

Hydrologische Modellierung zur Vorhersage von Niedrigwasser auf der Mikro- bis Mesoskala

Eine Evaluation am Beispiel des Kantons Aargau

Geo 511: Masterarbeit am Geographischen Institut der mathematisch-
naturwissenschaftlichen Fakultät, Universität Zürich

von

Isabelle Nanz

BSc in Geografie

09-705-419



Betreut durch

Prof. Dr. Jan Seibert
Institutsvertreter
Leiter Abteilung H2K
Winterthurerstr. 190
8057 Zuerich
jan.seibert@geo.uzh.ch

Dipl.-Hydrol. Maria Staudinger
Betreuerin
Doktorandin H2K
Winterthurerstr. 190
8057 Zuerich
maria.staudinger@geo.uzh.ch

Dr. Christophe Lienert
Co-Betreuer
Leiter Hydrometrie Kanton AG
Entfelderstrasse 22
5001 Aarau
christophe.lienert@ag.ch

Abgabe: 30. 6. 2014

Titelbild: Die Aare bei Brugg am 25. Mai 2011. Quelle: W. Lehmann, ALG Wasserbau, BVU, Kanton Aargau.

Zusammenfassung

An 36 Einzugsgebieten, deren Messstellen sich im Kanton Aargau befinden, wurde die Vorhersagefähigkeit des HBV-Modells für Niedrigwasserereignisse getestet. Dabei wurden nur kleine bis mittelgrosse Gewässer betrachtet; die untersuchten Gebiete weisen daher Flächen zwischen 0.6 und gut 350 km² auf.

In einer ersten Phase wurde die Standardversion des HBV-Modells verwendet, die in einem Split-Sample-Verfahren kalibriert und anhand einer Schwellenwertmethode für den Sommer 2003 und das Frühjahr 2011 ausgewertet wurde. Dabei erreichten zwei Drittel der untersuchten Gebiete ein gutes bis sehr gutes Ergebnis. Unbefriedigende Resultate wurden nur für Gebiete erreicht, die klein (< 30 km²) sind und zumindest einen gewissen Anteil an verkarsteten Flächen enthalten. Um auch für diese übrigen Gebiete zufriedenstellende Resultate zu erreichen, wurden zwei Dinge untersucht: Erstens wurde getestet, ob andere Modellstrukturen bessere Ergebnisse erzielen können, wobei der Untersuchungsfokus auf der Variante „Delay“ lag; zweitens wurde der Einfluss der potentiellen Evapotranspiration (PET) ausgelotet. Dabei liessen sich tendenziell leicht bessere Resultate mit dem Delay-Modell erzielen, einen sehr ausgeprägten Einfluss zeigten aber die PET-Werte. Für die meisten Gebiete konnte durch eine Erhöhung der PET-Werte bis zu einem Faktor von 2 eine markante Verbesserung der Modellgüte erzielt werden; verschiedene Gebiete erreichten dadurch den Bereich einer guten bis sehr guten Modellierung.

In einer zweiten Phase wurde untersucht, ob sich die Niedrigwasservorhersagen durch den zusätzlichen Einbezug von Grundwasserdaten verbessern lassen. Dieses Verfahren führte allgemein zu einer leichten Reduktion der Modellgüte und auch die „Probability Of Detection“ und die „Frequency Of Misses“ verschlechterten sich; eine visuelle Betrachtung der Kurven zeigte auf, dass der Einbezug von Grundwasserdaten in der Tendenz zu einer Erhöhung der Niedrigwassersimulationen führte. Je nach Ausgangslage bewirkte dies eine Verbesserung oder Verschlechterung der Vorhersagen. Es lässt sich aber annehmen, dass gleichzeitig eine bessere interne Modellstimmigkeit erzielt werden konnte und damit eine grössere Verlässlichkeit für Prognosen erreicht würde.

Das HBV-Modell scheint in den untersuchten Gebieten also durchaus Potential für aussagekräftige kurz- bis mittelfristige Niedrigwasserprognosen aufzuweisen. Der Einbezug von alternativen Modellstrukturen, von Grundwasserdaten und eine Erhöhung der PET-Werte können dabei auch in problematischen Gebieten deutliche Modellverbesserungen bewirken.

Danksagung

Zum Gelingen dieser Arbeit haben verschiedenste Personen beigetragen, denen an dieser Stelle Dank gebührt. Zunächst wäre diese Arbeit ohne den Anstoss durch Dr. Norbert Kräuchi, Leiter der Abteilung Landschaft und Gewässer des Kantons Aargau, nie zustande gekommen. Von Seiten des Kantons dienten mir ausserdem verschiedene Personen als Anlaufstelle und Informationsquelle. In diesem Sinne danke ich besonders Dr. Peter Berner, Silja Kempf, Dr. Daniel Schaub sowie Thomas Gebert. Besten Dank ebenfalls an Werner Lehmann und Hans-Peter Nussbaum für das Zustellen von passenden Fotos.

Ein besonderer Dank gebührt auch Marc Vis, der anlässlich dieser Arbeit die neue HBV-Variante mit Grundwasserdaten-Integration programmiert hat.

Einen grossen Beitrag zum Abschluss dieser Arbeit hat Susanne Nanz als Korrektorin geleistet; aber auch Katharina Zimmermann, Katharina Kaelin und Manuela Hasler John danke ich dafür, dass sie Teile der Arbeit gegengelesen haben.

Schliesslich danke ich meinen Betreuungspersonen: Allen voran meiner Betreuerin Maria Staudinger für die Unterstützung, die hilfreichen Anregungen und die Zeit, die sie in die Betreuung dieser Arbeit investiert hat. Dr. Christophe Lienert hat mir in verschiedenen datentechnischen Fragen als Ansprechperson gedient, vielen Dank dafür. Prof. Dr. Jan Seibert danke ich besonders für die zwischenzeitliche Vertretung von Maria Staudinger und darüber hinaus für die vielen Antworten und Anregungen und die immer wieder motivierenden Rückmeldungen.

Inhalt

1.	Einleitung.....	7
2.	Definitionen.....	9
2.1.	Trockenheit.....	9
2.2.	Niedrigwasser.....	10
2.2.1.	Q ₃₄₇ und Mindestrestwassermenge	10
2.3.	Evapotranspiration	11
2.4.	Skalengrößen.....	12
3.	Modell und Gütekriterien.....	13
3.1.	Warum HBV?	13
3.2.	Modellstruktur und Parameter	14
3.3.	Modellannahmen	17
3.4.	Benötigte Input-Dateien.....	18
4.	Datengrundlage	19
4.1.	Abflussdaten.....	19
4.1.1.	Messart.....	19
4.1.2.	Qualitätsprüfung	19
4.1.3.	Einzugsgebiete: Auswahl und Charakteristika.....	20
4.1.4.	Quellen möglicher Fehlinformationen	25
4.2.	Grundwasserdaten	26
4.2.1.	Stationsauswahl und Datenaufbereitung.....	26
4.2.2.	Korrelationen und Dateninterpolation.....	29
4.3.	Niederschlags- und Temperaturdaten	30
4.4.	Evapotranspiration	32
4.5.	Höhenkorrektur	33
4.6.	Stationarität der Datenreihen	33
5.	Methodik	38
5.1.	Vorhersagbarkeit	38
5.2.	Schwellenwert-Methode.....	38
5.3.	Kontingenztabellen.....	40
5.4.	Auswertungsperioden	40
5.4.1.	Sommertrockenheit 2003.....	40
5.4.2.	Frühjahrestrockenheit 2011	41
5.4.3.	Vergleich der beiden Ereignisse	42
5.5.	Modellkalibrierung und Gütekriterien	42
5.5.1.	Genetic Algorithm mit Powell-Optimierung (GAP)	43
5.5.2.	Zielfunktionen.....	44
5.5.3.	Modellparameter	47

5.6.	Parameter-Sensitivität.....	49
5.7.	Alternative Modellstrukturen für Karst-Gebiete.....	50
5.8.	Grundwasserdaten zur Verbesserung der Vorhersage	54
5.8.1.	Motivation zur Multi-Kriterien-Analyse	54
5.8.2.	Rangkorrelation nach Spearman als zusätzliche Zielfunktion	55
6.	Ergebnisse und Auswertungen.....	57
6.1.	Standard-Modell.....	57
6.1.1.	Modellgüte	57
6.1.2.	Trockenheits-Masse	57
6.1.3.	Kontingenz-Tabellen.....	69
6.1.4.	Zusammenhang zwischen Modellgüte, Karstanteil und Einzugsgebietsgrösse	69
6.1.5.	Modellgüte im Verlauf der GAP-Berechnung.....	70
6.2.	Sensitivitäts-Analyse (Standard-Modell).....	72
6.2.1.	Analyse 1: Vergleich von 10 Kalibrierungen.....	72
6.2.2.	Analyse 2: Vergleich mit Gütemass-Variation	73
6.2.3.	Einfluss des Split-Sample-Test-Verfahrens.....	74
6.3.	Alternative Modellstrukturen für Karst-Gebiete.....	76
6.4.	Einfluss der potentiellen Evapotranspiration (PET).....	82
6.5.	Grundwasser-Co-Kalibrierung	86
6.5.1.	Trockenheitsmasse und Abflussverhalten.....	86
6.5.2.	Auswertungen der Kontingenztabellen.....	92
6.5.3.	Delay-Modell mit Grundwasser-Co-Kalibrierung	93
7.	Diskussion.....	96
7.1.	Vorhersagen des Standard-Modells	96
7.2.	Sensitivitätsanalyse	98
7.3.	Karstgebiete und Modellvarianten.....	99
7.4.	Tests zur potentiellen Evapotranspiration	100
7.5.	Co-Kalibrierung auf Grundwasserdaten	101
7.6.	Weitere Unsicherheiten	102
7.7.	Anwendbarkeit und Möglichkeiten zur Weiterentwicklung	104
8.	Schlussfolgerung und Ausblick	106
	Literatur.....	107
	Abbildungsverzeichnis.....	113
	Tabellenverzeichnis	116
	Verzeichnis vermerkter Internetquellen	117
	Anhang	118

1. Einleitung

Im Zuge des Klimawandels weisen die Prognosen darauf hin, dass in der Schweiz insbesondere im Sommer künftig zunehmend mit Trockenperioden zu rechnen ist (CH2011, 2011; BAFU, 2012). Bereits gibt es erste Anzeichen für diese Entwicklung: Der Sommer 2003 blieb als Hitzesommer in Erinnerung und auch das Frühjahr 2011 zeigte sich aussergewöhnlich trocken. Im Kanton Aargau musste die Abteilung Landschaft und Gewässer aufgrund der niedrigen Pegelstände an verschiedenen Gewässern Nutzungsbeschränkungen erlassen. Hiervon betroffen waren hauptsächlich Gemüsebauern, die zur Bewässerung ihrer Felder über Entnahmebewilligungen an Gewässern verfügen. In Zukunft dürften solche Zustände häufiger und gravierender auftreten; gleichzeitig erhöht sich mit der Trockenheit aber der Bewässerungsbedarf (Fuhrer, 2010), was die Situation weiter verschärfen dürfte.

Aufgrund dieser Entwicklungen besteht von Seiten der Kantonsverantwortlichen ein Interesse, Niedrigwasserstände prognostizieren zu können, um in der Lage zu sein, Anpassungsmassnahmen frühzeitig einzuleiten.¹ Dasselbe Ziel wird schweizweit verfolgt: Der Bericht zum „Umgang mit lokaler Wasserknappheit in der Schweiz“ des BAFU (2012b) führt die Entwicklung von Früherkennungs- und Warnsystemen als eine der anzustrebenden Massnahmen zur Trockenheits-Prävention auf. Von Seite der Forschung wurde aufgrund des allgemeinen Wissensbedarfs um die künftige Entwicklung der schweizerischen Wassersituation das Nationale Forschungsprojekt (NFP) 61 lanciert und in dessen Rahmen die Internet-Testplattform drought.ch (auch www.trockenheit.ch) zur Früherkennung von Trockenheiten aufgebaut (z.B. Seneviratne et al., 2013; Kruse, Seidl, & Stähli, 2010).

Diese Ausgangslage bildet die Motivation für die vorliegende Masterarbeit. Ihr Ziel ist, am Beispiel des Kantons Aargau und des hydrologischen Modells HBV die Möglichkeiten zur Vorhersage von Niedrigwasserständen auf kantonaler Ebene zu evaluieren. Der Fokus wird dabei auf die kleineren bis mittleren Gewässer im Kanton Aargau gelegt. Die grossen Gewässer wie Rhein, Reuss, Limmat oder Aare wurden nicht berücksichtigt, weil diese bisher nie den Bereich kritischer Pegelstände erreicht haben. Zudem sind diese grösseren Gewässer sehr stark durch Wasserkraftanlagen und andere Flussregulierungen beeinflusst, was eine Niederschlags-Abfluss-Modellierung schwierig gestalten würde.

Für die Modellierung werden nur beobachtete Daten und keine Prognosedaten verwendet. Prognosedaten haben in den letzten Jahren zwar deutliche Fortschritte in zeitlicher und räumlicher Hinsicht verzeichnet, weisen aber immer noch grosse Unsicherheiten auf (Sene, 2010: 68). Verschiedene Studien, die den Ursprung von Unsicherheiten in Modellergebnissen untersucht haben, kamen zum Schluss, dass Prognosedaten einen grossen oder sogar den Hauptanteil der Unsicherheiten ausmachen, sowohl in kurz- bis mittelfristigen Prognosen (z.B. Olsson & Lindström, 2008; Bosshard et al., 2013), wie auch in langfristigen (z.B. Wilby & Harris, 2006; Gardelin, Bergström, Carlsson, Graham, & Lindström, 2002). Dabei ist zu berücksichtigen, dass diese Unsicherheit grösser wird, je länger der Prognose-Zeitraum ist (Sene, 2010: 4; Addor, Jaun, Fundel, & Zappa, 2011).

Von Seiten des Kantons Aargau wären grundsätzlich zwei Prognose-Zeiträume von Interesse: Einerseits die mittelfristige Prognose auf ein bis drei Wochen, um kurzfristige Anpassungsmassnahmen einleiten zu können, zum anderen wäre aber auch eine saisonale Prognose hilfreich, damit die betroffenen Bauern ihren Anbau entsprechend anpassen könnten.

Zumindest für die kleineren Gewässer erscheint eine saisonale Prognose aber eher als schwierig, da diese Gewässer stark von lokalen Ereignissen beeinflusst werden. Gemäss dem Prinzip der

¹ Verwiesen sei hierzu beispielhaft auf den Zeitungsartikel „Kanton stellt Bauern das Wasser ab“, erschienen in der Aargauer Zeitung am 11. Mai 2011, verfasst von Marco Wölfli (Wölfli, 2011).

Verknüpfung von Raum und Zeit bedingt eine räumlich kleine Skala nämlich eine zeitlich kleine bzw. kurze Skala (Buttle, 1998: 35). Für kurzfristige Prognosen existieren allerdings bereits operierende Systeme: Für Niedrigwasser ist momentan die im Rahmen des NFP 61 entwickelte Plattform www.drought.ch in der Testphase; für Hochwasser existiert beispielsweise das Frühwarnsystem der Stadt Zürich für die Sihl (Addor et al., 2011). Solche Prognosen scheinen daher im Bereich des Machbaren zu liegen.

Dennoch werden für diese Arbeit keine Prognosedaten verwendet, denn in einer Evaluierung sollte klar unterschieden werden zwischen der Unsicherheit, die durch das hydrologische Modell erzeugt wird, und der Unsicherheit, welche die Prognosedaten einbringen. In der vorliegenden Arbeit wird der Fokus auf das hydrologische Modell selbst gelegt und daher nur mit meteorologischen Messdaten und keinen Prognosedaten gearbeitet.

Das Vorgehen in dieser Arbeit entspricht damit dem, was in der englischen Literatur als „hindcasting“ (Beven & Young, 2013; Dawson, Abrahart, & See, 2007) bezeichnet wird: Eine Periode in der Vergangenheit wird simuliert, als ob sie in der Zukunft läge. Die entsprechende Periode wird dazu aus dem Kalibrierungszeitraum entfernt („Split Sample-Test“ gemäss Klemes, 1986) und die Messdaten dieses Zeitraums übernehmen die Rolle von Prognosedaten. Es wird also keine echte Prognose in die Zukunft erstellt, das Modell wird aber ausserhalb der Kalibrierungsperiode ausgewertet, was an das Modell dieselben Anforderungen stellt. Um nicht zu suggerieren, dass meteorologische Prognosedaten verwendet wurden, wird im Weiteren nicht der Begriff „Prognose“, sondern der Begriff „Vorhersage“ benutzt. Vorteilhaft am Split Sample-Verfahren ist, dass die entsprechenden Abflussmesswerte für die Modellauswertung bereits zur Verfügung stehen.

Diese Arbeit ist in zwei Phasen aufgebaut: Im ersten Schritt wird die Fähigkeit des HBV-Modells, die Niedrigwasserperioden 2003 und 2011 vorherzusagen, getestet. Dabei wird ein besonderes Augenmerk auf die Karstgebiete im Aargauer Jura gelegt. Dazu werden alternative Modellstrukturen für diese Gebiete getestet und zusätzlich wird die Evapotranspiration als weiterer Faktor mit potentiell grossem Einfluss auf die Modellgüte untersucht.

In einem zweiten Schritt wird geprüft, ob sich die Modellierungen durch den Einbezug von Informationen über die Grundwasserstände verbessern lassen. Motiviert wird dieser Schritt durch die Tatsache, dass Oberflächengewässer, die keine Rückhalte wie Seen oder Gletscher besitzen, bei Niedrigwasser hauptsächlich durch Grundwasser gespeist werden (Smakhtin, 2001: 148). Ausserdem sind Trockenheiten Ereignisse, die sich langsam entwickeln. Grundwasser reagiert seinerseits stärker verzögert auf Niederschlagsdefizite als Oberflächengewässer (Sene, 2010: 241; Mishra & Singh, 2010: 206) und kann daher genauere Informationen über das Ausmass einer Trockenheit enthalten, als sie aus den Pegeln von Oberflächengewässern allein zu gewinnen wären.

Das Ziel ist also, die Fähigkeit des HBV-Modells zu evaluieren, Niedrigwasserprognosen für mittel- bis kleinskalige Einzugsgebiete im Kanton Aargau zu erstellen. Dazu wird zunächst das Standardmodell verwendet, anschliessend werden alternative Modellstrukturen sowie eine Co-Kalibrierung von Abfluss- und Grundwasserdaten getestet.

2. Definitionen

2.1. Trockenheit

Die verschiedenen hydrometeorologischen und auch sozioökonomischen Gegebenheiten haben weltweit zu einer Sammlung von verschiedenen Definitionen für das Phänomen Trockenheit geführt (Mishra & Singh, 2010: 205). Die gebräuchlichsten, u.a. die Definitionen nach WMO, UN oder FAO, sind in Mishra & Singh (2010: 206) zusammengestellt. Für diese Arbeit wird jedoch auf die etwas komplexere Definition nach Tallaksen & van Lanen (2004: 6) zurückgegriffen. Trockenheit ist demnach:

“... a sustained and regional extensive occurrence of below average natural water availability”.

„... ein anhaltendes und regional flächenhaftes Auftreten von unterdurchschnittlich tiefer natürlicher Wasserverfügbarkeit“.

Damit wird hervorgehoben, dass Trockenheit sowohl eine zeitliche als auch eine räumliche Komponente aufweist. Als temporäre Abweichung vom Normalzustand muss Trockenheit von Aridität, die eine Eigenschaft des lokalen Klimas darstellt, unterschieden werden (Mishra & Singh, 2010: 203). In dieser Arbeit werden daher auch die Begriffe Trockenperiode und Trockenheitsereignis verwendet.

Weiter ist zu betonen, dass die obige Definition nur die natürliche Dynamik einschliesst. Für eine erste Analyse von Trockenheiten und die Festlegung von Grenzwerten macht diese Definition durchaus Sinn. In den hier verwendeten Daten kann aber – wie vermutlich in den meisten praktischen Anwendungen – nicht mehr klar zwischen der natürlichen Wasserverfügbarkeit und der durch menschliche Nutzung veränderten bzw. reduzierten Verfügbarkeit unterschieden werden. In der Realität kann der Einfluss des Menschen Trockenperioden abschwächen oder verstärken und im Extremfall sogar auslösen (Van Lanen et al., in: Tallaksen & van Lanen, 2004: 347ff).

Ausgangspunkt für Trockenheitsereignisse bildet üblicherweise fehlender Niederschlag (in gemäßigten bis kalten Klimaten spielen Schnee und Eis eine modifizierende Rolle (Van Loon & Van Lanen, 2012); dieser Punkt spielt im untersuchten Zeitraum von 2011 eine Rolle). Dieses Defizit propagiert nach und nach in das gesamte hydrologische System (engl. *propagation of drought*): Zunächst reagieren die Böden und trocknen aus, anschliessend sinkt der Pegel der Oberflächen-Gewässer und schliesslich zeigt sich auch eine Reaktion in den Grundwasservorkommen (Van Loon & Van Lanen, 2012). Aufgrund dieser Abfolge werden vier Hauptkategorien von Trockenheit unterschieden (Tallaksen & van Lanen, 2004: 6; Mishra & Singh, 2010: 206):

- Die **meteorologische Trockenheit** (engl. *meteorological drought*) bezeichnet ein Niederschlagsdefizit. Sie bildet den Ausgangspunkt für alle weiteren Formen der Trockenheit.
- Die **Bodenfeuchte-Trockenheit** (engl. *soil moisture drought*) fokussiert auf die Austrocknung der Böden. Sie ist identisch mit dem Begriff *agricultural drought* (landwirtschaftliche Trockenheit), wenn damit die fehlende Bodenfeuchte zur Erhaltung von Feldkulturen bezeichnet wird.
- Die **hydrologische Trockenheit** (engl. *hydrological drought*) bezieht sich auf die verfügbaren Wasserressourcen. Hier sollte unterschieden werden zwischen der „Trockenheit der Oberflächengewässer“ (engl. *streamflow drought*) und der „Grundwasser-Trockenheit“ (engl. *groundwater drought*); während die Oberflächengewässer relativ bald auf fehlenden Niederschlag und trockene Böden reagieren, zeigen Grundwasservorkommen normalerweise

nur sehr langsam eine Reaktion. Entsprechend dauert es auch länger, bis sich Grundwasservorkommen nach einer Trockenheit wieder erholt haben.

- Die **sozio-ökonomische Trockenheit** (engl. *socio-economic drought*) bezeichnet die Auswirkungen von Trockenheit (unter Umständen gemildert durch Anpassungsmassnahmen) auf die Bevölkerung und die Wirtschaft. Verluste entstehen vor allem in landwirtschaftlichen Kulturen und durch verminderten Zugang zu Trink- und Nutzwasser.

An dieser Stelle kann alternativ auch die „ökologische Trockenheit“ (engl. *ecological drought*), also die Auswirkungen auf das (aquatische und nicht-aquatische) Ökosystem genannt werden. Je nach Untersuchungszweck ist dieser Aspekt wichtiger als der sozio-ökonomische.

Die Vielfalt an Begriffen illustriert, dass Trockenheit ein relatives Konzept ist: Je nach angewandtem Fokus und gewähltem Schwellenwert wird Trockenheit anders definiert.

Obwohl die am wenigsten direkt zugängliche Definition, sollte der vierte Punkt nicht unterschätzt werden. Zwar denken die wenigsten Menschen bei Naturkatastrophen zuerst an Trockenheit, tatsächlich ist aber weltweit gesehen Trockenheit die Naturgefahr mit den grössten Auswirkungen. (Tallaksen & van Lanen, 2004; Mishra & Singh, 2010; Wilhite & Glantz, 1985; Van Loon & Van Lanen, 2012: 1915)

Hisdal et al. (in: Tallaksen & van Lanen, 2004: 139) empfehlen, zur Analyse von Trockenperioden grundsätzlich zwischen Sommer- und Wintertrockenheiten zu unterscheiden, da ihnen unterschiedliche Ursachen zur Grunde liegen: Im Sommer steht die Kombination von Niederschlagsmangel mit hoher Evapotranspiration im Vordergrund, im Winter ist hingegen Frost der entscheidende Faktor. Für Van Loon & Van Lanen (2012: 1937) greift diese Unterscheidung jedoch zu kurz, weil sie besonders für die Wintertrockenheit verschiedene Prozesse als Auslöser ausmachen. Sie schlagen daher eine Klassierung in sechs Grundtypen vor. Eine Analyse der Untersuchungszeiträume gemäss dieser Einteilung ist in Kap. 5.4.1 und 5.4.2 vorgenommen worden. In der vorliegenden Arbeit interessieren grundsätzlich jedoch nur die Trockenheiten ausserhalb des Winters, da im Winter kein Bewässerungsbedarf für landwirtschaftliche Kulturen besteht. Es soll hier aber angemerkt sein, dass das Modell auf das gesamte Jahr kalibriert wird.

2.2. Niedrigwasser

Von den oben genannten Begriffen ist für die vorliegende Arbeit nur die hydrologische Trockenheit bzw. die Trockenheit der Oberflächengewässer von Bedeutung, da nur der Pegel der Fliessgewässer darüber entscheidet, ob Massnahmen durch den Kanton eingeleitet werden müssen. Obwohl gemäss Smakhtin (2001) klar zwischen den Begriffen Trockenheit (i.e. ein bestimmtes Ereignis, das aus einem Niederschlagsdefizit resultiert) und Niedrigwasser (i.e. ein charakteristischer Teil des Fliessregimes) unterschieden werden muss, wird der Begriff Niedrigwasser in dieser Arbeit synonym zu Trockenheit der Oberflächengewässer verwendet.

2.2.1. Q_{347} und Mindestrestwassermenge

Niedrigwasser muss über einen Schwellenwert definiert werden. Für den Kanton Aargau sind von Gesetzes wegen zwei Schwellenwerte relevant: Zum einen das Q_{347} , zum anderen die Mindestrestwassermenge. Ab Erreichen des Q_{347} werden die Gewässer konstant überwacht und die Wasserbezüger und Gemeinden über die kritische Lage informiert. Bei Erreichen der Mindestrestwassermenge werden Nutzungsbeschränkungen bzw. -verbote erlassen (Berner, 2011).

Q_{347} entspricht dem 95%-Quantil, also der Abflussmenge, „die, gemittelt über 10 Jahre, durchschnittlich während 347 Tagen des Jahres erreicht oder überschritten wird und die durch Stauung, Entnahme oder Zuleitung von Wasser nicht wesentlich beeinflusst ist“ (Aschwanden & Kan, 1999: 19). Andersherum betrachtet, wird Q_{347} im Durchschnitt an 18 Tagen (=5%) im Jahr unterschritten. Wie bereits angetönt, lassen die Klimaprognosen erwarten, dass dieser Grenzwert in Zukunft häufiger (bis doppelt so häufig) unterschritten wird (BAFU, 2012: 62). In der praktischen Umsetzung wird das Q_{347} aufgrund von verschiedenen Punktmessungen statistisch bestimmt und auf das gesamte Gewässer (oder zumindest Abschnitte davon) extrapoliert. Diese Verallgemeinerung bringt relativ grosse Ungenauigkeiten mit sich. Weil aber die Entnahmen grösstenteils zur Feldbewässerung durch viele verschiedene Kleinbauern geschehen, wurde dieses Vorgehen vom Bund dennoch empfohlen, um Aufwand und Ertrag in einem angemessenen Verhältnis zu halten (BAFU, 1997: 5).

Die Mindestrestwassermenge ist im Gewässerschutzgesetz (GSchG) Art. 31, Abs. 1 und Abs. 2 ausgehend von Q_{347} festgelegt. Hierbei gelten beispielsweise folgende Werte: Bei einem Q_{347} von 60'000 l/s beträgt die Mindestrestwassermenge 10'000 l/s; bei einem viel geringeren Q_{347} von 60 l/s ist die Mindestrestwassermenge auf 50 l/s festgesetzt. Abs. 2 verlangt, dass diese Werte unter gewissen Bedingungen erhöht werden, beispielsweise um in einem sehr breiten und flachen Gerinne die freie Fischwanderung zu gewährleisten. Zum Schutz des aquatischen Ökosystems dürfen diese Werte nicht unterschritten werden; daher werden jegliche Entnahmebewilligungen an Gewässern sistiert, sobald dieser Wert erreicht ist (Berner, 2011).

Für alle bis auf drei Einzugsgebiete konnte auf die operationell verwendeten Werte des Kantons für Q_{347} und die Mindestrestwassermengen zurückgegriffen werden (die Werte sind in Tab. 1 zusammengestellt). Für die übrigen drei Einzugsgebiete (345, 352 und 371) wurden die Werte aufgrund der vorliegenden Abflussmesswerte berechnet. Per Definition wird Q_{347} über ein 10-jähriges Mittel bestimmt – nur fällt das Resultat dabei unterschiedlich aus, je nachdem, welches 10-Jahres Fenster verwendet wird. Aus diesem Grund wurden hier die gesamten vorliegenden Messreihen zur Berechnung verwendet, da längere Zeitreihen zu robusteren Ergebnissen führen (unter der Annahme von stationären Bedingungen).

Zu den Mindestrestwassermengen ist speziell erwähnenswert, dass 17 der 36 Einzugsgebiete ein Q_{347} unter dem Grenzwert von 60 l/s aufweisen und damit die Mindestrestwassermenge auf den Wert von 50 l/s festgesetzt wird. Dies bedeutet, dass der Grenzwert für die Mindestrestwassermenge grösser als Q_{347} wird, sobald Q_{347} unter 50 l/s fällt. Bei den Gewässern mit Q_{347} grösser als 60 l/s wird die Mindestrestwassermenge gemäss GSchG Art. 31, Abs. 1 berechnet.

2.3. Evapotranspiration

In der Literatur werden die Begriffe „Evaporation“ und „Evapotranspiration“ (sowie „Transpiration“) nicht einheitlich behandelt, wie McMahon, Peel, Lowe, Srikanthan, & McVicar (2013: 1334) aufzeigen. Dies gilt auch für die Literatur, die sich mit HBV beschäftigt: Während z.B. Bergström (1976) und Gardelin et al. (2002) nur den Begriff „Evapotranspiration“ verwenden, benutzen beispielsweise Seibert (1997 & 2000), Seibert & Vis (2012) und Uhlenbrook, Seibert, Leibundgut, & Rodhe (1999) nur den Begriff „Evaporation“. Das Programm selbst ist ebenfalls uneinheitlich: Während die Benutzeroberfläche die Beschriftung „Evapotranspiration“ aufweist, ist in der Hilfe nur „Evaporation“ zu finden. Lindström, Johansson, Persson, Gardelin, & Bergström (1997) benutzen beide Begriffe, ohne sie genau abzugrenzen, scheinen aber zwischen „Evaporation“ als Prozess über offenen Wasserflächen und der Kombination mit pflanzlicher „Transpiration“ zu „Evapotranspiration“ zu unterscheiden. McMahon et al. (2013: 1332) definieren die drei Begriffe folgendermassen:

- Evaporation bezeichnet verschiedene Prozesse, durch die Wasser (von offenen Wasseroberflächen, von Bodenoberflächen oder von Interzeption) aus der flüssigen in die Gasphase überführt wird.
- Transpiration bezeichnet die Verdunstung von Wasser durch Pflanzen. Das Wasser wird dabei aus dem Pflanzeninneren durch die Stomata an die Luft abgegeben.
- Evapotranspiration bezeichnet die Kombination der beiden Prozesse.

In dieser Arbeit wird fortan mit dem Begriff „Evapotranspiration“ gearbeitet, erstens um dem Umstand Rechnung zu tragen, dass die HBV-Benutzeroberfläche den Begriff „Evapotranspiration“ verwendet, zweitens um herauszustreichen, dass HBV nicht zwischen den einzelnen Teilprozessen unterscheidet, und drittens weil die verwendete Berechnungsmethode (siehe Kap. 4.4) ebenfalls diesen Begriff benutzt. Allerdings ist dabei weiter zu unterscheiden zwischen der „potentiellen Evapotranspiration“ (PET), die den theoretischen Wert bei unlimitierter Wasserverfügbarkeit bezeichnet, und der „eigentlichen Evapotranspiration“ (engl. *actual evapotranspiration*, AET), die den unter realen Bedingungen tatsächlich auftretenden Wert beschreibt (definiert gemäss McMahon et al., 2013).

2.4. Skalengrössen

Da die verwendeten Einzugsgebiete eine Grösse von 0.6 bis gut 350 km² aufweisen (die meisten davon liegen im Bereich von 5 bis 40 km²), befinden wir uns gemäss Schultz (1994) auf der Ebene der Mesoskala. Hingegen unterscheidet Buttle (1998) nur zwischen kleinskalig und grossskalig und definiert ersteren Begriff für den Bereich von 10⁻² bis 10² km², wonach alle ausser den beiden grössten untersuchten Einzugsgebieten als kleinskalig zu bezeichnen wären. Da die Abgrenzung zwischen klein- und mittelskalig schwierig zu ziehen ist, verbleibt diese Arbeit dabei, dass sie sich im Übergangsbereich zwischen der Mikro- und der Mesoskala befindet.

3. Modell und Gütekriterien

3.1. Warum HBV?

Das HBV (Hydrologiska Byråns Vattenbalansavdelning)-Modell (Bergström, 1976), entwickelt am Schwedischen Meteorologischen und Hydrologischen Institut (SMHI), ist ein einfach zu handhabendes und vielseitig einsetzbares hydrologisches Modell. Es modelliert den Abfluss eines Fließgewässers aufgrund des Niederschlags und ist damit ein klassisches Niederschlags-Abfluss (engl. *rainfall-runoff*)-Modell. HBV ist ein konzeptionelles bzw. semi-empirisches Modell: Die einzelnen Teile des Modells (hier „Routinen“ von engl. *routines* genannt) basieren zwar auf hydrologischen Prozessen, simulieren diese aber nicht auf physikalisch detaillierter Ebene. Beispielsweise enthält das Modell eine Boden-Routine, die den Wasserspeicher im Boden simuliert – dieser entspricht aber nicht direkt den im Feld messbaren Werten wie Saugspannung und Bodendurchlässigkeit. Die Modellstruktur ist dadurch wesentlich einfacher und überblickbarer als in physikalisch-basierten Modellen, wodurch HBV auch deutlich weniger Rechenleistung beansprucht und sich so einen breiten Einsatzhorizont eröffnet.

Physikalisch basierte Modelle sind zwar näher am Stand des Verständnisses der hydrologischen Prozesse, sind aber nicht per se für jeden Zweck vorteilhaft:

“However, if the main interest simply lies in the estimation of stream flow at the catchment scale, then simpler conceptual or data-driven models often perform well and the high complexity of physically based models is not required“ (Solmatine & Wagener, 2011: 440)

Wie dem auch sei, wenn das Hauptinteresse nur auf der Abflussberechnung auf Einzugsgebietsebene liegt, dann bewähren sich einfachere, konzeptionelle oder datengetriebene Modelle oft gut und die hohe Komplexität von physikalisch basierten Modellen wird gar nicht benötigt.

Dasselbe Prinzip gilt nicht nur für die Komplexität der implementierten Modellformeln, sondern auch für den räumlichen Aspekt. Grundsätzlich existieren in der Hydrologischen Modellierung zwei Extreme an Modellen: Zum einen die „lumped models“, die das gesamte Einzugsgebiet als undifferenzierte Einheit verwenden, und zum anderen die „distributed models“, die das Modellergebnis für jede einzelne Zelle des Einzugsgebietes separat berechnen (Beven, 2012: 16).

Von ihrer Natur her sind physikalisch basierte Modelle meist *distributed*, während empirische Modelle eher *lumped* aufgebaut sind. Obwohl der höhere räumliche Detaillierungsgrad der *distributed models* eine grössere Genauigkeit suggeriert, müssen hier Vorbehalte angebracht werden:

„ ... due to difficulties in calibrating distributed models, in many cases, conceptual models are in fact more accurate in reproducing the resulting catchment stream flow than the distributed ones. “ (Solmatine & Wagener, 2011: 441)

... Aufgrund von Schwierigkeiten in der Kalibrierung von *distributed models* sind konzeptionelle Modelle tatsächlich in vielen Fällen genauer in der Beschreibung des resultierenden Abflusses als *distributed models*.

Kalibrierung ist ein Schlüsselthema der Modellierung allgemein und damit auch der hydrologischen Modellierung: Da Modelle per se niemals vollständige Repräsentationen der Realität sind, enthalten sie immer, wie detailliert ihr Aufbau auch sei, gewisse Fehler bzw. Ungenauigkeiten. Hinzu kommt in

der Hydrologie ein kaum umgehbarer Skaleneffekt: Während hydrologische Messmethoden die Parameter normalerweise auf lokaler Ebene und zu einem bestimmten Zeitpunkt messen, verwendet das Modell über das gesamte Einzugsgebiet und langfristig gemittelte Werte (bei *distributed models* gilt dieser Vorbehalt zwar für die zeitliche, nicht unbedingt aber für die räumliche Skala). Wir treffen dabei implizit die Annahme, dass von der lokal-instantanen Skala auf die regional-langfristige Skala geschlossen werden kann – was keine vorbehaltlos berechtigte Annahme ist (Beven, 2012: 18).

Daher ist es so gut wie unmöglich, hydrologische Modellierungen ohne einen Kalibrierungs-Vorgang durchzuführen. Bei Modellen mit grossen Parametermengen, wie dies bei physikalisch basierten Modellen der Fall ist, stösst man unter Umständen auf das Problem der Über-Parametrisierung: Das Modell verfügt über so viele veränderbare Parameter, dass nahezu jegliches Verhalten, unabhängig davon, ob hydrologisch sinnvoll oder nicht, erzeugbar ist (Solmatine & Wagener, 2011). Bereits 1972 zeigte Kuchment (zit. in Bergström, 1976: 20) auf, dass eine grössere Anzahl von Parametern zwar eine bessere Anpassung an die Beobachtungen des Kalibrierungszeitraumes ermöglicht, eine Anwendung ausserhalb des Kalibrierungszeitraumes, also beispielsweise die Erstellung einer Prognose, aber schlechtere Resultate liefert als bei einem Modell mit weniger Parametern.

HBV scheint in diesem Konflikt einen sinnvollen Kompromiss zu bieten:

„The number of free parameters, subject to calibration, is moderate, but large enough to give freedom for adjustment under different conditions. ” (Lindström, Johansson, Persson, Gardelin, & Bergström, 1997:274/275)

Die Anzahl der wählbaren Parameter, die der Kalibrierung unterliegen, ist moderat, aber dennoch gross genug, um die Anpassung an unterschiedliche Bedingungen zu ermöglichen.

Die Autoren ergänzen diese Aussage mit der Feststellung, dass sich HBV in einem breiten Skalenbereich bewährt habe.

HBV-light (die in dieser Arbeit verwendete Version von HBV, beschrieben in Seibert & Vis, 2012) ist allerdings weder *lumped*, noch *distributed*, sondern beschreitet einen Kompromissweg und wird als „semi-distributed“ bezeichnet. Es erlaubt nämlich den Einbezug verschiedener Teilgebiete, die beispielsweise über die Höhenlage, den Vegetationstyp und/oder die Hangexposition definiert werden können (siehe auch Kap. 4.5).

Desweiteren hat sich HBV bereits in der operationellen Anwendung bewährt: Seit Juli 2004 nutzt das SMHI HBV operationell für tägliche 9-Tages-Prognosen in 50 Einzugsgebieten der Grössenordnung von 8 - 6'000 km² (Olsson & Lindström, 2008). Es sprechen also verschiedene Gründe für die Verwendung des HBV-Modells, insbesondere für den nicht rein forschungsspezifischen, sondern anwendungsorientierten Ansatz dieser Arbeit.

3.2. Modellstruktur und Parameter

Das Modell wird hier mit Blick auf seine Funktionsweise beschrieben. Die ursprüngliche Modellvariante wurde erstmals durch Bergström (1976) vorgestellt. Die hier verwendete Version HBV-light wurde von Seibert & Vis (2012) dokumentiert und ist auf der Webseite der Abteilung Hydrologie und Klima des Geographischen Instituts der Universität Zürich unter folgendem [Link](#)² frei verfügbar; für detaillierte Beschreibungen der Formeln sei auf die genannten Autoren verwiesen.

Das Modell arbeitet standardmässig in Tagesschritten, kann aber auch in anderen Zeitschritten benutzt werden.

² <http://www.geo.uzh.ch/en/units/h2k/services/hbv-model> (letzter Zugriff: 16.6.2014)

Grundsätzlich ist das Modell in vier Bereiche, Routinen genannt, unterteilt (siehe Abb. 1; eine etwas detailliertere, schematischere Darstellung ist in Abb. 2 zu finden):

- *Snow Routine* (Schneeberechnung, hier „Schnee-Routine“ genannt)
- *Soil Routine* (Berechnung der Bodenfeuchte, hier „Boden-Routine“ genannt)
- *Groundwater Routine* (auch *Response Routine*; Grundwasser-Berechnung, hier „Grundwasser-Routine“ genannt)
- *Routing Routine* (Abflussbildung, hier „Abfluss-Routine“ genannt)

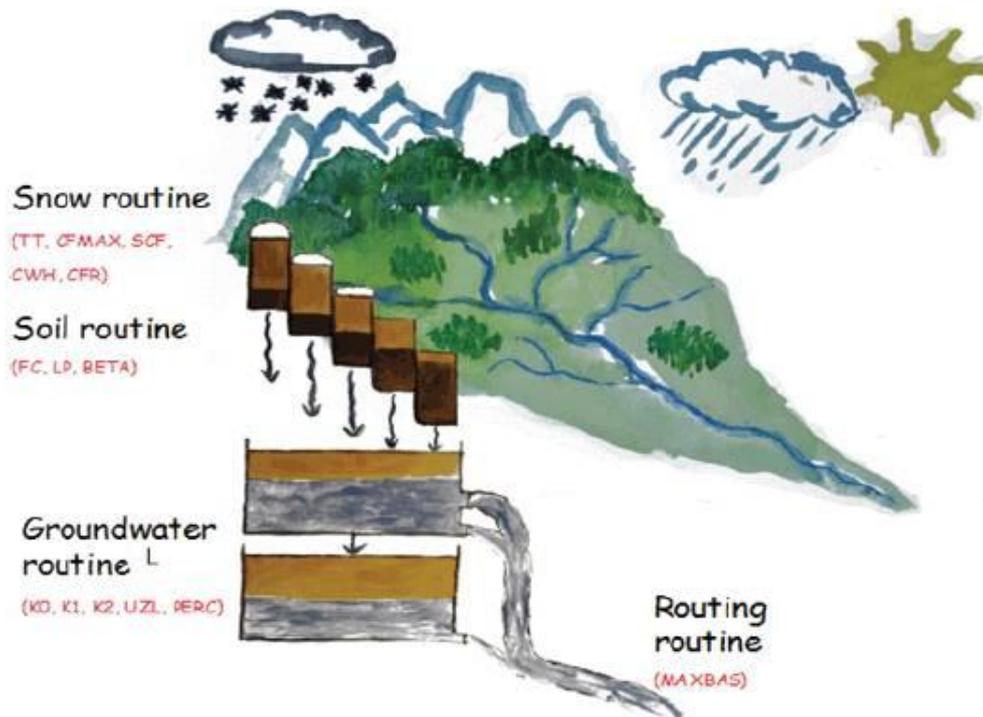


Abb. 1: Darstellung der HBV-Struktur bzw. der verschiedenen Routinen und deren Parameter. Quelle: <http://www.geo.uzh.ch/en/units/h2k/services/hbv-model> (Zugriff 28.1.2014)

Die **Schnee-Routine** berechnet die Schneeakkumulation und -ablation basierend auf einem Grenzwert TT [°C] und der Gradtag-Methode. Der Gradtagfaktor CFMAX [mm/(d°C)] bestimmt die Intensität der Schneeschmelze; der Korrekturfaktor SCF (engl. *Snowfall Correction Factor*) [-] korrigiert die systematische Unterschätzung der Schneefallmenge an einer Messstation durch den Einfluss von Wind; der Parameter CWH (engl. *Water Holding Capacity*) [-] simuliert die Sättigung des Schneepaketes mit Schmelzwasser und verzögert so die Abflussgenerierung bei Überschreiten von TT, bis der Schnee übersättigt ist; der Parameter CFR (engl. *Refreezing Coefficient*) [-] kommt schliesslich zum Zug, wenn Schmelzwasser im Schneepaket wieder gefriert.

In der **Boden-Routine** wird berechnet, welcher Anteil des Wasserinputs aus Niederschlag oder Schneeschmelze in den Boden infiltriert (i.e. dem Bodenspeicher zugeführt wird), wie viel weiter in die Grundwasser-Routine geht und welcher Anteil verdunstet wird. Diese Anteile sind abhängig vom Wassergehalt des Bodenspeichers zu einem bestimmten Zeitpunkt im Verhältnis zum maximal möglichen Wassergehalt. Dabei stellt der Parameter FC [mm] den maximal möglichen Wassergehalt im Bodenspeicher dar (nicht zu verwechseln mit der tatsächlichen Feldkapazität des Bodens!); der Parameter BETA [-] formt eine Verhältniskurve, die bestimmt, bei welchem Wassergehalt des Bodenspeichers ein wie grosser Anteil des Niederschlagsinputs ins Grundwasser weitergeleitet wird

und wie viel im Bodenspeicher zurückgehalten wird. Der Bodenspeicher erfährt dabei gleichzeitig Verluste durch Evapotranspiration. Die tatsächliche Evapotranspiration (AET) zu einem bestimmten Zeitpunkt berechnet sich aus der potentiellen Evapotranspiration (PET) und dem Verhältnis des aktuellen Bodenspeicher-Wassergehalts zum maximalen Wassergehalt FC, wobei der maximale Wassergehalt FC mit dem Faktor LP [-] multipliziert wird. LP bestimmt damit den Grenzwert, über dem die AET den Wert der PET annimmt. Die PET wiederum wird von HBV basierend auf langjährigen Mittelwerten (siehe Kap. 4.4) berechnet. Dafür verwendet das Modell die Differenz zwischen der Tagestemperatur und dem langjährigen Temperaturmittel und korrigiert diese mit dem Faktor C_{et} [$^{\circ}\text{C}^{-1}$].

In Ausblick auf die Modellkalibrierung ist erwähnenswert, dass Zhang & Lindström (1997) für die drei Parameter FC, BETA und LP der Boden-Routine eine starke gegenseitige Abhängigkeit festgestellt haben.

Was vom Wasserinput nicht in den Bodenspeicher eingeleitet wurde, gelangt in die **Grundwasser-Routine**. Diese ist unterteilt in die obere Grundwasserbox (engl. *Upper Groundwater Box*, SUZ) und die untere Grundwasserbox (engl. *Lower Groundwater Box*, SLZ), die oberflächennahe respektive tieferliegende Grundwasservorkommen repräsentieren. Die obere Box kann allerdings auch für gesättigte Böden stehen. Die beiden Gefässe sind über den Parameter PERC [mm/d] verbunden, der die maximale Perkolationsrate von der oberen in die untere Box vorgibt. In der HBV-Standardversion entstehen aus diesen beiden Gefässen drei lineare Ausflüsse, die über die Parameter K_0 , K_1 und K_2 (alle in der Einheit [d^{-1}]) bestimmt werden. Liegt der Wassergehalt der oberen Box über dem Schwellenwert UZL [mm], so tritt der oberste Ausfluss (oberer Ausfluss der oberen Box), definiert über K_0 , in Kraft. Üblicherweise ist K_0 grösser als K_1 oder K_2 und steht damit für schnellen und verhältnismässig grossen Abfluss, der die Abflussspitzen massgeblich formt. Liegt der Wassergehalt in der oberen Box unter dem Schwellenwert UZL, ist in der oberen Box nur der untere Ausfluss aktiv. Dieser wird über den Parameter K_1 definiert. Gleichzeitig tritt auch Ausfluss aus der unteren Box auf, der über K_2 bestimmt wird.

Zur Konstellation der beiden Boxen ist erwähnenswert, dass die obere Box aufgrund der Perkolationsrate in die untere Box vollkommen austrocknen kann – im Gegensatz zur unteren Box, deren Ausfluss als Produkt ihres Inhalts mit dem Parameter K_2 berechnet wird. Daher kann sich der Wassergehalt dieser Box – und damit auch der Abfluss – bei extremer Trockenheit zwar asymptotisch an null annähern, fällt aber nie auf den Wert null. HBV kann also ein Versiegen des Abflusses nicht simulieren.

Die beiden Boxen SUZ und SLZ verfügen über keinerlei Grenzwerte; ihre Werte ergeben sich aus den Modellgleichungen.

Der eigentliche Abfluss wird schliesslich in der **Abfluss-Routine** berechnet. Hier werden die Ausflüsse aus der Grundwasser-Routine zu einer Abflusssimulation vereint. Der Parameter MAXBAS [d] bestimmt die zeitliche Verzögerung des Abflusses; bildlich gesprochen steht dieser Parameter für die Zeit, die Wasser benötigt, um durch das Einzugsgebiet hinab bis zur Messstation zu gelangen. Weil die dafür benötigte Zeit je nach Ausgangspunkt im Einzugsgebiet unterschiedlich ist, wird keine lineare, sondern eine Dreiecksfunktion verwendet, wobei MAXBAS den Maximalwert dieses Dreiecks definiert (siehe Abb. 2); die Ausflüsse aus der Grundwasser-Routine werden mit dieser Dreiecksfunktion verrechnet. Daraus ergibt sich der berechnete Abfluss [mm/d], der vom Modell dargestellt und ausgewertet wird.

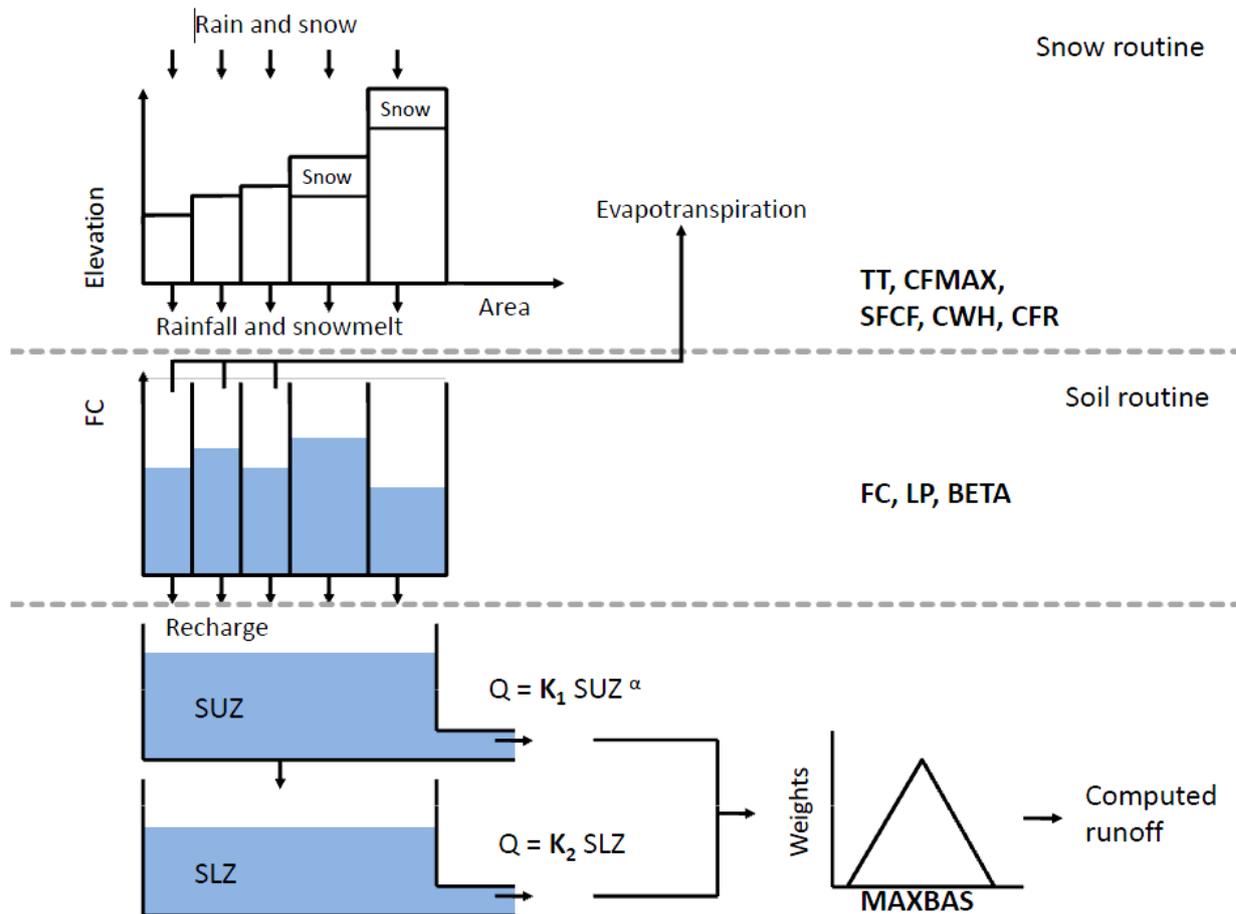


Abb. 2: HBV-Modellstruktur in systematischer Darstellung.

© M. Staudinger

Die durch Seibert & Vis (2012) beschriebene neuere Version HBV-light wurde insbesondere auf Benutzerfreundlichkeit ausgelegt. Zusätzlich wurden einige Ergänzungen implementiert, wie eine Gletscher-Routine, die das Modell auch auf teil-vergletscherten Einzugsgebieten anwendbar macht. Für die vorliegende Arbeit waren zwei weitere Neuerungen von Bedeutung: Zum einen verwendet HBV-light eine Aufwärmphase (engl. *Warm-up Period*), die an Stelle der früher nötigen Definition der Ausgangszustände (engl. *initial states*) trat. Desweiteren ermöglicht es die neue Version, für die Parameterkalibrierung eine Kombination verschiedener Gütemasse zu verwenden.

3.3. Modellannahmen

Gemäss Beven (2012: 4) ist es essentiell, die Annahmen und Vereinfachungen eines Modells explizit festzuhalten. Zum HBV-Modell sind gemäss Bergström (1976) in dieser Beziehung folgende Punkte zu nennen:

- HBV kennt keinen Oberflächen-Abfluss. Dies ist in der Tatsache begründet, dass HBV für Schwedische Einzugsgebiete erstellt wurde, die sehr hohe Versickerungsraten und relativ schwache Regenintensitäten aufweisen. Oberflächenabfluss tritt daher kaum auf. Bezogen auf die Schweiz bzw. auf den Aargau könnte diese Tatsache beispielsweise die korrekte Simulation des Abflusses bei Niederschlag auf sehr trockene Böden oder die Modellierung von Abflussspitzen bei Starkniederschlägen erschweren.
- Die Trennung der Boden-Routine von der oberen Box der Grundwasser-Routine ist zwar für die Berechnung angenehm, entspricht aber nicht unbedingt dem natürlichen System. In der Realität steht Grundwasser normalerweise durch kapillaren Aufstieg im Kontakt mit dem

Boden; die Verbindung zwischen Grundwasser und Bodenfeuchte wird stärker, je näher der Grundwasserspiegel an der Oberfläche liegt.

- Der Bodenspeicher ist keine direkte Berechnung der Bodenfeuchte, sondern eher ein Index dafür, da in ihm bzw. in der berechneten AET verschiedene Prozesse vermischt werden, wie beispielsweise Interzeption, Evaporation und Transpiration.
- Ob Niederschlag in Form von Regen oder Schnee fällt, wird in HBV durch eine simple Berechnung über die Schwellenwerttemperatur TT definiert. Nicht berücksichtigt wird hierbei beispielsweise die Bodentemperatur, die ein Schmelzen von frisch fallendem Schnee verursachen kann. Ebenso wenig wird der Einfluss von gefrorenem Boden berücksichtigt.
- Die Parameter TT und CFMAX sind stark voneinander abhängig. Es ist daher oft schwierig, eindeutige Schlüsse über ihre Werte zu ziehen. Dies ist ein klassischer Fall der durch Beven (2006) beschriebenen „Equifinality“ von Parametern (dasselbe Konzept wurde im Übrigen bereits von Bergström (1976: 39) als „nonuniqueness of processes“, Nicht-Einheitlichkeit der Prozesse, beschrieben).
- Der Parameter SFC beeinflusst das Gesamtvolumen an Wasser, das dem System zugeführt wird, und ist damit ein entscheidender Faktor.
- HBV korrigiert Temperatur und Niederschlag für verschiedene Höhenstufen (wenn diese angegeben werden), nicht aber die potentielle Evapotranspiration.

Diese Punkte zeigen auf, dass HBV, genau wie jedes Modell, keine exakte Beschreibung der Wirklichkeit ist. Dennoch haben die vielen Anwendungen des Modells bewiesen, dass aussagekräftige Simulationen mit dem erstaunlich einfach gehaltenen Modell erstellt werden können (Bergström, 1976: 107). Bekräftigt wird dies durch die Ergebnisse von Lindström et al. (1997), die die gesamte Struktur des ursprünglichen HBV-Modells unter die Lupe nahmen und verschiedenste Erweiterungen testeten. Nur in wenigen Fällen konnten sie damit Verbesserungen erzielen; so ergaben beispielsweise der Ausbau der Snow Routine oder die explizite Berechnung der Interzeption keine nennenswerten Verbesserungen. Die Struktur des HBV-Modells kann also aus guten Gründen als robust bezeichnet werden.

3.4. Benötigte Input-Dateien

HBV-light benötigt folgende Input-Daten:

- Eine „PTQ“-Datei, die die Zeitreihen von Niederschlag, Temperatur und gemessenem Abfluss enthält (es sind grundsätzlich auch andere Zeitschritte als eine Tagesauflösung möglich). Der gemessene Abfluss wird zur Auswertung der Modellgüte verwendet. Das Modell rechnet mit dem spezifischen Gebietsabfluss [mm/d]. Gemessene Abflusswerte [meist in m^3/s] müssen also über die Fläche des Einzugsgebiets normiert und in Millimeter pro Tag umgewandelt werden (um in Tagesschritten zu rechnen).
- Eine zusätzliche Temperaturdatei, die langjährige, monatliche Mittelwerte enthält.
- Eine Datei mit Evapotranspirationswerten, ebenfalls pro Monat aus langjährigen Zeitreihen gemittelt.
- Eine Datei, die die Flächenanteile der verschiedenen Höhenstufen bestimmt.

Die Herkunft und Eigenschaften der Daten, die zur Erstellung dieser Input-Dateien verwendet wurden, sind in den Kapiteln 4.1 bis 4.5 beschrieben.

4. Datengrundlage

4.1. Abflussdaten

Ein gutes Modellierungsergebnis kann nur erzielt werden, wenn die Daten für den Modellinput eine gute Qualität aufweisen. Daher betonen Solmatine & Wagener (2011: 451), dass neben der Modellstruktur selbst die Güte der Daten für die gesamte Unsicherheit der Modellierung viel entscheidender als die Parameter-Unsicherheit sei. In diesem Sinne ist es umso erfreulicher, dass für diese Arbeit auf qualitätsgeprüfte Abflussdaten in langjährigen Messreihen zurückgriffen werden konnte.

Die verwendeten Abflussdaten stammen hauptsächlich vom Kanton Aargau, ergänzt durch die beiden BAFU-Stationen der Murg bei Murgenthal und der Wigger bei Zofingen. Alle Daten wurden vom Kanton Aargau bzw. dem BAFU qualitätsgeprüft und im jeweiligen hydrologischen Jahrbuch veröffentlicht (online verfügbar unter www.ag.ch/hydrometrie bzw. www.hydrodaten.admin.ch).

Für die vorliegende Masterarbeit wurden Tagesmittelwerte aus so langen Zeitreihen verwendet, wie sie jeweils für die Stationen bis zum 31.12.2012 verfügbar waren.

Die Angaben der Kapitel 4.1.1 und 4.1.2 beruhen für den Kanton Aargau auf mündlichen Angaben von Ch. Lienert vom 16. 10. 2013 und S. Kempf vom 25.9.2013 und für die Daten des BAFU auf schriftlichen Auskünften von M. Oberhänsli vom 25.10.2013.

4.1.1. Messart

Im Kanton Aargau stehen mit Radarmessung und Drucksonde zwei verschiedenen Messtechniken im Einsatz. Bei den meisten Stationen wird redundant mit beiden Methoden gemessen. Ausgehend vom Pegel wird über eine P-Q- (Pegel – Abfluss) Beziehungskurve, die für jede Station einzeln ermittelt ist, das Abflussvolumen bestimmt. Dafür sollte das Gerinne über einen fixen und genau bestimmbar Querschnitt verfügen, weshalb an fast allen Stationen zwischen 2007 und 2011 neue Messschwellen eingerichtet wurden. Die Justierung der Pegel-Abfluss-Beziehung ist eine Daueraufgabe: Zur Verbesserung der Kurve werden laufend neue Eichmessungen durchgeführt.

Ursprünglich wurden die Pegelstände mit analogen Limnigraphen festgehalten, deren Messwerte durch periodische Digitalisierung in die Datenbank übertragen wurden. Hierbei, sowie am Ende des Jahres beim Sichten der Jahreskurven, wurden die Aufzeichnungen auf Plausibilität geprüft. Nach dem Jahr 2000 wurden alle Messstationen nach und nach digitalisiert. Seit dieser Umstellung wird im 10-Minuten-Rhythmus gemessen, wobei die Messung selbst 2-minütlich erfolgt, der Wert aber nur alle 10 Minuten als Mittel festgehalten wird.

Über die Messinstrumentierung der BAFU-Stationen liegen keine genaueren Angaben vor.

4.1.2. Qualitätsprüfung

Hoch- und Niedrigwasserbereich von Abfluss-Messkurven sind aufgrund messtechnischer Schwierigkeiten und der Seltenheit solcher Ereignisse jeweils mit den grössten Unsicherheiten behaftet (Morgenschweis, 2010: 389). Umso grössere Bedeutung gewinnt bei der Verarbeitung von Niedrigwasserdaten die vorhergehende Qualitätsprüfung.

Im Kanton Aargau geschieht diese auf zwei Ebenen: Zum einen im Büro, wo die Abflusskurven visuell auf Plausibilität geprüft werden und allfällige systematische Fehler oder Messlücken auskorrigiert werden; zum anderen wird im Feld jede Messstation regelmässig und mindestens zweiwöchentlich überprüft und die Anlage gereinigt. Ohne eine solch sorgfältige und regelmässige Kontrolle können sich Messfehler aus unterschiedlichen Gründen ergeben: So konnte es bei den früher verwendeten, offenen Trichtern der Radarmessgeräte vorkommen, dass Spinnen im Trichter ihr Netz aufspannten

und so Fehlmessungen verursachen; ebenso entscheidend ist besonders bei Niedrigwasser die regelmässige Entfernung des Algenbewuchses auf der Messschwelle (siehe Abb. 3).

Ähnlich geht das BAFU bei der Sicherstellung der Datenqualität vor: Auch hier wird die Qualität sowohl im Feld als auch im Büro laufend und zusätzlich nach Jahresabschluss geprüft.



Abb. 3: Algenbewuchs an der Messschwelle der Sissle bei Eiken (EZ 331) bei Niedrigwasser im Sommer 2006.

Quelle: W. Lehmann, ALG Wasserbau, BVU, Kanton Aargau.

4.1.3. Einzugsgebiete: Auswahl und Charakteristika

Die Ausgangslage zur Auswahl der Einzugsgebiete (EZ) wurde durch das Hydrometrie-Netz des Kantons Aargau vorgegeben (alle Stationen sind unter www.ag.ch/hydrometrie aufrufbar). Ausgewählt wurden schliesslich alle Stationen, die erstens an keinem der grossen Flüsse liegen (weil diese bisher nie in der Gefahr standen, den gesetzlichen Grenzwert zu unterschreiten), zweitens genügend lange Zeitreihen aufweisen und insbesondere die Jahre 2003 und 2011 enthielten (eine Station wies im Sommer 2003 einen Messausfall auf), und drittens nicht durch einen regulierten See (Baldegger-, Hallwiler- oder Sempachersee) beeinflusst waren. Auf diese Weise wurden 34 Stationen, die durch den Kanton Aargau betreut werden, sowie 2 Stationen des BAFU ausgewählt – insgesamt also 36 Stationen. Ein Überblick zu den wichtigsten Eigenschaften ist in Tab. 1 zusammengestellt. Abb. 6 verschafft einen geografischen Überblick.

Das kleinste Einzugsgebiet (345) weist eine Fläche von 0.6 km² auf, das grösste 368 km² (2450). Die Fläche der meisten Gebiete (26 von 36) liegt zwischen 5 und 40 km², was sich in einem Mittelwert von 45,8 km², aber einem Median von 20.2 km² widerspiegelt.

Alle ausgewählten Einzugsgebiete weisen ein pluviales Regime auf. Genauer kann unterschieden werden zwischen einem Regime „pluvial jurassien“ für die Einzugsgebiete, die nördlich bzw. westlich der Aare im Aargauer Jura liegen, und „pluvial inférieure“ (siehe Abb. 4) für alle übrigen Einzugsgebiete (BAFU, 1992). Diese Einteilung wird anhand der Pardé-Koeffizienten (mittlerer Monatsabfluss dividiert durch den mittleren Jahresabfluss) vorgenommen; in Abb. 5 sind die Pardé-Koeffizienten für alle untersuchten Einzugsgebiete zusammengestellt.

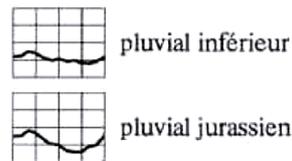


Abb. 4: Die beiden hier relevanten Abflussregime (gemäss BAFU, 1992).

Die Tatsache, dass alle Einzugsgebiete pluvial bestimmt sind, bietet eine gute Ausgangslage zur Modellierung mit HBV, da HBV ursprünglich für die Modellierung pluvial-nivaler Einzugsgebiete entwickelt wurde. Weniger geeignet wäre die bewusst einfach gehaltene, konzeptionelle HBV-Modellstruktur für Berechnungen in komplexem Gebirgsterrain.

Geologisch ist das untersuchte Gebiet ausserhalb des Juras hauptsächlich durch Moränenmaterial, Obere Süsswassermolasse (OSM) und Obere Meeresmolasse (OMM, am Südrand des Aaretals von Lenzburg Richtung Langenthal aufgeschlossen) charakterisiert; die Böden der grossen Täler sind mit fluvioglazialen Schottern verfüllt, welche mächtige Aquifere bilden; Reste von älteren fluvioglazialen Schottern lassen sich verschiedenerorts als Deckenschotter ausmachen; im Aare- und Rheintal finden sich an wenigen Stellen (Rheinfelden, Würenlingen, Unterentfelden) Lössablagerungen. Das Juragebirge wird durch Jura- und Triasgesteine gebildet, die von einem Band aus OSM- und Moränen-Ablagerungen von Brugg Richtung Laufen durchzogen werden. (nach Swisstopo, 2005)

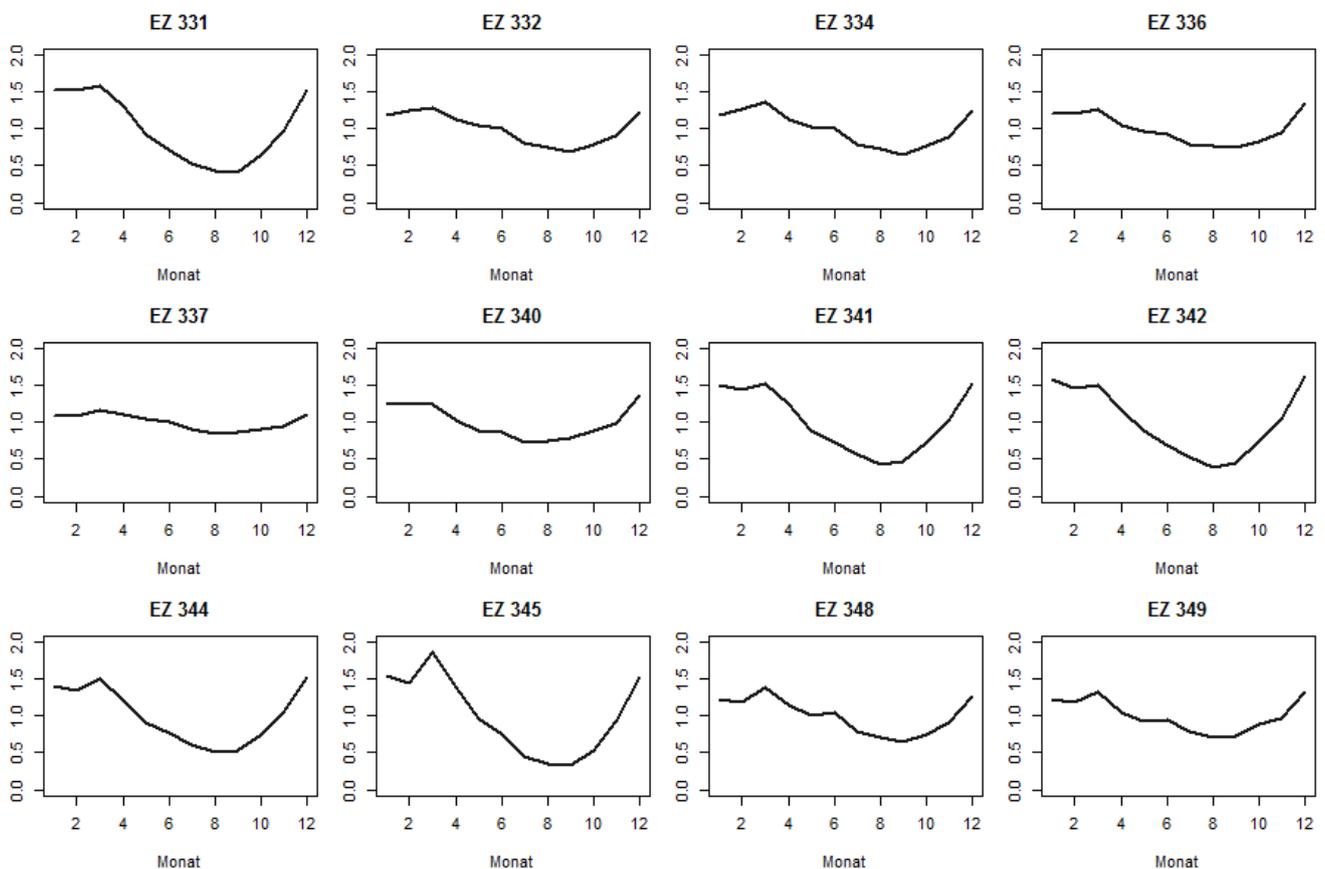
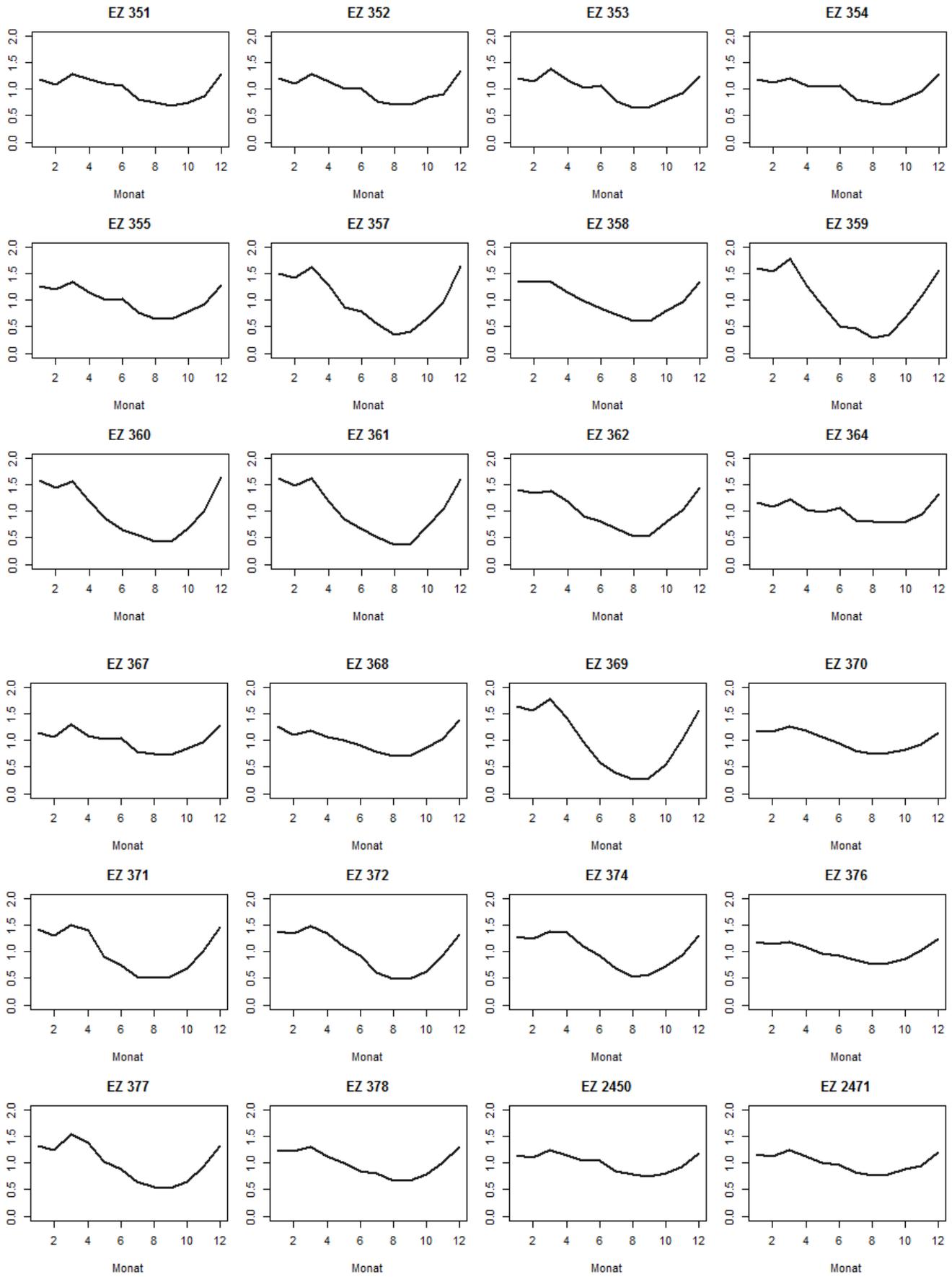


Abb. 5: Pardé-Koeffizienten aller untersuchten Einzugsgebiete. Anhand dieser Koeffizienten wird das hydrologische Regime charakterisiert.

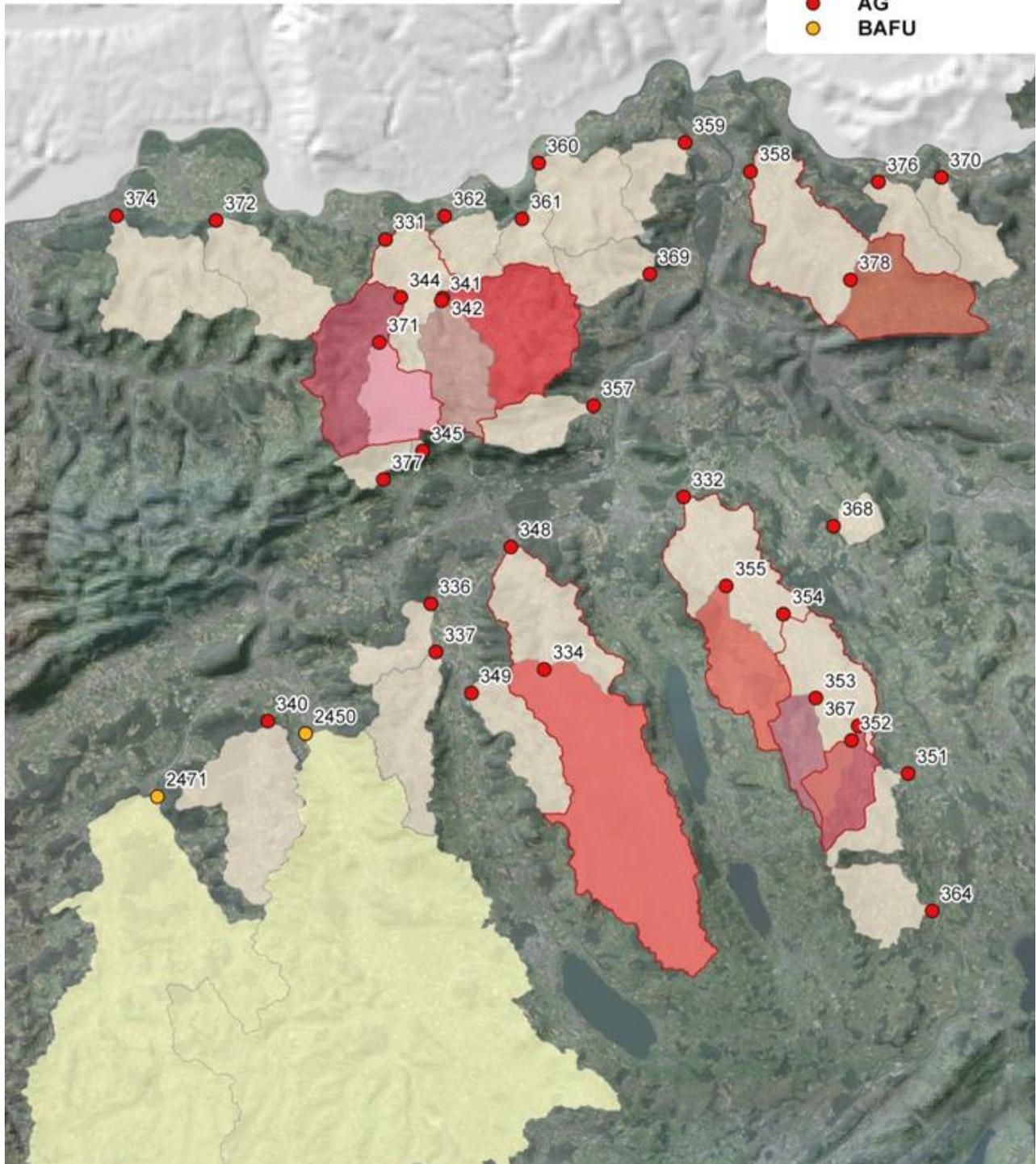


Fortsetzung Abb. 5

Übersicht Einzugsgebiete

Hydrometriestationen
TYP

- AG
- BAFU



0 3 6 12 km

Quellen: AGIS, BAFU, Swisstopo

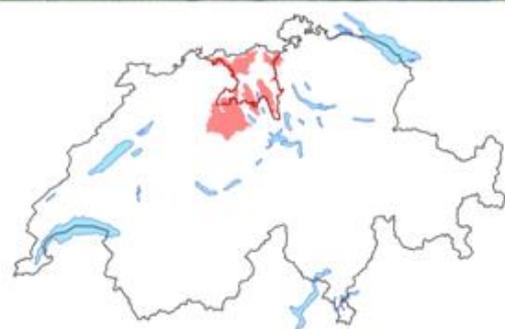


Abb. 6 (vorherige Seite): Übersicht der untersuchten Einzugsgebiete. Mehrere Gebiete sind Teilgebiete anderer Einzugsgebiete, daher sind diejenigen Gebiete, die andere Einzugsgebiete enthalten, fein rot umrandet. Teilgebiete sind zusätzlich farblich abgehoben. Die Einzugsgebiete, die durch den Kanton Aargau bemessen werden, sind rötlich eingefärbt, die beiden durch das BAFU bemessenen Gebiete sind gelblich eingefärbt.

Tab. 1: Liste aller verwendeten Stationen. Befinden sich mehrere Stationen am selben Gewässer, so sind sie in derselben Farbe eingefärbt. Die Mindestrestwassermenge ist gelb markiert, wenn sie durch die Untergrenze gemäss GSchG Art. 31, Abs. 1 festgesetzt ist. Achtung: Für HBV müssen $[m^3/s]$ in $[mm/d]$ umgerechnet werden.

NR	Ort	Gewässer	Grösse EZ $[km^2]$	Besteht seit	Höhe Station	Mittlere EZ-Höhe	Karst [%]	Q347 $[m^3/s]$	RWmin $[m_3/s]$
331	Eiken	Sissle	123.11	1977	314	529.57	54.0	0.095	0.078
332	Othmarsingen	Bünz	111.33	1977	390	531.70	0	0.458	0.261
334	Unterkulm	Wyna	91.95	1977	455	644.02	0	0.423	0.246
336	Kölliken	Köllikerbach	10.57	1979	421	489.18	0	0.026	0.050
337	Holziken	Uerke	24.98	1979	438	574.47	0	0.266	0.177
340	Vordemwald	Pfaffnern	39.07	1979	417	517.98	0	0.247	0.168
341	Frick	Staffeleggbach	21.00	1979	358	536.77	72.7	0.018	0.050
342	Hornussen	Sissle	37.15	1979	365	524.69	43.1	0.032	0.050
344	Gipf-Oberfrick	Bruggbach	44.78	1979	356	576.63	55.4	0.121	0.099
345	Küttigen	Wilenbergbach	0.61	1979	650	722.75	100.0	0.0005	0.050
348	Suhr	Wyna	120.84	1980	394	612.56	0	0.423	0.246
349	Schöftland	Ruederchen	19.04	1980	463	611.39	0	0.091	0.075
351	Merenschwand	Wissenbach	9.90	1980	392	565.76	0	0.031	0.050
352	Muri	Katzbach	5.89	1980	460	642.49	0	0.023	0.050
353	Boswil	Wissenbach	11.81	1980	460	685.48	0	0.047	0.050
354	Wohlen	Bünz	52.64	1980	421	554.56	0	0.220	0.156
355	Villmergen	Holzbach	23.74	1980	418	580.39	0	0.089	0.073
357	Schinznach-Dorf	Talbach	13.46	1980	358	559.03	91.7	0.008	0.050
358	Döttingen	Surb	67.25	1980	335	508.09	13.3	0.298	0.191
359	Leuggern	Guntenbach	10.31	1982	338	468.77	6.8	0.007	0.050
360	Etzgen	Etzgerbach	25.36	1980	308	477.42	39.6	0.061	0.051
361	Sulz	Sulzerbach	8.29	1980	352	497.94	45.9	0.013	0.050
362	Kaisten	Kaisterbach	11.94	1980	321	459.89	82.1	0.043	0.050
364	Sins	Sinserbach	16.38	1981	415	567.69	0	0.044	0.050
367	Muri (Hasli)	Bünz	15.65	1981	449	612.33	0	0.060	0.050
368	Künten (Sulz)	Küntenerbach	5.22	1981	381	585.20	0	0.017	0.050
369	Remigen	Schmittenbach	13.24	1982	385	522.66	43.3	0.002	0.050
370	Fisibach	Fisibach	15.02	1982	379	515.89	0	0.083	0.068
371	Witnau	Wölflinswilerbach	17.35	1982	395	593.05	88.8	0.045	0.050
372	Zeiningen	Möhlinbach	26.32	1982	338	513.15	60.8	0.066	0.055
374	Rheinfelden	Magdenerbach	33.09	1982	300	483.78	18.9	0.150	0.122
376	Wislikofen	Tägerbach	13.84	1982	390	547.27	0	0.092	0.076
377	Erlinsbach	Erzbach	6.72	1982	460	687.21	69.1	0.041	0.050
378	Ehrendingen	Surb	27.61	1991	424	539.28	5.9	0.217	0.155
2450	Zofingen	Wigger	368	1980	426	657.53	0	2.10	0.776
2471	Murgenthal, Walliswil	Murg	207	1981	419	654.44	0	1.60	0.621

4.1.4. Quellen möglicher Fehlinformationen

Bevor Daten in ein Modell eingespeist werden, sollte immer hinterfragt werden, ob sie nicht Fehlinformationen enthalten, die das Modellergebnis fehlleiten könnten (Beven, 2012: 270).

Für die vorliegenden Daten aus dem Kanton Aargau standen in dieser Beziehung zwei Themen im Vordergrund: Kleinkraftwerke und Hochwasserrückhaltebecken.

An den untersuchten 36 Gewässern liegen 19 Kleinwasserkraftwerke. Im Gegensatz zu grossen Flusskraftwerken ist ihr Einfluss auf die Abflusslinie aber sehr kurzfristig und nur im Tagesgang bei einer stündlicher oder gar minütiger Auflösung deutlich erkennbar. Im Tagesmittel, wie sie für diese Arbeit verwendet wurden, ist ihr Einfluss kaum merklich. Aus diesem Grund wurden Kleinwasserkraftwerke nicht weiter berücksichtigt.

Für die Hochwasserrückhaltebecken gestaltete sich die Abwägung deutlich schwieriger. In den betrachteten Einzugsgebieten befinden sich fünf Hochwasserrückhaltebecken, eines davon ist seit 1984 im Betrieb, die übrigen vier wurden zwischen 2002 und 2011 in Betrieb genommen. Eine Zusammenstellung ist auf der Webseite des Kantons Aargau unter folgendem [Link](#)³ verfügbar.

Grundsätzlich darf der Einfluss von Hochwasserrückhaltebecken in einer Modellierung nicht vernachlässigt werden, da sie das Abflussregime beeinflussen. Schliesslich wurde aber entschieden, ihren Einfluss aus folgenden Gründen dennoch zu ignorieren:

- Wie ihr Name besagt, verändern Hochwasserrückhaltebecken die Hochwasserhydrologie. In dieser Arbeit wird der Fokus aber auf Niedrigwasser gelegt und daher das Modell nicht spezifisch auf Hochwasserspitzen kalibriert. Die Verfälschung der Modellergebnisse dürfte daher gering sein.
- Die Becken wurden in der Messperiode nur sehr selten überhaupt aktiviert, und wenn, dann nur für ein oder zwei Tage. Betrachtet man die Abflusswerte der betroffenen Gewässer, so zeigt sich deutlich, dass an diesen Tagen der Abfluss kurzfristig deutlich – im Vergleich zu den Messwerten am Tag zuvor und danach sogar um ein Mehrfaches – erhöht ist. Werden diese Werte nun entfernt, wird damit ein essentieller Teil der Dynamik aus der Abflusskurve entfernt. Dies wäre für das Modell vermutlich die grössere Fehlinformation, als wenn die Abflussspitzen nicht ganz korrekt wiedergegeben sind.
- Es ist anzunehmen, dass die Abflussspitzen nur schwach verfälscht sind. Die Becken dienen oft nur über einige Stunden, selten für ein bis zwei Tage als Rückhalteraum. Daher ist anzunehmen, dass das Tagesmittel kaum verfälscht ist. Bei Daten mit einer stündlichen oder noch höheren Auflösung wäre der Einfluss hingegen nicht vernachlässigbar. Diese Annahmen wurden durch den Kantonsverantwortlichen bestätigt (schriftliche Auskunft durch Th. Gebert vom 22.1.2014).
- Der Einfluss der Hochwasserrückhaltebecken auf den Abfluss an der Messstation ist kaum abschätzbar, wenn das Becken nicht nahezu direkt vor der Messstation liegt. Ist die Messstation aber weiter entfernt, müsste der Abfluss so skaliert werden, dass er den Abflusswerten am Ort des Beckens entspricht. Gleichzeitig müssten auch die dazwischen einmündenden Zuflüsse berücksichtigt werden, über welche aber keine Informationen vorliegen.
- Grundsätzlich kann Hochwasser-Rückhalt auch auf natürliche Weise durch Gerinneblockaden auftreten. Solche Vorkommnisse zählen zum Unschärfbereich einer Modellierung. Bei der Kalibrierung werden solche Ereignisse vom Modell schliesslich als Charakteristik des

³https://www.ag.ch/de/bvu/umwelt_natur_landschaft/hochwasserschutz/hochwasserschutzmassnahmen/hochwasserrueckhaltebecken/hochwasserrueckhaltebecken_1.jsp (letzter Zugriff: 16.6.2014)

Einzugsgebietes betrachtet und die Parameter entsprechend angepasst. Problematisch wäre dies nur, wenn aus den kalibrierten Modell-Parametern direkte Rückschlüsse auf die hydrologischen Eigenschaften des Einzugsgebietes gezogen werden wollten. Weil dies in dieser Arbeit aber nicht bezweckt wird, ist ein solches Vorgehen gut vertretbar.

- Bis auf eines wurden alle Rückhaltebecken erst nach 2001 erstellt. In Anbetracht dessen, dass die für die Modellkalibrierung verwendeten Zeitreihen an den betroffenen Gewässern 31 bis 35 Jahre umfassen, scheint die allfällige Verfälschung durch die Hochwasserrückhaltebecken nicht gewichtig zu sein.

Aufgrund der hier aufgeführten Überlegungen wurde schliesslich entschieden, den Einfluss der Hochwasserrückhaltebecken nicht zu berücksichtigen. Ausschlaggebend waren dabei vor allem Argumente der Einfachheit und der Unsicherheit.

4.2. Grundwasserdaten

Der Grossteil der verwendeten Grundwasserdaten stammen vom Kanton Aargau⁴. Für das Einzugsgebiet der Wigger (EZ 2450) und der Murg (EZ 2471) wurden zusätzlich Daten des Kantons Luzern⁵ bzw. des Kantons Bern⁶ im Tagesmittel verwendet.

Die Messungen des Grundwasserpegels werden an Trinkwasserfassungen durchgeführt. Dies bedeutet, dass die Messwerte durch die Trinkwasserentnahme beeinflusst werden. Zumindest bis zu einem gewissen Grad wird Vergleichbarkeit aber dadurch erreicht, dass die Daten jeweils zum gleichen Zeitpunkt täglich bzw. wöchentlich erfasst werden. Die grafische Darstellung der Daten bestätigt, dass die Grundwasserspiegel ein plausibles Verhalten zeigen. Die Daten werden allerdings im Kanton Aargau nach der Erfassung nicht qualitätsgeprüft. Diesen Daten gegenüber sind daher deutlich grössere Vorbehalte angebracht als bei den Abflussdaten.

Im Kanton Aargau werden die Messungen nach wie vor manuell durchgeführt (mündliche Auskunft durch D. Schaub vom 14.8.2013). Im Kanton Luzern ist die Pegelmessung heute digitalisiert; von den verwendeten Stationen sind alle mit Drucksonden ausgestattet, mit Ausnahme einer Station, die über ein Schwimmersystem verfügt. Auch im Kanton Bern wurden die ursprünglich alle mit Schwimmer ausgestatteten Stationen innerhalb der letzten fünf bis sechs Jahre digitalisiert und sind heute mit einer Drucksonde ausgerüstet. Die Daten werden täglich überprüft, allfällige Datenlöcher korrigiert und die Messungen aller Stationen einmal monatlich per Lichtlot überprüft (mündliche Auskunft durch B. Baumann vom 3. 6. 2014).

4.2.1. Stationsauswahl und Datenaufbereitung

Aus verschiedenen, anschliessend erläuterten Gründen konnten bei Weitem nicht alle Grundwasser-Messstationen, die sich innerhalb der Einzugsgebiete befinden, verwendet werden. Zur Auswertung standen schliesslich 13 Einzugsgebiete mit insgesamt 33 Grundwassermessstellen (22 des Kantons Aargau, 6 des Kantons Luzern, 5 des Kantons Bern) zur Verfügung; Abb. 7 bietet einen geografischen Überblick und in Tab. 2 sind die Stationen mit ihren wichtigsten Eigenschaften pro Einzugsgebiet zusammengestellt. Da verschiedene Einzugsgebiete Teilgebiete anderer Einzugsgebiete darstellen, gelten manche Grundwasser-Messstationen für mehrere Einzugsgebiete. Für vier Einzugsgebiete

4

https://www.ag.ch/de/bvu/umwelt_natur_landschaft/umweltinformationen/wasser_2/grundwasser_2/grundwasser_1.jsp (letzter Zugriff 21.5.2014)

⁵ <http://www.uwe.lu.ch/index/themen/gewaesser/hydrometrie/grundwasserstand.htm> (letzter Zugriff: 21.5.2014)

⁶ <http://www.bve.be.ch/de/index/wasser/wasser/grundwasser.html> (letzter Zugriff: 21.5.2014)

wurde zusätzlich zum Standard-Modell die Delay-Variante (siehe Kap. 5.7) kalibriert. Von diesen vier Gebieten besitzen zwei Anteile an Karstgebieten (Nr. 341 und 360), die anderen beiden sind nicht verkarstet (Nr. 348 und 353).

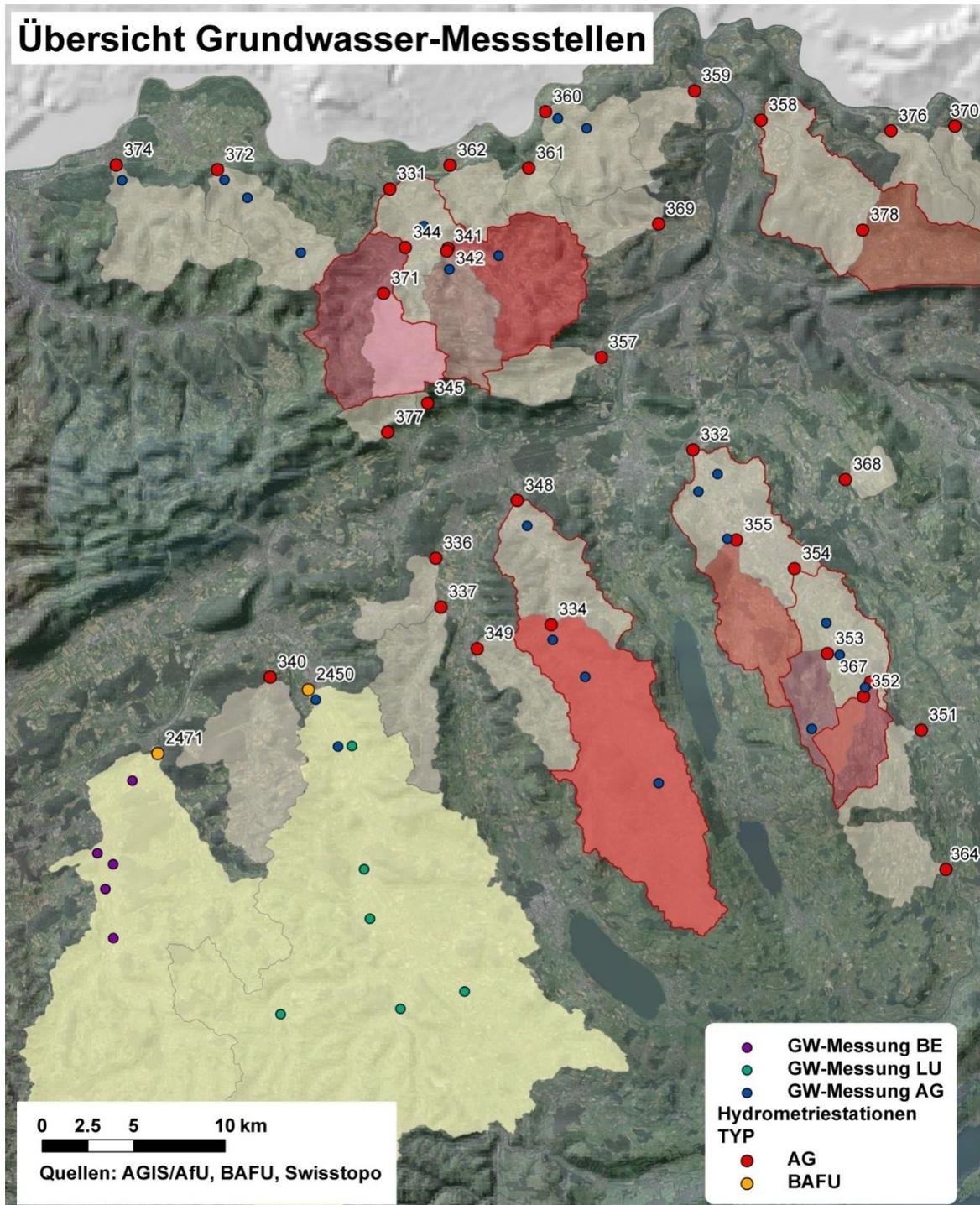


Abb. 7: Lage der verwendeten Grundwasser-Messstationen.

Zwölf Messstationen fielen aufgrund deutlich kürzerer Messperioden als die für den Abfluss verwendeten weg. Weitere Gründe für das Nichtberücksichtigen von Messstationen waren:

- Der Grundwasserspiegel steigt zeitweise über die Oberkante (Stationen 334 / 348).
- Die Daten weisen einen deutlichen Trend auf (Station 161).
- Der Spiegel scheint nach oben begrenzt zu sein, obwohl er nicht die Oberkante erreicht (Station 129).

- Der Pegel zeigt ein ungewöhnliches Verhalten und weist eine sehr schwache Korrelation zu den übrigen Messstellen auf. Aus solchen Messreihen können keine hilfreichen Informationen gewonnen werden (Stationen G267, 114).
- Die Daten zeigen ein deutliches Streifenmuster, was auf Qualitätsprobleme schliessen lässt (Stationen 220, 216).

Aufgrund solcher Datenprobleme konnte für die Einzugsgebiete 344, 349, 355, 358, 367, 378 trotz eigentlich vorhandenen Messstellen keine Co-Kalibrierung auf Grundwasser durchgeführt werden.

An drei Messreihen mussten zusätzlich Korrekturen vorgenommen werden: Bei Station 178 (Gebiete 332 und 354) musste die gesamte Periode anfangs Januar bis anfangs März 2006 entfernt werden, weil dieser Zeitraum mit einem Versatz nach unten behaftet war. Bei Stationen 159 und 160 (Gebiet 360) wurden deutliche Ausreisser entfernt. Allerdings bleibt eine gewisse Skepsis gegenüber der Datenqualität bestehen. Auch bei Messstelle 190 (Gebiete 332 / 354) mussten Ausreisser entfernt werden.

Tab. 2: Zusammenstellung der verwendeten Einzugsgebiete und ihrer Grundwasser-Messstationen. Angegeben sind neben Nummer, Name und Standortgemeinde der Stationen die Höhenlage der Terrain-Oberkante, die euklidische Distanz zum nächsten Gewässer (gemäss dem Gewässernetz von VECTOR25 von swisstopo), der Spearman-Korrelationskoeffizient zwischen Grundwasserpegel und Abfluss sowie die euklidische Distanz zur Abfluss-Messstation. Messstationen des Kantons Luzern sind in blauer Schrift, solche des Kantons Bern in roter Schrift markiert.

Nr.	Name	Gemeinde	Terrain-Oberkante [m.ü.M.]	Distanz zum Gewässer [m]	Korrelation zum Abfluss [-]	Distanz zur Q-Messstation [m]
EZ 331						
137	Mühlematt	Bözen	393.3	29.89	0.579	6'993.71
165	Blackimatt	Ueken	377.5	48.28	0.658	5'506.51
229	Neumatt 2	Frick	340.6	162.23	0.344	2'756.75
EZ 332						
178	Lippertswies I und II	Muri	451.0	42.02	0.405	16'149.41
186	Femmatten	Boswil	431.0	181.90	0.359	12'016.91
173	Grundacker	Hendschiken	415.7	182.44	0.470	2'320.47
189	Buchen	Buttwil	761.7	247.60	0.343	16'716.09
131	Unterzelg	Villmergen	424.2	333.83	0.428	5'258.25
190	Grundächer	Boswil	443.9	463.75	0.383	13'877.01
130	Schützenhaus	Dottikon	419.0	604.48	0.324	1'895.06
EZ 334						
153	Russirain	Menziken	557.9	13.99	0.375	10'540.76
148	Neumatt	Oberkulm	492.4	54.12	0.419	3'434.54
149	Hueb	Unterkulm	462.2	75.70	0.281	844.99
EZ 341						
165	Blackimatt	Ueken	377.5	48.28	0.675	1'025.88
EZ 342						
137	Mühlematt	Bözen	393.3	29.89	0.614	2'789.44
EZ 348						
153	Russirain	Menziken	557.9	13.99	0.335	17'411.03
105	Töndler	Gränichen	404.2	30.48	0.392	1'504.05
148	Neumatt	Oberkulm	492.4	54.12	0.256	10'421.63
149	Hueb	Unterkulm	462.2	75.70	0.586	7'931.65
EZ 353						
189	Buchen	Buttwil	761.7	247.60	0.350	4'246.01
EZ 354						
178	Lippertswies I und II	Muri	451.0	42.02	0.412	7'619.91
186	Femmatten	Boswil	431.0	181.90	0.266	3'466.78
189	Buchen	Buttwil	761.7	247.60	0.332	8'899.53
190	Grundächer	Boswil	443.9	463.75	0.409	5'371.28

EZ 360						
159	Kirchmatt	Mettauertal-Will	355.6	20.89	0.696	2'424.68
160	Brunnmatte	Mettauertal-Mettau	327.1	51.26	0.644	769.68
EZ 372						
202	Unter Reben 1	Zeiningen	343.2	8.70	0.637	676.76
198	Thalmatt	Wegenstetten	431.6	12.53	0.678	6'458.02
192	Niederhofen	Zuzgen	367.0	16.29	0.825	2'250.69
EZ 374						
193	Enge	Magden	311.5	24.78	0.646	899.60
EZ 2450						
207	PW II Trinerplatten	Zofingen	427.1	75.39	0.489	688.77
211	Pritschmatte	Brittnau	445.4	129.92	0.518	3'531.16
503	RE 503	Reiden	450.8	255.34	0.493	3'912.54
6541	NE 6541	Nebikon	484.8	67.68	0.438	10'368.73
112	SZ 112 / LUG 05	Schötz	505.7	586.06	0.540	13'081.53
701	ZE 701	Zell	627.3	107.46	0.379	17'974.38
103	WL 103	Willisau-Stadt	549.1	42.37	0.319	18'326.65
102	GR 102 / LUG 10	Grosswangen	534.9	86.29	0.490	18'721.30
EZ 2471						
G250	MAB01	Madiswil	536.7	49.97	0.417	10'475.73
G252	LOB21	Lotzwil	507.1	67.25	0.674	8'012.13
G253		Lotzwil	494.0	118.37	0.379	6'574.54
G161	ROB51	Roggwil	454.6	220.30	0.489	2'044.76
G256	LTB20	Langenthal	486.2	256.01	0.479	6'423.48

Fortsetzung von **Tab. 2.**

4.2.2. Korrelationen und Dateninterpolation

Für eine erste Abschätzung, ob die Grundwasserdaten brauchbar und hilfreich sind, wurde die Korrelation zwischen den Grundwasserpegeln und dem Abfluss sowie zwischen den Grundwasserstationen untereinander untersucht (siehe dazu auch Kapitel 5.8). Zwischen Abfluss und den Grundwasserpegeln wurde über alle Stationen und Einzugsgebiete hinweg eine mittlere Korrelation von 0.47 mit einer Standardabweichung von 0.15 beobachtet (die einzelnen Werte sind in Tab. 2 aufgeführt). Die Korrelationskoeffizienten der Grundwasser-Stationen untereinander sollen hier nur zusammenfassend erwähnt werden: Über alle Einzugsgebiete und Messstationen hinweg zeigen die Grundwasserpegel untereinander eine mittlere Korrelation von 0.605 (Median 0.629) mit einer Standardabweichung von 0.22. Die maximal erreichte Korrelation beträgt 0.977, die minimale -0.05. Eine solche negative Korrelation wurde nur einmal erreicht, die nächst geringste Korrelation beträgt 0.1. Solch geringe Werte wurden in Kauf genommen, falls sie nur zwischen einzelnen Stationen desselben Einzugsgebietes auftraten. Zeigte eine Station hingegen gegenüber allen anderen Stationen besonders geringe Korrelationswerte, wurde diese entfernt.

Problematisch an den Grundwasserdaten ist, dass die Messungen im Kanton Aargau meist nicht täglich, sondern wöchentlich bis zweiwöchentlich, teilweise unregelmässig sowie zeitlich nicht koordiniert festgehalten wurden. Es kann daher sein, dass an zwei Messstellen in einem Einzugsgebiet zwar regelmässig wöchentlich, aber jeweils um einige Tage versetzt gemessen wurde. Daher müssen die Daten erst interpoliert werden, bevor sie verrechnet werden können, um eine mittlere Grundwasser-Reaktionskurve für jedes Einzugsgebiet zu erhalten. Die Grundwasserstände wurden daher in Abhängigkeit der zeitlichen Abstände linear interpoliert (R-Funktion ‚na.approx‘ im Paket ‚zoo‘). Dabei wurde maximal eine Lücke von 30 Tagen interpoliert (R-Parameter ‚maxgap‘). Als Alternative wurde die Spline-Methode (berechnet mit Cubic Convolution) betrachtet; diese erzeugte jedoch nervös zitternde Resultate, die über die Messwerte hinaus schwankten – für Grundwasser wäre dies ein unrealistisches Verhalten.

Weil die Grundwasserspiegel der verschiedenen Messstellen eines Einzugsgebietes auf unterschiedlichen Niveaus unter dem Terrain liegen und unterschiedliche Amplituden aufweisen, wurden die Messungen gemäss Gleichung (1) standardisiert, bevor sie miteinander verrechnet wurden (R-Funktion ‚scale‘).

$$(1) \quad Z = \frac{X - \mu}{\sigma}$$

Anschliessend wurde pro Einzugsgebiet der Mittelwert aus den standardisierten Grundwasser-Werten berechnet. Dabei wurden teilweise sehr grosse Lücken (maximal 30 Tage) aufgrund recht weniger Werte interpoliert. Um den Fehler aufgrund möglicherweise falsch interpolierter Daten zu reduzieren, wurden im Anschluss an die Mittelwertberechnung nur diejenigen Zeitpunkte extrahiert, an denen tatsächlich eine Messung an irgendeiner der Messstellen im Einzugsgebiet durchgeführt worden war. Alle übrigen Werte werden als NAs („nicht verfügbar“, engl. „not available“) behandelt und dafür mit dem Wert „-9999“ bezeichnet. Dieser wird von HBV als NA erkannt und fliesst damit nicht in die Gütemassberechnung mit ein. Einen Eindruck der Ergebnisse dieser Datenaufbereitung gibt Abb. 8.

Fünf Einzugsgebiete (341, 342, 353, 374, 378) enthalten nur eine einzige Grundwassermessstelle, was eine Interpolierung, Standardisierung und Mittelwertberechnung natürlich überflüssig macht.

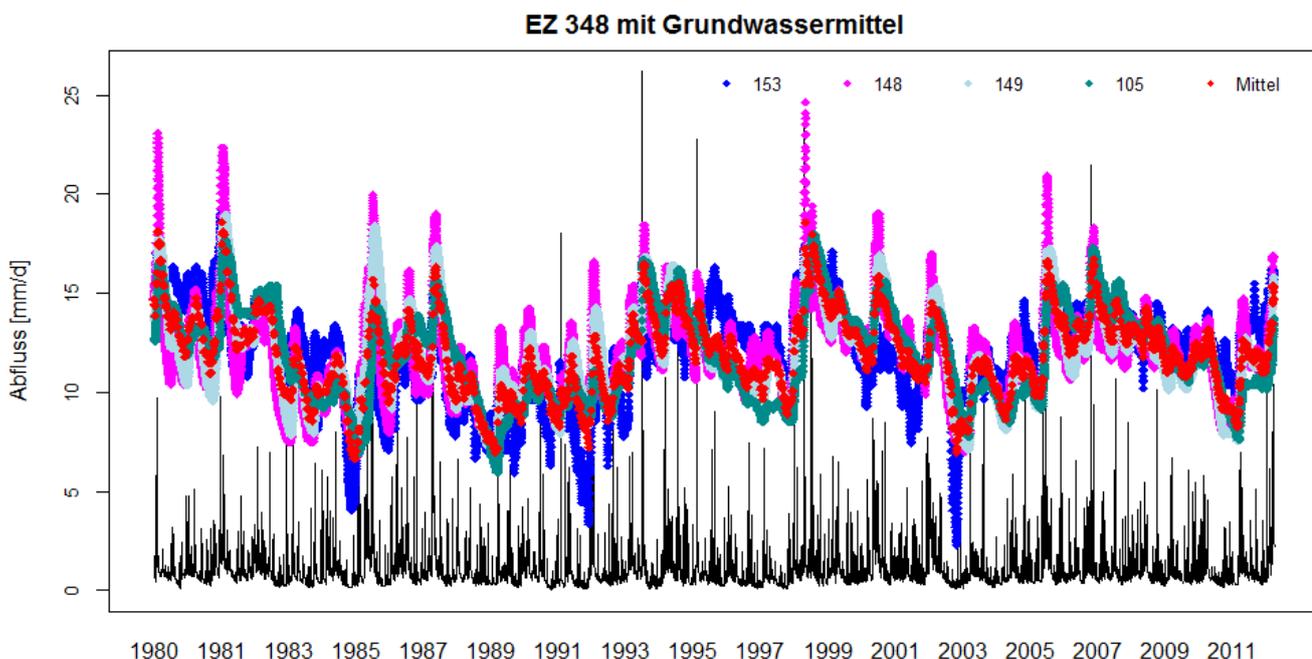


Abb. 8: Einzugsgebiet 348 mit seinen Grundwasser-Messstationen. In Rot ist die berechnete mittlere Grundwasserkurve dargestellt. In schwarz hinterlegt ist die Abflusskurve. Die Grundwasserspiegel sind normiert, um visuelle Vergleichbarkeit zu erhalten; die Skala fehlt daher.

4.3. Niederschlags- und Temperaturdaten

Die in dieser Arbeit verwendeten Niederschlags- und Temperatur-Messdaten stammen von MeteoSchweiz. Die Daten wurden einem bereits von der MeteoSchweiz interpolierten Rasterdatensatz entnommen und über die Fläche jedes Einzugsgebietes gemittelt.

Verwendet wurden Datensätze von Tagesmittelwerten, die eine räumliche Auflösung von 0.0208° aufweisen, was in West-Ost-Richtung einer Rasterbreite von ca. 1.6 km und in Nord-Süd-Richtung ca. 2.3 km entspricht. Die von MeteoSchweiz verwendete Interpolationsmethode für Niederschlagswerte berücksichtigt Topographie-Effekte und basiert auf einem langjährig gemittelten

klimatologischen Referenzfeld, um insbesondere die geringe Anzahl von Messstationen in höheren Lagen zu korrigieren und damit das Risiko für systematische Fehler zu verringern. Detailliertere Informationen zur Datenkorrektur und Interpolationsmethode sind der Datendokumentation unter folgendem [Link⁷](#) zu entnehmen.

Ähnlich wie die Niederschlagsdaten werden auch die Temperaturdaten von MeteoSchweiz homogenisiert und topographieabhängig interpoliert. Das Verfahren ist ebenfalls in der entsprechenden Datendokumentation (verfügbar unter folgendem [Link⁸](#)) beschrieben.

Ein möglicher Stolperstein für erfolgreiche Modellierungen könnte die im Verhältnis zu der Grösse der Einzugsgebiete geringe räumliche Auflösung der Daten darstellen, denn eine Rastergitterzelle weist eine Fläche von knapp 3.7 km² auf. Dies ist insbesondere für das kleinste der verwendeten Einzugsgebiete (345) mit einer Fläche von 0.6 km² problematisch – das Einzugsgebiet ist also deutlich kleiner als eine Rastergitterzelle. Vermutlich aus diesem Grund hat die automatisierte Extraktion der Niederschlags- und Temperaturwerte für dieses Gebiet nicht funktioniert. Daher mussten die Werte des direkt benachbarten Gebietes 377 verwendet werden. Alle übrigen Gebiete verfügen über Flächen, die grösser als eine Rasterzelle sind.

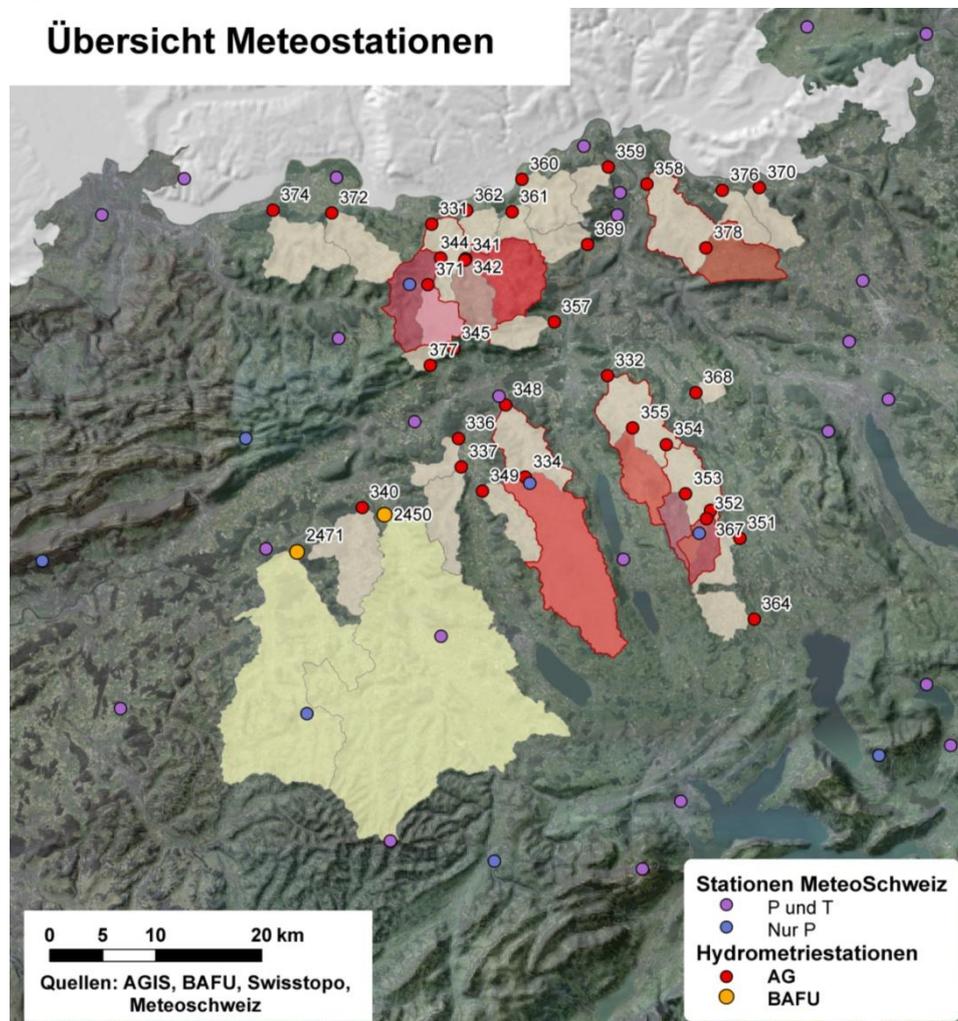


Abb. 9: Lage der meteorologischen Messstationen im und um das Untersuchungsgebiet.

⁷ <http://www.meteoschweiz.admin.ch/web/de/services/datenportal/gitterdaten/precip/rhiresd.html> (letzter Zugriff: 27.2. 2014)

⁸ <http://www.meteoschweiz.admin.ch/web/de/services/datenportal/gitterdaten/temp/tabsd.html> (letzter Zugriff: 27.2. 2014)

Gewisse Vorbehalte bezüglich der Genauigkeit der Niederschlagswerte sind auch angebracht, weil die Messstationen im Untersuchungsgebiet unregelmässig und teilweise sehr spärlich angeordnet sind (Abb. 9 bietet einen Überblick über die Verteilung der meteorologischen Messstationen). Beispielsweise ist Einzugsgebiet 368 (5.2 km²) weit von den nächstgelegenen Messstationen entfernt. Gerade ein solch kleines Einzugsgebiet wird bei konvektiver Wetterlage massgeblich durch lokale Niederschläge geprägt. Je weiter die nächstgelegenen Messstationen entfernt sind, desto geringer ist die Wahrscheinlichkeit, dass sie solche Ereignisse korrekt erfassen.

Des Weiteren sind Vorbehalte insbesondere in den östlichen Jurahöhen gegenüber Temperatur und Niederschlag angebracht, die nur aufgrund von Messungen aus den Talböden im Rhein- und Aaretal berechnet werden. Zwar führt MeteoSchweiz, wie bereits erwähnt, Topographie-abhängige Korrekturen durch, dennoch ist ein gewisser Bias zu erwarten.

Zu Abb. 9 ist anzumerken, dass die Angaben auf dem Stand der Homepage des SwissMetNet⁹ vom 28. Mai 2014 beruhen. Das Messnetz unterliegt aber zeitlichen Veränderungen: Neue Stationen werden hinzugefügt, alte werden aufgehoben, zudem können Messausfälle auftreten. Auf der Homepage werden nur bereits digitalisierte Stationen aufgeführt und es kann nicht direkt zurückverfolgt werden, zu welchem Zeitpunkt welche Stationen wie stark in die Berechnung der interpolierten Datensätze einfließen.

4.4. Evapotranspiration

HBV benötigt als Input Angaben zur PET im langjährigen Monatsmittel. Diese ist allerdings schwierig zu bestimmen, insbesondere als repräsentativer Wert für ein gesamtes Einzugsgebiet: Messungen werden jeweils nur an einem bestimmten Punkt durchgeführt und sind daher nicht unbedingt allgemeingültig für ein Gebiet. Daher wird die PET für eine Abfluss-Modellierung meistens berechnet. Dafür bieten sich eine ganze Anzahl verschiedener Formeln in unterschiedlichen Komplexitätsgraden an. Häufig benutzt werden die Penman- (Penman, 1948) oder die Penman-Monteith- (Monteith, 1965) Formeln, die beide auf physikalischen Zusammenhängen basieren, aber relativ komplex sind und einen entsprechend umfangreichen Input an meteorologischen Daten (relative Luftfeuchtigkeit, Windgeschwindigkeit, Temperatur und Sonneneinstrahlung) verlangen.

Oudin et al. (2005) haben jedoch aufgezeigt, dass für die Abflussmodellierung aus Niederschlag einfachere PET-Berechnungen gleich gute oder gar bessere Resultate erzielen als kompliziertere Varianten. Unter anderem bestätigten Kay & Davies (2008) dieses Ergebnis. Oudin et al. (2005) empfehlen, die Wahl der Formel zur Berechnung der PET auf zwei Kriterien zu basieren: Auf der erzielten Modelleffizienz und bei gleicher Effizienz auf der Einfachheit der Formel. Aus zwei sich sehr ähnlichen Methoden von Jensen & Haise (1963) und McGuinness & Bordne (1972) bestimmten Oudin et al. (2005: 302) eine einfache Formel (2), die als Ausgangsdaten nur die Tagestemperatur T_a , gemittelt über eine langjährige Zeitreihe verlangt:

$$(2) \quad PET = \frac{Re * (T_a + 5)}{\lambda * \rho * 100} \quad \text{wenn } T_a + 5 > 0, \text{ sonst } PET = 0$$

Re steht dabei für die kosmische Solarstrahlung, die aufgrund des Julianischen Datums und der Breitenlage bestimmt wird. λ bezeichnet den latenten Wärmefluss (2.45 MJ*kg⁻¹) und ρ die Dichte von Wasser (1'000 kg*m⁻³). Obwohl diese Berechnung hauptsächlich temperaturbasiert zu sein

⁹ <http://www.meteoschweiz.admin.ch/web/de/klima/messsysteme/boden/swissmetnet.html> (letzter Zugriff: 30. 5. 2014)

scheint, beruht die Berechnung wegen der Verwendung des Julianischen Datums und der Breitenlage mindestens ebenso stark auf der Solarstrahlung (Shaw & Riha, 2011: 1467).

Ausgehend von den Erkenntnissen von Oudin et al. (2005) wurde diese relativ simple Berechnung der PET in dieser Arbeit angewandt. Diese einfache Berechnung wird weiter gerechtfertigt durch die relativ simpel gehaltene Berechnung der AET aus der PET in HBV.

4.5. Höhenkorrektur

HBV erlaubt den Einbezug von verschiedenen Höhenstufen, um die Berechnung der Niederschlags- und Temperaturverteilung (jedoch nicht die potentielle Evapotranspiration) und damit vor allem der Schneeakkumulation und –ablation realitätsgetreuer widerzugeben. Für die Modellierungen wurden Höhenstufen in 100m-Schritten verwendet..

Zur Korrektur wurde ein Temperaturgradient von -0.6°C pro 100m (Parameter TCALT) und zur Niederschlagskorrektur ein Gradient von 0 bis 25% pro 100m (Parameter PCALT) verwendet. Diese Korrektur wird vom Modell ausgehend von den Parametern Telev und Pelev, die eigentlich die mittlere Höhenlage der meteorologischen Messstationen bezeichnen, berechnet. Aus praktischen Gründen wurde hier jedoch die mittlere Einzugsgebietshöhe verwendet (siehe auch Kap. 5.5.3).

Die Hangexposition wurde nicht mit einbezogen. In gebirgigem Gelände wäre dies sicherlich notwendig, im Mittelland wird der Einfluss der Exposition auf die Hydrologie aber als nicht so stark angenommen, dass es sich lohnen würde, die Höhenstufen zusätzlich mit der Exposition zu kombinieren.

4.6. Stationarität der Datenreihen

Streng genommen sind jegliche Prognosen nur unter der Bedingung von Stationarität gültig, was in Bezug auf den Abfluss bedeutet, dass das durchschnittliche Abflussverhalten über die Zeit konstant bleibt. Diese Bedingung fällt stärker ins Gewicht, je weiter in die Zukunft extrapoliert wird.

Stationarität kann aus verschiedenen Gründen gebrochen werden: Aufgrund des Klimawandels, durch Urbanisierung und Bodenversiegelung, durch Nutzungsänderungen etc. (Robson, Bardossy & Kundzewicz, in: WMO, 2000: 58/59).

In den Daten macht sich Nicht-Stationarität auf verschiedene Arten bemerkbar (Pegram, in: WMO, 2000: 104):

- Als steter Trend
- Als Änderung des Mittelwertes (engl. *Shift, Step-change*)
- Als Veränderung der Streuung (Heteroskedastizität)

oder als eine Kombination dieser Punkte.

Nicht-stationäre Daten sind für Prognosen ein beträchtliches Hindernis, denn Nicht-Stationarität bedeutet, dass aus den Beobachtungen der Vergangenheit nicht - oder zumindest nicht direkt - auf die Zukunft geschlossen werden kann. Nicht-stationäre Datenreihen sind daher streng genommen nicht für Prognosezwecke geeignet. Allerdings können in der Praxis Daten, die nur unter Vorbehalten zu betrachten sind, immer noch hilfreicher als keine Daten sein – insbesondere im Zeitalter des Klimawandels, der Stationarität per se fraglich macht und Prognosen allgemein mit einer grösseren Unsicherheit behaftet (z.B. Milly et al., 2008; Milly, Dunne, & Vecchia, 2005; Vaze et al., 2010; Kundzewicz et al., 2008). Daher kann argumentiert werden, dass die Datenreihen dennoch verwendet, aber mit der nötigen zusätzlichen Vorsicht interpretiert werden sollten. In dieser Arbeit werden die Daten so gehandhabt.

Weil die Frage der Stationarität für Prognosen ein nicht zu vernachlässigendes Thema darstellt, werden die zur Modellierung verwendeten Abflussdaten an dieser Stelle auf Stationarität getestet. Diese Untersuchung ist jedoch nicht Teil des eigentlichen Fokus dieser Arbeit, sondern dient lediglich der Überprüfung der Dateneignung, daher werden die Ergebnisse hier anstatt in Kap. 6 aufgeführt. Es wird ausserdem auf eine Erläuterung der Begriffe und Konzepte verzichtet; für eine Übersicht wird Pfaff (2008) empfohlen.

Zur Überprüfung von Stationarität werden in der Literatur verschiedenste Testverfahren vorgeschlagen (z.B. Pfaff, 2008: 105); die meisten davon entstammen der Ökonometrie. Als Ausgangspunkt zur Bestimmung von Verfahren, die sich zur Analyse hydrologischer Zeitreihen eignen, diente die Auflistung der WMO (2000: 56ff). Weil Nicht-Stationarität in verschiedenen Formen auftreten kann, werden normalerweise verschiedene Verfahren kombiniert; oft wird erst ein Test zur Detektion von Trends und/oder Shifts angewandt und anschliessend mit Einheitswurzel-Tests weiter untersucht (z.B. Rutkowska & Ptak, 2012; Wang, van Gelder, & Vrijling, 2003). Das hier gewählte Verfahren lehnt sich an dieses Vorbild an: Es werden Tests auf Trend und auf Einheitswurzel bzw. Stationarität durchgeführt; die verwendeten Verfahren werden im Folgenden kurz vorgestellt.

Vorauszuschicken ist, dass alle Tests an den mittleren Jahresabflüssen (arithmetisches Mittel) durchgeführt wurden. Jahresmittel wurden verwendet, um die starke Autokorrelation der Daten möglichst zu eliminieren und robuste, aussagekräftige Daten zu erzeugen.

Mann-Kendall-Test: Dieses nicht-parametrische Verfahren geht auf Mann (1945) und Kendall (1938) zurück und untersucht, ob die Datenreihe einen Trend aufweist. Wie der verwandte Spearman's Rho (ρ)-Test ist er rangbasiert, setzt damit keine Verteilung voraus, verwendet aber im Unterschied zur Spearman-Methode das nicht-parametrische Korrelationsmass Kendall's Tau (τ). Wie Spearman's Rho ist dieser Test robust gegenüber Extremwerten und Abweichungen von einer linearen Beziehung. (WMO, 2000; Wang & Swail, 2001; Yue, Pilon, Phinney, & Cavadas, 2002)

Die Test-Statistik kann WMO (2000: 72/73) oder Wang & Swail (2001: 2219) entnommen werden. Der Test arbeitet mit folgenden Hypothesen:

H_0 : Die Daten sind zufällig (unabhängig und gleichverteilt).

H_1 : Die Daten weisen einen (monotonen) Trend auf.

Die Alternativhypothese weist darauf hin, dass ein detektierter Trend nicht linear, aber monoton sein muss. Zur Berechnung wurde die R-Funktion ‚cor.test‘ mit der Methode ‚kendall‘ und der Alternative ‚two-sided‘ verwendet. In Tab. 3 sind die erhaltenen P-Werte zusammengestellt.

Ein Problem des Mann-Kendall-Tests ist jedoch, dass er sensibel auf Autokorrelation in den Daten reagiert. Autokorrelierte Daten bewirken, dass der Test die Nullhypothese häufiger verwirft, als dies das Signifikanzniveau vorgibt; der Test ist damit nicht zuverlässig (z.B. Wang & Swail, 2001). Von Storch (1995: 17) empfiehlt daher ein sogenanntes „Prewhitening“ der Daten gemäss Formel (3). Dabei wird der Korrelationskoeffizient bei Zeitschritt (engl. lag) 1 (r_1) verwendet, um die Daten zu korrigieren.

$$(3) \quad x'_t = x_t - r_1 * x_{t-1}$$

Yue & Wang (2002) zeigen auf, dass Prewhitening zwar die Varianz der Daten erhöht, aber auch einen Teil des Trends entfernt und damit das Testresultat falsch beeinflussen kann. Sie halten Prewhitening daher für ungeeignet zur Datenaufbereitung und raten (unter bestimmten Bedingungen) zu einer Verwendung der rohen Daten. Zhang & Zwiers (2004) halten dem entgegen, dass durch ein iteratives Verfahren nach Wang & Swail (2001) Autokorrelation angemessen berücksichtigt und korrigiert werden kann.

Für diese Arbeit wurde eine vereinfachte Kompromissvariante gewählt: Der Test wurde für alle Gebiete, die eine signifikante (das 95%-Konfidenzintervall überschreitende) Autokorrelation aufweisen, zweimal durchgeführt, einmal ohne und einmal mit Prewhitening gemäss Formel (3). Dabei bestätigte sich, dass das Prewhitening die Stärke des Trends reduziert. Daraus wurde folgende Interpretationsregel abgeleitet: Wenn der P-Wert vor dem Prewhitening das 5%-Signifikanzniveau unterschreitet und auch nach dem Prewhitening noch unter einem Signifikanzniveau von 10% liegt, wird der Trend als statistisch signifikant interpretiert. Dieses vereinfachte Verfahren wird hier für berechtigt gehalten, weil diese Stationaritätsberechnungen nicht im eigentlichen Fokus der Arbeit liegen, sondern nur eine ergänzende Untersuchung der Daten darstellen. In Tab. 3 sind daher in der Spalte ohne Prewhitening alle Werte unter dem Signifikanzniveau von 5% rot eingefärbt, in der Spalte mit Prewhitening alle Werte unter 10%. Ergänzend sind in ersteren Spalte alle Werte unter dem 10%-Signifikanzniveau gelb markiert.

Phillips-Perron-Test (PP-Test): Dieses Verfahren geht auf Phillips & Perron (1988) zurück und ist eine nicht-parametrische Abwandlung des Dickey-Fuller-Tests. Es verwendet eine nicht-parametrische Korrektur, die leichte Autokorrelation und Heterogenität der Fehler erlaubt (Zivot & Wang, 2006). Die Hypothesen lauten wie folgt:

H_0 : Die Daten weisen eine Einheitswurzel (nicht-stationärer Prozess) auf.

H_1 : Die Daten weisen Stationarität auf.

Eine detaillierte Beschreibung der Teststatistik findet sich beispielsweise in Pfaff (2008: 94ff).

In der Literatur wird häufiger ein Dickey-Fuller (DF) bzw. Augmented Dickey-Fuller-Test (ADF; berücksichtigt allfällige Autokorrelation) aufgeführt. Ein Vorteil des PP-Tests gegenüber dem ADF ist jedoch, dass kein Lag für den Einbezug der Autokorrelation definiert werden muss (Zivot & Wang, 2006: 127). In Tab. 3 sind die erhaltenen Ergebnisse (P-Werte) aufgeführt (berechnet mit der R-Funktion ‚pp.test‘ des Pakets ‚tseries‘ mit der Alternativhypothese ‚stationary‘). Eine Eigenheit beider Tests (PP und ADF) ist, dass Stationarität die Alternativhypothese darstellt. Ein Verwerfen der Nullhypothese weist also nicht etwa auf einen kritischen (nicht-stationären) Zustand der Daten, sondern im Gegenteil auf die erwünschte Stationarität hin. In Tab. 3 wurden alle Werte, die über dem Signifikanzniveau von 5% liegen und daher mit grosser Wahrscheinlichkeit nicht-stationär sind, rot eingefärbt.

Kwiatkowski-Phillips-Schmidt-Shin-Test (KPSS-Test): Dieses Verfahren bietet sich als Ergänzung zum PP-Test an und wurde von Kwiatkowski, Phillips, Schmidt & Shin (1992) vorgeschlagen. Hier sind die Hypothesen vertauscht: Stationarität bildet die Nullhypothese. Der KPSS-Test beschreitet damit das übliche Verfahren, bei dem das Merkmal von Interesse als Alternativhypothese gewählt wird. Dieser Test ist damit aussagekräftiger als der PP- oder ADF-Test. Im Fall, dass diese beiden auf einen Einheitswurzel-Prozess hinweisen, der KPSS-Test aber Stationarität anzeigt, sollte dem letzteren mehr Gewicht beigemessen werden Pfaff (2008: 103). Der KPSS-Test bietet ferner die Möglichkeit, zwei verschiedene Nullhypothesen zu testen: Entweder Trend- oder (echte) Niveau-Stationarität. Damit kann genauer als beim PP-Test unterschieden werden – schliesslich ist auch die Trend-Stationarität nicht erwünscht. Die entsprechenden Hypothesen lauten:

H_0 : Stationarität (entweder Trend- oder Niveau-Stationarität)

H_1 : Einheitswurzel (nicht-stationärer Prozess)

Die entsprechende Teststatistik kann beispielsweise Pfaff (2008: 103ff) oder Rutkowska & Ptak (2012: 60) entnommen werden. Die berechneten P-Werte sind in Tab. 3 zusammengestellt (erstellt mit der R-Funktion ‚kpss.test‘ des Pakets ‚tseries‘ mit der Nullhypothese ‚Level‘ bzw. ‚Trend‘ und der

Standardeinstellung ‚short = TRUE‘). Rot markiert sind Werte, die das 5%-Signifikanzniveau unterschreiten und daher unter grossem Rückhalt als nicht-stationär bezeichnet werden können. Gelb markiert sind Werte, die ein Signifikanzniveau von 10% unterschreiten. Als Vorbehalt muss angefügt werden, dass Kwiatkowski et al. (1992: 170) für eine (kleine) Stichprobengrösse von 30, also für die Grössenordnung der hier verwendeten Daten, Verzerrungen der Teststatistik festgestellt haben.

Viele Einzugsgebiete zeigen beim KPSS-Test sowohl für die Nullhypothese einer Niveau-Stationarität als auch einer Trend-Stationarität keine signifikanten Werte. Dies mag zunächst verwirrend erscheinen, da eine Zeitreihe ja nicht zugleich Trend- und Niveau-stationär sein kann. Allerdings kann ein Trend auch eine sehr geringe Steigung aufweisen. Vermutlich zeigt ein solches Ergebnis daher an, dass die Streuung der Daten gross genug ist, dass auch ein schwacher Trend eingepasst werden könnte. Bei Einzugsgebieten, die einen signifikanten Wert für Trend-Stationarität und einen nicht-signifikanten für Niveau-Stationarität zeigen, konnte meist eine leichte U-förmige Krümmung in den Datenreihen beobachtet werden, die wahrscheinlich Auslöser für dieses Ergebnis ist.

Von besonderem Interesse sind jedoch diejenigen Fälle, in denen die Nullhypothese der Niveau-Stationarität verworfen werden muss (Einzugsgebiete 331, 337, 340, 341 und 355). Dies ist nämlich diejenige Art von Stationarität, die anzustreben ist, denn Trend-„Stationarität“ ist keine echte Stationarität. Unterstützt wird dieses Urteil durch die Ergebnisse des Mann-Kendall-Tests, der für diese Gebiete ebenfalls einen Trend anzeigt. Diese fünf Gebiete können also mit grosser Sicherheit als nicht-stationär bezeichnet werden. Der Mann-Kendall-Test zeigt ausserdem signifikante Werte für die Gebiete 357, 359, 369, 370, 371 und 2471.

Der PP-Test weist für die Gebiete 336, 337, 349, 359 und 378 auf Nicht-Stationarität hin. Der KPSS-Test ergibt ausser für die bereits erwähnten fünf Gebiete auch für das Einzugsgebiet 357 einen signifikanten Wert. Bei einem Signifikanzniveau von 10% führt dieser Test ausserdem bei den Gebieten 336, 369 und 370 zu einem Verwerfen der Nullhypothese.

Als zusammenfassendes Ergebnis der verschiedenen Tests können folgende Gebiete als kritisch bezeichnet werden: 331, 336, 337, 340, 341, (349), 355, 357, 359, 369, 370, 371, (378) und 2471 – insgesamt also ganze 14 Einzugsgebiete!

Es wäre kaum sinnvoll, all diese Gebiete aus der Analyse zu entfernen. Viel interessanter dürfte sein, diese Information schliesslich mit den erzielten Modellierungsergebnissen zu vergleichen.

Tab. 3: Ergebnisse (P-Werte) der verschiedenen durchgeführten Stationaritäts-Tests. Autokorrelierte Daten (siehe Spalte „Auto-Korr.“) wurden zusätzlich mit einem Prewhitening (PW) korrigiert.

Anmerkung zum KPSS-Test: R gibt keine Werte über 0.1 an, da solche Werte grundsätzlich als nicht-signifikant betrachtet werden. Gleichermassen werden bei keinem Test Werte unter 0.01 angezeigt; 0.01 steht also für hoch-signifikante Werte.

EZ	Auto-Korr.	Mann-Kendall-Test		PP-Test	KPSS-Test	
		Kein PW	PW		Trend	Level
331		0.01		0.01	0.1	0.04
332		0.09		0.018	0.1	0.1
334		0.65		0.01	0.1	0.1
336	X	0.08	0.103	0.098	0.025	0.071
337	X	0.00037	0.037	0.096	0.1	0.016
340	X	0.0035	0.053	0.023	0.1	0.02
341		0.017		0.01	0.1	0.01
342		0.393		0.01	0.1	0.1
344		0.747		0.01	0.1	0.1
345		0.131		0.01	0.1	0.1
348		0.94		0.011	0.1	0.1
349	X	0.549		0.088	0.1	0.1
351		0.79		0.01	0.1	0.1
352		0.747		0.02	0.053	0.1
353		0.747		0.028	0.1	0.1
354		0.678		0.01	0.1	0.1
355	X	0.005	0.028	0.012	0.1	0.013
357		0.017	0.052	0.011	0.044	0.024
358		0.2		0.01	0.1	0.1
359	X	0.048	0.155	0.166	0.047	0.1
360		0.866		0.017	0.1	0.1
361		0.508		0.01	0.1	0.1
362		0.223		0.01	0.1	0.1
364		0.53		0.011	0.1	0.1
367		0.736		0.01	0.1	0.1
368		0.595		0.038	0.012	0.1
369		0.037		0.01	0.1	0.0503
370		0.01		0.012	0.1	0.055
371		0.041		0.01	0.1	0.1
372		0.071		0.01	0.1	0.1
374		0.418		0.01	0.1	0.1
376	X	0.155	0.362	0.026	0.1	0.1
377		1		0.015	0.1	0.1
378		0.656		0.091	0.099	0.1
2450		0.262		0.01	0.1	0.1
2471	X	0.028	0.169	0.027	0.1	0.67

5. Methodik

5.1. Vorhersagbarkeit

Der Methodik lässt sich die Frage voranschicken, eine wie gute Vorhersagbarkeit bei Trockenheitsereignissen überhaupt zu erwarten ist. Trockenheit wird häufig als ein „schleichendes Phänomen“ bezeichnet (z.B. Mishra & Singh, 2010: 205). Im Gegensatz zu Hochwasser, das sich innerhalb von wenigen Stunden entwickeln kann, entsteht Trockenheit durch ein anhaltendes Niederschlagsdefizit, das sich in den Gewässern schliesslich als Niedrigwasser auswirkt. Aufgrund dieser langsamen Entwicklung sowie des Gedächtniseffekts verschiedener Speicher wie Bodenfeuchte und Grundwasservorkommen ist Niedrigwasser potentiell gut vorhersagbar (Zappa, Bernhard, Fundel, & Jörg-Hess, 2012); weil gleichzeitig aber verschiedenste Faktoren zu einer Trockenheit beitragen, ist sie ein komplexes Phänomen, was eine Prognose wiederum erschwert (Seneviratne et al., 2013). Dabei spielt allerdings der Vorhersagezeitraum eine entscheidende Rolle: Zappa (2011) berichtet von einer erfolgreichen Durchführung einer 5-Tages-Prognose an der Thur; die Prognose auf längere Zeiträume wurde anschliessend anvisiert. Mishra & Singh (2011) halten fest, dass es nahezu unmöglich sei, Trockenheiten Monate bis Jahre im Voraus zu prognostizieren (wie es oft erwünscht wäre). Sie verweisen dabei auf die raum-zeitliche Variabilität der hydro-meteorologischen Variablen (insbesondere Niederschlag) sowie auf die aktuell beobachtbare Intensivierung des hydrologischen Kreislaufes durch den Globalen Klimawandel, die Zukunfts-Prognosen aufgrund von Beobachtungen in der Vergangenheit grundsätzlich in Frage stellt.

Prinzipiell lässt sich Niedrigwasser dennoch leichter vorhersagen als Hochwasser: Fundel et al. (2013) konnten für einen Prognose-Zeitraum von einem Monat aussagekräftige Prognosen erstellen. Für den Kanton Aargau lassen diese Ergebnisse vermuten, dass eine kurz- bis mittelfristige Prognose gut durchführbar wäre. Kurzfristig ist gemäss WMO (2010) als ein Prognosezeitraum von 1 bis 3 Tagen, mittelfristig als bis zu 10 Tage zu verstehen.

Da diese Arbeit sich aber nicht mit Prognosedaten beschäftigt, werden die Herausforderungen der Niedrigwasservorhersage insbesondere darin bestehen, geeignete Gütemasse auszuwählen, um das Modell auf Niedrigwasser zu kalibrieren (siehe hierzu Kap. 5.5.2). Die Modellierung von Niedrigwasser hängt entscheidend davon ab, ob das Modell das Zusammenspiel zwischen Boden, Grundwasser und Abfluss korrekt widerspiegelt (Seibert, 1997).

5.2. Schwellenwert-Methode

Zur Auswertung von hydrologischer Trockenheit existieren verschiedene Ansätze. Die gesetzlich vorgeschriebenen Grenzwerte geben aber klar eine Schwellenwert-Methode vor. Die Ursprünge dieser Methode sind bei Yevjevich (1967) zu finden. Durch den Schwellenwert wird die Trockenperiode von Nicht-Trockenperioden unterschieden (Hisdal et al., in: Tallaksen & van Lanen, 2004: 164/165).

Von Fleig, Tallaksen, Hisdal, & Demuth (2006) werden verschiedene Trockenheitsmasse vorgeschlagen, von denen für diese Arbeit folgende ausgewählt wurden (siehe Abb. 10):

- Die **Dauer** (engl. *duration*) einer Trockenheit ist definiert durch einen Startzeitpunkt, zu dem der Abfluss den Grenzwert unterschreitet, und einen Endzeitpunkt, zu dem der Abfluss den Grenzwert wieder übersteigt.
- Das **Defizit** (engl. *severity* oder *deficit volume*) einer Trockenheit beschreibt das fehlende Wasservolumen. Grafisch gesprochen entspricht dies der Fläche zwischen der Abflusskurve und dem Grenzwert, wenn der Abfluss den Grenzwert unterschreitet.

- Die **Intensität** (engl. *intensity* oder *magnitude*) einer Trockenheit ist ein normiertes Mass und wird berechnet als das Defizit dividiert durch die Dauer. Damit werden zwei Trockenheits-Ereignisse, die gleich lang dauern, aber ein unterschiedlich grosses Defizit aufweisen (der Abfluss im einen Ereignis also tiefer abfällt) unterscheidbar. Andererseits weist die Intensität als Quotient von Defizit und Dauer aber auch denselben Wert für eine kurze und schwache Trockenheit wie für eine länger anhaltende Trockenheit mit grösserem Defizit auf, sofern Dauer und Defizit um denselben Faktor grösser sind.

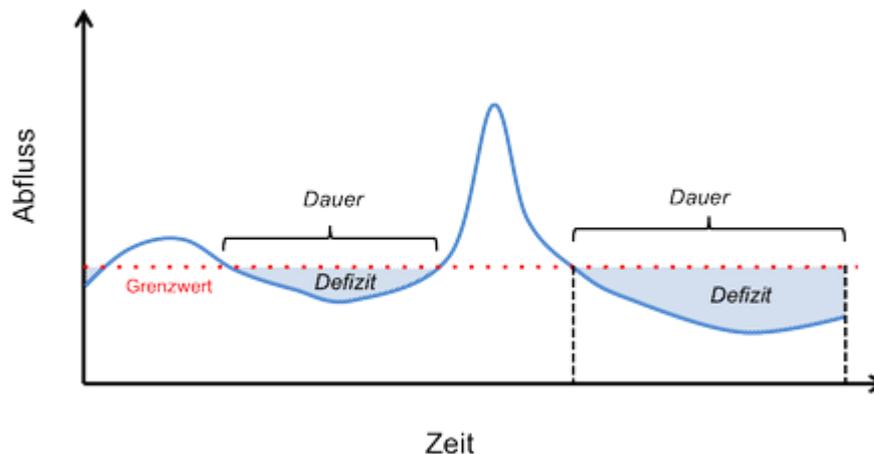


Abb. 10: Illustration der Begriffe Dauer und Defizit einer Trockenperiode. Quelle: http://www.drought.ch/Klima/Runoff/index_DE (Zugriff: 21.1.2014)

Diese drei Werte werden auch von Smakhtin (2001: 158) als die wichtigsten drei Niedrigwasser-Charakteristiken bezeichnet.

Im Verlaufe der Auswertungen zeigte sich, dass zwei weitere Masse wertvolle Informationen enthalten:

- Der **Zeitpunkt des Einsetzens** der Trockenheit (engl. *drought onset*), der ebenfalls durch den Schwellenwert definiert wird.
- Der **Minimalabfluss** jeder Auswertungsperiode.

Das Einsetzen der Trockenheit (im Folgenden mit „Onset“ bezeichnet) wurde zusätzlich ausgewählt, weil dieser Zeitpunkt in der Praxis für das Erlassen von allfälligen Nutzungsbeschränkungen entscheidend ist (Smakhtin, 2001: 158; Sene, 2010: 135). Der Minimalabfluss wurde als weiteres Mass zur Abschätzung des Realitätsgehalts des Modells ergänzt (siehe auch Fleig et al., 2006: 537). Diese beiden Kennwerte sind gemäss WMO (2008: 71) ebenfalls gebräuchliche Trockenheitsmasse.

In der Literatur werden weitere Charakteristiken zur Analyse vorgeschlagen, wie beispielsweise die zur Regenerierung benötigte Zeitspanne (z.B. Luo & Wood, 2007). Grundsätzlich wäre auch denkbar, den Schwellenwert variabel, beispielsweise den saisonalen Bedürfnissen entsprechend, zu wählen (z.B. Fleig et al., 2006). Die durch den Kanton bzw. das Gewässerschutzgesetz vorgegebenen Schwellenwerte sind jedoch konstant festgesetzt.

Für alle Kennwerte wurden relative bzw. absolute Abweichungen und deren Mittelwerte berechnet, um einen Eindruck der durchschnittlichen Modellgüte über das gesamte Untersuchungsgebiet hinweg zu erhalten. *False Alarms*, *Missed Events* und *No Event* werden in diese Mittelwertberechnung aber nicht mit einbezogen.

5.3. Kontingenztabelle

In der Praxis ist für das Vertrauen, das man einem Modell entgegen bringt, entscheidend, wie häufig das Modell Ereignisse nicht anzeigt (engl. *missed events*) oder fälschlicherweise Alarm schlägt (engl. *false alarms*). Zur Auswertung solcher Fehlvorhersagen wird in der Literatur häufig (und meistens angewandt auf Hochwasserprognosen) eine Kontingenztabelle aus Treffern (engl. *hits*), *False Alarms*, *Missed Events* und keinem Ereignis (engl. *no event*) verwendet, wie in

Tab. 4 zusammengestellt (z.B. Sene, 2010; Cloke & Pappenberger, 2009; Addor et al., 2011; Demirel, Booi, & Hoekstra, 2013; Wilks, 2006: 260ff). Aus praktischen Gründen werden fortan diese englischen Begriffe verwendet.

Von der Kontingenztabelle lassen sich verschiedene Masse für die Zuverlässigkeit einer Vorhersage ableiten. Am häufigsten verwendet sind die POD (4) („Probability Of Detection“, Wahrscheinlichkeit, dass ein Ereignis erkannt wird) und die FAR (5) („False Alarm Ratio“, Wahrscheinlichkeit, dass ein falscher Alarm geschlagen wird). Weil aber eigentlich die *Missed Events* die kritische Grösse darstellen, wird hier als weiterer Wert die FOM (6) („Frequency Of Misses“, Wahrscheinlichkeit, dass ein Ereignis fälschlicherweise nicht vorhergesagt wird) gemäss Doswell III, Davies-Jones, & Keller (1990: 578) ergänzt. Anzumerken ist, dass die FAR-Berechnung das korrekte Nicht-Anzeigen von Ereignissen (Kombination *No Event* simuliert mit *No Event* beobachtet) nicht mit einbezieht.

2003 Q347		Obs	
		Event	No Event
Sim	Event	A	B
	No Event	C	D

Tab. 4: Prinzip der Kontingenztabelle.

$$(4) \quad \text{POD} = \frac{A}{A+C}$$

$$(5) \quad \text{FAR} = \frac{B}{A+B}$$

$$(6) \quad \text{FOM} = \frac{C}{A+C}$$

Diese Arbeit hält sich bewusst an diese einfachen und verständlichen Masse, obwohl in der Literatur eine Vielzahl weiterer möglicher Masse zur Beurteilung der Modell-Fähigkeiten vorgeschlagen werden. Doch verschiedene Autoren (z.B. Wilks, 2006: 268; Cloke & Pappenberger, 2009: 617) betonen, dass das bestgeeignete Mass zur Bewertung von Vorhersagen nicht existiert – dazu bietet die Kontingenztabelle zu viele Kombinationsmöglichkeiten bzw. dieselbe Vorhersage kann in verschiedenen Sichtweisen unterschiedlich gut abschneiden. Aus Sicht der Autorin enthalten die drei Masse POD, FAR und FOM die hier essentiellen Informationen und bleiben dabei gut interpretierbar.

5.4. Auswertungsperioden

5.4.1. Sommertrockenheit 2003

Im Sommer 2003 erlebte der Europäische Kontinent extrem hohe Temperaturen, kombiniert mit unterdurchschnittlichen Niederschlägen. Ursache dafür war ein Hochdruckgebiet, das über den westlichen Landmassen Europas stationär blieb, heisse Luft aus Nordafrika nach Europa trug und gleichzeitig Tiefdruckgebiete, die Niederschlag hätten eintragen können, über dem Atlantik zurückhielt. Dabei traten bereits im Frühjahr ungewöhnlich häufig Hochdruck-Lagen auf (Wilhite, in: Tallaksen & van Lanen, 2004: 41).

In der Schweiz brachen die Temperaturen wie in weiten Gebieten Mitteleuropas alle bisherigen Rekorde: Das langjährige Mittel für die Monate Juni, Juli und August wurde um zirka 3°C

überschritten, was statistisch gesehen einem Ausreisser gleichkommt (Schär et al., 2004). Die Monate Februar bis September waren zudem extrem trocken: Verbreitet fielen nur gerade etwa die Hälfte der sonst üblichen Niederschlagsmengen (BAFU, 2012b).

Aufgrund dieser Lage verzeichneten verschiedene Flüsse in Europa extreme Niedrigstände. In der Schweiz wurden stark verminderte Abflüsse bei Gewässern mit unvergletschertem Einzugsgebiet festgestellt, hingegen hohe Wasserstände bei Gewässern mit hochalpinen Einzugsgebieten (BAFU, 2004).

In der Typologie von Van Loon & Van Lanen (2012) liegt hier eine „klassische Trockenheit wegen Niederschlagsdefizit“ (engl. *classical rainfall deficit drought*) vor. Weil das Niederschlagsdefizit sich aber bereits im normalerweise feuchten Frühling wegen der überdurchschnittlich häufig auftretenden Hochdruckgebiete aufzubauen begann, klingt das Ereignis aber nahezu an ein „Regenzeit-zu-Trockenzeit“-Ereignis (engl. *wet-to-dry-season drought*) an – obwohl die Schweiz als Land der gemässigten Klimate natürlich keine eigentliche Trockenzeit kennt. Die Betrachtung der Niederschlags- und Abflusswerte zeigte, dass auf Oktober starker Niederschlag einsetzte und die Abflüsse daraufhin anstiegen. Daher wurde die Zeitspanne für die Auswertung vom 1. Juni bis zum 30. September festgelegt.

5.4.2. Frühjahrstrockenheit 2011

Im Jahr 2011 traten sowohl im Frühjahr als auch im Herbst lange Perioden mit unterdurchschnittlich geringen Niederschlagsmengen auf. Die geringsten Abflusswerte bzw. das grösste Defizit trat allerdings im Mai auf, weshalb die vorliegende Arbeit den Fokus nur auf das Frühjahrs-Ereignis legt. Die Abflüsse fielen im Frühjahr auf besonders tiefe Pegel, weil bereits im Winter wenig Schnee gefallen war und daher der Abfluss aus der Schneeschmelze besonders gering ausfiel. Dank dem schliesslich sehr nassen Einsetzen des Frühsommers erholten sich die Gewässer bereits bis im Juli vom Frühjahrs-Defizit (Zappa, 2011).

Seit Jahresbeginn bis anfangs Mai fielen im Mittelland nur gerade 25 bis 40% des üblichen Niederschlags (Schlegel, Kunz, & Gerstgasser, 2011). Hinzu kamen ein überdurchschnittlich warmer März und bereits frühlommerliche Temperaturen im April (BAFU, 2011a). Obwohl die Gewässer mit Normalstand ins Jahr starteten, erreichten sie daher anfangs Mai teils so niedrige Wasserstände wie noch nie in dieser Jahreszeit, insbesondere auch im Mittelland. Die Auswirkungen zeigten sich nicht nur an den Oberflächengewässern, sondern auch die Grundwasserstände im Mittelland befanden sich auf sehr tiefem Niveau (BAFU, 2011b).

In der Typologie von Van Loon & Van Lanen (2012) kann dieses Ereignis als „kombinierte Trockenheit“ (engl. *composite drought*) bezeichnet werden. Hier folgt nämlich auf eine „Trockenheit wegen warmer Schnee-Saison“ (engl. *warm snow season drought*) eine „klassische Trockenheit wegen Niederschlagsdefizit“ (siehe Kap. 5.4.1). Die Trockenheit begann also mit einem verfrühten Einsetzen der Schneeschmelze, welche wegen des geringen Schneevolumens schnell abflaute, und verstärkte sich mit dem anhaltend defizitären Niederschlag.

Basierend auf diesen Informationen wurden zur Festlegung des Auswertungszeitraumes für das Jahr 2011 die Niederschlags- und Abflusswerte mehrerer Einzugsgebiete betrachtet. Allgemein war eine leichte Pegelabnahme ab März zu beobachten. Ab Ende Mai traten zwar vermehrt Niederschläge auf, dennoch waren die Abflüsse auch im Juni oft noch niedrig. Der Auswertungs-Zeitraum wurde daher vom 1. März bis zum 30. Juni festgesetzt.



Abb. 11: Die Aare bei Brugg am 8. April 2011. Grosse Flüsse wie die Aare werden in dieser Arbeit zwar nicht berücksichtigt, das Ausmass der Trockenheit wird an ihnen aber sehr eindrücklich sichtbar.

Quelle: W.Lehmann, ALG Wasserbau, BVU, Kanton Aargau.

5.4.3. Vergleich der beiden Ereignisse

Stellt man die Ereignisse 2003 und 2011 einander gegenüber, so ist zu betonen, dass sie nicht in derselben Jahreszeit auftraten. Die Zuweisung zur Typologie von Van Loon & Van Lanen (2012) zeigt, dass den beiden Ereignissen zumindest bis zu einem gewissen Mass unterschiedliche Prozesse zugrunde liegen. Gemäss Smakhtin (2001) treten Niedrigwasserstände normalerweise immer wieder in derselben Saison auf. Für die Schweiz besonders unüblich ist dabei die Trockenheit im Frühjahr 2011, da aufgrund der Schneeschmelze zu dieser Jahreszeit normalerweise Hoch- anstatt Niedrigwasser zu erwarten ist.

Im Frühjahr 2011 bestanden bereits Befürchtungen, dass ein solcher Frühling einen Sommer wie 2003 nach sich ziehen könnte (BAFU, 2011b). Dies trat zu aller Erleichterung nicht ein; grundsätzlich wäre ein solches Ereignis aber nicht ausgeschlossen.

5.5. Modellkalibrierung und Gütekriterien

Das Ziel jeder Modellierung ist, eine möglichst perfekte Nachahmung des natürlichen Verhaltens einer bestimmten Sache zu erzielen. Modellierungen sind aber immer mit Unsicherheiten aus verschiedenen Quellen behaftet:

“As with many other types of models, hydrological forecasting models are subject to many sources of uncertainty, including uncertainties in the input data, model parameters, initial conditions, and boundary conditions. The model structure may also be inappropriate or less than ideal for the situation under consideration.”

(Sene, 2010: 104)

„Wie verschiedene andere Arten von Modellen sind auch hydrologische Vorhersage-Modelle vielen Quellen von Unsicherheit unterworfen, einschliesslich Unsicherheiten der Inputdaten, der Modell-Parameter, der Ausgangszustände und der Randbedingungen. Auch die Modellstruktur könnte unpassend oder zumindest nicht optimal für die betrachtete Situation sein.“

Ein Modell muss daher immer validiert werden (Krause, Boyle & Bäse, 2005). Validierung bedeutet, das Modellergebnis anhand bestimmter Kriterien auszuwerten. Daher plädieren Beven & Young (2013) dafür, diese Methode als bedingte Validierung (engl. *conditional validation*) zu bezeichnen. Erfüllt ein Modell die gestellten Bedingungen in zufriedenstellendem Masse, kann es als für den Zweck geeignet (engl. *fit for purpose*) bezeichnet werden.

Dazu müssen allerdings die Bedingungen, die das Modell erfüllen soll und anhand derer die Modellgüte bemessen wird, definiert werden. Die Gütekriterien, die in dieser Arbeit verwendet werden, sind im Kapitel 5.5.2 genauer erläutert.

Zur Evaluierung der Vorhersagekraft des Modells wird ein „Split-Sample-Test“ (Klemes, 1986) verwendet. Dabei werden die Zeitreihen der beobachteten Daten in zwei Perioden aufgeteilt: Die eine wird zur Kalibrierung des Modells verwendet, die andere zur Simulation eines Zeitraumes ausserhalb der Kalibrierungsperiode. Erstere umfasst üblicherweise eine möglichst lange Zeitreihe; in dieser Arbeit wurden die gesamten verfügbaren Zeitreihen mit Ausnahme der Auswertungsperioden gemäss Kap. 5.4 verwendet, die normalerweise 30 bis 35 Jahre und bei einer Station 21 Jahre umfassten. Ziel dieser Methode ist, die Fähigkeit des Modells zu testen, das Abflussverhalten auch ausserhalb der Kalibrierungsperiode erfolgreich zu simulieren.

Der erste Schritt zur Vorhersage besteht in der Kalibrierung der Modell-Parameter. Aufgrund verschiedener Limitierungen in der Realitätsnähe eines Modells ist dieser Schritt für jede Form von Modellen – auch für komplexe, physikalisch basierte – unumgänglich (Beven, 2012: 231).

Kalibrierung zielt üblicherweise darauf ab, das optimale Set an Parametern zu finden. In der Realität ist man hierbei mit dem bereits erwähnten Problem der „Equifinality“ (Beven, 2006) konfrontiert: Verschiedene Sets von Parametern können ein gleich gutes Modellergebnis erzeugen. Parameter sind daher immer mit einer gewissen Unsicherheit behaftet; diese könnte reduziert werden, indem man zusätzliche Informationen über das hydrologische System, beispielsweise Daten aus Tracer-Studien oder zur Schneebedeckung, Bodenfeuchte und Grundwasserständen, bezieht (z.B. Seibert, 1997; Seibert & McDonnell, 2002; Hartmann et al., 2013; Son & Sivapalan, 2007). Aus diesem Grund werden im zweiten Schritt dieser Arbeit Grundwasser-Messdaten in die Modellierung mit einbezogen.

Kalibrierungen wurden ursprünglich von Hand durchgeführt. Diese Methode ist zwar sehr subjektiv, ermöglicht aber dem Modellierer, ein Gefühl für das Verhalten seines Modells zu entwickeln (z.B. Bergström, 1976). Vor allem dank der massiv vergrösserten Leistungsfähigkeit der Computer werden heute fast ausschliesslich automatisierte Methoden zur Optimierung der Parameter verwendet.

In HBV-light sind zwei Methoden zur automatisierten Parameter-Kalibrierung implementiert: Die Monte Carlo Simulation oder die GAP-Optimierung. In der vorliegenden Arbeit wurde die zweite Variante verwendet.

5.5.1. Genetic Algorithm mit Powell-Optimierung (GAP)

Auf der Suche nach dem globalen Optimum unter den möglichen Parameterkombinationen bilden lokale Optima ein Problem für viele Such-Algorithmen: Einmal ein solches erreicht, sind klassische „Hill-Climbing“-Algorithmen darin gefangen. Um dennoch das globale Optimum zu finden, sind sehr viele Versuche aus verschiedenen Ausgangspositionen nötig. Ähnlich funktioniert die häufig verwendete Monte Carlo-Simulation, die einfach eine sehr grosse Anzahl von Parameterkombinationen testet und die besten Ergebnisse zusammenstellt. Bereits bei der Parameteranzahl von HBV bedeutet dies aber einen beträchtlichen Rechenaufwand. (Beven, 2012: 233ff)

Die in HBV-light implementierte Variante *Genetic Algorithm* mit *Powell Optimization* (GAP) zielt darauf ab, das globale Optimum bei verringertem Rechenaufwand zu finden. *Genetic Algorithms* („genetische Algorithmen“) basieren auf einem Prinzip, das die biologische Evolution zu imitieren versucht. Erstmals beschrieben wurde die Idee von Holland (1975). Zu Beginn wird eine Population von verschiedenen Individuen (d.h. bestimmten Parameterkombinationen) zufällig ausgewählt, die sich über Generationen hinweg durch Mutation und Rekombination entwickeln dürfen. Dabei erhalten Parameter-Sets, die eine höhere Güte erzielen, auch eine höhere Chance, neue Sets zu erzeugen. In HBV-light wurde der Genetische Algorithmus schliesslich mit einer Powell-Optimierung ("Powell's quadratically convergent method", siehe Press, Teukolsky, Vetterling, & Flannery, 1992) zur lokalen Feinabstimmung der Parameterwerte kombiniert. Eine detailliertere Beschreibung der beiden Methoden findet sich in Seibert (2000).

5.5.2. Zielfunktionen

Für eine Kalibrierung müssen im Voraus die Kriterien definiert werden, anhand derer die Modellgüte bestimmt werden soll. Dazu werden bei einer automatisierten Kalibrierung bestimmte Zielfunktionen (engl. *objective functions*) verwendet, die ein objektives Mass für die Modellgüte bzw. die Übereinstimmung der Simulation mit den Messwerten berechnet. Wie Beven (2012: 241) betont, existiert unglücklicherweise kein allgemeingültiges und universal anwendbares Gütemass. Die Wahl dieser Gütemasse muss daher in Hinblick auf die bezweckte Anwendung erfolgen.

Die in der hydrologischen Modellierung meist verwendete Zielfunktion ist das durch Nash & Sutcliffe (1970) beschriebene Effizienz-Kriterium R^2 (7) (in HBV-light als *Reff* bezeichnet, sonst häufig auch *NSE* („Nash-Sutcliffe-Efficiency“) genannt), das auf der Fehler-Varianz basiert.

$$(7) \quad Reff = 1 - \frac{\sum(Q_{obs} - Q_{sim})^2}{\sum(Q_{obs} - \overline{Q_{obs}})^2}$$

Q_{sim} steht hierbei für den simulierten, Q_{obs} für den beobachteten Abfluss und $\overline{Q_{obs}}$ für den Mittelwert des beobachteten Abflusses. Bei einer perfekten Übereinstimmung von beobachtetem und simuliertem Abfluss erreicht der *Reff* einen Wert von 1.

Interpretiert kann der *Reff*-Wert wie folgt werden: Im Nenner wird die Abweichung vom Mittelwert zu einem bestimmten Zeitpunkt berechnet. Der Mittelwert kann als ein Modell ohne Vorwissen über das System interpretiert werden. Der *Reff*-Wert bezeichnet also die Güte eines Modells im Vergleich zu diesem Modell ohne Vorwissen, bzw. zu einer Simulation durch den Mittelwert. Ein *Reff*-Wert von 0 besagt, dass das Modell keine bessere Aussagekraft als eine Beschreibung durch den Mittelwert hat, ein Wert unter 0 zeigt eine schlechtere Simulation an; die Werteskala ist dabei nach unten offen (Beven, 2012: 239ff; Schaefli & Gupta, 2007: 2076).

Bereits Bergström (1976: 14) stellte allerdings fest, dass der *Reff*-Wert durch den Vergleich mit dem Abflussmittelwert stark von der Saisonalität eines Einzugsgebietes abhängig ist: Weist der Abfluss starke saisonale Schwankungen wie etwa bei nivalen Einzugsgebieten auf, ist es leichter, einen hohen *Reff* zu erzielen, weil die Ausgangs-Fehlervarianz bereits entsprechend gross ist. Zuverlässigere Aussagen lassen sich aus dem *Reff*-Wert daher für Einzugsgebiete mit geringer Saisonalität bzw. pluvialem Regime ziehen (Schaefli & Gupta, 2007; Krause et al., 2005).

Obwohl ein häufig verwendetes Mass, entspricht der *Reff* jedoch nicht dem hier angestrebten Ziel: Da er auf dem quadrierten Modellfehler basiert, werden Abflussspitzen stärker gewichtet als tiefe Abflüsse. Zur Auswertung der Modellgüte bei Niedrigwasser ist er daher nicht geeignet (Pushpalatha, Perrin, Moine, & Andréassian, 2012; Krause et al., 2005) – andere Gütekriterien sind also gefragt.

Pushpalatha et al. (2012) und Krause et al. (2005) besprechen verschiedene Masse, die für diesen Zweck vorgeschlagen wurden. Aus dieser Bandbreite sowie mit Blick auf die in HBV-light implementierten Funktionen wurden folgende drei Zielfunktionen ausgewählt:

- **LogReff (8):** Wie verschiedene andere zur Niedrigwasser-Auswertung vorgeschlagene Masse ist auch dieser Wert eine Abwandlung des Reff. LogReff (Oudin, Andréassian, Mathevet, Perrin, & Michel, 2006) verwendet die logarithmierten Abflusswerte, schwächt damit den Einfluss der Abflussspitzen und legt mehr Gewicht auf niedrigere Abflüsse.

$$(8) \quad \text{LogReff} = 1 - \frac{\sum(\ln Q_{obs} - \ln Q_{sim})^2}{\sum(\ln Q_{obs} - \ln Q_{obs})^2}$$

Pushpalatha et al. (2012) zeigen auf, dass der LogReff das Hauptaugenmerk auf die mittleren und niedrigen Abflüsse legt, aber dennoch auch auf hohe Abflüsse sensibel ist. Dadurch wird er zu einem guten Mass, um eine Anpassung der Simulation an die gesamte Variationsbreite der Abflusskurve zu erreichen. Ihre Untersuchungen belegen ausserdem, dass das Logarithmieren der Abflusswerte den Informationsgehalt grundsätzlich verändert, wodurch Reff und LogReff sich unterschiedlich verhalten. Für ephemere Gewässer müsste im Übrigen berücksichtigt werden, dass die Berechnung des LogReff mathematisch nicht möglich ist, wenn der Abfluss auf null sinkt; in HBV ist die Funktion allerdings so implementiert, dass Abflusswerte von 0 als fehlende Daten behandelt und damit ignoriert werden. Ein Trockenfallen von Gewässern bildet somit kein Problem.

Ergänzend soll hier erwähnt sein, dass weitere Variationen des Reff existieren, die noch stärker auf Niedrigwasser fokussieren als der LogReff: Pushpalatha et al. (2012) empfehlen, den Reff der invertierten Abflusswerte zu verwenden; Krause et al. (2005) fanden mit einer Form des Reff relativ zu den beobachteten Werten ein Gütemass, das keine Reaktion mehr auf hohe Abflüsse zeigt. Beide empfohlenen Masse kommen für die vorliegende Arbeit aber aus dem einfachen Grund nicht in Frage, dass sie nicht in HBV-light implementiert sind. Zusätzlich kann argumentiert werden, dass das Ziel dieser Modellierungen zwar in erster Linie darin besteht, Niedrigwasser möglichst gut vorhersagen zu können, dies aber nicht auf Kosten eines ausserhalb des Niedrigwasserbereichs unzuverlässigen hydrologischen Modells erreicht werden soll. Auch wenn der Fokus auf Niedrigwasser liegt, sollten die übrigen Bereiche der Abflusskurve nicht komplett vernachlässigt werden – ansonsten muss das Modell aus hydrologischer Sicht als wertlos bezeichnet werden. In diesem Sinne ist es sinnvoll, mit dem LogReff ein Mass zu verwenden, das der ganzen Bandbreite der Abflusskurve ein gewisses Gewicht beimisst.

- **MARE (9):** Der „Mittlere absolute Relativfehler“ (engl. *Mean Absolute Relative Error*) (Dawson et al., 2007) berechnet das Mittel aus dem absoluten Fehler, der durch den beobachteten Wert relativiert ist:

$$(9) \quad \text{MARE} = 1 - \frac{1}{n} \sum \frac{|Q_{ob} - Q_{sim}|}{Q_{obs}}$$

Dabei steht n für die Anzahl Tage. Da der MARE im Gegensatz zum Reff nicht mit quadratischen Fehlern arbeitet, ist der MARE stärker auf Fehler im Bereich der niedrigen Abflüsse sensibel.

Im Gegensatz zur Anwendung bei Riad, Mania, Bouchaou, & Najjar (2004) wird der MARE in HBV-light so berechnet, dass die Werte identisch mit denjenigen des Reff sind: So steht auch hier der Wert 1 für eine perfekte Übereinstimmung zwischen Messung und Simulation und die Skala ist nach unten offen.

Weil zur Berechnung des MARE durch den Beobachtungswert dividiert wird, ist auch hier die Berechnung mathematisch nicht möglich, wenn der beobachtete Abfluss auf 0 sinkt. Wie der LogReff ist die Funktion in HBV allerdings so implementiert, dass Messwerte von 0 ignoriert werden; damit ist die Berechnung des MARE auch für zwischenzeitlich trockenfallende Gewässer möglich.

- Als drittes Mass sollte schliesslich eines mit einbezogen werden, das den Volumen-Fehler des Modells bewertet. Damit wird sichergestellt, dass die Wasserbilanz des Modells mit derjenigen der Messwerte übereinstimmt. Für diesen Zweck wurden zwei Masse in Betracht gezogen: Der relative Volumenfehler (Dawson et al., 2007), in HBV **VolumeError** (10) genannt, berechnet den totalen Fehler relativ zu den gemessenen Werten und bezeichnet daher die Abweichung des Volumens bzw. der Gesamtbilanz zwischen Modell und Messungen. Im Gegensatz zur Formulierung bei Dawson et al. (2007), wo eine perfekte Übereinstimmung durch den Wert 0 bezeichnet wird, ist der relative Volumenfehler in HBV-light wie folgt implementiert:

$$(10) \quad \text{VolumeError} = 1 - \frac{|\sum(Q_{obs} - Q_{sim})|}{\sum(Q_{obs})}$$

Eine perfekte Übereinstimmung wird also durch den Wert 1 wiedergegeben; auch hier ist die Skala nach unten offen. Ein Wert von 1 besagt allerdings nicht, dass das Modell die Abflusskurve korrekt simuliert, sondern nur, dass die Gesamtbilanz stimmig ist.

Eine alternative Möglichkeit zur Einbeziehung des Volumenfehlers bietet das Lindström-Mass (Lindström et al., 1997), in HBV **LindströmMeasure** (11) genannt, das ebenfalls in HBV-light zur Verfügung steht. Dieses bildet eine Kombination aus dem Reff und dem relativen Volumenfehler (in HBV-light mit dem Faktor 0.1 gewichtet).

$$(11) \quad \text{LindströmMeasure} = \text{Reff} - 0.1 \frac{|\sum(Q_{obs} - Q_{sim})|}{\sum(Q_{obs})}$$

Lindström et al. (1997) kreierte dieses Mass, um den Volumenfehler, der eine Kalibrierung nach Reff üblicherweise mit sich bringt, zu eliminieren und fanden so ein hilfreiches Mass, um einen nahezu ebenso guten Reff zu erzielen und gleichzeitig den Volumenfehler zu minimieren. Auch hier steht der Wert 1 für eine perfekte Übereinstimmung, nach unten ist die Skala offen.

Um sich für eines der beiden Masse zu entscheiden, wurden an drei Test-Einzugsgebieten unterschiedlicher Grösse (Nr. 2471, 340 und 367) Kalibrierungen in verschiedenen Parameter-Kombinationen durchgeführt. Dabei zeigte sich, dass eine Kombination mit dem VolumeError zwar eine etwas schlechtere Anpassung der Abflussspitzen mit sich brachte, gleichzeitig aber der Volumenfehler des Modells deutlich reduziert und die allgemeine Abflussdynamik („Zähnung“ der Kurve) dennoch besser als mit dem LindströmMeasure dargestellt wurde. Zudem wurden die Niveaus von Niedrigwasserständen in gewissen Perioden besser getroffen.

Diese Beobachtungen führten schliesslich zu einer Kombination der Masse LogReff, MARE und VolumeError, die im Weiteren als „kombiniertes Gütemass“ (von HBV *weighted objective function* genannt) bezeichnet wird. Um diese drei Zielfunktionen zu einem Mass verrechnen zu können, musste eine passende Gewichtung gefunden werden. Die schlussendlich verwendete Kombination ist in Tab. 5 zusammengestellt.

Zielfunktion	MARE	LogReff	VolumeError
Gewichtung	0.5	0.4	0.1

Tab. 5: Zur Modellkalibrierung verwendete Zielfunktionen und ihre Gewichtung.

Zur Bestimmung dieser Gewichtung waren folgende (visuelle) Beobachtungen an den bereits erwähnten Testgebieten ausschlaggebend:

- Sowohl MARE wie LogReff resultierten in einem relativ guten Abbild der gesamten Abflussdynamik.
- Eine höhere Gewichtung von MARE erzielte eine bessere Darstellung der Dynamik (bessere „Zähnung“ der Abflusskurve).
- MARE brachte die Dynamik im Niedrigwasserbereich besser zur Geltung als LogReff. Hingegen traf LogReff teilweise die durchschnittlichen Niveaus von Niedrigwasser besser.
- Der Volumenfehler wird nur schwach gewichtet, weil er die allgemeine Abflussdynamik sehr schlecht wiedergibt. Sein Einbezug ist sinnvoll zur Korrektur der Gesamtbilanz, zur korrekten Kurvenform selbst trägt er aber nur wenig bei.

Die gewählte Kombination erzielte im Vergleich zu anderen Gewichtungen eine bessere Wiedergabe von Niedrigwasser-Niveaus sowie der gesamten Abflussdynamik bei einem geringen Volumenfehler. Eine Kombination verschiedener Gütemasse zur Kalibrierung wird von Abebe, Ogden & Pradhan (2010) grundsätzlich empfohlen, um ein Modell nicht einseitig auf einen bestimmten Aspekt zu eichen. Die hier verwendete Kombination legt bewusst mehr Gewicht auf den Niedrigwasser-Bereich. In der Literatur sind keine objektiven Angaben zu finden, ab welchen Werten eine Modellierung als akzeptabel angesehen werden kann, insbesondere nicht für das hier neu kombinierte Gütemass. Für diese Arbeit wurden daher Werte über 0.7 als „gut“ und Werte über 0.8 als „sehr gut“ definiert.

5.5.3. Modellparameter

Anhand erster Test-Kalibrierungen wurde die Spannweite der Parameterwerte so festgelegt, wie in Tab. 6 aufgeführt und für die meisten Gebiete auch so verwendet. Wenn ein berechneter Parameter allerdings einen Grenzwert annimmt, wird der erlaubte Wertebereich bei einer Kalibrierung normalerweise schrittweise vergrössert, sofern kein logischer bzw. physikalisch basierter Grund vorliegt, warum der Bereich nicht erweitert werden darf.

Auch in den hier durchgeführten Kalibrierungen wurden daher die ursprünglich festgesetzten Spannweiten verschiedentlich erweitert; Maximalwerte sind in Tab. 6 in Klammern vermerkt. Für den Gradtagfaktor CFMAX wurde die Obergrenze in drei Fällen sehr stark auf 25 bzw. 20 mm/(d°C) erhöht. Dieser Wert liegt weit über den üblicherweise in hydrologischen Modellierungen verwendeten Werten (z.B. legt Seibert (1997: 11) den Maximalwert auf 10 mm/(d°C) fest). Ausserhalb der Niederschlags-Abflussmodellierung sind in der Literatur jedoch sehr unterschiedliche Werte für den Gradtagfaktor zu finden, wobei er im späten Frühjahr deutlich grösser als im Winter ist. Kirnbauer (1993, zitiert in Asztalos, 2004: 16) nennt gar eine Schwankungsbreite von 0.6 bis 47 mm/(d°C). Bestimmt wird der Gradtagfaktor üblicherweise aus historischen Messungen von Temperatur und Ablation. Solche Daten liegen für die hier untersuchten Gebiete leider nicht vor.

Die Kalibrierung wurde mit jeweils 300 Parameter-Sets durchgeführt, die vom *Genetic Algorithm* zufällig ausgewählt und über 50'000 Durchgänge sowie 1'000 Powell-Durchläufe optimiert werden. Die übrigen verwendeten GAP Einstellungen („Reproduction Settings“) entsprechen den Standard-Einstellungen und sind der Abb. 12 zu entnehmen. Die schliesslich für die Vorhersagen verwendeten Parameter sind Tab. 16 im Anhang zu entnehmen.

Tab. 6: Zur Kalibrierung verwendete Parameter-Grenzen. In Klammern sind einzeln verwendete Minimal- bzw. Maximalwerte aufgeführt.

Snow Routine	Min	Max	Soil Routine	Min	Max	GW Routine	Min	Max	Routing Routine	Min	Max
TT [°C]	-2	2 (4)	FC [mm]	50	800 (1000)	PERC [mm/d]	0	7	MAXBAS [d]	1	7
CFMAX [mm/(d°C)]	0.5	12 (25)	LP [-]	0.2 (0)	1	UZL [mm]	0	100	Weitere	Min	Max
SFCF [-]	0.2 (0)	1.5	BETA [-]	0.5	8 (20)	K0 [1/d]	0.01	0.8 (0.999)			
CFR [-]	0	1	Cet [1/°C]	0	0.5	K1 [1/d]	0.01	0.6	PCALT [%/100m]	0	25 (30)
CWH [-]	0	0.5 (0.8)				K2 [1/d]	10 ⁻⁶	0.15			

Anmerkung: Pelev und Telev wurden für jedes Einzugsgebiet separat festgelegt.

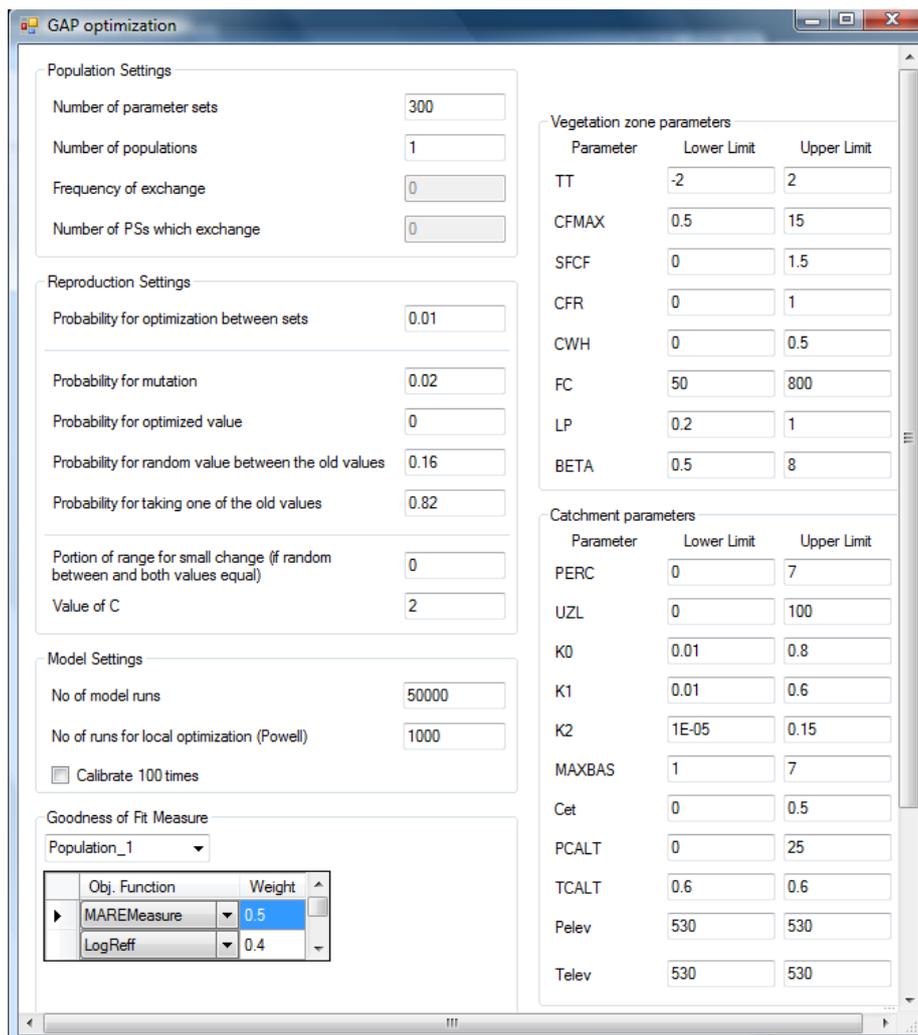


Abb. 12: GAP-Einstellungen am Beispiel des Einzugsgebietes 331.

Parameter, die bereits bei der Justierung der Parametergrenzen eine starke Auswirkung auf die Modellgüte zeigten (siehe auch Kap. 6.2) waren Pelev und Telev in Verbindung mit Pcalt, CFMAX und SFCF. Pelev bzw. Telev stehen eigentlich für die Höhe der meteorologischen Messstationen, wurden aber grundsätzlich als mittlere Höhe des Einzugsgebietes festgelegt. Die Erklärung für diese Wahl liegt in der Tatsache, dass kaum nachvollzogen werden kann, welche Messstation zu welchem Zeitpunkt einen wie starken Einfluss auf die interpolierten Werte eines bestimmten Einzugsgebietes ausgeübt hat. Pelev und Telev lassen sich daher aufgrund der umliegenden Messstationen nicht präzise bestimmen. Die mittlere Einzugsgebietshöhe zu verwenden erscheint daher als eine einfache und sinnvolle Alternative.

Insbesondere für Gebiete im östlichen Jura wurden aber grössere Abweichungen zwischen mittlerer Einzugsgebietshöhe und mittlerer Höhe der umliegenden meteorologischen Messstationen festgestellt, da sich alle Messstationen in den umgebenden Tallagen und nicht auf den Jurahöhen befinden. Vor allem für Pcalt und CFMAX resultierten daraus unrealistische Werte. Durch eine Anpassung von Pelev und Telev an die Höhenlage der meteorologischen Messstationen konnten sinnvolle Werte in der Kalibrierung dieser Parameter erreicht werden.

5.6. Parameter-Sensitivität

Klassischerweise werden Sensitivitätsanalysen in Studien der hydrologischen Modellierung durchgeführt, indem eine bestimmte Anzahl an Monte-Carlo-Simulationen berechnet und ausgewertet werden (Solmatine & Wagener, 2011: 451). Auch die in der hydrologischen Modellierung häufig verwendete GLUE-Methode (Beven & Binley, 1992) zur Quantifizierung der Parameter-Unsicherheit basiert auf Monte-Carlo-Simulationen. Häufig werden die Ergebnisse der Monte-Carlo-Simulationen in Streudiagrammen dargestellt (z.B. Seibert, 1997; Son & Sivapalan, 2007). Dieses Vorgehen kann hier aus zwei Gründen nicht angewandt werden: Zum einen wurde zur Modellkalibrierung nicht die Monte-Carlo-Simulation, sondern die GAP verwendet, die im Gegensatz zur Monte-Carlo-Variante als Ergebnis einer Kalibrierung nur ein einzelnes, bestes Resultat liefert. Eine Monte-Carlo-Kalibrierung ergibt hingegen verschiedene ähnlich gute Parameter-Kombinationen, die direkt für eine Sensitivitäts-Analyse verwendet werden können. Zum anderen können in einem Streudiagramm nur die Ergebnisse eines einzelnen Einzugsgebietes dargestellt werden. Für diese Arbeit wurde hingegen nach einer Variante gesucht, die einen Überblick über die Parameter-Streuung über alle Einzugsgebiete hinweg verschafft.

Um diese Vergleichbarkeit zu ermöglichen, wurde als relatives Mass für die Parameter-Streuung pro Einzugsgebiet der Variationskoeffizient v gewählt. Dieser berechnet sich gemäss Formel (12):

$$(12) \quad |v| = \frac{\sigma}{\bar{x}}$$

Damit, dass die Standardabweichung durch den Mittelwert normiert wird, wird die Streuung der verschiedenen Parameter, die unterschiedliche Wertebereiche aufweisen, vergleichbar. Dabei besagt ein Wert über 1, dass die Standardabweichung grösser als der Mittelwert ist. Nicht möglich war diese Berechnung allerdings für TT: Weil dieser Wert um 0 schwankt, ist es nicht sinnvoll, den Variationskoeffizienten zu berechnen. Für diesen Parameter wurden daher die Werte für jedes Einzugsgebiet in einem eindimensionalen Scatterplot dargestellt.

Zur Auswertung wurde ein Vorgehen gewählt, bei dem die Parameter-Sensitivität aus zwei unterschiedlichen Blickwinkeln (Analyse 1 und Analyse 2) untersucht wurde:

Analyse 1: Pro Einzugsgebiet wurden 10 Kalibrierungen mit dem regulären Gütemass durchgeführt und deren Parameterstreuung ausgewertet. 10 Kalibrierungen können zwar nicht als statistisch aussagekräftig betrachtet werden, ermöglichen aber dennoch einen Eindruck des Parameter-Verhaltens.

Analyse 2: Für diese Analyse wurden die einzelnen Zielfunktionen neu gewichtet und kombiniert und zu unterschiedlichen, aber ähnlichen neuen Gütemassen gemäss Tab. 7 kombiniert. Für jedes Gütemass wurde ein Kalibrierungslauf durchgeführt, insgesamt also 5 Kalibrierungen für Analyse 2. Die so erzielte Parameterstreuung wurde gleich wie in Analyse 1 ausgewertet. Die Kombinationen wurden so gewählt, dass jeweils eine der drei Zielfunktionen MARE, LogReff und Reff mit dem Volumenfehler in derselben Gewichtung wie im regulären Gütemass kombiniert wurden. Dadurch soll eine gewisse Vergleichbarkeit erhalten bleiben. Als Gegenpole zu MARE und LogReff wurden der Reff und das Lindström-Measure, die stärker auf Abflussspitzen fokussieren, neu mit einbezogen. Man mag gegen dieses Vorgehen argumentieren, dass verschiedene Gütemasse unterschiedliche Ergebnisse erzielen, die nicht direkt verglichen werden können. Dem kann aber entgegengehalten werden, dass das eigentliche Ziel jeder Kalibrierung sinnvolle Modellierungen im allgemeinen Sinne sind und diese mit unterschiedlichen Gütemassen erreicht werden können. Daher ist nicht nur von Interesse, die Parameterstreuung innerhalb eines einheitlichen Gütemasses zu untersuchen (dies entspricht Analyse 1), sondern auch die Parameterstreuung unter leicht veränderten Kriterien zu betrachten. Damit wird indirekt auch die Sensitivität des Modells auf der Wahl der Zielfunktionen und deren Gewichtungen untersucht.

	MARE	LogReff	Reff	Volume-Error	Lindström-Measure
Gütemass 1	0.9			0.1	
Gütemass 2		0.9		0.1	
Gütemass 3			0.9	0.1	
Gütemass 4					1

Tab. 7: Zur Analyse der Parameter-Sensitivität (Analyse 2) verwendete kombinierte Gütemasse und ihre Gewichtung.

Um den Rechenaufwand und damit den Zeitaufwand zu reduzieren, wurden alle Kalibrierungen mit 20'000 GAP-Runs (anstatt 50'000) und 200 Parameter-Sets (statt 300) durchgeführt. Damit konnte der Zeitaufwand für 10 Kalibrierungsläufe (Analyse 1) pro Einzugsgebiet auf etwa 5 Stunden, für 5 Kalibrierungsläufe (Analyse 2) entsprechend auf 2.5 Stunden reduziert werden.

Als Parameter-Grenzen wurden dieselben wie für die Hauptkalibrierung verwendet (siehe Tab. 6).

5.7. Alternative Modellstrukturen für Karst-Gebiete

Niedrigwasser wird stark von der zu Grunde liegenden Hydrogeologie beeinflusst (Clausen & Rasmussen, 1993: 298). Verkarstung ist dabei ein Faktor, der das hydrologische Regime entscheidend bestimmt. Karst ist ein heterogener Aquifer: Einerseits besitzt er die niedrige Matrix-Porosität des eigentlichen Gesteins, andererseits die hohe Porosität, die durch die verkarsteten Gänge und grossen Risse im Gestein verursacht werden (Hartmann et al., 2013: 3345). Bildlich wird von einem System kommunizierender Röhren gesprochen (Sponholz, in: Gebhardt, Glaser, Radtke, & Reuber, 2007:

309). Wasser wird dadurch zu einem Teil sehr schnell, zum andern Teil sehr langsam durch den Untergrund hindurch geleitet.

Im Laufe der Auswertungen wurde festgestellt, dass schlechte Kalibrierungs-Ergebnisse nur bei Gebieten auftraten, die sowohl klein (< 30 km²) als auch teilweise verkarstet sind. Allerdings zeigten nicht alle Gebiete mit Karstanteil schlechte Ergebnisse. Hierfür sind vermutlich hauptsächlich zwei Gründe verantwortlich:

- Verkarstung ist nicht gleich Verkarstung: Sie kann sehr unterschiedlich stark ausgeprägt sein und verschiedene Formen annehmen. Damit wirkt sie sich auch unterschiedlich auf die Hydrologie aus. Wo neben verkarsteten Gebieten ausserdem solche mit ausreichend gut entwickelten Böden vorhanden sind, dürfte der Karst-Effekt stark abgeschwächt werden. Um diesem Punkt genauer auf den Grund zu gehen, wären detaillierte Felduntersuchungen notwendig.
- Den Karstanteil der Einzugsgebiete zu ermitteln war schwieriger, als zunächst erwartet. Beispielsweise kann einer geologischen Karte nicht direkt entnommen werden, wo verkarstete Gebiete vorkommen, da geologische Einheiten normalerweise nicht einheitlich aus einem einzigen Material aufgebaut sind. Die Verkarstungsfähigkeit des Gestein wird aber durch die Reinheit des Kalziumkarbonates bedingt (Sponholz, in: Gebhardt, Glaser, Radtke, & Reuber, 2007: 309).

Für die Berechnung wurde schliesslich ein GIS-Layer des Kantons Aargau verwendet, auf dessen Grundlage die Verwendung von triazinhaltigen Herbiziden eingeschränkt wird. In Abb. 13 ist die geografische Lage dieser Karstgebiete innerhalb der verwendeten Einzugsgebiete dargestellt. Die Flächen entsprechen gemäss der Datenbeschreibung¹⁰ zwar nicht der rein geologischen Definition, sie wurden aber mit Fokus auf die Hydrologie aufgrund der geologischen Karte in Kombination mit Felderkundungen erstellt und bilden damit erstens die beste auffindbare Datengrundlage und entsprechen zweitens durchaus dem Verwendungszweck dieser Arbeit.

Gerade aufgrund der erwähnten Heterogenität von Karstmaterial betonen Hartmann et al. (2012: 130), dass es besonders wichtig ist, für die Modellierung von Karst adäquate Modellstrukturen zu verwenden. Aus diesem Grund wurden für diejenigen Gebiete, die einen Karstanteil aufweisen und nicht zufriedenstellende Modellgütewerte erreicht hatten, alternative HBV-Strukturen getestet.

HBV-light bietet dazu verschiedenen Varianten, die bereits standardmässig implementiert sind. Drei davon, die Varianten „Delay“, „One GW Box“ und „Three GW Boxes“ wurden hier in Betracht gezogen. Die englischen Namen illustrieren das jeweilige Hauptcharakteristikum der drei Varianten: „Delay“ beinhaltet eine Verzögerung des Abflusses, „One GW-Box“ berechnet die Grundwasserroutine in nur einer Box, wodurch eine Kommunikation zwischen oberen und unteren Grundwasservorkommen ermöglicht wird, und „Three GW Boxes“ verwendet eine zusätzliche Unterteilung der Grundwasserroutine in drei Boxen. Die entsprechenden Strukturen sind in Abb. 14 bis Abb. 16 dargestellt. Eine Untersuchung zum Verhalten dieser drei Modell-Varianten wurde von Uhlenbrook, Seibert, Leibundgut & Rodhe (1999) durchgeführt.

¹⁰ Verfügbar unter <https://www.ag.ch/geoport/geodatenshop/datensuche.aspx> (letzter Zugriff: 1. 6. 2014). Besagter Layer ist unter dem Namen „Karstgebiete (Pflanzenschutz)“ zu finden.

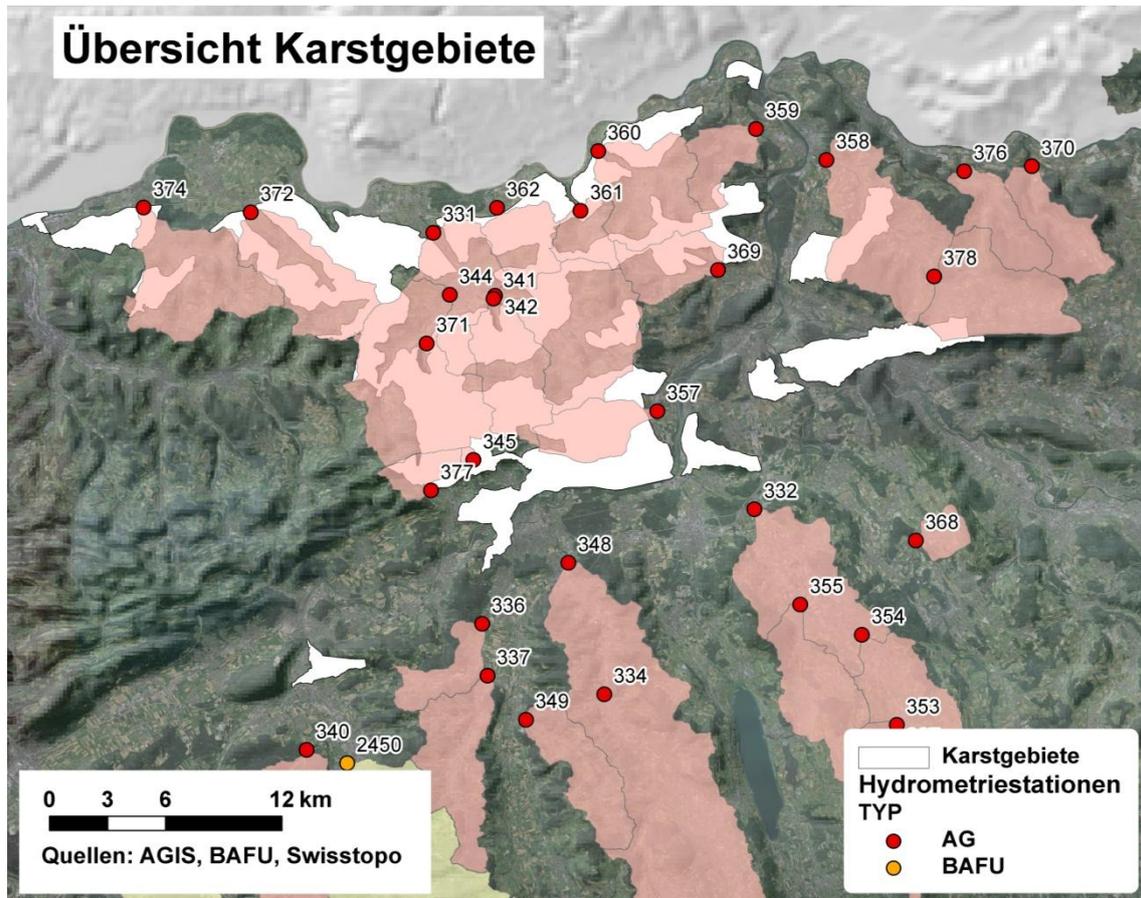


Abb. 13: Lage der verkarsteten Gebiete gemäss GIS-Layer des Kantons Aargau, der die Grundlage für Einschränkungen im Einsatz triazinhaltiger Herbizide bildet.

Die Delay-Variante stand aus zwei Gründen im Fokus der Untersuchungen: Erstens scheint sie von ihrem Aufbau her – der Input in die Grundwasser-Routine wird in einen schnellen Abfluss und einen verzögerten Abfluss aufgeteilt – am ehesten der unterschiedlichen Durchlässigkeit von Karstgestein zu entsprechen; zweitens unterscheidet sie sich am grundlegendsten von der HBV-Standardversion, worin eine grössere Chance besteht, dass sie ein andersartiges und damit vielleicht passenderes Abflussverhalten erzeugt. Auch bei Uhlenbrook et al. (1999) erbrachte die Delay-Version diejenigen Ergebnisse, die sich am stärksten von den übrigen unterschieden – in ihrem Fall zwar schlechtere, im vorliegenden Fall wird aber gerade dieser Unterschied als vielversprechend gewertet. Allerdings sehen Seibert (2000: 219) und Van Loon & Van Lanen (2012: 1919) die Delay-Variante eher als Repräsentation für Einzugsgebiete mit langsam reagierendem, tiefliegendem Aquifer; ihre Eignung für die Simulation von Karststrukturen wird in der Literatur nicht besprochen.

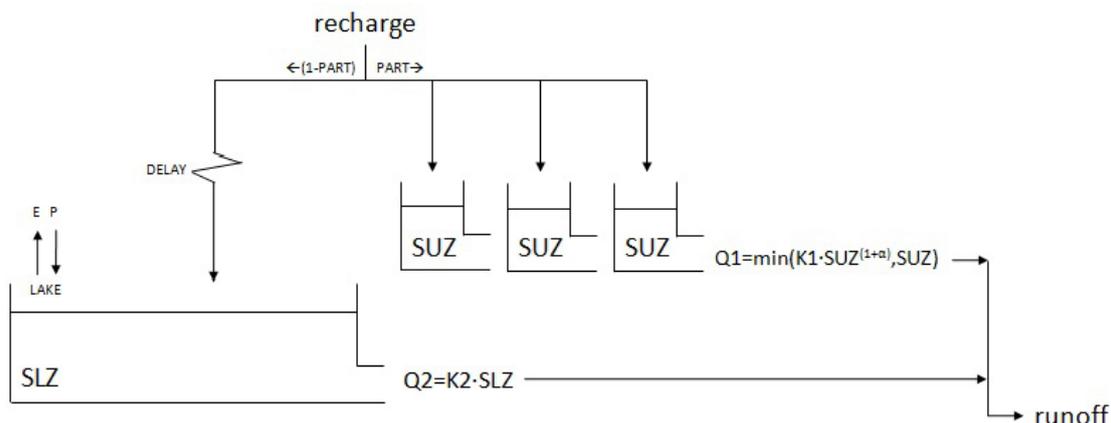


Abb. 14: Struktur der Modellvariante “Response Routine With **Delay**”. Der Anteil an Wasser, der in die Grundwasser-Routine gelangt, wird in zwei Teile aufgespalten: Der eine Teil (bestimmt durch den Parameter PART [-]) bildet den „schnellen“ Abfluss, der andere Teil wird über einen bestimmten Zeitraum (Parameter DELAY [d]) verzögert und kommt daher erst später zum Abfluss. Die Aufteilung in verschiedene SUZ-Boxen entsteht durch die Verwendung verschiedener räumlicher Einheiten wie z.B. Höhenstufen oder Vegetationszonen. Quelle: HBV-Help

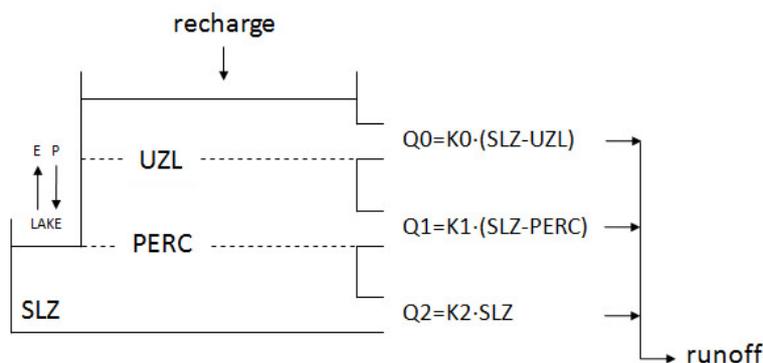


Abb. 15: Struktur der Modellvariante “**One GW Box**”. Anstelle von zwei getrennten Boxen wie im Standardmodell wird hier für die Grundwasser-Routine nur eine einzige Box mit drei Ausflüssen verwendet. Uhlenbrook et al. (1999) sehen diese Struktur als mögliche Repräsentation eines nahe an der Oberfläche liegenden Aquifers, der mit den Oberflächengewässern in engem Kontakt steht. Quelle: HBV-Help

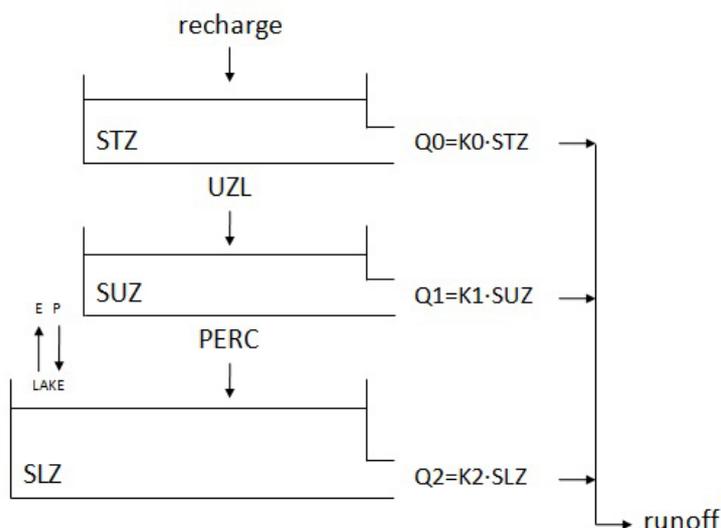


Abb. 16: Struktur der Modellvariante “**Three GW Boxes**”. Uhlenbrook et al. (1999) sehen diese Struktur als mögliche Repräsentation für mehrere, übereinander liegende Aquifere und hangende Grundwasservorkommen. Quelle: HBV-Help

5.8. Grundwasserdaten zur Verbesserung der Vorhersage

5.8.1. Motivation zur Multi-Kriterien-Analyse

Im Normalfall steht Grundwasser in stetigem, mehr oder weniger engem Kontakt mit den Oberflächengewässern. Bei ausbleibendem Niederschlag und einsetzender Bodentrockenheit sind es, wo vorhanden, die Grundwasserreserven, die die Oberflächengewässer speisen (Rees et al., in: Tallaksen & van Lanen, 2004: 110; Smakhtin, 2001). Daher sind Grundwasserstände insbesondere für die Untersuchung von Niedrigwasser ein wichtiger und hilfreicher Aspekt.

Daher wurde in dieser zweiten Phase getestet, ob die Niedrigwasser-Vorhersagen durch eine Co-Kalibrierung des Modells auf Abfluss- **und** Grundwasserdaten verbessert werden können. Dabei wurde angenommen, dass die Grundwasserspiegel direkte Indikatoren für den Zufluss von Grundwasser in die Oberflächengewässer darstellen. Der Anstoss für diesen Schritt wurde durch Seibert (2000) gegeben; zusätzlich zeigten Van Loon & Van Lanen (2012), dass GW-Daten wertvoll für Simulationen sein können bzw. dass HBV gut zur Simulation von Grundwasserpegeln verwendet werden kann (auch wenn diese Autoren keine Co-Kalibrierung durchführten).

In der bis hierhin verwendeten Version von HBV-light ist eine Kalibrierung auf Grundwasserdaten nicht möglich. Anlässlich dieser Arbeit wurde daher von M. Vis eine neue HBV-Version, die eine Multi-Kriterien-Kalibrierung ermöglicht, programmiert. Weil hier nur die zusätzlich verfügbare Grundwasserkalibrierung verwendet wird, wird dieses Modell im Weiteren mit „HBV-GW“ bezeichnet.

Das Ziel einer Multi-Kriterien-Kalibrierung besteht darin, das interne Modellverhalten zu verbessern, i.e. einen realitätsnäheren Verlauf zu erzielen. Es ist anzunehmen, dass damit die Zuverlässigkeit des Modells verbessert wird (siehe z.B. Seibert, 2000; Seibert, Rodhe, & Bishop, 2003; Son & Sivapalan, 2007; Hartmann et al., 2013). Konkret soll hier erreicht werden, dass die Grundwasser-Routine ein möglichst realitätsgetreues Verhalten annimmt. Das Modell wird also schon einen Schritt vor der Abfluss-Routine kalibriert, was sich ebenfalls auf die davor liegenden Routinen auswirkt. Gemäss Van Loon & Van Lanen (2012: 1922) ist zu erwarten, dass diese Co-Kalibrierung nicht für alle Einzugsgebiete gleich vorteilhaft für die Simulation von Niedrigwasser ist:

„In fast reacting systems (...), discharge drought characteristics are more comparable to those of precipitation (more and shorter); in slowly reacting systems (...), discharge drought characteristics are more comparable to those of groundwater storage (fewer and longer).“

“In schnell reagierenden Systemen sind Niedrigwasser-Eigenschaften stärker mit jenen einer meteorologischen Trockenheit (häufig und kurz) vergleichbar, in langsam reagierenden Systemen sind Niedrigwasser-Eigenschaften hingegen stärker mit jenen von Grundwasser-Trockenheiten (seltener und länger anhaltend) vergleichbar.“

Überdies spielt es eine Rolle, wo die Grundwassermessstationen relativ zum Fließgewässer gelegen sind: Seibert, Bishop, Rodhe & McDonnell (2003) haben aufgezeigt, dass die Korrelation zwischen Grundwasser und (Oberflächen-) Abfluss mit zunehmender Distanz zum Fließgewässer hangaufwärts abnimmt. Entsprechend unterscheidet sich der Informationsgehalt, der aus Messungen aus unterschiedlicher Distanz gewonnen werden kann. Ebenso zeigten Grundwassermessstellen untereinander eine hohe Korrelation, wenn sie sich in ähnlicher Distanz zum Gewässer befanden, wobei die Korrelation grösser wurde, wenn weiter vom Fließgewässer entfernte Stationen betrachtet wurden. In Tab. 2 sind daher die Korrelation jeder Station zum Abfluss sowie die Distanz zum Fließgewässer und die Distanz zur Abflussmessstation mit aufgeführt.

Weil im Voraus aber nicht gesagt werden kann, welche Messpunkte die für HBV vorteilhafteren Informationen enthalten bzw. weil HBV sozusagen einen über das gesamte Einzugsgebiet gemittelten Grundwasserspeicher simuliert, werden in der vorliegenden Arbeit die Grundwasserdaten pro Einzugsgebiet zu einem arithmetischen Mittel verrechnet (siehe dazu Kapitel 4.2.2).

5.8.2. Rangkorrelation nach Spearman als zusätzliche Zielfunktion

Zur Kalibrierung der Grundwasser-Boxen des Modells auf die gemessenen Grundwasserpegel wird eine zusätzliche Zielfunktion benötigt, weil simuliertes und gemessenes Grundwasser nicht gleich direkt vergleichbar sind wie simulierter und gemessener Abfluss. Modellboxen und Grundwasser-Messwerte befinden sich nicht in derselben Einheit: HBV berechnet die Grundwasser-Boxen in Millimeter (Wassersäule), Messwerte werden in Meter über Meer angegeben; um die gemessenen Grundwasserpegel in dieselbe Einheit umrechnen zu können, müssten die Ausmasse des Aquifers bekannt sein. Dies ist aber gar nicht nötig, weil das Ziel nur darin besteht, ein übereinstimmendes Verhalten zwischen Modellboxen und Messwerten zu erzielen. Dafür kann eine Korrelationsfunktion verwendet werden; die Einheit muss dazu nicht übereinstimmen. Es wird dabei angenommen, dass im Modell die Summe von SUZ (obere Grundwasser-Box) und SLZ (untere Grundwasser-Box) die Entsprechung zu den Grundwasser-Messwerten bildet.

Abb. 17 zeigt den Zusammenhang zwischen gemessenem Grundwasserspiegel und simuliertem Grundwasser aus der Standardkalibrierung (links) und nach der Co-Kalibrierung auf Grundwasser (rechts) am Beispiel des Einzugsgebietes 341. Daran wird deutlich, dass eine monotone, aber nicht unbedingt eine lineare Korrelation angenommen werden kann (vgl. hierzu Seibert, Bishop, et al. (2003), die eine nicht-lineare Korrelation in Flussnähe und eine lineare in Hanglagen beobachteten). Aus diesem Grund wird als Zielfunktion nicht die Pearson-Korrelation, sondern die robuste Rangkorrelation nach Spearman verwendet (z.B. Wilks, 2006: 55ff). In HBV-GW wird diese Zielfunktion mit „GroundWaterSR“ bezeichnet.

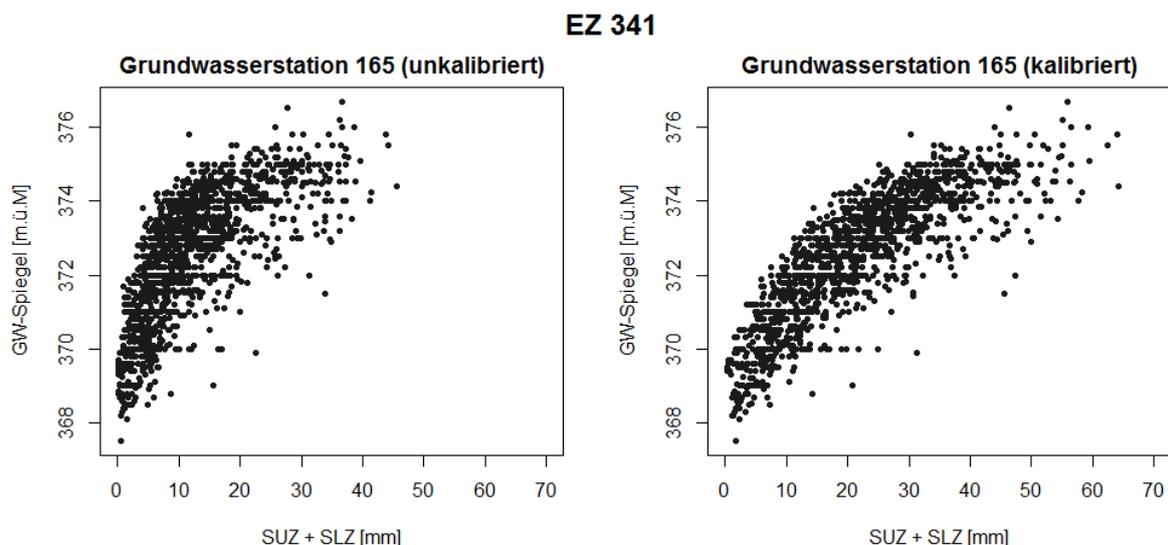


Abb. 17: Korrelation zwischen gemessenem Grundwasserspiegel [m.ü.M.] und den HBV-Grundwasser-Boxen SUZ und SLZ [mm]. Links sind die Ergebnisse aus der Standard-Kalibrierung abgebildet, rechts nach der Co-Kalibrierung auf die Grundwasserdaten. Die Korrelation wurde dadurch sichtlich verstärkt.

Mit der zusätzlichen Zielfunktion mussten die Gewichtungen des kombinierten Gütemasses neu verteilt werden. Die Schlüsselfrage hierbei ist, wie stark die neue Zielfunktion zur Grundwasser-Kalibrierung gewichtet werden soll.

An den beiden Gebieten 331 und 341 wurden daher drei verschiedene Gewichtungen getestet: Erstens Abfluss und Grundwasser zu je 50% gewichtet, zweitens Abfluss zu 60% und Grundwasser zu 40% gewichtet, und drittens Abfluss zu 80% und Grundwasser zu 20% gewichtet. Die Grundwasser-Kalibrierung mit mehr als 50% zu gewichten, wurde nicht für sinnvoll erachtet, schliesslich stellt die Grundwasser-Routine nur einen Teil des Modells dar und darf daher nicht höher gewichtet werden als die Abflusssimulation, die sozusagen das Modellfazit darstellt.

Die Beurteilung, welche Gewichtung am passendsten ist, wurde in erster Linie visuell vorgenommen. Dabei liess sich in der Tendenz erkennen, dass sich das Rezessionsverhalten unter Einbezug der Grundwasserdaten verbesserte. Bei einer 50%-Gewichtung zeigten sich Niedrigwasserstände aber eher überschätzt. Es schien daher sinnvoll, die Grundwasserkalibrierung mit weniger als 50% zu gewichten. Werden die Grundwasserdaten aber mit nur 20% gewichtet, bringt dies relativ wenig neue Information ins Modell ein. Das zeigt sich auch in den Werten der Zielfunktion *GroundWaterSR*: Für das Gebiet 341 wurde ohne Grundwasser-Co-Kalibrierung der Wert 0.783 erreicht, hingegen 0.815 für eine Gewichtung mit 80 zu 20 und 0.86 für die Gewichtung 60 zu 40; für das Gebiet 331 wurde 0.719 ohne Co-Kalibrierung, 0.73 mit der Gewichtung 80 zu 20 und 0.753 mit der Gewichtung 60 zu 40 erreicht. In beiden Einzugsgebieten ist damit eine deutliche Verbesserung des Gütemasses bei einer Grundwassergewichtung von 40% im Vergleich zu 20% zu erreichen. Aus diesen Gründen wurde schliesslich die Gewichtung 60% Abfluss und 40% Grundwasser gewählt.

Damit der Abfluss aber nach denselben Kriterien wie bisher kalibriert wurde, musste diese neue Gewichtung mit der bisherigen verknüpft werden. Die Gewichtungen des bisherigen kombinierten Gütemasses wurden daher mit dem Faktor 0.6 multipliziert. Die sich daraus ergebende neue Gewichtung ist in Tab. 8 zusammengestellt.

Gütemass	MARE	LogReff	VolumeError	GroundWaterSR
Gewichtung	0.3	0.24	0.06	0.4

Tab. 8: Gewichtung der verschiedenen Zielfunktionen für die Co-Kalibrierung auf Abfluss und Grundwasser.

6. Ergebnisse und Auswertungen

6.1. Standard-Modell

6.1.1. Modellgüte

Mit dem Standardmodell wurde in der Kalibrierung eine durchschnittliche Modellgüte (kombiniertes Gütemass gemäss Tab. 5) von 0.681 (arithmetisches Mittel) mit einem Median von 0.719 erreicht. Die Ausreisser, welche die Erniedrigung des Mittelwertes im Vergleich zum Median bewirken, sind im Boxplot (Abb. 18) deutlich ersichtlich. Die Werte der einzelnen Einzugsgebiete sind im Anhang in Tab. 16 zusammengestellt.

Mit Blick auf die selbst festgesetzten Gütegrenzen erreichen oder überschreiten 23 der 36 Gebiete, also knapp zwei Drittel, den Wert 0.7 und sind damit als „gut“ einzustufen, 2 dieser 23 Gebiete überschreiten sogar die Marke von 0.8 und fallen in die Kategorie „sehr gut“.

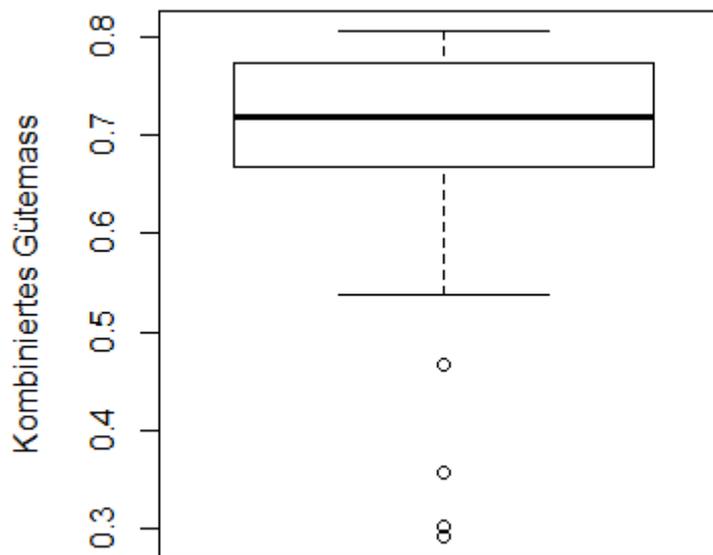


Abb. 18: Boxplot der Modellgüten aller Einzugsgebiete.

In Abb. 19 sind die Ergebnisse der Kalibrierung mit der HBV-Standard-Modellversion zusätzlich räumlich ausgewertet dargestellt. Dabei fällt auf, dass alle drei Gebiete mit den schlechtesten Ergebnissen der Modellgüte sich im Aargauer Jura mit Ausfluss gegen die Aare hin befinden. Aber auch eines der beiden Gebiete, die ein sehr gutes Resultat erzielten, liegt in den (westlichen) Jurahöhen.

6.1.2. Trockenheits-Masse

Die mit der Standardversion des HBV-Modells berechneten Werte für Dauer, Defizit und Intensität sind für die beiden Schwellenwerte Q_{347} und RW_{min} in Abb. 20, Abb. 22 resp. Abb. 26 grafisch dargestellt; die detaillierten Werte finden sich für Dauer, Defizit und Intensität im Anhang in Tab. 17 (Auswertungsperiode 2003) und Tab. 18 (Auswertungsperiode 2011); die Werte für Onset und Minimalabfluss folgen in den Tab. 19 (2003) Tab. 20 (2011); die grafischen Auswertungen zu Minimum und Onset finden sich in Abb. 29 und Abb. 31. Anzumerken ist, dass in den Grafiken für Defizit und Intensität das Gebiet 345 für den Schwellenwert RW_{min} fehlt. Weil dieses Gebiet so klein ist, liegt der Grenzwert RW_{min} extrem hoch, wodurch die Werte von Defizit und Intensität um zwei Größenordnungen über den Werten der übrigen Gebiete liegen.

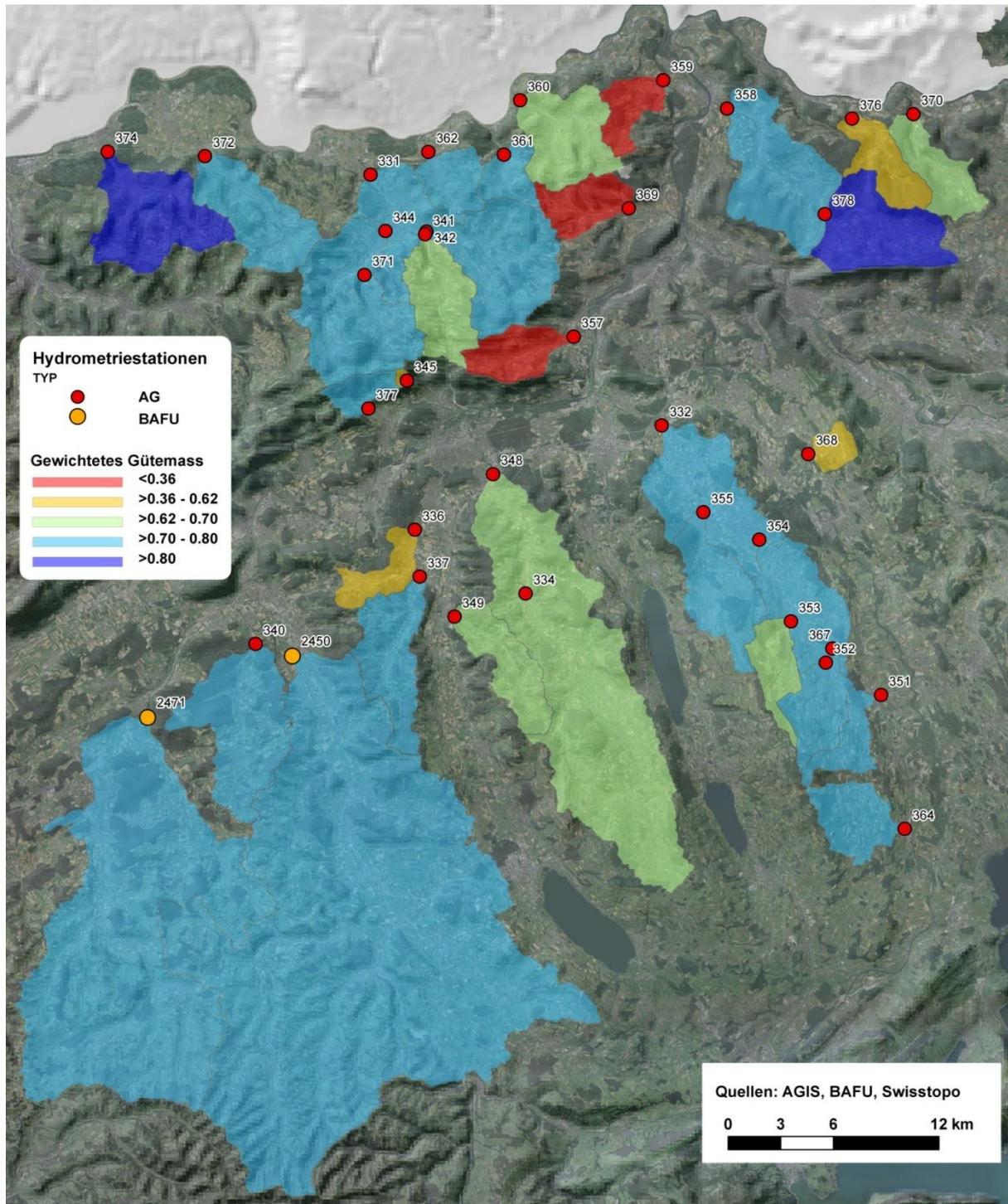


Abb. 19: Räumliche Darstellung der mit der Standardversion erzielten Modellgüten. Auffällig ist, dass sich die drei Gebiete mit den schlechtesten Güteergebnissen alle im Aargauer Jura mit Ausfluss gegen die Aare hin befinden.

Missed Events und *False Alarms* sind in den Abb. 20, Abb. 22 und Abb. 26 anhand der Säulenfarbe leicht identifizierbar: Die Säulen sind halbtransparent dargestellt und in diesen Fällen zeigt sich allein das Grau der beobachteten Werte oder das Blau der simulierten; in Tab. 17 bis Tab. 20 sind *Missed Events* und *False Alarms* schwierig zu erkennen und wurden daher farblich hervorgehoben (*False Alarms* in hellem Orange, *Missed Events* in hellem Rot). Auf diese Fälle wird in Kap. 6.1.3 näher eingegangen. In diesen Tabellen sind zusätzlich zur Auswertung der Trockenheits-Masse prozentuale

Werte als Differenz zwischen simuliertem und beobachtetem Wert dividiert durch den beobachteten Wert angegeben.

Es würde zu weit führen, die Ergebnisse jedes Einzugsgebietes hier einzeln zu erörtern. Einige Ergebnisse sollen aber beispielhaft erläutert werden.

Den Resultatbetrachtungen vorauszuschicken ist, dass 17 der 36 Einzugsgebiete ein Q_{347} unter dem absoluten Mindestwert für RW_{min} von 50 l/s aufweisen. Das bedeutet, dass der RW_{min} -Grenzwert höher als das Q_{347} liegt. Bei den kleinsten Einzugsgebieten kann dies zur Folge haben, dass die RW_{min} -Schwelle so hoch liegt, dass die gesamte Auswertungsperiode unterhalb dieses Schwellenwertes liegt (insbesondere im Jahr 2003).

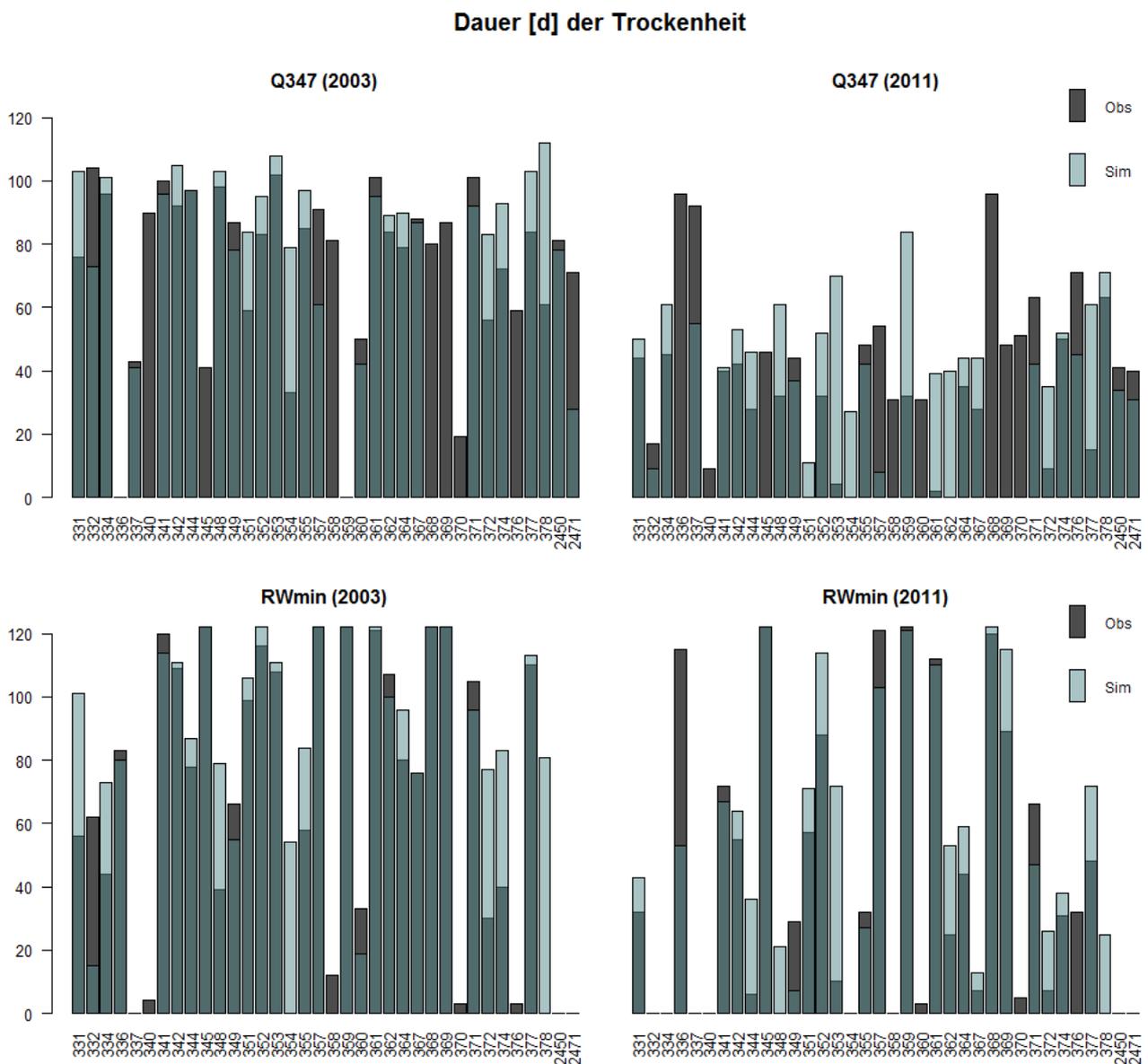


Abb. 20: Vergleich der Simulationen mit den beobachteten Werten für die Trockenheits-Dauer im 2003 (links) und 2011 (rechts) für die beiden Schwellenwerte Q_{347} (oben) und RW_{min} (unten).

6.1.2.1. Dauer

In der Auswertungsperiode 2003 wurde die Dauer sowohl für Q_{347} wie für RW_{min} im Durchschnitt leicht überschätzt (Q_{347} : um 11.6%, RW_{min} : um 17.35%). Im 2011 ist die Tendenz zur Überschätzung jedoch markant stärker (Q_{347} : um 173.7%, RW_{min} : um 74.22%).

Allerdings zeigt sich beim Betrachten der Abflusskurven ein weniger gravierendes Bild, als dies die Zahlen vermuten lassen: Vergleicht man beispielsweise die Gebiete 353 und 361, welche die Trockenheits-Dauer am stärksten überschätzen, so zeigt Nr. 353 einen systematischen Versatz nach unten (Abb. 21), Gebiet Nr. 361 hingegen zeigt eine sehr gute Simulation, sinkt aber im Sommer etwas zu stark ab. Weil gerade in diesem Bereich der Schwellenwert liegt, resultiert ein sehr schlechter Kennwert.

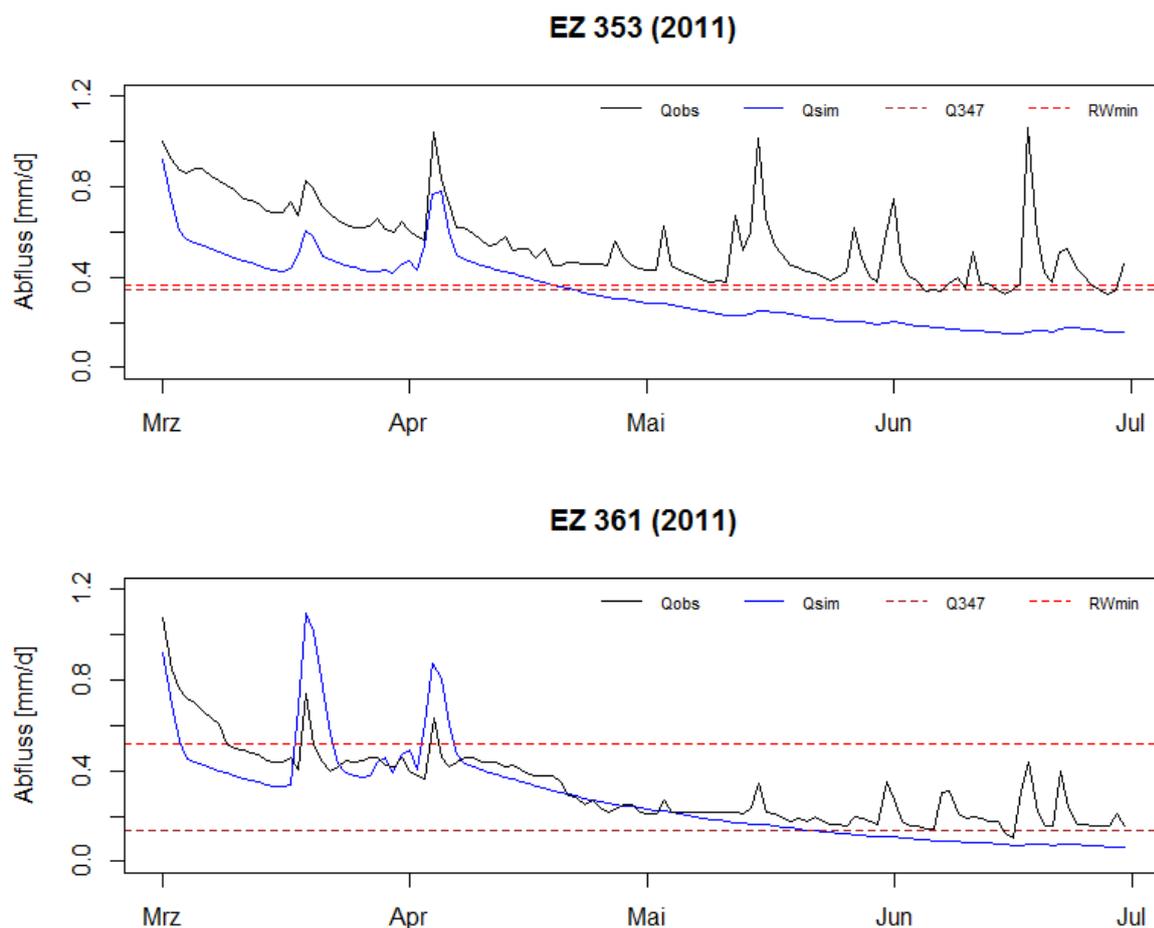


Abb. 21: Gebiete Nr. 353 (oben) und 361 (unten) in der Auswertungsperiode 2011.

Für einige Gebiete mit hohem RW_{min} wurden Abweichungen von 0 erreicht, da sich der Abfluss bereits bei Beginn der Auswertungsperiode von 122 Tagen unter dem Schwellenwert befindet. In Abb. 20 (unten) sind daher mehrere gleich hohe Balken erkennbar, die 122 Tage erreichen. Dennoch können auch aus diesem Kennwert wertvolle Informationen gewonnen werden: Würde das Modell hier versagen, würde dies größere Fehler aufdecken.

6.1.2.2. Defizit

Auch das Trockenheits-Defizit wird vom Modell gemittelt über alle Einzugsgebiete deutlich überschätzt, prozentual gesehen in der Auswertungsperiode 2003 (Q_{347} : um 125%, RW_{min} : um 127.3%) etwas weniger stark als in der Periode 2011 (Q_{347} : um 908,1%, RW_{min} : um 385.5%).

Für Q_{347} im 2011 sind es die bereits in Abb. 21 dargestellten beiden Gebiete 353 und 361, die sehr extreme Überschätzungen liefern. Hinzu kommt Gebiet 377, das ähnlich 353 einen systematischen Versatz nach unten und damit eine starke Überschätzung der Trockenheit zeigt. Für RWmin zeigt ausserdem Gebiet 344 prozentual gesehen grosse Simulationsschwierigkeiten. Ein Blick auf die Abb. 23 zeigt die Ursache des Problems: In den Sommermonaten sackt die Simulation zu stark ab (dasselbe Bild zeigt sich für 344 in der Auswertung 2003) und weil das beobachtete Defizit sehr gering ist, ist die relative Abweichung entsprechend gross.

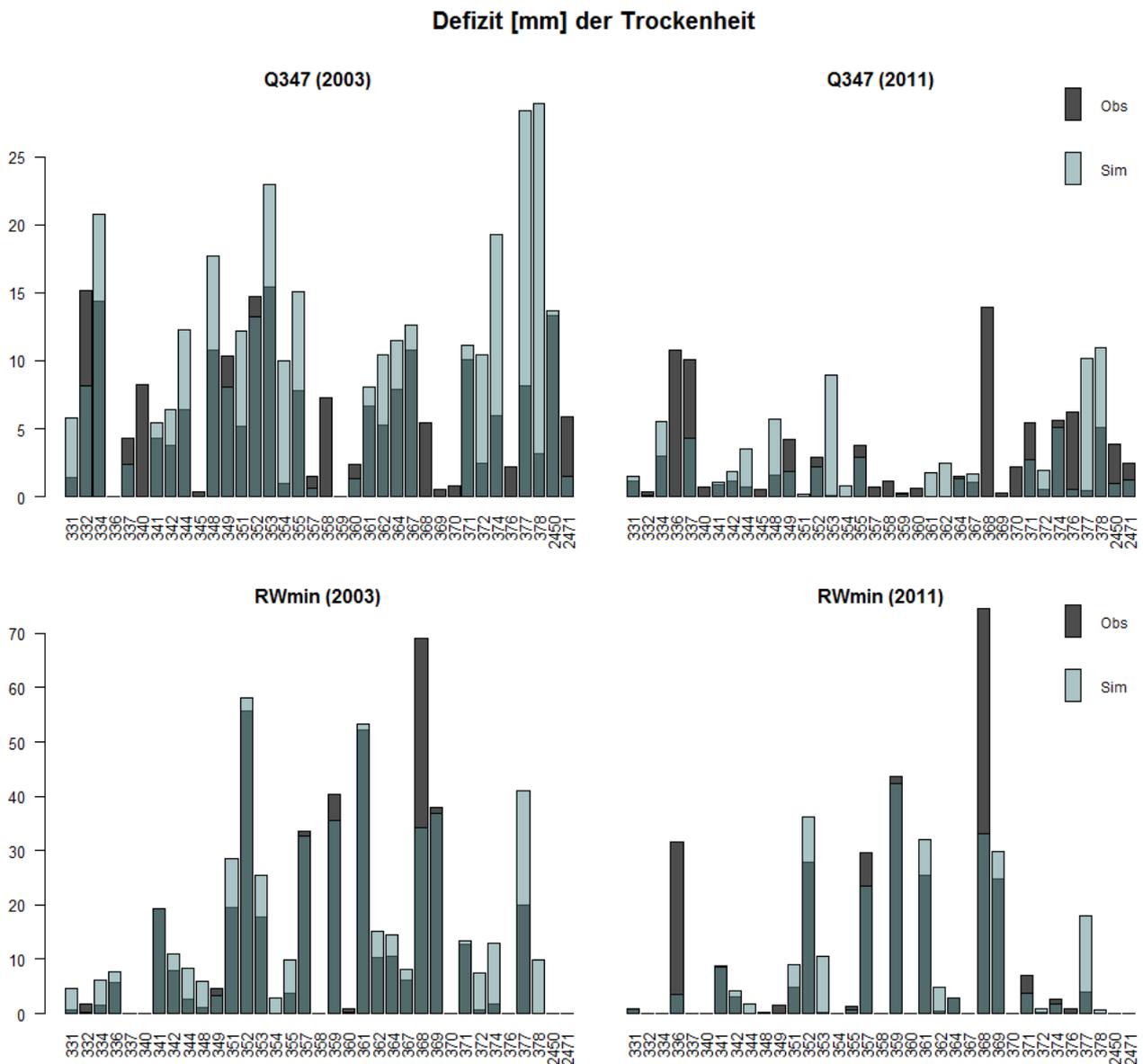


Abb. 22: Vergleich der Simulationen mit den beobachteten Werten für das Trockenheits-Defizit im 2003 (links) und 2011 (rechts). Achtung: Auf den Grafiken für den Schwellenwert RWmin (unten) fehlt aus Gründen der Darstellbarkeit das Einzugsgebiet 345.

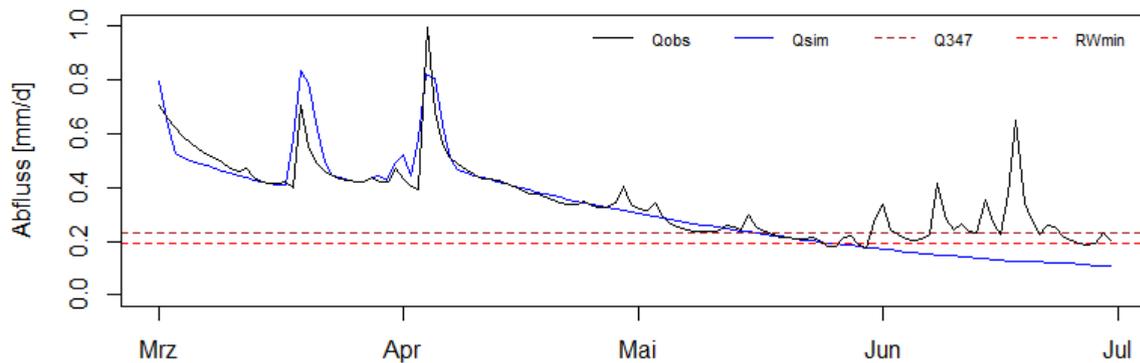
EZ 344 (2011)

Abb. 23: Gebiet 344 in der Auswertungsperiode 2011. Im Sommer sackt die Simulation zu stark ab.

Das gegenteilige Verhalten zeigt beispielsweise das Gebiet 358 (siehe Abb. 24). Hier bleibt die Simulation auf konstant zu hohem Niveau. Zu diesen beiden Gebieten ist erwähnenswert, dass sie gute Ergebnisse in der Kalibrierung erreichten (358 mit 0.711 und 344 mit 0.776).

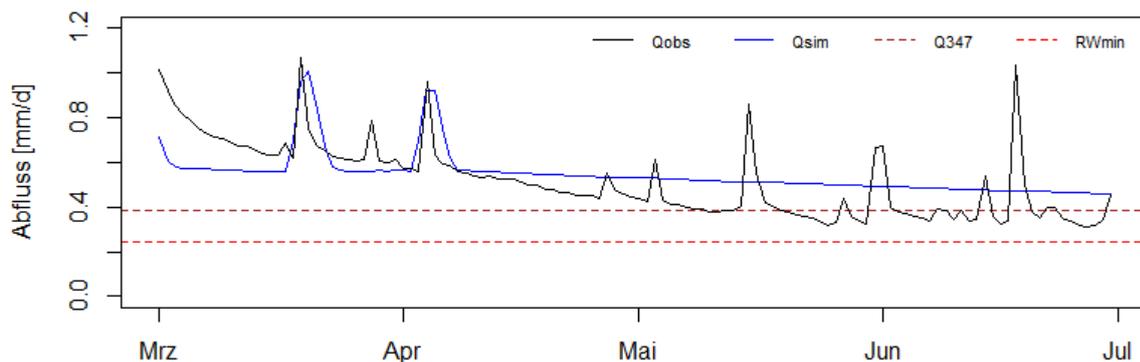
EZ 358 (2011)

Abb. 24: Gebiet 344 in der Auswertungsperiode 2011. Die Simulation bleibt auf zu hohem Niveau.

In der Auswertungsperiode 2003 ist ein bei mehreren Gebieten auftretendes Phänomen, dass der Abfluss über die Sommermonate und im Herbst zu stark abfällt (und die feinen Schwankungen nicht modelliert werden können). Als Beispiel ist in Abb. 25 die Simulation für das Einzugsgebiet 378 dargestellt. Ein möglicher Grund für dieses Verhalten wird in Kap. 7.1 besprochen.

Wie bei der Trockenheits-Dauer stehen bei der Betrachtung der absoluten Werte für RWmin auch hier besonders das Gebiet 345 (Achtung: Ist aufgrund seiner um zwei Größenordnungen höheren Werte nicht abgebildet), aber auch die Gebiete 357, 359, 361, 368 und 369 mit sehr hohen Defizit-Werten ins Auge. Diese werden wiederum durch den hohen RWmin-Wert verursacht. Sie zeigen also nicht eine eigentliche Trockenheit an, nützen aber als Zeiger, ob die Größenordnung der Simulation korrekt ist.

EZ 378 (2003)

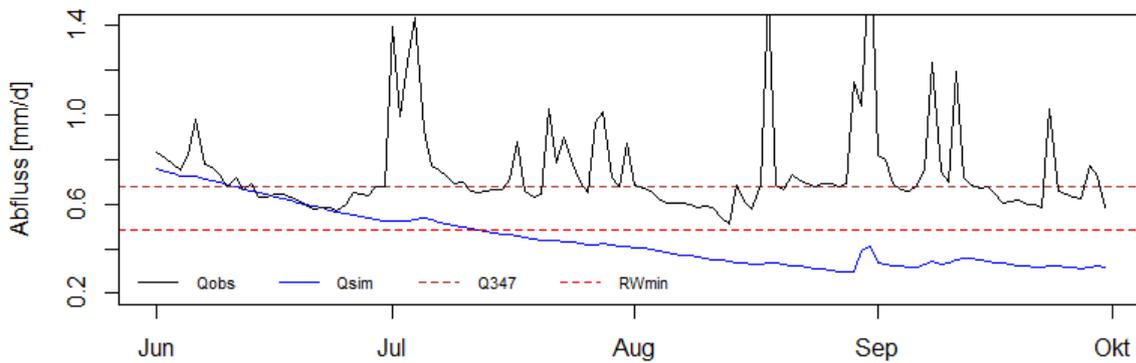


Abb. 25: Gebiet 378 in der Auswertungsperiode 2003. Die Simulation sinkt ab Ende Juni zu stark ab.

Intensität [mm/d] der Trockenheit

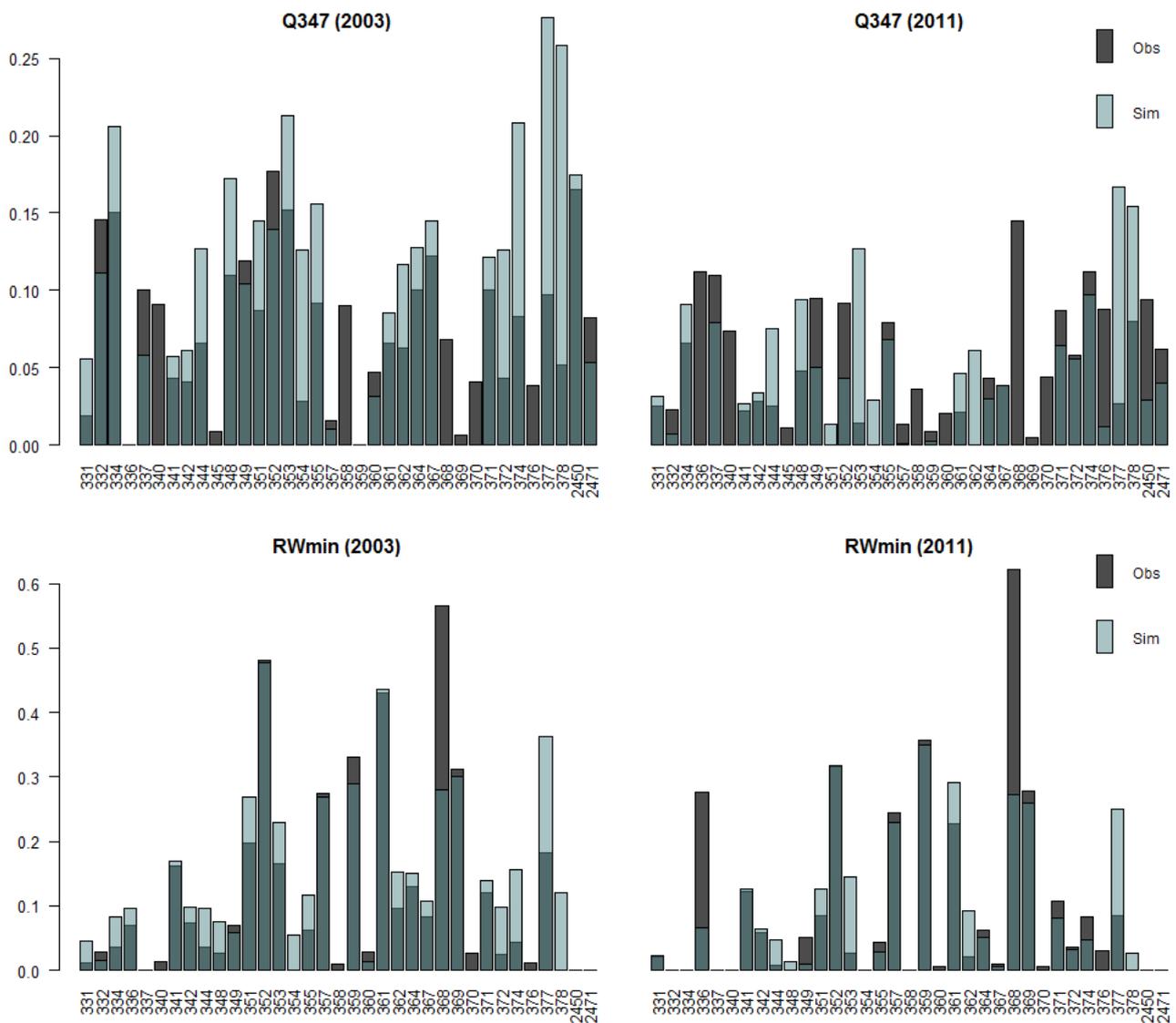


Abb. 26: Vergleich der Simulationen mit den beobachteten Werten für die Trockenheits-Intensität im 2003 (links) und 2011 (rechts). Achtung: Auf den Grafiken für den Schwellenwert RWmin (unten) fehlt aus Gründen der Darstellbarkeit das Einzugsgebiet 345.

6.1.2.3. Intensität

Prozentual betrachtet wird die Intensität durch das Modell besser vorhergesagt als das Defizit, aber dennoch im Schnitt um gut die Hälfte überschätzt, wobei die Überschätzung im 2003 (Q_{347} : um 70.3%, RWmin: um 54%) ist leicht höher als im 2011 (Q_{347} : um 51.3%, RWmin: um 52.2%) ist.

Der Grund für diese Verbesserung vom Defizit zur Intensität liegt in folgendem Zusammenhang: Wie bereits festgestellt, können die HBV-Simulationen die feinen Abflussschwankungen im Niedrigwasserbereich nicht nachbilden. Wenn der beobachtete Abfluss nun Schwankungen über den Schwellenwert aufweist, wird dadurch sowohl das Defizit wie auch die Dauer der beobachteten Trockenheit verkleinert. Bleibt der simulierte Abfluss gleichzeitig auf tiefem Niveau, weist er ein grösseres Defizit, aber auch eine längere Dauer auf. Und weil sich die Intensität als Quotient aus Defizit und Dauer berechnet, kann nun die Auswertung für die Intensität relativ gesehen besser als diejenige für das Defizit ausfallen (genauso wäre aber auch eine Verschlechterung möglich). Dies ist beispielsweise für das Gebiet 342 im Frühjahr 2011 sowie für das Gebiet 364 im Sommer 2003 der Fall (Abb. 27).

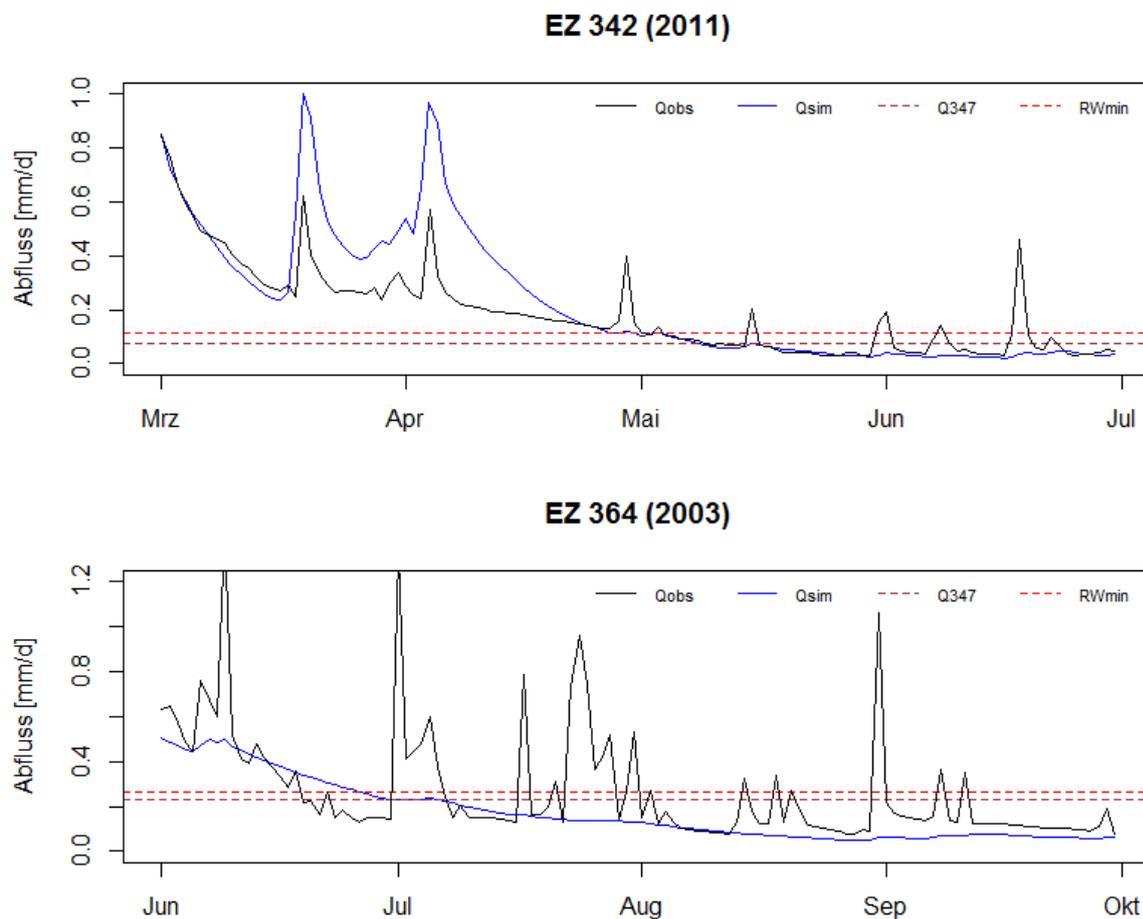


Abb. 27: Gebiet 342 (oben) in der Auswertungsperiode 2011 und Gebiet 364 (unten) in der Auswertungsperiode 2003. Beide zeigen für die Intensität eine bessere Vorhersage als für das Defizit.

Ein weiterer interessanter Effekt der Quotient-Bildung kann am Gebiet 372 (Abb. 28) in der Auswertungsperiode 2011 beobachtet werden: Während das Defizit klar überschätzt wird, ist die Intensität leicht unterschätzt, weil die Dauer entsprechend zu lang simuliert ist.

EZ 372 (2011)

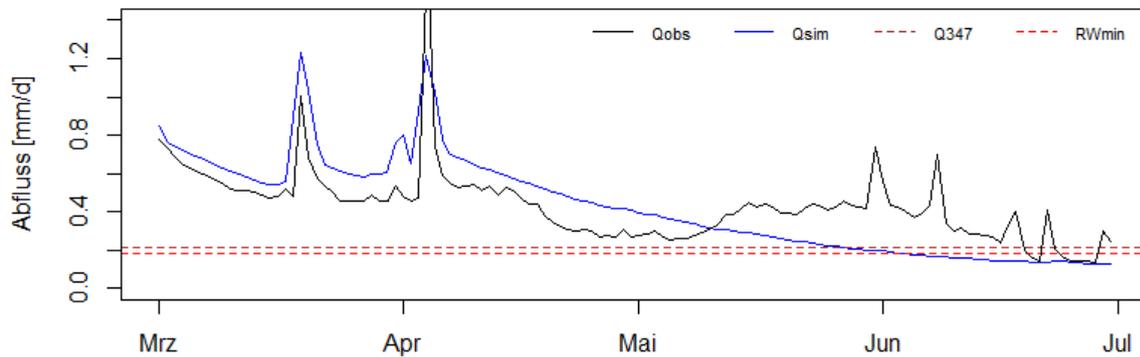


Abb. 28: Gebiet 372 in der Auswertungsperiode 2011. Hier zeigen Defizit und Intensität ein interessantes Verhalten.

Onset-Differenz [d] Sim - Obs

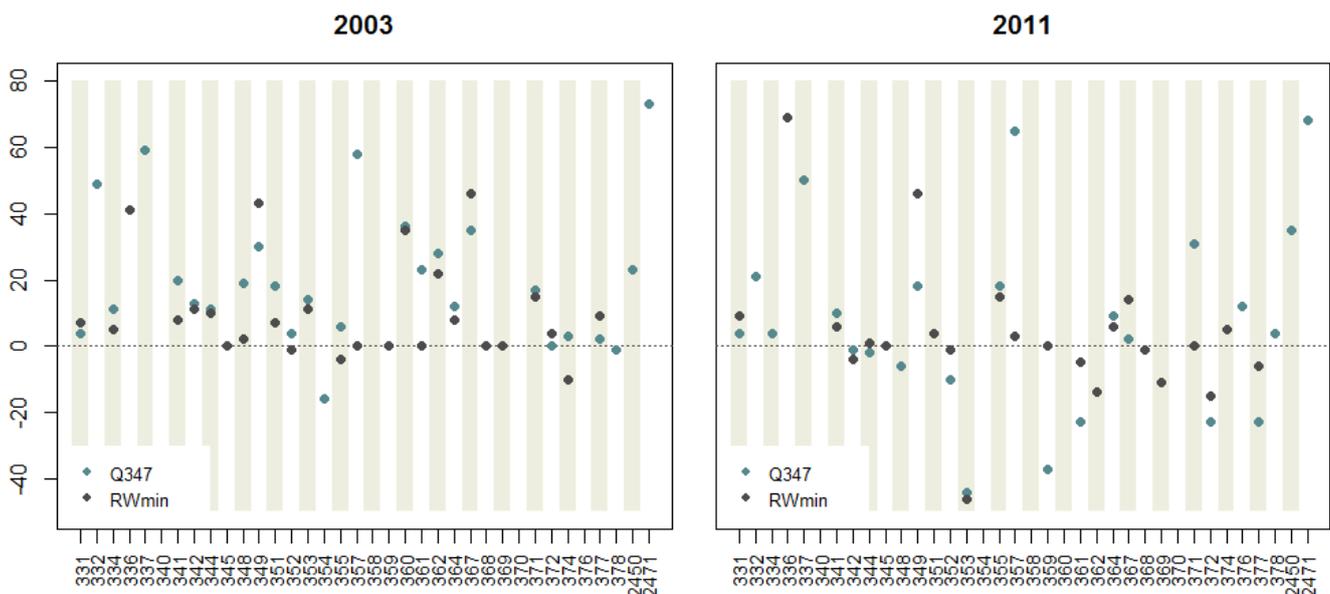


Abb. 29: Vergleich der Differenzen im Onset für die beiden Schwellenwerte Q347 und RWmin. Die Differenz in Anzahl Tagen wurde berechnet als simuliertes Datum – beobachtetes Datum. Positive Werte stehen damit für einen verspäteten simulierten Onset, negative für einen verfrühten.

6.1.2.4. Onset

Das Einsetzen der Trockenheit wird im Sommer 2003 im Schnitt gut 20 Tage für Q_{347} und gut 13 Tage für RWmin zu spät simuliert. Für die Auswertungsperiode 2011 wird der Onset deutlich besser, aber ebenfalls im Schnitt etwas zu spät erfasst: Für Q_{347} um gut 7 Tage und für RWmin um gut 3 Tage.

Aus Abb. 29 ist diese Tendenz deutlich ablesbar: Im 2003 treten fast ausschliesslich verspätete Simulationen auf, in drei Fällen sogar mit einem Versatz von zwei Monaten und mehr. Im 2011 sind verfrühte und verspätete Simulationen stärker ausgeglichen. Die grössten Versätze von verfrühten Simulationen haben zugenommen, diejenigen der verspäteten Simulationen dafür abgenommen. Die häufiger auftretenden Abweichungen von 0 für RWmin sind wiederum darauf zurückzuführen, dass sich in mehreren Fällen die gesamte Auswertungsperiode unter dem Schwellenwert befindet. Diese Vorhersage ist für das Modell relativ einfach. Dies betrifft die Gebiete 345 und 357 in beiden

Auswertungsperioden sowie die Gebiete 359, 361, 368 und 369 im 2003. Deren Werte wurden in die Mittelwert-Berechnung mit einbezogen und verbessern damit den Mittelwert. Die Durchschnittswerte fallen deshalb für RWmin in beiden Auswertungsperioden besser aus als für Q_{347} . Diese Größe wird dadurch aber nicht belanglos: Wie bereits für Dauer und Defizit erwähnt, würde hier eine Fehlvorhersage grobe Modellfehler anzeigen. Ausserdem bedeutet ein bei Beginn der Auswertungsperiode weit unter RWmin liegender Abfluss nicht, dass dieser den Schwellenwert RWmin über den ausgewerteten Zeitraum nie überschreiten könnte, wie das Beispiel des Gebietes 361 (Abb. 30) zeigt. Hier steigt im August 2003 der Abfluss für kurze Zeit über den Grenzwert RWmin an. Abb. 30 ist im Übrigen ein weiteres typisches Beispiel dafür, dass es die Simulationsfähigkeit von HBV insbesondere im Sommer übersteigt, die feinen Schwankungen im Niedrigwasserbereich nachzuzeichnen.

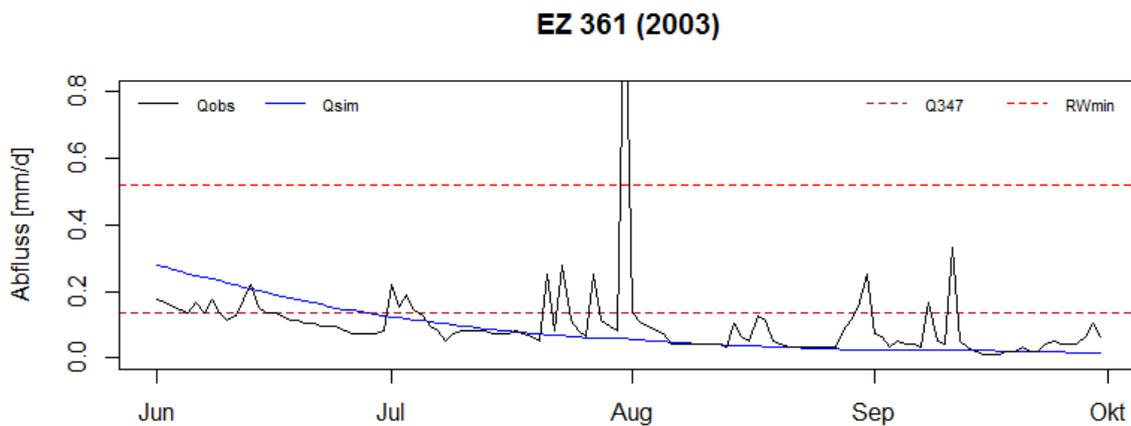


Abb. 30: Gebiet 361 in der Auswertungsperiode 2003. RWmin liegt allgemein deutlich über dem Abfluss, wird aber von einer Abflussspitze im August überschritten.

Minimalabfluss [mm/d] der Trockenheit

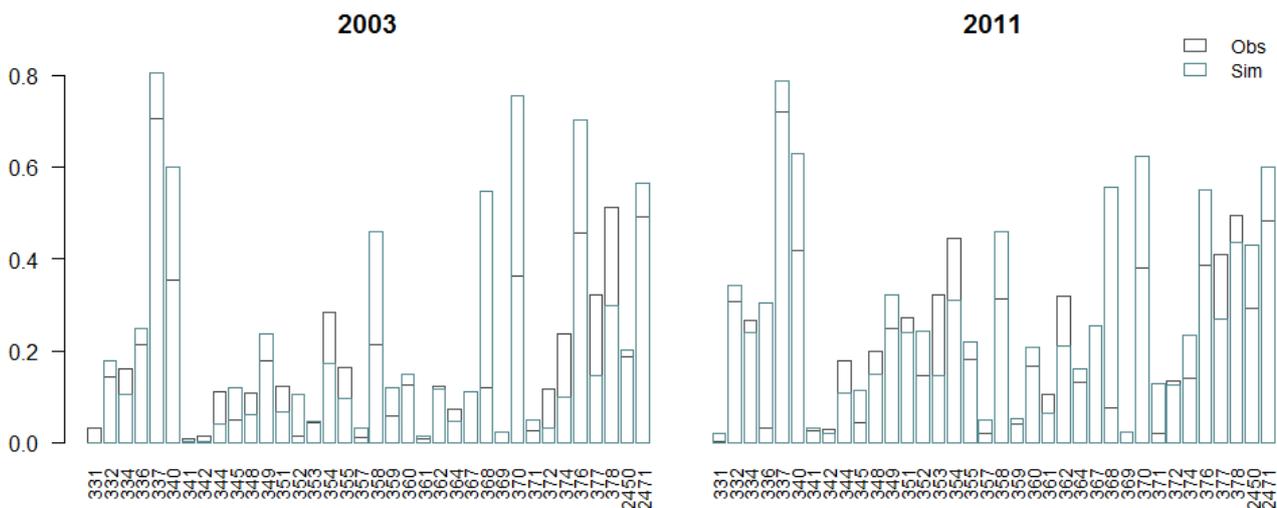


Abb. 31: Vergleich der Simulationen mit den beobachteten Werten für den **minimalen Abfluss** in den beiden Auswertungsperioden 2003 (links) und 2011 (rechts).

6.1.2.5. Minimum

Der letzte betrachtete Kennwert liefert schliesslich besonders erfreuliche Resultate (siehe Abb. 31). Sowohl im 2003 wie im 2011 werden die minimalen Abflusswerte zwar tendenziell überschätzt, die Abweichungen sind im Schnitt aber nur sehr gering: Im Sommer 2003 wurde der Abfluss um durchschnittlich 0.03 mm/d und im 2011 um 0.05 mm/d überschätzt.

Grosse Überschätzungen des Minimalwertes werden durch diejenigen Simulationen erzeugt, die versagt haben (*Missed Events*). Die einzige Simulation, die im 2003 deutlich nach unten ausschlägt, ist bereits in Abb. 25 dargestellt; im 2011 ist es die in Abb. 21 (oben) dargestellte Simulation des Einzugsgebietes 353, die am stärksten nach unten ausschlägt.

Im 2003 und im 2011 wird die grösste Überschätzung durch das Einzugsgebiet 368 erzeugt, das sehr schlechte Simulationen in Form eines über die gesamte Auswertungsperiode hinweg konstanten Abflusswertes lieferte. In Hinblick auf das bereits schlechte Kalibrierungsergebnis dieses Gebietes verwundert dieses Resultat allerdings nicht sonderlich.

6.1.2.6. Weitere Beobachtungen

Bei diesen Auswertungen konnte allgemein beobachtet werden, dass so gut wie alle Einzugsgebiete einen deutlichen Unterschied zwischen ihrer Simulationsgüte im 2003 und im 2011 zeigen. Als Beispiel ist in Abb. 32 das Einzugsgebiet 332 dargestellt (kombiniertes Gütemass = 0.79). Im Jahr 2003 ist die Simulation im Juni und Juli systematisch zu hoch, im August nähert sie sich den beobachteten Werten an und ab September stimmt das Niveau der Simulation sehr gut, nur die feinen Schwankungen werden nicht abgebildet. Im Jahr 2011 hingegen ist die Simulation ausserordentlich gut, nur gerade im April und Mai ist ein leichter Versatz nach oben zu erkennen.

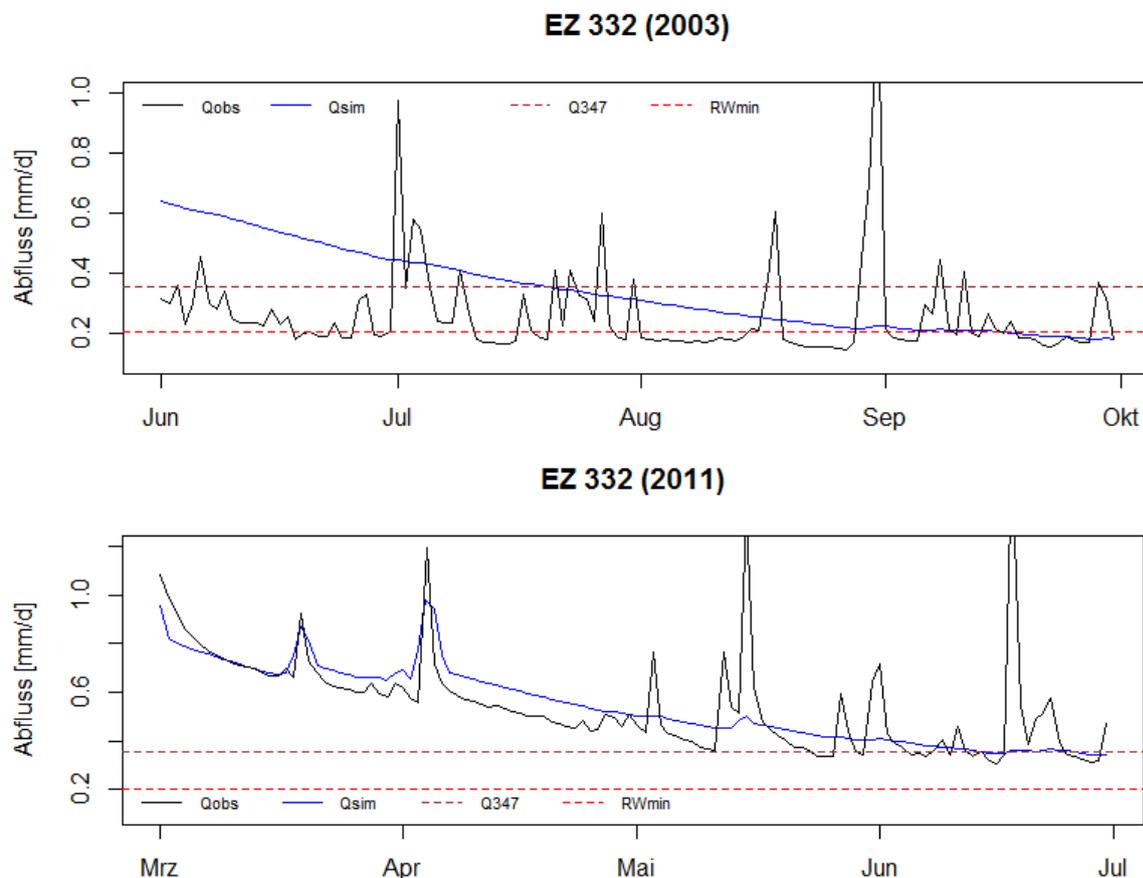


Abb. 32: Simulationen des Gebietes 332 für die Auswertungsperioden 2003 (oben) und 2011 (unten). Die beiden Grafiken zeigen deutlich, wie unterschiedlich sich die Simulation in den beiden Perioden verhält.

Interessant ist, dass man aufgrund der visuellen Betrachtung die Simulationen im 2011 tendenziell besser einstufen würde als die Simulationen im 2003. Ein Blick auf Tab. 17 bis Tab. 20 im Anhang suggeriert aber eher das Gegenteil.

Ausserdem konnte beobachtet werden, dass ein gutes Kalibrierungsergebnis nicht unbedingt eine gute Simulation der beiden Auswertungsperioden sicherstellt. So versagte die Vorhersage des Gebietes 340, das eine Modellgüte von 0.785 erreicht hatte, in beiden Jahren komplett (im Sommer 2003 war die Simulation systematisch leicht zu hoch, im Frühjahr 2011 war der Fehler zwar gering, aber die Simulation unterschritt den Schwellenwert im Gegensatz zum Messwert knapp nicht). Auch der soeben erläuterte Fall des Gebietes 332 (Abb. 32) zeigt im 2011 ein ähnliches Verhalten, wobei die Simulation bei diesem Gebiet aber keinen *Missed Event* erzeugte.

Interessant ist ein Blick auf die Gebiete 374 und 378, welche die besten Kalibrierungsergebnisse erreichten: Die Vorhersage für 378 im Jahr 2003 (Abb. 25) zeigt ein zu starkes Absacken der Simulation; für 2011 ist die Simulation deutlich besser, fällt aber ebenfalls leicht zu tief, sinkt dabei knapp unter den *RWmin*-Schwellenwert und löst so einen *False Alarm* aus. Beim Gebiet 374 (Abb. 33) sinkt die Simulation ebenfalls im Sommer zu stark ab, während die Simulation im 2011 die beobachteten Werte nahezu perfekt trifft.

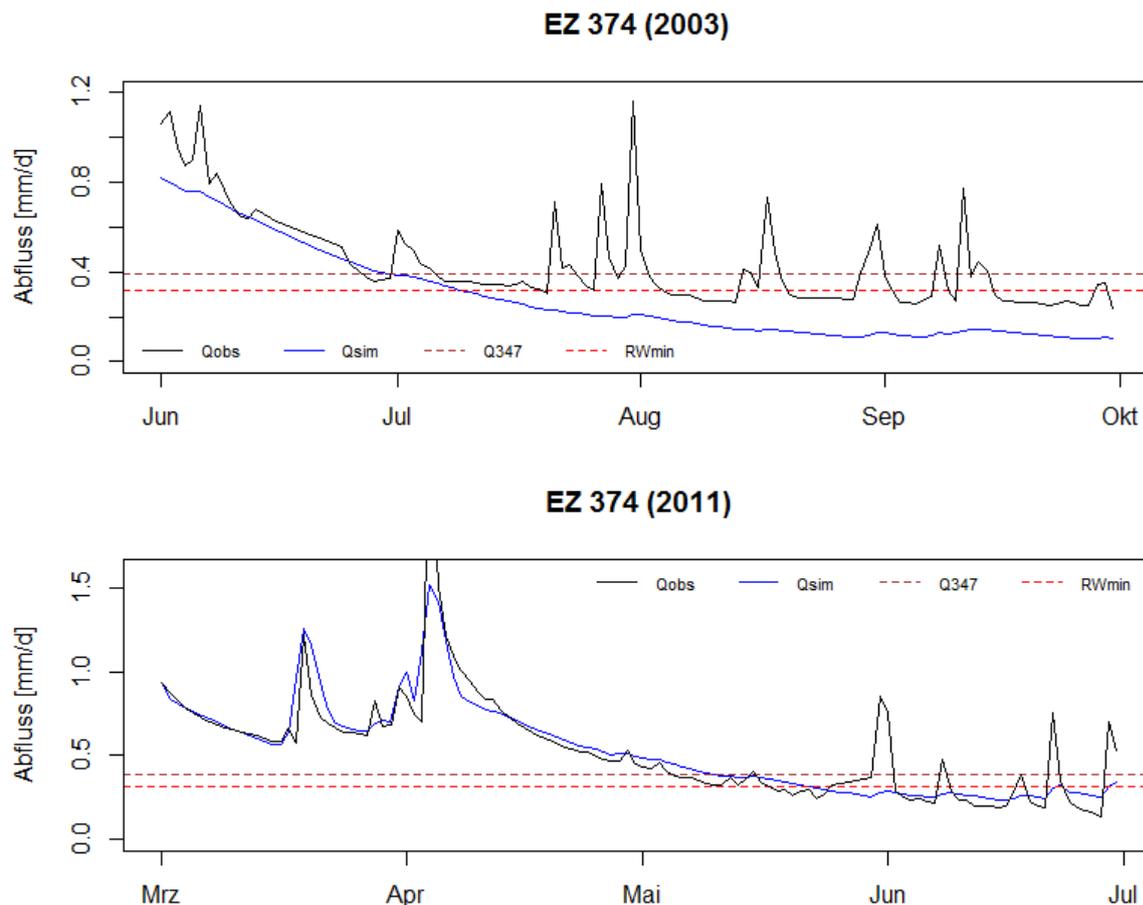


Abb. 33: Simulationen des Gebietes 374 für die Auswertungsperioden 2003 (oben) und 2011 (unten). Die Simulation 2003 sinkt zu stark ab, während im 2011 die beobachteten Werte nahezu perfekt getroffen werden.

6.1.3. Kontingenz-Tabellen

Die beiden Auswertungsperioden im 2003 und 2011 wurden zusätzlich in Kontingenz-Tabellen ausgewertet. Die Werte für die Periode im Jahr 2003 und die Schwellenwerte Q_{347} und RWmin sind in Tab. 9 zusammengestellt.

Tab. 9: Kontingenztabellen für die Auswertungsperiode im 2003. E= Event, NE = No Event.

2003 Q347		Obs	
		E	NE
Sim	E	27	0
	NE	7	2

POD = 79.4 %

FAR = 0 %

FOM = 20.6 %

2003 RWmin		Obs	
		E	NE
Sim	E	27	2
	NE	4	3

POD = 87.1 %

FAR = 6.9 %

FOM = 12.9 %

Für die Auswertungsperiode 2011 finden sich analog die Werte in Tab. 10.

Tab. 10: Kontingenztabellen für die Auswertungsperiode im 2011. E= Event, NE = No Event.

2011 Q347		Obs	
		E	NE
Sim	E	25	3
	NE	8	0

POD = 75.8 %

FAR = 10.7 %

FOM = 24.2 %

2011 RWmin		Obs	
		E	NE
Sim	E	23	2
	NE	3	8

POD = 88.5 %

FAR = 8 %

FOM = 11.6 %

In allen vier Tabellen fällt zunächst positiv auf, dass die POD hoch und die FAR hingegen tief liegt. Die FOM befindet sich auf nicht übermässig hohem Niveau, dennoch wäre hier eine Modellverbesserung erstrebenswert.

6.1.4. Zusammenhang zwischen Modellgüte, Karstanteil und Einzugsgebietsgrösse

Unter den Ergebnissen der Kalibrierung interessierte besonders, ob ein Zusammenhang der erreichten Modellgüte einerseits mit dem Karstanteil, andererseits mit der Grösse der Einzugsgebiete besteht. Die Exploration der Daten zeigte deutlich, dass sehr schlechte Ergebnisse nur für Einzugsgebiete erreicht wurden, die sowohl klein ($< 30 \text{ km}^2$) sind als auch zumindest einen gewissen Anteil an verkarsteten Gebieten aufweisen. Sowohl der Karstanteil wie die Einzugsgebietsgrösse alleine reichten nicht zur Erklärung schlechter Kalibrierungsergebnisse aus: Verschiedene Gebiete erreichten auch mit einem hohen Karstanteil gute Ergebnisse für das kombinierte Gütemass (siehe Abb. 34), ebenso wie auch kleine Gebiete ein gutes Ergebnis erzielen konnten (siehe Abb. 35). So erreichte beispielsweise das Einzugsgebiet 377 mit einer Fläche von 6.7 km^2 eine Güte von 0.796. Allerdings kann in beiden Fällen der Umkehrschluss gezogen werden: Erstens erreichten grössere Einzugsgebiete durchwegs recht bis sehr gute Ergebnisse; unter den Gebieten mit einem Gütemass kleiner als 0.7 finden sich ein grosses (Nr. 348 mit 120.8 km^2) sowie ein mittleres (Nr. 360 mit 25.6 km^2) Einzugsgebiet, die Fläche aller übrigen ist kleiner als 20 km^2 . Zweitens erreichten Gebiete ohne Karstanteil zwar teilweise nur sehr mässig gute Ergebnisse, aber niemals sehr schlechte (< 0.5).

Daraus kann also der Schluss gezogen werden, dass Karst ein Hindernis für eine erfolgreiche Kalibrierung sein kann, es aber nicht sein muss. Aus diesem Grund wird in Kap. 6.3 ein spezielles Augenmerk auf diese Gebiete gelegt.

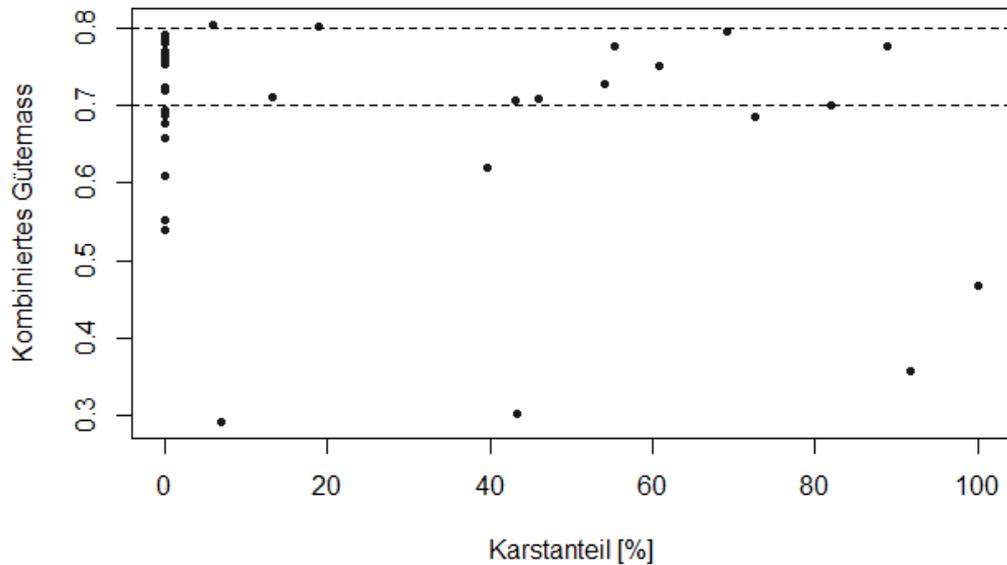


Abb. 34: Erreichtes kombiniertes Gütemass im Vergleich mit dem Anteil an Karstflächen im Einzugsgebiet.

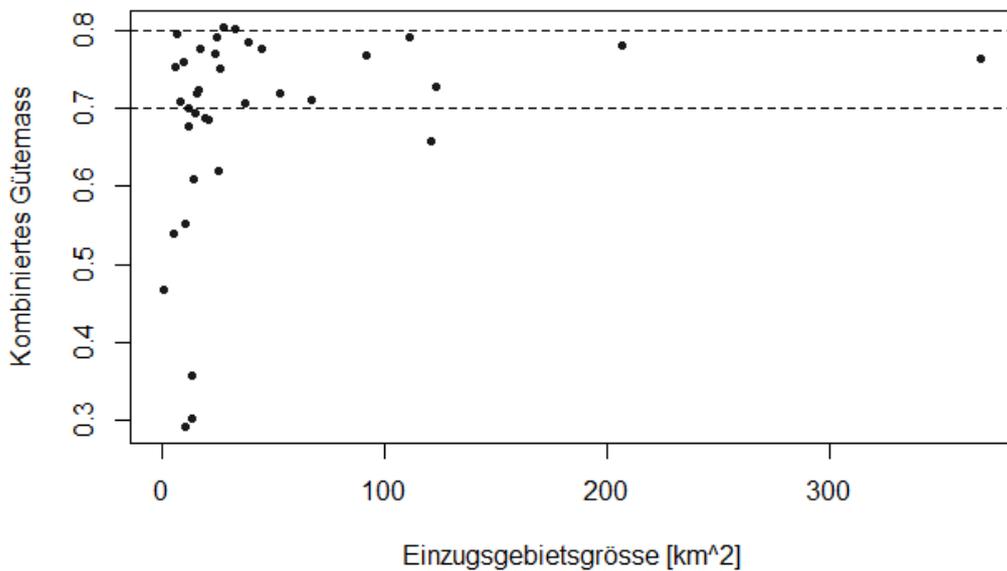


Abb. 35: Erreichtes kombiniertes Gütemass im Vergleich mit der Grösse des Einzugsgebietes.

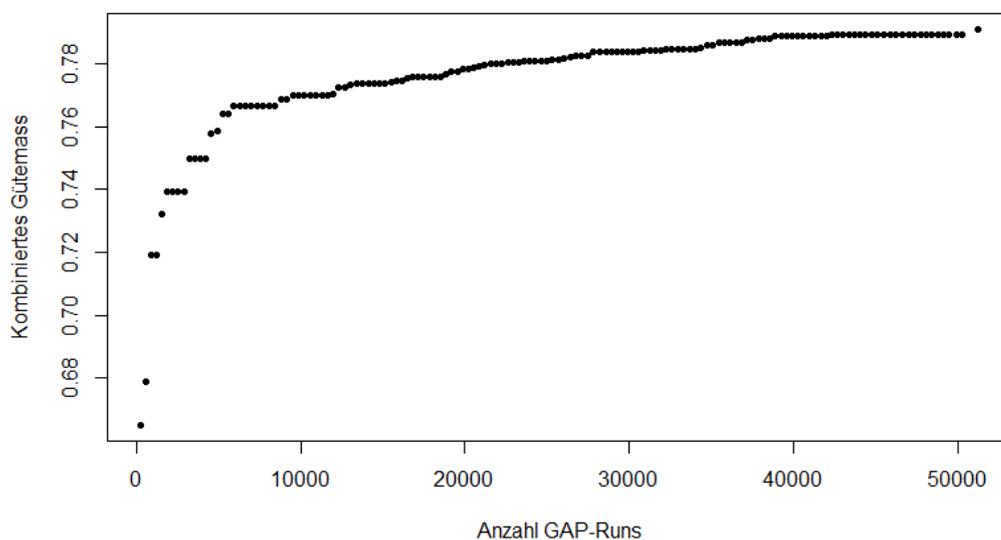
6.1.5. Modellgüte im Verlauf der GAP-Berechnung

Um einen Eindruck über die Modellgüte-Entwicklung im Verlauf der GAP-Berechnung zu erhalten, wurde die Modellgüte (kombiniertes Gütemass gemäss Tab. 5) der Nummer der GAP-Durchläufe (engl. *runs*) gegenübergestellt. Eine solche Grafik kann leicht aus der Output-Textdatei der GAP

abgeleitet werden. In dieser Datei hält HBV die Entwicklung der Berechnung fest, indem für jede berechnete „Generation“ das beste Zwischenresultat festgehalten wird.

Abb. 36 zeigt diesen Zusammenhang am Beispiel der Einzugsgebiete 332 und 352. Diese beiden Gebiete sind exemplarisch für unterschiedliche Entwicklungen der Modellgüte: Bei Gebiet 332 flacht die Kurve bei 40'000 Runs nahezu komplett ab, Gebiet 352 zeigt hingegen bis zum Ende der Berechnung bei 50'000 Runs eine Zunahme der Modellgüte. Im Vergleich zu den übrigen Einzugsgebieten ist aber eher die Entwicklung von Gebiet 332 typisch, da die meisten Gebiete ab etwa 20'000 Runs nur noch eine moderate und visuell oft kaum mehr erkennbare Gütesteigerung zeigen. Als deutlich abgesetzter Punkt fällt allerdings in beiden Grafiken der letzte Berechnungsschritt auf: Dieser steht für die Powell-Optimierung (siehe Kap. 5.5.1). Dass dieser letzte Punkt sich markant vom übrigen Verlauf abhebt, unterstreicht die Bedeutung der Powell-Optimierung für die Feinabstimmung, wie sie bereits von Wang (1991) vorgeschlagen wurde.

Anstieg der Modellgüte mit Anzahl GAP-Runs (EZ 332)



Anstieg der Modellgüte mit Anzahl GAP-Runs (EZ 352)

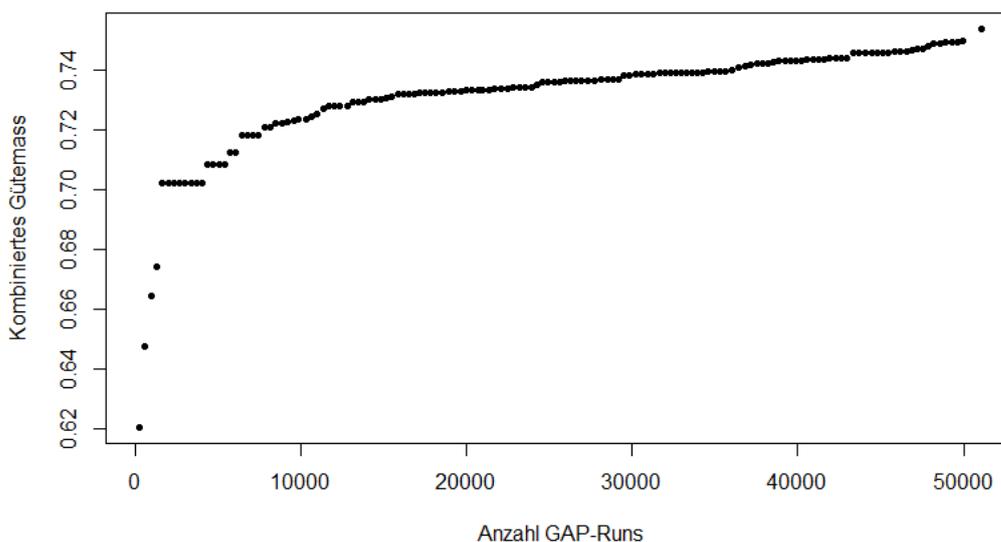


Abb. 36: Entwicklung der Modellgüte mit dem Verlauf der GAP-Berechnung für die Einzugsgebiete 332 (oben) und 352 (unten).

6.2. Sensitivitäts-Analyse (Standard-Modell)

6.2.1. Analyse 1: Vergleich von 10 Kalibrierungen

Die Ergebnisse von Analyse 1 (Variations-Koeffizient für 10 Kalibrierungen) sind in Abb. 37 und Abb. 38 dargestellt. Abb. 37 lässt gut erkennen, welche Parameter am stärksten streuen: Dies sind PCALT sowie die Parameter der Schnee-Routine, insbesondere CFR und CWH. Da PCALT die Schnee-Routine massgeblich mitbeeinflusst, wird deutlich, dass die Schneeberechnung offensichtlich den unsichersten Teil des Modells darstellt. Ausserdem zeigen K2, Cet und LP zwar einzelne Ausreisser (in einzelnen Einzugsgebieten sind diese Parameter also schlecht eingrenzbar), aber ansonsten relativ einheitliche Werte. Die übrigen Parameter zeigen im Vergleich nur eine geringe Streuung und scheinen damit durch das Modell recht gut definierbar zu sein. Am besten eingrenzbar ist der Wertebereich von MAXBAS, der so gut wie keine Streuung aufweist.

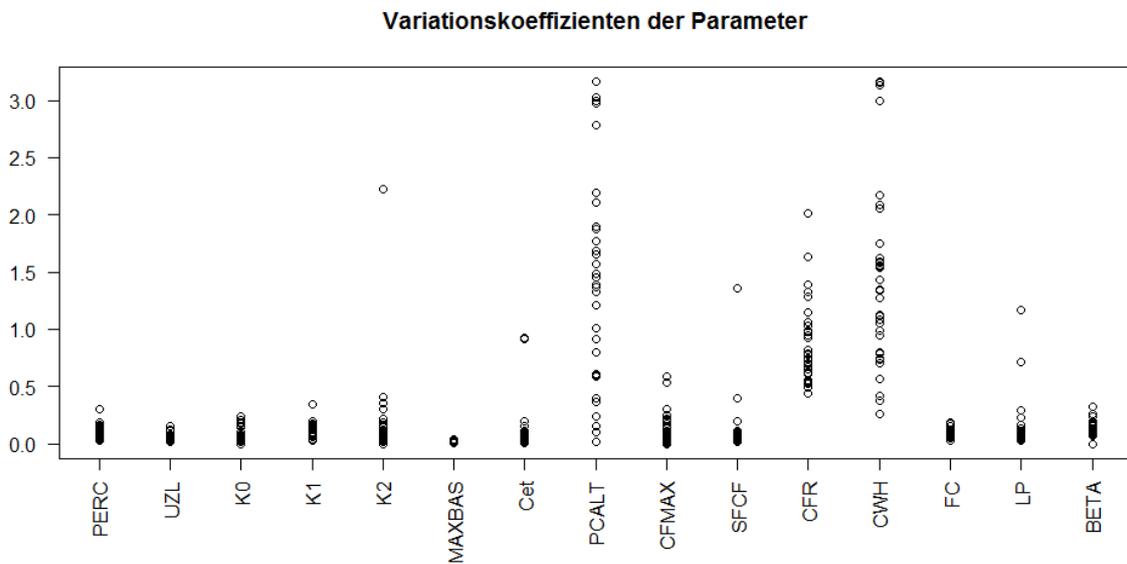


Abb. 37: Parameterstreuung für 10 Kalibrierungen (Analyse 1). Jeder Punkt repräsentiert den Variationskoeffizienten eines Einzugsgebietes.

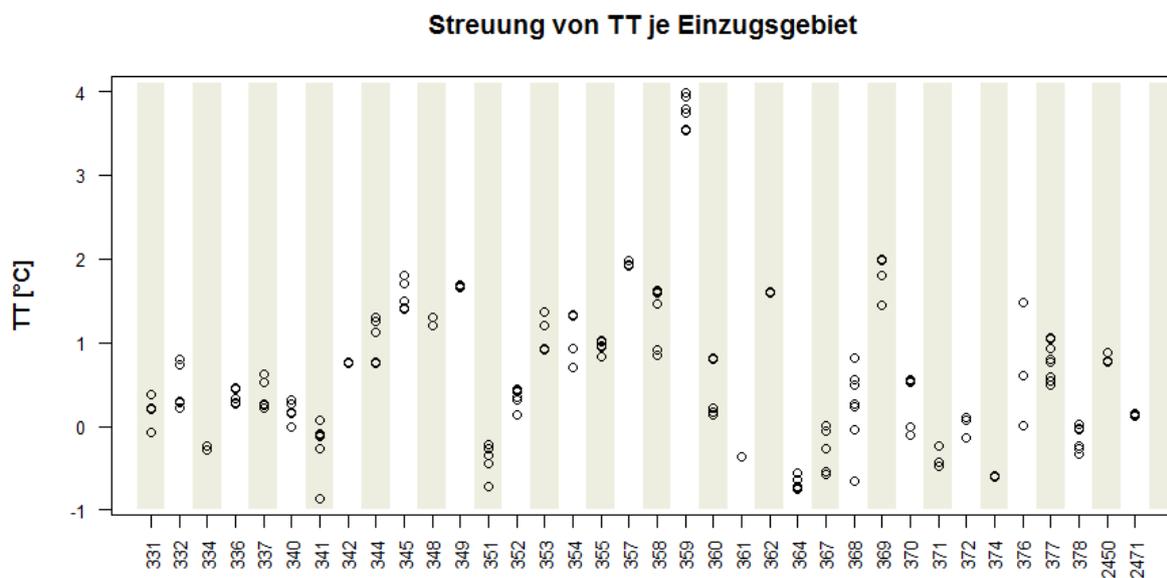


Abb. 38: Streuung der TT-Werte für 10 Kalibrierungen (Analyse1) pro Einzugsgebiet.

Bei TT (Abb. 38) fällt als erstes auf, dass das Gebiet 359 über alle Kalibrierungsläufe hinweg sehr hohe TT-Werte zwischen 3.5°C und 4°C aufweist. Ansonsten schwanken die Werte zwischen -1°C und 2°C. Einige Gebiete (334, 342, 349, 361, 362, 374 und 2471) zeigen einen sehr stark eingegrenzten, d.h. gut definierbaren Wert für TT. Die grösste Streuung weisen hingegen die Einzugsgebiete 368 und 376 auf (beides Gebiete, die relativ schlechte Modellgüte-Ergebnisse in der Kalibrierung erzielten).

6.2.2. Analyse 2: Vergleich mit Gütemass-Variation

Im Vergleich hierzu sind in Abb. 39 und Abb. 40 die Ergebnisse der Analyse 2 dargestellt. Merke, dass Abb. 39 nicht dieselbe y-Skala wie Abb. 37 aufweist. Die für den Variationskoeffizienten maximal erreichten Werte (Parameter PCALT und CWH) sind also geringer. Ja nach Blickwinkel konnte dies erwartet werden oder erstaunte dies: Da hier nur fünf verschiedene Kalibrierungen einander gegenüber gestellt sind, konnte erwartet werden, dass die Streuung geringer ausfällt als bei Analyse 1 mit zehn Kalibrierungen. Andererseits wäre aber mit grösseren Streuungen als in Analyse 1 zu rechnen gewesen, weil hier unterschiedliche Gütemasse verglichen werden, von denen zu erwarten wäre, dass ihre Ergebnisse weiter auseinanderklaffen als eine Analyse anhand desselben Gütemasses. In diesem Sinne ist es ein besonders erfreuliches Resultat, dass die Streuung nicht grösser ausfällt, und spricht für die Robustheit des Modells.

Dennoch ist erkennbar, dass die allgemeine Streuung der Parameter zugenommen hat und nun ausser für MAXBAS der Wert 0.5 von vielen Gebieten überschritten wird. Wie bereits erwähnt, war dies zu erwarten, weil hier unterschiedliche Gütemasse verwendet wurden. Allerdings zeigen sich auch in Analyse 2 die Parameter der Schnee-Routine als die unsichersten; MAXBAS ist dagegen wiederum der bestdefinierte Parameter.

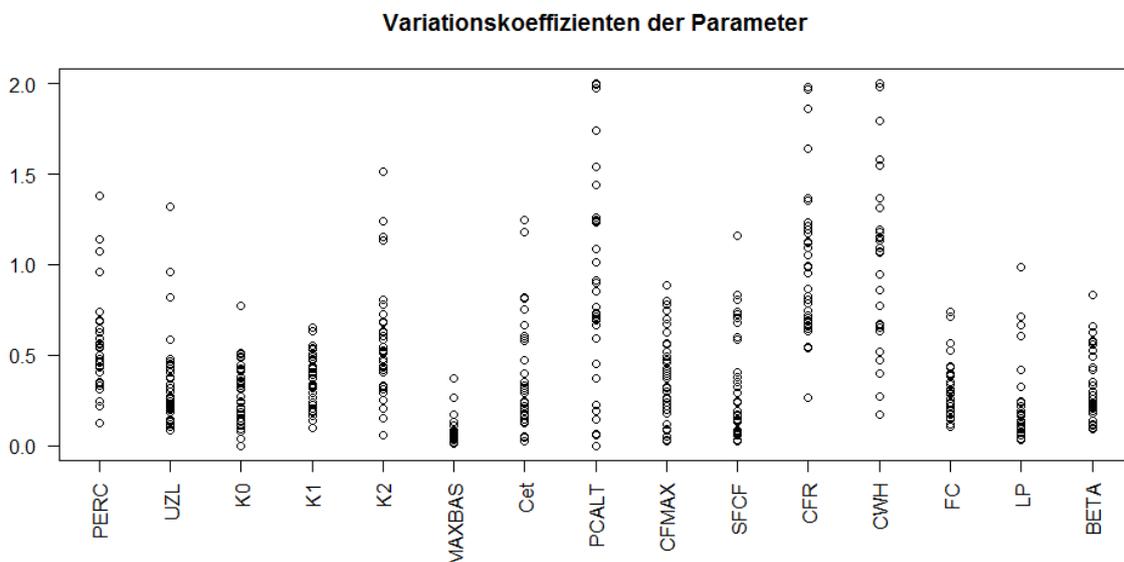


Abb. 39: Parameter-Variation für 5 Kalibrierungen (Analyse 2). Jeder Punkt repräsentiert den Variationskoeffizienten eines Einzugsgebietes.

In Abb. 40 wird ersichtlich, dass die Streuung der TT-Werte im Vergleich zu Analyse 1 zugenommen hat. Die TT-Werte schwanken dabei aber immer noch um in etwa dieselben Mittelwerte. Die Abweichungen haben also zwar etwas zugenommen, die Wertebereiche haben sich aber nicht grundsätzlich verschoben. Dies spricht einmal mehr für die Robustheit des Modells bzw. für eine relativ gute Parameter-Identifizierbarkeit.

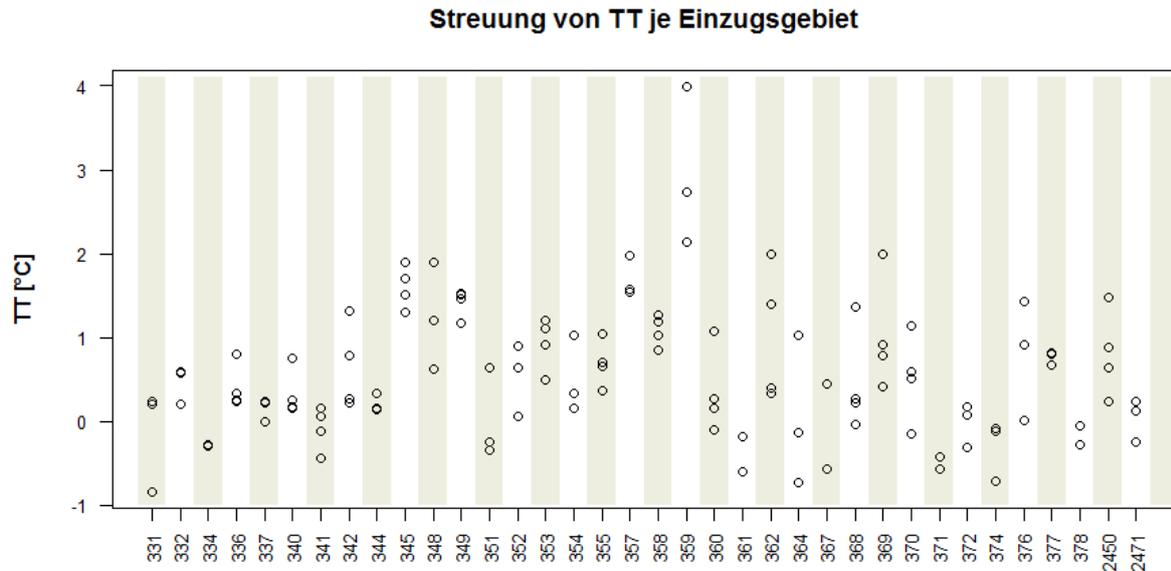


Abb. 40: Streuung der TT-Werte für 5 Kalibrierungen (Analyse2) pro Einzugsgebiet.

6.2.3. Einfluss des Split-Sample-Test-Verfahrens

Gemäss dem Prinzip eines Split-Sample-Tests (Klemes, 1986) wurden für die Kalibrierung die viermonatigen Zeiträume in den Jahren 2003 und 2011, die schliesslich zur Modellauswertung verwendet wurden, entfernt. Da diese Zeiträume auf Trockenheits-Extremereignisse fokussieren, wird damit das Modellvermögen getestet, aussergewöhnliche Ereignisse zu simulieren.

Nun ist es aber von Interesse, zu untersuchen, welchen Einfluss die Extraktion dieser Zeiträume auf das Modell- bzw. Kalibrierungsergebnis hat. Dies wurde beispielhaft am Einzugsgebiet 337 untersucht. Für diesen Vergleich wurden je 100 Kalibrierungen mit und ohne Einbezug der besagten Auswertungsperioden durchgeführt, insgesamt also 200 Kalibrierungen. Wäre eine solche Kalibrierung mit den für die Hauptkalibrierung verwendeten Einstellungen (50'000 GAP-Runs mit 300 Parameter-Sets) durchgeführt worden, wäre der Zeitaufwand zu gross gewesen. Daher wurden die Einstellungen wie für die Sensitivitätsanalyse auf 20'000 GAP-Runs mit 200 Parameter-Sets reduziert. Mit dem schnellsten für diese Arbeit zur Verfügung stehenden Rechner dauerten 100 Kalibrierungen auf diese Weise zirka 50 h. Aus zeitlichen Gründen wurde daher auf weitere Vergleichskalibrierungen verzichtet.

Abb. 41 stellt die Ergebnisse für Januar bis September 2003 und 2011 grafisch dar: In hellblau sind Kalibrierungsergebnisse ohne Einbezug der Auswertungsperioden dargestellt, in dunkelblau solche, die mit den Auswertungsperioden berechnet wurden. Um die Unterschiede zwischen den verschiedenen Simulationen besser hervorzuheben, wurde der Abfluss im Zehnerlogarithmus dargestellt. Erkennbar ist eine leichte, aber systematische Verschiebung nach unten für die Kalibrierungen inklusive der Auswertungsperioden (in dunkelblau). Äusserst bemerkenswert ist dabei allerdings, dass sich der Versatz in beiden Fällen erst ab Mai bemerkbar macht, obwohl im Jahr 2011 der Zeitraum von Anfang März bis Ende Juni entfernt wurde. Der beobachtbare Versatz zieht sich aber über diese Periode hinaus. Offensichtlich tritt also nur im Niedrigstwasserbereich, der in diesem Gebiet im 2011 bis in den Juli hinein andauerte, ein merklicher Unterschied auf – und das auch ausserhalb der nicht-kalibrierten Periode. Das lässt vermuten, dass die leichte Modellverbesserung, die durch einen Miteinbezug des Sommers 2003 in die Kalibrierung erreicht werden könnte, sich auch auf den Sommer 2011 auswirken würde.

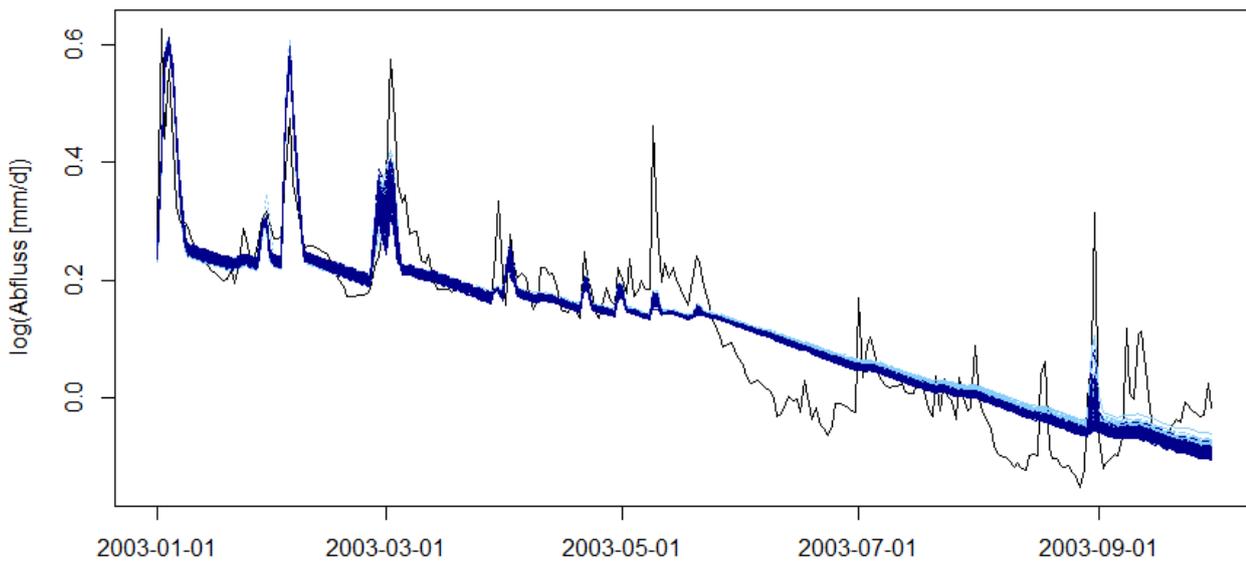
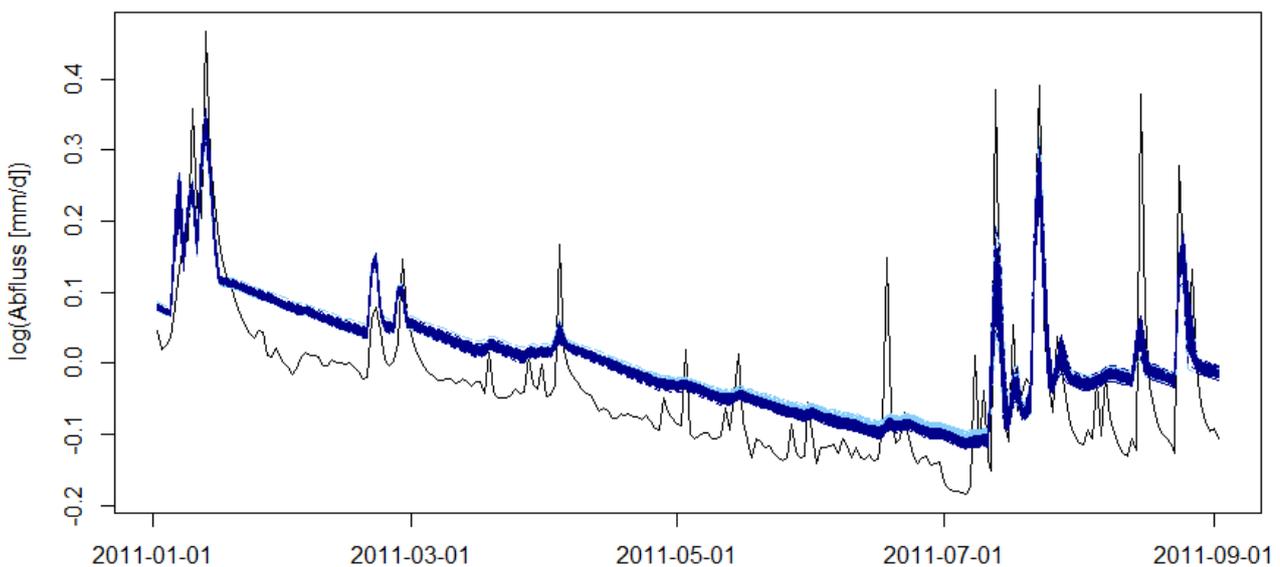
EZ 337 (2003)**EZ 337 (2011)**

Abb. 41: Vergleich zwischen je 100 Kalibrierungen mit (in dunkelblau) und ohne (in hellblau) Einbezug der viermonatigen Auswertungsperioden im 2003 und 2011 (Einzugsgebiet 337) für Januar bis September 2003 (oben) und 2011 (unten) exemplarisch dargestellt. In schwarz ist der gemessene Abfluss abgebildet. Zu beachten ist die logarithmische y-Achse!

Als sehr erfreuliches Resultat zeigt dieser Vergleich aber auch, dass die GAP-Kalibrierung offensichtlich sehr einheitliche Resultate liefert, denn die Streuung innerhalb der 200 Kalibrierungen ist bemerkenswert gering. Dies spricht sowohl für die Robustheit des HBV-Modells als auch für die Eignung der GAP als Kalibrierungsmethode.

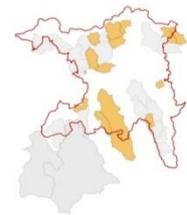
	Mittelwert	Min	Max	Verwendet
Ohne	0.7919	0.7875	0.7972	0.7914
Mit	0.7957	0.7921	0.7999	-

Tab. 11: Gütemasse der je 100 Kalibrierungen mit und ohne Einbezug der besagten Auswertungsperioden.

Dieses positive Ergebnis wird ferner unterstrichen, wenn man die Gütemasse dieser je 100 Kalibrierungen betrachtet (Tab. 11): Das Gütemass schwankt nur um 0.01. Die Ergebnisse zeigen sich also auch numerisch betrachtet als sehr einheitlich bzw. robust.

Aus Abb. 41 lässt sich entnehmen, dass die Simulationen unter Einbezug des gesamten verfügbaren Kalibrierungszeitraums leicht verbessert werden könnten. Das Modellverhalten wird zwar nicht grundsätzlich verändert – der Versatz nach oben im Vergleich zum Messwert bleibt erhalten und feine Schwankungen werden nicht besser abgebildet –, aber das Niveau der Vorhersage könnte leicht verbessert werden. Dies spiegelt sich auch im durchschnittlich etwas verbesserten Wert der Modellgüte wider: Die 100 Kalibrierungen ohne die Auswertungsperioden erreichten im Durchschnitt eine Güte von 0.7919, währendem die 100 Kalibrierungen unter Einbezug derselben einen Schnitt von 0.7957 erreichten. Diese Werte wie auch Abb. 41 zeigen aber auch, dass die dadurch erreichbare Verbesserung der Simulationen nur gering wäre.

6.3. Alternative Modellstrukturen für Karst-Gebiete



Für die 7 karsthaltigen Gebiete, die zusätzlich mit der Delay-Variante von HBV kalibriert wurden, hat sich im Laufe der Modellierungen bestätigt, dass die Delay-Variante ein sich deutlich unterscheidendes Abflussverhalten erzeugt, das sich in den untersuchten Gebieten mehrheitlich als passender erwiesen hat. Die meisten Gebiete zeigten zwar nur eine geringe Verbesserung oder gar eine leichte Verschlechterung der Modellgüte, bei einem grafischen Vergleich der beiden Simulationen zeigte die Delay-Variante aber jeweils das realitätsnähere Verhalten. Abb. 42 illustriert ein typisches Beispiel, wie das Rezessionsverhalten einen natürlicher anmutenden Verlauf annimmt und das Minimalniveau besser getroffen wird. Die Ausnahme bildet das Einzugsgebiet Nr. 359, für das mit der Standardversion die geringfügig besseren (aber dennoch unbefriedigenden) Ergebnisse erzielt wurden.

Aufgrund dieser Beobachtung wurden schliesslich auch weitere, nicht-verkarstete Einzugsgebiete mit

Tab. 12: Ergebnisse der Kalibrierungen mit der HBV-Delay-Variante im Vergleich zu den Standard-Simulationen.

EZ-Nr.	Modellgüte Standard	Modellgüte Delay
Echte Karstgebiete		
341	0.684	0.699
345	0.466	0.426
357	0.356	0.326
359	0.291	0.144
360	0.620	0.629
362	0.700	0.737
369	0.302	0.166
Keine Karstgebiete		
336	0.552	0.578
348	0.658	0.66
349	0.686	0.722
353	0.676	0.666
368	0.538	0.639
370	0.693	0.750
376	0.608	0.723

unbefriedigendem Kalibrierungs-Ergebnis zusätzlich mit der Delay-Variante kalibriert. Bei diesen untersuchten Gebieten konnte für alle eine leichte bis markante Verbesserung festgestellt werden. Insgesamt wurden 14 Gebiete zusätzlich mit der HBV-Delay-Variante kalibriert. Die Ergebnisse dieser Kalibrierungen sind in Tab. 12 zusammen- und den Resultaten der Standardversion gegenübergestellt.

Erstaunlich ist, dass diese Nicht-Karstgebiete die grösseren Verbesserungen als die echten Karstgebiete zeigten, also stärker auf diese Modellvariante ansprachen.

Für einzelne Gebiete wurden zusätzlich die „One GW Box“-Version und die „Three GW Boxes“-Version getestet. Häufig resultierten diese Modell-Varianten zwar in einem sich nochmals deutlich unterscheidenden Abflussverhalten, beide Modelle überzeugten aber nicht im Vergleich zur Delay-Variante. Es konnte zwar vorkommen, dass die Niedrigststände besser getroffen wurden, dann lagen aber gleichzeitig jeweils markante Fehler bei den höheren Abflüssen vor. In einem solchen Fall simuliert das Modell

insgesamt gesehen ein hydrologisch nicht sinnvolles Verhalten und muss daher als weniger vertrauenswürdig angesehen werden. Aus diesem Grund wurden diese beiden Varianten im weiteren Vorgehen nicht mehr berücksichtigt. Beispielhaft ist in Abb. 42 zusätzlich zur Delay- und der Standard-Simulation auch der Abfluss dargestellt, der durch die Modellvariante „Three GW-Boxes“ berechnet wurde.

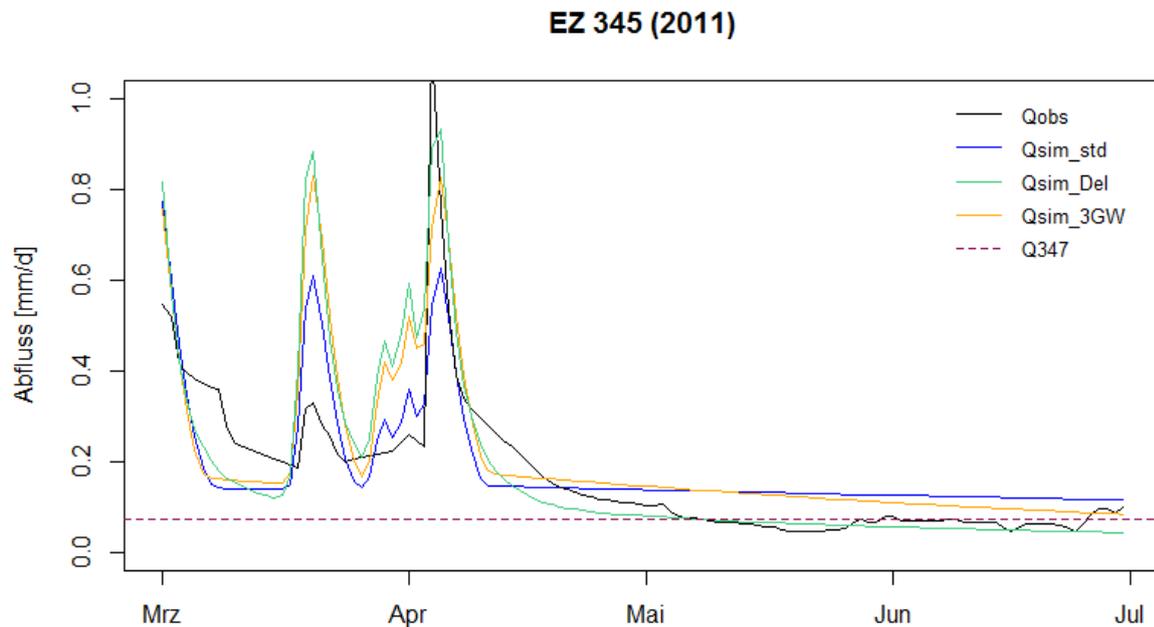


Abb. 42: Einzugsgebiet 345 in der Auswertungsperiode 2011. Qobs = gemessener Abfluss, Qsim_std = Standard-HBV-Modell simulierten Abfluss, Qsim_Del für die Simulation durch die Delay-Variante und Qsim_3GW für die Simulation durch die Variante mit drei Grundwasserboxen („Three GW Boxes“).

In Abb. 43 bis Abb. 47 sind die Ergebnisse der Schwellenwert-Auswertungen für die Delay-Variante grafisch ausgewertet, die detaillierten Werte sind im Anhang in Tab. 21 bis Tab. 25 zusammengestellt; die Ergebnisse der Standardversion wurden jeweils als Vergleichswert ergänzt. In den Tabellen sind *Missed Events* und *False Alarms* wiederum farblich hervorgehoben. Weil die Verbesserungen bzw. Verschlechterungen kaum in aussagekräftige Zahlen gefasst werden können, wurden die Resultate zudem qualitativ bewertet. Dabei bezeichnet \equiv eine gleich gute Simulation (wobei eine Überschätzung ebenso gut wie eine Unterschätzung um denselben Betrag ist), $+$ steht für eine leichte bis deutlich Verbesserung (der simulierte Wert hat sich dem beobachteten angenähert) und $-$ für eine leichte bis deutliche Verschlechterung. Wenn ein vormalig übergangenes Ereignis (*Missed Event*) neu erkannt wurde, wurde dies mit $++$ bewertet (ausser wenn die neue Simulation sehr schlecht ausfiel), ebenso wenn ein vormalig *False Alarm* nun nicht mehr ausgelöst oder eine sehr markante Modellverbesserung im Bereich von mindestens einer Grössenordnung erzielt wurde. Eine Bewertung von $-$ wurde hingegen erteilt, wenn ein vormalig erkanntes Ereignis neu einen *Missed Event* auslöste. Sehr markante Verschlechterungen im Bereich mindestens einer Grössenordnung im Vergleich zur Standardsimulation wurden nie erzeugt.

Bei einer Betrachtung dieser Bewertungen in Tab. 21 bis Tab. 25 fällt auf, dass die Resultate sich stark zwischen den beiden Auswertungsperioden und Schwellenwerten unterscheiden. In der Auswertungsperiode 2003 überwiegen für die Mindestrestwassermenge (RWmin) die negativen Entwicklungen, hingegen schneidet die Delay-Simulation in der Auswertungsperiode 2011 für den Schwellenwert Q_{347} deutlich besser ab. Weiter unterscheiden sich die Ergebnisse für die verschiedenen Kennwerte. Beispielsweise zeigt das Gebiet 353 im 2003 für RWmin eine leichte

Dauer [d] der Trockenheit

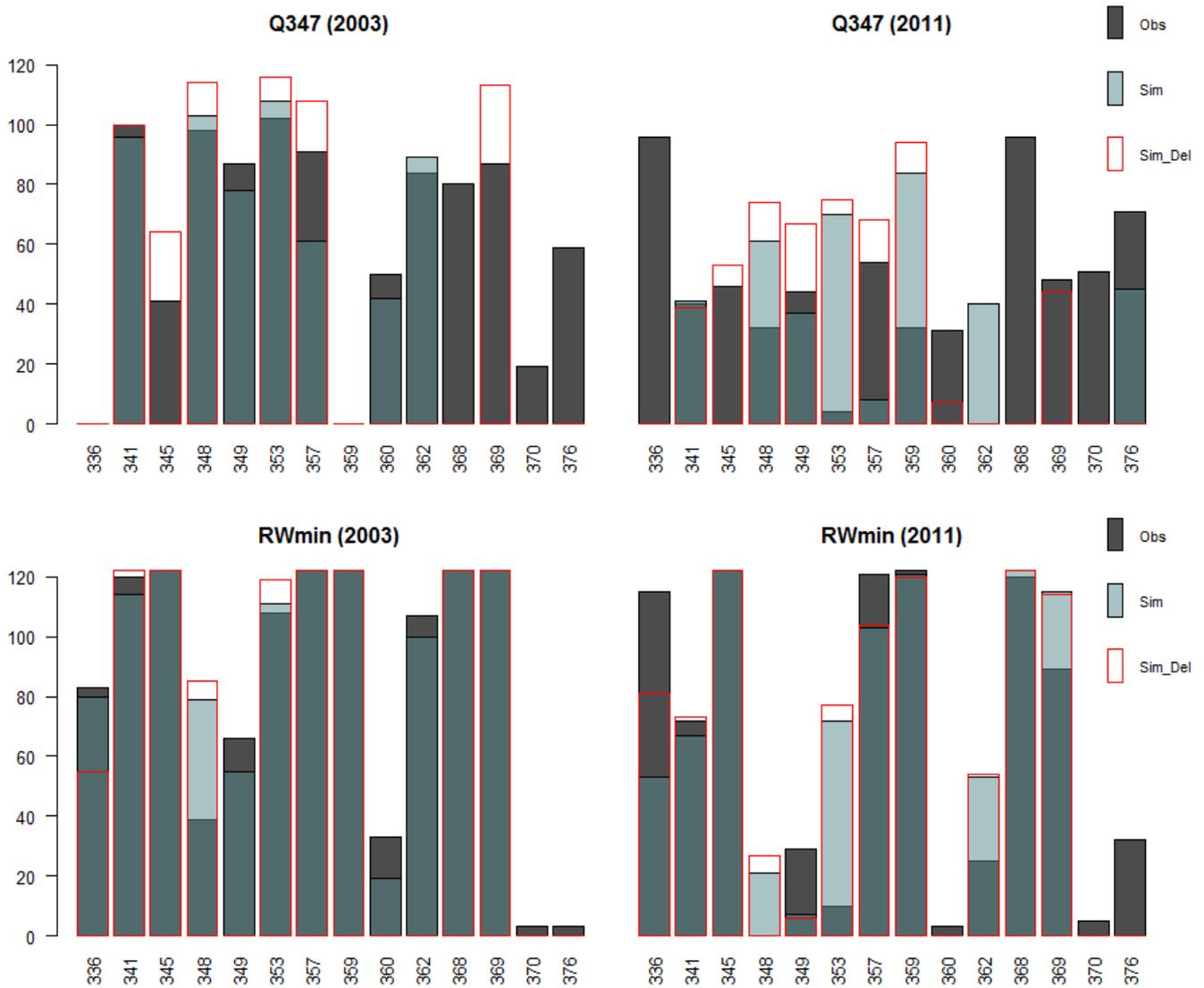


Abb. 43: Vergleich der Simulationen des **Delay-Modells** (Sim_Del) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für die Trockenheits-Dauer im 2003 (links) und 2011 (rechts) für die beiden Schwellenwerte Q_{347} (oben) und RW_{min} (unten).

Verschlechterung in der Vorhersage der Dauer; für Defizit und Intensität schneidet die Delay-Simulation aber leicht besser und für den Onset deutlich besser ab als das Standardmodell.

Für den Minimalabfluss (Tab. 25) schneidet das Delay-Modell in der Auswertungsperiode 2011 deutlich besser ab als das Standardmodell; im 2003 hingegen halten sich Verbesserungen und Verschlechterungen die Waage.

In Abb. 43 lässt sich für die Dauer der Trockenheit eine klare Tendenz zur Überschätzung durch das Delay-Modell erkennen, insbesondere in der Auswertungsperiode im Jahr 2003.

Beim Defizit (Abb. 44) zeigt sich hingegen ein heterogenes Bild: Für Q_{347} im 2003 sind die Delay-Simulationen im Schnitt leicht besser als die Simulationen des Standardmodells; im 2011 überschätzt das Delay-Modell das Defizit allerdings deutlich stärker als das Standardmodell (*Missed Events* nicht berücksichtigt). Für die beiden Auswertungen an RW_{min} nähert sich das Delay-Modell tendenziell besser den beobachteten Werten an als das Standardmodell. Die durchschnittlich höheren Defizitwerte für den Schwellenwert RW_{min} als für Q_{347} weisen darauf hin, dass hier mehrheitlich sehr kleine Gebiete mit einem entsprechen hohen RW_{min} -Schwellenwert ausgewertet wurden. Aus

Defizit [mm] der Trockenheit

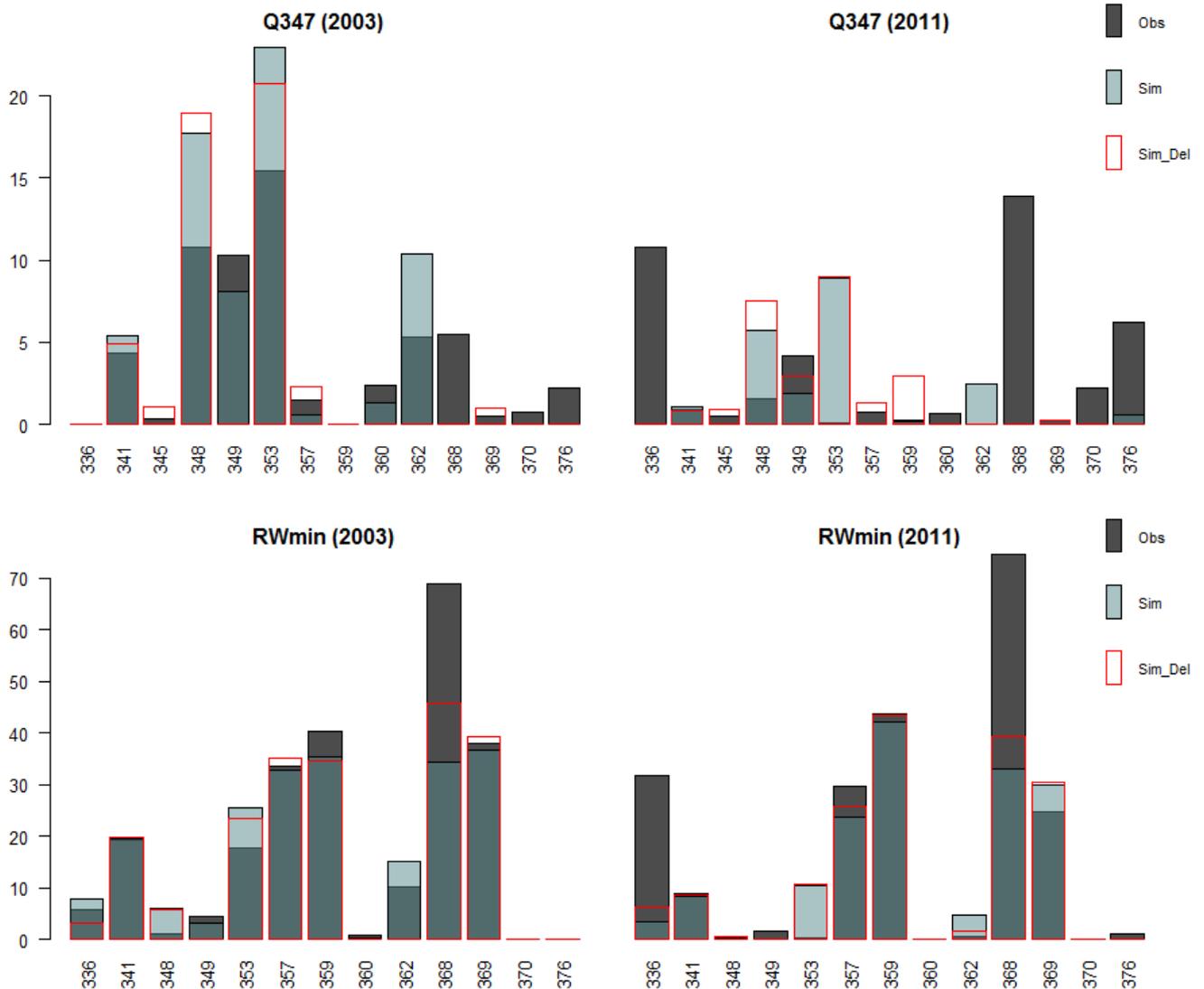


Abb. 44: Vergleich der Simulationen des **Delay-Modells** (Sim_Del) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für das Trockenheits-**Defizit** im 2003 (links) und 2011 (rechts). Achtung: Auf den Grafiken für den Schwellenwert RWmin (unten) fehlt aus Gründen der Darstellbarkeit das Einzugsgebiet 345.

demselben Grund fehlt wiederum das Einzugsgebiet 345 in den Grafiken zu Defizit (Abb. 44) und Intensität (Abb. 45).

Ein Blick auf die grafischen Auswertungen der Intensitäts-Simulationen (Abb. 45) hinterlässt hingegen den deutlichen Eindruck, dass sich das Delay-Modell besser bewährt als das Standardmodell. Dieser Eindruck wird durch die Bewertungen in Tab. 23 bestätigt, bei denen die Verbesserungen deutlich überwiegen. Auch hier weisen die allgemein höheren Werte für RWmin auf die geringe Grösse der meisten Einzugsgebiete hin.

Für den Minimalabfluss (Abb. 46) zeigt sich wiederum ein zweigeteiltes Bild: Im 2003 tendiert das Delay-Modell im Durchschnitt (0.274 mm/d) zu einer stärkeren Überschätzung des beobachteten Wertes (0.152 mm/d) als das Standardmodell (0.202 mm/d). Im 2011 erreicht das Delay-Modell (0.229 mm/d) dagegen die durchschnittlich besseren Werte als das Standardmodell (0.253 mm/d) im Vergleich zu den gemessenen Werten (0.198 mm/d). Zu Abb. 46 ist anzumerken, dass beispielsweise für Gebiet 359 die Werte für die Standard- und die Delay-Simulation identisch sind und daher in der Grafik nicht unterschieden werden können. Alle Werte sind aber in Tab. 25 zu finden.

Für den Onset der Trockenheit (Abb. 47) ergibt sich ein relativ deutliches Muster im Vergleich von Q_{347} und RW_{min} : Für Q_{347} erzeugt das Delay-Modell mehrheitlich Verbesserungen im Vergleich zum Standardmodell – mit Ausnahme von vier neu erzeugten *Missed Events*. Für RW_{min} überwiegen im 2003 hingegen eher die negativen Entwicklungen, im 2011 halten sich Verbesserungen und Verschlechterungen die Waage. Besonders auffällig ist das Gebiet 353, das im 2011 ausschert und die mit Abstand am stärksten verfrühten Onsets über alle Simulationen hinweg zeigt; genauso weist es

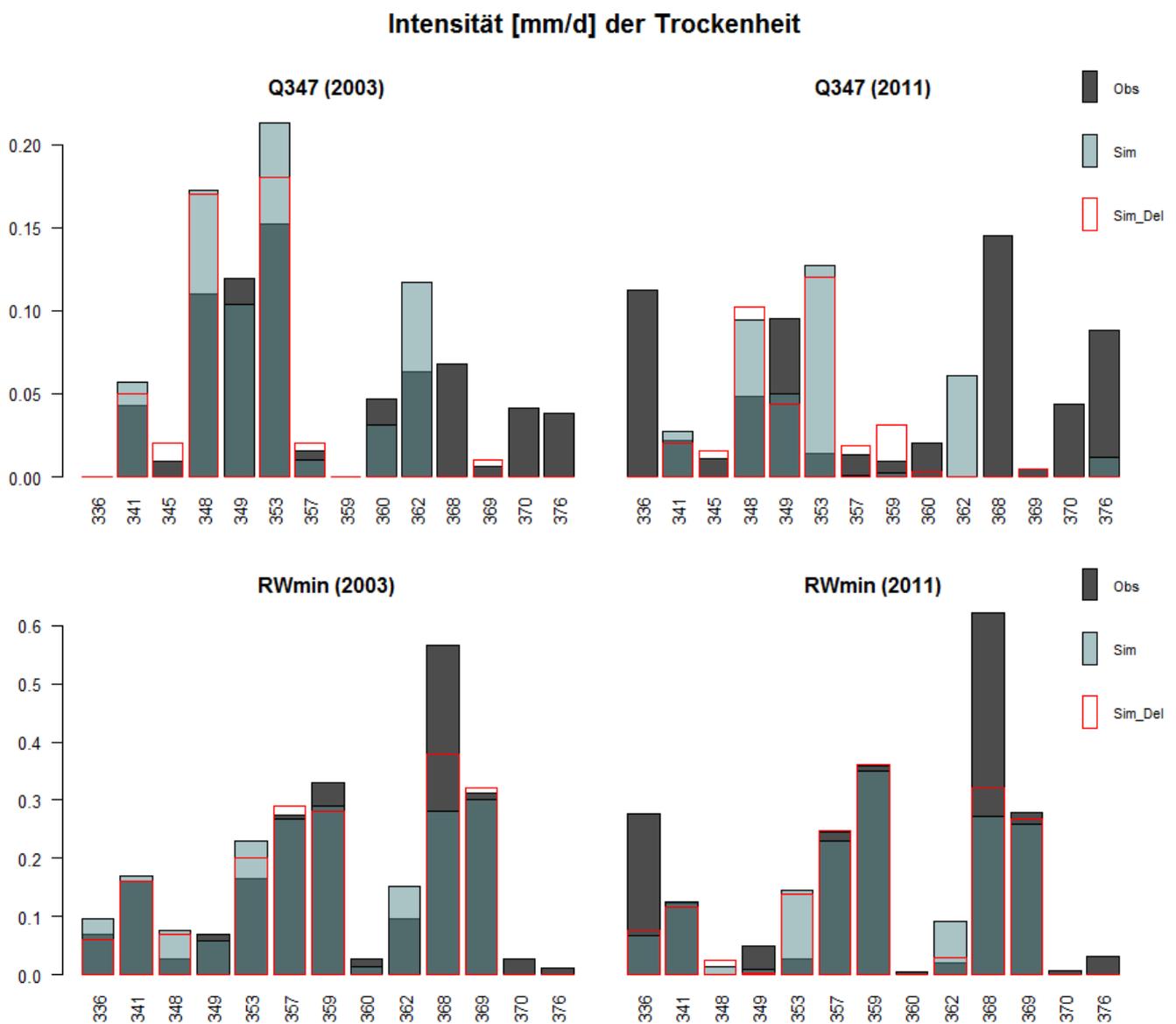


Abb. 45: Vergleich der Simulationen des **Delay-Modells** (Sim_Del) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für die Trockenheits-**Intensität** im 2003 (links) und 2011 (rechts) Achtung: Auf den Grafiken für den Schwellenwert RW_{min} (unten) fehlt aus Gründen der Darstellbarkeit das Einzugsgebiet 345.

im 2003 einen systematischen Shift über alle Auswertungen hinweg auf, zeigt hier aber einen durchwegs verspäteten Onset. Die übrigen Gebiete weisen grundsätzlich grössere Streuungen ihrer Onset-Punkte auf. Zu beachten ist in Abb. 47, dass fehlende Punkte von *Missed Events*, *False Alarms* und *No Events* herrühren – beispielsweise sind bei Gebiet 370 kein Punkte abgebildet, weil sowohl das Standard- wie das Delay-Modell in beiden Zeiträumen und für beide Schwellenwerte mit *Missed Events* aufwarten.

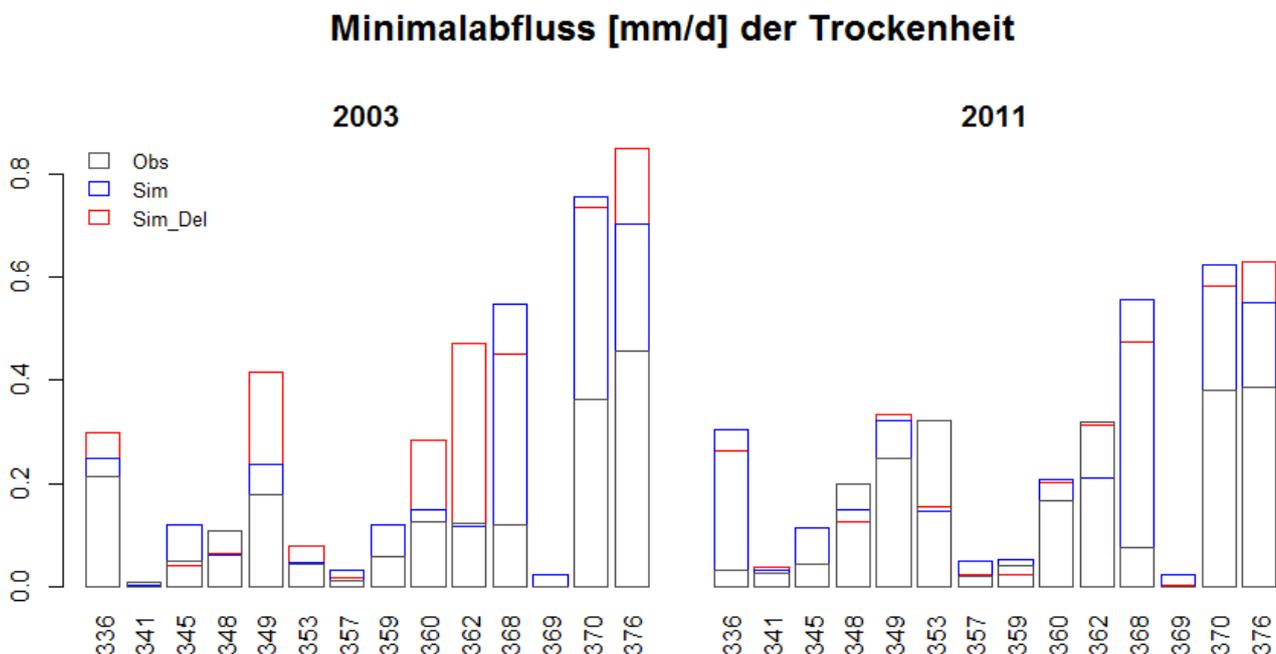


Abb. 46: Vergleich der Simulationen des **Delay-Modells** (Sim_Del) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für den **minimalen Abfluss** in den beiden Auswertungsperioden 2003 (links) und 2011 (rechts).

In Bezug auf die *False Alarms* lässt sich beobachten, dass in den Delay-Simulationen eines von zwei solcher Ereignisse wegfällt. In der Auswertung der Kontingenztabelle (Tab. 13) fällt daher die FAR für Q_{347} im 2011 auf 0%. Bei den *Missed Events* werden über beide Zeiträume und beide Schwellenwerte hinweg fünf *Missed Events* beim Delay-Modell aufgehoben, dafür werden sieben neue *Missed Events* erzeugt (Gebiete 349, 362, 369 im 2003 und 376 im 2011). Die FOM nimmt daher für Q_{347} im 2011 ab (verbessert sich also), steigt aber dafür für beide Schwellenwerte im 2003 (verschlechtert sich). Dieser Unterschied zwischen 2003 und 2011 lässt sich auch für die POD beobachten: Während sie sich für beide Schwellenwerte im 2003 verschlechtert, bleibt sie im 2011 konstant für RWmin und verbessert sich bzw. steigt für Q_{347} .

Tab. 13: Auswertung der Kontingenztabelle des **Delay-Modells** im Vergleich zum Standardmodell. Die Werte des Standardmodells beziehen sich nur auf die Gebiete, für die auch das Delay-Modell verwendet wurde.

		Q347 2003	RWmin 2003	Q347 2011	RWmin 2011
Std-Modell	POD	58.3 %	85.7 %	53.8 %	76.9 %
	FAR	0 %	0 %	12.5 %	9.1 %
	FOM	41.7 %	14.3 %	46.2 %	23.1 %
Delay-Modell	POD	50 %	64.3 %	69.2 %	76.9 %
	FAR	0 %	0 %	0 %	9.1 %
	FOM	50 %	35.7 %	30.8 %	23.1 %

Onset-Differenz [d] Sim - Obs

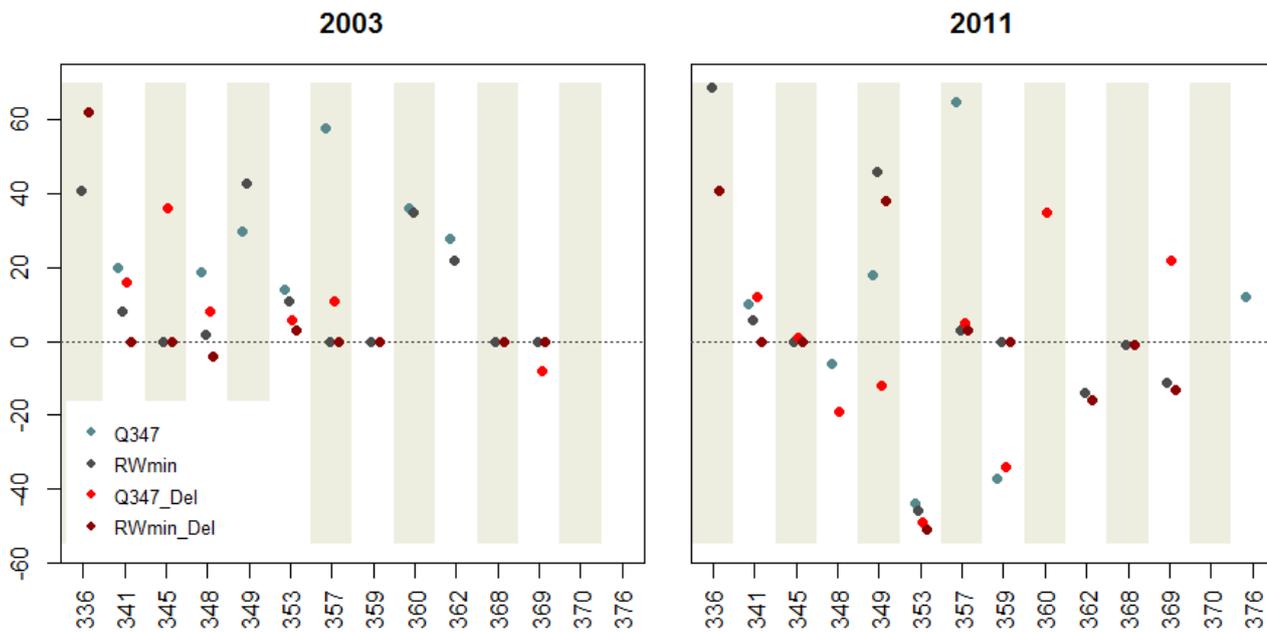


Abb. 47: Vergleich der Differenzen im **Onset** zwischen dem **Delay-Modell** bzw. dem Standard-Modell und dem beobachteten Wert für die beiden Schwellenwerte Q347 und RWmin. Die Differenz in Anzahl Tagen wurde als simuliertes Datum – beobachtetes Datum berechnet. Positive Werte stehen damit für einen verspäteten simulierten Onset, negative für einen verfrühten.

Zusammenfassend kann nicht eindeutig gesagt werden, ob sich die Delay-Variante besser für die Niedrigwasser-Simulation in den 14 betrachteten Gebieten eignet. Die grafischen Auswertungen der Modellierungen weisen darauf hin, dass sich das Delay-Modell für die hier untersuchten Gebiete besser eignet, die Schwellenwert-Auswertung zeigt allerdings sehr uneinheitliche Ergebnisse. Auch die Kontingenz-Tabelle zeigt sowohl positive wie negative Entwicklungen. Es kann im Gegenzug aber ebenso wenig gesagt werden, dass sich die Delay-Variante schlechter als das Standard-Modell eigne. Aus diesem Grund wird im letzten Schritt, der Co-Kalibrierung mit Grundwasserständen (siehe Kap. 6.5) mit beiden Modellen, dem Standard-Modell und (wo in diesem Kapitel getestet) der Delay-Variante, gearbeitet.

6.4. Einfluss der potentiellen Evapotranspiration (PET)

Weil die Modellierungen mit der Delay-Variante nicht die erhofften deutlichen Verbesserungen in der Modellgüte erbrachten, wurden weitere Untersuchungen nach der Ursache, die für schlechte Modellergebnisse oder gar ein Modellversagen verantwortlich sein könnten, unternommen. Dabei wurde ein interessanter und unerwarteter Zusammenhang entdeckt: Abb. 48 zeigt einen klaren Zusammenhang zwischen der erreichten Modellgüte und dem Wert Q50. Im Gegensatz zu Q_{347} steht 50 hier nicht für die Anzahl Tage, sondern für 50%; damit gemeint ist also der Abflusswert, der in 50% aller Fälle erreicht oder überschritten wird. Anders ausgedrückt bezeichnet Q50 den mittleren Abfluss (Medianwert) (z.B. Smakhtin, 2001: 155). Abb. 48 zeigt nun auf, dass die Modellierung des Abflusses problematischer wird, je kleiner der Q50-Wert eines Einzugsgebietes ist.



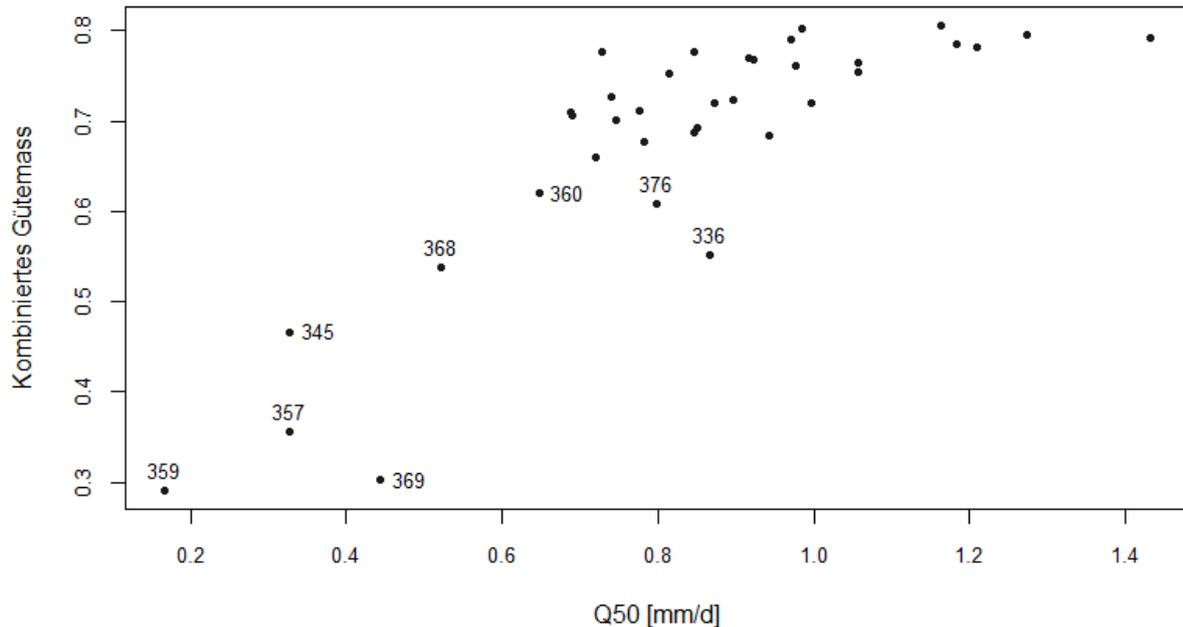


Abb. 48: Zusammenhang zwischen erreichter Modellgüte (kombiniertes Gütemass) und dem Wert Q50.

Dieser Zusammenhang könnte einerseits darauf zurückzuführen sein, dass bei einem sehr tiefen Wert für Q50 – die Abflusswerte also über einen grossen Zeitraum extrem niedrig sind – schlicht die Grenzen der Messgenauigkeit erreicht werden. Abb. 49 zeigt am Beispiel des Einzugsgebietes 369 in der Auswertungsperiode 2011, dass die Messwerte im Niedrigwasser-Bereich eine Treppung aufweisen. Dasselbe Phänomen ist auch bei Gebiet 359 sehr deutlich erkennbar und bei den Gebieten 357 und 345 ansatzweise. Für alle übrigen Gebiete (mit $Q50 > 0.5$ mm/Tag) ist keine solche Treppung zu erkennen. Gemäss WMO (2008: 30ff) ist dieser Effekt ein typischer Ausdruck der begrenzten Messauflösung im Niedrigwasserbereich. Wenn die Messinstrumente an ihre Genauigkeitsgrenze stossen, kommt zwangsläufig auch das Modell an seine Grenzen, denn das Modell kann nicht besser als seine Input-Daten sein.

An dieser Stelle ist zu ergänzen, dass derselbe Zusammenhang auch gegen Q_{347} und den Minimalabfluss (der gesamten verwendeten Zeitreihen) getestet wurde, mit der Erwartung, eine ähnliche Korrelation zu finden. Jedoch konnte für diese beiden Masse kein ebenso deutlicher Zusammenhang festgestellt werden. Worin genau der Grund für den Zusammenhang zwischen Q50 und der erreichten Modellgüte liegt, konnte bislang nicht restlos geklärt werden.

Zusammen mit dem Wissen, dass in den Kalibrierungen mehrmals hohe Werte für den Parameter BETA erreicht wurden, weckte das Ergebnis von Abb. 48 den Verdacht, dass eine Verbindung zur Evapotranspiration bestehen könnte. Konkret wurde vermutet, dass die verwendeten Werte zu niedrig sein könnten.

Einen Hinweis in diese Richtung geben Kay & Davies (2008), die aufgezeigt haben, dass die Oudin-Methode die PET im Jahresschnitt leicht unterschätzt (mit einer Tendenz zur Unterschätzung im Winter und Überschätzung im Sommer).

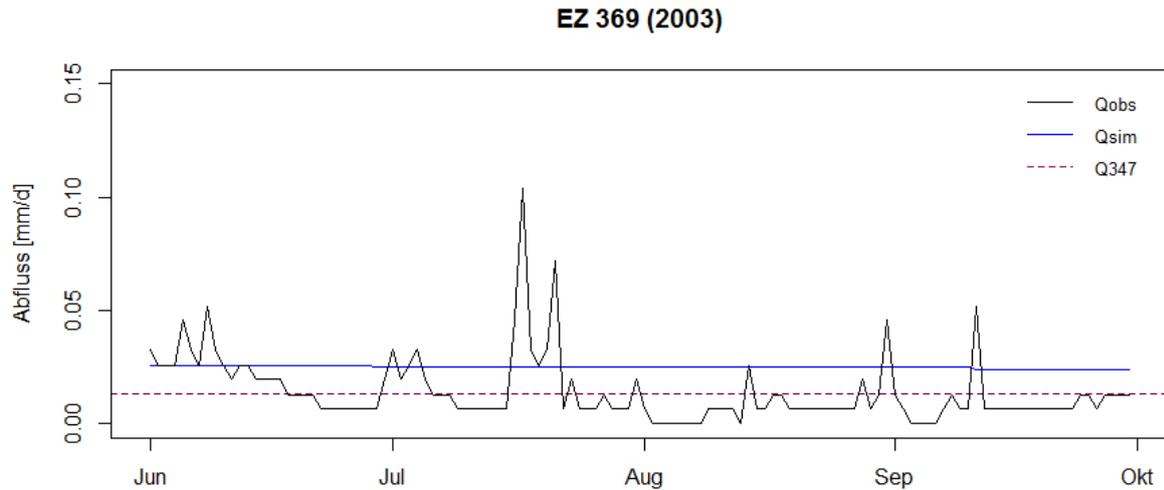


Abb. 49: Simulation (Qsim) durch das Standard-HBV-Modell in der Auswertungsperiode 2003. Der Abfluss befindet sich auf extrem niedrigen Werten und fällt zeitweise sogar auf 0. In den Messwerten (Qobs) ist eine deutlich Plafonierung erkennbar, die vermutlich auf die Grenzen der Messtechnik zurückzuführen sind.

Um dieser Vermutung auf den Grund zu gehen, wurden die PET-Werte in einem ersten Schritt um je 20% erhöht und reduziert und das Modell (für diese Tests wurde nur das Standard-Modell verwendet) mit den neuen PET-Werten neu kalibriert. Für diese Kalibrierungen wurden dieselben Einstellungen wie in den Standard-Modellierungen verwendet. Einzig für die beiden Gebiete 337 und 2471 musste die Parameter-Untergrenze für BETA auf 0.001 erniedrigt werden, um zu verhindern, dass BETA an der unteren Limite von 0.5 blockiert wird. Die Einstellungen weichen nur in diesem Punkt von den in Kap. 6.1 verwendeten Werten ab.

Untersucht wurden insgesamt 15 Einzugsgebiete, von denen 13 entweder ein schlechtes Modellgüte-Ergebnis in der Kalibrierung erzielt hatten und/oder hohe BETA-Werte (> 6) aufwiesen. Das Gebiet 337 weist mit 1.9 den tiefsten überhaupt erreichten BETA-Wert auf und wurde daher als Referenzwert ergänzt. Im selben Sinne wurde das Gebiet 2450, das über einen mittleren BETA-Wert (3.97) verfügt, hinzugenommen.

Die ersten Versuche zeigten, dass eine Erhöhung der PET zu einer Verbesserung der Modellgüte führt. Daher wurde die PET systematisch weiter um die Faktoren 1.5, 2, 4, 8 und 16 erhöht (jeweils ausgehend vom Originalwert). Die PET-Werte wurden derart stark erhöht, weil vermutet wurde, dass ab einem bestimmten Punkt die maximal erreichbare Modellgüte überschritten ist und das kombinierte Gütemass bei einer weiteren Erhöhung der PET abnehmen würde. Allerdings konnten selbst bei einer Erhöhung um den Faktor 16 nur leichte Abnahmen der Modellgüte festgestellt werden; der erwartete deutliche Abfall der Modellgüte blieb aus. Die Ergebnisse sind in Abb. 50 zusammen gestellt. Die Grafik unterscheidet zwischen Gebieten, die in der Standardkalibrierung einen BETA-Wert über 5 aufwiesen (obere Grafik) und solchen, deren BETA-Wert unter 5 lag (untere Grafik). Es wurde nämlich vermutet, dass diese Gebiete unterschiedlich auf die Erhöhung der PET reagieren würden. Dies hat sich auch bestätigt: Die Gebiete der unteren Grafik starten bereits auf relativ hohem Niveau und zeigen nur mässige bis sehr schwache Reaktionen. Um den Unterschied deutlicher zu machen, wurde für beide Grafiken dieselbe Skalierung der y-Achse verwendet. Anders hingegen bei der oberen Grafik, wo vor allem die besonders problematischen Gebiete 345, 353, 359 und 369 mit einer sehr ausgeprägten Erhöhung der Modellgüte auf eine Erhöhung der PET reagieren. Eine Ausnahme unter den beobachtbaren Verhalten bildet das Gebiet 368, das eine schwankende Kurve zeigt.

In allen Gebieten ausser 368 lässt sich eine Abnahme der Modellgüte bei einer Verkleinerung der PET-Werte sowie eine teils sehr ausgeprägte, teils nur schwache Verbesserung der Modellgüte bei einer Erhöhung der PET-Wert bis zu Faktor 2 erkennen. Bei Faktoren über 2, spätestens ab Faktor 8, zeigen alle Gebiete (ausser 368) eine Tendenz zur Abnahme der Modellgüte.

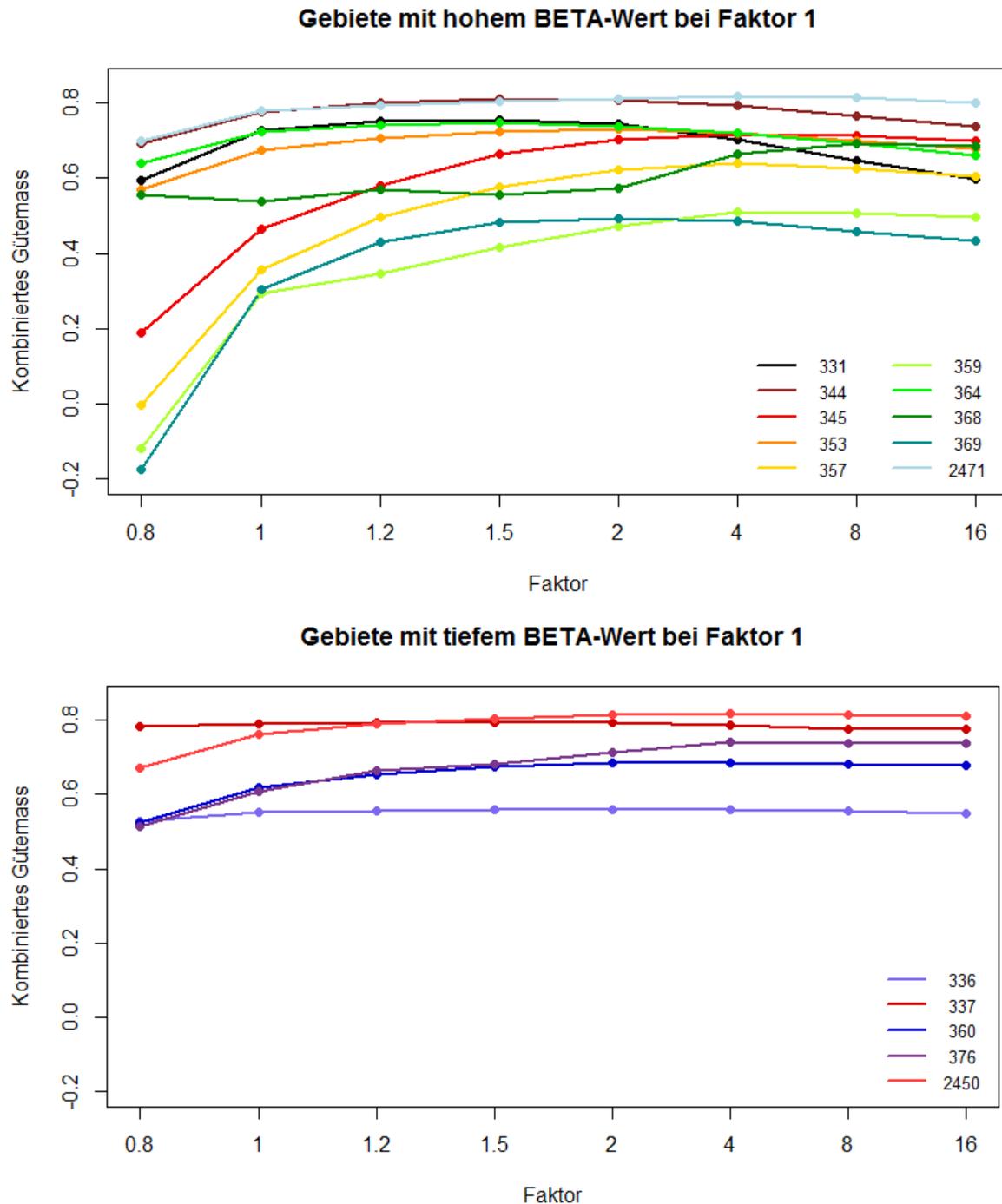


Abb. 50: Beeinflussung der Modellgüte durch Variieren der potentiellen Evapotranspiration (PET) an 15 Einzugsgebieten. Die x-Achse gibt den Faktor an, mit welchem die ursprüngliche PET verändert wurde (Faktor 1 entspricht also dem Ausgangswert). Die obere Grafik zeigt Gebiete, deren BETA-Wert bei der Standardkalibrierung über 5 lagen; die BETA-Werte der Gebiete der unteren Grafik lagen alle unter 5; den in der Standardkalibrierung tiefsten überhaupt erreichten BETA-Wert zeigte EZ 337. Zur besseren Vergleichbarkeit entsprechen sich die y-Skalen.

6.5. Grundwasser-Co-Kalibrierung

6.5.1. Trockenheitsmasse und Abflussverhalten

Wie gemäss Seibert (2000) erwartet, ergibt sich mit der Co-Kalibrierung in allen Gebieten eine leichte Verschlechterung des kombinierten Gütemasses. Das in der Grundwasser-Co-Kalibrierung neu berechnete kombinierte Gütemass, das zusätzlich die Zielfunktion *GroundWaterSR* beinhaltet, kann allerdings nicht direkt mit der Standardsimulation verglichen werden; das bisherige Gütemass musste erst mit der Gewichtung nach Tab. 5 berechnet werden. So zeigte sich, dass das Gütemass in den 13 Gebieten im Durchschnitt um 0.013 fällt, wobei die grösste Abnahme 0.028 bei Gebiet 360 beträgt (Tab. 14).



Für die vier Gebiete, die ebenfalls in den Tests zum Delay-Modell verwendet worden waren, wurde zusätzlich dieses Modell mit den Grundwasserdaten co-kalibriert. Auch hier zeigt ein Vergleich der Gütemasse eine Erniedrigung des kombinierten Gütemasses. Die grösste und tatsächlich markante Verschlechterung des Gütemasses um -0.134 tritt bei Gebiet 353 auf.

Tab. 14: Ergebnisse der Kalibrierungen mit der HBV-GW-Variante im Vergleich zum Standard-Modell.

EZ-Nr.	Modellgüte	Modellgüte
	Standard	HBV-GW
Standard-Modell		
331	0.727	0.716
332	0.791	0.781
334	0.768	0.742
341	0.684	0.672
342	0.707	0.702
348	0.659	0.634
353	0.677	0.652
354	0.719	0.719
360	0.620	0.592
372	0.752	0.748
374	0.802	0.798
2450	0.764	0.753
2471	0.781	0.772
Delay-Modell		
341	0.703	0.699
348	0.664	0.653
353	0.673	0.539
360	0.631	0.576

In Tab. 26 bis Tab. 30 im Anhang sind die Ergebnisse der Grundwasser-Co-Kalibrierung anhand der Schwellenwertmethode ausgewertet. Die Auswertungen wurden wie für die Delay-Kalibrierungen qualitativ vorgenommen, weil ein quantitativer Vergleich unter Einbezug von *Missed Events* und *False Alarms* schlecht möglich ist. Der Bewertungsschlüssel entspricht demjenigen in Kap. 6.3, ausser dass die Wertung von  zusätzlich für markante Verschlechterungen (im Bereich von mindestens einer Grössenordnung) verwendet wurde.

Ein Blick auf die Wertung der simulierten Dauer (Tab. 26) zeigt, dass Verschlechterungen tendenziell überwiegen, es treten aber auch Verbesserungen auf. Grundsätzlich simuliert die Co-Kalibrierung mit Grundwasserdaten kürzere Trockenheitsereignisse als das Standard-Modell. Weil das Standard-Modell die Trockenheit tendenziell überschätzt, führt dies durchschnittlich (unter Nicht-Einbezug von

Missed Events und *False Alarms*) eher zu einer Modellverbesserung. Nur im 2003 lässt sich für Q_{347} deutlich erkennen, dass die HBV-GW-Simulation durchschnittlich markant zu tief liegt. Dabei fällt das Gebiet 360 mit konsistent über alle vier Auswertungen hinweg auftretenden Überschätzungen der Dauer auf. In allen vier Fällen leicht verbessert hat sich hingegen das Gebiet 331.

Dauer [d] der Trockenheit

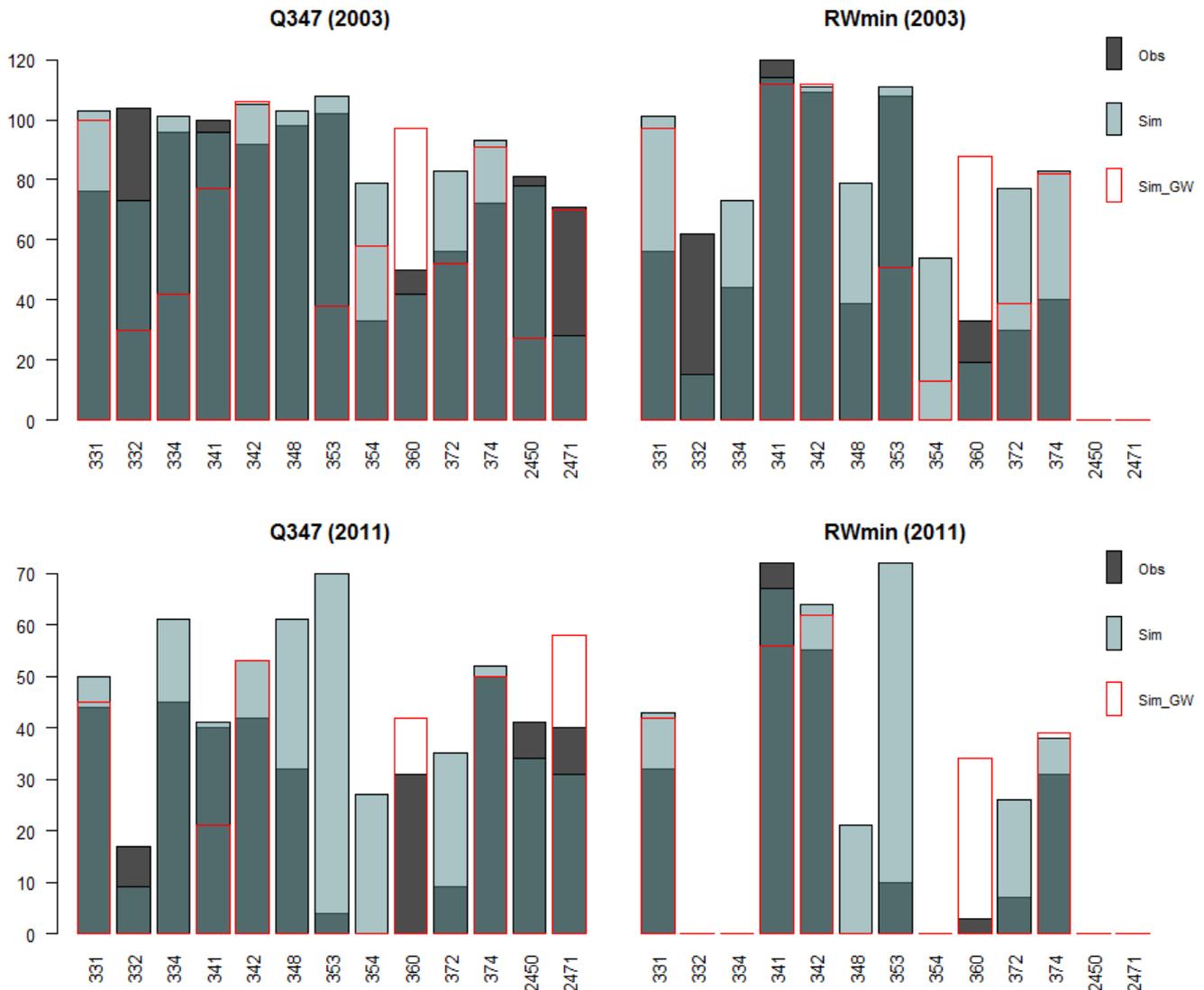


Abb. 51: Vergleich der Simulationen des **HBV-GW-Modells** (Sim_GW) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für die Trockenheits-Dauer im 2003 (oben) und 2011 (unten) für die beiden Schwellenwerte Q_{347} (links) und RW_{min} (rechts).

Für die Defizit-Simulation zeigt sich noch deutlicher eine allgemeine Tendenz zur Verschlechterung (Tab. 27). Ebenso setzt sich die Tendenz der Verminderung fort: Insbesondere im 2003 berechnet das HBV-GW-Modell für Q_{347} mit Ausnahme weniger Gebiete deutlich geringere Defizitwerte (Abb. 26). Auch hier sticht das Gebiet 360 mit deutlichen Überschätzungen des Defizits ins Auge (Abb. 26 und Abb. 57); ebenso verhält sich Gebiet 2471 für Q_{347} . Gebiet 331 zeigt hier in der HBV-GW-Simulation nur sehr schwache Abweichungen von der Standardsimulation, wobei HBV-GW im 2003 jeweils eine leichte Modellverbesserung, im 2011 aber je eine Verschlechterung bewirkt. Andere Gebiete (z.B. 332, 334, 32 und 354 im 2003 für Q_{347}) weisen sehr grosse Unterschiede zwischen beobachtetem Wert, Standardsimulation und HBV-GW-Simulation auf.

Für Q_{347} im 2003 lässt sich für die Intensität eine durchschnittliche Tendenz zur Unterschätzung ausmachen; für die übrigen Auswertungen lässt sich kein klarer Trend feststellen (Abb. 53). Wiederum stechen die Gebiete 360 und 2471 mit markanten Überschätzungen hervor. Ein Blick auf Tab. 28 zeigt einige Verbesserungen mehr im Vergleich zum Defizit, aber immer noch überwiegend Verschlechterungen.

Dasselbe Bild zeigt sich im Vergleich der simulierten Minimalabflüsse (Abb. 54 und Tab. 30): Grundsätzlich überwiegen die Verschlechterungen, wobei im 2011 mehr Verbesserungen als im 2003 erreicht wurden. In beiden Auswertungsperioden führt die HBV-GW-Simulation zu einer Überschätzung des minimalen Abflusses und ist damit in beiden Fällen im Durchschnitt schlechter als die Standard-Simulation. Obwohl die Standard-Simulation die Werte im 2003 tendenziell unterschätzt, ist die durchschnittliche Abweichung geringer als die Überschätzung durch das HBV-GW-Modell.

Defizit [mm] der Trockenheit

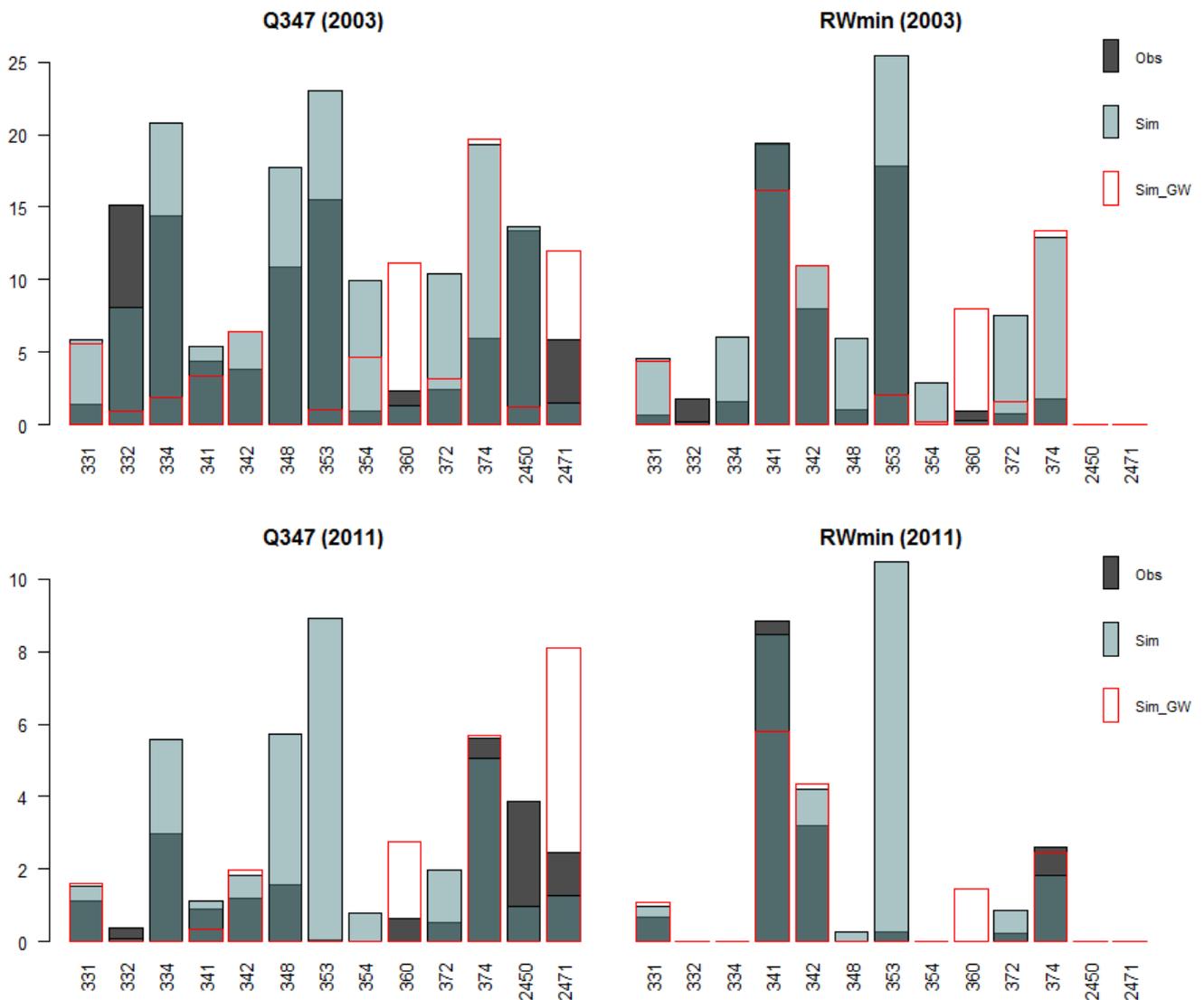


Abb. 52: Vergleich der Simulationen des **HBV-GW-Modells** (Sim_GW) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für das Trockenheits-Defizit im 2003 (oben) und 2011 (unten).

Intensität [mm/d] der Trockenheit

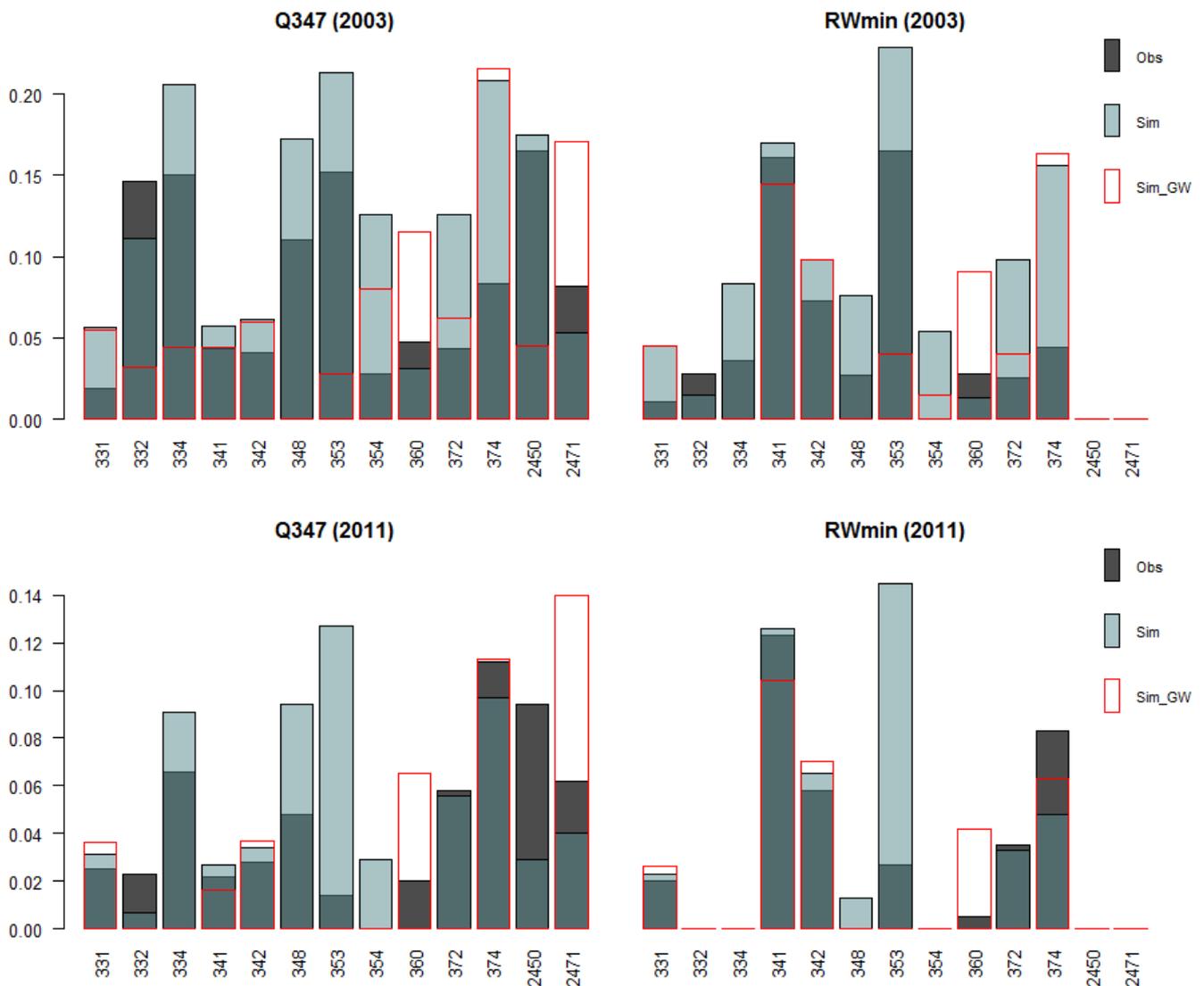


Abb. 53: Vergleich der Simulationen des **HBV-GW-Modells** (Sim_GW) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für die Trockenheits-Intensität im 2003 (oben) und 2011 (unten).

Bei der Auswertung der Onset-Simulationen zeigt sich wiederum ein heterogenes Bild: Die Gebiete reagieren sehr unterschiedlich, es überwiegen aber die Verschlechterungen. Einmal mehr sticht das Gebiet 353 ins Auge, für das die HBV-GW-Simulation im 2003 markante Verschlechterungen und im 2011 *Missed Events* erzeugt. Einige Gebiete (EZ 331, 342, 374) zeigen nur geringe Veränderungen ihrer Vorhersagen, andere (EZ 332, 334, 360, 2450) resultieren in beträchtlichen Abweichungen.

Insgesamt fällt zu den negativen Entwicklungen auf, dass insbesondere *Missed Events* neu häufiger auftreten. Wiederum lohnt es sich aber, einen genaueren Blick auf die Verschlechterung zu werfen: Beispielsweise zeigt Abb. 56 des Gebietes 353 in der Auswertungsperiode 2011, dass die HBV-GW-Simulation visuell beurteilt markant besser als die Standard- und die Delay-Simulation ist. Weil sie aber knapp über dem Schwellenwerten verbleibt, schneidet sie als *Missed Event* schlechter ab als die beiden anderen Simulationen.

Minimalabfluss [mm/d] der Trockenheit

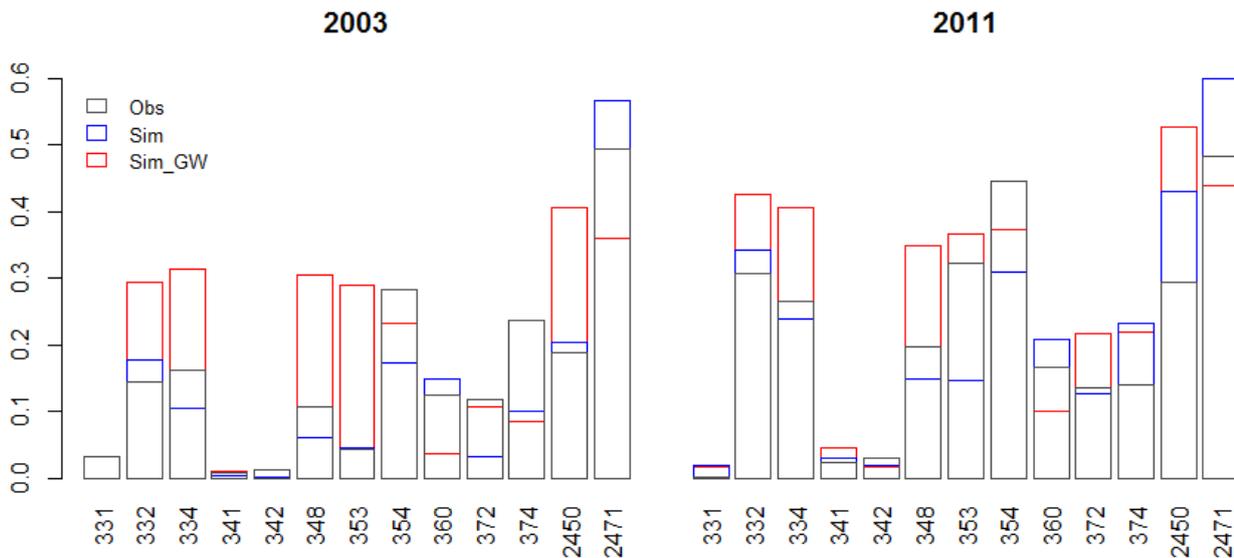


Abb. 54: Vergleich der Simulationen des **HBV-GW-Modells** (Sim_GW) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für den **minimalen Abfluss** in den beiden Auswertungsperioden 2003 (links) und 2011 (rechts).

Onset-Differenz [d] Sim - Obs

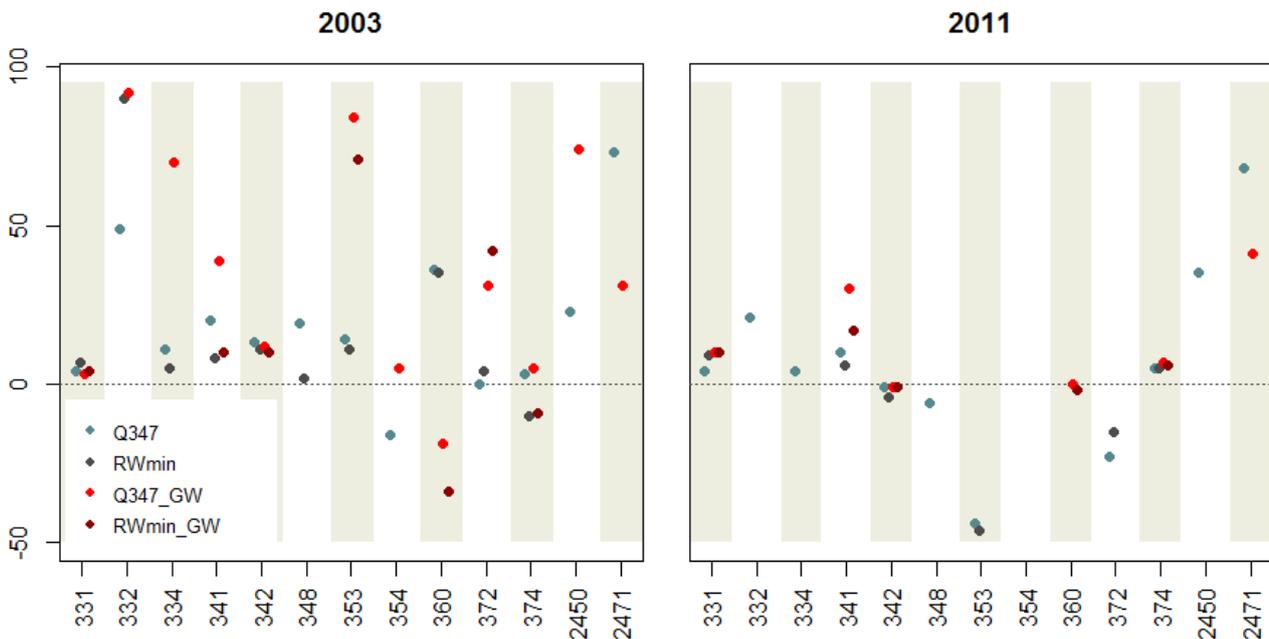


Abb. 55: Vergleich der Differenzen im **Onset** zwischen dem **HBV-GW-Modell** (mit Suffix _GW) bzw. dem Standard-Modell (ohne Suffix) und dem beobachteten Wert für die beiden Schwellenwerte Q347 und RWmin. Die Differenz in Anzahl Tagen wurde als simuliertes Datum – beobachtetes Datum berechnet. Positive Werte stehen damit für einen verspäteten simulierten Onset, negative für einen verfrühten.

Ein interessantes Phänomen ist bei Gebiet 360 im 2003 zu beobachten (Abb. 57): Während die Simulierung durch das Standardmodell das mittlere Niveau der Beobachtungen gut trifft, liegt die HBV-GW-Simulation gerade etwa so viel zu tief, wie die Delay-Simulation zu hoch liegt. Für dieses Einzugsgebiet sinkt die HBV-GW-Simulation auch in der Auswertungsperiode 2011 zu stark ab, trifft dabei allerdings das Einsetzen der Trockenheit perfekt (siehe Tab. 29). Das Gegenteil ist für das

Gebiet 372 im 2003 der Fall: Obwohl das Einsetzen der Trockenheit viel zu spät vorhergesagt wird, ist die Simulation an sich deutlich besser als die Standardvariante.

Allgemein kann beobachtet werden, dass co-kalibrierte Simulationen im Niedrigwasserbereich tendenziell höhere Abflüsse erzeugen als das Standardmodell (und das Delay-Modell). Je nach Ausgangslage kann dies eine Modellverbesserung oder Verschlechterung bewirken; die neu auftretenden *Missed Events* werden jedenfalls durch diese Eigenschaft verursacht. Eine Ausnahme bilden die Gebiete 342 und 374, für die beide Modelle nahezu identische Ergebnisse erreichen, sowie die Gebiete 360 (siehe Abb. 57) und 2471, für die HBV-GW tiefere Abflüsse berechnet. Für alle übrigen Einzugsgebiete simuliert HBV-GW leicht bis deutlich höhere Niedrigwasserniveaus. Abb. 58 illustriert, dass dies unterschiedliche Auswirkungen haben kann: Im Juni (Rezessionsbereich) ist das Niveau dadurch etwas zu hoch, ab Mitte Juli bis Ende August aber sehr gut und deutlich besser als die Standardsimulation, ab September liegen beide Modelle zu tief, wobei die HBV-GW-Variante deutlich besser, da höher ist; sie fällt aber ebenfalls fälschlicherweise unter den RWmin-Grenzwert. Nach der Schwellenwertmethode wird sie daher zwar ebenfalls als *False Alarm* ausgewertet, nähert sich aber ansonsten in den ausgewerteten Kennwerten den Sollwerten an.

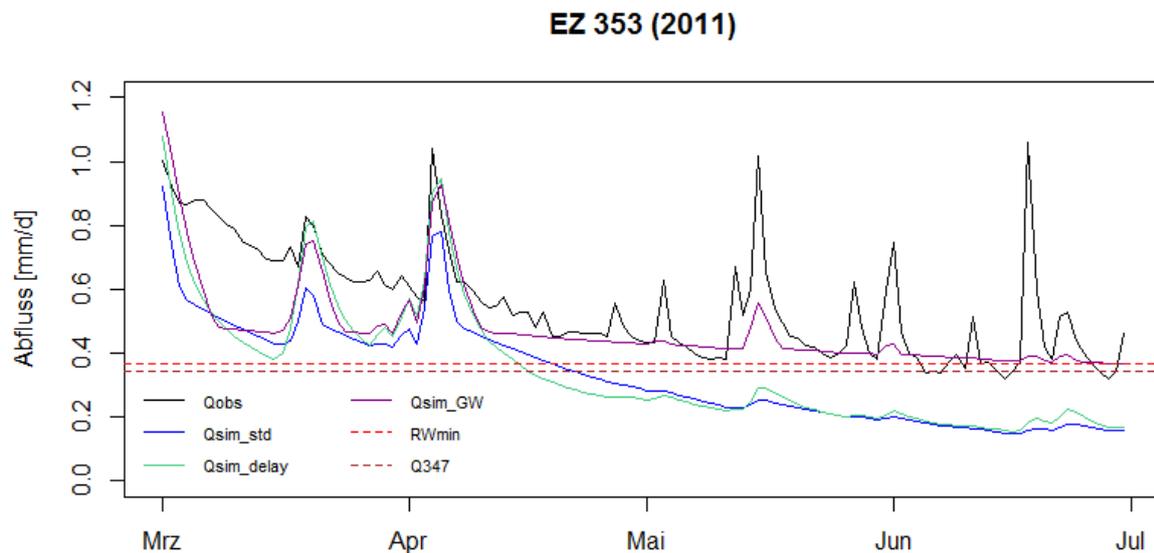


Abb. 56: HBV-GW-Simulation (Q_{sim_GW}) für das Einzugsgebiet 353 in der Auswertungsperiode 2011 im Vergleich zur Standard-Simulation (Q_{sim_std}) und der Delay-Simulation (Q_{sim_delay}). Die HBV-GW-Simulation ist deutlich besser, bleibt aber knapp über den Schwellenwerten und schneidet daher als *Missed Event* schlechter ab als die anderen Simulationen.

EZ 360 (2003)

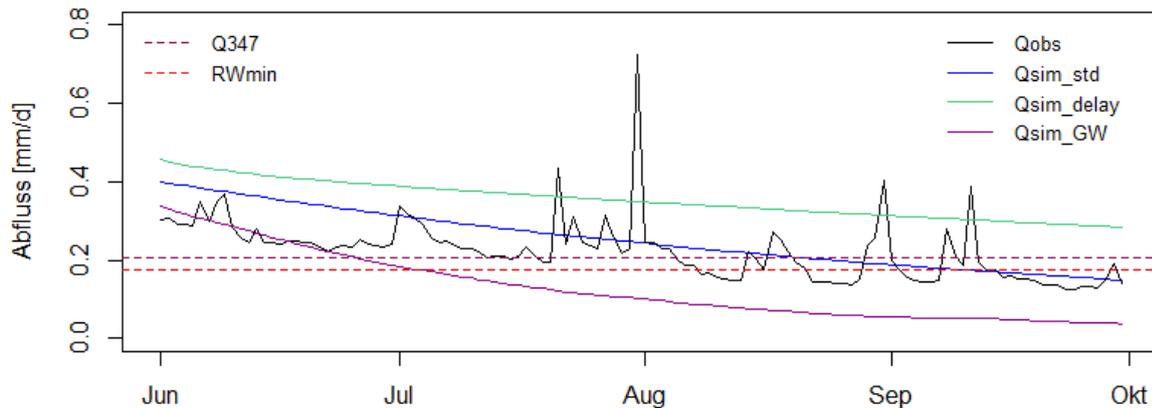


Abb. 57: HBV-GW-Simulation für das Einzugsgebiet 360 in der Auswertungsperiode 2003. Dies ist eines von zwei Gebieten, in denen HBV-GW tiefere Niedrigwasserabflüsse als das Standardmodell erzeugt.

EZ 354 (2003)

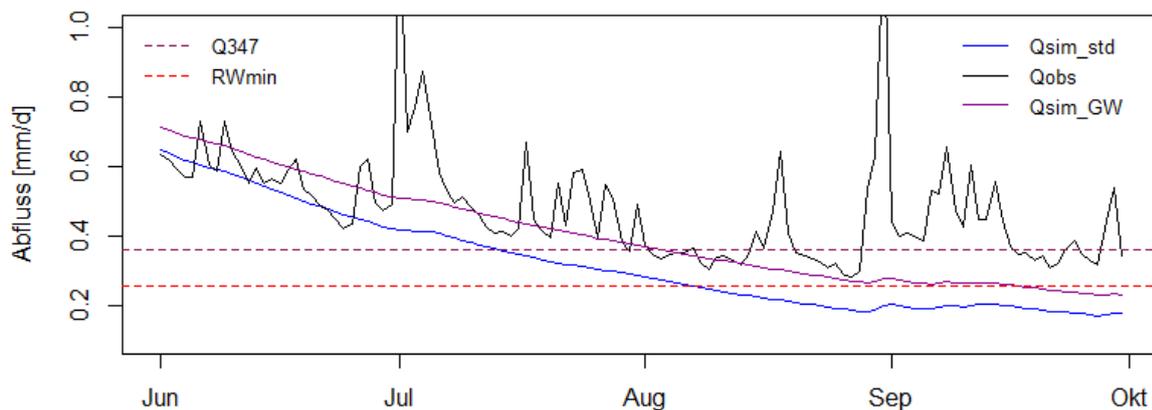


Abb. 58: HBV-GW-Simulation für das Einzugsgebiet 354 in der Auswertungsperiode 2003. Diese Simulation ist ein typisches Beispiel dafür, dass HBV-GW in fast allen Gebieten höhere Niedrigwasserabflüsse berechnet als das Standardmodell.

6.5.2. Auswertungen der Kontingenztabellen

Das häufigere Auftreten von *Missed Events* spiegelt sich klar in der FOM wieder (siehe Tab. 15): Bei den Simulationen des Standardmodells erreicht sie maximal knapp 15% und liegt in beiden Fällen von 2003 sogar auf 0%; für die Simulationen des HBV-GW-Modells steigt die FOM in allen Fällen, für Q_{347} im 2011 sogar auf 50%.

Im Gegenzug sinkt die FAR: In der Auswertungsperiode 2011 verschwinden die beiden in der Standard-Simulation auftretenden *False Alarms*, womit die FAR auf 0 sinkt. Für RWmin im 2003 verschlechtert sich die FAR scheinbar, allerdings liegt nach wie vor nur ein einziger Fall eines *False Alarm* vor; weil die Anzahl *Hits* aber abgenommen hat, steigt der Wert der FAR. Die POD zeigt auf ganzer Linie eine leichte (Q_{347} im 2003) bis markante (Q_{347} im 2011) Verschlechterung.

Bezüglich den Auswertungen der Kontingenztabellen muss angemerkt werden, dass für die Werte des Standardmodells nur diejenigen Gebiete, die auch mit Grundwasserdaten co-kalibriert wurden, verwendet wurden. Im Gegensatz zu den für die Delay-Tests verwendeten Gebieten sind die hier verwendeten grösstenteils unproblematisch. Die Werte für das Standardmodell sind daher besser als

in der Gesamtauswertung (Tab. 9 und Tab. 10), ganz im Gegensatz zu den Werten des Delay-Modells (Tab. 13), wo ausschliesslich problematische Gebiete betrachtet wurden.

Tab. 15: Auswertung der Kontingenztabellen des HBV-Standardmodells und des HBV-GW-Modells im Vergleich.

		Q347 2003	RWmin 2003	Q347 2011	RWmin 2011
Std-Modell	POD	100%	100 %	91.6 %	85.7 %
	FAR	0 %	9.1 %	8.3 %	14.3 %
	FOM	0 %	0 %	8.3 %	14.3 %
HBV-GW	POD	92.3 %	70 %	50 %	71.4 %
	FAR	0 %	12.5 %	0 %	0 %
	FOM	7.7 %	30 %	50 %	28.6 %

6.5.3. Delay-Modell mit Grundwasser-Co-Kalibrierung

Für vier Gebiete, für die sowohl Grundwasserdaten zur Verfügung stehen und für die ebenfalls bereits das Delay-Modell getestet wurde, wurde schliesslich eine Co-Kalibrierung auf Grundwasserdaten mit dem Delay-Modell durchgeführt. Zwei dieser vier Gebiete sind echte karsthaltige Gebiete (341 und 360), die anderen beiden (348 und 353) enthalten keinen Karst, sondern wurden wegen schlechten Modellergebnissen mit der Standardversion mit der Delay-Variante getestet. Die Resultate dieser Simulationen sind an Tab. 26 bis Tab. 30 (Anhang) ergänzt. Wie erwartet, verschlechtert sich auch in diesen Fällen das kombinierte Gütemass, wobei die grösste Differenz beim Gebiet 353 auftritt (Tab. 14). Die Werte zeigen einen Trend zur Modellverbesserung, allerdings nicht für alle Kennwerte: Während für die Dauer Verbesserungen überwiegen, werden für Onset und Minimum in der Auswertungsperiode 2003 nur Verschlechterungen erzielt.

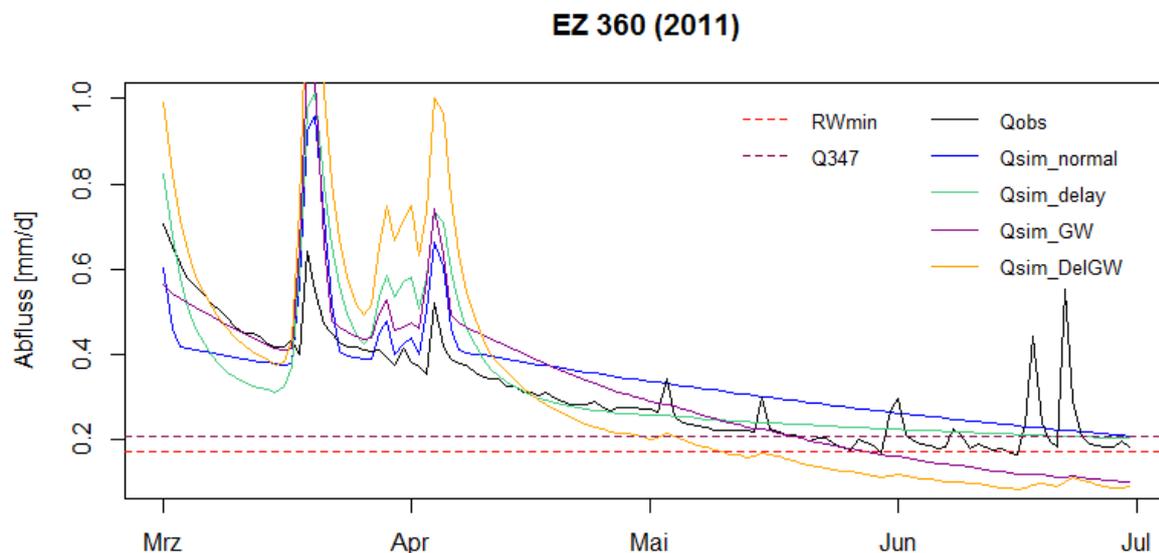


Abb. 59: Gebiet 360 in der Auswertungsperiode 2011. Vergleich der Standard-Simulation, der Delay-Simulation, und der entsprechenden Simulationen, die auf Grundwasser co-kalibriert wurden. Qsim_DelGW steht für das co-kalibrierte Delay-Modell und ist bei diesem Gebiet deutlich schlechter als die übrigen Simulationen.

Erfreulich ist allerdings, dass alle in der Delay-Simulation auftretenden *Missed Events* sowie der eine *False Alarm* durch die Co-Kalibrierung auf Grundwasser ausgemerzt werden können. Damit steigt die POD für diese vier Gebiete auf 100%, während die FAR und die FOM auf 0% fallen. Es ist aber zu bedenken, dass diese Werte auf nur vier Gebieten basieren.

Die visuelle Entwicklung der Abflusskurven zeigt sich sehr heterogen: Während Gebiet 341 kaum einen Unterschied mit der Co-Kalibrierung erkennen lässt, zeigt Gebiet 360 eine deutlich schlechtere Kurve, weil sie noch tiefer abfällt als die bereits zu tiefe Grundwasser-Co-Kalibrierung mit dem Standardmodell (Abb. 59). Bei Gebiet 348 zeigt sich die Delay-Grundwasser-Kalibrierung deutlich besser als die Standardmodell-Grundwasser-Kalibrierung; im 2011 ist sie ebenfalls deutlich besser als die einfache Standardmodell-Simulation (Abb. 60) und die Delay-Simulation; im 2003 ist es hingegen schwierig zu beurteilen, ob eine Verbesserung oder eine Verschlechterung vorliegt.

EZ 348 (2011)

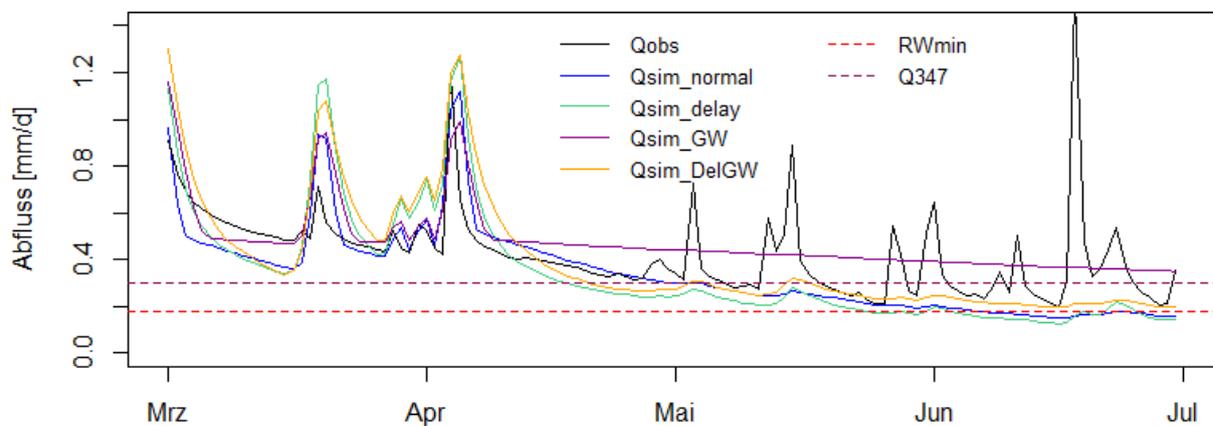


Abb. 60: Vergleich der Standard- und Delay-Simulationen mit und ohne Grundwasser-Co-Kalibrierung für das Gebiet 348 in der Auswertungsperiode 2011. In diesem Fall bildet die Grundwasser-co-kalibrierte Variante die beste Modellierung.

EZ 353 (2003)

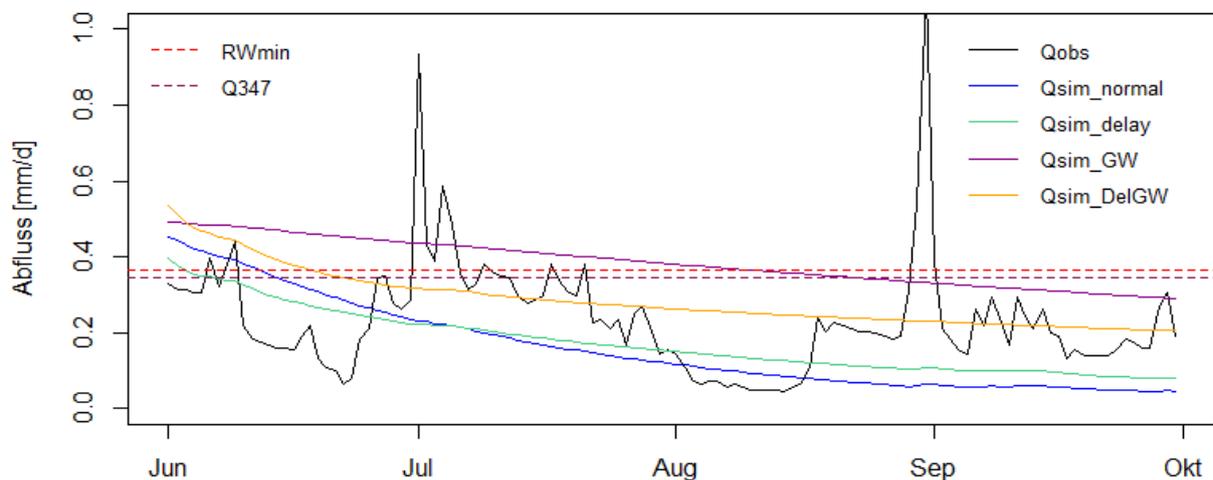


Abb. 61: Vergleich der Standard- und Delay-Simulationen mit und ohne Grundwasser-Co-Kalibrierung für das Gebiet 353 in der Auswertungsperiode 2003. Hier ist eine Entscheidung, welche Variante am besten ist, schwierig.

Für das Gebiet 353 ist die Entscheidung, welche Simulation am besten ist, visuell eher schwierig: Im 2011 ist die Delay-mit-Co-Kalibrierung ziemlich eindeutig am besten, im 2003 ist die Entscheidung aber stark davon abhängig, welche Aspekte am stärksten gewichtet werden. Die Delay-co-kalibrierte Variante schätzt zwar den niedrigsten Abfluss zu hoch, ist ansonsten aber nicht unbedingt schlechter als die Delay-Simulation und die Standardsimulation. In der Schwellenwertmethode schneidet sie aber ausser für die Trockenheits-Dauer in allen Kennwerten schlechter ab als die Delay-Simulation. Bei diesen vier Einzugsgebieten lässt sich die Tendenz beobachten, dass die Co-Kalibrierung eine Erhöhung der Delay-Simulation bewirkt, dass dieses Resultat aber tiefer liegt als die Ergebnisse der Co-Kalibrierung mit dem Standardmodell. Die Ausnahme bildet das Gebiet 360.

7. Diskussion

7.1. Vorhersagen des Standard-Modells

In der Auswertung der Simulationen des Standard-Modells fällt bei den Trockenheitsmassen auf, dass sie die Intensität tendenziell besser vorhersagen als das Defizit, wobei aber beide Werte überschätzt werden. Dies zeigt das Zusammenspiel zwischen Dauer und Defizit auf, das in der Intensität zusammengefasst wird. Weil sowohl Dauer als auch Defizit durchschnittlich überschätzt werden, kann für die Intensität wiederum ein besseres Ergebnis resultieren.

Für die Trockenheitsmasse wurden jeweils Durchschnittswerte aller 36 Gebiete angegeben. Bei dieser Verrechnung ist zu bedenken, dass über die Simulationen verschiedener Einzugsgebiete gemittelt wird, deren Modellierungen sehr unterschiedliche Gütemasse und unterschiedliche Stärken und Probleme aufweisen. Eine Mittelung gibt daher zwar einen allgemeinen Eindruck der Eignung des Modells für die gesamte betrachtete Region, ist aber durch Veränderungen bzw. Verbesserungen in den Modellen einzelner Gebiete stark beeinflussbar.

Bei der Auswertung der Kontingenztabelle sticht im Vergleich zu anderen Studien besonders die in allen vier Fällen sehr tiefe FAR heraus, die sich beispielsweise bei Addor et al. (2011: 2338) auf 50 bis 70% belief, allerdings für Hochwasserprognosen. Eine solch hohe FAR schmälert in der Praxis das Vertrauen in die Modellvorhersage deutlich, auch wenn die Folgen von falschen Alarmen im Gegensatz zu *Missed Events* weniger gravierend sind. Die wohl am direktesten vergleichbaren Resultate für Niedrigwasser finden sich bei Kempf (2012: 37). Bei diesem Vergleich schneiden die hier erreichten Resultate ebenfalls gut ab, wobei erwähnt werden muss, dass Kempf (2012) mit Prognosedaten arbeitete, wodurch grössere Unsicherheiten ins Modell eingebracht werden.

Zu der hier durchgeführten Kontingenz-Berechnungen müssen allerdings zwei Vorbehalte angebracht werden: Erstens sind 36 Einzugsgebiete statistisch nicht besonders aussagekräftig, und zweitens gilt auch hier der Einwand, dass Modellierungen verschiedener Einzugsgebiete miteinander verrechnet werden und damit nicht ganz einheitliche Dinge miteinander verglichen werden. Üblicherweise wird eine solche Auswertung in Kontingenztabelle für ein einziges Einzugsgebiet durchgeführt. Dass hier hingegen alle Gebiete direkt miteinander verrechnet werden, wird wiederum mit dem Argument begründet, dass ein Mass für die allgemeine Eignung des HBV-Modells zur Vorhersage von Trockenereignissen im Untersuchungsgebiet erhalten werden soll. Es muss aber weiter bedacht werden, dass hier nur zwei Perioden ausgewertet werden, von denen bekannt ist, dass Trockenheiten auftraten. Es ist davon auszugehen, dass dieser Umstand vor allem die FAR beeinflusst. Werden vollständige Zeitreihen ausgewertet, ist es nämlich so, dass das *No Event*-Feld eine sehr hohe Frequenz aufweist und die *Hits*, *Misses* und *False Alarms* im Vergleich dazu sehr selten sind, was ein Problem für die statistische Signifikanz darstellt. Dabei treten die *False Alarms* üblicherweise häufiger auf als *Hits* und *Misses* (Cloke & Pappenberger, 2009: 617).

In den Auswertungen mittels Schwellenwertmethode sowie in der visuellen Betrachtung der Kurven liessen sich in vielen Gebieten deutliche Unterschiede in der Vorhersage-Güte zwischen den beiden Auswertungsperioden 2003 und 2011 erkennen. Abb. 32 ist ein gutes Beispiel dafür, dass oft die eine Periode deutlich bessere Simulationen ergab als die andere. Die grafischen Darstellungen suggerieren, dass eher für die Auswertungsperiode 2011 bessere Ergebnisse erzielt wurden, die Schwellenwertmethode ergibt aber eher für 2003 die besseren Resultate. Das Modell scheint die beiden Ereignisse jedenfalls nicht gleich gut erfassen zu können. Diese Differenz kann mit den in Kap. 5.4 erläuterten Unterschieden in den zugrundeliegenden Prozessen erklärt werden: Die Trockenheit 2003 ist eine „klassische Trockenheit wegen Niederschlagsdefizit“, hingegen entsteht das trockene

Frühjahr 2011 durch das Zusammenspiel eines warmen Winters mit einem Niederschlagsdefizit, das bereits im Winter einsetzt („kombinierte Trockenheit“).

Ein Phänomen des Spätsommers 2003, das in verschiedenen Gebieten beobachtet werden konnte (Abb. 25), ist ein zu starkes Absinken der Simulationen. Der Grund dafür könnte im ungewöhnlich starken Austrocknen der Böden liegen: Sehr trockene Böden weisen eine reduzierte Infiltrationskapazität auf, im Extremfall (in ariden und semi-ariden Gebieten) werden sie als zeitweilig hydrophob bezeichnet. Ihre Infiltrationskapazität ist dann so stark reduziert, dass sie bei kurzfristigem Niederschlag kaum Wasser aufnehmen können – Hortonscher Oberflächenabfluss tritt auf (Davie, 2008: 81/82). In der Abflusskurve würde das bewirken, dass ein grösserer Anteil des Niederschlags als sonst üblich direkt in den Abfluss gelangt und nicht von den Böden aufgenommen und verdunstet wird. Da HBV keinen Hortonschen Oberflächenabfluss kennt und keine Informationen über die Infiltrationskapazität der Böden berücksichtigt bzw. immer mit derselben Rate an Wasser, das in die Boden-Routine gelangt, rechnet, würde es in diesem Fall den Abfluss unterschätzen, denn es erwartet, dass aufgrund der hohen Temperaturen starke Evapotranspiration auftritt und damit wenig Wasser zum Gewässerabfluss beiträgt. Damit könnten solche systematisch zu tief ausfallenden Simulation erklärt werden. Dabei kommt auch die Trennung zwischen der Boden-Routine und der oberen Grundwasserbox (SUZ) zum Tragen: In der Realität würden austrocknende Böden oberflächennahen Grundwasservorkommen Wasser entziehen. HBV lässt diese Koppelung aber nicht zu.

Ein weiteres Hindernis der Modellstruktur für eine korrekte Niedrigwasser-Simulation ist die in Kap. 3.2 erwähnte Tatsache, dass der simulierte Abfluss nie auf null fallen kann, weil die untere Grundwasserbox nicht komplett austrocknen kann. Bei vier der untersuchten Gebiete (331, 357, 359 und 369) kommt es innerhalb der gesamten Zeitreihe vor, dass das Gewässer völlig austrocknet (siehe Abb. 62). In den beiden Auswertungsperioden geschieht dies allerdings nur in Gebiet 369, in der Periode 2003 während insgesamt 12 Tagen und im 2011 während insgesamt 8 Tagen. Diese Werte kann HBV also grundsätzlich nicht simulieren. Obwohl HBV zur Berechnung der Modellgüte Werte von ≤ 0 ignoriert, könnte dies das schlechte Kalibrierungsergebnis des Gebietes 369 zumindest teilweise erklären. Das Ignorieren der Nullwerte in der Güteberechnung bewirkt nämlich, dass diese Extremwerte aus der Information, die das Modell erhält, entfernt werden. Das Modell dürfte daher die übrigen sehr niedrigen Abflüsse tendenziell überschätzen – was tatsächlich der Fall ist.

Weiter konnte in verschiedenen Gebieten ein konstanter Versatz zwischen Simulation und Messkurve beobachtet werden (Abb. 21, oben). Ein solcher systematischer Fehler könnte in der Praxis durch Anwendung einer Echtzeit-Aufdatierung (eng. *data assimilation* oder *real-time updating*) der Prognosen verbessert oder gar ausgemerzt werden. Mit diesem Verfahren werden die Modellprognose an die aktuellen Messwerte angeglichen. In der Forschung wird dieses Vorgehen unterdessen häufig angewandt (z.B. Wöhling, Lennartz, & Zappa, 2006; Sene, 2010: 124ff; Cloke & Pappenberger, 2009).

In verschiedenen Grafiken (z.B. Abb. 25 und Abb. 30) wurde ebenfalls aufgezeigt, dass die feinen Schwankungen im Niedrigwasserbereich die Simulationsfähigkeit von HBV übersteigen. Was eigentlich eine grosse Modellschwäche ist, muss aber nicht per se ein Nachteil für Prognosen sein: Gerade dadurch, dass die feinen Schwankungen „ausgemittelt“ werden, kann die Prognose robuster und damit für eine operationelle Anwendung wertvoller werden. Denn für das Erlassen von Nutzungsbeschränkungen, die nicht alle paar Tage aufgehoben und neu erlassen werden können, sondern längerfristig angelegt sein müssen, ist es schwieriger, mit stark schwankenden Abflussprognosen als mit „ausgemittelten“ Werten zu arbeiten.

Es kann also gesagt werden, dass die Resultate dieser Arbeit zwar nur unter bestimmten Vorbehalten zu betrachten sind, sie aber dennoch darauf hinweisen, dass HBV in den untersuchten Gebieten durchaus Potential zur Erstellung von zuverlässigen Niedrigwasserprognosen besitzt. Im Durchschnitt aller Einzugsgebiete hat sich das Standardmodell nämlich bewährt und eine Güte um 0.7 erreicht (Mittelwert gerade so viel darunter, wie der Median darüber liegt). Damit befindet sich dieser Durchschnitt an der Grenze zu einer Beurteilung mit „gut“. Dabei muss bedacht werden, dass einige problematische Gebiete diesen Durchschnitt zwar stark beeinflussen, jedoch für viele dieser Gebiete im Verlauf dieser Arbeit Möglichkeiten zu Modellverbesserung – besonders durch Erhöhung der PET-Werte (Kap. 6.4) – aufgezeigt werden konnten.



Abb. 62: Die Sissle bei Eiken (EZ 331) am 1. Juli 2008. Wie der Schmittenbach (EZ 369) ist auch die Sissle mehrmals im Verlauf der gesamten Zeitreihen komplett ausgetrocknet.

Quelle: W. Lehmann, ALG Wasserbau, BVU, Kanton Aargau

7.2. Sensitivitätsanalyse

Die Sensitivität der Parameter wurde auf zwei verschiedene Arten untersucht. Im Vergleich von Analyse 1 und 2 wäre zu erwarten gewesen, dass Analyse 2 eine deutlich grössere Parameterstreuung erzeugt als Analyse 1. Die Ergebnisse zeigen aber, dass die Streuung der Parameterwerte zwar etwas zugenommen hat und die Werte der Variationskoeffizienten demzufolge leicht erhöht wurden, dass aber die maximalen Werte der Variationskoeffizienten leicht abnahmen. Dies spricht für die Stabilität des HBV-Modells und zeigt, dass mit anderen Kombinationen bzw. Gewichtungen der Zielfunktionen nicht grundsätzlich andere Parameter zu erwarten sind. Allerdings wurden in Analyse 1 zehn Kalibrierungsläufe durchgeführt und in Analyse 2 pro neue Kombination nur je ein Lauf, insgesamt also fünf Kalibrierungen. Es ist davon auszugehen, dass die Streuung bei mehr Durchgängen zunehmen würde. Dennoch ergibt ein Vergleich von Analyse 2 mit Analyse 1 erfreulich positive Resultate.

Die Sensitivitätsanalyse hat ausserdem den in den Kalibrierungen gewonnen Eindruck bestätigt, dass die sensibelsten Parameter in der Schnee-Routine liegen. Die Parameter der Grundwasser-Routine haben sich als relativ gut definiert erwiesen. Dieses Ergebnis lässt hinterfragen, wie sinnvoll die in dieser Arbeit durchgeführte Co-Kalibrierung auf Grundwasserdaten ist. Es liesse sich vermuten, dass eine Co-Kalibrierung auf Schneemessdaten für das Modell wertvoller als auf Grundwasserdaten wäre. Es ist allerdings anzunehmen, dass Schneedaten deutlich spärlicher erhältlich sind als

Grundwasserdaten, welche verbreitet durch Behörden erfasst werden und damit meistens gut zugänglich sind. Eine Co-Kalibrierung der Grundwasser-Routine wirkt sich jedoch nicht nur auf diese Routine, sondern ebenfalls auf den ganzen Berechnungsprozess davor aus. Wie stark der Einfluss der durchgeführten Co-Kalibrierung auf die Streuung der Parameter, insbesondere auch derjenigen der Schnee-Routine, tatsächlich ist, müsste in einer erneuten Sensitivitätsanalyse untersucht werden.

Für die Stabilität des HBV-Modells sprechen darüber hinaus auch die Ergebnisse aus den 200 Kalibrierungen mit und ohne die ausgewerteten Zeiträume der Jahre 2003 und 2011. Die Abweichungen zwischen diesen Kalibrierungen sind erstaunlich gering – die grössten und dennoch geringen Schwankungen treten bei den Abflussspitzen auf (in Abb. 41 ist die logarithmische Achse zu beachten). Der Vergleich zeigte aber auch, dass das Entfernen der beiden Auswertungsperioden in der Kalibrierung zu einer leichten, aber systematischen Überschätzung der Niedrigstflüsse geführt hat. Daraus lässt sich schliessen, dass der Einbezug der gesamten Zeitreihe in die Kalibrierung, wie es für eine Anwendung in der Praxis sicherlich durchgeführt würde, Prognosen von extrem geringen Abflüssen leicht verbessern könnte.

Aufgrund der Ergebnisse der Sensitivitätsanalyse und der Untersuchung des Einflusses des Split-Sample-Verfahrens sowie unter der Annahme, dass sich das Einzugsgebiet 337 repräsentativ für die übrigen Einzugsgebiete verhält – was aufgrund der Erfahrungen aus den übrigen Untersuchungen angenommen werden kann –, kann also davon ausgegangen werden, dass HBV mit der GAP-Kalibrierung sehr einheitliche und in diesem Sinne zuverlässige Resultate liefert.

7.3. Karstgebiete und Modellvarianten

Ausgehend von den karsthaltigen Gebieten wurden drei Modellvarianten getestet, von denen sich aber die beiden Strukturen „One GW Box“ und „Three GW Boxes“ nicht bewährt haben. Der Fokus wurde daher auf die „Delay“-Variante gelegt. Schliesslich wurden 14 Einzugsgebiete, 7 davon mit Karstanteil, die übrigen 7 ohne Karst, zusätzlich mit der Modellvariante „Delay“ kalibriert. Die Ergebnisse haben gezeigt, dass nicht eindeutig gesagt werden, ob sich die Delay-Variante besser für die Niedrigwasser-Simulation eignet. Die grafischen Auswertungen suggerieren, dass das Delay-Modell sich für die untersuchten Gebiete besser bewährt bzw. ein natürlicheres Abflussverhalten ergibt, die Schwellenwert-Auswertung und die Kontingenz-Tabellen zeigen hingegen uneinheitliche Ergebnisse.

Das Modell scheint sich zudem nicht wie erhofft spezifisch für Karstgebiete zu eignen, da die grösseren Verbesserungen in der Modellgüte für nicht-verkarstete Gebiete als für karsthaltige erreicht wurden. Vermutlich ist dies darauf zurückzuführen, dass das Delay-Modell, wie es der Name sagt (engl. für „Verzögerung“), den Niederschlag nicht schneller, sondern verzögert in den Abfluss bringt. Modelle für Karstgebiete müssten aber eine schnell-reagierende Komponente enthalten – so zumindest gemäss dem Systemverständnis. HBV verfügt über keine solche Komponente, es kennt ja auch keinen (schnellen) Oberflächenabfluss. Das Delay-Modell scheint diesen Nachteil höchstens geringfügig besser als das Standardmodell auskorrigieren zu können.

Es muss jedoch angefügt werden, dass die Möglichkeit von Differenzen zwischen der tatsächlichen Karstverbreitung und der verwendeten Karstkarte besteht. Verkarstungen im Untergrund, die einen grossen Einfluss auf die Hydrologie eines Gebietes haben, sind nicht direkt zugänglich und damit ohne detaillierte Untersuchungen des Untergrundes schwer abschätzbar.

In Bezug auf das Delay-Modell lässt sich hingegen die Frage stellen, ob es sich dabei um eine Struktur handeln könnte, die sich besonders für kleine Einzugsgebiete eignet, denn alle nicht-verkarsteten Einzugsgebiete weisen eine Fläche kleiner als der Median aller 36 Einzugsgebiete (20 km^2) auf. Bei den verkarsteten Einzugsgebieten weisen zwei Gebiete (341 und 360) eine Fläche knapp über 20 km^2

auf (21 resp. 25.4 km²), die übrigen Gebiete befinden sich ebenfalls unter der Grenze von 20 km². Eine Kalibrierung verschiedener grösserer Einzugsgebiete zur Überprüfung dieser Hypothese wurde hier jedoch nicht durchgeführt.

Aufgrund der Aussagen von Seibert (2000: 219) und Van Loon & Van Lanen (2012: 1919) liesse sich ebenfalls vermuten, dass die Eigenschaften von Grundwasservorkommen eine entscheidende Rolle für die Eignung des Delay-Modells spielen. Eine Übereinstimmung zwischen erzielter Modellverbesserung und dem Vorkommen bzw. der Mächtigkeit von Aquiferen konnte allerdings nicht festgestellt werden.

7.4. Tests zur potentiellen Evapotranspiration

An 15 Gebieten wurden die Auswirkungen einer Veränderung der PET-Werte getestet. Die Neukalibrierungen zeigten deutlich, dass eine Verringerung der PET-Werte nahezu ausnahmslos zu einer deutlichen Modellverschlechterung führt. Eine Erhöhung der PET bis um den Faktor 2 bewirkt dagegen eine leichte bis sehr markante Verbesserung der Modellgüte (Abb. 50). Bei einer weiteren Erhöhung der PET-Werte kann ein allgemeiner Trend zur Modell-Verschlechterung festgestellt werden. Der erwartete starke Abfall der Modellgüte wurde allerdings auch mit dem sehr hohen Faktor 16 nicht erzielt. Eine Ausnahme bildet das Gebiet 368, das keinen einheitlichen Trend, sondern eine schwankende Modellgüte zeigt.

Ausserdem reagieren nicht alle Einzugsgebiete gleich stark auf eine Veränderung der PET-Werte. Gebiete, die bereits bei der ursprünglichen Kalibrierung einen tiefen BETA-Wert aufwiesen, sind auf eine Veränderung der PET-Werte eher wenig sensibel. Das Extrembeispiel hierfür bietet das Gebiet 337, das den tiefsten in den Standard-Kalibrierungen überhaupt erreichten BETA-Wert aufweist; auf eine Veränderung der PET reagiert es über den untersuchten Bereich so gut wie gar nicht (die Differenz zwischen minimal und maximal erreichter Modellgüte beträgt 0.018). Besonders empfindlich reagieren dagegen diejenigen Gebiete, deren Q50 unter 0.5 mm/d liegt (i.e. 359, 345, 357 und 369 (Ausnahme: 368), vergleiche Abb. 48). Für eine weitere Untersuchung dieser Ergebnisse würde sich empfehlen, das Zusammenspiel der verschiedenen Parameter der Boden-Routine bei veränderter PET zu analysieren.

Bei diesen Auswertungen wurde mit besonderem Interesse das Verhalten der beiden Gebiete 2471 und 2450 beobachtet. In diesen beiden Gebieten wird heute noch Wässermatten-Wirtschaft (z.B. Bieri, 1949; Leibundgut, 1976; Bieri, 1975) betrieben, was die Evapotranspiration beeinflussen dürfte. Im Wiggertal (Gebiet 2450) sind es zwar nur noch 15 ha, im Langete- und Rottal (Gebiet 2471) dagegen noch ca. 100 ha, die als Wässermatten bewirtschaftet werden (BAK, 2012). Ein aussergewöhnliches Verhalten dieser beiden Gebiete konnte allerdings nicht beobachtet werden. Die Wässermattenflächen stellen aber einen möglichen Auslöser für den hohen BETA-Wert des Gebietes 2471 dar.

Aufgrund der in Abb. 50 dargestellten Ergebnisse lässt sich ableiten, dass HBV besser mit über- als unterschätzten PET-Werten umgehen kann. Es würde sich daher anbieten, die mit den üblichen Methoden berechneten PET-Werte um den Faktor 1.5 bis 2 zu erhöhen, um bessere Modellergebnisse zu erzielen.

Eine weitere Überlegung ist wert, ob mit einer detaillierter gestalteten Berechnung der AET in HBV bessere Ergebnisse erzielt werden könnten. HBV nutzt standardmässig die PET zur Berechnung der AET in der Boden-Routine, wobei allerdings verschiedene Komponenten der Evapotranspiration nicht unterschieden werden – Prozesse wie Interzeption, Transpiration und Evaporation über offenen Wasserflächen werden also integriert betrachtet. Auch etwaige Interaktionen mit Grundwasser oder Windgeschwindigkeiten werden nicht berücksichtigt. Aus diesem Grund testeten Lindroth (1993) und

Joukainen (2000) detailliertere Formen der Evapotranspirations-Implementierung in HBV. Ihre Versuche, anstelle der PET direkt die AET zu verwenden, verlangte die Einführung vieler zusätzlicher Parameter, erzielte aber keine merkliche Modellverbesserung. Diese Resultate erhärten die Vermutung, dass eine Modellierung auf Einzugsgebietsebene durch eine allzu detaillierte physikalische Beschreibung eher behindert als gefördert wird.

7.5. Co-Kalibrierung auf Grundwasserdaten

Im letzten Schritt dieser Arbeit wurden 13 Gebiete mit Abfluss- und Grundwasserdaten co-kalibriert. Das damit erzielte (kombinierte) Gütemass fällt zwar in allen Gebieten etwas tiefer aus, die Abnahme ist aber derart gering, dass die dadurch erreichte bessere interne Modellstimmigkeit als ebenso gewichtig angenommen werden kann.

Allgemein kann beobachtet werden, dass co-kalibrierte Simulationen im Niedrigwasserbereich tendenziell höhere Abflüsse erzeugen als das Standardmodell (und das Delay-Modell). Damit werden Dauer und Defizit tendenziell unterschätzt, der Minimalabfluss überschätzt und der Onset zu spät vorhergesagt. Die Auswertungen der Schwellenwertmethode weisen daher auf eine nicht zu vernachlässigende Modellverschlechterung hin; ebenso zeigen die Auswertungen der Kontingenztabellen eine deutliche Verschlechterung der POD und FOM, hingegen eine Verbesserung der FAR. Weil in der Praxis die POD und insbesondere die FOM aber stärker ins Gewicht fallen als die FAR, weisen diese Resultate darauf hin, dass eine Co-Kalibrierung mit Grundwasserdaten für die hier untersuchten Ziele nicht vorteilhaft ist – mit Ausnahme einzelner Gebiete (z.B. 354). Betrachtet man allerdings die Abflusskurven visuell, so zeigt sich das Bild nicht so deutlich negativ wie mit der Schwellenwertmethode. Ausserdem darf nicht ignoriert werden, dass die tendenziell höher liegende co-kalibrierte Simulation bei einer zu tief liegenden Simulation des Standard-Modells zu einer Modellverbesserung führt (z.B. Abb. 58). Es bleibt daher die Frage, ob die vermutlich gewonnene Modellstimmigkeit den Güteverlust aufzuwiegen vermag – schliesslich sollte eine grössere interne Konsistenz des Modells die Zuverlässigkeit von Prognosen erhöhen.

Als möglicher Ausweg aus diesem Dilemma kann erwogen werden, die Grundwasserdaten schwächer zu gewichten, beispielsweise mit 20 % oder 30 % an Stelle der hier verwendeten 40 %. Dadurch könnte unter Umständen eine bessere Modellstimmigkeit erreicht werden, ohne dass die Abflussdynamik zu stark verschlechtert wird. Allerdings wurde in dieser Arbeit, wie in Kap. 5.8.2 erläutert, eine Gewichtung mit nur 20 % verworfen, weil der zusätzlich ins Modell eingebrachte Informationsgehalt damit entsprechend gering wäre. Eine zu schwache Gewichtung lässt die Grundwasser-Co-Kalibrierung überflüssig werden.

Beim Delay-Modell scheint eine Co-Kalibrierung mit Grundwasserdaten eher vorteilhaft zu sein. Deutliche Verbesserungen zeigen sich dabei vor allem bei den beiden nicht-verkarsteten Gebieten. Dies deutet wiederum darauf hin, dass die Delay-Variante nicht spezifisch zur Simulierung von Karstlandschaften geeignet ist, sondern dass andere Gebietseigenschaften für die Eignung des Delay-Modells entscheidender sein müssen. Dieses Ergebnis könnte aber ebenfalls bedeuten, dass in karsthaltigen Gebieten ein Einbezug von Grundwasserdaten nicht sinnvoll ist.

Ein wichtiger Vorbehalt bei der Co-Kalibrierung mit Grundwasser war die Datenqualität. Der Grossteil der Grundwasserdaten stammt vom Kanton Aargau, wo sie leider nicht qualitätsgeprüft werden. Viele Stationen konnten daher wegen Qualitätsproblemen nicht verwendet werden. Bei den Verbleibenden wurden einzelne klar erkennbare Ausreisser und Abschnitte mit Versatz entfernt. Dennoch sind gegenüber der Datenqualität gewisse Vorbehalte angebracht. Im Einzugsgebiet 2450 wurden allerdings zwei Stationen des Kantons Aargau mit sechs qualitätsgeprüften Stationen des Kantons Luzern kombiniert. Die Stationen zeigten untereinander starke Korrelationen, wobei die

Messwerte des Kantons Aargau sich gut in das Verhalten der Luzerner Stationen einpassten. Die Daten können also ebenso wenig als schlecht bezeichnet werden, was ihre Verwendung durchaus rechtfertigt.

Wie in Kap. 7.4 muss auch hier auf die Wässermatten-Bewirtschaftung in den beiden Gebieten 2471 und 2450 hingewiesen werden. Die Wässerung der Wiesenflächen bewirkt nämlich eine direkte Reaktion der Grundwasserspiegel (z.B. Leibundgut, 1976; Leibundgut, 1980). Diese „Verfälschung“ des Grundwasserspiegels könnte dafür verantwortlich sein, dass in diesen beiden Gebieten kaum Verbesserungen durch den Einbezug der Grundwasserdaten erzielt werden konnten. Die visuelle Betrachtung der Grundwasserpegelkurven gab allerdings keinen Anlass, die Qualität bzw. Eignung der Daten anzuzweifeln.

Bereits mehrmals wurde angesprochen, was auch bei den Auswertungen der Co-Kalibrierungen festgestellt wurde: Die Schwellenwertmethode kann in gewissen Fällen sehr stark über die tatsächliche Güte der Simulation täuschen. Ein markantes Beispiel dafür ist Abb. 56. Es lässt sich daraus ableiten, dass die Schwellenwertmethode zwar durchaus zur Bewertung eines Modells hilfreich sein kann, aber nicht als einzige Methode angewandt werden sollte.

Wie bereits in Kap. 5.8.1 erläutert, besteht der eigentliche Zweck einer Multi-Kriterien-Kalibrierung darin, die interne Konsistenz des Modells zu verbessern. Um beurteilen zu können, ob dies mit der Co-Kalibrierung auf Abfluss- und Grundwasserdaten tatsächlich erreicht wurde, müsste hier eine erneute Sensitivitätsanalyse durchgeführt werden. Aus zeitlichen Gründen war dies leider nicht möglich. Daher verbleibt diese Arbeit gestützt auf die Literatur bei der Annahme, dass die Parameterstreuung eingegrenzt werden konnte. Auf jeden Fall hat die Co-Kalibrierung bewirkt, dass sich die Grundwasser-Routine realitätsnaher verhält.

Die Multi-Kriterien-Analyse liesse sich jedoch in verschiedene Richtungen weiterentwickeln: Neben der hier durchgeführten Co-Kalibrierung von Abfluss- und Grundwasserdaten wurde bereits der Einbezug von Schneemessdaten als weitere, sinnvolle Möglichkeit angesprochen. Das neu programmierte HBV-GW-Modell würde diese Möglichkeit bereits anbieten. Eine weitere Möglichkeit stellt die Berücksichtigung von Bodenfeuchte- bzw. Saugspannungswerten dar: Van Loon & Van Lanen (2012) verglichen HBV-Simulationen mit Messwerten der Bodenfeuchte und erhielten eine zufriedenstellende Übereinstimmung zwischen gemessenen und simulierten Werten (ohne dass eine Co-Kalibrierung durchgeführt wurde). Eine Co-Kalibrierung könnte das Modellverhalten in diesem Bereich verbessern. Dabei könnte künftig für die untersuchten Gebiete auf das Bodenmessnetz Nordwestschweiz (www.bodenmessnetz.ch) zurückgegriffen werden. Im Kanton Aargau wurden die ersten Messstationen im Rahmen dieses Projektes Ende 2012 in Betrieb genommen. Weil aber keine älteren Daten vorliegen, wurde diese Variante der Multi-Kriterien-Analyse für vorliegende Arbeit nicht in Betracht gezogen.

7.6. Weitere Unsicherheiten

Die Ergebnisse dieser Arbeit sind mit verschiedenen Unsicherheiten behaftet, die von der Wahl der Methoden, vom Modell und von den verwendeten Daten herrühren. Beispielsweise kann die Güte des verwendeten kombinierten Gütemasses hinterfragt werden. Es hat sich nämlich gezeigt, dass ein gutes Kalibrierungs-Ergebnis noch keine perfekten Simulationen garantiert (siehe Kap. 6.1.2.6). Im Gegenzug haben aber Gebiete mit einer schlechten Modellgüte in der Kalibrierung oft auch in den Auswertungen der Schwellenwertmethode versagt. Beispielsweise wurde die grösste Überschätzung des Minimums in beiden Auswertungsperioden durch das Einzugsgebiet 368 erzeugt, dessen Simulationen mit dem Standardmodell in beiden Fällen einen konstanten Abflusswert lieferten. Dieses Resultat verwundert aber nicht in Hinblick auf das Kalibrierungsergebnis von nur gerade

0.538. Es konnte aber auch ein Zusammenhang zwischen der erreichten Modellgüte, der Einzugsgebietsgrösse und dem Karstanteil festgestellt werden: Alle grösseren Einzugsgebiete erreichten recht bis sehr gute Gütewerte und kein Gebiet ohne Karstanteil erzielte hingegen ein sehr schlechtes Kalibrierungsergebnis. Die Umkehrschlüsse sind hingegen nicht gültig.

Zur Verbesserung des Gütemasses könnten der von Pushpalatha et al. (2012) empfohlene Reff der invertierten Abflusswerte oder die von Krause et al. (2005) als gut befundene Form des Reff relativ zu den Messwerten erwogen werden. Diese beiden Masse würden ein noch stärkeres Gewicht auf Niedrigwasser legen, sind allerdings in HBV bisher nicht implementiert.

Ferner kann die Eignung der Schwellenwertmethode zur Beurteilung der Modellgüte hinterfragt werden. Bereits mehrmals wurde erwähnt, dass der visuelle Eindruck der Simulationskurven sich von Ergebnissen der Schwellenwertmethode unterscheidet. Beispielsweise würde man aufgrund der visuellen Betrachtung die Simulationen des Standard-Modells im 2011 tendenziell als besser einstufen als die Simulationen im 2003, die Trockenheitsmasse und Kontingenztabellen suggerieren aber eher das Gegenteil. Für die der Schwellenwertmethode schwierig zu erfassen sind Situationen, in denen beobachtete Werte und Simulationen direkt um den Schwellenwert schwanken. Ein Grundproblem beim Vergleich verschiedener Simulationen ist, bei *Missed Events* und *False Alarms* Verbesserungen oder Verschlechterungen bestimmen zu können – schliesslich ist es möglich, dass die Simulation sich zwar besser an den beobachteten Wert annähert, aber dennoch knapp ober- bzw. unterhalb des Schwellenwertes verbleibt und so einen *Missed Events* oder *False Alarms* simuliert. Es stellt sich daher die grundsätzliche Frage, ob den Kennwerten der Schwellenwertmethode oder dem visuellen Eindruck mehr Gewicht beigemessen werden soll. Es erscheint jedenfalls empfehlenswert, eine Schwellenwert-Auswertung mit einer visuellen Inspektion der Abflusskurven zu ergänzen.

Unsicherheiten sind auch bei den verwendeten Daten vorhanden: Zu den Abflussdaten wurden in Kap. 4.1.4 Kleinkraftwerke und Hochwasserrückhaltebecken als Quellen möglicher Fehlinformationen besprochen, die allerdings als vernachlässigbar eingestuft wurden. Bei den Grundwasserdaten des Kantons Aargau sind gewisse Vorbehalte zur Datenqualität anzubringen. Eine weitere entscheidende Quelle für Unsicherheiten stellen die meteorologischen Daten dar. Auffällige Punkte sind hier die räumliche Auflösung der Rasterzellen, die im Vergleich zu den kleinsten untersuchten Einzugsgebieten sehr gering ist (für das Gebiet 345 mussten gar die Werte des benachbarten Einzugsgebietes verwendet werden, weil das Gebiet kleiner als eine Rasterzelle ist und die Datenextraktion vermutlich aus diesem Grund nicht funktionierte) und die räumliche Anordnung der Messstationen. So befindet sich beispielsweise das Gebiet 368 weit von der nächsten Messstation entfernt, was gerade für ein solch kleines Gebiet (5.2 km²) grossen Einfluss bei konvektiven Wetterlagen mit lokalen Niederschlagsereignissen haben kann. Fraglich ist zudem die Güte der meteorologischen Daten im östlichen Jura, wo sich sämtliche umliegende Messstationen im Talboden befinden. Ein Bias der Daten ist daher trotz topografischer Korrektur zu erwarten.

In Zusammenhang mit den meteorologischen Daten stehen ferner die beiden Parameter Pelev und Telev, die eine starke Auswirkung auf die Modellgüte zeigten. Pelev und Telev sind eng verbunden mit den Parametern Pcalt, CFMAX und SFCF, die ebenfalls grossen Einfluss auf die Modellgüte nahmen. Pelev bzw. Telev stehen eigentlich für die Höhe der meteorologischen Messstationen, wurden aber als mittlere Höhe des Einzugsgebietes (mit einigen Anpassungen) festgelegt. Die Erklärung für diese Wahl liegt in der Tatsache, dass kaum nachvollzogen werden kann, welche Messstation zu welchem Zeitpunkt wie stark in die Berechnung des Rasterdatensatzes einfluss. Das Messnetz ist zudem über die Zeit nicht konstant. Pelev und Telev lassen sich daher aufgrund der umliegenden Messstationen nicht präzise bestimmen. Die mittlere Einzugsgebietshöhe wird deshalb als gute Alternative betrachtet.

Ein weiterer, grundsätzlich wichtiger Punkt in Bezug auf die Daten ist die Frage nach der Stationarität der Zeitreihen (Kap. 4.6). Für 14 Gebiete wurden signifikante Werte in einem oder mehreren der durchgeführten Tests auf Trend, Einheitswurzel und (Niveau-)Stationarität erzielt und damit als nicht-stationär identifiziert. Für das Gebiet 340 beispielsweise könnte diese Eigenschaft das komplette Versagen der Simulationen trotz überdurchschnittlich gutem Kalibrierungsergebnis erklären. Sechs dieser Gebiete (336, 349, 357, 359, 369, 370) erreichten dagegen ein nicht befriedigendes Kalibrierungs-Resultat – die Nicht-Stationarität der Datenreihen könnte der Auslöser dafür sein. Bei verschiedenen dieser Gebiete kommen überdies andere in dieser Arbeit untersuchte kritische Merkmale hinzu: So zeigen die Gebiete 357, 359 und 369 neben dem schlechten Kalibrierungsergebnis und den signifikanten Werten für Nicht-Stationarität auch ein tiefes Q50 und einen hohen BETA-Wert. Gebiet 359 fällt zusätzlich mit sehr hohen TT-Werten und einem ungewöhnlichen Muster der Abfluss-Jahresmittelwerte auf.

Signifikante Werte für Nicht-Stationarität können allerdings nicht in allen Fällen als Erklärung dienen: Gebiet 368 wies keine signifikanten Werte auf, zeigte aber in beiden Auswertungsperioden eine Simulation auf konstantem Niveau. Auffällig sind hier der hohe BETA-Wert und das tiefe Q50. Die grosse Entfernung zur nächstgelegenen meteorologischen Messstation könnte bei diesem Gebiet aber den entscheidenden Grund für das Modellversagen darstellen. Auch das Gebiet 353 fiel in den Auswertungen immer wieder negativ auf, zeigt aber als einzige Auffälligkeiten hohe Werte für BETA und Pcalt. Schliesslich erwies sich das Gebiet 345 als nicht unproblematisch, was aufgrund der geringen Einzugsgebietsgrösse und der 100-prozentigen Verkarstung nicht erstaunlich ist; hinzu kommen ein geringer Q50- und ein hoher BETA-Wert.

Wenn nicht-stationäre Datenreihen zur Modellierung verwendet werden, muss bedacht werden, dass die resultierenden Parameter gemittelte Werte und keine festen Charakteristiken des Einzugsgebietes darstellen. Auslöser für die Untersuchungen auf Stationarität bildete nämlich die Beobachtung an verschiedenen Einzugsgebieten, dass „Qdiff“ (die kumulierte Differenz zwischen simuliertem und beobachtetem Abfluss) systematische Fehler aufwies. An Einzugsgebiet 331 konnte beispielsweise beobachtet werden, dass Qdiff vor dem Jahr 2000 konstant stark positiv, nach 2000 aber konstant negativ war (Gebiet 331 zeigte auch signifikante Werte für Nicht-Stationarität). Diese Beobachtung weist auf eine Veränderung der zugrundeliegenden hydrologischen Prozesse hin. Der Einbezug des Volumenfehlers in die Kalibrierung des Modells bewirkt hier, dass der Fehler über den gesamten Kalibrierungszeitraum ausgemittelt wird. Dies bedeutet wiederum, dass auch die Parameter in „gemittelten“ Werten resultieren. Ob solche Parameter dennoch eine gute Prognose erzeugen können, hängt von der Grösse der Abweichungen ab.

7.7. Anwendbarkeit und Möglichkeiten zur Weiterentwicklung

Zum Abschluss stellt sich daher die Frage nach der Anwendbarkeit des Modells für operationelle Prognosen. Um die Fähigkeit des HBV-Modells zu testen, Vorhersagen ausserhalb des Kalibrierungszeitraumes zu erstellen, hat diese Arbeit ein Split-Sample-Verfahren angewandt. Die verwendeten Messdaten symbolisieren dabei den Idealfall von fehlerfreien Prognosedaten. Wenn aber mit echten Prognosedaten gearbeitet würde, ist eine Reduktion der Modellverlässlichkeit aufgrund der neu eingebrachten Unsicherheiten zu erwarten. Wie zuverlässig die Abflussprognosen dann ausfallen würden, ist stark von der Güte der meteorologischen Prognosedaten abhängig. Alternativ könnten statistische Verfahren zur Nachbesserung der Prognosen angewandt werden (siehe z.B. Hemri, Fundel, & Zappa, 2013).

Bei einer operationellen Nutzung des Modells könnte allerdings ein Teil der durch die meteorologischen Prognosen zusätzlich eingeführten Unsicherheit mit der bereits erwähnten

Echtzeit-Aufdatierung ausgeglichen werden (siehe z.B. Wöhling, Lennartz, & Zappa, 2006; Sene, 2010: 124ff; Cloke & Pappenberger, 2009). Damit würden die Prognosen deutlich an Wert gewinnen, denn die Qualität der Ausgangsbedingungen ist für die Güte von Prognosen von entscheidender Bedeutung (Zappa et al., 2012: 26; Fundel et al., 2013: 404). Durch dieses Verfahren würden gleichzeitig systematische Shifts, wie sie in den Simulationen dieser Arbeit beobachtet werden konnten, auskorrigiert.

Die Ergebnisse dieser Arbeit lassen erwarten, dass die Erstellung von guten kurz- bis mittelfristigen Prognosen im Bereich des Möglichen liegt. Hingegen ist anzunehmen, dass längerfristige Abflussprognosen (gemäss WMO (2010) wird zwischen „extended-range“ mit bis zu 30 Tagen und „long-range“ mit bis zu zwei Jahren unterschieden) für mittel- bis kleinskalige Einzugsgebiete, wie sie hier untersucht wurden, nicht sinnvoll ausgeführt werden können, da Raum und Zeit voneinander abhängig sind und räumliche Kleinskaligkeit auch zeitliche Kleinskaligkeit bedingt (Buttle, 1998: 35). Das Potential hierfür müsste jedoch separat getestet werden.

8. Schlussfolgerung und Ausblick

Aufgrund der Ergebnisse dieser Arbeit kann davon ausgegangen werden, dass HBV das Potential für aussagekräftige Prognosen in den untersuchten Gebieten aufweist und damit einen hilfreichen Beitrag im Umgang mit Niedrigwasserereignissen leisten könnte. Eine Umsetzung von kurz- bis mittelfristigen Prognosen erscheint möglich; hingegen ist anzunehmen, dass längerfristige Abflussprognosen für mittel- bis kleinskalige Einzugsgebiete, wie sie hier untersucht wurden, nicht durchführbar sind. Das Potential hierfür müsste jedoch separat getestet werden.

Allerdings muss betont werden, dass diese Arbeit sozusagen den Idealfall von fehlerfreien Prognosedaten getestet hat – sobald aber mit echten Prognosedaten gearbeitet wird, ist eine Reduktion der Modellzuverlässigkeit aufgrund der zusätzlich eingebrachten Unsicherheiten zu erwarten. Wie verlässlich die Abflussprognosen dann ausfallen würden, ist stark von der Güte der meteorologischen Prognosedaten abhängig. Mit statistischen Verfahren und/oder Echtzeit-Aufdatierung der Abflussprognosen könnten diese zusätzlich eingebrachten Unsicherheiten aber mindestens teilweise ausgeglichen werden.

Unabhängig der verwendeten Modellvariante wurden deutliche Unterschiede in der Güte der Niedrigwasservorhersagen zwischen den beiden Auswertungsperioden beobachtet. Dies erstaunt nicht angesichts der nicht identischen Charakteristiken dieser beiden Ereignisse. Die Beobachtung wirft allerdings die Vermutung auf, dass saisonal implementierte Parameter die Modellgüte deutlich verbessern könnten. Diese Option ist in HBV-light bisher nicht verfügbar. Zu bedenken ist aber, dass sich mit einer saisonal variablen Parameterimplementierung die Anzahl der zu kalibrierenden Variablen mit einem Schlag vervierfachen würde. HBV punktet aber gerade damit, dass es nicht überparametrisiert ist. Ob eine solche Modellentwicklung daher erstrebenswert ist, lässt sich diskutieren.

In Bezug auf die Multi-Kriterien-Analyse liessen sich weitere Möglichkeiten ausmachen: Eine Co-Kalibrierung mit Schneedaten lässt erwarten, die Parameter-Unsicherheit stark eingrenzen können. Die Verfügbarkeit der benötigten Daten dürfte dabei allerdings problematisch sein. Eine weitere Möglichkeit böte die Co-Kalibrierung auf Bodenfeuchte- bzw. Saugspannungswerte. Dies sind verschiedene denkbare Möglichkeiten. Ob es sinnvoller ist, nur eine einzelne all dieser Grössen für eine Co-Kalibrierung mit einzubeziehen oder ob eine echte „Multi“-Kriterien-Kalibrierung mit verschiedenen dieser Werte bessere Resultate erzielt, müsste untersucht werden. Auf jeden Fall wäre es wichtig, diejenige Messgrösse, die den für HBV wertvollsten Informationsgehalt aufweist, zu identifizieren.

Weiteres Potential zur Verbesserung der Modellvorhersagen läge in einer höheren räumlichen Auflösung der meteorologischen Daten. Hier könnte auch die Nutzung von Radardaten erwogen werden.

In diesem Sinne vermittelt diese Arbeit einen ersten Eindruck der Vorhersagefähigkeit des HBV-Modells für Niedrigwasserereignisse in den untersuchten Gebieten. Wie aufgezeigt gibt es dabei verschiedene Aspekte, die weiter verfolgt oder genauer untersucht werden sollten. Ein Potential zur Erstellung von aussagekräftigen operationellen Prognosen ist aufgrund der Ergebnisse dieser Arbeit zu erwarten. Möglichkeiten zur Entwicklung und Verbesserung in verschiedener Hinsicht sind vorhanden und wurden besprochen. Als Fazit lässt sich damit eine positive Bilanz ziehen.

Literatur

- Abebe, N. A., Ogden, F. L., & Pradhan, N. R. (2010). Sensitivity and uncertainty analysis of the conceptual HBV rainfall-runoff model: Implications for parameter estimation. *Journal of Hydrology*, 389(3-4), 301–310. doi:10.1016/j.jhydrol.2010.06.007
- Addor, N., Jaun, S., Fundel, F., & Zappa, M. (2011). An operational hydrological ensemble prediction system for the city of Zurich (Switzerland): skill, case studies and scenarios. *Hydrology and Earth System Sciences*, 15(7), 2327–2347. doi:10.5194/hess-15-2327-2011
- Aschwanden, H., & Kan, C. (1999). Die Abflussmenge Q347 - Eine Standortbestimmung. In *Hydrologische Mitteilungen LHG*. Bern: Bundesamt für Umwelt BAFU. Aufgerufen von <http://www.bafu.admin.ch/publikationen/publikation/00397/index.html?lang=de>
- Asztalos, J. (2004). *Ein Schnee- und Eisschmelzmodell für vergletscherte Einzugsgebiete*. Technische Universität Wien, Technische Universität Wien.
- BAFU (Bundesamt für Umwelt). (1992). *Hydrologischer Atlas der Schweiz*. (Landeshydrologie und -geologie, Ed.). Bern: EDMZ.
- BAFU (Bundesamt für Umwelt). (1997). Restwassermengen in Fließgewässern - Wasserentnahmen, die insbesondere der Bewässerung dienen. *Mitteilungen Zum Gewässerschutz*, 24, 47.
- BAFU (Bundesamt für Umwelt). (2004). Auswirkungen des Hitzesommers 2003 auf die Gewässer. *Umwelt*, 369.
- BAFU (Bundesamt für Umwelt). (2011a). Mai 2011: Anhaltende Trockenheit - Wasserstände bleiben trotz Regen-Intermezzo aussergewöhnlich tief. 18.05.2011. Aufgerufen am 23. Januar 2014, von <http://www.bafu.admin.ch/hydrologie/01834/02041/11404/index.html?lang=de>
- BAFU. (2011b). Trockenheit Frühling 2011: Auswirkungen der Trockenheit auf Natur und Umwelt. 19.05.2011. Aufgerufen am 23. Januar 2014, von <http://www.bafu.admin.ch/dokumentation/fokus/11095/11332/index.html?lang=de>
- BAFU (Bundesamt für Umwelt). (2012a). *Auswirkungen der Klimaänderung auf Wasserressourcen und Gewässer - Synthesebericht zum Projekt «Klimaänderung und Hydrologie in der Schweiz» (CCHydro)* (p. 78).
- BAFU (Bundesamt für Umwelt). (2012b). *Umgang mit lokaler Wasserknappheit in der Schweiz* (p. 87). Aufgerufen von <http://www.bafu.admin.ch/dokumentation/medieninformation/00962/index.html?lang=de&msgid=46701>
- BAK (Bundesamt für Kultur). (2012). Wässermatten. *Lebendige Traditionen*.
- Bergström, S. (1976). Development and Application of a Conceptual Runoff Model for Scandinavian Catchments. *SMHI Rapport / Hydrologi Och Oceanografi*, 7.
- Berner, P. (2011). Zuständigkeiten und Vorgehen für Massnahmen des Kantons bei lang anhaltender Trockenheit und Wassermangel in landwirtschaftlichen Kulturen. Kantonsinternes Papier.
- Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, 320, 18–36. doi:10.1016/j.jhydrol.2005.07.007
- Beven, K. (2012). *Rainfall-Runoff Modelling: The Primer* (Second Edi., p. 457). John Wiley&Sons Ltd.
- Beven, K., & Binley, A. (1992). The Future of Distributed Models: Model Calibration and Uncertainty Prediction. *Hydrological Processes*, 6, 279–298.
- Beven, K., & Young, P. (2013). A guide to good practice in modeling semantics for authors and referees. *Water Resources Research*, 49, 1–7. doi:10.1002/wrcr.20393
- Bieri, W. (1949). Die Wässermatten von Langenthal. *Mitteilungen Der Naturforschenden Gesellschaft in Bern. Neue Folge*, 6, 115–126.
- Bieri, W. (1975). Wässermatten - Reminiszenzen. *Jahrbuch Des Oberaargaus*, 18, 138–142.

- Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., & Schär, C. (2013). Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections. *Water Resources Research*, *49*, 1–14. doi:10.1029/2011WR011533
- Buttle, J. M. (1998). Fundamentals of Small Catchment Hydrology. In C. Kendall & J. J. McDonnell (Eds.), *Isotope Tracers in Catchment Hydrology* (p. 839). Amsterdam: Elsevier Science B.V.
- CH2011. (2011). *Swiss Climate Change Scenarios CH2011* (p. 88). Zurich, Switzerland: C2SM, MeteoSwiss, ETH, NCCR Climate, and OcCC. ISBN: 978-3-033-03065-7
- Clausen, B., & Rasmussen, K. R. (1993). Low Flow Estimation and Hydrogeology in a Chalk Catchment. *Nordic Hydrology*, *24*, 297–308.
- Cloke, H. L., & Pappenberger, F. (2009). Ensemble flood forecasting: A review. *Journal of Hydrology*, *375*(3-4), 613–626. doi:10.1016/j.jhydrol.2009.06.005
- Davie, T. (2008). *Fundamentals of Hydrology* (2nd Edition). Oxon and New York.
- Dawson, C. W., Abrahart, R. J., & See, L. M. (2007). HydroTest: A web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. *Environmental Modelling & Software*, *22*(7), 1034–1052. doi:10.1016/j.envsoft.2006.06.008
- Demirel, M. C., Booij, M. J., & Hoekstra, A. Y. (2013). Effect of different uncertainty sources on the skill of 10 day ensemble low flow forecasts for two hydrological models. *Water Resources Research*, *49*(7), 4035–4053. doi:10.1002/wrcr.20294
- Doswell III, C. A., Davies-Jones, R., & Keller, D. L. (1990). On Summary Measures of Skill in Rare Event Forecasting Based on Contingency Tables. *Weather and Forecasting*, *5*, 576 – 585.
- Fleig, A. K., Tallaksen, L. M., Hisdal, H., & Demuth, S. (2006). A global evaluation of streamflow drought characteristics. *Hydrology and Earth System Sciences*, *10*, 535–552.
- Fuhrer, J. (2010). Bewässerungsbedürftigkeit heute und in Zukunft. *Geomatik Schweiz*, *10*, 448–449.
- Fundel, F., Jörg-Hess, S., & Zappa, M. (2013). Monthly hydrometeorological ensemble prediction of streamflow droughts and corresponding drought indices. *Hydrology and Earth System Sciences*, *17*, 395–407. doi:10.5194/hess-17-395-2013
- Gardelin, M., Bergström, S., Carlsson, B., Graham, L. P., & Lindström, G. (2002). Climate change and water resources in Sweden - analysis of uncertainties. In M. Beniston (Ed.), *Climatic Change: Implications for the Hydrological Cycle and for Water Management* (pp. 189–207). Kluwer Academic Publishers.
- Gebhardt, H., Glaser, R., Radtke, U., & Reuber, P. (Eds.). (2007). *Geographie - Physische Geographie und Humangeographie*. Heidelberg: Spektrum Akademischer Verlag.
- Hartmann, A., Lange, J., Vivó Aguado, À., Mized, N., Smiatek, G., & Kunstmann, H. (2012). A multi-model approach for improved simulations of future water availability at a large Eastern Mediterranean karst spring. *Journal of Hydrology*, *468-469*, 130–138. doi:10.1016/j.jhydrol.2012.08.024
- Hartmann, A., Wagener, T., Rimmer, A., Lange, J., Brielmann, H., & Weiler, M. (2013). Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. *Water Resources Research*, *49*, 3345–3358. doi:10.1002/wrcr.20229
- Hemri, S., Fundel, F., & Zappa, M. (2013). Simultaneous calibration of ensemble river flow predictions over an entire range of lead times. *Water Resources Research*, *49*(10), 6744–6755. doi:10.1002/wrcr.20542
- Holland, J. H. (1975). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence* (First Edit., p. 183). Oxford, England: University of Michigan Press.
- Jensen, M. E., & Haise, H. R. (1963). Estimating evapotranspiration from solar radiation. *Journal of Irrigation and Drainage Division, ASCE* *89*(LR4), 15–41.
- Joukainen, S. (2000). Improving the calculation of potential evapotranspiration of the HBV model: application to the Ounasjoki watershed. In T. Nilsson (Ed.), *Nordic hydrological conference 2000. Nordic Hydrological Programme Report* (pp. 347–354). Uppsala: Swedish Hydrological Council.

- Kay, A. L., & Davies, H. N. (2008). Calculating potential evaporation from climate model data: A source of uncertainty for hydrological climate change impacts. *Journal of Hydrology*, 358(3-4), 221–239. doi:10.1016/j.jhydrol.2008.06.005
- Kempf, S. (2012). Evaluating the use of forecast and observed climatological data for simulating hydrological droughts in Switzerland. *Master Thesis*.
- Kendall, M. G. (1938). A New Measure of Rank Correlation. *Biometrika*, 30(1/2), 81–93.
- Klemes, V. (1986). Operational testing of hydrological simulation models. *Journal of Hydrological Science*, 31(1), 13–24.
- Krause, P., Boyle, D. P., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 5, 89–97.
- Kruse, S., Seidl, I., & Stähli, M. (2010). Informationsbedarf für die Früherkennung von Trockenheit in der Schweiz - Die Sicht potenziell betroffener Nutzergruppen. *Wasser Energie Luft*, 102(4), 305–308.
- Kundzewicz, Z. W., Mata, L. J., Arnell, N. W., Döll, P., Jimenez, B., Müller, K., Oki, T., Sen, Z., Shiklomanov, I. (2008). The implications of projected climate change for freshwater resources and their management management. *Hydrological Sciences Journal*, 53(1), 3–10. doi:http://dx.doi.org/10.1623/hysj.53.1.3
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54, 159–178. doi:10.1016/0304-4076(92)90104-Y
- Leibundgut, C. (1976). Zum Wasserhaushalt des Oberaargaus und zur hydrologischen Bedeutung des landwirtschaftlichen Wiesenbewässerungssystems im Langetental. *Beiträge Zur Geologie Der Schweiz - Hydrologie*, 23, 107. Aufgerufen von http://chy.scnatweb.ch/downloads/Nr.23_000.pdf
- Leibundgut, C. (1980). Wässermatten und Grundwasserspeisung - Hydrologische Folgen einer Nutzungsänderung und ihre Folgen für die Regionalplanung. *Jahrbuch Des Oberaargaus*, 23, 205–234.
- Lindroth, A. (1993). Potential Evaporation - A Matter of Definition. *Nordic Hydrology*, 24, 359–364.
- Lindström, G., Johansson, B., Persson, M., Gardelin, M., & Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, 201, 272–288.
- Luo, L., & Wood, E. F. (2007). Monitoring and predicting the 2007 U.S. drought. *Geophysical Research Letters*, 34(22), L22702. doi:10.1029/2007GL031673
- Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, 13(3), 245–259.
- McGuinness, J. L., & Bordne, E. F. (1972). *A comparison of lysimeter- derived potential evapotranspiration with computed values*. (Agricultural Research Service, Ed.) *Technical Bulletin* (Technical ., Vol. 1452). Washington, DC: US Department of Agriculture.
- McMahon, T. a., Peel, M. C., Lowe, L., Srikanthan, R., & McVicar, T. R. (2013). Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis. *Hydrology and Earth System Sciences*, 17(4), 1331–1363. doi:10.5194/hess-17-1331-2013
- Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Zbigniew, W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity Is Dead: Whither Water Management? *Science*, 139, 573–574.
- Milly, P. C. D., Dunne, K. a, & Vecchia, a V. (2005). Global pattern of trends in streamflow and water availability in a changing climate. *Nature*, 438(7066), 347–50. doi:10.1038/nature04312
- Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, 391(1-2), 202–216. doi:10.1016/j.jhydrol.2010.07.012
- Mishra, A. K., & Singh, V. P. (2011). Drought modeling – A review. *Journal of Hydrology*, 403(1-2), 157–175. doi:10.1016/j.jhydrol.2011.03.049
- Monteith, J. L. (1965). *Evaporation and the environment - The state and movement of water in living organisms. XIXth symposium*. Swansea: Cambridge University Press.
- Morgenschweis, G. (2010). *Hydrometrie - Theorie und Praxis der Durchflussmessung in offenen Gerinnen* (p. 582). Berlin: Springer-Verlag.

- Nash, J. E., & Sutcliffe, J. V. (1970). River Flow Forecasting through Conceptual Models. Part I - A Discussion of Principles. *Journal of Hydrology*, *10*, 282–290.
- Olsson, J., & Lindström, G. (2008). Evaluation and calibration of operational hydrological ensemble forecasts in Sweden. *Journal of Hydrology*, *350*, 14–24. doi:10.1016/j.jhydrol.2007.11.010
- Oudin, L., Andréassian, V., Mathevet, T., Perrin, C., & Michel, C. (2006). Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations. *Water Resources Research*, *42*(W07410), 10. doi:10.1029/2005WR004636
- Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., & Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model? Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. *Journal of Hydrology*, *303*(1-4), 290–306. doi:10.1016/j.jhydrol.2004.08.026
- Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London*, *193*, 120–145.
- Pfaff, B. (2008). *Analysis of Integrated and Cointegrated Time Series with R* (Second Edi., p. 188). New York, NY: Springer Science+Business Media, LLC.
- Phillips, P., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335–346.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical Recipes in C++: The Art of Scientific Computing Second Edition* (2nd Editio.). Cambridge: Cambridge University Press.
- Pushpalatha, R., Perrin, C., Moine, N. Le, & Andréassian, V. (2012). A review of efficiency criteria suitable for evaluating low-flow simulations. *Journal of Hydrology*, *420-421*, 171–182. doi:10.1016/j.jhydrol.2011.11.055
- Riad, S., Mania, J., Bouchaou, L., & Najjar, Y. (2004). Predicting catchment flow in a semi-arid region via an artificial neural network technique. *Hydrological Processes*, *18*, 2387–2393. doi:10.1002/hyp.1469
- Rutkowska, A., & Ptak, M. (2012). On Certain Stationarity Tests For Hydrologic Series. *Studia Geotechnica et Mechanica*, *XXXIV*(1), 51–63.
- Schaefli, B., & Gupta, H. V. (2007). Do Nash values have value ? *Hydrological Processes*, *21*, 2075–2080. doi:10.1002/hyp.6825
- Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Häberli, C., Liniger, M. A., & Appenzeller, C. (2004). The role of increasing temperature variability in European summer heatwaves. *Nature*, *427*(January), 332–336. doi:10.1038/nature02230.1.
- Schlegel, T., Kunz, H., & Gerstgasser, D. (2011, May 5). Aktuelles zum Wettergeschehen - Trockenheit Jahresbeginn 2011. *MeteoSchweiz*.
- Schultz, G. A. (1994). Meso-scale modelling of runoff and water balances using remote sensing and other GIS data. *Hydrological Sciences Journal*, *39*(2), 121–142. doi:10.1080/02626669409492729
- Seibert, J. (1997). Estimation of Parameter Uncertainty in the HBV Model. *Nordic Hydrology*, *28*(4/5), 247–262.
- Seibert, J. (2000). Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. *Hydrology and Earth System Sciences*, *4*(2), 215–224. doi:doi:10.5194/hess-4-215-2000
- Seibert, J., Bishop, K., Rodhe, A., & McDonnell, J. J. (2003). Groundwater dynamics along a hillslope: A test of the steady state hypothesis. *Water Resources Research*, *39*(1), 1014. doi:10.1029/2002WR001404
- Seibert, J., & McDonnell, J. J. (2002). On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. *Water Resources Research*, *38*(11), 23–1–23–14. doi:10.1029/2001WR000978
- Seibert, J., Rodhe, A., & Bishop, K. (2003). Simulating interactions between saturated and unsaturated storage in a conceptual runoff model. *Hydrological Processes*, *17*, 379–390. doi:10.1002/hyp.1130
- Seibert, J., & Vis, M. J. P. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences*, *16*(9), 3315–3325. doi:10.5194/hess-16-3315-2012

- Sene, K. (2010). *Hydrometeorology - Forecasting and Applications* (p. 356). Dordrecht, Heidelberg, London, New York: Springer Science+Business Media B.V.
- Seneviratne, S. I., Orth, R., Jörg-Hess, S., Kruse, S., Seidl, I., Stähli, M., Zappa, M., Seibert, J., Staudinger, M., Stahl, K., Weiler, M. (2013). Trockenheit in der Schweiz. *AQUA & GAS*, 9, 38–47.
- Shaw, S. B., & Riha, S. J. (2011). Assessing temperature-based PET equations under a changing climate in temperate, deciduous forests. *Hydrological Processes*, 25(9), 1466–1478. doi:10.1002/hyp.7913
- Smakhtin, V. . (2001). Low Flow Hydrology: A Review. *Journal of Hydrology*, 240, 147–186. doi:10.1016/S0022-1694(00)00340-1
- Solmatine, D. P., & Wagener, T. (2011). 2.16 Hydrological Modeling. In P. Wilderer (Ed.), *Reference Module in Earth Systems and Environmental Sciences - Treatise on Water Science* (pp. 435–457). Oxford: Elsevier B.V. Retrieved from <http://dx.doi.org/10.1016/B978-0-444-53199-5.00044-0>
- Son, K., & Sivapalan, M. (2007). Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data. *Water Resources Research*, 43(W01415), 18. doi:10.1029/2006WR005032
- Swisstopo. (2005). Geologische Karte der Schweiz 1 :500 000.
- Tallaksen, L. M., & van Lanen, H. A. J. (2004). *Hydrological Drought - Processes and Estimation Methods for Streamflow and Groundwater*. (L. M. Tallaksen & H. A. J. van Lanen, Eds.) *Hydrological Drought - Processes and Estimation Methods for Streamflow and Groundwater. Development in Water Science (48)* (Vol. 48). Amsterdam, San Diego, Oxford, London: Elsevier B.V.
- Uhlenbrook, S., Seibert, J., Leibundgut, C., & Rodhe, A. (1999). Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrological Sciences Journal*, 44(5), 779–797. doi:10.1080/02626669909492273
- Van Loon, A. F., & Van Lanen, H. A. J. (2012). A process-based typology of hydrological drought. *Hydrology and Earth System Sciences*, 16(7), 1915–1946. doi:10.5194/hess-16-1915-2012
- Vaze, J., Post, D. a., Chiew, F. H. S., Perraud, J.-M., Viney, N. R., & Teng, J. (2010). Climate non-stationarity – Validity of calibrated rainfall–runoff models for use in climate change studies. *Journal of Hydrology*, 394, 447–457. doi:10.1016/j.jhydrol.2010.09.018
- Von Storch, H. (1995). Misuses of Statistical Analysis in Climate. In H. V. Storch & A. Navarra (Eds.), *Analysis of Climate Variability: Applications of Statistical Techniques* (pp. 11–26). New York: Springer-Verlag.
- Wang, Q. J. (1991). The Genetic Algorithm and Its Application to Calibrating Conceptual Rainfall-Runoff Models. *Water Resources Research*, 27(9), 2467–2471. doi:10.1029/91WR01305
- Wang, W., van Gelder, P. H. A. J. M., & Vrijling, J. K. (2003). Detection of changes in streamflow series in western Europe over 1901 – 2000. *Water Science and Technology: Water Supply*, 5(6), 289–299.
- Wang, X. L., & Swail, V. R. (2001). Changes of Extreme Wave Heights in Northern Hemisphere Oceans and Related Atmospheric Circulation Regimes. *Journal of Climate*, 14(10), 2204–2221. doi:10.1175/1520-0442(2001)014<2204:COEWHI>2.0.CO;2
- Wilby, R. L., & Harris, I. (2006). A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK. *Water Resources Research*, 42, 1–10. doi:10.1029/2005WR004065
- Wilhite, D. A., & Glantz, M. H. (1985). Understanding the Drought Phenomenon : The Role of Definitions. *Water International*, 10(3), 111–120.
- Wilks, D. S. (2006). *Statistical Methods in the Atmospheric Sciences* (Second Edi., p. 630). Elsevier Inc.
- WMO (World Meteorological Organisation). (2000). Detecting Trend and Other Changes in Hydrological Data. In Z. W. Kundzewicz & A. Robson (Eds.), *World Climate Programme -Data and Monitoring (WMO/TD-No., p. 158)*. Geneva.
- WMO (World Meteorological Organisation). (2008). Manual on Low-flow Estimation and Prediction. *Operational Hydrology Report, No. 50*(WMO-No. 1029), 136.
- WMO (World Meteorological Organisation). (2010). Manual on the Global Data-processing and Forecasting System. *Wheater, Climate, Water, WMO-No. 48*.

- Wöhling, T., Lennartz, F., & Zappa, M. (2006). Technical Note: Real-time updating procedure for flood forecasting with conceptual HBV-type models. *Hydrology and Earth System Sciences Discussions*, 10, 783–788. doi:10.5194/hessd-3-925-2006
- Wölfli, M. (2011). Kanton stellt Bauern das Wasser ab. *Aargauer Zeitung*, p. 29.
- Yevjevich, V. (1967). An Objective Approach to Definitions and Investigations of Continental Hydrologic Droughts. *Hydrology Papers, Colorado State University*, (August).
- Yue, S., Pilon, P., Phinney, B., & Cavadias, G. (2002). The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological Processes*, 16(9), 1807–1829. doi:10.1002/hyp.1095
- Yue, S., & Wang, C. Y. (2002). Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test. *Water Resources Research*, 38(6), 4–1–4–7. doi:10.1029/2001WR000861
- Zappa, M. (2011). Aussergewöhnliche Trockenheit im 2011 - Neue Vorhersagesysteme für Wasserressourcen. *Newsletter SLF*, (3), 4–5.
- Zappa, M., Bernhard, L., Fundel, F., & Jörg-Hess, S. (2012). Vorhersage und Szenarien von Schnee- und Wasserressourcen im Alpenraum. *Forum Für Wissen*, 19–27.
- Zhang, X., & Lindström, G. (1997). Development of an Automatic Calibration Scheme for the HBV Hydrological Model. *Hydrological Processes*, 11, 1671–1682.
- Zhang, X., & Zwiers, F. W. (2004). Comment on “Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test” by Sheng Yue and Chun Yuan Wang. *Water Resources Research*, 40(3), n/a–n/a. doi:10.1029/2003WR002073
- Zivot, E., & Wang, J. (2006). Unit Root Tests. In *Modeling Financial Time Series with S-PLUS* (2nd Edition, pp. 111–139). New York, NY: Springer Science+Business Media, Inc. Aufgerufen von <http://faculty.washington.edu/ezivot/index.htm>

Abbildungsverzeichnis

Abb. 1: Darstellung der HBV-Struktur bzw. der verschiedenen Routinen und deren Parameter. Quelle: http://www.geo.uzh.ch/en/units/h2k/services/hbv-model (Zugriff 28.1.2014)	15
Abb. 2: HBV-Modellstruktur in systematischer Darstellung. © M. Staudinger	17
Abb. 3: Algenbewuchs an der Messschwelle der Sissle bei Eiken (EZ 331) bei Niedrigwasser im Sommer 2006. 20	
Abb. 4: Die beiden hier relevanten Abflussregime (gemäss BAFU, 1992).	21
Abb. 5: Pardé-Koeffizienten aller untersuchten Einzugsgebiete. Anhand dieser Koeffizienten wird das hydrologische Regime charakterisiert.	21
Abb. 6 (vorherige Seite): Übersicht der untersuchten Einzugsgebiete. Mehrere Gebiete sind Teilgebiete anderer Einzugsgebiete, daher sind diejenigen Gebiete, die andere Einzugsgebiete enthalten, fein rot umrandet. Teilgebiete sind zusätzlich farblich abgehoben. Die Einzugsgebiete, die durch den Kanton Aargau bemessen werden, sind rötlich eingefärbt, die beiden durch das BAFU bemessenen Gebiete sind gelblich eingefärbt. 24	
Abb. 7: Lage der verwendeten Grundwasser-Messstationen.	27
Abb. 8: Einzugsgebiet 348 mit seinen Grundwasser-Messstationen. In Rot ist die berechnete mittlere Grundwasserkurve dargestellt. In schwarz hinterlegt ist die Abflusskurve. Die Grundwasserspiegel sind normiert, um visuelle Vergleichbarkeit zu erhalten; die Skala fehlt daher.	30
Abb. 9: Lage der meteorologischen Messstationen im und um das Untersuchungsgebiet.	31
Abb. 10: Illustration der Begriffe Dauer und Defizit einer Trockenperiode. Quelle: http://www.drought.ch/Klima/Runoff/index_DE (Zugriff: 21.1.2014)	39
Abb. 11: Die Aare bei Brugg am 8. April 2011. Grosse Flüsse wie die Aare werden in dieser Arbeit zwar nicht berücksichtigt, das Ausmass der Trockenheit wird an ihnen aber sehr eindrücklich sichtbar.	42
Abb. 12: GAP-Einstellungen am Beispiel des Einzugsgebietes 331.	48
Abb. 13: Lage der verkarsteten Gebiete gemäss GIS-Layer des Kantons Aargau, der die Grundlage für Einschränkungen im Einsatz triazinhaltiger Herbizide bildet.	52
Abb. 14: Struktur der Modellvariante "Response Routine With Delay". Der Anteil an Wasser, der in die Grundwasser-Routine gelangt, wird in zwei Teile aufgespalten: Der eine Teil (bestimmt durch den Parameter PART [-]) bildet den „schnellen“ Abfluss, der andere Teil wird über einen bestimmten Zeitraum (Parameter DELAY [d]) verzögert und kommt daher erst später zum Abfluss. Die Aufteilung in verschiedene SUZ-Boxen entsteht durch die Verwendung verschiedener räumlicher Einheiten wie z.B. Höhenstufen oder Vegetationszonen. Quelle: HBV-Help	53
Abb. 15: Struktur der Modellvariante "One GW Box". Anstelle von zwei getrennten Boxen wie im Standardmodell wird hier für die Grundwasser-Routine nur eine einzige Box mit drei Ausflüssen verwendet. Uhlenbrook et al. (1999) sehen diese Struktur als mögliche Repräsentation eines nahe an der Oberfläche liegenden Aquifers, der mit den Oberflächengewässern in engem Kontakt steht. Quelle: HBV-Help	53
Abb. 16: Struktur der Modellvariante "Three GW Boxes". Uhlenbrook et al. (1999) sehen diese Struktur als mögliche Repräsentation für mehrere, übereinander liegende Aquifere und hangende Grundwasservorkommen. Quelle: HBV-Help	53
Abb. 17: Korrelation zwischen gemessenem Grundwasserspiegel [m.ü.M.] und den HBV-Grundwasser-Boxen SUZ und SLZ [mm]. Links sind die Ergebnisse aus der Standard-Kalibrierung abgebildet, rechts nach der Co-Kalibrierung auf die Grundwasserdaten. Die Korrelation wurde dadurch sichtlich verstärkt.	55
Abb. 18: Boxplot der Modellgüten aller Einzugsgebiete.	57
Abb. 19: Räumliche Darstellung der mit der Standardversion erzielten Modellgüten. Auffällig ist, dass sich die drei Gebiete mit den schlechtesten Güteergebnissen alle im Aargauer Jura mit Ausfluss gegen die Aare hin befinden.	58
Abb. 20: Vergleich der Simulationen mit den beobachteten Werten für die Trockenheits-Dauer im 2003 (links) und 2011 (rechts) für die beiden Schwellenwerte Q_{347} (oben) und RW_{min} (unten).	59
Abb. 21: Gebiete Nr. 353 (oben) und 361 (unten) in der Auswertungsperiode 2011.	60
Abb. 22: Vergleich der Simulationen mit den beobachteten Werten für das Trockenheits-Defizit im 2003 (links) und 2011 (rechts). Achtung: Auf den Grafiken für den Schwellenwert RW_{min} (unten) fehlt aus Gründen der Darstellbarkeit das Einzugsgebiet 345.	61

Abb. 23: Gebiet 344 in der Auswertungsperiode 2011. Im Sommer sackt die Simulation zu stark ab.....	62
Abb. 24: Gebiet 344 in der Auswertungsperiode 2011. Die Simulation bleibt auf zu hohem Niveau.	62
Abb. 25: Gebiet 378 in der Auswertungsperiode 2003. Die Simulation sinkt ab Ende Juni zu stark ab.	63
Abb. 26: Vergleich der Simulationen mit den beobachteten Werten für die Trockenheits-Intensität im 2003 (links) und 2011 (rechts). Achtung: Auf den Grafiken für den Schwellenwert RWmin (unten) fehlt aus Gründen der Darstellbarkeit das Einzugsgebiet 345.	63
Abb. 27: Gebiet 342 (oben) in der Auswertungsperiode 2011 und Gebiet 364 (unten) in der Auswertungsperiode 2003. Beide zeigen für die Intensität eine bessere Vorhersage als für das Defizit.	64
Abb. 28: Gebiet 372 in der Auswertungsperiode 2011. Hier zeigen Defizit und Intensität ein interessantes Verhalten.	65
Abb. 29: Vergleich der Differenzen im Onset für die beiden Schwellenwerte Q347 und RWmin. Die Differenz in Anzahl Tagen wurde berechnet als simuliertes Datum – beobachtetes Datum. Positive Werte stehen damit für einen verspäteten simulierten Onset, negative für einen verfrühten.	65
Abb. 30: Gebiet 361 in der Auswertungsperiode 2003. RWmin liegt allgemein deutlich über dem Abfluss, wird aber von einer Abflussspitze im August überschritten.....	66
Abb. 31: Vergleich der Simulationen mit den beobachteten Werten für den minimalen Abfluss in den beiden Auswertungsperioden 2003 (links) und 2011 (rechts).	66
Abb. 32: Simulationen des Gebietes 332 für die Auswertungsperioden 2003 (oben) und 2011 (unten). Die beiden Grafiken zeigen deutlich, wie unterschiedlich sich die Simulation in den beiden Perioden verhält...	67
Abb. 33: Simulationen des Gebietes 374 für die Auswertungsperioden 2003 (oben) und 2011 (unten). Die Simulation 2003 sinkt zu stark ab, während im 2011 die beobachteten Werte nahezu perfekt getroffen werden.	68
Abb. 34: Erreichtes kombiniertes Gütemass im Vergleich mit dem Anteil an Karstflächen im Einzugsgebiet.	70
Abb. 35: Erreichtes kombiniertes Gütemass im Vergleich mit der Grösse des Einzugsgebietes.	70
Abb. 36: Entwicklung der Modellgüte mit dem Verlauf der GAP-Berechnung für die Einzugsgebiete 332 (oben) und 352 (unten).	71
Abb. 37: Parameterstreuung für 10 Kalibrierungen (Analyse 1). Jeder Punkt repräsentiert den Variationskoeffizienten eines Einzugsgebietes.	72
Abb. 38: Streuung der TT-Werte für 10 Kalibrierungen (Analyse1) pro Einzugsgebiet.....	72
Abb. 39: Parameter-Variation für 5 Kalibrierungen (Analyse 2). Jeder Punkt repräsentiert den Variationskoeffizienten eines Einzugsgebietes.	73
Abb. 40: Streuung der TT-Werte für 5 Kalibrierungen (Analyse2) pro Einzugsgebiet.....	74
Abb. 41: Vergleich zwischen je 100 Kalibrierungen mit (in dunkelblau) und ohne (in hellblau) Einbezug der viermonatigen Auswertungsperioden im 2003 und 2011 (Einzugsgebiet 337) für Januar bis September 2003 (oben) und 2011 (unten) exemplarisch dargestellt. In schwarz ist der gemessene Abfluss abgebildet. Zu beachten ist die logarithmische y-Achse!.....	75
Abb. 42: Einzugsgebiet 345 in der Auswertungsperiode 2011. Qobs = gemessener Abfluss, Qsim_std = Standard-HBV-Modell simulierten Abfluss, Qsim_Del für die Simulation durch die Delay-Variante und Qsim_3GW für die Simulation durch die Variante mit drei Grundwasserboxen („Three GW Boxes“).	77
Abb. 43: Vergleich der Simulationen des Delay-Modells (Sim_Del) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für die Trockenheits-Dauer im 2003 (links) und 2011 (rechts) für die beiden Schwellenwerte Q ₃₄₇ (oben) und RWmin (unten).	78
Abb. 44: Vergleich der Simulationen des Delay-Modells (Sim_Del) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für das Trockenheits-Defizit im 2003 (links) und 2011 (rechts). Achtung: Auf den Grafiken für den Schwellenwert RWmin (unten) fehlt aus Gründen der Darstellbarkeit das Einzugsgebiet 345.....	79
Abb. 45: Vergleich der Simulationen des Delay-Modells (Sim_Del) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für die Trockenheits-Intensität im 2003 (links) und 2011 (rechts) Achtung: Auf den Grafiken für den Schwellenwert RWmin (unten) fehlt aus Gründen der Darstellbarkeit das Einzugsgebiet 345.....	80

Abb. 46: Vergleich der Simulationen des Delay-Modells (Sim_Del) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für den minimalen Abfluss in den beiden Auswertungsperioden 2003 (links) und 2011 (rechts).....	81
Abb. 47: Vergleich der Differenzen im Onset zwischen dem Delay-Modell bzw. dem Standard-Modell und dem beobachteten Wert für die beiden Schwellenwerte Q347 und RWmin. Die Differenz in Anzahl Tagen wurde als simuliertes Datum – beobachtetes Datum berechnet. Positive Werte stehen damit für einen verspäteten simulierten Onset, negative für einen verfrühten.....	82
Abb. 48: Zusammenhang zwischen erreichter Modellgüte (kombiniertes Gütemass) und dem Wert Q50.	83
Abb. 49: Simulation (Qsim) durch das Standard-HBV-Modell in der Auswertungsperiode 2003. Der Abfluss befindet sich auf extrem niedrigen Werten und fällt zeitweise sogar auf 0. In den Messwerten (Qobs) ist eine deutlich Plafonierung erkennbar, die vermutlich auf die Grenzen der Messtechnik zurückzuführen sind.	84
Abb. 50: Beeinflussung der Modellgüte durch Variieren der potentiellen Evapotranspiration (PET) an 15 Einzugsgebieten. Die x-Achse gibt den Faktor an, mit welchem die ursprüngliche PET verändert wurde (Faktor 1 entspricht also dem Ausgangswert). Die obere Grafik zeigt Gebiete, deren BETA-Wert bei der Standardkalibrierung über 5 lagen; die BETA-Werte der Gebiete der unteren Grafik lagen alle unter 5; den in der Standardkalibrierung tiefsten überhaupt erreichten BETA-Wert zeigte EZ 337. Zur besseren Vergleichbarkeit entsprechen sich die y-Skalen.	85
Abb. 51: Vergleich der Simulationen des HBV-GW-Modells (Sim_GW) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für die Trockenheits-Dauer im 2003 (oben) und 2011 (unten) für die beiden Schwellenwerte Q ₃₄₇ (links) und RWmin (rechts).	87
Abb. 52: Vergleich der Simulationen des HBV-GW-Modells (Sim_GW) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für das Trockenheits-Defizit im 2003 (oben) und 2011 (unten).....	88
Abb. 53: Vergleich der Simulationen des HBV-GW-Modells (Sim_GW) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für die Trockenheits-Intensität im 2003 (oben) und 2011 (unten).....	89
Abb. 54: Vergleich der Simulationen des HBV-GW-Modells (Sim_GW) mit den Simulationen des Standardmodells (Sim) und den beobachteten Werten (Obs) für den minimalen Abfluss in den beiden Auswertungsperioden 2003 (links) und 2011 (rechts).	90
Abb. 55: Vergleich der Differenzen im Onset zwischen dem HBV-GW-Modell (mit Suffix _GW) bzw. dem Standard-Modell (ohne Suffix) und dem beobachteten Wert für die beiden Schwellenwerte Q347 und RWmin. Die Differenz in Anzahl Tagen wurde als simuliertes Datum – beobachtetes Datum berechnet. Positive Werte stehen damit für einen verspäteten simulierten Onset, negative für einen verfrühten.	90
Abb. 56: HBV-GW-Simulation (Qsim_GW) für das Einzugsgebiet 353 in der Auswertungsperiode 2011 im Vergleich zur Standard-Simulation (Qsim_std) und der Delay-Simulation (Qsim_delay). Die HBV-GW-Simulation ist deutlich besser, bleibt aber knapp über den Schwellenwerten und schneidet daher als <i>Missed Event</i> schlechter ab als die anderen Simulationen.....	91
Abb. 57: HBV-GW-Simulation für das Einzugsgebiet 360 in der Auswertungsperiode 2003. Dies ist eines von zwei Gebieten, in denen HBV-GW tiefere Niedrigwasserabflüsse als das Standardmodell erzeugt.....	92
Abb. 58: HBV-GW-Simulation für das Einzugsgebiet 354 in der Auswertungsperiode 2003. Diese Simulation ist ein typisches Beispiel dafür, dass HBV-GW in fast allen Gebieten höhere Niedrigwasserabflüsse berechnet als das Standardmodell.	92
Abb. 59: Gebiet 360 in der Auswertungsperiode 2011. Vergleich der Standard-Simulation, der Delay-Simulation, und der entsprechenden Simulationen, die auf Grundwasser co-kalibriert wurden. Qsim_DelGW steht für das co-kalibrierte Delay-Modell und ist bei diesem Gebiet deutlich schlechter als die übrigen Simulationen.	93
Abb. 60: Vergleich der Standard- und Delay-Simulationen mit und ohne Grundwasser-Co-Kalibrierung für das Gebiet 348 in der Auswertungsperiode 2011. In diesem Fall bildet die Grundwasser-co-kalibrierte Variante die beste Modellierung.	94

Abb. 61: Vergleich der Standard- und Delay-Simulationen mit und ohne Grundwasser-Co-Kalibrierung für das Gebiet 353 in der Auswertungsperiode 2003. Hier ist eine Entscheidung, welche Variante am besten ist, schwierig..... 94

Abb. 62: Die Sissle bei Eiken (EZ 331) am 1. Juli 2008. Wie der Schmittenbach (EZ 369) ist auch die Sissle mehrmals im Verlauf der gesamten Zeitreihen komplett ausgetrocknet. 98

Tabellenverzeichnis

Tab. 1: Liste aller verwendeten Stationen. Befinden sich mehrere Stationen am selben Gewässer, so sind sie in derselben Farbe eingefärbt. Die Mindestrestwassermenge ist gelb markiert, wenn sie durch die Untergrenze gemäss GSchG Art. 31, Abs. 1 festgesetzt ist. Achtung: Für HBV müssen $[m^3/s]$ in $[mm/d]$ umgerechnet werden. 24

Tab. 2: Zusammenstellung der verwendeten Einzugsgebiete und ihrer Grundwasser-Messstationen. Angegeben sind neben Nummer, Name und Standortgemeinde der Stationen die Höhenlage der Terrain-Oberkante, die euklidische Distanz zum nächsten Gewässer (gemäss dem Gewässernetz von VECTOR25 von swisstopo), der Spearman-Korrelationskoeffizient zwischen Grundwasserpegel und Abfluss sowie die euklidische Distanz zur Abfluss-Messstation. Messstationen des Kantons Luzern sind in blauer Schrift, solche des Kantons Bern in roter Schrift markiert..... 28

Tab. 3: Ergebnisse (P-Werte) der verschiedenen durchgeführten Stationaritäts-Tests. Autokorrelierte Daten (siehe Spalte „Auto-Korr.“) wurden zusätzlich mit einem Prewhitening (PW) korrigiert. 37

Tab. 4: Prinzip der Kontingenztabelle. 40

Tab. 5: Zur Modellkalibrierung verwendete Zielfunktionen und ihre Gewichtung..... 47

Tab. 6: Zur Kalibrierung verwendete Parameter-Grenzen. In Klammern sind einzeln verwendete Minimal- bzw. Maximalwerte aufgeführt. 48

Tab. 7: Zur Analyse der Parameter-Sensitivität (Analyse 2) verwendete kombinierte Gütemasse und ihre Gewichtung. 50

Tab. 8: Gewichtung der verschiedenen Zielfunktionen für die Co-Kalibrierung auf Abfluss und Grundwasser... 56

Tab. 9: Kontingenztabelle für die Auswertungsperiode im 2003. E= Event, NE = No Event. 69

Tab. 10: Kontingenztabelle für die Auswertungsperiode im 2011. E= Event, NE = No Event. 69

Tab. 11: Gütemasse der je 100 Kalibrierungen mit und ohne Einbezug der besagten Auswertungsperioden..... 75

Tab. 12: Ergebnisse der Kalibrierungen mit der HBV-Delay-Variante im Vergleich zu den Standard-Simulationen. 76

Tab. 13: Auswertung der Kontingenztabelle des Delay-Modells im Vergleich zum Standardmodell. Die Werte des Standardmodells beziehen sich nur auf die Gebiete, für die auch das Delay-Modell verwendet wurde. 81

Tab. 14: Ergebnisse der Kalibrierungen mit der HBV-GW-Variante im Vergleich zum Standard-Modell. 86

Tab. 15: Auswertung der Kontingenztabelle des HBV-Standardmodells und des HBV-GW-Modells im Vergleich. 93

Tab. 16: Resultate (Parameterwerte und Gütemasse) der Kalibrierung mit dem Standard-Modell 118

Tab. 17: Zusammenstellung der Schwellenwert-Auswertung für die Auswertungsperiode 2003..... 120

Tab. 18: Zusammenstellung der Schwellenwert-Auswertung für die Auswertungsperiode 2011..... 121

Tab. 19: Auswertung des berechneten Einsetzens der Trockenheit sowie des absoluten Minimums für die Auswertungsperiode 2003. 122

Tab. 20: Auswertung des berechneten Einsetzens der Trockenheit sowie des absoluten Minimums für die Auswertungsperiode 2011. 123

Tab. 21: Auswertung der Trockenheits-Dauer für das Delay-Modell, je für die Auswertungsperioden 2003 und 2011 und für die Schwellenwerte Q347 und Mindestrestwassermenge (RWmin). Obs = beobachteter Wert, Std Sim = Modellierung mit dem Standard-Modell (Vergleichswert), Delay sim = Modellierung mit dem Delay-Modell. Zusätzlich sind jeweils die relative Abweichung zwischen der Delay-Simulation und dem beobachteten Wert sowie eine qualitative Bewertung im Vergleich zur Standardsimulation angegeben. . 124

Tab. 22: Auswertung des Trockenheits-Defizits für das Delay-Modell. Für Erläuterungen siehe Tab. 21. 125

Tab. 23: Auswertung der Trockenheits-Intensität für das Delay-Modell. Für Erläuterungen siehe Tab. 21..... 126

Tab. 24: Auswertung des Trockenheits-Onsets für das Delay-Modell. Für Erläuterungen siehe Tab. 21. Die Abweichung zum beobachteten Wert ist hier absolut angegeben.	127
Tab. 25: Auswertung des Abfluss-Minimums für das Delay-Modell. Für Erläuterungen siehe Tab. 21. Die Abweichung zum beobachteten Wert ist hier absolut angegeben.	128
Tab. 26: Auswertung der Trockenheits-Dauer für das HBV-GW-Modell, je für die Auswertungsperioden 2003 und 2011 und für die Schwellenwerte Q347 und Mindestrestwassermenge (RWmin). Obs = beobachteter Wert, Sim = Modellierung mit dem Standard-Modell (Vergleichswert), GW sim = Modellierung mit Co-Kalibrierung auf Grundwasserstände. Zusätzlich sind jeweils die relative Abweichung zwischen der GW-Simulation und dem beobachteten Wert sowie eine qualitative Wertung der GW-Simulation im Vergleich zur Standardsimulation angegeben.....	129
Tab. 27: Auswertung des Trockenheits-Defizits für das HBV-GW-Modell. Für Erläuterungen siehe Tab. 26.....	130
Tab. 28: Auswertung der Trockenheits-Intensität für das HBV-GW-Modell. Für Erläuterungen siehe Tab. 26..	131
Tab. 29: Auswertung des Trockenheits-Onsets für das HBV-GW-Modell. Für Erläuterungen siehe Tab. 26. Die Abweichung zum beobachteten Wert ist hier in Tagen angegeben.	132
Tab. 30: Auswertung des Trockenheits-Minimums für das HBV-GW-Modell. Für Erläuterungen siehe Tab. 26. Die Abweichung zum beobachteten Wert ist hier absolut angegeben.....	133

Verzeichnis vermerkter Internetquellen

Testplattform Drought CH: www.drought.ch / www.trockenheit.ch (letzter Zugriff: 16.6.2014).....	8
Zugriff HBV-Modell: http://www.geo.uzh.ch/en/units/h2k/services/hbv-model (letzter Zugriff: 16.6.2014).....	15
Abflussdaten des BAFU: www.hydrodaten.admin.ch (letzter Zugriff: 16.6.2014).....	19
Abflussdaten des Kantons Aargau: www.ag.ch/hydrometrie (letzter Zugriff: 16.6.2014).....	19
Informationen zu Hochwasserrückhaltebecken: https://www.ag.ch/de/bvu/umwelt_natur_landschaft/hochwasserschutz/hochwasserschutzmassnahmen/hochwasserrueckhaltebecken/hochwasserrueckhaltebecken_1.jsp (letzter Zugriff: 16. 6. 2014).....	25
Grundwasserdaten Kanton Aargau: https://www.ag.ch/de/bvu/umwelt_natur_landschaft/umweltinformationen/wasser_2/grundwasser_2/grundwasser_1.jsp (letzter Zugriff 21.5.2014).....	26
Grundwasserdaten Kanton Luzern: http://www.uwe.lu.ch/index/themen/gewaesser/hydrometrie/grundwasserstand.htm (letzter Zugriff: 21.5.2014).....	26
Grundwasserdaten Kanton Luzern: http://www.bve.be.ch/de/index/wasser/wasser/grundwasser.html (letzter Zugriff: 21.5.2014).....	26
Datendokumentation der Niederschlagsdaten von MeteoSchweiz: http://www.meteoschweiz.admin.ch/web/de/services/datenportal/gitterdaten/precip/rhiresd.html (letzter Zugriff: 27.2.2014).....	31
Datendokumentation der Temperaturdaten von MeteoSchweiz: http://www.meteoschweiz.admin.ch/web/de/services/datenportal/gitterdaten/temp/tabsd.html (letzter Zugriff: 27.2.2014).....	31
Stationen des SwissMetNet: http://www.meteoschweiz.admin.ch/web/de/klima/messsysteme/boden/swissmetnet.html (letzter Zugriff: 30.5.2014).....	32
Geodatenportal Kanton Aargau: https://www.ag.ch/geoportal/geodatenshop/datensuche.aspx (letzter Zugriff: 1.6.2014).....	51
Bodenmessnetz Nordwestschweiz (Kantone Solothurn, Aargau und Baselland) www.bodenmessnetz.ch (letzter Zugriff: 16.6.2014).....	102

Anhang

Tab. 16: Resultate (Parameterwerte und Gütemasse) der Kalibrierung mit dem Standard-Modell

NR	Reff	LogReff	VolumeError	MAREMeasure	Weighted_Obj_Function	PERC	UZL	K0	K1	K2	MAXBAS	Cet
331	0.752422	0.802463	0.990766	0.614072	0.727097657	0.828818	21.58907	0.378168	0.174518	0.078289	2.085003	0.0558
332	0.680586	0.768205	0.989051	0.769309	0.790841558	1.512523	17.12811	0.425826	0.135244	0.013689	2.075005	0.067878
334	0.65734	0.769074	0.969433	0.727148	0.768146692	1.436114	17.84159	0.549448	0.189803	0.02422	2.1221	0.04935
336	0.623499	0.526688	0.950125	0.492981	0.552178026	1.334808	8.687277	0.539774	0.281442	0.009501	2.045172	0.120405
337	0.342024	0.676841	0.999978	0.841309	0.79138891	4.170523	25.87684	0.762266	0.133721	0.005538	1.964949	0.127718
340	0.655499	0.734909	0.999802	0.782585	0.785236303	2.120759	18.82368	0.68672	0.137248	0.005027	2.065018	0.096382
341	0.748136	0.732227	0.999824	0.583151	0.684448584	0.771658	25.80361	0.407529	0.14952	0.057008	2.090832	0.121397
342	0.698464	0.7935	0.921901	0.594514	0.706847172	1.349711	16.1187	0.48731	0.218948	0.082591	2.284998	0.037166
344	0.744989	0.811615	0.966131	0.709685	0.776101397	0.865018	14.81498	0.401794	0.150358	0.019033	2.164997	0.036428
345	0.041516	0.588997	0.71828	0.317562	0.466207891	0.264518	7.956261	0.6478	0.193387	0.002928	2.715005	0.096924
348	0.396493	0.652037	0.869035	0.622372	0.658904404	1.257976	16.05326	0.644146	0.222051	0.025608	2.235004	0.059532
349	0.213654	0.627208	0.924154	0.687103	0.68684994	1.338506	15.95243	0.707951	0.130559	0.008674	2.091406	0.067465
351	0.676193	0.758408	0.999943	0.713558	0.760085116	1.453968	17.10071	0.440932	0.156499	0.032869	2.185022	0.044938
352	0.676992	0.746373	0.999774	0.710669	0.753860711	1.349919	16.16567	0.592938	0.143235	0.024374	2.145056	0.070556
353	0.260734	0.674224	0.901017	0.634281	0.676931899	1.129154	13.83268	0.499993	0.136357	0.024216	2.10558	0.027993
354	0.680774	0.675475	0.976959	0.702191	0.718981452	1.676833	17.55091	0.55485	0.125145	0.015959	2.125082	0.041283
355	0.703764	0.754768	0.99963	0.736682	0.77021096	1.909706	18.82524	0.355359	0.150343	0.018037	2.125123	0.101597
357	-1.52051	0.58356	0.557777	0.13467	0.356536853	0.107474	13.86949	0.672076	0.209936	0.007313	2.174994	0.10282
358	0.358195	0.664574	0.917571	0.707108	0.711140743	1.102535	15.42345	0.585917	0.116539	0.003031	2.094936	0.049414
359	-0.46279	0.318386	0.981837	0.132144	0.291610513	0.273225	6.341825	0.57755	0.266397	0.001159	2.524966	0.173687
360	0.318674	0.643394	0.926939	0.540032	0.620067311	0.862959	13.65433	0.675493	0.172688	0.008264	2.214971	0.056206
361	0.649143	0.771857	0.946307	0.611027	0.708887121	0.746142	15.41328	0.637549	0.188527	0.028846	2.057173	0.055468
362	0.199341	0.682269	0.860397	0.682848	0.700371626	1.564274	13.57494	0.626923	0.184514	0.013035	2.115026	0.055263
364	0.599855	0.752274	0.940726	0.657085	0.723524918	0.980414	14.82771	0.711521	0.169547	0.035694	2.380449	0.043577
367	0.644909	0.71291	0.999999	0.669119	0.719723107	1.11532	14.42853	0.505601	0.214784	0.01763	2.044959	0.06677
368	0.003131	0.333322	1	0.610135	0.538396271	3.472196	57.72157	0.989967	0.037368	1.00E-05	1.87501	3.22E-07
369	-0.44088	0.510394	0.816439	0.033871	0.302737001	0.032546	15.14645	0.674609	0.188512	0.000537	2.394983	0.082633
370	0.109097	0.536602	0.998872	0.757148	0.693102162	1.59895	17.7863	0.944394	0.113571	0.001372	2.074374	0.175258
371	0.763708	0.808461	0.999659	0.704685	0.775692681	1.056728	17.71219	0.405117	0.167593	0.022898	2.07606	0.077073
372	0.729085	0.770163	0.998384	0.687683	0.751745017	1.612585	21.60931	0.2201	0.180465	0.02593	2.044927	0.119908
374	0.770127	0.813918	0.99999	0.753646	0.802389178	1.433986	18.53931	0.218256	0.126386	0.02907	2.008619	0.020846
376	-0.03792	0.37627	0.987774	0.718319	0.608444612	1.763368	14.90791	0.724804	0.109314	0.00332	2.094904	0.08617
377	0.740906	0.799311	0.999177	0.75214	0.795712022	1.410795	34.4724	0.314036	0.09113	0.021398	2.13622	0.052852
378	0.715053	0.782395	1	0.784427	0.805171554	1.948733	16.9892	0.557697	0.146159	0.014073	2.009445	0.056461
2450	0.529135	0.732984	0.928889	0.754923	0.763544217	1.675532	18.50568	0.367226	0.13139	0.016353	2.104954	0.052516
2471	0.351932	0.71257	0.968797	0.79868	0.781247535	2.785664	26.86068	0.591688	0.110828	0.006059	1.824968	0.076839

Fortsetzung Tab. 16

NR	PCALT	Pelev	Telev	TT	CFMAX	SFCF	CFR	CWH	FC	LP	BETA
331	11.9897	530	530	0.576051	3.51639	0.423133	0.184467	1.82E-09	270.5103	0.702646	6.731788
332	1.06E-08	530	530	0.29736	10.79159	0.344728	0.474609	0.286317	206.1686	0.754013	3.966715
334	1.76E-10	450	450	-0.28954	2.22426	0.364815	0.231644	3.69E-06	209.1563	0.759567	4.871478
336	4.88E-06	400	400	0.279982	6.340853	0.323669	0.150599	0.499749	180.358	0.658106	2.845537
337	14.60691	575	575	0.630489	3.499642	0.569094	0.003931	0.025416	172.4607	0.756193	1.909133
340	4.177959	450	450	0.164795	3.561609	0.520861	0.037828	0.051007	219.1213	0.832059	3.747157
341	17.69879	535	535	0.069345	4.143345	0.696194	0.426097	1.17E-10	269.3186	0.702966	5.467348
342	10	520	520	0.760002	3.415101	0.366299	0.449214	3.92E-07	243.6876	0.702514	5.952087
344	18.59456	575	575	0.75716	9.960557	0.258913	0.559111	5.12E-09	190.9058	0.677599	7.770256
345	7.71E-09	450	450	1.700005	14.99883	0.134897	0.316544	2.45E-05	303.5823	0.61823	19.99996
348	3.90E-07	500	500	1.30003	24.88997	0.17352	0.382894	0.000107	182.3227	0.607267	4.278817
349	2.47E-09	610	610	1.665161	11.03642	0.227122	0.5313	0.387695	163.4275	0.668745	4.943758
351	8.55E-06	500	500	-0.46593	1.503973	0.810478	0.190591	4.06E-07	237.9008	0.785589	5.152639
352	2.177474	640	640	0.139457	1.991212	0.77137	0.410992	1.75E-06	136.5306	0.811601	4.145098
353	19.66766	685	685	1.21	11.6173	0.215431	0.648785	0.725369	210.0891	0.677495	7.478746
354	17.8065	555	555	0.930016	14.95941	0.388401	0.601688	4.19E-05	166.9706	0.713912	4.75773
355	10.1078	580	580	0.952001	11.00161	0.287632	0.99749	0.263935	170.8563	0.697719	4.367072
357	3.52E-07	420	420	1.795301	11.99858	0.087158	0.359237	1.78E-05	200.745	0.485976	8
358	3.42E-06	450	450	0.854802	19.99826	0.287729	0.198667	3.50E-06	298.1184	0.667123	4.922362
359	3.49E-07	340	340	3.940006	11.99999	0.034996	0.496	7.52E-05	383.9274	0.200005	11.99998
360	4.44E-07	350	350	0.210402	11.99898	0.166346	0.450009	1.09E-05	189.4528	0.505596	4.579395
361	4.29E-06	350	350	-0.36963	2.497661	0.410419	0.346491	1.02E-09	208.7996	0.670613	4.987656
362	2.70E-10	400	400	1.593901	14.64182	0.152385	0.219344	0.000277	138.5221	0.520522	3.161526
364	3.35E-10	500	500	-0.62961	1.426429	0.596911	0.278717	0.005432	172.514	0.730199	6.088779
367	2.14E-06	450	450	0.013605	2.33374	0.494848	0.388275	2.98E-06	210.2017	0.807977	5.327618
368	0.328374	500	500	0.229978	8.311587	0.261613	0.525327	0.009013	321.1148	0.005912	11.99979
369	7.73E-10	350	350	1.984001	11.99984	0.117204	0.370055	5.09E-06	181.5806	0.175501	11.99999
370	3.84E-09	450	450	0.526115	14.99878	0.335261	0.033063	3.21E-05	666.8043	0.612501	3.328855
371	5.22E-05	595	595	-0.42167	1.747769	0.557521	0.467922	6.24E-09	176.5329	0.676092	4.080666
372	8.11E-09	500	500	0.118745	1.796786	0.471663	0.014229	0.112226	253.1983	0.624811	4.154684
374	20.18035	485	485	-0.59515	2.863302	0.808972	0.978884	0.009195	273.2127	0.763531	4.10279
376	3.27E-07	420	420	0.611136	18.3459	1.51E-09	0.395622	0.527969	955.9509	0.512917	2.155262
377	11.24641	685	685	0.479964	2.700888	1.145458	0.0063	0.05357	197.6999	0.768528	3.558424
378	24.2334	540	540	-0.048	6.280044	1.065102	0.281457	1.16E-06	178.0092	0.78638	2.675961
2450	1.35E-06	580	580	0.885701	24.64609	0.258627	0.241635	9.30E-10	179.5775	0.649252	3.974639
2471	26.30148	655	655	0.154391	13.42481	0.400242	0.000214	0.125208	272.9293	0.525308	19.99969

Tab. 17: Zusammenstellung der Schwellenwert-Auswertung für die Auswertungsperiode 2003.

EZ NR	Dauer						Defizit						Intensität					
	Q347			Rwmin			Q347			Rwmin			Q347			Rwmin		
	Obs [d]	Sim [d]	+/- [%]	Obs [d]	Sim [d]	+/- [%]	Obs [mm]	Sim [mm]	+/- [%]	Obs [mm]	Sim [mm]	+/- [%]	Obs [mm/d]	Sim [mm/d]	+/- [%]	Obs [mm/d]	Sim [mm/d]	+/- [%]
331	76	103	35.53	56	101	80.36	1.413	5.805	310.82	0.621	4.588	639.37	0.019	0.056	203.13	0.011	0.045	309.95
332	104	73	-29.81	62	15	-75.81	15.140	8.114	-46.41	1.741	0.219	-87.41	0.146	0.111	-23.65	0.028	0.015	-47.94
334	96	101	5.21	44	73	65.91	14.422	20.803	44.25	1.599	6.066	279.44	0.150	0.206	37.11	0.036	0.083	128.71
336	0	0	0.00	83	80	-3.61	0.000	0.000	0.00	5.817	7.703	32.42	0.000	0.000	0.00	0.070	0.096	37.38
337	43	41	-4.65	0	0	0.00	4.301	2.386	-44.52	0.000	0.000	0.00	0.100	0.058	-41.82	0.000	0.000	0.00
340	90	0	-100.00	4	0	-100.00	8.217	0.000	-100.00	0.054	0.000	-100.00	0.091	0.000	-100.00	0.014	0.000	-100.00
341	100	96	-4.00	120	114	-5.00	4.332	5.431	25.39	19.328	19.406	0.41	0.043	0.057	30.61	0.161	0.170	5.69
342	92	105	14.13	109	111	1.83	3.772	6.386	69.32	7.994	10.915	36.53	0.041	0.061	48.35	0.073	0.098	34.07
344	97	97	0.00	78	87	11.54	6.411	12.304	91.92	2.734	8.393	206.98	0.066	0.127	91.92	0.035	0.096	175.23
345	41	0	-100.00	122	122	0.00	0.351	0.000	-100.00	851.255	846.347	-0.58	0.009	0.000	-100.00	6.978	6.937	-0.58
348	98	103	5.10	39	79	102.56	10.826	17.690	63.39	1.048	5.967	469.57	0.110	0.172	55.46	0.027	0.076	181.18
349	87	78	-10.34	66	55	-16.67	10.323	8.112	-21.41	4.551	3.224	-29.15	0.119	0.104	-12.35	0.069	0.059	-14.98
351	59	84	42.37	99	106	7.07	5.140	12.206	137.46	19.519	28.537	46.20	0.087	0.145	66.79	0.197	0.269	36.54
352	83	95	14.46	116	122	5.17	14.729	13.252	-10.03	55.745	58.202	4.41	0.177	0.139	-21.39	0.481	0.477	-0.73
353	102	108	5.88	108	111	2.78	15.479	22.995	48.56	17.807	25.397	42.63	0.152	0.213	40.30	0.165	0.229	38.77
354	33	79	139.39	0	54	undef.	0.925	9.960	976.68	0.000	2.925	undef.	0.028	0.126	349.75	0.000	0.054	undef.
355	85	97	14.12	58	84	44.83	7.847	15.092	92.32	3.672	9.793	166.67	0.092	0.156	68.53	0.063	0.117	84.13
357	91	61	-32.97	122	122	0.00	1.493	0.602	-59.65	33.488	32.675	-2.43	0.016	0.010	-39.80	0.274	0.268	-2.43
358	81	0	-100.00	12	0	-100.00	7.304	0.000	-100.00	0.117	0.000	-100.00	0.090	0.000	-100.00	0.010	0.000	-100.00
359	0	0	0.00	122	122	0.00	0.000	0.000	0.00	40.429	35.436	-12.35	0.000	0.000	0.00	0.331	0.290	-12.35
360	50	42	-16.00	33	19	-42.42	2.370	1.305	-44.96	0.922	0.246	-73.28	0.047	0.031	-34.47	0.028	0.013	-53.59
361	101	95	-5.94	121	122	0.83	6.672	8.047	20.61	52.103	53.224	2.15	0.066	0.085	28.23	0.431	0.436	1.31
362	84	89	5.95	107	100	-6.54	5.285	10.405	96.87	10.258	15.210	48.28	0.063	0.117	85.81	0.096	0.152	58.66
364	79	90	13.92	80	96	20.00	7.918	11.519	45.48	10.439	14.506	38.96	0.100	0.128	27.70	0.130	0.151	15.80
367	88	87	-1.14	76	76	0.00	10.746	12.630	17.54	6.220	8.129	30.69	0.122	0.145	18.89	0.082	0.107	30.69
368	80	0	-100.00	122	122	0.00	5.445	0.000	-100.00	68.987	34.289	-50.30	0.068	0.000	-100.00	0.565	0.281	-50.30
369	87	0	-100.00	122	122	0.00	0.496	0.000	-100.00	38.039	36.749	-3.39	0.006	0.000	-100.00	0.312	0.301	-3.39
370	19	0	-100.00	3	0	-100.00	0.770	0.000	-100.00	0.081	0.000	-100.00	0.041	0.000	-100.00	0.027	0.000	-100.00
371	101	92	-8.91	105	96	-8.57	10.118	11.138	10.08	12.708	13.486	6.12	0.100	0.121	20.85	0.121	0.140	16.07
372	56	83	48.21	30	77	156.67	2.423	10.430	330.52	0.751	7.537	903.06	0.043	0.126	190.47	0.025	0.098	290.80
374	72	93	29.17	40	83	107.50	5.943	19.315	224.98	1.758	12.938	635.94	0.083	0.208	151.60	0.044	0.156	254.67
376	59	0	-100.00	3	0	-100.00	2.223	0.000	-100.00	0.037	0.000	-100.00	0.038	0.000	-100.00	0.012	0.000	-100.00
377	84	103	22.62	110	113	2.73	8.142	28.454	249.47	19.965	40.976	105.24	0.097	0.276	185.00	0.182	0.363	99.79
378	61	112	83.61	0	81	undef.	3.144	28.947	820.82	0.000	9.813	undef.	0.052	0.258	401.52	0.000	0.121	undef.
2450	81	78	-3.70	0	0	0.00	13.366	13.674	2.31	0.000	0.000	0.00	0.165	0.175	6.25	0.000	0.000	0.00
2471	71	28	-60.56	0	0	0.00	5.849	1.483	-74.64	0.000	0.000	0.00	0.082	0.053	-35.70	0.000	0.000	0.00

Tab. 18: Zusammenstellung der Schwellenwert-Auswertung für die Auswertungsperiode 2011.

EZ NR	Dauer						Defizit						Intensität					
	Q347			Rwmin			Q347			Rwmin			Q347			Rwmin		
	Obs [d]	Sim [d]	+/- [%]	Obs [d]	Sim [d]	+/- [%]	Obs [mm]	Sim [mm]	+/- [%]	Obs [mm]	Sim [mm]	+/- [%]	Obs [mm/d]	Sim [mm/d]	+/- [%]	Obs [mm/d]	Sim [mm/d]	+/- [%]
331	44	50	13.64	32	43	34.38	1.113	1.527	37.21	0.654	0.981	50.04	0.025	0.031	20.75	0.020	0.023	11.66
332	17	9	-47.06	0	0	0.00	0.386	0.066	-82.93	0.000	0.000	0.00	0.023	0.007	-67.76	0.000	0.000	0.00
334	45	61	35.56	0	0	0.00	2.983	5.565	86.54	0.000	0.000	0.00	0.066	0.091	37.61	0.000	0.000	0.00
336	96	0	-100.00	115	53	-53.91	10.787	0.000	-100.00	31.689	3.475	-89.03	0.112	0.000	-100.00	0.276	0.066	-76.20
337	92	55	-40.22	0	0	0.00	10.120	4.337	-57.15	0.000	0.000	0.00	0.110	0.079	-28.32	0.000	0.000	0.00
340	9	0	-100.00	0	0	0.00	0.666	0.000	-100.00	0.000	0.000	0.00	0.074	0.000	-100.00	0.000	0.000	0.00
341	40	41	2.50	72	67	-6.94	0.892	1.099	23.21	8.859	8.471	-4.39	0.022	0.027	20.20	0.123	0.126	2.75
342	42	53	26.19	55	64	16.36	1.178	1.812	53.89	3.194	4.181	30.92	0.028	0.034	21.95	0.058	0.065	12.51
344	28	46	64.29	6	36	500.00	0.704	3.471	393.13	0.045	1.729	3736.25	0.025	0.075	200.17	0.008	0.048	539.37
345	46	0	-100.00	122	122	0.00	0.509	0.000	-100.00	843.485	841.245	-0.27	0.011	0.000	-100.00	6.914	6.895	-0.27
348	32	61	90.63	0	21	undef.	1.549	5.731	269.94	0.000	0.278	undef.	0.048	0.094	94.07	0.000	0.013	undef.
349	44	37	-15.91	29	7	-75.86	4.188	1.841	-56.05	1.460	0.069	-95.25	0.095	0.050	-47.73	0.050	0.010	-80.32
351	0	11	undef.	57	71	24.56	0.000	0.138	undef.	4.847	9.033	86.37	0.000	0.013	undef.	0.085	0.127	49.62
352	32	52	62.50	88	114	29.55	2.933	2.212	-24.59	27.831	36.251	30.25	0.092	0.043	-53.59	0.316	0.318	0.55
353	4	70	1650.00	10	72	620.00	0.057	8.909	15427	0.269	10.468	3792.65	0.014	0.127	787.28	0.027	0.145	440.65
354	0	27	undef.	0	0	0.00	0.000	0.793	undef.	0.000	0.000	0.00	0.000	0.029	undef.	0.000	0.000	0.00
355	48	42	-12.50	32	27	-15.63	3.781	2.855	-24.48	1.382	0.782	-43.38	0.079	0.068	-13.69	0.043	0.029	-32.90
357	54	8	-85.19	121	103	-14.88	0.710	0.011	-98.48	29.602	23.579	-20.35	0.013	0.001	-89.72	0.245	0.229	-6.43
358	31	0	-100.00	0	0	0.00	1.109	0.000	-100.00	0.000	0.000	0.00	0.036	0.000	-100.00	0.000	0.000	0.00
359	32	84	162.50	122	121	-0.82	0.285	0.193	-32.47	43.638	42.230	-3.23	0.009	0.002	-74.28	0.358	0.349	-2.43
360	31	0	-100.00	3	0	-100.00	0.627	0.000	-100.00	0.014	0.000	-100.00	0.020	0.000	-100.00	0.005	0.000	-100.00
361	2	39	1850.00	112	110	-1.79	0.042	1.793	4171.50	25.429	32.117	26.30	0.021	0.046	119.05	0.227	0.292	28.60
362	0	40	undef.	25	53	112.00	0.000	2.440	undef.	0.508	4.794	843.29	0.000	0.061	undef.	0.020	0.092	353.51
364	35	44	25.71	44	59	34.09	1.519	1.325	-12.78	2.779	2.966	6.73	0.043	0.030	-30.62	0.063	0.050	-20.40
367	28	44	57.14	7	13	85.71	1.053	1.689	60.40	0.061	0.080	31.39	0.038	0.038	2.07	0.009	0.006	-29.25
368	96	0	-100.00	120	122	1.67	13.918	0.000	-100.00	74.606	33.134	-55.59	0.145	0.000	-100.00	0.622	0.272	-56.32
369	48	0	-100.00	89	115	29.21	0.238	0.000	-100.00	24.753	29.831	20.51	0.005	0.000	-100.00	0.278	0.259	-6.73
370	51	0	-100.00	5	0	-100.00	2.220	0.000	-100.00	0.030	0.000	-100.00	0.044	0.000	-100.00	0.006	0.000	-100.00
371	63	42	-33.33	66	47	-28.79	5.471	2.691	-50.81	7.084	3.796	-46.42	0.087	0.064	-26.22	0.107	0.081	-24.76
372	9	35	288.89	7	26	271.43	0.525	1.972	275.68	0.243	0.856	252.60	0.058	0.056	-3.40	0.035	0.033	-5.07
374	50	52	4.00	31	38	22.58	5.617	5.054	-10.02	2.585	1.808	-30.06	0.112	0.097	-13.48	0.083	0.048	-42.95
376	71	45	-36.62	32	0	-100.00	6.232	0.554	-91.11	1.000	0.000	-100.00	0.088	0.012	-85.97	0.031	0.000	-100.00
377	15	61	306.67	48	72	50.00	0.399	10.194	2453.90	4.011	17.974	348.09	0.027	0.167	528.01	0.084	0.250	198.73
378	63	71	12.70	0	25	undef.	5.069	10.924	115.52	0.000	0.668	undef.	0.080	0.154	91.24	0.000	0.027	undef.
2450	41	34	-17.07	0	0	0.00	3.870	0.982	-74.61	0.000	0.000	0.00	0.094	0.029	-69.38	0.000	0.000	0.00
2471	40	31	-22.50	0	0	0.00	2.460	1.249	-49.24	0.000	0.000	0.00	0.062	0.040	-34.51	0.000	0.000	0.00

Tab. 19: Auswertung des berechneten Einsetzens der Trockenheit sowie des absoluten Minimums für die Auswertungsperiode **2003**.

EZ NR	Onset						Minimum		
	Q347			Rwmin			Obs [mm/d]	Sim [mm/d]	+/- [mm/d]
	Obs	Sim	+/- [d]	Obs	Sim	+/- [d]			
331	20.06.2003	24.06.2003	4	22.06.2003	29.06.2003	7	0.032	0.001	-0.031
332	01.06.2003	20.07.2003	49	18.06.2003	16.09.2003	90	0.144	0.178	0.034
334	11.06.2003	22.06.2003	11	15.07.2003	20.07.2003	5	0.162	0.105	-0.057
336	--	--	--	02.06.2003	13.07.2003	41	0.213	0.249	0.036
337	21.06.2003	19.08.2003	59	--	--	--	0.706	0.806	0.1
340	07.06.2003	--	--	09.08.2003	--	--	0.354	0.6	0.246
341	07.06.2003	27.06.2003	20	01.06.2003	09.06.2003	8	0.008	0.004	-0.004
342	05.06.2003	18.06.2003	13	01.06.2003	12.06.2003	11	0.014	0.002	-0.012
344	15.06.2003	26.06.2003	11	26.06.2003	06.07.2003	10	0.112	0.04	-0.072
345	23.06.2003	--	---	01.06.2003	01.06.2003	0	0.05	0.121	0.071
348	01.06.2003	20.06.2003	19	12.07.2003	14.07.2003	2	0.107	0.062	-0.045
349	15.06.2003	15.07.2003	30	25.06.2003	07.08.2003	43	0.177	0.237	0.06
351	21.06.2003	09.07.2003	18	10.06.2003	17.06.2003	7	0.122	0.067	-0.055
352	24.06.2003	28.06.2003	4	02.06.2003	01.06.2003	-1	0.015	0.105	0.09
353	01.06.2003	15.06.2003	14	01.06.2003	12.06.2003	11	0.044	0.046	0.002
354	30.07.2003	14.07.2003	-16	--	08.08.2003	--	0.284	0.173	-0.111
355	20.06.2003	26.06.2003	6	13.07.2003	09.07.2003	-4	0.164	0.096	-0.068
357	04.06.2003	01.08.2003	58	01.06.2003	01.06.2003	0	0.013	0.033	0.02
358	18.06.2003	--	--	16.07.2003	--	--	0.215	0.459	0.244
359	--	--	--	01.06.2003	01.06.2003	0	0.059	0.12	0.061
360	15.07.2003	20.08.2003	36	08.08.2003	12.09.2003	35	0.126	0.149	0.023
361	05.06.2003	28.06.2003	23	01.06.2003	01.06.2003	0	0.01	0.016	0.006
362	06.06.2003	04.07.2003	28	01.06.2003	23.06.2003	22	0.123	0.117	-0.006
364	19.06.2003	01.07.2003	12	19.06.2003	27.06.2003	8	0.074	0.048	-0.026
367	01.06.2003	06.07.2003	35	01.06.2003	17.07.2003	46	0.11	0.11	0
368	09.06.2003	--	--	01.06.2003	01.06.2003	0	0.12	0.546	0.426
369	18.06.2003	--	--	01.06.2003	01.06.2003	0	0	0.024	0.024
370	24.06.2003	--	--	23.08.2003	--	--	0.362	0.754	0.392
371	14.06.2003	01.07.2003	17	12.06.2003	27.06.2003	15	0.025	0.051	0.026
372	10.07.2003	10.07.2003	0	12.07.2003	16.07.2003	4	0.118	0.033	-0.085
374	27.06.2003	30.06.2003	3	20.07.2003	10.07.2003	-10	0.238	0.1	-0.138
376	12.06.2003	--	--	17.06.2003	--	--	0.456	0.701	0.245
377	18.06.2003	20.06.2003	2	01.06.2003	10.06.2003	9	0.321	0.145	-0.176
378	12.06.2003	11.06.2003	-1	--	12.07.2003	--	0.513	0.298	-0.215
2450	22.06.2003	15.07.2003	23	--	--	--	0.188	0.203	0.015
2471	22.06.2003	03.09.2003	73	--	--	--	0.493	0.566	0.073

Tab. 20: Auswertung des berechneten Einsetzens der Trockenheit sowie des absoluten Minimums für die Auswertungsperiode **2011**.

EZ NR	Onset						Minimum		
	Q347			Rwmin			Obs [mm/d]	Sim [mm/d]	+ / - [mm/d]
	Obs	Sim	+ / - [d]	Obs	Sim	+ / - [d]			
331	07.05.2011	11.05.2011	4	10.05.2011	19.05.2011	9	0.002	0.019	0.017
332	24.05.2011	14.06.2011	21	--	--	--	0.307	0.343	0.036
334	25.04.2011	29.04.2011	4	--	--	--	0.265	0.24	-0.025
336	11.03.2011	--	--	01.03.2011	09.05.2011	69	0.033	0.303	0.27
337	18.03.2011	07.05.2011	50	--	--	--	0.719	0.788	0.069
340	17.05.2011	--	--	--	--	--	0.418	0.63	0.212
341	11.05.2011	21.05.2011	10	19.04.2011	25.04.2011	6	0.025	0.031	0.006
342	10.05.2011	09.05.2011	-1	01.05.2011	27.04.2011	-4	0.03	0.02	-0.01
344	18.05.2011	16.05.2011	-2	25.05.2011	26.05.2011	1	0.178	0.108	-0.07
345	08.05.2011	--	--	01.03.2011	01.03.2011	0	0.045	0.115	0.07
348	07.05.2011	01.05.2011	-6	--	08.06.2011	--	0.198	0.149	-0.049
349	07.05.2011	25.05.2011	18	09.05.2011	24.06.2011	46	0.25	0.323	0.073
351	--	10.05.2011	--	15.04.2011	19.04.2011	4	0.271	0.239	-0.032
352	11.05.2011	01.05.2011	-10	07.03.2011	06.03.2011	-1	0.147	0.244	0.097
353	05.06.2011	22.04.2011	-44	05.06.2011	20.04.2011	-46	0.322	0.147	-0.175
354	--	04.06.2011	--	--	--	--	0.446	0.31	-0.136
355	02.05.2011	20.05.2011	18	20.05.2011	04.06.2011	15	0.182	0.22	0.038
357	19.04.2011	23.06.2011	65	02.03.2011	05.03.2011	3	0.019	0.049	0.03
358	10.05.2011	--	--	--	--	--	0.312	0.459	0.147
359	07.04.2011	01.03.2011	-37	01.03.2011	01.03.2011	0	0.042	0.054	0.012
360	20.05.2011	--	--	30.05.2011	--	--	0.167	0.209	0.042
361	15.06.2011	23.05.2011	-23	09.03.2011	04.03.2011	-5	0.104	0.065	-0.039
362	--	22.05.2011	--	24.05.2011	10.05.2011	-14	0.318	0.21	-0.108
364	26.04.2011	05.05.2011	9	23.04.2011	29.04.2011	6	0.132	0.161	0.029
367	07.05.2011	09.05.2011	2	26.05.2011	09.06.2011	14	0.254	0.256	0.002
368	20.03.2011	--	--	02.03.2011	01.03.2011	-1	0.077	0.556	0.479
369	26.04.2011	--	--	15.03.2011	04.03.2011	-11	0	0.023	0.023
370	10.04.2011	--	--	24.05.2011	--	--	0.38	0.623	0.243
371	19.04.2011	20.05.2011	31	18.04.2011	18.04.2011	0	0.02	0.13	0.11
372	19.06.2011	27.05.2011	-23	20.06.2011	05.06.2011	-15	0.135	0.127	-0.008
374	05.05.2011	10.05.2011	5	17.05.2011	22.05.2011	5	0.141	0.233	0.092
376	19.04.2011	01.05.2011	12	19.05.2011	--	--	0.387	0.551	0.164
377	24.05.2011	01.05.2011	-23	26.04.2011	20.04.2011	-6	0.411	0.27	-0.141
378	17.04.2011	21.04.2011	4	--	30.05.2011	--	0.494	0.435	-0.059
2450	19.04.2011	24.05.2011	35	--	--	--	0.293	0.43	0.137
2471	24.03.2011	31.05.2011	68	--	--	--	0.483	0.6	0.117

Tab. 21: Auswertung der Trockenheits-Dauer für das **Delay-Modell**, je für die Auswertungsperioden 2003 und 2011 und für die Schwellenwerte Q347 und Mindestrestwassermenge (RWmin). Obs = beobachteter Wert, Std Sim = Modellierung mit dem Standard-Modell (Vergleichswert), Delay sim = Modellierung mit dem Delay-Modell. Zusätzlich sind jeweils die relative Abweichung zwischen der Delay-Simulation und dem beobachteten Wert sowie eine qualitative Bewertung im Vergleich zur Standardsimulation angegeben.

Die Gebiete mit Karst-Anteil sind in blauer Schrift hervorgehoben.

EZ NR	2003										2011									
	Q347					RWmin					Q347					RWmin				
	Obs [d]	Std Sim [d]	Delay sim [d]	+/- [%]	Wertung	Obs [d]	Std Sim [d]	Delay sim [d]	+/- [%]	Wertung	Obs [d]	Std Sim [d]	Delay sim [d]	+/- [%]	Wertung	Obs [d]	Std Sim [d]	Delay sim [d]	+/- [%]	Wertung
336	0	0	0	--	=	83	80	55	-33.73	-	96	0	0	-100	=	115	53	81	-29.57	+
341	100	96	100	0	+	120	114	122	1.67	+	40	41	39	-2.5	=	72	67	73	1.39	+
345	41	0	64	56.10	++	122	122	122	0	=	46	0	53	15.22	++	122	122	122	0	=
348	98	103	114	16.33	-	39	79	85	117.95	-	32	61	74	131.25	-	0	21	27	undef.	-
349	87	78	0	-100	--	66	55	0	-100	--	44	37	67	52.27	-	29	7	6	-79.31	-
353	102	108	116	13.73	-	108	111	119	10.19	-	4	70	75	1775	-	10	72	77	670	-
357	91	61	108	18.68	+	122	122	122	0	=	54	8	68	25.93	++	121	103	104	-14.05	+
359	0	0	0	--	=	122	122	122	0	=	32	84	94	193.75	-	122	121	120	-1.64	-
360	50	42	0	-100	--	33	19	0	-100	--	31	0	7	-77.42	+	3	0	0	-100	=
362	84	89	0	-100	--	107	100	0	-100	--	0	40	0	0	++	25	53	54	116	-
368	80	0	0.00	-100	=	122	122	122	0	=	96	0	0	-100	=	120	122	122	1.67	=
369	87	0	113	29.89	++	122	122	122	0	=	48	0	44	-8.33	++	89	115	114	28.09	+
370	19	0	0	-100	=	3	0	0	-100	=	51	0	0	-100	=	5	0	0	-100	=
376	59	0	0	-100	=	3	0	0	-100	=	71	45	0	-100	--	32	0	0	-100	=

Tab. 22: Auswertung des Trockenheits-Defizits für das Delay-Modell. Für Erläuterungen siehe Tab. 21.

EZ NR	2003										2011									
	Q347					RWmin					Q347					RWmin				
	Obs [mm]	Std Sim [mm]	Delay sim [mm]	+/- [%]	Wertung	Obs [mm]	Std Sim [mm]	Delay sim [mm]	+/- [%]	Wertung	Obs [mm]	Std Sim [mm]	Delay sim [mm]	+/- [%]	Wertung	Obs [mm]	Std im [mm]	Delay sim [mm]	+/- [%]	Wertung
336	0	0	0	--	=	5.817	7.70	3.15	-45.87	-	10.787	0	0	-100	=	31.689	3.48	6.24	-80.30	+
341	4.332	5.431	4.87	12.35	+	19.328	19.41	19.87	2.82	-	0.892	1.10	0.79	-11.21	+	8.859	8.47	8.57	-3.31	+
345	0.351	0	1.05	200.45	+	851.255	846.35	854.80	0.42	+	0.509	0	0.87	70.44	++	843.485	841.25	842.13	-0.16	+
348	10.826	17.690	18.98	75.29	-	1.048	5.97	5.83	456.05	+	1.549	5.73	7.53	385.99	-	0	0.28	0.64	undef.	-
349	10.323	8.112	0	-100	--	4.551	3.22	0	-100	--	4.188	1.84	2.97	-29.16	+	1.460	0.07	0.02	-98.70	-
353	15.479	22.995	20.79	34.31	+	17.807	25.40	23.38	31.30	+	0.057	8.91	9.00	15'593.30	-	0.269	10.47	10.68	3'869.98	-
357	1.493	0.602	2.25	50.90	+	33.488	32.68	35.01	4.55	-	0.710	0.01	1.29	81.68	+	29.602	23.58	25.81	-12.80	+
359	0	0	0	--	=	40.429	35.44	34.50	-14.67	-	0.285	0.19	2.93	928.58	--	43.638	42.23	43.34	-0.68	+
360	2.370	1.305	0	-100	--	0.922	0.25	0	-100	--	0.627	0	0.02	-97.01	+	0.014	0	0	-100	=
362	5.285	10.405	0	-100	--	10.258	15.21	0	-100	--	0	2.44	0	0	++	0.508	4.79	1.55	204.92	+
368	5.445	0	0	-100	=	68.987	34.29	45.80	-33.62	+	13.918	0	0	-100	=	74.606	33.13	39.26	-47.38	+
369	0.496	0	0.97	95.14	+	38.039	36.75	39.16	2.95	+	0.238	0	0.24	-1.34	++	24.753	29.83	30.49	23.17	-
370	0.770	0	0	-100	=	0.081	0	0	-100	=	2.220	0	0	-100	=	0.030	0	0	-100	=
376	2.223	0	0	-100	=	0.037	0	0	-100	=	6.232	0.55	0	-100	--	1.000	0	0	-100	=

Tab. 23: Auswertung der Trockenheits-Intensität für das **Delay-Modell**. Für Erläuterungen siehe Tab. 21.

EZ NR	2003										2011									
	Q347					RWmin					Q347					RWmin				
	Obs [mm/ d]	Std Sim [mm/d]	Delay sim [mm/d]	+ / - [%]	Wer- tung	Obs [mm/d]	Std Sim [mm/d]	Delay sim [mm/d]	+ / - [%]	Wer- tung	Obs [mm/d]	Std Sim [mm/d]	Delay sim [mm/d]	+ / - [%]	Wer- tung	Obs [mm/d]	Std Sim [mm/d]	Delay sim [mm/d]	+ / - [%]	Wer- tung
336	0	0	0	--	=	0.070	0.096	0.06	-18.32	+	0.112	0	0	-100	=	0.276	0.066	0.077	-72.03	+
341	0.043	0.057	0.05	12.35	+	0.161	0.170	0.16	1.14	+	0.022	0.027	0.020	-8.94	+	0.123	0.126	0.117	-4.63	-
345	0.009	0	0.02	92.47	+	6.978	6.937	7.01	0.42	+	0.011	0	0.016	47.93	++	6.914	6.895	6.903	-0.16	+
348	0.110	0.172	0.17	50.69	+	0.027	0.076	0.07	155.13	+	0.048	0.094	0.102	110.16	-	0	0.013	0.024	undef.	-
349	0.119	0.104	0	-100	--	0.069	0.059	0	-100	--	0.095	0.050	0.044	-53.48	-	0.050	0.010	0.003	-93.70	-
353	0.152	0.213	0.18	18.10	+	0.165	0.229	0.20	19.16	+	0.014	0.127	0.120	736.98	+	0.027	0.145	0.139	415.58	+
357	0.016	0.010	0.02	27.15	+	0.274	0.268	0.29	4.55	-	0.013	0.001	0.019	44.27	+	0.245	0.229	0.248	1.46	+
359	0	0	0	--	=	0.331	0.290	0.28	-14.67	-	0.009	0.002	0.031	250.16	--	0.358	0.349	0.361	0.97	+
360	0.047	0.031	0	-100	--	0.028	0.013	0	-100	--	0.020	0	0.003	-86.74	+	0.005	0	0	-100	=
362	0.063	0.117	0	-100	--	0.096	0.152	0	-100	--	0	0.061	0	--	++	0.020	0.092	0.029	41.17	++
368	0.068	0	0	-100	=	0.565	0.281	0.38	-33.62	+	0.145	0	0	-100	=	0.622	0.272	0.322	-48.24	+
369	0.006	0	0.01	50.24	++	0.312	0.301	0.32	2.95	+	0.005	0	0.005	7.62	++	0.278	0.259	0.267	-3.84	+
370	0.041	0	0	-100	--	0.027	0	0	-100	=	0.044	0	0	-100	=	0.006	0	0	-100	=
376	0.038	0	0	-100	--	0.012	0	0	-100	=	0.088	0.012	0	-100	--	0.031	0	0	-100	=

Tab. 24: Auswertung des Trockenheits-Onsets für das **Delay-Modell**. Für Erläuterungen siehe Tab. 21. Die Abweichung zum beobachteten Wert ist hier absolut angegeben.

EZ NR	2003										2011									
	Q347					RWmin					Q347					RWmin				
	Obs	Std Sim	Delay sim	+/ - [d]	Wer- tung	Obs	Std Sim	Delay sim	+/ - [d]	Wer- tung	Obs	Std Sim	Delay sim	+/ - [d]	Wer- tung	Obs	Std Sim	Delay sim	+/ - [d]	Wer- tung
336	--	--	--	--	=	02.06.	13.07.	03.08.	62	-	11.03.	--	--	--	=	01.03.	09.05.	11.04.	41	+
341	07.06.	27.06.	23.06.	-16	+	01.06.	09.06.	01.06.	0	+	11.05.	21.05.	23.05.	12	-	19.04.	25.04.	19.04.	0	+
345	23.06.	--	29.07.	-36	++	01.06.	01.06.	01.06.	0	=	08.05.	--	09.05.	1	++	01.03.	01.03.	01.03.	0	=
348	01.06.	20.06.	09.06.	-8	+	12.07.	14.07.	08.07.	-4	-	07.05.	01.05.	18.04.	-19	-	--	08.06.	25.05.	--	=
349	15.06.	15.07.	--	--	--	25.06.	07.08.	--	--	--	07.05.	25.05.	25.04.	-12	+	09.05.	24.06.	16.06.	38	+
353	01.06.	15.06.	07.06.	-6	+	01.06.	12.06.	04.06.	3	+	05.06.	22.04.	17.04.	-49	-	05.06.	20.04.	15.04.	-51	-
357	04.06.	01.08.	15.06.	-11	+	01.06.	01.06.	01.06.	0	=	19.04.	23.06.	24.04.	5	++	02.03.	05.03.	05.03.	3	=
359	--	--	--	--	=	01.06.	01.06.	01.06.	0	=	07.04.	01.03.	04.03.	-34	+	01.03.	01.03.	01.03.	0	=
360	15.07.	20.08.	--	--	--	08.08.	12.09.	--	--	--	20.05.	--	24.06.	35	+	30.05.	--	--	--	=
362	06.06.	04.07.	--	--	--	01.06.	23.06.	--	--	--	--	22.05.	--	--	++	24.05.	10.05.	08.05.	-16	-
368	09.06.	--	--	--	=	01.06.	01.06.	01.06.	0	=	20.03.	--	--	--	=	02.03.	01.03.	01.03.	-1	=
369	18.06.	--	10.06.	8	++	01.06.	01.06.	01.06.	0	=	26.04.	--	18.05.	22	+	15.03.	04.03.	02.03.	-13	-
370	24.06.	--	--	--	=	23.08.	--	--	--	=	10.04.	--	--	--	=	24.05.	--	--	--	=
376	12.06.	--	--	--	=	17.06.	--	--	--	=	19.04.	01.05.	--	--	--	19.05.	--	--	--	=

Tab. 25: Auswertung des Abfluss-**Minimums** für das **Delay-Modell**. Für Erläuterungen siehe Tab. 21. Die Abweichung zum beobachteten Wert ist hier absolut angegeben.

EZ NR	2003					2011				
	Obs [mm/d]	Std Sim [mm/d]	Delay sim [mm/d]	+ / - [mm/d]	Wertung	Obs [mm/d]	Std Sim [mm/d]	Delay sim [mm/d]	+ / - [mm/d]	Wertung
336	0.213	0.249	0.298	0.085	-	0.033	0.303	0.263	0.23	+
341	0.008	0.004	0.008	0	+	0.025	0.031	0.039	0.014	-
345	0.05	0.121	0.041	-0.009	++	0.045	0.115	0.043	-0.002	++
348	0.107	0.062	0.064	-0.043	+	0.198	0.149	0.126	-0.072	-
349	0.177	0.237	0.415	0.238	-	0.25	0.323	0.334	0.084	-
353	0.044	0.046	0.079	0.035	-	0.322	0.147	0.154	-0.168	+
357	0.013	0.033	0.017	0.004	+	0.019	0.049	0.022	0.003	+
359	0.059	0.12	0.12	0.061	=	0.042	0.054	0.023	-0.019	-
360	0.126	0.149	0.284	0.158	-	0.167	0.209	0.203	0.036	+
362	0.123	0.117	0.472	0.349	-	0.318	0.21	0.313	-0.005	+
368	0.12	0.546	0.45	0.33	+	0.077	0.556	0.473	0.396	+
369	0	0.024	0.001	0.001	++	0	0.023	0.004	0.004	++
370	0.362	0.754	0.734	0.372	+	0.38	0.623	0.581	0.201	+
376	0.456	0.701	0.849	0.393	-	0.387	0.551	0.629	0.242	-

Tab. 26: Auswertung der Trockenheits-Dauer für das **HBV-GW-Modell**, je für die Auswertungsperioden 2003 und 2011 und für die Schwellenwerte Q347 und Mindestrestwassermenge (RWmin). Obs = beobachteter Wert, Sim = Modellierung mit dem Standard-Modell (Vergleichswert), GW sim = Modellierung mit Co-Kalibrierung auf Grundwasserstände. Zusätzlich sind jeweils die relative Abweichung zwischen der GW-Simulation und dem beobachteten Wert sowie eine qualitative Wertung der GW-Simulation im Vergleich zur Standardsimulation angegeben.

An die Tabelle angefügt sind ausserdem die Werte der vier Einzugsgebiete, die zusätzlich mit dem Delay-Modell kalibriert wurden. In diesem Fall steht Sim für die Simulation des Delay-Modells und GW sim für die Ergebnisse der Co-Kalibrierung des Delay-Modells mit Grundwasserdaten. Weiter sind die relative Abweichung des simulierten Werts vom beobachteten Wert sowie eine qualitative Wertung im Vergleich zur einfachen Delay-Simulation angegeben.

EZ NR	2003										2011									
	Q347					Rwmin					Q347					Rwmin				
	Obs [d]	Sim [d]	GW sim [d]	+/- [%]	Wertung	Obs [d]	Sim [d]	GW sim [d]	+/- [%]	Wertung	Obs [d]	Sim [d]	GW sim [d]	+/- [%]	Wertung	Obs [d]	Sim [d]	GW sim [d]	+/- [%]	Wertung
331	76	103	100	31.58	+	56	101	97	73.21	+	44	50	45	2.27	+	32	43	42	31.25	+
332	104	73	30	-71.15	-	62	15	0	-100	--	17	9	0	-100	--	0	0	0	0	=
334	96	101	42	-56.25	-	44	73	0	-100	--	45	61	0	-100	--	0	0	0	0	=
341	100	96	77	-23.00	-	120	114	112	-6.67	-	40	41	21	-47.50	-	72	67	56	-22.22	-
342	92	105	106	15.22	-	109	111	112	2.75	-	42	53	53	26.19	=	55	64	62	12.73	+
348	98	103	0	-100	--	39	79	0	-100	--	32	61	0	-100	--	0	21	0	0	++
353	102	108	38	-62.75	-	108	111	51	-52.78	-	4	70	0	-100	--	10	72	0	-100	--
354	33	79	58	75.76	+	0	54	13	0	+	0	27	0	0	++	0	0	0	0	=
360	50	42	97	94.00	-	33	19	88	166.67	-	31	0	42	35.48	++	3	0	34	1'033.33	+
372	56	83	52	-7.14	+	30	77	39	30.00	+	9	35	0	-100	--	7	26	0	-100	--
374	72	93	91	26.39	+	40	83	82	105.00	+	50	52	50	0	+	31	38	39	25.81	-
2450	81	78	27	-66.67	-	0	0	0	0	=	41	34	0	-100	--	0	0	0	0	=
2471	71	28	70	-1.41	+	0	0	0	0	=	40	31	58	45.00	-	0	0	0	0	=
Simulationen mit Delay-Modell																				
341	100	100	98	-2.00	-	120	122	122	1.67	=	40	39	37	-7.50	-	72	73	71	-1.39	=
348	98	114	96	-2.04	+	39	85	1	-97.44	-	32	74	67	109.38	+	0	27	0	0	++
353	102	116	100	-1.96	+	108	119	104	-3.70	+	4	75	70	1'650	+	10	77	73	630	+
360	50	0	113	126.00	++	33	0	108	227.27	++	31	7	60	93.55	+	3	0	52	1633.33	+

Tab. 27: Auswertung des Trockenheits-Defizits für das HBV-GW-Modell. Für Erläuterungen siehe Tab. 26

EZ NR	2003										2011									
	Q347					Rwmin					Q347					Rwmin				
	Obs [mm]	Sim [mm]	GW sim [mm]	+ / - [%]	Wer tung	Obs [mm]	Sim [mm]	GW sim [mm]	+ / - [%]	Wer tung	Obs [mm]	Sim [mm]	GW sim [mm]	+ / - [%]	Wer tung	Obs [mm]	Sim [mm]	GW sim [mm]	+ / - [%]	Wer tung
331	1.41	5.81	5.54	292.00	+	0.62	4.59	4.36	602.47	+	1.11	1.53	1.60	43.92	-	0.65	0.98	1.08	65.69	-
332	15.14	8.11	0.96	-93.66	--	1.74	0.22	0	-100	--	0.39	0.07	0	-100	--	0	0	0	0	=
334	14.42	20.80	1.85	-87.20	--	1.60	6.07	0	-100	--	2.98	5.56	0	-100	--	0	0	0	0	=
341	4.33	5.43	3.38	-22.03	+	19.33	19.41	16.20	-16.20	-	0.89	1.10	0.33	-62.88	-	8.86	8.471	5.81	-34.41	-
342	3.77	6.39	6.41	69.88	-	7.99	10.91	10.97	37.18	-	1.18	1.81	1.98	68.24	-	3.19	4.18	4.35	36.35	-
348	10.83	17.69	0	-100	--	1.05	5.97	0	-100	--	1.55	5.73	0	-100	--	0	0.28	0	0.00	++
353	15.48	23.00	1.06	-93.14	--	17.81	25.40	2.05	-88.51	--	0.06	8.91	0	-100	--	0.27	10.47	0	-100	--
354	0.93	9.96	4.64	401.29	+	0.00	2.92	0.19	0.00	+	0	0.79	0	0	++	0	0	0	0	=
360	2.37	1.30	11.17	371.23	--	0.92	0.25	8.00	767.29	--	0.63	0	2.74	337.91	+	0.01	0	1.44	9'980.29	+
372	2.42	10.43	3.20	32.13	+	0.75	7.54	1.56	108.19	+	0.52	1.97	0	-100	--	0.24	0.86	0	-100	--
374	5.94	19.32	19.66	230.78	-	1.76	12.94	13.33	658.49	-	5.62	5.05	5.67	0.91	+	2.59	1.81	2.47	-4.36	+
2450	13.37	13.67	1.22	-90.91	-	0	0	0	0	=	3.87	0.98	0	-100	--	0	0	0	0	=
2471	5.85	1.48	11.94	104.08	-	0	0	0	0	=	2.46	1.25	8.10	229.30	-	0	0	0	0	=
Simulationen mit Delay-Modell																				
341	4.33	4.87	4.79	10.58	+	19.33	19.87	19.77	2.29	+	0.89	0.79	0.75	-15.95	-	8.86	8.57	8.13	-8.23	-
348	10.83	18.98	7.17	-33.77	+	1.05	5.83	0.0009	-99.91	+	1.55	7.53	4.1	164.66	+	0	0.64	0	0	++
353	15.48	20.79	8.65	-44.12	-	17.81	23.38	10.9	-38.79	-	0.06	9.00	4.12	7'080.60	+	0.27	10.68	5.68	2'012.18	+
360	2.37	0	14.96	531.18	+	0.92	0	11.19	1'113.82	+	0.63	0.02	4.80	666.13	-	0.01	0	2.92	20'374.23	+

Tab. 28: Auswertung der Trockenheits-Intensität für das HBV-GW-Modell. Für Erläuterungen siehe Tab. 26.

EZ NR	2003										2011									
	Q347					Rwmin					Q347					Rwmin				
	Obs [mm/d]	Sim [mm/d]	GW sim [mm/d]	+ / - [%]	Wertung	Obs [mm/d]	Sim [mm/d]	GW sim [mm/d]	+ / - [%]	Wertung	Obs [mm/d]	Sim [mm/d]	GW sim [mm/d]	+ / - [%]	Wertung	Obs [mm/d]	Sim [mm/d]	GW sim [mm/d]	+ / - [%]	Wertung
331	0.019	0.056	0.055	197.92	+	0.011	0.045	0.045	305.55	=	0.025	0.031	0.036	40.72	-	0.020	0.023	0.026	26.24	-
332	0.146	0.111	0.032	-78.01	-	0.028	0.015	0	-100	--	0.023	0.007	0	-100	--	0	0	0	0	=
334	0.150	0.206	0.044	-70.73	-	0.036	0.083	0	-100	--	0.066	0.091	0	-100	--	0	0	0	0	=
341	0.043	0.057	0.044	1.26	+	0.161	0.170	0.145	-10.22	-	0.022	0.027	0.016	-29.30	-	0.123	0.126	0.104	-15.67	-
342	0.041	0.061	0.060	47.45	+	0.073	0.098	0.098	33.50	=	0.028	0.034	0.037	33.32	-	0.058	0.065	0.070	20.96	-
348	0.110	0.172	0	-100	--	0.027	0.076	0	-100	--	0.048	0.094	0	-100	--	0	0.013	0	0	++
353	0.152	0.213	0.028	-81.58	--	0.165	0.229	0.040	-75.66	--	0.014	0.127	0	-100	--	0.027	0.145	0	-100	--
354	0.028	0.126	0.080	185.22	+	0	0.054	0.015	0	+	0	0.029	0	0	++	0	0	0	0	=
360	0.047	0.031	0.115	142.90	-	0.028	0.013	0.091	225.23	-	0.020	0	0.065	223.22	++	0.005	0	0.042	789.44	+
372	0.043	0.126	0.062	42.29	+	0.025	0.098	0.040	60.14	+	0.058	0.056	0	-100	--	0.035	0.033	0	-100	--
374	0.083	0.208	0.216	161.72	-	0.044	0.156	0.163	269.99	-	0.112	0.097	0.113	0.91	+	0.083	0.048	0.063	-23.98	+
2450	0.165	0.175	0.045	-72.72	-	0	0	0	0	=	0.094	0.029	0	-100	--	0	0	0	0	=
2471	0.082	0.053	0.171	106.99	-	0	0	0	0	=	0.062	0.040	0.140	127.10	-	0	0	0	0	=
Simulationen mit Delay-Modell																				
341	0.043	0.049	0.049	13.12	+	0.161	0.163	0.162	0.58	+	0.022	0.020	0.020	-10.34	=	0.123	0.117	0.115	-6.54	-
348	0.110	0.166	0.075	-32.11	-	0.027	0.069	0.0009	-96.65	+	0.048	0.102	0.061	26	+	0	0.024	0	0	++
353	0.152	0.179	0.087	-42.67	-	0.165	0.196	0.105	-36.32	-	0.014	0.120	0.059	311.32	+	0.027	0.139	0.078	190.05	+
360	0.047	0	0.132	178.46	++	0.028	0	0.104	272.28	++	0.020	0.003	0.080	295.84	-	0.005	0	0.056	1'077.97	+

Tab. 29: Auswertung des Trockenheits-Onsets für das **HBV-GW-Modell**. Für Erläuterungen siehe Tab. 26. Die Abweichung zum beobachteten Wert ist hier in Tagen angegeben.

EZ NR	2003										2011									
	Q347					Rwmin					Q347					Rwmin				
	Obs	Sim	GW sim	+ / - [d]	Wertung	Obs	Sim	GW sim	+ / - [d]	Wertung	Obs	Sim	GW sim	+ / - [d]	Wertung	Obs	Sim	GW sim	+ / - [d]	Wertung
331	20.06.	24.06.	23.06.	3	+	22.06.	29.06.	26.06.	4	+	07.05.	11.05.	17.05.	10	-	10.05.	19.05.	20.05.	10	-
332	01.06.	20.07.	01.09.	92	--	18.06.	16.09.	--	--	--	24.05.	14.06.	--	--	--	--	--	--	--	=
334	11.06.	22.06.	20.08.	70	--	15.07.	20.07.	--	--	--	25.04.	29.04.	--	--	--	--	--	--	--	=
341	07.06.	27.06.	16.07.	39	-	01.06.	09.06.	11.06.	10	-	11.05.	21.05.	10.06.	30	-	19.04.	25.04.	06.05.	17	-
342	05.06.	18.06.	17.06.	12	+	01.06.	12.06.	11.06.	10	+	10.05.	09.05.	09.05.	-1	=	01.05.	27.04.	30.04.	-1	+
348	01.06.	20.06.	--	--	--	12.07.	14.07.	--	--	--	07.05.	01.05.	--	--	--	--	08.06.	--	--	++
353	01.06.	15.06.	24.08.	84	--	01.06.	12.06.	11.08.	71	--	05.06.	22.04.	--	--	--	05.06.	20.04.	--	--	--
354	30.07.	14.07.	04.08.	5	+	--	08.08.	18.09.	--	=	--	04.06.	--	--	++	--	--	--	--	=
360	15.07.	20.08.	26.06.	-19	+	08.08.	12.09.	05.07.	-34	+	20.05.	--	20.05.	0	++	30.05.	--	28.05.	-2	++
372	10.07.	10.07.	10.08.	31	-	12.07.	16.07.	23.08.	42	-	19.06.	27.05.	--	--	--	20.06.	05.06.	--	--	--
374	27.06.	30.06.	02.07.	5	-	20.07.	10.07.	11.07.	-9	+	05.05.	10.05.	12.05.	7	-	17.05.	22.05.	23.05.	6	-
2450	22.06.	15.07.	04.09.	74	--	--	--	--	--	=	19.04.	24.05.	--	--	--	--	--	--	--	=
2471	22.06.	03.09.	23.07.	31	+	--	--	--	--	=	24.03.	31.05.	04.05.	41	+	--	--	--	--	=
Simulationen mit Delay-Modell																				
341	07.06.	23.06.	25.06.	18	-	01.06.	01.06.	01.06.	0	=	11.05.	23.05.	25.05.	14	-	19.04.	19.04.	21.04.	2	-
348	01.06.	09.06.	27.06.	26	-	12.07.	08.07.	30.09.	80	-	07.05.	18.04.	21.04.	-16	+	--	25.05.	--	--	++
353	01.06.	07.06.	23.06.	22	-	01.06.	04.06.	19.06.	18	-	05.06.	17.04.	22.04.	-44	+	05.06.	15.04.	19.04.	-47	+
360	15.07.	--	10.06.	-35	+	08.08.	--	15.06.	-54	-	20.05.	24.06.	30.04.	-20	+	30.05.	--	10.05.	-20	+

Tab. 30: Auswertung des Trockenheits-**Minimums** für das **HBV-GW-Modell**. Für Erläuterungen siehe Tab. 26. Die Abweichung zum beobachteten Wert ist hier absolut angegeben.

EZ NR	2003					2011				
	Obs [mm/d]	Sim [mm/d]	GW sim	+ / - [mm/d]	Wer- tung	Obs [mm/d]	Sim [mm/d]	GW sim	+ / - [mm/d]	Wer- tung
331	0.032	0.001	0.001	-0.031	=	0.002	0.019	0.017	0.015	+
332	0.144	0.178	0.294	0.15	-	0.307	0.343	0.426	0.119	-
334	0.162	0.105	0.313	0.151	-	0.265	0.24	0.406	0.141	-
341	0.008	0.004	0.01	0.002	+	0.025	0.031	0.046	0.021	-
342	0.014	0.002	0.002	-0.012	=	0.03	0.02	0.018	-0.012	-
348	0.107	0.062	0.305	0.198	-	0.198	0.149	0.35	0.152	-
353	0.044	0.046	0.29	0.246	-	0.322	0.147	0.367	0.045	+
354	0.284	0.173	0.233	-0.051	+	0.446	0.31	0.374	-0.072	+
360	0.126	0.149	0.037	-0.089	-	0.167	0.209	0.102	-0.065	-
372	0.118	0.033	0.107	-0.011	+	0.135	0.127	0.218	0.083	-
374	0.238	0.100	0.086	-0.152	-	0.141	0.233	0.22	0.079	+
2450	0.188	0.203	0.407	0.219	-	0.293	0.43	0.527	0.234	-
2471	0.493	0.566	0.359	-0.134	-	0.483	0.6	0.438	-0.045	+
Simulationen mit Delay-Modell										
341	0.008	0.008	0.007	-0.001	-	0.025	0.039	0.039	0.014	=
348	0.107	0.064	0.175	0.068	-	0.198	0.126	0.196	-0.002	+
353	0.044	0.079	0.206	0.162	-	0.322	0.154	0.238	-0.084	+
360	0.126	0.284	0.03	-0.096	-	0.167	0.203	0.085	-0.082	-

Persönliche Erklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig verfasst und die den verwendeten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Ort und Datum

(Isabelle Nanz)
