TUM School of Engineering and Design

# Multistep Flood Inundation Forecast with Artificial Neural Network

Qing Lin

Vollständiger Abdruck der von der TUM School of Engineering and Design der Technischen Universität München zur Erlangung des akademischen Grades eines
**Doktors der Ingenieurwissenschaften (Dr.-Ing.)**
genehmigten Dissertation.

**Vorsitzender:** Prof. Dr. rer. nat. habil. Gabriele Chiogna

**Prüfer der Dissertation**

1. Prof. Dr.-Ing. Markus Disse
2. Prof. Dr. phil. Jorge Leandro

Die Dissertation wurde am 29.06.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Engineering and Design am 06.09.2022 angenommen.

# Abstract

Floods are the leading natural disaster in many countries in terms of damage and losses per year. Early prediction of such events can help prevent some of those losses. An early flood warning system requires a short response time and relatively accurate flood hazard precautions. Artificial neural networks (ANN) show a solid ability to deal quickly with large amounts of measured data. For early warnings of floods, an ANN for outputting flood inundation maps is developed based on multiple discharge inputs with a high grid resolution (4 m × 4 m). After testing different neural network training algorithms and network structures, we found resilience backpropagation to perform best. Furthermore, by introducing clustering for preprocessing discharge curves before training, the prediction quality could be improved. Synthetic flood events are used for the training and validation of the ANN. Historical events were additionally used for further validation. The results show that the developed ANN can predict the maximum flood inundation extent. The mean squared error in more than 98 and 86% of the total area is smaller than 0.2 $m^2$ in the prediction of synthetic events and historical events, respectively.

Besides the maximum inundation extent of flood events, a dynamic profile of the flood inundation area is also interested in flood precaution. It gives more details of the development and propagation of the flood areas. A forecast system is planned to perform multistep forecasts for 1 — 5 h after the flooding event has been triggered by a forecast threshold value. Thus, an ANN developed for the real-time forecast of flood inundation with a high spatial resolution (4 m × 4 m) is extended to allow multiple forecasts. After training with 120 synthetic flood events, the ANN was first tested with 60 synthetic events for verifying the forecast performance for 3h, six 6h, nine 9h, and 12h lead time. The model produces good results, as shown by more than 81% of all grids having an RMSE below 0.3 m. The ANN is then applied to the three historical flood events to test the multistep inundation forecast. For the historical flood events, the results show that the ANN outputs have a good forecast accuracy of the water depths for (at least) the 3h forecast with over 70% accuracy (RMSE within 0.3 m), and a moderate accuracy for the subsequent forecasts with (at least) 60

Hydrodynamic modeling is a widely applied method for detailed urban flood events. Meanwhile, the development of hydrodynamic modeling enables the possibility of including different drainage components in urban flood modeling, which also adds to the model structure uncertainties. The extreme flash flood event of Simbach 2016 is investigated with seven different model structures of modeled drainage components. First, the 2D hydrodynamic model PD-Wave is applied to model the surface flow. The 1D hydrodynamic model EPA-SWMM to model the inflows boundary conditions and urban drainage systems. The maximum inundation depths from different model structures are compared. By comparing the model performance, adding the sewer network could improve the model fit to the measured water depth data from the gauge stations. Adding the drainage pumps and water treatment plants

could slightly improve the model performance. The model performance drops slightly as the structure includes the dam erosions, resulting from simplifying the breach processes. Overall, the model structure uncertainty from the drainage components would have the larger in the rivers and the downstream polders, with an uncertainty of water depth up to 0.5 m.

In hydraulic urban pluvial flood modeling, the incorporation of different drainage components and means of coupling of different satellite areas have a high impact on the overall model structure uncertainty. Probabilistic flood maps can help visualize the uncertainties from multiple model structures and integrate the results from multiple model structures. Therefore, two Bavarian cities are chosen for studying the model structure uncertainty. The probability flood maps generated by three different weighting methods are applied to visualize the model structure uncertainty in both cities. Two new methods based on nearest map weights (NCW) and inverted weighted distances (IWD) are proposed for generating more informative probabilistic flood maps than the method considering overall equal weights (OW) of available checkpoints. Based on the case studies, uncertainty in the flood inundation arising from model couplings (20% – 70%) is more significant than drainage components (only up to 10%). Therefore, if the number of the checkpoints is limited OW method is a good alternative. NCW is more appropriate for localized analysis. IWD is more suitable for implementing a smooth transition across the whole area while still accounting for the impact of localized checkpoints.

# Zusammenfassung

Überschwemmungen sind in vielen Ländern die größte Naturkatastrophe, was die Schäden und Verluste pro Jahr angeht. Eine frühzeitige Vorhersage solcher Ereignisse kann dazu beitragen, einen Teil dieser Verluste zu verhindern. Ein Hochwasserfrühwarnsystem erfordert eine kurze Reaktionszeit und eine relativ genaue Vorhersage der Hochwassergefahr. Künstliche neuronale Netze (ANN) zeigen eine solide Fähigkeit, schnell mit großen Mengen von Messdaten umzugehen. Für die Hochwasserfrühwarnung wird ein ANN zur Ausgabe von Überschwemmungskarten entwickelt, das auf mehreren Abflussdaten mit einer hohen Gitterauflösung (4 m × 4 m) basiert. Nach dem Testen verschiedener Trainingsalgorithmen für neuronale Netze und Netzstrukturen haben wir festgestellt, dass die Resilienz-Backpropagation am besten funktioniert. Darüber hinaus konnte die Vorhersagequalität durch die Einführung von Clustering zur Vorverarbeitung der Abflusskurven vor dem Training verbessert werden. Synthetische Hochwasserereignisse werden für das Training und die Validierung des ANN verwendet. Zur weiteren Validierung wurden zusätzlich historische Ereignisse herangezogen. Die Ergebnisse zeigen, dass das entwickelte ANN das maximale Überschwemmungsausmaß vorhersagen kann. Der mittlere quadratische Fehler ist bei der Vorhersage von synthetischen Ereignissen und historischen Ereignissen in mehr als 98 bzw. 86% der Gesamtfläche kleiner als 0,2 $m^2$.

Neben dem maximalen Überschwemmungsausmaß von Hochwasserereignissen ist für die Hochwasservorsorge auch ein dynamisches Profil des Überschwemmungsgebietes von Interesse. Es gibt nähere Auskunft über die Entwicklung und Ausbreitung der Überschwemmungsgebiete. Ein Vorhersagesystem soll mehrstufige Vorhersagen für 1 – 5 h nach Auslösung des Hochwasserereignisses durch einen Vorhersageschwellenwert durchführen. Dazu wird ein ANN, das für die Echtzeitvorhersage von Überschwemmungen mit hoher räumlicher Auflösung (4 m × 4 m) entwickelt wurde, so erweitert, dass Mehrfachvorhersagen möglich sind. Nach dem Training mit 120 synthetischen Hochwasserereignissen wurde das ANN zunächst mit 60 synthetischen Ereignissen getestet, um die Vorhersageleistung für 3h, sechs 6h, neun 9h und 12h Vorlaufzeit zu überprüfen. Das Modell liefert gute Ergebnisse, was sich darin zeigt, dass mehr als 81% aller Raster einen RMSE unter 0,3 m aufweisen. Das ANN wird dann auf die drei historischen Hochwasserereignisse angewendet, um die mehrstufige Überschwemmungsvorhersage zu testen. Für die historischen Hochwasserereignisse zeigen die Ergebnisse, dass die ANN-Ausgänge eine gute Vorhersagegenauigkeit der Wassertiefen für (mindestens) die 3-Stunden-Vorhersage mit über 70% Genauigkeit (RMSE innerhalb von 0,3 m) und eine mäßige Genauigkeit für die nachfolgenden Vorhersagen mit (mindestens) 60% Genauigkeit aufweisen.

Die hydrodynamische Modellierung ist eine weit verbreitete Methode zur detaillierten Beschreibung städtischer Hochwasserereignisse. Die Entwicklung der hydrodynamischen Modellierung ermöglicht inzwischen die Einbeziehung verschiedener Abflusskomponenten in die

städtische Hochwassermodellierung, wodurch sich auch die Unsicherheiten in der Modellstruktur erhöhen. Das extreme Sturzflutereignis von Simbach 2016 wird mit sieben verschiedenen Modellstrukturen von modellierten Abflusskomponenten untersucht. Zunächst wird das 2D-Hydrodynamikmodell PD-Wave zur Modellierung der Oberflächenströmung eingesetzt. Das 1D-Hydrodynamikmodell EPA-SWMM zur Modellierung der Zuflussrandbedingungen und der städtischen Entwässerungssysteme. Die maximalen Überschwemmungstiefen der verschiedenen Modellstrukturen werden verglichen. Der Vergleich der Modellleistung zeigt, dass das Hinzufügen des Kanalisationsnetzes die Anpassung des Modells an die gemessenen Wassertiefendaten der Pegelstationen verbessern könnte. Das Hinzufügen von Entwässerungspumpen und Wasseraufbereitungsanlagen könnte die Modellleistung leicht verbessern. Die Modellleistung sinkt leicht, wenn die Struktur die Dammerosionen einschließt, was auf eine Vereinfachung der Bruchprozesse zurückzuführen ist. Insgesamt wäre die Modellstrukturunsicherheit durch die Entwässerungskomponenten in den Flüssen und den flussabwärts gelegenen Poldern am größten, mit einer Unsicherheit der Wassertiefe von bis zu 0,5 m.

Bei der hydraulischen Hochwassermodellierung in städtischen Überschwemmungsgebieten haben die Einbeziehung verschiedener Entwässerungskomponenten und die Art der Kopplung verschiedener Satellitengebiete einen großen Einfluss auf die Gesamtmodellstrukturunsicherheit. Probabilistische Hochwasserkarten können dabei helfen, die Unsicherheiten aus mehreren Modellstrukturen zu visualisieren und die Ergebnisse aus mehreren Modellstrukturen zu integrieren. Daher wurden zwei bayerische Städte für die Untersuchung der Modellstrukturunsicherheit ausgewählt. Die Wahrscheinlichkeits-Hochwasserkarten, die mit drei verschiedenen Gewichtungsmethoden erstellt wurden, werden zur Visualisierung der Modellstrukturunsicherheit in beiden Städten verwendet. Zwei neue Methoden, die auf den Gewichten der nächstgelegenen Karte (NCW) und den invertierten gewichteten Abständen (IWD) basieren, werden vorgeschlagen, um informativere probabilistische Hochwasserkarten zu erstellen als die Methode, die insgesamt gleiche Gewichte (OW) der verfügbaren Kontrollpunkte berücksichtigt. Basierend auf den Fallstudien ist die Unsicherheit in der Überschwemmung, die sich aus Modellkopplungen (20% - 70%) ergibt, bedeutender als die Abflusskomponenten (nur bis zu 10%). Daher ist die OW-Methode eine gute Alternative, wenn die Anzahl der Kontrollpunkte begrenzt ist. NCW ist besser geeignet für eine lokalisierte Analyse. Die IWD-Methode ist besser geeignet, um einen sanften Übergang über das gesamte Gebiet zu implementieren und gleichzeitig die Auswirkungen der lokalen Kontrollpunkte zu berücksichtigen.

# List of Publications

## Journal Articles

Prediction of maximum flood inundation extents with resilient backpropagation neural network: Case Study of Kulmbach
Qing Lin, Jorge Leandro, Wenrong Wu, Punit Bhola, Markus Disse, *Frontiers in Earth Science*, Vol. 8, 2020, p332. DOI: 10.3389/feart.2020.00332.

Multistep flood inundation forecasts with resilient backpropagation neural networks: Kulmbach Case Study
Qing Lin, Jorge Leandro, Stefan Gerber, Markus Disse, *Water*, 12(12), 2020, 3268. DOI: 10.3390/w12123568.

Probabilistic Flood Maps Using Multiple Urban Flood Model Structures
Qing Lin, Jorge Leandro, Markus Disse, submitted to *Hydrological Sciences Journal*.

Optimization of Artificial Neural Network (ANN) for Maximum Flood Inundation Forecasts
Hongfei Zhu, Jorge Leandro, Qing Lin, *Water*, 13(16), 2021, 2252. DOI: 10.3390/w13162252.

Inclusion of narrow flow paths between buildings in coarser grids for urban flood modeling: virtual surface links
Sebastian Ramsauer, Jorge Leandro, Qing Lin, *Water*, 2021, 13(19), 2629. DOI: 10.3390/w13192629.

## Conferences

Impact of the model structure uncertainty in flood simulations: Simbach case study
Qing Lin, Jorge Leandro, Daniel Sturm, Markus Disse, *EGU General Assembly 2020*, Online, 4–8 May 2020, EGU2020-15216. DOI: 10.5194/egusphere-egu2020-15216.

Uncertainty 1D-2D coupled models for flash flood simulation with different rainfall boundary conditions (Kulmbach case study)
Qing Lin, Jorge Leandro, Markus Disse, *EGU General Assembly 2019*, Vienna, 2019, vol.21, EGU2019-14976-1. Web: https://meetingorganizer.copernicus.org/EGU2019/EGU2019-14976-1.pdf.

Flash flood modeling on HPC systems
Lin et al., *Enviroinfo*, Leibnitz-Rechenzentrum Garching, 4.9.2018. doi: 10.13140/RG.2.2.28893.46565

Das Projekt HiOS - Erstellung einer Hinweiskarte für Oberflächenabfluss und Sturzfluten für

bayerische Gemeinden
Karl Broich, Maria Kaiser, Qing Lin, Johannes Mitterer, Hai Nguyen, Thomas Pflugbeil, Fabian von Trentini, Florian Willkofer, Markus Disse, Ralf Ludwig, *Seminar Starkregen und Sturzfluten - Erfassen, Erforschen, Evaluieren*, Nov. 2018. doi: 10.3243/kwe2018.11.005.

# Acronyms and Abbreviations

| | |
|---|---|
| 1D | one-dimensional |
| 2D | two-dimensional |
| ANN | artificial neural network |
| CA | California |
| CET | central European time |
| CGF | conjugate gradient |
| CK | checkpoint |
| CNN | convolutional neural network |
| CSI | critical success index |
| DBM | data based mechanistic |
| DEM | digital elevation model |
| DWD | Deutscher Wetterdienst, German Weathter Service |
| EPA | Environmental Protection Agency |
| FCM | fuzzy C-means clustering |
| FAR | false alarm ratio |
| FNN | forward-feed neural network |
| GB | gigabyte |
| GIS | geographical information system |
| GLUE | generalized likelihood uncertainty estimation |
| GPU | graphical processing unit |
| HEC-RAS | Hydrologic Engineering Center's River Analysis System |
| IWD | inverted weighted distance |

| | |
|---|---|
| LARSIM | Large Area Runoff Simulation Model |
| LfU | Landesamt für Umwelt, Environmental Agency |
| MLP | multilayer perceptron |
| MS | model structure |
| MSE | mean-square error |
| NCW | nearest checkpoint weight |
| NSE | Nash-Sutcliffe efficiency |
| OW | overall weight |
| PC | personal computer |
| PCA | principal component analysis |
| PD-Wave | Parallel Diffusive Wave |
| POD | possibility of detection |
| RAM | ramdom-access memory |
| RMSD | root-mean-square deviation |
| RMSE | root-mean-square error |
| RP | resilient backpropagation |
| SUFI | sequential uncertainty fitting |
| SVM | support vector machine |
| SWMM | Storm Water Management Model |
| UK | United Kingdom |
| USA | United States of America |
| WBANN | wavelet bootstrap artificial neural network |
| WTTP | wastewater treatment plant |

# Acknowledgment

First, I would like to give my most enormous thanks to Prof. Dr. -Ing. Markus Disse. During all my Ph.D. studies. Prof. Disse supported me in all kinds of material and scientific means. Thanks for offering me the precious research position in the HiOS project and his support on the research topic; therefore, I had a chance to get the project funding, which relieved my pressure on pursuing my Ph.D. degree. Thus, I have an opportunity to access enough data and materials from the HiOS Project for my research. Furthermore, Prof. Disse gave me plenty of suggestions on academic conferences and instructions on topic choices, which lit up my future directions many times. Besides, I would also like to thank Prof. Disse for the chair organization and the social events during my Ph.D. These efforts have built up an excellent platform for communication inside our faculty and exchanging research thoughts with colleagues. I also thank my second supervisor and mentor, Prof. Dr. Jorge Leandro. Prof. Leandro gave me great help with this work. I benefitted greatly from the discussion with Prof. Leandro in all aspects. It is a great honor to work with Prof. Leandro, with his expertise in urban floods and his passion for daily work. I can still remember countless times of our lunchtime together and the friendly exchange of ideas from time to time. Besides, I would like to thank Prof. Leandro for his kind support in all my submissions, from journal articles to conference submissions. Prof. Leandro guided me with countless suggestions and extraordinary patience. Meanwhile, I have also gained plenty of experience in student co-supervision with Prof. Leandro. From this, I would like to thank Prof. Leandro again for the communication techniques and the great practice of inspiring students toward further goals.

I want to thank more people for offering diverse help on my way to this academic work. I want to give my further profound thanks to our colleagues in the HiOS project, Dr. Kaiser Maria, Dr. Karl Broich, and Thomas Pflugbeil, as well as other members from LMU and LRZ. Dr. Kaiser Maria, as the leader of the HiOS project, manages the project efficiently, which eases the data and methods exchanges and boosts my Ph.D. studies. Dr. Karl Broich and Thomas Pflugbeil offered me great help on hydrodynamic modeling and gave valuable comments on our research outputs. I also want to express my gratitude to Dr. Punit Bhola for my benefit from his results and dataset from the project FloodEvac. I would also thank our secretary, Christiane Zach-Cretaine, for supporting chair organization and administration issues. Besides, the other colleagues from the chair of hydrology and river basin management have also offered great help. From the Ph.D. seminar to the everyday talks, the exchange of different topics further widened my research perspectives and inspired me on research methods.

I would also like to give my deepest thanks to my parents. Thanks to their great effort to raise me and provide me with a good education, which paves a solid fundamental for my academic road. During my Ph.D. studies, they understood and supported me with my life both mentally and materially. Especially in the global COVID, when the international travel bans

stopped us from seeing each other, their understanding and love still gave me great strength to concentrate on my dissertation. I would also like to thank my girlfriend, Lu. Thanks to her emotional and scientific support in recent years. Her advice is indispensable for refining this academic work, and her optimistic attitude makes our daily life more joyful.

Finally, I thank the Bavarian Environmental Agency and Bavarian State Ministry of the Environment and Consumer Protection for the project supervision and financial support. It was a remarkable scientific collaboration, and I want to express my gratitude again for the funding support.

# Contents


1 Introduction ..... 17
1.1 Background ..... 17
1.2 Problem Statement ..... 18
1.3 Research Objectives, Questions, and Hypothesis ..... 20
1.4 Dissertation Structure ..... 21
2 State of the Art ..... 22
2.1 Modeling Methods of Urban Flood Inundation ..... 22
2.2 Data-driven Models for Flood Forecast ..... 23
2.3 Model Structures for Urban Flood Modelling ..... 25
2.4 Model Uncertainties in Urban Flood Modelling ..... 26
3 Study Areas ..... 28
3.1 Kulmbach ..... 28
3.2 Simbach am Inn ..... 29
3.3 Baiersdorf ..... 29
4 Prediction of Maximum Flood Inundation Extents ..... 31
4.1 Methodology ..... 31
4.1.1 Artificial Neural Network and Resilient Backpropagation Algorithm ..... 31
4.1.2 ANN Network Input and Output ..... 32
4.1.3 Fuzzy C-Means Clustering and Principal Component Analysis Clustering 33
4.1.4 Model Evaluation ..... 35
4.1.5 Hydrodynamic Model HEC-RAS 2D ..... 35
4.2 Data Preparation ..... 36
4.3 Results and Discussion ..... 36
4.3.1 Training Algorithm ..... 37
4.3.2 Number of Hidden Layers and Neurons ..... 38
4.3.3 Grid Resolution ..... 39
4.3.4 Fuzzy C-Means (FCM) Clustering ..... 40
4.3.5 Maximum Flood Inundation Prediction for Synthetic Events ..... 43
4.3.6 Maximum Flood Inundation Prediction for Historical Events ..... 44
5 Prediction of Multi-Step Flood Inundation ..... 46
5.1 Methodology ..... 46
5.1.1 Data and Network Structure ..... 46
5.1.2 Hyperparameter Tuning in ANN ..... 47
5.1.3 Prediction of the First Interval of Flood Events ..... 48
5.1.4 Real-Time Forecasting for Multistep Forecast Intervals ..... 49
5.1.5 Model Evaluation ..... 49
5.2 Results and Discussion ..... 50
5.2.1 Water Depth Prediction of the First Intervals ..... 51
5.2.2 Multistep Water Depth Predictions for Realtime Forecast ..... 53
5.2.3 Forecast of Inundation Extents ..... 56

6 Model Structure Uncertainty ..... 64
6.1 Methodology ..... 64
6.1.1 1D Dynamic Rainfall-Runoff Model SWMM ..... 65
6.1.2 2D Dynamic Rainfall-Runoff Model PD-Wave ..... 67
6.1.3 1D-2D Model Coupling of SWMM and P-DWave ..... 69
6.1.4 Rainfall Data ..... 70
6.1.5 1D Catchment Inflow Modelling ..... 71
6.1.6 Model Structures of 1D/2D Modelling ..... 72
6.1.7 1D Sewer Network Modelling for Simbach Case ..... 73
6.1.8 2D Surface Flow Modelling for Simbach Case ..... 73
6.1.9 Uncertainty Assessment in Model Structures ..... 74
6.2 Data for Model Performance Evaluation ..... 75
6.3 Results and Discussion ..... 76
6.3.1 Rainfall-Runoff Model Calibration ..... 76
6.3.2 Maximum Water Depth in Coupled Model Structures and Structural Uncertainty ..... 77
6.3.3 Inundation extent ..... 82
6.3.4 Impact of Sewer System and Dam Erosion on Maximum Water Depth ..... 83
7 Probabilistic Flood Map ..... 86
7.1 Methodology ..... 86
7.1.1 Probabilistic Inundation Map ..... 86
7.1.2 Probabilistic Inundation Map Extent ..... 86
7.1.3 Probabilistic Inundation Map by Overall Weights (OW) ..... 87
7.1.4 Probabilistic Inundation Map by Nearest Checkpoint Weights (NCW) ..... 87
7.1.5 Probabilistic Inundation Map by Inverted Weighted Distances (IWD) ..... 88
7.1.6 Taylor Diagram ..... 89
7.1.7 Hydraulic Models for Urban Flood Modelling ..... 89
7.1.8 Study Area and Materials ..... 90
7.2 Results and Discussion ..... 92
7.2.1 Probabilistic flood inundation map of uncertainty on modeled components 92
7.2.2 Probabilistic flood inundation map of uncertainty on model couplings ..... 97
7.2.3 The Advantages and Disadvantages of Probabilistic Inundation Maps ..... 102
8 Summary and Conclusions ..... 103
8.1 ANN for Flood Inundation Forecast ..... 103
8.2 Probabilistic Flood Maps for Model Structure Uncertainty Quantification ..... 105
8.3 Outlook and Future Work ..... 107

# List of Figures


Figure 1 Study area of Kulmbach ..... 28
Figure 2 Study area of Simbach ..... 30
Figure 3 Study area of Baiersdorf ..... 30
Figure 4 Neural network setups for maximum inundation prediction ..... 33
Figure 5 Model performances by two training algorithms in testing dataset ..... 37
Figure 6 Model performances by the two training algorithms in the training dataset..... 38
Figure 7 Model performance by numbers of ANNs in testing dataset..... 39
Figure 8 Determine cluster numbers and numbers of principal components ..... 40
Figure 9 Clustering of discharge curves in training dataset by FCM ..... 41
Figure 10 Clustering of discharge curves in training dataset by PCA-FCM..... 42
Figure 11 Example of maximum flood inundation prediction in testing dataset..... 44
Figure 12 The error of ANN prediction of the Historical Event Feb 2005 ..... 45
Figure 13 The error of ANN prediction of the Historical Event May 2013..... 45
Figure 14 The forward-feed neural network setup for the multistep forecast ..... 46
Figure 15 Training of ANN forecast model using first intervals ..... 48
Figure 16 Shift of ANN forecast models for multistep forecasts..... 49
Figure 17 Hydrograph of Event 2006 ..... 52
Figure 18 Predicted inundation maps of the first intervals of Event 2006 ..... 58
Figure 19 Hydrograph of Flood Event 2013 ..... 59
Figure 20 Predicted inundation maps of the first intervals of Flood Event 2013..... 60
Figure 21 Hydrographs of Event 2005..... 61
Figure 22 Predicted inundation maps of the first intervals of Flood Event 2005..... 62
Figure 23 Performance of the forecast of inundation extent growths by three indices ..... 63
Figure 24 Model coupling scheme ..... 64
Figure 25 Spatial distribution of the 1D and the coupled 1D/2D model ..... 65
Figure 26 Rainfall intensities on the 1st of June 2016..... 70
Figure 27 Sub-catchments and open channels for the rainfall-runoff simulation ..... 71
Figure 28 Observed discharge and precipitation between 01.06.2015 and 03.06.2015.... 71
Figure 29 Observed discharge at the gauge station of Simbach am Inn between 31.5.2016 and 02.06.2016 (LfU 2020) ..... 72

Figure 30 Sewer network with important features in Simbach ..... 74
Figure 31 Specific locations in Simbach ..... 75
Figure 32 Calibrated discharge input curve of Simbach 2016 and comparison to observation ..... 76
Figure 33 Maximum water depths MS 1 ..... 77
Figure 34 Maximum water depths MS 2 ..... 78
Figure 35 Maximum water depths MS 3a ..... 79
Figure 36 Maximum water depths MS 3b ..... 79
Figure 37 Maximum water depths MS 3c ..... 80
Figure 38 Maximum water depths MS 4a ..... 81
Figure 39 Maximum water depths MS 4b ..... 81
Figure 40 Flood extents of MS 3c and MS 4b ..... 83
Figure 41 Flood extents of MS 3c and MS 4b, upstream ..... 83
Figure 42 Flood extents MS 3c and MS 4b, downstream ..... 84
Figure 43 Impact of the sewer system on the maximum depth (MS 3c minus MS 1) ..... 84
Figure 44 Impact of the sewer system on the maximum depth (MS 4b minus MS 4a) ..... 84
Figure 45 Probabilistic inundation map of Simbach weighted by the overall weights (OW) 93
Figure 46 Probabilistic inundation map of Simbach by the nearest checkpoint weights (NCW) ..... 94
Figure 47 Probabilistic inundation map of Simbach weighted by Inverted Weighted Distances (IWD) ..... 94
Figure 48 Histograms of the percentage of inundation probabilities ..... 95
Figure 49 Taylor Diagram of the different model structures compared to the measured water levels in Simbach ..... 96
Figure 50 Histograms of uncertain water depths in the seven model structures, Simbach 96
Figure 51 Probabilistic inundation map of Baiersdorf weighted by the overall weights (OW) ..... 98
Figure 52 Probabilistic inundation map of Baiersdorf by the nearest checkpoint weights (NCW) ..... 98
Figure 53 Probabilistic inundation map of Baiersdorf weighted by Inverted Weighted Distances (IWD) ..... 99
Figure 54 Histograms of inundation probabilities Baiersdorf ..... 100

Figure 55 Taylor Diagram of different model structures compared to the measured values, Baiersdorf ..... 101
Figure 56 Histograms of uncertain water depths from seven model structures Baiersdorf 102

# List of Tables


Table 1 Number of grids in each combination (number of layers and neurons) which outperform the others ..... 38

Table 2 Impact of ANN number on training ..... 39

Table 3 Integral of the bandwidth of 90% confidence intervals of discharge inputs with FCM and PCA-FCM ..... 43

Table 4 Sign of error difference between PCA-FCM clustered 100 events, original 120 events, and the randomly clustered 100 events ..... 43

Table 5 Numbers of grids with large errors and their percentages ..... 51

Table 6 Numbers of wet grids and accurate grid percentage for event 2006 ..... 52

Table 7 Numbers of wet grids and accurate grid percentages for the flood event in 2013 54

Table 8 Numbers of wet grids and accurate grid percentages for the flood event in 2005 54

Table 9 Sub-sequential forecast accuracy percentages for the flood event in 2006 ..... 55

Table 10 Sub-sequential forecast accuracy percentages rate for the flood event in 2013 . 55

Table 11 Sub-sequential forecast accuracy percentages for the flood event in 2005 ..... 56

Table 12 Rainfall-runoff parameter calibration ..... 72

Table 13 Summary of the conducted simulation scenarios ..... 73

Table 14 Goodness of fit of seven model structures Simbach ..... 82

Table 15 Summary of the model structures of Simbach ..... 92

Table 16 Summary of model structures of Baiersdorf ..... 93

# 1. Introduction

## 1.1. Background

Flood is one of the most dangerous natural hazards hitting settlements which threatens the safety of civilians and the integrity of infrastructures (Berz, 2001). Flooding is the leading cause of damage and losses in many countries in the world (Kron, 2005). Furthermore, as a result of climate and land-use changes, the flood vulnerability of some regions is expected to rise greatly (Vogel et al., 2011).

Flood is the primary cause of loss and damage among natural phenomena in many regions worldwide. According to Kron (2005), flood is defined as "temporary covering of land by water as a result of surface waters escaping from their normal confines or as a result of heavy precipitation." By different characteristics, a flood could be further classified as river flood, flash flood, storm surge, etc. River flood results from long-lasting precipitation for days and weeks, with the precipitation volume exceeding the soil water capacity. As the water cannot be infiltrated into the soil, it enters directly into the river, causing a rise in the water level in the river. The overspilling water from the river channel can cause large inundated areas in flat valleys. In narrow valleys, the rising water level brings higher momentum, damaging the infrastructures with solid forces and sometimes transported sediments. Flash flood happens when precipitation with high intensity falls quickly when the precipitation volume exceeds the infiltration volume greatly. As flash flood is caused only by precipitation, it can occur on any type of topography, including flat areas without river creeks. In this case, many flash floods propagate rapidly, with the flooded area difficult foreseeable. Storm surges happen along the coast or by the big lakes. Following tsunamis, storm surges can cause huge losses of lives and properties. Overall, flood often leads to thousands of deaths in less developed countries. Besides, it also results in massive economic losses. Since the 1990s, significant flood hazards, e.g., the Yangtze in 1998 and 1996, Mississippi in 1993, Elbe and Danube in 2002, all make the economic losses over tens of thousand dollars. Due to the scattered flood hazard worldwide, better flood protections are a global demand rather than a regional issue.

Human activities have been one of the driving factors for the increase in flood occurrences. In some areas, the rising human settlements occupy the natural flood retention areas, enhancing the flood risks locally. The fast growth of population also raises the value of infrastructures per area, which raises the total loss in a flood hazard. From the perspective of land use, most floodplains are flat areas with easy access to water resources, suitable for transportation and development.

Flood damages are also increasing dramatically, especially in recent decades, with the grow-

ing intensity and frequency of flood events. The damage can cause even higher losses in highly populated regions and countries. With the enhanced mobility and growing population, the human settlements expanded to more expansive areas. Expanding human settlements to those newly developed areas might contain unknown flood risks. Due to this reason, the flood vulnerability is also rising dramatically in densely urbanized areas after the rapid urbanization processes in many developing countries.

More dams and channels are built for flood drainage against the growing flood risks, which requires a proper flood design. To fulfill this demand, an estimation of flood hazard-prone areas with flood modeling could help evaluate and prevent losses.

## 1.2. Problem Statement

Flood modeling is practical for flood hazard quantification. While applying flood modeling to its application, two aspects are worth paying attention to. The first is the rising demand for real-time flood forecasting. Due to the high uncertainties of flood development processes, the time for applying the proper flood mitigation measures is usually fixed into a tight time frame to secure human life and finance. From this objective, a real-time flood forecast system with a short running time to give an above-moderate accuracy would meet the growing need for flood early warnings. The second is the uncertainty quantification for flood modeling. In flood studies, flood models inevitably contain uncertainty due to simplifying modelings and data errors. Typically, some uncertainties that originated from the choice of the model are referred to as the model structure uncertainty. Therefore, the model results can be better quantified and transferred by analyzing model structure uncertainties, which helps in further decision-making in flood protections.

Based on these modeling techniques, flood risk approaches are evaluated historically in a deterministic manner (Di Baldassarre and Montanari, 2009). As the models become complex, the uncertainty from the dynamic, stochastic, and uncertain nature of the climate, hydrological, and river processes are omitted or simplified in the modeling processes. Uncertainty in flood modeling rises from the model inputs, model parameters, and model structures (Savage et al., 2016). In many cases, when modeling the same study area with different model structures, the model with the best fit may not always be the one with the most complex model structure (Butts et al., 2004). Therefore, a proper choice of the model structure would reconstruct the historical event with better accuracy.

The modeling uncertainty in hydrological and hydrodynamic simulation can originate from various sources. According to Butts et al. (2004), four sources of uncertainties occurs in the flow modeling: random or systematic errors in the model inputs or boundary condition data, random or systematic errors in the recorded output data, uncertainty due to sub-optimal parameter values, and errors due to incomplete or biased model structures. In the past studies, much research has been done on the first three kinds of parameter uncertainties. In

contrast, the influence of the model structure on model performance and uncertainty is rarely investigated.

There are several motivations to consider the model structures in flood simulations. First, it is a practical simplification to consider the main flow processes while ignoring the other processes in flood modeling. Second, as the flood modeling proceeds, new processes are integrated into the earlier models to represent the flow mechanisms. From this perspective, a quantify of the new model performance on the feedback of the more complex model structure is helpful for the new model evaluation (Atkinson et al., 2002). Third, many model structures are justified based on a specific dataset and data quality. As a new dataset becomes available, the evaluation of the adjusted model structures for the new dataset shows the impacts of the new data set on the simulation results. Finally, analyses of model structures are essential in model development and improvement.

The model structures can vary in different perspectives in flood modeling depending on the modeler's choice and assumptions. The commons differences are like (Butts et al., 2004):

- different processes
- different couplings
- different numerical discretization
- different spatial representation of grids and sub-catchments
- different interpretation and representation of raw data
- etc.

Specifically, the model processes are different at different levels in urban flood modeling. Flood events can be modeled using surface runoff, one-dimensional hydrodynamic, two-dimensional, or three-dimensional hydrodynamic models for different goals. In the context of urban flood modeling, the sewer network drains massive amounts of water in urban flood events. Therefore, the dual-drainage concept is implemented in most of the urban flood models (Djordjević et al., 2005) (Maksimovi et al., 2007) (Nasello and Tucciarelli, 2005). Generally, the sub-surface drainage infrastructures are modeled independently from the surface systems (Chen et al., 2016). Thus, the sub-surface sewer drainage system and drainage pumps, and water treatment plants could be included in an urban flood model as model components. The modeling of model component combinations produces various model structures with different complexity. For coupled flood models, model couplings could also impact the results. For the modeler, it would be helpful to understand this uncertainty from model couplings and the pros and cons for the model coupling simplification on the boundary conditions.

For the representation and analysis of flood damages in flood modeling, flood maps are a valuable tool to represent and analyze the spatial distribution of flood hazards from the output

of flood models, visualizing the flood vulnerabilities and the risks. Flood maps can be classified into three categories for their contents: flood hazard maps, flood vulnerability maps, and flood risk maps (Merz et al., 2007). Flood hazard maps show the flood intensity and the exceedance probability; flood vulnerability maps present the consequences of floods on both the economy and society; flood risk maps illustrate the spatial distribution of flood risks, the losses multiplied by the consequences. Generally, flood inundation maps represent the spatial occurrences of flooding areas. The inundation maps could be generated either in deterministic or probabilistic approaches (Stephens et al., 2012). As for deterministic inundation maps, the set-up is usually based on a two-dimensional (2D) hydraulic model calibrating some specific historical flood events to optimize the best-fit model parameters. Due to the model simplification, significant model structure uncertainties are inevitable in the simulation model, which should also be estimated to avoid the wrong assessment of the flood hazards. The probabilistic approaches of flood maps have been developed recently to overcome the limitation of a single model structure. The probabilistic inundation maps are generated from multiple hydraulic models based on ensemble runs with sensitivity analysis.

## 1.3. Research Objectives, Questions, and Hypothesis

One of the objectives of this study is to set up an efficient neural network model to replace the conventional 2D hydraulic model for producing flood inundation maps for fluvial floods. For an intermediate urban area, a flood simulation using hydrodynamic modeling takes several hours or days to evaluate the flood hazard, which is time-consuming for precautions for flood events. Luckily, with the help of a neural network model, a forecast result can be given within several minutes, enabling an early estimation of the future flood damages. Thus, the related peoples and authorities would have enough time to react with proper countermeasures. Therefore, the main objectives are:

- To choose the network structure and proper algorithm for network tuning
- To validate the forecast results with a testing dataset
- To apply the network for historical flood event forecast

Specific questions will be addressed in work:

- Which algorithm performs better at the neural network training for flood forecast?
- At which accuracy can the neuron network forecast reach? Can the forecast results from the neuron network be compared with the simulation results from hydraulic models?

The model structure uncertainty is further included in this study to quantify the uncertainty in flood modeling. The uncertainty analysis generally gives the confidence range of the flood model results. The case will be analyzed with further details for the flood risk based on the analyzed uncertainty range. Therefore, the main objectives are:

- To model flood events with different model structures
- To quantify the model performance using suitable statistical variables
- To represent the model structure uncertainty in visualization

Specific questions will be discussed in work:

- Which type of model structures has higher impacts on the flood simulation?
- How large is the model structure uncertainty in different simulation outputs?

## 1.4. Dissertation Structure

This dissertation is organized into eight chapters. The first chapter introduces the research issues regarding the flood forecast and flood modeling uncertainties, the objectives and the hypothesis are presented. Chapter 2 summarizes the state of the art of research background, from different approaches for flood forecast to various studies about the flood model uncertainties. Chapter 3 gives an overview of the three study areas, namely Kulmbach, Simbach, and Baiersdorf, respectively, for flood inundation forecast using artificial neural network and model structure uncertainty studies. Chapter 4 presents the study of artificial neural networks, focusing on the network setting for the forecast. Chapter 5 elaborates the similar methodology for multi-step flood forecast compared the results among different forecast time frames. Chapter 6 introduces different approaches for constructing hydrodynamic urban flood models by modeling different drainage components and model couplings. Chapter 7 analyses the model structure uncertainties from drainage components and model couplings and visualizes the results using probabilistic maps. Finally, Chapter 8 summarizes the majority of findings of the dissertation and points out the possibilities for future works.

# 2. State of the Art

## 2.1. Modeling Methods of Urban Flood Inundation

Various numerical models are widely used for urban flood modeling (Henonin et al., 2013). The typical approach for flood models is through modeling of physical flooding processes. The urban flood inundation modeling can generally be classified into several types: rapid flood spreading, one-dimensional sewer, one-dimensional overland, two-dimensional overland, coupling sewer overland (Bulti and Abebe, 2020). The choice of the hydrodynamic urban flood inundation modeling is considered counting several factors, such as the representation capability of the flooding process, the output forms and quality, the simulation time, and the requirement for input data.

The rapid flood spreading model takes the total flood volume as the input data (Liu and Pender, 2013). The model simulates the expansion of the flood water by equalizing the water levels in neighboring cells of the terrain (Yang et al., 2015). The model calculates pre-calculation routines and inundation routines. First, it delineates the terrain topography for determining the flow routing. Second, it spreads the water volume among the neighboring grids from the lowest elevation to the highest elevation. The simplification of the model enables a simulation time of fewer than two minutes (Liu et al., 2015). However, the model can only produce the final state of the inundation, which lacks the development of the flooding process (Fritsch et al., 2016).

One-dimensional sewer model simulates the stormwater by simplifying the system for simulating the surcharge or drainage of the underground drainage network (Mark et al., 2004). It models conduits as links and manholes as nodes. The one-dimensional Saint-Venant flow equations are applied to simulate the procedure, using the sewer network and the terrain data. In the calculation, the virtual storage volumes in the sewer network need to be estimated, which adds to the model uncertainty (Jiang et al., 2015). Nevertheless, the precision of the one-dimensional sewer model is suitable for emergency cases as well as early warnings (Gharbi et al., 2016).

A one-dimensional surface model simulates the surface flow in one dimension. It models the floodplain as a set of nodes and the flow paths as the links in between (Leandro et al., 2009). First, the automatic tool delineates the terrain data to set up the one-dimensional model. Then, the same equation in the one-dimensional sewer model solves the system. The model produces some inundation characteristics within the delineated flow channels with this approach. Unfortunately, the model cannot deal with multiple flow directions or flows surcharging upon street curbs, which jump over the pre-set flow paths.

A two-dimensional surface model considers the two orthogonal flow directions and simulates the process of flow propagation. First, it generates the mesh of the flood plain based on the terrain data. Then, The two-dimensional shallow water equations in a discretized form solve the hydraulic process. Overall, the two-dimensional surface model can generate the flood inundation areas and the flow dynamics, e.g., water depth and velocity. The results from the two-dimensional surface model are relatively higher. Usually, the digital terrain model used here should include more building features and urban details, while a resolution of 5 meters or less is recommended for this demand (Mark et al., 2004). However, the 2D surface model fails to represent the channel flow with a width below the resolution of the terrain model, ignoring some stormwater drainages in this case. On the other side, the 2D surface model also requires high resolution of the digital terrain model. Besides, the computation expense is relatively higher up to hours or days.

The coupling of the different hydrodynamic models expands the model capabilities in urban flood simulation. 1D – 1D sewer – surface model joins a one-dimensional surface model with a one-dimensional sewer network model, which enables the model of the flows between the surface water and the water in the drainage system. The boundary conditions at the coupling interfaces can be set as bi-directional links (Mark et al., 2004) or uni-directional inlets (Leandro et al., 2009). The 1D – 1D sewer – surface model usually has a low computation time of less than one hour. However, it fails to calculate the inundation of sewer surcharges beyond the pre-defined flow pathways (Henonin et al., 2013). A 1D – 2D sewer – surface model interconnects the underground 1D sewer network with the 2D surface. With the integration of the sewer networks, the results from the 1D – 2D model are usually more accurate than the 2D surface model, as the behavior of the sewer drainage system could be simulated (Hankin et al., 2008). A complete flood profile could be calculated through this model, including sewer surcharge points, water depths, flood extents, and flow velocity. The high data requirements and the relatively long computation time would be the restrictions for applying the 1D – 2D sewer – surface model (Bamford et al., 2008).

## 2.2. Data-driven Models for Flood Forecast

All approaches require field measurements for defining the model parameters. Besides, the prohibitive high computation costs and the requirement for detailed datasets further restrict the application for real-time forecasting (Vogel et al., 2011). With the advances in high-performance computing, graphics processing units (GPU) nowadays are capable of faster 2D simulation in much larger areas (Kalyanapu et al., 2011). Although these scalability techniques reduce the simulation time significantly, it is still unacceptably high in many cases for real-time early warning systems.

The development of computer technologies enables the access and analysis of a large amount of data, prompting the development of data-driven models. Data-driven approaches can be a feasible alternative for established flood simulation models (Mosavi et al., 2018). Unlike

conventional numerical models, data-driven models require input/output data only. The fast-growing trend of data-driven models has shown their high performance even for non-linear problems (Maniak, 2016). Unlike physical-based models, data-driven models do not require field measurements for determining (physically based) model parameters, which alleviates the burden on the users for data gathering and model setup. ANN can be a helpful tool for modeling if properly applied. Indeed some of the pitfalls are the likelihood of over-fitting or under-fitting the data and insufficient length of the data sets, which may lead to erroneous model results (Zhang, 2007).

Data-driven models are being applied more widely in flood precaution and forecast. For example, various short and long-term flood forecasts have been developed using neuro-fuzzy (Dineva et al., 2014), support vector machine (SVM) (Bermúdez et al., 2019), support vector regression (Gizaw and Gan, 2016) (Taherei Ghazvinei et al., 2018) and artificial neural network (ANN) (Kasiviswanathan et al., 2016).

An artificial neural network is a popular approach in flood prediction (Abbot and Marohasy, 2015) (Elsafi, 2014), to cope quantitatively with the complicated behavior of flood parameters. Some works have successfully applied ANN for forecasting water levels. Dawson and Wilby (2001) applied ANN to conventional hydrological models in flood-prone catchments in the UK in 1998. Since then, many studies about flood forecasts in catchment scales arose (Chang et al., 2018a) (Yu et al., 2006). Thirumalaih and Deo (1998) compared the water level forecast results along a river using backpropagation, conjugate gradient as well as cascade correlation. Coulibaly et al. (2000) combined Levenberg-Marquardt Backpropagation (Marquardt, 1963) with cross-validation to prevent the under-fitting and overfitting in daily reservoir inflow forecasting. Sattari et al. (2012) applied a backpropagation network and a time lag recurrent network, having reached a similar forecast accuracy in reservoir inflow. Humphrey et al. (2016) joined a conceptual rainfall-runoff model with a Bayesian artificial neural network for improving the accuracy of the neural network. Sit and Demir (2019) used discretized neural networks for the entire river network in Iowa. By including more location information, they could enhance the forecasting results. Bustami et al. (2007) applied a backpropagation ANN model for forecasting water level at gauging stations. Tiwari and Chatterjee (2010) compared different types of ANN predictions of water levels at gauging stations, namely a wavelet-based, a bootstrap-based, and a hybrid wavelet-bootstrap-ANN (WBANN). They showed that the WBANN model was more accurate and reliable than the other three ANNs. For flood inundation forecast, Simon Berkhahn et al. (2019) trained an ANN with synthetic events of spatial rainfall data for 2D urban pluvial inundation. Chang et al. (2018b) applied a mix of SVM and GIS analysis to expand point forecasts to flooded areas at a sub-catchment scale. Chu et al. (2020) proposed an ANN-based framework for flood inundation prediction based on single inflow data for a 20 m × 20 m grid resolution.

## 2.3. Model Structures for Urban Flood Modelling

Various hydraulic models have been developed for urban pluvial flood simulation in the recent 50 years, from one-dimensional to three-dimensional, simple to complex (Nkwunonwo et al., 2020). Based on these modeling techniques, flood risk approaches are evaluated historically in a deterministic manner (Di Baldassarre and Montanari, 2009). As the models become complex, the uncertainty from the dynamic, stochastic, and uncertain nature of the climate, hydrological, and river processes are omitted or simplified in the modeling processes. Uncertainty in flood modeling rises from the model inputs, model parameters, and model structures (Savage et al., 2016). In many cases, when modeling the same study area with different model structures, the model with the best fit may not always be the one with the most complex model structure (Butts et al., 2004). Therefore, a proper choice of the model structure would reconstruct the historical event with better accuracy.

There are several motivations to consider the model structures in flood simulations. First, it is a practical simplification to consider the main flow processes while ignoring the other processes in flood modeling. Second, as the flood modeling proceeds, new processes are integrated into the earlier models to represent the flow mechanisms. From this perspective, a quantify of the new model performance on the feedback of the more complex model structure is helpful for the new model evaluation (Atkinson et al., 2002) (Farmer et al., 2003).

Specifically, in urban flood modeling, the model processes are different in many ways. In the context of urban flood modeling, the sewer network drains massive amounts of water in urban flood events. Therefore, the dual-drainage concept is implemented in most of the urban flood models (Djordjević et al., 2005) (Maksimović et al., 2009)(Nasello and Tucciarelli, 2005)). Generally, the sub-surface drainage infrastructures are modeled independently from the surface systems (Chen et al., 2016). The modeling of model component combinations produces various model structures with different complexity.

The sewer drainage system has an important influence on urban flood modeling. By integrating the sewer drainage system, an urban flood model also models the processes of the channel flow in urban drainage systems. Therefore, the model considers the water-transporting and surcharging of the drainage system in flood events by implementing the dual drainage concept (Djordjević et al., 1999). Classified by the relation to the drainage types, the urban area can be divided into several types (Schmitt et al., 2004):

- surface areas
- single drainage areas
- distinct surface drainage components
- closed underground sewers

Surface areas behave differently in a flood event. Single drainage areas turn rainfall into surface runoff, affected by surface features. Distinct surface drainage components put the surface runoff to the underground sewer systems. Closed underground sewer represents sewer network facilities under the urban area. By connecting the surface and underground sewer networks, bi-directional connections allow water exchange between the surface and the sewer network.

Besides the dual-drainage concept, the 1D hydrodynamic model is also applied for modeling the inflow upstream in urban flood modeling. Compared to the urban area, the upstream catchment usually has a much larger area. As flood details in the upstream are far from the constructed area, detailed modeling with 2D would undoubtedly lead to an unaffordable computational expanse. The modeling of the upstream catchment with 1D modeling hence becomes an alternative. However, this approach gives the discharge volume at the 1D–2D interface as the inflow boundary condition of the 2D models. Consequently, the model structure uncertainty also increases as the model structure grows.

## 2.4. Model Uncertainties in Urban Flood Modelling

Uncertainties in hydrology studies are inevitable due to the unknown dynamics in surfaces and sub-surfaces processes (Mirzaei et al., 2015). Therefore, two steps are introduced into urban flood modeling, namely identification of model uncertainties and quantification of model uncertainties (Beven et al., 2015).

According to Merz and Thieken (2005), the model uncertainties could be classified into natural and epistemic uncertainties. Natural uncertainty originates from stochastic processes, while epistemic uncertainty arises from a lack of understanding of the modeled processes.

Data and boundary condition uncertainties appear widely in the hydrological modelings. River bed profiles and flow paths are usually simplified in mathematical representation. As a result, the boundary and initial conditions can hardly be measured and remain unknown. Besides, many hydrological cases have data scarcity problems, with some assumptions and interpolation to fulfill the model requirement. All these cases cause disagreement of the model from reality. The model uncertainty quantification is an exciting topic among researchers, as it helps to understand better the outputs from the models. Especially in flood risk assessment, five sources of epistemic uncertainties are significant. They, therefore, should be considered: the assessment of rainfall or snowmelt input, the flood occurrence frequency, the percentage of the contribution that the flood inputs to the flood runoff, the outline of the flood inundation, and the flood damage assessment (Beven et al., 2017).

In the previous studies, various model uncertainties have been investigated, e.g. model structures (Apel et al., 2009), model parameters (Bates et al., 2004), model inputs (Abily et al., 2016), and validation data (Stephens et al., 2012). In a study of conceptual models, Enge-

land et al. (2005) showed that model structure uncertainty has a more significant impact on the results than the model parameter uncertainty. Kavetski et al. (2002), and Chowdhury and Sharma (2007) studied model uncertainty by adding noise to input data and set up an empirical formulation for parameter errors.

Statistical probability methods and the quantification of uncertainty sources are applied in the modeling processes (Beven et al., 2015). One typical approach, sensitivity analysis, is to quantify the uncertain parameters with a more significant impact on the results (Hall et al., 2005). The sensitivity analysis deals with a single parameter at a time around a fixed value, while the outputs are inspected for their changes (Tsubaki and Kawahara, 2013). Another approach, the "global method," samples the parameters varying in the whole feasible spaces, while the outputs are checked for their ranges and distributions (Abily et al., 2016). Many previous studies on model uncertainty focus on the uncertainty from model parameters, especially the most significant uncertainty caused by frictions (Hall et al., 2005). However, since the uncertainty definition relies on statistical assumptions, the assessed uncertainties could also vary due to the different methods (Pappenberger et al., 2008).

In the bayersian approach, an input function takes a sampling over the model parameters to assess the uncertainty on the output. The method uses formal probability distribution for input samplings and introduces formal likelihood measures to obtain confidence intervals of the model outputs (Hall et al., 2011). Alternatively, the Generalized Likelihood Uncertainty Estimation (GLUE) applies a more flexible likelihood to obtain the confidence intervals of the model outputs (Beven and Binley, 1992). A lot of previous methods for probability flood inundation predictions have been developed using GLUE (Aronica et al., 2002) (Romanowicz and Beven, 2003).

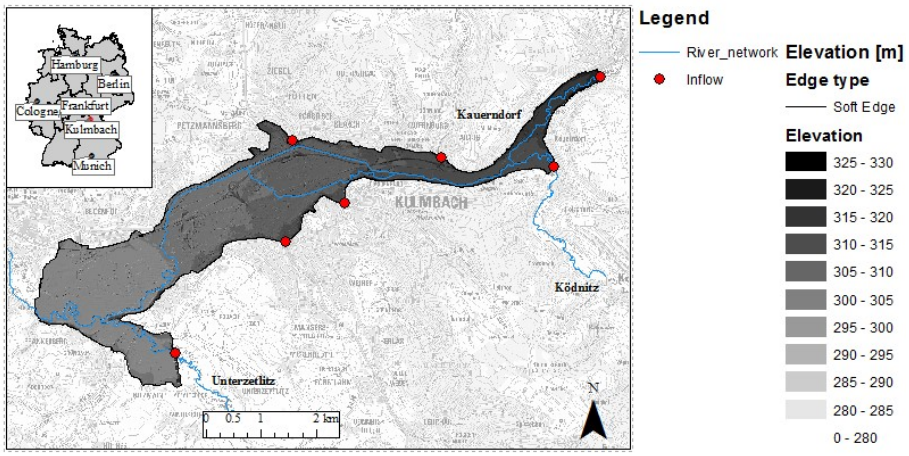
For the errors in the input and the model structure, an error model for temporal correlation of the residuals is also introduced for uncertainty quantifications (Yang et al., 2007), and Schaefli et al. (2007)). By introducing improved likelihood functions for input and model structure errors, integrated Bayesian uncertainty estimation (Ajami et al., 2007).

By comparison with five uncertainty analysis methods using Nash–Sutcliffe coefficient, Yang et al.(2008) concluded GLUE produces a good prediction of uncertainty bands. The GLUE approach was originated from the sensitivity analysis method by Hornberger and Spear (1981). Afterward, due to its easy concept and implementation, the GLUE is applied to many studies, e.g. flood inundation (Aronica et al., 2002) (Chau et al., 2005) (Goodarzi et al., 2011), rainfall-runoff modelling (Beven and Binley, 1992) (Freer et al., 1996) (Lamb et al., 1998), ground radar-rainfall estimation (Tadesse and Anagnostou, 2005), and distributed hydrologic modeling (McMichael et al., 2006) (Muleta and Nicklow, 2005). In a word, the GLUE is becoming a popular tool for accessing model uncertainties.

# 3. Study Areas

## 3.1. Kulmbach

Kulmbach is located in the middle of Upper Franconia in Bavaria by the river Main. The city consists of northern and southern parts split by the White Main crossing it. About 27 thousand inhabitants live in this city. The city has a population density of 292 inhabitants per $km^2$. The study area for flood prediction in Figure 1 has the area of 11.5 $km^2$. On May 28, 2006, Kulmbach was heavily flooded from the river and streams nearby. This event triggered decision-makers to review the initiatives of flood prevention for the city.




Blue curves: the river network
Shaded region: the study area
Red points: the seven inflows (three rivers and four smaller streams)

**Figure 1** Study area of Kulmbach

The Kulmbach catchment has an area of 1166.70 $km^2$ (by Gauge Mainleus, including Gauge Ködnitz 313.60 $km^2$, Gauge Kauerndorf 247.20 $km^2$, Gauge Unterzettlitz 496.30 $km^2$) with an elevation between 288 m and 529 m. The catchment has four land-use types: 80.7% vegetation area, 11.8% settlement area, 6.2% traffic area, and 1.3% water area. Seven streams contribute to this area, namely the Red Main, Schorgast, Dobrach, White Main, Kinzelsbach, Kohlenbach, and Mühlbach. A canal, Mühlbach, has been constructed in the city center for flood protection. On the upstream, Schorgast is the major tributary joining with White Main. In the downstream, Red Main and White Main are joined in the south of city Kulmbach, forming the river Main, the longest tributary of the Rhine. The White Main has an annual average discharge of 7.7 $m^3/s$ (Wasserwirtschaftamt Hof, 2012).

## 3.2. Simbach am Inn

Simbach am Inn is located in the district of Rottal Inn of the administrative region of Lower Bavaria in south-eastern Bavaria next to Austria. The river Simbach converges into the river Inn in the city with an annual average discharge of 0.3 $m^3/s$, which forms the local German-Austrian border (Wasserwirtschaftamt Deggendorf, 2012). The population of Simbach city is 9813 (LfStat, 2019). The city Simbach covers an area of 47.33 km $^2$ with a population density of 210 inhabitants per square kilometer. The flash flood event in 2016 has caused significant damages and losses to the city, which triggered the local authority's attention on flood prevention. For the event reconstruction, the measurement data collected by the University of Natural Resources and Life Sciences Vienna are used in the following studies (Hübl et al., 2017). In Simbach city, the water levels at 14 checkpoints are used to validate the different model structures for the goodness of fit to the historical event.

The Simbach catchment has an area of 36.38 $km^2$, with an elevation between 551 m and 334 m. In our study, the Tributary Area Antersdorfer Bach and Kirchberger Bach of 26.92 $km^2$ are modelled with 1D surface runoff model and the Simbach city center of 9.73$km^2$ is modelled with 1D-2D dual drainage model. The catchment has four different land-use types: 83.1% vegetation area, 12.1% settlement area, 4.5% traffic area, and 0.3% water area. The city lies in the south of the catchment. The catchment has into two tributaries: the Antersdorfer Bach and Kirchberger Bach. The average channel slope of Antersdorfer Bach is 1.3%, and the average channel slope of Kirchberger Bach is 1.2%. The two streams converge to the river Simbach. As a comparison, the average slope of Simbach in the city area has an average slope of 1%. In the south of the city, two water pumps control the floodplain of River Simbach, and one water pump controls the floodplain of River Erlach.

## 3.3. Baiersdorf

Baiersdorf lies in Middle Franconia in the northwest of Bavaria, between Erlangen and Forchheim. River Regnitz flows through the west of the city of Baiersdorf.

The Baiersdorf catchment has an area of 49.45 $km^2$ with the inhabitants more than 15 thousand. Its elevation varies between 256 m (west) and 548 m (east). The catchment has four land-use types: 77.32% vegetation, 13.85% settlement, 6.58% traffic, and 2.25% water. Within the Baiersdorf catchment, several satellite urban areas are distributed in the east, e.g., Poxdorf, Effeltrich, Hagenau, Igelsdorf, Langensendelbach, Bräuningshof, and Adlitz, while the city is in the south of the catchment. Only the flows from the adjacent hills contribute to the flooding and further leave the catchment into the River Regnitz.

An extreme flood event hit Baiersdorf on 21st, July 2007, causing significant disruption in transportation and a total loss of 86 million euros. Since then, Baiersdorf city has invested

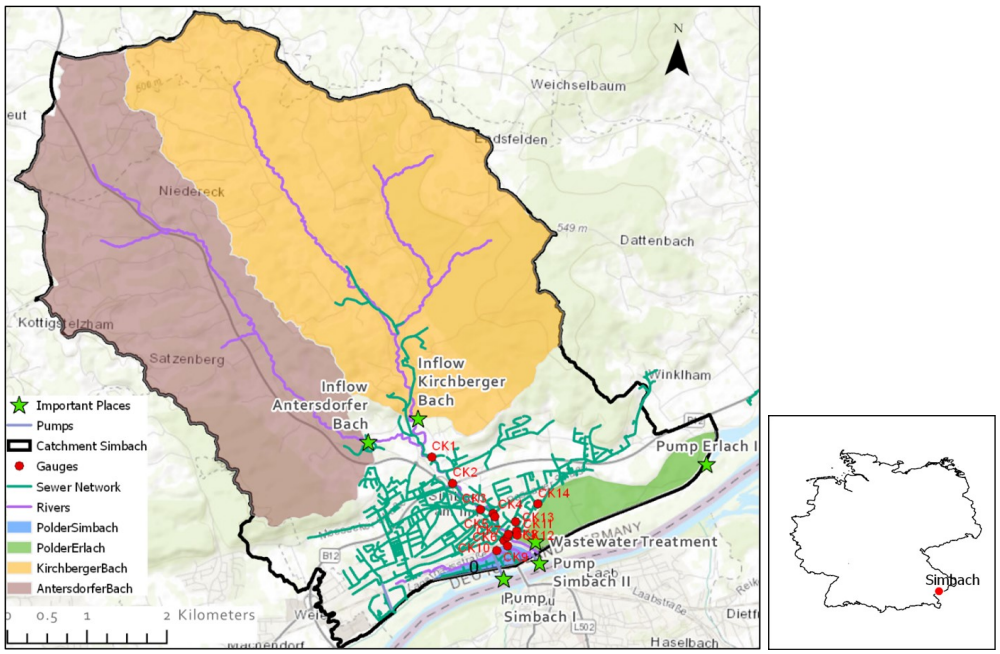



**Figure 2** Study area of Simbach

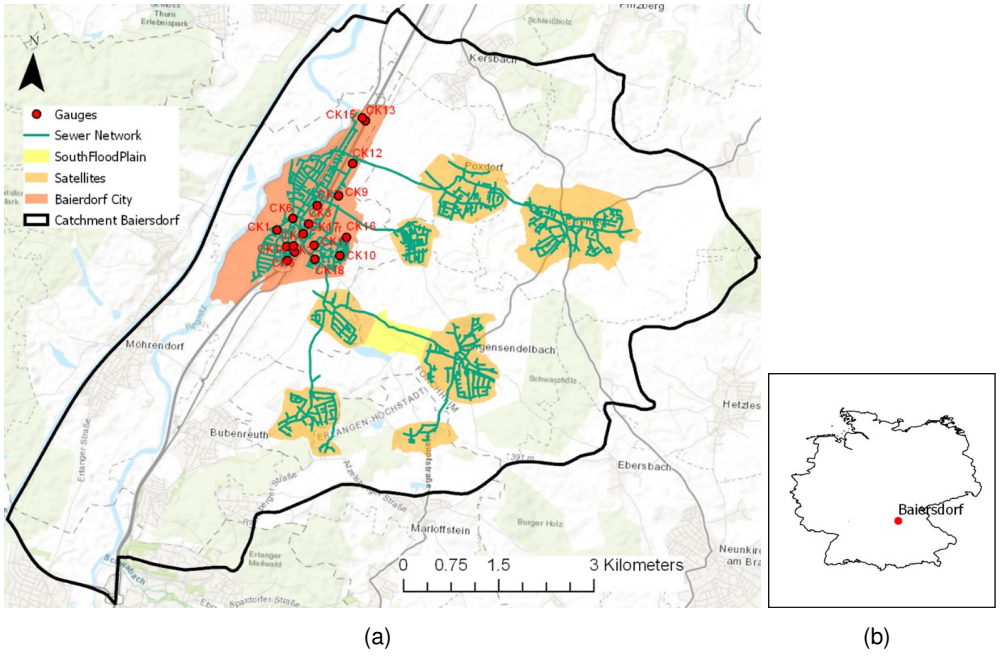



**Figure 3** Study area of Baiersdorf

more in flood protection measures. In Baiersdorf city, the water levels at 18 checkpoints from the flood event 2007 are used to validate the different model structures for the goodness of fit to the flood event.

# 4. Prediction of Maximum Flood Inundation Extents

## 4.1. Methodology

This section introduces the methodology for maximum water depth prediction of flood events. The methodology would start from the algorithm for neural network training. After that, we introduce the input and output patterns. Next, clustering methods will be discussed as a pre-processing for the training event database to test for better model performance. Finally, the model evaluation method is presented at the end of this section.

### 4.1.1. Artificial Neural Network and Resilient Backpropagation Algorithm

The ANN applied in this work for modeling the study area is a forward-feed neural network (FNN) (Nawi et al., 2007), producing and transmitting the data in a network structure. The fundamental element of the neural network is the neuron. Each neuron collects values from the previous layer by summing up the values multiplying the weight on each input arc, and storing the results on itself. Through multiple layered neurons, information is proceeded by the weights and transferred over the network, finally reaching the output layer. Seven inflows give the input layer upstream contributing to the urban area of Kulmbach from the event database (more details in Section 4.2). The output layer generates from the raster flood inundation map from the event database.

Backpropagation is an algorithm widely applied in neural network studies for optimizing the weights in forward-feed neural networks (Nawi et al., 2007). The procession consists of two phases: the training phase collects a part of data from the existing database, tuning the model by changing the weights on input arcs to minimize the bias on the output layer; the recalling phase produces the new outputs for the testing inputs. The rest individuals in the training dataset are used to evaluate the network performance. The total bias between the output of ANN and the observed values defines as the error function. The weights are modified automatically using the function below to reduce the error function in each iteration. The chain rule is applied here for minimizing the biases, namely written as:

$$\frac{\partial L}{\partial w_{ij}} = \frac{\partial L}{\partial O_i} \cdot \frac{\partial O_i}{\partial \mathrm{net}_i} \cdot \frac{\partial \mathrm{net}_i}{\partial w_{ij}} \tag{4.1}$$

where

$L$ is the is error function of the model.
$w_{ij}$ is weight from $i$ 'th neuron to $j$ 'th neuron.
$O_i$ is output of the model.
$\mathrm{net}_i$ is weighted sum of the inputs of neuron $i$.

$$w_{ij}(t+1) = w_{ij}(t) - \epsilon \cdot \frac{\partial L}{\partial w_{ij}} \tag{4.2}$$

where
$\epsilon$ is the learning rate taken as 0.01 in our training.

The learning rate scales the gradient in each iteration of the weight update. Therefore, it is critical to pick up the correct value. A high learning rate will miss the optimal point, while a low learning rate would slow the training process. Herein, we apply the gradient descent algorithm to calculate the update of the weights. To speed up the convergence of the iteration in Formula 4.2, resilient backpropagation as defined in (Saini, 2008) is applied, which treats the update of weights differently depending on the derivative of the error function. A higher alternative learning rate $\eta^+$ could speed up the iterations if the error gradient remains in the same direction in neighboring time-steps and a lower alternative learning rate $\eta^-$ when approaching the optimal weights.

$$\Delta_{ij}(t) = \begin{cases} \eta^+ \cdot \Delta_{ij}(t-1) & , \frac{\partial L}{\partial w_{ij}}(t) \cdot \frac{\partial L}{\partial w_{ij}}(t-1) > 0 \\ \eta^- \cdot \Delta_{ij}(t-1) & , \frac{\partial L}{\partial w_{ij}}(t) \cdot \frac{\partial L}{\partial w_{ij}}(t-1) < 0 \\ \Delta_{ij}(t-1) & , \text{else} \end{cases} \tag{4.3}$$

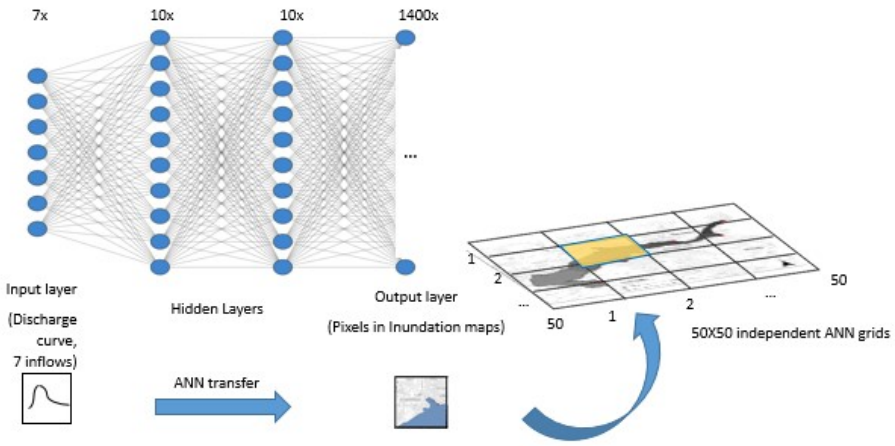
$$w_{ij}(t) = \begin{cases} w_{ij}(t-1) + \Delta_{ij}(t) & , \frac{\partial L}{\partial w_{ij}}(t) < 0 \\ w_{ij}(t-1) - \Delta_{ij}(t) & , \frac{\partial L}{\partial w_{ij}}(t) > 0 \\ 0 & , \text{else} \end{cases} \tag{4.4}$$

in which $0 < \eta^- < 1 < \eta^+$. In our study these were set constant and equal to $\eta^- = 0.5, \eta^+ = 1.2$.

Due to the total number of pixels (resolution of 4 m by 4 m) in the city of Kulmbach, a single hidden layer would exceed 365 thousand elements. Therefore, for a lower storage requirement and a shorter ANN model training time, the study area is subdivided into 50 × 50 squared grids, each grid having its independent ANN (the output layer has 1400 elements) (Figure 4). A similar strategy was used by Berkhahn et al. (2019) for an ANN for flood prediction having rainfall as input.

### 4.1.2. ANN Network Input and Output

The inflows of the seven streams, taken from the boundary conditions for the hydraulic model from the synthetic event database, are inputted for ANN training. The training–validation of




The study area is divided into 50 × 50 raster grids
Input layer: seven input hydrographs
Output layer: flood inundation extent in each grid

**Figure 4** Neural network setups for maximum inundation prediction

the 50 × 50 ANN aims to replace the hydraulic processes within the marked study area (see Figure 1). The ANN receives the same input discharge inputs, namely the seven hourly discharge inflow volumes. Each ANN aims to generate the inundation map for one sub-divided area, after training with the same inflows the inundation map within the sub-divided areas from the synthetic event database at a high resolution. Only the water depth is to be generated as the outputs.

### 4.1.3. Fuzzy C-Means Clustering and Principal Component Analysis Clustering

We apply clustering to the discharges training dataset to further enhance the ANN behavior. Therefore, we can reduce the size of the training dataset while still keeping the main representative events. This approach can reduce the training time and minimize the overfitting effects. Fuzzy C-means clustering (FCM) (Tilson et al., 1988) is a widely used clustering method, which avoids the deficit of the sub-clusters with certain similarities within its components (Mukerji et al., 2009). In FCM, every event is given a membership u, indicating the relation between the event and a certain cluster. If a membership is equal to zero, the event has nothing in common with a specific cluster; if the membership is one, the event is at the cluster center. Once a cluster is set up, the membership $u$ can be calculated by the following equations and the events' distances. For cluster $i$ and event $j$, the membership $u_{ij}$ is to quantify distances between events and cluster centres.

$$\begin{cases} u_{ij} \in [0,1] \\ \sum_{i=1}^{n} u_{ij} = 1, \quad 1 < j \leq n \end{cases} \tag{4.5}$$

$$u_{ij}^{(k)} = \frac{1}{\sum_{r=1}^{c} \left( \frac{d_{ij}^{(k)}}{d_{rj}^{(k)}} \right)^2} \tag{4.6}$$

$$d_{ij} = |x_j - v_i|^2 \tag{4.7}$$

$$v_i^{(k+1)} = \frac{\sum_{j=1}^{n} u_{ij}^{(k)} x_j}{\sum_{j=1}^{n} u_{ij}^{(k)}} \tag{4.8}$$

where
$c$ is number of clusters, $2 \leq c \leq n-1$.
$v_i$ is centroid of $i$ 'th cluster.
$d_{ij}$ is Euclidean distance between event $j$ and its corresponding centroid.

The total sum of distances between events and the cluster centroids has to be the minimum possible for optimal clustering. Therefore, the following objective function needs to be optimized:

$$J_m(U,V) = \sum_{j=1}^{n} \sum_{i=1}^{c} (u_{ij})^2 d_{ij} \tag{4.9}$$

Two approaches are applied for deciding the clustering parameters: (a) conventional clustering (by pre-selected hydrograph characteristic parameters); (b) dimension reduction methods. In the former, the clustering variables chosen were $P$(peak discharge value), $T$(peak time), $V$(total volume), $V_{24}$ (volume in the first 24 hours). These can be applied individually or combined. The latter clustering method is based on principal component analysis among the hydrographs. The data are projected to several principal eigenvectors for dimensionality reduction via PCA for further clustering with FCM. To determine the optimal number of clustering $c$, we define the clustering performance index $L(c)$.

$$L(c) = \frac{\sum_{i=1}^{c} \left( \sum_{j=1}^{n} u_{ij}^2 \right) \parallel v_i - \overline{x} \parallel^2}{\sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^2 \parallel v_i - x_j \parallel^2} \cdot \frac{n-c}{c-1} \tag{4.10}$$

The optimal cluster number c can be determined by the maximum $L(c)$.

### 4.1.4. Model Evaluation

To evaluate the performance of the ANN prediction of maximum flood inundation in the study area is based on the mean squared error (MSE) of each independent ANN. The inundation maps from the synthetic events produced using dynamic model HEC-RAS are the observed values for ANN validation. The synthetic events have been produced from FloodEvac-Tool (Bhola et al., 2018). The model has been validated (Bhola et al., 2019). As each grid has its independent training network, the MSE is evaluated using all the pixels in each grid.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(T-S)^2 \tag{4.11}$$

where
$T$ is the predicted value.
$S$ is the observed value.

To evaluate the overall behavior of the model across the training and validation data sets, the average MSE is evaluated for the average accuracy.

$$\mathrm{avg.\,MSE}^m = \frac{1}{N}\sum_{n=1}^{N}\mathrm{MSE}_n^m \tag{4.12}$$

where
$m$ is the grid index.
$n$ is the pixel index in a grid.
$N$ is the total number of pixels.

### 4.1.5. Hydrodynamic Model HEC-RAS 2D

The hydrodynamic model applied as a benchmark for ANN prediction network is HEC-RAS 2D. The model, originally developed by U.S. Army Corps of Engineers, has been applied broadly in diverse studies (Quirogaa et al., 2016) (Patel et al., 2017). 2D Saint-Venant and diffusive wave equations are integrated in the solver of HEC-RAS 2D. For generating the event database for the ANN predictions, the diffusive wave is chosen for its higher stability.The mass and momentum are calculated using Eq. 4.13 and 4.14, 4.15. The barotropic pressure gradient, equal to the bottom friction, drives the flow movement (Brunner, 2016).

$$\frac{\partial H}{\partial t} + \frac{\partial hu}{\partial x} + \frac{\partial hv}{\partial y} + q = 0 \tag{4.13}$$

$$g\frac{\partial H}{\partial x} + c_f u = 0 \tag{4.14}$$

$$g\frac{\partial H}{\partial y} + c_f v = 0 \tag{4.15}$$

$$c_f = \frac{g|V|}{M^2 R^{4/3}} \tag{4.16}$$

where $H$ is surface elevation [m].
$h$ is water depth [m].
$u$ and $v$ are the velocity in x-, y-direction [m/s].
$q$ is the source/sink.
$g$ is gravitational acceleration [m/(s$^2$)].
$c_f$ is bottom friction coefficient [/s].
$R$ is hydraulic radius [m].
$|V|$ is magnitude of velocity [m/s].
$n$ is Manning roughness coefficient [s/m $^{1/3}$].

## 4.2. Data Preparation

This study develops a method for predicting the maximum flood inundation in an urban area by backpropagation networks based on multiple inflow data for a grid resolution of 4 m × 4 m. Unlike most previous studies, this work focuses on applying ANN in an urbanized area to produce high-resolution flood inundation maps from river flooding. Thus, only the real-time discharges of the upstream catchments are needed to predict maximum flood inundation.

The synthetic event database is generated by the 2D hydraulic model HEC-RAS (Hydrologic Engineering Center - River Analysis System, Davis, CA, USA) for various rainfall intensities, distribution, duration (Bhola et al., 2018). The synthetic events are built following two stages. First, the hydrologic model LARSIM (Large Area Runoff Simulation Model) (Ludwig and Bremicker, 2006) calculates the discharge hydrographs into the city area. LARSIM is a hydrological model applied for flood forecasting at the Bavarian Environment Agency (Disse et al., 2018). Afterward, the 180 convective and advective events are simulated with the 2D hydraulic model HEC-RAS 2D to generate the flood inundation map database. The maps are generated with high temporal resolution (15 min) and projected to a spatial resolution of 4 m by 4 m. For further details of the generation of synthetic events, please refer to Bhola et al. (2018).

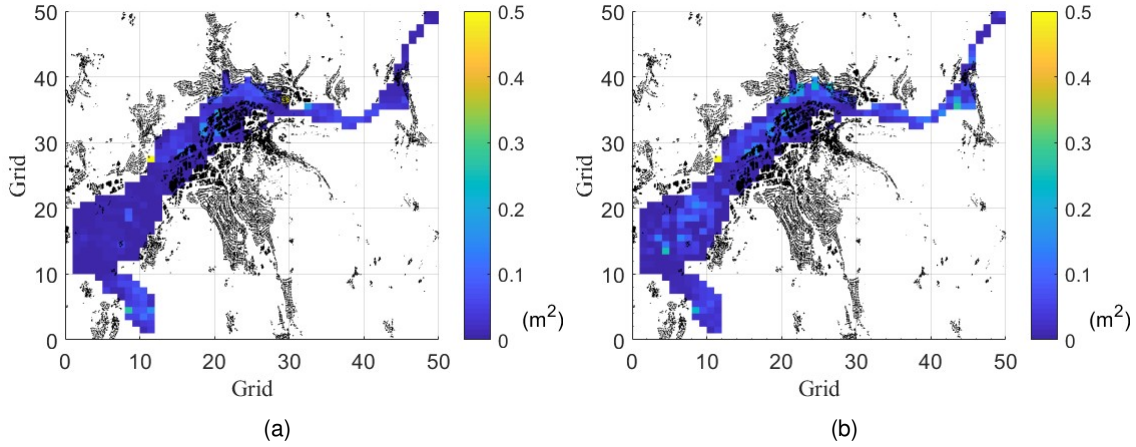
## 4.3. Results and Discussion

This section compares different network settings, and the better combinations are selected based on the model evaluation criteria given in Section 4.1.4. Different training algorithms,

the numbers of hidden layers and neurons, grid resolution, are tested for improving network performance. The results with and without pre-processing using FCM have been compared afterward with the selected network settings. The configured networks are finally applied to predict the synthetic and historical events.

### 4.3.1. Training Algorithm

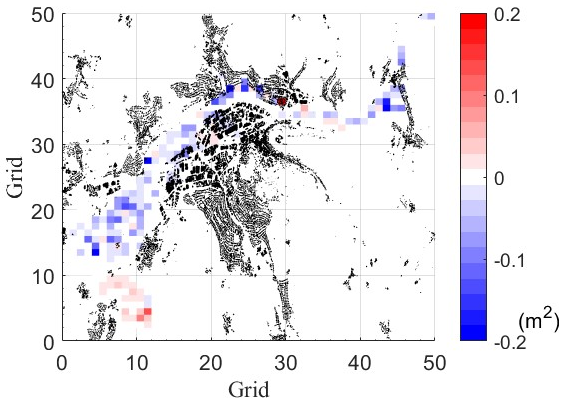
Two training algorithms are applied for training the ANN model using the same training dataset (Event #1~#120): resilient backpropagation (RP) and the conjugate gradient (CGF). After that, both generated models are evaluated using the MSE over the remaining runs (60) in the testing dataset (Event #121~#180) (see Figure 5). Figure 6 shows the MSE evaluated over the training dataset just for comparison purposes. In Figure 5(a)(b), most grids have the MSE lower than 0.2 $m^2$, showing that both RP and CGF networks behave well in general. Figure 6 shows the MSE from RP is mostly better than that of CGF.




Each grid is an ANN with houses in black (a) Average MSE among events by resilient backpropagation (RP) (b) Average MSE among events by the conjugate gradient (CGF)

**Figure 5** Model performances by two training algorithms in testing dataset

This section discusses the two training algorithms, resilient backpropagation and conjugate gradient. Figure 5 and Figure 6 support that both algorithms do not show overfitting. Indeed, similar MSEs from the training and testing datasets are visible. However, we still observe a few grids with higher MSEs, suggesting that increasing the size of the training datasets could further improve the performance. Figure 6 shows that the resilient backpropagation has a lower standard deviation of the MSE for the testing dataset compared to that of the conjugate gradient. Therefore, we choose resilient backpropagation as the training algorithm. This observation is in line with other researchers who also described resilient backpropagation as efficient with forward-feed neural networks (Bustami et al., 2007) (Chibueze and Nonyelum, 2009).




Each grid is an ANN with houses in black. Negative values indicate RP performs better than CGF with smaller MSE. 149 grids have positive values and 335 grids have negative values.

**Figure 6** Model performances by the two training algorithms in the training dataset

### 4.3.2. Number of Hidden Layers and Neurons

For better performance of the neural networks, the numbers of network layers and neurons per layer are optimized using the error function with the training dataset. The layer number is tested between two and six, while the neuron number is 10 to 30.

**Table 1** Number of grids in each combination (number of layers and neurons) which outperform the others

| | **Number of Layers** | | | | |
|---|---|---|---|---|---|
| **Number of Neurons** | **2** | **3** | **4** | **5** | **6** |
| **10** | 16.72% | 4.83% | 1.21% | 2.59% | 3.62% |
| **15** | 0.17% | 7.0% | 1.72% | 2.24% | 2.93% |
| **20** | 1.55% | 6.21% | 1.38% | 0.69% | 1.90% |
| **25** | 2.41% | 13.10% | 3.28% | 0.69% | 0.86% |
| **30** | 1.55% | 17.07% | 5.00% | 0.00% | 1.21% |

Gray: the optimal grid percentage larger than 13%.

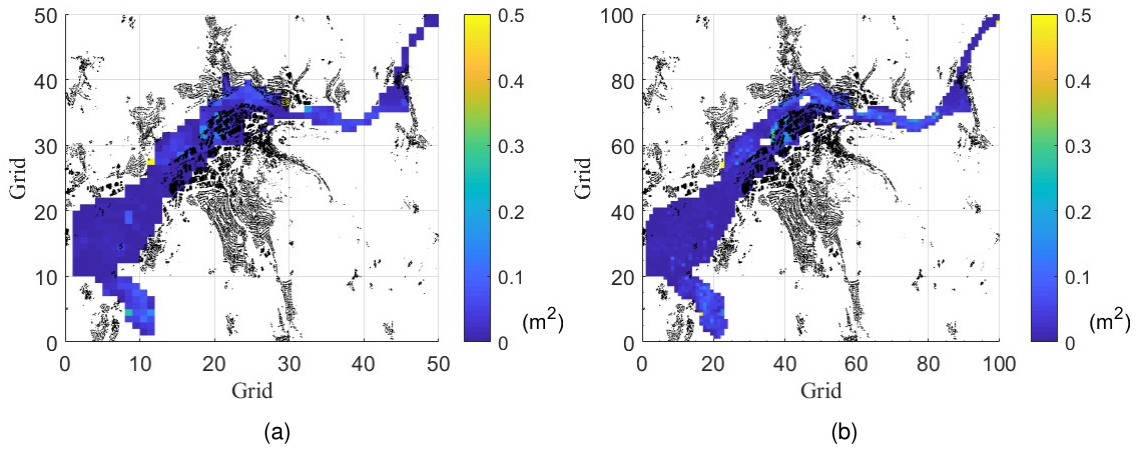
Table 1 shows the number of grids in each combination (number layers and neurons) which outperform all the others; it shows that 70% of all grids fall within the number of layers equal to two or three layers.

The number of hidden layers and the number of neurons have a decisive impact on the effectiveness of the neural network (Xu and Chen, 2008). Due to their critical importance, different combinations of hidden layer numbers and neuron numbers were tested for an optimal combination. The number of layers varies among two, three, four, five, and six, and the number of

neurons among 10, 15, 20, 25, and 30, amounting to a total number of combinations tested of 25. In Table 1, a grid is 'optimal', once the error function reaches the minimum. From the results, it is possible to conclude that most grids (70%) behave better with two or three hidden layers. However, no general trend is visible between the number of neurons and the number of hidden layers. The 'optimal' occurs over all the combinations of hidden layers and the number of neurons. When the layer number takes four, five, or six, it is possible to find widespread 'optimal' neuron numbers. For most cases (70%, where the layer number takes 2 or 3), the network with fewer neurons behaves better with only two hidden layers. However, the network with more neurons behaves better with three hidden layers. Overall, the complex dependency between hidden layers and neurons reflects the complexity in the input training datasets. Our study used two hidden layers with ten neurons per layer.

### 4.3.3. Grid Resolution

The grid resolution comparison aims to verify if a finer grid improves the prediction performances. Each grid is an individual ANN for the inundation prediction.




Black elements: houses
(a) Average MSE among events in 50 × 50 grids by RP (b) Average MSE among events in 100 × 100 grids by RP

**Figure 7** Model performance by numbers of ANNs in testing dataset

**Table 2** Impact of ANN number on training

| Grid size | Wet grid | Average MSE > 0.1 | (Average MSE > 0.1) % |
|---|---|---|---|
| **50 × 50** | 485 | 19 | 3.92% |
| **100 × 100** | 1679 | 82 | 4.88% |

Wet grid: number of grids which have water
Average MSE > 0.1: number of grids with average MSE > 0.1

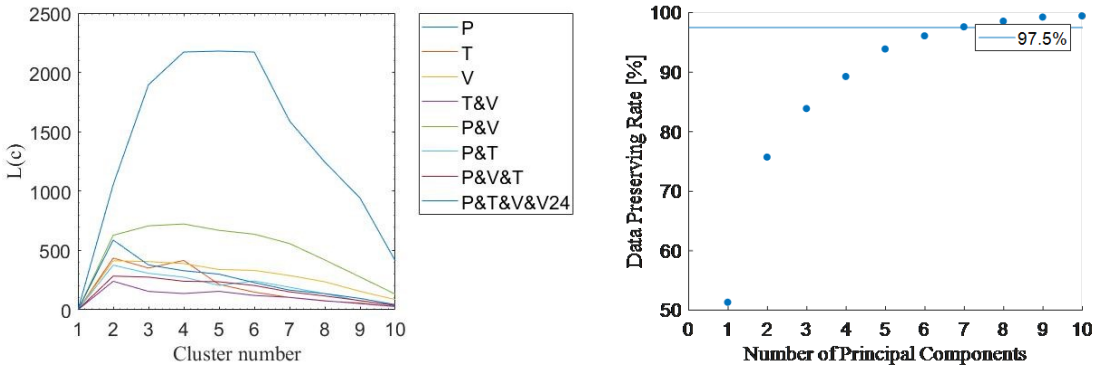
Two grid sizes are tested, namely one with 50 × 50 (each grid has 1400 pixels) and another

with 100 × 100 (each grid as 350 pixels) grids (Figure 7). The former has 2500 ANN networks, while the latter has 10000.

The results from the two grid size settings show the capability of the ANN to perform predictions with small MSE (Figure 7). However, from the comparison of error, the 50 × 50 performs better (see Table 2) and is computationally more efficient. Concerning the processing time, the 100 × 100 (finer grids) takes two hours more only to initialize the network due to its higher number of ANN. Furthermore, the 100 × 100 grid tends to have a larger MSE during the testing phases, indicating more overfitting than the solution with 50 × 50. Therefore, testing for different higher grids is unnecessary for over-fitting. On the other hand, a decrease in grids produces less precise results. Therefore, giving more weight to the training efficiency with less overfitting, our study decided to proceed with the 50 × 50 grid.

### 4.3.4. Fuzzy C-Means (FCM) Clustering

Here different results from clustering different sets of parameters obtained from the hydrographs are evaluated using the index $L(c)$ (Tilson et al., 1988). More significant numbers indicate better parameter settings.

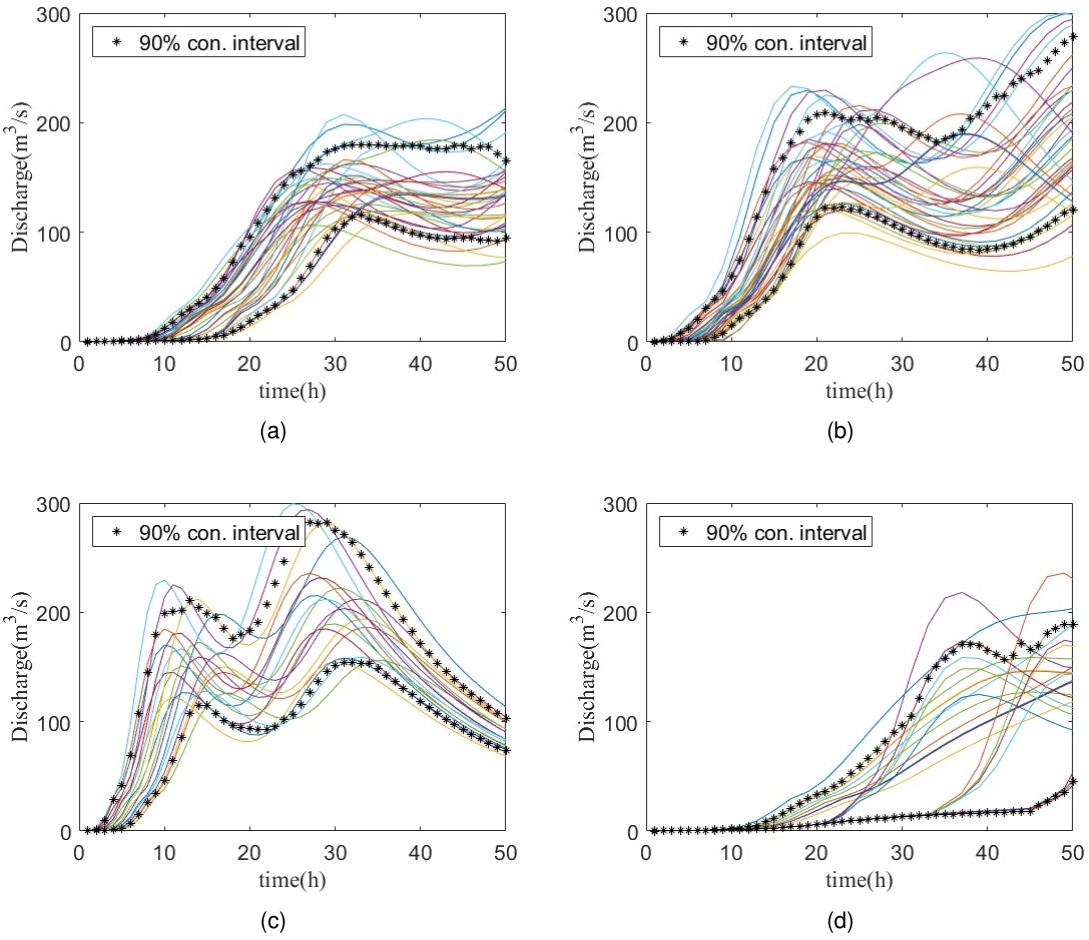



(a) (b)

(a) Conventional FCM criteria and their corresponding $L(c)$ values. $P$ (peak discharge value), $T$ (peak time), $V$ (total volume), $V_{24}$ (volume in the first 24 hours). (b) Data preserving rate in relation to the numbers of principal components. A minimum Data-preserving rate of 97.5% was selected as a good representation of the training database

**Figure 8** Determine cluster numbers and numbers of principal components

Figure 8(a) shows the relationship between the index $L(c)$ and the clustering number, with the selected parameters for fuzzy c-means clustering. Table 3 shows the spreads of the 90% confidence intervals of the clusters. According to Table 3, the clustering by four parameters ($P$, $V$, $V_{24}$, $T$) produces the minimum spread, which is the best clustering parameter combination for our studies. Figure 9 also shows that the 90% confidence interval of the four clusters by conventional FCM according to the parameter combination of ($P$, $V$, $V_{24}$, $T$) is the best choice. Besides the conventional FCM, FCM is also quantified based on principal component analysis (PCA-FCM). By choosing seven components for clustering, we can represent more than 97% of the original data (Figure 8(b)). Figure 10 shows the clustering results by

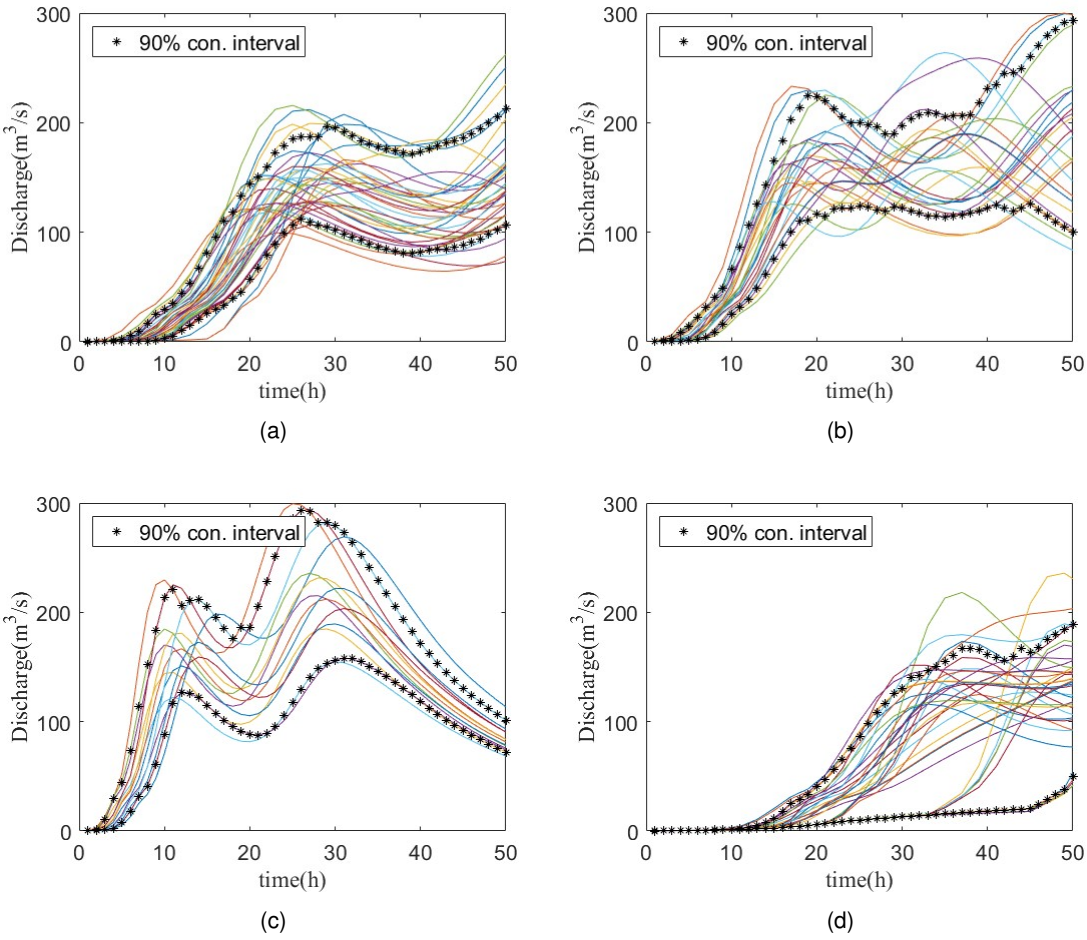



Stream Red Main (the biggest inflow) are clustered into four clusters (a–d)
The clustering is based on the combination of $P$(peak discharge value), $T$ (peak time), $V$ (total volume), $V_{24}$ (volume in the first 24 hours)
Asterisks represent the 90% confidence intervals.

**Figure 9** Clustering of discharge curves in training dataset by FCM

PCA-FCM. The comparison in Table 3 shows that PCA-FCM generates smaller integrals of the bandwidth area (i.e. the 90% confidence interval is shown in Figure 10) than those of the conventional FCM. The integral of the bandwidth area is a measure of the spread of the discharge curves in each cluster. An efficient clustering strategy will have a small spread. Therefore, this study introduces the PCA-FCM clustering. Table 4 summarizes the sign of the differences among the training results from 100 clustered events to those from original unclustered 120 events and a random unclustered 100 events.

In the conventional FCM by hydrograph characteristic parameters of $P$(peak value), $T$(peak time), $V$(total volume) and $V_{24}$ (volume in the first 24 hours), and their combinations were tested. Figure 8(a) and Table 3 show that the potent combination of all four parameters creates the most compact clusters evaluated by the performance index. In the second approach of PCA-FCM, we conducted PCA over the discharge curves (50 dimensions) to reduce their




Stream Red Main (the biggest inflow) are clustered into four clusters (a–d).
Asterisks represent the 90% confidence intervals

**Figure 10** Clustering of discharge curves in training dataset by PCA-FCM

dimension to the first ten orthogonal eigenvectors. The discharge inputs have a preserving rate of 97% (Figure 8(b)) when represented with the first seven eigenvectors. Therefore, we choose the first seven principal components for clustering into four groups in our research. Compared to conventional FCM, we observe that five out of seven streams have a smaller clustering spread (except Upper Schorgast and White Main), while the rest two are only slightly larger.

It is crucial to verify that the clustering strategy is efficient. Therefore, we compare it with three other strategies using the MSE:

- the original training dataset, consisting of the original 120 events in the training dataset from the synthetic database (RP120)
- randomly select the 100 events from the 120 events (RP100)

**Table 3** Integral of the bandwidth of 90% confidence intervals of discharge inputs with FCM and PCA-FCM

| Clustering Parameters | Red Main | Upper Schorgast | Dobrach | White Main | Kinzelsbach | Kohlenbach | Mühlbach |
|---|---|---|---|---|---|---|---|
| $P$ | 20886 | 11916 | 327 | 14275 | 327 | 327 | 656 |
| $T$ | 19808 | 10954 | 277 | 12624 | 277 | 277 | 551 |
| $V$ | 18298 | 10056 | 229 | 11337 | 229 | 229 | 457 |
| $P, V$ | 20208 | 10799 | 285 | 12744 | 285 | 285 | 569 |
| $T, V$ | 19358 | 10844 | 275 | 12433 | 275 | 275 | 550 |
| $P, T$ | 18826 | 10447 | 283 | 12353 | 283 | 283 | 565 |
| $P, V, T$ | 17964 | 9494 | 271 | 11452 | 271 | 271 | 539 |
| $P, V, V_{24}, T$ | 17844 | 9405 | 244 | 11030 | 244 | 244 | 489 |
| **PCA** | 17311 | 9799 | 221 | 11378 | 221 | 221 | 443 |

**Table 4** Sign of error difference between PCA-FCM clustered 100 events, original 120 events, and the randomly clustered 100 events

| 50h-Prediction | Number of Grids | |
|---|---|---|
| Avg.MSE RP100C-Avg.MSE RP120 | Positive | Negative |
| | 231 | 253 |
| Avg.MSE RP100C-Avg.MSE RP100 | Positive | Negative |
| | 188 | 296 |

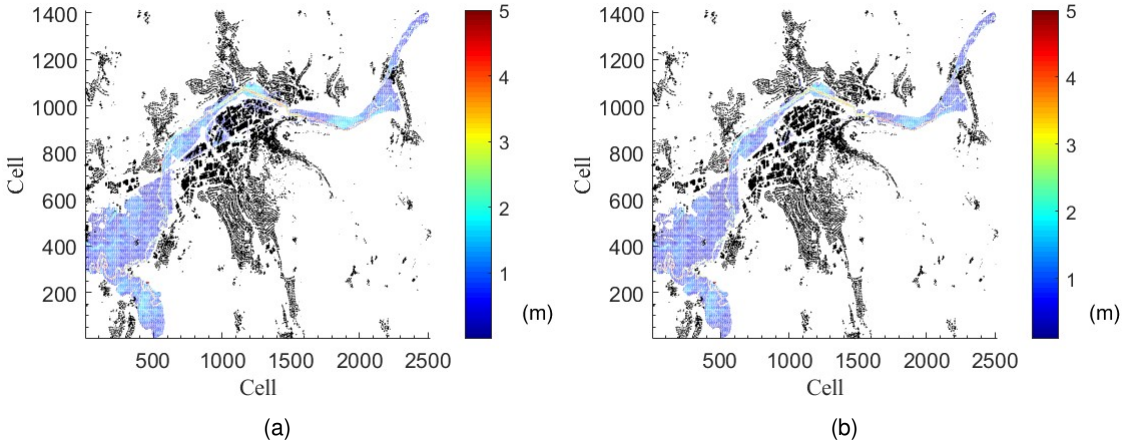
- randomly select 100 events from the four clusters (RP100C)

The results show that the RP100C behaves slightly better than RP120, which means we can achieve similar good predictions even with a clustered dataset with a smaller size. However, the RP100C behaves much better than RP100; this shows that clustered individuals perform better than random individuals with the same training database.

### 4.3.5. Maximum Flood Inundation Prediction for Synthetic Events

For the sake of representation of results, Figure 11 shows one example of the comparison of the flood inundation maps of one single event, Event 180. It is visible that the flood inundation maps from the ANN and the HEC-RAS database are very similar. The average and standard deviation of MSE of every testing event in the whole area is calculated to study the overall performance for 60 events in the testing set. In Figure 5(a), most of the area is displayed blue, showing that the MSE is close to 0.1 $m^2$. Overall, only 1.21% (7 out of 580) of total grids have their MSE over 0.2 $m^2$.

Figure 11 shows the validation of the results in the synthetic events using the whole testing




The plot shows an example for the prediction of Event # 180
Black parts: houses
(a) Inundation prediction from ANN model (b) Inundation prediction from the database

**Figure 11** Example of maximum flood inundation prediction in testing dataset

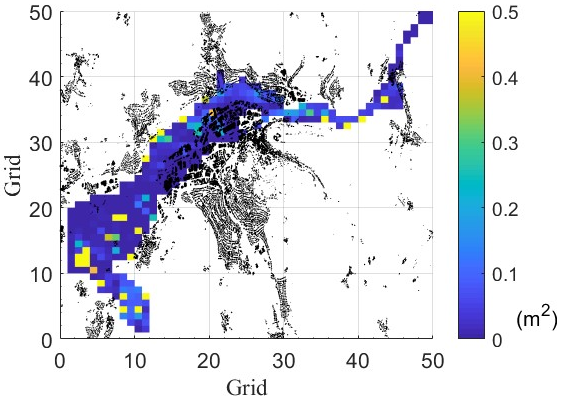
datasets. The majority of the predictions are accurate with MSE around 0.1 $m^2$. In any case, there are still more than 1.21% of grids with MSE larger than 0.2 $m^2$. Since the terrain elevation is relatively flat (city center), the impact of a highly variable terrain in the ANN predictions drops. This observation could also add to the good agreement found in our results. In any case, the results show clearly that the ANN prediction is bounded by the local topography, displaying a very similar inundation extent as the hydraulic model. Despite that, more significant errors can occur in the water depths, particularly in the southwest of our study area. However, the southwest is the farthest away from all the seven inflows (model inputs). Hence, we can anticipate that this area would be more challenging to predict by the ANN model.

### 4.3.6. Maximum Flood Inundation Prediction for Historical Events

The historical discharges from the historical events are taken from Bavarian Hydrological Services (Bhola et al., 2018). Two representative events are selected to validate the ANN from the historical events. Feb 2005 is an example of advective precipitation with lower peaks and longer duration, with an intensity of 2 – 3 mm/h. May 2013 is an example of convective precipitation with higher peaks and shorter duration, with an intensity of 5 – 60 mm/h.

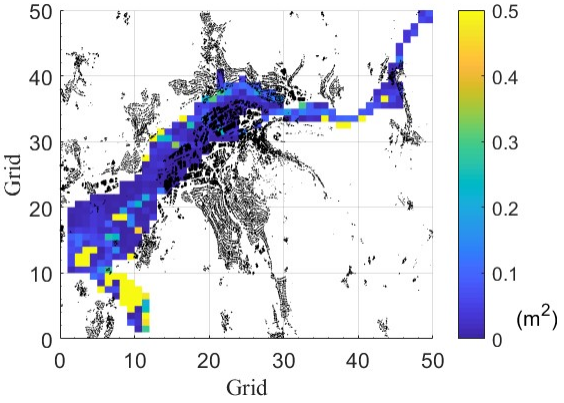
Figure 12 and Figure 13 show the MSE obtained for the prediction of the historical event in Feb 2005 and May 2013. In Figure 12, the large MSE occurs mainly in the ponding area to the southwest. On the other hand, figure 13 shows a larger MSE in the southwest than that in Feb 2005.

Finally, the developed ANN predicts two actual events, 2005 and 2013. In the results, the grids with larger MSE than 0.2 $m^2$ are 8.97% in 2005 and 13.62% in 2013. It shows that our ANN provides an accurate prediction on water depth for more than 85% of the grids




The hydrodynamic result is taken as the benchmark for comparison
Each grid is an ANN with houses in black

**Figure 12** The error of ANN prediction of the Historical Event Feb 2005




The hydrodynamic result is taken as the benchmark for comparison
Each grid is an ANN with houses in black

**Figure 13** The error of ANN prediction of the Historical Event May 2013

for the actual events. In both of the synthetic events, we observed that the large MSE part occurs in the southwest of the study area (see Figure 12 & Figure 13). Similar behavior also occurs during the testing phase (Figure 5), since it is the area further away from the significant inflows (see Figure 1). As the distance to the inflows (model inputs) increases, the growing uncertainty causes the water depths prediction to deviate from the observed data. However, the inundation extent is always well predicted. As in the results of the synthetic events, the ANN can accommodate the flooded volume within very similar topographic limits as the hydraulic model.

# 5. Prediction of Multi-Step Flood Inundation

## 5.1. Methodology

This section describes the methodology for a multi-step flood inundation forecast. The data and network structures are first presented. The method for hyperparameter tuning is then explained. Finally, the artificial neural network is first trained and tested for the first interval of the event (the forecast interval starting from time zero).

### 5.1.1. Data and Network Structure

Artificial neural networks are algorithms applied to map features into a series of outputs. Through a structure of the input, output, and intermediate hidden layers, artificial neural networks can learn data relationships between input and output data (Sit et al., 2020). A feed-forward neural network is applied in this work for modeling the study area, proceeding, and transmitting data in a network structure (Nawi et al., 2007). One of the most widely used ANN is the multilayer perceptron (MLP) (Sankaranarayanan et al., 2019). The MLP consists of highly interconnected neurons organized in layers to process information. The neurons in one layer are fully connected to the next layer. Each connection is then assigned a weight. Each neuron collects values from the previous layer by summing up the values multiplying the weight on each input arc, and storing the results on itself. An activation function transfers the results from the hidden layers to the output layer, and a loss function is applied to measure the fit of the neural network to a set of input-output data pairs.

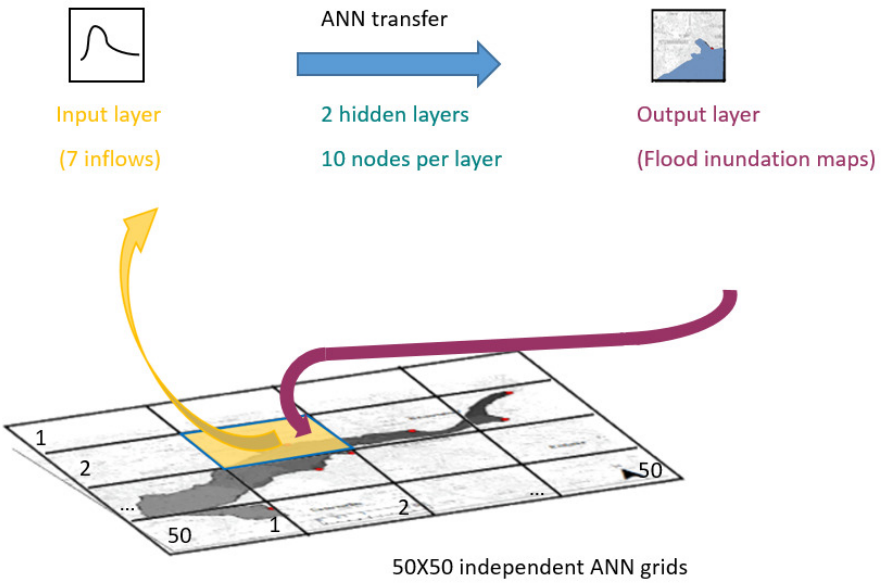



**Figure 14** The forward-feed neural network setup for the multistep forecast

The input layer collects the seven inflows to the urban area of Kulmbach, given the hourly

discharge intensity. The output layer takes the hourly raster inundation maps with a 4 m × 4 m from the event database. The ANN has two hidden layers with ten nodes per layer between the input and the output layers. The network training uses One hundred twenty synthetic events from the event database (see Section 4.2). Afterward, the model validates the other 60 events in the event database. The $2/3$ of the date for training and $1/3$ for testing is a division found in the literature (Kohavi, 1995). Finally, the model is applied to forecast three historical events. The activation function takes the widely applied sigmoid function for the neural network (Jhong et al., 2018). Due to the high-resolution of the map (4 m by 4 m), weights between the last hidden layer and the output layer would have been 1 GB RAM with a dimension of the problem of more than 30 million. The optimization of these weights is very time-consuming, even with the latest optimization techniques (Wang et al., 2013). Hence, a "divide and conquer" strategy enables calculation in a single PC. In principle, the results of the two strategies should be the same. Some alternative artificial network structures, such as a convolutional neural network (CNN), could not be applied in this study. The network size would require a massive number of hyper-parameters, which was beyond the memory capacity of a personal computer for forecasting purposes (Kabir et al., 2020).

Furthermore, the time for training can be reduced with our strategy due to parallelization. The estimated time for training all networks in parallel with four cores is six h. For shorter training time and lower memory requirements, the study area is divided into 50 × 50 squared grids (see Figure 14). A similar idea of splitting has been applied to a former study (Berkhahn et al., 2019). Each grid had four independent ANNs for intervals (3 h, 6 h, 9 h, and 12 h). In total, 10,000 ANNs are trained to produce multistep forecasts.

### 5.1.2. Hyperparameter Tuning in ANN

To optimize weights in ANN, resilient backpropagation is a widely applied effective algorithm (Saini, 2008).

According to Shamim et al. (2007) and Panda et al. (2010), backpropagation neural networks outperform other methods in flood forecasting studies for their more efficiency and higher robustness. Berkhahn et al. (2019) compared the training algorithms for hyperparameter tuning. The study showed that resilient backpropagation is more efficient than backpropagation and Levenberg–Marquardt for maximum flood inundation prediction. The process has two stages: the training stage gathers information from the flood event database, changing the weights between layers to minimize the error on the output layer; the recalling stage generates the forecast for the rest of the events in the database for testing the model.

Formula 5.17 and Formula 5.17 show the scheme of a resilient backpropagation. The gradient descent algorithm is applied to calculate the update of the network weights $w_{ij}$ from $i$'th neuron to $j$ 'th neuron, . It distinguishes the update of weights upon the derivative of the model's loss function $L$. The loss function $L$ takes the mean square error (MSE). The iteration stops once the loss function reaches its minimum (chosen $10^{-6}$ in this case).

$$\Delta_{ij}(t) = \begin{cases} \eta^+ \cdot \Delta_{ij}(t-1) & , \frac{\partial L}{\partial w_{ij}}(t) \cdot \frac{\partial L}{\partial w_{ij}}(t-1) > 0 \\ \eta^- \cdot \Delta_{ij}(t-1) & , \frac{\partial L}{\partial w_{ij}}(t) \cdot \frac{\partial L}{\partial w_{ij}}(t-1) < 0 \\ \Delta_{ij}(t-1) & , \text{else} \end{cases} \tag{5.17}$$

$$w_{ij}(t) = \begin{cases} w_{ij}(t-1) + \Delta_{ij}(t) & , \frac{\partial L}{\partial w_{ij}}(t) < 0 \\ w_{ij}(t-1) - \Delta_{ij}(t) & , \frac{\partial L}{\partial w_{ij}}(t) > 0 \\ 0 & , \text{else} \end{cases} \tag{5.18}$$

The learning rate is to scale the speed in each weight updating iteration. The larger alternative learning rate $\eta^+$ is chosen when the error gradient in the same signal in neighbouring iterations and lower alternative learning rate $\eta^-$ when the loss function is close to zero, fulfilling $0 < \eta^- < 1 < \eta^+$. In our study, these were set constant and equal to $\eta^- = 0.5$, $\eta^+ = 1.2$.

### 5.1.3. Prediction of the First Interval of Flood Events

The ANN model is trained with the first 120 events in the synthetic flood event database (more details in Section 4.2). The training extracts the time series of each event (starting from time 0).

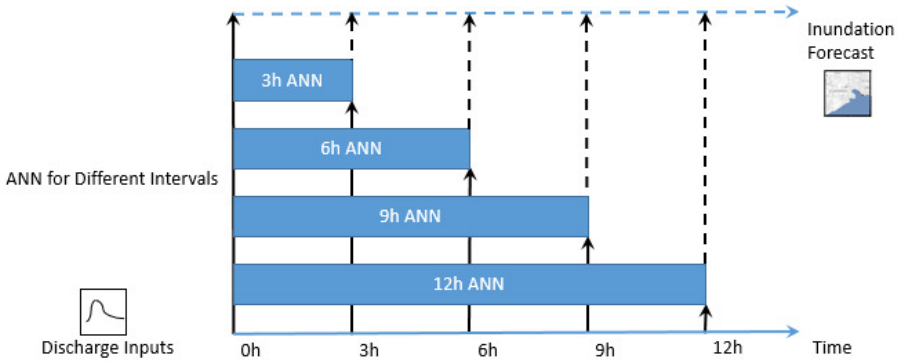


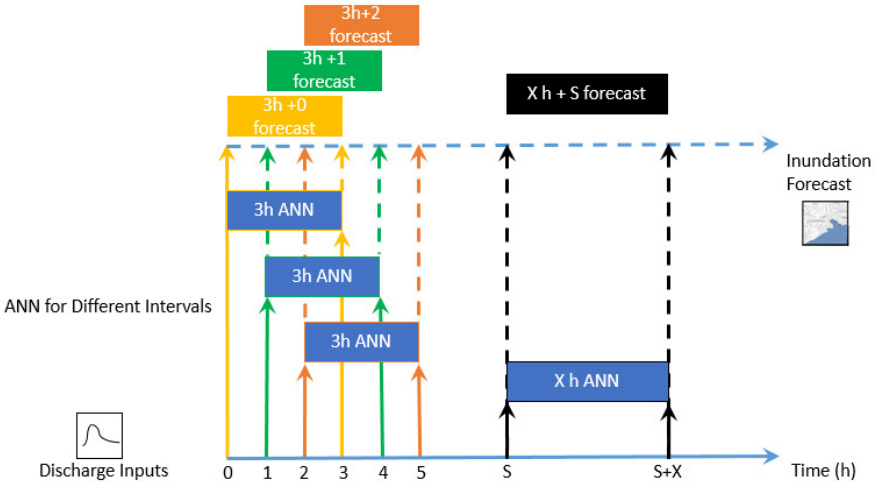

**Figure 15** Training of ANN forecast model using first intervals

The input inflow discharges are extracted from time 0 to $X$ h ($X$ takes the values from 3, 6, 9, 12). The output layer takes the output inundation maps at $X$ h ($X$ takes the values from 3, 6, 9, 12). The training of the networks uses the intervals of 3 h, 6 h, 9 h, and 12 h of the flood events with the exact forecast lead times (see Figure 15). The ANN models only consider the input flow values from the initial time step (blue bars of the events in Figure 15), but not from the previous time steps. The continuous forecast concept was similar to the approach in the framework FloodEvac, which successfully produced forecasts based on the selection of pre-recorded flood maps (Bhola et al., 2018).

After the training, the first interval forecast tests the rest of 60 events in the synthetic database.

### 5.1.4. Real-Time Forecasting for Multistep Forecast Intervals

The flood forecast starts when a specific discharge reaches the forecast threshold in this work. If the start point occurs sometime later at time $x$, the beginning of prediction is also shifted to time $x$ accordingly. Finally, if all the discharge inflows fall below the forecast threshold, the forecast is stopped. With this setup, the forecast can run in continuous mode.




Yellow: Forecast 0 – 3 h
Green: Forecast 1 h – 4 h
Orange: Forecast 2 h – 5 h
Black: General case of forecast $S$ h – $X$ h $+S$

**Figure 16** Shift of ANN forecast models for multistep forecasts

The ANN receives the corresponding discharge inputs of an interval, just as in real-time forecasts. After the forecast is complete for a particular step, the discharge forecast repeats one hour ahead. The forecasts are done with the same ANN model, starting one hour later, taking the discharge inputs from the next interval. This procedure repeats many times to enable the continuous mode of flood forecasting. In this case study, the real-time forecast proceeds with the ANN models trained to forecast at multiple steps of 1—5 h. Figure 16 shows the forecast from time 0 and the shift forward of the forecast intervals by one hour and two hours.

For a more straightforward interpretation of the different forecast groups, we call each forecast the "$X$ h $+S$" forecast. The "$X$ h" indicates the forecast interval of $X$ hours, and the "+ $S$" behind it shows the start time of the forecast.

### 5.1.5. Model Evaluation

The root-mean-square error (RMSE) is applied to access the ANN forecast performance in the study area. The forecasts of the ANN are compared against the inundation maps produced by the 2D dynamic model (see Section 4.2). Hence the 2D dynamic model results

are assumed as the observed values to enable the evaluation of the ANN. The events in the database have been processed by the FloodEvac tool (Bhola et al., 2018) and validated (Bhola et al., 2019).

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(T-S)^2} \tag{5.19}$$

where
$T$ is the predicted value, water depth from the ANN model in our case. $S$ is the observed value, water depth from the hydraulic model (HEC-RAS) in our case. To assess the general conduct of the model over the training and validation dataset, the average RMSE is also calculated for the average accuracy in the testing dataset.

The following indices measure the correspondence between the ANN model and the hydraulic model to quantify the forecast of inundation extent growth, namely probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI) (Li et al., 2009).

$$\mathrm{POD} = \frac{\mathrm{hits}}{\mathrm{hits} + \mathrm{misses}} \tag{5.20}$$

$$\mathrm{FAR} = \frac{\mathrm{false\ alarms}}{\mathrm{hits} + \mathrm{false\ alarms}} \tag{5.21}$$

$$\mathrm{CSI} = \frac{\mathrm{hits}}{\mathrm{hits} + \mathrm{misses} + \mathrm{false\ alarms}} \tag{5.22}$$

A pixel with water depths under 10 cm is defined as a dry pixel, while over 10 cm as a wet pixel. Hits count the wet pixels by the ANN forecast and the hydraulic simulation. Misses counting the predicted pixels dry by the ANN model but simulated wet by the hydraulic model. False alarms count the predicted wet by ANN model but simulated dry by hydraulic model.

## 5.2. Results and Discussion

Section 5.2.1 focuses on assessing the results for the same time interval for ANN training. Section 5.2.2 focuses on the subsequent multistep forecast intervals. The aim is to verify the hypothesis that the same ANNs can be used to forecast subsequent multistep intervals successfully even though they are from the first interval.

### 5.2.1. Water Depth Prediction of the First Intervals

Synthetic Flood Events

The ANN model is tested with the 60 synthetic flood events from the FloodEvac tool. First, the ANN model for predicting the first intervals of flood events is set up for 3 h, 6 h, 9 h, and 12 h, using the discharge within the same time as the model input. After this, the RMSE with the testing dataset (Event # 121 to Event # 180) evaluates the prediction performance using the first intervals of flood events. Finally, the averaged RMSE was calculated for different prediction times (3 h, 6 h, 9 h, and 12 h) to quantify the prediction performance of each ANN. Table 5 shows the percentage of the accurate prediction ANNs, classified by RMSE of 0.2 m, 0.3 m, and 0.4 m. For example, in Table 5, with the error threshold of 0.3 m, the accuracy can be considered excellent with values above 80% for all the prediction durations.

**Table 5** Numbers of grids with large errors and their percentages

| Prediction time [h] | Wet ANN grid | 1 Doorstep Height | | 1.5 Doorsteps Height | | 2 Doorsteps Height | |
|---|---|---|---|---|---|---|---|
| | | ANN Grid with average RMSE > 0.2 m | ANN Grid% with average RMSE $\leq$ 0.2 m | ANN Grid with average RMSE > 0.3 m | ANN Grid% with average RMSE $\leq$ 0.3 m | ANN Grid with average RMSE > 0.4 m | ANN Grid% with average RMSE $\leq$ 0.4 m |
| 3 | 300 | 47 | 84.33% | 18 | 94.00% | 10 | 96.67% |
| 6 | 417 | 174 | 58.27% | 78 | 81.29% | 27 | 93.53% |
| 9 | 474 | 106 | 77.64% | 37 | 92.19% | 15 | 96.84% |
| 12 | 483 | 50 | 89.65% | 12 | 97.52% | 7 | 98.55% |

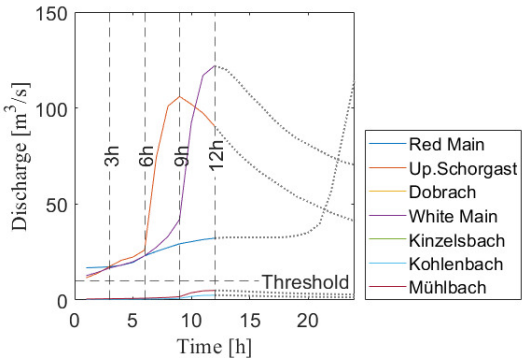
Gray: percentages larger than 70%.

Overall, the prediction performance of the first intervals of flood events shows that more than 80% of the grids have errors smaller than 0.3 m. The accuracy table (see Table 5) shows that the ANN has good accuracy in the water depth prediction of first intervals for all the synthetic events in the testing dataset. This test validated the network structure and the resilient backpropagation for solving the ANN. Since the network is initially trained with the 120 events from the synthetic event database and validated with the rest 60 events, the training events and the validation events bear more similarities, which could explain the good performance of the prediction of the first interval of the ANN.

Historical Flood Events

After testing with the synthetic events, the ANN model performance is further tested with the historical flood events. Afterward, the grid RMSE is calculated to evaluate prediction accuracy on the three historical flood events of their first intervals. A value of 10 $m^3/s$ was selected as the forecast threshold to initiate the forecasts in all three historical events. The forecast threshold is chosen slightly over the average discharge of 9.2 $m^3/s$ of White Main

(LfU, 2020) to avoid the low discharges from triggering flood warnings.




Seven discharge curves of three rivers and four streams are shown in different colors
Time 0: the start of the prediction
Dash lines: different discharge sections for prediction inputs

**Figure 17** Hydrograph of Event 2006

**Table 6** Numbers of wet grids and accurate grid percentage for event 2006

| Prediction time [h] | Wet ANN grid | 1 Doorstep Height | | 1.5 Doorsteps Height | | 2 Doorsteps Height | |
|---|---|---|---|---|---|---|---|
| | | ANN Grid with average RMSE > 0.2 m | ANN Grid% with average RMSE ≤ 0.2 m | ANN Grid with average RMSE > 0.3 m | ANN Grid% with average RMSE ≤ 0.3 m | ANN Grid with average RMSE > 0.4 m | ANN Grid% with average RMSE ≤ 0.4 m |
| 3 | 280 | 46 | 83.57% | 20 | 92.86% | 6 | 97.86% |
| 6 | 405 | 84 | 79.26% | 42 | 89.63% | 25 | 93.83% |
| 9 | 474 | 134 | 71.73% | 64 | 86.50% | 36 | 92.41% |
| 12 | 483 | 157 | 67.49% | 85 | 82.40% | 47 | 90.27% |

Gray: percentages > 70%
Wet grid: water level > 0.1 m

Historical Flood Event 2006 Figure 17 shows the discharge inputs for the historical flood event in 2006. The first 3 h, 6 h, 9 h, and 12 h discharge curves are given to the trained ANN as in Section 5.2.1. Figure 18 compares the prediction of the inundation map of the first intervals of 3 h, 6 h, 9 h, and 12 h with the inundation map from the hydraulic model of the historical flood event 2006. Table 6 shows the performance of the prediction for historical event 2006, evaluated by average RMSE for each ANN. As the forecast interval increases from 3 h to 12 h, the prediction accuracy drops, evaluated by grid percentages of RMSE.

From Table 6, 83% of grids have RMSE smaller than 0.2 m, and the rest of the grids have RMSE around zero in 3 h prediction. For the 6 h prediction, 79% of grids have RMSE less than 0.2 m. In the 9 h and 12 h predictions, the area with significant errors grows slightly

(see Table 6). From Figure 18, the inundation maps from 3 h and 6 h predictions match well with the hydraulic inundation maps. The 9 h and 12 h are less accurate, especially in the southwest of the study area, which is further away from the location of the discharge inflows and is, thus, likely less sensitive to the changes in the discharge inputs. In brief, the prediction for flood event 2006 is precise, with more than 82% of grids having RMSE less than 0.3 m.

Historical flood event 2013 Figure 19 shows the discharge inputs for the historical flood event in 2013. It shows that the initial discharge curves are below the forecast threshold of 10 $m^3/s$; therefore, the start of the prediction at 9 h is marked with a red line when one discharge hits the forecast threshold. Figure 20 compares the prediction of the inundation map of the first intervals of 3 h, 6 h, 9 h, and 12 h with the inundation map from the hydraulic model of the historical flood event 2013. Table 7 shows the performance of the prediction for historical event 2013, evaluated by average RMSE for each ANN. The forecast performance is slightly better than that of the event 2006.

For the historical flood event 2013, the discharge forecast threshold for the forecast start was reached later, signaling that the discharge forecast threshold is indeed effective in starting and stopping the forecast. Therefore, the start of the forecast is picked up at a later moment in time once one discharge crosses the forecast threshold of 10 $m^3/s$ for a second time. The red line in Figure 19 marks the new start of the forecast (red line). This event is nine hours later after the first forecast signaled by the ANN. Table 7 shows that the ANN model achieved high accuracy for the flood event in 2013. For the 3 h prediction, 96% of the grids have RMSE less than 0.2 m. 6 h prediction has 82% grids with RMSE less than 0.2 m. From 3 h and 6 h prediction of event 2013, the ANN performs better than the event 2006. Overall, the event of 2013 is also well predicted, with over 78% of grids having RMSE less than 0.3 m. Similar to event 2006, the predicted flood inundation maps of 3 h and 6 h intervals are similar to the hydraulic inundation simulations (see Figure 22).

Historical flood event 2005 Figure 21 shows the discharge inputs for the historical flood event in 2005. Figure 22 compares the prediction of the inundation map of the first intervals of 3 h, 6 h, 9 h, and 12 h with the inundation map from the hydraulic model of the historical flood event 2005. Table 8 shows the prediction performance of historical event 2005, evaluated by average RMSE for each ANN. As the forecast interval increases from 3 h to 12 h, the prediction accuracy drops. The grid percentage of RMSE can evaluate this.

### 5.2.2. Multistep Water Depth Predictions for Realtime Forecast

Historical flood event 2006 Table 9 shows the forecast for multistep forecast intervals of the event in 2006. The forecast for the event in 2006 has good accuracy for all the intervals.

Table 9 shows the forecast accuracy of the historical flood event in 2006. The forecast of

**Table 7** Numbers of wet grids and accurate grid percentages for the flood event in 2013

| Prediction time [h] | Wet ANN grid | 1 Doorstep Height | | 1.5 Doorsteps Height | | 2 Doorsteps Height | |
|---|---|---|---|---|---|---|---|
| | | ANN Grid with average RMSE > 0.2 m | ANN Grid% with average RMSE $\leq$ 0.2 m | ANN Grid with average RMSE > 0.3 m | ANN Grid% with average RMSE $\leq$ 0.3 m | ANN Grid with average RMSE > 0.4 m | ANN Grid% with average RMSE $\leq$ 0.4 m |
| 3 | 285 | 9 | 96.84% | 2 | 99.30% | 2 | 99.30% |
| 6 | 405 | 72 | 82.22% | 27 | 93.33% | 8 | 98.02% |
| 9 | 474 | 134 | 71.73% | 65 | 86.29% | 25 | 94.73% |
| 12 | 483 | 175 | 63.77% | 104 | 78.47% | 56 | 88.41% |

Gray: percentages > 70%
Wet grids: water levels > 0.1 m

**Table 8** Numbers of wet grids and accurate grid percentages for the flood event in 2005

| Prediction time [h] | Wet ANN grid | 1 Doorstep Height | | 1.5 Doorsteps Height | | 2 Doorsteps Height | |
|---|---|---|---|---|---|---|---|
| | | ANN Grid with average RMSE > 0.2 m | ANN Grid% with average RMSE $\leq$ 0.2 m | ANN Grid with average RMSE > 0.3 m | ANN Grid% with average RMSE $\leq$ 0.3 m | ANN Grid with average RMSE > 0.4 m | ANN Grid% with average RMSE $\leq$ 0.4 m |
| 3 | 280 | 65 | 76.79% | 36 | 87.14% | 19 | 93.21% |
| 6 | 405 | 165 | 59.26% | 115 | 71.60% | 74 | 81.73% |
| 9 | 474 | 216 | 54.43% | 148 | 68.78% | 93 | 80.38% |
| 12 | 483 | 244 | 49.48% | 168 | 65.22% | 107 | 77.85% |

Highlighted in gray: the percentages > 70%
Wet grid: water level > 0.1 m

Event 2006 shows a good accuracy in 3 h, 6 h, and most of 9 h (over 70% grid with RMSE < 0.3 m). However, it is visible that as the multistep forecast shifts further away from the original start used for the model training ($X$ h + 5 h forecasts in Table 9), the ANN model performance decreases.

Historical flood event 2013 Table 10 shows the forecast for multistep forecast intervals of the event in 2013. The forecast of Event 2013 has good accuracy for all the intervals, with a similar performance as Event 2006.

The discharge forecast threshold of the flood event from 2013 exceeds shortly at the beginning. Hence, the forecast is deactivated and reactivated again when the discharge exceeds

**Table 9** Sub-sequential forecast accuracy percentages for the flood event in 2006

| Starting point[h] | Prediction interval [h] | | | |
|---|---|---|---|---|
| | 3 | 6 | 9 | 12 |
| +1 | 98.93% | 92.10% | 83.12% | 79.09% |
| +2 | 98.94% | 90.86% | 77.85% | 79.92% |
| +3 | 96.94% | 89.38% | 76.58% | 78.26% |
| +4 | 95.11% | 86.95% | 70.89% | 75.36% |
| +5 | 86.60% | 69.29% | 66.88% | 68.12% |

Gray: percentages > 70%

**Table 10** Sub-sequential forecast accuracy percentages rate for the flood event in 2013

| Starting point[h] | Prediction interval [h] | | | |
|---|---|---|---|---|
| | 3 | 6 | 9 | 12 |
| +1 | 99.30% | 92.59% | 84.18% | 72.67% |
| +2 | 98.95% | 89.88% | 78.90% | 70.39% |
| +3 | 96.30% | 89.17% | 75.74% | 67.91% |
| +4 | 94.17% | 82.51% | 68.78% | 67.29% |
| +5 | 91.28% | 77.34% | 66.46% | 66.67% |

Gray: percentages > 70%.

the forecast threshold of 10 $m^3/s$ for the second time, namely 9 hours after time 0 (the first time the forecast window was activated). From the second starting point, all other forecasts for every forecast for $X$ h + 1 h to $X$ h + 5 h. Table 10 shows the forecast accuracy of the historical flood event of 2013. From this table, the ANN model performs similarly to the flood event in 2006. The forecast of the 2013 flood event has good results in 3 h, 6 h, and most of 9 h (over 70% of the grid with RMSE < 0.3 m).

Historical flood event 2005 Table 11 shows the forecast of multistep forecast intervals of the event 2005. The 3 h forecast of event 2005 still has a good accuracy of over 70%. However, the ANN model produces less accurate results for the 6 h, 9 h, 12 h forecasts.

Table 11 shows the accuracy percentage of the grids evaluated by the RMSE less than 0.3 m. With the changing of the forecast starting point, all the forecasts of different intervals have similar RMSE as the forecast of the first intervals. It is noticeable that the model provides a good forecast (over 70% grids with RMSE < 0.3 m) for 3 h intervals for all starting points.

**Table 11** Sub-sequential forecast accuracy percentages for the flood event in 2005

| Starting point[h] | Prediction interval [h] | | | |
|---|---|---|---|---|
| | 3 | 6 | 9 | 12 |
| +1 | 83.74% | 69.95% | 66.89% | 63.15% |
| +2 | 81.67% | 68.23% | 64.77% | 61.70% |
| +3 | 76.31% | 67.00% | 63.08% | 60.25% |
| +4 | 74.85% | 66.50% | 60.97% | 60.25% |
| +5 | 70.97% | 61.08% | 58.65% | 60.25% |

Gray: percentages > 70%

This observation shows that 3 h ANN trained for the first interval is applicable to forecast subsequent intervals with a slight drop in the overall accuracy. However, the forecasts of 6 h, 9 h, and 12 h show a poor performance (Table 11). Similar to the other events, most of the errors occur in the southwest of the study area (Figure 22). However, in this particular event, the errors are substantially more significant at the southwest; hence it causes poor performance of the ANNs.

In all three historical events, the forecast accuracy decreases as the forecast interval increases from 3 h to 12 h (see Tables 9 – 11). One exception occurs in the event 2006 between 9 h and 12 h, where the 12 h forecast has higher accuracy than the 9 h forecast. From the discharge curve (see Figure 17), unlike in other events, the two major discharges are falling after the peak value, which could be the reason for the higher accuracy at 12 h in this case.

### 5.2.3. Forecast of Inundation Extents

Figure 23 evaluates the ANN performance of the forecast of inundation extent (water depths over 0.1 m) growth with three indices. The status of wet/dry produces the following three indices for the likelihood between ANN and the hydraulic model. The probability of detection (POD) represents how well the ANN forecasted the same inundation extent as the hydraulic model. The false alarm ratio (FAR) measures the discrepancy of the ANN forecast to the hydraulic model. The critical success index (CSI) is the ratio between the correct forecasted inundation and the join of both the inundations (hits + misses + false alarms), showing the general correctness of the flood extent forecast of the ANN model. According to the verification criteria in another study (Bernhofen et al., 2018), CSI over 0.7 (see Figure 23) is considered a good fit for the benchmark, and over 0.5 (see Figure 23) is a sufficient fit. The lines in the figures show how these three indices change as the multistep forecasts advance in time.

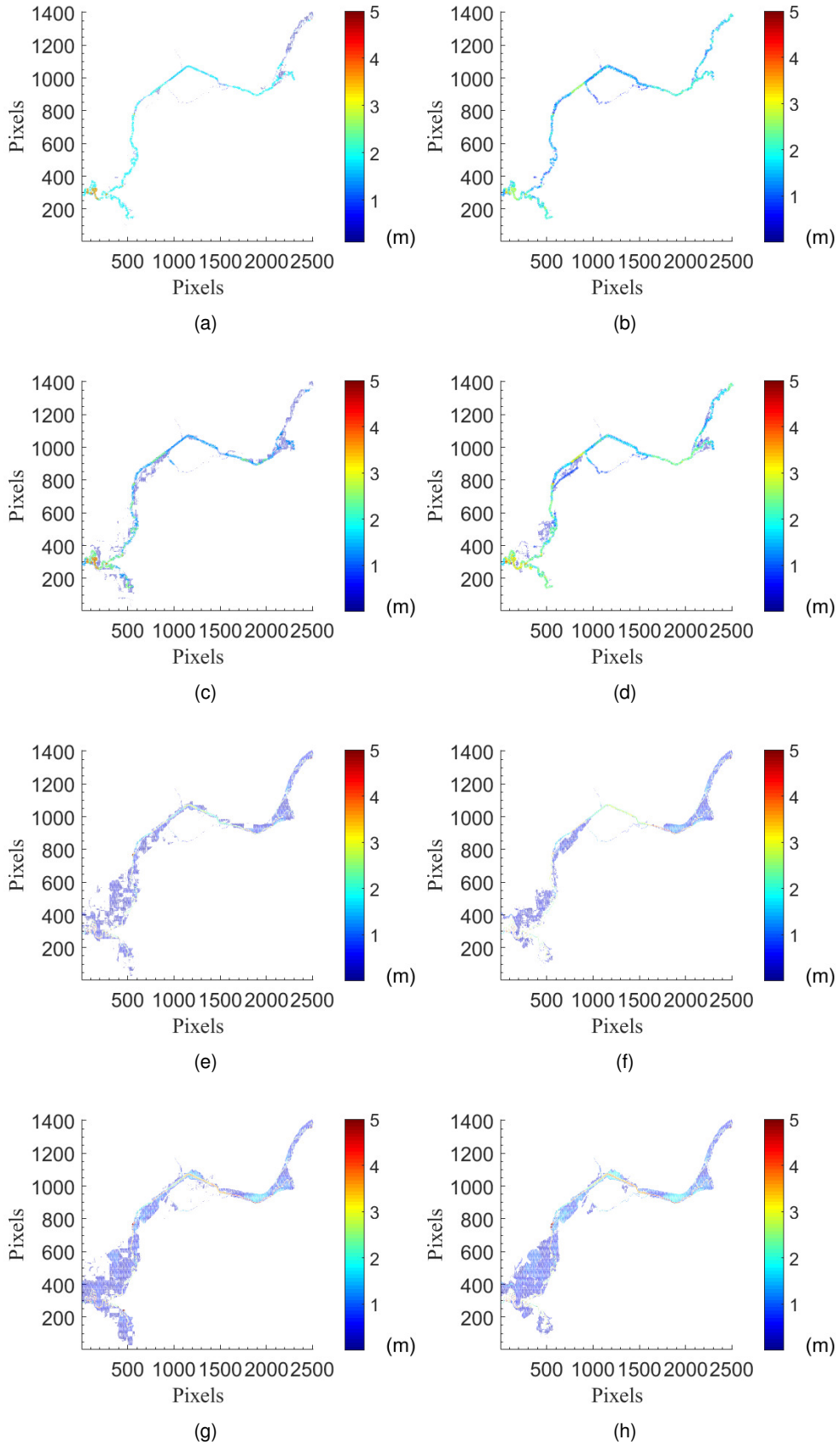
The statistical analysis proposed by Li et al. (2009) evaluates the forecasts of flood inundation

extent growths. Figure 23 shows three indices, POD, FAR, and CSI, for measuring the forecast performance of the flood inundation extent. Analyzing the POD index (see Figure 23 a, d, g), it is clear that for the 3 h ANN forecast, the accuracy decreases slightly as forecasts proceeds from 3 h + 0 h to 3 h + 5 h. In other words, the accuracy of the 3 h ANN network is more sensitive to the shift of the forecast intervals than the 6 h, 9 h, and 12 h ANN networks. The 3 h network achieves the best forecast performance for the first interval (training interval same as the forecast intervals). When moving forward for multistep forecast, shifting each hour decreases the POD by a value that varies between 0.08 to 0.1. This value means that 8% to 10% of the inundation extent displayed by the hydraulic model is missing in the ANN forecast. In any case, except for the event of 2005, the POD exhibits values above 70% for the first 2 hours of the forecast.

The FAR index (see Figure 23 b, e, h) indicates the false-alarm percentage of the ANN forecasted flood inundation extents. In all three events, it is noticeable that the area percentage with false-alarms decreases in all the forecast networks when the forecast interval moves forward. It shows the ANN forecast produces more percentage of false alarms at the early stage in a flood event, concerning the flood inundation is relatively small at the beginning. As a result, the number of outer pixels is more significant than that of inner pixels, causing a higher number of false alarms. Moreover, the decreasing trends of POD and FAR show that the ANN model tends to change from overestimation to underestimation when the forecast starts to shift from 0 — 5 h.

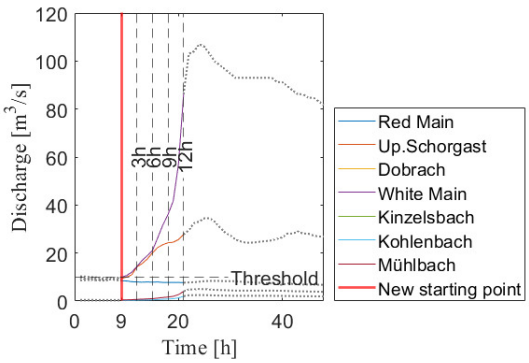
The CSI index (see Figure 23 c, f, i) shows the percentage of agreement of the ANN forecasts of the flood inundation extent to the hydraulic model. The ANN predicts better the flood inundation extents for two events of 2006 and 2013 than for 2005. The CSI from 2006 and 2013 is close to 0.6. For 2005, the CSI is around 0.4 showing a poor accuracy forecast in the flood inundation extent forecast.

It is noteworthy to mention that the ANN shows an expectable performance of water depth prediction decreasing with lead time (see Tables 9 – 11). However, if we focus on the inundation extent, it seems contradictory, as the 12 h prediction shows good performance (see CSI and POD in Figure 23). The latter apparent contraction could be due to the flood inundation extent confined by the topography. The topography limits the size of the inundation, making it easier for the ANN to predict it better.



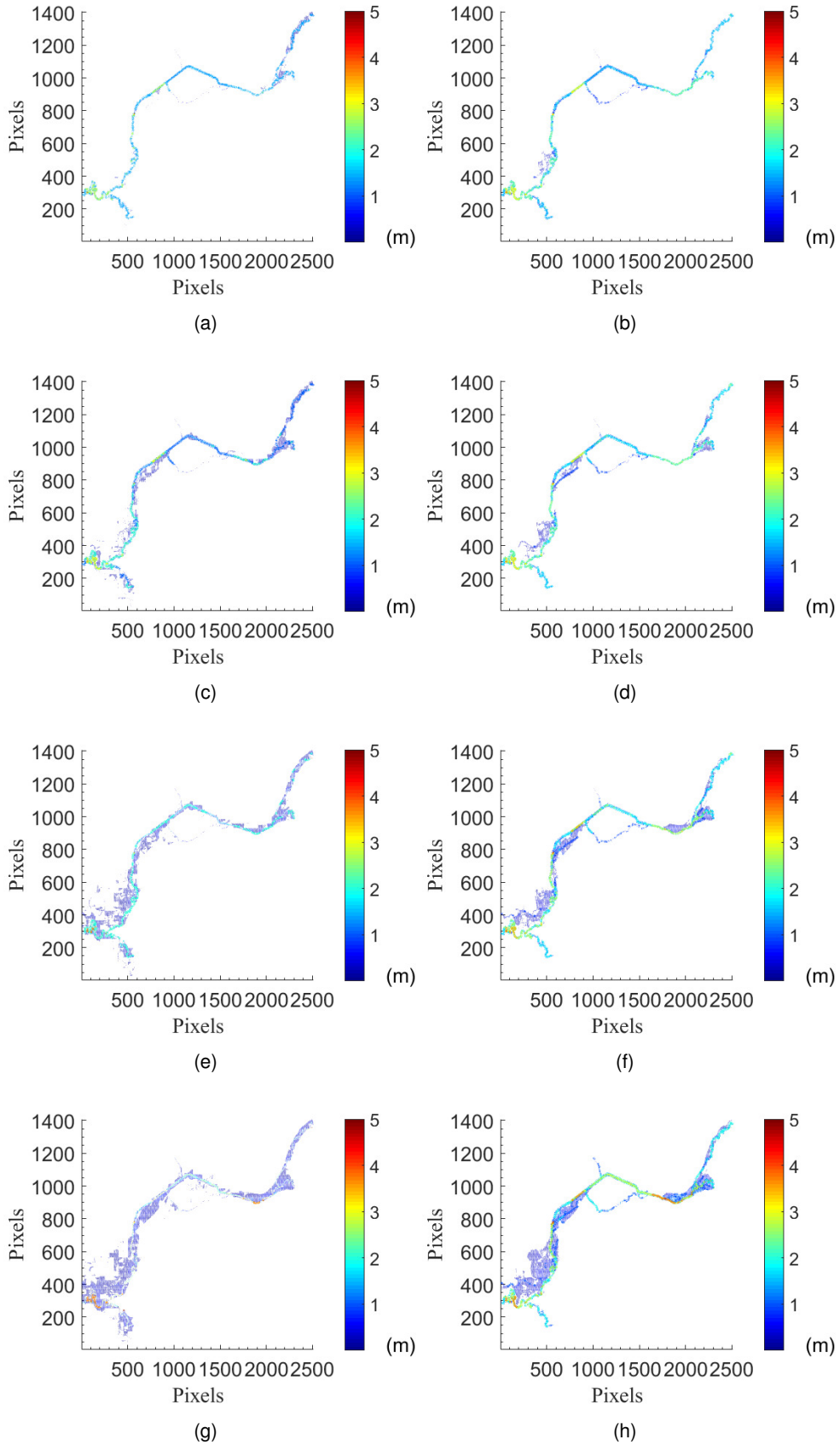

(a) ANN 3 h (b) hydrodynamic 3 h (c) ANN 6 h (d) hydrodynamic 6 h (e) ANN 9h (f) hydrodynamic 9 h (g) ANN 12 h (h) hydrodynamic 12 h

**Figure 18** Predicted inundation maps of the first intervals of Event 2006



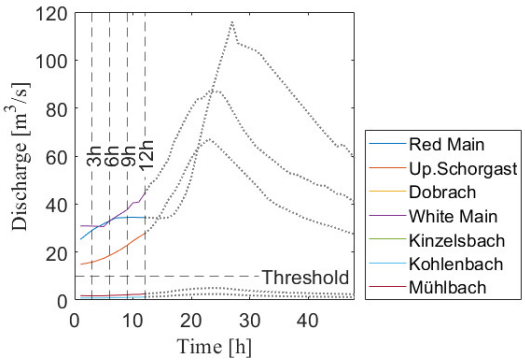

Seven discharge curves of three rivers and four streams are shown in different colors
Red line: new start of the prediction
Dash lines: different discharge sections for prediction inputs

**Figure 19** Hydrograph of Flood Event 2013



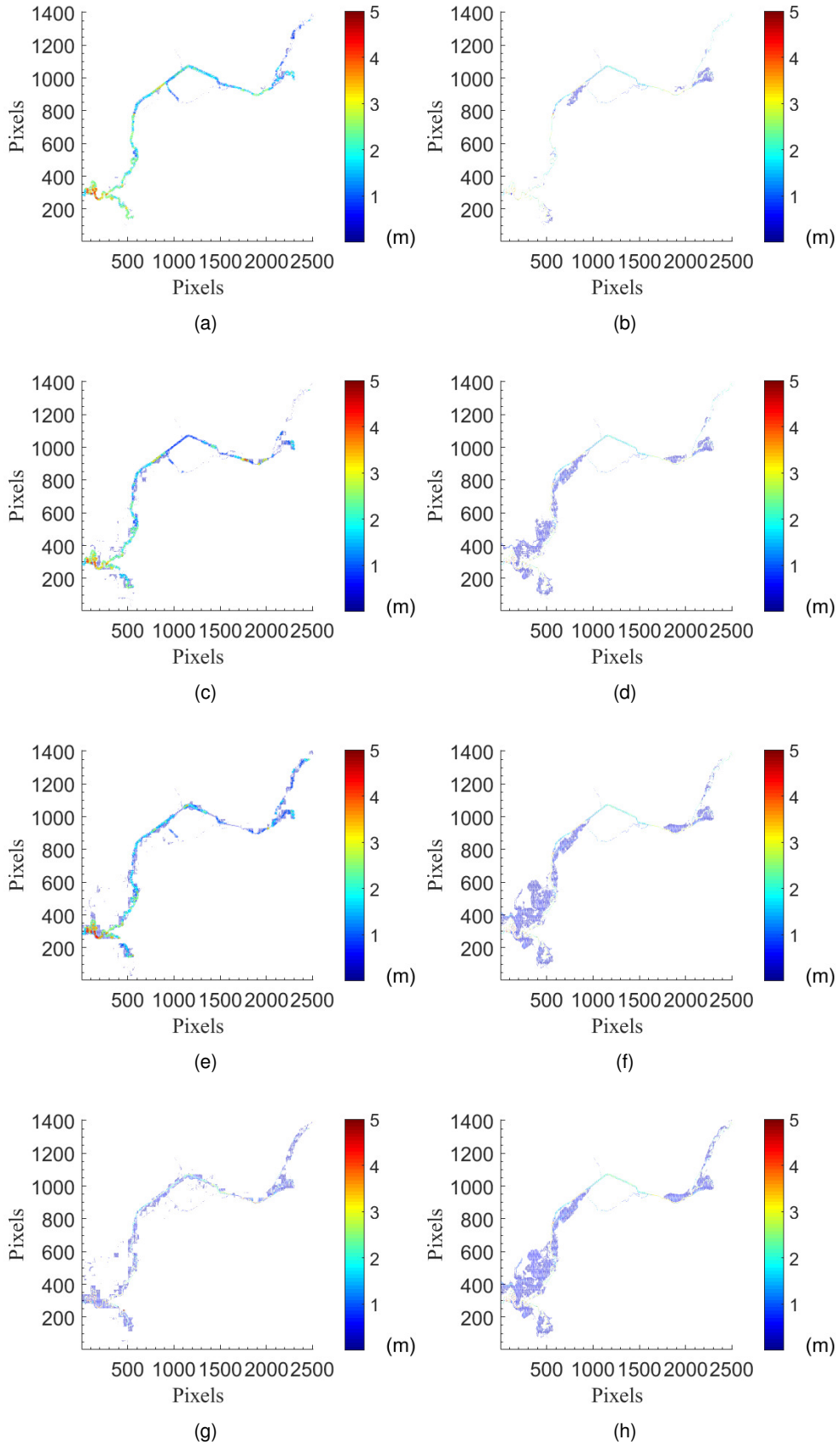

(a) ANN 3h (b) hydrodynamic 3h (c) ANN 6h (d) hydrodynamic 6h (e) ANN 9h (f) hydrodynamic 9h (g) ANN 12h (h) hydrodynamic 12h

**Figure 20** Predicted inundation maps of the first intervals of Flood Event 2013



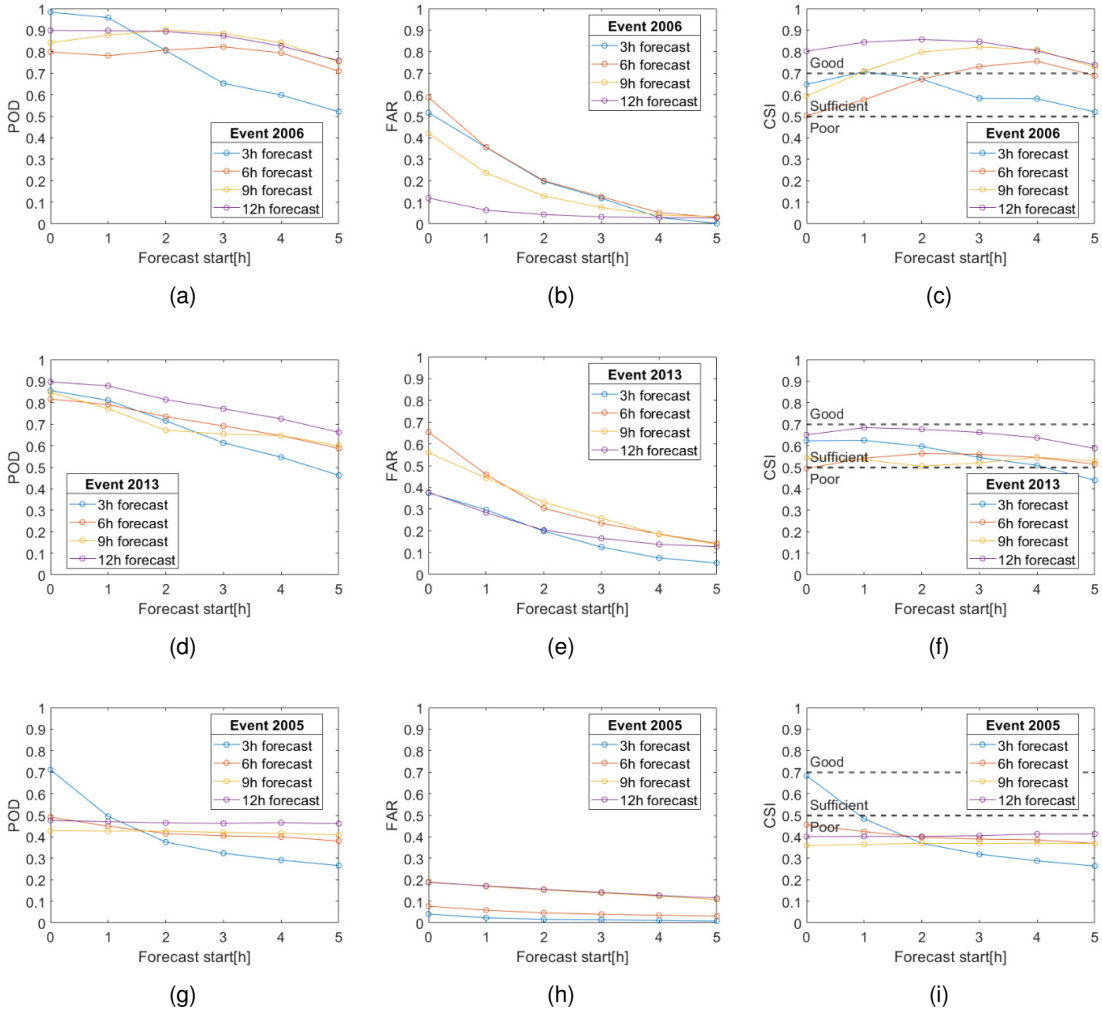

Seven discharge curves of three rivers and four streams are shown in different colors.
Time 0: the start of the prediction.
The dash lines: different discharge sections for prediction inputs.

**Figure 21** Hydrographs of Event 2005




(a) ANN 3 h (b) hydrodynamic 3 h (c) ANN 6 h (d) hydrodynamic 6 h (e) ANN 9 h (f) hydrodynamic 9 h (g) ANN 12 h (h) hydrodynamic 12 h

**Figure 22** Predicted inundation maps of the first intervals of Flood Event 2005




(a) POD in Event 2006 (b) FAR in Event 2006 (c) CSI in Event 2006 (d) POD in Event 2013 (e) FAR in Event 2013 (f) CSI in Event 2013 (g) POD in Event 2005 (h) FAR in Event 2005 (i) CSI in Event 2005

**Figure 23** Performance of the forecast of inundation extent growths by three indices

# 6. Model Structure Uncertainty

## 6.1. Methodology

This section introduces the mathematical principles of the model coupling of 1D SWMM and the 2D overland flow model P-DWave. Conceptually, SWMM generates the inflow hydrographs for the coupled model of SWMM and P-DWave of the urban area. First, the two models are coupled uni-directionally at the coupling boundary. Afterward, the evaluation criteria and the model structure uncertainty quantification are introduced.

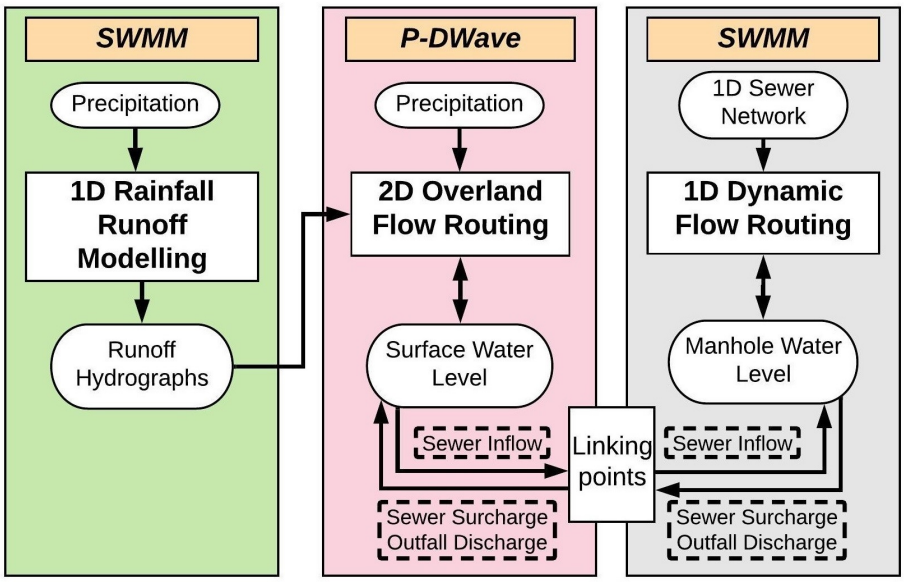
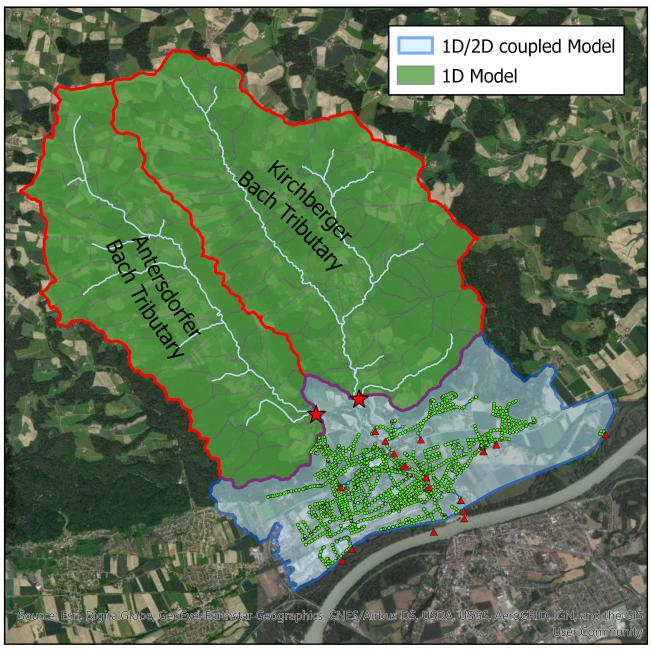



**Figure 24** Model coupling scheme

Figure 24 shows the primary workflow and the interactions between the model components. The green module represents the 1D SWMM model for upstream inflow calculation; the pink module represents the 2D P-DWave for 2D overland flow routing; the violet module represents the underground sewer network modeled by the 1D SWMM model. Uni-directional coupling is implemented at the inflow points, and bi-directional couplings occur at all the manholes.

Figure 25 shows the spatial localization of the different models in the Simbach catchment. The tributary Antersdorfer Bach and Kirchberger Bach (with a total area of 26.92 $km^2$. The urban area of 9.43 $km^2$ is modeled by coupled 1D/2D model using P-DWave for the city center. The two stars on the map represent the coupling junctions between the models. Implementing the 1D modeling for the upper stream area reduces the 2D modeled area; thus, the computational time is reduced dramatically for this case study.




Stars: inflow locations generated by 1D model

**Figure 25** Spatial distribution of the 1D and the coupled 1D/2D model

### 6.1.1. 1D Dynamic Rainfall-Runoff Model SWMM

The one-dimensional rainfall-runoff model applied in this study is the stormwater management model developed by the United States Environment Protection Agency (USEPA). The model is based on Saint-Venant Equation with routing in pipes, channels, pumps, and storage systems.

Hydrology

SWMM considers the rainfall-runoff process into the surface runoff by rainfall and snowmelt, with infiltration, evaporation, and initial abstraction. The mass conservation is written as:

$$\frac{\partial d}{\partial t} = i - e - f - q \tag{6.23}$$

where $d$ is the depth [m], $t$ the time [s], $i$ the rate of rainfall and snowmelt [m/s], $e$ the evaporation rate [m/s], $f$ the infiltration rate [m/s] and $q$ the runoff rate [m/s].

Several infiltrations are implemented in the SWMM model: Horton's method, a modified Horton method, the Green-Ampt method, and the Curve Number method (Rossmann and Huber, 2016). In this study, Horton's method is applied, written as:

$$f_p = f_\infty + (f_0 - f_\infty)e^{-k_d t} \tag{6.24}$$

$$F(t_p) = \int_0^\infty f_p dt = f_\infty t_p + \frac{f_0 - f_\infty}{k_d}(1 - e^{-k_d t_p}) \tag{6.25}$$

where $f_p$ is the infiltration capacity of the soil [mm/s], $f_\infty$ is the minimum infiltration capacity [mm/s], $f_0$ is the initial infiltration capacity [mm/s], $t$ is the time [s], $k_d$ is the decay coefficient [$s^{-1}$] and $F(t_p)$ is the cumulative infiltration capacity at time $t_p$ [s]. Due to numerical stabilities, the original form of Horton's method in Equation 6.24 is replaced by its integral form in Equation 6.25 in SWMM. Further information about the infiltration methods of SWMM are available in Rossmann et al. (2016).

Hydraulics

The routing process in SWMM model is calculated in pipes or channels. The model offers kinematic wave or dynamic wave for this calculation. In each step, the momentum is updated in all the channels by:

$$\frac{dQ}{dt} = 2V\frac{dA}{dt} - V^2\frac{dA}{dx} + gA\frac{dH}{dx} - gAS_f \tag{6.26}$$

where $x$ is the distance [m], $t$ is the time [s], $A$ is the flow cross-sectional area [m], $V$ is the velocity, $Q$ is the flow rate [$m^3/s$], $H$ is the hydaulic head of water in the conduit $(Z + Y)$ [m], $Y$ is the conduit water depth [m], $S_f$ is the friction slope and $g$ is the acceleration of the gravity [$m/s^2$]. In the finite difference form, the spatial step $dx$ takes the length of the link $L$. In the continuity equation:

$$\frac{dH}{dt} = \frac{\sum Q}{\sum A_S + A_{\text{Store}}} \tag{6.27}$$

where $H$ is the water head at the node [m], $A_S$ is the surface area at the nodes and the half-length of each link [$m^2$], $A_{\text{Store}}$ is the node storage [$m^2$] (Rossmann, 2017).

To solve the water head $H$ and the node discharge $Q$ in Equation 6.27 and Equation 6.26, SWMM uses an implicit backward Euler scheme. The time step should fulfill the Courant-Friedrichs-Lewy condition to ensure the numerical stability (Rossmann, 2017):

$$\Delta t \leq \frac{L}{|V|}\Big(\frac{\mathrm{Fr}}{1+\mathrm{Fr}}\Big)\mathrm{Cr} \tag{6.28}$$

where $\mathrm{Fr}$ is the Froud number and $\mathrm{Cr}$ is the Courant number.

When a node is surcharged, the Equation 6.26 has no value to update. Therefore, Equation 6.27 is used for the water head update:

$$\Delta H = \frac{-\sum Q}{\sum \frac{dQ}{dH}} \tag{6.29}$$

If the surcharging is over the node's maximum depth $H_{\max}$, the node becomes flooded with the overflow rate $Q_{\mathrm{ovfl}}$:

$$Q_{\mathrm{ovfl}} = 0.5\Big(\sum Q^{t} + \sum Q^{t+\Delta t}\Big) \tag{6.30}$$

Further details of the SWMM models are available in Rossmann et al. (2017).

### 6.1.2. 2D Dynamic Rainfall-Runoff Model PD-Wave

P-DWave is a two-dimensional diffusive model for rainfall-runoff simulations. The detailed information about the model is published in Leandro et al. (2014).

The model is based on the two-dimensional shallow water equation. The continuity equation and the momentum equation are written as:

$$\frac{dh}{dt} = \nabla(\boldsymbol{u}h) = R \tag{6.31}$$

$$\frac{d\boldsymbol{u}}{dt} = (\boldsymbol{u}\nabla)\boldsymbol{u} + \frac{\nu_t}{h}(h\nabla\boldsymbol{u}) + g\nabla(h+z) = g\boldsymbol{S}_f \tag{6.32}$$

where $\boldsymbol{u} = [u_x \quad u_y]^T$ is the depth-averaged flow velocity, $h$ is the water depth, $g$ is the gravitational acceleration, $z$ is the height of the river bed, $\nu_t$ is the eddy viscosity, $R$ is the source/sink term and $\boldsymbol{S}_f = [S_{fx} \quad S_{fy}]^T$ is the bed friction vector.

The P-DWave model neglects the forces with only the gravity and the friction left. Thus, the momentum equation 6.32 is modified to:

$$g\nabla(h+z) = g\boldsymbol{S}_f \tag{6.33}$$

The bed friction is calculated with Manning's formula:

$$\boldsymbol{S}_f = \begin{bmatrix} \frac{n^2|\boldsymbol{u}|u_x}{h^{4/3}} \\ \\ \frac{n^2|\boldsymbol{u}|u_y}{h^{4/3}} \end{bmatrix} \tag{6.34}$$

where the $\nabla(h+z) = [S_{wx} \quad S_{wy}]^T$ is the water-level surface gradient vector, with $\boldsymbol{S}_w = [d(h+z)/dx \quad d(h+z)/dy]^T$. The modulus of the averaged flow velocity over depth is written as:

$$|\boldsymbol{u}| = \frac{h^{2/3}\sqrt{I_m}}{n} \tag{6.35}$$

$$I_m^2 = S_{wx}^2 + S_{wy}^2 \tag{6.36}$$

The continuity equation 6.31 is discretized with an explicit first-order finite volume scheme on a regular grid. Finally, the cell-centered controlled volume is applied for the spatial domain.

$$\frac{h_i^{t+1} - h_i^t}{\Delta t} + \frac{1}{A_i}\sum\nolimits_{j=1}^{4} h_{ij}u_{ij}L_{ij} = R \tag{6.37}$$

$A_i$ is the cell area, $L_{ij}$ is the contact area between the cells, $u_{ij}$ and $h_{ij}$ are the flow velocity and water depth which is calculated by the following equations:

$$h_{ij} = \frac{h_i + h_j}{2} \tag{6.38}$$

$$u_{ij} = \frac{h_{ij}^{4/3}}{n_{ij}^2|\boldsymbol{u}_{ij}|}I_{n,ij} \tag{6.39}$$

$u_{ij}$ is the normal flow velocity to the cell contact surfaces. $I_{n,ij} = (S_{wx}\tilde{n_x} + S_{wy}\tilde{n_y})$ is the vector of the surface gradient multiplied by the normal vector at the contact surface $\tilde{\boldsymbol{n}} = [\tilde{n_x} \quad \tilde{n_y}]^T$. For the numerical stability of the explicit discretization scheme, the condition for

the time step must be fulfilled. For this purpose, the equation is reformulated by replacing $a_{ij} = u_{ij}L_{ij}$. Thus, Equation 6.37 is rewritten as:

$$h_i^{t+1} = h_i^t\left(1 - \frac{\Delta t}{2A_i}\sum_{j=1}^{4} a_{ij}\right) + \frac{\Delta t}{2A_i}\sum_{j=1}^{4} a_{ij}h_j^t \tag{6.40}$$

$$a_{ij} = \frac{h_{ij}^{2/3}}{n_{ij}}\frac{I_{n,ij}}{\sqrt{I_{m,ij}}}L_{ij} \tag{6.41}$$

For a regular mesh with $A_i = \Delta x^2$, the limitation for $\Delta t$ is written as:

$$\Delta t \leq \max\left\{\min_{i,j}\left\{2\Delta x^2 n_{ij}\frac{\sqrt{S_{wx,ij}}}{h_{ij}^{5/3}}\right\}, \min_{i,j}\left\{2\Delta y^2 n_{ij}\frac{\sqrt{S_{wy,ij}}}{h_{ij}^{5/3}}\right\}, \Delta t_{\min}\right\} \tag{6.42}$$

During the flood event, the cell states are changing between the wet state and the dry state, causing a movement of the boundary conditions. For the sake of it, the parameter $\phi$ is introduced to the continuity equation:

$$\frac{h_i^{t+1} - h_i^t}{\Delta t} + \frac{1}{A_i}\sum_{j=1}^{4}\phi_j h_{ij}u_{ij}L_{ij} = R \tag{6.43}$$

$\phi$ takes the value 1 if the water depth is greater than 0 in the next time step. $\phi$ takes a value between 0 and 1 to avoid the water depth from getting negative values.

$$\phi = \phi_{1\ldots4} = \begin{cases} 1 & , h_i^{t+1} \geq 0 \\ \dfrac{\Delta x^2}{\Delta t}\dfrac{h_i^t + R\Delta t}{\sum_{j=1}^{4} h_{ij}u_{ij}L_{ij}} & , h_i^{t+1} < 0 \end{cases} \tag{6.44}$$

For more details of mathematical descriptions of the PD-Wave model, please refer to the article Leandro et al. (2014).

### 6.1.3. 1D-2D Model Coupling of SWMM and P-DWave

SWMM 5.1 integrated the Dynamic Link Libraries (DLL), which enables the linage to other models (Leandro et al., 2016). This implementation can attach the 2D overland model with the standard SWMM model for bi-directional communication. In this way, the coupled 1D/2D model can exchange the water depth between the P-DWave model and the SWMM model at all the linked nodes at each time step.

A detailed description of the SWMM-P-DWave can be found in the publication (Chen et al., 2016). The first routine links 1D-SWMM and 2D-P-DWave models bi-directionally. Depending on the hydraulic heads, the discharge is determined by either a weir equation or an orifice equation.

$$Q = \begin{cases} c_w w h_{2D} \sqrt{2 g h_{2D}} & \text{, if} \quad h_{1D} \leq Z_{2D} \\ c_0 A_{mh} \sqrt{2 g (h_{2D} + Z_{2D} - h_{1D})} & \text{, if} \quad h_{1D} > Z_{2D} \end{cases} \tag{6.45}$$

where $Q$ is the discharge between the 1D and 2D models, $c_w$ is the weir discharge coefficient, $w$ is the weir crest width [m], $A_{mh}$ is the manhole area [m$^2$]. Surcharging discharge is calculated by the following equation:

$$Q = -c_0 A_{mh} \sqrt{2 g (h_{1D} - Z_{1D} - h_{2D})} \tag{6.46}$$

The second routine synchronizes the two models. From the consideration on efficiency, the time step in 1D model $\Delta t_{1D}$ is larger than the time step in the 2D model $\Delta t_{2D}$. To synchronize the models, the synchronize time $T_{sync}$ is set to the 1D time step $\Delta t_{1D}$ and the 2D time step $\Delta t_{2D}$ is adjusted at every synchronization.

$$\Delta t_{2D} = \min \left\{ \left( T_{sync} \Delta t_{1D} \sum \Delta t_{2D} \right), \Delta t_{2D} \right\} \tag{6.47}$$

### 6.1.4. Rainfall Data

The rainfall data is obtained from the YW-RADOLAN data from the German Weather Service(DWD)(Winterrath et al., 2018)organized per gauge with a spatial resolution of about 1km and a temporal resolution of 5 minutes.

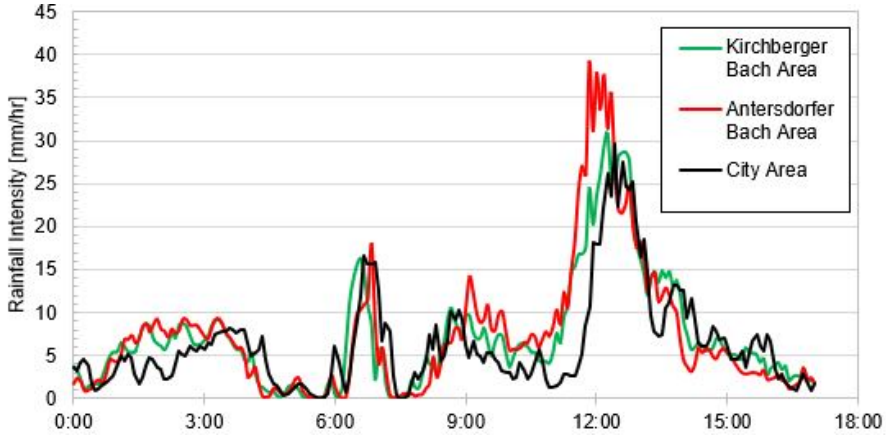



**Figure 26** Rainfall intensities on the 1st of June 2016

Figure 26 shows the rainfall intensity on 1st, Jul. 2016 of the event. The peak intensity of the whole catchment, Simbach city, the Antersdorfer Bach, and the Kirchberger Bach are 35.04 mm/hr, 29.58 mm/hr, 39.11 mm/hr and 30.77 mm/hr, between 11:50 and 12.25 (CET).

### 6.1.5. 1D Catchment Inflow Modelling

The 1D SWMM model models the discharge inflows entering the urban area. Due to the data scarcity on event 2016, the 1D catchment modeling is calibrated using an extreme rainfall event in 2015. After that, the model is further tested for 2016 before reconstructing the historical event.

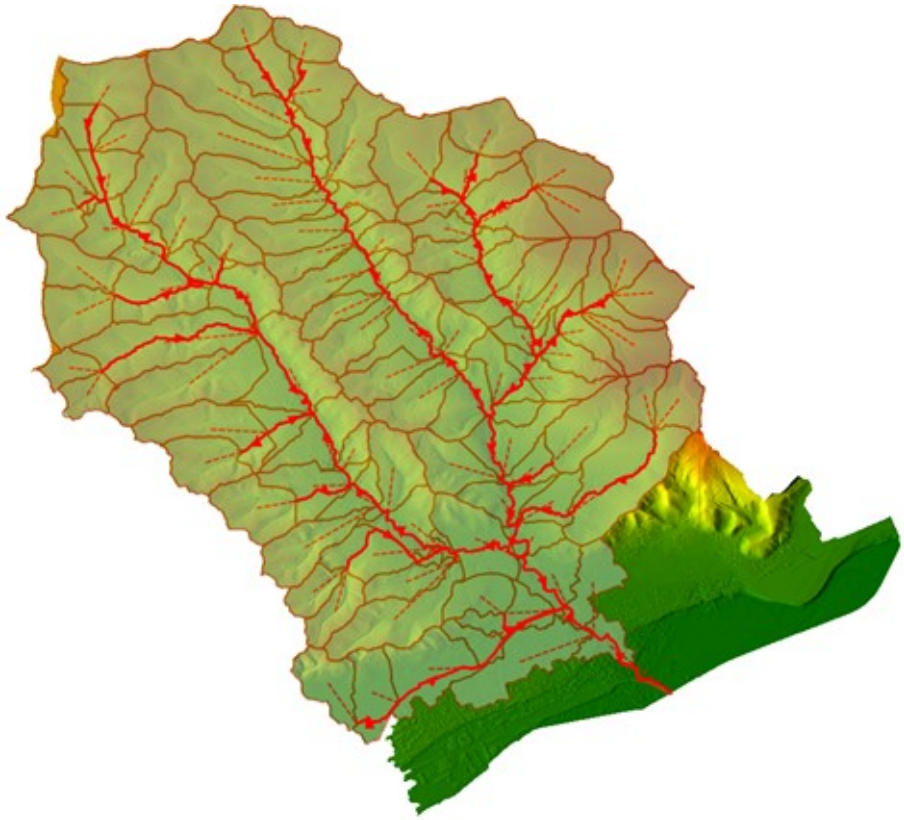


**Figure 27** Sub-catchments and open channels for the rainfall-runoff simulation

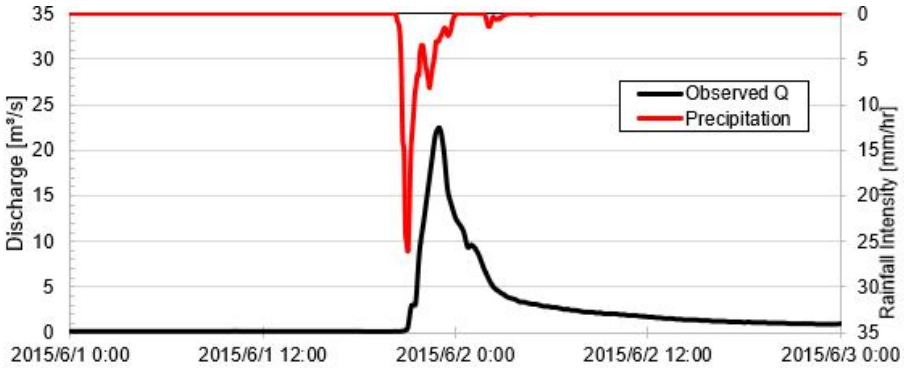



**Figure 28** Observed discharge and precipitation between 01.06.2015 and 03.06.2015

Figure 27 shows the splitting of the catchment Simbach am Inn into 103 smaller sub-catchments for flow path routing. The flows are modeled as open channels with 15m and a depth of 10

m to ensure the water stays within the channel profile even in extreme events (Leandro et al., 2009).

Figure 28 presents the documented precipitation and discharge on a gauge station from Jun 1st, 2015, for the model calibration. The discharge is accessible on the LfU website (LfU, 2020).

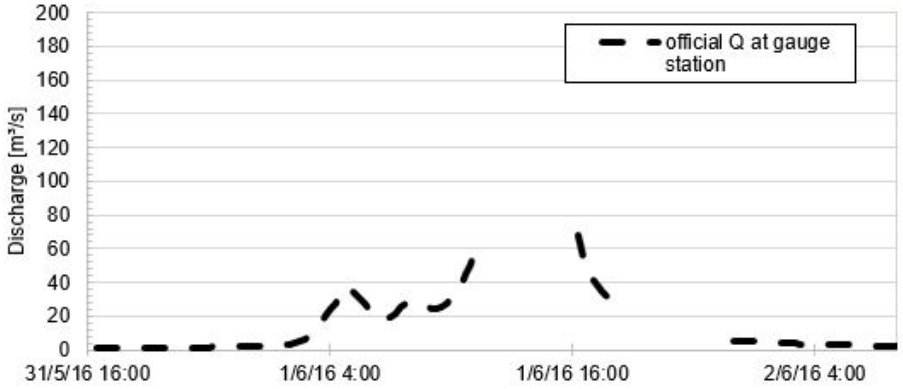



**Figure 29** Observed discharge at the gauge station of Simbach am Inn between 31.5.2016 and 02.06.2016 (LfU 2020)

**Table 12** Rainfall-runoff parameter calibration

| Parameter | Default | Event 2015 | Event 2016 |
|---|---|---|---|
| Min. Infiltration Rate | 0.5 | 5 | 0.462 |
| Max. Infiltration Rate | 3 | 11 | 16 |
| Decay constant $k_D$ | 4 | 2 | 0.416 |
| Pervious Area Roughness | 0.1 | 0.075 | 0.065 |
| Conduit Roughness | 0.01 | 0.068 | 0.105 |
| Impervious Area % | 25 | 4 | 4 |

Figure 29 shows the missing documented gauging station hydrograph of the 2016 event. The parameters are calibrated for the 2015 rainfall event and adjusted for 2016. Using distributed precipitation data, an ensemble of 150 samples runs with different parameter sets for the realistic inflows. The parameter set best fitting the discharge hydrograph at the gauging station is used for the following studies, see Table 12.

### 6.1.6. Model Structures of 1D/2D Modelling

In the urban area modeling, seven different model structures are set up to assess the model structure uncertainty. The model structures include sewer systems, dam erosion, rain pumps, pump failure, open wastewater treatment plant (WTTP), and street erosions.

Table 13 lists the seven model structures for Simbach floods. MS1 only considers the inflows for 2D surface runoff; MS2 adds the sewer system to the first structure; MS3a adds the rainwater pumps in Folder Erlach and Simbach to MS2. MS3b applies the pump failure in

**Table 13** Summary of the conducted simulation scenarios

| | Inflow | Sewer | Rainwater Pumps | Pump failure | WTTP Open | Street Erosion |
|---|---|---|---|---|---|---|
| MS1 | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ |
| MS2 | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ |
| MS3a | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ |
| MS3b | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| MS3c | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ |
| MS4a | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ |
| MS4b | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

Simbach at 13:00. Scenario 3c includes the modeling of the wastewater treatment plant (WTTP) in Polder Erlach with 50 $m^2$. Finally, MS4a models the erosion of the street curb at Schulstrasse based on M1. MS4b models all the mentioned components, including the sewer system, pumps, and WTTP. All the runs start from 00:00 and end at 20:00.

### 6.1.7. 1D Sewer Network Modelling for Simbach Case

Figure 30 shows the SWMM sewer network for 1D/2D model coupling in Simbach city. According to the information from the local operator, the rainwater pumps Simbach I, Simbach II, and Erlach I are modeled with pump rates of 600 l/s, 1250 l/s, and 1500 l/s. The entire sewer network consists of 1525 manholes and 51.8 km of conduits. In addition, 36 outfalls in the urban area let the water surface.

### 6.1.8. 2D Surface Flow Modelling for Simbach Case

As a Balance between the model efficiency and model accuracy, a cell resolution of 5 m × 5 m is applied for the study (Mark et al., 2004) (Gallegos et al., 2009). The underpasses, the B12, the bridge, and the railway underpass are modeled as open channels. The heights of all the streets in the DEM are reduced by 0.1 m as the street canalization. The buildings are raised for 5 m in the DEM to create the blockage for the flows (Mark et al., 2004) (Schubert and Sanders, 2012). For the DEM erosion modeling, 14 DEMs are generated in series according to the documented data (Hübl et al., 2017). As the dam break process remains undocumented, the height reduction is assumed to change linearly during the process.

Boundary Conditions The study area includes the whole catchment Simbach (see Figure 25). The river Inn bounds the southeast boundary of the catchment Simbach as a free outflow. The rest areas end by the catchment boundary; therefore, a close boundary is assigned.

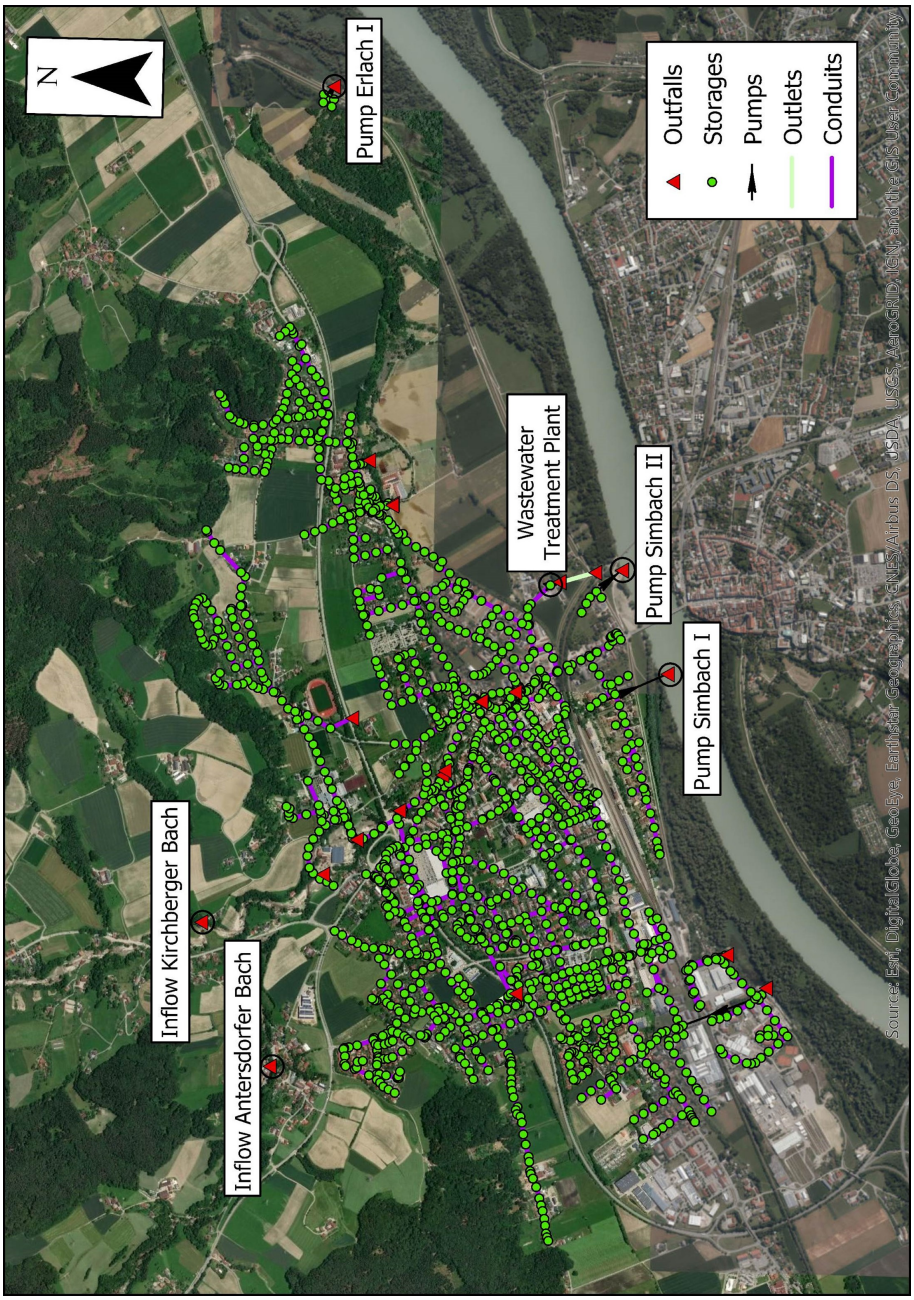



**Figure 30** Sewer network with important features in Simbach

### 6.1.9. Uncertainty Assessment in Model Structures

For quantifying model structure uncertainties, the maximum water depth, the flood propagation speed, the flooding extent, and the discharge at the gauge stations/sewers are compared to show the model differences.

The simulation values are compared with the documented maximum water depths from the historic event to show the model's goodness of fit at the gauge stations. Three indices are used to quantify the model fits: the Nash-Sutcliffe-Efficiency (NSE) (Nash and Sutcliffe, 1970), the Root-Mean-Square-Error (RMSE), and the Mean-Average-Error (MAE).

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n}(y_i' - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{6.48}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i' - y_i)^2}{n}} \tag{6.49}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n}|y_i' - y_i|}{n} \tag{6.50}$$

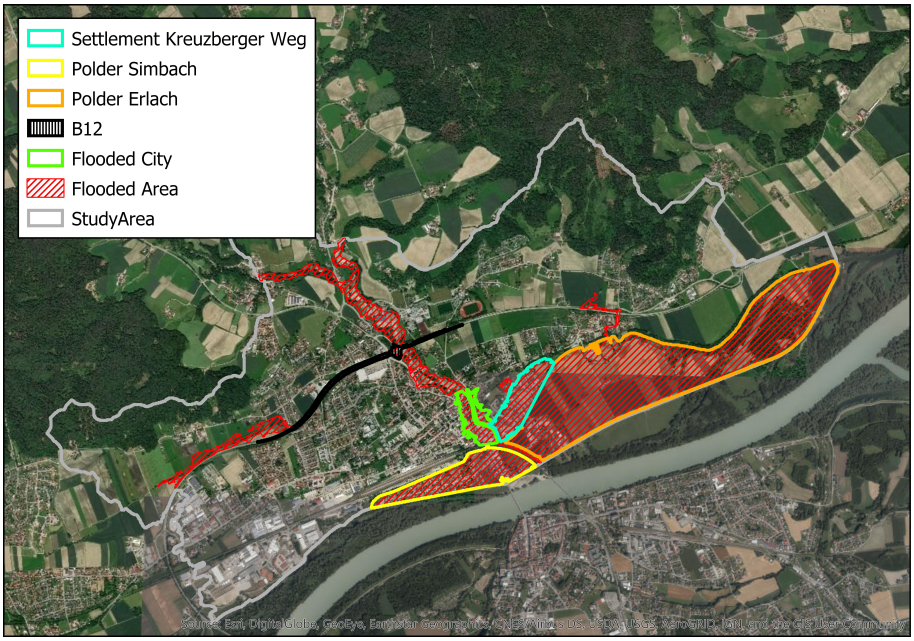



**Figure 31** Specific locations in Simbach

The average water depth within a particular area is compared. This part analyzes the following areas: the whole area, the total area flooded by the flash flood, the city center orographic left (Kreuzberger Weg), and right of the Simbach, Polder Simbach, Polder Erlach, and the B12.

Besides, the discharge volumes at the gauge stations are also checked to quantify the model uncertainty. Since the P-DWave model uses the regular square mesh for discretization, the discharges at the gauge station are measured either vertically or horizontally across the cell. This assumption would likely underestimate the discharge volume, but it is still enough to compare different model structures.

## 6.2. Data for Model Performance Evaluation

The documented water levels at 14 points in Simbach city are used for the model evaluation.

Figure 2 shows the location of the 14 checkpoints. Checkpoint 1 lies upstream of the street dam at Schulstrasse. Checkpoint 2 is on the top of Highway B12. Checkpoint 3 is in the river channel before the artificial channel. After the artificial channel, checkpoints 4 and 5 are by the bridge at Passauerstrasse. Checkpoints 6, 7, and 8 lie north of the railway inside the city. Checkpoints 9 and 10 are in the Polder Simbach. Checkpoints 11 to 14 are distributed between Kreuzberger Weg and Wilhelm-Diess-Strasse in the north of the Polder Erlach.

## 6.3. Results and Discussion

### 6.3.1. Rainfall-Runoff Model Calibration

The model is calibrated according to the documented discharge at the gauge station with 100 sample runs. In addition, due to the missing extreme discharge peak value, the model is also calibrated according to the event 2015 with 15 sample simulation runs.

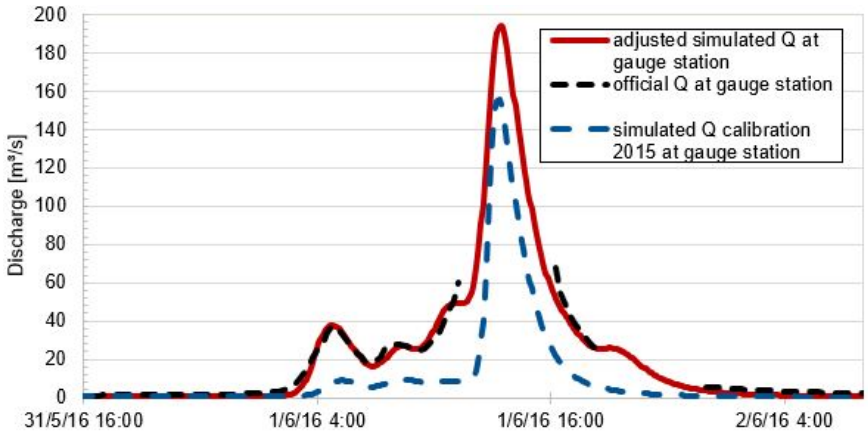



**Figure 32** Calibrated discharge input curve of Simbach 2016 and comparison to observation

Figure 32 shows the discharge curve after the calibration on SWMM parameter. Figure 32 demonstrates the newly chosen hydrograph at the gauge station as compared to the hydrograph from the 2015 calibration. The peak discharge of 194 $m^3/s$ at the gauge station occurs at 13:25. The hydrographs are used as the input inflows for the sewer-surface coupled model.

After the rainfall model calibration by the rainfall event 2015, the model validation with the extreme event in 2016 reaches an NSE over 0.7. The high NSE is because a long-term calibration is omitted due to the extreme computation run time. Therefore, the incomplete dataset and the underestimated peak volume of 2016 limits the goodness of fit. Besides, the rainfall duration and the rainfall intensity have exceeded the 2015 event by volume. Furthermore, the rainfall-runoff model does not consider the dam erosion at 12:40 into consideration, limiting the simulation results of the total discharge volume.

### 6.3.2. Maximum Water Depth in Coupled Model Structures and Structural Uncertainty

This section compares seven 1D/2D coupled model structures for Event 2016 in Simbach with dam erosion and sewer system couplings. The model structures are evaluated by comparing them to the measured parameters from the event.

Model Structure 1

The first model structure includes neither dam erosion nor sewer networks. To omit the random oscillation and for better visualization, areas with water depth higher than 0.1 m are displayed as the inundation area. The maximum depth at Schulstrasse (CK 1) is 0.2 m higher than the observed water depth. The water depth on the B12 (CK 2) is underestimated by 0.4 m. The maximum water depths at the Passauer Bridge (CK 4 and 5) are overestimated by 0.7 m and 0.5 m. On the north side of the railway, CK 6, CK 7, and CK 8 are underestimated by 0.4 m to 0.8 m. The water levels in Polder Simbach (CK 9 and CK 10) are underestimated by 0.9 to 1.1 m. CK 11, 12, 13 are underestimated by 0.2 to 0.3 m. CK 14 is slightly overestimated by 5 cm. The RMSE, NSE, and MAE are 0.559, 0.908, and 0.464, respectively. Figure 33 shows the maximum depths of Model Structure 1.

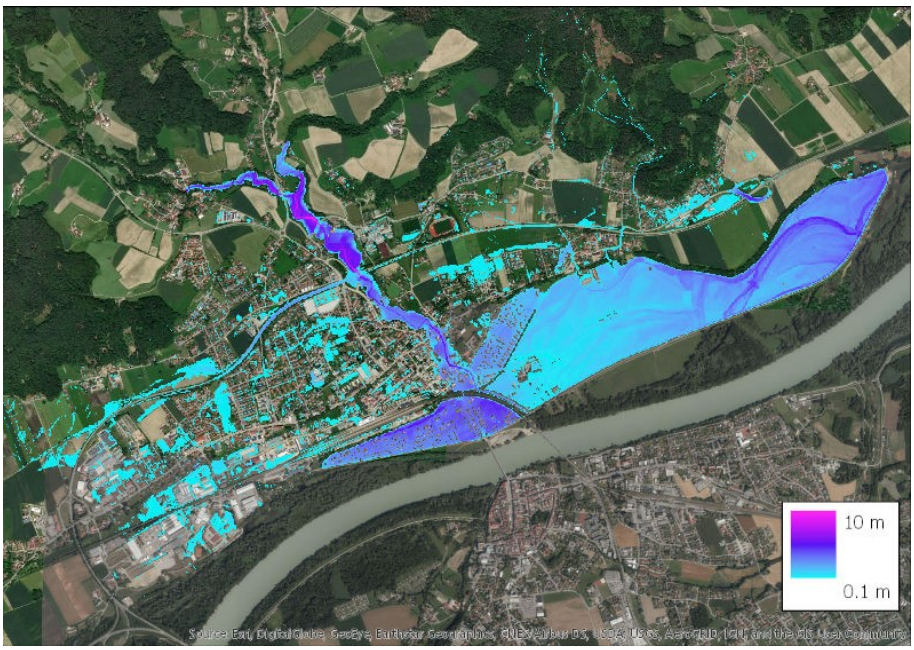



**Figure 33** Maximum water depths MS 1

Model Structure 2

Model Structure 2 added the sewer system based on Model Structure 1. Therefore the system applied the dual drainage concept, enabling the water exchange between the surface and underground sewer networks.

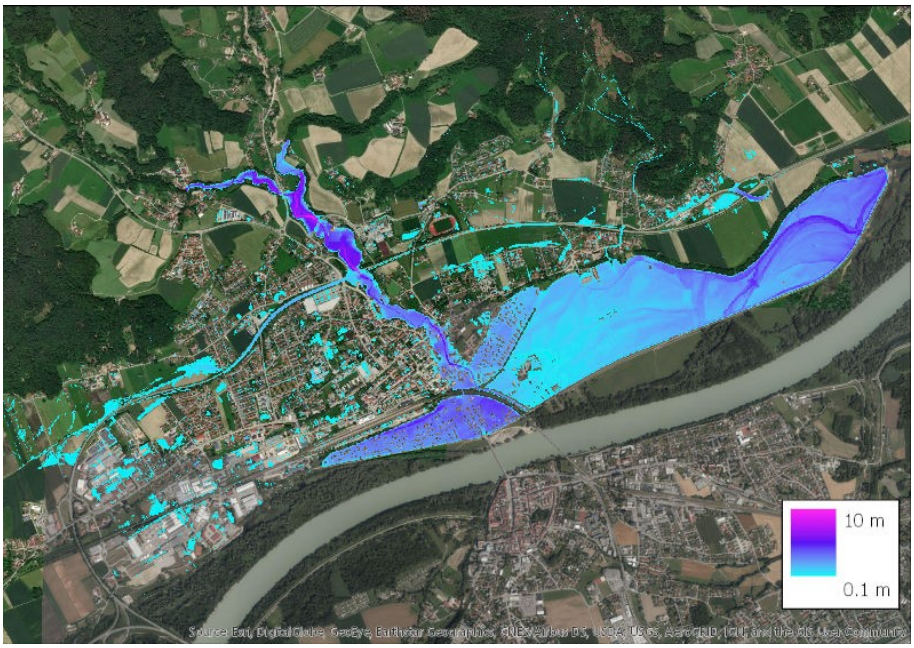



**Figure 34** Maximum water depths MS 2

Figure 34 shows the maximum water depth in Model Structure 2. In the urban area, the inundation areas are significantly smaller than Model Structure 1, showing that the sewer network drains surface water underground. Schulstrasse (CK 1) is overestimated by 0.2 m. B12 (CK 2) is underestimated by 0.4 m. The drainage channel is overestimated by 0.1 m. The maximum water depths at the Passauer Bridge are overestimated by 0.5 to 0.7 m. Similar to Model Structure 1, CK 6, CK 7, and CK 8 are underestimated by 0.4 m to 0.8 m. The water depth in Polder Simbach (CK 9 and CK 10) is overestimated with a reduced error from 0.7 to 0.9 m. The gauge stations by the river (CK 11, 12, 13) are underestimated by 0.3 m to 0.4 m. CK 14 accurately matches the observed value with a difference of 4 cm.

### Model Structure 3a

Model Structure 3a combines the rainwater pumps in Polder Simbach and Polder Erlach with the sewer system.

Figure 35 shows the maximum depths of the model structure. Compared to Model Structure 1 and 2, the overestimated water depth in Polder Simbach is reduced (Model Structure 1: 0.9 m to 1.1 m; Model Structure 2: 0.7 m to 0.9 m; MS 3: 0.6 m to 0.7 m). As a result, the errors of the water level in the east are around 10 cm less.

### Model Structure 3b

Model Structure 3b models the pump failure at 13:00. The checkpoints in the polder show a water depth increase of 2 cm and 3 cm.

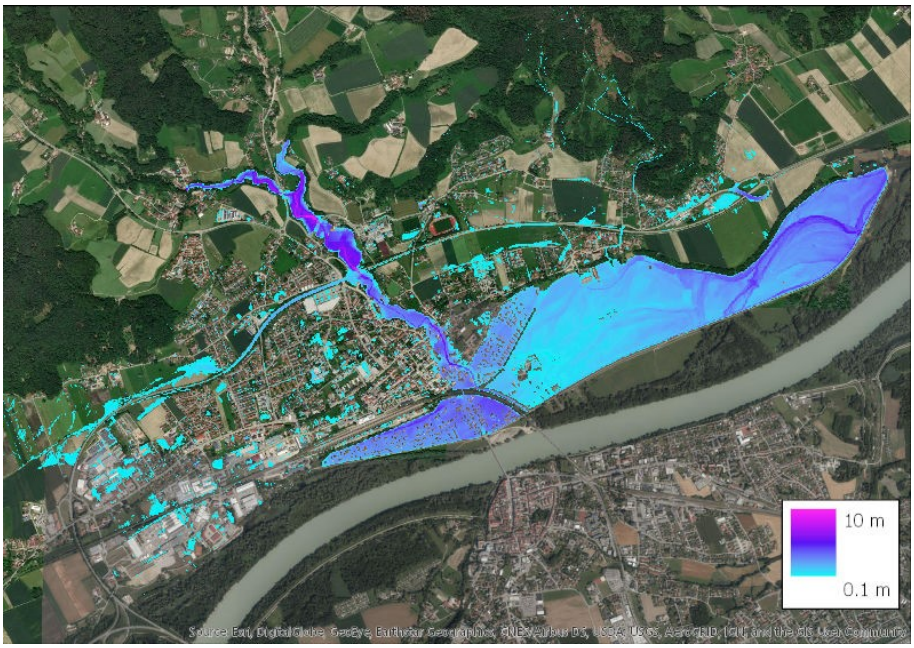



**Figure 35** Maximum water depths MS 3a

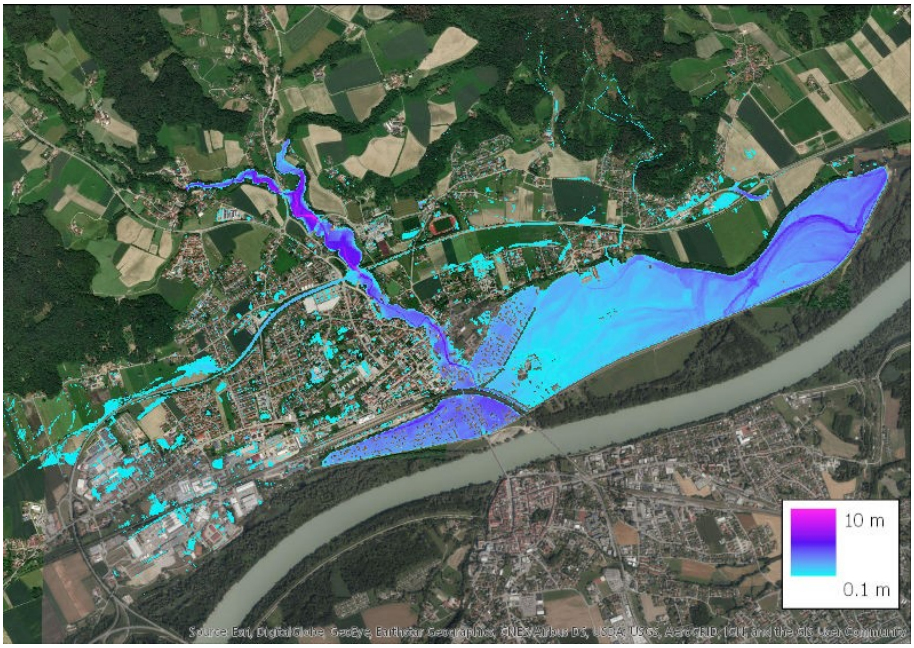



**Figure 36** Maximum water depths MS 3b

Figure 36 shows the maximum depths of the Model Structure 3b.

Model Structure 3c

Model Structure 3c implements the pump failure at 13:00 and a water treatment plant.

Figure 37 shows the maximum depths of the Model Structure 3c. Compared to Model Struc-

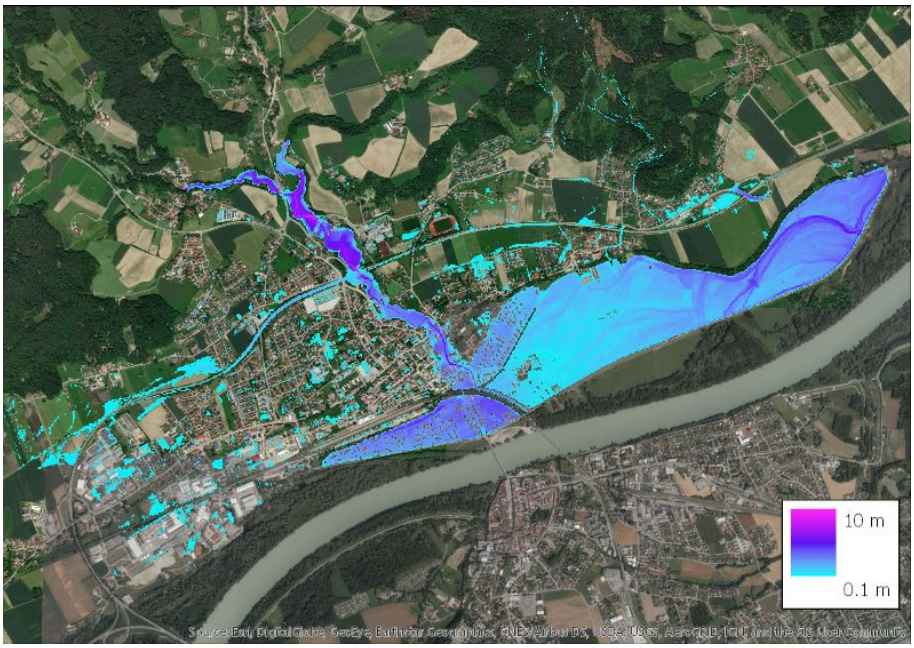



**Figure 37** Maximum water depths MS 3c

ture 3b, the errors in maximum depths are reduced by 0.7 m and 0.5 m at the Passauerstrasse (CK 4 and 5). Model Structure 3c performs similarly to Model Structure 3b near the railway (CK 6, 7, and 8). The water depths in the eastern of the city are lower in 2 cm to 10 cm compared to 3b. The water treatment plant drains this portion of the water to the sewer drainage system.

### Model Structure 4a

Model Structure 4a implements the dam erosion at Schulstrasse without a sewer system or pumps.

Figure 38 shows the maximum water depth of Model Structure 4a. The maximum water depth at the dam erosion (Schulstrasse) is underestimated by 0.22 m and 0.5 m at B12. The gauges at the open channel upstreams (CK 3), at the Passauerstrasse bridge (CK 4), and nearby (CK 5) are overestimated by 0.5 m, 0.9 m, and 0.7 m. The inundation depths in the west side of the river (CK6, CK7, and CK8) are underestimated by 0.39 m to 0.75 m. As the sewer and pump systems are omitted in the model structure, the water depths in Polder Simbach (CK9, CK10) are 0.9 to 1.1 m higher than the measured values. CK11, CK12, CK13, and CK14 match the measured values with an error smaller than 0.17 m.

### Model Structure 4b

This model structure combines the sewer network, the pump system, the wastewater treatment plant, and the dam erosion.

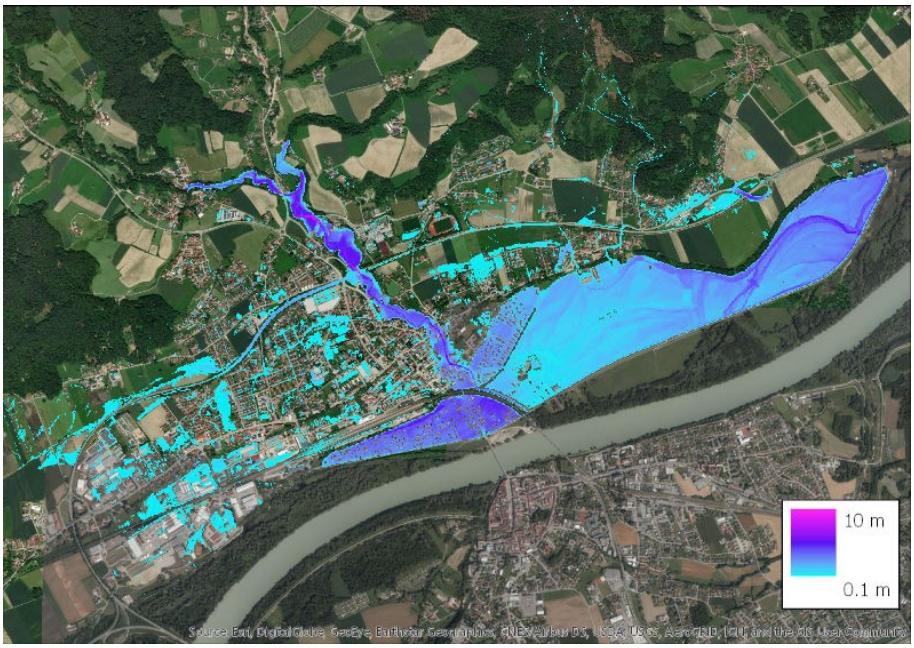



**Figure 38** Maximum water depths MS 4a

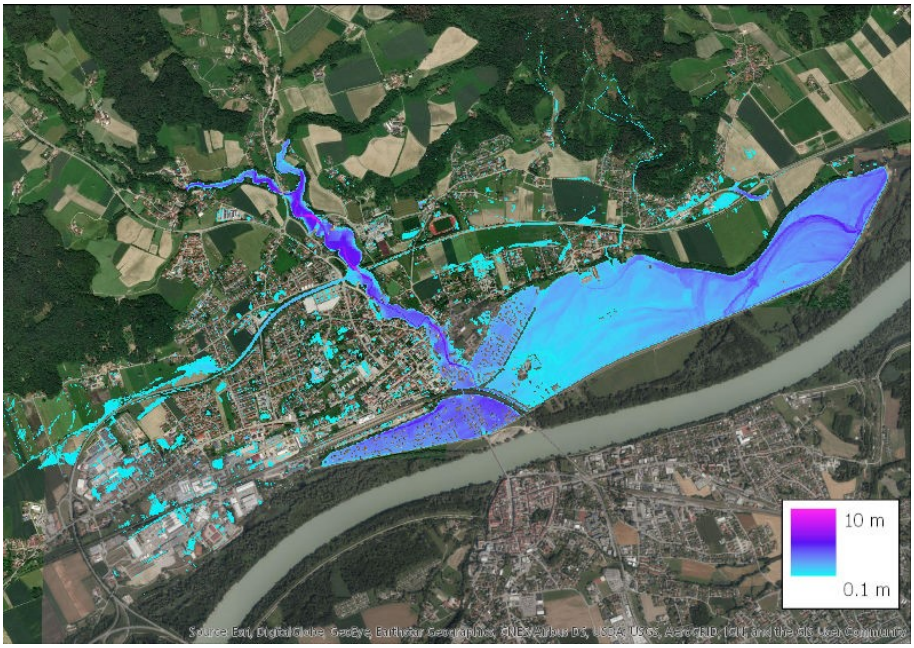



**Figure 39** Maximum water depths MS 4b

Figure 39 shows the maximum water depths. The water depth on the B12 is underestimated by 0.36 m compared to the measured data. CK 3, 4, and 5 overestimated the results by 47 cm, 97 cm, and 0.69 cm. The checkpoints in the city center (CK 6, CK 7, CK 8) are underestimated by 0.4 m to 0.8 m. Compared to MS 4a, CK 9 and 10 are 25 cm and 28 cm less estimated. The modeled sewer network reduces the water depths in the eastern of the city; thus, the maximum depths there (CK 11, 12, 13, and 14) are underestimated by 14 cm to 33 cm.

Summary of Model Structures

Over the seven structures reach the largest average RMSE and MAE of 0.594 m and 0.496 m and the smallest NSE of 0.896. The best-performed model structure has an RMSE, MAE, and NSE of 0.493 m, 0.429 m, and 0.928 from the MS 3c, with the complete sewer system coupled, but no dam erosion included.

**Table 14** Goodness of fit of seven model structures Simbach

| | **RMSE** | **NSE** | **MAE** |
|---|---|---|---|
| MS1 | 0.559 | 0.908 | 0.464 |
| MS2 | 0.535 | 0.916 | 0.463 |
| MS3a | 0.496 | 0.928 | 0.429 |
| MS3b | 0.511 | 0.923 | 0.438 |
| MS3c | 0.493 | 0.928 | 0.429 |
| MS4a | 0.594 | 0.896 | 0.489 |
| MS4b | 0.568 | 0.905 | 0.496 |

The maximum depth in the retention area upstream is underestimated and delayed in the simulation of the model structure with dam erosions. The reason could be the model of the street embankment and the breach processes. The model structure models the streets as open channels with curbs on both sides, allowing the water to flow out. Thus the water would less cumulate before the dam erosion. As a simplification, the dam erosion is assumed to follow a linear interpolation over time, inducing an error in the model structure.

Generally, the water depth in the city area in the north of the railways is underestimated, while the Polder Simbach is overestimated. The water level in the eastern part of Simbach is slightly overestimated due to the orthogonal grids, which include only the perpendicular flow in the discharge volume. Nevertheless, the most complex Model Structure 4b achieves the best performance overall. The possible reason could be that the drainage system is not accurate enough to represent the reality, or other factors, such as the mud blockage.

### 6.3.3. Inundation extent

The area with a maximum water depth of over 0.1 m is defined as the inundation area. The outlines of the inundation area are compared to show the model structure differences.

Figure 40 shows the extent difference of MS 3c (with dam erosion) and MS 4b (without dam erosion).

Figure 41 and Figure 42 show the detailed comparison in the upper- and downstream of River Simbach in the city. In Figure 41, the differences in the flood extent significantly occur in the underpass of the PAN4, the street curbs along Schulstrasse, and the B12 underpass.

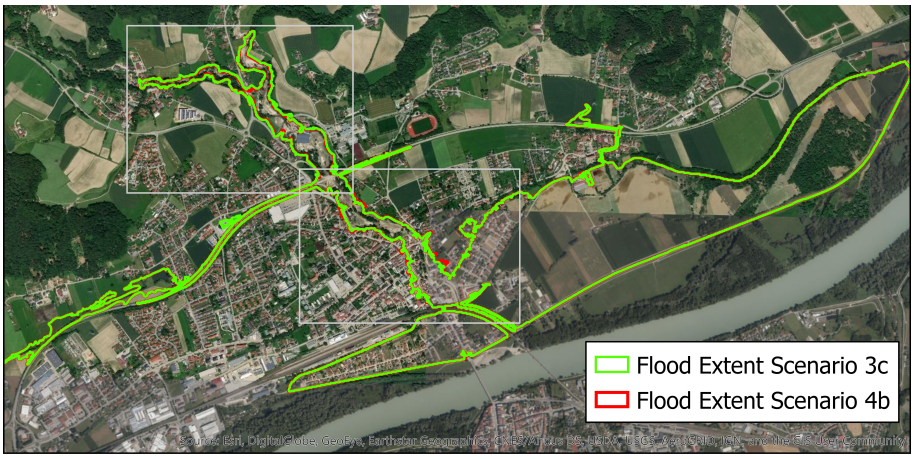



**Figure 40** Flood extents of MS 3c and MS 4b

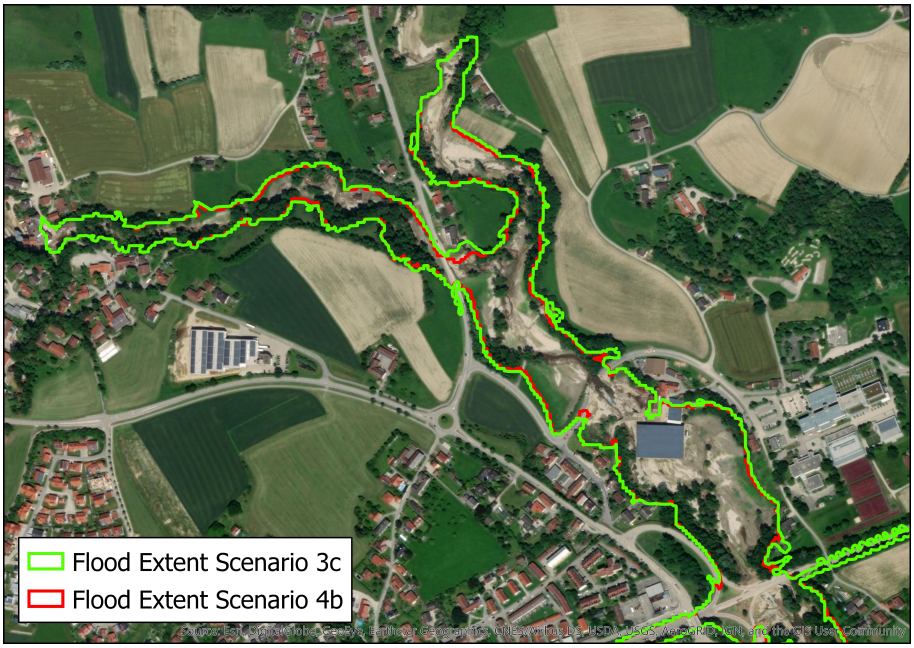



**Figure 41** Flood extents of MS 3c and MS 4b, upstream

Overall, Model Structure 4b has a slightly wider extent than Model Structure 3c. However, in most areas, two model structures produce similar outlines.

### 6.3.4. Impact of Sewer System and Dam Erosion on Maximum Water Depth

The maximum depth differences of Simbach city are presented here.

Figure 43 presents the difference in the maximum water depths between MS 3c and MS 1. Figure 44 shows the difference in the maximum water depth between MS 4b and MS 4a. The two couplings with the sewer network and the pump system increase their drainage capacities with a decrease of 0.38 m and 0.27 m in the urban area. By checking the maximum depth in MS 2, the difference of 0.38 m between MS 3c and MS 1 is subdivided by 20 cm through the rainfall pumps (the difference between MS 3 and MS 2) and 18 cm through the

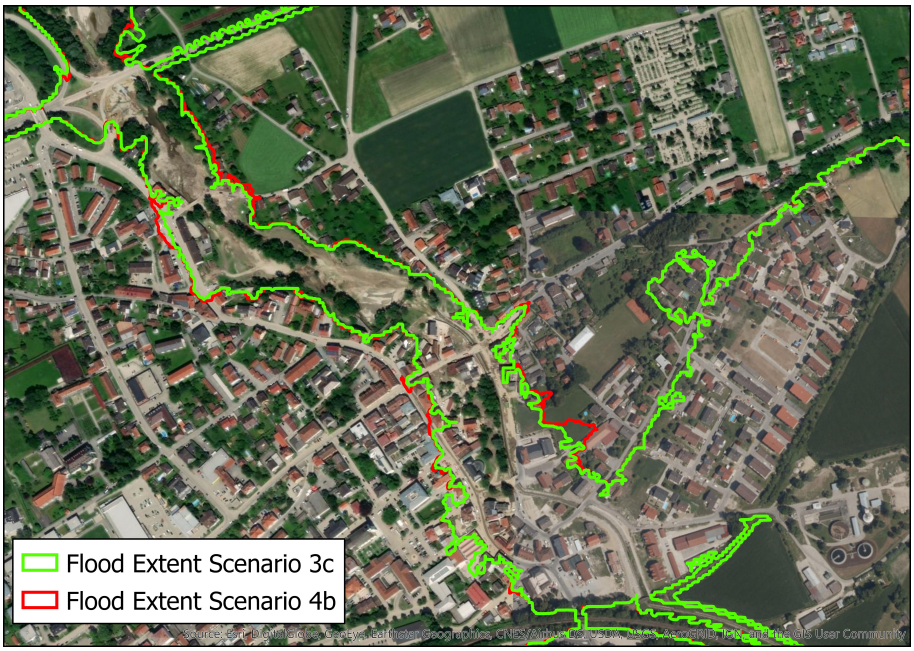



**Figure 42** Flood extents MS 3c and MS 4b, downstream

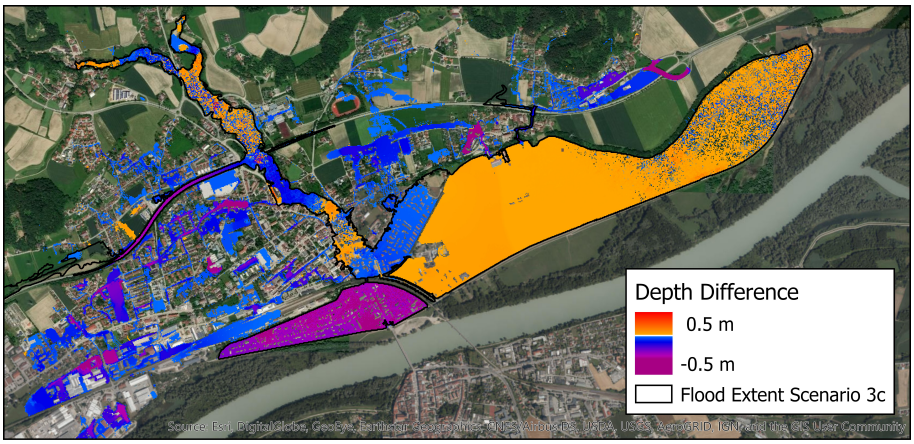



**Figure 43** Impact of the sewer system on the maximum depth (MS 3c minus MS 1)

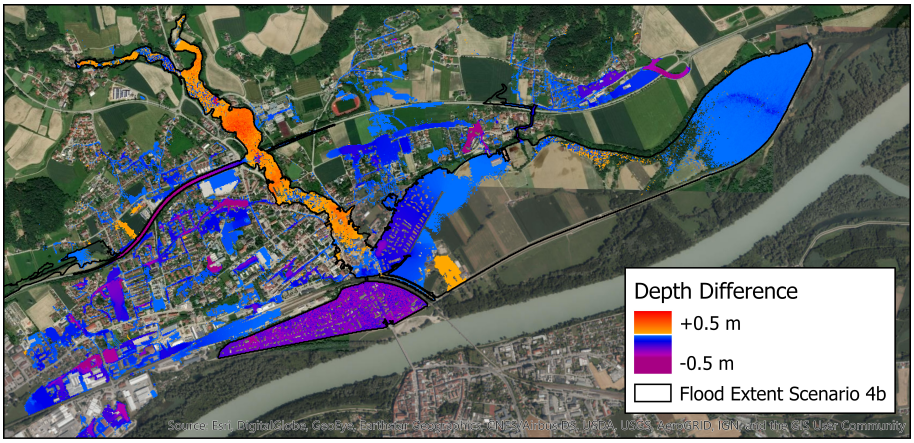



**Figure 44** Impact of the sewer system on the maximum depth (MS 4b minus MS 4a)

sewer system.

# 7. Probabilistic Flood Map

In hydraulic urban pluvial flood modeling, different choice of modeled components and modeling couplings makes a great variety of modeling the same area, causing a difference in model structure uncertainties. A probabilistic flood map is a helpful tool for visualizing the uncertainties from multiple model structures for quantifying the uncertainties and integrating the results from multiple model structures. In this work, two Bavarian cities are chosen for different sources of model structure uncertainty, Simbach is modeled with seven different modeled components, and Baiersdorf with seven model couplings. The probability flood maps generated by three different weighting methods are applied to visualize the model structure uncertainty in both cities. From the result, all three probabilistic flood maps show more potential flooding areas; thus perform better than a single deterministic flood map. Based on the study cases, the uncertainties from model couplings (up to 22%) are more significant than those from modeled components (up to 10%).

## 7.1. Methodology

### 7.1.1. Probabilistic Inundation Map

This study applies the probabilistic flood map proposed by Di Baldassarre (Di Baldassarre et al., 2010). The procedure tests the simulation from different 2D flood models based on the historical flood data to measure fit and combines the multiple models with their fits as a new model. This approach would reduce the uncertainty from a specific model structure. Thus, to avoid the incorrect assessment of flood hazard events.

In this study, seven model structures model the drainage components in Simbach, and another seven model structures model the coupling details in Baiersdorf. Both the study cases are analyzed with their probabilistic inundation maps by the three methods for the two historical flood events (see Section 7.1.8 for details). All the weights for generating the probabilistic inundation maps apply the model goodness of fit of the simulated maximum water level to the measured maximum water level on the checkpoints.

### 7.1.2. Probabilistic Inundation Map Extent

The probabilistic inundation map is the weighted results of flood maps of different model structures.

$$C_j = \frac{\sum_i L_{ij} w_{ij}}{\sum_i L_{ij}} \tag{7.51}$$

where $w_{ij}$ is the simulation result of Cell $j$ in the $i$'th model structure, as each cell with bit inundation information (0 for dry and 1 for inundated); $L_{ij}$ is taken from the different weighting

functions defined in Section 7.1.3, 7.1.4 and 7.1.5; $C_j$ is the inundation probability in Cell $j$ after weighting. All these cell values in the probabilistic inundation map range from 0 to 1, representing the inundation probabilities of evaluated model structures.

### 7.1.3. Probabilistic Inundation Map by Overall Weights (OW)

The first method evaluates model structure uncertainty following the suggestion from (Di Baldassarre and Montanari, 2009) the Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992):

$$L_{ij} = L_i = \frac{F_i - \min(F_i)}{\max(F_i) - \min(F_i)} \tag{7.52}$$

where the $F_i$ is the measure of fit of Model Structure $i$, with a value representing a better model fit. $\max(F_i)$ and $\min(F_i)$ take the maximum and the minimum of the model fit from the model structures. Therefore, after normalization by Formula 7.52, the likelihood weight $L_{ij}$ takes the homogeneous value $L_i$ for all the cells, ranging $[0, 1]$.

This study chooses the Nash-Sutcliffe efficiency coefficient (NSE) to measure the fit on the measured maximum water depth at the checkpoints.

$$F_i = 1 - \frac{\sum_{k=1}^{p} (h_m^{i,k} - h_o^k)^2}{\sum_{k=1}^{p} (h_o^k - \overline{h_o})^2} \tag{7.53}$$

For Model Structure $i$, where $p$ is the number of checkpoints, $h_m^{i,k}$ is the simulated maximum water depth at checkpoint $k$, $h_o^k$ is the observed maximum water depth at checkpoint $k$, and $\overline{h_o})$ is the mean of the observed maximum water depth overall checkpoints. The NSE ranges from negative infinite to one when a value closer to 1 represents a better fit of the model structure to the measured water depth at the checkpoints.

### 7.1.4. Probabilistic Inundation Map by Nearest Checkpoint Weights (NCW)

For a flood near a checkpoint, it is reasonable to rank the better-fit model structure higher by that checkpoint. Thus, the weighing of the probabilistic inundation map follows the rank of a model structure fit inside each patch for the nearest checkpoint.

The error is normalized to let the normalized performance (NP) range between 0 and 1 to coordinate the errors from different checkpoints, while 0 for the best-fit model. The normalized performance of model structure i at checkpoint k is calculated by:

$$\mathrm{NP}^{i,k} = 1 - \frac{e^{i,k} - \min_t e^{i,k}}{\max_t e^{i,k} - \min_t e^{i,k}} \tag{7.54}$$

where $\min_t e^{i,k}$ and $\max_t e^{i,k}$ are the minimum and maximum water level differences at checkpoint $k$ from all the model structures. $e^{i,k}$ is the water level differences at checkpoint $k$ of model structure $i$, calculated by the difference of simulated water depth $h_s^{i,k}$ from Model Structure $i$ at Checkpoint $k$ and the observed water depth $h_o^k$ at Checkpoint $k$, defined by:

$$e^{i,k} = |h_s^{i,k} - h_o^k| \tag{7.55}$$

The minimum and maximum errors are calculated among all the different model structures. The weighting function takes the normalized performance at the nearest checkpoint.

$$L_{ij} = \mathrm{NP}^{i,\hat{k}(j)} \tag{7.56}$$

The nearest checkpoint $\hat{k}(j)$ of any raster cell j is calculated:

$$\hat{k}(j) = \mathrm{argmin}_k d(j,k) \tag{7.57}$$

where $d(j,k)$ represents the Euclidean distance between a raster cell $j$ and a checkpoint $k$.

### 7.1.5. Probabilistic Inundation Map by Inverted Weighted Distances (IWD)

To combine the information of the model fits from different checkpoints, the third method considers the model structure goodness of fit on all checkpoints. By rescaling the normalized performance with the inverse of the distance to the checkpoint, the model gives a higher impact if one cell is closer to a checkpoint. Performance is measured by:

$$L_{ij} = \begin{cases} \dfrac{\sum_{k=1}^{p} \dfrac{\mathrm{NP}^{i,k}}{d(j,k)}}{\sum_{k=1}^{p} \dfrac{1}{d(j,k)}}, & \text{if } j \text{ is not a checkpoint} \\ \mathrm{NP}^{i,k}, & \text{if } j \text{ is Checkpoint } k \end{cases} \tag{7.58}$$

### 7.1.6. Taylor Diagram

For visualization of the model performance, three measures, namely, correlation, root-mean-square (RMSE), and the ratio of variance, are applied to the Taylor diagram for data representation (Taylor, 2001). The Taylor diagram helps visualize the relative merits between model structures and track the statistical significances as model structure changes. The Taylor diagram displays the model performance at the gauge stations relative to the observed maximum water depth in this work. The measured water depth was taken as the reference value in the diagram to quantify the model performance. In addition, the diagram allows us to visualize the uncertainty range and similarities among the different model structures.

### 7.1.7. Hydraulic Models for Urban Flood Modelling

In recent decades, different hydraulic models have been developed for urban flooding, with growing quality and complexity in the model structures (Mignot et al., 2019). The shallow water equations are widely applied in the hydraulic simulations of the physical processes of the water flows in urban floods (Henonin et al., 2013).

#### 1D Hydraulic Models

One-dimensional hydraulic models solve the Saint-Vernant continuity and depth-averaged equation in the longitudinal direction (Dimitriadis et al., 2016). Therefore, 1D models have higher computation efficiency, requiring less effort for model construction and calibration. Furthermore, the relatively low computational complexity enables the 1D model for extensive floodplains. Therefore, 1D hydraulic models are preferred for the most water in the catchment flow within the given flow path, such as street profiles or pre-defined river/sewer channels (Mark et al., 2004).

Among all the 1D hydraulic models, the Storm Water Management Model (SWMM) has been widely used for designing urban flood estimation since the 1970s (Rawls et al., 1980). The US Environmental Protection Agency developed the open-source program in 1971 for hydrology and water quality models in the world (Obropta et al., 2008). The model has been successfully applied in flooding analyses in various scales, from street blocks to catchments/mega-cities (Gironás et al., 2010).

#### 2D Surface Hydraulic Models

Two-dimensional (2D) hydraulic models solve the Saint-Vernant continuity and depth-averaged equation in the longitudinal and latitude directions. According to Hunter et al. (2008), 2D models are suitable to model complex flow paths with high-speed shallow-water flows in urban areas. Although 2D hydraulic models are computationally more expensive compared to the 1D model (Leandro et al., 2009), 2D models can produce the flood profiles resulting from flow ponding and rerouting in the urban area. Thus, 2D models are preferred for heavy flood

surface simulation in urban areas with streets and other complex infrastructures.

The 2D Parallel Diffusive Wave Model (P-DWave) (Leandro et al., 2014) is applied here. The model solves simplified shallow water equations by neglecting the inertial term and uses an explicit numerical form for discretization. Together with the scalability methods, the equation simplification increases the model performance in urban flooding simulation, enabling the model to deal with refined resolutions (less than 10 m).

Coupling of 1D and 2D Hydraulic Models

According to the impact of rain on the city, urban floods may cause (a) directly a large amount of precipitation above the city beyond the capacity of urban structures or (b) the precipitation on the large upstream area exceeding the infiltration capacity. Considering the water for the input to the urban hydraulic model in different forms, the water from case (a) could be the meteorological input to the hydraulic model, while the water from the case (b) could be calculated from a preliminary hydrologic or hydraulic model, for the hydrographs at the city boundary (Paquier et al., 2015). Different concepts of modeling a catchment could build different model structures for the same study area.

The dual drainage concept is widely implemented in urban flood simulations. The model based on this concept consists of the surface and the underground parts. The surface part contains channels, flow paths, and retention basins, and the underground part contains sewer systems of manholes, inlets, and control structures (Djordjević et al., 1999). The coupling of 1D models with 2D surface runoff models enables the simulation of surcharges of storm sewers as well as estimation of inundation depth (Huong and Pathirana, 2013). Besides, the bi-directional links built at the manholes allow the water to enter the drainage network and surcharge when the drainage system is saturated.

In this study, the 2D model P-DWave and the 1D model SWMM are coupled in some of the model structures within the urban area. The water levels are exchanged dynamically at manholes to coincide with the 2D surface runoff and 1D sewer drainage models.

### 7.1.8. Study Area and Materials

In this work, two study areas, Simach and Baiersdorf, are chosen to study model structure uncertainties. Simbach suffered from an extreme flash flood event in 2016. The city has several infrastructures, e.g., sewers, water treatment plants, pump systems, which could be included or excluded in a model, generating different model structures. Thus, Simbach is used to analyze model structure uncertainty from different modeled components. Furthermore, seven satellite cities surround Baiersdorf on its upper catchment within 5 km. Open drainage channels and sewer drainages connect the satellites and the Baiersdorf City. The less concerned flood inundations between the cities could be modeled by 1D instead of 2D to simplify the computational complexity. Thus, Baiersdorf is used to analyze model structure

uncertainties from model couplings.

Simbach am Inn 2016

The flash flood event in 2016 has caused significant damages and losses to the city, which triggered the local authorities' attention on flood prevention (Hübl et al., 2017). The measurement data collected by the University of Natural Resources and Life Sciences Vienna are used in the following studies (Hübl et al., 2017). In Simbach city, the water levels at 14 checkpoints validate the different model structures for the goodness of fit to the historical event (see red points in Figure 2).

Seven different model structures (MS) are implemented and calibrated to model the Simbach flash flood event. MS 1 models the urban area with a 2D surface runoff model. MS 2 implements the 1D/2D model in the urban area as 1D/2D dual drainage model. MS 3 adds rainwater pumps in Polder Erlach and Simbach based on MS 2. Furthermore, MS 4 adds the pump failure based on MS 3. MS 5 includes the wastewater treatment plant based on MS 4. MS 6 simulates the erosion of the street embankment based on MS 2. Finally, MS 7 includes all the mentioned model parts, the 1D/2D coupled model with pumps, pump failure, wastewater treatment plant, and street erosion modeled. Table 15 shows the summary of the seven model structures for studying the uncertainty with different modeled components.

Baiersdorf 2007

Figure 3 shows the study area Baiersdorf for the model coupling studies. The colored area represents Baiersdorf City, the satellites, and the south flood plain. Besides, the map also shows the sewer network within the catchment. Data from the eighteen gauge stations evaluate and weigh the model structures, plotted as red points on the map.

The following seven different model structures (MS) are constructed to study the model structure uncertainty from model couplings. The MS differ in the surface runoff by the model dimension for the urban and the upstream area, 1D-sewer networks, and the connections between urban satellites. MS1 models the upstream catchment as a 1D model and the urban area as a 1D/2D model, a coupling model of urban surface (2D), and the urban sewer system (1D). One-way links connect the upstream and urban parts. MS1, MS2 (1D & 1D/2D model — Satellites 2D Independent) further models the satellite villages upstream with a 2D surface model with one-directional connections on their boundaries. Different from MS2, in MS3 (1D & 1D/2D model -– Satellites 2D Dependent), the satellite villages in the upstream are simulated with the 2D surface model with bi-directional connections on the boundaries. Based on MS2, MS4 (1D & 1D/2D model — Satellites & South Floodplain 2D Independent) increases the 2D model area in the south between the satellites. Based on MS3, MS5 (1D/1D & 1D/2D model -– Satellites & South Floodplain 2D Dependent) also increases the 2D model area in

**Table 15** Summary of the model structures of Simbach

| | Inflows | Sewers | Rainwater Pumps | Pump Failure | WTTP Open | Dam Erosion |
|---|---|---|---|---|---|---|
| **MS 1 (2D)** | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ |
| **MS 2 (1D/2D)** | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ |
| **MS 3 (1D/2D)** | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ |
| **MS 4 (1D/2D)** | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ |
| **MS 5 (1D/2D)** | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ |
| **MS 6 (2D)** | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ |
| **MS 7 (1D/2D)** | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

the south between the satellites. MS6 (1D/2D model –– Whole extent 2D) models the whole extent by 2D surface runoff models coupled with the sewer network. Finally, MS7 (2D model — Whole extent 2D) models the whole study area using the 2D surface runoff model. Table 16 presents the summary of all the seven model structures of Baiersdorf.

## 7.2. Results and Discussion

### 7.2.1. Probabilistic flood inundation map of uncertainty on modeled components

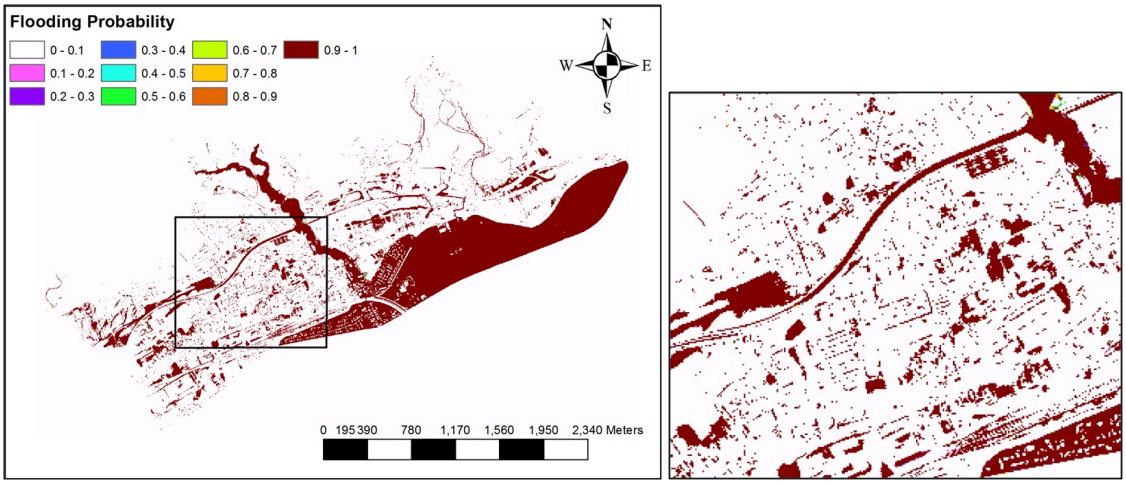
This section shows the uncertainty in the flood map from the different drainage components of Simbach 2016, based on 7 model structures from Table 15. Two methods (NCW, IWD) are applied for generating the probabilistic flood inundation map of the Simbach flood event 2016 according to the goodness of fit of the models, compared to the method (OW).

The first probabilistic inundation map (Figure 45) is calculated based on the overall weights (OW). The NSE values of the simulated water depth to the observed data from 14 checkpoints weigh the flood maps from the seven model structures. In Figure 45, the white pixels (with probability 0) are dry in all the model structures, while the brown pixels (with the probability

**Table 16** Summary of model structures of Baiersdorf

| | City | Sewers | Satellites | South Floodplain | Upstream |
|---|---|---|---|---|---|
| **MS 1** | 2D | ✓ | 1D | 1D | 1D |
| **MS 2** | 2D | ✓ | 2D independent | 1D | 1D |
| **MS 3** | 2D | ✓ | 2D dependent | 1D | 1D |
| **MS 4** | 2D | ✓ | 2D independent | 2D | 1D |
| **MS 5** | 2D | ✓ | 2D dependent | 2D | 1D |
| **MS 6** | 2D | ✓ | 2D | 2D | 2D |
| **MS 7** | 2D | ✗ | 2D | 2D | 2D |

2D independent: one-directional inflow on the satellites' boundary
2D dependent: bi-directional connection on the satellites' boundary



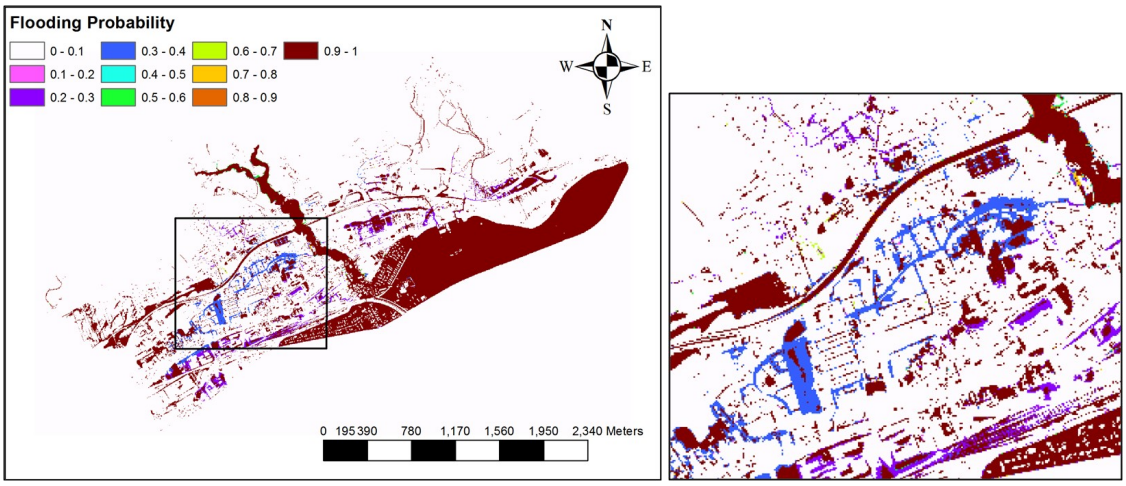

(a) (b)

**Figure 45** Probabilistic inundation map of Simbach weighted by the overall weights (OW)

of 1) are wet in all the model structures.

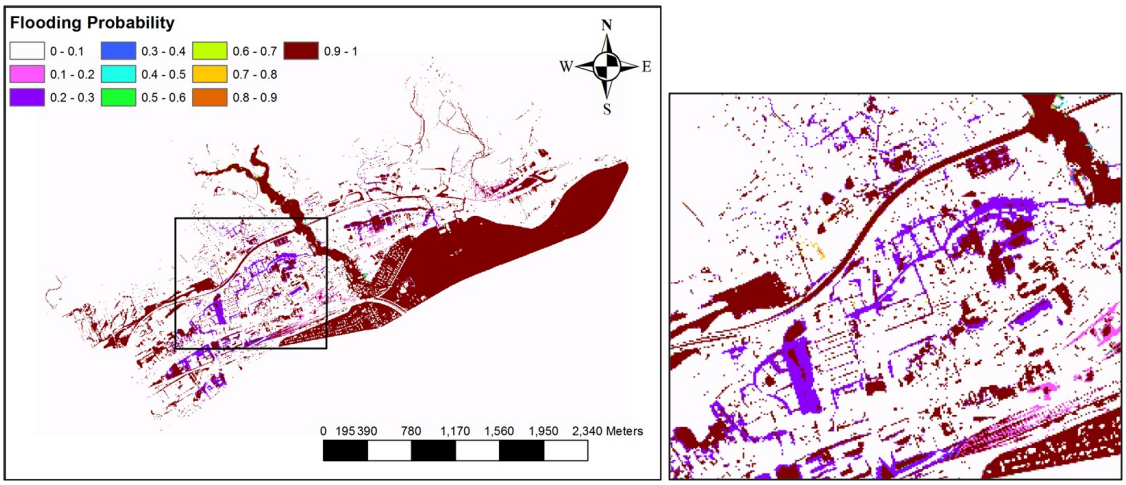
The second probabilistic inundation map (Figure 46) is calculated based on nearest checkpoint weights (NCW). The NCW normalizes the weighting factor of the goodness of fit at the nearest checkpoint. The goodness of fit takes 1 for the best and 0 for the worst.

The third probabilistic inundation map of Simbach (Figure 47) is calculated based on inverted weighted distance (IWD). Instead of weighting by the model fit from the nearest checkpoints, the weights here take the goodness of fit at all checkpoints inverted weighted by inverted




(a) (b)

**Figure 46** Probabilistic inundation map of Simbach by the nearest checkpoint weights (NCW)



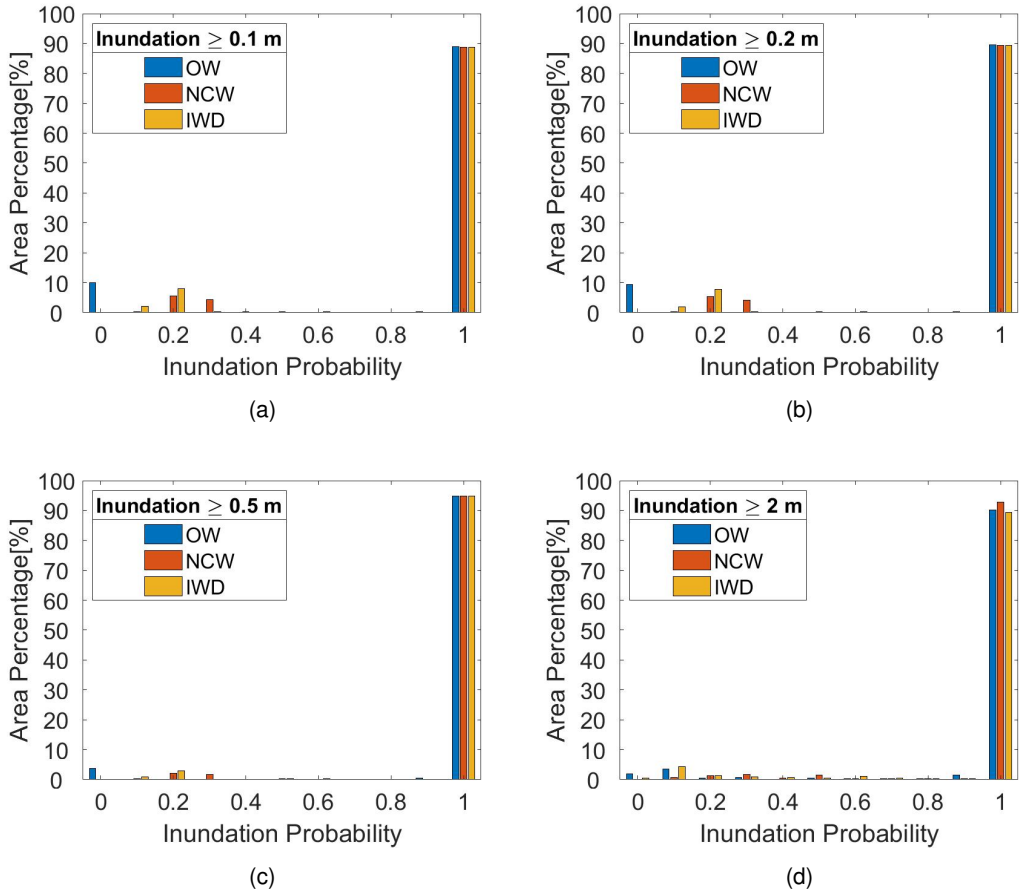

(a) (b)

**Figure 47** Probabilistic inundation map of Simbach weighted by Inverted Weighted Distances (IWD)

weighted distance.

Comparing the three different methods to display the flood inundation maps in Simbach, OW, NCW, and IWD (see Figure 45, 46 and 47), it is possible to observe that these are similar. The only exceptions are the two large uncertain flood areas in method NCW (Figure 46) and one large uncertain flood area in method IWD (Figure 47), which are visible in the city center. At the same time, method NCW (Figure 46) and method IWD (Figure 47) show uncertain flood areas in the west of Simbach city (the zoom-in area), the first method OW does not. These two methods can better display the uncertainty of the flooded area arising from the model structure of drainage components. Generally, for the uncertain inundation areas, NCW (Figure 46) shows a higher inundation probability than OW (Figure 45), which shows that the map weighted by the nearest checkpoints can capture a larger uncertainty than the first method (OW). Comparing the impact on the inundation depths by different drainage compo-

nents differs locally. If the weighting function depends solely on the overall weighting (OW), local differences disappear because a constant factor weights the map. The definition of the weighting function also explains the differences found for the inundation uncertainty area between NCW and IWD: NCW weighting uses the goodness of fit of the nearest checkpoints, highlighting, therefore, more local differences of each model structure.




(a) Inundation $\leq$ 0.1 m. (b) Inundation $\leq$ 0.2 m. (c) Inundation $\leq$ 0.5 m. (d) Inundation $\leq$ 2 m.

**Figure 48** Histograms of the percentage of inundation probabilities

Figure 48 shows the histograms of the percentage of inundation probabilities in the study area by three different methods to estimate the probability of inundation. A probability equal to one means that all the model structures agree on the wet areas. The inundation probability more negligible than one represents the uncertain inundation extent, dry in some of the model structures and wet in other model structures. The percentage of the uncertain extents reveals the model structure uncertainties.

Figure 48 shows the Taylor Diagram of the seven model structures for flood modeling of event 2016 in Simbach. The Taylor diagram presents the standard deviation, the correlation coefficient, and the root-mean-square standard deviation (RMSD) of the seven model structures. From the diagram, all the model structures have a similar correlation coefficient around 0.96,

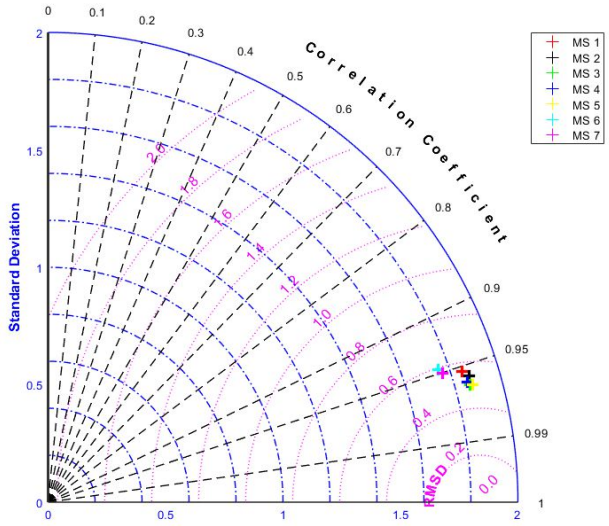



**Figure 49** Taylor Diagram of the different model structures compared to the measured water levels in Simbach

a standard deviation around 0.5 m, and an RMSD around 0.9 m. The seven model structures have a similar performance based on overall goodness of fit, validated by the observed data.

In Figure 48, around 90% of the inundation area has an inundation probability of 1. Therefore, the uncertainty from the drainage components could affect at most 10% of the inundated area for Simbach. For detailed analysis within the urban area, consideration of the drainage components impacts the simulation results. Figure 49 shows the goodness of fit from seven model structures at the 14 checkpoints in a Taylor diagram. The correlation and the standard deviation to the observed maximum water level show that all the model structures have similar overall performance. In this extreme flood event of Simbach, the flood volume exceeds the urban drainage capacity greatly. The drainage system carries a small percentage of the total flooded volume (0.9 – 5.2%). The overwhelmed status of the drainage component and the lower drainage efficiency is responsible for the low impact of the model structures in their performance (Figure 49), as well as in the flood inundation uncertainty (Figure 48).

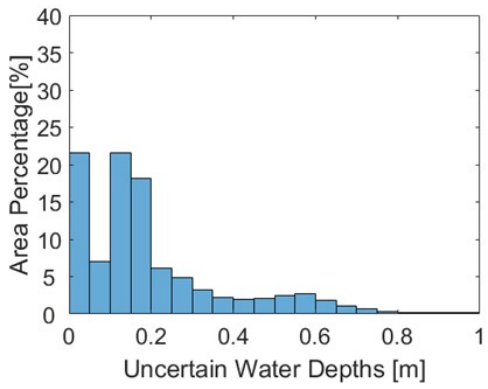



**Figure 50** Histograms of uncertain water depths in the seven model structures, Simbach

Figure 50 shows the percentage of uncertain water depths of the seven model structures with different drainage components for the event 2016 in Simbach. The uncertain water depths are defined as the maximum water depths minus the minimum water depths among the seven model structures to represent the spread of the structure uncertainties. From the histogram, 68.17% of the inundation area has a small uncertain water depth within 0.2 m. 28.49% of the inundated area is independent of the drainage components (uncertain inundation impact less than 0.1 m). Besides, 11.29% of the inundation area strongly depends on the drainage components, where the different drainage components can result in water depths of more than 0.5 m.

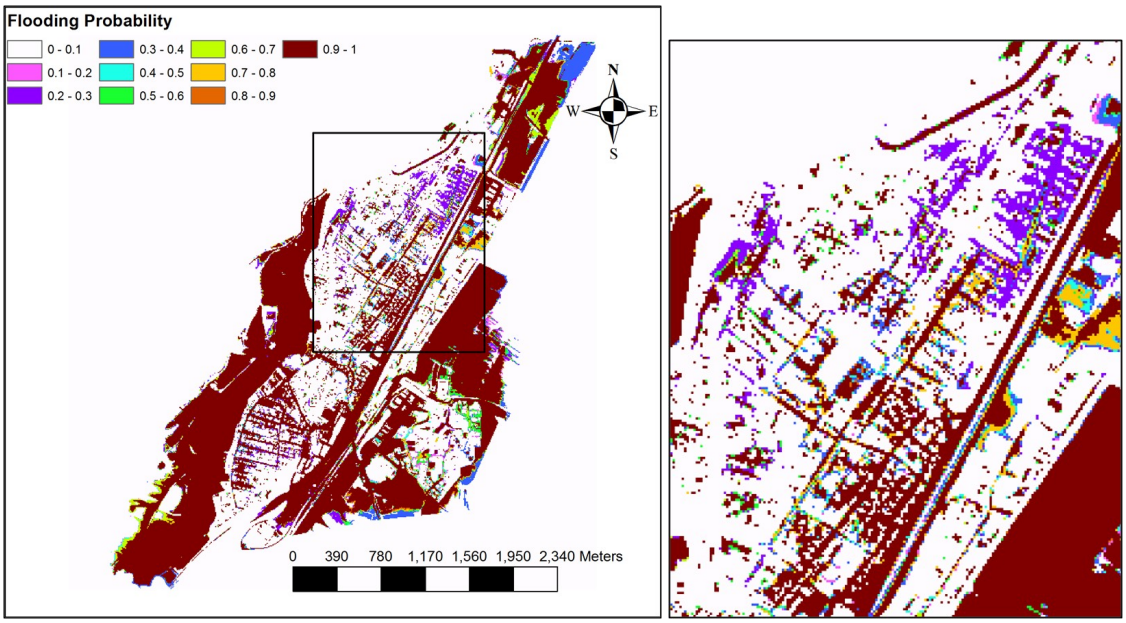
The observation is further quantified in Figure 48, among all the four inundation levels (0.1 m, 0.2 m, 0.5 m, 2 m), around 90% of the inundation area has the inundation probability of 1. In other words, The uncertainty from the modeled components could affect about 10% of the inundated area, with a probability between 0 and 1 on the maps as the model structures differ only in the sewer system, dam erosion, and pipe system (see details in Table 15). Therefore, modeling those drainage components only has local impacts in the urban area. Even though, Figure 48(d) shows inundation depth over 2 meters has a wide probability distribution between 0 and 1. It shows that the drainage components also cause a water depth difference within these areas. Therefore, a proper choice of the modeled components within the urban area is essential for accurate simulation results of these water depths. Figure 49 shows the goodness of fit from seven model structures at the 14 checkpoints in a Taylor diagram. By checking the correlation and the standard deviation to the observed maximum water level, it is noticeable that all the model structures have the similar overall performance to the observed values.

### 7.2.2. Probabilistic flood inundation map of uncertainty on model couplings

This section shows the uncertainty flood map of the different modeled coupling, based on 7 model structures from Table 16. Three different methods generate the probabilistic flood inundation map of the Baiersdorf flood event 2007 according to the goodness of fit of the models.

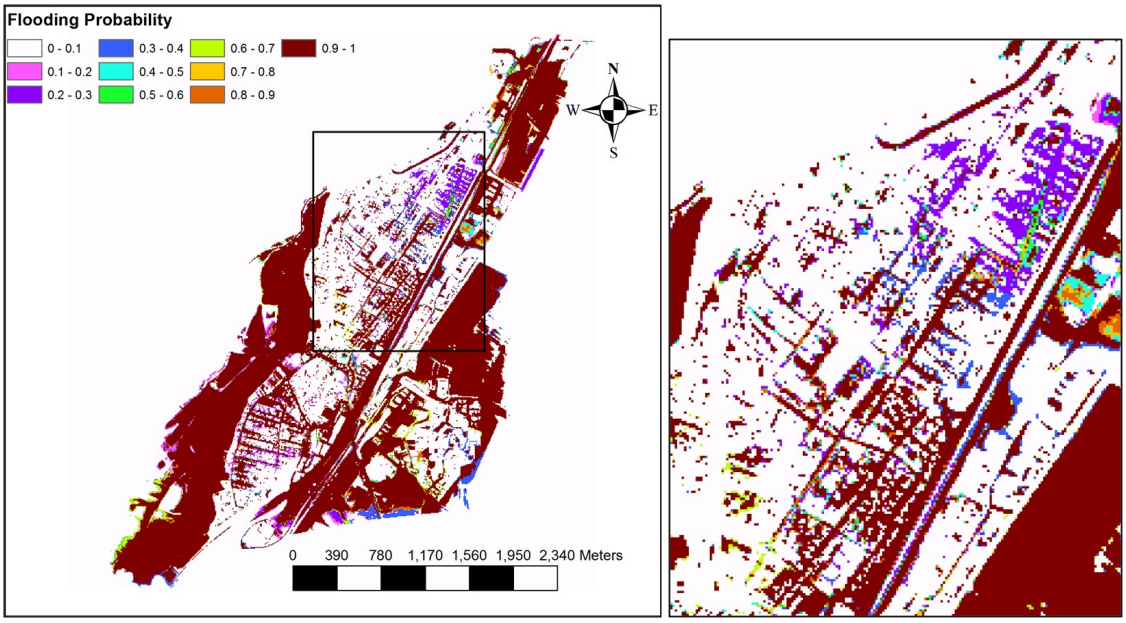
Figure 51 is calculated based on the overall weights (OW). The NSE values of the simulated water depth to the observed data from 18 checkpoints weight the flood maps from the seven model structures. In Figure 51, the white pixels (with probability 0) are dry in all the model structures, while the brown pixels (with probability 1) as inundated in all the model structures.

The second probabilistic inundation map (Figure 52) is calculated based on nearest checkpoint weights (NCW). For each pixel in the area, NCW normalizes the weighting factor to let the model structure with the best fit at the nearest checkpoint weight one and the model structure with the worst fit at the nearest checkpoint weight 0.




(a) (b)

**Figure 51** Probabilistic inundation map of Baiersdorf weighted by the overall weights (OW)

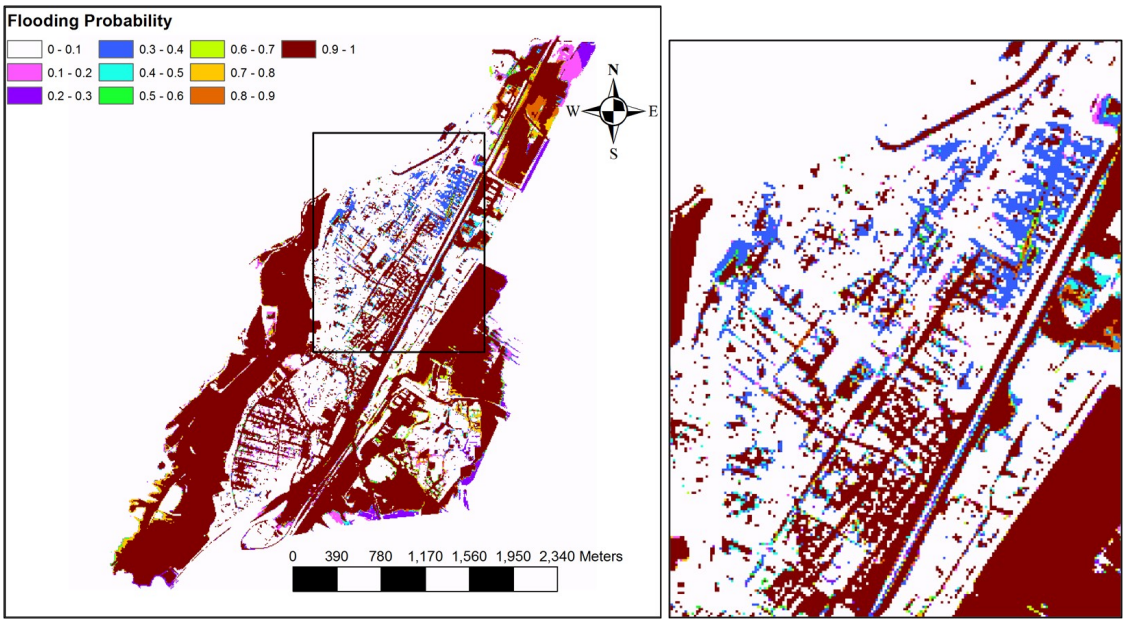



(a) (b)

**Figure 52** Probabilistic inundation map of Baiersdorf by the nearest checkpoint weights (NCW)

The third probabilistic inundation map of Baiersdorf (Figure 53) is calculated based on inverted weighted distance (IWD). The weights here take the goodness of fit from all the checkpoints with inverted weighted distance.

The three different methods, OW, NCW, and IWD, provide probabilistic flood inundation maps for the event in Baiersdorf with clearer differences compared to the case of Simbach. The



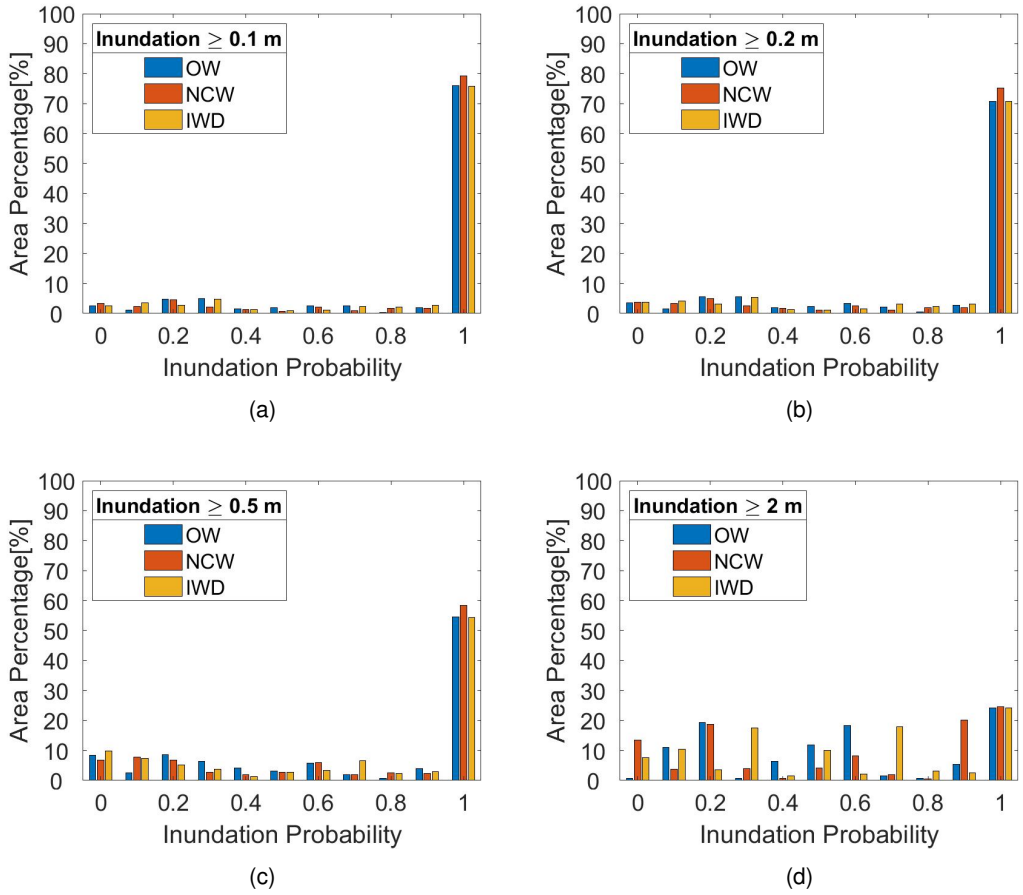

(a) (b)

**Figure 53** Probabilistic inundation map of Baiersdorf weighted by Inverted Weighted Distances (IWD)

method NCW (Figure 52) shows more uncertainty in the flood inundation than IWD (Figure 53), and IWD (Figure 53) shows more uncertainty than the OW (Figure 51). The method NCW (Figure 52) shows more inundation mainly in the urban area in the city of Baiersdorf than the method IWD (Figure 53). The method IWD (Figure 53) shows a larger inundated area than the method OW (Figure 51) in the southwest downstream.

While the inundation uncertainty steaming from the model components is mainly visible in the city center, the uncertainty from the model coupling is visible in both city centers and the surroundings. From all three different methods, OW, NCW, and IWD, the areas with a higher uncertain inundation probability (probability less than 1) distribute along or near the city boundary. In Baiersdorf, the highest probability appears near the main city boundary and in the zoom area (see orange areas in Figure 51, 52 and 53). The main city boundary corresponds to the outer limit of the 2D model area. Since different model coupling structures produce slightly different discharges inputs into the 2D model area, higher uncertainty appears first in the main city boundary (see blue areas in Figure 51, 52 and 53). These differences propagate then further into the 2D model and become visible in areas away from the boundaries (see purple and orange areas in Figure 51, 52 and 53).

Figure 54 shows the histograms of the percentage of inundation probabilities in the study area by three different probability maps. The inundation probability equals one represents the area inundated in all the model structures. The inundation probability more negligible than one represents the uncertain extent, dry in the sum of the model structures and wet in other model structures. The percentage of the uncertain extents reveals the model structure uncertainties.




(a) Inundation $\leq$ 0.1 m. (b) Inundation $\leq$ 0.2 m. (c) Inundation $\leq$ 0.5 m. (d) Inundation $\leq$ 2 m.

**Figure 54** Histograms of inundation probabilities Baiersdorf

The inundation probabilities further quantify the uncertain flood inundation areas in Figure 54. By looking at the percentage of the inundated area having no uncertainty (with the inundation probability of 1), the model structure uncertainty from model coupling increases with the increase of the flood depth. Especially for the inundation depths over 2 m, more than 75% of the area has a higher uncertainty arising from the different coupling structures. This observation shows (in these two examples) that the uncertainty from model coupling is several times larger than that from the drainage components (It now varies from 20% to 70% depending on the method applied.) Therefore, the choice of the model coupling should be planned carefully due to its higher impact on the overall flood inundation uncertainty.

Figure 55 shows the Taylor Diagram of the seven model structures for flood modeling of event 2016 in Baiersdorf. The Taylor diagram presents the standard deviation, the correlation coefficient, and the root-mean-square standard deviation (RMSD) of the seven model structures. From the diagram, all the model structures have a similar correlation coefficient of around 0.7.

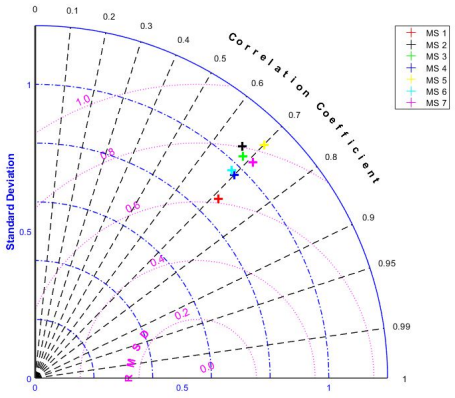



**Figure 55** Taylor Diagram of different model structures compared to the measured values, Baiersdorf

The inundation probabilities further quantify the uncertainty of inundation areas in Figure 54. By looking at the percentage of the inundated area having no uncertainty (with the inundation probability of 1), the model structure uncertainty from model couplings emerges with the growth of inundation depths. Especially for the inundation deeper than 2 m, more than 75% of the area is uncertain due to the model structure of different couplings. This observation shows, compared to the model structure uncertainty from the modeled components (with 90% uncertain area at all the inundation depths), the uncertainty from model couplings is several times larger. Therefore, it is essential to consider the choice of the model coupling due to its significant impacts on the inundation area. On the other side, to visualize the uncertainty from the model couplings, the probability flood inundation maps are helpful to join the information from different coupling model structures, thus giving a more precise estimation of the inundation area.

Besides, the spread of uncertainty probabilities in model couplings is also broader in all the inundation depths (see Figure 54). Therefore, the choice of proper model coupling is critical for its wide-ranging impacts in the study area over all the water depths. In the Taylor diagram (Figure 55), the model structure performance varies when constructed with different couplings. The standard deviation of the seven model structures varies from 0.6 to 0.8. The choice of more suitable model coupling could improve the goodness of fit to the observed water depths on the checkpoints, significantly larger than modeling more components in the study area with the same model coupling.

Figure 56 shows the percentage of uncertain water depths of the seven model structures with different model couplings of event 2007 in Baiersdorf. The uncertain water depths are defined as the maximum water depths minus the minimum water depths among all the seven model structures to represent the spread of the structure uncertainties. From the histogram, 71.86% of the inundation area has a small uncertain water depth within 0.2 m. 28.4% of the inundated area is independent of the model couplings (uncertain inundation impact less than 0.1 m). Besides, 4.47% of the area strongly depends on the model couplings, where the different model couplings can result in water depths more than 0.5 m.

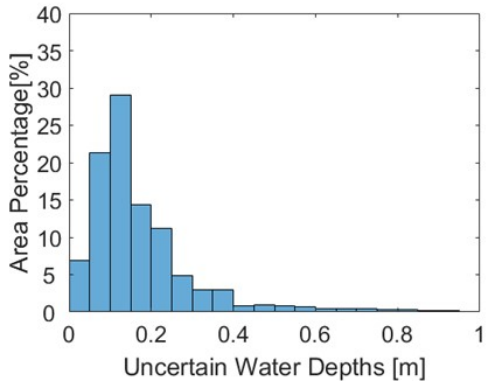



**Figure 56** Histograms of uncertain water depths from seven model structures Baiersdorf

### 7.2.3. The Advantages and Disadvantages of Probabilistic Inundation Maps

From all the probabilistic flood inundation maps, the certain flood areas (with the inundation probability of 1) in brown are calculated as the inundation by all the model structures. The other colors show the flood areas with uncertainty (with the inundation probability between 0 and 1), the weighted probabilities of inundated from some model structures, and non-inundated from other model structures. The probabilistic flood map gives more inundation information beyond a single flood inundation map from both study cases. From the Simbach study case, it is also visible that the probabilistic map weighted by overall NSE does not show much of the uncertain flood area, while the other two methods represent the uncertain area. In this case, we can conclude that the weighting factors considering local errors would be more suitable to identify some potentially inundated areas than overall weighting factors in generating probabilistic inundation maps. Finally, the probabilistic flood maps based on multiple model structures could overcome the limitation of a single flood model and improve the flood extent simulation results.

On the other side, unlike the probability inundation map in other studies based on ensemble runs (Papaioannou et al., 2017), the number of model structures in a study is usually small, which causes the probability inundation map to be sensitive to the outlier of incorrect model structures. Worse performance of one model structure would lead to the bias of the outcome probabilistic inundation map in some areas. Therefore, a model performance test is necessary before a model structure generates probabilistic inundation maps.

# 8. Summary and Conclusions

## 8.1. ANN for Flood Inundation Forecast

A forward-feed network structure was applied and set up with a training dataset with 120 synthetic events and a testing dataset of 60 events to produce the maximum inundation depth and compare it to the 2D hydraulic models. The two popular algorithms compared here, namely resilient backpropagation and conjugate gradient can provide a relatively accurate prediction of maximum inundation depths by evaluating the MSE in the study area. Besides, Resilient backpropagation performed better than the conjugate gradient with a smaller MSE on average.

The ANN topology selects two hidden layers and ten neurons per layer. The dependencies from the interaction of these two parameters are complex since there is no single setting that outperforms all the networks in this study. Therefore, it was impossible to find a clear trend over the ANN with a simple set of network layers and neurons. However, 70% of our networks performed better with two or three hidden layers. The comparison indicates that the prediction of flood inundation extents by inflow hydrographs is more likely to be precise at a low hidden layer number.

The network size for sub-dividing the area is also tested for performance evaluation. Both settings, network sizes of 50 × 50 and 100 × 100 for the study area, produced minor errors. However, the 50 × 50 grids have slightly smaller MSE with much less model tuning time, hence being chosen in this study.

Two clustering methods, conventional FCM and PCA-FCM, were tested for reducing the test dataset for performance. Both clustering methods capture the characteristics in each cluster (by the trend of curves and confidence intervals). By checking the 90% confidence interval over all the clusters, it could be inferred that the cluster spread of PCA-FCM was smaller than the spread with the conventional FCM clustering. Hence, the former was preferred. The MSE difference map of the clustering strategy showed that this strategy efficiently reduces the size of the training set. Thus, clustering is helpful as the preprocessing of the training dataset. The network could cover a broader range of inputs with the clustered data and avoid overfitting by similar training data. Overall, in our case study, clustering enhanced the performance of the ANN training by reducing the size of the training set and slightly improved the prediction of maximum flood inundation.

The prediction results on the testing dataset are excellent. The prediction of maximum flood inundation shows no visible difference from the synthetic events in flood extent and water depth. By comparing the MSE, only 1.21% of the wet grids have values larger than 0.2 $m^2$,

suggesting that the prediction is successful over 98% ANN. Tests on historic events showed that the prediction results of the flood inundation are still outstanding but with some localized disagreements in the maximum water depths. Overall the prediction by grids, 91.03% in 2005 and 86.38% in 2013, was good, for which the MSE was smaller than 0.2 $m^2$. However, it was also noticeable that the model prediction quality decreased as the forecast area was further away from the inputs.

The work above proved that resilient backpropagation networks could replace the 2D hydraulic model for flood inundation prediction, requiring only the discharge inflows as inputs, which are easier to measure in real-time than the physical-based parameters of the hydraulic model.

In the next step, the neural network performs multistep forecasts for one, two, three, four, and five hours after the flooding event has started. The results show that it is possible to use different ANNs for the first forecast interval (time 0) to issue the multistep forecasts. First, however, the quality of the forecast regarding the water-depths or the flood inundation extent should be distinguished. The overall forecast performance of the water-depths was found slightly better than the flood inundation extents. The performance was mainly affected by the flood event from 2005, particularly close to the southwest end, far away from the location where the input inflows are.

The flood predictions of the different first intervals were tested in synthetic and historical flood events, namely that of 2005, 2006, and 2013. Resilient backpropagation networks are set up and trained for forecast intervals of three, six, nine, and twelve hours. The model produces good results for the 60 synthetic flood events, as over 81% of the grids have RMSE below 0.3 m. Concerning the water depths, the ANN model has a good performance for the Flood Events 2006 and 2013 and a sufficient performance for the Flood Events 2005 in predicting water depths. Generally, the error increases gradually as the forecast interval increases by checking the inundation maps. For the far-end area away from the inflow inputs, the long-distance may be responsible for a decrease in the forecast performance; therefore, it is likely that the model requires other information than those of discharge to enhance the forecast accuracy for those areas.

The ANN model was applied for the real-time forecast in 2006, 2013, and 2005. For this purpose, the same ANN model was used for the forecast. The input discharge inputs were replaced by the shifted intervals for 1 – 5 h after the event beginnings. The forecast shows promising results in the flood events 2006 and 2013 for the real-time forecasts, with over 70% of grids with RMSE less than 0.3 m. The forecast shows worse results in flood event 2005, with only over 58% grids with RMSE less than 0.3 m. Overall, the forecast accuracy drops as the forecast interval increases from 3 h to 12 h. The forecast accuracy also decreases as the forecast progresses from $X$ h +1 h to $X$ h + 5 h. For the three historic flood events, the three hours forecast is good, with more than 70% of the grids accurately forecasted. However, the

quality of 6 h or longer intervals was more event-dependent.

Based on the analysis of POD, FAR, and CSI indices, the multistep ANN flood forecast provides good results at the beginning and decreases as the forecast progresses. The forecast switches from overestimating to underestimating when the forecast proceeds from 0 h to 5 h. In our case, except for the event 2005, the three hours ANN trained by the first interval improved the performances slightly with the sub-sequential forecast. The six, nine, and twelve hours ANN trained by the first interval for sub-sequential interval forecasts would have an accuracy depending on the actual flood events.

## 8.2. Probabilistic Flood Maps for Model Structure Uncertainty Quantification

To better model the growing occurrences of urban flooding hazards, the coupled 1D/2D hydrodynamic models are conducted for investigating the model structure impacts the 2016 Simbach flash flood event. The sewer network, the dam erosion, the pump system (rainwater pumps, pump failure), and the water treatment plant are modified in seven model structures to test their modeling impacts. Meanwhile, a 1D rainfall-runoff model models the inflow to Simbach city.

The statistical evaluation of all the model structures of model couplings showed satisfactory results, with a maximum RMSE and MAE of 0.594 m and 0.496 m and a minimum NSE of 0.896. Therefore, the 1D rainfall-runoff model for the inflow generation and the model grid size of 5 m by 5 m are validated suitable for the study. Furthermore, including dam erosion into the model structure successfully improved discharge rates and wider flooding extent in the city.

The sewer system reduced the average maximum water depth by 4.65%. The sewer system drains 0.9% – 5.2% of the total network inflow. Meanwhile, 69% – 73% of the total volume stayed on the surface, while 26% – 28% routed to the boundary of the study area. The rainwater pumps drain 3.9% – 4.3% of the total net inflow out of the folder, which reduces the maximum water depth in the Simbach Polder by up to 35 cm. For the model structure with dam erosion, the dam water-path reduces the efficiency of the sewer system. For the area only affected by local precipitation, the sewer network relocates a large amount of water to the riverbed of Simbach and the Erlach Polder.

Besides the advantages of the sewer system, the stormwater drainage and emergency flooding outlets also added to a massive volume of water relocation, which increases the water level inside the riverbed and the inundation area.

It is crucial to realize that the current implementation of the model structure study added great

uncertainty when introducing the different model structures. In our case, the wide spreads of averaged maximum water depth and volume significantly impact the small changes in modeled components. Furthermore, in extreme flash flood events such as Simbach 2016, the discharge data are usually scarce. Consequently, the model validation would rely more on the eyewitness observation, which adds to more uncertainties from the measurements.

Other similar case studies should also consider the necessity of drainage components modeling. In our case, the model still gives good results on the water depth and the inundation extent without modeling the dam break. Nevertheless, if the interest focuses more on the complex interaction between the 2D overland flow and the 1D sewer system, dam erosion, and pump systems modeling could increase the modeled system's accuracy and increase model complexity.

A probabilistic inundation map is a helpful tool for analyzing the model structure uncertainty of hydraulic flood models. The model structure uncertainties herein are studied in two sources: from modeled components and model couplings, respectively in the flood event Simbach 2016 and the flood event Baiersdorf 2007. Seven model structures of modeled components and seven models of model couplings are set up and calibrated for the probabilistic inundation maps. Sixteen water gauges from Simbach and eighteen from Baiersdorf provide the maximum water depths in the flood events, determining the weights in probabilistic flood inundation maps.

Three methods for calculating the probabilistic flood inundation maps weigh the goodness of fit for the model structures: overall fit, fit at the nearest checkpoints and fit at all checkpoints. To further quantify the model structure uncertainty of the original model structure, Taylor diagrams help to show the performance of modal structures at the checkpoint data by root-mean-square errors, standard deviations, and correlation coefficients. Changing the model couplings could significantly impact the model fit at the observed checkpoint data from the diagram. In contrast, the change in the modeled components did not affect much on the results. In our case, improving the model couplings could shrink the standard deviation of the errors to a quarter, which points out the importance of the coupling strategy in hydraulic urban flood simulation. On the other side, modifying drainage components did not affect the model performance much, probably due to the complexity of the model structures. Adding a model component did not change much on the model performance, which could result from either the local impacts or the extra uncertainties from adding the components themselves.

The probabilistic flood maps give more inundation information over the simulated flood events than from a single deterministic flood map in the two study cases. With the different model structures of the modeled components, about 10% of the inundation area is uncertain (with a probability between 0 and 1), which is possible to neglect when applying a deterministic model structure. In the study of model couplings, more than 20% of the inundation area is uncertain. Based on these results, using multiple model structures and constructing probabilistic flood

maps in studies are essential means to overcome the limitation of a single model and obtain a more accurate flood extent.

## 8.3. Outlook and Future Work

As the objective of the real-time flood forecast, the forecast model could give accurate water depth profiles. However, due to the great variety of artificial neural network models, it is possible to have a better choice of other neural network models. Future work could include other neural network models, e.g., a recurrent neural network with long short-term memory, to include the water depth information acquired from previous forecast steps for a multi-step forecast. To further split the forecast time interval for finer temporal multi-steps could also be another possibility to enhance the accuracy of the forecast. For instance, a Convolutional Neural Network (CNN), a Generative Adversarial Network, and other networks could be alternatives. On the other side, it is meaningful to include more dynamic processes when processing multi-step inundation forecasts.

The model structure uncertainty analysis in this work proceeded on the two specific case studies on drainage components and model couplings. Therefore, the model uncertainties are strongly impacted by the local topologic features, the hydrologic characteristics, and the local drainage infrastructures. Consequently, the model variants are also limited to the drainage/couplings in the cases. Future works should extend to other areas with different drainage infrastructures and different coupling from urban planning to overview the two sources of structure uncertainties. By comparison results from different study cases, the uncertainty from drainage components and model couplings could be better quantified and concluded. Besides, it is necessary to include and analyze other sources of model structure uncertainties, e.g., different discretization of the urban terrain model, numerical schemes, formations of one-dimensional surface models, and different land cover representations.

The precipitations applied to different areas are worth considering for flood inundation modeling in multiple urban areas. The historical precipitation intensity is input for the simulation to integrate the gauge station data. Thus, the inequivalent precipitation intensities in different urban areas would also make a difference in comparing uncertainties. Furthermore, applying the same return period's precipitation would be interesting to assess the uncertainty at the same hazard intensity for multiple urban areas.

Due to the heavy computational expenses of two-dimensional surface models, the model structure could only evaluate the limited possible model structures at a time. Suppose the time-consuming two-dimensional part could be modified. In that case, the simulation time for a single model structure could be much shorter, which increases the number of model structures in our uncertainty analysis study. Consequently, future work could include replacing the two-dimensional model with less expensive models. For example, it would be helpful to include only the rapid flood spreading model or some replicable indices from a one-dimensional

model to reduce the computation efforts. Besides, if the AI-based model could achieve higher accuracy in simulating the 2D surface model, it could be replaced for model structure uncertainty analysis. The potential capability of generating hundreds or thousands of ensembles enables the complete variants among model structures, making the uncertainty analysis less biased. Besides, different visualization methods should be applied to more study cases for further comparisons with more complex structural differences.

It is also necessary to include other measurements for the model performance in future studies. The performance evaluation depends on the water levels at the gauge stations; in our case, fewer than twenty for both cases. Most cities can have a similar issue, that the total number of gauge stations is limited, causing a high uncertainty in the measured water level in an extreme flood event. Moreover, the gauge stations' distribution is usually within the city center, which reduces a lower weighing of the model fit in the suburban areas. Finally, a more comprehensive evaluation of the inundation area could include inundation information available from drones and cameras.

# List of Supervisions

| | | | |
|---|---|---|---|
| Abdullah Fadeen Sawari | A Study on Influence Factors of Flash Floods | master thesis | 11.02.2021 |
| Aitzaz Ahsan | A Study on the Influence Factors of Flash Floods by 2D Hydrodynamic Model – A Case Study in Baiersdorf, Bavaria | study project | 01.02.2021 |
| Akshaya Arivalagan | 1D (SWMM) – 2D (P-D Wave) coupled modelling for flood inundation analysis in Kulmbach | study project | 29.11.2019 |
| Benedikt Bauer | Analyse des Einflussfaktors Gebäude auf das Verhalten von Sturzfluten mit Hilfe des hydrodynamischen 2D-Modells "P-DWave" Kulmbach Sued | study project | 04.09.2020 |
| Birgit Bell | 1D Model set up and validation of the sewer network of the city of Simbach using PCSWMM | study project | 24.04.2019 |
| Daniel Sturm | A Study of the Simbach Flash Flood of 2016 using a 2D parallel diffusive wave model (P-DWave) | study project | 04.06.2019 |
| Daniel Sturm | A Study on the Simbach Flash Flood of 2016 using a coupled 1D/2D Model (SWMM/P-DWave) | master thesis | 09.03.2020 |
| Felix Schmid | 1D – Überflutungsanalyse des Kanalnetzes Baiersdorf mittels PCSWMM | study project | 19.12.2019 |
| Felix Schmid | An artificial intelligence framework for fore-casting time series of sewer surcharge | master thesis | 29.01.2021 |
| Federica Scholari | A Study on Influence Factors of Flash Floods by 2D Hydrodynamic Model A Case Study in Simbach am Inn | study project | 18.12.2019 |
| Hongfei Zhu | Optimization of Artificial Neural Network (ANN) based Flood Inundation Forecast: Case Study of Kulmbach | study project | 18.12.2020 |
| Ignacio Mollá Carcaño | A Study on the Influence of Buildings in Flash Floods in the City of Baiersdorf using a 2D parallel diffusive wave model (P-DWave) | study project | 04.02.2020 |

| | | | |
|---|---|---|---|
| Ignacio Mollá Carcaño | Flood risk management: A preliminary study: Los Alcázares. Región de Murcia. Spain. | master thesis | 04.02.2020 |
| Laura Dendorfer | Analyse des Einflussfaktors Gebäude auf das Verhalten von Sturzfluten mit Hilfe eines hydrodynamischen 2D-Modells | study project | 17.03.2020 |
| Lena Cora Weisgerber | Analyse des Einflussfaktors Gebäude auf das Verhalten von Sturzfluten mit Hilfe des hydrodynamischen 2D-Modells "P-DWave" Kulmbach Nord | study project | 09.11.2020 |
| Leena Julia Warmedinger | A Study on Influencing Factors of Flash Floods with the Hydrodynamic Model P-DWave | study project | 31.01.2020 |
| Mengmeng Chen | The Impact of the Sewer System and Climate Change on the Flood Resilience Index for the Maxvorstadt 1D-2D Coupled Model | master thesis | 12.11.2019 |
| Ragini Bal Mahesh | Physics Informed Neural Network for Spatial Temporal Flood Forecasting | master thesis | 17.12.2020 |
| Sayed Ahmed Fawad Sadat | The Impact of Average Influence Factors of Flash Floods on Inundation Prediction with ANN | master thesis | 11.02.2021 |
| Sebastian Tobias Ramsauer | Inclusion of Buildings in Coarser Grids for Urban Flood Modeling | master thesis | 08.03.2021 |
| Stefan Gerber | Real Time Flood Forecasting Using Artificial Neural Networks (ANN), Kulmbach Case Study | master thesis | 27.02.2020 |
| Teresa Quintana | Evaluation of Urban Flood Model Structures to Assess the Sewer Network Impact during Flash Floods | master thesis | 11.09.2020 |
| Yu Chen Liu | Flood Forecasting Using Support Vector Machine for the Gaoping River Watershed in Taiwan | master thesis | 15.09.2020 |
| Veronika Zwirglmaier | Flood Inundation Prediction with Bootstrap Based Artificial Neural Networks (BANNs) — A Case Study in Kulmbach | study project | 23.04.2019 |

| | | | |
|---|---|---|---|
| Wenrong Wu | Prediction of Flood Inundation Extents with Artificial Neural Networks (ANN): Case Study of Kulmbach | master thesis | 24.04.2019 |

# Reference


Abbot, J. and Marohasy, J. (2015). Improving monthly rainfall forecasts using artificial neural networks and single-month optimisation: a case study of the Brisbane catchment, Queensland, Australia. *Water Resources Management VIII*, 1:3–13.

Abily, M., Bertrand, N., Delestre, O., Gourbesville, P., and Duluc, C. M. (2016). Spatial Global Sensitivity Analysis of High Resolution classified topographic data use in 2D urban flood modelling. *Environmental Modelling and Software*, 77:183–195.

Ajami, N. K., Duan, Q., and Sorooshian, S. (2007). An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resources Research*, 43(1):1–19.

Apel, H., Aronica, G. T., Kreibich, H., and Thieken, A. H. (2009). Flood risk analyses—how detailed do we need to be? *Natural Hazards*, 49(1):79–98.

Aronica, G., Bates, P. D., and Horritt, M. S. (2002). Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. *Hydrological Processes*, 16(10):2001–2016.

Atkinson, S. E., Woods, R. A., and Sivapalan, M. (2002). Climate and landscape controls on water balance model complexity over changing timescales. *Water Resources Research*, 38(12):50–1–50–17.

Bamford, T., Balmforth, D., Digman, C., Waller, S., and Hunter, N. (2008). Modelling flood risk assessment, an evaluation of different methods. In *WaPUG autumn conference*.

Bates, P. D., Horritt, M. S., Aronica, G., and Beven, K. (2004). Bayesian updating of flood inundation likelihoods conditioned on flood extent data. *Hydrological Processes*, 18(17):3347–3370.

Berkhahn, S., Fuchs, L., and Neuweiler, I. (2019). An ensemble neural network model for real-time prediction of urban floods. *Journal of Hydrology*, 575(February):743–754.

Bermúdez, M., Cea, L., and 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(S1):1–14.

Bernhofen, M. V., Whyman, C., Trigg, M. A., Sleigh, P. A., Smith, A. M., Sampson, C. C., Yamazaki, D., Ward, P. J., Rudari, R., Pappenberger, F., Dottori, F., Salamon, P., and Winsemius, H. C. (2018). A first collective validation of global fluvial flood models for major floods in Nigeria and Mozambique. *Environmental Research Letters*, 13(10).

Berz, G. (2001). Flood disasters: Lessons from the past-worries for the future. *Water Management*, 148(1):57–58.

Beven, K., Almeida, S., Aspinall, W., Bates, P., Blazkova, S., Borgomeo, E., Goda, K., Hall, J., Phillips, J., Simpson, M., Smith, P., Stephenson, D., Wagener, T., Watson, M., and Wilkins, K. (2017). Epistemic uncertainties and natural hazard risk assessment. 1. A review of different natural hazard areas. *Natural Hazards and Earth System Sciences Discussions*, pages 1–53.

Beven, K. and Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6(3):279–298.

Beven, K., Lamb, R., Leedal, D., and Hunter, N. (2015). Communicating uncertainty in flood inundation mapping: A case study. *International Journal of River Basin Management*, 13(3):285–295.

Bhola, P. K., Leandro, J., and Disse, M. (2018). Framework for offline flood inundation forecasts for two-dimensional hydrodynamic models. *Geosciences (Switzerland)*, 8(9).

Bhola, P. K., Nair, B. B., Leandro, J., Rao, S. N., and Disse, M. (2019). Flood inundation forecasts using validation data generated with the assistance of computer vision. *Journal of Hydroinformatics*, 21(2):240–256.

Brunner, G. W. (2016). *HEC-RAS River Analysis System Hydraulic Reference Manual*. Number February.

Bulti, D. T. and Abebe, B. G. (2020). A review of flood modeling methods for urban pluvial flood application. *Modeling Earth Systems and Environment*, 6(3):1293–1302.

Bustami, R., Bessaih, N., Bong, C., and Suhaili, S. (2007). Artificial Neural Network for Precipitation and Water Level Predictions of Bedup River. *IAENG International Journal of Computer Science*, 34(2):228–233.

Butts, M. B., Payne, J. T., Kristensen, M., and Madsen, H. (2004). An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of Hydrology*, 298(1-4):242–266.

Chang, L. C., Amin, M. Z. M., Yang, S. N., and Chang, F. J. (2018a). Building ANN-based regional multi-step-ahead flood inundation forecast models. *Water (Switzerland)*, 10(9):1–18.

Chang, M. J., Chang, H. K., Chen, Y. C., Lin, G. F., Chen, P. A., Lai, J. S., and Tan, Y. C. (2018b). A support vector machine forecasting model for typhoon flood inundation mapping and early flood warning systems. *Water (Switzerland)*, 10(12).

Chau, K. W., Wu, C. L., and Li, Y. S. (2005). Comparison of Several Flood Forecasting Models in Yangtze River. *Journal of Hydrologic Engineering*, 10(6):485–491.

Chen, A. S., Leandro, J., and Djordjević, S. (2016). Modelling sewer discharge via displacement of manhole covers during flood events using 1D/2D SIPSON/P-DWave dual drainage simulations. *Urban Water Journal*, 13(8):830–840.

Chibueze, T. and Nonyelum, F. (2009). Feed-forward Neural Networks for Precipitation and River Level Prediction. *Adv. in Nat. Appl. Sci*, 3(33):350–356.

Chowdhury, S. and Sharma, A. (2007). Mitigating parameter bias in hydrological modelling due to uncertainty in covariates. *Journal of Hydrology*, 340(3-4):197–204.

Chu, H., Wu, W., Wang, Q. J., Nathan, R., and Wei, J. (2020). An ANN-based emulation modelling framework for flood inundation modelling: Application, challenges and future directions. *Environmental Modelling and Software*, 124:104587.

Coulibaly, P., Anctil, F., and Bobée, B. (2000). Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *Journal of Hydrology*, 230(3-4):244–257.

Dawson, C. W. and Wilby, R. L. (2001). Hydrological modelling using artificial neural networks. *Progress in Physical Geography*, 25(1):80–108.

Di Baldassarre, G. and Montanari, A. (2009). Uncertainty in river discharge observations: A quantitative analysis. *Hydrology and Earth System Sciences*, 13(6):913–921.

Di Baldassarre, G., Schumann, G., Bates, P. D., Freer, J. E., and Beven, K. J. (2010). Cartographie de zone inondable: Un examen critique d'approches déterministe et probabiliste. *Hydrological Sciences Journal*, 55(3):364–376.

Dimitriadis, P., Tegos, A., Oikonomou, A., Pagana, V., Koukouvinos, A., Mamassis, N., Koutsoyiannis, D., and Efstratiadis, A. (2016). Comparative evaluation of 1D and quasi-2D hydraulic models based on benchmark and real-world applications for uncertainty assessment in flood mapping. *Journal of Hydrology*, 534:478–492.

Dineva, A., Várkonyi-Kóczy, A. R., and Tar, J. K. (2014). Fuzzy expert system for automatic wavelet shrinkage procedure selection for noise suppression. *INES 2014 - IEEE 18th International Conference on Intelligent Engineering Systems, Proceedings*, pages 163–168.

Disse, M., Konnerth, I., Bhola, P. K., and Leandro, J. (2018). Unsicherheitsabschätzung für die Berechnung von dynamischen Überschwemmungskarten – Fallstudie Kulmbach. In Heimerl, S., editor, *Vorsorgender und nachsorgender Hochwasserschutz: Ausgewählte Beiträge aus der Fachzeitschrift WasserWirtschaft Band 2*, pages 350–357. Springer Fachmedien Wiesbaden, Wiesbaden.

Djordjević, S., Prodanović, D., and Maksimović, Č. (1999). An approach to simulation of dual drainage. *Water Science and Technology*, 39(9):95–103.

Djordjević, S., Prodanović, D., Maksimović, Č., Ivetić, M., and Savić, D. (2005). SIPSON - Simulation of interaction between pipe flow and surface overland flow in networks. *Water Science and Technology*, 52(5):275–283.

Elsafi, S. H. (2014). Artificial Neural Networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan. *Alexandria Engineering Journal*, 53(3):655–662.

Engeland, K., Xu, C. Y., and Gottschalk, L. (2005). Estimation bayésienne des incertitudes au sein d'une modé lisation conceptuelle de bilan hydrologique. *Hydrological Sciences Journal*, 50(1):45–63.

Farmer, D., Sivapalan, M., and Jothityangkoon, C. (2003). Climate, soil, and vegetation controls upon the variability of water balance in temperate and semiarid landscapes: Downward approach to water balance analysis. *Water Resources Research*, 39(2):1–21.

Freer, J., Beven, K., and Ambroise, B. (1996). Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resources Research*, 32(7):2161–2173.

Fritsch, K., Assmann, A., and Tyrna, B. (2016). Long-term experiences with pluvial flood risk management. *E3S Web of Conferences*, 7:04017.

Gallegos, H. A., Schubert, J. E., and Sanders, B. F. (2009). Two-dimensional, high-resolution modeling of urban dam-break flooding: A case study of Baldwin Hills, California. *Advances in Water Resources*, 32(8):1323–1335.

Gharbi, M., Soualmia, A., Dartus, D., and Masbernat, L. (2016). Comparison of 1D and 2D Hydraulic Models for Floods Simulation on the. *Journal of Materials and Environmental Science*, 7(8):3017–3026.

Gironás, J., Niemann, J. D., Roesner, L. A., Rodriguez, F., and Andrieu, H. (2010). Evaluation of methods for representing Urban Terrain in storm-water modeling. *Journal of Hydrologic Engineering*, 15(1):1–14.

Gizaw, M. S. and Gan, T. Y. (2016). Regional Flood Frequency Analysis using Support Vector Regression under historical and future climate. *Journal of Hydrology*, 538:387–398.

Goodarzi, E., Mirzaei, M., Shui, L. T., and Ziaei, M. (2011). Evaluation dam overtopping risk based on univariate and bivariate flood frequency analysis. *Hydrology and Earth System Sciences Discussions*, 8(6):9757–9796.

Hall, J. W., Manning, L. J., and Hankin, R. K. (2011). Bayesian calibration of a flood inundation model using spatial data. *Water Resources Research*, 47(5):1–14.

Hall, J. W., Tarantola, S., Bates, P. D., and Horritt, M. S. (2005). Distributed Sensitivity Analysis of Flood Inundation Model Calibration. *Journal of Hydraulic Engineering*, 131(2):117–126.

Hankin, B., Waller, S., Astle, G., and Kellagher, R. (2008). Mapping space for water: screening for urban flash flooding. *Journal of Flood Risk Management*, 1(1):13–22.

Henonin, J., Russo, B., Mark, O., and Gourbesville, P. (2013). Real-time urban flood forecasting and modelling - A state of the art. *Journal of Hydroinformatics*, 15(3):717–736.

Hornberger, G. M. and Spear, R. C. (1981). An approach to the preliminary analysis of environmental systems. *Journal of Environmental Management*, 12(1):7–18.

Hübl, J., Heiser, M., Braito, S., Tscharner, S., Kuntner, K., Schraml, K., Falkensteiner, M., and Rabanser, E. (2017). Ereignisdokumentation und Ereignisanalyse Rottal-Inn 2016, Band 1: Ereignisdokumentation. Technical report, Universität für Bodenkultur Wien Department Bautechnik und Naturgefahren Institut für Alpine Naturgefahren (IAN), Vienna, Austria.

Humphrey, G. B., Gibbs, M. S., Dandy, G. C., and Maier, H. R. (2016). A hybrid approach to monthly streamflow forecasting: Integrating hydrological model outputs into a Bayesian artificial neural network. *Journal of Hydrology*, 540(June):623–640.

Hunter, N. M., Bates, P. D., Neelz, S., Pender, G., Villanueva, I., Wright, N. G., Liang, D., Falconer, R. A., Lin, B., Waller, S., Crossley, A. J., and Mason, D. C. (2008). Benchmarking 2D hydraulic models for urban flooding. *Proceedings of the Institution of Civil Engineers: Water Management*, 161(1):13–30.

Huong, H. T. L. and Pathirana, A. (2013). Urbanization and climate change impacts on future urban flooding in Can Tho city, Vietnam. *Hydrology and Earth System Sciences*, 17(1):379–394.

Jhong, Y. D., Chen, C. S., Lin, H. P., and Chen, S. T. (2018). Physical hybrid neural network model to forecast typhoon floods. *Water (Switzerland)*, 10(5):1–17.

Jiang, L., Chen, Y., and Wang, H. (2015). Urban flood simulation based on the SWMM model. *IAHS-AISH Proceedings and Reports*, 368(August 2014):186–191.

Kabir, S., Patidar, S., Xia, X., Liang, Q., Neal, J., and Pender, G. (2020). A deep convolutional neural network model for rapid prediction of fluvial flood inundation. *Journal of Hydrology*, 590(September):125481.

Kalyanapu, A. J., Shankar, S., Pardyjak, E. R., Judi, D. R., and Burian, S. J. (2011). Assessment of GPU computational enhancement to a 2D flood model. *Environmental Modelling and Software*, 26(8):1009–1016.

Kasiviswanathan, K. S., He, J., Sudheer, K. P., and Tay, J. H. (2016). Potential application of wavelet neural network ensemble to forecast streamflow for flood management. *Journal of Hydrology*, 536:161–173.

Kavetski, D., Binning, P., and Sloan, S. W. (2002). Adaptive backward Euler time stepping with truncation error control for numerical modelling of unsaturated fluid flow. *International Journal for Numerical Methods in Engineering*, 53(6):1301–1322.

Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In *IJCAI*.

Kron, W. (2005). Flood risk = hazard • values • vulnerability. *Water International*, 30(1):58–68.

Lamb, R., Beven, K., and Myrabø, S. (1998). Use of spatially distributed water table observations to constrain uncertainty in a rainfall-runoff model. *Advances in Water Resources*, 22(4):305–317.

Leandro, J., Chen, A. S., Djordjević, S., and 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):495–504.

Leandro, J., Chen, A. S., and Schumann, A. (2014). A 2D parallel diffusive wave model for floodplain inundation with variable time step (P-DWave). *Journal of Hydrology*, 517:250–259.

Leandro, J., Schumann, A., and Pfister, A. (2016). A step towards considering the spatial heterogeneity of urban key features in urban hydrology flood modelling. *Journal of Hydrology*, 535:356–365.

LfStat (2019). Statistik kommunal 2018 Stadt Simbach a.Inn 09 277 145. Technical report, Bayerisches Landesamt für Statistik, Nuremberg, Germany.

LfU (2020). Runoff: Master data Simbach / Simbach.

Li, L., Hong, Y., Wang, J., Adler, R. F., Policelli, F. S., Habib, S., Irwn, D., Korme, T., and Okello, L. (2009). Evaluation of the real-time TRMM-based multi-satellite precipitation analysis for an operational flood prediction system in Nzoia Basin, Lake Victoria, Africa. *Natural Hazards*, 50(1):109–123.

Liu, L., Liu, Y., Wang, X., Yu, D., Liu, K., Huang, H., and Hu, G. (2015). Developing an effective 2-D urban flood inundation model for city emergency management based on cellular automata. *Natural Hazards and Earth System Sciences*, 15(3):381–391.

Liu, Y. and Pender, G. (2013). Carlisle 2005 urban flood event simulation using cellular automata-based rapid flood spreading model. *Soft Computing*, 17(1):29–37.

Ludwig, K. and Bremicker, M. (2006). *The Water Balance Model LARSIM: Design, Content and Applications*. Freiburger Schriften zur Hydrologie. Inst. für Hydrologie der Univ.

Maksimovi, Č., Prodanovi, D., and Djordjevi, S. (2007). Urban Pluvial Flooding : Development of GIS Based Pathway Model for Surface Flooding and Interface with Surcharged Sewer Model. *Novatech*, pages 481–488.

Maksimović, Č., Prodanović, D., Boonya-Aroonnet, S., Leitão, J. P., Djordjević, S., and Allitt, R. (2009). Overland flow and pathway analysis for modelling of urban pluvial flooding. *Journal of Hydraulic Research*, 47(4):512–523.

Maniak, U. (2016). Hydrologie und Wasserwirtschaft. *Hydrologie und Wasserwirtschaft*.

Mark, O., Weesakul, S., Apirumanekul, C., Aroonnet, S. B., and Djordjević, S. (2004). Potential and limitations of 1D modelling of urban flooding. *Journal of Hydrology*, 299(3-4):284–299.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. INDUST. APPL. MATH.*, 11(2):431–441.

McMichael, C. E., Hope, A. S., and Loaiciga, H. A. (2006). Distributed hydrological modelling in California semi-arid shrublands: MIKE SHE model calibration and uncertainty estimation. *Journal of Hydrology*, 317(3-4):307–324.

Merz, B., Thieken, A., and Gocht, M. (2007). Flood Risk Mapping At The Local Scale: Concepts and Challenges. In *Risk Management*, volume 23, pages 231–251. Springer.

Merz, B. and Thieken, A. H. (2005). Separating natural and epistemic uncertainty in flood frequency analysis. *Journal of Hydrology*, 309(1-4):114–132.

Mignot, E., Li, X., and Dewals, B. (2019). Experimental modelling of urban flooding: A review. *Journal of Hydrology*, 568:334–342.

Mirzaei, M., Huang, Y. F., El-Shafie, A., and Shatirah, A. (2015). Application of the generalized likelihood uncertainty estimation (GLUE) approach for assessing uncertainty in hydrological models: a review. *Stochastic Environmental Research and Risk Assessment*, 29(5):1265–1273.

Mosavi, A., Ozturk, P., and Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. *Water (Switzerland)*, 10(11):1–40.

Mukerji, A., Chatterjee, C., and Raghuwanshi, N. S. (2009). Flood Forecasting Using ANN, Neuro-Fuzzy, and Neuro-GA Models. *Journal of Hydrologic Engineering*, 14(6):647–652.

Muleta, M. K. and Nicklow, J. W. (2005). Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. *Journal of Hydrology*, 306(1-4):127–145.

Nasello, C. and Tucciarelli, T. (2005). Dual Multilevel Urban Drainage Model. *Journal of Hydraulic Engineering*, 131(9):748–754.

Nash, J. and Sutcliffe, J. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, 10(3):282–290.

Nawi, N. M., Ransing, R. S., and Ransing, M. R. (2007). An improved Conjugate Gradient based learning algorithm for back propagation neural networks. *International Journal of Computational Intelligence*, 4(1):46–55.

Nkwunonwo, U. C., Whitworth, M., and Baily, B. (2020). A review of the current status of flood modelling for urban flood risk management in the developing countries. *Scientific African*, 7:e00269.

Obropta, C., Kallin, P., Mak, M., and Ravit, B. (2008). Modeling Urban Wetland Hydrology for the Restoration of a Forested Riparian Wetland Ecosystem. *Urban Habitats*, 5(1):183–198.

Panda, R. K., Pramanik, N., and Bala, B. (2010). Simulation of river stage using artificial neural network and MIKE 11 hydrodynamic model. *Computers and Geosciences*, 36(6):735–745.

Papaioannou, G., Vasiliades, L., Loukas, A., and Aronica, G. T. (2017). Probabilistic flood inundation mapping at ungauged streams due to roughness coefficient uncertainty in hydraulic modelling. *Advances in Geosciences*, 44:23–34.

Pappenberger, F., Beven, K. J., Ratto, M., and Matgen, P. (2008). Multi-method global sensitivity analysis of flood inundation models. *Advances in Water Resources*, 31(1):1–14.

Paquier, A., Mignot, E., and Bazin, P. H. (2015). From hydraulic modelling to urban flood risk. *Procedia Engineering*, 115(i):37–44.

Patel, D. P., Ramirez, J. A., Srivastava, P. K., Bray, M., and Han, D. (2017). Assessment of flood inundation mapping of Surat city by coupled 1D/2D hydrodynamic modeling: a case application of the new HEC-RAS 5. *Natural Hazards*, 89(1):93–130.

Quirogaa, V. M., Kurea, S., Udoa, K., and Manoa, A. (2016). Application of 2D numerical simulation for the analysis of the February 2014 Bolivian Amazonia flood: Application of the new HEC-RAS version 5. *Ribagua*, 3(1):25–33.

Rawls, W. J., Stricker, V., and Wilson, K. (1980). Review and evaluation of urban flood flow frequency procedures.

Romanowicz, R. and Beven, K. (2003). Estimation of flood inundation probabilities as conditioned on event inundation maps. *Water Resources Research*, 39(3):1–12.

Rossmann, L. A. (2017). *Storm Water Management Model Reference Manual Volume II - Hydraulics*, volume 2. U.S. Environmental Protection Agency, Washington, DC.

Rossmann, L. A. and Huber, W. C. (2016). *Storm Water Management Model Reference Manual Volume I – Hydrology (revised)(EPA/600/R-15/162A)*, volume 1. U.S. Environmental Protection Agency, Cincinnati, jan. 2016 edition.

Saini, L. M. (2008). Peak load forecasting using Bayesian regularization, Resilient and adaptive backpropagation learning based artificial neural networks. *Electric Power Systems Research*, 78(7):1302–1310.

Sankaranarayanan, S., Prabhakar, M., Satish, S., Jain, P., Ramprasad, A., and Krishnan, A. (2019). Flood prediction based on weather parameters using deep learning. *Journal of Water and Climate Change*, pages 1–18.

Savage, J. T. S., Bates, P., Freer, J., Neal, J., and Aronica, G. (2016). When does spatial resolution become spurious in probabilistic flood inundation predictions? *Hydrological Processes*, 30(13):2014–2032.

Schaefli, B., Talamba, D. B., and Musy, A. (2007). Quantifying hydrological modeling errors through a mixture of normal distributions. *Journal of Hydrology*, 332(3-4):303–315.

Schmitt, T. G., Thomas, M., and Ettrich, N. (2004). Analysis and modeling of flooding in urban drainage systems. *Journal of Hydrology*, 299(3-4):300–311.

Schubert, J. E. and Sanders, B. F. (2012). Building treatments for urban flood inundation models and implications for predictive skill and modeling efficiency. *Advances in Water Resources*, 41:49–64.

Sit, M. and Demir, I. (2019). Decentralized Flood Forecasting Using Deep Neural Networks.

Sit, M., Demiray, B. Z., Xiang, Z., Ewing, G. J., Sermet, Y., and Demir, I. (2020). A comprehensive review of deep learning applications in hydrology and water resources. *Water Science and Technology*.

Stephens, E. M., Bates, P. D., Freer, J. E., and Mason, D. C. (2012). The impact of uncertainty in satellite data on the assessment of flood inundation models. *Journal of Hydrology*, 414-415:162–173.

Tadesse, A. and Anagnostou, E. N. (2005). A statistical approach to ground radar-rainfall estimation. *Journal of Atmospheric and Oceanic Technology*, 22(11):1720–1732.

Taghi Sattari, M., Yurekli, K., and Pal, M. (2012). Performance evaluation of artificial neural network approaches in forecasting reservoir inflow. *Applied Mathematical Modelling*, 36(6):2649–2657.

Taherei Ghazvinei, P., Darvishi, H. H., Mosavi, A., Bin Wan Yusof, K., Alizamir, M., Shamshirband, S., and Chau, K. W. (2018). Sugarcane growth prediction based on meteorological parameters using extreme learning machine and artificial neural network. *Engineering Applications of Computational Fluid Mechanics*, 12(1):738–749.

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research Atmospheres*, 106(D7):7183–7192.

Thirumalaiah and Deo (1998). RIVER STAGE FORECASTING USING ARTIFICIAL NEURAL NETWORKS By Konda Thirumalaiah 1 and M. C. Deo 2. *Journal of Hydrologic Engineering*, 3(1):26–32.

Tilson, L. V., Excell, P. S., and Green, R. J. (1988). A generalisation of the Fuzzy c-Means clustering algorithm. *Remote sensing. Proc. IGARSS '88 symposium, Edinburgh, 1988. Vol. 3*, 10(2):1783–1784.

Tiwari, M. K. and Chatterjee, C. (2010). Development of an accurate and reliable hourly flood forecasting model using wavelet-bootstrap-ANN (WBANN) hybrid approach. *Journal of Hydrology*, 394(3-4):458–470.

Tsubaki, R. and Kawahara, Y. (2013). The uncertainty of local flow parameters during inundation flow over complex topographies with elevation errors. *Journal of Hydrology*, 486:71–87.

Vogel, R. M., Yaindl, C., and Walter, M. (2011). Nonstationarity: Flood magnification and recurrence reduction factors in the united states. *Journal of the American Water Resources Association*, 47(3):464–474.

Wang, H., Rahnamayan, S., and Wu, Z. (2013). Parallel differential evolution with self-adapting control parameters and generalized opposition-based learning for solving high-dimensional optimization problems. *Journal of Parallel and Distributed Computing*, 73(1):62–73.

Wasserwirtschaftamt Deggendorf (2012). Statistik Simbach / Simbach.

Wasserwirtschaftamt Hof (2012). Statistik Kauerndorf / Schorgast.

Winterrath, T., Brendel, C., Hafer, M., Junghänel, T., Klameth, A., Lengfeld, K., Walawender, E., Weigl, E., and Becker, A. (2018). RADKLIM Version 2017.002: Reprocessed quasi gauge-adjusted radar data, 5-minute precipitation sums (YW).

Xu, S. and Chen, L. (2008). A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining. In *5th International Conference on Information Technology and Applications, ICITA 2008*, number Icita, pages 683–686.

Yang, J., Reichert, P., Abbaspour, K. C., Xia, J., and Yang, H. (2008). Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *Journal of Hydrology*, 358(1-2):1–23.

Yang, J., Reichert, P., Abbaspour, K. C., and Yang, H. (2007). Hydrological modelling of the Chaohe Basin in China: Statistical model formulation and Bayesian inference. *Journal of Hydrology*, 340(3-4):167–182.

Yang, T. H., Chen, Y. C., Chang, Y. C., Yang, S. C., and Ho, J. Y. (2015). Comparison of different grid cell ordering approaches in a simplified inundation model. *Water (Switzerland)*, 7(2):438–454.

Yu, P. S., Chen, S. T., and Chang, I. F. (2006). Support vector regression for real-time flood stage forecasting. *Journal of Hydrology*, 328(3-4):704–716.

Zhang, G. P. (2007). Avoiding pitfalls in neural network research. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 37(1):3–16.