

DWD

REGNIE:

**Regionalisierte Niederschläge
Verfahrensbeschreibung und Nutzeranleitung**



Abteilung Hydrometeorologie



1.	Einleitung	2
2.	Produktpalette	3
3.	Datengrundlage	3
4.	Methode.....	4
4.1	Datenqualität	4
4.2	Interpolationsverfahren	4
5.	Gitter.....	7
6.	Formatbeschreibung	7
7.	Bekannte Probleme	8
8.	Wichtige Hinweise zur Nutzung.....	8
9.	Bezug der Daten	9
10.	Zitat.....	9



1. Einleitung

Rasterdatensätze von Messungen klimatologischer Variablen, wie z. B. Niederschlag sind von großer Bedeutung. Nur mit ihnen ist es möglich Klimamodelle zu validieren oder hydrologische Modelle zu betreiben. Dabei ist vor allem für letzteres eine möglichst hohe horizontale Auflösung auf regionaler Ebene notwendig. Um die Rasterdatensätze zu berechnen, werden die Messwerte von irregulär im Raum verteilten Stationen mit Hilfe von geeigneten Regionalisierungsverfahren auf ein reguläres Gitter von 1 km² interpoliert, wodurch u. a. die Bestimmung von Gebietsmitteln der einzelnen Variablen vereinfacht wird.

Das unter dem Namen REGNIE (Regionalisierung der Niederschlagshöhen) laufende Verfahren umfasst die Regionalisierung aller beim DWD vorhandenen Niederschlagsdaten Deutschlands ab 1891. Wobei das Hauptprodukt, die Rastertageswerte, seit 1931 zur Verfügung steht.

Im Folgenden soll ein Überblick über das Verfahren REGNIE und dessen Produktpalette gegeben werden, sowie eine Anleitung zur praktischen Nutzung der Daten.

2. Produktpalette

- Raster- und Gebietswerte der täglichen Niederschlagshöhe ab 1931 für die alten Bundesländer und ab 1951 für die neuen Bundesländer
- Raster- und Gebietswerte der täglichen korrigierten Niederschlagshöhe ab 1961, bei denen die durch Benetzungs-, Verdunstungs- und Überwehungseffekte systematisch auftretenden Niederschlagsverluste rechnerisch ausgeglichen wurden
- Raster- und Gebietswerte der monatlichen und jährlichen Niederschlagshöhe ab 1891
- Raster- und Gebietswerte der monatlichen und jährlichen Niederschlagshöhe, bezogen auf den 30-jährigen Zeitraum 1961 bis 1990 und 1971 bis 2000
- Raster- und Gebietswerte der monatlichen und jährlichen Anzahl der Tage mit Niederschlag oberhalb eines Schwellenwertes

3. Datengrundlage

Für den Datensatz REGNIE-Deutschland werden alle haupt- und nebenamtlichen Stationen des DWD sowie alle verfügbaren Länderstationen herangezogen. Dabei ist zu beachten, dass die Anzahl der Stationen von Jahr zu Jahr schwankt. So basieren die Rasterwerte jeweils auf einer sehr unterschiedlichen Anzahl an echten Beobachtungen. Besonders in den Jahren vor 1940 und in den Kriegsjahren 1944 und 1945 ist die Anzahl der Stationen deutlich geringer als nach 1950. Seit 1995 nimmt die Anzahl der Stationen wieder leicht ab. Zurzeit (Stand 2011) fließen gut 2000 Stationen in die REGNIE-Datensätze ein.

Die für das Verfahren notwendige Höhenrasterdatei (DHM M475) stammt vom Amt für Geoinformationswesen der Bundeswehr, das in einer 1km-Auflösung verwendet wurde.

4. Methode

Im Folgenden werden kurz die verwendeten Verfahren beschrieben. In REGNIE werden Niederschlagsdaten verwendet, die bereits im DWD eine Datenprüfung durchlaufen haben.

4.1 Datenqualität

Die Daten des DWD werden durch automatische und halbautomatische Programmsysteme (Qualimet und Qulako) qualitätsgeprüft. Diese Methoden beinhalten eine:

- Vollständigkeitsprüfung
- klimatologische Konsistenzprüfung
- zeitliche Konsistenzprüfung
- innere Konsistenzprüfung
- räumliche Konsistenzprüfung

Als Ergebnis dieser Teilprüfungen werden die durch die Prüfalgorithmen angezweifelten Mess- und Beobachtungswerte markiert und müssen vom Bearbeiter überprüft werden. Dabei sind folgende Aktionen grundsätzlich möglich:

- Daten durchlaufen die Prüfungen ohne Beanstandung.
- Daten werden korrigiert.
- Vorhandene Daten werden gelöscht.
- Fehlende Daten werden eingefügt.
- Angezweifelte Daten werden bestätigt.
- Datenlücken werden bestätigt.

4.2 Interpolationsverfahren

Das Regionalisierungsverfahren REGNIE zur Regionalisierung von beobachteten täglichen Niederschlagshöhen basiert im Wesentlichen auf der Interpolation von Anomalien in Bezug auf langjährige Mittelwerte und besteht aus zwei Teilschritten: Berechnung der Hintergrundfelder und Berechnung der Niederschlagsfelder.



Berechnung der Hintergrundfelder

Zunächst basiert die Berechnung der Hintergrundfelder auf dem Ansatz der multiplen linearen Regression (MLR):

$$y_i = a_0x_{i0} + a_1x_{i1} + a_2x_{i2} + a_3x_{i3} + a_4x_{i4} + a_5x_{i5} + \varepsilon_i$$

mit

x_{ik} : k Einflussgrößen an der Station i mit $x_{i1} \equiv 1$

y_i : Zielgröße an der Station i

a_{ik} : Regressionskoeffizient an der Station i für die Einflussgröße k

ε_i : zufälliger Anteil von y_i bzw. Residuum der Station i

Die 5 Einflussgrößen (k) setzen sich aus:

- geographischen Länge und Breite,
- Höhe über NN und
- Richtung und des Betrags der Exposition

einer Station zusammen. Die Geländeexposition bezieht sich auf die Neigungsrichtung von Hängen, welche über Höhenangaben benachbarter Punkte (bzw. Raster) bestimmt werden kann, und sollte gerade bei hoher Auflösung (1km²) in der Regression verwendet werden. Die Ausrichtung der Hänge relativ zur Anströmung der Luftmassen kann für die Niederschlagsmenge relevant sein. Diese Abhängigkeit ist mit zwei Einflussgrößen in der multiplen linearen Regression berücksichtigt. Die Zielgröße ist der monatliche Bezugswert einer Station für den Zeitraum 1961 bis 1990, der auf Basis von als homogen befundenen Zeitreihen berechnet wurde. Für das Hintergrundfeld kommen also nur die Stationen in Betracht, die über homogene Monatsreihen verfügen.

Die Einflussgrößen und Zielgröße einer Station sind bekannt. Die Regressionskoeffizienten sind unbekannt, können aber nach der sogenannten Methode der kleinsten Quadrate durch die Lösung des Normalgleichungssystems berechnet werden. Bei der Berechnung werden alle verfügbaren Stationen (mindestens 10) in einem Gebiet verwendet, so dass man einen Satz Koeffizienten pro Gebiet und pro betrachteten Zeitraum (z. B. Kalendermonat, Jahr) als Lösung des Gleichungssystems erhält.

Um das Hintergrundfeld zu erhalten, werden die Residuen der MLR auf die nicht besetzten Rastermittelpunkte mit der nachfolgenden Gleichung entfernungsgewichtet interpoliert. Berücksichtigt werden hierbei alle Stationen in einer Entfernung von 20 km zum Zielraster. Können in diesem Radius keine Station gefunden wird, wird der Radius auf 30 km vergrößert.

$$\tilde{z}_i = \frac{\sum_{k=1}^n \frac{z_k}{d_{ki}^2}}{\sum_{k=1}^n \frac{1}{d_{ki}^2}}$$

mit

\tilde{z}_i : interpolierter Wert für unbesetztes Raster i

z_k : Residuum für (besetztes) Raster k

d_{ki} : Abstand zwischen den Rastern k und i bezogen auf die Rastermittelpunkte

n : Anzahl der nächstgelegenen Stationen, die für die Interpolation herangezogen werden

Damit hat man für alle Raster entweder das „Original-Residuum“ vorliegen, wenn eine Station in dem Raster liegt, oder ein interpoliertes Residuum, wenn keine Station in dem Raster liegt. Schließlich wird für jeden Rastermittelpunkt der Regressionswert zum Residuum aufaddiert. Auf diese Weise entsteht das Hintergrundfeld.

Für die Regressionsrechnungen spielt die Gebietseinteilung eine wesentliche Rolle. Auf Grundlage eines Einzugsgebietsdatensatz werden nach und nach die Gebiete so zusammengelegt, dass die Güte der Regression ausgedrückt durch die erklärte Varianz optimal wird. Gründe für ein Zusammenlegen von Gebieten sind zum einen, dass in einem Gebiet weniger als 10 Stationen liegen und die Regression somit instabil ist und zum anderen dass die erklärte Varianz verglichen mit einer anderen Gebietskombination zu niedrig ist. Insgesamt treten in vielen Gebieten erstaunlich hohe Werte von 60 % bis 90 % für die erklärte Varianz auf. In einigen Gebieten sind es immerhin noch 40% bis 50% (Norddeutschland).

Berechnung der Niederschlagsfelder

Nach der Methode REGNIE werden tägliche, monatliche und jährliche Rasterdatensätze und Gebietsniederschlagshöhen in verschiedenen Auflösungen unter Verwendung der berechneten Hintergrundfelder in drei Schritten berechnet:

- Der Niederschlagswert eines Einzeltages, Einzelmonats oder Einzeljahres einer Station wird auf den nächstgelegenen Rastermittelpunkt abgebildet. Dann wird der Quotient aus diesem Niederschlagswert und dem Wert des Hintergrundfeldes (des gewählten Kalendermonats oder Ganzjahres) gebildet.
- Die dimensionslosen Quotienten werden entfernungsgewichtet auf nicht besetzte Rastermittelpunkte interpoliert, wobei die vier nächstgelegenen Stationen zum Zielraster verwendet werden.
- Die Berechnung der räumlichen Niederschlagsverteilungen erfolgt durch Multiplikation der nun pro Raster vorliegenden dimensionslosen Quotienten mit den Rasterwerten des entsprechenden Hintergrundfeldes.



Die Quantifizierung der Gebietsniederschlagshöhen erfolgt je nach Bedarf durch Summation und Mittelbildung der jetzt in REGNIE-Rasterung vorliegenden Niederschlagshöhen in täglicher, monatlicher und jährlicher Auflösung.

5. Gitter

Das Gitter besteht aus 611 Rasterpunkten in West/Ost-Richtung und 971 Rasterpunkten in Nord/Süd-Richtung.

Die Auflösung beträgt 60 geogr. Sekunden längenparallel und 30 geogr. Sekunden breitenparallel, d.h. 1 Längengrad setzt sich aus 60 Gitterpunkten und 1 Breitengrad aus 120 Gitterpunkten zusammen.

Die geographischen Koordinaten der Rasterpunkte lassen sich wie folgt berechnen:

Schrittweite längenparallel (xdelta) : 60 geographische Sekunden
 $xdelta = 1.0 * (1./60.)$ (Grad)

Schrittweite breitenparallel (ydelta) : 30 geographische Sekunden
 $ydelta = 1.0 * (1./120.)$ (Grad)

Die geographischen Koordinaten (x,y) eines beliebigen Rasterpunktes R(m,n) :

$$x(m,n) = (6. - 10.*xdelta) + (n - 1)*xdelta$$
$$y(m,n) = (55. + 10.*ydelta) - (m - 1)*ydelta,$$

wobei x in West/Ost-Richtung (n von 1 bis 611) und y in Nord/Süd-Richtung (m von 1 bis 971) verläuft.

In den Daten sind die Rasterfelder zeilenweise von Nord nach Süd und West nach Ost abgespeichert so dass sie sich wie folgt mit FORTRAN auslesen:

```
INTEGER NDH(971,611)
DO I=1,971
  WRITE (2,'(611i4)') (NDH(I,J),J=1,611)
ENDDO
```

Der Bezugspunkt der Daten ist immer der Rastermittelpunkt.

6. Formatbeschreibung

Nichtbesetzte Rasterpunkte sind mit Fehlwerten -999 besetzt. Die Dimension der monatlichen und jährlichen Niederschlagshöhen beträgt mm, die der täglichen mm/10.



Die täglichen REGNIE-Raster werden jeweils pro Tag in einer ASCII-Datei (rayymmdd) abgelegt (Gitterdefinition siehe 4.3). Die Einzeljahre werden in einem TAR-Archiv (rayyyym.tar) zusammengefasst. Die einzelnen Tage sind zusätzlich gezippt. Beide Komprimierungen können mit Standardtools unter LINUX und WINDOWS entpackt werden.

Die monatlichen REGNIE-Raster werden jeweils pro Monat in einer ASCII-Datei (rasayymm) abgelegt. (Gitterdefinition siehe 4.3). Die Einzeljahre werden in einem TAR-Archiv (rasyy.tar) zusammengefasst. Zusätzlich enthält das Archiv noch das 2 Jahresraster bei dem die Monatskennung 13 ist, das eine die absoluten Jahreswerte (rasayy13) sind und das andere die relativen Werte im Vergleich zu 1961-1990 (rasryy13) sind. Alle Einzelraster sind zusätzlich auch noch gezippt.

7. Bekannte Probleme

1. Die Anzahl der Stationen schwankt von Jahr zu Jahr und beeinflusst die Qualität der Rasterdaten, da Niederschlagsfelder häufig, vor allem im Sommer, räumlich sehr inhomogen sind.
2. Ein Knackpunkt der Herstellung der Rasterdaten ist die Berechnung der Hintergrundfelder. Von deren Qualität hängt die des Endproduktes im Wesentlichen ab. Dabei ist vor allem in den Randbereichen die Bestimmung der Regressionskoeffizienten durch die nicht vorhandenen Daten außerhalb Deutschlands unsicher.

8. Wichtige Hinweise zur Nutzung

- Messzeitraum: 07:30 GZ bis 07:30 GZ (gesetzliche Zeit) Folgetag, d.h. jeweils MEZ bzw. MSEZ im Winter bzw. Sommer
- Ab dem 01.01.1971 ist die Niederschlagshöhe auf den Tag bezogen, an dem der Niederschlag gefallen ist (ein Tag bevor dieser gemessen wurde). Vor 1971 ist der Tag Bezugstag, an dem gemessen wurde. Um eine kontinuierliche Zeitreihe zu bekommen, müssen die Raster vor 1971 um einen Tag zurückdatiert werden.
- Es wird routinemäßig täglich ein vorläufiges Produkt erstellt (dieses hat den Zeitbezug 5:50 UTC bis 5:50 UTC), das alle automatisch meldenden Stationen enthält. Die übrigen Daten werden nach Eingang und Qualitätskontrolle nachträglich eingepflegt und damit die endgültigen Rasterdaten mit dem Zeitbezug (07:30-07:30 GZ) erstellt.



9. Bezug der Daten

Die REGNIE-Rasterdaten (Tages-, Monats und Jahreswerte) sind für Behörden sowie Forschung und Lehre nach Bezahlung eines Zugangsentgelts in WebWerdis (<http://www.dwd.de/webwerdis>) entgeltfrei zu erhalten. Im WebWerdis stehen zurzeit die täglichen Daten nach endgültiger Erstellung direkt zur Verfügung. Die weitere Produktpalette sowie alle REGNIE-Produkte für sonstigen Kunden sind auf Anfrage bei hydromet@dwd.de erhältlich. Bei weiterführenden Fragen wenden Sie sich bitte auch an diese Adresse.

10. Zitat

Die Methode REGNIE ist inzwischen wissenschaftlich veröffentlicht, bitte folgendes zitieren:

Rauthe, M., H. Steiner, U. Riediger, A. Mazurkiewicz and A. Gratzki, 2013: A Central European precipitation climatology – Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS), Vol. 22(3), p 235–256, DOI:10.1127/0941-2948/2013/0436.

Des Weiteren kann auch dieser interne Bericht zitiert werden:

Abteilung Hydrometeorologie: REGNIE (REGionalisierte NIEederschläge): Verfahrensbeschreibung & Nutzeranleitung, interner Bericht im DWD, Offenbach 2017.