

Entwicklung eines Vorhersagemodells für pluviale Überflutungen in urbanen Gebieten auf Basis von Maschinellen Lernverfahren

Zur Erlangung des akademischen Grades
Doktor der Ingenieurwissenschaften
(Dr.-Ing.)

von dem Institut für Wasserbau und Wasserwirtschaft
der Universität Duisburg Essen (UDE) genehmigte

Dissertation

vorgelegt von
Benjamin Burrichter, M.Sc.

Gutachter: Prof. Dr.-Ing. André Niemann, Universität Duisburg-Essen
Institut für Wasserbau und Wasserwirtschaft
Prof. Dr.-Ing. Markus Quirnbach, Hochschule Ruhr West
Lehrgebiet für Siedlungswasserwirtschaft, Hydrologie und Wasserbau

Tag der Einreichung: 12. März 2024

Tag der mündlichen Prüfung: 25. Juni 2024

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub | universitäts
bibliothek

Diese Dissertation wird via DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt und liegt auch als Print-Version vor.

DOI: 10.17185/duepublico/82134

URN: urn:nbn:de:hbz:465-20240731-090838-9

Alle Rechte vorbehalten.

Dank

Die vorliegende Arbeit ist während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Lehr- und Forschungsgebiet Siedlungswasserwirtschaft, Hydrologie und Wasserbau am Institut Bauingenieurwesen der Hochschule Ruhr West entstanden. Es handelt sich um eine kooperative Promotion, die in enger Zusammenarbeit mit dem Institut für Wasserbau und Wasserwirtschaft der Universität Duisburg-Essen angefertigt wurde. Die Idee dazu entstand im Rahmen des vom Ministerium für Innovation, Wissenschaft und Forschung des Landes NRW (MIWF NRW) geförderten Projektes „Smart Water – Chancen und Risiken einer digitalisierten Wasserwirtschaft“. Der Großteil der anschließenden inhaltlichen Arbeit erfolgte im Rahmen des Projektes „KIWaSuS –KI-basiertes Warnsystem vor Starkregen und urbanen Sturzfluten“, dass durch das Bundesministerium für Bildung und Forschung (BMBF) gefördert wurde. Auf diesem Wege bedanke ich mich recht herzlich beim MIWF NRW und dem BMBF für die finanzielle Unterstützung sowie allen Projektbeteiligten für die gute Zusammenarbeit.

Ich danke meinem Erstgutachter Herrn Prof. Dr.-Ing. André Niemann für die Möglichkeit an seinem Institut kooperativ zu promovieren. Vielen Dank auch für die regelmäßigen Gespräche, den geleisteten Input und die angenehme Art der Betreuung. Zudem ermöglichte seine Zusage über die Betreuung dieser externen Arbeit überhaupt erst die Durchführung einer Promotion.

Mein besonderer Dank gilt Herrn Prof. Dr.-Ing. Markus Quirnbach, ohne den diese Arbeit nie entstanden wäre. Ich danke für das mir entgegen gebrachte Vertrauen, die Schaffung der Rahmenbedingungen meiner Arbeit und die unermüdliche Unterstützung in allen Belangen während der Anfertigung dieser Dissertation. Die vielen Gespräche auf fachlicher und auch persönlicher Ebene, haben mich immer weitergebracht und werden mir in guter Erinnerung bleiben.

Meinen Kollegen am Institut Bauingenieurwesen der HRW danke ich für die ausgezeichnete Zusammenarbeit in den vergangenen Jahren und die vielen schönen Momente. Hervorzuheben sind hierbei Juliana, Saphira, Henrike, Piet und Kristian vom „Team Wasser“, vielen Dank für die vielen Gespräche, Diskussionen und das Gegenlesen der Arbeit.

Mein größter Dank gilt jedoch meiner Familie. Meiner Frau Julia, die für mich immer einen unverzichtbaren Rückhalt darstellte und mir genügend Freiraum für meine Dissertation schaffte. Meinen Kindern Felina und Milo, die mir immer vor Augen geführt haben, dass es auch wichtigere Dinge im Leben gibt. Meinen Eltern, die mich bei meinem Werdegang immer unterstützt haben und mir mit Rat und Tat zur Seite standen.

Kurzfassung

Starkregenereignisse und die daraus resultierenden Sturzfluten führen insbesondere in den hochverdichteten urbanen Gebieten immer wieder zu einer Überlastung der Entwässerungssysteme und in der Folge zu Überschwemmungen. Da derzeit keine geeigneten Frühwarnsysteme für solche Ereignisse existieren, kann die konkrete Überflutungssituation aufgrund überlasteter Kanalnetze und oberflächlich abfließender Wassermaßen erst beim Eintreten eines Ereignisses beschrieben werden. Dadurch ist die zielgerichtete Warnung der betroffenen Bevölkerung nur unzureichend möglich. Zudem wird die Handlungsfähigkeit der beteiligten Akteure des Krisenmanagements stark beeinträchtigt und beschränkt sich häufig nur noch auf die Schadensbeseitigung.

Da konvektive Starkregenereignisse nur für kurze Zeiträume vorhergesagt werden können, die Berechnung der resultierenden Überflutungssituation mit den verfügbaren Modellen aber sehr rechenintensiv ist, ist eine effektive Frühwarnung derzeit nicht realisierbar. Um dieser Problematik zu begegnen, wurde im Rahmen der vorliegenden Arbeit ein Ansatz entwickelt, mit dem auf Basis einer Niederschlagsprognose das resultierende Überflutungsmaß möglichst verzögerungsarm berechnet werden kann. Ziel war es dabei, die bevorstehende Überflutungssituation als dynamische Starkregengefahrenkarte für die kommenden Zeitschritte darzustellen. Hierfür wurde auf Maschinelle Lernverfahren (ML) zurückgegriffen, deren Training zwar sehr rechen- und damit zeitintensiv ist, mit denen im Anschluss an den Trainingsprozess aber eine schnelle Berechnung der gewünschten Zielgröße möglich ist.

Zur Berechnung der bevorstehenden Überflutungssituation wurde ein zweigeteilter Systemaufbau, bestehend aus einem ML-Modell zur Vorhersage von Überstauganglinien an den Schächten eines Kanalnetzes und einem ML-Modell zur Vorhersage der resultierenden Überflutungsflächen an der Geländeoberfläche entwickelt. Da für das Training der beiden Modelle keine geeigneten Beobachtungsdaten über das Überflutungsmaß historischer Ereignisse zur Verfügung standen, wurde ein umfangreicher künstlicher Trainingsdatensatz mit einem hydrodynamischen Berechnungsmodell erstellt. Mit diesem Datensatz wurden anschließend die beiden Vorhersagemodelle für Überstau und für Überflutungsflächen entwickelt.

Im Rahmen des Entwicklungsprozesses wurden verschiedene ML-Verfahren und Modellaufbauten getestet und hinsichtlich der Prognosegüte miteinander verglichen. Zudem wurde der Einfluss unterschiedlicher Kombinationen von Eingabedaten auf das Prognoseergebnis analysiert. Die besten Modelle wurden zusätzlich verschiedenen Sensitivitätsanalysen unterzogen, um die Genauigkeit der entwickelten Modelle weiter zu optimieren sowie deren Skalierbarkeit und Übertragbarkeit zu verbessern. Abschließend erfolgte einerseits die Kopplung der beiden Vorhersagemodelle zur Bewertung der gemeinsamen Prognosegüte, andererseits wurde die Integration der beiden Modelle in die Vorhersagekette eines Frühwarnsystems aufgezeigt.

Abstract

Heavy rainfall events and the resulting flash floods repetitively overload the capacity of drainage systems and lead to flooding, particularly in highly urbanised areas. As there are currently no appropriate early warning systems for such events, the specific flooding situation due to the overload of the sewer networks and to the generated surface runoff can only be determined when an event occurs. As a result, the targeted warning of the affected population is only insufficiently possible. In addition, the ability of crisis management actors to respond during these extreme events is severely impaired, being often limited to damage repair.

Since convective rainfall events can only be predicted for short lead times, and calculating the resulting flooding situation with the available models is, on the other hand, very computationally intensive, an adequate early warning is currently not feasible. To overcome this problem, an approach to calculate the resulting flood extent based on a precipitation nowcast with as little temporal delay as possible was developed in this study. The aim was to visualise the upcoming flooding situation as a dynamic heavy rain hazard map for the subsequent time steps. For this purpose, machine learning (ML) methods were applied, whose training is very computationally intensive and consequently time-consuming, but which allow a rapid calculation of the desired target value, once the training process has been completed.

To estimate the upcoming flooding situation, a system structure consisting of two sub-models was developed, one ML model for predicting overflow hydrographs at the manholes of a sewer network and one ML model for predicting the resulting flooded areas at the terrain surface. As no suitable monitoring data about the flooding situation of historical events was available for the training of these two models, a comprehensive artificial training data set was created with a hydrodynamic simulation model. This data set was then used to develop the two forecasting models for overflow and flooded areas.

As part of the development process, various ML methods and modelling setups were tested and compared with each other in terms of their forecasting performance. The influence of different combinations of input data on the forecast result was also analysed. Additionally, the best models were subjected to various sensitivity analyses to further optimise the accuracy of the developed models and improve their scalability and transferability. Finally, the two forecasting models were coupled to evaluate the combined forecasting performance and the integration of the two models into the forecasting chain of an early warning system was demonstrated.

Anmerkungen

Teilergebnisse der vorliegenden Arbeit wurden in den folgenden Veröffentlichungen vorab publiziert:

Artikel in Fachzeitschriften

Burrichter, B.; Koltermann da Silva, J.; Niemann, A.; Quirnbach, M. (2024): A Temporal Fusion Transformer Model to Forecast Overflow from Sewer Manholes during Pluvial Flash Flood events. *Hydrology* 11(3), S. 41. DOI: 10.3390/hydrology11030041

Burrichter, B.; Hofmann, J.; Koltermann da Silva, J.; Niemann, A.; Quirnbach, M. (2023): A Spatiotemporal Deep Learning Approach for Urban Pluvial Flood Forecasting with Multi-Source Data. *Water* 15 (9), S. 1760. DOI: 10.3390/w15091760.

Frentrup S., Schultheis H., Quirnbach M., **Burrichter B.**, et al. (2022): Intelligentes Management von Datenströmen und KI-Anwendungen in KIWaSuS. *KA Korrespondenz Abwasser, Abfall* 69 (4), S. 264-270.

Burrichter, B.; Quirnbach, M. (2021): Projekt KIWaSuS –mit künstlicher Intelligenz Frühwarnsysteme entwickeln und die Bevölkerung rechtzeitig warnen. *abwasserreport 21* (4), S. 15–16.

Burrichter, B.; Quirnbach, M.; Oelmann, M.; Niemann, A. (2021): Künstliche Intelligenz in der Wasserwirtschaft. *KA Korrespondenz Abwasser, Abfall* 68 (2), S. 94–101.

Konferenzbeiträge

Koltermann da Silva, J.; **Burrichter, B.**; Quirnbach, M. (2023): Application of Artificial Intelligence in Rainfall Nowcasting and Flash Floods Forecast in an urban catchment: First results from Research Project KIWaSuS in Germany. Novatech 2023. Lyon, France, 3. - 7. July 2023. Online verfügbar unter hal-04176936.

Burrichter, B.; Koltermann da Silva, J.; Quirnbach, M. (2023): Räumliche und zeitliche Vorhersage pluvialer Überflutungen unter Berücksichtigung verschiedener Eingabedatenströme: Ein Deep Learning Ansatz. Vortrag zum Tag der Hydrologie 2023. Bochum, 21. - 23. März 2023.

Burrichter, B.; Quirnbach, M. (2023): KI-basierte Überflutungsvorhersage im urbanen Raum. In: *IWARU Institut für Infrastruktur · Wasser · Ressourcen · Umwelt (Hrsg.), Klimawandel – Trockenheit und Starkregen im urbanen Raum*. 7. Wassertage Münster. Münster, 14./15. Februar 2023, S. 49-55. ISBN: 978-3-947263-34-9

Burrichter, B.; Koltermann da Silva, J.; Quirnbach, M. (2022): KI-basierte Vorhersage kanalinduzierter Überflutungen. In: *Grün statt grau. Mit Blau-Grünen Infrastrukturen gemeinsam die Siedlungsentwässerung der Zukunft planen*. Aqua Urbanica. Glattfelden, Schweiz, 14. /15. November 2022, S. 218–225.

Koltermann da Silva, J.; **Burrichter, B.**; Quirnbach, M. (2022): *Vorhersagemodelle in KIWa-SuS: Vorverarbeitungsschritte für die Entwicklung von KI-Modellen*. In: Disse, M.; Ludwig, R.; Reisenbüchler, M. (Hrsg.): *Forum für Hydrologie und Wasserbewirtschaftung. Im Wandel - Klima, Wasser und Gesellschaft: Prozesse - Methoden - Kommunikation*. Tag der Hydrologie. München, 22./23. März 2022. Heft 43.22, S. 81–90.

Im konkreten Fall wurden die entsprechenden Stellen in der vorliegenden Arbeit durch Zitate gekennzeichnet.

Inhaltsverzeichnis

| | |
|---|-------------|
| Inhaltsverzeichnis | IX |
| Abbildungsverzeichnis | XIII |
| Tabellenverzeichnis | XIX |
| Abkürzungs- und Symbolverzeichnis | XXI |
| 1 Einleitung..... | 1 |
| 1.1 Veranlassung | 1 |
| 1.2 Zielsetzung und Schwerpunkt der Arbeit | 2 |
| 1.3 Aufbau der Arbeit | 3 |
| 2 Überflutungsmodellierung und -vorhersage in urbanen Gebieten | 5 |
| 2.1 Abgrenzung des Begriffs der Überflutungsmodellierung im Sinne dieser Arbeit | 5 |
| 2.2 Allgemeine Grundlagen zur Modellierung | 6 |
| 2.2.1 Modellkonzepte..... | 6 |
| 2.2.2 Maschinelles Lernen | 10 |
| 2.3 Abfluss- und Überflutungsmodellierung in der Siedlungsentwässerung..... | 12 |
| 2.3.1 Hintergrund und allgemeine Grundlagen..... | 12 |
| 2.3.2 Niederschlag-Abfluss-Modellierung..... | 14 |
| 2.3.3 Analyse der Überflutungsgefährdung | 18 |
| 2.4 Abfluss- und Überflutungsvorhersage für den Echtzeiteinsatz..... | 20 |
| 2.4.1 Physikalisch-basierte und konzeptionelle Modelle..... | 21 |
| 2.4.2 Datengetriebene Ansätze..... | 24 |
| 2.5 Ansatzpunkte, Untersuchungsmethodik und thematische Abgrenzung..... | 27 |
| 3 Material und Methoden | 30 |
| 3.1 Beschreibung des Untersuchungsgebietes | 30 |
| 3.2 Verwendete Datengrundlage..... | 32 |
| 3.2.1 Historische Niederschlagsdaten..... | 32 |
| 3.2.2 Geländemodell..... | 33 |
| 3.2.3 Kanalnetzdaten | 34 |
| 3.2.4 Daten zur Oberflächenbeschaffenheit | 34 |

| | | |
|----------|--|-----------|
| 3.3 | Künstlich Neuronale Netze | 35 |
| 3.3.1 | Grundlagen | 35 |
| 3.3.2 | Netzarchitekturen | 40 |
| 3.3.2.1 | Rekurrente Neuronale Netze (<i>Recurrent Neural Networks</i> , RNN) | 40 |
| 3.3.2.2 | Konvolutionelle Neuronale Netze (<i>Convolutional Neural Networks</i> , CNN) | 42 |
| 3.3.2.3 | <i>Graph Convolutional Networks</i> (GCN) | 44 |
| 3.3.2.4 | <i>Generative Adversarial Networks</i> (GAN) | 47 |
| 3.3.2.5 | <i>Transformer</i> | 48 |
| 3.4 | Verwendete Software und Berechnungsansätze | 50 |
| 3.4.1 | Physikalisch-basierte Überflutungsmodellierung | 50 |
| 3.4.2 | Implementierung der ML-Verfahren und Datenvorverarbeitungsroutinen | 54 |
| 3.4.3 | Sonstige Software | 56 |
| 3.5 | Gütekriterien | 56 |
| 4 | Datengenerierung und -aufbereitung | 59 |
| 4.1 | Generierung von Niederschlagsreihen | 60 |
| 4.1.1 | Modellregenreihen | 60 |
| 4.1.2 | Natürliche Regenreihen | 62 |
| 4.2 | Aufbau des gekoppelten 1D/2D-Überflutungsmodells | 63 |
| 4.3 | Modellkalibrierung | 67 |
| 4.4 | Hydrodynamische Simulationen (Datengenerierung) | 70 |
| 4.5 | Zusammenfassung und Bewertung der Ergebnisse | 71 |
| 5 | Modellentwicklung und -analyse: Überflutungsflächenvorhersage | 74 |
| 5.1 | Problembeschreibung | 74 |
| 5.2 | Datenvorverarbeitung und -aufteilung | 77 |
| 5.3 | Untersuchungen zum Modellaufbau | 79 |
| 5.3.1 | Untersuchung 1: Vergleich unterschiedlicher Eingabegrößen | 81 |
| 5.3.2 | Untersuchung 2: Vergleich unterschiedlicher Vorverarbeitung der Überstaudaten | 83 |
| 5.3.3 | Untersuchung 3: Vergleich der bisher besten Modellstruktur (T-GCN) mit einem cGAN | 85 |
| 5.4 | Sensitivitätsanalysen | 88 |
| 5.4.1 | Einfluss der Trainingsdatenquantität und -verteilung | 88 |

| | | |
|----------|--|------------|
| 5.4.2 | Übertragbarkeit des Modells auf andere Gebiete | 94 |
| 5.4.3 | Analyse des Einflusses von Hyperparametern (Hyperparameter tuning)..... | 96 |
| 5.5 | Beurteilung der Vorhersagegenauigkeit..... | 99 |
| 5.6 | Zusammenfassung und Bewertung der Ergebnisse | 104 |
| 6 | Modellentwicklung und -analyse: Überstauvorhersage | 107 |
| 6.1 | Problembeschreibung | 107 |
| 6.2 | Datenvorverarbeitung und -aufteilung | 108 |
| 6.3 | Untersuchungen zum Modellaufbau | 110 |
| 6.3.1 | Auswahl und Vergleich von ML-Verfahren/Netzarchitekturen..... | 110 |
| 6.3.2 | Untersuchungen zum Einfluss von Messungen im Kanal..... | 113 |
| 6.3.3 | Vergleich der Ergebnisse | 115 |
| 6.4 | Sensitivitätsanalysen..... | 118 |
| 6.4.1 | Einfluss der Trainingsdatenquantität und -verteilung..... | 119 |
| 6.4.2 | Übertragbarkeit des Modells auf andere Gebiete | 122 |
| 6.4.3 | Analyse des Einflusses von Hyperparametern (Hyperparameter tuning)..... | 122 |
| 6.5 | Beurteilung der Vorhersagegenauigkeit..... | 124 |
| 6.6 | Zusammenfassung und Bewertung der Ergebnisse | 127 |
| 7 | Zusammenführung des Gesamtsystems..... | 130 |
| 7.1 | Gekoppelte Betrachtung Überstau- und Überflutungsflächenvorhersage | 130 |
| 7.2 | Integration der Vorhersagemodelle in einen Prognose- und Warnprozess | 134 |
| 8 | Zusammenfassung und weiterer Forschungsbedarf..... | 136 |
| 8.1 | Zusammenfassung der wesentlichen Erkenntnisse | 136 |
| 8.2 | Limitierungen und weiterer Forschungsbedarf..... | 138 |
| | Literaturverzeichnis | 143 |
| | Anhang..... | 161 |

Abbildungsverzeichnis

| | | |
|----------------|--|----|
| Abbildung 1-1: | Wissenschaftlicher Aufbau der Arbeit..... | 4 |
| Abbildung 2-1: | Konzepte zur Abbildung von Prozessen in Modellen..... | 7 |
| Abbildung 2-2: | Ansätze zur Kombination von <i>White-Box</i> - mit <i>Black-Box</i> -Ansätzen zu hybriden Modellen..... | 9 |
| Abbildung 2-3: | Vergleich prozessbasierter Modellierung und ML-basierter Modellierung (verändert nach Abbildung 1, Burrichter et al., 2021) | 10 |
| Abbildung 2-4: | Unterteilung von gängigen ML-Verfahren in Abhängigkeit von Lernstil und -aufgabe..... | 11 |
| Abbildung 2-5: | Übersicht über die bei der Niederschlag-Abfluss-Modellierung in der Siedlungsentwässerung berücksichtigten Komponenten, Prozesse und Berechnungsansätze | 14 |
| Abbildung 2-6: | Methoden und Verfahren zur Analyse der Überflutungsgefährdung (verändert nach Tabelle 2, DWA M-119 (2016))..... | 19 |
| Abbildung 2-7: | Übergeordneter Systemaufbau und Untersuchungsmethodik | 28 |
| Abbildung 3-1: | Übersicht über das Untersuchungsgebiet: a) Übersichtskarte und Detailansicht des Untersuchungsgebiets mit Darstellung des Entwässerungssystems sowie den zugehörigen Einzugsgebietsflächen (Datenquelle: AGG (2021) und EGLV (2021b)) b) Ausgeschilderte Überschwemmungsgefahr vor einer Unterführung im Untersuchungsgebiet c) Ausmaß der Überflutung nach einem Starkregenereignis am 03.07.2009 (Quelle: Screenshot aus einem YouTube Video (Steffi86w 2009)) | 31 |
| Abbildung 3-2: | Übersicht über die berücksichtigten Stationen von Emschergenossenschaft und Lippeverband und deren Lage zum Untersuchungsgebiet, Datenquelle: EGLV (2021a)..... | 32 |
| Abbildung 3-3: | Schematische Darstellung eines mehrschichtigen Perzeptrons mit Detailansicht eines einzelnen Perzeptrons nach Frochte (2021)..... | 36 |
| Abbildung 3-4: | Übersicht über drei gängige Aktivierungsfunktionen..... | 37 |
| Abbildung 3-5: | Schematische Darstellung des <i>Overfittings</i> (Woernle 2008), links: Darstellung einer überangepassten Modellfunktion (gestrichelte Linie) und einer Modellfunktion mit höherem Fehler aber besserer Generalisierung (durchgezogene Linie), rechts: Entwicklung der Fehler von Trainings- und Testdaten mit Fortschritt des Trainings | 40 |
| Abbildung 3-6: | Schematische Darstellung eines RNNs mit Rückkopplung (links) und eines RNNs im ausgerollten Zustand (rechts) (Raschka und Mirjalili | |

| | | |
|-----------------|---|----|
| | 2018)..... | 40 |
| Abbildung 3-7: | Vergleich einer RNN- und einer LSTM-Zelle (verändert nach Christopher Olah, 2015 und Raschka und Mirjalili, 2018) | 41 |
| Abbildung 3-8: | Schematische Darstellung einer 2D-Faltung in einer Faltungsschicht eines CNNs..... | 43 |
| Abbildung 3-9: | Genereller Aufbau eines CNNs mit einer Abwechselnden Anordnung von Faltungs- und Poolingschichten, sowie zwei vollständig verbundenen Schichten, die die Netzausgabe erzeugen (Raschka und Mirjalili 2018)..... | 44 |
| Abbildung 3-10: | Aggregation der Informationen von den Nachbarknoten in einem GCN: Der dargestellte Zielknoten A aggregiert die Informationen der direkten Nachbarknoten B, C und D. Die Informationen von B, C und D basieren wiederum auf den aggregierten Informationen aus deren Nachbarschaften und so weiter (verändert nach Hamilton et al., 2017)..... | 46 |
| Abbildung 3-11: | Exemplarischer Aufbau eines <i>Generative Adversarial Networks</i> . Zur Generierung der Abbildung wurde auf den MNIST-Datensatz (LeCun et al. 2010) zurückgegriffen..... | 47 |
| Abbildung 3-12: | <i>Transformer</i> -Architektur mit Aufmerksamkeitsmechanismus in Form von <i>Multi-Head Attention</i> -Schichten (Vaswani et al. 2017)..... | 49 |
| Abbildung 3-13: | Darstellung der <i>Temporal Fusion Transformer</i> -Architektur (links) sowie Detailansicht der <i>Gated Residual Network</i> -Blöcke und der <i>Variable Selection Network</i> -Blöcke (rechts) (Lim et al. 2019)..... | 50 |
| Abbildung 4-1: | Schematische Darstellung der berücksichtigten Modellregen sowie der gewählten Dauerstufen, Modellregenarten und Wiederkehrzeiten/Szenarien. Bei den Szenarien S 1,5 und S 4,0 gibt die Nummer jeweils den Erhöhungsfaktor an, mit dem die Werte der 100-jährigen Modellregen multipliziert wurden | 61 |
| Abbildung 4-2: | Häufigkeitsverteilung der Ereignisse je Wiederkehrintervall getrennt für die berücksichtigten Dauerstufen und zusammengefasst für die maximale Wiederkehrzeit der Ereignisse. Die Auswertung wurde erstellt auf Basis von EGLV (2021a)..... | 63 |
| Abbildung 4-3: | Übersicht über das 1D-Kanalnetzmodell mit den zugehörigen (Teil-) Einzugsgebieten sowie dem Bereich des Oberflächenmodells, Datenquelle: AGG (2021) und EGLV (2021b) | 65 |
| Abbildung 4-4: | Schematische Darstellung der Datengenerierung mit den relevanten Ein- und Ausgaben..... | 71 |
| Abbildung 5-1: | Geplanter Modellaufbau mit allen potentiellen Eingaben (links) und der gesuchten Zielgröße (rechts) | 75 |

| | | |
|-----------------|---|----|
| Abbildung 5-2: | Darstellung eines Ausschnitts von den als räumliche Eingabedaten berücksichtigten Größen | 77 |
| Abbildung 5-3: | Umwandlung der Daten in ein überwachtetes Lernproblem am Beispiel eines einzelnen Trainingspärchens | 79 |
| Abbildung 5-4: | Ausgangsarchitektur für die nachfolgenden Untersuchungen (Die Zahlen über den einzelnen Schichten geben jeweils die Anzahl der Merkmale bzw. Merkmalskarten an, die durch die Anzahl der berücksichtigten Filter der jeweils zuvor durchgeführten Faltungsfunktion (als Pfeil dargestellt) festgelegt wird. Bei der ersten Schicht bezieht sich der Wert auf die Anzahl der Eingabemerkmale.) | 80 |
| Abbildung 5-5: | Ausgangsarchitektur mit den jeweiligen Eingabepfaden für die Überstauvorhersage und die räumlichen Informationen (Die Zahlen über den einzelnen Schichten geben jeweils die Anzahl der Merkmale bzw. Merkmalskarten an, die durch die Anzahl der berücksichtigten Filter der jeweils zuvor durchgeführten Faltungsfunktion (als Pfeil dargestellt) festgelegt wird. Bei den ersten Schichten bezieht sich der Wert auf die Anzahl der Eingabemerkmale.) | 82 |
| Abbildung 5-6: | Ausgangsarchitektur mit den verschiedenen Eingabepfaden für die unterschiedlichen Formate der Überstauvorhersage (Die Zahlen über den einzelnen Schichten geben jeweils die Anzahl der Merkmale bzw. Merkmalskarten an, die durch die Anzahl der berücksichtigten Filter der jeweils zuvor durchgeführten Faltungsfunktion (als Pfeil dargestellt) festgelegt wird. Bei der ersten Schicht bezieht sich der Wert auf die Anzahl der Eingabemerkmale.) | 84 |
| Abbildung 5-7: | Schematische Vorgehensweise zur Erzeugung der Überstaurastersequenzen | 85 |
| Abbildung 5-8: | Aufbau und Funktionsweise des conditional GANs (durch den gestrichelten Pfeil soll verdeutlicht werden, dass der Diskriminator abwechselnd eine durch den Generator generierte Überflutungskarte und eine „reale Überflutungskarte“ als Eingabe erhält) | 86 |
| Abbildung 5-9: | Aufbau des Diskriminators | 87 |
| Abbildung 5-10: | Verteilung der Metriken über alle 26 Ereignisse im Testdatensatz für die Varianten Alle (gesamter Datensatz), MR (Modellregenreihen) und NR (natürliche Regenreihen) | 89 |
| Abbildung 5-11: | Übersicht über die fünf berücksichtigten Varianten zur Auswahl von Ereignissen aus den natürlichen Regenreihen (nicht berücksichtigte Ereignisse sind transparent dargestellt) | 90 |
| Abbildung 5-12: | Verteilung der Metriken über alle 26 Ereignisse im Testdatensatz für die fünf berücksichtigten Varianten zur Auswahl von Ereignissen aus den natürlichen Regenreihen. Zur besseren Darstellung werden beim RMSE | |

| | | |
|-----------------|---|-----|
| | die unteren drei Grenzwerte zusätzlich vergrößert. | 91 |
| Abbildung 5-13: | Übersicht über die fünf berücksichtigten Varianten zur Auswahl von Ereignissen aus den Modellregenreihen (nicht berücksichtigte Ereignisse sind transparent dargestellt) | 91 |
| Abbildung 5-14: | Verteilung der Metriken über alle 26 Ereignisse im Testdatensatz für die fünf berücksichtigten Varianten zur Auswahl von Ereignissen aus den Modellregenreihen. Zur besseren Darstellung wurden beim RMSE die unteren drei Grenzwerte zusätzlich vergrößert..... | 92 |
| Abbildung 5-15: | Vergleich der besten Varianten aus den vorangegangenen Untersuchungen mit einer Kombination aus gleichmäßig ausgewählten Modellregen (Variante 5 Abbildung 5-13) und gleichverteilt ausgewählten natürlichen Regenreihen (Variante 2 Abbildung 5-11) | 93 |
| Abbildung 5-16: | Untersuchungsgebiet mit den maximalen im Datensatz enthaltenen Wassertiefen und dem verwendeten Raster für die Erzeugung von <i>Patches</i> . Das Testgebiet ist rot umrandet, die übrigen Gebiete wurden für das Training verwendet..... | 95 |
| Abbildung 5-17: | Vergleich unterschiedlicher Längen des berücksichtigten Eingabefensters für die Niederschlagsinformationen..... | 97 |
| Abbildung 5-18: | Übersicht über die RMSE-Werte für die einzelnen Hyperparameter-Kombinationen (die drei besten Ergebnisse sind rot eingekreist) | 98 |
| Abbildung 5-19: | Verteilung der Metriken über alle 26 Ereignisse im Testdatensatz. Die Bezeichnung der Modelle ergibt sich jeweils aus den Werten der Hyperparameter in der Reihenfolge Netztiefe-Basisfilter-Kernelgröße-Verlustfunktion | 99 |
| Abbildung 5-20: | Verteilung der Metriken in Abhängigkeit verschiedener Wiederkehrintervalle (die blaue, gestrichelte Linie markiert jeweils das Optimum) | 100 |
| Abbildung 5-21: | 2D-Histogramm mit dem pixelweisen Vergleich der vorhergesagten und simulierten Wasserstände (links) und dem pixelweisen Vorhersagefehler in Abhängigkeit der simulierten Wasserstände (rechts) | 100 |
| Abbildung 5-22: | Niederschlagsganglinie mit markiertem Vorhersagezeitpunkt zu Beginn des Ereignisses vom 3. Juli 2009 in Gelsenkirchen (oben) und Vorhersageergebnis und -evaluierung für drei Zeitschritte der Vorhersage. Der nachfolgend in Abbildung 5-23 dargestellte Detailausschnitt ist rot markiert | 102 |
| Abbildung 5-23: | Detailansicht des Vorhersageergebnisses und der zugehörigen Evaluierung für einen Ausschnitt mit einer überfluteten Unterführung | 103 |
| Abbildung 6-1: | Darstellung des zu lösenden Lernproblems..... | 108 |

| | | |
|-----------------|---|-----|
| Abbildung 6-2: | Umwandlung der Daten in ein überwachtes Lernproblem | 109 |
| Abbildung 6-3: | Exemplarische Darstellung der berücksichtigten Modellaufbauten (i) LSTM, (ii) CNN, (iii) Seq2Seq und (iv) DA-RNN. Die Eingaben x und Ausgaben y können zu jedem Zeitpunkt i auch mehrere Merkmale umfassen. Der grüne Pfeil bei den LSTM-Schichten stellt jeweils die rekurrente Verbindung dar, über die die Zustände jeweils an den folgenden Zeitpunkt weitergegeben werden..... | 111 |
| Abbildung 6-4: | Übersicht über das Kanalnetz und das künstliche Sensornetz im Untersuchungsgebiet, Datenquelle: AGG (2021) | 114 |
| Abbildung 6-5: | Violinen-Plot der Metriken für die besten drei Modelle der Variante mit einer Messtation und Füllungsklassen als Messsignal | 117 |
| Abbildung 6-6: | Einfluss der einzelnen Merkmale der Encoder und Decoder Eingaben für a) die Variante mit 20 Stationen und b) die Variante mit einer Station (die Merkmale mit M zu Beginn geben jeweils Messungen der Füllstandsklassen mit der jeweiligen Haltungs-ID an)..... | 118 |
| Abbildung 6-7: | Violinen-Plot der Metriken für die drei berücksichtigten Varianten (berücksichtigt werden nur Prognosen, bei denen das kumulierte Volumen des HD-Modells oder des ML-Modells > 100 l liegt) | 120 |
| Abbildung 6-8: | Violinen-Plot der Metriken für die berücksichtigten Varianten zur Auswahl von Ereignissen aus den Modellregenreihen (berücksichtigt werden nur Prognosen, bei denen das kumulierte Volumen des HD-Modells oder des ML-Modells > 100 l liegt) | 121 |
| Abbildung 6-9: | Violinen-Plot der Metriken für die Modellregenreihen und die natürliche Regenreihen gleichverteilt ausgedünnt. Als Vergleichsgrößen sind die Variante mit dem gesamten Datensatz sowie den gleichmäßig ausgedünnten Modellregenreihen dargestellt (berücksichtigt werden nur Prognosen, bei denen entweder das kumulierte Volumen des HD-Modells oder des ML-Modells > 100 l liegt) | 122 |
| Abbildung 6-10: | Violinen-Plot der Metriken für den ursprünglichen und den optimierten TFT (berücksichtigt werden nur Prognosen, bei denen entweder das kumulierte Volumen des HD-Modells oder des ML-Modells > 100 l liegt) | 124 |
| Abbildung 6-11: | Volumenfehler (VE), Maximalwertabweichung (PE) und zeitliche Maximalwertabweichung (PTE) in Abhängigkeit der Wiederkehrzeit der Ereignisse (die blaue, gestrichelte Linie markiert jeweils das Optimum).. | 125 |
| Abbildung 6-12: | Scatterplot mit dem Vergleich des vorhergesagten und simulierten Überstauvolumens je Vorhersage (links) und dem Vorhersagefehler in Abhängigkeit des simulierten Überstauvolumens (rechts) | 125 |
| Abbildung 6-13: | Vorhersageergebnis an sechs Schächten für die historischen Ereignisse | |

| | | |
|----------------|--|-----|
| | vom 03. Juli 2009 und 03. Juli 2010 in Gelsenkirchen..... | 126 |
| Abbildung 7-1: | Verteilung der Metriken über alle 26 Ereignisse im Testdatensatz für die Variante mit durch das HD-Modell simulierten Überstauganglinien (Simulierte Vorhersage) und die mit dem Überstauvorhersagemodell gekoppelte Variante (Gekoppelte Vorhersage)..... | 130 |
| Abbildung 7-2: | 2D-Histogramm mit dem pixelweisen Vergleich der vorhergesagten und simulierten Wasserstände (links) und dem pixelweisen Vorhersagefehler in Abhängigkeit der simulierten Wasserstände (rechts) | 131 |
| Abbildung 7-3: | Niederschlagsganglinie mit markiertem Vorhersagezeitpunkt zu Beginn des Ereignisses vom 3. Juli 2009 in Gelsenkirchen (oben) und Vorhersageergebnis und -evaluierung für drei Zeitschritte der Vorhersage (unten). Der nachfolgend in Abbildung 7-4 dargestellte Detailausschnitt ist rot markiert | 132 |
| Abbildung 7-4: | Detailansicht des Vorhersageergebnisses und der zugehörigen Evaluierung für einen Ausschnitt mit einer überfluteten Unterführung | 133 |
| Abbildung 7-5: | Exemplarische Darstellung der einzelnen am Prognoseprozess beteiligten Komponenten (verändert nach Frentrup et al., 2024)..... | 135 |

Tabellenverzeichnis

| | | |
|--------------|--|----|
| Tabelle 3-1: | Übersicht über die berücksichtigten Stationen und die jeweiligen Messzeiträume..... | 33 |
| Tabelle 3-2: | Kennzahlen des verwendeten Kanalnetzmodells | 34 |
| Tabelle 3-4: | Übersicht über unterschiedliche Arten von Graphen sowie den zugehörigen Adjazenzmatrizen (Die Zeilen und Spalten sind hier jeweils alphabetisch geordnet)..... | 45 |
| Tabelle 3-3: | Übersicht über die hydrologischen Modelle zur Oberflächenabflussberechnung in MIKE+ (verändert nach DHI (2015)) | 52 |
| Tabelle 3-5: | Übersicht der verwendeten Bibliotheken für die Datenverarbeitung und Modellerstellung | 54 |
| Tabelle 4-1: | Parametersätze der einzelnen Flächennutzungen für den verwendeten Modellansatz C2 | 64 |
| Tabelle 4-2: | Flächenanteile im Untersuchungsgebiet..... | 66 |
| Tabelle 4-3: | Flächenspezifische Parametrisierung des Oberflächenmodells..... | 66 |
| Tabelle 4-4: | Berücksichtigte Niederschlagsereignisse | 68 |
| Tabelle 4-5: | Einordnung der Abweichungsmaße in ein Bewertungsschema (DWA-AG ES-2.6, zitiert nach DWA-M 165 (2021)) | 68 |
| Tabelle 4-6: | Bewertung der Modellgüte für die beste Modellvariante | 69 |
| Tabelle 4-7: | Bewertung der Modellgüte für das kalibrierte Modell | 70 |
| Tabelle 5-1: | Aufteilung der Datensätze in Trainings-, Validierungs- und Testdatensätze | 78 |
| Tabelle 5-2: | Ergebnis der Evaluierung zum Vergleich unterschiedlicher Modelleingaben (für jede Untersuchung und Metrik wurde das beste Ergebnis fett markiert)..... | 83 |
| Tabelle 5-3: | Ergebnis der Evaluierung zum Vergleich unterschiedlicher Formate der Überstauvorhersage (für jede Untersuchung und Metrik wurde das beste Ergebnis fett markiert)..... | 85 |
| Tabelle 5-4: | Ergebnis der Evaluierung zum Vergleich unterschiedlicher Modellarchitekturen (für jede Untersuchung und Metrik wurde das beste Ergebnis fett markiert)..... | 87 |
| Tabelle 5-5: | Übersicht über die Anzahl der berücksichtigten Ereignisse und Samples sowie deren Anteil am Ausgangsdatsatz | 93 |
| Tabelle 5-6: | Ausgewählte Hyperparameter und berücksichtigte Werte | 98 |

| | | |
|--------------|---|-----|
| Tabelle 6-1: | Übersicht der bei den einzelnen Modellaufbauten berücksichtigten Hyperparameter | 113 |
| Tabelle 6-2: | Ergebnis der Evaluierung aller Modelle und Varianten (für jede Variante und Metrik wurde das beste Ergebnis fett markiert, das beste Gesamtergebnis jeder Metrik wurde zusätzlich unterstrichen)..... | 116 |
| Tabelle 6-3: | Mittlerer Fehler und Parametrisierung der besten fünf Optimierungsläufe | 123 |

Abkürzungs- und Symbolverzeichnis

| Abkürzung | Erläuterung |
|------------------|---|
| $A_{E,k}$ | Kanalisierte Einzugsgebietsfläche [ha] |
| ALS | Flugzeuggestütztes Laserscanning (<i>Airborne Laserscanning</i>) |
| ANN | Künstlich Neuronales Netz (<i>Artificial Neural Network</i>) |
| BPTT | <i>Backpropagation-through-time</i> |
| CA | Zelluläre Automaten (<i>Cellular Automata</i>) |
| cGAN | <i>conditional Generative Adversarial Network</i> |
| CNN | Konvolutionelle Neuronale Netze (<i>Convolutional Neural Network</i>) |
| CPU | <i>Central Processing Unit</i> |
| CSI | <i>Critical Success Index</i> |
| D | Dauerstufe [min] |
| DA-RNN | <i>Dual-Stage Attention-Based Recurrent Neural Network</i> |
| DGM | Digitales Geländemodell |
| DL | <i>Deep Learning</i> |
| EGLV | Emschergenossenschaft und Lippeverband |
| GAN | <i>Generative Adversarial Network</i> |
| GCN | <i>Graph Convolutional Network</i> |
| GIS | Geografisches Informationssystem |
| GNN | <i>Graph Neural Network</i> |
| GPU | <i>Graphical Processing Units</i> |
| GRU | <i>Gated Recurrent Units</i> |
| HD | Hydrodynamisch |
| IDW | Inverse Distanzgewichtung (<i>Inverse Distance Weighting</i>) |
| KI | Künstliche Intelligenz |
| KIWaSuS | KI-basiertes Warnsystem vor Starkregen und urbanen Sturzfluten (BMBF-Verbundforschungsvorhaben) |
| Log Cosh | Logarithmus des Kosinus hyperbolicus (<i>Logarithm of the hyperbolic cosine</i>) |
| LSTM | <i>Long Short-Term Memory</i> |
| MAE | Mittlerer Absoluter Fehler (<i>Mean Absolute Error</i>) |
| ML | Maschinelles Lernen |
| MLP | Mehrschichtiges Perzeptron (<i>Multilayer-Perceptron</i>) |
| MRSE | <i>Relative Mean Squared Error</i> |
| MSE | Mittlerer Quadratischer Fehler (<i>Mean Squared Error</i>) |
| NARX | <i>Nonlinear Autoregressive with Exogenous Inputs</i> |
| PE | Maximalwertabweichung (<i>Peak Error</i>) |

| | |
|---------|---|
| PTE | Zeitliche Abweichung des Maximalwertes (<i>Peak Time Error</i>) |
| RMSE | <i>Root Mean Squared Error</i> |
| ReLU | Rektifizierte Lineare Einheit (<i>Rectified Linear Unit</i>) |
| RNN | Rekurrente Neuronale Netze (<i>Recurrent Neural Network</i>) |
| Seq2Seq | <i>Sequence to Sequence</i> |
| SVR | <i>Support Vector Regression</i> |
| T | Wiederkehrzeit [a] |
| T-GCN | <i>Temporal Graph Convolutional Network</i> |
| TFT | <i>Temporal Fusion Transformer</i> |
| VE | Volumenfehler (<i>Volume Error</i>) |
| VSN | <i>Variable Selection Networks</i> |
| WMS | <i>Web Map Service</i> |

1 Einleitung

1.1 Veranlassung

Pluviale Überflutungen können theoretisch immer und überall auftreten und stellen insbesondere für hochverdichtete urbane Gebiete, in denen der Großteil des Niederschlages direkt abflusswirksam wird, ein hohes Sicherheitsrisiko dar. Auslöser sind Starkregenereignisse, deren Anzahl und Intensität gemäß dem sechsten Bericht des *Intergovernmental Panel on Climate Change* (IPCC) (Seneviratne et al. 2021) in den vergangenen Jahren zugenommen haben und aufgrund der anhaltenden globalen Erwärmung mit hoher Wahrscheinlichkeit weiter zunehmen werden. In Kombination mit der anhaltenden Urbanisierung und Nachverdichtung von Städten ist künftig auch mit einer Zunahme der Häufigkeit von Überflutungen sowie dem daraus entstehenden Risiko zu rechnen (Rosenzweig et al. 2018). Da pluviale Überflutungen gegenüber fluvialen Überschwemmungen immer und überall auftreten können, ist ein umfassender Schutz aus technischer und ökonomischer Sicht nicht möglich. Umso wichtiger sind daher frühzeitige Warnmeldungen, um proaktives Handeln im Ereignisfall zu ermöglichen. Gleichzeitig stellen Akteure des kommunalen Krisenmanagements als potentielle Anwender von Echtzeitwarnsystemen hohe Anforderungen an die Verlässlichkeit der Warnmeldungen. Dies erfordert räumlich und zeitlich präzise Vorhersagen, um ein effizientes Handeln im Krisenfall zu ermöglichen und Fehlentscheidungen weitestgehend zu vermeiden. Zusammenfassend ergeben sich somit die beiden in Zhao et al. (2021) aufgeführten konträren Anforderungen an die in Echtzeitwarnsystemen verwendeten Vorhersagemodelle:

- Eine zeitlich und räumlich hochauflösende Vorhersage der Überflutungssituation
- Ausreichend Vorlaufzeit zwischen Vorhersagezeitpunkt und Ereigniseintritt

Zur zeitlich und räumlich hochauflösenden Darstellung der Überflutungssituation infolge eines Niederschlagsereignisses stellen hydrodynamische (HD) Berechnungsmodelle ein weitverbreitetes Werkzeug dar. Während der Fokus zu Beginn noch auf der Modellierung des Kanalnetzes lag, hat in den vergangenen Jahren die Betrachtung der Ereignisse an der Oberfläche immer mehr an Bedeutung gewonnen und auch die gekoppelte Modellierung beider Systeme gehört mittlerweile zum Stand der Technik. Die Ausgabe dieser Modelle sind Überstauganglinien an den Schächten des Kanalnetzes und hochauflösende 2D-Wasserstandskarten, aus denen die Überflutungsgefährdung hervorgeht. Verschiedene Untersuchungen zur Validierung von Berechnungsergebnissen beispielsweise anhand von Bildern aus sozialen Netzwerken (Wang et al. 2018; Chaudhary et al. 2019), Aufnahmen von Überwachungskameras (Moy de Vitry et al. 2019) oder gemeldeten Versicherungsschäden (Zischg et al. 2018) zeigen die gute Qualität dieser Modelle. Demgegenüber steht jedoch ein hoher Rechenaufwand (Henonin et al. 2013), wodurch die Rechendauern in Abhängigkeit der Untersuchungsgebietsgröße für ein einzelnes Ereignis mehrere Stunden bis Tage betragen können.

Gegenüber Sturzfluten in natürlichen Einzugsgebieten, die durch advective Niederschlagsereignisse ausgelöst werden und sich mit numerischen Wettermodellen für mehrere Stunden

oder Tage vorhersagen lassen (Lin et al. 2005; Sun et al. 2014), werden pluviale Sturzfluten in urbanen Gebieten meist durch konvektive Niederschlagsereignisse ausgelöst. Diese Ereignistypen lassen sich mit den derzeitigen Vorhersagemodellen lediglich mit Vorlaufzeiten von bis zu zwei Stunden mit einer adäquaten Genauigkeit vorhersagen (Faure et al. 1999; Quirmbach 2003; Sun et al. 2014; Jasper-Tönnies et al. 2017b). Bedingt durch die kurzen Vorlaufzeiten sind hydrodynamische Berechnungsmodelle aufgrund der langen Rechenzeiten nicht ohne Weiteres für die Erzeugung von Überflutungsprognosen im Echtzeitbetrieb geeignet. Das Anwendungsgebiet beschränkt sich daher meist auf die Simulation einzelner Szenarien zur Ausweisung der generellen Starkregengefährdung. Es besteht somit Forschungsbedarf darin, geeignete Ansätze zu finden, um die Dauer für die Berechnung der Überflutungssituation massiv zu reduzieren, bei weiterhin möglichst hoher räumlicher und zeitlicher Auflösung der Berechnungsergebnisse.

1.2 Zielsetzung und Schwerpunkt der Arbeit

Ziel der vorliegenden Arbeit ist die Entwicklung eines zeitlich und räumlich hochauflösenden Vorhersagemodells für pluviale Sturzfluten in urbanen Gebieten. Mit dem Modell soll eine möglichst verzögerungsarme Übersetzung einer Niederschlagsprognose in die daraus resultierende Überflutungssituation erfolgen. So soll der Weg von den bisherigen statischen Starkregengefahrenkarten hin zu einer dynamischen Starkregengefahrenkarte ermöglicht werden. Damit ist gemeint, dass analog zu den Regenradaren in gängigen Wetterportalen (u. a. Kachelmannwetter.com, Wetteronline.com) die Prognose als Sequenz für die kommenden Zeitschritte erfolgt. Für jeden Zeitschritt wird dann eine Starkregengefahrenkarte mit den prognostizierten Wasserständen erzeugt.

Wie im vorherigen Kapitel beschrieben sind die derzeit für die Berechnung von Starkregengefahrenkarten verwendeten physikalisch-basierten hydrodynamischen Berechnungsmodelle sehr rechenintensiv und damit nicht echtzeitfähig. Um diese Problematik zu lösen, werden Maschinelle Lernverfahren herangezogen, um ein sogenanntes Ersatzmodell (*surrogate model*) zu entwickeln. Ersatzmodelle kommen in der Wasserwirtschaft in verschiedenen Bereichen zur Anwendung (Razavi et al. 2012), mit dem Ziel, die Ergebnisse rechenintensiver hochgenauer Simulationsmodelle mit einfacheren und schnelleren Modellen zu approximieren. Im vorliegenden Fall werden also Maschinelle Lernverfahren verwendet, um den Zusammenhang zwischen einer Niederschlagsprognose und den Berechnungsergebnissen eines hydrodynamischen Modells abzubilden. Damit Maschinelle Lernverfahren solche Aufgaben lösen können, ist zunächst ein umfangreicher Trainingsprozess erforderlich, in dem das Verfahren anhand einer Vielzahl an Beispielen von Niederschlagsprognosen und der jeweils dazugehörigen Überflutungssituation den zugrunde liegenden Zusammenhang erlernt. Zwar ist dieser Trainingsprozess sehr rechenintensiv, einmal trainiert sind Maschinelle Lernverfahren dann aber in der Lage, für neue Ereignisse die zugehörige Zielgröße innerhalb weniger Sekunden zu berechnen.

Den Schwerpunkt der Arbeit stellt somit die Entwicklung eines Ersatzmodells zur Überflutungs-

vorhersage auf Basis von Maschinellen Lernverfahren dar. Dazu werden die folgenden wesentlichen Schritte durchgeführt:

- Erzeugung eines Trainingsdatensatzes bestehend aus Niederschlagsereignissen und den für diese Ereignisse mittels hydrodynamischen Modells berechneten Überflutungssituationen.
- Entwicklung eines Ersatzmodells zur Vorhersage von pluvialen Überflutungen mit Hilfe geeigneter Maschineller Lernverfahren.
- Evaluierung der Prognosegüte des entwickelten Vorhersagemodells.

Im Rahmen der Arbeit werden die verwendeten Methoden sowie die gewählte Vorgehensweise zur Durchführung der einzelnen Schritte beschrieben. Damit soll es ermöglicht werden, das beschriebene Vorhersagemodell für andere Gebiete aufzubauen. Gleichzeitig werden aber auch ungeeignete Verfahren und Methodiken herausgestellt, um den Entwicklungsprozess für nachfolgende Untersuchungen zu erleichtern.

1.3 Aufbau der Arbeit

Der Aufbau der vorliegenden Dissertation ist in Abbildung 1-1 dargestellt. Insgesamt ist die Arbeit in acht Kapitel untergliedert. Verbindungen zwischen den einzelnen Kapiteln sind anhand von Pfeilen dargestellt. Nach der Einleitung werden in Kapitel 2 die theoretischen Grundlagen der Überflutungsmodellierung und -vorhersage dargestellt. Dabei erfolgt zunächst eine Abgrenzung des Begriffs der Überflutungsmodellierung im Sinne dieser Arbeit. Anschließend werden allgemeine Grundlagen zu Modellkonzepten erläutert, bevor dann der Stand der Wissenschaft und Technik im Bereich der Überflutungsmodellierung und -vorhersage aufgeführt wird. Abschließend werden anhand des abgeleiteten Forschungsbedarfs Ansatzpunkte identifiziert sowie die in den nachfolgenden Kapiteln verwendete Untersuchungsmethodik beschrieben.

In Kapitel 3 werden zu Beginn das Untersuchungsgebiet sowie die verwendete Datengrundlage beschrieben. Ebenfalls werden sämtliche für die Generierung des Trainingsdatensatzes sowie die Entwicklung der Vorhersagemodelle entwickelten Berechnungsansätze und Softwareanwendungen vorgestellt. Dazu zählen insbesondere das zur Datengenerierung verwendete hydrodynamische Berechnungsmodell sowie die als Maschinelle Lernverfahren verwendeten Künstlich Neuronale Netze.

Die Kapitel 5 und 6 umfassen mit der Entwicklung der ML-basierten Vorhersagemodelle den Kern der Arbeit. Hier erfolgt getrennt für ein Vorhersagemodell für Überflutungsflächen (Kapitel 5) und für Überstau (Kapitel 6) die Beschreibung aller im Rahmen der Modellentwicklung und -analyse durchgeführten Schritte. Begonnen wird dabei jeweils mit der Beschreibung der zugrunde liegenden Lernprobleme sowie den erforderlichen Datenvorverarbeitungsschritten. Anschließend folgen die Entwicklung und der Vergleich verschiedener Modellarchitekturen, bevor jeweils mit der besten Modellarchitektur verschiedene Sensitivitätsanalysen durchgeführt werden. Die Modellgüte der finalen Aufbauten zur Prognose von Überstau und Überflutungsflä-

chen wird abschließend bewertet, bevor eine Zusammenfassung und Bewertung der Ergebnisse folgt.

Während die Modelle in Kapitel 5 und 6 jeweils noch getrennt voneinander entwickelt werden, erfolgt in Kapitel 7 die Zusammenführung und Bewertung der gekoppelten Prognosegüte. Zudem wird auf theoretischer Basis die Integration der beiden Vorhersagemodelle in einen Prognose- und Warnprozess für den späteren operationellen Betrieb beschrieben. Zum Abschluss der Arbeit werden in Kapitel 8 die gesammelten Erkenntnisse sowie Limitierungen der entwickelten Modelle aufgezeigt und daraus der weitere Forschungsbedarf abgeleitet.

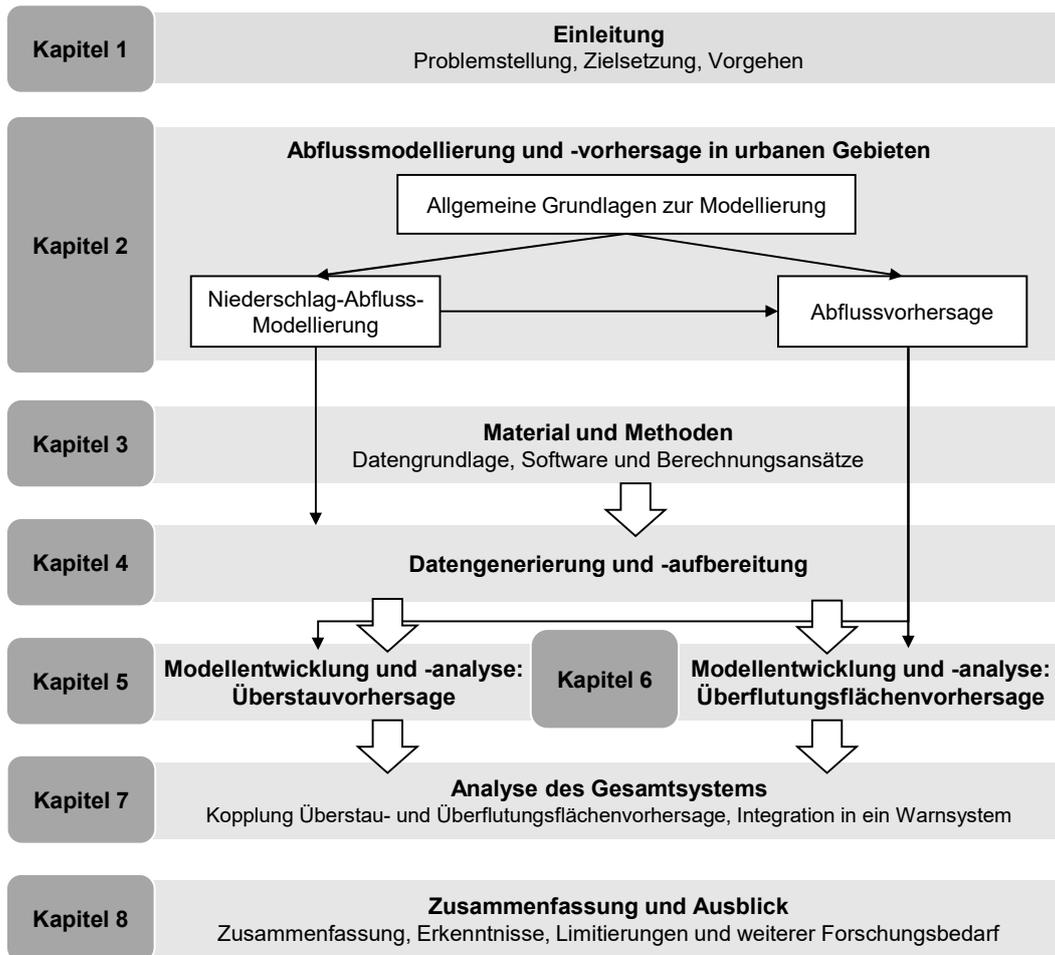


Abbildung 1-1: Wissenschaftlicher Aufbau der Arbeit

2 Überflutungsmodellierung und -vorhersage in urbanen Gebieten

Im folgenden Kapitel werden zunächst die Grundlagen zur Modellierung im Allgemeinen und spezifischer zur Abfluss- und Überflutungsmodellierung beschrieben, bevor der Stand der Wissenschaft und Technik im Bereich der Überflutungsvorhersage aufgeführt wird. Das Kapitel schließt mit dem daraus abgeleiteten Forschungsbedarf sowie der für diese Arbeit gewählten Untersuchungsmethodik.

2.1 Abgrenzung des Begriffs der Überflutungsmodellierung im Sinne dieser Arbeit

Abflüsse und Überflutungen können in unterschiedlicher Form und in unterschiedlichen Systemen auftreten. So kann es in Folge von Sturmfluten an Küsten, Hochwasser an Fließgewässern oder urbanen Sturzfluten zur Ausprägung von Überflutungs- und Überschwemmungsflächen kommen. Die vorliegende Arbeit beschäftigt sich mit der Vorhersage von urbanen Sturzfluten, die nach DWA-M 119 (2016) wie folgt definiert sind:

„kurzfristig auftretende, große oder sehr große Oberflächenabflüsse innerhalb eines Siedlungsgebiets aufgrund lokal auftretender Starkregen“

Gemäß des LUBW-Leitfadens zum Starkregenrisikomanagement in Baden-Württemberg (LUBW 2016), werden Überflutungen in Folge von Starkregenereignissen von den Überschwemmungen ausgehend von Gewässern abgegrenzt. Dennoch wird darauf hingewiesen, dass bei den Überflutungen in Folge von lokal auftretenden Starkregenereignissen auch kleinere Gewässer berücksichtigt werden. Das ist im Rahmen der hier durchgeführten Untersuchungen nicht der Fall. Der Untersuchungsumfang lässt sich daher weiter eingrenzen auf Überflutungen, die gemäß DWA-M 119 (2016) als

„Zustand, bei dem Schmutzwasser und/oder Niederschlagswasser aus einem Entwässerungssystem entweichen oder nicht in dieses eintreten können und entweder auf der Oberfläche verbleiben oder in Gebäude eindringen“

definiert werden. Bei dem berücksichtigten Entwässerungssystem handelt es sich konkret um die Kanalisation, aus der bei einer Überlastung Niederschlagswasser oder Mischwasser in Form von Überstau auf die Geländeoberfläche austritt. Sowohl Überstau als auch die daraus in Kombination mit oberflächlichem Abfluss entstehenden Überflutungsflächen sollen hier konkret vorhergesagt werden. Somit eignet sich neben der Definition von Überflutungen gemäß DWA-M 119 (2016) hier auch der mit der DIN EN 752:2017 eingeführte Begriff der kanalindizierten Überflutung, der wie folgt definiert ist und den im Rahmen dieser Arbeit vorherzusagenden Zustand besonders gut umreißt:

„Zustand, bei dem Abwasser aus einem Entwässerungssystem entweichen oder nicht in dieses eintreten kann und entweder auf der Oberfläche verbleibt oder von der Oberfläche her in Gebäude eindringt.“

Ähnlich wird dieser Zustand im jüngst überarbeiteten und veröffentlichten DWA-A 118 (2024) definiert, dort jedoch um die Betrachtung des entstandenen Schadens ergänzt:

„Zustand, bei dem Wasser aus einem Entwässerungssystem entweicht oder nicht in dieses eintreten kann und dies zu Schäden oder relevanten Funktionsstörungen führt.“

In der nachfolgenden Arbeit werden die Begriffe urbane Sturzflut, Überflutung und kanalindizierte Überflutung synonym verwendet und umschreiben das zuvor abgesteckte Themenfeld.

2.2 Allgemeine Grundlagen zur Modellierung

2.2.1 Modellkonzepte

Modelle werden in nahezu allen Bereichen der Siedlungswasserwirtschaft und insbesondere auch in der Siedlungsentwässerung eingesetzt, um Systeme und die dazugehörigen Prozesse abzubilden. Ziel bei der Anwendung eines Modells ist es, einen Zusammenhang zwischen einer Eingangsgröße und einer gesuchten Ausgabe herzustellen. Die eingesetzten Modelle lassen sich in Abhängigkeit ihrer Struktur und Berechnungsansätze verschiedenen Kategorien zuordnen. Zunächst wird zwischen deterministischen und stochastischen Modellen unterschieden, wobei die in dieser Arbeit zur Abflussmodellierung angewendeten Verfahren deterministischer Natur sind. Das bedeutet, dass eine eindeutige Beziehung zwischen Ursache und Wirkung hergestellt wird und dementsprechend der gleiche Eingabeparametersatz auch immer zu der gleichen Ausgabe führt.

Darüber hinaus lassen sich Modelle in Abhängigkeit des Systemverständnisses in *White-Box*- (oder physikalisch-basierte/prozessorientierte Modelle) und *Black-Box*-Modelle (oder datengetriebene Modelle) unterteilen (Price und Vojinovic 2011; Nützmann und Moser 2016; Becker 1995) (vgl. Abbildung 2-1). Bei Ersteren sind die wesentlichen Komponenten eines Systems sowie die zugrundeliegenden Mechanismen in Form von physikalischen Gesetzmäßigkeiten bekannt und können mit Gleichungssystemen beschrieben werden. Dabei handelt es sich in der Regel um Differentialgleichungen, die durch analytische oder numerische Lösungsverfahren nach der gesuchten Variablen aufgelöst werden können. Das Modell ist somit in der Lage, für eine beliebige Modelleingabe eine Ausgabe zu berechnen. Streng genommen müsste ein *White-Box*-Modell die gesuchte Zielvariable komplett fehlerfrei ausgeben, was in der Praxis jedoch so gut wie nie der Fall ist. Ein wesentlicher Grund hierfür ist, dass es sich bei Modellen immer um eine Vereinfachung (Abstraktion) der Realität handelt (Mehlhorn 1998). Gleiches gilt auch bei *White-Box*-Modellen, bei denen notwendigerweise Annahmen getroffen werden, um physikalischen Prozesse mathematisch abbilden zu können. Ein Beispiel hierfür sind die Saint-Venant-Gleichungen (vgl. Kapitel 3.4.1), bei denen zur vereinfachten Berechnung von Wasserstand und Durchfluss in einem System die Strömungsgleichungen auf eine eindimensionale Strömung reduziert werden.

Bei *Black-Box*-Modellen kann das innere Systemverhalten dagegen nicht nachvollzogen werden. Der Fokus liegt hier vielmehr darauf, mit Hilfe eines mathematischen Konstrukts aus deterministischen oder stochastischen Funktionen eine Beziehung zwischen Eingabe- und Aus-

gabewerten herzustellen. In der Niederschlags-Abfluss-Modellierung könnte das beispielsweise die Abbildung der Beziehung zwischen gefallenem Niederschlag und Abfluss durch eine einzige Funktion sein, wie im Fall der Einheitsimpulsantwort (Maniak 2016). Aber auch statistische Modelle wie autoregressive Modelle (AR-Modelle) oder Verfahren des Maschinellen Lernens (vgl. Kapitel 0) zählen zu diesem Modellkonzept. Voraussetzung für das Erstellen von *Black-Box*-Modellen sind Paare aus Eingabe- und Ausgabevariablen, aus denen der Zusammenhang abgeleitet werden kann. Vor diesem Hintergrund wird dieses Modellkonzept teilweise auch als datengetriebene Modellierung (*Data-Driven Modelling*) bezeichnet (Solomatine und Ostfeld 2008). In Bezug auf die Bestimmung der Einheitsimpulsantwort entspricht das den Niederschlags- und dazugehörigen Abflussdaten.

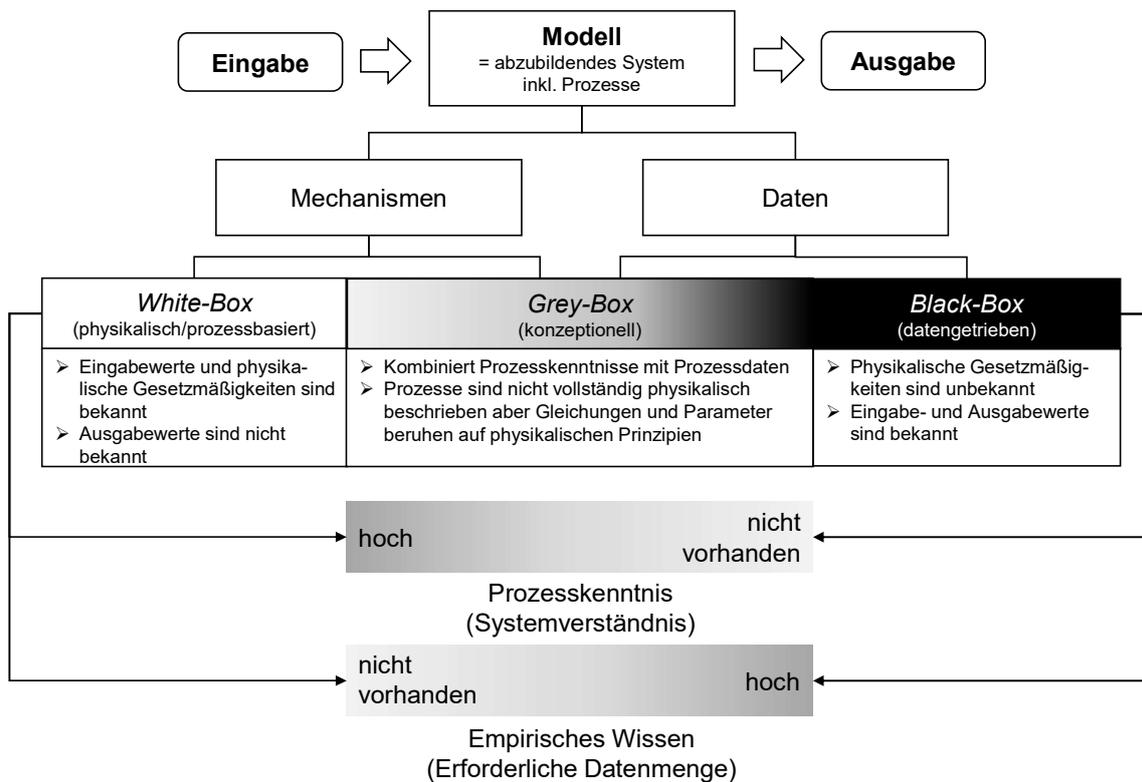


Abbildung 2-1: Konzepte zur Abbildung von Prozessen in Modellen

Neben *Black-Box*- und *White-Box*-Modellen gibt es zudem *Grey-Box*-Modelle als dritte Gruppe. Dabei handelt es sich um eine Mischung von *Black-Box*- und *White-Box*-Modellen. Das Vorwissen über die in einem System stattfindenden Mechanismen wird dabei mit empirischem Wissen aus Messdaten kombiniert. Die für die Abbildung der physikalischen Prozesse verwendeten Ansätze sind vereinfacht, beruhen aber dennoch auf physikalischen Prinzipien (Price und Vojinovic 2011). Die entsprechenden Modellparameter sind physikalisch interpretierbar oder werden wie bei *Black-Box*-Modellen aus Messdaten bestimmt (Mehlhorn 1998). Bei *Grey-Box*-Modellen handelt es sich um das in der Praxis am häufigsten angewendete Modellkonzept (Nacken 2010). Ein Beispiel für dieses Modellkonzept ist das durch den *US Soil Conservation Service* (SCS) entwickelte SCS-Verfahren (vgl. Maniak 2016), zur Berechnung der abflusswirksamen Anteile eines Niederschlagsereignisses.

In Grundmann (2010) wird in Bezug auf die Niederschlag-Abfluss-Modellierung darauf hin-

gewiesen, dass aufgrund diverser Unsicherheiten (u. a. bei den Anfangsbedingungen oder der großen Heterogenität der Gebietseigenschaften) selbst bei physikalisch-basierten Modellen nach wie vor eine Kalibrierung von Modellparametern an Messwerten erforderlich ist, weshalb auch diese Modelle einen gewissen Grad an Konzeptionalität in der Prozessbeschreibung und Parameterbestimmung besitzen. Auf der anderen Seite werden *Black-Box*-Modelle häufig unterstützend zur Abbildung von Teilprozessen verwendet oder auf Basis von Simulationsergebnissen physikalisch-basierter Modelle erstellt. Dementsprechend sind die zugrundeliegenden Prozesse nicht gänzlich unbekannt und die Modelle lassen sich wiederum der Gruppe der *Grey-Box*-Modelle zuordnen. Das Konzept der *Grey-Box*-Modelle kann dementsprechend weniger als eine feste Klasse, sondern vielmehr als ein Bereich zwischen *White-Box* und *Black-Box*-Modellen angesehen werden. In Abhängigkeit der Prozesskenntnis und des empirischen Wissens, das über ein System vorliegt und zur Modellerzeugung verwendet wird, können sich *Grey-Box*-Modelle irgendwo im Bereich der grauen Zone in Abbildung 2-1 befinden.

Modelle lassen sich in der Siedlungsentwässerung zudem hinsichtlich der räumlichen Heterogenität des abgebildeten Systems einteilen (Becker und Serban 1990). Dabei wird zwischen räumlich verteilten Modellen und sogenannten Blockmodellen (*Lumped Models*) unterschieden. Während räumlich verteilte Modelle ein System in viele einzelne Komponenten unterteilen, für die die Prozesse einzeln abgebildet werden, erfolgt bei Blockmodellen keine oder nur eine grobe räumliche Unterteilung. In Bezug auf die Niederschlag-Abfluss-Modellierung kommen beide Detaillierungsgrade zum Einsatz. Im Bereich der Kanalnetzmodellierung, einer Sonderform der Niederschlag-Abfluss-Modellierung, wird beispielsweise das Entwässerungssystem zur Überprüfung von überstaugefährdeten Bereichen haltungsweise abgebildet. Bei Schmutzfrachtmodellen zur Ermittlung von Mischwasserentlastungen oder des Zulaufs zur Kläranlage erfolgt die Modellierung des Entwässerungssystems dagegen deutlich gröber. Je nach Anwendungsfall wird das gesamte Einzugsgebiet als ein „Block“ abgebildet, ohne dass dabei zwischen unterschiedlichen Flächennutzungen und Haltungen unterschieden wird. Während physikalisch-basierte Modelle sich hinsichtlich der räumlichen Heterogenität als räumlich verteilt und datengetriebene Modelle meist als „Block“ einordnen lassen, können konzeptionelle Modelle beide Charakteristiken aufweisen.

Für die Abbildung der Systeme und Prozesse im Bereich der Siedlungsentwässerung werden die beschriebenen Modellkonzepte häufig auch kombiniert. So werden im Bereich der Kanalnetzmodellierung häufig konzeptionelle Modelle (*Grey-Box*) zur Abbildung des Abflussbildungs- und -konzentrationsprozesses verwendet, während der anschließende Abflusstransport im Kanalnetz mit Hilfe physikalisch-basierter Modelle (*White-Box*) erfolgt.

Die beschriebenen Modellkonzepte zeichnen sich durch verschiedene Vor- und Nachteile aus. Die gängigsten werden in Price und Vojinovic (2011) aufgeführt. Während datengetriebene Modelle lediglich versuchen, eine möglichst genaue Verbindung zwischen Eingabe- und Ausgabewerten herzustellen, bilden physikalisch-basierte Modelle den Prozess detailliert ab. Dieser hohe Detaillierungsgrad physikalischer-basierter Modelle geht jedoch meist gleichzeitig mit einer hohen Modellkomplexität und damit einem hohen Rechenaufwand einher. Teilweise lassen sich auch physikalische Prozesse nur unzureichend beschreiben, es liegt aber Prozess-

kenntnis in Form von (Mess-) Daten vor. In diesen Fällen kann ein hybrider Modellierungsansatz hilfreich sein, um vorhandene Prozesskenntnis mit Prozesswissen zu kombinieren. Ziel dieser Modellierungsart ist es, durch die Kombination verschiedener Teilmodelle mit unterschiedlichen Modellkonzepten die Nachteile des jeweiligen Einzelmodells für den vorliegenden Anwendungsfall zu kompensieren (Halfmann und Holzmann 2003). Datengetriebene Modelle können dabei als Ergänzung oder als Ersatz physikalisch-basierter Modelle dienen (Solomatin und Ostfeld 2008). In Abbildung 2-2 werden Möglichkeiten zur Kopplung physikalisch-basierter Modelle mit datengetriebenen Modellen dargestellt, die in Halfmann und Holzmann (2003) genauer beschrieben werden. Bei den dargestellten gekoppelten Modellen, die alleamt zu dem *Grey-Box*-Konzept zählen, werden

- datengetriebene Modelle mit den Ergebnissen von physikalisch-basierten Modellen angelernt (Serielle Modelle),
- datengetriebene Modelle zur Schätzung veränderlicher Parameter des physikalisch-basierten Modells eingesetzt (Parameterschätzung) oder
- die Ergebnisse beider Modelle kombiniert (Parallele Modelle).

Zu den seriellen Modellen gehört auch der in dieser Arbeit verwendete Ansatz (vgl. Kapitel 2.5), bei dem physikalisch-basierte Kanalnetz- und Oberflächenabflussmodelle verwendet werden, um Trainingsdaten für das Anlernen Maschinellem Lernverfahren zu generieren.

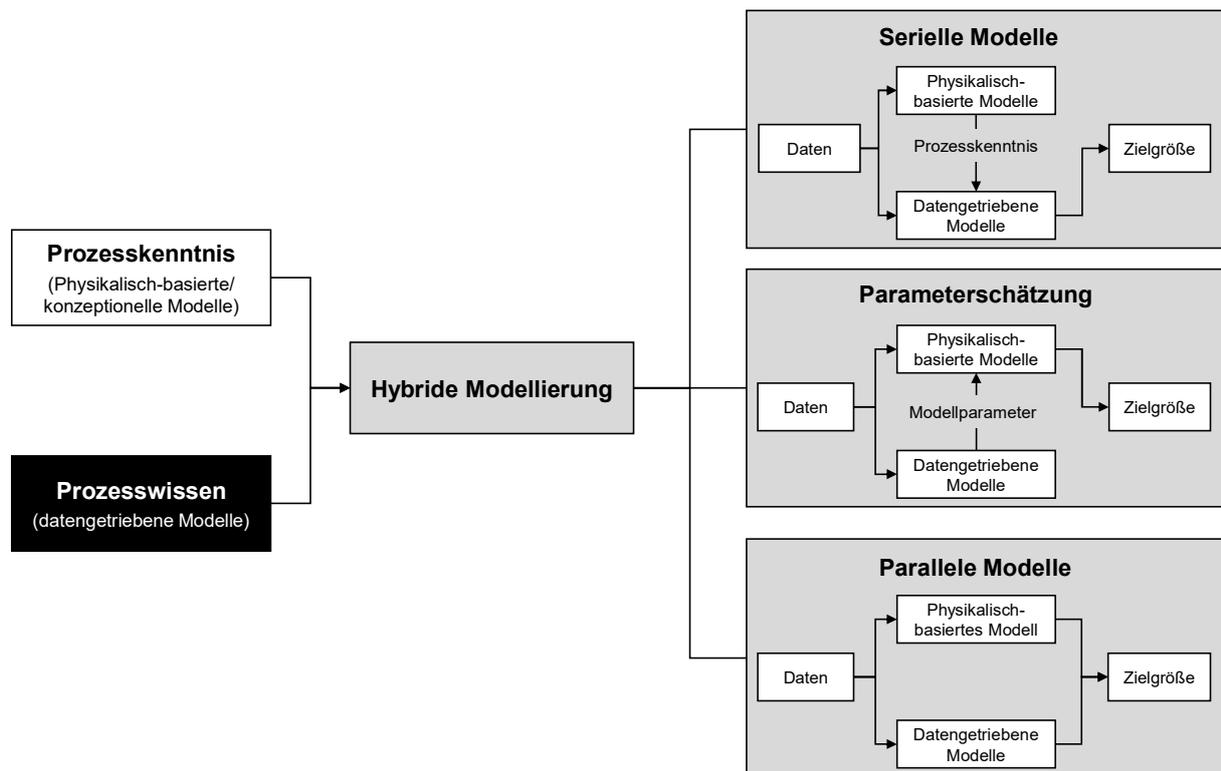


Abbildung 2-2: Ansätze zur Kombination von *White-Box*- mit *Black-Box*-Ansätzen zu hybriden Modellen

2.2.2 Maschinelles Lernen

Maschinelles Lernen (*Machine Learning*, ML) stellt ein großes Teilgebiet der Künstlichen Intelligenz dar, dem verschiedene Verfahren zur Datenverarbeitung angehören, auf die in dieser Arbeit zurückgegriffen wird. Die Verfahren des ML zeichnen sich dadurch aus, automatisiert Zusammenhänge in Daten zu erlernen. Damit lassen sie sich den datengetriebenen Modellen zuordnen und weisen größtenteils eine *Black-Box*-Charakteristik auf. Gegenüber den in Kapitel 2.3 beschriebenen prozessbasierten oder konzeptionellen Berechnungsansätzen sind die physikalischen Gesetzmäßigkeiten und Parameter bei Modellen, die auf ML-Verfahren beruhen, unbekannt. Vielmehr lernt das Modell auf Grundlage empirischer Beobachtungen (z.B. Messdaten), den Zusammenhang zwischen Eingabedaten und Zielvariablen abzubilden.

Gegenüber den prozessbasierten Modellierungsansätzen ergibt sich so eine völlig andere Herangehensweise in der Modellentwicklung. Bei der prozessorientierten Modellierung werden dem Modell vom Anwender oder Entwickler Daten sowie Regeln, wie diese zu verarbeiten sind, vorgegeben. Dadurch, dass das Modell die entsprechenden Verarbeitungsregeln kennt, ist es in der Lage, die gewünschten Zielgrößen zu berechnen (vgl. Abbildung 2-3 links). Beim Maschinellen Lernen werden hingegen zunächst im Rahmen eines Trainingsprozesses die Zusammenhänge (Regeln) zwischen Daten und dazugehörigen Zielgrößen erlernt. Im Anschluss an den Trainingsprozess ist das Modell dann in der Lage, die erlernten Zusammenhänge auf andere, während des Trainingsprozesses nicht berücksichtigte Daten zu verallgemeinern und so die gesuchten Zielgrößen zu bestimmen (vgl. Abbildung 2-3 rechts).

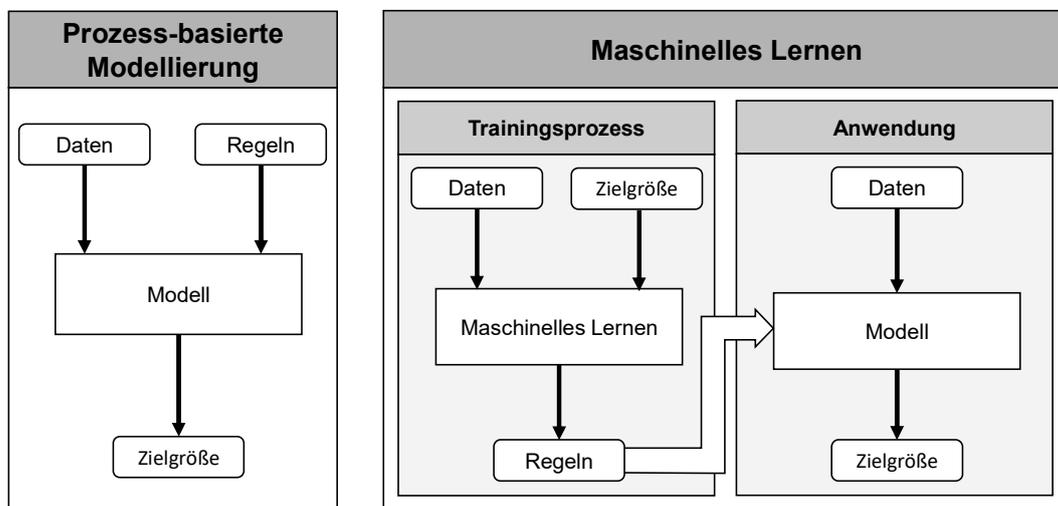


Abbildung 2-3: Vergleich prozessbasierter Modellierung und ML-basierter Modellierung (verändert nach Abbildung 1, Burrichter et al., 2021)

Bei der Art und Weise wie Maschinelle Lernverfahren Zusammenhänge und Muster in Daten erlernen, wird zwischen unterschiedlichen Lernstilen unterschieden, denen sich wiederum unterschiedliche Lernaufgaben zuordnen lassen. Eine Übersicht über die gängigsten Lernstile und -aufgaben sowie eine Auswahl zugehöriger Verfahren kann Abbildung 2-4 entnommen werden.

Bei den Lernstilen wird im Wesentlichen zwischen überwachtem, unüberwachtem und bestärkendem Lernen unterschieden. Am weitesten verbreitet sind dabei das überwachte Lernen

(*supervised learning*), das auch in der vorliegenden Arbeit zur Anwendung kommt und das unüberwachte Lernen (*unsupervised learning*). Beide Lernstile unterscheiden sich vorrangig dadurch, dass beim überwachten Lernen dem ML-Algorithmus gekennzeichnete Trainingsdaten vorgegeben werden, während beim unüberwachten Lernen keine Kennzeichnung der Eingabedaten erfolgt. Unter der Kennzeichnung von Daten ist die Zuordnung der gesuchten Zielvariablen als Feedback zu verstehen. Das ML-Verfahren wird beim überwachten Lernen darauf trainiert, den Zusammenhang zwischen Eingabe- und Zielvariablen zu erlernen. Diese Art des Lernens eignet sich insbesondere, um ein unbekanntes Ergebnis vorherzusagen. Das kann einerseits die Vorhersage einer festen Klasse (Klassifikation) oder eines stetigen Wertes (Regression) sein. Dabei handelt es sich zugleich um die gängigsten Lernaufgaben, denen sich die Verfahren des überwachten Lernens zuordnen lassen.

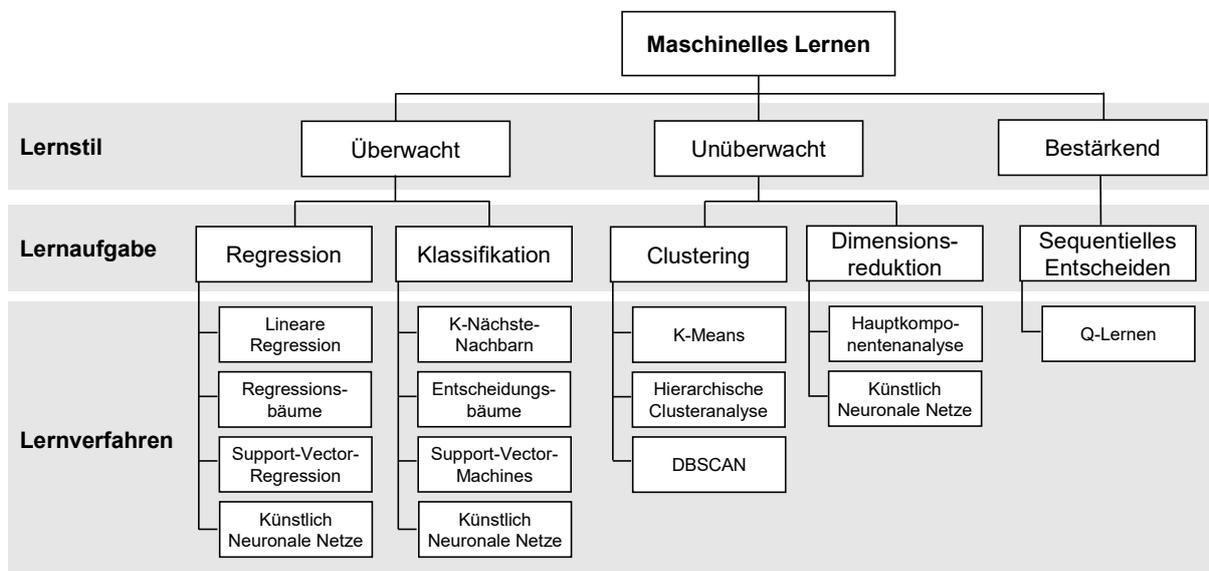


Abbildung 2-4: Unterteilung von gängigen ML-Verfahren in Abhängigkeit von Lernstil und -aufgabe

Das unüberwachte Lernen zielt gegenüber dem überwachten Lernen nicht auf die Vorhersage einer Zielvariablen ab, sondern vielmehr darauf, Muster in Datensätzen zu identifizieren. Dem ML-Verfahren werden dabei meist große unstrukturierte Datenmengen vorgegeben, um Strukturen und Unterschiede in den Daten zu erkennen (Fraunhofer-Gesellschaft 2018). Die Verfahren des unüberwachten Lernens lassen sich insbesondere den Lernaufgaben Clustering und Dimensionsreduktion zuordnen. Das Clustering hat zum Ziel, den Datensatz in einzelne Cluster (Gruppen) mit ähnlichen Eigenschaften aufzuteilen. Manchmal wird diese Aufgabe auch als „unüberwachte Klassifizierung“ bezeichnet (Raschka und Mirjalili 2018), wobei die Klassenbezeichnung der einzelnen Gruppen hier gegenüber der eigentlichen Klassifikation unbekannt ist. Die Dimensionsreduktion hat hingegen zum Ziel, Datensätze zu vereinfachen ohne dabei allzu viele Informationen zu verlieren (Géron 2019). Hierzu gibt es verschiedene Algorithmen, die beispielsweise darauf abzielen, korrelierende Merkmale in einem Datensatz zu einem Merkmal zusammenzuführen, um so die Dimensionen des Datensatzes zu reduzieren. Aufgrund dieser Eigenschaften werden Verfahren zur Dimensionsreduktion in der Regel anderen ML-Algorithmen (überwacht oder unüberwacht) vorgeschaltet.

Der dritte in Abbildung 2-4 aufgeführte Lernstil ist das bestärkende Lernen (*reinforcement learning*). Dabei handelt es sich um einen grundlegend anderen Lernstil, bei dem ein häufig auch als „Agent“ bezeichnetes System durch Interaktion mit seiner Umgebung lernt (Raschka und Mirjalili 2018). Dabei führt der Agent eine Aktion durch und erhält in Abhängigkeit von dieser ein Feedback von seiner Umgebung. In Abhängigkeit des Feedbacks erlernt der Agent dann, seine Interaktionen mit der Umgebung für zukünftige Aktionen zu verbessern. Dieser Lernstil eignet sich dazu, virtuelle (Bots) oder physische Maschinen (Roboter) darauf zu trainieren, bestimmte Aufgaben in ihrer Umgebung durchzuführen. Im Rahmen der hier durchgeführten Arbeiten kommt das bestärkende Lernen nicht zur Anwendung und wird lediglich der Vollständigkeit halber aufgeführt.

Über die letzten Jahre hinweg wurde eine Vielzahl an verschiedenen Verfahren des ML entwickelt. Eines der bekanntesten Verfahren des Maschinellen Lernens sind die Künstlich Neuronale Netze, deren Beschreibung in Kapitel 3.3 folgt und die in dieser Arbeit fast ausschließlich verwendet wurden. Für eine detaillierte Beschreibung weiterer Lernstile, -aufgaben und -verfahren wird an dieser Stelle auf die entsprechende Fachliteratur verwiesen (Raschka und Mirjalili 2018; Géron 2019; Richter 2019). Zum besseren Verständnis der Abbildung 2-4 sei an dieser Stelle vorab darauf hingewiesen, dass Künstlich Neuronale Netze sich in Abhängigkeit ihrer Struktur weiter differenzieren lassen und dadurch zur Lösung unterschiedlicher Lernaufgaben geeignet sind.

2.3 Abfluss- und Überflutungsmodellierung in der Siedlungsentwässerung

Abflussmodelle werden in der Siedlungsentwässerung zur quantitativen und qualitativen Berechnung des Abflusses eingesetzt. Da der Fokus der vorliegenden Arbeit auf der quantitativen Abflussmodellierung und -vorhersage liegt, werden hier auch nur die für diesen Anwendungsfall eingesetzten Niederschlag-Abfluss-Modelle beschrieben. Dafür werden zunächst die Grundlagen der Niederschlag-Abfluss-Modellierung zur Simulation des Abflussgeschehens in urbanen Gebieten erläutert. Neben Hintergrund und Anwendungsbereichen werden zudem die wichtigsten Komponenten und Prozesse dargestellt, die Einfluss auf den Niederschlag-Abfluss-Prozess in der Siedlungsentwässerung haben. Eine Beschreibung der qualitativen Abflussberechnung mit Schmutzfrachtmodellen findet sich u. a. bei Butler und Davies (2004) oder Price und Vojinovic (2011).

2.3.1 Hintergrund und allgemeine Grundlagen

Historische Hauptziele von Entwässerungssystemen sind die Aufrechterhaltung hygienischer Verhältnisse in Siedlungsgebieten sowie der Schutz von Menschen und Gütern vor Überflutungen. Auch wenn mit der Einführung des DWA-A 100 (2006c) der Schutz der Gewässer als gleichrangiges Ziel hinzugekommen ist, stellt die sichere und überflutungsfreie Entwässerung von Siedlungsgebieten weiterhin eine maßgebende Zielvorgabe dar. Eine wesentliche Voraussetzung dafür, dass diese Zielvorgabe erreicht werden kann, stellt die richtige Bemessung von Bauwerken der Siedlungsentwässerung dar. Ein Rahmen für die ordnungsgemäße hydraulische Bemessung von Entwässerungssystemen wurde für Deutschland bereits 1956 mit dem

Arbeitsblatt ATV-A 118 veröffentlicht, das in der aktuellen Fassung als DWA-A 118 (2024) immer noch gültig ist und gerade neu überarbeitet wurde. Während zunächst noch einfache Berechnungsverfahren wie das Zeitbeiwertverfahren zur Bemessung von Kanälen eingesetzt wurden, haben im Zuge der Computerisierung Niederschlag-Abfluss-Modelle (vgl. Kapitel 0) immer mehr an Bedeutung gewonnen. Mit der Einführung des ATV-M 165 wurde bereits 1994 ein Regelwerk veröffentlicht, das die Grundlagen und Anwendungsmöglichkeiten von Niederschlag-Abfluss-Modellen beschreibt. Gemäß der ersten überarbeiteten Fassung des Merkblatts ATV-DVWK-M 165 (2004a) wurden Niederschlag-Abfluss-Modelle als Stand der Technik angesehen und werden auch im DWA-A 118 (2006a) als Berechnungsmethodik aufgeführt. Mit der Veröffentlichung des DWA-M 165-1 (2021), einer umfangreich überarbeiteten Fassung des ATV-DVWK-M 165, sind Niederschlags-Abfluss-Modelle mittlerweile gemeinsam mit Schutzfrachtmodellen vollends in den Fokus gerückt und stellen den Kern des Merkblattes dar. Auch die neuste Fassung des DWA-A 118 (2024) unterstreicht diese Entwicklung. Darüber hinaus werden dort vereinfachte Berechnungsverfahren wie das Zeitbeiwert- und Fließzeitverfahren für Überstau und Überflutungsberechnungen als ungeeignet herausgestellt.

Mit der Einführung der europäischen Norm DIN EN 752 im Jahr 1996 wird eine Überflutungsprüfung für kommunale Entwässerungssysteme explizit gefordert. Hierzu wurden gleichzeitig Bemessungsregen- und Überflutungshäufigkeiten empfohlen, die als Bewertungskriterien für die Neuplanung und Sanierung von Kanälen berücksichtigt werden sollen, soweit keine abweichenden länderspezifischen Vorgaben gelten. Beide Bewertungskriterien wurden auch in das DWA-A 118 (2006a) übernommen. Zudem wurden im DWA-A 118 Überstauhäufigkeiten für die rechnerische Nachweisführung als weiteres Kriterium ergänzt, das auch noch in der überarbeiteten Fassung DWA-A 118 (2024) besteht. Für die Ermittlung von Überstauhäufigkeiten bzw. die Nachweisführung einer ausreichenden hydraulischen Leistungsfähigkeit von Entwässerungssystemen haben sich Niederschlag-Abfluss-Modelle in der Praxis etabliert.

Zur Sicherstellung der nach DIN EN 752:2017 geforderten Überflutungssicherheit ist für Bereiche mit rechnerischem Überstau eine zusätzliche Überflutungsprüfung durchzuführen. Hier wird im DWA-A 118 (2006a) aufgrund der Schwierigkeiten, den oberflächigen Überflutungsvorgang modelltechnisch abzubilden, noch eine Bewertung der Überflutungsgefährdung vor Ort gefordert. Mit der gestiegenen Rechenleistung und Datenverfügbarkeit in den letzten Jahren sowie der Entwicklung geeigneter Berechnungstools (u. a. MIKE+, HYSTEM-EXTRAN 2D oder GeoCPM) stellt eine detaillierte computergestützte Überflutungsmodellierung heute jedoch kaum noch ein Problem dar. Mit der Veröffentlichung des DWA-M 119 (2016) wird dieser Entwicklung erstmals Rechnung getragen ein Merkblatt mit detaillierteren Empfehlungen zur kommunalen Überflutungsvorsorge veröffentlicht. Hier werden auch verschiedene Methoden zur computergestützten Ermittlung von Überflutungsrisiken beschrieben, die eine wesentlich detailliertere Überflutungsprüfung ermöglichen, als dies im Rahmen von reinen Vor-Ort-Begehungen möglich ist. Auch hier stellen Niederschlag-Abfluss-Modelle eine weitverbreitete Methodik dar und kommen in verschiedenen Konstellationen und unterschiedlichen Detaillierungsgraden zum Einsatz (vgl. Kapitel 0). In der letzten Aktualisierung des DWA-M 165-1 (2021) und des DWA-A 118 (2024) werden dementsprechend auch dort Erläuterungen zum Einsatz von Niederschlag-Abfluss-Modellen in der Überflutungsberechnung aufgeführt.

2.3.2 Niederschlag-Abfluss-Modellierung

Zur vollständigen Beschreibung des Niederschlag-Abfluss-Prozesses in der Siedlungsentwässerung sind verschiedene Prozesse abzubilden. Zunächst wird dabei zwischen

- Oberflächenabfluss und
- Abflusstransport im Kanalnetz

unterschieden. Der Oberflächenabfluss beschreibt dabei den Prozess vom Aufprall des Niederschlags auf die Gebietsoberfläche bis zum Eindringen in die Kanalisation und lässt sich wiederum in die Teilprozesse

- Abflussbildung und
- Abflusskonzentration

unterteilen. Bei den in der Siedlungsentwässerung gemäß dem Stand der Technik eingesetzten Niederschlag-Abfluss-Modellen handelt es sich in der Regel um eine Kombination aus mehreren konzeptionellen Modellen oder physikalisch-basierten und konzeptionellen Modellen zur Abbildung der einzelnen Teilprozesse. Eine Übersicht über die nach DWA-M 165-1 (2021) im Rahmen der Niederschlag-Abfluss-Berechnung zu berücksichtigenden Komponenten und Prozesse sowie gängige Berechnungsansätze (nicht vollständig) sind in Abbildung 2-5 dargestellt. Dabei handelt es sich auch um die Verfahren, auf denen die gängigen nationalen (u.a. ++Systems und Hystem-Extran) sowie internationalen (u.a. SWMM und MIKE+) Softwareanwendungen zur Niederschlags-Abfluss-Modellierung basieren. Im Folgenden werden die einzelnen Teilprozesse kurz erläutert, für eine detaillierte Beschreibung der einzelnen Berechnungsansätze wird an dieser Stelle auf die entsprechende Fachliteratur verwiesen (Butler und Davies 2004; Price und Vojinovic 2011; Aigner und Bollrich 2015; Maniak 2016).

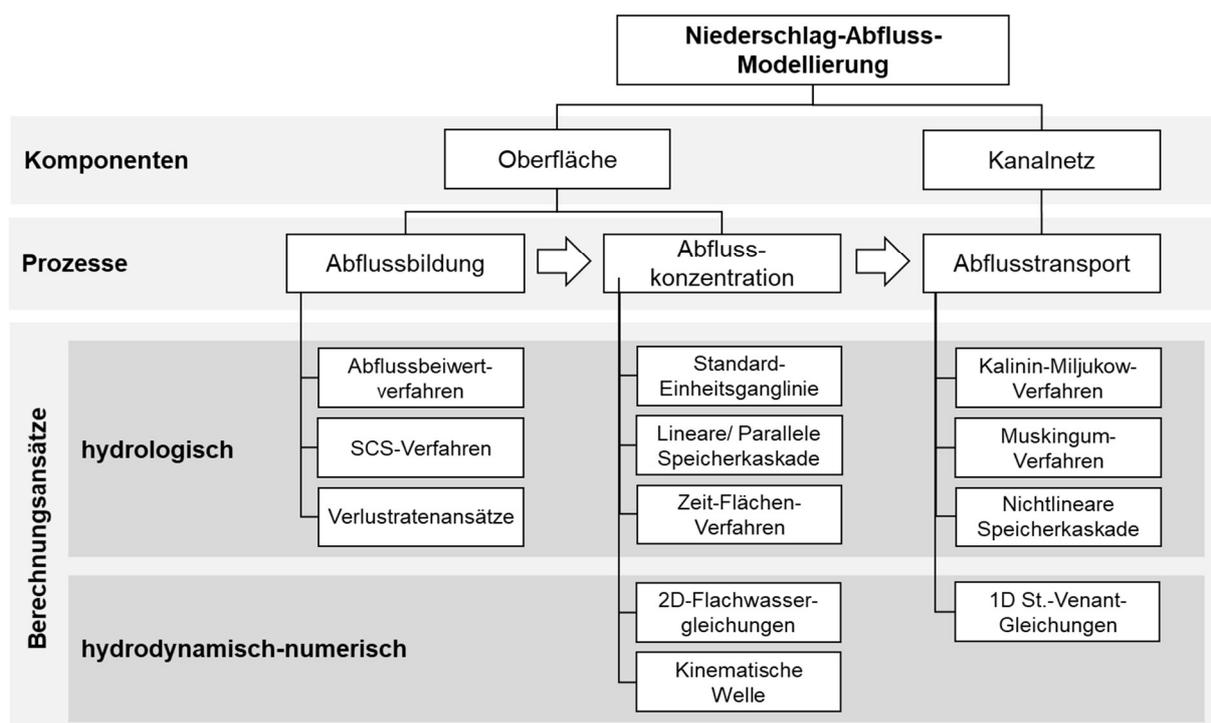


Abbildung 2-5: Übersicht der bei der Niederschlag-Abfluss-Modellierung in der Siedlungsentwässerung berücksichtigten Komponenten, Prozesse und Berechnungsansätze

Abflussbildung

Der Prozess der Abflussbildung beschreibt die Umwandlung des gefallenen Niederschlags in den abflusswirksamen Anteil des Niederschlags und legt somit das Volumen des Oberflächenabflusses fest. Dazu können verschiedene Verlustansätze berücksichtigt werden, die dazu führen, dass nur ein Teil des Niederschlags, der auf die Gebietsoberfläche fällt, zum Abfluss kommt. Bei der Art von Verlustansätzen wird in Anfangsverluste und Dauerverluste unterschieden. Die Anfangsverluste werden zu Beginn eines Ereignisses vollständig von der gefallenen Niederschlagsmenge abgezogen und führen zu einer Verzögerung der Abflusskurve. Hierzu zählen insbesondere Benetzungs- und Muldenverluste. Dauerverluste werden hingegen über die gesamte Ereignisdauer als konstante Rate oder als über die Zeit veränderliche Verlustgrößen berücksichtigt. Die Größe des abflusswirksamen Niederschlags hängt nach DWA-A 118 (2006a) vor allem von folgenden Einflussfaktoren ab:

- Anteil der befestigten Flächen,
- Art der Flächenbefestigung,
- Geländeneigung,
- Regenstärke und -dauer,
- Bodenart und Bewuchs bei durchlässigen Flächen.

Die aufgeführten Einflussfaktoren stellen auch zugleich die Randbedingungen der gängigen Berechnungsmodelle dar.

Bei der Abflussbildung wird nach DWA-M 165-1 (2021) aufgrund des unterschiedlichen Abflussverhaltens in Abhängigkeit des Befestigungsgrades zwischen durchlässigen und undurchlässigen Flächen unterschieden. Für beide Flächen werden dann in Berechnungsmodellen spezifische Verlustansätze gewählt. So ist bei durchlässigen Flächen beispielsweise zusätzlich die Versickerung in den Boden zu berücksichtigen, während diese bei undurchlässig befestigten Flächen entfällt. Teildurchlässige Flächen, wie beispielsweise Schotterwege oder Pflasterbeläge, sind durch geeignete Verfahren zu berücksichtigen. Dies kann unter anderem wie in DWA-M 165-1 (2021) beschrieben durch die Zuordnung zu den undurchlässig befestigten oder den nicht befestigten Flächen je nach Grad der Durchlässigkeit erfolgen.

Abflusskonzentration

Die Abflusskonzentration beschreibt den Prozess der Umwandlung des abflusswirksamen Niederschlags an der Oberfläche des Einzugsgebietes in Zuflussganglinien an den Eintrittspunkten in das Entwässerungssystem und legt somit die zeitliche Verteilung des Oberflächenabflusses fest. Dabei kommt es aufgrund der Fließvorgänge an der Oberfläche sowohl zu einer zeitlichen Verzögerung (Translation) als auch einer Dämpfung (Retention) der Ganglinie des abflusswirksamen Niederschlags (DWA 2021).

Wie bei der Abflussbildung wird auch bei der Abflusskonzentration bei den gängigen Modellen zwischen durchlässig und undurchlässig befestigten Flächen unterschieden. Hintergrund ist die bei durchlässigen Flächen deutlich größere Abflussverzögerung. Der Detaillierungsgrad der betrachteten Flächen kann dabei von einer haltungsweisen Betrachtung der angeschlossenen Flächen bis hin zur Abbildung ganzer Einzugsgebiete eines Kanalnetzes als eine Fläche

stark variieren (Rauch et al. 2010).

Zur Beschreibung des Abflusskonzentrationsprozesses wird im Rahmen der Kanalnetzmodellierung in der Regel auf die exakte physikalische Beschreibung unter Berücksichtigung von Translations- und Retentionseffekten verzichtet (DWA 2021). Vielmehr kommen vereinfachte hydrologische Berechnungsansätze zum Einsatz, die mit Hilfe einer Übertragungsfunktion den abflusswirksamen Niederschlag in eine Abflussganglinie überführen (HSGSim 2008). Hierzu zählen unter anderem Einheitsganglinienverfahren oder verschiedene Speichermodelle wie Einzellinearspeicher oder lineare Speicherkaskaden.

Neben den gängigen hydrologischen Ansätzen können auch hydrodynamische Ansätze zur Abbildung des Abflusskonzentrationsprozesses gewählt werden. Dabei wird die Annahme getroffen, dass der Oberflächenabfluss dem Abfluss in einem Rechteckgerinne mit großer Breite und geringer Fließtiefe ähnelt. Zu diesen Ansätzen zählt die kinematische Welle, die sich durch Vereinfachung der Saint-Venant-Gleichungen ergibt, die wiederum auch zur Berechnung des Abflusstransportprozesses im Kanal verwendet werden. Dieser Ansatz ist in der Urbanhydrologie nicht ganz so stark verbreitet (Rauch et al. 2010), ist aber unter anderem in der Modellierungssoftware Mike+ (DHI 2022a) zur Auswahl hinterlegt.

Einen weiteren hydrodynamischen Ansatz stellen zweidimensionale hydrodynamische Oberflächenabflussmodelle zur Abbildung des Fließverhaltens auf der Oberfläche dar, die insbesondere zur Analyse der Überflutungsgefährdung eingesetzt werden (HSGSim 2008). Diese Modelle basieren auf den Flachwassergleichungen zur Abflussberechnung, die wiederum die Reduzierung der Navier-Stokes-Gleichungen und der Kontinuitätsgleichung auf ein ebenes Strömungsproblem mit freier Oberfläche darstellen (Aigner und Bollrich 2015). Die Grundlage von 2D-Oberflächenabflussmodellen stellen Rechennetze dar, die mit Höheninformationen aus 3D-Punktmessungen (i. d. R. flugzeuggestütztes Laserscanning (*Airborne Laserscanning*, ALS)) versehen werden und die Geländeoberfläche abbilden. Auf diesen dreiecks- oder rasterbasierten Rechennetzen erfolgt dann die Simulation des Abflussverhaltens. 2D-Modelle zur Oberflächenabflusssimulation können über geeignete Ansätze mit Kanalnetzmodellen gekoppelt werden und so zwischen zwei aufeinanderfolgenden Berechnungsschritten Wasservolumina zwischen Oberfläche und Kanalnetz austauschen. Möglichkeiten zur Kombination und Kopplung von hydrodynamischen 1D-Kanalnetzmodellen und 2D-Oberflächenmodellen werden ausführlich in HSB (2017) beschrieben.

Abflusstransport im Kanalnetz

Dem Oberflächenabfluss schließt sich beim Eindringen in die Kanalisation der Abflusstransportprozess im Kanalnetz an. Die über das Oberflächenmodell berechneten Zuflussganglinien an den Eintrittspunkten (u.a. Schächte, Straßeneinläufe und Hausanschlüsse), Fremdwasseranteile sowie beim Mischsystem der zusätzliche Schmutzwasserabfluss stellen die Belastung dar. Im Gegensatz zur Oberfläche handelt es sich beim Kanalnetz aufgrund der vorgegebenen Geometrien von Haltungen und Bauwerken um ein eindeutig definiertes System. Dadurch lassen sich die Abflussvorgänge detailliert beschreiben.

Nach DWA-M 165-1 (2021) lassen sich beim Transportprozess einer Abflusswelle durch eine Haltung zwei Effekte beobachten, die es modelltechnisch abzubilden gilt:

- Translation (Wellenverschiebung)
- Retention (Wellendämpfung).

Für die Berechnung des Transportprozesses kommen sowohl hydrologische als auch hydrodynamische Berechnungsmethoden zum Einsatz. Während hydrologische Berechnungsansätze zur Ermittlung des Abflusses lediglich die Massenbilanz (Kontinuitätsgleichung) berücksichtigen, können hydrodynamische Modelle darüber hinaus auch die Energiebilanz abbilden. Dadurch sind diese Modelle in der Lage, hydraulisch überlastete Abschnitte und dadurch entstandene Rückstaueffekte sowie überstaute Schächte zu identifizieren. Zur Durchführung von Überstaunachweisen nach DIN EN 752 (DIN EN 752:2017) oder DWA-A 118 (2024) sind daher nur hydrodynamische Modelle geeignet.

Hydrodynamische Ansätze basieren auf den physikalischen Gesetzmäßigkeiten des Fließvorgangs, beschrieben durch die Saint-Venant-Gleichungen für eindimensionale Strömung. Diese bestehen aus der Kontinuitäts- und der Bewegungsgleichung (vgl. Kapitel 3.4.1). Dabei handelt es sich um partielle, nicht-lineare Differentialgleichungen, die analytisch nicht lösbar sind (ATV-DWWK 2004a). Daher wird für deren Lösung auf numerische Verfahren zurückgegriffen, die das kontinuierliche Abflussgeschehen zwischen den Haltungen in Zeitintervallen diskretisiert betrachten und iterativ lösen. Um den Rechenaufwand bei der Strömungsberechnung zu reduzieren, können einzelne Terme der Bewegungsgleichung vernachlässigt werden. Dadurch ergeben sich die Diffusionswelle bei Vernachlässigung der lokalen und konvektiven Beschleunigung oder die kinematische Wellenapproximation bei zusätzlicher Vernachlässigung der Druckdifferenz. Hier ist jedoch zu berücksichtigen, dass aufgrund der fehlenden Druckdifferenz mit dem kinematischen Wellenansatz keine Rückstau- oder Dämpfungseffekte abgebildet werden können. Für eine detaillierte Beschreibung der Saint-Venant-Gleichungen sowie der möglichen Vereinfachungen sei an dieser Stelle auf das DWA-A 110 (2006b) verwiesen.

Im Gegensatz zu hydrodynamischen Modellen wird zur Abbildung von Retentionseffekten bei hydrologischen Modellen anstelle der physikalisch begründeten Energiegleichung auf konzeptionelle Verfahren zurückgegriffen (DWA 2021). Dabei wird auf Grundlage empirischer Beobachtungen eine Beziehung zwischen Zu- und Abflussganglinien hergestellt. Durch die fehlende Berücksichtigung der Energiegleichung geht jedoch die Information über die zeitliche Veränderung der Wasserstand-Abflussbeziehung verloren, sodass unter anderem Rückstaueffekte nicht abgebildet werden können. Der Vorteil dieser Vereinfachung ist, dass sich das Gleichungssystem direkt lösen lässt (DWA 2021) und somit gegenüber hydrodynamischen Berechnungsansätzen deutliche kürzere Rechenzeiten erzielt werden. Neben dem vereinfachten Berechnungsansatz wird bei hydrologischen Modellen häufig auch auf eine detaillierte räumliche Differenzierung der Modellkomponenten verzichtet. Anwendungsgebiete sind daher insbesondere die Bemessung und Nachweisführung von zentralen Bauwerken der Siedlungs-entwässerung wie beispielsweise Mischwasserentlastungen sowie Schmutzfrachtnachweise.

2.3.3 Analyse der Überflutungsgefährdung

Starkregenereignisse und die dadurch resultierenden extremen Wassermassen führen deutschlandweit immer wieder zu einer Überlastung von städtischen Entwässerungsstrukturen und in der Folge zu Überflutungen. Durch die anhaltende Nachverdichtung urbaner Gebiete in Kombination mit der klimawandelbedingten Zunahme von Starkregenereignissen hat die Relevanz von Überflutungsereignissen stark zugenommen. Aus diesem Grund wurden verschiedene Leitfäden (u. a. DWA 2013; LUBW 2016; MULNV 2018) sowie mit dem DWA-M 119 (2016) ein Regelwerk zum kommunalen Starkregenrisikomanagement veröffentlicht. Die Grundlage stellt dabei jeweils eine Analyse der Überflutungsgefährdung dar. Während der gemäß DWA-A 118 (2024) geforderte Überstaunachweis bereits einen Rückschluss auf die hydraulische Leistungsfähigkeit des Kanalnetzes liefert und überlastete Punkte im Kanalnetz ausweist, ist eine Aussage zur Überflutungsgefährdung anhand der Ergebnisse nur eingeschränkt möglich. Hierfür ist neben der Betrachtung der Abflussvorgänge im Kanal auch eine detaillierte Betrachtung der Abflussvorgänge auf der Oberfläche erforderlich (DWA 2013).

Die verfügbaren Methoden zur Analyse der Überflutungsgefährdung sind vielfältig und unterscheiden sich hinsichtlich des Detaillierungsgrades und der Komplexität. Nach DWA-M 119 (2016) wird zunächst in belastungsunabhängige und belastungsabhängige Methoden unterschieden. Zu den belastungsunabhängigen Methoden zählen topografische Analysen der Geländeoberfläche, die eine qualitative Gefährdungsbeurteilung u. a. anhand von Fließwegen ermöglichen. Belastungsabhängige Methoden beziehen hingegen eine Niederschlagsbelastung mit in die Gefährdungsbeurteilung ein. Dadurch wird auch eine ereignisspezifische Gefährdungsbeurteilung ermöglicht.

Darüber hinaus erfolgt im DWA-M 119 (2016) eine beispielhafte Einteilung der gängigen methodischen Ansätze zur Analyse der Überflutungsgefährdung, die in Abbildung 2-6 dargestellt ist. Der Detaillierungsgrad, aber gleichzeitig auch die Komplexität und somit der Bearbeitungsaufwand der aufgeführten Methoden nehmen von oben nach unten zu. Das detaillierteste Verfahren stellt dabei die gekoppelte 1D/2D-Abflusssimulation dar, die auf den in Kapitel 0 beschriebenen hydrodynamischen Berechnungsmodellen beruht. Dabei erfolgt, ausgelöst durch einen Niederschlag als Systembelastung, eine 1D-Simulation im Kanalnetz und eine 2D-Simulation des Abflussverhaltens auf der Geländeoberfläche. Durch Kopplungspunkte an Schächten und/oder Sinkkästen findet zudem ein Volumenaustausch zwischen dem 1D-Kanalnetzmodell und dem 2D-Oberflächenmodell statt. Eine auf den Ausführungen im DWA-M 119 (2016) aufbauende ausführliche Beschreibung der Verfahren zur detaillierten Überflutungsberechnung erfolgt im „Praxisleitfaden zur Ermittlung von Überflutungsgefahren mit vereinfachten und detaillierten hydrodynamischen Modellen“ (HSB 2017). Es werden zudem die unterschiedlichen Möglichkeiten zur Abflussbildung im Rahmen einer gekoppelten Berechnung erläutert, die über das Kanalnetzmodell, das Oberflächenmodell oder einer Kombination aus beidem erfolgen kann (vgl. HSB 2017).

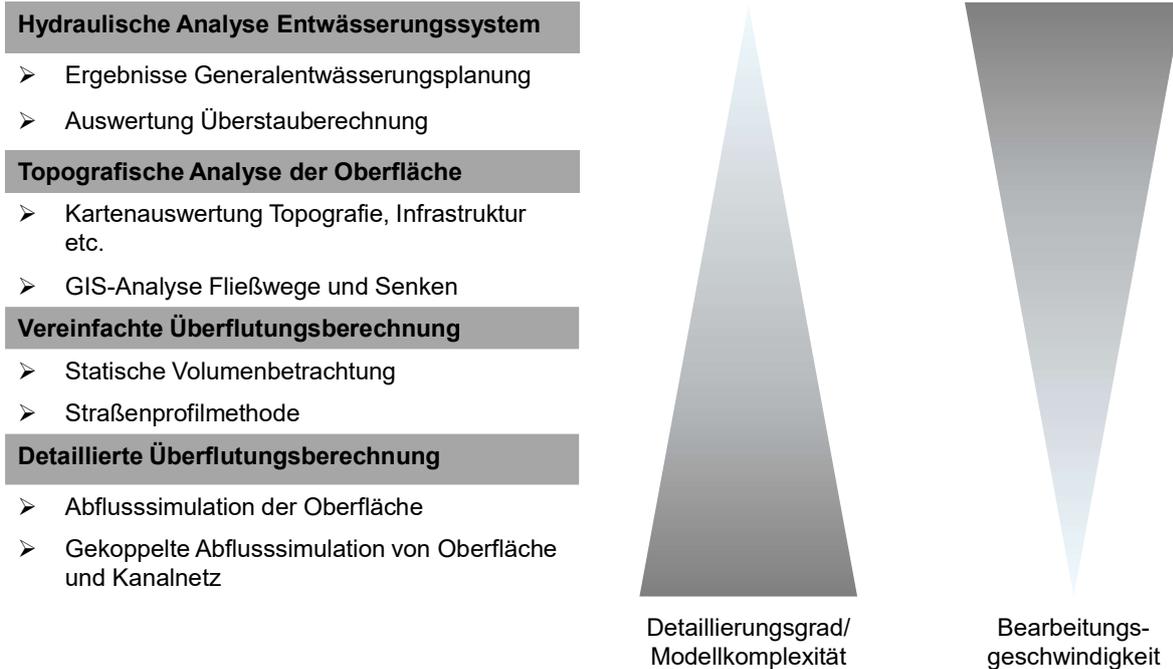


Abbildung 2-6: Methoden und Verfahren zur Analyse der Überflutungsgefährdung (verändert nach Tabelle 2, DWA M-119 (2016))

Während sich zu Beginn der 2000er Jahre die detaillierte modelltechnische Nachbildung der Überflutungsvorgänge auf der Oberfläche noch als schwierig erwies (DWA 2006a) und auf Sonderfragestellungen beschränkte (ATV-DWWK 2004b) sind stadtgebietsweite Starkregengefahrenkarten mit detaillierter Darstellung von Überflutungsflächen und Wasserständen mittlerweile keine Seltenheit mehr, wie eine Umfrage der DWA Arbeitsgruppe ES-2.5 zeigt (Krieger 2018). Darüber hinaus wurde für Nordrhein-Westfalen (NRW) 2021 eine flächendeckende Starkregengefahrenhinweiskarte veröffentlicht (BKG 2021), die auf den Berechnungen eines hydrodynamischen 2D-Oberflächenmodells beruht und sukzessive auf ganz Deutschland ausgeweitet werden soll (BKG 11.05.2023). Der Hintergrund hierfür ist insbesondere eine gestiegene Verfügbarkeit der für die Ermittlung von Überflutungsflächen erforderlichen Datenbasis. So stehen hochaufgelöste digitale Geländemodelle (> 1 Messpunkt/m²) sowie Daten zu abflussbeeinflussenden Elementen (u. a. Gebäude) mittlerweile in vielen Regionen kostenlos und hochaufgelöst zur Verfügung. Demgegenüber steht ein erhöhter Rechenaufwand, der mit der hohen Modellkomplexität einhergeht. Im Falle der gekoppelten 1D/2D-Modellierung können Rechendauern bei größeren Untersuchungs- oder ganzen Stadtgebieten schnell mehrere Stunden oder sogar Tage beanspruchen. Aus diesem Grund wird häufig ein abgestuftes Vorgehen zur Gefährdungsanalyse empfohlen (Angermaier et al. 2012; DWA 2013, 2016; MULNV 2018). Dabei werden zunächst mit weniger detaillierten Verfahren überflutungsauslösende Regenereignisse sowie Überflutungs-Hot-Spots ermittelt, für die dann anschließend eine detaillierte Betrachtung mittels gekoppelten 1D/2D-Berechnungen erfolgt.

Mit dem zunehmenden Einsatz von gekoppelten 1D/2D-Berechnungsmodellen wird auch den weiter konkretisierten Anforderungen der DIN EN 752:2017 an die Überflutungsprüfung Rechnung getragen. Während in der Vorgängerversion (DIN EN 752:2008) zwar bereits der Einsatz von zeitveränderlichen Bemessungsregen und computergestützten Modellen zur Abflusssimu-

lation gefordert wurde, erfolgten die Empfehlungen für Überflutungshäufigkeiten als Bemessungskriterien noch Anhand vereinfachter Gebietstypisierungen. In der aktuellen Version wird hingegen eine zunehmend objektspezifischere Betrachtung unter Berücksichtigung des Schadenspotentials gefordert. Darüber hinaus werden für Orte mit hohem Schadens- und Gefährdungspotential umfassende Untersuchungen unter Berücksichtigung von Fließwegen und des Einflusses von Oberflächeneigenschaften wie Bordsteinen empfohlen.

2.4 Abfluss- und Überflutungsvorhersage für den Echtzeiteinsatz

Im Bereich der Überflutungsvorhersage haben sich Vorhersagesysteme für flussbedingte Überschwemmungen in großen Einzugsgebieten seit Jahren etabliert und auch für kleinere Einzugsgebiete wurden Lösungen in Forschungsvorhaben entwickelt (Jasper-Tönnies et al. 2017a; HoWa-innovativ o. J.; HAPLUS o. J.). Die Vorhersage erfolgt dabei häufig auf Grundlage von oberhalb liegenden Pegelständen der vergangenen Zeitschritte und/oder hydrologischen oder hydrodynamischen Modellrechnungen. Die Vorhersage urbaner Sturzfluten infolge einer Überlastung der Entwässerungskapazität von Kanalnetzen stellt dagegen weiterhin eine Herausforderung dar und wird weder in Regelwerken noch in Leitfäden zum Starkregenerisikomanagement berücksichtigt (vgl. DWA, 2016; LUBW, 2016; MULNV, 2018). Aufgrund der geringeren Einzugsgebietsgröße und des hohen Versiegelungsgrades gegenüber natürlichen Einzugsgebieten reagieren urbane Einzugsgebiete deutlich schneller auf ein Niederschlagsereignis. Meist betragen die Reaktionszeiten nur wenige Minuten, weshalb ein Vorhersagemodell auf alleiniger Grundlage aktueller Abflussmessungen nicht zielführend ist.

Bisher untersuchte Ansätze zur Vorhersage von Sturzfluten in kleinräumigen urbanen Einzugsgebieten konzentrieren sich daher darauf, die bevorstehende Überflutungssituation aus Niederschlagsvorhersagen abzuleiten oder zu berechnen. Da Sturzfluten in der Regel durch konvektive Niederschlagsereignisse ausgelöst werden, die mit den derzeitigen Vorhersagemodellen lediglich mit Vorlaufzeiten von bis zu zwei Stunden mit einer adäquaten Genauigkeit vorhergesagt werden können (Faure et al. 1999; Quirnbach 2003; Sun et al. 2014; Jasper-Tönnies et al. 2017b), erfordert die Vorhersage urbaner Sturzfluten schnelle Berechnungsansätze. Aus diesem Grund sind die in Kapitel 2.3 beschriebenen gekoppelten hydrodynamischen 1D/2D Berechnungsmodelle, deren Rechenzeiten mehrere Stunden oder Tage betragen, nicht ohne Weiteres für die Überflutungsvorhersage geeignet. Eine Anwendung in Echtzeitwarnsystemen ist daher nur unter starken Einschränkungen bzw. Vereinfachungen möglich. Gleichzeitig werden aus Sicht der Akteure des kommunalen Krisenmanagements als potentielle Anwender von Echtzeitwarnsystemen hohe Anforderungen an die Verlässlichkeit der Warnmeldungen gestellt. Es gilt somit Methoden zu finden, um ausgehend von einer Niederschlagsprognose möglichst verzögerungsarm die resultierende Überflutungssituationen zeitlich und räumlich hochauflösend zu berechnen.

In den vergangenen Jahren konnten einige Fortschritte darin erzielt werden, die Rechendauer von Überflutungsmodellen zu optimieren. In den nachfolgenden Kapiteln wird hierzu der aktuelle Stand von Wissenschaft und Technik aufgeführt. Dabei werden insbesondere die Ansätze berücksichtigt, die auf eine quantitative Darstellung der Überflutungssituation abzielen (z. B.

Darstellung von ereignisspezifischen Senken und Wasserständen). Qualitative Warnsysteme wie beispielsweise Warnungen, die allein in Abhängigkeit prognostizierter Niederschlagsmengen ausgegeben werden, werden hier hingegen nicht weiter berücksichtigt. Viele Fortschritte konnten mit dem Ziel erreicht werden, größere Gebiete oder mehrere Varianten zu modellieren, diese Entwicklungen tragen dazu bei, dass die Modelle für den Einsatz zur Echtzeitvorhersage interessanter werden. Deshalb wurden nicht nur die Ansätze aufgeführt, die sich explizit darauf konzentrieren, eine Vorhersage zu ermöglichen, sondern allgemein Ansätze, die auf eine verkürzte Rechenzeit abzielen.

2.4.1 Physikalisch-basierte und konzeptionelle Modelle

Eine Übersicht über die Eignung gängiger physikalisch-basierter und konzeptioneller Modelle für den Einsatz in Echtzeitwarnsystemen liefern Henonin et al. (2013). Die Übersicht zeigt, dass hydrodynamische 2D-Oberflächenmodelle und gekoppelte 1D/2D-Modelle zwar den besten Überblick über das tatsächliche Überflutungsausmaß liefern, jedoch zu langsam für den Echtzeiteinsatz sind. Um der Problematik der immensen Rechendauern physikalisch-basierter Modelle zu begegnen, wurden in den vergangenen Jahren verschiedene Ansätze untersucht. Diese Ansätze lassen sich nach Jamali et al. (2018) und Zhao et al. (2021) in die folgenden Kategorien unterteilen:

- Modellvereinfachung
- Reduzierung des Detaillierungsgrades
- Maximale Ausnutzung von Rechenressourcen

Diese Kategorisierung wird auch hier zur Beschreibung des Stands der Wissenschaft und Technik verwendet. Bei den beschriebenen Ansätzen handelt es sich, falls nicht explizit anders erwähnt, um 2D-Überflutungsberechnungen, die im Rahmen der gekoppelten hydrodynamischen Modellierung den größten Rechenaufwand erfordern.

Modellvereinfachung

Ein Großteil des Rechenaufwandes bei der Überflutungsmodellierung ist auf die numerische Lösung der vollständigen 2D-Flachwassergleichungen (vgl. Kapitel 3.4.1) zur Beschreibung des oberflächigen Fließverhaltens zurückzuführen. Mehrere Ansätze, die Rechendauer zu verkürzen, konzentrieren sich daher darauf, die Komplexität dieser Gleichungen durch die Approximation oder Vernachlässigung einzelner Terme zu reduzieren (Hunter et al. 2007; Bates et al. 2010). Andere Ansätze wiederum konzentrieren sich darauf, die Dimension des Oberflächenmodells auf ein 1D-Modell aus Fließwegen und Senken zu reduzieren (Mark et al. 2004; Maksimović et al. 2009; Leandro et al. 2009), in dem das Fließverhalten wie in 1D-Kanalnetzmodellen mittels Saint-Venant-Gleichungen (vgl. Kapitel 3.4.1) beschrieben werden kann. Um die Genauigkeit der Ergebnisse von gekoppelten 1D/1D-Modellen zu erhöhen, haben Leandro et al. (2009) das Modell zudem an den Ergebnissen eines 1D/2D-Berechnungsmodells kalibriert.

Modellvereinfachungen kommen häufig auch in Kombination mit dem Konzept der zellulären Automaten (*cellular Automata*, CA) zur Anwendung. Dabei handelt es sich um eine Berechnungsmethodik zur Modellierung räumlich diskreter dynamischer Systeme. Die Berechnung

erfolgt auf Grundlage eines Netzes aus gleichmäßigen Zellen, wie einem Rasterdatensatz. Jede Zelle verfügt dabei über einen Status, der für jeden Zeitschritt von dem eigenen Status und dem der Nachbarzellen im vorangegangenen Zeitschritt abhängig ist. Der Austausch zwischen den einzelnen Zellen erfolgt in Abhängigkeit von einfachen vordefinierten Übergangsregeln. Im Rahmen der Überflutungsmodellierung kommen dabei hydrodynamische Gleichungen in stark vereinfachter Form (z. B. Manning-Strickler-Gleichung) oder empirischer Zusammenhänge zum Einsatz (Jamali et al. 2019). Der Einsatz von CA-Modellen wurde sowohl zur 1D-Simulation des Fließverhaltens im Kanal (Austin et al. 2014) als auch zur 2D-Simulation an der Geländeoberfläche (Dottori und Todini 2011; Ghimire et al. 2013; Guidolin et al. 2016) untersucht.

Noch kürzere Rechenzeiten können durch konzeptionelle Modelle wie RFSM (Lhomme et al. 2008) und RUFIDAM (Jamali et al. 2018) erreicht werden. Bei diesen nach Zhao et al. (2021) als statische 1D-Modelle bezeichneten Überflutungsmodellen werden die Berechnungen nicht auf einem Rechennetz durchgeführt, sondern erfolgen in einem Netzwerk aus Senken mit den dazugehörigen Einzugsgebieten. Dabei handelt es sich um ein Verfahren, das der im DWA-M 119 (2016) beschriebenen statischen Volumenbetrachtung sehr ähnlich ist. Ein wesentlicher Unterschied ist allerdings, dass gegenüber der statischen Volumenbetrachtung auch die Weitergabe von Abflussvolumina an angrenzende Senkengebiete über sogenannte Überlaufpunkte möglich ist. Beaufschlagt werden diese Modelle mit dem Direktabflussvolumen für ein Niederschlagsereignis oder wie im Falle von RUFIDAM durch das Überstauvolumen aus einem gekoppelten Kanalnetzmodell (Jamali et al. 2018). Dadurch ist zwar eine ereignisspezifische Betrachtung der Überflutungssituation möglich, aufgrund der fehlenden dynamischen Modellierung der zeitlichen Variation der Ereignisse kann jedoch nur das endgültige Ausmaß der Überschwemmung bestimmt werden (Guidolin et al. 2016; Zhao et al. 2021).

Reduzierung des Detaillierungsgrades

Ein weiterer Ansatz, die Rechendauer zu reduzieren, liegt in der Beschränkung des Untersuchungsgebietes auf Teileinzugsgebiete mit relevanten Hotspots (Koelsch et al. 2020; Zhao et al. 2021) oder in der Reduzierung der Auflösung des Rechennetzes (René et al. 2015). Beide Ansätze wurden von Hofmann und Schüttrumpf (2020) für die Anwendung in einem Echtzeitwarnsystem kombiniert. Dabei wurden im Rahmen einer Hotspot-Analyse zunächst besonders gefährdete Einzugsgebiete ermittelt, für die dann anschließend die Auflösung des Rechennetzes iterativ so weit angepasst wurde, dass eine Vorhersage möglich war. Neben der Reduzierung der Rechenelemente können bei einer gröberen Auflösung auch längere Simulationszeitschritte gewählt werden, ohne dass numerische Instabilitäten auftreten, wodurch sich der Rechenaufwand weiter reduziert.

Zwar reichen gröbere Auflösungen meist aus, um Überflutungshotspots zu ermitteln (Stapel et al. 2019), gleichzeitig können so aber kritische Höheninformationen verwischen, wodurch Fließwege oder die Ausbreitung von Überflutungsflächen beeinflusst werden können (Fewtrell et al. 2011). So haben verschiedene Untersuchungen gezeigt, dass die Simulationsgenauigkeit durch hochauflösende Oberflächenmodelle erheblich verbessert werden kann (Fewtrell et al. 2008; Sampson et al. 2012; Meesuk et al. 2015; Almeida et al. 2018). Bei den Untersu-

chungen wurde teilweise neben Aufnahmen aus dem flugzeuggestützten Laserscanning (*Airborne Laserscanning*, ALS) auch auf terrestrische Laserscans als Grundlage für die Rechenetze zurückgegriffen, um Auflösungen bis in den Dezimeterbereich zu ermöglichen. Für den Einsatz in Echtzeitwarnsystemen in Kombination mit hydrodynamischen 2D- oder gekoppelten 1D/2D-Modellen werden solch feine Auflösungen allerdings derzeit aufgrund des hohen Rechenaufwands noch als ungeeignet angesehen (Zhao et al. 2021).

Maximale Ausnutzung von Rechenressourcen

Eine weitere Entwicklung, die auch zur Optimierung von Überstau- und Überflutungsberechnung zum Einsatz kommt, ist die parallele Datenverarbeitung. Während in der Hardwareentwicklung ursprünglich Leistungssteigerungen durch höhere Taktraten einzelner Prozessoren (*Central Processing Unit*, CPU) erreicht wurden, konzentrieren sich Hersteller seit einigen Jahren auf den Einbau mehrerer Prozessoren, um die Rechenleistung zu verbessern. Diese Änderung in der Hardwarearchitektur erfordert aber gleichzeitig auch Änderungen in der Programmierung von Algorithmen, die für eine parallele Datenverarbeitung ausgelegt sein müssen, um die Vorteile der neuen Technologie nutzen zu können (Burger et al. 2014b). Sowohl für die 1D-Kanalnetzrechnung (Tandler 1994; Burger et al. 2014a) als auch für die 2D-Oberflächenabflussrechnung (Neal et al. 2010; Leandro et al. 2014) wurden bereits entsprechende Ansätze entwickelt. Durch die parallele Datenverarbeitung finden neben den klassischen CPUs auch Grafikkarten (*Graphical Processing Units*, GPU) immer mehr Anwendung im Bereich des wissenschaftlichen Rechnens. Die Nutzung von GPUs setzt jedoch einen erheblichen Parallelisierungsgrad des verwendeten Algorithmus voraus, weshalb die Nutzung nicht ohne Weiteres für alle Probleme geeignet ist (Burger et al. 2014b).

Algorithmen zur parallelen Datenverarbeitung haben sich mittlerweile immer mehr in der Praxis etabliert und kommen in der Überflutungsmodellierung in immer mehr kommerziellen Softwareanwendungen zum Einsatz, wie ein Vergleich von Néelz und Pender (2013) zeigt. Auch die in Deutschland weitverbreiteten Anwendungen zur hydrodynamischen Modellierung, wie ++SYSTEMS oder MIKE+, ermöglichen die parallele Berechnung, wobei letztere sogar die Berechnung mittels GPU zulässt. Auch das 2D-Berechnungsmodell HiPIMS (Liang und Smith 2015) lässt eine Berechnung mittels Multi-Core-CPU und GPU zu und eignet sich zur hochauflösenden Modellierung großflächiger Einzugsgebiete. So wurde für das gesamte Gebiet des Landes Nordrhein-Westfalen eine Starkregengefahrenhinweiskarte mit einem 1 x 1 m-Raster erstellt, die im Oktober 2021 durch das Bundesamt für Kartographie und Geodäsie (BKG) veröffentlicht wurde. Ming et al. (2020) testeten den Einsatz von HiPIMS zudem in einem Echtzeitwarnsystem für fluviale Überflutungen. Das dabei berücksichtigte Einzugsgebiet mit einer Fläche von 2.500 km² wurde mit einem Rechenetz mit einer Auflösung von 10 m modelliert. Im Echtzeitbetrieb werden so Vorlaufzeiten von 34 Stunden erreicht, wenn die Wettervorhersagen 36 Stunden vor einem Ereignis verfügbar sind.

Neben der Parallelisierung von Prozessen können durch das verteilte Rechnen die Rechenressourcen erhöht und somit die Rechenzeiten weiter reduziert werden. Hierbei wird anstelle eines einzelnen Rechners auf einen Rechnerverbund zur Lösung von Berechnungsaufgaben zurückgegriffen. Je nach Infrastruktur wird dabei beispielsweise über lokale Netzwerke (*Clus-*

ter-Computing) oder das Internet (*Cloud-Computing*) auf zusätzliche Rechenressourcen zurückgegriffen. Letzteres wird unter anderem durch die Software CityCat (Glenis et al. 2018) im Rahmen der Überflutungsmodellierung ermöglicht. Auch die zuvor genannte Software HiPIMS wurde zur Erzeugung der Starkregengefahrenhinweiskarte für NRW in einer Cloud ausgeführt. Bei den in Deutschland weitverbreiteten kommerziellen Anwendungen bietet HYSTEM-EXTRAN eine Erweiterung mit den Berechnungen auf mehreren Rechnern im Unternehmensnetzwerk durchgeführt werden können (itwh GmbH 2022). Mike+ kann dagegen mittlerweile auch in der Microsoft Azure-Cloud betrieben werden (DHI 2022b).

2.4.2 Datengetriebene Ansätze

Neben physikalisch-basierten und konzeptionellen Ansätzen zur Überflutungsmodellierung wurden in den vergangenen Jahren auch zunehmend datengetriebene Ansätze untersucht. Ein Großteil dieser Ansätze basiert auf Maschinellen Lernverfahren und häufig auch auf Künstlich Neuronalen Netzen (*Artificial Neural Network*, ANN). Zwar besteht eine Stärke von datengetriebenen Ansätzen darin, Informationen aus Messdaten zu extrahieren und diese anschließend für die Modellierung und Vorhersage von Prozessen zu nutzen, dies setzt allerdings insbesondere im Bereich des Maschinellen Lernens große Datensätze voraus. Im Bereich der Siedlungsentwässerung stehen Messdaten im Bereich des Kanalnetzes meist nur an wenigen zentralen Punkten wie Sonderbauwerken oder aus zeitlich begrenzten Messkampagnen zur Verfügung. Messwerte zum oberflächigen Abfluss in urbanen Gebieten fehlen dagegen in der Regel sogar gänzlich. Es gibt zwar verschiedene Ansätze, Informationen über den Oberflächenabfluss zu erfassen, diese ermöglichen jedoch keine flächendeckende und gleichzeitig zeitlich hochauflösende Erfassung der Überflutungssituation.

Aufgrund der eingeschränkten vorhandenen Datenbasis greifen datengetriebene Ansätze zur Überflutungsmodellierung meist auf physikalisch-basierte oder konzeptionelle Modelle zurück, und werden mit deren Ergebnissen trainiert. Bei den Ansätzen lässt sich zwischen Untersuchungen unterscheiden, die Wasserstands-, Durchfluss- oder Überstauganglinien im Bereich des Kanalnetzes vorhersagen und Untersuchungen, bei denen Überflutungsflächen oder Wasserstände auf der Geländeoberfläche berechnet werden. Nachfolgend werden verschiedene Ansätze getrennt nach den beiden Gruppen aufgeführt. Gekoppelte 1D/2D Ansätze wie im Bereich der physikalisch-basierten Modellierung sind zum aktuellen Zeitpunkt nicht bekannt.

Wasserstands-, Durchfluss- und Überstauganglinien

Bei der Vorhersage von Ganglinien gibt es zum einen datengetriebene Ansätze, die darauf ausgelegt sind, Vorhersagen bis auf Schachtebene zu erzeugen (Duncan et al. 2013; Abou Rjeily et al. 2017; Kilsdonk et al. 2022; Palmitessa et al. 2022; Schmid und Leandro 2023; Zhu et al. 2023) und zum anderen Ansätze, die sich auf zentrale Bauwerke der Kanalisation, einzelne Messstationen oder Ausläufe konzentrieren (Chang et al. 2014; Zhang et al. 2018; She und You 2019; Piadeh et al. 2023). Bei Letzteren wurden sowohl Möglichkeiten untersucht, datengetriebene Modelle auf Basis von Messdaten (Chang et al. 2014; Zhang et al. 2018; Piadeh et al. 2023) als auch von Simulationsergebnissen (She und You 2019) zu trainieren. Zwar stand bei diesen Untersuchungen nicht immer die Überflutungsvorhersage im Vorder-

grund, dennoch werden diese Ansätze hier aufgeführt, da die Methodik sich auf diesen Anwendungsfall übertragen lässt.

Bei der datengetriebenen Modellierung von Ganglinien bis auf Schachtebene werden derzeit ausschließlich Simulationsergebnisse von physikalisch-basierten Modellen als Zielgröße für den Trainingsprozess verwendet. Hintergrund hierfür ist die eingangs beschriebene geringe räumliche Auflösung bestehender Messnetze. So entwickelten Duncan et al. (2013) und Abou Rjeily et al. (2017) ein Vorhersagemodell auf Basis von Künstlich Neuronale Netzen. Während in Duncan et al. (2013) ein mehrschichtiges Perzeptron (*Multilayer-Perceptron*, MLP) verwendet wurde, kommt in Abou Rjeily et al. (2017) ein nichtlineares autoregressives exogenes (*nonlinear autoregressive with exogenous inputs*, NARX) Neuronales Netz zum Einsatz. Dabei handelt es sich um ein dynamisches Neuronales Netz, das eine rekurrente Verbindung der Ausgabeschicht zulässt (Hang Xie et al. 2009). Als Eingabe nutzen Duncan et al. (2013) Niederschlagsmessungen, wodurch der maximale Vorhersagehorizont auf die Konzentrationszeit des Einzugsgebietes begrenzt ist. Es wird aber bereits darauf hingewiesen, dass durch die Berücksichtigung von Niederschlagsvorhersagen der Zeitraum ausgedehnt werden kann, wie dies auch bei Abou Rjeily et al. (2017) der Fall ist.

Auch Chang et al. (2014) nutzten in ihren Untersuchungen zur Wasserstandsvorhersage in einem Speicherbecken eines Pumpwerks ein NARX Neuronales Netz. Sie verglichen die Leistungsfähigkeit mit einem Rekurrenten Neuronale Netz (*Recurrent Neural Network*, RNN) und einem vorwärtsgerichteten ANN mit einer verdeckten Schicht. Im Ergebnis lieferte das NARX neuronale Netz die besten Vorhersagen. She und You (2019) kombinieren die Vorteile zweier Netzarchitekturen und entwickelten ein gekoppeltes Vorhersagemodell für den Durchfluss an einem Auslaufpunkt eines Kanalnetzes. Dabei wird ein radiale Basisfunktionen (*radial basis function*) Neuronales Netz zur Vorhersage von monoton steigenden und fallenden Teilen der Durchflussganglinie genutzt und ein NARX Neuronales Netz für die Vorhersage des Scheitels. Neuere Untersuchungen greifen insbesondere auf Rekurrente Neuronale Netze mit *Long Short-Term Memory* (LSTM)-Zellen zurück (Zhang et al. 2018; Kilsdonk et al. 2022; Zhu et al. 2023). So verglichen Zhang et al. (2018) zur Vorhersage von Mischwasserentlastungen ein RNN mit LSTM-Zellen mit einem MLP und konnten herausstellen, dass insbesondere bei längeren Vorhersagehorizonten LSTM-Netze deutliche bessere Ergebnisse erzielen als das MLP. Kilsdonk et al. (2022) und Zhu et al. (2023) nutzten hingegen LSTM-Netze zur Vorhersage von Überstauereignissen an sämtlichen Schächten eines Untersuchungsgebiets.

Piadeh et al. (2023) nutzten ein Ensemble an ML-Modellen zur Prognose von Wasserstandsklassen. Durch die Berücksichtigung eines Ensembles aus schwachen Lernern (*weak learners*) konnte die Vorhersagequalität gegenüber der Nutzung einzelner Verfahren verbessert werden. Auch Schmid und Leandro (2023) nutzten zur Überstauvorhersage ein Ensemble bestehend aus unterschiedlich parametrisierten Neuronale Netzen, um modellbedingte Unsicherheiten bei der Prognose zu berücksichtigen. Palmitessa et al. (2022) stellten einen weiteren Ansatz vor, bei dem ein physikalisch geleitetes Modell verwendet wird. Hierbei berechnet ein Neuronales Netz zunächst die Zu- und Abflüsse eines Knotens, mit denen anschließend über eine Massenbilanz das Überstauvolumen ermittelt wird. Überstau tritt dementsprechend immer auf, wenn der Zufluss zum Knoten größer als der Abfluss ist. Da die Massenbilanz als

Zwischenschicht in das Netz integriert wurde, fließt das Ergebnis in die Fehlerfunktion ein und trägt so zur Netzoptimierung während des Trainingsprozesses bei.

Überflutungsflächen

Zur datengetriebenen Modellierung von Überflutungsflächen geben Mosavi et al. (2018) einen generellen Überblick über bestehende ML-Ansätze und Bentivoglio et al. (2022) speziell über tiefe Neuronale Netze. Trainiert werden diese Modelle in der Regel analog zu den Überstauvorhersagemodellen, mit den Ergebnissen von hydrodynamischen Berechnungsmodellen. Ziel ist es dabei, möglichst identische Ergebnisse in einem Bruchteil der Berechnungszeit zu erzielen. Die untersuchten Ansätze unterscheiden sich teilweise deutlich hinsichtlich der Herangehensweise. So nutzten beispielsweise Bermúdez et al. (2018) ein ANN zur Vorhersage der maximalen Überstauvolumina in Teileinzugsgebieten für ein Niederschlagsereignis. In Abhängigkeit der prognostizierten Überstauvolumina wird dann aus einem Ergebniskatalog von vor-simulierten Ereignissen eine passende Überflutungssituation ausgewählt. Einen ähnlichen Ansatz verfolgen Lin et al. (2013) und Jhong et al. (2017) mit einem auf *Support Vector Regression* (SVR) basierten Modell. Das SVR-Modell wurde dort zunächst dafür genutzt, um die Wasserstandsganglinie auf der Geländeoberfläche an verschiedenen Referenzpunkten im Untersuchungsgebiet vorherzusagen. Anschließend wird ein Klassifikationsmodell genutzt, um in Abhängigkeit der prognostizierten Wasserstände an diesen Punkten in Kombination mit geografischen Informationen die Überflutungsflächen zu ermitteln.

Auch Bermúdez et al. (2019) entwickelten ein SVR-Modell zur Vorhersage von Überflutungsflächen. Die Vorhersage der Wasserstände wird dabei direkt für 25.000 Punkte anstelle von wenigen Referenzpunkten durchgeführt, sodass direkt eine räumliche Information über das Überflutungsgeschehen vorliegt. Das gleiche Vorgehen nutzten auch Berkahn et al. (2019), griffen aber auf ein MLP als Netzarchitektur zurück. Der Fokus lag dabei darauf, für ein bevorstehendes Niederschlagsereignis mit einer Dauer von bis zu zwei Stunden die resultierenden maximalen Wasserstände zu prognostizieren. Zur Ermittlung der Überflutungssituation wurde ein Ensemble aus vollständig verbundenen Netzen mit gleicher Netzarchitektur aber unterschiedlich initialisierten Gewichten verwendet. Hierdurch sollte der Einfluss von Unsicherheiten auf das Ergebnis reduziert werden. Ein Nachteil von vollständig verbundenen Netzen ist jedoch die hohe Anzahl an zu trainierenden Gewichten (Parametern) und damit einhergehend ein hoher Rechen- und Arbeitsspeicherbedarf. Anders ist dies bei *Convolutional Neural Networks* (CNN), die sich die Gewichtsmatrix im Raum teilen, wodurch sich die Anzahl der Gewichte stark reduziert (LeCun et al. 2015). Guo et al. (2020) nutzten CNNs in einer *Autoencoder*-Architektur und konnten so mit einem Modell die Überflutungssituation für ganze Städte berechnen. Auch Hofmann und Schüttrumpf (2021) nutzten CNNs, ordneten diese aber in einem *Generative Adversarial Network* (GAN) an. Der Ansatz von Hofmann und Schüttrumpf (2021) nutzt zudem räumliche Niederschlagsinformationen als Eingabedaten, um so auch eine ungleichmäßige Überregnung des Einzugsgebietes zu berücksichtigen.

Um eine Übertragbarkeit der trainierten Modelle auf andere Gebiete zu ermöglichen, nutzten Guo et al. (2020) und Löwe et al. (2021) räumliche Informationen, wie die Geländehöhe, -neigung oder die Abflussakkumulation, als zusätzliche Eingabegröße. Seleem et al. (2023) nutzten ebenfalls räumliche Eingaben und konnten im Hinblick auf die Übertragbarkeit die gute

Performance einer CNN-Architektur gegenüber einem *Random-Forrest-Algorithmus* herausstellen. Darüber hinaus nutzten do Lago et al. (2023) räumliche Informationen in Kombination mit einem GAN, um das mit einem hydrologischen Modell ermittelte Niederschlagabflussvolumen in einem Untersuchungsgebiet zu verteilen. Bentivoglio et al. (2023) nutzten mit einem *Graph Neural Network* (GNN) eine Graph-basierte Netzstruktur, bei der die Knoten des Graphs die Zellen eines Rechnernetzes und die Kanten die Verbindung der einzelnen Zellen abbilden. Während die zuvor aufgeführten CNN- oder GAN-basierten Netzarchitekturen nur rasterbasierte Daten verarbeiten können, ist mit dem GNN auch die direkte Verarbeitung der Berechnungsergebnisse von Simulationen mit dreiecksbasierten Rechnernetzen möglich.

Auch bei ähnlichen Aufgabenstellungen wie der Überflutungsempfindlichkeitsmodellierung (Zhao et al. 2020; Tien Bui et al. 2020; Seleem et al. 2022) oder der Vorhersage fluvialer Überflutungen (Syed Kabir et al. 2020; Lin et al. 2020) konnten tiefe Neuronale Netze vielversprechende Ergebnisse erzielen. Alle aufgeführten Ansätze führen gegenüber hydrodynamischen Modellen zu einer deutlich reduzierten Rechendauer. Je nach berücksichtigter Untersuchungsgebietsgröße und der vorhandenen Rechenressourcen werden so nur Rechendauern von wenigen Minuten oder sogar Sekunden zur Ergebniserzeugung benötigt.

2.5 Ansatzpunkte, Untersuchungsmethodik und thematische Abgrenzung

Der aktuelle Stand von Wissenschaft und Technik zeigt, dass gerade in der jüngeren Vergangenheit zahlreiche Ansätze zur schnellen Überflutungsberechnung untersucht wurden. Insbesondere mit Maschinellen Lernverfahren und speziell Künstlich Neuronale Netzen konnten dabei besonders vielversprechende Ergebnisse erzielt werden. Das bezieht sich zum einen auf die Reduktion der Rechendauern auf wenige Minuten oder sogar Sekunden im Vergleich zu mehreren Stunden bis Tagen bei hydrodynamischen Berechnungsmodellen. Zum anderen weisen die Ergebnisse meist dennoch eine hohe räumliche Auflösung auf, die im Bereich der Berechnung von Überflutungsflächen in der Regel sogar der von hydrodynamischen Modellen entspricht. Dadurch bilden diese Verfahren einen vielversprechenden Ansatz für die Prognosegenerierung in Echtzeitwarnsystemen und wurden daher auch in der vorliegenden Arbeit für die entwickelten Vorhersagemodelle verwendet.

Der Fokus der bisherigen Untersuchungen zur ML-basierten Berechnung von Überflutungsflächen lag auf der Ausgabe eines Rasters mit den maximalen Wasserständen für ein Niederschlagsereignis (vgl. Kapitel 2.4.2). Damit ist zwar bekannt, welche Bereiche potentiell gefährdet sind, nicht jedoch der Zeitpunkt, an dem die Gefährdung auftritt. Insbesondere für die Akteure des kommunalen Krisenmanagements wie Feuerwehr und Rettungsdienst, sind dies aber wichtige Informationen, um eine möglichst effiziente Einsatzplanung vorzunehmen. Da hohe Wasserstände auch Rettungswege versperren können, ist es besonders wichtig zu wissen, wann kritische Punkte versperrt oder wieder passierbar sind, um stets auf dem schnellsten Weg zum Einsatzort zu gelangen und so wertvolle Zeit zu sparen. Vor diesem Hintergrund ist das Ziel der vorliegenden Arbeit, einen Ansatz zu entwickeln, der nicht nur eine statische, sondern eine ereignisspezifische und dynamische Starkregengefahrenkarte ermöglicht. Gegenüber den übrigen Untersuchungen soll dabei das Abflussverhalten sowohl im Kanalnetz

als auch an der Oberfläche berücksichtigt werden, um eine möglichst hohe Prognosegüte zu erreichen.

Um dieses Vorhaben zu ermöglichen, wird ein zweistufiger Ansatz verfolgt. Dieser besteht aus einem Modell zur Vorhersage der Überstauvolumina an den einzelnen Schächten im Einzugsgebiet (Modell 1) und einem zweiten Modell zur Ermittlung der Überflutungsflächen (Modell 2). Beide Modelle werden verbunden, indem die prognostizierten Überstauganglinien aus dem Modell 1 als Eingangsgröße in das Modell 2 angesetzt werden. Beide Modelle erhalten zudem als Eingabe den gefallen und prognostizierten Niederschlag. Darüber hinaus erhält das Modell 1 Messungen über den aktuellen Füllstand im Kanalnetz. Ziel ist es, so den aktuellen Systemzustand im Kanalnetz mit in die Prognose einfließen zu lassen. Da es für die Erfassung der Zielgrößen Überstau und Überflutungsflächen derzeit noch keine Messdaten gibt, werden diese mit Hilfe eines physikalisch-basierten Ansatzes in Form eines hydrodynamischen 1D/2D-Berechnungsmodells erzeugt. Einen Überblick über den geplanten Modellaufbau sowie die beschriebene Untersuchungsmethodik liefert Abbildung 2-7.

An dieser Stelle sei darauf hingewiesen, dass bei dem gewählten Modellaufbau die ML-basierten Prognosemodelle lediglich so gut werden können wie die hydrodynamischen Berechnungsmodelle. Der Fokus der Untersuchungen lag daher darauf, die Rechenzeit durch den Einsatz von datengetriebenen Modellen in Form von ML-Verfahren drastisch zu reduzieren und gleichzeitig die Abweichung der Ergebnisse zu hydrodynamischen Modellen möglichst gering zu halten. Die Prognosegüte des vorgestellten Systemaufbaus hängt damit maßgeblich von der Qualität des zur Datengenerierung verwendeten hydrodynamischen Modells ab.

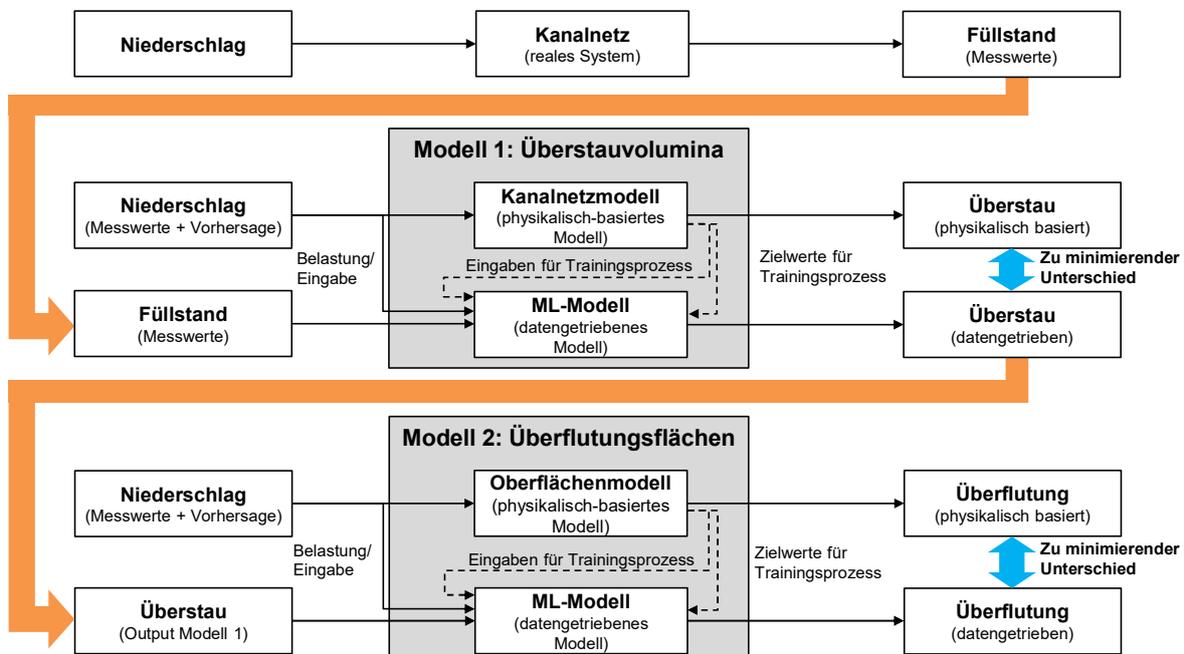


Abbildung 2-7: Übergeordneter Systemaufbau und Untersuchungsmethodik

Die wichtigsten Beiträge dieser Arbeit lassen sich wie folgt zusammenfassen und von den in Kapitel 2.4 beschriebenen Ansätzen abgrenzen:

1. Das entwickelte Modell ist in der Lage, die räumliche und zeitliche Entwicklung der Überflutungssituation zu prognostizieren. Gegenüber anderen Studien, die den Einsatz von ML-Verfahren zur Vorhersage pluvialer Überflutungen untersucht haben (Berkhahn et al. 2019; Guo et al. 2020; Hofmann und Schüttrumpf 2021; Löwe et al. 2021; Seleem et al. 2023; do Lago et al. 2023) wird hier eine Überflutungssequenz für die kommenden Zeitschritte anstelle der maximalen Wasserstände ausgegeben. Die Genauigkeit der Ergebnisse soll dabei weitestgehend an die von physikalisch-basierten Modellen heranreichen, bei gleichzeitig drastisch reduzierten Rechenzeiten.
2. Im Vergleich zu existierenden Studien zur Vorhersage von pluvialen Sturzfluten mit ML-Verfahren, die entweder nur das Kanalnetz (Duncan et al. 2013; Abou Rjeily et al. 2017; Kilsdonk et al. 2022; Schmid und Leandro 2023; Zhu et al. 2023) oder die Oberfläche (Berkhahn et al. 2019; Guo et al. 2020; Hofmann und Schüttrumpf 2021; Löwe et al. 2021; Seleem et al. 2023; do Lago et al. 2023; Bentivoglio et al. 2023) betrachten, erfolgt hier eine gekoppelte Betrachtung beider Systeme.
3. Durch die Integration von Messungen im Kanalnetz in den Modellaufbau wird im Vergleich zu bisherigen Ansätzen auch die Vorbelastung der Entwässerungsstruktur mitberücksichtigt. Obwohl die Messungen nicht unmittelbar als Eingabe für das Überflutungsflächenvorhersagemodell dienen, erfolgt durch die Einbeziehung der Überstauprognose eine indirekte Berücksichtigung.
4. Es werden unterschiedliche Modellaufbauten miteinander verglichen. Das bezieht sich zum einen auf die berücksichtigten Modelleingaben und deren Formate sowie die davon abhängige Modellarchitektur. Zum anderen werden verschiedene moderne, im Bereich der Wasserwirtschaft noch wenig verbreitete *Deep Learning*-Architekturen wie *Graph Neural Networks*, *Generative Adversarial Networks* oder *Transformer* bei den Untersuchungen berücksichtigt und miteinander verglichen.
5. Es wurde ein umfangreicher Trainingsdatensatz erstellt, der mit 253 Ereignissen gegenüber den meisten anderen Studien deutlich mehr Ereignisse umfasst. Zudem beinhaltet der Datensatz sowohl historische Starkregenereignisse als auch Modellregen, um die in der Realität potentiell möglichen Ereignisverläufe und -intensitäten möglichst repräsentativ abzudecken.

3 Material und Methoden

3.1 Beschreibung des Untersuchungsgebietes

Das verwendete Untersuchungsgebiet (vgl. Abbildung 3-1) liegt in Ückendorf, einem Stadtteil im Gelsenkirchener Süden. Der Großteil des Gebietes ist urban geprägt, im Nordosten befinden sich Industrie- und Entwicklungsflächen, darunter mit dem Alma-Gelände ein Landschaftsschutzgebiet. Die an das Kanalnetz angeschlossene befestigte und durchlässig befestigte Einzugsgebietsfläche $A_{E,K}$ beträgt 211 ha (2,11 km²). Die Entwässerung erfolgt größtenteils im Mischsystem. Zwei kleine Neubaugebiete sind zwar bereits im Trennsystem erschlossen, Schmutzwasser- und Regenwasserleitungen werden am Auslass der Gebiete aber wieder gemeinsam ans Mischsystem angeschlossen. Der Großteil des Untersuchungsgebiets entwässert in ein Pumpwerk der Emschergenossenschaft. Von dort aus wird das Abwasser aus dem Einzugsgebiet über den Emscherkanal in Richtung Kläranlage Bottrop geleitet. Die Lage des Gebietes sowie das Gebiet selbst sind in Abbildung 3-1 a) in der Übersicht und einer Detailansicht dargestellt.

Quer durch das Einzugsgebiet verläuft eine Bahnstrecke, die das Einzugsgebiet in einen nördlichen und südlichen Bereich unterteilt. Beide Bereiche werden durch zwei Unterführungen verbunden, die potentiell überflutungsgefährdet sind und bei denen dies teilweise sogar ausgeschildert ist (vgl. Abbildung 3-1 b)). Das Untersuchungsgebiet verfügt über keine stärkeren Geländeneigungen und insgesamt über eine eher flache Topographie. Zudem fließen weder Flüsse noch Bachläufe durch das Untersuchungsgebiet. In der unmittelbaren Umgebung zum Untersuchungsgebiet befinden sich zwar kleinere Bachläufe, jedoch weder Flüsse noch Hänge, die in Richtung des Untersuchungsgebiets entwässern. Eine zusätzliche Gefährdung ist dementsprechend weder durch fluviale Überschwemmungen noch durch Hangabfluss gegeben.

In der Vergangenheit war das Untersuchungsgebiet mehrfach von Starkregen und daraus resultierenden Überflutungen betroffen. Besonders gravierend waren die beiden Ereignisse am 03.07.2009 und 03.07.2010, bei denen an der Niederschlagsstation Gelsenkirchen-Altstadt in nächster Nähe zum Untersuchungsgebiet (vgl. Kapitel 3.2.1) Niederschlagsmengen von über 50 mm und über 70 mm innerhalb einer Stunde gemessen wurden. Gemäß einer erzeugten Niederschlagsstatistik für die Station Gelsenkirchen-Altstadt entspricht das für das Ereignis 2009 bei einer Dauerstufe von 60 Minuten einer Wiederkehrzeit von über 200 Jahren und für das Ereignis 2010 sogar einer Wiederkehrzeit von über 1000 Jahren. Abbildung 3-1 c) zeigt exemplarisch das Ausmaß des Ereignisses von 2009 für einen Straßenzug im Südwesten des Untersuchungsgebiets. Als besonders gefährdet haben sich erwartungsgemäß die Unterführungen erwiesen. Je nach Ereignis werden hier Wasserstände von mehr als zwei Metern erreicht, wie auch die Starkregengefahrenkarte der Stadt Gelsenkirchen (Stadt Gelsenkirchen 2019) und die Starkregengefahrenhinweiskarte des Landes Nordrhein-Westfalen (BKG 2021) bestätigen.

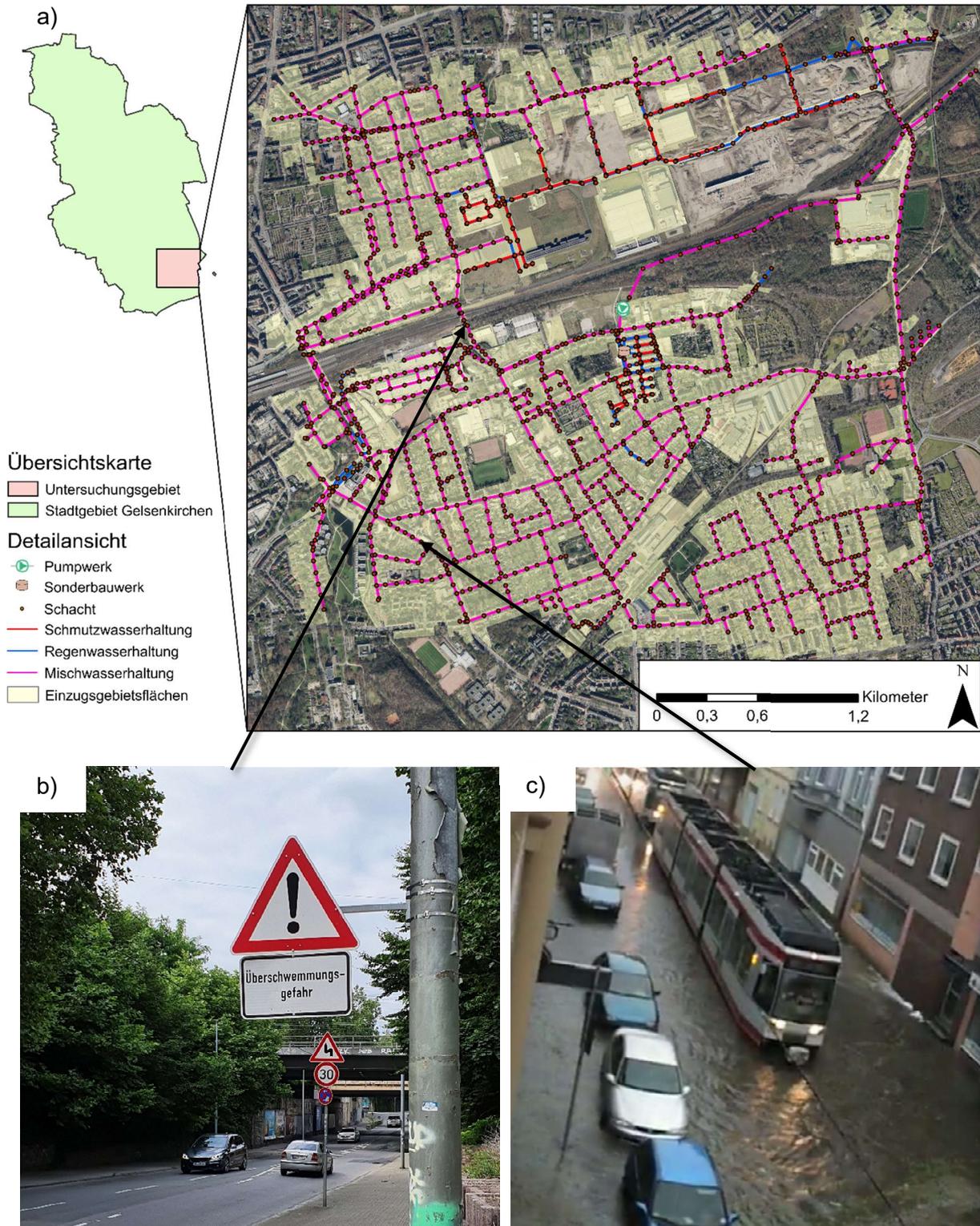


Abbildung 3-1: Übersicht über das Untersuchungsgebiet: a) Übersichtskarte und Detailansicht des Untersuchungsgebiets mit Darstellung des Entwässerungssystems sowie den zugehörigen Einzugsgebietsflächen (Datenquelle: AGG (2021) und EGLV (2021b)) b) Ausgeschilderte Überschwemmungsgefahr vor einer Unterführung im Untersuchungsgebiet c) Ausmaß der Überflutung nach einem Starkregenereignis am 03.07.2009 (Quelle: Screenshot aus einem YouTube Video (Steffi86w 2009))

3.2 Verwendete Datengrundlage

Nachfolgend werden die für das gekoppelte 1D/2D-Überflutungsmodell (vgl. Kapitel 4.2) sowie für die ML-basierten Vorhersagemodelle (vgl. Kapitel 5 und 6) benötigten Datengrundlagen beschrieben. Die zur Generierung der für das spätere Modelltraining erforderlichen Datensätze durchgeführten Schritte werden in Kapitel 4 erläutert.

3.2.1 Historische Niederschlagsdaten

Für die Generierung von Trainingsdatensätzen sowie das spätere Training der Maschinellen Lernverfahren werden Niederschlagsdaten als Systembelastung benötigt. In der vorliegenden Arbeit wurde hierzu auf terrestrische Niederschlagsmessungen von Emschergenossenschaft und Lippeverband (EGLV) zurückgegriffen (EGLV 2021a). Insgesamt wurden acht Stationen berücksichtigt, für die

- lückenfreie Daten für einen Zeitraum > 60 Jahre in 1-minütiger Auflösung vorliegen,
- die in der Nähe zum Untersuchungsgebiet liegen und
- für die extreme Starkregenereignisse bekannt sind.

Eine Übersicht der berücksichtigten Stationen sowie deren Lage zum Untersuchungsgebiet liefert Abbildung 3-2.

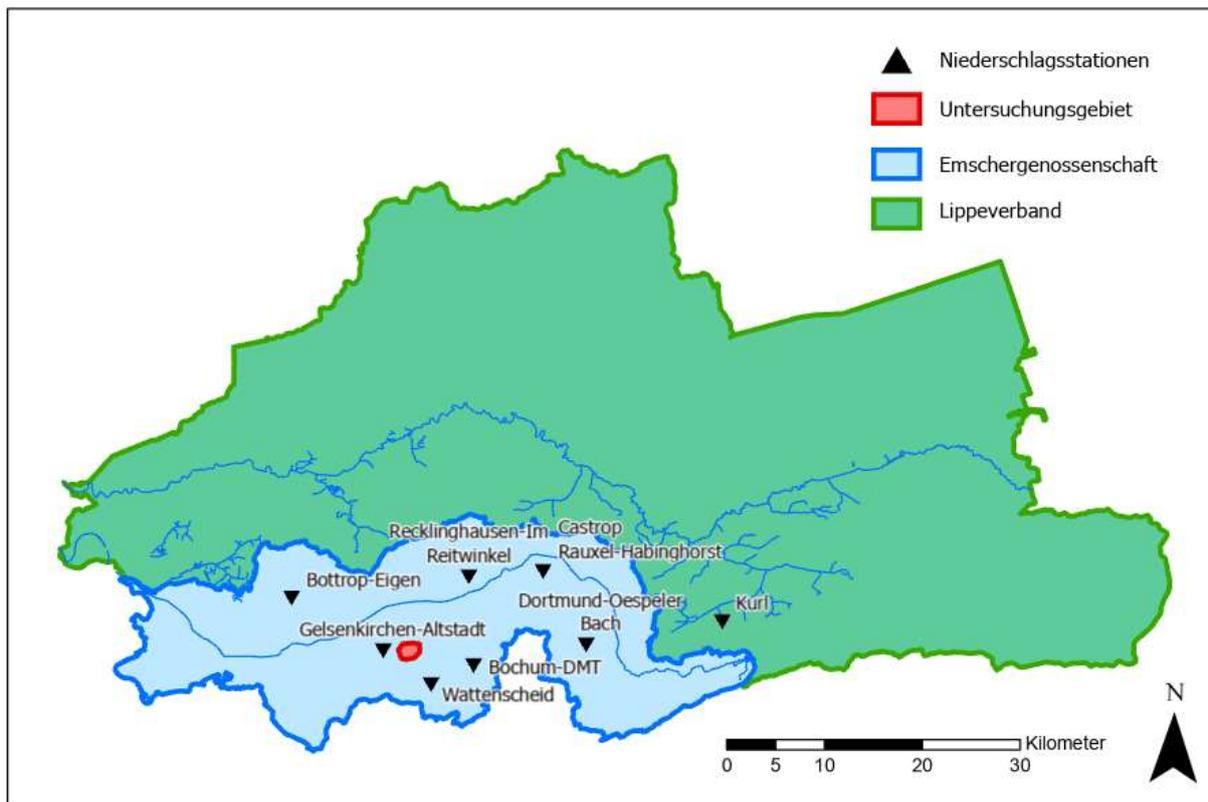


Abbildung 3-2: Übersicht über die berücksichtigten Stationen von Emschergenossenschaft und Lippeverband und deren Lage zum Untersuchungsgebiet, Datenquelle: EGLV (2021a)

Tabelle 3-1 gibt ergänzend die Zeiträume an, für die lückenfreie Daten vorliegen. Keine der

berücksichtigten Stationen liegt unmittelbar im Untersuchungsgebiet. Mit der Station Gelsenkirchen-Altstadt liegt die nächste Station 2,5 km vom Zentrum des Untersuchungsgebietes entfernt. In dieser Arbeit gemachte Aussagen zu Wiederkehrzeiten beziehen sich, soweit nicht anders angegeben, auf eine statistische Auswertung dieser Zeitreihe nach dem DWA-A 531 (2012).

Tabelle 3-1: Übersicht über die berücksichtigten Stationen und die jeweiligen Messzeiträume, Datenquelle: EGLV (2021a)

| Name | Stations-Nr. | Lückenfreie Daten | | |
|--------------------------------|--------------|-------------------|------------|-----------|
| | | von | bis | Dauer [a] |
| Dortmund - Oespeler Bach | 2241 | 01.01.1937 | 31.12.2020 | 84 |
| Castrop-Rauxel - Habinghorst | 2342 | 01.01.1931 | 31.12.2020 | 90 |
| Recklinghausen - Im Reitwinkel | 2368 | 01.01.1937 | 31.12.2020 | 84 |
| Gelsenkirchen - Altstadt | 2582 | 01.01.1931 | 31.12.2020 | 90 |
| Wattenscheid | 2587 | 01.06.1940 | 31.12.2020 | 80,5 |
| Bottrop - Eigen | 2652 | 01.01.1952 | 31.12.2020 | 69 |
| Bochum - DMT | 4630 | 01.01.1931 | 31.12.2020 | 90 |
| Dortmund - Kurl | 8767 | 01.01.1936 | 31.12.2020 | 85 |

3.2.2 Geländemodell

Die Grundlage für die Erstellung eines 2D-Oberflächenmodells zur Überflutungssimulation bildet ein digitales Geländemodell (DGM). In NRW werden digitale Geländemodelle durch die Bezirksregierung Köln kostenlos als 1 m x 1 m Raster (DGM 1) zur Verfügung gestellt. Ferner werden klassifizierte 3D-Messdaten aus dem flugzeuggestützten Laserscanning als unstrukturierte Punktwolke zur Verfügung gestellt, die gleichzeitig auch die Grundlage für die Erstellung des DGM 1 bilden. Die mittlere Punktdichte der 3D-Messdaten liegt bei 4 – 10 Pkt./m² bei einer Lagegenauigkeit von +/- 30 cm und einer Höhengenaugigkeit von +/- 15 cm (Bezirksregierung Köln 2020). Aufgrund der Klassifizierung der Messpunkte ist es problemlos möglich, Geländepunkte zu filtern und ein Geländemodell zu erstellen.

Für die Überflutungsberechnung ist zudem die Berücksichtigung abflussrelevanter Strukturen von besonderer Bedeutung. Hierzu zählen insbesondere Gebäudeumrisse, die ein Fließhindernis darstellen. Auf der anderen Seite gibt es aber auch Strukturen wie Durchlässe oder Brücken, die in Geländemodellen häufig fälschlicherweise als Fließhindernis abgebildet sind. Daher ist hier eine Vorverarbeitung des Geländemodells erforderlich, bei denen Gebäude erhöht und Durchlässe sowie Brücken geöffnet werden. Während Gebäude in der verwendeten hydrodynamischen Modellierungssoftware MIKE+ (vgl. Kapitel 3.4.1) im Oberflächenmodell automatisch um einen vorgegebenen Wert erhöht werden, ist das Öffnen von Durchlässen und Brücken vorab durchzuführen. Eine Überprüfung des DGM 1 der Bezirksregierung ergab, dass dieser Schritt nur teilweise durchgeführt wurde. Daher wurde auf die 3D-Messdaten als Datengrundlage zurückgegriffen und zunächst Messpunkte in den Bereichen von Brücken und Durchlässen entfernt sowie Bruchkanten eingefügt, bevor anschließend mittels Interpolation durch Inverse Distanzwichtung (*Inverse Distance Weighting*, IDW) ein DGM 1 mit einem Meter

Gitterweite erstellt wurde. Weitere den Oberflächenabfluss beeinflussende Strukturen wie Bordsteinkanten, Mauern oder Hofdurchfahrten konnten aufgrund fehlender Daten nicht berücksichtigt werden. Es wurde aber im Rahmen einer Vor-Ort-Begehung geprüft, dass keine besonders relevanten Strukturen fehlen und erste Berechnungsergebnisse plausibilisiert.

3.2.3 Kanalnetzdaten

Für das Untersuchungsgebiet wurde von der Abwassergesellschaft Gelsenkirchen mbH, dem Kanalnetzbetreiber der Stadt Gelsenkirchen, ein Kanalnetzmodell zur Verfügung gestellt (AGG 2021). Das Kanalnetzmodell wurde mit der Software Mike Urban (DHI 2017b) erstellt. Somit konnten alle Bauwerksdaten problemlos in die für die vorliegende Arbeit verwendete Software MIKE+ (DHI 2021b) übernommen werden. Wesentliche Kennzahlen zum Kanalnetz sind in Tabelle 3-2 dargestellt.

Tabelle 3-2: Kennzahlen des verwendeten Kanalnetzmodells, Datenquelle: AGG (2021)

| Bezeichnung | Anzahl/Größe | | |
|---|--------------|-------------|---------------|
| | Mischwasser | Regenwasser | Schmutzwasser |
| Schächte | 1.458 | 183 | 125 |
| Haltungen | 1.459 | 186 | 123 |
| Wehre | 1 | - | - |
| Pumpen | 5 | - | - |
| Summe Leitungslänge [m] | 48.623 | 7.130 | 4.522 |
| Summe Leitungsvolumen [m ³] | 19.695 | 3.873 | 378 |

3.2.4 Daten zur Oberflächenbeschaffenheit

Daten zur Oberflächenbeschaffenheit bilden im Rahmen der gekoppelten Überflutungsberechnung eine wichtige Grundlage für die Bestimmung verschiedener Modellparameter. Hierzu zählen einerseits Parameter, die die Abflussbildung beeinflussen, wie der Befestigungsgrad oder Versickerungsraten, andererseits aber auch Parameter, die den Abflusskonzentrationsprozess beeinflussen, wie die Oberflächenrauheit im 2D-Modell. Im Rahmen der Bearbeitung wurde auf die folgenden beiden Datensätze zurückgegriffen:

- Bodenkarte von Nordrhein-Westfalen im Maßstab 1:50.000 (IS BK 50)
- Befestigte Flächendaten der Emschergenossenschaft (EGLV 2021b)

Die Bodenkarte liefert Informationen zum anstehenden Boden und den Grundwasserverhältnissen, die für das Versickerungsverhalten maßgebend sind. Bei den Daten über befestigte Flächen der Emschergenossenschaft handelt es sich um ein Kataster, das für das gesamte Verbandsgebiet erstellt und fortlaufend aktualisiert wird. Enthalten sind verschiedene undurchlässig oder teildurchlässig befestigte Flächen, die in die folgenden sechs Klassen unterteilt sind:

- Dachflächen
- Öffentliche Verkehrsflächen
- Private saubere Flächen

- Privat schmutzige Flächen
- Halden und Deponien in Betrieb
- Halden und Deponien rekultiviert

Da im Untersuchungsgebiet keine Halden und Deponien vorhanden sind, werden nur die ersten vier Klassen im weiteren Verlauf berücksichtigt. Die übrigen Bereiche werden als eine zusätzliche fünfte Klasse „durchlässige Flächen“ berücksichtigt. Dabei handelt es sich um Gärten, Kleingärten, Parkanlagen, Sportplätze, Grün-, Wald- und Ackerflächen oder sonstige unentwickelte Flächen.

3.3 Künstlich Neuronale Netze

Künstlich Neuronale Netze (*Artificial Neural Networks*, ANN) stellen das in dieser Arbeit vorwiegend verwendete Maschinelle Lernverfahren dar. Nachfolgend werden zunächst allgemeine Grundlagen wie der Aufbau von Neuronalen Netzen und der Trainingsprozess beschrieben. Anschließend folgt die Beschreibung der wesentlichen in dieser Arbeit verwendeten Netzarchitekturen.

3.3.1 Grundlagen

In diesem Kapitel werden die wichtigsten Grundlagen zu Künstlich Neuronalen Netzen erläutert. Darüber hinaus werden die in dieser Arbeit verwendeten Netzarchitekturen kurz dargestellt. Für vertiefte Erläuterungen zu ANN wird auf die entsprechende Fachliteratur verwiesen (Rojas 1996; Goodfellow et al. 2016; Raschka und Mirjalili 2018; Chollet 2018; Géron 2019; Frochte 2021), auf der auch die hier dargestellten Ausführungen im Wesentlichen basieren.

Aufbau eines Neuronalen Netzes

Künstlich Neuronale Netze gelten als eine besonders effiziente Methode des Maschinellen Lernens. Inspiriert durch das Nervensystem des menschlichen Gehirns, bestehen ANN aus hintereinander geschalteten Schichten (*Layern*) aus künstlichen Neuronen, auch als Knoten oder Einheiten bezeichnet, die mit Hilfe von Graphen („künstliche Synapsen“) verbunden sind. Bevor in jedem Neuron über die Weitergabe von Informationen entschieden wird, werden die Eingaben des Neurons durch Gewichte angepasst. Diese Gewichte gilt es, in einem Trainingsprozess an die zu lösende Aufgabe anzupassen, um möglichst genaue Ergebnisse zu erzielen.

Die Entwicklung von Künstlich Neuronalen Netzen reicht bis in die 1940er Jahre zurück, als McCulloch und Pitts ein erstes mathematisches Modell eines Neurons entwickelten (McCulloch und Pitts 1943). Aufbauend auf diesem gemäß seiner Erfinder benannten McCulloch-Pitts-Neuron, entwickelte Rosenblatt einige Jahre später mit dem Perzeptron (Rosenblatt 1958) das erste lernfähige ANN. Auch heute noch stellt das Perzeptron die Grundlage verschiedener Arten von Künstlich Neuronaler Netzen dar. Hierzu zählt auch das mehrschichtige Perzeptron (*Multilayer-Perceptron*, MLP), das in Abbildung 3-3 dargestellt ist und an dem nachfolgend exemplarisch die generelle Funktionsweise eines ANNs erläutert wird.

Die Abbildung zeigt den generellen Aufbau eines Neuronalen Netzes. Darin stellen die grauen

Knoten die Neuronen und die Verbindungen die Graphen mit den trainierbaren Gewichten dar. Die Neuronen sind dabei immer in Schichten organisiert. Ein MLP besteht jeweils mindestens aus einer Eingabe- und einer Ausgabeschicht, über welche die Eingabevariablen entgegengenommen und die berechneten Zielgrößen ausgegeben werden. Darüber hinaus können zwischen Eingabe- und Ausgabeschicht beliebig viele verdeckte Schichten angeordnet werden, in denen die Weiterverarbeitung der Eingabedaten erfolgt. MLPs werden von der Eingabe bis zur Ausgabe vorwärtsgerichtet durchlaufen. Daher werden diese Netze häufig auch als vorwärtsgerichtete Neuronale Netze (*Feedforward Neural Network*) bezeichnet. Je größer die Anzahl der hintereinander geschalteten Neuronenschichten in einem ANN ist, desto komplexere Zusammenhänge können erlernt werden. Lange Zeit fehlten jedoch geeignete Algorithmen, um auch tiefe Netze mit vielen Neuronenschichten effizient zu trainieren. Dies hat sich erst zu Beginn der 2010er Jahre durch Deep Learning geändert. Deep Learning kann als eine Algorithmenmenge aufgefasst werden, die dafür entwickelt wurde, ANNs mit vielen Schichten effizient zu trainieren (Raschka und Mirjalili 2018).

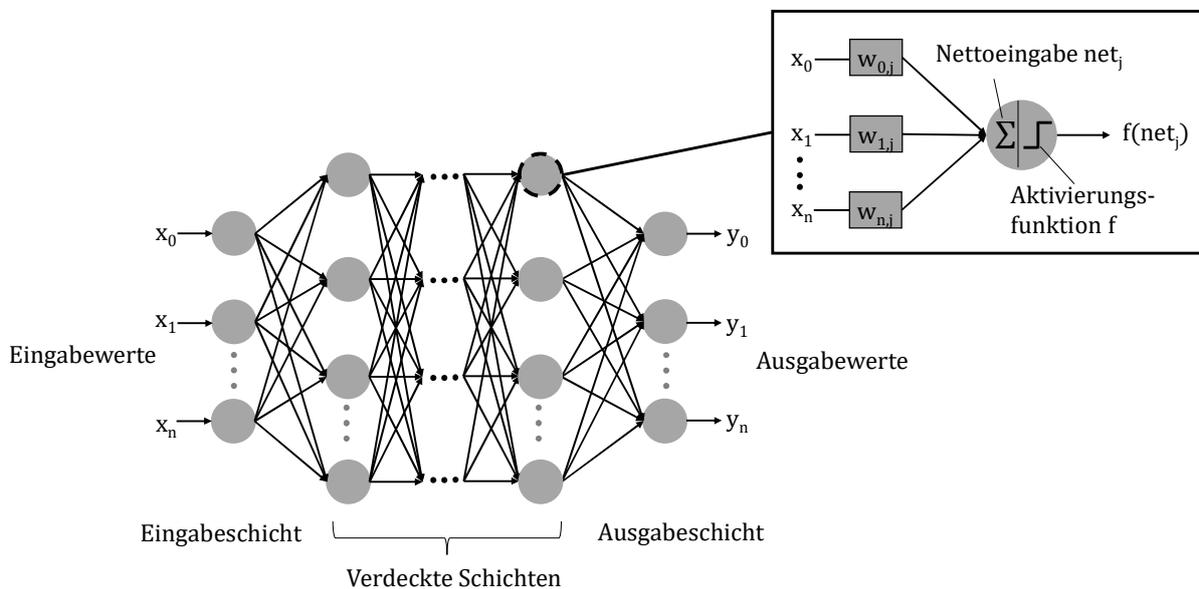


Abbildung 3-3: Schematische Darstellung eines mehrschichtigen Perzeptrons mit Detailansicht eines einzelnen Perzeptrons nach Frochte (2021)

In einem ANN stellen sogenannte künstliche Neuronen die kleinste Verarbeitungseinheit dar, in der Informationen entgegengenommen, verarbeitet und weitergeleitet werden. Die generelle Struktur eines künstlichen Neurons ist in Abbildung 3-3 in der Detailansicht exemplarisch für ein Perzeptron dargestellt. Als Eingabe erhält ein künstliches Neuron einen Merkmalsvektor x mit einer Reihe von Werten x_1, \dots, x_n aus denen es dann die Ausgabe y berechnet. Auf dem Weg zum Neuron werden die Eingabewerte mit den jeweiligen Gewichten w_1, \dots, w_n multipliziert und aufsummiert. Das Ergebnis wird häufig auch als Nettoeingabe net eines Neurons bezeichnet und lässt sich wie folgt beschreiben:

$$net_j = \sum_{i=0}^n w_{ij} \cdot x_i \quad (3-1)$$

Die so gebildete gewichtete Summe net_j ist wiederum der Eingangswert in eine Aktivierungsfunktion $f(net_j)$, die darüber entscheidet, wie die Information weitergeleitet wird. Eine Übersicht über drei gängige Aktivierungsfunktionen liefert Abbildung 3-4. Eine der einfachsten Aktivierungsfunktionen stellt die dort abgebildete Sprungfunktion dar, die der Funktionsweise eines menschlichen Neurons nachempfunden ist. Dabei wird die gewichtete Summe net_j mit einem Schwellenwert θ verglichen, bei dessen Unterschreiten eine 0 und beim Überschreiten eine 1 ausgegeben wird. Im Bereich des Schwellenwerts weist die Sprungfunktion eine Unstetigkeit auf und ist daher nicht differenzierbar. Dadurch ist sie für einen Großteil der Optimierungsverfahren, die im Trainingsprozess zur Anwendung kommen, nicht geeignet. Hier wird stattdessen auf weitere Funktionen wie die Sigmoidfunktion (Han und Moraga 1995) oder die rektifizierte lineare Einheit (*Rectified Linear Unit*, ReLU) (Nair und Hinton 2010) zurückgegriffen.

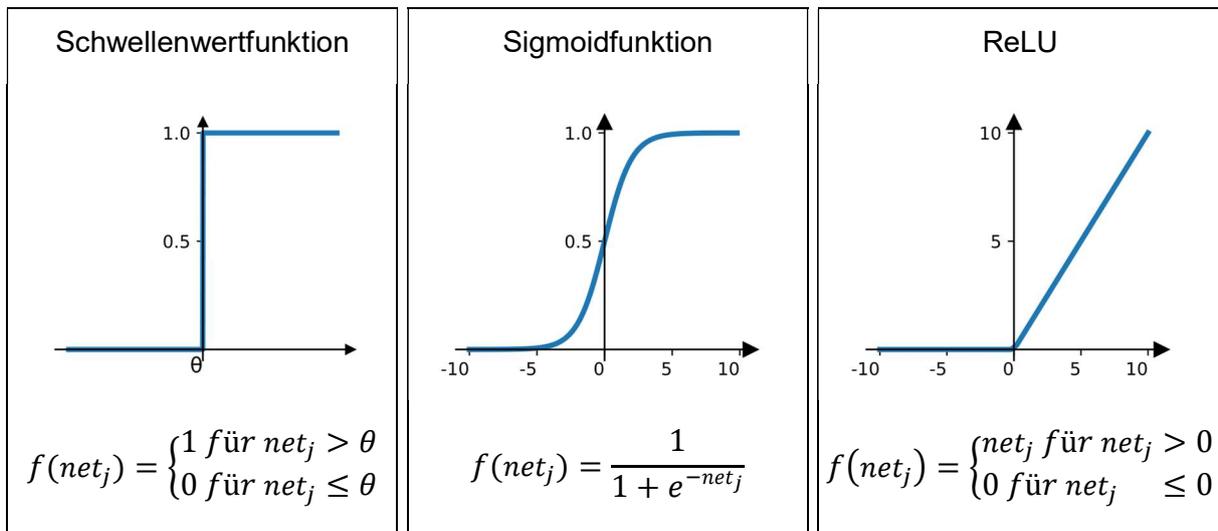


Abbildung 3-4: Übersicht über drei gängige Aktivierungsfunktionen

Das beschriebene einzelne Perzeptron eignet sich gut zur Lösung einfacher linearer Probleme, kommt bei komplexeren nichtlinearen Problemstellungen jedoch schnell an seine Grenzen. Hier ist die Einführung größerer Netze mit mehreren Neuronen und vor allem mehreren Schichten, wie bei dem in Abbildung 3-3 aufgeführten MLP, erforderlich. Der Ablauf bleibt aber der gleiche. Für jedes Neuron in einem MLP wird die Nettoeingabe durch Matrixmultiplikation der Eingaben mit den Gewichtungen ermittelt und anschließend mit Hilfe einer Aktivierungsfunktion die Ausgabe berechnet. Die Eingaben sind dabei entweder die an der Eingabeschicht übergebenen Werte oder, bei tieferen Schichten, die Ausgaben der vorherigen Schicht.

Trainingsprozess

Damit ein ANN für einen Eingabevektor x den gewünschten Ausgabevektor y bestimmen kann, müssen die Gewichte auf die Aufgabenstellung hin angepasst werden. Dieser Prozess wird als Trainingsprozess bezeichnet und erfordert einen Trainingsdatensatz bestehend aus Paaren von Eingaben und Ausgaben. Im Falle des Perzeptrons lässt sich der Trainingsprozess exemplarisch an der Perzeptron-Lernregel beschreiben. Dabei werden zu Beginn die einzelnen Gewichte des Perzeptrons zufällig initialisiert. Anschließend werden die Gewichte iterativ angepasst. Dabei wird je Iterationsschritt für eine Eingabe und die aktuellen Gewichtungen eine Ausgabe berechnet. Anschließend wird der Fehler E zwischen berechneter Ausgabe

$f(net_j)$ und erwarteter Ausgabe y wie folgt ermittelt:

$$E = y - f(net_j) \quad (3-2)$$

Mit Hilfe des Fehlers lässt sich dann der Wert Δw_j berechnen, der wiederum zur Aktualisierung der Gewichte w_j dient:

$$\Delta w_j = \eta \cdot E \cdot x_j \quad (3-3)$$

$$w_j^{neu} = w_j^{alt} + \Delta w_j \quad (3-4)$$

Der Wert η bezeichnet hier die Lernrate, einen Faktor der darüber entscheidet, wie stark die Gewichtungen angepasst werden. Ein hoher Wert führt zwar zu einem schnelleren Lernfortschritt, birgt aber gleichzeitig die Gefahr, dass die Werte zu stark springen und das Minimum nicht erreicht wird. Letzteres kann durch einen niedrigen Wert vermieden werden, der aber gleichzeitig zu einem langsamen Trainingsfortschritt führt. Nachdem die Gewichte aktualisiert wurden, beginnt mit dem nächsten Iterationsschritt der Vorgang von vorn. Der Trainingsprozess ist beendet, wenn entweder der Fehler einen Grenzwert unterschreitet oder eine festgelegte Anzahl an Iterationsschritten erreicht ist.

Bei größeren Netzen wie einem MLP ist der Trainingsprozess etwas komplizierter. In solchen Netzen können je nach Anzahl der Neuronen und Schichten schnell größere Mengen von mehreren Tausend oder sogar Millionen Gewichtungen entstehen, die es zu trainieren gilt. Hier ist der *Backpropagation*-Algorithmus eines der meistgenutzten Lernverfahren, der insbesondere durch die Arbeit von Rumelhart et al. (1986) populär wurde. Dabei handelt es sich um einen rechentechnisch sehr effizienten Trainingsalgorithmus, der gegenüber einigen anderen Lernalgorithmen auch das Training von verdeckten Schichten tiefer Neuronaler Netze zulässt. Damit stellt er eine wichtige Grundlage für das *Deep Learning* dar. Das Ziel des *Backpropagation*-Algorithmus ist es, Fehler eines Netzes zu minimieren. Der Name *Backpropagation* leitet sich daraus ab, dass der Algorithmus hierzu den Fehler an der Ausgabeschicht berechnet und anschließend rückwärts, ausgehend von der Ausgabeschicht, den Fehleranteil der einzelnen Verbindungen ermittelt und diese anpasst. Im Wesentlichen lässt sich das Vorgehen des *Backpropagation*-Algorithmus durch die folgenden Schritte beschreiben:

1. Berechnung der Netzausgabe für die Eingabe eines Trainingsobjektes
2. Ermittlung des Fehlers zwischen berechneter Netzausgabe und Soll-Ausgabe mit Hilfe einer Fehlerfunktion
3. Ermittlung des Fehlerbeitrags der einzelnen Verbindungen eines Netzes
4. Aktualisierung der Gewichte der Verbindung, um den Fehler zu minimieren

Eine für die Ermittlung des Fehlers häufig eingesetzte Funktion ist die Summe der quadratischen Abweichungen, mit der sich der Fehler für ein Trainingsobjekt i wie folgt berechnen lässt:

$$E = \frac{1}{2} \sum_i (y^i - f(net_j))^2 \quad (3-5)$$

Der Index j ist hier mit der Ausgabeschicht gleich zu setzen. Diese Fehlerfunktion gilt es nun

in Richtung eines Minimums zu optimieren. Zur Optimierung der Fehlerfunktion greift der *Backpropagation*-Algorithmus auf Gradientenverfahren zurück, die die Fehlerfunktion schrittweise in Richtung des negativen Gradienten, also in Richtung des Minimums optimieren. Die für die Optimierung erforderlichen Gradienten werden durch den *Backpropagation*-Algorithmus mittels partieller Ableitung bestimmt. Bei der Fehlerfunktion E handelt es sich bei ANNs mit mehreren Schichten um eine verschachtelte Funktion, da die Ausgaben der Ausgabeschicht von den Ausgaben der verdeckten Schichten abhängig sind. Um die partielle Ableitung der Fehlerfunktion E nach den Gewichtungen zwischen den Neuronen i und j bilden zu können, greift der Backpropagation-Algorithmus wie folgt auf die Kettenregel zurück:

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial f(net_j)} \cdot \frac{\partial f(net_j)}{\partial w_{ij}} \quad (3-6)$$

Dadurch lässt sich dann der Fehlerbeitrag δ_j der einzelnen Neuronen berechnen. Dabei wird unterschieden, ob es sich bei j um ein verdecktes oder ein Ausgabeneuron handelt:

$$\delta_j = \begin{cases} f'(net_j) \cdot (y_j - f(net_j)) & \text{falls } j \text{ Ausgabeneuron} \\ f'(net_j) \cdot \sum_k \delta_k \cdot w_{jk} & \text{falls } j \text{ verdecktes Neuron} \end{cases} \quad (3-7)$$

Der Index k bezieht sich hier auf die Nachfolgerneuronen von j . Mit dem berechneten Fehlerbeitrag eines Neurons lässt sich dann wiederum die Aktualisierung der Gewichtungen Δw_{ij} berechnen, mit der anschließend die Gewichte in die negative Richtung des Gradienten angepasst werden können.

$$\Delta w_{ij} = -\eta \cdot f'(net_i) \cdot \delta_j \quad (3-8)$$

Im Anschluss an die Aktualisierung der Gewichtungen startet der Vorgang in eine neue Iterationsschleife. Die Aktualisierung der Gewichtungen nach jedem Trainingspärichen (*Sample*) eines Datensatzes ist jedoch sehr rechenintensiv. Daher ist es üblich, zunächst den Fehler für mehrere Trainingspärichen eines Stapels (*Batches*) zu berechnen, diesen zu mitteln und auf Grundlage des Mittelwertes für den Stapel die Gewichtungen zu aktualisieren. Jede Iteration über alle Pärichen und Stapel eines Trainingsdatensatzes wird als Epoche bezeichnet. Der gesamte Trainingsprozess eines Neuronalen Netzes dauert in der Regel mehrere Epochen an.

Wenn das Netz gut lernt, nimmt mit zunehmenden Trainingsfortschritt der am Trainingsdatensatz ermittelte Fehler immer weiter ab, was aber nicht zwingend eine verbesserte Leistungsfähigkeit bedeutet. Vielmehr ist es ein wichtiges Ziel Maschinelles Lernverfahren wie ANNs, ein Modell zu entwickeln, das gut generalisieren kann und auch bei nicht zuvor gesehenen Datensätzen gute Ergebnisse erzielt. Bei zu starker Anpassung des Modells an die Trainingsdaten kann es jedoch passieren, dass die Generalisierungsleistung des Netzes verloren geht. Dieses Phänomen wird als Überanpassung (*Overfitting*) bezeichnet und ist in Abbildung 3-5 dargestellt. Um dieses Problem zu vermeiden, wird häufig ein Validierungsdatsatz verwendet, für den für jeden Trainingsschritt parallel der Fehler berechnet wird, der jedoch nicht zum Aktualisieren der Gewichte genutzt wird. Dieser Fehler sinkt mit zunehmenden Trainingsfortschritt zunächst ab, nimmt aber dann irgendwann zu, wenn es zur Überanpassung kommt. Das Training sollte dann an dieser Stelle beendet werden.

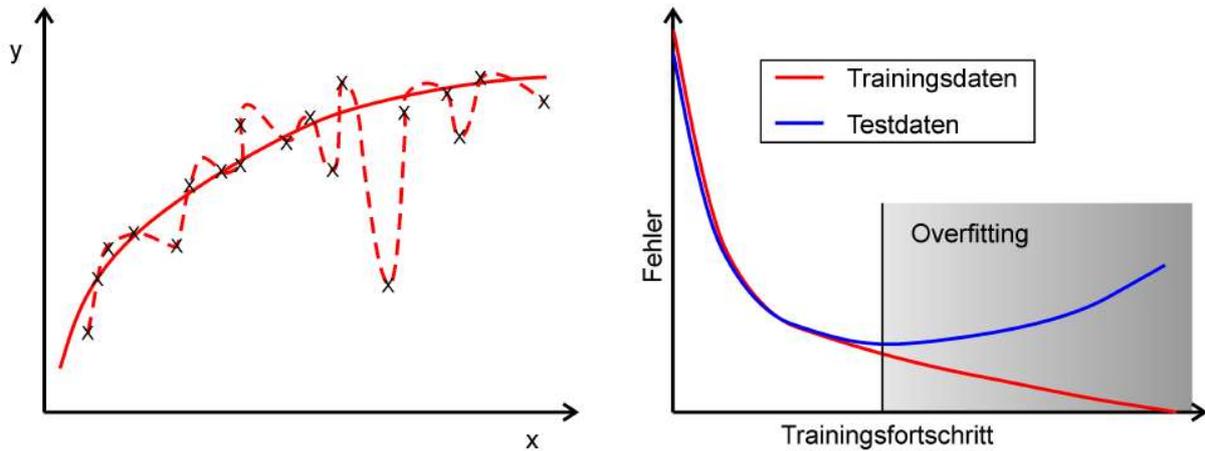


Abbildung 3-5: Schematische Darstellung des *Overfittings* (Woernle 2008), links: Darstellung einer überangepassten Modellfunktion (gestrichelte Linie) und einer Modellfunktion mit höherem Fehler aber besserer Generalisierung (durchgezogene Linie), rechts: Entwicklung der Fehler von Trainings- und Testdaten mit Fortschritt des Trainings

3.3.2 Netzarchitekturen

Neben dem im Kapitel 3.3.1 aufgeführten MLP haben sich in den letzten Jahren verschiedene weitere Netzarchitekturen als besonders effizient für einzelne Lernaufgaben erwiesen. Nachfolgend werden einige Architekturen erläutert, die in dieser Arbeit zur Entwicklung der Vorhersagemodelle Verwendung finden.

3.3.2.1 Rekurrente Neuronale Netze (*Recurrent Neural Networks, RNN*)

RNNs verfügen gegenüber MLPs über Rückkopplungen, über die die Ausgaben des Modells wieder Eingang in selbiges finden. Das bedeutet, dass bei einem dreischichtigen ANN die verdeckte Schicht nicht nur Eingaben aus der Eingabeschicht erhält, sondern auch von der verdeckten Schicht der vorangegangenen Eingabe (vgl. Abbildung 3-6). Die Rückkopplung bewirkt somit, dass Informationen während der gesamten Verarbeitungsdauer einer Sequenz im Netz verbleiben. Dadurch ist diese Netzarchitektur besonders gut geeignet, um sequenzielle Daten wie Text oder Zeitreihen zu modellieren. Diese Fähigkeit wird häufig auch als „Gedächtnis“ bezeichnet (Chollet 2018; Raschka und Mirjalili 2018).

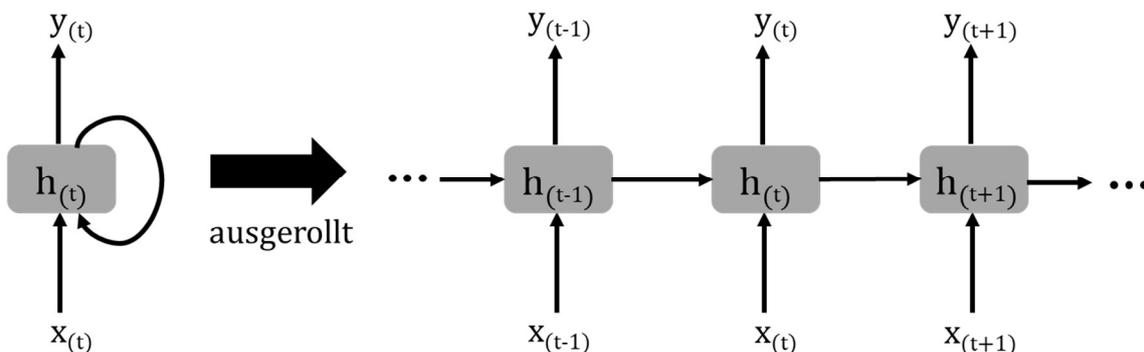


Abbildung 3-6: Schematische Darstellung eines RNNs mit Rückkopplung (links) und eines RNNs im ausgerollten Zustand (rechts) (Raschka und Mirjalili 2018)

Die Ausgabe y eines rekurrenten Neurons zum Zeitpunkt t ist eine Funktion der Eingaben $x_{(t)}$ und der Ausgaben aus dem vorherigen Zeitschritt $h_{(t-1)}$ sowie der jeweiligen Gewichtsvektoren W_x und W_h . Durch die folgende Formel lässt sich die Aktivierung eines rekurrenten Neurons beschreiben (b ist der *Bias*-Term und f die Aktivierungsfunktion wie z. B. die Sigmoidfunktion vgl. Abbildung 3-4):

$$h_{(t)} = y_{(t)} = f(W_x x_{(t)} + W_h h_{(t-1)} + b) \quad (3-9)$$

Trainiert werden RNNs durch eine spezielle Art des *Backpropagation*-Algorithmus, dem *Backpropagation-through-time* (BPTT)-Lernalgorithmus. Dadurch, dass die Neuronen in einer verdeckten Schicht neben der aktuellen Eingabe $x_{(t)}$ auch die Ausgabe aus dem vorangegangenen Schritt $h_{(t-1)}$ als Eingabe erhalten, muss der BPTT-Algorithmus neben dem Gradienten der Fehlerfunktion der Eingabewerte auch den Gradienten der Fehlerfunktion der Ausgaben der Neuronen aus den vorangegangenen Schritten berechnen. Bei der Verarbeitung langer Sequenzen muss das an der Ausgabeschicht berechnete Fehlersignal daher sehr viele Schichten durchlaufen. Dabei tritt häufig die Problematik auf, dass der Gradient zu den niedrigen Schichten hin immer kleiner wird und verschwindet oder explosionsartig anwächst. Diese auch als verschwindender und explodierender Gradient bezeichneten Problematiken führen dazu, dass das Netz in den tieferen Schichten entweder nicht weiter konvergiert oder das Training instabil wird (Bengio et al. 1994). Daher wird diesen Netzen auch ein „Kurzzeitgedächtnis“ nachgesagt, welches zu der Problematik führt, dass längerfristige Abhängigkeiten nicht oder kaum berücksichtigt werden. Um dieser Problematik zu begegnen, haben Hochreiter und Schmidhuber (1997) die sogenannten *Long Short-Term Memory* (LSTM) Netze entwickelt. Dabei kommen anstelle von RNN-Zellen sogenannte LSTM-Zellen in den verdeckten Schichten eines RNN zum Einsatz. Abbildung 3-7 stellt eine RNN-Zelle einer LSTM-Zelle gegenüber.

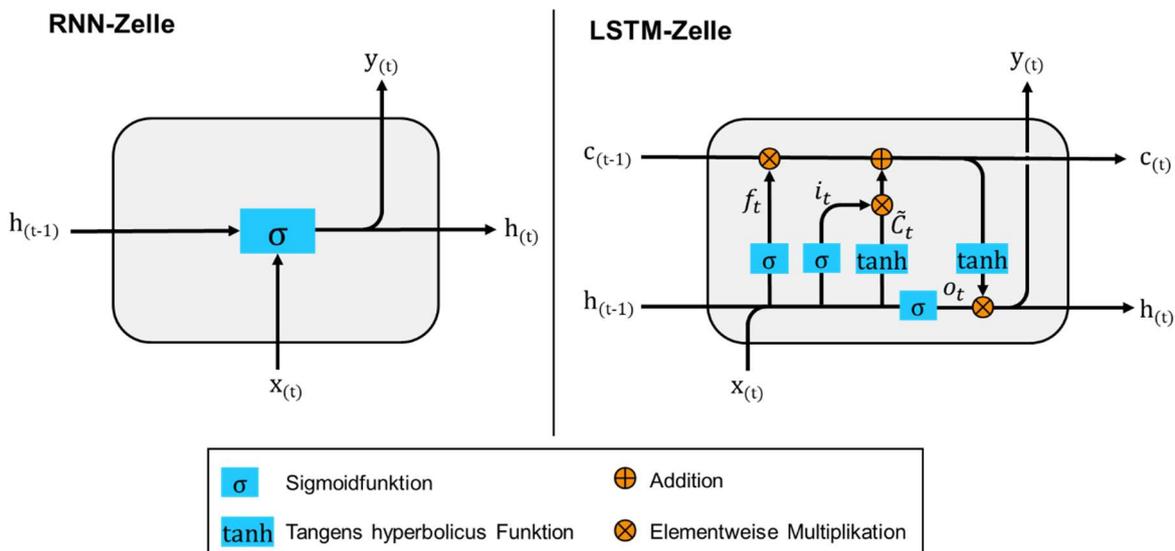


Abbildung 3-7: Vergleich einer RNN- und einer LSTM-Zelle (verändert nach Christopher Olah, 2015 und Raschka und Mirjalili, 2018)

Im Wesentlichen unterscheidet sich eine LSTM-Zelle durch folgende Eigenschaften von der RNN-Zelle:

- Es wird der Zellzustand c als weitere Größe ergänzt

- Es werden sogenannte Tore (*Gates*) eingeführt, über die der Informationsfluss innerhalb einer Zelle gesteuert wird

Der Zellzustand $c_{(t)}$ ermöglicht es der LSTM-Zelle, auch längerfristige Abhängigkeiten mitzuführen. Dabei ist $c_{(t)}$ abhängig von dem Zellzustand $c_{(t-1)}$ der innerhalb der LSTM-Zelle in Abhängigkeit der aktuellen Eingabe $x_{(t)}$ und der rekurrenten Eingabe $h_{(t-1)}$ angepasst wird. Dieser Vorgang wird über das *Forget-Gate* f_t und das *Input-Gate* i_t gesteuert. Das durch Gers et al. (2000) nachträglich ergänzte *Forget-Gate* f_t entscheidet darüber, welche Informationen des Zellstands $c_{(t-1)}$ gelöscht werden. Das *Input-Gate* i_t kontrolliert hingegen, welche Informationen von der aktuellen Zelleingabe \tilde{c}_t dem Zellzustand hinzugefügt werden. Die Zelleingabe \tilde{c}_t ist mit der Ausgabe $h_{(t)}$ eines RNNs vergleichbar (vgl. Formel (3-9)). Ein drittes Gate, das *Output-Gate* o_t , entscheidet dann darüber, wie der Zellzustand die Ausgabe $h_{(t)}$ beeinflusst, die wiederum die rekurrente Eingabe für den nächsten Schritt der Sequenz darstellt. Alle drei Gates verfügen über eine Sigmoid-Aktivierungsfunktion σ , wodurch die Ausgaben zwischen 0 und 1 liegen. Durch die anschließende Verwendung der Ausgaben in einer elementweisen Multiplikation \circ (siehe Formel 3-14) wird dann gesteuert, wie stark das mit der Gate-Ausgabe multiplizierte Signal weitergeleitet wird. Mathematisch lassen sich die Vorgänge mit den folgenden Formeln beschreiben:

$$f_t = \sigma(W_{xf}x_{(t)} + W_{hf}h_{(t-1)} + b_f) \quad (3-10)$$

$$i_t = \sigma(W_{xi}x_{(t)} + W_{hi}h_{(t-1)} + b_i) \quad (3-11)$$

$$\tilde{c}_t = \tanh(W_{x\tilde{c}}x_{(t)} + W_{h\tilde{c}}h_{(t-1)} + b_{\tilde{c}}) \quad (3-12)$$

$$o_t = \sigma(W_{xo}x_{(t)} + W_{ho}h_{(t-1)} + b_o) \quad (3-13)$$

$$c_{(t)} = (c_{(t-1)} \circ f_t) + (i_t \circ \tilde{c}_t) \quad (3-14)$$

$$h_{(t)} = y_{(t)} = o_t \circ \tanh(c_{(t)}) \quad (3-15)$$

Neben LSTM-Zellen haben sich auch die durch Cho et al. (2014) vorgestellten *Gated Recurrent Units* (GRU) als deutlich effizienterer Ersatz für die RNN-Zellen erwiesen. Das Grundprinzip ist ähnlich dem der LSTM-Zellen, jedoch sind GRU-Zellen etwas einfacher aufgebaut und verfügen über weniger trainierbare Parameter. Das hat den Vorteil, dass sich der Trainingsprozess schneller gestaltet, kann aber gleichzeitig auch zu einer geringeren Leistungsfähigkeit führen.

3.3.2.2 Konvolutionelle Neuronale Netze (*Convolutional Neural Networks*, CNN)

Ein CNN ist eine maßgeblich durch die Arbeit von LeCun et al. (1989) entwickelte Netzarchitektur, die sich insbesondere in der Bilderkennung als besonders effektiv erwiesen hat. CNNs bestehen meist aus mehreren Faltungsschichten (*Convolutional Layer*) und sogenannten *Pooling*-Schichten, denen abschließend wie beim MLP vollständige verknüpfte Schichten folgen.

Im Rahmen der Bildverarbeitung betrachtet eine Faltungsschicht das Eingabebild als eine 2D-Ebene. Die Faltungsschicht besteht aus einem kleinen Sichtfeld von $n \times n$ Pixeln, das auch

als rezeptives Feld bezeichnet wird (Raschka und Mirjalili 2018) und einem sogenannten Filter (häufig auch als Kernel bezeichnet) mit der gleichen Größe, der die erlernbaren Parameter enthält. Bei der Verarbeitung eines Bildes wird das Sichtfeld mit einer vordefinierten Schrittweite über das Eingabebild verschoben und die Summe des elementweisen Produkts der Werte der Eingabepixel und der Filterwerte gebildet. Das Ergebnis wird für jede Position als einzelner Pixel auf einer neuen Merkmalskarte abgebildet, die gleichzeitig die Ausgabe der Schicht darstellt (vgl. Abbildung 3-8).

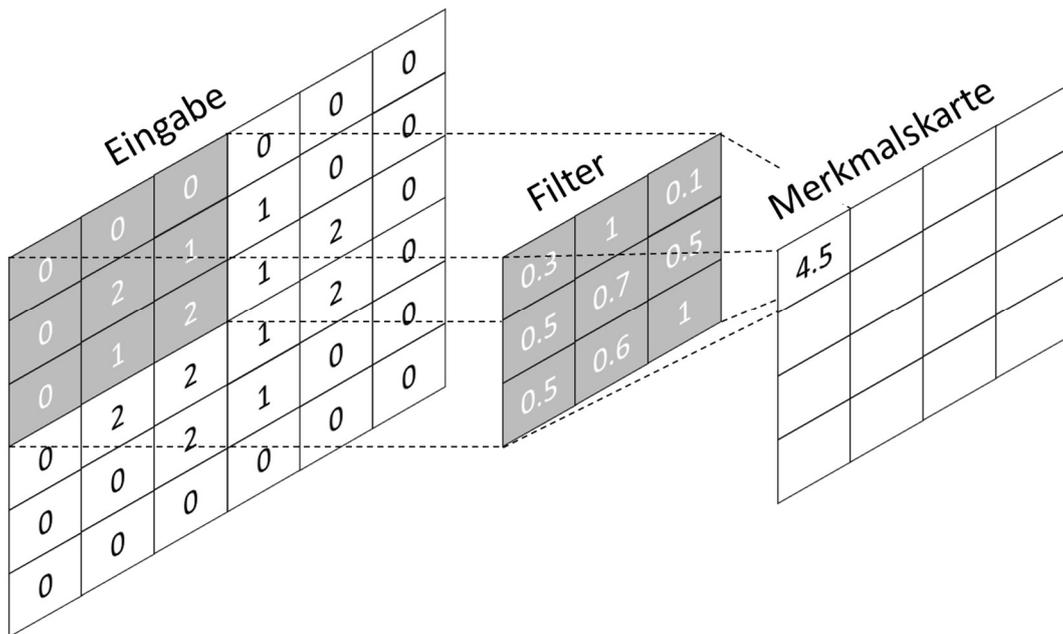


Abbildung 3-8: Schematische Darstellung einer 2D-Faltung in einer Faltungsschicht eines CNNs

Mathematisch lässt sich die Faltungsoperation für ein zweidimensionales Bild I nach Goodfellow et al. (2016) wie folgt beschreiben:

$$S_{i,j} = \sum_m \sum_n I_{(i+m,j+n)} \cdot K_{(m,n)} \quad (3-16)$$

$S_{(i,j)}$ beschreibt hier die Faltungsoperation für den Pixel der i -ten Reihe und der j -ten Spalte einer Merkmalskarte. Die Indices m und n geben Reihe und Spalte des Filters K an, über die die Summe des elementweisen Produkts zwischen Eingabewert und Filterwert gebildet wird.

Das Besondere gegenüber einem MLP ist, dass der Filter an den verschiedenen Stellen des Bildes die gleiche Gewichtsmatrix nutzt. Dementsprechend wird die Anzahl der zu erlernenden Parameter drastisch reduziert. Zudem konzentrieren sich CNNs dadurch auf die Erkennung besonders relevanter Merkmale und sind in der Lage, diese an verschiedenen Orten in einem Bild zu erkennen. Um verschiedene Merkmale in einem Bild zu erkennen, kann jede Faltungsschicht mehrere Filter beinhalten, die jeweils über eine eigene Gewichtsmatrix verfügen und eine eigene Merkmalskarte ausgeben.

In einem CNN können auch mehrere Faltungsschichten gestapelt werden. Die jeweils nachfolgende Faltungsschicht greift dann auf die Merkmalskarte der vorherigen Schicht zurück und

führt die gleichen Operationen durch. Um bei CNNs die Rechenlast weiter zu reduzieren, können zwischen einzelnen Faltungsschichten sogenannte *Pooling*-Schichten angeordnet werden. Diese funktionieren ähnlich wie die Faltungsschichten, nur dass diese über keine erlernbaren Parameter verfügen. Vielmehr werden alle in einem Sichtfeld liegenden Pixel entweder gemittelt oder jeweils nur der Maximalwert ausgewählt und auf einer Merkmalskarte abgebildet. Die nachfolgende Faltungsschicht muss dann nur auf eine kleinere Merkmalskarte zurückgreifen, dementsprechend müssen im Trainingsprozess auch weniger Parameter angepasst werden. Darüber hinaus bietet das Pooling weitere Vorteile, wie beispielsweise, dass die Netze robuster gegenüber einer Überanpassung oder einem Rauschen in den Eingangsdaten werden.

Abbildung 3-9 zeigt einen üblichen Aufbau eines mehrschichtigen CNNs zur Bildklassifizierung. Die Dimension der dargestellten Faltungsschichten setzen sich aus der Kernelgröße (hier 5×5 Pixel) sowie der Anzahl an Filtern zusammen. Bei den *Pooling*-Schichten stehen die Dimensionen hingegen für die Größe des Sichtfeldes, dessen Pixel zu einem einzelnen Pixel in der nachfolgenden Merkmalskarte zusammengefasst werden. In Abhängigkeit der Größe des Sichtfeldes reduziert sich dementsprechend die Größe der folgenden Merkmalskarten. Die Ausgabe der letzten *Pooling*-Schicht wird zunächst in eine andere Dimension umgeformt, um anschließend von vollständig verbundenen Schichten verarbeitet werden zu können, die dann in Abhängigkeit der herausgearbeiteten Merkmale die Klassifizierung durchführen.

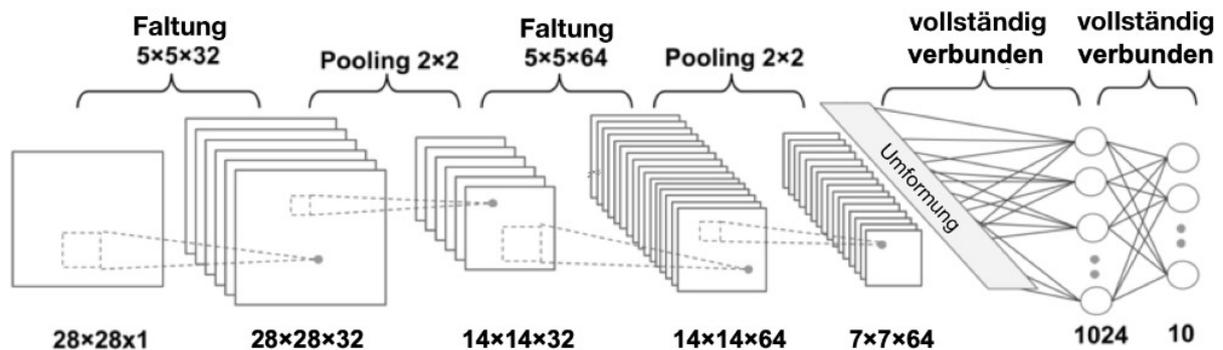


Abbildung 3-9: Genereller Aufbau eines CNNs mit einer Abwechselnden Anordnung von Faltungs- und Poolingschichten, sowie zwei vollständig verbundenen Schichten, die die Netzausgabe erzeugen (Raschka und Mirjalili 2018).

Neben der Bilderkennung eignen sich CNNs unter anderem auch für die Verarbeitung von Zeitreihen. Dabei werden Zeitreihen als ein eindimensionales Bild betrachtet und können so durch CNNs verarbeitet werden. Auch hier bieten CNNs die Möglichkeit, lokale Muster zu erlernen und diese dann an verschiedenen Stellen der Zeitreihe zu erkennen.

3.3.2.3 Graph Convolutional Networks (GCN)

Die zuvor beschriebenen Netzarchitekturen eignen sich für die Verarbeitung von kategorischen, sequentiellen oder als Raster strukturierten Daten. Viele Datensätze liegen allerdings auch als Netzwerke in Form von Graphen vor, bei denen die Verbindung einzelner Objekte zu anderen Objekten eine wichtige Rolle spielt. Hierzu zählen beispielsweise soziale Netzwerke,

Moleküle, Verkehrsnetze oder eben auch Kanalnetze. Für diese Art von Daten wurde eine spezielle Art von Neuronalen Netzen entwickelt, die sogenannten *Graph Neural Networks* (GNN). Diese Architektur wurde bereits vor über einer Dekade von Scarselli et al. (2009) vorgestellt und insbesondere durch die Arbeiten von Defferrard et al. (2016) sowie Kipf und Welling (2016) populär, die GNNs mit CNNs zu *Graph Convolutional Networks* (GCN) kombiniert haben. Mit dieser Architektur konnte die Leistungsfähigkeit von Neuronalen Netzen für viele Problemstellungen optimiert werden. Dazu zählen unter anderem Verkehrsprognosen (Yu et al. 2017), Prognosen des Drucks in Trinkwasserversorgungsnetzen (Hajgató et al. 2021) oder die Prognose des COVID-19-Infektionsgeschehens (Fritz et al. 2022).

Die Grundlage von GNNs stellen Graphen G dar, die sich in der einfachsten Form als $G = (V; E)$ darstellen lassen. Wobei V für die Knoten und E für die Kanten steht. Eine Kante vom Knoten $v_i \in V$ zum Knoten $v_j \in V$ lässt sich dabei als $(v_i, v_j) \in E$ beschreiben. Für die effiziente Verarbeitung von Graphen in ML-Anwendungen werden diese in der Regel als Matrix abgebildet. Eine Möglichkeit dazu stellen Adjazenzmatrizen (*adjacency matrices*) $A \in \mathbb{R}^{N \times N}$ dar, die aus einer $N \times N$ Matrix bestehen in der für jede Position i, j ($1 \leq i \leq N$; $1 \leq j \leq N$) gilt:

$$a_{i,j} = \begin{cases} 1 & \text{falls } (v_i, v_j) \in E, \\ 0 & \text{sonst.} \end{cases} \quad (3-17)$$

Weiterhin lassen sich Graphen in gerichtet und ungerichtet unterscheiden. Während bei gerichteten Graphen die Kanten zwischen zwei Knoten nur in eine Richtung abgebildet werden, wird bei ungerichteten Graphen die Verbindung in beide Richtungen dargestellt. Die Adjazenzmatrix ist dementsprechend bei letzteren symmetrisch. Darüber hinaus können die Kanten in Graphen gewichtet sein, wodurch die Einträge in der Adjazenzmatrix zusätzlich mit einem gegebenen Gewicht multipliziert werden. Die nachfolgende Tabelle 3-3 verdeutlicht die beschriebenen Zusammenhänge.

Tabelle 3-3: Übersicht über unterschiedliche Arten von Graphen sowie den zugehörigen Adjazenzmatrizen (die Zeilen und Spalten sind hier jeweils alphabetisch geordnet)

| | Ungewichtet | Gewichtet |
|-------------|---|---|
| Ungerichtet | $\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & 10 & 0 & 0 & 0 \\ 10 & 0 & 20 & 0 & 50 \\ 0 & 20 & 0 & 10 & 0 \\ 0 & 0 & 10 & 0 & 0 \\ 0 & 50 & 0 & 0 & 0 \end{pmatrix}$ |
| Gerichtet | $\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 10 & 0 & 0 & 0 & 0 \\ 0 & 20 & 0 & 0 & 0 \\ 0 & 0 & 10 & 0 & 0 \\ 0 & 50 & 0 & 0 & 0 \end{pmatrix}$ |

GNNs können eingesetzt werden, um Merkmale auf Ebene der Knoten, Kanten oder des ganzen Graphs zu prognostizieren. Weit verbreitet ist dabei die Vorhersage auf Ebene der Knoten, bei der neben der Adjazenzmatrix eine Merkmalsmatrix (*feature matrix*) $X \in \mathbb{R}^{N \times D}$ als Eingabe berücksichtigt wird. Dabei beschreibt N die Anzahl der Knoten und D die Anzahl der Eingabemerkmale je Knoten. Die Ausgabe an den Knoten $Z \in \mathbb{R}^{N \times F}$ (F steht hier für die Anzahl der Ausgabemerkmale) ist dem entsprechend eine Funktion f der Adjazenz- und Merkmalsmatrix:

$$Z = f(A; X) \quad (3-18)$$

Die Berechnung von Z kann mit verschiedenen GNN-Architekturen erfolgen. Viele Untersuchungen greifen dabei auf die von Kipf und Welling (2016) beschriebenen *Graph Convolutional Networks* zurück (Wu et al. 2019). GCNs übertragen die von CNNs bekannte Faltungsoperation von Raster- auf Graphdaten. Die wesentliche Idee dahinter ist, dass die Repräsentation eines Knotens immer von den vorherigen eigenen Merkmalen und den Merkmalen der Nachbarknoten abhängig ist. Die Funktionsweise eines GCNs basiert dabei auf den folgenden wesentlichen Schritten:

- Nachbarschaftsaggregation: Führt für jeden Knoten die Informationen der Nachbarknoten zusammen, indem beispielsweise die Summe, der Mittelwert oder der Maximalwert aller Nachbarknoten gebildet wird.
- Aktualisierung der Knotenrepräsentation: Aktualisiert die Werte der Knoten, indem die aggregierten Informationen der Nachbarknoten mit den aktuellen Werten der Knoten kombiniert werden.

Dieser Vorgang ähnelt dem der Faltungsoperation in einem CNN und ist exemplarisch in Abbildung 3-10 dargestellt.

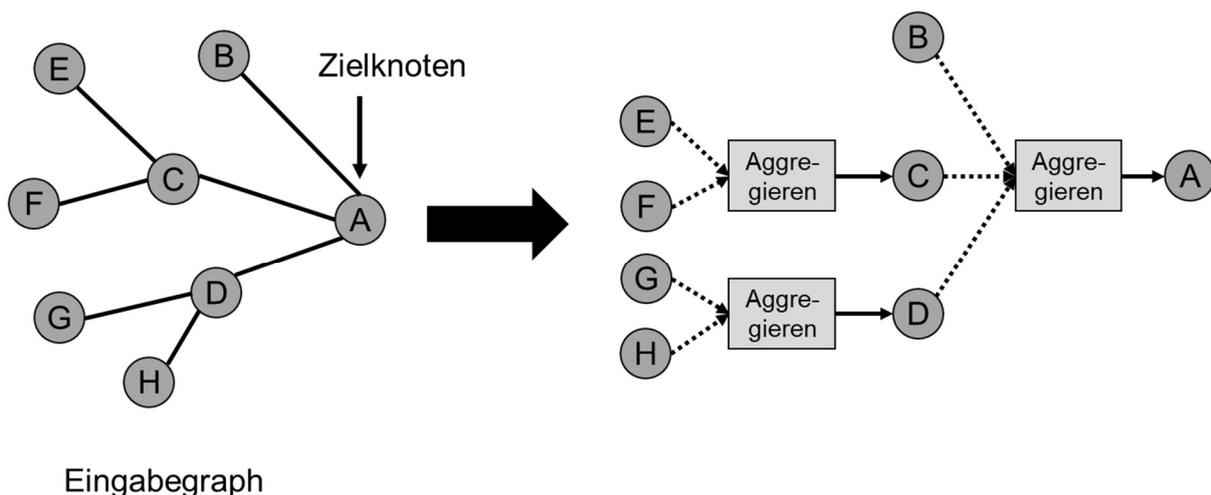


Abbildung 3-10: Aggregation der Informationen von den Nachbarknoten in einem GCN: Der dargestellte Zielknoten A aggregiert die Informationen der direkten Nachbarknoten B, C und D. Die Informationen von B, C und D basieren wiederum auf den aggregierten Informationen aus deren Nachbarschaften und so weiter (verändert nach Hamilton et al., 2017)

Bei vielen Problemstellungen variieren die Werte der Verbindungen oder Knoten des Graphens zusätzlich im Laufe der Zeit. Für diese Art von Problemen wurden *Temporal Graph*

Convolutional Networks (T-GCN) entwickelt, die beispielsweise im Bereich der Verkehrsprognose eingesetzt werden (Yu et al. 2017; Zhao et al. 2018) und die zusätzliche Berücksichtigung der zeitlichen Dynamik eines Graphens ermöglichen.

3.3.2.4 Generative Adversarial Networks (GAN)

Generative Adversarial Networks (GAN) ist eine erstmals in Goodfellow et al. (2014) vorgestellte Architektur, bestehend aus zwei Teilmodellen (vgl. Abbildung 3-11). Zum einen ein Generator G , der auf Basis einer Zufallsverteilung z Datensätze generiert, die denen im Trainingsdatensatz ähneln. Zum anderen ein als Diskriminator D bezeichneter Klassifikator, der die Wahrscheinlichkeit berechnet, ob ein Datenbeispiel x aus dem Trainingsdatensatz oder vom Generator stammt. Beide Modelle haben entgegengesetzte Ziele und treten beim Training in einem Nullsummenspiel gegeneinander an. Ziel des Generators ist es, Ausgaben zu erzeugen, bei denen der Diskriminator nicht unterscheiden kann, ob es sich um „reale“ Inhalte aus einem Datensatz handelt oder durch den Generator erzeugte „falsche“ Ausgaben. Der Diskriminator zielt dagegen darauf ab, die Ausgaben des Generators mit einer möglichst hohen Wahrscheinlichkeit als „falsche“ Inhalte zu klassifizieren.

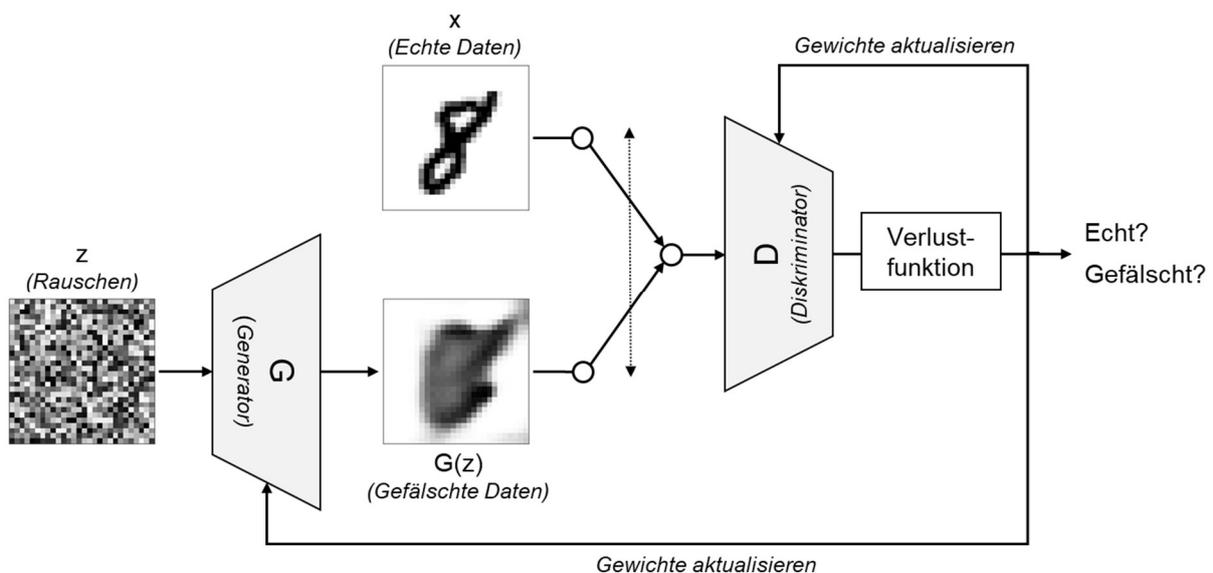


Abbildung 3-11: Exemplarischer Aufbau eines Generative Adversarial Networks. Zur Generierung der Abbildung wurde auf den MNIST-Datensatz (LeCun et al. 2010) zurückgegriffen

Beide Modelle werden gleichzeitig trainiert und nutzen die gleiche Verlustfunktion L , die die Wahrscheinlichkeit angibt, ob ein Eingabedatensatz „echt“ oder „gefälscht“ ist. Während die Parameter des Diskriminators so angepasst werden, dass diese Wahrscheinlichkeit maximiert wird, werden die Parameter des Generators mit dem Ziel der Minimierung angepasst. Damit ergibt sich die folgende in Goodfellow et al. (2014) dargestellte Verlustfunktion:

$$\min_G \max_D L(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (3-19)$$

Eines der Hauptanwendungsgebiete von GANs ist die Generierung von neuen Inhalten wie beispielsweise Bildern (Radford et al. 2015), Text (Yu et al. 2016) oder Audiosignalen (Donahue et al. 2018). Generator und Diskriminator können beliebige Architekturen enthalten, die

für das jeweilige Lernproblem geeignet sind. Da ein Hauptanwendungsgebiet von GANs die Bildverarbeitung ist, basieren viele Modelle auf Architekturen aus CNN-Schichten, dies ist aber keine Voraussetzung.

Bei den klassischen GAN-Modellen, die lediglich eine Zufallsverteilung als Eingabe erhalten, gibt es keine Möglichkeit, die Art und Weise wie Daten erzeugt werden zu kontrollieren (Mirza und Osindero 2014). Vor diesem Hintergrund haben Mirza und Osindero (2014) mit *conditional* GANs (cGANs) eine abgeänderte Variante entwickelt, die neben einer Zufallsverteilung z auch von zusätzlichen Informationen y abhängig ist. Dadurch erhält das Modell einen gewissen Kontext, mit dem sich die Ausgabe beeinflussen lässt. Damit lassen sich beispielsweise aus Konturen oder Labeln Bilder generieren (Isola et al. 2017) oder eben auch aus einer Niederschlagsvorhersage die resultierende Überflutungssituation (Hofmann und Schüttrumpf 2021).

3.3.2.5 Transformer

Transformer stellen eine weitere spezielle Architektur von Künstlich Neuronalen Netzen dar, die für die Verarbeitung von Sequenzen entwickelt wurde und hier speziell im Bereich der Sprachverarbeitung (*Natural Language Processing*, NLP) für Aufsehen gesorgt hat. So greifen verschiedene Sprachmodelle wie beispielsweise das von Open AI entwickelte *Generative Pre-trained Transformer* (GPT)-Modell auf diese Architektur zurück (Radford et al. 2018). Erstmals vorgestellt wurde die Architektur in Vaswani et al. (2017) mit dem in Abbildung 3-12 dargestellten Aufbau. Das Besondere an dieser Architektur ist der Einsatz sogenannter Aufmerksamkeitsmechanismen (*attention-mechanism*). Die Kernidee von Aufmerksamkeitsmechanismen liegt darin, sich auf die relevantesten Merkmale der Eingabedaten zu konzentrieren. Dies geschieht, indem jedem Wert (z. B. ein Wort) in einer Eingabesequenz Gewichtungen in Abhängigkeit von dessen Relevanz für eine Ausgabe in der Zielsequenz zugewiesen werden. Durch die Tatsache, dass das Netz für jede Ausgabe auf alle Punkte der Eingabesequenz zugreift, kann dem Effekt des Kurzzeitgedächtnisses bei rekurrenten Netzen entgegengewirkt werden. Dadurch sind *Transformer* in der Lage, besonders lange Abhängigkeiten zu erlernen. Der Kernprozess von Aufmerksamkeitsmechanismen ähnelt dem der Matrixmultiplikationen bei einer einzelnen Neuronenschicht in einem MLP. Im Detail gibt es leichte Ergänzungen, die auch in Abhängigkeit des verwendeten Aufmerksamkeitsmechanismus variieren können. Verbreitete Mechanismen sind beispielsweise die in Bahdanau et al. (2014) oder Luong et al. (2015) vorgestellten Varianten sowie die in der *Transformer*-Architektur verwendete *Multi-Head Attention* (Vaswani et al. 2017).

Neben Aufmerksamkeitsmechanismen basiert die in Abbildung 3-12 dargestellte *Transformer*-Architektur auf einer *Encoder-Decoder*-Struktur. Bei dieser im Bereich des *Deep Learnings* und insbesondere der Verarbeitung von Sequenzen sehr weit verbreiteten Grundstruktur lässt sich das Netz in zwei Teile untergliedern: zum einen dem Encoder, der aus den Eingabedaten relevante Informationen extrahiert und zum anderen dem Decoder, der aus den extrahierten Informationen die Zielgröße generiert. Gegenüber den meisten anderen Architekturen zur Verarbeitung von Sequenzen greift die ursprüngliche *Transformer*-Architektur nicht auf rekurrente Schichten zurück. Stattdessen werden vollständig verbundene vorwärtsgerichtete Schichten

verwendet, wie sie auch im MLP zur Anwendung kommen. Dadurch verläuft das Training deutlich schneller, stabiler und lässt sich zudem besser parallelisieren. Um dennoch die Reihenfolge einer Sequenz zu berücksichtigen, werden die Positionen der Eingaben mit Hilfe sogenannter *Positional Encodings* kodiert (vgl. Géron 2019).

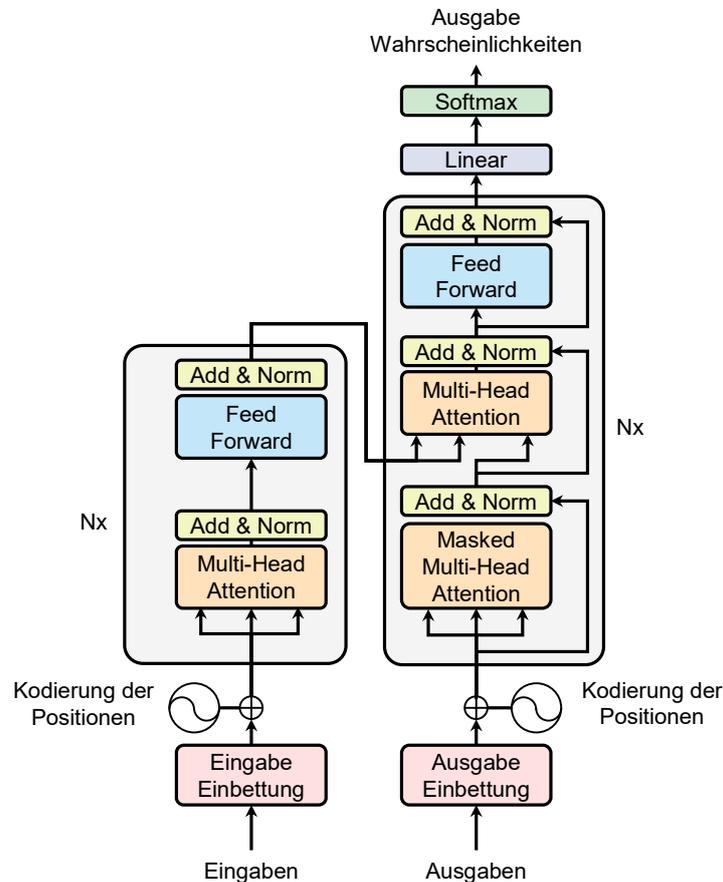


Abbildung 3-12: Transformer-Architektur mit Aufmerksamkeitsmechanismus in Form von Multi-Head Attention-Schichten (Vaswani et al. 2017)

Aufgrund ihrer Eignung zur Modellierung von Sequenzen liegt neben dem Bereich NLP auch die Anwendung von *Transformern* im Bereich der Zeitreihenanalyse nahe. Eine umfassende Übersicht über die verschiedenen Anwendungsgebiete und Studien liefern Wen et al. (2023). Im Bereich der Zeitreihenvorhersage haben sich die durch Lim et al. (2019) entwickelten *Temporal Fusion Transformer* (TFT) als besonders effizient erwiesen (siehe Abbildung 3-13). TFT zeichnen sich durch ihre Flexibilität hinsichtlich der zu berücksichtigenden Eingabedaten aus, wodurch sie nicht nur bekannte Eingaben vergangener und zukünftiger Zeitschritte, sondern auch statische Variablen berücksichtigen können. Dadurch eignen sich TFT für eine Vielzahl von Problemstellungen.

Wie die ursprüngliche *Transformer*-Architektur nutzen auch TFT eine *Encoder-Decoder*-Struktur, greifen jedoch auf rekurrente LSTM-Schichten anstelle von MLP-Schichten zurück. Die Kombination der auf LSTM-Zellen basierten *Encoder-Decoder*-Struktur mit der im originalen *Transformer* verwendeten *Multi-Head Attention* ermöglicht es TFT, sowohl langfristige als auch kurzfristige zeitliche Abhängigkeiten zu erlernen. Ein weiterer Bestandteil der TFT-Archi-

tektur sind *Gated Residual Network* (GRN)-Blöcke, die sich an verschiedenen Stellen der Gesamtarchitektur wiederfinden. Diese Blöcke ermöglichen das Überspringen ungenutzter Komponenten, was es dem Netz ermöglicht, seine Tiefe und Komplexität an unterschiedliche Aufgaben und Datensätze anzupassen. Zusätzlich umfassen TFTs *Variable Selection Networks* (VSN)-Blöcke, die Netzeingaben entgegennehmen und die wesentlichen Eingangsgrößen identifizieren, während weniger relevante Eingaben ignoriert werden. Dazu werden, in Abhängigkeit des Einflusses der Eingabemerkmale auf die Zielgröße, Wichtungen berechnet, die zusätzlich eine gewisse Interpretierbarkeit des Modells zulassen.

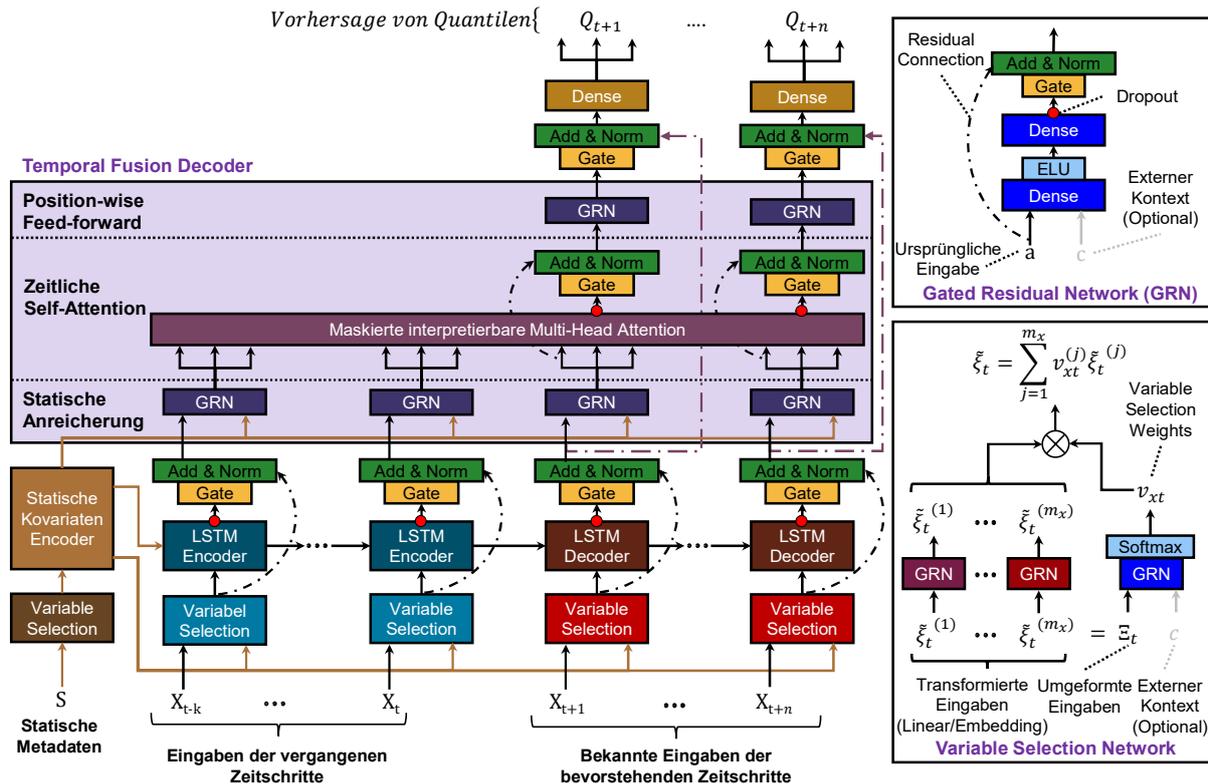


Abbildung 3-13: Darstellung der *Temporal Fusion Transformer*-Architektur (links) sowie Detailansicht der *Gated Residual Network*-Blöcke und der *Variable Selection Network*-Blöcke (Lim et al. 2019)

Der in Lim et al. (2019) vorgestellte TFT wird mit der in Wen et al. (2017) beschriebenen quantilen Verlustfunktion trainiert. Dadurch erfolgt die Vorhersage auf Basis von Intervallen, die den Bereich der wahrscheinlichen Zielwerte für einen Prognosezeitpunkt angeben. Die Begrenzung der Intervalle erfolgt durch Angabe von zu berücksichtigenden Quantilen beim Training. Durch die Prognose von Quantilen können Modellunsicherheiten bei der Prognose besser berücksichtigt werden.

3.4 Verwendete Software und Berechnungsansätze

3.4.1 Physikalisch-basierte Überflutungsmodellierung

Für die Erzeugung der künstlichen Trainingsdatensätze werden die in der Modellierungssoftware MIKE+ (DHI 2021b) integrierten hydrodynamischen Berechnungsmodelle MIKE 1D und

MIKE 21 verwendet. Während mittels MIKE 1D die 1D-Abflussberechnung im Kanalnetz erfolgt, bietet MIKE 21 die Möglichkeit, das 2D-Abflussverhalten auf der Geländeoberfläche zu simulieren. Beide Modelle können über verschiedene Ansätze gekoppelt werden, sodass ein bidirektionaler Austausch zwischen Oberfläche und Kanalnetz möglich ist. Nachfolgend werden die einzelnen Modelle sowie die Kopplung beider Modelle zu einem bidirektional gekoppelten 1D/2D-Abflussmodell beschrieben. Die nachfolgenden Beschreibungen nehmen, wenn nicht anders gekennzeichnet, Bezug auf die durch das Danish Hydraulic Institute (DHI) herausgegebenen wissenschaftlichen Dokumentationen zu MIKE 1D (DHI 2017a) und MIKE 21 (DHI 2021a).

Berechnung des Abflusses im Kanalnetz

Zur Belastung des Kanalnetzes stehen verschiedene hydrologische Oberflächenabflussmodelle zur Auswahl, von denen sich drei für den Einsatz in der Kanalnetzmodellierung eignen. Tabelle 3-4 liefert eine Übersicht über die jeweils verwendeten Ansätze und Methoden zur modelltechnischen Abbildung des Abflussbildungs- und des Abflusskonzentrationsprozesses. Alle drei Modellansätze bieten die Möglichkeit, flächenspezifische Parametersätze festzulegen und so unterschiedliche Oberflächenbeschaffenheiten zu berücksichtigen. Für die detaillierte Erläuterung der Funktionsweise und der mathematischen Beschreibung der aufgeführten Ansätze und Methoden sei an dieser Stelle auf die wissenschaftliche Dokumentation verwiesen (DHI 2017a).

Der für die hydrodynamische Kanalnetzmodellierung in MIKE+ integrierte Rechenkern MIKE 1D ermöglicht die Berechnung von Abflussvorgängen in Kanalnetzen unter instationären Bedingungen. Die Berechnung des Abflustransportes erfolgt dabei auf Basis der Saint-Venant-Gleichungen, die sich durch verschiedene Vereinfachungen aus den Navier-Stokes-Gleichungen ableiten lassen (Price und Vojinovic 2011). Diese bestehen in allgemeingültiger Form aus der Kontinuitätsgleichung

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = q \quad (3-20)$$

und der eindimensionalen Bewegungsgleichung in Fließrichtung (DWA 2006b):

$$\frac{1}{g} \cdot \frac{\partial v}{\partial t} + \frac{v}{g} \cdot \frac{\partial v}{\partial x} + m \cdot \frac{v \cdot q}{g \cdot A} + \frac{\partial h}{\partial x} = I_S - I_R \quad (3-21)$$

mit:

| | | |
|-----|---|---------------------|
| Q | Durchfluss | [m ³ /s] |
| x | Wegkoordinate in Fließrichtung | [m] |
| A | Fließquerschnitt | [m ²] |
| t | Zeitkoordinate | [s] |
| g | Erdbeschleunigung | [m/s ²] |
| v | Mittlere Geschwindigkeit in Fließrichtung | [m/s] |

| | | |
|-------|--|-------------------------|
| m | Faktor unter Einschluss der Zusatzverluste | [-] |
| q | Seitlicher Zufluss je Längeneinheit in Fließrichtung | [m ³ /(s·m)] |
| h | Füllhöhe im Profil bzw. Wassertiefe (normal zur Sohle) | [m] |
| I_S | Sohlgefälle | [-] |
| I_R | Reibungsgefälle | [-] |

Die numerische Lösung der Saint-Venant-Gleichungen erfolgt in MIKE 1D implizit und mit Hilfe der Finite-Differenzen-Methode. Da die Saint-Venant-Gleichungen nur für Freispiegelabfluss gültig sind, sich bei Vollenfüllung der Rohrleitung in einem Kanalnetz aber Druckabfluss einstellt, wird in MIKE 1D das Konzept des *Preissmann Slots* eingeführt. Dabei handelt es sich um einen fiktiven Schlitz mit geringer Breite im Scheitel der Haltungen. Kommt es zur Vollenfüllung, kann der Wasserstand in diesem Schlitz weiter steigen und erzeugt einen hydrostatischen Druck. Dadurch lässt sich auch Abfluss unter Druck simulieren.

Tabelle 3-4: Übersicht über die hydrologischen Modelle zur Oberflächenabflussberechnung in MIKE+ (verändert nach DHI (2015))

| Bezeichnung | | Modell A | Modell B | Modell C | |
|-----------------------------|------------------|--------------------------|-----------------------|--------------------|-------------------------|
| | | | | C1 | C2 |
| Abflussbildung | | | | | |
| Verlustart | Kategorie | | | | |
| Benetzung | einmalig | Anfangs- verlust | Benetzung | Anfangsverlust | |
| Interzeption | einmalig | | Speicher- modell | | |
| Muldenverlust | einmalig | | | | |
| Infiltration | kontinuierlich | Reduzierungs- faktor | Ansatz nach Horton | Ansatz nach Horton | |
| Evapo- transpiration | kontinuierlich | | N/A | N/A | Reduzierungs- faktor |
| Abflusskonzentration | | | | | |
| Routing-Methode | | Zeitflächen- diagramm | Kinematische Welle | Linearspeicher | |

Berechnung des Oberflächenabflusses

Die hydrodynamische Oberflächenabflussmodellierung mit MIKE 21 erfolgt auf Basis der 2D-Flachwassergleichungen (auch als 2D-Saint-Venant-Gleichungen bekannt). Diese stellen eine Reduzierung der Navier-Stokes-Gleichungen und der Kontinuitätsgleichung auf ein ebenes Strömungsproblem mit freier Oberfläche dar (Aigner und Bollrich 2015). Dabei wird die Fließgeschwindigkeit in vertikaler Richtung vernachlässigt und in horizontaler Richtung als Mittelwert über die Wassertiefe definiert. Die 2D-Flachwassergleichungen in der in Aigner und Bollrich (2015) dargestellten Form setzen sich wie folgt aus der Kontinuitätsgleichung,

$$\frac{\partial h}{\partial t} + \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} = 0 \quad (3-22)$$

der Flachwassergleichung in x-Richtung

$$\frac{\partial q_x}{\partial t} + \frac{\partial}{\partial x} \left(\frac{q_x^2}{h} + g \cdot \frac{h^2}{2} - h \cdot v_T \cdot \frac{\partial v_x}{\partial x} \right) + \frac{\partial}{\partial y} \left(\frac{q_x \cdot q_y}{h} - h \cdot v_T \cdot \frac{\partial v_x}{\partial y} \right) + g \cdot h \cdot (I_{Rx} - I_{Sx}) = 0 \quad (3-23)$$

und der Flachwassergleichung in y-Richtung zusammen:

$$\frac{\partial q_y}{\partial t} + \frac{\partial}{\partial x} \left(\frac{q_x \cdot q_y}{h} - h \cdot v_T \cdot \frac{\partial v_y}{\partial x} \right) + \frac{\partial}{\partial y} \left(\frac{q_y^2}{h} + g \cdot \frac{h^2}{2} - h \cdot v_T \cdot \frac{\partial v_y}{\partial y} \right) + g \cdot h \cdot (I_{Ry} - I_{Sy}) = 0 \quad (3-24)$$

mit:

| | | |
|-------|--|---------------------|
| q_x | Spezifischer Abfluss in x Richtung ($v_x \cdot h$) | [m ² /s] |
| q_y | Spezifischer Abfluss in y Richtung ($v_y \cdot h$) | [m] |
| v_T | Viskosität | [m ²] |
| I_S | Sohlgefälle | [-] |
| I_R | Reibungsgefälle | [-] |

Die Lösung der dargestellten partiellen Differentialgleichungen erfolgt in MIKE 21 numerisch mittels Finite-Volumen-Methode. Dies setzt neben einer zeitlichen auch eine räumliche Diskretisierung der Oberfläche voraus, die durch die Aufteilung der Oberfläche in ein Rechennetz erfolgt. In MIKE+ können Rechennetze rasterbasiert als „grid“ oder trianguliert als „flexible Mesh“ erstellt werden. Mittels Interpolation können die einzelnen Elemente des Netzes mit Höheninformationen aus einem DGM versehen werden. Daraus lässt sich dann das für die Berechnung erforderliche Sohlgefälle ermitteln. Für die Ermittlung des Reibungsgefälles sind dem Rechennetz ergänzend Rauigkeiten in Abhängigkeit der Oberflächenbeschaffenheit zuzuweisen. Zur Darstellung von abflussrelevanten Strukturen wie Gebäuden oder Straßenzügen gibt es zudem verschiedene Ansätze, um diese beispielsweise durch feste Höhenunterschiede in entsprechenden Bereichen zu berücksichtigen.

Die Belastung des Rechennetzes kann bei der hydrodynamischen Oberflächenabflussberechnung entweder direkt durch einen Effektivniederschlag erfolgen oder durch ein angesetztes Niederschlagsereignis in Kombination mit entsprechenden Verlustansätzen berechnet werden. Als Verluste können Evaporation und Infiltration durch konstante oder zeitlich variierende Spenden sowie, im Falle der Infiltration, durch Speichermodelle berücksichtigt werden. Anfangs- und Muldenverluste werden hingegen nicht gesondert berücksichtigt, da sich insbesondere Muldenverluste durch die Ansammlung von Wasser in Geländesenken bei der 2D-Modellierung automatisch ergeben. Als wesentliches Ergebnis einer MIKE 21-Berechnung können Wasserstände und Fließgeschwindigkeiten für jedes Oberflächenelement als Ganglinie oder als Maximalwert ausgegeben werden.

Modellkopplung und Abflussbildungsansatz

Die Berechnungsmodelle MIKE 1D und MIKE 21 lassen sich über verschiedene Ansätze bidirektional miteinander koppeln, um den Austausch von Wasser zwischen beiden Modellen zu

ermöglichen. Das bedeutet, dass sowohl der Eintritt von Oberflächenabfluss in das Kanalnetz als auch der Austritt von Wasser in Form von Überstau auf das Oberflächenmodell möglich ist. Die Kopplung folgt dabei seitens des Kanalnetzmodells über Knotenelemente wie Schächte, Ausläufe oder Straßeneinläufe. Die Knotenelemente können entweder mit dem direkt darüber befindlichen Element des Oberflächenmodells oder mit allen in einem definierten Quadrat liegenden Elementen des Rechnernetzes gekoppelt werden.

Die Abflussbildung und -konzentration können bei einer gekoppelten Berechnung sowohl über die hydrologischen Ansätze des Kanalnetzmodells als auch über das Oberflächenmodell erfolgen. Zudem gibt es die Möglichkeit, beide Ansätze zu kombinieren und den Abfluss für an das Kanalnetz angeschlossene Flächen über das Kanalnetzmodell zu generieren und für alle anderen Flächen das Oberflächenmodell direkt zu berechnen. Für eine detaillierte Beschreibung möglicher Varianten von gekoppelten Berechnungen sei an dieser Stelle auf den „Praxisleitfaden zur Ermittlung von Überflutungsgefahren mit vereinfachten und detaillierten hydrodynamischen Modellen“ (HSB 2017) verwiesen. Die dort aufgeführten Modellkombinationen lassen sich auch in MIKE+ erstellen.

3.4.2 Implementierung der ML-Verfahren und Datenvorverarbeitungsroutinen

Die ML-basierten Vorhersagemodelle sowie die dazu erforderlichen Datenvorverarbeitungsroutinen wurden in der Programmiersprache Python verwirklicht. Python verfügt über eine Vielzahl an Bibliotheken, die vorgefertigte Funktionen enthalten und auf die im Rahmen der Bearbeitung zurückgegriffen wurde. Tabelle 3-5 bietet einen Überblick über die wichtigsten in dieser Arbeit verwendeten Bibliotheken sowie deren Version und Aufgabenbereich. Nachfolgend werden die wichtigsten Eigenschaften der einzelnen Bibliotheken kurz beschrieben.

Tabelle 3-5: Übersicht der verwendeten Bibliotheken für die Datenverarbeitung und Modellerstellung

| Bibliothek | Version | Aufgabenbereich |
|---------------------|---------|--|
| TensorFlow | 2.10.0 | Entwicklung Künstlich Neuronaler Netze (Überflutungsflächenvorhersage) |
| PyTorch | 2.0.1 | Entwicklung Künstlich Neuronaler Netze (Überstauvorhersage) |
| PyTorch Forecasting | 1.0.0 | Entwicklung <i>Temporal Fusion Transformer</i> (Überflutungsflächenvorhersage) |
| scikit-learn | 1.0.2 | Datenskalierung/Test Maschinelles Lernverfahren |
| Pandas | 1.3.5 | Datenanalyse und -manipulation |
| NumPy | 1.21.5 | Numerische Operationen |
| Matplotlib | 3.5.1 | Grafische Datenvisualisierung |
| GDAL | 3.4.1 | Verarbeitung von Rasterdaten |
| GeoPandas | 0.11.0 | Verarbeitung von Vektordaten |
| NetworkX | 2.7.1 | Verarbeitung von Graphdaten |

Die Entwicklung und das Training Künstlich Neuronaler Netze für das Überflutungsflächenvorhersagemodell erfolgte mit TensorFlow (Abadi et al. 2015) und hier mit dem Modul Keras. Bei

TensorFlow handelt es sich um eine durch das Unternehmen Google entwickelte Open-Source-Bibliothek, die insbesondere für das *Deep Learning*, also der Entwicklung tiefer Neuroner Netze, zum Einsatz kommt. Keras ist eine ursprünglich eigenständig entwickelte Bibliothek, die darauf abzielt, die Entwicklung von *Deep Learning*-Modellen so einsteiger- und nutzerfreundlich wie möglich zu machen. Dadurch wird es möglich, auch sehr komplexe Modelle mit meist wenigen Zeilen Code zu entwickeln. Keras greift beim Ausführen auf verschiedene Bibliotheken wie unter anderem TensorFlow zurück. Mittlerweile ist Keras in TensorFlow als eigenes Modul implementiert und kann von dort ausgeführt werden.

Im Rahmen der Entwicklung des Überstauvorhersagemodells wurde auf PyTorch (Paszke et al. 2019) und PyTorch Forecasting (Breitner 2020) für die Implementierung Künstlich Neuroner Netzarchitekturen zurückgegriffen. PyTorch wurde durch Meta AI entwickelt und wie TensorFlow open-source veröffentlicht. Beide Bibliotheken haben viele Gemeinsamkeiten und zielen auf die Entwicklung von tiefen Neuronalen Netzen ab. PyTorch Forecasting ist eine speziell für die Vorhersage von Zeitreihen mittels *Deep Learning* entwickelte Bibliothek. Wie am Namen bereits zu erkennen, basiert PyTorch Forecasting auf PyTorch, ermöglicht durch vorgefertigte Netzarchitekturen dem Anwender allerdings eine viel schnellere Implementierung auch komplexer Modelle.

Zum Testen weiterer ML-Verfahren wurde auf die Bibliothek scikit-learn (Pedregosa et al. 2011) zurückgegriffen. Scikit-learn beinhaltet eine Vielzahl an gebräuchlichen ML-Algorithmen mit und ermöglicht es, diese sehr einfach aber dennoch effizient anzuwenden. Zudem bietet scikit-learn eine einheitliche Programmierschnittstelle, die es dem Anwender bei der Modellentwicklung sehr einfach ermöglicht, zwischen verschiedenen ML-Verfahren zu wechseln. Neben verschiedenen ML-Algorithmen bietet scikit-learn auch Werkzeuge für die Datenvorverarbeitung, Modellauswahl oder Modellevaluierung.

Zur Vorverarbeitung der Daten wurde insbesondere auf NumPy (Harris et al. 2020) und Pandas (Jeff Reback et al. 2021) zurückgegriffen. NumPy eignet sich für numerische Berechnungen zwischen Vektoren, Matrizen oder mehrdimensionalen Arrays. Die Funktionen von NumPy sind dabei extra für numerische Berechnungen optimiert und erlauben somit eine deutlich schnellere Ausführung von komplexen Berechnungen gegenüber reinen Python-Algorithmen. In der vorliegenden Arbeit kommt NumPy bei der Verarbeitung von Zeitreihen, Rastern und Raster-Sequenzen zum Einsatz. Pandas baut auf NumPy auf und bietet die Verarbeitung von Daten in dem sogenannten *DataFrame*-Format, das sich durch ein tabellarisches Format mit Bezeichnungen für Zeilen und Spalten auszeichnet. Die Bibliothek stellt verschiedene Funktionen für die Datenanalyse und -manipulation zur Verfügung. Ein weitverbreitetes Einsatzgebiet von Pandas ist die Verarbeitung von Zeitreihen, die auch hier den Einsatzschwerpunkt darstellt. Die Visualisierung der Daten und Ergebnisse erfolgt mit matplotlib (Hunter 2007), die vielfältige mathematische Darstellungen ermöglicht.

Für die Verarbeitung von Informationen mit geografischem Bezug wurde auf GDAL und GeoPandas zurückgegriffen. Während GDAL auf die Verarbeitung von Rasterdaten ausgelegt ist, eignet sich GeoPandas für die Verarbeitung von Vektordaten. GDAL kam daher schwerpunktmäßig für die Vorverarbeitung von Geländemodellen oder Überflutungsrastern zum Einsatz. Geopandas wurde hingegen zur Verarbeitung der Kanalnetzdaten genutzt, die als Punkt-

(Knoten) und Linieninformationen (Haltungen) vorlagen. Als dritte Bibliothek zur Verarbeitung von räumlichen Informationen wurde auf NetworkX zurückgegriffen, die speziell auf die Verarbeitung von als Graph strukturierter Daten ausgelegt ist. Damit eignet sich NetworkX zur Vorverarbeitung von Daten für die Modellierung mittels *Graph Neural Networks*.

3.4.3 Sonstige Software

Niederschlagsdatenaufbereitung

Für die Aufbereitung der Niederschlagsdaten wurde die Software AquaZIS (aqua_plan mbH 2019) verwendet. Dabei handelt es sich um ein Messdatenmanagementsystem gemäß DWA-M 151, das neben der reinen Erfassung und Verwaltung von hydrologischen Messdaten auch verschiedene Auswertungen ermöglicht. Dazu zählen unter anderem die Erzeugung von Niederschlagsstatistiken nach DWA-A 531 (2012), Modellregen und partiellen Starkregenserien nach DWA-M 165-1 Anhang A und B (2021).

Verarbeitung und Visualisierung von räumlichen Informationen

Die Verarbeitung von räumlichen Informationen erfolgte überwiegend durch die Software ArcGIS Pro der Firma Esri (ESRI Inc. 2020). Dabei handelt es sich um die Desktopanwendung eines geografischen Informationssystems (GIS), die verschiedene Geoverarbeitungswerkzeuge in einer grafischen Benutzeroberfläche zur Verfügung stellt. Im Rahmen der Arbeit wurde ArcGIS Pro insbesondere für die Vorverarbeitung von räumlichen Informationen verwendet, die anschließend zum Aufbau des gekoppelten hydrodynamischen 1D/2D-Berechnungsmodells verwendet wurden. Hierzu zählt beispielsweise das Öffnen von Brücken und Durchlässen im Geländemodell. ArcGIS Pro wurde aber auch für die Visualisierung von Kartenmaterial des Untersuchungsgebiets und die Durchführung kleinerer Vorabanalysen, wie einer Senken- und Fließweganalyse, verwendet.

3.5 Gütekriterien

Um die Genauigkeit der entwickelten Vorhersagemodelle während des Entwicklungsprozesses sowie in den späteren Sensitivitätsanalysen zu evaluieren, wurden verschiedene Gütekriterien herangezogen. Dabei wurde unterschieden zwischen dem Überstauvorhersagemodell, das Zeitreihen als Zielgröße ausgibt, und dem Überflutungsflächenvorhersagemodell mit Rastersequenzen als Ausgabe. Die für die beiden Anwendungsfälle berücksichtigten Metriken werden nachfolgend nacheinander beschrieben.

Zur Bewertung der Überstauvorhersage werden die prognostizierten (ML-Modell) und simulierten (HD-Modell) Überstauganglinien miteinander verglichen. Bei der Wahl der Kriterien zur Bewertung der Modellgüte wurde insbesondere der spätere Anwendungsfall, nämlich die Übergabe der Überstauvorhersage als Eingabe in das Überflutungsflächenvorhersagemodell, in den Vordergrund gestellt. Wichtig ist dabei, eine möglichst genaue Übereinstimmung des Überstauvolumens sowie die zeitliche und mengenmäßige Überlagerung des Maximalwertes, um die daraus resultierenden Wasserstände an der Geländeoberfläche richtig darzustellen.

Auch die Berücksichtigung von Kriterien zur Bewertung der generellen Anpassung von prognostizierter an die simulierte Überstauganglinie wurden geprüft. Hier ergibt sich bei den in der Hydrologie gängigen Metriken wie der Nash-Sutcliffe-Effizienz (Nash und Sutcliffe 1970) oder der Kling-Gupta-Effizienz (Gupta et al. 2009) jedoch die Problematik, dass fehlerhafte Überstauprognosen an Zeitpunkten, an denen es in der Simulation nicht zu Überstau kam, nicht in die Bewertung einbezogen werden können. Andernfalls würde bei der Berechnung der Metrik für diese Zeitpunkte durch null geteilt werden. Vor diesem Hintergrund und aufgrund der Tatsache, dass die Prognosen nur für kurze Horizonte erzeugt werden und infolgedessen auch die absolute und zeitliche Maximalwertabweichung Aufschluss über die Anpassung geben, wurde auf die Berücksichtigung dieser Metriken verzichtet.

Als Bewertungskriterien werden der absolute Volumenfehler (*Volume Error*, VE), die absolute Maximalwertabweichung (*Peak Error*, PE) sowie die zeitliche Abweichung des Maximalwertes (*Peak Time Error*, PTE) als Gütekriterien herangezogen. Sowohl bei der Berechnung des Volumenfehlers als auch bei der Maximalwertabweichung werden anders als in der Wasserwirtschaft üblich (u. a. DWA-M 165-1, (2021)) nicht die relativen, sondern die absoluten Fehler berechnet. Dadurch soll der Fehler von großen Überstauungen, die auch einen stärkeren Einfluss auf die Überflutungssituation haben, stärker zur Geltung kommen. So wäre etwa ein relativer Volumenfehler von 100 % bei einem Überstauvolumen von wenigen Litern nicht so relevant wie 50 % Abweichung bei mehreren Kubikmetern.

Der absolute Volumenfehler vergleicht die Abweichung des akkumulierten Überstauvolumens zwischen zwei Zeitreihen und wird wie folgt berechnet:

$$VE = \sum_{i=1}^n y_i^{NN} - y_i^{HD} \quad (3-25)$$

Dabei steht n für die Anzahl der verglichenen Zeitschritte und y_i für die jeweiligen mit dem Neuronalen Netz NN und dem hydrodynamischen Modell HD ermittelten Werte der einzelnen Zeitschritte. Der VE kann Werte im Bereich $[0, \infty]$ annehmen, wobei der Wert 0 der geringsten Abweichung entspricht.

Die Maximalwertabweichung PE vergleicht die Scheitel zweier Überstauganglinien miteinander, liefert allerdings keine Aussage hinsichtlich der zeitlichen Überlagerung. Hierzu wird die zeitliche Abweichung des Maximalwertes PTE ergänzend berücksichtigt. Beide Kriterien lassen sich mit den folgenden Formeln berechnen:

$$PE = y_{peak}^{NN} - y_{peak}^{HD} \quad (3-26)$$

$$PTE = y_{t,peak}^{NN} - y_{t,peak}^{HD} \quad (3-27)$$

Hier stehen y_{peak} und $y_{t,peak}$ jeweils für den Scheitelwert und den Zeitpunkt des Auftretens in den mit dem Neuronalen Netz NN und dem hydrodynamischen Modell HD ermittelten Überstauganglinien.

Zur Bewertung der Vorhersageergebnisse des Überflutungsflächenvorhersagemodells wer-

den die prognostizierten (ML-Modell) und simulierten (HD-Modell) Überflutungsraster zellenweise miteinander verglichen. Hierzu werden mit dem *Root Mean Squared Error* (RMSE) und dem *Critical Success Index* (CSI) zwei unterschiedliche Arten von Gütekriterien herangezogen.

Der RMSE ist ein kontinuierlicher Index, der die genauen Wasserstände miteinander vergleicht und die durchschnittliche Abweichung bewertet:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^{NN} - y_i^{HD})^2} \quad (3-28)$$

Dabei steht n für die Anzahl der verglichenen Zellen und y_i für die jeweiligen mit dem Neuralen Netz NN und dem hydrodynamischen Modell HD ermittelten Werte der einzelnen Zellen. Der RMSE kann Werte im Bereich $[0, \infty]$ annehmen, wobei der Wert 0 der optimalen Anpassung entspricht. Als Ergebnis wird der absolute Fehler angegeben. Andere Metriken zur Ermittlung des relativen Fehlers, wie der *Relative Mean Squared Error* (MRSE) wurden ebenfalls getestet. Es zeigte sich allerdings, dass Pixel mit niedrigen Wasserständen teilweise zu extremen prozentualen Abweichungen führten. Bei der anschließenden Mittelwertbildung über alle Zellen eines Überflutungsrasters, führte diese Problematik zu schlechten Ergebnissen, obwohl die betroffenen Zellen nur ein geringes Gefährdungspotential aufweisen und damit gegenüber Zellen mit hohen Wasserständen von untergeordneter Relevanz sind.

Der CSI ist ein kategorischer Index zur Bewertung der Lagegenauigkeit, der ein weitverbreitetes Maß zur Bewertung von Extremereignissen darstellt und sowohl bei der Vorhersage von Niederschlag (u. a. Shi et al. 2015; Ayzel et al. 2019; Ravuri et al. 2021) als auch Sturzfluten (Jamali et al. 2019; Löwe et al. 2021) eingesetzt wird. Gegenüber anderen kategorischen Indizes, wie der *Hit Rate* (HR) oder der *False Alarm Rate* (FAR) berücksichtigt der CSI sowohl verpasste als auch fälschlicherweise prognostizierte Überflutungen. Zur Ermittlung des CSI erfolgt zunächst eine binäre Klassifizierung der Zellen. Im vorliegenden Fall erfolgt die Klassifizierung der Pixel nach „überflutet“ und „nicht überflutet“. Anschließend lässt sich der CSI wie folgt berechnen:

$$CSI = \frac{TP}{TP + FN + FP} \quad (3-29)$$

Dabei steht TP für die Anzahl der korrekt als überflutet prognostizierten Zellen, FP bezeichnet die fälschlicherweise als überflutet prognostizierten Zellen und FN gibt die Anzahl der Zellen an, die fälschlicherweise als nicht überflutet prognostiziert wurden. Es wird somit sowohl auf verpasste als auch fehlerhafte Alarme reagiert. Damit ist der CSI besonders gut für die vorliegende Problemstellung geeignet, da sowohl fehlende als auch fälschlicherweise prognostizierte Überflutungsflächen im Einsatzfall gleichermaßen ungünstig sind. Die Werte des CSI liegen im Intervall $[0, 1]$, wobei 1 dem besten Ergebnis entspricht.

4 Datengenerierung und -aufbereitung

Um Maschinelle Lernverfahren und insbesondere Künstlich Neuronale Netze für die Lösung eines Problems verwenden zu können, ist zunächst ein Trainingsprozess erforderlich. Neben dem verwendeten Lernverfahren sowie den verfahrensspezifischen Parametern hängt die Leistungsfähigkeit von ML-Modellen in erster Linie von der für den Trainingsprozess zugrunde liegenden Datenbasis ab. Im vorliegenden Anwendungsfall sollen ML-Modelle zur Vorhersage von Überstau und Überflutungsflächen in Folge von Starkregenereignissen verwendet werden. Damit das verwendete Modell in der Lage ist, die gesamte Variabilität der potentiellen Überstau- und Überflutungseignisse im Einzugsgebiet adäquat vorherzusagen, muss ein möglichst repräsentativer Datensatz im Trainingsprozess zur Verfügung gestellt werden. Das vorhandene Messnetz ist in Kanalnetzen erfahrungsgemäß sehr grob und beschränkt sich in erster Linie auf einzelne Messungen an Sonderbauwerken und ggf. ergänzende temporäre Messkampagnen an wenigen Punkten im Kanalnetz. Dementsprechend reichen reale Messdaten für eine detaillierte Abbildung der Überstaucharakteristik in einem Einzugsgebiet derzeit nicht aus. An der Oberfläche fehlen Messungen über vergangene Überflutungssituationen sogar vollständig und es liegen, wenn überhaupt, nur Schätzungen von Wasserständen vergangener Ereignisse an einzelnen Punkten vor. Zudem sind gemäß dem im DWA-A 118 (2024) empfohlenen Überstaunachweis Kanalnetze in urbanen Gebieten für Überstauhäufigkeiten im Bereich von 2 bis 10 Jahren auszulegen. Selbst bei einer hohen Messnetzdichte und langen Messreihen würde die Anzahl der erfassten Ereignisse nur einen geringen Teil der möglichen Verläufe abbilden.

Physikalisch-basierte Modelle sind dagegen in der Lage, auch unbekannte Überstauverläufe und Überflutungssituationen zu berechnen. Aufgrund der prozessorientierten Modellierung besitzen diese Modelle zudem die Fähigkeit zur Extrapolation. Dementsprechend können für ein Einzugsgebiet auch extreme, bisher nicht an diesem Ort aufgetretene Ereignisse simuliert werden, solange entsprechende Eingabedaten zur Verfügung stehen. In der vorliegenden Arbeit soll diese Fähigkeit auf ein ML-Modell übertragen werden, um die Vorteile beider Modellierungsparadigmen zu kombinieren. Hierzu wird mit einem physikalisch-basierten Modell ein künstlicher Trainingsdatensatz erzeugt, der eine möglichst hohe Variabilität an Niederschlagszenarien und den daraus resultierenden Überflutungseignissen abbildet. Grundlage hierfür sind die in Kapitel 3.2 beschriebene Datengrundlage und die in Kapitel 3.4.1 beschriebene Software zur physikalisch-basierten Überflutungsmodellierung. Wie in Kapitel 2.5 beschrieben, können bei dem gewählten Systemaufbau die ML-basierten Vorhersagemodelle maximal so gut wie das zur Generierung der Trainingsdaten verwendete physikalisch-basierte Modell werden. Dementsprechend sorgfältig sollte das physikalisch-basierte Modell aufgebaut und kalibriert werden, um die Unsicherheiten möglichst gering zu halten. Das im Rahmen dieser Arbeit verwendete Vorgehen zur Generierung von Niederschlagsszenarien, der Aufbau und die Kalibrierung des physikalisch-basierten Modells sowie die finale Generierung des Trainingsdatensatzes werden in diesem Kapitel beschrieben.

4.1 Generierung von Niederschlagsreihen

Um einen möglichst repräsentativen Trainingsdatensatz zu erzeugen, ist bereits bei der Auswahl der verwendeten Niederschlagszeitreihen darauf zu achten, möglichst viele potentielle Ereignistypen zu berücksichtigen. Da insbesondere extreme Starkregenereignisse mit besonders verheerendem Überflutungsmaß nur selten auftreten, werden die Untersuchungen nicht nur auf die historischen Ereignisse im Untersuchungsgebiet beschränkt. Stattdessen werden Untersuchungen mit einem Datensatz aus verschiedenen Modellregen und einem Datensatz mit Starkregenserien von allen acht in Kapitel 3.2.1 aufgeführten Niederschlagsstationen durchgeführt. Mit beiden Datensätzen wird das hydrodynamische 1D/2D-Modell belastet, um die Trainingsdatensätze zu erstellen. Für die Erzeugung der Niederschlagsreihen wurde die Software AquaZIS (aqua_plan mbH 2019) verwendet (vgl. Kapitel 3.4.3).

Ein wesentlicher Bestandteil der vorliegenden Arbeit sind Untersuchungen zur erforderlichen Datenmenge und -qualität, um ein ML-basiertes Vorhersagemodell effizient trainieren zu können. Während natürliche Regenreihen das tatsächliche Niederschlagsgeschehen besser abbilden, bieten Modellregen die Möglichkeit, auch Niederschlagsereignisse mit hohen Wiederkehrzeiten repräsentativ abzudecken. Aufgrund der geringen Größe des Untersuchungsgebietes wurden bei der Auswahl der natürlichen Ereignisse als auch der Modellregen Dauerstufen im Bereich von 15 bis 240 Minuten berücksichtigt. Ereignisse mit längeren Dauerstufen wurden hingegen vernachlässigt, da diesen unter Berücksichtigung der relativ geringen Größe des Untersuchungsgebiets nur eine untergeordnete Bedeutung zukommt. Ferner haben Untersuchungen gezeigt, dass bei Modellregen vom Euler Typ II mit Dauerstufen im Bereich von 5 bis 120 Minuten erst ab einer Wiederkehrzeit von 10 Jahren vereinzelt Überstau im betrachteten Untersuchungsgebiet auftritt. Daher werden bei den Modellregen nur Wiederkehrzeiten ab 10 Jahren berücksichtigt. Bei den natürlichen Regenreihen werden zum Teil auch Ereignisse ab einer Wiederkehrzeit von 5 Jahren berücksichtigt, da hier auch ungünstigere Verläufe auftreten können, die bereits bei einer geringeren Wiederkehrzeit zu Überstau führen. Weitere spezifische Eigenheiten der beiden Datensätze werden in den nachfolgenden Kapiteln getrennt voneinander beschrieben.

4.1.1 Modellregenreihen

Für die Erzeugung der Modellregen wurde auf Niederschlagsmessungen der Station Gelsenkirchen-Altstadt zurückgegriffen (vgl. Kapitel 3.2.1). Für die Station liegen geprüfte Messdaten für den Zeitraum 01.01.1931 bis 31.12.2020 in einer zeitlichen Auflösung von einer Minute vor. Mit Hilfe der Messdaten wurde zunächst eine Niederschlagsstatistik gemäß DWA-A 531 (2012) erstellt, mit der anschließend die Modellregen konstruiert wurden. Da die vollständige Berücksichtigung der Messreihe aus rechentechnischen Gründen nicht möglich war, wurde die Statistik für den Zeitraum 1980 bis 2020 erstellt. Damit wird die nach DWA-M 165-1 (2021) geforderte Dauer von mindestens 30 Jahren eingehalten. Zudem wird so eine eventuell vorhandene klimawandelbedingte Zunahme der Anzahl und Intensitäten von Starkregenereignissen stärker berücksichtigt.

Um eine möglichst große Bandbreite an potentiellen Starkregenverläufen abzudecken, wurden

Modellregen vom Typ Euler mit unterschiedlichen Intensitätsverläufen, Dauerstufen und Wiederkehrzeiten erzeugt. Bei den Verläufen wurden Euler-Regen vom Typ I, II und III berücksichtigt. Die Konstruktion der Regen vom Typ Euler I und II erfolgte in Anlehnung an das DWA-M 165-1 (2021). Bei dem Modellregen vom Euler Typ III wurde die Intensitätsspitze nach der Hälfte der Dauer berücksichtigt. Für alle drei Euler-Regen wurden die Dauerstufen 15, 30, 60, 120 und 240 Minuten sowie Wiederkehrzeiten von 10, 20, 30, 50 und 100 Jahren berücksichtigt. Zusätzlich zu den abgedeckten Wiederkehrzeiten wurden in Anlehnung an Schmitt (2017) sogenannte Erhöhungsfaktoren berücksichtigt, um auch extreme Niederschläge jenseits einer Wiederkehrzeit von 100 Jahren abzudecken. Als Erhöhungsfaktoren wurden die Werte 1,5 und 4,0 berücksichtigt. Der Wert 4,0 wurde in Anlehnung an die Erkenntnisse aus Untersuchungen zur Bestimmung „Maximierter Gebietsniederschlagshöhen für Deutschland (MGN)“ festgelegt. Dabei handelt es sich um eine physikalisch-empirisch basierte Schätzung der vermutlich größten physikalisch möglichen Niederschlagshöhen (Meon et al. 2009). Da die MGN-Werte mit Bezug auf die DWD-KOSTRA Datenbasis etwa das 3,5- bis 4-fache der 100-jährlichen Niederschlagshöhe betragen (Schmidt, 1997, zitiert nach Schmitt 2017) wurde diese Obergrenze auch hier als maximale Niederschlagshöhe berücksichtigt. Durch die Kombination der drei Modellregenarten mit den fünf Dauerstufen sowie den sieben Wiederkehrzeiten bzw. Erhöhungsfaktoren (vgl. Abbildung 4-1) ergeben sich insgesamt 105 Ereignisse, die im Datensatz enthalten sind.

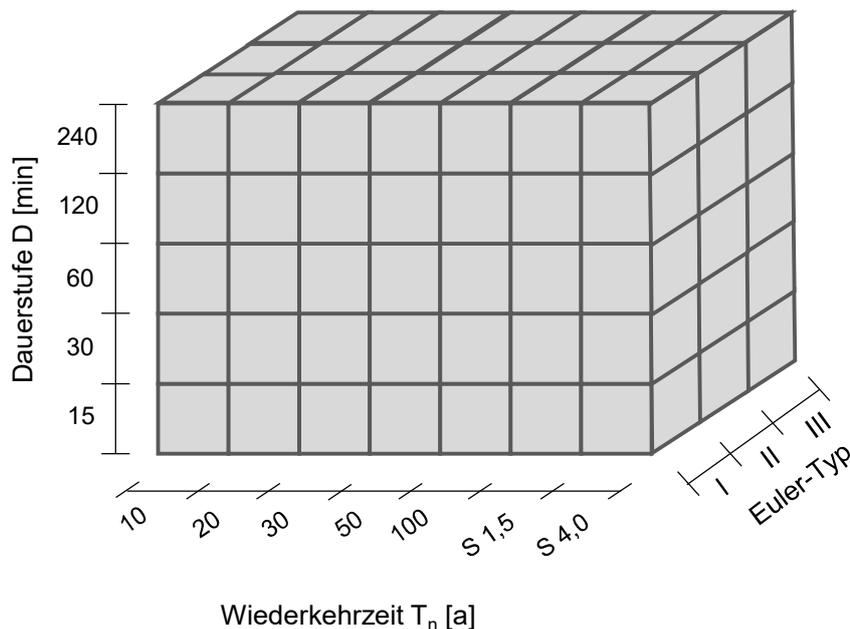


Abbildung 4-1: Schematische Darstellung der berücksichtigten Modellregen sowie der gewählten Dauerstufen, Modellregenarten und Wiederkehrzeiten/Szenarien. Bei den Szenarien S 1,5 und S 4,0 gibt die Nummer jeweils den Erhöhungsfaktor an, mit dem die Werte der 100-jährigen Modellregen multipliziert wurden

4.1.2 Natürliche Regenreihen

Zur Berücksichtigung natürlicher Niederschlagsereignisse bei der Erzeugung der Trainingsdatensätze wurde auf alle acht in Kapitel 3.2.1 aufgeführten Messstationen zurückgegriffen. Ziel ist es, dadurch gegenüber der ausschließlichen Berücksichtigung der nächstgelegenen Station, die Anzahl von Starkregenereignissen mit besonders hohen Wiederkehrzeiten im Trainingsdatensatz zu erhöhen. Aus den Regenreihen wurden anschließend partielle Starkregenserien erstellt. Dabei wurden in Anlehnung an die Empfehlungen im Anhang B des DWA-M 165-1 (2021) folgende Randbedingungen berücksichtigt:

- Relevante Dauerstufen: 15, 30, 60, 120 und 240 Minuten
- Zeitlicher Mindestabstand zwischen zwei Ereignissen: 4 h
- Mindestintensität: 0,1 mm in 5 min

Die Zuweisung von Wiederkehrzeiten bei der Erzeugung der Starkregenserien stützt sich jeweils auf die Niederschlagsstatistik der Station Gelsenkirchen-Altstadt, die auch für die Erzeugung der Modellregen verwendet wurde. Wie eingangs erwähnt, haben Voruntersuchungen gezeigt, dass erst bei Niederschlagsereignissen mit einer Wiederkehrzeit > 5 Jahre mit Überstau und auch der Ausbildung relevanter Überflutungsflächen zu rechnen ist. Daher wurde die Anzahl der zu berücksichtigenden Ereignisse über die Festlegung eines Jährlichkeitsbereichs reduziert. Hierzu wurde bei den Stationen Gelsenkirchen-Altstadt und Bochum-DMT eine minimale Jährlichkeit von 5 Jahren berücksichtigt. Dadurch konnten im Jährlichkeitsbereich zwischen 5 und 10 Jahren bereits durch die Regenreihen der beiden Stationen sämtliche berücksichtigte Dauerstufen umfangreich abgedeckt werden. Um unnötige Rechenzeiten zu vermeiden, wurde bei den übrigen Stationen die minimale Jährlichkeit der zu berücksichtigenden Ereignisse auf 10 Jahre festgelegt. Bezüglich der maximalen Jährlichkeit erfolgt hingegen bei allen Stationen keine Begrenzung. Insgesamt wurden aus den Messreihen aller acht Stationen 153 Ereignisse identifiziert, die für mindestens eine der berücksichtigten Dauerstufen in den entsprechenden Jährlichkeitsbereichen liegen. Anhang A-2 gibt einen Überblick über die jeweiligen Ereignisse, deren Dauer, Niederschlagsmenge und Wiederkehrzeit für die bei der Auswahl berücksichtigten Dauerstufen.

Abbildung 4-2 liefert einen Überblick über die Verteilung der Ereignisse im Datensatz in Abhängigkeit der Wiederkehrzeit für die fünf berücksichtigten Dauerstufen. Zusätzlich ist die Verteilung der Ereignisse in Abhängigkeit der maximalen Wiederkehrzeit hervorgehoben dargestellt. Insgesamt zeigt sich, dass alle Bereiche abgedeckt sind, jedoch die Anzahl der Ereignisse für die geringen Wiederkehrzeiten erwartungsgemäß höher ausfällt. Zwar steigt die Anzahl der Ereignisse für das Intervall > 100 wieder leicht an, dies begründet sich jedoch durch die deutlich größere Ausdehnung des Intervalls. Hier werden im Rahmen der Sensitivitätsanalysen Untersuchungen durchgeführt, inwieweit gegebenenfalls Anpassungen am Datensatz durchzuführen sind, um eine Überanpassung des Modells an weniger starke Ereignisse zu vermeiden (vgl. Kapitel 5.4.1 und 6.4.1).

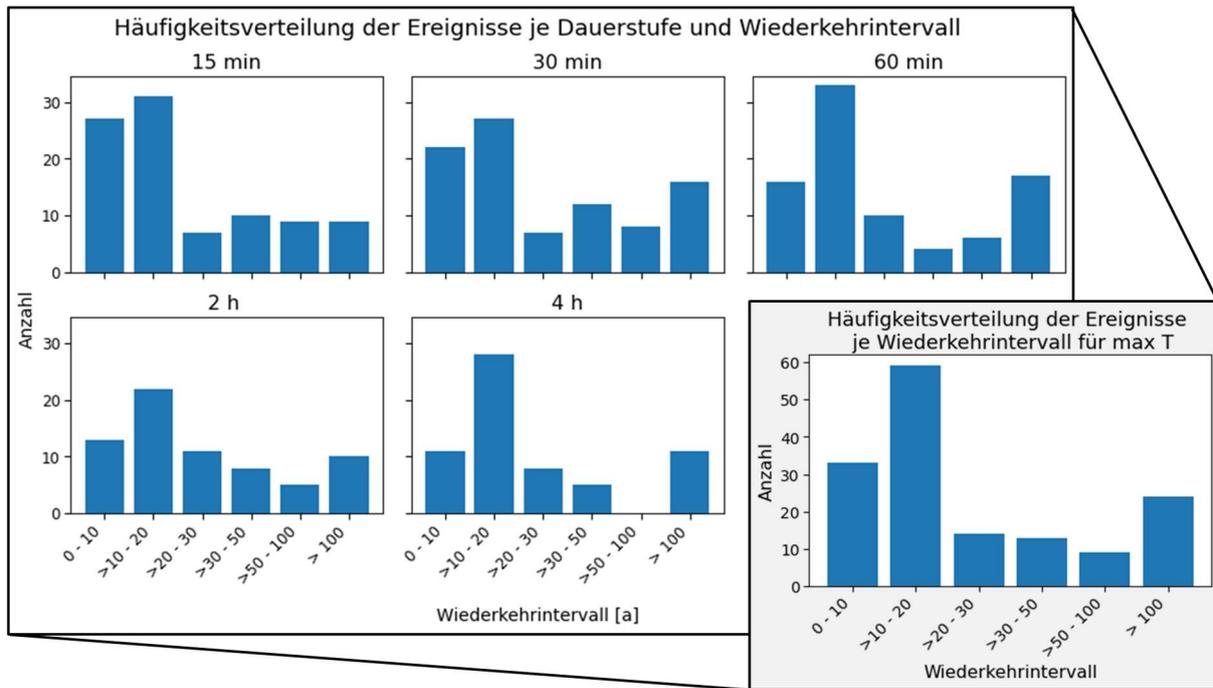


Abbildung 4-2: Häufigkeitsverteilung der Ereignisse je Wiederkehrintervall getrennt für die berücksichtigten Dauerstufen und zusammengefasst für die maximale Wiederkehrzeit der Ereignisse. Die Auswertung wurde erstellt auf Basis von EGLV (2021a)

4.2 Aufbau des gekoppelten 1D/2D-Überflutungsmodells

Das in dieser Arbeit für die Datengenerierung verwendete 1D/2D-Berechnungsmodell wurde mit Hilfe der in Kapitel 3.2 beschriebenen Datengrundlage in der Software MIKE+ erstellt (vgl. Kapitel 3.4.1). Hierzu wurden die folgenden wesentlichen Schritte durchgeführt, nach denen sich auch die nachfolgende Beschreibung gliedert:

- Aufbau des 1D-Kanalnetzmodell
- Aufbau 2D-Oberflächendmodell
- Modellkopplung und Abflussbildungsansatz

Aufbau 1D-Kanalnetzmodell

Während die Netzdaten des zur Verfügung gestellten Kanalnetzmodells (AGG 2021) bereits vollständig und in einem geeigneten Format vorlagen, konnten die angeschlossenen Flächen nicht ohne Weiteres übernommen werden. Grund hierfür war die zu grobe räumliche Aufteilung der einzelnen Flächen, die zwar für die reine 1D-Berechnung ausreichend, für eine gekoppelte Betrachtung jedoch zu grob war. Stattdessen wurde auf die Daten zur Oberflächenbeschaffenheit (vgl. Kapitel 3.2.4) zurückgegriffen, um die Zuweisung von Flächen an die einzelnen Knoten vorzunehmen.

Die Zuordnung der einzelnen Flächen zu den Schächten im Kanalnetz erfolgte auf Flurstücksebene. Hierzu wurde auf Flurstückdaten des amtlichen Liegenschaftskatasterinformationssystem (ALKIS) zurückgegriffen (Bezirksregierung Köln 2022), die mit Hilfe eines in MIKE+ integrierten Algorithmus den Schächten des Kanalnetzes zugeordnet wurden. Im Anschluss daran

wurde die Information zu der Einzugsgebietszuordnung der Flurstücke auf die jeweiligen im Flurstück enthaltenen Teileinzugsgebiete übertragen. Somit kann sichergestellt werden, dass alle Flächen eines Flurstückes an den gleichen Schacht angeschlossen sind. Zusätzlich zur automatisierten Zuordnung erfolgte eine Plausibilisierung und manuelle Nachbearbeitung. Dies gilt insbesondere für Flurstücke mit einer Fläche größer 5000 m², die nicht automatisiert, sondern vollständig manuell zugeordnet wurden. Durchlässige Flächen wurden dabei vernachlässigt, da hier die Abflussbildung vollständig über das Oberflächenmodell erfolgt.

Als nächstes wurden den einzelnen Flächen Ansätze zur Abflussbildung und -konzentration zugewiesen. In Anlehnung an die Ergebnisse der Untersuchungen von Wistuba (2021) wurde der in MIKE+ hinterlegte Modellansatz C2 gewählt (vgl. Kapitel 3.4.1). Bei der Parametrisierung wurden zunächst die durch die Software vorgegebenen Standardparameter gewählt, die in Tabelle 4-1 abgebildet sind. Da bei allen berücksichtigten Flächennutzungen der geschätzte Befestigungsgrad größer als 20 % ist, wird bei dem verwendeten Modellansatz keine Infiltration über den Horton-Ansatz berücksichtigt (DHI 2017a). Falls erforderlich, kann die Infiltration aber vereinfacht über den Reduzierungsfaktor als konstanter Dauerverlust abgebildet werden.

Tabelle 4-1: Parametersätze der einzelnen Flächennutzungen für den verwendeten Modellansatz C2

| Flächennutzung | Abflussbildung | | Abflusskonzentration |
|-----------------------------|------------------------|---------------------------|----------------------------|
| | Anfangsverlust [mm] | Reduzierungsfaktor [-] | Speicherkonstante [min] |
| Dachflächen | 0,5 | 0,9 | 10 |
| Öffentliche Verkehrsflächen | 0,5 | 0,9 | 10 |
| Privat schmutzig | 1 | 0,9 | 10 |
| Privat sauber | 1 | 0,9 | 10 |

Aufbau 2D-Oberflächenmodell

Das Oberflächenmodell basiert auf einem rasterbasierten Rechennetz, dem durch Interpolation die Höheninformationen aus dem in Kapitel 3.2.2 beschriebenen Geländemodell hinzugefügt wurden. Bei diesem Vorgang wurden Gebäude als Abflusshindernisse im Rechennetz pauschal um 10 m erhöht. Als Auflösung wurde eine Kantenlänge von 2 m je Rasterfeld gewählt, als optimaler Kompromiss zwischen möglichst hoher Genauigkeit der Ergebnisse und gleichzeitig möglichst geringer Rechendauer. Voruntersuchungen haben gezeigt, dass sich gegenüber einer Auflösung von 1 m kaum Unterschiede in den Berechnungsergebnissen ergeben, die Rechendauer aber um einen Faktor > 4 reduziert wird. Eine Auflösung von 4 m stellte sich hingegen in manchen Bereichen als zu grob heraus, um auch kleinflächige Details abzubilden. Eine erste Plausibilitätsprüfung hat gezeigt, dass die Überflutungsflächen des entwickelten Modells sehr gut mit der aktuellen Starkregengefahrenkarte für Gelsenkirchen (Stadt Gelsenkirchen 2019) übereinstimmen.

Die Ausdehnung des Oberflächenmodells im Untersuchungsgebiet wurde im Rahmen einer

GIS-basierten Fließweg- und Senkenanalyse ermittelt. So konnten vorab Überflutungs-Hotspots sowie die zugehörigen Einzugsgebiete ermittelt werden, die zur Bestimmung der Ausdehnung des Rechnetzes genutzt wurden. Das finale Rechnetz wurde über eine Fläche von 3,1 km² mit insgesamt 772.415 Elementen aufgespannt (vgl. Abbildung 4-3).

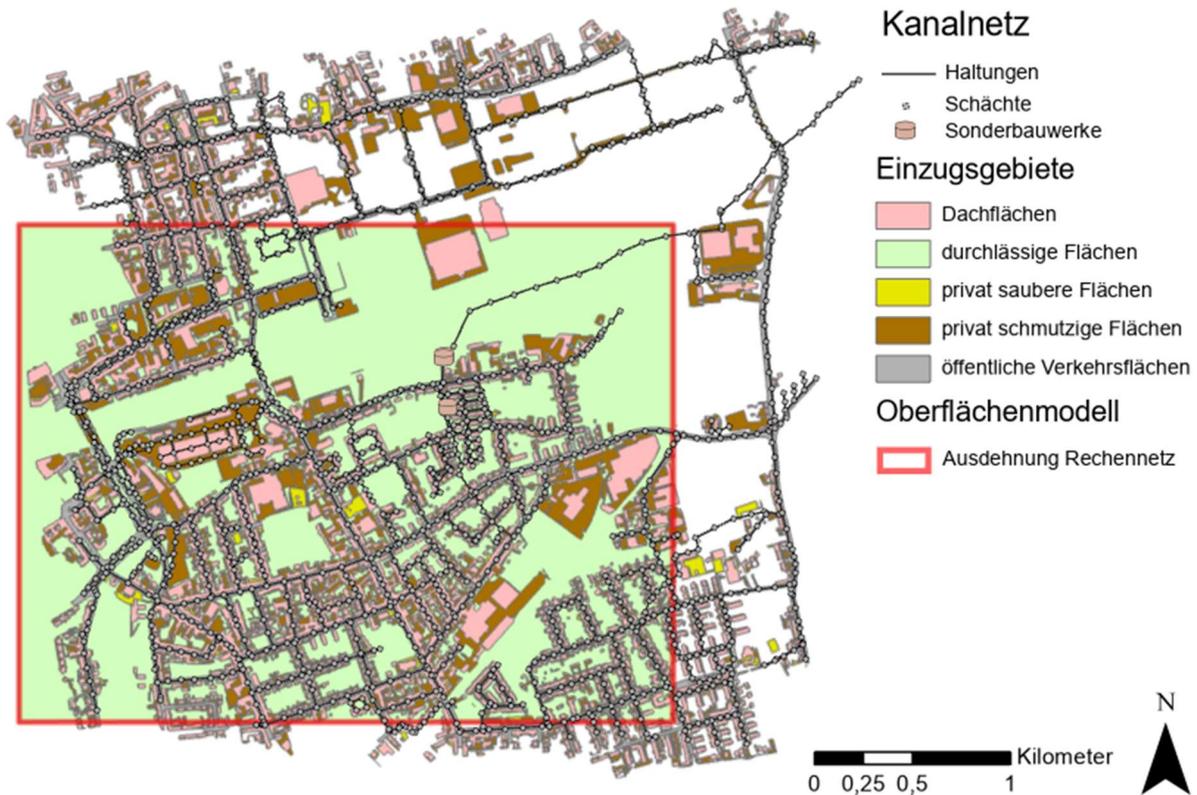


Abbildung 4-3: Übersicht über das 1D-Kanalnetzmodell mit den zugehörigen (Teil-) Einzugsgebieten sowie dem Bereich des Oberflächenmodells, Datenquelle: AGG (2021) und EGLV (2021b)

Die nachfolgenden Untersuchungen konzentrieren sich auf den gekoppelt modellierten Bereich. Da die umliegenden Bereiche durch Rückstaueffekte einen Einfluss auf den gekoppelt modellierten Bereich nehmen können, werden diese weiterhin als 1D-Kanalnetzmodell berücksichtigt. Aufgrund der Tatsache, dass nur befestigte und teildurchlässige Flächen an das Kanalnetzmodell angeschlossen sind, werden durchlässige Flächen in diesen Bereichen nicht berücksichtigt. Insgesamt ergeben sich so die in Tabelle 4-2 aufgeführten Flächenanteile. Durch die Eingrenzung des Oberflächenmodells kann der für die hier durchgeführten Untersuchungen benötigte Rechenaufwand weiter reduziert werden, ohne relevante Informationsverluste zu erleiden.

Zur möglichst realitätsnahen Abbildung des Abflussgeschehens an der Oberfläche können in MIKE+ flächendifferenzierte Verluste und Oberflächenrauheiten angegeben werden. Bei den Verlusten wurden Infiltrationsverluste über ein in MIKE+ implementiertes vereinfachtes Speichermodell berücksichtigt (vgl. (DHI 2021a)). Dabei wird die ungesättigte Zone als Speicher angesehen, dessen Speicherfähigkeit über die Tiefe, die Porosität und den Wassergehalt parametrisiert werden kann. Über eine Infiltrationsrate kann der Befüllungsvorgang und über eine Leckagerate der Entleerungsvorgang festgelegt werden. Zur Parametrisierung wurde zunächst die Bodenkarte von NRW herangezogen. Hieraus geht hervor, dass die Hauptbodenart

im Untersuchungsgebiet nach Bundes-Bodenschutz- und Altlastenverordnung (BBodSchV) „Lehm/Schluff“ ist. Weiterhin geht aus der Bodenkarte hervor, dass Grundwasser erst in größeren Tiefen (1,3 bis 2 m) anzutreffen ist. Hierin begründet sich die in Tabelle 4-3 aufgeführte Anfangsparametrisierung der Speicherparameter. Die Infiltrationsraten und Leckageraten wurden anhand von Literaturwerten (u. a. Illgen 2010) geschätzt. Auf eine zusätzliche Berücksichtigung von Verdunstungsverlusten wurde aufgrund der geringen Bedeutung bei urbanen Sturzfluten verzichtet. Die Oberflächenrauheiten werden zur Ermittlung des Reibungsgefälles für die 2D-Flachwassergleichungen benötigt (vgl. Gleichung (3-23) und (3-24)). Die Rauheit kann in MIKE+ als Rauheitsbeiwert nach Manning-Strickler angegeben werden und beeinflusst den Abflusstransport. Wie bei den Infiltrationsraten wurden auch hier flächenspezifische Werte in Anlehnung an Literaturwerte (Chow 1959, zitiert nach Hürter 2018) festgelegt, die in Tabelle 4-3 dargestellt sind.

Tabelle 4-2: Flächenanteile im Untersuchungsgebiet

| Flächenart/-nutzung | | Gekoppelter Bereich (1D/2D-Modell) | | Umliegende Bereiche (1D-Modell) | | Gesamt | |
|-----------------------------|-----------------------------|------------------------------------|------|---------------------------------|------|--------------------|------|
| | | [km ²] | [%] | [km ²] | [%] | [km ²] | [%] |
| Befestigt/ Teil-durchlässig | Dachflächen | 0,60 | 41,8 | 0,29 | 38,8 | 0,89 | 47,4 |
| | Öffentliche Verkehrsflächen | 0,36 | 25,2 | 0,22 | 29,1 | 0,58 | 30,8 |
| | Privat schmutzig | 0,38 | 26,5 | 0,19 | 25,2 | 0,57 | 30,3 |
| | Privat sauber | 0,09 | 6,5 | 0,05 | 7,0 | 0,15 | 7,8 |
| | Summe | 1,44 | 100 | 0,74 | 100 | 1,87 | 100 |
| Durchlässige Flächen | | 1,87 | 100 | - | - | 1,87 | 100 |
| Summe | | 3,31 | | 0,74 | | 4,05 | |

Tabelle 4-3: Flächenspezifische Parametrisierung des Oberflächenmodells

| Flächennutzung | Infiltrationsmodell | | | | | Rauheit [m ^{1/3} /s] |
|-----------------------------|--------------------------|--------------------|-----------|---------------|------------------|-------------------------------|
| | Infiltrationsrate [mm/h] | Leckagerate [mm/h] | Tiefe [m] | Porosität [-] | Wassergehalt [%] | |
| Dachflächen | 0 | 0 | 0 | 0 | 0 | 3 |
| Öffentliche Verkehrsflächen | 0 | 0 | 0 | 0 | 0 | 70 |
| Privat schmutzig | 4,5 | 0,5 | 2 | 0,3 | 50 | 50 |
| Privat sauber | 4,5 | 0,5 | 2 | 0,3 | 50 | 50 |
| Durchlässig | 18 | 1,8 | 2 | 0,3 | 50 | 35 |

Modellkopplung und Abflussbildung

Beide Modelle wurden abschließend über die Schächte bidirektional miteinander gekoppelt. Die Kopplung erfolgt dabei jeweils mit dem oberhalb des Schachtes liegenden Rasterelement. Im Rahmen der gekoppelten Modellierung mit MIKE+ kann die Abflussbildung sowohl über die Einzugsgebietsflächen des Kanalnetzmodells als auch über das Oberflächenmodell erfolgen. Darüber hinaus ist auch die Kombination von beidem möglich, indem Flächen wie Dächer, die

bevorzugt direkt an das Kanalnetz angeschlossen sind, über das Kanalnetzmodell belastet werden, während alle übrigen Flächen oberflächlich belastet werden. Da durchlässige Flächen in der Regel nicht direkt an das Kanalnetz angeschlossen sind, wurden diese Flächen über das Oberflächenmodell belastet. Dachflächen wurden hingegen über das Kanalnetzmodell berücksichtigt. Die übrigen Flächen wurden zwar zunächst dem Kanalnetzmodell zugeordnet, die genaue Zuweisung ist aber im Rahmen der Modellkalibrierung zu prüfen und anzupassen.

4.3 Modellkalibrierung

Da Modelle immer eine Abstraktion der Realität unter Annahme diverser Vereinfachungen darstellen, sind diese mit verschiedenen Unsicherheiten verbunden. Dies gilt auch für Niederschlag-Abfluss-Modelle, weshalb immer davon auszugehen ist, dass es zu Abweichungen zwischen Simulationsergebnissen und Messungen kommen kann. Neben den verwendeten Verfahren sind auch die zugrunde liegenden Modellparameter mit Unsicherheiten verbunden, da sich diese wie im Falle der abflusswirksamen Fläche oder der Oberflächenrauheiten messtechnisch nicht eindeutig bestimmen lassen (Muschalla et al. 2015). Die Modellkalibrierung hat daher die Aufgabe, diese Modellparameter so weit anzupassen, dass das Modell die Realität mit einer adäquaten Genauigkeit abbildet. Dieser Vorgang erfolgt in der Regel durch den Abgleich von Simulationsergebnissen mit Messwerten (DWA 2021). Ferner stellt die Modellkalibrierung eine Möglichkeit dar, Fehler in der Modellerstellung zu detektieren. Aufgrund der großen Bedeutung der Modellkalibrierung wurde diese in der aktuellen Fassung des Merkblattes DWA-M 165 (2021) als eigenes Kapitel aufgenommen. Dort erfolgt auch eine detaillierte Vorgehensbeschreibung, nach der sich auch die vorliegende Arbeit richtet.

Für die Kalibrierung standen Niederschlagsmessungen der Station Gelsenkirchen-Altstadt und Durchflussmessungen von zwei temporären Messstellen mit Messgeräten der Firma NIVUS zur Verfügung. Messungen im Bereich der Geländeoberfläche lagen nicht vor, weshalb sich die Kalibrierung auf den Bereich des Kanalnetzes beschränkte. Durch die Kopplung beider Modelle erfolgt aber zumindest eine indirekte Kalibrierung des Oberflächenmodells, auch wenn zusätzliche Daten an der Geländeoberfläche wünschenswert gewesen wären. Insgesamt wurden Messdaten für den Zeitraum vom 01.07.2021 bis 31.01.2022 verwendet. In einem ersten Schritt wurden anhand der Niederschlagsmessdaten mit einem Trennkriterium von vier Stunden alle Niederschlagsereignisse im Messzeitraum identifiziert. Anschließend wurden alle Ereignisse herausgefiltert, die folgende Kriterien erfüllen:

- Mindestniederschlagshöhe: 5 mm
- Mindestniederschlagsintensität: 2,5 mm/h

Mit diesem Vorgehen konnten zunächst sieben Ereignisse identifiziert werden. Bei drei dieser Ereignisse fehlten entweder aufgrund von Messausfällen die Durchflussmessungen oder der Faktor des akkumulierten Durchflussvolumens zwischen beiden Durchflussmessstellen wies auf Unstimmigkeiten hin. Um dadurch verursachte Probleme bei der Kalibrierung zu vermeiden, wurden diese Ereignisse nicht weiter berücksichtigt. Die visuelle Kontrolle der Ereignis-

ganglinien zeigte zudem, dass ein Ereignis vom 14.07.2021 zwei geeignete Zeiträume mit einer niederschlagsfreien Zeit von mehr als drei Stunden dazwischen aufwies. Dieses Ereignis wurde zwar zusammenhängend simuliert, für die spätere Überprüfung der Modellgüte aber in zwei Ereignisse aufgeteilt. Hierdurch wird eine Überschätzung der Modellgüte durch eine gute Anpassung in niederschlagsfreien Zeiträumen vermieden. Insgesamt wurden so für die Kalibrierung die in Tabelle 4-4 aufgeführten Ereignisse berücksichtigt. Ereignisränder mit geringen Intensitäten wurden dabei bereits entfernt, um die Simulationsdauer moderat zu halten. Da nur Niederschlagsdaten für eine Station vorhanden waren, konnte eine gleichmäßige Überregung an den berücksichtigten Ereignissen nicht überprüft werden. Diese Problematik wurde aber bei der Bewertung der Simulationsergebnisse während der Modellkalibrierung als mögliche Unsicherheitsquelle berücksichtigt.

Tabelle 4-4: Berücksichtigte Niederschlagsereignisse

| ID | Beginn | Ende | Dauer | Summe [mm] | max. Intensität [mm/h] |
|----|------------------|------------------|-------|------------|------------------------|
| 1 | 13.07.2021 23:35 | 14.07.2021 10:20 | 11:45 | 25,9 | 7,2 |
| 2 | 14.07.2021 15:30 | 14.07.2021 22:05 | 06:35 | 35,2 | 14,1 |
| 3 | 27.07.2021 16:25 | 27.07.2021 18:20 | 01:55 | 21,5 | 21,5 |
| 4 | 01.08.2021 19:40 | 01.08.2021 23:10 | 03:30 | 8,0 | 5,5 |
| 5 | 12.10.2021 12:45 | 12.10.2021 14:00 | 01:15 | 6,0 | 6,0 |

Zur objektiven Bewertung der Qualität der Kalibrierung wurden die in DWA-M 165 (2021) aufgeführten folgenden Gütekriterien verwendet:

- Volumenabweichung VOL (%)
- Maximalwertabweichung DYMAX (%)
- Effizienzkoeffizient E2 (-)

Neben weiteren Erläuterungen und den Gleichungen zur Berechnung der Kriterien wird in dem Merkblatt auch das in Tabelle 4-5 aufgeführte Bewertungsschema bereitgestellt. Dieses Schema wurde auch in dieser Arbeit zur Bewertung der Ergebnisse herangezogen und wird nachfolgend durch die ergänzte Farbgebung bei der Ergebnisdarstellung berücksichtigt.

Tabelle 4-5: Einordnung der Abweichungsmaße in ein Bewertungsschema (DWA-AG ES-2.6, zitiert nach DWA-M 165, (2021))

| Prüfgröße | Einheit | Bewertung | | |
|-----------|---------|-----------|-----------|----------------|
| | | gut | mittel | unbefriedigend |
| VOL | % | < 10 | 10 - 20 | > 20 |
| DYMAX | % | < 15 | 15 - 25 | > 25 |
| E2 | - | > 0,7 | 0,4 - 0,7 | < 0,4 |

Zu Beginn der Kalibrierung wurden verschiedene Modellvarianten untersucht, bei denen jeweils die befestigten und teildurchlässigen Flächenanteile im Untersuchungsgebiet in unterschiedlichen Kombinationen oberflächlich beregnet oder über das Kanalnetz berücksichtigt wurden. Die Bewertung der Simulationsergebnisse mit den zuvor aufgeführten Gütekriterien

zeigte, dass die besten Ergebnisse erzielt werden, wenn die Abflussbildung bei

- Dachflächen,
- öffentlichen Verkehrsflächen und
- privat schmutzigen Flächen

über das Kanalnetzmodell und bei

- privat sauberen Flächen und
- durchlässigen Flächen

über das Oberflächenmodell erfolgt. Die Ergebnisse für diese Modellvariante sind in Tabelle 4-6 dargestellt und zeigen für alle fünf Ereignisse die an den beiden berücksichtigten Messstellen ermittelten Gütekriterien. Anhand der Einfärbung ist zu erkennen, dass die Ergebnisse teilweise in einem unbefriedigenden Bereich liegen und hier noch weiteres Optimierungspotential durch die Kalibrierung der Modellparameter besteht.

Tabelle 4-6: Bewertung der Modellgüte für die beste Modellvariante

| ID | Niederschlags- summe [mm] | Δ Volumen [%] | | Δ Maximalwert [%] | | E2 (NSE) [-] | |
|-------------------------|---------------------------------|----------------------|-------|--------------------------|--------|--------------|-------|
| | | M1 | M3 | M1 | M3 | M1 | M3 |
| 1 | 25,9 | 4,21 | 21,70 | -14,14 | 0,64 | 0,45 | 0,90 |
| 2 | 35,2 | -11,66 | -0,46 | 9,26 | 22,23 | 0,71 | 0,78 |
| 3 | 21,5 | 1,30 | 7,56 | -16,36 | 15,37 | 0,76 | 0,88 |
| 4 | 8,9 | 47,68 | 2,36 | -2,99 | -31,61 | 0,55 | -0,39 |
| 5 | 6,0 | 52,06 | 44,65 | 47,35 | 22,73 | 0,54 | -0,15 |
| Mittelwert | | 18,72 | 15,16 | 4,63 | 5,87 | 0,60 | 0,40 |
| Mittelwert (abs) | | 23,38 | 15,35 | 18,02 | 18,52 | 0,60 | 0,62 |

Die Kalibrierung des Modells wurde manuell durchgeführt, indem die in Tabelle 4-1 und Tabelle 4-3 aufgeführten Parametersätze iterativ angepasst wurden. Im Laufe der Kalibrierung hat sich gezeigt, dass nicht bei allen Gütekriterien für alle Ereignisse ein mittleres oder gutes Ergebnis erzielt werden kann. Dies lässt sich beispielsweise, wie bei dem Maximalwert, durch eine konträre Abweichung zwischen den unterschiedlichen Ereignissen erklären. Eine Verbesserung bei dem einen Ereignis führt somit gleichzeitig zu einer Verschlechterung bei einem anderen. Diese Problematik kann durch verschiedene Unsicherheitsquellen entstanden sein, die mit der zur Verfügung stehenden Datenlage nicht weiter nachvollzogen werden können. Sowohl beim Mittelwert als auch beim Mittelwert der absoluten Abweichungen konnten aber mit den in Anhang A-3 aufgeführten Parametersätzen mittlere bis gute Ergebnisse erzielt werden, weshalb die Kalibrierung an diesem Punkt beendet wurde. Das finale Kalibrierungsergebnis ist in Tabelle 4-7 dargestellt, die zugehörigen Ganglinien sind in Anhang A-3 aufgeführt.

Tabelle 4-7: Bewertung der Modellgüte für das kalibrierte Modell

| ID | Niederschlags- summe [mm] | Δ Volumen [%] | | Δ Maximalwert [%] | | E2 (NSE) [-] | |
|-------------------------|---------------------------------|----------------------|--------|--------------------------|--------|--------------|------|
| | | M1 | M3 | M1 | M3 | M1 | M3 |
| 1 | 25,9 | 0,73 | 25,47 | -17,27 | 11,58 | 0,42 | 0,65 |
| 2 | 35,2 | -13,17 | 5,53 | 4,72 | 27,12 | 0,73 | 0,87 |
| 3 | 21,5 | -4,93 | 1,43 | -25,21 | 9,54 | 0,80 | 0,63 |
| 4 | 8,9 | 16,87 | -12,43 | -30,53 | -43,26 | 0,72 | 0,05 |
| 5 | 6,0 | 15,73 | 19,27 | 10,56 | -0,33 | 0,93 | 0,93 |
| Mittelwert | | 3,05 | 7,85 | -11,55 | 0,93 | 0,72 | 0,63 |
| Mittelwert (abs) | | 10,29 | 12,83 | 17,66 | 18,37 | 0,72 | 0,63 |

4.4 Hydrodynamische Simulationen (Datengenerierung)

Mit dem kalibrierten 1D/2D-Überflutungsmodell sowie den generierten Niederschlagsreihen als Modellbelastung wurden im nächsten Schritt Simulationen durchgeführt, um die für die Entwicklung der KI-basierten Vorhersagemodelle benötigten Trainingsdatensätze zu erzeugen. Der Niederschlag wurde dabei als räumlich homogen angesetzt und für jedes Ereignis wurde eine zusätzliche Nachlaufzeit von 120 Minuten berücksichtigt, damit auch das Ab-schwellen von Überflutungen abgebildet wird. Als Ergebnis lassen sich in MIKE+ sowohl verschiedene Ganglinien als auch Überflutungsflächen als Sequenz ausgeben. Die für die im Rahmen dieser Arbeit zu entwickelnden Vorhersagemodelle benötigten Größen sind dabei:

- Überstauganglinien,
- Durchflussganglinien und
- Wasserstandsganglinien

im Bereich des Kanalnetzes sowie

- Sequenz der Überflutungsflächen

an der Geländeoberfläche. Der schematische Prozess der Datengenerierung ist in Abbildung 4-4 dargestellt. Die gesamte Simulationsdauer für alle 153 natürlichen Niederschlagsereignisse und die 105 Modellregen betrug auf einer Workstation mit einem AMD Ryzen Threadripper 2920X 12-Core Prozessor etwa zwei Monate. In Hinblick auf die Skalierbarkeit der Vorhersagemodelle wird daher die Relevanz von Untersuchungen zur Reduktion der benötigten Trainingsdatenmenge und zur Übertragbarkeit der Modelle deutlich. Wesentlichen Einfluss auf die Berechnungsdauer hat neben dem Überflutungsausmaß insbesondere auch die Länge der zu berechnenden Ereignisse. Als erste Maßnahme zur Reduktion der Rechendauer wurden daher bei den natürlichen Regenreihen bereits Zeiträume zu Beginn und Ende eines Ereignisses beschnitten, die nur über vernachlässigbar geringe Niederschlagsmengen verfügten.

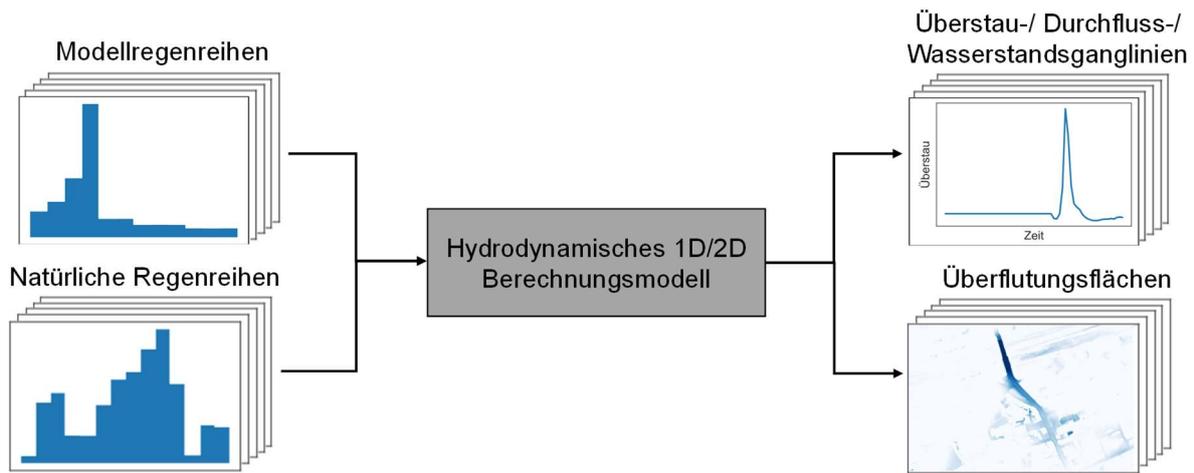


Abbildung 4-4: Schematische Darstellung der Datengenerierung mit den relevanten Ein- und Ausgaben

Die Ausgabe sämtlicher Ganglinien erfolgt in MIKE+ an allen Knoten (Überstau) sowie Haltungen (Durchfluss und Wasserstand) und wird in einer gemeinsamen Datei gespeichert. Diese Eigenschaft, sämtliche Ganglinien automatisiert und gebündelt in einer Datei zu speichern, ist bei Softwareprodukten zur Kanalnetzmodellierung nicht zwingend gegeben, stellt jedoch eine Grundvoraussetzung für das weitere Vorgehen dar. Bei den Überflutungsflächen erfolgt neben der Ausgabe eines Rasters mit den maximalen Wasserständen an der Geländeoberfläche auch die Ausgabe eines multidimensionalen Rasterdatensatzes. Dabei handelt es sich um eine zeitliche Abfolge von Rastern mit den Wasserständen zum jeweiligen Zeitpunkt des simulierten Ereignisses. Dieser Datensatz bietet die Möglichkeit, ein ML-Verfahren darauf zu trainieren, die zeitliche Entwicklung eines bevorstehenden Ereignisses zu prognostizieren und wurde in dieser Arbeit für die weiteren Schritte verwendet.

Die Ausgabe der Berechnungsergebnisse erfolgt bei MIKE+ in den softwarespezifischen Datenformaten `res1d` (Zeitreihen) und `dfs2` (Rasterdatensätze). Dabei handelt es sich um Formate, die nicht ohne Weiteres mit den für die Entwicklung von Maschinellen Lernverfahren in Python verwendeten Modulen kompatibel sind. Daher wurden zunächst die Ganglinien in das CSV-Format und die Rasterdatensätze in das Esri-ASCII-Grid Format konvertiert. Hierzu wurden die beiden Python-Module `MIKE IO 1D` und `MIKE IO` (Andersson et al. 2022) verwendet, die geeignete Konvertierungswerkzeuge zur Verfügung stellen. Anschließend konnten die weiteren Vorverarbeitungsschritte mit den gängigen, in Kapitel 3.4.2 aufgeführten Python-Bibliotheken durchgeführt werden.

4.5 Zusammenfassung und Bewertung der Ergebnisse

Um einen möglichst umfangreichen und repräsentativen Datensatz zu erzeugen, wurden als Belastungsszenarien Modellregenreihen und natürliche Regenreihen von insgesamt acht Stationen mittels partieller Starkregenserien erstellt. Die Modellregenereignisse dienen dabei insbesondere dazu, Extremereignisse mit hohen Wiederkehrzeiten ausreichend abzudecken. Bei den natürlichen Regenereignissen liegt der Fokus hingegen darauf, reale Ereignisverläufe möglichst repräsentativ zu berücksichtigen. Insgesamt umfasst der generierte Datensatz 258

Ereignisse (105 Modellregen, 153 natürliche Regenerereignisse), für die anschließend mit einem hydrodynamischen 1D/2D-Berechnungsmodell die resultierenden Überstauganglinien und Überflutungsflächen berechnet wurden. Dank einer geeigneten Schnittstelle der verwendeten Berechnungssoftware MIKE+ lassen sich sowohl Überstauganglinien als auch die Überflutungsflächen gebündelt und in ein für die weitere Verarbeitung geeignetes Datenformat konvertieren.

Wie bereits von Hofmann und Schüttrumpf (2021) herausgestellt, ist die Generierung der Trainings-, Validierungs- und Testdatensätze mit hydrodynamischen Modellen sehr rechen- und zeitintensiv. Für das in dieser Arbeit verwendete vergleichsweise kleine Testgebiet (ca. 3,1 km²) dauerte die Berechnung aller 258 berücksichtigten Ereignisse auf einer leistungsfähigen Workstation etwa zwei Monate. Bei einer späteren Skalierung auf das gesamte Stadtgebiet von Gelsenkirchen (ca. 105 km²) würden die Rechenzeiten bei Annahme einer linearen Zunahme in einem nicht mehr handhabbaren Bereich liegen. Dementsprechend hoch ist der Bedarf, den Rechenaufwand bei der Datengenerierung zu minimieren. Ansätze hierzu werden in den Kapiteln 5.4 und 6.4 untersucht.

Zudem können bei dem gewählten Systemaufbau (vgl. Kapitel 2.5) die entwickelten Vorhersagemodelle maximal so gut werden wie das für die Datengenerierung verwendete physikalisch-basierte Modell. Dementsprechend wichtig ist bereits eine möglichst hohe Qualität des hydrodynamischen Modells, um die in den Entwicklungsprozess der ML-Modelle eingehenden Unsicherheiten möglichst gering zu halten. Im Bereich des Kanalnetzes wurde zur Verbesserung der Modellgenauigkeit eine Kalibrierung anhand von zwei Messstellen aus einer temporären Messkampagne mit Messungen über einen Zeitraum von sieben Monaten durchgeführt. In Anlehnung an das Bewertungsschema nach DWA-M 165-1 (2021) konnte ein mittleres bis gutes Ergebnis erzielt werden. Es ist aber davon auszugehen, dass mit weiteren Messstellen und längeren Messzeiträumen die Qualität weiter verbessert werden könnte. Hier gilt es, zwischen gewünschter Modellgenauigkeit und erforderlichem Messaufwand einen Kompromiss zu finden.

Im Bereich der Geländeoberfläche waren keine Beobachtungsdaten vorhanden, wodurch eine zusätzliche Kalibrierung nicht möglich war. Daten in diesem Bereich wären besonders wünschenswert, da im Rahmen der Gefährdungsanalyse für Sturzfluten dem Überflutungsmaß an der Geländeoberfläche eine höhere Bedeutung gegenüber der Belastung im Kanalnetz zukommt. Die fehlende Datengrundlage in diesem Bereich stellt jedoch eine generelle Problematik dar, da Systeme zur systematischen Erfassung der Überflutungssituation an der Geländeoberfläche derzeit noch Gegenstand der Forschung sind. Erste Ansätze konzentrieren sich beispielsweise auf die Erfassung von Wasserständen anhand von Bildern aus sozialen Medien (Wang et al. 2018; Chaudhary et al. 2019) oder Aufnahmen von Überwachungskameras (Moy de Vitry et al. 2019). Darüber hinaus wäre eine flächenhafte Information über das Überflutungsmaß wünschenswert, beispielsweise durch die Extraktion aus Satellitendaten (Serpico et al. 2012; Mateo-Garcia et al. 2021). Aufgrund der groben zeitlichen und räumlichen Auflösung beschränkt sich dieser Ansatz momentan aber eher auf fluviale Überschwemmungen oder Sturmfluten. Durch weitere technische Entwicklungen könnte der Ansatz in Zukunft

aber auch für die pluviale Überflutungsmodellierung interessant werden. Dadurch würde einerseits eine Datenbasis zur räumlichen Kalibrierung von hydrodynamischen Modellen an der Geländeoberfläche geschaffen und andererseits könnten diese Daten ggf. sogar direkt zum Training von ML-Modellen eingesetzt werden.

5 Modellentwicklung und -analyse: Überflutungsflächenvorhersage

Gemäß dem in dieser Arbeit berücksichtigten übergeordneten Systemaufbau (vgl. Abbildung 2-7) wurde ein Modell zur Vorhersage von Überstau (Modell 1) und eines zur Vorhersage von Überflutungsflächen (Modell 2) entwickelt. Begonnen wurde mit letzterem, um so zunächst zu prüfen, inwieweit sich die zusätzliche Prognose von Überstau positiv auf die Genauigkeit der Überflutungsflächenvorhersage auswirkt. Da bei der Entwicklung des Modells 2 noch keine Überstauvorhersagen aus dem Modell 1 vorlagen, wurden die Überstaumengen aus der hydrodynamischen Kanalnetzmodellierung als Eingabegröße verwendet. Nachfolgend werden zunächst das zugrunde liegende Lernproblem und die verwendete Datenbasis beschrieben. Anschließend erfolgt eine Beschreibung der Datenvorverarbeitungsschritte, bevor Untersuchungen zum Modellaufbau, Sensitivitätsanalysen und die abschließende Beurteilung der Vorhersagegenauigkeit folgen.

5.1 Problembeschreibung

In diesem Abschnitt wird das dem Vorhersagemodell für Überflutungsflächen zugrundeliegende Lernproblem näher erläutert. Ziel des Modells ist es, aus einer Starkregenprognose die resultierende Überflutungssituation als Zielgröße möglichst verzögerungsarm abzuleiten. Hierzu werden verschiedene potentielle Eingabedaten sowie die prognostizierten Überflutungsflächen aus der hydrodynamischen Kanalnetzmodellierung berücksichtigt, die in den nachfolgenden unterschiedlichen Datenformaten vorliegen:

- 1D-Zeitreihen (Niederschlagsinformationen und Überstauvorhersage):
Hierbei handelt es sich um Zeitreihen, deren Werte auf der zeitlichen Achse variieren, die aber über die räumliche Ausdehnung des Untersuchungsgebietes als konstant angenommen werden.
- 2D-Raster (Räumliche Informationen):
Es handelt sich hierbei um Rasterdatensätze wie beispielsweise ein digitales Geländemodell (DGM), deren Werte über die räumliche Ausdehnung des Untersuchungsgebietes variieren, die über die Zeit jedoch als konstant angenommen werden.
- 3D-Raster-Sequenz (Prognostizierte Überflutungsflächen):
Hierbei handelt es sich um Sequenzen aus Rastern, die das gleiche Format wie Videosequenzen aufweisen. Die Werte variieren sowohl räumlich als auch zeitlich.

Die Berücksichtigung unterschiedlicher Datenarten stellt eine besondere Anforderung an das zu entwickelnde ML-Modell. Damit ist das Lernproblem gegenüber dem Modell zur Überstauvorhersage, bei dem nur 1D-Zeitreihen betrachtet werden (vgl. Kapitel 6), deutlich komplexer. Dadurch schränkt sich die Anzahl der geeigneten ML-Verfahren stark ein. Ferner müssen die Verfahren in der Lage sein, Bild- (2D-Raster) sowie insbesondere auch Videodaten (3D-Raster-Sequenz) effizient zu verarbeiten und darin enthaltene Strukturen zu erkennen. Aus diesem Grund wird der Fokus dieser Arbeit auf Künstlich Neuronale Netze gelegt, die sich bei ähnlichen Problemstellungen wie dem Niederschlags-Nowcasting (u. a. Shi et al. 2015; Ayzel

et al. 2020; Ravuri et al. 2021) oder verschiedenen Verkehrsprognosen (u. a. Ma et al. 2017; Bao et al. 2019; Zhang et al. 2020) als besonders effizient erwiesen haben. Abbildung 5-1 stellt den geplanten Modellaufbau mit den potentiellen Eingaben sowie der prognostizierten Überflutungssituation als Zielgröße dar.

Der Modellaufbau wird als ein überwachtes Lernproblem behandelt, da Eingaben und dazugehörige Zielgrößen klar definiert sind. Darüber hinaus soll das Modell die Zielgrößen als Sequenz von Rastern mit stetigen Werten voraussagen, somit handelt es sich um ein Regressionsproblem. Die genauen Merkmale sowie die Struktur der berücksichtigten Eingaben und der Zielgröße werden nachfolgend im Detail beschrieben.

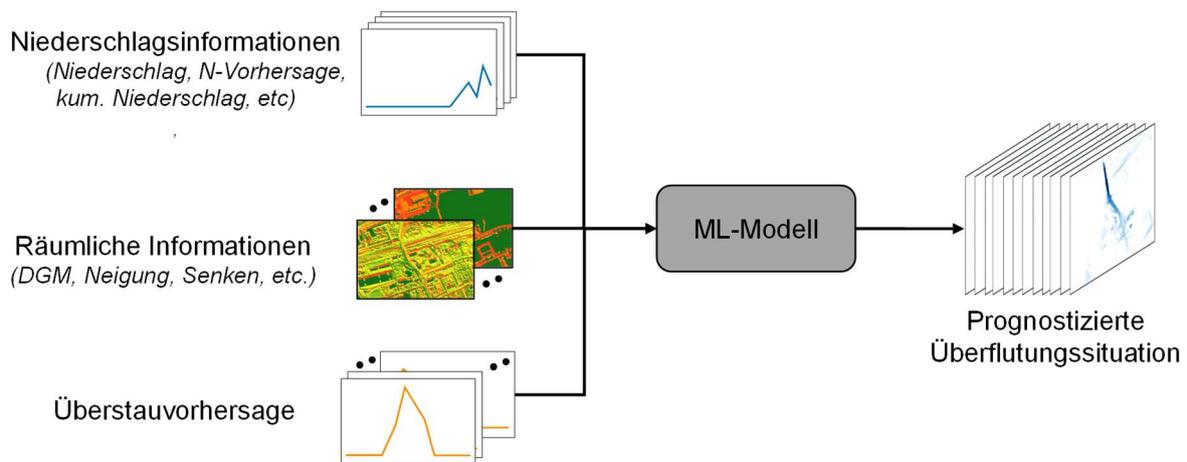


Abbildung 5-1: Geplanter Modellaufbau mit allen potentiellen Eingaben (links) und der gesuchten Zielgröße (rechts)

Niederschlagsinformationen

Bei den Niederschlagsinformationen gehen verschiedene Zeitreihen in das Modell als Eingabe ein. Im Rahmen der Modellentwicklung sind das:

- **Gemessener Niederschlag:**
Die gemessene Niederschlagsintensität der letzten G Zeitschritte in mm/h.
- **Niederschlagsprognose:**
Die prognostizierte Niederschlagsintensität der kommenden H Zeitschritte in mm/h. Dabei handelt es sich im Rahmen der Modellentwicklung zunächst um die gemessene Niederschlagsganglinie der nachfolgenden H Zeitschritte. Im späteren operationellen Betrieb wird diese „perfekte“ Prognose dann durch die Prognose eines Vorhersagemodells ersetzt.
- **Kumulierter Niederschlag:**
Zeitreihe mit dem seit Ereignisbeginn aufsummierten Niederschlag in mm. Der Eingabehorizont umfasst analog zum gemessenen Niederschlag die letzten G Zeitschritte.
- **Verstrichene Zeit:**
Zeitreihe mit der seit Ereignisbeginn erfassten Dauer in Minuten. Der Eingabehorizont umfasst analog zum gemessenen Niederschlag die letzten G Zeitschritte.

Überstauvorhersage

Als Überstauvorhersage wurden in dieser Arbeit zunächst die im Rahmen der Datengenerierung mit dem HD-Modell erzeugten Überstauganglinien verwendet, die die Überstauintensität in l/s angeben. Analog zur Niederschlagsprognose wurde die Überstauintensität für die kommenden H Zeitschritte als Eingabe berücksichtigt. Insgesamt umfasst das Untersuchungsgebiet 975 Schächte, an denen die Überstauganglinien berücksichtigt werden.

Räumliche Informationen

Der Einsatz räumlicher Information als Eingabe bei der ML-basierten Ermittlung von hochwassergefährdeten Gebieten wurde bereits in verschiedenen Studien untersucht (u. a. Pham et al. 2021; Wang et al. 2020). Guo et al. (2020) haben diesen Ansatz auf die ML-basierte Gefährdungsanalyse für urbane Sturzfluten übertragen und verschiedene Geländeinformationen als zusätzliche Eingabe neben Niederschlag verwendet. Löwe et al. (2021) haben auf diesen Ansatz aufgebaut und neben Geländeinformationen auch weitere räumliche Informationen wie den Befestigungsgrad berücksichtigt. Insgesamt wurden in den Untersuchungen elf räumliche Größen berücksichtigt, die sich mit einfachen GIS-basierten Werkzeugen zur Prozessierung von Rasterdaten aus einem digitalen Geländemodell oder Flächennutzungsdaten ableiten lassen. Im Rahmen der Untersuchungen konnte herausgestellt werden, dass mit den Größen Ausrichtung, Krümmung, Befestigungsgrad, Senkentiefe und neigungsgewichtete Abflussakkumulation die besten Ergebnisse erzielt werden. Die Berücksichtigung weiterer Größen führte in den Untersuchungen von Löwe et al. (2021) zu keiner weiteren Verbesserung.

Es muss damit gerechnet werden, dass in Abhängigkeit der örtlichen Gegebenheiten in anderen Untersuchungsgebieten ggf. auch Kombinationen mit mehreren oder anderen Größen etwas bessere Ergebnisse erzielen. Da es sich aber bei dem hier verwendeten Untersuchungsgebiet auch um ein urbanes Einzugsgebiet mit ähnlicher Topografie handelt, ist davon auszugehen, dass dieselben räumlichen Informationen wie in den Untersuchungen von Löwe et al. (2021) besonders relevant sind. Ferner beinhalten die aufgeführten Größen alle für den Anwendungsfall relevanten Faktoren wie Fließwege, Senken aber auch Kanten, die ein Abflusshindernis darstellen. Daher werden in den nachfolgenden Untersuchungen die gleichen räumlichen Informationen als potentielle Modelleingabe genutzt (vgl. Abbildung 5-2). Zur Erzeugung der Datensätze wurden das Geländemodell (vgl. Kapitel 3.2.2) und die Daten zur Oberflächenbeschaffenheit (vgl. Kapitel 3.2.4) sowie die Software ArcGIS Pro (ESRI Inc. 2020) zur Ermittlung der anderen genannten Größen verwendet.

Überflutungsflächenvorhersage

Die Überflutungsflächenvorhersage umfasst die im Rahmen der Datengenerierung erzeugten Überflutungskarten, die im Rasterformat für die einzelnen Zeitschritte eines Ereignisses ausgegeben werden. Die Raster haben wie das im Oberflächenmodell berücksichtigte Rechenetz eine Auflösung von 2 m x 2 m und bestehen aus 772.415 Feldern. Wie bei den anderen Vorhersagen wird auch bei der Überflutungsvorhersage der Horizont mit H Zeitschritten festgelegt.

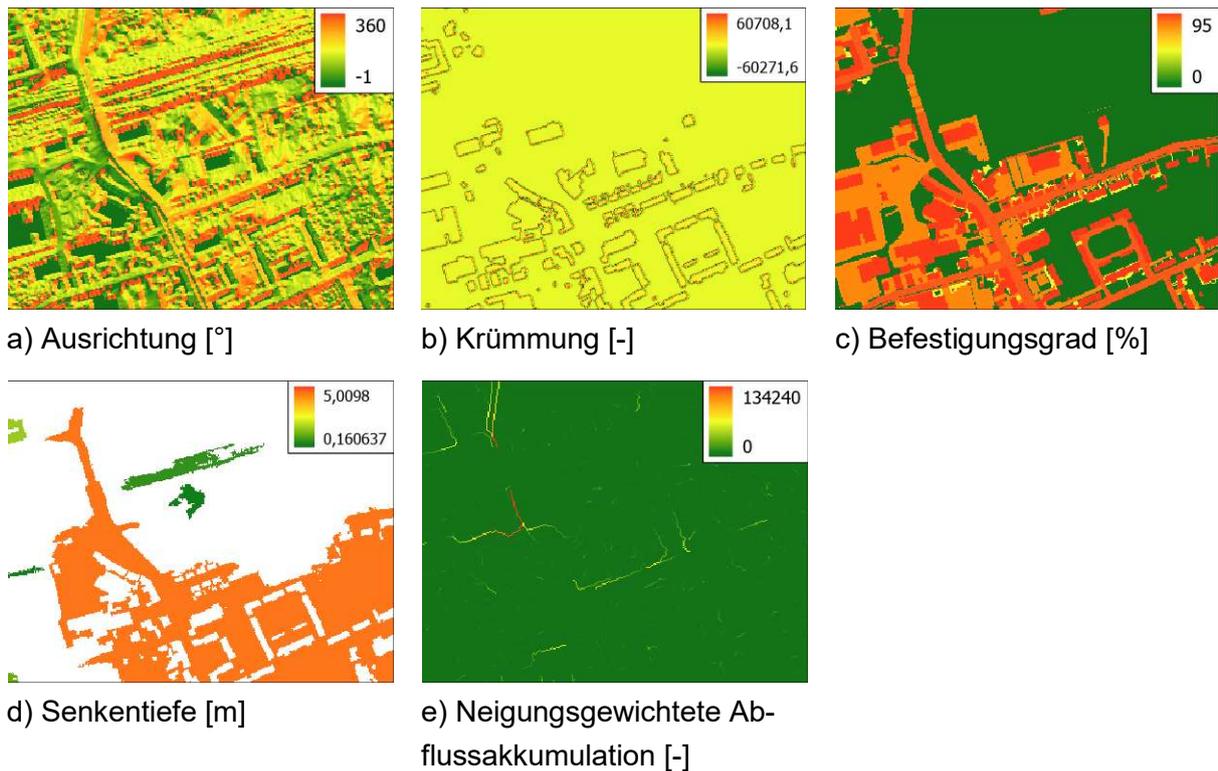


Abbildung 5-2: Darstellung eines Ausschnitts von den als räumliche Eingabedaten berücksichtigten Größen

5.2 Datenvorverarbeitung und -aufteilung

Bevor mit der Modellentwicklung begonnen wurde, erfolgte zunächst die Aufteilung des Datensatzes in Trainings-, Validierungs- und Testdatensatz. Anschließend erfolgte die Vorverarbeitung der Daten, um diese in ein für das Training der ML-Modelle erforderliches Format zu bringen. Das Vorgehen dabei wird nachfolgend nacheinander beschrieben.

Aufteilung des Datensatzes

Für die Modellentwicklung wurde der Trainingsdatensatz in einen Trainings-, einen Validierungs- und einen Testdatensatz aufgeteilt. Der Trainingsdatensatz wird zum Anpassen der Modellparameter während des Trainingsprozesses verwendet. Der Validierungsdatensatz dient hingegen zum Überwachen des Trainingsfortschritts. Dadurch wird in festgelegten Intervallen der Fehler auf dem Validierungsdatensatz ermittelt. Stagniert der Fehler oder steigt er sogar an, wird das Training abgebrochen, um eine Überanpassung des Modells zu vermeiden. Der Testdatensatz wird im Anschluss an den Trainingsprozess zur Evaluierung des trainierten Modells verwendet. Dazu werden Vorhersagen erzeugt, mit denen anschließend die in Kapitel 3.5 aufgeführten Gütekriterien ermittelt werden. So lässt sich die Generalisierbarkeit der Modelle auf einen vom Trainingsprozess unabhängigen Datensatz prüfen.

Im Rahmen der durchgeführten Untersuchungen wurden die natürlichen Regenereignisse der Station Gelsenkirchen-Altstadt als Testdatensatz verwendet. Die Ereignisse der übrigen Stationen sowie die Modellregenereignisse wurden in einem Verhältnis von 90 % Trainingsdaten

und 10 % Validierungsdaten aufgeteilt. Dadurch ergibt sich die in Tabelle 5-1 aufgeführte Anzahl an Ereignissen in den jeweiligen Datensätzen. Um eine möglichst gleichmäßige und zufällige Verteilung der starken und schwachen Ereignisse in Test- und Validierungsdatensatz zu gewährleisten, wurden die Ereignisse vor der Aufteilung durchmischt.

Tabelle 5-1: Aufteilung der Datensätze in Trainings-, Validierungs- und Testdatensätze

| Datensatz | Anzahl der Ereignisse | | | |
|------------------|-----------------------|----------|-------------|--------|
| | Gesamt | Training | Validierung | Testen |
| Modellregen | 105 | 94 | 11 | - |
| Natürliche Regen | 153 | 138 | 15 | 26 |

Skalierung der Daten

Neuronale Netze profitieren beim Trainingsprozess davon, wenn die Daten einheitlich skaliert sind. Das ist insbesondere dann der Fall, wenn die einzelnen verwendeten Merkmale unterschiedliche Einheiten aufweisen, wie bei dem hier vorliegendem Lernproblem. Die Verwendung der Originalskala kann den Merkmalen mit einem großen Wertebereich in diesem Fall mehr Gewicht verleihen, während der Einfluss von Merkmalen mit einem niedrigen Wertebereich verschwindend klein ist. Zur Skalierung der Daten kommen bevorzugt zwei Verfahren zum Einsatz. Zum einen die Standardisierung der Daten und zum anderen die Normalisierung der Daten durch Skalierung auf das Intervall $[0,1]$. Im Rahmen der Bilderverarbeitung ist die Normalisierung der Daten üblich und kam aufgrund der ähnlichen Datenformate auch hier zur Anwendung. Zur Ermittlung des normalisierten Wertes x' wird dabei von einem betrachteten Wert x der kleinste im Datensatz enthaltene Wert abgezogen. Das Ergebnis wird anschließend durch die Differenz von Minimal- $\min(x)$ und Maximalwert $\max(x)$ geteilt:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (5-1)$$

Diese Art der Normalisierung wurde manuell für die verschiedenen Eingabemerkmale sowie die Zielgröße umgesetzt. Dazu wurden Minimal- und Maximalwert am Trainingsdatensatz bestimmt und anschließend für die Skalierung des gesamten Datensatzes gemäß Formel (5-1) verwendet.

Erzeugung von Trainingsbeispielen

Wie eingangs erwähnt, wird das Vorhersageproblem als überwachtes Lernproblem betrachtet. Bei dieser Art von Lernproblem erfolgt das Training immer auf Grundlage von Pärchen aus Eingaben und dazugehörigen Zielgrößen. Für die Eingaben werden im Trainingsprozess Ausgaben berechnet, für die dann gemeinsam mit den Zielgrößen das Fehlersignal berechnet wird, das zur Anpassung der Modellparameter dient. Die Berücksichtigung der räumlichen Informationen beim Bilden der Pärchen ist recht einfach, da diese statisch sind und sich nicht verändern. Für die Zeitreihen und Rastersequenzen, die sich beide entlang der zeitlichen Dimension dynamisch ändern, gilt das dagegen nicht. Hier wird auf einen gleitenden Fenster-Ansatz zurückgegriffen. Dabei wird für jeden Betrachtungszeitpunkt t eines Ereignisses ein Fenster über die vergangenen G Zeitschritte und die kommenden H Zeitschritte aufgespannt,

sodass sich die Intervalle $[t_{-G+1}, \dots, t]$ für die vergangenen und $[t+1, \dots, t_H]$ für die prognostizierten Zeitschritte ergeben. G und H werden bei den hier durchgeführten Untersuchungen jeweils mit 60 Minuten festgelegt, was bei der gewählten zeitlichen Auflösung von fünf Minuten zwölf Zeitschritten entspricht. Das Vorgehen zur Erzeugung der Trainingspärchen P ist exemplarisch für einen Betrachtungszeitpunkt in Abbildung 5-3 dargestellt. Wenn ein Ereignis aus n Zeitschritten besteht, kann dementsprechend das erste Trainingspärchen zum Zeitpunkt $t = G$ und das letzte Trainingspärchen zum Zeitpunkt $t = n - H$ gebildet werden. Bei einer festgelegten Schrittweite von 1 ergeben sich so $m = n - (G + H)$ Trainingsbeispiele. Die Überstauganglinien werden dabei an k Schächten zur Verfügung gestellt.

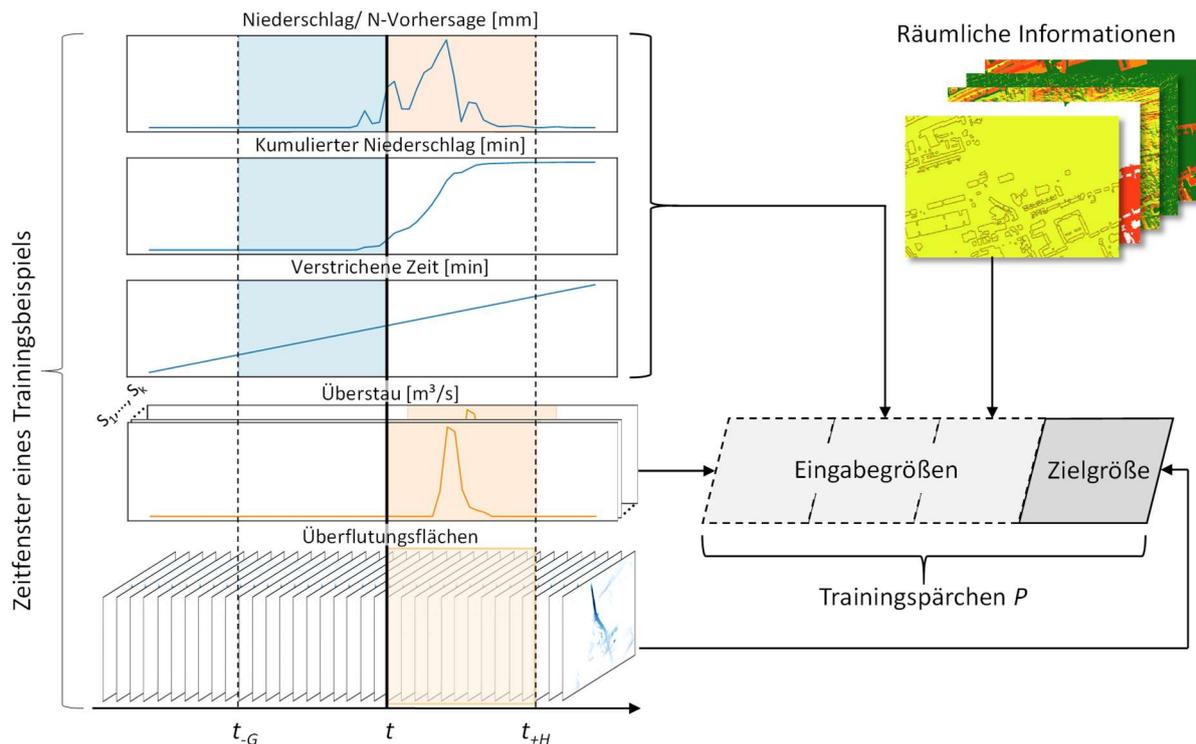


Abbildung 5-3: Umwandlung der Daten in ein überwachtes Lernproblem am Beispiel eines einzelnen Trainingspärchens

5.3 Untersuchungen zum Modellaufbau

Im Rahmen des Modellentwicklungsprozesses wurden verschiedene Untersuchungen durchgeführt, um die optimale Modellarchitektur zu ermitteln. Wie im Rahmen der Problembeschreibung erwähnt (vgl. Kapitel 5.1), wurden Künstlich Neuronale Netze als einziges Lernverfahren berücksichtigt. Die durchgeführten Untersuchungen waren daher nicht auf den Vergleich unterschiedlicher Lernverfahren, sondern vielmehr auf die Wahl geeigneter Eingabedatenströme, Datenformate und Netzarchitekturen ausgelegt und werden nachfolgend in Anlehnung an (Burrichter et al. 2023) beschrieben. Um den Lesefluss nicht zu unterbrechen, wird im Text auf eine detaillierte Erläuterung der verschiedenen Netzbestandteile, Hyperparameter und anderer Begriffe im Kontext von Neuronalen Netzen verzichtet. Stattdessen wird an dieser Stelle auf Anhang A-1 verwiesen.

Zu Beginn der Untersuchungen wurde die in Abbildung 5-4 dargestellte Architektur als Basis

gewählt. Dabei wurden zunächst nur Niederschlagsinformationen als Eingabe berücksichtigt, um eine Sequenz aus Überflutungsrastern vorherzusagen. Die Architektur wurde inspiriert durch die Arbeit von Guo et al. (2020), weist aber diverse Änderungen auf. Die Verarbeitung der Niederschlagsinformationen erfolgt zunächst mit zwei *Convolutional-1D*-Schichten zur Extraktion von Merkmalen, bevor eine vollständig verbundene Schicht und eine Umformschicht folgen. Letztere haben die Aufgabe, die Daten in ein Format zu bringen, das sich auf das Ausgabeformat hochskalieren lässt. Anschließend folgt ein *Decoder*, bestehend aus vier *Deconvolutional-3D*-Schichten, die aus den extrahierten Merkmalen die Überflutungsrastersequenz erzeugen. Andere Architekturen, wie LSTM-Schichten oder vollständig verbundene Schichten zur Merkmalsextraktion oder *Convolutional-3D*-Schichten in Kombination mit *Up-sampling*-Schichten zum Dekodieren wurden ebenfalls getestet, führten aber zu schlechteren Ergebnissen. Allen *Convolutional*- und *Deconvolutional*-Schichten außer der letzten folgt eine *Batch*-Normalisierungsschicht (Ioffe und Szegedy 2015) zur Stabilisierung des Trainingsprozesses und um höhere Lernraten zu ermöglichen sowie eine ReLU-Aktivierungsfunktion (vgl. Abbildung 3-4). Der letzten *Deconvolutional*-Schicht folgt eine Sigmoid-Aktivierungsfunktion (vgl. Abbildung 3-4) ohne *Batch*-Normalisierung.

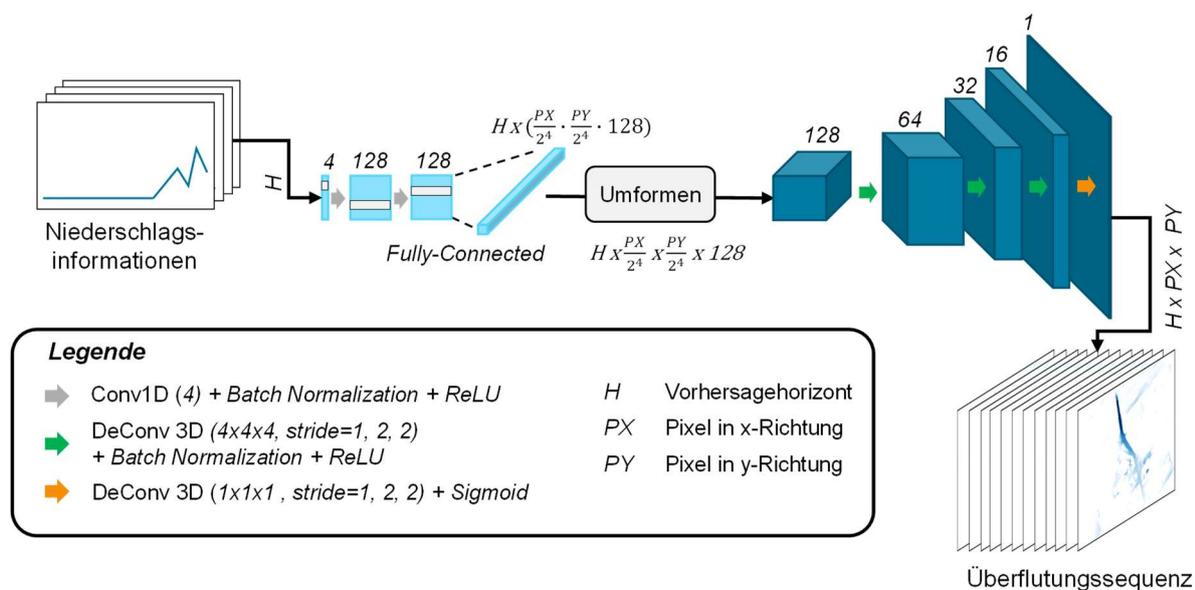


Abbildung 5-4: Ausgangsarchitektur für die nachfolgenden Untersuchungen (Die Zahlen über den einzelnen Schichten geben jeweils die Anzahl der Merkmale bzw. Merkmalskarten an, die durch die Anzahl der berücksichtigten Filter der jeweils zuvor durchgeführten Faltungsfunktion (als Pfeil dargestellt) festgelegt wird. Bei der ersten Schicht bezieht sich der Wert auf die Anzahl der Eingabemerkmale.)

Für das Training nutzen alle Modelle, soweit nicht anders beschrieben, den *Mean-Squared-Error* als Zielfunktion und wurden mit dem *Adaptive Moment Estimation* (ADAM)-Optimierungsalgorithmus (Kingma und Ba 2014) für 100 Epochen trainiert. Der ADAM-Algorithmus basiert auf dem in Kapitel 3.3.1 beschriebene Gradientenverfahren, dass zur Optimierung der Fehlerfunktion während des Trainingsprozesses zum Einsatz kommt, ist jedoch deutlich effizienter und ermöglicht eine schnellere Konvergenz. Die Größe der Eingabestapel wurde auf 8 Pärchen festgelegt, da größere Stapel teilweise zu einer Überlastung des GPU-Speichers führen. Als Lernrate wurde jeweils ein Wert von 0,001 festgelegt, der zuvor in Anlehnung an das

in Smith (2015) beschriebene Vorgehen ermittelt wurde. Um eine Überanpassung zu vermeiden, wurde das Modell nicht nach der letzten Epoche gespeichert, sondern nach der Epoche mit dem geringsten Fehler für den Validierungsdatensatz.

5.3.1 Untersuchung 1: Vergleich unterschiedlicher Eingabegrößen

In einer ersten Untersuchung wurde überprüft, mit welcher Kombination der potentiellen Eingaben das Modell die besten Ergebnisse liefert. Dabei wurden die Niederschlagsinformationen, die auch die Niederschlagsvorhersage enthalten, als zwingend erforderlich angesehen. Die übrigen beiden Eingaben wurden jeweils in allen möglichen Kombinationen variiert, sodass folgende Modelle mit den jeweiligen Eingaben miteinander verglichen wurden:

- Modell 1: Niederschlag
- Modell 2: Niederschlag + Überstauvorhersage
- Modell 3: Niederschlag + Räumliche Informationen
- Modell 4: Niederschlag + Überstauvorhersage + Räumliche Informationen

Abbildung 5-5 zeigt die Ausgangsarchitektur mit den zusätzlich berücksichtigten Eingabepfaden zur Merkmalsextraktion. Die Überstauvorhersage wird analog zu den Niederschlagsinformationen verarbeitet und nach der Umformschicht über eine *Concatenate*-Schicht mit der Ausgangsarchitektur verbunden, bevor dann dort der Dekodier-Pfad folgt. An der gleichen Stelle werden auch die räumlichen Informationen in das Netz eingebunden. Hier erfolgt die Merkmalsextraktion allerdings mit einer *Encoder*-Struktur bestehend aus mehreren *Convolutional-2D*-Schichten. Zur Verringerung der Auflösung der Eingaberaster (*downsampling*) wird bei den einzelnen Schichten die Schrittweite (*stride*) auf zwei gesetzt. Um vor der *Concatenate*-Schicht die gleichen Dimensionen zu erhalten, wird die Ausgabe der letzten *Convolutional-2D*-Schicht *H*-mal gestapelt.

Zur Ergebnisbewertung wurden der CSI und der RMSE herangezogen (vgl. Kapitel 3.5). Beide Metriken wurden im Rahmen der Auswertung nur für Pixel berechnet, in denen für den jeweiligen Vorhersagezeitpunkt entweder das HD-Modell oder das Neuronale Netz Wasserstände über einen Grenzwert aufweist. In Anlehnung an das Vorgehen in Löwe et al. (2021) oder wie auch bei der Niederschlagsvorhersage üblich (Shi et al. 2017) wurden für den CSI mehrere Grenzwerte berücksichtigt, um die Lagegenauigkeit in Abhängigkeit verschiedener Wasserstände zu bewerten. Das gleiche Vorgehen wurde hier auch für den RMSE verwendet, um auch die Abweichung in Abhängigkeit unterschiedlicher Wasserstände berücksichtigen zu können.

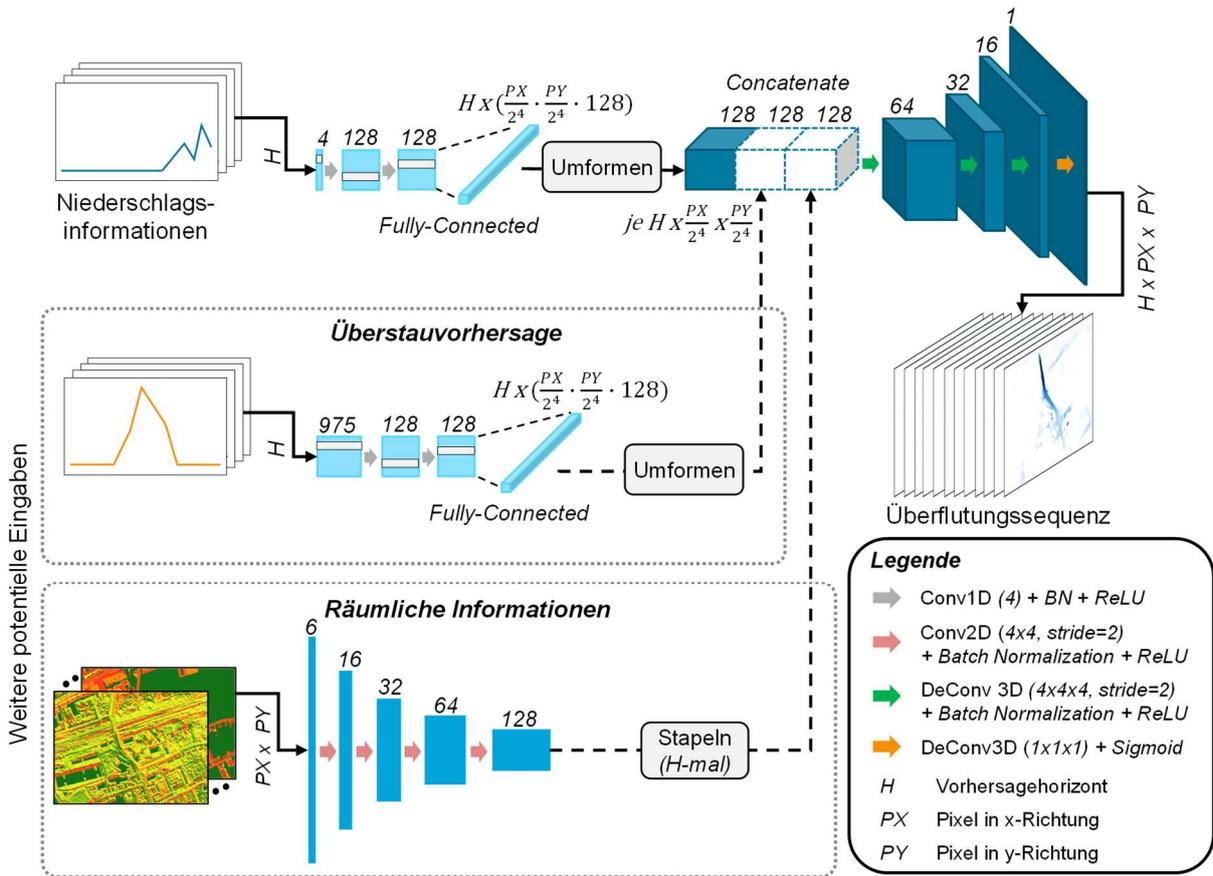


Abbildung 5-5: Ausgangsarchitektur mit den jeweiligen Eingabepfaden für die Überstauvorhersage und die räumlichen Informationen (Die Zahlen über den einzelnen Schichten geben jeweils die Anzahl der Merkmale bzw. Merkmalskarten an, die durch die Anzahl der berücksichtigten Filter der jeweils zuvor durchgeführten Faltungsfunktion (als Pfeil dargestellt) festgelegt wird. Bei den ersten Schichten bezieht sich der Wert auf die Anzahl der Eingabemerkmale.)

Tabelle 5-2 fasst die Ergebnisse zusammen. Die Metriken für die einzelnen Grenzwerte d [m] wurden dabei jeweils als Mittelwert aller *Samples* der 26 Testereignisse und aller Vorhersagezeitpunkte gebildet. Es zeigt sich deutlich, dass die Berücksichtigung der Überstauvorhersage (Modell 2 und 4) zu besseren Ergebnissen führt und insbesondere hohe Wasserstände so deutlich verlässlicher vorhersagt werden können. Die zusätzliche Berücksichtigung von räumlichen Informationen führt zu einer etwas höheren Genauigkeit bei Grenzwerten $\leq 0,2$ m. Bei höheren Grenzwerten schneidet hingegen das Modell 2 ohne räumliche Informationen am besten ab. Insgesamt ist der Unterschied zwischen beiden Modellen gering. Vor dem Hintergrund, dass die räumlichen Informationen als statische Variable ohne Änderungen zwischen einzelnen Trainingspärchen berücksichtigt werden, ist das Ergebnis nicht überraschend. Damit fungiert der Datensatz eher als Maske und hat keinen wesentlichen Einfluss auf das für den Trainingsfortschritt erforderliche Fehlersignal. Da verschiedene Untersuchungen aber gezeigt haben, dass räumliche Informationen die Übertragbarkeit trainierter Modelle ermöglichen können (Löwe et al. 2021; Guo et al. 2022), wurde Modell 4 im weiteren Verlauf berücksichtigt.

Tabelle 5-2: Ergebnis der Evaluierung zum Vergleich unterschiedlicher Modelleingaben (für jede Untersuchung und Metrik wurde das beste Ergebnis fett markiert)

| Modell | RMSE [m]↓ | | | | | CSI [-]↑ | | | | |
|--|-------------------|--------------|--------------|--------------|--------------|-------------------|--------------|--------------|--------------|--------------|
| | Grenzwert d [m] | | | | | Grenzwert d [m] | | | | |
| | ≥0,02 | ≥0,05 | ≥0,1 | ≥0,2 | ≥0,5 | ≥0,02 | ≥0,05 | ≥0,1 | ≥0,2 | ≥0,5 |
| Modell 1 (Eingaben: N) | 0,039 | 0,052 | 0,074 | 0,140 | 0,553 | 0,504 | 0,488 | 0,399 | 0,299 | 0,122 |
| Modell 2 (Eingaben: $N + \ddot{U}$) | 0,028 | 0,037 | 0,052 | 0,096 | 0,096 | 0,538 | 0,543 | 0,495 | 0,414 | 0,768 |
| Modell 3 (Eingaben: $N + RI$) | 0,037 | 0,050 | 0,074 | 0,144 | 0,547 | 0,510 | 0,466 | 0,384 | 0,293 | 0,157 |
| Modell 4 (Eingaben: $N + RI + \ddot{U}$) | 0,026 | 0,035 | 0,051 | 0,092 | 0,118 | 0,595 | 0,574 | 0,511 | 0,421 | 0,746 |
| mit: | | | | | | | | | | |
| N Niederschlag | | | | | | | | | | |
| \ddot{U} Überstau | | | | | | | | | | |
| RI Räumliche Informationen | | | | | | | | | | |

5.3.2 Untersuchung 2: Vergleich unterschiedlicher Vorverarbeitung der Überstaudaten

In einem zweiten Schritt wurde untersucht, in welchem Format die Überstaudaten am besten in das Modell eingebunden werden können. Ursprünglich liegen diese als Ganglinien für alle Knoten im Einzugsgebiet vor. Hier stellt sich die Frage, inwieweit das Modell in der Lage ist, aus diesen mehreren 100 Ganglinien ohne zusätzlichen räumlichen Bezug einen Mehrwert zu generieren. Vor diesem Hintergrund wurden neben den unstrukturierten Überstauganglinien (Variante a) zwei zusätzliche Varianten untersucht, die Überstaudaten als Rastersequenz (Variante b) oder als *spatio-temporal Graph* (Variante c) in das Modell zu geben. Abbildung 5-6 liefert einen Überblick über die möglichen Architekturen.

Die Rastersequenzen in Variante b) wurden durch die Verschneidung der Überstauganglinien mit einem Senkeneinzugsgebiets-Raster erstellt (vgl. Abbildung 5-7). Im Ergebnis entsteht so eine Sequenz aus Rastern mit den akkumulierten Überstauvolumina aller Schächte je Zeitschritt und je Senkeneinzugsgebiet. Die Rastersequenzen werden mit einer *Encoder*-Struktur bestehend aus *Convolutional*-Schichten mit einer Schrittweite von zwei verarbeitet. Wie in dem Dekodier-Pfad folgt auch hier nach jeder Schicht eine *Batch*-Normalisierungsschicht und eine ReLU-Aktivierungsfunktion. Die Ausgabe der letzten Schicht wird dann mit der Ausgangsarchitektur verbunden und dort weiterverarbeitet.

Bei Variante c) wird die Überstauvorhersage mit einem T-GCN (vgl. Kapitel 3.3.2.3) in Anlehnung an Yu et al. (2017) und Zhao et al. (2018) verarbeitet. Dabei wird eine *Graph Convolutional*-Schicht mit einer rekurrenten Schicht kombiniert. Während die *Graph Convolutional*-Schicht die räumlichen Abhängigkeiten des Kanalnetzes berücksichtigt, erfasst die rekurrente Schicht die zeitliche Dynamik des Überstauvorgangs an den einzelnen Schächten. Damit wird die Modellierung des hier vorliegenden raumzeitlichen Lernproblems ermöglicht. Die *Graph Convolutional*-Schicht erhält als Eingabe eine Merkmalsmatrix mit den Überstauganglinien und eine Adjazenzmatrix. Die Adjazenzmatrix (vgl. Kapitel 3.3.2.3) bildet das Kanalnetz in Form

eines ungewichteten und gerichteten Graphen ab. Als rekurrente Schicht wird hier eine LSTM-Schicht verwendet. Die Ausgabe des T-GCN-Blocks wird dann analog zu den Niederschlagsinformationen an eine vollständig verbundene Schicht, gefolgt von einer Umformschicht übergeben, bevor diese dann mit der Ausgangsarchitektur verbunden wird.

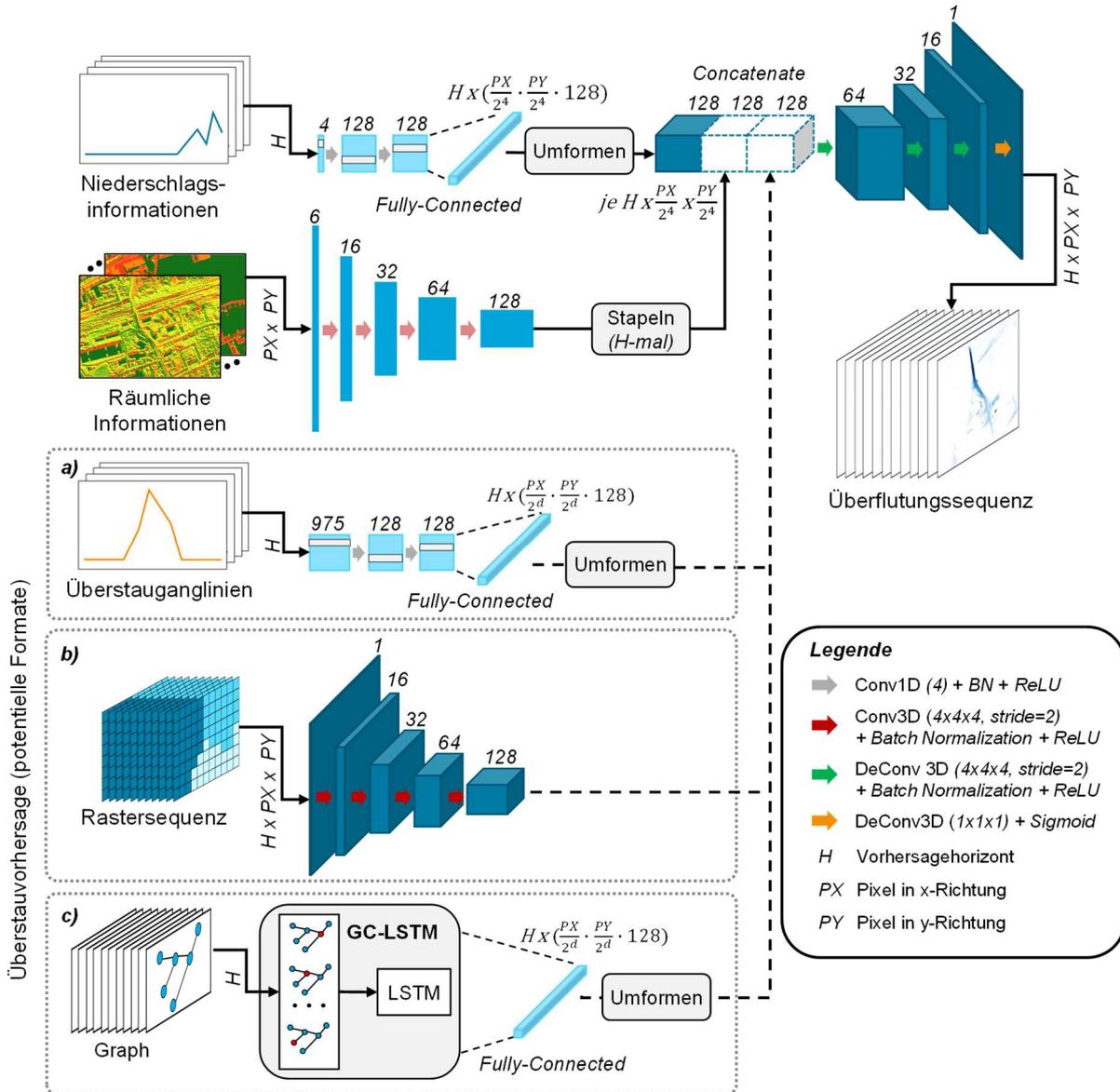


Abbildung 5-6: Ausgangsarchitektur mit den verschiedenen Eingabepfaden für die unterschiedlichen Formate der Überstauvorhersage (Die Zahlen über den einzelnen Schichten geben jeweils die Anzahl der Merkmale bzw. Merkmalskarten an, die durch die Anzahl der berücksichtigten Filter der jeweils zuvor durchgeführten Faltungsfunktion (als Pfeil dargestellt) festgelegt wird. Bei der ersten Schicht bezieht sich der Wert auf die Anzahl der Eingabemerkmale.)

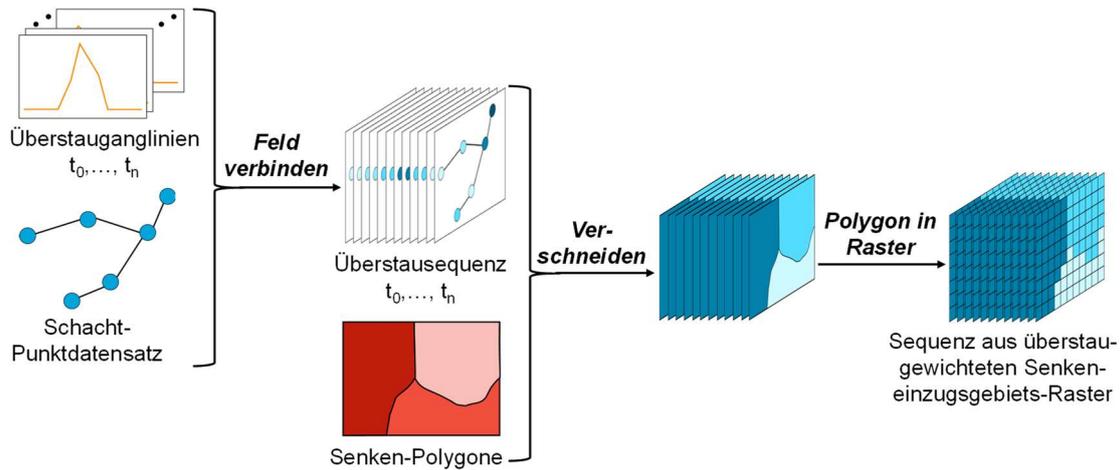


Abbildung 5-7: Schematische Vorgehensweise zur Erzeugung der Überstaurastersequenzen

Der Vergleich der unterschiedlichen Modellaufbauten zeigt, dass die Berücksichtigung der Überstauinformation als Rastersequenz zu den schlechtesten Ergebnissen führt (vgl. Tabelle 5-3). Weiterhin führt das Format zu einem deutlich größeren Speicherbedarf beim Modeltraining und fast doppelt so langen Rechenzeiten. Die unstrukturierte Eingabe der Überstauganglinien schneidet bei den niedrigen Grenzwerten etwas besser ab, wobei die Unterschiede insbesondere beim RMSE nur marginal sind. Bei den höheren Grenzwerten, die hier besonders relevant sind, liefert die Eingabe als Graph die besten Ergebnisse. Daher wird das Modell 3 in die letzte Untersuchung mitgeführt.

Tabelle 5-3: Ergebnis der Evaluierung zum Vergleich unterschiedlicher Formate der Überstauvorhersage (für jede Untersuchung und Metrik wurde das beste Ergebnis fett markiert)

| Modell | RMSE [m]↓ | | | | | CSI [-]↑ | | | | |
|-----------------------------|-------------------|--------------|--------------|--------------|--------------|-------------------|--------------|--------------|--------------|--------------|
| | Grenzwert d [m] | | | | | Grenzwert d [m] | | | | |
| | ≥0,02 | ≥0,05 | ≥0,1 | ≥0,2 | ≥0,5 | ≥0,02 | ≥0,05 | ≥0,1 | ≥0,2 | ≥0,5 |
| Modell 1 (Ungeordnet) | 0,026 | 0,035 | 0,051 | 0,092 | 0,118 | 0,595 | 0,574 | 0,511 | 0,421 | 0,746 |
| Modell 2 (Rastersequenz) | 0,030 | 0,040 | 0,058 | 0,115 | 0,148 | 0,548 | 0,514 | 0,434 | 0,340 | 0,679 |
| Modell 3 (Graph) | 0,026 | 0,036 | 0,052 | 0,092 | 0,081 | 0,575 | 0,557 | 0,492 | 0,424 | 0,788 |

5.3.3 Untersuchung 3: Vergleich der bisher besten Modellstruktur (T-GCN) mit einem cGAN

In einer dritten Untersuchung wurde die Performance des zuvor am besten bewerteten Modell, dass nachfolgend als T-GCN bezeichnet wird, einem *conditional Generative Adversarial Network* (cGAN) gegenübergestellt. Der Aufbau des cGANs ist in Abbildung 5-8 dargestellt und ist inspiriert durch die Arbeiten von Isola et al. (2017) und Hofmann und Schüttrumpf (2021). Letztere haben die Architektur bereits erfolgreich zur Überflutungsvorhersage eingesetzt. Anders als ein normales GAN erhält das cGAN neben Rauschen zusätzlichen Kontext als Eingabe. In den vorliegenden Untersuchungen wurde in Anlehnung an die Erkenntnisse aus Isola et al. (2017) sogar vollständig auf Rauschen als Eingabe verzichtet und lediglich Kontext in

Form der potentiellen Modelleingaben aus Untersuchung 1 berücksichtigt. Zudem wurde analog zu dem Vorgehen in Isola et al. (2017) und Hofmann und Schüttrumpf (2021) eine *Mean Absolute Error*-Fehlerfunktion (L1 loss) in die Zielfunktion integriert. Dadurch zielt der Generator nicht nur darauf ab, den Diskriminator zu täuschen, sondern gleichzeitig auch den Fehler zwischen der Netzausgabe und der Zielgröße, den „realen Überflutungsflächen“, die hier der Ausgabe des HD-Modells entsprechen, zu minimieren.

Als Architektur für den Generator wird mit dem T-GCN das beste Modell aus den Untersuchungen 1 und 2 verwendet. Als Diskriminator wird eine zur Klassifizierung geeignete Netzstruktur verwendet, die zunächst analog zu den übrigen Modellstrukturen aus den einzelnen Eingaben die Merkmale extrahiert und anschließend mit einer *Concatenate*-Schicht zusammenführt (vgl. Abbildung 5-9). Anschließend folgt noch ein *Convolutional 3D*-Block mit ReLU und *Batch*-Normalisierung sowie eine *Convolutional 3D*-Schicht gefolgt von einer Sigmoid-Aktivierungsfunktion. Die Ausgabe entspricht einer binären Klassifizierung. Anders als im Generator wurde im Diskriminator *dropout* (Srivastava et al. 2014) mit einer *dropout*-Rate von 0,5 zur Regularisierung verwendet. Dadurch konnten der Trainingsprozess stabilisiert und bessere Ergebnisse erzielt werden.

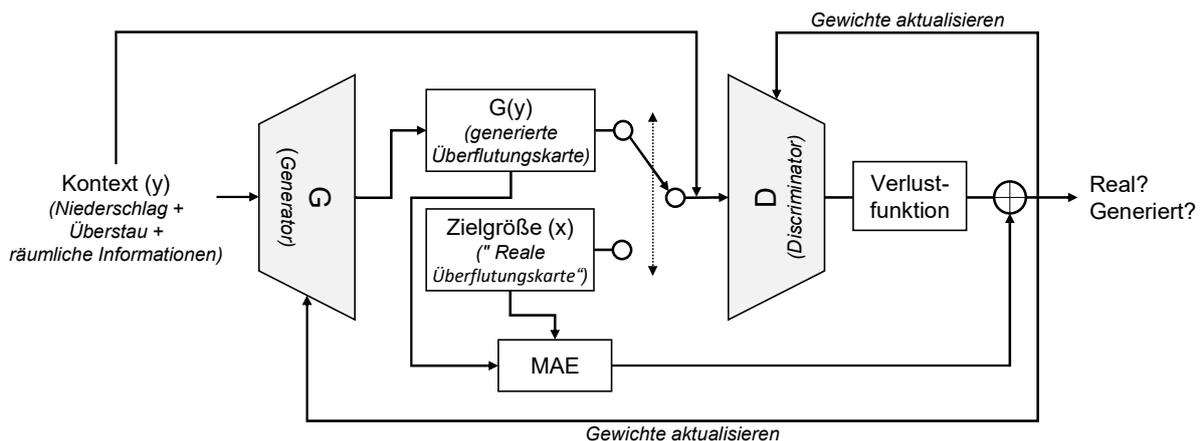


Abbildung 5-8: Aufbau und Funktionsweise des conditional GANs (durch den gestrichelten Pfeil soll verdeutlicht werden, dass der Diskriminator abwechselnd eine durch den Generator generierte Überflutungskarte und eine „reale Überflutungskarte“ als Eingabe erhält)

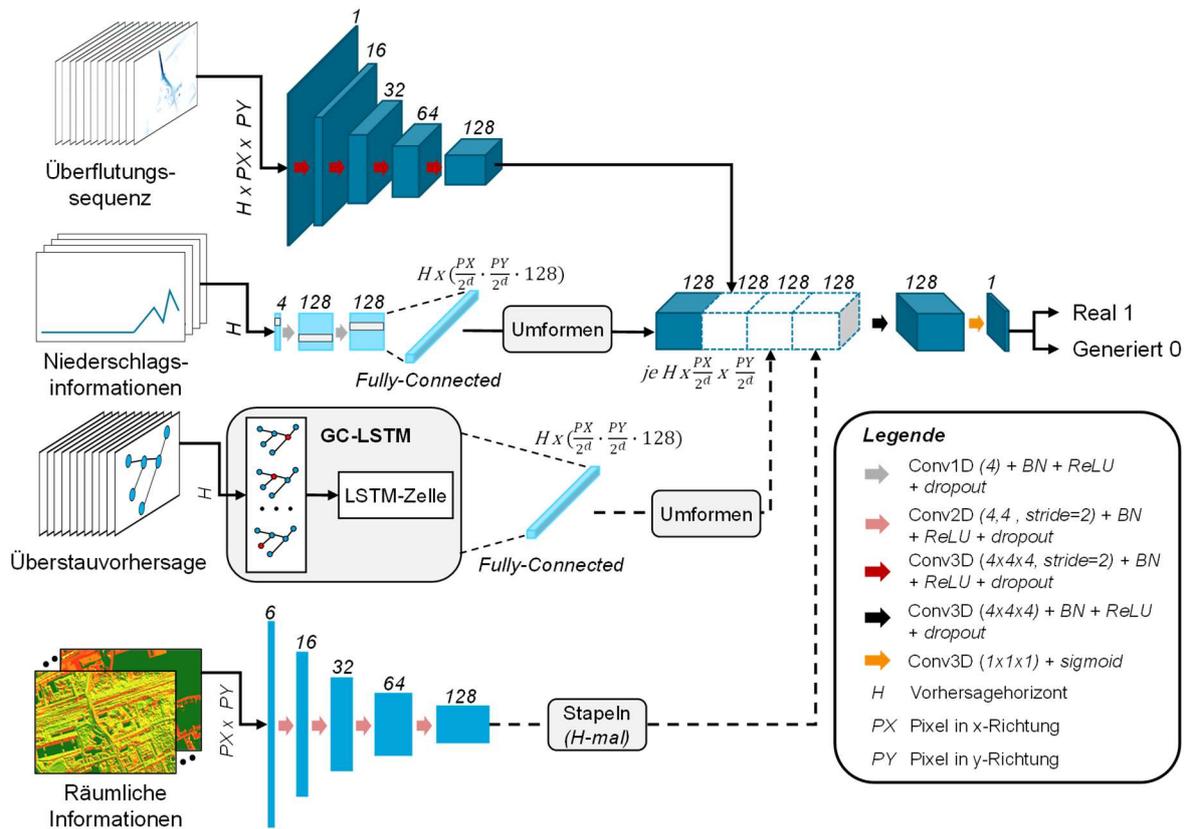


Abbildung 5-9: Aufbau des Diskriminators

Beim Vergleich beider Modelle in Tabelle 5-4 zeigt sich eine hohe Lagegenauigkeit (CSI) bei dem Modellaufbau als cGAN (Modell 2). Das T-GCN (Modell 1) zeigt hingegen eine höhere Genauigkeit beim RMSE. Ferner weist das T-GCN auch beim CSI für den höchsten Grenzwert von $\geq 0,5$ m ein etwas besseres Ergebnis auf. Damit schneidet das Modell bei der Evaluierung insgesamt betrachtet etwas besser ab. Zudem ist das Training von cGANs gegenüber vorwärtsgerichteten Netzen größeren Instabilitäten ausgesetzt. Diese ergeben sich aufgrund verschiedener Faktoren, für die an dieser Stelle auf die entsprechende Literatur verwiesen wird (Goodfellow et al. 2014; Salimans et al. 2016; Arjovsky et al. 2017). Aus diesen beiden Gründen wird für die nachfolgende Beurteilung der Vorhersagegenauigkeit und die nachfolgenden Sensitivitätsanalysen auf das T-GCN als das am besten geeignete Modell zurückgegriffen.

Tabelle 5-4: Ergebnis der Evaluierung zum Vergleich unterschiedlicher Modellarchitekturen (für jede Untersuchung und Metrik wurde das beste Ergebnis fett markiert)

| Modell | RMSE [m]↓ | | | | | CSI [-]↑ | | | | |
|---------------------|-------------------|--------------|--------------|--------------|--------------|-------------------|--------------|--------------|--------------|--------------|
| | Grenzwert d [m] | | | | | Grenzwert d [m] | | | | |
| | $\geq 0,02$ | $\geq 0,05$ | $\geq 0,1$ | $\geq 0,2$ | $\geq 0,5$ | $\geq 0,02$ | $\geq 0,05$ | $\geq 0,1$ | $\geq 0,2$ | $\geq 0,5$ |
| Modell 1 (T-GCN) | 0,026 | 0,036 | 0,052 | 0,092 | 0,081 | 0,575 | 0,557 | 0,492 | 0,424 | 0,788 |
| Modell 2 (cGAN) | 0,027 | 0,037 | 0,055 | 0,113 | 0,158 | 0,623 | 0,602 | 0,545 | 0,440 | 0,723 |

5.4 Sensitivitätsanalysen

Die Praxistauglichkeit des entwickelten Modells hängt maßgeblich von dessen Skalierbarkeit ab. In diesem Kapitel wird daher untersucht, inwieweit sich die Trainingsdatenmenge reduzieren und sich das trainierte Modell übertragen lässt, um die Skalierbarkeit des finalen Modells zu verbessern. Zudem folgen anschließend Untersuchungen zum Einfluss von Hyperparametern, um die Genauigkeit des Vorhersagemodells weiter zu optimieren. Analog zu Kapitel 5.3 wird auch hier für tiefergehende Erläuterungen zu spezifischen Begriffen im Kontext von Neuronalen Netzen auf Anhang A-1 verwiesen.

5.4.1 Einfluss der Trainingsdatenquantität und -verteilung

Der in Kapitel 4.4 beschriebene Prozess der Datengenerierung hat einen maßgeblichen Einfluss auf die erforderlichen Rechenressourcen. Die benötigte Rechendauer von etwa zwei Monaten zur Simulation aller berücksichtigten Niederschlagsereignisse mit dem HD-Modell ist für das berücksichtigte Untersuchungsgebiet (3,1 km² Fläche) noch akzeptabel. Die Berechnung des gesamten Stadtgebietes von Gelsenkirchen (105 km² Fläche) würde unter der Annahme, dass die Rechendauer linear mit der Fläche anwächst und identische Rechenressourcen verwendet werden, über fünf Jahre dauern. Damit ist der vorgestellte Ansatz derzeit nicht praxistauglich. Auch wenn davon auszugehen ist, dass die Rechenleistung in den kommenden Jahren weiter zunehmen wird, verdeutlicht das Beispiel den Bedarf, die verwendete Datenmenge möglichst gering zu halten. Zudem wird mit einem reduzierten Trainingsdatensatz und der damit einhergehenden reduzierten Rechendauer auch eine effizientere Hyperparameteroptimierung ermöglicht (vgl. Kapitel 5.4.3). Aus diesem Grund wurde im Rahmen der Arbeit der Einfluss der Trainingsdatenquantität auf die Berechnungsergebnisse untersucht. Gleichzeitig wurde untersucht, welchen Einfluss die Verteilung der berücksichtigten Ereignisse auf das Trainingsergebnis hat. Ziel der Untersuchungen ist es, einen guten Kompromiss zwischen berücksichtigter Datenmenge und Modellgenauigkeit zu erzielen. Im Rahmen der Untersuchungen wurden dazu mit unterschiedlichen Vorgehensweisen Ereignisse aus dem Trainingsdatensatz ausgewählt. Insgesamt wurden die folgenden Untersuchungen in der aufgeführten Reihenfolge berücksichtigt:

- Vergleich des gesamten Datensatzes mit den natürlichen Regenreihen und Modellregenreihen
- Vergleich der gesamten natürlichen Regenreihen mit Teilmengen dieser Reihen
- Vergleich der gesamten Modellregenreihen mit Teilmengen dieser Reihen
- Kombination von Teilmengen der natürlichen Regenreihen und Modellregenreihen

In der ersten Untersuchung wurde zunächst der Einfluss der beiden unterschiedlichen Niederschlagsdatensätze auf die Modelleleistung überprüft. Hierzu wurde das T-GCN zum einen mit dem Datensatz aller natürlichen Regenreihen und zum anderen mit dem Datensatz aller Modellregen trainiert, wobei jedes Niederschlagsereignis aus unterschiedlich vielen Zeitschritten und damit auch *Samples* besteht. Das T-GCN aus Kapitel 5.3, das mit dem gesamten Trai-

ningsdatensatz trainiert wurde, wurde als Benchmark herangezogen. Somit wurden die folgenden drei Varianten miteinander verglichen:

- Variante 1 (Alle): Gesamter Datensatz (258 Ereignisse; 8026 *Samples*)
- Variante 2 (MR): Modellregenreihen (105 Ereignisse; 3318 *Samples*)
- Variante 3 (NR): Natürliche Regenreihen (153 Ereignisse; 4708 *Samples*)

Zur Evaluierung der Modellgüte wurden die Gütekriterien ereignisweise für die 26 Ereignisse des Testdatensatzes ermittelt. Abbildung 5-10 zeigt den Median und die Streuung der Ergebnisse für alle drei Varianten. An dieser Stelle sei darauf hingewiesen, dass nur vier der 26 Ereignisse zu Wasserständen $> 0,5$ m führen, weshalb die hierfür ermittelten Metriken nicht repräsentativ sind. Insgesamt zeigt sich, dass die Varianten MR und NR gegenüber dem vollständigen Datensatz erwartungsgemäß schlechtere Ergebnisse aufweisen. Vor dem Hintergrund, dass die Datensätze nur etwa 40 % und 60 % der Ereignisse beinhalten, fällt die Abweichung sowohl beim Median als auch bei der Streuung der Ergebnisse für Grenzwerte $\leq 0,2$ m moderat aus. Bei einem Grenzwert von 0,5 m sind die Abweichungen deutlicher, hier ist allerdings auch die nicht repräsentative Anzahl der Ereignisse zu berücksichtigen. Es ist zudem zu erkennen, dass die Variante mit den Modellregen für die übrigen Grenzwerte bei beiden Metriken meist eine etwas größer Streuung der Ergebnisse aufweist. Diese Tatsache könnte sich dadurch begründen, dass die Anzahl der berücksichtigten Ereignisse um etwa ein Drittel geringer ist. Eine genaue Aussage dazu ist allerdings an dieser Stelle noch nicht möglich. Aufgrund der zufriedenstellenden Ergebnisse für die einzelnen Datensätze wurden weitere Untersuchungen durchgeführt, um die Sensitivität des Modells gegenüber einer weiteren Ausdünnung der einzelnen Datensätze zu untersuchen.

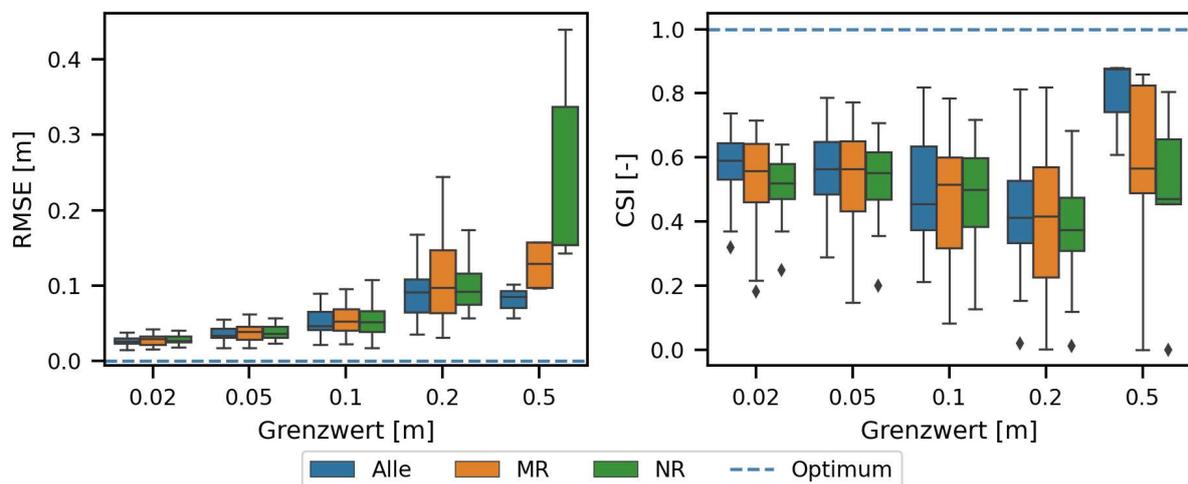


Abbildung 5-10: Verteilung der Metriken über alle 26 Ereignisse im Testdatensatz für die Varianten Alle (gesamter Datensatz), MR (Modellregenreihen) und NR (natürliche Regenreihen)

Als nächstes wurden Untersuchungen mit den natürlichen Regenreihen durchgeführt. Insgesamt wurden die fünf in Abbildung 5-11 aufgeführten Varianten untersucht. Die Vergleichsgröße für die Untersuchungen stellt die Variante 1 mit allen natürlichen Regenreihen dar. Für die Auswahl der zu berücksichtigten Ereignisse bei den übrigen Varianten wurden alle Ereignisse in Abhängigkeit der Wiederkehrintervalle in Klassen eingeteilt. Anschließend wurden für

die einzelnen Varianten Stichproben aus den einzelnen Klassen entnommen. Bei Variante 2 wurde zunächst eine gleichverteilte Stichprobe entnommen. Die Anzahl der Ereignisse pro Klasse wurde dazu mit neun Ereignissen auf die Größe der kleinsten Klasse festgelegt. Bei den folgenden Varianten wurden sukzessive weitere Ereignisse berücksichtigt. Unter der Annahme, dass insbesondere extreme Ereignisse mit hohen Wiederkehrzeiten relevant für den Trainingsprozess sind, wurden dabei erst die besonders seltenen Ereignisse überproportional berücksichtigt.

Mit den verschiedenen Datensätzen wurde erneut das T-GCN trainiert. In Abbildung 5-12 ist die Modellgüte für die einzelnen Varianten gegenübergestellt. Hier nehmen die Abweichungen deutlich zu. Abgesehen von Variante 5 kann in etwa eine Verdopplung des Medians bei den RMSE-Werten und eine Halbierung des Medians bei den CSI-Werten festgestellt werden. Variante 5 erzielt die besten Ergebnisse, beinhaltet aber mit 126 Ereignissen auch die meisten Ereignisse im Trainingsdatensatz. Gegenüber Variante 1 wurde die Anzahl lediglich um 27 Ereignisse (ca. 18 %) reduziert. Dementsprechend wird trotz deutlich sichtbarem Leistungsrückgang der Aufwand für die Datengenerierung in einem eher geringen Umfang reduziert. Insgesamt zeigen die Ergebnisse, dass eine gleichverteilte Berücksichtigung aller Klassen oder eine überproportionale Berücksichtigung der Extremereignisse nicht ausreichen, um an die Genauigkeit des am gesamten Datensatz trainierten Modells heranzureichen. Die zu Beginn der in diesem Kapitel durchgeführten Untersuchungen aufgestellte Annahme bezüglich der überproportionalen Berücksichtigung von Extremereignissen lässt sich somit an dieser Stelle nicht bestätigen.

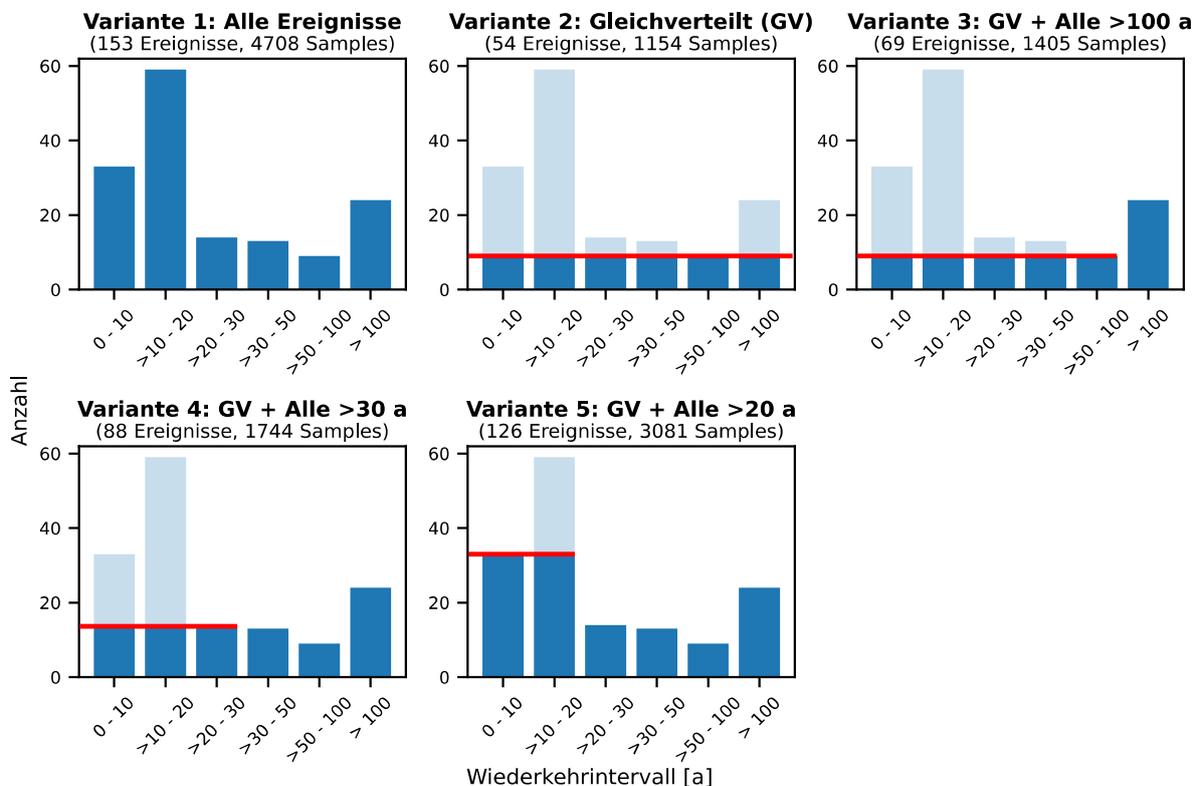


Abbildung 5-11: Übersicht über die fünf berücksichtigten Varianten zur Auswahl von Ereignissen aus den natürlichen Regenreihen (nicht berücksichtigte Ereignisse sind transparent dargestellt)

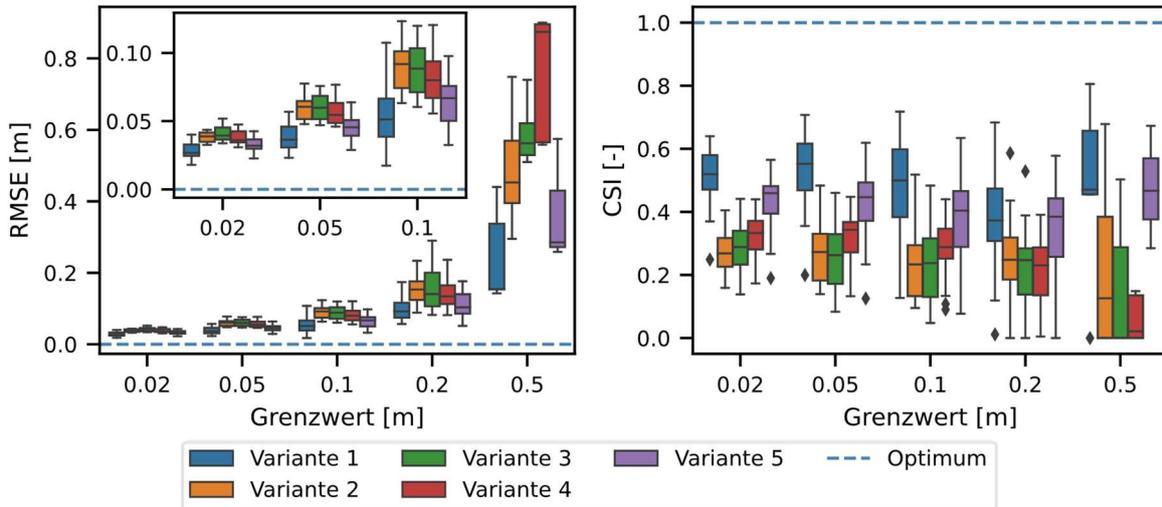


Abbildung 5-12: Verteilung der Metriken über alle 26 Ereignisse im Testdatensatz für die fünf berücksichtigten Varianten zur Auswahl von Ereignissen aus den natürlichen Regenreihen. Zur besseren Darstellung werden beim RMSE die unteren drei Grenzwerte zusätzlich vergrößert.

Als Nächstes wurden ähnliche Untersuchungen mit den Modellregenreihen durchgeführt. Auch hier wurde zunächst der gesamte Modellregendatensatz in der Variante 1 als Vergleichsgröße herangezogen. Anschließend wurden Ereignisse einzelner Wiederkehrzeiten, Dauerstufen oder Modellregentypen aus dem Datensatz entnommen, um weitere Varianten zu erstellen. In einer letzten Variante, der Variante 5, wurde gleichmäßig über alle Wiederkehrzeiten, Dauerstufen und Modellregentypen hinweg die Hälfte aller Ereignisse entnommen. Die einzelnen Varianten sind in Abbildung 5-13 dargestellt.

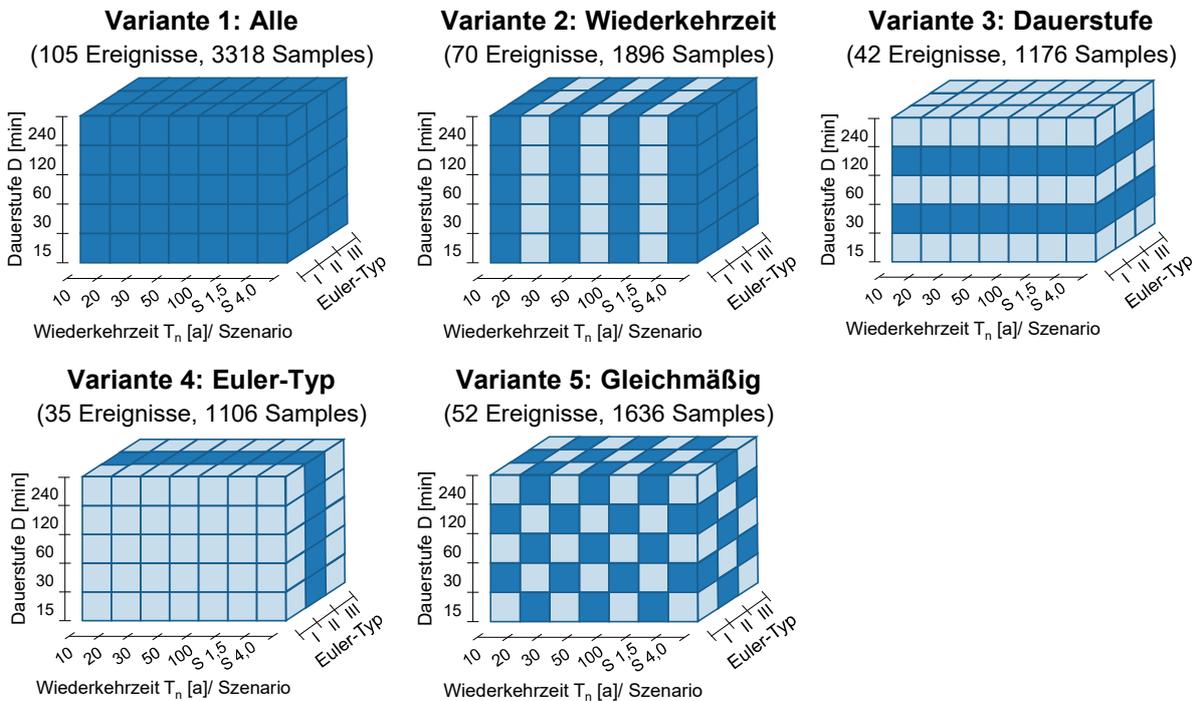


Abbildung 5-13: Übersicht über die fünf berücksichtigten Varianten zur Auswahl von Ereignissen aus den Modellregenreihen (nicht berücksichtigte Ereignisse sind transparent dargestellt)

Die Ergebnisse zeigen für die Varianten 2 bis 4 ähnlich hohe Abweichungen wie bei den Untersuchungen mit den natürlichen Regenreihen. Variante 5 mit gleichmäßig ausgewählten Ereignissen liefert hingegen gute Ergebnisse und zeigt gegenüber der Variante 1 nur geringe Abweichungen. Dadurch wird die hohe Bedeutung eines repräsentativen Datensatzes noch einmal verdeutlicht. Trotz einer halbierten Datenmenge können gegenüber den anderen Varianten gute Ergebnisse erzielt werden. Die etwas bessere Performance bei Variante 5 gegenüber dem Ausgangsdatsatz lässt sich abermals durch die nicht repräsentative Anzahl an Ereignissen für den Grenzwert 0,5 m begründen. Gegenüber dem Ausgangsdatsatz bestehend aus Modellregen und natürlichen Niederschlagsereignissen kann die Anzahl der berücksichtigten Ereignisse bei moderatem Genauigkeitsverlust so um 206 Ereignisse (80 %) reduziert werden.

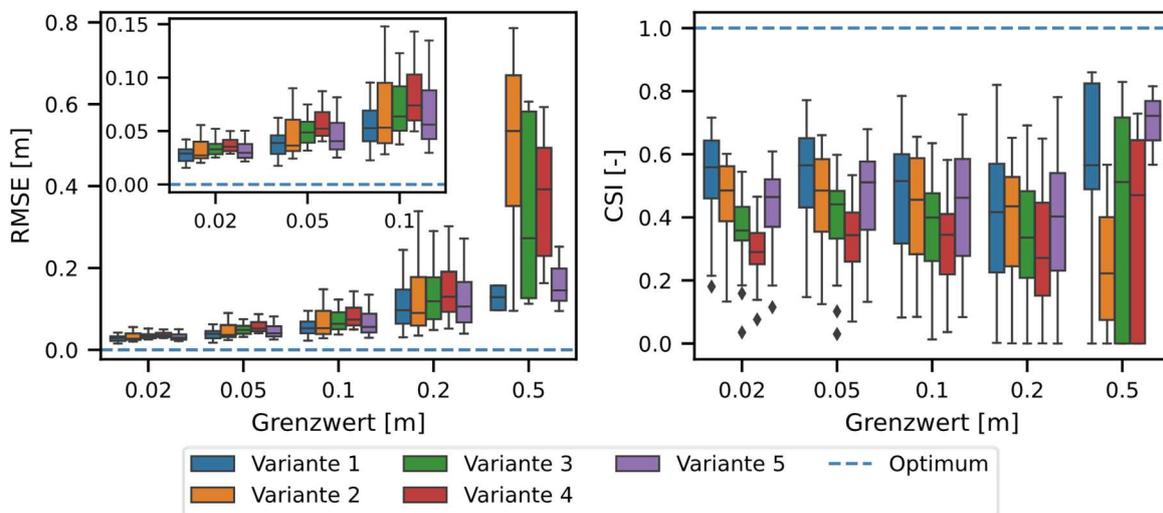


Abbildung 5-14: Verteilung der Metriken über alle 26 Ereignisse im Testdatensatz für die fünf berücksichtigten Varianten zur Auswahl von Ereignissen aus den Modellregenreihen. Zur besseren Darstellung wurden beim RMSE die unteren drei Grenzwerte zusätzlich vergrößert.

In einer letzten Analyse wurde eine Kombination aus ausgedünnten Modellregenreihen und natürlichen Regenreihen untersucht. Hierzu wurden die Variante mit den gleichmäßig ausgewählten Modellregenreihen und die Variante mit den gleichverteilt ausgewählten natürlichen Regenreihen kombiniert. Zwar führte letztere bei den Untersuchungen mit den natürlichen Regenreihen zu einem deutlichen Genauigkeitsverlust, sie beinhaltete aber auch die geringste Anzahl an Ereignissen. Ziel war es so, durch die Kombination beider Datensätze die Datenmenge immer noch möglichst gering zu halten, gleichzeitig aber die Modelleistung deutlich zu steigern. Der kombinierte Datensatz wurde mit dem Ausgangsdatsatz (Alle Ereignisse), den Modellregenreihen, den natürlichen Regenreihen sowie der Variante mit den gleichmäßig ausgewählten Modellregen verglichen. Tabelle 5-5 liefert zunächst einen Überblick über die enthaltenen Ereignisse und *Samples* der einzelnen Datensätze sowie deren Anteil am Ausgangsdatsatz.

Tabelle 5-5: Übersicht über die Anzahl der berücksichtigten Ereignisse und Samples sowie deren Anteil am Ausgangsdatensatz

| Variante | Ereignisse | | Samples | |
|---|------------|------------|------------|------------|
| | Anzahl [-] | Anteil [%] | Anzahl [-] | Anteil [%] |
| Alle | 258 | 100 | 8.026 | 100 |
| Natürliche Regenreihen (NR) | 153 | 59,3 | 4.708 | 58,7 |
| Modellregenreihen (MR) | 105 | 40,7 | 3.318 | 41,3 |
| Modellregenreihen gleichmäßig (MR GV) | 52 | 20,2 | 1.636 | 20,4 |
| Modellregenreihen gleichmäßig + Natürliche Regenreihen gleichverteilt (MR GV + NR GV) | 106 | 41,1 | 2.790 | 34,8 |

Die Ergebnisse der Modelle sind in Abbildung 5-15 dargestellt. Die Kombination beider Datensätze weist für einen Grenzwert von 0,5 m für beide Metriken gegenüber den übrigen Varianten mit einzelnen oder ausgedünnten Datensätzen bessere Ergebnisse auf. Für die übrigen Grenzwerte liegen die Ergebnisse aller Varianten, ausgenommen der Ausgangsvariante mit allen Ereignissen, nahe beieinander. Es ist auch zu erkennen, dass die einzelnen Varianten bei den verschiedenen Grenzwerten unterschiedlich gut abschneiden. So weist die Variante mit den natürlichen Regenreihen für niedrige Grenzwerte verhältnismäßig gute Werte auf, führt bei einem Grenzwert von 0,5 m aber zu dem schlechtesten Ergebnis. Hier ist davon auszugehen, dass sich die repräsentative Berücksichtigung der verschiedenen Wiederkehrzeiten und Dauerstufen auch bei extremen Niederschlagszenarien bei den Modellregen positiv auf die Modellgenauigkeit auswirkt.

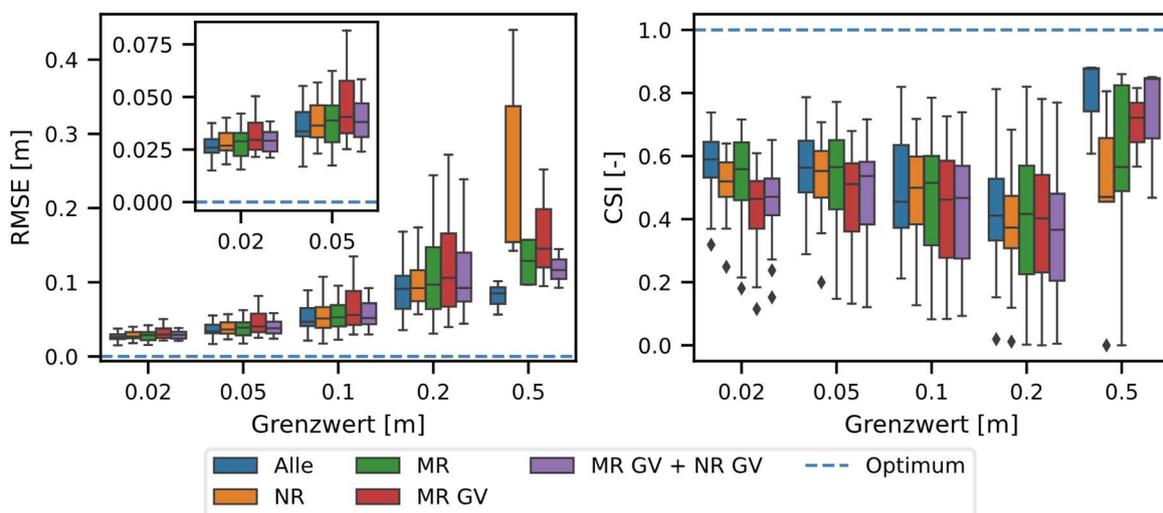


Abbildung 5-15: Vergleich der besten Varianten aus den vorangegangenen Untersuchungen mit einer Kombination aus gleichmäßig ausgewählten Modellregen (Variante 5 Abbildung 5-13) und gleichverteilt ausgewählten natürlichen Regenreihen (Variante 2 Abbildung 5-11)

Generell zeigt sich somit, dass nicht die Menge der Daten allein, sondern eine Kombination aus Datenmenge und Repräsentativität ausschlaggebend für hohe Genauigkeiten ist. Insbesondere die Varianten mit gleichmäßig ausgedünnten Modellregenreihen sowie einer Kombination aus gleichmäßig ausgedünnten Modellregen und gleichverteilt ausgewählten natürlichen Regenreihen zeigen, dass trotz einer erheblichen Reduktion der Trainingsdatenmenge

mit dem Modell noch gute Ergebnisse erzielt werden können. Hier bleibt auch zu berücksichtigen, dass in Hinblick auf die Ausweitung des entwickelten Ansatzes auf größere Gebiete ein Kompromiss zwischen Genauigkeit und erforderlicher Rechendauer gefunden werden muss, um die Praxistauglichkeit herzustellen.

5.4.2 Übertragbarkeit des Modells auf andere Gebiete

Eine weitere Möglichkeit zur Skalierung des Modells auf ein ganzes Stadtgebiet liegt darin, die Übertragbarkeit eines trainierten Modells herzustellen. Ein Modell könnte dann für ein Teilgebiet entwickelt und anschließend auch für die Prognose in den übrigen Teilen des Stadtgebietes genutzt werden. Guo et al. (2020) und Löwe et al. (2021) nutzten topografische Informationen als zusätzliche Eingabe, um die physikalischen Systemeigenschaften mit zu berücksichtigen und so die Übertragbarkeit der Modelle herzustellen. Das Untersuchungsgebiet wird bei diesem Ansatz in Ausschnitte (*Patches*) mit einer definierten Kantenlänge unterteilt. Anschließend wird das Modell mit Ausschnitten aus den unterschiedlichen Bereichen des Untersuchungsgebiets trainiert, um die unterschiedlichen Gebietscharakteristiken und deren Überflutungsverhalten zu erlernen. Im Anschluss an das Training soll das Modell dann möglichst gut auf andere Gebiete, die nicht im Trainingsdatensatz enthalten sind, generalisieren.

Dieser Ansatz kann zusätzlich mit *Transfer Learning*-Techniken kombiniert werden, um mit einem geringen zusätzlichen Trainingsaufwand die Ergebnisse für das Zielgebiet weiter zu verbessern (Seleem et al. 2023). Dabei handelt es sich um eine Technik, bei der die Fähigkeiten eines Modells auf ein anderes übertragen werden. Im Bereich von Künstlich Neuronalen Netzen werden dazu meist bei trainierten Modellen die ersten Schichten „eingefroren“ und nur die hinteren Schichten werden auf ein neues Problem hintrainiert. Dies bietet sich insbesondere an, um ein mit großen Datenmengen trainiertes, gut funktionierendes Modell auf eine ähnliche Aufgabe mit wenigen vorhanden Trainingsbeispielen anzupassen (Goodfellow et al. 2016). Im vorliegenden Beispiel bietet der Ansatz somit den Vorteil, dass ein Modell für neue Gebiete nur mit einer geringen Trainingsdatenmenge trainiert werden muss, wodurch sich der Aufwand für die Datengenerierung reduziert.

Auch im Rahmen der vorliegenden Arbeit wurden aufbauend auf den Ansätzen von Guo et al. (2020) und Löwe et al. (2021) Untersuchungen durchgeführt. Durch die zusätzliche Berücksichtigung von Überstau als Eingabegröße reicht hier die alleinige Berücksichtigung von topografischen Informationen allerdings nicht aus. Vielmehr ist die Kombination oder Verschneidung der Überstauprognose mit weiteren physikalischen Systemeigenschaften als zusätzliche Eingabe erforderlich, um eine Übertragbarkeit zu ermöglichen. Eine Möglichkeit dies zu tun, besteht in der in Untersuchung 2 (vgl. Kapitel 5.3.2) durchgeführten Verschneidung der Überstauprognosen mit den Senkeneinzugsgebieten. Einen ähnlichen Ansatz schlagen Löwe et al. (2021) vor, indem in einem Rasterdatensatz Fließwege durch angrenzende Überstauvolumina gewichtet werden. Auch wenn die Untersuchungen in Kapitel 5.3.2 gezeigt haben, dass bei der Verwendung der Überstauprognose als Rastersequenzen gegenüber dem T-GCN mit Einbußen bei der Vorhersagegenauigkeit zu rechnen ist, wurde dieser Ansatz für erste Tests herangezogen. Dies insbesondere vor dem Hintergrund, dass beim T-GCN die Darstellung des

räumlichen Kontextes als Graphen nicht ohne Weiteres übertragbar ist, da die Anzahl der Knoten sowie deren Anordnung und Abstände zwischen den einzelnen Gebieten variieren.

Zur Erzeugung eines Trainingsdatensatzes wurden zunächst *Patches* mit einem gleichmäßigen Raster erzeugt (siehe Abbildung 5-16). Als *Patch*-Größe wurden 256×256 Pixel gewählt, das entspricht bei der hier gewählten Auflösung von 2×2 m einer Kantenlänge von 512 m. Wie bereits in Guo et al. (2022) und (Seleem et al. 2023) erwähnt, reicht die Größe der Ausschnitte nicht immer aus, um Senken und deren Einzugsgebiete vollständig abzudecken. Durch die Berücksichtigung der verschiedenen Geländeinformationen und hydrologischen Eigenschaften als Eingabedatensatz wird versucht, dem entgegenzuwirken. Insgesamt lässt sich das vorliegende Untersuchungsgebiet so in zwölf *Patches* unterteilen, wofür zunächst elf für das Training und einer zum Testen verwendet wurden. Die geringe Anzahl der *Patches* führt allerdings dazu, dass insbesondere stark überflutete Bereiche nicht repräsentativ mit dem Datensatz abgedeckt werden (vgl. Abbildung 5-16). Durchgeführte Tests mit einem trainierten Modell haben dies bestätigt und keine zufriedenstellenden Ergebnisse geliefert. Der in (Seleem et al. 2023) vorgestellte *Transfer Learning*-Ansatz konnte die Ergebnisse etwas verbessern, die Abweichungen waren jedoch weiterhin so groß, dass die Untersuchungen an dieser Stelle abgebrochen wurden. Auf eine Darstellung der Ergebnisse wird aufgrund der mangelhaften Prognosegenauigkeit ebenfalls verzichtet.

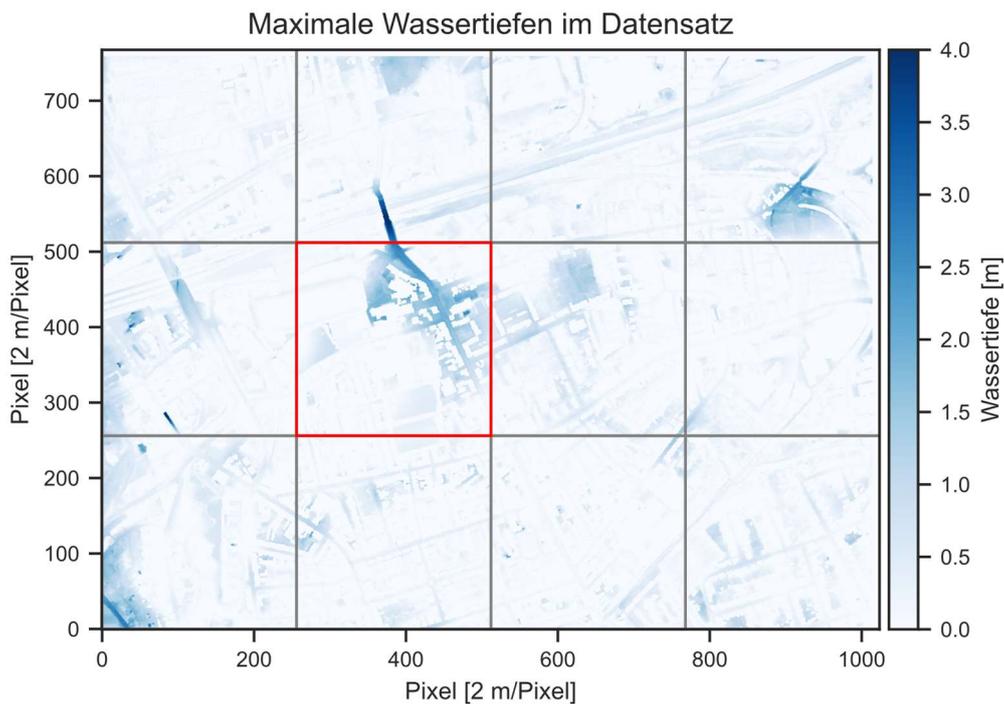


Abbildung 5-16: Untersuchungsgebiet mit den maximalen im Datensatz enthaltenen Wassertiefen und dem verwendeten Raster für die Erzeugung von *Patches*. Das Testgebiet ist rot umrandet, die übrigen Gebiete wurden für das Training verwendet.

Es bleibt somit festzuhalten, dass aufgrund der geringen Größe des berücksichtigten Untersuchungsgebietes die gängigen in der Literatur aufgeführten Ansätze zur Herstellung der Übertragbarkeit eines trainierten Modells nicht funktionieren. Ein übertragbares Vorhersagemodell birgt jedoch nach wie vor erhebliches Potenzial für eine spätere Ausweitung auf ganze Stadtgebiete oder sogar auf Landesebene. Um die generelle Eignung der Ansätze zu prüfen, sollten

daher künftig weitere Untersuchungen an größeren Gebieten durchgeführt werden.

5.4.3 Analyse des Einflusses von Hyperparametern (Hyperparameter tuning)

Neuronale Netze verfügen je nach Aufbau über diverse Hyperparameter, die es festzulegen gilt. Dabei handelt es sich um nicht trainierbare Parameter, mit denen die Netzstruktur und der Lernprozess beeinflusst wird. Hierzu zählen beispielsweise die Lernrate, die Anzahl an Schichten eines Netzes und deren Neuronen, die berücksichtigte Aktivierungsfunktion oder verschiedene Regularisierungsparameter wie die *Dropout*-Rate. Im Rahmen der Modellentwicklung wurden die verschiedenen Hyperparameter zunächst in Anlehnung an Erkenntnisse aus verschiedenen Literaturquellen (u. a. (Löwe et al. 2021; Hofmann und Schüttrumpf 2021; Guo et al. 2020; Guo et al. 2022)) und ersten Voruntersuchungen festgelegt. Um die Modelleleistung weiter zu optimieren, werden nachfolgend ausgewählte Hyperparameter mit einem systematischen Vorgehen angepasst. In einem ersten Schritt wird dabei der Einfluss des berücksichtigten Eingabefensters bei den Niederschlagsinformationen analysiert. Dieser Parameter kann im Rahmen der Vorverarbeitung der verwendeten Datensätze angepasst werden. In einer zweiten Untersuchung erfolgt dann mittels der unter anderem in Géron (2019) beschriebenen Rastersuche (*Grid Search*) die automatische Optimierung verschiedener Hyperparameter, die die gewählte Netzarchitektur beeinflussen. Die Ergebnisse beider Untersuchungen werden im Folgenden nacheinander aufgeführt. Da das Training des T-GCN sehr rechen- und zeitintensiv ist, wurde jeweils auf die in Kapitel 5.4.1 aufgeführte gleichmäßig ausgewählte Stichprobe aus den Modellreihen als Trainingsdatensatz zurückgegriffen.

Einfluss des Eingabefensters

Im Rahmen der Untersuchungen zum Einfluss des Eingabefensters wurde überprüft, wie sich die Länge der Eingabesequenz der Niederschlagsinformationen auf die Prognosegenauigkeit auswirkt. Die Niederschlagsprognose blieb hiervon unberücksichtigt, bei allen übrigen Informationen (gefallener Niederschlag, kumulierte Niederschlagssumme und verstrichene Zeit) wurde die Länge des Eingabefensters mit 5, 15, 30 und 60 Minuten variiert. Hierdurch wurde geprüft, ob ein möglichst langes Zeitfenster das Modell positiv beeinflusst oder ggf. überflüssige Informationen liefert. Die Ergebnisse der Untersuchung sind in Abbildung 5-17 dargestellt. Es ist klar zu erkennen, dass die Vorhersagegenauigkeit sowohl im Falle des RMSEs als auch des CSIs gemeinsam mit der Länge des Eingabefensters zunimmt. Das ursprünglich verwendete Eingabefenster von 60 Minuten liefert mit Abstand die besten Ergebnisse. Es ist möglich, dass mit noch längeren Eingabefenstern die Ergebnisqualität weiter verbessert werden kann. Zudem zeigt das Ergebnis die hohe Sensitivität des Modells gegenüber den Niederschlagsinformationen der vergangenen Zeitschritte. Es könnte daher zielführend sein, in künftigen Untersuchungen weitere Merkmale aus den Niederschlagsinformationen abzuleiten und dem Modell als zusätzliche Eingabe zur Verfügung zu stellen.

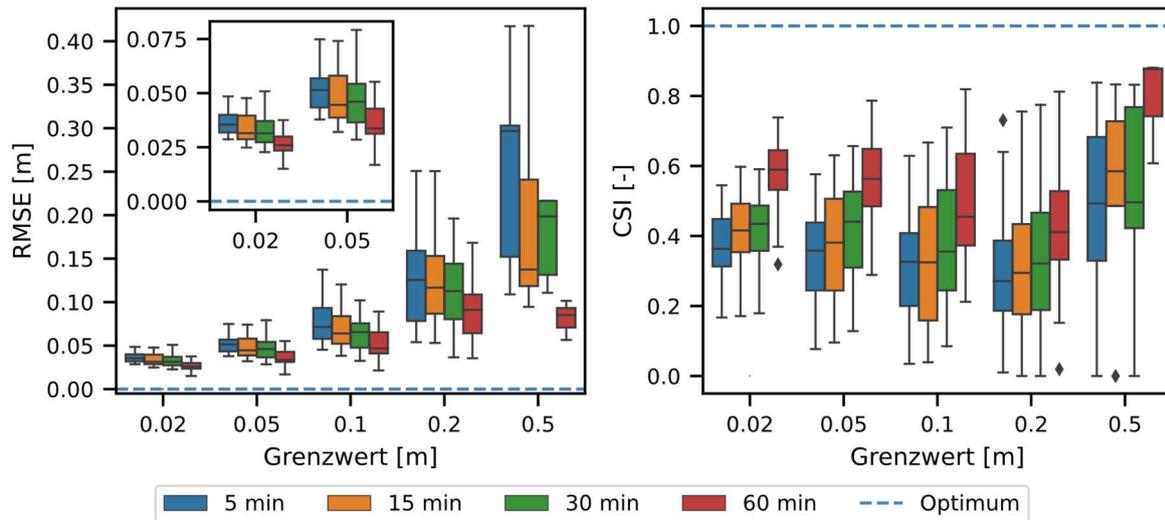


Abbildung 5-17: Vergleich unterschiedlicher Längen des berücksichtigten Eingabefensters für die Niederschlagsinformationen

Einfluss der gewählten Netztiefe, Filteranzahl, Kernelgröße und Verlustfunktion

Im Vergleich zur linearen Regression, bei der nur wenige oder sogar bloß ein einzelner Hyperparameter variiert werden können, gibt es bei Neuronalen Netzen je nach gewähltem Netzaufbau eine Vielzahl an Hyperparametern. Die automatisierte Optimierung sämtlicher Hyperparameter im hier verwendeten Netzaufbau würde bei der vorhandenen Rechenkapazität mehrere Monate oder sogar Jahre dauern. Deswegen wurden in der vorliegenden Arbeit mit der Netztiefe, Filteranzahl, Kernelgröße und Verlustfunktion vier Parameter ausgewählt, die optimiert wurden. Die Filteranzahl bezieht sich jeweils auf die Anzahl der Filter in der ersten Schicht des Netzes und wird nachfolgend als Basisfilter bezeichnet. Bei der Netztiefe handelt es sich um die Anzahl der Schichten vor und nach der *Concatenate*-Schicht, also der Mitte des Netzes. Die gesamte Anzahl an Schichten ist dementsprechend doppelt so hoch.

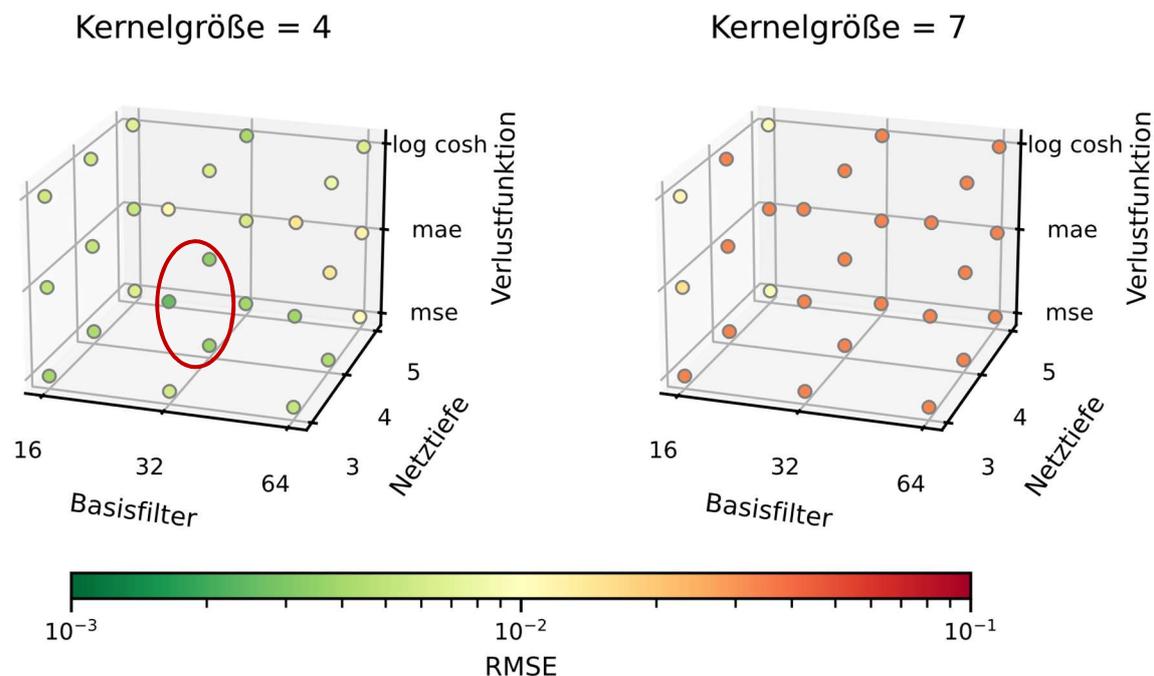
Weitere relevante Hyperparameter wie die Lernrate konnten beispielsweise bereits mit geeigneten Algorithmen bestimmt (Vorgehen nach Smith (2015)) oder wie im Falle der gewählten Art der Schichten durch manuelle Voruntersuchungen festgelegt werden. Darüber hinaus wurde auf eine Optimierung der Aktivierungsfunktion, *Batch*-Normalisierungsschichten oder *Dropout*-Schichten auf Erfahrungen aus anderen Studien aufgebaut (u. a. Hofmann und Schüttrumpf 2021; Löwe et al. 2021; Guo et al. 2022).

Für die vier Parameter wurde mittels Rastersuche aus den in Tabelle 5-6 aufgeführten Werten die optimale Kombination ermittelt. Die Rastersuche ist ein gängiges Verfahren zur Hyperparameteroptimierung, bei dem Modelle mit sämtlichen Parameterkombinationen trainiert und evaluiert werden. Mit steigender Anzahl an Hyperparametern nimmt der Berechnungsaufwand dabei exponentiell zu. Daher ist das Verfahren in Hinsicht auf die Rechendauer nicht besonders effizient. Angesichts dessen wurden auch modernere und weniger rechenintensive Methoden wie der in Keras implementierte Hyperband-Algorithmus (Li et al. 2018) getestet. Diese führten jedoch zu einer Überlastung des GPU-Speichers und konnten daher nicht verwendet werden.

Tabelle 5-6: Ausgewählte Hyperparameter und berücksichtigte Werte

| Hyperparameter | Werte |
|----------------------------|--|
| Netztiefe | 3, 4, 5 |
| Basisfilter (Filteranzahl) | 16, 32, 64 |
| Kernelgröße | 4, 7 |
| Verlustfunktion | MSE, MAE, Log Cosh |
| mit: | |
| MSE | Mittlerer Quadratischer Fehler (<i>Mean Squared Error</i>) |
| MAE | Mittlerer Absoluter Fehler (<i>Mean Absolute Error</i>) |
| Log Cosh | Logarithmus des Kosinus hyperbolicus (<i>Logarithm of the hyperbolic cosine</i>) |

Das Ergebnis der Rastersuche ist in Abbildung 5-18 dargestellt. Es ist klar zu erkennen, dass die Modelle mit einer Kernelgröße von vier besser funktionieren als mit sieben. Es zeigt sich aber auch, dass bei der gewählten Grundarchitektur die Genauigkeit nicht mit der Basisfilteranzahl oder Netztiefe zunimmt. Vielmehr wiesen die besten drei Parameterkombinationen eine Basisfilteranzahl von 32 und eine Netztiefe von drei oder vier auf (siehe roter Kreis in Abbildung 5-18). Bei der Verlustfunktion konnten mit dem MSE und MAE die besten Ergebnisse erzielt werden. Es ist allerdings zu beachten, dass sich das Ergebnis durch die Variation weiterer, hier nicht berücksichtigter Parameter, verschieben könnte. So könnte unter anderem die für das ursprünglich in Kapitel 5.3 entwickelte Modell festgelegte Lernrate für Netze mit einer steigenden Anzahl an Parametern zu hoch sein, sodass der Trainingsalgorithmus divergiert und das globale Minimum nicht findet. Auf weitere Untersuchungen hierzu wurde aufgrund der eingangs beschriebenen langen Rechendauern im Rahmen dieser Arbeit verzichtet.

**Abbildung 5-18: Übersicht über die RMSE-Werte für die einzelnen Hyperparameter-Kombinationen (die drei besten Ergebnisse sind rot eingekreist)**

Mit den drei besten Parameterkombinationen wurden in einer abschließenden Untersuchung

Modelle erstellt. Diese wurden am gesamten Datensatz für 100 Epochen trainiert. Anschließend wurden zur Evaluierung der Modellgüte wieder die Metriken RMSE und CSI für einen Grenzwert von 0,05 m ereignisweise am Testdatensatz ermittelt. Die Ergebnisse sind in Abbildung 5-19 dargestellt. Es ist zu erkennen, dass das Modell mit dem MSE als Verlustfunktion hinsichtlich des RMSE für alle Grenzwerte am besten abschneidet. Beim CSI liefert das Modell für den Grenzwert von 0,02 m ein etwas schlechteres Ergebnis, zeigt bei den übrigen Grenzwerten aber im Vergleich zu den anderen Modellen kaum Unterschiede. Da die höheren Grenzwerte für die Verwendung der Überflutungsvorhersage im Krisenmanagement besonders relevant sind und aufgrund der besseren Ergebnisse beim RMSE, wird folgende Model-larchitektur für die nachfolgenden Untersuchungen ausgewählt:

- Netztiefe: 4 Schichten
- Basisfilter: 32
- Kernelgröße: 4
- Verlustfunktion: MSE (*Mean Squared Error*)

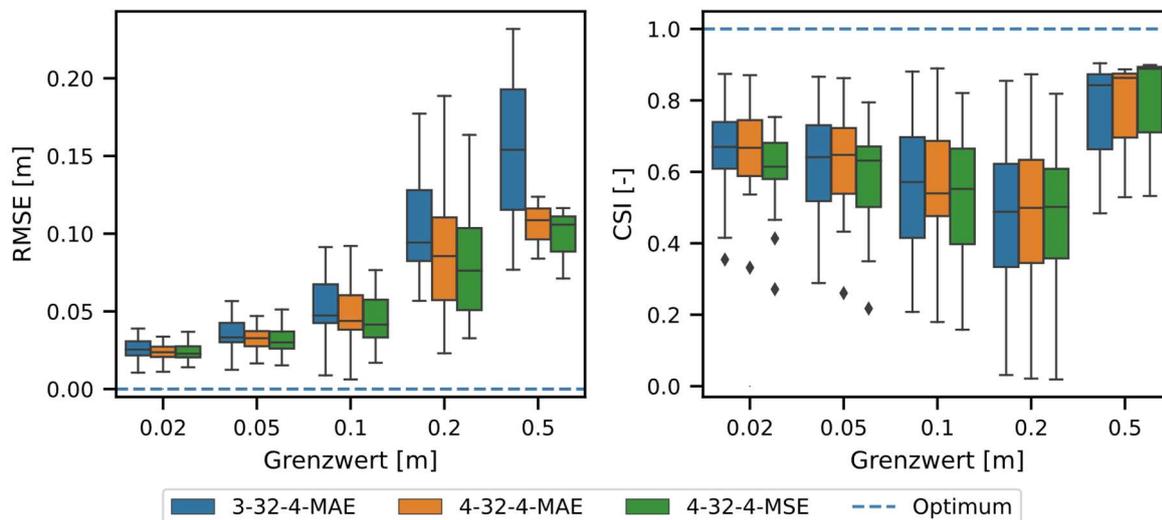


Abbildung 5-19: Verteilung der Metriken über alle 26 Ereignisse im Testdatensatz. Die Bezeichnung der Modelle ergibt sich jeweils aus den Werten der Hyperparameter in der Reihenfolge Netztiefe-Basisfilter-Kernelgröße-Verlustfunktion

5.5 Beurteilung der Vorhersagegenauigkeit

Zur abschließenden Beurteilung der Vorhersagegenauigkeit des T-GCNs mit den optimierten Hyperparametern wurden weitere differenzierte Untersuchungen durchgeführt. In einem ersten Schritt wurde getestet, wie sich die Vorhersagegenauigkeit in Abhängigkeit der Wiederkehrzeit der einzelnen Ereignisse verhält. Hierdurch soll geprüft werden, ob das Modell insbesondere extreme Ereignisse mit einer hinreichenden Genauigkeit vorhersagt. Das Ergebnis ist in Abbildung 5-20 dargestellt und zeigt für einen Grenzwert von 0,05 m den RMSE und den CSI der 26 Ereignisse in Abhängigkeit von deren Wiederkehrzeiten. Hier zeigt sich, dass das Modell auch für die besonders relevanten sehr seltenen Ereignisse mit $T > 100$ a gute Ergebnisse liefert. Beim CSI liegen die Ergebnisse sogar im oberen Bereich. Auch die Ergebnisse

beim RMSE sind vor dem Hintergrund, dass hier die absoluten Abweichungen in die Berechnung eingehen, sehr positiv anzusehen. Das Modell sticht bei keinem der berücksichtigten Wiederkehrintervalle besonders heraus, die Ergebnisse sind ähnlich.

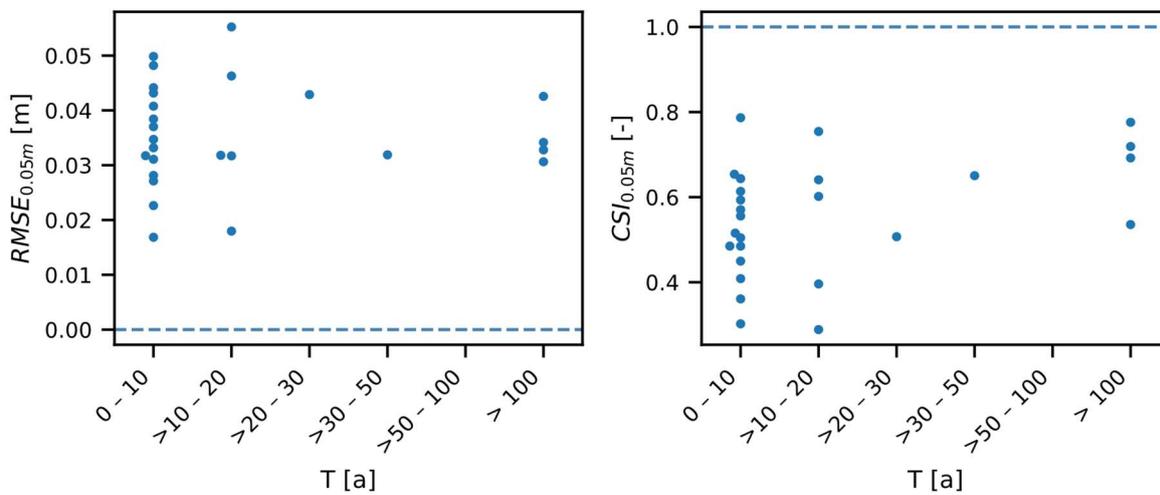


Abbildung 5-20: Verteilung der Metriken in Abhängigkeit verschiedener Wiederkehrintervalle (die blaue, gestrichelte Linie markiert jeweils das Optimum)

In einem zweiten Schritt wurde geprüft, ob das Modell Wasserstände über- oder unterschätzt. Hierzu wurden die prognostizierten und die simulierten Wasserstände in Abbildung 5-21 als 2D-Histogramm gegenübergestellt. Berücksichtigt wurden alle Pixel über einem Grenzwert von 0,05 m aus allen Vorhersagerastern der 26 Ereignisse. Die gestrichelte Linie gibt jeweils die ideale Übereinstimmung zwischen Vorhersage und HD-Simulation an. Die Abweichungen schwanken relativ gleichmäßig um die gestrichelte Linie, mit einer leichten Tendenz des Vorhersagemodells zum Unterschätzen der Wasserstände. Es ist aber keine starke Systematik erkennbar, weshalb an dieser Stelle kein zwingender Bedarf gesehen wird, das Modell dahingehend anzupassen. Zur Bewertung der Ergebnisse sei an dieser Stelle auch darauf hingewiesen, dass für die farbliche Darstellung der Pixelanzahl eine logarithmische Skala verwendet wurde. Abweichungen von mehr als einem Meter treten dementsprechend nur sehr selten auf und der Großteil der Fehler liegt nahe null.

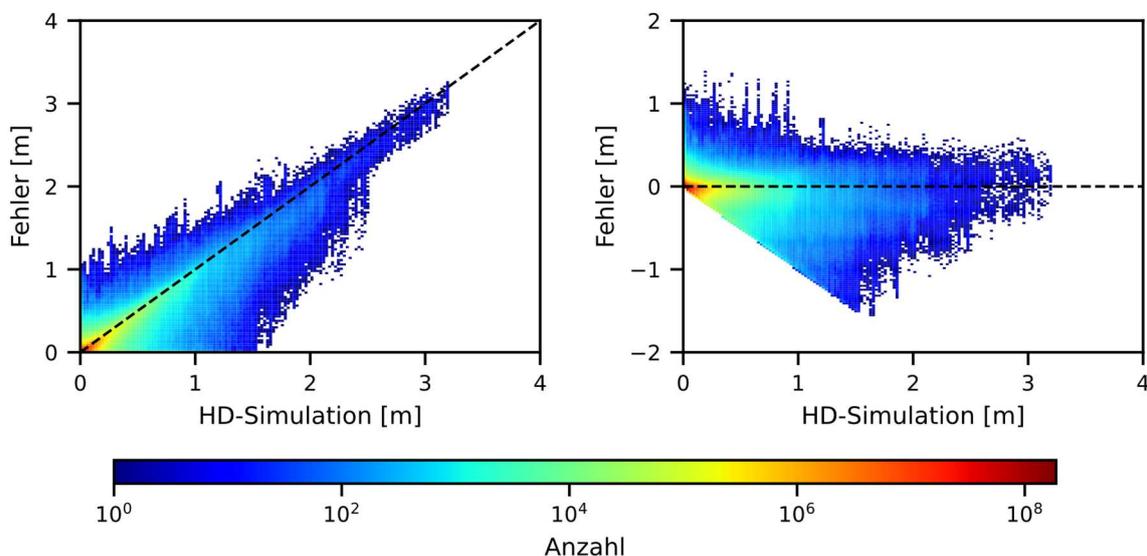


Abbildung 5-21: 2D-Histogramm mit dem pixelweisen Vergleich der vorhergesagten und simu-

lierten Wasserstände (links) und dem pixelweisen Vorhersagefehler in Abhängigkeit der simulierten Wasserstände (rechts)

Zur besseren Einschätzung der Vorhersagegenauigkeit wurden abschließend für verschiedene historische Ereignisse Abbildungen der Überflutungsflächen erzeugt. Neben unterschiedlichen Ereignissen wurden dabei auch unterschiedliche Zeitpunkte innerhalb der Ereignisse berücksichtigt. Als Vorhersagehorizont wurden jeweils 60 Minuten berücksichtigt und es wurden jeweils Abbildungen für die Zeitschritte $t = 15, 30$ und 60 Minuten erzeugt.

Abbildung 5-22 zeigt exemplarisch die Niederschlagsganglinie und das Prognoseergebnis für ein historisches Starkregenereignis vom 3. Juli 2009. Die Prognose wurde zu Ereignisbeginn erstellt. Für den prognostizierten Zeitraum beträgt die Niederschlagssumme 46 mm und für einzelne Dauerstufen wurden in diesem Zeitraum Wiederkehrzeiten von mehr als 200 Jahren erreicht. Die Abbildung stellt neben den Ergebnissen der HD-Simulation und des Prognosemodells auch die Lagegenauigkeit und die Abweichung der Wasserstände für die einzelnen Zeitpunkte dar. Zur detaillierten Bewertung der Prognosegüte wird zudem in Abbildung 5-23 ein Detailausschnitt abgebildet. Der visuelle Vergleich zeigt eine sehr hohe Übereinstimmung, es werden keine unrealistischen Überflutungsmuster ausgegeben. Die Ergebnisse zu weiteren Ereignissen sind im Anhang A-4 dargestellt. Auch dort zeigt sich für die unterschiedlichen Ereignisse und Vorhersagezeitpunkte weitestgehend eine hohe Übereinstimmung zwischen HD- und Prognosemodell. Ansonsten lassen sich anhand der Abbildungen die folgenden beiden Eigenschaften erkennen:

1. Vorhersagezeitpunkte mit geringen Wassertiefen und wenigen überfluteten Pixeln führen häufig zu großen Fehlern. Diese Problematik wird bei den dargestellten Ereignissen insbesondere zu Ereignisbeginn deutlich, wo sich zunächst nur geringe Wasserstände bilden. Der CSI liegt dann nahe 0 , dem schlechtesten möglichen Ergebnis. Zudem zeigt sich die Problematik bei weniger extremen Ereignissen, bei denen über die gesamte Ereignisdauer nur geringe Wasserstände auftreten (vgl. Ereignis-ID 81 in Anhang A-4). Die gleiche Problematik wurde auch von Löwe et al. (2021) festgestellt. Vorhersagezeitpunkte mit vielen überfluteten Pixeln weisen hingegen meist eine hohe Genauigkeit auf, wie dies auch in Abbildung 5-22 für die Zeitpunkte $t = +30$ min und $+60$ min der Fall ist. Die für das Krisenmanagement besonders relevanten Überflutungsmuster werden dementsprechend mit einer hohen Lagegenauigkeit prognostiziert.
2. In der Mitte des in Abbildung 5-23 und im Anhang dargestellten Ausschnitts befindet sich eine Unterführung, in der mit bis zu 25 cm oder vereinzelt sogar 40 cm die größten Abweichungen auftreten. Vor dem Hintergrund, dass die Wasserstände dort teilweise über zwei Meter liegen, relativiert sich der Fehler jedoch. Der relative Fehler würde in dem Fall im Bereich von etwa $10-20\%$ und somit in einem vertretbaren Rahmen liegen. Dementsprechend weisen besonders extreme Ereignisse aufgrund der Berücksichtigung von absoluten Wasserständen trotz hoher Lagegenauigkeit teilweise auch etwas höhere RMSE-Werte auf (vgl. insbesondere Ereignis-ID 102 in Anhang A-4).

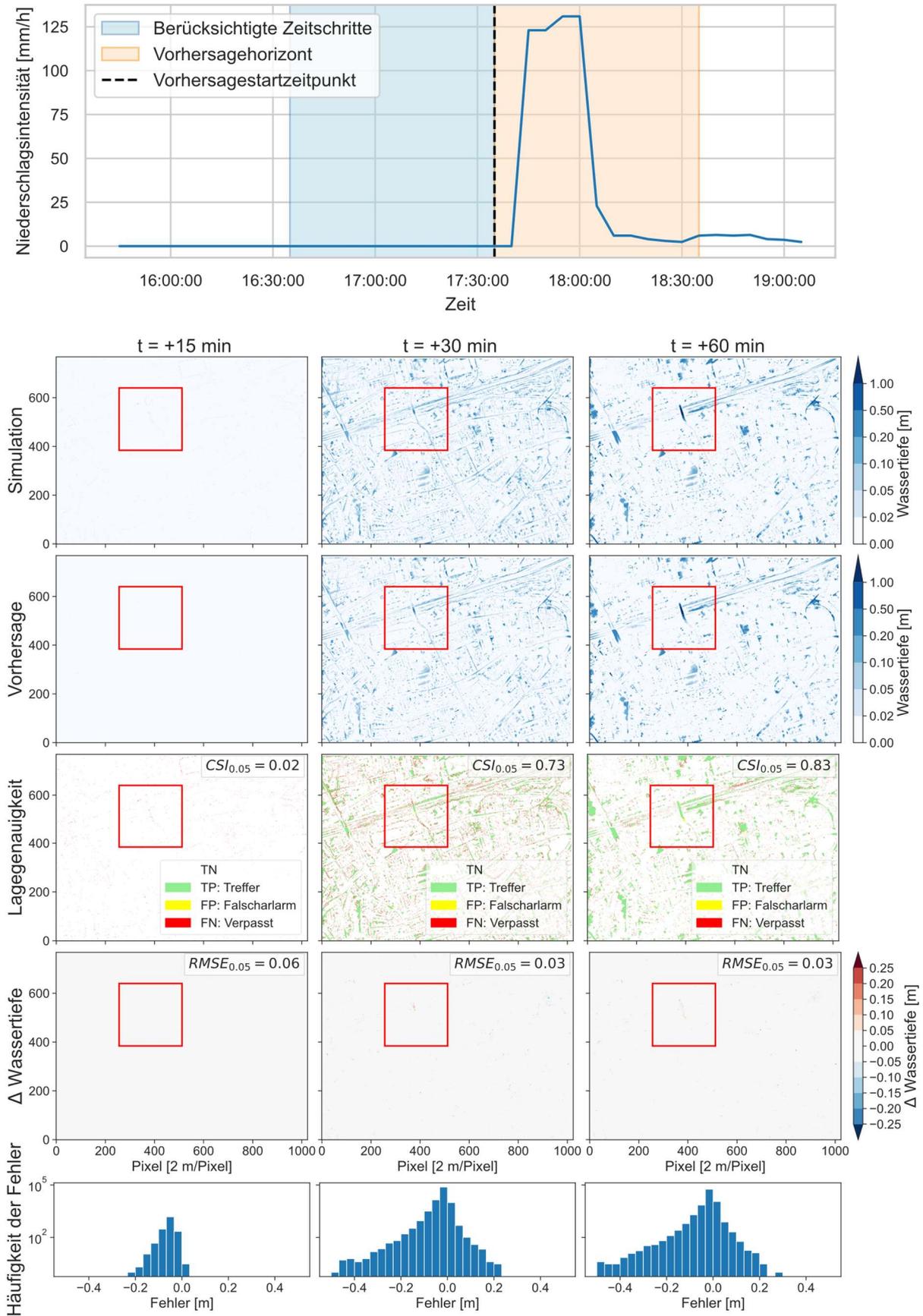


Abbildung 5-22: Niederschlagsganglinie mit markiertem Vorhersagezeitpunkt zu Beginn des Ereignisses vom 3. Juli 2009 in Gelsenkirchen (oben) und Vorhersageergebnis und -evaluierung für drei Zeitschritte der Vorhersage. Der nachfolgend in Abbildung 5-23 dargestellte Detailausschnitt ist rot markiert

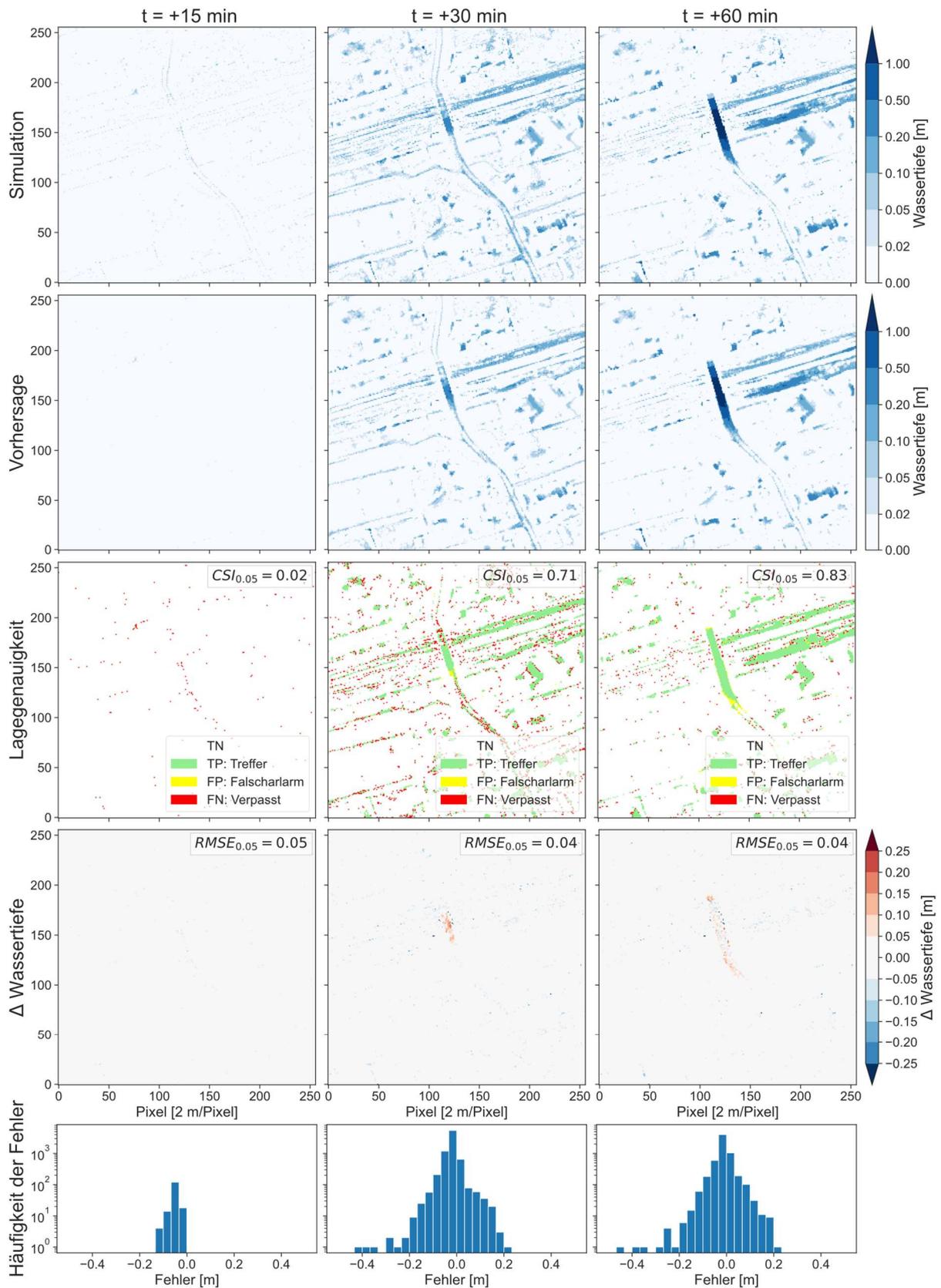


Abbildung 5-23: Detailansicht des Vorhersageergebnisses und der zugehörigen Evaluierung für einen Ausschnitt mit einer überfluteten Unterführung

5.6 Zusammenfassung und Bewertung der Ergebnisse

Im Rahmen der durchgeführten Untersuchungen wurde mit Künstlich Neuronalen Netzen ein geeignetes ML-Verfahren für das gegebene Lernproblem ausgewählt. Als Modellarchitektur wurde eine *Encoder-Decoder*-Struktur gewählt, bei der zunächst Merkmale aus den Eingabedaten extrahiert und anschließend aus den Merkmalen die Überflutungssequenz für die kommenden Zeitschritte als Zielgröße generiert wird. Untersuchungen zum Modellaufbau haben gezeigt, dass die besten Ergebnisse unter Berücksichtigung der Niederschlagsinformationen, der Überstauvorhersage und räumlichen Informationen als Eingabegrößen erzielt werden. Bei der Architektur hat es sich als hilfreich erwiesen, wenn die Überstauprognose als Graph abgebildet und verarbeitet wird. Die Verwendung eines cGANs als Modellansatz hat sich hingegen als nicht zielführend erwiesen. Die Ergebnisse konnten so gegenüber einem vorwärtsgerichteten Ansatz nicht verbessert werden und gleichzeitig verläuft das Training dabei nicht so stabil. Als finales Modell wurde daher das T-GCN ausgewählt (vgl. Kapitel 5.3.3).

Das T-GCN zeigt eine gute Übereinstimmung der prognostizierten Überflutungsflächen mit den Ergebnissen des HD-Modells. Der visuelle Abgleich zeigt nur geringfügige Unterschiede, die für den Einsatz in Warnsystemen oder als Entscheidungsgrundlage für die Akteure des Krisenmanagements nur eine untergeordnete Rolle spielen dürften. Die Werte des RMSE und des CSI liegen in einem ähnlichen Rahmen wie in anderen Untersuchungen (Hofmann und Schüttrumpf 2021; Löwe et al. 2021), bei denen *Deep Learning*-Modelle zur Berechnung von Überflutungskarten eingesetzt wurden. An dieser Stelle sei jedoch darauf hingewiesen, dass der Vergleich mit anderen Studien aufgrund der abweichenden Aufgabe des Vorhersagemodells (Rastersequenzen anstatt einzelner Raster) sowie den unterschiedlichen Eigenschaften der berücksichtigten Untersuchungsgebiete (Topografie, Versiegelungsgrad etc.) nur bedingt aussagekräftig ist. Zudem ist an dieser Stelle zu berücksichtigen, dass in den Untersuchungen sowohl für das Training als auch die Auswertung der Modelle immer perfekte Niederschlags- und Überstauvorhersagen berücksichtigt wurden. In der Realität sind beide Vorhersagen mit Unsicherheiten behaftet, die sich dann auch auf das hier entwickelte Modell auswirken. In Hinblick auf den operationellen Betrieb sind dementsprechend weitere Untersuchungen mit Ergebnissen aus Vorhersagemodellen für Niederschlag und Überstau erforderlich, um die Auswirkungen der Unsicherheiten auf die prognostizierten Überflutungsflächen zu evaluieren. Letzteres erfolgt bereits in Kapitel 7 der vorliegenden Arbeit.

Das finale Modell benötigt für die Berechnung der Überflutungssequenz für die kommenden 60 Minuten lediglich wenige Sekunden. Auch bei einer Ausweitung auf größere Gebiete ist davon auszugehen, dass die Rechenzeit unter einer Minute liegt, womit das Modell für den Echtzeitbetrieb geeignet ist. Dieses Ergebnis deckt sich mit den Erkenntnissen aus anderen Untersuchungen (Berkhahn et al. 2019; Hofmann und Schüttrumpf 2021; Guo et al. 2022). Eine Ausweitung des Vorhersagehorizonts führt aus Sicht des entwickelten Modells zwar zu einem höheren Rechenbedarf, wäre ansonsten aber technisch problemlos möglich. Eingeschränkt wird der mögliche Vorhersagehorizont jedoch durch die Unsicherheiten der Niederschlagsvorhersage. Diese nehmen bei Extremereignissen und den derzeit verfügbaren Vorhersagemodellen bereits nach wenigen Minuten extrem zu und lassen dadurch keine aussagekräftigen Vorhersagen für mehr als zwei Stunden zu (Sun et al. 2014).

Wie bereits in Kapitel 4.5 beschrieben, ist die Generierung der Trainings- und Testdatensätze sehr rechen- und damit zeitintensiv. Bei einer Ausweitung auf das gesamte Stadtgebiet würden die Rechendauern ohne zusätzliche Rechenkapazitäten mehrere Jahre betragen, wodurch der Ansatz nicht mehr praktikabel ist. Vor diesem Hintergrund wurden im Rahmen einer Sensitivitätsanalyse weitere Untersuchungen in Hinblick auf die Skalierbarkeit des Modells durchgeführt. Eine Analyse hinsichtlich der erforderlichen Datenquantität und -verteilung hat gezeigt, dass die Reduzierung der Anzahl an berücksichtigten Ereignissen erwartungsgemäß zu einem Rückgang der Modellgenauigkeit führt. Es konnte aber auch nachgewiesen werden, dass je nach Auswahlmethodik der Datensatz bei moderatem Genauigkeitsverlust um bis zu 80 % reduziert werden kann. Damit kann eine erhebliche Reduktion der Rechendauer für die Datengenerierung erreicht werden und gleichzeitig wird die Skalierbarkeit des entwickelten Ansatzes erhöht. Für weitere Untersuchungen oder auch einer späteren Umsetzung in der Praxis ist dementsprechend zu prüfen, welche Genauigkeit erforderlich und aufgrund der vorhandenen Rechenressourcen erreichbar ist.

Hinsichtlich der Auswahl der berücksichtigten Ereignisse haben die Untersuchungen die hohe Relevanz einer möglichst repräsentativen Abdeckung sämtlicher Ereignisverläufe und -intensitäten aufgezeigt. Es ist daher möglich, dass durch die Berücksichtigung weiterer Wiederkehrzeiten, Dauerstufen und Modellregentypen bei den Modellregenreihen bei einer gleichen Anzahl an Ereignissen, bessere Ergebnisse gegenüber dem Ausgangsdatsatz (natürliche Regenreihen und Modellregenreihen) erzielt werden können.

Durch die Reduktion des Trainingsdatensatzes auf 20 % der Ereignisse kann zwar bereits Rechendauer eingespart werden, am Beispiel der Stadt Gelsenkirchen würde der Prozess der Datengenerierung aber immer noch knapp ein Jahr dauern. Aus diesem Grund wurde in einer zweiten Untersuchung geprüft, inwieweit durch einen geeigneten Trainingsansatz die Übertragbarkeit eines trainierten Modells hergestellt werden kann. Hierzu wurden Untersuchungen in Anlehnung an das Vorgehen in Guo et al. (2020) und Löwe et al. (2021) durchgeführt. Dieser Ansatz bietet eine deutlich höhere Skalierbarkeit des Modells. Darüber hinaus könnten bei Änderungen im Einzugsgebiet einfach die Eingabedaten angepasst werden, ohne dass ein erneutes Training erforderlich ist. Erste Tests haben aber bereits gezeigt, dass das in dieser Arbeit verwendete Untersuchungsgebiet dafür zu klein ist. Dadurch ist es nicht möglich einen repräsentativen Datensatz zu erzeugen, der es dem Modell ermöglicht, räumlich zu generalisieren. Auch die Kombination mit einem *Transfer Learning*-Ansatz in Anlehnung an Seleem et al. (2023) führte zu keinen zufriedenstellenden Ergebnissen. Hier ergibt sich dementsprechend weiterer Untersuchungsbedarf mit einem größeren Untersuchungsgebiet.

Neben den erforderlichen Rechenressourcen für die Datengenerierung wird die Skalierbarkeit des Modells durch den hohen Speicherbedarf beim Modelltraining eingeschränkt. Aufgrund der hohen Dimensionalität der als Zielgröße verwendeten Überflutungssequenzen kommt es bei einer Ausweitung auf größere Gebiete und damit einer höheren Anzahl an Rasterzellen schnell zu einer Überschreitung des verfügbaren GPU-Speichers. Dies macht ein Training unter Berücksichtigung der derzeit verfügbaren Technik ab einer gewissen Datenmenge nicht mehr ohne Weiteres möglich. Um dennoch ganze Stadtgebiete abzudecken, ist ein Ansatz die Unterteilung des Berechnungsgebiets in Teilmodelle mit anschließender Zusammenführung

der Ergebnisse, wie in Berkahn et al. (2019) beschrieben. Ein weiterer Ansatz ist die Verwendung eines rekursiven Modells, wie dies in Berkahn und Neuweiler (2024) der Fall ist. Dabei werden die Überflutungsraster nacheinander berechnet, wodurch gegenüber dem hier gewählten Ansatz (zwölf Ausgabebilder) die Modellgröße deutlich reduziert werden kann. Zusätzlich könnte das parallele Training von Modellen auf mehreren GPUs eine Lösung sein. Ein Ansatz hierzu stellt die Python-Bibliothek Mesh-TensorFlow dar (Shazeer et al. 2018), die es ermöglicht, große Modelle mit enormen Speicherbedarf zu entwickeln und diese parallel auf mehreren GPU-Einheiten zu trainieren. Weiterhin würde sich die Problematik durch den zuvor beschriebenen übertragbaren Modellansatz nach Guo et al. (2020) und Löwe et al. (2021) lösen.

Abschließend konnten durch ein im Rahmen der Sensitivitätsanalysen durchgeführtes Hyperparameter-tuning weitere Erkenntnisse gesammelt und die Vorhersagegenauigkeit noch leicht optimiert werden. Einerseits konnte der positive Einfluss eines möglichst langen Eingabefensters bei den Niederschlagsinformationen herausgestellt werden. Andererseits hat sich die Anzahl der Schichten und der Neuronen pro Schicht noch einmal leicht geändert, da so bessere Ergebnisse erzielt werden konnten. Auch hier hat sich allerdings eine hohe Rechendauer als problematisch herausgestellt. Daher musste die Anzahl der berücksichtigten Parameter eingeschränkt werden, um den ganzen Prozess noch handhabbar zu halten. Mit steigenden Rechenressourcen oder effizienteren Optimierungsalgorithmen ergibt sich daher künftig weiteres Optimierungspotential. Gleichzeitig wurde aus Gründen der Vergleichbarkeit das Modelltraining auf maximal 100 Epochen je berücksichtigter Variante begrenzt. Die Lernkurven zeigen jedoch weiter nach unten und es ist noch keine Überanpassung ersichtlich. Auch an dieser Stelle ist daher noch Verbesserungspotential vorhanden.

6 Modellentwicklung und -analyse: Überstauvorhersage

Aufgrund der positiven Auswirkungen auf die Überflutungsflächenvorhersage wurde gemäß dem geplanten übergeordneten Systemaufbau (vgl. Abbildung 2-7) ein Überstauvorhersagemodell entwickelt, um das Überstauvolumen an den einzelnen Schächten im Untersuchungsgebiet zu prognostizieren. Nachfolgend werden wieder zunächst das zugrunde liegende Lernproblem und die verwendete Datenbasis beschrieben. Anschließend erfolgt eine Beschreibung der Datenvorverarbeitungsschritte, bevor dann Untersuchungen zum Modellaufbau, Sensitivitätsanalysen und die abschließende Beurteilung der Vorhersagegenauigkeit folgen.

6.1 Problembeschreibung

In diesem Abschnitt wird zunächst wieder das zugrunde liegende Lernproblem näher erläutert. Ziel des Modells ist es, den Zusammenhang zwischen den Überstauganglinien an den k Schächten S im Einzugsgebiet als Zielgröße sowie den erklärenden Größen der Niederschlagsinformationen (gemessen und prognostiziert) und der Messungen an den l Messstellen M im Kanal zu erlernen. Es handelt sich dementsprechend um multivariate Zeitreihen, wobei die Zielgrößen ausschließlich durch exogene Variablen beschrieben werden. Das Modell soll so im Anschluss an den Trainingsprozess anhand der letzten G Zeitschritte der erklärenden Merkmale die Zielgrößen für den gewählten Vorhersagehorizont H ermitteln. Da der Vorhersagehorizont $H > 1$ gewählt wird, handelt es sich um eine Vorhersage von mehreren Zeitschritten. Ben Taieb et al. (2012) beschreiben in ihrer Arbeit verschiedene Strategien für diesen Anwendungsfall. Die dort aufgeführte *Multi-Input Multi-Output* (MIMO) Strategie kommt auch hier zur Anwendung. Dabei erfolgt die Vorhersage mehrstufig anstelle von mehreren einstufigen Prognosen für jeden Zeitschritt, wie es bei rekursiven oder direkten Vorhersagestrategien üblich ist. Zu jedem Zeitpunkt t erfolgt dementsprechend für alle Schächte S_i eine Vorhersage für die kommenden H Werte $[S_{i,t+1}, \dots, S_{i,t+H}]$ (vgl. Abbildung 6-1). Das schränkt zwar die Flexibilität ein, da das Modell für jeden Zeitschritt dieselbe Struktur aufweist, bietet aber auch Vorteile: Einerseits bleiben gegenüber der direkten Strategie die stochastischen Abhängigkeiten zwischen den einzelnen Zeitschritten erhalten, zum anderen werden die Vorhersagefehler nicht wie bei der rekursiven Strategie kumuliert (Ben Taieb et al. 2012). Der Modellaufbau wird als ein überwachtes Lernproblem behandelt, da Eingaben und dazugehörige Zielgrößen klar definiert sind. Weiterhin soll das Modell die Zielgrößen als stetige Werte vorhersagen, weshalb es sich um ein Regressionsproblem handelt. Bei der Auswahl der ML-Verfahren wurden dementsprechend nur Regressionsverfahren berücksichtigt.

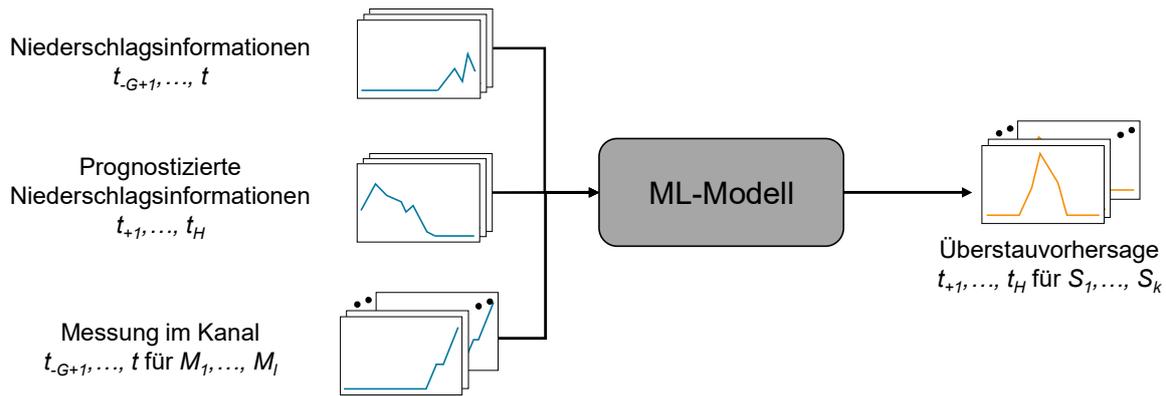


Abbildung 6-1: Darstellung des zu lösenden Lernproblems

Bei den Niederschlagsinformationen handelt es sich um die gleichen Größen, die auch beim Überflutungsflächenvorhersagemodell berücksichtigt wurden (vgl. Kapitel 5.1). Gleiches gilt für die Überstauvorhersage für die 975 Schächte im Untersuchungsgebiet aus der hydrodynamischen Kanalnetzmodellierung, die hier im Vergleich zum Überflutungsflächenvorhersagemodell allerdings als Zielgröße berücksichtigt wurde. Neu hinzu kommen die Messungen im Kanal, durch die der aktuelle Systemzustand des Kanalnetzes bei der Prognoseerzeugung berücksichtigt werden soll. Messungen wurden jeweils für die vergangenen $G-1$ Zeitschritte sowie den Prognosezeitpunkt berücksichtigt. Die Anzahl an Messstellen M sowie die gewählte Messgröße sind Gegenstand der durchgeführten Untersuchungen.

6.2 Datenvorverarbeitung und -aufteilung

Aufteilung des Datensatzes

Die Aufteilung des Datensatzes erfolgte analog zu dem Vorgehen bei dem Überflutungsflächenvorhersagemodell und ist in Kapitel 5.2 beschrieben.

Skalierung der Daten

Wie beim Überflutungsflächenvorhersagemodell wurden auch die Daten für das Überstauvorhersagemodell einheitlich skaliert. Abweichend vom Überstauvorhersagemodell kam hier die Standardisierung zum Einsatz, da dieses Verfahren robuster gegen Ausreißer ist (Raschka und Mirjalili 2018). Zur Ermittlung des standardisierten Wertes x' wird dabei gemäß Formel (6-1) von dem zu standardisierenden Wert x der Mittelwert μ abgezogen und das Ergebnis durch die Standardabweichung σ dividiert:

$$x' = \frac{x - \mu}{\sigma} \quad (6-1)$$

Dadurch sind die standardisierten Werte x' um den Mittelwert 0 mit der Standardabweichung 1 verteilt. Dieser Vorgang wird für alle Merkmale getrennt durchgeführt. Die Standardisierung der Daten wurde in Python mit der *StandardScaler*-Funktion aus dem scikit-learn Paket (vgl. Kapitel 3.4.2) durchgeführt. Die Parameter μ und σ wurden dabei an dem Trainingsdatensatz geschätzt. Anschließend wurden neben dem Trainingsdatensatz auch der Validierungs- und Testdatensatz mit den gleichen Parametern standardisiert.

Erzeugung von Trainingsbeispielen

Da es sich um ein überwachtes Lernproblem handelt, sind für den Trainingsprozess wieder Pärchen aus den verschiedenen Eingaben und Zielgrößen zu bilden. Die Trainingspärchen P werden wie beim Überflutungsflächenvorhersagemodell mit einem gleitenden Fensteransatz für jeden Zeitpunkt t eines Ereignisses erzeugt. Dabei werden die Eingabefenster für die letzten G Zeitschritte über das Intervall $[t_{-G+1}, \dots, t]$ und die Fenster der Zielgrößen für die nächsten H Zeitschritte über das Intervall $[t_{+1}, \dots, t_{+H}]$ aufgespannt. Für den gesamten Datensatz gilt auch hier wieder, dass die Anzahl der Trainingspärchen bei einer Ereignisdauer n und einer festgelegten Schrittweite von 1 für jedes berücksichtigte Niederschlagsereignis $m = n - (G + H)$ ist. Der für die Erzeugung der Trainingspärchen verwendete gleitende Fensteransatz ist in Abbildung 6-2 dargestellt und wurde so auf alle Ereignisse im Datensatz angewandt. Das Eingabefenster der vergangenen Zeitschritte G und der Vorhersagehorizont H wurden auch hier wieder auf eine Stunde festgelegt. Die erforderlichen Ganglinien werden an den l Messstellen (Durchfluss) sowie den k Schächten (Überstau) berücksichtigt.

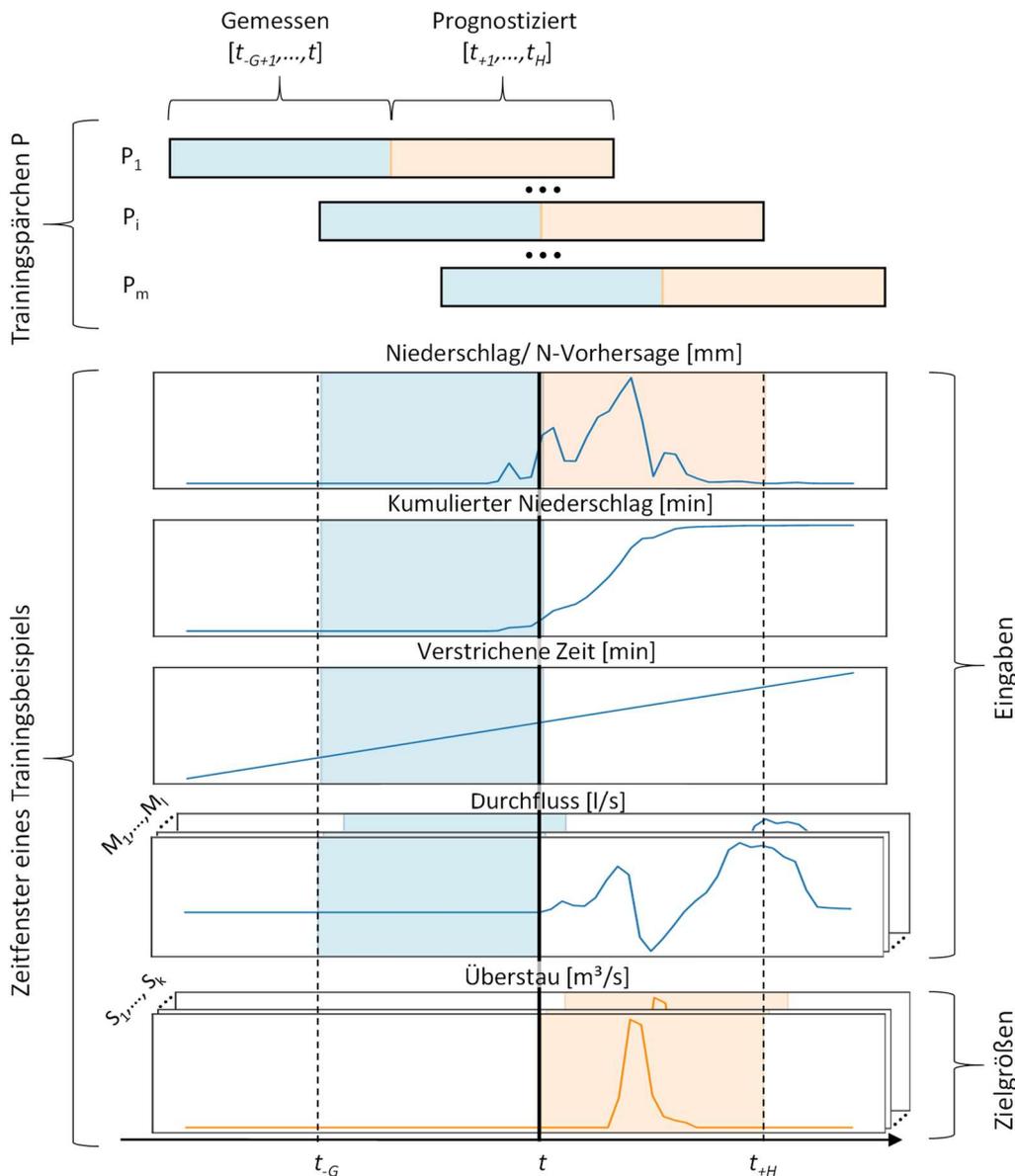


Abbildung 6-2: Umwandlung der Daten in ein überwachtes Lernproblem

6.3 Untersuchungen zum Modellaufbau

Zur Wahl eines geeigneten Modellaufbaus wurden diverse Untersuchungen durchgeführt. Einerseits wurden verschiedene für das vorliegende Lernproblem geeignete ML-Verfahren miteinander verglichen. Weiterhin wurde der Einfluss eines Messnetzes im Kanalnetz als zusätzliche Eingabegröße und dabei konkret der Einfluss der Anzahl an Sensoren sowie des berücksichtigten Messsignals untersucht. Für die unterschiedlichen ML-Verfahren und Messsignale sowie Variationen der Anzahl an Sensoren wurden in allen möglichen Kombinationen Modelle erstellt und deren Leistungsfähigkeit miteinander verglichen. Die durchgeführten Untersuchungen werden nachfolgend in Anlehnung an Burrichter et al. (2024) beschrieben. Um den Lesefluss nicht zu unterbrechen, wird analog zu den Untersuchungen im Rahmen der Prognose von Überflutungsflächen auf eine detaillierte Erläuterung der verschiedenen Netzbestandteile, Hyperparameter und anderer Begriffe im Kontext von Neuronalen Netzen im Text verzichtet. Stattdessen wird an dieser Stelle auf Anhang A-1 verwiesen.

6.3.1 Auswahl und Vergleich von ML-Verfahren/Netzarchitekturen

In einem ersten Schritt wurde untersucht, welche ML-Verfahren sich für das vorliegende Lernproblem am besten eignen. Dabei wurden verschiedene Architekturen von Neuronalen Netzen getestet, mit unterschiedlichen Komplexitätsgraden von einfachen CNNs bis hin zu einem *Temporal Fusion Transformer*. Die Komplexität der Modelle lässt sich dabei unter anderem an der Anzahl an Modellparametern festmachen, mit deren Anzahl häufig die Leistungsfähigkeit, aber gleichzeitig auch der erforderliche Rechenaufwand zunimmt. Dementsprechend ist es das Ziel, hier einen guten Kompromiss zu finden. Zu den berücksichtigten Modellen zählt (i) ein *Long Short-Term Memory* (LSTM)-Network, (ii) ein *Convolutional Neural Network* (CNN), (iii) ein *Sequence to Sequence* (Seq2Seq) Modell, (iv) ein *Dual-Stage Attention-Based* Rekurrentes Neuronales Netz (DA-RNN) sowie (v) ein *Temporal Fusion Transformer*. Darüber hinaus wurde (vi) ein naiver Ansatz als Benchmark herangezogen, der für jeden Prognosezeitpunkt fehlenden Überstau annimmt.

Die Implementierung des TFT erfolgte in PyTorch Forecasting (Breitner 2020), die übrigen Modelle wurden in PyTorch (Paszke et al. 2019) implementiert. Weitere Maschinelle Lernverfahren wie der *Random Forest*-Algorithmus wurden ebenfalls in scikit-learn (Pedregosa et al. 2011) implementiert und getestet. In Anbetracht der Vielzahl an berücksichtigten Zeitreihen haben sich allerdings aufgrund des fehlenden GPU-Supports trotz paralleler Berechnung auf mehreren CPU-Kernen extreme Rechenzeiten ergeben und das Training konnte nicht abgeschlossen werden. Erste Tests an einem kleinen Teil des Untersuchungsgebietes haben aber bereits gezeigt, dass die Leistungsfähigkeit sonstiger ML-Verfahren nicht an die komplexer Neuronaler Netzarchitekturen heranreicht. Aufgrund der beiden gesammelten Erkenntnisse wurde daher im weiteren Verlauf auf die Berücksichtigung klassischer ML-Verfahren verzichtet. Die im Rahmen der Untersuchungen berücksichtigten Netzarchitekturen werden nachfolgend beschrieben. Die Architektur des TFTs wurde bereits in Abbildung 3-13 dargestellt, die übrigen Netzarchitekturen sind exemplarisch in Abbildung 6-3 aufgeführt.

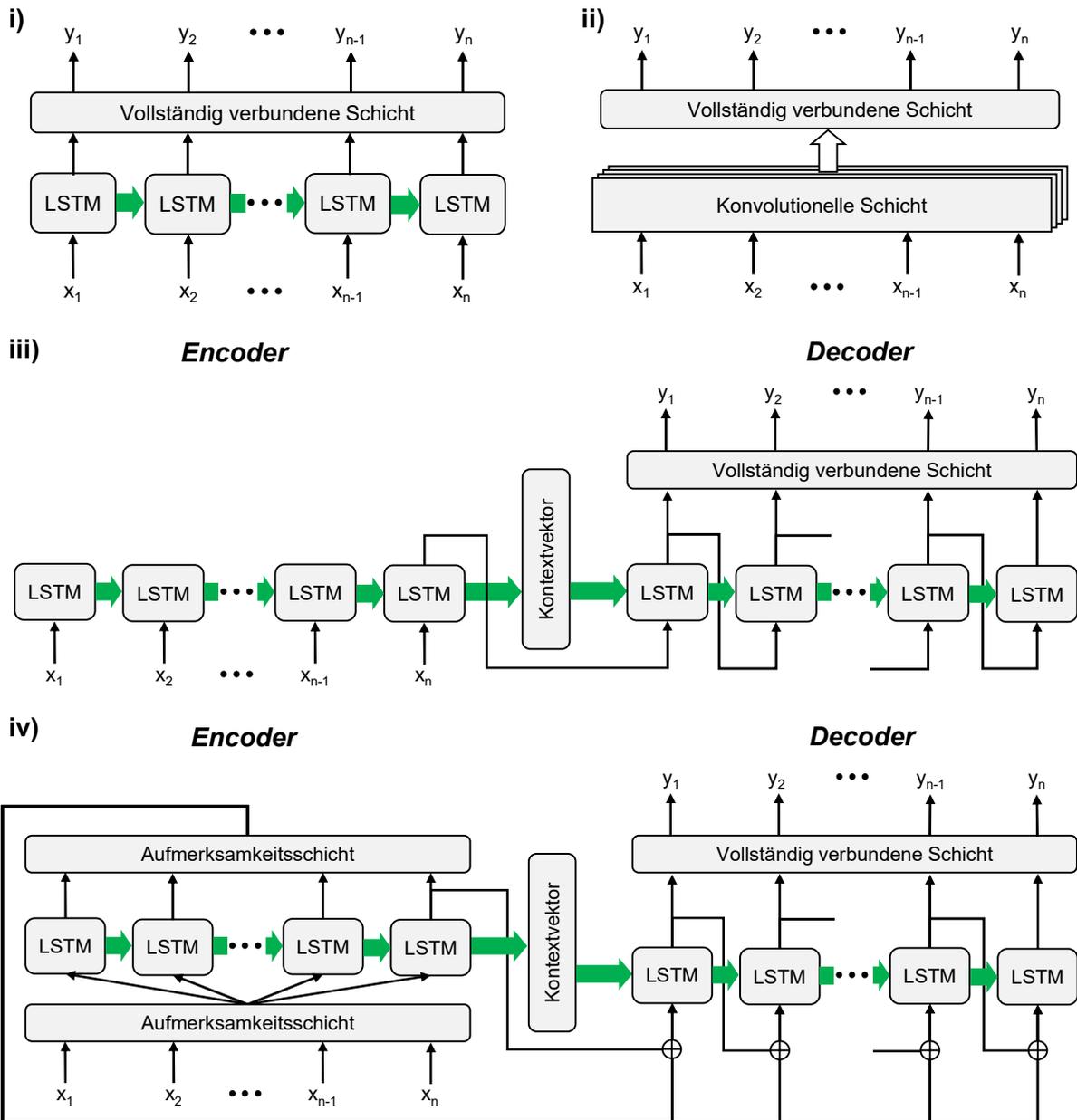


Abbildung 6-3: Exemplarische Darstellung der berücksichtigten Modellaufbauten (i) LSTM, (ii) CNN, (iii) Seq2Seq und (iv) DA-RNN. Die Eingaben x und Ausgaben y können zu jedem Zeitpunkt i auch mehrere Merkmale umfassen. Der grüne Pfeil bei den LSTM-Schichten stellt jeweils die rekurrente Verbindung dar, über die die Zustände jeweils an den folgenden Zeitpunkt weitergegeben werden.

Für die Beschreibung der Kernkonzepte der berücksichtigten Netzarchitekturen wird an dieser Stelle auf das Kapitel 3.3.2.2 verwiesen. Zusätzlich erforderliche Spezifikationen, die zur Anpassung der Architekturen an das vorliegende Lernproblem erforderlich waren, werden nachfolgend aufgeführt:

- CNN:
Im vorliegenden Anwendungsfall werden zunächst *Convolutional*-Schichten verwendet, um kurzfristige Abhängigkeiten und lokale Muster in den Eingabezeitreihen zu identifizieren. Die extrahierten Muster werden anschließend genutzt, um daraus mit einer vorwärtsgerichteten, vollständig verbundenen Schicht die Überstauprognose für

die kommenden Zeitschritte zu generieren. Die Anzahl der Neuronen wurde dementsprechend mit der Anzahl der Ausgabezeitschritte festgelegt.

- LSTM:
Wie beim CNN-Modell wurde auch bei dem hier verwendeten LSTM-Modell eine vorwärtsgerichtete, vollständig verbundene Schicht als Ausgabeschicht angeordnet, um so eine *Multi-Step*-Vorhersage zu erzeugen.
- Sequence to Sequence (Seq2Seq):
Die in Sutskever et al. (2014) vorgestellten Seq2Seq-Modelle stellen eine spezielle Netzarchitektur für die Verarbeitung von sequenziellen Daten dar, die auch rekurrente Schichten umfasst. Gegenüber dem zuvor beschriebenen LSTM-Modell werden die rekurrenten Schichten in einer *Encoder-Decoder*-Struktur angeordnet. Der *Encoder* verarbeitet dabei die Eingaben und erzeugt daraus einen Kontextvektor. Aus diesem generiert der Decoder anschließend eine Ausgabesequenz. Seq2Seq-Modelle können zudem eine Vorhersage für mehrere Zeitschritte erzeugen, ohne auf zusätzliche vorwärtsgerichtete Schichten zurückzugreifen. Als rekurrente Schichten werden bei dem in dieser Arbeit verwendeten Seq2Seq-Modell ebenfalls LSTM-Zellen eingesetzt.
- Dual-Stage Attention-Based Rekurrentes Neuronales Netz (DA-RNN):
Ein DA-RNN besteht aus einem Seq2Seq-Modell, das um einen zweistufigen Aufmerksamkeitsmechanismus ergänzt wurde (Qin et al. 2017). Dabei handelt es sich um Mechanismen, die auch in *Transformern* zur Anwendung kommen (vgl. Kapitel 3.3.2.5). Die Aufmerksamkeitsmechanismen werden dabei vor und hinter dem Encoder angeordnet und dienen dazu, alle Zeitpunkte der Eingabesequenz zu berücksichtigen und diese in Abhängigkeit der Relevanz auf das Prognoseergebnis zu wichten. Im Rahmen der vorliegenden Arbeit wurde auf die Bahdanau-*Attention* (Bahdanau et al. 2014) als Aufmerksamkeitsmechanismus zurückgegriffen.
- Temporal Fusion Transformer (TFT):
Die Flexibilität der TFT-Architektur gegenüber der berücksichtigten Eingabemerkmale ermöglicht eine problemlose Übertragung auf den vorliegenden Anwendungsfall. Da das geplante Überstauvorhersagemodell nur kontinuierliche Merkmale berücksichtigt, entfallen lediglich die Schichten zur Verarbeitung von statischen Metadaten (vgl. Abbildung 3-13).

Die wichtigsten bei der Erzeugung der Modellarchitekturen verwendeten Hyperparameter sind in Tabelle 6-1 aufgeführt. Die Stapelgröße (*Batch Size*) wurde bei der Generierung des für das Training verwendeten Datensatzes mit 16 festgelegt. Zusätzlich wurde einheitlich für alle Modelle die Anzahl der Epochen für das Training auf 100 gesetzt. Um eine Überanpassung zu vermeiden, wurde das Training abgebrochen, wenn es über 20 Epochen hinweg zu keiner Verbesserung des Validierungsfehlers kam.

Tabelle 6-1: Übersicht der bei den einzelnen Modellaufbauten berücksichtigten Hyperparameter

| Modell | Parameter | Wert |
|---|---|---------------------------------------|
| CNN | Conv.-Schichten | 2 |
| | Filter pro Schicht | 128 |
| | Kernelgröße/ Schrittweite/ <i>Padding</i> | 3/ 1/ <i>Same</i> |
| | Aktivierungsfunktion | ReLU |
| | Verlustfunktion | <i>Mean Squared Error</i> |
| | Optimierungsalgorithmus | ADAM (Kingma und Ba 2014) |
| Rekurrente Modelle: LSTM/ Seq2Seq/ DA-RNN | Lernrate | 0,0002 |
| | LSTM-Schichten | 2 |
| | Einheiten pro Schicht | 128 |
| | Verlustfunktion | <i>Mean Squared Error</i> |
| | Optimierungsalgorithmus | ADAM (Kingma und Ba 2014) |
| | Lernrate | 0,0002 |
| TFT | LSTM-Schichten | 2 |
| | Einheiten pro Schicht | 128 |
| | Einheiten pro kont. Schicht | 128 |
| | <i>Attention Head Size</i> | 2 |
| | Verlustfunktion | Quantile Loss |
| | Quantile | 0,02; 0,1; 0,25; 0,5; 0,75; 0,9; 0,98 |
| | Optimierungsalgorithmus | Ranger (Wright 2019) |
| | Lernrate | 0,02 |

6.3.2 Untersuchungen zum Einfluss von Messungen im Kanal

Im Rahmen des Forschungsvorhabens KIWaSuS (Quirnbach 2021) in dem die hier beschriebenen Untersuchungen durchgeführt wurden, waren als zusätzliche Datenquelle für das zu entwickelnde Vorhersagemodell *LowCost*-Sensorsysteme für Messungen im Kanal vorgesehen. Diese konnten aber während der Projektlaufzeit nicht im hinreichenden Umfang umgesetzt werden. Um dennoch den Mehrwert eines Messnetzes überprüfen zu können, wurden Untersuchungen mit künstlich generierten Messstellen durchgeführt. Dazu wurden an festgelegten Sensorstandorten die Simulationsergebnisse aus dem hydrodynamischen Kanalnetzmodell abgegriffen, um diese als zusätzliche Modelleingabe beim Training zu berücksichtigen. Ziel der Untersuchungen war es, den Einfluss der räumlichen Auflösung des Messnetzes sowie des verwendeten Messsignals zu evaluieren. Hierzu wurden die räumliche Auflösung sowie das verwendete Signal des Messnetzes variiert und mit den so erzeugten Varianten Modelle erstellt, deren Leistung anschließend miteinander verglichen wurde.

In Bezug auf die räumliche Auflösung wurde zunächst das in Abbildung 6-4 dargestellte Messnetz bestehend aus 20 fiktiven Sensoren als Variante (i) berücksichtigt. Verglichen wurde diese Variante mit Variante (ii), die nur Messungen am Gebietsauslass berücksichtigt und Variante (iii), die gar keine Messungen im Kanal berücksichtigt. Bei der Wahl der Sensorstandorte wurde darauf geachtet, dass diese räumlich möglichst gleichmäßig über das Untersuchungsgebiet verteilt sind. Zudem wurden nur Standorte berücksichtigt, zu denen mehrere Haltungen hin entwässern. An dieser Stelle gilt es zu beachten, dass bei der gewählten Vorgehensweise Wasserstand und Durchfluss an allen Messstellen aus der Belastungsgröße Niederschlag über klar vorgegebene Rechenvorschriften im hydrodynamischen Kanalnetzmodell berechnet

werden. Dadurch korrelieren die berechneten „Messwerte“ stark und es wird beim Variantenvergleich gezielt geprüft, inwieweit durch eine größere Anzahl an Sensoren ein Mehrwert erzeugt werden kann oder ob sich die hohe Anzahl sogar negativ auf das Ergebnis auswirkt.

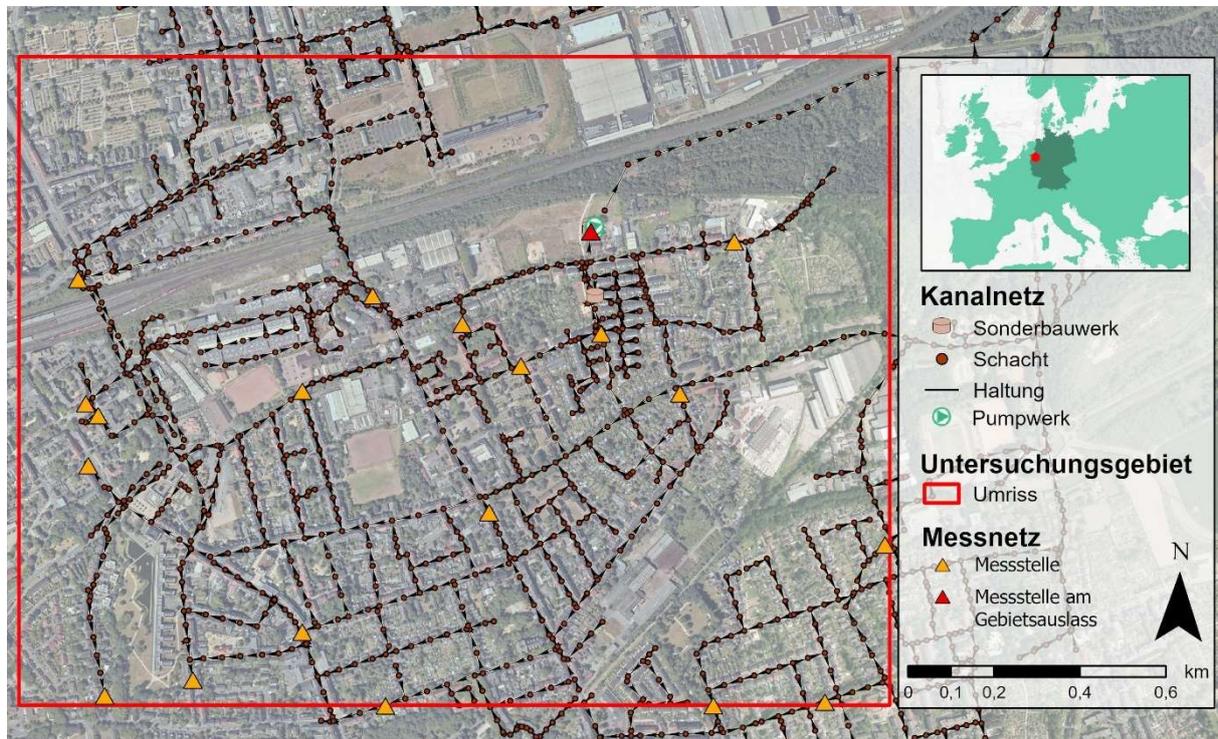


Abbildung 6-4: Übersicht über das Kanalnetz und das künstliche Sensornetz im Untersuchungsgebiet, Datenquelle: AGG (2021)

Gegenüber kostenintensiven Präzisionsmessgeräten, die Durchfluss mit einer hohen Genauigkeit messen können, greifen *LowCost*-Sensoren meist auf andere Größen zurück. Im Falle des Projektes KIWaSuS wurde ein Sensor entwickelt, der akustische Messsignale in Füllstandsklassen umrechnet (Frentrup et al. 2022). Dadurch wird ein robustes und kostengünstiges Messverfahren verwendet, das aber gleichzeitig auch nur ein minderwertiges Messsignal liefert. Um zu prüfen, inwieweit sich die geringere Auflösung der Messungen auf die Vorhersageergebnisse auswirkt, wurden (i) Durchfluss- und (ii) Wasserstandsmessungen sowie Messungen des (iii) Füllungsgrades mit dem Ansatz der (iv) Füllstandsklassen verglichen. Durchfluss- und Wasserstandsmessungen wurden im Modell direkt an den Sensorstandorten für die simulierten Ereignisse abgegriffen. Der Füllungsgrad wurde hingegen aus dem Wasserstand berechnet und gibt die prozentuale Füllung der Haltungen an. Anhand des Füllungsgrades wurden wiederum die Füllstandsklassen wie folgt bestimmt:

$$f(x) = \begin{cases} 1, & x < 25 \% \\ 2, & 25 \% \leq x < 50 \% \\ 3, & 50 \% \leq x < 75 \% \\ 4, & 75 \% \leq x < 100 \% \\ 5, & x \geq 100 \% \end{cases} \quad (6-2)$$

6.3.3 Vergleich der Ergebnisse

Wie Eingangs erwähnt, wurden für die berücksichtigten Netzarchitekturen, Variationen der Anzahl an Sensoren sowie die unterschiedlichen Messsignale Modelle in allen möglichen Kombinationen erstellt und deren Leistungsfähigkeit miteinander verglichen. Die herangezogenen Gütekriterien Volumenfehler (VE), Maximalwertabweichung (PE) und zeitliche Maximalwertabweichung PTE (vgl. Kapitel 3.5) wurden jeweils für alle *Samples* der 26 Testereignisse und an allen Schächten ermittelt. Tabelle 6-2 fasst jeweils für alle Modelle die Mittelwerte der Gütekriterien über alle 26 Ereignisse und alle berücksichtigten Schächte zusammen. Dabei fließen nur die Prognoseergebnisse für *Samples* und Schächte ein, an denen das Überstauvolumen für den betrachteten Prognosezeitraum entweder im HD-Modell oder dem ML-Modell > 500 l liegt. Ziel ist es, so nur relevante Überstauereignisse bei der Bewertung zu berücksichtigen. Im Falle des TFT erfolgte die Berechnung der Metriken anhand des Fehlers für das 0,5-Quantil, das dem mittleren absoluten Fehler entspricht. Es ist zu beachten, dass der naive Ansatz in allen Varianten dasselbe Ergebnis liefert, da er stets keinen Überstau prognostiziert. Zudem variieren die Ergebnisse bei den Varianten ohne Messstationen nicht zwischen den berücksichtigten Messsignalen, da keine Messungen als Eingabe einbezogen wurden.

Anhand der Tabelle können folgende Rückschlüsse hinsichtlich der einzelnen Untersuchungen gezogen werden:

- Vergleich der Modellarchitekturen:
Der Vergleich der unterschiedlichen Modellarchitekturen untereinander zeigt zunächst einmal, dass der naive Ansatz sowie CNN und LSTM deutlich schlechtere Prognosen gegenüber den übrigen Modellarchitekturen liefern. CNN und LSTM weisen dabei teilweise sogar schlechtere Ergebnisse als der naive Ansatz auf. Auch wenn der TFT nicht über alle berücksichtigten Varianten am besten abschneidet, so wird das beste Ergebnis für jede Metrik jeweils mit einem TFT erzielt. Die Ergebnisse für das Seq2Seq-Modell und den TFT liegen jedoch meist nahe beieinander.
- Vergleich der Anzahl an Sensoren:
Die Anzahl der Sensoren hat keinen positiven Einfluss auf die Ergebnisse. So wurden bei den Varianten mit 20 Sensoren tendenziell eher schlechtere Ergebnisse erzielt. Die Varianten mit einem Sensor und keinem Sensor liegen hingegen in den meisten Fällen eng beieinander. Die besten Gesamtergebnisse für alle drei Metriken konnten bei einer Variante mit einem Sensor erzielt werden.
- Vergleich des berücksichtigten Messsignals:
Bei den berücksichtigten Messsignalen ist keine klare Tendenz zu einer Größe zu erkennen. Bei den Varianten mit einer Station stechen lediglich die Ergebnisse für den Wasserstand bei den beiden besten Modellarchitekturen Seq2Seq und TFT negativ heraus. Der geringste Volumenfehler sowie die geringste zeitliche Maximalwertabweichung werden bei der Messung des Füllungsgrads erreicht. Die geringste Maximalwertabweichung tritt bei dem Ansatz mit fünf Füllstandsklassen auf.

Tabelle 6-2: Ergebnis der Evaluierung aller Modelle und Varianten (für jede Variante und Metrik wurde das beste Ergebnis fett markiert, das beste Gesamtergebnis jeder Metrik wurde zusätzlich unterstrichen)

| Modell | 20 Stationen | | | 1 Station | | | Keine Station | | |
|------------------------------|--------------|--------------|-------------|--------------|--------------|-------------|---------------|--------------|-------------|
| | VE | PE | PTE | VE | PE | PTE | VE | PE | PTE |
| | [l] | [l/s] | [min] | [l] | [l/s] | [min] | [l] | [l/s] | [min] |
| Durchfluss [l/s] | | | | | | | | | |
| Naïve Zero | 138,65 | 199,30 | 27,50 | 138,65 | 199,30 | 27,50 | 138,65 | 199,30 | 27,50 |
| CNN | 235,58 | 148,54 | 20,32 | 228,65 | 169,03 | 25,51 | 208,87 | 152,20 | 20,70 |
| LSTM | 147,61 | 179,65 | 10,55 | 174,80 | 246,92 | 7,84 | 169,46 | 181,34 | 7,84 |
| Seq2Seq | 90,73 | 85,41 | 7,13 | 49,07 | 83,81 | 5,96 | 50,51 | 74,04 | 6,12 |
| DA-RNN | 133,41 | 154,09 | 5,08 | 85,79 | 99,92 | 6,11 | 96,08 | 112,13 | 8,01 |
| TFT | 92,84 | 136,83 | 15,82 | 44,25 | 79,77 | 2,53 | 61,13 | 91,45 | 4,04 |
| Wasserstand [m] | | | | | | | | | |
| Naïve Zero | 138,65 | 199,30 | 27,50 | 138,65 | 199,30 | 27,50 | 138,65 | 199,30 | 27,50 |
| CNN | 200,96 | 137,53 | 22,10 | 259,23 | 154,80 | 22,07 | 208,87 | 152,20 | 20,70 |
| LSTM | 157,30 | 170,15 | 8,20 | 137,15 | 178,62 | 9,29 | 169,46 | 181,34 | 7,84 |
| Seq2Seq | 77,10 | 76,65 | 5,18 | 89,24 | 96,58 | 4,03 | 50,51 | 74,04 | 6,12 |
| DA-RNN | 101,18 | 117,02 | 4,52 | 62,94 | 84,53 | 7,34 | 96,08 | 112,13 | 8,01 |
| TFT | 62,05 | 101,91 | 2,12 | 57,76 | 101,35 | 2,64 | 61,13 | 91,45 | 4,04 |
| Füllungsgrad [%] | | | | | | | | | |
| Naïve Zero | 138,65 | 199,30 | 27,50 | 138,65 | 199,30 | 27,50 | 138,65 | 199,30 | 27,50 |
| CNN | 225,52 | 160,67 | 19,54 | 270,09 | 168,48 | 26,15 | 208,87 | 152,20 | 20,70 |
| LSTM | 147,70 | 168,95 | 10,29 | 157,77 | 184,91 | 7,36 | 169,46 | 181,34 | 7,84 |
| Seq2Seq | 97,99 | 98,01 | 9,56 | 43,39 | 90,01 | 2,16 | 50,51 | 74,04 | 6,12 |
| DA-RNN | 80,77 | 100,03 | 4,58 | 108,75 | 103,39 | 8,65 | 96,08 | 112,13 | 8,01 |
| TFT | 43,19 | 78,49 | 3,10 | 36,93 | 77,42 | 2,07 | 61,13 | 91,45 | 4,04 |
| Füllstandsklassen [-] | | | | | | | | | |
| Naïve Zero | 138,65 | 199,30 | 27,50 | 138,65 | 199,30 | 27,50 | 138,65 | 199,30 | 27,50 |
| CNN | 271,46 | 171,68 | 19,61 | 300,05 | 178,35 | 25,37 | 208,87 | 152,20 | 20,70 |
| LSTM | 151,27 | 170,61 | 10,43 | 160,64 | 176,63 | 6,67 | 169,46 | 181,34 | 7,84 |
| Seq2Seq | 78,38 | 91,00 | 10,31 | 59,15 | 78,35 | 4,08 | 50,51 | 74,04 | 6,12 |
| DA-RNN | 102,93 | 126,97 | 7,80 | 65,84 | 93,26 | 5,05 | 96,08 | 112,13 | 8,01 |
| TFT | 86,68 | 125,77 | 4,90 | 44,25 | 73,72 | 2,71 | 61,13 | 91,45 | 4,04 |

Aufgrund der Tatsache, dass das gewählte Messsignal nur geringe Auswirkungen auf die Modellgüte hat, wurden für die weiteren Untersuchungen nur Modelle mit Füllstandsklassen als Eingabe berücksichtigt. Der Hintergrund dafür ist, dass es sich dabei um das minderwertigste Messsignal handelt, das am einfachsten und somit am ehesten mit LowCost-Sensorik erfasst werden kann. Zudem wurde im weiteren Verlauf die Variante mit einer Station berücksichtigt, da mit dieser Variante jeweils das beste Gesamtergebnis erzielt werden konnte.

Im weiteren Verlauf wurde zunächst die Modellgüte der drei besten Modelle Seq2Seq, DA-RNN und TFT für die Variante mit Füllstandsklassen an einer Messstation tiefergehend betrachtet. Hierzu wurde der in Abbildung 6-5 dargestellte Violinen-Plot für die einzelnen Metriken und Modelle erstellt. Berücksichtigt sind jeweils die zur Mittelwertbildung in Tabelle 6-2 verwendeten Metriken für die einzelnen Vorhersagezeitpunkte an den überstauten Schächten

im Untersuchungsgebiet. Anhand des Violinen-Plots werden Streuung und Dichte der Ergebnisse sichtbar. Hier zeigt sich die bessere Performance des TFT, der für alle Metriken die geringste Streuung aufweist und bei dem die größte Dichte jeweils in der Nähe des Optimums liegt, etwas deutlicher. Große Ausreißer wie bei den Seq2Seq- und DA-RNN-Modellen treten beim TFT lediglich bei der zeitlichen Abweichung des Maximalwertes auf. Diese ergeben sich bei langen gleichmäßigen Überstauereignissen, wodurch deren Relevanz wieder etwas relativiert wird. Gegenüber den anderen Modellen unterschätzt der TFT beim Volumen und Maximalwert geringfügig. Gerade beim Volumen sind aber vor dem Hintergrund, dass bei der Evaluierung nur Prognosen mit > 500 l Überstauvolumen berücksichtigt wurden und es sich um Extremereignisse handelt, die Abweichungen vertretbar. Da die Ergebnisse bei anderen Messsignalen ähnlich aussehen, wurde auf eine zusätzliche Darstellung verzichtet.

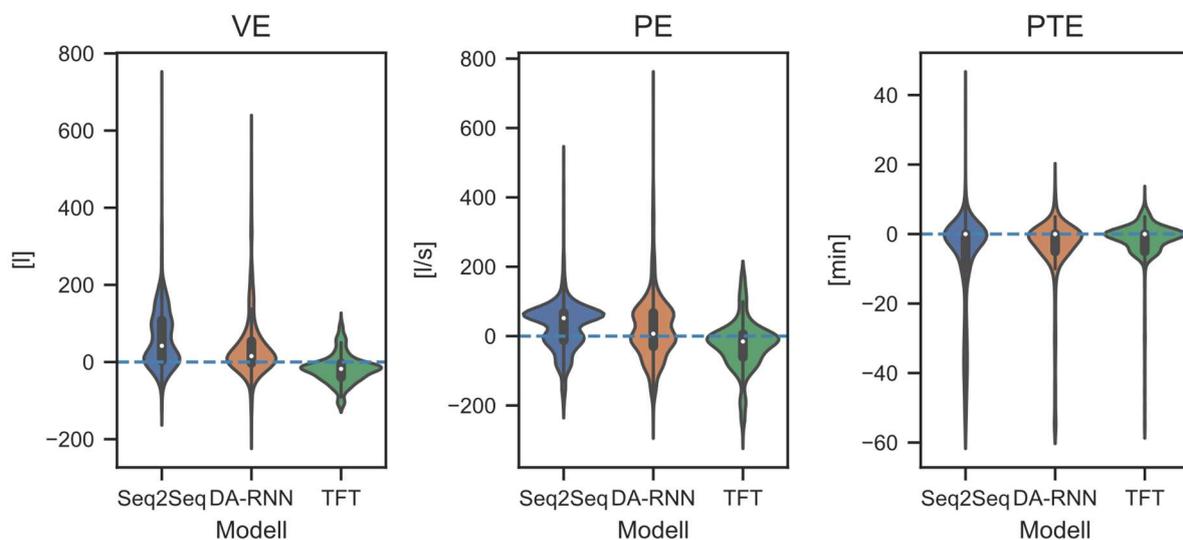


Abbildung 6-5: Violinen-Plot der Metriken für die besten drei Modelle der Variante mit einer Messstation und Füllungsklassen als Messsignal

In einem weiteren Schritt wurde der Einfluss der unterschiedlichen Eingabesignale auf das Prognoseergebnis des TFTs untersucht. Ziel war es insbesondere, herauszustellen, weshalb die Variante mit 20 Messstationen bei den Untersuchungen zu schlechteren Ergebnissen führte. Aufgrund der in Kapitel 3.3.2.5 beschriebenen VSN-Blöcke ist es beim TFT möglich, den Einfluss der einzelnen Eingabemerkmale auf die Ergebnisgenerierung auszugeben. Abbildung 6-6 zeigt die prozentuale Bedeutung der einzelnen Eingabemerkmale des *Encoders* und des *Decoders* bei der Prognoseerzeugung. Aufgeführt sind die Varianten mit 20 und einer Messstation unter Berücksichtigung von Füllstandsklassen als Messsignal. An dieser Stelle sei darauf hingewiesen, dass der *Encoder* jeweils Merkmale der vergangenen Zeitschritte und der *Decoder* Merkmale der bevorstehenden Zeitschritte als Eingabe erhält. Bei beiden Modellen zeigt sich erwartungsgemäß die hohe Relevanz der Niederschlagsprognose, die als Eingabe beim *Decoder* berücksichtigt wird. Beim *Encoder* zeigt sich, dass bei dem Ansatz mit 20 Stationen nur wenigen Merkmalen eine relevante Bedeutung zukommt, während bei der Variante mit einer Station die meisten Merkmale Einfluss auf die Prognoseergebnisse haben. Ein Grund dafür könnte die hohe Korrelation der einzelnen Messungen untereinander, aber auch mit dem gefallenen Niederschlag sein. Diese Korrelation ergibt sich in dem vorliegenden Fall aufgrund der Tatsache, dass ein virtuelles Messnetz mit simulierten Messwerten verwendet

wurde und wird in Kapitel 6.6 weiter diskutiert. Dadurch ist der Mehrwert der vielen Messungen gering und führt zu unnötiger Komplexität des Modells. Da der verwendete Trainingsdatensatz zudem weniger als 10.000 *Samples* umfasst und damit für das Training von *Deep Learning*-Modellen und insbesondere *Transformern* eher als klein einzustufen ist, wird es für das Modell bei vielen Merkmalen umso schwerer, die besonders wichtigen Merkmale und deren zugrunde liegenden Muster zu erlernen.

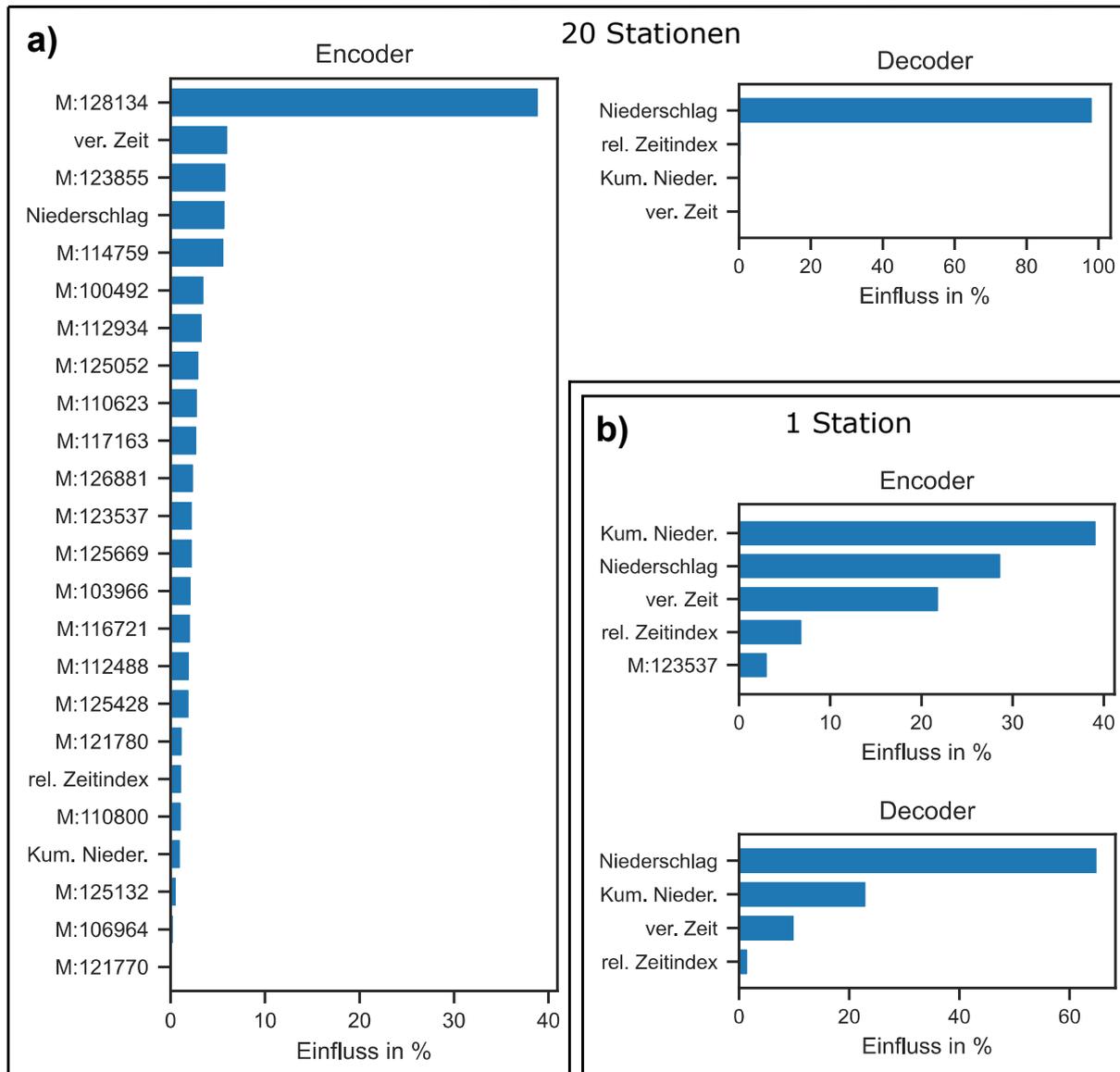


Abbildung 6-6: Einfluss der einzelnen Merkmale der Encoder und Decoder Eingaben für a) die Variante mit 20 Stationen und b) die Variante mit einer Station (die Merkmale mit M zu Beginn geben jeweils Messungen der Füllstandsklassen mit der jeweiligen Haltungs-ID an)

6.4 Sensitivitätsanalysen

Auch beim Überstauprognosemodell wurden verschiedene Sensitivitätsanalysen durchgeführt. Das Vorgehen dabei wurde in Anlehnung an Kapitel 5.4 gewählt und der Fokus wieder auf die Skalierbarkeit der finalen Modelle gelegt. Da der rechentechnische Aufwand bei der

Modellentwicklung maßgeblich von der Erzeugung des Trainingsdatensatzes abhängig ist, wurden zunächst wieder Untersuchungen mit dem Ziel durchgeführt, dessen Umfang zu reduzieren. Ferner wurde geprüft, inwieweit sich das trainierte Modell übertragen lässt und welche Anpassungen ggf. durchzuführen sind. Darüber hinaus wurde abschließend ein Hyperparameter-tuning durchgeführt, um die Genauigkeit des finalen Modells weitestgehend zu optimieren. Als Modellvariante wurde in allen Untersuchungen der TFT mit einer Messstation und Füllstandsklassen als berücksichtigtes Messsignal verwendet. Hyperparameter, die nicht im Text erläutert werden, sind in Anhang A-1 beschrieben.

6.4.1 Einfluss der Trainingsdatenquantität und -verteilung

Wie in Kapitel 4.4 beschrieben ist der Prozess der Datengenerierung extrem rechenintensiv, weshalb es in Hinblick auf die Skalierbarkeit wünschenswert ist, den Umfang des Trainingsdatensatzes möglichst gering zu halten. Aufbauend auf den Erkenntnissen zur Minimierung des Trainingsdatensatzes im Bereich der Vorhersage von Überflutungsflächen (vgl. Kapitel 5.4.1) wurden auch hier Untersuchungen durchgeführt, um den Umfang des Datensatzes zu reduzieren. Die im Rahmen der durchgeführten Untersuchungen berücksichtigten Varianten werden im Folgenden nur kurz aufgezählt. Für eine detaillierte Beschreibung der jeweils berücksichtigten Anzahl an Ereignissen sowie die Verteilung der Ereignisse innerhalb der einzelnen Varianten wird an dieser Stelle auf das Kapitel 5.4.1 verwiesen.

In einer ersten Untersuchung wurde wieder damit begonnen, die Leistungsfähigkeit des Modells nach dem Training am gesamten Datensatz (Variante Alle) mit den Varianten der natürlichen Regenreihen (NR) und der Modellregen (MR) zu vergleichen. Die Ergebnisse sind in Abbildung 6-7 dargestellt und zeigen ein ähnliches Bild wie bei dem Überflutungsflächenvorhersagemodell. Auch hier führt die Variante NR zu den schlechtesten Ergebnissen. Zwar führt ein Ausreißer bei der Variante MR zu einer etwas größeren Streuung des Volumenfehlers und der Maximalwertabweichung, die Dichte der Ergebnisse im Bereich des Optimums ist jedoch eindeutig größer. Der Median der Variante MR liegt sowohl beim Volumenfehler als auch bei der Maximalwertabweichung sogar etwas näher am Optimum als beim Training am gesamten Datensatz. Beim Volumenfehler und der Maximalwertabweichung fällt zudem auf, dass die Variante NR eher überschätzt, während die beiden übrigen Varianten leicht unterschätzen. Eine mögliche Ursache hierfür könnte das unterschiedliche Verhältnis zwischen Ereignissen mit Wiederkehrzeiten < 100 a und > 100 a in den jeweiligen Datensätzen sein. Während bei der Variante NR die Anzahl der Ereignisse < 100 a deutlich größer ist, finden durch die alleinige oder zusätzliche Berücksichtigung von Modellregenereignissen auch Ereignisse > 100 a eine stärkere Berücksichtigung. Insgesamt zeigt sich also, dass die Variante mit Modellregen trotz extremer Reduktion der Datenmenge nur geringfügig von der Variante mit allen Ereignissen abweicht. Auch wenn die Streuung der Ergebnisse aufgrund einzelner Ausreißer bei der Variante MR etwas höher ist, lässt sich an dieser Stelle hinsichtlich der Ergebnisgüte keine eindeutige Entscheidung zugunsten einer Variante fällen.

Da Abbildung 6-7 die Erkenntnisse aus Kapitel 5.4.1 bestätigt und zeigt, dass mit der alleinigen

Berücksichtigung von natürlichen Regenreihen gegenüber Modellregenreihen schlechtere Ergebnisse erzielt werden, wurde auf eine weitere Ausdünnung verzichtet. Stattdessen wurden die Untersuchungen mit den Modellregenreihen fortgeführt und der Datensatz gemäß Abbildung 5-11 nach den folgenden Varianten ausgedünnt:

- Variante 1 (Alle Modellregenereignisse): 105 Ereignisse, 3318 Trainingspärchen
- Variante 2 (Wiederkehrzeit): 70 Ereignisse, 1896 Trainingspärchen
- Variante 3 (Dauerstufe): 42 Ereignisse, 1176 Trainingspärchen
- Variante 4 (Euler Typ): 35 Ereignisse, 1106 Trainingspärchen
- Variante 5 (Gleichmäßig): 52 Ereignisse, 1636 Trainingspärchen

Die Ergebnisse sind in Abbildung 6-8 dargestellt und spiegeln die Erkenntnisse der Analysen des Überflutungsflächenvorhersagemodells wider. Auch hier führt nach der Variante 1 die Variante 5 mit der gleichmäßigen Ausdünnung des Datensatzes zu den besten Ergebnissen. Zwar fällt die Streuung der Variante 5 sogar geringer als bei der Variante 1 aus, dies hängt jedoch wie weiter oben bereits beschrieben von einzelnen Ausreißern ab und sollte nicht über die Gesamtleistung hinwegtäuschen. Anhand der Ergebnisse wird noch einmal verdeutlicht, dass die Modellgüte weniger von der Anzahl, sondern vielmehr von der repräsentativen Verteilung der Trainingspärchen abhängt.

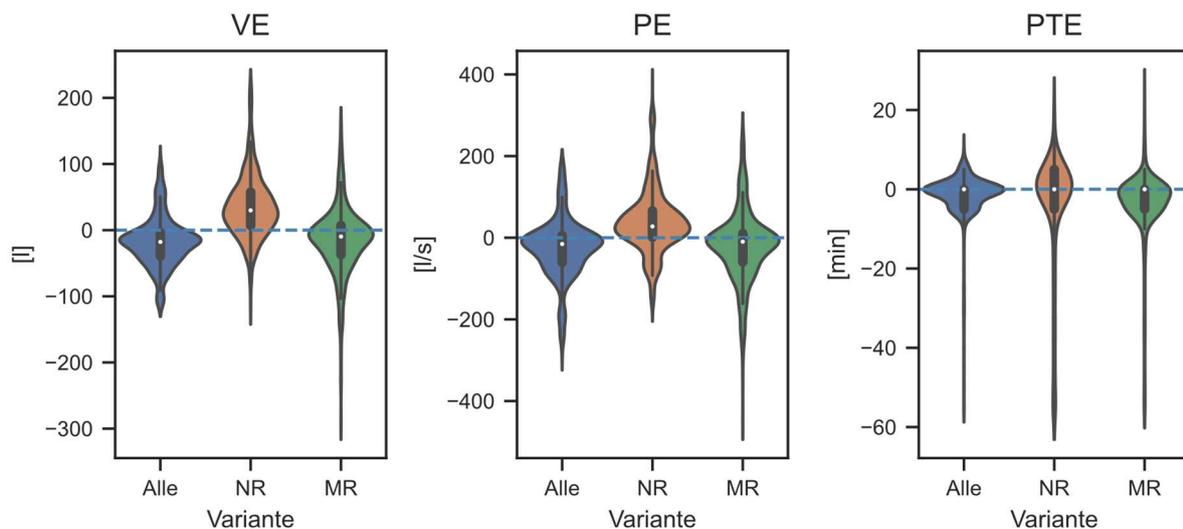


Abbildung 6-7: Violinen-Plot der Metriken für die drei berücksichtigten Varianten (berücksichtigt werden nur Prognosen, bei denen das kumulierte Volumen des HD-Modells oder des ML-Modells > 100 l liegt)

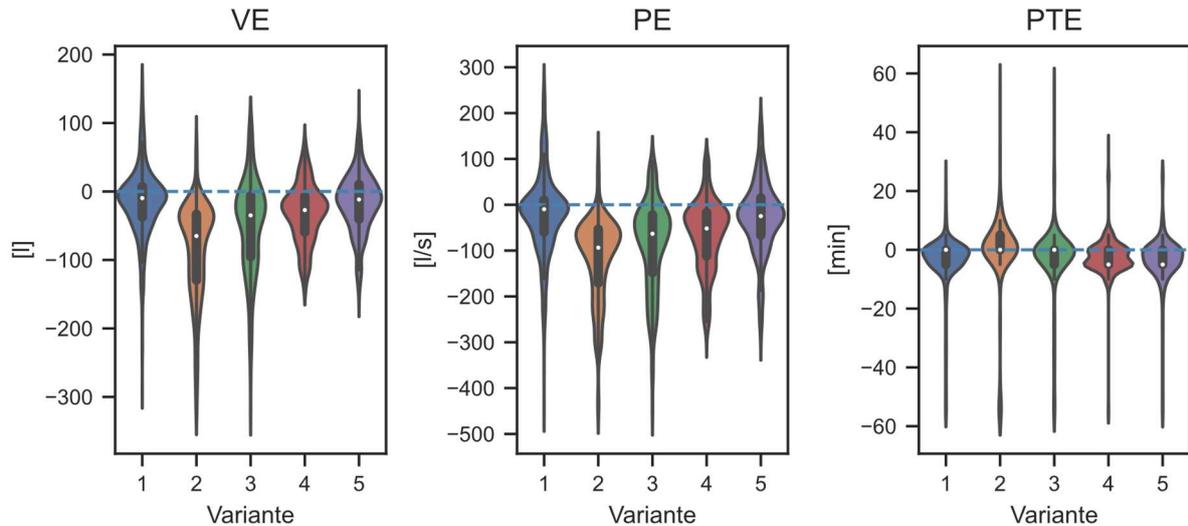


Abbildung 6-8: Violinen-Plot der Metriken für die berücksichtigten Varianten zur Auswahl von Ereignissen aus den Modellregenreihen (berücksichtigt werden nur Prognosen, bei denen das kumulierte Volumen des HD-Modells oder des ML-Modells > 100 l liegt)

Da bei den Untersuchungen in Kapitel 5.4.1 eine Kombination aus gleichmäßig ausgedünnten Modellregenreihen und gleichverteilt ausgedünnten natürlichen Regenreihen bei der Prognose von hohen Wasserständen am besten abgeschnitten hat, wurde diese Variante auch hier berücksichtigt. Die Ergebnisse sind in Abbildung 6-9 dargestellt. Es zeigt sich, dass sich die zusätzliche Berücksichtigung von natürlichen Regenreihen negativ auf die Ergebnisse auswirkt. Zwar ist die zeitliche Abweichung zum Maximalwert im Mittel etwas geringer, die Maximalwertabweichung selbst und der Volumenfehler werden aber deutlich unterschätzt. Das gilt sowohl im Vergleich zu der Variante mit sämtlichen Ereignissen als auch zu der Variante mit den gleichmäßig ausgedünnten Modellregenreihen. An dieser Stelle ist allerdings zu erwähnen, dass bei dem verwendeten Ansatz zwar die Anzahl der Ereignisse je Wiederkehrzeitintervall vorgegeben wird, die Ereignisse selbst aber zufällig ausgewählt werden. Es kann daher sein, dass andere Stichproben zu anderen Ergebnissen führen würden. Aufgrund der deutlich schlechteren Ergebnisse der Variante mit natürlichen Regenreihen im Vergleich zur Variante mit Modellregen (vgl. Abbildung 6-7) und der Tatsache, dass selbst die Variante mit gleichmäßig ausgedünnten Modellregen nur zu geringfügigen Leistungsverlusten führt, wurde an dieser Stelle auf weitere Untersuchungen verzichtet.

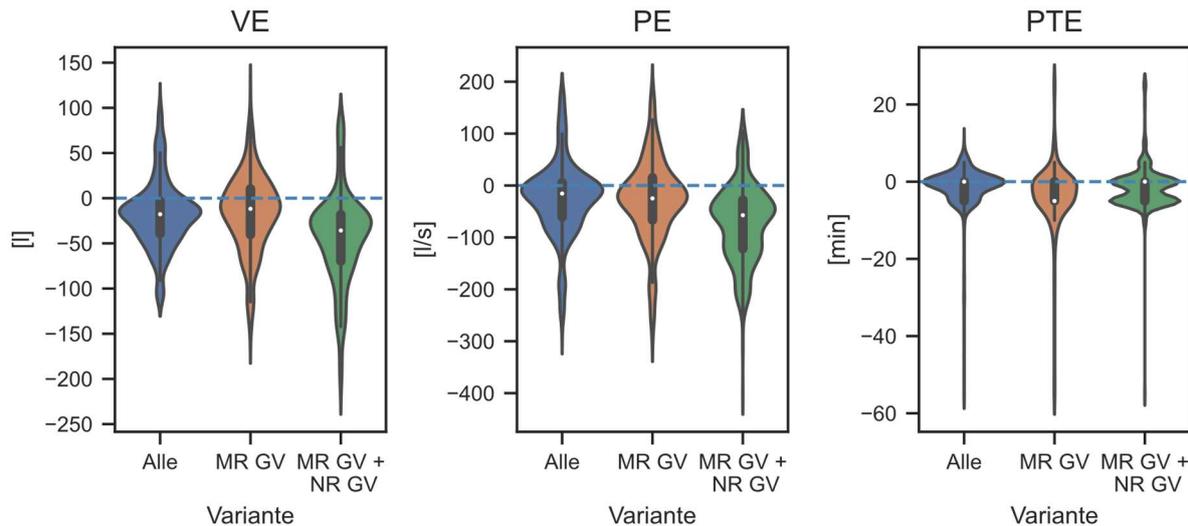


Abbildung 6-9: Violinen-Plot der Metriken für die Modellregenreihen und die natürliche Regenreihen gleichverteilt ausgedünnt. Als Vergleichsgrößen sind die Variante mit dem gesamten Datensatz sowie den gleichmäßig ausgedünnten Modellregenreihen dargestellt (berücksichtigt werden nur Prognosen, bei denen entweder das kumulierte Volumen des HD-Modells oder des ML-Modells > 100 l liegt)

6.4.2 Übertragbarkeit des Modells auf andere Gebiete

Analog zum Überflutungsflächenvorhersagemodell wurde geprüft, inwieweit sich das trainierte Modell übertragen lässt. Dabei ergibt sich jedoch die Problematik, dass im gewählten Modellaufbau keine Informationen über den physikalischen Systemaufbau berücksichtigt werden. Vielmehr werden lediglich Eingaben über den aktuellen Systemzustand sowie die anstehende Belastung berücksichtigt. Anders als beim Modell zur Vorhersage von Überflutungsflächen, das den Systemaufbau über räumliche Informationen als Eingaben berücksichtigt, ist ohne grundlegende Änderungen am Modellaufbau beim Überstauvorhersagemodell keine Übertragbarkeit möglich. Hierfür müssten Informationen über das Kanalnetz wie beispielsweise Lage, Gefälle und Länge von Haltungen als zusätzliche Eingabe berücksichtigt werden. Ein Ansatz, solche Informationen zu berücksichtigen, könnte der Einsatz von physikalisch-geleiteten (Palmitessa et al. 2022) oder physikalisch-informierten (Mahesh et al. 2022) Neuronalen Netzen sein. Auf dem Gebiet besteht allerdings noch erheblicher Forschungsbedarf und zum aktuellen Zeitpunkt sind keine Studien bekannt, die bereits ein übertragbares Modell vorstellen. Auf weitere Untersuchungen wird daher an dieser Stelle verzichtet.

6.4.3 Analyse des Einflusses von Hyperparametern (Hyperparameter tuning)

Als abschließende Analyse wurde der Einfluss verschiedener Hyperparameter untersucht. Dazu wurde der TFT mit unterschiedlichen Parametersätzen initialisiert, trainiert und die jeweiligen Validierungsfehler wurden miteinander verglichen. Da das Training vieler Modellaufbauten sehr rechen- und damit zeitintensiv ist, wurde anstelle des gesamten Datensatzes auf die gleichmäßig ausgedünnten Modellregen für das Training zurückgegriffen. Zusätzlich wurden nur die Überstauganglinien von 34 Schächten berücksichtigt, an denen im Testdatensatz ein Grenzwert von > 100 l Überstau für den gewählten Prognosehorizont mindestens einmal

überschritten wurde. Durch diese Anpassungen konnte die Dauer des Modelltrainings auf unter eine Stunde reduziert werden. Damit war gegenüber dem Überflutungsflächenvorhersagemodell der Einsatz von Optimierungsalgorithmen möglich. Vor diesem Hintergrund wurde zur Hyperparameteroptimierung direkt auf die in PyTorch-Forecasting für TFTs implementierte Methodik zurückgegriffen, die auf dem in Bergstra et al. (2011) vorgestellten *Tree-structured Parzen Estimator* (TPE) basiert. Dabei handelt es sich um einen Algorithmus, der gegenüber der Rastersuche nicht den gesamten Parameterraum in einem vordefinierten Raster absucht, sondern neue Parameterkombinationen auf Grundlage von Erkenntnissen vorangegangener Parameterkombinationen auswählt. Dadurch können bei gleicher Anzahl an berücksichtigten Parameterkombinationen mehr Kombinationen getestet werden, die zu guten Ergebnissen führen und schlechte Kombinationen ausgeschlossen werden. Zur detaillierten Beschreibung der Funktionsweise sei an dieser Stelle auf Bergstra et al. (2011) verwiesen.

Für die Parameterstudie wurden die folgenden Hyperparameter mit den angegebenen Wertebereichen berücksichtigt:

- Einheiten pro Schicht: [16; 256]
- Einheiten pro kont. Schicht: [16; 256]
- *Attention Head Size*: [1; 4]
- *Dropout-Rate*: [0; 0,3]
- Lernrate: [0,005; 0,1]
- *Gradient Clip-Wert*: [0,01; 0,3]

Insgesamt wurden 100 Optimierungsläufe durchgeführt. Die Ergebnisse der besten fünf Läufe sind in Tabelle 6-3 aufgeführt.

Tabelle 6-3: Mittlerer Fehler und Parametrisierung der besten fünf Optimierungsläufe

| ID | Einheiten pro Schicht | Einheiten pro kont. Schicht | <i>Attention Head Size</i> | <i>Dropout-Rate</i> | Lernrate | <i>Gradient Clip-Wert</i> | MAE |
|----|-----------------------|-----------------------------|----------------------------|---------------------|----------|---------------------------|--------|
| 17 | 182 | 40 | 3 | 0,166 | 0,052 | 0,044 | 10,496 |
| 83 | 184 | 23 | 3 | 0,154 | 0,038 | 0,033 | 10,833 |
| 52 | 174 | 26 | 3 | 0,192 | 0,058 | 0,048 | 11,114 |
| 77 | 213 | 27 | 3 | 0,217 | 0,037 | 0,034 | 11,149 |
| 99 | 152 | 18 | 3 | 0,156 | 0,045 | 0,038 | 11,440 |

Mit der besten Parameterkombination wurde abschließend ein TFT am gesamten Datensatz trainiert. Anschließend wurde die Prognosegüte evaluiert und mit der ursprünglichen Variante aus Kapitel 6.3 verglichen. Das Ergebnis ist in Abbildung 6-10 dargestellt und zeigt zwar, dass die Dichte im Bereich des Optimums leicht zugenommen hat, der Median, die Form der Box und Whisker sowie die Streuung der Ergebnisse aber in etwa gleichbleiben. Dementsprechend werden entweder bereits mit der auf Erfahrungswerten basierenden Parameterkombination gute Ergebnisse erzielt oder der Einfluss der berücksichtigten Hyperparameter spielt bei der gewählten Architektur nur eine untergeordnete Rolle. In beiden Fällen ist davon auszugehen, dass auch mit weiteren Optimierungsläufen keine wesentlichen Verbesserungen mehr erzielt werden können. Daher wurden die Untersuchungen an dieser Stelle beendet.

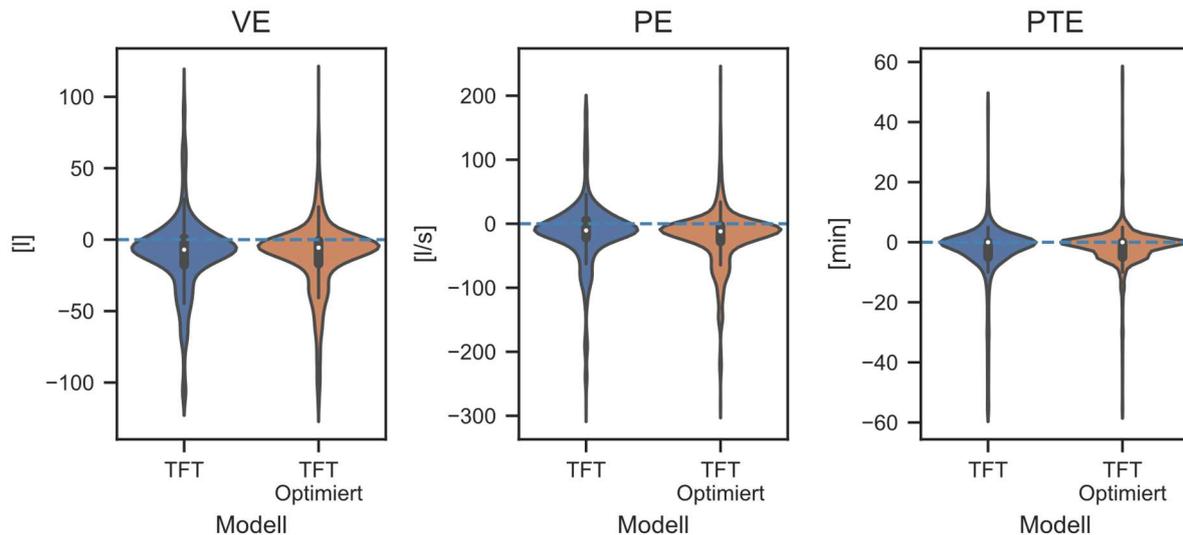


Abbildung 6-10: Violinen-Plot der Metriken für den ursprünglichen und den optimierten TFT (berücksichtigt werden nur Prognosen, bei denen entweder das kumulierte Volumen des HD-Modells oder des ML-Modells > 100 l liegt)

6.5 Beurteilung der Vorhersagegenauigkeit

Zur Beurteilung der Vorhersagegenauigkeit des finalen und optimierten TFTs wurden verschiedene Untersuchungen durchgeführt. Zunächst wurde wieder die Vorhersagegenauigkeit in Abhängigkeit der Wiederkehrzeit eines Ereignisses überprüft. Hierzu wurden für die 26 Ereignisse des Testdatensatzes die berücksichtigten Gütekriterien ermittelt. Dabei wurde der Mittelwert ereignisweise über alle Schächte und Prognosezeitpunkte gebildet, bei denen das akkumulierte Überstauvolumen für den Prognosezeitraum beim HD-Modell oder ML-Modell mehr als 100 l beträgt. Insgesamt wurde dieser Wert nur an 13 der 26 Ereignisse an mindestens einem Schacht überschritten. Die Mittelwerte der einzelnen Metriken für diese 13 Ereignisse sind in Abbildung 6-11 aufgeführt. Es zeigt sich, wie bereits beim Violinen-Plot ersichtlich, dass das Modell tendenziell eher unterschätzt. So liegen bei allen drei berücksichtigten Metriken mehr Mittelwerte der Ereignisse unter dem Optimum als darüber. Zudem zeigt sich, dass, anders als beim Überflutungsflächenvorhersagemodell, bei extremen Ereignissen die Abweichung teilweise etwas zunimmt. Bei der Einordnung der Ergebnisse ist allerdings auch zu beachten, dass es sich bei allen Metriken um die absoluten Abweichungen handelt. Die Darstellung der prozentualen Fehler wurde ebenfalls geprüft, zeigte aber insbesondere bei den häufigeren Ereignissen ($T > 20$ a) meist extreme Werte. Verursacht werden diese Werte, wie in Kapitel 3.5 beschrieben, überwiegend durch Prognosezeitpunkte, an denen nur geringer Überstau auftritt, das Modell aber keinen Überstau prognostiziert oder andersherum. Dementsprechend ist eine sinnvolle Bewertung nicht möglich, weshalb auf eine Darstellung der Ergebnisse verzichtet wurde.

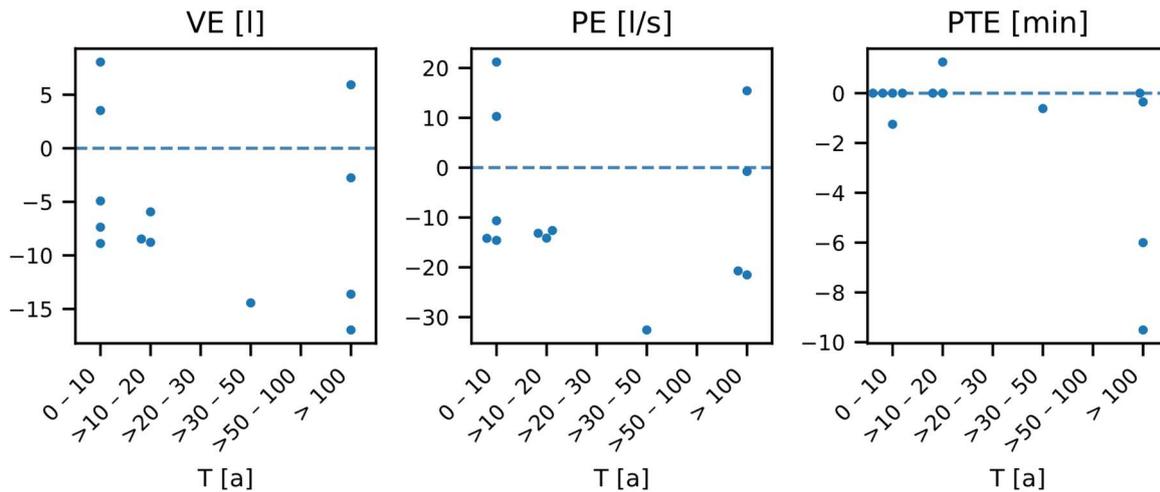


Abbildung 6-11: Volumenfehler (VE), Maximalwertabweichung (PE) und zeitliche Maximalwertabweichung (PTE) in Abhängigkeit der Wiederkehrzeit der Ereignisse (die blaue, gestrichelte Linie markiert jeweils das Optimum)

Zur detaillierten Betrachtung der Modellgüte in Abhängigkeit des prognostizierten Überstauvolumens wurden die in Abbildung 6-12 dargestellten Scatterplots erstellt. In den Abbildungen wurde zum einen das prognostizierte dem simulierten Überstauvolumen gegenübergestellt. Zum anderen ist der Prognosefehler in Abhängigkeit der simulierten Überstauvolumina dargestellt. Wie bei der Berechnung der Metriken wurde das Volumen wieder pro Vorhersagezeitpunkt für den Prognosehorizont von 60 Minuten berechnet. Anhand der Abbildung ist abermals zu erkennen, dass das Modell tendenziell eher unterschätzt. Zudem zeigt sich, dass die Streuung der Prognosefehler in etwa gleichmäßig über alle Volumina verläuft. Damit zeigt sich auch hier, dass die Modellgüte bei den besonders relevanten Extremereignissen nicht abnimmt.

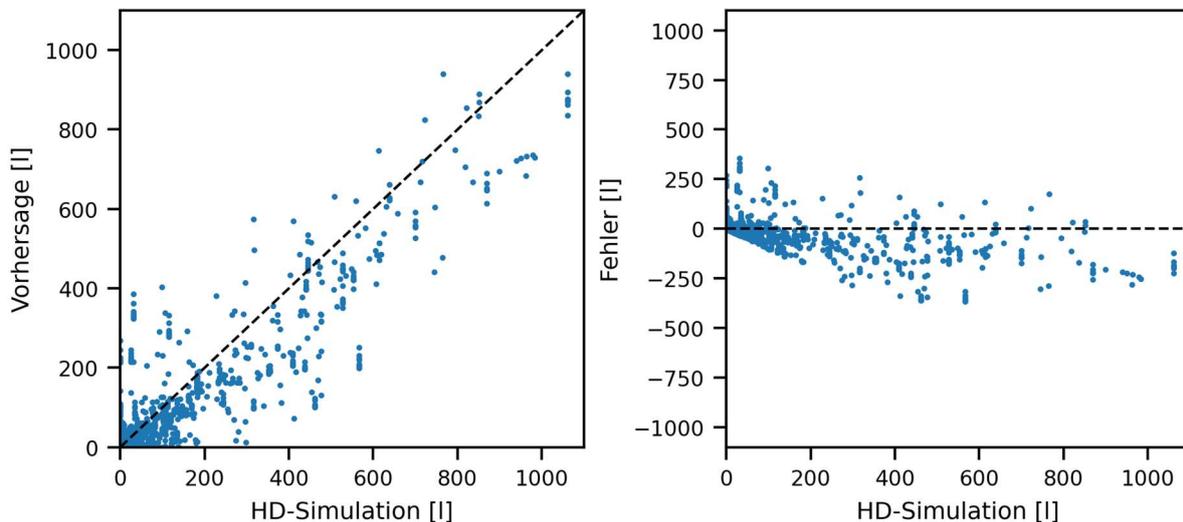


Abbildung 6-12: Scatterplot mit dem Vergleich des vorhergesagten und simulierten Überstauvolumens je Vorhersage (links) und dem Vorhersagefehler in Abhängigkeit des simulierten Überstauvolumens (rechts)

Zur abschließenden Bewertung wurde das TFT-Modell anhand historischer Starkregenereignisse getestet. Abbildung 6-13 zeigt Überstauganglinien der Ereignisse vom 03.07.2009 und 03.07.2010, die maximale Wiederkehrzeiten T_{max} von > 200 a und > 1000 a Jahren aufweisen.

Dargestellt sind die Niederschlagsprognose sowie die dafür prognostizierten Überstauganglinien an sechs Schächten im Einzugsgebiet. Die Ergebnisse zu weiteren Ereignissen und Schächten sind in Anhang A-5 dargestellt.

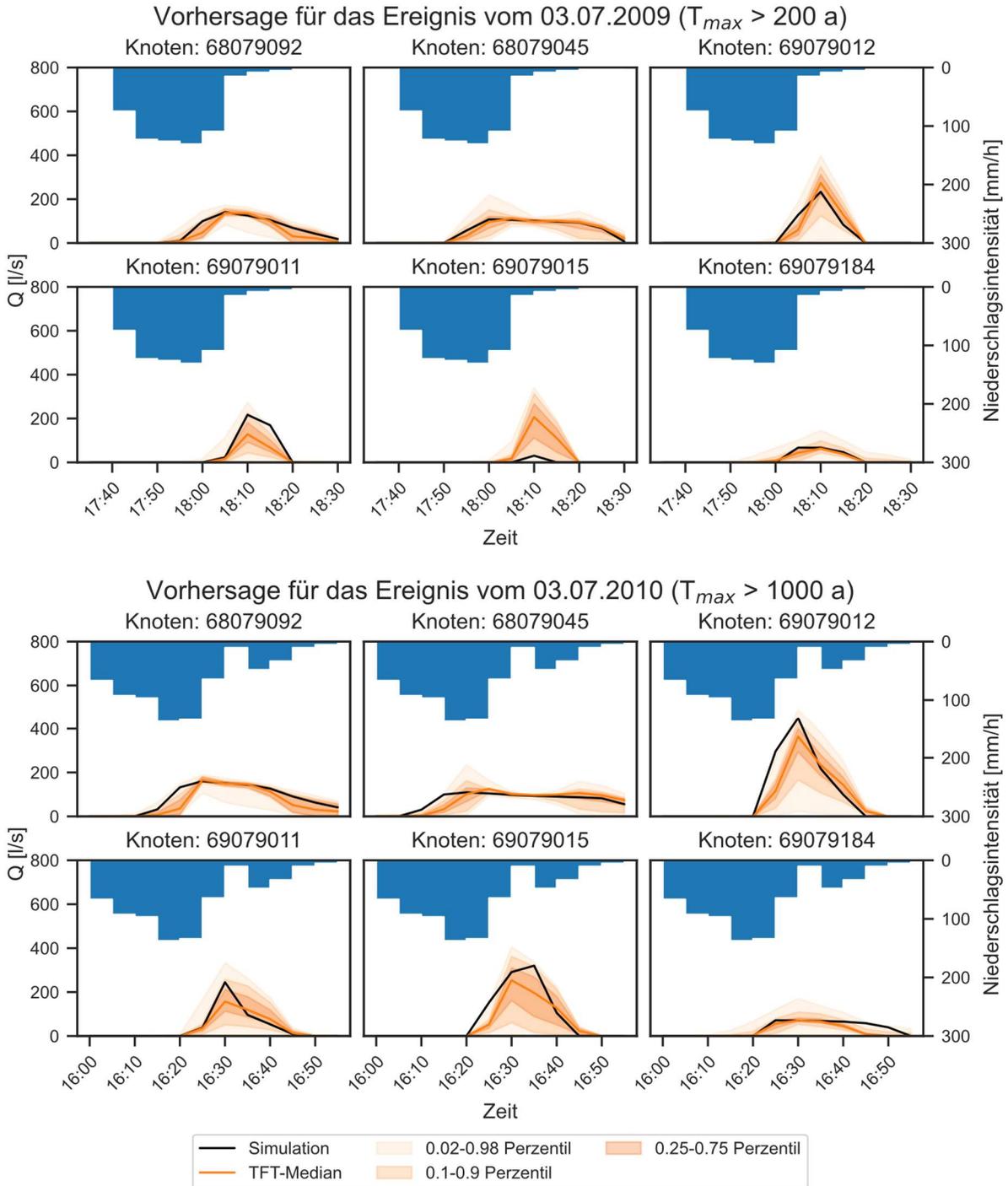


Abbildung 6-13: Vorhersageergebnis an sechs Schächten für die historischen Ereignisse vom 03. Juli 2009 und 03. Juli 2010 in Gelsenkirchen

Anhand der Abbildungen lassen sich die folgenden Erkenntnisse ableiten:

1. Das 0,5-Quantil stimmt teilweise sehr gut mit der simulierten Zielgröße überein, zum Teil ergeben sich aber auch deutliche Abweichungen. Zudem zeigt sich, dass bei größeren Abweichungen auch der Unsicherheitsbereich zwischen den 0,02- und 0,98-Quantilen zunimmt.

2. Längere Überstauereignisse können mit einer hohen Genauigkeit prognostiziert werden, während sich bei kurzen Peaks teils extreme Abweichungen von $> 100\%$ beim Maximalwert, aber auch dem resultierenden Überstauvolumen ergeben.
3. Auch wenn das Modell insgesamt die Überstaumengen eher unterschätzt, ist eine Tendenz zur konsequenten Unter- oder Überschätzung an den aufgeführten Schächten nicht zu erkennen. Diese Erkenntnis spiegelt sich auch an den übrigen Schächten im Einzugsgebiet wider.

6.6 Zusammenfassung und Bewertung der Ergebnisse

Im Rahmen der Untersuchungen wurde die Leistungsfähigkeit verschiedener Künstlich Neuronaler Netzarchitekturen im Bereich der Überstauprognose evaluiert. Die Ergebnisse zeigen, dass die Ergebnislänge mit zunehmender Modellkomplexität ebenfalls zunimmt. Die besten Ergebnisse konnten mit dem komplexesten Modell, einem TFT erzielt werden. Insbesondere die Streuung der Ergebnisse fällt deutlich geringer aus und im Vergleich zu den Benchmark-Modellen kommt es zu keinen extremen Überschätzungen. Die Prognosen des finalen Modells weisen teilweise eine sehr gute Übereinstimmung mit den Simulationsergebnissen auf. Gerade bei kurzen Überstauereignissen kommt es aber teils auch zu großen Abweichungen. Dies hängt aber unter anderem auch damit zusammen, dass die Vorhersage von extremen Werten, die über kurze Zeitspannen auftreten, generell eine anspruchsvolle Aufgabe darstellt. Die gleiche Problematik konnte auch bei ähnlichen Untersuchungen (Kilsdonk et al. 2022; Palmitessa et al. 2022) festgestellt werden. Als ein möglicher Lösungsansatz wird in Palmitessa et al. (2022) der Überstau nicht direkt berechnet. Vielmehr werden zunächst Zu- und Abflüsse eines Knotens berechnet, mit denen dann über eine als zusätzliche Schicht eingefügte Massenbilanz der Überstau berechnet wird.

Ansonsten bleibt bei den Ergebnissen zu beachten, dass in den Untersuchungen sowohl für das Training als auch die Auswertung der Modelle immer perfekte Niederschlagsmessungen und -vorhersagen sowie Messungen im Kanal berücksichtigt wurden. In der Realität sind diese Daten mit Unsicherheiten behaftet, die sich dann zusätzlich auf das hier entwickelte Modell auswirken. In Hinblick auf den operationellen Betrieb sind dementsprechend weitere Untersuchungen mit realen Messungen und Prognosen erforderlich, um die Auswirkungen der Unsicherheiten auf die Prognoseergebnisse zu evaluieren. Gleichzeitig soll die Überstauprognose als Eingabe in das Überflutungsflächenvorhersagemodell dienen. Um den potentiellen Mehrwert des Modells aufzuzeigen, folgt eine gemeinsame Evaluierung in Kapitel 7.1.

Der Anlass für die Entwicklung des hier beschriebenen Überstauvorhersagemodells war die Einbindung in ein Frühwarnsystem. Aufgrund der Problematik, dass die Unsicherheiten bei Niederschlagsprognosen für Starkregenereignisse bereits nach ein bis zwei Stunden extrem zunehmen (Sun et al. 2014; Koltermann da Silva et al. 2023b), wurde der berücksichtigte Vorhersagehorizont auf eine Stunde festgelegt. Um diesen Vorhersagehorizont bestmöglich auszunutzen, sind kurze Rechenzeiten eine wesentliche Anforderung an das Überstauvorhersagemodell. Alle in den Untersuchungen berücksichtigten Modelle erfüllen diese Anforderung und liefern Prognoseergebnisse innerhalb weniger Sekunden. Auch bei einer Ausweitung auf

größere Gebiete ist davon auszugehen, dass die Rechenzeit selbst bei geringer Rechenleistung unter einer Minute liegt. Andernfalls besteht die Möglichkeit, anhand eines Extremereignisses zunächst alle potentiell überflutungsgefährdeten Schächte herauszufiltern und nur diese beim Modellaufbau zu berücksichtigen, um überflüssigen Rechenbedarf zu vermeiden. So hat sich beispielsweise bei den im Rahmen dieser Arbeit berücksichtigten Testereignisse gezeigt, dass lediglich 45 der 977 Schächte überhaupt zu Überstau führen.

Etwas überraschend war die schlechte Leistung der Modelle mit einem Messnetz von 20 Stationen. Es ist wahrscheinlich, dass aufgrund der hohen Korrelation der Messungen untereinander beim gewählten Systemaufbau kein Mehrwert durch ein Messnetz geschaffen werden kann. Vielmehr konnte das beste Ergebnis mit einem Modell erzielt werden, das auf eine zentrale Messung am Gebietsauslass zurückgreift. Für den Einsatz in der Praxis bietet dies den Vorteil, dass je nach Untersuchungsgebiet keine zusätzlich installierte Sensorik benötigt wird, sondern auf vorhandene Messstellen an zentralen Punkten der Kanalisation wie Regenrückhaltebecken oder Pumpwerken zurückgegriffen werden kann. Diese Messstellen müssen beispielsweise in Nordrhein-Westfalen gemäß Verordnung zur Selbstüberwachung von Abwasseranlagen (SüwVO Abw) (MKULNV 2013) ohnehin betrieben werden und verursachen so keine zusätzlichen Kosten.

Bei der Bewertung der Ergebnisse ist jedoch zu berücksichtigen, dass der fehlende Mehrwert eines Messnetzes nur dann nachgewiesen werden konnte, wenn perfekte Niederschlagsmessungen und -prognosen und keine Einschränkungen im Kanalnetz, wie beispielsweise verstopfte Einläufe, vorlagen. In der Realität liegt dieser Fall aber selten vor und dementsprechend ist damit zu rechnen, dass sich durch die Berücksichtigung vieler Sensoren und der damit einhergehenden Schwarmintelligenz Vorteile ergeben können. Eine weitere Möglichkeit, einen Mehrwert durch das Messnetz zu generieren, ist die Einbindung der Messungen zur Korrektur der Modellergebnisse, wie dies in Zhu et al. (2023) durchgeführt wurde. Möglicherweise bietet auch die Verarbeitung der Sensormessungen als räumlich strukturierte Graph Sequenz anstelle von ungeordneten Zeitreihen weiteres Optimierungspotential. Die Kombination von *Transformern* mit *Graph Neural Networks* hat sich beispielsweise im Bereich der Verkehrsprognose als leistungsstark erwiesen (Yu et al. 2020; Xu et al. 2020).

Da der zum Training verwendete Datensatz mit dem gekoppelten hydrodynamischen 1D/2D-Berechnungsmodell erfolgte, besteht hinsichtlich der langen Rechendauern bei der Datengenerierung die gleiche Problematik wie bei dem Modell zur Vorhersage von Überflutungsflächen. Daher wurde auch für das Überstauvorhersagemodell geprüft, ob und wie weit sich die berücksichtigte Datenmenge reduzieren lässt. Die Ergebnisse zeigen, analog zu den Ergebnissen beim Überflutungsflächenvorhersagemodell, dass je nach Auswahlmethodik der Datensatz bei moderatem Genauigkeitsverlust um bis zu 80 % reduziert werden kann. Ebenfalls wurde die hohe Relevanz einer möglichst repräsentativen Abdeckung sämtlicher Ereignisverläufe und -intensitäten aufgezeigt. Die Berücksichtigung der natürlichen Regenreihen spielt dagegen auch hier eine eher untergeordnete Rolle, weshalb bei nachfolgenden Untersuchungen der Fokus eher auf einen noch repräsentativeren Datensatz aus Modellregen gelegt werden sollte.

Eine wesentliche Limitierung des vorgestellten Ansatzes ist die fehlende Generalisierbarkeit.

Durch die fehlende Berücksichtigung von Systeminformationen über das Kanalnetz lässt sich das trainierte Modell nicht ohne Weiteres auf andere Gebiete übertragen und es ist ein erneutes Modelltraining erforderlich, sobald sich Änderungen am Netz ergeben. Dadurch ergibt sich noch erheblicher Forschungsbedarf, um das entwickelte Modell praxistauglich zu gestalten. Ein Ansatz in die Richtung könnte der Einsatz von physikalisch-geleiteten (Palmitessa et al. 2022) oder physikalisch-informierten (Mahesh et al. 2022) Neuronalen Netzen sein. Studien, die bereits ein übertragbares ML-basiertes Modell zur Prognose von Überstau vorstellen, sind jedoch zum aktuellen Zeitpunkt nicht bekannt.

7 Zusammenführung des Gesamtsystems

7.1 Gekoppelte Betrachtung Überstau- und Überflutungsflächenvorhersage

Die Untersuchungen in Kapitel 5 haben gezeigt, dass Maschinelle Lernverfahren generell zur Prognose von Überflutungsflächen geeignet sind und je nach Verfahren sogar hohe Genauigkeiten erzielt werden können. Bei den Untersuchungen wurden allerdings keine Unsicherheiten bei den Eingabedaten berücksichtigt, die sich im späteren Betrieb durch die verwendeten Messnetze und Vorhersagemodelle ergeben. Im vorliegenden Kapitel wird daher untersucht, welchen Einfluss die Unsicherheiten der Überstauvorhersage auf die Prognosegüte des Überflutungsflächenvorhersagemodells haben. Dazu wurde das finale T-GCN herangezogen und einerseits wie in Kapitel 5 mit simulierten Überstauganglinien und andererseits mit prognostizierten Überstauganglinien belastet. Zur Prognose der Überstauganglinien wurde wiederum der finale TFT aus Kapitel 6 herangezogen. Die Untersuchungen wurden an dem Testdatensatz aus dem vorherigen Kapitel durchgeführt.

Im Rahmen einer ersten Analyse wurden zunächst wieder ereignisweise die Gütekriterien RMSE und CSI für unterschiedliche Grenzwerte ermittelt. In Abbildung 7-1 sind die Ergebnisse der mit dem Überstauvorhersagemodell gekoppelten Variante und der Variante mit durch das HD-Modell simulierten Überstauganglinien gegenübergestellt. Es zeigt sich, dass die Unsicherheiten bei der Überstauvorhersage sich erwartungsgemäß auch auf die Überflutungsflächenprognose auswirken. Mit steigenden Grenzwerten nimmt dabei auch der Fehler zu, was sowohl für den RMSE als auch den CSI gilt. Dieser Trend war zu erwarten, da erst extreme Ereignisse, die zu hohen Wasserständen führen, auch Überstausituationen im Kanalnetz hervorrufen. Daher sind diese Ereignisse besonders von den Unsicherheiten in der Überstauvorhersage betroffen. Wie in Kapitel 5 bereits erwähnt, sind die Ergebnisse für einen Grenzwert $> 0,5$ m nur begrenzt belastbar, da lediglich vier der 26 Ereignisse des Testdatensatzes diesen Grenzwert überschreiten.

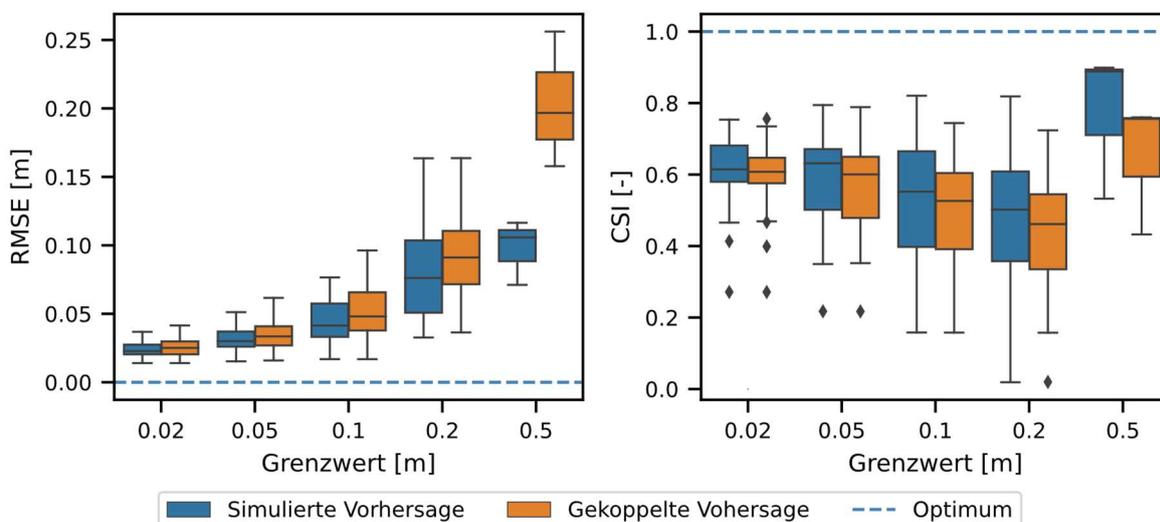


Abbildung 7-1: Verteilung der Metriken über alle 26 Ereignisse im Testdatensatz für die Variante

mit durch das HD-Modell simulierten Überstauganglinien (Simulierte Vorhersage) und die mit dem Überstauvorhersagemodell gekoppelte Variante (Gekoppelte Vorhersage)

Im weiteren Verlauf wurde der pixelweise Fehler bei der gekoppelten Vorhersage überprüft und dem Fehler der simulierten Vorhersage gegenübergestellt (vgl. Abbildung 7-2). Ähnlich dem Modell mit den simulierten Überstauganglinien schwanken die Werte relativ gleichmäßig um das Optimum (vgl. Abbildung 5-21) und neigen eher zum Unterschätzen der Wasserstände. Auch an dieser Stelle sei darauf hingewiesen, dass zur Darstellung der Pixelanzahl eine logarithmische Skala verwendet wurde. Abweichungen von mehr als einem Meter treten entsprechend nur sehr selten auf und es zeigt sich auch beim gekoppelten Modell, dass der Prognosefehler bei höheren Wasserständen abnimmt. Damit werden diese, für den späteren Anwendungsfall besonders relevanten Bereiche, mit einer hohen Genauigkeit vorhergesagt.

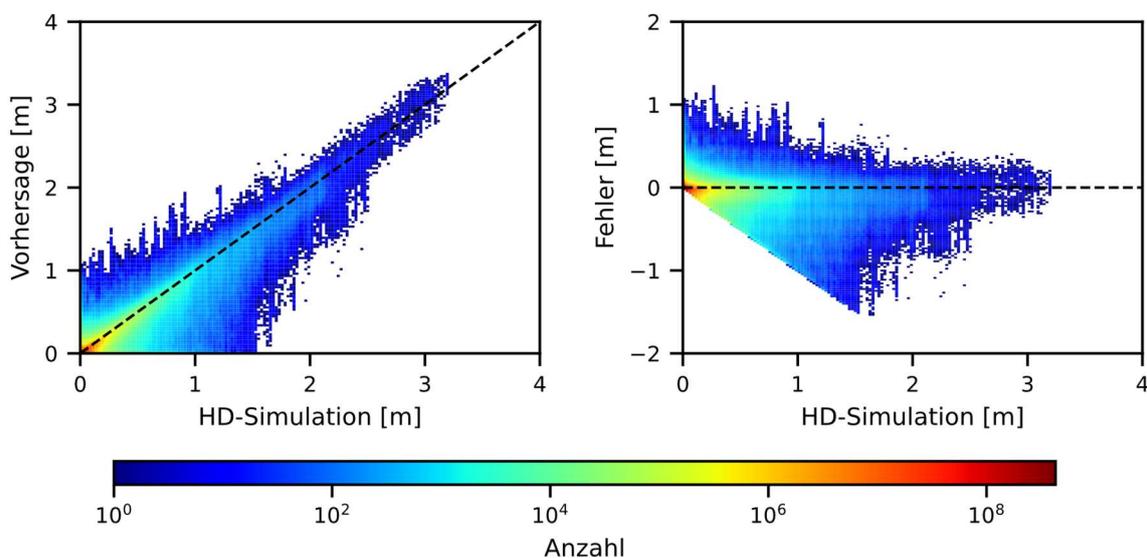


Abbildung 7-2: 2D-Histogramm mit dem pixelweisen Vergleich der vorhergesagten und simulierten Wasserstände (links) und dem pixelweisen Vorhersagefehler in Abhängigkeit der simulierten Wasserstände (rechts)

Zur abschließenden Beurteilung der Prognosegüte wurde wieder das Ereignis vom 3. Juli 2009 herangezogen ($T > 200$ a). Abbildung 7-3 und Abbildung 7-4 zeigen die Niederschlagsganglinie des Ereignisses sowie die Prognoseergebnisse für drei Zeitpunkte einer einzelnen Prognose. Gegenüber der Prognose mit simulierten Überstauganglinien (vgl. Abbildung 5-22 und Abbildung 5-23) zeigt sich eine etwas stärkere Abweichung der Überflutungssituation, was unter anderem an den schlechteren Werten beim CSI und RMSE deutlich wird. Während bei den simulierten Überstauganglinien visuell kaum ein Unterschied zwischen simulierten und prognostizierten Überflutungsflächen zu erkennen war, sind bei der gekoppelten Vorhersage insbesondere in der herangezoomten Abbildung 7-4 Unterschiede zu erkennen. Gleichwohl zeigt insbesondere die Detailansicht der Unterführung, dass die Wasserstände an den relevanten Punkten nahezu identisch sind. Im vorliegenden Beispiel würden die Unsicherheiten dementsprechend nicht zu abweichenden Entscheidungen durch die Akteure des Krisenmanagements führen. Die Ergebnisse zu weiteren Ereignissen zeigen in den meisten Fällen ein ähnliches Bild und sind im Anhang A-6 dargestellt. An dieser Stelle kann somit festgehalten werden, dass die Unsicherheiten der Überstauvorhersage sich im berücksichtigten Untersuchungsgebiet zwar auf die Vorhersage der Überflutungsflächen auswirken, der Umfang in den

meisten Fällen aber in einer vertretbaren Größenordnung liegt.

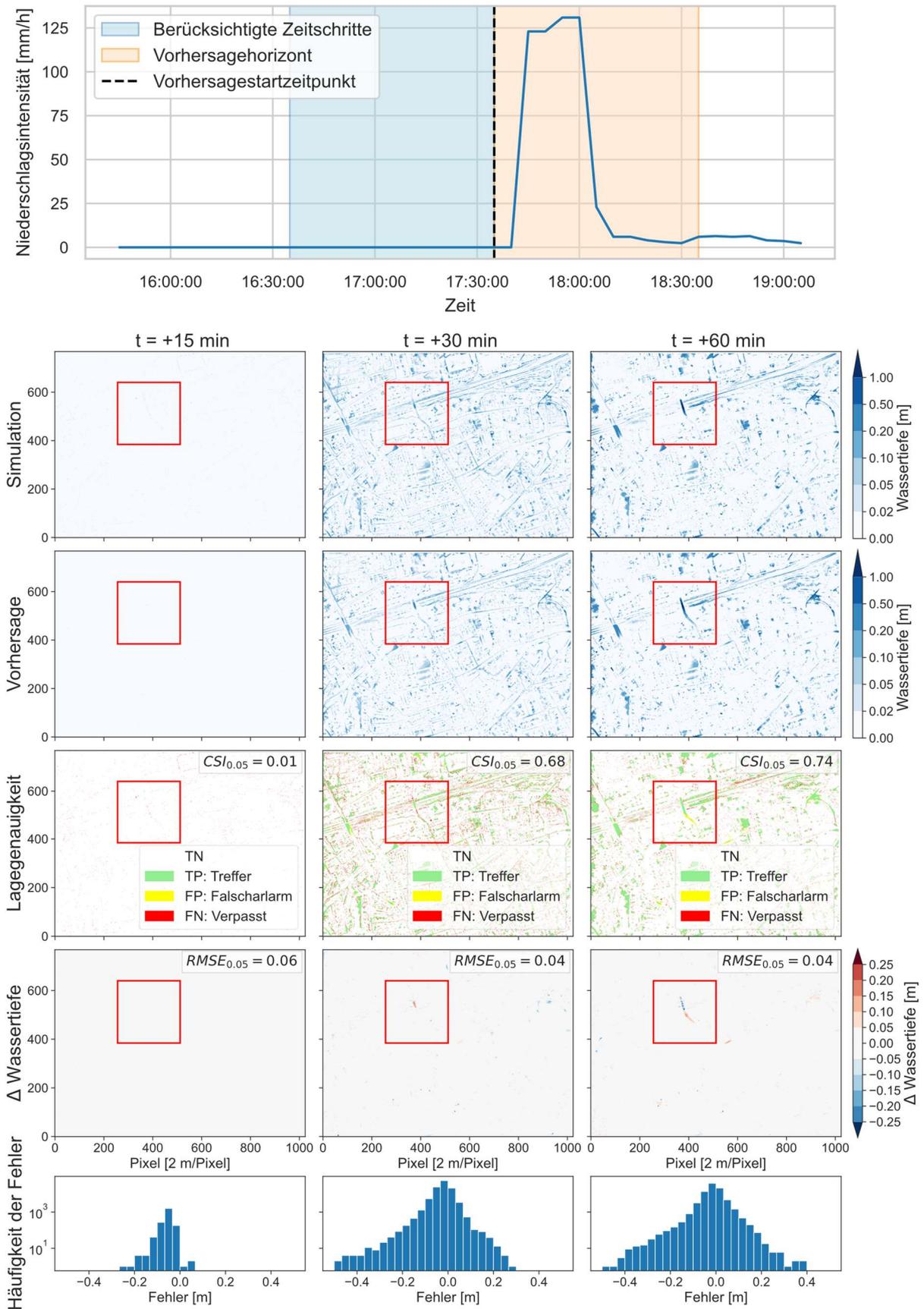


Abbildung 7-3: Niederschlagsganglinie mit markiertem Vorhersagezeitpunkt zu Beginn des Ereignisses vom 3. Juli 2009 in Gelsenkirchen (oben) und Vorhersageergebnis

und -evaluierung für drei Zeitschritte der Vorhersage (unten). Der nachfolgend in Abbildung 7-4 dargestellte Detailausschnitt ist rot markiert

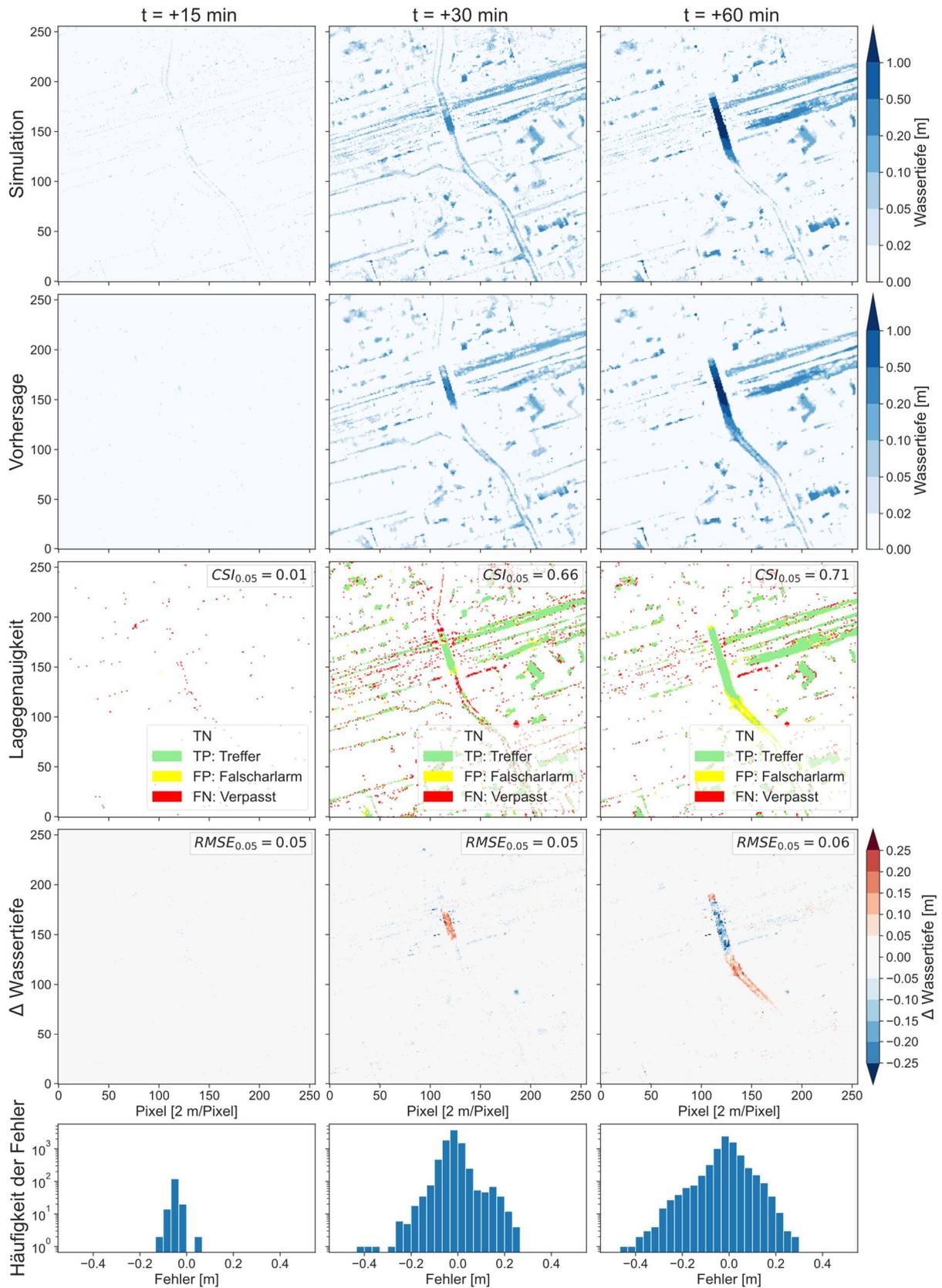


Abbildung 7-4: Detailansicht des Vorhersageergebnisses und der zugehörigen Evaluierung für einen Ausschnitt mit einer überfluteten Unterführung

7.2 Integration der Vorhersagemodelle in einen Prognose- und Warnprozess

Für die Anwendung der Vorhersagemodelle im Krisenmanagement ist die Einbindung in einen Prognose- und Warnprozess erforderlich. Im Forschungsvorhaben KIWaSuS, in dem auch die in dieser Arbeit vorgestellten Modelle entwickelt wurden, ist ein Demonstrator für diesen Zweck entwickelt worden. Abbildung 7-5 bildet die wichtigsten am Prognoseprozess beteiligten Komponenten ab. Diese lassen sich in die übergeordneten Bereiche der Datenerfassung, Prognoseerzeugung und Bereitstellung unterteilen. Unter den Bereich der Datenerfassung fallen drei Quellen. Zum einen sind das die Radarmessungen des Deutschen Wetterdienstes, die als flächige Niederschlagsmessung kostenlos über ein Open-Data-Portal (DWD o. J.) zur Verfügung gestellt werden. Zum anderen wurden im Untersuchungsgebiet Niederschlagssensoren sowie Sensoren im Kanal installiert. Letztere erfassen ein akustisches Signal, das anschließend in Füllstandsklassen umgerechnet wird. Die so erfassten Daten dienen als Eingabe für die verschiedenen Prognosemodelle. Bei den Radarmessungen erfolgt zunächst eine Korrektur, bevor diese anschließend an ein KI-basiertes Niederschlagsvorhersagemodell weitergeleitet werden (vgl. Koltermann da Silva et al. 2023a). Die resultierende Niederschlagsprognose dient dann als Eingabe für die Prognosemodelle für Überstau und Überflutungsflächen. Das Überstauvorhersagemodell erhält als zusätzliche Eingaben die terrestrischen Messungen des gefallenen Niederschlags sowie den Füllstand im Kanalnetz. Die erzeugte Überstauprognose wird wiederum als zusätzliche Eingabe an das Überflutungsflächenvorhersagemodell weitergeleitet.

Sämtliche Prognosen können durch den Anwender entweder über eine Visualisierungsoberfläche direkt eingesehen oder über einen *Web Map Service* (WMS) als Schnittstelle verwendet werden. Letzterer eignet sich insbesondere zur Einbindung in vorhandene Geoportale, in denen die Prognosen dann mit dort hinterlegten Informationen, wie Elementen der kritischen Infrastruktur überlagert werden können. Die Prognoseergebnisse können zusätzlich mit einem Warnprozess gekoppelt werden, der beim Überschreiten kritischer Grenzwerte Betroffene wie insbesondere die Akteure des kommunalen Krisenmanagements alarmiert. Als Grenzwerte können dabei beispielsweise Niederschlagshöhen, Wiederkehrzeiten oder kritische Wasserstände an der Geländeoberfläche festgelegt werden. Auch ein abgestufter Warnprozess wie im Bereich der Hochwasserwarnung über das länderübergreifende Hochwasserportal (LHP) (LfU und LUBW o. J.) ist möglich. So kann etwa die Feuerwehr ihre Einsatzkräfte beim Überschreiten einer ersten Warnstufe zunächst in Bereitschaft versetzen, bevor beim Überschreiten einer kritischen Warnstufe dann konkrete Maßnahmen ergriffen werden. Aufgrund des Pilotcharakters im Projekt KIWaSuS wurden für den Warnprozess zunächst nur die Akteure des kommunalen Krisenmanagements berücksichtigt. Nach einer erfolgreichen Testphase ist auch die Kopplung mit Warn-Apps wie NINA oder KATWARN denkbar, um auch die Bevölkerung im Krisenfall zu informieren.

Im Rahmen des Prognoseprozesses fallen große Datenmengen an, die in Echtzeit verarbeitet werden müssen. Klassische Datenbanken allein reichen für diesen Anwendungsfall nicht aus, vielmehr ist eine Datenplattform erforderlich, in der sämtliche Datenströme effizient zusammengeführt, gespeichert und bereitgestellt sowie die Prognosemodelle betrieben werden kön-

nen. Wichtig ist dabei eine hohe Zuverlässigkeit und Skalierbarkeit der Plattform, um im Ereignisfall einen ordnungsgemäßen Betrieb gewährleisten zu können. Welche weiteren Anforderungen an eine solche Plattform gestellt werden sowie die Struktur der im Projekt KIWaSuS entwickelten Plattform ist in Fretrup et al. (2022) beschrieben.

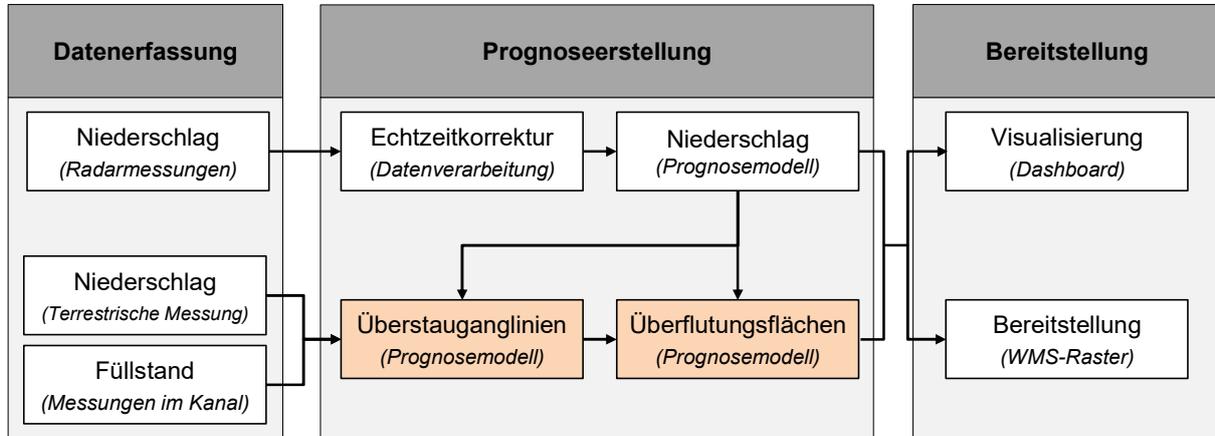


Abbildung 7-5: Exemplarische Darstellung der einzelnen am Prognoseprozess beteiligten Komponenten (verändert nach Fretrup et al., 2024)

8 Zusammenfassung und weiterer Forschungsbedarf

Im Rahmen der vorliegenden Arbeit wurde ein Vorhersagemodell für pluviale Überflutungen in urbanen Gebieten entwickelt. Ziel war es dabei eine ereignisspezifische und dynamische Prognose der bevorstehenden Überflutungssituation in Form einer Sequenz aus Überflutungsras-tern (dynamische Starkregengefahrenkarte) zu ermöglichen. Während die in der Praxis weitverbreiteten hydrodynamischen Berechnungsmodelle für die Generierung von statischen Starkregengefahrenkarten gut geeignet sind, verhindern die hohen Rechenzeiten aber den Einsatz als Vorhersagemodell in Echtzeitwarnsystemen. Um diese Problematik zu lösen, wurden Maschinelle Lernverfahren als Ersatzmodell eingesetzt, die anhand von Berechnungsergebnissen hydrodynamischer Modelle trainiert wurden. Ziel ist es, so Ergebnisse mit möglichst ebenwertiger Genauigkeit in deutlich kürzerer Zeit zur Verfügung zu stellen und so den Einsatz in Warnsystemen zu ermöglichen. Die Arbeit stellt alle für dieses Vorhaben relevanten Grundlagen, die entwickelten Modellaufbauten und zugehörige Evaluierungsergebnisse dar. Nachfolgend werden zunächst die wesentlichen Erkenntnisse der Arbeit zusammengefasst, bevor anschließend die Limitierungen des entwickelten Ansatzes sowie der weiterhin bestehende Forschungsbedarf aufgezeigt werden.

8.1 Zusammenfassung der wesentlichen Erkenntnisse

Das übergeordnete Ziel der Arbeit bestand in der Entwicklung eines echtzeitfähigen Vorhersagemodells für pluviale Überflutungen. Um den späteren Einsatz des entwickelten Modells in der Praxis zu ermöglichen, sind verschiedene Voraussetzungen zu erfüllen. Dazu zählen insbesondere (i) kurze Rechenzeiten bei der Prognoseerstellung, um den Echtzeiteinsatz zu ermöglichen. Zudem sind (ii) möglichst hohe Genauigkeiten erforderlich, um effizient handeln zu können und Fehlalarme zu vermeiden. Weiterhin müssen (iii) alle für die Modellentwicklung und den späteren Betrieb erforderlichen Daten verfügbar sein. Im Folgenden werden die gesammelten Erkenntnisse für jeden der aufgeführten Punkte zusammengefasst. Dabei werden sowohl die im Rahmen der Entwicklung des Vorhersagemodells für Überflutungsflächen als auch für Überstau gesammelten Erkenntnisse gemeinsam aufgeführt.

Rechengeschwindigkeit

Aufgrund der im Rahmen der Arbeit mehrfach erwähnten Problematik, dass die Unsicherheiten bei Niederschlagsprognosen für Starkregenereignisse bereits nach ein bis zwei Stunden extrem zunehmen, wurde der berücksichtigte Vorhersagehorizont auf 60 Minuten festgelegt. Die Untersuchungen haben gezeigt, dass bei dem gewählten Vorhersagehorizont sowohl das Überstau- als auch das Überflutungsflächenvorhersagemodell Rechendauern < 10 Sekunden aufweisen. Damit steht beinahe der gesamte Vorhersagehorizont einer Niederschlagsprognose auch für die Überflutungsprognose zur Verfügung. Gegenüber dem für die Datengenerierung verwendeten physikalisch-basierten Modell, das für die Berechnung eines 60-minütigen Ereignisses beim gleichen Untersuchungsgebiet und auf derselben Hardware etwa zwei Stunden benötigt, konnte die Rechenzeit massiv reduziert werden. Auch bei einer Ausweitung auf größere Gebiete ist davon auszugehen, dass die Rechenzeit im Bereich von Sekunden

oder wenigen Minuten liegt, womit das Modell für den Echtzeitbetrieb geeignet ist.

Genauigkeit

Bei der Genauigkeit erfolgte zunächst die Beurteilung der Vorhersagemodelle für Überstau und Überflutungsflächen getrennt voneinander, bevor dann abschließend eine gekoppelte Evaluierung erfolgte. Die getrennte Betrachtung zeigte bei der Vorhersage von Überflutungsflächen eine hohe Güte der Ergebnisse. Sowohl die Lage von Überflutungsflächen als auch die enthaltenen Wasserstände stimmten größtenteils mit denen der physikalisch-basierten HD-Simulation überein. Auch bei der Vorhersage der Überstauganglinien konnte teilweise eine hohe Übereinstimmung mit den Ergebnissen der HD-Simulation erzielt werden. Dies war insbesondere bei gleichmäßigen und lang anhaltenden Überstauereignissen der Fall. Gerade bei kurzen Überstauereignissen kommt es aber teils auch zu großen Abweichungen. Dies hängt unter anderem aber auch damit zusammen, dass die Vorhersage von extremen Werten, die über kurze Zeitspannen auftreten, generell eine anspruchsvolle Aufgabe darstellt. Im Rahmen der gekoppelten Betrachtung konnte aber gezeigt werden, dass sich die Unsicherheiten der Überstauvorhersage zwar auf die Ergebnisse der Prognose von Überflutungsflächen auswirken, in den meisten Fällen aber nicht in einem für das Krisenmanagement entscheidungsrelevanten Rahmen liegen.

Datenbedarf und -verfügbarkeit

Für die Entwicklung der in dieser Arbeit beschriebenen ML-basierten Vorhersagemodelle ist zunächst ein Trainingsdatensatz erforderlich. Da derzeit weder für die flächendeckende Erfassung von Überflutungsflächen noch der Überstauganglinien an den Schächten eines Untersuchungsgebietes geeignete Sensorik oder Messnetze zur Verfügung steht, lagen auch keine Datensätze für das Modelltraining vor. Um diese Problematik zu umgehen, wurde mit einem physikalisch-basierten Modell ein künstlicher Datensatz generiert. Bei dem verwendeten Modell handelt es sich um ein gekoppeltes hydrodynamisches 1D/2D-Berechnungsmodell, das aus einem Kanalnetz- und einem Oberflächenmodell besteht. Die Erzeugung des Oberflächenmodells basiert dabei im Wesentlichen auf einem Geländemodell und Flächennutzungsdaten. Im Rahmen der vorliegenden Arbeit wurde das Geländemodell aus den für NRW flächendeckend und kostenlos vorliegenden 3D-Messdaten erstellt. Die verwendeten Flächennutzungsdaten wurden von Emschergenossenschaft und Lippeverband bezogen. Ähnliche Daten werden in NRW in einem etwas geringeren Detaillierungsgrad über das Amtliche Liegenschaftskatasterinformationssystem kostenlos und flächendeckend zur Verfügung gestellt. Kanalnetzmodelle und die zugrundeliegenden Daten werden hingegen in der Regel durch Kommunen im Rahmen der Generalentwässerungsplanung erstellt und fortgeschrieben. Da die Daten nicht frei zur Verfügung gestellt werden, sind diese dort zu erfragen.

Als Modellbelastung wurden Niederschlagsdaten der Emschergenossenschaft verwendet, die nicht frei verfügbar sind. Ähnliche Datensätze werden aber kostenlos über das Open-Data-Portal des Deutschen Wetterdienstes (DWD o. J.) bereitgestellt und stehen somit für alle Regionen Deutschlands zur Verfügung. Zusätzlich haben im Rahmen der Arbeit durchgeführte Sensitivitätsanalysen gezeigt, dass auch mit der alleinigen Berücksichtigung von Modellregensreihen gute Ergebnisse erzielt werden können. Diese Modellregen können entweder auf

Grundlage von eigenen statistischen Auswertungen oder anhand des KOSTRA-DWD-2020 des Deutschen Wetterdienstes selber erstellt werden. Abschließend kann somit festgehalten werden, dass zwar aktuell noch keine geeigneten Messdaten zum Überstau- und Überflutungsverhalten für das Modelltraining vorliegen, jedoch alle sonstigen für die Erzeugung eines künstlichen Trainingsdatensatzes erforderlichen Daten entweder frei zur Verfügung stehen oder in den meisten Fällen bei Kommunen vorhanden sind.

8.2 Limitierungen und weiterer Forschungsbedarf

Mit den im Rahmen der Arbeit entwickelten Modellen konnte das Ziel erreicht werden, die bevorstehende Überflutungssituation zeitlich und räumlich hochauflösend vorherzusagen. Im Rahmen des Entwicklungsprozesses sowie der durchgeführten Sensitivitätsanalysen konnten aber auch Einschränkungen und Schwachstellen der Modelle identifiziert werden. Nachfolgend werden daher wesentliche Limitierungen offengelegt und der sich daraus ergebene Forschungsbedarf aufgezeigt.

Skalierbarkeit und Übertragbarkeit der Modelle

Eine wesentliche Limitierung der entwickelten Modelle ist deren eingeschränkte Skalierbarkeit. Die durchgeführten Untersuchungen haben gezeigt, dass die Generierung eines künstlichen Trainingsdatensatzes sehr rechen- und damit zeitintensiv ist. Zwar lag die Berechnungsdauer für das berücksichtigte Untersuchungsgebiet mit zwei Monaten noch in einem handhabbaren Rahmen, bei einer Ausweitung des Untersuchungsgebietes kommt dieser Ansatz aber schnell an seine Grenzen. Mit den durchgeführten Sensitivitätsanalysen konnte gezeigt werden, dass eine Reduzierung der erforderlichen Datenmenge um 80 % bei moderatem Genauigkeitsverlust möglich ist. Dennoch ist eine Ausweitung auf Stadtgebiets- oder sogar Landesebene bei den derzeitigen hardwaretechnischen Möglichkeiten nicht oder nur sehr eingeschränkt umsetzbar.

Neben der Dauer der Datengenerierung wird die Skalierbarkeit des Überflutungsflächenvorhersagemodells zusätzlich durch den hohen erforderlichen Speicherbedarf beim Trainingsprozess eingeschränkt. Aufgrund der hohen Dimensionalität der als Zielgröße verwendeten Überflutungssequenzen kommt es bei einer Ausweitung auf größere Gebiete und damit einer höheren Anzahl an Rasterzellen schnell zu einer Überschreitung des verfügbaren GPU-RAMs. Zwar lässt sich mit verschiedenen Ansätzen wie der Verwendung mehrerer parallel betriebener Modelle (Berkhahn et al. 2019) oder einem rekursiven Modell (Berkhahn und Neuweiler 2024) hier gegensteuern, die Betrachtung größerer Städte oder sogar Bundesländer bleibt aber dennoch problematisch.

Ein weiteres Problem der entwickelten Modelle stellt die fehlende Übertragbarkeit dar. Während diese im Bereich der Überflutungsflächenvorhersage durch die zusätzliche Berücksichtigung von Systeminformationen (hier räumliche Informationen) noch teilweise gegeben ist, fehlen im Bereich der Überstauvorhersage sämtliche Systeminformationen auf der Seite der Modelleingaben. Der Einsatz des Modells in einem anderen Untersuchungsgebiet setzt dementsprechend immer auch das erneute Modelltraining einschließlich der Generierung eines neuen

Trainingsdatensatzes voraus. Gleiches gilt, sobald sich Änderungen an einem betrachteten Kanalnetz selbst ergeben. Dies können einerseits bauliche Änderungen an dem Kanalnetz, die Zuweisung neuer Flächen oder aber auch bereits das Ändern von Einstellungen an den Steuerungen zentraler Bauwerke sein. In einem gewissen Maße können die dadurch entstehenden Unsicherheiten sicher toleriert werden, spätestens bei größeren Änderungen ist das erneute Training aber unumgänglich.

Da erst bei der Abdeckung größerer Gebiete mit einem Einsatz in der Praxis zu rechnen ist, ergibt sich in diesem Bereich der größte Forschungsbedarf. Neben der Entwicklung leistungsfähigerer Hardware, weniger rechenintensiven Modellen zur Datengenerierung sowie einer Verringerung des Speicherbedarfs der ML-basierten Vorhersagemodelle sind daher insbesondere auch geeignete Methoden zu entwickeln, um eine Übertragbarkeit bereits trainierter Vorhersagemodelle herzustellen. Da mit einem übertragbaren Modell auch die Skalierbarkeit eines Modells verbessert wird, würden durch einen solchen Ansatz auch die zuvor aufgeführten Probleme entschärft. Ansätze könnten physikalisch-informierte oder physikalisch-geleitete Neuronale Netze sein, die bei entsprechendem Aufbau eine Übertragbarkeit auf andere Gebiete zulassen. Da die Übertragbarkeit des Überstauvorhersagemodells derzeit am meisten eingeschränkt ist, könnte auch die Kopplung von HD-Kanalnetzmodellen mit dem Überflutungsflächenvorhersagemodell eine Lösung sein. Da HD-Kanalnetzmodelle weniger rechenintensiv sind, könnte je nach Größe des Betrachtungsgebietes auch mit der derzeit verfügbaren Hardware ein Echtzeitbetrieb möglich sein.

Gekoppelte Betrachtung und Berücksichtigung von Unsicherheiten

Im Rahmen der durchgeführten Untersuchungen konnte eine hohe Genauigkeit bei der Prognose der Überflutungssituation aufgezeigt werden, dies allerdings weitestgehend unter Verwendung von historischen Messungen oder Simulationsergebnissen als Modelleingaben. Zwar erfolgte eine gekoppelte Betrachtung der Überstau- und Überflutungsflächenprognose, weitere Eingangsdaten wie insbesondere die Prognose des bevorstehenden Niederschlags wurden jedoch nur über historische Messwerte berücksichtigt. Dabei handelt es sich um die identischen Werte, die auch zur Generierung der als Grundwahrheit verwendeten Überstauganglinien und Überflutungsflächen verwendet wurden. Dementsprechend wurde eine „perfekte Prognose“ des Niederschlags angenommen. Im operationellen Betrieb sind Niederschlagsprognosen jedoch mit diversen Unsicherheiten belastet, die sich dann auch auf die Prognose der Überflutungssituation auswirken. Gleiches gilt für die als Eingabe angesetzten Messungen des gefallenen Niederschlags und der Situation im Kanal. Auch hier wurden historische Messungen oder Simulationsergebnisse verwendet und damit keine Messunsicherheiten betrachtet. Um die Praxistauglichkeit der entwickelten Vorhersagemodelle für Überstau und Überflutungsflächen abschließend zu beurteilen, ist daher noch eine Evaluation mit realen Prognosen und Messungen als Modelleingabe sowie unter Berücksichtigung der zugrundeliegenden Unsicherheiten erforderlich.

Anhand der vorherigen Erläuterungen wird deutlich, dass im aktuellen Entwicklungszustand keine Erfassung und Kommunikation von Unsicherheiten entlang der Vorhersagekette erfolgt. Lediglich im Bereich der Überstauvorhersage erfolgt durch die Darstellung von Quantilen in

einem gewissen Maße die Berücksichtigung von Modellunsicherheiten bei der Prognose. Weitere Unsicherheiten bei Messungen im Kanal oder des gefallenen Niederschlags, der Prognose von Niederschlag und Überflutungsflächen oder aber der Generierung von Trainingsdaten mittels physikalisch-basierter Modelle werden weder erfasst noch in der Darstellung der Ergebnisse berücksichtigt. Sowohl bei der Erfassung und Quantifizierung als auch der Kommunikation von Unsicherheiten ergibt sich somit künftig noch Forschungsbedarf.

Weiterentwicklung der Modellarchitekturen

Die Entwicklung im Bereich des Maschinellen Lernens und speziell auch *Deep Learnings* hat in den letzten Jahren extrem an Geschwindigkeit aufgenommen. Jedes Jahr wird eine Vielzahl an neuen Architekturen veröffentlicht, die die Leistungsfähigkeit der bestehenden Ansätze erneut überbieten. Dies zeigt sich auch an den in dieser Arbeit verwendeten Verfahren. Während zu Beginn des Bearbeitungszeitraums viele Publikationen im Bereich der Verarbeitung von sequentiellen Daten noch auf klassische LSTM-Netze bauten, haben in der jüngsten Vergangenheit *Transformer* in diesem Bereich extrem an Relevanz gewonnen. Mit dem im Bereich der Überstauvorhersage verwendeten *Temporal Fusion Transformer* konnte gezeigt werden, dass gegenüber einem klassischen LSTM-Netz eine deutliche Steigerung der Prognosegüte erzielt werden kann. Dementsprechend sollten bei anstehenden Untersuchungen auch immer die dann aktuellen Modellansätze auf ihre Leistungsfähigkeit in Hinblick auf die vorliegende Problemstellung geprüft werden.

Weiterhin wird der Einsatz von Maschinellen Lernverfahren in der Praxis aufgrund von deren *Black-Box*-Charakteristik von den Anwendern häufig noch als kritisch angesehen. Dies ist insbesondere dann der Fall, wenn in Abhängigkeit von Modellergebnissen Entscheidungen getroffen werden, die Auswirkungen auf den Menschen haben. Immer mehr Studien konzentrieren sich daher auf die Entwicklung erklärbarer Modellarchitekturen oder Techniken, um Modellergebnisse interpretierbar zu machen (Adadi und Berrada 2018). Auch hier ist die im Rahmen der Überstauvorhersage verwendete *Temporal Fusion Transformer*-Architektur als Beispiel zu nennen, mit der durch die Anordnung spezieller Schichten die Relevanz der Eingabemerkmale auf die Zielgröße ermittelt und ausgegeben werden kann.

Verbesserung der vorhandenen Datengrundlage

Neben dem gewählten ML-Verfahren ist die Qualität der Ergebnisse maßgeblich von der für das Modelltraining zur Verfügung stehenden Datengrundlage abhängig. So hängt bei dem gewählten Aufbau die Qualität einerseits von der Genauigkeit der im Echtzeitbetrieb zur Verfügung stehenden Eingabedaten, wie insbesondere der Niederschlagsprognose ab. Andererseits ist die Prognosegüte von der Genauigkeit des hydrodynamischen Modells abhängig, das die Zielgrößen für den Trainingsprozess liefert. Aufgrund der Tatsache, dass hydrodynamische Modelle wenn überhaupt nur an wenigen punktuellen Messungen von kurzen Zeiträumen (ca. 1–2 Jahre) kalibriert werden, sind die Berechnungsergebnisse bereits mit Unsicherheiten verbunden. An der Geländeoberfläche fehlen Beobachtungsdaten vergangener Ereignisse dagegen häufig sogar komplett, wodurch keine Modellkalibrierung oder Ergebnisvalidierung in diesem Bereich möglich ist.

Gleichzeitig kann das Modell bei dem gewählten Aufbau maximal so gut wie das hydrodynamische Modell werden. Zusätzliche Messungen der Zielgrößen Überstau und Überflutungsflächen würden vor diesem Hintergrund die Möglichkeit bieten, künftig mit *Transfer Learning*-Techniken die Qualität der Prognosemodelle über die des HD-Modells hinaus anzuheben. Beim Vorliegen einer umfangreichen Datenbasis könnte irgendwann ggf. sogar komplett auf die Generierung von künstlichen Trainingsdaten verzichtet werden. Während Überläufe an zentralen Bauwerken teilweise messtechnisch erfasst werden, ist die Messung von Überstau bis auf Schachtebene derzeit nicht wirtschaftlich umsetzbar. Gleiches gilt für die zeitlich und räumlich hochauflösende Erfassung von Überflutungsflächen. Geeignete Methoden zur Erfassung des Abflussverhaltens im Kanal und an der Oberfläche wären daher wünschenswert, um künftig die Qualität der Trainingsdaten und der Vorhersagemodelle selbst zu erhöhen.

Literaturverzeichnis

- Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G. S.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Goodfellow, I.; Harp, A.; Irving, G.; Isard, M. et al. (2015): TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. Online verfügbar unter <https://www.tensorflow.org/>, Zugriff 25.01.2022.
- Abou Rjeily, Y.; Abbas, O.; Sadek, M.; Shahrour, I.; Hage Chehade, F. (2017): Flood forecasting within urban drainage systems using NARX neural network. *Water science and technology* 76 (9), S. 2401–2412. DOI: 10.2166/wst.2017.409.
- Adadi, A.; Berrada, M. (2018): Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access* 6, S. 52138–52160. DOI: 10.1109/access.2018.2870052.
- AGG (2021): Datenlieferung der Abwassergesellschaft Gelsenkirchen mbH (AGG), Kanalnetzmodell des Einzugsgebietes PW Alma in Gelsenkirchen, bereitgestellt am 11.05.2021.
- Aigner, D.; Bollrich, G. (2015): Handbuch der Hydraulik. Für Wasserbau und Wasserwirtschaft. 1. Auflage, ISBN 978-3-410-21341-3, Beuth Verlag GmbH, Berlin.
- Andersson, H.; Mariegaard, J.; Ridler, M. (2022): MIKE IO: input/output of MIKE files in Python. Version 1.2.4. Online verfügbar unter <https://dhi.github.io/mikeio/index.html>, zuletzt geprüft am 29.11.2022.
- Angermaier, G.; Braunschmidt, S.; Obermayer, A. (2012): Entwicklung eines Vorgehensmodells zur Führung eines Überflutungsnachweises in urbanen Gebieten. *KA Korrespondenz Abwasser, Abfall* 59 (5), S. 414–425.
- aqua_plan mbH (2019): AquaZIS – Messdatenmanagementsystem für die Wasserwirtschaft. Aachen. Online verfügbar unter https://aquaplan.de/downloads/dokus/aquazis_aktuell.pdf, zuletzt geprüft am 05.10.2020.
- Arjovsky, M.; Chintala, S.; Bottou, L. (2017): Wasserstein GAN. Online verfügbar unter <http://arxiv.org/pdf/1701.07875v3>, zuletzt geprüft am 05.10.2023.
- ATV-DVWK (2004a): Anforderungen an Niederschlag-Abfluss-Berechnungen in der Siedlungsentwässerung. Merkblatt ATV-DVWK-M 165. Januar 2004. Hennef, Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall (ATV-DVWK).
- ATV-DVWK (2004b): Bewertung der hydraulischen Leistungsfähigkeit bestehender Entwässerungssysteme. Arbeitsbericht der DWA-AG E S-2.1. *KA Korrespondenz Abwasser, Abfall* 51 (1), S. 69–79.
- Austin, R. J.; Chen, A. S.; Savić, D. A.; Djordjević, S. (2014): Quick and accurate Cellular Automata sewer simulator. *Journal of Hydroinformatics* 16 (6), S. 1359–1374. DOI: 10.2166/hydro.2014.070.
- Ayzel, G.; Heistermann, M.; Sorokin, A.; Nikitin, O.; Lukyanova, O. (2019): All convolutional neural networks for radar-based precipitation nowcasting. *Procedia Computer Science* 150, S. 186–192. DOI: 10.1016/j.procs.2019.02.036.

- Ayzel, G.; Scheffer, T.; Heistermann, M. (2020): RainNet v1.0: a convolutional neural network for radar-based precipitation nowcasting. *Geoscientific Model Development* 13 (6), S. 2631–2644. DOI: 10.5194/gmd-13-2631-2020.
- Bahdanau, D.; Cho, K.; Bengio, Y. (2014): *Neural Machine Translation by Jointly Learning to Align and Translate*. Online verfügbar unter <https://arxiv.org/pdf/1409.0473.pdf>, zuletzt geprüft am 09.03.2024.
- Bao, J.; Liu, P.; Ukkusuri, S. V. (2019): A spatiotemporal deep learning approach for citywide short-term crash risk prediction with multi-source data. *Accident Analysis & Prevention* 122, S. 239–254. DOI: 10.1016/j.aap.2018.10.015.
- Bates, P. D.; Horritt, M. S.; Fewtrell, T. J. (2010): A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology* 387 (1-2), S. 33–45. DOI: 10.1016/j.jhydrol.2010.03.027.
- Becker, A. (1995): Problems and progress in macroscale hydrological modelling. In: *Reinder A. Feddes (Hg.): Space and time scale variability and interdependencies in hydrological processes: Cambridge University Press*, S. 135–144.
- Becker, A.; Serban, P. (1990): Hydrological models for water-resources system design and operation. *Operational Hydrology Report (WMO)* 34.
- Ben Taieb, S.; Bontempi, G.; Atiya, A. F.; Sorjamaa, A. (2012): A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. *Expert Systems with Applications* 39 (8), S. 7067–7083. DOI: 10.1016/j.eswa.2012.01.039.
- Bengio, Y.; Simard, P.; Frasconi, P. (1994): Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks* 5 (2), S. 157–166. DOI: 10.1109/72.279181.
- Bentivoglio, R.; Isufi, E.; Jonkman, S. Nicolaas; Taormina, R. (2022): Deep Learning Methods for Flood Mapping: A Review of Existing Applications and Future Research Directions. *Hydrology and Earth System Sciences* 26 (16), S. 4345–4378. DOI: 10.5194/hess-26-4345-2022.
- Bentivoglio, R.; Isufi, E.; Jonkman, S. Nicolas; Taormina, R. (2023): Rapid spatio-temporal flood modelling via hydraulics-based graph neural networks. *Hydrology & Earth System Sciences* 27 (23), S. 4227–4246. DOI: 10.5194/hess-27-4227-2023.
- Bergstra, J.; Bardenet, R.; Bengio, Y.; Kégl, B. (2011): *Algorithms for Hyper-Parameter Optimization*. *Advances in Neural Information Processing Systems* 24. Online verfügbar unter <https://proceedings.neurips.cc/paper/4443-algorithms-for-hyper-parameter-optimization>, zuletzt geprüft am 09.03.2024.
- Berkhahn, S.; Fuchs, L.; Neuweiler, I. (2019): An ensemble neural network model for real-time prediction of urban floods. *Journal of Hydrology* 575, S. 743–754. DOI: 10.1016/j.jhydrol.2019.05.066.
- Berkhahn, S.; Neuweiler, I. (2024): Data driven real-time prediction of urban floods with spatial

- and temporal distribution. *Journal of Hydrology X*, S. 100167. DOI: 10.1016/j.hydroa.2023.100167.
- Bermúdez, M.; Cea, L.; Puertas, J. (2019): A rapid flood inundation model for hazard mapping based on least squares support vector machine regression. *Journal of Flood Risk Management* 12, S. e12522. DOI: 10.1111/jfr3.12522.
- Bermúdez, M.; Ntegeka, V.; Wolfs, V.; Willems, P. (2018): Development and Comparison of Two Fast Surrogate Models for Urban Pluvial Flood Simulations. *Water Resources Management* 32 (8), S. 2801–2815. DOI: 10.1007/S11269-018-1959-8.
- Bezirksregierung Köln (2020): 3D-Messdaten. Online verfügbar unter https://www.bezreg-koeln.nrw.de/brk_internet/geobasis/hoehenmodelle/3d-messdaten/index.html, zuletzt geprüft am 20.07.2021.
- Bezirksregierung Köln (2022): ALKIS Grundrissdaten. Online verfügbar unter https://www.bezreg-koeln.nrw.de/brk_internet/geobasis/liegenschaftskataster/aktuell/alkis-produkte-sekundaerdatenbestand/grundrissdaten/index.html, zuletzt geprüft am 22.04.2022.
- BKG (2021): Starkregengefahrenhinweise Nordrhein-Westfalen. Hg. v. Bundesamt für Kartographie und Geodäsie (BKG). Online verfügbar unter https://geoportal.de/Info/tk_04-starkregengefahrenhinweise-nrw, zuletzt geprüft am 09.03.2022.
- BKG (11.05.2023): Erfolgreicher Auftakt: Bundesweite Kartierung über die Gefahren durch Starkregen gestartet. Bundesamt für Kartographie und Geodäsie (BKG). Lukas Wimmer. Online verfügbar unter https://www.bkg.bund.de/SharedDocs/Pressemitteilungen/BKG/DE/PM_2023/230511-Kickoff-Starkregen.html, zuletzt geprüft am 07.03.2024.
- Breitner, J. (2020): *PyTorch Forecasting Documentation — pytorch-forecasting documentation*. Online verfügbar unter <https://pytorch-forecasting.readthedocs.io/en/stable/#>, zuletzt geprüft am 30.11.2023.
- Burger, G.; Kleidorfer, M.; Rauch, W. (2014a): Kanalnetzberechnung – die nächste Generation?. *KA Korrespondenz Abwasser, Abfall* 61 (3), S. 202–209.
- Burger, G.; Sitzenfrei, R.; Kleidorfer, M.; Rauch, W. (2014b): Parallel flow routing in SWMM 5. *Environmental Modelling & Software* 53, S. 27–34. DOI: 10.1016/j.envsoft.2013.11.002.
- Burricher, B.; Hofmann, J.; Koltermann da Silva, J.; Niemann, A.; Quirnbach, M. (2023): A Spatiotemporal Deep Learning Approach for Urban Pluvial Flood Forecasting with Multi-Source Data. *Water* 15 (9), S. 1760. DOI: 10.3390/w15091760.
- Burricher, B.; Koltermann da Silva, J.; Niemann, A.; Quirnbach, M. (under review): A Temporal Fusion Transformer Model to Forecast Overflow from Sewer Manholes during Pluvial Flash Flood events. *Hydrology* 11(3), S. 41. DOI: 10.3390/hydrology11030041
- Burricher, B.; Quirnbach, M.; Oelmann, M.; Niemann, A. (2021): Künstliche Intelligenz in der Wasserwirtschaft. *KA Korrespondenz Abwasser, Abfall* 68 (2), S. 94–101. DOI: 10.3242/kae2021.02.001.

- Butler, D.; Davies, J. W. (2004): Urban Drainage. CRC Press, London.
- Chang, F.-J.; Chen, P.-A.; Lu, Y.-R.; Huang, E.; Chang, K.-Y. (2014): Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control. *Journal of Hydrology* 517, S. 836–846. DOI: 10.1016/j.jhydrol.2014.06.013.
- Chaudhary, P.; D’Aronco, S.; Moy de Vitry, M.; Leitão, J. P.; Wegner, J. D. (2019): Flood-Water level estimation from social media images. *SPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 4 (2/W5), S. 5–12. DOI: 10.5194/isprs-annals-IV-2-W5-5-2019.
- Cho, K.; van Merriënboer, B.; Bahdanau, D.; Bengio, Y. (2014): On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. Online verfügbar unter <https://arxiv.org/pdf/1409.1259>, zuletzt geprüft am 09.03.2024.
- Chollet, François (2018): Deep learning with Python. Manning Publications Company, Shelter Island New York.
- Chow, V. T. (1959): Open-channel hydraulics. McGraw-Hill Book Company, New York.
- de Almeida, G.; Bates, P.; Ozdemir, H. (2018): Modelling urban floods at submetre resolution: challenges or opportunities for flood risk management?. *Journal of Flood Risk Management* 11, S. 855-S865. DOI: 10.1111/jfr3.12276.
- Defferrard, M.; Bresson, X.; Vandergheynst, P. (2016): Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. *Advances in neural information processing systems* 29.
- DHI (2015): Storm Water Runoff from Green Urban Areas. Modellers’ Guideline. Hg. v. Danish Hydraulic Institute (DHI).
- DHI (2017a): MIKE 1D. DHI Simulation Engine for 1D river and urban modelling - Reference Manual. Hg. v. Danish Hydraulic Institute (DHI).
- DHI (2017b): MIKE URBAN - Collection System. Modelling of storm water drainage networks and sewer collection systems. Danish Hydraulic Institute (DHI).
- DHI (2021a): MIKE 21 Flow Model FM - Hydrodynamic and Transport module. Scientific Documentation. Hg. v. Danish Hydraulic Institute (DHI).
- DHI (2021b): MIKE+. Release 2021 Update 1. Hg. v. Danish Hydraulic Institute (DHI). Hørsholm, Denmark.
- DHI (2022a): MIKE+. Collection System - User Guide. Hg. v. Danish Hydraulic Institute (DHI).
- DHI (2022b): Get the most out of MIKE Powered by DHI software on Azure. Online verfügbar unter <https://www.mikepoweredbydhi.com/products/mike-on-azure>, geprüft am 31.05.2022.
- DIN EN 752:2008: Entwässerungssysteme außerhalb von Gebäuden.
- DIN EN 752:2017: Entwässerungssysteme außerhalb von Gebäuden - Kanalmanagement.
- do Lago, C. A.F.; Giacomoni, M. H.; Bentivoglio, R.; Taormina, R.; Gomes, M. N.; Mendiondo,

- E. M. (2023): Generalizing rapid flood predictions to unseen urban catchments with conditional generative adversarial networks. *Journal of Hydrology* 618, S. 129276. DOI: 10.1016/j.jhydrol.2023.129276.
- Donahue, C.; McAuley, J.; Puckette, M. (2018): Adversarial Audio Synthesis. Online verfügbar unter <https://arxiv.org/pdf/1802.04208>, zuletzt geprüft am 09.03.2024.
- Dottori, F.; Todini, E. (2011): Developments of a flood inundation model based on the cellular automata approach: Testing different methods to improve model performance. *Physics and Chemistry of the Earth, Parts A/B/C* 36 (7-8), S. 266–280. DOI: 10.1016/j.pce.2011.02.004.
- Duncan, A.; Chen, A. S.; Kedwell, E. C.; Djordjević, S.; Savić, D. A. (2013): RAPIDS: Early Warning System for Urban Flooding and Water Quality Hazards. *Machine Learning in Water Systems: part of AISB Annual Convention*. University of Exeter, 3-5 April 2013.
- DWA (2006a): Hydraulische Bemessung und Nachweis von Entwässerungssystemen. Arbeitsblatt DWA-A 118. März 2006. Hennef, Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall (DWA), DWA-Regelwerk.
- DWA (2006b): Hydraulische Dimensionierung und Leistungsnachweis von Abwasserleitungen und -kanälen. Arbeitsblatt DWA-A 110. Oktober 2006, auf Aktualität geprüft 2017, korrigierte Fassung: November 2018. Hennef, Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall (DWA), DWA-Regelwerk.
- DWA (2006c): Leitlinien der integralen Siedlungsentwässerung (ISiE). Arbeitsblatt DWA-A 100. Dezember 2006. Hennef, Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall (DWA), DWA-Regelwerk.
- DWA (2012): Starkregen in Abhängigkeit von Wiederkehrzeit und Dauer. Arbeitsblatt DWA-A 531. September 2012. Hennef: Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall (DWA), DWA-Regelwerk.
- DWA (2013): Methoden der Überflutungsberechnung. Arbeitsbericht der DWA-Arbeitsgruppe ES-2.6 „Abfluss- und Schmutzfrachtsimulation“. *KA Korrespondenz Abwasser, Abfall* 60 (6), S. 506–511.
- DWA (2016): Risikomanagement in der kommunalen Überflutungsvorsorge für Entwässerungssysteme bei Starkregen. Merkblatt DWA-M 119. Hennef, Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall (DWA), DWA-Regelwerk.
- DWA (2021): Niederschlag-Abfluss- und Schmutzfrachtmodelle in der Siedlungsentwässerung – Teil 1: Anforderungen. Merkblatt DWA-M 165-1. Mai 2021. Hennef, Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall (DWA), DWA-Regelwerk.
- DWA (2024): Bewertung der hydraulischen Leistungsfähigkeit von Entwässerungssystemen. Arbeitsblatt DWA-A 118. 1. Auflage. Hennef, Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall (DWA), DWA-Regelwerk.
- DWD (o. J.): Wetter und Klima - Deutscher Wetterdienst - Leistungen - Open Data. Hg. v. Deutscher Wetterdienst. Online verfügbar unter <https://www.dwd.de/DE/leistungen/>

opendata/opendata.html, zuletzt geprüft am 14.12.2023.

- EGLV (2021a): Datenlieferung von EGLV: Langjährige Niederschlagsmessungen von acht Bodenstation aus dem EGLV-Gebiet in 1-minütiger Auflösung inklusive Shapefile der Standorte, bereitgestellt am 10.02.2021.
- EGLV (2021b): Datenlieferung von EGLV: Shapefile der befestigten Flächen für das Stadtgebiet Gelsenkirchen, bereitgestellt am 08.06.2021.
- ESRI Inc. (2020): ArcGIS Pro. Version 2.7: ESRI Inc. Online verfügbar unter <https://www.esri.com/de-de/arcgis/products/arcgis-pro/overview>, zuletzt geprüft am 30.11.2022.
- Faure, D.; Schmitt, P.; Auchet, P. (1999): Limits of radar rainfall forecasting for sewage system management: results and application in Nancy. In: *Proceedings of the 8th International Conference on Urban Storm Drainage*. Vol. 1. Sydney, Australia, 30 August 1999.
- Fewtrell, T. J.; Bates, P. D.; Horritt, M.; Hunter, N. M. (2008): Evaluating the effect of scale in flood inundation modelling in urban environments. *Hydrological Processes* 22 (26), S. 5107–5118. DOI: 10.1002/hyp.7148.
- Fewtrell, T. J.; Duncan, A.; Sampson, C. C.; Neal, J. C.; Bates, P. D. (2011): Benchmarking urban flood models of varying complexity and scale using high resolution terrestrial LiDAR data. *Physics and Chemistry of the Earth, Parts A/B/C* 36 (7-8), S. 281–291. DOI: 10.1016/j.pce.2010.12.011.
- Fraunhofer-Gesellschaft (2018): Maschinelles Lernen – Kompetenzen, Forschung, Anwendung. mit Mitteln des BMBF finanzierte Studie. Online verfügbar unter www.bigdata.fraunhofer.de/ml-studie, zuletzt geprüft am 27.04.2020.
- Frentrup, S.; Kümmel, M.; Da Koltermann Silva, J.; Burrichter, B.; Schultheis, H.; Quirnbach, M. (2024): Künstliche Intelligenz in der Prognose von extremen Wetterereignissen. In: V. Just, M. Bodemann, A. Paul und J. Weber (Hg.): *Innovative Nachhaltigkeit oder Nachhaltige Innovation*. 1. Aufl. Berlin, Heidelberg: Springer Gabler.
- Frentrup, S.; Schultheis, H.; Quirnbach, M.; Burrichter, B.; Koltermann da Silva, J.; Clemens, C.; Niemann, A.; Kunze, J. Erik; Dillhardt, M.; Jörissen, D. (2022): Intelligentes Management von Datenströmen und KI-Anwendungen in KIWaSuS. *KA Korrespondenz Abwasser Abfall* 69 (4), S. 264–270. DOI: 10.3242/kae2022.04.001.
- Fritz, C.; Dorigatti, E.; Rügamer, D. (2022): Combining graph neural networks and spatio-temporal disease models to improve the prediction of weekly COVID-19 cases in Germany. *Scientific reports* 12 (1), S. 3930. DOI: 10.1038/s41598-022-07757-5.
- Frochte, J. (2021): Maschinelles Lernen – Grundlagen und Algorithmen in Python. 3., aktualisierte Auflage. Hanser, München.
- Géron, A. (2019): Hands-on machine learning with Scikit-Learn, Keras and TensorFlow: concepts, tools, and techniques to build intelligent systems. 2nd Edition. O'Reilly.
- Gers, F. A.; Schmidhuber, J.; Cummins, F. (2000): Learning to forget: continual prediction with

- LSTM. *Neural computation* 12 (10), S. 2451–2471. DOI: 10.1162/089976600300015015.
- Ghimire, B.; Chen, A. S.; Guidolin, M.; Keedwell, E. C.; Djordjević, S.; Savić, D. A. (2013): Formulation of a fast 2D urban pluvial flood model using a cellular automata approach. *Journal of Hydroinformatics* 15 (3), S. 676–686. DOI: 10.2166/hydro.2012.245.
- Glenis, V.; Kutija, V.; Kilsby, C. G. (2018): A fully hydrodynamic urban flood modelling system representing buildings, green space and interventions. *Environmental Modelling & Software* 109, S. 272–292. DOI: 10.1016/j.envsoft.2018.07.018.
- Goodfellow, I. J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. (2014): Generative Adversarial Networks. Online verfügbar unter <https://arxiv.org/pdf/1406.2661>, zuletzt geprüft am 09.03.2024.
- Goodfellow, I.; Bengio, Y.; Courville, A. (2016): Deep learning. MIT Press. Online verfügbar unter <http://www.deeplearningbook.org/>, zuletzt geprüft am 09.03.2024.
- Grundmann, J. (2010): Analyse und Simulation von Unsicherheiten in der flächendifferenzier-ten Niederschlags-Abfluss-Modellierung. Dissertation, Technische Universität Dresden, Fakultät Forst-, Geo- und Hydrowissenschaften.
- Guidolin, M.; Chen, A. S.; Ghimire, B.; Keedwell, E. C.; Djordjević, S.; Savić, D. A. (2016): A weighted cellular automata 2D inundation model for rapid flood analysis. *Environmental Modelling & Software* 84, S. 378–394. DOI: 10.1016/j.envsoft.2016.07.008.
- Guo, Z.; Leitão, J. P.; Simões, N. E.; Moosavi, V. (2020): Data - driven flood emulation: Speeding up urban flood predictions by deep convolutional neural networks. *Journal of Flood Risk Management* 14 (1). DOI: 10.1111/jfr3.12684.
- Guo, Z.; Moosavi, V.; Leitão, J. P. (2022): Data-driven rapid flood prediction mapping with catchment generalizability. *Journal of Hydrology* 609, S. 127726. DOI: 10.1016/j.jhydrol.2022.127726.
- Gupta, H. V.; Kling, H.; Yilmaz, K. K.; Martinez, G. F. (2009): Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology* 377 (1-2), S. 80–91. DOI: 10.1016/j.jhydrol.2009.08.003.
- Hajgató, G.; Gyires-Tóth, B.; Paál, G. (2021): Reconstructing nodal pressures in water distribution systems with graph neural networks. Online verfügbar unter <https://arxiv.org/pdf/2104.13619>, zuletzt geprüft am 09.03.2024.
- Halfmann, C.; Holzmann, H. (2003): *Adaptive Modelle für die Kraftfahrzeugdynamik*. Springer Berlin, Heidelberg.
- Hamilton, W. L.; Ying, R.; Leskovec, J. (2017): Representation Learning on Graphs: Methods and Applications. Online verfügbar unter <http://arxiv.org/pdf/1709.05584.pdf>, zuletzt geprüft am 09.03.2024.
- Han, J.; Moraga, C. (1995): The influence of the sigmoid function parameters on the speed of backpropagation learning. In: *Gerhard Goos, Juris Hartmanis, Jan Leeuwen, José Mira und Francisco Sandoval (Hg.): From Natural to Artificial Neural Computation*, Bd. 930.

- Springer, Berlin Heidelberg, S. 195–201.
- HAPLUS (o. J.): Hierarchisches Frühwarn und Alarmierungssystem für plötzliche Sturzfluten nach Starkregenereignissen. BMBF – Fördermaßnahme (2017-2020). Online verfügbar unter <http://www.haplus.de/>, zuletzt geprüft am 12.10.2020.
- Harris, C. R.; Millman, K. Jarrod; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; Kern, R.; Picus, M.; Hoyer, S.; van Kerkwijk, M. H.; Brett, M. et al. (2020): Array programming with NumPy. *Nature* 585 (7825), S. 357–362. DOI: 10.1038/s41586-020-2649-2.
- Henonin, J.; Russo, B.; Mark, O.; Gourbesville, P. (2013): Real-time urban flood forecasting and modelling – a state of the art. *Journal of Hydroinformatics* 15 (3), S. 717–736. DOI: 10.2166/hydro.2013.132.
- Hochreiter, S.; Schmidhuber, J. (1997): Long short-term memory. *Neural computation* 9 (8), S. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.
- Hofmann, J.; Schüttrumpf, H. (2020): Risk-Based and Hydrodynamic Pluvial Flood Forecasts in Real Time. *Water* 12 (7), S. 1895. DOI: 10.3390/w12071895.
- Hofmann, J.; Schüttrumpf, H. (2021): floodGAN: Using Deep Adversarial Learning to Predict Pluvial Flooding in Real Time. *Water* 13 (16), S. 2255. DOI: 10.3390/w13162255.
- HoWa-innovativ (o. J.): Hochwasserfrühwarnung für kleine Einzugsgebiete mit innovativen Methoden der Niederschlagsmessung und -vorhersage. BMBF – Fördermaßnahme (2018-2021). Online verfügbar unter <http://www.howa-innovativ.sachsen.de>, zuletzt geprüft am 20.10.2020.
- HSB (2017): Ermittlung von Überflutungsgefahren mit vereinfachten und detaillierten hydrodynamischen Modellen. Praxisleitfaden, erstellt im Rahmen des DBU-Forschungsprojekts „KLAS II“. Lehrgebiet Siedlungswasserwirtschaft, Hochschule Bremen.
- HSGSim (2008): Integrierte Modellierung von Kanalnetz, Kläranlage und Gewässer. HSG-Leitfaden der Arbeitsgruppe Integrierte Modellierung. Hg. v. Hochschulgruppe „Erfahrungsaustausch Dynamische Simulation in der Siedlungswasserwirtschaft“ (HSGSim). Arbeitsgruppe Integrierte Modellierung. Online verfügbar unter <http://hsgsim.org/de/>, zuletzt geprüft am 20.05.2020.
- Hunter, J. D. (2007): Matplotlib: A 2D Graphics Environment. *Computing in science & engineering* 9 (3), S. 90–95. DOI: 10.1109/MCSE.2007.55.
- Hunter, N. M.; Bates, P. D.; Horritt, M. S.; Wilson, M. D. (2007): Simple spatially-distributed models for predicting flood inundation: A review. *Geomorphology* 90 (3-4), S. 208–225. DOI: 10.1016/j.geomorph.2006.10.021.
- Hürter, H. (2018): Erarbeitung gebietsspezifischer Anwendungsempfehlungen für bi-direktional gekoppelte 1D-2D-Überflutungsberechnungen. Dissertation. Kaiserslautern, Technische Universität Kaiserslautern, Schriftenreihe Wasser Infrastruktur Ressourcen, TU Kaiserslautern, Band 2.

- Illgen, M. (2010): Das Versickerungsverhalten durchlässig befestigter Siedlungsflächen und seine urbanhydrologische Quantifizierung. Dissertation. Technische Universität Kaiserslautern. Fachgebiet Siedlungswasserwirtschaft.
- Ioffe, S.; Szegedy, C. (2015): Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. Online verfügbar unter <http://arxiv.org/pdf/1502.03167v3>, zuletzt geprüft am 09.03.2024.
- Isola, P.; Zhu, J.-Y.; Zhou, T.; Efros, A. A. (2017): Image-to-Image Translation with Conditional Adversarial Networks. In: *IEEE (Hg.): Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA, 21.-27 July. Online verfügbar unter <http://arxiv.org/pdf/1611.07004v3>, zuletzt geprüft am 09.03.2024.
- itwh GmbH (2022): HYSTEM-EXTRAN Add-on Verteiltes Rechnen - itwh GmbH. Online verfügbar unter <https://itwh.de/de/softwareprodukte/desktop/hystem-extran-modul-verteiltes-rechnen/>, zuletzt geprüft am 31.05.2022.
- Jamali, B.; Bach, P. M.; Cunningham, L.; Deletic, A. (2019): *A Cellular Automata Fast Flood Evaluation (CA-ffé) Model. Water Resources Research*. DOI: 10.1029/2018WR023679.
- Jamali, B.; Löwe, R.; Bach, P. M.; Urich, C.; Arnbjerg-Nielsen, K.; Deletic, A. (2018): A rapid urban flood inundation and damage assessment model. *Journal of Hydrology* 564, S. 1085–1098. DOI: 10.1016/j.jhydrol.2018.07.064.
- Jasper-Tönnies, A.; Einfalt, T.; Strehz, A. (2017a): Sicherstellung der Entwässerung küstennaher, urbaner Räume unter Berücksichtigung des Klimawandels – das BMBF-Forschungsprojekt Stuck. Zwischenbericht von hydro & meteo GmbH & Co. KG zum Arbeitspaket 1. hydro & meteo GmbH & Co. KG. Lübeck.
- Jasper-Tönnies, A.; Hellmers, S.; Einfalt, T.; Strehz, A.; Fröhle, P. (2017b): Ensembles of radar nowcasts and COSMO-DE-EPS for urban flood management. *Water science and technology* 1, S. 27–35. DOI: 10.2166/wst.2018.079.
- Jhong, B.-C.; Wang, J.-H.; Lin, G.-F. (2017): An integrated two-stage support vector machine approach to forecast inundation maps during typhoons. *Journal of Hydrology* 547, S. 236–252. DOI: 10.1016/j.jhydrol.2017.01.057.
- Kachelmannwetter.com (2023): Radar HD+, Regenradar Deutschland. Hg. v. Kachelmann GmbH. Online verfügbar unter <https://kachelmannwetter.com/de/regenradar/deutschland#play-0-34-5>, zuletzt geprüft am 13.12.2023.
- Kilsdonk, R. A. H.; Bomers, A.; Wijnberg, K. M. (2022): Predicting Urban Flooding Due to Extreme Precipitation Using a Long Short-Term Memory Neural Network. *Hydrology* 9 (6), S. 105. DOI: 10.3390/hydrology9060105.
- Kingma, D. P.; Ba, J. (2014): Adam: A Method for Stochastic Optimization. Online verfügbar unter <http://arxiv.org/pdf/1412.6980v9>, zuletzt geprüft am 09.03.2024.
- Kipf, T. N.; Welling, M. (2016): Semi-Supervised Classification with Graph Convolutional Networks. Online verfügbar unter <https://arxiv.org/pdf/1609.02907>, zuletzt geprüft am 09.03.2024.

- Koelsch, D.; Sauer, D.; Scheid, C.; Neumann, J.; Engel, T. (2020): Die widerstandsfähige intelligente Stadt – Mit Sensorik und Modellen pluviale Überflutungen des Straßenraums erfassen und gezielt managen. Poster zum Verbundvorhaben der BMBF - Fördermaßnahme SENSARE. Online verfügbar unter <https://sensare.infralab.berlin/tag-der-hydrologie-2020/>, zuletzt geprüft am 09.03.2024.
- Koltermann da Silva, J.; Burrichter, B.; Quirnbach, M. (2023a): KI-basiertes Vorhersagemodell für Kurzzeit-Vorhersagen von Starkregen. *Wasserwirtschaft* 113 (7-8), S. 60–63. DOI: 10.1007/s35147-023-1875-6.
- Koltermann da Silva, J.; Burrichter, B.; Quirnbach, M. (2023b): Application of Artificial Intelligence in Rainfall Nowcasting and Flash Floods Forecast in an urban catchment: First results from Research Project KIWaSuS in Germany. Novatech 2023. Lyon, France, 03.07.2023. Online verfügbar unter <https://hal.science/hal-04176936/>.
- Krieger, K. (2018): Risikomanagement in der kommunalen Überflutungsvorsorge: Erfahrungen mit der Umsetzung von Gefährdungs- und Risikoanalysen für Entwässerungssysteme bei Starkregen nach DWA-M 119. *Regenwasser in urbanen Räumen - aqua urbana trifft Regenwasser Tage*. Unter Mitarbeit von T. G. Schmitt. Aqua Urbanica 2018. Landau i. d. Pfalz, 18.-19. Juni 2018. Kaiserslautern: Technische Universität Kaiserslautern, 43-53.
- Leandro, J.; Chen, A. S.; Djordjević, S.; Savić, D. A. (2009): Comparison of 1D/1D and 1D/2D Coupled (Sewer/Surface) Hydraulic Models for Urban Flood Simulation. *Journal of Hydraulic Engineering* 135 (6), S. 495–504. DOI: 10.1061/(ASCE)HY.1943-7900.0000037.
- Leandro, J.; Chen, A. S.; Schumann, A. (2014): A 2D parallel diffusive wave model for floodplain inundation with variable time step (P-DWave). *Journal of Hydrology* 517, S. 250–259. DOI: 10.1016/j.jhydrol.2014.05.020.
- LeCun, Y.; Bengio, Y.; Hinton, G. (2015): Deep learning. *Nature* 521 (7553), S. 436–444. DOI: 10.1038/nature14539.
- LeCun, Y.; Boser, B.; Denker, J. S.; Henderson, D.; Howard, R. E.; Hubbard, W.; Jackel, L. D. (1989): Backpropagation Applied to Handwritten Zip Code Recognition. *Neural computation* 1 (4), S. 541–551. DOI: 10.1162/neco.1989.1.4.541.
- LeCun, Y.; Cortes, C.; Burges, C. J. (2010): MNIST handwritten digit database. *ATT Labs* [Online]. Online verfügbar unter <http://yann.lecun.com/exdb/mnist>, zuletzt geprüft am 23.02.2024.
- LfU und LUBW (o. J.): Länderübergreifendes Hochwasser Portal. Hg. v. Bayerisches Landesamt für Umwelt (LfU) und Landesanstalt für Umwelt Baden-Württemberg (LUBW). Online verfügbar unter <https://www.hochwasserzentralen.de/warnungen>, zuletzt geprüft am 14.12.2023.
- Lhomme, J.; Sayers, P.; Gouldby, B.; Wills, M.; Mulet-Marti, J. (2008): Recent development and application of a rapid flood spreading method. In: *FLOODrisk 2008*, 30 September - 2 October 2008, Keble College, Oxford, UK.
- Li, L.; Jamieson, K.; DeSalvo, G.; Rostamizadeh, A.; Talwalkar, A. (2018): Hyperband: A Novel

- Bandit-Based Approach to Hyperparameter Optimization. *Journal of Machine Learning Research* 18. Online verfügbar unter <https://arxiv.org/pdf/1603.06560>, zuletzt geprüft am 09.03.2024.
- Liang, Q.; Smith, L. S. (2015): A high-performance integrated hydrodynamic modelling system for urban flood simulations. *Journal of Hydroinformatics* 17 (4), S. 518–533. DOI: 10.2166/hydro.2015.029.
- Lim, B.; Arik, S. O.; Loeff, N.; Pfister, T. (2019): Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting. Online verfügbar unter <https://arxiv.org/pdf/1912.09363.pdf>, zuletzt geprüft am 09.03.2024.
- Lin, C.; Vasić, S.; Kilambi, A.; Turner, B.; Zawadzki, I. (2005): Precipitation forecast skill of numerical weather prediction models and radar nowcasts. *Geophysical Research Letters* 32 (14). DOI: 10.1029/2005GL023451.
- Lin, G.-F.; Lin, H.-Y.; Chou, Y.-C. (2013): Development of a real-time regional-inundation forecasting model for the inundation warning system. *Journal of Hydroinformatics* 15 (4), S. 1391–1407. DOI: 10.2166/hydro.2013.202.
- Lin, Q.; Leandro, J.; Wu, W.; Bholá, P.; Disse, M. (2020): Prediction of Maximum Flood Inundation Extents With Resilient Backpropagation Neural Network: Case Study of Kulmbach. *Frontiers in Earth Science* 8. DOI: 10.3389/feart.2020.00332.
- Löwe, R.; Böhm, J.; Jensen, D. Getreuer; Leandro, J.; Rasmussen, S. Højmark (2021): U-FLOOD – Topographic deep learning for predicting urban pluvial flood water depth. *Journal of Hydrology* 603, S. 126898. DOI: 10.1016/j.jhydrol.2021.126898.
- LUBW (Hg.) (2016): Leitfaden Kommunales Starkregenrisikomanagement in Baden-Württemberg. Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg (LUBW). Karlsruhe.
- Luong, M.-T.; Pham, H.; Manning, C. D. (2015): Effective Approaches to Attention-based Neural Machine Translation. Online verfügbar unter <https://arxiv.org/pdf/1508.04025.pdf>, zuletzt geprüft am 09.03.2024.
- Ma, X.; Dai, Z.; He, Z.; Ma, J.; Wang, Y.; Wang, Y. (2017): Learning Traffic as Images: A Deep Convolutional Neural Network for Large-Scale Transportation Network Speed Prediction. *Sensors* 17 (4). DOI: 10.3390/s17040818.
- Mahesh, R. Bal; Leandro, J.; Lin, Q. (2022): *Physics Informed Neural Network for Spatial-Temporal Flood Forecasting*. In: *Climate Change and Water Security: Select Proceedings of VCDRR 2021*. Springer, Singapore, S. 77–91.
- Maksimović, Č.; Prodanović, D.; Boonya-Aroonnet, S.; Leitão, J. P.; Djordjević, S.; Allitt, R. (2009): Overland flow and pathway analysis for modelling of urban pluvial flooding. *Journal of Hydraulic Research* 47 (4), S. 512–523. DOI: 10.1080/00221686.2009.9522027.
- Maniak, Ulrich (2016): *Hydrologie und Wasserwirtschaft*. Springer, Berlin, Heidelberg.

- Mark, O.; Weesakul, S.; Apirumanekul, C.; Aroonnet, S. Boonya; Djordjević, S. (2004): Potential and limitations of 1D modelling of urban flooding. *Journal of Hydrology* 299 (3-4), S. 284–299. DOI: 10.1016/j.jhydrol.2004.08.014.
- Mateo-Garcia, G.; Veitch-Michaelis, J.; Smith, L.; Oprea, S. Vlad; Schumann, G.; Gal, Y.; Baydin, A. Güneş; Backes, D. (2021): Towards global flood mapping onboard low cost satellites with machine learning. *Scientific reports* 11 (1), S. 7249. DOI: 10.1038/s41598-021-86650-z.
- McCulloch, W. S.; Pitts, W. (1943): A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5 (4), S. 115–133. DOI: 10.1007/BF02478259.
- Meesuk, V.; Vojinovic, Z.; Mynett, A. E.; Abdullah, A. F. (2015): Urban flood modelling combining top-view LiDAR data with ground-view SfM observations. *Advances in Water Resources* 75, S. 105–117. DOI: 10.1016/j.advwatres.2014.11.008.
- Mehlhorn, J. (1998): Tracerhydrologische Ansätze in der Niederschlags-Abfluß-Modellierung: Institut für Hydrologie (Freiburger Schriften zur Hydrologie, 8).
- Meon, G.; Stein, K.; Förster, K.; Riedel, G. (2009): Abschlussbericht zum Forschungsprojekt "Untersuchung starkregengefährdeter Gebiete". Braunschweig.
- Ming, X.; Liang, Q.; Xia, X.; Li, D.; Fowler, H. J. (2020): Real - Time Flood Forecasting Based on a High-Performance 2-D Hydrodynamic Model and Numerical Weather Predictions. *Water Resources Research* 56 (7). DOI: 10.1029/2019WR025583.
- MKULNV (2013): Verordnung zur Selbstüberwachung von Abwasseranlagen - Selbstüberwachungsverordnung Abwasser (SüwVO Abw), vom 17.10.2013. Ministerium für Klimaschutz, Umwelt, Landwirtschaft, Natur- und Verbraucherschutz des Landes Nordrhein-Westfalen (MKULNV). Online verfügbar unter https://recht.nrw.de/lmi/owa/br_bes_text?anw_nr=2&bes_id=24944&aufgehoben=N.
- Mirza, M.; Osindero, S. (2014): *Conditional Generative Adversarial Nets*. Online verfügbar unter <http://arxiv.org/abs/1411.1784>, zuletzt geprüft am 09.03.2024.
- Mosavi, A.; Ozturk, P.; Chau, K. (2018): Flood Prediction Using Machine Learning Models: Literature Review. *Water* 10 (11), S. 1536. DOI: 10.3390/w10111536.
- Moy de Vitry, M.; Kramer, S.; Wegner, J. Dirk; Leitão, J. P. (2019): Scalable flood level trend monitoring with surveillance cameras using a deep convolutional neural network. *Hydrology and Earth System Sciences* 23 (11), S. 4621–4634. DOI: 10.5194/hess-23-4621-2019.
- MULNV (Hg.) (2018): Arbeitshilfe kommunales Starkregenrisikomanagement. Hochwasserisikomanagementplanung in NRW. Ministerium für Umwelt, Landwirtschaft, Natur- und Verbraucherschutz des Landes Nordrhein-Westfalen (MULNV).
- Muschalla, D.; Leimgruber, J.; Maier, R.; Tscheikner-Gratl, F.; Kleidorfer, M.; Rauch, W.; Ertl, T.; Kretschmer, F.; Sulzbacher, R. Maria; Neunteufel, R. (2015): Auf effizientem Wege von den Daten zum Model (DATMOD). Sanierungs- und Anpassungsplanung von kleinen

- und mittleren Kanalnetzen. Hg. v. Bundesministerium für Land- und Forstwirtschaft, Regionen und Wasserwirtschaft. Wien.
- Nacken, H. (2010): Wasserwirtschaftliche Modellierung 1: Modelltypen. Hg. v. Vorlesungsskript zur gleichnamigen Vorlesung an der RWTH Aachen. Aachen.
- Nair, V.; Hinton, G. E. (2010): Rectified Linear Units Improve Restricted Boltzmann Machines. In: *Proceedings of the 27th International Conference on Machine Learning*. Madison, WI, USA: Omnipress (ICML'10), S. 807–814.
- Nash, J. E.; Sutcliffe, J. V. (1970): River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology* 10 (3), S. 282–290. DOI: 10.1016/0022-1694(70)90255-6.
- Neal, J. C.; Fawcett, T. J.; Bates, P. D.; Wright, N. G. (2010): A comparison of three parallelisation methods for 2D flood inundation models. *Environmental Modelling & Software* 25 (4), S. 398–411. DOI: 10.1016/j.envsoft.2009.11.007.
- Néelz, S.; Pender, G. (2013): Benchmarking the latest generation of 2D Hydraulic Modelling Packages. Report – SC120002. Hg. v. Environment Agency. Bristol, UK.
- Nützmann, G.; Moser, H. (Hg.) (2016): Elemente einer analytischen Hydrologie. Springer Fachmedien, Wiesbaden.
- Olah, C. (2015): Understanding LSTM Networks. Online verfügbar unter <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>, zuletzt geprüft am 03.01.2022.
- Palmitessa, R.; Grum, M.; Engsig-Karup, A. Peter; Löwe, R. (2022): Accelerating hydrodynamic simulations of urban drainage systems with physics-guided machine learning. *Water research* 223, S. 118972. DOI: 10.1016/j.watres.2022.118972.
- Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimpel, N.; Antiga, L.; Desmaison, A.; Köpf, A.; Yang, E.; DeVito, Z.; Raison, M. et al. (2019): PyTorch – An Imperative Style, High-Performance Deep Learning Library. Online verfügbar unter <https://arxiv.org/pdf/1912.01703.pdf>, zuletzt geprüft am 09.03.2024.
- Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M. et al. (2011): Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, S. 2825–2830.
- Pham, B. Thai; Luu, C.; van Phong, T.; Trinh, P. Trong; Shirzadi, A.; Renoud, S.; Asadi, S.; van Le, H.; Meding, J. von; Clague, J. J. (2021): Can deep learning algorithms outperform benchmark machine learning algorithms in flood susceptibility modeling? *Journal of Hydrology* 592, S. 125615. DOI: 10.1016/j.jhydrol.2020.125615.
- Piadeh, F.; Behzadian, K.; Chen, A. S.; Kapelan, Z.; Rizzuto, J. P.; Campos, L. C. (2023): Enhancing Urban Flood Forecasting in Drainage Systems Using Dynamic Ensemble-based Data Mining. *Water research*, S. 120791. DOI: 10.1016/j.watres.2023.120791.

- Price, Roland K.; Vojinovic, Zoran (2011): *Urban Hydroinformatics: Data, Models and Decision Support for Integrated Urban Water Management*. IWA Publishing, London.
- Qin, Y.; Song, D.; Chen, H.; Cheng, W.; Jiang, G.; Cottrell, G. (2017): A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction. Online verfügbar unter <https://arxiv.org/pdf/1704.02971.pdf>, zuletzt geprüft am 09.03.2024.
- Quirnbach, M. (2003): *Nutzung von Wetterradardaten für Niederschlags- und Abflussvorhersagen in urbanen Einzugsgebieten*. Dissertation, Lehrstuhl für Hydrologie, Wasserwirtschaft und Umwelttechnik, Ruhr-Universität Bochum (Schriftenreihe Hydrologie / Wasserwirtschaft, 19).
- Quirnbach, M. (2021): KIWaSuS – KI-basiertes Warnsystem vor Starkregen und urbanen Sturzfluten. *KW Korrespondenz Wasserwirtschaft* 14 (6), S. 349–350.
- Radford, A.; Metz, L.; Chintala, S. (2015): Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. Online verfügbar unter <https://arxiv.org/pdf/1511.06434/1000>, zuletzt geprüft am 09.03.2024.
- Radford, A.; Narasimhan, K.; Salimans, T.; Sutskever, I. (2018): Improving language understanding by generative pre-training. Online verfügbar unter <https://www.mikecapptain.com/resources/pdf/gpt-1.pdf>, zuletzt geprüft am 09.03.2024.
- Raschka, S.; Mirjalili, V. (2018): *Python machine learning – Machine learning and deep learning with Python, scikit-learn, and TensorFlow*. Second edition, Packt Publishing, Birmingham, Mumbai.
- Rauch, W.; Kleidorfer, M.; Fach, S. (2010): Vom Bleistift zum Prozessor: Wandel der Modelle in der Siedlungsentwässerung. *Österreichische Wasser- und Abfallwirtschaft* 62 (3-4), S. 43–50. DOI: 10.1007/s00506-010-0172-7.
- Ravuri, S.; Lenc, K.; Willson, M.; Kangin, D.; Lam, R.; Mirowski, P.; Fitzsimons, M.; Athanasiadou, M.; Kashem, S.; Madge, S.; Prudden, R.; Mandhane, A.; Clark, A.; Brock, A.; Simonyan, K. et al. (2021): Skilful precipitation nowcasting using deep generative models of radar. *Nature* 597 (7878), S. 672–677. DOI: 10.1038/s41586-021-03854-z.
- Razavi, S.; Tolson, B. A.; Burn, D. H. (2012): Review of surrogate modeling in water resources. *Water Resources Research* 48 (7). DOI: 10.1029/2011WR011527.
- Reback, J.; McKinney, W.; Van den Bossche, J.; Augspurger, T.; Cloud, P.; Klein, A.; Hawkins, S.; Roeschke, M.; Petersen, T.; Hoefler, P.; Tratner, J.; She, C. et al. (2021): pandas-dev/pandas: Pandas 1.3.5. Zenodo.
- René, J.-R.; Djordjević, S.; Butler, D.; Mark, O.; Henonin, J.; Eisum, N.; Madsen, H. (2015): A real-time pluvial flood forecasting system for Castries, St. Lucia. *Journal of Flood Risk Management* 11, 269-283. DOI: 10.1111/jfr3.12205.
- Richter, Stefan (2019): *Statistisches und maschinelles Lernen – Gängige Verfahren im Überblick*. Springer Spektrum, Berlin.
- Rojas, Raúl (1996): *Neural Networks*. Springer, Berlin, Heidelberg.

- Rosenblatt, F. (1958): The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review* 65 (6), S. 386–408. DOI: 10.1037/h0042519.
- Rosenzweig, B. R.; McPhillips, L.; Chang, H.; Cheng, C.; Welty, C.; Matsler, M.; Iwaniec, D.; Davidson, C. I. (2018): Pluvial flood risk and opportunities for resilience. *WIREs Water* 5 (6). DOI: 10.1002/wat2.1302.
- Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. (1986): Learning representations by back-propagating errors. *Nature* 323 (6088), S. 533–536. DOI: 10.1038/323533a0.
- Salimans, T.; Goodfellow, I.; Zaremba, W.; Cheung, V.; Radford, A.; Chen, X. (2016): Improved Techniques for Training GANs. Online verfügbar unter <http://arxiv.org/pdf/1606.03498v1>, zuletzt geprüft am 09.03.2024.
- Sampson, C. C.; Fewtrell, T. J.; Duncan, A.; Shaad, K.; Horritt, M. S.; Bates, P. D. (2012): Use of terrestrial laser scanning data to drive decimetric resolution urban inundation models. *Advances in Water Resources* 41, S. 1–17. DOI: 10.1016/j.advwatres.2012.02.010.
- Scarselli, F.; Gori, M.; Tsoi, A. Chung; Hagenbuchner, M.; Monfardini, G. (2009): The graph neural network model. *IEEE Transactions on Neural Networks* 20 (1), S. 61–80. DOI: 10.1109/TNN.2008.2005605.
- Schmid, F.; Leandro, J. (2023): An ensemble data-driven approach for incorporating uncertainty in the forecasting of stormwater sewer surcharge. *Urban Water Journal* 20 (9), S. 1140–1156. DOI: 10.1080/1573062X.2023.2240309.
- Schmitt, T. G. (2017): Ortsbezogene Regenhöhen im Starkregenindexkonzept SRI12 zur Risikokommunikation in der kommunalen Überflutungsvorsorge. *KA Korrespondenz Abwasser, Abfall* 64 (4), S. 294–300. DOI: 10.3242/kae2017.04.001.
- Seleem, O.; Ayzel, G.; Bronstert, A.; Heistermann, M. (2023): Transferability of data-driven models to predict urban pluvial flood water depth in Berlin, Germany. *Natural Hazards and Earth System Sciences* 23 (2), S. 809–822. DOI: 10.5194/nhess-23-809-2023.
- Seleem, O.; Ayzel, G.; Souza, A. Costa Tomaz de; Bronstert, A.; Heistermann, M. (2022): Towards urban flood susceptibility mapping using data-driven models in Berlin, Germany. *Geomatics, Natural Hazards and Risk* 13 (1), S. 1640–1662. DOI: 10.1080/19475705.2022.2097131.
- Seneviratne, S. I.; Zhang, X.; Adnan, M.; Badi, W.; Dereczynski, C.; Di Luca, A.; Ghosh, S.; Iskandar, I.; Kossin, J.; Lewis, S.; Otto, F.; Pinto, I.; Satoh, M.; Vicente-Serrano, S. M.; Wehner, M. et al. (2021): *Weather and Climate Extreme Events in a Changing Climate*. In: V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger et al. (Hg.): *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, Cambridge University Press, United Kingdom and New York, NY, USA. S. 1513–1766.
- Serpico, S. B.; Dellepiane, S.; Boni, G.; Moser, G.; Angiati, E.; Rudari, R. (2012): Information Extraction From Remote Sensing Images for Flood Monitoring and Damage Evaluation.

- Proc. IEEE* 100 (10), S. 2946–2970. DOI: 10.1109/JPROC.2012.2198030.
- Shazeer, N.; Cheng, Y.; Parmar, N.; Tran, D.; Vaswani, A.; Koanantakool, P.; Hawkins, P.; Lee, H.; Hong, M.; Young, C.; Sepassi, R.; Hechtman, B. (2018): Mesh-TensorFlow: Deep Learning for Supercomputers. Online verfügbar unter <https://arxiv.org/pdf/1811.02084>, zuletzt geprüft am 09.03.2024.
- She, L.; You, X. (2019): A Dynamic Flow Forecast Model for Urban Drainage Using the Coupled Artificial Neural Network. *Water Resources Management* 33 (9), S. 3143–3153. DOI: 10.1007/s11269-019-02294-9.
- Shi, X.; Chen, Z.; Wang, H.; Yeung, D.-Y.; Wong, W.; Woo, W. (2015): Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. Online verfügbar unter <https://arxiv.org/pdf/1506.04214>, zuletzt geprüft am 09.03.2024.
- Shi, X.; Gao, Z.; Lausen, L.; Wang, H.; Yeung, D.-Y.; Wong, W.; Woo, W. (2017): Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model. Online verfügbar unter <https://arxiv.org/pdf/1706.03458>, zuletzt geprüft am 09.03.2024.
- Smith, L. N. (2015): Cyclical Learning Rates for Training Neural Networks. Online verfügbar unter <https://arxiv.org/pdf/1506.01186>, zuletzt geprüft am 09.03.2024.
- Solomatine, D. P.; Ostfeld, A. (2008): Data-driven modelling: some past experiences and new approaches. *Journal of Hydroinformatics* 10 (1), S. 3–22. DOI: 10.2166/hydro.2008.015.
- Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. (2014): Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *The journal of machine learning research* 15 (1), S. 1929–1958.
- Stadt Gelsenkirchen (2019): *Starkregengefahrenkarte der Stadt Gelsenkirchen*. Online verfügbar unter https://gdi.gelsenkirchen.de/mapapps/resources/apps/UN_003/index.html?lang=de, zuletzt geprüft am 08.07.2019.
- Stapel, C.; Träckner, J.; Kachholz, F. (2019): Urban Flood Modeling – Impact of Digital elevation model (DEM). Novatech 2019. Lyon, 01.07.2019.
- Steffi86w (2009): Ein Fluss in Ge-Ückendorf. Video: Youtube. Online verfügbar unter https://www.youtube.com/watch?v=W_BgAjzvs_Q, zuletzt geprüft am 14.10.2021.
- Sun, J.; Xue, M.; Wilson, J. W.; Zawadzki, I.; Ballard, S. P.; Onvlee-Hooimeyer, J.; Joe, P.; Barker, D. M.; Li, P.-W.; Golding, B.; Xu, M.; Pinto, J. (2014): Use of NWP for Nowcasting Convective Precipitation: Recent Progress and Challenges. *Bulletin of the American Meteorological Society* 95 (3), S. 409–426. DOI: 10.1175/BAMS-D-11-00263.1.
- Sutskever, I.; Vinyals, O.; Le, Q. V. (2014): Sequence to Sequence Learning with Neural Networks. Online verfügbar unter <http://arxiv.org/pdf/1409.3215.pdf>, zuletzt geprüft am 09.03.2024.
- Kabir, S.; Patidar, S.; Xia, X.; Liang, Q.; Neal, J.; Pender, G. (2020): A deep convolutional neural network model for rapid prediction of fluvial flood inundation. *Journal of Hydrology* 590, S. 125481. DOI: 10.1016/j.jhydrol.2020.125481.

- Tandler, R. (1994): Ansätze für eine parallele Überstauberechnung von Kanalnetzen. *KA Korrespondenz Abwasser, Abfall* 41 (10).
- Tien Bui, D.; Hoang, N.-D.; Martínez-Álvarez, F.; Ngo, P.-T. Thi; Hoa, P. Viet; Pham, T. Dat; Samui, P.; Costache, R. (2020): A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area. *The Science of the total environment* 701, S. 134413. DOI: 10.1016/j.scitotenv.2019.134413.
- Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; Polosukhin, I. (2017): Attention Is All You Need. Online verfügbar unter <https://arxiv.org/pdf/1706.03762.pdf>, zuletzt geprüft am 09.03.2024.
- Wang, Y.; Chen, A. S.; Fu, G.; Djordjević, S.; Zhang, C.; Savić, D. A. (2018): An integrated framework for high-resolution urban flood modelling considering multiple information sources and urban features. *Environmental Modelling & Software* 107, S. 85–95. DOI: 10.1016/j.envsoft.2018.06.010.
- Wang, Y.; Fang, Z.; Hong, H.; Peng, L. (2020): *Flood susceptibility mapping using convolutional neural network frameworks*. *Journal of Hydrology* 582, S. 124482. DOI: 10.1016/j.jhydrol.2019.124482.
- Wen, Q.; Zhou, T.; Zhang, C.; Chen, W.; Ma, Z.; Yan, J.; Sun, L. (2023): Transformers in Time Series: A Survey. Online verfügbar unter <http://arxiv.org/pdf/2202.07125v5>, zuletzt geprüft am 09.03.2024.
- Wen, R.; Torkkola, K.; Narayanaswamy, B.; Madeka, D. (2017): A Multi-Horizon Quantile Recurrent Forecaster. Online verfügbar unter <https://arxiv.org/pdf/1711.11053.pdf>, zuletzt geprüft am 09.03.2024.
- Wetteronline.com (o. J.): RegenRadar Deutschland - WetterOnline. Hg. v. WetterOnline GmbH. Online verfügbar unter <https://www.wetteronline.de/regenradar?wro=true>, zuletzt geprüft am 13.12.2023.
- Wistuba, N. (2021): Kalibrierung eines gekoppelten 1D-2D Überflutungsmodells am Beispiel des im Forschungsprojekt KIWaSuS verwendeten Untersuchungsgebiets. Bachelorarbeit. Hochschule Ruhr West, Mülheim an der Ruhr. Lehrgebiet für Siedlungswasserwirtschaft, Hydrologie und Wasserbau.
- Woernle, M. Inga (2008): Anwendbarkeit künstlicher neuronaler Netze zur Untergrundbewertung in der oberflächennahen Geothermie. Dissertation. Universität Fridericiana zu Karlsruhe, Karlsruhe. Fakultät für Bauingenieur-, Geo- und Umweltwissenschaften.
- Wright, L. (2019): Ranger - a synergistic optimizer: GitHub. Online verfügbar unter <https://github.com/lessw2020/Ranger-Deep-Learning-Optimizer>, zuletzt geprüft am 09.03.2024.
- Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Yu, P. S. (2019): A Comprehensive Survey on Graph Neural Networks. *IEEE transactions on neural networks and learning systems* 32 (1). Online verfügbar unter <https://arxiv.org/pdf/1901.00596>, zuletzt geprüft am 09.03.2024.

- Xie, H.; Tang, H.; Liao Y. H. (2009): Time series prediction based on NARX neural networks: An advanced approach. In: *2009 International Conference on Machine Learning and Cybernetics. 2009 International Conference on Machine Learning and Cybernetics 3*, IEEE, S. 1275–1279.
- Xu, M.; Dai, W.; Liu, C.; Gao, X.; Lin, W.; Qi, G.-J.; Xiong, H. (2020): Spatial-Temporal Transformer Networks for Traffic Flow Forecasting. Online verfügbar unter <https://arxiv.org/pdf/2001.02908.pdf>, zuletzt geprüft am 09.03.2024.
- Yu, B.; Yin, H.; Zhu, Z. (2017): Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. Online verfügbar unter <https://arxiv.org/pdf/1709.04875>, zuletzt geprüft am 09.03.2024.
- Yu, C.; Ma, X.; Ren, J.; Zhao, H.; Yi, S. (2020): Spatio-Temporal Graph Transformer Networks for Pedestrian Trajectory Prediction. Online verfügbar unter <https://arxiv.org/pdf/2005.08514.pdf>, zuletzt geprüft am 09.03.2024.
- Yu, L.; Zhang, W.; Wang, J.; Yu, Y. (2016): SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient. Online verfügbar unter <https://arxiv.org/pdf/1609.05473>, zuletzt geprüft am 09.03.2024.
- Zhang, D.; Lindholm, G.; Ratnaweera, H. (2018): Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring. *Journal of Hydrology* 556, S. 409–418. DOI: 10.1016/j.jhydrol.2017.11.018.
- Zhang, J.; Zheng, Y.; Sun, J.; Qi, D. (2020): Flow Prediction in Spatio-Temporal Networks Based on Multitask Deep Learning. *IEEE Transactions on Knowledge and Data Engineering* 32 (3), S. 468–478. DOI: 10.1109/TKDE.2019.2891537.
- Zhao, G.; Balstrøm, T.; Mark, O.; Jensen, M. B. (2021): Multi-Scale Target-Specified Sub-Model Approach for Fast Large-Scale High-Resolution 2D Urban Flood Modelling. *Water* 13 (3), S. 259. DOI: 10.3390/w13030259.
- Zhao, G.; Pang, B.; Xu, Z.; Peng, D.; Zuo, D. (2020): Urban flood susceptibility assessment based on convolutional neural networks. *Journal of Hydrology* 590, S. 125235. DOI: 10.1016/j.jhydrol.2020.125235.
- Zhao, L.; Song, Y.; Zhang, C.; Liu, Y.; Wang, P.; Lin, T.; Deng, M.; Li, H. (2018): T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction. *IEEE transactions on intelligent transportation systems* 21 (9), S. 3848–3858. DOI: 10.1109/TITS.2019.2935152.
- Zhu, W.; Tao, T.; Yan, H.; Yan, J.; Wang, J.; Li, S.; Xin, K. (2023): An optimized long short-term memory (LSTM)-based approach applied to early warning and forecasting of ponding in the urban drainage system. *Hydrology and Earth System Sciences* 27 (10), S. 2035–2050. DOI: 10.5194/hess-27-2035-2023.
- Zischg, A. P.; Mosimann, M.; Bernet, D. B.; Röthlisberger, V. (2018): Validation of 2D flood models with insurance claims. *Journal of Hydrology* 557, S. 350–361. DOI: 10.1016/j.jhydrol.2017.12.042.

Anhang

Anhang A-1: Netzbestandteile, Hyperparameter und sonstige Begriffe

Anhang A-2: Ergebnis der partiellen Serie

Anhang A-3: Kalibrierungsergebnis

Anhang A-4: Vorhersageergebnisse für historische Ereignisse: Überflutungsflächenvorhersage

Anhang A-5: Vorhersageergebnisse für historische Ereignisse: Überstauvorhersage

Anhang A-6: Vorhersageergebnisse für historische Ereignisse: Gekoppelte Betrachtung

A-1: Netzbestandteile, Hyperparameter und sonstige Begriffserläuterungen

Nachfolgend werden ergänzend zu den in Kapitel 3.3.2 beschriebenen Netzarchitekturen weitere Netzbestandteile, Hyperparameter und sonstige Begriffe beschrieben, die im Rahmen der Modellentwicklung in den Kapiteln 5 und 6 berücksichtigt wurden. Die Erläuterungen basieren weitestgehend auf Géron (2019) und Raschka und Mirjalili (2018).

ADAM-Algorithmus: Der *Adaptive Moment Estimation* (ADAM)-Algorithmus ist ein im Bereich des Trainings von Neuronalen Netzen weitverbreiteter Optimierungsalgorithmus. Dieser basiert auf dem Gradientenverfahren, ist gegenüber diesem aber robuster und effizienter. Der Hauptunterschied zwischen dem ADAM-Algorithmus und dem herkömmlichen Gradientenverfahren besteht darin, dass die Lernrate während des Trainingsprozesses adaptiv anhand von Momentenschätzungen angepasst wird. Dadurch wird eine bessere Leistung bei der Optimierung von Neuronalen Netzen erreicht und diese konvergieren schneller.

Attention Head Size: Ein wichtiger Bestandteil von *Transformer*-Architekturen sind sogenannte Aufmerksamkeitsmechanismen, die häufig in Form der *Multi-Head Attention* umgesetzt werden. Dabei handelt es sich um mehrere Aufmerksamkeitsmechanismen, deren Ergebnisse kombiniert werden. Die *Attention Head Size* gibt als Hyperparameter an, wie viele Aufmerksamkeitsmechanismen berücksichtigt werden. Mehrere Aufmerksamkeitsmechanismen bieten den Vorteil, dass verschiedene Aspekte der Eingaben gleichzeitig berücksichtigt werden können, führen aber auch zu mehr Parametern und dementsprechend einem höheren Trainingsaufwand.

Batch-Normalisierung: Dabei handelt es sich um eine Technik, bei der kurz vor der Aktivierungsfunktion einer Schicht in einem Neuronalen Netz die Eingaben normalisiert werden. Damit kann das Training stabilisiert und der Problematik des schwindenden/explodierenden Gradienten entgegengewirkt werden.

Concatenate-Schicht: Die *Concatenate*-Schicht ermöglicht es, Vektoren, Matrizen oder multidimensionale Datensätze (häufig als *arrays* bezeichnet) innerhalb eines Neuronalen Netzes miteinander zu verbinden/konkateneren.

Convolutional-Schicht: Eine *Convolutional*-Schicht (siehe auch Kapitel 3.3.2.2) stellt die Grundlage eines CNNs dar. Bei der Verarbeitung von Eingabedaten führt eine *Convolutional*-Schicht eine Faltungsoperation durch (vgl. Abbildung 3-8). Dabei gleitet ein sogenannter Kernel, der die Gewichtungen enthält, mit einer definierten Größe (Kernelgröße) und einer vorgegebenen Schrittweite über die Eingabedaten. An jeder Position wird aus den aktuellen Eingabewerten ein Ausgabewert durch Matrixmultiplikation der Gewichtungen des Filters berechnet. Durch diese Funktion werden die gleichen Gewichtungen an unterschiedlichen Stellen der Eingabedaten verwendet und können so gegenüber vollständig verbundenen Schichten relevante Muster an verschiedenen Orten erkennen und extrahieren. In dem Python-Framework TensorFlow werden *Convolutional*-Schicht-

ten in Abhängigkeit der Eingabedimension als 1D-, 2D- oder 3D-Schicht weiter spezifiziert.

Deconvolutional-Schicht: Im Wesentlichen führt die *Deconvolutional*-Schicht die umgekehrte Funktion einer *Convolutional*-Schicht (vgl. Kapitel 3.3.2.2) durch. Während die *Convolutional*-Schicht darauf ausgelegt ist, Merkmale aus Eingabedaten zu extrahieren und dabei in der Regel die räumliche Dimension verkleinert, ist die *Deconvolutional*-Schicht darauf fokussiert, aus den extrahierten Merkmalen die Ausgaben zu generieren und die räumliche Dimension wieder zu vergrößern. Damit ähnelt die Schicht einer *Upsampling*-Schicht, verfügt aber über zusätzliche trainierbare Gewichtungen.

Dropout: *Dropout* stellt eine beim *Deep Learning* weitverbreitete Regularisierungstechnik dar, bei der einzelne Gewichtungen während eines Trainingsschritts ignoriert werden. Dieser Prozess kann über die *Dropout*-Rate gesteuert werden. Dabei handelt es sich um einen Hyperparameter, der die Wahrscheinlichkeit angibt, mit der Neuronen ausgelassen werden. Ziel ist es, so eine Überanpassung an die Trainingsdaten zu vermeiden.

Encoder-Decoder-Struktur: Dabei handelt es sich um eine zweigeteilte Netzstruktur, bei der zunächst ein erstes Netz, der sogenannte *Encoder*, aus den Eingaben die wichtigsten Merkmale extrahiert. Anschließend generiert ein zweites Netz, der sogenannte *Decoder*, die Ausgaben aus den Merkmalen. Im Bereich der Bildverarbeitung bietet diese Art der Datenverarbeitung unter anderem die Möglichkeit, durch die Anordnung von *Pooling*-Schichten im *Encoder* die Dimension der Daten zu reduzieren, um so tiefere Netze trainieren zu können. Bei der Verarbeitung von sequentiellen Daten bietet diese Struktur den Vorteil, dass der *Encoder* bereits auf die gesamte Eingabesequenz zurückgreift und eine Repräsentation der wichtigsten Merkmale an den *Decoder* weiterleitet. Damit ist zu jedem Vorhersagezeitpunkt bereits die gesamte Eingabesequenz bekannt.

Epoche: Beim Training von Neuronalen Netzen wird jede Iteration über den Trainingsdatensatz als Epoche bezeichnet. Während einer Epoche durchläuft das Modell den Trainingsdatensatz einmal und berechnet für die enthaltenen *Samples* den Prognosefehler, der anschließend zur Optimierung der Modellparameter/Gewichtungen genutzt wird.

Gradient Clipping: Hierbei handelt es sich um eine speziell beim Training von rekurrenten Netzen verwendete Technik, bei der die Gradienten während des Optimierungsvorgangs beim Überschreiten eines Grenzwertes beschnitten werden. Ziel ist es, so der Problematik des schwindenden/explodierenden Gradienten entgegenzuwirken. Der sogenannte *Gradient Clip Value* gibt dabei den Schwellenwert an, ab dem und auf den die Gradienten beschnitten werden.

Graph Convolutional-Schicht: *Graph Convolutional*-Schichten (siehe auch Kapitel 3.3.2.3) wurden speziell für die Verarbeitung von als Graphen strukturierte Daten entwickelt. Gegenüber einem klassischen GNN wendet eine *Graph Convolutional*-Schicht die von *Convolutional*-Schichten bekannte Faltungsoperation auf Graphen an. Bei der Verarbeitung von Graphen auf Knotenebene hängt somit die Ausgabe für einen Knoten nicht nur von dessen aktuellem Wert selbst, sondern auch den Werten der Nachbarknoten ab.

Hyperparameter und Hyperparametertuning: Als Hyperparameter werden im Bereich des

maschinellen Lernens nicht trainierbare Parameter bezeichnet, mit denen die Modellstruktur und der Lernprozess beeinflusst werden. Hierzu zählen beispielsweise bei denen in dieser Arbeit verwendeten Neuronalen Netzen die Lernrate, die Anzahl an Schichten eines Netzes und deren Neuronen, die berücksichtigte Aktivierungsfunktion oder verschiedene Regularisierungsparameter wie die *Dropout*-Rate. Der Prozess der systematischen Anpassung der Hyperparameter, mit dem Ziel die Modellgenauigkeit zu erhöhen, wird als Hyperparametertuning bezeichnet.

Kernelgröße: Die Kernelgröße bezieht sich auf die Ausdehnung des Kernels in einer *Convolutional*-Schicht, der häufig auch als Filter bezeichnet wird und die Gewichte enthält. Im Rahmen der Faltungsoperation einer *Convolutional*-Schicht (vgl. Abbildung 3-8) entscheidet die Kernelgröße somit, wie viele Eingabewerte bei der Berechnung eines Ausgabewertes berücksichtigt werden. Beträgt die Kernelgröße bei der Verarbeitung eines Bildes beispielsweise 3 x 3 Pixel, so werden immer neun Pixel für die Berechnung jeden Wertes auf der Merkmalskarte berücksichtigt.

Lernrate: Die Lernrate bezeichnet einen Faktor, der darüber entscheidet, wie stark die Gewichtungen während eines Trainingsschrittes angepasst werden (siehe auch Formel (3-3)). Ein hoher Wert führt zwar zu einem schnelleren Lernfortschritt, birgt aber gleichzeitig die Gefahr, dass die Werte zu stark springen und das Minimum nicht erreicht wird. Letzteres kann durch einen niedrigen Wert vermieden werden, der aber gleichzeitig zu einem langsameren Trainingsfortschritt führt.

LSTM-Schicht: Eine LSTM-Schicht (siehe auch Kapitel 3.3.2.1) ist eine spezielle Form einer rekurrenten Schicht, die durch verschiedene Komponenten so erweitert wurde, dass auch die Modellierung längerfristiger Abhängigkeit in einer Sequenz ermöglicht wird.

Padding: Dabei handelt es sich um das zusätzliche Hinzufügen von Elementen zu einem Datenbereich. So kann *Padding* beispielsweise verwendet werden, um durch das Hinzufügen von Pixeln Bilder in einem Datensatz auf eine einheitliche Größe zu bringen. Im Bereich von CNNs kommt *Padding* häufig zur Anwendung, um durch das zusätzliche Hinzufügen von Pixeln zu einem Eingabebild die gleiche Größe beim Ausgabebild zu erhalten. Dieser Modus wird in dem verwendeten Python-Framework TensorFlow als „*same*“ bezeichnet.

Pooling-Schicht: Um in einem Neuronalen Netz die Rechenlast zu reduzieren, können zwischen einzelnen Schichten sogenannte *Pooling*-Schichten angeordnet werden. Diese funktionieren ähnlich wie die Faltungsschichten eines CNNs, nur dass diese über keine erlernbaren Parameter verfügen. Vielmehr werden alle in einem Sichtfeld liegenden Pixel entweder gemittelt oder jeweils nur der Maximalwert ausgewählt und auf einer Merkmalskarte abgebildet. Die nachfolgende Faltungsschicht muss dann nur auf eine kleinere Merkmalskarte zurückgreifen und dementsprechend müssen im Trainingsprozess auch weniger Parameter angepasst werden.

Rekurrente Schicht: Rekurrente Schichten stellen im Kontext von Neuronalen Netzen (siehe auch Kapitel 3.3.2.1) speziell für die Verarbeitung von sequentiellen Daten entwickelte Schichten dar. Gegenüber vorwärtsgerichteten Schichten wie beispielsweise in einem

MLP verfügen rekurrente Schichten über eine Rückkopplung. Dadurch erhalten diese neben dem aktuellen Wert auch den Zustand aus dem vorangegangenen Schritt als Eingabe.

Stapel (Batch): Beim Training von Neuronalen Netzen ist die Aktualisierung der Gewichtungen nach jedem Trainingspäarchen eines Datensatzes sehr rechenintensiv. Daher ist es üblich, zunächst den Fehler für mehrere Trainingspäarchen eines Stapels zu berechnen, diesen zu mitteln und auf Grundlage des Mittelwertes für den Stapel die Gewichtungen zu aktualisieren. Die Stapelgröße stellt dabei einen Hyperparameter dar, der bei größeren Werten zu einem schnelleren Trainingsfortschritt, dafür aber auch zu einem erhöhten Speicherbedarf und unter Umständen zu Instabilitäten führen kann.

Schrittweite (stride): Die Schrittweite gibt an, wie das Sichtfeld in einer *Convolutional*-Schicht über den Eingabedatensatz gleitet. Bei einer Schrittweite von eins wird das Sichtfeld bei einem Eingabebild immer um einen Pixel weitergeschoben, bei größeren Schrittweiten dementsprechend um mehrere Pixel. Größere Schrittweiten führen daher auch zu einer räumlichen Reduktion der Eingabedaten.

Trainingspäarchen (Sample): Beim Maschinellen Lernen werden den Verfahren während des Trainingsprozesses Paare aus Eingaben und dazugehörigen Ausgaben vorgegeben, die als Trainingspäarchen oder häufig wie im Englischen als *Sample* bezeichnet werden.

Upsampling-Schicht: Eine *Upsampling*-Schicht stellt das Gegenstück zur *Pooling*-Schicht dar und wird verwendet, um die räumliche Dimension innerhalb eines Netzes zu erhöhen. Eine Möglichkeit ist hierbei, fehlende Werte in der vergrößerten Dimension durch Nächste-Nachbarn-Interpolation auszufüllen.

Vollständig verbundene (Fully-Connected)-Schicht: Der Begriff vollständig verbundene Schicht gibt an, dass alle Neuronen einer Schicht mit allen Neuronen der nachfolgenden Schicht verbunden sind. Dieser Begriff kommt häufig im Bereich von MLPs zur Anwendung, bei denen sämtliche Neuronen der angrenzenden Schichten miteinander verbunden sind.

A-2: Ergebnis der partiellen Serie

Tabelle A2-1: Übersicht über die berücksichtigten Ereignisse und der zugehörigen maximalen Wiederkehrzeit ($T_{\max} \geq 5$ a) für die fünf berücksichtigten Dauerstufen. Der Stationsname lässt sich anhand der Stations-Nr. und Tabelle 3-1 zuordnen. Die Auswertung wurde auf Basis von EGLV (2021a) erstellt.

| ID | Stations-Nr. | Ereignisdauer | | | Ereignis Summe | Dauerstufe [min] | | | | |
|----|--------------|------------------|------------------|-------|----------------|------------------|------|------|------|------|
| | | Von | Bis | Dauer | | 15 | 30 | 60 | 120 | 240 |
| 1 | 4630 | 06.09.1949 20:58 | 07.09.1949 06:43 | 09:45 | 65,8 mm | | | | | 16,8 |
| 2 | 4630 | 30.09.1956 20:26 | 01.10.1956 02:31 | 06:05 | 39,8 mm | | 6,8 | 22 | 10,4 | 6,5 |
| 3 | 4630 | 11.07.1958 21:23 | 11.07.1958 23:53 | 02:30 | 33,1 mm | 8,1 | 7,3 | 9,6 | 6,2 | |
| 4 | 4630 | 11.07.1959 15:17 | 11.07.1959 16:17 | 01:00 | 29,5 mm | | 11,6 | 9,2 | | |
| 5 | 4630 | 19.05.1960 10:04 | 19.05.1960 18:24 | 08:20 | 44,9 mm | | | | | 5 |
| 6 | 4630 | 25.08.1960 17:46 | 25.08.1960 21:13 | 03:27 | 30,5 mm | 70 | 21 | 7,9 | | |
| 7 | 4630 | 15.09.1964 13:38 | 15.09.1964 16:58 | 03:20 | 33,6 mm | | | | 6,2 | |
| 8 | 4630 | 14.06.1966 14:11 | 14.06.1966 17:36 | 03:25 | 34,0 mm | 7,1 | 9,9 | 6,7 | 5,1 | |
| 9 | 4630 | 15.06.1968 16:55 | 15.06.1968 23:00 | 06:05 | 48,2 mm | | | | 15,5 | 15,9 |
| 10 | 4630 | 30.08.1968 14:29 | 31.08.1968 14:19 | 23:50 | 92,5 mm | 12,4 | 40 | 29 | 25 | 24 |
| 11 | 4630 | 28.07.1970 23:47 | 29.07.1970 07:17 | 07:30 | 28,3 mm | 9,2 | 7,5 | 5,6 | | |
| 12 | 4630 | 18.06.1975 18:37 | 18.06.1975 19:37 | 01:00 | 25,9 mm | 6,1 | 8,7 | 5,3 | | |
| 13 | 4630 | 23.06.1975 16:09 | 23.06.1975 20:29 | 04:20 | 37,3 mm | 5,2 | 9,5 | 6,6 | 7,1 | |
| 14 | 4630 | 03.09.1975 13:28 | 03.09.1975 20:08 | 06:40 | 38,3 mm | 9,5 | 7,2 | | | |
| 15 | 4630 | 04.08.1982 17:32 | 04.08.1982 23:12 | 05:40 | 67,8 mm | 2501 | 5706 | 1641 | 403 | 170 |
| 16 | 4630 | 14.07.1985 17:48 | 14.07.1985 22:53 | 05:05 | 48,9 mm | 7,4 | | | 5,7 | 12,1 |
| 17 | 4630 | 17.06.1986 15:10 | 17.06.1986 17:35 | 02:25 | 34,0 mm | 7,1 | 14 | 13,6 | 7,2 | |
| 18 | 4630 | 02.06.1992 20:32 | 02.06.1992 22:17 | 01:45 | 21,6 mm | 5,7 | | | | |
| 19 | 4630 | 12.07.1995 15:52 | 12.07.1995 23:32 | 07:40 | 50,8 mm | 5,1 | 5,5 | | | 6,3 |
| 20 | 4630 | 03.08.2000 14:28 | 03.08.2000 19:33 | 05:05 | 24,6 mm | 5,4 | 5 | | | |
| 21 | 4630 | 27.08.2000 05:08 | 27.08.2000 07:28 | 02:20 | 26,6 mm | 6,3 | | | | |
| 22 | 4630 | 16.07.2003 22:40 | 17.07.2003 01:50 | 03:10 | 32,1 mm | 7 | 13,7 | 12,5 | 5,7 | |
| 23 | 4630 | 05.07.2006 15:29 | 05.07.2006 19:09 | 03:40 | 31,1 mm | 8,3 | 18,1 | 8,5 | | |
| 24 | 4630 | 18.09.2006 00:51 | 18.09.2006 11:21 | 10:30 | 53,1 mm | | | | | 5,7 |
| 25 | 4630 | 14.06.2007 14:06 | 14.06.2007 18:31 | 04:25 | 21,4 mm | | 5 | | | |
| 26 | 4630 | 03.08.2008 21:42 | 04.08.2008 03:42 | 06:00 | 37,1 mm | | | | | |
| 27 | 4630 | 03.07.2009 13:59 | 03.07.2009 19:48 | 05:49 | 46,0 mm | 45 | 40 | 15,9 | 8,3 | 8,3 |
| 28 | 4630 | 05.07.2012 12:16 | 05.07.2012 14:01 | 01:45 | 27,8 mm | | 5,6 | 5,9 | | |
| 29 | 4630 | 20.06.2013 11:51 | 20.06.2013 14:01 | 02:10 | 44,0 mm | 11,2 | 65 | 68 | 27 | 11,8 |
| 30 | 4630 | 08.09.2013 04:29 | 08.09.2013 08:34 | 04:05 | 38,2 mm | | | | | 6 |
| 31 | 4630 | 09.06.2014 20:27 | 09.06.2014 21:12 | 00:45 | 32,4 mm | 26 | 42 | 14,4 | 6 | |
| 32 | 4630 | 14.08.2020 19:48 | 14.08.2020 21:33 | 01:45 | 29,6 mm | | | 7,4 | | |
| 33 | 2652 | 08.07.1952 18:46 | 09.07.1952 07:26 | 12:40 | 62,2 mm | 11 | 10,4 | | | |
| 34 | 2652 | 14.08.1954 23:32 | 15.08.1954 13:27 | 13:55 | 97,4 mm | | | | | 46 |
| 35 | 2652 | 05.06.1960 17:23 | 05.06.1960 18:48 | 01:25 | 27,9 mm | | 10,3 | | | |
| 36 | 2652 | 25.08.1960 02:21 | 25.08.1960 02:56 | 00:35 | 28,8 mm | 14,7 | 22 | | | |
| 37 | 2652 | 17.09.1963 17:43 | 17.09.1963 18:18 | 00:35 | 23,5 mm | 15,3 | | | | |
| 38 | 2652 | 10.06.1966 15:00 | 10.06.1966 16:40 | 01:40 | 35,5 mm | | | 12,3 | | |

| | | | | | | | | | | |
|----|------|------------------|------------------|-------|----------|------|------|------|-------|-------|
| 39 | 2652 | 19.07.1966 16:55 | 19.07.1966 18:55 | 02:00 | 33,3 mm | 15,4 | | | | |
| 40 | 2652 | 28.07.1967 21:48 | 29.07.1967 03:03 | 05:15 | 44,5 mm | 10,5 | 17,5 | 14,8 | 11,3 | 10,4 |
| 41 | 2652 | 29.05.1980 14:29 | 29.05.1980 16:39 | 02:10 | 26,1 mm | 11,9 | | | | |
| 42 | 2652 | 25.06.1981 19:06 | 25.06.1981 22:11 | 03:05 | 29,8 mm | | 10 | | | |
| 43 | 2652 | 07.06.1991 16:52 | 07.06.1991 19:37 | 02:45 | 34,8 mm | | 14,4 | 15,1 | | |
| 44 | 2652 | 02.05.1998 14:44 | 02.05.1998 16:54 | 02:10 | 38,3 mm | | | 13,2 | 12,5 | |
| 45 | 2652 | 20.08.2002 20:34 | 21.08.2002 23:24 | 02:50 | 56,1 mm | | 18,4 | 12,4 | | |
| 46 | 2652 | 08.06.2003 12:54 | 08.06.2003 19:54 | 07:00 | 39,9 mm | 11,8 | 15,5 | 17 | | |
| 47 | 2652 | 27.06.2009 17:48 | 27.06.2009 19:28 | 01:40 | 32,4 mm | | 20 | 11,2 | | |
| 48 | 2652 | 27.07.2011 16:04 | 27.07.2011 18:34 | 02:30 | 35,9 mm | 27 | 34 | 16,1 | | |
| 49 | 2652 | 23.05.2012 14:40 | 23.05.2012 21:10 | 06:30 | 46,1 mm | 25 | 57 | 60 | 27 | 14,5 |
| 50 | 2652 | 09.06.2014 20:38 | 09.06.2014 21:58 | 01:20 | 29,5 mm | 40 | 21 | | | |
| 51 | 2652 | 27.05.2018 13:55 | 27.05.2018 18:45 | 04:50 | 51,5 mm | 136 | 602 | 182 | 61 | 26 |
| 52 | 2342 | 08.07.1952 18:22 | 09.07.1952 07:22 | 13:00 | 69,3 mm | 197 | 82 | 33 | 19,7 | 10,1 |
| 53 | 2342 | 05.08.1954 16:17 | 05.08.1954 17:07 | 00:50 | 46,2 mm | 714 | 575 | 120 | 36 | 15,3 |
| 54 | 2342 | 11.07.1958 21:33 | 12.07.1958 00:08 | 02:35 | 40,1 mm | | 33 | 22 | 14,3 | |
| 55 | 2342 | 27.05.1964 14:53 | 27.05.1964 23:13 | 08:20 | 39,0 mm | | | | 11,1 | |
| 56 | 2342 | 15.06.1968 17:15 | 15.06.1968 23:15 | 06:00 | 50,7 mm | | | | | 17,8 |
| 57 | 2342 | 27.06.1968 15:16 | 27.06.1968 20:51 | 05:35 | 42,7 mm | | | 13,3 | 16,5 | |
| 58 | 2342 | 02.06.1978 13:54 | 02.06.1978 16:59 | 03:05 | 49,2 mm | 72 | 140 | 130 | 50 | 22 |
| 59 | 2342 | 12.07.1995 14:35 | 13.07.1995 00:00 | 09:25 | 52,6 mm | | | | | 19,9 |
| 60 | 2342 | 15.06.1998 17:30 | 16.06.1998 04:55 | 11:25 | 44,4 mm | | | 15,5 | 12,8 | |
| 61 | 2342 | 27.06.2001 15:16 | 27.06.2001 19:26 | 04:10 | 39,1 mm | | | 17 | 13,9 | |
| 62 | 2342 | 16.07.2003 17:16 | 17.07.2003 00:51 | 07:35 | 51,5 mm | 38 | 215 | 95 | 47 | 25 |
| 63 | 2342 | 26.07.2008 14:25 | 26.07.2008 19:25 | 05:00 | 46,0 mm | 19,8 | 11,9 | | | 11,9 |
| 64 | 2342 | 03.07.2009 14:44 | 03.07.2009 20:14 | 05:30 | 50,9 mm | | | | | 18,9 |
| 65 | 2342 | 18.08.2011 18:58 | 18.08.2011 21:03 | 02:05 | 31,4 mm | 15,3 | | | | |
| 66 | 2342 | 20.06.2013 12:29 | 20.06.2013 14:04 | 01:35 | 48,7 mm | 51 | 87 | 129 | 50 | 20 |
| 67 | 2241 | 09.06.1949 13:26 | 10.06.1949 11:21 | 21:55 | 112,5 mm | | | | 46 | 372 |
| 68 | 2241 | 08.07.1952 15:30 | 09.07.1952 06:10 | 14:40 | 52,0 mm | 31 | 27 | 12,9 | | |
| 69 | 2241 | 18.06.1953 15:24 | 18.06.1953 16:14 | 00:50 | 29,4 mm | 11,8 | 18,8 | | | |
| 70 | 2241 | 30.09.1956 20:07 | 01.10.1956 02:28 | 06:21 | 50,3 mm | 60 | 51 | 112 | 41 | 21 |
| 71 | 2241 | 27.05.1964 14:09 | 27.05.1964 18:09 | 04:00 | 52,9 mm | 13 | 55 | 67 | 45 | 33 |
| 72 | 2241 | 15.06.1968 17:01 | 15.06.1968 22:56 | 05:55 | 98,7 mm | 35 | 71 | 76 | 3176 | 4318 |
| 73 | 2241 | 27.07.1971 12:13 | 27.07.1971 18:23 | 06:10 | 34,3 mm | 47 | 34,3 | | | |
| 74 | 2241 | 05.07.2006 14:22 | 05.07.2006 19:09 | 04:47 | 42,6 mm | | 14,6 | 11,1 | 12,9 | |
| 75 | 2241 | 26.07.2006 21:57 | 26.07.2006 23:12 | 01:15 | 39,3 mm | 10 | 58 | 39 | 14,8 | |
| 76 | 2241 | 14.06.2007 14:01 | 14.06.2007 20:16 | 06:15 | 29,7 mm | 18,3 | | | | |
| 77 | 2241 | 26.07.2008 14:04 | 26.07.2008 19:24 | 05:20 | 113,8 mm | 90 | 357 | 7014 | 52846 | 28867 |
| 78 | 2241 | 18.08.2011 19:08 | 18.08.2011 22:13 | 03:05 | 36,7 mm | 74 | 25 | 13,3 | | |
| 79 | 2241 | 20.06.2013 12:11 | 20.06.2013 13:41 | 01:30 | 49,1 mm | 26 | 35 | 108 | 52 | 21 |
| 80 | 2241 | 07.06.2018 16:57 | 07.06.2018 19:32 | 02:35 | 32,7 mm | | 19,2 | 13,5 | | |
| 81 | 2582 | 04.06.1947 23:38 | 05.06.1947 03:28 | 03:50 | 20,3 mm | 5,7 | | | | |
| 82 | 2582 | 09.06.1949 00:05 | 09.06.1949 04:05 | 04:00 | 69,6 mm | 10,7 | 48 | 252 | 391 | 231 |
| 83 | 2582 | 04.05.1951 17:44 | 04.05.1951 21:44 | 04:00 | 46,6 mm | | | | 11,5 | 16,1 |
| 84 | 2582 | 14.08.1954 23:33 | 15.08.1954 13:53 | 14:20 | 79,3 mm | | | | | 9,4 |

| | | | | | | | | | | |
|-----|------|------------------|------------------|-------|----------|------|------|------|-------|--------|
| 85 | 2582 | 27.08.1954 18:06 | 27.08.1954 19:11 | 01:05 | 27,6 mm | 5,6 | 7,2 | 6,8 | | |
| 86 | 2582 | 17.09.1963 15:34 | 17.09.1963 19:54 | 04:20 | 28,9 mm | 12,2 | 7 | | | |
| 87 | 2582 | 19.07.1967 22:41 | 19.07.1967 23:06 | 00:25 | 18,6 mm | 7,6 | | | | |
| 88 | 2582 | 15.06.1968 17:04 | 15.06.1968 23:24 | 06:20 | 39,5 mm | | | | 5,2 | 6,3 |
| 89 | 2582 | 29.03.1974 16:22 | 29.03.1974 20:22 | 04:00 | 41,3 mm | 8,3 | 36 | 28 | 14,4 | 8,7 |
| 90 | 2582 | 03.09.1975 12:39 | 03.09.1975 18:24 | 05:45 | 55,0 mm | 5,1 | 6,8 | | | 6 |
| 91 | 2582 | 04.08.1982 18:14 | 04.08.1982 22:09 | 03:55 | 28,1 mm | 6,3 | 5,8 | | | |
| 92 | 2582 | 14.07.1985 19:40 | 14.07.1985 22:30 | 02:50 | 25,8 mm | 5,3 | | | | |
| 93 | 2582 | 22.08.1987 20:50 | 22.08.1987 21:40 | 00:50 | 22,1 mm | | 5,1 | | | |
| 94 | 2582 | 28.06.1988 17:19 | 28.06.1988 18:14 | 00:55 | 18,0 mm | 5,2 | | | | |
| 95 | 2582 | 12.07.1995 19:10 | 12.07.1995 23:40 | 04:30 | 65,5 mm | | | 9,5 | 55 | 104 |
| 96 | 2582 | 10.08.1996 14:35 | 10.08.1996 19:45 | 05:10 | 31,6 mm | 5,8 | 6,8 | 8,4 | 5,2 | |
| 97 | 2582 | 02.07.2000 21:29 | 03.07.2000 04:44 | 07:15 | 44,1 mm | | | 17,5 | 21 | 11,7 |
| 98 | 2582 | 16.09.2000 02:18 | 16.09.2000 03:33 | 01:15 | 24,6 mm | | 5,4 | | | |
| 99 | 2582 | 27.06.2009 17:32 | 27.06.2009 19:47 | 02:15 | 23,4 mm | 5 | | | | |
| 100 | 2582 | 03.07.2009 17:41 | 03.07.2009 20:01 | 02:20 | 54,1 mm | 216 | 905 | 248 | 100 | 39 |
| 101 | 2582 | 07.10.2009 17:20 | 08.10.2009 09:20 | 16:00 | 59,5 mm | | | 6,4 | 7,1 | 7,4 |
| 102 | 2582 | 03.07.2010 15:22 | 03.07.2010 17:47 | 02:25 | 81,2 mm | 158 | 1309 | 6061 | 3324 | 891 |
| 103 | 2582 | 18.08.2011 18:45 | 18.08.2011 22:10 | 03:25 | 32,2 mm | | 10,7 | 7,8 | 5,7 | |
| 104 | 2582 | 09.06.2014 20:25 | 09.06.2014 22:55 | 02:30 | 25,8 mm | 10,5 | 9,6 | | | |
| 105 | 2582 | 28.07.2014 16:01 | 28.07.2014 16:36 | 00:35 | 23,0 mm | 9,1 | 6,6 | | | |
| 106 | 2582 | 30.05.2016 03:33 | 30.05.2016 18:13 | 14:40 | 49,0 mm | 6,9 | 5,8 | 15,9 | 11,3 | 12,6 |
| 107 | 8767 | 12.06.1940 19:40 | 13.06.1940 17:25 | 21:45 | 56,9 mm | | | 10 | 24 | 27 |
| 108 | 8767 | 02.08.1948 16:23 | 02.08.1948 17:18 | 00:55 | 42,5 mm | 113 | 263 | 68 | 22 | |
| 109 | 8767 | 11.07.1951 16:40 | 11.07.1951 18:30 | 01:50 | 24,2 mm | 18 | | | | |
| 110 | 8767 | 31.07.1951 18:19 | 01.08.1951 00:19 | 06:00 | 49,1 mm | 48 | 297 | 107 | 38 | 20 |
| 111 | 8767 | 15.06.1968 16:45 | 15.06.1968 22:30 | 05:45 | 128,8 mm | 56 | 339 | 3490 | 69863 | 136263 |
| 112 | 8767 | 04.08.1982 16:43 | 04.08.1982 22:38 | 05:55 | 33,3 mm | 51 | 43 | 14,8 | | |
| 113 | 8767 | 24.06.1983 16:00 | 24.06.1983 20:45 | 04:45 | 71,0 mm | 1705 | 1921 | 1930 | 558 | 268 |
| 114 | 8767 | 04.09.1987 13:55 | 04.09.1987 17:15 | 03:20 | 43,1 mm | | | 22 | 22 | 10,7 |
| 115 | 8767 | 11.07.1995 18:04 | 11.07.1995 20:34 | 02:30 | 32,2 mm | 15,6 | 16,2 | | | |
| 116 | 8767 | 29.08.1996 14:02 | 31.08.1996 12:12 | 22:10 | 72,6 mm | | | 18,8 | 17,7 | 10,2 |
| 117 | 8767 | 12.07.1999 15:20 | 12.07.1999 16:35 | 01:15 | 65,4 mm | 20 | 147 | 1172 | 440 | 143 |
| 118 | 8767 | 13.07.1999 16:42 | 13.07.1999 18:12 | 01:30 | 25,1 mm | 16,9 | | | | |
| 119 | 8767 | 18.06.2002 17:27 | 18.06.2002 18:42 | 01:15 | 31,7 mm | | | 10,9 | | |
| 120 | 8767 | 27.07.2005 17:43 | 28.07.2005 04:48 | 11:05 | 25,4 mm | 11,4 | | | | |
| 121 | 8767 | 22.07.2006 13:58 | 22.07.2006 16:33 | 02:35 | 34,1 mm | | | 14,1 | | |
| 122 | 8767 | 18.08.2011 19:25 | 18.08.2011 22:40 | 03:15 | 34,6 mm | 19,8 | 15,3 | | | |
| 123 | 8767 | 20.06.2013 12:57 | 20.06.2013 13:42 | 00:45 | 24,5 mm | 17,9 | | | | |
| 124 | 8767 | 11.07.2019 16:43 | 11.07.2019 17:28 | 00:45 | 26,1 mm | 21 | 12,1 | | | |
| 125 | 8767 | 15.08.2020 14:42 | 15.08.2020 15:52 | 01:10 | 39,0 mm | 45 | 136 | 40 | 14,2 | |
| 126 | 2368 | 09.06.1949 00:15 | 09.06.1949 04:25 | 04:10 | 48,3 mm | | | | 13,5 | 19,3 |
| 127 | 2368 | 21.07.1964 14:27 | 21.07.1964 15:27 | 01:00 | 36,1 mm | | 16,7 | 25 | | |
| 128 | 2368 | 19.07.1966 17:00 | 19.07.1966 19:50 | 02:50 | 23,3 mm | 11,5 | | | | |
| 129 | 2368 | 27.06.1968 15:03 | 27.06.1968 20:03 | 05:00 | 39,4 mm | | | | 12,1 | |
| 130 | 2368 | 10.07.1981 17:04 | 11.07.1981 00:24 | 07:20 | 52,4 mm | 64 | 50 | 20 | | |

| | | | | | | | | | | |
|-----|------|------------------|------------------|-------|---------|------|------|------|------|------|
| 131 | 2368 | 14.09.1984 15:18 | 14.09.1984 18:58 | 03:40 | 36,4 mm | | 14,3 | 21 | | |
| 132 | 2368 | 17.06.1986 14:44 | 17.06.1986 20:09 | 05:25 | 46,8 mm | | 28 | 47 | 27 | |
| 133 | 2368 | 03.06.2003 19:20 | 03.06.2003 20:30 | 01:10 | 36,0 mm | 10,8 | | 20 | | |
| 134 | 2368 | 03.07.2009 13:17 | 03.07.2009 20:09 | 06:52 | 71,5 mm | 323 | 1075 | 476 | 120 | 155 |
| 135 | 2368 | 07.10.2009 17:11 | 08.10.2009 07:51 | 14:40 | 66,5 mm | 19,4 | 18,6 | 21 | 26 | 19,7 |
| 136 | 2368 | 20.06.2013 12:33 | 20.06.2013 14:13 | 01:40 | 54,5 mm | 47 | 172 | 313 | 106 | 40 |
| 137 | 2368 | 08.09.2013 04:31 | 08.09.2013 12:36 | 08:05 | 43,8 mm | | | | | 10,7 |
| 138 | 2368 | 12.07.2014 19:27 | 13.07.2014 03:42 | 08:15 | 56,5 mm | 19,9 | 15,9 | 28 | 86 | 48 |
| 139 | 2368 | 18.09.2014 13:08 | 19.09.2014 00:48 | 11:40 | 50,9 mm | | | | | 11,3 |
| 140 | 2587 | 03.06.1951 15:51 | 03.06.1951 17:41 | 01:50 | 26,4 mm | 12,4 | 10,7 | | | |
| 141 | 2587 | 15.08.1952 13:39 | 15.08.1952 15:48 | 02:09 | 37,6 mm | 26 | 12,7 | 10,1 | 11,6 | |
| 142 | 2587 | 14.08.1954 23:36 | 15.08.1954 14:06 | 14:30 | 80,6 mm | | | | | 12,5 |
| 143 | 2587 | 26.09.1956 20:25 | 27.09.1956 10:40 | 14:15 | 71,2 mm | | | | | 18,4 |
| 144 | 2587 | 19.07.1967 22:23 | 20.07.1967 01:03 | 02:40 | 25,2 mm | 35 | 11,4 | | | |
| 145 | 2587 | 15.06.1968 17:07 | 15.06.1968 23:27 | 06:20 | 52,0 mm | | | 15,9 | 26 | 26 |
| 146 | 2587 | 24.06.1975 13:39 | 24.06.1975 15:54 | 02:15 | 35,8 mm | 25 | 24 | 13,6 | | |
| 147 | 2587 | 20.05.1978 14:01 | 20.05.1978 16:11 | 02:10 | 43,7 mm | | | | 26 | 11,4 |
| 148 | 2587 | 13.07.1979 16:26 | 14.07.1979 00:41 | 08:15 | 49,7 mm | | | 23 | 15,2 | 12,5 |
| 149 | 2587 | 20.07.1980 10:42 | 21.07.1980 16:37 | 05:55 | 65,4 mm | 14 | | | | |
| 150 | 2587 | 01.08.2002 08:03 | 01.08.2002 12:28 | 04:25 | 45,9 mm | | | 14,3 | 19 | 14,6 |
| 151 | 2587 | 16.07.2003 22:38 | 16.07.2003 23:58 | 01:20 | 32,2 mm | | | 13,1 | | |
| 152 | 2587 | 28.07.2011 15:39 | 28.07.2011 17:09 | 01:30 | 30,4 mm | | 12,9 | | | |
| 153 | 2587 | 28.07.2014 15:47 | 28.07.2014 17:27 | 01:40 | 33,6 mm | 11,6 | 44 | 17,2 | | |

A-3: Kalibrierungsergebnis

A 3-1: Finale Parametersätze

Tabelle A3-1: Flächenspezifische Parametrisierung des Kanalnetzmodells unter Berücksichtigung des verwendeten Modellansatzes C2

| Flächennutzung | Abflussbildung | | Abflusskonzentration |
|-----------------------------|------------------------|---------------------------|----------------------------|
| | Anfangsverlust [mm] | Reduzierungsfaktor [-] | Speicherkonstante [min] |
| Dachflächen | 2 | 1 | 12 |
| Öffentliche Verkehrsflächen | 2 | 0,95 | 12 |
| Privat schmutzig | 2 | 0,8 | 12 |
| Privat sauber | - | - | - |

Tabelle A3-2: Flächenspezifische Parametrisierung des Oberflächenmodells

| Flächennutzung | Infiltrationsmodell | | | | | Rauheit [m ^{1/3} /s] |
|-----------------------------|-----------------------------|------------------------|--------------|------------------|---------------------|----------------------------------|
| | Infiltrationsrate [mm/h] | Leckage Rate [mm/h] | Tiefe [m] | Porosität [-] | Wassergehalt [%] | |
| Dachflächen | 0 | 0 | 0 | 0 | 0 | 3 |
| Durchlässig | 25 | 1,8 | 1,1 | 0,1 | 90 | 20 |
| Privat sauber | 4,5 | 1,8 | 1,1 | 0,1 | 90 | 40 |
| Privat schmutzig | 4,5 | 1,8 | 1,1 | 0,1 | 90 | 50 |
| Öffentliche Verkehrsflächen | 0 | 0 | 0 | 0 | 0 | 70 |

A 3-2: Gemessene und simulierte Durchflussganglinien für die unterschiedlichen Ereignisse und Messstationen

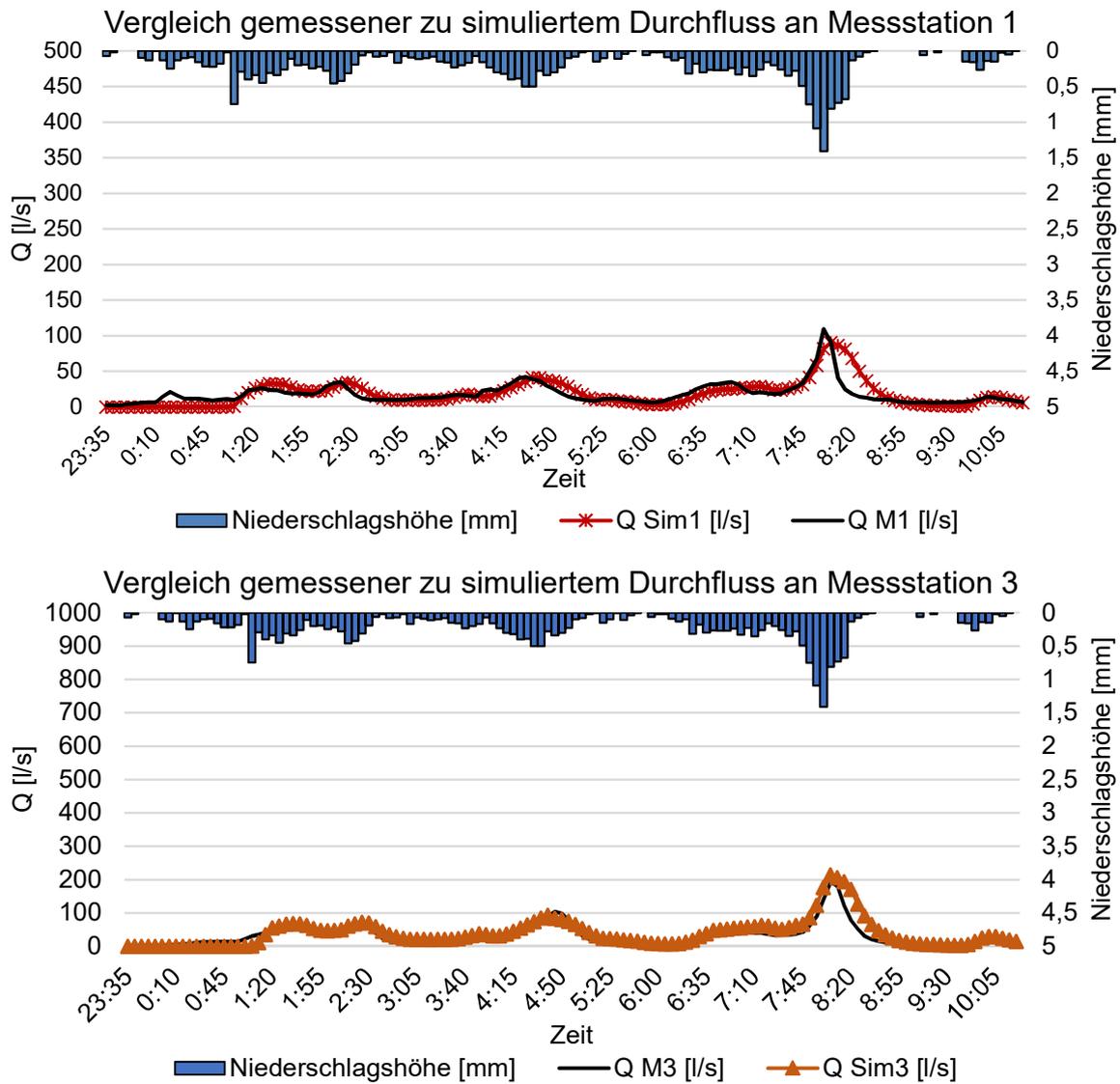


Abbildung A3-1: Kalibrierungsergebnis für das Ereignis 1 an den beiden Messstationen

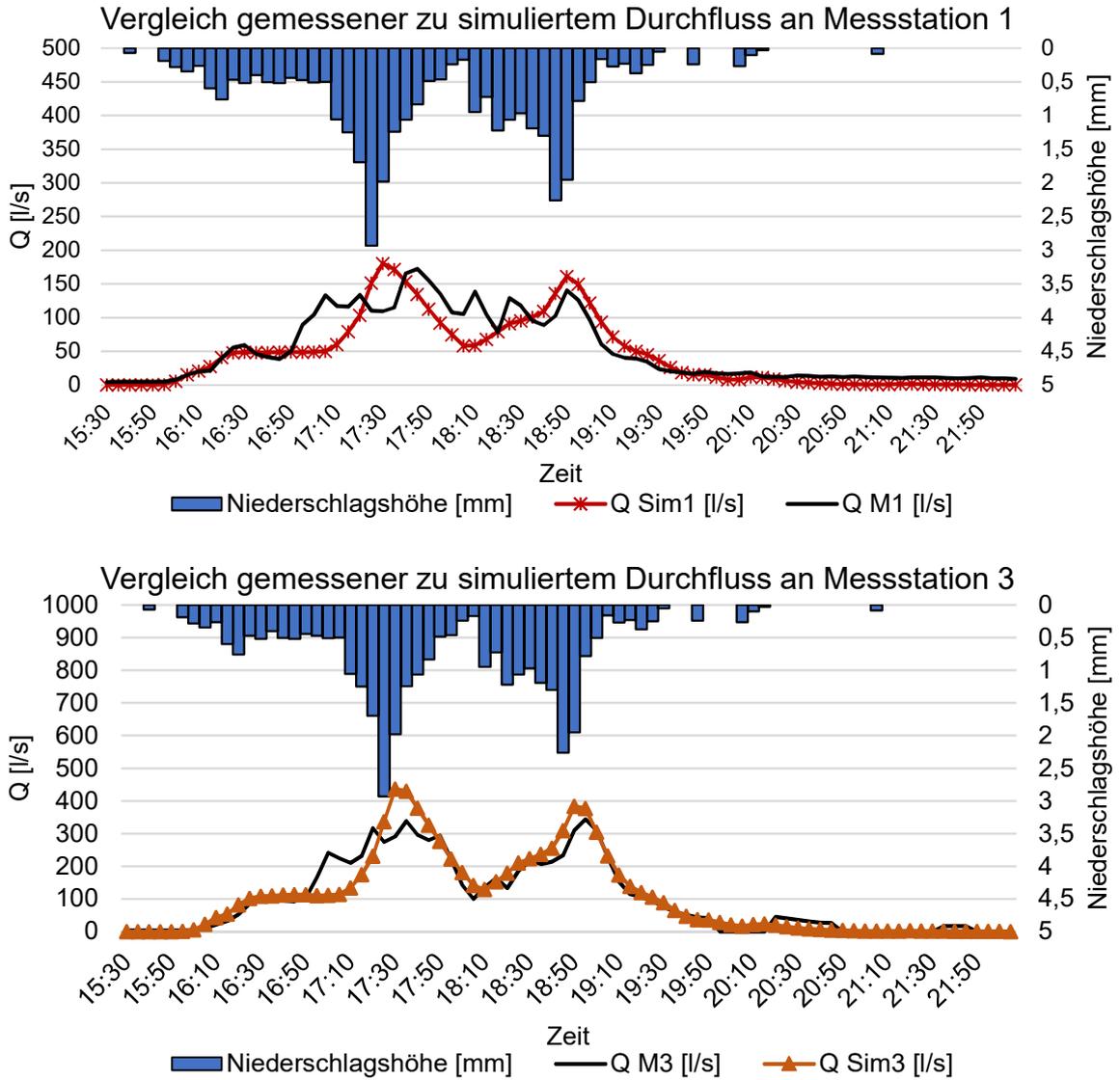


Abbildung A3-2: Kalibrierungsergebnis für das Ereignis 2 an den beiden Messstationen

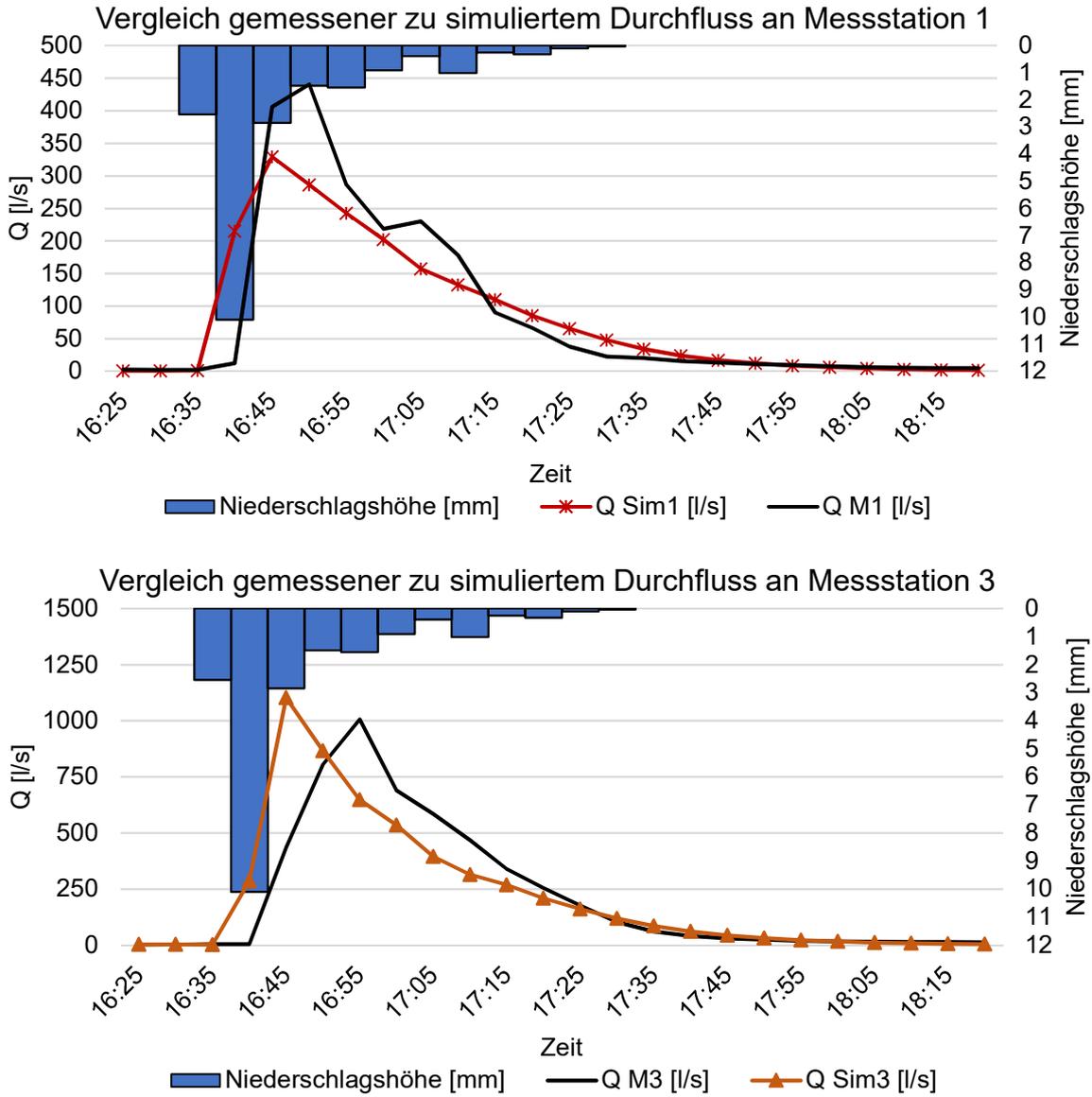


Abbildung A3-3: Kalibrierungsergebnis für das Ereignis 3 an den beiden Messstationen

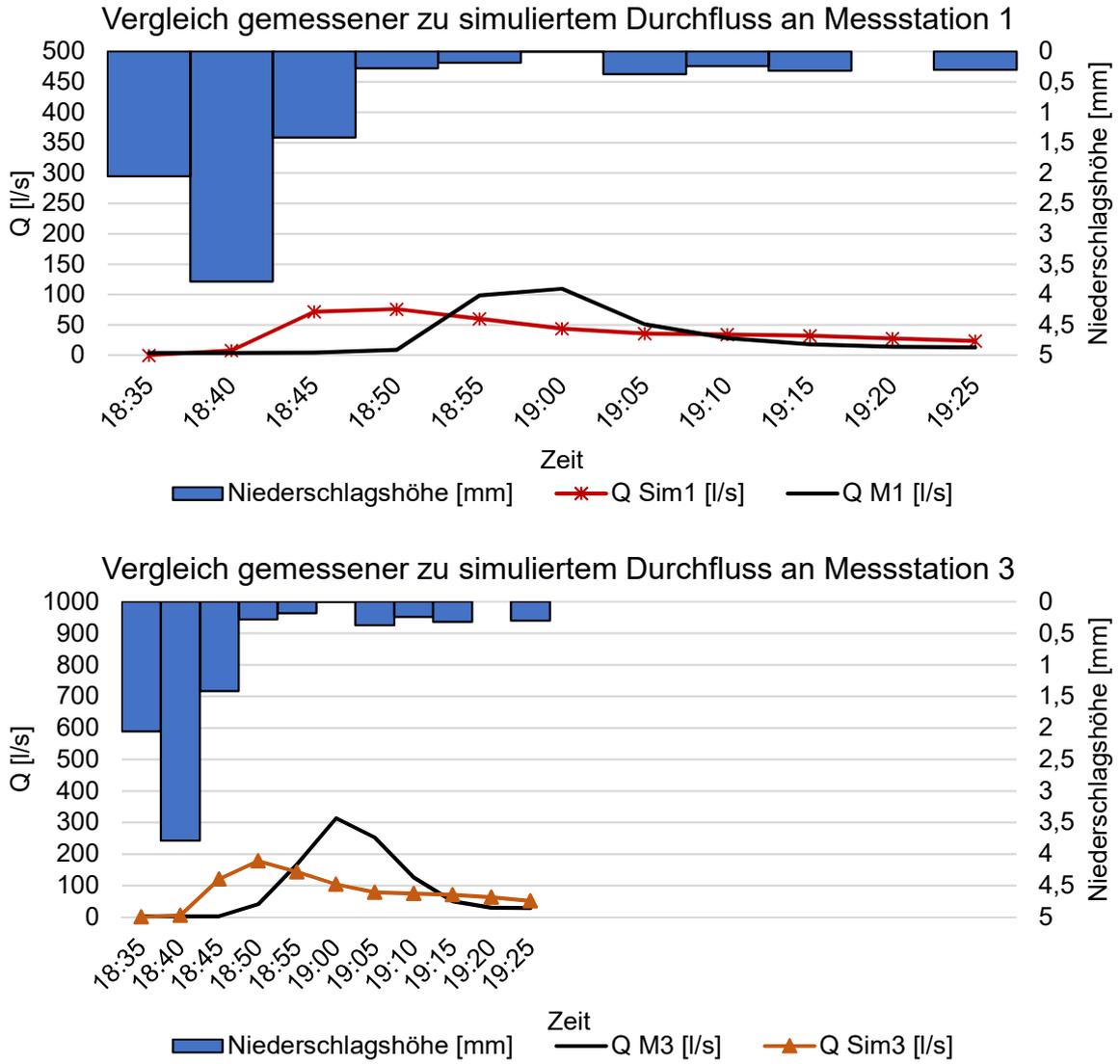


Abbildung A3-4: Kalibrierungsergebnis für das Ereignis 4 an den beiden Messstationen

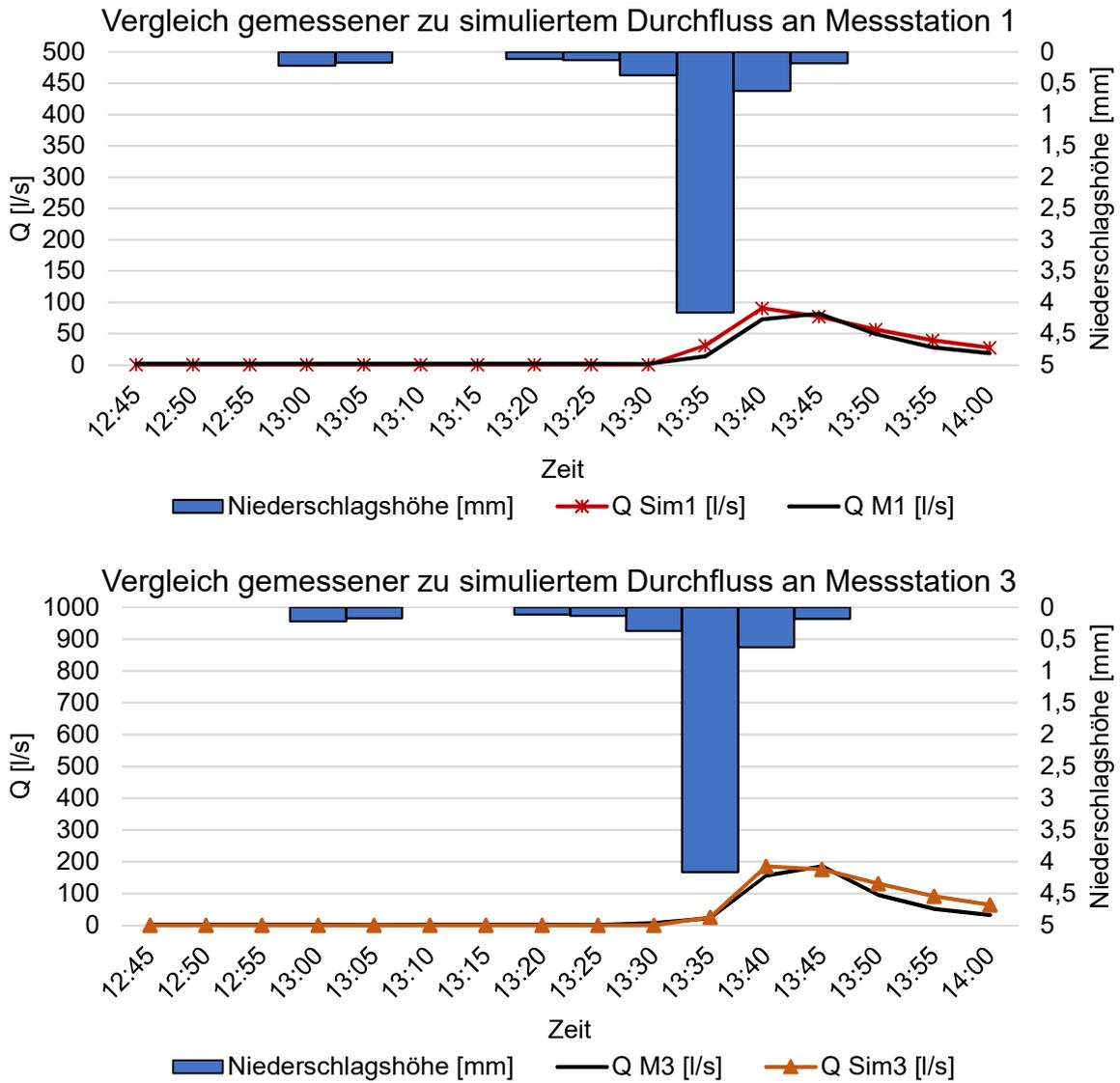


Abbildung A3-5: Kalibrierungsergebnis für das Ereignis 5 an den beiden Messstationen

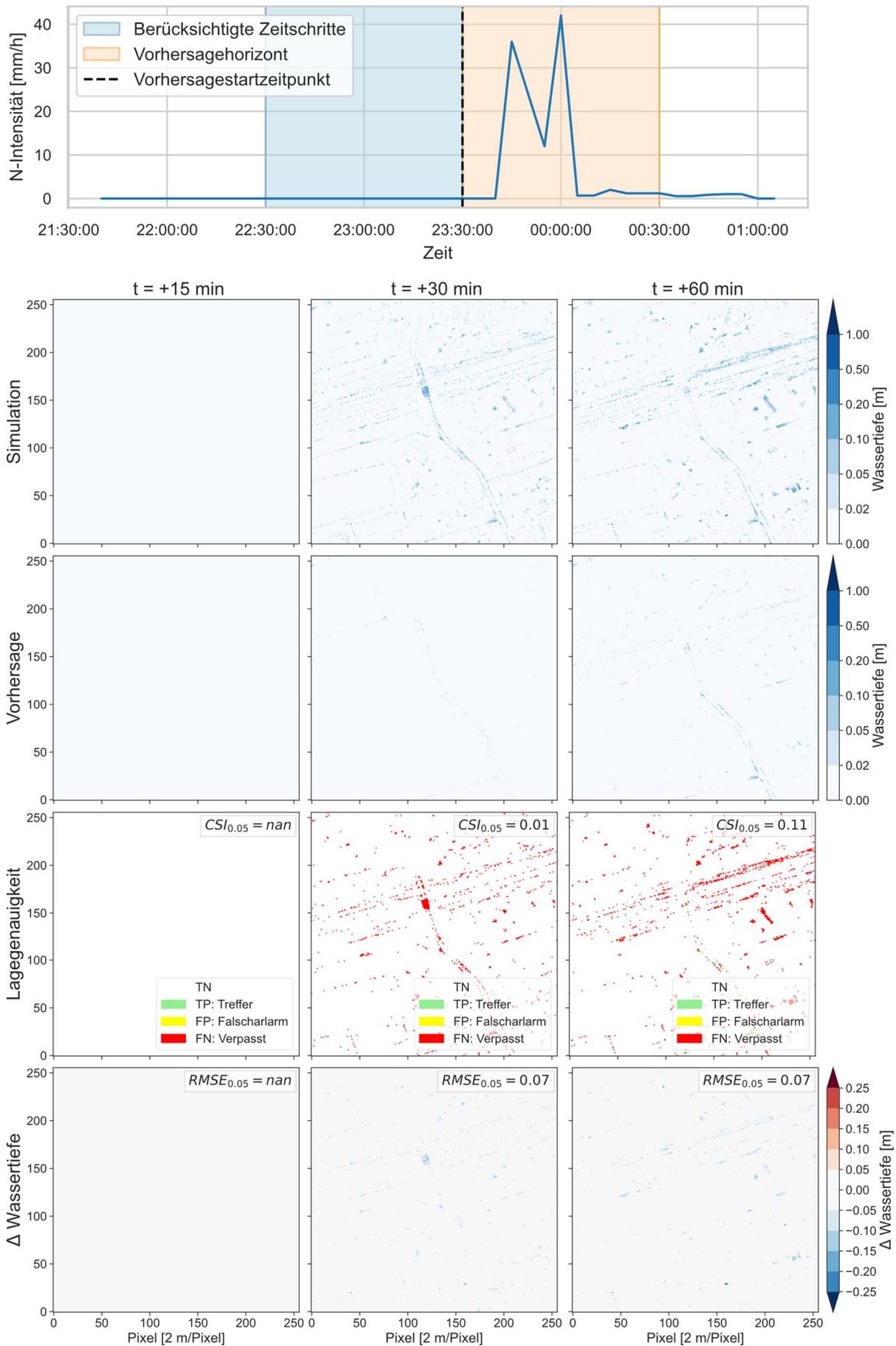
A-4: Vorhersageergebnisse für historische Ereignisse: Überflutungsflächenvorhersage

Nachfolgend werden für die in Tabelle A4-1 aufgeführten Ereignisse der Station Gelsenkirchen-Altstadt (Stations-Nr. 2582) Vorhersageergebnisse für Vorhersagen zu Beginn, Mitte und Ende des jeweiligen Ereignisses dargestellt. Bei der Auswahl der Ereignisse aus dem Testdatensatz wurden die drei Ereignisse mit den höchsten Wiederkehrzeiten (ID 82, 100, 102), ein Ereignis mit einer mittleren Wiederkehrzeit (ID 89) und ein Ereignis mit einer geringen Wiederkehrzeit (ID 81) ausgewählt. Letzteres führte bei der Evaluierung zu den schlechtesten Ergebnissen.

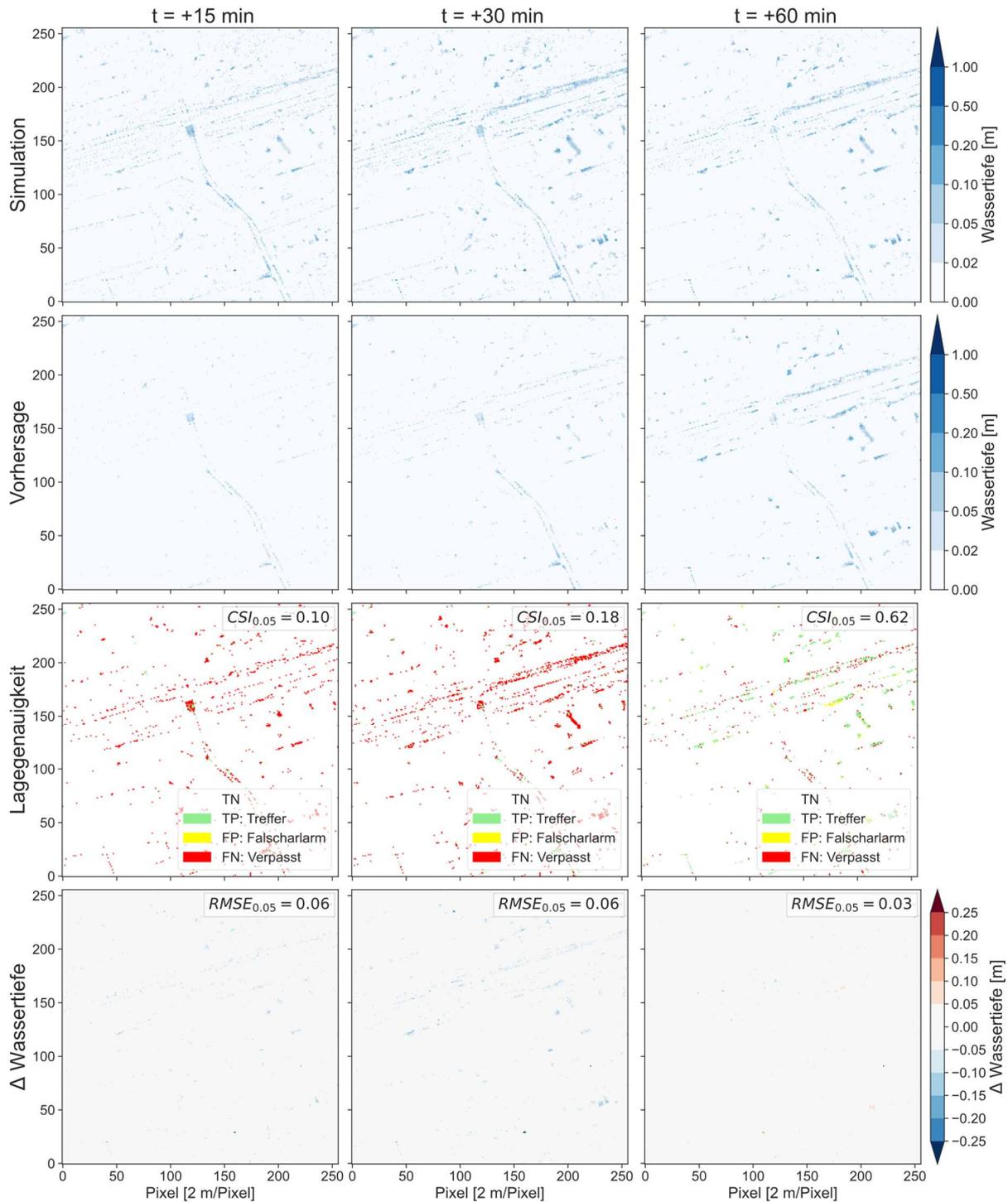
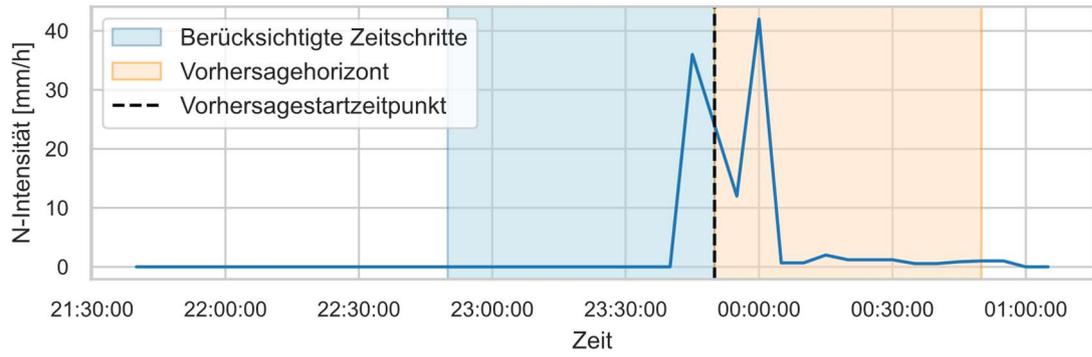
Tabelle A4-1: Übersicht über die aus dem Testdatensatz ausgewählten historischen Ereignisse

| ID | Ereignisdauer | | | Ereignis Summe | T [a] für Dauerstufe [min] | | | | |
|-----|------------------|------------------|-------|-------------------|----------------------------|------|------|------|-----|
| | Von | Bis | Dauer | | 15 | 30 | 60 | 120 | 240 |
| 81 | 04.06.1947 23:38 | 05.06.1947 03:28 | 03:50 | 20,3 mm | 5,7 | | | | |
| 82 | 09.06.1949 00:05 | 09.06.1949 04:05 | 04:00 | 69,6 mm | 10,7 | 48 | 252 | 391 | 231 |
| 89 | 29.03.1974 16:22 | 29.03.1974 20:22 | 04:00 | 41,3 mm | 8,3 | 36 | 28 | 14,4 | 8,7 |
| 100 | 03.07.2009 17:41 | 03.07.2009 20:01 | 02:20 | 54,1 mm | 216 | 905 | 248 | 100 | 39 |
| 102 | 03.07.2010 15:22 | 03.07.2010 17:47 | 02:25 | 81,2 mm | 158 | 1309 | 6061 | 3324 | 891 |

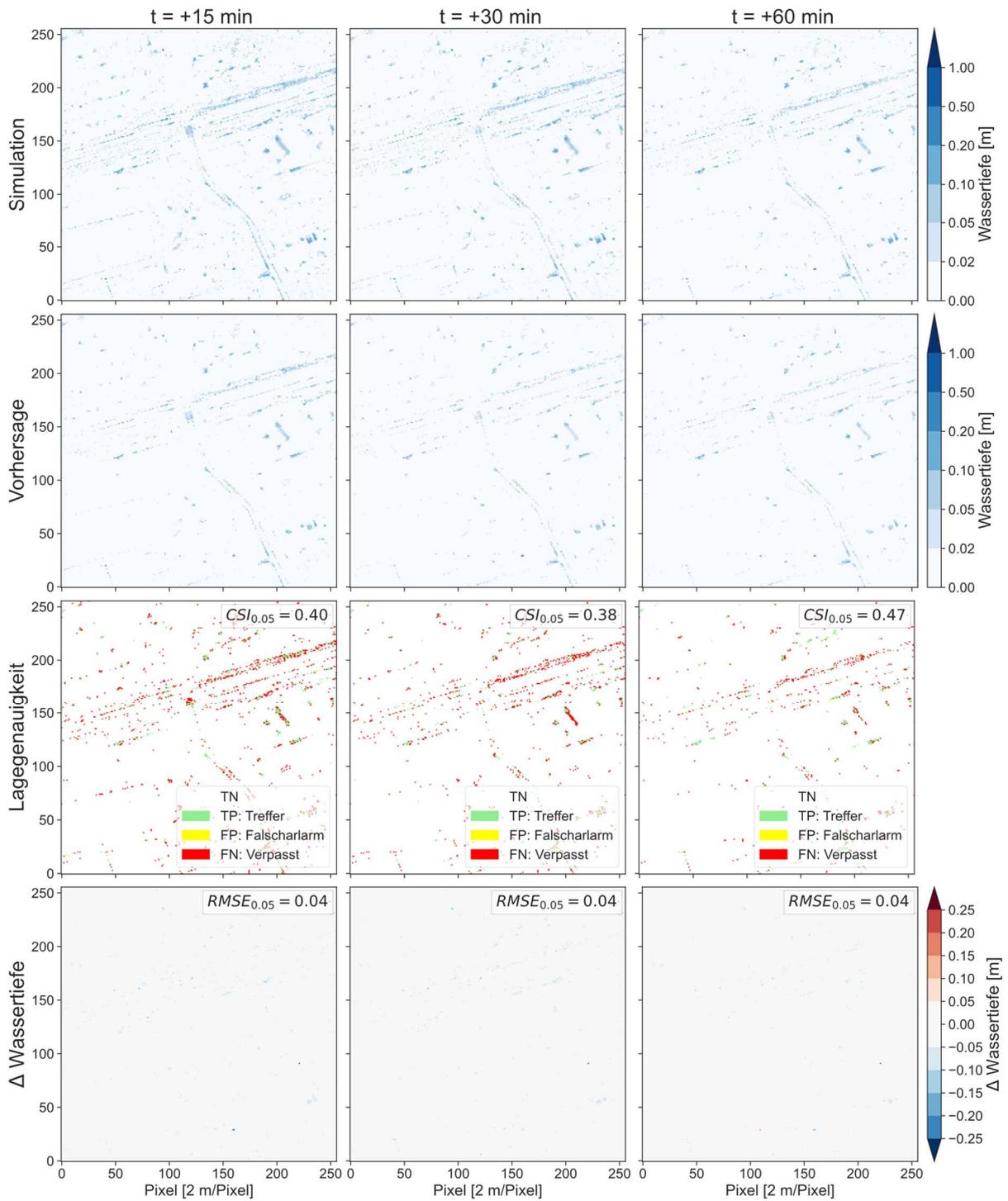
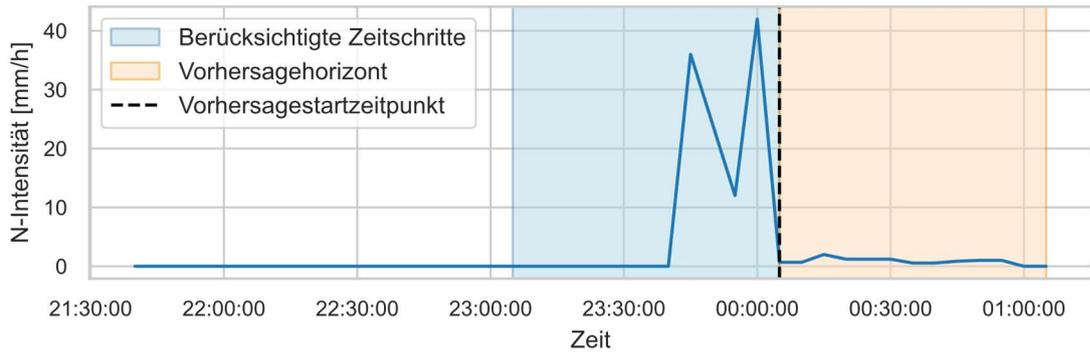
A 4-1: Ereignis-ID 81: Prognose zum Ereignisbeginn



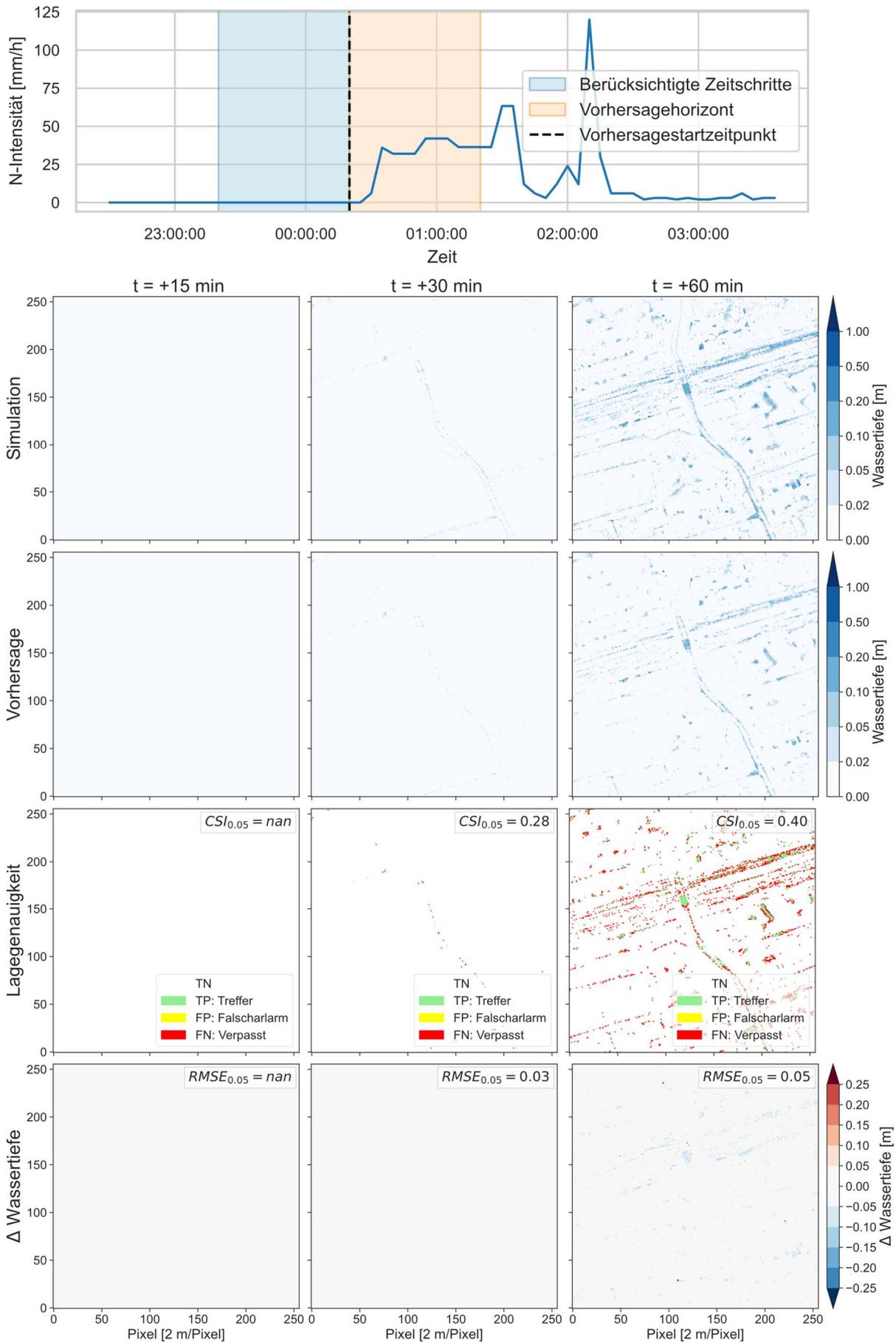
A 4-2: Ereignis-ID 81: Prognose zur Ereignismitte



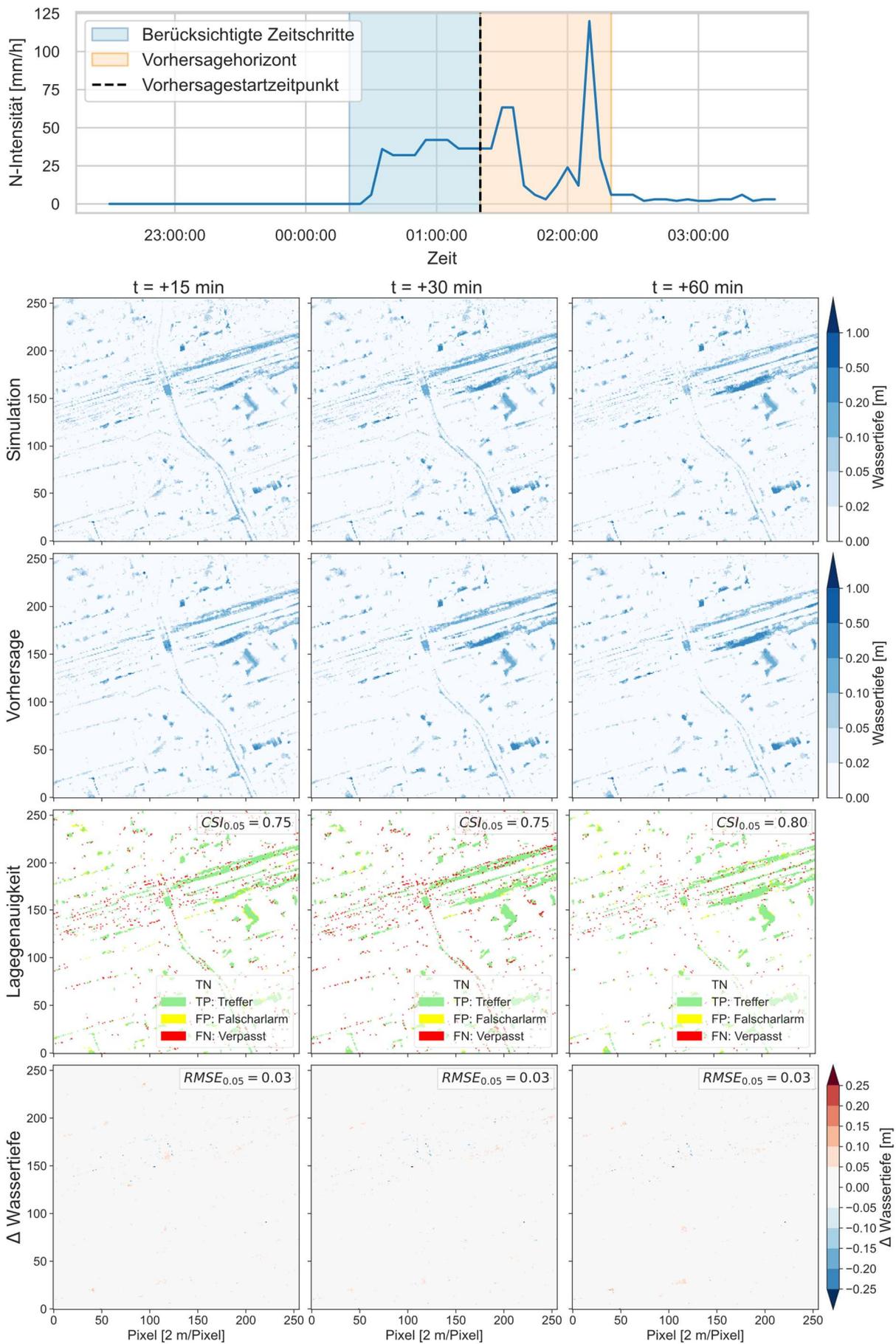
A 4-3: Ereignis-ID 81: Prognose zum Ereignisende



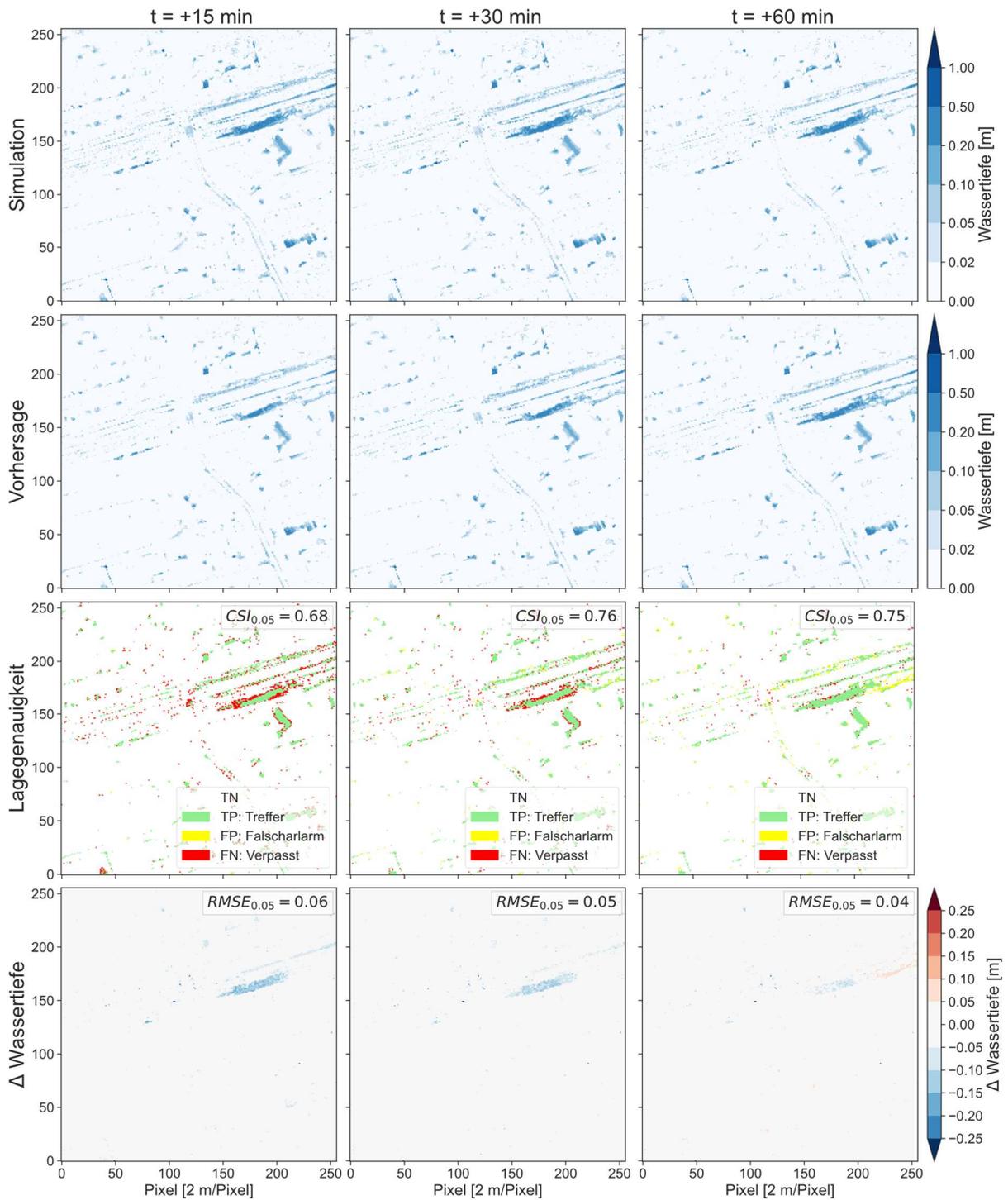
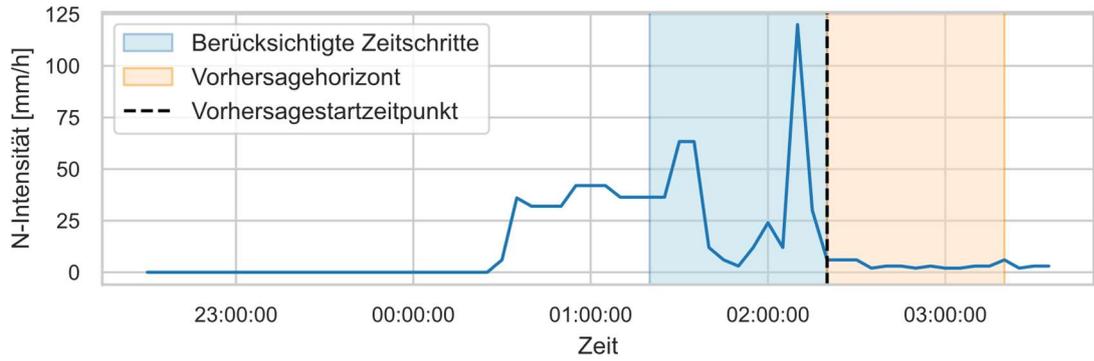
A 4-4: Ereignis-ID 82: Prognose zum Ereignisbeginn



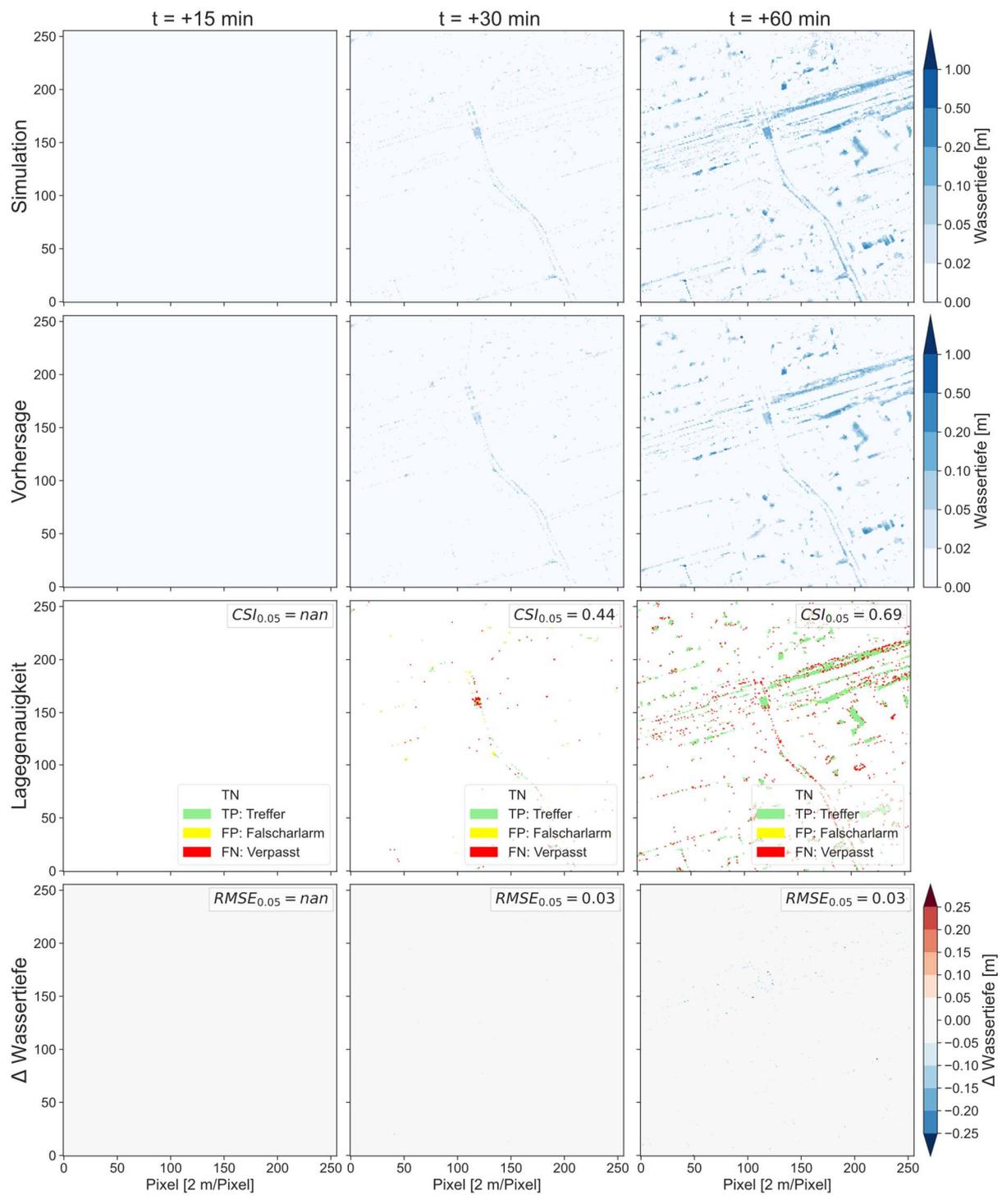
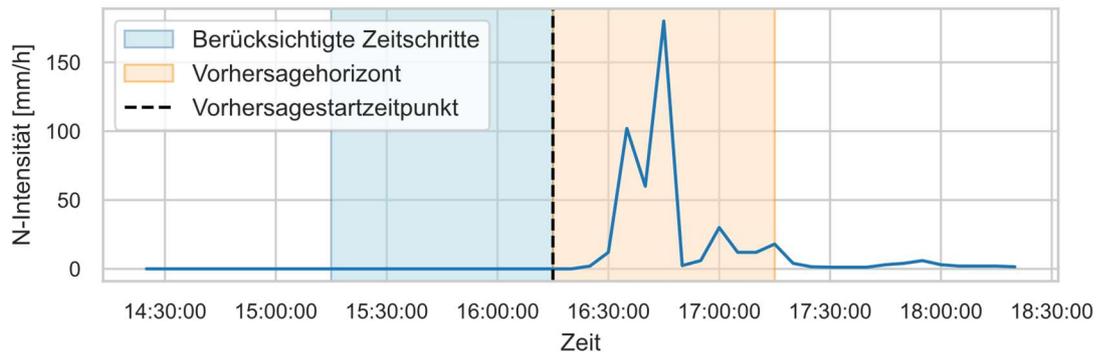
A 4-5: Ereignis-ID 82: Prognose zur Ereignismitte



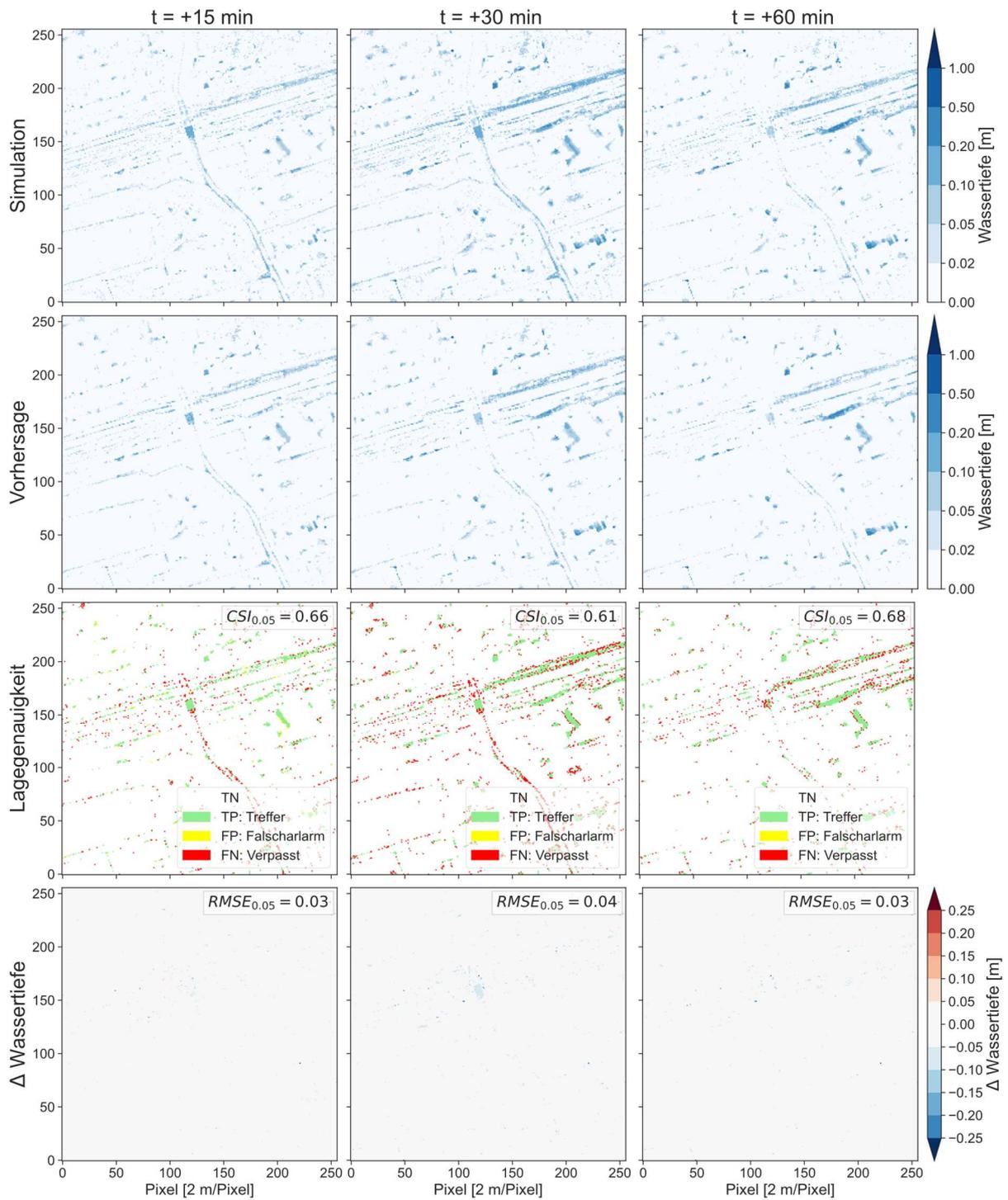
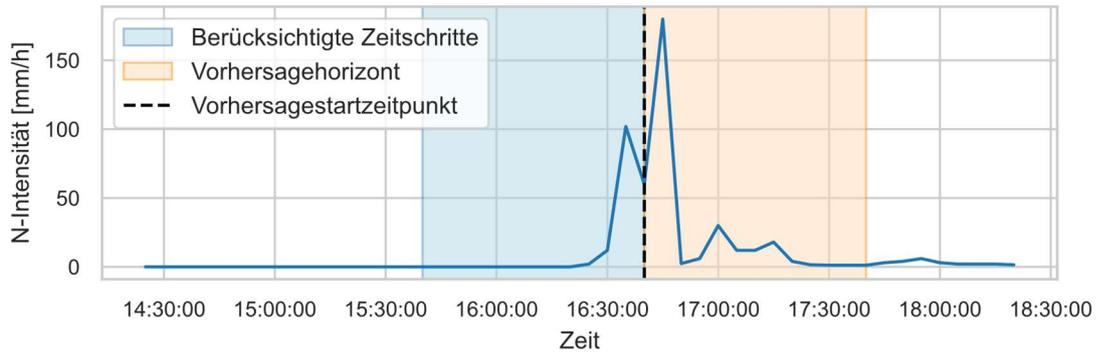
A 4-6: Ereignis-ID 82: Prognose zum Ereignisende



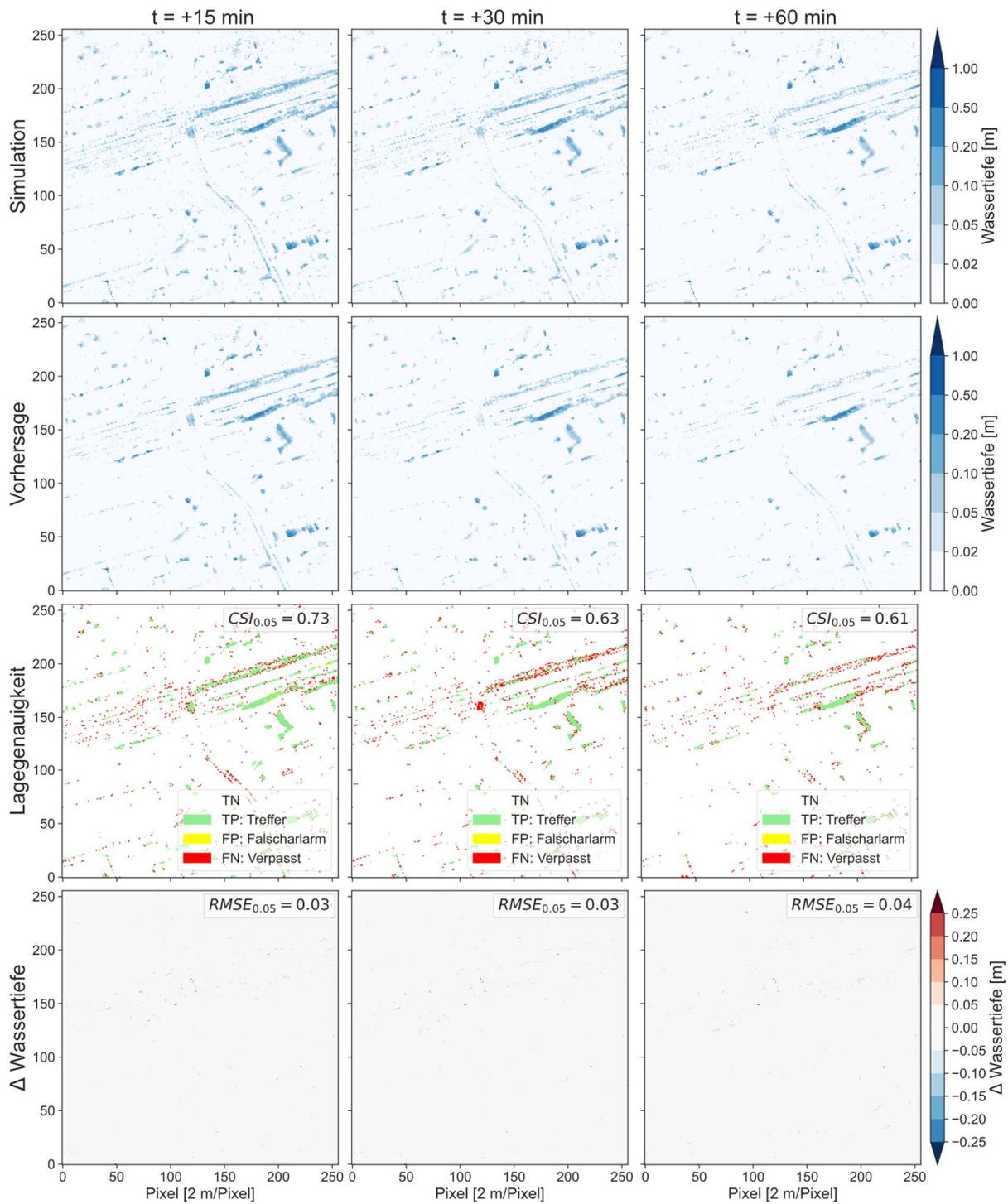
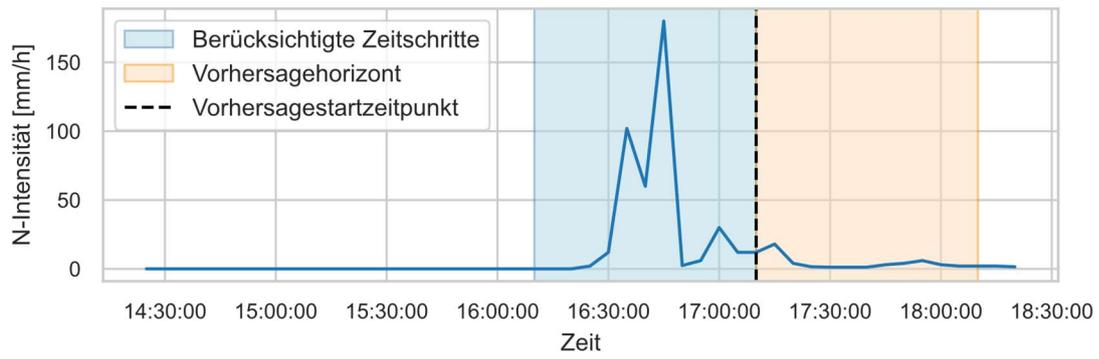
A 4-7: Ereignis-ID 89: Prognose zum Ereignisbeginn



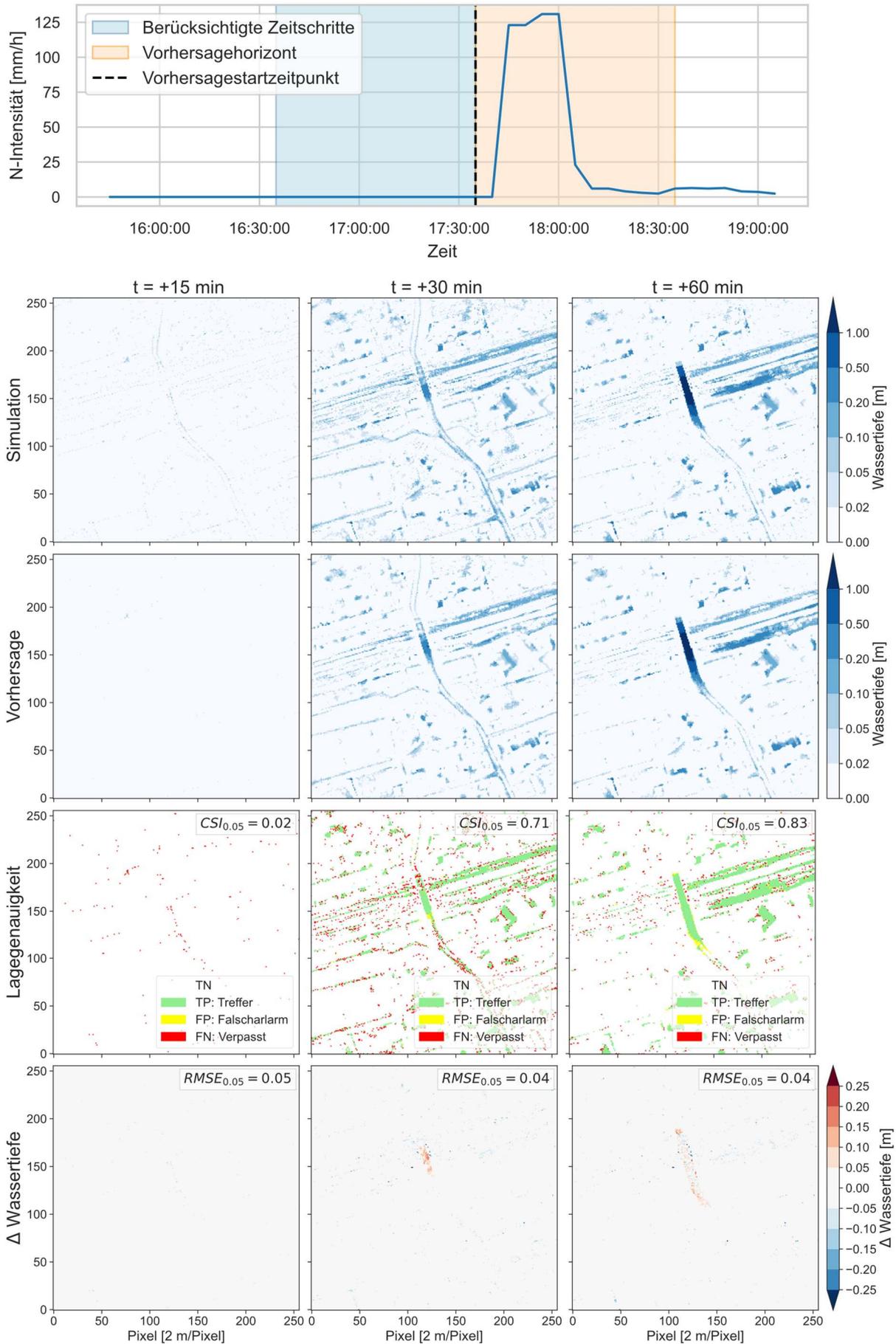
A 4-8: Ereignis-ID 89: Prognose zur Ereignismitte



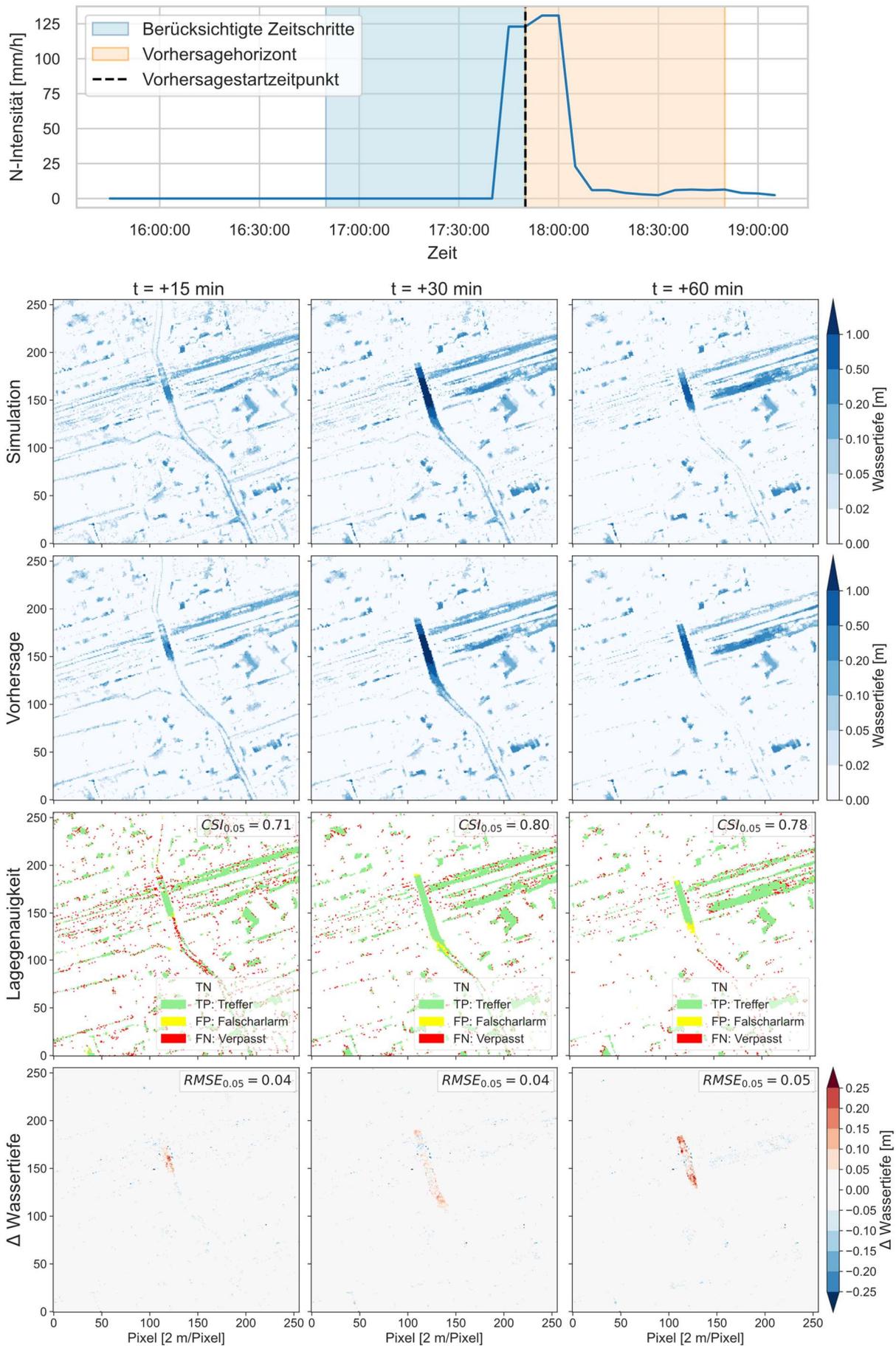
A 4-9: Ereignis-ID 89: Prognose zum Ereignisende



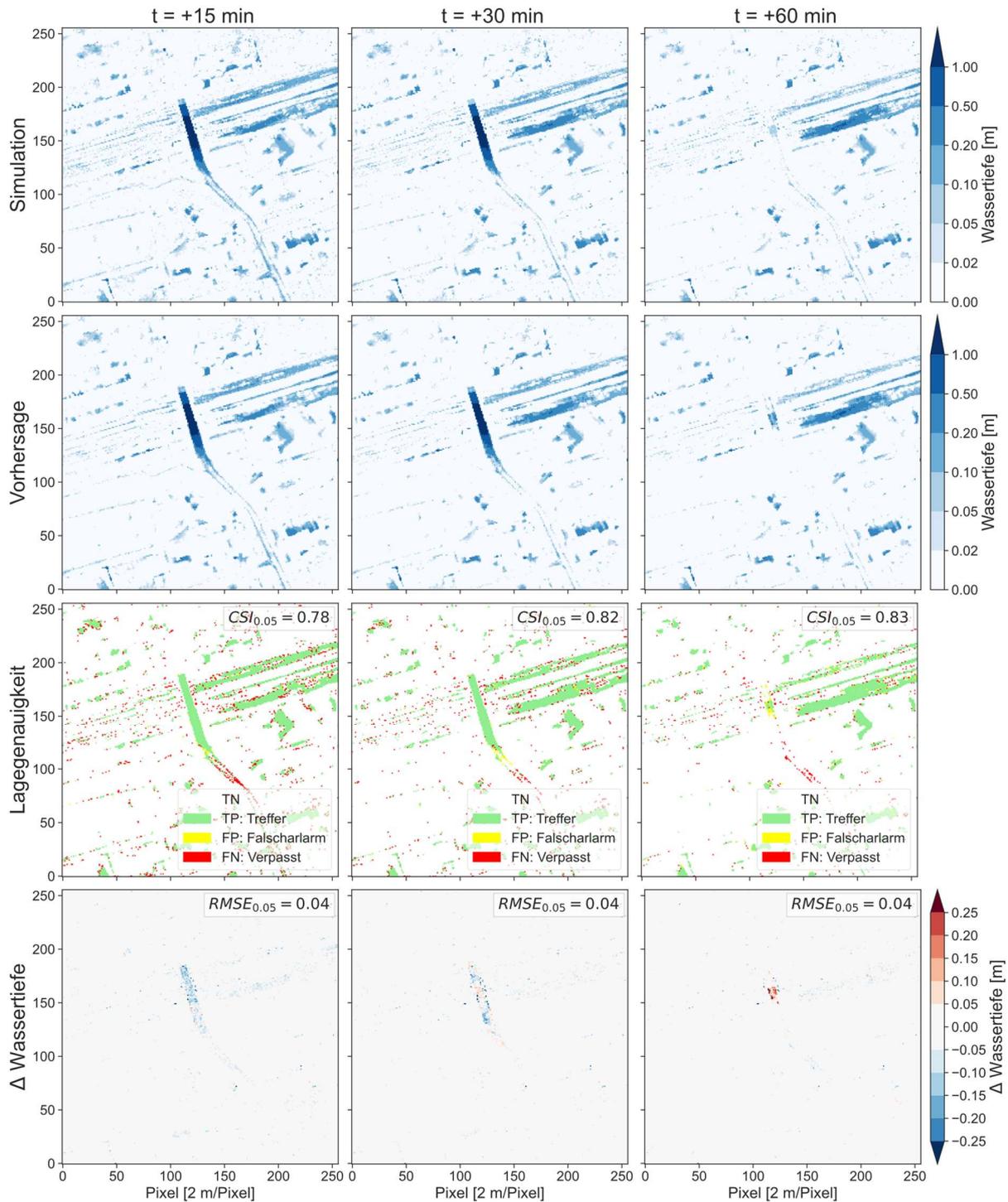
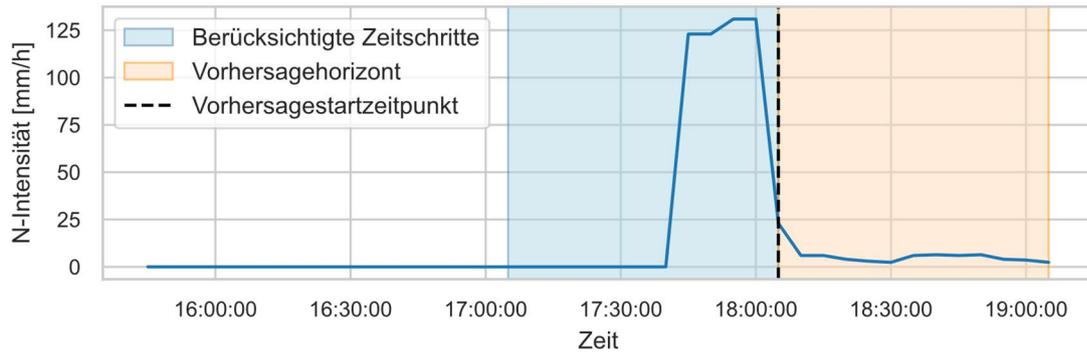
A 4-10: Ereignis-ID 100: Prognose zum Ereignisbeginn



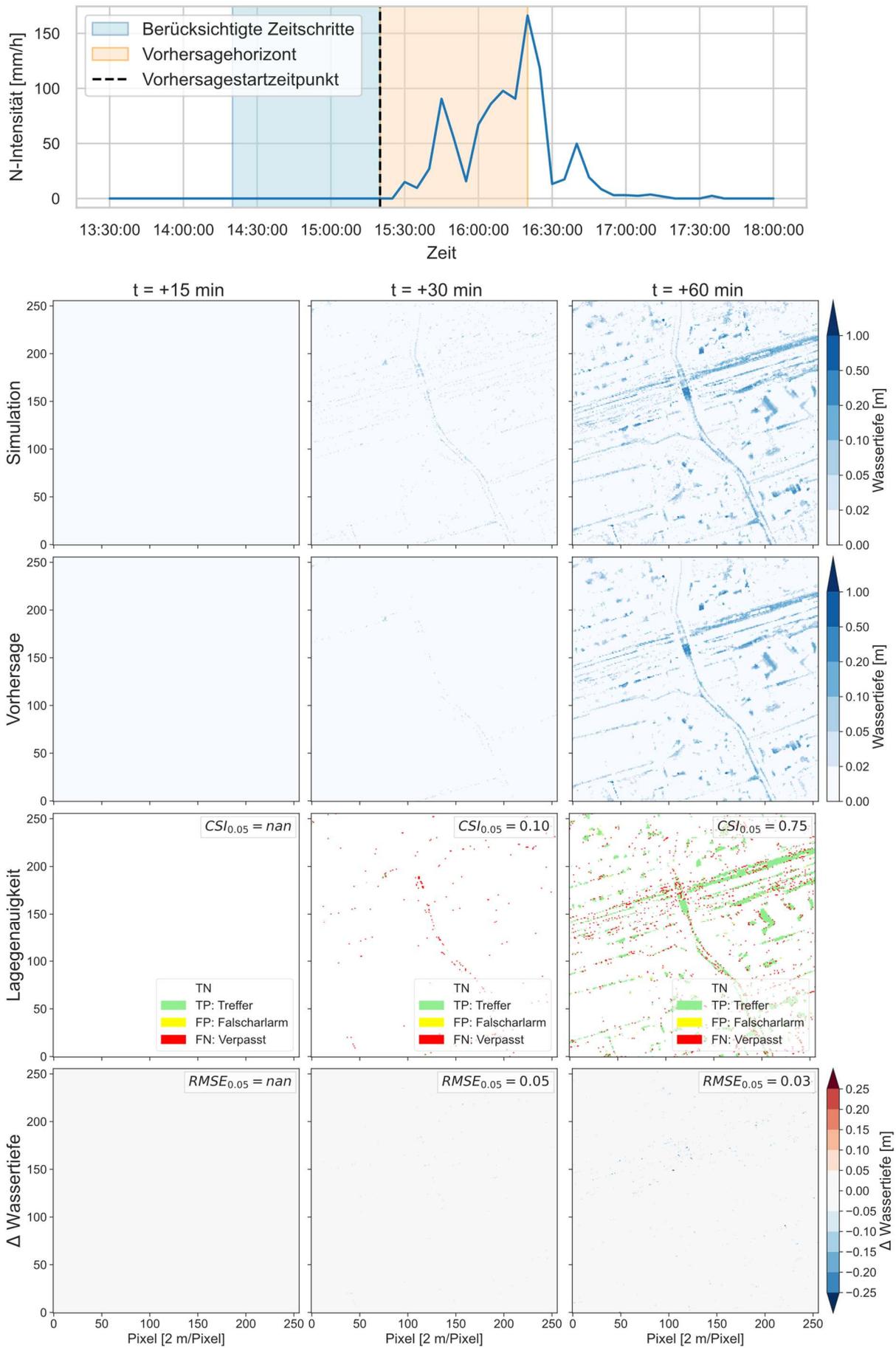
A 4-11: Ereignis-ID 100: Prognose zur Ereignismitte



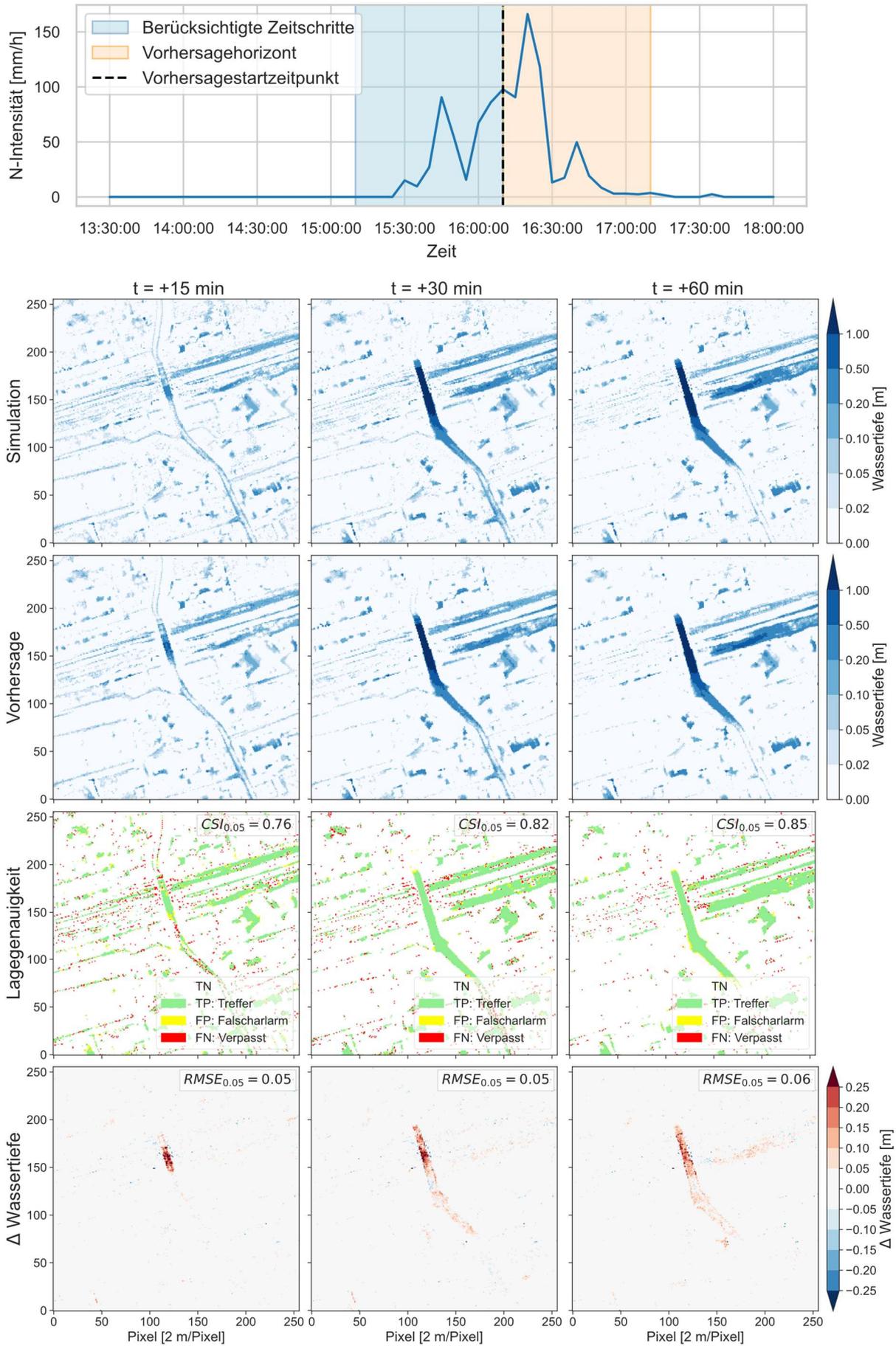
A 4-12: Ereignis-ID 100: Prognose zum Ereignisende



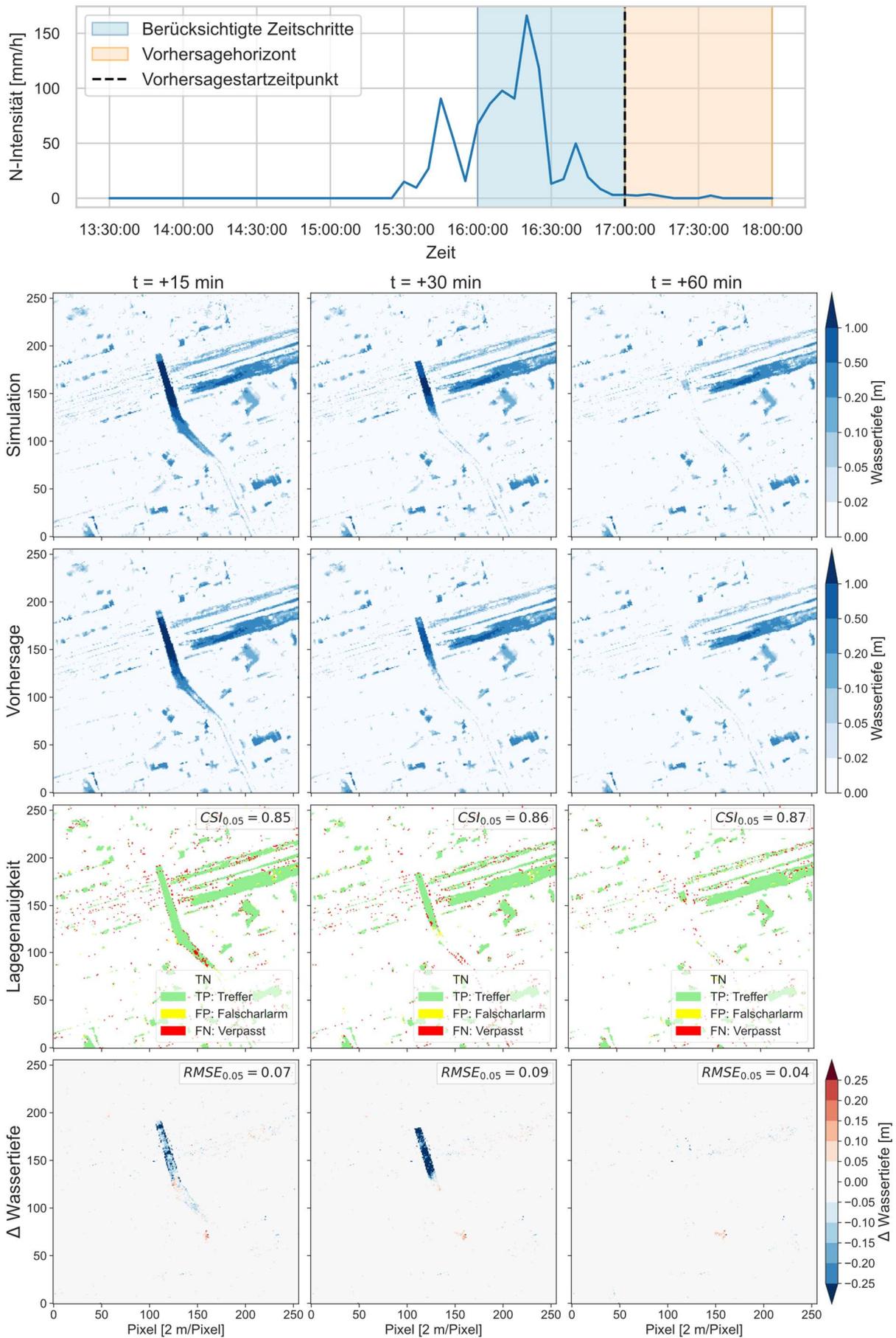
A 4-13: Ereignis-ID 102: Prognose zum Ereignisbeginn



A 4-14: Ereignis-ID 102: Prognose zur Ereignismitte



A 4-15: Ereignis-ID 102: Prognose zum Ereignisende



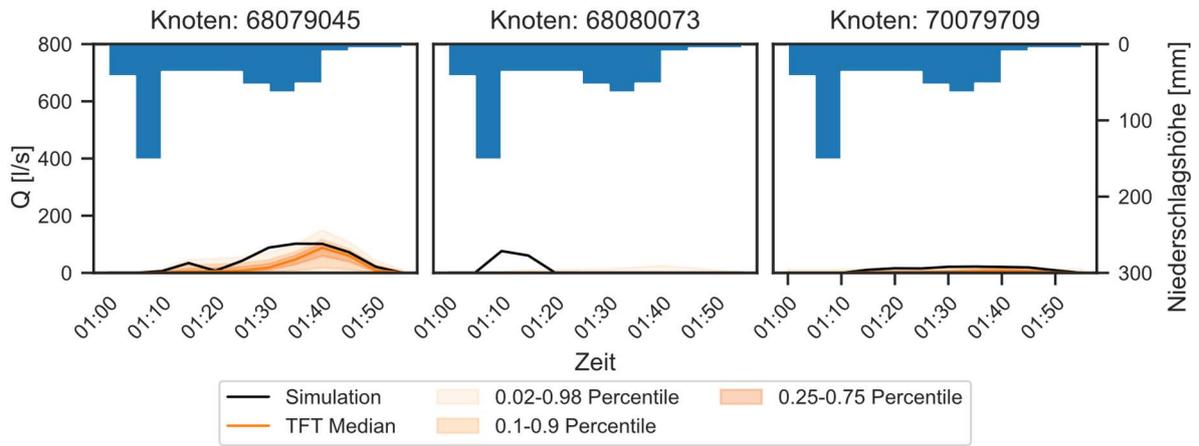
A-5: Vorhersageergebnisse für historische Ereignisse: Überstauvorhersage

Nachfolgend werden für die in Tabelle A5-1 aufgeführten Ereignisse der Station Gelsenkirchen-Altstadt (Stations-Nr. 2582) mit dem TFT Vorhersageergebnisse für alle Schächte im Einzugsgebiet mit mindestens 100 l Überstau dargestellt. Bei den aufgeführten Ereignissen handelt es sich um alle Ereignisse im Testdatensatz, die zu größeren Überstaumengen führten. Bei der Wahl der Prognosezeitpunkte wurde jeweils darauf geachtet, dass der Prognosezeitraum möglichst den Zeitraum eines Ereignisses abdeckt, an dem es zu Überstau kommt. Auf eine Darstellung der Prognoseergebnisse zu Ereignisbeginn, -mitte und -ende wurde verzichtet, da sich die Form der Überstauganglinien nur geringfügig ändert. Der Hintergrund hierfür ist, dass im Rahmen der Untersuchungen immer die Niederschlagsmessungen als Prognose angenommen wurden und damit keine zusätzlichen Unsicherheiten in Abhängigkeit des Prognosehorizonts hinzukommen.

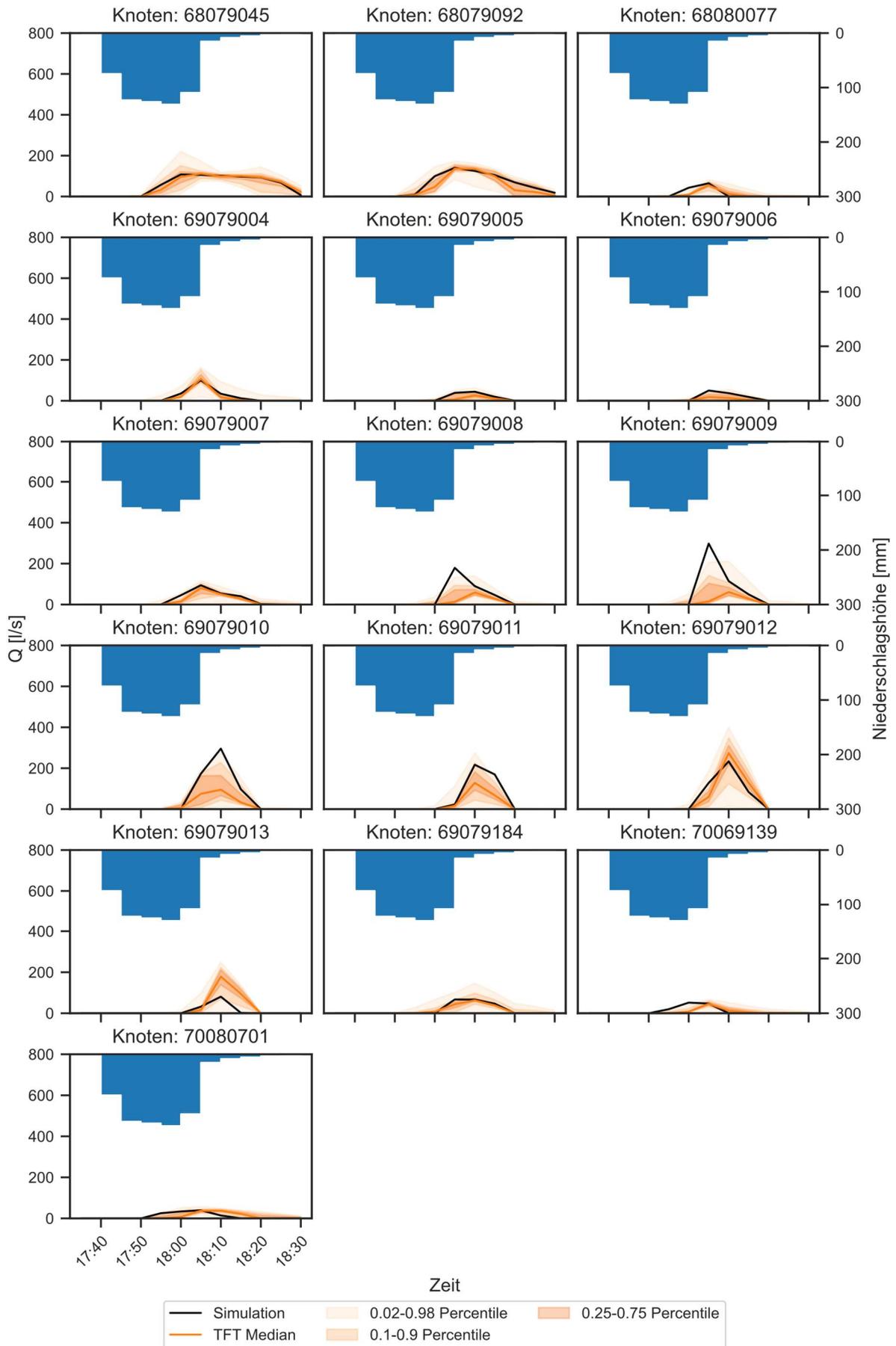
Tabelle A5-1: Übersicht über die aus dem Testdatensatz ausgewählten historischen Ereignisse

| ID | Ereignisdauer | | | Ereignis Summe | T [a] für Dauerstufe [min] | | | | |
|-----|------------------|------------------|-------|-------------------|----------------------------|------|------|------|-----|
| | Von | Bis | Dauer | | 15 | 30 | 60 | 120 | 240 |
| 82 | 09.06.1949 00:05 | 09.06.1949 04:05 | 04:00 | 69,6 mm | 10,7 | 48 | 252 | 391 | 231 |
| 100 | 03.07.2009 17:41 | 03.07.2009 20:01 | 02:20 | 54,1 mm | 216 | 905 | 248 | 100 | 39 |
| 102 | 03.07.2010 15:22 | 03.07.2010 17:47 | 02:25 | 81,2 mm | 158 | 1309 | 6061 | 3324 | 891 |

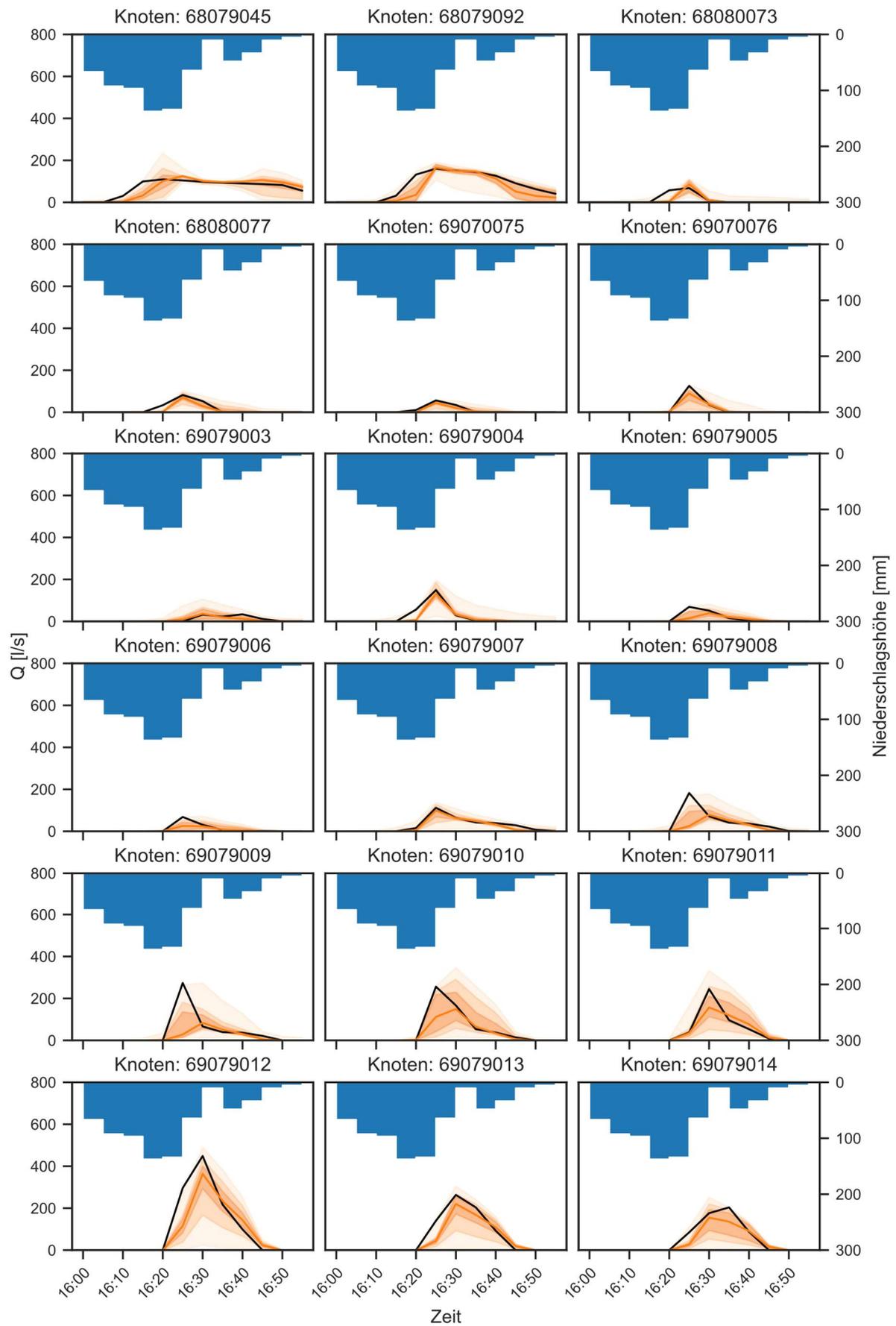
A 5-1: Ereignis-ID 82

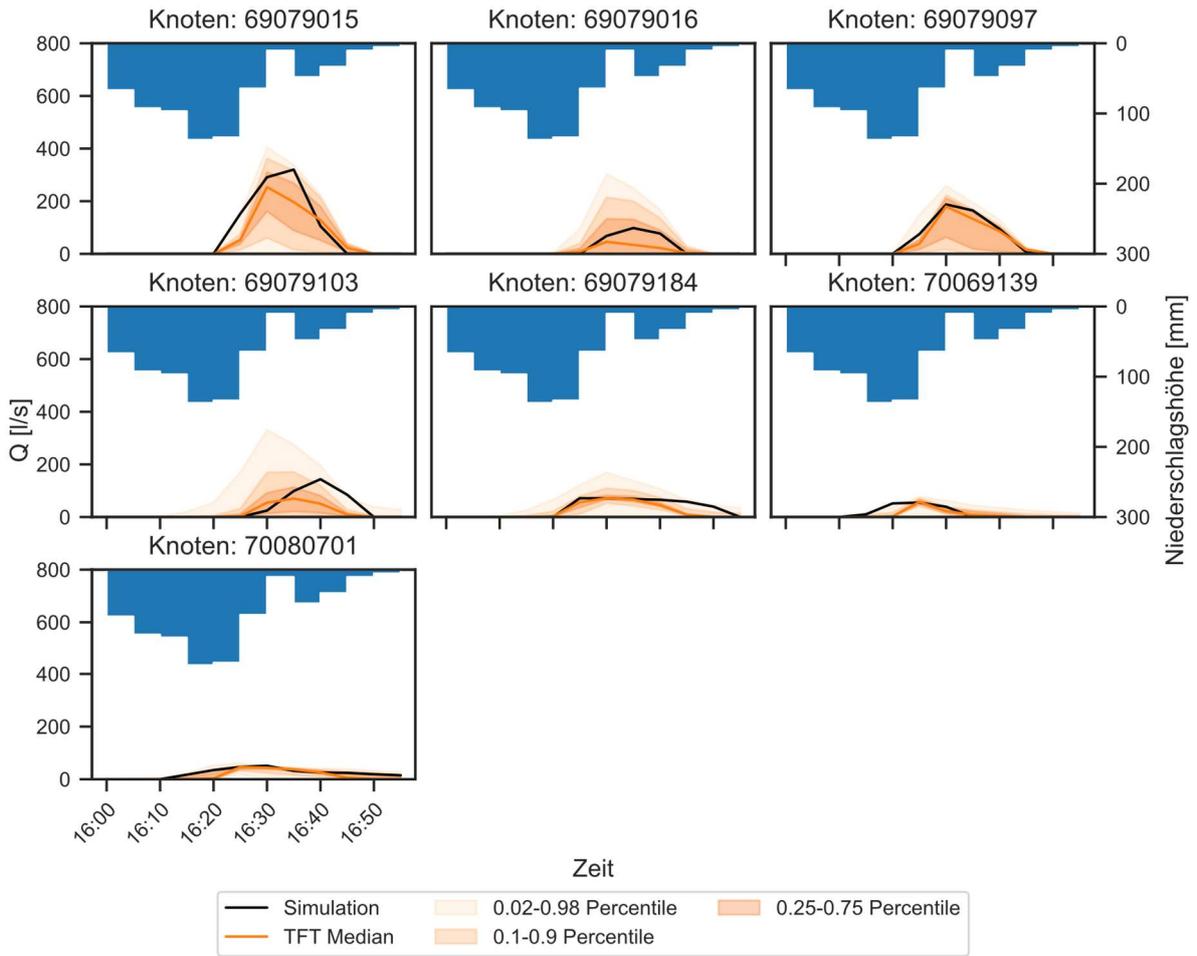


A 5-2: Ereignis ID 100



A 5-3: Ereignis ID 102





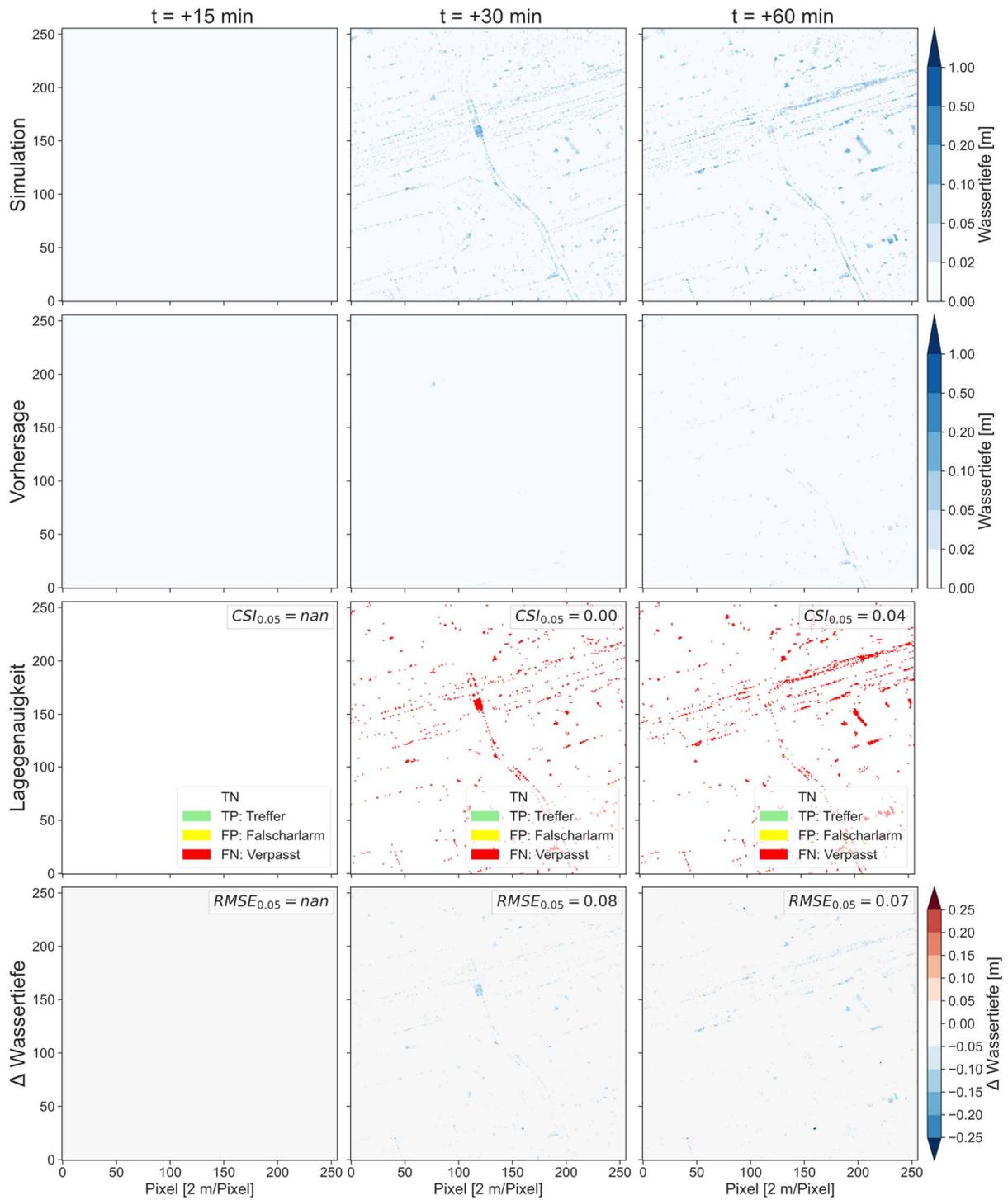
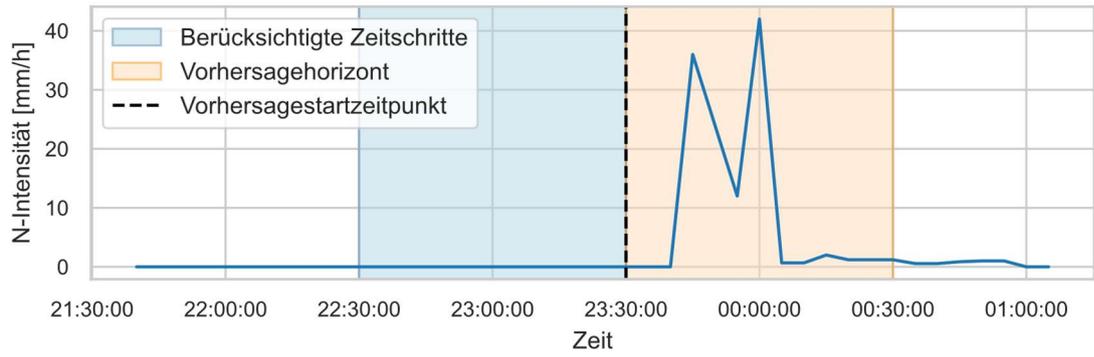
A-6: Vorhersageergebnisse für historische Ereignisse: Gekoppelte Betrachtung

Analog zur Darstellung der Vorhersageergebnisse für die Überflutungsflächenvorhersage (vgl. Anhang 4) werden nachfolgend für die in Tabelle A6-1 aufgeführten Ereignisse der Station Gelsenkirchen-Altstadt (Stations-Nr. 2582) Vorhersageergebnisse für Vorhersagen zu Beginn, Mitte und Ende des jeweiligen Ereignisses dargestellt. Bei der Auswahl der Ereignisse aus dem Testdatensatz wurden die drei Ereignisse mit den höchsten Wiederkehrzeiten (ID 82, 100, 102), ein Ereignis mit mittlerer Wiederkehrzeit (ID 89) und ein Ereignis mit einer geringen Wiederkehrzeit (ID 81) ausgewählt. Letzteres führte bei der Evaluierung zu den schlechtesten Ergebnissen.

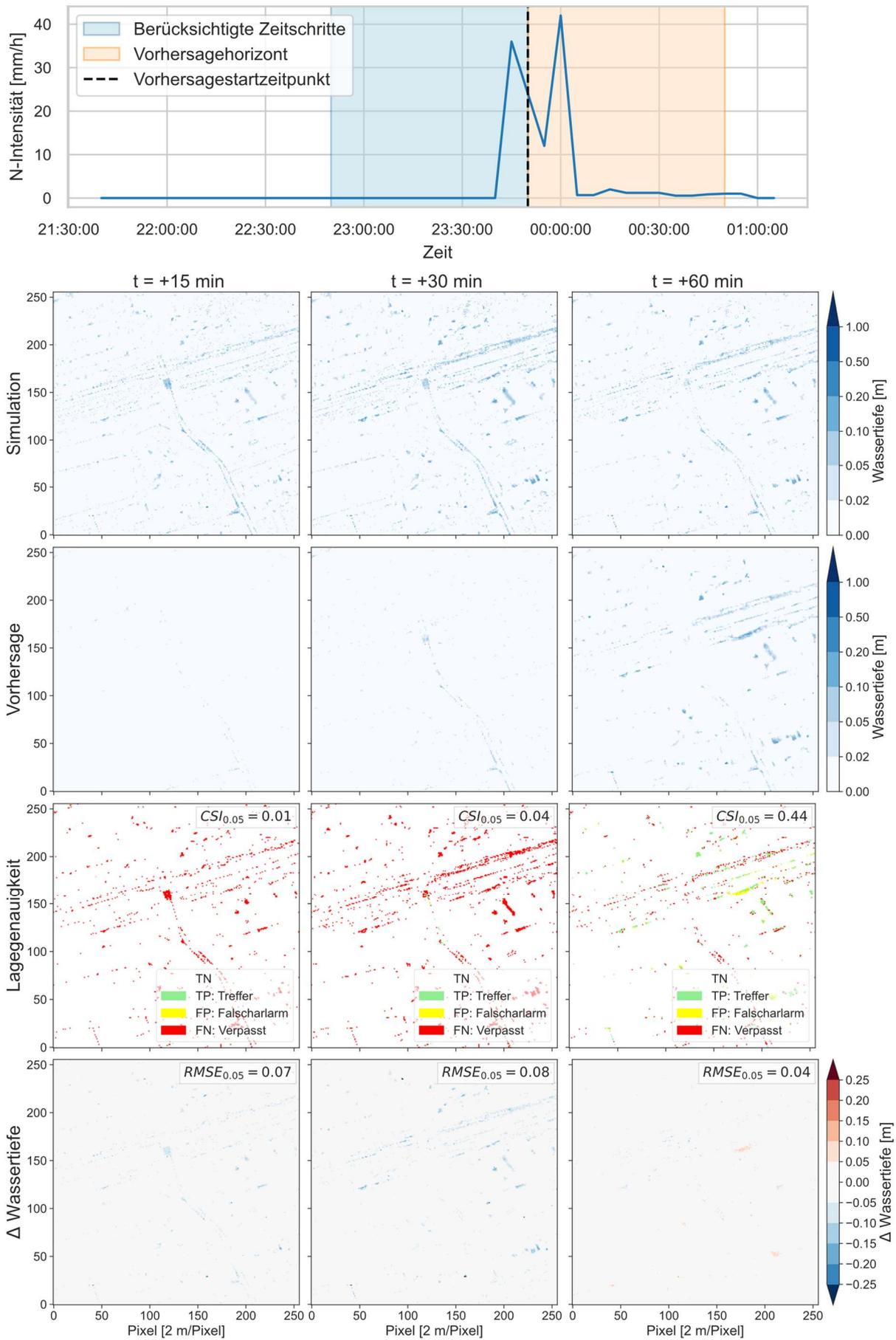
Tabelle A6-1: Übersicht über die aus dem Testdatensatz ausgewählten historischen Ereignisse

| ID | Ereignisdauer | | | Ereignis Summe | T [a] für Dauerstufe [min] | | | | |
|-----|------------------|------------------|-------|-------------------|----------------------------|------|------|------|-----|
| | Von | Bis | Dauer | | 15 | 30 | 60 | 120 | 240 |
| 81 | 04.06.1947 23:38 | 05.06.1947 03:28 | 03:50 | 20,3 mm | 5,7 | | | | |
| 82 | 09.06.1949 00:05 | 09.06.1949 04:05 | 04:00 | 69,6 mm | 10,7 | 48 | 252 | 391 | 231 |
| 89 | 29.03.1974 16:22 | 29.03.1974 20:22 | 04:00 | 41,3 mm | 8,3 | 36 | 28 | 14,4 | 8,7 |
| 100 | 03.07.2009 17:41 | 03.07.2009 20:01 | 02:20 | 54,1 mm | 216 | 905 | 248 | 100 | 39 |
| 102 | 03.07.2010 15:22 | 03.07.2010 17:47 | 02:25 | 81,2 mm | 158 | 1309 | 6061 | 3324 | 891 |

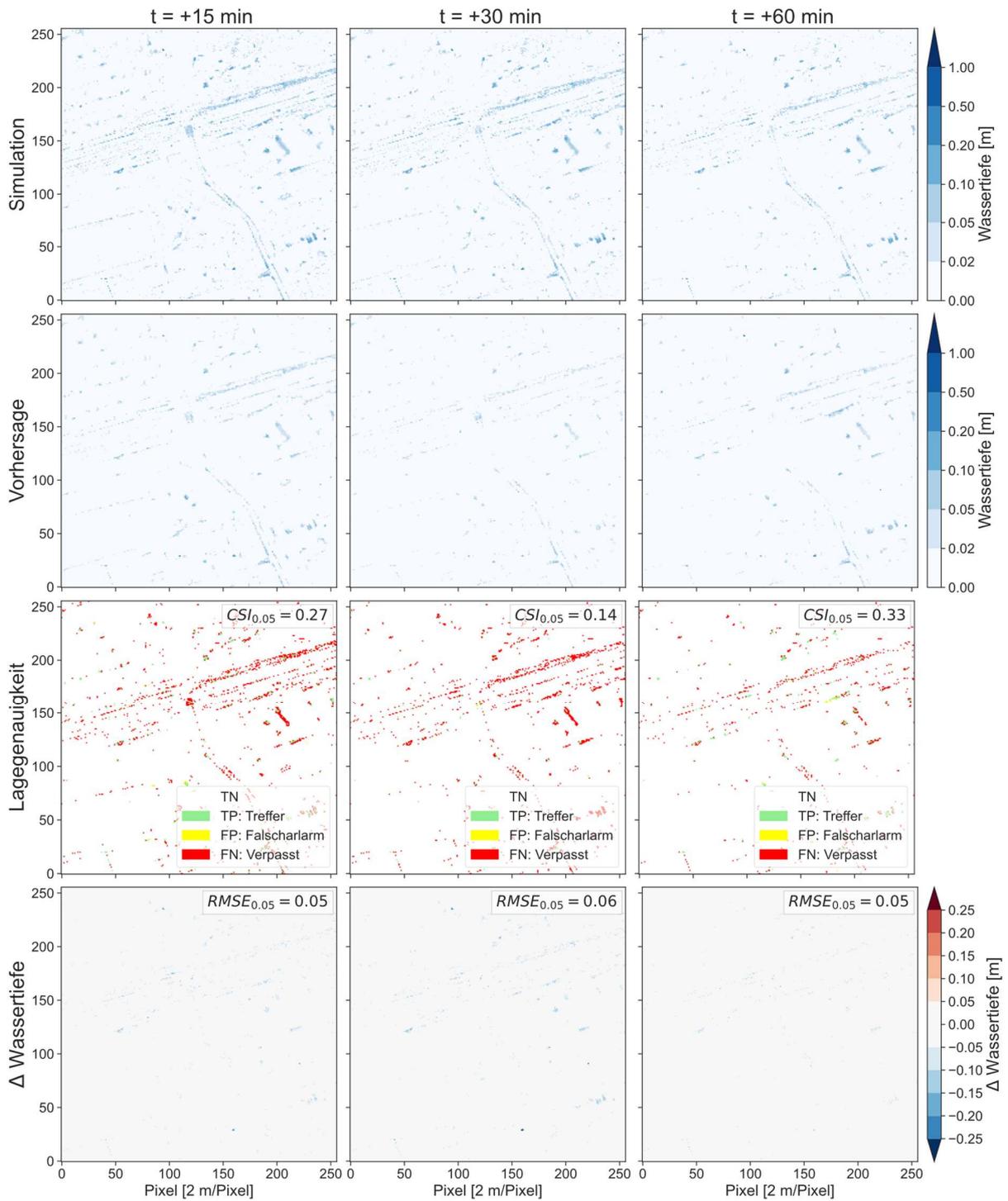
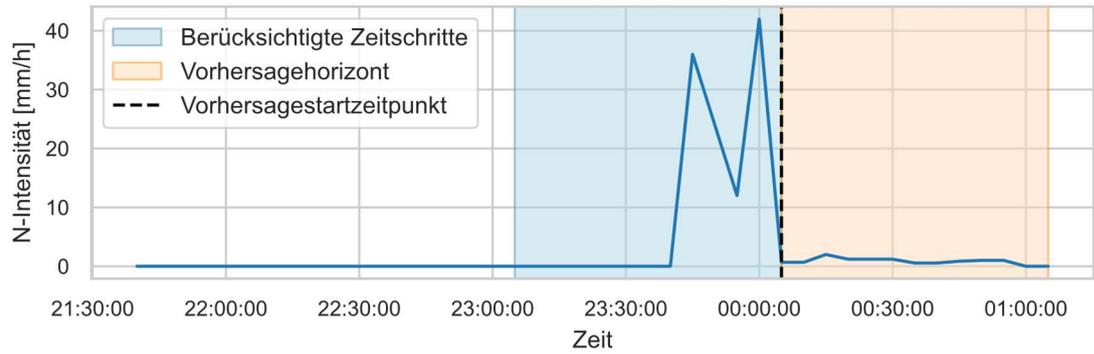
A 6-1: Ereignis ID 81: Prognose zur Ereignisbeginn



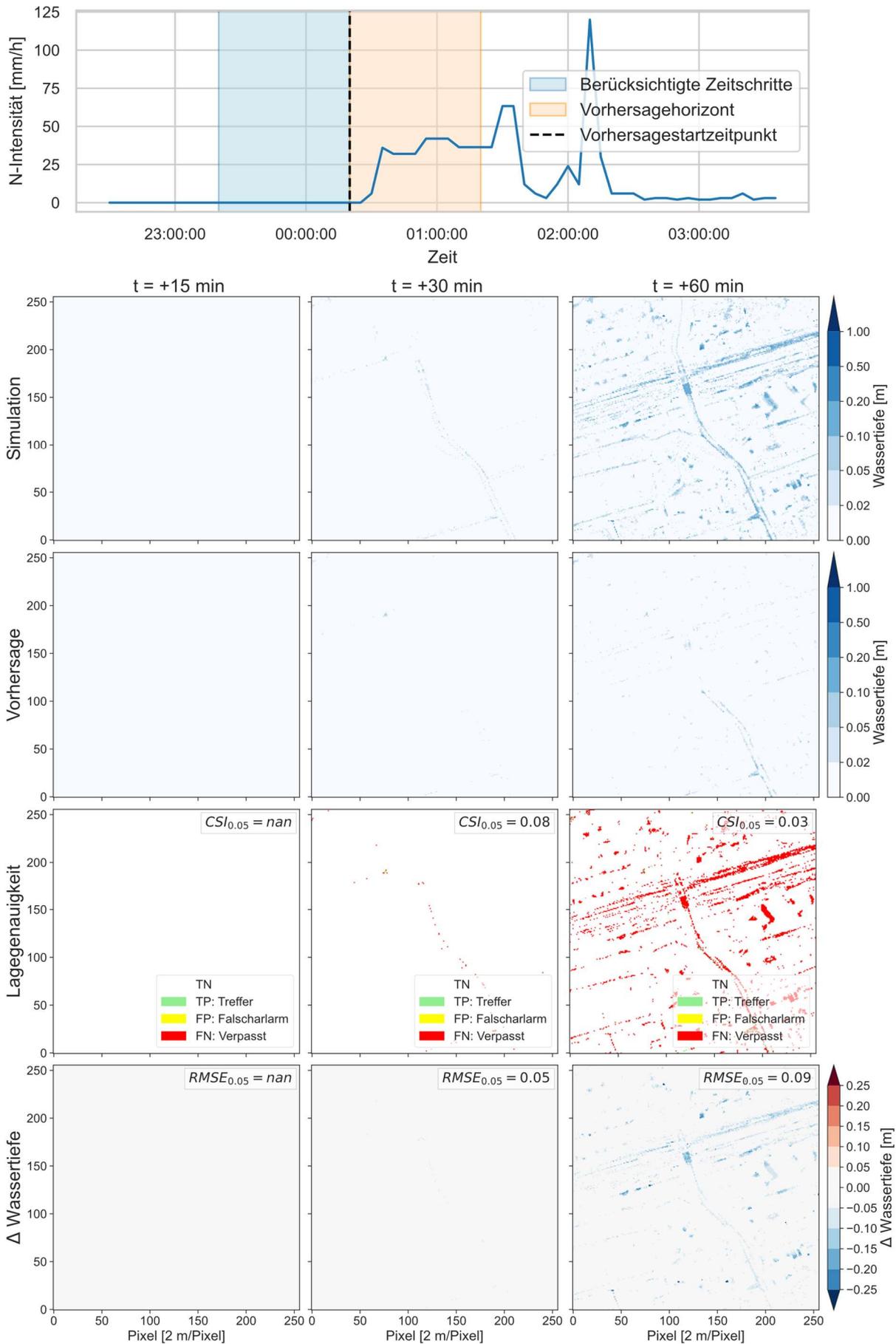
A 6-2: Ereignis ID 81: Prognose zur Ereignismitte



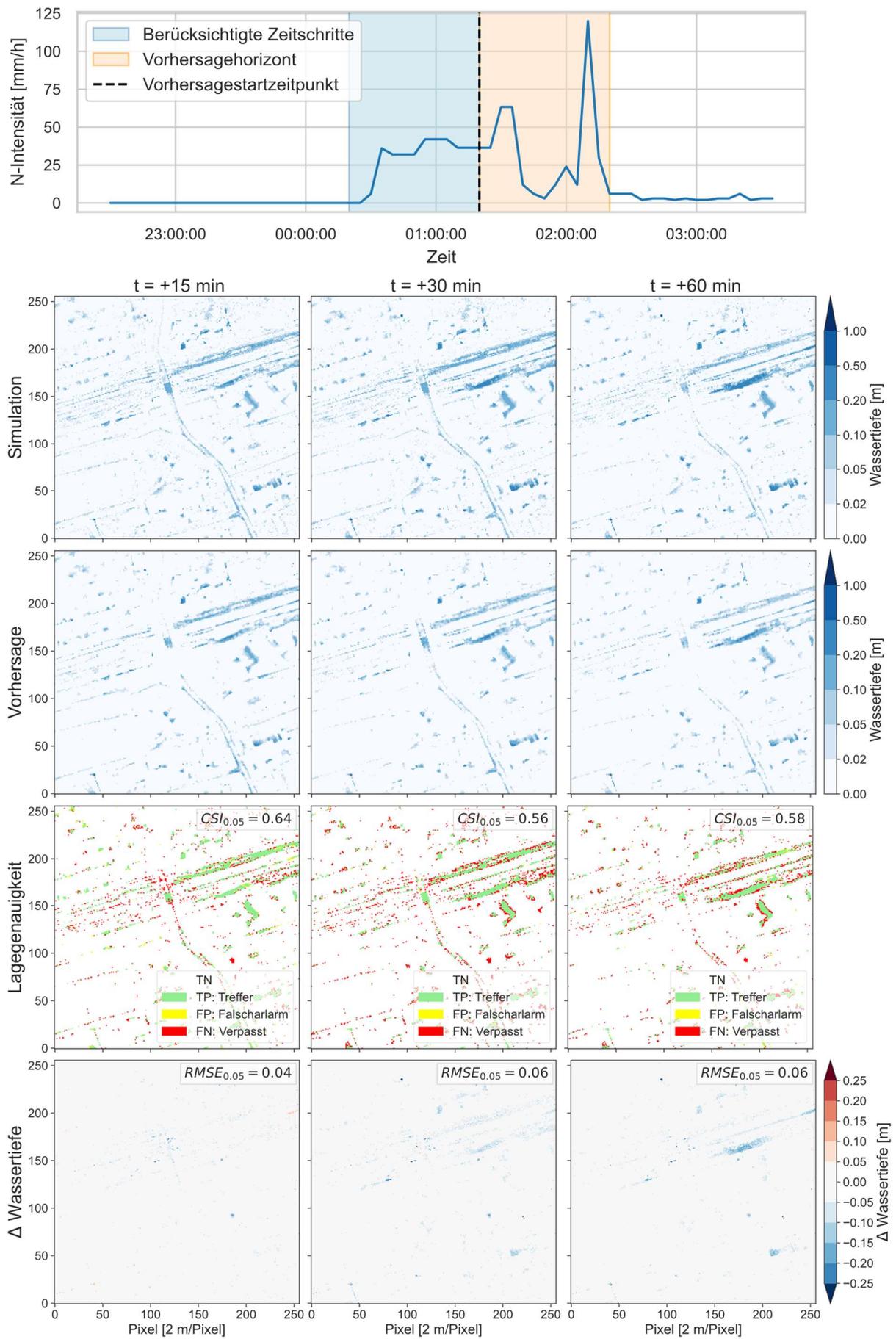
A 6-3: Ereignis ID 81: Prognose zum Ereignisende



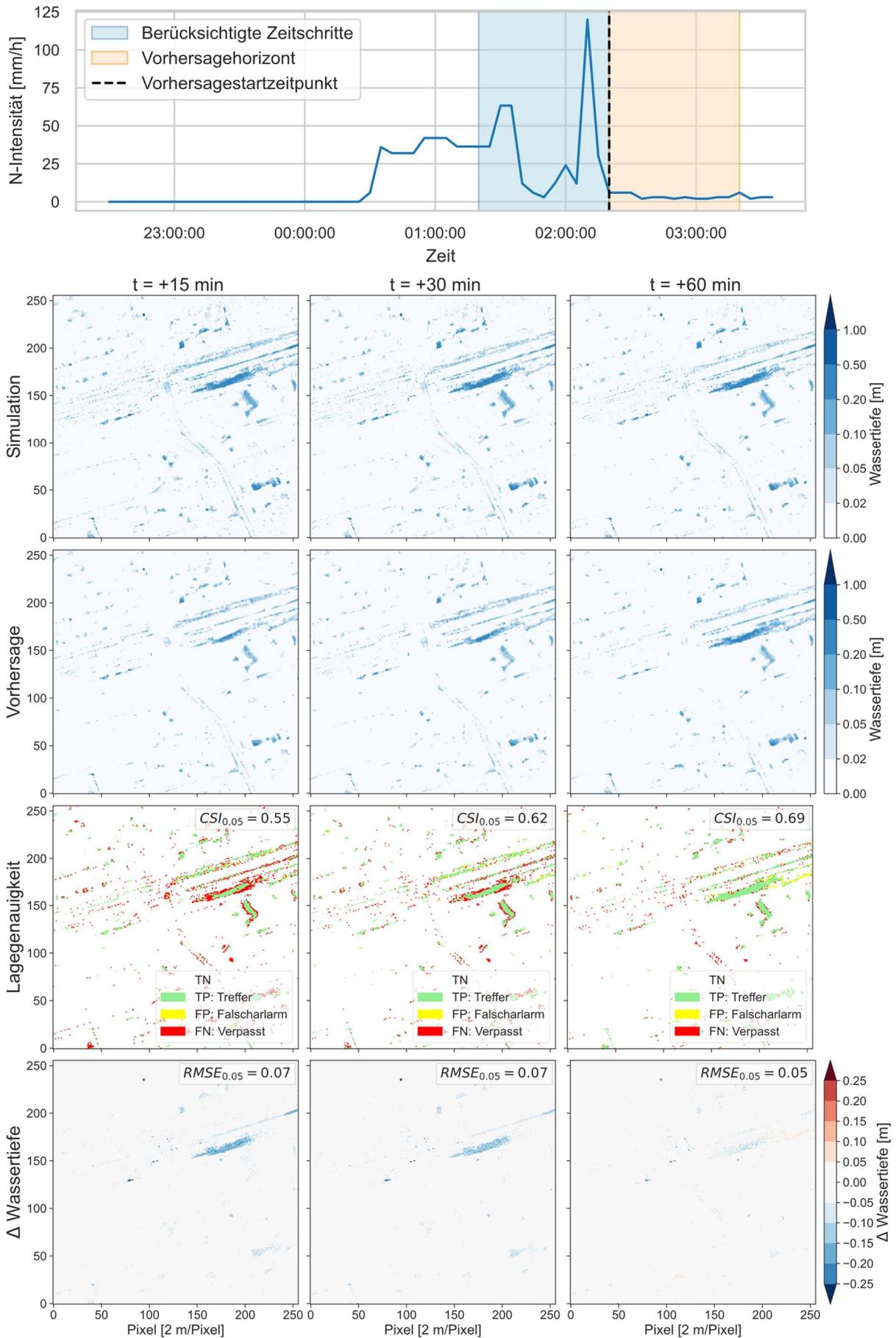
A 6-4: Ereignis ID 82: Prognose zum Ereignisbeginn



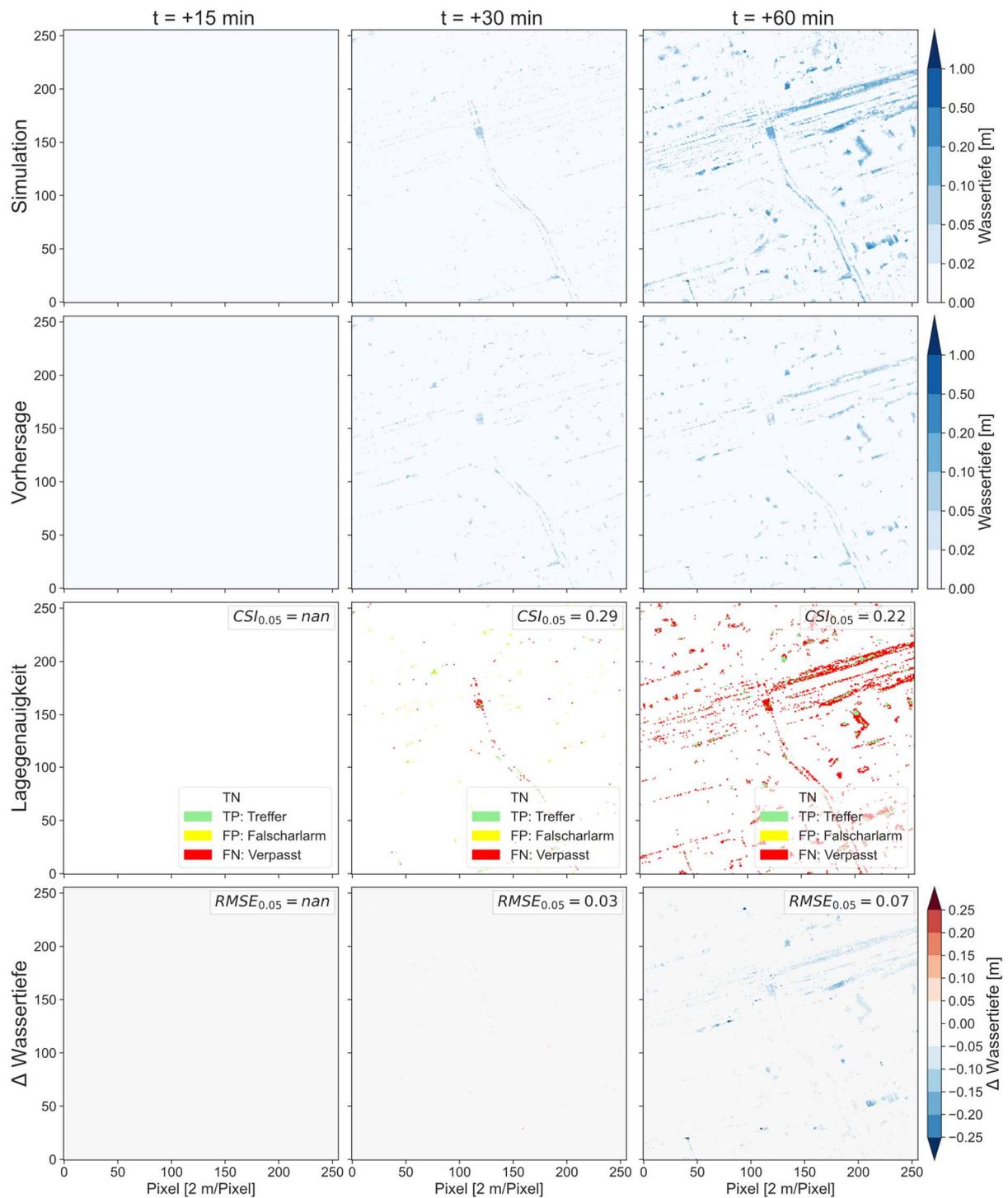
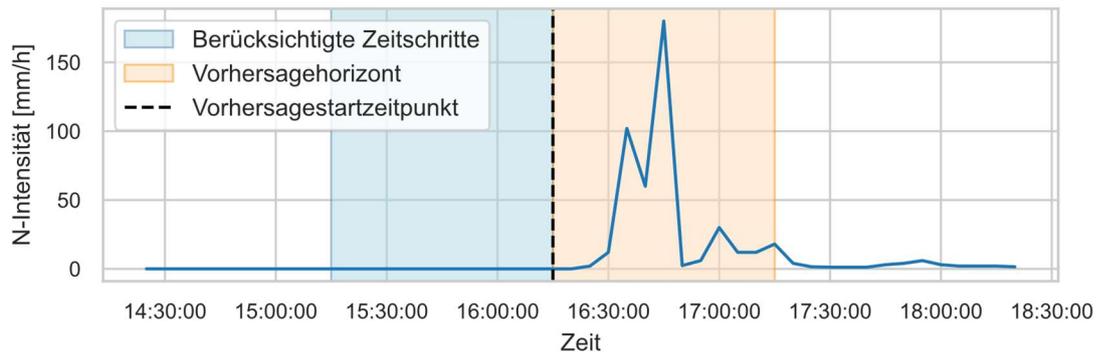
A 6-5: Ereignis ID 82: Prognose zur Ereignismitte



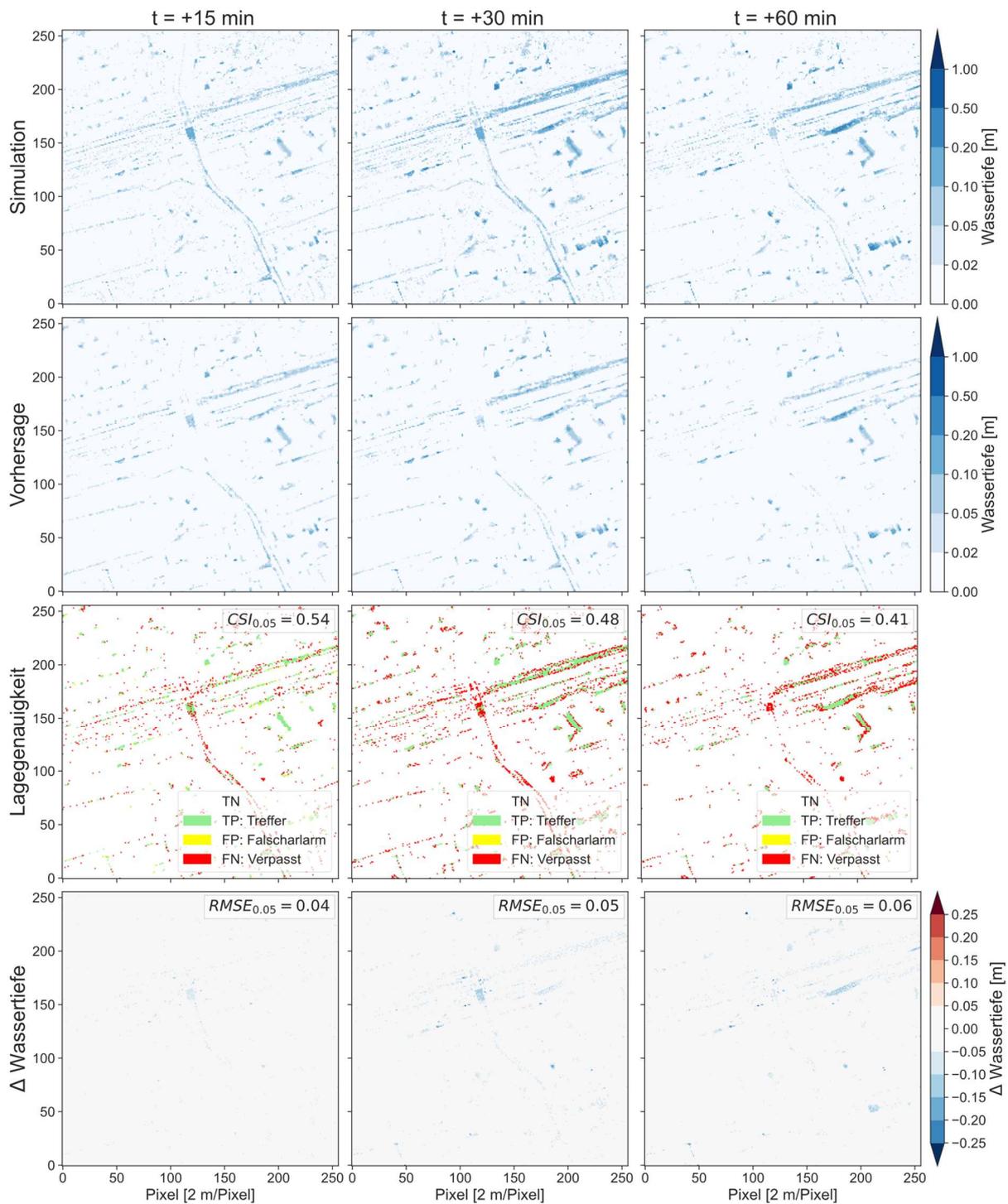
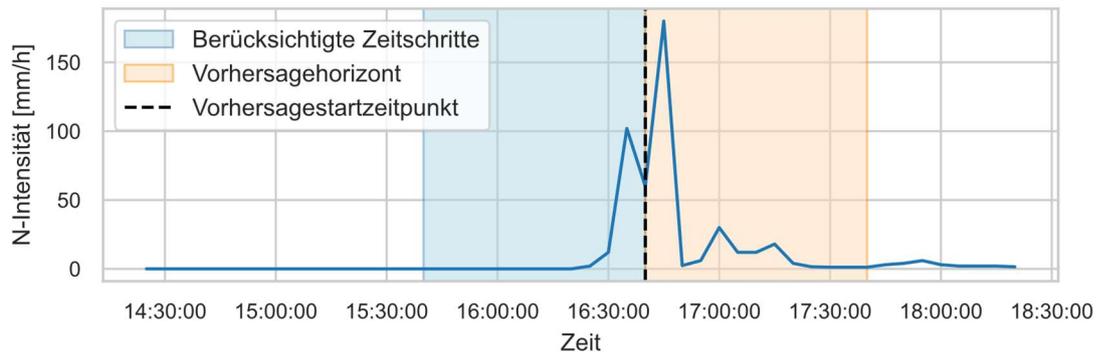
A 6-6: Ereignis ID 82: Prognose zum Ereignisende



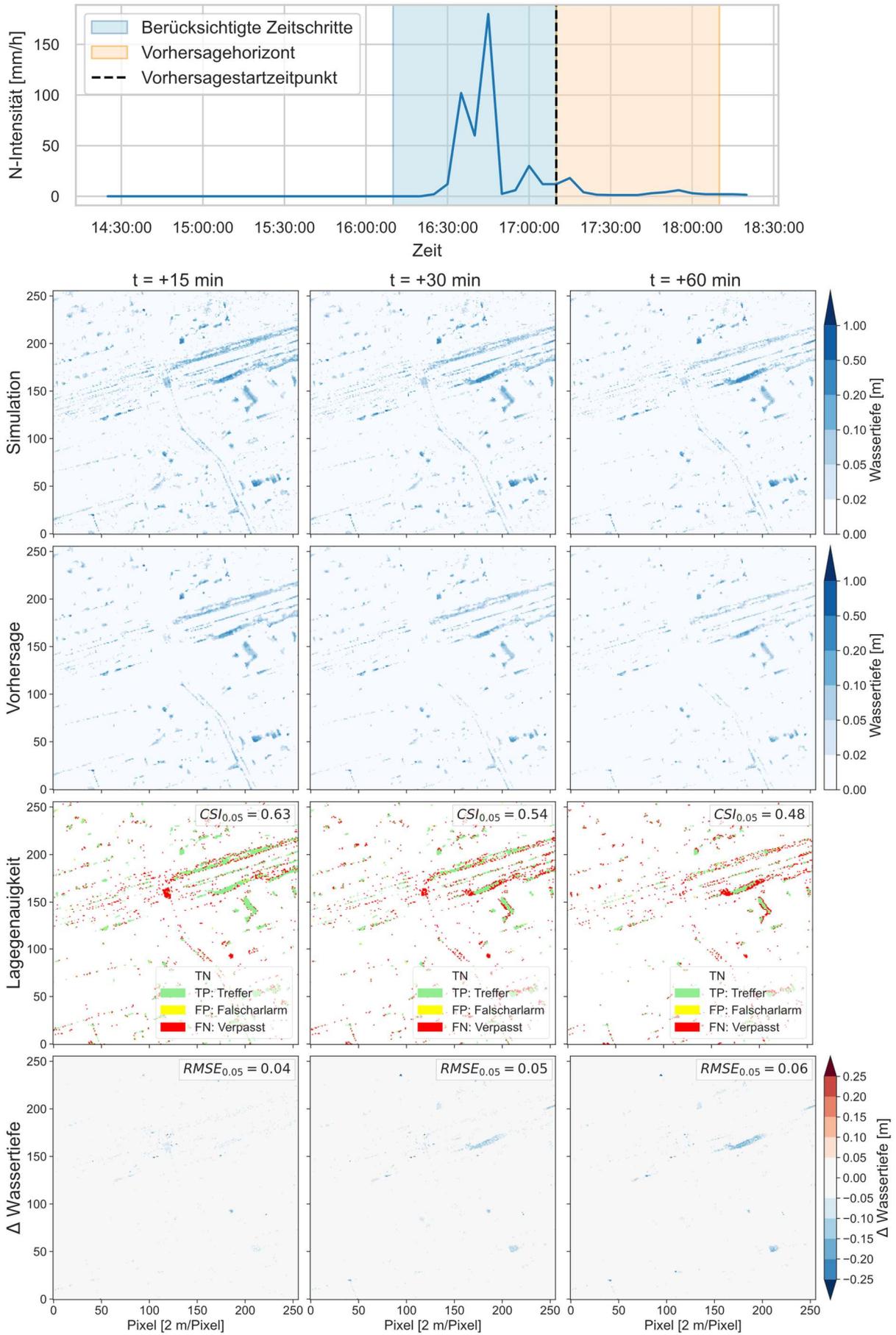
A 6-7: Ereignis ID 89: Prognose zum Ereignisbeginn



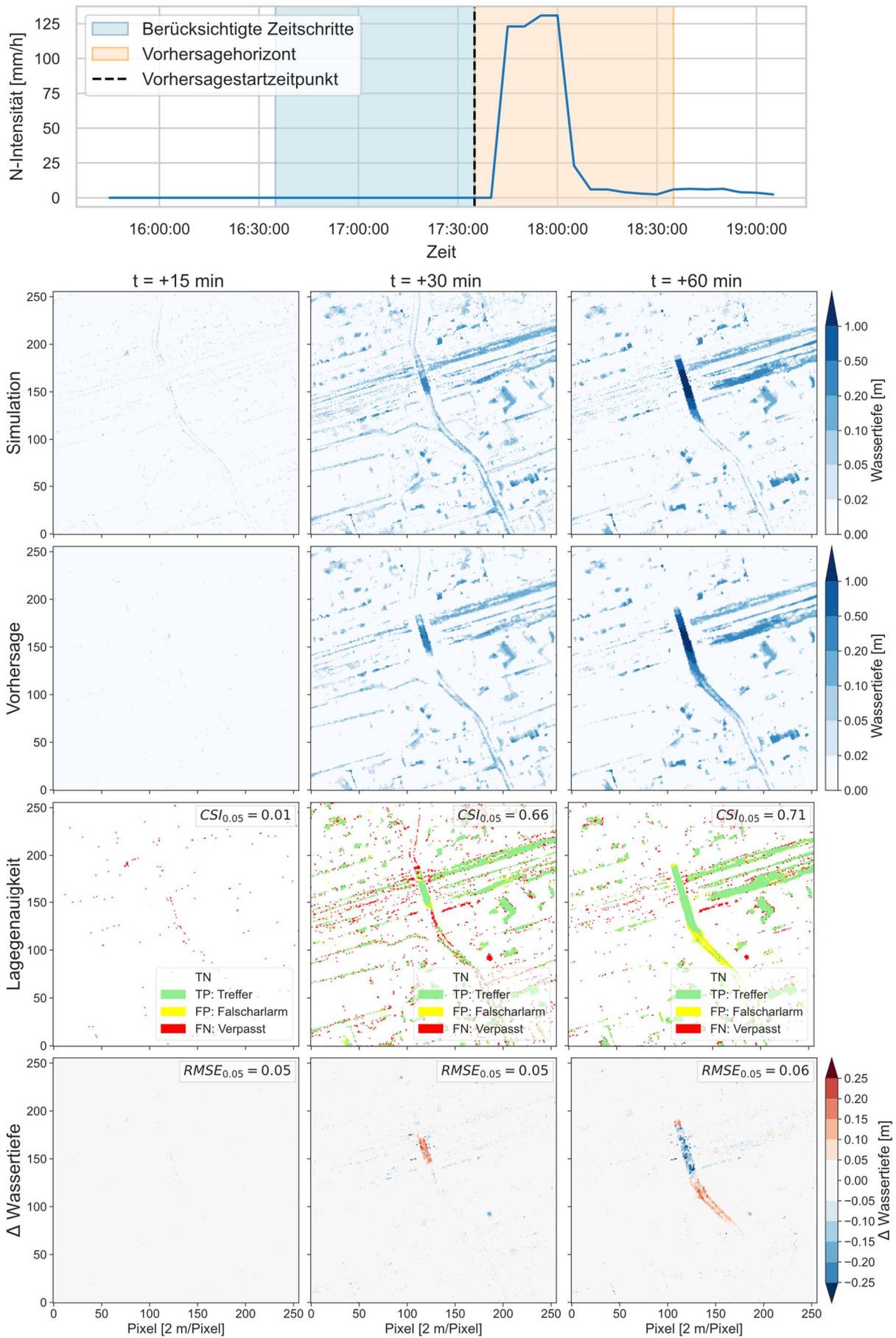
A 6-8: Ereignis ID 89: Prognose zur Ereignismitte



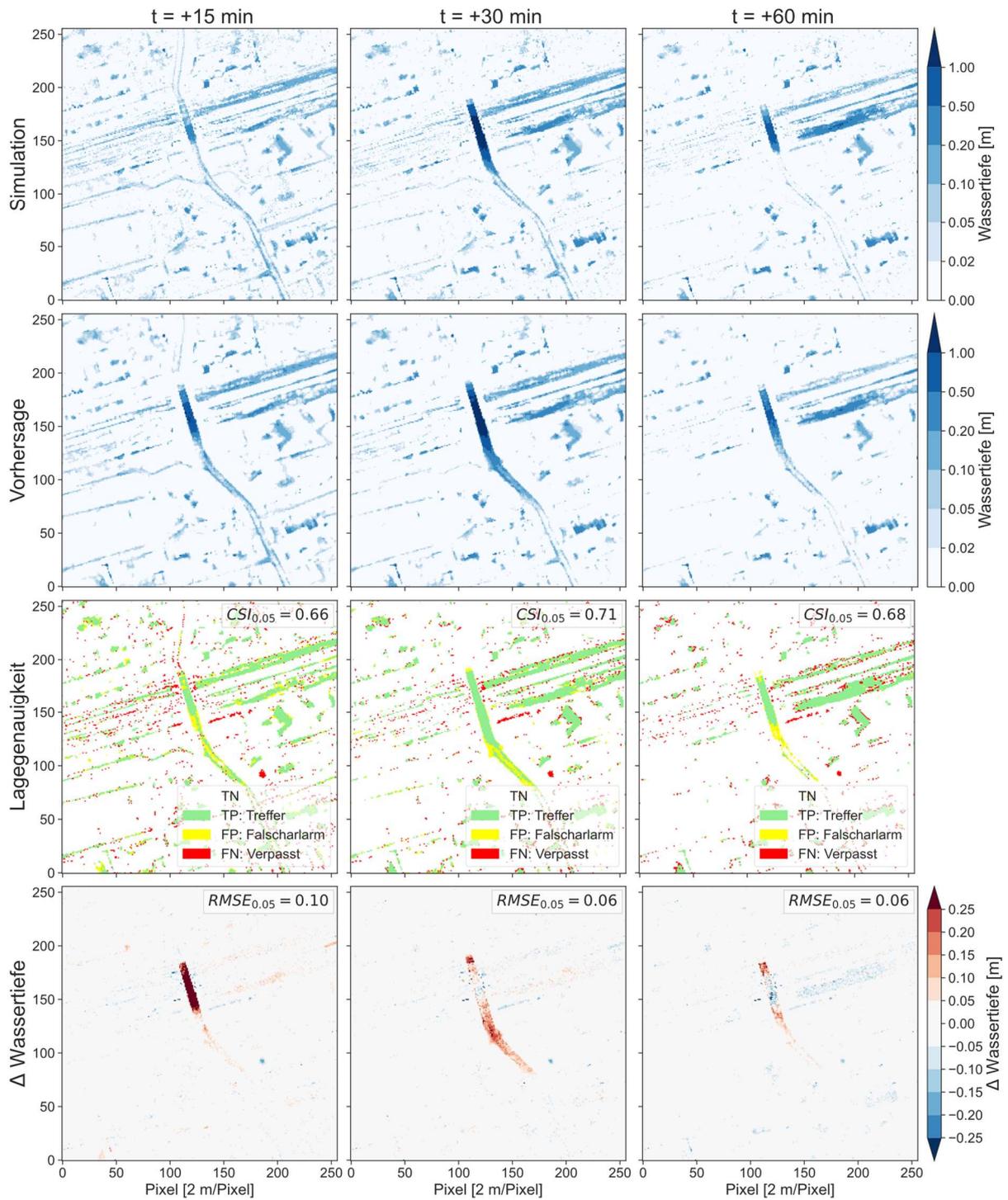
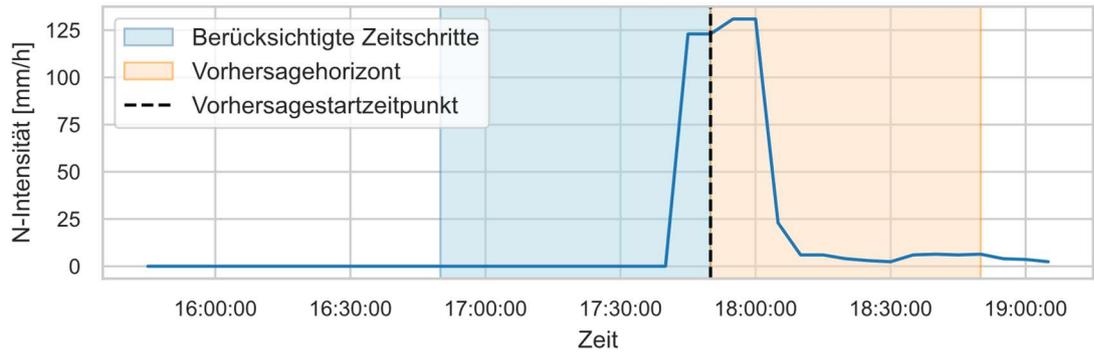
A 6-9: Ereignis ID 89: Prognose zum Ereignisende



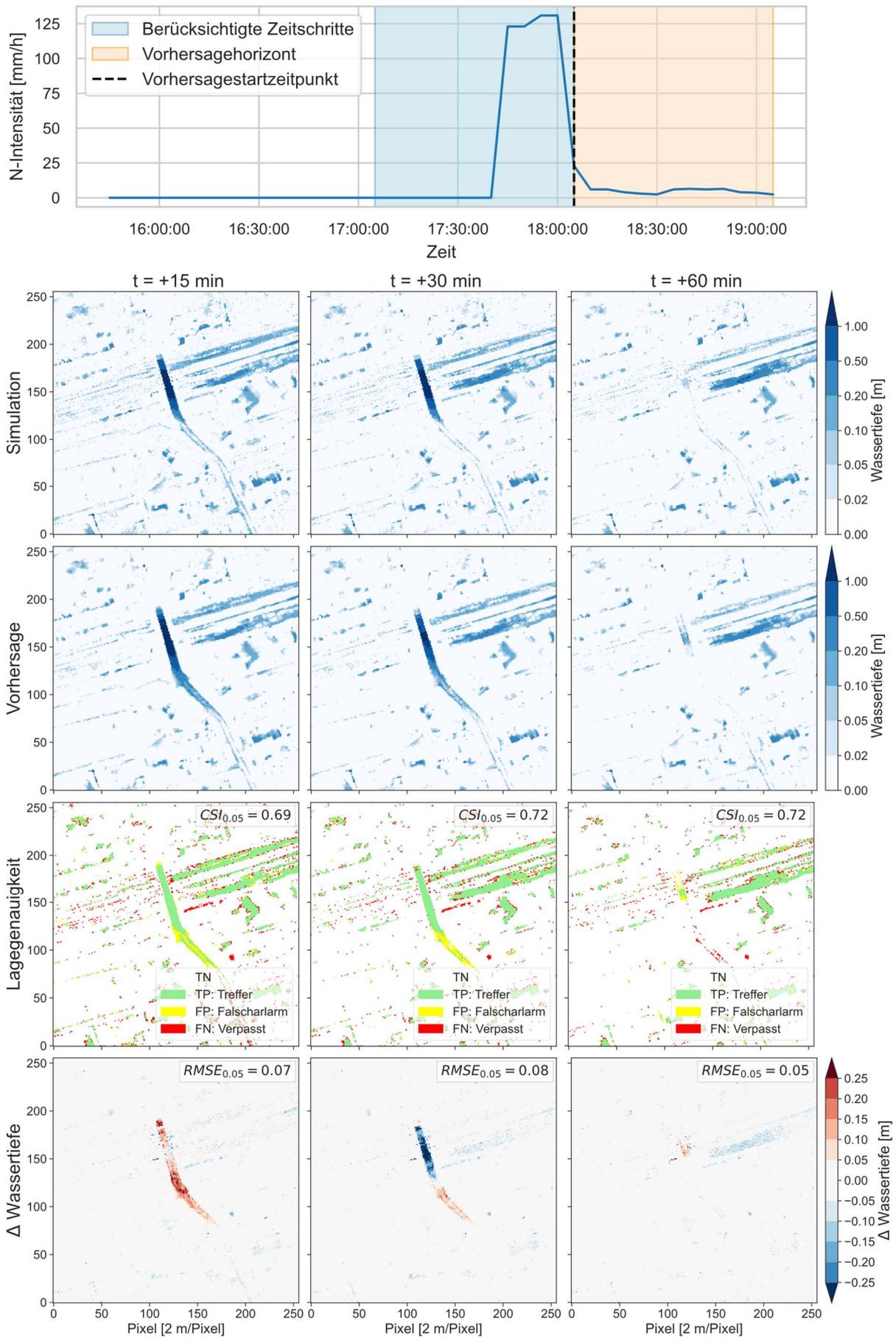
A 6-10: Ereignis ID 100: Prognose zum Ereignisbeginn



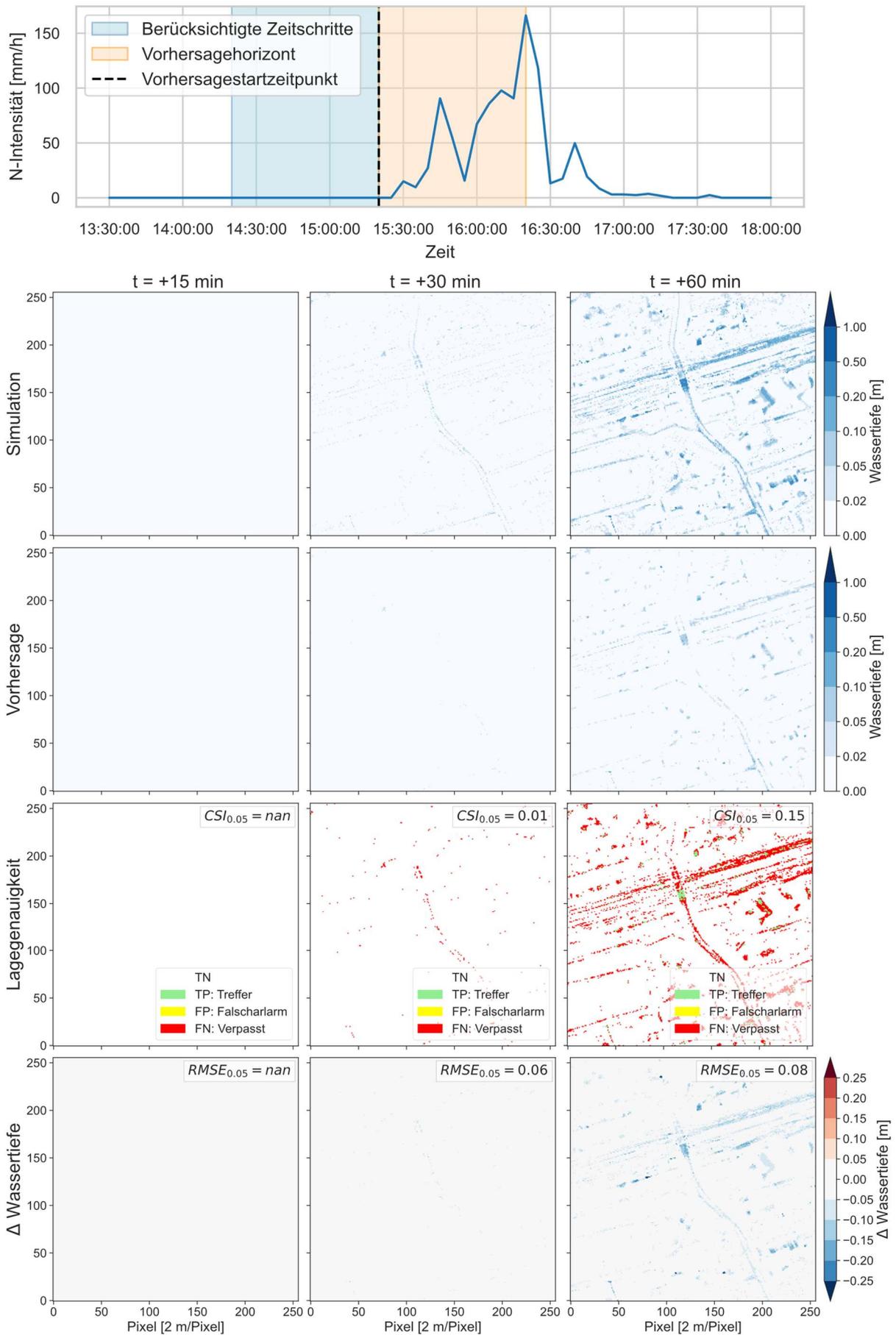
A 6-11: Ereignis ID 100: Prognose zur Ereignismitte



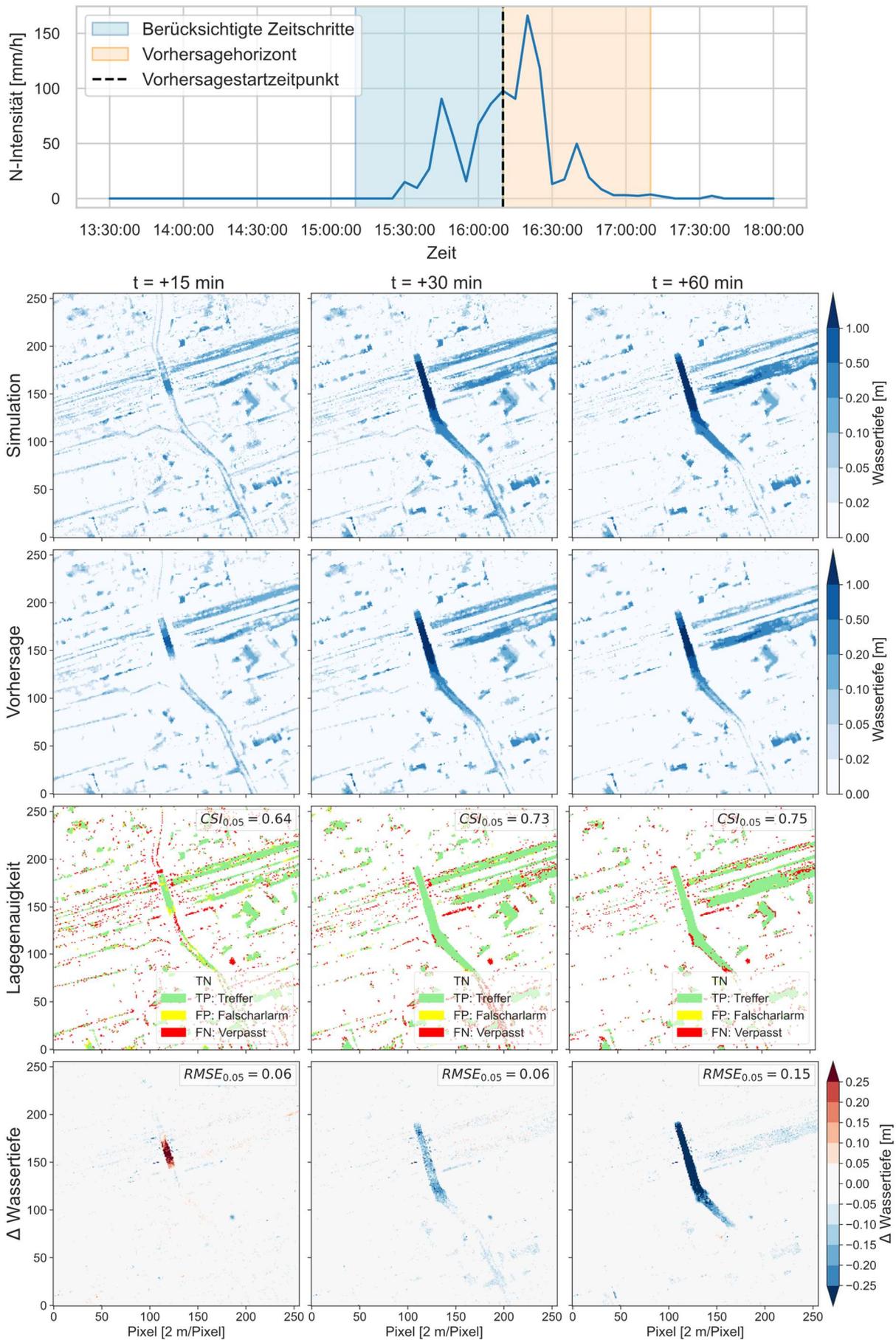
A 6-12: Ereignis ID 100: Prognose zum Ereignisende



A 6-13: Ereignis ID 102: Prognose zum Ereignisbeginn



A 6-14: Ereignis ID 102: Prognose zur Ereignismitte



A 6-15: Ereignis ID 102: Prognose zum Ereignisende

