Journal of Hydrology 610 (2022) 127958

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Improving rating curve accuracy by incorporating water balance closure at river bifurcations

M.R.A. Gensen [a,c,*], J.J. Warmink [a], K.D. Berends [b], F. Huthoff [a,c], S.J.M.H. Hulscher [a]

[a] Department of Marine and Fluvial Systems, University of Twente, P.O. Box 217, 7500 AE, Enschede, The Netherlands
[b] Department of River Dynamics and Inland Navigation, Deltares, Boussinesqweg 1, 2629 HV, Delft, The Netherlands
[c] HKV, P.O. Box 2120, 8203 AC, Lelystad, The Netherlands

ARTICLE INFO

Keywords:
Rating curves
River bifurcation
Water balance
Uncertainty analysis
Rhine river

ABSTRACT

Accurate discharge records are essential for flood frequency analyses, hydraulic model calibration and flood forecasting. Discharge records are often obtained via a transformation of water levels to discharges using a rating curve. For accurate rating curves, a physical basis is important, particularly in the extrapolation domain towards extreme discharges. In this study, physical processes and constraints are incorporated in a rating curve model: water balance closure at a bifurcation and bed level degradation. The aim is to assess the effect of incorporating these physical processes and constraints for rating curves at two bifurcations of the Rhine river in the Netherlands. Intermittent gaugings are available for a 31 year period at these bifurcations. Bayesian inference and Markov Chain Monte Carlo sampling is used to estimate the posterior distributions of the rating curves. If rating curves are constructed independently, they show a large water balance error at bifurcations of up to 10%. Incorporating bed level degradation is required for accurate rating curves as it reduces the residual errors by up to 50%. If explicitly accounting for water balance closure, the water balance error can be reduced to 1%, while residual errors remain equally small. As water balance closure is a physical constraint at a river bifurcation, the rating curves that account for water balance closure are expected to be more physically realistic. Therefore, it is recommended to specifically gauge at river bifurcations and confluences, such that the constraint of water balance closure can be used to improve the accuracy of rating curves and discharge records.

1. Introduction

Discharge records in streams and rivers are essential for water resources management of both low flow and flood conditions. At a river bifurcation, knowledge on the distribution of discharges over the branches is crucial (Dong et al., 2020; Gensen et al., 2020). However, continuously measuring the discharge is not easily done. Therefore, discharges are often derived by converting continuously measured water levels at gauging stations to discharges using a rating curve (e.g. ISO 1100, 2010; Rantz, 1982). A rating curve is a relationship between water levels and discharges at a specific cross-section of the stream or the river, which is established using earlier gaugings of discharges and water levels. These gauged discharges are often based on measured flow velocities and cross-sectional geometry.

Rating curve errors can lead to significant errors in the analyses in which the discharge records are used, such as flood frequency analysis (Lang et al., 2010; Steinbakk et al., 2016), hydraulic or hydrological model calibration (Domeneghetti et al., 2012; Peña Arancibia et al., 2014; Sikorska and Renard, 2017) and flood forecasting (Ocio et al., 2017). The largest errors in derived discharges arise when the rating curve is used during extremely high flow conditions (Domeneghetti et al., 2012; Pappenberger et al., 2006). In this domain, discharge gaugings are scarce or even unavailable and the discharge is thus derived by mere extrapolation of the rating curve. Rating curve uncertainty in the lower domain is amplified in the extrapolation domain, possibly leading to very large errors (Di Baldassarre et al., 2009). Still, the extrapolation of rating curves is often required to obtain knowledge on the flood conditions in a river (Pappenberger et al., 2006).

Physical constraints can be imposed to increase the accuracy of rating curves, especially in the extrapolation domain. In literature, several methods have shown to enable using additional physical basis to construct more accurate rating curves, through e.g. using knowledge on hydraulic controls at structures (e.g. weirs and dams, Le Coz et al., 2014), using physically realistic parameters in rating curve formulations (e.g. Le Coz et al., 2014) and using hydraulic modelling (Di Baldassarre and Claps, 2011; Lang et al., 2010). Furthermore, knowledge on the

* Corresponding author.

https://doi.org/10.1016/j.jhydrol.2022.127958
Received 12 August 2021; Received in revised form 13 May 2022; Accepted 17 May 2022
Available online 27 May 2022

physical explanations for rating curve changes can be used to improve rating curve accuracy. Especially, if a long record of discharge gaugings is available, rating curve changes due to non-stationarities, are a significant source of errors (Pappenberger et al., 2006) and should therefore be incorporated if possible. Often, rating curve changes are the result of morphological development of the river bed (i.e. bed level degradation or aggradation; e.g. Mansanarez et al., 2019), but may also be the result of hydraulic roughness changes (e.g. Gensen et al., 2020; Perret et al., 2021) and human intervention (Berends et al., 2019).

Water balance considerations is sometimes used in hydrological studies for more accurate rainfall-runoff estimates (Sebok et al., 2016; Beven, 2019) and can practically be used to physically constrain rating curves in the high discharge domain (Hollaway et al., 2018). At a river bifurcation, where a river splits into two or more branches, a water balance could be imposed as physical constraint. The water balance is closed if at a bifurcation the incoming discharge equals the sum of all outgoing discharges. A non-closed water balance is a direct indication of uncertainty in the respective rating curves. To our current knowledge, water balance considerations at a river bifurcation have not yet been addressed for the assessment of rating curves.

There is broad scientific consent that the uncertainty in discharges derived with rating curves must be better assessed (Di Baldassarre et al., 2009; McMillan et al., 2017; Pappenberger et al., 2006). In literature, many methods have been used to estimate this uncertainty, such as quantification from residual errors (Herschy, 1999; ISO 1100, 2010), Generalized Likelihood Uncertainty Estimation (Guerrero et al., 2012) and error assessment using hydraulic models (Di Baldassarre and Montanari, 2009; Domeneghetti et al., 2012; Lang et al., 2010). In recent years, combining Bayesian inference with Markov Chain Monte-Carlo has shown to be an effective method for rating curve establishment with which uncertainty assessment is easily performed (Le Coz et al., 2014; Perret et al., 2021; Mansanarez et al., 2019; Moyeed and Clarke, 2005). Bayesian inference allows the inclusion of prior knowledge of the hydraulics in a river, while updating that prior knowledge based on water level and discharge gaugings.

In this study, gaugings at two bifurcations of the Dutch Rhine river are used to study the effect of incorporating two physical constraints in the rating curve construction. Those two constraints are the gradual bed level degradation that occurs close to the first bifurcation and the water balance closure at the two bifurcations. The following research question is addressed:

To what extent does the incorporation of bed level degradation and water balance closure affect the rating curves at two major bifurcations of the Dutch Rhine river?

The paper is structured as follows. Section 2 describes relevant features of the Dutch Rhine river and its two major bifurcations, and the available discharge and water level gaugings. Section 3 describes the standard rating curve model, how the rating curve parameters are estimated using Bayesian inference and Markov Chain Monte Carlo sampling, and how bed level degradation and water balance closure are incorporated into the standard rating curve model. Section 4 describes the results for the two bifurcation and shows the potential added value of incorporating these physical constraints for the two analysed bifurcations. Section 5 discusses the results. Finally, conclusions are drawn in Section 6.

2. Study area and available gaugings

2.1. Study area

The Rhine river enters the Netherlands near Lobith (then called the Bovenrijn) after which it splits at the Pannerdensche Kop (Fig. 1). Here, the discharge of the Bovenrijn is distributed in a ratio of approximately 2/3rd and 1/3rd over the Waal and Pannerdensch Kanaal respectively. After 11 km, the Pannerdensch Kanaal splits at the IJsselkop into the Nederrijn and IJssel in again ratio of 2/3rd and 1/3rd, respectively. The current planform of the two bifurcation exists since the 18th century (Kleinhans et al., 2011). Since then, the bifurcations have been morphologically relatively stable and the discharge distribution has been fairly constant. The Rhine distributaries all have compound cross-sections with a main channel and floodplains on both sides. Groynes narrow the flow widths under low and moderate discharges. Embankments demarcate the main channels from the floodplains, and prevent flooding of the floodplains until a Bovenrijn discharge of approximately 5,000 m^3/s. For lower discharges, movable weirs in the Nederrijn increase the water levels for better navigability.

All of the Rhine branches are protected by dikes which are designed to withstand Bovenrijn discharges of 16,000 m^3/s. In 2017, the Netherlands have adopted a new flood risk policy under which the river system is designed based on local flood risks along the branches instead of a single design discharge (Kok et al., 2017). Multiple failure mechanisms are accounted for in the calculation of the probability of flooding. Therefore, design conditions of dikes are determined by the entire range of possible discharges along the downstream branches, for which the discharge distribution at the bifurcations is important. The maximum attainable discharge in the Bovenrijn is estimated to be 18,000 m^3/s (Bomers et al., 2019) and is thus the highest discharge that is accounted for in flood safety assessment.

In 1995, the highest gauged discharge event of approximately 12,000 m^3/s occurred. No major flooding occurred, but it was still the

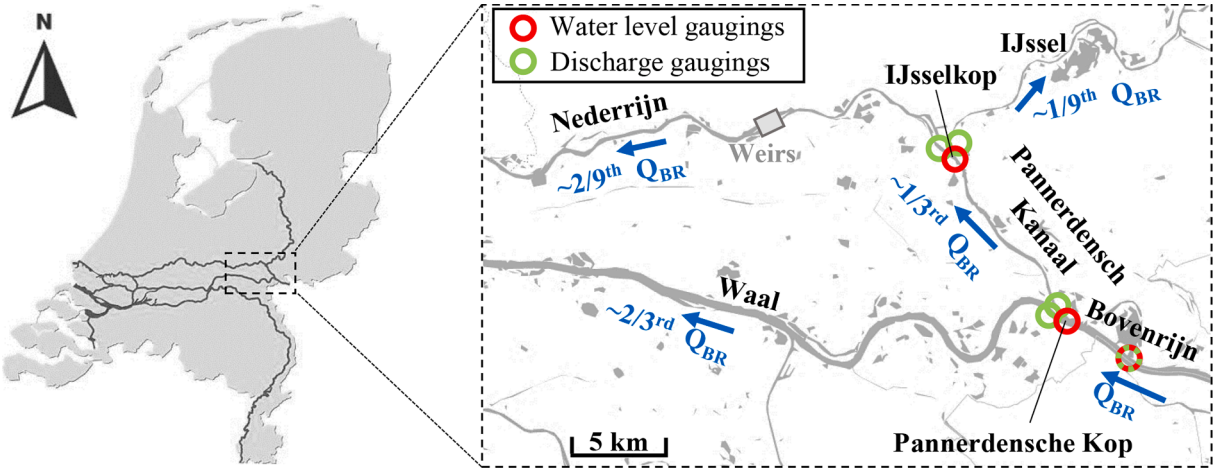

Fig. 1. Map of the bifurcation area, including the discharge distribution over the distributaries, the location of the weirs in the Nederrijn and the gauging locations. Figure adapted from Gensen et al. (2020).

reason to set up the 'Room for the River' program. Under this program, 34 major projects were implemented along the Rhine distributaries. Besides dike reinforcements, many projects aimed at lowering water levels under flood discharges, through e.g. side-channels, floodplain excavations and dike relocations. Most of these projects were completed between 2010 and 2015.

2.2. Available discharge and water level gaugings

Discharges have been intermittently estimated by Rijkswaterstaat (Executive agency of the Dutch Ministry of Infrastructure and Water Management) from 1988 onwards at five locations close to the bifurcations (Fig. 1). These discharge estimates are derived from cross-sectional flow velocity profiles, measured with mechanical hydrometric current meters up to 2002 and after that with an Acoustic doppler current profiler (ADCP). Generally, high discharge events are covered in the observational records, e.g. the events in 1993 ($Q_{Bovenrijn}$ ≈11,000 m^3/s) and 1995 ($Q_{Bovenrijn}$ ≈12,000 m^3/s). The most gaugings are available for the Bovenrijn and the two distributaries from the Pannerdensche Kop bifurcation (Table 1). For the Nederrijn and IJssel distributaries much less gaugings are available, especially in the past decade.

Water levels are continuously measured at Lobith and at both bifurcation points using accurate automatic boat-driven shaft encoders. These measurements are available for the entire period of discharge gaugings. Water levels are given relative to the ordnance level *NAP*.

3. Methodology

Rating curves are constructed by Bayesian inference for the three branches at each of the two bifurcations. First, the available data is processed (Section 3.1). Then, a standard rating curve model is introduced in Section 3.2, of which the parameters are estimated by Bayesian inference using Markov Chain Monte Carlo sampling (Section 3.3). Subsequently, bed level degradation is incorporated into the standard rating curve model aiming to improve rating curve accuracy (Section 3.4). Finally, the model is adapted to improve the water balance closure at the bifurcation (Section 3.5).

3.1. Data processing

The available data is processed in two steps: 1) transform the locations of gaugings such that they coincide at the bifurcation and 2) remove gaugings that are influenced by the weirs.

For step 1: At both bifurcations, the inflowing discharge is measured less than 10 km upstream from the bifurcation (Fig. 1). To make use of the physical constraint of water balance closure and water level continuity at the bifurcation, the upstream gaugings must be relocated to the bifurcation: Lobith to Pannerdensche Kop and Pannerdensche Kop to IJsselkop. The available discharge gaugings are paired with coinciding day-averaged water levels. It is assumed that by day-averaging, the error in water level gaugings caused by the time delay between the two locations is negligibly small. The Pannerdensch Kanaal discharge gaugings are thus paired with water levels at both the Pannerdensche Kop (where it is the outgoing discharge) and IJsselkop (where it is the incoming discharge).

Table 1
Total and used number of gaugings at the Pannerdensche Kop and IJsselkop between 1988 and 2018.

Location	Branch	Total gaugings	Used gaugings
Pannerdensche Kop	Bovenrijn	1303	727
	Waal	1202	1202
	Pan. Kanaal	1535	747
IJsselkop	Pan. Kanaal	1535	675
	Nederrijn	561	364
	IJssel	868	868

For step 2: In the lower discharge domain, the weirs in the Nederrijn influence the water levels at the gauging locations upstream. Furthermore, the weirs affect the discharge distribution in this domain, with relatively more discharge diverted towards the Waal at the Pannerdensche Kop and towards the IJssel at the IJsselkop. Even though it is possible to include the effects of the weirs in the rating curve model (e.g. Le Coz et al., 2014), it is chosen to exclude the gaugings in this domain from the rating curve construction as the main interest is in higher discharges. Therefore, after visual inspection, gaugings are excluded that have a water level below 9.00 m + NAP and 8.50 m + NAP at the Pannerdensche Kop and IJsselkop, respectively (Fig. 2 and Table 1).

3.2. Standard rating curve model

The standard form of the rating curve model is given in Eq. (1) and is based on the Manning–Strickler formula for steady and uniform flow in a wide, rectangular cross-section. For complex cross-sections, such as those of the Rhine branches, a rating curve often exists of successive or additive segments (Le Coz et al., 2014). In this study, the rating curve of the Rhine branches is described by two segments, as presented in Eq. (1), roughly representing the main channel (mc) and the floodplains (fp). Using a single segment does not represent the complexity of the cross-section sufficiently (a deterministic fit resulted in much larger residual errors for a single segment then for two segments), while three segments may result in overfitting of the rating curve (see Section 5.1).

$$Q(h) = \begin{cases} 0, & \text{if } h \le b_{mc} \\ a_{mc}(h - b_{mc})^{p_{mc}}, & \text{if } h > b_{mc} \text{ and } h \le b_{fp} \\ a_{mc}(h - b_{mc})^{p_{mc}} + a_{fp}(h - b_{fp})^{p_{fp}} & \text{if } h > b_{fp} \end{cases} \quad (1)$$

In this equation, Q is the branch discharge [m^3/s], h is the water level [m + NAP], a_{mc} and a_{fp} are terms that mainly depend on the width, bed slope and hydraulic roughness of the sub-sections [$m^{4/3}/s$], b_{mc} and b_{fp} roughly represent the bed level (i.e. when there is no flow) and the floodplain level respectively [m + NAP], p_{mc} and p_{fp} are the hydraulic exponents [-] which should be close to 5/3 according to the Manning–Strickler equation for a wide, rectangular cross-section.

3.3. Bayesian inference and Markov Chain Monte Carlo sampling

Using the rating curve model from Eq. (1), the discharge at location i can be modelled, ($\widehat{Q_i}$), given the model parameters $\boldsymbol{\theta_i}$ [$a_{mc,i}$, $a_{fp,i}$, $b_{mc,i}$, $b_{fp,i}$, $p_{mc,i}$, $p_{fp,i}$]:

$$\widehat{Q_i} = f_1(h_i|\boldsymbol{\theta_i}) \quad (2)$$

To account for gauging and model errors, a structural error term is added:

$$\begin{aligned} Q_i &= \widehat{Q_i} + \epsilon_{rc,i} \\ \epsilon_{rc,i} &= \epsilon_{gauging,i} + \epsilon_{model,i} \sim \mathcal{N}\left(0, \varsigma_{rc,i}\widehat{Q_i}\right) \end{aligned} \quad (3)$$

where for the structural error $\epsilon_{rc,i}$ a normal distribution is assumed without bias and a variance that linearly depends on the modelled discharge and the structural error parameter $\varsigma_{rc,i}$.

The available gaugings $\mathcal{O}_i$ each consisting of discharge, water level and relative time, for every location i are used for Bayesian inference to estimate the model parameters $\boldsymbol{\theta_i}$ and unknown structural error $\varsigma_{rc,i}$. Hereby, it is assumed that the gaugings (j = 1:N) are independent. Measurement errors are included in the structural error term, such that the gaugings represent the true value. The likelihood equation is as follows:

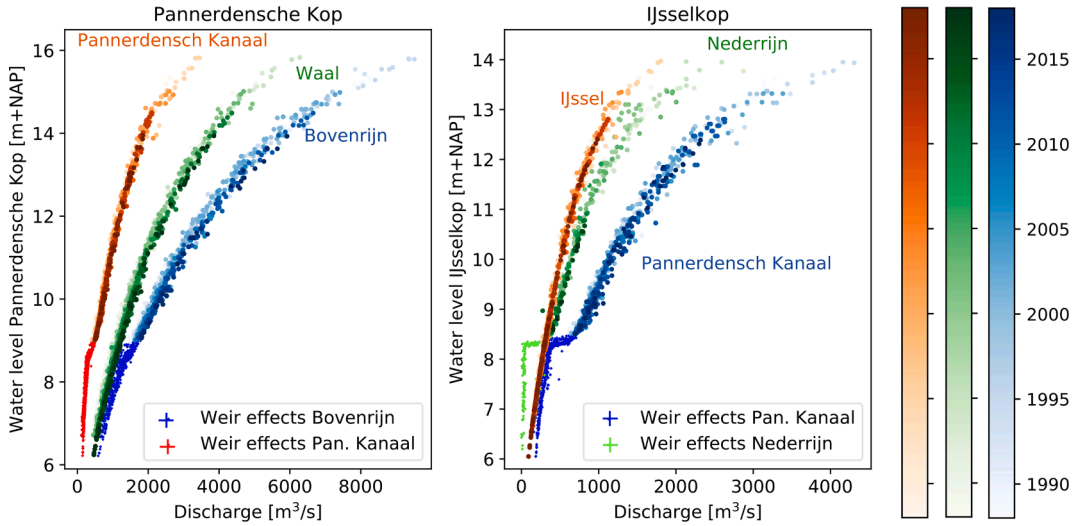

Fig. 2. Discharge-water level gaugings at the 6 locations. The colorbars indicate the year of the gauging, making it insightful that the water level for given discharge at the Pannerdensche Kop decreases over time.

$$\mathscr{L}_i = p\left(\widetilde{\mathbf{Q_i}} \middle| \boldsymbol{\theta_i}, \varsigma_{rc,i}\right) = \prod_{j=1}^{N} p_{\mathcal{N}}\left(\widehat{Q_{ij}}, \varsigma_{rc,i}\widehat{Q_{ij}}\right) \tag{4}$$

The prior distributions of the rating curve parameters are based on a deterministic optimization process. In this process, the hydraulic exponents $p_{mc,i}$ and $p_{fp,i}$ were set to a value of 5/3 (similar to the exponent for a rectangular cross-section in the Manning–Strickler formula). Subsequently, the optimal values of the other parameters in Eq. (1) were determined with a least-squares error approach. The prior distributions are centered around these deterministic values. The prior distributions of $a_{mc,i}$, $a_{fp,i}$, $p_{mc,i}$ and $p_{fp,i}$ are weakly informative following a normal distribution, while $b_{mc,i}$ and $b_{fp,i}$ have non-overlapping uniform prior distributions. The weakly informative priors are necessary for convergence of the posterior distributions with physically realistic results. For the structural error parameter $\varsigma_{rc,i}$, a weakly informative half-Cauchy prior distribution is used (Gelman, 2006). The prior distributions are summarized in Table 2.

A Markov Chain Monte Carlo (MCMC) algorithm is used to sample from the posterior distribution as solving the posterior is non-trivial. MCMC is done using the No U-Turn Sampler, which is a hamiltonian sampling method that efficiently explores the parameter space in a non-random path (NUTS; see details in Hoffman and Gelman, 2014). The MCMC contains 14,000 samples divided over 2 chains. The first 2000 samples of each trace are discarded. Correct MCMC convergence is checked visually for chain similarity, autocorrelation and prior expectations for the posterior distributions.

3.4. DegRC model: incorporating bed level degradation

The upper reaches of the Dutch Rhine branches, most prominently the Waal branch, are experiencing gradual bed degradation at a rate of approximately 2 cm/year (Ylla Arbós et al., 2020). This bed level degradation has been observed to cause a linear decrease of water levels over time (Berends et al., 2021). As this non-stationarity potentially gives a large rating curve uncertainty, the bed level degradation is incorporated in the standard rating model. The new rating curve model ("DegRC") is time-dependent.

The rate of bed level degradation β is incorporated into the standard rating curve model Eq. (1) as an additional stochastic parameter and where the date relative to 1-jan-1988, expressed as t [year], determines the total amount of bed level change. The DegRC model is as follows:

Table 2
Prior distributions of the unknown variables in the rating curve model for each of the 6 locations. The parameter β_i is used to incorporate bed level degradation (Section 3.4). The parameter ς_{WBE} is used to incorporate water balance closure (Section 3.5). $\mathcal{N}(m,s)$ corresponds to a normal distribution with mean m and standard deviation s. $\mathcal{U}(min,max)$ corresponds to a uniform distribution with boundaries at min and max. $HC(\mu,\sigma)$ corresponds to a half-Cauchy distribution with location μ and scale σ.

	Pannerdensche Kop		
Parameter	Bovenrijn	Waal	Pan. Kanaal
$a_{mc,i}$ [$m^{4/3}$/s]	$\mathcal{N}$(140, 25)	$\mathcal{N}$(78, 25)	$\mathcal{N}$(53, 25)
$b_{mc,i}$ [m + NAP]	$\mathcal{U}$(2.0, 6.0)	$\mathcal{U}$(1.2, 5.2)	$\mathcal{U}$(2.6, 6.6)
$p_{mc,i}$ [-]	$\mathcal{N}$(1.67, 0.1)	$\mathcal{N}$(1.67, 0.1)	$\mathcal{N}$(1.67, 0.1)
$a_{fp,i}$ [$m^{4/3}$/s]	$\mathcal{N}$(417, 25)	$\mathcal{N}$(190, 25)	$\mathcal{N}$(326, 25)
$b_{fp,i}$ [m + NAP]	$\mathcal{U}$(9.7, 15.7)	$\mathcal{U}$(8.9, 14.9)	$\mathcal{U}$(10.6, 16.6)
$p_{fp,i}$ [-]	$\mathcal{N}$(1.67, 0.1)	$\mathcal{N}$(1.67, 0.1)	$\mathcal{N}$(1.67, 0.1)
$\varsigma_{rc,i}$ [-]	HC(0, 2)	HC(0, 2)	HC(0, 2)
β_i [m/year]	$\mathcal{U}$(-0.03, 0)	$\mathcal{U}$(-0.03, 0)	$\mathcal{U}$(-0.03, 0)
ς_{WBE} [-]	HC(0, 2)	HC(0, 2)	HC(0, 2)
	IJsselkop		
Parameter	Pan. Kanaal	Nederrijn	IJssel
$a_{mc,i}$ [$m^{4/3}$/s]	$\mathcal{N}$(56, 25)	$\mathcal{N}$(42, 25)	$\mathcal{N}$(24, 25)
$b_{mc,i}$ [m + NAP]	$\mathcal{U}$(1.8, 5.8)	$\mathcal{U}$(2.8, 6.8)	$\mathcal{U}$(1.8, 5.8)
$p_{mc,i}$ [-]	$\mathcal{N}$(1.67, 0.1)	$\mathcal{N}$(1.67, 0.1)	$\mathcal{N}$(1.67, 0.1)
$a_{fp,i}$ [$m^{4/3}$/s]	$\mathcal{N}$(326, 25)	$\mathcal{N}$(331, 25)	$\mathcal{N}$(123, 25)
$b_{fp,i}$ [m + NAP]	$\mathcal{U}$(8.6, 14.6)	$\mathcal{U}$(9.2, 15.2)	$\mathcal{U}$(8.6, 14.6)
$p_{fp,i}$ [-]	$\mathcal{N}$(1.67, 0.1)	$\mathcal{N}$(1.67, 0.1)	$\mathcal{N}$(1.67, 0.1)
$\varsigma_{rc,i}$ [-]	HC(0, 2)	HC(0, 2)	HC(0, 2)
β_i [m/year]	$\mathcal{U}$(-0.03, 0)	$\mathcal{U}$(-0.03, 0)	$\mathcal{U}$(-0.03, 0)
ς_{WBE} [-]	HC(0, 2)	HC(0, 2)	HC(0, 2)

$$Q(h,t) = \begin{cases} 0, & \text{if } h \le b_{mc} \\ a_{mc}(h-(b_{mc}+\beta t))^{p_{mc}}, & \text{if } h > b_{mc} \text{ and } h \ge b_{fp} \\ a_{mc}(h-(b_{mc}+\beta t))^{p_{mc}} + a_{fp}(h-b_{fp})^{p_{fp}} & \text{if } h > b_{fp} \end{cases} \tag{5}$$

A constant rate of bed level degradation β is assumed for the entire observation period, in line with the linear decrease of water levels observed by Berends et al. (2021). In Section 5.1, it is discussed that splitting up the observational record into two periods and deriving independent rating curves for these periods, does not improve rating curve accuracy.

Bayesian inference and estimation of the posterior distribution of the

model parameters $\boldsymbol{\theta_i}$ and structural error parameter $\varsigma_{rc,i}$ is similar as before, but now the unknown bed level degradation β_i is inferred as well. It is expected that the posterior distributions of model parameter $b_{mc,i}$ and β_i are correlated. Therefore, a weakly informed uniform prior distribution for β_i is chosen that matches the weakly informed uniform prior distribution of $b_{mc,i}$. The uniform prior distribution for β_i is bounded between −0.03 and 0.0 m/year because it is expected to be around −0.02 m/year or less.

3.5. BifRC model: improving water balance closure

In a next step, water balance closure is incorporated in the derivation of the rating curves. The bifurcation rating curves ("BifRC") model is constructed in which the rating curves at a bifurcation are derived all at once and in which the water level error is calculated. In this set-up, the rating curves of the branches are dependent on each other. The BifRC model aims at reducing the water balance error at a bifurcation from rating curves at the individual locations, while maintaining accurate rating curves for each branch.

The BifRC model consists of the rating curve models at each location, similar to those in the DegRC model Eq. (5) deterministic water balance error calculation Eq. (6):

$$WBE\left(h_{bif},t_{bif}\right)=\frac{Q_0\left(h_{bif},t_{bif}\right)-Q_1\left(h_{bif},t_{bif}\right)-Q_2\left(h_{bif},t_{bif}\right)}{Q_0\left(h_{bif},t_{bif}\right)} \tag{6}$$

where *WBE* is the deterministic water balance error relative to the model upstream discharge Q_0 for water level h_{bif} and the time t_{bif} in years relative to 01–01-1988 (determining the total amount of bed level degradation in the branches), and Q_1 and Q_2 are the model discharges in the distributaries. The model discharges are defined by Eq. (5).

Using this equation for the deterministic water balance error, the water balance error can be modelled for given water levels h_{bif} and relative times t_{bif}:

$$\widehat{WBE}=f_2\left(h_{bif},t_{bif}\middle|\widehat{\theta}_0,\widehat{\theta}_1,\widehat{\theta}_2\right) \tag{7}$$

To account for the combined errors of the three rating curves, an error term is added, bringing the WBE close to 0:

$$WBE=\widehat{WBE}+\epsilon_{WBE}\approx 0. \qquad \epsilon_{WBE}\sim\mathcal{N}(0,\varsigma_{WBE}) \tag{8}$$

ϵ_{WBE} partly describes the same residual error between the rating curves and the gaugings as the rating curve error parameters $\epsilon_{rc,i}$. However, the errors do not add up, as WBE is determined by the modelled discharges $\widehat{Q_i}$.

The BifRC model combines the equations for the random values Q_0, Q_1, Q_2 and *WBE*. The likelihood equation for the BifRC model is defined such that it calculates the likelihood of the joint distribution of those random values. By looking at the joint distribution, the water balance error is accounted for in the derivation of the individual rating curves for each branch and a reduction of the water balance error can be achieved by an adjustment in any of the rating curves. The equation is a summation of the log-likelihoods (see Eq. (4)) for each location at the bifurcation and a log-likelihood term based on water balance closure, whereby it is assumed that the gaugings from the three branches are independent. In this formulation, the minimum likelihood depends on the accuracy of the individual rating curves as well as on the degree of water balance closure. The likelihood equation reads:

$$\ell_{bif}=\log\mathscr{L}_0+\log\mathscr{L}_1+\log\mathscr{L}_2+\log p\left(\widetilde{\mathcal{O}_{\mathbf{WBE}}}\middle|\boldsymbol{\theta_0},\boldsymbol{\theta_1},\boldsymbol{\theta_2}\right) \tag{9}$$

The gaugings, $\mathcal{O}=[\mathcal{O}_0,\mathcal{O}_1,\mathcal{O}_2]$ that are used in the individual likelihood terms ($\mathscr{L}_0$, $\mathscr{L}_1$ and $\mathscr{L}_2$) do not cover the entire discharge domain for which the water balance error is aimed to be reduced (up to 18.0 m + NAP at the Pannerdensche Kop, see Fig. 3. Therefore, a hypothetical record $\widetilde{\mathcal{O}_{\mathbf{WBE}}}$ is used, consisting of water balance errors $\widetilde{\mathbf{WBE}}$ and water levels $\widetilde{\mathbf{h_{bif}}}$ (Fig. 3). Furthermore, each data point has a random relative time $\widetilde{\mathbf{t_{bif}}}$ in years between the start and end of the period of interest (January 1, 1988 to December 31, 2018). The $\widetilde{\mathbf{WBE}}$ samples are randomly drawn from a normal distribution $\mathcal{N}(0,var)$. The variance in this distribution is a constant value and is calculated from all gauged discharges. This assumption means that the water balance error samples do not increase in magnitude as function of observed or modelled discharges and water levels. The water level samples ($\widetilde{\mathbf{h_{bif}}}$) are linearly spaced in the domain of interest: 9.0–18.0 m + NAP at the Pannerdensche Kop and 8.5–17.0 m + NAP at the IJsselkop. The hypothetical record contains 727 and 364 data points for the Pannerdensche Kop and IJsselkop locations, respectively. The amount of data points determines the relative weight of the water balance term in the combined likelihood equation Eq. (9). To ensure that the individual rating curves remain accurate, the amount of data points in the hypothetical record is chosen to be equal to the shortest record from the three branches at a bifurcation. The sensitivity analysis in section 5.1 shows that the results are marginally influenced by the amount of data points in the hypothetical record.

The prior distributions of the rating curve parameters for each location are the same as for the individual rating curves, see Table 2. For the water balance error parameter ϵ_{WBE}, a weakly informative half-Cauchy prior distribution is used (Gelman, 2006). In total, the BifRC model consists of 25 parameters, 8 parameters for the rating curve of each branch ($\boldsymbol{\theta_i}$ and $\varsigma_{rc,i}$) and ς_{WBE} for the water balance.

4. Results

The results are shown for the branches connected to the Pannerdensche Kop (Section 4.1) and to the IJsselkop (Section 4.2) separately.

4.1. Pannerdensche Kop

Standard rating curves are constructed for the three branches connected to the Pannerdensche Kop (Section 4.1.1). Then, bed level degradation (Section 4.1.2) and water balance closure (Section 4.1.3) are incorporated and the resulting differences between the rating curves or rating curve parameters is shown.

4.1.1. Rating curves without physical constraints at the Pannerdensche Kop

The modelled rating curves for the branches connected to the Pannerdensche Kop matches the trends of the available gaugings (Fig. 4). For low water levels (<11.0 m + NAP), the modelled credibility intervals do not entail all data. The posterior distributions of the structural error parameter $\boldsymbol{\varsigma_{rc}}$ (Fig. 5) is dominated by the large spread in data at this lower domain, which then also results in wide credibility intervals for higher water levels, as the structural error $\boldsymbol{\epsilon_{rc}}$ scales linearly with the modelled discharge (see Eq. (3)).

The effect of the second segment in the rating curve model, roughly describing the discharge contribution of the floodplains, is observed in all branches as the slope of the rating curve changes relatively abruptly. The water level at which this occurs differs in each branch, indicated by the median values of $b_{fp,i}$: 12.60 m + NAP in the Bovenrijn, 12.10 m + NAP in the Waal and 13.58 m + NAP in the Pannerdensch Kanaal (Fig. 5). The differences between these values are consistent with the elevations of the embankments that demarcate the main channel from the floodplains, which are lower in the Waal river than in the Pannerdensch Kanaal.

4.1.2. Rating curves with bed level degradation at the Pannerdensche Kop

Fig. 6 shows that the credibility intervals of the rating curves are narrower for a given year when bed level degradation is accounted for. This shows that a large amount of spread in the gaugings can be explained by a mean trend in water levels, which can be attributed to

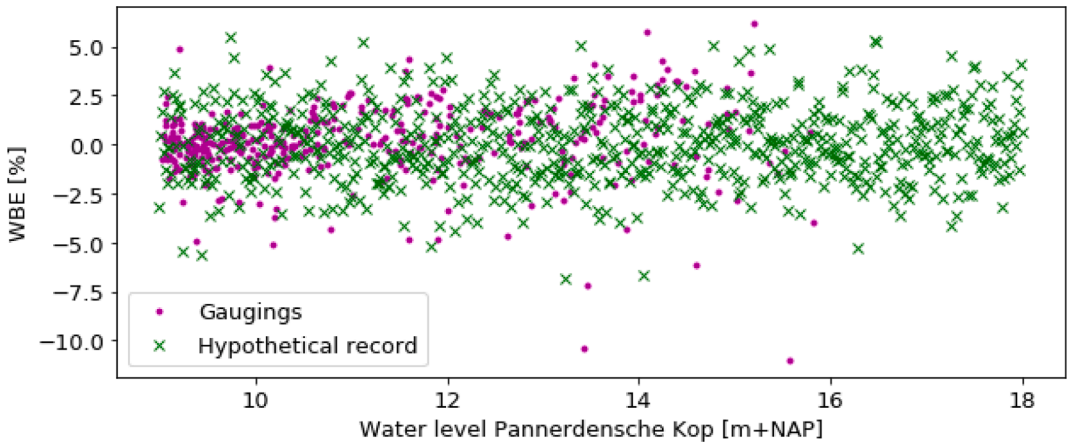

Fig. 3. Hypothetical record at the Pannerdensche Kop of water balance errors for water balance closure (green) and the water balance errors of same-day gaugings (purple).

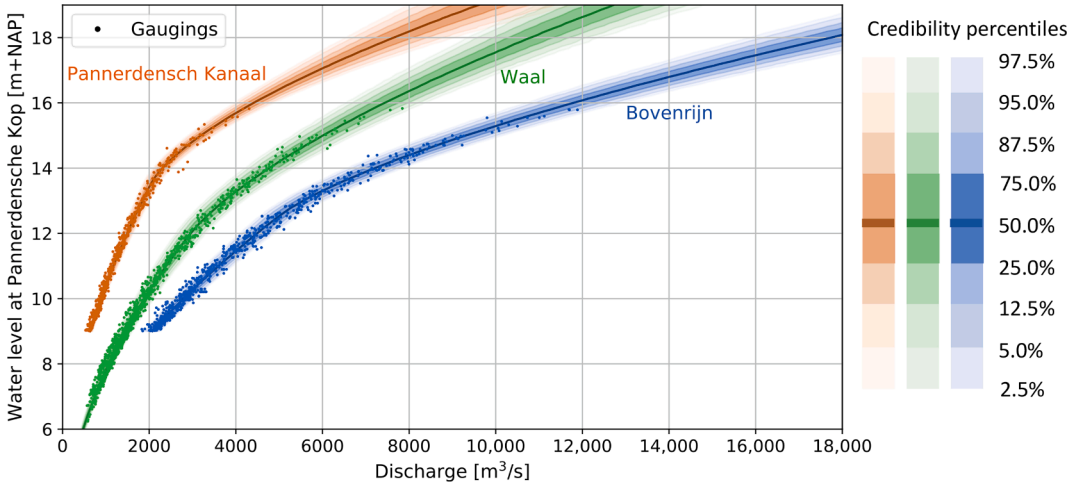

Fig. 4. The mean and credibility intervals modelled with the Standard rating curve model for the Bovenrijn (Blue), Waal (Green) and the Pannerdensch Kanaal (Orange) at the Pannerdensche Kop bifurcation.

bed level degradation. Fig. 5 shows that the narrowing of the credibility intervals can be attributed to a narrowing of the posterior distributions of some of the rating curve parameters in the first segment ($a_{mc,i}$, $b_{mc,i}$, $p_{mc,i}$) as well as to large reductions of the structural error parameters ς_{rc} with respect to the StaRC model: 30% reduction in the Bovenrijn, 50% reduction in the Waal and 20 % reduction in the Pannerdensch Kanaal.

The modelled bed level degradation rates β (Fig. 5) are in the same order as observed values (Ylla Arbós et al., 2020), with median values of −1.7, −2.0 and −1.4 cm/year in the Bovenrijn, Waal and Pannerdensch Kanaal, respectively. Also consistent with observed values of bed degradation rates by Ylla Arbós et al. (2020), the modelled bed degradation rate in the Waal is higher than in the Pannerdensch Kanaal. This also affects the discharge distribution, with in time increasingly more discharge diverted towards the Waal over time.

4.1.3. Bifurcation rating curves with improved water balance closure at the Pannerdensche Kop

If water balance closure is accounted for, the calculated water balance error ($\widehat{WBE}$) at the Pannerdensche Kop bifurcation is strongly reduced in the extrapolation domain (Fig. 7A). In the StaRC and DegRC models, the Bovenrijn discharge is smaller than the sum of the Waal and Pannerdensch Kanaal discharges in the extrapolation domain, resulting in a negative water balance error. The rating curves constructed with the BifRC model result in a better closed water balance, with a slight bias towards a positive water balance error (Bovenrijn > Waal + Pannerdensch Kanaal). This bias may exist, as besides closing the water balance, the likelihood function in the BifRC model still includes the accuracy of the individual rating curves. While not shown, water balance closure is also maintained for different years.

In the BifRC model, a low water balance error is also achieved in the tails of the posterior distributions of the modelled discharges, indicated by the narrow 95% credibility interval of $\widehat{WBE}$ (Fig. 7A). In the StaRC and DegRC models, the modelled discharges are independent between the branches, which may result in large water balance errors. Instead, modelled discharges are dependent in the BifRC model, where the discharges compensate each other to result in a low water balance error.

Incorporating the water balance does not negatively impact the accuracy of the individual rating curves at the Pannerdensche Kop, as the structural error parameters ς_{rc} only slightly increase compared to the DegRC model. This indicates that the residual discharge errors are almost equal. The improved water balance closure in the BifRC model can mainly be attributed to a decrease in modelled Pannerdensch Kanaal discharges (Fig. 7D). In the Pannerdensch Kanaal, fewer gaugings are available for higher water levels in comparison to the other two branches, such that a changed Pannerdensch Kanaal rating curve in the extrapolation domain does not go at the expense of its accuracy.

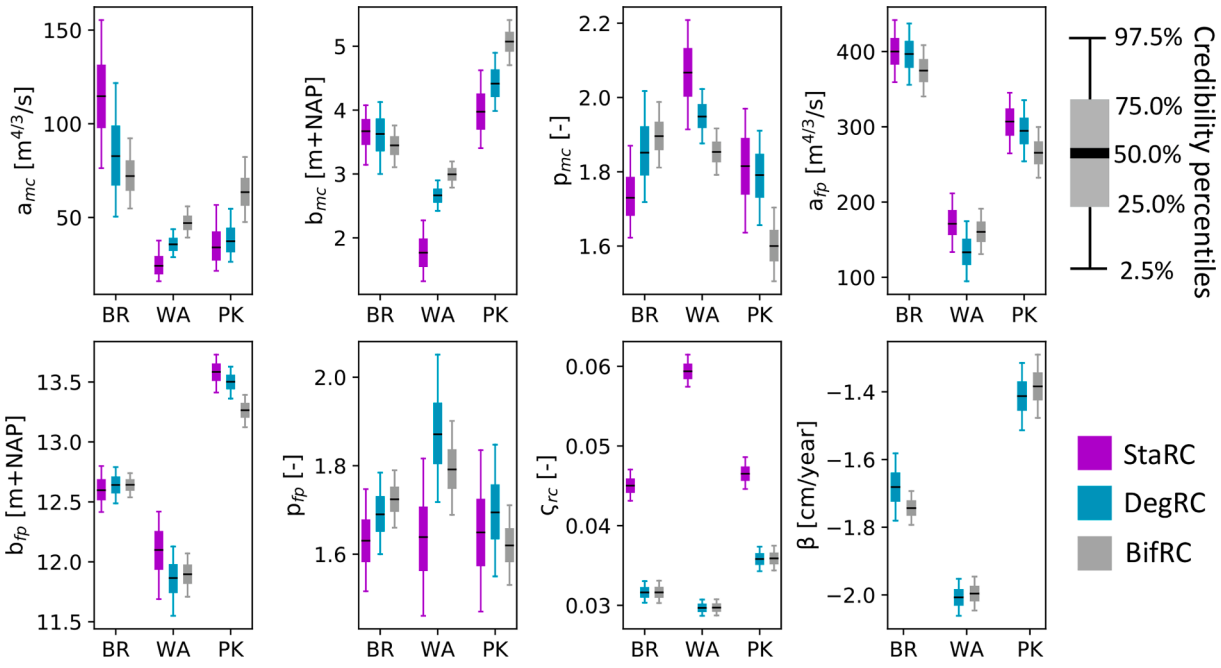

Fig. 5. Posterior distributions of the rating curve parameters for the rating curves at the Pannerdensche Kop in the standard rating curve model (StaRC), the rating curve model with bed level degradation (DegRC) and the bifurcation rating curve model (BifRC).

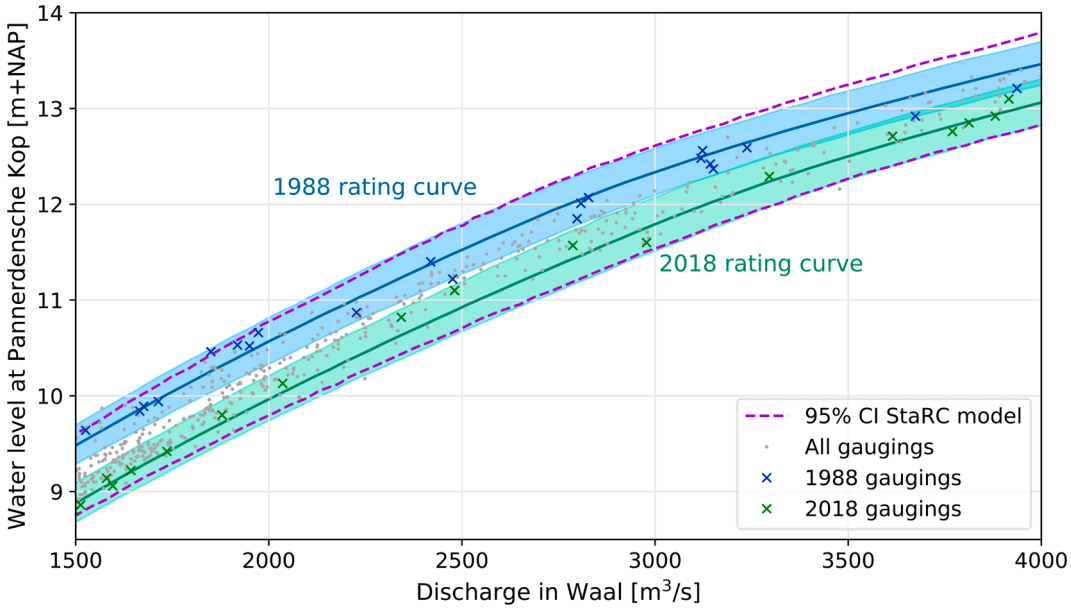

Fig. 6. Rating curves in the Waal for the years 1988 and 2018 modelled with the DegRC model in comparison with the StaRC model (StaRC). The shaded areas mark the 95% credibility intervals (CI) and the continuous lines mark the median rating curves.

4.2. IJsselkop

The rating curves modelled with the StaRC model for the branches connected to the IJsselkop are shown in Section 4.1.1. Subsequently, bed level degradation (Section 4.1.2) and water balance closure (Section 4.1.3) are incorporated and the resulting differences in rating curves and their parameters are shown.

4.2.1. Rating curves without physical constraints at the IJsselkop

The modelled rating curves without physical constraints of the branches connected to the IJsselkop match the available gaugings well (Fig. 8). The gaugings that lie outside the shown credibility intervals are not specifically related to a certain discharge domain, indicating a valid assumption of the linear increase of the structural discharge errors with the modelled discharge itself. Still, the credibility intervals of the rating curves are wide, especially in the Nederrijn branch. Fig. 9 shows that this can be attributed to both high structural error parameters ς_{rc} (mainly in the Nederrijn) and wide posterior model parameter distributions in the first rating curve segment (mainly in the Pannerdensch Kanaal and Nederrijn) or second rating curve segment (mainly in the IJssel). Gaugings in the high discharge domain are scarce, especially recent gaugings (after 2002; Fig. 2). Recent gaugings are expected to have less measurement uncertainty due to the use of ADCP measurement equipment and would thus likely show less spread compared to older

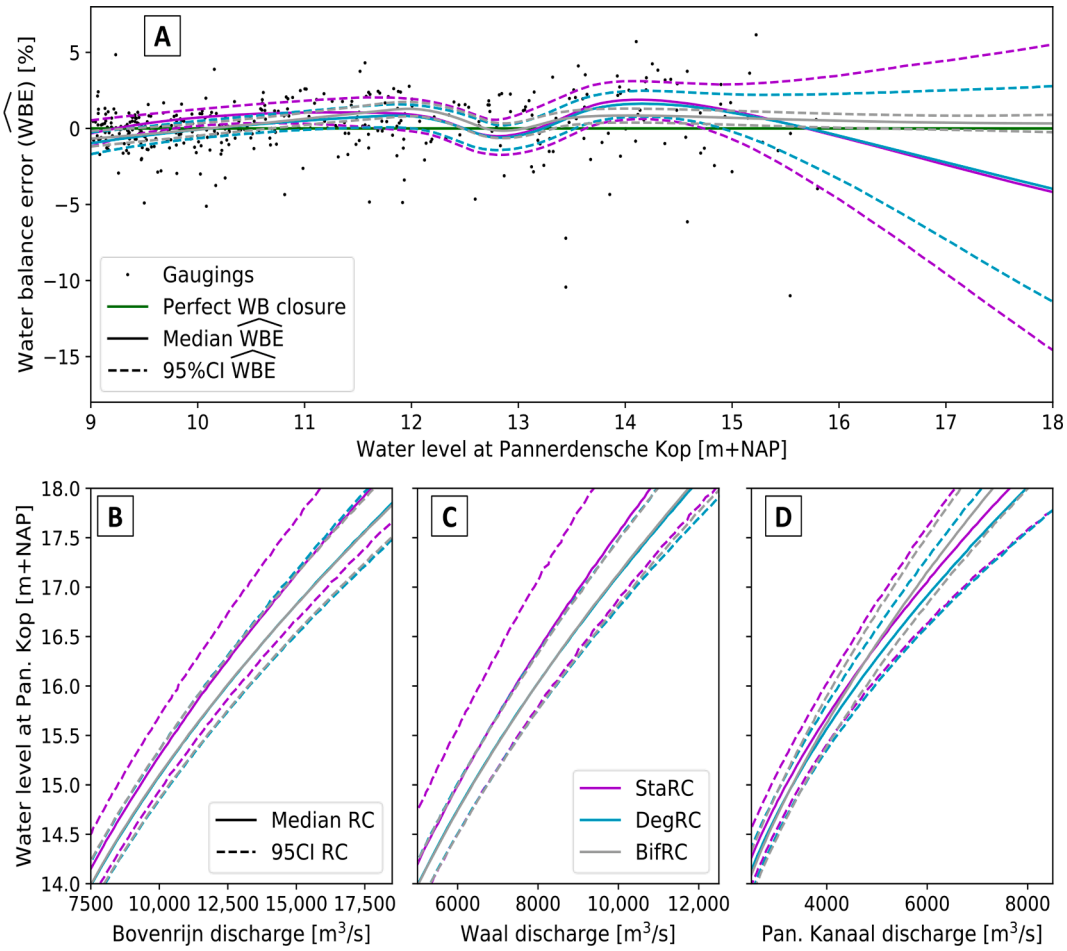

Fig. 7. Comparison of the StaRC, DegRC and BifRC models at the Pannerdensche Kop. A) Water balance error ($\widehat{WBE}$), as defined in Eq. (6), based on the modelled discharges ($\widehat{Q_i}$) in 2018. The markers indicate the WBE calculated with same-day gaugings in each branch (400 gaugings). (B,C,D): Extrapolation domain of the median rating curves and their 95% credibility intervals of the Pannerdensch Kanaal of the Bovenrijn (B), Waal (C) and Pannerden.sche Kanaal (D).

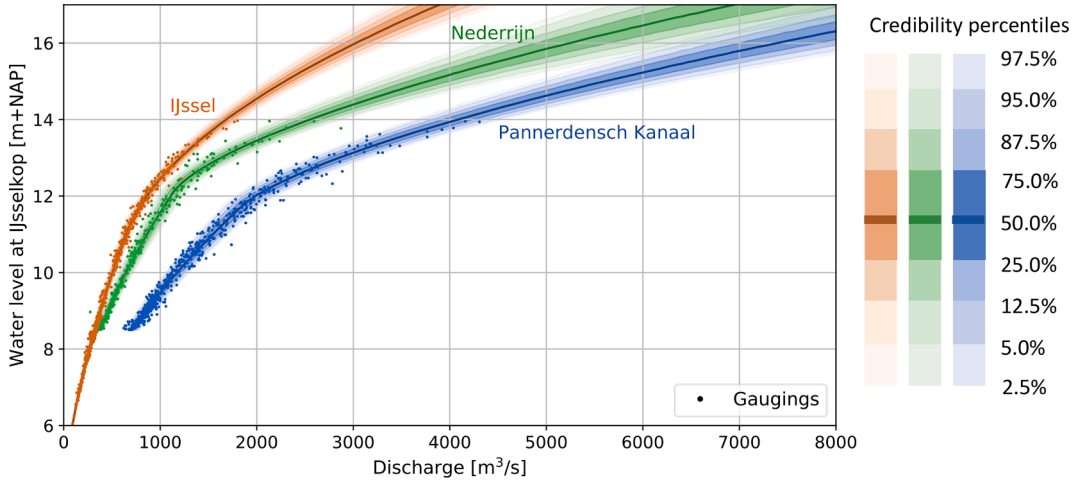

Fig. 8. Rating curves for the Pannerdensch Kanaal (blue), Nederrijn (green) and IJssel (orange) at the IJsselkop bifurcation constructed with the standard rating curve model. The markers show all available gaugings for the branches.

gaugings. Future gaugings during high discharges could narrow the credibility intervals in the second segment of the rating curves.

4.2.2. *Rating curves with bed level degradation at the IJsselkop*

Incorporating bed level degradation for the rating curves at the IJsselkop only marginally increases their accuracy. The median bed degradation rates are −0.9, −0.2 and −0.4 cm/year for the three branches, respectively (Fig. 9). The structural error parameters ς_{rc} only slightly reduce under the DegRC model, as only a small amount of spread in the gaugings can be explained by the bed level degradation. As a

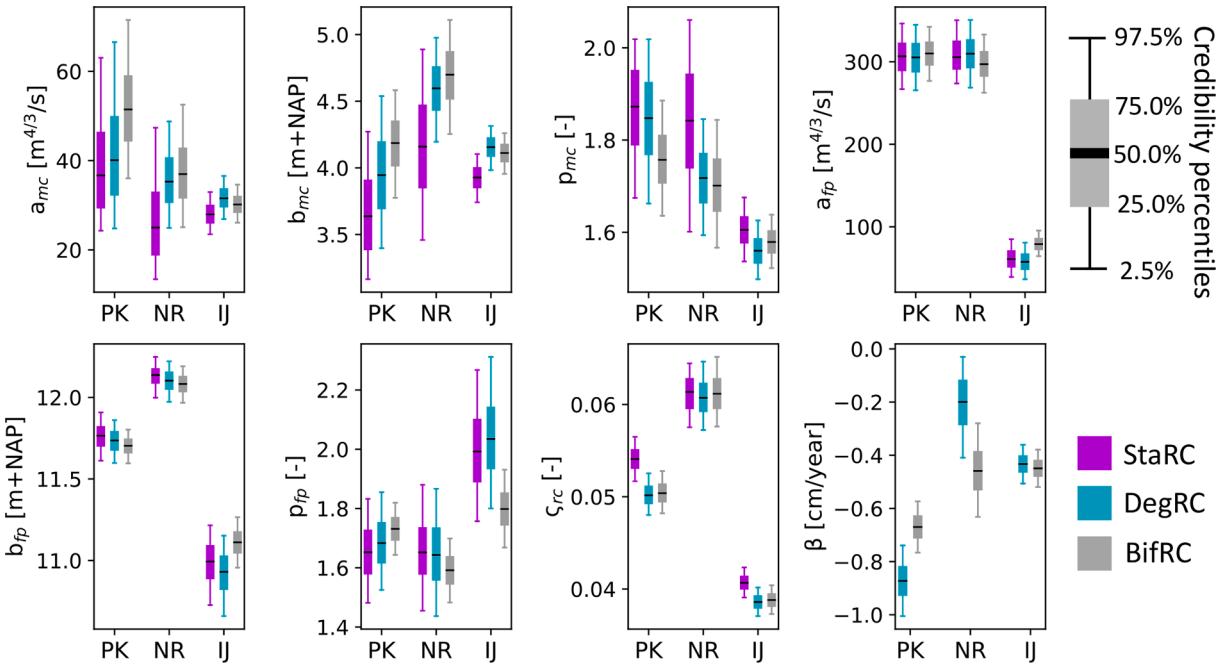

Fig. 9. Posterior distributions of rating curve parameters for the rating curves at the IJsselkop in the standard rating curve model (StaRC), the rating curve model with bed level degradation (DegRC) and the bifurcation rating curve model (BifRC).

result, credibility intervals of the modelled discharges only marginally narrow (Fig. 10B–D).

4.2.3. Bifurcation rating curves with improved water balance closure at the IJsselkop

Even though the water balance error under the StaRC and DegRC model at the IJsselkop is larger than at the Pannerdensche Kop, the BifRC model still minimizes the water balance error (Fig. 10A). A slight bias (~1%) is observed, which may indicate a more structural and biased error in the discharge gaugings. The likelihood term balances the accuracy of the underlying rating curves and the water balance term. The positive bias in the WBE suggests that possibly the Pannerdensch Kanaal discharge is generally overestimated or that the Nederrijn or IJssel discharge is underestimated.

The reduction of the water balance error is attributed to all of the branches (Fig. 10B–D). The Nederrijn and IJssel discharges are lower and the Pannerdensch Kanaal discharge is higher with the BifRC model. All IJsselkop branches have little gaugings for high water levels, such that the changes in rating curves are proportionally divided over all branches. Rating curve accuracy is maintained, indicated by the minor change of the structural error parameters ς_{rc} (Fig. 9).

5. Discussion

5.1. Varying the physical constraints

Generally, the rating curves are little sensitive to the modelling scenarios in physical constraints, with modelled discharges in the extrapolation domain showing little variation between the various modelling choices (Table 3).

In this study, a rating curve model with two segments was chosen, with the segments roughly representing the main channel and floodplains. However, a third segment can be considered, which would represent the contribution of flow over the groyne fields for water levels exceeding the groyne heights. If a third segment is added to the DegRC and BifRC models, the values of the structural error parameter ς_{rc} slightly reduce (Table 3), which indicates that the third segment explains a slight amount of residual errors from the model with two segments. The addition of the third segment increases the discharges for high water levels, which would match the observed underprediction of the modelled discharge for the flood events of 1993 and 1995 (see Section 5.2). However, the prior distributions of the model parameters dominate the posterior distributions, resulting in much wider credibility intervals (Table 3), as very little gaugings are available for very high water levels (>14 m + NAP). Overfitting of the rating curve model can occur when adding more segments to the model (Sörengård and Di Baldassarre, 2017). Therefore, it may be concluded that currently the gaugings do not support the inclusion of a third segment in the rating curve model.

The results showed that incorporating bed level degradation into the model strongly improved the rating curves, indicated by the large reduction in the residual errors. That this single extra parameter strongly improves the results, indicates that it is an important parameter and that overfitting does not occur. Splitting the observational record into two sub-periods and fitting the DegRC and BifRC models to the observations in those periods separately, also explains a small portion of the residual errors (Table 3). Table 3 also shows that modelled Waal discharges in the extrapolation domain are much lower in the second sub-period (i.e. in 2018). This is equivalent to an increase in water levels for given discharges over time. However, this is not as expected, as water-level-lowering interventions have been implemented along the Waal river during this second sub-period (Berends et al., 2021). The lower than expected discharge is likely explained by the lack of recent gaugings of very high discharges in this period. So, while splitting up the observational record into multiple periods may match the available gaugings better, it does not necessarily increase rating curve accuracy and is likely the result of overfitting.

The choice for a certain length of the hypothetical record of water balance errors (see Section 3.5) finds the balance between the rating curve accuracy (i.e. low residual errors) and the reduction of the water balance errors (WBE). In the BifRC model, this length was set equal to the amount of gaugings in the least gauged branch. With a shorter record, the WBE increases, but is still smaller in comparison to the models in which the water balance is not considered (StaRC and DegRC models). Oppositely, with a longer record, the WBE can be reduced slightly more, but at the expense of a further increase of the residual errors between the

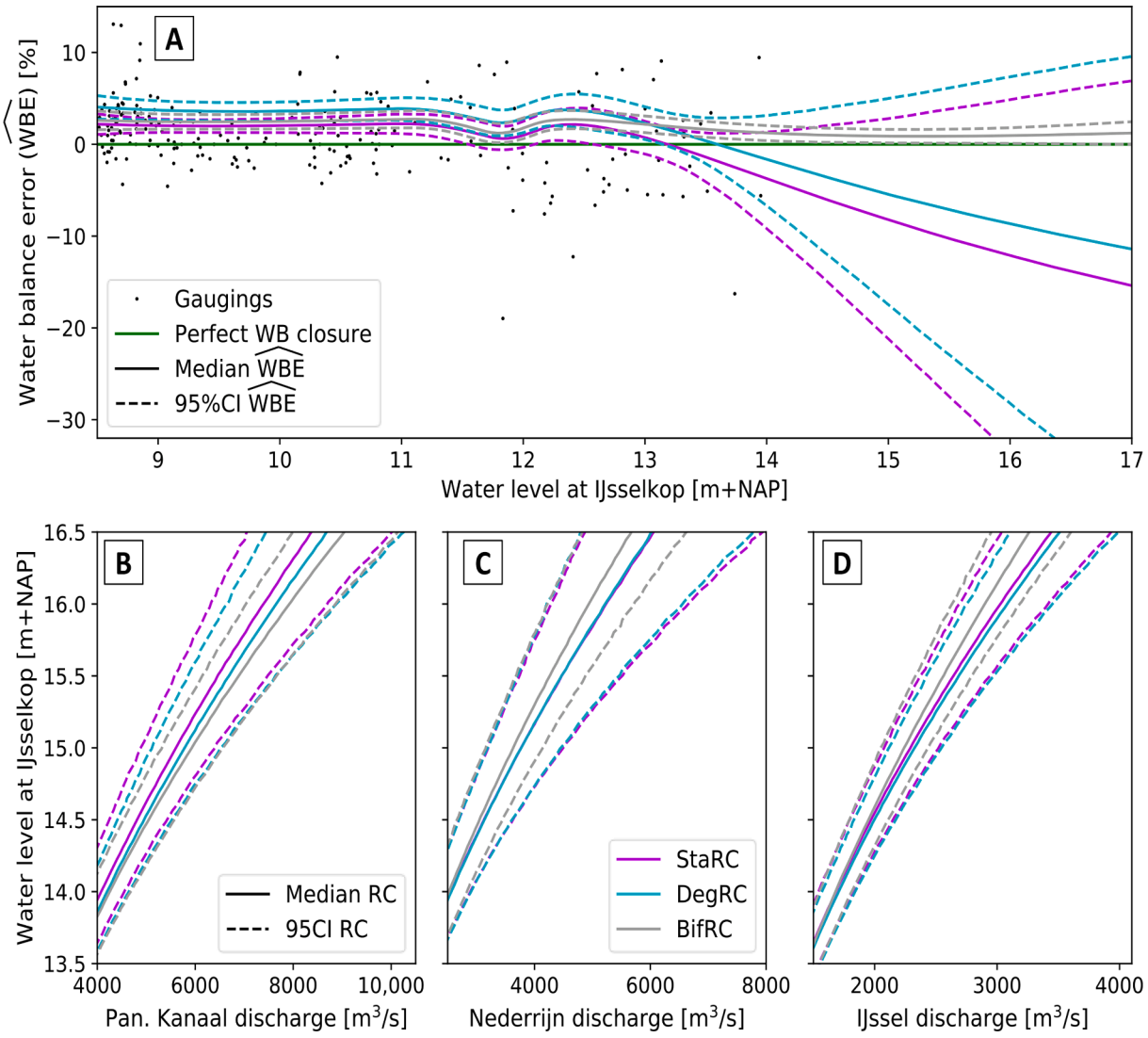

Fig. 10. Comparison of the StaRC, DegRC and BifRC models at the IJsselkop. A) Water balance error ($\widehat{WBE}$), as defined in Eq. (6), based on the modelled discharges ($\widehat{Q_i}$) in 2018. The markers indicate the WBE calculated with same-day gaugings in each branch (215 gaugings). (B,C,D): Extrapolation domain of the median rating curves and their 95% credibility intervals of the Pannerdensch Kanaal (B), Nederrijn (C). and IJssel (D).

Table 3
Relevant rating curve extrapolation results (for a water level of 17 m + NAP in 2018) for the Waal branch under various modelling scenarios. These results are 1) the median discharge ($Q_{50\%}$), 2) the width of the 95% credibility interval of the model discharge $\widehat{Q}$, 3) the median value of the structural error parameter ς_{rc} and 4) the median water balance error based on Eq. (6). The considered modelling scenarios are the addition of a third segment in the DegRC and BifRC models, two sub-periods in the DegRC and BifRC model with a constant ς_{rc} between the two periods, and a different length of the hypothetical record of observed water balance errors (see Section 3.5).

Scenarios (Waal)	$Q_{50\%}$	W95%CI	$\varsigma_{rc,50\%}$	WBE
StaRC model	9,040 m^3/s	1,047 m^3/s	0.0594	−2.5%
DegRC model	9,767 m^3/s	634 m^3/s	0.0297	−2.3%
BifRC model	9,742 m^3/s	442 m^3/s	0.0297	+0.4%
3 segments in DegRC	10,179 m^3/s	1,579 m^3/s	0.0284	−5.3%
3 segments in BifRC	9,646 m^3/s	644 m^3/s	0.0284	+0.4%
2 periods in DegRC	8,054 m^3/s	578 m^3/s	0.0290	+5.8%
2 periods in BifRC	8,697 m^3/s	1,320 m^3/s	0.0294	+0.5%
1/4x WBE samples	9,596 m^3/s	486 m^3/s	0.0296	+1.1%
4x WBE samples	9,903 m^3/s	477 m^3/s	0.0299	+0.3%

modelled rating curve and the gaugings, reflected in a higher ς_{rc}.

5.2. Sensitivity of rating curves to single flood events

The flood events of 1993 and 1995 were important for flood risk management in the Netherlands in multiple aspects. During the 1995 event, the highest observed discharge was gauged in the Bovenrijn branch (i.e. 11,885 m^3/s). Daily gaugings are available for each of the branches. The event is therefore generally used for the calibration of the hydraulic models used in flood risk analyses in the Netherlands. These calibrated hydraulic models are widely used to estimate the water levels for design conditions (former design discharge of the Rhine river system: $Q_{Bovenrijn} = 16.000\ m^3/s$). Therefore, errors in the discharge gaugings of this 1995 high discharge event may have a large effect on estimates of design water levels.

Also in this study, the gaugings of the 1995 flood event have a strong influence on the results. Fig. 11 shows that by excluding the gaugings from the 1995 flood event, modelled rating curves would show lower discharges for given water levels. With the DegRC model, the rating curve change is larger than with the BifRC. This shows that adding the physical constraint of water balance closure reduces the dependency of

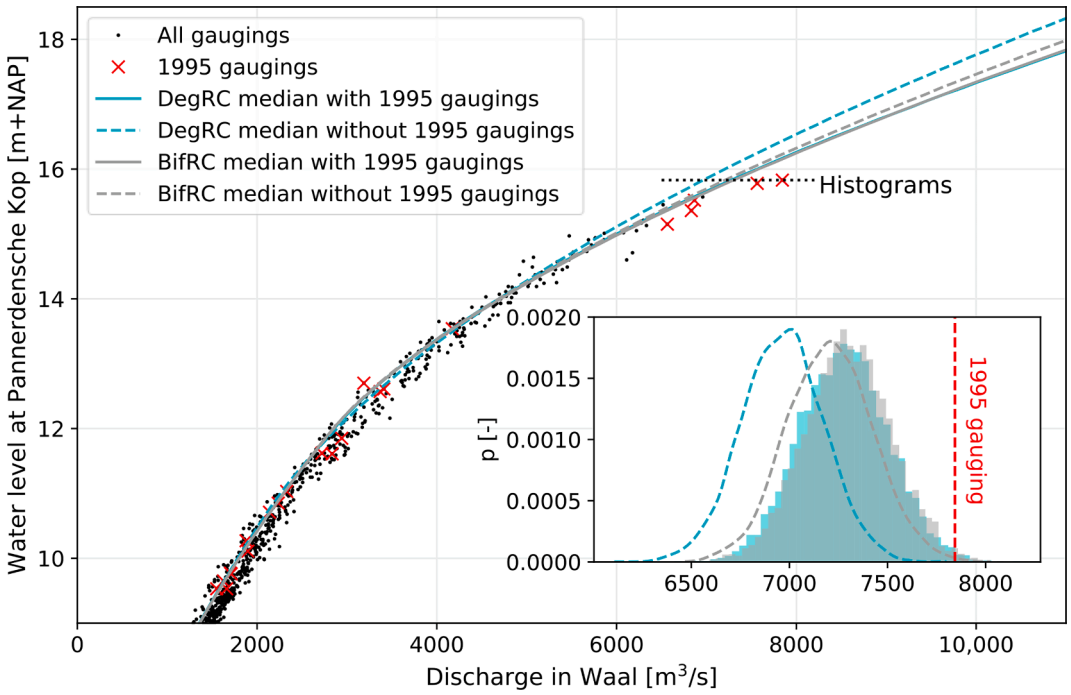

Fig. 11. Median Waal rating curves in 1995 with the DegRC and BifRC models when including (continuous lines; are nearly equal) or excluding (dotted lines) the 1995 gaugings from the observational record. For the highest observed water level (15.83 m + NAP), the probability distributions of the discharges is shown.

rating curves on a single high discharge event.

5.3. Discharge distribution at the bifurcations

In a bifurcating river system, accurate predictions on the distribution of discharge over the downstream branches are crucial for flood risk management (Gensen et al., 2020). With the modelled rating curves using the BifRC model, the discharge distribution over the branches can be estimated while excluding or including the structural error ϵ_{rc}:

$$\begin{aligned} \widehat{DD} &= \frac{\widehat{Q_1}}{\widehat{Q_0}} \qquad \left(\text{excl. } \epsilon_{rc}\right) \\ DD &= \frac{Q_1}{Q_0} \qquad \left(\text{incl. } \epsilon_{rc}\right) \end{aligned} \tag{10}$$

Where DD is the (modelled) discharge distribution, and Q_0 and Q_1 the (modelled) discharges in the incoming branch and one of the outgoing branches, respectively.

The estimated discharge distribution at the Pannerdensche Kop (Fig. 12) is consistent with observations and hydraulic model results

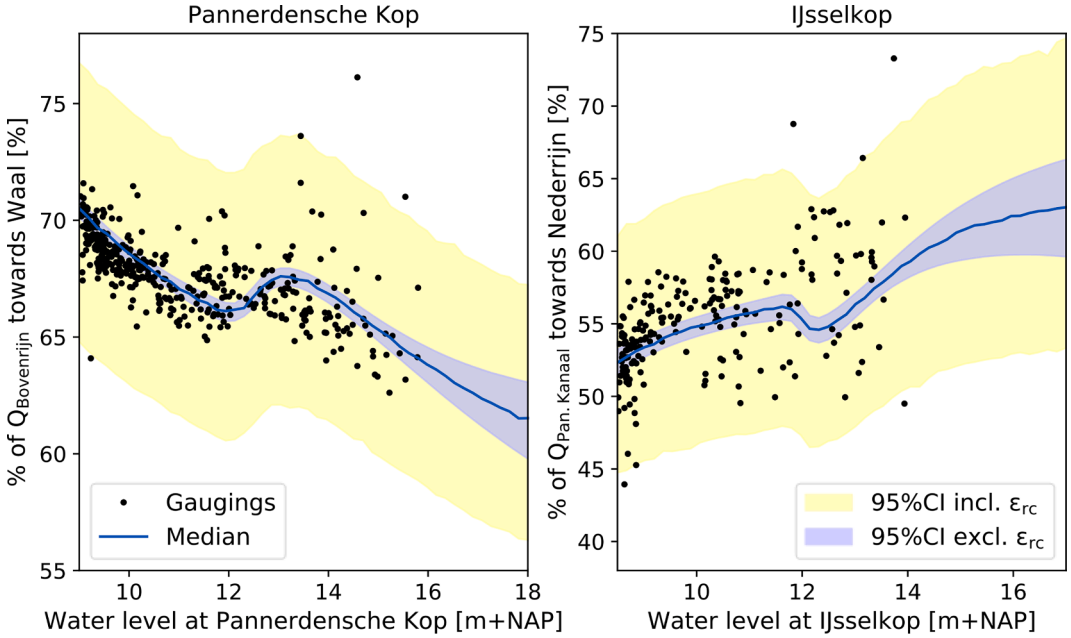

Fig. 12. Estimated discharge distributions and their 95% credibility intervals (CI) using the BifRC model at the Pannerdensche Kop (A) and IJsselkop (B). The figure shows that the structural error dominates the uncertainty in discharge distribution.

(Gensen et al., 2020), with less discharge diverted towards the Waal branch for higher water levels. The change in this trend, between water levels of 12 m + NAP and 14 m + NAP, is caused by the difference in water levels for which the second segment of the rating curve model becomes active (reflected in the value of b_1, see Fig. 5). The embankments that form the division between the main channel and the floodplains are lower in the Waal than in the Pannerdensch Kanaal, which causes relatively more discharge diverted towards the Waal in this domain.

The estimated discharge distribution at the IJsselkop also matches observed values well (Fig. 12). The increasing percentage of discharge diverted towards the Nederrijn for unobserved water levels is inconsistent with hydraulic model results (Gensen et al., 2020). Instead, they predict a decreasing fraction of discharge towards the Nederrijn for increasing water levels. This inconsistency is likely the result of the lack of (recent) gaugings in the high water level domain.

Fig. 12 shows that the uncertainty in discharge distribution is dominated by the structural error ϵ_{rc}. In the BifRC model, these structural errors are assumed independent between the branches. However, it can be expected that the structural errors are to some degree correlated between the branches. Such correlation may be caused by e.g. hysteresis or biased gaugings. Therefore, the uncertainty in discharge distribution is expected to be smaller than the green shaded area, but still larger than the blue shaded area.

5.4. Using the physical constraint of water balance closure for accurate rating curves in river systems

This study has shown that the physical constraint of water balance closure at a river bifurcation can be incorporated into rating curve construction and that it leads to physically more realistic rating curves. Particularly in the extrapolation domain, rating curves are constrained, resulting in better water balance closure and narrower credibility intervals. At a river confluence, the same constraint of water balance closure holds, which can be used to improve rating curve accuracy and thus discharge estimates. As rating curve accuracy can be improved, it can be beneficial to construct rating curves specifically at bifurcations and confluences instead of at another location along a single river branch. These locations must be chosen such that between the rating curve locations the storage of discharge, in for instance floodplains, is negligible. Additionally, the most accurate rating curves may be obtained by gauging in each branch at the bifurcation or confluence during one measurement campaign instead of multiple campaigns in a single branch. When limited resources are available, it is thus recommend to plan measurement campaigns at bifurcations or tributaries at the same time.

6. Conclusions

This study showed that the physical constraints water balance closure and bed level degradation can be incorporated in the construction of rating curves at two river bifurcations. Accounting for bed level degradation strongly improves the accuracy of the rating curves at the bifurcation which experiences bed level degradation, with a reduction of up to 50% of the residual errors from the modelled rating curves with respect to the standard rating curve model. To improve water balance closure, a bifurcation rating curves model was established in which the rating curves of the separate branches at the bifurcation become dependent by adding an explicit water balance term to the likelihood function of the Bayesian inference. The results showed that water balance closure is significantly improved at both of the analysed river bifurcations of the Rhine river, while rating curve accuracy is maintained with residual errors only marginally increasing. As water balance closure is a physical constraint at a bifurcation, the newly constructed rating curves are not only more accurate, they are also expected to be more physically realistic. It is expected that the presented approach in which the water balance is explicitly incorporated can also improve rating curve accuracy and discharge predictions at other river bifurcations or at river confluences.

Funding

This work is part of the Perspectief research programme All-Risk with project number P15–21, which is (partly) financed by NWO Domain Applied and Engineering Sciences, in collaboration with the following private and public partners: the Dutch Ministry of Infrastructure and Water Management (Rijkswaterstaat), Deltares, STOWA, HKV consultants, Natuurmonumenten and the regional water authorities Noorderzijlvest, Vechtstromen, it Fryske Gea, HHNK.

CRediT authorship contribution statement

M.R.A. Gensen: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Visualization. **J.J. Warmink:** Conceptualization, Writing - review & editing, Supervision, Project administration, Funding acquisition. **K.D. Berends:** Conceptualization, Methodology, Software, Writing - review & editing. **F. Huthoff:** Writing - review & editing, Supervision. **S.J.M.H. Hulscher:** Writing - review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Jetze Twijnstra for his work on this topic.

The NUTS sampler, which was used for Markov Chain Monte Carlo sampling, was implemented in the PyMC3 package (Salvatier et al., 2016).

The authors thank the anonymous reviewers for the helpful comments they provided on earlier drafts of the manuscript.

References

Berends, K.D., Gensen, M.R.A., Warmink, J.J., Hulscher, S.J.M.H., 2021. Multidecadal analysis of an engineered river system reveals challenges for model-based design of human interventions. CivilEng 2, 580–598. https://doi.org/10.3390/civileng2030032.

Berends, K.D., Straatsma, M.W., Warmink, J.J., Hulscher, S.J.M.H., 2019. Uncertainty quantification of flood mitigation predictions and implications for interventions. Nat. Hazards Earth Syst. Sci. 19, 1737–1753. https://doi.org/10.5194/nhess-19-1737-2019.

Beven, K., 2019. Towards a methodology for testing models as hypotheses in the inexact sciences. https://doi.org/10.1098/rspa.2018.0862.

Bomers, A., Schielen, R.M.J., Hulscher, S.J.M.H., 2019. Consequences of dike breaches and dike overflow in a bifurcating river system. Nat. Hazards 97, 309–334. https://doi.org/10.1007/s11069-019-03643-y.

Di Baldassarre, G., Claps, P., 2011. A hydraulic study on the applicability of flood rating curves. Hydrol. Res. 42, 10–19. https://doi.org/10.2166/nh.2010.098.

Di Baldassarre, G., Laio, F., Montanari, A., 2009. Design flood estimation using model selection criteria. Phys. Chem. Earth, Parts A/B/C 34, 606–611. https://doi.org/10.1016/j.pce.2008.10.066.

Di Baldassarre, G., Montanari, A., 2009. Uncertainty in river discharge observations: a quantitative analysis. Hydrol. Earth Syst. Sci. 13, 913–921. https://doi.org/10.5194/hess-13-913-2009.

Domeneghetti, A., Castellarin, A., Brath, A., 2012. Assessing rating-curve uncertainty and its effects on hydraulic model calibration. Hydrol. Earth Syst. Sci 16, 1191–1202. https://doi.org/10.5194/hess-16-1191-2012.

Dong, T.Y., Nittrouer, J.A., McElroy, B., Il'icheva, E., Pavlov, M., Ma, H., Moodie, A.J., Moreido, V.M., 2020. Predicting water and sediment partitioning in a delta channel network under varying discharge conditions. Water Resour. Res. 56, e2020WR027199 https://doi.org/10.1029/2020WR027199.

Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models. Bayesian Anal. 1, 515–534. https://doi.org/10.1214/06-BA117A.

Gensen, M.R.A., Warmink, J.J., Huthoff, F., Hulscher, S.J.M.H., 2020. Feedback mechanism in bifurcating river systems: the effect on water-level sensitivity. Water 12, 1915. https://doi.org/10.3390/w12071915.

Guerrero, J.-L., Westerberg, I.K., Halldin, S., Xu, C.-Y., Lundin, L.-C., 2012. Temporal variability in stage-discharge relationships. J. Hydrol. 446–447, 90–102. https://doi.org/10.1016/j.jhydrol.2012.04.031.

Herschy, R., 1999. Hydrometry. Wiley, Principles and Practices.

Hoffman, M.D., Gelman, A., 2014. The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J. Mach. Learn. Res. 15, 1593–1623. URL: http://mcmc-jags.sourceforge.net.

Hollaway, M.J., Beven, K.J., Benskin, C.M.W., Collins, A.L., Evans, R., Falloon, P.D., Forber, K.J., Hiscock, K.M., Kahana, R., Macleod, C.J., Ockenden, M.C., Villamizar, M.L., Wearing, C., Withers, P.J., Zhou, J.G., Barber, N.J., Haygarth, P.M., 2018. A method for uncertainty constraint of catchment discharge and phosphorus load estimates. Hydrol. Process. 32, 2779–2787. https://doi.org/10.1002/hyp.13217.

ISO 1100–2, 2010. Hydrometry measurement of liquid flow in open channels–Part 2: determination of the stage-discharge relationship. Geneva, Switzerland: International Organization for Standardization.

Kleinhans, M.G., Cohen, K.M., Hoekstra, J., IJmker, J.M., 2011. Evolution of a bifurcation in a meandering river with adjustable channel widths, Rhine delta apex, The Netherlands. Earth Surface Processes and Landforms, 36, 2011–2027. doi: 10.1002/esp.2222.

Kok, M., Jongejan, R., Nieuwjaar, M., Tanczos, I., 2017. Fundamentals of Flood Protection. Technical Report Ministery of Infrastructure and the Environment and Expertise Network for Flood Protection (ENW) Breda, Netherlands.

Lang, M., Pobanz, K., Renard, B., Renouf, E., Sauquet, E., 2010. Extrapolation of rating curves by hydraulic modelling, with application to flood frequency analysis. Hydrol. Sci. J. 55, 883–898. https://doi.org/10.1080/02626667.2010.504186.

Le Coz, J., Renard, B., Bonnifait, L., Branger, F., Le Boursicaud, R., 2014. Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: a Bayesian approach. J. Hydrol. 509, 573–587. https://doi.org/10.1016/j.jhydrol.2013.11.016.

Mansanarez, V., Renard, B., Le Coz, J., Lang, M., Darienzo, M., 2019. Shift happens! Adjusting stage-discharge rating curves to morphological changes at known times. Water Resour. Res. 55, 2876–2899. https://doi.org/10.1029/2018WR023389.

McMillan, H., Seibert, J., Petersen-Overleir, A., Lang, M., White, P., Snelder, T., Rutherford, K., Krueger, T., Mason, R., Kiang, J., 2017. How uncertainty analysis of streamflow data can reduce costs and promote robust decisions in water management applications. Water Resour. Res. 53, 5220–5228. https://doi.org/10.1002/2016WR020328.

Moyeed, R.A., Clarke, R.T., 2005. The use of Bayesian methods for fitting rating curves, with case studies. Adv. Water Resour. 28, 807–818. https://doi.org/10.1016/j.advwatres.2005.02.005.

Ocio, D., Le Vine, N., Westerberg, I., Pappenberger, F., Buytaert, W., 2017. The role of rating curve uncertainty in real-time flood forecasting. Water Resour. Res. 53, 4197–4213. https://doi.org/10.1002/2016WR020225.

Pappenberger, F., Matgen, P., Beven, K.J., Henry, J.-P., Pfister, L., De Fraipont, P., 2006. Influence of uncertain boundary conditions and model structure on flood inundation predictions. Adv. Water Resour. 29, 1430–1449. https://doi.org/10.1016/j.advwatres.2005.11.012.

Peña Arancibia, J.L., Zhang, Y., Pagendam, D.E., Viney, N.R., Lerat, J., Van Dijk, A.I.J.M., Vaze, J., Frost, A.J., 2014. Streamflow rating uncertainty: characterisation and impacts on model calibration and performance. Environ. Model. Softw. 63, 32–44. https://doi.org/10.1016/j.envsoft.2014.09.011.

Perret, E., Renard, B., Le Coz, J., 2021. A rating curve model accounting for cyclic stage-discharge shifts due to seasonal aquatic vegetation. Water Resour. Res. 57 https://doi.org/10.1029/2020WR027745.

Rantz, S.E., 1982. Measurement and Computation of Streamflow: Volume 1. Measurement of Stage and Discharge. Technical Report United States Geological Survey Washington, D.C., USA.

Salvatier, J., Wiecki, T.V., Fonnesbeck, C., 2016. Probabilistic programming in Python using PyMC3. PeerJ Comput. Sci. 2 https://doi.org/10.7717/peerj-cs.55 url:https://github.com/pymc-devs/pymc3.

Sebok, E., Refsgaard, J.C., Warmink, J.J., Stisen, S., Jensen, K.H., 2016. Using expert elicitation to quantify catchment water balances and their uncertainties. Water Resour. Res. 52, 5111–5131. https://doi.org/10.1002/2015WR018461.

Sikorska, A.E., Renard, B., 2017. Calibrating a hydrological model in stage space to account for rating curve uncertainties: general framework and key challenges. Adv. Water Resour. 105, 51–66. https://doi.org/10.1016/j.advwatres.2017.04.011.

Sörengård, M., Di Baldassarre, G., 2017. Simple vs complex rating curves: accounting for measurement uncertainty, slope ratio and sample size. Hydrol. Sci. J. 62, 2072–2082. https://doi.org/10.1080/02626667.2017.1367397.

Steinbakk, G.H., Thorarinsdottir, T.L., Reitan, T., Schlichting, L., Hølleland, S., Engeland, K., 2016. Propagation of rating curve uncertainty in design flood estimation. Water Resour. Res. 52, 6897–6915. https://doi.org/10.1002/2015WR018516.

Ylla Arbós, C., Blom, A., Acevedo Goldaracena, F., Van Vuren, S., Schielen, R.M.J., 2020. Bed level change in the Upper Rhine Delta and Niederrhein. In: Uijttewaal, W., Franca, M.J., Valero, D., Chavarrias, V., Ylla Arbós, C., Schielen, R.M.J., Crosato, A. (Eds.), River Flow 2020; Proceedings of the 10th Conference on Fluvial Hydraulics (pp. 680–684). Delft; Netherlands.