

Mittel- und Langfristvorhersagen für die Bundeswasserstraßen

Barbara Frielingsdorf, Bastian Klein, Dennis Meißner
Bundesanstalt für Gewässerkunde

Regionales Delft-FEWS-Anwendertreffen
Essen, 03 – 04 Juli 2025



Delft-FEWS-DE
Vorhersagesystem des Bundes

BfG Bundesanstalt für Gewässerkunde

Im Auftrag von:
Bundesministerium für Digitales und Verkehr

WSV.de

Partner:
Deltares

Inhalt

- Mittel- und Langfristvorhersage bei der BfG
- Qualitätssicherung
- Kommunikation der Vorhersagegüte



Mittel- und Langfristvorhersage bei der BfG

Zeitskalen



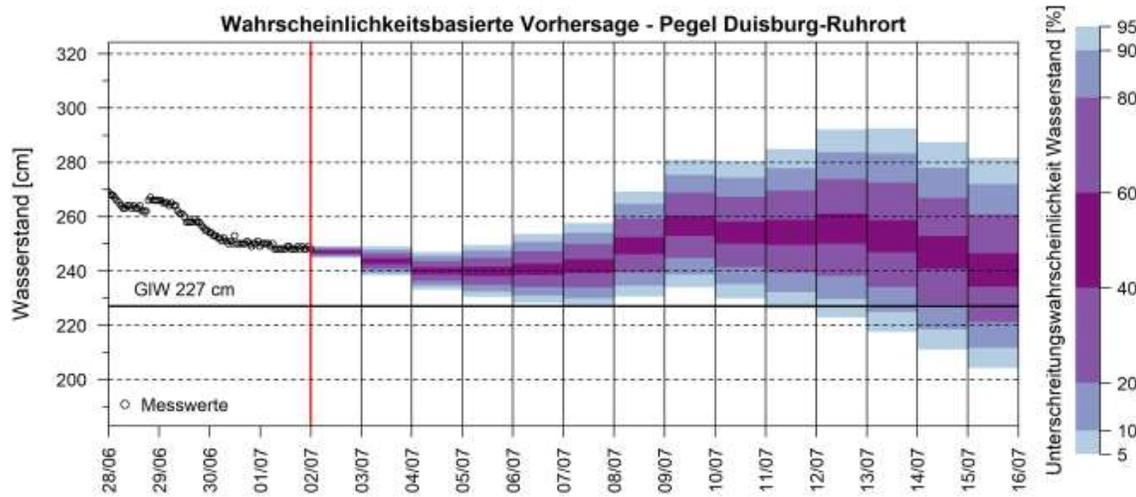
Veröffentlichung



Wahrscheinlichkeitsbasierte 14-Tage-Vorhersage Pegel Duisburg-Ruhrort

Vorhersage vom: 02.07.2025 00:00 MEZ

Die in der Grafik und in den Tabellen angegebenen Unter- bzw. Überschreitungswahrscheinlichkeiten definierter Wasserstände basieren auf Tagesmittelwerten (jeweils 00:00 bis 00:00 des Folgetages).



Unterschreitungswahrscheinlichkeiten definierter Wasserstände [%]

	02.07	03.07	04.07	05.07	06.07	07.07	08.07	09.07	10.07	11.07	12.07	13.07	14.07	15.07
GIW: 227 cm	0	0	0	1	3	5	2	1	3	5	8	12	19	28
247 cm	49	86	95	89	79	71	42	25	32	34	34	40	50	61
267 cm	100	100	100	100	100	100	93	77	80	76	70	73	80	86
287 cm	100	100	100	100	100	100	100	98	98	96	92	92	95	97
307 cm	100	100	100	100	100	100	100	100	100	100	99	99	99	100

Überschreitungswahrscheinlichkeiten HW-Marke I und HW-Marke II (HSW) [%]

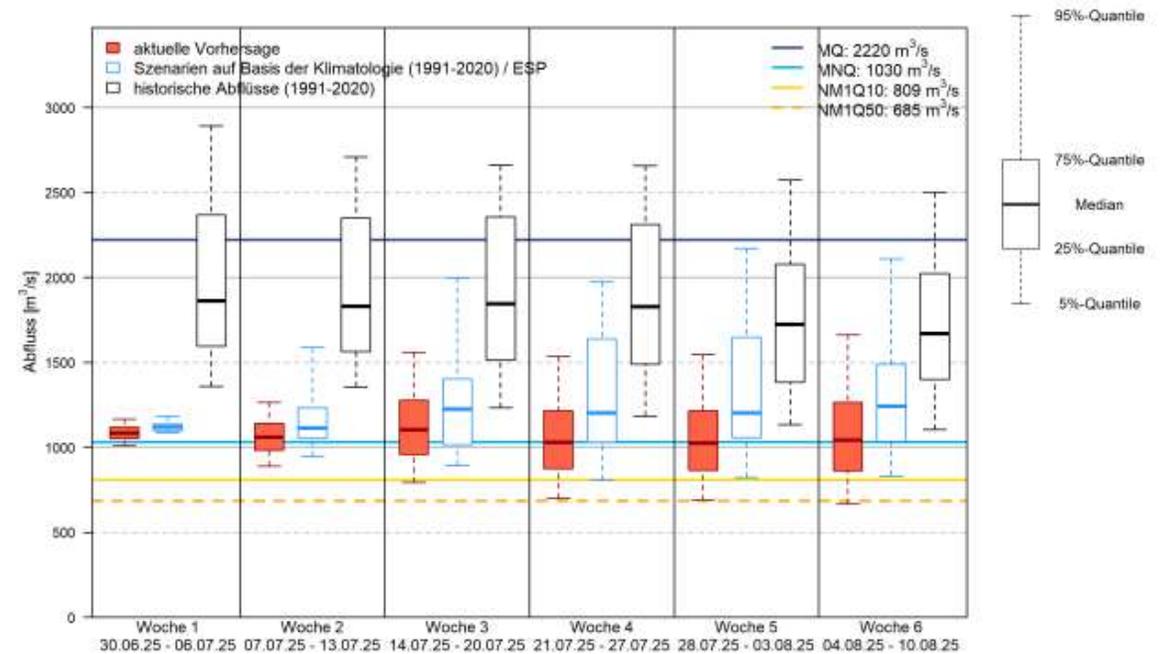
	02.07	03.07	04.07	05.07	06.07	07.07	08.07	09.07	10.07	11.07	12.07	13.07	14.07	15.07
Marke I: 930 cm	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Marke II: 1130 cm	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Hydrologische 6-Wochen-Vorhersage Pegel Duisburg-Ruhrort

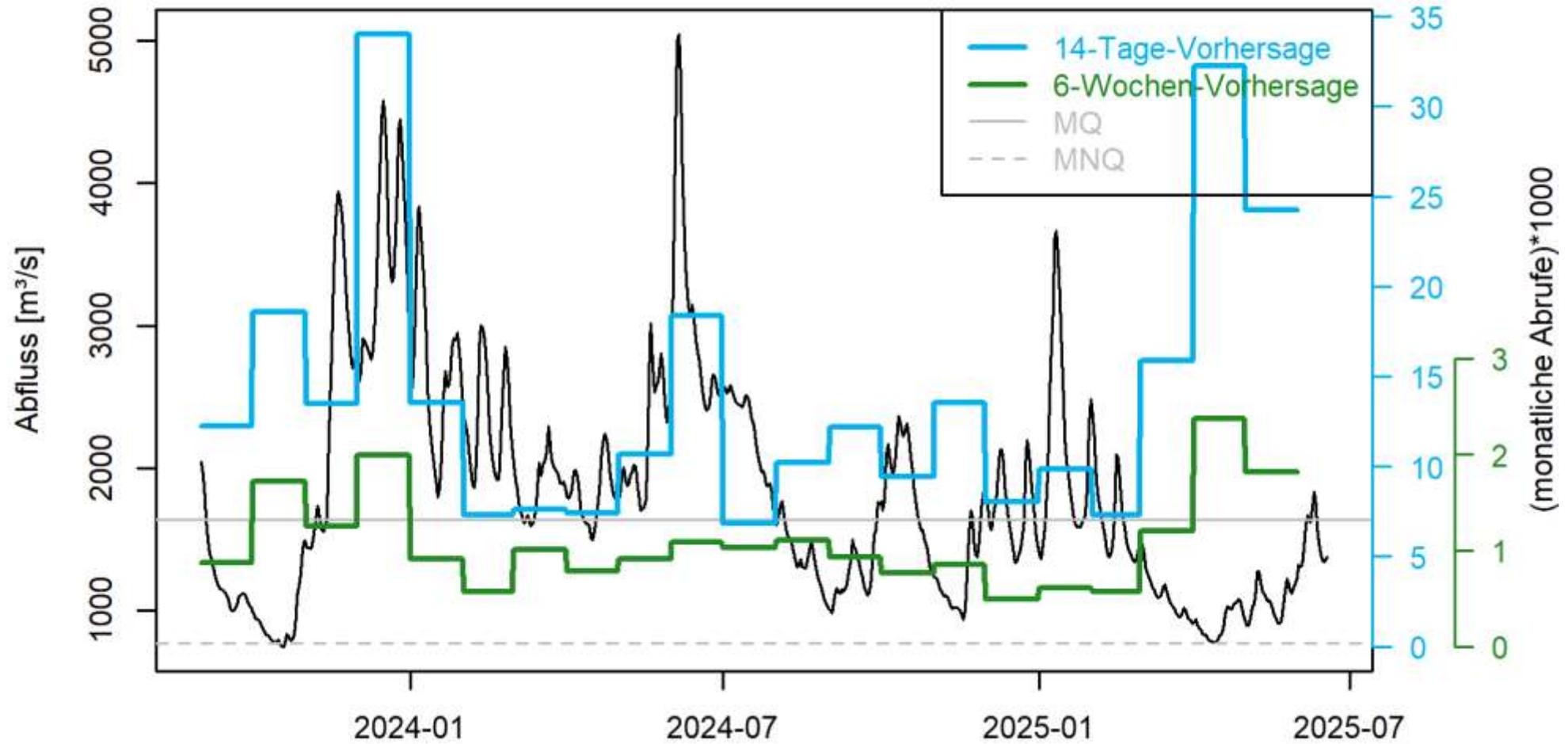
Vorhersage vom: 30.06.2025



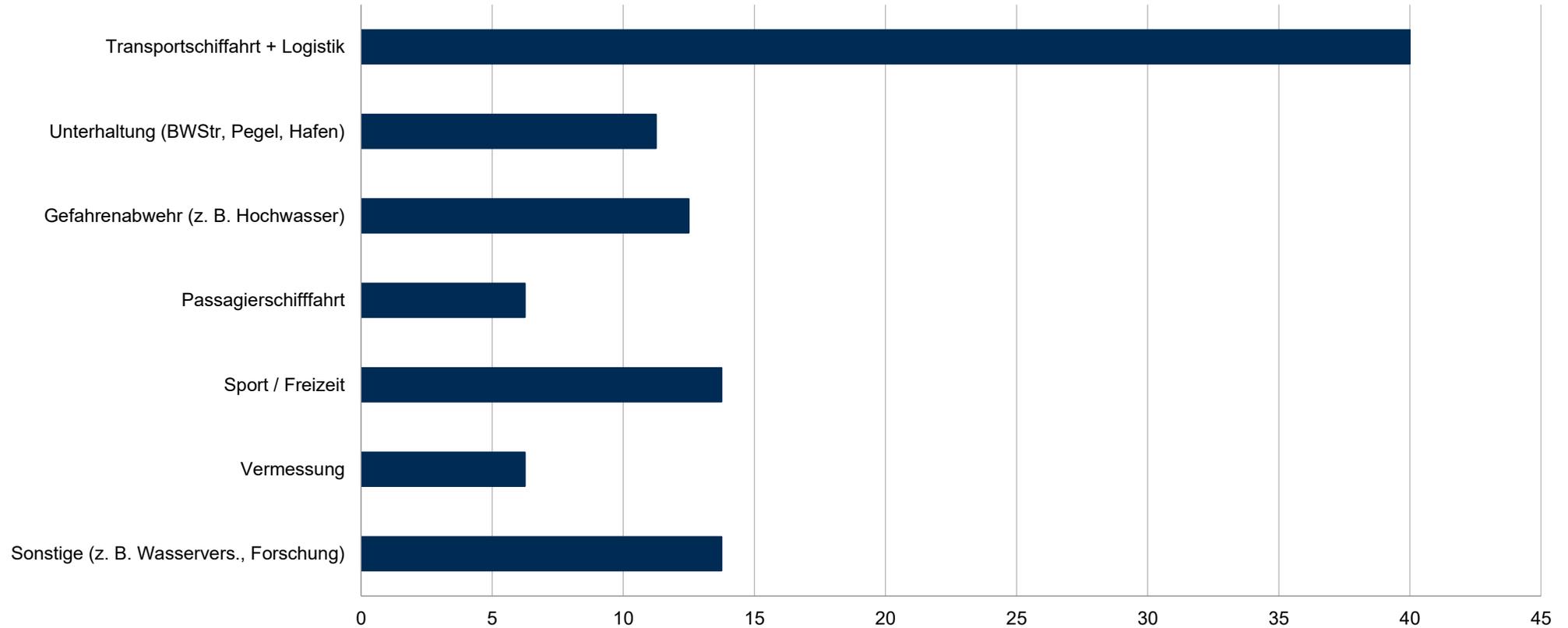
Die Box-Plot Diagramme stellen die Wahrscheinlichkeitsverteilung der Abflussentwicklung der kommenden Wochen dar (Wochenmittelwerte). Neben der aktuellen Vorhersage (rot) sind zu ihrer Einordnung die Verteilung einer auf der Klimatologie basierenden Prognose (blau) sowie der historischen Abflüsse im entsprechenden Zeitraum der Jahre 1991-2020 dargestellt.



Kaub (Rhein)



Nutzungsbereiche der Mittel- und Langfristvorhersagen in %

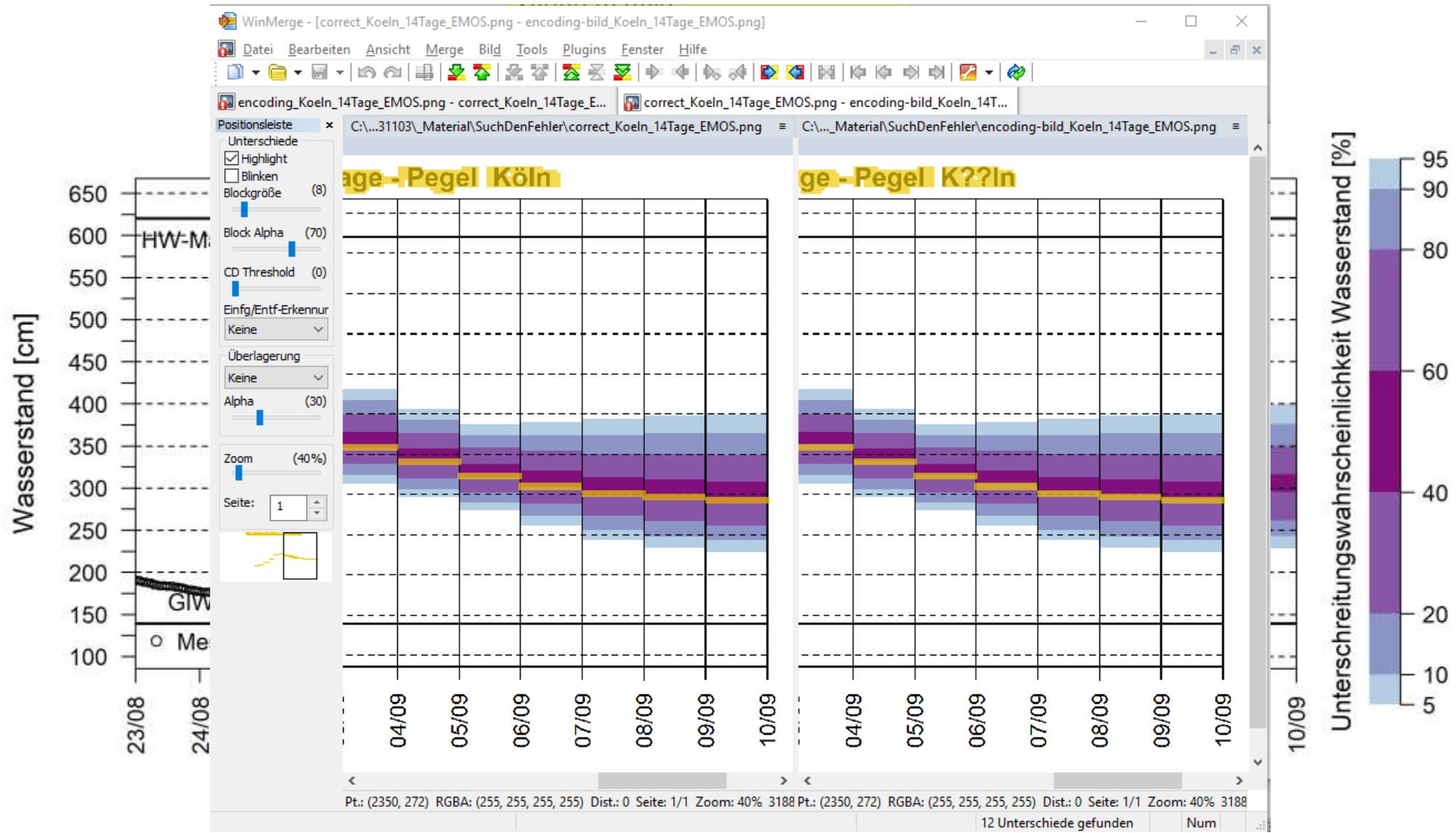


Qualitätssicherung

Risiken:

- Anpassung von Workflows
- Änderung der Methodik
- Neue Datengrundlage
- Systemupdates





Technische Um

Anforder

- Delft-FEWS parallel u
- Konfigura wird nich
- localDat
- Ordnerir temp, ...
- Testen m Ergebnis

```
<?xml version="1.0" encoding="UTF-8"?>
<wftrEntireConfig xmlns="de.bjoernsen.fews.wftr.config" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="de.bjoernsen.fews.wftr.config ../../de.bjoernsen.fews.wftr.config/etc/schema/fewsWftrTestConfig.xsd">
  <wftrSingleFewsConfig>
    <regionHome>..\FEWSDE_GIT\FewsDE_SA</regionHome>
    <wftrFewsExe>..\FEWSDE_GIT\bin\fews-stable-202102-108652-bin\windows\Delft-FEWS.exe</wftrFewsExe>
    <wftrFewsPatchJar>*202102*patch.jar</wftrFewsPatchJar>
    <wftrStartBat>_wftrStart.bat</wftrStartBat>

    <wftrConfig>
      <name>000_ProofOfConcept</name>
      <keepFinalDirs>>true</keepFinalDirs>
      <wftrBaseDir>FewsDE_SA\000_ProofOfConcept<\</wftrBaseDir>
      <wftrAdditionalBaseDir>M:\000_ProofOfConcept<\</wftrAdditionalBaseDir>
      <wftrXmlFile>FewsDE_SA\000_ProofOfConcept\000_ProofOfConcept_WftrTest.xml</wftrXmlFile>
    </wftrConfig>

    <wftrConfig>□
    <wftrConfig>□

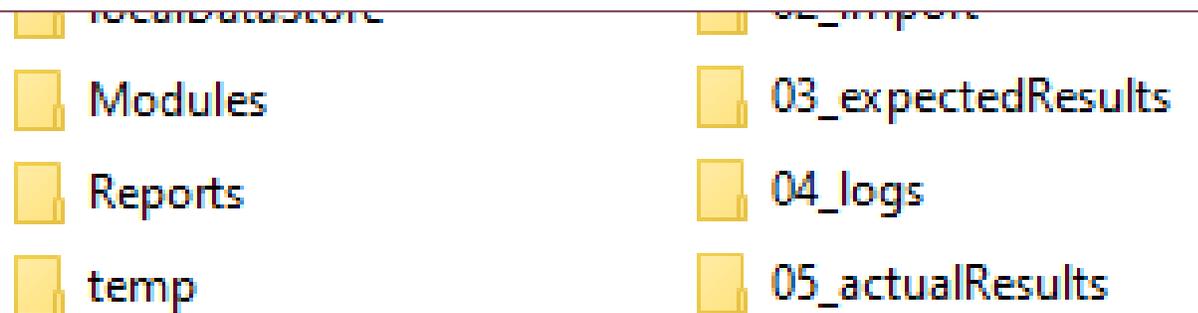
  </wftrSingleFewsConfig>
</wftrEntireConfig>
```

<wftrXmlFile>

- Ordnerstruktur der Workflows
- Kopiervorgänge
- Workflows (inkl. T0, Kaltstart, etc.)

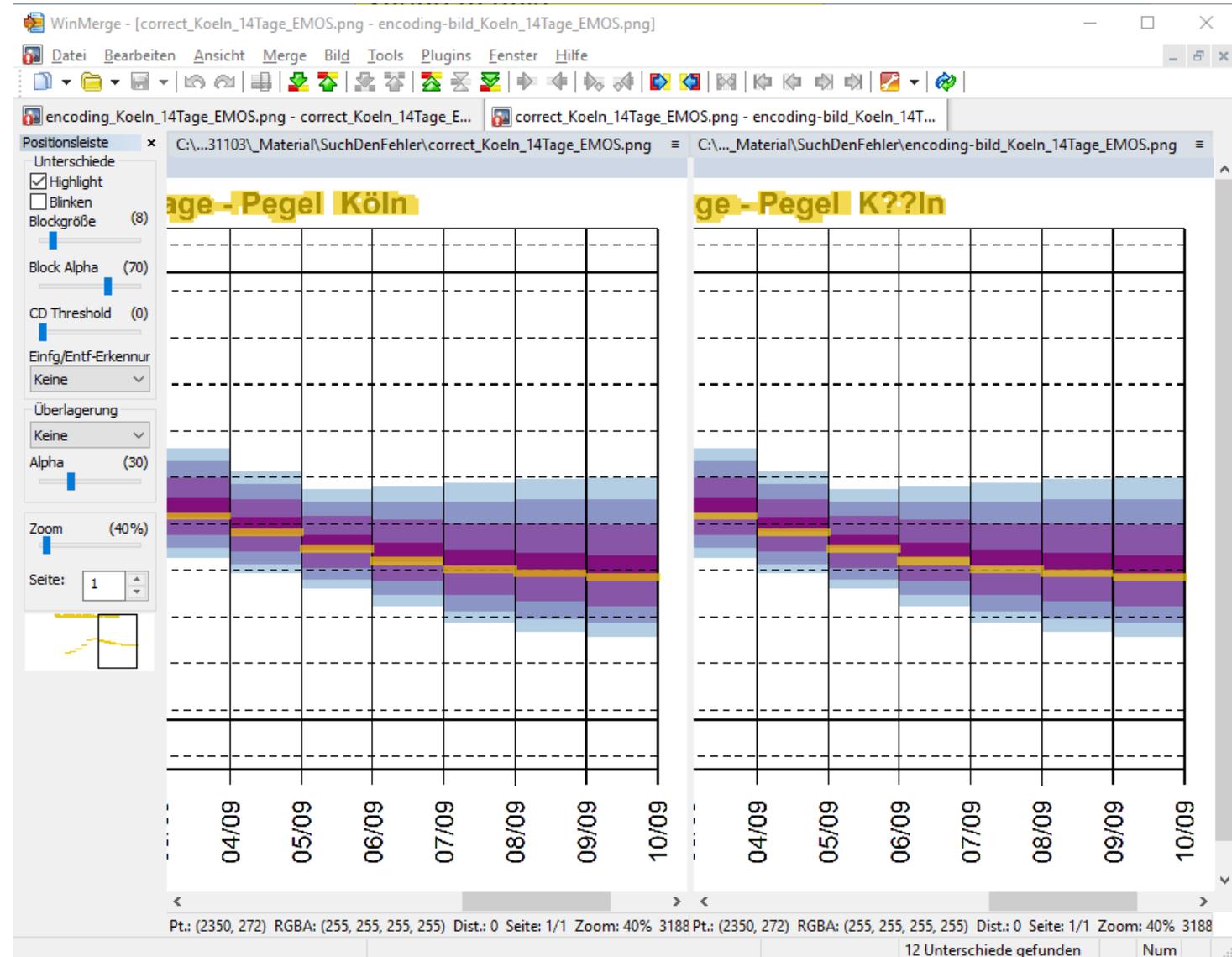
=> Konfiguration via workflowTestRun

- Reproduzierbare Ergebnisse: Speichern von Vergleichsergebnissen



Testablauf

- Ordner des SA sichern (LocalDataStore, etc.)
- Reproduzierbarer Anfangszustand: kopiere vorbereitete Ordner (Import, LocalDataStore)
- Start Delft-FEWS workflowTestRun
- Vergleichbarkeit der Ergebnisse: Postprozessierung wie z. B. löschen von Einträgen, wie <creationDate> und <creationTime> in PI-XML-Dateien
- Für die Fehlersuche: kopiere die Ordner localDataStore und generalAdapterRun
- Clean Up
- Start von verschiedenen Vergleichtools



Regelmäßiges Testen schützt vor bösen Überraschungen!

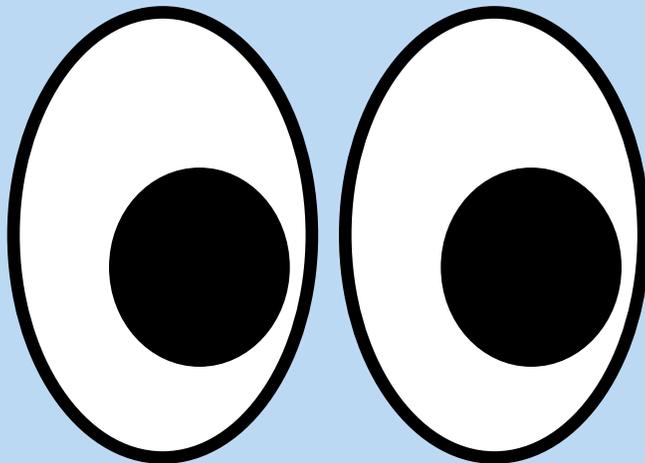
Testen bedeutet mehr Aufwand, sorgt aber auch für mehr Entspannung.



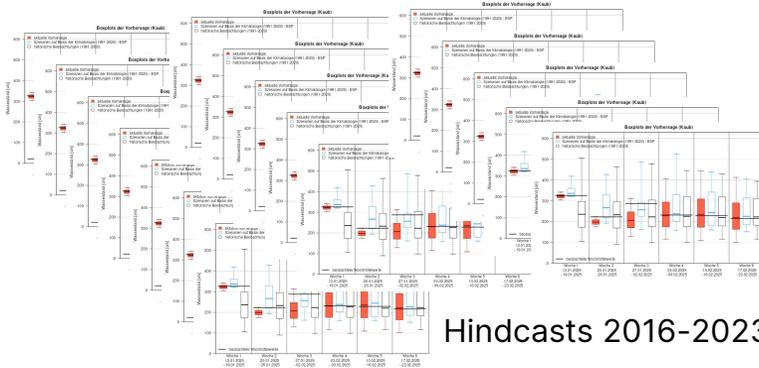
Kommunikation der Vorhersagegüte (6-Wochen-Vorhersage)



Wie wird die Vorhersagegüte berechnet?



Berechnung der Vorhersagegüte

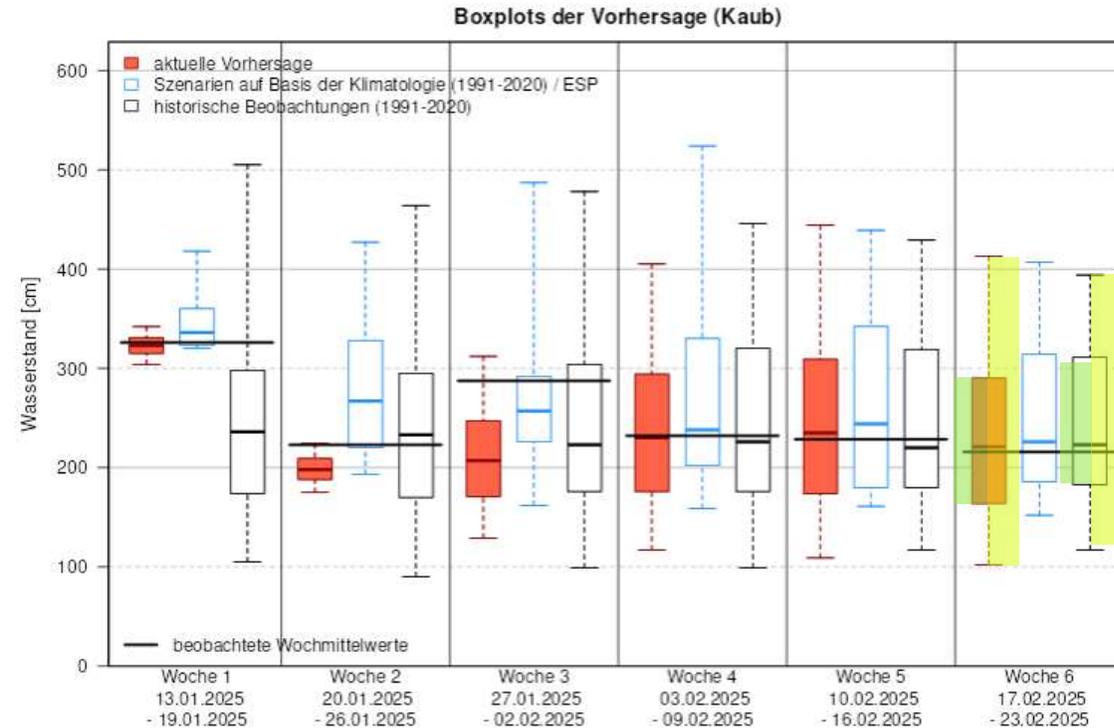
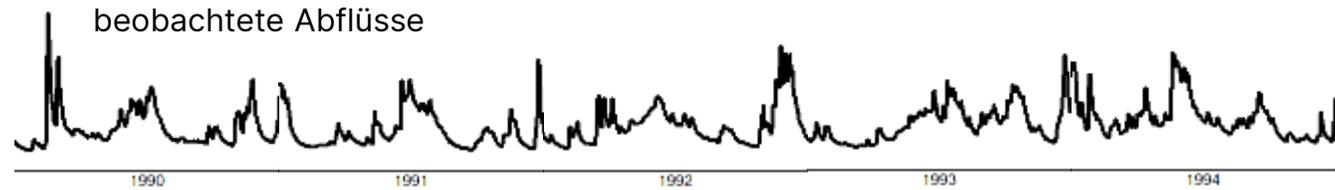


Hindcasts 2016-2023

Continuous Ranked Probability Score (CRPS)
 =
 mittlere absolute Abweichung der Vorhersage
 von der Beobachtung

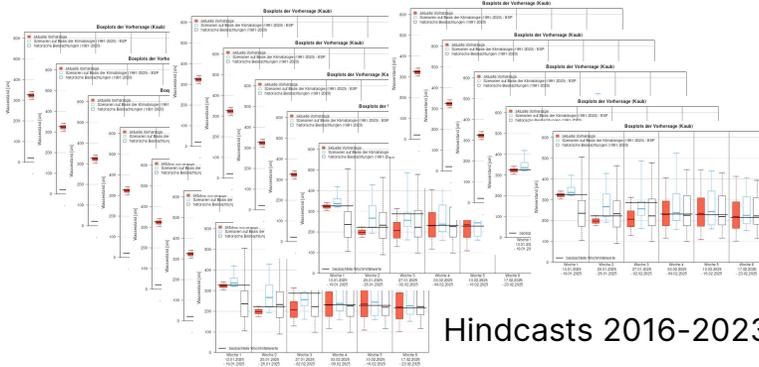


Fehlende Vergleichbarkeit
 Fehlende Bewertung und Einordnung



Unsicherheits-
 band
 Bereich höchster
 Wahrscheinlichkeit

Berechnung der Vorhersagegüte



Hindcasts 2016-2023

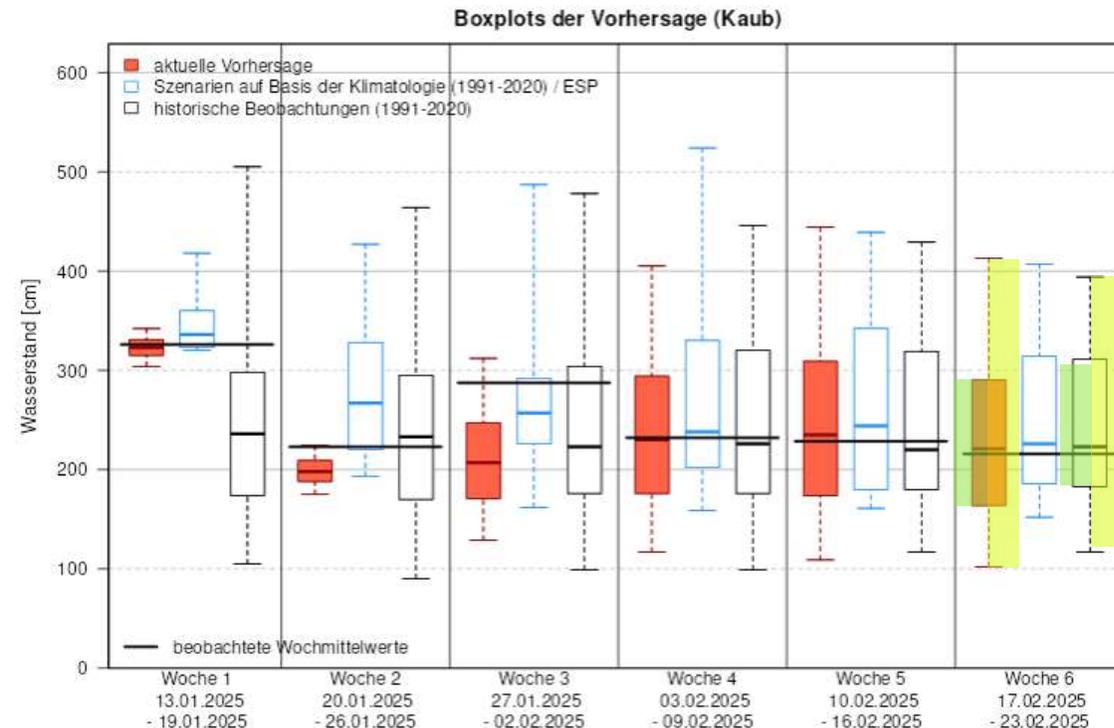
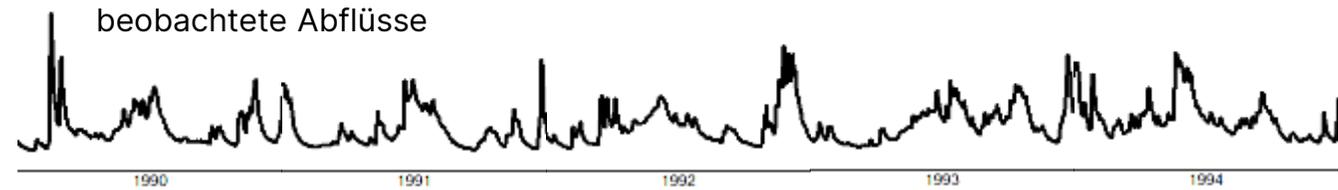
Boxplotvorhersage: Continuous Ranked Probability Skill Score (CRPSS)

5-Klassen-Vorhersage: Ranked Probability Skill Score (RPSS)

Unterschreitungsvorhersage: Brier Skill Score (BSS)

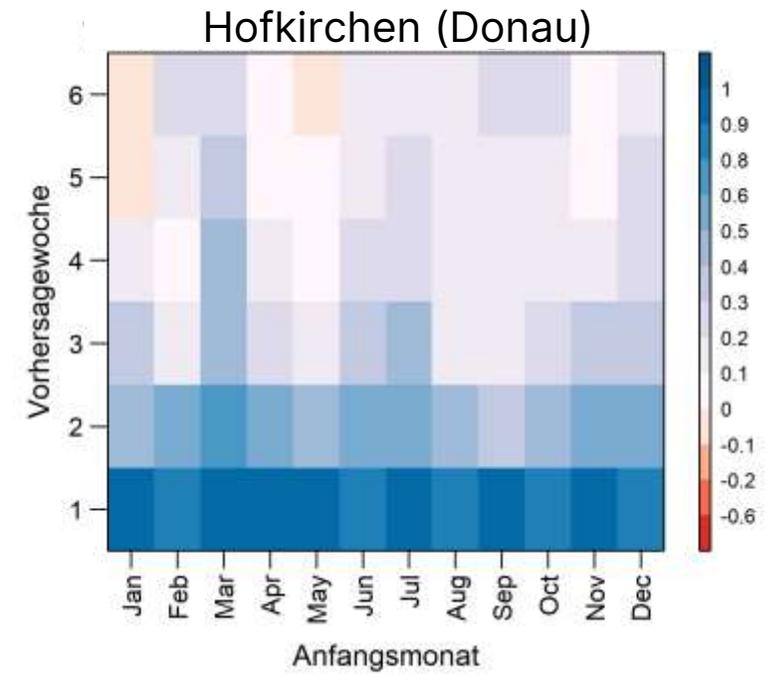
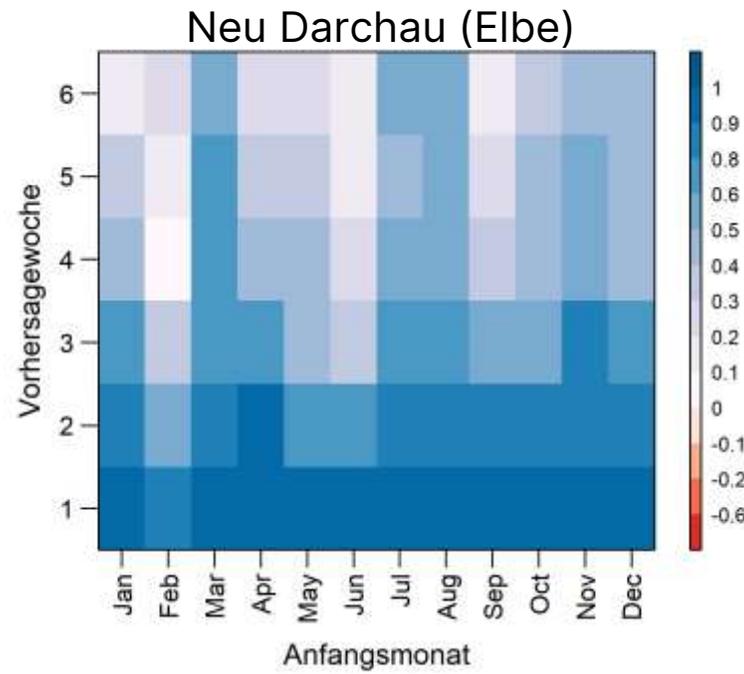
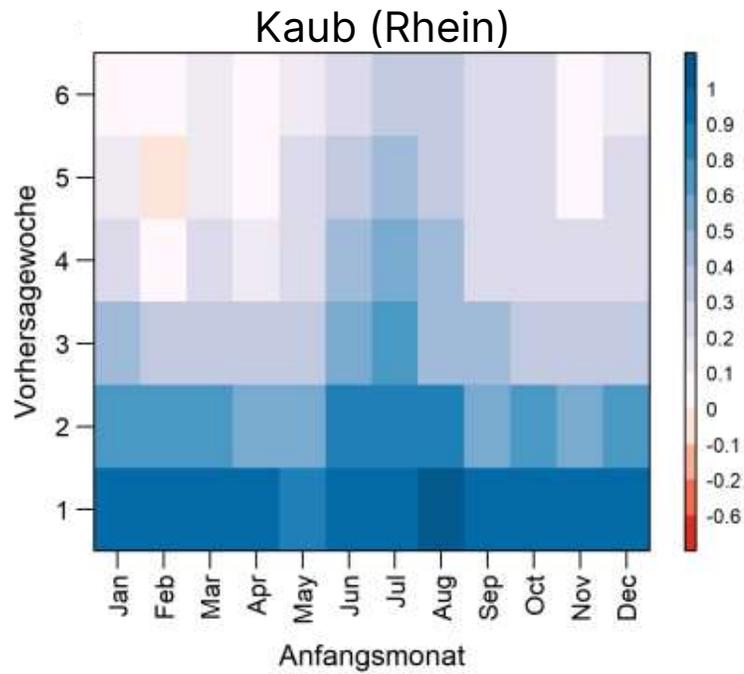
Für alle gilt:

- 1 = perfekte Vorhersage
- > 0 = Vorhersage besser
- < 0 = Referenzvorhersage besser



Unsicherheitsband
Bereich höchster Wahrscheinlichkeit

Zusatzinfos Gütemaße: <https://confluence.ecmwf.int/display/FUG/Section+12.B+Statistical+Concepts+-+Probabilistic+Data>



Im Jahresmittel haben alle 6 Vorhersagewochen Skill.

Vorhersagen: 2016-2023
Referenzvorhersage:
Beobachtungen von 1991-2020

Verschiedene Ampelphasen

 sehr gute Vorhersagequalität

Gütemaß $> 0,75$.

Die besten 10-15% der Vorhersagen der Vergangenheit aus dem Evaluierungszeitraum (2016-2023) erreichen diesen Wert.

=> Die roten Boxplots haben eine sehr gute Vorhersagequalität.

 geringe Vorhersagequalität

Gütemaß > 0 , aber nicht signifikant

Gütemaß < 0 , aber nicht signifikant

=> Die roten und schwarzen Boxplots zeigen die gleiche Vorhersagequalität.

 gute Vorhersagequalität

Gütemaß > 0 und $< 0,75$

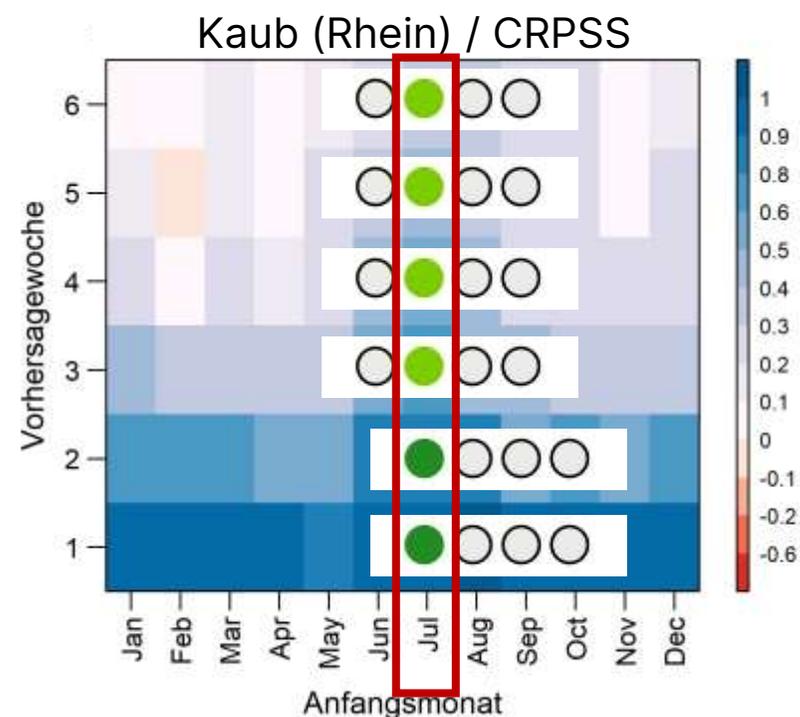
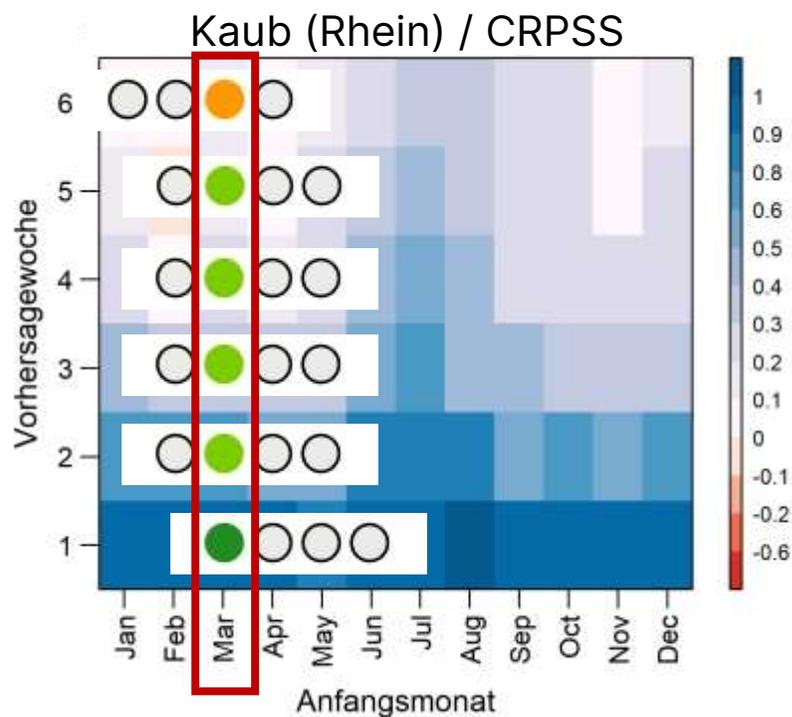
=> Die roten Boxplots sind den schwarzen Boxplots klar überlegen und zeigen eine gute Vorhersagequalität.

 schlechte Vorhersagequalität

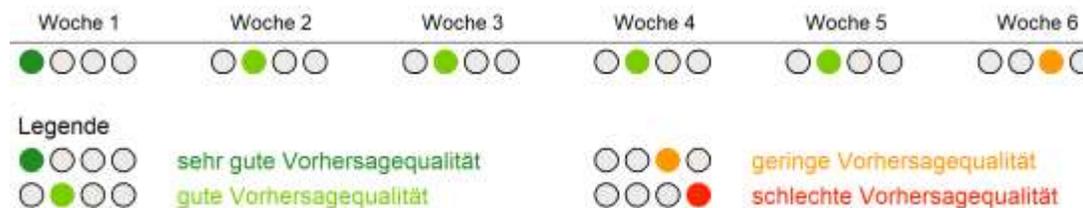
Gütemaß < 0

=> Die schwarzen Boxplots bieten statistisch gesehen die bessere Vorhersagequalität als die roten Boxen.

Wie entsteht die Ampel



Güte-Ampel der Boxplotvorhersage
(Vorhersagezeitpunkt im März, Evaluierungszeitraum: 2016-2023)



6-Wochen-Vorhersage Güte-Ampel



Barbara Frielingsdorf
Bundesanstalt für Gewässerkunde
Wasserhaushalt, Vorhersage und Prognosen (M2)

BfG Federal Institute of Hydrology

Maßnahmen oder Speicherregulierungen, die sich zeigen.

Fehlende oder unsichere Beobachtungsdaten:
Wenige oder fehlerhafte hydrologische und meteorologische Daten, die den Ist-Zustand im Modell korrekt abzubilden. Wenn diese Flächen vorhanden sind, wie zum Beispiel das Wasserverdunstung oder die Bodenfeuchte, werden diese in die Berechnung einbezogen, was ebenfalls zu Unsicherheiten für die Anfangsbedingungen im Modell eine entscheidende Rolle spielt.

Wettervorhersage:
Auch die numerischen Wettervorhersagen der Modellkette werden am betrachteten Ort in die Berechnung der 6-Wochen-Vorhersage einbezogen. Zur Berechnung der Güte der 6-Wochen-Vorhersage werden die Mittelwerte der 6-Wochen-Vorhersage mit den Mittelwerten der Beobachtungen verglichen.

Berechnung der Güte:
Um trotz der Unsicherheiten eine Vorhersage zu machen, ist es notwendig, ein Modell zu verwenden. In der Vergangenheit wurde ein probabilistisches Vorhersagemodell in der Vorhersage verwendet, das auch die zukünftigen Ereignisse abbildet, da viele Faktoren wie Schneeschmelze, Wasserstände, etc. das ganze Jahr über einflussreich sind.

BfG – Hintergründe der 6-Wochen-Vorhersage Güte-Ampel

BfG Bundesanstalt für Gewässerkunde

6-Wochen-Vorhersage Güte-Ampel – Wie entsteht die Güte-Ampel und wie ist sie zu verstehen?

Zusätzlich zur bisherigen Vorhersageabbildung wird nun auch die Güte der Vorhersagen mit Hilfe einer Güte-Ampel dargestellt. Diese zeigt die Vorhersagegüte abhängig vom Vorhersagezeitpunkt im Jahr. Die Ampel ändert sich monatlich, abhängig vom Monat, in dem der Vorhersagezeitpunkt liegt. Neben der Darstellung der Saisonalität der Güte, bietet Sie ebenfalls eine Vergleichs- und Einordnungsmöglichkeit der Qualität der hydrologischen 6-Wochen-Vorhersage von Pegel zu Pegel.

Güte-Ampel der Boxplotvorhersage (Vorhersagezeitpunkt im März, Evaluierungszeitraum: 2016-2023)



Legende:
 ●○○○ sehr gute Vorhersagequalität
 ○●○○ gute Vorhersagequalität
 ○○○● geringe Vorhersagequalität
 ○○○○ schlechte Vorhersagequalität

Abbildung 1: Beispiel einer Güte-Ampel, wie sie im PDF-Bericht abgebildet wird. Die Güte-Ampel ändert sich abhängig vom Monat in dem der Vorhersagezeitpunkt liegt. Hier handelt es sich um eine Vorhersage aus dem März.

Einflussfaktoren der Vorhersagequalität

Jede Vorhersage ist mit Unsicherheiten behaftet. Diese ist, wie die Vorhersage selbst, dynamisch. Grundsätzlich gilt eine Zunahme der Unsicherheit über den Vorhersagezeitraum.

Die Einflussfaktoren sind dabei variabel:

Eigenschaften des Einzugsgebietes:
Hydrologische Vorhersagen für größere Einzugsgebiete sind in der Regel genauer, als für kleinere, da sich hier Fehler der numerischen Wettervorhersage und Anfangsbedingungen eher ausgleichen und mehr meteorologische und hydrologische Messdaten für das Gebiet vorhanden sind. Eine Abflussvorhersage in Einzugsgebieten mit längeren Fließzeiten aufgrund von geringerem Gefälle ist ebenfalls einfacher, als in steilen Einzugsgebieten, da die Niederschlagsmessung für einen längeren Zeitraum die Abflussbildung bestimmt und erst später die Unsicherheit der Wettervorhersage eine Rolle spielt. Die Qualität der hydrologischen Modellierung ist außerdem abhängig von weiteren, komplexen Gebieteigenschaften, wie z. B. Karst, wasserbaulichen

BfG – Hintergründe der 6-Wochen-Vorhersage Güte-Ampel

Legende:
 ●○○○ sehr gute Vorhersagequalität
 ○●○○ gute Vorhersagequalität
 ○○○● geringe Vorhersagequalität
 ○○○○ schlechte Vorhersagequalität



Abbildung 2: Zusammenfassung der hydrologischen Vorhersage für den Zeitraum 2016-2023. Zeigt Boxplots für Beobachtung und Vorhersage.

Die bei der BfG verwendeten Vorhersagen auf Niedrig- und Mittelwasserabflüsse werden mit den Beobachtungen verglichen. Zur Berechnung der Güte der 6-Wochen-Vorhersage werden die Mittelwerte der 6-Wochen-Vorhersage mit den Mittelwerten der Beobachtungen verglichen.

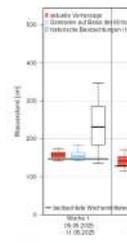


Abbildung 3: Rückblick auf die Vorhersage am Rheingel Pegel vom 13.01.2025. Zeigt Boxplots für Beobachtung und Vorhersage.

Da es sich um eine wahrscheinlichkeitsbasierte Vorhersage handelt, kann eine Wasserstandsganglinie mit einer anderen abgeglichen werden. Für die Güteberechnung probabilistischer Vorhersagen gibt es verschiedene Gewichte unter anderem statistisch abbildbar, ob die Messwerte im Bereich der Wahrscheinlichkeit der Vorhersage liegen und ob das Unsicherheitsband abgebildet wird.

Da die Berechnung einer absoluten Vorhersagegüte nur die Aussage zu der Vorhersage im Mittel von der Beobachtung abweicht, aber keine Werte dazwischen, wird für die Güte-Ampel eine relative Vorhersagegüte berechnet, die auf Grundlage vergangener Vorhersagen von 2016-2023. Hier absolute Vorhersagegüte (die mittlere Abweichung der roten Boxplots in der Abbildung 3) mit der absoluten Güte einer Referenzvorhersage verglichen. In diesem Fall wird die Verteilung der hydrologischen Beobachtungen (schwarze Boxplots in der Abbildung 3) als Referenzvorhersage genutzt. Sie bildet das natürliche Verhalten des Flusses auf Jahresgang und Maximalabflüsse ab.

BfG – Hintergründe der 6-Wochen-Vorhersage Güte-Ampel

Referenzvorhersage in schwarz. Die Güte-Ampel zeigt die Qualität der Vorhersage ab. Diese ist über das gesamte Jahr verteilt, da sich die Grenzen der Güte, von Woche zu Woche ändern.

Die Güte-Ampel wird in der Vorhersage verwendet. Sie zeigt die Qualität der Vorhersage (CRPSS) und die Unsicherheit (SS).

Sie hier: [Statistical+Concepts+-](#)

Die Güte-Ampel erstellt. Sie zeigt, wie gut die Vorhersage ist.

Für alle verwendeten Gütemaße gilt: Bei einem Wert von 1 ist eine perfekte Vorhersage erreicht. Bei Werten zwischen 1 und 0 ist die aktuelle Vorhersage das beste Werkzeug, bei 0 kann weder die Vorhersage noch die Referenzvorhersage als besser eingestuft werden und bei Werten kleiner 0 ist die Referenzvorhersage der Vorhersage überlegen.

Zusätzlich zur Güteberechnung wird statistisch geprüft, wie wahrscheinlich eine Über- oder Unterschreitung des Grenzwertes 0 ist, also ob diese signifikant ist. Ist sie nicht signifikant, kann sie statistisch gesehen zufällig aufgetreten sein und wird damit nicht bestätigt.

Es ergeben sich folgende Kriterien für die Farben der Ampel (Abbildung 4):
 Sehr gute Vorhersagequalität (dunkelgrün)
 Das Gütemaß zeigt einen Wert größer als 0,75. Diese Grenze ist so gewählt, dass die besten 10-15% der Vorhersagen der Vergangenheit aus dem Evaluierungszeitraum (2016-2023) dies erreichen.
 => Die roten Boxplots haben eine sehr gute Vorhersagequalität.

BfG – Hintergründe der 6-Wochen-Vorhersage Güte-Ampel

Fazit:

- Mittel- und Langfristvorhersagen sind für verschiedene Zwecke gefragt
- Qualitätssicherung ist Arbeit, aber bei komplexen FEWS-Workflows unverzichtbar
- Je komplexer die Vorhersage und je länger der Vorhersagezeitraum, desto wichtiger sind die Informationen zur Vorhersagegüte



Danke für Ihre Aufmerksamkeit!



Dr. Bastian Klein
Bundesanstalt für
Gewässerkunde (BfG)
+49 261 1306-5256
klein@bafg.de



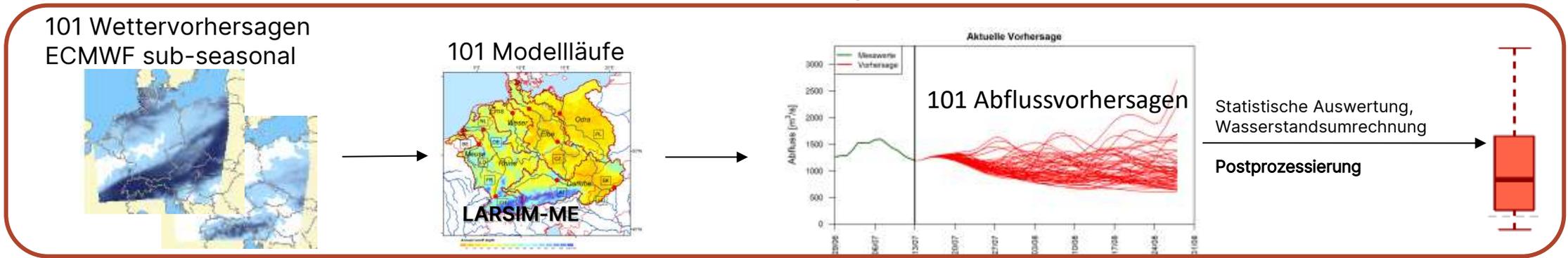
Barbara Frielingsdorf
Bundesanstalt für
Gewässerkunde (BfG)
+49 261 1306-5406
frielingsdorf@bafg.de



Dennis Meißner
Bundesanstalt für
Gewässerkunde (BfG)
+49 261 1306-5183
meissner@bafg.de

Vorhersageworkflow

Aktuelle Vorhersage



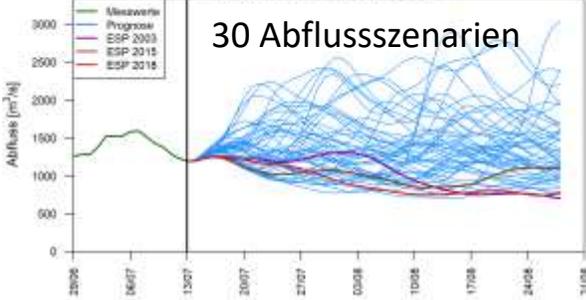
30 meteorologische
Messungen zur gleichen
Zeit im Jahr (1991-2020)



30 Modellläufe



Klimatologische-basierte Prognose / ESP



Statistische Auswertung,
Wasserstandsumrechnung

Postprozessierung

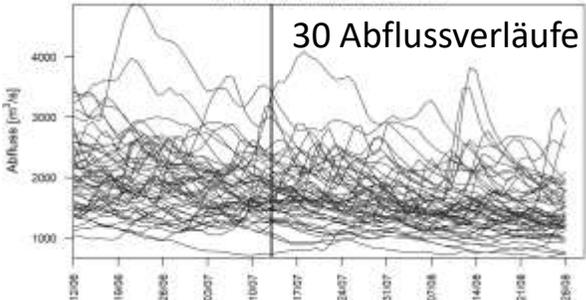


30 Abflussmessungen zur
gleichen Zeit im Jahr
(1991-2020)

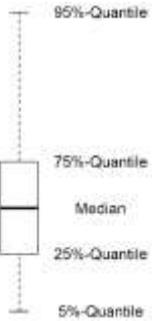


Kein Modelllauf
notwendig

30 Abflussverläufe



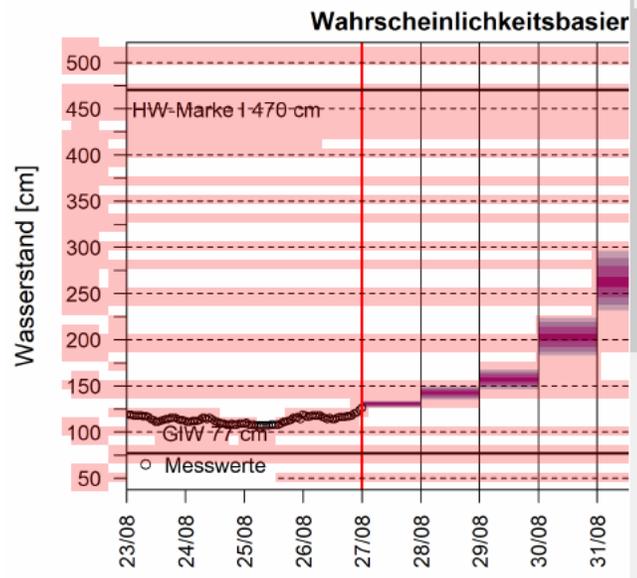
Statistische Auswertung,
Wasserstandsumrechnung



Wahrscheinlichkeitsbasierte 14-Tage Pegel Koblenz

Vorhersage vom: 27.08.2023 00:00 MEZ

Die in der Grafik und in den Tabellen angegebene definierte Wasserstände basieren auf Tagesmit



Unterschreitungswahrscheinlichkeiten definierter

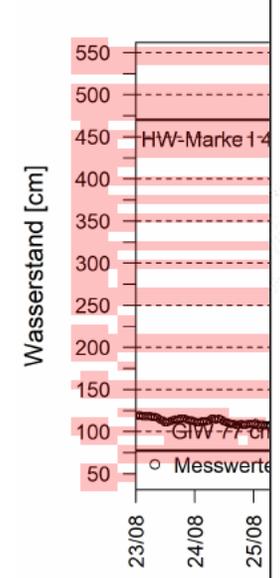
	27.08	28.08	29.08	30.08	31.08	01.09	02.09	03.09
GIW: 77 cm	0	0	0	0	0	0	0	0
97 cm	0	0	0	0	0	0	0	0
117 cm	0	0	0	0	0	0	0	0
137 cm	100	8	0	0	0	0	0	0
157 cm	100	100	50	0	0	0	0	0

Überschreitungswahrscheinlichkeiten HW-Marke

Wahrscheinlichkeitsbasierte 14-Tage Pegel Koblenz

Vorhersage vom: 27.08.2023 00:00 MEZ

Die in der Grafik und in den Tabellen angegebene definierte Wasserstände basieren auf Tagesmit

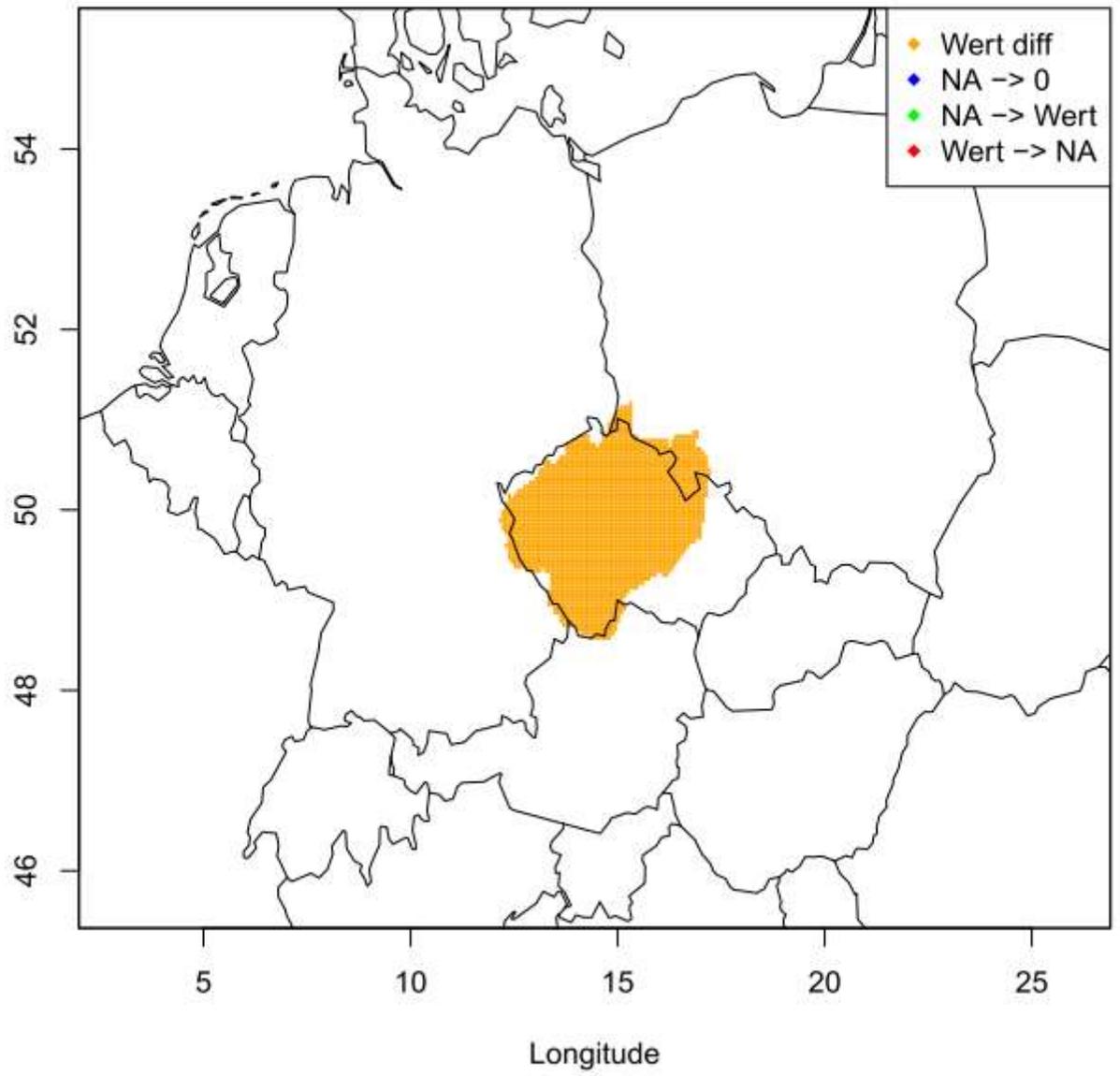


Unterschreitungswahrscheinlichkeiten definierter

	27.08	28.08
GIW: 77 cm	0	0
97 cm	0	0
117 cm	0	0
137 cm	95	3
157 cm	100	100

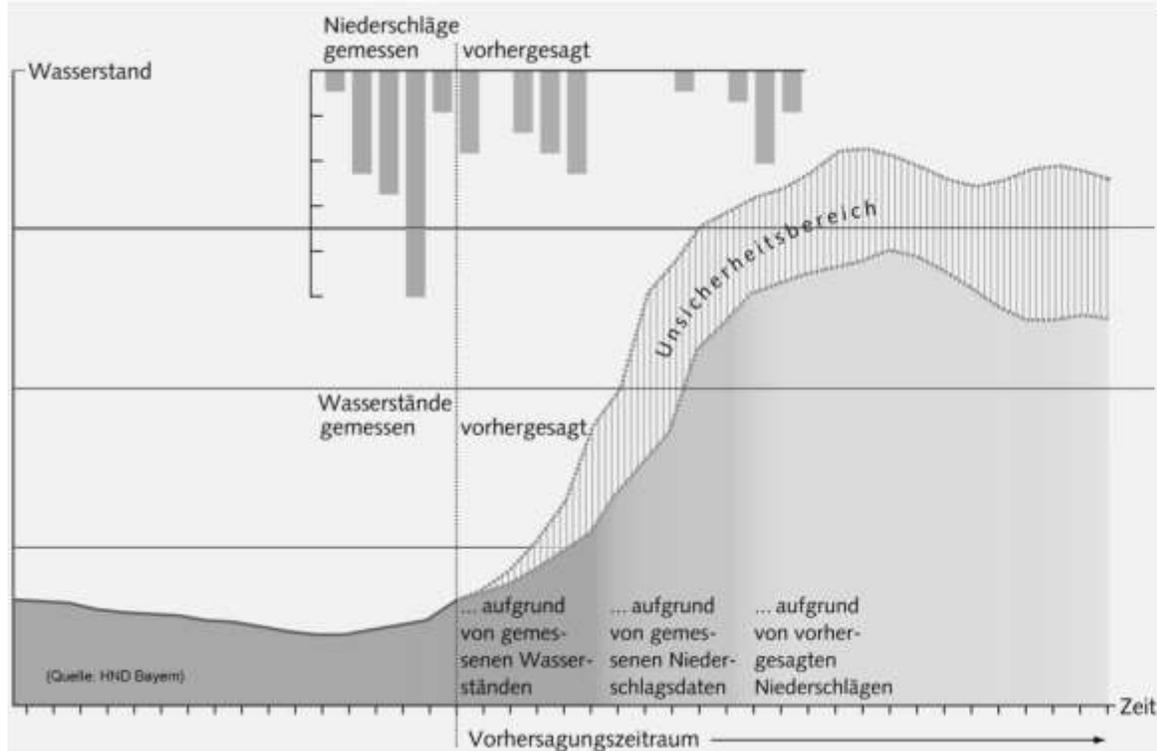
Überschreitungswahrscheinlichkeiten HW-Marke

Zeitschritt 4; Variable = hurs Abweichungen in 4.02% aller Zellen 0 neue NA-Werte

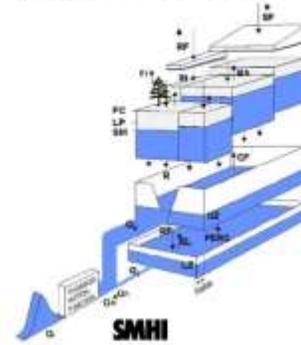


Wieso überhaupt eine Güte-Ampel?

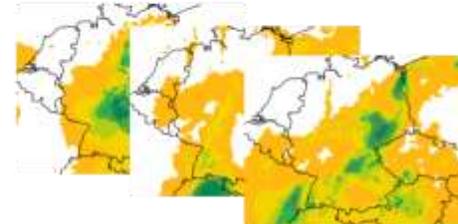
Keine Vorhersage ohne Unsicherheiten



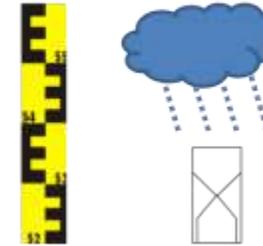
Modellunsicherheiten



Meteorologische Vorhersageunsicherheiten



Messunsicherheiten



Unsicherheit der Anfangszustände



- Die Unsicherheit ist – wie die Vorhersage selbst – dynamisch.
- Grundsätzlich: Zunahme der Unsicherheit über den Vorhersagezeitraum
(Einflussfaktoren dabei variabel: Wetterlage, Abflusssituation, Vorhersagezeitraum, Pegel, ...)
- Unterschiedliche Vorhersageprodukte für unterschiedliche Vorhersageskalen
(Kurzfrist-, Mittelfrist-, Langfristvorhersagen), um der Unsicherheit Rechnung zu tragen