



**TECHNISCHE
UNIVERSITÄT
DRESDEN**

Institut für Hydrologie und Meteorologie Lehrstuhl für Hydrologie

Ronny Peters

**KÜNSTLICHE NEURONALE NETZE ZUR
BESCHREIBUNG DER
HYDRODYNAMISCHEN PROZESSE
FÜR DEN HOCHWASSERFALL UNTER
BERÜCKSICHTIGUNG DER
NIEDERSCHLAGS-ABFLUSS-PROZESSE
IM ZWISCHENEINZUGSGEBIET**

Dresden, Juli 2008



Dresdner Schriften zur Hydrologie

Heft 7

**Künstliche neuronale Netze zur Beschreibung der
hydrodynamischen Prozesse für den
Hochwasserfall unter Berücksichtigung der
Niederschlags-Abfluß-Prozesse im
Zwischeneinzugsgebiet**

Dissertation

zur Erlangung des akademischen Grades
Doctor rerum naturalium

vorgelegt der Fakultät Forst-, Geo- und Hydrowissenschaften
Der Technischen Universität Dresden

von Diplom-Hydrologe Ronny Peters

Gutachter:

Prof. Dr.-Ing. Habil. Schmitz, Gerd H., TU Dresden
Prof. Dr. rer. nat. Liedl, Rudolf, TU Dresden
Prof. Dr. rer. nat. Dr.-Ing. habil. Bárdossy, András, Universität
Stuttgart

Dresden, März 2007

Zusammenfassung

Aus den Mängeln bisher verwendeter Modelle zur Abbildung des Wellenablaufes zu Prognosezwecken im Hochwasserfall wird in dieser Arbeit eine Methodik entwickelt, die die Schnelligkeit und Robustheit künstlicher neuronaler Netze mit der Zuverlässigkeit hydrodynamisch-numerischer Modellierung verbindet.

Ein eindimensionales hydrodynamisches Modell beinhaltet die genaue Kenntnis der Geometrie des Flußlaufes und der Vorländer und berücksichtigt die physikalischen Prozesse des Wellenablaufes. Mit diesem deterministischen Modell ist eine Grundlage für umfangreiche Szenarienrechnungen zur Erstellung einer Datenbasis geschaffen, die die weite Spanne theoretisch möglicher Hochwasserereignisse abdeckt. Mit dieser Datenbasis können dann künstliche neuronale Netze trainiert werden, die auch im Bereich extremer Hochwasserereignisse zuverlässige Prognosen liefern. In dieser Arbeit werden mit Multilayer-Feedforward-Netzen und selbstorganisierenden Merkmalskarten zwei Netztypen als Vertreter überwacht und unüberwacht lernender neuronaler Netze auf ihre diesbezügliche Eignung untersucht und beurteilt.

Desweiteren wurde die Methodik auf die Einbeziehung von Merkmalen für die Niederschlags-Abfluß-Prozesse im unbeobachteten Zwischengebiet zur Berücksichtigung lateraler Zuflüsse entlang der modellierten Fließstrecken erweitert. Die Datenbasis wurde hierfür mit einem Niederschlags-Abfluß-Modell erstellt.

Ein Hauptschwerpunkt liegt in der Überführung der Eingangsdaten in charakteristische Merkmale zur Abbildung der Zielgrößen, in diesem Falle des Durchflusses und Wasserstandes am Zielpegel. So dienen die deterministischen Modelle nicht nur zur Erstellung einer verlässlichen Datenbasis für das Training der Netze, sondern ermöglichen – sowohl für die Niederschlags-Abfluß-Prozesse, als auch für die hydrodynamischen Prozesse – Analysen betreffs der Sensitivität der Modellergebnisse infolge von Änderungen der Inputdaten. Mit Hilfe dieser Analysen werden wichtige Informationen zur Findung der relevanten Merkmale erlangt.

Ein Schlüssel für die erfolgreiche Eingliederung der Niederschlags-Abfluß-Prozesse in das Prognosenetz ist die Einführung eines einzigen Zustandsmerkmals, welches die gesamte meteorologische Vorgeschichte des Ereignisses zur Charakterisierung des Gebietszustandes vereinigt.

Die entwickelte Methodik wurde anhand des Einzugsgebietes der Freiburger Mulde erfolgreich getestet.

0 Inhalt

0	Inhalt	1
1	Einleitung	3
2	Künstliche neuronale Netze zur Durchflußvorhersage	6
3	Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen	11
3.1	Grundprinzipien der entwickelten Methodik zur Hochwasservorhersage.....	11
3.2	Neuronale Berechnungsmodelle	18
3.2.1	Grundlegende Konzepte von Lernregeln.....	18
3.2.2	Multilayer feedforward networks.....	20
3.2.3	Self organizing feature maps.....	23
3.3	Datengenerierung als Trainingsgrundlage	26
3.3.1	Die eindimensionale hydrodynamisch-numerische Modellierung	26
3.3.2	Die Niederschlags-Abfluß-Modellierung	31
3.4	Szenarien der Einzugsgebietsgliederung	33
3.5	Merkmalsselektion für die hydrodynamischen Prozesse	34
3.6	Merkmalsselektion für die NA-Prozesse im Zwischengebiet	37
3.6.1	Ereignismerkmale	39
3.6.2	Das Zustandsmerkmal	41
3.6.3	Räumliche und zeitliche Struktur der Eingangsdaten und NA-Merkmale	44
4	Versuchsgebiet und Erstellung der Trainingsdaten.....	50
4.1	Allgemeine Angaben zum Untersuchungsgebiet	50
4.2	Untergliederung in Teileinzugsgebiete für die deterministische Modellierung	55
4.3	Einzugsgebietsgliederung der ANN-Szenarien.....	57
4.4	Daten zur Gerinnegeometrie in HEC-RAS.....	59
4.5	Pegeldaten zur Eichung des hydrodynamischen Modells.....	59
4.6	Datenbasis für die Szenarienrechnung	61
4.6.1	Niederschlag und sonstige Wetterdaten.....	61
4.6.2	Zuflüsse zu den hydrodynamisch modellierten Flußabschnitten	63
4.7	Ergebnisse der physikalisch begründeten Modellierung	63
4.7.1	Erstellung des hydrodynamisch-numerischen Modelles.....	63
4.7.2	Niederschlags-Abfluß-Modellierung mit WaSim-ETH.....	68
4.7.3	Berechnung der Hochwasserszenarien.....	69
5	Ergebnisse der Prognose-ANN	73
5.1	Prognose-ANN ohne Merkmale für NA-Prozesse im Zwischeneinzugsgebiet.....	73
5.1.1	MLFN für die HD-Szenarien	73
5.1.2	SOM für die HD-Szenarien.....	76
5.2	Prognose-ANN mit Merkmalen für NA-Prozesse im Zwischeneinzugsgebiet	80
5.2.1	Das Zustandsmerkmal	80

Inhalt

5.2.2	Test der Netztypen für die NA-Prozesse	86
5.2.3	Kombination der Subnetze mit dem NA-Netz	94
5.2.4	MLFN für die HDNA-Szenarien	96
5.2.5	SOM für die HDNA-Szenarien	101
5.3	Vergleich der Prognose-MLFN	102
6	Diskussion	105
7	Zusammenfassung	108
	Symbolverzeichnis	110
	Abbildungen	111
	Tabellen	115
	Literaturverzeichnis	116
	Anhang	120

1 Einleitung

Hydrologische Extremereignisse charakterisieren und prägen die menschliche Geschichte wie sonst kaum andere Naturkatastrophen. Bereits die ersten bekannten großen Zivilisationen wurden an Flußläufen gegründet, da diese zum einen als Transportwege, aber auch und vor allem zum Zwecke der Bewässerung genutzt werden konnten. Seither ist die menschliche Siedlungsgeschichte eng verknüpft mit dem Gewässernetz des Planeten. Mit der Nutzbarmachung der Wasserkraft und dem Wasserbedarf der Industrie kamen weitere Standortfaktoren hinzu, die die Nutzung der Flußtäler als menschlichen Siedlungsraum förderten. Und seit jeher ist der Mensch gezwungen mit den damit verbundenen Risiken und Nebenwirkungen zu leben.

In heutiger Zeit ist dieses Thema aktuell wie eh und je. Wenngleich der technische Fortschritt in der jüngeren Geschichte Neuerungen, wie zum Beispiel Hochwasserrückhaltebecken, Wetterprognosen oder moderne Kommunikationsmedien hervorgebracht hat, die sehr hilfreich im Umgang mit Hochwasserereignissen sein können, hat sich nicht viel an der Brisanz der Problematik geändert. Eines der Hauptprobleme von Naturkatastrophen jeglicher Art ist ihre Seltenheit. Ein Hochwasser vom Ausmaße des Ereignisses im August 2002 im Erzgebirge hat keiner der Augenzeugen vorher je selbst erlebt. Die Erkenntnisse, die man daraus gewinnt, werden möglicherweise in den betroffenen Gebieten erst wieder von Nutzen sein, wenn das Hochwasser in Vergessenheit geraten ist. Allerdings spielen auch Faktoren eine Rolle, die das traditionelle Problem verschärfen. Besonders hydrologische Extremereignisse werden heute wie nie zuvor immer wieder im Zusammenhang mit der globalen Klimaänderung diskutiert. Wenngleich es unmöglich ist, das einzelne Ereignis eindeutig dem Klimawandel anzulasten, wird ein Zusammenhang zwischen der Auftretenswahrscheinlichkeit eines Extremhochwassers und dem Wirken der modernen Zivilisation kaum mehr bestritten. Doch unser Einfluß beschränkt sich nicht nur auf das globale Klima und die meteorologischen Inputs in das System Einzugsgebiet. Der moderne Mensch ist befähigt, gewollt – aber leider auch ungewollt – in immer kürzerer Zeit immer größere Eingriffe in der Landschaft vorzunehmen, die sich in der Ausprägung und Verformung der Hochwasserwellen niederschlagen. Wasserbauliche Maßnahmen an den Flußläufen, aber gleichfalls großflächige Änderungen der Landnutzung, zum Beispiel durch das Waldsterben in den

Einleitung

Kammlagen des Osterzgebirges in den siebziger und achtziger Jahren des letzten Jahrhunderts, sind als Beispiele dafür anzuführen.

Um einerseits diese Gefahren richtig einzuschätzen und andererseits die zerstörerische Wirkung solcher Hochwasserereignisse, wie jenes im August 2002, welches 12 Menschenleben kostete und Schäden im Wert von einer Milliarde Euro hinterließ (LfUG 2004), zu mindern, bedarf es großer Anstrengungen auf den Gebieten der Hochwasservorsorge und auch des operativen Hochwasserschutzes. Insbesondere letzterer kann um so effektiver erfolgen, je schneller und präziser entsprechende Hochwasserprognosen und -warnungen möglich sind.

Für die Hochwasservorhersage in großen, schiffbaren Gewässern wie Elbe oder Rhein finden heute hydrodynamisch-numerische Modelle Anwendung. Damit ist man in der Lage, in adäquater Weise Rückstaueffekte an Zusammenflüssen, Ausuferungen und Retentionseffekte auf Überflutungsflächen oder den Einfluß von Deichen bei Modellierung und Vorhersage einzubeziehen. Im Prognosefall kann hierbei auf genaue Meßreihen an oberstromig gelegenen Pegeln zurückgegriffen werden. Zuflüsse aus unbeobachteten Zwischeneinzugsgebieten sind aufgrund der Größenverhältnisse meist vernachlässigbar. Für die Prognose in schnellreagierenden Mittelgebirgseinzugsgebieten werden demgegenüber für die Wellenablaufprozesse größtenteils konzeptionelle Ansätze verwendet. Zur nachträglichen Analyse von Hochwasserereignissen werden unter anderem hydrodynamische Modelle genutzt, die den ablaufenden Prozessen in Flußabschnitten, an denen die Flüsse ausuferen oder sich an Bauwerken und Zusammenflüssen aufstauen, gerecht werden. Leider sind deren Einsatzmöglichkeiten zur Prognose hier stark limitiert, da der Rechenaufwand doch beträchtlich ist. Einbeziehung von Unsicherheiten durch Monte-Carlo-Simulationen und umfangreiche Szenarienrechnungen sind damit undenkbar.

In zunehmendem Maße gibt es Versuche, Methoden der künstlichen Intelligenz, wie künstliche neuronale Netze oder Neuro-Fuzzy-Systeme hydrologischen Anwendungen und im speziellen der Durchflußvorhersage nutzbar zu machen. Unübersehbarer Vorteil solcher Ansätze ist deren Flexibilität, Robustheit und Schnelligkeit. Während ein hydrodynamisch-numerisches Modell in Abhängigkeit der räumlichen Auflösung mit Zeitschritten von einigen Sekunden die Ausgabegrößen für den gesamten Flußlauf kontinuierlich berechnet, kann beispielsweise ein Multilayer-Feedforward-Netz mit Hilfe einiger Matrizenoperationen die interessante Zielgröße direkt aus den relevanten Inputgrößen in Sekundenbruchteilen ermitteln. Doch die Vorhersagegenauigkeit von Extremereignissen im speziellen ist mit solchen datengetriebenen Ansätzen durchaus

mit Vorsicht zu betrachten, da sich die Prognosefähigkeit nur auf den Bereich, der zum Anlernen des Systems verwendet wurde, beschränkt.

In diesem Spannungsfeld findet sich die vorliegende Arbeit wieder. Hierin wird für die unteren Flußabschnitte schnellreagierender Mittelgebirgseinzugsgebiete eine Methodik entwickelt, die die Vorteile hydrodynamischer und hydrologischer Modellierung mit denen künstlicher neuronaler Netze verbindet. Diese Methodik soll am Einzugsgebiet Freiburger Mulde getestet werden.

Gemeinsam mit der Dissertation von Cullmann (2007), welche den Fokus auf die Beschreibung der Niederschlags-Abfluß-Prozesse unter Berücksichtigung der räumlichen Niederschlagsverteilung in Mittelgebirgseinzugsgebieten mit künstlichen neuronalen Netzen legt, zeigt die vorliegende Arbeit eine komplette Methodik zur Hochwasservorhersage mit künstlichen neuronalen Netzen in schnellreagierenden Mittelgebirgseinzugsgebieten auf.

2 Künstliche neuronale Netze zur Durchflußvorhersage

Künstliche neuronale Netze (artificial neural networks, ANN) sind sehr leistungsfähige, vielseitig einsetzbare Berechnungsmodelle. Das gemeinsame Konzept hinter der Vielfalt an unterschiedlichen Architekturen bildet das Vorbild biologischer neuronaler Netze, die in Sekundenbruchteilen komplizierteste Aufgaben – wie z.B. das Lesen, motorische Steueraufgaben, oder das Wiedererkennen von Dingen – zu lösen in der Lage sind. Die Nervenzellen, bzw. die biologischen Neuronen, nehmen dabei Informationen über die Dendriten auf, verarbeiten sie im Zellkörper und geben sie über das Axon und die Synapsen als elektrische Impulse zu den benachbarten Neuronen weiter. Ein Teil dieser Nervenstruktur existiert bereits bei der Geburt. Andere Verbindungen werden durch Lernprozesse mit der Zeit unterschiedlich stark entwickelt oder verkümmern.

Die künstlichen Neuronen sind parallel arbeitende Berechnungseinheiten, die durch einen Input stimuliert werden und diesen in einen Output transformieren, wobei ihre Verbindungen zu benachbarten Neuronen in Form von Gewichten unterschiedlich stark ausgeprägt sind. Das Lernen in Bezug auf künstliche neuronale Netze besteht in der Anpassung eben dieser Gewichte. Entsprechend ihrem biologischen Pendant eignen sich ANN ideal für Klassifikationsaufgaben, zur Mustererkennung oder zur Approximation nichtlinearer funktionaler Zusammenhänge.

Die Geschichte künstlicher neuronaler Netze wurde von kreativen Menschen verschiedenster Wissenszweige geprägt. Einige Grundlagen wurden bereits ab dem Ende des 19. Jahrhunderts in interdisziplinären Arbeiten in Physik, Psychologie und Neurophysiologie von Wissenschaftlern wie Hermann von Helmholtz, Ernst Mach und Ivan Pavlow gelegt. Diese frühen Arbeiten begannen Theorien über das Lernen, das Sehen und Erlernen von Reiz-Reaktions-Mustern zu entwickeln. In den vierziger Jahren des 20. Jahrhunderts zeigten Warren McCulloch und Walter Pitts (McCulloch, Pitts, 1943), daß Netzwerke künstlicher Neuronen prinzipiell arithmetische und logische Funktionen berechnen können.

In den späten fünfziger Jahren wurde von Frank Rosenblatt das Perceptron-Netz erfunden (Rosenblatt, 1958) und die Fähigkeit der Mustererkennung gezeigt. Man stieß aber schnell an die Grenzen der Möglichkeiten des Perceptrons. Die Entwicklung komplexerer Netztypen und neuer Lernalgorithmen scheiterte jedoch zunächst. In den siebziger Jahren entwickelten Teuvo Kohonen (Kohonen, 1972) und James Anderson

Künstliche neuronale Netze zur Durchflußvorhersage

(Anderson, 1972) neue, assoziative Netztypen, die Beziehungen zwischen In- und Outputvektoren erlernen konnten. Stephen Grossberg (Grossberg, 1976) arbeitete an selbstorganisierenden Netzwerken, die in ihrer Funktionsweise dem menschlichen Sehen und der Bildverarbeitung nachempfunden waren. In den Achtziger Jahren war die Rechentechnik auf einem Niveau angelangt, welches die Forschung auf dem Gebiet der künstlichen neuronalen Netze außerordentlich beschleunigte. Ein Schlüsselereignis war die Entwicklung des Backpropagation-Algorithmus (Rumelhart, McClelland, 1986), was das Training von Multilayer-Perceptron-Netzen (oder im Zusammenhang mit der Funktionsapproximation meist bezeichnet als Multilayer-Feedforward-Networks, MLFN) ermöglichte. In den Folgejahren konnten ANN immer neue Anwendungsgebiete finden (nach Hagan et al., 1996).

Bereits 1996 benutzten Solomatine und Avila Torres ANN erfolgreich als einfaches und vor allem schnelles Berechnungsmodell, um ein hydrodynamisches Modell (MIKE11) zu ersetzen. Der Hintergrund war der Versuch, die Bewirtschaftung eines Systems dreier Stauseen in Venezuela in bezug auf Schiffbarkeit zu optimieren. Die Motivation für diese Arbeit resultierte hauptsächlich aus zwei Gründen: Die Berechnungsmodule des komplexen hydrodynamischen Modells extern aufzurufen, stellte sich als unmöglich heraus und die Rechenzeit der hydrodynamischen Komponente war zu lang für eine Einbettung in einen Optimierungsalgorithmus. Es wurden MLFN mit einer verdeckten Schicht implementiert. Die ANN wurden auf Wochenmittel von Wasserständen an verschiedenen Zielquerschnitten trainiert. Die Netzinputs bestanden aus einzelnen Abflußwerten von den Teileinzugsgebieten und den Abgaben aus den Stauseen. Die besten Resultate wurden erreicht, wenn MLFN für jeden einzelnen Zielquerschnitt trainiert wurden, d.h. mit nur einem Outputneuron.

Unter Verwendung derselben Datensätze testeten Dibike et al. (1999) Variationen von MLFN in bezug auf Trainingsalgorithmus und Transferfunktionen, sowie verschiedene Netzarchitekturen (Radial Basis Function Networks, RBFN und Recurrent Neural Networks, RNN). Die Resultate dieser Arbeit wurden verglichen hinsichtlich unterschiedlicher Parameter von Rechenaufwand beim Netztraining und Performance der Netze. Die Performance betreffend, gab es nur geringe Differenzen. Die rekurrenten neuronalen Netze zeigten leicht bessere Anpassungen als die RBFN. Doch insgesamt konnte für alle betrachteten Netztypen gezeigt werden, daß sie die Fähigkeit besitzen, die Resultate des hydrodynamisch-numerischen Modells recht gut zu reproduzieren.

Die meisten Anwendungen künstlicher neuronaler Netze in der Hydrologie sind jedoch reine datengetriebene Ansätze. Hinsichtlich der Modellerstellung für Extremereignisse, wie z.B. zur Hochwasserprognose, ist die Verwendung von ausschließlich gemessenen Daten kaum zielführend, da Extremereignisse ihrer Natur nach in erster Linie selten sind. So trainieren Shrestha et al (2005) MLFN anhand eines einzigen Hochwasserereignisses an der Mosel, je ein weiteres wird als Validierungs- und Testdatensatz verwendet.

Chen und Adams (2006, 1) verwenden neuronale Netze in Kombination mit konzeptionellen Niederschlags-Abfluß-Modellen. Dabei wird mit den konzeptionellen Ansätzen die Abflußbildung und -konzentration in drei oberstromigen Teileinzugsgebieten beschrieben, welche separat nicht validiert werden können, da die Gebietsauslässe unbeobachtet sind. Ein MLFN füllt die Lücke zu einem beobachteten unterstromigen Pegel, an dem beobachtete Ganglinien vorliegen. Somit erfüllt das MLFN eine ähnliche Funktion wie in dieser Arbeit, indem es die Beschreibung des Wellenablaufes mit dem Einfluß der NA-Prozesse des Zwischeneinzugsgebietes verbindet. Das als „hybride Form eines NA-Modells“ bezeichnete Modellsystem wird jedoch von Jain (2006) stark kritisiert, da eine quasi gemeinsame Eichung in Reihe geschalteter konzeptioneller und Black-Box-Ansätze nicht besonders zielführend sein kann.

Künstliche neuronale Netze fanden ebenfalls bereits Eingang in die Niederschlags-Abfluß-Modellierung. Dibike und Solomatine (2001) verglichen für das oben erwähnte venezolanische Einzugsgebiet MLFN mit RBFN und einem konzeptionellen Tank-Modell. Mit den beiden ANN-Ansätzen konnten bessere Ergebnisse erzielt werden, als mit dem konzeptionellen Modell, wengleich darauf hingewiesen wird, daß diese Tatsache nicht verallgemeinerbar ist. Zudem wurde hier der Schwerpunkt nicht auf die Betrachtung von Extremereignissen gelegt. Die Autoren konstatieren weiterhin, daß die Anwendung neuronaler Netze nur erfolgversprechend sein kann, wenn die relevanten Inputgrößen an vergangenen Regenwerten und Durchflüssen an Oberliegerpegeln vorher sorgfältig ausgewählt wurden. Die MLFN zeigten ferner leicht bessere Anpassungen als die RBFN, obgleich der Trainingsaufwand für RBFN leicht geringer war.

Neben der Anwendung von ANN sind in der Literatur auch Beispiele für die Anwendung von Neuro-Fuzzy-Systemen zu finden. Dieser Typ von Berechnungsmodellen, die zu „künstlicher Intelligenz“ zählen, soll hier nicht näher

Künstliche neuronale Netze zur Durchfluvorhersage

vorgestellt werden. Es sollen nur einige Artikel zitiert werden, die den Einsatz von ANN und Neuro-Fuzzy-Systemen als Black-Box-Modelle vergleichen.

Bernd et al. (1999) verglichen diese beiden Modelltypen künstlicher Intelligenz auf ihre generelle Eignung für die Black-Box-Modellierung anhand zweier mathematischer und dreier komplexer technischer Beispiele nichtlinearer Systeme. NFS (neuro fuzzy systems) beruhen auf der Grundidee, daß nichtlineare Systeme durch weich ineinander übergehende lineare Systeme angenähert werden können. Das Fuzzysystem hat strukturell betrachtet ein Erscheinungsbild ähnlich dem der neuronalen Netze; deswegen verwendet man die Bezeichnungen Neuro-Fuzzy-System oder auch englisch „fuzzy network“. Dieses Konzept ist bekannt von Multi-Modell-Anwendungen, welche allerdings übergangslos („crisp“) zwischen den Modellen umschalten. Funktionale Fuzzysysteme sind vorteilhaft, wenn die Identifikationsaufgabe durch lineare Funktionen angenähert werden kann, während Neuronale Netze als Punkt-zu-Punkt-Projektion mit nichtlinearer Interpolation dazwischen verstanden werden können. ANN benötigen eine ausreichende Anzahl Parameter für eine entsprechende Anpassungsgüte, währenddessen die Qualität von NFS qualitativ nicht unbedingt durch eine Vergrößerung der Parameterzahl verbessert werden kann, weil sich nicht jede nichtlineare Funktion vernünftig in quasilineare Abschnitte unterteilen läßt. Wenn dynamische Prozesse modelliert werden, so fassen die Autoren zusammen, benötigen sowohl ANN als auch FNS eine sinnvolle Auswahl der Inputdaten und effiziente Lernmethoden, um ein hochwertiges Input-Output-Modell zu erstellen. Wenn man ein interpretierbares Modell benötigt, sind Fuzzysysteme zu bevorzugen. Für die betrachteten Beispiele wurden vergleichbare Anpassungen gefunden, so daß sich beide Methoden generell für Vorhersagezwecke eignen.

Stübner und Gemmar (1997) testeten MLFN und Sugeno-Neurofuzzy-Systeme (SNFS) für Wasserstandsvorhersagen an der Mosel im westlichen Deutschland auf der Grundlage 14 beobachteter Hochwasserereignisse in Stundenauflösung. Dabei wurden jeweils ANN bzw. SNFS für eine 6-Stunden-Prognoseweite für den Wellenablauf zwischen mehreren oberstromigen Zulaufpegeln und dem jeweiligen Zielpegel (Trier und Cochem) erstellt. 11 der 14 Ereignisse wurden für das Netztraining verwendet. Ein Netz mit zwei verdeckten Schichten mit je 10 Neuronen konnte den Zielpegel am besten treffen. Trotz der Tatsache, daß nur zwei Ereignisse für die Kalibrierung der SNFS verwendet wurden, sind die Validierungsergebnisse besser als für das MLFN. Die Autoren bemerken, daß der gegebene Umfang an Trainingsereignissen für einen Black-Box-Ansatz wie ANN unzureichend für die

Erstellung eines zuverlässigen Prognosetools ist. In Bezug auf die SNFS wird konstatiert, daß diese durch die Möglichkeit der Einbeziehung von Fachwissen bei deren Gestaltung für ein sehr begrenztes Angebot an Trainingsdaten geeigneter sind.

Alvisi et al. (2006) präsentierten und diskutierten drei datengetriebene Wasserstandsvorhersagemodelle auf der Basis von Niederschlagswerten für das Einzugsgebiet des Flusses Reno bei Bologna mit einer Fläche von reichlich 1000 km² am Pegel in Casalecchio. Sie verglichen MLFN mit einer einzelnen verdeckten Schicht mit Fuzzy-Logic-Ansätzen des Mamdani- und des Takagi-Sugeno-Typen (MNFS und SNFS). Dabei wurden die räumlich und zeitlich hochaufgelösten Niederschlagswerte auf zwei Weisen in Inputdatensätze umgewandelt: Ein erster (ARI, aggregated rainfall information) reduziert die Information auf wenige räumlich und zeitlich zusammengefaßte Werte, während der zweite, umfangreichere Datensatz mehr Information über Regenverteilung und -verlauf beibehält (DRI, distributed rainfall information). Es wurde gezeigt, daß die Fuzzy-Ansätze bessere Ergebnisse zeigen, wenn die Zahl der Inputgrößen limitiert ist (ARI). Demgegenüber konnten die MFLN ihre Performance stark verbessern, wenn die detailliertere Information angeboten wird. Für Kurzzeitprognosen konnten beide Modelltypen vergleichbar gute Ergebnisse liefern. Bei größeren Vorhersagehorizonten verschlechterte sich die Performance bei der Verwendung der ARI insbesondere hinsichtlich des ansteigenden Astes. Die Fuzzy-Netze bestätigten die Tatsache (siehe auch Abebe et al., 2000), daß sie nur mit einer begrenzten Anzahl an Inputvariablen umgehen und damit detaillierte Informationen schlechter verarbeiten können.

Eine interessante Kopplung dieser Methoden findet sich bei Corani und Guariso (2005): Die datengetriebene Abflußmodellierung mit neuronalen Netzen wird eingebettet in ein Neuro-Fuzzy-System als Rahmenwerk. Dabei übernimmt das NFS zunächst eine Klassifizierung der Vorfeuchte in Form eines Vorregenindexes und bildet danach ein gewichtetes Mittel aus verschiedenen vorfeuchtespezifisch trainierten ANN. Die Autoren stellen eine deutliche Verbesserung gegenüber der klassischen Anwendung von ANN zur datengetriebenen Niederschlags-Abfluß-Modellierung fest.

3 Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

3.1 Grundprinzipien der entwickelten Methodik zur Hochwasservorhersage

Der instationäre Prozeß des Wellenablaufes in offenen Gerinnen wird durch die beiden als Saint-Venant-Gleichungen bekannten Differentialgleichungen beschrieben (Dyck und Peschke, 1989). Die Kontinuitätsgleichung berücksichtigt die Massenerhaltung. Die dynamische Grundgleichung (oder auch Impulsgleichung, Bewegungsgleichung) approximiert die Energiebilanz des Abflußprozesses und gliedert sich in die Terme der Trägheit, des Druckes, sowie des Reibungs- und Gerinnegefälles (ausführlichere Beschreibung unter 3.3.1).

Unter bestimmten Annahmen kann diese dynamische Grundgleichung vereinfacht werden. So kann unter Vernachlässigung von Rückstauwirkungen auf den Trägheitsterm verzichtet werden. Daraus resultiert der Diffusionswellenansatz, der für die Beschreibung des Wellenablaufes in verschiedenen Niederschlags-Abfluß-Modellen integriert ist.

Doch speziell im Hochwasserfall sind mit zunehmender Konzentration des Abflusses in den weniger steilen Unterläufen schnell reagierender Mittelgebirgseinzugsgebiete Überflutungen von Vorländern und Rückstauwirkungen an Zusammenflüssen zu beobachten. Die Saint-Venant-Gleichungen sind im allgemeinen jedoch nicht analytisch lösbar, und die gängigen Modelle, die auf numerischen Lösungen beruhen, sind sehr rechenintensiv und kompliziert. Nicht selten ist der Anwender mit numerischen Instabilitäten konfrontiert. Damit ist eine mögliche Verwendung solcher Modelle bei der operationellen Hochwasservorhersage stark limitiert, insbesondere, wenn beispielsweise durch Ensemblerechnungen Unsicherheiten der Eingangsdaten Rechnung getragen werden soll.

Wie im vorhergehenden Kapitel beschrieben, sind Methoden der künstlichen Intelligenz generell in der Lage, jegliche Input-Output-Beziehung abzubilden. Unter anderem gehören künstliche neuronale Netze zu den vielseitig einsetzbaren, robusten und flexiblen datengetriebenen Berechnungsmodellen, die in der Lage sind, Datenvektoren zu klassifizieren und Muster oder beliebige funktionale Zusammenhänge zu erlernen und wiederzugeben. Da sich die notwendigen mathematischen Operationen bei der Anwendung auf einfache Matrizenoperationen

Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

beschränken, ist der Rechenaufwand sehr gering. In jedem Falle ist ein ANN jedoch ein Black-Box-Modell, welches nur in der Lage ist, Zusammenhänge in dem Bereich richtig wiederzugeben, der ihm durch den Lernprozess bekannt ist (Abbildung 1). Im Zusammenhang mit der NA-Modellierung mit neuronalen Netzen konstatierten Minns and Hall (1996): "ANN are a prisoner of their training data". Im Falle von Extremereignissen kann jedoch davon ausgegangen werden, daß für nur äußerst wenige extreme Hochwasserereignisse Daten vorliegen. Das heißt, die Zahl der tatsächlich verfügbaren Stichproben aus der großen Menge theoretisch möglicher Hochwasserereignisse ist zu gering, um damit zuverlässig ein Neuronales Netz zu trainieren (Abbildung 1).

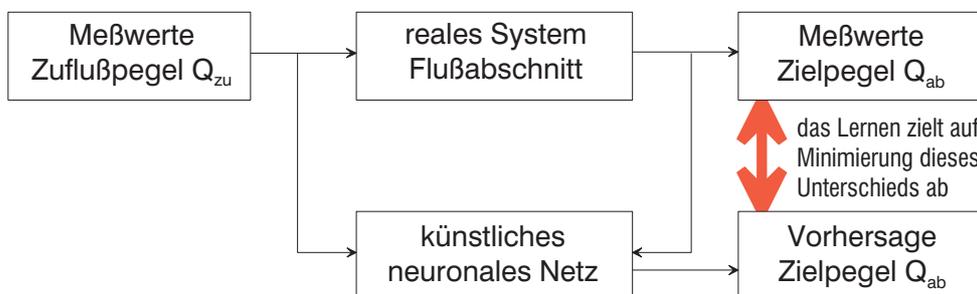


Abbildung 1: Lernen bei datengetriebener Modellierung (nach Solomatine, 2002)

Ziel dieser Arbeit ist nun, die Vorteile hydrodynamisch-numerischer Modellierung und künstlicher Neuroner Netze zu kombinieren und eine Methodik zu entwickeln, die sich unter Beibehaltung der Zuverlässigkeit physikalisch begründeter Modelle der Einfachheit und Robustheit neuronaler Netze bedient. Dabei soll ein kalibriertes und validiertes hydrodynamisch-numerisches Modell dazu genutzt werden, eine ausreichende Datenbasis an Input-Output-Paaren zu erstellen, die sämtliche theoretisch denkbaren Konstellationen umfaßt, ANN zuverlässig zu trainieren. Damit wird die generelle Methodik zum Training eines künstlichen neuronalen Netzes (Abbildung 1) um die Modellebene ergänzt, wie in Abbildung 2 dargestellt. Die Vorgehensweise der Netzerstellung soll sich dabei grob betrachtet in folgende Schritte gliedern:

- Erstellung, Kalibrierung und Validierung eines hydrodynamisch-numerischen Modells für ausgewählte Flußabschnitte

Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

- Generierung von Durchflußganglinien an den oberstromigen Zuflußpegeln für das hydrodynamische Modell, Berücksichtigung der unbeobachteten lateralen Zuflüsse entlang der modellierten Fließstrecken durch Niederschlags-Abfluß-Modell
- Berechnung der gesamten Palette theoretisch möglicher Szenarien mit dem hydrodynamischen Modell und Erstellung einer Datenbasis an Input-Output-Paaren für das Training eines künstlichen neuronalen Netzes
- Training und Test des neuronalen Netzes

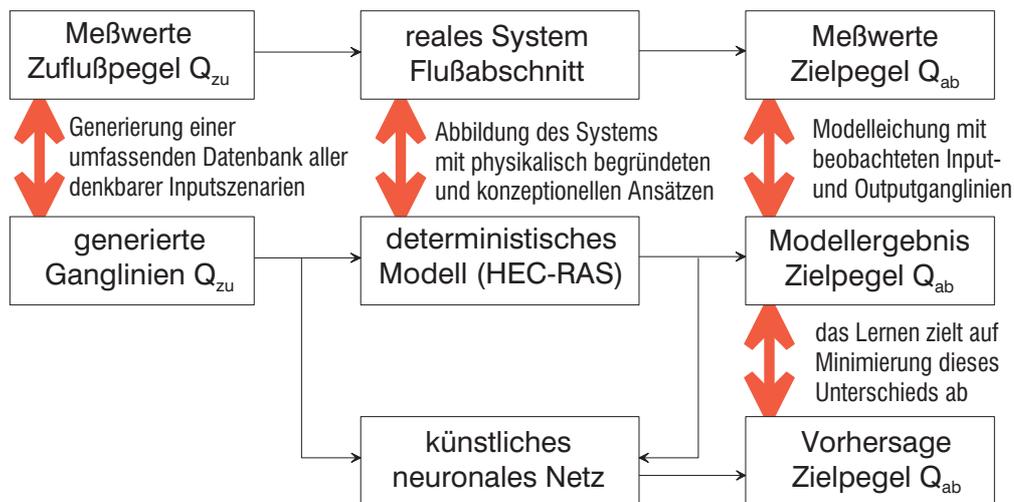


Abbildung 2: Aus dem Defizit an Trainingsdaten abgeleitete allgemeine Methodik

Für die Modellierung des Wellenablaufes – sowohl mit einem hydrodynamischen Modell, als auch mit künstlichen neuronalen Netzen – besteht natürlich die primäre Eingangsinformation aus den Ganglinien an den oberstromigen Rändern der entsprechenden Flußabschnitte, für die im Anwendungsfall Pegelbeobachtungen oder Prognosen verfügbar sind. Doch ebenso wie die Hochwasservorhersage für die Quellgebiete und die oberen Teile eines Einzugsgebietes (z. B. Cullmann, 2007) neben der Akzentuierung der Niederschlags-Abfluß-Prozesse auch nicht auf eine Berücksichtigung von Wellenablaufprozessen verzichten kann, verlangt die Beschreibung der hydrodynamischen Prozesse auch die Einbeziehung lateraler Zuflüsse aus dem Zwischeneinzugsgebiet. Da für diese im allgemeinen keine separaten Ganglinien aus Pegeldaten oder Vorhersagemodellen existieren, müssen die zugrundeliegenden Niederschlags-Abfluß-Prozesse vom Vorhersagewerkzeug berücksichtigt werden. Abbildung 3 illustriert diesen Sachverhalt.

→ **Zwischeneinzugsgebiet, Zwischengebiet:** Teil des gesamten Einzugsgebietes, welcher in die mit dem hydrodynamisch-numerischen Modell beschriebenen Flußabschnitte entwässert und für den keine expliziten Hochwasservorhersagemodelle oder Pegelbeobachtungen existieren.

Die oberen Flußabschnitte in Abbildung 3, bei denen der Wellenablauf mit einfachen, in Niederschlag-Abfluß-Modellen enthaltenen Ansätzen zur Beschreibung des Wellenablaufes modellierbar ist, werden mit hier als NA-ANN bezeichneten künstlichen neuronalen Netzen beschrieben, die das Hauptgewicht auf die Abflußbildung und -konzentration legen.

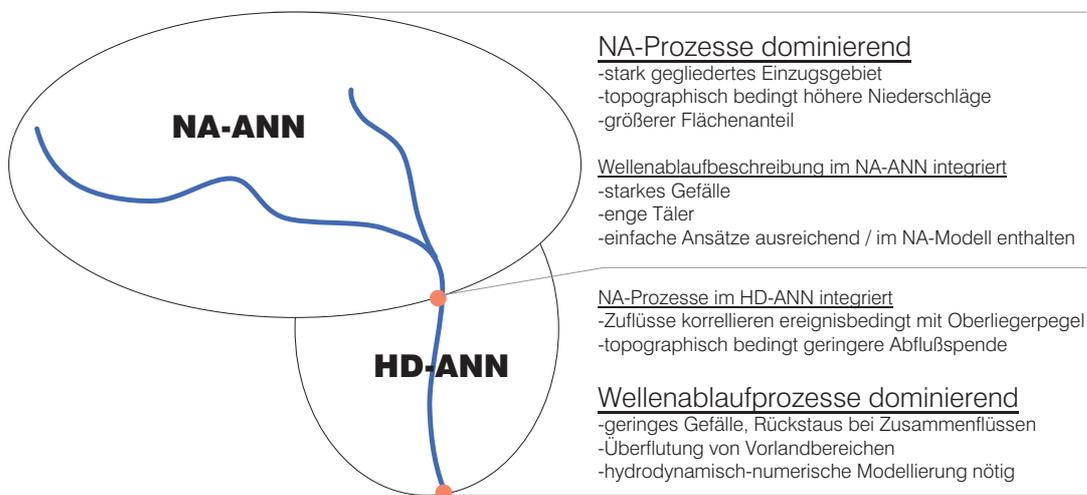


Abbildung 3: räumliche Gliederung des Gesamteinzugsgebietes nach dominanten Prozessen, schematisch

→ **NA-ANN:** künstliche neuronale Netze zur Beschreibung der Niederschlags-Abfluß-Prozesse (NA)

→ **HD-ANN:** künstliche neuronale Netze zur Beschreibung der hydrodynamischen (HD-) Prozesse ohne Verwendung expliziter Eingangsgrößen für die Berücksichtigung der Niederschlags-Abfluß-Prozesse im Zwischengebiet

→ **HDNA-ANN:** künstliche neuronale Netze zur Beschreibung der hydrodynamischen Prozesse mit Verwendung expliziter Eingangsgrößen für die Berücksichtigung der Niederschlags-Abfluß-Prozesse im Zwischengebiet

Für die neuronalen Netze zur Beschreibung komplizierter hydrodynamischer Prozesse, für deren Erstellung hier eine Methodik aufgezeigt werden soll, werden zwei Bezeichnungen eingeführt. Die Unterscheidung beruht in der angewandten Strategie

Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

zur Berücksichtigung des Einflusses der oben beschriebenen unbeobachteten Zuflüsse aus dem Zwischeneinzugsgebiet.

Die Verwendung von HD-ANN unterstellt also, daß die Information der Zuflußganglinien an den Schnittstellen zu den NA-ANN für die Prognose von Wasserstand und Durchfluß am unteren Gebietsauslaß ausreichend ist. Diese Annahme ist sicherlich gerechtfertigt, wenn der Einfluß des Zwischeneinzugsgebietes gering ist und die unbeobachteten Zuflüsse abhängig vom Zufluß am oberstromigen Zuflußpegel sind.

Die HDNA-ANN benutzen demgegenüber zusätzliche Informationen, welche die Niederschlags-Abfluß-Prozesse im Zwischeneinzugsgebiet explizit beschreiben.

Für das Verständnis der vorgestellten Methodik werden bereits hier Definitionen folgender im Zusammenhang künstlicher neuronaler Netze gebräuchter Begriffe nötig:

- **Merkmal:** *Eingangsgröße für das künstliche neuronale Netz, meteorologische oder hydrologische Meßgröße mit zeitlichem Bezug zum Targetwert oder aus mehreren aggregierter Wert*
- **Zustandsmerkmal:** *Merkmal zur Beschreibung des Systemzustandes bei Ereignisbeginn*
- **Ereignismerkmal:** *Merkmal zur Charakterisierung des Ereignisverlaufs und seiner Dynamik*
- **Merkmalsvektor, Inputvektor:** *Gesamtheit der verwendeten Merkmale in einer vordefinierten Reihenfolge*
- **Outputvektor:** *Ergebnisvektor des neuronalen Netzes, Netzoutput*
- **Targetvektor, Targetwert:** *Beim Training des neuronalen Netzes angestrebter Netzoutput*
- **Netztraining:** *Anpassung der Netzparameter (Gewichte) zur Minimierung der Differenz zwischen Target und Netzoutput*

In Abbildung 4 sind die Methodik der Netzerstellung und die nötigen Schritte zur Netzanwendung schematisch abgebildet. Die spezifischen Grundprinzipien dieser Methodik lassen sich folgendermaßen formulieren:

- ein einziges Prognosenetz für eine variable Prognoseweite
- Merkmalsdefinition physikalisch fundiert mit Hilfe der deterministischen Modelle
- Methodik unabhängig vom verwendeten Modell und untersuchten Einzugsgebiet
- Erstellung der ANN unter Verwendung von allgemein verbreiteter Standardsoftware möglich (MATLAB)

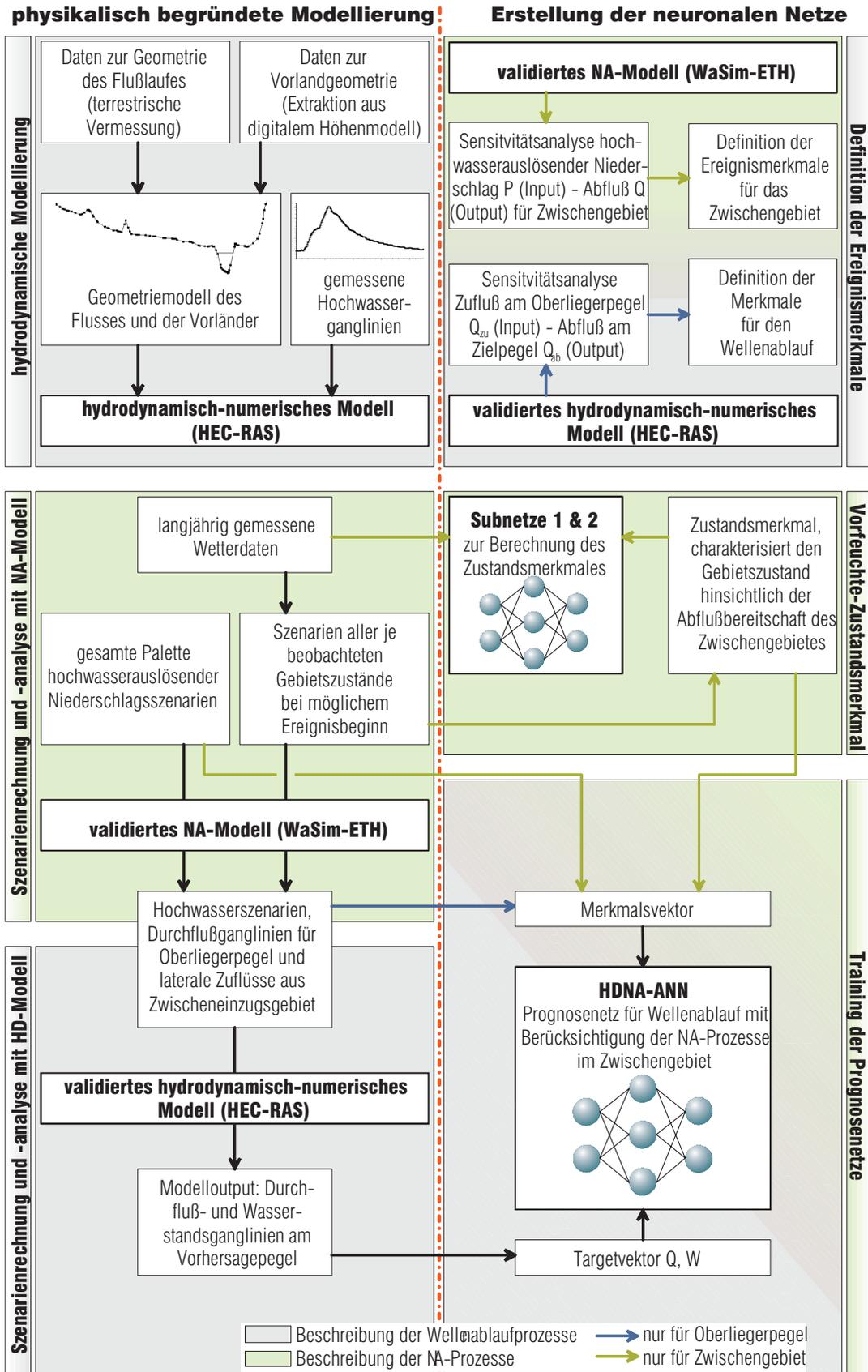


Abbildung 4: Gesamtmethodik der Einzugsgebietsmodellierung und Erstellung eines ANN für den Wellenablauf unter Berücksichtigung des Zwischeneinzugsgebietes

Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

Mit dem Aufbau eines konventionellen hydrodynamischen Modells zur Beschreibung des Wellenablaufs im Untersuchungsgebiet ist das Werkzeug zu einer umfassenden Szenarienrechnung für die Erstellung einer Trainingsdatenbank gegeben. Die Eingangsgrößen für die Berechnung der alle denkbaren Hochwassersituationen umfassenden Trainingsdaten sind die Durchflußganglinien, die mit dem Niederschlags-Abfluß-Modell aus generierten Niederschlags- und natürlichen Vorfeuchteszenarien ermittelt wurden. Eine Analyse der Sensitivität der Ganglinie am Prognosepegel auf Einzelwerte aus der Ganglinie am Zuflußpegel mit dem hydrodynamischen Modell führt zur Auswahl der relevanten Merkmale.

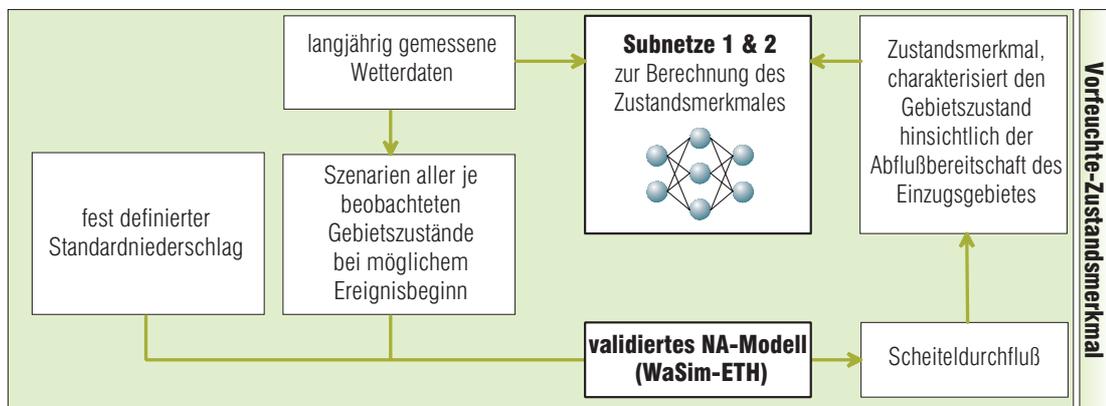


Abbildung 5: Methodik der Erstellung der Subnetze für die Berechnung des Zustandsmerkmals

Im Falle der expliziten Berücksichtigung der Niederschlags-Abfluß-Prozesse im Zwischeneinzugsgebiet (Abbildung 5) durch das Prognose-ANN (HDNA-ANN) wird die Abhängigkeit der Vorfeuchte von der Wettergeschichte in Form eines einzelnen Zustandsmerkmals beschrieben. Dieses charakterisiert die Vorfeuchte des Einzugsgebietes und faßt als solche die relevante Wettergeschichte in einem integralen Wert zusammen.

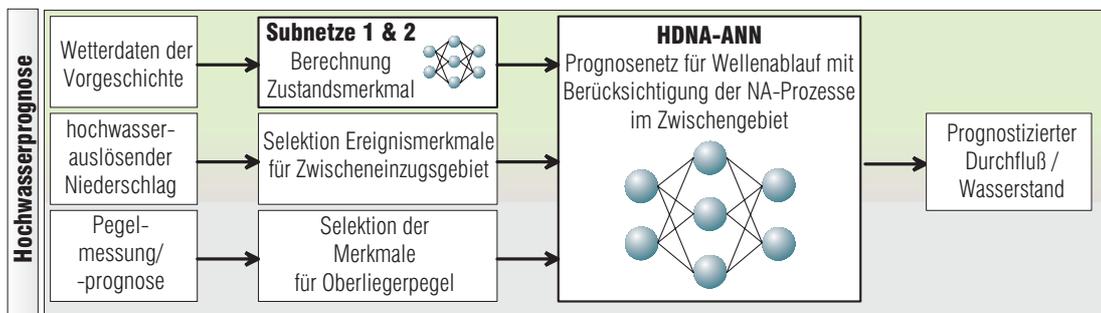


Abbildung 6: Gesamtmethodik der Netzanwendung für HDNA-ANN

Das Zustandsmerkmal berechnet sich als Abflußscheitelhöhe Q_{\max} der Gebietsreaktion auf einen Standardniederschlag im Niederschlags-Abfluß-Modell. Zur Reproduktion des Zustandsmerkmals aus der Vorgeschichte werden die beiden Subnetze (Abschnitt 3.6.2) erstellt, welche separat trainiert werden. Die Einzelwerte des hochwasserauslösenden Niederschlags im Zwischengebiet werden in Abhängigkeit der zeitlichen Entfernung zum Vorhersagezeitschritt zu Ereignismerkmalen für die Niederschlags-Abfluß-Prozesse zusammengefaßt.

Die Anwendung des Netzes im Prognosefall – schematisch in Abbildung 6 dargestellt – erfordert als vorbereitenden Schritt nur die Anwendung der Vorschriften zur Transformation der Zuflußganglinien und der meteorologischen Eingangsdaten (Niederschlag, Globalstrahlung, Temperatur, Windgeschwindigkeit und relative Luftfeuchte) in die definierten Merkmale. Im Falle der expliziten Berücksichtigung des Zwischeneinzugsgebietes durch Verwendung eines HDNA-ANN schließt das eine Verwendung der beiden Subnetze zur Berechnung des Zustandsmerkmals aus den meteorologischen Daten der Ereignisvorgeschichte ein.

Im folgenden wird die Methodik detailliert entwickelt und die verwendeten neuronalen Berechnungsmodelle werden vorgestellt.

3.2 Neuronale Berechnungsmodelle

Entsprechend der unterschiedlichen Vorbilder an biologischen neuronalen Netzen (Kapitel 2) und Aufgaben der Datenverarbeitung existieren die unterschiedlichsten Netztypen. In dieser Arbeit soll die Untersuchung der Eignung spezieller Architekturen nur anhand zweier Beispiele erfolgen, die sich in ihrem Lernverhalten deutlich unterscheiden.

3.2.1 Grundlegende Konzepte von Lernregeln

Verschiedene Konzepte des Verständnisses der Informationsverarbeitung mit biologischen neuronalen Netzen haben die Entwicklung unterschiedlicher Typen künstlicher neuronaler Netze inspiriert. Spezielle Netztypen können in ihrer Grundidee als Modelle konkreter neuronaler Strukturen in höherentwickelten Lebewesen verstanden werden, bzw. ist ihre Entwicklung davon motiviert und inspiriert gewesen. Grundsätzlich ist bereits die Vorstellung des Neurons als kleinste autonome Einheit und Rechenbaustein die Basis sowohl biologischer als auch künstlicher neuronaler

Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

Netze. Die Informationsflüsse und die Funktionalität des Netzwerkes werden in der Biologie durch unterschiedlich stark ausgeprägte Verbindungen durch Synapsen zu benachbarten Neuronen bestimmt, die bei künstlichen neuronalen Netzen durch unterschiedliche Gewichte repräsentiert werden. Das biologische Lernen bedeutet nichts anderes als das Ausbilden dieser Verbindungen. Ein Teil der neuronalen Struktur ist bereits bei der Geburt gegeben, andere Teile entwickeln sich durch das Lernen. Dabei werden manche Verbindungen geknüpft oder gestärkt, andere gehen verloren. Dieses Lernen geschieht vor allem im Kindesalter und in der Jugend, ist aber nie abgeschlossen.

Das Anlernen eines künstlichen neuronalen Netzes bedeutet gleichfalls eine Verstärkung oder Abschwächung der Verbindung zwischen verschiedenen Neuronen durch eine Veränderung der Gewichte, die die Wirkung des Outputs des einen Neurons auf das benachbarte Neuron bestimmen. Prinzipiell unterscheidet man zwischen überwachtem und unüberwachtem Lernen (Hagan et al., 1996).

Beim überwachten Lernen wird dem Netz ein Trainingsdatensatz mit Beispielen für das richtige Verhalten des Netzes gezeigt. Dieser Trainingsdatensatz besteht aus einem Inputvektor und einem dazugehörigem Targetvektor. Die Lernregel (oder auch der Trainingsalgorithmus) sorgt dann für eine Anpassung der Gewichte, um die Netzoutputs näher an die Targets heranzubewegen.

Unüberwachtes Lernen benötigt demgegenüber keine vordefinierten Targetwerte. Alle Größen werden im Inputvektor zusammengefaßt. Die Gewichte werden durch die Lernregel dergestalt verändert, daß die Inputs in eine endliche Anzahl von Klassen eingeteilt werden. Die Merkmalsvektoren werden also gewissermaßen nach ihrer Lage im Raum geclustert. Eine beliebige Größe, die Bestandteil des Merkmalsvektors ist, kann bei der Anwendung des trainierten Netzes durch Wiedererkennung eines bekannten Musters aus den anderen Größen bestimmt werden.

Einen klassischen Netztyp für überwachtes Lernen stellt das Multilayer-feedforward-Netz dar, welches sich sowohl zur Klassifizierung als auch zur Funktionsapproximation eignet. Aufgrund seiner Flexibilität und auch seines erfolgreichen Einsatzes bei verschiedenen Anwendungen (z. B. Schütze, 2005) wird dieses als Vertreter der überwacht lernenden Netze verwendet. Cullmann (2007) verwendet für die Beschreibung der Niederschlags-Abfluß-Prozesse Polynomnetze (PoNN), die ebenso überwacht angelern werden. Die Merkmalsauswahl ist dabei praktischerweise in den Trainingsprozeß integriert, findet andererseits aber auf rein empirische Weise statt. In dieser Arbeit sollen die Merkmale mit Hilfe des deterministischen Modells

entsprechend der physikalisch ablaufenden Prozesse definiert werden. Damit entfällt der Vorteil der PoNN und es werden die flexibleren MLFN verwendet.

Ein leistungsfähiger Netztyp, welcher unüberwacht angelernt wird, ist die selbstorganisierende Merkmalskarte (SOM, self-organizing feature map). Schütze (2005) verwendet diesen Netztyp mit großem Erfolg zur Bestimmung von Bewässerungsparametern. Wenngleich die Anzahl der abhängigen Parameter einer Input-Output-Beziehung bei der Beschreibung hydrodynamischer Prozesse weitaus höher und damit die Klassifizierung der Merkmalsvektoren komplizierter ist, soll dieser Netztyp auf eine Eignung für diese Anwendung getestet werden. In dieser Arbeit sollen MLFN und SOM auf ihre Eignung zur Hochwasserprognose untersucht werden.

3.2.2 Multilayer feedforward networks

MLFN sind künstliche neuronale Netze, die sich für Mustererkennung oder Funktionsapproximation eignen. Das Anlernen des Netzes ist in jedem Falle überwacht, d. h. das Netz wird auf einen gegebenen Targetvektor oder -wert, dem Input \vec{p} entsprechenden angestrebten Output \vec{y} trainiert.

$$\vec{y} = f(\vec{p}) \quad (4)$$

Das Netz besteht aus einer Anzahl Schichten (engl. layers) von Neuronen. Alle Neuronen einer Schicht verarbeiten parallel zueinander die gleichen Informationen und liefern die Eingangsgrößen für die nächste Schicht, bzw., in der letzten Schicht, die Netzoutputs. Das Neuron, wie in Abbildung 7 dargestellt, ist die kleinste Berechnungseinheit des neuronalen Netzes. In einem Berechnungsschritt transformiert jedes Neuron eine Anzahl Inputwerte in einen einzigen Outputwert.

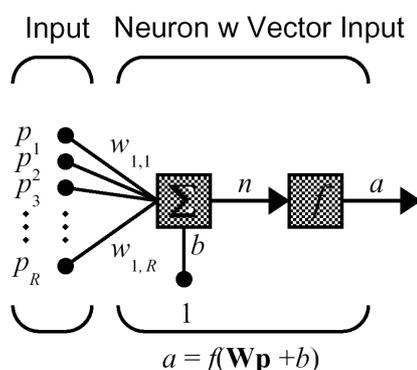


Abbildung 7: Neuron mit R Inputs (aus Hagan et al., 1996)

Wie sehr ein Input auf das Neuron wirkt (oder wie stark also das Neuron auf einen Reiz des benachbarten Neurons reagiert), wird durch einen Gewichtungsfaktor determiniert. Ein Neuron mit R Inputs p_i besitzt also R Gewichte w_i . Bzw. ein Neuron, das einen Inputvektor \vec{p} der Dimension R verarbeitet, besitzt einen Gewichtsvektor \vec{w} gleicher Dimension. Die erste Operation, die von dem Neuron durchgeführt wird, besteht aus der Addition der gewichteten Inputs, also der skalaren Multiplikation von Input- und Gewichtsvektor. Desweiteren besitzt das Neuron einen Biaswert b (quasi ein Gewicht für einen zusätzlichen Input = 1), der ebenso zu dieser Summe hinzugefügt wird (Gl. 1)

$$n = \vec{w} \cdot \vec{p} + b \quad (1)$$

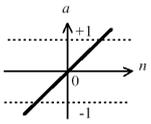
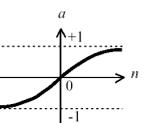
Die zweite Operation ist die Anwendung einer Transferfunktion f . Der skalare Wert n wird durch diese Funktion in den Outputwert a des Neurons überführt (Gl. 2)

$$a = f(n) \quad (2)$$

In dieser Arbeit verwendete Transferfunktionen sind in Tabelle 1 dargestellt. Zusammengefaßt für das Neuron ergibt sich aus Gl. 1 und Gl. 2:

$$a = f(\vec{w} \cdot \vec{p} + b) \quad (3)$$

Tabelle 1: Transferfunktionen

Bezeichnung	Input-Output-Beziehung	Funktion
Linear	$a = n$	
Tangens-Hyperbolicus-Sigmoid	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}} = \tanh(n)$	
Kompetitiv	$a = 1$ (Neuron mit größtem n) $a = 0$ (alle anderen Neuronen)	

Im Allgemeinen gilt: Befindet sich das Neuron in der ersten Schicht, sind die Inputs identisch mit den Netzinputs. In jeder weiteren Schicht sind die Inputs gleich den Outputs aller Neuronen der vorhergehenden Schicht. Grundsätzlich haben alle Neuronen einer Schicht die gleichen Inputs.

Die Dimension des Outputvektors bestimmt die Anzahl der Neuronen in der letzten Schicht, der Outputschicht. Dabei steht jedes Neuron für eine Größe. In dieser Schicht wird die lineare Transferfunktion verwendet. Die anderen Neuronenschichten des neuronalen Netzes bezeichnet man als verdeckte Schichten (hidden layers). Die verdeckten Schichten bedienen sich der Tangens-Hyperbolicus-sigmoiden Transferfunktion. Ein solches Netz mit nur einer verdeckten Schicht kann scheinbar jede beliebige Funktion in beliebiger Genauigkeit nachbilden, vorausgesetzt, es gibt ausreichend verdeckte Neuronen (Hornik et al, 1989).

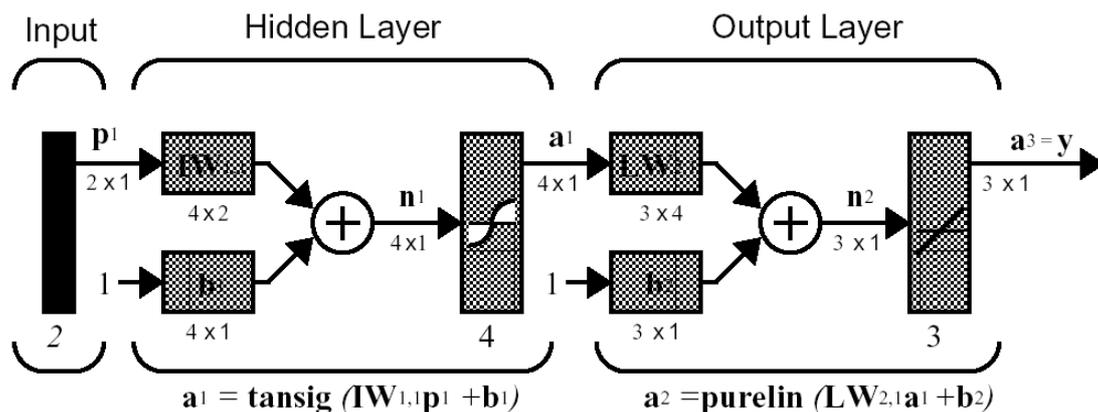


Abbildung 8: Multilayer-feedforward-Netz nach Hagan et al (1996)

Abbildung 8 zeigt ein MLFN mit einer verdeckten Schicht schematisch. Gegenüber Abbildung 7, in der ein einzelnes Neuron dargestellt ist, umfassen die horizontalen Klammern hier eine ganze Neuronenschicht. Wird das Neuron von einem R -dimensionalen Gewichtsvektor \vec{w} und einem Biaswert b charakterisiert, ist eine Neuronenschicht von S Neuronen durch die $R \times S$ -dimensionale Gewichtsmatrix W und den S -dimensionalen Biasvektor \vec{b} bestimmt. Der Output der Schicht ist dann ein S -dimensionaler Vektor \vec{a} .

Die Funktionsweise eines MLFN mit einer verdeckten Schicht mit der Gewichtsmatrix W_{HIDDEN} und dem Biasvektor \vec{b}_{HIDDEN} und der Outputschicht mit W_{OUTPUT} und \vec{b}_{OUTPUT} ist damit beschrieben durch:

$$\hat{y} = W_{OUTPUT} \cdot \tanh(W_{HIDDEN} \cdot \vec{p} + \vec{b}_{HIDDEN}) + \vec{b}_{OUTPUT} \quad (5)$$

Für das Training der Multilayer-Feedforward-Netze wird hier der Levenberg-Marquardt-Algorithmus verwendet. (Hagan et al, 1996)

Unter den verschiedenen Netzarchitekturen sind Multilayer-feedforward-Netze (oder Multilayer-Perceptron) eindeutig dominierend bei hydrologischen Anwendungen. Der Anteil dieser Netze in Publikationen beläuft sich auf ca. 90 % (Schütze, 2005). Speziell für die NA-Modellierung und Beschreibung des Wellenablaufes werden nahezu ausschließlich MLFN verwendet.

3.2.3 Self organizing feature maps

Selbstorganisierende Merkmalskarten zählen zu den kompetitiven Netzen. Sie besitzen nur eine, kompetitive Neuronenschicht, deren Neuronen bestimmte Prototypen von Inputvektoren verkörpern. Ein Neuron kann also als Vektor im Inputraum verstanden werden. Das Neuron, welches nun den Input am besten trifft (best matching unit, BMU), sozusagen den geringsten Abstand zum Eingangsvektor besitzt, wird angeregt und bringt den Output 1. Alle anderen Neuronen werden gehemmt und geben den Output 0 („The winner takes all“).

Unter den kompetitiven Netzen besitzen SOM die Eigenart, daß die Outputneuronen miteinander in Beziehung stehen. Ein Neuron besitzt „benachbarte“ Neuronen, welche im Inputraum dem Neuron am nächsten sind. Dazu sind die Neuronen der kompetitiven Schicht in einer (meist zweidimensionalen) Struktur oder Merkmalskarte angeordnet. Grundsätzlich gibt es zwei Strukturen, die die Beziehung der Neuronen untereinander definiert: die hexagonale mit jeweils sechs Nachbarn erster Ordnung und die viereckige (englisch: rectangular) mit jeweils 4 unmittelbaren Nachbarn, wie in Abbildung 9 dargestellt.

Dieser Netztyp ist direkt von neurophysiologischen Kenntnissen beeinflusst. Eine solche Grundstruktur von vernetzten Nervenzellen läßt sich in der Hirnrinde wiederfinden. Die dortigen biologischen neuronalen Netze sind typischerweise auch in zweidimensionalen Schichten angeordnet. Ein Reiz stimuliert dabei jedoch eine ganze Gruppe von Neuronen. Die Stärke des Reizes nimmt ab, je weiter sich das Neuron vom Reizzentrum entfernt befindet.

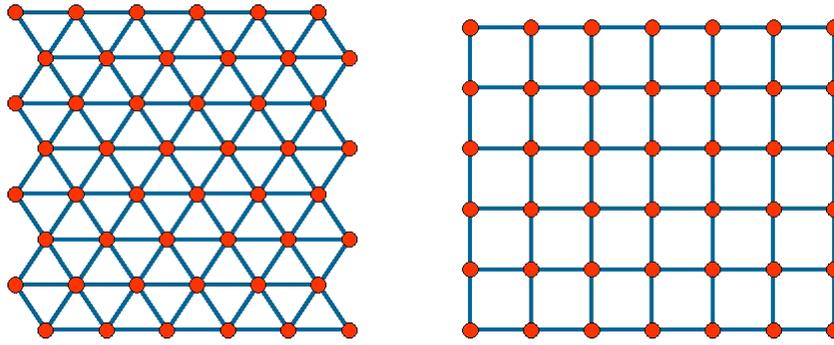


Abbildung 9: SOM, zweidimensionale Netztopologien: links hexagonal, rechts viereckig

Das Training eines solchen Netzes bedeutet de facto eine Organisation dieser Karten im Inputraum. Dabei werden die best matching unit (BMU) – das Gewinnerneuron i – repräsentiert durch den Vektor ${}_i\vec{w}$, und die benachbarten Neuronen dem jeweiligen Inputvektor \vec{p} um eine Lernrate α angenähert (Kohonen-Regel):

$${}_i\vec{w}(q) = {}_i\vec{w}(q-1) + \alpha(\vec{p}(q) - {}_i\vec{w}(q-1)) \quad (6)$$

$${}_i\vec{w}(q) = (1-\alpha) {}_i\vec{w}(q-1) + \alpha \vec{p}(q) \quad (7)$$

Die Lernrate ist α dabei für das Gewinnerneuron größer als für die Nachbarn. Die Karte sortiert sich während des Lernprozesses dergestalt, daß benachbarte Neuronen ähnliche Vektoren repräsentieren. Dadurch bildet die Merkmalskarte im Inputraum eine Fläche nach, die möglichst nah an den Trainingsvektoren liegt. SOM benötigen keinen Output zum Training, das Netztraining wird daher als unüberwacht bezeichnet. Eine SOM eignet sich hervorragend, Daten zu analysieren und Beziehungen zu visualisieren. Merkmalskarten sind anschaulich und im Gegensatz zu den Gewichtsmatrizen eines MLFN interpretierbar. Für die Funktionsapproximation eignen sich SOM indirekt. Fehlt eine beliebige Größe, die Bestandteil (oder Dimension) des Inputvektors ist, kann diese durch die Lage der Merkmalskarte im Raum bestimmt werden. Auch mit dieser fehlenden Dimension ist die Bestimmung der BMU – und damit der fehlenden Größe – möglich.

Damit beim Training der SOM – also der Verringerung der Distanz zwischen dem jeweiligen Inputvektor und dem nächsten Neuron – die einzelnen Merkmale des

Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

Inputvektors gleichbehandelt werden, werden die Eingangsgrößen über ihren Wertebereich normiert. Anderenfalls würde die Distanz des Merkmalsvektors von den einzelnen Neuronen von jenen Merkmalen dominiert, deren Wertebereich am größten ist.

Nach der Normierung sind diese Merkmale in bezug auf den Trainingsprozeß zunächst also gleichwertig. Jedoch ist die Anwendung des Netzes klarerweise auf die Bestimmung der Größe Durchfluß (beziehungsweise Wasserstand) am Zielpegel ausgerichtet, welche hier entsprechend der Besonderheit des Netztypes nicht explizit als Targetwert definiert werden kann. Um der herausgehobenen Stellung dieses Merkmales gerecht zu werden, soll ein Outputdimensionsgewicht (*ODW*) definiert werden, mit welchem die normierten Durchfluß- oder Wasserstandswerte am Zielpegel multipliziert werden. Der Wertebereich wird also in dieser Dimension von ursprünglich (durch die Normierung vorgegeben) zwischen -1 und 1 auf zwischen $-ODW$ und $+ODW$ gestreckt. Der Abstand eines Inputvektors von einem Neuron hängt damit in stärkerem Maße von dieser „Outputdimension“ ab.

Die Verwendung von SOM birgt den Nachteil in sich, daß nur die diskreten Werte als Lösungen möglich sind, die in den Lagevektoren der Neuronen gespeichert sind. Um diesen Schönheitsfehler zu umgehen, können beispielsweise die Neuronen um lokale lineare Abbildungen erweitert werden (SOM-LLM, nach Ritter et al, 1992). Jedem Neuron werden dabei lokale Felder (Voronoi-Gebiete) zugeordnet. Wird die Lage der BMU im Raum durch den Merkmalsvektor nicht genau getroffen, so wird das fehlende Merkmal mittels einer für jedes Neuron angelernten Funktion linear korrigiert. Einen anderen Weg beschreibt Schütze (2005). Die SOM wird um eine variable Input-Output-Funktion erweitert (SOM-MIO). Dabei werden anstelle der Bestimmung nur einer best matching unit die n nächsten Neuronen als Stützstellen im Merkmalsraum bestimmt und anhand der Abstände des Merkmalsvektors zu diesen Neuronen wird der komplette Outputvektor aus den n Stützstellen interpoliert.

Bei einer ausreichenden Anzahl von Neuronen – und damit möglicher Werte für die Ausgabegröße – ist dieser Effekt in Anbetracht der zu erwartenden Ungenauigkeiten durch die hohe Dimension des Merkmalsraumes wohl zu vernachlässigen und in den berechneten Ganglinien nicht festzustellen. Auf eine Anwendung dieser Erweiterungen zur Ausgabe des fehlenden Merkmals soll hier verzichtet werden.

3.3 Datengenerierung als Trainingsgrundlage

Um künstliche neuronale Netze als zuverlässige Werkzeuge zur Hochwasservorhersage zu erschließen, muß gewährleistet sein, daß sie im Trainingsprozeß mit der gesamten Variabilität möglicher Hochwasserereignisse bekannt gemacht wurden. Dazu reichen Beobachtungsdaten im allgemeinen nicht aus, denn Extremhochwässer treten naturgemäß höchst selten auf und es existieren demzufolge sehr wenige Datensätze zum Anlernen der Netze. Demgegenüber besitzen physikalisch begründete, prozeßbeschreibende Modelle durchaus die Fähigkeit, mit einiger Sicherheit gänzlich unbekannte Szenarien zu berechnen. Um die Fähigkeit zur Extrapolation, die durch prozeßorientierte Modellierung gewährleistet ist, auf die ANN zu übertragen, sollen deterministische Modelle zur Erstellung einer umfassenden Trainingsdatenbank mit allen denkbaren Szenarien benutzt werden.

Grundlage dieser Szenarienrechnung ist eine Analyse der hochwasserauslösenden Wettersituationen und die Generierung einer Basis an meteorologischen Inputszenarien. Unter Verwendung dieser meteorologischen Inputszenarien wird ein kalibriertes und validiertes Niederschlags-Abfluß-Modell (WaSim-ETH, siehe Cullmann, 2007) betrieben, welches die Randbedingungen für das hydrodynamische Modell errechnet. Da aufgrund fehlender Beobachtungsdaten für die lateralen Zuflüsse eine direkte Parametrisierung des Niederschlags-Abfluß-Modelles für das Zwischeneinzugsgebiet nicht möglich ist, werden hierfür Parametersätze aus benachbarten Teilgebieten übertragen. Das kalibrierte und validierte hydrodynamische Modell simuliert nun die Ganglinien am Zielquerschnitt.

Damit sind die Grundlagen für eine umfangreiche Trainingsdatenbank gelegt. Da es sich hierbei jedoch um kontinuierliche Prozesse handelt, bei denen ein konkreter Durchflußwert am Zielquerschnitt aus einer mehr oder weniger langen Vorgeschichte resultiert, ergibt sich das Problem einer Auswahl der für die zu simulierende Ganglinie relevanten Merkmale, aus denen sich der vorherzusagende Wert hinreichend genau berechnen läßt.

3.3.1 Die eindimensionale hydrodynamisch-numerische Modellierung

Modellanforderungen

Der Abfluß von Hochwasserwellen in den Fließstrecken mit geringem Sohlgefälle ist geprägt von Überflutungen der Vorländer und Rückstauerscheinungen an Zusammenflüssen der Hauptgewässer. Die dadurch hervorgerufenen Retentions- und

Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

Verzögerungseffekte sind nichtlinearer Natur und abhängig vom aktuellen Wasserstand, beziehungsweise Durchfluß. Befindet sich der zu prognostizierende Zielquerschnitt in einem Bereich, der von diesen Vorgängen direkt beeinflusst ist, treten beispielsweise nicht vernachlässigbare Hystereseeffekte in der Wasserstands-Durchfluß-Beziehung auf.

Um solchen Prozessen gerecht zu werden, ist eine instationäre hydrodynamische Modellierung unter Ausschöpfung einer detaillierten Geländeinformation erforderlich. Die Anwendung eines eindimensionalen hydrodynamisch-numerischen Modellsystems ist hierbei sinnvoll. Das zu verwendende Modell soll für die Erstellung der Modellgeometrie über eine Schnittstelle zu geographischen Informationssystemen verfügen, da üblicherweise die Vermessungen der Fließquerschnitte für die Abbildung der im Hochwasserfall überfluteten Vorländer nicht ausreichen und durch Informationen aus digitalen Geländemodellen erweitert werden müssen. Darüber hinaus ermöglicht die GIS-gestützte Erstellung der Modellgeometrie eine Abschätzung der Rauigkeitsbeiwerte der Vorländer unter Einbeziehung digitaler Daten der Landnutzung in den Flußbauen.

Desweiteren sollte die Modellierung von Zusammenflüssen mehrerer Fließstrecken und die Einbindung lateraler Zuflüsse im Modellsystem in adäquater Weise implementiert sein, um eine mehr oder weniger automatisierte Abarbeitung einer größeren Anzahl von Berechnungsszenarien zu ermöglichen.

Ein Programmpaket, daß diesen Anforderungen genügt und frei verfügbar ist, ist HEC-RAS des Hydrologic Engineering Center des US Army Corps of Engineers.

HEC-RAS

HEC-RAS ist ein Programmpaket des US Army Corps of Engineers zur eindimensionalen hydrodynamisch-numerischen Modellierung. Als Public-domain-Software ist es frei erhältlich. HEC-RAS besitzt eine Graphische Benutzeroberfläche (Abbildung 10), die die Datenverwaltung, Visualisierung und Auswertung unterstützt.

Die instationäre Strömungsberechnung basiert auf einer numerischen Lösung der Saint-Venant-Gleichungen.

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} - q_l = 0 \quad (7)$$

$$\frac{\partial Q}{\partial t} + \frac{\partial QV}{\partial x} + g A \left(\frac{\partial z}{\partial x} + S_f \right) = 0 \quad (8)$$

Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

Gleichungen (7) und (8) bezeichnet man auch als Kontinuitäts- und Impulsgleichung. Die Kontinuitätsgleichung steht für die Massenerhaltung am betrachteten Fließquerschnitt. Die Impulsgleichung leitet sich aus dem zweiten Newtonschen Axiom ab und erfaßt die wirkenden Druck-, Gravitations- und Reibungskräfte.

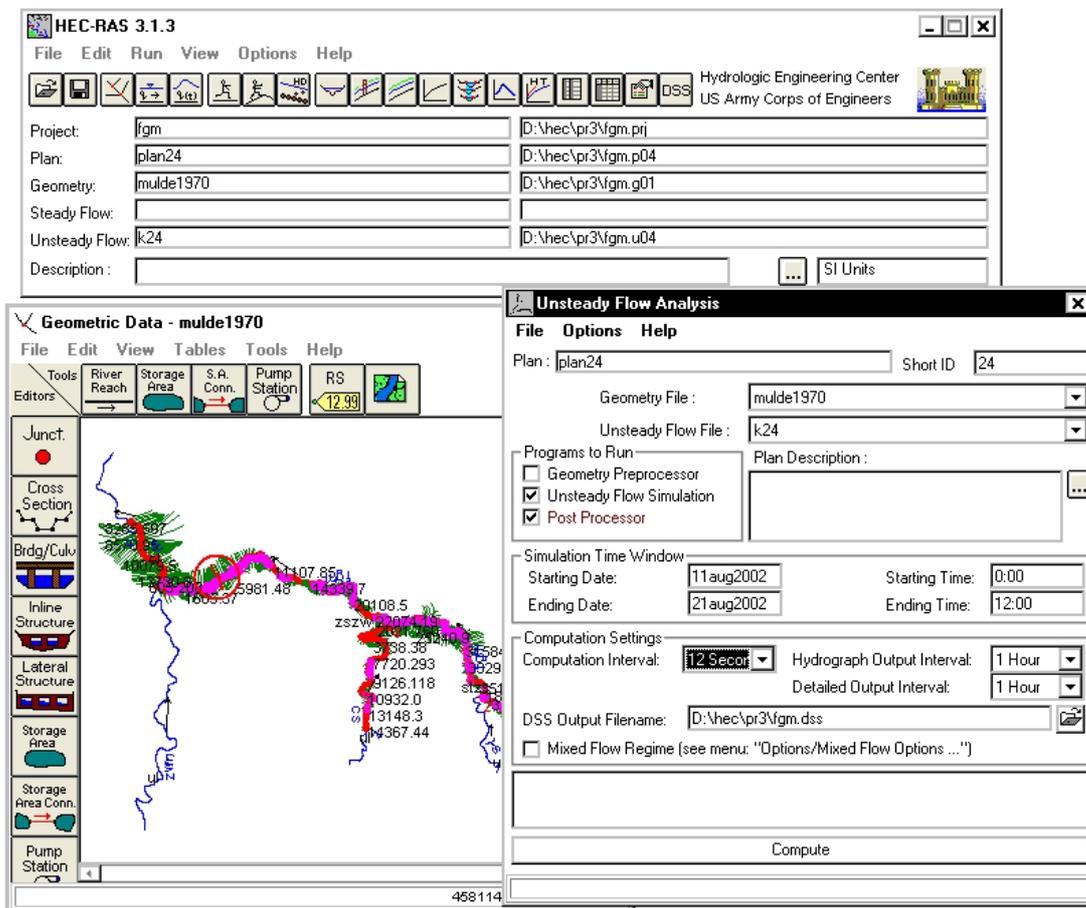


Abbildung 10: HEC-RAS, Erscheinungsbild

Die Datenhaltung in HEC-RAS ist für intensive Szenarienrechnungen oder die Kopplung mit anderen Modellen, beispielsweise einem Niederschlags-Abfluß-Modell, bedingt geeignet. Zuflußganglinien werden über Dialogfenster eingegeben und in ASCII-Dateien in einem speziellen Format gespeichert. Der automatisierte Zugriff auf diese Dateien ist prinzipiell möglich, jedoch müssen dafür externe Werkzeuge programmiert werden. Sämtliche Modellergebnisse werden in Binärdateien abgelegt. Ein externer Zugriff erscheint unmöglich – der Datenexport hat manuell über die Benutzeroberfläche zu erfolgen.

Für eine realitätsnahe Modellierung mit einem hydrodynamischen Modell wie HEC-RAS ist es unerlässlich, die Geometrie des Flußbettes und der Vorländer so exakt wie

Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

möglich im Modell zu berücksichtigen. Mit der ArcView-Extension HEC-GeoRAS (US-ACE, public domain software) ist eine hervorragende Schnittstelle zu GIS gegeben, die die Verwendung digitaler Geländedaten erleichtert.

Erstellung der Modellgeometrie mit HEC-GeoRAS

Die ArcView-Extension HEC-GeoRAS erleichtert die GIS-gestützte Erstellung der Flußgeometrie in ArcView. Die Information über die geodätische Höhe muß dazu in einem digitalen Höhenmodell als TIN (Triangulated Irregular Network) zusammengefaßt vorliegen. Mit HEC-GeoRAS ist die automatisierte Extraktion der manuell eingefügten Quer- und Längsschnitte, Fließwege, Uferlinien etc. möglich. Desweiteren können mit Hilfe von Landnutzungskarten den verschiedenen Vorlandbereichen entsprechende Rauigkeiten zugeordnet werden. HEC-GeoRAS unterstützt ebenso den Export der relevanten Geometriedaten im HEC-RAS-Format (USACE-I, 2002)

Randbedingungen und räumliche und zeitliche Datenstruktur

Die Wahl der Randbedingungen muß alle veränderlichen Größen berücksichtigen, die das gewünschte Modellergebnis beeinflussen. Das Modellsystem HEC-RAS fordert die Definition der oberen Randbedingung am Hauptgewässer und der Randbedingungen aller anderen oberstromigen Enden der Nebenflüsse als Wasserstands- oder / und Durchflußganglinien. Als untere Randbedingung kann außer der Bereitstellung von Ganglinien auch ein konstantes Reibungsgefälle oder eine Schlüsselkurve verwendet werden. Desweiteren kann an jedem beliebigen Querschnitt ein konstanter Zufluß, eine Zuflußganglinie oder ein grundwasserstandsabhängiger Zufluß definiert werden.

Zuflüsse aus dem Zwischeneinzugsgebiet entlang der modellierten Fließstrecken müssen bei einer adäquaten Modellierung berücksichtigt werden. Sind diese Zuflüsse durch Pegel repräsentiert, können diese Ganglinien am entsprechend nächsten Querprofil als lateraler Zufluß definiert werden.

Da in dieser Arbeit das hydrodynamische Modell mit einem NA-Modell gekoppelt wird, ist eher die Gliederung des Einzugsgebietes in Teileinzugsgebiete ausschlaggebend. Da das NA-Modell Ganglinien für die Teilgebietsauslässe ausgibt, werden diese Zuflüsse an den entsprechenden Querschnitten positioniert. Die oberen Ränder der modellierten Teilflüsse inklusive des oberen Modellrandes sollten mit der Teilgebietsgliederung des NA-Modells abgestimmt sein.

Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

Die Ganglinien für den Prognosepegel sollen hier simuliert werden und scheiden somit als untere Randbedingung aus. Schlüsselkurven sind speziell im Hochwasserfall sehr fehlerbehaftet und berücksichtigen keine Hystereseeffekte. Da darüber hinaus das Reibungsgefälle weder bekannt, noch im Verlaufe eines Hochwasserereignisses konstant ist, eignet sich der eigentliche Zielquerschnitt nicht für die Wahl der unteren Randbedingung. Vielmehr sollte die Lage der unteren Randbedingung genügend weit unterhalb des Zielquerschnittes gewählt werden, so daß der Hochwasserverlauf am Zielquerschnitt davon unbeeinflusst bleibt. Ob die gewählte Entfernung (beispielsweise eines konstanten Reibungsgefälles) ausreichend ist, bedarf einer Sensitivitätsanalyse.

Modellkalibrierung und -validierung

Ziel der Kalibrierung ist, eine gute Übereinstimmung der gemessenen und modellierten Ganglinien am Zielquerschnitt zu erreichen.

Die Kalibrierung des hydrodynamischen Modells erfolgt durch eine Anpassung der Rauigkeitsbeiwerte (Manningbeiwerte). Dabei kann die Anpassung für die verschiedenen Fließzonen (linkes und rechtes Vorland, Hauptkanal) separat durchgeführt werden. Auch kann eine noch genauere Gliederung unternommen werden. Die Startwerte ergeben sich bei der Erstellung der Geometriedaten mit Hilfe von HEC-GeoRAS aus der digitalen Landnutzungskarte durch Zuordnung von Manningbeiwerten.

Die Kalibrierung des Modellsystems – insbesondere anhand gemessener Hochwasserereignisse – ist jedoch von besonderen Schwierigkeiten geprägt, die sich aus der Ungenauigkeit der gemessenen Ganglinien ergeben. Bei einer Verwendung der Wasserstandsganglinien zur Kalibrierung des Modells sind Fehler unausweichlich, da für eine Veränderung der Geometrie am Meßquerschnitt im Laufe der Zeit meist keine Daten vorliegen. Währenddessen eine WQ-Beziehung meist nach jedem größeren Hochwasser aktualisiert wird und die Durchflußganglinien damit auf die neuen Gegebenheiten angepaßt werden, ist eine Neuvermessung der Gerinnegeometrie nur äußerst selten.

Bei der Verwendung von Durchflußganglinien wiederum ist nicht unbedingt vorauszusetzen, daß die Massenbilanz der Hochwasserwellen eindeutig bestimmbar ist, da Extremhochwässer die empirisch bestimmten Wasserstands-Durchfluß-Beziehungen in jedem Falle im Extrapolationsbereich treffen.

Da hier für diese Anwendung die meisten Daten nur als Durchflußganglinien vorlagen, wird die Kalibrierung und die Validierung mit diesen durchgeführt. Bei der

Szenarienrechnung sind die Zuflußganglinien auch in Form von Durchflüssen als Resultate der NA-Modellierung gegeben.

Wegen der beschriebenen Unzulänglichkeiten der Eingangsdaten ist eine Abweichung der modellierten von der gemessenen Ganglinie zu erwarten, die durch eine sinnvolle Anpassung der Rauigkeitsbeiwerte nicht zu eliminieren sein wird. Erfahrungen für das gewählte Pilotgebiet (LfUG, 2006) stützen diese Vermutung.

Eine gute Übereinstimmung der Scheitellaufzeit und eine optisch gute Anpassung der Form der Ganglinie für die wenigen mit Meßdaten belegten höchsten gemessenen Hochwässer sollen bei sorgfältiger Erarbeitung der Modellgeometrie als ausreichend betrachtet werden.

3.3.2 Die Niederschlags-Abfluß-Modellierung

WaSim-ETH

Die Modellierung der Niederschlags-Abfluß-Prozesse erfolgte mit dem Wasserhaushaltsmodell WaSim-ETH (Schulla und Jasper, 1999). Das Modell ist modular aufgebaut. In der verwendeten Form wurden folgende Module benutzt:

- Potentielle Verdunstung (Penman-Monteith)
- Reale Verdunstung (saugspannungsabhängige Reduktion der potentiellen Verdunstung)
- Schneemodul (Tagesgradverfahren)
- Interzeptionsmodell (Überlaufspeicher)
- Infiltrationsmodell (nach Green and Ampt, 1911)
- Bodenmodell (1D-Richardsgleichung)
- Abflußkonzentration (Flächen-Laufzeit-Verfahren)
- Abflußrouting (Translations-Diffusionsansatz)

Der Kern des Modells ist das Bodenmodul für die Berechnung des vertikalen Wassertransportes im Boden. Dafür stehen zwei verschiedene Ansätze zur Verfügung: der Topmodel-Ansatz und – wie bereits erwähnt – der hier verwendete Richards-Ansatz. Dabei wird für die Modellierung der Wasserflüsse in der ungesättigten Bodenzone die Richards-Gleichung (Gleichung 9) in einer für jede Gridzelle schichtweise diskretisierten Bodensäule in eindimensional vertikaler Form angewendet.

$$\frac{\Delta \Theta}{\Delta t} = \frac{\Delta Q}{\Delta z} = q_{in} - q_{out} \quad (9)$$

Eine Übersicht der Modellstruktur nach Schulla und Jasper (1999) ist im Anhang beigefügt.

In dieser Arbeit wurde eine Parametrisierung des Modells nach Cullmann (2007) für das Pilotgebiet verwendet bzw. auf nichtparametrisierte Teilgebiete übertragen.

Räumliche und zeitliche Datenstruktur und Datenfluß des NA-Modells

WaSim-ETH arbeitet auf Rasterbasis. Die Modellinputs unterteilen sich in geographische Daten, die in binärer Form als Raster bereitzustellen sind und Zeitreihen der meteorologischen Eingangsdaten, die lagebezogen gegeben sein müssen und dann durch interne Interpolationsalgorithmen auf das Raster übertragen werden. Zu ersteren zählen digitale Karten der Landnutzung und Bodenart und Karten topologischer Natur, die mit dem Preprocessing-Tool Tanalys (Schulla und Jasper, 1999) aus einem digitalen Höhenmodell erstellt werden.

Die meteorologischen Eingangsdaten umfassen Niederschlag, Globalstrahlung, relative Luftfeuchte, Lufttemperatur und Windgeschwindigkeit, wobei auch alternative Feuchtemaße und anstelle der Globalstrahlung auch die Sonnenscheindauer genutzt werden können. Diese Daten müssen in ihrer zeitlichen Auflösung mit dem gewählten Rechenzeitschritt übereinstimmen. Dieser kann vom Anwender definiert werden; für die Anwendung des Programmes zur Modellierung von Hochwasserwellen wird sinnvollerweise ein Rechenzeitschritt von einer Stunde angesetzt, der der üblichen zeitlichen Auflösung der verwendeten meteorologischen Eingangsdaten entspricht. Eine Berechnung der Vorgeschichte der Ereignisse zur Determinierung des Gebietszustandes bei Ereignisbeginn in einer geringeren zeitlichen Auflösung ist nicht möglich.

Der Datenfluß soll an dieser Stelle nur soweit beschrieben werden, wie es für das Verständnis dieser Arbeit vonnöten ist. Der Niederschlag geht zunächst gemeinsam mit der Temperatur für jeden Zeitschritt in das Schneemodell ein. Der Schneespeicher füllt sich oder entleert sich entsprechend der Temperatur und Niederschlagsmenge. Der Output des Schneemodells stellt den Effektivniederschlag für die Weiterverarbeitung in den Modulen für Interzeption, Infiltration und Boden. Gleichzeitig berechnet sich aus den meteorologischen Inputdaten Globalstrahlung, relative Luftfeuchte, Temperatur und Windgeschwindigkeit rasterweise die potentielle

Verdunstung. Das Bodenmodell berechnet den vertikalen Wassertransport mit dem Input Niederschlag und den Verlustgrößen Abfluß (in den Komponenten langsamer und schneller Basisabfluß, Interflow und Direktabfluß) und der realen Verdunstung, welche sich wiederum aus dem aktuellen Feuchtezustand und der potentiellen Verdunstung ergibt. Abschließend wird der Abfluß aus der Rasterzelle über eine aus dem digitalen Höhenmodell im Preprocessing bestimmte Flächenlaufzeitbeziehung zum Teilgebietsauslaß geroutet. Mit Hilfe eines Translations-Diffusionsansatzes wird der Wellenablauf über die Fließgewässer durch die unterstromigen Teileinzugsgebiete realisiert.

Die Daten werden für die einzelnen Wasserhaushaltskomponenten, wie Effektivniederschlag nach dem Schneemodell, Inhalt des Schneespeichers, potentielle und reale Verdunstung oder die verschiedenen Abflußkomponenten teileinzugsgebietsweise als Zeitreihen ausgegeben. Der Gesamtabfluß wird für das jeweilige Teileinzugsgebiet separat und als Summe mit den gerouteten Abflüssen der oberstromigen Teileinzugsgebiete dargebracht.

Die das Gebiet repräsentierenden Zustandsgrößen wie Füllung des Schneespeichers oder Feuchte der Bodenschichten werden als Raster für den aktuellen Zustand gespeichert.

3.4 Szenarien der Einzugsgebietsgliederung

Die Berücksichtigung des Zwischeneinzugsgebietes in Form lateraler Zuflüsse bei der Erstellung künstlicher neuronaler Netze soll in zwei verschiedenen Varianten erfolgen. Einerseits wird unterstellt, daß der Einfluß des Zwischengebietes gering ist und die Information der Ganglinien an den Oberliegerpegeln zur Abbildung des Zielpegels ausreicht. Diese Variante wird mit HD-ANN bezeichnet, da hier aus den oberstromigen Ganglinien ausschließlich Merkmale für die hydrodynamischen Prozesse abgeleitet werden. Andererseits sollen diesen HD-ANN künstliche neuronale Netze gegenübergestellt werden, die die unbeobachteten lateralen Zuflüsse aus dem Zwischeneinzugsgebiet in Form von Merkmalen für die Niederschlags-Abflußprozesse in diesem Teilgebiet berücksichtigen. Diese sollen hier als HDNA-ANN bezeichnet werden.

Das Resultat einer Gegenüberstellung dieser Ansätze ist mit einiger Sicherheit abhängig von der Größe des Zwischengebietes. Daher werden drei Fälle hinsichtlich der Größe des Zwischeneinzugsgebietes betrachtet. Entsprechend der Bezeichnung

der neuronalen Netze mit und ohne Verwendung von Merkmalen für die NA-Prozesse im Zwischeneinzugsgebiet und der Größe der Zwischengebiete werden im folgenden die in Tabelle 2 aufgeführten Bezeichnungen verwendet.

Tabelle 2: Szenarienbezeichnung

Anteil des Zwischeneinzugsgebietes am Gesamteinzugsgebiet	Nur Merkmale für hydrodynamische Prozesse im Flußlauf	Merkmale für HD-Prozesse und NA-Prozesse im Zwischeneinzugsgebiet
10%	HD10	HDNA10
20%	HD20	HDNA20
40%	HD40	HDNA40

3.5 Merkmalsselektion für die hydrodynamischen Prozesse

Bei der Definition der Merkmale wird zwischen der Festlegung der Ereignismerkmale und Zustandsmerkmale unterschieden. Der Zustand des Systems ist hierbei ausschließlich durch die Geometrie des Gewässers charakterisiert. Unter der Voraussetzung, daß die Geometrie inklusive der Rauigkeit des Fließgewässers und der Vorländer bei der hydrodynamisch-numerischen Modellierung als konstant angenommen wird, kann auf die Verwendung von Zustandsmerkmalen für die reinen Wellenablaufprozesse verzichtet werden.

Oberstromig sind die zu modellierenden Flussabschnitte von einem oder mehreren Oberliegerpegeln, für welche ein kalibriertes und validiertes Niederschlags-Abfluß-Modell existiert, begrenzt. Die Hochwasserwelle ist durch Zuflußganglinien an den Oberliegerpegeln charakterisiert. Die entsprechenden Ereignismerkmale können diesen Ganglinien entnommen werden. Dabei wird davon ausgegangen, daß für einen konkreten Wasserstands- beziehungsweise Durchflußwert am Zielquerschnitt (Targetwert) jedem Oberliegerpegel eine Einflußspanne zugeordnet werden kann. Die Abschnitte der Zuflußganglinien, die sich in den Einflußspannen befinden, bestimmen diesen Targetwert. Der Durchflußverlauf außerhalb dieser Einflußspannen wirkt sich nicht auf den Targetwert aus.

Betrachtet man jeden einzelnen Oberliegerpegel separat, so kann für diese je eine Einflußspanne von $t_{inf,max}$ bis $t_{inf,min}$ definiert werden. Für die Berechnung einer Ganglinie am Zielpegel finden sich für jeden einzelnen Targetwert die Merkmale in den korrespondierenden Zeitfenstern für die jeweiligen Zuflüsse mit dem zeitlichen Abstand

von $t_{inf,max}$ bis $t_{inf,min}$. Abbildung 11 veranschaulicht dieses Prinzip, wobei die mit roten Quadraten gekennzeichneten diskreten Werte der Inputganglinie die für den mit dem roten Dreieck gekennzeichneten Outputwert charakteristischen Ereignismerkmale darstellen.

→ $t_{inf,min}$: minimale Einflußzeit

→ $t_{inf,max}$: maximale Einflußzeit

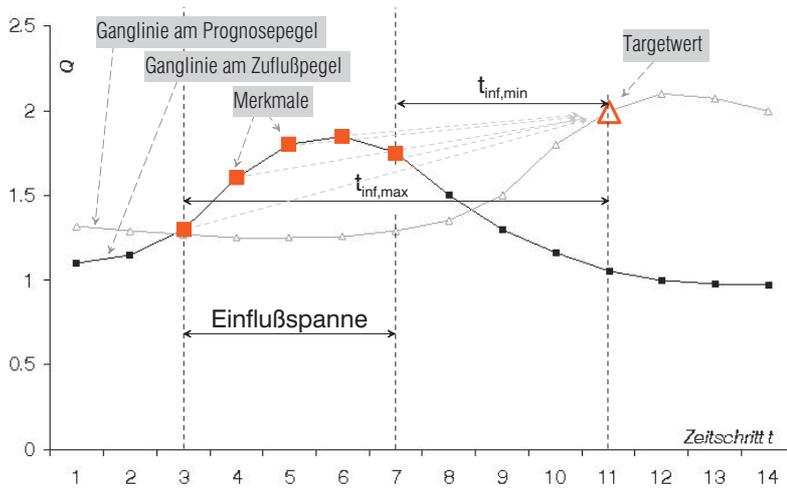


Abbildung 11: Einflußbereich der Durchflußganglinie am Oberliegerpegel auf einen Wert der Ganglinie am Zielpegel

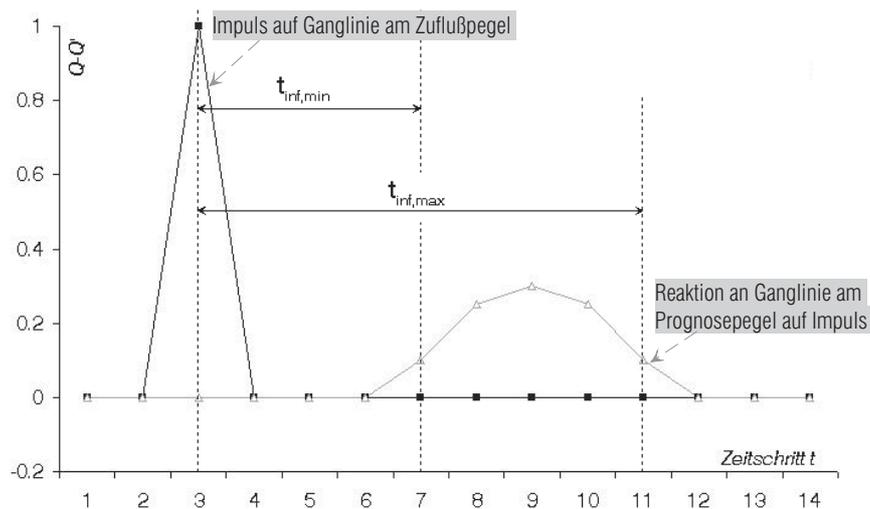


Abbildung 12: Bestimmung der Parameter $t_{inf,min}$ und $t_{inf,max}$ der Einflußspanne

Beginn und Ende dieser Einflußspannen werden invers und für jede der Zuflußganglinien einzeln bestimmt. Dazu wird das hydrodynamisch-numerische Modell verwendet. Die Zuflußganglinie wird dabei mit einem negativen oder positiven Impuls beaufschlagt, der einen Zeitschritt andauert und nacheinander an verschiedenen Stellen einer Ganglinie plaziert wird: am ansteigenden Ast, am Scheitel und am abfallenden Ast. Dies wird für verschiedene Hochwasserwellen durchgeführt. Eine signifikante Differenz der Ganglinien am Zielquerschnitt ohne und mit Impuls weist die Einflußspanne dieses einen Inputwertes auf die Ganglinie am Zielquerschnitt aus (Abbildung 12).

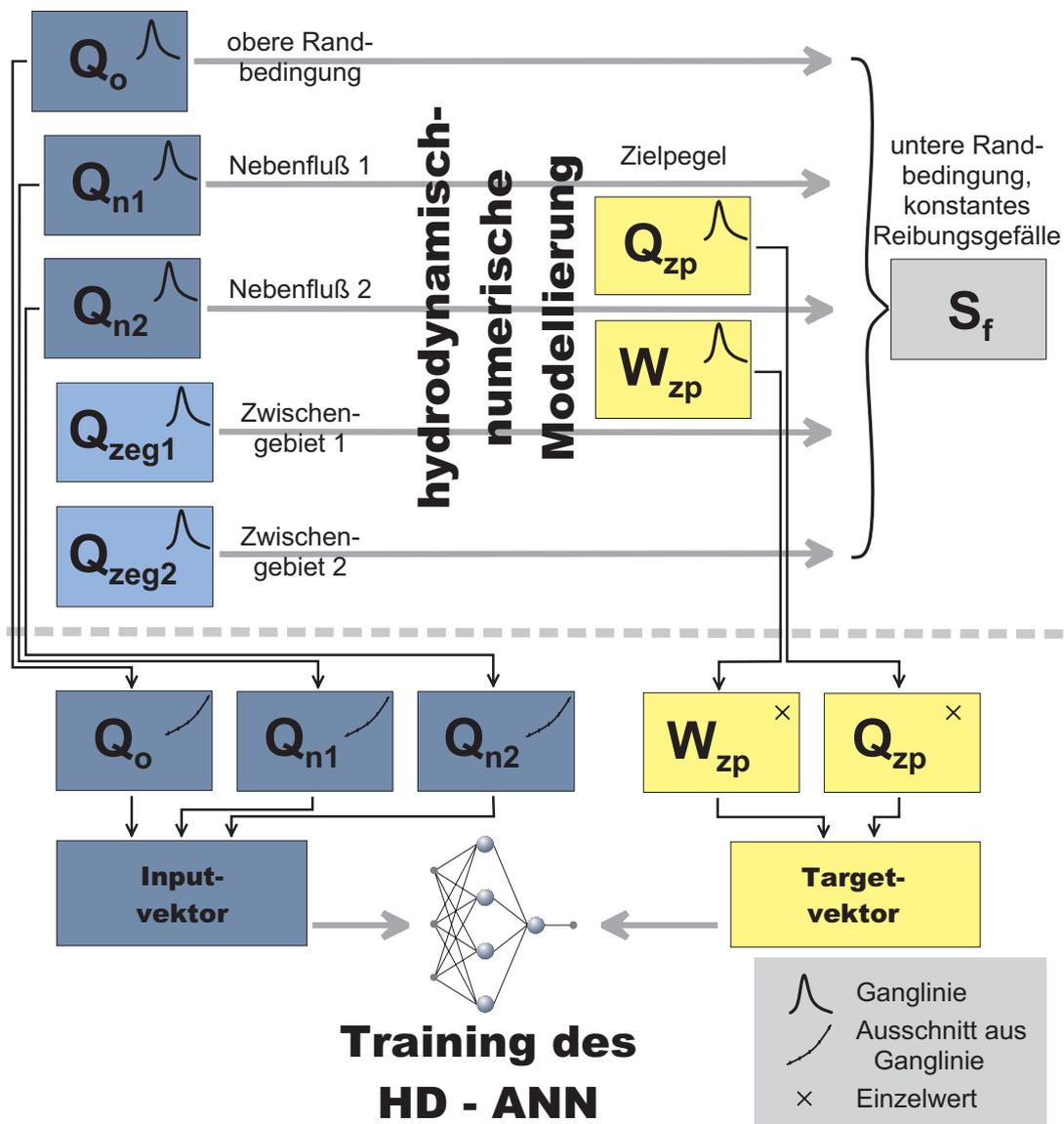


Abbildung 13: Datenfluß bei der Erstellung des Trainingsdatensatzes für ein neuronales Netz

Abbildung 13 zeigt zusammengefaßt den Datenfluß für die Erstellung einer Datenbasis für das Training eines künstlichen neuronalen Netzes. Auf der Ebene des hydrodynamisch-numerischen Modells liegen die Inputdaten ausschließlich als Ganglinien vor. Ebenso sind die Ergebnisse für den Zielpegel kontinuierliche Zeitreihen. Bei der Umwandlung der Ganglinien in Merkmale für das neuronale Netz wird jedem einzelnen aus der modellierten Ganglinie am Zielpegel herausgelösten Targetwert eine Reihe von Inputwerten aus der Durchflußganglinie der oberen Randbedingung Q_o und den (hier zwei) Ganglinien Q_{n1} und Q_{n2} , welche die Nebenflüsse repräsentieren, entsprechend den Einflußspannen zugeordnet.

Die lateralen Zuflüsse (in Abbildung 13, Q_{ZEG1} und Q_{ZEG2}) aus dem Zwischeneinzugsgebiet bleiben bei den Szenarien HD10, HD20 und HD40 im Merkmalsatz unberücksichtigt. Bei den Szenarien HDNA10, HDNA20 und HDNA40 werden sie durch Merkmale für die zugrundeliegenden meteorologischen Größen repräsentiert (siehe dazu 3.6).

Diese Methodik ist unabhängig vom gewählten Modellsystem, da die Merkmale direkt aus den Inputganglinien für die hydrodynamische Modellierung abgeleitet werden. Die Anzahl der Merkmale hängt einzig vom Grad der Wellenverformung – und damit von Faktoren wie Länge der Fließstrecke, Gefälle und Beeinflussung durch Retentions- und Rückstaueffekte – ab. Mit einer stärkeren Transformation der Welle erhöht sich die Länge der Einflußspannen und damit die Dimension des Merkmalsvektors.

3.6 Merkmalsselektion für die NA-Prozesse im Zwischengebiet

Die in Abbildung 13 als Q_{ZEG1} und Q_{ZEG2} bezeichneten lateralen Zuflüsse aus dem Zwischeneinzugsgebiet wurden als Zuflußganglinien für die hydrodynamisch-numerische Modellierung mit dem Niederschlags-Abfluß-Modell WaSim-ETH berechnet. Im operativen Fall bei der Anwendung des neuronalen Prognosewerkzeuges liegen diese Ganglinien jedoch nicht vor, da diese Zuflüsse nicht durch Pegelbeobachtungen oder Vorhersagen repräsentiert werden. Damit ist eine Definition der Merkmale entsprechend der in 3.5 vorgestellten Methodik nicht möglich. Um diese Zuflüsse als Folge von Niederschlags-Abfluß-Prozessen im Zwischeneinzugsgebiet für die Szenarien HDNA10, HDNA20 und HDNA40 (Tabelle 2) zu berücksichtigen, sollen zusätzliche Merkmale aus den zugrundeliegenden

meteorologischen Daten extrahiert werden. Die Methodik zur Bestimmung dieser Merkmale soll auch hierbei Allgemeingültigkeit besitzen, daß heißt, unabhängig vom gewählten NA-Modell und Einzugsgebiet einsetzbar sein. Die Verwendung von Merkmalen nach der bei Cullmann (2007) entwickelten Methodik ist nicht möglich, da dort Netze für verschiedene Vorhersageweiten mit unterschiedlichen Merkmalen erstellt werden.

Die Identifizierung dieser Merkmalsgruppe erfolgt zunächst testweise für ein reines Niederschlags-Abfluß-ANN, da hierbei die Merkmale aus den diesbezüglich relevanten meteorologischen Eingangsgrößen, wie dem hochwasserauslösenden Niederschlag, am besten identifiziert werden können. Dabei soll sich auf die Kenntnis der ablaufenden und im deterministischen Modell abgebildeten Prozesse gestützt werden. Für die Erstellung des NA-ANN soll eine vereinfachende Annahme getroffen werden. Der Niederschlag wird im Zwischengebiet als räumlich gleichverteilt angenommen, es werden keine flächenmäßig differenzierten Merkmale definiert. Diese Annahme ist dadurch gerechtfertigt, daß es sich hier um die Charakterisierung des Beitrages eines im Vergleich zum gesamten Einzugsgebiet kleinen Teilgebietes handelt. Diese hier im folgenden beschriebene Methodik kann leicht zur Erstellung eines eigenständigen NA-ANN durch die Einführung einer räumlichen Differenzierung erweitert werden.

In einem zweiten Schritt fließen die Zustands- und Ereignismerkmale für die Beschreibung der Abflußbildung und -konzentration in den Szenarien HDNA10, HDNA20 und HDNA40 mit den Ereignismerkmalen für die Wellenablaufprozesse zusammen zu einem Inputvektor.

Die Abflußbildung wird grundsätzlich und unabhängig vom verwendeten Modell von zwei Faktoren charakterisiert. Zum einen ist das der unmittelbare hochwasserauslösende Input an Niederschlags- bzw. Schmelzwasser. Daraus werden die Ereignismerkmale, wie in 3.6.1 detailliert beschrieben, abgeleitet. Da im Verlaufe eines Extremhochwassers Wasserhaushaltsprozesse wie die Verdunstung vernachlässigbar sind, werden hierfür nur Niederschlagswerte berücksichtigt.

Zum anderen ist für die Ausprägung der Hochwasserwelle der Gebietszustand bezüglich der Vorfeuchte oder Speicherfüllung bei Ereignisbeginn wesentlich. Dieser wird durch die Wettergeschichte von bis zu mehreren Monaten durch Verdunstung und Niederschlag und die Gebietseigenschaften beeinflusst. Wenn beispielsweise ein Niederschlagsereignis in einem trockenen Sommer eintritt, ist die zu erwartende Hochwasserwelle bedeutend geringer, als wenn der gleiche Regen im März nach gerade abgeschlossener Schneeschmelze fällt. Wie weit die dahingehend relevante

Vorgeschichte zurückreicht, ist Gegenstand einer Sensitivitätsanalyse. Die Größe, die dies im Merkmalsvektor widerspiegelt, ist hier das Zustandsmerkmal. Dieses wird hier als Scheiteldurchfluß einer aus einem vordefinierten Standardniederschlag resultierenden Hochwasserwelle festgelegt. Das Zustandsmerkmal ist damit ausschließlich von den Speicherfüllungen bei Ereignisbeginn abhängig und somit eine Funktion der Wettergeschichte. Ferner mißt es die Vorgeschichte explizit an ihrer Wirkung auf ein stattfindendes Hochwasserereignis. Für die Bestimmung dieses Zustandsmerkmals aus den Wetterdaten im Anwendungsfall werden zwei Subnetze trainiert. Boden- und Landnutzungsparameter und andere bei der deterministischen Modellierung als konstant betrachtete Größen sind davon analog zur Geometrie und Rauigkeit des Fließgewässers ausgenommen. Eine Beschreibung des Zustandsmerkmals und der zur seiner Berechnung entwickelten Subnetze findet sich unter 3.6.2.

In 3.6.3 soll die Anwendung der Methodik praktisch anhand der Kopplung mit dem Niederschlags-Abfluß-Modell und die notwendige Organisation der Daten dargestellt werden.

3.6.1 Ereignismerkmale

Um die Dynamik des Hochwassers dem Modell entsprechend gut nachzubilden, wird zunächst die zeitliche Auflösung der Eingangsdaten übernommen. Es werden also alle Stundenwerte des Niederschlags ab Ereignisbeginn bis zum Vorhersagezeitpunkt im Merkmalsvektor berücksichtigt. Da hier ein Netz alle Vorhersageweiten ab Ereignisbeginn abdecken soll, muß eine entsprechende systematische Aggregation der Daten erfolgen.

Dabei kann entsprechend der Methodik für die hydrodynamischen Prozesse (Kapitel 3.5) die Idee einer Einflußspanne aufgegriffen werden. Da jedoch jeder einzelne Niederschlagswert zwischen Ereignisbeginn und Vorhersagezeitpunkt mehr oder weniger relevant ist, soll hier eine zentrale Einflußspanne definiert werden. Innerhalb dieser Einflußspanne wird jeder einzelne Wert in den Merkmalsvektor aufgenommen, außerhalb davon werden die Werte mit zunehmendem Abstand zu dieser zentralen Einflußspanne immer stärker zusammengefaßt. Abbildung 14 verdeutlicht dieses Prinzip.

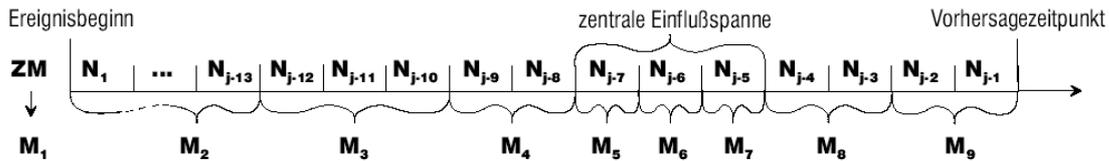


Abbildung 14: Prinzip der Zusammenfassung der Niederschläge zu Ereignismerkmalen (ZM = Zustandsmerkmal nach 3.6.2)

Um die zentrale Einflußspanne zu finden, soll hier eine andere Vorgehensweise gewählt werden als bezüglich der Merkmale für den Wellenablauf. Das wird mit der Tatsache begründet, daß sich der Einfluß eines einzelnen Niederschlagswertes auf den Abfluß am Pegel weniger gut aus dem Zusammenhang des Niederschlagsverlaufes und der Gesamtsumme des gefallenen Niederschlags lösen läßt. Während die hydrodynamischen Prozesse ausschließlich zu einer gut beobachtbaren Verformung des Impulses führen, hängt die Reaktion des Einzugsgebietes auf den Niederschlagsimpuls stark von der Gebietsfeuchte ab, die sich im Laufe des Ereignisses kontinuierlich ändert. Stellt man einen einzelnen diskreten Niederschlagswert in den Zusammenhang eines realistischen, mehrere Stunden andauernden Ereignisses, wirkt dieser nicht nur direkt auf die Abflußbildung, sondern beeinflusst mit seinem Beitrag zu einer Veränderung der Gebietsfeuchte ebenso die Reaktion des Gebietes auf die darauffolgenden Niederschlagswerte.

Aus diesem Grunde soll die Korrelation der Niederschlagswerte zum Zeitpunkt t_i mit den verschiedenen darauffolgenden Abflußwerten zum Zeitpunkt t_j mit $t_0 < t_i < t_j$ (Ereignisbeginn zum Zeitpunkt t_0) untersucht werden. Damit läßt sich der Vielzahl der möglichen Varianten an Gebietsfeuchte, Niederschlagsdauer und -intensität besser gerecht werden.

Außerhalb der zentralen Einflußspanne werden mehrere Niederschlagswerte zu Summen zusammengefügt (Gleichung 10).

$$M = \sum_{t_i} N(t_i) \quad (10)$$

Ab einem bestimmten Abstand $t_j - t_i$ zum Vorhersagezeitpunkt t_j werden alle Werte bis hin zum Ereignisbeginn zusammengefaßt (Merkmal M_2 in Abbildung 14). Dieses Merkmal kann auch als erweitertes Vorfeuchtemerkmal interpretiert werden, da damit

bei großen Vorhersageweiten gewissermaßen der Gebietszustand um den länger zurückliegenden, nicht mehr direkt am Abfluß zum Vorhersagezeitpunkt beteiligten Niederschlag aktualisiert wird. Der konkrete zeitliche Bezug der in diesem Merkmal zusammengefaßten Niederschläge ist kaum noch gegeben.

Bei kürzeren Vorhersagezeiträumen können einzelne Merkmale Niederschlagswerte vor dem definierten Zeitpunkt des Ereignisbeginns t_0 beschreiben. Diese werden null gesetzt, da sämtliche meteorologische Größen vor t_0 im Zustandsmerkmal zusammengefaßt sind. Abbildung 14 zeigt das Prinzip für die Erstellung des Merkmalsvektors mit einer sehr kurzen zentralen Einflußspanne.

3.6.2 Das Zustandsmerkmal

Es wird unterstellt, daß sich die Speicherfüllungen in den verschiedenen Bodenschichten und in ihrer räumlichen Verteilung im Einzugsgebiet in einem Merkmal zusammenfassen lassen, welches ausschließlich Aussagekraft über die Wirkung des Gebietszustandes auf eine entstehende Hochwasserwelle besitzt. Um diese abzuschätzen, werden Modellrechnungen durchgeführt, welche auf mehrere reale Gebietszustände ein und das selbe künstliche Standardereignis treffen lassen. Dabei wird ein einheitlicher 24h-Regen mit einer Intensität von 4 Millimetern pro Stunde auf das Einzugsgebiet gegeben. Das Zustandsmerkmal ist die Scheitelhöhe der Hochwasserwelle am Gebietsauslaß. Damit mißt das Zustandsmerkmal den Gebietszustand nicht anhand interner Parameter, sondern direkt an ihrer potentiellen Wirkung. Diese scharfe Trennung der Beschreibung der Vorgeschichte und des eigentlichen Hochwasserereignisses ist hydrologisch sinnvoll, da damit das Problem der unterschiedlichen zeitlichen Skalen der Wasserhaushaltsprozesse und der wesentlich dynamischeren Vorgänge während des Hochwassers umgangen werden kann. Die Aktualisierung der Gebietsfeuchte während des Ereignisses kann auf die Beschreibung der langsamen Verdunstungsprozesse verzichten und ausschließlich auf Niederschlagswerten während des Ereignisses beruhen (siehe 3.6.1). Das Zustandsmerkmal ergibt sich nicht aus prognostizierten Daten, womit sich die datenbezogenen Unsicherheiten gering halten. Ensembleprognosen für ein Hochwasserereignis können damit – im Gegensatz zu den dabei variierten Niederschlagszeitreihen für die Ereignismerkmale – auf ein konstant bleibendes Zustandsmerkmal zurückgreifen, welches die gesicherte Information der Vorgeschichte speichert.

Die Vorteile dieser Vorgehensweise für die Erstellung des Prognosenetzes sind:

- ein einziges Zustandsmerkmal hält die Größe des Inputvektors des Prognosenetzes gering und vereinfacht damit das Training
- die Methodik ist unabhängig von der Wahl des verwendeten NA-Modells bzw. der benutzten Module
- das Zustandsmerkmal ist direkt relevant für Hochwasserausprägung

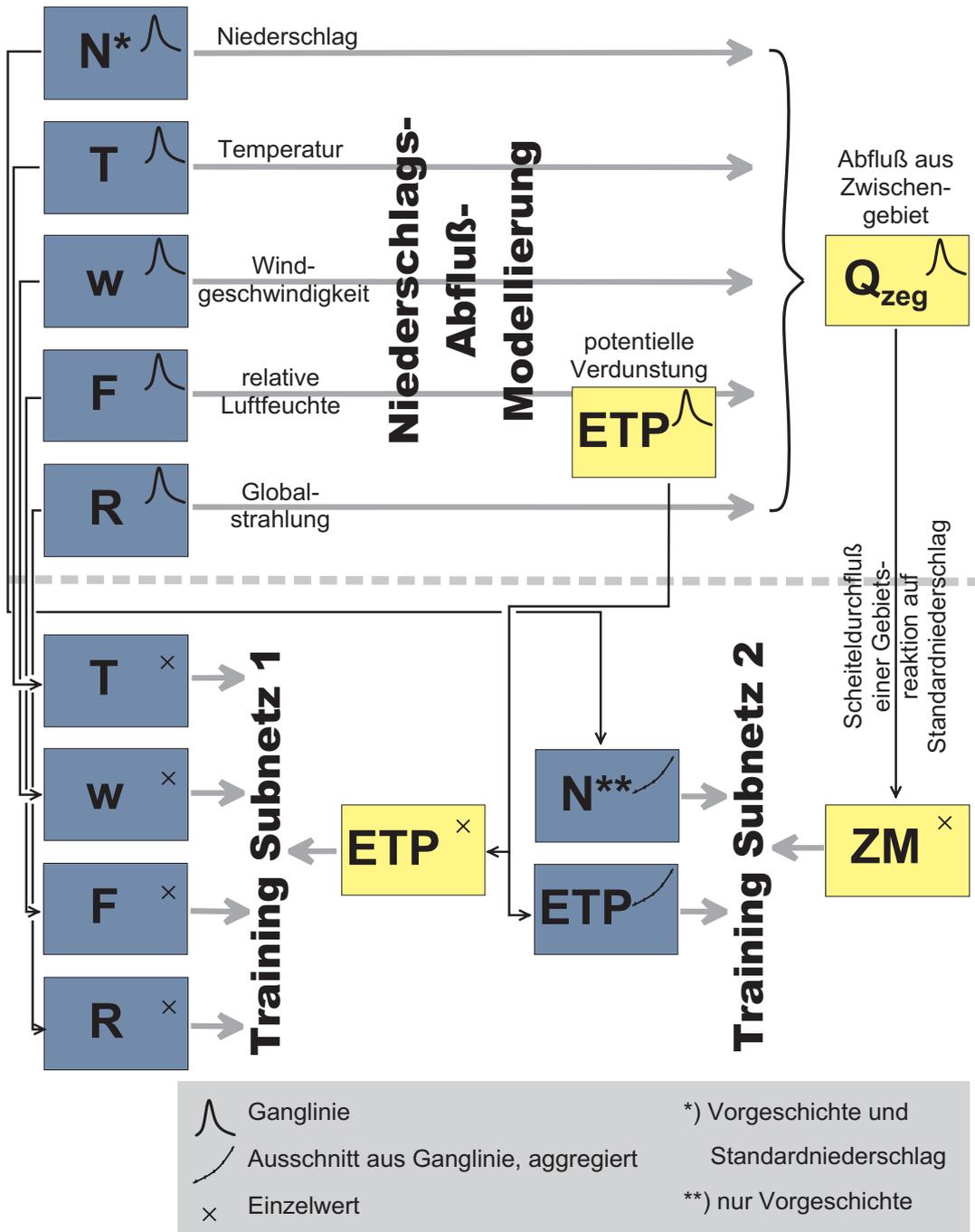


Abbildung 15: Datenfluß in WaSim-ETH (vereinfachtes Schema) und Unterteilung in Subnetze zur Berechnung des Zustandsmerkmals ZM

Im Gegenzug ist es aber nötig, ein zusätzliches neuronales Netz zu trainieren, um das Zustandsmerkmal aus den zurückliegenden Wetterdaten zu ermitteln. Da sich Speicher bei Wasserinput relativ schnell füllen, aber eher langsam wieder entleeren, ergeben sich bei einer trockenen Vorgeschichte auf dieses Zustandsmerkmal recht lange Einflußzeiten von mehreren Monaten. Um diese Fülle an Daten auf eine handhabbare Menge mit vertretbarem Verlust an Genauigkeit zu reduzieren, werden Erkenntnisse aus einer Analyse der Modellstruktur hinzugezogen. (Abbildung 15)

Der durch das Zustandsmerkmal zu charakterisierende Gebietszustand wird im Modell durch die Speicherfüllungen repräsentiert. Diese werden durch Niederschläge bzw. Schmelzwasser aus dem Schneemodell gefüllt und durch Abfluß ins Gewässer und Verdunstung entleert. Schneeschmelzprozesse sollen hier durch die Verwendung von Effektivniederschlägen grundsätzlich ausgeklammert bleiben. (Künftig sind vom DWD Effektivniederschläge als aktuelle Werte und Prognosen zu erwarten.) Während der Abfluß nicht direkt von meteorologischen Inputgrößen sondern von den Speicherfüllungen selbst abhängt, wird die reale Verdunstung *ETR* aus dem Wasserdargebot der Speicherfüllungen und der potentiellen Verdunstung *ETP* berechnet. Die reale Verdunstung und die Speicherfüllung beeinflussen sich also gegenseitig und damit ist die reale Verdunstung ebenso abhängig von der Vorgeschichte. Demgegenüber wird die potentielle Verdunstung (nach PENMAN-MONTEITH) wiederum jeweils aus den aktuellen Werten für Temperatur, Globalstrahlung, Windgeschwindigkeit, relativer Luftfeuchte und, um die Zugehörigkeit zu den Vegetationsperioden zu bestimmen, indirekt auch dem Datum für jeden Rechenzeitschritt berechnet. Die potentielle Verdunstung eines Zeitschrittes ergibt sich also nur aus Eingangsgrößen des gleichen Zeitschrittes. Verwendet man nur Effektivniederschläge (nach Durchlauf des Schneemodells), so münden sämtliche verwendeten Eingangsdaten mit Ausnahme des Niederschlages ausschließlich in der potentiellen Verdunstung *ETP*. Aus diesem Grunde bietet sich diese Größe an, in einem ersten Schritt für eine deutliche Reduktion der Datenmenge mit geringstmöglichem Informationsverlust genutzt zu werden. Dazu wird ein erstes Subnetz (Subnetz 1) auf Tageswertbasis trainiert.

Das zweite Subnetz (Subnetz 2) muß nun mit der Vorgeschichte aus der potentiellen Verdunstung *ETP* und dem Effektivniederschlag als Inputs und der Scheitelhöhe des Standardereignisses als Target erstellt werden. Die Dauer der dabei berücksichtigten Historie wird in einer Sensitivitätsanalyse ermittelt. Die unmittelbare Vorgeschichte erfordert eine stärkere Wichtung als länger zurückliegende Zeiträume. So sollen als

Merkmale für das Subnetz 2 für die jüngste Vergangenheit Effektivniederschläge und potentielle Verdunstungswerte in Form von Tagessummen eingehen. Je länger die Zeiträume zurückliegen, desto größer wird die Zeitspanne, für die die Eingangsgrößen zusammengefaßt werden. Für die am weitesten zurückliegende noch relevante Vergangenheit werden die Inputwerte als Monatssummen zusammengefaßt.

Wenngleich das Subnetz 1 zur Reduktion der Datenmenge anhand des hierin verwendeten Modells WaSim-ETH hergeleitet wurde, so ist die Wirkungsweise der prozeßbeschreibenden Module in physikalisch begründeten Modellen naturgemäß sehr ähnlich. So ist die Methodik unabhängig davon, ob es sich um ein rasterbasiertes Modell wie WaSim-ETH oder hydrotopbezogene Modelle handelt. Die Modellunabhängigkeit der Methodik wird dadurch nicht beeinträchtigt.

3.6.3 Räumliche und zeitliche Struktur der Eingangsdaten und NA-Merkmale

Dieser Abschnitt behandelt detailliert den Datenfluß und die Aggregation der Eingangs- und Ergebnisdaten des Niederschlags-Abfluß-Modells zu Inputmerkmalen und Targetgrößen für das Training eines künstlichen neuronalen Netzes. Dies erfolgt praktisch am Beispiel des verwendeten Niederschlags-Abfluß-Modells WaSim-ETH, was jedoch die Anwendbarkeit der Methodik auf andere Modellsysteme nicht einschränkt.

Zunächst soll zusammenfassend die räumliche und zeitliche Datenstruktur des deterministischen Modells WaSim-ETH den Erfordernissen an die Trainingsdaten der beschriebenen künstlichen neuronalen Netze gegenüber gestellt werden. Die Eingangsgrößen für die Anwendung des NA-Modells WaSim-ETH sind für die verwendete Parametrisierung generell räumlich und zeitlich hoch aufgelöst. Dabei entspricht die zeitliche Gliederung der Eingangsdaten dem Rechenzeitschritt, der bei der Betrachtung von Hochwasserereignissen in Mittelgebirgen üblicherweise eine Stunde beträgt. Die räumliche Gliederung ergibt sich aus der verwendeten Rasterweite, die die topographischen und gebietscharakteristischen Eigenschaften in ausreichender Genauigkeit abbilden kann. Die Ausgabe sämtlicher Größen inklusive des Abflusses erfolgt räumlich entsprechend der vorgegebenen Gliederung in Teileinzugsgebiete und zeitlich ebenfalls stundenweise.

Aufgrund der Fülle von Daten, die für die Niederschlags-Abfluß-Modellierung benötigt wird, ist für die Erstellung der neuronalen Netze eine Datenaggregation notwendig und soll nach folgenden Gesichtspunkten durchgeführt werden:

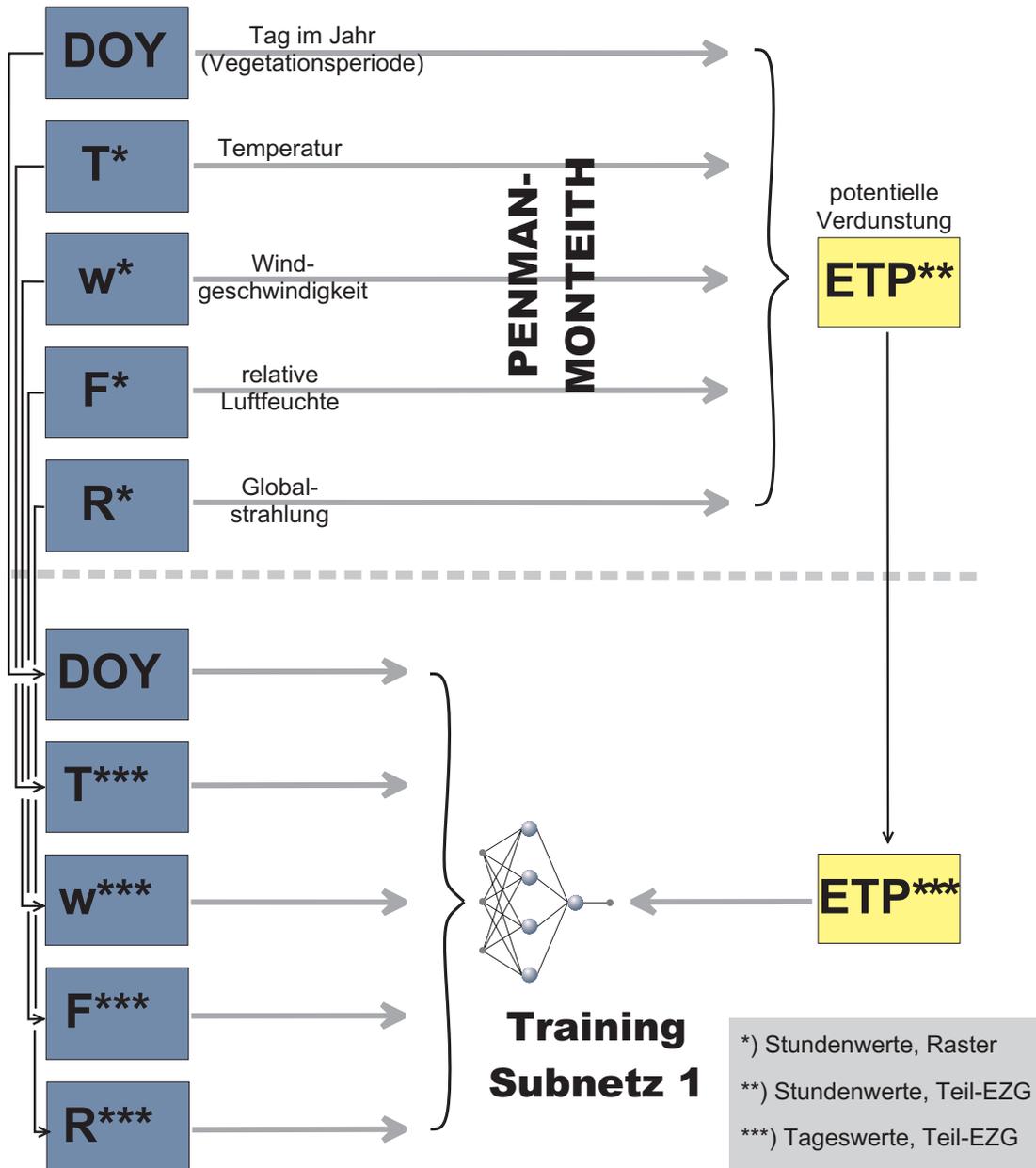


Abbildung 16: Datenfluß für Erstellung der Trainingsdaten und Training des Subnetzes 1

Wie im vorhergehenden Kapitel beschrieben, ist die Vorfeuchte in erster Linie von weniger dynamisch verlaufenden Wasserhaushaltsprozessen geprägt. Es wird für die Bestimmung des Zustandsmerkmals davon ausgegangen, daß eine zeitliche Gliederung der meteorologischen Inputdaten in mindestens Tagesschritte ausreichend ist. Für das Subnetz 2, welches aus der Vorgeschichte an potentieller Verdunstung und Niederschlag das Zustandsmerkmal bestimmt, werden die Eingangsdaten für länger zurückliegende Zeitspannen entsprechend ihres zeitlichen Abstandes zum Ereignis stärker zusammengefaßt (bis zu 30-Tages-Mitteln). Die letzten Tage vor dem

hochwasserauslösenden Niederschlag gehen als Tageswerte ein. Für die Berechnung der Hochwasserwellen (Ereignismarkere) wird die Auflösung der Niederschlagswerte von einer Stunde für die zentrale Einflußspanne beibehalten. Das Prognosenetz für eine variable Vorhersageweite wird für die Targetgrößen Durchfluß und Wasserstand für einzelne Stundenwerte trainiert. Desweiteren soll die Berechnung des Abflusses für das gesamte Zwischeneinzugsgebiet wie oben beschrieben ohne weitere räumliche Untergliederungen durchgeführt werden. Somit werden die WaSim-intern errechneten und separat ausgegebenen Gebietsmittel für Eingangsdaten und Ausgabegrößen für das Training der Netze verwendet.

Im folgenden wird der Datenfluß im Detail und netzweise beschrieben:

Abbildung 16 zeigt den Datenfluß für die Erstellung des Inputvektors und des Targetwertes bezüglich des Subnetzes 1, welches die Fülle an Eingangsdaten reduziert, indem die Zwischengröße potentielle Verdunstung *ETP* bestimmt wird. Das Niederschlags-Abfluß-Modell verwertet die natürlichen historischen Daten rasterweise und im Stundenschritt. Es erfolgen Ausgaben in Stundenwerten als Mittelwerte für das Zwischeneinzugsgebiet für sämtliche Eingangs- und Ausgabegrößen. Für die Übertragung der Daten zu Trainingspaaren für das Subnetz 1 werden die meteorologischen Eingangsgrößen Temperatur, Windgeschwindigkeit, relative Luftfeuchte und Globalstrahlung zu gebietsgemittelten Tageswerten zusammengefaßt und um das Datum als Nummer des Tages im Jahr ergänzt. Die Targetgröße potentielle Verdunstung wird als Tagessumme und Gebietsmittel verwendet. Da die potentielle Verdunstung unabhängig von einer Vorgeschichte ist und nur von Momentanwerten abhängt, tauchen diese Eingangsgrößen im Inputvektor nur einmalig als aktuelle Werte auf. Als Netztyp wird ein MLFN ausgewählt.

In Abbildung 17 ist die Datenaggregation für das Subnetz 2 dargestellt, wobei sich der Modellinput aus gemessenen kontinuierlichen Datenreihen von mehreren Monaten und einem anschließendem Standardniederschlag von 4 mm h^{-1} in 24 Stunden zusammensetzt. Dieser Standardniederschlag ist in allen Fällen gleich und dient dazu, eine Gebietsreaktion entsprechend dem aktuellen Feuchtezustand des Gebietes hervorzurufen. Der Output des Niederschlags-Abfluß-Modells ist eine Abflußganglinie am Gebietsauslaß. Aus diesem wird allein der Scheiteldurchfluß der Hochwasserwelle, die durch den Standardniederschlag ausgelöst wurde, als Zustandsmerkmal und damit als Targetwert für das Training des Subnetzes 2 übernommen. Die Netzinputs setzen sich aus den meteorologischen Größen, die zu diesem Gebietszustand vor dem Standardniederschlag geführt haben, zusammen. Diese sind hier der (Effektiv-)

Niederschlag und die potentielle Verdunstung, welche mit Hilfe von Subnetz 1 aus den in das NA-Modell eingehenden Meßgrößen ermittelt wurde. Der Niederschlag wird äquivalent zur potentiellen Verdunstung zu Tagessummen und Gebietsmitteln zusammengefaßt.

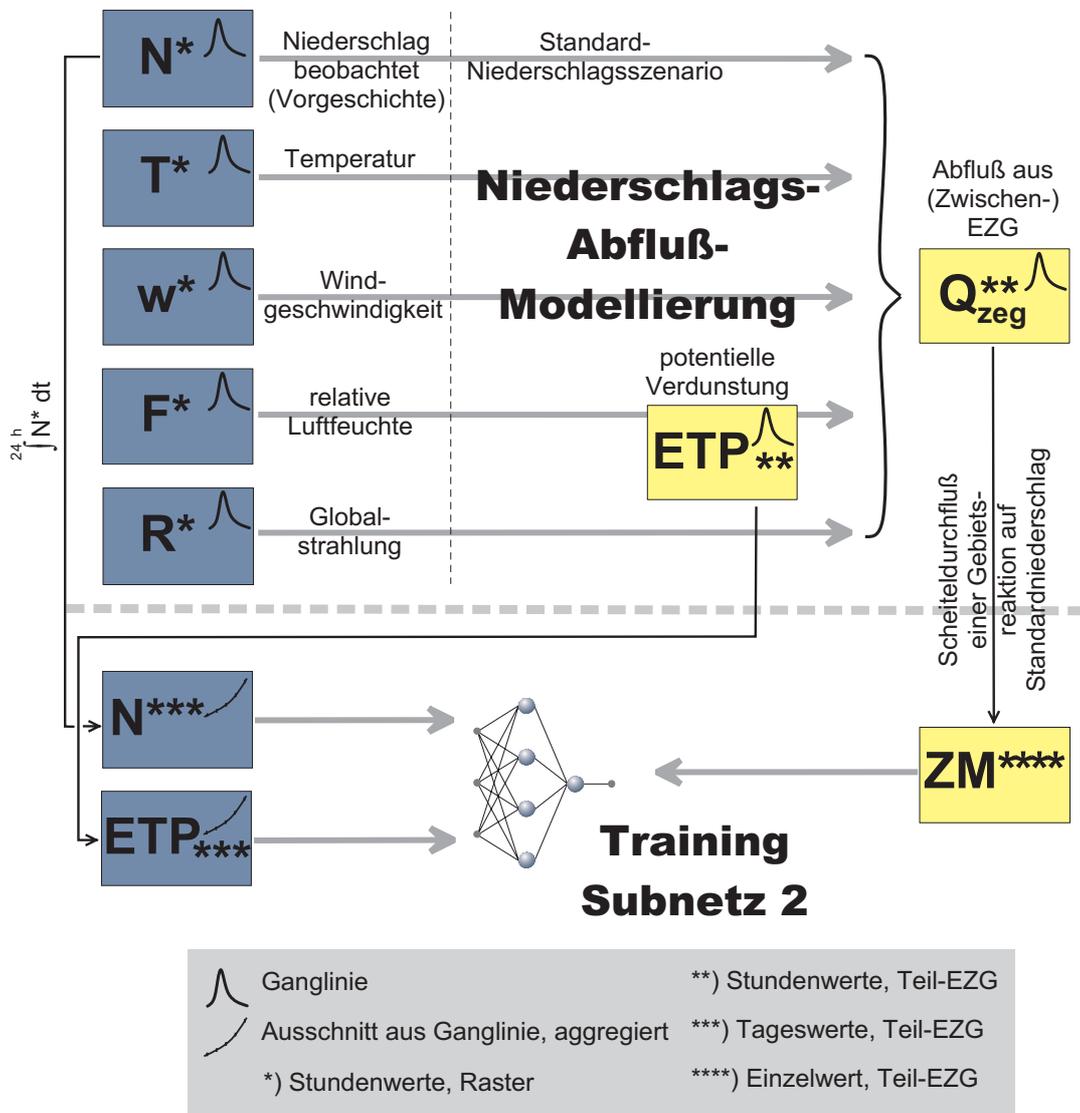


Abbildung 17: Datenfluß für Erstellung der Trainingsdaten und Training des Subnetzes 2

Weil die Gebietsfeuchte auf einer langen Vorgeschichte basiert, bilden die Tageswerte von Effektivniederschlag und potentieller Verdunstung dieser gesamten Vorgeschichte den Input des Netzes. Da diese relevante Vorgeschichte mehrere Monate zurückreicht, ist eine Aggregation der Daten erforderlich. Dazu werden diese Größen in einem zweiten Schritt dann entsprechend ihrer Ereignisferne weitergehend zusammengefaßt.

Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

Tabelle 3 zeigt die genaue räumliche und zeitliche Strukturierung der Daten im deterministischen Modell und für das künstliche neuronale Netz. Dabei wird zur Verdeutlichung des Prinzips von einer relevanten Vorgeschichte von 150 Tagen ausgegangen, was dem konkreten Beispiel des im folgenden betrachteten Einzugsgebietes der Freiburger Mulde entspricht.

Tabelle 3: Räumliche und zeitliche Diskretisierung der Daten

	WaSim-ETH		ANN	
	räumlich	zeitlich	räumlich	zeitlich
<u>Subnetz1</u>				
T, w, F, R	1km	1h	Zwischengebiet	1d
DOY	-	1d	-	1d
ETP	Zwischengebiet	1h	Zwischengebiet	1d
<u>Subnetz2</u>				
N _{eff}	Zwischengebiet	1h	Zwischengebiet	1d****
ETP	Zwischengebiet	1h	Zwischengebiet	1d****
Q _{ges} */ZM**	Zwischengebiet	1h	Zwischengebiet	Einzelwert
<u>Prognosenetz</u>				
N _{eff} ***	Zwischengebiet	1h	Zwischengebiet	1h
T, w, F, R*/ZM**	1km	1h	Zwischengebiet	Einzelwert

* nur WaSim-ETH

** nur ANN

*** für ANN: explizit nur für Ereignis, für Vorgeschichte in ZM enthalten

**** letzte 150 Tageswerte gemittelt für 2x30, 2x15, 4x8, 4x4, 4x2, 4x1 Tage

Die Trainingsdaten für das eigentliche Prognosenetz werden mit dem NA-Modell auf ähnliche Weise gewonnen, wie die Daten für das Subnetz 2. Der Standardniederschlag wird dabei durch unterschiedliche künstliche Niederschlagsszenarien ersetzt (Abbildung 18). Da die natürliche Vorgeschichte bereits durch das Zustandsmerkmal aus dem Subnetz 2 repräsentiert wird, sind für das neuronale Netz zur Abbildung der Hochwasserganglinie darüber hinaus nur Ereignismerkmale notwendig, d. h. die Niederschlagswerte zwischen eigentlichem Ereignisbeginn und vorherzusagendem Durchflußwert. Es werden demnach alle Stundenwerte des Niederschlages für das Ereignis als Basis für die Merkmale und die dazugehörigen Durchflußwerte als Targets übernommen. Eine Aggregation der Inputdaten zu Ereignismerkmalen erfolgt entsprechend der Relevanz der einzelnen Niederschlagswerte für den Durchflußwert. Niederschlagswerte mit einer zeitlichen Differenz, die sich innerhalb der zentralen Einflußspanne finden, werden als Einzelwerte in den Merkmalsvektor übernommen. Außerhalb davon werden Summen über mehrere Stunden zu Ereignismerkmalen zusammengefaßt (siehe 3.6.1). Da ein Netz alle Vorhersageweiten abdecken soll,

Synthese physikalisch begründeter Modellierung mit künstlichen neuronalen Netzen

müssen für länger andauernde Ereignisse die zeitlich weiter zurückliegenden, aber nach Ereignisbeginn (für welchen das Zustandsmerkmal gilt) aufgetretenen Niederschlagswerte in einen Summenwert einmünden, der den Beitrag dieses Niederschlags zur Aufwechtung des Gebietes beschreibt. Für kürzere Vorhersageweiten beschreiben diese und möglicherweise andere Merkmale Niederschlagssummen vor Ereignisbeginn, die in diesem Falle null gesetzt werden.

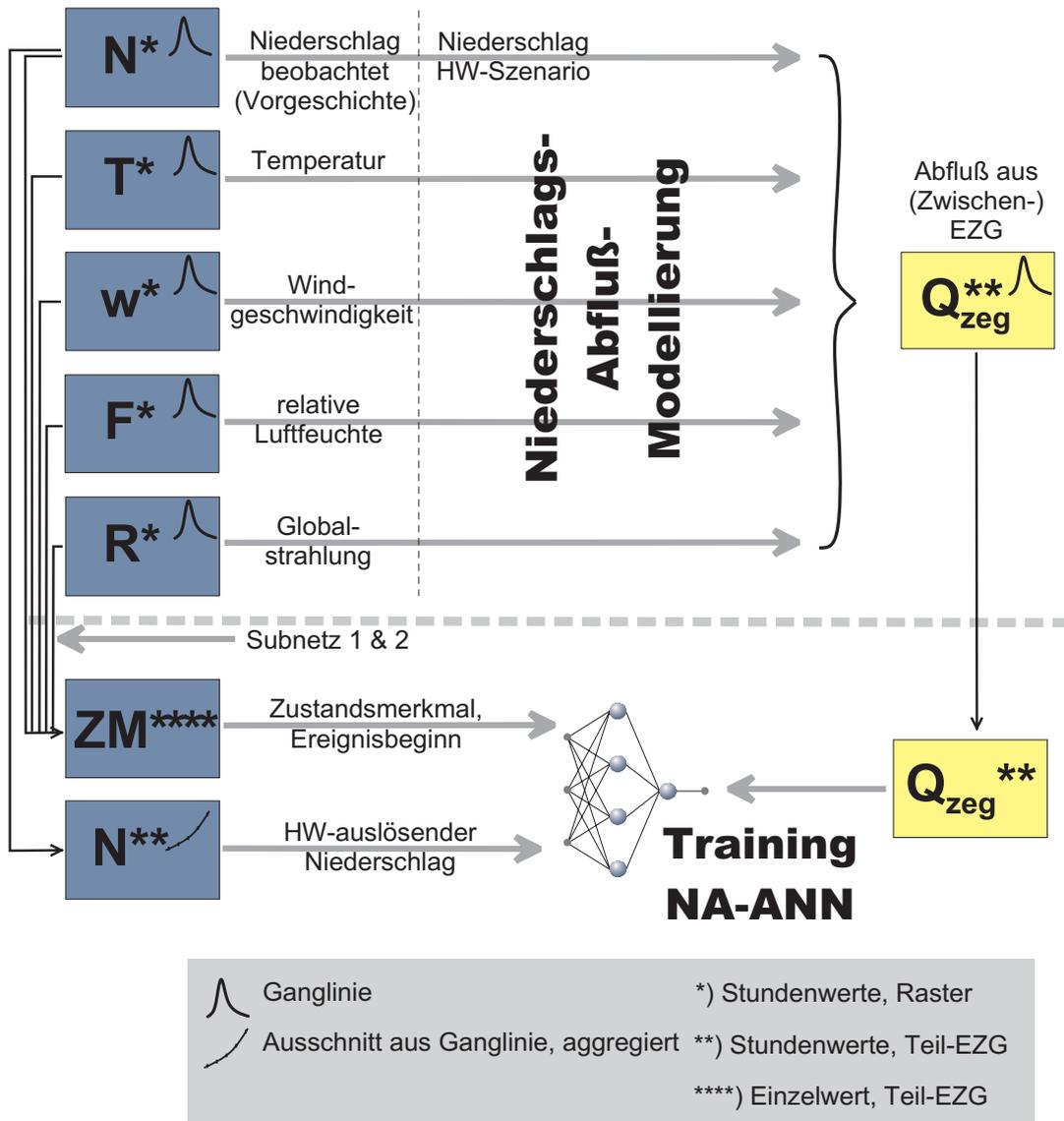


Abbildung 18: Datenfluß für das Prognosenetz

Die Merkmale aus den zeitlich hochaufgelösten Effektivniederschlägen ab Ereignisbeginn und das Zustandsmerkmal zur Beschreibung der Gebietsfeuchte bilden die Merkmale für die Niederschlags-Abfluß-Prozesse im Zwischeneinzugsgebiet.

4 Versuchsgebiet und Erstellung der Trainingsdaten

4.1 Allgemeine Angaben zum Untersuchungsgebiet

Lage und Geomorphologie

Das 2983 km² große Einzugsgebiet der Freiburger Mulde erstreckt sich zwischen den Kammlagen des mittleren Erzgebirges im Süden mit der höchsten Erhebung am Fichtelberg von 1244 m NN bis zum Rand der Leipziger Tieflandsbucht im Nordwesten mit einer geodätischen Höhe von 132 m NN am Gebietsauslaß am Pegel Erlln. Daraus resultiert eine Höhendifferenz von 1112 m im Gebiet. Die Nord-Süd-Ausdehnung beträgt ca. 94 km. Ca. 10% der Einzugsgebietsfläche befinden sich auf tschechischem Staatsgebiet.

Das heutige Landschaftsbild reflektiert die Erdgeschichte des Erzgebirges, wonach der gebietsrelevante Bereich durch Hebungs- und Senkungsprozesse nach Nordwesten hin schräggestellt wurde. Dabei sind im westlichen Teil des Einzugsgebietes, der von der Zschopau entwässert wird, zumeist steile Kerbtäler anzufinden, der östliche Teil ist eher durch sanftere Muldentäler geprägt. Das mittlere Hanggefälle beträgt 5,5 ° mit Maxima von >15 ° unmittelbar an Gewässerläufen. Abbildung 19 stellt die Höhengliederung und die Hangneigung im Einzugsgebiet dar.

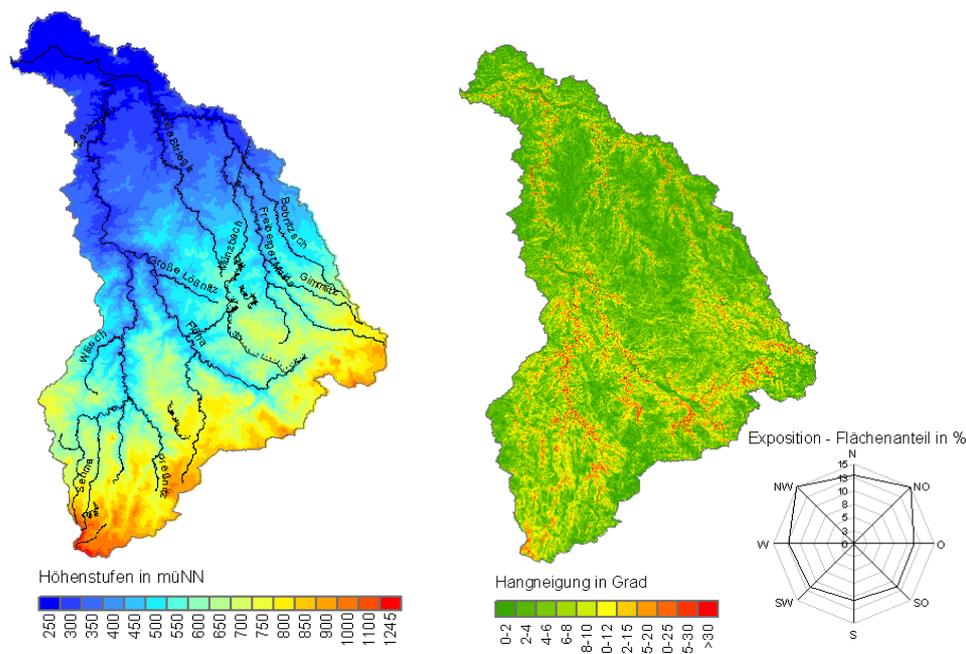


Abbildung 19: Höhen und Gefälle im Untersuchungsgebiet

Klima

Nach dem Meteorologischen Dienst der DDR (1973) ist das Erzgebirge der gemäßigten Klimazone mit kontinentalem Einfluß zuzuordnen. Es besteht eine orographische Abhängigkeit der Klimagrößen. Das Wetter wird von Luftmassen aus vorrangig westlicher bis norwestlicher Richtung bestimmt. In geschützten Tälern herrschen oft mildere Verhältnisse als in den Höhenlagen.

Sonnenscheindauer und Globalstrahlung erreichen im Mai ihr Maximum, schwanken aber im Sommer aufgrund unsteter Konvektionsbewölkung. Im Jahresmittel beträgt die Sonnenscheindauer ca. 32 % der maximal möglichen Dauer auf dem Fichtelberg und 35 % in Annaberg-Buchholz.

Die mittlere Jahreslufttemperatur nimmt vom Tiefland bis zum Kamm von 7,6 °C in Freiberg (398 m NN) auf unter 2,7 °C auf dem Keilberg (1244 m NN) ab. Die Monatsmittel sinken von -1,2 °C auf -5,6 °C für Januar und für Juli von 16,6 °C auf 11,3 °C.

Das Einzugsgebiet liegt insbesondere bei Westwind im Lee-Bereich vorgelagerter Gebirge (Thüringer Wald, Rothaargebirge, Harz), weshalb die mittleren Jahresniederschläge jene anderer deutscher Mittelgebirge unterschreiten. Gleichfalls ist von West nach Ost eine Abnahme der Jahresniederschläge zu verzeichnen. Das Maximum des Niederschlags findet sich im Sommer. Mit der Höhe steigen die mittleren Jahresniederschläge von 650 mm auf 1100 mm, dabei im Monatsmittel Januar von 40 mm auf 100 mm und im Juli von 90 mm auf 130 mm. Lokale Effekte wirken sich z.T. erheblich aus, so sind Spitzenwerte an exponierten Stellen wie dem Fichtelberg zu verzeichnen.

Die Anzahl der Tage mit Schneedecke steigt ebenfalls deutlich mit der Höhe von 60 d auf über 150 d am Fichtelberg. Alle mittleren und höheren Lagen weisen mehr als 50 % der Tage eine Schneedecke auf.

Geologie und Hydrogeologie

Im Gebiet dominieren metamorphe, wenig klüftige und somit kaum Grundwasser führende Gneise. Die Magmatite besitzen oft eine gute Klüftung und entsprechend auch eine geringfügige Grundwasserleitung. Ferner führen stärkere tertiäre Sedimente in weiteren Flusstälern sowie lokale Sand- und Kiesschichten im mit Löß bedeckten Tiefland etwas Grundwasser.

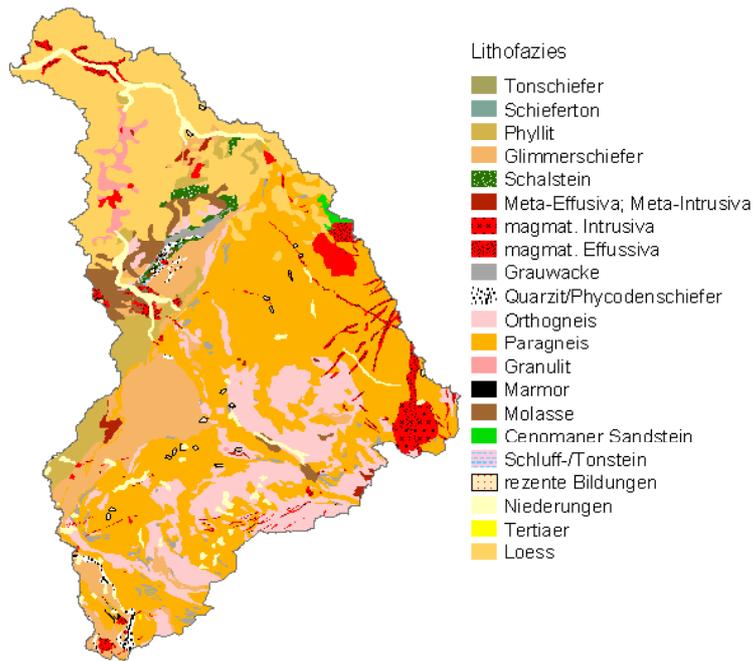


Abbildung 20: Lithofazieseinheiten im Untersuchungsgebiet

Bodentypen und Bodenarten

Im Einzugsgebiet der Freiburger Mulde sind vor allem saure, flachgründige (0,2 - 0,6 m) Braunerden und Lößgemische verbreitet (Abbildung 21). Die Bodenmächtigkeit nimmt mit steigendem Hanggefälle ab. Dieser Umstand begünstigt gemeinsam mit dem hohen Skelettanteil eine extrem schnelle Reaktion auf Niederschläge und damit vor allem lateralen, hochwasserwirksamen hypodermischen Abfluss.

Über den metamorphen Gneisen ist Hangsandlehmbraunerde entstanden. Dieser Boden wird reliefabhängig bis Höhen von etwa 550 m NN landwirtschaftlich, sonst forstwirtschaftlich genutzt. In den Niederungen bis ca. 350 m NN, an Gewässerläufen und in Senken der Hochlagen treten grund- und stauwasserbeeinflusste Böden (Gleye, Moore) auf. Charakteristisch für diese Böden sind periodisch wechselnde Vernässung und Austrocknung.

Die Spezifikation der Bodentypen in Bodenarten zur Schätzung bodenhydraulischer Eigenschaften für die hydrologische Modellierung orientiert sich an der Bodenübersichtskarte von Sachsen 1:400.000 und ist in Abbildung 21 definiert.

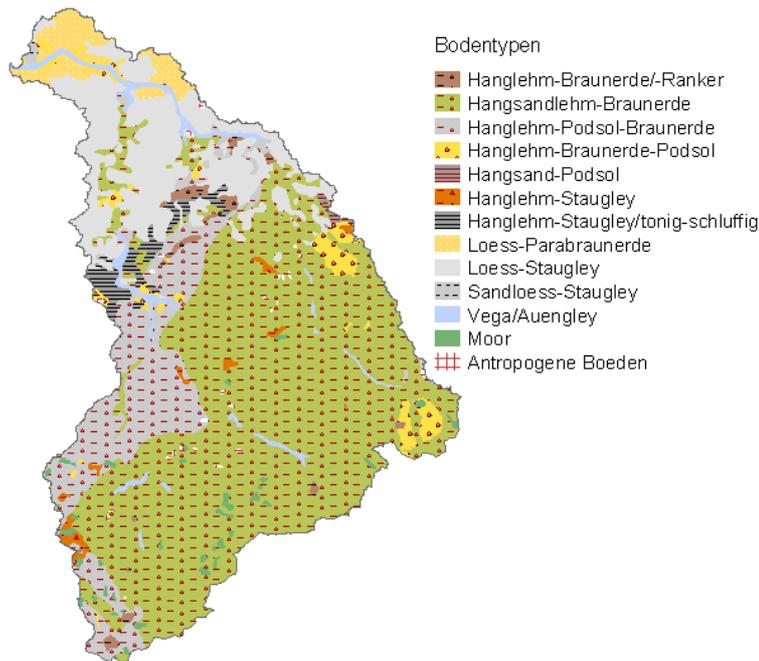


Abbildung 21: Bodenkarte des Untersuchungsgebietes

Das skelettreiche (bis 50 % Steine) oberflächennah anstehende Gestein begünstigt einen schnellen Wassertransport in tiefere Bodenschichten und bei ausreichendem Hanggefälle schnellen unterirdischen Abfluss. Im Tiefland steigt durch Erosion und Sedimentation der Feinkornanteil im Oberboden stark. Die Horizontfolge ist lehmiger Schluff bis schluffiger Lehm (IU, uL) über kiesig-sandigem Lehm bis Lehm (L) über kiesigem Sand und verwittertem Gestein. In den Flusstälern dominiert lehmiger Sand (IS) über Sand, Kies oder Schotter.

Landnutzung

Die heutige Flächennutzung ist ein Abbild der Siedlungsgeschichte des Erzgebirges.

Nach Lauterbach (2000) entspricht die potentielle Vegetation dem Waldbild der älteren Nachwärmezeit (600-1300 n.Chr.) mit einer Buchen-Tannen-Stufe von 500 bis 600 m NN und darüber einer Buchen-Tannen-Fichten-Stufe. Heute gilt das touristisch erschlossene Erzgebirge bis in die Kammlagen als eines der am dichtesten besiedelten Gebirge Europas. Große Waldflächen beschränken sich auf Steilhänge und Regionen oberhalb der klimatischen Landwirtschaftsgrenze. Es dominieren Fichten-Monokulturen und einzelne Zwergstrauchheiden. Wegen stark klimaabhängiger Ackererträge wird Viehhaltung bevorzugt.

Versuchsgebiet und Erstellung der Trainingsdaten

Die Verteilung der Landnutzung im Einzugsgebiet veranschaulicht Abbildung 22. Mit 57 % dominiert der Ackerbau. Der Flächenanteil urbaner Strukturen ist mit 6 % gering und beschränkt sich weitgehend auf Lagen unterhalb der Nadelwaldgrenze bis ca. 700 m NN. Ein Anteil von 40 % des Einzugsgebietes besteht aus bewaldeten Flächen, 6 % werden von Heiden und Moorheiden und 5 % von Wiesen und Weiden bedeckt, sowie 29 % als nicht bewässertes Ackerland klassifiziert. Mit zunehmender Besiedlung in Richtung Norden dominiert landwirtschaftliche Nutzung.

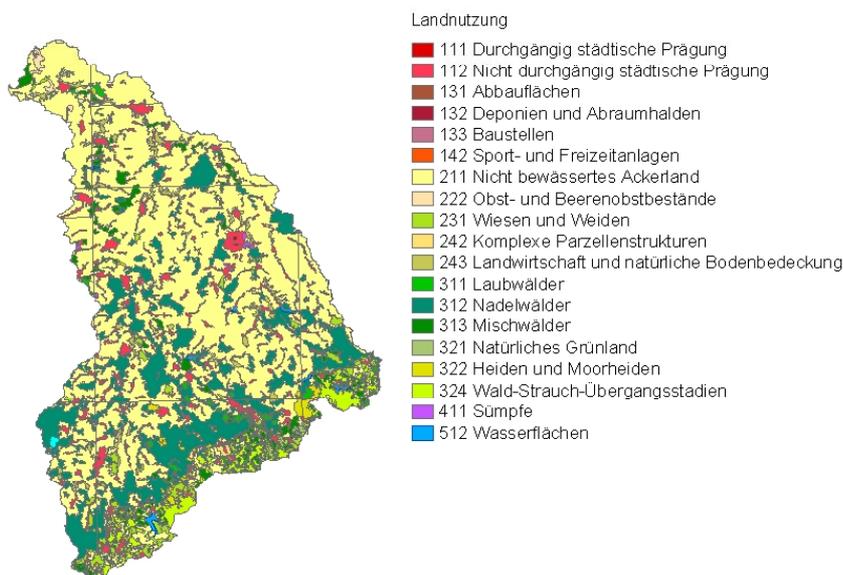


Abbildung 22: Landnutzung des Einzugsgebietes der Freiburger Mulde

Gewässer

Die Freiburger Mulde entspringt am östlichsten Zipfel des Einzugsgebietes auf tschechischem Territorium am Kamm des Erzgebirges und fließt auf einem Fließweg von ungefähr 120 km bis zum Pegel am Gebietsauslaß in Erlbn. Am Pegel Nossen 1 (km 47), dem Eintritt in die hydrodynamisch modellierte Fließstrecke, konzentriert sich der Abfluß von 585 km² Einzugsgebietsfläche. Dieser Teil des Einzugsgebietes ist eher schmal und umfaßt einen sehr kleinen Teil der Kammlagen des Erzgebirges.

Am Flußkilometer 37 mündet die Striegis mit einer Einzugsgebietsflächen von 283 km² in die Freiburger Mulde ein. Dieses Gebiet liegt ausschließlich im flacheren Teil und hat keinen Anteil am Erzgebirgskamm. Der Pegel Niederstriegis befindet sich knapp 200 Meter von der Einmündung entfernt.

Versuchsgebiet und Erstellung der Trainingsdaten

Am Flußkilometer 24 befindet sich die Einmündung der Zschopau in die Freiburger Mulde. 14 Kilometer stromauf befindet sich der Pegel Kriebstein, der mit 1757 km² weit über die Hälfte des Gesamteinzugsgebietes abdeckt. Zusammen mit deren größten Nebenfluß Flöha, die wenige Kilometer neben der Quelle der Freiburger Mulde im Osten des Einzugsgebietes entspringt, entwässert die Zschopau fast den gesamten Erzgebirgskamm. Aufgrund dieser hydrologischen Gegebenheiten werden bei der Modellierung der Unterlauf der Zschopau und die Freiburger Mulde zwischen der Zschopaumündung und dem Pegel ErlIn als Hauptgewässer angesehen und die Freiburger Mulde stromauf der Zschopaumündung als Nebenfluß. Eine komplette graphische Darstellung des Gewässernetzes ist in Anhang B zu finden.

Tabelle 4 faßt die hydrologischen Kenndaten der erwähnten Pegel nach dem Deutschen Gewässerkundlichen Jahrbuch des Elbegebietes (Teil I) auf dem Stand von 2003 zusammen. Da der Pegel Kriebstein nicht im DGJ aufgeführt ist, wurde außerdem der benachbarte Pegel Lichtenwalde in die Tabelle aufgenommen. Außerdem finden die Pegel Hopfgarten und Borstendorf, die letzten Pegel von Zschopau und Flöha vor deren Zusammenfluß, Eingang.

Tabelle 4: Gewässerkundliche Hauptwerte nach DGJ (2003)

Gewässer	Pegel	Mündung [km]	Fläche [km ²]	Gewässerkundliche Hauptwerte [m ³ /s]			
				MQ	MHQ	HHQ	Datum
Freiberger Mulde	ErlIn	1,5	2982	35	300	610	09.12.1974
Freiberger Mulde	Nossen1	47,2	585	6.81	68.7	690	13.08.2002
Striegis	Niederstriegis	0,2	283	2.68	27.1	69.1	06.07.1958
Zschopau	Kriebstein	14,4	1757	- (kein Jahrbuchpegel)			
Zschopau	Lichtenwalde	45,6	1575	21.8	218	1250	13.08.2002
Zschopau	Hopfgarten	82,8	529	7.94	81.2	420	13.08.2002
Flöha	Borstendorf	23,6	644	9.15	91.3	540	13.08.2002

4.2 Untergliederung in Teileinzugsgebiete für die deterministische Modellierung

Das gesamte Einzugsgebiet trägt zum Abfluß am Pegel ErlIn bei und muß somit vom Niederschlags-Abfluß-Modell abgedeckt sein. Neben der Verfügbarkeit der Ganglinien und der Validierbarkeit des Modells an Pegelstationen dient die Untergliederung in Teileinzugsgebiete zur Trennung in Flächen der Abflußbildung und -konzentration zum Gebietsauslaß und Fließstrecken, auf denen der Wellenablauf durch die darauffolgenden Teilgebiete beschrieben wird. Dies gilt sowohl für die in WaSim-ETH

Versuchsgebiet und Erstellung der Trainingsdaten

integrierte Wellenablaufberechnung mit dem Translations-Diffusions-Ansatz als auch für die externe hydrodynamische Modellierung mit HEC-RAS, bei der die Teilgebietsauslässe als laterale Zuflüsse eingehen.

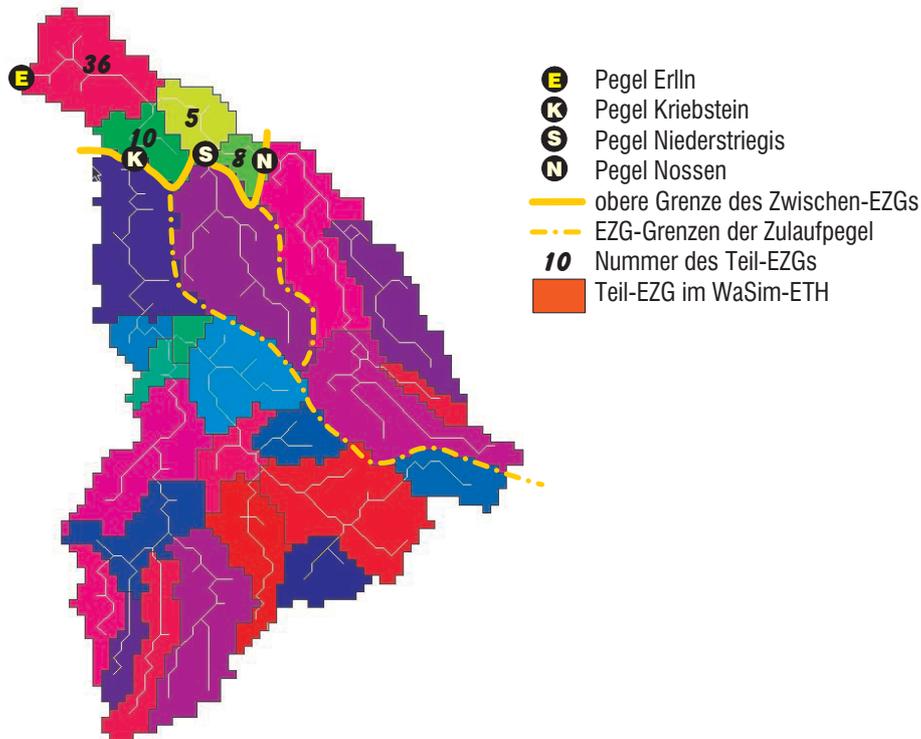


Abbildung 23: Untergliederung des Einzugsgebietes

Abbildung 23 zeigt die Untergliederung des Einzugsgebietes. In dieser Arbeit wird eine Parametrisierung von WaSim-ETH nach Cullmann (2007) für die Niederschlags-Abfluß-Modellierung der Hochwasserszenarien an den Pegeln Kriebstein, Niederstriegis und Nossen (mit einem weißen Großbuchstaben markiert) verwendet. In den Einzugsgebieten dieser Pegel wird die Unterteilung der Einzugsgebiete nach Cullmann (2007) beibehalten.

Da im Zwischeneinzugsgebiet unterhalb der bei Cullmann (2007) modellierten Flächen keine Pegel zur Zwischenbilanzierung zu finden sind, wird eine Gliederung entsprechend der natürlichen Gewässerstruktur übernommen. Die Auslässe der Teileinzugsgebiete werden unmittelbar an den Zusammenflüssen der Hauptflüsse Freiburger Mulde, Striegis und Zschopau platziert. Das Zwischeneinzugsgebiet entlang der hydrodynamisch modellierten Fließstrecken wird also in folgende vier Teilgebiete unterteilt:

- Freiburger Mulde von Pegel Nossen bis Einmündung der Striegis (Teilgebiet 8 in Abbildung 23)
- Freiburger Mulde zwischen den Einmündungen der Striegis und der Zschopau (Teilgebiet 5 in Abbildung 23)
- Freiburger Mulde von Einmündung der Zschopau bis Pegel ErlIn (Teilgebiet 36 in Abbildung 23)
- Zschopau von Pegel Kriebstein bis Einmündung in Freiburger Mulde (Teilgebiet 10 in Abbildung 23)

Das dabei nicht berücksichtigte Stück der Striegis ist mit 0,2 Kilometern wesentlich kleiner als die Rasterweite von einem Kilometer, so daß es dem obersten Teilgebiet der Freiburger Mulde zugeschlagen wird. Die Abgrenzung der Teilgebiete erfolgt mit dem WaSim-Tool Tanalys (Schulla et al, 1999). Tanalys ist ein Preprocessing-Werkzeug, welches die Fließwege, das Gefälle, die Parameter für den Translations-Diffusions-Ansatz zur Beschreibung des Wellenablaufes, Flächenlaufzeiten der Abflußkonzentration und andere geländeabhängige Größen für das Niederschlags-Abfluß-Modell aus dem digitalen Höhenmodell ableitet.

Damit definieren sich die Zuflüsse für das hydrodynamische Modell: Die Hauptflüsse Zschopau, Striegis und Freiburger Mulde werden durch Ganglinien an den Pegeln Kriebstein, Niederstriegis und Nossen repräsentiert, das Zwischeneinzugsgebiet durch die Teilgebiete 8, 5, 36 und 10.

4.3 Einzugsgebietsgliederung der ANN-Szenarien

Nach 3.4 soll der Einfluß der Verwendung von Merkmalen für den Beitrag des Zwischeneinzugsgebietes anhand dreier Szenarien bezüglich des flächenmäßigen Anteiles dieses Zwischengebietes untersucht werden (Tabelle 2). Bezogen auf das Einzugsgebiet der Freiburger Mulde am Pegel ErlIn entspricht die originale Gliederung nach Abbildung 23 dem Szenario mit 10 Prozent Flächenanteil (HD10 / HDNA10). Für die beiden anderen Varianten soll das originale Zwischeneinzugsgebiet um die Einzugsgebiete des Pegels Niederstriegis bzw. der Pegel Niederstriegis und Nossen erweitert werden (Abbildung 24, Tabelle 5).

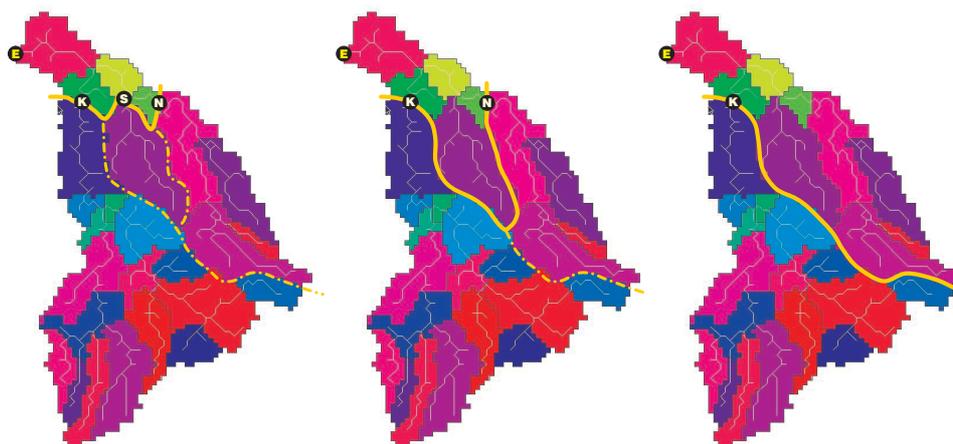


Abbildung 24: Einzugsgebietgliederung entsprechend der Fälle HD10/HDNA10 (links), HD20/HDNA20 (mitte), HD40/HDNA40 (rechts)

Tabelle 5: Flächenanteile der Einzugsgebiete der Zuflußpegel und des Zwischeneinzugsgebietes entsprechend der Szenarien

	Zuflußpegel mit Einzugsgebietsfläche (Prozent der Gesamtfläche)	Zwischeneinzugsgebiet (Prozent der Gesamtfläche)
HD10/HDNA10	Kriebstein, 1757 km ² (58,9%) Nossen, 585 km ² (19,6%) Niederstriegis, 283 km ² (9,5%)	358 km ² (12,0%)
HD20/HDNA20	Kriebstein, 1757 km ² (58,9%) Nossen, 585 km ² (19,6%)	641 km ² (21,5%)
HD40/HDNA40	Kriebstein, 1757 km ² (58,9%)	1226 km ² (41,1%)

Für die Pegel Kriebstein, Niederstriegis und Nossen liegen validierte Parametrisierungen des Modells WaSim-ETH vor (Cullmann, 2006). Für das Zwischeneinzugsgebiet unterhalb dieser Pegel wird eine Modellparametrisierung benachbarter Teilflächen entsprechend der geringsten Entfernung übernommen. Dies entspricht dem Szenario mit einem Flächenanteil des Zwischeneinzugsgebietes am Gesamtgebiet von 10 Prozent. Aber auch für die anderen Fälle, bei denen die Einzugsgebiete Niederstriegis bzw. Nossen und Niederstriegis dem (als unbeobachtet angenommenen) Zwischeneinzugsgebiet zugeschlagen werden, wird die validierte Parametrisierung beibehalten. Dies ist insofern sinnvoll, da hier untersucht werden soll, inwiefern die mit dem deterministischen Modell beschriebenen Prozesse mit künstlichen neuronalen Netzen nachgebildet werden könne. Um vergleichbare Ergebnisse zu erhalten, sollen auf Modellebene keine Unterschiede bestehen. Wie gut

Parametersätze des deterministischen Modells auf benachbarte Teileinzugsgebiete übertragbar sind, ist nicht Gegenstand dieser Arbeit und soll hier ausgeklammert bleiben.

4.4 Daten zur Gerinnegeometrie in HEC-RAS

Die Daten der Gerinnegeometrie charakterisieren die Querprofile der Flußabschnitte der Freiburger Mulde, der Zschopau und der Striegis unterhalb der Pegel Nossen, Kriebstein UP und Niederstriegis, sowie die unteren Ausläufer der Zwickauer Mulde und 7 km des Oberlaufes der Vereinigten Mulde. Die Höheninformationen für die entsprechenden Querprofile entstammen terrestrischen Vermessungen der Flüsse nach dem Augsthochwasser 2002. Diese und detaillierte Daten zur Verlängerung der Querprofile in die im Hochwasserfall von Überschwemmungen betroffenen Vorlandbereiche in Form von aktuellen Laserscanningdaten wurden vom Regierungspräsidium Chemnitz und der Landestalsperrerverwaltung des Freistaates Sachsen zur Verfügung gestellt. Für eine primäre Abschätzung der Rauigkeiten in den Flußauen wurden digitale Landnutzungsdaten verwendet und entsprechende Tabellen in der Literatur (Preißler und Bollrich, 1980; USACE1) zugeordnet.

4.5 Pegeldata zur Eichung des hydrodynamischen Modells

Zur Eichung des hydrodynamischen Modells wurden Durchflußganglinien der Oberliegerpegel und des Zielpegels in ErlIn verwendet. Diese Daten wurden vom sächsischen Landesamt für Umwelt und Geologie zum Teil in digitaler Form und zum Teil als Pegelbögen zur Verfügung gestellt. Die bereits digitalisierten Daten lagen teilweise in der gewünschten zeitlichen Auflösung von einer Stunde, aber auch als 2- oder 3-Stundendaten vor. Als problematisch stellte sich heraus, daß nur für relativ wenige Ereignisse Daten für alle 4 benötigten Pegel vorlagen. Tabelle 6 zeigt die Scheiteldurchflüsse der Ereignisse an allen vier Pegeln. Bei den in Abbildung 25 dargestellten Ganglinien der Hochwasserereignisse von 1974 und 1983 bzw. 1987 wird offensichtlich, daß die Massenbilanz in den Daten nicht gewahrt wird. Der Pegel Kriebstein, der knapp 60 % der Fläche des Einzugsgebietes des Pegels ErlIn abdeckt, zeigt im Dezember fast den gleichen Scheiteldurchfluß wie der Pegel ErlIn, im Januar 1987 noch reichlich die Hälfte. Die Summe der Scheiteldurchflüsse der drei Oberliegerpegel erreicht bei den dargestellten Daten nur für Dezember 1974 und Juli

Versuchsgebiet und Erstellung der Trainingsdaten

1996 den Scheitel am Pegel ErlIn. Wenngleich der Einfluß des Zwischeneinzugsgebietes bei dieser Betrachtung vernachlässigt wird, kann auf fehlerbehaftete WQ-Beziehungen geschlossen werden (LfUG, 2006).

Tabelle 6: Scheiteldurchflüsse der Hochwasserereignisse zur Kalibrierung/ Validierung des hydrodynamisch-numerischen Modells

Scheiteldurchfluß	ErlIn [m ³ /h]	Kriebstein [m ³ /h]	Nossen 1 [m ³ /h]	Niederstriegis [m ³ /h]
Dezember 1974	608	598	87	34
August 1983	569	417	88	60
Januar 1987	444	246	82	48
Juni 1995	433	288	62	35
September 1995	453	298	78	36
Juli 1996	241	206	36	14
März 1998	334	204	62	24

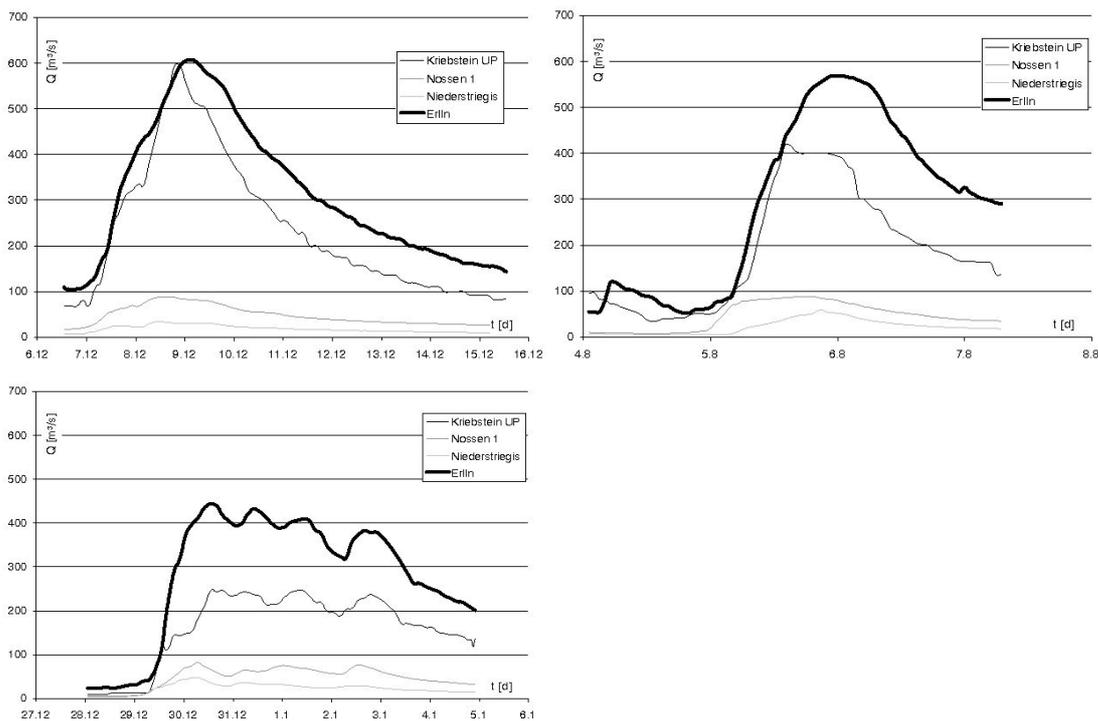


Abbildung 25: Ganglinien der Hochwasser im Dezember 1974, August 1983 und Januar 1987 an den Pegeln

Für die Modellierung mit HEC-RAS wurde der Einfluß des Zwischengebietes in Form von lateralen Zuflüssen mit dem Modell WaSim-ETH mit übertragenen Parametersätzen und Wetterdaten simuliert. Eine Verifizierung dieser Ganglinien war

allerdings unmöglich, da diese Gebiete nicht mit Pegeln ausgestattet sind und der Anteil am Gesamtabfluß in ErlIn in der Größenordnung des Fehlers durch Inkonsistenz der Pegel im Einzugsgebiet liegt.

4.6 Datenbasis für die Szenarienrechnung

Für die Szenarienrechnungen mit dem hydrodynamischen Modell und das Training der neuronalen Netze werden Ganglinien benötigt, die in Verlauf und Dimension dem gleichen Ereignis zugeordnet werden können. Darüber hinaus sind für alle Teilgebiete, die nach Abbildung 24 als Zwischengebiet in den verschiedenen Szenarien definiert werden, zum Training der neuronalen Netze Wetterdaten zur Verfügung zu stellen. Dies ist nur zu gewährleisten, wenn die gesamten Trainingsdaten mit dem Niederschlags-Abfluß-Modell generiert werden. Das bedeutet, Grundlage der Erstellung von Trainingsdatensätzen ist die Generierung meteorologischer Datensätze, mit Hilfe derer und des parametrisierten NA-Modells die entsprechenden Ganglinien simuliert werden können

4.6.1 Niederschlag und sonstige Wetterdaten

Die durch das Zustandsmerkmal repräsentierten Gebietszustände zu Ereignisbeginn sollen ebenso wie die generierten hochwasserauslösenden Niederschläge die gesamte Palette der praktisch möglichen Varianten umfassen. Für die Berechnung dieser Gebietszustände mit dem Niederschlags-Abfluß-Modell wurden gemessene meteorologische Datenreihen verwendet, die für die Jahre 1953-1999 vorlagen und mittels eines Kriging-Verfahrens (external drift kriging) interpoliert wurden. Da die generierten Hochwasserszenarien zeitlich über den gesamten Zeitraum verteilt wurden, treten sie sowohl in trockenen Perioden als auch in direkter Folge zu tatsächlich abgelaufenen, in den historischen Niederschlagsreihen enthaltenen Hochwasserereignissen auf (siehe auch 4.7.3). Damit ist die gesamte Bandbreite möglicher Gebietsfeuchteszenarien durch die historischen Daten abgedeckt.

Als Zustandsmerkmal ZM zur Charakterisierung der Gebietsfeuchte wird nach 3.6.2 der Scheiteldurchfluß einer durch den Standardniederschlag ausgelösten Hochwasserwelle verwendet. Dieser Standardniederschlag beträgt 4 mm h^{-1} für 24 Stunden. Unter der Annahme, daß während des Hochwassers im allgemeinen – und im Falle eines Extremhochwassers im speziellen – Verdunstungsprozesse eine vernachlässigbare Rolle spielen, kann ab Ereignisbeginn ein beliebiger Datensatz für

Versuchsgebiet und Erstellung der Trainingsdaten

diese Größen verwendet werden, solange dieser nicht zur Simulation von Schneefall führt.

Gleiches gilt für die Wetterdaten (außer Niederschlag) bei der Berechnung der simulierten Hochwasserszenarien. Die generierten Niederschlagsszenarien müssen die Bedingung erfüllen, die ganze Spanne denkbarer Ereignisse von gewöhnlichen bis hin zu extremen abzudecken. Die Niederschläge fallen gleichmäßig über die jeweilige Teileinzugsgebietsfläche der Zuflußpegel und der Zwischeneinzugsgebietsteile verteilt. Insgesamt wird eine Überzugsgeschwindigkeit der Niederschlagsfront von mindestens 5 Kilometer pro Stunde angenommen und die Niederschlagshöhe für die einzelnen Teilgebiete um flächenbezogen 30 Prozent variiert. (Sollen Netze ausschließlich für die NA-Prozesse in einem Einzugsgebiet erstellt werden, bei dessen Größe die räumliche Variabilität des Niederschlages innerhalb der einzelnen Teilgebiete nicht vernachlässigbar ist, ist dies bei der Generierung der Wetterszenarien zu berücksichtigen.)

Die konkrete Realisierung der Generierung von Niederschlagsszenarien erfolgt mit dem Tool FnetDeveloper von W. Görner (Cullmann, 2007). Dabei werden 200 verschiedene Niederschlagsereignisse bis zum doppelten des hundertjährigen Niederschlages nach KOSTRA (DWD, 1997) mit Variationen der Parameter für Dauer, Verteilung, Schiefe und Rauschen generiert. Durch Kombination von je 15 Niederschlagsszenarien mit den verschiedenen Szenarien der Vorfeuchte ergibt sich eine Datenbasis von 21150 unterschiedlichen Ereignissen. Da die neuronalen Netze für eine flexible Vorhersageweite erstellt werden, ergeben sich für jedes Ereignis je nach Dauer im Mittel 24 Input-Output-Paare für das Training der ANN.

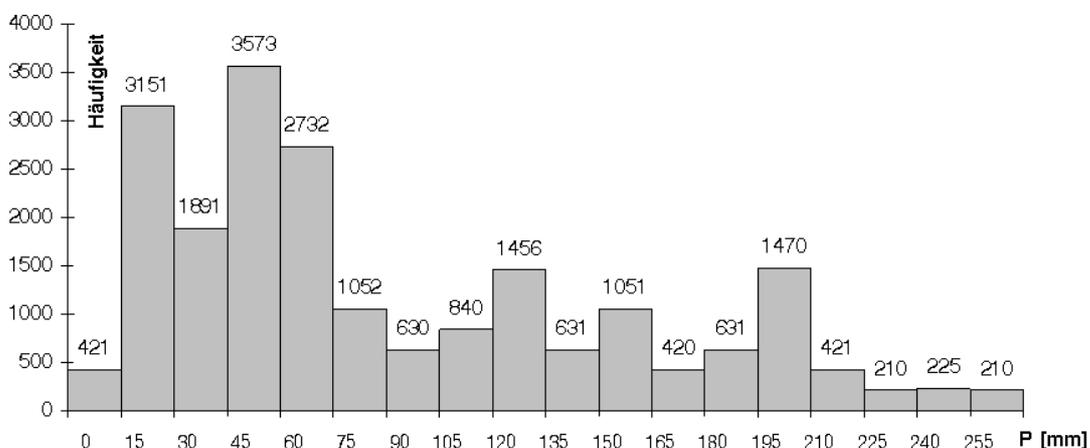


Abbildung 26: Szenarienrechnung, Histogramm für die Niederschlagsmenge

Abbildung 26 zeigt die Verteilung der Niederschlagssummen in den verschiedenen hier verwendeten Niederschlagsszenarien unabhängig von der Niederschlagsdauer. Es ist zu sehen, daß die geringen Niederschläge, die zu keinen extremen Hochwasserereignissen führen, nicht unterrepräsentiert sind, aber die Spanne alle Bereiche bis hin zu sehr extremen Niederschlagsereignissen abdeckt. Zum Vergleich sei erwähnt, daß der maximale hier für das gesamte Einzugsgebiet des Pegels ErlIn generierte Niederschlag den beim bisherigen Rekordhochwasser im August 2002 im am stärksten betroffenen Teilgebiet (Pegel Nossen) leicht übersteigt.

4.6.2 Zuflüsse zu den hydrodynamisch modellierten Flußabschnitten

Die Zuflußganglinien als Randbedingungen für das hydrodynamisch-numerische Modell HEC-RAS sind durch die Berechnung der 21150 Hochwasserszenarien mit WaSim-ETH gegeben (detaillierte Beschreibung in 4.7.2).

4.7 Ergebnisse der physikalisch begründeten Modellierung

In diesem Kapitel sind die mit HEC-RAS und WaSim-ETH gewonnenen Ergebnisse dargestellt. Dabei wird in bezug auf die hydrodynamische Modellierung ausführlich auf die GIS-gestützte Modellerstellung und Kalibrierung und Validierung anhand gemessener Ganglinien eingegangen. Da für die Szenariengenerierung bereits Parametrisierungen des NA-Modells WaSim-ETH nach Cullmann (2007) vorlagen, wird dieser Bereich nur im Hinblick auf geringfügige Veränderungen betrachtet.

4.7.1 Erstellung des hydrodynamisch-numerischen Modelles

Aufbau der Flussgeometrie mit HEC-GeoRAS

Die vorliegenden Daten der Flußgeometrie (Kapitel 4.4) wurden zunächst im Arc-GIS zu einem digitalen Geländemodell im TIN-Format (Triangulated Irregular Network) verschnitten. Mit Hilfe der Extension HEC-GeoRAS wurde daraus ein HEC-RAS-Importfile mit den erforderlichen Geometriedaten erstellt. Dazu mußten die Uferlinien, die Deiche, die Fließwege im Hauptkanal und im rechten und linken Vorland, sowie die Querprofile eingezeichnet werden. Die Uferlinien und Deiche konnten später im Geometrie-Editor des HEC-RAS nachjustiert werden.

Die Festlegung der Fließwege dient der Bestimmung der Abstände zwischen den einzelnen Querprofilen für den Hauptkanal und das rechte beziehungsweise linke

Versuchsgebiet und Erstellung der Trainingsdaten

Vorland. In Flußbiegungen ist dabei der Hauptfließweg ebenso wie die Größe der überfluteten Fläche wasserstandsabhängig. Die Festlegung auf eine bestimmte Linie ist daher stark subjektiv beeinflusst. Jedoch wird im Zweifelsfall der Fließweg in der Nähe des Gerinnes geführt, da in diesen Bereichen der Durchfluß in der Regel am größten sein dürfte.

Die Lage der Querprofile ist durch die Lage der eingemessenen Querschnitte determiniert. Bei der Verlängerung dieser Querprofile durch die Vorländer ist zu beachten, daß der Schnitt an jeder Stelle senkrecht zur Fließrichtung verläuft. Die Breite der Querprofile muß die gesamte möglicherweise überflutete Fläche umfassen. Der Abstand der Profile beträgt 150 Meter. Diese räumliche Auflösung wird als ausreichend angesehen; es werden keine weiteren Profile durch Interpolation erstellt.

Neben der Möglichkeit der GIS-gestützten Kombination verschiedener Geländedaten bietet HEC-GeoRAS die Möglichkeit der Vorabschätzung der Rauigkeitsbeiwerte anhand digitaler Landnutzungsdaten (CORINE land cover 1990). Dabei wurden die in den Flußtälern anzufindenden Landnutzungen mit den bei Preißler und Bollich (1980) und USACE I aufgeführten Rauigkeitsbeiwerten belegt (Tabelle 7).

Tabelle 7: Landnutzungsspezifische Rauigkeitsbeiwerte

Bereich	Corine-Code	Charakteristik/Landnutzung	minimal	normal	maximal
Flußbett	-	sauber, nicht klüftig, gerade	0.025	0.03	0.033
	-	etwas steinig oder bewachsen	0.03	0.035	0.04
Vorland	112	nicht durchgehend städtische Prägung	0.08	0.1	0.15
	211	nicht bewässertes Ackerland	0.03	0.035	0.05
	231	Wiesen und Weiden	0.03	0.035	0.05
	243	Landwirtsch. u. natürl. Bodenbedeckung	0.025	0.04	0.05
	311	Laubwälder	0.08	0.1	0.12
	313	Mischwälder	0.08	0.1	0.15

Wahl der Randbedingungen

Die Randbedingungen, welche die Zuflüsse repräsentieren, ergeben sich aus der Strukturierung des Einzugsgebietes nach Abbildung 23 als Schnittstellen zu den Auslässen der mit WaSim-ETH modellierten Teilgebiete. Der Pegel Kriebstein repräsentiert das größte Teilgebiet und steuert den größten Beitrag am Gesamtabfluß in ErlIn bei. Darüberhinaus stellt er in allen Szenarien bezüglich der Gliederung des Einzugsgebietes einen Zulaufpegel dar. Somit wird an diesem die obere Randbedingung in Form einer Durchflußganglinie positioniert. Zusätzliche Randbedingungen – gleichfalls Durchflußganglinien – erfordern die ebenso

Versuchsgebiet und Erstellung der Trainingsdaten

hydrodynamisch-numerisch modellierten Nebenflüsse Freiburger Mulde am Pegel Nossen und die Striegis am Pegel Niederstriegis. Desweiteren ergeben sich aus den vier Teileinzugsgebieten, die seitlich in die modellierten Fließstrecken entwässern, vier laterale Zuflußganglinien. Diese werden aus Sicht des NA-Modells am betreffenden Teilgebietsauslaß, beziehungsweise im hydrodynamischen Modell jeweils oberhalb der Zusammenflüsse und des Pegels in ErlIn eingebunden.

Als untere Randbedingung wird ein konstantes Reibungsgefälle entsprechend dem Sohlgefälle von 0.0012 in einer Entfernung von 8 km unterhalb des Pegels ErlIn gewählt. Eine Sensitivitätsanalyse zeigt, daß sowohl der Durchfluß als auch der Wasserstand in ErlIn davon nicht beeinflusst werden.

Modelleichung – Anpassung der Rauigkeitsbeiwerte

Die Eichung des hydrodynamischen Modells erfolgt durch Anpassung der Rauigkeitsbeiwerte. Wasserspiegellagenfixierungen für eine stationäre Schichteichung lagen hierbei nicht vor. Doch aufgrund der Vorabschätzung entsprechend der Landnutzung in den Vorländern nach Tabelle 7 (die mittleren Werte wurden angenommen) konnte bei den für die Modelleichung ausgewählten Hochwassern von Dezember 1974 und Juli 1996 bereits eine sehr gute Übereinstimmung der gemessenen und modellierten Ganglinien festgestellt werden.

Nach einer globalen leichten Erhöhung der Rauigkeitsbeiwerte um 0.002 konnten die in Abbildung 27 gezeigten Ergebnisse erzielt werden. Die starke Abweichung für das Hochwasser 1974 um den Hochwasserscheitel ist eindeutig auf fehlerhafte Wasserstands-Abfluß-Beziehungen zurückzuführen, da durch das Modell die Massenbilanz gewahrt bleibt und hier deutlich im Extrapolationsbereich der Schlüsselkurven eine systematische Abweichung feststellbar ist. Die Scheiteleintrittszeit und die Bereiche niedrigerer Durchflüsse werden hier hervorragend getroffen. Das Hochwasser 1996, welches eher klein ist, aber eine ausgeprägte Dynamik aufweist, wird vom Modell in jeder Hinsicht hervorragend abgebildet.

Da die Eichparameter des Modells nach einer physikalisch fundierten Vorabschätzung nur marginal verändert wurden, kann davon ausgegangen werden, daß das Modell die Gegebenheiten in der Natur sehr gut widerspiegelt.

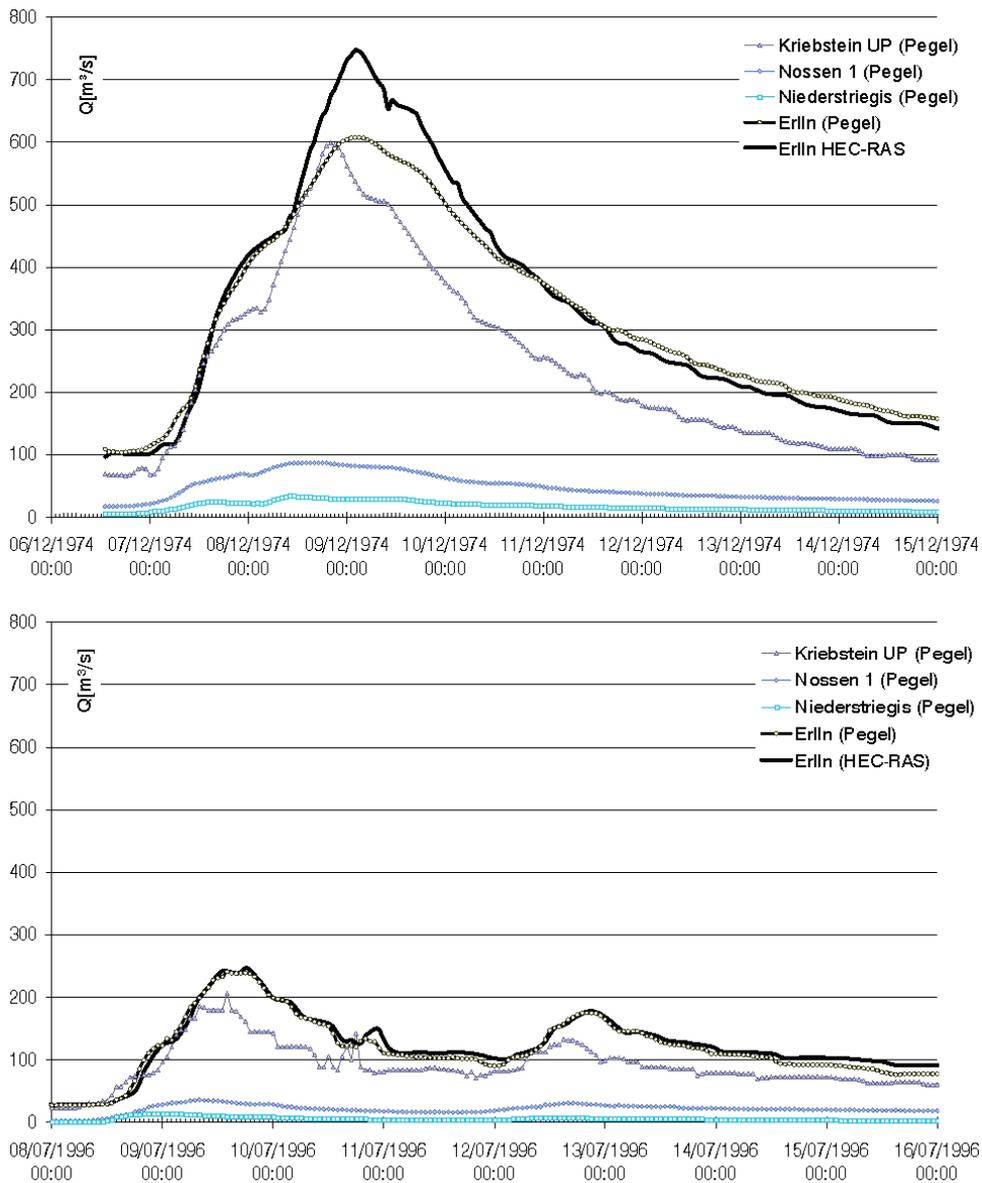


Abbildung 27: Hochwasserereignisse Dezember 1974 und Juli 1996, Eichung

Modellvalidierung

Für die Validierung des hydrodynamischen Modells wurden die übrigen fünf Hochwasserereignisse mit verfügbaren Daten für alle beteiligten Pegelstationen verwendet. Abbildung 28 zeigt die Hochwasserereignisse von August 1983 und März 1998, die übrigen finden sich in Anhang C.

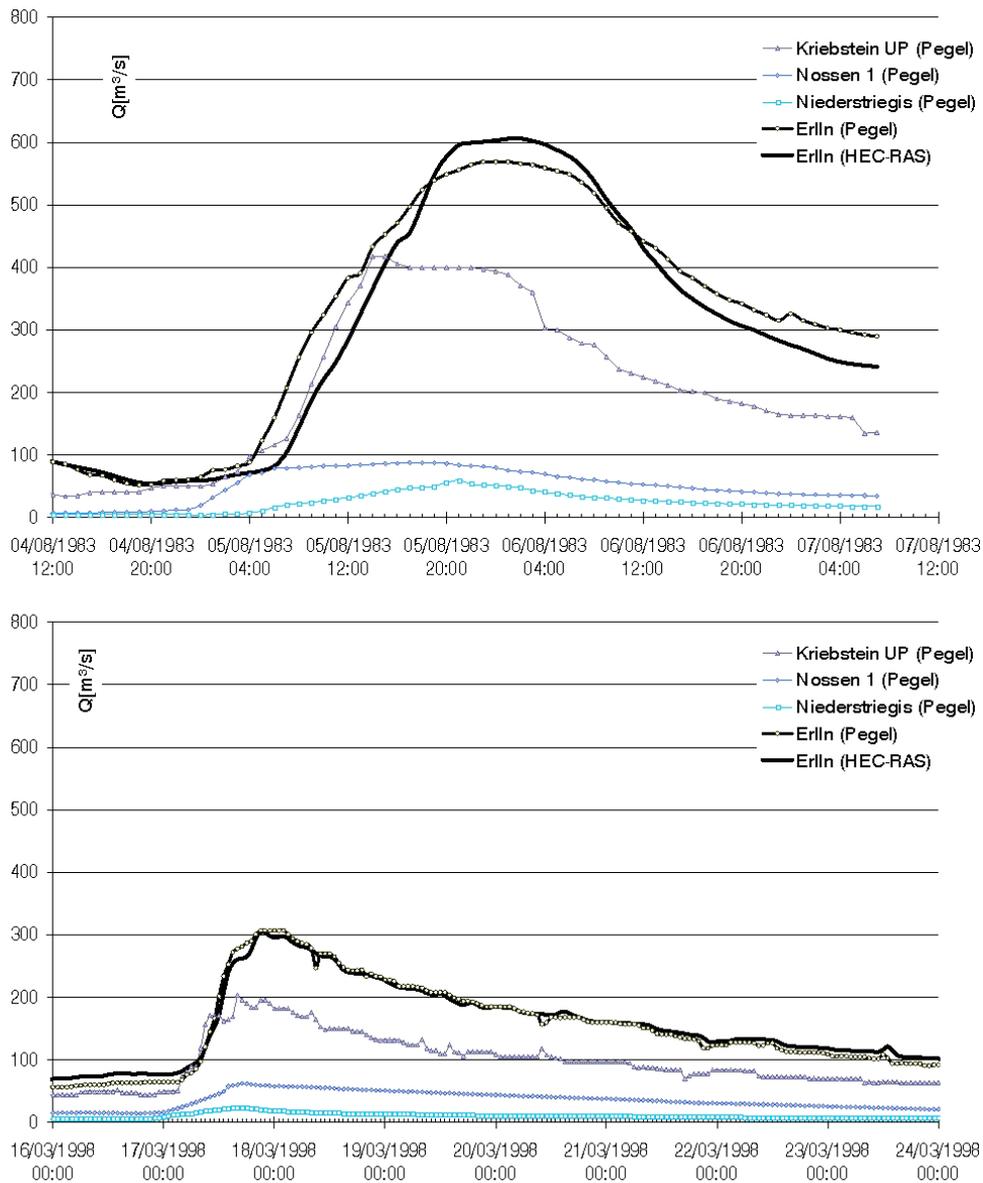


Abbildung 28: Hochwasserereignisse August 1983 und März 1998, Validierung

Die Verzögerung im Anstieg für das Hochwasser 1983 ist mit Fehlern des Modells nicht erklärbar, da der steile Anstieg der Welle am Pegel Erlin bereits vor dem Anstieg der Welle in Kriebstein beginnt. Wenn man eine mögliche zeitliche Verschiebung der gemessenen gegenüber der modellierten Welle berücksichtigt, treten, wie auch beim Hochwasser 1974 zu beobachteten war, für gleiche Durchflüsse sehr ähnliche Fehler auf. Der Grund für diese Abweichungen wird auch hier in den nicht hochwassertauglichen Wasserstands-Durchfluß-Beziehungen gesehen. Darauf lassen auch die im Anhang dargestellten Hochwasserwellen schließen, bei denen allein die Differenz in der Fülle der modellierten und gemessenen Hochwasserwellen auf

fehlerbehaftete Schlüsselkurven deutet. Demgegenüber ist das kleine Hochwasser von 1998 (Abbildung 28) hervorragend getroffen. Es kann konstatiert werden, daß das hydrodynamische Modell das System sehr gut abbildet. Dabei kann das Modell aufgrund der berücksichtigten Massenerhaltung – im Gegensatz zu den Pegeldaten – zu den Zuflußganglinien konsistente Wellen für den Zielpegel berechnen, die unter der Annahme korrekter Zuflußganglinien die Realität sogar besser widerspiegeln.

4.7.2 Niederschlags-Abfluß-Modellierung mit WaSim-ETH

Für WaSim-ETH lagen Parametrisierungen für die Einzugsgebiete der Pegel Kriebstein UP, Nossen 1 und Niederstriegis nach Cullmann (2007) vor. Die Parametersätze für das Einzugsgebiet der Striegis wurden auf das Zwischeneinzugsgebiet unterhalb der genannten Pegel übertragen, da deren Naturräume die größte Ähnlichkeit aufweisen. Beide haben keinen Anteil an den Kammlagen des Erzgebirges.

Im Detail soll auf die Parametrisierungen hier nur insofern eingegangen werden, daß für jedes Einzugsgebiet je zwei Parametersätze verwendet werden, die sich in ihrem Gültigkeitsbereich in bezug auf die Abflußmenge unterscheiden. Cullmann (2007) verwendet eine sigmoide Wichtungsfunktion nach den Gleichungen (11) und (12) zur Berechnung des Abflusses aus den beiden unterschiedlichen Modellergebnissen, um einen abrupten Übergang der simulierten Ganglinien zu vermeiden. Der Wichtungsfaktor w berechnet sich dabei aus dem Abfluß Q und den Parametern t_h und τ . Der Parameter t_h verringert mit zunehmender Größe den maximalen Anstieg der Sigmoiden und τ steht für den Grenzwert der Gültigkeitsbereiche der beiden Parametersätze von 1.8 mm h^{-1} . Der resultierende Abflußwert Q_w ergibt sich mit dem Wichtungsfaktor w aus den Modellergebnissen mit der Parametrisierung für extreme Ereignisse Q_g und der normalen Parametrisierung Q_k .

$$w = \frac{1}{1 + e^{\frac{t_h - Q}{\tau}}} \quad (11)$$

$$Q_w = w \cdot Q_g + (1 - w) \cdot Q_k \quad (12)$$

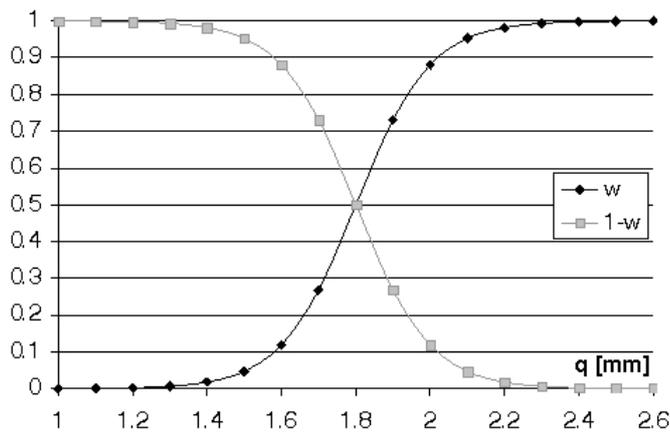


Abbildung 29: sigmoide Wichtung der Abflüsse für die verschiedenen Modellparametrisierungen

Abweichend von Cullmann (2007) wurde eine steilere Funktion gewählt, da Tests mit den verwendeten neuronalen Netzen zeigten, daß diese sich auch unter Verwendung der steileren Funktion problemlos anpassen lassen. Der Parameter t_h wurde entsprechend statt 0.9 mm h^{-1} auf 0.1 mm h^{-1} gesetzt. Abbildung 29 zeigt die Wichtungsfaktoren w und $1-w$ in Abhängigkeit des Abflusses. Alle mit dem Modell berechneten Abflußwerte, eingeschlossen das Zustandsmerkmal, resultieren aus der Kombination der beiden Modellparametrisierungen mit der Sigmoiden.

4.7.3 Berechnung der Hochwasserszenarien

Die Gebietszustände bei Ereignisbeginn und das Zustandsmerkmal

Für die Hochwasserszenarien wurden unterschiedliche Systemzustände als Basis benötigt. Die gemessene Geschichte der Wetterdaten in dem relativ großen Zeitraum von 1953 bis 1999 bildete die Grundlage, auf der diese Systemzustände mit dem kalibrierten NA-Modell simuliert wurden. Der organisatorische Aufwand bei der Speicherung sämtlicher relevanter Zustandsdaten, die das NA-Modell benötigt, um darauf aufbauend verschiedene Niederschlagsszenarien zu verarbeiten, limitierte die Anzahl dieser Zustandsszenarien. Desweiteren erforderte die Methodik der Berechnung des Zustandsmerkmals mit dem NA-Modell schneefreie Verhältnisse im Einzugsgebiet bei der Abspeicherung des Systemzustandes, da ansonsten der Effektivniederschlag durch Schmelzwasser vom zugeführten Einheitsniederschlag abweichen würde und die Systemantworten nicht mehr vergleichbar wären (siehe 3.6).

Versuchsgebiet und Erstellung der Trainingsdaten

Für solche Szenarien ist das Zustandsmerkmal mit dem NA-Modell also unbestimmbar.

Aus diesen beiden Einschränkungen ergab sich die Wahl der Zeitpunkte, an denen der Zustand des Einzugsgebietes abgespeichert wurde. In jedem simulierten Jahr wurden 30 Zeitpunkte festgelegt, was zu einer Basis von 1410 Vorfeuchteszenarien führte. Es wurden jeweils der 7., 14., 21. und 28. Tag des Monats zwischen dem 28. 3. und 28. 10. jedes Jahres und zusätzlich der Jahreswechsel ausgewählt. Es wurde für jeden Zeitpunkt getestet, ob tatsächlich Schneefreiheit gegeben war und das Szenario gegebenenfalls von der Berechnung des Zustandsmerkmals ausgeschlossen. Es ist davon auszugehen, daß mit dieser Vorgehensweise die Bandbreite der möglichen Systemzustände ausreichend abgedeckt wird. Unter diesen Szenarien sind sehr feuchte, hochwasserfördernde Gegebenheiten zu finden, da die natürlichen Daten bereits diverse stärkere Hochwasserereignisse enthalten und diese hier ausschließlich als Vorgeschichte simulierter Hochwasserereignisse verwendet wurden.

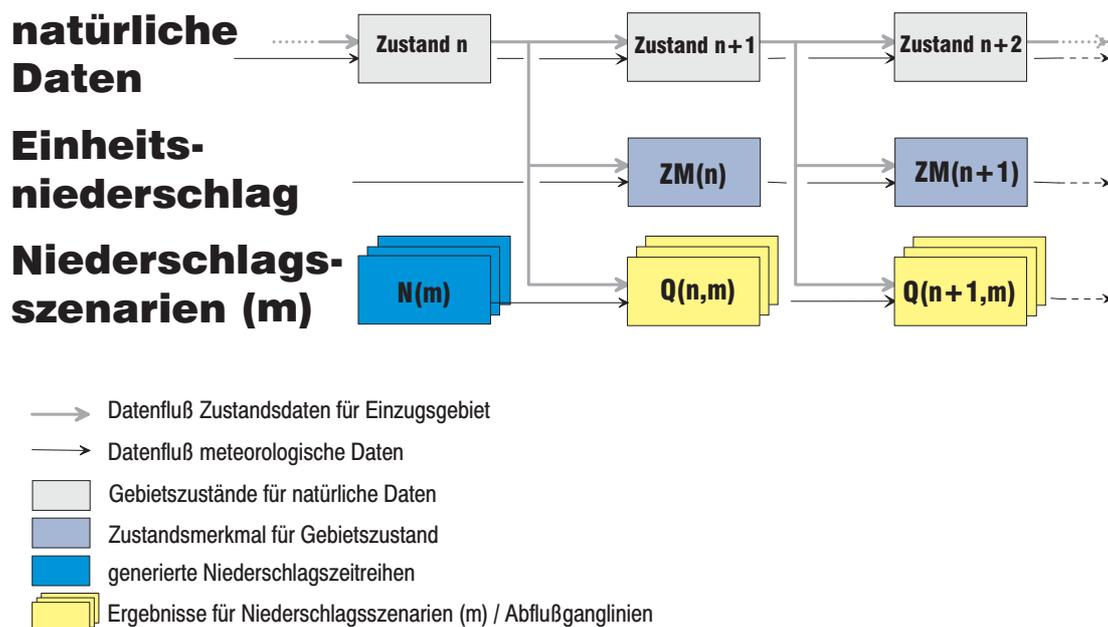


Abbildung 30: Berechnung der Zustandsmerkmale und Hochwasserszenarien mit WaSim-ETH

Die Simulation wurde technisch folgendermaßen organisiert: Die Wasserhaushalts-simulation mit dem Modell WaSim-ETH wurde an den definierten Zeitpunkten gestoppt und die wesentlichen Zustandsdateien für die mehrfache Weiterverarbeitung separat gespeichert. Aufbauend auf diesen Systemzuständen wurde einerseits mit den natürlichen Daten zum nächsten Termin weitergerechnet. Andererseits wurde das

Versuchsgebiet und Erstellung der Trainingsdaten

Zustandsmerkmal als Hochwasserscheiteldurchfluß der Gebietsreaktion auf den Standardniederschlag und die einzelnen Hochwasserszenarien simuliert (Abbildung 30).

Die Abbildung 31 stellt die Häufigkeit des Zustandsmerkmals in der Trainingsdatenbank dar. (Jeder Wert erscheint in der Abbildung jeweils 15fach, da jeder Gebietszustand mit 15 Niederschlagsszenarien kombiniert wurde.)

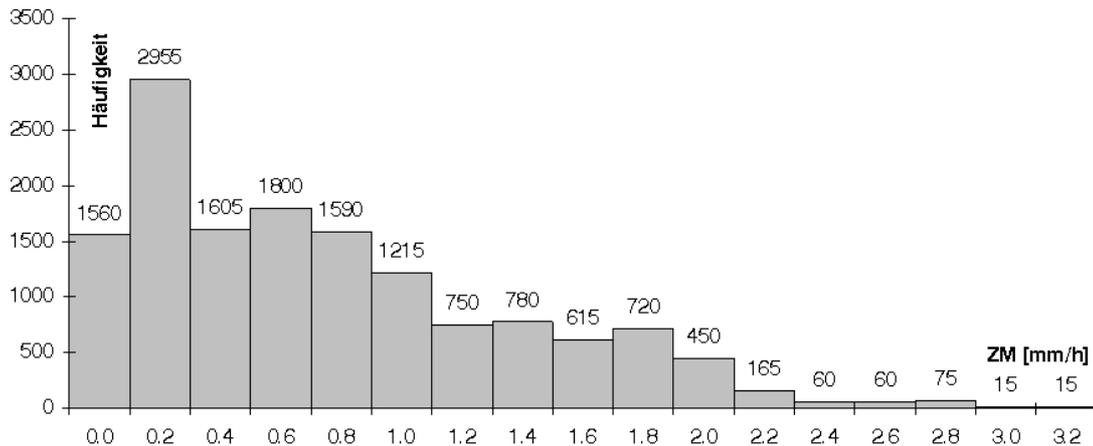


Abbildung 31: Histogramm für das Zustandsmerkmal

Um das bisher größte gemessene Hochwasserereignis im August 2002 in diesen Zusammenhang zu stellen, sei an dieser Stelle bemerkt, daß sich für den Gebietszustand vor diesem Hochwasserereignis ein Zustandsmerkmal von $0,38 \text{ mm h}^{-1}$ ergibt ($>0,2$, zweite Feuchtekategorie in Abbildung 31). Der Scheitel der Gebietsreaktion auf den Standardniederschlag (94 mm in 24 h) ist also – im Vergleich aller auf historischen Daten beruhenden Werte – für den August 2002 relativ klein. Das bedeutet, daß der Scheiteldurchfluß der realen Hochwasserwelle ebenfalls – gemessen am gefallenem Niederschlag – relativ gering ausgefallen ist.

Die Hochwasserszenarien

Abbildung 32 stellt die Ergebnisse der Simulationen in Form der Verteilung der Hochwasserscheiteldurchflüsse exemplarisch für den Pegel Nossen 1 dar. Aus der Kombination der Niederschlagsszenarien (Abbildung 26) mit den realistischen Gebietszuständen (Abbildung 31) resultiert eine breite Spanne von Hochwasserereignissen, die bis zu einem Maximum von über 13 mm h^{-1} reichen. Zum Vergleich sei das Hochwasser im August 2002 genannt, welches bei einer ungefähren

Versuchsgebiet und Erstellung der Trainingsdaten

Niederschlagssumme von 250 mm (vgl. Abbildung 26) am Pegel Nossen 1 einen Scheitelabfluß von 4.3 mm h^{-1} (entspricht $690 \text{ m}^3 \text{ s}^{-1}$, nach DGL, 2003) ergab.

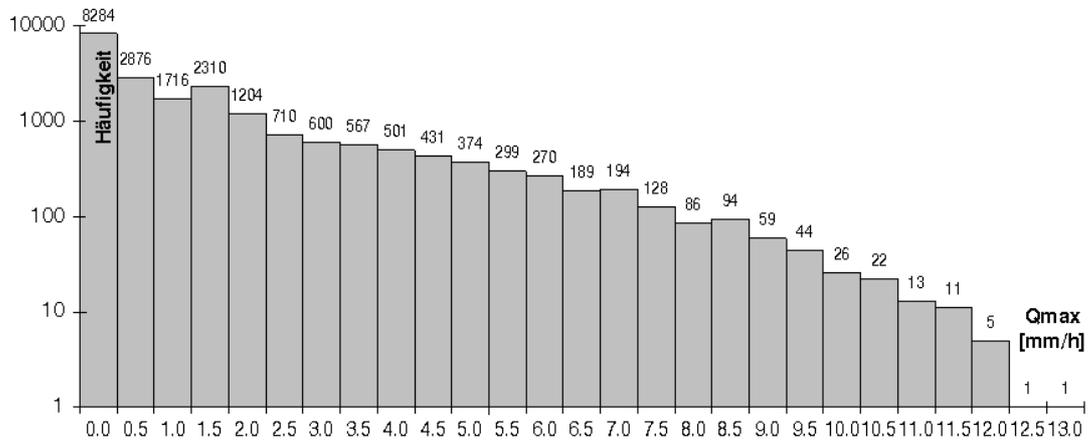


Abbildung 32: Histogramm für den Scheiteldurchfluß der Hochwasserszenarien am Pegel Nossen, logarithmische Darstellung

5 Ergebnisse der Prognose-ANN

5.1 Prognose-ANN ohne Merkmale für NA-Prozesse im Zwischeneinzugsgebiet

5.1.1 MLFN für die HD-Szenarien

Die Merkmalsselektion für die hydrodynamischen Prozesse wurde entsprechend der unter 3.5 beschriebenen Methodik für die drei Pegel Kriebstein UP, Nossen 1 und Niederstriegis separat durchgeführt. Dabei wurden die Impulse jeweils am aufsteigenden Ast, auf dem Scheitel und am abfallenden Ast der Hochwasserganglinien von 1974 und 1998 plaziert. Daraus resultierten die in Tabelle 8 dargestellten Einflußspannen für die jeweiligen Pegel. Für die Szenarien HD10, HD20 und HD40 ergeben sich daraus entsprechend der berücksichtigten Pegel 13, 11 beziehungsweise 7 Merkmale.

Tabelle 8: Einflußspannen für die Pegel Kriebstein UP, Nossen 1 und Niederstriegis

	Kriebstein UP	Nossen 1	Niederstriegis
Fläche des EZG [km ²]	1757	585	283
Fließstrecke bis Pegel ErlIn [km]	36.9	45.7	35.6
Δz [m]	49.8	70.7	48.4
t_{min} [h]	<u>3</u>	<u>6</u>	<u>6</u>
t_{max} [h]	<u>9</u>	<u>9</u>	<u>7</u>
Anzahl der Merkmale [-]	7	4	2

Für diese Szenarien wurden MLFN mit bis zu 25 verdeckten Neuronen mit je fünf verschiedenen Initialisierungen der Gewichte für die beiden Outputgrößen Wasserstand und Durchfluß am Pegel ErlIn trainiert. In Abbildung 33 sind die Trainingsergebnisse für die besten der jeweils fünf Realisierungen dargestellt. Erwartungsgemäß sind die Anpassungen generell besser, je kleiner der Anteil des als unbeobachtet angenommenen Zwischeneinzugsgebietes gewählt wird. Ebenso geht tendentiell für größere Zwischengebiete die Schere zwischen der Performance der Trainings- und der Testdaten weiter auseinander. Der Grund hierfür besteht klarerweise in der größeren Information, die dem Netz bei der Kenntnis aller drei oberstromigen Pegel zur Verfügung gestellt wird. Desweiteren verwenden die Netze für größere Zwischengebiete kleinere Inputvektoren, welche tendentiell mit weniger verdeckten Neuronen optimale Anpassungen in bezug zur Testperformance zeigen.

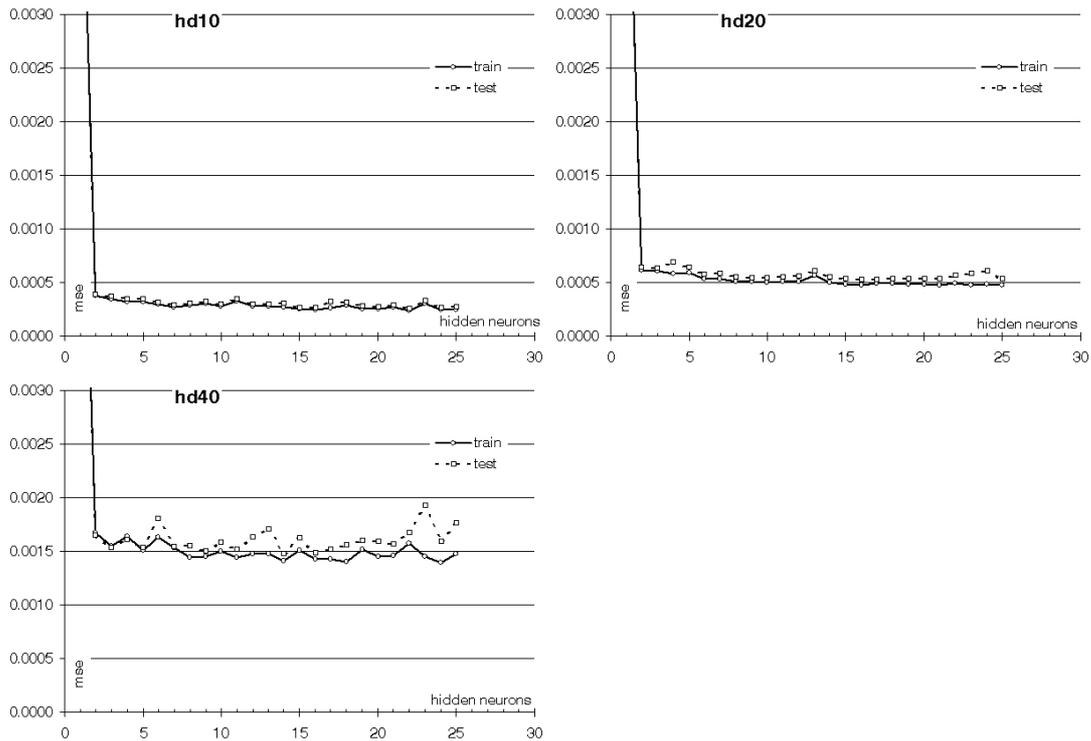


Abbildung 33: : Performance der MLFN für die Szenarien HD10, HD20 und HD40 in Abhängigkeit von der Anzahl der verdeckten Neuronen, mittlerer quadratischer Fehler für normierte Daten

Folgerichtig wurden die besten Anpassungen hinsichtlich der Testdaten für 22, 16 und 14 verdeckte Neuronen (HD10, HD20 und HD40) erzielt (Abbildung 34).

In Abbildung 34 wird weiterhin sichtbar, daß die Anpassung für den Durchfluß leicht besser ist, als für den Wasserstand. Die größten Abweichungen treten dabei für geringere Wasserstände auf. Daß sich Fehler in diesem Bereich stärker auf den Wasserstand als auf den Durchfluß auswirken, ist insofern nicht verwunderlich, da der Anstieg der Schlüsselkurve allgemein mit steigendem Durchfluß abnimmt, und damit die Sensitivität des Wasserstandes auf den Durchfluß im unteren Bereich am größten ist. Abbildung 35 zeigt einen Ausschnitt aus der Testreihe. Dabei wird offenbar, worin die Ursache für diese Abweichungen liegt. Am jeweiligen Anfang einer Hochwasserwelle tritt meist eine Unterschätzung der Targetwerte auf. Die Zuflußganglinien, die den Inputvektor bestimmen, reagieren aufgrund der Einzugsgebietsgröße und der größeren mittleren Fließzeit erst später in vollem Umfang auf den hochwasserauslösenden Niederschlag als das Zwischeneinzugsgebiet, welches hier nicht im Merkmalsvektor repräsentiert ist. Das führt zu einem Anstieg am Zielpegel, der mit den verwendeten Merkmalen nicht erklärt werden kann. Dieser Effekt ist desto größer, je größer die Fläche des Zwischeneinzugsgebietes ist. Die

Ergebnisse der Prognose-ANN

Verwendung von Merkmalen für die NA-Prozesse im Zwischeneinzugsgebiet verspricht auch in dieser Hinsicht eine bessere Bestimmbarkeit der Zielgrößen.

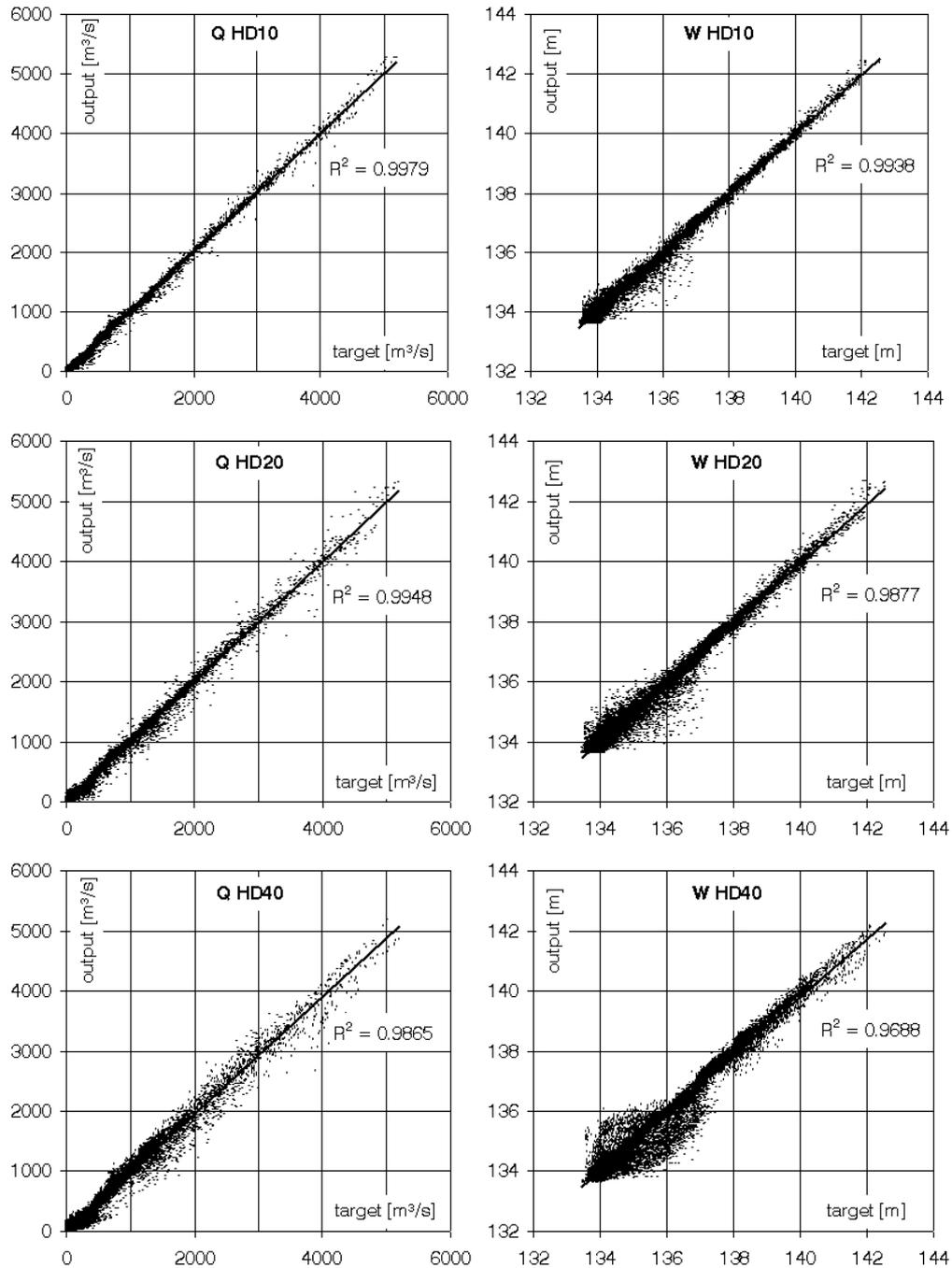


Abbildung 34: Testdaten für die HD-Netze: HD10 mit 22, HD20 mit 16 und HD40 mit 14 verdeckten Neuronen, Durchfluß und Wasserstand

Dieses Phänomen beeinträchtigt die Prognosefähigkeit des Netzes jedoch nur unwesentlich, da die Hochwasserscheitel davon nicht beeinflusst werden.

Ergebnisse der Prognose-ANN

Grundsätzlich ist in Abbildung 35 aber auch erkennbar, daß die Anpassung in diesem Ausschnitt (wie auch in den nichtabgebildeten) insbesondere für die Szenarien HD10 und HD20 recht gute Ergebnisse zeigt. Scheitelertrittszeit und -durchfluß, vor allem aber der Scheitelwasserstand werden sehr gut wiedergegeben.

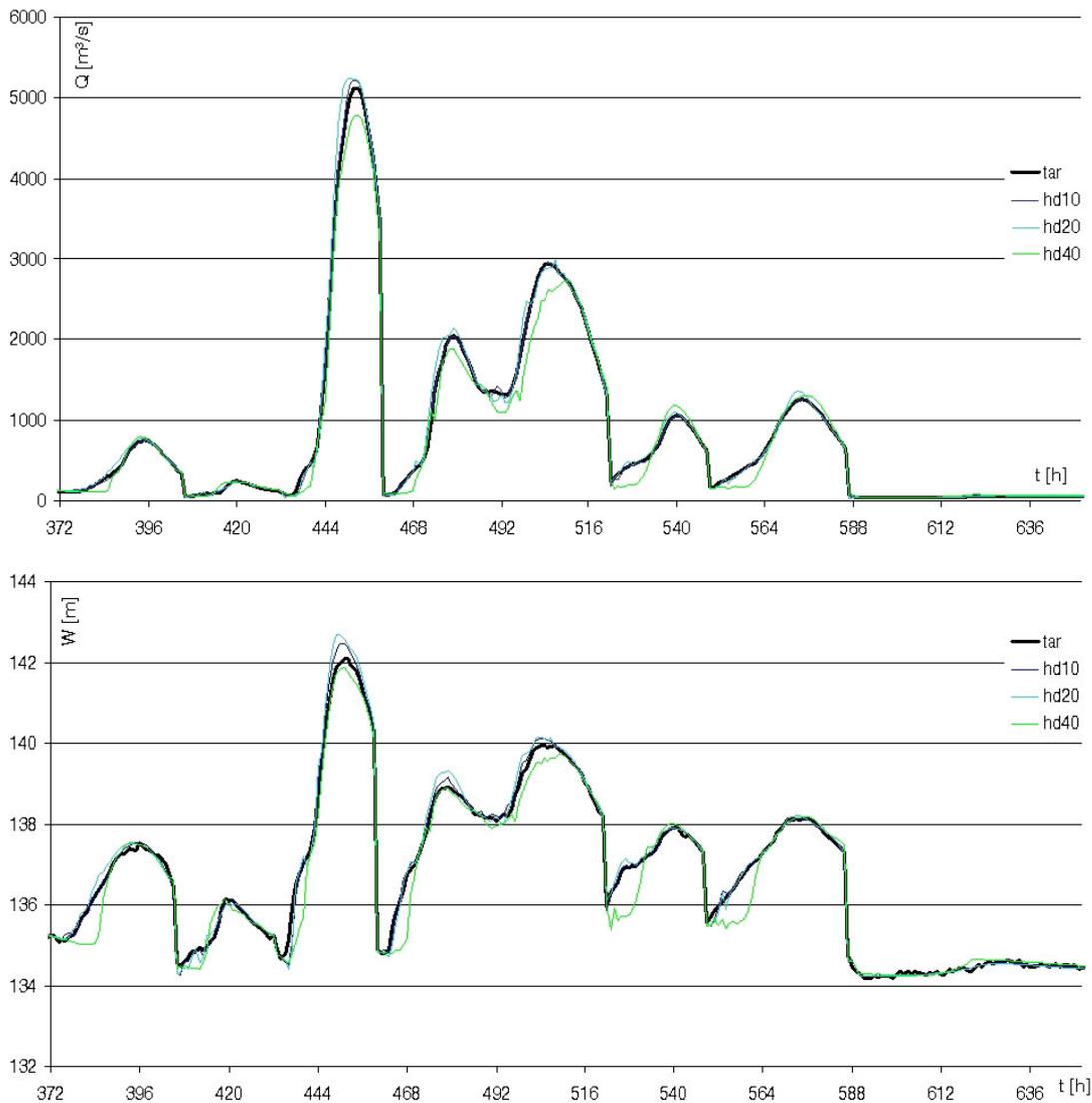


Abbildung 35: HD-MLFN - Darstellung einzelner Szenarien für Durchfluß und Wasserstand

5.1.2 SOM für die HD-Szenarien

Die gleichen Merkmalsvektoren wie für die Erstellung der MLFN, erweitert um den Durchfluß (beziehungsweise Wasserstand), wurden für das Training der selbstorganisierenden Merkmalskarten verwendet. Es wurden hier für Wasserstand und Durchfluß unterschiedliche Netze trainiert. Der Merkmalsvektor für eine SOM

Ergebnisse der Prognose-ANN

besitzt demnach jeweils eine Dimension mehr, als der Inputvektor der MLFN, also 14 (HD10), 12 (HD20) oder 8 Dimensionen (HD40). Um die Bedeutung des Zielwertes unter diesen beim Training der SOM gleichberechtigten Größen herauszuheben, wurde die Verwendung eines Outputdimensionengewichtes *ODW* getestet. Die Ergebnisse sind in Abbildung 36 dargestellt. Dabei wird deutlich, daß die Anpassungen für *ODW* = 1 am besten sind, und die Einführung des Outputdimensionengewichtes damit nicht die erhoffte Verbesserung gebracht hat.

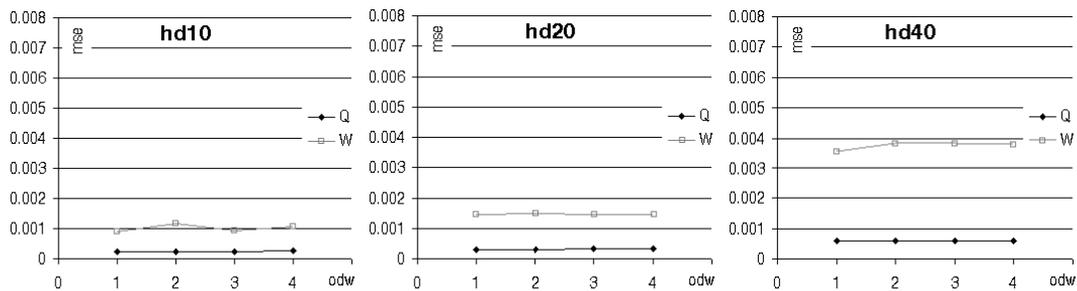


Abbildung 36: Performance der SOM (150 x 40) für die Szenarien HD10, HD20 und HD40 in Abhängigkeit vom Outputdimensionengewicht, mittlerer quadratischer Fehler für normierte Daten.

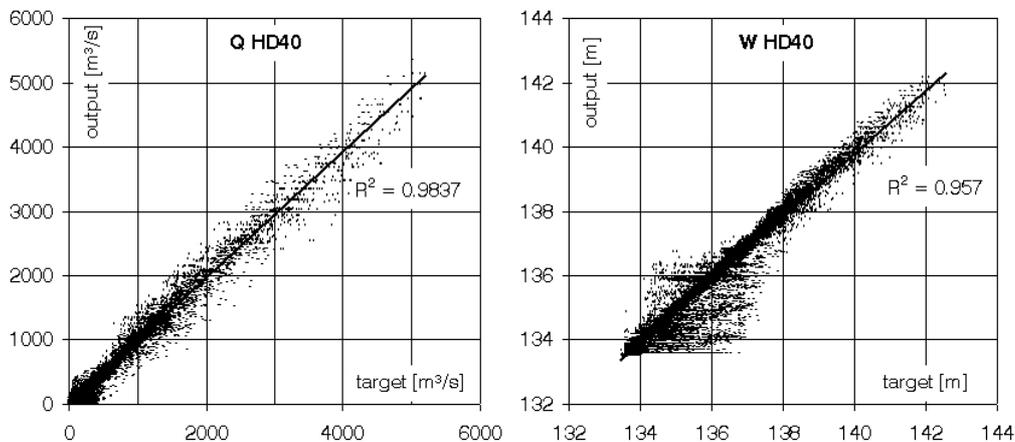


Abbildung 37: Testdaten für die HD-Netze: HD10 mit 22, HD20 mit 16 und HD40 mit 14 verdeckten Neuronen, Durchfluß und Wasserstand

Darüberhinaus ist auch hier die Anpassung für den Durchfluß offenbar besser als für den Wasserstand, jedoch auch hier in erster Linie im Bereich niedriger Wasserstände. Analog zu Abbildung 34 ist in Abbildung 37 die Korrelation der Netzoutputs mit den Targetwerten für das Szenario HD40 dargestellt. (Eine komplette Darstellung aller Szenarien findet sich in Anhang D.) Abgesehen von einer leicht schlechteren

Ergebnisse der Prognose-ANN

Performance der SOM, ist hier der gleiche Effekt wie bei den MLFN zu beobachten: Besonders bei niedrigen Wasserständen treten die größten Abweichungen auf. Die Ursache dafür wurde bereits oben diskutiert, das Auftreten des gleichen Effektes bei der Verwendung dieses gänzlich verschiedenen Netztypes ist lediglich ein weiteres Indiz für die Richtigkeit dieser Aussage.

Abbildung 38 zeigt die gleiche Sequenz von generierten Hochwasserereignissen für die SOM wie Abbildung 35 für die MLFN. Prinzipiell zeigt sich auch hier eine Eignung der trainierten Netze insbesondere für die Szenarien HD10 und HD20. Die Dynamik der Wellen ist ähnlich gut getroffen, nur sind die Ganglinien nicht so glatt, wie die mit dem Modell oder den MLFN errechneten.

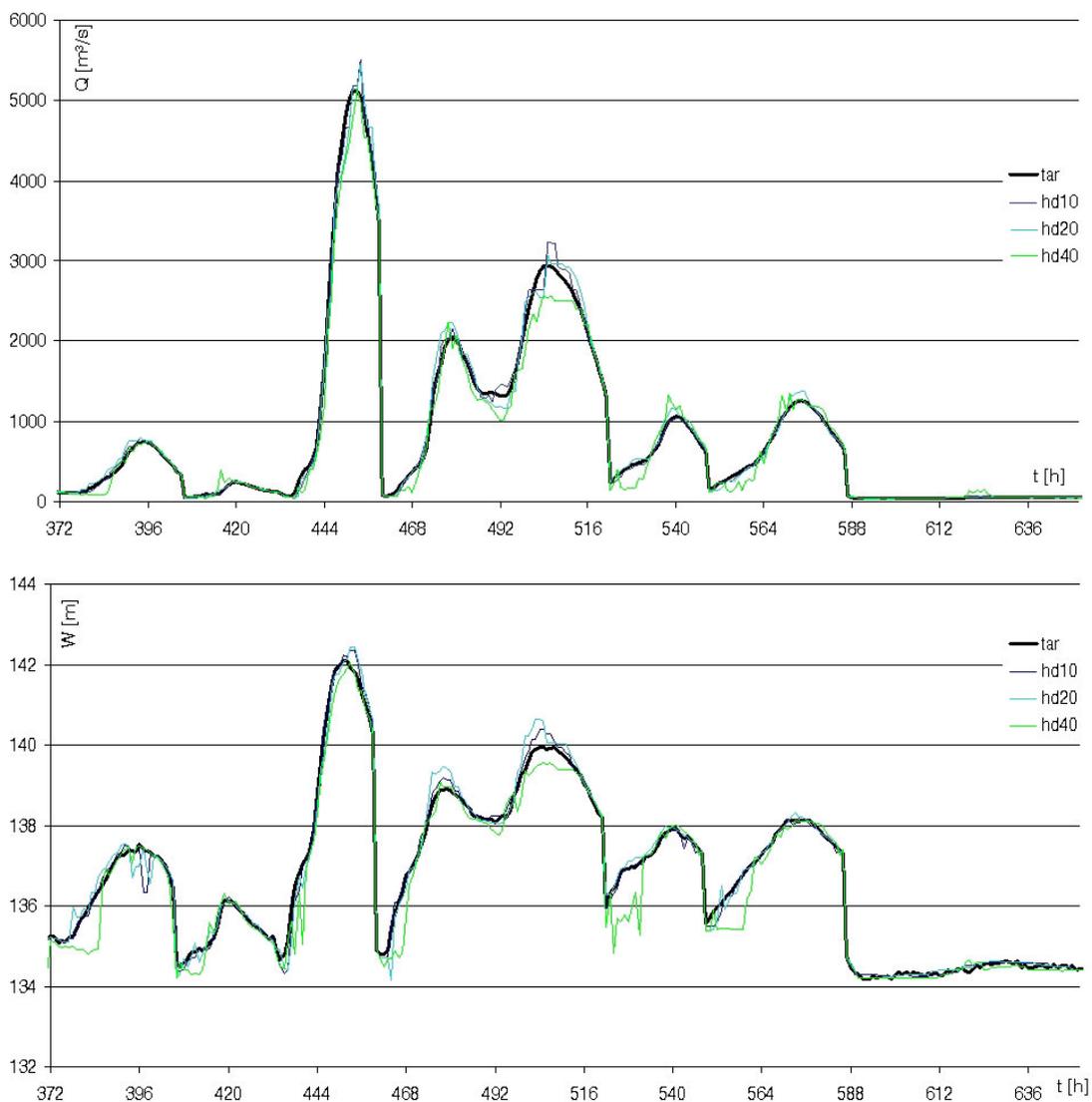


Abbildung 38: HD-SOM - Darstellung einzelner Szenarien für Durchfluß und Wasserstand

Ergebnisse der Prognose-ANN

In Abbildung 39 sind drei Ansichten der SOM für das Szenario HD20 dargestellt. Die Karten bilden die Lage des Netzes im Merkmalsraum ab, wobei jede Darstellung eine Dimension darstellt. Analog zur geodätischen Höhe auf Landkarten sind hohe Werte orange bis rot dargestellt, blaue Flächen bedeuten kleine Werte in der dargestellten Dimension. Ein Pixel steht für ein Neuron. Dessen Koordinaten ergeben sich aus den Einfärbungen in den verschiedenen Ansichten der Karte.

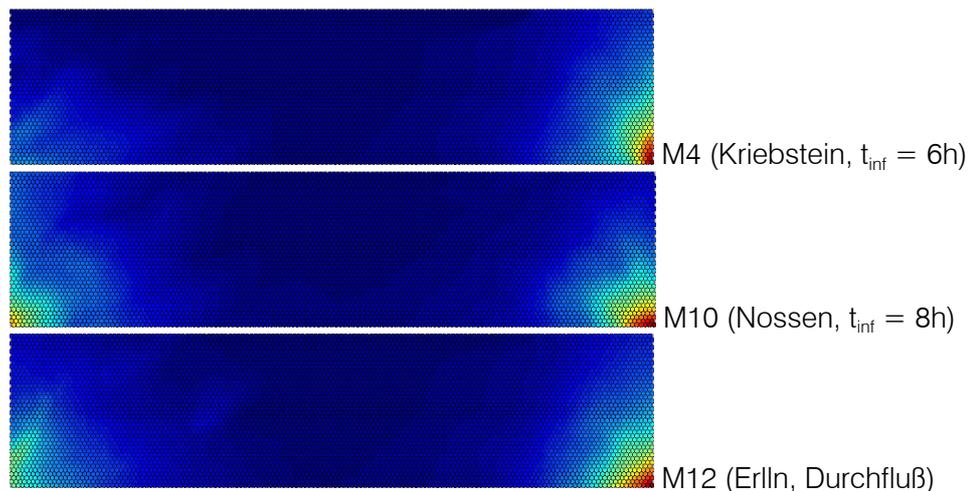


Abbildung 39: SOM für Durchfluß, Szenario HD20, Ansicht für die Merkmale M4, M10 und M12, braun: maximal, blau: minimal

Die Merkmale M4 und M10 entsprechen jeweils einem zentral in der Einflußspanne gelegenen Merkmal für jeden der zwei Zuflußpegel Kriebstein UP und Nossen 1. Die Einflußzeiten für diese Merkmale liegen also annähernd bei der Scheitellaufzeit. Es ist gut zu erkennen, daß die Maxima erwartungsgemäß in der gleichen Region der Karte auftreten. Darüberhinaus ist ein zweites lokales Maximum auf der linken Seite ausgeprägt, welches Situationen klassifiziert, die in Nossen höhere Durchflüsse in das System einspeist. Im Merkmal für Kriebstein ist dieses lokale Maximum allerdings eher schwach ausgeprägt. Folgerichtig zeigt sich in der Ansicht für das Merkmal M12 für den Durchfluß in ErlIn auch ein lokales Maximum, welches bedingt durch die relativ kleine Einzugsgebietsfläche des Pegels Nossen 1 allerdings sehr schwach ausfällt. Eine komplette Ansicht dieser SOM in allen Dimensionen ist in Anhang E eingefügt. Dabei ist – speziell für die Merkmale des dominierenden Pegels Kriebstein – sehr gut zu erkennen, wie zwar alle Merkmale auch ein deutliches Maximum an der rechten unteren Ecke der Karte besitzen, was die Relevanz dieser Merkmale unterstreicht, aber sich auch für die am Rande der Einflußspanne befindlichen Merkmale Bereiche hoher

Werte ausgebildet haben, die nicht mehr mit dem Maximum für ErlIn übereinstimmen. Das weist darauf hin, daß die Wahl der Parameter $t_{\text{inf,max}}$ und $t_{\text{inf,min}}$ für die Einflußspanne von den Daten bestätigt wird.

Die SOM können nicht die sehr gute Performance der MLFN erreichen, jedoch in noch zufriedenstellender Art und Weise die Beziehung zwischen den Merkmalen abbilden. Aus der reinen Klassifikation der Merkmalsvektoren ergibt sich jedoch, trotz der Unzulänglichkeiten in bezug zur Prognosegüte, eine interpretierbare Visualisierung des Merkmalsraumes.

5.2 Prognose-ANN mit Merkmalen für NA-Prozesse im Zwischeneinzugsgebiet

5.2.1 Das Zustandsmerkmal

Nach der unter 3.6 beschriebenen Methodik wird die den Gebietszustand bei Ereignisbeginn betreffende Vorgeschichte in einem Zustandsmerkmal erfaßt, um den gesamten Umfang der Merkmale im eigentlichen Prognosenetz begrenzt zu halten. Um die Fülle der Einflußfaktoren auf den Gebietszustand entsprechend des Datenflusses im Modell zu reduzieren, wurde als erster Schritt das Subnetz 1 eingeführt, welches alle meteorologischen Eingangsgrößen mit Ausnahme des Niederschlags zu einer Größe auf Tageswertbasis zusammenfaßt: der potentiellen Evapotranspiration.

Subnetz1

Das Subnetz 1 faßt die Größen Temperatur, Windgeschwindigkeit, relative Luftfeuchte, Globalstrahlung und Tag des Jahres (doy) zur potentiellen Verdunstung zusammen. Dazu wurden die Inputgrößen und die Targetgröße zu Tagesmitteln zusammengefaßt. Kalibrierungs-, Validierungs- und Testdatensatz wurden je 5000, 1575 und 5478 Input-Output-Paare zugeteilt. Als Netztyp wurde ein MLFN mit einer verdeckten Schicht gewählt. Die Anzahl der Neuronen in der verdeckten Schicht wurde beginnend mit zwei solange erhöht, bis scheinbar keine substantielle Verbesserung mehr zu erwarten war. Dabei wurde jede Variante fünf mal mit unterschiedlichen Anfangsgewichten initialisiert. Die mittleren quadratischen Abweichungen der besten Parametrisierungen der jeweiligen Netze sind in Abbildung 40 für die beim Training nicht verwendeten Testdaten dargestellt (Daten auf das Intervall von -1 bis 1 normiert).

Ergebnisse der Prognose-ANN

Die Variante mit 11 verdeckten Neuronen (grau gekennzeichnet) wurde als beste befunden.

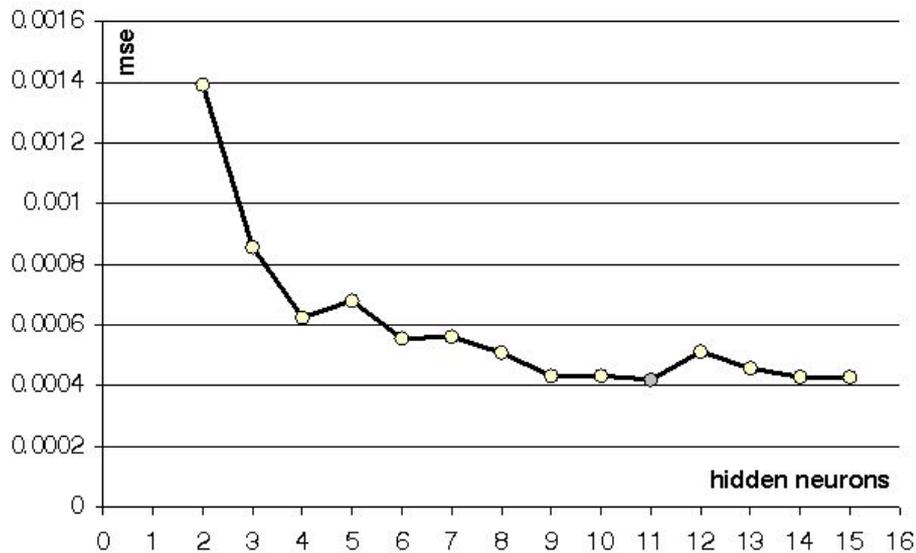


Abbildung 40: Performance des Subnetzes 1 in Abhängigkeit von der Anzahl der Neuronen in der verdeckten Schicht für die Testdaten, mittlerer quadratischer Fehler für normierte Daten.

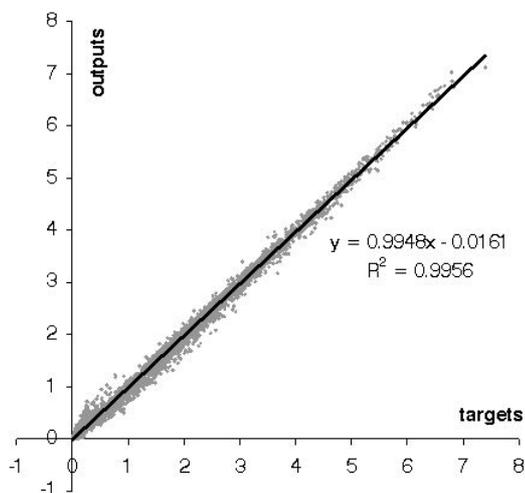


Abbildung 41: Testdaten für Subnetz1, Variante mit 11 Neuronen

Abbildung 41 zeigt die Korrelation der Netzoutputs mit den Targets für dieses Netz. Es kann festgestellt werden, daß die Zusammenfassung des Informationsflusses auf der Ebene der potentiellen Verdunstung sehr zufriedenstellend funktioniert. Im deterministischen Modell werden Daten verwendet, die im Raster von einem

Ergebnisse der Prognose-ANN

Quadratkilometer sich beispielsweise schon durch unterschiedliche Höhenwerte, Landnutzungen und Temperaturen unterscheiden. Mit Blick auf die starke räumliche Kompression für das gesamte Zwischeneinzugsgebiet (hier für HDNA40, 1200 km²) und der Summierung zu Tagesdaten sowohl der Input- als auch der Targetwerte ist die Performance dieses Subnetzes 1 hervorragend.

Subnetz 2

Das Subnetz 2 errechnet das Zustandsmerkmal aus der Wettergeschichte vor dem eigentlichen Ereignis. Dazu soll ebenso ein MLFN verwendet werden. Das Zustandsmerkmal ist die Scheitelhöhe eines Hochwassers, das aus einem 24 Stunden dauernden Einheitsniederschlag von 4 mm h⁻¹ resultiert. Da Einheitsimpulse auf eine natürliche Wettergeschichte in die schneefreie Zeit plaziert und, um unterschiedliche Muster aufzuweisen, mit einem Abstand von mindestens einer Woche versehen wurden, ist der Umfang der Stichprobe gering. Kalibrierungs-, Validierungs- und Testdatensatz umfassen jeweils 356, 54 und 300 verschiedene Gebietszustände. Die relevanten Merkmale beschränken sich zwar auf die Größen Effektivniederschlag und potentielle Verdunstung, doch ist an dieser Stelle eine konkrete Untersuchung nötig, wieweit die relevante Historie zurückreicht, um gegebenenfalls noch einen Einfluß auf den Gebietszustand bei Ereignisbeginn auszuüben.

Dazu soll mit dem Einzugsgebietsmodell eine Sensitivitätsanalyse durchgeführt werden. Natürliche Ganglinien der Temperatur, Windgeschwindigkeit, relativen Luftfeuchte und Globalstrahlung (beziehungsweise der daraus resultierenden potentiellen Verdunstung) wurden dazu mit künstlichen Ganglinien des Niederschlags kombiniert. Damit wurde der Einfluß des Niederschlags in der Vorgeschichte getestet, da dessen unmittelbarer Effekt auf die Bodenfeuchte größer ist als der der potentiellen Verdunstung. Dabei wurde das Gebiet mit 0,1 mm h⁻¹ (entspricht 876 mm a⁻¹) gleichmäßig beregnet, um eine Austrocknung des Gebietes durch das Fehlen natürlicher Niederschläge zu verhindern, aber trotzdem natürliche Impulse auszuschalten. Außer diesem „Nullszenario“ wurden Szenarien mit in unterschiedlicher zeitlicher Distanz (2 Wochen, 1, 2, 3, 4 und 5 Monate) zum hochwasserauslösenden Einheitsniederschlag und in unterschiedlicher Intensität (1, 2 und 4 mm h⁻¹ für 24 h) gesetzten Niederschlagsimpulsen gerechnet. Abbildung 42 zeigt die Ganglinien des Abflusses für den Einheitsniederschlag bezogen auf den stärksten Impuls von gleichfalls 4 mm h⁻¹ in der jeweiligen Vorgeschichte.

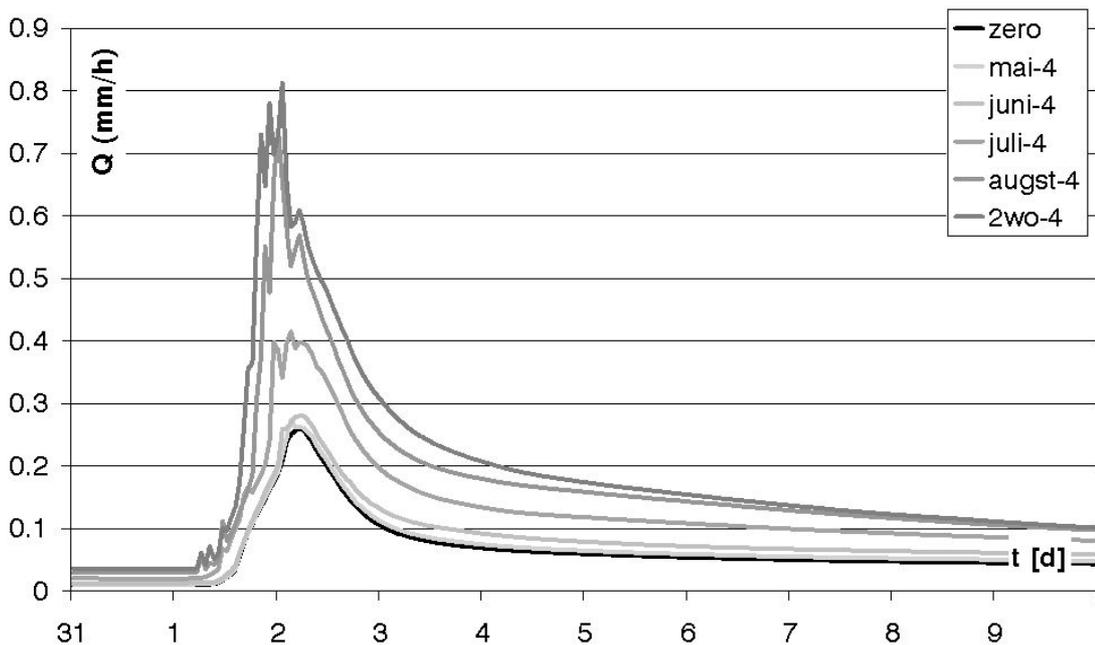


Abbildung 42: Sensitivität eines Hochwassers (1. September) auf zurückliegende Niederschlagsimpulse von 4 mm h^{-1} für 24 h (1. Mai, 1. Juni, 1. Juli, 1. August und 18. August)

Tabelle 9 faßt die Abhängigkeit der Scheitelhöhe von einem Ereignis in der Geschichte zusammen. Dabei ist zu beachten, daß die Vorfeuchte für selbst das Nullszenario nicht außergewöhnlich trocken ist. Ein Regenereignis, welches 3 Monate zurückliegt, beeinflusst noch deutlich die Gebietsreaktion auf einen aktuellen Niederschlag.

Tabelle 9: Scheitelhöhe des Standardhochwassers am Pegel Nossen in Abhängigkeit von Intensität und zeitlichem Abstand eines vergangenen Niederschlagsimpulses (in mm h^{-1} und Prozent des Vergleichsscheitels ohne Niederschlagsimpuls von $0,2606 \text{ mm h}^{-1}$)

	24h x 1mm/h		24h x 2mm/h		24h x 4mm/h	
	mm/h	%	mm/h	%	mm/h	%
4 Monate	0.2629	100.9	0.2629	100.9	0.2629	100.9
3 Monate	0.2745	105.3	0.2781	106.7	0.2807	107.7
2 Monate	0.2943	112.9	0.3819	146.5	0.4156	159.4
1 Monat	0.3136	120.3	0.5009	192.2	0.7380	283.1
2 Wochen	0.3246	124.5	0.5685	218.1	0.8135	312.1

Szenarienrechnungen ohne die gleichmäßige Beregnung mit $0,1 \text{ mm h}^{-1}$ führten bei einem Impuls von $24\text{h} \times 4\text{mm}^{-1}$ mit vier Monaten Abstand zum Einheitsimpuls noch zu einer Scheitelerhöhung auf 150% im Vergleich zur Nullvariante. Somit kann die Einflußspanne für trockenere Jahre im NA-Modell sogar noch größer sein.

Ergebnisse der Prognose-ANN

Versuche, die Auswirkung der zeitlichen Distanz eines vergangenen Niederschlagsimpulses auf den Gebietszustand in einem gewichteten Mittel der Niederschlagsgeschichte zu berücksichtigen, schlugen fehl, da sich zeigte, daß der Einfluß eines lange zurückliegenden Niederschlagsereignisses durch folgende Ereignisse stark überprägt wird. Die Relevanz eines solchen steigt allerdings, wenn darauf eine trockene Periode folgt.

Aus diesem Grunde werden die Niederschläge für verschiedene Zeitspannen nach Tabelle 10 aufsummiert und diese Summen einzeln als Merkmale verwendet.

Tabelle 10: Niederschlagsmerkmale für das Subnetz 2

Zeitraum vor Ereignis	zeitl. Auflösung	Anzahl der Merkmale
1-4 Tage	1 d	4
5-12 Tage	2 d	4
13-28 Tage	4 d	4
29-60 Tage	8 d	4
61-90 Tage	15 d	2
91-150 Tage	30 d	2
		Σ 20

Außer den Niederschlagsmerkmalen benötigt das Subnetz 2 noch Informationen über die potentielle Verdunstung. Da deren Einfluß indirekter Natur ist, hat es wenig Sinn, die Sensitivität der potentiellen Verdunstung in Form von Impulsen zu testen. Im Modell greift die potentielle Verdunstung mit dem aktuellen Gebietszustand und dem Niederschlag ineinander und ergibt die neue Gebietsfeuchte. Die Wirkung verdunstungsfördernder Verhältnisse hängt maßgeblich von der Wasserverfügbarkeit ab. Danach soll die potentielle Verdunstung genau wie der Niederschlag als Summe über den gleichen Zeitraum eingehen und steuert damit ebenso 20 Merkmale für das Subnetz 2 bei.

Der Stichprobenumfang und die Dimensionierung des Inputvektors stehen allerdings in einem ungünstigen Verhältnis: Ein Netz mit vielen voneinander unabhängigen Inputs kann daraus hervorragend Funktionen mit wenigen Stützstellen approximieren, aber die Gefahr einer „Überanpassung“ (engl. overtraining) ist groß. Das bedeutet, daß das Netz alle zufälligen Abweichungen mit erlernt, aber dadurch an Generalisierbarkeit verliert. Damit würde die wesentliche Input-Output-Beziehung verwässert.

Es wurden wieder mehrere Varianten von Neuronenzahlen mit je 5 Initialisierungen in der verdeckten Schicht getestet. Abbildung 43 zeigt die besten Ergebnisse aus diesen Trainingsläufen. Trotz der Verwendung eines Validierungsdatensatzes, welcher eine

Ergebnisse der Prognose-ANN

Überanpassung verhindern soll, zeigt sich, daß Netze mit einer größeren Anzahl verdeckter Neuronen besser an die Trainingsdaten angepaßt werden können. Bei unbekanntem Datensätzen nimmt die Performance deutlich ab, die Generalisierbarkeit des Netzes leidet. Als beste Anpassung wird das Netz mit zwei verdeckten Neuronen gewählt.

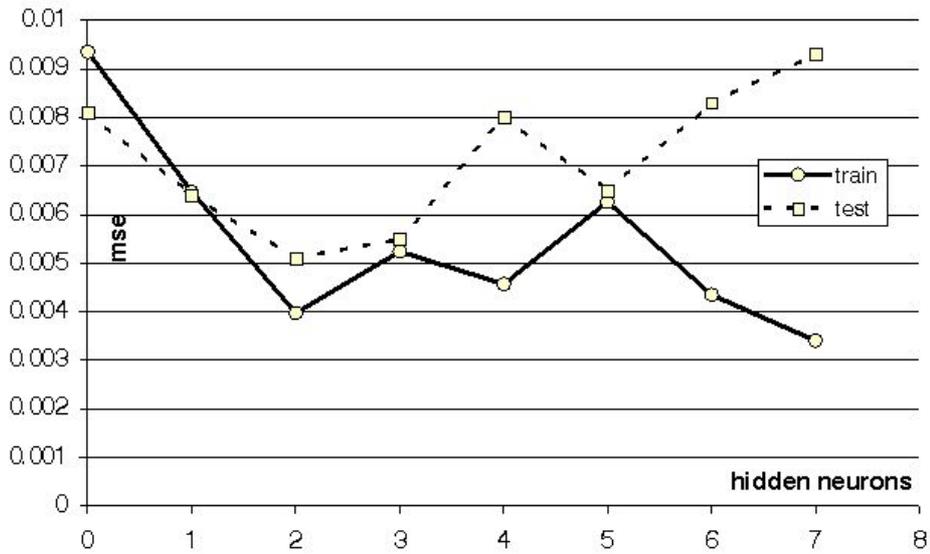


Abbildung 43: Performance des Subnetzes 2 in Abhängigkeit von der Anzahl der Neuronen in der verdeckten Schicht, mittlerer quadratischer Fehler für normierte Daten.

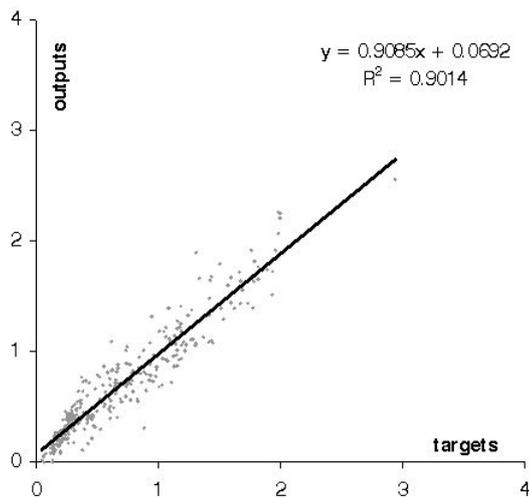


Abbildung 44: Testdaten für Subnetz 2, Variante mit 2 Neuronen

Tests mit Datensätzen, die eine um einen Monat kürzere Vorgeschichte verwenden, zeigten durchweg schlechtere Anpassungen ($mse > 0.0068$ für den Testdatensatz), was belegt, daß die relevante Geschichte durchaus so weit zurückreicht.

Abbildung 44 zeigt die Performance des Netzes mit zwei verdeckten Neuronen für die Testdaten. Es kann konstatiert werden, daß das Netz durchaus sehr gut in der Lage ist, das gewählte Zustandsmerkmal aus der Geschichte des Wetters in dieser Diskretisierung zu treffen. Die Streuung in dem Bereich mit einer hohen Vorfeuchte (große Werte für das Zustandsmerkmal), der für Extremereignisse interessant ist, ist nicht größer als im trockeneren Bereich.

5.2.2 Test der Netztypen für die NA-Prozesse

Um die für die NA-Prozesse relevanten Merkmale herauszufiltern und die Prognosegüte eines nach dieser Methodik erstellten Netzes in bezug auf diese Prozesse testen und beurteilen zu können, wird ein reines NA-Netz für den Pegel Nossen und für räumlich gleichverteilte Niederschläge (mit Berücksichtigung einer Höhenabhängigkeit nach KOSTRA) erstellt. Es soll außer dem Zustandsmerkmal noch die gesamte relevante Information für das Ereignis in adäquater zeitlicher Auflösung umfassen.

Entsprechend der in 3.6 beschriebenen Methodik muß grundsätzlich ein einziges Netz für alle Prognoseweiten erstellt werden, da nur so eine flexible Einarbeitung der Merkmale in das HD-ANN möglich ist. Es muß eine Zeitspanne gefunden werden, innerhalb derer die Niederschlagswerte den Targetwert des Abflusses am Gebietsauslaß maßgeblich beeinflussen und außerhalb derer die Information zunehmend komprimiert in den Merkmalsvektor einbezogen werden kann. Dazu wurde für alle berechneten Szenarien jeder Niederschlagswert zum Zeitschritt t_i (mit dem Ereignisbeginn t_0 , für welchen das Zustandsmerkmal gilt) mit jedem Durchflußwert am Gebietsauslaß zum Zeitpunkt t_j korreliert. Abbildung 45 zeigt die Ergebnisse, wobei die Abszisse den Zeitunterschied zwischen Niederschlagswert und Abflußwert $t = t_j - t_i$ abbildet. Der Korrelationskoeffizient R auf der Ordinate zeigt den Zusammenhang zwischen den Größen. Die verschiedenen Realisierungen stehen für verschiedene Niederschlagszeitpunkte $t_i \leq 15$. Dabei zeigt sich eine tendenzielle Abhängigkeit von t_i selbst. Diese Abhängigkeit ist allerdings schwer zu quantifizieren.

Ergebnisse der Prognose-ANN

Die Ursache dafür ist wohl in der unterschiedlichen Dauer der Niederschlagsereignisse zu sehen, so daß für größere t_i ein Teil der Ereignisse bereits beendet ist. Ein Niederschlag von null verlagert den Peak aber tendenziell zeitlich nach hinten, wo der Durchfluß wieder abnimmt. Für die Entscheidung über die zentrale Einflußspanne sollen der Maximalwert und der Mittelwert aus diesen einzelnen Graphen maßgeblich sein.

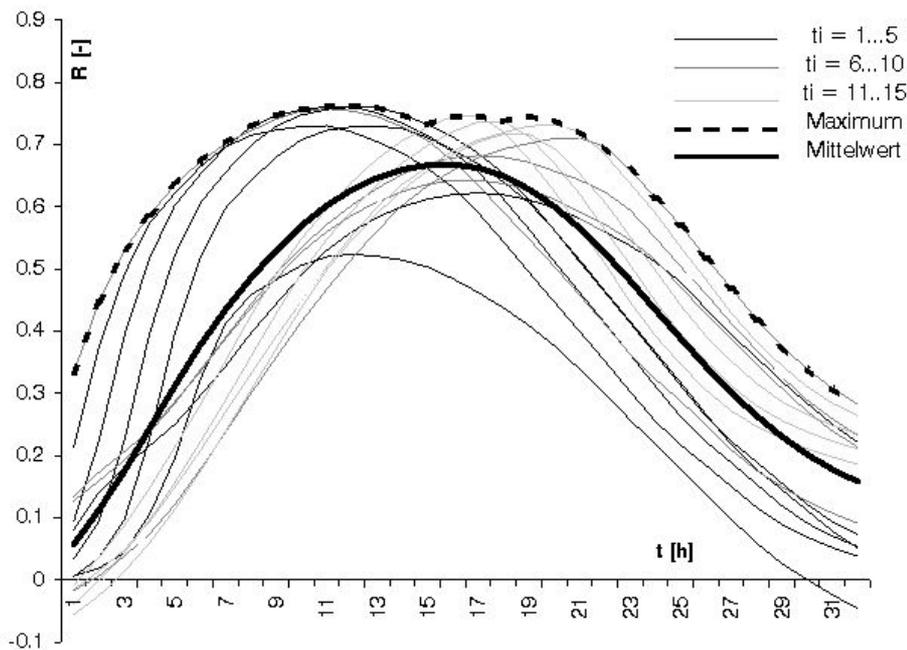


Abbildung 45: Korrelation eines Niederschlagswertes mit im Abstand t darauffolgenden Abflußwerten, Realisierungen für verschiedene Zeitpunkte ab Ereignisbeginn, Maximum (gestrichelt fett) und Mittelwert (fett)

Tabelle 11: Merkmalsvektor für das NA-Netz

Merkmalsvektor	Bereich	Merkmalsvektor	Bereich
M_1	- (ZM)	M_{11}	$t_j - t_i = 15 \text{ h}$
M_2	$0 \text{ h} \leq t_j - t_i \leq 2 \text{ h}$	M_{12}	$t_j - t_i = 16 \text{ h}$
M_3	$3 \text{ h} \leq t_j - t_i \leq 5 \text{ h}$	M_{13}	$t_j - t_i = 17 \text{ h}$
M_4	$6 \text{ h} \leq t_j - t_i \leq 7 \text{ h}$	M_{14}	$t_j - t_i = 18 \text{ h}$
M_5	$8 \text{ h} \leq t_j - t_i \leq 9 \text{ h}$	M_{15}	$t_j - t_i = 19 \text{ h}$
M_6	$t_j - t_i = 10 \text{ h}$	M_{16}	$20 \text{ h} \leq t_j - t_i \leq 21 \text{ h}$
M_7	$t_j - t_i = 11 \text{ h}$	M_{17}	$22 \text{ h} \leq t_j - t_i \leq 23 \text{ h}$
M_8	$t_j - t_i = 12 \text{ h}$	M_{18}	$24 \text{ h} \leq t_j - t_i \leq 26 \text{ h}$
M_9	$t_j - t_i = 13 \text{ h}$	M_{19}	$27 \text{ h} \leq t_j - t_i \leq 29 \text{ h}$
M_{10}	$t_j - t_i = 14 \text{ h}$	M_{20}	$30 \text{ h} \leq t_j - t_i$

Ergebnisse der Prognose-ANN

Nach Abbildung 45 wurde diese Spanne mit $10h \leq t_j - t_i \leq 19h$ festgesetzt; das heißt, in diesem Bereich wurden die Stundenwerte komplett in den Merkmalsvektor übernommen. Außerhalb dieser Spanne wurden die Werte entsprechend Tabelle 11 zusammengefaßt. Dabei wurde als erstes Merkmal M_1 das Zustandsmerkmal ZM verwendet. Alle weiteren berechneten sich als Niederschlagssummen entsprechend der angegebenen Zeitbereiche. Liegen diese Zeitbereiche ganz oder zum Teil zeitlich vor dem durch das Zustandsmerkmal determinierten Ereignisbeginn, sind die Niederschlagswerte in jedem Falle gleich null.

Da das resultierende Netz für alle Vorhersagezeiträume gültig ist und damit für ein Ereignis mehrere Input-Output-Paare aus den Niederschlags- und Abflußganglinien herausgefiltert werden, müssen für jedes Ereignis Vorhersagebeginn und -ende definiert werden. Aus den Betrachtungen über den zentralen Einflußbereich resultieren die folgenden Kriterien für den trainierten Zeitraum:

- Beginn: $t_j = 6$ (12)

- Ende: $\sum_{i=2}^{15} M_i > 0$ (13)

Damit beträgt die minimale Vorhersageweite 6 Stunden (6 Stunden nach Beginn des hochwasserauslösenden Niederschlages). Die maximale Vorhersageweite ist ereignisabhängig und mit 20 Stunden nach Niederschlagsende festgesetzt. Daraus ergibt sich eine umfangreiche Datenbasis mit 200 000, 79 445 und 162 558 Input-Output-Paaren für Trainings-, Validierungs- und Testdatensatz.

MLFN

Es wurden wiederum verschiedene Neuronenzahlen in der verdeckten Schicht mit je 5 Initialisierungen getestet. Abbildung 46 zeigt die Güte der besten trainierten Netze anhand des mittleren quadratischen Fehlers für den Trainings- und den Testdatensatz (Daten zwischen -1 und 1 normiert). Die Testperformance des Netzes mit 9 verdeckten Neuronen konnte von Netzen mit mehr Neuronen in dieser Schicht nicht übertroffen werden.

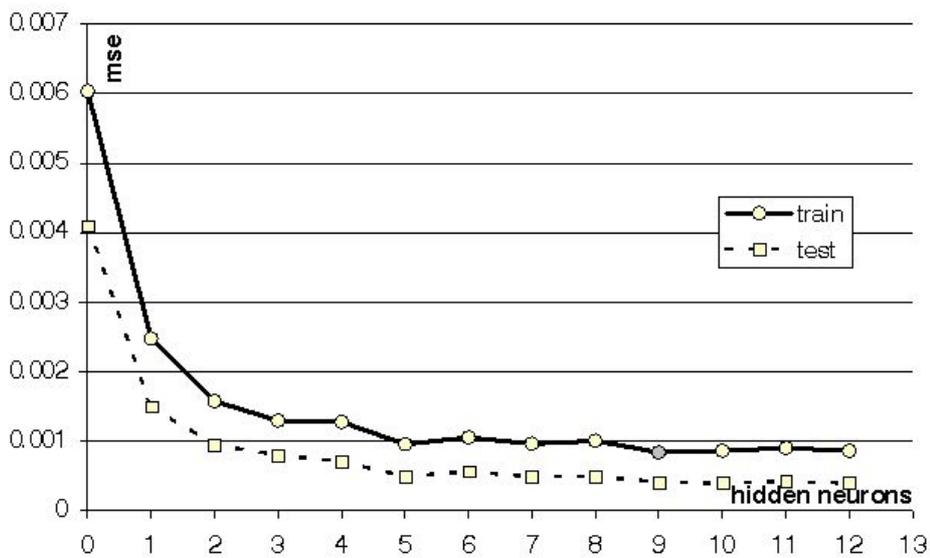


Abbildung 46: Performance des MLFN für die NA-Prozesse in Abhängigkeit von der Anzahl der Neuronen in der verdeckten Schicht, mittlerer quadratischer Fehler für normierte Daten.

Bemerkenswert ist die Tatsache, daß die Performance der Netze sämtlicher Varianten für die Testdaten durchgängig deutlich besser ist, als für die Trainingsdaten. Bei der großen Anzahl an Trainingspaaren ist eine Überanpassung des Netzes sehr unwahrscheinlich. Daß die Performance für den Testdatensatz aber zufällig merklich besser ist als für die Trainingsdaten, ist der Aufteilung der Ereignisse geschuldet, da die Grundlage der Testdaten reale Gebietszustände ab 1985 und der Trainings- und Validierungsdaten vor 1985 sind.

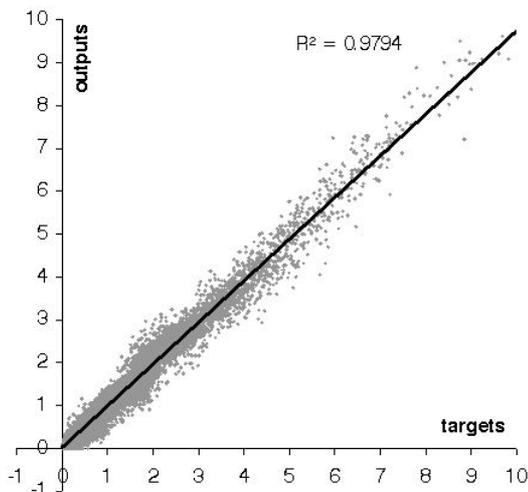


Abbildung 47: Testdaten für das NA-MLFN, Abfluß in mm h^{-1} , Variante mit 9 verdeckten Neuronen

Ergebnisse der Prognose-ANN

Die Performance des ausgewählten Netzes mit 9 verdeckten Neuronen ist für die Testdaten in Abbildung 47 dargestellt.

SOM

Selbstorganisierende Merkmalskarten sind klassifizierende Netzwerke. Wie in 3.2.3 beschrieben, gibt es keine Input-Output-Paare, sondern den Trainingsvektoren im n -dimensionalen Raum wird eine Merkmalskarte angepaßt, die die Lage der Trainingsvektoren so gut wie möglich abbildet. Die Targetwerte bilden also eine zusätzliche Dimension des Merkmalsraumes, - und damit eine weitere Größe im Inputvektor. Dabei werden alle Werte zunächst ebenso zwischen -1 und 1 normiert. Da der Vektorraum hier 21-dimensional ist und die Lage der Merkmalskarte im Raum a priori keine Dimension bevorzugt, wurde der Wertebereich in Richtung der „Outputdimension“ um ein Outputdimensionsgewicht ODW gestreckt.

Um die Vergleichbarkeit mit MLFN herzustellen, wurden die gleichen Gütekriterien verwendet und der mittlere quadratische Fehler auch wiederum auf die ursprünglichen (zwischen -1 und 1 normierten) Daten bezogen.

Die hier dargestellten SOM besitzen eine Dimension von 59 x 38 Neuronen in einem hexagonalen Raster. Das Training eines signifikant größeren Rasters ist rechentechnisch nicht mehr möglich. Vergleichbare SOM mit Rechteckraster liefern äquivalente Ergebnisse.

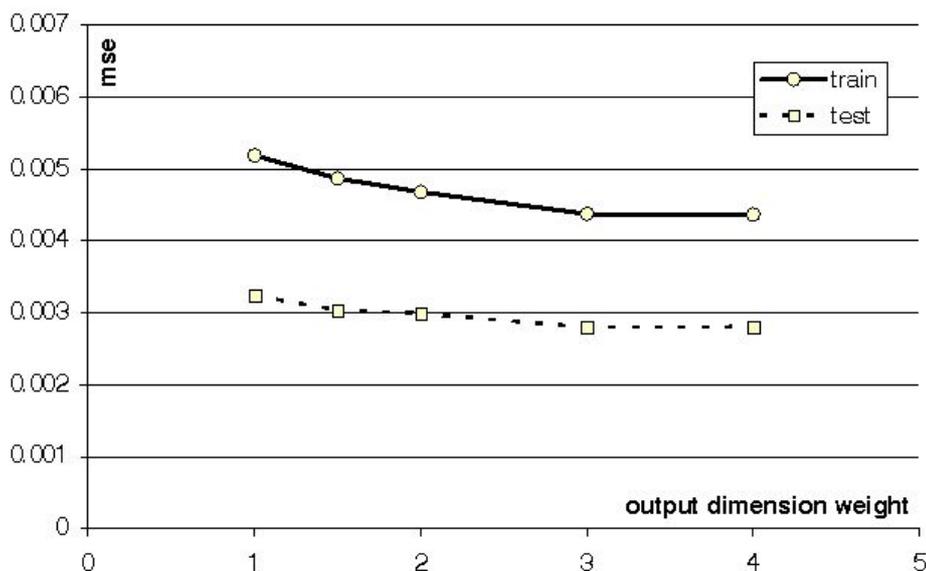


Abbildung 48: Performance des SOM (59 x 38) für die NA-Prozesse in Abhängigkeit vom Outputdimensionsgewicht, mittlerer quadratischer Fehler für normierte Daten.

Ergebnisse der Prognose-ANN

Wie in Abbildung 48 zu sehen ist, hat die Einführung des Outputdimensionengewichtes durchaus einen positiven Effekt auf die Performance des Netzes, wenngleich die exzellenten Ergebnisse des MLFN (Abbildung 46) hier bei weitem nicht erreicht werden können. Abbildung 49 zeigt die Korrelation der mit dem SOM berechneten Abflußwerte mit den Targets für ein Outputdimensionengewicht von $ODW=3$. Die Ergebnisse decken zwar den gleichen Wertebereich ab, jedoch zeigt sich eine große Streuung – besonders im Bereich höherer Abflüsse. Augenfällig ist hier auch die für SOM typische diskrete Verteilung der Werte, die aus der endlichen Anzahl von Neuronen resultiert. Jedoch ist der daraus resultierende Fehler zweifellos vernachlässigbar.

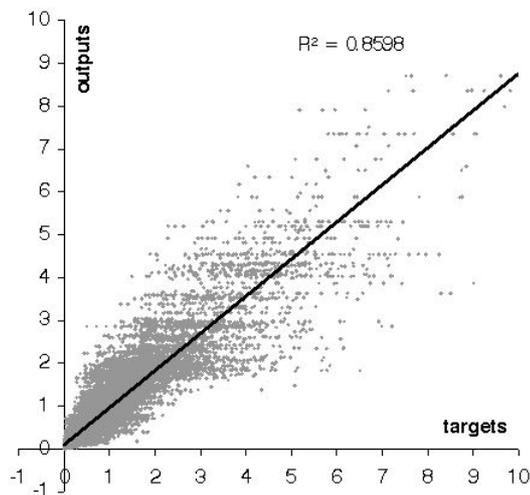


Abbildung 49: Testdaten für das NA-SOM, Variante $ODW=3$, Abfluß in $mm\ h^{-1}$

Eine der herausragenden Eigenschaften selbstorganisierender Merkmalskarten ist die Möglichkeit der Visualisierung des Zusammenhanges verschiedener Größen. In Abbildung 50 sind die Karten für das Zustandsmerkmal und den Abfluß gegenübergestellt. Hohe Werte sind durch rot bis braun gefärbte Waben und niedrige durch blaue Waben gekennzeichnet. Jede Wabe entspricht einem Neuron.

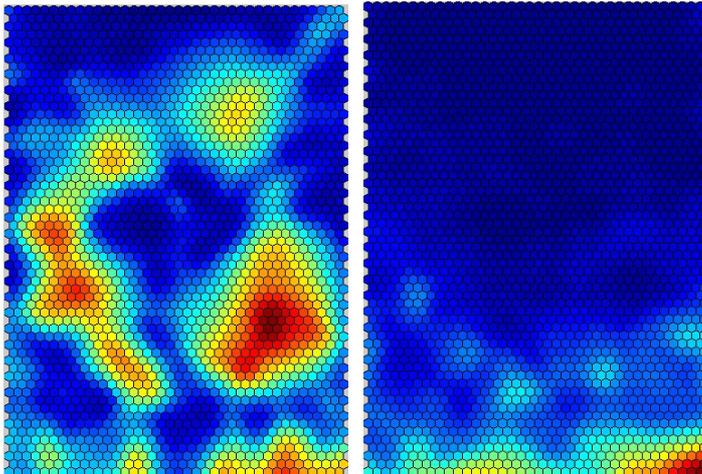


Abbildung 50: Ansicht der SOM für die Merkmale M1 (links, Zustandsmerkmal) und M21 (rechts, Abfluß), braun: maximal, blau: minimal

Es ist gut erkennbar, daß Neuronen, die einen hohen Abfluß verkörpern, tendenziell einer erhöhten Vorfeuchte entsprechen. Eine komplette Darstellung der Karte in all ihren Dimensionen ist in Anhang F abgebildet. Daraus geht hervor, daß vor allem die Niederschlagsmerkmale M_5 bis M_{10} auch Maxima im Bereich des maximalen Abflusses aufweisen. Hiermit bestätigt sich die Interpretation von Abbildung 45, wonach die späteren Peaks der Kurven für höhere t_i eher weniger die realen Zusammenhänge darstellen. Um diese Vermutung zu überprüfen, soll die zentrale Einflußspanne verkleinert und damit die Anzahl der Merkmale testweise verringert werden. Damit ergeben sich Merkmale nach Tabelle 12.

Tabelle 12: verkleinerter Merkmalsvektor für das NA-Netz (SOM)

Merkmals	Bereich	Merkmals	Bereich
M_1	- (ZM)	M_8	$t_j - t_i = 12 \text{ h}$
M_2	$0 \text{ h} \leq t_j - t_i \leq 2 \text{ h}$	M_9	$13 \text{ h} \leq t_j - t_i \leq 14 \text{ h}$
M_3	$3 \text{ h} \leq t_j - t_i \leq 5 \text{ h}$	M_{10}	$15 \text{ h} \leq t_j - t_i \leq 16 \text{ h}$
M_4	$6 \text{ h} \leq t_j - t_i \leq 7 \text{ h}$	M_{11}	$17 \text{ h} \leq t_j - t_i \leq 19 \text{ h}$
M_5	$8 \text{ h} \leq t_j - t_i \leq 9 \text{ h}$	M_{12}	$20 \text{ h} \leq t_j - t_i \leq 23 \text{ h}$
M_6	$t_j - t_i = 10 \text{ h}$	M_{13}	$24 \text{ h} \leq t_j - t_i$
M_7	$t_j - t_i = 11 \text{ h}$	M_{14}	- ($Q(t_i)$)

Die Performance der sich daraus ergebenden SOM ist in Abbildung 51 als mittlerer quadratischer Fehler für normierte Eingangsdaten aufgetragen. Im Vergleich zu den SOM für den kompletten Merkmalsatz (Abbildung 48) ist zu sehen, daß die

Ergebnisse der Prognose-ANN

Anpassungen leicht besser, aber nahezu identisch sind. Eine stärker komprimierte Form der Information kann sich demnach positiv auf die Performance selbstorganisierender Merkmalskarten auswirken. Dieses Ergebnis ist insofern als positiv zu bewerten, da für die Erstellung eines HD-Netzes eine größere Anzahl an Merkmalen zu berücksichtigen ist. In Anhang F findet sich eine detaillierte Darstellung der SOM für $ODW=3$.

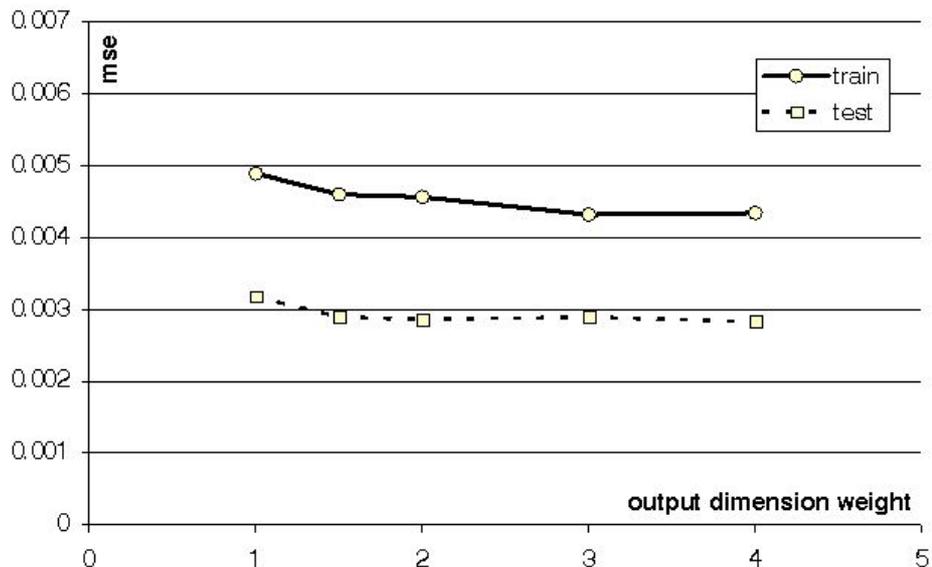


Abbildung 51: Performance der SOM für den reduzierten Merkmalsatz (14) für die NA-Prozesse in Abhängigkeit vom Outputdimensionsgewicht, mittlerer quadratischer Fehler für normierte Daten

Zusammenfassend kann bestätigt werden, daß sich die Methodik der Merkmalsselektion für die NA-Prozesse eignet, die Modellergebnisse mit künstlichen neuronalen Netzen zu reproduzieren. Multilayer-feedforward-Netze sind selbstorganisierenden Merkmalskarten bei der Abbildung komplizierter funktionaler Zusammenhänge mit einer differenzierten Eingangsinformation deutlich überlegen. Trainierte SOM hingegen liefern auch und gerade bei der Verwendung hochdimensionaler Merkmalsvektoren interpretierbare Ergebnisse, die zu einer sinnvollen, weitergehenden Komprimierung der Eingangsdaten verwendet werden können.

5.2.3 Kombination der Subnetze mit dem NA-Netz

Um die Eignung der Methodik zur Bereitstellung der Merkmale für die NA-Prozesse abschließend zu beurteilen, müssen das NA-MLFN und die NA-SOM in Kombination mit den Subnetzen betrieben werden. Das heißt, das Zustandsmerkmal im Merkmalsvektor wird nun direkt mit den Subnetzen simuliert.

Äquivalent zu Abbildung 47 ist in Abbildung 52 (a und b) die Korrelation der Netzoutputs des MLFN mit den über das NA-Modell tatsächlich berechneten Werten (Targets) für die Testperiode dargestellt. Dabei berücksichtigt Abbildung 52 (a) alle Ereignisse in der Testperiode. In Abbildung 52 (b) sind die Ereignisse ausgelassen, bei denen Schneeschmelze auftritt, da für diese das Zustandsmerkmal mit WaSim-ETH nicht bestimmbar ist und diese Vorfeuchteszenarien nicht für das Training des Subnetzes 2 verwendet werden konnten. Eine Berechnung des Zustandsmerkmals mit dem trainierten Subnetz 2 ist demgegenüber immer möglich. Hier zeigt sich, daß die Netze auch diesen Fällen eine ebenso gute Performance zeigen.

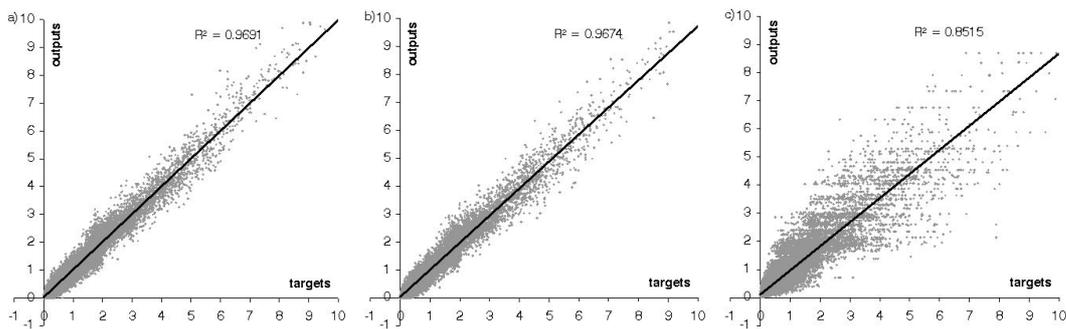


Abbildung 52: Performance der NA-Netze in Verbindung mit den Subnetzen (a: MLFN, b: MLFN-schneefreie Historie, c: SOM)

Abbildung 52 (c) zeigt die Performance des SOM in Verbindung mit den Subnetzen. Während die Verwendung des mit den Subnetzen berechneten Zustandsmerkmals bei den sehr gut funktionierenden MLFN erwartungsgemäß zu einer leicht vergrößerten Abweichung führt, wird die nicht so gute Prognosequalität des SOM im Vergleich zu Abbildung 49 nicht weiter beeinträchtigt.

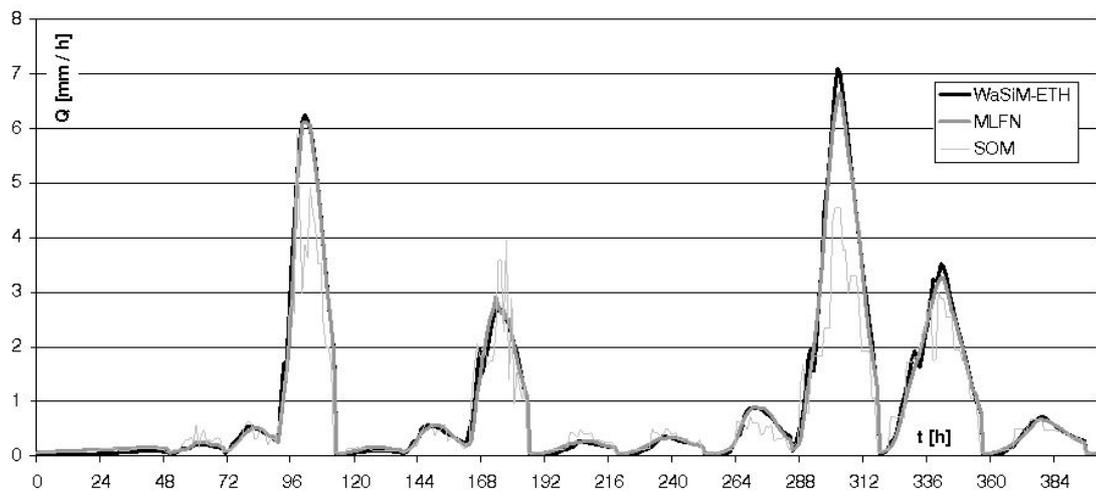


Abbildung 53: NA-Netze - Darstellung einzelner Szenarien

Während bisher die Netze nur an der Prognose einzelner Werte ohne Berücksichtigung ihrer zeitlichen Abfolge gemessen wurden, stellt Abbildung 53 die Ergebnisse zurück in den Zusammenhang einzelner Hochwasserganglinien. Dabei handelt es sich nicht um eine fortlaufende Ganglinie, sondern um eine Aneinanderreihung der Ausschnitte der Ganglinien, die den Hauptteil der Hochwasserwelle enthalten. Die Länge des Ausschnittes ist abhängig von der Dauer des hochwasserauslösenden Niederschlages, beträgt aber ungefähr 24 Stunden. Es ist zu erkennen, daß sehr hohe Scheitel mit einem Abfluß von zirka 7 mm h^{-1} mit den MLFN ebensogut getroffen werden wie Ereignisse, die gar keinen sichtbaren Scheitel hervorrufen. Dabei wird die Form der Welle im Vergleich zu den Modellergebnissen geglättet, ohne daß die Höhe des Hauptscheitels merklich beeinträchtigt wird. Die Scheiteleintrittszeiten sind von den MLFN hervorragend getroffen. Die SOM zeigen demgegenüber ziemlich große Abweichungen betreffs der Scheitelhöhe. Auch die Dynamik kann nicht so gut wiedergegeben werden, zum Teil fehlt sogar eine deutliche Ausprägung eines einzelnen Scheitels.

Diese Ergebnisse lassen vermuten, daß MLFN für die Beschreibung der hydrodynamischen Prozesse gut detaillierte und aussagekräftige Merkmale für die NA-Prozesse im Zwischengebiet integrieren können. Die Dimensionierung der reinen NA-MLFN (9 verdeckte Neuronen) hält sich in überschaubaren Grenzen, was die Komplexität der Aufgabe, die NA-Prozesse in das HD-Netz zu integrieren, beherrschbar erscheinen läßt.

Die Anwendung der SOM als Prognosewerkzeug ist weniger geeignet. Wenngleich die erhöhte Wichtung des Abflußwertes im Merkmalsvektor die Performance der SOM verbessern konnte, ist das Prinzip des nichtüberwachten Lernens für dieses Problem scheinbar ungeeignet, da der Merkmalsvektor zu viele gleichgestellte Dimensionen aufweist und im Grunde genommen eine Klassifizierung der Datensätze ausführt. Ein MLFN mit einer ausgewiesenen Outputgröße ist für eine solche Fragestellung offenbar wesentlich geeigneter.

5.2.4 MLFN für die HDNA-Szenarien

Die Basis der Merkmalsvektoren der Netze für die HDNA-Szenarien bilden die Merkmale aus den Zuflußganglinien für die Oberliegerpegel, also die Merkmale aus den HD-Szenarien. Zusätzlich werden das Zustandsmerkmal und die Ereignismerkmale für die Niederschlags-Abfluß-Prozesse im Zwischengebiet aufgenommen. Mit zunehmender Größe des Zwischeneinzugsgebietes steigt zum einen der Anteil des Beitrages der Niederschlags-Abfluß-Prozesse am Gesamtabflußvolumen am Zielquerschnitt. Andererseits wirkt sich der als räumlich gleichverteilt angenommene Niederschlag eines Zeitschrittes über einen größeren Zeitraum auf den Abfluß aus. Daraus resultiert der Bedarf an einer größeren Anzahl von NA-Merkmalen. Demgegenüber sinkt die Anzahl der Zuflüsse, was folgerichtig zu einer Reduktion der HD-Merkmale führt. Tabelle 13 zeigt die entsprechend der Methodik ermittelten und hier verwendeten Merkmale für die betrachteten Szenarien.

Die Gesamtanzahl der Merkmale ist für die Szenarien also mehr oder weniger gleich, jedoch variiert der Anteil der HD- bzw. NA-Merkmale entsprechend der Zwischengebietsgröße. Damit wurden wiederum MLFN mit ein bis fünfundzwanzig verdeckten Neuronen trainiert, wobei Abbildung 54 jeweils die besten von je fünf Initialisierungen zeigt. Generell ist festzustellen, daß auch hier kleinere Zwischengebiete zu besseren Anpassungen mit einem geringeren Unterschied zwischen Trainings- und Testperformance führen. Hierfür ist jedoch im Gegensatz zu den HD-Szenarien nicht vordergründig ein geringerer Informationsgehalt der Merkmale verantwortlich zu machen. Der Abfluß aus dem Zwischeneinzugsgebiet resultiert hier aus Niederschlagswerten, die sowohl durch das Niederschlags-Abfluß-Modell als auch durch das hydrodynamische Modell transformiert werden, während der Zufluß in das Zwischengebiet hier bereits das NA-Modell durchlaufen hat und im operativen Fall einem eigenen Prognosefehler unterliegt. Eine größere Zwischengebietsfläche muß

Ergebnisse der Prognose-ANN

also aufgrund der mit dem neuronalen Netz beschriebenen Prozesse naturgemäß zu einem größeren Fehler führen. Die hier als HD-Merkmale bezeichneten Inputgrößen unterliegen also – für den Fall, daß auch sie einer Prognose aus gemessenen Niederschlägen entstammen – einer zusätzlichen modellabhängigen Unsicherheit gegenüber den Niederschlagswerten, beziehungsweise NA-Merkmalen.

Tabelle 13: Merkmale für die HDNA-ANN, HD-Merkmale grau unterlegt

Merkmal	HDNA10	HDNA20	HDNA40
M1	K3	K3	K3
M2	K4	K4	K4
M3	K5	K5	K5
M4	K6	K6	K6
M5	K7	K7	K7
M6	K8	K8	K8
M7	K9	K9	K9
M8	N6	N6	ZM
M9	N7	N7	0..2
M10	N8	N8	3..5
M11	N9	N9	6..7
M12	S6	ZM	8..9
M13	S7	0..2	10
M14	ZM	3..5	11
M15	0..2	6..7	12
M16	3..4	8..9	13
M17	5..6	10	14
M18	7..8	11	15
M19	9	12	16
M20	10	13..14	17
M21	11..12	15..16	18
M22	13..14	17..19	19
M23	15..17	20..23	20..21
M24	18...	24...	22..23
M25	-	-	24..26
M26	-	-	27..29
M27	-	-	30...

Kx...Durchfluß Pegel Kriebstein x Stunden vor Prognosezeitschritt

Nx...Durchfluß Pegel Nossen x Stunden vor Prognosezeitschritt

Sx...Durchfluß Pegel Niederstregis x Stunden vor Prognosezeitschritt

ZM...Zustandsmerkmal für NA-Prozesse

x..y...Summe der Gebietsniederschläge für Zeitpunkte $x \leq t_j - t_i \leq y$

Ergebnisse der Prognose-ANN

Die besten Anpassungen in bezug auf die Testperformance zeigen die Netze mit 18, 23 und 24 verdeckten Neuronen für die Szenarien HDNA10, HDNA20 und HDNA40. Die Testperformance weist bei allen drei Szenarien zwischen 20 und 25 verdeckten Neuronen keinen deutlichen Trend mehr auf, so daß keine größeren Netze mehr trainiert wurden.

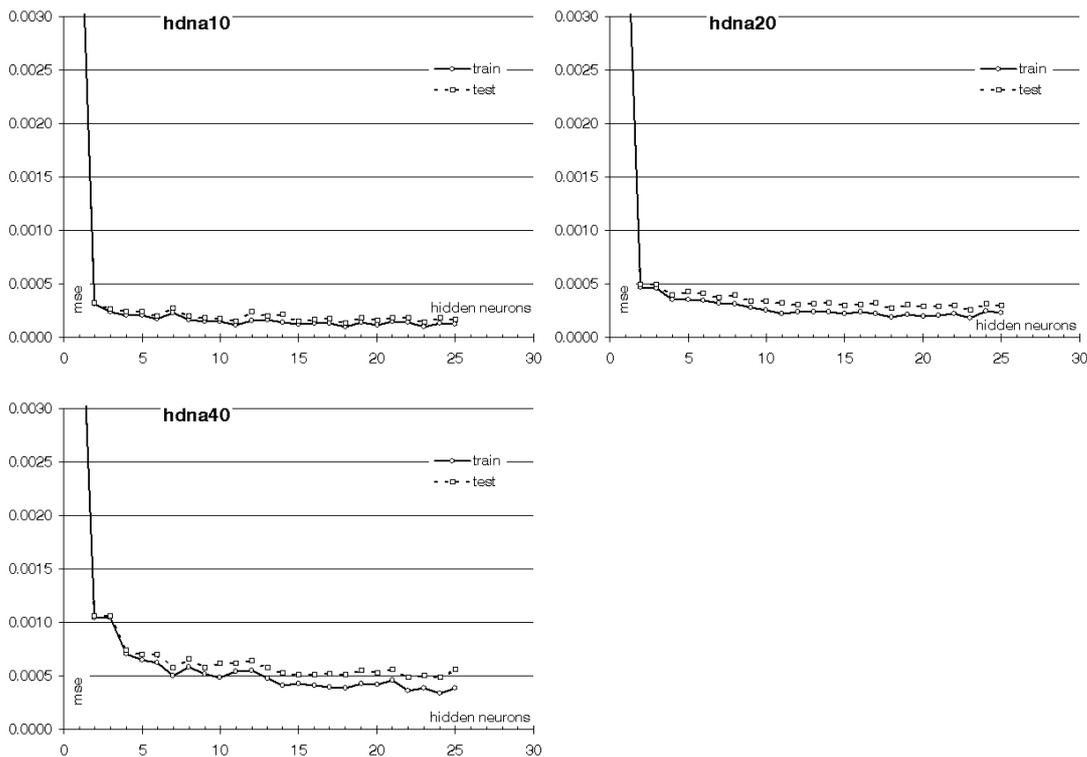


Abbildung 54: Performance der MLFN für die Szenarien HDNA10, HDNA20 und HDNA40 in Abhängigkeit von der Anzahl der verdeckten Neuronen, mittlerer quadratischer Fehler für normierte Daten

Abbildung 55 zeigt die Korrelation der Targetwerte und Netzoutputs der bestangepaßten MLFN für Durchfluß und Wasserstand am Pegel ErlIn. Auch hierbei zeigt sich, daß der Durchfluß von den neuronalen Netzen leicht besser wiedergegeben wird. Die größten Abweichungen treten allerdings wiederum bei niedrigeren Wasserständen auf, was für die Hochwasserprognose nicht nachteilig ist. Dieser Effekt ist allerdings wesentlich geringer als bei den HD-Netzen (Abbildung 34). Bei der hydrodynamischen Simulation mit den aneinandergfügten Ganglinienausschnitten können durch Sprünge in den Zuflußganglinien Ungenauigkeiten im Anfangsbereich der resultierenden Ganglinie am Zielpiegel auftreten, die das neuronale Netz so nicht wiedergibt. Eine Berücksichtigung der Zuflüsse aus dem Zwischeneinzugsgebiet mit

Ergebnisse der Prognose-ANN

geringeren Laufzeiten zum Zielpegel bei den HDNA-Netzen dämpft diesen Effekt. Diese Ungenauigkeiten treten vor allem am Anfang eines Hochwasserereignisses auf, wo die niedrigen Durchflüsse und Wasserstände zu finden sind.

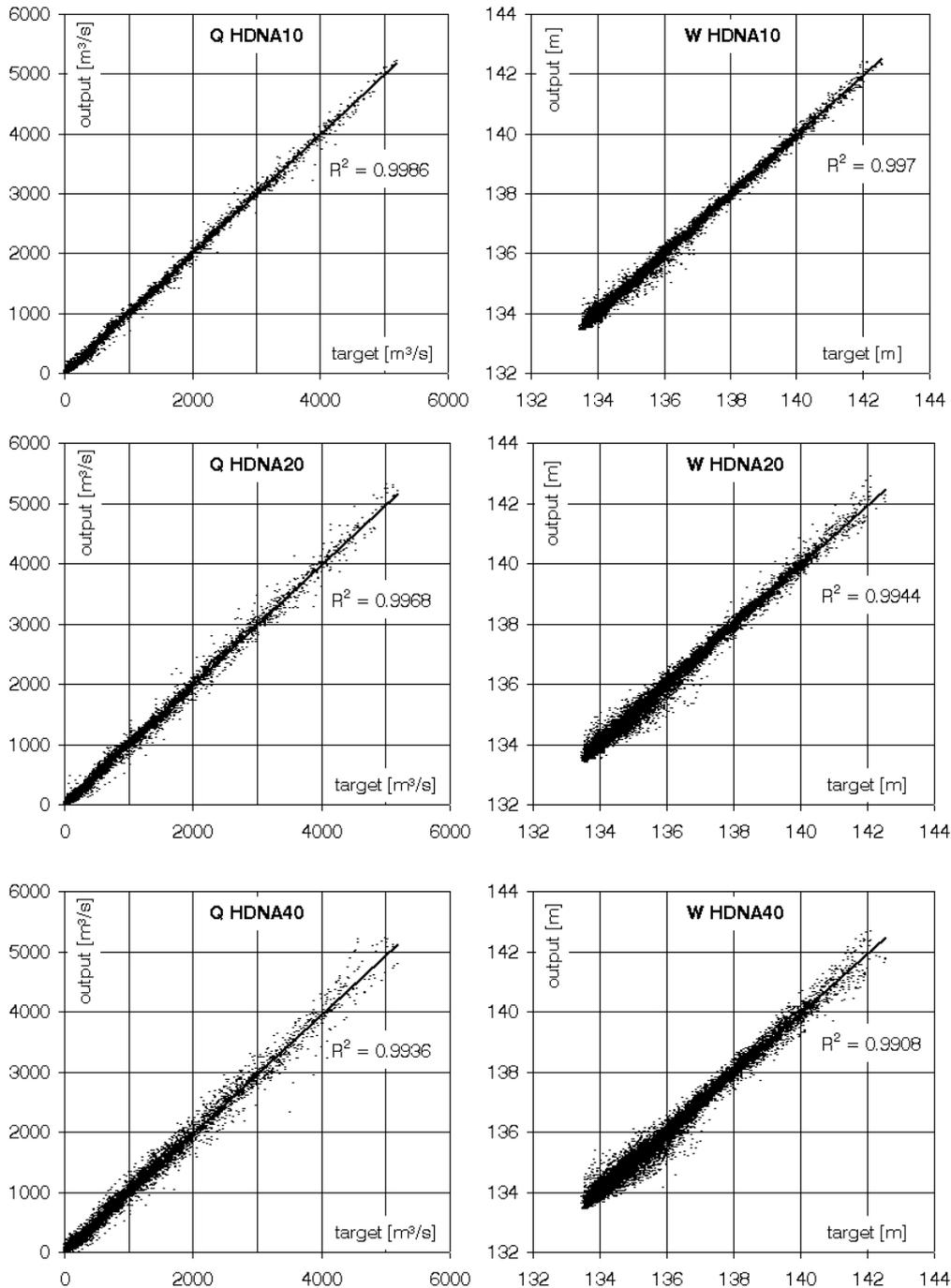


Abbildung 55: Testdaten für die HDNA-Netze: HDNA10 mit 18, HDNA20 mit 23 und HDNA40 mit 24 verdeckten Neuronen, Durchfluß und Wasserstand

Die Darstellung der Sequenz künstlicher Hochwasserereignisse in Abbildung 56 bestätigt die Prognosegüte der HDNA-Netze. Bei dieser Auswahl ist der Effekt der bei

Ergebnisse der Prognose-ANN

den HD-Netzen beobachteten systematischen Unterschätzung der Wasserstände und Durchflüsse zu Beginn des Anstieges der Hochwasserwelle gänzlich verschwunden. Speziell der Durchflußverlauf des hydrodynamischen Modells ist mit den MLFN für alle drei Szenarien hervorragend getroffen. Der Fehler für den Scheitelwasserstand ist bis auf das höchste hier dargestellte Hochwasser jeweils kleiner als 20 Zentimeter. (Das Hochwasser im August 2002 liegt in der Größenordnung des dritthöchsten dargestellten Peaks bei ca. $t=480h$.)

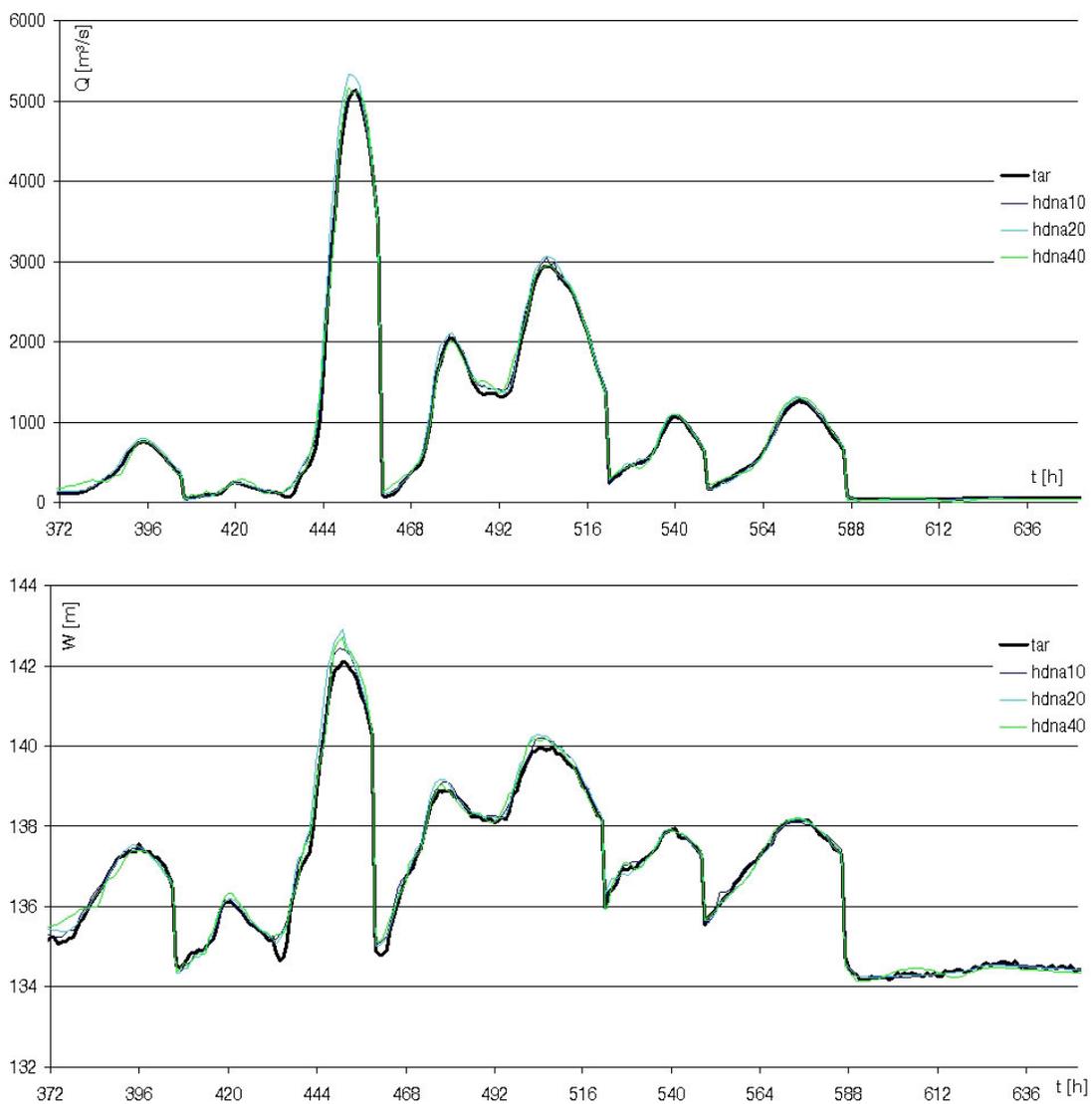


Abbildung 56: HDNA-MLFN - Darstellung einzelner Szenarien für Durchfluß und Wasserstand

5.2.5 SOM für die HDNA-Szenarien

Die Merkmalsvektoren der SOM setzen sich zusammen aus den Inputvektoren der HDNA-MLFN und dem korrespondierenden Durchfluß- beziehungsweise Wasserstandswert am Pegel ErlIn. Abbildung 57 zeigt die Performance der trainierten Netze in Abhängigkeit vom Outputdimensionsgewicht. Dabei zeigt sich für alle SOM ein Optimum für $ODW = 2$. Jedoch ist der Effekt dieses Gewichtsfaktors geringer als bei einem NA-SOM (Abbildung 48 und Abbildung 51). Vergleicht man Abbildung 57 mit den Ergebnissen des Trainings der SOM für die HD-Szenarien (Abbildung 36), zeigt sich für alle Zwischengebietsgrößen ein deutlich größerer mittlerer quadratischer Fehler. Die Hinzunahme der Merkmale für die Niederschlags-Abfluß-Prozesse bringt demnach keine Verbesserung. Offenbar liegt die Ursache darin, daß SOM klassifizierende Netze sind. Der Zusammenhang kann nur als Muster abgespeichert werden. Die größere Wichtung des zu bestimmenden Wertes durch die Einführung des Outputdimensionsgewichtes ist dabei offenbar nicht ausreichend. Jedoch kann dieser Effekt nicht ausschließlich mit der Anzahl der Dimensionen des Merkmalsvektors erklärt werden, da die Performance der SOM ausschließlich für Niederschlags-Abfluß-Prozesse (Abbildung 48 und Abbildung 51) bei einer geringeren Anzahl an Merkmalen noch schlechter ausfällt. Offensichtlich findet eine SOM unter den HD-Merkmalen zuverlässiger wiederkehrende Muster. Eine Niederschlagsreihe auf Stundenwertbasis ist im allgemeinen stark verrauscht und unterliegt kaum einer Autokorrelation. Die daraus durch eine Transformation im Einzugsgebiet resultierenden einzelnen Durchflußwerte einer Hochwasserganglinie zeigen untereinander einen viel klareren Zusammenhang. Eine Hochwasserprognose aus Niederschlagsdaten ist mit einer Klassifikation und Mustererkennung also schwer möglich. Auf diesem Feld besitzen überwacht angelegte neuronale Netze zur Funktionsapproximation deutliche Vorzüge.

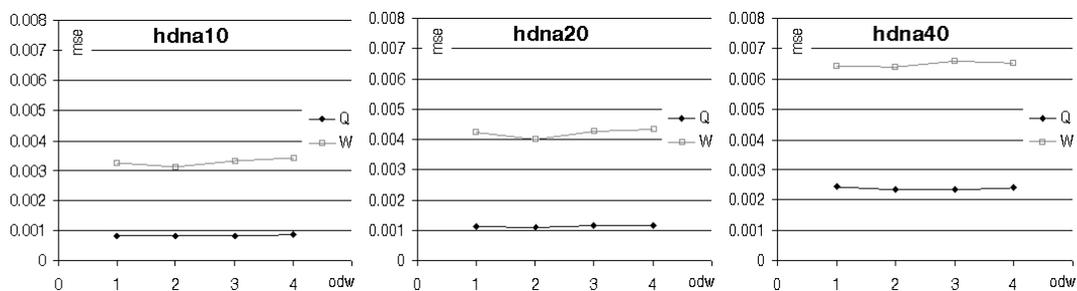


Abbildung 57: : Performance der SOM (150 x 40) für die Szenarien HDNA10, HDNA20 und HDNA40 in Abhängigkeit vom Outputdimensionsgewicht, mittlerer quadratischer Fehler für normierte Daten

Ergebnisse der Prognose-ANN

In Abbildung 58 ist die Ausgabe der SOM für die bekannten Szenarien abgebildet. Die trainierten SOM bilden die Ganglinien in ihrer groben Form recht zuverlässig ab, im Detail jedoch ist die Prognosegüte nicht vergleichbar mit den Anpassungen der MLFN.

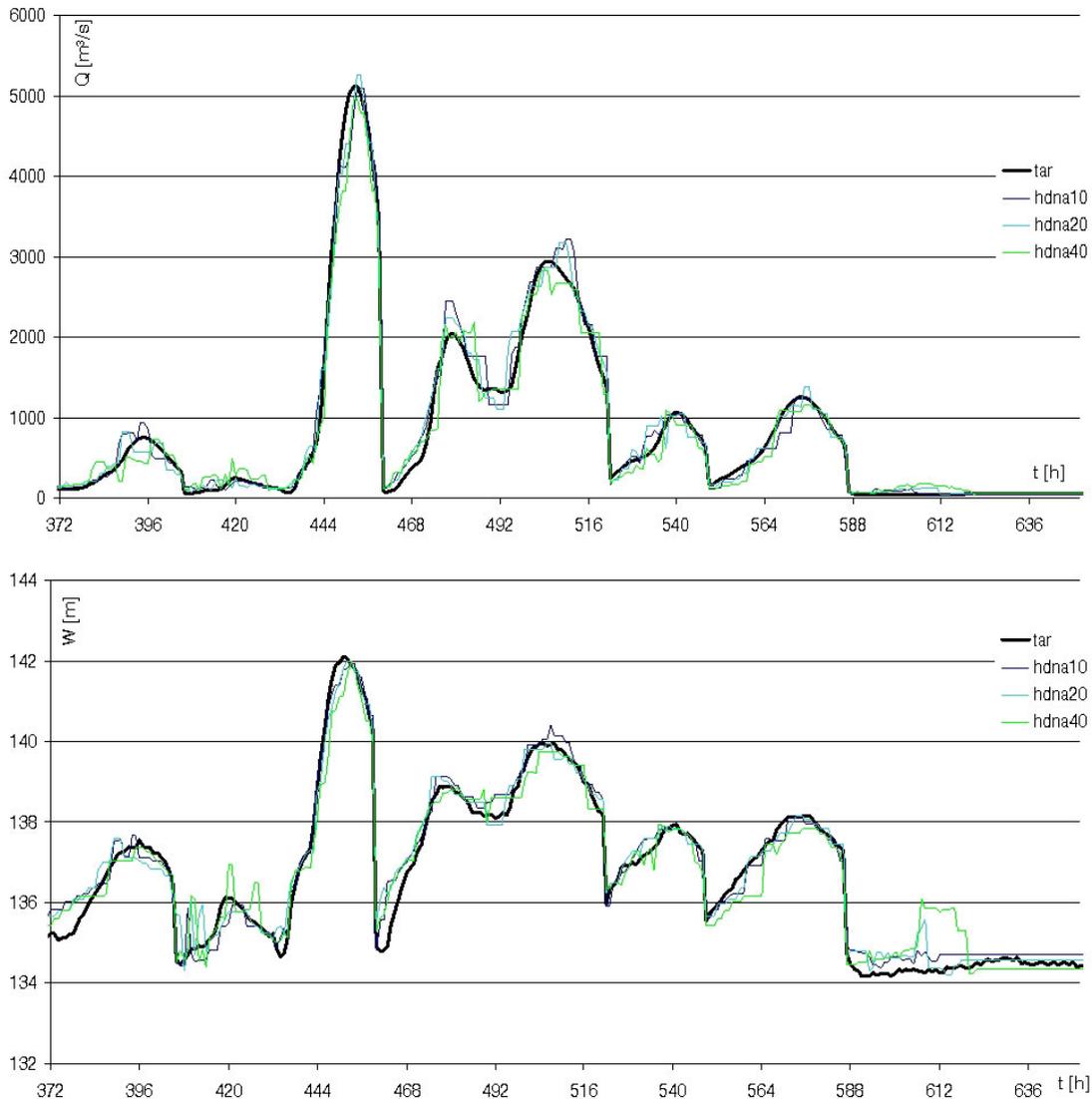


Abbildung 58: HDNA-SOM - Darstellung einzelner Szenarien für Durchfluß und Wasserstand

5.3 Vergleich der Prognose-MLFN

Vergleicht man die Performance der jeweils besten Anpassungen für die sechs Szenarien (Abbildung 59), so fällt auf, daß der Fehler ohne die Verwendung der NA-Merkmale überproportional zur Zwischengebietsgröße steigt. Prinzipiell ist die Differenz zwischen den HD- und HDNA-Szenarien darauf zurückzuführen, daß bei den HD-Szenarien für einen Teil des gesamten Einzugsgebietes des Pegels ErlIn keine expliziten Informationen bereitgestellt werden. Die Güte der Anpassung hängt davon

Ergebnisse der Prognose-ANN

ab, inwieweit der Output aus diesen Flächen mit den gegebenen Abflüssen korreliert. Ist der Anteil dieser Flächen am Gesamtgebiet sehr klein und stark an die Netzinputs gekoppelt, wovon nur bei kleinen Zwischeneinzugsgebieten auszugehen ist, kann auf die NA-Merkmale für diesen Teil verzichtet werden. Der mittlere quadratische Fehler für das HD40-Szenario ist mit einigem Abstand am größten. Dies ist insofern nicht überraschend, da hierbei der Pegel ErlIn ausschließlich aus der Durchflußganglinie des Pegels Kriebstein prognostiziert wird. Dabei ist der absolute Fehler am stärksten von der räumlichen Variabilität der generierten meteorologischen Inputs abhängig. Demgegenüber sind die Fehler für die HDNA-MLFN davon unabhängig – freilich unter der Bedingung, daß die räumliche Variabilität des Niederschlages innerhalb des Zwischengebietes klein genug ist, daß die Verwendung eines Gebietsmittels sinnvoll ist.

Nach Abbildung 46 beträgt der mittlere quadratische Fehler für ein reines NA-Netz für den Pegel Nossen ohne eine räumliche Untergliederung (Einzugsgebietsfläche entspricht annähernd dem Szenario HDNA20) für die Testdaten ungefähr 0.0004 (ohne Berücksichtigung des Fehlers des Zustandsmerkmals). Das MLFN für das Szenario HDNA20 zeigt also auch eine deutlich bessere Prognosegüte, als ein reines NA-Netz für ein Gebiet dieser Größe.

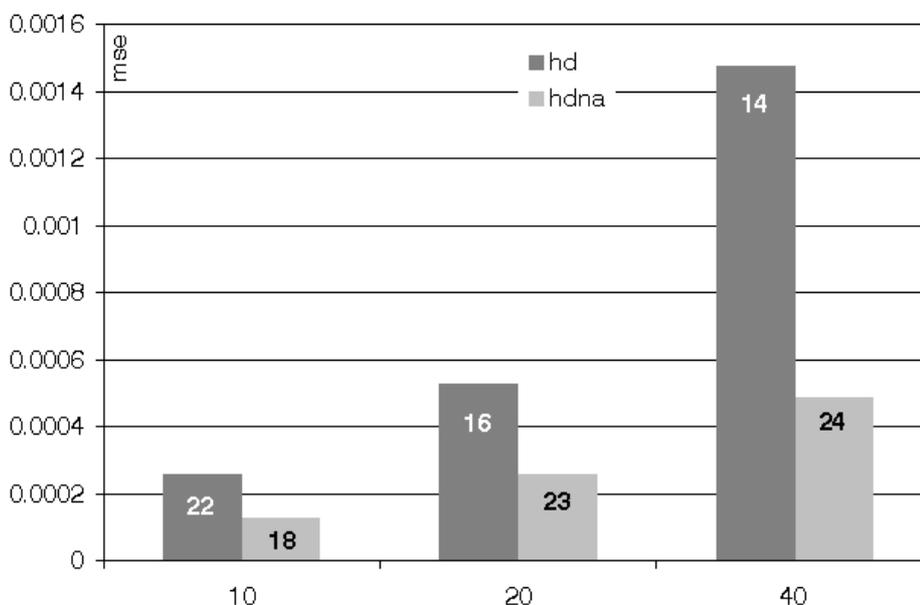


Abbildung 59: Vergleich der Performance der bestangepaßten MLFN für die Szenarien (Anzahl der verdeckten Neuronen am Säulenende), mittlerer quadratischer Fehler für normierte Daten

Ergebnisse der Prognose-ANN

Die trainierten und getesteten MLFN wurden auf das Hochwasserereignis im Juli 1996 (Abbildung 60) angewendet. Dieses Ereignis wurde ausgewählt, da es zwei aufeinanderfolgende Scheitel aufweist. Dies stellt insbesondere für die HDNA-ANN eine besondere Herausforderung dar, da die Ausprägung des zweiten Scheitels maßgeblich durch eine Aufwechtlung des Gebietes während des Ereignisses beeinflusst wird. In Abbildung 60 ist zu erkennen, daß der zweite Scheitel mehr als 4 Tage nach dem durch das Zustandsmerkmal charakterisierten Ereignisbeginn eintritt. Somit kann auch eine Aussage über die Prognosegüte bei größeren Vorhersageweiten getroffen werden.

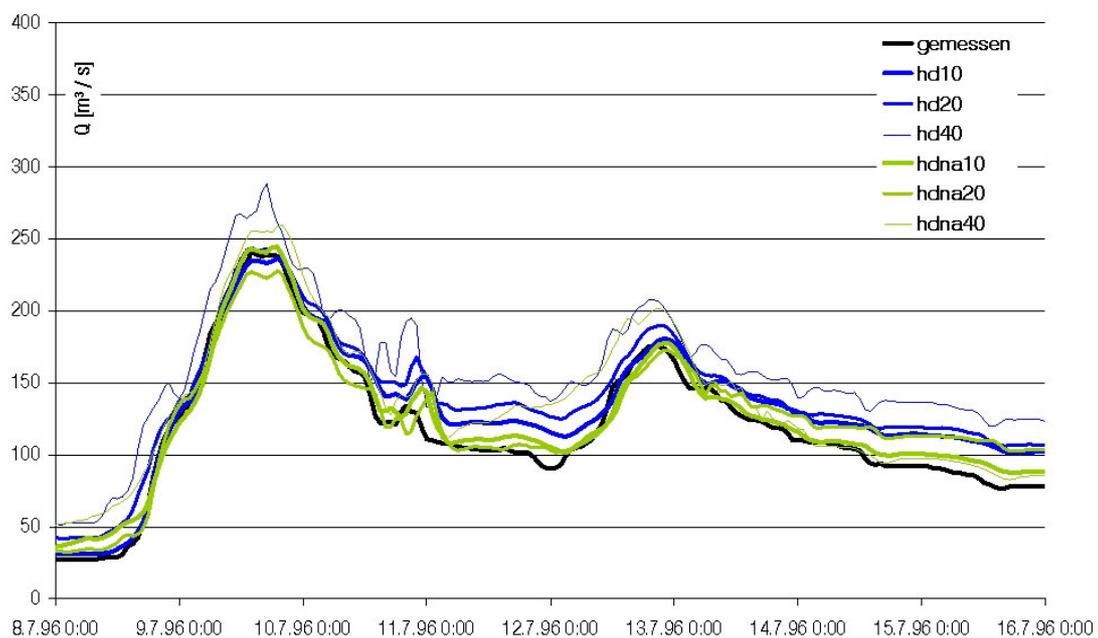


Abbildung 60: Validierung für das Hochwasserereignis im Juli 1996

Alle Netze bilden den Verlauf des Hochwassers sehr zufriedenstellend ab. Das MLFN für das HD40-Szenario, welches den Pegel Erlin ausschließlich aus der Ganglinie des Pegels Kriebstein vorhersagt, ist erwartungsgemäß am stärksten fehlerbehaftet. Eine Einbeziehung der Niederschlags-Abfluß-Prozesse im Zwischeneinzugsgebiet (40 % der Gesamtfläche) mit dem HDNA40-Szenario zeigt hier eine deutliche Verbesserung. Die anderen Anpassungen der HD-Szenarien (HD10 und HD20, keine Merkmale für die NA-Prozesse im Zwischengebiet) bilden den Durchflußverlauf hervorragend ab. Die Einbeziehung der Merkmale für die NA-Prozesse im Zwischengebiet (HDNA20 und HDNA40) hat hier kaum einen Einfluß auf die Prognosegüte.

6 Diskussion

Da die Anzahl der realen Hochwasserereignisse zur Verifizierung der gesamten Methodik anhand des gewählten Einzugsgebietes relativ begrenzt ist, soll an dieser Stelle eine theoretische Bewertung und Diskussion erfolgen.

An erster Stelle muß konstatiert werden, daß ein mit dieser Methode erstelltes neuronales Netz niemals besser sein kann als ein prozeßbasiertes Modell. Das hydrodynamisch-numerische Modell ist in hohem Grade physikalisch fundiert, da sämtliche relevante Annahmen für Geometrie und Rauigkeiten gemessenen Größen entspringen und nur eine geringfügige Kalibrierung stattfand. Die Abweichungen von gemessenen Ganglinien am Zielpegel in ErlIn bewegen sich innerhalb der Unsicherheiten der Eingangsdaten durch fehlerhafte Wasserstands-Abfluß-Beziehungen. Durch die Berücksichtigung der Massenerhaltung in Kombination mit relativ kurzzeitigen Retentionseffekten wirken sich Fehler in den Meßdaten wenig auf das parametrisierte Modell aus. Damit ist die Beurteilung der Modellgüte insofern schwierig, da davon ausgegangen werden kann, daß im Bereich extrem hoher Abflüsse die Modellfehler geringer sind als die Meßfehler bzw. Fehler durch die WQ-Beziehung.

Die Güte des Niederschlags-Abfluß-Modells wurde hier nicht gesondert untersucht, da eine Parametrisierung nach Cullmann (2007) übernommen wurde. Aber bei der Beurteilung der Prognosegüte der neuronalen Netze kann dieser Faktor natürlich nicht ausgeblendet werden. In Anbetracht der geringen Anzahl der mit Daten belegten extremen Hochwasserereignisse und der Komplexität der unterschiedlichen ineinandergreifenden Prozesse ist grundsätzlich von einer geringeren Anpassungsgüte im Hochwasserbereich gegenüber normalen Abflüssen auszugehen. Die Massenbilanz wird zwar bei der NA-Modellierung auch gewahrt (mit leichten Einschränkungen bedingt durch den Wechsel zwischen den beiden Parametrisierungen), jedoch ist die Menge des während Hochwasserereignisses abfließenden Wassers durch längere Speicherung, langsame Abflußkomponenten und Verdunstung nicht so eng an den Systeminput gekoppelt, wie bei der hydrodynamischen Modellierung. Damit wirken sich fehlerhafte Eichdaten für die seltenen Extremereignisse stärker auf die Güte der Modellparametrisierung aus.

Da der Schwerpunkt dieser Arbeit auf der Abbildung der deterministischen Modelle durch künstliche neuronale Netze liegt, sollen die Ergebnisse bezüglich der Größe des

Zwischeneinzugsgebietes auch auf dieser Ebene vergleichbar sein. Aus diesem Grunde wurden für die Einzugsgebiete der Pegel Nossen und Niederstregis auch für den Fall, daß sie Bestandteil des quasi unbeobachteten Zwischengebietes waren, die bekannten Parametrisierungen nach Cullmann (2007) verwendet. Anderenfalls wäre das Ergebnis wenig aufschlußreich für die Güte der Methodik zur Einbeziehung der Niederschlags-Abfluß-Merkmale in das künstliche neuronale Netz. Vielmehr hätte es auf die Übertragbarkeit von Parametersätzen eines NA-Modells auf unbeobachtete Einzugsgebiete schließen lassen. Der Fehler der Parametrisierungen des NA-Modells für die Szenarien mit 20% und 40% der Zwischeneinzugsgebietsgröße ist also wesentlich geringer, als er bei einer Übertragung von Parametersätzen auf unbeobachtete Einzugsgebiete in der Tat zu veranschlagen wäre.

Relevant für die Beurteilung der Methodik sind jedoch die Auswahl der Merkmale und die Repräsentanz der Trainingsdaten. Die Aussagekraft der Merkmale läßt sich am besten durch die Performance der erstellten Netze einschätzen, worauf im vorhergehenden Kapitel ausführlich eingegangen wurde. Darüberhinaus ist die Methodik der Merkmalsauswahl entsprechend der ablaufenden Prozesse theoretisch und in Abstimmung mit den deterministischen Modellen begründet und somit auf andere Einzugsgebiete übertragbar.

Die Trainingsdaten resultieren aus meteorologischen Inputszenarien. Diese wiederum basieren zunächst auf realen Vorfeuchteszenarien. Jeder generierte, hochwasserauslösende Niederschlag hat dabei eine ausschließlich natürliche Vorgeschichte, was garantiert, daß kein simuliertes Hochwasserereignis von vorhergehenden simulierten Ereignissen beeinflusst wird. Diese Vorfeuchteszenarien entstammen einer siebenundvierzigjährigen Reihe gemessener Daten. Es kann davon ausgegangen werden, daß damit die ganze Bandbreite möglicher Zustände abgedeckt ist. Da dabei auch Gebietszustände nach realen Hochwasserereignissen eingeschlossen sind, beinhaltet diese Grundlage auch extrem hochwasserfördernde Konstellationen.

Die generierten, hochwasserauslösenden Niederschläge umfassen in Dauer und Summe (Abbildung 26) alle realistischen bis hin zu extremen Varianten. Die daraus resultierenden Hochwasserwellen an den einzelnen Gebietsauslässen decken damit auch den real möglichen Bereich ab. Da hier ein unbeobachtetes Zwischengebiet angenommen wurde, konnte allerdings weder eine räumliche Differenzierung der Niederschlagssumme (mit Ausnahme einer invariablen Skalierung nach KOSTRA), noch eine mögliche Zugbewegung des Niederschlagsfeldes über ein

Diskussion

Teileinzugsgebiet betrachtet werden. Da die Niederschlags-Abfluß-Prozesse nur eine untergeordnete Rolle spielen, ist dies zumindest für Zwischeneinzugsgebiete sicherlich zulässig. Da die Methodik für die Festlegung der NA-Merkmale prinzipiell durch die Einführung einer räumlichen Differenzierung auf die Erstellung eines reinen Niederschlags-Abfluß-ANN erweiterbar ist, sollte dieser Punkt jedoch nicht unerwähnt bleiben.

Die einzelnen Teilgebiete wurden gleichmäßig überregnet, aber – auf das gesamte Einzugsgebiet des Pegels ErlIn bezogen – insgesamt mit einer Überzugsgeschwindigkeit von mindestens 5 km h^{-1} in alle Richtungen und einer Variation in der Summe von 30% untereinander versehen. Wenngleich diese Annahmen sinnvoll erscheinen, kann nicht bewiesen werden, daß damit möglicherweise realistische Niederschlagskonstellationen ausgeschlossen werden. In Ermangelung einer ausreichenden natürlichen Datenbasis ist es allerdings schwer möglich, die Unsicherheiten, die aus diesen oder jeglichen anderen Annahmen hinsichtlich der Generierung der Niederschlagsszenarien resultieren, zu quantifizieren.

In dieser Arbeit wurde vereinfachend davon ausgegangen, daß bereits Effektivniederschläge als Prognose vorliegen. Ist dies in der Praxis nicht der Fall, muß eine gesonderte Betrachtung des Schneespeichers erfolgen. Die aktuelle Schneespeicherung in den Wintermonaten hängt wie die Gebietsfeuchte sehr von der Vorgeschichte ab und ist darüber hinaus durch die Abhängigkeit von der Morphologie und der geodätischen Höhe noch stärker räumlich gegliedert. Das erschwert den Einsatz von ANN für die Abbildung des Schneespeichers. Da das hier von WaSiM-ETH verwendete Tagesgradverfahren ein relativ einfacher und wenig rechenintensiver Ansatz ist, bringt eine Verwendung von ANN hierfür auch kaum Vorteile mit sich.

7 Zusammenfassung

In Mitteleuropa zählen Hochwasserereignisse zu den schwersten Naturkatastrophen. Speziell in schnell reagierenden Mittelgebirgseinzugsgebieten stellt die Hochwasserprognose noch immer eine große Herausforderung dar. In Flußabschnitten, die durch Rückstauinflüsse und Retentionseffekte in überfluteten Vorländern geprägt sind, kommt eine adäquate hydrodynamische Modellierung bislang nur zur Rekonstruktion abgelaufener Ereignisse zum Einsatz. Der Grund dafür sind die noch immer großen Anforderungen derartiger Modellsysteme an die Rechenleistung. Dies gilt insbesondere für Versuche der Einbeziehung der Unsicherheiten -- beispielsweise durch Monte-Carlo-Simulationen. Eine Anwendung rein datengetriebener Ansätze zur Hochwasserprognose unter alleiniger Verwendung beobachteter Hochwasserereignisse ist nicht zielführend, da die natürliche Datenbasis für Extremereignisse in jedem Falle zu gering ist.

Diese Arbeit zeigt einen Ausweg aus diesem Dilemma auf, der darauf basiert, physikalisch begründete Modellierung mit künstlichen neuronalen Netzen zu kombinieren und die Defizite der beiden Herangehensweisen zu überwinden. Ein eindimensionales hydrodynamisches Modell beinhaltet die genaue Kenntnis der Geometrie des Flußlaufes und der Vorländer und berücksichtigt die physikalischen Prozesse des Wellenablaufes. Damit ist eine Grundlage für die Erstellung einer Datenbasis, die die weite Spanne theoretisch möglicher Hochwasserereignisse abdeckt, durch umfangreiche Szenarienrechnungen geschaffen. Mit dieser Datenbasis können dann künstliche neuronale Netze trainiert werden, die auch im Bereich extremer Hochwasserereignisse zuverlässige Prognosen liefern.

Desweiteren wurde die Methodik auf die Einbeziehung von Merkmalen für die Niederschlags-Abfluß-Prozesse im unbeobachteten Zwischengebiet zur Berücksichtigung lateraler Zuflüsse entlang der modellierten Fließstrecken erweitert. Die Datenbasis wurde hierfür mit einem Niederschlags-Abfluß-Modell erstellt.

Die Verwendung von deterministischen Modellen garantiert jedoch nicht nur eine verlässliche Datenbasis für das Training der Netze, sondern ermöglicht – sowohl für die Niederschlags-Abfluß-Prozesse, als auch für die hydrodynamischen Prozesse – Analysen betreffs der Sensitivität der Inputdaten auf das Modellergebnis. Mit Hilfe dieser werden wichtige Informationen zur Findung der relevanten Merkmale erlangt.

Zusammenfassung

Ein Schlüssel für die erfolgreiche Eingliederung der Niederschlags-Abfluß-Prozesse in das Prognosenetz ist die Einführung eines einzigen Zustandsmerkmals, welches die gesamte meteorologische Vorgeschichte des Ereignisses zur Charakterisierung des Gebietszustandes vereinigt. Dieses Merkmal ergibt sich als Scheiteldurchfluß der Gebietsreaktion auf ein einheitliches Niederschlagsereignis. Damit mißt dieses Zustandsmerkmal den Gebietszustand direkt an seiner Wirkung auf ein Hochwasserereignis und ist gleichzeitig unabhängig vom hochwasserauslösenden Niederschlag. Für die Berechnung des Zustandsmerkmals werden keine prognostizierten Daten verwendet, womit sich die datenbezogenen Unsicherheiten gering halten. Ensembleprognosen für ein Hochwasserereignis können damit – neben den variierenden Niederschlagszeitreihen für die Ereignismerkmale – auf ein Zustandsmerkmal zurückgreifen, welches die gesicherte Information der Vorgeschichte speichert.

Diese Methodik wurde unter Verwendung der deterministischen Modelle WaSim-ETH zur Beschreibung der Niederschlags-Abfluß-Prozesse und HEC-RAS für die Wellenablaufmodellierung am Einzugsgebiet der Freiburger Mulde für den Pegel Erlin erprobt. Dabei wurden verschiedene Szenarien bezüglich der Zwischeneinzugsgebietsgröße und der Berücksichtigung von Merkmalen für die Niederschlags-Abfluß-Prozesse betrachtet.

Desweiteren wurden zwei hinsichtlich ihrer Art des Lernens verschiedene Typen künstlicher neuronaler Netze getestet. Bei der relativ hohen Anzahl von Merkmalen zeigen Multilayer-Feedforward-Netze als Vertreter des überwachten Lernens durch die eindeutige Ausweisung eines Targetwertes deutlich bessere Anpassungen als die unüberwacht angelegten selbstorganisierenden Merkmalskarten, die keinen expliziten Output, sondern nur gleichberechtigte Merkmale kennen. Jedoch können SOM Aufschlüsse über den Zusammenhang der verschiedenen Größen geben und damit bei der Auswahl oder Zusammenfassung von Merkmalen behilflich sein.

Die Methodik für die Einbeziehung der Niederschlags-Abfluß-Merkmale kann prinzipiell auch auf die Erstellung eines reinen Niederschlag-Abfluß-ANN erweitert werden. Hierzu wäre es nötig, die räumliche Differenzierung des Niederschlags und gegebenenfalls des Gebietszustandes zu berücksichtigen.

Symbolverzeichnis

α	Lernrate	Q_{\max}	Scheiteldurchfluß [$\text{m}^3 \text{s}^{-1}$, mm h^{-1}]
τ	Grenzparameter der Sigmoiden	Q_n	Zufluß aus Nebenfluß [$\text{m}^3 \text{s}^{-1}$]
Θ	Wassergehalt [-]	Q_o	Zufluß, obere Randbedingung [$\text{m}^3 \text{s}^{-1}$]
a	Outputwert des Neurons	Q_w	gewichteter Durchfluß [$\text{m}^3 \text{s}^{-1}$, mm h^{-1}]
\bar{a}	Output einer Neuronenschicht	Q_{ZEG}	Zufluß aus Zwischeneinzugsgebiet [$\text{m}^3 \text{s}^{-1}$]
A	durchströmte Fläche [m^2]	R	Globalstrahlung [W m^{-2}]
b	Biaswert eines Neurons	S_f	Reibungsgefälle [-]
\bar{b}	Biasvektor	t	Zeit [d, h]
\bar{b}_{HIDDEN}	Biasvektor der verdeckten Schicht	t_0	Ereignisbeginn [d, h]
\bar{b}_{OUTPUT}	Biasvektor der Outputschicht	t_h	Steilheitsparameter der Sigmoiden
DOY	Tag des Jahres	t_i	Zeitpunkt eines Niederschlagswertes [h]
ETP	potentielle Verdunstung [mm , mm h^{-1}]	$t_{\text{inf,max}}$	maximale Einflußzeit [h]
ETR	reale Verdunstung [mm , mm h^{-1}]	$t_{\text{inf,min}}$	minimale Einflußzeit [h]
f	Transferfunktion	t_j	Prognosezeitschritt [h]
F	relative Luftfeuchte [%]	T	Temperatur [$^{\circ}\text{C}$]
g	Gravitationskonstante [m s^{-2}]	w	Wichtungsfaktor
n	gewichtete Summe der Neuroneninputs	w	Windgeschwindigkeit [m s^{-1}]
N	Niederschlag [mm , mm h^{-1}]	\vec{w}	Gewichtsvektor eines Neurons
N_{eff}	Effektivniederschlag [mm , mm h^{-1}]	w_i	i-ter Gewichtungsfaktor im Neuron
ODW	Outputdimensionsgewicht	${}_i\vec{w}$	Ortsvektor des i-ten Neurons einer SOM
p_i	i-ter Inputwert des Neurons	V	Volumen [m^3]
\vec{p}	Inputvektor eines Neurons oder künstlichen neuronalen Netzes	W	Gewichtsmatrix
q	Trainingsschritt	W	Wasserstand [m]
q	spezifischer Fluß [m s^{-1}]	W_{HIDDEN}	Gewichtsmatrix der verdeckten Schicht
q_{in}	Zufluß in Bodenschicht [m s^{-1}]	W_{OUTPUT}	Gewichtsmatrix der Outputschicht
q_l	lateraler Zufluß [$\text{m}^2 \text{s}^{-1}$]	x	Fließweg [m]
q_{out}	Abfluß aus Bodenschicht [m s^{-1}]	\vec{y}	Targetvektor
Q	Durchfluß [$\text{m}^3 \text{s}^{-1}$, mm h^{-1}]	z	geodätische Höhe [m]
Q_g	Durchfluß nach Parametersatz für extreme HQ [$\text{m}^3 \text{s}^{-1}$, mm h^{-1}]	ZM	Zustandsmerkmal
Q_k	Durchfluß nach Parametersatz für normale HQ [$\text{m}^3 \text{s}^{-1}$, mm h^{-1}]		

Abbildungen

Abbildung 1: Lernen bei datengetriebener Modellierung (nach Solomatine, 2002).....	12
Abbildung 2: Aus dem Defizit an Trainingsdaten abgeleitete allgemeine Methodik	13
Abbildung 3: räumliche Gliederung des Gesamteinzugsgebietes nach dominanten Prozessen, schematisch	14
Abbildung 4: Gesamtmethodik der Einzugsgebietsmodellierung und Erstellung eines ANN für den Wellenablauf unter Berücksichtigung des Zwischeneinzugsgebietes.....	16
Abbildung 5: Methodik der Erstellung der Subnetze für die Berechnung des Zustandsmerkmals.....	17
Abbildung 6: Gesamtmethodik der Netzanwendung für HDNA-ANN	17
Abbildung 7: Neuron mit R Inputs (aus Hagan et al., 1996)	20
Abbildung 8: Multilayer-feedforward-Netz nach Hagan et al (1996)	22
Abbildung 9: SOM, zweidimensionale Netztopologien: links hexagonal, rechts viereckig.....	24
Abbildung 10: HEC-RAS, Erscheinungsbild.....	28
Abbildung 11: Einflußbereich der Durchflußganglinie am Oberliegerpegel auf einen Wert der Ganglinie am Zielpiegel.....	35
Abbildung 12: Bestimmung der Parameter $t_{inf,min}$ und $t_{inf,max}$ der Einflußspanne	35
Abbildung 13: Datenfluß bei der Erstellung des Trainingsdatensatzes für ein neuronales Netz.....	36
Abbildung 14: Prinzip der Zusammenfassung der Niederschläge zu Ereignismerkmalen (ZM = Zustandsmerkmal nach 3.6.2).....	40
Abbildung 15: Datenfluß in WaSim-ETH (vereinfachtes Schema) und Unterteilung in Subnetze zur Berechnung des Zustandsmerkmals ZM	42
Abbildung 16: Datenfluß für Erstellung der Trainingsdaten und Training des Subnetzes 1	45
Abbildung 17: Datenfluß für Erstellung der Trainingsdaten und Training des Subnetzes 2.....	47
Abbildung 18: Datenfluß für das Prognosenetz.....	49
Abbildung 19: Höhen und Gefälle im Untersuchungsgebiet.....	50
Abbildung 20: Lithofaziesseinheiten im Untersuchungsgebiet	52
Abbildung 21: Bodenkarte des Untersuchungsgebietes.....	53

Abbildung 22: Landnutzung des Einzugsgebietes der Freiburger Mulde	54
Abbildung 23: Untergliederung des Einzugsgebietes	56
Abbildung 24: Einzugsgebietsgliederung entsprechend der Fälle HD10/HDNA10 (links), HD20/HDNA20 (mitte), HD40/HDNA40 (rechts)	58
Abbildung 25: Ganglinien der Hochwasser im Dezember 1974, August 1983 und Januar 1987 an den Pegeln	60
Abbildung 26: Szenarienrechnung, Histogramm für die Niederschlagsmenge	62
Abbildung 27: Hochwasserereignisse Dezember 1974 und Juli 1996, Eichung	66
Abbildung 28: Hochwasserereignisse August 1983 und März 1998, Validierung	67
Abbildung 29: sigmoide Wichtung der Abflüsse für die verschiedenen Modellparametrisierungen.....	69
Abbildung 30: Berechnung der Zustandsmerkmale und Hochwasserszenarien mit WaSim-ETH	70
Abbildung 31: Histogramm für das Zustandsmerkmal.....	71
Abbildung 32: Histogramm für den Scheiteldurchfluß der Hochwasserszenarien am Pegel Nossen, logarithmische Darstellung..	72
Abbildung 33: : Performance der MLFN für die Szenarien HD10, HD20 und HD40 in Abhängigkeit von der Anzahl der verdeckten Neuronen, mittlerer quadratischer Fehler für normierte Daten.....	74
Abbildung 34: Testdaten für die HD-Netze: HD10 mit 22, HD20 mit 16 und HD40 mit 14 verdeckten Neuronen, Durchfluß und Wasserstand.....	75
Abbildung 35: HD-MLFN - Darstellung einzelner Szenarien für Durchfluß und Wasserstand	76
Abbildung 36: Performance der SOM (150 x 40) für die Szenarien HD10, HD20 und HD40 in Abhängigkeit vom Outputdimensionsgewicht, mittlerer quadratischer Fehler für normierte Daten.....	77
Abbildung 37: Testdaten für die HD-Netze: HD10 mit 22, HD20 mit 16 und HD40 mit 14 verdeckten Neuronen, Durchfluß und Wasserstand.....	77
Abbildung 38: HD-SOM - Darstellung einzelner Szenarien für Durchfluß und Wasserstand	78
Abbildung 39: SOM für Durchfluß, Szenario HD20, Ansicht für die Merkmale M4, M10 und M12, braun: maximal, blau: minimal.....	79
Abbildung 40: Performance des Subnetzes 1 in Abhängigkeit von der Anzahl der Neuronen in der verdeckten Schicht für die Testdaten, mittlerer quadratischer Fehler für normierte Daten.	81

Abbildung 41: Testdaten für Subnetz1, Variante mit 11 Neuronen	81
Abbildung 42: Sensitivität eines Hochwassers (1. September) auf zurückliegende Niederschlagsimpulse von 4 mm h^{-1} für 24 h (1. Mai, 1. Juni, 1. Juli, 1. August und 18. August).....	83
Abbildung 43: Performance des Subnetzes 2 in Abhängigkeit von der Anzahl der Neuronen in der verdeckten Schicht, mittlerer quadratischer Fehler für normierte Daten.	85
Abbildung 44: Testdaten für Subnetz 2, Variante mit 2 Neuronen	85
Abbildung 45: Korrelation eines Niederschlagswertes mit im Abstand t darauffolgenden Abflußwerten, Realisierungen für verschiedene Zeitpunkte ab Ereignisbeginn, Maximum (gestrichelt fett) und Mittelwert (fett)	87
Abbildung 46: Performance des MLFN für die NA-Prozesse in Abhängigkeit von der Anzahl der Neuronen in der verdeckten Schicht, mittlerer quadratischer Fehler für normierte Daten.	89
Abbildung 47: Testdaten für das NA-MLFN, Abfluß in mm h^{-1} , Variante mit 9 verdeckten Neuronen.....	89
Abbildung 48: Performance des SOM (59 x 38) für die NA-Prozesse in Abhängigkeit vom Outputdimensionsgewicht, mittlerer quadratischer Fehler für normierte Daten.	90
Abbildung 49: Testdaten für das NA-SOM, Variante ODW=3, Abfluß in mm h^{-1}	91
Abbildung 50: Ansicht der SOM für die Merkmale M1 (links, Zustandsmerkmal) und M21 (rechts, Abfluß), braun: maximal, blau: minimal	92
Abbildung 51: Performance der SOM für den reduzierten Merkmalsatz (14) für die NA-Prozesse in Abhängigkeit vom Outputdimensionsgewicht, mittlerer quadratischer Fehler für normierte Daten	93
Abbildung 52: Performance der NA-Netze in Verbindung mit den Subnetzen (a: MLFN, b: MLFN- schneefreie Historie, c: SOM)	94
Abbildung 53: NA-Netze - Darstellung einzelner Szenarien	95
Abbildung 54: : Performance der MLFN für die Szenarien HDNA10, HDNA20 und HDNA40 in Abhängigkeit von der Anzahl der verdeckten Neuronen, mittlerer quadratischer Fehler für normierte Daten	98
Abbildung 55: Testdaten für die HDNA-Netze: HDNA10 mit 18, HDNA20 mit 23 und HDNA40 mit 24 verdeckten Neuronen, Durchfluß und Wasserstand	99
Abbildung 56: HDNA-MLFN - Darstellung einzelner Szenarien für Durchfluß und Wasserstand	100

Abbildungen

Abbildung 57: : Performance der SOM (150 x 40) für die Szenarien HDNA10, HDNA20 und HDNA40 in Abhängigkeit vom Outputdimensionsgewicht, mittlerer quadratischer Fehler für normierte Daten..... 101

Abbildung 58: HDNA-SOM - Darstellung einzelner Szenarien für Durchfluß und Wasserstand 102

Abbildung 59: Vergleich der Performance der bestangepaßten MLFN für die Szenarien (Anzahl der verdeckten Neuronen am Säulenende), mittlerer quadratischer Fehler für normierte Daten 103

Abbildung 60: Validierung für das Hochwasserereignis im Juli 1996 104

Tabellen

Tabelle 1: Transferfunktionen.....	21
Tabelle 2: Szenarienbezeichnung	34
Tabelle 3: Räumliche und zeitliche Diskretisierung der Daten	48
Tabelle 4: Gewässerkundliche Hauptwerte nach DGJ (2003).....	55
Tabelle 5: Flächenanteile der Einzugsgebiete der Zuflußpegel und des Zwischeneinzugsgebietes entsprechend der Szenarien	58
Tabelle 6: Scheiteldurchflüsse der Hochwasserereignisse zur Kalibrierung/ Validierung des hydrodynamisch-numerischen Modells	60
Tabelle 7: Landnutzungsspezifische Rauigkeitsbeiwerte	64
Tabelle 8: Einflußspannen für die Pegel Kriebstein UP, Nossen 1 und Niederstriegis	73
Tabelle 9: Scheitelhöhe des Standardhochwassers am Pegel Nossen in Abhängigkeit von Intensität und zeitlichem Abstand eines vergangenen Niederschlagsimpulses (in mm h^{-1} und Prozent des Vergleichsscheitels ohne Niederschlagsimpuls von $0,2606 \text{ mm h}^{-1}$)	83
Tabelle 10: Niederschlagsmerkmale für das Subnetz 2	84
Tabelle 11: Merkmalsvektor für das NA-Netz.....	87
Tabelle 12: verkleinerter Merkmalsvektor für das NA-Netz (SOM).....	92
Tabelle 13: Merkmale für die HDNA-ANN, HD-Merkmale grau unterlegt.....	97

Literaturverzeichnis

- Abebe, A. J., Solomatine, D. P., Venneker, R. G. W., „Application of adaptive fuzzy rule-based models for reconstruction of missing precipitation events”, *Hydrological Sciences Journal*, vol. 45 (3), pp. 425-436, 2000
- Alvisi, S., Mascellani, G., Franchini, M., Bárdossy, A., “Water level forecasting through fuzzy logic and artificial neural network approaches”, *Hydrology and Earth System Sciences*, 10, pp. 1-17, 2006
- Anderson, J. A., „A simple neural network generating an interactive memory”, *Mathematical Biosciences*, vol. 14, pp. 197-220, 1972
- Bernd, T., Kleutges, M., Kroll, A., „Nonlinear Black Box Modelling – Fuzzy Networks versus Neural Networks”, *Neural Computing and Applications*, 8, pp. 151-162, 1999
- Chen, J., Adams, B.J., “Integration of artificial neural networks with conceptual models in rainfall-runoff modeling”, *Journal of Hydrology*, vol. 318, pp. 232-249, 2006
- Chen, J., Adams, B.J., Reply to the comments of Dr. Ashu Jain on “Integration of artificial neural networks with conceptual models in rainfall-runoff modeling”, *Journal of Hydrology*, vol. 329, pp. 277-280, 2006
- Coriso, G., Guariso, G., “Coupling Fuzzy Modelling and Neural Networks for River Flood Prediction”, *IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and Reviews*, vol. 35, no. 3, 2005
- Cullmann, J., *Online flood forecasting in fast responding catchments on the basis of a synthesis of artificial neural networks and process models*, Dissertation am Institut für Hydrologie und Meteorologie, TU Dresden, 2007
- Deutsches Gewässerkundliches Jahrbuch, Elbegebiet, Teil I: Von der Grenze zur CR bis zur Havelmündung, Landesamt für Hochwasserschutz und Wasserwirtschaft Sachsen-Anhalt, Magdeburg, 2003

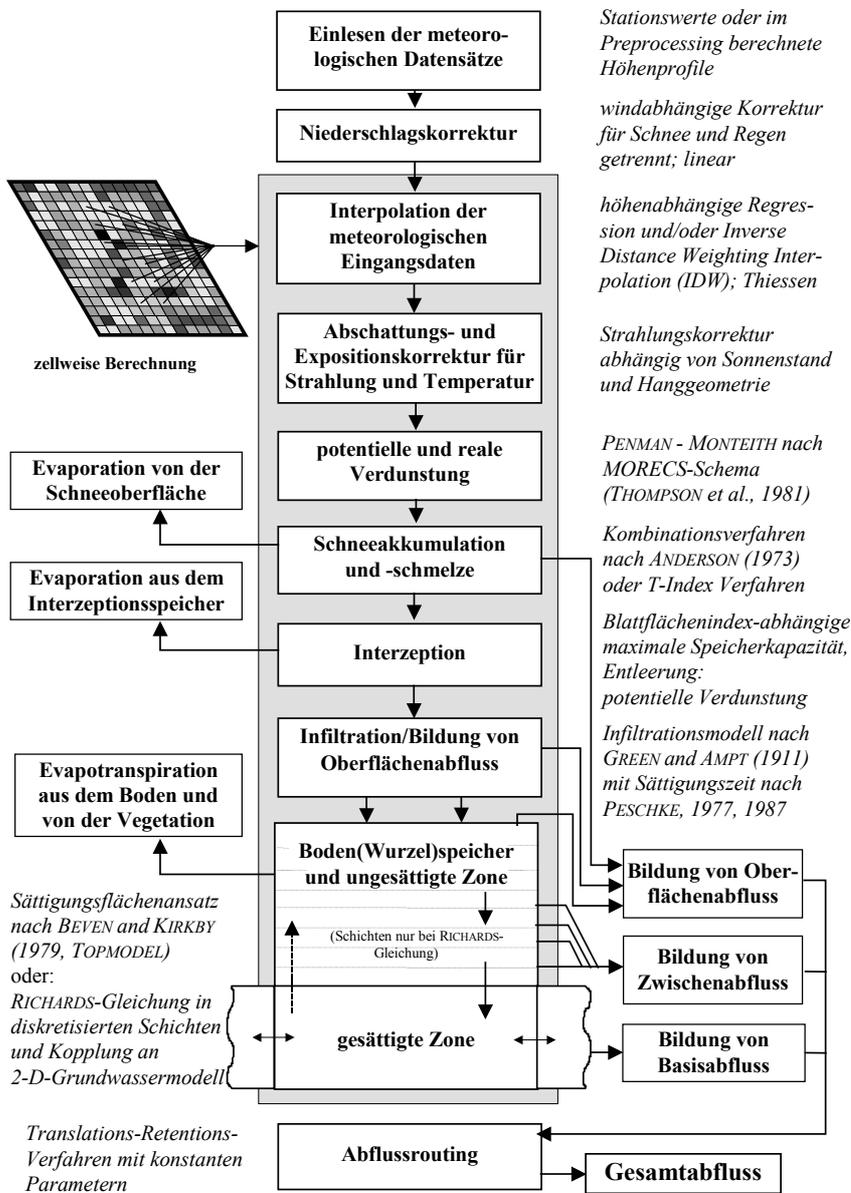
Literaturverzeichnis

- Dibike, Y. B., Solomatine, D., "River Flow Forecasting Using Artificial Neural Networks", *Phys. Chem. Earth (B)*, vol. 26, no. 1, pp. 1-7, 2001
- Dibike, Y. B., Solomatine, D., Abbott, M. B., "On the encapsulation of numerical-hydraulic models in artificial neural network", *Journal of hydraulic research*, vol. 37, no. 2, 1999
- DWD, *Starkniederschlagshöhen für die Bundesrepublik Deutschland*, Selbstverlag des Deutschen Wetterdienstes, Offenbach am Main, 1997
- Dyck, S., Peschke, G., *Grundlagen der Hydrologie.*, 2. Aufl., Verlag für Bauwesen, Berlin, 1989
- Green, W. H., Ampt, G.A., "Studies on on Soil Physics: I. The flow of air and water trough soils", *Journal of Agricultural Sciences*, 4, 1-24. 1911
- Grossberg, S., „Adaptive pattern classification and universal recording: I. Parallel development and coding of neural feature detectors", *Biological Cybernetics*, vol. 23, pp. 121-134, 1976.
- Hagan, M. T., Demuth, H. B., Beale, M., *Neural Network Design*, Boston, MA: PWS Publishing Company, 1996
- Hornik, K. M., Stinchcombe, M., White, H., "Multilayer feedforward networks are universal approximators", *Neural Networks*, vol.2, no.5, pp. 359-366, 1989
- Jain, A., Comments on "Integration of artificial neural networks with conceptual models in rainfall-runoff modeling", *Journal of Hydrology*, vol. 329, pp. 274-276, 2006
- Kohonen, T., „Correlation matrix memories", *IEEE Transaction on Computers*, vol. 21, pp. 353-359, 1972.
- Landesamt für Umwelt und Geologie des Freistaates Sachsen, *Ereignisanalyse Hochwasser August 2002 in den Osterzgebirgsflüssen*, <http://www.umwelt.sachsen.de/>, 2004
- Landesamt für Umwelt und Geologie des Freistaates Sachsen, *Konsistenzprüfung für eine Überarbeitung zur Berechnung von*

- Hochwasserscheiteldurchflüssen mit Wiederkehrintervall $HQ(T)$ für unbeobachtete Flußabschnitte*, Abschlußbericht, 2006
- Lauterbach, G., *Wasser- und Stoffhaushalt dreier Waldökosysteme des Osterzgebirges*, Dissertation, Universität Göttingen, online-Publikation, webdoc.sub.gwdg.de/diss/2001/lauterbach/lauterbach.pdf
- McCulloch, W., Pitts, W., „A logical calculus of the ideas immanent in nervous activity“, *Bulletin of Mathematical Biophysics*, vol. 5, pp.115-133, 1943.
- Minns, A.W., Hall, M.J., “Artificial neural networks as rainfall-runoff models”, *Hydrological Sciences*, vol. 41, pp. 399-417, 1996
- Pegolverzeichnis des Freistaates Sachsen, (Hydrologisches Handbuch, Teil 1), Sächsisches Landesamt für Umwelt und Geologie, 2001
- Preißler, Bollrich, *Technische Hydromechanik / 1*, Berlin, VEB Verlag für Bauwesen, 1980
- Ratke, R., Zielke, W., „Prozeßsimulation“, Skript zur Vorlesung, Institut für Strömungstechnik, Universität Hannover, 2006
- Ritter, H., Martinetz, T., Schulten, K., *Neural Computation and Self-Organizing Maps: An Introduction*, Addison Wesley, 1992
- Rosenblatt, F., „The perceptron: A probabilistic model for information storage and organization in the brain“, *Psychological Review*, vol. 65, pp. 386-408, 1958.
- Rumelhart, D. E., McClelland, J. L., eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1, Cambridge, MA: MIT Press 1986
- Schütze, N., „Neue Methoden zur Steuerung der Wassergabe mit Neuronalen Netzen in der Bewässerungswirtschaft“, *Dresdner Schriften zur Hydrologie*, Heft 3, Dresden, 2005
- Schulla, J., Jasper K., *Model description WaSiM-ETH*, Institute of Geography, ETH, Zürich, 1999

- Shresta, R. R., Theobald, S., Nestmann, F., "Simulation of flood flow in a river system using artificial neural networks", *Hydrology and Earth System Sciences*, 9 (4), pp. 313-321, 2005
- Solomatine, D. P., "Data-driven modelling: paradigm, methods, experiences", *proc. 5th International Conference on Hydroinformatics*, Cardiff, 2002
- Solomatine, D. P., Avila Torres, L. A., "Neural Network Approximation of a Hydrodynamic Model in Optimizing Reservoir Operation", *proc. International Conference on Hydroinformatics*, Zürich, 1996
- Stübner, M., Gemmar, P., „An Approach for Data Analysis and Forecasting with Neuro Fuzzy Systems – demonstrated on Flood Events at River Mosel, *Reprints: International Conference on Computational Intelligence, 5th Fuzzy Days*, Dortmund, 1997
- USACE-I, United States Army Corps of Engineers, „HEC-GeoRAS, An extension support of HEC-RAS using ArcView, users manual”, version 3.1, <http://www.hec.usace.army.mil/software/hec-ras/>, 2002
- USACE-II, United States Army Corps of Engineers, „HEC-RAS Hydraulic Reference Manual”, version 3.1, <http://www.hec.usace.army.mil/software/hec-ras/>, 2002

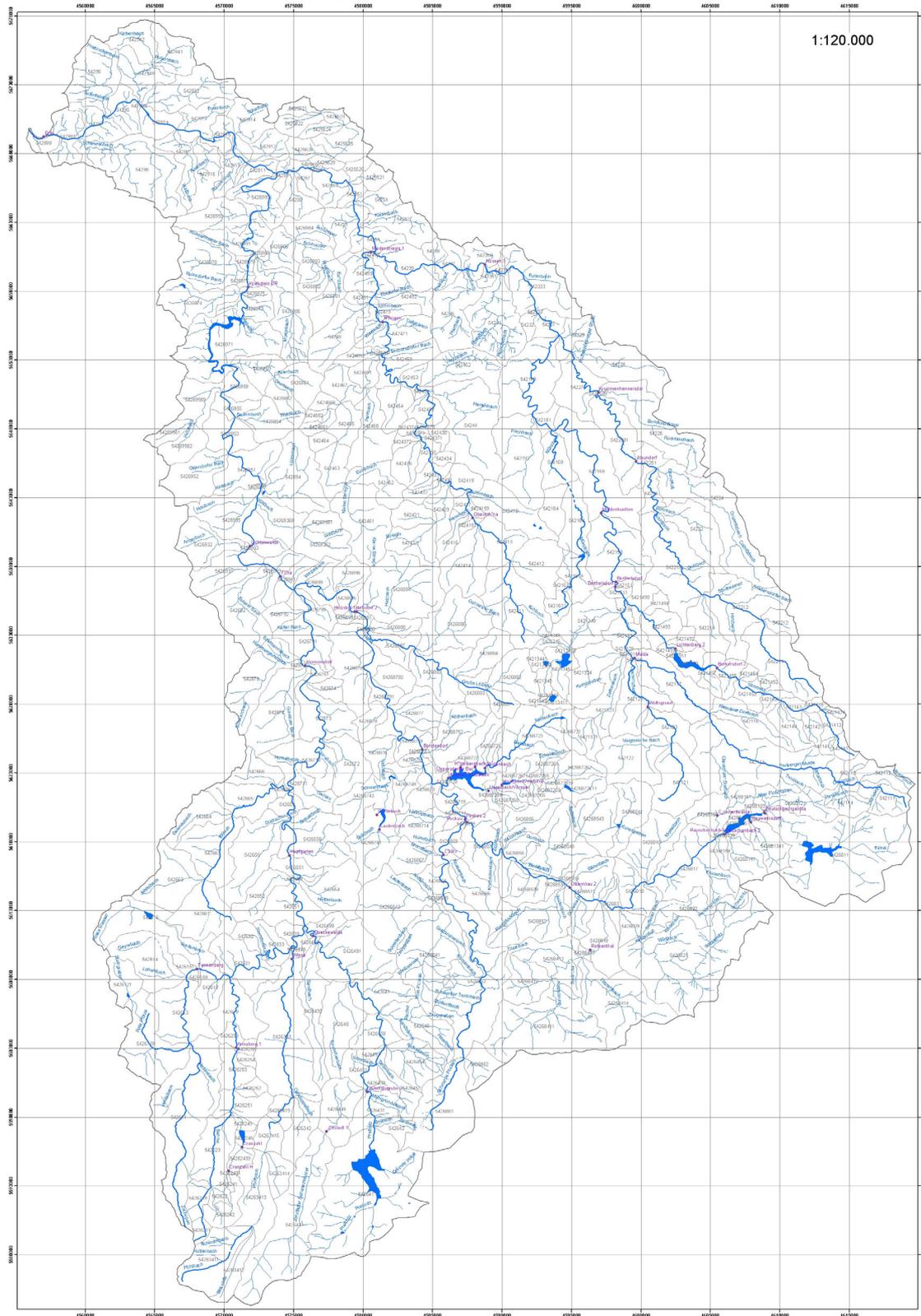
ANHANG A



Modellstruktur des WaSiM-ETH (nach Schulla und Jasper, 1999)

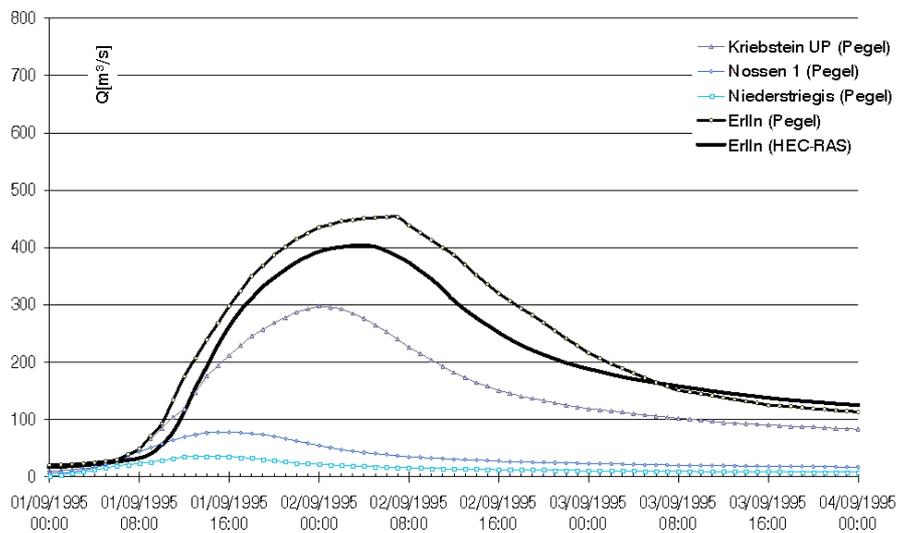
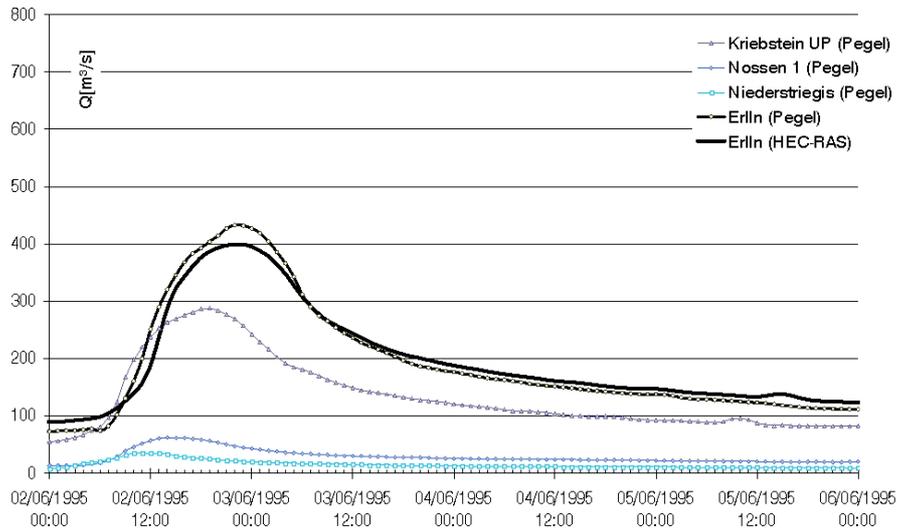
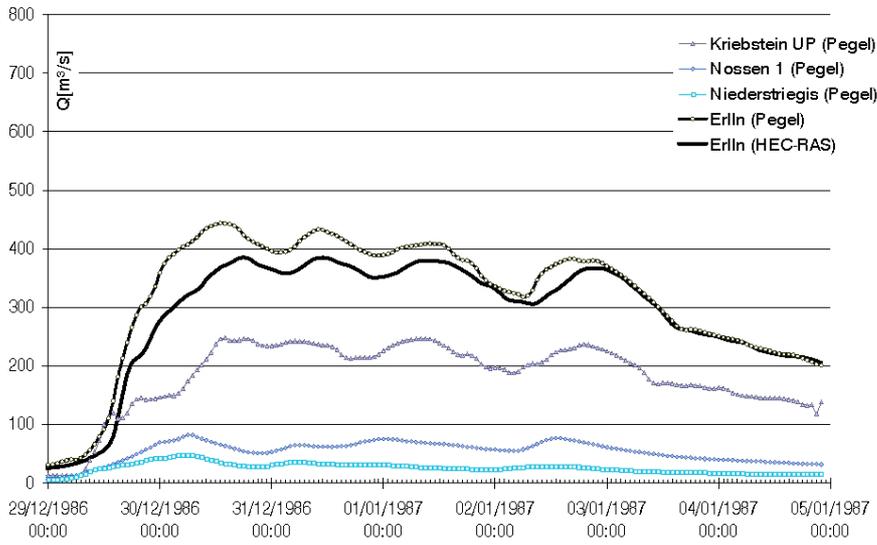
ANHANG B

EG Freiberger Mulde



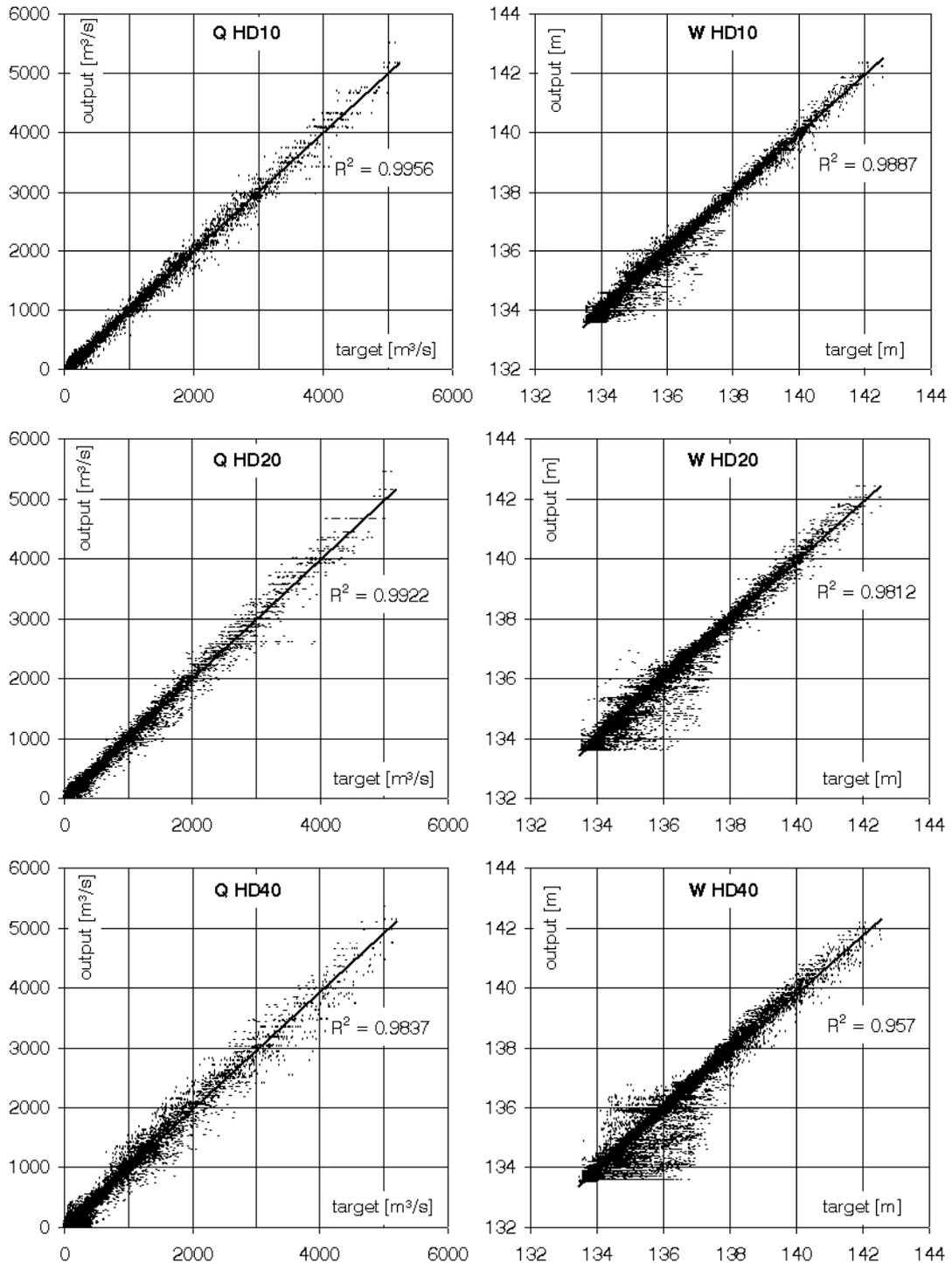
Einzugsgebiet der Freiberger Mulde

ANHANG C



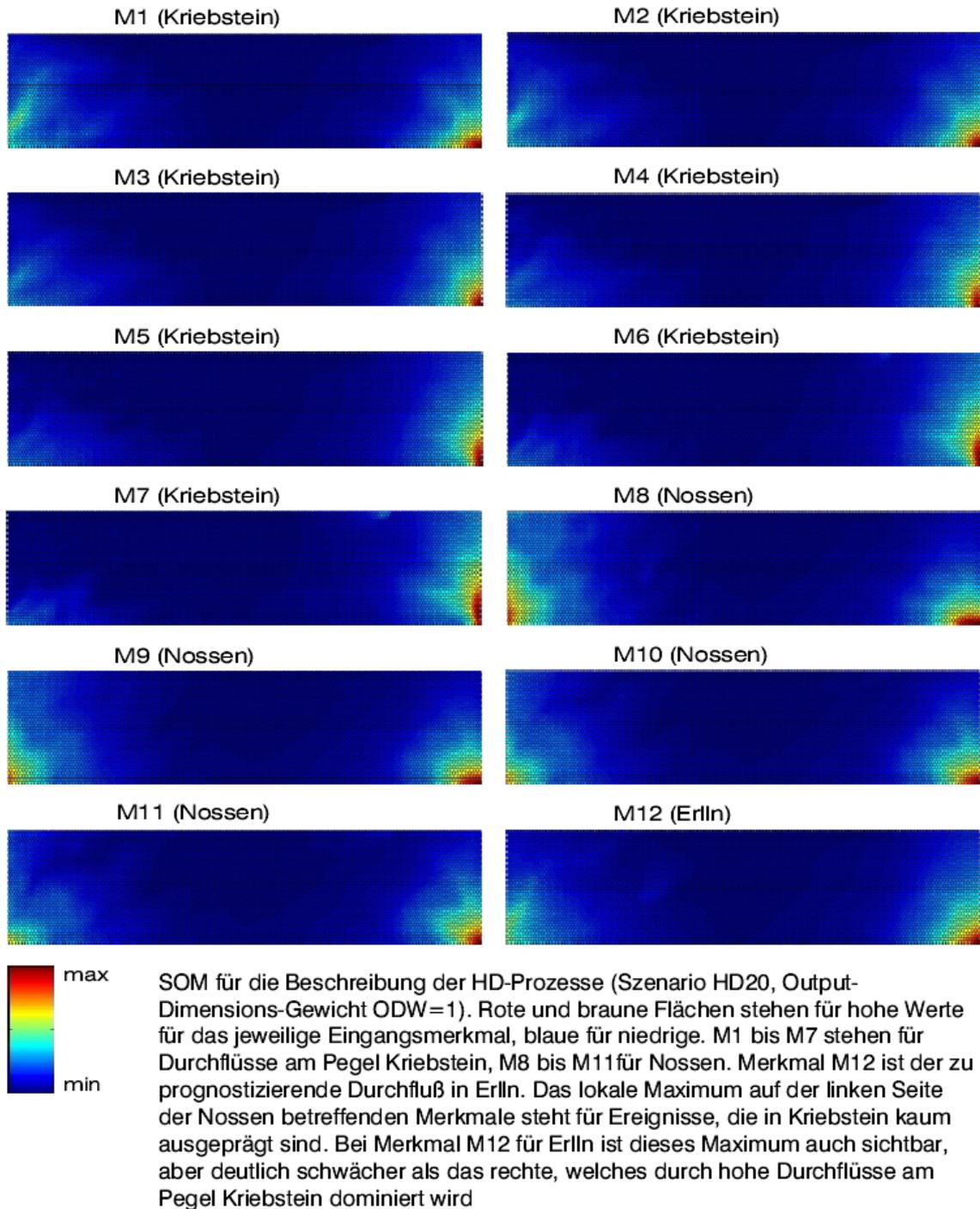
Validierung des hydrodynamischen Modells für Dezember 1986 und Juni und September 1995

ANHANG D



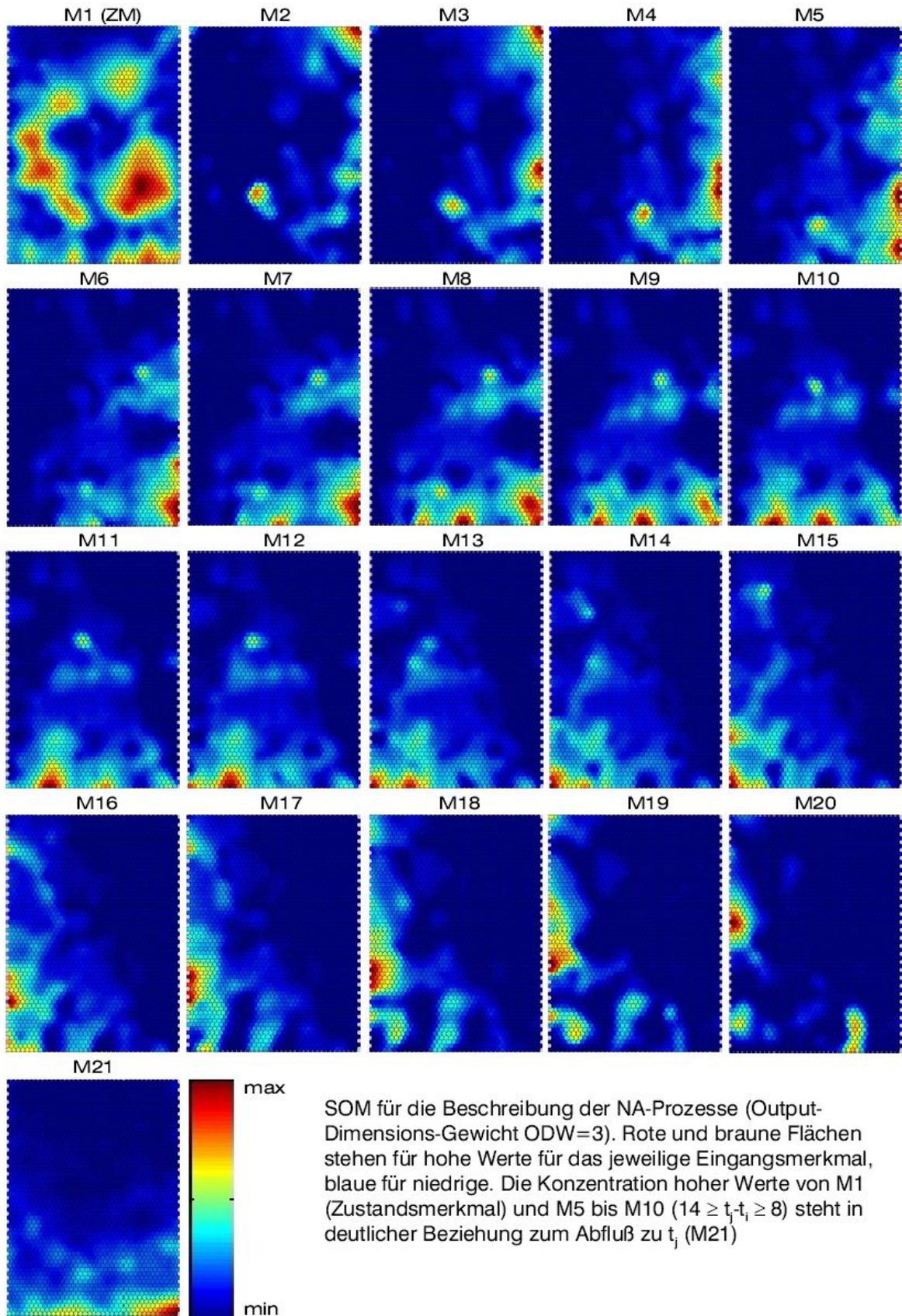
Anpassung der SOM für Durchfluß (links) und Wasserstand (rechts) für die Szenarien HD10, HD20 und HD40

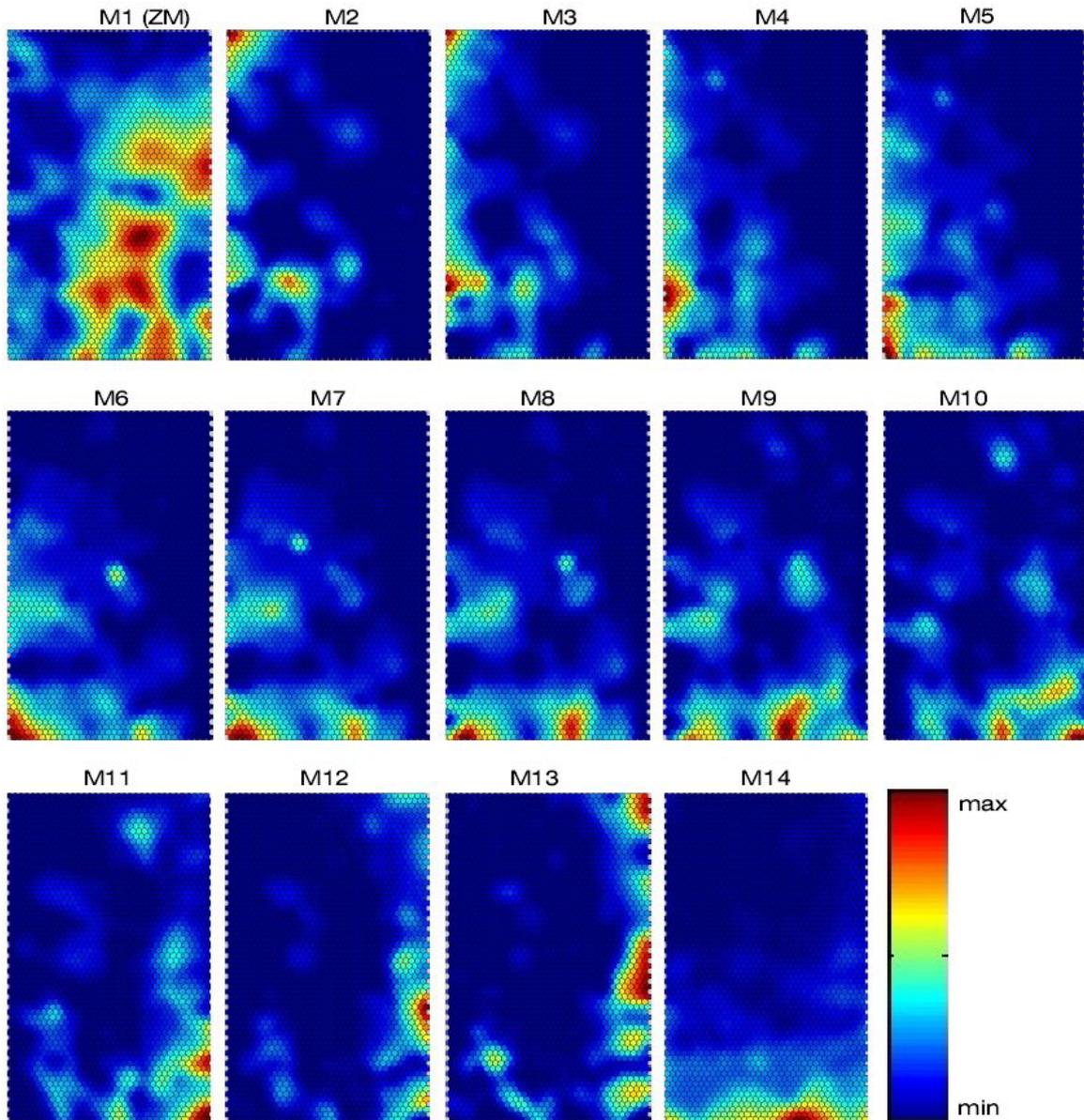
ANHANG E



Die Karten bilden die Lage des Netzes im Merkmalsraum ab, wobei jede Darstellung eine Dimension darstellt. Analog zur geodätischen Höhe auf Landkarten sind hohe Werte orange bis rot dargestellt, blaue Flächen bedeuten kleine Werte in der dargestellten Dimension. Ein Pixel steht für ein Neuron. Dessen Koordinaten ergeben sich aus den Einfärbungen in den verschiedenen Ansichten der Karte.

ANHANG F





SOM für die Beschreibung der NA-Prozesse (Output-Dimensions-Gewicht $ODW=3$) mit reduziertem Merkmalsatz. Rote und braune Flächen stehen für hohe Werte für das jeweilige Eingangsmerkmal, blaue für niedrige. Die Konzentration hoher Werte von M1 (Zustandsmerkmal) und M5 bis M10 ($16 \geq t_j - t_i \geq 8$) steht in deutlicher Beziehung zum Abfluß zu t_j (M14)