Geophysical Research Letters

Supporting Information for

Turning lakes into river gauges using the LakeFlow algorithm

Ryan M. Riggs[1], George H. Allen[2], Craig B. Brinkerhoff[3], Md. Safat Sikder[4], Jida Wang[4]

(1)Department of Geography, Texas A&M University, College Station, TX, USA, (2)Department of Geosciences, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, (3)Department of Civil and Environmental Engineering, University of Massachusetts Amherst, Amherst, MA, USA, (4)Department of Geography and Geospatial Sciences, Kansas State University, Manhattan, KS, USA.

Contents of this file

Introduction

This supporting information contains text describing the development of the synthetic dataset and details on the ancillary datasets used in the analysis. In addition, a table of performance metrics is included as well as figures that support the results discussed in the manuscript.

Text S1: Developing SWOT-like synthetic dataset

To develop the synthetic SWOT dataset, we leverage gauge records from the USGS (U.S. Geological Survey, 2022), Landsat-based Global Surface Water (GSW) dataset (Pekel et al., 2016), and prior channel attributes in the SWOT River Database (SWORD) (Altenau et al., 2021). For each studied lake, we obtain daily gauge based water volume (V) and surface level (L) and calculate surface area (SA) as

$$SA = dV/dL, \tag{S1}$$

where dV and dL represent daily reservoir volume and level changes, respectively.

To add in SWOT-like errors, we corrupt the surface area (SA_c) using 15% relative errors and corrupt water level (L_c) using 10-cm error, which corresponds to the mission science requirements of SWOT (Biancamaria et al., 2016). It is worth noting that the error budgets required by the mission science are the expected baseline of SWOT performance, and the actual measurement errors may often be smaller than the science requirement errors applied here (Desrochers et al., 2021). Thus, our LakeFlow assessment using the science requirement error budgets may result in a conservative accuracy. Following this step, we calculate corrupted volume (V_c) using

$$V_c = SA_c \times L_c. \tag{S2}$$

For the synthetic river dataset, we rely on the GSW dataset, *in situ* discharge, and SWORD attributes. To produce river width (W), we develop a width-discharge power law relationship (Leopold and Maddock, 1953) by pairing minimum width-discharge and maximum width-discharge to develop a linear regression in log-log space. The minimum and maximum discharges are retrieved from the gauge records in our five-year testing period for each lake, whereas the minimum and maximum widths are manually approximated from the GSW water

occurrence map assuming the latter can represent river width variability within the testing period. To reduce the possible bias and error in manual approximation, the minimum and maximum widths are further calibrated by the difference between their mean and the channel width recorded in the corresponding SWORD reach. Using this power law relationship, we invert *in situ* discharge to produce a synthetic river width and then we corrupt river width (W_c) with 10.6% relative errors (Biancamaria et al., 2016). To produce cross-sectional area changes (dA) we assume a simple trapezoidal shape (Tuozzolo et al., 2019) and calculate it as follows:

$$dA = 0.5(W + W_{min}) \times (L - L_{min}), \quad \text{(S3)}$$

where the subscript min denotes the minimum measurement of a parameter. We corrupt cross-sectional area (dA_c) using eq. S3, but substituting in the corrupted width (W_c) and corrupted level (L_c) values. We calculate A_0 using eq. 1, where we assume a Manning's n value based on local geology and channel geomorphology (Manning's n values ranging from 0.030 to 0.035 $s/m^{1/3}$) (Brinkerhoff et al., 2020; Chow, 1959; Durand et al., 2014), a slope corresponding to the reach-specific slope value from SWORD, and use the median observed value for the remaining variables. Following this, we consider A_0 and n to be constants in time and combine them with the rest of our uncorrupted estimates values in eq. 1 to solve for time varying slope (S). We then corrupt slope (S_c) using the recommended 1.7 cm/km error (Biancamaria et al., 2016). We limit both the synthetic dataset to one observation each week as SWOT will not provide daily observations. One observation each week is approximately equal to the number of SWOT overpasses for each of our lakes during SWOT's 21-day orbit.

Text S2: Ancillary data: Lateral inflow data and lake surface evaporation

For the non-SWOT observed tributary inflow discharge and evaporation, we use modeled daily discharge data from the Global Reach-Level A Priori Discharge Estimates for SWOT

(GRADES) (Lin et al., 2019) and monthly lake surface evaporation estimates from the Global Lake Evaporation Volume (GLEV) dataset (Zhao et al., 2022). GRADES contains daily discharge estimates for 2.94 million river reaches globally from 1979-2014. The discharge is based on runoff simulation from the Variable Infiltration Capacity (VIC) land surface model (Liang et al., 1994), using meteorological forcing that merges gauge-, reanalysis-, and satellite-based data (Beck et al., 2019). The simulated runoff is then routed by the Routing Application for Parallel computation of Discharge (RAPID) model (David et al., 2011) to give discharge for the ~3 million reaches, which offer adequate hydrographic details to represent lateral tributaries to our studied reservoirs. To be consistent with the evaporation dataset, we calculate mean monthly discharge estimates for each reach in GRADES and use these values to estimate a total mean monthly lateral discharge for each reservoir which is denoted as Q_l in eq. 2.

GLEV contains monthly lake evaporation estimates from 1.42 million lakes. GLEV is developed by pairing monthly lake surface area measurements from the Landsat derived GSW dataset (Zhao and Gao, 2018) with monthly meteorological data from several sources (Abatzoglou et al., 2018; Rodell et al., 2004; Xia et al., 2018). Where these meteorological data overlap, the mean monthly value is used to reduce uncertainty. The evaporation rate is then modeled using the Penman Equation (Penman, 1948) with consideration of lake heat storage (Zhao and Gao, 2019). Monthly evaporation loss from each reservoir is calculated as the modeled evaporation rate multiplied by the GSW derived lake surface area and is denoted by E in eq. 2.

Table S1. Error metrics used in this study.

Nash-Sutcliffe efficiency	$NSE = 1 - \sum_{i=1}^{N}(Q_i - \hat{Q}_i)^2 / \sum_{i=1}^{N}(Q_i - \bar{Q})^2$	Where N is the number of timesteps, Q_i is *in situ* discharge at time *i*, $\bar{Q}$ is mean *in situ* discharge, and $\hat{Q}_i$ is LakeFlow estimated discharge at time *i*.
Relative bias	$rBias = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat{Q}_i - Q_i}{Q_i}\right)$	
Normalized root-mean-square error	$NRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{Q_i - \hat{Q}_i}{\bar{Q}}\right)^2}$	
Mean absolute error	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert X_i - \hat{X}_i\rvert$	Where N is the number of observations, X_i is an *in situ* or synthetic value at time *i* and $\hat{X}_i$ is LakeFlow or SoS estimated value at time *i*.

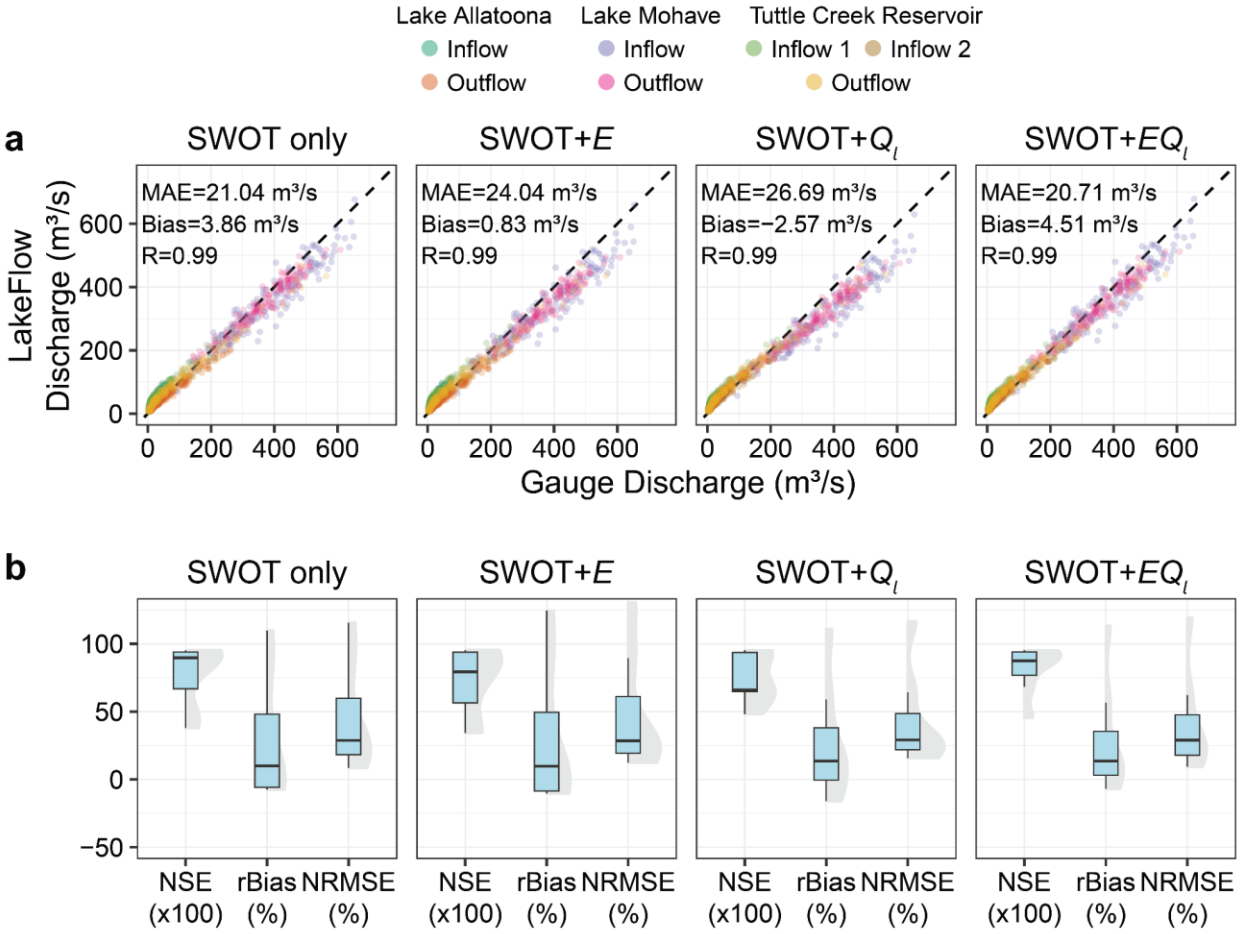

Figure S1. LakeFlow discharge performance for four combinations of input data. Synthetic SWOT data with various ancillary data included from no ancillary data ("SWOT only"), only evaporation data ("SWOT+E"), only lateral inflow data ("SWOT+Q_l"), and both evaporation and lateral inflow data ("SWOT+EQ_l"). (a) Scatterplots of same-day gauge discharge vs. LakeFlow estimated discharge across all reaches. (b) Boxplots and half violin plots of LakeFlow discharge performance metrics across all reaches: NSE (scaled by 100), rBias (%), and NRMSE (%).

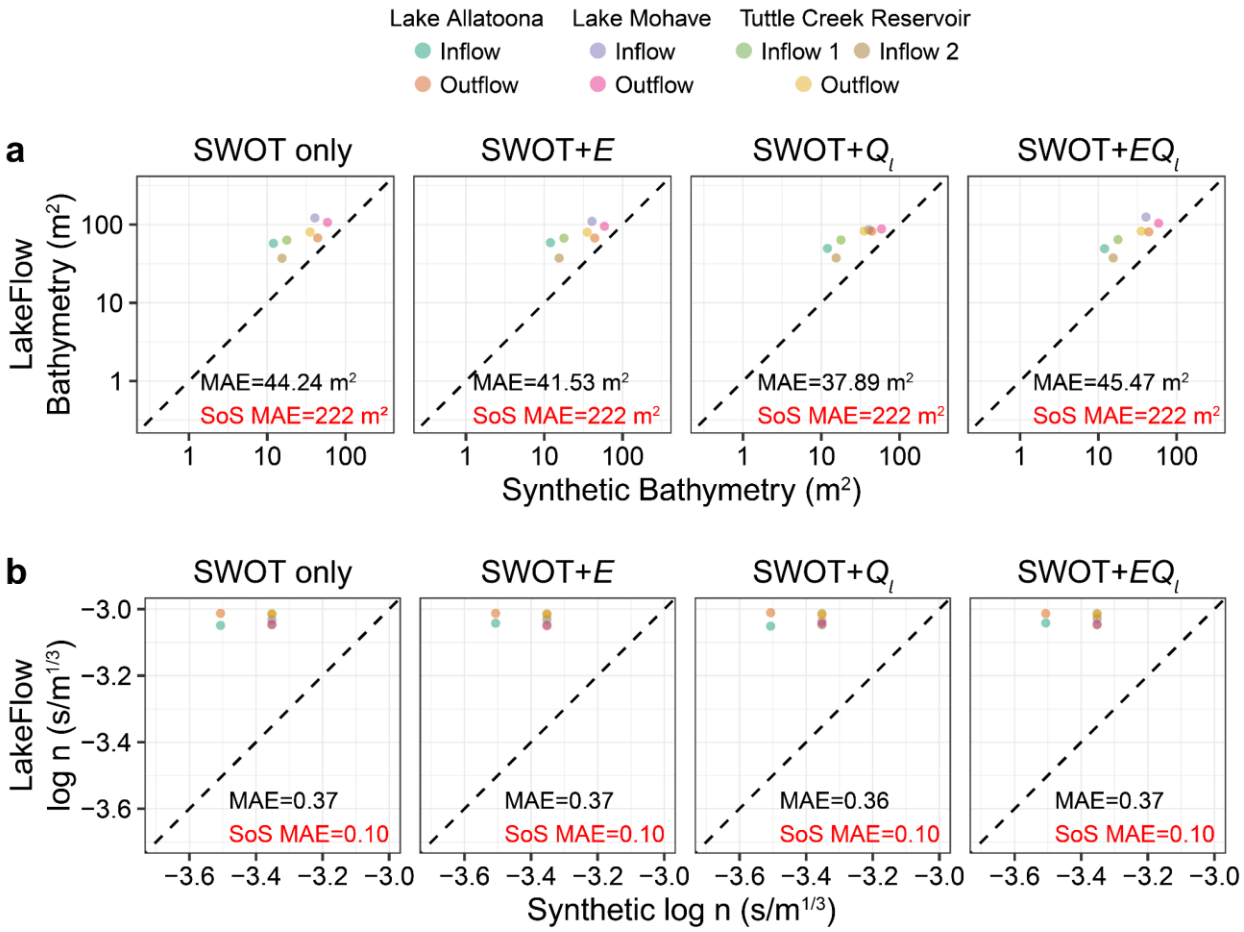

Figure S2. The LakeFlow algorithm's performance of unknown variables for four combinations of input data. Synthetic SWOT data with various ancillary data included from no ancillary data ("SWOT only"), only evaporation data ("SWOT+E"), only lateral inflow data ("SWOT+Q_l"), and both evaporation and lateral inflow data ("SWOT+EQ_l"). MAE for LakeFlow (black) and SoS priors (red) shown. (a) Scatterplots of true bathymetry vs. LakeFlow estimated bathymetry across all reaches. (b) Scatterplots of log true Manning's n vs. log LakeFlow estimated Manning's n across all reaches.

References

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Sci. Data 5, 170191. https://doi.org/10.1038/sdata.2017.191

Altenau, E.H., Pavelsky, T.M., Durand, M.T., Yang, X., Frasson, R.P. de M., Bendezu, L., 2021. The Surface Water and Ocean Topography (SWOT) Mission River Database (SWORD): A Global River Network for Satellite Data Products. Water Resour. Res. 57, e2021WR030054. https://doi.org/10.1029/2021WR030054

Beck, H.E., Wood, E.F., Pan, M., Fisher, C.K., Miralles, D.G., van Dijk, A.I.J.M., McVicar, T.R., Adler, R.F., 2019. MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. Bull. Am. Meteorol. Soc. 100, 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1

Biancamaria, S., Lettenmaier, D.P., Pavelsky, T.M., 2016. The SWOT Mission and Its Capabilities for Land Hydrology. Surv. Geophys. 37, 307–337. https://doi.org/10.1007/s10712-015-9346-y

Brinkerhoff, C.B., Gleason, C.J., Feng, D., Lin, P., 2020. Constraining Remote River Discharge Estimation Using Reach-Scale Geomorphology. Water Resour. Res. 56, e2020WR027949. https://doi.org/10.1029/2020WR027949

Chow, V.T., 1959. Open-channel hydraulics. McGraw-Hill, New York.

David, C.H., Maidment, D.R., Niu, G.-Y., Yang, Z.-L., Habets, F., Eijkhout, V., 2011. River Network Routing on the NHDPlus Dataset. J. Hydrometeorol. 12, 913–934. https://doi.org/10.1175/2011JHM1345.1

Desrochers, N.M., Trudel, M., Biancamaria, S., Siles, G., Desroches, D., Carbonne, D., Leconte, R., 2021. Effects of Aquatic and Emergent Riparian Vegetation on SWOT Mission Capability in Detecting Surface Water Extent. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 12467–12478. https://doi.org/10.1109/JSTARS.2021.3128133

Durand, M., Neal, J., Rodríguez, E., Andreadis, K.M., Smith, L.C., Yoon, Y., 2014. Estimating reach-averaged discharge for the River Severn from measurements of river water surface elevation and slope. J. Hydrol. 511, 92–104. https://doi.org/10.1016/j.jhydrol.2013.12.050

Leopold, L.B., Maddock, T., 1953. The Hydraulic Geometry of Stream Channels and Some Physiographic Implications. U.S. Government Printing Office.

Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. J. Geophys. Res. Atmospheres 99, 14415–14428. https://doi.org/10.1029/94JD00483

Lin, P., Pan, M., Beck, H.E., Yang, Y., Yamazaki, D., Frasson, R., David, C.H., Durand, M., Pavelsky, T.M., Allen, G.H., Gleason, C.J., Wood, E.F., 2019. Global Reconstruction of Naturalized River Flows at 2.94 Million Reaches. Water Resour. Res. 55, 6499–6516. https://doi.org/10.1029/2019WR025287

Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418–422. https://doi.org/10.1038/nature20584

Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass. Proc. R. Soc. Lond. Ser. Math. Phys. Sci. 193, 120–145. https://doi.org/10.1098/rspa.1948.0037

Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J.K., Walker, J.P., Lohmann, D., Toll, D., 2004. The Global Land Data Assimilation System. Bull. Am. Meteorol. Soc. 85, 381–394. https://doi.org/10.1175/BAMS-85-3-381

Tuozzolo, S., Langhorst, T., de Moraes Frasson, R.P., Pavelsky, T., Durand, M., Schobelock, J.J., 2019. The impact of reach averaging Manning's equation for an in-situ dataset of water surface elevation, width, and slope. J. Hydrol. 578, 123866. https://doi.org/10.1016/j.jhydrol.2019.06.038

U.S. Geological Survey, 2022. USGS Current Water Data for the Nation [WWW Document]. US Geol. Surv. URL https://waterdata.usgs.gov/nwis/rt (accessed 5.1.21).

Xia, X., Zhang, S., Li, S., Zhang, Liwei, Wang, G., Zhang, Ling, Wang, J., Li, Z., 2018. The cycle of nitrogen in river systems: sources, transformation, and flux. Environ. Sci. Process. Impacts 20, 863–891. https://doi.org/10.1039/C8EM00042E

Zhao, G., Gao, H., 2018. Automatic Correction of Contaminated Images for Assessment of Reservoir Surface Area Dynamics. Geophys. Res. Lett. 45, 6092–6099. https://doi.org/10.1029/2018GL078343

Zhao, G., Li, Y., Zhou, L., Gao, H., 2022. Evaporative water loss of 1.42 million global lakes. Nat. Commun. 13, 3686. https://doi.org/10.1038/s41467-022-31125-6