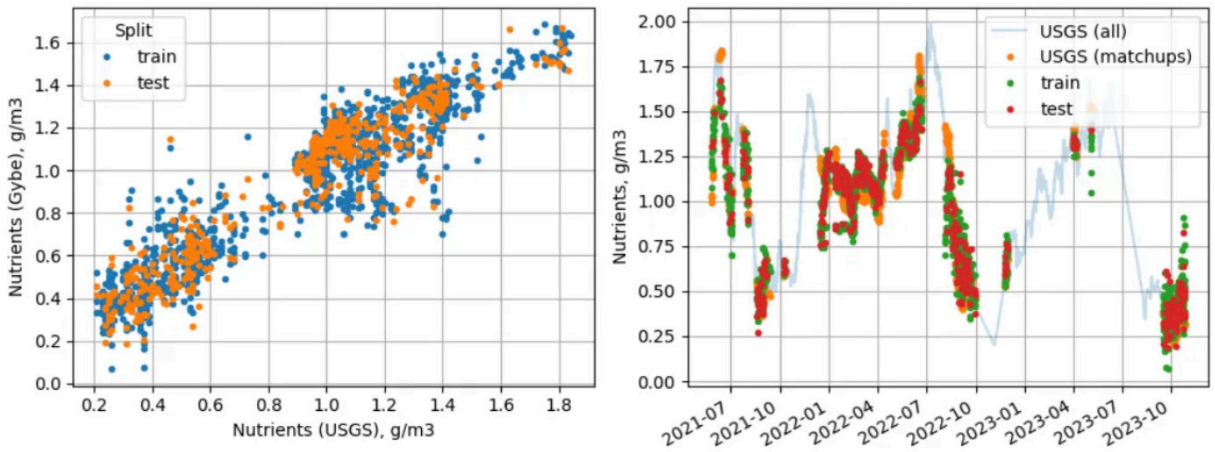
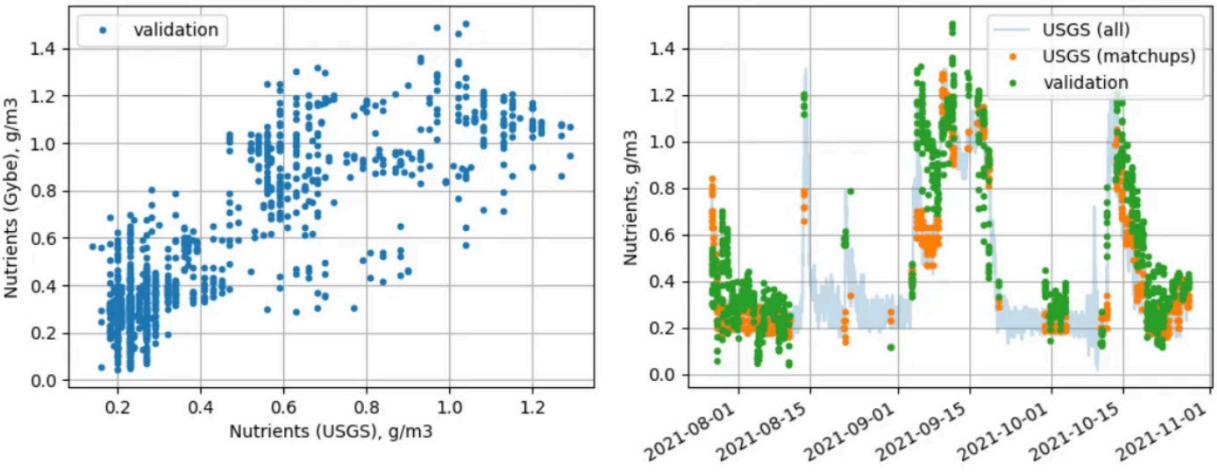
Machine Learning & Artificial Intelligence based Nutrient Modeling.

We developed a neural network-based approach to derive nitrate levels from hyperspectral reflectance observations. The network was trained on two years of matching Gybe sensor reflectances and USGS nutrient readings (1926 records) at Baton Rouge (Louisiana). A season worth of match-ups (1136 records) from a second site (De Soto, Kansas) were used to validate prediction performance and assess model transferability. Performance for the test-split (data not used during training) at Baton Rouge was convincing (R^2)=0.87, p<0.01, MAPE=0.17) and only slightly degraded when applied directly to our installation at De Soto (R^2)=0.72, p<0.01, MAPE=0.41) co-located with a different USGS nitrate site. We anticipate that additional model parameter tuning and more training data will further improve performance and robustness of these approaches, and we're actively seeking to validate against additional nutrient & hyperspectral reflectance datasets.

Our machine learning approach is designed to pick up on non-linear correlations between hyperspectral reflectance inputs and nutrients as target variable. In the current iteration of the model, we trained on data from a single site (at Baton Rouge, Louisiana), which transferred exceedingly well to observations at De Soto, Kansas. The next logical step is to train on a consolidated data set from as many possible sites and varied water types to demonstrate scalability. If such a generalized model performs equal or even better than site-specific models, we would be able to confidently derive nutrients for extended transects across a large range of water types and qualities along the Mississippi River and other basins. Tying these models to multispectral satellite-derived reflectances would enable mapping of quantified nutrient concentrations over large basin extents and updated as frequently as daily.



Nutrients predicted from Gybe hyper-spectral reflectance measurement wes, compared against USGS reference readings. The model was trained on 75 % of the data set and tested on the remaining 25 % at the USGS site in **Baton Rouge, Louisiana**.



Validation of the model results on a completely independent data set collected at **De Soto, Kansas**.