On communicating cost-effectiveness of flood-mitigation schemes

Onno Bokhove

Leeds Institute for Fluid Dynamics E-mail: o.bokhove@leeds.ac.uk
University of Leeds, Leeds, UK

Flood risk protection measures are designed to reduce intensity, frequency and extent of feared events. For any category of measures ranging from classical civil engineered measures to Nature-Based Solutions (NBS), being able to assess their physical and technical capacity remains a starting key issue and requirement. It is essential both to design effective solutions and also to analyze their reliability during their lifetime. For hydraulic applications, the analysis of this capacity consists in checking that proposed solutions are able to evacuate flood water discharge or to store water volume. The protocol described in this paper provides an easy-to-understand framework to assess and represent the effect of measures on the considered flood event and to compare it with their relative costs. It can therefore be considered as a basis to help decision-making within the risk management process and also as a contribution to the analysis of the safety and reliability of planned measures. The protocol enables rapid *a priori*, as well as thorough *a posteriori*, comparisons to be made of the efficacy of various flood-mitigation options and scenarios. We have considered a concept called "dynamic flood-excess volume" (dFEV or FEV) and revisited it in a three-panel graph comprised of the (measured) in-situ river-level as function of time, the rating curve and the hydrograph, including critical flooding thresholds and error estimates. FEV is the amount of water in a river system that cannot be contained by existing flood defences. The new tool deliberately eschews equations and scientific jargon and instead uses a graphical display with FEV displayed as a (dynamic) hypothetical square lake two metres deep. This square-lake graphic is overlaid with the various mitigation measures necessary to capture the floodwaters and how much each option will cost. The tool is designed to help both the public and policymakers grasp the headline options and trade-offs inherent in flood-mitigation schemes. It has already led to better understanding and decision-making regarding flood defences in the UK, Slovenia and France, particularly where a number of alternatives are being considered. Three realistic cases –from the UK, Slovenia and France– will be reviewed, including insights on dealing with uncertainty and on the communication of multiple benefits of Nature-Based Solutions, followed by a Socratic-method dialogue.

Keywords: flood-mitigation assessment, cost-effectiveness analysis, decision-making, flood-excess volume

1. Introduction

The Wetropolis flood demonstrator is a portable, live set-up visualising to a general audience what a return period is for extreme flooding and rainfall events (Bokhove et al., 2020). Given that one cannot wait on average for a $1:100$years Wetropolis flood to occur or, alternatively, wait for one with an annual exceedance probability of $AEP = 1\%$, time and space have been scaled down to a 10s "Wetropolis" day and a $1.2 \times 1.2\text{m}^2$ river-moor landscape, in which an extreme flood is designed to have a $6{:}06\text{min}$ return period. Rainfall is supplied randomly at two locations, none or both, in four rainfall amounts via an adapted Galtonboard in which the trajectory of a steel ball determines the outcome every 10s. One of these $4 \times 4 = 16$ combinations then leads to extreme flooding of a miniature river in a conceptual city with a Wetropolis-daily exceedance probability of $DEP = 7/256$ –giving a return period of $256 \times 10\text{s}/7 = 6{:}06\text{min}$. The set-up has been appreciated by the public, drawn the attention of flood professionals and inspired us to define a new protocol for assessing and communicating flood-mitigation plans and their cost-effectiveness. The protocol stimulates exploration and discussion of flood-mitigation scenarios with the aim to facilitate improved decision-making. In what follows we review it from various perspectives.

Both Wetropolis and the protocol were in part triggered by Leeds' Boxing Day floods of 2015, during which the Yorkshire area in the UK experienced widespread flooding with substantial damage to property. These floods led to the development of new flood-mitigation plans by Leeds City Council and works aimed to mitigate such flood damage have been accelerated since 2015. The analysis of these plans led to a generalisation into a new graphical tool or protocol (Bokhove et al., 2020), which has been used to create and/or assess other flood-mitigation plans across the world.

Proceedings of the 31st European Safety and Reliability Conference
Edited by Bruno Castanier, Marko Cepin, David Bigaud, and Christophe Berenguer

ISBN: 978-981-18-2016-8; doi:10.3850/978-981-18-2016-8_134-cd

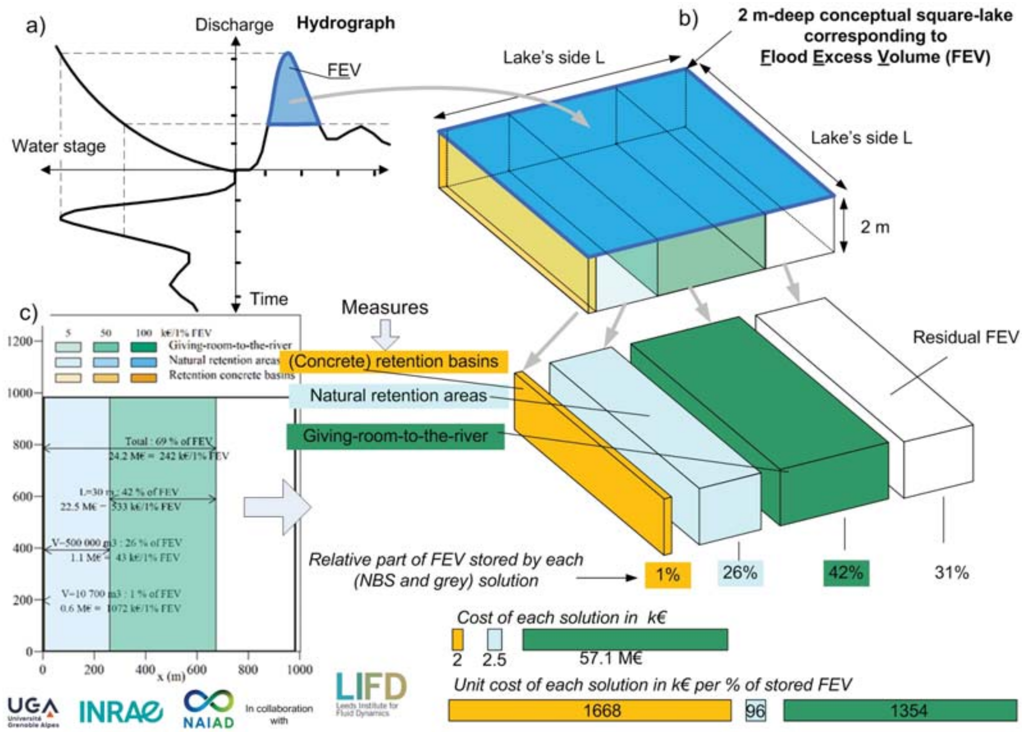

Fig. 1. Schematic of square-lake cost-effectiveness based on FEV: a) three-panel graph resulting in a hydrograph with flood-excess volume (FEV) indicated via shading; b) FEV is expressed as a two-metre deep square lake, then partioned into segments representing each flood-mitigation measure; c) finally resulting in the square-lake graphs with costings and information of each measure, as well as total costs, indicated. Figure courtesy: J.-M. Tacnet.

1.1. *Schematic summary*

The entire protocol is summarised succinctly in Fig. 1. It is important to stress that the protocol is what meteorologists could call a diagnostic tool and decision-makers an executive summary. It is not a predictive method in the sense of being a solution of a relevant system of (coupled) ordinary and partial differential equations that aim to forecast precipitation and river (flood) levels. Rather, the protocol starts with data of the river level or discharge of a river flood at a critical location along the river as function of time. These data can be acquired via measurements taken, simulations performed, a combination, or preferably even an ensemble of such measurements or simulations. For simplicity, we will start with data from one particular flood event. In Fig. 1a), a three-panel graph shows such data: in the bottom-left panel river level $h = h(t)$ (horizontal axis going left) is displayed versus time t (vertical axis going down); in the top-left panel the rating curve displays the discharge $Q(h) = Q(h(t)) = Q(t)$ (vertical axis going upwards) versus river level h (same horizontal axis going left); in the top-right panel we find the discharge (same vertical axis going upwards) versus time (new horizontal axis going right). The reason for this new three-panel amalgamation is that generally river levels are the basic frequently-measured data, including errors therein, with the hydrograph a derivation based on combining river levels with a rating curve, with larger errors therein. When the data arise from simulations, river levels and discharge rates are available in tandem with the rating curve a resultant outcome, which can be compared with in-situ velocity-river-level measurements. Flooding can be defined to occur when the river level exceeds a certain threshold h_T, in which h_T can be chosen depending on a threshold for minor, intermediate or major flooding. Given the rating curve and threshold h_T a matching threshold Q_T can be found (as indicated graphically), including errors therein. Flood-excess volume (FEV) is defined as the integrated discharge volume V_e with $Q(t) > Q_T$, a volume coloured in the top-right panel of Fig. 1a). It is the volume that caused the flood damage. Note that V_e is an effective or dynamic volume, dFEV, and not a static volume since the river water is flowing with generally considerable speed through the river cross-section involved. When we divide V_e by a typical, human-size scale of $d = 2\mathrm{m}$ and take its square root, we have defined a square lake of two metres depth with lake side $L = \sqrt{V_e/d}$. Such a square lake is displayed in Fig. 1b). It is a dynamic lake given that it is filled with the flowing waters leading to the flood-excess volume V_e over the flood duration. For Leeds' Boxing Day flood of 2015, $V_e = (9.34 \pm 1.50)\mathrm{Mm}^3$ and the lake side $L = 2161\mathrm{m}$, while for the River Brague 2015 flood $V_e = (0.488 \pm 0.311)\mathrm{Mm}^3$ and $L = 494\mathrm{m}$ (Bokhove et al., 2019). The size L of this square lake can be compared with the dimensions of the river valley or catchment concerned in order to assess roughly whether certain flood-mitigation measures could fit within the catchment landscape. Each measure can be represented as a subvolume of this square lake, a subarea when viewed from above. In Fig. 1b) the lake is partitioned in rectangular blocks, indicated with different colours, each representing the (relative) portion of the measure. In the sketch, (concrete) retention basins (orange), natural retention areas (light blue), giving-room-to-the-river (GRR –green) and residual FEV (white) are exemplified. The residual FEV can for example be accommodated by building (higher) flood-defense walls. Finally, since the depth of the lake is generally negligible compared with its length, in Fig. 1c), the square-lake is viewed from above. For each measure, the costs, costs per percent mitigation and total costs can be indicated around the double-headed arrows. In the decision-making process, various scenarios can be considered via such square-lake graphs and compared on their relative merits. In addition, part of the decision-making process can include the exploration and design of alternative scenarios.

The use of our diagnostic tool can be split into roughly three categories, as follows:

(i) as consistency check of flood-mitigation plans, in order to understand the plans and place them into a data-based context;
(ii) as a priori investigation to scope out flood-mitigation scenarios using measured or simulated flood data to explore ideas and stimulate discussion; and,
(iii) as a posteriori executive summary to disseminate flood-mitigation plans and as consistency check of expert engineering simulations underpinning such plans.

Three realistic cases will be reviewed: (i) the Boxing Day 2015 flood of the River Aire in the UK and associated mitigation plans (Bokhove et al., 2020); (ii) scoping of Nature Based Solutions (NBS) for flood mitigation and ecological enhancement for the River Glinščica in the Ljubljana area of Slovenia, where focus was placed on communication and stakeholder involvement (Piton et al., 2018); and, (iii) setting up and communicating flood-mitigation plans to deal with River Brague floods in France (Bokhove et al., 2019), plans later augmented with extensive hydraulic simulations.

1.2. *Effectiveness and reliability*

To deal with effectiveness and reliability assessment in flood mitigation in the FEV approach, we will address the following questions in each of the reviewed cases.

(a) What is the connection to assets, infrastructures, protection works management and decision-making issues?
(b) What is the application context?
(c) How does the approach contribute to effectiveness and reliability assessment?
(d) Which feedbacks can be used to support and justify the FEV-approach?
(e) What is the link with real-case issues?
(f) How can the approach be used to assess the physical capacity of protection works?

In the remainder of the paper, reviews of the three cases will be given in §2, 3 and 4, followed by a discussion in §5.

2. Understanding the River Aire Boxing Day flood of 2015, UK

The River Aire in Yorkshire, UK, flooded on Boxing Day 2015 (26-12-2015) after $48\mathrm{hrs}$ hours of extreme rainfall. The entire Yorkshire area experienced extreme flooding with record rainfall and river levels, some of which record levels were subsequently eclipsed during Storm Ciara in 2019. The River Aire is during normal flow partitioned in various sections between man-made weirs and within the Leeds' boundaries flow is subcritical between weirs, except during low flow when there are also some minor natural rapids. Across weirs there is a transition to supercritical flow and back to subcritical flow below the weir. During flooding, the Dark Arches' weir under Leeds' railway station does not submerge, given the large height drop, and it controls the subcritical flow for a few miles upstream. Widespread flooding on Boxing Day 2015 occurred in the Kirkstall industrial estate upstream of this Dark Arches weir as well as downstream thereof, in the centre of Leeds. Due to familiarity with the area, the study by Bokhove et al. (2020) has focused on the Kirkstall estate upstream of the Dark Arches for which the Armley river gauge is relevant. The Boxing Day flood of 2015 was estimated to have a return period of approximately $1{:}250$ years or an $AEP \approx 0.4\%$. The council in Leeds made flood-mitigation plans for protection against floods of an $AEP > 0.5\%$ in the period 2016-2019 involving a series of measures including: raised flood defense walls, flood-plain storage by potentially stowing up waters higher than before using a moveable weir, with potential sites at various locations near and further upstream of Leeds, Natural Flood Managment (NFM) in the upper Aire catchment, and removal of obstacles as well as local widening of the river bed (collected under the heading giving-room-to-the-river –GRR).

2.1. *Cost-effectiveness analysis*

The three-panel plot of the stage-discharge relations of the Armley river gauge around its peak on 27-12-2015 is provided in Fig. 2. Based on local observations of high river levels mid December 2015 and photographic evidence of flood levels on 27-12-2015 as well as in February 2019 during Storm Ciara near a business on the Kirkstall estate, a suitable threshold emerged as $h_T \in [3.9, 4.2]\mathrm{m}$. However, knowing that flooding started earlier further upstream in a lower lying area, we took $h_T = 3.9\mathrm{m}$. Via the rating curve provided by the Environment Agency (EA), this yields a critical discharge threshold of $Q_T = 219.09\mathrm{m}^3/s$. Errors in the rating curves were estimated by the EA to be circa 5.5%, leading to a dFEV of $V_e = (9.34 \pm 1.50)\mathrm{Mm}^3$. The side length of the matching two-metres deep square lake is then $L = 2161\mathrm{m}$.

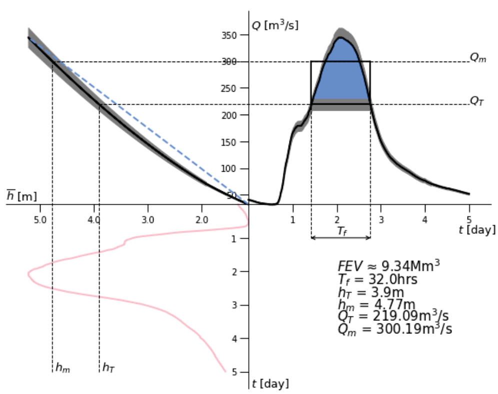

Fig. 2. Three-panel hydrograph for the 2015 River Aire flood at the Armley gauge station. The threshold river level $h_T = 3.9\mathrm{m}$, revealed by a black dashed line; the corresponding threshold flow rate $Q_T = 219.09\mathrm{m}^3/s$, obtained from the rating curve (with circa 5.5% error –grey shading). The flood duration is 32 hours and river-level measurements are made every 15min. The total volume of water that attributes to the flood (FEV) is $V_e = 9.34\mathrm{Mm}^3$, visualised by the blue shaded area. Extension of a graph in Bokhove et al. (2020) by Zheming Zhang, `https://github.com/Flood-Excess-Volume`

The square-lake graphs allow us to graphically display the proportion of flood-excess volume as well as the cost-effectiveness of each measure. In Fig. 3, such a graph for scenario "S1", corresponding closely to the approved flood-mitigation plans by Leeds, is shown. The associated costs displayed are partly estimates and partly based on some information provided by Leeds City Council. Moreover, we took the FEV of the Boxing Day event as proxy for the $1 : 200$ years protection offered in the plans, effectively taking a somewhat lower value of h_T. In the end, higher walls offer circa 85% of the required protection, 15% is obtained by both GRR and the dynamic flood-storage site in Calverley a few miles upstream of Leeds enhanced via dynamic weir. GRR and the walls are displayed with certainty as rectangular areas in the square-lake graph, with the flood-plain storage using the lower bound together with walls and GRR covering 100% of the square lake. The weir needs to be lowered optimally to reach the maximum available flood-plain storage with the lower bound an illustrative guess. The problem of optimal flood storage is a challenging control problem. Any flood-storage volume above this lower bound is taken as extra storage beyond the target 100%. In the graphical display this leads to quadrilateral or triangular shapes in the square-lake graph, displaying the uncertainty, with the least certain case at the bottom face of the lake and the most certain case at the (extended) top face of the lake. Hence, NFM with additional minimal protection offered by introducing about 85 beaver colonies in the upper Aire catchment is considered entirely as extra and uncertain storage beyond 100%. Here NFM consists of water storage behind (numerous) leaky woody debris dams and water retainment by extensive tree planting, planned in the upper Aire catchment. Note that the FEV-methodology lends itself for comparisons between different scenarios, i.e. that leads to various square-lake graphs for different measures each with different proportions, which can easily be compared visually, see Bokhove et al. (2020).

The FEV-approach in this case study appears to be spatially zero dimensional since it is based on measurements in time at one location. However, the river gauge value is representative for the relevant stretch of river because the threshold level h_T in those measurements pertains to a critical flooding level over that river segment during extreme flooding. That choice was based on several local observations that led to the decision that $h_T = 3.9\mathrm{m}$ on that gauge was representative for the Kirkstall industrial estate.

For the Aire River case study, the following questions raised in §1.2 are addressed as follows: (b) The application concerned both public empowerment by facilitating communication to the public with clear summary graphics and analysis of Leeds' flood-mitigation plans as a pre-

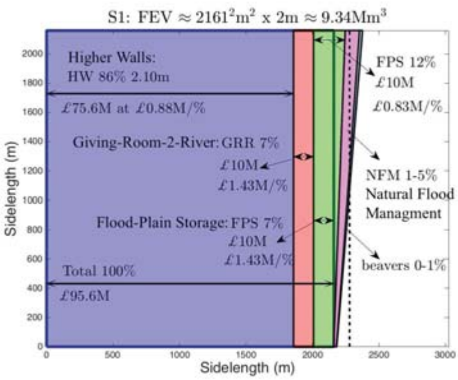

Fig. 3. The flood-excess volume $V_e = 9.34\mathrm{Mm}^3$ of the Leeds' Boxing Day 2015 floods is represented as a square lake (view from above) of two metres deep and 2161 metres side length. Each coloured segment represents the portion of the effective or dynamic square lake absorbed by each mitigation measure. Cost and cost per percentage covered per mitigation measure are displayed around the relevant double-sided arrow, in addition to the overall costs. Partially uncertain mitigation is offered by flood-plain storage (FPS, green), Natural Flood Management (NFM, pink) and 85 beaver colonies (the small silver sliver). For large floods, the contribution of NFM and 85 beaver colonies tends to be small-to-minute as well as uncertain: associated areas are small relative to areas of other measures, such as higher walls (HW, dark purple) and river-bed widening (GRR, red). Adapted from Bokhove et al. (2020) and OBokhove et al. (2020).

liminary consistency check. Some inconsistencies were found in Leeds' 2017 plans and duly reported to the authorities prior to publication of the 2018 report predating Bokhove et al. (2020). The FEV cost-effectiveness analysis enabled to reveal inconsistencies as a reliability assessment (c) for a real case (e). Finally, the physical capacity of the protection works (f) is expressed as percentage of the overall FEV. Higher walls can be related to a (dynamic) volume but are best viewed as the remaining portion of the FEV not covered by other measures, which effectively leads to choosing an increased threshold level h_T. Note that 100% flood protection by only higher walls then corresponds to bringing h_T to the maximum flood level involved, even though we can expect (minor) changes in the rating curve. GRR is more complex and directly relates to a change in the hydrograph. In general, in an event for which one has done systematic simulations, volumetric effects can be discerned to some extent as fraction of the original FEV by comparing or subtracting various hydrographs for situations with or without mitigation measures. Of course, some caution is required since combinations of mitigation measures can reinforce or weaken the effects of other measures.

3. A priori FEV analysis: assessment of NBS for River Glinščica, Slovenia

Flood management and ecological integrity were the primary goals in a catchment management investigation of the River Glinščica in the Ljubljana area of Slovenia. A demonstration of participatory catchment management with stakeholders was undertaken for NBS as most suitable solution to reach these primary goals, as part of the EU-funded project NAIAD (NAture Insurance value: Assessment and Demonstration, see Piton et al. (2018)). Predictions from coupled hydrological and hydraulic (groundwater and pluvial as well as fluvial) models were undertaken to analyse the impact of NBS effectiveness in producing benefits and co-benefits. The design process involved the identification of societal challenges connected to water-related risks, subsequent selection and ranking of risk-management goals, wherein the stakeholders were involved to find the most suitable NBS. An integrated effectiveness assessment then triggered a feedback loop till a satisfactory design was reached. The benefit was identified as reduction of flood extent with as co-benefits: reduction of infrastructural damages, improvement of ecosystems and biodiversity, improved safety of the local community and improved social value of the ecosystem. During the EU project, the coupling between the hydrological and hydraulic models was not feasible due to the multiple combinations of NBS to be tested and the FEV cost-effectiveness diagnostic was therefore used instead as a first and minimal data-based "modelling" approach. Meanwhile, the computational hydraulic analysis performed was used to determine representative river-level data in current conditions, a rating curve and hydrograph data, at a suitable location. In addition, the FEV-approach was more intuitive for stakeholder engagement.

Input data of the one-dimensional hydrological-hydraulic analysis were $5 \times 5\mathrm{m}^2$ LIDAR data, rainfall series, discharge series and measurements from field work. Unsteady flow simulations yielded the FEV at an identified critical location where a threshold $h_T = 2.39\mathrm{m}$ was established, in addition providing stage-discharge relationships based on a range of events, see Fig. (4) –top. Four flood-mitigation NBS were considered: urban wet retention areas at 9%, green roofs at 10%, dry retention areas at 66% and opening of flood plains at 16%, in total 101% with round-off errors, as visualised from left-to-right in Fig. (4) –bottom. Volume estimates were based on a 67% efficiency estimate except for the 80% efficiency estimate used for the green roofs. In conclusion, the FEV-

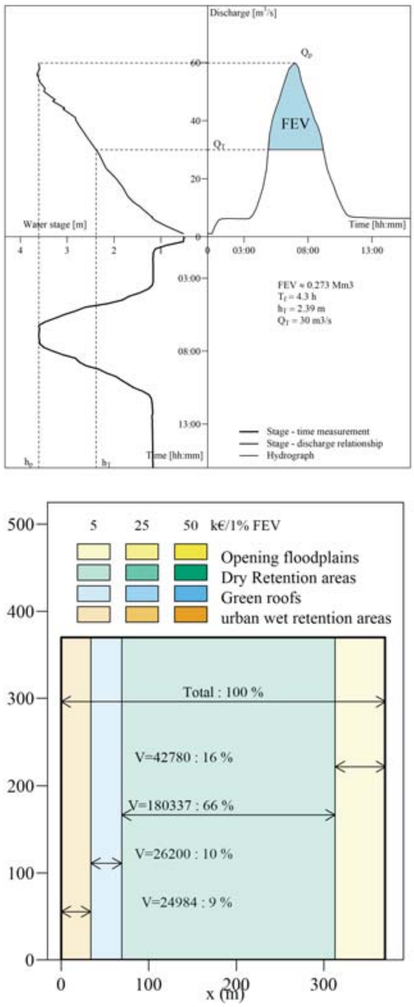

Fig. 4. Top: Integrated hydrograph of the River Glinščica at a critical location. The threshold river level $h_T = 2.39\mathrm{m}$, one of the dashed lines, and the corresponding threshold flow rate $Q_T = 30\mathrm{m}^3/s$, graphically obtained from the rating curve via that dashed line. The flood duration is 4.3hours. The total $V_e = 0.273\mathrm{Mm}^3$, visualised by the blue-shaded area. Bottom: Square-lake graphs with costings in € and the colour-shading intensity corresponding to cost-per-percent for the values of $(5, 25)$k€ and 50k€. Figure courtesy: A. Pagano and P. Pengal.

based approach provided a straightforward and effective assessment of the combined effects of NBS, which clear graphical outputs were appreciated and understood by the stakeholders involved. More dynamic output can augment the approach or the approach can be a summary of extensive a-posteriori dynamical analyses, as considered in the next section.

The distinction between benefits and co-benefits is not always so clear as in the case for the Glinščica catchment. In Perosa et al. (2021), flood protection was initially thought to be the dominant benefit but the study made clear that the main economic benefit was the recreational value of the NBS studied, which value was originally deemed a co-benefit, in case-study areas along the River Danube. Relevative merits of NBS were clearly expressed in histograms of the economic value of all benefits (Figs. 4 and 5 in Perosa et al. (2021)).

Finally, with respect to the questions raised in §1.2, the following: (a) in terms of infrastructure, the urban wet retention areas required $38.000\mathrm{m}^2$ of urban park area, green roofs concerned adaptation of 30% of the roofs in the area and dry retention areas required $530.000\mathrm{m}^2$ of land area; the interactive stakeholder involvement was aimed at improving participation and hence decision-making; (b) the FEV-approach was part of a larger scoping application on the use of NBS in the municipality; and, (c,d) part of the larger NBS-design process, in which the FEV approach was used, included feedback loops and integrated effectiveness assessment (Pagano et al., 2019).

4. A posteriori FEV analysis: flood protection for River Brague, France

The extreme flood event of October 3rd 2015 saw the River Brague burst it banks after torrential rainfall. The river flows through a hilly catchment of $69\mathrm{km}^2$ with rural and suburban developments in the French Riviera into the Mediterranean Sea. In the catchment, there were four casualties and over 200M€ of insured damages accrued. Both a zero-dimensional hydraulic analysis based on the FEV-approach was employed as well as two-dimensional hydraulic simulations to investigate the efficacy of NBS-type measures including natural retention areas, GRR and retention concrete basins. Cumulative retention volumes surpassed $1\mathrm{Mm}^3$ and GRR was employed by bank lowering and widening over $30\mathrm{m}$. Based on data of hydrographs across the catchment following hydraulic simulations, the effects of GRR can be explained graphically in the three-panel graph, see Fig. 5. Solid lines therein display the stage-time relationship (bottom left), the current rating curve based on an average of hydraulic simulations across the catchment, and the discharge-time relationship. Since the rating curve after the GRR-procedure moves up, leading to more through-flow for the same river level, with the same threshold h_T, the discharge threshold $Q_{T,GRR}$ is raised leading to a reduced dFEV (dashed lines). Key is that the discharge curve remains the same but that the rating curve is altered leading to higher threshold discharge $Q_{T,GRR} > Q_T$ and lower peak depth $h_{P,GRR} < h_P$. A corresponding square-lake graph with costings in € and lighter colours for cheaper-per-percentages measures is the one in Fig. 1c). The three measures cover 69% of the dFEV, with concrete basins at 1% represented by the thin sliver, natural retention areas at 26% being the cheapest per percent and GRR at 42%. The remaining 31% unprotected dFEV requires additional measures for the worst-case design event of $1:500$ years or an $AEP = 0.2\%$. In Piton et al. (2018) and Bokhove et al.

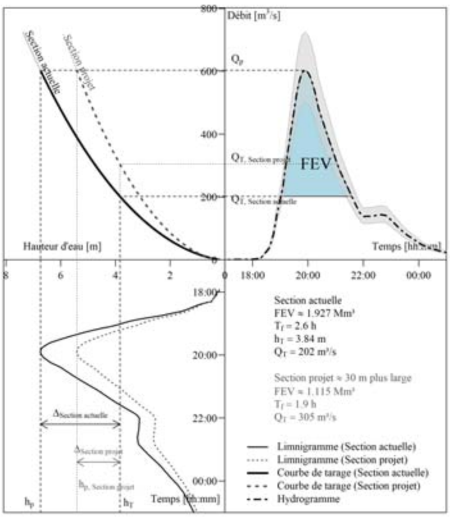

Fig. 5. Three-panel graph of the 2015 flood of the River Brague, France, solid-line curves, as well as a GRR-modified case, dashed curves. The threshold river level $h_T = h_{T,GRR} = 3.84\text{m}$, and the corresponding threshold flow rate $Q_T = 202\text{m}^3/s$, graphically obtained from the rating curve via that dashed line. The flood duration is 2.6hours. The total dFEV is $V_e = 1.93\text{Mm}^3$. Since the rating curve changes due to GRR, the discharge lowers and hence $V_{e,GRR} = 1.12\text{Mm}^3$ is lower. Figure courtesy Piton et al. (2018).

(2019), the two-dimensional hydraulic analysis was still in progress. The two-dimensional hydraulic analysis results followed later and were quite consistent with the FEV analysis found in Bokhove et al. (2019). However, that hydraulic analysis revealed two bottleneck sections (concerning highway culverts and a road bridge near the shore), which were driving the flood levels in two major areas of the floodplain. Backwater stowage effects played a major role here. The stage-discharge relationship used should in such case be the most critical one in the bottleneck sections, or both, or the bottleneck sections should be removed. The hydrograph data used in Fig. 5 were based on a (weighted) sum of all simulated hydrographs coming from various river branches, and neglected flood-plain buffering. In contrast, in Bokhove et al. (2019) the FEV-approach was used further upstream based on a field survey of the upstream hilly part of the catchment. We note that the cost-effectiveness analysis based on volumetric effects has the advantage that measures based on retention provide valid protection for the river downstream of these measures while higher walls and GRR only give local protection. Finally, torrential floods introduce a new parameter linking to sediment transport and, additionally, the Brague project is still in the design stage in which a final solution has not been chosen (as of June 2021).

5. Discussion

We observed already that the FEV-approach is formally zero dimensional, mostly based on time-series data from one location, while in practice it either covers a (flooded) river stretch for which the data at that critical location are relevant or it can concern an average of hydrographs across a (sub)catchment. Expert judgement based on local information and observations is often used in choosing the threshold h_T water level such that the analysis gains some spatial coverage. In the a posteriori case, extensive (ensemble) simulations based on expert engineering of hydrology and hydraulics can be used to created multivariate data sets in space, to which the FEV-cost-effectiveness analysis can be applied. That will lead to more statistical information and more complicated graphics, including averaged information, the latter which can be used to meaningfully communicate with stakeholders and the general public. Limitations of the FEV-approach have been discussed extensively in Bokhove et al. (2019, 2020); Piton et al. (2018). New elements in this discussion are presented next in a Socratic-method dialogue, based on questions raised.

How can a a modular way to create flood protection be made by using FEV rather than (stage and discharge) thresholds?

Both water level h_T and discharge Q_T thresholds have been used since temporal integration above a flow or water-level threshold (related via the rating curve) yields the FEV. FEV is not a static volume but has a dynamic origin –dFEV. It is an effective volume. All flood-mitigation measures used in the square-lake representations considered, for various international scenarios and rivers, were based on such dynamic flows.

FEV seems limited because it focusses on just one particular event?

In the road-map in Bokhove et al. (2020), we advocate the use of ensemble predictions for flood events leading to ensemble dFEVs and an ensemble of flood-mitigation scenarios. To exemplify the methodology, focus was deliberately first placed on particular worst-case events.

The FEV-approach focuses on storage-based solutions, rather than a suite of potential flood management measures, e.g. a suite improving conveyance, NFM, infiltration improvements, etc.

For various scenarios and river floods, the FEV approach included several flood-mitigation measures such as raised walls, GRR, flood-plain storage (using controllable weirs), NFM (including leaky dams, trees, beavers), and dynamic drawdown of reservoirs. Hence, consideration has been given to a suite of potential flood management measures, yielding a variety of segmented square-lake graphs. We have therein highlighted approaches in which each mitigation measure is less than 50% of the total FEV. Stakeholders have appreciated that expressing each mitigation measure as a volume conveys a sense of size of that measure, relative to another and the river catchment.

For extreme events, FEVs can change significantly on duration, storm coverage and rainfall intensity. How can this be factored in?

We advocate ensemble forecasting based on a variety of events. However, by mitigating against a rare event, either in terms of an AEP based on peak water level or FEV, one automatically also mitigates against an event with a (similar) high water level yet shorter duration. A mixture of GRR, floodwater storage, use of reservoirs and NFM could optimally mitigate against both types of events without large wall height increases.

Given the variety of events falling under one AEP, using FEV as a descriptor would change depending on the hydrological nature of the event. This would diffuse the straightforward message communicated via FEV since the emerging uncertainty will be more complicated to convey?

We advocate the use of both FEV and water levels to calculate return periods or the AEPs, leading to (slightly) different classifications. In combination with the use of ensemble forecasting this does cover a range of flooding events and scenarios. In the case of ensemble forecasting, the averaged results will still be similar to the ones displayed here but, additionally, one can convey the uncertainty arising from such ensemble predictions via the slanted lines already included in Fig. 3 and cases in Bokhove et al. (2019, 2020); see also Bokhove et al. (2018).

The FEV methodology is by itself and alone not a proper safety and reliability analysis approach. However, this kind of approach is an essential input in the whole chain. It indeed provides valuable inputs in global approaches dedicated to multifactorial analysis of flood protection measures' effectiveness, including technical and physical aspects but also environmental, social, human and economic features using complementary methods, such as proposed in the NAIAD H2020 project.

Acknowledgement

Input from Alessandro Pagano, Polona Pengal, Guillaume Piton and Jean-Marc Tacnet is much appreciated.

References

Bokhove, O., T. Hicks, W. Zweers, and T. Kent (2020). Wetropolis extreme rainfall and flood demonstrator: from mathematical design to outreach and research. *Hydrology and Earth System Sci. 24*, 2483–2503.

Bokhove, O., T. Kent, and M. Kelmanson (2018). On using flood-excess volume to assess natural flood management, exemplified for extreme 2007 and 2015 floods in Yorkshire. `https://eartharxiv.org/repository/view/1282/`.

Bokhove, O., T. Kent, M. Kelmanson, G. Piton, and J.-M. Tacnet (2019). Communicating (nature-based) solutions using flood-excess volume for 3 UK and French river floods. *River Res. Appl. 35*, 1402–1414.

Bokhove, O., T. Kent, M. Kelmanson, G. Piton, and J.-M. Tacnet (2020). A cost-effectiveness protocol for flood-mitigation plans based on Leeds' Boxing Day 2015 floods. *Water 12*, 652.

OBokhove, M. Kelmanson, and T. Kent (2020). A new tool for communicating cost-effectiveness of flood-mitigation schemes. Technical report, UK Government inquiry on flooding. `https://committees.parliament.uk/writtenevidence/9641/pdf/`.

Pagano, A., I. Pluchinotta, P. Pengal, B. Cokan, and R. Giordano (2019). Engaging stakeholders in the assessment of NBS effectiveness in flood risk reduction: a participatory system dynamics model for benefits and co-benefits evaluation. *Science of Total Envir. 690*, 543–555.

Perosa, F., M. Gelhaus, V. Zwirglmaier, L. Airas-Rodriguez, A. Zingraff-Hamed, B. Cyffka, and M. Disse (2021). Integrated valuation of nature-based solutions using TESSA: three floodplain restoration studies in the Danube catchment. *Sustainability 13*, 1482.

Piton, G., S. Dupire, P. Arnaud, A. Mas, R. Marchal, D. Moncoulon, T. Curt, and J. Tacnet (2018). DEL6.2 From hazards to risk: models for the DEMOs –Part 3: France: Brague catchment DEMO. Technical report, NAIAD H2020 project GA nº 730497. `http://naiad2020.eu/wp-content/uploads/2019/02/D6.2_REV_FINAL.pdf`.

Piton, G., A. Pagano, R. Basile, B. Cokan, and S. Lesjak (2018). DEL6.2 From hazards to risk: models for the DEMOs –Part 5: Slovenia –Glinščica catchment DEMO. Technical report, NAIAD H2020 project GA nº 730497. `http://naiad2020.eu/wp-content/uploads/2019/02/D6.2_REV_FINAL.pdf`.