

FH-DGGVArbeitskreis

Kalibrierung und Prognosefähigkeit
von Grundwassermodellen

Leitfaden zur Kalibrierung und Prognosefähigkeit von Grundwassermodellen

**Praxisbezogener Bewertungsansatz für Hydrogeologische Modelle sowie für die
Kalibrierungsgüte und die Prognosefähigkeit von numerischen Grundwas-
serströmungs- und -stofftransportmodellen**

Datum: Oktober 2024



Inhaltsverzeichnis (Text)

Vorbemerkungen	VI
1. Zielsetzung des Leitfadens	1
2. Bewertung des Hydrogeologischen Modells (HGM)	3
2.1 Bewertungsansatz	3
2.2 Bewertung des HGM	6
2.3 Berechnungsschema	9
3. Bewertung der Kalibrierungsgüte	11
3.1 Aufgabenrelevanz und Datengrundlage	11
3.2 Maßgebliche Variablen und Volumenströme	13
3.3 Anpassung von Modellparametern	13
3.4 Überkalibrierung	13
3.5 Sensitivitätsanalyse	16
3.6 Bewertung des Erfüllungsgrades der Kalibrierungskriterien	17
3.6.1 Bewertungsansatz	17
3.6.2 Bewertung der Grundwasserbilanz	19
3.6.3 Bewertung des Erfüllungsgrades der Kalibrierungskriterien	21
3.6.3.1 Stationäre Kalibrierung anhand von GwStänden	21
3.6.3.2 Instationäre Kalibrierung anhand kurzer Zeitreihen (GwGanglinien < 1 Jahr) und/oder hydraulischer Versuche	22
3.6.3.3 Instationäre Kalibrierung anhand langer Zeitreihen (GwGanglinien > 1 Jahr) und/oder Langzeitpumpversuchen	22
3.6.3.4 Stationäre Bilanzkontrolle anhand von Abflussdaten bzw. Quellschüttungsdaten	23
3.6.3.5 Instationäre Bilanzkontrolle anhand langjähriger Quellschüttungs- und/oder Abflussganglinien bei Ansatz der zeitlichen Variabilität der GwNeubildung	23
3.6.3.6 Plausibilitätskontrolle zur Dynamik des GwUmsatzes anhand von Beschaffenheitsdaten	24
3.6.3.7 Kalibrierung anhand der GwGeschwindigkeit und -fließrichtung nach Markierungsversuchsergebnissen	24
3.6.3.8 Kalibrierung anhand der GwGeschwindigkeit und -fließrichtung anhand von Beschaffenheitsdaten	25
3.6.3.9 Kalibrierung auf Tracer-Durchbruchskurven, Konzentrationsganglinien oder Stoff-/Temperaturausbreitungsfahnen	25
3.7 Anwendungshinweise	26
3.8 Bewertung der Gesamtgüte des Grundwassermodells	28
4. Bewertung der Prognosefähigkeit des Grundwassermodells	29



Vorwort

Grundwassermodelle gehören seit Jahrzehnten zu den wichtigsten Werkzeugen der Hydrogeologie, sowohl in der Forschung als auch in der beruflichen Praxis. Deshalb gehört die Grundwassermodellierung auch zu den wichtigsten Themen der Fachsektion Hydrogeologie, was sich unter anderem in Sessions auf unseren Tagungen, Fortbildungskursen, Arbeitskreisen und Leitfäden zu diesem Thema manifestiert. Die beiden Leitfäden über „Hydrogeologische Modelle“, die 2002 und 2010 von Arbeitsgruppen der FH-DGGV veröffentlicht wurden, gehören bis heute zu den am stärksten nachgefragten und am meisten verwendeten Publikationen unserer Fachsektion.

Der Arbeitskreis „Kalibration und Prognosefähigkeit von Grundwassermodellen“ hat nun einen weiteren Leitfaden zu diesem Themenfeld erarbeitet, der einen praxistauglichen Bewertungsansatz für die Kalibrierungsgüte und Prognosefähigkeit von numerischen Modellen zur Simulation von Grundwasserströmung und Stofftransport präsentiert. Dazu gehört auch ein Excel-Tool, das für die praktische Umsetzung der vorgeschlagenen Kriterien verwendet werden kann.

Als Vorsitzender der FH-DGGV danke ich der Sprecherin, Frau Dr.-Ing. Angela Prein, und allen Mitgliedern dieses Arbeitskreises für ihre sorgfältige und fundierte Arbeit und wünsche diesem Leitfaden eine ebenso weitreichende und nachhaltige Verbreitung und Umsetzung wie seinen beiden oben genannten Vorgängern. Ein Kollege war übrigens an allen drei Leitfäden über Grundwassermodelle als Autor beteiligt: Herr Dr. Johannes Riegger, ein Ehrenmitglied unserer Fachsektion, dem ich an dieser Stelle für sein langanhaltendes Engagement danken möchte.

Prof. Dr. Nico Goldscheider
Vorsitzender der FH-DGGV
Oktober 2024



Mitglieder des Arbeitskreises

Dr. Dirk Brehm
BGU Dr. Brehm & Grünz GbR, Bielefeld

Andreas Graf
Bayerisches Landesamt für Umwelt, Hof

Dr. Bernd Hanauer
HG Büro für Hydrogeologie und Umwelt GmbH, Gießen

Dr. Bernhard Huber
HydroConsult GmbH, Augsburg

Dr.-Ing. Ulrich Lang
Ingenieurgesellschaft Prof. Kobus und Partner GmbH, Leinfelden-Echterdingen

Dr. Ulf Mohrlök
Karlsruher Institut für Technologie KIT, Karlsruhe

Dr.-Ing. Angela Prein (Sprecherin)
Hessisches Landesamt für Naturschutz, Umwelt und Geologie, Wiesbaden

Alexander Renz
Bundesgesellschaft für Endlagerung mbH, Peine

Dr. Johannes Riegger (Kapitel 2)
Universität Stuttgart, Institut für Wasser- und Umweltsystemmodellierung, Stuttgart

Dr.-Ing. Wolfgang Schäfer
Grundwassermodellierung, Wiesloch

Axel Voss
GeoFIRM Ronschke & Voss Hydrogeologie GbR, Bad Nenndorf



Anhang

Excel-Tool „Bewertung Grundwassermodell“ mit Vorlagen/ Beispielen für

- (1) die Bewertung des Hydrogeologischen Modells (HGM)
- (2) die Güte der Kalibrierung
- (3) die Bewertung der Wasserbilanz

Download über Webseite der FH-DGGV: <https://fh-dggv.de/kalpro-gw-modelle/>

Verzeichnis der verwendeten Unterlagen

- /1/ Aufbau und Anwendung numerischer Grundwassermodelle in Wassergewinnungsgebieten
Technische Regel – Arbeitsblatt, DVGW W 107 (A), DVGW, Bonn, Februar 2016
- /2/ Hydrogeologische Modelle – Ein Leitfaden mit Fallbeispielen
Schriftenreihe der Deutschen Geologischen Gesellschaft – Hydrogeologische Beiträge, Heft 24,
Hannover, 2002
- /3/ Hydrogeologische Modelle – Bedeutung des Hydrogeologischen a priori-Wissens
Schriftenreihe der Deutschen Geologischen Gesellschaft – Heft 70, Hannover, 2010
- /4/ FH-DGGV – Kursunterlagen Angewandte Grundwassermodellierung IV. Kalibration und Parameteroptimierung. Johannes Riegger



Verzeichnis verwendeter Abkürzungen

B	Themenblock
BB	Themenblock Bewertung
BG	Blockgewicht
c	Konzentration
DB	Datenbewertung
DGM	Digitales Geländemodell
DP	Datenpunkt
E	Erfüllungsgrad
g	global
GOK/POK	Geländeoberkante/Pegeloberkante
Gw...	Grundwasser...
h	Wasserstand, Piezometerhöhe
HB	HGM-Bewertung
HGM	Hydrogeologisches Modell
KP	Kriterienpunkte
KPV	Kurzpumpversuch
l	lokal
LPV	Langzeitpumpversuch
MNQ	Mittlerer Niedrigwasserabfluss
Q _{ges.}	Gesamtvolumenstrom
Q	Volumenstrom (z. B. Quellschüttung, MNQ)
REV	Repräsentatives Elementarvolumen
RGB	Relevanzgewicht Themenblock
RGT	Relevanzgewicht Thema
RP	(Aufgaben-)Relevanzpunkte
TG	Themengewicht
TP	Themenpunkte
t	Temperatur
v	GwFließgeschwindigkeit

Tabellenblätter des Excel-Tools

HGM-Bewertung

HGM-Gewichtung

Bewertung der Wasserbilanz

Zielgrößen der Kalibrierung

Güte der Kalibrierung



Vorbemerkungen

Der Leitfaden „Kalibrierung und Prognosefähigkeit von Grundwassermodellen“ mit dem zugehörigen Excel-Tool „Bewertung Grundwassermodell“ kann über die Website der FH-DGGV unter dem Arbeitskreis „KalPro GwModelle“ als PDF und Excel-Arbeitsmappe (mit Makros) heruntergeladen werden. Die Verweise auf die Tabellenblätter des Excel-Tools sind farblich hervorgehoben.

Wir möchten Modellanwender und Modellprüfer¹ gerne dazu aufrufen, die vorgelegten Bewertungsansätze des Leitfadens – insbesondere zur Kalibrierung, Gesamtgüte und Prognosefähigkeit des numerischen Grundwassermodells – anhand aktueller Projekte auf ihre Verständlichkeit, Angemessenheit und Praxistauglichkeit zu überprüfen.

Der Leitfaden-Entwurf wurde im März 2024 im Rahmen der FH-DGGV-Tagung in Aachen dem Fachpublikum vorgestellt und zusammen mit wesentlichen Rückmeldungen diskutiert. Die entsprechenden Anregungen und Diskussionsbeiträge wurden in den vorliegenden Leitfaden eingearbeitet.

Wir bedanken uns an dieser Stelle für die Unterstützung bei der Fortschreibung des vorliegenden Leitfadens und bei dem Ziel, eine transparente und aufgabenspezifisch geeignete Bewertungsgrundlage zur Güte des Hydrogeologischen Modells, des numerischen Grundwassermodells und deren Prognosefähigkeit zur einheitlichen Anwendung zu bringen.

Ein besonderer Dank der Mitglieder des Arbeitskreises „Kalibrierung und Prognosefähigkeit von Grundwassermodellen“ geht an die beiden Reviewer, die sich der Mühe unterzogen haben, alle Aspekte des Leitfadens kritisch zu prüfen und konstruktive Anmerkungen zu machen.

¹ Zur besseren Lesbarkeit wird in diesem Leitfaden das generische Maskulinum verwendet. Die in dieser Arbeit verwendeten Personenbezeichnungen beziehen sich – sofern nicht anders kenntlich gemacht – auf alle Geschlechter.

1. Zielsetzung des Leitfadens

Grundwassermodelle sind wichtige Werkzeuge zur Beantwortung vieler wasserrechtlicher Fragestellungen. Reine Strömungsmodelle unterstützen hierbei die Bemessung von Schutzgebieten, die Dimensionierung von Wasserhaltungen oder die Bestimmung der Auswirkungen von Baumaßnahmen auf das Grundwasserregime. Transportmodelle helfen bei der Berechnung der Stoff- oder Wärmeausbreitung im Untergrund, unter anderem im Zusammenhang mit Grundwasser-Schadensfällen, -Sanierungsmaßnahmen und geothermischen Nutzungen. Dabei bestehen von Seiten der Auftraggeber, Behörden und der Öffentlichkeit hohe Erwartungen und qualitative Anforderungen an die Aussagefähigkeit solcher Modelle. Dies gilt insbesondere dann, wenn eine quantitative Beschreibung des Systems durchgeführt werden muss. Diesen Erwartungen steht eine komplexe Untergrundstruktur gegenüber, für die häufig auch bei hohem Erkundungsaufwand nur eine verhältnismäßig geringe Informationsdichte zur Verfügung steht.

Die Güte und die Prognosefähigkeit eines numerischen Grundwassermodells (GwModell) werden in der Praxis oftmals hinterfragt. Grundlage für ein numerisches GwModell ist stets ein Hydrogeologisches Modell (HGM). Daher wurde der vorliegende Leitfaden mit dem Ziel erstellt, eine Orientierungshilfe für die Bewertung der Qualität des HGM, der Güte der Kalibrierung von numerischen Grundwassermodellen und deren anwendungsbezogener Prognosefähigkeit zu geben.

Bei der Kalibrierung werden die Parameter eines Modells dahingehend optimiert, dass das Modell vergangene (bekannte) Systemzustände (z. B. gemessene Grundwasserspiegel) ausreichend gut abbildet. Der Sinn der Kalibrierung ist es, die Prognosefähigkeit zu erreichen oder zu erhöhen. Dies gelingt unter der Annahme, dass mit der guten Reproduktion vergangener (bekannter) Systemzustände auch die Prognose zukünftiger (unbekannter) Systemzustände plausibel abgebildet wird.

Die Qualität und Prognosefähigkeit von Grundwassermodellen hängen entscheidend vom HGM ab. Zum Aufbau des HGM wird auf das Heft 24 „Hydrogeologische Modelle“ /2/ sowie das Heft 70 „Hydrogeologische Modelle – Bedeutung des Hydrogeologischen a priori-Wissens“ /3/ der Schriftenreihe der Deutschen Geologischen Gesellschaft verwiesen. Das HGM beschreibt das Grundwassersystem hinsichtlich Geometrie, (Hydro-)Geologie, Hydraulik und maßgeblicher Randbedingungen und bildet somit die wesentliche Grundlage für numerische Grundwassermodelle. Daher ist die Bewertung des HGM von zentraler Bedeutung für die Beurteilung der Qualität von Grundwassermodellen.

Der Arbeitskreis „Kalibrierung und Prognosefähigkeit von Grundwassermodellen“ hat praxisbezogene, quantitative Bewertungskriterien für die Güte von hydrogeologischen Modellen entwickelt, welche eine quantitative Einordnung zu den Qualitätsklassen bzw. Modellkategorien des DVGW-Arbeitsblattes W 107 „Aufbau und Anwendung numerischer Grundwassermodelle in Wassergewinnungsgebieten“ (Prinzip-, Planungs- und Bewirtschaftungsmodelle) /1/ erlauben.

Im vorliegenden Leitfaden wird ausgehend von der Bewertung des HGM eine Möglichkeit zur orientierenden Bewertung der Kalibrierungsgüte von Planungs- und Bewirtschaftungsmodellen aufgezeigt. Darauf aufbauend wird ein Ansatz zur Bewertung der Prognosefähigkeit eines numerischen Grundwassermodells dargestellt.

Bewertet werden Planungs- und Bewirtschaftungsmodelle. Ausgenommen von dieser Bewertung sind Prinzipmodelle, mit denen grundlegende hydrogeologisch-geohydraulische Zusammenhänge auf der Basis plausibler Ansätze untersucht bzw. verifiziert werden können. Eine Kalibrierung von Prinzipmodellen erfolgt üblicherweise nicht, sodass sie im strengen Sinne auch nicht prognosefähig sind.

Die Güte von numerischen Grundwassermodellen hängt wesentlich von der Güte des zugrundeliegenden HGM ab. Für die Gütebewertung des HGM als auch die darauf aufbauende Bewertung der Kalibrierungsgüte und der Prognosefähigkeit numerischer Grundwassermodelle wurde das Excel-Tool „Bewertung Grundwassermodell“ entwickelt, das in Teilen als Bewertungstool für das HGM bereits seit 2019 auf der



FH-DGGV-Website zur Verfügung steht¹. Im vorliegenden Leitfaden sind sowohl Aufbau und Wirkungsweise dieses Excel-Tools zur Gütebewertung des HGM, als auch die darauf aufbauende Bewertung der grundlegenden Wasserbilanz und der Kalibrierungsgüte numerischer Grundwassermodelle dargestellt. Das vollständige Excel-Tool ist auf der FH-DGGV-Website bereitgestellt und kann dort heruntergeladen werden.

Der vorliegende Leitfaden bzw. dessen Anwendung in Form des Excel-Tools „Bewertung Grundwassermodell“ soll sowohl dem Modellierer als auch dem Modellprüfer und/oder dem Auftraggeber Hilfestellung für die Einschätzung der Modellgüte entsprechend der jeweiligen Aufgabenstellung geben. Wesentlich ist hierbei ein systematisches, transparentes und gut dokumentiertes Vorgehen als Grundlage für die Diskussion und Bewertung von Modellergebnissen.

¹ Download über Internetseiten der FH-DGGV: 2023-03_FH-DGGV_Bewertung_Grundwassermodell.xlsm

2. Bewertung des Hydrogeologischen Modells (HGM)

2.1 Bewertungsansatz

Die Qualität des HGM hängt maßgeblich von der Datenlage und deren Relevanz für Modellaussagen ab und diese wiederum von der geplanten Modellanwendung. Zudem ist die räumliche Skala von Datenlage und Modellanwendung wichtig für die Qualität der Modellaussagen. Das hier vorgestellte Bewertungsschema berücksichtigt, dass sowohl Datenlage als auch Relevanz der jeweiligen Themen bzw. Themenblöcke von der späteren Modellanwendung und von der räumlichen Skala abhängen.

Für den praktischen Einsatz des HGM-Bewertungsschemas, etwa in Ingenieurbüros oder bei Behörden, wurde ein Vereinfachungsansatz entwickelt, der sich in dem Tabellenblatt „HGM-Bewertung“ des Excel-Tools realisieren lässt.

Folgende Vereinfachungen werden getroffen:

- 3 Anwendungsklassen in Kombination von Anwendung und Lage des Aquifers (tiefe oder flache großräumige Grundwassersysteme, kleinräumige Systeme für Grundwassersanierungen)
- 2 Vorhersagebereiche: global (über das gesamte Modellgebiet), lokal (über Teilbereiche des Modellgebiets bzw. Aussagegebiet)
- 3 Relevanzklassen: (0) irrelevant, (1) relevant, (2) hoch relevant
- 4 Datenklassen als Zusammenfassung von Datendichte und Datenrepräsentativität

Themen und Themenblöcke

Die Bewertung des HGM ist im Tabellenblatt „HGM-Bewertung“ des Excel-Tools thematisch strukturiert, indem die zu bewertenden Aspekte in einzelne Themen aufgeschlüsselt und folgenden Themenblöcken zugeordnet werden.

- Hydrogeologisches a priori-Wissen und Ergebnisberichte:
 - o Abgrenzung des Bilanz- und Modellraumes
 - o Strukturierung des Modellraumes, hydrostratigrafisch (vertikal, lateral)
 - o Grundwasserhydraulik, Randbedingungen
 - o Grundwasserbeschaffenheit
- Schichtgeometrien/-verbreitung:
 - o Schichtflächen (Verbreitung, Verläufe, Höhen, Mächtigkeiten)
 - o Verlauf, Versatz und hydraulische Wirksamkeit von Störungszonen
- Geohydraulische Parameter:
 - o Transmissivität(en), Durchlässigkeit(en), horizontal
 - o Durchlässigkeit(en), vertikal (Leakage-Faktor)
 - o Speicherkoeffizient(en)
 - o Leakage-Faktor(en) entlang von Oberflächengewässern
 - o durchflusswirksame Porosität (advektiver Transport)



- o Dispersionsparameter
- Messwerte und Versuchsdaten:
 - o Piezometerhöhen, Vorflutniveaus
 - o Konzentrationen/Temperaturen
 - o Abfluss-/Schüttungsdaten
- Hydraulische Randbedingungen:
 - o Herleitung und Lage von Randzuflüssen
 - o Herleitung und Lage von Randstromlinien
 - o Herleitung und Lage von höhenbezogenen Randbedingungen (Festpotenzial-, Leakage-, Drainagerandbedingungen)
 - o Herleitung und Lage von Zufluss- (Recharge) / Abflusszonen (Discharge) zu benachbarten Aquiferen
- Quellen und Senken:
 - o Grundwasserneubildungsrate, flächenhaft
 - o Entnahme- / Infiltrationsraten, punkt- oder linienhaft
 - o Stoffquellen / Stoffsenken

Anwendungsfälle

Für die unterschiedlichen Anwendungsfälle sind die Gewichtung und die Relevanz der Themen unterschiedlich festgelegt. Diese Anwendungsfälle sind wie folgt charakterisiert:

- Bilanzmodell für tiefe großräumige Grundwassersysteme: Großräumige Grundwassersysteme, für die nur wenige Messinformationen vorliegen, mit Strömungsverhältnissen, die sich nicht ausschließlich anhand von Piezometerhöheninformationen charakterisieren lassen. Die Einbindung von Wasserbeschaffenheits- bzw. Isotopeninformationen ist zwingend erforderlich. Typische Anwendungsbereiche und Fragestellungen sind:
 - o Austausch zwischen den Grundwasserleitern
 - o Geothermie
 - o Grundwasseralter
 - o Klimafolgenabschätzung
 - o Mineralwasser
 - o nutzbares Grundwasserdargebot
 - o Wasserbilanz
- Strömungsmodell für oberflächennahe großräumige Grundwassersysteme: Grundwassersysteme, die eine direkte Anbindung an oberirdische Vorflutsysteme haben und über die Grundwasserneubildung aus Niederschlag gespeist werden. Typische Anwendungsbereiche und Fragestellungen sind:
 - o Bemessung von Wasserschutzgebieten und Schutzzonen
 - o Grundwasserbewirtschaftung
 - o Trinkwassergewinnung

- o Klimafolgenabschätzung
- o Prognose von Grundwasserspiegeländerungen
- Transportmodell für kleinräumige Grundwassersysteme (Sanierung): Grundwassersysteme, in denen die Ausbreitung von Schadstoffen untersucht werden soll und hydraulische oder reaktive Sanierungsverfahren zur Anwendung kommen. Hier spielen insbesondere lokale Heterogenitäten der Aquiferkennwerte und ggf. auch von Reaktionsparametern eine dominante Rolle für die Aussageschärfe. Typische Anwendungsbereiche und Fragestellungen sind:
 - o Transport von Wasserinhaltsstoffen und Wärme
 - o Sanierung und Sicherung von Schadstofffahnen
 - o Identifizierung von natürlichen Abbauprozessen

Relevanz der Themen (Relevanzklasse vorgegeben)

Die Bedeutung jedes einzelnen Themas wird für die verschiedenen Modellanwendungen und die räumliche Skala (global - lokal) durch die Relevanzklassen (0, 1, 2) beschrieben. Für die Gesamtwertung des HGM werden die Themenblöcke entsprechend ihrer Bedeutung für die jeweilige Modellanwendung in ihrer Relevanz gewichtet. Die Themenblöcke umfassen jeweils verschiedene Themen mit unterschiedlicher Bedeutung für den betreffenden Themenblock und werden deshalb innerhalb des Themenblocks relativ gewichtet.

Die Gewichtung der Relevanz basiert auf der Erfahrung der Mitglieder des Arbeitskreises und ist für den Anwender nicht zugänglich, jedoch im Tabellenblatt „HGM Gewichtung“ des Excel-Tools dokumentiert.

Datenlage (Datentypgewicht vorgegeben)

Die Bewertung der Datensituation sollte die Datendichte sowie die Repräsentativität der Daten im Hinblick auf die Systemkomplexität berücksichtigen und deren potenziellen Einfluss auf die Modellgenauigkeit bewerten.

Die Informationen zu den jeweiligen Themen können verschiedenen Datentypen entstammen, die unterschiedlichen Informationsgehalt besitzen.

In der Tabelle 2-1 (exemplarisch, hier für das Thema Schichtflächen) sind zu jedem Thema die möglichen Datentypen und Informationen aufgelistet, die alternativ stehen oder sich ergänzen können. Diese sind entsprechend ihrer Bedeutung mit relativen Datentypgewichten (DTG) versehen, die sich für das Thema zu 100% aufsummieren. Das Gewicht, mit dem der entsprechende Datentyp zur Datenqualität des jeweiligen Themas beiträgt, ist vom Arbeitskreis vorgegeben und dokumentiert (Tabellenblatt „HGM Gewichtung“ des Excel-Tools).

Tabelle 2-1: Datentyp und Datentypgewicht

Thema	Datentyp	Datentypgewicht
Schichtflächen (Verbreitung, Verläufe, Höhen, Mächtigkeiten)		
	Bohrungen	50%
	Geophysik	30%
	hydro- / geologische Karten und Schnitte	20%
	Summe	100%

Datenbewertung (Datenklasse vom Anwender einzustufen)

Aufgabe des Anwenders bei der Bewertung des HGM ist es, zu jedem Thema und für jeden Datentyp eine Bewertung der Datensituation anhand der vorgegebenen Datenklassen vorzunehmen. Eine exemplarische Beschreibung für den Datentyp Bohrungen ist in der nachfolgenden Tabelle 2-2 beschrieben.

Die zugehörige Beschreibung ist für jeden Datentyp über den Button „Eingabe“ am oberen Rand des Tabellenblatts „HGM-Bewertung“ verfügbar.

Tabelle 2-2: Themenbezogene Beschreibung der Datenklasse

Datentyp	Datenklasse +2	Datenklasse +1	Datenklasse -1	Datenklasse -2
Bohrungen	umfangreich, gut verteilt, hohe Bohrqualität	umfangreich, ungleichmäßig verteilt, unterschiedliche Bohrqualität	wenig, gut verteilt, mäßige Bohrqualität	wenig, schlecht verteilt

In Ausnahmefällen können einzelne Datentypen für die gegebene Aufgabenstellung irrelevant sein, wie etwa fehlende Temperaturdaten bei einer nicht Dichte-induzierten Strömungsmodellierung. In diesem Fall kann in dem Tabellenblatt „HGM-Bewertung“ für den betreffenden Datentyp die Datenklasse 0 angesetzt werden.

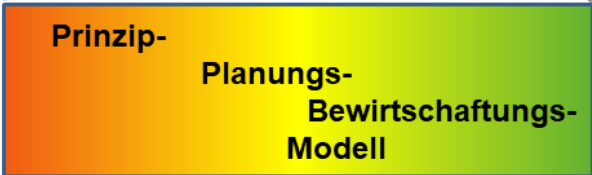
2.2 Bewertung des HGM

Aus den vorgegebenen Relevanzpunkten und -gewichten sowie aus den anwenderdefinierten Datenpunkten und den vorgegebenen Datengewichten wird im Excel-Tool für jeden Themenblock eine Punktzahl errechnet, die dem gewichteten Mittel der Datenpunkte bzw. umgerechnet einer Maßzahl zwischen 0 % und 100 % entspricht (siehe Abbildung 1).

Analog kann aus den Themen-Blockbewertungen mit den Blockgewichten eine Gesamtbewertung für das HGM mit einer Maßzahl zwischen 0 % und 100 % bestimmt werden.

Die gesamte Struktur ist im Tabellenblatt des Excel-Tools über den Button „Detailansicht“ sichtbar.

Themen (Daten)	Relevanz		Datenqualität		Bewertung
	Gewichte	Punkte (0,1,2)	Gewichte	Punkte (-2,-1,1,2)	
Hydrogeologisches a priori Wissen: ...					
...					
Schichtgeometrien / -verbreitung: ...					
Geohydraulische Parameter: ...					
Messwerte und Versuchsdaten: ...					
Hydraulische Randbedingungen: ...					
Quellen und Senken: ...					
Summe					0 ... 100 %



**Prinzip-
Planungs-
Bewirtschaftungs-
Modell**

Abbildung 1: Schematische Darstellung des Berechnungsschemas zum HGM, Struktur des Tabellenblattes „HGM-Bewertung“ des Excel-Tools mit Berechnungsschritten

Das Ergebnis der Bewertung für die einzelnen Themen und Themenblöcke ist im Excel-Tool über den Button „Bewertung“ sichtbar dargestellt. Die Grafik der Blockbewertung (Abbildung 2) gibt einen schnellen Überblick und visualisiert deutlich Stärken und Schwächen in der Datenlage der einzelnen Themenblöcke. Sie ist in die Zwischenablage kopierbar. Eine Differenzierung nach der Modellskala, global bzw. lokal, ist ebenfalls gegeben.

Mit Hilfe dieses Bewertungsschemas können auch die Auswirkungen möglicher Maßnahmen zur Verbesserung der Datenqualität hinsichtlich der Auswirkungen auf die Qualität und Prognosefähigkeit eines Grundwassermodells eingeschätzt werden.

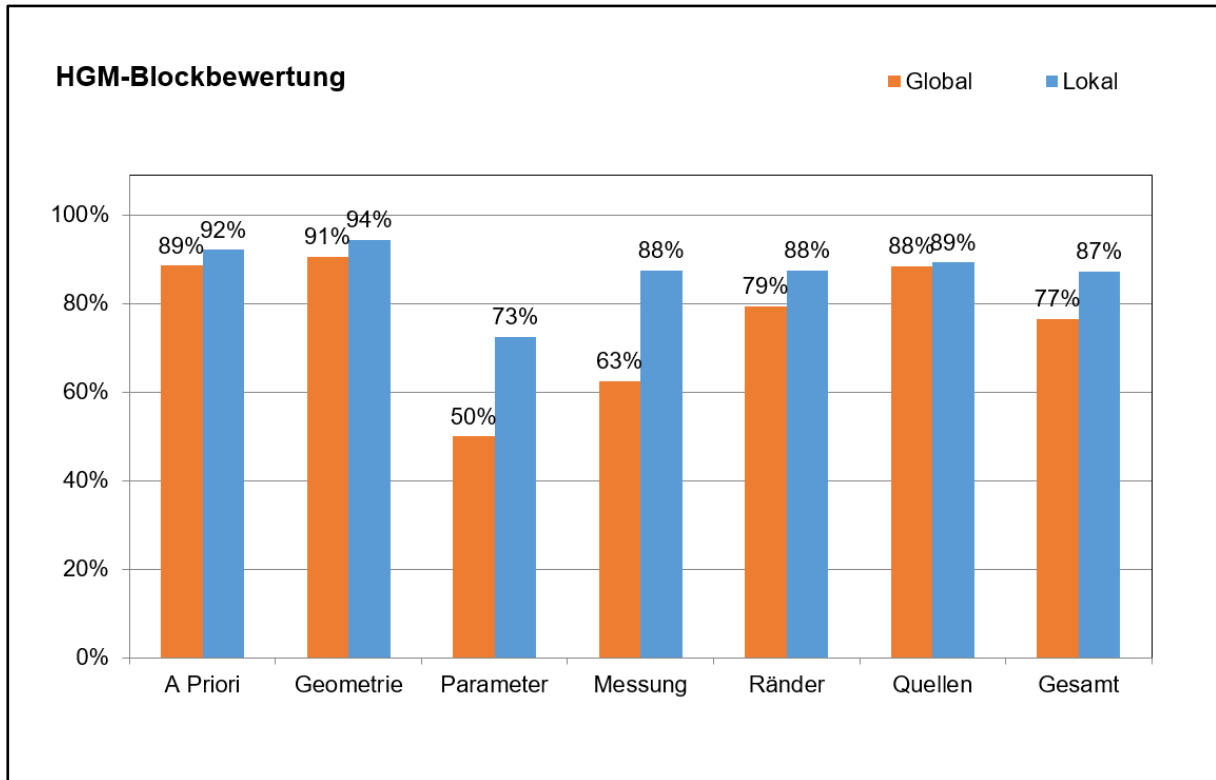


Abbildung 2: Beispielhaftes Ergebnis der Bewertungen der einzelnen Themenblöcke und des HGM insgesamt

Die Gesamtbewertung des HGM mit einer Maßzahl zwischen 0 % und 100 % erlaubt eine quantitative Einordnung des HGM in die Qualitätsklassen bzw. Modellkategorien des DVGW-Arbeitsblattes W 107. Es werden vom Arbeitskreis folgende Grenzen vorgeschlagen, wobei die Grenzen als Übergangsbereiche zu verstehen sind:

Tabelle 2-3: Einordnung der Modelle nach der Gesamtbewertung des HGM

Prinzipmodell		Planungsmodell		Bewirtschaftungsmodell	
Schlecht	→ gut	Schlecht	→ gut	Schlecht	→ gut
0 %	33 %	33 %	66 %	66 %	100 %

2.3 Berechnungsschema

Das Bewertungsschema im Excel-Tool basiert im Einzelnen auf folgenden Berechnungsschritten:

Relevanz der Themen

Zur Festlegung der Relevanz sind die Themenblöcke mittels der Blockgewichte (BG) relativ zueinander gewichtet. Die Themen innerhalb der Blöcke sind mittels Themenpunkten (TP) gewichtet. Daraus ergibt sich für die Gewichtung des Themas das Themengewicht (TG):

$$TG = BG \times \frac{TP}{\sum_{T\text{-Block}} TP}$$

In Abhängigkeit vom Anwendungsfall ist die Relevanz eines Themas separat sowohl für das Modellgebiet (global) als auch das Aussagegebiet (lokal) durch den Relevanzpunkt (RP_{g,l}) festgelegt. Damit ergibt sich für das Modellgebiet (global) als auch das Aussagegebiet (lokal) das Relevanzgewicht eines Themas:

$$RGT_{g,l} = TG \times RP_{g,l},$$

bzw. eines Themenblocks:

$$RGB_{g,l} = \sum_{T\text{-Block}} RGT_{g,l}.$$

Datenlage

Die Gewichtungen und die Relevanzpunkte sind vorgegeben und können vom Anwender nicht verändert werden. Die Relevanzpunkte sind mittels der 3 Relevanzklassen vergeben:

0 - irrelevant, 1 - relevant, 2 - hochrelevant

Die Datenqualität ist durch Datenpunkte (DP_{g,l}) definiert, die für jeden Datentyp jeweils für das Modellgebiet (global) und das Aussagegebiet (lokal) durch den Anwender vergeben werden. Die Datenpunkte werden in 4 Klassen eingeteilt und beinhalten die Dichte und Repräsentativität der Daten für den jeweiligen Datentyp (s. o. Beispiel Datentyp "Bohrungen").

Darüber hinaus werden in jedem Thema die einzelnen Datentypen entsprechend ihrer Bedeutung für das HGM mittels des Datentypgewichts (DTG) relativ zueinander gewichtet (s. o. Beispiel Thema "Schichtflächen").

Innerhalb des Themas erfolgt damit jeweils global und lokal die Datenbewertung (DB):

$$DB_{g,l} = \sum_{\text{Thema}} DTG \times DP_{g,l}$$

bzw. die Themenbewertung (TB):

$$TB_{g,l} = \sum_{T\text{-Block}} RGB_{g,l} \times DB_{g,l}.$$



Bewertung des HGM (gesamt, Themenblöcke)

Damit ist es nun möglich, eine Gesamtbewertung für das HGM (HB) in Prozent zu generieren, wieder jeweils global und lokal:

$$HB_{g,l} = \frac{1}{4} \times \frac{\sum_{HGM} TB_{g,l}}{\sum_{HGM} RGB_{g,l}} + \frac{1}{2} .$$

$$BB_{g,l} = \frac{1}{4} \times \frac{TB_{g,l}}{\sum_{HGM} RGB_{g,l}} + \frac{1}{2} .$$

Für die einzelnen Themenblöcke wird eine entsprechende Bewertung (BB) angegeben.

Anmerkung:

Die Formeln berücksichtigen die Gewichtungen der Themenblöcke ($RGB_{g,l}$) und der Themenbewertung ($TB_{g,l}$) und erzeugen so eine prozentuale Gesamtbewertung für das HGM sowohl auf globaler als auch auf lokaler Ebene.

3. Bewertung der Kalibrierungsgüte

3.1 Aufgabenrelevanz und Datengrundlage

Bei der Modellkalibrierung ist es erforderlich, alle für den jeweiligen Fall aufgabenrelevanten Aspekte einzubeziehen. Voraussetzung hierfür ist, dass eine entsprechende Datengrundlage besteht oder gegebenenfalls durch entsprechende (Erkundungs-) Maßnahmen geschaffen wurde. Andernfalls schränkt dies die Kalibrierungsgüte und somit auch grundlegend die Prognosefähigkeit des numerischen Modells möglicherweise entscheidend ein. Für die Einschätzung der Kalibrierungsgüte ist zunächst eine entsprechende – notwendigerweise subjektive – Erstbewertung dahingehend durchzuführen, ob die aufgabenrelevanten Aspekte im Rahmen der Kalibrierung hinreichend behandelt wurden bzw. behandelt werden konnten. Diese Vorprüfung kann zu folgenden Ergebnissen und Schlussfolgerungen führen:

Tabelle 3.1-4: Vorprüfung – Erstbewertung der aufgabenrelevanten Aspekte der Kalibrierung

Fallbeschreibung	Schlussfolgerung
a) Eine ausreichende Datengrundlage ist vorhanden und wurde aufgabenrelevant und vollständig bei der Modellkalibrierung berücksichtigt.	Die Modellkalibrierung ist, ggf. vorbehaltlich der Quantifizierung der Modellungenauigkeiten, hinreichend präzise und das Modell ist prognosefähig.
b) Eine ausreichende Datengrundlage ist vorhanden, wurde allerdings im Sinne der Aufgabenstellung nicht oder nicht vollständig bei der Modellkalibrierung berücksichtigt.	Die Modellkalibrierung ist im Sinne der Aufgabenstellung (noch) nicht vollständig, und das Modell ist nicht oder nur eingeschränkt prognosefähig. Es ist darzustellen, welche Daten zur Erreichung der notwendigen Kalibrierungsgüte bzw. der Prognosefähigkeit zusätzlich im Kalibrierungsprozess zu erheben und zu berücksichtigen sind. Ebenso sind die darauf basierende Kalibrierungsstrategie und die anzustrebenden Anpassungsziele darzulegen.
c) Es ist keine ausreichende Datengrundlage vorhanden.	Eine Modellkalibrierung ist im Sinne der Aufgabenstellung nicht möglich und das Modell ist nicht prognosefähig. Es ist zu prüfen, welche Daten zum Erreichen der notwendigen Kalibrierungsgüte hierzu bzw. zur Erreichung der Prognosefähigkeit erforderlich sind. Die darauf basierende Kalibrierungsstrategie und die anzustrebenden Anpassungsziele sind darzulegen.

Im **Fall b** ist denkbar, dass die umfassende Berücksichtigung der verfügbaren relevanten Daten zu einer ausreichenden Kalibrierungsgüte im Sinne der Aufgabenstellung führt. Sollte dies nicht der Fall sein, sind sehr wahrscheinlich ergänzende Untersuchungs- bzw. Erkundungsmaßnahmen zur Schaffung der erforderlichen Datengrundlagen notwendig.

Im **Fall c** ist darzustellen, welche Untersuchungs- bzw. Erkundungsmaßnahmen zur Schaffung der erforderlichen Datengrundlagen zur Erreichung einer adäquaten Kalibrierungsgüte notwendig sind.

Bei den Fällen b und c ist zu berücksichtigen, dass eine frühzeitige Modellierung dazu beitragen kann, die aufgabenspezifischen Datenlücken/Erkenntnisdefizite in einem sehr frühen Modellierungsstadium sichtbar zu machen. Das Modell stellt in diesen Fällen ein effizientes Analyseinstrument dar, auf dessen Basis bei Realisierung der notwendigen Erkundungsmaßnahmen eine Datengrundlage geschaffen werden kann, die dem Fall a entspricht. Nach Klärung entsprechender Datendefizite und ggf. Ergänzung der Datengrundlage kann eine sachgerechte Fortführung der Modellierung erfolgen.

Der folgende Ablaufplan zeigt die typische Vorgehensweise für die Kalibrierung eines numerischen Grundwassermodells.

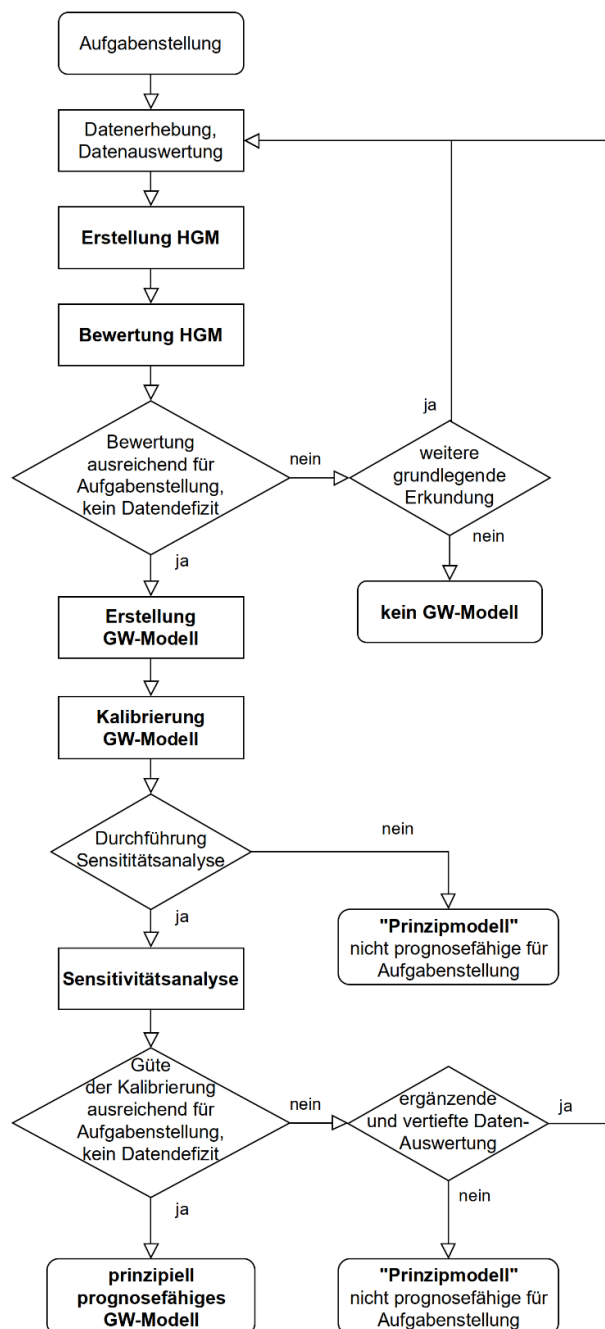


Abbildung 3: Ablaufplan – Typische Vorgehensweise bei der Modellkalibrierung

3.2 Maßgebliche Variablen und Volumenströme

Intensive Variablen sind Grundwasserstände, Temperaturen oder Konzentrationen. Sie ändern sich bei unterschiedlicher Größe des betrachteten Systems nicht, sind beobachtbar und bestimmen den Volumen- bzw. analog den Energie- und den Massenstrom.

Extensive Variablen sind Volumen, Energie und Masse. Sie ändern sich mit der Größe des betrachteten Systems und sind nicht direkt oder nur schwierig zu beobachten. Die extensiven Variablen beschreiben die Volumen- bzw. analog die Energie- und Massenstromdichte (Volumen- bzw. analog Energie- und Massenstrom pro Flächeneinheit). Diese Ströme ergeben sich aus dem Gradienten der intensiven Variablen und der räumlichen Verteilung der Parameterwerte (z. B. Durchlässigkeit und Aquifergeometrie).

Volumenströme sind alle Zu- und Abströme über die Modellränder, und zwar sowohl über die das Modell oben und unten begrenzenden Flächen als auch über die lateralen Modellränder.

Über die Kalibrierung wird das globale Verhalten des Systems bestimmt und damit der Zusammenhang zwischen extensiven und intensiven Variablen festgelegt. Dadurch werden der globale Volumenstrom bzw. analog der globale Energie- und Massenstrom quantifiziert. Die Summe der Volumenströme bzw. analog der Energie- und Massenströme (Zu- und Abflüsse) ergibt die Bilanz.

Die im HGM definierten Bilanzglieder bzw. einzelnen Volumenströme sind maßgeblich für die Kalibrierung und müssen für alle Modellränder aufgeschlüsselt werden.

Wenn die Volumenströme zweier Modelle für den gleichen Modellraum nicht in der gleichen Größenordnung liegen, ist ein weitergehender Klärungsbedarf gegeben, da sich diese Differenzen auf die Prognosegenauigkeit auswirken. Daher ist es notwendig die Bestimmungsgenauigkeit bzw. Unsicherheit der einzelnen Volumenströme anzugeben.

3.3 Anpassung von Modellparametern

Ist ein Modell in seiner globalen Wasser- bzw. Stoffbilanz hinreichend bestimmt, ist lokal eine Anpassung der relevanten Parameter vorzunehmen. Hier werden in einem Iterationsprozess gemessene intensive Variablen mit berechneten intensiven Variablen verglichen. D. h., das relative Verhältnis der Transmissivitäten oder das relative Verhältnis der Volumenströme kann unter Beachtung der Gesamtbilanz verändert werden. Der Anpassungserfolg kann über Scatterplots, Fehler und Streuung gemessen werden. Der hierbei zu verwendende Zielwert sollte an die erwartbare Abweichung zwischen Modell und Beobachtung im Zuge eines Iterationsprozesses angepasst sein, um Überkalibrierung zu vermeiden.

3.4 Überkalibrierung

In der Praxis wird die Annahme, dass mit einer Kalibrierung die Prognosefähigkeit verbessert wird, meist nicht hinterfragt. Es wird somit implizit davon ausgegangen, dass aus der Fähigkeit eines Modells bekannte (meist vergangene) Zustände zu reproduzieren, auf die Fähigkeit unbekannte (zumeist in der Zukunft liegende) Zustände vorherzusagen geschlossen werden kann. In den meisten Fällen mit sachgerechter händischer Kalibrierung ist dies auch der Fall.

Weniger bekannt ist, dass die obige Annahme nur bis zu einem bestimmten Grad der Anpassungsgüte gilt. Wird dieser Zustand bei der Kalibrierung überschritten, nimmt trotz Verbesserung der Anpassungsgüte die Prognosefähigkeit des Modells ab. Dieser (kontraintuitive) Effekt wird als Überkalibrierung (engl. *overfitting*)

bezeichnet und kann insbesondere (aber nicht nur) dann eintreten, wenn automatische Optimierer zur Kalibrierung verwendet werden.

Die Ursache für den Effekt der Überkalibrierung beruht auf der Tatsache, dass ein Modell immer inhärente Schwächen aufweist. Es ist immer eine Vereinfachung der Realität und basiert auf Daten, die mit einer Unsicherheit behaftet sind. Im Falle einer Überkalibrierung werden unzulängliche Vereinfachungen (wie z. B. Verletzung des REV¹-Konzepts, unbekannte Entnahmen, unzulängliche Anbindung von Fließgewässern usw.) und ungenaue Daten (wie z. B. GwStände, Entnahmeraten, GwNeubildungsverteilungen) durch bei der Kalibrierung optimierte Parameterwerte kompensiert. Dadurch wird das beobachtete Systemverhalten zwar besser, doch die Eigenschaften des Systems schlechter beschrieben. Diese schlechtere Beschreibung der Systemeigenschaften verringert potenziell die Prognosefähigkeit. Dies führt dann beispielsweise trotz besserer Modellanpassung bei der Kalibrierung zu einer schlechteren Modellanpassung im Rahmen der Validierung bzw. Prognose.

Auch wenn das Erkennen einer Überkalibrierung schwierig ist, folgt ihr Auftreten aber bekannten Mustern:

- Generell wird zu Beginn des Kalibrierungsvorganges die Modellabweichung aus den (noch unkalibrierten) Parameterwerten herrühren. Die Anpassung der Parameterwerte führt in dieser Phase normalerweise immer zu einer Verbesserung der Kalibrierungsgüte und der Prognosefähigkeit. Mit zunehmender Anpassung der Parameterwerte wird der Großteil der aus ihnen herrührenden Modellabweichungen beseitigt.
- Im Verlauf der Kalibrierung kann ein Punkt erreicht werden, ab dem vor allem Modellschwächen auskalibriert werden. Parameterwerte in Modell und in Realität entfernen sich wieder voneinander (das Modell wird „unphysikalischer“). Dieser Punkt wird umso früher erreicht, je mehr (unbekannte) Modellschwächen vorhanden sind.
- Eine Verbesserung der Prognosefähigkeit mit zunehmend besserer Kalibrierung ist ab diesem Punkt an folgende Bedingung geknüpft: Sind sich Prognose und Kalibrierung ähnlich (wird z. B. die Art der Aussagegrößen beibehalten und die Stresslevel werden nicht wesentlich geändert), wirkt die Auskalibrierung von Modellschwächen noch eine Zeitlang positiv auf die Prognosefähigkeit.
- Ist das Optimum der Prognosefähigkeit erreicht, führt eine weitere Kalibrierung (Reduktion des RMS) zu einer abnehmenden Prognosefähigkeit. Das Modell ist überkalibriert. Das Erreichen dieses Optimums ist nicht deterministisch bestimmbar, da hierzu die Fehler des Modells bekannt sein müssten (dann würde man das Modell aber entsprechend korrigieren können).

Die folgende Abbildung stellt den Prozess grafisch dar. Grün dargestellt ist die Abweichung von Modell und Messung, welche mit zunehmendem Fortschritt der Kalibrierung (entlang der Rechtsachse) immer weiter sinkt. Die Prognosefähigkeit ist als Abweichung eines Validierungsdatensatzes illustriert, welche an einem bestimmten Punkt optimal (minimale Abweichung) wird. Ab diesem Punkt sinkt die Prognosefähigkeit wieder, da die aus Modellschwächen resultierenden Modellfehler (gepunktet) nun stärker werden, bis sie ab einem gewissen Punkt dominieren.

¹ Zum REV-Konzept siehe G. Matthes (Hrsg.) (1993), Lehrbuch der Hydrogeologie. Band 3 Geohydraulik. 3. Auflage, Berlin

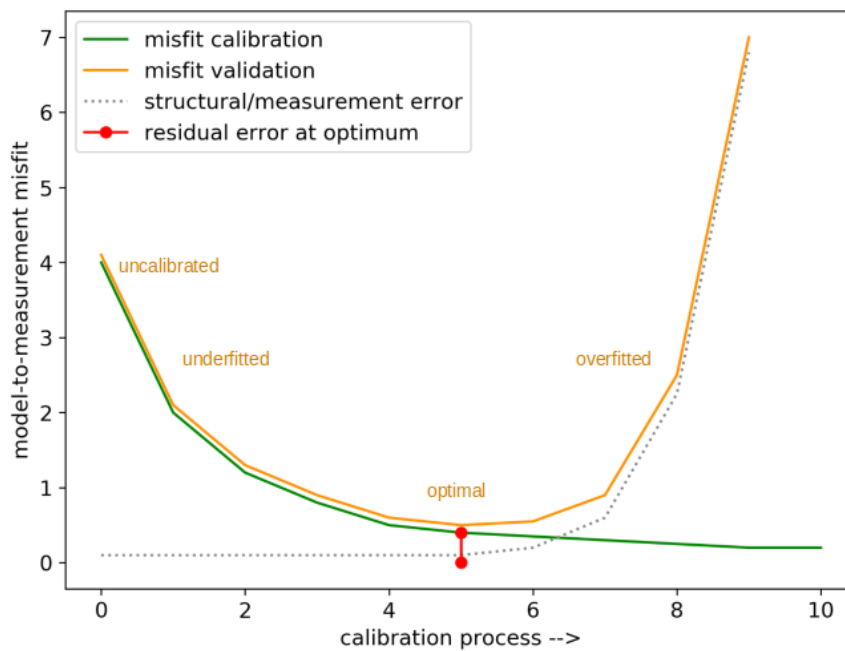


Abbildung 4: Typisches Verhältnis von Kalibriergüte (grün) und Prognose/Validierung (orange)

Das Erkennen von Überkalibrierung setzt Erfahrung voraus, insbesondere ein realistisches Verständnis, bis zu welcher Abweichung ein Modell die Messwerte abbilden kann (inkl. möglicher Messfehler). Diese kumuliert sich aus den Unsicherheiten, mit denen alle Modellschritte behaftet sind, z. B. Stratigrafie und Schichtgrenzen im HGM, Diskretisierung und unveränderbare Parameter (z. B. DGM, Randbedingungen, zeitliche Auflösung von Messreihen) im numerischen Modell. Eine beginnende Überkalibrierung kann auch dadurch angezeigt werden, dass die über die Kalibrierung ermittelten Parameterwerte unplausibel werden und nicht mehr den im HGM definierten Bandbreiten entsprechen. Dies kann zusammen mit einem guten Systemverständnis dazu genutzt werden, um Modellschwächen zu identifizieren und zu korrigieren.

Automatische Optimierer können die Plausibilität der Parameterwerte durch die Verwendung von Regularisierungsmethoden berücksichtigen und in die Optimierung mit einbeziehen. Bei der Regularisierung wird dem Optimierer eine zweite Zielfunktion vorgegeben, welche die Plausibilität der Parameterwerte beschreibt.

Die Auswirkungen einer Überkalibrierung werden an folgendem Beispiel erläutert:

Der Einfluss der Überkalibrierung soll an einem Modell illustriert werden, in welchem eine Grundwassermessstelle im Einflussbereich eines Vorfluters kalibriert werden soll. Für dieses Beispiel einer Modellschwäche soll hier angenommen werden, dass der Wasserspiegel des Vorfluters aufgrund von Interpolationen örtlich nur mit einer Genauigkeit von 5 cm bekannt sei und die Messung nur halbjährlich erfolgte. Weiterhin wurde aufgrund des niedrigen Flurabstandes vereinfachend eine Vollsättigung des Grundwasserleiters angenommen, wodurch es zu einer leichten Überschätzung der Transmissivität des Grundwasserleiters im Modell kommt. Der einzige Kalibrierungsparameter sei die hydraulische Durchlässigkeit zwischen Vorfluter und Messstelle.

Ein optimales Kalibrierergebnis kann generell auf folgende Weise abgeschätzt werden:

Allen Modellschritten liegen Unsicherheiten zu Grunde, die sich zu einer Abweichung von gemessenen und simulierten Daten kumulieren. Im HGM sind z. B. Stratigrafie und Schichtgrenzen mit Unsicherheiten behaftet, die im numerischen Modell zusätzlich meist vereinfacht abgebildet werden. Die Diskretisierung im numerischen Modell begrenzt die Auflösung der räumlichen und zeitlichen Beschreibung des Systemverhaltens. Weiterhin werden im numerischen Modell unveränderbare Parameter vorgegeben, wie z. B. DGM,

Randbedingungen, nicht ausreichend zeitlich aufgelöste Messwertaufnahmen. Sollten diese z. B. bei einer Vorhersage des Grundwasserstandes einen RMS-Wert von 20 cm (an modellierten Messstellen) erwarten lassen, sollte das RMS-Ziel der Kalibrierung nicht geringer angesetzt werden, um eine optimale Prognosefähigkeit erreichen zu können.

Für das Beispiel bedeutet dies somit Folgendes:

Im unkalibrierten Zustand würde die Modellabweichung bei ca. 50 cm liegen, da sowohl die hydraulische Durchlässigkeit im Modell als auch die Modellschwächen wie das Potenzial des Vorfluters zu einer Abweichung vom realen Verhalten führen.

Die hydraulische Durchlässigkeit würde nun so eingestellt, dass die Modellabweichung sinkt. Dabei wird sich der Parameterwert im Modell zunächst dem realen Wert annähern. Würde nun der exakte Wert genau „getroffen“ (in diesem Fall wäre die Prognosefähigkeit des Modells optimal), würde allerdings die Modellabweichung aufgrund der Modellschwächen immer noch größer Null sein; z. B. würde allein das fehlerbehaftete Vorfluterpotenzial bereits einen Beitrag von 5 cm liefern.

Tatsächlich dürfte es nun möglich sein, den Modellparameter weiter so anzupassen, sodass die Modellabweichung unter 5 cm fällt und das Modell noch besser kalibriert erscheint. Spätestens ab diesem Punkt würde aber der Modellparameter beginnen, den (nicht kalibrierbaren) Modellfehler zu kompensieren und sich damit wieder von seinem optimalen (dem realen Wert entsprechenden) Parameterwert entfernen.

Das Auskalibrieren des fehlerhafteten Vorfluterpotenzials könnte so zu einer Verschlechterung der Prognosefähigkeit führen. Zwar würde im Messpunkt der Wasserspiegel im Vorfluter besser dargestellt. Da mit der falschen hydraulischen Durchlässigkeit aber auch das Grundwassergefälle falsch berechnet würde, würde sich die Prognose für andere Bereiche im Modell verschlechtern.

Die vernachlässigte ungesättigte Zone wirkt sich in dieser Hinsicht möglicherweise vorteilhaft aus. Die optimale hydraulische Durchlässigkeit wird dadurch zwar von der Realität abweichen. Dies wird aber dazu führen, dass die Gesamttransmissivität des Aquifers (gesättigte und ungesättigte Zone) besser angenähert wird und das System evtl. sogar besser beschrieben wird. Bei Prognosen, für welche eine korrekte Transmissivität wichtiger ist als die hydraulische Durchlässigkeit, wäre dies von Vorteil.

3.5 Sensitivitätsanalyse

Im Anschluss an die Modellkalibrierung ist eine Sensitivitätsanalyse durchzuführen. Sie dient der Überprüfung, ob und wie sich durch Variation der kalibrierten Parameter und/oder der Randbedingungen die Anpassungsgüte des Modells verändert. Im Falle, dass sich die Anpassungsgüte des Modells hierbei verschlechtert, liegt mit hoher Wahrscheinlichkeit eine Eindeutigkeit der gefundenen Parameterverteilung und/oder ein hinreichend präziser Ansatz der Randbedingungen im kalibrierten Modell vor.

Je höher die Kalibrierungsgüte nach Tabelle 3.6-5 bewertet wird, umso geringer kann der Aufwand für eine Sensitivitätsanalyse gehalten werden. Ist die Bandbreite der bei der Sensitivitätsanalyse einsetzbaren Parameterwerte groß, schmälert dies die Prognosefähigkeit des kalibrierten Modells.

Es sind in der Sensitivitätsanalyse die Parameter und/oder Randbedingungen zu berücksichtigen, die in der Kalibrierung variiert wurden. Unter Berücksichtigung der Aufgabenstellung ist der Schwerpunkt auf jene Parameter bzw. Modellvariablen zu legen, deren mögliche Schwankungsbreite für das Prognoseergebnis im Aussagegebiet relevant sind.

Es kann sich anbieten, bereits im Prozess der Kalibrierung eine Dokumentation der Parametervariationen durchzuführen und damit die Anforderungen an eine Sensitivitätsanalyse zu erfüllen.

Eine Sensitivitätsanalyse sollte bei Planungs- und Bewirtschaftungsmodellen generell durchgeführt werden. Bei Prinzipmodellen kann sie zum besseren Systemverständnis beitragen.

3.6 Bewertung des Erfüllungsgrades der Kalibrierungskriterien

3.6.1 Bewertungsansatz

Die erzielbare Güte der Kalibrierung hängt maßgeblich von der Güte des HGM und der Nutzung der verfügbaren Daten im Kalibrierungsprozess sowie bei der Parameteranpassung ab.

Wesentlich für jede Modellkalibrierung und letztendlich entscheidend für die Prognosefähigkeit ist die Güte der grundlegenden GwBilanz. Diese GwBilanz (Modellbilanz) führt alle wesentlichen Bilanzglieder des Zu- und Abstroms eines Modells auf. Die Qualität der GwBilanzierung fließt mit einer hohen Wichtung, entsprechend dem Erfüllungsgrad E (siehe unten), über die Bewertungstabelle „Güte der Kalibrierung“ in das Excel-Tool „Bewertung Grundwassermodell“ ein.

Vor der Gütebewertung der Kalibrierung ist zunächst die einzelfallbezogene Aufgabenrelevanz zu klären; hierfür gelten folgende Relevanzpunkte (RP):

- Relevanz der GwBilanz ist in jedem Fall gegeben: RP = 2 (hoch relevant)
- Aufgabenrelevanz gegeben: RP = 1 (relevant)
- Aufgabenrelevanz nicht gegeben: RP = 0 (irrelevant)

Die entsprechenden Relevanzpunkte (RP) sind in der folgenden Tabelle 3.6-5 im Einklang mit den Bewertungen des HGM vorzugeben. Diese Tabelle ist unter dem Tabellenblatt „Güte der Kalibrierung“ des Excel-Tools um die Spalte „Bemerkung“ ergänzt, um eine Erläuterung zur Wahl eines Kriteriums zu ermöglichen.

Anschließend ist der Erfüllungsgrad (E) des betreffenden, aufgabenrelevanten Kriteriums im Einzelfall zu bewerten (s. u.). Der Erfüllungsgrad kann einen Wert von 0 bis 1 annehmen. Damit soll einer angestrebten Eindeutigkeit praktikabel Rechnung getragen werden.

Als Produkt aus dem Betrag für die Aufgabenrelevanz (Relevanzpunkte [RP] 0 oder 1) und dem Erfüllungsgrad (E) (≥ 0 bis ≤ 1) resultieren Kriterienpunkte, die in Verbindung mit der HGM-Bewertung in der Summe die Kalibrierungsgüte wiedergeben.

Tabelle 3.6-5: Bewertungstabelle Güte der Kalibrierung

Zeile	Kriterium	Parameter	Relevanzpunkte (RP)	Erfüllungsgrad (E)	Kriterienpunkte (KP)
				Faktor 0 bis 1	KP = $R_p \cdot E$
1	Plausibler Ansatz der GwBilanz (alle wesentlichen Zu- und Abflüsse / Volumenströme) für das Modellgebiet; Zustrom (GwNeubildung, lateraler Zustrom etc.) Abstrom (Entnahmen etc.)	$Q_{ges.}$	2		
2	Stationäre Kalibrierung anhand von GwStandsmessungen (Scatterplot, Fehler, Streuung, räumliche Verteilung der Abweichungen)	h	0 oder 1		
3	Instationäre Kalibrierung anhand von kurzen Zeitreihen (GwGanglinien < 1 Jahr) und/oder hydraulischen Versuchen	h	0 oder 1		
4	Instationäre Kalibrierung anhand von langen Zeitreihen (GwGanglinien > 1 Jahr) und/oder Langzeitpumpversuchen	h	0 oder 1		
5	Stationäre Bilanzkontrolle anhand von Abflussdaten (MNQ) bzw. Quellschüttungsdaten (Mittel- oder Einzelwerte)	Q	0 oder 1		
6	Instationäre Bilanzkontrolle anhand von langjährigen Quellschüttungsganglinien bei Ansatz der zeitlichen Variabilität der GwNeubildung	Q	0 oder 1		
7	Plausibilitätskontrolle zur Dynamik des GwUmsatzes anhand von hydrochemischen Daten und/oder Daten zur GwAltersstruktur	Q	0 oder 1		
8	Kalibrierung der GwFließgeschwindigkeit und GwFließrichtung anhand von Markierungsversuchsergebnissen	v	0 oder 1		
9	Kalibrierung der GwFließgeschwindigkeit und GwFließrichtung anhand von Beschaffenheitsdaten	v	0 oder 1		
10	Kalibrierung auf Tracer-Durchbruchskurven, Konzentrationsganglinien oder Stoff-/Temperaturausbreitungsfahnen (Stoff- bzw. Wärmetransportmodellierung)	v, c, t	0 oder 1		
11	Summe (Zeilen 1 bis 10)		(Σ RP)		(Σ KP)

Die Kriterien der Tabelle 3.6-5 (Zeilen 1 bis 10) werden im Folgenden erläutert.

Eine schlüssige und konsistente GwBilanz gemäß Kriterium 1 ist für jedes Modell essenziell. Ein plausibler Ansatz der maßgeblichen GwBilanzgrößen für das Modellgebiet ist daher in jedem Fall relevant. Aufgrund der großen Bedeutung der GwBilanz erhält dieses Kriterium generell zwei Relevanzpunkte (RP = 2) und wird also als einziges, hoch relevantes Kriterium doppelt gewichtet.

Die Relevanz der Kriterien 2 bis 10 (siehe Kap. 3.6.3) hängt dagegen von der jeweiligen Aufgabenstellung (Aufgabenrelevanz) ab. Hierbei gilt, dass neben der GwBilanz/Gesamtvolumenstrom (Zeile 1) bei jedem numerischen GwModell mindestens ein weiteres, die Parameteranpassung beschreibendes Kriterium, aufgabenrelevant sein muss (siehe Beispiel in Tabelle 3.6-6). Des Weiteren darf die Tabelle 3.6-5 nicht im Widerspruch zur Bewertung des HGM stehen bzw. muss alle verfügbaren Daten des HGM berücksichtigen.

Tabelle 3.6-6: Beispielhafte Zuordnung von zu bewertenden Kriterien aus Tabelle 3.5-6 zur betreffenden Aufgabenstellung

Exemplarische Aufgabenstellung	Kriterien aus Tabelle 3.5-6 („/“ = und/oder)				
	Q _{ges.}	h	Q	v	v, c, t
Wasserrecht	1	2 / 3 / 4	5 / 6 / 7	-	-
Großräumige GwBilanzen	1	2 / 3 / 4	5 / 6 / 7	-	-
Bauwasserhaltung	1	2 / 3 / 4	-	-	-
WSG	1	2 / 3 / 4	5 / 7	8 / 9	10
Sanierungen	1	2 / 3 / 4	-	8 / 9	10
Großanlagen Geothermie	1	3 / 4	-	-	10

3.6.2 Bewertung der Grundwasserbilanz

Die Erarbeitung einer schlüssigen GwBilanz (Zeile 1 in Tabelle 3.6-5) bzw. die Durchführung und Bewertung einer GwBilanzrechnung für das Modellgebiet ist von grundlegender Bedeutung. Deren Güte bzw. Präzision geht als genereller Wichtungsfaktor entsprechend dem Erfüllungsgrad (E) in die Bewertung der Kalibrierungsgüte ein.

Der Erfüllungsgrad E für die GwBilanz wird in Tabelle 3.6-5 bewertet. Als Orientierung gelten folgende Wertebereiche, wobei die Konsistenz zum HGM zu wahren ist:

- E = 0 bis 0,3: keine schlüssige (0) bis grob abgeschätzte GwBilanz (< 0,3)
- E = 0,3 bis 0,6: schlüssige GwBilanz auf der Basis einer begrenzten (≥ 0,3) bis befriedigenden (< 0,6) Datenlage und -auswertung
- E = 0,6 bis 1: präzise GwBilanz auf der Basis einer guten (≥ 0,6) bis optimalen (1) Datenlage und -auswertung

Die GwBilanz und die darauf basierende Herleitung des maßgeblichen Bilanzfaktors sind im Einzelfall darzulegen und zu dokumentieren, inkl. einer Bewertung/Einschätzung der daraus resultierenden Konsequenzen für die Modellgüte und die Prognosefähigkeit/-genauigkeit.

Besonders hohe Bedeutung hat die GwBilanz für Neubildungsgetriebene Modelle.

Bei anreicherungsgetriebenen Modellen ist es meist schwierig, eine abgesicherte GwBilanz zu erreichen. Hier sind noch weitere Aspekte wie hydrochemische Daten, Pumpversuchsdaten oder Temperatur einzu-beziehen.

Für eine nachvollziehbare Einstufung der Bilanzgüte anhand des maßgeblichen Bilanzfaktors wird empfohlen, die einzelnen zu- und abstromigen Bilanzkomponenten quantitativ inkl. Bestimmungsgenauigkeit bzw. Unsicherheit zu dokumentieren und hinsichtlich der relativen Streubreite der Zustrom- und Abstrombilanz zu bewerten. Die nachfolgende Tabelle 3.6-7 inkl. Erläuterung zeigt ein Beispiel anhand einer fiktiven stationären Modellwasserbilanz auf. Diese Tabelle ist unter dem Tabellenblatt „Bewertung der Wasserbilanz“ des Excel-Tools „Bewertung Grundwassermodell“ enthalten.

Tabelle 3.6-7: Beispielhafte Bewertung einer fiktiven stationären Modellwasserbilanz

A	B	C	D	E	F	G	H	I	J	K	L
Bilanzgrößen (Volumenströme)	Randbedingung, Quellen/Senken	Vorgabe aus HGM	Beispiel für Schwankung	Kommentar des Modellierers/Prüfers zur Schwankungsbreite	Vorgabe aus HGM	Vorgabe aus HGM	Modellbilanz	Streubreite	relativer Beitrag	relative Streubreite	relative Streubreite
		Mittelwert		Erläuterung	Schwankung +/-	Schwankung +/-	berechnet im Modell	95% Konf.intervall 2 x Spalte F	Anteil an Zustrom/Abstrom Bilanz	Wurzel (Summe Spalte G²)	
		L/s	%		%	L/s	L/s	%		L/s	
Neubildung	q	1.490	> 15		20	298,1	1.490	40	62,1%	324,5	13,5%
Rand-Zustrom	q,h	600	> 15		20	120,0	450	40	18,8%		
Gew. Infiltration Rand	leak	300	> 30		15	45,0	459	30	19,1%		
vertikaler Zustrom	leak,q	0	> 30		15	0,0	0	30	0,0%		
Gew. Infiltration Modellgebiet	leak	0	> 10		50	0,0	0	100	0,0%		
Punkt-Quellen	q	0	> 10		15	0,0	0	30	0,0%		
zus. Komponente 1	?					0,0		0	0,0%		
zus. Komponente 2	?					0,0		0	0,0%		
Zustrom-Bilanz		2.390		Hinweis: erwartete Streuung in Spalte L kann kleiner sein als die Streuung eines Einzelwertes		463	2.399		100%		
Rand-Abstrom	q,h	0	> 25		15	0,0	0	30	0,0%		
Gew. Exfiltration Rand	leak	930	> 30		50	465,0	939	100	39,1%		
vertikaler Abstrom	leak,q	0	> 30		15	0,0	0	30	0,0%		
Punkt- Senken	q	950	> 10		5	47,5	939	10	39,1%		
Gew. Exfiltration Modellgebiet	leak	400	> 30		25	100,0	387	50	16,1%		
Drainagen/ Quellen	leak	110	> 10		50	55,0	134	100	5,6%		
Grundwasserzehrung durch Verdunstung	q	0	> 15		15	0,0	0	30	0,0%		
zus. Komponente 3	?					0,0	0	0	0,0%		
zus. Komponente 4	?					0,0	0	0	0,0%		
Abstrom-Bilanz		2.390				668	2.399		100%		

kleinere relative Streubreite aus Zustrom- oder Abstrombilanz: 13,5%

Erfüllungsgrad (0 - 1) -> 0,9

Erfüllungsgrad (Bewertungstext) präzise bis optimal

Kalibrierziel im Beispiel

Erläuterung der Tabelleninhalte Tabelle 3.6-7:

Die Zuordnung der Bilanzgrößen (Volumenströme) hinsichtlich Modellansatz, Modellergebnis und Kalibrierungsziel sind modellspezifisch. Sie sind durch den Modellierer aufgabenorientiert vorzugeben.

Die Bilanzgrößen (Mittelwerte und Sigma-Werte) sind aus der vorläufigen GwBilanz im HGM abzuleiten und wie folgt zu übernehmen (Vorgaben für die numerische Modellierung):

- Kalibrierziel: bekannter Wert, der mit dem Modell nachvollzogen werden muss

- Sigma: Schätzfehler/Abweichung aufgrund von Bestimmungsgenauigkeiten bzw. Unsicherheiten
- 2*Sigma: Streubreite der Unsicherheiten

Die relative Streubreite der Zustrom- bzw. der Abstrombilanz ist die Wurzel aus der Summe der Quadrate der Schätzfehler/Abweichungen.

Der Erfüllungsgrad (E) resultiert aus der Bilanzseite mit der kleineren relativen Streubreite (Zustrom in Beispiel der Tabelle 3.6-7). Er wird entsprechend der Tabelle 3.6-8 eingestuft. Für das Beispiel in Tabelle 3.6-7 ist der Zustrom mit relativer Streubreite von 16,6 % maßgeblich, was nach der Umrechnungstabelle 3.6-8 einen Erfüllungsgrad von 0,8 ergibt. Ein Erfüllungsgrad = 1 bedeutet die Abwesenheit von Unsicherheiten, ist aber unrealistisch.

Tabelle 3.6-8: Umrechnungstabelle für den Erfüllungsgrad aus der Bilanzseite mit der kleineren relativen Streubreite

Summe der relativen Streubreite aus Tabelle 3.6-7	Erfüllungsgrad	Beschreibung
<10	1	ideal
10 - 40	0,9 - 0,6	präzise bis optimal
41 - 70	0,5 - 0,3	schlüssig bis befriedigend
>70	0,1	nicht schlüssig bis grob abgeschätzt

Die Bilanzseite mit der größeren relativen Streubreite (Abstrom aus Beispiel in Tabelle 3.6-7) ist durch die Massenerhaltung bestimmt und daher nicht maßgeblich.

3.6.3 Bewertung des Erfüllungsgrades der Kalibrierungskriterien

3.6.3.1 Stationäre Kalibrierung anhand von GwStänden

Der Erfüllungsgrad (E) in Zeile 2 der Tabelle 3.6-5 ist optimal mit 1 zu bewerten, wenn – ausgehend von der Bewertung der Datenlage zu den GwMessstellen und GwStandsmessungen im HGM hinsichtlich Messstellendichte/-verteilung und eindeutig definiertem hydrologischen Zustand – folgende Kriterien vollständig erfüllt sind:

- Gutes Fehler-Histogramm (ideal: Gauß-Verteilung, keine Schiefe, Mittelwert ~0).
- Mittlerer relativer Fehler ist klein gegenüber der natürlichen, bekannten oder gemessenen GwSpiegelschwankungsbreite im Aussagegebiet (Abweichung < 20 %) (lässt sich ggfs. für regionale Modelle nicht einhalten)
- Mittlerer relativer Fehler ist klein gegenüber der maximalen GwStandsdifferenz im Aussagegebiet (Abweichung < 5 %). Der Scatterplot zeigt keine Auffälligkeit hinsichtlich der Position und Lage der Abweichungen.
- numerischer Fehler in der Modellbilanz < 1 %.
- Plausible flächenhafte Verteilung der Abweichungen im Modellgebiet.

Ist eines der Kriterien a) bis e) nicht erfüllt, mindert dies den Erfüllungsgrad jeweils um den Betrag 0,2.

3.6.3.2 Instationäre Kalibrierung anhand kurzer Zeitreihen (GwGanglinien < 1 Jahr) und/oder hydraulischer Versuche

Der Erfüllungsgrad (E) in Zeile 3 der Tabelle 3.6-5 ist optimal mit 1 zu bewerten, wenn folgende Kriterien vollständig erfüllt sind:

- a) Hydrogeologisch eindeutige Anfangsbedingungen sind gegeben und/oder werden durch eine stationäre Kalibrierung mit einem Erfüllungsgrad von $E \geq 0,8$ bestätigt.
- b) Saisonale Variationen der GwNeubildung und evtl. weiterer systemtreibender Randbedingungen (z. B. infiltrierende Gewässer), Quellen (Infiltrationen) und Senken (GwEntnahmen) werden quantitativ umgesetzt.
- c) Die Modellergebnisse repräsentieren die gemessene Dynamik des GwSystems adäquat (visueller Vergleich von gemessenen und gerechneten Ganglinien und Verlaufsdiagrammen von hydraulischen Versuchen, z. B. von mittlerem GwStandsniveau, Amplitude und synchrones Abbild der GwStandsdynamik (geringe Phasenverschiebung).
- d) Numerischer Fehler in der Modellbilanz $< 1 \%$.

Ist eines der Kriterien von a) bis d) nicht oder nur eingeschränkt erfüllt, mindert dies den Erfüllungsgrad jeweils um den Betrag 0,2. Ist keines dieser Kriterien erfüllt, führt dies trotzdem zu $E = 0,20$, wenn die Anfangsbedingungen plausibel geschätzt wurden. Ist diese Schätzung jedoch unplausibel, führt dies zu $E = 0$.

3.6.3.3 Instationäre Kalibrierung anhand langer Zeitreihen (GwGanglinien > 1 Jahr) und/oder Langzeitpumpversuche

Der Erfüllungsgrad (E) in Zeile 4 der Tabelle 3.6-5 ist optimal bzw. mit 1 zu bewerten, wenn folgende Kriterien erfüllt sind:

- a) Hydrogeologisch eindeutige Anfangsbedingungen sind gegeben und/oder werden durch eine stationäre Kalibrierung mit einem Erfüllungsgrad von $E \geq 0,8$ bestätigt.
- b) Langjährige oder saisonale Variationen der GwNeubildung und evtl. weiterer systemtreibender Randbedingungen (z. B. infiltrierende Gewässer), Quellen und Senken werden quantitativ umgesetzt.
- c) Relative Potenzialänderungen durch hydraulische Versuche im Aussagegebiet werden treffend wiedergegeben, auch deren räumliche Verteilung.
- d) Die Modellergebnisse repräsentieren die gemessene Dynamik des GwSystems adäquat (visueller Vergleich von Ganglinien und Verlaufsdiagramme von hydraulischen Langzeitversuchen; Anpassungskriterien: mittleres GwStandsniveau, Amplitude und zeitlicher Verlauf der GwStandsdynamik, Leerlaufverhalten, Trends, räumliche Verteilung der Abweichungen für geeignete Stichtage (z. B. für NW- und MW-Verhältnisse und quasistationäre Absenkungszustände).
- e) Numerischer Fehler in der Modellbilanz $< 1 \%$.

Ist eines der Kriterien von a) bis e) nicht oder nur eingeschränkt erfüllt, mindert sich der Erfüllungsgrad jeweils um den Betrag 0,2; ist keines dieser Kriterien erfüllt, führt dies zu $E = 0$. Anfangsbedingungen können in Einzelfällen nur geschätzt werden. Ist diese Schätzung jedoch unplausibel, führt dies ebenfalls zu $E = 0$.

3.6.3.4 Stationäre Bilanzkontrolle anhand von Abflussdaten bzw. Quellschüttungsdaten

Der Erfüllungsgrad (E) in Zeile 5 der Tabelle 3.6-5 ist optimal mit 1 zu bewerten, wenn folgende Kriterien vollständig erfüllt sind:

- a) Die zugrundeliegenden Werte repräsentieren mittlere Abflussbedingungen (\leq MNQ) oder einen hydrologisch eindeutig definierten Zustand.
- b) Die (Quell-) Abflussdaten sind eindeutig abgrenzbaren Teileinzugsgebieten zuzuordnen, die vollständig innerhalb des Modellgebietes liegen oder mindestens eindeutig bilanzierbare Zuflusskomponenten enthalten.
- c) Die hydrostratigrafische Zuordnung der Abflussdaten/Quellschüttungen bzw. eine eindeutige Definition des entwässerten GwStockwerks ist gegeben.

Ist eines der Kriterien von a) bis c) nicht erfüllt, mindert dies den Erfüllungsgrad jeweils um den Betrag 0,25. Ist keines der Kriterien erfüllt, sind gegebenenfalls begründete Schätzwerte zugrunde zu legen, und es gilt dann $E = 0,25$. Ist diese Schätzung jedoch unplausibel, führt dies zu $E = 0$.

3.6.3.5 Instationäre Bilanzkontrolle anhand langjähriger Quellschüttungs- und/oder Abflussganglinien bei Ansatz der zeitlichen Variabilität der GwNeubildung

Der Erfüllungsgrad (E) in Zeile 6 der Tabelle 3.6-5 ist optimal mit 1 zu bewerten, wenn folgende Kriterien vollständig erfüllt sind:

- a) Hydrogeologisch eindeutige Anfangsbedingungen sind gegeben und/oder werden durch eine instationäre Kalibrierung anhand langjähriger Betriebs- und Messdaten von Brunnenfeldern und/oder Langzeitpumpversuchen oder anhand von GwGanglinien über einen langjährigen Zeitraum gemäß Tab. 3.6-5, Zeile 4. mit einem Erfüllungsgrad von $E \geq 0,8$ bestätigt.
- b) Die Dynamik der Quellschüttungsganglinien und/oder Abflussganglinien (für Zeiträume mit vorwiegend grundwasserbürtigem Basisabfluss) wird adäquat abgebildet (visueller Vergleich von Ganglinien; Anpassungskriterien: Mittelwert und dynamisches Verhalten, Auslaufverhalten, Trends).
- c) Die hydrostratigrafische Zuordnung der Abflussdaten/Quellschüttungen ist gegeben.

Ist eines der Kriterien von a) bis c) nicht erfüllt, mindert dies den Erfüllungsgrad jeweils um den Betrag 0,25. Ist keines der Kriterien erfüllt, sind gegebenenfalls begründete Schätzwerte zugrunde zu legen, und es gilt dann $E = 0,25$. Ist diese Schätzung jedoch unplausibel, führt dies zu $E = 0$.

3.6.3.6 Plausibilitätskontrolle zur Dynamik des GwUmsatzes anhand von Beschaffenheitsdaten

Der Erfüllungsgrad (E) in Zeile 7 der Tabelle 3.6-5 ist optimal mit 1 zu bewerten, wenn folgendes Kriterium vollständig erfüllt ist:

- a) Hydrochemische Befunde und/oder beobachtete GwAltersstruktur stehen nicht im Widerspruch zur berechneten Fließrichtung und der berechneten Abstandsgeschwindigkeit (siehe nachfolgende Beispiele).

Zum Kriterium a): Das Auftreten anthropogener Stoffe in GwLeitern und/oder Daten zur GwAltersstruktur liefern Informationen zur Dynamik des GwUmsatzes, die auch plausibel mit den Ergebnissen eines kalibrierten GwModells übereinstimmen müssen. Hier ist die Aufenthaltszeit in der ungesättigten Zone mit den dort unter Umständen stattfindenden Prozessen zu beachten. Hierzu einige Beispiele:

- Ist Grundwasser in einem tieferen GwStockwerk tritiumfrei, so muss die modelltheoretische Fließzeit (unter Berücksichtigung der Passage der ungesättigten Bodenzone) > ca. 70 Jahre betragen; bei Vorliegen von FCKW- oder SF₆- Daten sind Fließzeiten im Grundwasser < 50 Jahre beurteilbar.
- Treten abwasserbürtige Stoffe in einem Brunnen mit anteiliger Uferfiltrat-Förderung auf, muss im GwModell ein Fließweg vom entsprechend markierten infiltrierenden Gewässer zum Brunnen resultieren.
- In tiefen Systemen mit sehr alten Grundwässern (hohe ¹⁴C-Modellalter) ist die erdgeschichtliche Entwicklung des GwStrömungsfeldes zu berücksichtigen. Beispielsweise erfolgte im Pleistozän über mehrere tausend Jahre in Periglazialgebieten (nahezu) keine GwNeubildung bzw. Zusicke- rung aufgrund des Permafrosts. Evtl. räumliche Differenzierungen bei hohen ¹⁴C-Modellaltern und daraus resultierende Überlegungen zur GwDynamik können daher höchstens eingeschränkt ein Bewertungskriterium für modellrechnerische Fließgeschwindigkeiten darstellen, die für das aktuelle regionale GwStrömungsfeld gelten.

Bei Erfüllung des Kriteriums a) ist das kalibrierte Modell plausibel (E = 1), wenn dies die entsprechenden Messergebnisse bestätigen oder die Rechenergebnisse zumindest diesen nicht widersprechen.

Sind die Modellergebnisse teilweise erfüllt, so ist der Erfüllungsgrad im jeweiligen Anwendungsfall hydro- geologisch-geohydraulisch zu bewerten (E = 0,5). Werden die Messergebnisse nicht ansatzweise vom kalibrierten Modell wiedergegeben, gilt E = 0.

3.6.3.7 Kalibrierung anhand der GwGeschwindigkeit und -fließrichtung nach Markierungsversuchsergebnissen

Der Erfüllungsgrad (E) in Zeile 8 der Tabelle 3.6-5 ist optimal mit 1 zu bewerten, wenn folgende Kriterien vollständig erfüllt sind:

- a) Der durch den Markierungsversuch belegte GwFließweg wird durch das kalibrierte Modell schlüs- sig nachvollzogen.
- b) Die durch den Markierungsversuch belegte GwFließgeschwindigkeit (Abstandsgeschwindigkeit, definiert durch 50 % Tracerdurchgang [Median]) wird durch das kalibrierte Modell nachvollzogen.

Ist eines der Kriterien von a) bis b) nicht erfüllt, mindert sich der Erfüllungsgrad jeweils um den Betrag 0,5. Ist keines dieser Kriterien erfüllt, gilt E = 0, da das Markierungsversuchsergebnis in keiner Weise mit dem Modell nachvollzogen werden kann.

3.6.3.8 Kalibrierung anhand der GwGeschwindigkeit und -fließrichtung anhand von Beschaffenheitsdaten

Der Erfüllungsgrad (E) in Zeile 9 der Tabelle 3.6-5 ist optimal mit 1 zu bewerten, wenn folgende Kriterien vollständig erfüllt sind:

- a) Der durch den bzw. die bewertungstauglichen Stoff(e) belegte GwFließweg wird durch das kalibrierte Modell schlüssig nachvollzogen.
- b) Sollte die Datenaufzeichnung auch Hinweise/Belege, z. B. aus Schadensfällen, für die GwFließgeschwindigkeit (Abstandsgeschwindigkeit) liefern, so wäre dies analog zu dem oben beschriebenen Umgang mit Markierungsversuchsergebnissen zu sehen und durch das kalibrierte Modell adäquat nachzuvollziehen. Hinweis: die Datendichte ist in solchen Fällen i. d. R. nicht mit planmäßig durchgeführten Markierungsversuchen vergleichbar.

Ist eines der Kriterien von a) bis b) nicht erfüllt, mindert dies den Erfüllungsgrad jeweils um den Betrag 0,5. Ist keines der Kriterien erfüllt, gilt $E = 0$, da der gemessene Fließweg des betreffenden Stoffs in keiner Weise mit dem Modell nachvollzogen werden kann.

3.6.3.9 Kalibrierung auf Tracer-Durchbruchskurven, Konzentrationsganglinien oder Stoff- /Temperaturausbreitungsfahnen

Der Erfüllungsgrad (E) in Zeile 10 der Tabelle 3.6-5 ist optimal mit 1 zu bewerten, wenn folgende Kriterien vollständig erfüllt sind:

- a) Der durch die Tracer-Durchbruchskurven, Konzentrationsganglinien oder Stoff-/ Temperaturausbreitungsfahnen belegte GwFließweg wird durch das kalibrierte Modell schlüssig nachvollzogen. Dispersion, Retardation und Stoffabbau sind adäquat berücksichtigt.
- b) Die durch die Tracer-Durchbruchskurven, Konzentrationsganglinien oder Stoff-/ Temperaturausbreitungsfahnen belegte Stoff-/ Wärmetransportgeschwindigkeit wird durch das kalibrierte Modell ausreichend nachvollzogen.

Ist eines der Kriterien von a) bis b) nicht erfüllt, mindert dies den Erfüllungsgrad jeweils um den Betrag 0,5. Ist keines dieser Kriterien erfüllt, gilt $E = 0$, da das gemessene Ausbreitungsverhalten des betreffenden Stoffs in keiner Weise mit dem Modell nachvollzogen werden kann.

3.7 Anwendungshinweise

Zur Erreichung von Aussagen zu den Zeilen 7 – 9 in Tabelle 3.6-5 kann mit der Berechnung von (rein advektiven) Bahnlinien, Stromlinien und/oder Fließzeiten eine ausreichende Bewertungsgrundlage geschaffen werden. Zur Erreichung von Aussagen zu Zeile 10 ist zwingend eine Transportmodellierung erforderlich.

Je nach Anwendungsfall kann die Aufgabenrelevanz der einzelnen Kriterien unterschiedlich sein. Dies ist im Einzelfall festzulegen, entsprechend mit 0 oder 1 zu bewerten und die Summe dieser aufgabenspezifischen Relevanzpunkte (Summe RP, Zeile 11 in Tabelle 3.6-5) zu bilden.

Anschließend ist der Erfüllungsgrad (E) nach den oben beschriebenen Einzelbewertungen zu ermitteln. Es resultieren Beträge für E zwischen 0 und 1. Aus dem Produkt der **Relevanzpunkte (RP)** und des jeweiligen **Erfüllungsgrades (E)** ergeben sich **Kriterienpunkte (KP)** für das betreffende aufgabenrelevante Kriterium, die ebenfalls aufzusummieren sind (**Summe KP**, Zeile 11 in Tabelle 3.6-5).

Unabhängig vom Anwendungsfall muss neben dem zwingend notwendigen 1. Kriterium (GwBilanz; Zeile 1) für ein kalibriertes Modell mindestens ein weiteres aufgabenrelevantes Kriterium aus Zeile 2 bis 10 in Tabelle 3.6-5 erfüllt sein. Ist das 1. Kriterium (GwBilanz) nicht erfüllt, ist ein unzureichend kalibriertes Modell oder bestenfalls ein Prinzipmodell festzustellen.

Je nach Anwendungsfall kann es erforderlich sein, dass mehrere aufgabenrelevante Kriterien erfüllt sein müssen, um eine akzeptable Kalibrierung zu erreichen. Dies ist im Einzelfall zu bewerten und begründet darzulegen. In jedem Fall resultieren nach Tabelle 3.6-5 einzelfallspezifische Summen der Relevanzpunkte (Summe RP) und der Kriterienpunkte (Summe KP).

Für den jeweiligen Anwendungsfall ist die Summe der Kriterienpunkte (Summe KP) mit der Summe der Relevanzpunkte (Summe RP) ins Verhältnis zu setzen, wobei der **Quotient Summe (Σ) KP / Summe (Σ) RP** maximal den Betrag 1 annehmen kann. Dieser Quotient beschreibt die **Kalibrierungsgüte** in Anlehnung an das DVGW-Arbeitsblatt W 107 wie folgt:

- < 0,33 Prinzipmodell
- 0,33 bis 0,66 Planungsmodell
- > 0,66 Bewirtschaftungsmodell

Diese Grenzen sind als orientierend und nicht als absolut zu verstehen.

Durch Anwendung des Tabellenblatts „Zielgrößen der Kalibrierung“ des Excel-Tools können die Kalibrierungsziele transparent dargestellt werden. Dies zeigt die folgende Tabelle beispielhaft für ein regionales Grundwasserströmungsmodell.

Tabelle 3.7-1: Beispielhafte Anwendung des Tabellenblatts „Zielgrößen der Kalibrierung“ des Excel-Tools – Darstellung der Kalibrierungsziele für ein regionales GwStrömungsmodell

Zeile	Kriterien für die Modellkalibrierung	RP	Ansätze, Kalibrierungsziele
1	Plausibler Ansatz der GwBilanz (alle wesentlichen Zu- und Abflüsse / Volumenströme) für das Modellgebiet; Zustrom (GwNeubildung, lateraler Zustrom etc.); Abstrom (Entnahmen etc.)	2	Umsetzung der GwNeubildung aus Niederschlag (stationär und instationär) auf Basis eines entsprechenden Datensatzes der Fachbehörden
2	Stationäre Kalibrierung anhand von Standardrohrspiegelhöhen/Piezometerhöhen (Scatterplot, Fehler, Streuung, räumliche Verteilung der Abweichungen)	1	Stationäre Kalibrierung anhand der GwStandsmessungen bei mittleren Bedingungen und Einbeziehung aller GWM im Modellgebiet
3	Instationäre Kalibrierung anhand von kurzen Zeitreihen (GwGanglinien < 1 Jahr) und/oder hydraulischen Versuchen	1	Reproduktion des 2-monatigen Pumpversuchs an den relevanten Brunnen
4	Instationäre Kalibrierung anhand von langen Zeitreihen (GwGanglinien > 1 Jahr) und/oder Langzeitpumpversuchen	1	Reproduktion der langjährigen GwGanglinien aus dem Gewinnungsgebiet und von amtlichen Messstellen im Modellgebiet
5	Stationäre Bilanzkontrolle anhand von Niedrigwasser-Abflussdaten bzw. Quellschüttungsdaten (Mittel- oder Einzelwerte)	1	Bilanzkontrolle anhand der Abflussdaten (MNQ) vom Vorfluter X, dessen Einzugsgebiet vollständig im Modellgebiet liegt
6	Instationäre Bilanzkontrolle anhand von langen Quellschüttungsganglinien bei Ansatz der zeitlichen Variabilität der GwNeubildung	0	Nicht erforderlich und wegen fehlender Quellschüttungsdaten auch unmöglich.
7	Plausibilitätskontrolle zur Dynamik des GwUmsatzes anhand von hydrochemischen Daten und/oder Daten zur GwAltersstruktur	1	Der modellrechnerische Uferfiltratanteil aus der Infiltrationsrandbedingung (Gewässer Y) muss plausibel sein mit den Spurenstoffanalysen der Flachbrunnen (ca. 50 % Uferfiltratanteil im Förderwasser).
8	Kalibration der GwFließgeschwindigkeit und -fließrichtung anhand von Markierungsversuchsergebnissen	1	Reproduktion der GwFließgeschwindigkeiten im Gewinnungsgebiet gemäß Pump- und Markierungsversuch des Jahres 199X
9	Kalibration der GwFließgeschwindigkeit und -fließrichtung anhand von Beschaffenhheitsdaten	0	-
10	Kalibrierung auf Tracer-Durchbruchskurven, Konzentrationsganglinien oder Stoff-/Temperaturausbreitungsfahnen (Stoff-bzw. Wärmetransportmodellierung)	0	-

3.8 Bewertung der Gesamtgüte des Grundwassermodells

Durch die Verrechnung des Quotienten Summe KP/Summe RP mit dem Ergebnis der Bewertung des HGM resultiert die Gesamtgüte des numerischen Modells wie folgt (Wichtung der HGM-Bewertung anhand der Kalibrierungsgüte):

$$\text{Bewertung HGM (\%)} \times (\text{Summe } [\Sigma] \text{ KP}) / (\text{Summe } [\Sigma] \text{ RP}) = \text{Gesamtgüte numerisches GwModell (\%)}$$

Dieses Ergebnis wäre – analog der HGM-Bewertung - der Einstufung in Anlehnung an das DVGW-Arbeitsblatt W 107 wie folgt zu unterziehen:

- < 33 % Prinzipmodell
- 33 % bis 66 % Planungsmodell
- > 66 % Bewirtschaftungsmodell

Diese Grenzen sind als **Orientierungswerte** zu verstehen.

Für die o. g. Modellkategorien ergeben sich die Geltungsbereiche der Kalibrierung, die in der folgenden Tabelle stichwortartig beschrieben sind:

Tabelle 3.8-9: Modellkategorien in Anlehnung an DVGW-Arbeitsblatt W 107 und Geltungsbereiche der Kalibrierung

Modellkategorie	Zeitdiskretisierung	Kalibrierung	Geltungsbereich der Kalibrierung
Prinzipmodell	stationär/ instationär	keine	○ Prinzipmodelle sind nicht kalibriert
Planungsmodell	stationär	Kalibrierung an historischen Daten	○ Aussagegebiet nach W 107 ○ Stichtagsmessung/Mittelwert einer definierten Zeitspanne
	instationär		○ Aussagegebiet nach W 107 ○ Zeitspanne der Kalibrierung ○ Stress-Situationen aus der Kalibrierung (z. B. Trocken-, Nassperioden, Pumpversuch)
Bewirtschaftungsmodell	instationär	Kalibrierung wird begleitend zur Bewirtschaftung und langfristigen Maßnahmen wiederholt verbessert	○ Aussagegebiet nach W 107 ○ Zeitspanne der Kalibrierung, saisonale und mehrjährige Zyklen beinhaltend ○ Stress-Situationen aus der Kalibrierung

4. Bewertung der Prognosefähigkeit des Grundwassermodells

Prognosefähigkeit meint die Eignung zur Anwendung eines Modells im Rahmen der Wertebereiche der kalibrierten Parameter. Ausgehend von der Gesamtgüte des numerischen Grundwassermodells ist dessen Prognosefähigkeit aufgabenspezifisch wie folgt zu bewerten:

- Die Prognosefähigkeit ist bei **Prinzipmodellen** grundsätzlich nicht gegeben.
- Bei **Planungsmodellen** ist die Prognosefähigkeit in der Regel gegeben, wenn die entsprechende Einstufung des Modells in Anlehnung an das DVGW-Arbeitsblatt W 107 erfolgt bzw. erfolgen kann.
- Bei **Bewirtschaftungsmodellen** ist die Prognosefähigkeit grundsätzlich immer gegeben.

Grundwassermodelle dienen in den meisten Fällen der Berechnung von Prognosen. Dabei ist im Vorfeld zu klären, auf welche Zielgrößen sich die Prognose bezieht; hierauf sollte auch die Kalibrierung ausgerichtet sein (siehe Tabelle 3.7-1 „Zielgrößen der Kalibrierung“). Quantitative Aussagen zu Veränderungen bekannter Zustände sind bei Modellprognosen eher erreichbar als absolute Aussagen.

Wenn auch Prinzipmodelle grundsätzlich nicht prognosefähig sind, dienen sie doch der Klärung von grundsätzlichen Zusammenhängen und haben dort ihre Berechtigung.

Zu berechnende Prognosen basieren auf zu untersuchenden Szenarien oder Varianten. Das können veränderte Lokationen oder veränderte Lastfälle (z. B. veränderte Entnahmen) sein. Die berechneten Prognosen zeigen die Auswirkungen der Szenarien/Varianten/Lastfälle und werden als Entscheidungshilfe bei Planungen und Vorhaben verwendet.

Prognosen sind naturgemäß mit Unsicherheiten behaftet. Es ist wichtig, diese Unsicherheiten zu quantifizieren und bei der Interpretation der Prognosen zu berücksichtigen. Sensitivitätsanalysen können dabei helfen, die Auswirkungen von Unsicherheiten in den Modellparametern und Prozessen zu verstehen.

Veränderungen, die durch den Klimawandel bedingt sind, werden in der Klimaforschung durch Projektionen (Verwendung vieler Modelle und vieler Szenarien) abgebildet. Deshalb sollte in diesem Fall nicht von Prognosen gesprochen werden. Bei Berechnungen von Projektionen kann sich der Stress außerhalb dessen bewegen, was in der Vergangenheit bereits stattfand. Dabei kann eine solche Simulation über die Grenze der in der Kalibrierung verwendeten Wertebereiche hinausgehen und es ist gegebenenfalls eine Begründung für eine trotzdem bestehende Prognosefähigkeit zu erbringen.

Bei der Modellanwendung ist stets zu begründen, wofür ein Modell prognosefähig ist. Geht eine Prognose über die Grenze der in der Kalibrierung verwendeten Wertebereiche hinaus, ist gegebenenfalls eine Begründung für eine trotzdem bestehende Prognosefähigkeit bei der durchgeführten Modellanwendung zu erbringen. Diese lässt sich für Planungs- und Bewirtschaftungsmodelle wie folgt bewerten:

Tabelle 4-10: Bewertung der Prognosefähigkeit von Planungs- und Bewirtschaftungsmodellen

Modellart	Kalibrierung	Prognosefähigkeit
Planungsmodell	stationär an historischen Daten kalibriert	Prognosefähigkeit für mittlere Verhältnisse ist gegeben
Planungsmodell	instationär an historischen Daten kalibriert	Kalibriergüte und Prognosefähigkeit sind gleichwertig, wenn sich die Prognose innerhalb der Wertebereiche der Kalibrierung bewegt
Bewirtschaftungsmodell	begleitend zur Bewirtschaftung und zu langfristigen Maßnahmen	Prognosefähigkeit ist grundsätzlich gegeben, wenn sich die Prognose innerhalb der Wertebereiche der Kalibrierung bewegt



Berechnungen, die über den Gültigkeitsbereich der Kalibrierung hinausgehen, sind als Simulationen bzw. Projektionen zu verstehen (z. B. worst-case-Szenarien).

Neben diesen Aspekten ist zu beachten, dass die Prognosefähigkeit eines Modells auch durch andere Prozesse und die Heterogenität des modellierten Systems eingeschränkt werden kann. Eine vertiefte Darstellung zu den Einflüssen auf die Prognosefähigkeit und -genauigkeit ist beispielsweise den FH-DGGV-Kursunterlagen /4/ zu entnehmen.