergebirge (Skala 6 Monate) - alle Vorhersagen SPI

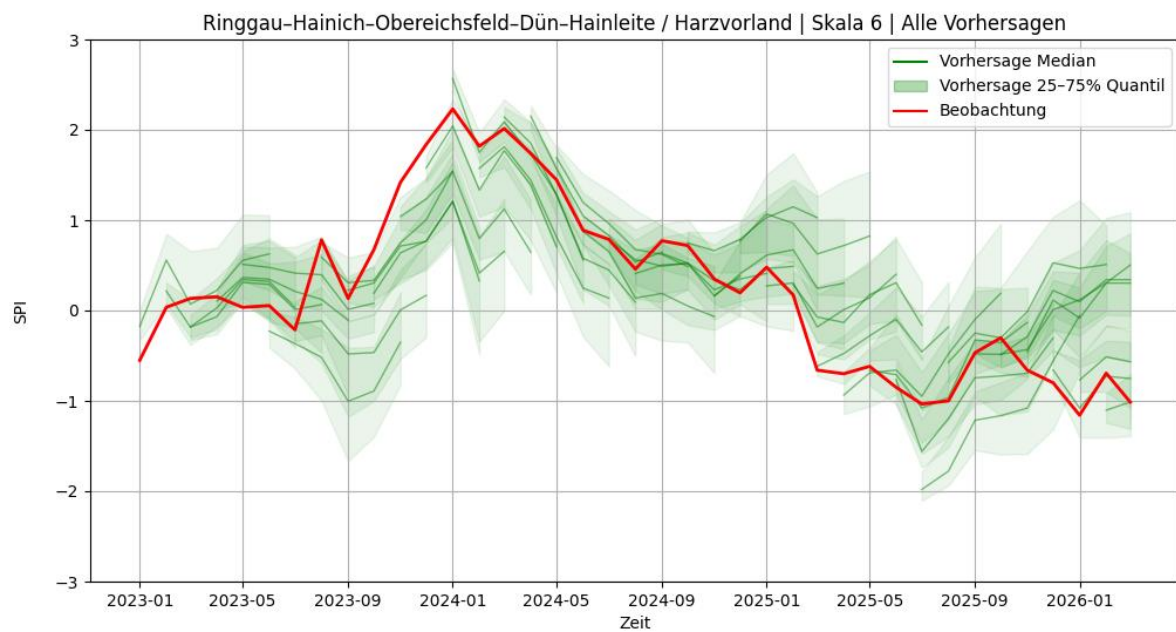


Abbildung 9: Gebiet Ringgau / Harzvorland (Skala 6 Monate) - alle Vorhersagen SPI

In Abbildung 10 und Abbildung 11 sind die SPI-Vorhersagen für einen Vorhersagehorizont von zwei Monaten ersichtlich. Im Gebiet Thüringer Wald / Schiefergebirge konnten bei diesem Vorhersagehorizont einzelne Vorhersagen 2023 und vor allem die Vorhersage für das Frühjahr 2025 die Messdaten teilweise nicht besonders gut treffen. Im Gebiet Unstrut / Thüringer Becken waren vor allem die Vorhersagen für das Frühjahr 2025 im Vergleich deutlich besser.

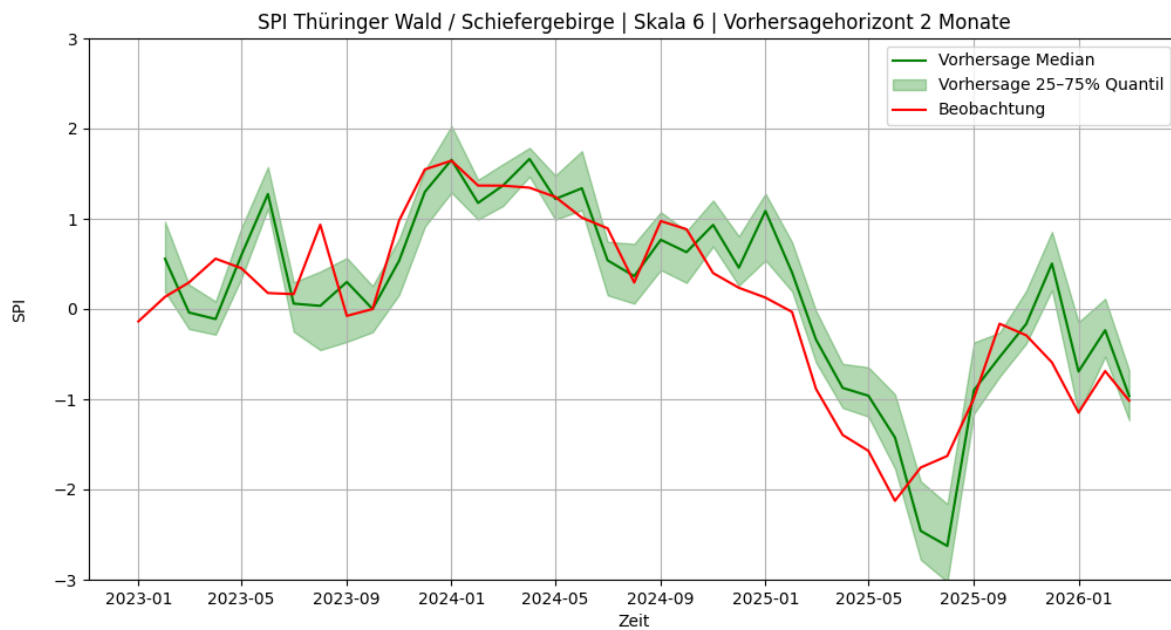


Abbildung 10: Thüringer Wald / Schiefergebirge (Skala 6 Monate) - Vorhersagehorizont 2 Monate SPI

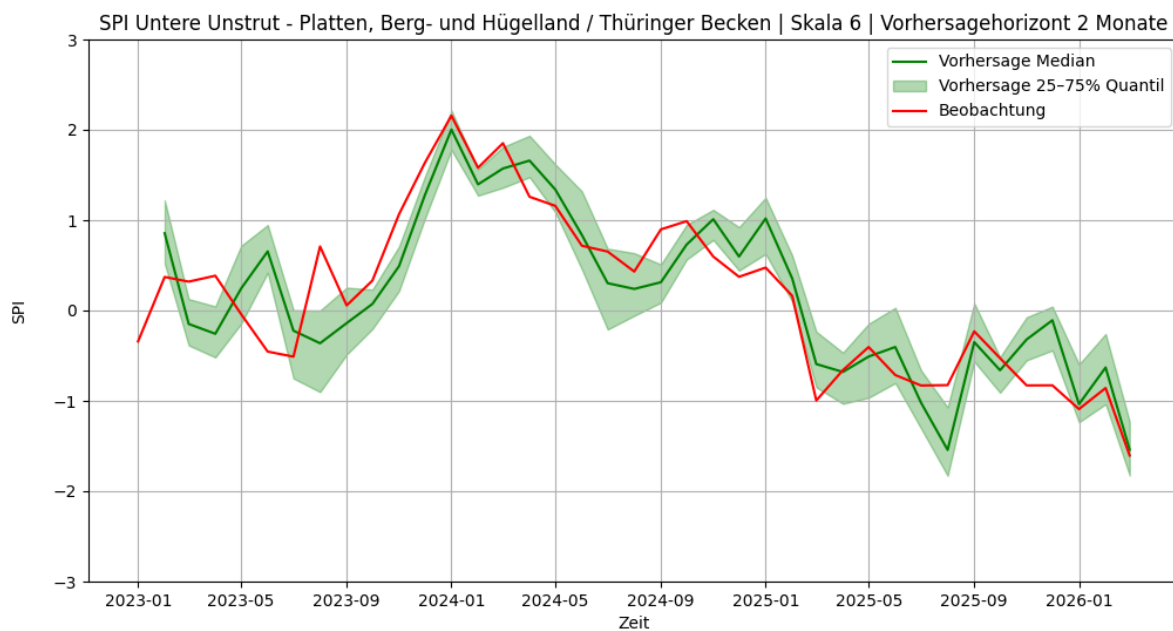


Abbildung 11: Unstrut / Thüringer Becken (Skala 6 Monate) - Vorhersagehorizont 2 Monate SPI

Die beiden Abbildungen Abbildung 12 und Abbildung 13 zeigen die SPI-Vorhersagen für einen Vorhersagehorizont von vier Monaten. Die beiden Grafiken bestätigen das zuvor beobachtete Muster: Im Gebiet Thüringer Wald / Schiefergebirge funktionierten die Vorhersagen im Winter 2024 besser, während sie im Frühjahr 2025 – insbesondere im Vergleich zu anderen Regionen, vor allem im Norden – eine geringere Güte aufwiesen. Die Unsicherheiten und Fehler in den Vorhersagen sind dabei erwartungsgemäß stärker ausgeprägt als beim Vorhersagehorizont von zwei Monaten.

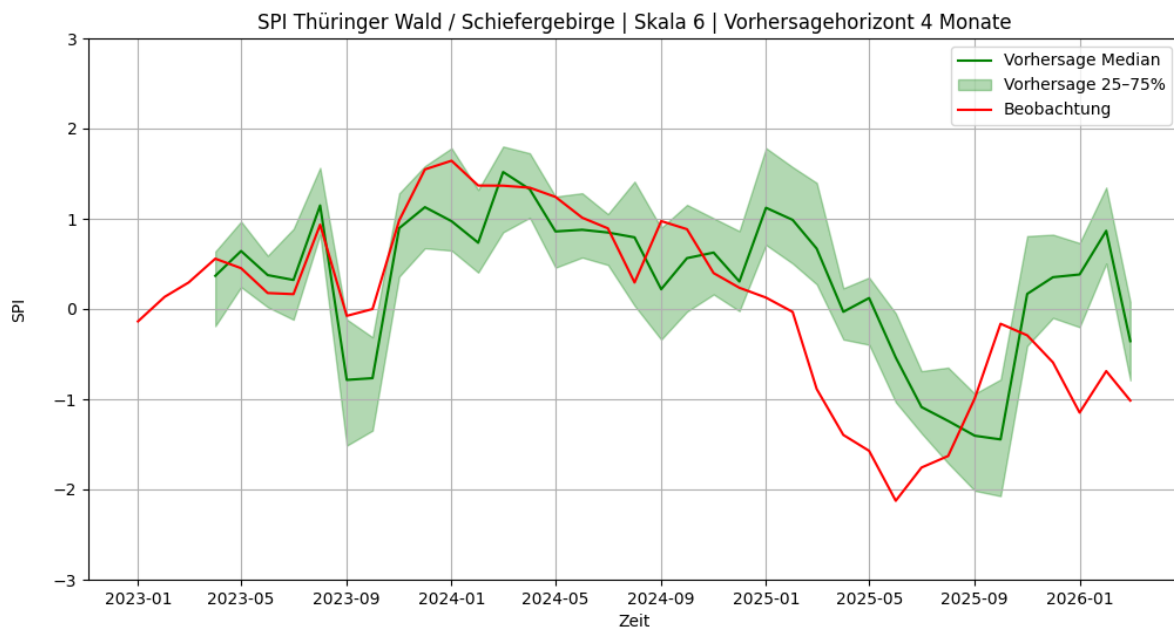


Abbildung 12: Thüringer Wald / Schiefergebirge (Skala 6 Monate) - Vorhersagehorizont 4 Monate SPI

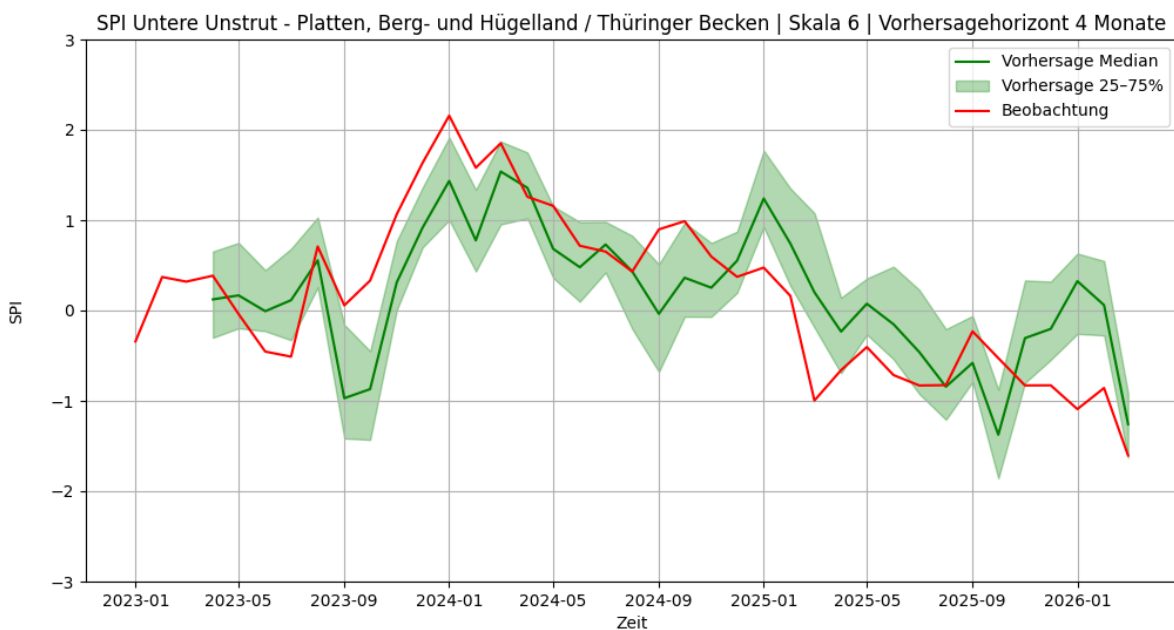


Abbildung 13: Unstrut / Thüringer Becken (Skala 6 Monate) - Vorhersagehorizont 4 Monate SPI

3.3 Quantitative Analyse

Um die Erkenntnisse aus der qualitativen Analyse zu bestätigen und noch genauer zu untersuchen, ist eine quantitative Analyse der Vorhersagen anhand der zuvor beschriebenen Metriken erforderlich. Diese erfolgt in den folgenden Kapiteln.

3.3.1 Genauigkeit / Accuracy

Das für probabilistische Vorhersagen am häufigsten genutzte Gütemaß ist der „continuous ranked probability score“ (CRPS). Der CRPS misst den Abstand zwischen der vorhergesagten Wahrscheinlichkeitsverteilung (z. B. aus einem Ensemble) und dem tatsächlich beobachteten Wert. Der CRPS kann als Pendant zu dem für deterministische Vorhersagen genutzten mittleren absoluten Fehlers (MAE) betrachtet werden. Ein niedrigerer CRPS-Wert weist auf eine bessere Vorhersage hin. Der CRPS nimmt die Einheit des analysierten Werts an, welcher in diesem Fall der SPI ist. Dabei zeigt der CRPS den Fehler der gesamten Verteilung an und ist wie folgt definiert:

$$CRPS(F, y) = \int_{-\infty}^{\infty} (F(x) - 1(x \geq y))^2 dx$$

wobei $F(x)$ die kumulative Verteilungsfunktion (CDF) der Vorhersage und y den tatsächlich beobachtete SPI darstellt. Der CRPS lässt sich für den SPI so interpretieren, dass ein CRPS-Wert von z.B. 0,5 angibt, dass die vorhergesagte Wahrscheinlichkeitsverteilung im Mittel um 0,5 SPI-Einheiten von der idealen Verteilung abweicht.

In Abbildung 14 sind die CRPS-Werte des SPI für die Skalen von drei, sechs und zwölf Monaten über alle Vorhersagehorizonte hinweg dargestellt. Dabei wird deutlich, dass der SPI.12M die niedrigsten und SPI.3M die höchsten CRPS für alle Vorhersagehorizonte aufweist. Dies ist darauf zurückzuführen, dass bei längeren Skalen mehr historische Monate in die Berechnung einfließen als bei kürzeren Zeiträumen. Aufgrund der Korrelation dieser Ergebnisse, wird im Folgenden auf einen direkten Vergleich unterschiedlicher Skalen verzichtet.

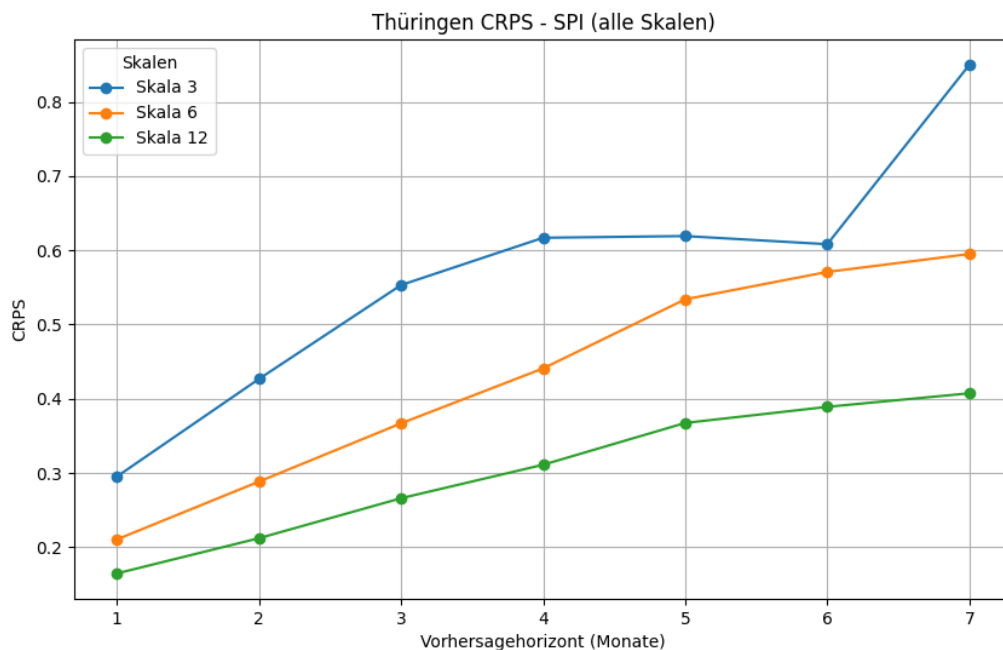
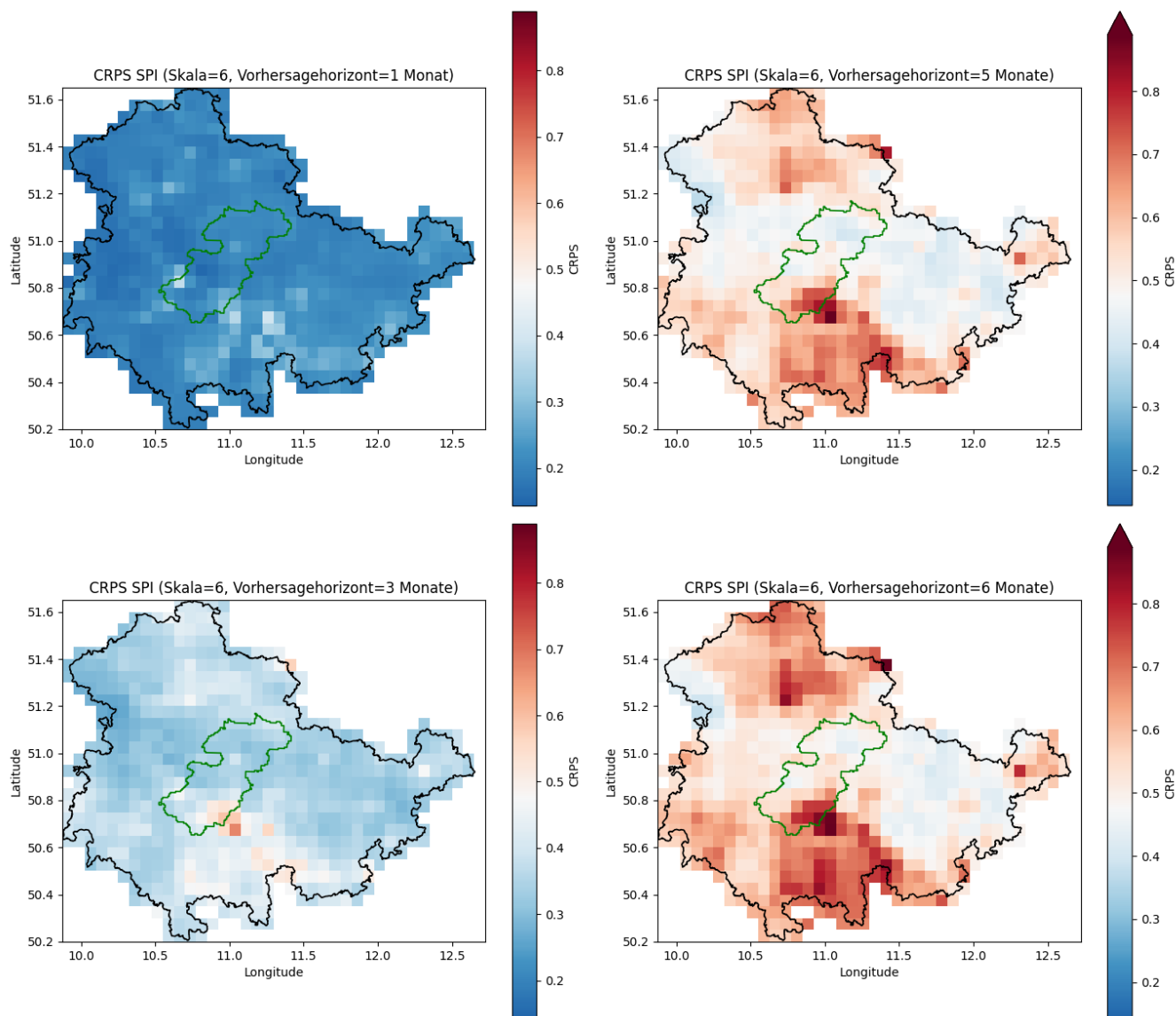


Abbildung 14: Thüringen CRPS - SPI alle Skalen

Abbildung 16 zeigt vier Karten mit CRPS-Werten für ganz Thüringen für den SPI.6M und die Vorhersagehorizonte von einem, drei, fünf und sechs Monaten. Erwartungsgemäß verschlechtert sich die Vorhersagequalität mit zunehmendem Vorhersagehorizont. Gleichzeitig werden räumliche Unterschiede sichtbar: Im Bereich des Thüringer Walds sowie in Teilen des Südwestens, Südostens und Nordens Thüringens stimmen die Vorhersagen weniger gut mit den Beobachtungswerten überein als in anderen Regionen.



In Abbildung 16 sind die CRPS-Ergebnisse der verschiedenen Regionen nach Zielmonat ersichtlich. In
 Abbildung 15: Thüringen CRPS - Karten SPI.6M (Vorhersagehorizonte 1, 3, 5 und 6 Monate)

der Grafik sind die Vorhersagehorizont zwei, vier und sechs Monate für den SPI.6M zu sehen. Dabei werden deutliche räumliche und zeitliche Unterschiede sichtbar. Die Vorhersage für Juni 2023 funktionierte beispielsweise bei einem Vorhersagehorizont von zwei Monaten im Thüringer Wald sowie im Innerthüringer Ackerland, auf der Saale-Platte und in der Region Rhön / Meiningen weniger gut. In Nordthüringen konnten hingegen bessere Ergebnisse erzielt werden. Wie bereits in der qualitativen Analyse beschrieben, funktionierten die Vorhersagen im Winter 2023 / Frühjahr 2024 lediglich im Nordthüringer Hügelland und Ringgau / Harzvorland schlechter, während in den übrigen Regionen bessere Ergebnisse erzielt wurden.

Ebenso ist im Frühjahr 2025 eine räumliche Konzentration hoher CRPS-Werte bei längeren Vorhersagehorizonten erkennbar. Auch hier schnitten die Vorhersagen im Thüringer Wald sowie in der Region Rhön / Meiningen schlechter ab. Für Januar 2026 weisen Vorhersagehorizonte von mehr als vier Monaten in allen Regionen ähnlich hohe CRPS-Werte auf. Bei vielen weiteren Vorhersagen erzielten alle Regionen vergleichbar gute CRPS-Werte. Insbesondere die Vorhersagen von Januar 2024 bis in das Frühjahr 2025 lieferten über alle Regionen hinweg stabile und insgesamt gute Ergebnisse.

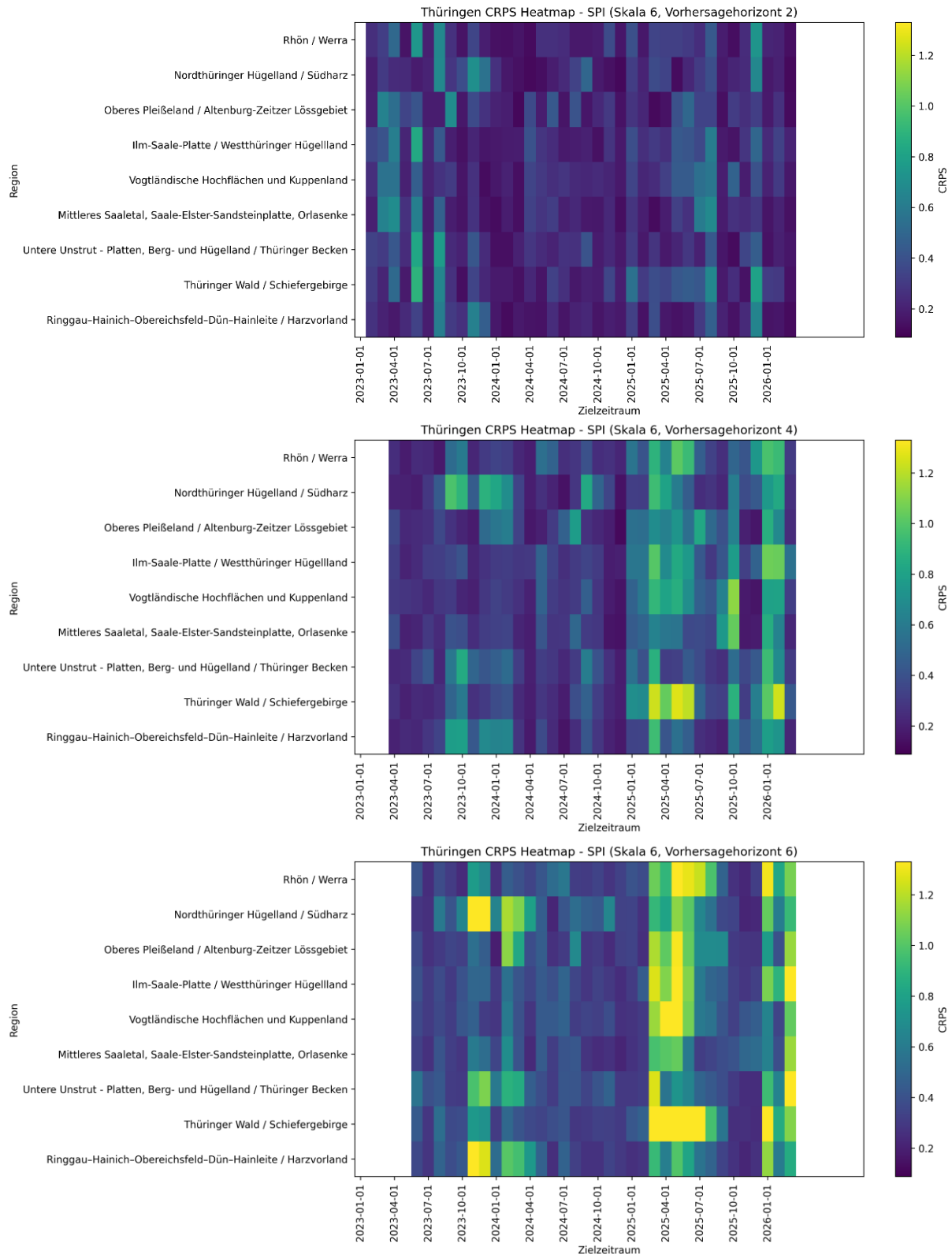


Abbildung 16: Thüringen CRPS Regionen - Heatmap nach Zielzeitraum SPI.6M (Vorhersagehorizonte 2, 4 und 6 Monate)

In Abbildung 17 sind die CRPS-Werte für den Index SMA dargestellt. Hierbei zeigt sich ein ähnliches Bild, wie beim SPI und SSI, dessen Ergebnisse in Abbildung 18 ersichtlich ist. Insbesondere der Süden Thüringens, vor allem die Gebiete Thüringer Wald / Schiefergebirge sowie Oberes Pleißeland / Altenburg-Zeitzer Lössgebiet, weist eine geringere Vorhersagequalität auf.

Dennoch zeigen sich insbesondere beim SMA Unterschiede hinsichtlich der Vorhersagehorizonte. Während sich die Vorhersagequalität des SPI aufgrund der längeren Skalen und der Kombination aus historischen Daten und Vorhersagewerten mit zunehmendem Vorhersagehorizont tendenziell verschlechtert, ist dieses Muster beim SMA nicht durchgehend erkennbar. In einzelnen Regionen kann sich die Vorhersagequalität bei längeren Vorhersagehorizonten sogar wieder verbessern, da hier historische Daten eine geringere Rolle spielen.

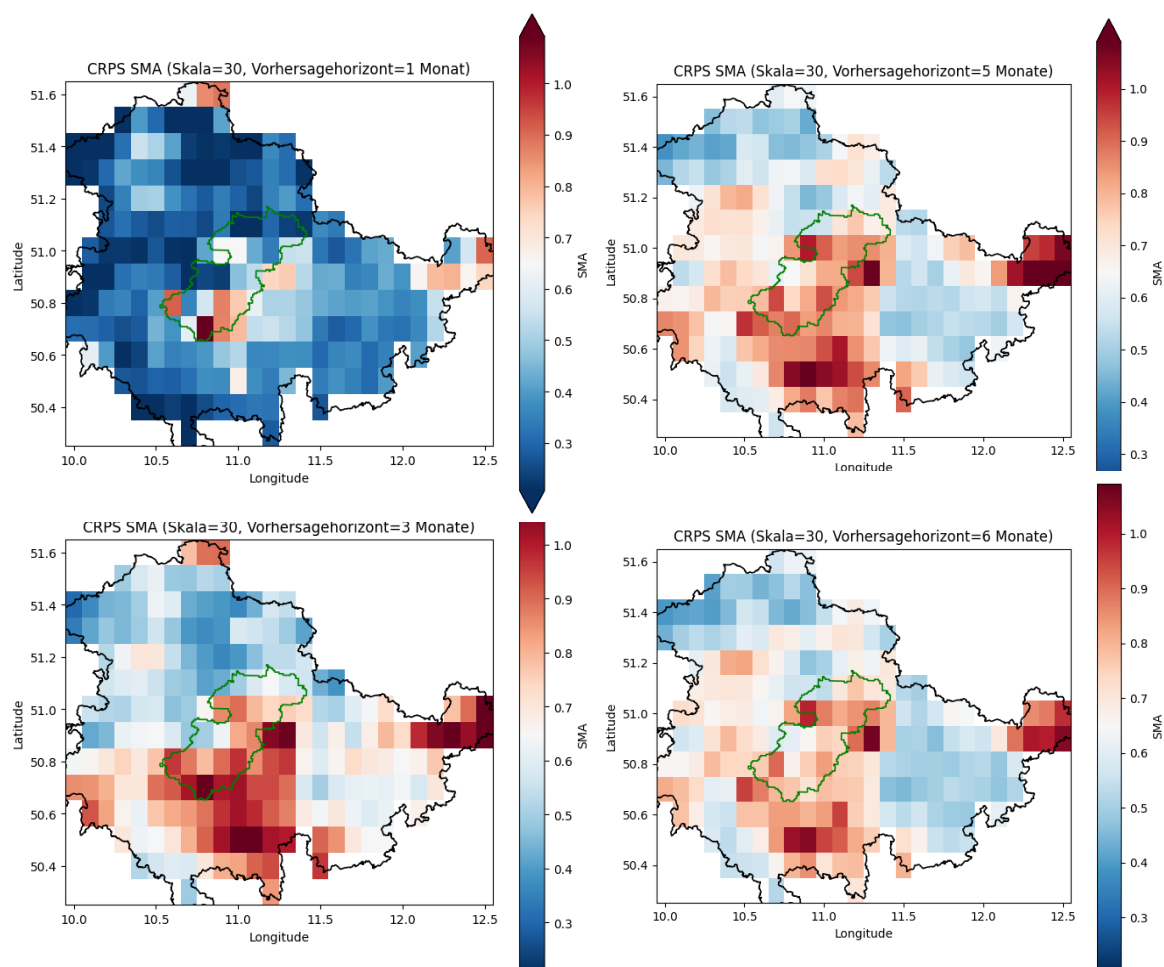


Abbildung 17: Thüringen CRPS - Karten SMA.30d (Vorhersagehorizonte 1, 3, 5 und 6 Monate)

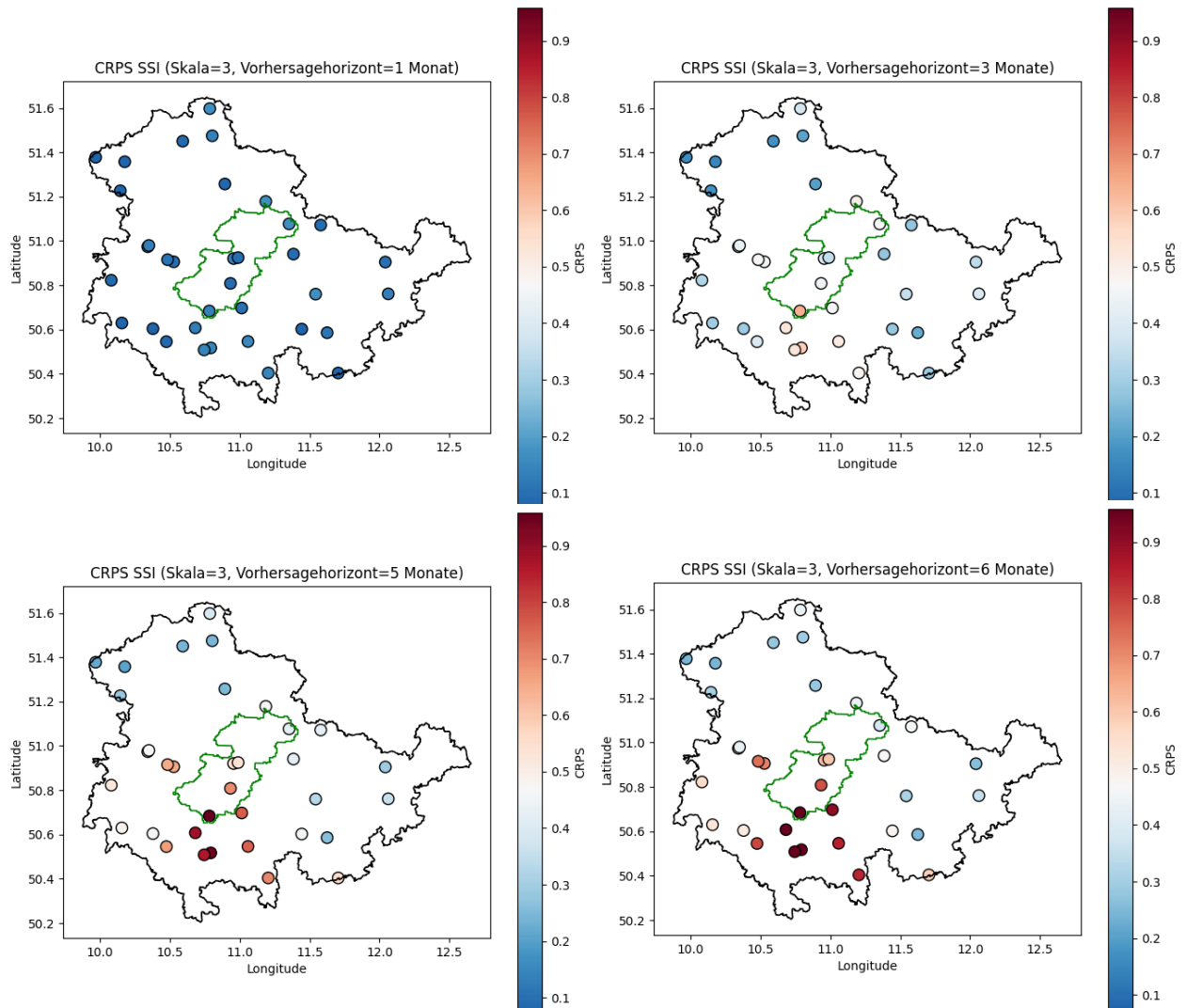


Abbildung 18: Thüringen CRPS - Karten SSI.3M (Vorhersagehorizonte 1, 3, 5 und 6 Monate)

3.3.1.1 Brier Score

Der Brier Score zählt zu den am häufigsten verwendeten Gütemaßen zur Bewertung der Vorhersagegenauigkeit hinsichtlich des Eintretens oder Nichteintretens eines Ereignisses, beispielsweise der Unter- oder Überschreitung eines definierten Schwellenwerts. Bei Ensemblevorhersagen wird dabei analysiert, wie viele der Ensemble-Mitglieder einen Schwellenwert unter- oder überschritten haben und, ob das entsprechende Ereignis (Über-/Unterschreitung des Schwellenwerts) tatsächlich eingetreten ist.

Der Brier Score kann Werte von 0 (perfekte Vorhersage) bis 1 annehmen und ist wie folgt definiert:

$$BS = \frac{1}{N} \sum_{i=1}^N (p_i - o_i)^2$$

wobei p_i die vorhergesagte Wahrscheinlichkeit und o_i die Beobachtung (0 oder 1) ist.

Ein Brier Score von 0,25 bedeutet, dass 50 % der Ensemblemitglieder korrekt eine Über- oder Unterschreitung des Schwellenwerts vorhergesagt haben. Werte unter 0,25 deuten darauf hin, dass mehr als die Hälfte der Ensemblemitglieder das Ereignis korrekt vorhergesagt haben, während Werte über 0,25 bedeuten, dass weniger als die Hälfte korrekt lag. Beispielsweise zeigt ein Brier Score von 0,1,

dass etwa 68 % der Ensemblemitglieder das Ereignis korrekt vorhergesagt haben, während ein Wert von 0,5 darauf hindeutet, dass rund 30 % der Ensemblemitglieder richtig lagen.

In Abbildung 20 ist, auf Basis der Beobachtungsdaten, die Anzahl der Schwellenwertunterschreitungen für die Schwellenwerte -2 (extreme Dürre), -1 (moderate Dürre) und 0 für jede Rasterzelle in Thüringen dargestellt. Der Schwellenwert -2 wurde lediglich in wenigen Gebieten im Süden Thüringens unterschritten. Der Schwellenwert -1 trat hingegen in ganz Thüringen auf, wurde jedoch nur im Südosten häufiger als zwölfmal unterschritten. Der Schwellenwert 0 wurde im gesamten Untersuchungsgebiet vergleichsweise häufig unterschritten. Auf Grundlage dieser Ergebnisse konzentriert sich die weitere Analyse auf die Schwellenwerte 0 und -1 . Auf eine Analyse des Schwellenwerts -2 wird aufgrund der geringen Anzahl an Unterschreitungen verzichtet.

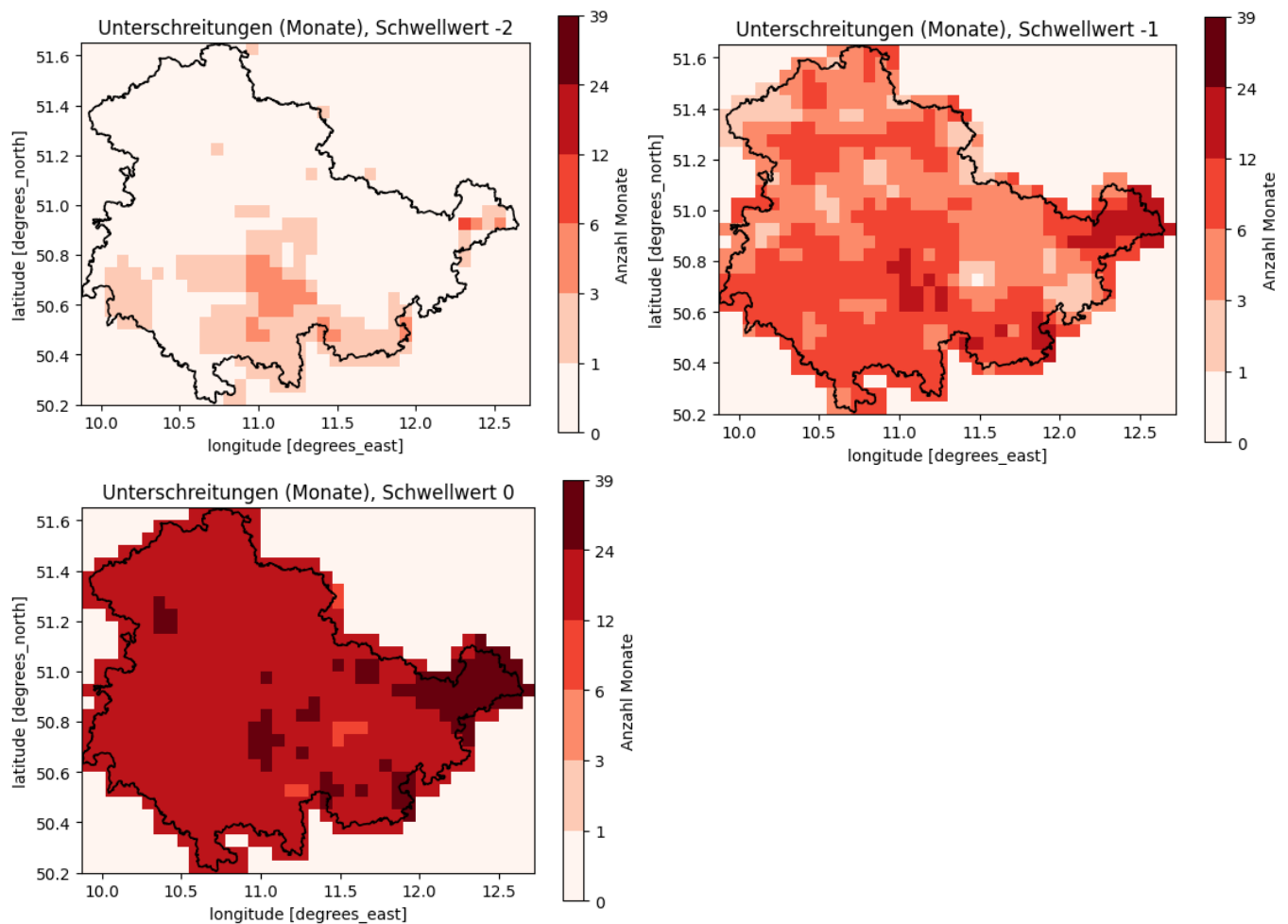


Abbildung 19: Anzahl Schwellenwertunterschreitungen Januar 2023 - März 2026

Abbildung 20 zeigt den über Thüringen gemittelten Brier Score für den Schwellenwert 0. Hierbei ist zu sehen, dass die Skalen von sechs und zwölf Monaten bei allen Vorhersagehorizonten unter einem Brier Score von 0,25 liegen (Ausnahme: Skala 6 Monate, Vorhersagehorizont sieben Monate). Bei der Skala von drei Monaten liegen hingegen nur die ersten drei Vorhersagehorizonte unter einem Brier Score von 0,25. Für Vorhersagehorizonte mit Werten unter 0,25 konnten somit mehr als 50 % der Ensemblemitglieder eine Schwellenwertunterschreitung korrekt prognostizieren.

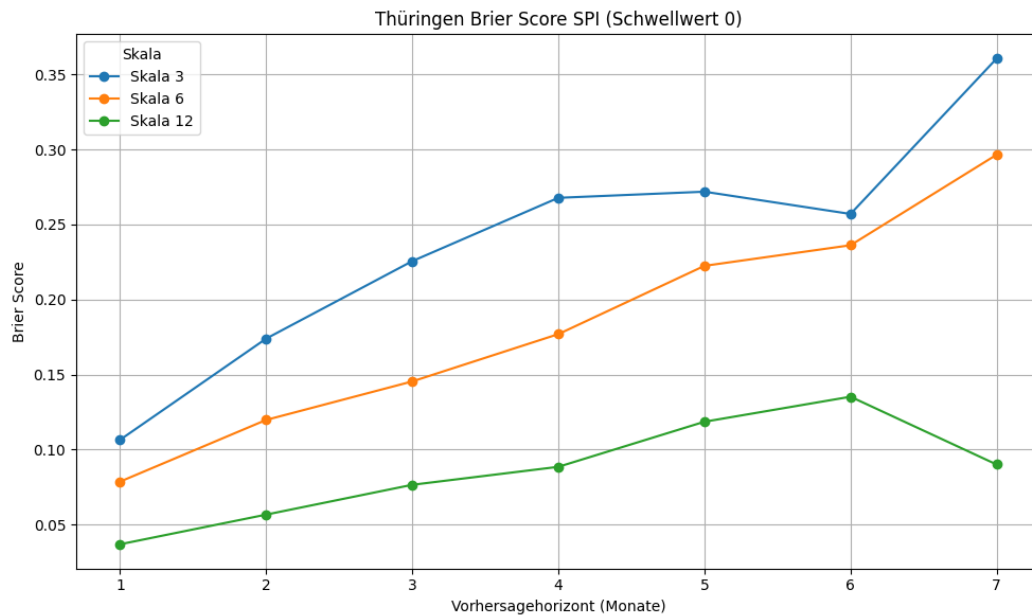


Abbildung 20: Thüringen Brier Score - Vergleich Skalen (Schwellwert 0)

Abbildung 21 zeigt den Vergleich der Brier-Score-Werte verschiedener Skalen für die Unterschreitung des Schwellenwerts -1 . Dabei zeigt sich, dass alle drei Skalen für die ersten sechs Vorhersagehorizonte eine gute Vorhersagegüte hinsichtlich des Eintretens dieses Ereignisses aufweisen. Einschränkend muss jedoch erwähnt werden, dass dieser Schwellenwert in vielen Monaten nicht unterschritten wurde. Dadurch wird die Bewertung des Brier Scores beeinflusst, da seltene Ereignisse schwieriger zu bewerten sind. Zudem zeigt die in Kapitel 3.3.2 *Verlässlichkeit (Reliability) und Trennschärfe (Discrimination)* durchgeführte Analyse, dass die Vorhersagen insbesondere bei längeren Vorhersagehorizonten dazu tendieren, Unterschreitungen extremer Werte zu unterschätzen. Der niedrige Brier-Score-Wert ist somit auch darauf zurückzuführen, dass in den meisten Monaten eine Nicht-Unterschreitung vorhergesagt wurde. Der Brier Score wird durch die hohe Zahl an Monaten ohne Schwellenwertunterschreitung somit tendenziell günstig beeinflusst.

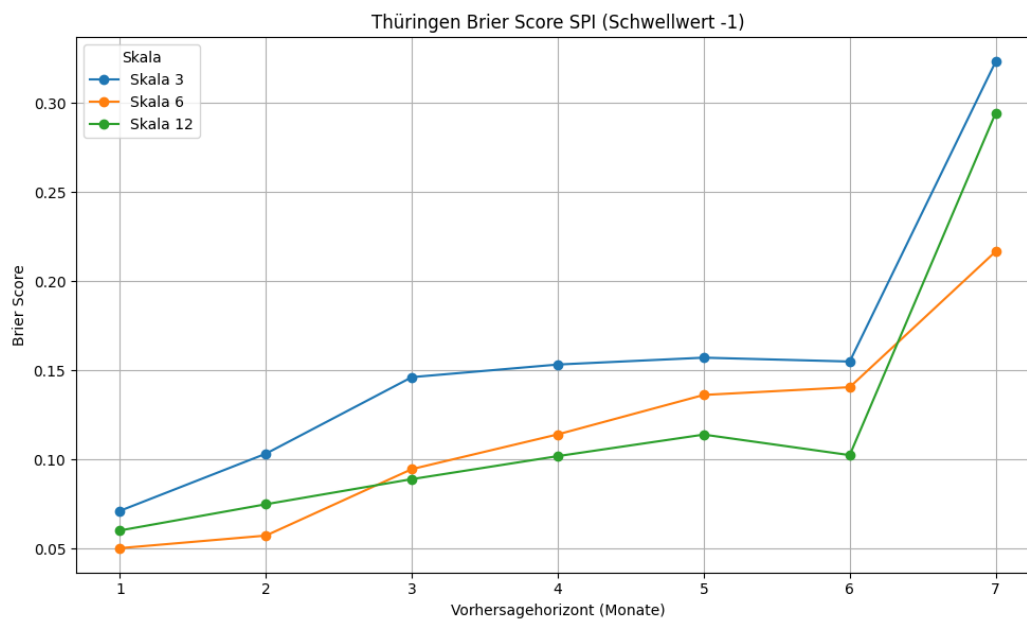


Abbildung 21: Thüringen Brier Score - Vergleich Skalen (Schwellwert -1)

In Abbildung 22 ist der Brier Score für Thüringen für den SPI.6M, den Schwellwert -1 und die Vorhersagehorizonte von einem, drei, fünf und sechs Monaten ersichtlich. Es zeigt sich ein ähnliches Bild wie bei den CRPS-Werten, wenn auch weniger stark ausgeprägt. Mit steigendem Vorhersagehorizont nimmt die Vorhersagegüte von Schwellwertunterschreitungen ab. Zudem wurden insbesondere im Südosten Thüringens sowie im Gebiet des Thüringer Walds bei höheren Vorhersagehorizonten schlechtere Ergebnisse erzielt. Im Norden Thüringens zeigen sich hingegen bessere Ergebnisse als beim CRPS. Dies könnte jedoch auf die geringere Anzahl an Schwellenwertunterschreitungen in dieser Region zurückzuführen sein.

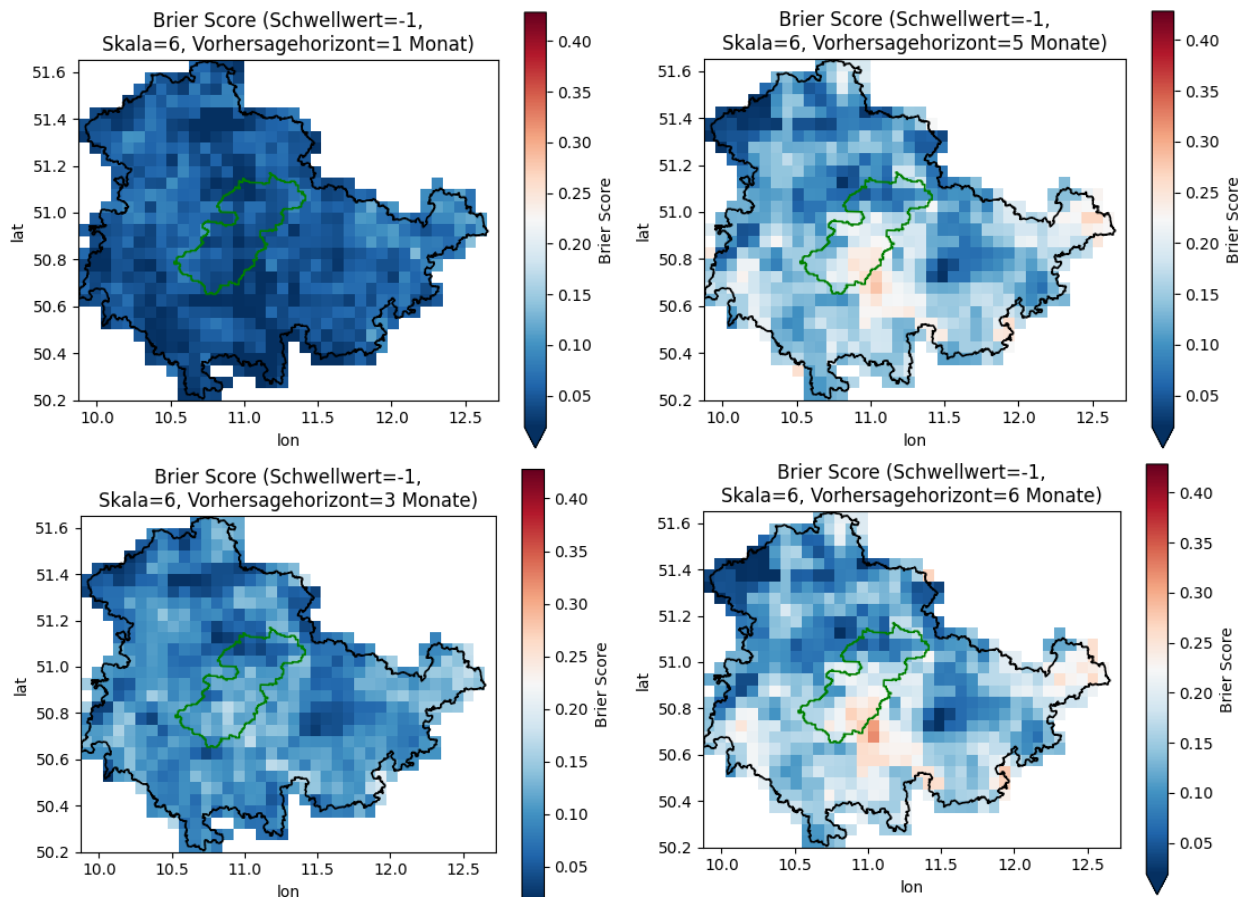


Abbildung 22: Thüringen Brier Score - Karten SPI.6M (Schwellwert -1; Vorhersagehorizonte 1, 3, 5 und 6 Monate)

3.3.2 Verlässlichkeit (Reliability) und Trennschärfe (Discrimination)

Da die Analysen der Verlässlichkeit und Trennschärfe auf die Vorhersagegüte von Schwellenwertunterschreitungen abzielen und der Schwellenwert -1 nur in wenigen Gebieten unterschritten wurde, wird in Abbildung 23 der Schwellenwert 0 betrachtet. Die x-Achse zeigt die vorhergesagte Wahrscheinlichkeit einer Schwellenwertunterschreitung in fünf Kategorien (0-20%, 20-40%, 40-60%, 60-80%, 80-100%). In den Abbildungen wird die von den Ensemblevorhersagen prognostizierte Wahrscheinlichkeit einer Schwellenwertunterschreitung mit der tatsächlichen Häufigkeit ihres Auftretens verglichen. Die Vorhersagewahrscheinlichkeit wird aus den Ensemblemitgliedern abgeleitet: Sagen beispielsweise 40 von 50 Ensemblemitgliedern eine Unterschreitung des Schwellenwerts voraus, ergibt sich eine Vorhersagewahrscheinlichkeit von 80 %.

In Abbildung 23 sind Reliability-Diagramme für das Gebiet Thüringer Wald / Schiefergebirge dargestellt. Für den Vorhersagehorizont von zwei Monaten funktionierte die Vorhersage der Schwellwertunterschreitung sehr gut. So konnte in einigen Fällen die Linie der perfekten Zuverlässigkeit annähernd erreicht werden. Auch die Vorhersagehorizonte von vier, fünf und sechs Monaten zeigen insgesamt gute Ergebnisse. Allerdings ist insbesondere bei den Vorhersagehorizonten von vier und sechs Monaten eine Unterschätzung der Schwellwertunterschreitungen zu erkennen. Dies bedeutet, dass Schwellwertunterschreitungen seltener vorhergesagt wurden, als sie tatsächlich eingetreten sind. Ein Beispiel hierfür ist die Wahrscheinlichkeitskategorie von 20–40 % bei einem Vorhersagehorizont von vier Monaten: Obwohl lediglich eine geringe Eintrittswahrscheinlichkeit prognostiziert wurde, trat das Ereignis tatsächlich in mehr als 40 % der Fälle ein.

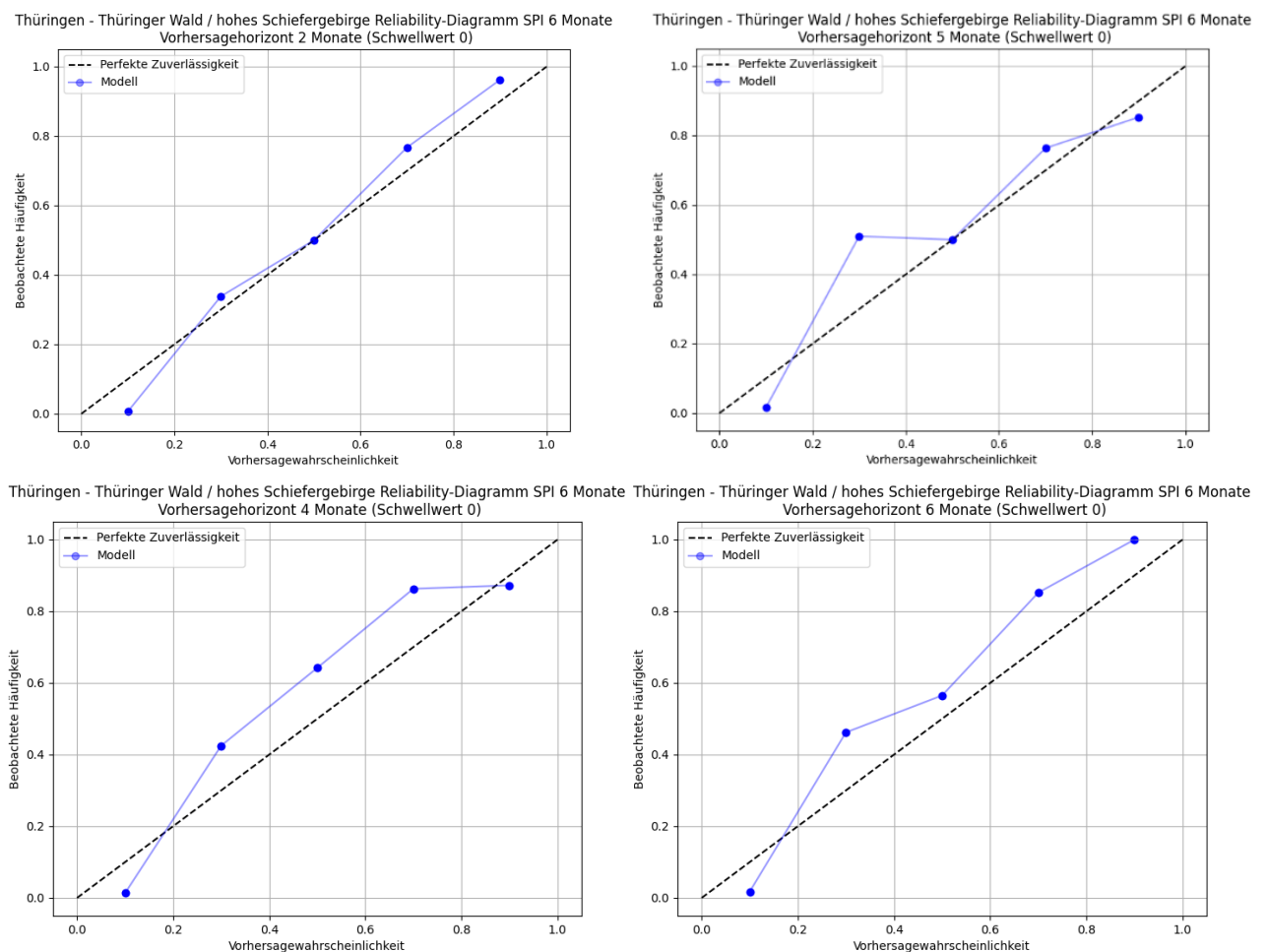


Abbildung 23: Thüringer Wald / Schiefergebirge Reliability-Diagramm - SPI.6M Schwellwert 0 (Vorhersagehorizonte 2, 4, 5 und 6 Monate)

Im Folgenden werden Discrimination-Diagramme gezeigt. Die x-Achse gibt die vorhergesagte Wahrscheinlichkeit der Schwellwertunterschreitung in fünf Kategorien an (0–20 %, 20–40 %, 40–60 %, 60–80 %, 80–100 %). Die blauen Balken repräsentieren den Anteil der beobachteten Fälle, bei denen der Schwellwert **nicht** unterschritten wurde, während die roten Balken den Anteil der Fälle zeigen, bei denen der Schwellwert tatsächlich unterschritten wurde.

Abbildung 24 – das Discrimination-Diagramm für Nordthüringen und SPI.12M – dient als Beispiel für eine hohe Trennschärfe. Der blaue Balken in der Kategorie 0–20 % umfasst alle Fälle, in denen der Schwellenwert 0 nicht unterschritten wurde. Folglich wurden in diesem Fall alle Beobachtungen, wo keine Schwellwert-Unterschreitung stattfand, korrekt mit einer niedrigen Eintrittswahrscheinlichkeit von 0 – 20 % vorhergesagt. Das Gegenstück dazu bildet der rote Balken in der Kategorie 80–100 %, der den Großteil jener Fälle zeigt, in denen der Schwellenwert tatsächlich unterschritten wurde. Rund 90 % aller Schwellenwertunterschreitungen wurden somit mit einer Ensemblewahrscheinlichkeit von 80–100 % korrekt prognostiziert. Die kleineren roten Balken in den mittleren Kategorien repräsentieren Ereignisse, bei denen der Schwellenwert zwar unterschritten wurde, die prognostizierte Eintrittswahrscheinlichkeit jedoch geringer als 80 % ausfiel.

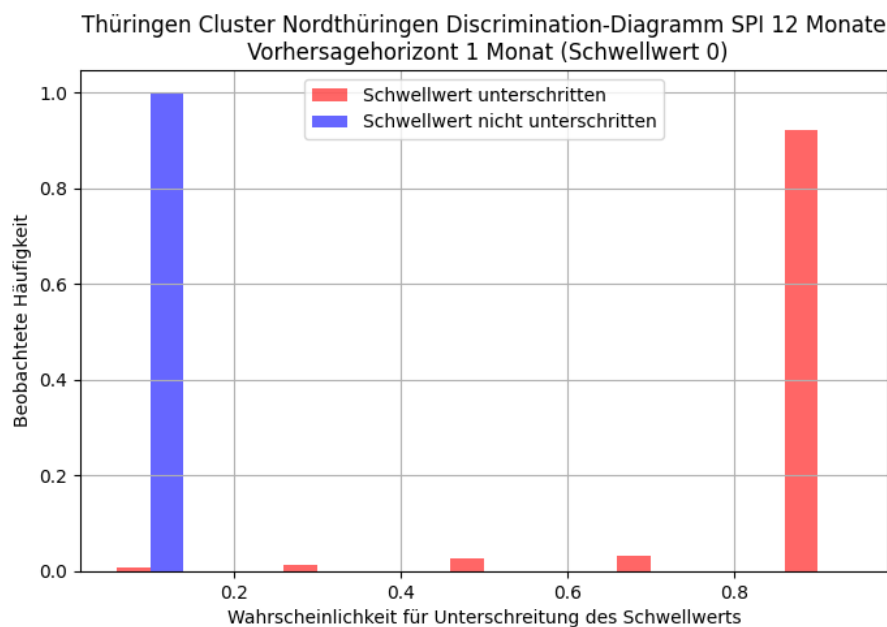


Abbildung 24: Nordthüringen Discrimination-Diagramm - SPI.12M Schwellwert 0 (Vorhersagehorizont 1 Monat)

Das Gebiet Altenburg / Ronneburg wies im Untersuchungszeitraum mehrere Unterschreitungen des Schwellenwerts -1 auf und wurde deshalb für die Discrimination-Diagramme in Abbildung 25 ausgewählt. Dabei zeigt sich ein bekanntes Phänomen von Indexvorhersagen: Insbesondere bei längeren Vorhersagehorizonten werden Schwellenwertunterschreitungen beziehungsweise Extremereignisse tendenziell unterschätzt. Die Vorhersagen nähern sich mit zunehmendem Vorhersagehorizont stärker dem klimatologischen Normalbereich an.

Während bei einem Vorhersagehorizont von einem Monat mehr als 50 % der tatsächlichen Schwellenwertunterschreitungen mit einer hohen Wahrscheinlichkeit von 80–100 % vorhergesagt wurden, treten bei Vorhersagehorizonten von fünf und sechs Monaten keine Vorhersagen mehr in Wahrscheinlichkeitskategorien über 60 % auf. Mit anderen Worten: Bei längeren Vorhersagehorizonten werden Schwellenwertunterschreitungen nur noch selten mit hoher Wahrscheinlichkeit prognostiziert, wodurch die Vorhersage insgesamt konservativer erscheint.

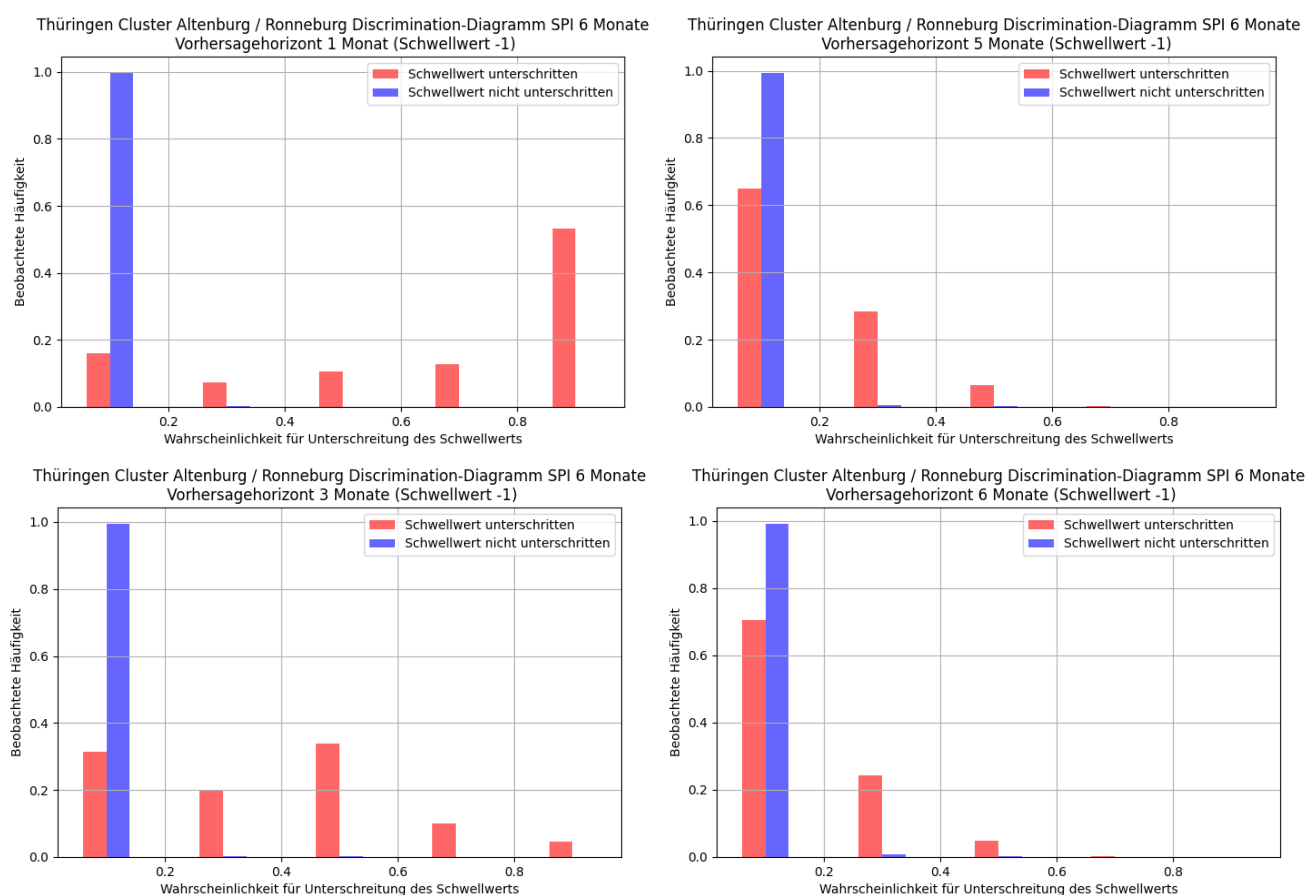


Abbildung 25: Altenburg / Ronneburg Discrimination-Diagramm - SPI.6M Schwellwert -1 (Vorhersagehorizonte 1, 3, 5 und 6 Monate)

3.3.3 Skill

Der Skill bewertet die relative Vorhersagegüte, also die Leistungsfähigkeit einer Vorhersage im Vergleich zu einer Referenz. In der vorliegenden Analyse dient eine Kombination aus Persistenz und Klimatologie – definiert als das langjährige Mittel des Zeitraums 2000–2025 – als Referenz. Bei kurzen Vorhersagehorizonten ist der Einfluss der Persistenz, also des letzten Messwerts vor Beginn der Vorhersage, besonders hoch. Mit zunehmendem Vorhersagehorizont nimmt dieser Einfluss ab, während die Bedeutung der Klimatologie steigt. Im ersten Vorhersagemonat wird ausschließlich die Persistenz als Referenz verwendet. Mit jedem zusätzlichen Vorhersagemonat sinkt ihr Gewicht um 12,5

Prozentpunkte, während der Anteil der Klimatologie entsprechend zunimmt. Bei einem Vorhersagehorizont von sieben Monaten beträgt der Anteil der Klimatologie somit 75 %, während die Persistenz nur noch einen Anteil von 25 % hat.

Positive Werte des Brier Skill Score (BSS) beziehungsweise des CRPS Skill Score (CRPSS) zeigen an, dass das Vorhersagemodell bessere Ergebnisse liefert als die Referenz. Ein Wert von 0 bedeutet, dass Vorhersagemodell und Referenz gleich gut abschneiden, während negative Werte darauf hinweisen, dass die Referenz eine höhere Vorhersagegüte aufweist. Ein BSS von 0,5 weist darauf hin, dass das Vorhersagemodell um 50 % besser abschneidet als die Referenz, während ein Wert von –0,5 darauf hinweist, dass das Vorhersagemodell um 50 % schlechter abschneidet als die Referenz. Ein Wert von 1 gibt an, dass die Vorhersage die Messdaten perfekt trifft. Ein Wert von –1 bedeutet, dass der Fehler des Vorhersagemodells doppelt so groß ist wie der der Referenz – anders als der positive Wertebereich, der bei 1 nach oben begrenzt ist, ist der BSS nach unten nicht begrenzt. Der BSS ist wie folgt definiert:

$$BSS = 1 - \frac{BS_{Vorhersage}}{BS_{Referenz}}$$

In Abbildung 26 ist der BSS für die Skala von drei Monaten, die Vorhersagehorizonte von einem, drei und fünf Monaten und Schwellwert -1 dargestellt. Blaue Bereiche kennzeichnen Regionen, in denen die Vorhersage bessere Ergebnisse erzielt als die Kombination aus Persistenz und Klimatologie, während rote Bereiche auf eine bessere Leistung der Referenz hinweisen. Bei der Skala von drei Monaten schneidet die Vorhersage in allen drei Horizonten zum Großteil besser ab als die Referenz. Lediglich beim Vorhersagehorizont von fünf Monaten gibt es größere Bereiche, wo die Referenz besser abschneidet als die Vorhersage.

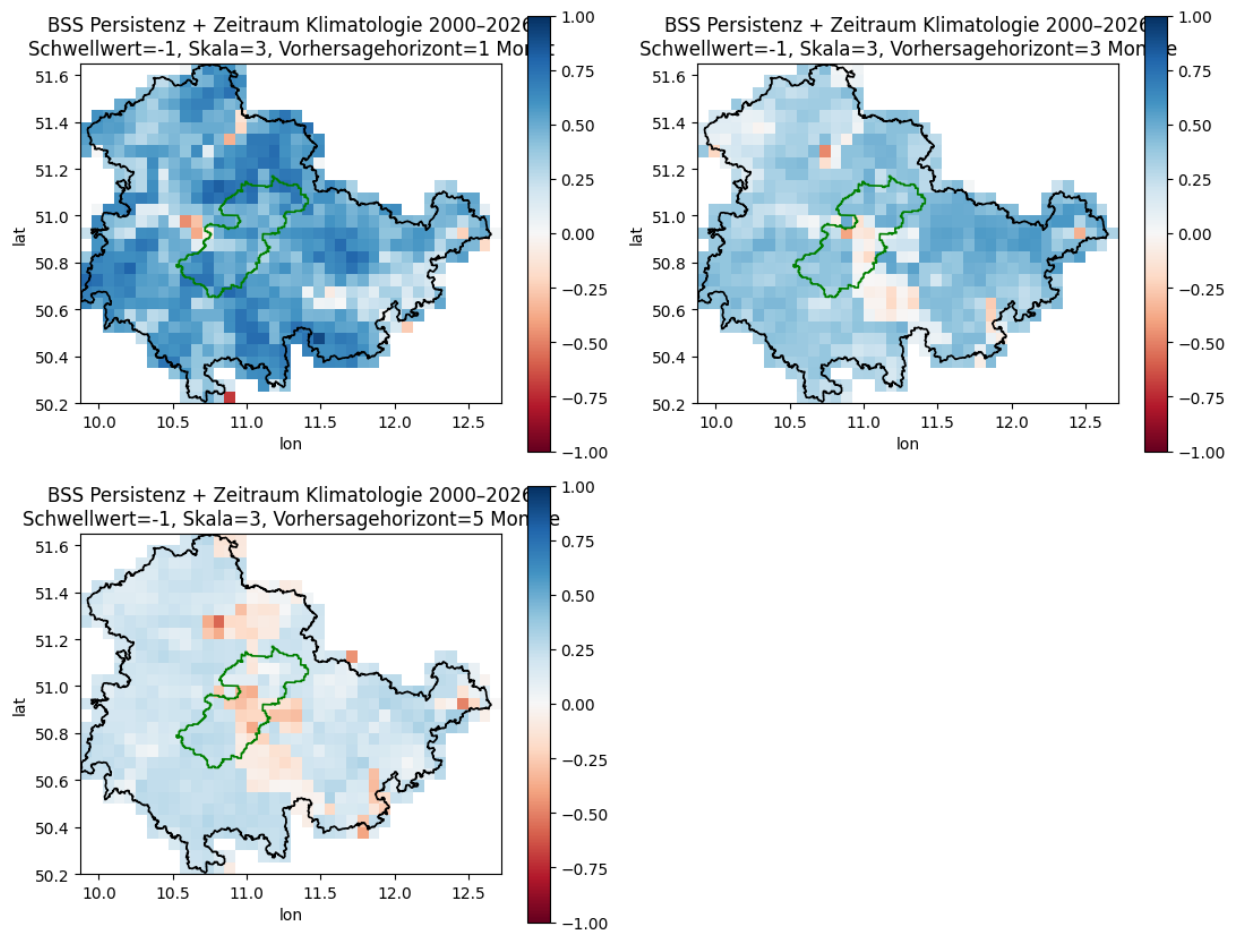


Abbildung 26: Thüringen Brier-Skill-Score - Karten SPI.3M (Schwellwert -1; Vorhersagehorizonte 1, 3 und 5 Monate)

Abbildung 27 zeigt den BSS für den Aggregationszeitraum von sechs Monaten und ansonsten die gleichen Parameter wie oben. Der Vergleich der Abbildungen verdeutlicht, dass die relative Vorhersagegüte bei einem Aggregationszeitraum von drei Monaten höher ausfällt als bei sechs Monaten. Damit ergibt sich ein anderes Bild als bei der absoluten Vorhersagegüte, bei der längere Aggregationszeiträume bessere Ergebnisse aufweisen konnten. Eine mögliche Erklärung hierfür liegt in der gewählten Referenz: Längere Aggregationszeiträume reagieren träger auf Veränderungen, wodurch die Persistenz auch für längere Vorhersagehorizonte eine vergleichsweise gute Referenz darstellt, während kürzere Aggregationszeiträume dynamischer reagieren.

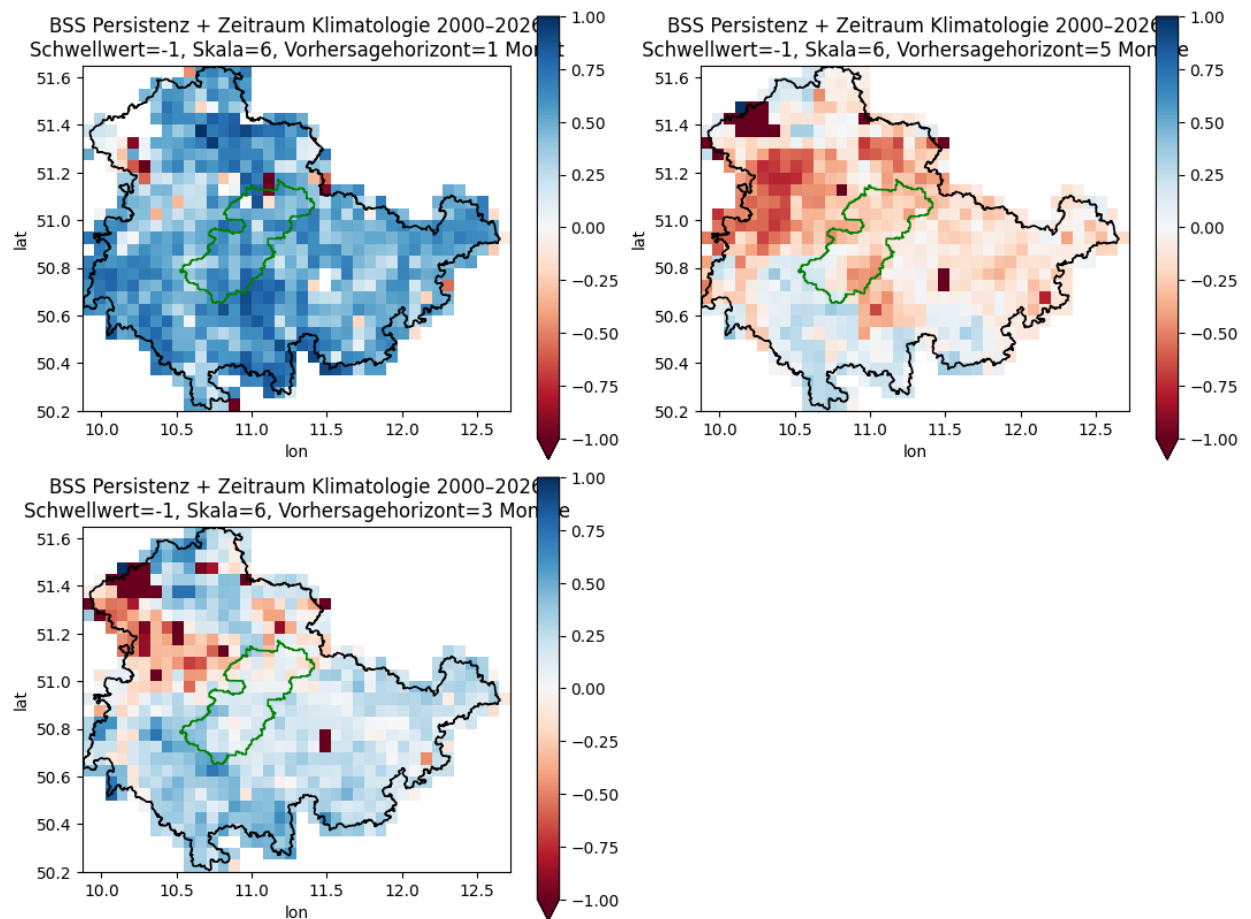


Abbildung 27: Thüringen Brier-Skill-Score - Karten SPI.6M (Schwellwert -1; Vorhersagehorizonte 1, 3 und 5 Monate)

Für den CRPS Skill Score (CRPSS) wurde dieselbe Referenz – die Kombination aus Persistenz und Klimatologie – verwendet. Der CRPSS vergleicht den CRPS-Wert der Vorhersage mit jenem der Referenz und ist wie folgt definiert:

$$CRPSS = 1 - \frac{CRPS_{Vorhersage}}{CRPS_{Referenz}}$$

Die Ergebnisse für den Aggregationszeitraum von sechs Monaten sind in Abbildung 28 dargestellt. Dabei zeigt sich, dass die Vorhersage in allen betrachteten Vorhersagehorizonten bessere Ergebnisse erzielt als die Referenz. Somit ergibt sich ein anderes Bild als beim Brier Skill Score. Während die Referenz bei der Vorhersage von Schwellwertunter- beziehungsweise -überschreitungen teilweise bessere Ergebnisse liefern kann, weist die Vorhersage hinsichtlich der allgemeinen relativen Vorhersagegüte insgesamt bessere Ergebnisse auf.

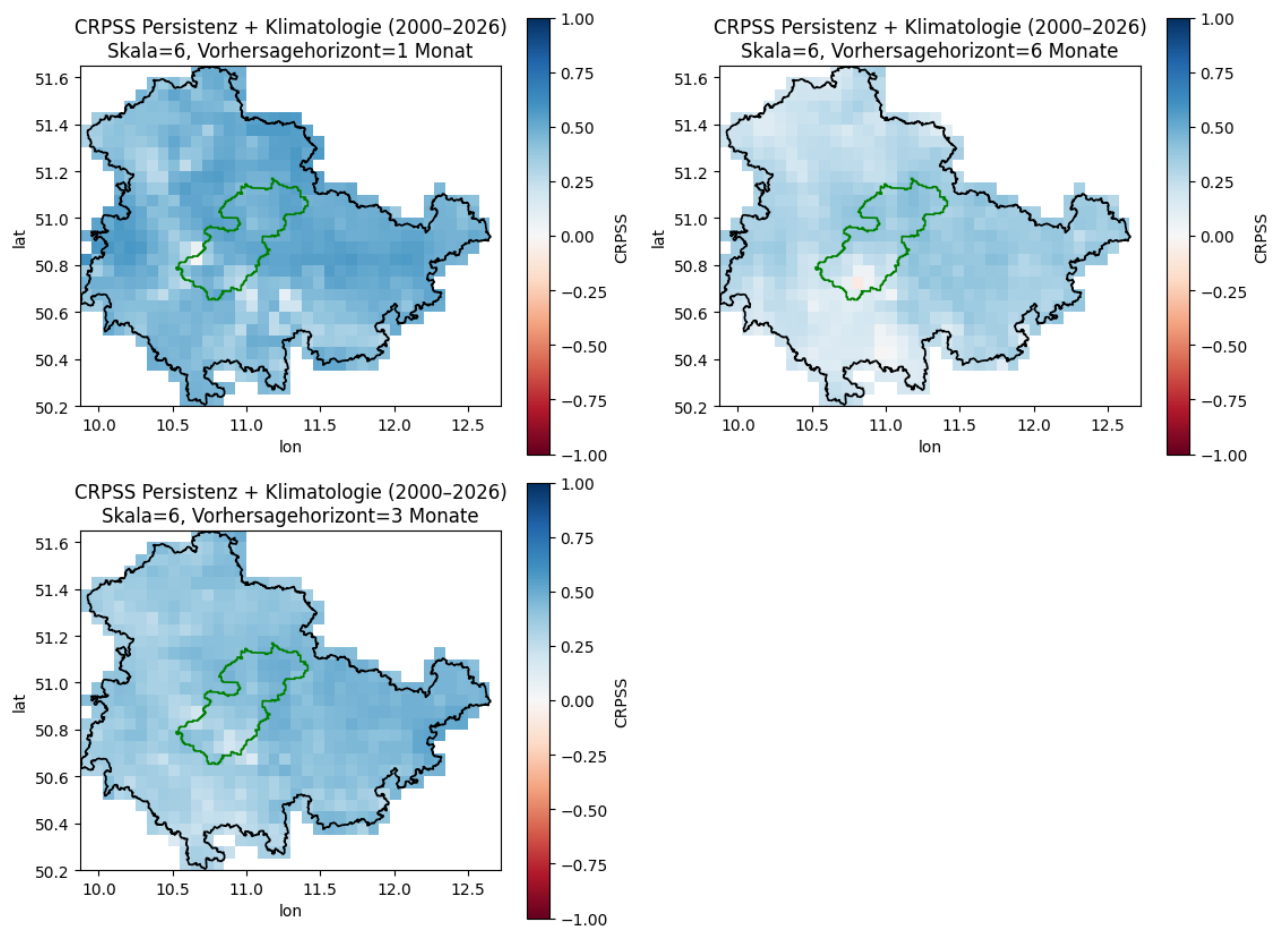


Abbildung 28: Thüringen CRPS-Skill-Score - Karten SPI.6M (Vorhersagehorizonte 1, 3 und 6 Monate)

3.3.4 Bias

Abbildung 29 zeigt Karten des Bias über den gesamten Untersuchungszeitraum. Der Bias ist definiert als Differenz zwischen dem Ensemblemittel der Vorhersage und der zugehörigen Beobachtung (Vorhersage – Messung). Positive Werte weisen auf einen größeren Index-Wert durch das Vorhersagemodell im Vergleich zur Beobachtung hin (damit wird der Index-Wert überschätzt), negative Werte auf einen kleineren. Dabei wird deutlich, dass im Norden Thüringens (Nordthüringer Hügelland / Südharz) sowie in Teilen des Südostens (Vogtländische Hochflächen und Kuppenland) tendenziell zu trockene Bedingungen vorhergesagt wurden, während in den übrigen Regionen – insbesondere im Thüringer Wald / Schiefergebirge – eher zu feuchte Bedingungen prognostiziert wurden. Mit zunehmendem Vorhersagehorizont verstärkt sich dieses Muster.

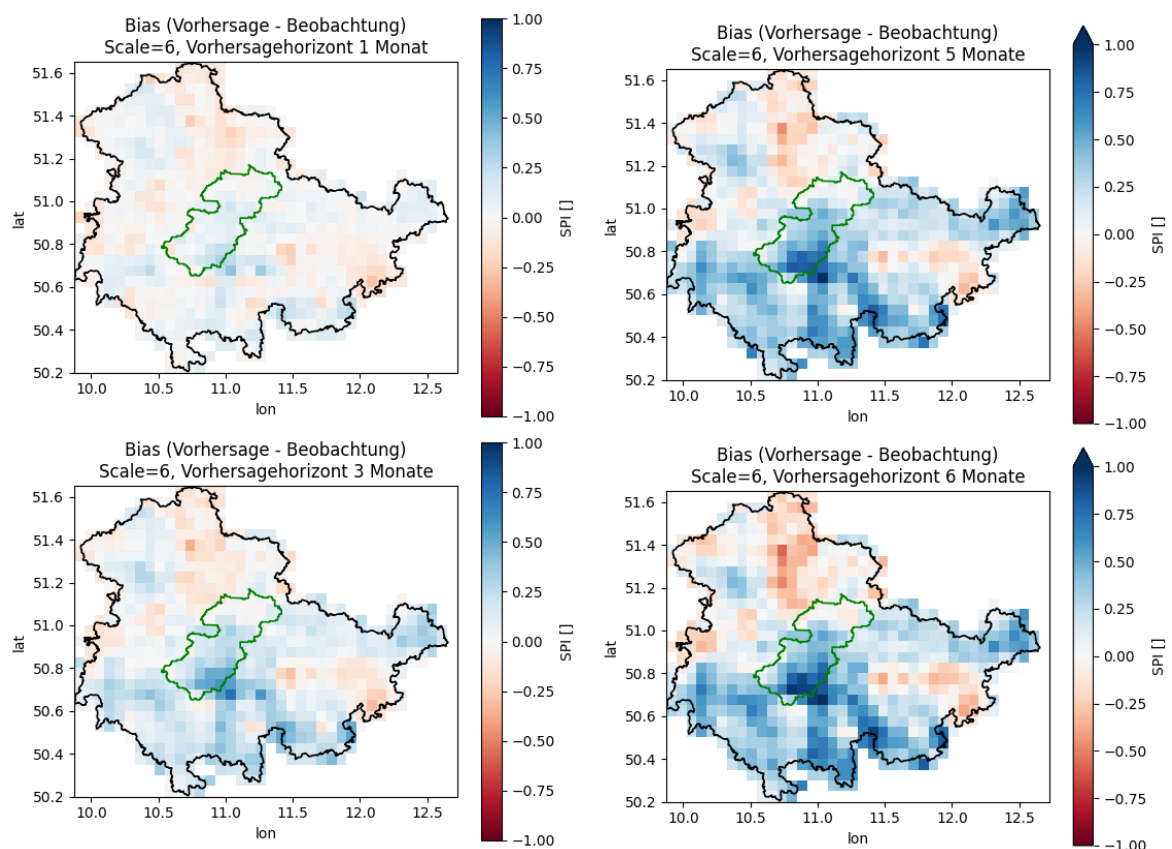


Abbildung 29: Thüringen Bias - Karten SPI.6M (Vorhersagehorizonte 1, 3, 5 und 6 Monate)

Dieses Bild wird durch die in Abbildung 30 dargestellten Ereignisse bestätigt. Dabei handelt es sich um vier Ereignisse, bei denen in einzelnen Regionen eine geringere Vorhersagequalität festgestellt wurde. Für den Zeitraum Juli bis September 2023 waren die gemittelten Vorhersagen zu trocken, was insbesondere im Norden Thüringens ausgeprägt war, während im Thüringer Wald zu feuchte Bedingungen vorhergesagt wurden. Auch in den Zeiträumen Februar bis April 2025 bzw. April bis Juni 2025 sowie Januar bis März 2026 waren die Vorhersagen im Thüringer Wald zu feucht.

Für das Gebiet Oberes Pleißeland / Altenburg-Zeitzer Lössgebiet ist ein differenzierteres Bild festzustellen. Hier waren die Vorhersagen in allen Gebieten im Mittel zu feucht, im Juli – September 2023 waren diese jedoch deutlich zu trocken. Auf der anderen Seite waren die Vorhersagen im April – Juni 2025 deutlich zu feucht.

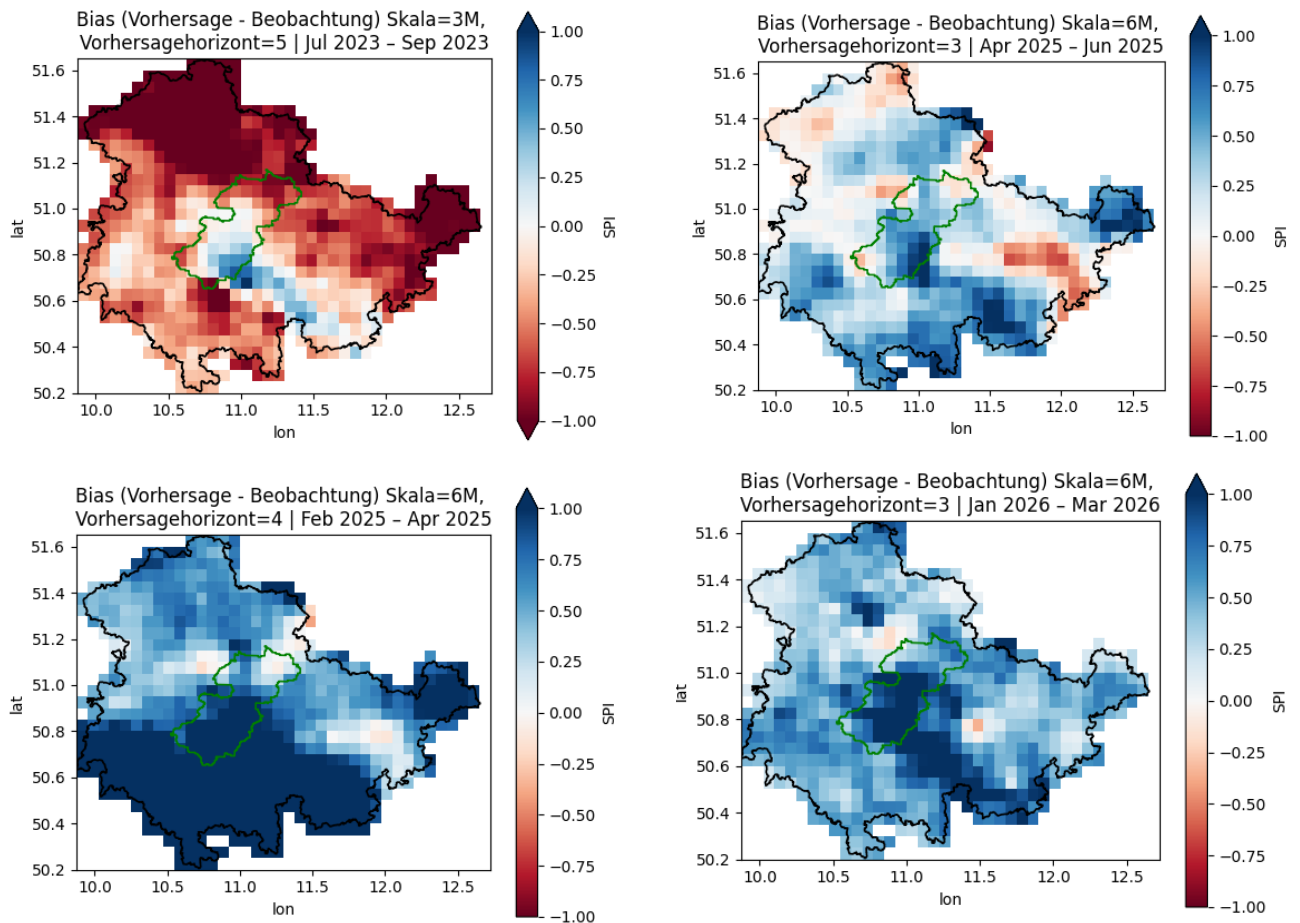


Abbildung 30: Thüringen Bias - Karten für einzelne Ereignisse

4 ZUSAMMENFASSUNG

In der Analyse der im Dashboard dargestellten Indizes lag der Fokus insbesondere auf dem SPI. Die vorliegende Untersuchung der Vorhersagen zeigt insbesondere bei kurzen Vorhersagehorizonten eine gute Übereinstimmung der prognostizierten Werte mit den Messdaten, die mit zunehmendem Vorhersagehorizont abnimmt. Es traten räumliche Unterschiede in der Vorhersagegüte auf, und einzelne Ereignisse wurden durch die Vorhersagen nur unzureichend erfasst – gerade im Hinblick auf Trockenheit sind dies im analysierten Zeitraum jedoch die entscheidenden Fälle. Vor allem bei kürzeren Skalen konnten trockene Bedingungen, wie z.B. im Frühjahr 2025 nur unzureichend vorhergesagt werden.

- Die Vorhersagequalität nimmt erwartungsgemäß mit steigendem Vorhersagehorizont ab.
- Insgesamt stimmen die Vorhersagen insbesondere mit einmonatigem Vorhersagehorizont (d.h. die Vorhersage von z.B. Anfang Juli für den Gesamtmonat Juli) überwiegend gut mit den Messdaten überein. Es gab einzelne Ereignisse, wo die Vorhersage die Messdaten weniger gut treffen konnte, wobei insbesondere das Frühjahr 2025 hervorsteicht.
- Zwischen den unterschiedlichen Skalen besteht ein Zusammenhang in der Vorhersagegüte. Bei der absoluten Güte zeigen kleinere Skalen (z. B. drei Monate) eine höhere Variabilität bzw. Unsicherheit als größere Skalen (z. B. zwölf Monate). Beim Brier Skill Score hingegen schneiden kürzere Skalen teilweise besser ab als längere, was auf den Einfluss der gewählten Referenz zurückzuführen sein kann.

- Es bestehen deutliche räumliche Unterschiede in der Vorhersagegüte. Die vergleichsweise schlechte Performance im Frühjahr 2025 zeigt sich insbesondere im Thüringer Wald / Schiefergebirge sowie in der Region Rhön / Werra.
- Die Vorhersagen verhalten sich bei höheren Vorhersagehorizonten konservativ hinsichtlich der Prognose von Schwellwertunterschreitungen. Während bei kurzen Horizonten noch Schwellwertunterschreitungen mit hoher Wahrscheinlichkeit (über 60 % der Ensemblemitglieder) für extreme (–2) und moderate (–1) Dürre vorhergesagt werden, treten solche hohen Wahrscheinlichkeiten bei längeren Horizonten kaum noch auf.
- Wenn eine Schwellenwertunterschreitung mit hoher Wahrscheinlichkeit prognostiziert wird, tritt dieses Ereignis in der Regel auch tatsächlich ein – unabhängig vom Vorhersagehorizont.
- Hinsichtlich der relativen Güte (Skill Scores) zeigt sich ein differenziertes Bild: Der Brier Skill Score weist bei höheren Vorhersagehorizonten eine geringere Güte auf, während der CRPS Skill Score über alle Vorhersagehorizonte hinweg durchgehend bessere Werte als die Referenz erzielt.
- Im Mittel liefern die betrachteten Naturräume ähnliche Ergebnisse. Lediglich der Thüringer Wald / Schiefergebirge weist über alle Vorhersagehorizonte hinweg eine tendenziell schlechtere Vorhersagegüte als die übrigen Regionen auf.

Die Vorhersagen eignen sich grundsätzlich dazu, allgemeine Trends der Dürreentwicklung kurzfristig und regionsspezifisch etwas unterschiedlich für eins, zwei bis drei Monate im Voraus zu erkennen. Vor allem die relative Vorhersagegüte weist insgesamt – auch für längere Vorhersagehorizonte – gute Ergebnisse auf. Insofern liefert die Vorhersage einen Mehrwert gegenüber der verwendeten Referenz, einer Kombination aus Persistenz und Klimatologie.

Hinsichtlich der Vorhersage von Schwellwertunterschreitungen moderater (–1) und extremer Dürre (–2) ist jedoch einschränkend festzustellen, dass diese Ereignisse bei höheren Vorhersagehorizonten nur selten mit hoher Wahrscheinlichkeit prognostiziert werden. Gleichzeitig zeigt die Analyse, dass bei einer vorhergesagten Wahrscheinlichkeit von mehr als 50 % die tatsächliche Eintrittswahrscheinlichkeit einer Schwellenwertunterschreitung hoch ist. Falsch prognostizierte Schwellwertunterschreitungen – also Fälle, in denen eine hohe Eintrittswahrscheinlichkeit vorhergesagt wurde, das Ereignis jedoch ausblieb – traten im Untersuchungszeitraum nicht auf. Folglich gibt es bei der saisonalen Vorhersage, vor allem für kürzere Aggregationszeiträume und längeren Vorhersagehorizonte, Schwierigkeiten hinsichtlich der Vorhersage sehr trockener Bedingungen.

Die Vorhersage von Extremwerten im Rahmen saisonaler Vorhersagen stellt eine bekannte Herausforderung dar¹². Insbesondere für die Prognose von Trockenphasen könnten Weiterentwicklungen des Vorhersagesystems zu einer Verbesserung der Vorhersagegüte beitragen. Mögliche Ansätze hierfür betreffen verbesserte Vorhersage-Produkte des DWD (statistisch selektierte Vorhersage) sowie die Anwendung weiterer Post-Processing-Schritte. Darüber hinaus kann die Implementierung und anschließende Analyse der Hindcasts zusätzliche Erkenntnisse über systematische Fehler und Schwachstellen des Vorhersagesystems liefern. Die Auswertung der Vorhersagegüte der Hindcasts wird zukünftig Aufschluss darüber geben, welche Güte das Vorhersagesystem erreicht und in welchem Umfang die Ergebnisse für operationelle Anwendungen herangezogen werden können.

¹² Lavaysse, C., Stockdale, T., McCormick, N., Vogt, J. (2020): Evaluation of a New Precipitation-Based Index for Global Seasonal Forecasting of Unusually Wet and Dry Periods. *Weather and Forecasting*, 35(4), 1189–1202. <https://doi.org/10.1175/WAF-D-19-0196.1>

5 ANHANG

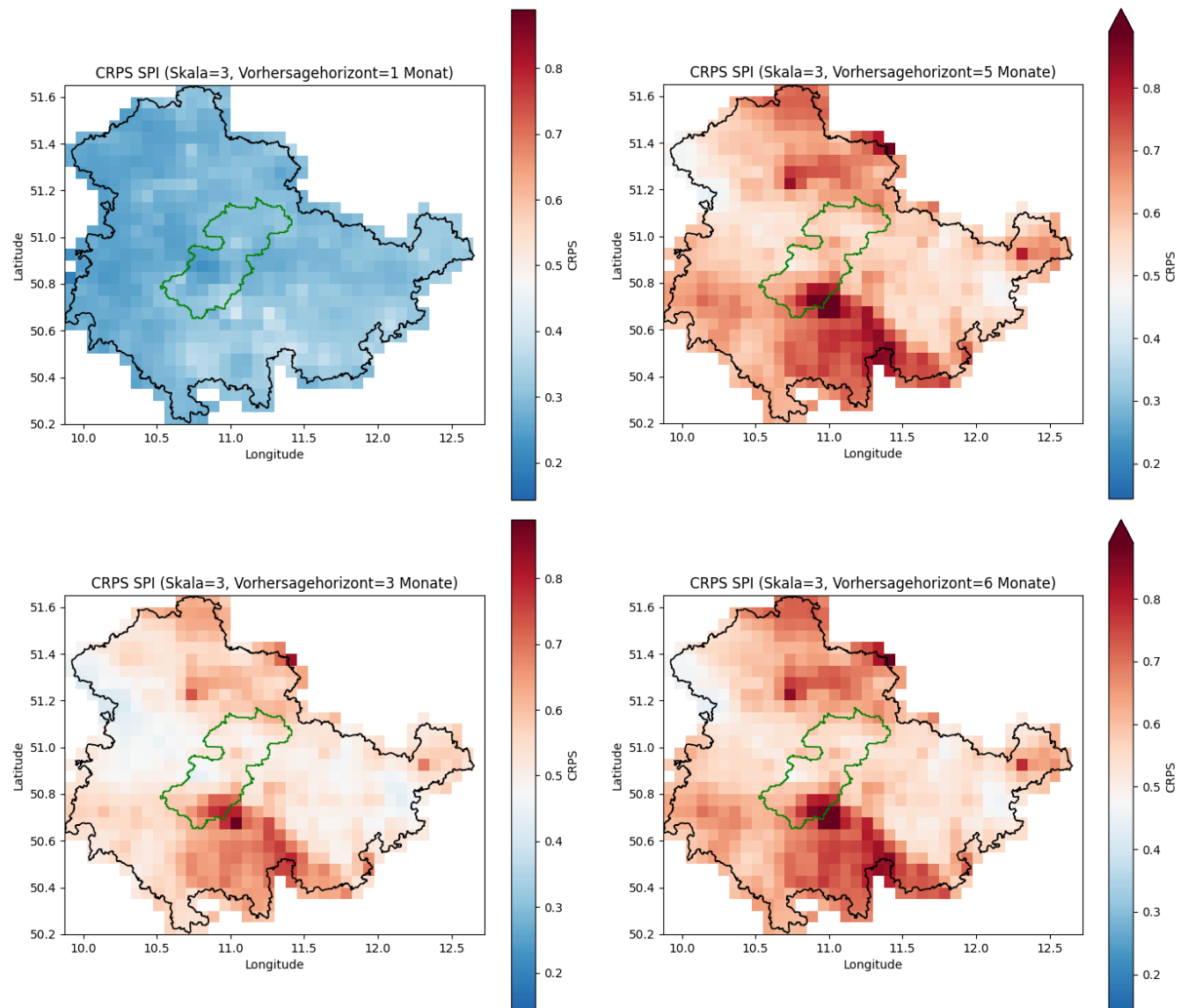


Abbildung 31: Thüringen CRPS - Karten SPI.3M (Vorhersagehorizonte 1, 3, 5 und 6 Monate)

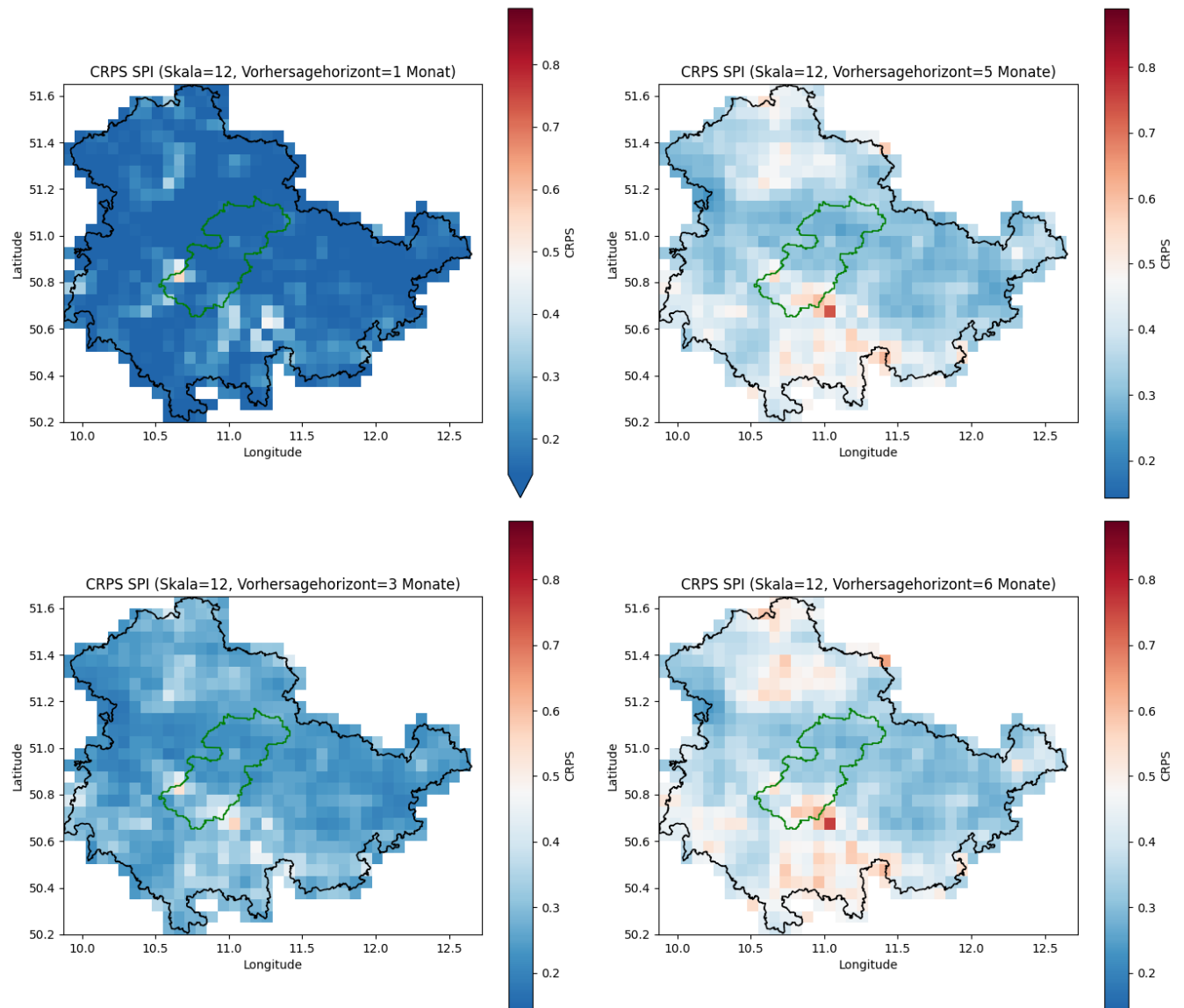


Abbildung 32: Thüringen CRPS - Karten SPI.12M (Vorhersagehorizonte 1, 3, 5 und 6 Monate)

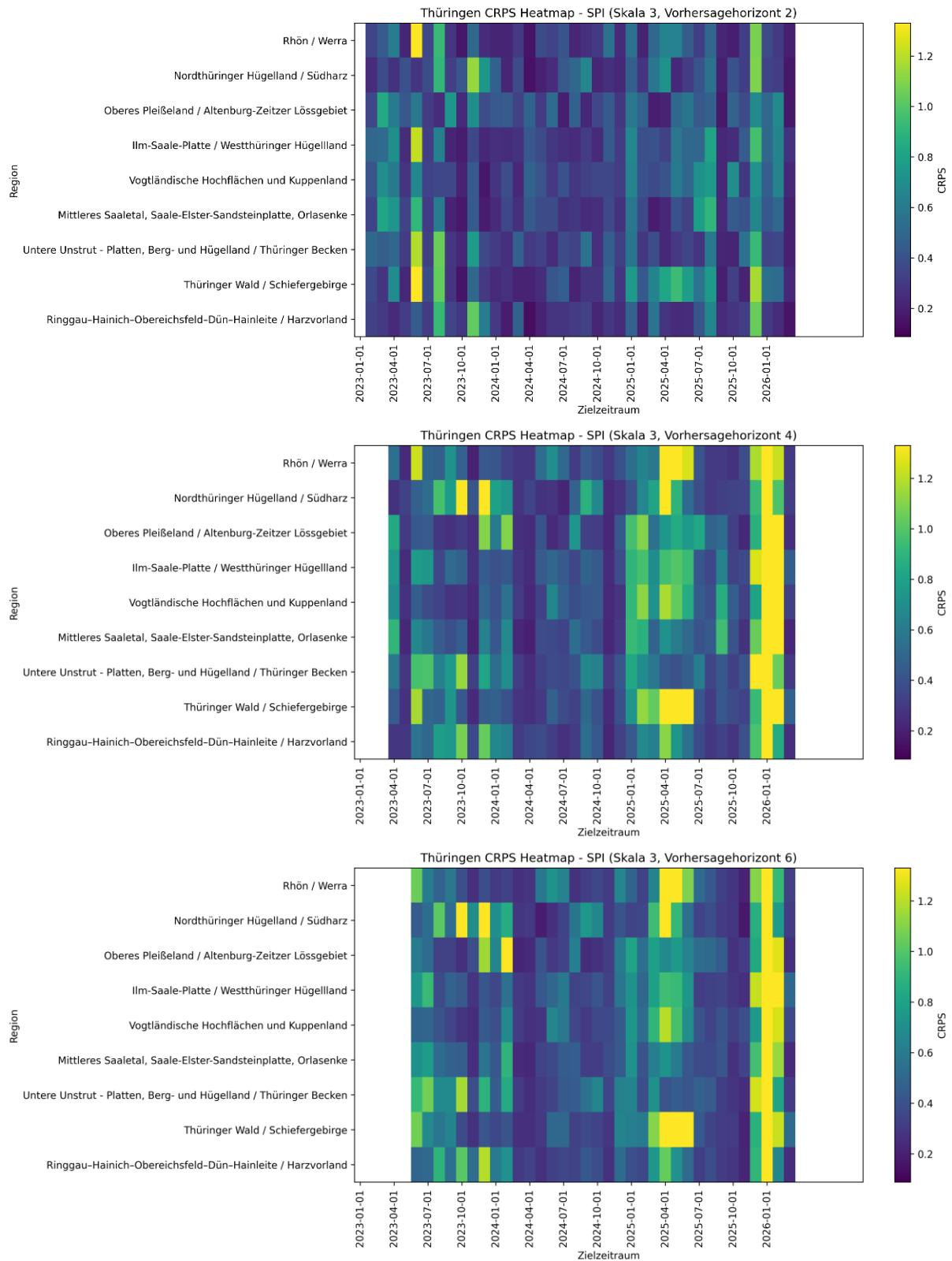


Abbildung 33: Thüringen CRPS Regionen - Heatmap nach Zielzeitraum SPI.3M (Vorhersagehorizonte 2, 4 und 6 Monate)

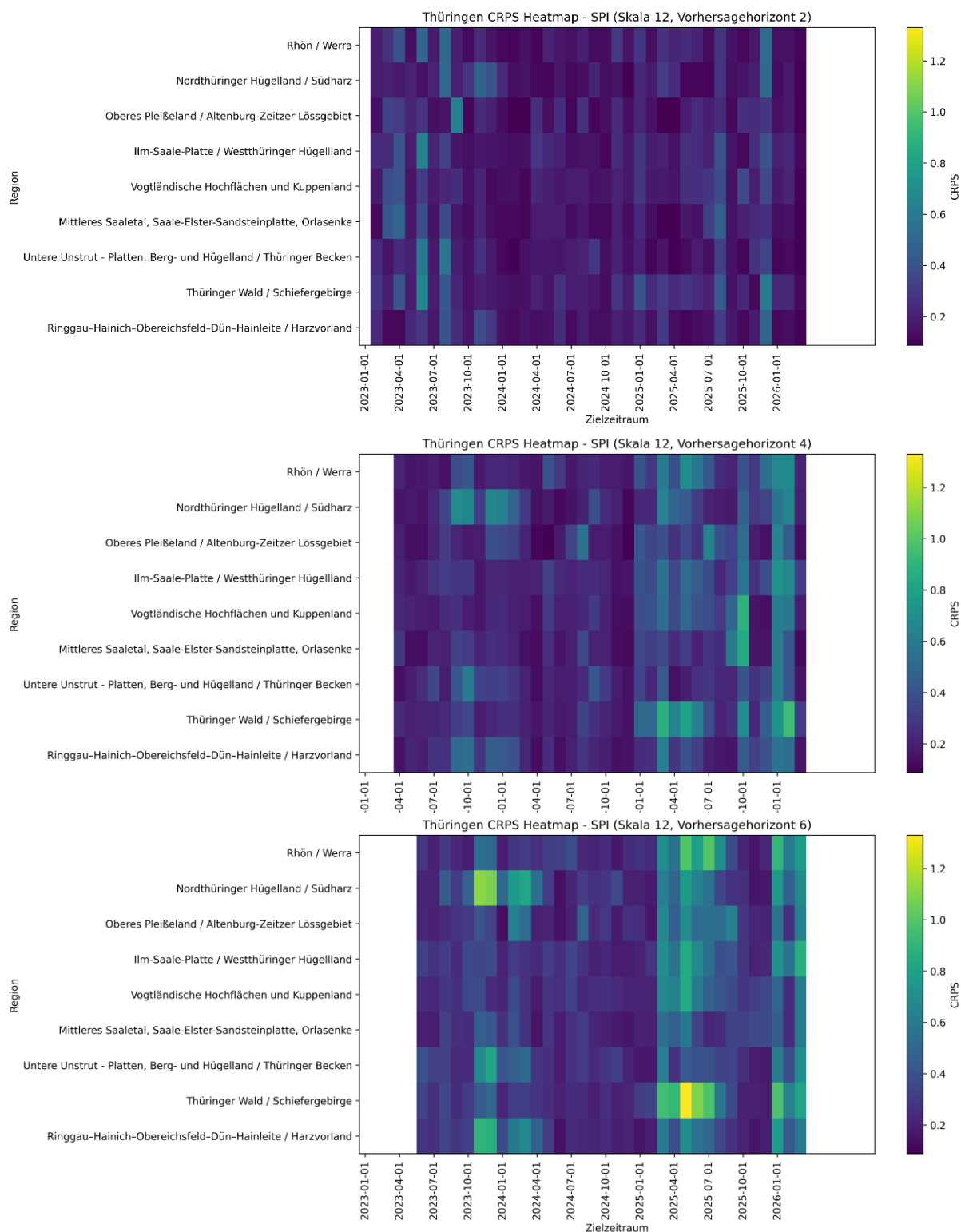


Abbildung 34: Thüringen CRPS Regionen - Heatmap nach Zielzeitraum SPI.12M (Vorhersagehorizonte 2, 4 und 6 Monate)

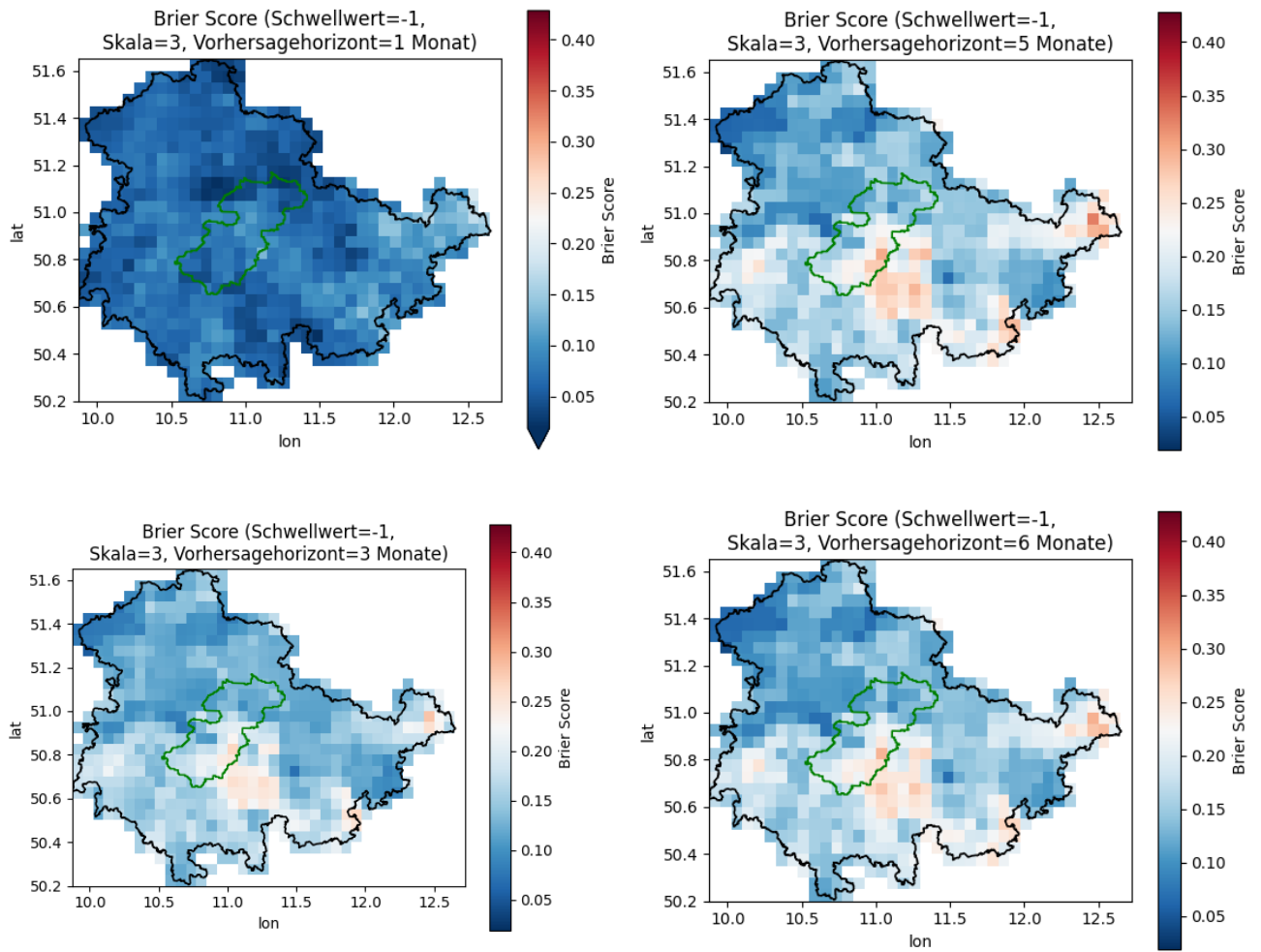


Abbildung 35: Thüringen Brier Score - Karten SPI.3M (Schwellwert -1; Vorhersagehorizonte 1, 3, 5 und 6 Monate)

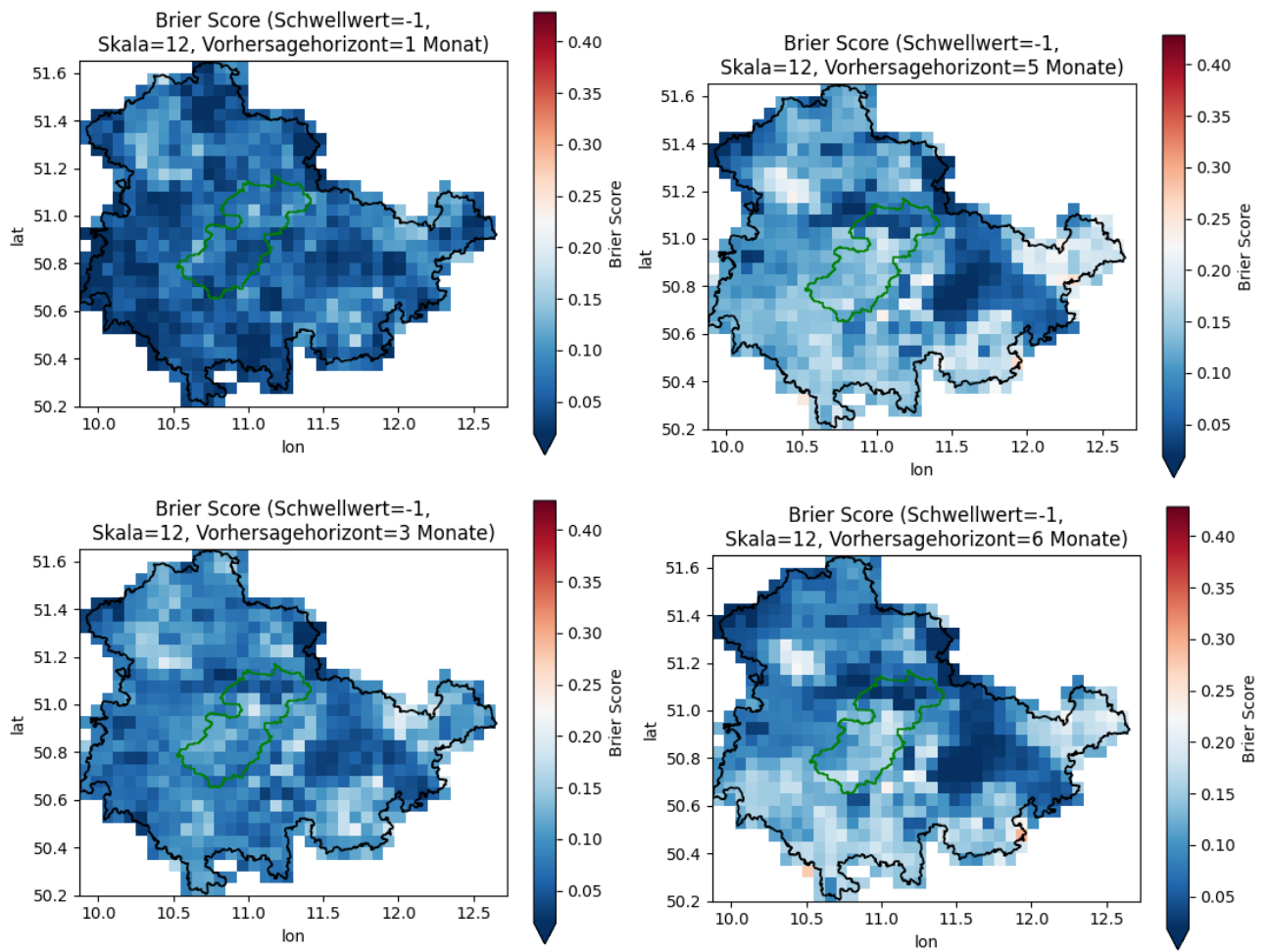


Abbildung 36: Thüringen Brier Score - Karten SPI.12M (Schwellwert -1; Vorhersagehorizonte 1, 3, 5 und 6 Monate)

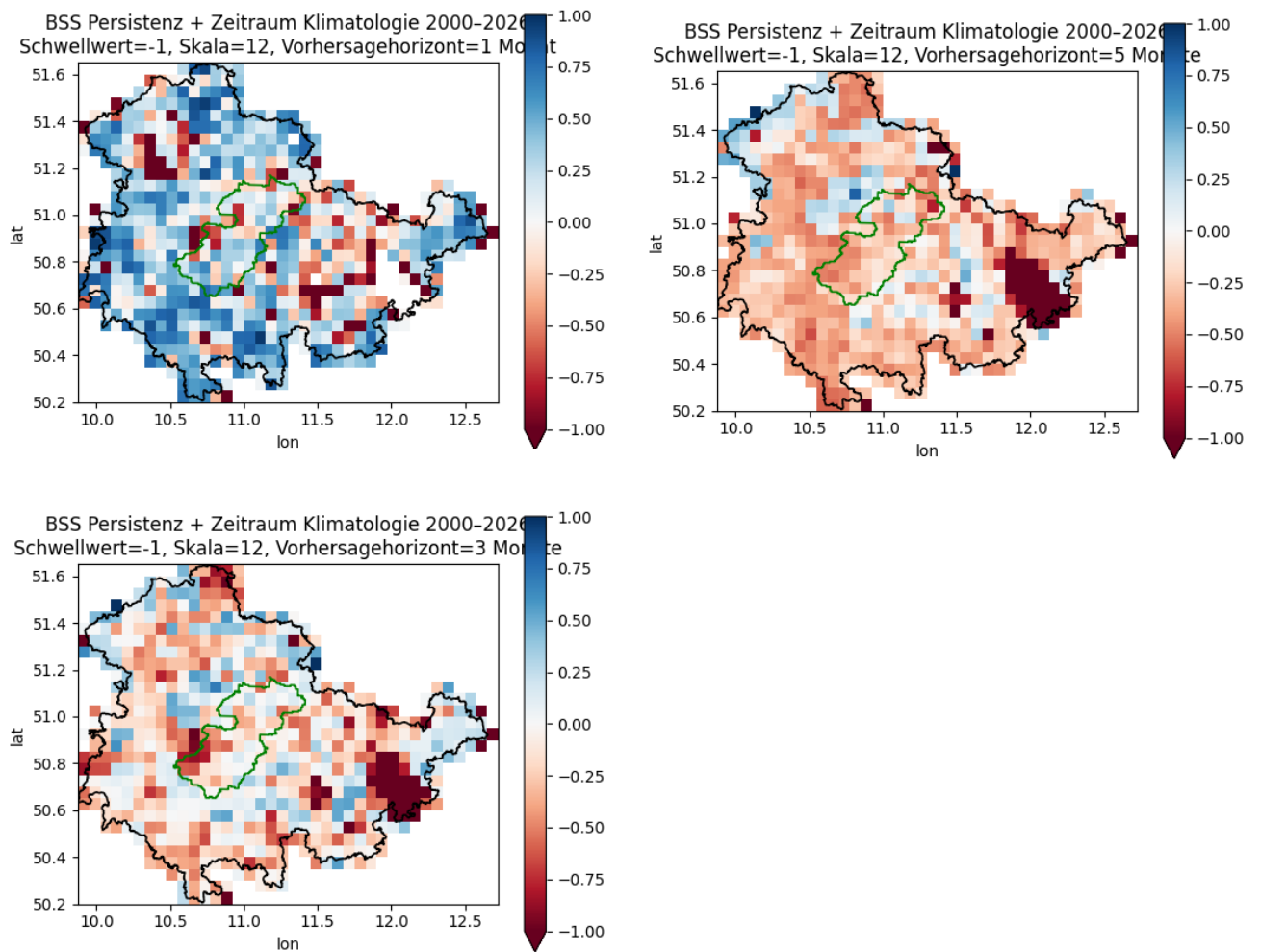


Abbildung 37: Thüringen Brier-Skill-Score - Karten SPI.12M (Schwellwert -1; Vorhersagehorizonte 1, 3 und 5 Monate)

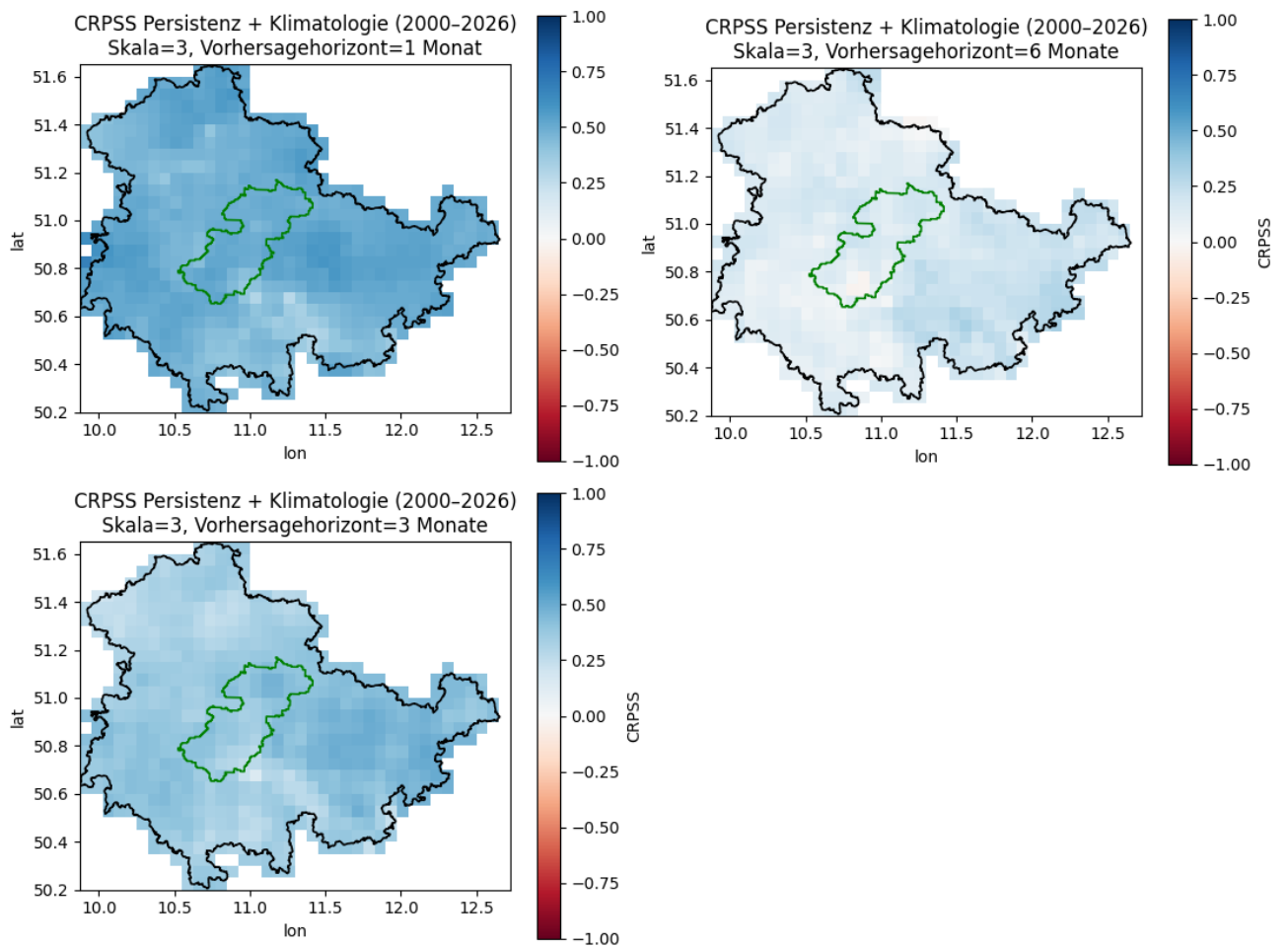


Abbildung 38: Thüringen CRPS-Skill-Score - Karten SPI.3M (Vorhersagehorizonte 1, 3 und 6 Monate)

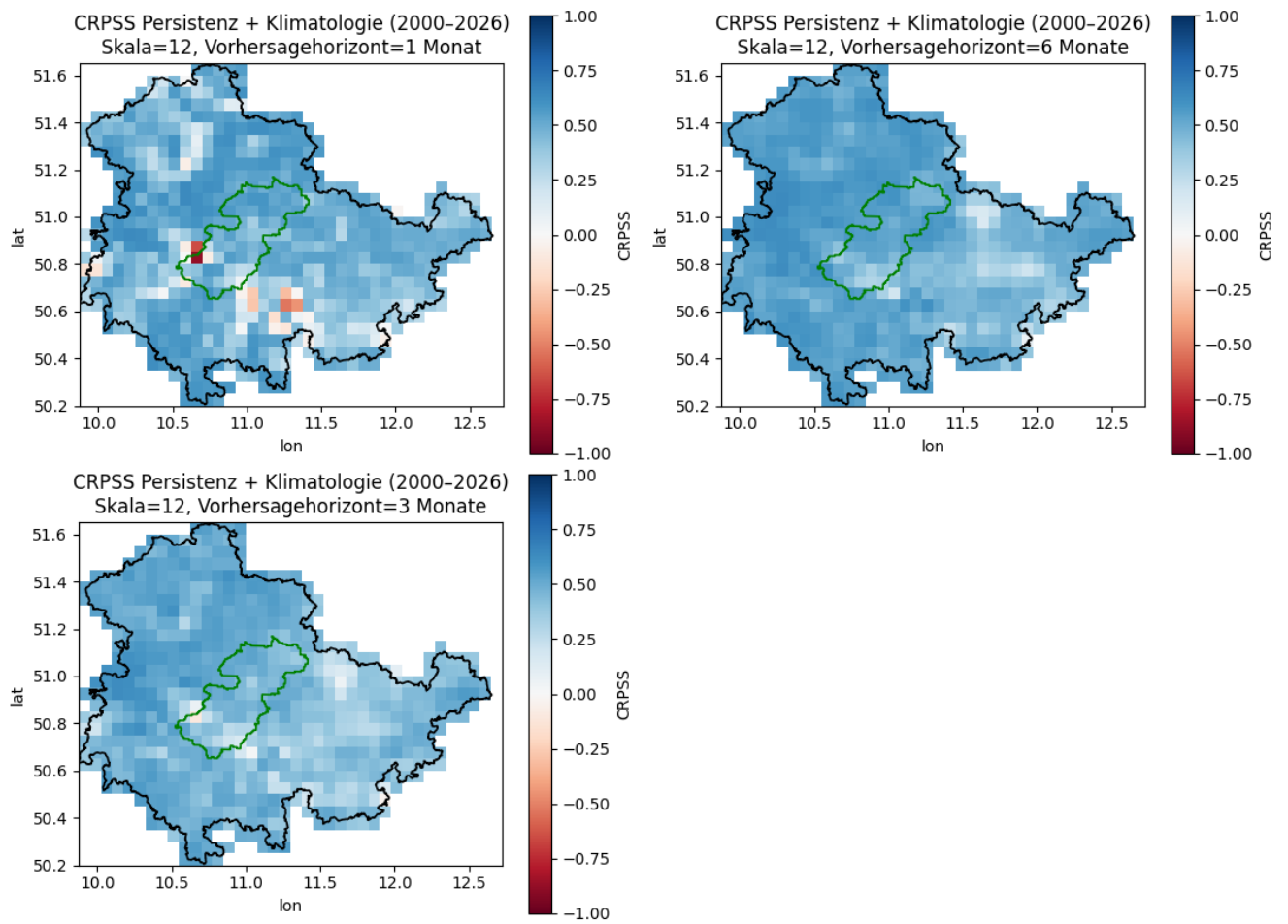


Abbildung 39: Thüringen CRPS-Skill-Score - Karten SPI.12M (Vorhersagehorizonte 1, 3 und 6 Monate)