Calibrating 1D hydrodynamic river models in the absence of cross-sectional geometry: A new parameterization scheme

Liguang Jiang^{1, 2}, Silja Westphal Christensen^{2, 3}, Peter Bauer-Gottwein²

¹School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, 518055, China

²Department of Environmental Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark ³Department of Applied Mathematics and Computer Science, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

Corresponding author: Liguang Jiang (jianglg@sustech.edu.cn)

Contents of this file

Text S1 to S3 Figures S1 to S11

Introduction

This document contains text and figures supporting the description and explanation in the main text. Below is a brief overview of the supporting information.

- Text S1 provides an overview of the Songhua river as well as the 1D hydrodynamic model.
- Text S2 describes the details of the objective function and optimization used to calibrate models.
- Text S3 gives the seven scenarios which are used to test the effect of calibration data on model performance.
- Figure S1 depicts the locations and settings of six rivers used to prove the hydraulic relationships.
- Figure S2 is similar to Figure 1, but a uniform Manning's number of 0.03 was used to calculate conveyance.
- Figure S3 is similar to Figure 2, but a uniform Manning's number of 0.03 was used to present the linearity of gamma ~ alpha and delta ~ beta when pooling all data together.
- Figure S4 depicts the linear relationships between gamma ~ alpha and delta ~ beta for six rivers separately.
- Figure S5 draws a detailed map of the Songhua river as well as the distribution of the 23 cross sections.
- Figure S6 displays the temporal and spatial distribution of water levels and widths, which are used to calibrate models.
- Figure S7 shows the effect of different configurations of calibration data sets on model simulation of water level at two gauging stations.
- Figure S8 compares the calibrated curves using three scenarios of width observations, i.e. all available widths and water levels (#1), one width observation per two-km reach and water levels (#2), and one width observation per five-km reach and water levels (#3).
- Figure S9 illustrates the performance of model simulated discharge compared against in situ discharge at to gauging stations.
- Figure S10 shows model simulated water level, width, and datum using different calibration data sets described in Text S3.
- Figure S11 shows how upstream boundary (rainfall-runoff model simulated discharge) affects model performance in terms of water level at four locations.

Text S1. Study site and MIKE Hydro River model

Songhua River is a major tributary of the Amur or Heilongjiang River. It is the third largest river in China in terms of discharge. The river has two sources, i.e. Nenjiang and Second Songhua rivers, originating from the Greater Khingan Range in the north and the Paektu Moutain in the south, respectively, and drains an area of 556,800 sq. km. At Sanchahe, two rivers merge into one, called Songhua river. It runs 840 km northeastward before emptying into the Amur River (Songliao River Conservancy Commission, 2004, 2015).

In this study, we focus on the middle reach of the main Songhua River, between Harbin and Jiamusi (Figure S5). The reason we selected this reach is twofold: firstly, it is wide enough to have good altimetry data as shown in previous study (Jiang et al., 2017) and secondly, we have access to in situ data of the several hydrometric stations across this region. This reach covers an area of 138500 sq. km and stretches 433 km long. See Figure S5 for a better understanding. The river is completely frozen during winter. The altimetry is not able to reflect the real WSE beneath the ice. Therefore, we only consider ice-free period in this study.

A 1D river model is built using the MIKE Hydro River (DHI, 2017). The river network is set using the center line of the reach, while cross sections are equally distributed along the 433 km reach (see Figure S5). To run the model, boundary conditions are needed. In this study, the upstream boundary is in situ discharge at the Harbin gauging station. The downstream boundary is set as normal depth that can be calculated based on Manning's equation. Lateral inflows are also included using either in situ or simulations from the Xinanjiang rainfall-runoff model (Jiang et al., 2019).

Text S2. Objective function and optimization

Considering the regularization terms as well as data uncertainty, the objective function is formulated as:

$$misfit = w \left(\frac{1}{N_{h}} \left(\frac{h_{s}-h_{o}}{\sigma_{h}}\right)^{2}\right) + (1-w) \left(\frac{1}{N_{b}} \left(\frac{b_{s}-b_{o}}{\sigma_{b}}\right)^{2}\right)$$
(S1)

$$reg = \lambda_{\gamma} \left(\frac{1}{N} \left(\frac{L\gamma}{\sigma_{\gamma}}\right)^{2}\right) + \lambda_{\delta} \left(\frac{1}{N} \left(\frac{L\delta}{\sigma_{\delta}}\right)^{2}\right) + \lambda_{p1} \left(\frac{p_{1}+1.4}{\sigma_{p1}}\right)^{2} + \lambda_{p2} \left(\frac{p_{2}-1}{\sigma_{p2}}\right)^{2} + \lambda_{p3} \left(\frac{p_{3}+0.7}{\sigma_{p3}}\right)^{2} + \lambda_{p4} \left(\frac{p_{4}-1}{\sigma_{p4}}\right)^{2} + (1-\lambda_{\gamma}-\lambda_{\delta}-\lambda_{p1}-\lambda_{p2}-\lambda_{p3}-\lambda_{p4}) \left(\frac{1}{N} \left(\frac{Lz}{\sigma_{z}}\right)^{2}\right)$$
(S2)

$$\phi(\mathbf{X}) = \lambda \operatorname{misfit} + (1 - \lambda) \operatorname{reg}$$
(S3)

Where, **X** is vector containing parameters of γ , δ , p_1 , p_2 , p_3 , p_4 , and z; w is a weighting factor balancing the fitness of water level and width; h_s , h_o , N_h , σ_h are simulated water level, observed water level, number of water levels, and the uncertainty of observed water level; b_s , b_o , N_b , σ_b are simulated width, observed width, number of widths, and the uncertainty of observed width; λ_γ , λ_δ , λ_{p1} , λ_{p2} , λ_{p3} , λ_{p4} are regularization parameters; σ_γ , σ_δ , σ_{p1} , σ_{p2} , σ_{p3} , σ_{p4} and σ_z are the standard deviations indicating how spread the parameters are; N is the number of cross sections; and L is the first-order regularization roughening matrix, which is a finite-difference approximation:

$$\boldsymbol{L} = \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 & \\ & & & & 1 & -1 \end{bmatrix}$$

In our case study, the weighting factor *w* is 0.5 by treating water level and river width observations equally important. The uncertainties of water level and width are 0.5 m and 99 m according to (Jiang et al., 2017) and (Yang et al., 2020). The standard deviations of σ_{γ} and σ_{δ} are 0.7 and 0.4, respectively, which are similar for relatively large rivers (see γ and δ distributions in Figure S2). The deviations of p_1 , p_2 , p_3 , p_4 are 0.02, 0.01, 0.02, and 0.01, respectively, given that those parameters vary slightly (see Figures S2 and S3). For the datum Z0, it is the sum of parameter z and a constant value which is estimated from the average water level subtracting the depth of 5 m. The deviation of z is 0.5 m. The regularization parameters, i.e. λ_{γ} , λ_{δ} , λ_{p1} , λ_{p2} , λ_{p3} , λ_{p4} are empirically set as 0.1, 0.1, 0.15, 0.15, 0.15, 0.15, respectively, based on their smoothness.

We use the Levenberg-Marquardt (LM) algorithm (Marquardt, 1963) to optimize the objective given above. The LM algorithm is an iterative process that steps its way toward a minimum solution of the objective. In each iterative step, the Jacobian of the objective function is calculated. However, computing the Jacobian in every iteration is computational expense and in some cases the Jacobian does not change and thus, evaluation of the Jacobian can be unnecessarily costly. Instead, Broyden's rank-one update of the Jacobian (Broyden, 1965) is more efficient (Madsen et al., 2004). We use the Immoptibox toolbox (Nielsen & Völcker, 2010) to optimize the objective function. Given that LM finds only a local minimum, an ensemble of 10 calibrations are carried out with different initial guesses to avoid convergence to a local minimum.

The implementation of calibration is in Matlab. We also wrote some C# script to modify and dump MIKE Hydro River parameters and simulations. Specifically, for each iteration of the optimization, the parameters generated by LM algorithm as well as the further calculated flow area and conveyance are passed to a C# script that overrides the setup of MIKE Hydro River model. Then the model is executed, and the results are dumped into Matlab. In such a way, the model is eventually calibrated.

Text S3. Scenario testing

As mentioned in the main text, to test the capability of different data sets to constrain model parameters, three basic scenarios are used based on the type of data sets. That is, **calibration #1 only uses altimetry derived WSE; calibration #2 only uses imagery derived width; and calibration #3 use both WSE and width**.

WSE is from the CryoSat-2 altimeter which has been collecting data since 2010. The reason we use CryoSat-2 is that the spatial sampling is very dense compared to other missions. And this spatial sampling is more useful as shown in the previous studies (Jiang et al., 2019; Schneider et al., 2018). Width is mainly derived from Landsat images using the RivWidthCloud algorithm (Yang et al., 2020) in Google Earth Engine. RivWidthCloud is an app developed by Yang et al. (2020). We followed the procedures described in Yang et al. and we refer the reader to that paper for a detailed introduction. We used Landsat 5 and Landsat 8 based on the cloudiness and dates. Specifically, if the river part of the image is cloud-free, we select it. Images collected in Dec to early April are excluded. Below are the image IDs we used.

LT05_L1TP_115027_20070925_20161112_01_T1	
LT05_L1TP_115027_20080522_20170613_01_T1	
LT05_L1TP_115027_20080607_20161031_01_T1	
LT05_L1TP_115027_20080623_20161030_01_T1	
LT05_L1TP_115027_20080725_20161030_01_T1	
LT05_L1TP_115027_20080927_20161029_01_T1	
LT05_L1TP_115027_20081013_20161029_01_T1	
LT05_L1TP_115027_20090728_20161023_01_T1	
LT05_L1TP_115027_20090930_20161019_01_T1	
LT05_L1TP_115027_20091016_20161019_01_T1	
LT05_L1TP_115027_20100901_20161013_01_T1	
LT05_L1TP_115027_20100917_20161013_01_T1	LC08_L1TP_115027_20130707_20170504_01_T1
LT05_L1TP_115027_20101019_20161012_01_T1	LC08_L1TP_115027_20130723_20170503_01_T1
LT05_L1TP_115027_20110616_20161009_01_T1	LC08_L1TP_115027_20130909_20170502_01_T1
LT05_L1TP_115027_20110904_20161006_01_T1	LC08_L1TP_115027_20131027_20170429_01_T1
LT05_L1TP_115027_20110920_20161006_01_T1	LC08_L1TP_115027_20140624_20170421_01_T1
LT05_L1TP_115027_20111006_20161005_01_T1	LC08_L1TP_115027_20140726_20170421_01_T1
LT05_L1TP_116027_20070527_20161114_01_T1	LC08_L1TP_115027_20140912_20170419_01_T1
LT05_L1TP_116027_20070714_20161112_01_T1	LC08_L1TP_115027_20140928_20170419_01_T1
LT05_L1TP_116027_20070831_20161111_01_T1	LC08_L1TP_116027_20130714_20170503_01_T1
LT05_L1TP_116027_20071002_20161112_01_T1	LC08_L1TP_116027_20130916_20170502_01_T1
LT05_L1TP_116027_20080427_20161101_01_T1	LC08_L1TP_116027_20131018_20170429_01_T1
LT05_L1TP_116027_20080513_20161031_01_T1	LC08_L1TP_116027_20140428_20170423_01_T1
LT05_L1TP_116027_20080630_20161030_01_T1	LC08_L1TP_116027_20140919_20170419_01_T1
LT05_L1TP_116027_20080817_20170613_01_T1	LC08_L1TP_116027_20141005_20170418_01_T1
LT05_L1TP_116027_20081004_20161029_01_T1	LC08_L1TP_116027_20141021_20170418_01_T1
LT05_L1TP_116027_20081020_20161029_01_T1	
LT05_L1TP_116027_20090719_20161023_01_T1	
LT05_L1TP_116027_20090804_20161023_01_T1	
LT05_L1TP_116027_20091007_20161024_01_T1	
LT05_L1TP_116027_20100503_20161015_01_T1	
LT05_L1TP_116027_20100908_20161013_01_T1	
LT05_L1TP_116027_20110522_20161_009_01_T1	
LT05_L1TP_116027_20110725_20161007_01_T1	
LT05_L1TP_116027_20110826_20161006_01_T1	
LT05_L1TP_116027_20110911_20161006_01_T1	
LT05_L1TP_116027_20111029_20161005_01_T1	

In total, there are 10022 observations of river width. To reduce the redundancy (30m interval), width is randomly selected for each 2 or 5 km reach regardless of the timing for each model calibration. Therefore, **three scenarios of width** (i.e. the number of width observations are 10022, 219 and 88, respectively) are also tested. Given that only 262 observations of WSE are available, no further exploration of the effect of WSE data is performed. We test seven scenarios of observations to calibrate models.

Scenario	Num. of WSE	Num. of width
All WSE only	262	0

All width only	0	10022
All WSE + all width	262	10022
All WSE + one width per 2 km	262	219
All WSE + one width per 5 km	262	88
One width per 2 km	0	219
One width per 5 km	0	88



Figure S1. Location and river setting of six rivers. Grey short lines indicate cross sections.



Figure S2. Relationship between logarithmic depth and logarithmic area and logarithmic conveyance. Similar to **Figure 1** but a uniform Manning's number of 0.03 was used to calculate conveyance. This results in a stronger linear relationship. However, a uniform Manning's number is not very realistic in natural rivers.



Figure S3. Linear relationships between gamma ~ alpha and delta ~ beta using data of all six rivers. Similar to **Figure 2**, but a uniform Manning's number of 0.03 was used.



Figure S4. Linear relationships between gamma ~ alpha and delta ~ beta for six rivers. Randomly generated Manning's number in range of $0.015 \sim 0.05$ for each cross section was used to calculate conveyance. The number of cross sections, coefficient of determination, and regression coefficients are labeled in each plot.



Figure S5. Study area. The reach studied is 433 km long between Harbin and Jiamusi. There are 23 cross sections evenly distributed along this reach as shown in the lower map.



Figure S6. Temporal and spatial distribution of water levels and widths, which are used to calibrate the model. Given that the river is frozen in cold season, only data in warm seasons are used. Landsat-5/8 images with low cloud cover (visual checked in Google Earth Engine) are selected to generate river width.



Figure S7. Water level validation at two gauging stations using different configurations of calibration data sets. Note, W indicate width observation; WSE indicate water level observation.



Figure S8. Calibrated curves using **three scenarios of width** observations, i.e. all available widths and water levels (#1), one width observation per two-km reach and water levels (#2), and one width observation per five-km reach and water levels (#3). Chainage is given in each plot.



Figure S9. Validation of simulated discharge at two gauging stations. Note, in each plot, results of the median and individual simulations of an ensemble of 30 calibrations (three scenarios of width observations, see Text S3) are shown.



Figure S10. Comparison of model simulated water level, width, and datum using different calibration data sets. (a) only water level data are used; (b-d) increasing number (one per 5-km, one per 2-km, and all width) of widths are used without any water level data; (e-f) increasing number (one per 5-km, one per 2-km, and all width) of widths are used besides water level data.



Figure S11. Similar to **Figure 4**, but upstream boundary is rainfall-runoff model simulation instead of in-situ discharge, which are shown in (e). Note that, NAM model was not calibrated at this location.

References

- Broyden, C. G. (1965). A class of methods for solving nonlinear simultaneous equations. *Mathematics of Computation*, 19(92), 577–577. https://doi.org/10.1090/S0025-5718-1965-0198670-6
- DHI. (2017). MIKE HYDRO River User Guide. Copenhagen: DHI.
- Jiang, L., Madsen, H., & Bauer-Gottwein, P. (2019). Simultaneous calibration of multiple hydrodynamic model parameters using satellite altimetry observations of water surface elevation in the Songhua River. *Remote Sensing of Environment*, 225, 229–247. https://doi.org/10.1016/j.rse.2019.03.014
- Jiang, L., Nielsen, K., Andersen, O. B., & Bauer-Gottwein, P. (2017). CryoSat-2 radar altimetry for monitoring freshwater resources of China. *Remote Sensing of Environment*, 200, 125–139. https://doi.org/10.1016/j.rse.2017.08.015
- Madsen, K., Nielsen, H. B., & Tingleff, O. (2004). *Methods for Non-Linear Least Squares Problems (2nd ed.)*.
- Marquardt, D. W. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, *11*(2), 431–441. https://doi.org/10.1137/0111030
- Nielsen, H. B., & Völcker, C. (2010). IMMOPTIBOX: A Matlab Toolbox for Optimization and Data Fitting. Technical University of Denmark, DTU Informatics, Building 321.
- Schneider, R., Tarpanelli, A., Nielsen, K., Madsen, H., & Bauer-Gottwein, P. (2018). Evaluation of multimode CryoSat-2 altimetry data over the Po River against in situ data and a hydrodynamic model. *Advances in Water Resources*, 112(August 2017), 17–26. https://doi.org/10.1016/j.advwatres.2017.11.027

Songliao River Conservancy Commission. (2004). The Songhua River. Changchun, China.

- Songliao River Conservancy Commission. (2015). *Water resources bulletin of Songhua & Liao River*. Changchun, China. Retrieved from http://www.slwr.gov.cn/
- Yang, X., Pavelsky, T. M., Allen, G. H., & Donchyts, G. (2020). RivWidthCloud: An Automated Google Earth Engine Algorithm for River Width Extraction From Remotely Sensed Imagery. *IEEE Geoscience* and Remote Sensing Letters, 17(2), 217–221. https://doi.org/10.1109/LGRS.2019.2920225