



# A Regression Model to Predict Total Suspended Solid Concentrations in the Severn River, MD

Midshipman 1/C Alexander V. Schuerch, USN, Class of 2022; U.S. Naval Academy, Oceanography Department, Annapolis, MD, USA

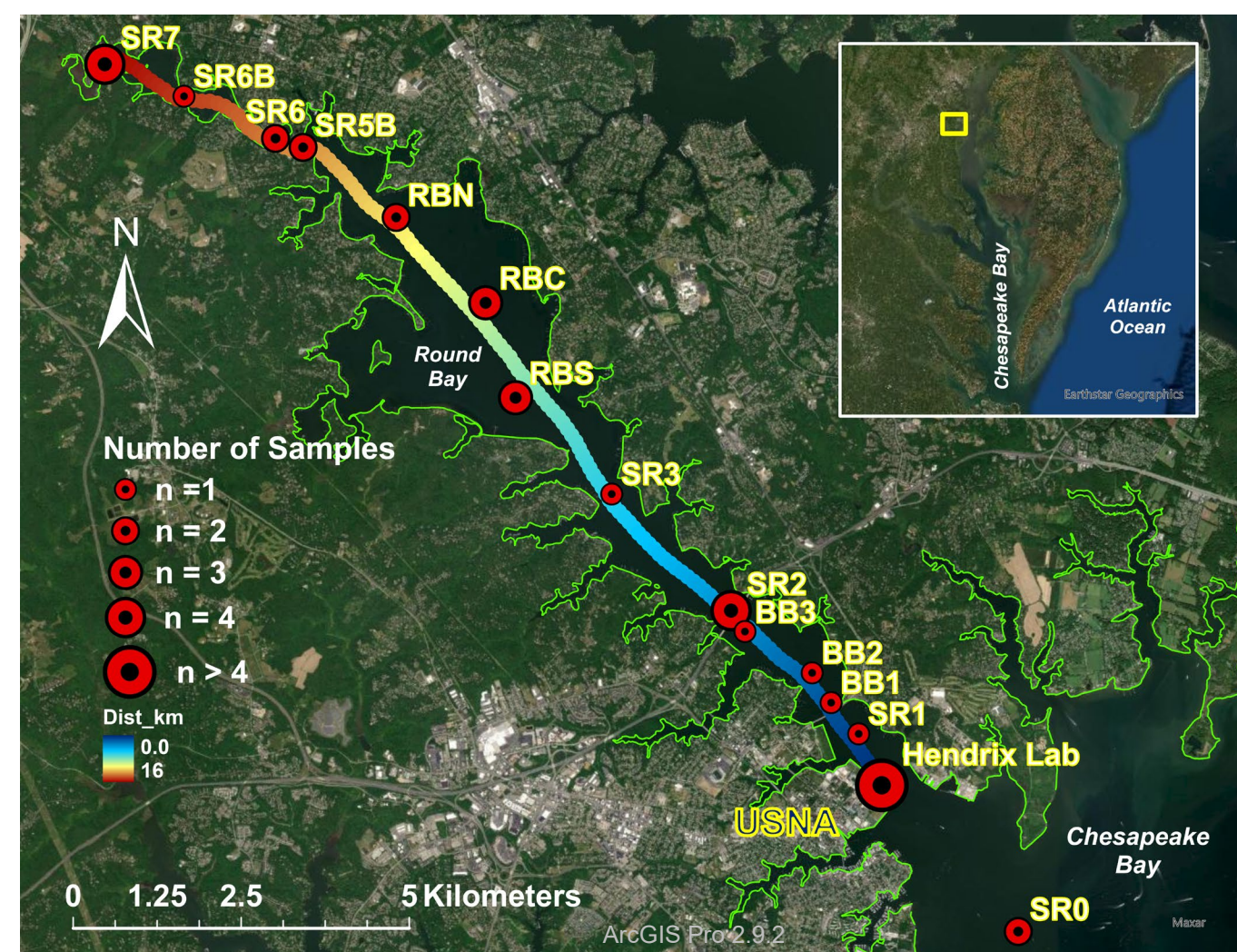
Advisers: Dr. Joseph P. Smith and Instructor Alexander R. Davies



## Background

The concentration of suspended particulate matter in an estuarine water column affects water clarity, water quality, the partitioning of reactive constituents between dissolved and particulate phases, and the scavenging of reactive constituents from surface waters. Total suspended solid (TSS) concentration is a quantitative measure of the mass of particulate matter (inorganics and organics,  $\geq 0.45 \mu\text{m}$  in effective diameter) in a fixed volume of water. In dynamic estuarine systems suspended particulate matter and TSS concentrations vary temporally and are spatially heterogeneous so quantifying TSS concentrations can be time consuming, laborious, and expensive. Turbidity is a measure of the amount of light scattered through a medium that can easily be measured *in situ* using optical probes. Since particles scatter light, the more particles present, the more light that will be scattered. In this way, turbidity can be used as a proxy for TSS concentration measurements but turbidity is also affected by factors other than suspended particulate matter such as particle size-specific concentrations, colored dissolved organic matter (CDOM), and photosynthetic pigments like chlorophyll-a (Chl-a) that, by function, absorb light at specific wavelengths. In this study discrete water samples were collected from the Severn River for TSS concentration measurements along with *in situ* salinity, and optical parameter data (turbidity, fluorescent dissolved organic matter (fDOM), and Chl-a over the period from August 2021 – March 2022. **Results were used to develop a regression model to predict TSS concentrations in the Severn River from *in situ* data collected using multi-parameter water quality sondes.**

## Study Area and Methods



**Figure 1.** Map of the Severn River, a tidal tributary to the mesohaline Chesapeake Bay (inset), showing locations where data and samples were collected ( $n = \#$  of samples) from August 2021 – March 2022. Also shown is the track for *in situ* data collected using a flow-through system on 04 March 2022. The graduated line shows distance up estuary from U.S. Naval Academy (USNA) Hendrix Oceanography Laboratory (HOL) at the mouth of the Severn (0 - 16 km). Sampling locations were chosen to reflect a diverse range of environmental conditions within the Severn River and Round Bay.

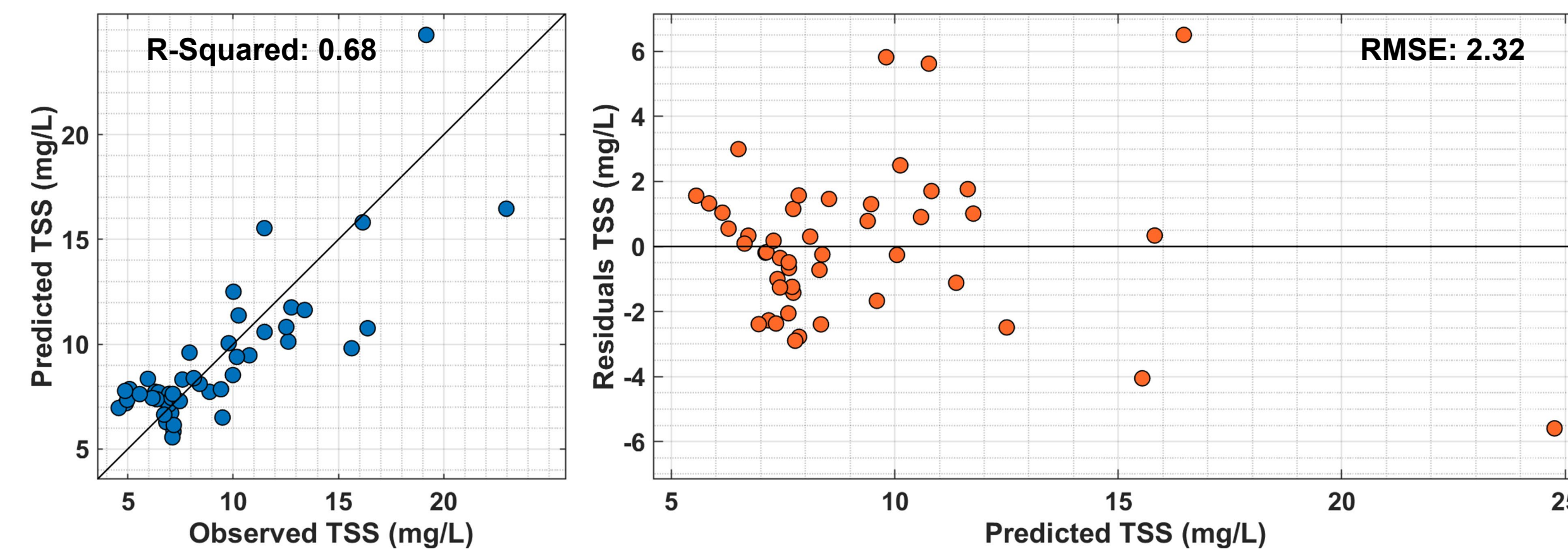
A total of 67 TSS concentration measurements were made at 14 sampling stations throughout the Severn River and Round Bay between August 2021 and March 2022 (Fig. 1). Water samples were collected with a Van Dorn bottle (Fig. 2a) at ~0.5 m depth. After collection, 250 mL was vacuum filtered through a pre-weighed  $0.45 \mu\text{m}$  nucleopore filter (Fig. 2b). The filter was dried and re-weighed to calculate the dry mass of particulate matter in the volume filtered (mg/L). A Xylem/YSI EXO2 Multiparameter Water Quality Sonde outfitted with the following pre-calibrated EXO SMART sensors: Conductivity & Temperature; Turbidity; Total Algae; and fDOM was used to collect coincident *in situ* measurements of water quality parameters. Water samples collected at HOL (Fig. 1) were matched with coincident calibrated water quality data from the Severn River Watershed Observatory at HOL, which employs the same above listed sensor suite.



**Figure 2.** a) Midshipman 1/C Alex Schuerch collecting a water sample from a small boat on the Severn River using a Van Dorn bottle and b) vacuum filtering a 250 ml water sample at HOL for gravimetric analysis of TSS concentrations in the Severn River.

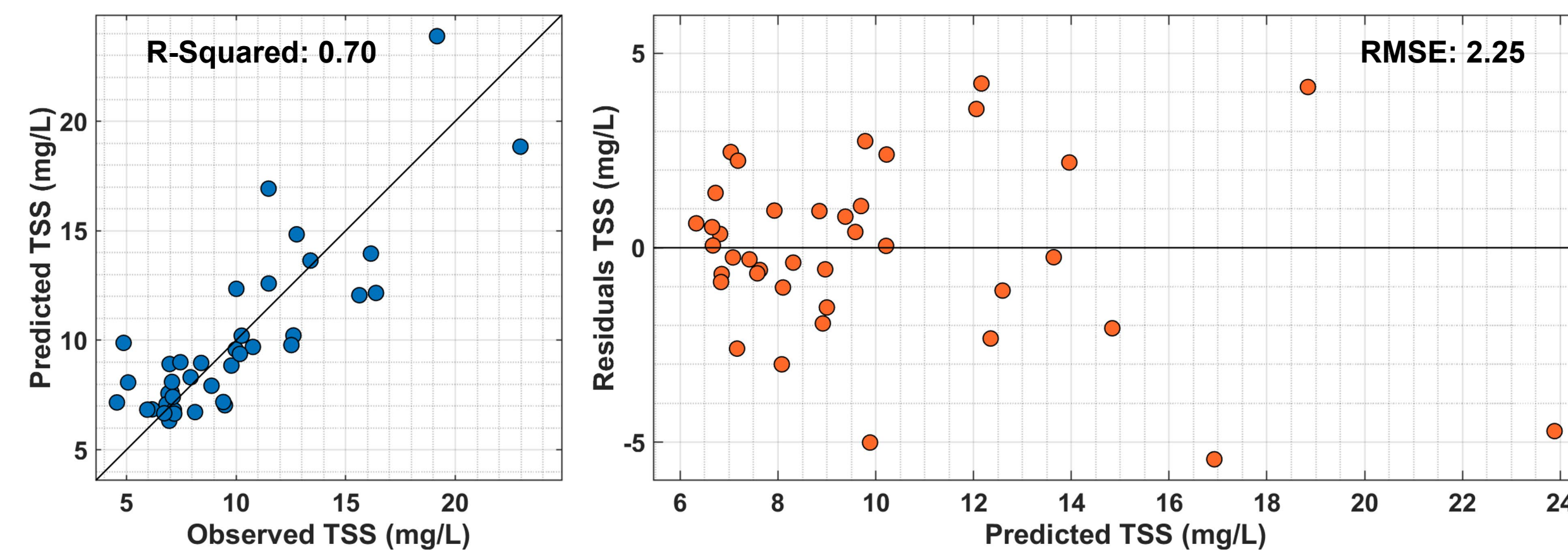
## Regression Model Development

Training - Linear (Turbidity Only)



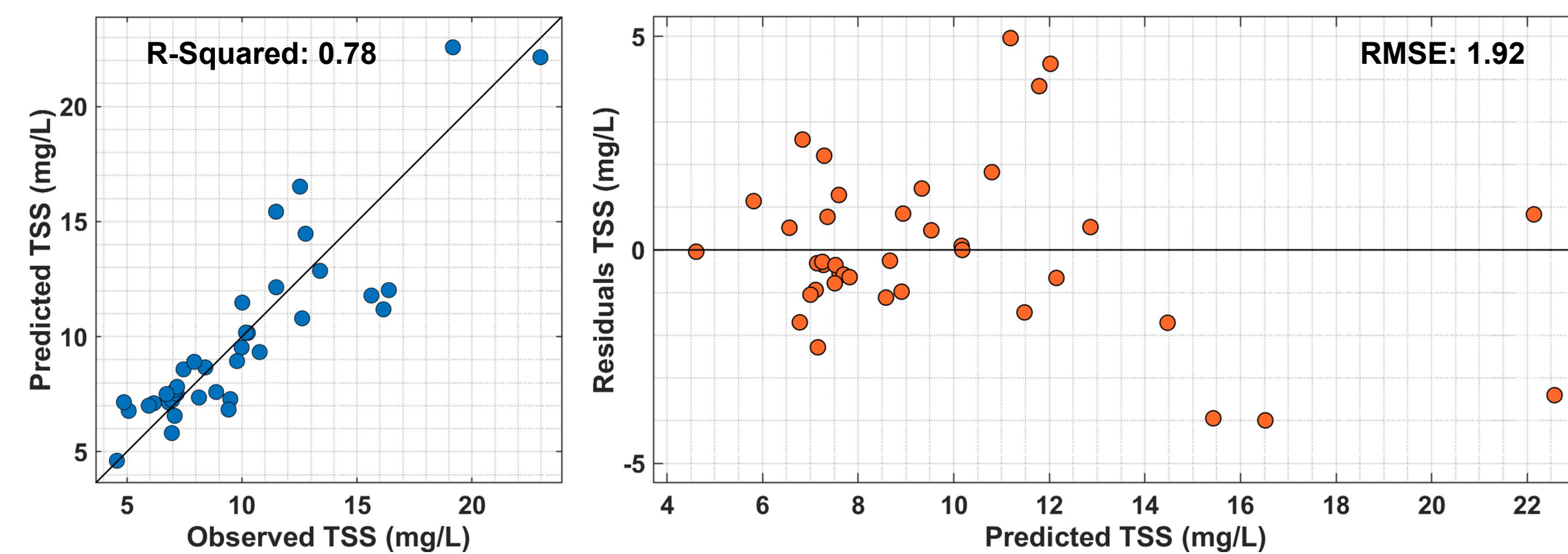
**Figure 3.** (Left) Predicted vs. observed TSS values (mg/L) from a linear machine learning model applied to a 70% training dataset using turbidity as the predictor variable. (Right) Residual vs. predicted TSS values (mg/L). The MATLAB R2021a regression learner toolbox was used to conduct all supervised modeling.

Training - Linear (All Variables)



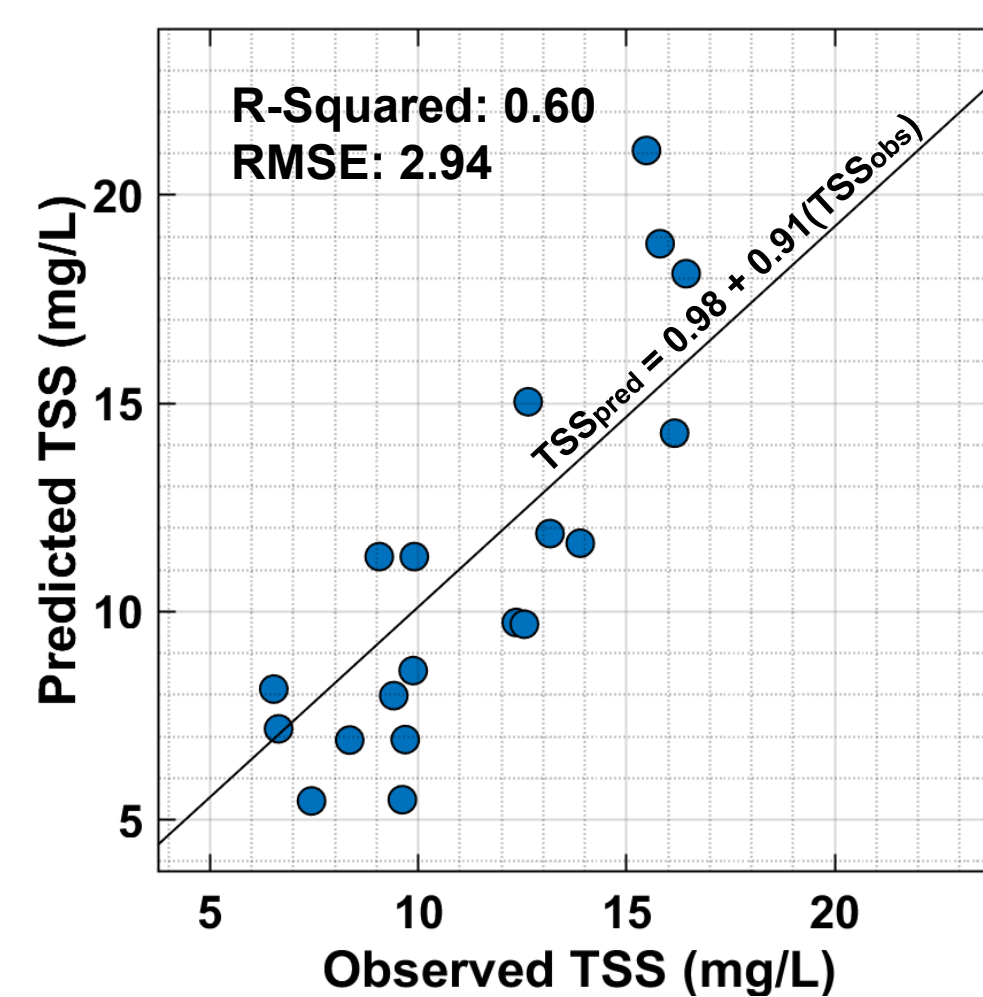
**Figure 4.** Same as Fig. 3 but using fDOM, turbidity, Chl-a, month, and salinity as predictor variables.

Training - Stepwise Linear (All Variables)



**Figure 5.** Same as Fig. 4 but using a step-wise linear regression machine learning model.

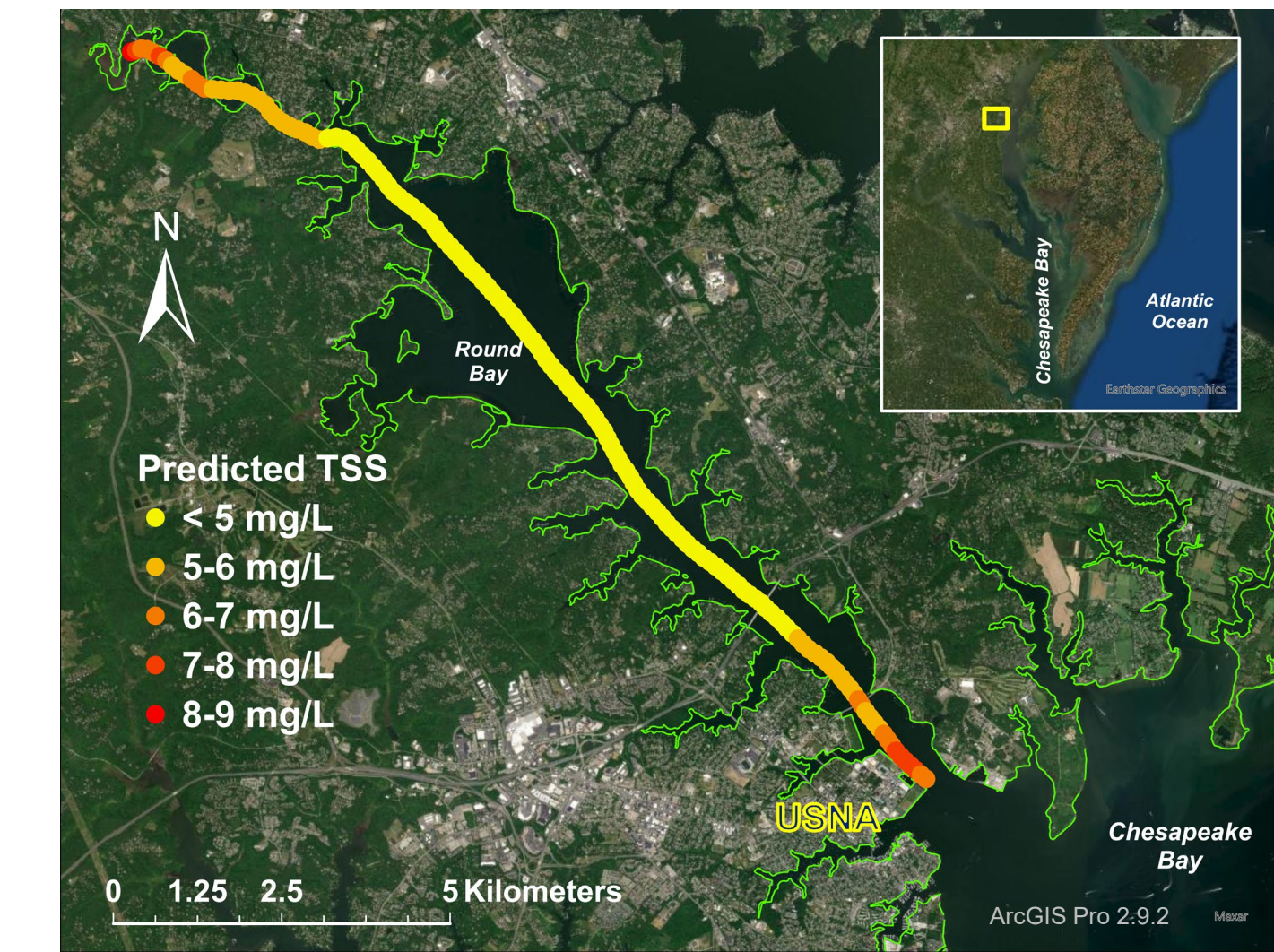
Testing - Stepwise Linear (All Variables)



**Figure 6.** The trained stepwise linear regression model (Fig. 5), tested against the randomized 30% testing data set and plotted against the observed TSS measurements. Linear equation of fit shown.

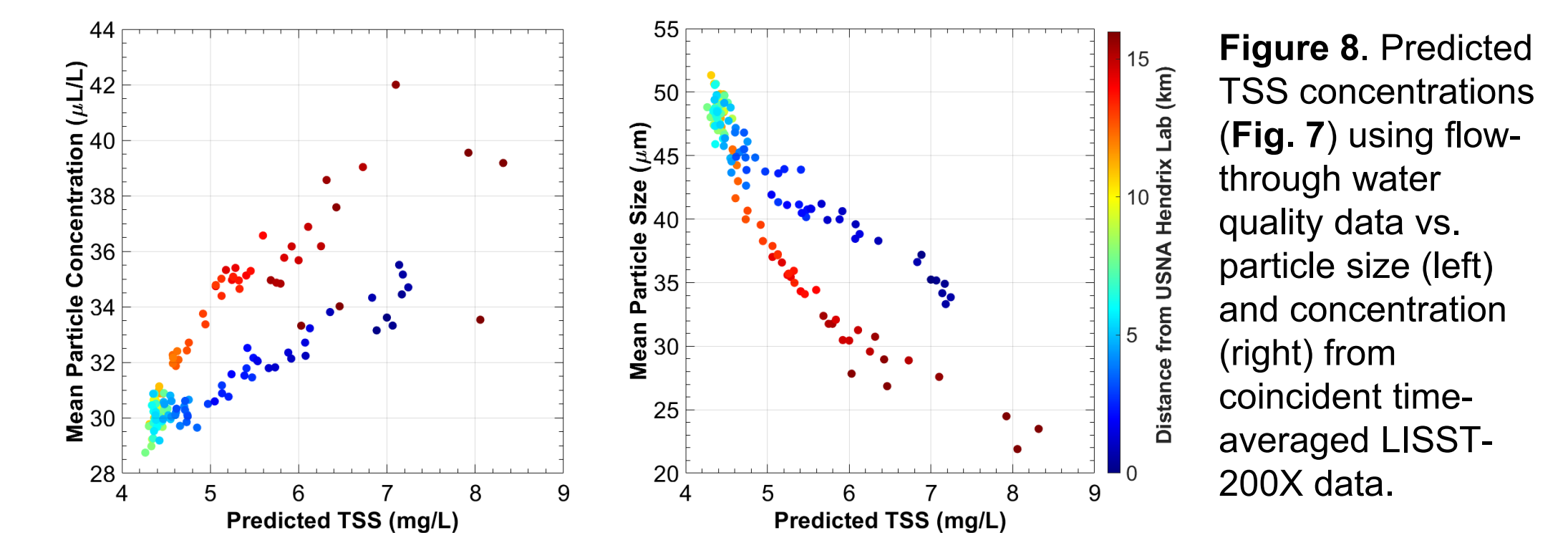
Measured TSS and *in situ* water quality parameter data was randomized into training (70%) and testing (30%) data sets for development of a linear machine learning model to predict TSS concentrations (Fig. 3-5). A linear regression model approach was used assuming a linear response of TSS to predictor variables. The initial training model (Fig. 3) used only turbidity as the predictor variable, as is commonly done to estimate TSS, resulting in an  $R^2$  value of 0.68 with a root mean square error (RMSE) of 2.32. Next, a linear model was trained with all predictor variables (Fig. 4), and the  $R^2$  value improved to 0.70 with RMSE of 2.25. The final trained linear model used a stepwise linear regression technique (Fig. 5) with an  $R^2$  value of 0.78 with a RMSE of 1.92. Stepwise linear regressions incrementally add or subtract predictor variable to maximize the model fit. The trained stepwise linear regression model (Fig. 5) performed well against the testing dataset (30%, Fig. 6) with an  $R^2$  value of 0.60 and RMSE of 2.94. Bias is not evident in the regression model residual data (Fig. 5).

## Discussion



**Figure 7.** Predicted TSS (mg/L) in the Severn River and Round Bay using the stepwise linear regression model (Fig. 5) applied to averaged, *in situ* water-quality data collected on 01 March 2022 using a surface water flow-through system that circulated surface waters through a Sequoia Scientific LISST-200X Submersible Particle Size Analyzer and a Xylem/YSI EXO2 Multiparameter Water Quality Sonde (Jordan et al., 2022).

The Xylem YSI EXO Turbidity SMART sensor uses white light (860 nm excitation wavelength) at a  $90^\circ$  detection angle (ISO7027 compliant; YSI/2019) eliminating direct interferences from CDOM and photosynthetic pigments like Chl-a during turbidity measurements. It is therefore not surprising that the TSS concentrations and *in situ* turbidity measurements in this study show a strong linear relationship (Fig. 3). Turbidity measurements at 860 nm are sensitive to variations in size-specific particle and floc concentrations and the inherent optical properties of particles present (i.e., organic vs. inorganic), especially in waters with low particle concentrations and/or with smaller sized-particles (Druine et al., 2018). Adding fDOM (the fluorescent fraction of the CDOM pool), salinity, Chl-a, and month as predictor variables to the linear model provides indirect proxies for factors like flocculation and increased particulate organic matter loading, improving the linear model fit and reducing the RMSE for TSS predictions (Fig. 4-6). Figure 7 shows predicted TSS concentrations using the stepwise linear regression machine learning model (Fig. 5) applied to temporally-averaged, *in situ* water-quality data collected using a surface water flow-through system on 01 March 2022. Comparing the TSS concentrations predicted along the axis of the Severn River to mean particle concentrations and size measured using the LISST-200X installed in the flow-through system (Fig. 8), there is a noticeable difference in the slope of predicted TSS vs. mean particle concentrations and size between the lower and upper Severn River and the uppermost regions of Round Bay where there are higher concentrations of smaller sized particles. This confirms that predictions of TSS concentrations from *in situ* turbidity data are highly dependent on site-specific environmental conditions influencing the particle size spectrum.



**Figure 8.** Predicted TSS concentrations (Fig. 7) using flow-through water quality data vs. particle size (left) and concentration (right) from coincident time-averaged LISST-200X data.

## Conclusions and Future Work

- Results demonstrate a regression model to predict TSS concentrations in the Severn River from *in situ* data collected using multi-parameter water quality sondes
- Future efforts should continue to test the model over multiple seasons and develop the model to work in other estuarine systems with a larger range of TSS concentrations, particle sizes, and turbidity values



**Acknowledgements:** This work was funded by the U.S. Defense Threat Reduction Agency (DTRA). Special thanks to Richard Fry and Bruce Trask (DTRA). Thanks to Mr. Benjamin Hickman (USNA) and Instructor Brianna Tracy (USNA) for data collection and TSS filtration assistance. Thanks to the Midshipman 1/C Deondre Williams, Daniel Taylor, and Brandon Jordan for sharing the flow-through data from the Severn River.