








Generation of harmonised pluvial flood hazard maps through decentralised analytics

Sascha Welten ^{a,*}, Adrian Holt ^{a,b}, Julian Hofmann ^b, Sven Weber ^a, Elena-Maria Klopries ^b,
Holger Schüttrumpf ^b and Stefan Decker ^a

^a Computer Science 5, RWTH Aachen University, Ahornstr. 55, Aachen 52074, Germany

^b Institute of Hydraulic Engineering and Water Resources Management, RWTH Aachen University, Mies-van-der-Rohe-Str. 17, Aachen 52074, Germany

*Corresponding author. E-mail: welten@dbis.rwth-aachen.de

 SW, 0000-0001-5570-9672; AH, 0000-0001-9401-1488; JH, 0000-0002-7156-7173; SW, 0000-0002-8518-9097; E-MK, 0000-0003-4752-6619;
HS, 0000-0002-0104-0499; SD, 0000-0001-6324-7164

ABSTRACT

Increasing extreme weather events pose significant challenges in hydrology, requiring tools for preparedness and prediction of intense rainfall impacts, especially flash floods. Current risk reduction measures for pluvial flood risk management rely on flood hazard maps, but inconsistencies in transregional standards that are used for risk assessment hinder cross-regional comparisons. While there are existing guidelines for the development of pluvial flood hazard maps, there is still a lack of holistic modelling systems that enable harmonised predictions of the impacts of heavy rainfall events. Furthermore, sensitive city data (e.g., critical infrastructure, sewer network) exist in many municipalities, which cannot be readily disclosed for modelling purposes. In this work, we propose an approach using distributed analytics to distribute computation commands to existing hydrodynamic models at different locations. In combination with harmonising model adapters, we enable the generation of harmonised pluvial flood hazard maps of different regions to tackle the inconsistencies and privacy concerns. We apply our approach to four adjacent urban areas in the Rhein-Sieg Kreis of North Rhine-Westphalia. Our results demonstrate the ability of our approach to produce cross-regional pluvial flood hazard maps, supporting disaster preparedness and management in regions prone to extreme weather events and flash floods.

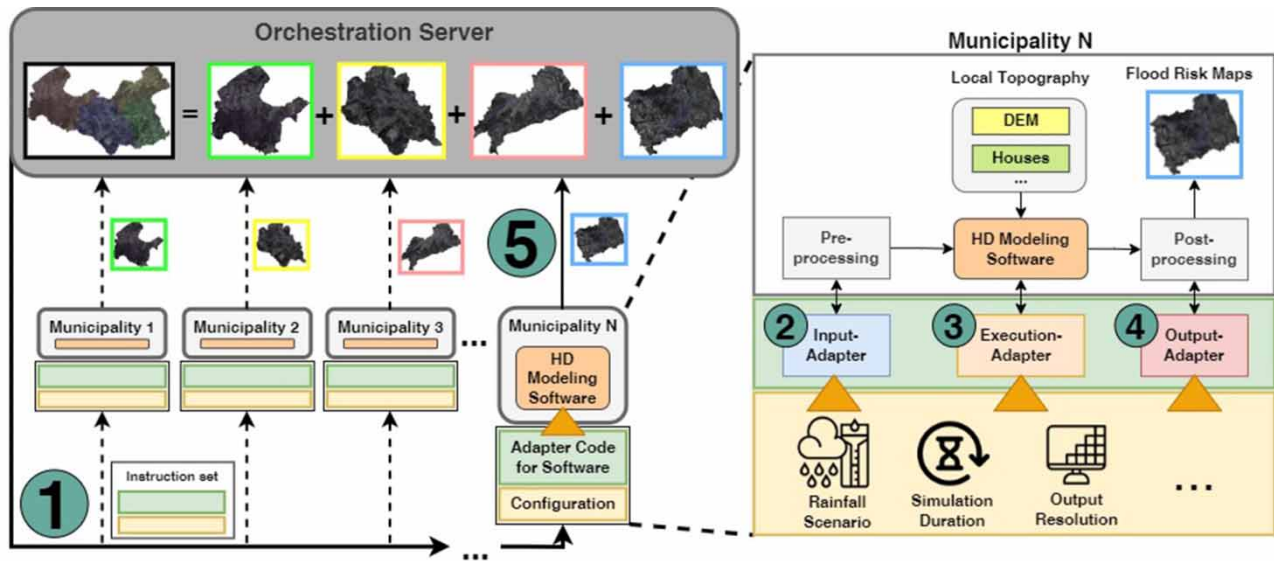
Key words: data harmonisation, disaster preparedness, distributed analytics, hydrodynamic modelling, hydroinformatics, pluvial flood hazard mapping

HIGHLIGHTS

- Decentralised flood modelling: Innovative approach using distributed analytics for harmonised pluvial flood hazard maps, addressing inconsistencies.
- Improved risk assessment: Enables standardised cross-regional comparisons, enhancing disaster preparedness.
- Real-world validation: Practical application in North Rhine-Westphalia validates the approach's feasibility and relevance for complex urban areas prone to flooding.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



INTRODUCTION

With the increasing number of extreme weather events, a significant current challenge within the field of hydrology revolves around preparing for, predicting, and addressing the impacts of intense rainfall occurrences (Westra *et al.* 2014; Hofmann & Schüttrumpf 2020). These extreme rainfall events often lead to flash floods, which pose a severe risk to human lives and infrastructure (Westra *et al.* 2014; Elmahdy *et al.* 2020). Recent extreme weather events, such as those witnessed in Germany, Pakistan, Greece, Bulgaria, Turkey, Libya, and the United States, underscore the need for tools for disaster preparedness (Davies 2021, 2022, 2023a, 2023b, 2023c). These tools are essential for protecting human lives and reducing the financial impacts of such events. The current methodologies for flood risk assessment predominantly incorporate pluvial flood hazard maps, which highlight locations likely to be flooded during heavy rainfall events (Hofmann & Schüttrumpf 2021; Mudashiru *et al.* 2021).

These maps are oftentimes created through two-dimensional (2D) hydrodynamic (HD) simulations, which incorporate the topography of a given area, specific rainfall scenarios (as well as other types of input data), and provide an assessment of flood-prone areas (Hofmann & Schüttrumpf 2021). Pluvial flood hazard maps are often created individually for specific urban areas, such as municipalities, often in isolation from one another (Rheinisch-Bergischer Kreis 2023; Stadt Bochum 2023; Stadt Dortmund 2023). Consequently, existing approaches used to create pluvial flood hazard maps suffer from inconsistencies in standards when crossing municipal boundaries (Rheinisch-Bergischer Kreis 2023; Stadt Bochum 2023; Stadt Dortmund 2023). These challenges are especially evident at the national level in Germany. German municipalities, influenced by factors such as historical practices, budgetary constraints, and local expertise, use a diverse range of software and data (Pyka 2020). This variety leads to inconsistencies, e.g., in data resolutions and approaches, further complicated by regional topographical and hydrological characteristics. This diverse landscape in standards, practices, or software poses significant challenges and complexities that hinder the development of uniform strategies for flood risk management or the comparisons of pluvial flood hazards across diverse geographical regions in Germany. We classify the present challenges into three categories:

- Variability in configuration and formats:** One challenge encompasses the lack of standardisation in pluvial flood hazard modelling parameters, including factors such as duration, intensity, modelling resolution, and output data format. This can make it difficult to compare and integrate data from different sources or regions, leading to potential incompatibilities in flood risk assessments (see Table 1).
- Boundaries:** Usually, bodies of water flow across several municipalities and national borders. In the traditional approach, municipalities create pluvial flood hazard maps only for their responsible area. This introduces a sharp cutoff at the boundaries of the map. The same disadvantage is present at the federal or national level.

Table 1 | The variability in pluvial flood hazard assessment parameters across several German municipalities

Location/municipality	Provider/software	SRI level & corresponding rainfall data	Modelling resolution	Coupled model (using sewer network data)
Kassel ^a	Geomer (FloodArea HPC)	SRI 5 (35 mm); SRI 7 (45 mm); SRI 10 (90 mm)	–	–
Cologne ^b	–	SRI 5 (43.48 mm); SRI 6 (47.52 mm); SRI 7 (53 mm); SRI 10 (58.48 mm)	1 m	No
Meckenheim ^c	Hydrotec (HydroAS)	SRI 7 (42 mm)	–	–
Rheinisch-Bergischen Kreis ^d	Hydrotec (HydroAS)	SRI 7 (55 mm)	1 m	No
Eppelborn ^e	Hydrotec (HydroAS)	SRI 6; SRI 7	–	–
Hamburg ^f	–	SRI 5 (29 mm); SRI 7 (36 mm); SRI 12 (100 mm)	–	Yes
Wuppertal ^g	Dr Pecher AG	SRI 6 (38.5 mm); SRI 7 (42 mm); SRI 10 (90 mm); SRI 12 (real event)	1 m	Simplified
Leipzig ^h	–	SRI 5 (43.5 mm); SRI 7 (53.6 mm); extreme (80 mm)	2 m	No

The comparison includes service providers, the statistical return periods (given by the SRI value), modelling resolution, and the incorporation of sewer network data in the models. This selection represents only a subset of the configurable parameters available to flood modelling. It is important to mention that certain details have not been made public (–).

^a<https://www.kasselwasser.de/abwasser/starkregen/starkregengefahrenkarte>

^b<https://www.hw-karten.de/index.html?Module=Starkregen>

^c<https://www.meckenheim.de/cms117/aktuelles/pressemitteilungen/artikel/78958/>

^d<https://www.hydrotec.de/starkregengefahrenkarten-rheinisch-bergischer-kreis/>

^e<https://www.eppelborn.de/starkregengefahrenkarten-fuer-die-gemeinde-eppelborn-simulationsmodelle-fa-hydrotec/>

^f<https://www.hamburg.de/starkregengefahrenkarte/>

^g<https://www.wuppertal.de/microsite/geoportal/topicmaps/contentseiten/starkregengefahrenkarte.php>

^h<https://www.leipzig.de/bauen-und-wohnen/bauen/starkregen#c193894>

3. Limited diversity in rainfall scenarios: Alongside the first challenge, it is worth noting that often, only limited sets of rainfall scenarios with minimal diversity as input are taken into account. Furthermore, these do not necessarily reflect realistic rainfall events (e.g., due to the assumption of spatially uniform rainfall over the whole study area), possibly leading to inaccuracies when assessing flood risks (Essenfelder *et al.* 2022). Hence, the amount of rainfall data and the scenarios modelled can vary significantly between regions (see Table 1).

It becomes evident that in the realm of flood risk assessment, heterogeneity is a common and persistent issue, especially in Germany, that must be faced to apply flood risk assessment on a larger scale. These challenges could be addressed by harmonising the parameters (challenges 1 and 2) and diversifying the range of rainfall scenarios (challenge 3) while leveraging the existing HD models to produce high-quality simulations of these standardised scenarios. The HD models created for simulating the risk of flooding for any specific urban area (e.g., municipalities) generally exist at differing locations. This is often because they are developed by different contractors or water authorities. Hence, in order to apply derived standards or combine their outputs, it becomes essential to either centralise their computation or execute it in a decentralised manner. Regarding the first, nevertheless, these HD models might be tightly integrated into a modelling environment or software platforms, each with its own unique data formats that may lead to challenges in centralising them for the purposes of harmonisation (Xia *et al.* 2019; DHI 2023). Moreover, in the water and wastewater systems sector, there is susceptibility to various types of attacks, especially when these models have local connections to sensitive data, such as sewer networks. This susceptibility poses a privacy-related challenge when attempting to centralise computation and sensitive data leaves institutional borders (Leandro *et al.* 2009; CISA 2023). As for the latter, a suitable approach is required that is capable of distributing computation commands to the corresponding locations, decentrally computing the pluvial flood hazard maps, and finally gathering and combining the results. Methods that can conduct such a decentralised approach include distributed analytics (DA), which represents an alternative to centralised solutions.

This work introduces a novel technological framework that directly addresses key challenges in hydrological modelling and flood risk assessment, specifically the generation of harmonised pluvial flood hazard maps. Our objective is to use a combination of a DA framework and model adapters to tackle the above-mentioned multi-faceted heterogeneity in the domain of

hydrology. In this setup, the HD models remain at their corresponding origins and utilise the institutions' local data, such as digital elevation models (DEMs) or sewer network data, without the need for external access to these data. The model adapters consist of the harmonised rainfall data input, along with instructions for model execution (including simulation parameters), and are distributed to the institutions via the DA framework. This approach ensures that every distributed HD model receives identical input (e.g., same rainfall scenario) and generates pluvial flood hazard maps that are compatible and can be compared. Once the model has been executed, the harmonised pluvial flood hazard maps are stored at the respective institutions and can be used for conducting local flood risk assessments. Alternatively, in a final step, the DA framework can initiate a transfer of the flood maps back to a central server, where they are aggregated into a single georeferenced flooding map, enabling interregional flood risk assessment. Through this DA-driven approach, we aim to seamlessly generate synchronised flood maps for diverse urban areas without centralising existing HD models. We finally evaluate our approach through a feasibility study encompassing four urban areas within the Rhein-Sieg Kreis region of North Rhine-Westphalia (NRW), Germany.

The remainder of this work is structured as follows. *Methods* presents our used materials, our DA-inspired methodology, and our proof of concept implementation. *Results and Discussion* shows our case study and its results. Finally, *Conclusion* concludes this work and gives an outlook for future work.

METHODS

In this study, our methodology can be divided into two main components. Initially, we provide background and materials on HD modelling (see *Hydrodynamic Modelling*), its applications, and current limitations with possible solutions. Then, we explore DA (see *Distributed Analytics*), as a means to address the above-mentioned challenges. Subsequently, we present our concept (see *Distributed Analytics applied to Hydrodynamic Modelling*) with the goal to decentrally perform flood risk analysis with HD modelling through DA.

Hydrodynamic modelling

Simulations for flood risk analysis are usually based on HD models, of which various types exist. One-dimensional (1D) simulations are, for example, used to model the processes in drainage networks (Leandro *et al.* 2009; Hofmann & Schüttrumpf 2019). However, when the capacity of the drainage networks is exceeded, or no drainage networks are present at all, they fail to model the water flow above the surface (Hofmann & Schüttrumpf 2019). 2D simulations are able to model the surface flow by taking into account the topography in addition to the specific rainfall scenario, returning 2D rasters of flooding values, water speed, and other parameters (Hofmann & Schüttrumpf 2019). Coupled with 1D simulations, the interaction with drainage networks can be modelled (Hofmann & Schüttrumpf 2019). At high modelling resolutions, 2D HD simulations provide an effective means to analyse flooding risk in urban areas (Hofmann & Schüttrumpf 2019).

The most fundamental part of an HD model is the topographical data in the form of DEMs, on which the fluid dynamics are simulated (Xia *et al.* 2019). DEMs are representations of the terrain's surface and can be represented as 2D raster data, where each raster point corresponds to a specific elevation above sea level (Xia *et al.* 2019). This topographical data, in combination with rainfall data and other input data such as land use data, can be used by HD modelling software to create pluvial flood hazard maps. Multiple software solutions for running 2D HD simulations exist, including MIKE +,¹ Infoworks-ICM,² and HydroAS.³ In addition to these commercial solutions, several open-source solutions exist, including HiPIMS,⁴ LIS-FLOOD-FP,⁵ and ANUGA.⁶ Although these solutions are primarily employed for a common purpose, they rely on different data types and formats. For example, the software HiPIMS requires data to be in a Georeferenced Tagged Image File Format (GeoTIFF) or American Standard Code for Information Interchange (ASCII) grid format, while MIKE+ uses the custom DFS2 or mesh-based formats (Xia *et al.* 2019; DHI 2023). HD models are the de-facto standard for creating detailed and accurate pluvial flood hazard maps and have been applied, for example, in multiple regions of Germany such

¹ MIKE + : <https://www.mikepoweredbydhi.com/>

² Infoworks-ICM: <https://www.innoaqua.de/software/infoworks-icm/>

³ HydroAS: <https://www.hydrotec.de/software/hydroas/>

⁴ HiPIMS: <https://github.com/HEMLab/hipims>

⁵ LISFLOOD-FP: <http://www.bristol.ac.uk/geography/research/hydrology/models/lisflood/>

⁶ ANUGA: <https://anuga.anu.edu.au/>

as Cologne,⁷ Bochum,⁸ Münster,⁹ and numerous other regions worldwide (Teng *et al.* 2017). Generally, pluvial flood hazard maps created for any specific area consider rainfall events of uniform intensity levels across the entire simulation time and spatial domain. Oftentimes, multiple simulations are executed, considering rainfall scenarios of differing return periods (e.g., SRI 7, a scenario that occurs statistically once every 100 years) (HW Karten 2023; Rheinisch-Bergischer Kreis 2023; Stadt Bochum 2023). However, the rainfall intensities usually vary, which might pose a challenge when the goal is to create uniform disaster response strategies. An overview – a selection – of the variations in the simulations across multiple regions is given in Table 1.

Recently, strides towards standardised pluvial flood hazard maps have been undertaken, e.g., through the creation of the Starkregenhinweiskarte NRW,¹⁰ covering the entire state of NRW with a uniform rainfall pattern. However, applying rainfall uniformly (the same rainfall value) over an entire study area is not realistic, as historical events show vastly different characteristics (e.g., intense rainfall only occurring over part of the study area) (Deutscher Wetterdienst 2023). Furthermore, the models that are used to create these cross-regional standardised maps are often less accurate, partly because they do not account for municipality-specific data such as sewer networks (Hofmann & Schüttrumpf 2019). For many urban areas, HD models have already been set up for flood risk assessment in the past. Consequently, these existing models can be utilised in cross-regional flood risk assessments, where standardised inputs incorporating a broad range of realistic rainfall scenarios are used. As a result, there is a need for a standardisation methodology. One potential way to achieve this is to centralise the existing models, which are currently hosted on servers belonging to various municipal water authorities, to a central location and apply the standardisation strategy there. However, this centralisation approach can pose various challenges (CISA 2023). Some HD models use input data that can be confidential, e.g., specialised high-resolution DEM models or sewer network data, which, if publicised, could compromise public safety (Leandro *et al.* 2009; CISA 2023). Furthermore, the logistics of extracting data from local databases might take time and effort. Given these challenges, a more feasible approach would be to leverage remote computing, allowing the HD models and their used data to remain in their origin (Welten *et al.* 2021). To accomplish this, one could employ DA by sending commands and instructions, including details about the desired rainfall scenario for simulation, directly to the model, rather than relocating the model to a central location (Welten *et al.* 2021).

Distributed analytics

DA enables analysis and remote computation on decentralised data by bringing the algorithmic code (e.g., machine learning (ML) model training or other analysis code) to the data when the reverse approach is impractical or infeasible (Welten *et al.* 2021). Reasons for infeasibility lie in the ethical, technical, privacy-related, or legal challenges of centralising data to a single location. One example of such a scenario is situations when dealing with sensitive patient data in the medical domain (Chang *et al.* 2018; Sheller *et al.* 2020; Welten *et al.* 2021). The impracticability of data centralisation can also occur due to possibly vast amounts of data at any participating institution leading to bottlenecks in the process (Chang *et al.* 2018). DA addresses these concerns by executing the algorithmic task where the data are located and ensures that privacy-sensitive assets remain within institutional boundaries (Welten *et al.* 2021). Therefore, DA has been beneficial for the application of, e.g., ML in the medical domain but has also been applied to other domains such as the field of hydrology (Chang *et al.* 2018; Sheller *et al.* 2020; Welten *et al.* 2022b, 2022c). Beyond ML, DA can be adapted to various other purposes such as specific tasks with pre-determined instruction sets at the data location. Multiple paradigms for executing DA workflows have been proposed, with the most relevant being Institutional Incremental Learning (IIL) and Federated Learning (FL).

The FL paradigm allows for parallel execution of algorithms at participating institutions with a central component responsible for aggregating the results. This approach was first introduced by McMahan *et al.* as a means to train ML models on data from mobile devices decentrally (McMahan *et al.* 2017). A standard FL workflow begins with a central server sending a copy of the algorithm to each participating institution (Welten *et al.* 2021). Once an institution has received the algorithmic task, it performs the task on its data, after which only the results of the computation are sent back to the central server for aggregation (Welten *et al.* 2021). After the central server receives the institutions' individual results of the computation, it performs an

⁷ Cologne: <https://www.hw-karten.de/index.html?Module=Starkregen>

⁸ Bochum: <https://www.bochum.de/Starkregengefahrenkarte>

⁹ Münster: <https://www.stadt-muenster.de/wasser/starkregengefahrenkarten>

¹⁰ Starkregenhinweiskarte NRW: https://geoportal.de/Info/tk_04-starkregengefahrenhinweise-nrw

aggregation step (Sheller *et al.* 2020; Welten *et al.* 2021). After the central server completes the aggregation, a single federated round is completed and may potentially be repeated. FL has been applied to various other domains such as healthcare, hydrology, and industrial engineering (Kim *et al.* 2017; Silva *et al.* 2019; Chen *et al.* 2020; Li *et al.* 2020; Sheller *et al.* 2020; Welten *et al.* 2022b). Moreover, Welten *et al.* specifically applied the FL method to train synthetic rainfall data generators using decentralised hydrological data and found that FL not only matched the performance of the centralised approach but also exhibited superior training stability (Welten *et al.* 2022b). The sequential IIL approach enables a workflow by transferring models from one institution to another (Sheller *et al.* 2020; Welten *et al.* 2021). The process of sending models from institution to institution in a pre-defined order can be repeated for several cycles (Sheller *et al.* 2020). Thus, no aggregation component is needed, as participating institutions perform their analysis task on the model they received from one prior institution. Just like FL, IIL has been used in various ML approaches, including brain tumour segmentation, breast cancer detection, mammography, retinal fundus classification, and synthetic rainfall data generation (Chang *et al.* 2018; Sheller *et al.* 2020; Welten *et al.* 2022a, 2022b). From a conceptual perspective, the principles of DA are capable of distributing generic code to various locations and perform computations decentrally. Therefore, we argue that the code transmitted through DA can act as a harmonising adapter among the locations, making it a suitable method to address the previously mentioned challenges associated with heterogeneity and the distributed nature of the setting. To standardise the varying data formats and enable the execution of different types of HD models, model adapters have been proposed (Werner *et al.* 2013). These model adapters are implemented separately for each HD model and consist of pre- and post-processing steps, transforming input data into HD models' native data format, and converting the output back to a standardised file format. In this study, our DA approach will leverage the idea of model adapters to flexibly accommodate various types of HD modelling software in a distributed fashion. In the context of our study, such adapters can include instruction sets, configuration parameters, and rainfall scenarios necessary for running the local HD models. How we build the synergistic integration of model adapters into DA will be covered in the upcoming section.

Distributed analytics applied to hydrodynamic modelling

We propose the use of model adapters incorporated into the code distributed through the paradigm of DA to overcome the aforementioned challenges (see Figures 1 and 2).

Due to the vast differences in how different HD modelling software operate (e.g., in terms of data format or execution steps), a model adapter that harmonises the data has to be implemented for each type of software. These model adapters

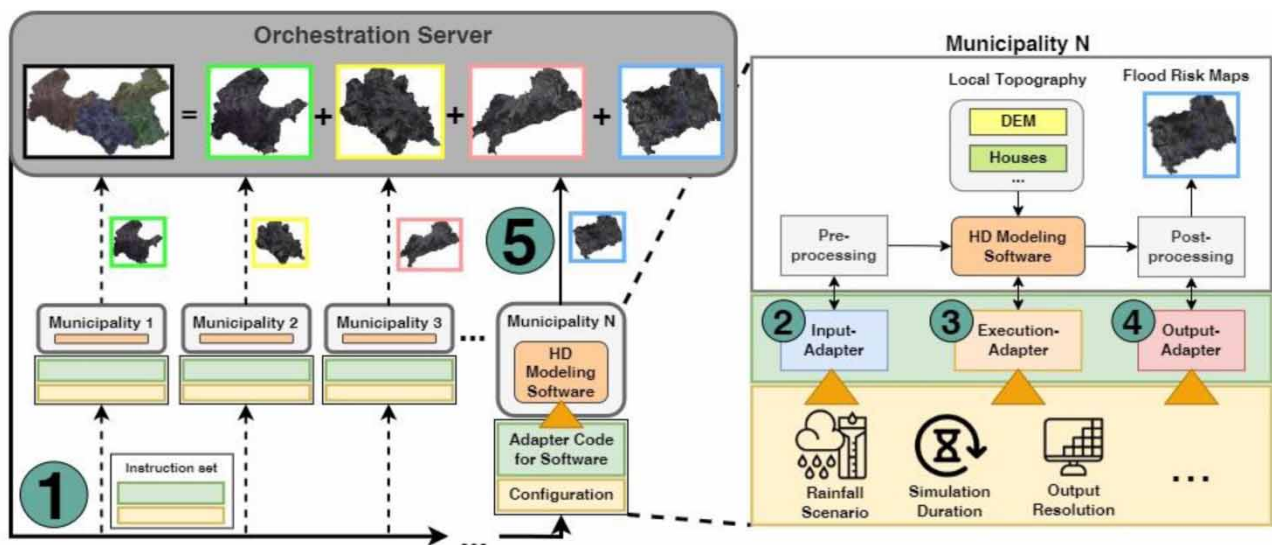


Figure 1 | Conceptualisation of our approach. First, an instruction set is distributed to each municipality (1). The instruction set defines configuration parameters that are used for the pluvial flood hazard mapping in order to guarantee harmonised outcomes and executable code for the simulation. Subsequently, a three-step (2–4) workflow involves the HD modelling software generating pluvial flood hazard maps based on the specified configuration parameters. In the fifth and last step (5), the produced maps are transmitted back to the central orchestration server, where the sub-maps are combined into one harmonised and cross-municipality pluvial flood hazard map.

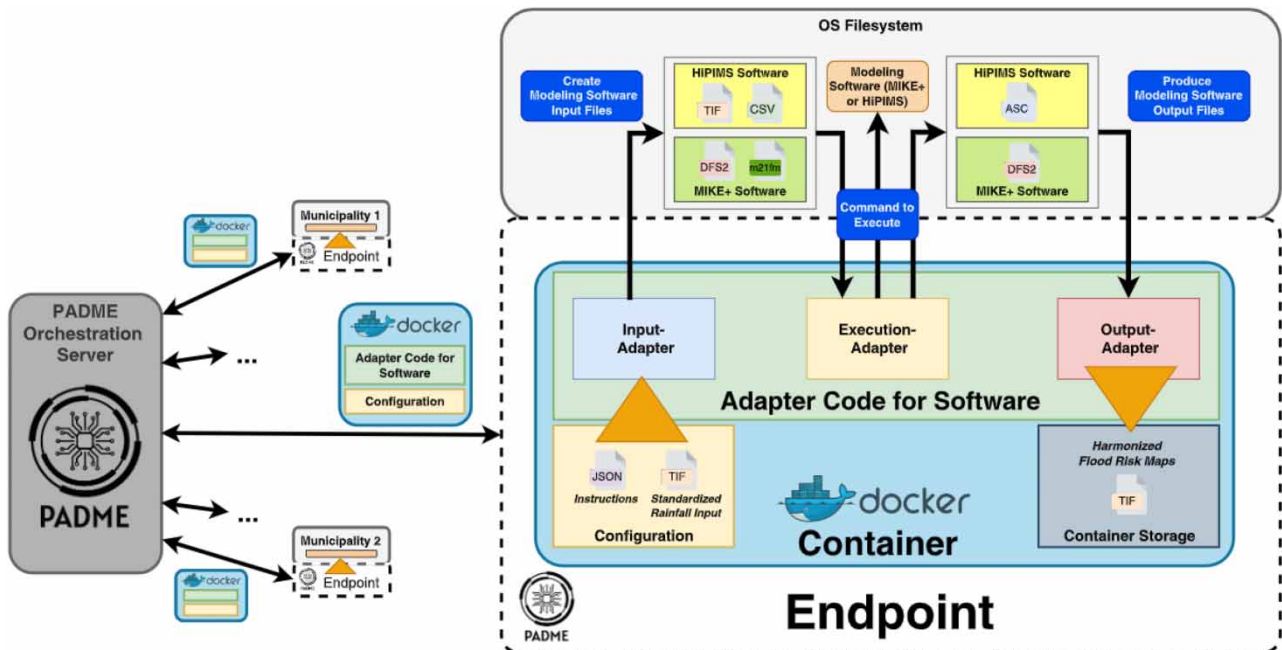


Figure 2 | Overview of the exemplary execution at an institution for both the modelling software HiPIMS and MIKE+. Our code for the HD model adapter contains three components consisting of an Input-Adapter, an Execution-Adapter and an Output-Adapter that are distributed to the institutions using the DA platform PADME. These adapters operate on a shared file system of the institution's underlying operating system and the Docker container. The Input-Adapter translates the rainfall data and instructions into the file format required by the institution's modelling software (GeoTIFF and CSV files for HiPIMS; DFS2 and m21fm files for MIKE+). The Execution-Adapter utilises these files to execute the corresponding HD modelling software. Finally, the produced pluvial flood hazard maps are harmonised and converted to the standardised GeoTIFF file format by the Output-Adapter. Once all three components have been executed, the harmonised pluvial flood hazard maps are transmitted back to the central server, where they are merged into an overarching map.

are then being utilised by DA to execute HD modelling software at the collaborating institutions that generate the pluvial flood hazard maps. We aim to streamline this implementation as much as possible through the following three-step workflow, including specific prerequisites. Note that we use the terms municipality and institutions interchangeably.

- 1. Unified adapter input/output (Prerequisites):** We determine that the input and the output of our adapter have a standardised file format for the rainfall input and output in form of the pluvial flood hazard map to simplify input data processing and enable interoperability between the pluvial flood hazard maps produced by each institution. In our study, we make use of the GeoTIFF file format for input and output (see Figure 2). Moreover, the output resolution is harmonised, e.g., to $8\text{ m} \times 8\text{ m}$, to account for potential variations in the resolutions at which different institutions' HD models operate. To instruct the model on how to process the input rainfall data (e.g., resolution, simulation time), we require an instruction set that defines these parameters and passes them on to the HD model. The instructions for the HD model are provided in JavaScript Object Notation (JSON) format. Just as the input rainfall data, this instruction set would be provided by the initiator prior to dispatching from the central location. While the GeoTIFF format is suited for portraying time series of raster data, it does not inherently support shape data. To incorporate such shape data, additional extensions would be necessary. However, the current concept of transmitting rainfall data and receiving pluvial flood hazard maps in return, does not necessarily require inclusion of shape data at this stage.
- 2. Input processing phase (Step 1):** Before starting an HD simulation, the rainfall data and instruction set must be processed locally at the institution to make them usable by each HD model. During the input processing phase, the rainfall data are transformed into the necessary format for the modelling software, generating input files that are compatible with the model, such as a rainfall mask, in advance of the actual execution of the simulation. Furthermore, the instruction set is parsed, and the model configuration is adjusted to reflect the requirements. In this initial step, all the necessary files and configurations for running the HD model are produced according to the number of rainfall events in the input

data. This facilitates the simulation of multiple pluvial flood hazard maps covering different rainfall scenarios. The Input-Adapter covers this phase of input processing (see Figure 2).

3. **Execution phase (Step 2):** During the execution phase, the HD modelling software is executed once for each input rainfall scenario. This process is conducted using the Execution-Adapter (see Figure 2) and is automated to the largest possible extent to minimise manual interaction.
4. **Output processing phase (Step 3):** The HD models' execution results in the generation of an equal number of pluvial flood hazard maps as there are input rainfall scenarios provided. During the subsequent output processing phase, an Output-Adapter (see Figure 2) converts the created pluvial flood hazard maps to the priorly specified standardised file format (see 1. Unified adapter input/output (Prerequisites)). Furthermore, other data may be returned by each institution, such as the local DEM, which the initiator can use for further inspection.

Combined, the phases of input, execution, and output processing compose our concept of the model adapter that is distributed to each institution using DA (see Figure 1). As indicated in section *Hydrodynamic Modelling* and Table 1, the technology stack varies among each participating institution, requiring us to develop tailored adapters for each unique situation. In the following sections, we provide an overview of our proof of concept implementation for the two well-established modelling software: HiPIMS and MIKE+.

Proof of concept implementation

In this section, we describe the general implementational details of the HD model adapters to match the characteristics of the HD modelling software: HiPIMS and MIKE+. Our implementation of the adapters is conducted with the Python programming language. The adapters are developed as independent modules to allow adaptability without requiring modifications to the HD models. The modules standardise model interfaces through a unified data format, acting as a common input/output for seamless data exchange across different systems, and abstracts operational interfaces by providing a set of generic operations (e.g., start simulation) applicable to various models. We design an adapter module customised for each type of HD model by implementing our three specified processing steps (see above): an Input-Adapter (for handling the input data), an Execution-Adapter (for executing the HD model), and an Output-Adapter (for harmonising the output flood inundation maps). Figure 2 gives a visual representation this workflow.

Input-Adapter

The Input-Adapter leverages the GeoTIFF file format, covering different rainfall scenarios as input. This rainfall input is spatially referenced and covers all regions. Thus, each Input-Adapter must geospatially manipulate these data, to sample the rainfall data corresponding to their coordinates. Regarding the modelling software in our study, we need to adjust the Input-Adapter according to their unique model input formats. In addition to the default availability of the DEM and other essential model components at each municipality, the HD model utilising HiPIMS relies on at least two additional inputs that must be generated by the Input-Adapter:

- A CSV file storing the rainfall values for differing parts of the coverage area and different points in time.
- A rainfall mask, a georeferenced raster file (we use GeoTIFF), that maps the rainfall values from the CSV file to locations inside of the coverage area.

Both input files are derived from the rainfall GeoTIFF provided by the central DA server. The Input-Adapter samples these data according to the municipality coordinates, after which each rainfall scenario is split into equal time slots (5 min). The rainfall values for each time slot are written into the CSV file, which is done for each rainfall scenario. The adapter also creates a mask (GeoTIFF file format) for each scenario that geographically references the rainfall values from the CSV (each value from the CSV has a unique identifier in the mask) to the areas inside the municipality's coverage area based on their geographic location.

The Input-Adapter for MIKE+ differs from the previous one, as this modelling software leverages the DFS2 file format for its input rainfall data. Initially, we once again resample the rainfall data to each municipality's bounds. For each rainfall scenario, the following steps are executed:

1. The data from the rainfall scenario are read into an array and written to a new file based on the reference file format used by MIKE+.

2. A simulation configuration file is created, with references to the aforementioned input file and a designated output directory.

Then, the execution of the MIKE+ simulation engine is triggered through sequential execution of the simulation configuration files.

Execution-Adapter

The Execution-Adapter uses the previously prepared inputs to run the HD model. The level of complexity that this phase requires differs depending on the software that is used. In our implementation, we provide the model with three parameters: Duration of the simulation (typically 2 h), flood output resolution in metres, and duration of rainfall during the simulation (usually 1 h, followed by 1 h of drainage time). For HiPIMS specifically, the execution process is straightforward since HiPIMS uses a Python API for model execution. We call the execution API with CSV and mask files, the local DEM, and the instruction set. Our adapter then runs simulations iteratively for each CSV and mask file combination. Running simulations with MIKE+ is more complicated because it lacks a Python API. Therefore, it involves one manual step, where a user must initiate the execution of the simulation configuration files located in a shared folder, which subsequently starts the simulations in a sequential and automated manner.

Output-Adapter

The Output-Adapter is responsible for combining the outputs of all simulations into a single file of standardised file format, which in our case is a GeoTIFF file. In addition to converting to the GeoTIFF file format, the data must be interpolated to the standardised resolution, effectively harmonising the output data. The output of each HiPIMS simulation consists of several ASCII files, of which we only require the file referring to the maximum flood inundation extent (the pluvial flood hazard maps) for our use case. This file is loaded for each simulation, interpolated to the resolution specified by the instructions, and added as a band to a GeoTIFF file. This GeoTIFF file has the same geospatial reference as the input rainfall file. The output of the MIKE+ simulation is stored in three DFS2 files: dynamic, statistics, and volume. In our specific scenario, we only need the statistics file, specifically the maximum flood inundation depth field from it. Each simulation's output is interpolated to the standardised resolution and saved to a GeoTIFF file.

DA framework

In order to implement our DA method, we need a compatible DA framework that manages the distribution and deployment/installation of our adapter code in each specific location. In our study, we utilise the DA framework known as Platform for Analytics and Distributed Machine Learning for Enterprises (PADME¹¹), which relies on containerisation (see Figure 2). Initially, we need to install the PADME endpoint at each of the participating municipalities (Welten *et al.* 2022c). From a technical perspective, the PADME endpoint includes a container engine that has the capability to run Docker¹² containers. Subsequently, we make the local HD models accessible through this endpoint and containerise our adapter code using Docker, which allows us to distribute the resulting container images across the DA network connecting the institutions. PADME uses a private–public key encryption protocol that guarantees that the containers (and their associated digital assets) remain encrypted during transmission and can only be accessed by the designated recipient. Specifically, in our case, when setting up each endpoint in every municipality, we assign both private and public keys to the endpoint for encryption and decryption purposes. As per this security protocol, the container is safely sent, decrypted, processed, and re-encrypted both at and among each endpoint. Once the adapter container image has been dispatched to each institution, it establishes a local connection to the HD model and initiates our pre-defined three-step workflow. One notable advantage of this approach is that it allows us to run the adapter code without being dependent on the host machine's operating system, as Docker is compatible with various OS environments. After addressing all technical requirements, our infrastructure is prepared for operation.

In conclusion, our proposed concept and its accompanying proof of concept implementation seek to address the current data diversity issues in conventional pluvial flood hazard assessment methods. It achieves this by facilitating the standardisation of outcomes from various HD model software and automating the simulation process. Based on a pre-defined input

¹¹ <https://padme-analytics.de>

¹² <https://www.docker.com>

format, the Input-Adapter prepares all files necessary for the Execution-Adapter to execute the HD model. Lastly, the Output-Adapter merges the results into a unified GeoTIFF file that is standardised and compatible with outputs from other simulations at other institutions. While the adapter implementations differ for each used software, the general workflow and structure (Input, Execution, Output) remains the same, enabling consistency between modelling results. Further, the DA framework PADME facilitates the collaboration of multiple institutions. As for the scope of this paper, this implementation enables the experiments of the following section.

RESULTS AND DISCUSSION

In this section, we present the use case to which we apply the combination of DA and HD modelling and showcase the results. Furthermore, we discuss our methodology and results including the benefits and limitations of our approach.

Study area

We apply the use case to a simulated study area of four adjacent urban areas located in the Rhein-Sieg Kreis of NRW, Germany, that have historically been susceptible to flooding (BPB 2021). We set up an HD model for each of these urban areas and treat them as distinct DA institutions located at different municipalities. To investigate the robustness of our approach, the institutions utilise different HD modelling software (MIKE+ and HiPIMS), and their models operate at differing DEM resolutions, ranging from 7 m × 7 m to 9 m × 9 m. This arrangement introduces the aforementioned heterogeneity (see section *Hydrodynamic Modelling*) into our scenario. We derived the DEMs from the open-data NRW product of 1 m resolution¹³ (downsampled to each institutions' resolution), and enhanced the DEMs through building location data, also derived from an open-data product.¹⁴ The institutions' models and data are hosted on four geographically distributed real-world machines to imitate and simulate a decentralised environment. We refer to Table 2 and Figure 3 for more information about these municipalities. Our objective in the scope of this feasibility study is to produce pluvial flood hazard maps with an 8 m × 8 m resolution using the decentralised HD models. In total, we aim to generate 50 pluvial flood hazard maps, each derived from a unique rainfall scenario. Finally, our goal is to centralise all these results into 50 combined pluvial flood hazard maps, each collectively representing the entire study area.

Results

Based on the above-mentioned setting, we initiate the dispatching of the containerised adapter code, equipped with the necessary input files, to the respective institutions. At each institution, we initiate the simulation that creates the pluvial flood hazard maps by starting the adapter container. Overall, our method was successful in generating the 50 harmonised pluvial flood hazard maps from different HD models hosted by the institutions, using both HiPIMS and MIKE+ software. Each institution received the same instruction set from the central server, processed the input data using the Input-Adapter and executed the HD model using the Execution-Adapter. Finally, the resulting pluvial flood hazard maps were processed using the Output-Adapter to ensure consistency in file format and resolution. After running all simulations, the PADME platform collects all produced outputs and combines them into joint maps that cover the whole study area. The total runtime for all 50 maps was around 10 h and 15 min, with Institution 1 submitting their results last. We applied GPU acceleration to all calculations. On average, each pluvial flood map took about 12 min and 18 s to complete. When extrapolated, the results of

Table 2 | Statistics about the simulated study area, consisting of four municipalities, each having their own local HD model

	Municipality name	Modelling software	Model resolution	Number of cells	Hardware (NVIDIA GPUs)
1	Sankt Augustin	MIKE +	9 m × 9 m	1,479,106	TITAN RTX 24 GB
2	Eitorf	HiPIMS	7 m × 7 m	1,429,000	GeForce RTX 3080 16 GB
3	Hennef	HiPIMS	8 m × 8 m	1,651,304	GeForce RTX 3090 Ti 24 GB
4	Ruppichteroth	HiPIMS	8 m × 8 m	957,735	GeForce RTX 4090 24 GB

The models vary in size, modelling software, and the resolution of their DEM, representing a realistic setting for our use case and its implemented HD model adapters.

¹³ DEM NRW 1m: <https://open.nrw/dataset/0c6796e5-9eca-4ae6-8b52-1fcc5ae5c481>

¹⁴ Hausumringe NW: <https://open.nrw/dataset/3f08a580-48ec-43c1-936d-d62f89c21cc9>

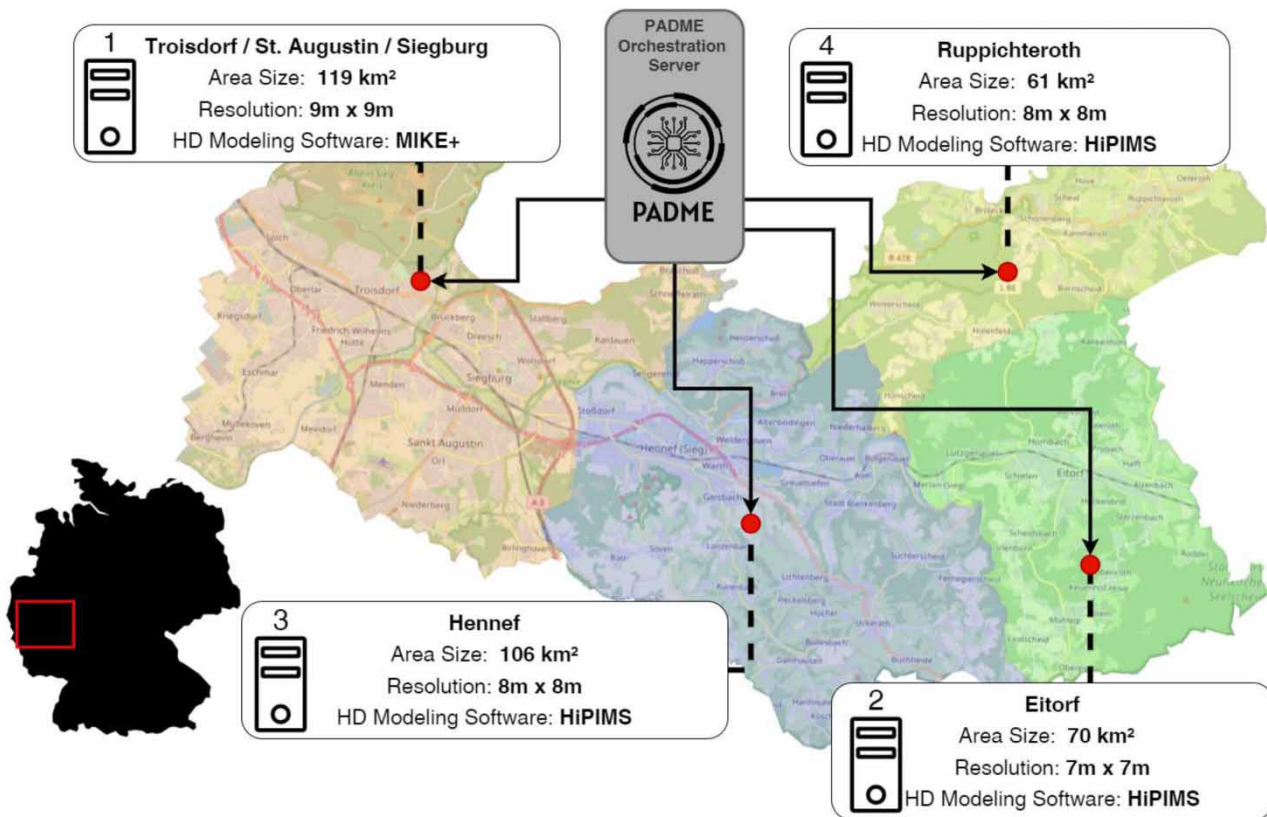


Figure 3 | Visualisation of the study area inside the Rhein-Sieg Kreis of NRW, Germany. The area consists of four adjacent municipalities, which we simulate to each have their own local HD model. These models are hosted on four geographically separate machines and differ in their modelling extent (number of cells), resolution, and modelling software. The goal is to access these models remotely using the DA framework PADME by sending instructions and rainfall data to each municipality, which returns the corresponding harmonised pluvial flood hazard maps that are aggregated at the central orchestration server.

our proof of concept align with the benchmarks set in the HiPIMS use cases, where 1.5 million cells were processed in 0.15 h using 8xK80 GPUs (Xia *et al.* 2019). We conduct a manual and visual examination of the 50 maps to identify any inconsistencies or unusual patterns that might suggest errors in data processing. The visual inspection of the output confirmed that the results are valid. We refer to Figure 4 for visualisation of one exemplary output map.

Discussion

The combination of DA and HD modelling for the decentralised generation of standardised pluvial flood hazard maps shows promising results. We derive the following advantages from our proof of concept:

- **Transregional comparison:** Our adapter approach harmonises configuration parameters across diverse technical conditions, fostering compatibility, and integration. This method effectively addresses the first and second challenges mentioned above. The uniformity across different regions allows for a common understanding of pluvial flood risks, as it accounts for regional variations while maintaining a baseline for comparison. Based on this transregional comparison, more effective flood mitigation strategies can be developed. For example, by understanding how different regions respond to similar hydrological conditions, one can devise more adaptable and efficient flood control measures.
- **Variety:** Concerning the third challenge, our approach can automatically generate a broader range of pluvial flood hazard maps produced under various rainfall scenarios that introduce additional viewpoints on a given scenario. These aspects contribute to the potential of our approach to facilitate a more comprehensive and multi-faceted view in flood risk assessments. Additionally, with our generative method, we have the ability to generate vast amounts of pluvial flood hazard maps for secondary use beyond its original intent. For example, the standardised and decentralised generation of pluvial flood

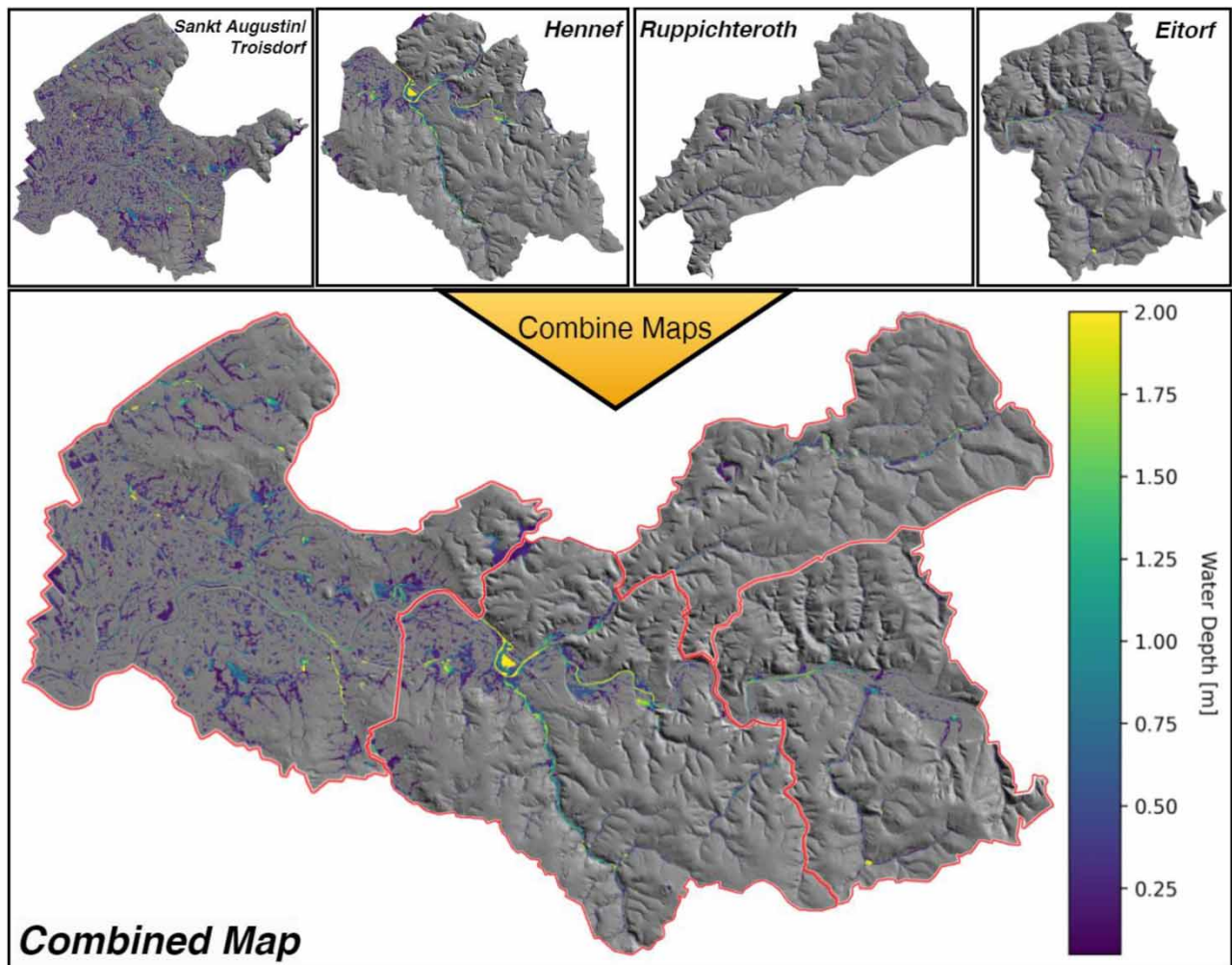


Figure 4 | Aggregated pluvial flood hazard maps created from our study area. After aggregation of the maps generated independently by the four municipalities, we combine these into an over-arching and harmonised map that can be used for further usage such as large-scale flood risk assessment.

hazard maps can also fuel the creation of deep learning (DL)-based flood models, of which multiple approaches have recently been proposed (Hofmann & Schüttrumpf 2021; Löwe *et al.* 2021). These models are designed to be trained on vast datasets of such pluvial flood hazard maps and can play an essential part in the early warning system against floods due to faster computation. Here, decentralised pluvial flood hazard mapping ensures that the training data are of high quality through use of existing HD models, while the scenario types are standardised.

- **Privacy:** Specifically, our method has demonstrated that HD models that are already present on geographically distributed servers can be executed in a decentralised manner without direct access to the underlying, possibly sensitive, data. For example, details about sewer networks, which are often classified due to security and privacy concerns, can remain confidential while still contributing to the overall accuracy of the HD models.
- **Granularity/precision:** In section *Hydrodynamic Modelling*, we have pointed out that the current state-of-the-art in pluvial flood hazard mapping on a larger scale follows unrealistic assumptions, e.g., spatially uniform rainfall, across a whole area. Our approach facilitates the creation of municipality-specific pluvial flood hazard maps, which can be aggregated to form a broader-scale view ('divide-and-conquer'). As a result, it considers the unique conditions and characteristics specific to each region. Hence, the aggregated map represents and includes the potential variations that can exist across multiple regions and incorporates information on a more fine-granular level.

- **Timeliness of information:** Our approach makes the integration of updates into the maps easier as centralisation is not required, ensuring that the maps accurately reflect the current landscape. Such updates can be software-alike, e.g., new HD models or data updates like changes in the topology due to both natural processes and human activities. Additionally, as environmental conditions and weather patterns evolve, our approach permits quick adjustments to various scenarios. This adaptability is crucial in accurately mapping pluvial flood hazards under changing circumstances. Hence, our approach ensures that the most current information is readily integrated into the mapping process.
- **Impulses for data spaces:** Recently, as part of the European strategy for data, a policy framework was outlined to develop common European Data Spaces (European Commission 2020). These include several data spaces, e.g., for environmental data (as part of the Green Deal Data Space) (European Commission 2020). Here, problems regarding data interoperability, governance, and data sovereignty were outlined, with these data spaces acting as one possible solution (European Commission 2020). In this context, our approach can be seen as a contribution to tackling the challenges associated with pluvial flood hazard data since it enables (data) interoperability across different sources and enables the seamless interaction between multiple HD models. As governance and data sovereignty are fundamental aspects in current data spaces, our approach also allows for responsible and secure data handling, management, and sharing since the HD models and the data coupled to them stay within institutional bounds. Hence, as a fundamental aspect of its design, our approach ensures that each collaborating institution retains control over its data.

Limitations

Our approach has certain limitations that need to be considered. First, developing model adapters requires considerable effort and experience with using HD modelling software and associated data formats. While we have implemented such adapters for two modelling software, the adapter concept might not be compatible with all existing modelling software. Furthermore, a limitation of our approach is the requisite expertise in containerisation technologies. The deployment of our model adapters, particularly in a decentralised setting, heavily relies on containerisation to ensure consistency, scalability, and ease of distribution. This necessitates a deep understanding of containerisation platforms and techniques, such as Docker, which are used to encapsulate the adapter environment and its dependencies into a portable and isolated container. This layer of complexity demands specialised skills in HD modelling, data format handling, and managing and orchestrating containers. Further, the decentralised nature of our approach could create bottlenecks, e.g., when the institutions' compute capacity varies. In such a case, the slowest institution determines the overall computation time, leading to possible delays. Nevertheless, as our aim has been to support hazard preparedness instead of addressing time-critical scenarios, potential delays might be less consequential.

CONCLUSIONS

In conclusion, our study introduces a novel approach to flood risk assessment that combines DA and HD modelling. By decentralising the creation of standardised pluvial flood hazard maps, we address the prevalent challenges of data heterogeneity and scarcity in the field of hydrology. Especially under the premise that the centralisation of the HD models is not possible, our method enables the utilisation of region-specific, high-quality HD models for decentral computation. We have demonstrated the practicality and flexibility of our approach by applying it to a study area with diverse HD models and configurations. These have been harmonised using a three-component adapter that is compatible with various technical conditions. Importantly, our method reduces human intervention through an (semi-)automated workflow, facilitating the generation of a more extensive range of pluvial flood hazard maps. In summary, we provide impulses to the field of hydrology in the form of a new method for harmonised pluvial flood hazard mapping that can enhance disaster preparedness and management efforts in regions prone to extreme weather events and flash floods. Future work could see the method applied to a wider variety of HD models and differing study areas. Additionally, we see the potential of our methodology to fuel DL approaches, as we have pointed out before, by generating increased quantities of pluvial flood hazard maps representing different or specific scenarios that can be employed in the training of these models.

ACKNOWLEDGEMENTS

This work has been part of the DeepWarn project and partially funded by the Federal Ministry of Education and Research (BMBF) and the Ministry of Culture and Science of the German State of NRW (MKW) under the Excellence Strategy of the Federal Government and the Länder.

AUTHOR CONTRIBUTIONS

Sascha Welten and Adrian Holt contributed to the conceptualisation, methodology, software development, and validation of this research. They actively participated in writing the original draft and reviewed and edited the manuscript. Julian Hofmann contributed to the conceptualisation and methodology, as well as reviewing the manuscript. Sven Weber played a role in reviewing the manuscript, alongside contributing to software development of the underlying DA platform. Elena-Maria Klopries reviewed this work. Holger Schüttrumpf and Stefan Decker supervised and reviewed this work.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

- Bundeszentrale für politische Bildung 2021 *Jahrhunderthochwasser 2021 in Deutschland*. Available from: <https://www.bpb.de/kurz-knapp/hintergrund-aktuell/337277/jahrhunderthochwasser-2021-in-deutschland/> (accessed 21 August 2023).
- Chang, K., Balachandar, N., Lam, C., Yi, D., Brown, J., Beers, A., Rosen, B., Rubin, D. L. & Kalpathy-Cramer, J. 2018 Distributed deep learning networks among institutions for medical imaging. *Journal of the American Medical Informatics Association* **25**, 945–954.
- Chen, Y., Qin, X., Wang, J., Yu, C. & Gao, W. 2020 Fedhealth: A federated transfer learning framework for wearable healthcare. *IEEE Intelligent Systems* **35**, 83–93.
- Cybersecurity and Infrastructure Security Agency 2023 *Water and Wastewater Sector*. Available from: <https://www.cisa.gov/topics/critical-infrastructure-security-and-resilience/critical-infrastructure-sectors/water-and-wastewater-sector> (accessed 9 October 2023).
- Davies, R. 2021 *Germany – More Than 100 fatalities Reported After Floods in Western States*. FloodList. Available from: <https://floodlist.com/europe/germany-floods-july-2021-update> (accessed 21 August 2023).
- Davies, R. 2022 *Pakistan – Almost 1,000 Dead, 33 Million Affected in Worst Floods in a Decade*. FloodList. Available from: <https://floodlist.com/asia/pakistan-floods-update-august-2022> (accessed 21 August 2023).
- Davies, R. 2023a *Greece, Turkey and Bulgaria – 11 Dead After Record Rainfall Triggers Catastrophic Floods*. FloodList. Available from: <https://floodlist.com/europe/greece-turkey-bulgaria-floods-september-2023> (accessed 10 October 2023).
- Davies, R. 2023b *Libya – Thousands Feared Dead After Floods in Eastern Region*. FloodList. Available from: <https://floodlist.com/africa/libya-floods-september-2023> (accessed 10 October 2023).
- Davies, R. 2023c *USA – State of Emergency After Deadly Floods in New York*. FloodList. Available from: <https://floodlist.com/america/usa/floods-new-york-july-2023> (accessed 10 October 2023).
- Deutscher Wetterdienst 2023 *Historical Hourly Radolan Grids Data*. Available from: https://opendata.dwd.de/climate_environment/CDC/grids_germany/hourly/radolan/historical/asc/ (accessed 21 August 2023).
- DHI 2023 *Mike 21 User Manual*. Available from: https://manuals.mikepoweredbydhi.help/2017/MIKE_21.htm (accessed 21 August 2023).
- Elmahdy, S., Ali, T. & Mohamed, M. 2020 Flash flood susceptibility modelling and magnitude index using machine learning and geohydrological models: A modified hybrid approach. *Remote Sensing* **12**, 2695.
- Essenfelder, A. H., Bagli, S., Mysiak, J., Pal, J. S., Mercogliano, P., Reder, A., Rianna, G., Mazzoli, P., Broccoli, D. & Luzzi, V. 2022 Probabilistic assessment of pluvial flood risk across 20 European cities: A demonstrator of the Copernicus disaster risk reduction service for pluvial flood risk in urban areas. *Water Economics and Policy* **8**, 2240007. doi:10.1142/S2382624X22400070.
- European Commission 2020 *A European Strategy for Data. Communication COM (2020) 66 Final*. European Commission. Brussels. Available from: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066>.
- Hofmann, J. & Schüttrumpf, H. 2019 Risk-based early warning system for pluvial flash floods: Approaches and foundations. *Geosciences* **9**, 127.
- Hofmann, J. & Schüttrumpf, H. 2020 Risk-based and hydrodynamic pluvial flood forecasts in real time. *Water* **12**, 1895.
- Hofmann, J. & Schüttrumpf, H. 2021 Floodgan: Using deep adversarial learning to predict pluvial flooding in real time. *Water* **13**, 2255.
- HW Karten 2023 *Starkregen*. Available from: <https://www.hw-karten.de/index.html?Module=Starkregen> (accessed 21 August 2023).
- Kim, Y., Sun, J., Yu, H. & Jiang, X. 2017 Federated tensor factorization for computational phenotyping. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 887–895.
- Leandro, J., Chen, A. S., Djordjevic, S. & Savic, D. A. 2009 Comparison of 1d/1d and 1d/2d coupled (sewer/surface) hydraulic models for urban flood simulation. *Journal of Hydraulic Engineering* **135**, 495–504.
- Li, L., Fan, Y., Tse, M. & Lin, K. Y. 2020 A review of applications in federated learning. *Computers & Industrial Engineering* **149**, 106854.
- Löwe, R., Böhm, J., Jensen, D. G., Leandro, J. & Rasmussen, S. H. 2021 U-flood-topographic deep learning for predicting urban pluvial flood water depth. *Journal of Hydrology* **603**, 126898.

- McMahan, B., Moore, E., Ramage, D., Hampson, S. & y Arcas, B. A. 2017 Communication-efficient learning of deep networks from decentralized data. In: *Artificial Intelligence and Statistics*, Vol. 54 (Aarti Singh and Jerry Zhu, eds.). PMLR, 20–22 April 2017, Fort Lauderdale, FL, USA, pp. 1273–1282.
- Mudashiru, R. B., Sabtu, N., Abustan, I. & Balogun, W. 2021 [Flood hazard mapping methods: A review](#). *Journal of Hydrology* **603**, 126846.
- Pyka, C. E. 2020 Ein GIS-basiertes Planungstool zur Unterstützung von Kommunen in ihrer Anpassung an Starkniederschläge, In: 50. IWASA Internationales Wasserbau Symposium Aachen 2020, 20. Starkregenforum, 2020-01-09 - 2020-01-10, Aachen, Germany [50. Internationales Wasserbau Symposium Aachen, Aachen (Germany), 9 Jan 2020 - 10 Jan 2020]. pp. 76–85. Available from: <https://publications.rwth-aachen.de/record/813863>.
- Rheinisch-Bergischer Kreis 2023 [Starkregengefahrenkarte](#). Available from: <https://www.rbk-direkt.de/starkregengefahrenkarte-fuer-den-rbk.aspx> (accessed 10 October 2023).
- Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., Milchenko, M., Xu, W., Marcus, D. Colen, R. R. & Bakas, S. 2020 [Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data](#). *Scientific Reports* **10**, 1–12.
- Silva, S., Gutman, B. A., Romero, E., Thompson, P. M., Altmann, A. & Lorenzi, M. 2019 Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, IEEE, pp. 270–274.
- Stadt Bochum 2023 [Starkregengefahrenkarte](#). Available from: <https://www.bochum.de/Starkregengefahrenkarte> (accessed 21 August 2023).
- Stadt Dortmund 2023 [Starkregengefahrenkarte](#). Available from: https://www.dortmund.de/de/leben_in_dortmund/planen_bauen_wohnen/stadtentwaesserung/grundstuecksentwaesserung/starkregen_und_hochwasser/starkregengefahrenkarte/index.html (accessed 10 October 2023).
- Teng, J., Jakeman, A. J., Vaze, J., Croke, B. F., Dutta, D. & Kim, S. 2017 [Flood inundation modelling: A review of methods, recent advances and uncertainty analysis](#). *Environmental Modelling & Software* **90**, 201–216.
- Welten, S., Neumann, L., Yediel, Y. U., da Silva Santos, L. O. B., Decker, S. & Beyan, O. 2021 [Dams: A distributed analytics metadata schema](#). *Data Intelligence* **3**, 528–547.
- Welten, S., Hempel, L., Abedi, M., Mou, Y., Jaberansary, M., Neumann, L., Weber, S., Tahar, K., Ucer Yediel, Y. Löbe, M., Decker, S., Beyan, O. & Kirsten, T. 2022a [Multi-institutional breast cancer detection using a secure on-boarding service for distributed analytics](#). *Applied Sciences* **12**, 4336.
- Welten, S., Holt, A., Hofmann, J., Schelter, L., Klopries, E. M., Wintgens, T. & Decker, S. 2022b [Synthetic rainfall data generator development through decentralised model training](#). *Journal of Hydrology* **612**, 128210.
- Welten, S., Mou, Y., Neumann, L., Jaberansary, M., Ucer, Y. Y., Kirsten, T., Decker, S. & Beyan, O. 2022c [A privacy-preserving distributed analytics platform for health care data](#). *Methods of Information in Medicine* **61**, e1–e11.
- Werner, M., Schellekens, J., Gijsbers, P., van Dijk, M., van den Akker, O. & Heynert, K. 2013 [The delft-fews flow forecasting system](#). *Environmental Modelling & Software* **40**, 65–77.
- Westra, S., Fowler, H. J., Evans, J. P., Alexander, L. V., Berg, P., Johnson, F., Kendon, E. J., Lenderink, G. & Roberts, N. 2014 [Future changes to the intensity and frequency of short-duration extreme rainfall](#). *Reviews of Geophysics* **52**, 522–555.
- Xia, X., Liang, Q. & Ming, X. 2019 [A full-scale fluvial flood modelling frame-work based on a high-performance integrated hydrodynamic modelling system \(HIPIMS\)](#). *Advances in Water Resources* **132**, 103392.

First received 4 November 2023; accepted in revised form 1 January 2024. Available online 22 January 2024