

# Development and comprehensive validation of SWOT river discharge algorithms from AirSWOT, simulator, and field measurements

---

A proposal submitted in response to NASA NRA A.10 Surface Water and Ocean Topography (SWOT) Science Team

PI: Michael Durand

School of Earth Science, The Ohio State University

Co-I: Colin J. Gleason

Department of Geography, UCLA

Co-I: David Bjerklie

New England Water Science Center, USGS

## Summary

The ability to provide accurate river discharge estimates across river networks is key to the transformational power of the SWOT mission for terrestrial hydrology. There has been considerable progress on SWOT discharge algorithms during the past several years, and the SWOT community is currently evaluating four such algorithms. While the project team has led an extensive effort to validate these algorithms for various river types using hydraulic models, this previous effort falls short of the “comprehensive validation” of discharge algorithms for a “range of different river types” called for in section 2.1.2 of the solicitation. We here propose to 1) perform comprehensive validation of the four existing algorithms using i) AirSWOT measurements, ii) field data, iii) SWOT instrument simulations, and iv) hydraulic model outputs, 2) develop a novel synergistic algorithm and provide an open source user platform to deliver discharge products, and 3) develop these algorithms for deployment at river basin scales. We will leverage significant ongoing work, existing datasets, and data available via collaboration with other groups to make this project feasible. First, we will form a core validation dataset consisting of the four dataset types described above. Field data would be utilized from Science Team proposals (if funded) by Dr. Laurence Smith (UCLA), Dr. Tamara Wood (USGS), and Dr. John Fulton (USGS). An additional field-based dataset would be collected under this proposal. AirSWOT data and hydraulic simulations from Dr. Wood’s proposal would also be utilized, along with all other existing AirSWOT and instrument simulator cases available to the PIs. These data ultimately yield a total of 102 separate discharge estimation cases to be examined for all four available algorithms. Second, we will develop a synergistic algorithm that combines portions of each candidate algorithm. This combined method should allow for decreased reliance on a priori flow data and achieve more accurate discharge retrievals. We will develop an open-source discharge estimation tool with graphic user interface allowing ST community members to produce discharge estimates from any of the four proposed algorithms, their ensemble mean, and the proposed synergistic algorithm. Third, the feasibility of global application of the synergistic algorithm will be explored. Specifically, the synergistic algorithm will be translated for High-Performance Computing (HPC) deployment and will be benchmarked for continental-to-global scale application and for estimation across river networks. It is our intent that the discharge algorithms developed and rigorously tested here will be fully operational and ready for SWOT launch upon completion of this award, thus fulfilling a major goal of the mission.

## 1 Introduction

The upcoming Surface Water and Ocean Topography (SWOT) mission measurements of river water surface elevation (WSE), top width, and free surface slope will allow estimation of river discharge for all rivers greater than 100 m in width, with a goal of estimating discharge for rivers greater than 50 m in width (SWOT science requirements document v1.14, requirement 1.2). Accurate discharge estimates from these measurements would enable a sea change in global hydrologic studies. The anticipated SWOT data have sparked new discharge estimation methods, which draw from decades of heritage in the fields of hydraulics, remote sensing, and fluvial geomorphology. Studies have indicated that SWOT-like observations can be used to accurately estimate discharge in ungauged areas in many cases.

Three major challenges remain before SWOT discharge algorithms are mature enough for production of scientific data post-SWOT launch. First, a clear guideline as to expected discharge accuracy for each SWOT-observable river must be provided prior to launch, as definitive algorithm performance based on river type cannot yet be predicted. Second, individual algorithm performance is not yet adequate to consistently meet the scientific communities' needs, primarily because the algorithms have not yet been combined and synergized. Third, a computational solution capable of processing the SWOT data globally for discharge calculations has not yet been developed and tested. Addressing each of these challenges, we will do the necessary groundwork to provide reliable, accurate discharge estimates of known quality from SWOT measurements.

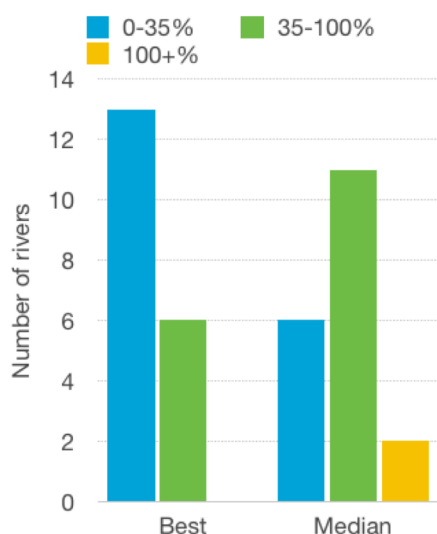
## 2 Objectives

1. Perform comprehensive validation of the existing algorithms using AirSWOT, field, simulator, and hydraulic model data over a range of river types, and assess performance in the context of river characteristics. 2. Develop a synergistic algorithm, which would capitalize on the best features of each approach. 3. Assess algorithm computational resource requirements, and optimize algorithms for deployment across river networks.

## 3 Background

Algorithm studies to date have produced four approaches for estimating river discharge from SWOT observables. The approaches differ in several ways: In terms of flow laws, AMHG is based on hydraulic geometry while the other three approaches use Manning's equation. AMHG and the Garambois algorithm focus on an estimation of discharge that generally relies on no prior discharge information by solving an inverse problem; the MFG relies completely on a priori discharge estimates of mean annual flow, while MetroMan follows the Bayesian approach and utilizes both a priori information and the inverse problem solution. The diversity of these approaches is a great strength of the discharge algorithm development work thus far. Algorithms are at different stages of development; published rRMSE values are generally less than 35% for all algorithms, except for a few types of rivers for which AMHG is not expected to function well.

One of the most telling findings of the recent intercomparison study is shown in Fig. 1. For 13 of the 19 rivers estimated, the best performing algorithm had an rRMSE less



**Fig. 1.** Preliminary algorithm intercomparison results. On the left, the rRMSE of the best performer, for each river. On the right, the rRMSE of the median discharge of the four estimates at each time.

than 35%. This is good news: at least one algorithm is usually performing well. However, we currently do not know why each algorithm performs better or worse in each case. The median across the four discharge estimates does not perform nearly as well as the best performer. This is bad news: we cannot yet predict algorithm performance based on a priori river and hydraulic characteristics. Prior to the upcoming SDT meeting in July 2015, the PIs and the DAWG will develop hypotheses to explain algorithm performance. But how will we test those hypotheses? The next few years will see a huge influx of data from AirSWOT, in situ measurements, simulator, and hydraulic models, developed by the Science Team. These new data will allow analysis of discharge hypotheses; this is the work proposed herein. If this is done, then the capacity will exist to deploy discharge algorithms of known accuracy and precision as SWOT launches and realize its transformational ability to provide discharge data to the global scientific community.

## 4 Technical Approach

The approach proposed herein is to take several critical steps to build upon the work performed during the Virtual Mission and the SDT. First, we will focus on validating river discharge algorithms utilizing new datasets to be made available to the Science Team in the near future. Second, we will implement new algorithm developments, in part informed by the results of the validation. Third, we will adapt algorithms for entire river networks, and for high-performance computing environments.

### 4.1 Validating and Assessing Algorithms

We propose to use four kinds of data to assess algorithm performance; each allows for assessment of different issues with the discharge algorithms. The four algorithms described in section 3.3.3 will be run for each case. Moreover, we will make every effort to include other new algorithmic approaches. For example, the proposal submitted by Dr. Pierre-Olivier Malaterre (IRSTEA) to TOSCA will explore 1-D variational approaches to  $Q$  estimation. If funded, Dr. Malaterre's student (Oubanas Hind) will visit PI Durand's laboratory to share datasets and compare methods, in person. This valuable interaction would be supported by university funds at Ohio State and by Malaterre's TOSCA budget. A total of 102 cases (summed across all data types) on a total 23 rivers are to be evaluated. A "case" is defined as a complete set of data for which discharge algorithms can be run, and performance evaluated, and typically consists of several river reaches, with a full suite of SWOT-like observations. Only one case is assumed for any given year of data, though this is conservative, since most algorithms can be run for each

hydraulic event that produces temporal dynamics. A total of 11 AirSWOT cases, 42 cases using field data, 24 instrument simulator cases, and 25 cases using hydraulic models will be utilized, as described below. Some of these cases represent data already in-hand, while others refer to data contingent upon other Science Team proposals being funded. Additional data will doubtless be collected as Science Team plans change; every effort will be made to include every possible case in algorithm comparison activities. River types include single channel (Sacramento), simple multi-channel (Seiene), braided (Tanana), significant side-channels and wetlands (Willamette), continuous permafrost (Mackenzie), significant floodplain interaction (Po), large (Amazon), small (Olentangy), tidal (Connecticut) and affected by low-head dams (Ohio).

## 4.2 Developing Algorithms

We propose to use ensemble mean and median products per observation, much like the IPCC uses in publishing climate projections. This should give more stable discharge estimates and dampen the effect of outlier retrievals. In the DAWG work to date, ensemble estimates have performed better than single algorithms. The second strategy for fusing algorithms is the development of a synergistic algorithm that combines aspects of each of the previously proposed approaches. A first step is to use either the AMHG or Garambois approach to provide a first estimate of flow then ingested by MetroMan or the MFG. By using AMHG or Garambois as first guess, reliance on a priori model data in MetroMan is eliminated, while still allowing for robust sensitivity analysis to measurement error – something not included in MFG or AMHG. We envision incorporating constraints on the synergistic algorithm based on the previous approaches, essentially running portions of each algorithm in parallel and flagging situations where discharge outputs violate either AHG or Manning's equation. These two approaches for algorithm synergy will form the core of our proposed SWOT discharge product.

## 5 Evaluation of Research Questions and Hypotheses

**Hypothesis 1a.** We hypothesize that at least one of the previously developed algorithms will be able to provide a hydrograph with less than 35% rRMSE for each river. **Hypothesis 1b.** We hypothesize that discharge anomaly will be predicted to within 25% rRMSE for each river. **Hypothesis 1c.** We hypothesize that performance will be slightly degraded for the cases using real data, as opposed to synthetic data, but that simulated layover and geolocation error from the instrument simulator will not lead to dramatic departure in discharge performance as compared to real cases. **Hypothesis 2.** We hypothesize that the synergistic algorithms (both ensemble means and the super-algorithm) will be more stable than individual algorithms, and will not have residuals greater than 30% at any point in the estimation time series for our test datasets. **Hypothesis 3.** We hypothesize that the Science Team component of the discharge algorithms will be challenging computationally, but doable using HPC. We hypothesize that estimation across the river network will help to eliminate bias in the Q estimates.