




# AHGestimation: An R package for computing robust, mass preserving hydraulic geometries and rating curves

J Michael Johnson <sup>1</sup>, Shahab Afshari <sup>2</sup>, and Arash Modaresi Rad <sup>1</sup>

<sup>1</sup> Lynker, NOAA/NWS Office of Water Prediction <sup>2</sup> University of Massachusetts Amherst

DOI: [10.21105/joss.06145](https://doi.org/10.21105/joss.06145)

## Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Chris Vernon](#)  

## Reviewers:

- [@mabesa](#)
- [@mengqi-z](#)

Submitted: 30 October 2023

Published: 13 April 2024

## License

Authors of papers retain copyright  
and release the work under a  
Creative Commons Attribution 4.0  
International License ([CC BY 4.0](#)).

## Summary

### Background

The behavior of a river channel at a given cross-section is often described using power law equations to relate top width (TW), mean depth (Y), and mean velocity (V) to discharge (Q). These equations collectively define the “at a station hydraulic geometry” (AHG), with coefficients (a, c, k) and exponents (b, f, m) for the Q-TW, Q-Y, and Q-V relationships ([Leopold & Maddock, 1953](#)). While each relation can be studied independently, together they represent the full hydraulic system.

$$TW = a \cdot Q^b \quad (1)$$

$$Y = c \cdot Q^f \quad (2)$$

$$V = k \cdot Q^m \quad (3)$$

$$Q = TW \cdot Y \cdot V \quad (4)$$

When using the traditional AHG fitting methods, power laws are estimated using Ordinary Least Squares Regression (OLS) on the logarithmic transformation of the variables, where the exponential of the intercept provides the power law coefficient, and the slope of the linear model provides the exponent. Several studies have highlighted the potential problem of using this approach given that when the log transformed values are back-transformed, the estimates are effectively medians, and not means of the estimates, resulting in a general low bias. AHG relations apply to within-bank flows at a specific location and assume the channel characteristics do not significantly change with discharge. Under these assumptions, two continuity conditions ensure mass conservation (equation 5-6).

$$b + f + m = 1 \quad (5)$$

$$a \times c \times k = 1 \quad (6)$$

Real-world data often fails to meet these conditions precisely, so an allowance is typically permitted. Violating these conditions leads to mass imbalances in computations, causing input streamflows to gain or lose mass in its translation to other states. Given the noisy nature of hydraulic data, tools are needed to prep input data and fit equations while balancing accuracy and mass preservation for hydrologic and hydraulic modeling.

## Statement of need

Hydrologic models that simulate streamflow are critical for forecasting water availability, drought, and flood inundation. A key aspect of these models is estimating the size and shape of channels, often achieved through hydraulic geometry relationships.

While extensively studied at local scales, these relationships remain unquantified for the majority of stream reaches globally, including in the United States. As a result, large-scale models often rely on incomplete approximations, leading to less accurate streamflow estimates (Fang et al., 2024; Heldmyer et al., 2022; Johnson et al., 2023) and flood forecasts (Johnson et al., 2019; Maidment, 2017; NOAA, 2023; Zheng et al., 2018). For instance, the National Oceanic and Atmospheric Administration National Water Model (Cosgrove et al., 2024) uses trapezoidal geometries (Gochis et al., 2020) that are in part derived from hydraulic geometry relationships and drainage area assumptions found in Bieger et al. (2015); Bieger et al. (2016); Blackburn-Lynch et al. (2017). Although these approximations are recognized as overly simplistic, advancing them requires integrating diverse observation systems and refining them into parameterized, mass-conserving relationships.

Several efforts have aimed to address this challenge in the United States, primarily relying on traditional OLS fitting methods and data preprocessing (Afshari et al., 2017; Afshari, 2019; Enzminger, 2023). However, these efforts are limited by the lack of shared software and source data, hindering the evolution, and interoperability, of their products.

## Software

AHG estimation has traditionally been conducted on a site-specific basis, with location-specific knowledge guiding data selection and output validation. However, as interest in large-scale model applications grows (Archfield et al., 2015), the importance of estimating these relationships from diverse and often noisy datasets becomes increasingly evident.

Towards this, AHGestimation is an R package (R Core Team, 2023) providing three capabilities:

1. **AHG Estimation** (`ahg_estimate`): This tool provides robust estimation techniques to estimate single and full system AHG relations that enforce flow continuity and minimizes total system error. This is accomplished by introducing a hybrid approach that supplements the traditional OLS approach, with a Nonlinear Least Square (NLS) regression, Evolutionary algorithm (NSGA-2) (Mersmann, 2020), and ensemble modeling approach.
2. **Outlier Filtering Methods** (`date_filter`, `qva_filter`, `mad_filter`, `nls_filter`): These methods allow users to filter outliers based on various criteria, including time, mass conservation, and statistical detection.
3. **Formalization of derived AHG Concepts** (`cross_section`, `compute_hydraulic_params`, `compute_n`): These functions formalize concepts from existing literature, such as the relationship between AHG coefficients and exponents and cross-section hydraulics and geometry. Formalize, in code, many of the concepts derived in (Dingman & Afshari, 2018) that relate AHG coefficients and exponents to cross-section hydraulics and geometry.

The package documentation includes several examples on the theory, design, and application of these tools.

The first stable version of AHGestimation was made available in 2019 and was applied to an aggregated dataset of USGS manual field measurements. Since then, it has been actively developed to better understand and quantify these fundamental relationships in the face of noisy, large, and disparate data sources. Applications of the software have been used to (1) demonstrate how improved flood forecasts could be delivered from the NOAA/NWS National Water Model (Johnson, Narock, et al., 2022), (2) help the NOAA/NWS Office of Water Prediction develop continental scale channel size and shape estimates to improve

flood prediction and hydraulic routing, and (3) bolster the co-agency USGS/NOAA National Hydrologic Geospatial Fabric and Next Generation Water Resource Modeling Framework Hydrofabric efforts ([Blodgett et al., 2021, 2023](#); [Bock et al., 2022](#); [Johnson, 2022](#)).

## Example of use

AHGESTimation is available on [GitHub](#) and can be installed as follows:

```
#install.packages(remotes)
remotes::install.packages("mikejohnson51/AHGESTimation")
```

This example illustrates how the package can be utilized to:

1. **Remove Data Outliers:** Outliers are filtered based on time criteria and an NLS envelope.
2. **Fit AHG Parameters:** AHG parameters are estimated using the hybrid modeling approach.
3. **Estimate and Plot Cross-Section Shape:** The shape of the associated cross-section is estimated and plotted with an area-depth relation.

The script to generate the plot found in [Figure 1](#) can be found [here](#), and the nwis data object is exported with the package to provides field measurements taken at [USGS site 01096500 on the Nashua River at East Pepperell in Massachusetts](#).

```
nwis
```

```
#> siteID      date      Q      Y      V      TW
#> 01096500 1984-11-14  9.7409954 0.5276645 0.652272 28.34640
#> 01096500 1985-01-04 11.8930757 0.6263473 0.682752 27.73680
#> 01096500 1985-05-06 10.8170356 0.5952226 0.563880 32.30880
```

```
nrow(nwis)
#> 245
```

```
# Keep only those observation made in the most recent 10 years,
# and that fall within the .5 nls envelope
(data = nwis |>
  dplyr::rename(Q = Q_cms, Y = Y_m, V = V_ms, TW = TW_m) |>
  date_filter(10, keep_max = TRUE) |>
  nls_filter(allowance = 0.5) )
```

```
# data reduced to 80 observations based on filters
nrow(data)
#> 85
```

The reduced clean data can then be used to fit an AHG relation and compute a set of hydraulic parameters:

```
# Fit AHG relations
ahg_fit = ahg_estimate(data)
t(ahg_fit[1,])
```

```
#> V_method "nls"
#> TW_method "nls"
#> Y_method "nls"
#> c1       "1.006"
#> c2       "1.001"
#> viable    "TRUE"
#> tot_nrmse "0.3234122"
```

```
#> V_nrmse "0.1337535"
#> TW_nrmse "0.1009869"
#> Y_nrmse "0.0886718"
#> V_coef "0.2905399"
#> TW_coef "18.23401"
#> Y_coef "0.1898499"
#> V_exp "0.3101059"
#> TW_exp "0.1756746"
#> Y_exp "0.5155772"
#> condition "bestValid"

# Use the AHG relations to compute hydraulic parameters
shape = compute_hydraulic_params(ahg_fit[1,])

#> r p d R bd fd md
#> 2.93 0.60 5.70 1.34 0.18 0.51 0.31
```

Finally, the max width and depth, paired with the derived *r* coefficient can be used to generate a cross section:

```
# Use the max width, max depth, and derived `r` to generate a cross section
# x: is the relative distance from the left bank
# Y: is the associated depth
# A: is the area associated with depth Y

cs = cross_section(r = shape$r,
                  TW = max(data$TW),
                  Ymax = max(data$Y))

#> ind x Y A
#> 1 0.000000 3.5525892702 1.337069e+02
#> 2 1.040816 3.1514749668 1.136999e+02
#> 3 2.081633 2.7814289966 9.601412e+01
```

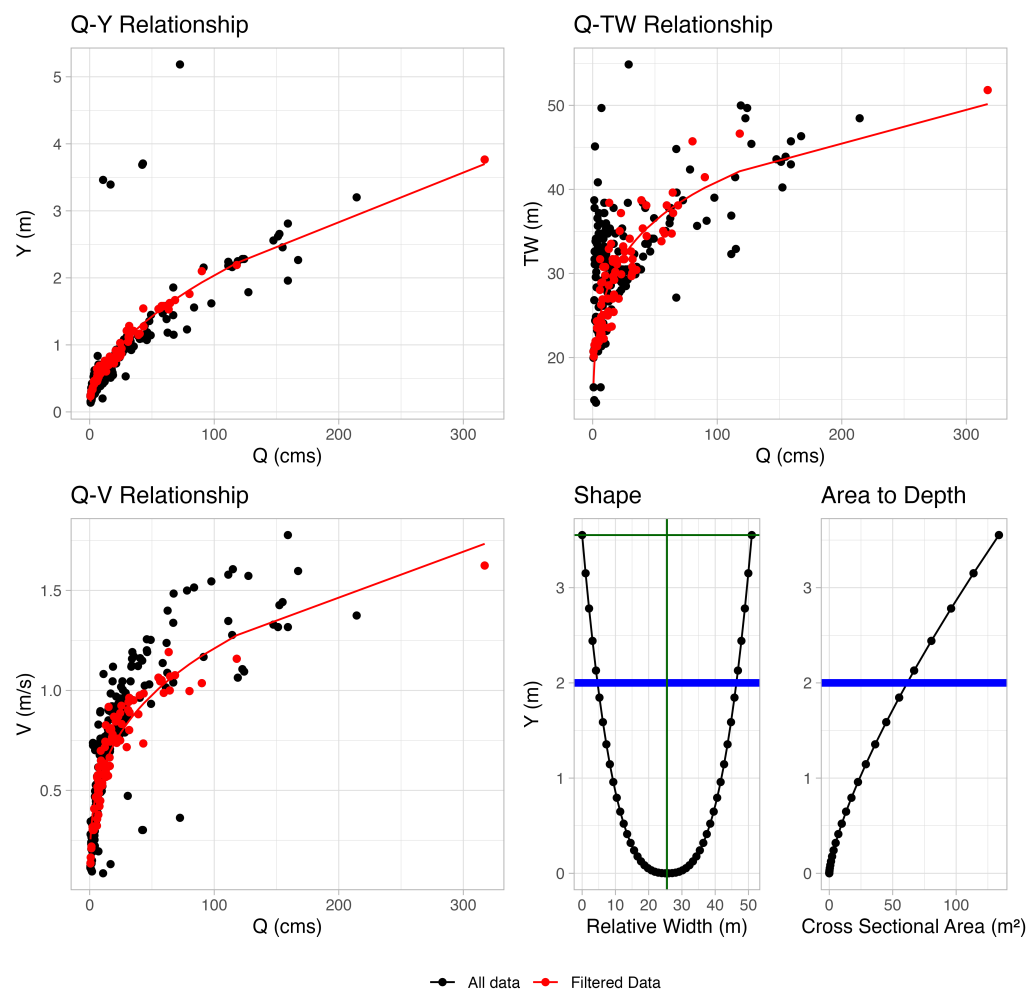


Figure 1: Faceted image with multiple views of the channel estimate.

As a proof of concept, this approach was applied to the synthetic rating curves generated by NOAA/NWS OWP's inundation mapping software (NOAA, 2023). The package allowed the size of their rating curve database to be reduced by 99.68% while maintaining average accuracy within 0.4% NRMSE of the source data (Johnson, Coll, et al., 2022). This reduction to a consistent AHG formulation not only enhances interoperability with efforts such as those by Enzminger (2023) or Afshari (2019), but also lays the groundwork for providing essential training data for advanced prediction methods that seek to estimate the shape on non-measured river segments.

## Acknowledgements

The development of this package began following the 2017 NOAA/NWS Office of Water Prediction Summer Institute (Johnson et al., 2017) where discussions highlighted the potential influence of channel shape representation on the performance of the National Water Model.

The algorithm and implementation began as a graduate school project between friends at UC Santa Barbara and UMass Amherst and has since evolved to provide an open source utility for robust large scale data synthesis and evaluation. Funding from the National Science Foundation (Grants 1937099, 2033607) provided time to draft (Johnson, Coll, et al., 2022) and apply an

early version of this software to the Continental Flood Inundation Mapping synthetic rating curve dataset (Liu & Maidment, 2020). Funding from the NOAA/NWS OWP supported the addition of data filtering and hydraulic geometry estimation, improved documentation, and code hardening. We are grateful to all involved.

## Disclaimer

The views expressed in this article do not necessarily represent the views of NOAA, the USGS, or the United States.

## References

- Afshari, S. (2019). *USGS Table AHG Parameters And Supplementary Data* (Version v1.0) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.2558565>
- Afshari, S., Fekete, B. M., Dingman, S. L., Devineni, N., Bjerklie, D. M., & Khanbilvardi, R. M. (2017). Statistical filtering of river survey and streamflow data for improving At-A-Station hydraulic geometry relations. *Journal of Hydrology*, 547, 443–454. <https://doi.org/10.1016/j.jhydrol.2017.01.038>
- Archfield, S. A., Clark, M., Arheimer, B., Hay, L. E., McMillan, H., Kiang, J. E., Seibert, J., Hakala, K., Bock, A., Wagener, T., & others. (2015). Accelerating advances in continental domain hydrologic modeling. *Water Resources Research*, 51(12), 10078–10091. <https://doi.org/10.1002/2015WR017498>
- Bieger, K., Rathjens, H., Allen, P. M., & Arnold, J. G. (2015). Development and evaluation of bankfull hydraulic geometry relationships for the physiographic regions of the United States. *JAWRA Journal of the American Water Resources Association*, 51(3), 842–858. <https://doi.org/10.1111/jawr.12282>
- Bieger, K., Rathjens, H., Arnold, J. G., Chaubey, I., & Allen, P. M. (2016). Development and comparison of multiple regression models to predict bankfull channel dimensions for use in hydrologic models. *JAWRA Journal of the American Water Resources Association*, 52(6), 1385–1400. <https://doi.org/10.1111/1752-1688.12460>
- Blackburn-Lynch, W., Agouridis, C. T., & Barton, C. D. (2017). Development of regional curves for hydrologic landscape regions (HLR) in the contiguous United States. *JAWRA Journal of the American Water Resources Association*, 53(4), 903–928. <https://doi.org/10.1111/1752-1688.12540>
- Blodgett, D., Johnson, J., & Bock, A. (2023). Generating a reference flow network with improved connectivity to support durable data integration and reproducibility in the coterminous US. *Environmental Modelling & Software*, 165, 105726. <https://doi.org/10.1016/j.envsoft.2023.105726>
- Blodgett, D., Johnson, J., Sondheim, M., Wiecek, M., & Frazier, N. (2021). Mainstems: A logical data model implementing mainstem and drainage basin feature types based on WaterML2 Part 3: HY Features concepts. *Environmental Modelling & Software*, 135, 104927. <https://doi.org/10.1016/j.envsoft.2020.104927>
- Bock, A. R., Blodgett, D. L., Johnson, J., Santiago, M., & Wiecek, M. E. (2022). *National Hydrologic Geospatial Fabric Reference and Derived Hydrofabrics*. U.S. Geological Survey software release. <https://www.sciencebase.gov/catalog/item/60be0e53d34e86b93891012b>
- Cosgrove, B., Gochis, D., Flowers, T., Dugger, A., Ogden, F., Graziano, T., Clark, E., Cabell, R., Casiday, N., Cui, Z., & others. (2024). NOAA's National Water Model: Advancing operational hydrology through continental-scale modeling. *JAWRA Journal of the American Water Resources Association*. <https://doi.org/10.1111/1752-1688.13184>

- Dingman, L. S., & Afshari, S. (2018). Field verification of analytical at-a-station hydraulic-geometry relations. *Journal of Hydrology*, 564, 859–872. <https://doi.org/10.1016/j.jhydrol.2018.07.020>
- Enzminger, M., Thomas L. (2023). *HyG: A hydraulic geometry dataset derived from historical stream gage measurements across the conterminous United States* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7868764>
- Fang, S., Johnson, J., Yeghiazarian, L., & Sankarasubramanian, A. (2024). Improved National-Scale Above-Normal Flow Prediction for Gauged and Ungauged Basins Using a Spatio-Temporal Hierarchical Model. *Water Resources Research*, 60(1). <https://doi.org/10.1029/2023WR034557>
- Gochis, D. J., M. Barlage, R., Cabell, M. C., Dugger, A., K. FitzGerald, J. M., M. McAllister, RafieeiNasab, A., L. Read, K. S., Yates, D., & Zhang, Y. (2020). The WRF-Hydro® modeling system technical description, (Version 5.2.0). *NCAR Technical Note*. <https://ral.ucar.edu/sites/default/files/public/projects/wrf-hydro/technical-description-user-guide/wrf-hydrov5.2technicaldescription.pdf>
- Heldmyer, A., Livneh, B., McCreight, J., Read, L., Kasprzyk, J., & Minear, T. (2022). Evaluation of a new observationally based channel parameterization for the National Water Model. *Hydrology and Earth System Sciences*, 26(23), 6121–6136. <https://doi.org/10.5194/hess-26-6121-2022>
- Johnson, J. (2022). *National hydrologic geospatial fabric (hydrofabric) for the next generation (NextGen) hydrologic modeling framework*. HydroShare. <https://www.hydroshare.org/resource/129787b468aa4d55ace7b124ed27dbde/>
- Johnson, J., Coll, J., Clarke, A., K. C., S., S., Saksena, & Yeghiazarian, L. (2022). Determining Feature Based Hydraulic Geometry and Rating Curves using a Physically Based, Computationally Efficient Framework. *Preprints*. <https://doi.org/10.20944/preprints202212.0390.v1>
- Johnson, J., Coll, J., Maidment, D., Cohen, S., Nelson, J., Ogden, F., Praskievicz, S., & Clark, E. (2017). National Water Center Innovators Program Summer Institute Report 2017. *Consortium of Universities for the Advancement of Hydrologic Science, Inc., Technical Report*, 14. <https://doi.org/10.4211/technical.20171009>
- Johnson, J., Fang, S., Sankarasubramanian, A., Rad, A. M., Kindl da Cunha, L., Jennings, K. S., Clarke, K. C., Mazrooei, A., & Yeghiazarian, L. (2023). Comprehensive analysis of the NOAA National Water Model: A call for heterogeneous formulations and diagnostic model selection. *Journal of Geophysical Research: Atmospheres*, 128(24). <https://doi.org/10.1029/2023JD038534>
- Johnson, J., Munasinghe, D., Eyelade, D., & Cohen, S. (2019). An integrated evaluation of the National Water Model (NWM)–Height Above Nearest Drainage (HAND) flood mapping methodology. *Natural Hazards and Earth System Sciences*, 19(11), 2405–2420. <https://doi.org/10.5194/nhess-19-2405-2019>
- Johnson, J., Narock, T., Singh-Mohudpur, J., Fils, D., Clarke, K., Saksena, S., Shepherd, A., Arumugam, S., & Yeghiazarian, L. (2022). Knowledge graphs to support real-time flood impact evaluation. *AI Magazine*, 43(1), 40–45. <https://doi.org/10.1002/aaai.12035>
- Leopold, L. B., & Maddock, T. (1953). *The hydraulic geometry of stream channels and some physiographic implications* (Vol. 252). US Government Printing Office. <https://doi.org/10.3133/pp252>
- Liu, T., Yan Y., & Maidment, D. R. (2020). Height Above Nearest Drainage (HAND) and Hydraulic Property Table for CONUS - Version 0.2.1. (20200601). In *Oak Ridge Leadership Computing Facility*. <https://cfim.ornl.gov/data/>



- Maidment, D. R. (2017). Conceptual Framework for the National Flood Interoperability Experiment. *JAWRA Journal of the American Water Resources Association*, 53(2), 245–257. <https://doi.org/10.1111/1752-1688.12474>
- Mersmann, O. (2020). *mco: Multiple Criteria Optimization Algorithms and Related Functions*. <https://CRAN.R-project.org/package=mco>
- NOAA, O. of W. P. (2023). Inundation Mapping. In *GitHub*. <https://github.com/NOAA-OWP/inundation-mapping>
- R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Zheng, X., Tarboton, D. G., Maidment, D. R., Liu, Y. Y., & Passalacqua, P. (2018). River channel geometry and rating curve estimation using height above the nearest drainage. *JAWRA Journal of the American Water Resources Association*, 54(4), 785–806. <https://doi.org/10.1111/1752-1688.12661>