Late Holocene flood magnitudes in the Lower Rhine river valley and upper delta resolved by a two-dimensional hydraulic modelling approach

Bas van der Meulen1 | Anouk Bomers2 | Kim M. Cohen1 | Hans Middelkoop1

1Faculty of Geosciences, Utrecht University, Princetonlaan 8a, Utrecht, CB, 3584, the Netherlands
2Faculty of Engineering Technology, University of Twente, Dienstweg 1, Enschede, ND, 7522, the Netherlands

Correspondence
Bas van der Meulen, Faculty of Geosciences, Utrecht University, Princetonlaan 8a, 3584. CB Utrecht, the Netherlands.
Email: b.vandermeulen@uu.nl

Abstract
Palaeoflood hydraulic modelling is essential for quantifying ‘millennial flood’ events not covered in the instrumental record. Palaeoflood modelling research has largely focused on one-dimensional analysis for geomorphologically stable fluvial settings because two-dimensional analysis for dynamic alluvial settings is time consuming and requires a detailed representation of the past landscape. In this study, we make the step to spatially continuous palaeoflood modelling for a large and dynamic lowland area. We applied advanced hydraulic model simulations (1D–2D coupled set-up in HEC-RAS with 950 channel sections and 108 x 103 floodplain grid cells) to quantify the extent and magnitude of past floods in the Lower Rhine river valley and upper delta. As input, we used a high-resolution terrain reconstruction (palaeo-DEM) of the area in early mediaeval times, complemented with hydraulic roughness values. After conducting a series of model runs with increasing discharge magnitudes at the upstream boundary, we compared the simulated flood water levels with an inventory of exceeded and non-exceeded elevations extracted from various geological, archaeological and historical sources. This comparison demonstrated a Lower Rhine millennial flood magnitude of approximately 14,000 m³/s for the Late Holocene period before late mediaeval times. This value exceeds the largest measured discharges in the instrumental record, but not the design discharges currently accounted for in flood risk management.

KEYWORDS
hydraulic model, Lower Rhine, millennial flood, palaeohydrology, Rhine delta

1 | INTRODUCTION

Palaeoflood analysis informs current and future flood risk (St. George et al., 2020; Wilhelm et al., 2019). Identifying past flooding patterns helps to determine areas at risk under different management scenarios (Alkema & Middelkoop, 2005; Remo et al., 2009), and quantifying past flood magnitudes helps to assess future discharge extremes (Benito & Thorndycraft, 2005; Schendel & Thongwichian, 2017). Many studies on past floods aim to quantify the peak discharge reached during specific events (Herget & Meurs, 2010; Hu et al., 2016; Toonen et al., 2013) or a non-exceeded discharge over a prolonged time period (Enzel et al., 1993; England et al., 2010). Traditionally, palaeoflood modelling studies have been conducted in geomorphologically stable fluvial settings, such as bedrock canyons,
where calculating a representative discharge is relatively straightforward (Kochel & Baker, 1982; overviews in Baker, 2008 and Benito & Díez-Herrero, 2015). In settings where the fluvial landscape has changed due to natural fluvial processes and anthropogenic activities, however, the floodplain topography and channel bathymetry must be reconstructed prior to calculations of palaeoflood discharges (Herget & Meurs, 2010; Machado et al., 2017; Toonen et al., 2013).

Palaeoflood reconstructions by hydraulic models have been largely limited to one-dimensional (1D) calculations, applied to a single or few consecutive valley cross-sections (Benito & Díez-Herrero, 2015; Busschers et al., 2011; Webb & Jarrett, 2002). Although major advantages of deploying 1D hydraulic models are computational speed and little required input, these models cannot resolve spatial aspects of flooding such as floodwater dispersal patterns over the floodplain. Therefore, 1D modelling produces first-order calculations of palaeoflood peak magnitudes rather than actual palaeoflood simulations. Introducing a two-dimensional (2D) model component lessens the assumptions in the modelling step of palaeoflood research and yields more realistic output than a 1D model (see discussion in Herget et al., 2014), providing stages and discharges at different points in time over the entire model area. Many studies discuss the performance of 2D hydraulic models for present landscape...
situations (Bomers et al., 2019b; Czuba et al., 2019; Dottori et al., 2013), and the importance of applying these in palaeoflood research is often emphasised, but the actual use is usually outside time and budget constraints (Benito & Díez-Herrero, 2015; England et al., 2010).

Besides computing power limits, the main problem associated with 2D modelling of past floods is a requirement for accurate high-resolution input data, most notably a reconstruction of the terrain. Generating this input involves the use of geological and geomorphological field data, and experimental GIS techniques (van der Meulen et al., 2020). A few studies have applied 2D models to simulate relatively recent floods, occurring in the nineteenth century (Bomers et al., 2019; Calenda et al., 2003; Hesselink et al., 2003; Skubics et al., 2016) and the twentieth century (Ballesteros et al., 2011; Bomers et al., 2019a; Denlinger et al., 2002; England et al., 2007; Masoero et al., 2013; Stamataki & Kjeldsen, 2020). For such recent events, abundant data are available for model schematisation and validation. Changes in morphology can be regarded as either insignificant or well documented and readily incorporated. Simulations of older floods in larger alluvial reaches, such as the Lower Rhine, are lacking due to terrain reconstruction demands. However, these reaches are usually densely populated, warranting detailed assessments of river flood risk. For such regions, it would be of interest to simulate extreme events in earlier historic or pre-historic times, thereby constraining peak discharges over longer time periods that are usually the focus of palaeoflood research.

The Lower Rhine is a large and economically important meandering river which transitions into the Rhine delta near the German–Dutch border (Figure 1). Here, it divides into three main distributaries: the Waal, the Nederrijn and the IJssel. These river branches have been embanked since late mediaeval times (Hesselink, 2002). Across the valley and delta, the river and floodplain have been heavily modified by anthropogenic activities (Hudson et al., 2008; Kalweit et al., 1993; Lammersen et al., 2002; van der Meulen et al., 2020). Throughout the Holocene, Lower Rhine floods have left scattered geological and geomorphological traces in the landscape (Gouw & Erkens, 2007; Minderhoud et al., 2016; Pierik et al., 2017; Toonen et al., 2013; Willemsse, 2019). In historic times, large floods were a recurring threat for the inhabitants of the area (Tol & Langen, 2000), and at present, considerable resources are dedicated to developing and maintaining flood protection structures. A major challenge lies in determining appropriate safety standards for river management works and underlying magnitude–frequency relationships of extreme events (Hegnauer et al., 2014; Hegnauer et al., 2015; van Alphen, 2016).

Previous palaeohydrological research in the delta apex region has provided estimates of Lower Rhine flood frequencies and relative magnitudes over the past centuries to millennia based on sedimentary archives obtained from oxbow-lake fills (Cohen et al., 2016; Toonen, 2013; Toonen et al., 2015; Toonen et al., 2017). These studies have shown that between 4.2 and 0.8 ka (kiloannum = thousands of years ago), that is, in the Late Holocene up to embankment along the downstream reaches, five flood events with recurrence times close to 1,000 years occurred. The largest of these was dated to 785 CE based on a tentative correlation between radiocarbon-dated flood deposits and historical sources (Cohen et al., 2016; Toonen, 2013). Recently, a palaeo-DEM has been constructed for the Lower Rhine river valley and upper delta in the first millennium CE (van der Meulen et al., 2020). The present research builds upon these studies by applying a hydraulic model to the palaeo-DEM, resolving absolute magnitudes of extreme floods in the Late Holocene period preceding the onset of river embankment in late mediaeval times.

The purpose of this study is to constrain past flooding patterns and magnitudes in the Lower Rhine river valley and upper delta (Figure 1). Specifically, we aim to quantify the peak discharge of the most extreme (‘millennial’) floods that occurred. Likely, this value is in between the largest measured flood discharge (~12,000 m³/s) and the largest discharges generated by stochastic weather simulations coupled to hydrological models (GRADE project: ~24,000 m³/s; Hegnauer et al., 2014; Hegnauer et al., 2015). We set up a 1D–2D coupled hydraulic model with reconstructed terrain and roughness as input to simulate the propagation of extreme discharge waves in the past landscape setting. Subsequently, we compare the output water levels with geological, geomorphological and archaeological observations informing on minimum and maximum palaeoflood levels in the study area. In addition to calculating a Late Holocene millennial flood magnitude for the Lower Rhine river, this study demonstrates for the first time the potential and the difficulties of spatially continuous palaeoflood modelling in large lowland areas. At the least, such research requires (1) an accurate reconstruction of the river and floodplain, (2) a numerical model with a two-dimensional component and (3) a collection of spatially distributed palaeoflood levels.

2 | MATERIALS AND METHODS

The overall ‘inverse modelling’ approach applied in this study comprises three major steps (Figure 2). The first step includes the reconstruction of the landscape, which encompasses floodplain topography, river position and bathymetry, and roughness values matching past land cover. Further, an inventory of palaeoflood levels was generated based on earlier geological and archaeological investigations. The second step consists of hydraulic modelling, using the past landscape as input. Next to terrain and roughness, the model set-up requires upstream and downstream boundary conditions, and a rasterization of the model area consisting of a 2D grid on the floodplain and 1D profiles in the channels. Key to the ‘inverse modelling’ approach is that multiple model runs are conducted with discharge waves of increasing magnitude as upstream boundary condition. In the third step, the output of the successive model runs is compared with the inventory of palaeoflood levels. The match between observed and simulated flood levels (m) is used to evaluate the model output resulting from different peak flows (m³/s), thus constraining the maximum discharge that occurred in the studied period.
The upstream boundary of the model area is at river km 638, between Andernach and Bonn, Germany, where the Rhine changes from a bedrock-confined river to an alluvial meandering river in a terraced valley floor (Figure 1). From here, the model area covers the entire floodplain of the Lower Rhine valley and the delta apex region. The downstream boundaries are placed at river km 908 (Waal), km 912 (Nederrijn) and km 950 (IJssel), where the delta plain widens substantially and water level slopes of the river branches are influenced by tides (Kleinhans et al., 2011; Stouthamer et al., 2011). Laterally, the area extends up to the edges of the valley floor and delta plain, where the terrain steeply rises (Figure 1), covering all fluvial geomorphological units of Holocene age and latest Pleistocene terrace surfaces. These lateral boundaries were selected to prevent simulations of hypothetically extreme floods from requiring additional outflow locations, even though realistic extreme floods did not inundate the full width of the area according to geological and historical evidence. In the present situation, the model area can be clearly subdivided into embanked floodplains along the river channels and flood-protected areas on the other sides of the dykes. Such a distinction cannot be made for early mediaeval times when dykes were not yet present and floodplains were much wider.

2.1 | Topography and bathymetry

The most important component of the model schematisation is the terrain, represented by a digital elevation model (DEM) of the past situation: a palaeo-DEM. We used a recently produced reconstruction of the topography, river position and channel bathymetry of the Lower Rhine valley and upper delta in early mediaeval times (circa 800 CE; Figure 3). This palaeo-DEM was constructed by stripping a modern light detection and ranging (LiDAR) DEM from all changes to the topography by anthropogenic activities (van der Meulen et al., 2020), after correcting for recent mining-induced subsidence (Harnischmacher & Zepp, 2014). In addition, the river position and channel bathymetry were reconstructed by combining available historical, archaeological and geological data, and the natural floodplain elevation along the river was restored using geomorphological interpolations (van der Meulen et al., 2020).

The early mediaeval target age of the palaeo-DEM pre-dates any major direct modifications to the terrain by humans. Therefore, the distal floodplain topography is representative for the entire Late Holocene period up to late mediaeval times (van der Meulen et al., 2020). The proximal floodplain topography and river position are considerably more variable through time, but within the Late Holocene the fluvial style and channel dimensions of the Rhine have been fairly stable (Klosterman, 1992; Erkens et al., 2011). Furthermore, local morphological change due to fluvial erosion and deposition likely did not alter the overall conveyance capacity for peak discharges at the scale of the entire valley. Similarly, indirect human impacts on the topography, most notably changes in sediment budget (Hoffmann et al., 2007, 2009; Middelkoop et al., 2010; Notebaert & Verstraeten, 2010), affected the area during the Late Holocene, but probably hardly influenced the propagation of extreme discharge waves over the full model area. As precisely these extremes are the focus of our simulations and analyses, we regard the simulation results produced using the early mediaeval palaeo-DEM representative for older Late Holocene times.

2.2 | Hydraulic roughness

The next component of the model schematisation is hydraulic roughness, expressed as Manning’s $n$ values. Vegetation distribution and land use are broadly known for early mediaeval times in the study area, but detailed constraints are scarce (Gouw-Bouman, in prep). Therefore, we developed a simplified land cover reconstruction. We distinguished five landscape classes based on distance to the river and relative floodplain elevation (with respect to a second-order polynomial trend surface), each associated with different natural vegetation and land use characteristics, and hence different average hydraulic roughness values (Figure 4; Table 1). We used standard values for
different vegetation types (Chow, 1959) as starting point to assign \( n \) values to landscape classes based on their estimated overall land cover composition.

The areas located more than 5 m above the trend surface (high grounds, class H) were mostly covered by forest, resulting in high \( n \) values. Note that class H covers only a small area (Table 1), because we cut off the model area where elevation significantly rises. The river itself (polygon obtained from palaeo-DEM, class R) consisted mostly of bare gravel and sand, resulting in low \( n \) values. The proximal floodplain (class P), which we defined as a 1-km-wide zone next to the river (Figure 4), had dispersed riparian vegetation of the type modern floodplain-ecologists aim to restore (van Oorschot, 2017) and was assigned relatively high \( n \) values. Realistically, the width of this zone greatly varied along the river, and thus our definition is a major simplification. However, given the large study area, a detailed reconstruction was not feasible. The higher parts of the distal floodplain (class DH) were used most intensively for settlement and agriculture (Pierik & van Lanen, 2019), and thus had little natural vegetation. Agricultural land in early mediaeval times consisted of smaller fields than today, with abundant hedges and local stands of trees. Therefore, we assigned higher \( n \) values to this class than used for agricultural land in models of present situations. The lower parts of distal floodplain (class DL) were, and still are, relatively wet and mostly used as meadows for grazing and hay production. In these areas, open grasslands prevailed, but locally swamp forests (carr) occurred as well (Gouw-Bouman et al., in prep). Grasslands have low \( n \) values, but the characteristically dense swamp forests have high \( n \) values. However, it is not possible to reconstruct the precise spatial distribution of grasslands and swamp forests within this zone. Overall, this class mainly consisted of grass; therefore, we assigned it low \( n \) values throughout the model area.

Besides average roughness values (\( n_{\text{best}} \)), we assigned lower (\( n_{\text{min}} \)) and upper estimates (\( n_{\text{max}} \)) to all classes, except for class H with its marginal occurrence (Table 1). The \( n_{\text{min}} \) and \( n_{\text{max}} \) were assigned to account for the uncertainties inherent in assigning roughness values (Bomers et al., 2019a; Chow, 1959; Warmink et al., 2013). All
landscape classes, expect class R, consist of a mixture of relatively open and forested parts. Because of the co-occurrence of different vegetation types, Manning’s n values beyond the lower and upper estimates given in Table 1 are unlikely to be correct for any Late Holocene situation. In the Rhine area, significant catchment-scale deforestation occurred since the Bronze Age circa 3.5 ka. There was a slight increase in natural vegetation cover in the Dark Ages around 1.5 ka, due to population decline and land abandonment following Roman times (Teunissen, 1990; Kaplan et al., 2009; Gouw-Bouman, in prep). For these reasons, lower n values (between n_min and n_best) may be appropriate for Middle Holocene to early Late Holocene times, whereas higher values (between n_best and n_max) may be typical for Roman times and for late mediaeval times. We used the n_best values in the model runs, unless otherwise stated.

### 2.3 Model set-up

We set up a 1D–2D coupled hydraulic model in HEC-RAS (v. 5.0.3). The main channels of the Rhine river and its distributaries were schematised by 500-m-spaced 1D profiles, whereas the floodplains were discretised on a 2D grid with a resolution of 200 × 200 m (Figure 5). Flexible grid shapes were used along the model domain boundaries and along the transition from reworked to inherited floodplain (van der Meulen et al., 2020; Figure 5). Rectangular grid cells were used in the remainder of the model domain. The 1D profiles were coupled with the 2D grid using the weir equation. The weir coefficient was set to a value corresponding to overland flow to enable correct prediction of the flow transfer and to keep the model stable. We used the full momentum equations to solve the system, because these resulted in more accurate model results compared with the simplified diffusive wave equations.

The upstream boundary condition consists of a discharge time series (Figure 6). An initial discharge of 1,000 m$^3$/s was used in all runs to avoid a dry channel at the start of the simulations. We selected the two largest measured discharge waves as input. These are the floods of December 1925/January 1926 (Bomers et al., 2019a) and January 1995. For the palaeoflood simulations, we used the shape of the 1926 discharge wave as a representative standard hydrograph because it closely resembles the average of modelled extreme flood waves of the Rhine near Andernach (Hegnauer et al., 2014; Hegnauer et al., 2015). By scaling this hydrograph, we obtained a series of input

---

**TABLE 1** Landscape classes and assigned Manning’s n values

| Class | Definition | % area | n_min | n_best | n_max |
|-------|------------|--------|-------|--------|-------|
| H     | High ground Area > trend surface + 5 m vertical | 4.1    | 0.1   | 0.1    | 0.1   |
| R     | River bed and banks River polygons + 100 m buffer | 5.0    | 0.025 | 0.03   | 0.045 |
| P     | Proximal floodplain Border of R + 1,000 m buffer | 14.7   | 0.06  | 0.07   | 0.08  |
| DH    | Distal floodplain, high Remaining area > trend surface | 38.6   | 0.04  | 0.05   | 0.06  |
| DL    | Distal floodplain, low Remaining area < trend surface | 37.6   | 0.035 | 0.04   | 0.055 |

---

**FIGURE 5** Model set-up of the 1D–2D coupled hydraulic model on top of the palaeo-DEM. The close-up, between Rees and Emmerich, visualises the nature of the 1D–2D coupled grid.
washes with different peak discharges, from 10,000 to 30,000 m$^3$/s with intervals of 2,000 m$^3$/s (Figure 6). We simplified the discharge waves of the Lower Rhine tributaries Sieg, Ruhr and Lippe to constant input values of 250, 500 and 250 m$^3$/s, respectively. Normal depths were used as downstream boundary conditions at the locations where water can potentially leave the model domain (Figure 5). These normal depths were calculated using Manning’s equation (Brunner, 2016).

In total, we conducted 15 simulations using the palaeoflood model set-up, one for each different peak discharge and two extra runs for both the 10,000 and 24,000 m$^3$/s input waves using the minimum and maximum Manning’s $n$ values for all classes (Table 1) to evaluate model sensitivity to roughness values. In addition to model runs on the Late Holocene pre-embanked landscape, we conducted model runs on the present landscape. For these we used a schematisation of river and floodplain geometry (‘Baseline’ dataset) provided by the Ministry of Infrastructure and Water Management (RWS) of the Netherlands and the Landesamt für Natur, Umwelt und Verbraucherschutz (LANUV) of North Rhine-Westphalia. The model set-up for the present landscape closely resembled the palaeoflood model set-up, but both the channel and the embanked floodplains were schematised by 1D profiles. The areas protected by embankments were discretised on a 2D grid (Bomers et al., 2019b).

### 3 | RESULTS

#### 3.1 | Inundation depth and extent

The effect of topography on flooded area is extensive. In the present landscape setting, an extreme flood such as that of January 1995 only inundates the embanked floodplains, amounting to about 450 km$^2$, given that no dyke breaches occur (Figure 7A). In contrast, a similar flood inundates about 2,370 km$^2$ in the pre-embanked landscape setting (Figure 7C). Conversely, inundation depths are much higher in the flooded area between the embankments in the present situation.
than in the past situation when overbank flow was unconfined (Figure 7D).

The water depths in the pre-embanked landscape are generally larger in the upstream part of the model area than in the downstream part (Figures 7C and 8). This is because elevation differences decrease in downstream direction, resulting in wider and relatively lower floodplains (Figure 1). This is especially clear in the delta apex where the floodplain divides, circa 50 km upstream of the channel bifurcations (Figure 8).

Both water depth and inundation extent increase with discharge (Figure 8; Figure 9). Interestingly, the increase in inundation extent is nearly linear for extreme discharges (Figure 9). Most of the model area consists of a terraced valley, where additional surfaces flood with each step in simulated discharge, although water depths at the newly flooded locations are small. This is different in the present embanked situation, where the inundation extent ceases to increase when all areas between embankments are flooded (Figure 7A). Despite this terrace effect, the graph does not appear step wise (Figure 9). The expected incremental trend is obscured due to the large size of the area (>4,500 km²) with many small elevation differences and local terrace relief (Figure 3).

When hydraulic roughness values are raised, the inundation extent increases (Figure 9). The total spread in extent resulting from varying the roughness in all landscape classes between \( n_{\text{min}} \) and \( n_{\text{max}} \) is approximately 15%. This value is rather stable for different discharges, amounting to 15.5% for a peak discharge of 10,000 m³/s, and 14.5% for a peak discharge of 24,000 m³/s (Figure 9).

### 3.2 Flood wave propagation

Simulated peak discharges in the pre-embanked landscape remain remarkably constant along the Lower Rhine (Figure 10). Retention in upstream parts of the area is apparently small and balanced by the
discharge contribution of tributaries, indicating that nearly the complete floodplain conveys water downstream. The rise in water levels with discharge is largest in the upstream part of the study area (Figure 11), which is related to the floodplain widening in downstream direction. An increase in peak discharge from 10,000 to 30,000 m$^3$/s leads to a 1.5 to 3 m rise in maximum water levels in the river valley (profiles 1 to 3 in Figure 11), but less than 1 m in the delta (profile 6 in Figure 11). Both extent and depth of inundation exhibit slightly declining (concave down) increasing trends with discharge (Figures 9 and 11), because flow velocities increase as water depths across the floodplain rise.

In all model runs, the westward route in the delta (Waal and Nederrijn Rivers and associated floodplain) carries more water than the northward route (Ussel valley). The division ratio is approximately 2:1 for a 10,000 m$^3$/s input discharge wave but reduces to approximately 3:2 for 18,000 m$^3$/s (Figure 10). For discharges lower than 10,000 m$^3$/s (not the focus of this study), an increasingly larger share of the flood water takes the westward route. After approximately 6,000 m$^3$/s enters the model area at the start of the modelled time period in the 18,000 m$^3$/s run, only one-fifth of total discharges flows northward (Figure 10). The Q–h curve in the upstream part of the westward route (profile 4 in Figure 11; ‘Gelderse Poort’) is similar to those in the upstream (profile 5 in Figure 11; ‘Oude Ussel’) and downstream (profile 7 in Figure 11; ‘Ussel’) parts of the northward route. However, in the downstream part of the westward route (profile 6 in Figure 11; ‘Betuwe’), the Q–h curve deviates and the rise in water levels...
level with discharge is significantly lower, owing to widening of the delta plain.

3.3 | Peak discharge values

The difference between pairs of MinElev and MaxElev values is about 3 to 5 m in the Lower Rhine valley and about 1 to 2 m in the delta (Supplementary Table). The elevations of the flood level markers roughly exhibit an overall downstream gradient of 0.25 m/km measured along the valley axis (Figure 12). The data show a large amount of spread, reflecting the variety of sources and uncertainty in their relation to palaeoflood levels (Supplementary Table). Somewhat in line with the observational data, the simulated water levels for large (10,000 m³/s) and extreme (30,000 m³/s) discharges differ in general by about 2 to 3 m in the upstream part of the study area, and by about 1 to 2 m in the downstream part (Figure 12). This indicates that water levels are less sensitive to discharge in the downstream and delta parts of the study area, which is attributable to the wider floodplain in these areas.

The large spread in the observational data implies that, for a simulated discharge, some MaxElev values are already exceeded whereas

---

**FIGURE 11** Maximum water depths, averaged over each profile, for the different input discharges. The data are plotted relative to the water level reached in the 10,000 m³/s run, which is about 10 m above the channel bottom in profiles 1, 2, 3 (Lower Rhine valley), 4 and 6 (westward route in delta) and about 2 m above the valley centre in profiles 5 and 7 (northward route in delta). The profile locations are indicated in Figure 1.

**FIGURE 12** Longitudinal profile of the study area with minimum and maximum flood levels extracted from observational data (blue and orange rectangles) and modelled output for the 10,000 and 30,000 m³/s palaeoflood simulations (yellow with black and light yellow with grey data points). The flood levels in the model output were read from 10,000 randomly distributed points across the study area. NAP (or NN) is the standard vertical datum in the Netherlands and Germany; 0 m is approximately mean sea level
some MinElev values are not (Figure 12). For example, around Düsseldorf, several MaxElev values plot below the model results of the 10,000 m$^3$/s simulation. These markers mostly derive from ‘flood-free’ terrace levels (Supplementary Table), which apparently at this location may not actually represent a non-exceeded elevation for the very largest floods that occurred. On the other hand, the MinElev values at some other sites plot disproportionally high, for example near Arnhem (Figure 12), where archaeological observations provided flood markers (Supplementary Table). Despite these uncertainties, we still trust the robustness of the comparison step in our approach as it employs the complete inventory consisting of a large amount of markers obtained over a large area and derived from various sources and various types of sources.

Obviously, the larger the discharge, the more elevation markers are exceeded by simulated water levels. Ideally, for the maximum flood that occurred, all MinElev values would plot above the modelled–observed 1:1 line, and all MaxElev values below it (Figure 13). However, this result is complicated by the spread in elevation marker data and by the relatively small sensitivity of water level to discharge, especially in the downstream part of the study area, and presumably also by inaccuracies and simplifications in model input and numerical calculations. For the smallest simulated extreme discharge (10,000 m$^3$/s), the flood water levels already exceed more than half MinElev values but hardly any MaxElev values (Figure 13A). This indicates that the lower end of our simulations realistically occurred. For the largest simulated discharge (30,000 m$^3$/s), the flood water levels exceed all but a few MinElev and MaxElev values (Figure 13B), indicating that the upper end of our simulations surpasses the largest floods that actually occurred in the Late Holocene.

To determine which simulated discharge best matches the observational data, we calculated the percentages of exceeded MinElev and non-exceeded MaxElev values for each simulation and used the average of these as a measure for the goodness-of-fit between modelled and observed water levels. This combined value peaks at a discharge of 14,000 m$^3$/s and drops considerably above 18,000 m$^3$/s (Figure 14). Thus, our results suggest a millennial flood magnitude around 14,000 m$^3$/s. Simulations with larger input discharges increasingly exceed the upper limits defined by observational data, and our results strongly suggest that floods exceeding 18,000 m$^3$/s did not occur in the period under investigation.

4 | DISCUSSION

4.1 | Late Holocene flooding patterns in the Lower Rhine valley and upper delta

Comparing flood simulations for past and present landscape settings allows to quantify combined effects of landscape changes and engineering impacts on flood characteristics (e.g. Bronstert et al., 2007; Remo et al., 2009). In the past, discharge wave propagation along the undivided Lower Rhine river took circa two days (Figure 10), which is remarkably similar to present floods (Hegnauer et al., 2014; Serinaldi et al., 2018). Large amounts of water flowed over the floodplain (Figure 7, Figure 8) but did not cause a downstream decrease in peak discharge up to the delta (Figure 10), because floodplain connectivity was high (van der Meulen et al., 2020). At present, upstream flooding can reduce the discharge peak as flood water is stored in embanked areas with up to 4,000 m$^3$/s between Bonn and Wesel for extreme floods (>12,000 m$^3$/s; Hegnauer et al., 2015). Such peak flow attenuation in the main river channel reduces the risks along downstream reaches (Lammersen et al., 2002; Skublics et al., 2016; te Linde et al., 2010). Although the floodplains along the Lower Rhine are currently seen as retention basins in flood risk management, these areas hardly served as such in the past situation and instead contributed to discharge routing.

**FIGURE 13** Observed versus modelled palaeoflood levels for input discharges of (A) 10,000 m$^3$/s and (B) 30,000 m$^3$/s. The observed MinElev and MaxElev values are provided in the Supplementary Table. The symbols plotted on top of the x-axis indicate non-flooded locations. NAP (or NN) is the standard vertical datum in the Netherlands and Germany; 0 m is approximately mean sea level.
In the Lower Rhine delta, the floodplain splits into two distributive systems (Figure 1), resulting in a larger area for floodwater dispersal than in the valley. In the present situation (Figure 7A), the water diverges where the river bifurcates, close to Lobith. However, in the past situation (Figure 7C), floodwater diversion occurred considerably further upstream, close to Wesel. Current river management in the Netherlands mainly accounts for water entering the country near Lobith, but potential dyke breaches in the German–Dutch border region and consequent floodwater diversion upstream of the river bifurcation can greatly affect the discharge distribution of the Rhine river branches (Bomers et al., 2019b; Parmet et al., 2001; te Linde et al., 2011). Our palaeoflood simulations underline this often overlooked importance of the delta apex area in flood risk analyses.

4.2 Palaeoflood magnitudes in the Lower Rhine valley and upper delta

Our ‘inverse modelling’ approach has enabled us to translate geological, archaeological and historical data on past flood levels into discharge values (Figures 2 and 12–14). Comparison of the palaeoflood simulations with observational data suggests a Late Holocene to early mediaeval millennial flood magnitude lower than 18,000 m³/s, most probably around 14,000 m³/s. This value is larger than any measured flood in the instrumental record (approximately 12,000 m³/s in 1925–1926 CE). Accordingly, a peak discharge of 14,000 m³/s is our best estimate for the geologically recognised event in 785 CE (Cohen et al., 2016; Toonen, 2013) as well as for the flood cluster episode (cf. Toonen et al., 2017) that is dendrologically identified (Jansma, 2020; Sass-Klaassen & Hanraets, 2006) and that is thought to have initiated the IJssel northward deltaic branch in the sixth to seventh century CE (Cohen et al., 2009; Groothedde, 2010; Makaske et al., 2008). Note that the IJssel river channel is absent in the palaeo-DEM and model schematisation, as it had not matured before late mediaeval times (circa 1,100 CE).

In late mediaeval times, the sedimentary record reveals one flood event of greater apparent magnitude than any in the millennia before (Cohen et al., 2016; Toonen, 2013). This event is linked to the year 1374 CE, which is recognised in historical sources as a year of extreme water levels and major flooding damage along the Lower Rhine (Buisman, 1996; Gottschalk, 1971; Weikinn, 1958). A cross-sectional calculation of the 1374 CE event near Cologne resulted in an estimated peak discharge of almost 24,000 m³/s (18,800 m³/s < Q < 29,000 m³/s; Herget & Meurs, 2010). Our findings suggest that only the very lower end of that estimate may be realistic, even though the event is outside the period covered by our simulations as it post-dates embankment in downstream parts of the study area.

In the study area, embankments and other river management works are currently designed to accommodate peak discharges with recurrence times varying from hundreds to thousands of years, with differences between German and Dutch safety standards (Hegnauer et al., 2015; te Linde et al., 2011). The associated peak discharge has been set to 16,000 m³/s at Lobith after the Lower Rhine flood of 1995 (Chbab, 1996). Although this value is a subject of discussion and the newly adopted risk approach no longer identifies a single peak flow as standard (Deltacommissie, 2008; van Alphen, 2016), the design discharges in the German–Dutch border region exceed our estimate of 14,000 m³/s for the millennial flood in the Lower Rhine river system. However, this does not necessarily imply that the current flood protection standards are too high, because the safety recurrence period set for some dyke sections exceeds the time period spanned by the Late Holocene. Furthermore, changes in landscape significantly affect discharge routing (section 4.1) and likely influence the magnitude of peak discharges. In addition, our methods cannot
account for the aggravating effects of future climate change on the magnitudes and recurrence times of millennial floods.

4.3 | Additional applications of model output

The model output supports our understanding of geological and geomorphological characteristics of the study area. For example, the coarsening of the river sediment since the onset of embankment (Fringers et al., 2009) can be explained by the greater water depths and flow velocities that arise when flood waters are concentrated between embankments (Figure 7). The principle of constraining simulated discharges by observational data such as flood deposits can be applied the other way around, that is, the model output provides the potential extent of Rhine flood deposits. For example, a Late Holocene clay cover is expected at topographic levels about 1 m below the potential extent of Rhine flood deposits. For instance, a Late Holocene clay cover is expected at topographic levels about 1 m below the potential extent of Rhine flood deposits. For example, a Late Holocene clay cover is expected at topographic levels about 1 m below the potential extent of Rhine flood deposits. For instance, a Late Holocene clay cover is expected at topographic levels about 1 m below the potential extent of Rhine flood deposits. For example, a Late Holocene clay cover is expected at topographic levels about 1 m below the potential extent of Rhine flood deposits. For instance, a Late Holocene clay cover is expected at topographic levels about 1 m below the potential extent of Rhine flood deposits. For example, a Late Holocene clay cover is expected at topographic levels about 1 m below the potential extent of Rhine flood deposits. For instance, a Late Holocene clay cover is expected at topographic levels about 1 m below the potential extent of Rhine flood deposits.

Future modelling efforts may incorporate sediment transport and morphological processes to explain overbank deposition patterns in the natural landscape setting, such as oxbow-lake infilling (Ishii & Hori, 2016; Toonen et al., 2012) and natural levee formation (Johnston et al., 2019; Pierik et al., 2017). However, modelling floodplain sedimentation in detail is largely limited to local studies for present landscape situations (Middelkoop & Van Der Perk, 1998; Nicholas et al., 2006; Thonon et al., 2007).

The palaeoflood model can contribute to palaeoecological research as vegetation types are related to number of inundation days per year in lowland river areas (van Oorschot, 2017; Gouw-Boom et al., in prep). Resulting land cover reconstructions may in turn be used to iteratively improve the currently oversimplified distribution of roughness classes (Figure 4). Varying the roughness values did not introduce excessive uncertainty (Figure 9), which is similar in palaeoflood studies applying 1D models (Machado et al., 2017). More extensive sensitivity analyses (Abu-Aly et al., 2014; Bomers et al., 2019a; Papaioannou et al., 2016) may focus on changing the roughness values of individual classes independently, stochastically varying the distribution of land cover within different classes (P, DH, DL), or changing the spatial definitions of the classes (Table 1, Figure 4). Such additions may not only improve the palaeoflood model but also contribute to improved understanding of natural vegetation patterns and early historical land use in the Lower Rhine region.

Insights into past flood extents and maximum water levels are valuable to test hypotheses in archaeological and historical research, for example regarding the preservation of auxiliary forts of the Roman Limes (first century BCE to third century CE) in relative proximity to the Lower Rhine. Long-known major sites of the Roman military border including legionary camps and urban centres (e.g. Cologne, Neuss, Xanten, Nijmegen; Figure 1) and the main connecting road were founded on terrains just outside floodable areas (thus providing MaxElev values for this study; Supplementary Table), which is similar in other parts of the Roman Empire (Obrocki et al., 2020; O’Shea & Lewin, 2020). Smaller Roman military installations including fortlets and watchtowers (e.g. Haus Buergel, Till-Moyland, Herwen) were located closer to the rivers. These sites are presently conserved within a topsoil that incorporates flood deposits of post-Roman age (thus providing MinElev values for this study; Supplementary Table). Their archaeological significance and conservation are receiving increased attention in light of pending UNESCO cultural heritage status. Flooding may have been both a taphonomic factor affecting preservation and an archaeological factor affecting post-Roman land use and settlement continuity.

The upper delta and especially the Betuwe area (between the Waal and Nederrijn rivers) flooded rapidly in all palaeoflood simulations before arrival of the discharge peak. This explains the habitation patterns that occurred in the first millennium CE, which were mostly limited to former natural levees (Pierik & van Lanen, 2019). Further, it illustrates the importance of ‘zijzwendes’ (local dykes perpendicular to the river), which were constructed already prior to dykes along the river and have continued to play a role in flood mitigation (Alkema & Middelkoop, 2005; van de Ven, 1993).

4.4 | Modelling advancements in palaeoflood hydrology

Our study shows that upscaling palaeoflood reconstructions to multiple dimensions provides important information that cross-sectional or longitudinal approaches cannot resolve (Baker, 2008; Benito & Diez-Herrero, 2015; Webb & Jarrett, 2002). Models with a 2D component can account for spatial flow patterns such as velocity and direction (Horrit & Bates, 2002; Liu et al., 2015; Tayefi et al., 2007). Especially in a delta setting, the incorporation of multiple dimensions is crucial, as discharge diversions can otherwise not be properly resolved. Another aspect not incorporated by 1D models is that the maximum water level is rarely horizontal in a direction perpendicular to the main flow. The two-dimensional nature of flood flows may cause some deposits to be placed above the cross-sectional average water surface, resulting in an often-ignored inaccuracy factor in palaeoflood studies (Benito & O’Connor, 2013). This offset is significant (up to 1 m; Figure 7D) and thus relevant when comparing observational levels with simulated output (Figure 12).

A major benefit of 1D–2D modelling compared with cross-sectional 1D analyses (Herget et al., 2014; Herget & Meurs, 2010; Toonen et al., 2013) is that not only the peak discharge but also the shape of the discharge wave may be varied (Figure 6). This is important in palaeoflood research since hydrograph characteristics such as volume affect the flooding patterns and water levels (Bomers et al., 2019; Vorogushyn et al., 2010). Choosing a discharge wave requires identifying a shape that is representative for an extreme flood in the river system under study, as we did in section 2.3.

Besides the reconstructed landscape and selected hydrograph, future research may focus on the resolution and shape of the model grid (e.g. Caviedes-Voullième et al., 2012; Costabile et al., 2020). We decided on an optimal coupled grid with 1D profiles in the rivers and 2D cell rasters over the floodplains. Comparable set-ups were used to model recent flood events of the Po River in 1951 (Masoero et al., 2018) and measure the level and extent of flooding (Frings et al., 2009).
et al., 2013) and 2005 (Morales-Hernández et al., 2016), demonstrating good agreement between modelled and observed inundation extents and timing. Effects of different grid sizes and shapes may be tested in a meaningful way by further aligning the grid to geomorphological features (Figure 5). Such detailed work, similar to the suggested studies into the floodplain geology of the area and human-landscape interactions related to past flood hydrology (section 4.3), warrant abundant further research.

Due to its large size, our study area covers different geomorphological settings, which respond differently to extreme floods. An increase in discharge results in a larger water level rise in the valley than in the delta (Figure 11). This is related both to the lowering of the gradient and to differences in cross-sectional morphology (visualised in Figure 12). Because of the low sensitivity of water level to discharge, determining palaeoflood magnitudes from geological or historical observations alone, without the use of a two-dimensional model, is practically impossible in a delta setting. Our results further indicate that geological-geomorphological, archaeological and historical inferences on past flood levels show a large range in values and may contain significant errors (Figure 13), and thus cannot individually constrain palaeoflood magnitudes in alluvial settings. Instead, such analyses require a collection of many observational points, as we used in this study (Figure 1; Supplementary Table). This again implies that accurately constraining palaeoflood magnitudes in lowland settings demands two dimensions, not only the hydraulic calculations but also the other steps in the ‘inverse modelling’ approach (Figure 2), including the comparison between model output and observational data.

5 | CONCLUSION

We substantially expand upon previous palaeoflood modelling approaches by making the step to spatially continuous 1D–2D coupled modelling of past floods in the large and dynamic Lower Rhine valley and upper delta. Major effects of landscape changes, notably anthropogenic elements, on flooding patterns necessitates reconstruction in palaeoflood modelling of lowland rivers. Moreover, 2D approaches account for variations in flood water levels in directions perpendicular to the main channel, significantly reducing the risks of incorrect correlations between observed flood markers and simulated flood levels, particularly for distal floodplain sites. In addition to resolving past floods, 2D palaeoflood modelling can substantially support further geological, archaeological, historical and palaeoecological studies.

A large discharge increase in the delta causes only a small increase in water level, which complicates the comparison between flood markers and model output. We have solved this issue by providing a large dataset (96 points at 55 locations) of various observational data, which do not indicate precise maximum flood levels, but exceeded or non-exceeded elevations at critical locations just within or outside of flooding range. According to our results, the largest floods of the Lower Rhine that occurred in the Late Holocene up to mediaeval times (before the first embankments) had a peak discharge lower than 18,000 m³/s, with a best estimate of 14,000 m³/s. This millennial flood magnitude is larger than the most extreme measured discharge in the instrumental record, but lower than the magnitudes accounted for in flood risk management of the modern, engineered Rhine.

ACKNOWLEDGEMENTS

This work is part of the research programme ‘Floods of the Past, Design for the Future’ with project number 14506, which is financed by the Dutch Research Council (NWO). We presented earlier versions of our results at FLAG 2018, Liège, and INQUA 2019, Dublin, and we thank everyone with whom we discussed our research following the talks at these conferences. In October 2019, we presented our results for a mixed Dutch and German audience at the Ministry for Infrastructure and Water Management of the Netherlands (Rijkswaterstaat, Arnhem), where we had useful exchanges with river managers from both countries. We thank Marjolein Gouw-Bouman for discussions on past vegetation patterns in the study area, which helped us to determine roughness classes and values. For general input and discussions on past floods of the Lower Rhine river, we thank Willem Toonen, Jürgen Herget, Roy Dierx, Ralph Schielen and Suzanne Hulscher. Two anonymous reviewers and the ESPL editor are thanked for their constructive commentary.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the supplementary material of this article.

ORCID

Bas van der Meulen https://orcid.org/0000-0002-3590-9500
Anouk Bomers https://orcid.org/0000-0002-1560-6828
Kim M. Cohen https://orcid.org/0000-0002-0095-3990
Hans Middelkoop https://orcid.org/0000-0002-9549-292X

REFERENCES

Abu-Aly, T.R., Pasternack, G.B., Wyrick, J.R., Barker, R., Massa, D. & Johnson, T. 2014 Effects of LiDAR-derived, spatially distributed vegetation roughness on two-dimensional hydraulics in a gravel-cobble river at flows of 0.2 to 20 times bankfull. Geomorphology, 206, 468–482. https://doi.org/10.1016/j.geomorph.2013.10.017
Alkema, D. & Middelkoop, H. 2005 The influence of floodplain compartmentalization on flood risk within the Rhine–Meuse Delta. Natural Hazards, 36(1-2), 125–145. https://doi.org/10.1007/s11069-004-4545-8
Baker, V.R. 2008 Palaeoflood hydrology: Origin, progress, prospects. Geomorphology, 101(1-2), 1–13. https://doi.org/10.1016/j.geomorph.2008.05.016
Ballesteros, J.A., Bodoque, J.M., Diez-Herrero, A., Sanchez-Silva, M. & Stoffel, M. 2011 Calibration of floodplain roughness and estimation of flood discharge based on tree-ring evidence and hydraulic modelling. Journal of Hydrology, 403(1-2), 103–115. https://doi.org/10.1016/j.jhydrol.2011.03.045
Benito, G. & Diez-Herrero, A. 2015 Palaeoflood hydrology: Reconstructing rare events and extreme flood discharges. In: Schroder, J., Paron, P., &
