Predicting discharge capacity of vegetated compound channels: uncertainty and identifiability of 1D process-based models

Adam Kiczko[1], Kaisa Västilä[2,3], Adam Kozioł[1], Janusz Kubrak[1], Elżbieta Kubrak[1], and Marcin Krukowski[1]

[1]Warsaw University of Life Sciences – SGGW, Institute of Environmental Engineering
[2]Department of Built Environment, Aalto University School of Engineering, Espoo, Finland
[3]Freshwater Centre, Finnish Environment Institute, Helsinki, Finland

Correspondence: Adam Kiczko (adam_kiczko@sggw.pl)

Abstract. Despite the development of advanced process-based methods for estimating the discharge capacity of vegetated river channels, most of the practical one-dimensional modeling is based on a relatively simple divided channel method (DCM) with the Manning's flow resistance formula. This study is motivated by the need to improve the reliability of modeling in practical applications while acknowledging the limitations on the availability of data on vegetation distributions and densities required by the process-based methods. We investigate whether the advanced methods can be applied to modeling vegetated compound channels by identifying the missing characteristics as parameters through the formulation of an inverse problem. We developed a new probabilistic approach for comparing six models of channel discharge capacity in respect of their uncertainty, with the model with the lowest uncertainty considered the most favorable. Calculations were performed for flume and field settings varying in floodplain vegetation submergence, density, and flexibility, and in hydraulic conditions. The output uncertainty, estimated on the basis of a quasi-Bayes approach, was analyzed for a varying number of observation points, demonstrating the significance of the parameter equifinality. The results showed that very reliable predictions with low uncertainties can be obtained for process-based methods with a large number of parameters. The equifinality affects the parameter identification but not the uncertainty of a model. The best performance for sparse, unsubmerged, rigid vegetation was obtained with the Mertens method and for dense, flexible vegetation with the generalized two-layer method combined with a description of the flexibility-induced reconfiguration. We found that the process-based methods are superior when applied for vegetative conditions they were developed for while the Manning based DCM seems to be the most flexible technique.

1 Introduction

Compound channels consisting of a main channel and vegetated floodplains are commonly observed both in natural and engineered settings. For instance, vegetated compound (two-stage) channels have been recently proposed as an environmentally preferable alternative to conventional dredging in flood and agricultural water management (e.g. Västilä and Järvelä, 2011).

Such nature-based solution (NBS) allow combining the technical needs, e.g. flow conveyance and stability, and the environmental requirements, e.g. improved water quality and biodiversity (Rowiński et al., 2018). Reliable predictions on the discharge capacity in such complex channels are required e.g. for river restoration, flood analyses, and for implementation of

25 nature-based solutions. Predictions on the discharge capacity using the conventionally applied methods (e.g. Posey, 1967) can be underpinned by high uncertainties caused by the complex cross-sectional geometry and the composite roughness resulting from regions with highly different flow resistance. Floodplain vegetation is the main factor complicating the predictions and causing uncertainty, particularly in small to medium-sized channels where up to 90 per cent of the flow resistance can be caused by plants (e.g. Västilä et al., 2016). Thus, there are needs to analyze the benefits and limitations of using more advanced

30 state-of-the-art methods for describing the influence of floodplain vegetation in practical engineering applications.

Since its formulation in 1960, the Divided Channel Method (DCM, Posey, 1967) with the Manning formula is the most widely used technique for predicting the discharge capacity in compound channels in one-dimensional flow routing models. This simple approach explains the flow separately in zones with differing flow resistance, usually the main channel and floodplains. The kinematic effect resulting from the momentum exchange between areas of different velocities is taken into account

35 through rough imaginary walls (Sellin, 1964; Kubrak et al., 2019b, a). Despite the development of more advanced methods, providing often much more detailed and physically based description of channel flow resistance, DCM is till this day found in the majority of practical models for flood hazard assessments, design of hydraulic structures or water management.

The limitations of the DCM are well-known. As it does not account for the momentum transfer between the main channel and floodplains, it overestimates main channel flows and as a result the total discharge (Myers, 1978). Furthermore, values of

40 the Manning coefficients, used to parameterize flow resistance, depend on the flow rate (Fread, 1989; Soong and DePue, 1996). It should be noted, that this relationship can be amplified with inadequacy of a flow model, as mentioned by Yen (1999).

A number of studies were devoted developing a more process-based description of channel flows (Yen, 2002). The attempts includes either, flow processes itself or interactions with obstacles, mostly in a form of vegetation. In the first category, the most sophisticated model of the channel capacity should be attributed to Shiono and Knight (1991), who on the basis of a turbulent

45 flow theory, derived equations for depth averaged velocities in the cross-section plane. Accompanied with an additional drag term, the method was successfully used to model flow in a channel with composite roughness consisting of vegetated and non-vegetated zones (Kalinowska et al., 2020).

Other scientific efforts focused on developing more physically-based description of the interactions between flow and vegetation. One of the most important methods was given by Pasche (Pasche and Rouvé, 1985). It is based on the detailed physical

50 description of the flow in zones with unsubmerged (emergent) vegetation. The kinematic effect is simulated with a rough imaginary wall, as in the DCM. Here however the wall friction depends implicitly on the flow velocities in the main channel and vegetated areas. In the original form the Pasche method consists of several implicit equations making it hard to implement. Simplified version of the method was proposed by Mertens (1989).

Recent progress in the modeling of compound channels aimed at developing methods for submerged and flexible vegetation

55 that undergoes marked streamlining and reconfiguration, reducing the apparent flow resistance Jalonen et al. (2015). A straight-

forward two-layer method was proposed by Luhar and Nepf (2013) while attempts to generalize the approach and provide basis for its parametrization are made by e.g. Västilä and Järvelä (2014); Jalonen and Järvelä (2015); Västilä and Järvelä (2018).

The most obvious reason why the process-based methods are unpopular for in practical applications is their complexity. More detailed representation of the flow processes increases the amount of necessary information on channel properties. Methods
60 like Pasche and Rouvé (1985) or Luhar and Nepf (2013) as well as the state-of-the-art vegetation description Västilä and Järvelä (2018) require specific data on vegetation, such as density, spacing, shape or species, and leaf area indices. In practical assessments for tens of kilometers long river reaches, such information is not readily available. With these practical limitations, the use of a roughness coefficient lumping all effects, such as the Manning coefficient, can be a reasonable solution (i.e. Marcinkowski et al., 2018, 2019).

65 The second, also important argument for simpler methods is that the roughness coefficients are often treated as "catch all parameters". It is a common practical approach of using the resistance parameters for adjusting the model fit to observations. In the result such parameters are often used beyond their physical interpretation, as disused by Yen (1999) in the response to Khatibi et al. (1997). Fr instance, higher or lower values of the Manning coefficient compared to the physical interpretation of the roughness could be used when bathymetric data do not account for the true complexity of the river geometry. One of such
70 examples is the modeling of steep-pool reaches with multiple sub- supercritical flow transitions by using very high values of roughness (Reid and Hickin, 2008).

Despite the recent developments of process-based methods for modeling vegetated flows, there is a lack of knowledge in whether the state-of-the-art methods with a significant number of parameters are reliable in common practical applications characterized by insufficient data on channel and vegetative properties. Any method can be widely applied only if all its
75 parameters can be identified as the solution to the inverse problem – a parameter calibration. This leads to an old dilemma, where a simple model with limited number of parameters is compared with a complex one with more parameters (Kuczera and Mroczkowski, 1998). The answer can be given in terms of the models predictive uncertainty. The better method is the one which for the same number of data points has a lower predictive uncertainty (Her and Chaubey, 2015; Her and Seong, 2018). As one of the first works, this paper evaluates the uncertainty of chosen state-of-art methods for predicting the discharge
80 capacity (understood as the dependency between water level and discharge) in vegetated compound channels, comparing to the Manning DCM (Posey, 1967). We have limited the scope to one-dimensional models, which despite development of two-dimensional ones, are still a practical solution for long river reaches. Most of previous studies on the channel discharge capacity were performed neglecting the problem of the parameter identification and associated uncertainty (Helmiö, 2002, 2005).

The first study that addressed the problem of the uncertainty associated with the parametrization of floodplain roughness was
85 Warmink et al. (2013) who compare the uncertainty of a 2D model for chosen methods of bed and vegetation resistance. The uncertainty analysis had an "explicit" form, as it was assumed that uncertainty sources were parameterized in a sense of their distributions. Outcomes confirmed that the uncertainty related to the vegetation and bed forms roughness is one of the most important sources of the uncertainty in flood routing. Interestingly from the point of the present study, Warmink et al. (2013) did not find a choice of a resistance formula as an important uncertainty factor. By contrast, Zinke et al. (2011) concluded that
90 for morphological modelling that the uncertainty is mainly related to the vegetation parameterization.

A similar approach, in terms of "explicit" uncertainty assessment was presented by Dalledonne et al. (2019), who compared several methods for vegetation induced resistance, including: Pasche (Pasche and Rouvé, 1985) and Järvelä Järvelä (2004) in a two-dimensional model. Here however, in terms of uncertainty, there were noticeable differences between the analyzed techniques, with the Järvelä approach found the most favorable.

95 In the study uncertainty is represented using the implicit method of of Generalized Likelihood Uncertainty Analysis (GLUE, Beven and Binley, 1992, 2014). As shapes of parameter distributions are estimated through an inverse problem using the Bayes theorem, our results have the benefit that they are less affected by the initial assumption on the parameter variability compared to the "explicit" approaches.

The overall goal of this paper is to compare the uncertainty, parameter identifiability and physical interpretation of the 100 parameters of discharge capacity methods characterized with different levels of parameterization. This work focuses on one-dimensional methods for compound channels with a significant share of the flow resistance generated by vegetation. The following methods were investigated: Manning based DCM, Pasche (Pasche and Rouvé, 1985) and Mertens (1989) methods designed for emergent rigid vegetation, and three versions of the two-layer model proposed by Luhar and Nepf (2013) as modified by Västilä and Järvelä (2018), designed for flexible submerged vegetation. All models were applied to vegetation 105 conditions differing in relative submergence (covering both submerged and emergent conditions) and density, as motivated by real cases where it is possible that e.g. a "rigid" vegetation model is applied for flexible vegetation because of lack of information on the vegetation properties. Parameter identification was conditioned on water depths instead of discharges to make the problem more similar to practical cases, like flood assessments, where a model outcome is usually the water level. It is out of the scope of the paper to provide a summary of all the available methods.

2 Methods

2.1 Parameter identification and uncertainty analysis

There are two approaches for parameter identification (Figure 1), of which the conservative approach is typically used. In a typical engineering task, the resistance term is usually poorly recognized and in the DCM Manning roughness coefficients are identified as an inverse solution, ensuring the best possible fit of the modeled and observed water levels or inundation 115 extents. For the more process-based methods, the conservative approach considers most of parameters, such as the vegetation properties, as an input. The model identification applies then only to several minor values, like surface roughness, as illustrated in Figure 1a. Instead of the conservative approach, this study considers additional inputs in terms of parameter identification, which is beneficial for being able to apply the advanced methods despite limited data (Figure 1b). For instance, in real applications detailed information on e.g. floodplain cover is not readily available, while for the DCM approaches it is sufficient, 120 at most, to know if dense vegetation is present or not. Thus, the concept of this study is to consider all additional inputs of the process-based methods as parameters that have to be identified. Such an approach was previously presented by Kiczko et al. (2017) and revealed that process-based methods of Pasche and Mertens can be applied in the same manner as DCM. The
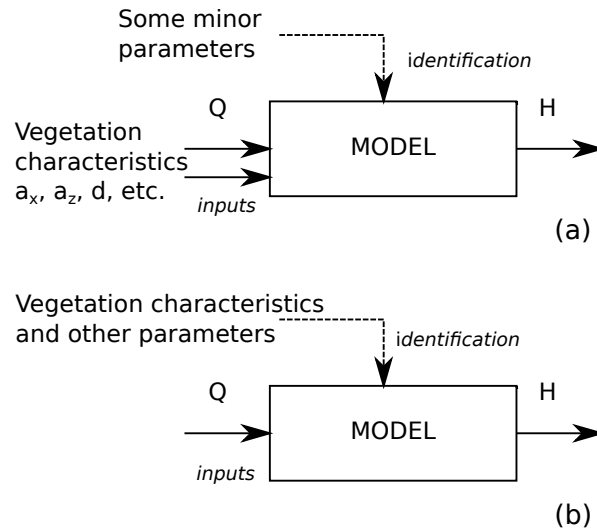
Figure 1. Parameter identification problem for process-based methods of channel discharge: (a) conservative approach, (b) adopted in the present study.

technical difference between different the methods is just the number of parameters. Identifiability of model parameters and the quality of the obtained solution determine then the applicability of a method.

125   The parameter identification problem was defined using Generalized Likelihood Uncertainty Estimation (GLUE) approach (Beven and Binley, 1992; Romanowicz and Beven, 2006), based on Bayes formula:

$$P\left(\theta/H\right) = \frac{L\left(H/\theta\right)P\left(\theta\right)}{\int L\left(H/\theta\right)P\left(\theta\right)d\theta} \qquad (1)$$

where θ stands for parameters, H water levels, $P\left(\theta\right)$ *a priori* parameter distribution, $P\left(\theta/H\right)$ *a posteriori* parameter distribution, $L\left(H/\theta\right)$ likelihood function.

130   In the practical application usually there is no strong theoretical grounds for the assumption of *a priori* parameter distribution and the shape of the likelihood function. In that case, as in the present study, a uniform/rectangular distribution is usually used. The likelihood function, necessary to transform *a priori* into *a posteriori* distribution, in the original GLUE approach is used along with selecting so-called behavioral simulations, above the given level of fit measures. This allows to adjust the variation of the estimated uncertainty. In the present study the Gauss shaped function was used, where output uncertainty depends on

135 the scaling factor κ (Romanowicz and Beven, 2006):

$$L\left(H/\theta\right) = \exp\left[\frac{-\sum_{i=1}^{n}\left(H_i - \hat{H}_i\right)^2}{\kappa\sigma^2}\right] \qquad (2)$$

with n standing for the number of observation points used in the parameter identification, σ^2 variation of model residuals, $\hat{H}$ and H observed and calculated water levels, respectively.

Determining the span of the model uncertainty variance is always an important part of the uncertainty estimation. A good
140 uncertainty model ensures that a desired number of observations is enclosed within uncertainty intervals (Blasone et al., 2008).
This is particularly important in the present study, where different methods are compared in respect of their uncertainty.
Confidence intervals should be sufficiently wide to cover the required number of observations but not wider. This condition
can be fulfilled with a sufficient variability range of model parameters, specified as *a priori* distribution $P(\theta)$ and appropriate
shape of the likelihood function $L(Y/\theta)$ depending here on the κ coefficient.

145 Parameter ranges can be usually found by trials-and-errors while the shape of the likelihood function should be determined
in respect of observations. In the present study, the shape coefficient κ (Equation 2) was computed on the basis of minimization
task:

$$\kappa = \arg\min_{\kappa} \left[\epsilon\kappa + \left| p - \frac{1}{n}\sum_{i=1}^{n} J\left(\hat{H}_i\right) \right| \right] \tag{3}$$

$$J\left(\hat{H}_i\right) = \begin{cases} 0 & \text{if} \quad \hat{H}_i \in [H_i^{q_L}, H_i^{q_u}] \\ 1 & \text{else} \end{cases} \tag{4}$$

150 where $H_i^{q_L}$, $H_i^{q_U}$ denote lower and upper quantile (q_L, q_U) of the calculated water levels from the *a posteriori* distribution
(Equation 1), obtained with the likelihood function (Equation 2); p stands for confidence interval, defined as: $p = q_U - q_L$. In
the present study 95% confidence intervals ($p = 0.95$) were used, with $q_L = 0.025$ and $q_U = 0.975$. ϵ is a small number as a
penalty for too wide confidence intervals of water levels H. The minimum of the function 3 should be the smallest value of κ
for which the last term in Eq. 3 equals zero:

$$155 \quad p - \frac{1}{n}\sum_{i=1}^{n} J\left(\hat{H}_i\right) \leq 0 \tag{5}$$

This is true when exactly n observations fall within the confidence intervals. For $p = 0.95$ and relatively small observation sets
of $n \sim 10$ in the present study, minimum is found when all observations are enclosed by intervals. In such a case, the sum term
is equal to 1 and the difference becomes negative. It should be noted, that for a poor model and/or inappropriate variability
ranges of its parameters, such a solution might not exist. Therefore it was necessary to control the solution of the minimization
160 in respect of Equation 5. If the constraint was not fulfilled, it was necessary to revise assumptions on the *a priori* parameter
distribution $P(\theta)$. For a poor model, a solution fulfilling the constraints of Eq. 5 might not exist within parameter ranges that
can be interpreted in terms of their physical characteristics. In such case, the model was considered as unidentifiable, i.e.,
inadequate for a given data set.

It is acknowledged that the parameter identification and associated uncertainty depend on the size of the observation data
165 set. To address this issue, the parameter identification (Eq. 1) was performed for a varying number of observation points
$n = 1, \ldots, N$, where N stands for the total size of a data set. The calculations include all possible combinations of observations
with the given n i.e. $\frac{N!}{n!(N-n)!}$. The total number of all combinations is then $2^N - 1$, excluding the empty set. Such an approach
allows eliminating the effect of non-representative observation samples. The method was discussed previously by Kiczko et al.
(2017).

170   Identification is considered as successful, if all n points used in the model identification are enclosed by confidence intervals. Remaining observation points $N-n$ act as a verification set. In this analysis, both the proportion of verification points that falls within estimated confidence intervals and the width of confidence intervals are used as measures of model performance. The more narrow the confidence bands and the less observation points falling outside them, the better a model is. On the opposite, a less adequate model requires larger spread of the solution, to enclose observations, as it wrongly explains their variability.

175   Because the different combinations of n points resulted in multiple uncertainty estimates, the results were presented in terms of statistical moments, as a function of n. For a detailed description of results box-plots were used, where the median is given as a horizontal line within a box, that spans over 25% and 75% quantile, whiskers indicate the result extent, excluding extreme values given with cross marks.

2.2 Discharge capacity formulas

180 2.2.1   Divided Channel Method

In the DCM approach (Posey, 1967), the channel cross section is divided in flow zones of similar hydraulic conditions, typically the main channel and floodplain. The interactions between the zones of significantly different mean velocities are reproduced with a rough imaginary wall, applied to the zone with the higher velocity, i.e. the main channel. In the present study, the roughness of the interface was assumed to equal the roughness of the channel banks next to the interface. Parameters of the

185   method are the roughness coefficients for each flow zone. In the present study, DCM was based on the Manning formula, with the common approach of having separate Manning coefficients for the main channel (n_c), and left (n_L) and right floodplain (n_R).

2.2.2   Pasche and Mertens methods

A brief concept of the Pasche method is provided by Pasche and Rouvé (1985) and a detailed description of the algorithm used

190   herein is provided in Kozioł et al. (2004). The model describes the discharge capacity of the compound cross section with rigid vegetation, derived for steady flow conditions. Similarly to DCM, the model divides the compound cross-section into regions of the main channel and floodplains, dominated by bottom and vegetation roughness, respectively. It accounts additionally for the transition region between these two main zones. As in the DCM, the interactions between the main channel and floodplains are modeled using an imaginary rough wall. For the resistance the Darcy-Weisbach formula is used.

195   The Darcy-Weisbach friction coefficients are determined using a set of semi-empirical equations for each zone and the imaginary wall, including transitional regions. The method explains the extent of the transition region within the vegetated region, affected by the higher flow velocity of the unvegetated main channel. The flow in the main channel depends on the apparent resistance of the imaginary wall. There is no general expression for the span of the transition region in the main channel, and it has to be established for each case.

200   Velocities in the flow zones and transitional regions are interrelated by the apparent resistance. Equations describing these dependencies have an implicit form that requires iterative methods for solving, so that the Pasche method has a very complex

numerical solution. Mertens (1989) attempted to improve the numerical efficiency of the Pasche concept by simplifying most of the demanding implicit formulas to less accurate but explicit ones, reducing the number of terms requiring iterative numerical solving.

205 In the Pasche and Mertens methods, a detailed parametrization of the channel, including plant properties, surface roughness and the extent of the interaction zone in the main channel, is used. Assuming that the modeler has only knowledge on the geometry of the cross-section, the following parameters have to be identified: a_x, a_y, longitudinal and horizontal spacing of plant stems; d_p average diameter of the stems; k_f, k_c roughness height of the floodplain and the main channel bed; b_{III}/B_c ratio of the interaction region width in the main channel (b_{III}) to the main channel width (B_c). Assuming that the channel is

210 symmetric, the total number of parameters is six. Modeling different properties of vegetation on left (subscript L) and right (subscript R) floodplains ($a_{x,L}$:$a_{x,R}$, $a_{z,L}$:$a_{z,R}$, $d_{p,L}$:$d_{P,R}$, $k_{f,L}$:$k_{f,R}$) increases the number of parameters up to ten.

2.2.3 Generalized and Simplified Two-Layer Model

In the present study, the two layer model of Luhar and Nepf (2013), generalized by Västilä and Järvelä (2018) is considered as the state-of-art approach for submerged vegetation. This Generalized Two-Layer Model (GTLM) is based on the momentum

215 balance with drag coefficients at the interfaces between vegetated and unvegetated areas of the channel cross section. Generalization proposed to the original model (Luhar and Nepf, 2013) by Västilä and Järvelä (2018) consists in replacing the channel width by the wetted perimeter (P) and water depth by the hydraulic radius (R).

 The channel discharge capacity is computed on the basis of equations for mean velocities in the unvegetated (u_0) and vegetated (u_v) parts of the cross section (Västilä and Järvelä, 2018):

$$\frac{u_0}{(gSR)^{1/2}} = \left[\frac{2P(1-B_X)}{C^*(L_b+L_v)}\right]^{1/2} \tag{6}$$

$$\frac{u_v}{(gSR)^{1/2}} = \left[\frac{2PB_X+C^*L_v(u_0^*)^2}{C_D a P R B_X}\right]^{1/2} \tag{7}$$

where g is the gravitational constant, S energy slope, $u_0^* = \frac{u_0}{(gSR)^{1/2}}$ dimensionless velocity in unvegetated zone, C^* the drag coefficient for shear stresses at the channel bed and at the interface between vegetated and unvegetated zones, L_b and L_v wetted lengths of the unvegetated channel margin and of the interface between vegetated and vegetated zones, respectively.

225 B_X denotes the vegetative blockage factor in the cross section, defined as the vegetated flow area divided by a total flow area. Physically, the drag coefficients for bed and the vegetation zone interface may take separate values. Following Luhar and Nepf (2013); Västilä and Järvelä (2018), it was herein assumed that the same value of C^* can be used for both regions.

 $C_d a$ is the vegetative drag per unit water volume, expressed conventionally as the product of a drag coefficient C_d and the frontal projected plant area per unit water volume a, assuming that plants are rigid simple-shaped objects. To account for

230 the presence of foliage and the flexibility of the plants inducing bending and streamlining, the vegetative drag per unit water
volume can be parameterized as (Västilä and Järvelä, 2018)

$$C_D a = C_{D_X,F} \left(\frac{u_C}{u_{X,F}} \right)^{\chi_F} \frac{A_L}{A_B h} + C_{D \chi S} \left(\frac{u_C}{u_{XS}} \right)^{\chi_S} \frac{A_S}{A_B h} \tag{8}$$

where u_C is a characteristic approach velocity, taken here as equal to the velocity in a vegetation layer: $u_C \approx u_v$. A_S denotes
total frontal projected areas of the plant stems and A_L the total one sided leaf area per unit ground area A_B. $C_{D_X,S}$ and $C_{D_X,F}$

235 represent constant coefficients for the drag of stems and foliage, respectively. The effect of streamlining and reconfiguration on
the drag is described using exponents χ_S and χ_F, for stems and foliage, respectively. $u_{X,F}$ and $u_{X,S}$ are reference velocities
needed for determining the drag and reconfiguration coefficients.

Equations 6 and 8 implicitly depend on each other and require numerical solving. In the conservative approach vegetation
parameters have to be known (Figure 1 a). The blockage factor B_X requires knowledge on the vegetation distribution and/or

240 height in the cross section. $\frac{A_S}{A_B}$ and $\frac{A_L}{A_B}$ ratios characterizing the plant structure can be measured or typical values for a certain
plant communities can be adopted. Drag coefficients $C_{D_X,S}$, $C_{D_X,F}$ and reconfiguration exponents χ_S and χ_F, along with
their reference velocities ($u_{X,F}$ and $u_{X,S}$), are plant species or plant type-specific factors and can be determined on the basis
of laboratory measurements. Their values have been published for common plant species (Västilä and Järvelä, 2014; Jalonen
and Järvelä, 2015; Västilä and Järvelä, 2018).

245 For channel flows with dense vegetation for which over 80 percent of the discharge is conveyed in the unvegetated regions,
the GTLM approach can be simplified by assuming that discharge in the vegetation layer is negligible with respect to the total
discharge: $u_v \approx 0 \, m/s$ (Luhar and Nepf, 2013; Västilä et al., 2016). The remaining Equation 6 does not require numerical solv-
ing. In the present study the above approach is referred as Simplified Two-Layer Model (STLM). By neglecting the Equation
7, the STLM requires five and GTLM nine parameters.

250 Parameters of GTLM and STLM, resulting from Equation 6 are the drag coefficient for shear stresses C^* and Blockage
Factor B_X. B_X depends on the area occupied by the vegetation in the cross section. It changes with the water level and
therefore should not be represented as a constant value but rather derived on the basis of the cross section and vegetation
geometric properties. In the present study B_X was described in terms of left-right extents l_L/L_L, l_R/L_R and the height h_L,h_R
of vegetation. L_L, L_R stand for the cross section width on the left and right side, respectively. l_L/L_L provides information of

255 the vegetation extent on the left side, starting form the top of the left bank towards the channel middle point: 0 stands for clean
bank, while 1 means that the vegetation cover extends over entire left side. Same applies for l_R/L_R, where it is assumed that
vegetation zones starts from the top of the right bank. The vertical range of the vegetation in the cross section is obtained by
adding h_L or h_R to the value of the ground elevation.

It should be noted, that by parameterizing the Blockage Factor, the parameter identification task is much more complicated

260 than in the conventional approaches. In the DCM the vegetation extent is equivalent to the division into main channel and
floodplains, which is known on the basis of the cross sectional geometry. Here, for GTLM and STLM it was considered as a
part of the parameter identification problem.

2.2.4   Practical Two-Layer Model

Luhar and Nepf (2013) proposed a formula for the Manning coefficient n:

$$265 \quad n\left(\frac{g^{1/2}}{KR^{1/6}}\right) = \frac{(gSR)^{\frac{1}{2}}}{U} = \left[\left(\frac{2}{C^*}\right)^{\frac{1}{2}}\left(1-\frac{h}{R}\right)^{\frac{3}{2}} + \left(\frac{2}{c_D ah}\right)^{\frac{1}{2}}\left(\frac{h}{R}\right)\right]^{-1} \tag{9}$$

where h stands for the vegetation height and $K = 1\ m^{1/3}s^{-1}$ to ensure correct dimensions of the equation. The formula is derived for shallow channels, lined with vegetation, where the blockage factor was approximated as $B_X \approx \frac{h}{H}$. In the presented form of the equation (9), following Västilä and Järvelä (2014), the water depth H was replaced with the hydraulic radius R.

The Luhar and Nepf (2013) formula 9 has a convenient form to be easily applied in practical cases, where usually the
270 Manning equation is used. In the present study this approach is named Practical Two-Layer Model (PTLM) and applied as a three parameter one, with the drag coefficient C^*, average vegetation height h in the cross section and $C_D a$.

2.3   Case studies

The analyses were conducted for a flume data set (Koziol, 2010) and a field data set (Västilä et al., 2016) collected from vegetated compound channels, interpreted herein as 5 distinct case studies, as detailed below. To our knowledge, the field
275 cases are one of the most thorough characterizations on the dependency between vegetation properties and discharge capacity in natural compound channels, including spatially-averaged values for vegetation height, blockage factor, and frontal area density in different seasons and flow conditions. The flume cases are representative of typical experimental arrangements where vegetation is simulated by rigid cylindrical elements at a uniform spacing.

2.3.1   Flume experiments.

280 The experiments were conducted at the Warsaw University of Life Sciences (WULS-SGGW) using a physical model of a compound channel with rigid cylinders simulating vegetation. A detailed description of the dataset can be i.e. found in Koziol (2010).

The modeled channel was straight, 16 m long with the compound trapezoidal cross section, $2.10\ m$ wide (Figure 2). The main channel bottom was made of smooth concrete with the estimated roughness height $k_s = 5 \cdot 10^{-5}\ m$. Floodplain vegetation
285 was simulated with rigid cylinders of a diameter $d_p = 0.008\ m$ and spacing $a_x = a_y = 0.1m$. There were two experimental variants of vegetation layout and floodplain roughness. In the first one (1) the floodplain bottom was made of the same smooth concrete as the main channel, with a single row of vegetation present also on channel bank (Figure 2a). In the second one (2), vegetation was constrained on the floodplain by removing the channel bank stems while floodplain surfaces were made rougher using a layer of terrazzo concrete of the grain size of 0.5 to $1\ cm$ (Figure 2b).

290 Experiments were performed for steady and uniform flow conditions. The water surface was kept parallel to the bed using a weir localized at the flume outflow. Water discharge was measured using a circular weir and water levels were recorded in the middle of the channel.

The data set, used in the present study, consisted of discharge and water level observations within the range of: 0.037-0.060 m^3/s (mean velocities: 0.2-0.4 m/s) and 0.2 - 0.3 m, respectively. The number of observation point in the first variant was nine ($N = 9$) and in the second one ten ($N = 10$).
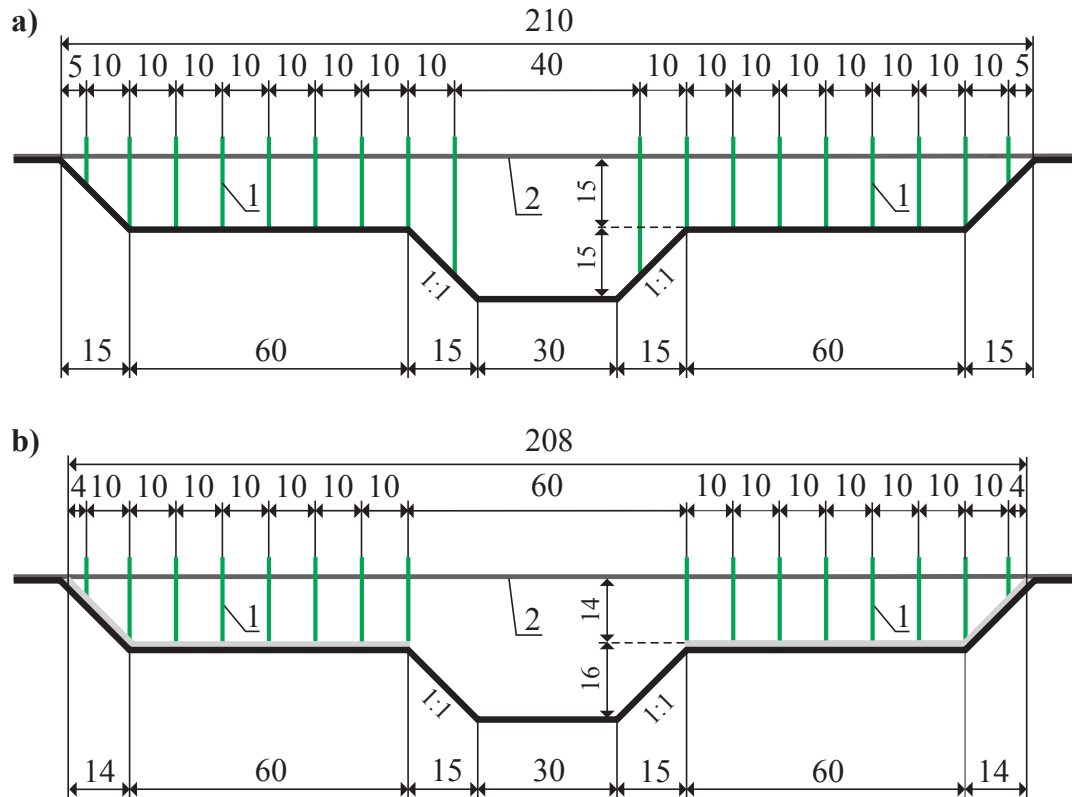
Figure 2. Laboratory channel cross section (dimensions in cm); 1 - rigid cylinders simulating vegetation; 2 - wooden strips supporting vegetation (Koziol, 2010); a) case 1; b) case 2.

2.3.2 Ritobacken field experiment

The field data was obtained from an 11 m wide compound channel, Ritobacken Brook (Finland, Figure 3), with seasonally and annually varying vegetation properties. The site and data are described by Västilä et al. (2016). Measurement series with vegetated floodplain flows were available for three seasons, with the number of observations given in brackets: Spring 2011 ($N = 6$), Autumn 2011 ($N = 12$) and Spring 2012 ($N = 11$). The respective mean floodplain vegetation heights were $h = 9$ cm, 47 cm and 24 cm while the vegetative blockage factor ranged at $B_X = 0.13 - 0.53$. The taller vegetation in Spring 2012 compared to Spring 2011 was explained by the fact that the floodplain was excavated in February 2010 and thus vegetation was under succession phase. Vegetation was submerged under all examined flows in Spring 2011 and under 42% and 64% of the flows in Autumn 2011 and Spring 2012, respectively.

305   The Manning coefficient of the narrow main channel as obtained from highest flows not inundating the floodplain was $n = 0.08 - 0.12\ m^{-1/3}s$ due to irregular main channel geometry, woody debris and some aquatic vegetation. For Autumn 2011 Västilä and Järvelä (2014) estimated values of plant drag coefficients and reconfiguration exponents: $C_{dx,F} = 0.14$, $C_{dx,S} = 0.93$, $\chi_F = -1.11$, $\chi_S = -0.26$.

The discharge capacity at different flow conditions was obtained from water level data recorded at 5-15 min intervals with
310   pressure transducers at the upstream and downstream ends of a 190 m long test reach. The discharge was obtained from a rating curve determined for a culvert at the downstream end of the test reach. At floodplain flows, discharge and floodplain water depth ranged at $0.19 - 1.59\ m^3/s$ and $0.10 - 0.67\ m$, respectively, with cross-sectional mean velocities of $0.11 - 0.30\ m/s$. Flow conditions were non-uniform, and in computations the measured energy grade slope was used.
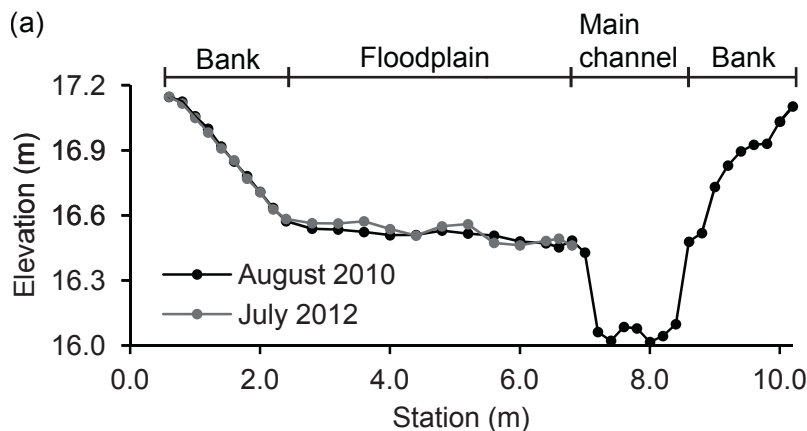
Figure 3. Ritobacken channel cross section (a) and a photography, Autumn 2011 (b)

3 Results

315   The size of the Monte Carlo sample (n_{mc}) was determined in each case by trial and error, to satisfy the convergence of the solution. In a similar way, the ranges of parameters for the *a priori* distributions $P(\theta)$ (Equation 1) were determined. It was done in respect of the parameter physical variability, by ensuring that observations will be enclosed by confidence intervals (Equation 5). Parameter bands with n_{mc} Monte Carlo sample sizes are provided in Tables 1 and 2, separately for flume and Ritobacken field experiments. Parameter ranges were often defined outside bands expected in the nature, to allow fulfilling the

320   constrain 5 by as many as possible models. For flume data sets calculations were performed for a symmetric channel, which allowed to reduce the number of parameters, as the same values were used for the left and right floodplain.

The numerical results were analyzed from three perspectives: (1) identifiability of the model for the given vegetation conditions; (2) width of estimated confidence intervals as a function of the number of the observation points; (3) the physical interpretation of the obtained parameter values.

325   ### 3.1   Computational output and general observations

The basic output of the computations which included Monte Carlo simulations using channel discharge models and parameter identification on the basis of Equations 1-4, were rating curves. They were derived with a different number of observation points n for the parameter identification, for all possible combinations (see Section 2.1).

Exemplary curves are presented to highlight some general observations (Figure 4). We show chosen solutions for a given

330   number of observation points used in the parameter identification, $n = 5$ for the two-layer approaches (GTLM, STLM, PTLM in Figure 4a-4c) developed for dense, submerged vegetation corresponding to the Ritobacken case study and $n = 4$ for the Pasche, Mertens and Manning based DCM models for rigid unsubmerged vegetation corresponding to the flume conditions (Figure 4d-4f). In this example, the parameters for discharge curves were identified at low flows, while the verification was conducted for high flows, which represents the common practical way of using hydraulic models to assess flood hazard at flows

335   higher than the ones the models were calibrated with. In terms of parameter identification results are considered as successful, as all n observation points were enclosed by the confidence intervals. Except STLM and DCM models, most of the remaining points, i.e. the verification set with $N - n$ points, are enclosed, indicating a good quality of the solutions. For the STLM and DCM (Figure 4b and 4f) the points used in the model identification are within confidence intervals (the condition given by Equation 5), but the verification points are outside notwithstanding the much wider confidence intervals. In the case of DCM,

340   the reason is, that for the flume data with rigid vegetation, the Manning formula with constant values of roughness coefficients is unable to correctly reproduce the rating curve and fulfill the constraint given by Formula 5, which is only possible by extending the confidence intervals. The STLM has narrow confidence intervals in the low flow region (identification set), but obtained parameters incorrectly explains water levels at higher flows.

Along with the probabilistic solution, Figure 4 presents a deterministic solution obtained as a computed rating curve with the

345   highest value of likelihood measure (Equation 2). The deterministic solution often deviates from the median of the probabilistic one, as in the case of the Pasche model (4a).
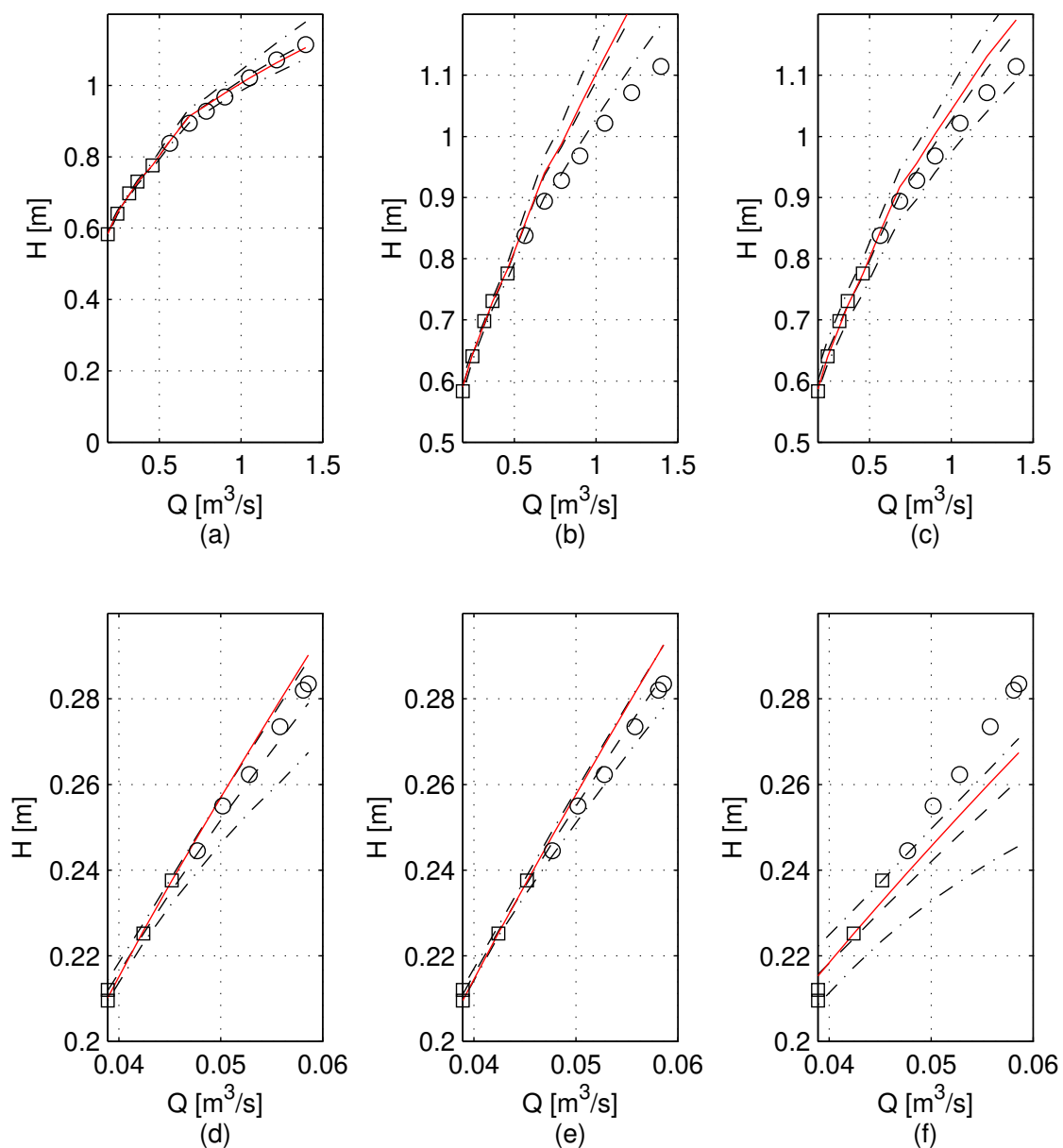
Figure 4. Exemplary rating curves, Ritobacken case study with $n = 5$: (a) GTLM, (b) STLM, (c) PTLM; the flume data set, case 2 with $n = 4$: (d) Pasche, (e) Mertens, (f) DCM. Confidence intervals and the median of the probabilistic solution are given with dashed lines, red line denotes the best simulation in the Monte Carlo ensemble. Observation points used for parameter identification are marked with squares (□), while verification data points are marked with circles (○).

Table 1. Parameter variability ranges (uniform $P(\theta)$ distribution) for flume experiments

Model	Parameter	Min. Value	Max. Value
DCM	$n_1 \ [m^{-1/3}s]$	0.012	0.06
$n_{mc} = 10^4$	$n_2 \ [m^{-1/3}s]$	0.012	0.12
	$d_p[m]$	0.004	0.072
	$a_x \ [m]$	0.05	0.9
Pasche and Mertens	$a_z \ [m]$	0.05	0.9
$n_{mc} = 4 \cdot 10^4$	$k_{ch} \ [m]$	2.5e-05	0.00045
	$k_{fp} \ [m]$	0.005	0.09
	$b_{iii}/B_{fp} \ [-]$	0.333	1
	$C_{dx,F}$	0.001	1.5
	$C_{dx,S} \ [-]$	0.001	1.5
	$\chi_F \ [-]$	-1.21	-0.97
	$\chi_S \ [-]$	-0.32	-0.2
GTLM	$A_l/A_b \ [-]$	0	3.2
$n_{mc} = 4 \cdot 10^4$	$A_s/A_b \ [-]$	0	0.16
	$C^* \ [-]$	0.001	1.03
	$l/L \ [-]$	0	1
	$h_v \ [m]$	0	0.3
STLM	C^*	0.08	3.09
$n_{mc} = 25 \cdot 10^3$	$l/L \ [-]$	0	1
	$h_v \ [m]$	0	2.15
PTLM	C^*	0.001	1.5
$n_{mc} = 4 \cdot 10^4$	$C_D a \ [-]$	0.001	1.03
	$h_v \ [m]$	0	0.3

On the basis of the rating curves computed for each combination of n observation points, it is possible to analyze the estimated average widths of confidence intervals as a function of observation points used in the identification. In the present study, the confidence widths were provided in relative sizes as W, normalized by the median of the probabilistic solution H^M and then averaged over computation points, corresponding to all n observation points:

$$W = \frac{1}{n} \sum_{i=1}^{N} \frac{H_i^{q_L} - H_i^{q_U}}{H_i^M} \tag{10}$$

where H^{q_L} and H^{q_U} stands for the estimates of lower and upper confidence intervals for calculated water level.

Table 2. Parameter variability ranges (uniform $P(\theta)$ distribution) for Ritobacken experiments, numerals in parameter symbols are used to distinguish properties on left (1) and right (1) channel side.

Model	Parameter	Min. Value	Max. Value
DCM $n_{mc} = 10^4$	$n_1\ [m^{-1/3}s]$	0.012	0.15
	$n_2\ [m^{-1/3}s]$	0.012	0.15
	$n_3\ [m^{-1/3}s]$	0.012	0.15
Pasche and Mertens $n_{mc} = 4 \cdot 10^4$	$d_p[m]$	0.004	0.072
	$a_{x1}, a_{x2}\ [m]$	0.05	0.9
	$a_{z1}, a_{z2}\ [m]$	0.05	0.9
	$k_{ch}[m]$	2.5e-05	0.00045
	$k_{fp1}, k_{fp2}\ [m]$	0.005	0.09
	$b_{iii}/B_{fp}\ [-]$	0.333	1
GTLM $n_{mc} = 25 \cdot 10^3$	$C_{dx,F}\ [-]$	0.09	0.2
	$C_{dx,S}\ [-]$	0.82	1.03
	$\chi_F\ [-]$	-1.21	-0.97
	$\chi_S\ [-]$	-0.32	-0.2
	$A_l/A_b\ [-]$	0	3.2
	$A_s/A_b\ [-]$	0	0.16
	$C^*\ [-]$	0.08	1.03
	$l_L/L_L, l_R/L_R\ [-]$	0	1
	$h_{v1}, h_{v2}\ [m]$	0	2.15
STLM $n_{mc} = 25 \cdot 10^3$	C^*	0.08	3.09
	$l_L/L_L, l_R/L_R\ [-]$	0	1
	$h_{v1}, h_{v2}\ [m]$	0	2.15
PTLM $n_{mc} = 10^4$	C^*	0.01	0.4
	$C_D a\ [-]$	0.08	1.03
	$h_v\ [m]$	0	2.15

Chosen results on the influence of the number of identification datapoints on the widths of the confidence intervals and the ratio of verification points included within the intervals are provided in Figures 5-7. In Figure 5 for GTLM applied for

355 Ritobacken case study for Autumn 2011 and also the Pasche model used for the flume data set in case 1 it can be noticed that: (1) the relative confidence interval widths (5a, 6a) are high for a small n as a result of the parameter equifinality; (2) with additional data points, the solution converges by reducing the span of intervals but also its variability due to different combination of observation points; (3) the width of confidence intervals for the full data set $n = N$ in both cases is below
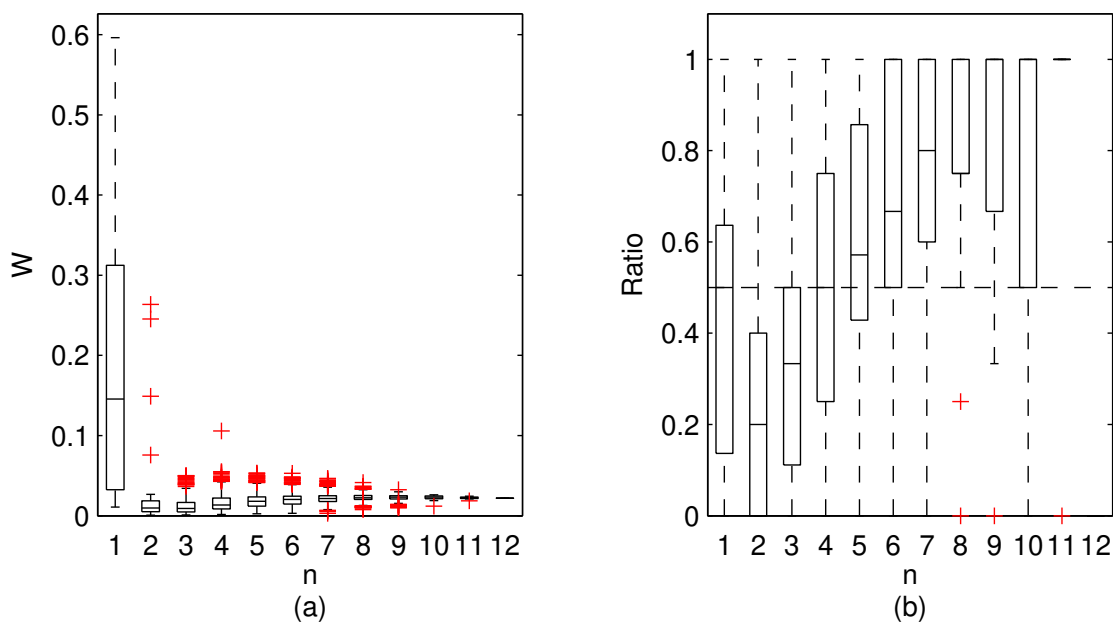
Figure 5. GTLM results for Ritobacken case study, Autumn 2011: (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)

5%; (4) the confidence intervals estimated for a low number of observations ($n < 4$) have poor predictive performance, as most of the observations in the verification sets fall outside (Figure 5b, 6b); (5) in both cases for $n > 4$ more than 50% of the verification set is enclosed with the estimated confidence intervals. Figure 7 shows an example of a model with a poor performance, indicating the model's inadequacy to the given case. The confidence intervals are extending with n (Figure 7a), which for $n > 4$ allows to enclose most of the verification set (Figure 7b).

3.2 Model identifiability

The model identifiability is understood here as the ability to determine the parameter *a posteriori* distribution that explains the model uncertainty in relation to observations. This is satisfied by meeting the constraint given in Equation 5, as for cases presented in Figure 4.The criterion of Eq. 5 might be fulfilled even for a poor model by extending the parameter variability ranges (1 and 2), specified with *a priori* distribution $P(\theta)$. The only limitation could be the physical meaning of the parameters.

Figure 8 shows exemplary results for a model that could not be identified for a given dataset. Values of J (Equation 4) were computed for observation points used in the parameter identification and averaged in respect of the ensemble count n. This model was unable to correctly reproduce the rating curve over the whole Monte Carlo ensemble of parameters. The computed

(a)   (b)
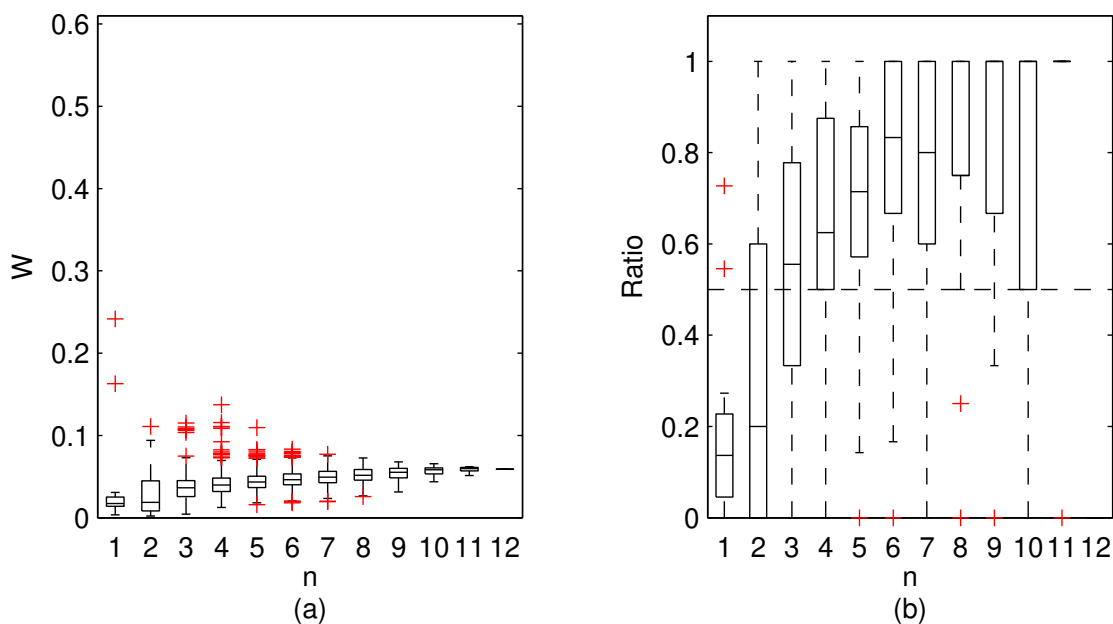
Figure 6. Pasche results for the flume data set, case 2: (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)

water levels did not follow the observed shape of the rating curve and as a result it was not possible to find such a solution of Equation 1 where identification data points would be enclosed by the confidence intervals (Equation 5). Only for $n = 1$ it was possible to find such κ (Equation 2) where almost all single point observations were enclosed by confidence intervals (Figure

375   8). For one extreme value, it was not possible to find a feasible solution at all. The reason is that the Monte Carlo sample of computed water levels did not cover that point. With an increasing number of n, the number of observation points enclosed by the confidence intervals depends on the combination of observation points. Some beneficial effects allow to fulfill the constraint 5, such as an extreme value of 1 for $n = 6$ whereas others enclose only a small share of observations. For $n = N = 11$, there is a single solution, in which about 60% observations were enclosed by confidence intervals. For an identifiable model, Figure 8

380   would consist of single horizontal lines between 0.95 and 1, indicating fulfillment of the constraint of Eq. 5 for all simulations.

The Pasche and Mertens models applied to the Ritobacken case study were not identifiable even with relatively large variability ranges of the parameters (Figure 8. This is explained by the fact that these methods were developed for rigid unsubmerged vegetation whereas the Ritobacken had immerse flexible vegetation. The two-layer approaches appeared to be more flexible and thus more universal. By applying large parameters variability for the GTLM and PTLM models, it was possible to meet

385   Equation 5 for the flume case study although these methods were not originally designed for such unsubmerged and rigid vegetation. The STLM model failed for the Ritobacken Spring 2011 case with sparse, low vegetation with h=9 cm, and for
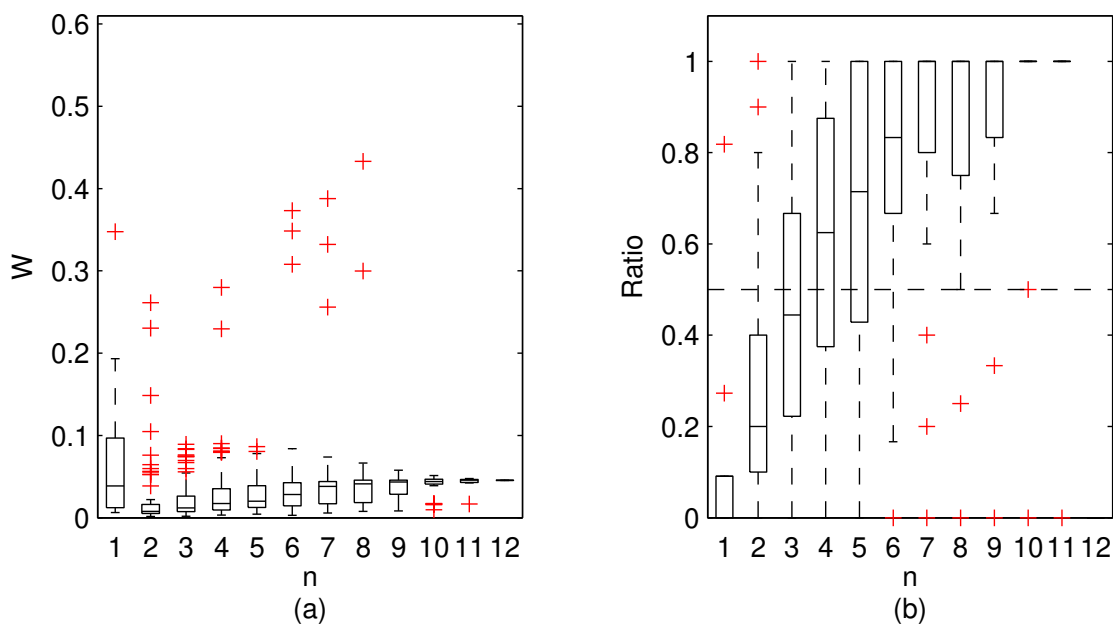
Figure 7. Manning based DCM results for the flume data set, case 2: (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)

both flume cases. For flume experiments, the STLM likely did not work because the assumption that >80% of flow should be conveyed in the non-vegetated zones was not fulfill. The rest of the models, including DCM for all cases, were identifiable.

3.3 Widths of confidence intervals and quality of uncertainty estimation

390 To compare the performance of the applied identifiable discharge prediction methods, we show bar plots of average values for the ratio of verification set points enclosed by confidence intervals and their relative widths as a function of observation points used in the model identification n (Figures 9-13). The averaged values correspond to the mean values of the box-plots in Figures 5-7. Separate bar plots were prepared for the different case studies: Figure 9 – Flume case 1; Figure 10 – Flume case 2; Figure 11 – Ritobacken Spring 2011; Figure 12 – Ritobacken Autumn 2011; Figure 13 – Ritobacken Spring 2012.

395 The values presented in Figures (9-13) are averaged over all uncertainty estimates at given number of observations n. Therefore, for $n = N - 1$, where there was always only one verification point, the ratio for verification points can be any value between $0 - 1$, not only 0 or 1. Averaged ratio of verification points enclosed within confidence intervals, together with their relative width W, should be considered as a two criteria measure on how well the obtained model reproduces the discharge curve. Narrow confidence intervals indicate that the model uncertainty, estimated using n observations, is small. The ratio of
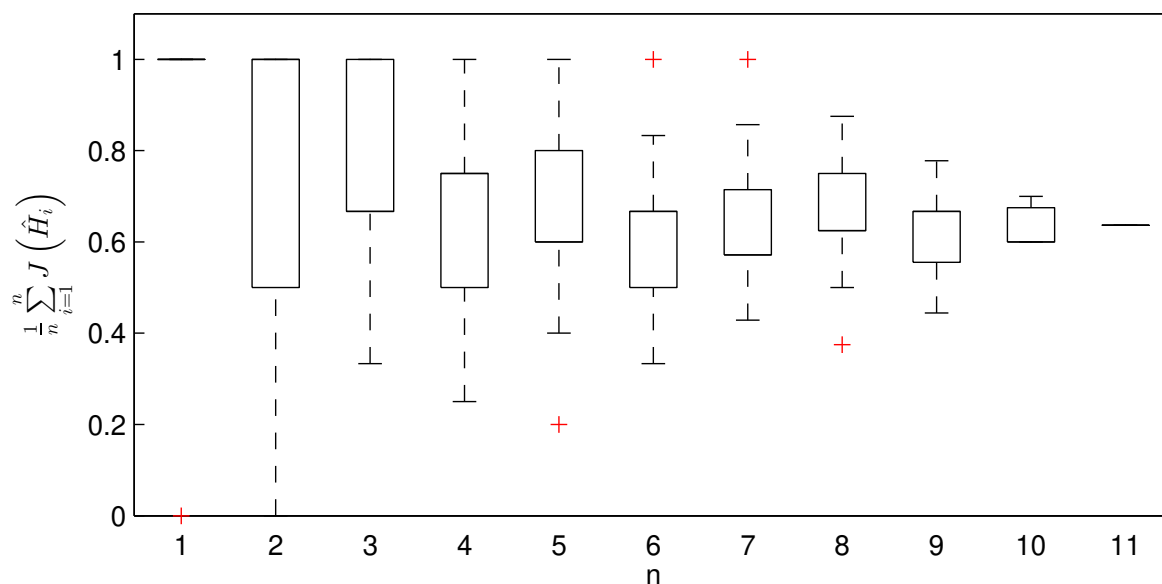
Figure 8. Portion of observation points within 95% confidence intervals for Pasche method in function of observation points used in parameter identification, presented in a form of box-plots; results for the unsuitable data set for the Pasche method of Ritobacken, Spring 2012.

400   observations from the verification set enclosed within these intervals informs how the estimated uncertainty is representative
for other data sets than these used for identification. A low ratio suggests that the probabilistic term incorrectly predicts the
model uncertainty for the verification set. Therefore, narrow confidence intervals for small n numbers resulting in small ratios
should be considered as unsuccessful, as the uncertainty analysis appears to be too optimistic. On the other hand, for larger n,
good ratios might be obtained with very wide confidence intervals, indicating a poor model. The best solution is that one, which
405   has the narrowest confidence intervals with satisfactory ratio of verification set enclosed within it. We interpret the results by
analyzing those both criteria together.

3.3.1   Flume data set, case 1

For the flume data in the case 1 (Figure 9), with rigid-high vegetation in floodplains and also channel banks, the best results
were obtained with the Mertens method. It is characterized with the narrowest confidence intervals W, having a good predictive
410   performance. Confidence intervals for $n > 1$ were below 5% and for $n > 3$ they already enclosed more than 50% of the
verification points. Almost similar performance was found for the DCM method, with slightly wider confidence intervals.

Surprisingly, both methods outperformed the Pasche model that is a very similar approach to the Mertens method, but with
a much more detailed description of the vegetation induced resistance. Estimated confidence intervals width was about three
times larger than for Mertens method and DCM, but enclosing a similar number of verification points. The reason could be
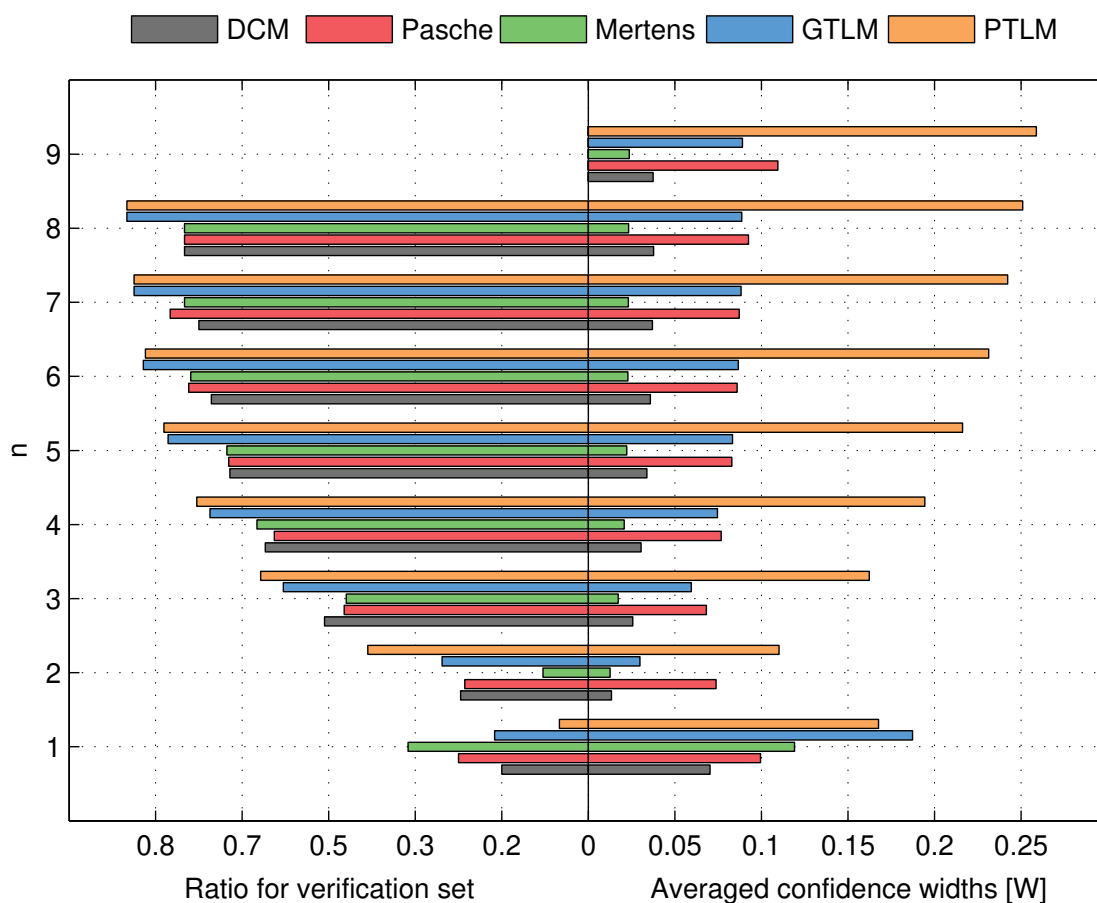
Figure 9. Ratio of verification set ($N - n$) enclosed by confidence intervals and average width of confidence intervals for different number of data points for model identification (n); flume data set, case 1.
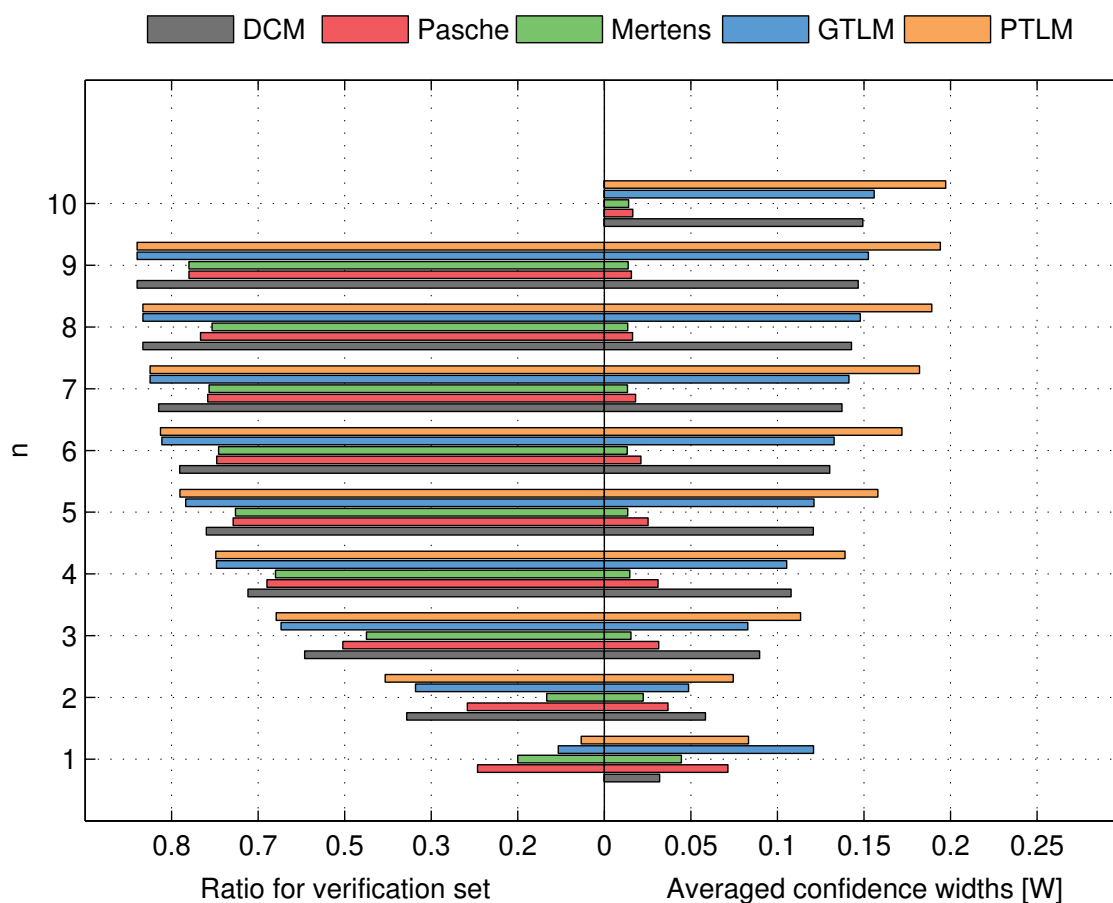
Figure 10. Ratio of verification set ($N - n$) enclosed by confidence intervals and average width of confidence intervals for different number of data points for model identification (n); flume data set, case 2.
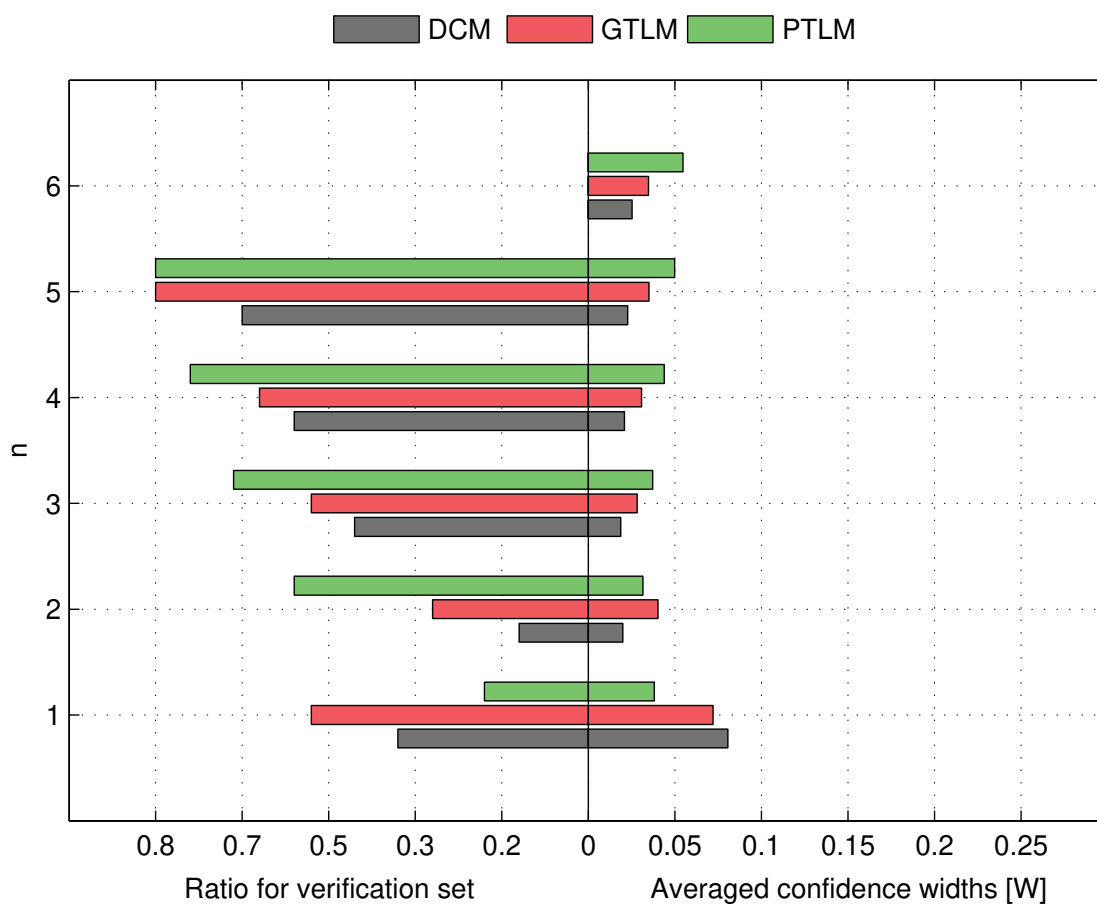
Figure 11. Ratio of verification set $(N - n)$ enclosed by confidence intervals and average width of confidence intervals for different number of data points for model identification (n); results shown for the identifiable models for Ritobacken, Spring 2011.
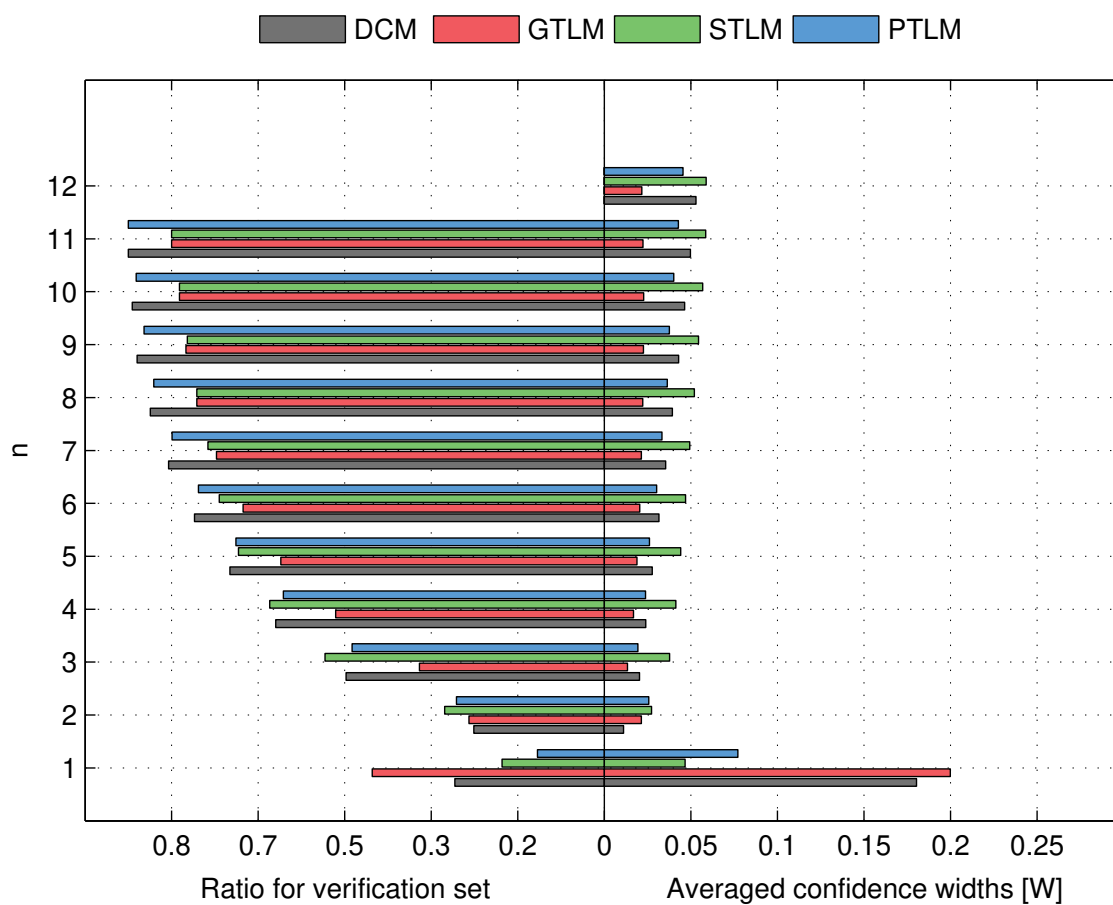
Figure 12. Ratio of verification set $(N - n)$ enclosed by confidence intervals and average width of confidence intervals for different number of data points for model identification (n); results shown for the identifiable models for Ritobacken, Autumn 2011.
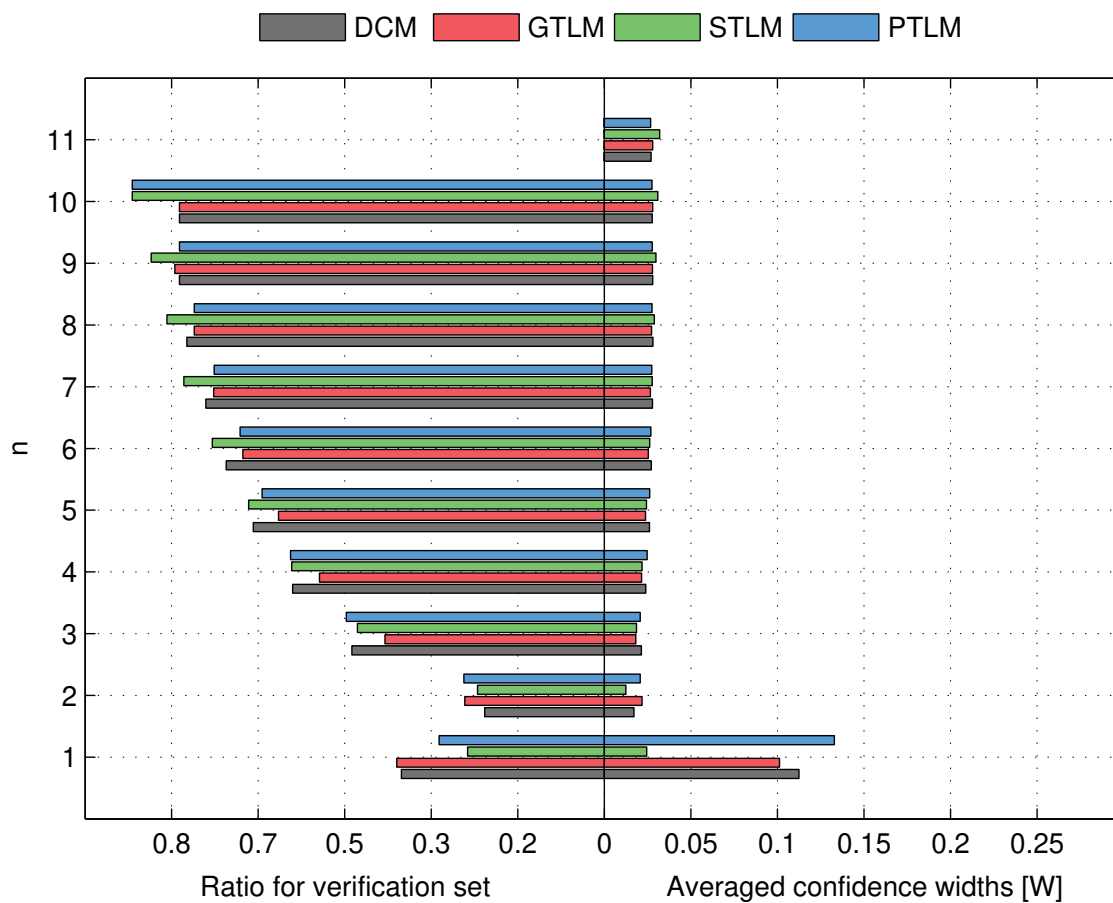
Figure 13. Ratio of verification set ($N - n$) enclosed by confidence intervals and average width of confidence intervals for different number of data points for model identification (n); results shown for the identifiable models for Ritobacken, Spring 2012.

415   the susceptibility of the Pasche method to numerical instabilities. Because of vegetation present on the channel banks, the
floodplain region was extended above geometrical channel banks. This introduces discontinuity to the hydraulic radius in
floodplains, as water levels slightly exceed geometrical banks. Probably, this might lead to numerical instability of implicit
formulas used in Pasche method, but not present in the Mertens method. GTLM confidence intervals were almost the same
as for the Pasche method and much more narrow than for PTLM. The GTLM and PTLM enclose the largest ratio of the
420   verification points because of wide confidence intervals.

3.3.2   Flume data set, case 2

For the flume case 2 (Figure 10), both the Pasche and Mertens methods appear to be the most appropriate. Estimated widths of
confidence intervals do not exceed 4-5% for $n > 1$ and fell bellow 2-3% for a sufficient number of observations ($n > 5$). The
predictive skills of the identified models are high, with around 70% of the verification set enclosed by the confidence intervals
425   at $n > 4$. The two-layer GTLM and PTLM have similar skills than in the flume case 1. GTLM has a similar uncertainty
performance as DCM while PTLM has slightly worse performance. For all these three models, the final confidence widths
for $n = N$ are about 20%. Because of their larger extent, the estimated intervals enclose slightly larger number of verification
points than with the Pasche and Mertens methods. The DCM has 8 times wider confidence intervals than for flume case 1. The
main difference between the flume cases 1 and 2 was the rough floodplain surface with the grain sizes of 0.5-1 cm for the case
430   2 compared to the smooth floodplain of case 1, indicating that the DCM was not able to perform reliably for the combination
of rough surface and emergent vegetation.

Figure (10) highlights the specific dependency of DCM, GTLM and PTLM on n. For a small number of data points for a
model identification at $n = 1$, confidence widths are high, because of the parameter equifinality effect. With additional points,
the equifinality effect is reduced, and for $n = 2$ the confidence interval widths are at their smallest but with poor predictive
435   skills. With increasing n the uncertainty estimates are corrected by additional data points. The same pattern is present but less
noticeably for the Pasche and Mertens methods and for the other cases.

3.3.3   Ritobacken, Spring 2011 case

The Spring 2011 case study refers to flow conditions with poorly developed vegetation 1 year after the floodplain excavation.
These conditions with low vegetation having a mean relative submergence (floodplain water depth divided by vegetation height)
440   of 3.3 is reflected in the computational output (Figures 11), with process-based methods for vegetation resistance characterized
with a relatively poor fit.

The PTLM had significantly higher ratios than DCM at $n > 2$, with GTLM falling between these. Among the two latter
methods, GTLM had only slightly wider confidence intervals than DCM: 3% vs 2%. For PTLM, with similar share of
verification points enclosed within confidence intervals, widths W were about 6%. At $n = 1$, GTLM had better performance
445   than PTLM and DCM, with notably higher ratio of points enclosed.

3.3.4 Ritobacken, Autumn 2011 and Spring 2012 cases

Ritobacken Autumn 2011 and Spring 2012 case studies reflect flow conditions in a channel of two phases of vegetation development. In Autumn 2011 vegetation was higher and denser than at the beginning of growing season in Spring 2012. This can be seen in the performance of the applied discharge methods. For the fully vegetated conditions of Autumn 2011 (Figure 12), all the 4 identified methods enclosed over 70% of the observations at $n > 5$ with $N = 12$. GTLM appeared the most appropriate method as it had half narrower confidence intervals (1-2 %) than the other methods, with only 10 % lower ratio of enclosed points than DCM and PTLM. PTLM with 4% interval width has, however, better predictive skills with 10 % more verification points enclosed. For Autumn 2011, good results were obtained also for STLM and DCM Manning, for which widths for $n = N$ were about 5%, so just slightly higher than for PTLM.

For the Spring 2012 (Figure 13), all methods have almost equal confidence widths and ratios of enclosed verification points. The overall measures are similar to those from Autumn 2011. At $n > 5$, STLM had a slightly higher ratio of verification data enclosed compared to the other methods. The confidence widths are about 3% and for $n > 5$ for all methods more than 70% of points fall within confidence intervals.

3.4 Physical interpretation of identified parameters

The obtained parameter values were compared with the measured ones for the two most complex models of Pasche and GTLM, for their most representative cases: the flume experiment, case 2 and Ritobacken Autumn 2011, respectively. In both cases, solutions for all observation points $n = N$ were used.

In the case of the Pasche method applied for controlled flume conditions, all parameter values were known. For GTLM there was no data for parameters of vegetation extent (l_L/L_L, l_R/L_L, h_L and h_R) but estimates of the blockage factor B_X itself, which values were used for comparison. To compare model identification outputs with observed values, we recalculated values of extent parameters l_L/L_L, l_R/L_L, h_L and h_R to B_X.

Figure 14 shows the cumulative distribution function for marginal *a posteriori* distribution $P(\theta/H)$ of parameters θ of the Pasche method. Measured values of parameters are provided with blue lines. Also a best solution from the Monte Carlo ensemble was given with red dashed lines. It can be noticed, that the strongest discrepancies between measured and identified values are present for the stem diameter d_p and longitudinal stem spacing a_x. A median of the probabilistic solution and also the best model fit for d_p is close to 0.04 m, while the real diameter was 0.008 m. In the case of a_x it is 0.6 m to 0.1 m. This has a clear physical sense, as in terms of the model identification, small stem diameters d_p at dense spacing with small a_x were equivalent to larger d_p and smaller a_x. This finding is supported, by much smaller discrepancies in other parameters. It should be noted, that the measured parameter values provide a fit close to the best one (Kiczko et al., 2017).

In the case of the two layer approaches, there was no direct data on vegetation extent parameters: l_L/L_L, l_R/L_L, h_L and h_R parameters values, for which only outcomes of the identification are available. As an example, in Figure 15 results for Ritobacken Autumn 2011 case are provided. It can be seen that, the identified values of the parameterization for flexible vegetation (Equation 8) had a fairly narrow distribution for the reconfiguration (X) and drag coefficient (C_{Dx}) of the foliage
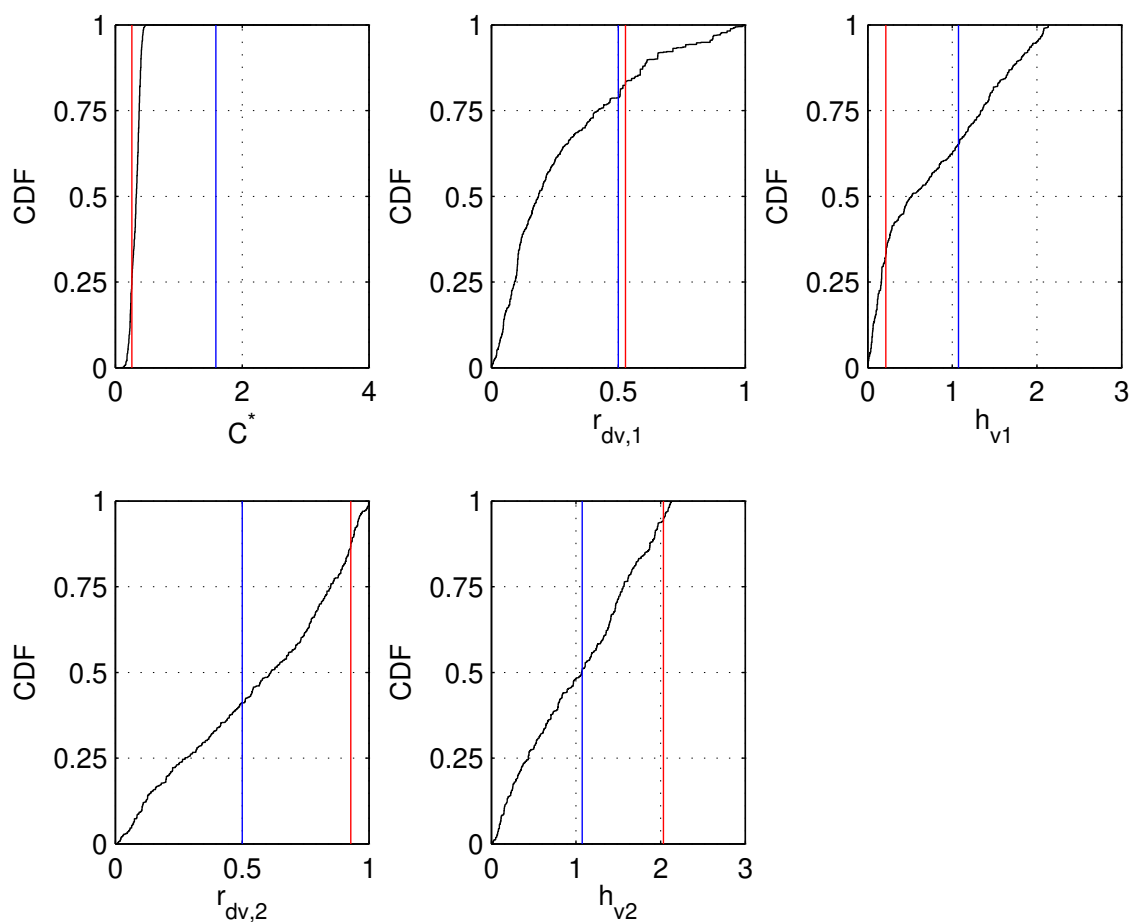
Figure 14. Marginal *a posteriori* distributions of Pasche model parameters, identified using $n = N$ observation points for flume experiment, case 2; measured parameters values were provided with blue lines, the best value in the Monte Carlo ensemble with red lines.
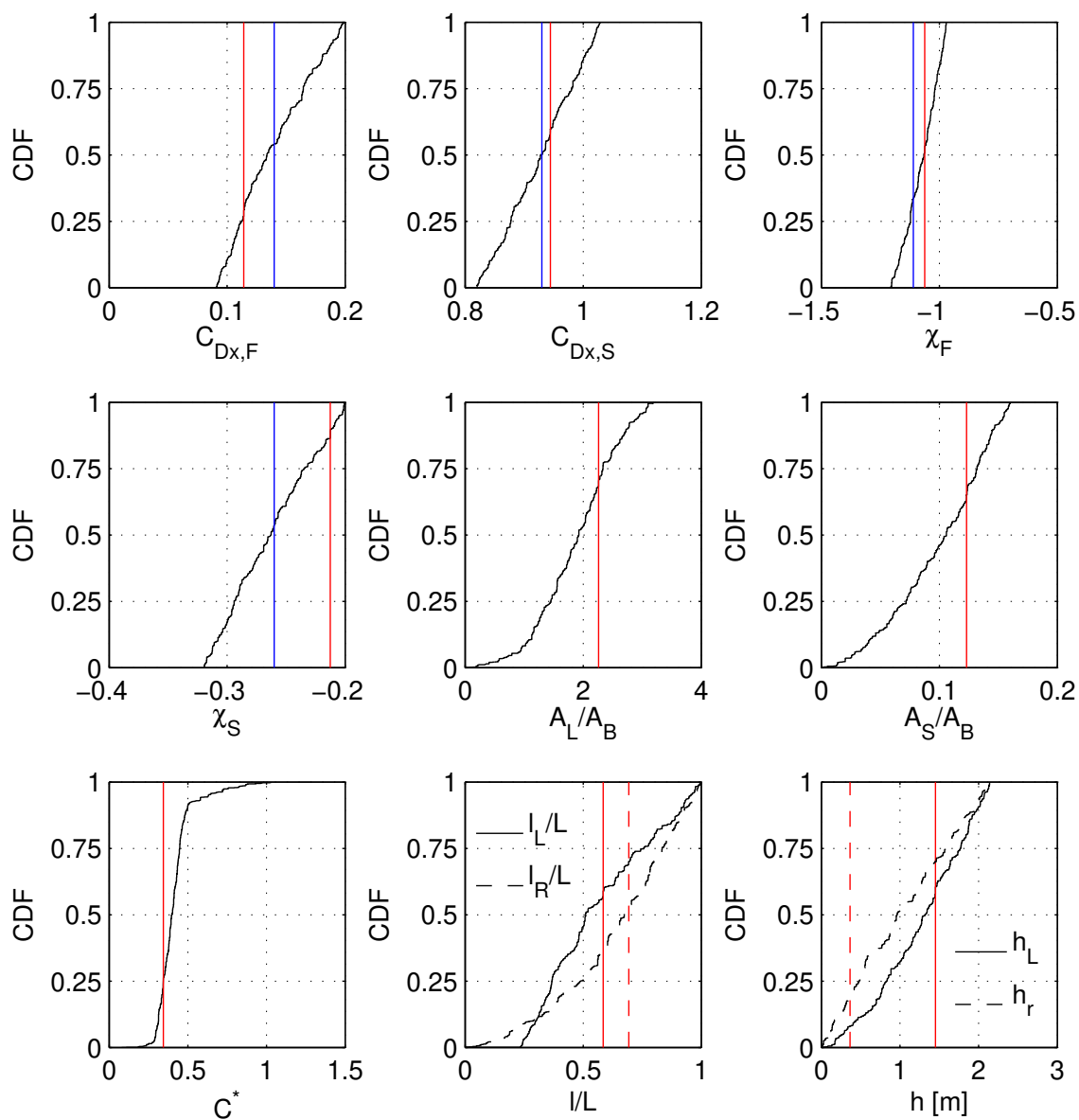
Figure 15. Marginal *a posteriori* distributions of GTLM model parameters, identified using $n = N$ observation points for Ritobacken Autumn 2011; dashed lines – confidence intervals and median of a probabilistic solution, red line – best simulation in the Monte Carlo ensemble; parameters values given by Västilä and Järvelä (2014) were provided with a blue lines
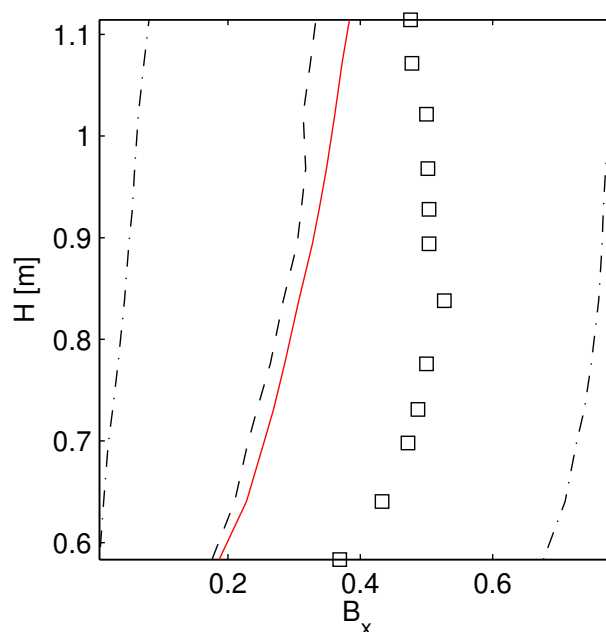
Figure 16. Blockage factor B_X measured in the field and determined as an inverse solution of GTLM for Ritobacken Autumn 2011 case study; squares denote measured values, dashed lines – confidence intervals and median of a probabilistic solution, red line – best simulation in the Monte Carlo ensemble.

and stem, which fell close to the values observed for willows and other woody species (e.g. Västilä and Järvelä, 2018). Wide

480 ranges for the vegetation heights h results from interactions with l/L and from the model insensitivity, when vegetation exceed the water level and there is no free flow zone above. The values in the distribution of the identified C^* were notably larger than the experimentally derived C^* value ($\sim 0.034 - 0.08$, Västilä et al. (2016)), which is compensated by the notably lower identified A_L/A_B and $C_D a$ compared to the measured vegetation densities at Ritobacken ($a \approx 10-25$ for the grassy vegetation, Västilä et al. (2016)). This is another example of the parameter equifinality that can result if all the vegetation properties have

485 to be identified because of lack of available measurement data.

The comparison of real and measured values on vegetation extent for the Ritobacken case study was possible for the blockage factor B_X. In the presented approach it probabilistic estimates can be calculated using values of h_L, h_R, l_L/L_L and l_R/L_R. In Figure 16, the measured and the identified blockage factor B_X of GTLM is given as a function of the water depth. It can be noted, that confidence intervals for the B_X are wide. The observed values are shifted from the median of a probabilistic

490 solution towards 0.6 quantile. A large spread of values for B_X, with very small variation of water levels for that solution (Figure 12), suggest a moderate model sensitivity on B_X.

4   Discussion

The present study is according to our knowledge the first one, where different discharge capacity methods were compared in
the respect of their uncertainty, estimated along with model parameters, using probabilistic formulation of the problem of the
parameter identification. It should be noted that noticeable focus was made to ensure that the uncertainty analysis was objective
and repeatable, which can be seen in the proposed technique for scaling the likelihood function. The novelty of the proposed
approach includes the analysis of obtained confidence widths, together with the ratio of independent observations explained by
them, with respect of the number of observations used in the model identification.

 Our results show that the number of parameters is not a factor precluding the use of a given method for predicting the channel
discharge. It was possible to identify a model with more than ten parameters (i.e. GTLM), almost as well as three parameter
ones (DCM). Parameter equifinality influenced the uncertainty estimates only when the number of observation points was
very small, independent of the number of parameters. Widths of confidence intervals stabilized close to the final extent at
about three-four observation points ($n > 3$, Figures 9-13). Equifinality of parameters was however present, which is shown by
the discrepancy between the identified and measured values of parameters, as well as their large variation (3.4). This agrees
with the finding of Her and Chaubey (2015), who reported similar effects of the parameter number on the equifinality and
uncertainty estimates for a horological runoff-model.

 Our results clearly demonstrate the influence of the number of observation points on uncertainty. For a small number of
observation points, the uncertainty estimates were for all methods relatively high. Such effect however was not reported in
Her and Chaubey (2015), where different lengths of time series were considered, all with large number of observation points.
Both studies demonstrate however a characteristic stabilization of uncertainty estimates for larger sets of observation points
(Figure 7). Her and Chaubey (2015) investigated also effect of additional data points, referring to other model derivatives than
the main output, such as information on the flow from sub-basins. For the present study, the analysis could be amended using
e.g. observations of velocities in the channel and floodplains. However, such data were not available for all cases and were thus
not included here.

 The results confirm previous findings of (Kiczko and Mirosław-Świątek, 2018; Kiczko et al., 2018; Romanowicz and Kiczko,
2016), that for discharge formulas the probabilistic solution differs form the deterministic one. This is evident from Figure 4
for calculated rating curves or parameter distributions in Figure 14. This obvious behavior of nonlinear models highlights the
needs for such uncertainty analyses.

 The more complex, process-based methods were usually better than the classical DCM, having more narrow confidence
intervals, enclosing larger ratios of observations, when applied to vegetative conditions they were developed for. This important
methodological finding suggests that it could be possible to choose an appropriate method on the basis of its fit measures and
uncertainty estimates. Thus, results show the advantages of process-based methods in the case of parametrization over simpler
ones. Dalledonne et al. (2019) come to similar conclusions, as they obtained the best uncertainty estimates for the most complex
model.

525   We found that the differences between the one-dimensional methods were notably larger than for the study of Dalledonne et al. (2019) focusing on a two dimensional model. Further, the Warmink et al. (2013) study did not consider the choice of the flow resistance parametrization method as crucial. The presently investigated flume and field cases had a notable portion of the cross-section covered by the floodplain vegetation, with Manning's n ranging at 0.017-0.150 $m^{-1/3}s$. Thus, our results indicate that the choice of the resistance formula is important for cases where vegetative resistance dominates. On the other
530   hand, one-dimensional models may be more sensitive to uncertainty related to the identification of the resistance parameters than are two-dimensional models.

The performance of a model depends on its adequacy for the given vegetative and flow conditions. For unsubmerged sparse rigid vegetation, the most reliable method was the Mertens model with mostly explicit formulas. Because of a simpler numerical form than in the Pasche method, the Mertens method was less vulnerable to numerical instabilities, which probably affected the
535   outcomes of the Pasche uncertainty estimation. In the case of dense flexible vegetation typically observed on natural floodplains (Figure 3), the most reliable performance with respect to uncertainty estimates was obtained with the two-layer approaches GTLM and PTLM that were performed well for both dense submerged and emergent vegetation (Figures 11-13).

The GTLM was in this paper amended with a vegetation parameterization (Eq. 8) that describes the influence of the plant streamlining and reconfiguration on flow resistance. The GTLM with (Eq. 8) performed particularly well when the vegetation
540   was high (Figure 12), appearing to be the most reliable method for predicting the discharge capacity during the most critical conditions when the vegetative flow resistance is high. The GTLM parameterized at low flows reliably predicted the water levels during high discharges, including the more rapid increase in discharge at water levels exceeding vegetation height (Figure4a). Although Eq. 8 has been developed for woody vegetation, it was applicable to the predominantly grassed vegetation at the field site. Field surveys indicated that much of the plants consisted of a main stem and more flexible leaves, conceptually similar
545   in structure to foliated woody vegetation. Eq. 8 describes the drag from stem and leaves and allows to set different values for the flexibility-induced reconfiguration for the stem and foliage.

It should be noted that the results for the DCM with constant values of the Manning coefficient were quite good except for flume case 2. In all cases it had worse performance than the process-based methods, but was applicable in all these cases. Based on the results, the process-based methods are expected to perform better than DCM when several important sources of flow
550   resistance, such as rough floodplain surface and vegetative drag, are present.

Despite the larger number of parameters, the process-based methods were less flexible than the Manning based DCM approach. Pasche and Mertens methods were only suitable for rigid unsubmerged vegetation, for which they were derived. The two-layer approaches GTML and PTML, although it was possible to identify them, had a very poor performance when applied to sparse emergent vegetation (Figure 2). Further, our findings confirmed that the STLM is strict about the assumption of neg-
555   ligible flow within vegetation (Section 3.2), and it had a more favorable performance during Spring 2012 (Figure 13) when the deflected height of the bent grasses was low, with an expected lower share of flow within vegetation compared to Autumn 2011 (Figure 12).

The most problematic issue in the proposed approach of identifying vegetation properties of process-based methods through the formulation of the inverse problem (Figure 1 b) is the physical interpretation of obtained parameters. The identified values

560   are different from the physical ones. The most obvious reason is equifinality, as in the case of the stem diameter and spacing
in Pasche and Mertens methods (Figure 14). In the case of two-layer approaches, the fit measures reveal a low sensitivity
of GTLM to the blockage factor: B_X has large variability, while variation of computed water levels was very small. The
application of process-based methods with numerous parameters seems to be inseparably connected with the problem of the
equifinality. Similar behavior was reported for the Shiono-Knight model by Knight et al. (2007). The parameter equifinality,

565   as over-parametrization is a basic assumption of the probabilistic approach in the parameter identification (Beven and Binley,
1992, 2014). Overall, the parameter identification is expected to result in more physically realistic values if at least some of the
required vegetation properties or the channel bed roughness can be directly measured and used as the input (see Figure 1a).

Discharge formulas analyzed in the study are usually only a part of the one-dimensional model. The uncertainty of such
models depends also on additional elements, like spatial variability of resistance and simplification of the channel geometry. It

570   should be also noted, that the investigated cases had fairly regular cross-section and homogeneous vegetation. Therefore, care
should be taken when attempting to generalize the presented findings to all one-dimensional approaches. In complex real-world
cases, it might be beneficial to include several discharge formulas through an ensemble approach, which is also used in other
fields, such as climate modeling.

5   Conclusions

575   In this study, six methods for estimating the channel discharge capacity were analyzed in terms of their uncertainty, for two
experiments: a flume experiment with rigid submerged vegetation and a field experiment with flexible vegetation under both
emergent and submerged conditions. The outcomes of the study are summarized as follows:

1. The numerical experiments showed that it is possible to identify parameters of process-based methods including a large
   number of parameters on the basis of the inverse problem with narrow uncertainty bands.

580   2. The number of parameters is not a factor determining the applicability of the method. It was possible to obtain similar
   uncertainty estimates for models with both a low and a high number of parameters.

3. The uncertainty related to the parameter equifinality is noticeable only when a small number of observations is used in
   parameter identification.

4. The parameters obtained through the identification differ from their measured physical values, which results from the
585   parameter equifinality. The equifinality does not, however, affect the uncertainty of a model.

5. Uncertainty estimates clearly indicate the applicability of a given model to the analyzed case. Unsuitable models, e.g.
   those developed for non-submerged vegetation but applied to submerged vegetation, have relatively wide uncertainty
   estimates or lack a probabilistic solution.

6. The best results in terms of the lowest uncertainty estimates were obtained with the Mertens method for the unsubmerged,
590   rigid vegetation case. For the dense flexible vegetation, the generalized two-layer method (GTLM) accompanied with

a parameterization describing the streamlining and reconfiguration of plants (Eq. 8) had the most reliable performance across different conditions, functioning under submerged and emeregent conditions. In most cases, the Manning based DCM had also satisfactory performance.

7. An open issue is the generalizability of the obtained results to spatially distributed one-dimensional models.

595 8. The proposed approach with the novelty of comparing different models in terms of their uncertainty along with the quality of the uncertainty estimation might be useful in other similar studies.

Author contributions. Adam Kiczko: manuscript text, implementation of discharge formulas in Matlab, numerical experiments; Kaisa Västilä: manuscript text, methodology for two layer models, Ritobacken field experiment, analysis of numerical outputs; Adam Kozioł, flume experiments and together with Janusz Kubrak: Pasche and Martens methodology including computation algorithm; Elżbieta Kubrak and Marcin
600 Krukowski: flume experiments, measurement data analysis, improving article text.

Competing interests. No competing interests are present.

Acknowledgements. The research was partly supported by National Science Centre (Poland), Program Miniatura 1, project no. 2017/01/X/ST10/00987, by Maa- ja vesitekniikan tuki ry (No 33271) and by Maj and Tor Nessling Foundation (No 201800045).

References

605 Beven, K. and Binley, A.: The future of distributed models: model calibration and uncertainty prediction, Hydrol. Process, 6, 279–298, 1992.

Beven, K. and Binley, A.: GLUE: 20 years on, Hydrological Processes, 28, 5897–5918, https://doi.org/10.1002/hyp.10082, 2014.

Blasone, R.-S., Vrugt, J. A., Madsen, H., Rosbjerg, D., Robinson, B. A., and Zyvoloski, G. A.: Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling, Adv. Water Resour., 31, 630–648, 2008.

Dalledonne, G. L., Kopmann, R., and Brudy-Zippelius, T.: Uncertainty quantification of floodplain friction in hydrodynamic models, Hy-
610 drology and Earth System Sciences, 23, 3373–3385, https://doi.org/10.5194/hess-23-3373-2019, https://www.hydrol-earth-syst-sci.net/23/3373/2019/, 2019.

Fread, D.: Flood routing models and the Manning n, in: International Conference for Centennial of Manning's Formula and Kuichling's Rational Formula, pp. 699–708, http://tsunamiready.noaa.gov/oh/hrl/hsmb/docs/hydraulics/papers{_}before{_}2009/hl{_}271.pdf, 1989.

Helmiö, T.: Unsteady 1D flow model of compound channel with vegetated floodplains, Journal of Hydrology, 269, 89–99,
615 https://doi.org/10.1016/S0022-1694(02)00197-X, 2002.

Helmiö, T.: Unsteady 1D flow model of a river with partly vegetated floodplains—application to the Rhine River, Environmental Modelling & Software, 20, 361–375, https://doi.org/https://doi.org/10.1016/j.envsoft.2004.02.001, http://www.sciencedirect.com/science/article/pii/S1364815204000714, 2005.

Her, Y. and Chaubey, I.: Impact of the numbers of observations and calibration parameters on equifinality, model performance, and output
620 and parameter uncertainty, Hydrological Processes, 29, 4220–4237, https://doi.org/10.1002/hyp.10487, 2015.

Her, Y. and Seong, C.: Responses of hydrological model equifinality, uncertainty, and performance to multi-objective parameter calibration, Journal of Hydroinformatics, 20, 864–885, https://doi.org/10.2166/hydro.2018.108, https://doi.org/10.2166/hydro.2018.108, 2018.

Jalonen, J. and Järvelä, J.: Estimation of drag forces caused by natural woody vegetation of different scales, J. Hydrodynam. B, 26, 319, 2015.

625 Jalonen, J., Järvelä, J., Virtanen, J.-P., Vaaja, M., Kurkela, M., and Hyyppä, H.: Determining characteristic vegetation areas by terrestrial laser scanning for floodplain flow modeling, Water, 7, 420–437, 2015.

Järvelä, J.: Determination of flow resistance caused by non-submerged woody vegetation, International Journal of River Basin Management, 2, 61–70, https://doi.org/10.1080/15715124.2004.9635222, 2004.

Kalinowska, M. B., Västilä, K., Kozioł, A. P., Rowiński, P. M., Kiczko, A., and Kubrak, J.: Modelling of velocity distribution in a channel
630 partly covered by submerged vegetation, in: Recent Trends in Environmental Hydraulics, edited by Kalinowska, M. B., Mrokowska, M. M., and Rowiński, P. M., vol. in press., Springer, Warsaw, geoplanet: edn., 2020.

Khatibi, R. H., Williams, J. J. R., and Wormleaton, P. R.: Identification Problem of Open-Channel Friction Parameters, Journal of Hydraulic Engineering, ASCE, 123, 1078–1088, https://doi.org/10.1061/(ASCE)0733-9429(1997)123:12(1078), 1997.

Kiczko, A. and Mirosław-Świątek, D.: Impact of Uncertainty of Floodplain Digital Terrain Model on 1D Hydrodynamic Flow Calculation,
635 Water, 10, 1308, 2018.

Kiczko, A., Kozioł, A., Kubrak, J., Krukowski, M., Kubrak, E., and Brandyk, A.: Identification of vegetation parameters for compound channel discharge as inverse problem, Annals of Warsaw University of Life Sciences–SGGW. Land Reclamation, 49, 255–267, https://doi.org/10.1515/sggw-2017-0020, 2017.

Kiczko, A., Szeląg, B., Kozioł, A., Krukowski, M., Kubrak, E., Kubrak, J., and Romanowicz, R. J.: Optimal Capacity of a Stormwater Reservoir for Flood Peak Reduction, Journal of Hydrologic Engineering, 23, 4018 008, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001636, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001636, 2018.

Knight, D. W., Omran, M., and Tang, X.: Modeling depth-averaged velocity and boundary shear in trapezoidal channels with secondary flows, Journal of Hydraulic Engineering, 133, 39–47, 2007.

Koziol, A.: Czasowa i przestrzenna makroskala turbulencji strumienia w dwudzielnym trapezowym korycie z drzewami na terenach zalewowych, Acta Scientiarum Polonorum. Formatio Circumiectus, 9, 2010.

Kozioł, A. P., Kubrak, J., and Ciepielowski, A.: A hydraulic model of discharge capacity for rivers with forest vegetation on flood lowland, in: Model application for wetlands hydrology and hydraulics, edited by Kubrak, J., Okruszko, T., and Ignar, S., pp. 103–111, Center of Excellence in Wetland Hydrology Wethydro. Warsaw : Warsaw Agricultural University Press, Warsaw, 2004.

Kubrak, E., Kubrak, J., Kozioł, A., Kiczko, A., and Krukowski, M.: Apparent Friction Coefficient Used for Flow Calculation in Straight Compound Channels, Water, 11, 745, 2019a.

Kubrak, E., Kubrak, J., Kuśmierczuk, K., Kozioł, A. P., Kiczko, A., Rowiński, P. M., Kozioł Adam Paweł, Kiczko, A., Kozioł, A. P., Kiczko, A., and Rowiński, P. M.: Influence of stream interactions on the carrying capacity of two-stage channels, Journal of Hydraulic Engineering, 145, 6019 003, https://doi.org/10.1061/(ASCE)HY.1943-7900.0001585, 2019b.

Kuczera, G. and Mroczkowski, M.: Assessment of hydrologic parameter uncertainty and the worth of multiresponse data, Water Resources Research, 34, 1481–1489, https://doi.org/10.1029/98WR00496, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98WR00496, 1998.

Luhar, M. and Nepf, H. M.: From the blade scale to the reach scale: A characterization of aquatic vegetative drag, Adv. Water Resour., 51, 305–316, 2013.

Marcinkowski, P., Kiczko, A., and Okruszko, T.: Model-based analysis of macrophytes role in the flow distribution in the anastomosing river system, Water, 10, 953, 2018.

Marcinkowski, P., Kiczko, A., and Okruszko, T.: Model-based evaluation of restoration measures efficiency in the anastomosing section of the River Narew, Ecological Engineering, 130, 213–227, https://doi.org/10.1016/J.ECOLENG.2019.02.022, https://www.sciencedirect.com/science/article/abs/pii/S0925857419300722, 2019.

Mertens, W.: Zur frage hydraulischer berechnungen naturnaher fliessgewasser, Wasserwirtschaft, 79, 170–179, 1989.

Myers, W.: Momentum Transfer In A Compound Channel, Journal of Hydraulic Research, 16, 139–150, https://doi.org/10.1080/00221687809499626, https://doi.org/10.1080/00221687809499626, 1978.

Pasche, E. and Rouvé, G.: Overbank Flow with Vegetatively Roughened Flood Plains, J. Hydraul. Eng., 111, 1262–1278, 1985.

Posey, C. J.: Computation of discharge including over-bank flow, Civil engineering, 37, 63, 1967.

Reid, D. E. and Hickin, E. J.: Flow resistance in steep mountain streams, Earth Surface Processes and Landforms, 33, 2211–2240, 2008.

Romanowicz, R. J. and Beven, K. J.: Comments on generalised likelihood uncertainty estimation, Reliability Engineering & System Safety, 91, 1315–1321, 2006.

Romanowicz, R. J. and Kiczko, A.: An event simulation approach to the assessment of flood level frequencies: risk maps for the Warsaw reach of the River Vistula, Hydrological Processes, 30, 2451–2462, https://doi.org/10.1002/hyp.10857, 2016.

Rowiński, P., Västilä, K., Aberle, J., Järvelä, J., and Kalinowska, M. B.: How vegetation can aid in coping with river management challenges: A brief review, Ecohydrology & Hydrobiology, 18, 345–354, 2018.

Sellin, R. H. J.: A laboratory investigation into the interaction between the flow in the channel of a river and that over its flood plain, La Houille Blanche, 20, 793–802, https://doi.org/10.1371/journal.pone.0116943, 1964.

Shiono, K. and Knight, D. W.: Turbulent open-channel flows with variable depth across the channel, J. Fluid Mech., 222, 617, 1991.

Soong, T. W. and DePue, P. M.: Variation of Manning's Coefficient with Channel Stage, Water Resources Center, University of Illinois,
680   U.S.A, 1996.

Västilä, K. and Järvelä, J.: Environmentally preferable two-stage drainage channels: considerations for cohesive sediments and conveyance, International journal of river basin management, 9, 171–180, 2011.

Västilä, K. and Järvelä, J.: Modeling the flow resistance of woody vegetation using physically based properties of the foliage and stem, Water Resour. Res., 50, 229–245, 2014.

685   Västilä, K. and Järvelä, J.: Characterizing natural riparian vegetation for modeling of flow and suspended sediment transport, Journal of Soils and Sediments, 18, 3114–3130, https://doi.org/10.1007/s11368-017-1776-3, 2018.

Västilä, K., Järvelä, J., and Koivusalo, H.: Flow-Vegetation-Sediment Interaction in a Cohesive Compound Channel, J. Hydraul. Eng., 142, 4015 034, 2016.

Warmink, J. J., Straatsma, M. W., Huthoff, F., Booij, M. J., and Hulscher, S. J. M. H.: Uncertainty of design water levels due to combined
690   bed form and vegetation roughness in the D utch R iver W aal, Journal of flood risk management, 6, 302–318, 2013.

Yen, B. C.: Identification Problem of Open-Channel Friction Parameters - discussion, Journal of Hydraulic Engineering, 125, 552–553, 1999.

Yen, B. C.: Open Channel Flow Resistance, J. Hydraul. Eng., 128, 20–39, 2002.

Zinke, P., Olsen, N. R. B., and Bogen, J.: Three-dimensional numerical modelling of levee depositions in a Scandinavian freshwater delta, Geomorphology, 129, 320–333, 2011.

695   **Appendix A: Box-plots for analyzed methods and cases**
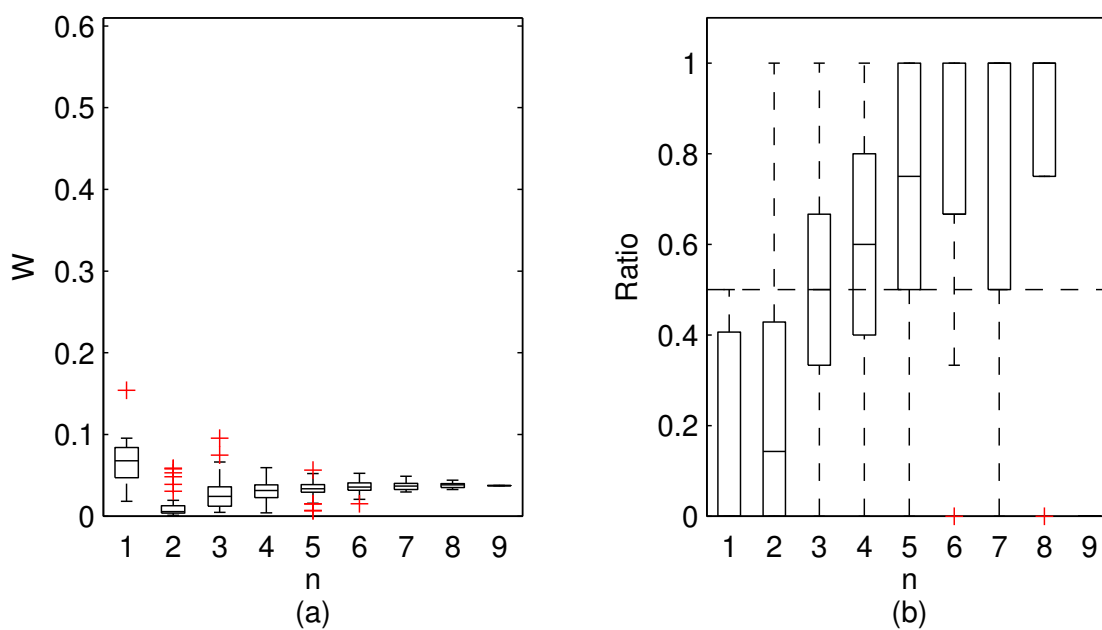
Figure A1. DCM Manning results for the flume case 1, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
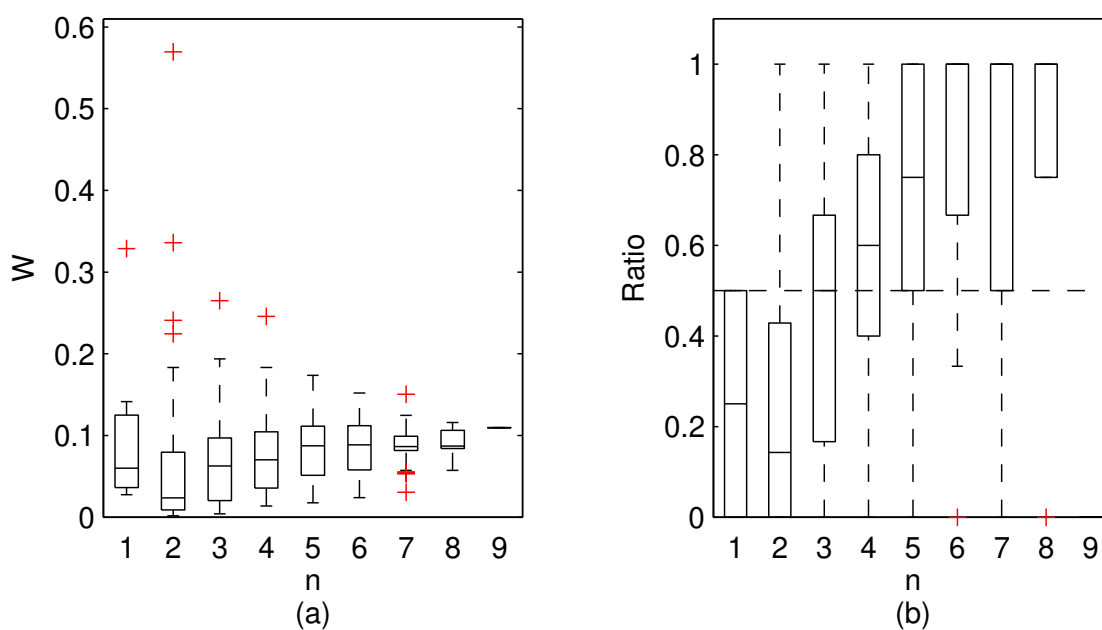
Figure A2. Pasche results for the flume case 1, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
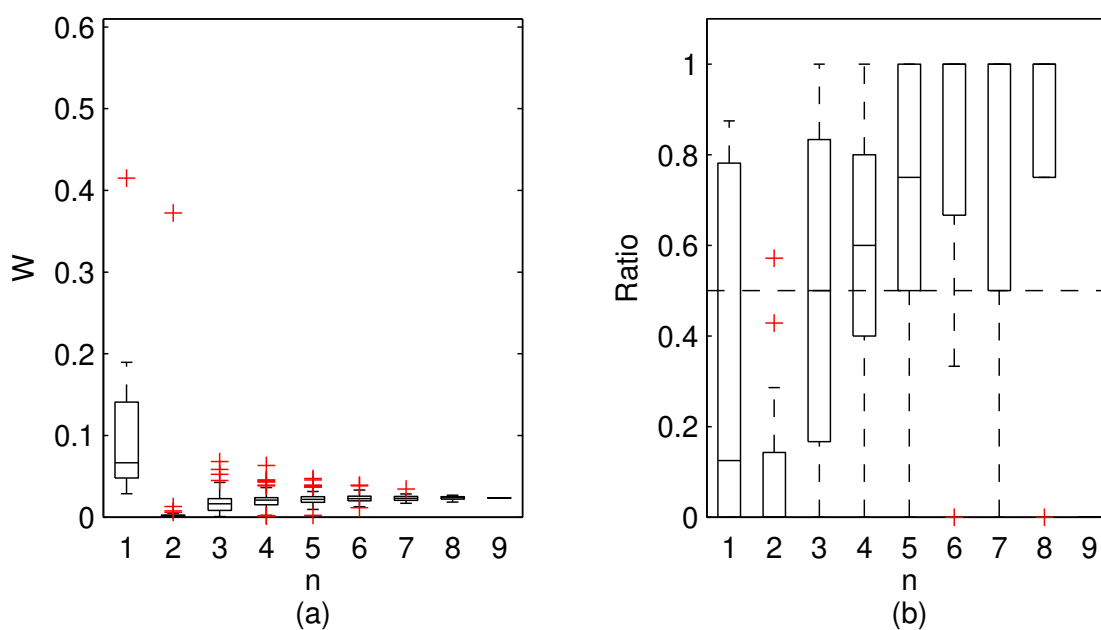
Figure A3. Mertens results for the flume case 1, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
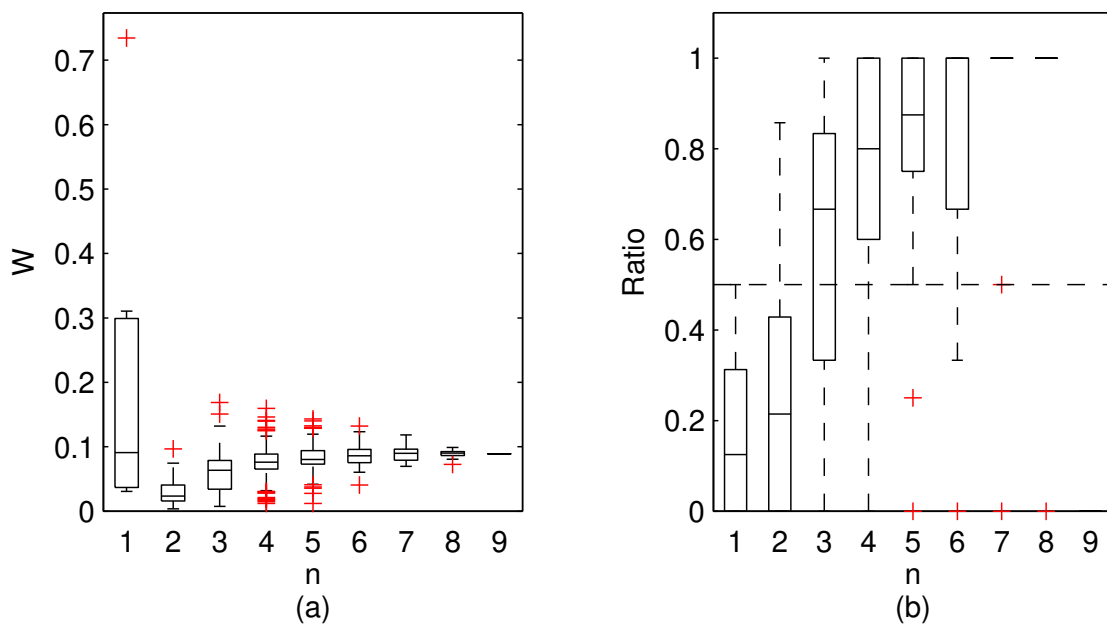
Figure A4. GTLM results for the flume case 1, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
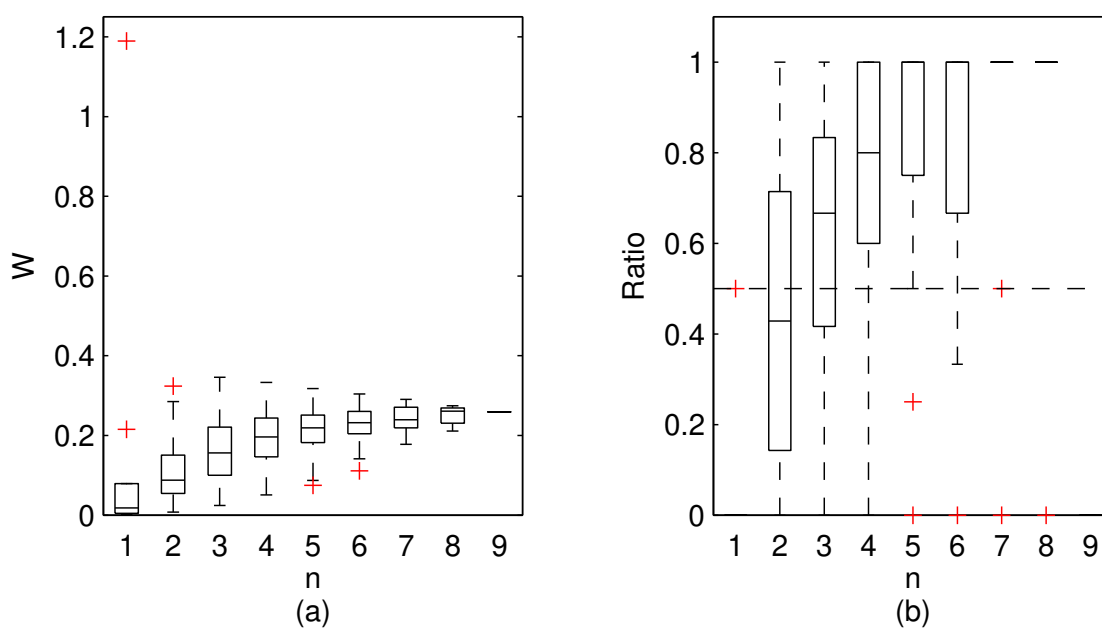
Figure A5. PTLM results for the flume case 1, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
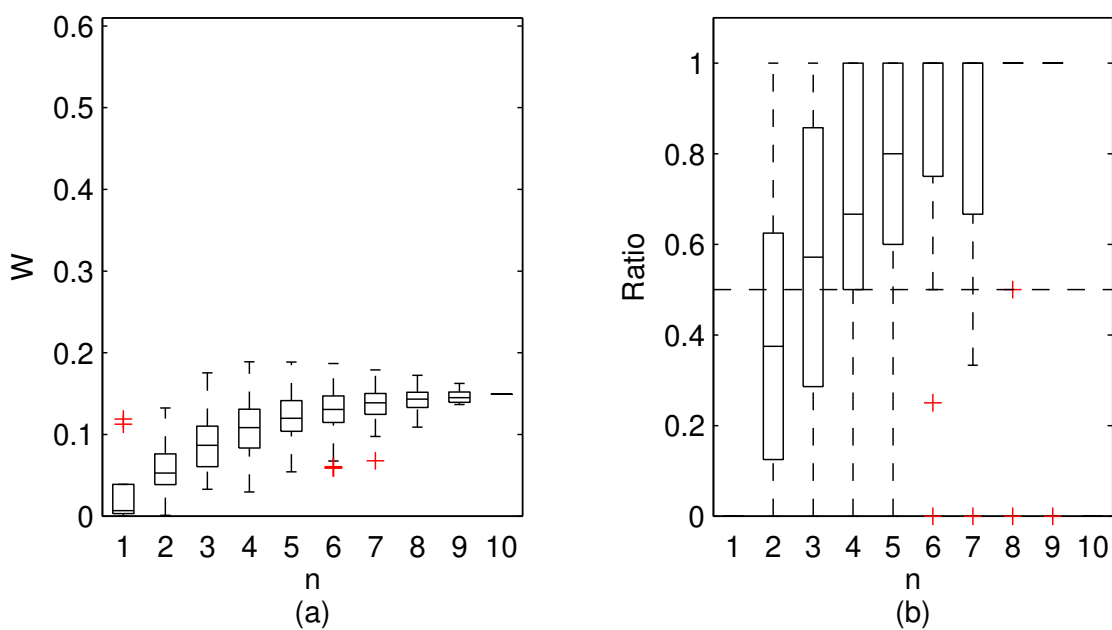
Figure A6. DCM Manning results for the flume case 1, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
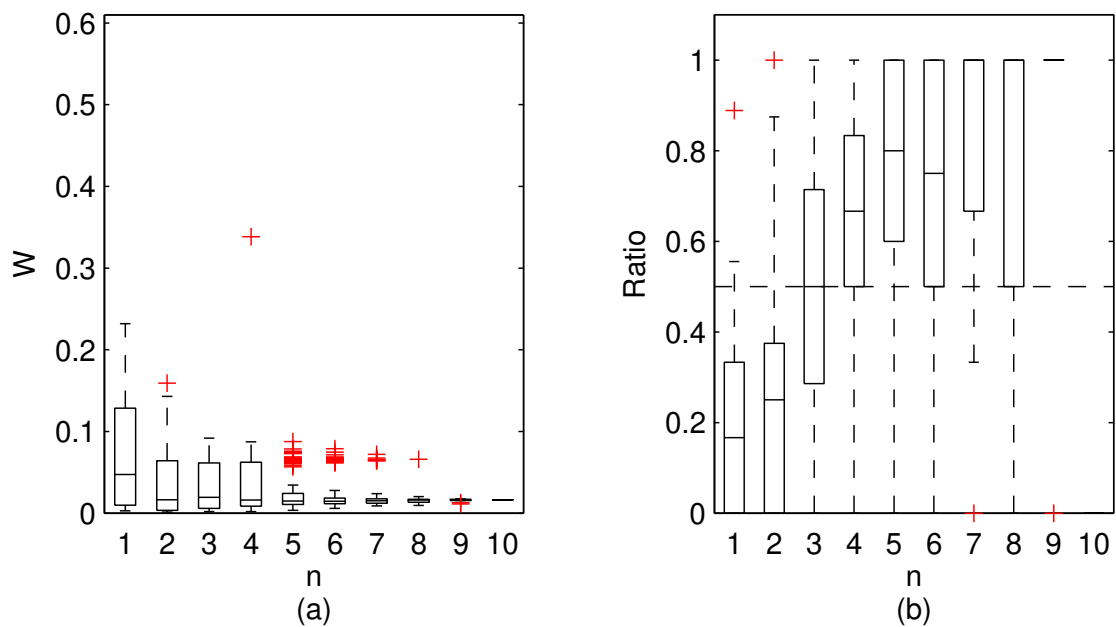
Figure A7. Pasche results for the flume case 1, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
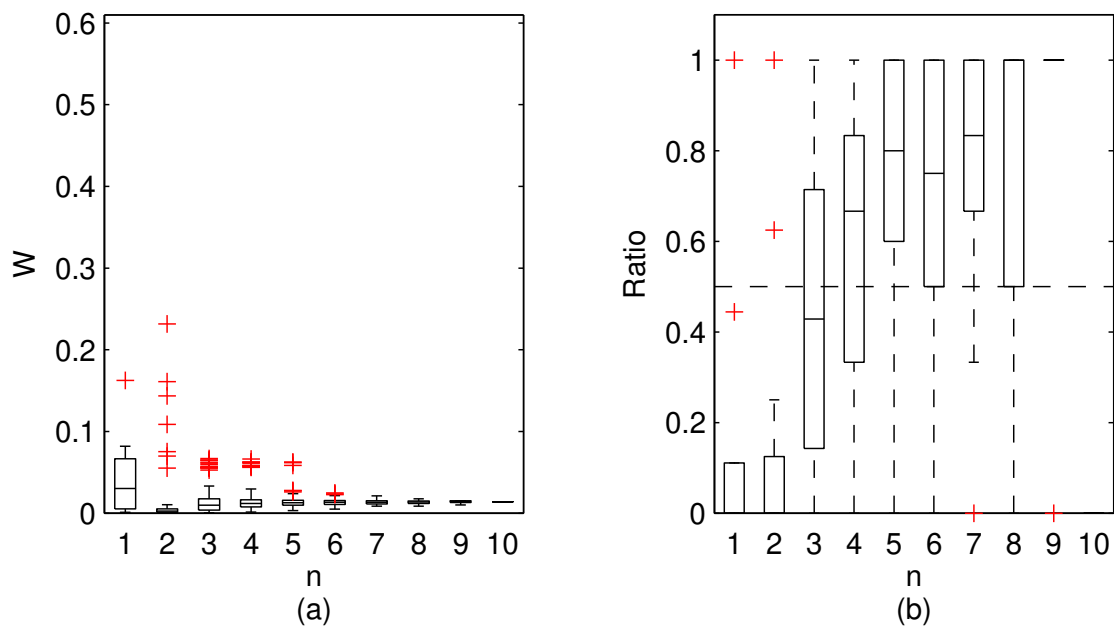
Figure A8. Mertens results for the flume case 1, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
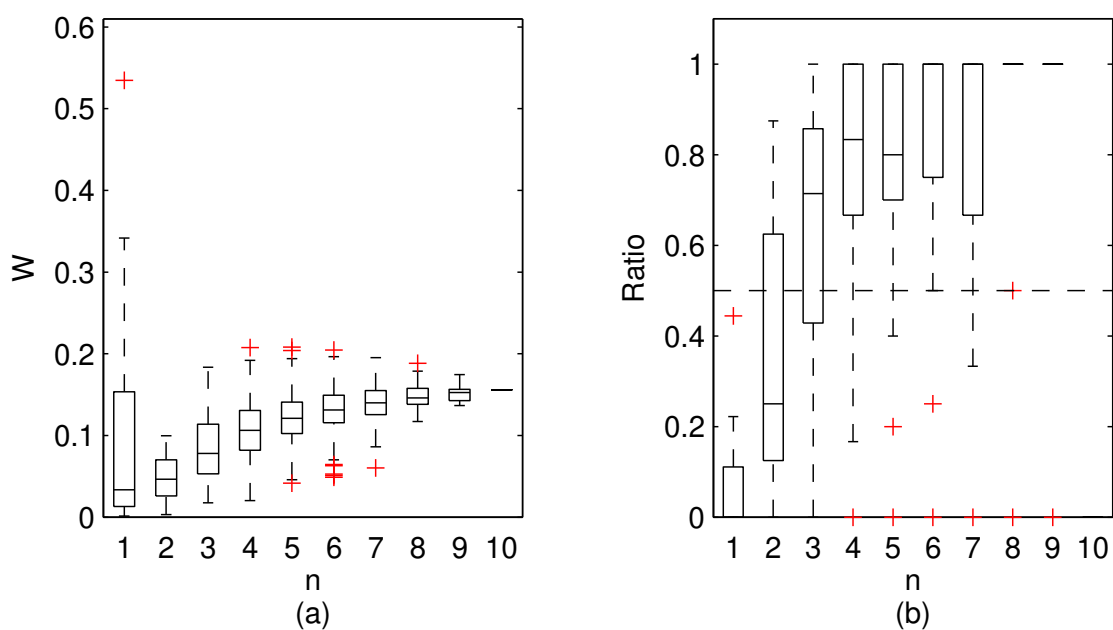
Figure A9. GTLM results for the flume case 1, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
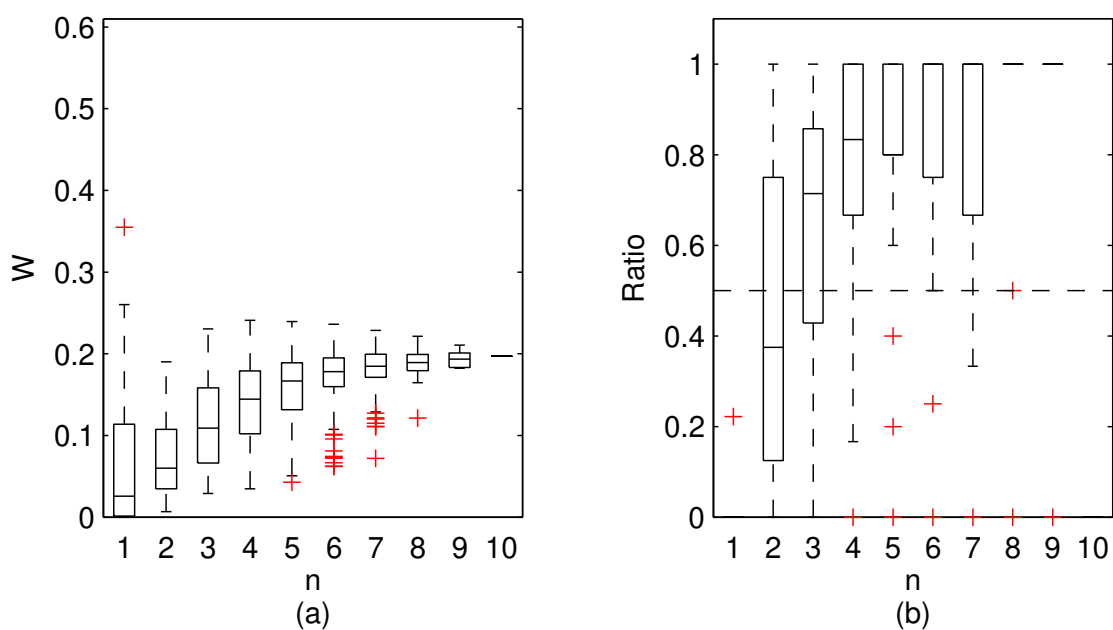
Figure A10. PTLM results for the flume case 1, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
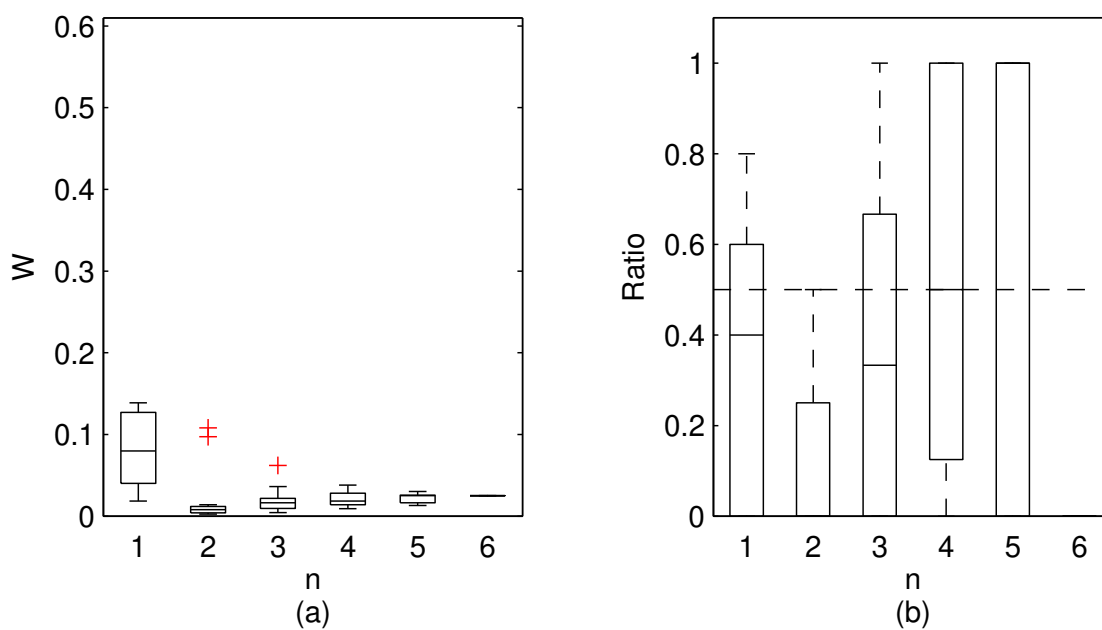
Figure A11. Manning DCM results for Ritobacken case study, Spring 2011, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
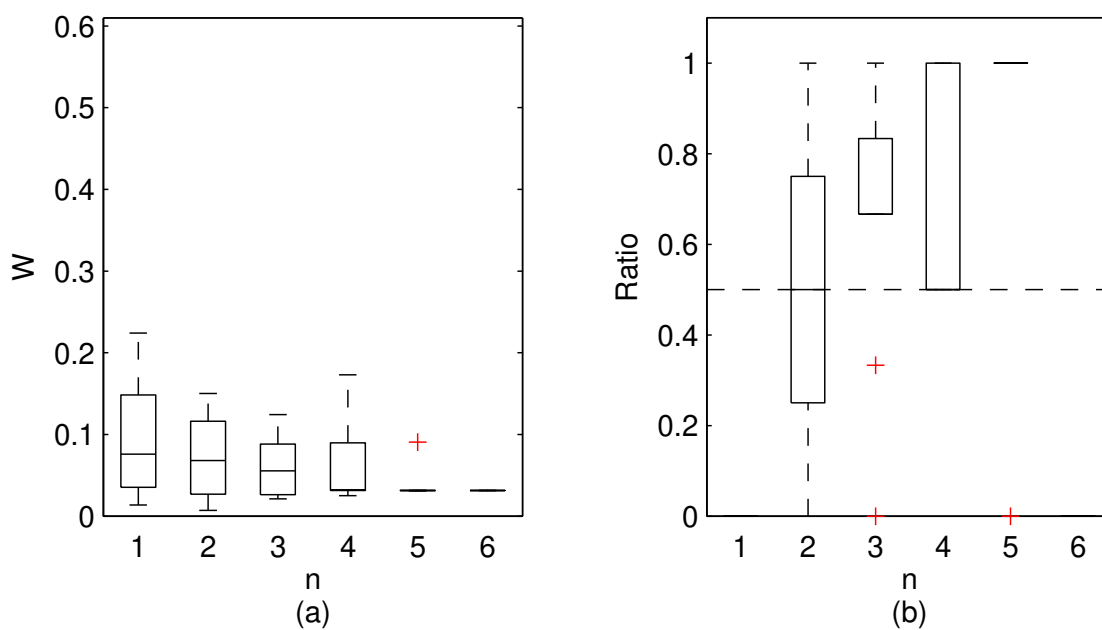
Figure A12. Mertens results for Ritobacken case study, Spring 2011, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
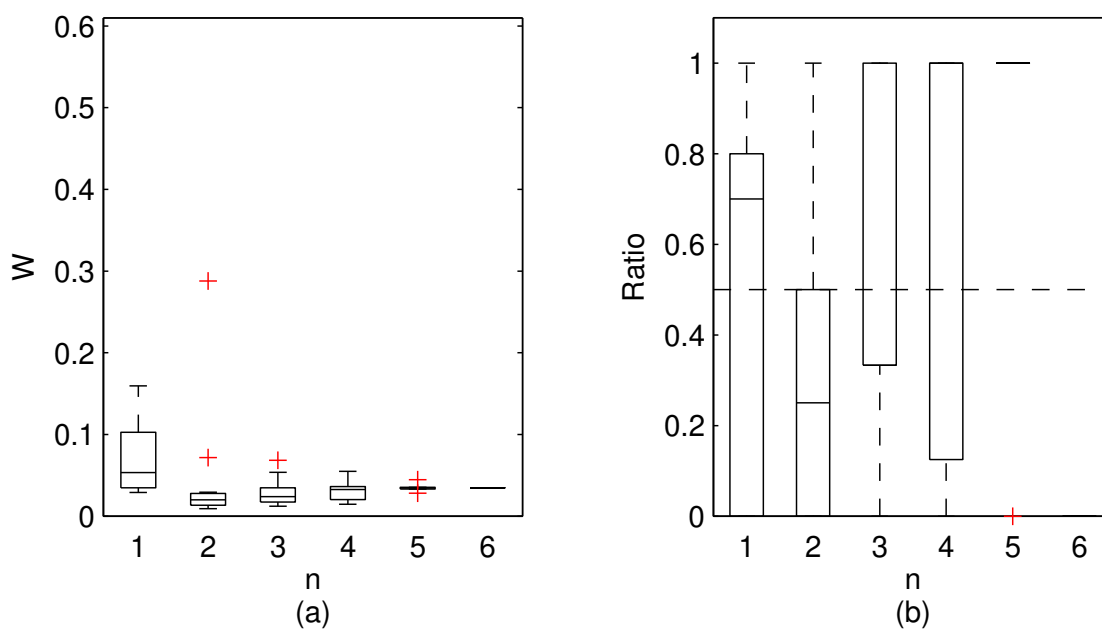
Figure A13. GTLM results for Ritobacken case study, Spring 2011, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
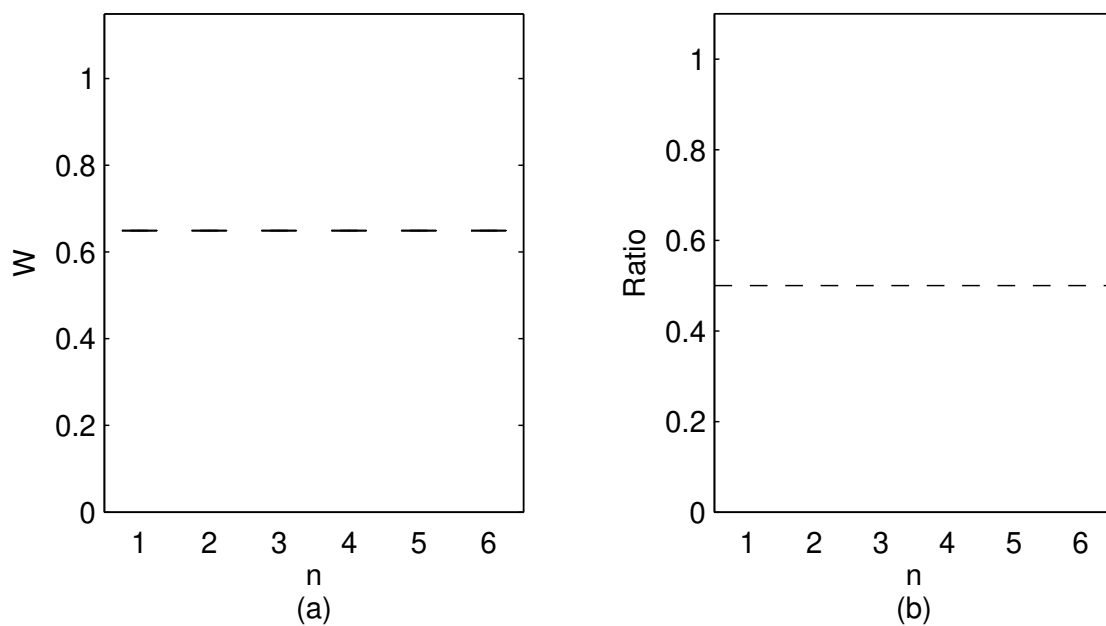
Figure A14. STLM results for Ritobacken case study, Spring 2011, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
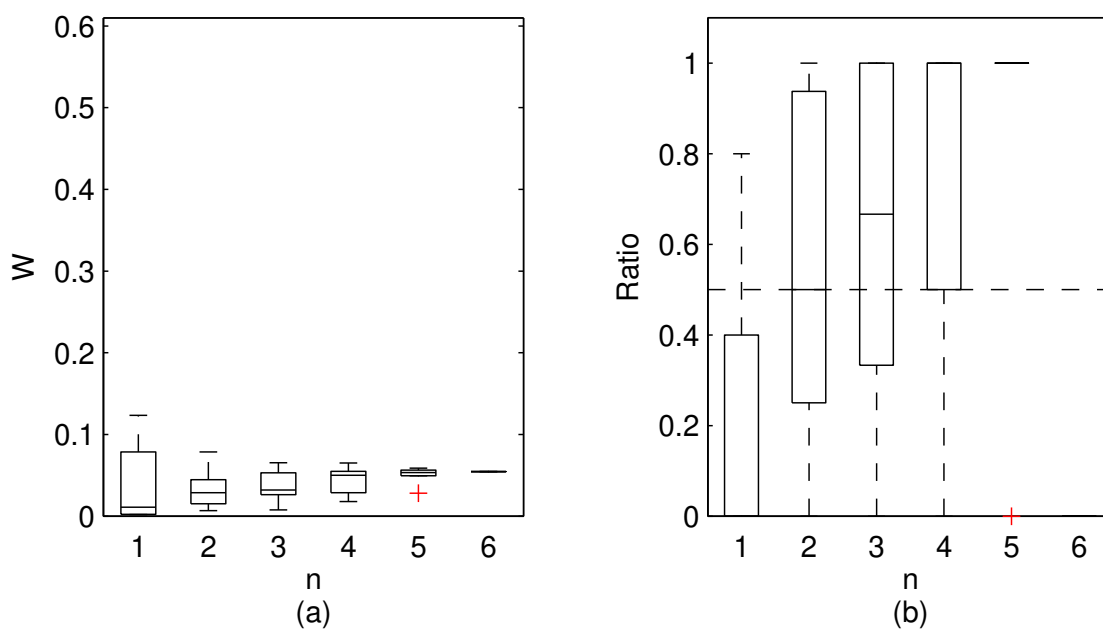
Figure A15. PTLM results for Ritobacken case study, Spring 2011, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
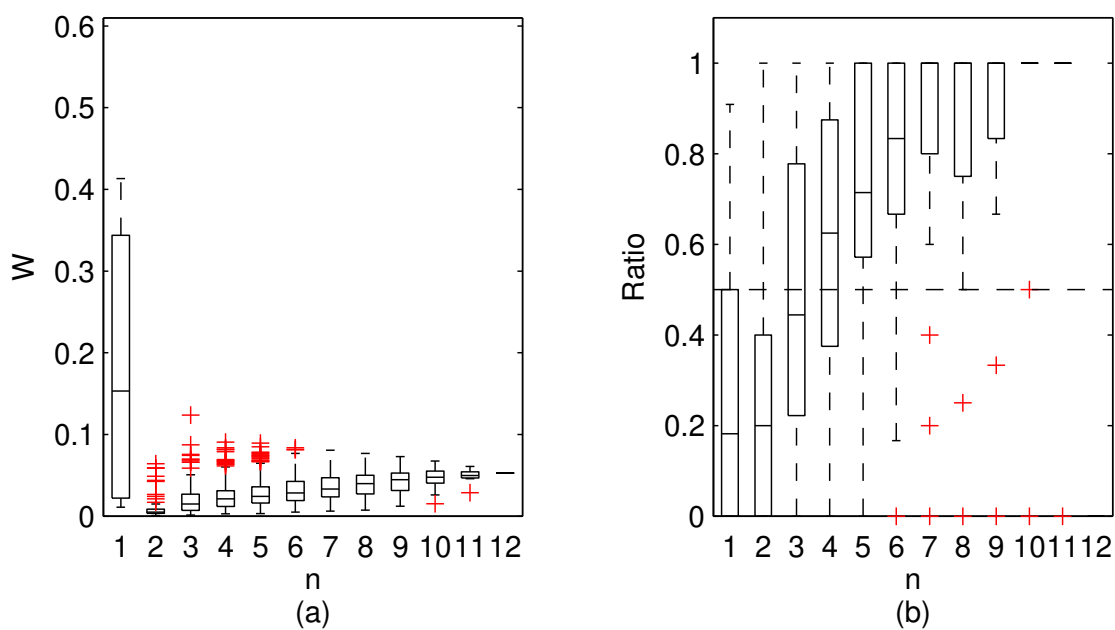
Figure A16. Manning DCM results for Ritobacken case study, Autumn 2011, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
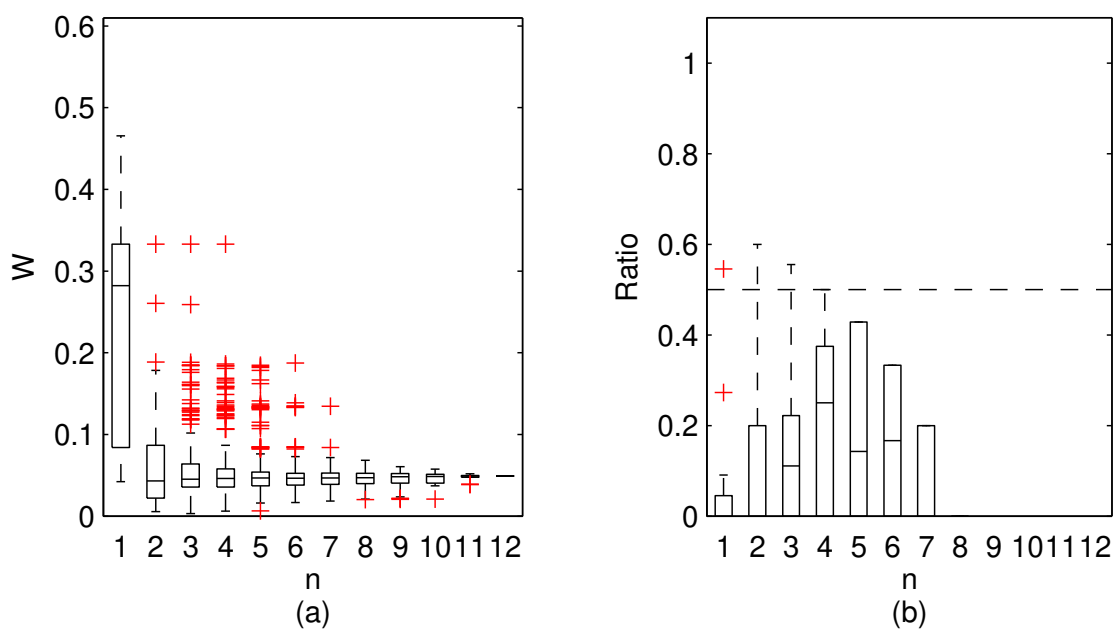
Figure A17. Mertens results for Ritobacken case study, Autumn 2011, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
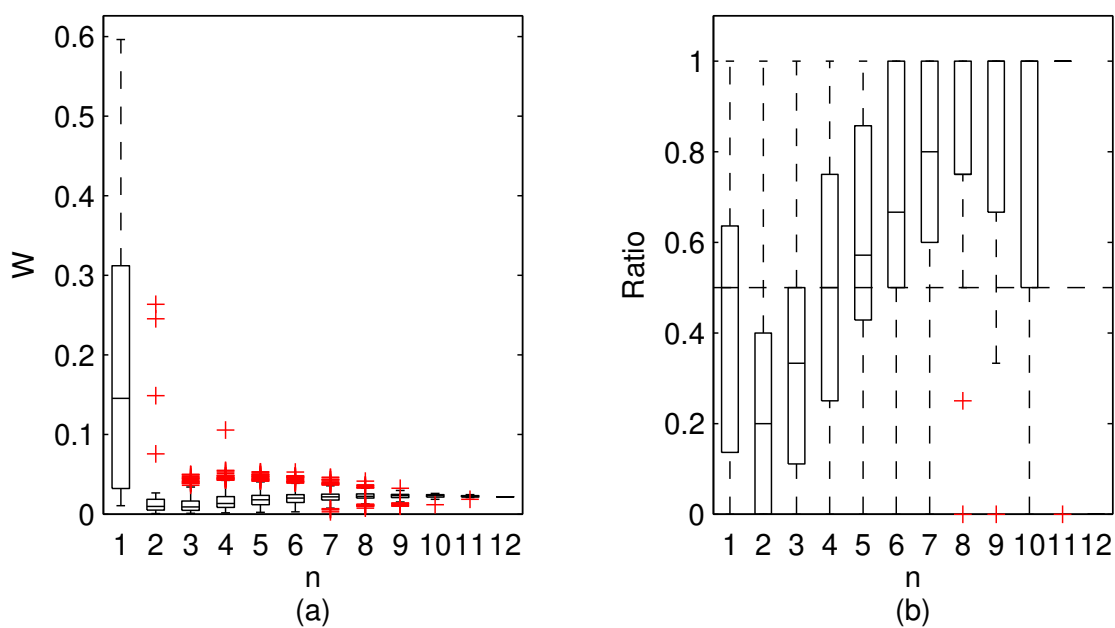
Figure A18. GTLM results for Ritobacken case study, Autumn 2011, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
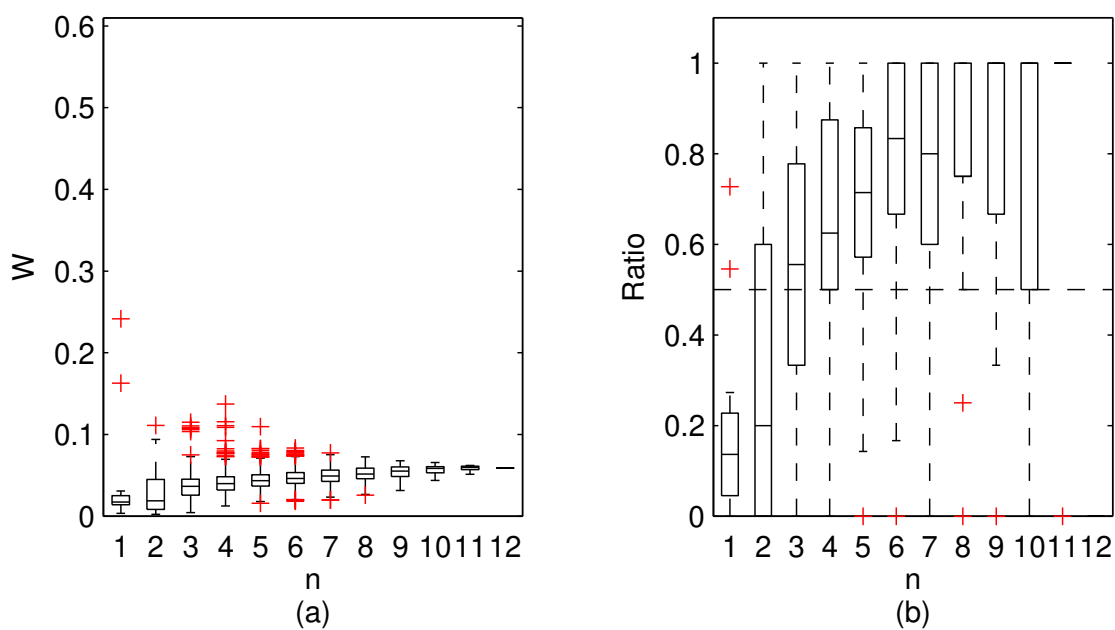
Figure A19. STLM results for Ritobacken case study, Autumn 2011, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
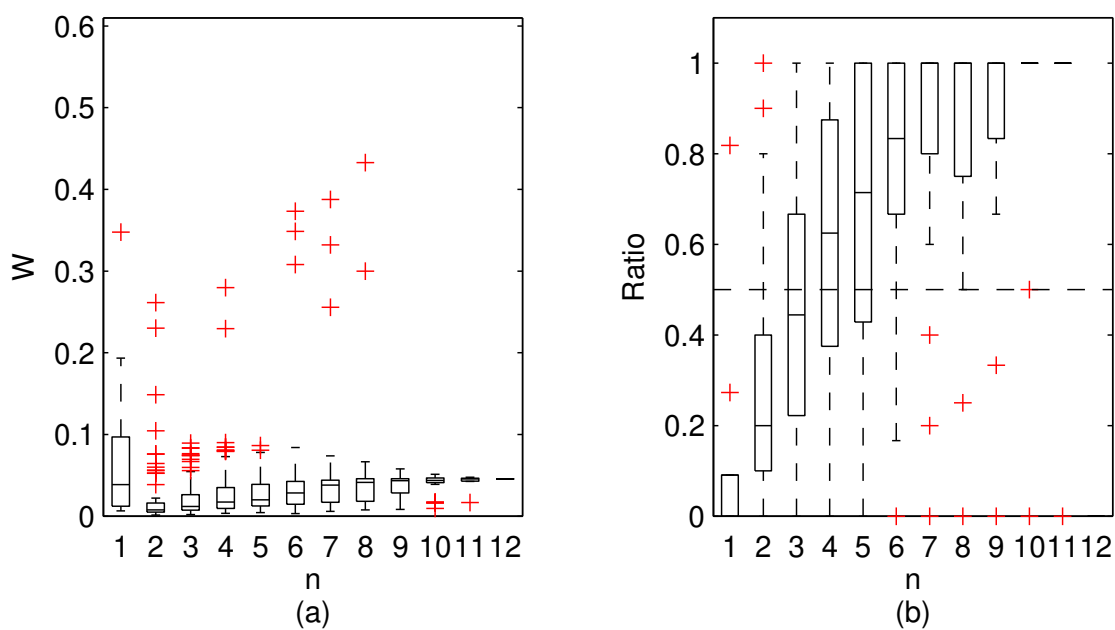
Figure A20. PTLM results for Ritobacken case study, Autumn 2011, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
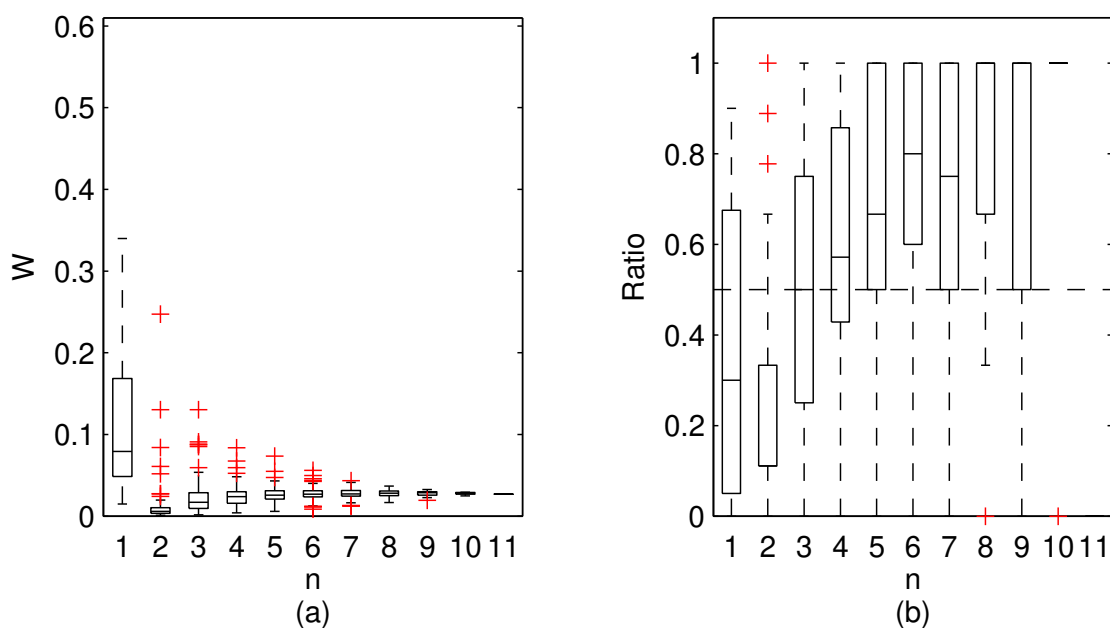
Figure A21. Manning DCM results for Ritobacken case study, Spring 2012, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
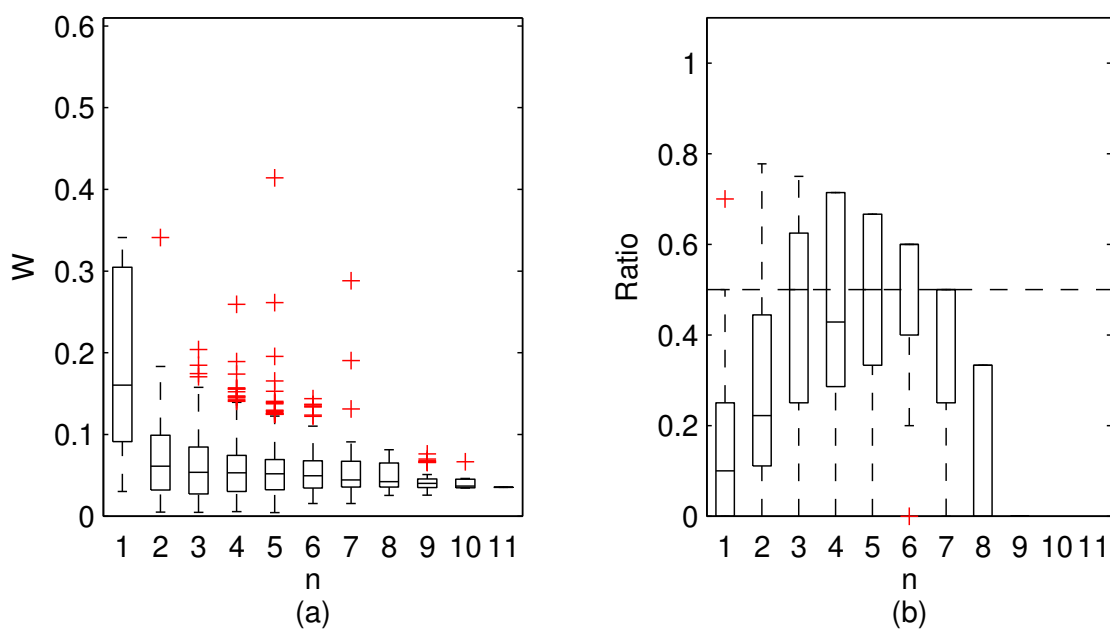
Figure A22. Mertens results for Ritobacken case study, Spring 2012, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
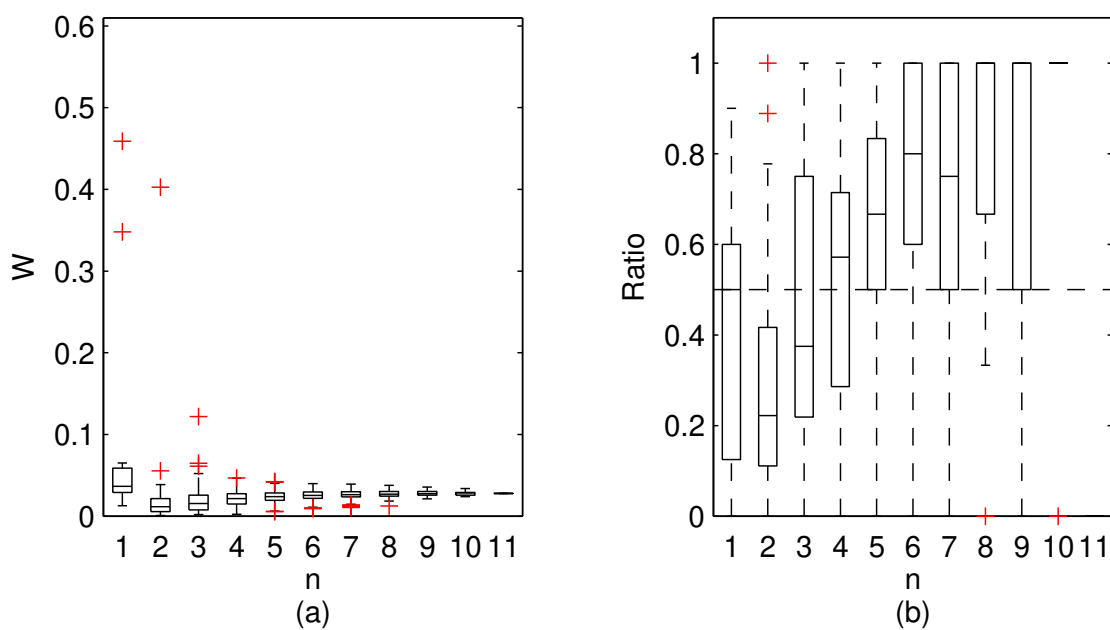
Figure A23. GTLM results for Ritobacken case study, Spring 2012, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
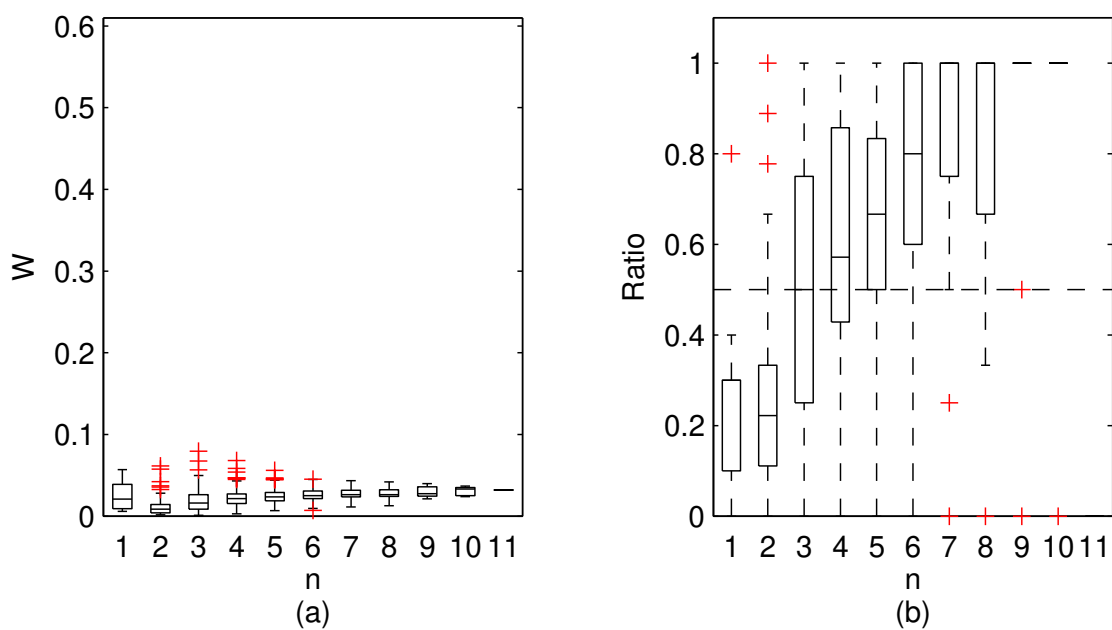
Figure A24. STLM results for Ritobacken case study, Spring 2012, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)
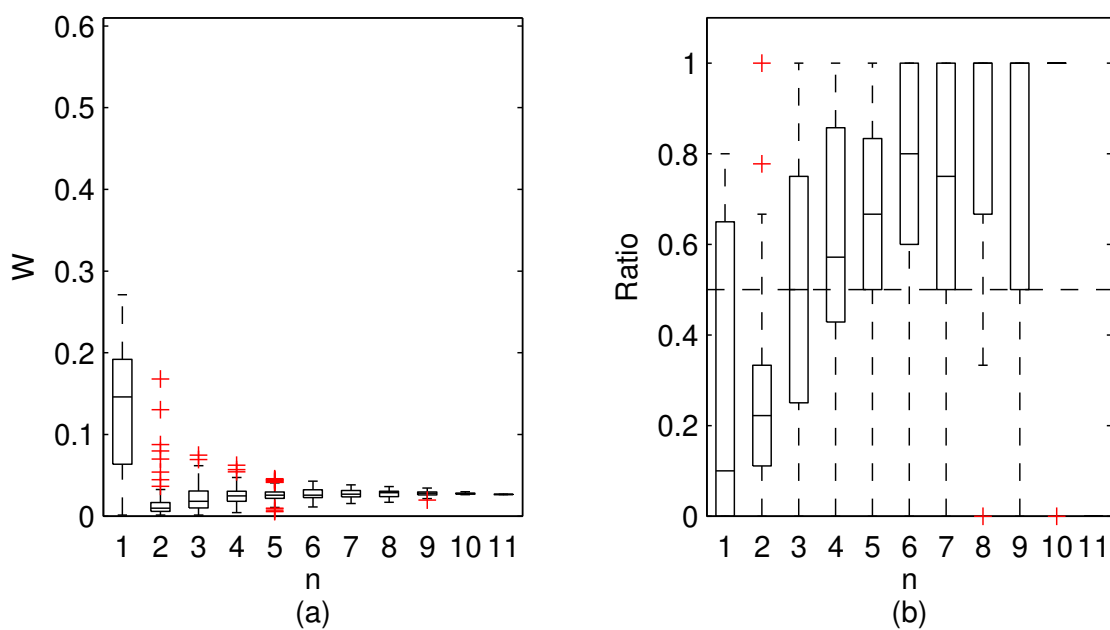
Figure A25. PTLM results for Ritobacken case study, Spring 2012, (a) Averaged relative confidence widths W as a function of observation set size n used for model identification; (b) Ratio of verification points enclosed by the confidence intervals (1 denotes all points within intervals, box spans over 25% and 75% quantile, median is given with horizontal line, whiskers indicate the result extent, cross marks are for extreme values)