FWP 54737

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

Pacific Northwest National Laboratory SFA Annual Report

River Corridor Hydrobiogeochemistry from Reaction to Basin Scale

June 2021

Laboratory Research Manager: Charlette Geffen

Principal Investigators: Tim Scheibe (PI), Xingyuan Chen (Co-PI), and James Stegen (Co-PI)

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <http://www.ntis.gov/about/form.aspx>
Online ordering: http://www.ntis.gov

This document was printed on recycled paper.
(8/2010)

PNNL SBR Scientific Focus Area Annual Report FY2021

River Corridor Hydrobiogeochemistry from Reaction to Basin Scale

2021 Annual Report

June 30, 2021

TABLE OF CONTENTS

I. PROGRAM OVERVIEW

The Pacific Northwest National Laboratory (PNNL) Environmental System Science (ESS) Science Focus Area (SFA) is transforming understanding of spatial and temporal dynamics in river corridor hydrobiogeochemistry from reaction to watershed and basin scales, enabling mechanistic representation of river corridor processes and their responses to disturbances in multiscale models of integrated hydrobiogeochemical function.

Our research is focused on understanding the controls on spatial and temporal variations in river corridor hydrobiogeochemistry, hydrobiogeochemical responses to wildfires and other disturbances, and representation of river corridor hydrobiogeochemistry in numerical models from reaction to basin scales. The project's goals are aligned with the objective of DOE's Office of Biological and Environmental Research (BER) to improve scientific understanding and prediction of the function of natural and managed watersheds and their responses to disturbances.

Our long-term vision is to mechanistically link the impacts of disturbance on hydrologic exchange flows (HEFs, the exchange of water between surface and subsurface environments in river corridors), molecular processes, and biogeochemistry to watershed function across CONUS basins. This vision is being achieved through distributed, basin-scale science pursued via extensive collaboration with the research community following ICON-FAIR principles[39]. These principles focus on doing science by-design that **I**ntegrates physical, chemical, and biological processes, is **C**oordinated via consistent methods from field to lab to analysis, uses **O**pen science methods such as making data **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable (FAIR), and is **N**etworked with the community to enable distributed data generation and modeling that are mutually beneficial to all. Our hypothesis-driven approach advances transferable understanding of coupled hydrobiogeochemical processes through integrated multiscale experiments, observations, and modeling. Enabling mechanistic representation of river corridor processes from reaction to watershed and basin scales will provide a foundation for developing the next generation of watershed models with enhanced predictive capacity to inform watershed management strategies aimed at solving the nation's environmental challenges in the face of extreme disturbances.

Progress toward this vision is achieved through four integrated research campaigns (RCs) integrated through a Multiscale Modex approach shown in Figure 1, in keeping with the concept of iterative model-driven experimentation and observation. SFA team members work on multiple campaigns, campaign activities are jointly coordinated by the PI team, and high-impact publications are targeted that integrate information across the four campaigns to address high-level project objectives.

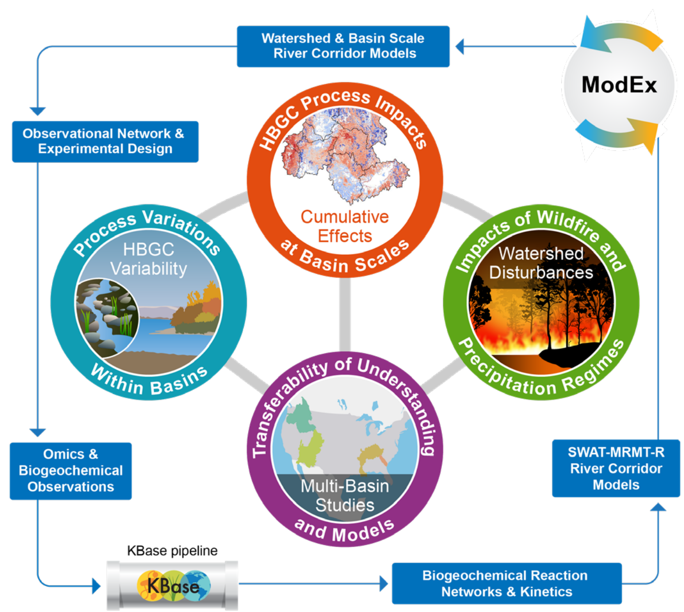


Figure 1. The SFA's Multiscale ModEx (MM) approach is essential for using knowledge and data to inform observations and models across scales and integrating across the four Research Campaigns. MM is an iterative data-model learning loop in which information is passed among models and experiments across scales.

- The *Cumulative Effects Campaign* (RC-1) aims to reveal the cumulative effects of river corridor processes and their appropriate representations in watershed- and basin-scale models.

- The *HGBC Variability Campaign* (RC-2) aims to elucidate interactions among hydrologic and molecular processes that control the cycling of nutrients (N, P), DOM, and inorganic contaminants (e.g., NO_3^-) in river corridors from reaction-to-basin scales.
- The *Watershed Disturbances Campaign* (RC-3) aims to reveal the mechanisms by which wildfires impact biogeochemical cycling in river corridors from reaction to basin scale.
- The *Multi-Basin Studies Campaign* (RC-4) aims to provide transferable principles that integrate DOM chemistry, microbial gene expression, biogeochemical function, and disturbance by combining existing global-scale data with WHONDRS-based data generation and numerical modeling distributed across CONUS basins.

II. KEY SCIENTIFIC OBJECTIVES

In keeping with the vision outlined above, we have expanded both the physical scale and complexity of our research while maintaining our focus on the study of HEFs, DOM chemistry, microbial activity, and associated biogeochemical processes in the river corridor and their cumulative impacts at watershed and basin scales. Our long-term objective is to expand the scope of our study to the scale of the full Columbia River Basin (CRB) encompassing more than 460,000 km of perennial streams in an area of 670,000 km^2 and to extend the transferability of our science by studying multiple basins across the CONUS. Going to larger scales provides new opportunities to study broadly distributed disturbances and their impacts across diverse environmental conditions. Our work will focus on impacts of wildfire and precipitation, while spanning gradients in climate, vegetation, land use, and other key watershed features. As an intermediate step toward this long-term objective, in FY21 we have focused on study of the Yakima River Basin (YRB), a major sub-basin of the CRB in which there exists a wide range of stream orders and physiographic watershed settings, allowing us to generalize site-specific findings from our previous studies to have broad applicability. Our research is structured around the scientific grand challenge defined below.

Scientific Grand Challenge: Understand and quantify processes governing the cumulative effects of HEFs, DOM chemistry, microbial activity, and disturbances on river corridor hydrobiogeochemical functions at watershed to basin scales.

Dynamic HEFs are a primary driver of river corridor biogeochemistry, which is highly sensitive to DOM chemistry and microbial activity (e.g., expressed metabolic pathways). Disturbances such as wildfire and extreme precipitation interact to create feedbacks among physical and biogeochemical processes. These concepts motivate both fundamental process studies and the development/implementation of a multiscale modeling framework, and lead to the following SFA-level science questions that guide our research plan:

Overarching Science Questions:

1. How do HEFs, DOM chemistry, microbial activity, and disturbances interactively influence river corridor hydrobiogeochemical function from reaction to basin scales?
2. How can mechanisms that govern river corridor hydrobiogeochemistry be efficiently and sufficiently represented in integrated land surface models at scales relevant to regional and national water challenges?

III. PROGRAM STRUCTURE

The PNNL SBR-SFA is led by a Principal Investigator (PI, Tim Scheibe) and two Co-PIs (Xingyuan Chen and James Stegen) (Figure 2). A Laboratory Research Manager (Charlette Geffen) provides laboratory oversight and guidance to the PI and Co-PIs and serves as the primary point-of-contact with the DOE Office of Biological and Environmental Research (BER). The SFA is organized around four Research Campaigns (RCs; Figure 1), each of which is led by one of the three PIs. Each RC comprises two to three major Activities; each Activity is organized into several Sub-Activities. The Multiscale ModEx (MM) cross-cutting activity serves an integrating role to ensure that observational, experimental, and modeling research are highly coordinated. Activity Leads are responsible for coordinating research within each Activity and are members of the SFA Leadership Team. Other key staff have responsibilities within one or more Activities or Sub-Activities and contributing staff (including funded collaborators) are disciplinary experts that lead or contribute to one or more RCs. SFA researchers are drawn from multiple research directorates at PNNL as dictated by the interdisciplinary nature of SFA research. The RCs, Activities, and Sub-Activities have been jointly designed and are closely coordinated by the Leadership Team to accomplish overall project scientific objectives.

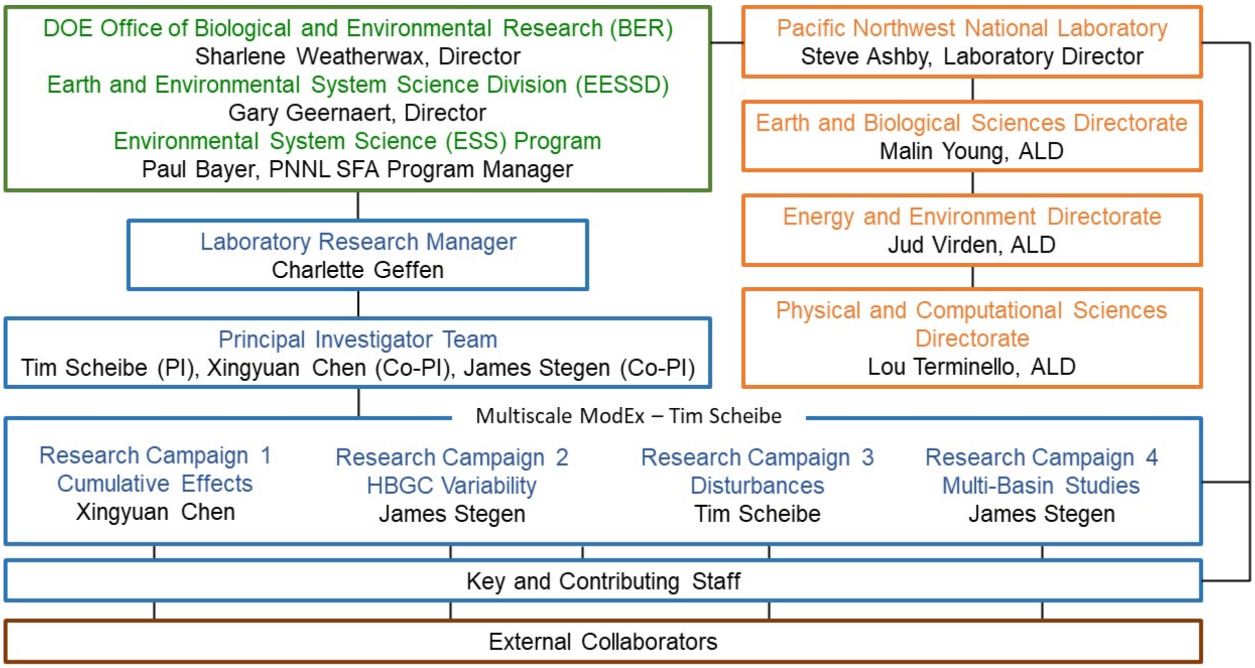


Figure 2. PNNL River Corridor SFA organizational structure.

Our management philosophy and process are characterized by 1) intentional, regular and transparent communication and documentation of progress and 2) emphasis on individual awareness of and accountability to team outcomes. This, in the context of our team values of scientific excellence, creativity, collaboration, and mentorship, leads to active engagement of the entire project team in pursuing the integrated objectives of the project. Primary avenues for team coordination and communication are: 1) regular meetings at a variety of organizational levels, and 2) extensive utilization of collaborative environments including Microsoft Teams (virtual meetings, file sharing, persistent chat), Atlassian Confluence(wiki), AirTable (task tracking), GatherTown (virtual workspace), and others.

IV. PERFORMANCE MILESTONES AND METRICS

Review of Scientific Progress Toward Program Objectives and Milestones

Research Campaign 1 (RC-1): Cumulative Effects

Overall Objectives: *The Cumulative Effects Campaign aims to reveal the cumulative effects of river corridor processes and their appropriate representations in watershed- and basin-scale models.*

- Quantify baseline cumulative effects: Quantify the cumulative effects of river corridor HEFs, DOM chemistry, and microbial activity on watershed and basin biogeochemical cycling, water quality (including temperature), contaminant mobility, and land surface fluxes. Use reaction network models and rate kinetics developed from lab and field studies in RC-2 to inform river corridor reactive transport model.
- Quantify cumulative effects of disturbances: Reveal the cumulative impacts of wildfire and precipitation on watershed- to basin-scale hydrobiogeochemical functions of river corridors. Leverage new reaction networks and kinetics, focusing on the effects of wildfire-affected DOM chemistry on river corridor metabolism, developed in RC-3 in coordination with RC-2.
- Guide field and lab experiments: Provide basin-scale river corridor model outputs of biogeochemical hot spots and hot moments to focus field and laboratory experiments of RC-2 and RC-3 in locations and time windows that are most effective at reducing uncertainties in watershed model predictions.

Key Contributions of this Research Campaign to the SFA: RC-1 is integrating numerical models developed at multiple scales with distributed monitoring in other RCs across the Yakima River Basin to advance predictive understanding of emergent system behaviors arising from complex hydrobiogeochemical interactions. We are building a watershed- to basin-scale river corridor model that links dynamic flow processes with variable temperatures and reaction kinetics (informed by molecular properties) to investigate water, energy, and mass fluxes across the river-groundwater interface. This model will be used to further quantify the cumulative impacts of HEFs, molecular properties, and disturbances on the watershed- and basin-scale biogeochemical cycling of key nutrients (C, N, P) and inorganic contaminants (e.g., NO_3^-).

FY21 Science Plan Milestone: Evaluate the impacts of groundwater and channel water chemistry on basin-scale biogeochemical function in river corridors using HEFs and residence time from NEXSS and nutrient inputs from SPARROW and SWAT. Outputs will guide RC-2 field activities. Couple 3D PFLOTRAN with SWAT. Set up, parameterize, and calibrate the SWAT and CLM-PFLOTRAN models in representative watersheds. Collect and analyze basin-scale water chemistry data and remote sensing data products for model parametrization, calibration, and validation.

Progress Brief for FY21

Our activities in FY21 are focused on quantifying the spatial and temporal variation of hyporheic denitrification and aerobic respiration in the Yakima River Basin (YRB) to guide RC-2 sensor placement and sampling design. We have also developed integrated hydrologic watershed models for a selected headwater catchment within the YRB using ATS and SWAT to quantify the watershed hydrologic inputs to river corridors and study the dynamics of watershed streamflow, evapotranspiration (ET), and snow melting. Machine learning methods have been applied to calibrate the SWAT models using the USGS gauge streamflow records and MODIS ET products. More specifically, the RC-1 team has achieved the following progress:

- We applied the SWAT-MRMT-R river corridor model to quantify the spatial and temporal variation of hyporheic denitrification and aerobic respiration in the Columbia River Basin. The spatially distributed nutrient loading was derived from the USGS SPARROW (SPAtially Referenced Regression On Watershed attributes) model, which estimates long-term average loading of a contaminant delivered downstream, based on existing monitoring data, location and strength of contaminant sources, and characteristics of the landscape. A random forest model revealed that the total river corridor denitrification assuming steady-state flow is primarily controlled by the riverbed physical characteristic and land use. A subset of the river corridor modeling results within the YRB was analyzed for representative classes of nutrient processing, which were used to selection monitoring and sampling locations for RC-2.
- Integrated watershed models have been developed for the American River watershed (ARW) within the YRB using ATS (with the newest evapotranspiration (ET) module) and SWAT models. Long

term (1997-2020) model simulations have been validated against USGS stream discharge at the outlet and MODIS ET and snow products. Reanalysis outputs from the National Water Model (NWM) were retrieved at the outlet ARW for comparison with the ATS and SWAT results. Before calibration, ATS has shown substantially better agreement with the observed streamflow, ET and snow water equivalent (SWE) than the empirical SWAT model. Significant performance improvement was achieved for the SWAT model after calibrating against the observed streamflow. NWM reanalysis outputs showed the best agreement with the observed streamflow.

- Remotely sensed data were used to inform, calibrate and validate watershed models. The ECOSTRESS-based (ET) estimates within the Priest Rapids Upper Columbia watershed were found to decrease with the distance from the Columbia River with a high correlation with the groundwater depth, confirming the groundwater dependence of the plant water use in the semi-arid regions. We also used MODIS ET and leaf area index (LAI) data products to improve the model parameterization of ATS and SWAT watershed models for ARW.
- We have developed a deep learning (DL)-based inverse modeling method to estimate the SWAT model parameters, leveraging a capability developed on the Exasheds project. Ensemble SWAT simulations were used to train, validate, and test DL models that inversely map model simulated system responses to model parameters. The trained DL models were then used to estimate the real parameters of the ARW using the real observations. When compared to the generalized likelihood uncertainty estimation (GLUE) method, the parameters estimated by the DL models led to more accurate predictions in streamflow for the ARW.
- We analyzed and modeled the chamber-based soil respiration data collected at the US-Hn1 Ameriflux site, using artificial intelligence (AI) techniques. DL techniques were used to model sub-hourly soil respiration using temporal sequences of ecosystem forcing and states as inputs. An exploratory analysis was performed to identify the key interactions between soil respiration and other environmental variables, guiding the design of DL models. A bi-directional Long Short Term Memory (biLSTM) model was built to predict the soil respiration from both soil temperature and wetness along with meteorological forcing. The biLSTM model outperformed the widely-adopted Q10 approach and other two benchmark DL models that do not explicitly transfer information between past and future states. Then, an explainable AI (XAI) technique, layerwise relevance propagation (LRP), was used to unravel the relations between soil respiration and biLSTM inputs learnt by the DL models. LRP analyses revealed two distinct controlling mechanisms under the dry and wet conditions, respectively.
- A workflow using Jupyter notebook has been leveraged and enhanced to generate unstructured meshes for watershed models using online data sources from USGS, NRCS, USDA and others. A new jupyter notebook workflow has been developed to download Daymet climate data for a given watershed and reformatted to conform with ATS input format.

Publication Highlight: River Dynamics Control Transit Time Distributions and Biogeochemical Reactions in a Dam-Regulated River Corridor. Dams upstream and downstream of the Hanford Reach of the Columbia River induce frequent variations in river stage. To understand how this variation affects hydrological exchange between river water and groundwater, and downstream nutrient cycling, we first used the extensive site characterization data available for this river corridor to build a baseline groundwater flow and transport model. Then we used a forward particle-tracking method (Figure 3) to estimate transit-time distributions for water flowing through the subsurface aquifer during a seven-year simulation window. We paired the estimated transit-time distributions with the rates of aerobic respiration and denitrification known to occur along this section of the river corridor to quantify rates and amounts of nutrients being processed. By evaluating the effects of dam operation on transit times and nutrient cycling rates and amounts, we found that dam-induced high-frequency variations in flow increased hydrologic exchanges between the river water and groundwater. These increases accounted for 44% of nutrient consumption in the river corridor along the Hanford Reach. The numerical particle-tracking approach developed in this study can be extended to other study sites that have robust site characterization data, and this approach can be very useful for extending nutrient cycling models from river reaches to larger-scale watersheds and basins.

Reference: Song, X., Chen, X., Zachara, J. M., Gomez-Velez, J. D., Shuai, P., & Ren, H., et al. (2020). River dynamics control transit time distributions and biogeochemical reactions in a dam-regulated river corridor. *Water Resources Research*. DOI: 10.1029/2019WR026470

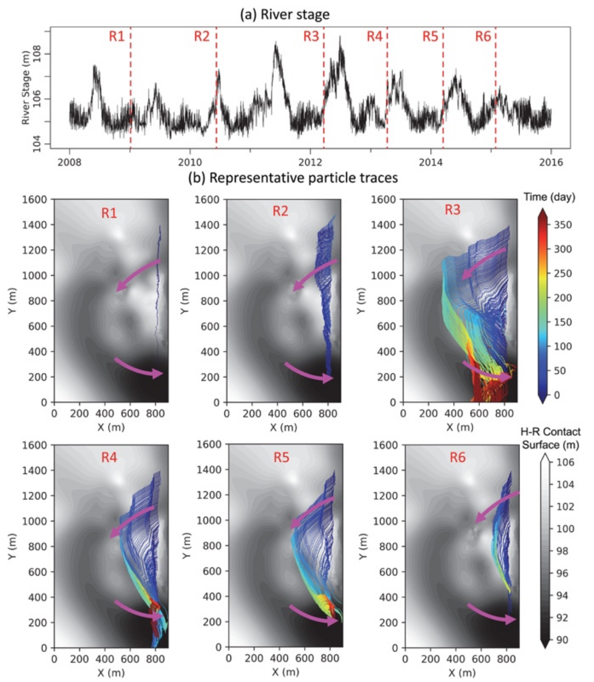


Figure 3. Flow paths of particles released at an elevation of 103.5 m: (a) river stage time series marked with six representative release times and (b) representative particle traces. The colors in (b) indicate the residence time of each particle in the aquifer.

Plans for FY22

In FY22, we will analyze the historical remotely sensed data to investigate the impacts of wildfire on vegetation, water quantity and quantity in the Yakima River Basin. We will use selected representative watersheds in the Yakima River Basin with varying degrees of burn severity to perform detailed modeling and evaluate how watershed processes are altered by the combined effects of wildfires and precipitation scenarios. The simulation results will guide monitoring and sampling efforts of RC-3. We will link the mechanistic understanding of wildfire-impacted DOM chemistry and hydrobiogeochemical regimes (informed by RC-3) with wildfire-impacted watershed hydrology to access the cumulative effects of wildfire impacts on river corridor biogeochemistry. In collaboration with IDEAS-Watersheds, we will lead watershed and river corridor model intercomparisons using the ARW as a testbed. We will collaborate with Exasheds to develop DL-enabled multifidelity watershed models for water quality and quantity by leveraging monitoring data and simulations by SWAT, ATS and NWM that will produce high-quality predictions with affordable computing cost.

<u>Research Campaign 2 (RC-2): Hydrobiogeochemical Variability</u>

Overall Objectives: *The HBGC Variability Campaign aims to elucidate interactions among hydrologic and molecular processes that control the cycling of nutrients (N, P), DOM, and inorganic contaminants (e.g.,* NO_3^-*) in river corridors from reaction-to-basin scales.*

- <u>Identify important places/times:</u> Use NEXSS and SPARROW predictions (existing and from RC-1.1) to guide placement of in situ sensors (FY21) to span a range of predicted respiration contributions of sediment-associated microbes (ERsed), relative to those in surface water. Use to identify where and when ERsed is disproportionately high.
- <u>Compare to predictions:</u> Compare field-estimated ERsed to NEXSS predictions (FY21) and use outcomes to inform structure of and parameterize basin-scale models in RC-1.1 (FY22).
- <u>Characterize variation:</u> Use field surveys (FY22–23) across reaches that differ in ERsed to characterize longitudinal and seasonal variation in DOM chemistry, microbial gene expression, and nutrient concentrations in surface and pore water. Contribute data to multiscale/basin analyses in RC-4.4.

- Inform models: Use field survey data (and data from RC-4.1,4.2,4.4) to inform 1-D reactive transport models (linked to RC-2.3 in FY23) that predict biogeochemical rates using explicit representation of DOM chemistry, microbial gene expression, and nutrient concentrations.
- Understand consequences: Use RC-1.1 models (developed in FY20–21) and field surveys in (RC-2.1 in FY21) to guide lab experiments to reveal influences of DOM chemistry, microbial gene expression, and nutrient concentrations on biogeochemical rates predicted by reactive transport models.
- Refine models: Use experimental outcomes to test and refine substrate/microbe-explicit reaction network models (from RC-2.2 in FY23), use these models to evaluate impacts of PyOM (with RC-3 in FY23), and integrate them into reactive transport models to inform RC-1.1.

Key Contributions of this Research Campaign to the SFA:

The SFA is designed to increase predictive understanding of the variations in river corridor hyrdobiogeochemical processes across stream orders as well as other climatic, ecological, and geographic settings. RC-2 has started to elucidate interactions among hydrologic and molecular processes that control the cycling of nutrients (N, P), DOM, and inorganic contaminants (e.g., $NO3-$) in river corridors by initiating a series of field-based sampling and monitoring campaigns across the YRB. As these campaigns generate datasets spanning both long periods of time and a broad variety of geographic settings, they provide the basis for both exploring the mechanistic linkages between process and setting, as well as iteration with numerical models to improve predictive ability.

FY21 Science Plan Milestone: Deploy sensors across low to high order reaches predicted to span a continuum of sediment contributions to system respiration. Compare outcomes to existing model predictions. Work with RC-1 to understand how deviations from model predictions inform need for additional process resolution in basin scale models. Work with RC-3 to use outcomes to inform disturbance-focused field campaigns.

Progress Brief for FY21: Our activities in FY21 are focused on laying the logistical groundwork for and initiating the process of collecting and analyzing a combination of both long time series and widely spatially distributed measurements of water parameters relevant to hydrobiogeochemical processes.

- RC-2's role in the RCSFA is centered around a pair of large-scale field monitoring campaigns. One campaign is designed to cover a smaller number of sites over an extended continuous time period (the temporal study), while the other is designed to cover a large range of field sites over a brief period of time (the spatial study). The logistical systems and methods development for both campaigns are significant undertakings. Some of the logistics and methods overlap, while others do not. Additionally, many of the logistical structures put in place are collaborative across RCs.
- The primary data sources for RC2 activities are long-term deployments of sensors, short term spot sensor readings ("dip sensors"), and samples collected in the field and returned to the main PNNL campus and MCRL for analysis. A great deal of progress has been made developing the sample collection and processing pipeline particularly with regard to samples analyzed at MCRL. Additionally, scripting to automate the processing of raw MCRL lab data and connection of field metadata to sensor data and lab-generated data has been developed to streamline data management and ensure that data products are well organized and ready to upload to ESS-DIVE. Additional technical staff have been brought on board at MCRL to run samples, supporting RC3 and RC4 as well as RC2.
- Through cross-RC meetings including both field staff and modelers, a sensor package was chosen that meets the requirements of RC2 while also supporting sensor needs of RC3. This allows scripting and data management practices to be coordinated across RCs, reducing the need for duplicative data management and providing cross-RC data compatibility.
- In addition to data logistics and sensor selection, RC2 has made significant progress with methods development. Several novel techniques and technologies are slated for deployment, including

"autochambers" for time-series metabolism estimation, "manual chambers" for single-measurement metabolism estimation, and kayak-based bathymetry transect collection systems. These systems have been prototyped, iterated on, and are either in late-phase testing/validation or are actively deployed.

- The landowner relationships and state and federal permits required to conduct large-scale fieldwork in the YRB have been cultivated and applied for. In particular, RC2 has reached out to the Yakama Nation, as well as a number of federal, state, and local agencies and stakeholders, including the United States Geological Survey, US Bureau of Reclamation, Washington State Department of Fish and Wildlife, Washington State Department of Ecology, and the Benton County Conservation District. These agencies and stakeholders are very interested in the work RC2 is engaged in, particularly with regard to increased understanding of the overall health of the YRB system and fisheries.
- RC2 and RC1 have collaborated heavily to use ModEX methodology to select RC2 field study sites for both the temporal and spatial studies. RC1 conducted analysis of the YRB models to determine the spatial variables that most strongly influenced predicted rates of metabolism, and then carried out clustering analyses with these spatially distributed variables. RC2 and RC1 iterated on these cluster maps, choosing field sites that satisfied access and logistical restrictions while also providing the best possible coverage across the clusters and underlying model-relevant variables.
- The temporal study has been initiated, with field staff collecting weekly water quality data at 6 sites across the YRB. An initial manuscript using data from these samples is under development as a collaboration between RC2 and RC3. It focuses on organic matter chemistry across stream orders and time. Deployment of autochambers will commence once permits are approved. In the meantime, RC2 is collaborating with Benton Conservation District and Washington State University to carry out a test deployment of the autochamber system as well as side-by-side comparison to the manual chamber system in the lower Yakima river, with a planned methods manuscript to follow.
- The spatial study is planned for the August-September timeframe, with deployment of the newly developed "manual chamber" system as well as a newly developed kayak-based bathymetric transect method.

Publication Highlight: Using metacommunity ecology to understand environmental metabolomes.

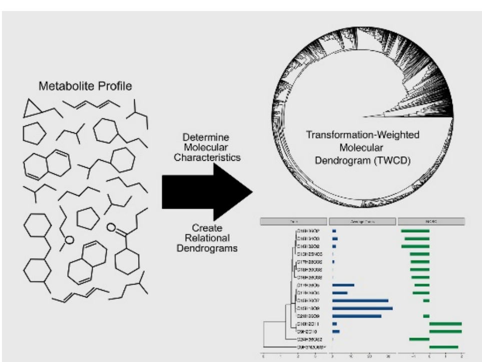


Figure 4. To study environmental metabolomes, we integrated trait-based dendrograms (right) with metabolomic data (left).

To better understand processes that constrain or promote variation in metabolomes of a given system, we integrated FTICR-MS data (from EMSL) with tools and concepts from community ecology. With these data, we developed several metabolite dendrograms (Figure 4) to group molecules based on common traits, such as elemental composition, structural features, and biochemical transformations. Next, we performed ecological null modeling, a common approach used in meta-community ecology but never before applied to organic metabolite assemblages. The null models quantified processes that governed the assembly of molecules into metabolomes. We found metabolites that were potentially biochemically active were more deterministically assembled than less active metabolites. Organic metabolites that are biochemically active and deterministically organized are most important to represent in mechanistic models. This approach provides a tool for modelers to winnow the enormously complex milieu of organic molecules down to a subset that is most important for enhancing predictive capacity. This tool is poised to be used with large-scale molecular analysis of environmental metabolomes, such as those from WHONDRS.

Reference: R.E. Danczak, et al. (2020) "Using metacommunity ecology to understand environmental metabolomes." *Nature Communications*. DOI: 10.1038/s41467-020-19989-y

Plans for FY22: Following the heavy investment in logistics, permitting, and data management of FY21, RC2 plans to operate within those structures for the majority of FY22. Data analysis and manuscript writing based on both spatial and temporal study results will be carried out. Time series water quality measurements will continue, as will time-series estimation of metabolism across the YRB. Cross-RC integration with RC1 for ModEx-based iterative study design will remain a priority, as will continued engagement with YRB collaborators and stakeholders. In addition, the logistical groundwork for FY23 will be laid. FY23 will involve sediment sample collection and monitoring of pore water chemistry, both of which will require significant cultural resources permitting, which will be pursued throughout FY22.

Research Campaign 3 (RC-3): Watershed Disturbances

Overall Objectives: *The Watershed Disturbances Campaign aims to reveal the mechanisms by which wildfires impact biogeochemical cycling in river corridors from reaction to basin scale.*

- Identify impacts of burn severity: Identify the impacts of burn severity on relationships between pyrogenic material (e.g., PyOM and inorganic nutrients) and river corridor biogeochemistry.
- Derive PyOM indicators: Reveal biochemical indicators of PyOM derived from different burn severities that can be used for multiscale characterization of PyOM distributions.
- Understand temporal trajectories: Advance the understanding of temporal trajectories of biogeochemical impacts of wildfires within river corridors and their relationship to precipitation.
- Develop PyOM reaction networks: Incorporate the processing of pyrogenic material into reaction network models.
- Relate pyrogenic impacts to watershed features: Lay a foundation for understanding dynamic relationships between pyrogenic materials and river corridor biogeochemistry across variation in watershed features (e.g., stream order and discharge, upland soil physical properties, evapotranspiration, slope of surrounding landscape, and burn area).

In accordance with reviews and BER guidance, the overall objectives for RC-3 have changed, resulting in a reduced budget by ~40%. *These changes have altered the RC-3 scope*, and those alterations are detailed below and throughout the document.

> "In response to reviewer concerns, we have made three major changes to RC-3 activities: 1) a shift in focus towards lab-based activities that leverage our expertise in biogeochemistry; 2) a reduction in field efforts to focus only on in-stream water column and sediment processes relating to two wildfires -- the Evans Canyon Fire and Cold Creek Fire […]; and 3) a large reduction of investigations into wildfire-precipitation impacts and rapid response efforts as a major focus of RC-3."
>
> – Response to Reviewer Comments Response to BER Submitted Oct 12, 2020

These impact the overall objectives in the following ways:

- Identify impacts of burn severity: No Change.
- Derive PyOM indicators: No Change.
- Understand temporal trajectories: Smaller spatial focus; emphasis on recent fires (Evans Canyon Fire, Cold Creek Fire, and Holiday Farm Fire).
- Develop PyOM reaction networks: No Change.
- Relate pyrogenic impacts to watershed features: Smaller spatial focus; emphasis on recent fires (Evans Canyon Fire, Cold Creek Fire, and Holiday Farm Fire).

Key Contributions of this Research Campaign to the SFA: The River Corridor SFA integrates understanding on the controls of spatial and temporal variations in river corridor hydrobiogeochemistry, its responses to disturbances, and representation in numerical models from reaction to basin scales. To robustly predict changes in watershed function in response to wildfire disturbances, and how those changes will affect water quality and ecosystem health, RC-3 is working on developing a mechanistic understanding of governing processes on the impacts of fires on river corridor hydrobiogeochemistry and will work with RC-1 to further improve model predictive capacity in watersheds impacted by fire disturbances, important for ascertaining the impact of fire on ecosystem structure and function.

FY21 Science Plan Milestone: Determine the chemical composition of burned substrates, derive indicators for PyOM, and begin laboratory experiments to determine their relationship to biogeochemical rates. Collect distributed samples for chemical characterization and analysis of PyOM indicators. Determine in situ biogeochemical rates. Begin determination of watershed features using a combination of field samples and remote sensing to be completed in FY22. Establish wildfire chronosequences and reference sites. Establish protocols and sampling kits for rapid response teams.

In accordance with the changes in budget and scope, the updated milestones for FY21 are as follows: Design and implement burn severity experiment to establish how signatures of PyOM relate to varying burning conditions/severities. Method development and analysis for chemical characterization of PyC, PyP, and PyN. Establish field protocol and sites – downstream of burn, in-burn and out-of-burn reference sites. Collect timeseries and storm samples for chemical characterization and analysis of biogeochemical parameters and PyOM indicators. Contextualize RC-3 efforts within the literature and SFA.

Progress Brief for FY21: Our activities in FY21 are focused on establishing the chemical composition of burned substrates in the context of burn severity, developing workflows and field protocols, and establishing routine sampling of field sites.

- Identify impacts of burn severity: We have designed and implemented a large-scale burn severity experiment in collaboration with Kevin Bladon and John Bailey at OSU. These experiments are one of the first of their kind, focused on recreating fire severity conditions that will result in burn severities observed in the natural environment. Fire burn severities are not well mapped to resultant PyOM creation, due in large to the lack of links between experimental work and empirical/*in-situ* observations. Charting such burn severity metrics to combustion conditions/fire type *and* the resultant impact on chemistry is difficult (Vega et al., 2013). For example, higher max temperatures with much shorter heating duration times result in greatest PyOM changes (e.g. Tmax 600˚C vs 300-500˚C) experienced during experimental forest burns compared to laboratory-based tests; open flaming fires with shorter duration are more realistic to wildfire conditions, despite most literature to characterize PyOM focuses on slow pyrolysis techniques more representative of biochar creation (Santin et al., 2017; Santin et al., 2015). We have explicitly designed this experiment to tackle these knowledge gaps and combined pre-modeling exercises to assess burn conditions to target those that will produce the most divergent modeled biogeochemical rates (see publication highlight below). We designed the vegetation to use in this experiment based on remote sensing of Yakima River Basin dominant vegetation coverage, in collaboration with MM and RC-1 staff. Finally, we have completed a proof-of-concept exercise for PyOM characterization via FT-ICR-MS, Orbitrap LC-MS, and ^{1}H NMR on burnt and unburnt hemlock vegetation. Full characterization on the chars will be completed early FY 22 and will provide information and material for several FY22 activities.
- Derive PyOM indicators: The development of methods to assess PyOM indicators, relevant for RC-1 modeling efforts, are well underway. In addition to the proof-of-concept of PyOM chemistry assessment detailed above, we have submitted an EMSL user proposal in March of 2021 for FY-22 that would enhance our assessment of PyOM indicators in an untargeted fashion, using Orbitrap LC-MS, ^{1}H and ^{31}P NMR, and FT-ICR-MS. PyP studies will also be enhanced with synchrotron experiments, with proposals being prepped to submit in fall 2020. At MCRL, development of the benzene polycarboxylic acid (BPCA) method (modification of (Wagner et al., 2017a)) to be used as PyOM indicator is underway. This well-established suite of PyOC indicators will be utilized with burn severity experiment samples and timeseries samples (detailed below). At MCRL the development of a workflow for analysis of PyN is ongoing, while at Richland campus the development of a workflow for analysis of PyP (organic and inorganic species) is also ongoing.
- Understand temporal trajectories: Main RC-3 sites in the Yakima River Basin are set up and have been sampled for surface waters routinely since December 2020. We are targeting Umtanum Creek

and Wenas Creek within the Evans Canyon Fire burn perimeter, and Cold Creek impacted by the Clear Creek Fire. Each stream is sampled inside the burn perimeter or the burn-impacted zone, downstream of the fire burn perimeter, and upstream of the burn perimeter, when/as sites are accessible. Routine parameters of analysis are complementary with those measured in RC-2, and the *in-situ* sondes purchased were done in complete consultation with RC-2's in-stream infrastructure design. Permitting for these deployments is underway.

- Develop PyOM reaction networks: This fiscal year, we have focused our PyOM reaction network work on literature assessment of PyOM reactivities, to be compared to empirical results from ongoing laboratory and field-based activities. A description of our modeling efforts can be found in the publication highlight.
- Relate pyrogenic impacts to watershed features: We have established a collaboration with the Eugene Water and Electric Board (EWEB) and Kevin Blandon at OSU in the McKenzie River watershed of the Willamette River Basin. The collaborators at EWEB have a long history and track record of capturing storm events for their water quality monitoring efforts. Many of the watersheds relevant to their drinking water program were impacted by the Holiday Farm Fire in 2020. We are analyzing these high-resolution temporal samples (1/hr over the first 24 hrs during a storm after the fire event) using advanced organic matter characterization techniques (to be submitted Sept 2021; (Roebuck et al., 2021)). This details our team's first collaborative effort to relate fire impacts on river corridor biogeochemistry and watershed features. We are continuing this collaboration in FY 22.
- Contextualize RC-3 efforts within the literature and SFA: As a new element of the River Corridor SFA, we have been conducting several efforts aimed at integrating the RC-3 campaign into both the scientific community and the greater RC SFA. The team is currently leading a community perspective paper, focused on the gaps in the community's knowledge on the biogeochemical impacts of fires on the river corridor (Myers-Pigg et al., *in prep*; sensu (Graham et al., 2021)), to be submitted early FY22, and team members are part of a dissolved PyOM in watersheds review effort (Abney et al., *in prep*) and a book chapter on OM dynamics (Moreland et al., 2021). RC-3 is also working closely with RC-2 on field site instrumentation choices.

Publication Highlight: Inferred bioavailability of pyrogenic organic matter compared to natural organic matter from global sediments and surface waters. Our team has recently completed a modeling experiment to evaluate PyOM bioavailability, leveraging a model that uses OM stoichiometry to predict aerobic respiration as part of the River Corridor SFA (Song et al., 2020). The predicted bioavailability for a diversity of PyOM compounds mined from the literature is comparable to natural OM pools in global surface waters and sediments (Figure 5). However, the model-derived carbon use efficiency (CUE) of PyOM varied dramatically, indicating a large range in PyOM's impact on ecosystem function. For instance, phenolic PyOM and 'Black Carbon' ('BC') molecules (defined here as condensed aromatic core structures poly-substituted with O-containing functionalities (Wagner et al., 2017b) had lower metabolic efficiency than other PyOM and NOM compounds, and 'BC' metabolism was also less negatively impacted by oxygen limitation. This work supports current theories that PyOM is more bioavailable than traditionally accepted, (e.g., (Bird et al., 2015)), supporting building evidence (e.g. (Myers-Pigg et al., 2015)) that it may be an underappreciated driver of river corridor biogeochemistry.

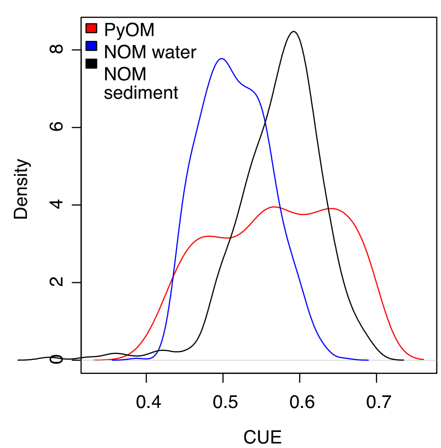


Figure 5. Carbon use efficiency (CUE) distribution in PyOM (red) and ubiquitous surface water (blue) and sediment (black) NOM molecules. Figure modified from: Graham et al., 2021.

Reference: Graham, E.B. H.-S. Song, S. Grieger, V.A. Garayburu-Caruso, J.C. Stegen, K.D. Bladon, A. N. Myers-Pigg (2021) "Inferred bioavailability of pyrogenic organic matter compared to natural organic matter from global sediments and surface waters." *EcoEvoRxiv*. DOI:10.32942/osf.io/6j9t5. In review at *Environmental Science & Technology Letters.*

Plans for FY22

In FY 22, we will continue to establish the chemical composition of burned substrates in the context of burn severity, focusing more deeply on PyOP and PyON. We will design and implement biogeochemical reactivity experiments of PyOM to compare reaction network model results of MM to RC-3's empirical laboratory and field-based measures of PyOM reactivity. We will use the process-based understanding generated herein to extend the suite of biogeochemical reactions considered in thermodynamic (Song et al., 2020) and reactive transport models (PFLOTRAN and SWAT-MRMT-R (Fang et al., 2020)) in collaboration with MM and RC-1, to further improve model predictive capacity in watersheds impacted by fire disturbances. We will increase the temporal nature of YRB sampling, based on RC-1 FY 21-22 modeling results, and inform RC-1 models of biogeochemically relevant PyOM chemistries.

Research Campaign 4 (RC-4): Multi-Basin Studies

Overall Objectives: *The Multi-Basin Studies Campaign aims to provide transferable principles that integrate DOM chemistry, microbial gene expression, biogeochemical function, and disturbance by combining existing global-scale data with WHONDRS-based data generation and numerical modeling distributed across CONUS basins.*

- Expand WHONDRS: Continue to develop WHONDRS as a resource to pursue community-based data analysis, interpretation, and publishing (FY21–24).
- Inform models: Couple DOM chemistry, microbial gene expression, and aerobic respiration in surface and pore water across globally distributed river corridors to inform models in RC-1,2,3 (FY21-22).
- Compare to predictions: Implement ICON-ModEx via crowdsourced sampling designed to test model predictions across CONUS basins and use outcomes to inform further model development (FY22-23).
- Characterize variation: Build on RC-2.2 with multiscale characterization of DOM chemistry and microbial gene expression from reach to multi-basin scales across the CONUS (FY23–24).

Key Contributions of this Research Campaign to the SFA: The RCSFA aims to develop predictive capabilities and a fundamental understanding of river corridor hydrobiogeochemistry that apply across basins. This is vital to informing large-scale models that span diverse basins within and beyond CONUS. RC4 is a key element of this long-term vision by developing transferable knowledge and models. Much of the scientific focus in RC4 builds from RC-1,2,3 while making more direct use of WHONDRS and expanding the community-based research model that WHONDRS embodies. RC4 is making use of existing WHONDRS data and growing WHONDRS as a community resource. For example, RC4 is building the Genome Resolved Open Watersheds (GROW) database. GROW will be an open resource for the community and used by the RCSFA to develop genome-informed mechanistic and reduced order models needed by RC1. In addition, RC4-generated data are being used by RC3 to contextual PyOM chemistry against a backdrop of natural OM. RC4 is also engaged in numerous crowdsourced sampling campaigns around the globe to generate transferable understanding of factors governing variation in OM chemistry and the biogeochemical impacts of that variation, ultimately to inform the RCSFA's mechanistic and AI models.

FY21 Science Plan Milestone: Using data from previous WHONDRS efforts to couple OM chemistry, microbial gene expression, and aerobic respiration across globally distributed river corridors; provide outcomes to models in RC-1,2,3. Analyze existing data on the transient metabolome to reveal influences of historical disturbances. Conduct community-enabled WHONDRS sample collection across diverse river corridor environments with more community ownership.

Progress Brief for FY21: Our activities in FY21 are focused on leveraging existing and new WHONDRS data to understand and model the global coupling among microbes, OM chemistry, biogeochemical function, and environmental features in river corridors while shifting the sampling paradigm to efforts guided more directly by the community. Note that the RC4 scope was modified following BER feedback, as detailed in the RCSFA's response to review document delivered to BER in late FY20.

- One focal area for RC4 in FY21 has been the development of sampling campaigns in which the collaborators taking the samples have more influence over the design and goals of the study. This is also commensurate with sampling campaigns that are each smaller scale than previous global efforts. This is called WHONDRS-Local, which has multiple instances in FY21, and emphasizes environmental contrasts through space and time. For example, samples are being taken through time in a small agricultural stream system in the Columbia River Basin across different hydrologic conditions. This adds to an existing research program studying stream hydrobiogeochemistry. In another instance, RC4 has partnered with EXCHANGE (the WHONDRS-like part of COMPASS) and the University of Quebec to study OM chemistry from source to sea along the St. Lawrence River. This effort emphasizes numerous environmental gradients and leverages a long-term research program. RC4 has also partnered with the LBNL Watershed Function SFA to support investigations of OM chemistry near the East River. There are a number of smaller scale efforts as well, with an emphasis on expanding to additional continents, with RC4 recently receiving the first sample from Africa.
- RC4 also continues to generate data from existing and new samples, while also publishing those data on ESS-DIVE. We have expanded data types from the 2019 global campaign to include elemental analysis of sediments, bulk C and N concentrations and both metagenomics and metatranscriptomics data via JGI. We submitted an EMSL proposal to further expand the data types from the 2019 campaign to small molecule metabolomics and metaproteomics. As data move through quality assurance they are added to ESS-DIVE data packages (e.g., ion data).
- RC4 has generated a number of science outcomes in FY21. For example, three recent papers used WHONDRS data and revealed new understanding of river corridor OM chemistry. Danczak et al. (2021) discovered a dynamic referred to as 'thermodynamic redundancy' whereby OM thermodynamics are relatively invariant despite significant changes in molecular composition of OM assemblages. This indicates that changes in the identity of organic molecules may not translate into functional differences associated with thermodynamics. Mueller et al. (2021) is the first paper using WHONDRS data that was not led by RCSFA team members. This study showed a strong influence of natural OM over transformations of organic contaminants in a wastewater influenced stream. Garayburu-Caruso et al. (2020) provided a global summary of OM chemogeography, with comparative analyses between surface water and sediments. Among many patterns, they found continental-scale gradients in OM chemistry within the CONUS, indicating large scale drivers of key chemical properties.
- Multiple RC4 studies are expected to be completed in FY21. For example, we are reviewing the state of river corridor microbiome science. This effort is revealing the primary gaps both in terms of where knowledge is missing and also the kinds of knowledge missing, in particular with respect to expressed microbial function in context of watershed disturbances such as fire and stream intermittency. The first major installment of the GROW database will be completed in FY21, with a focus on surface water microbial communities from global rivers. These data go beyond WHONDRS and include efforts from over 10 non-RCSFA investigators, which significantly expands the environmental breadth of the database. A manuscript using the database to summarize the biogeography of microbial metabolism is expected in FY21. Additional efforts span multiple manuscripts that leverage existing data, such as the integration of OM chemistry with microbial metabolic potential across global rivers (using WHONDRS 2018 data), development of a trait-based framework to study OM thermodynamics (collaboration with Stegen Early Career project using WHONDRS 2019 data), evaluation of microbial metabolic potential dynamics during dynamic hydrologic perturbations (using

WHONDRS 2018 data), using global natural OM data to understand PyOM chemistry (led by RC3 using WHONDRS 2019 data), and development of a pipeline from omics to reactive transport modeling (led by Subsurface Insights using RCSFA data).

- RC4 is continuing to change paradigm of open science through a globally crowdsourced effort using existing WHONDRS data to study OM chemistry. Going beyond our previous efforts in crowdsourced publishing, this current effort spans the entire research life cycle from idea generation to data analysis to interpretation and manuscript development. This effort was initially envisioned to result in one crowdsourced manuscript, but we held an initial, open workshop and the high level of interest made it clear that this effort will instead generate an entire special issue of crowdsourced manuscripts all using WHONDRS data. This effort is actively progressing with >100 participants from around the world.
- Community outreach is an important part of RC4, and there have been numerous associated efforts. One significant effort is launching the ICON Science institute, in partnership with EXCHANGE and likely other partners. The institute will enable the use of ICON principles across all domains of science and is poised to launch in summer 2021. We also engage in more traditional ways such as at AGU in December 2020 we held two town halls, had a virtual booth, and organized a river corridor session. In addition to numerous contributed presentations, we have given multiple of invited presentations such as at the recent ESS-DIVE community workshop and at a NEON workshop. Another example is RC4's contributions to a recent ASCR proposal on digital twinning at the basin-scale, in which we provided key elements of community engagement and data generation.

Publication Highlight: Using Community Science to Reveal the Global Chemogeography of River Metabolomes. OM transformations in aquatic ecosystems are a critical source of uncertainty in global biogeochemical cycles. Environmental metabolomics, or the analysis of organic molecules in environmental samples, help characterize the interactions of organisms within their environment. WHONDRS reported the first ultrahigh-resolution analysis of global river corridor metabolomes of both surface water and sediment across 97 global rivers spanning a wide range of sizes and ecosystem types (Figure 6). The publication describes the distribution of key chemical attributes of metabolomes, including East-West gradients, in many metabolomic features across the contiguous United States. It also shows that surface water metabolomes are more diverse than those in sediment, possibly suggesting a greater diversity of biological processes occurring in surface waters. This work provides a foundation for understanding global patterns in river corridor biogeochemical cycles. This publication was invited to be part of a Special Issue in the journal *Metabolites* on the Metabolome and Fluxomics.

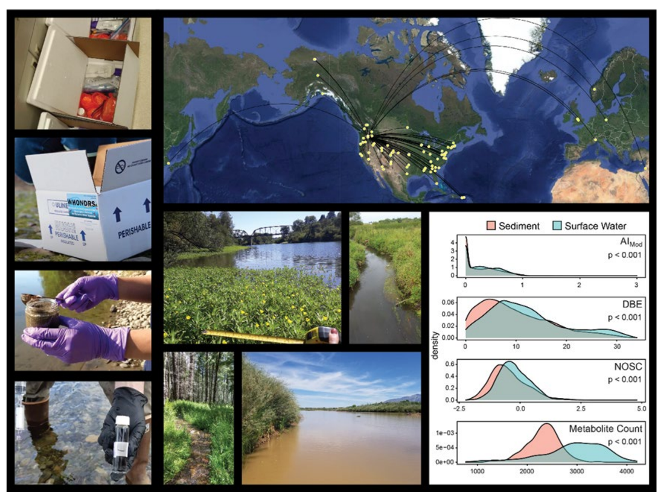


Figure 6. A global WHONDRS effort in 2019 used FTICR-MS to study metabolomes in surface waters and sediments across biomes.

Reference: V.A. Garayburu-Caruso, et al. (2020) "Using Community Science to Reveal the Global Chemogeography of River Metabolomes." *Metabolites.* DOI: 10.3390/metabo10120518.

Plans for FY22: In FY22 RC4 will focus on five primary goals. (1) Generating data and publishing those data on ESS-DIVE from samples collected in FY21. (2) Continue to advance the crowdsourced special collection using WHONDRS data. (3) Implement a crowdsourced sampling campaign based on the integration of ICON principles with a ModEx approach. This ICON-ModEx effort will focus on CONUS riverbed sediment sampling to test and enhance a suite of AI models developed via an SBIR partnership with ParallelWorks. (4) Continue building the GROW database by expanding it to include sediment-associated microbial data, while also facilitating the community to use GROW to generate manuscripts not led by the RCSFA. (5) Further develop mechanistic and AI models connecting microbes, OM chemistry, biogeochemistry, and hydrology for use in large scale models. The modeling work will be done in part via a new informal collaboration with LBNL and the international critical zone network of networks.

Multiscale ModEx (MM):

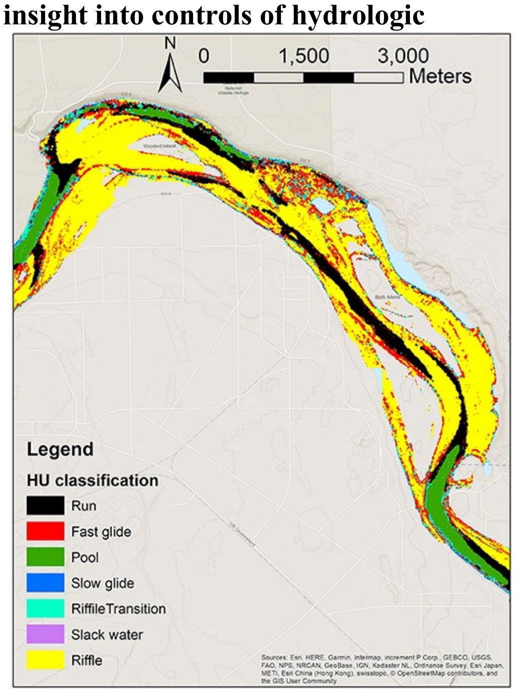
Overall Objectives: *The Multiscale ModEx crosscutting activity aims to coordinate model–data integration across RCs and scales, assuring optimal use of data and models and open distribution of SFA products.*

- Steward interactive data–model integration (ModEx) across scales and all RCs.
- Implement the Data Management Plan (DMP) and Software Productivity and Sustainability Plan (SW-PSP); provide tools for accessing data and software products.

Progress Brief for FY21

Our activities in FY21 are focused on 1) development of comprehensive nitrogen reaction networks to represent N cycling in river corridor and watershed models; 2) application of AI/ML approaches to develop transferable understanding of river corridor hydrobiogeochemistry (e.g., (Ren et al., 2021) and see highlight below); 3) implementation of our Data Management and Software Sustainability plans; and 4) development and application of fundamental river hydrodynamics simulations across scales (e.g., (Chen et al., 2020; Song et al., 2021)). This activity does not have specific associated milestones.

Publication Highlight: Machine learning analysis provides insight into controls of hydrologic exchange flows and travel time distributions: Two machine learning (ML) approaches, the Random Forest (RF) and Extreme Gradient Boosting (XGB) models, were applied to the outputs from reach-scale transient coupled surface/subsurface flow models. The objective was to improve understanding of the influences of physical, spatial and temporal factors on hydrologic exchange flows (HEFs) and subsurface transit time distributions (TTDs). The ML models achieved 70-80% predictive accuracy based on key controlling parameters: Thickness of the Hanford formation, flow regime, river velocity, and river depth. The ML models, based in part on hydromorphic classes (see figure to the right), provide a surrogate predictive capability that can be employed in other mainstem Columbia River reaches where computationally demanding coupled surface-subsurface flow simulations are not available.



Reference: Ren, H., X. Song, Y. Fang, Z. J. Hou, and T. D. Scheibe. Machine learning analysis of hydrologic exchange flows and transit time distributions in a large regulated river." 2021. *Frontiers in Artificial Intelligence.* doi: 10.3389/frai.2021.648071.

Publication Analysis

51 peer-reviewed journal articles and one peer-reviewed book chapter have been published or are in press during the current triennial period (2020 and 2021 to date - see Appendix A for a complete list), and an

additional 25 manuscripts have been submitted as of the writing of this document. As shown in Figure 4, this is closely comparable to the number of publications in the previous period for the last triennial cycle (2017-2019), and represents a significant increase over the period 2014-2017. Over the current period, the SFA published most frequently in a new journal, *Frontiers in Water* (8 papers), in part because of a special issue on riparian corridor biogeochemistry co-edited by the PNNL and LBNL ESS SFAs (D. Dwivedi and T. Scheibe). The second most papers (4) were published in *Water Resources Research*. Outside of these two journals, the remaining papers were published in 31 different journals, reflecting the diversity of SFA research and the wide audience reached by our publications. The SFA publishes in high-quality journals: Three publications are in DOE-designated high-impact journals (one in *Nature*, two in *Nature Communications*), 36% of the publications are in ISI-designated top ten journals in their respective fields, approximately three-quarters (72%) are in first quartile journals, and nearly all (97%) are in journals with impact factors above the median in their field. The average impact factor of the journals for which data are available (weighted by the number of SFA papers published in each) is 5.71, a 27% increase from the value of 4.49 reported in 2019.

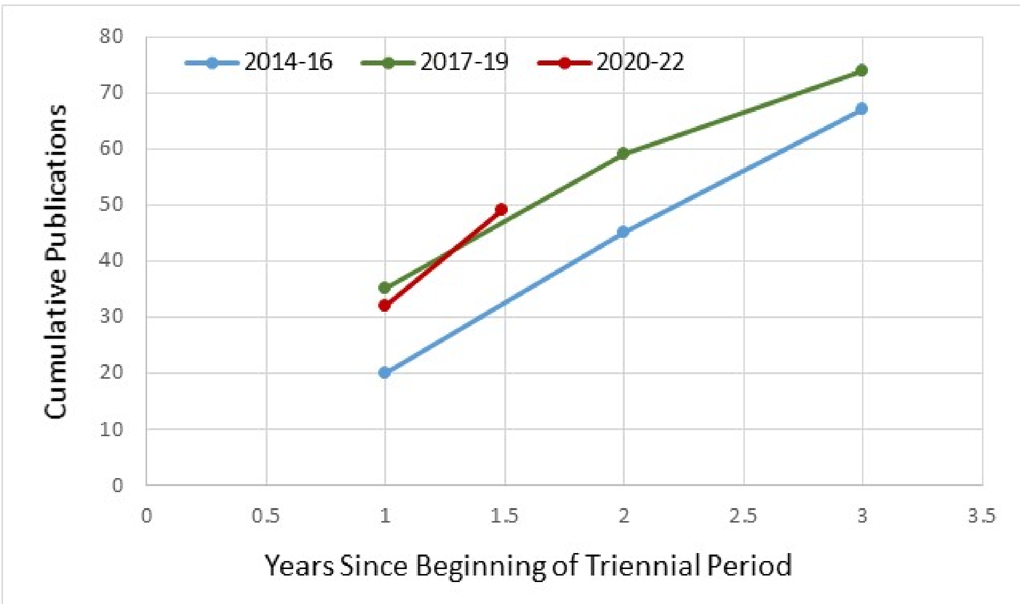


Figure 7. Comparison of number of publications to date in the current triennial period (2020-2022) with number of publications in the two previous triennial periods (2014-16 and 2017-19). Note that the point for 2021 is plotted at the elapsed time corresponding to late June when these statistics were compiled and does not include manuscripts that are in press or in review.

Future Scientific Goals, Vision, and Plans for Meeting Program Objectives

The Environmental System Science program examines complex ecological and hydro-biogeochemical processes within terrestrial and coastal systems to understand inherent and emergent properties of change to Earth and environmental system.[1] The long-term vision of the SFA, closely aligned with ESS objectives, will culminate in transferable understanding of critical hydrologic, biogeochemical, and ecological processes in river corridors, and in the creation of a new broadly applicable simulation capabilities, linked with other watershed hydrobiogeochemical system component models, to provide predictive understanding of watershed function and response to disturbances. In the current performance period (FY21-24), we are expanding our research to study the Yakima River Basin, which encompasses a broad range of physiographic and hydrobiogeochemical environments. We continue to strengthen our broad collaborations, e.g. through WHONDRS and other efforts, to support generalization of new

[1] https://ess.science.energy.gov/

knowledge, as embodied in numerical models, across the Columbia River Basin and the CONUS, and integration of those models with major community Earth system model frameworks such as E3SM and the National Water Model.

New Scientific Results that may Shift Current Research or Motivating Knowledge Gaps

FY21 is the first year of the current performance period for which our research plan was reviewed in 2020. The plan is currently being implemented largely as written, with the exception of modifications made in response to reviewer comments and BER guidance[1]. No near-term major changes in direction or activity plans have been identified at the current time. We intend to conduct a modest stocktaking workshop, including some external participants, near the beginning of FY22 to develop more detailed plans and timeline for FY22-24 and make decisions regarding potential expanded geographical scope based on preliminary results.

Collaborative Research Activities

The SFA is dedicated to implementing principles of open, collaborative, integrated team science. We have dramatically increased our collaborative footprint over the past four years through a number of avenues including (1) direct-funded (by subcontract) collaborations with universities, other national laboratories, and other federal agencies (e.g., USGS); (2) collaborative projects funded through SFA FOAs to university partners; (3) collaborative projects funded through the SBIR program; and (4) extensive community-oriented science activities such as the WHONDRS network. Lists of collaborative projects with co-PI names, institutions, and titles are provided in Appendix C.

SFA team members are also actively leading and involved in a number of community-level cross-cutting activities:

- Workshop Leadership: The SFA hosted the 2019 Watershed Community Workshop in Richland WA on Sept. 10-12, 2019 and the AI4ESP Ideation Workshop held virtually on Jan. 15, 2021. Tim Scheibe participated in the C-IHTM workshop held in November 2020.
- ESS Cyberinfrastructure Working Groups: Co-PI Xingyuan Chen represents the SFA on the Executive Committee and co-leads the Data-Model Integration working group, and other SFA team members participate in working group activities.
- ESS-Dive: The SFA has worked closely with the ESS-Dive team to apply new capabilities (e.g. ESS-DIVE portals[2]) and to develop sample identifiers (Damerow et al., 2021) and community data and metadata standards. SFA team member Amy Goldman led a community data standards project funded by ESS-DIVE to create standards for hydrological monitoring data[3].
- KBase and EMSL: SFA collaborator Hyun-Seob Song is working closely with KBase staff (Chris Henry of ANL) to develop new workflows to integrate multi-omics data into microbial community reaction networks. These networks are then formulated into elemental reaction pathways that can be incorporated into reactive transport simulations using the PFLOTRAN reaction sandbox. Tim Scheibe organized the 2020 EMSL Summer School on Multiscale Microbial Dynamics Modeling[4], which was co-sponsored by the SFA, EMSL and KBase. Over 500 community participants learned about the KBase-PFLOTRAN modeling pipeline as applied to WHONDRS datasets.
- IDEAS: Co-PI Xingyuan Chen is the PNNL lead for the IDEAS-Watershed collaborative project. Under this project the SFA is co-funding a post-doctoral associate working on development of community cyberinfrastructure.

[1] PNNL SBR SFA Reponses to the Triennial Review Comments, Transmitted to BER October 12, 2020,
[2] https://data.ess-dive.lbl.gov/portals/WHONDRS
[3] https://ess-dive.lbl.gov/community-projects/
[4] https://www.kbase.us/multiscale-microbial-dynamics-modeling/

- <u>ExaSheds:</u> Co-PI Xingyuan Chen is the PNNL lead for the ExaSheds project, which is developing next-generation high-performance watershed modeling capabilities that are integrated with and informed by artificial intelligence methodologies.
- <u>WHONDRS Network:</u> The ***Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems*** (WHONDRS) is a global consortium of researchers and other interested parties that aims to understand coupled hydrologic, biogeochemical, and microbial function, from local to global scales, within river corridors experiencing recurring, episodic, or chronic hydrologic perturbations. WHONDRS is coordinated by the SFA, and information on current activities is described above under RC-4. WHONDRS is linked with the Genome-Resolved Open Watershed (GROW[1]) network led by Kelly Wrighton (Colo. State Univ.).
- <u>Crowdsourced Paper:</u> Kayla Borton is leading a globally crowdsourced effort that uses ecological principles to better understand organic matter chemistry in rivers. This was initiated by a workshop in April 2021 that generated a number of science questions that could be addressed using the WHONDRS organic matter (OM) chemistry dataset. This crowdsourced effort is expected to lead to a collection of manuscripts published as a journal special issue.
- <u>ICON/FAIR Collection (Goldman):</u> The SFA is organizing a special collection hosted by the AGU open access journal *Earth and Space Science* (Goldman et al., 2021). The collection, entitled "*The Power of Many: Opportunities and Challenges of Integrated, Coordinated, Open, and Networked (ICON) Science to Advance Geosciences,*" will comprise commentary articles representing different geoscience disciplines as represented by up to 25 different AGU sections.
- <u>PyOM Community Review Paper:</u> Allison Myers-Pigg is leading the development of a community review paper on the topic of "*Shaping the future of wildfire science: Top priorities and unanswered questions on the cascading watershed impacts of fires.*" The development of the manuscript will engage a broad spectrum of researchers using a crowdsourced approach.

[1] https://jgi.doe.gov/csp-2020-microbial-genomes-across-world-rivers/

APPENDIX A: PROJECT PUBLICATIONS

2021

Published:

1. Conner A., Gooseff M. N., Chen X., Arntzen E. and Garayburu-Caruso V. (2021) Groundwater inflows to the Columbia River along the Hanford reach and associated nitrate concentrations. *Frontiers in Water,* **3**; doi: 10.3389/frwa.2021.574684.
2. Damerow J., Varadharajan C., Boye K., Brodie E. L., Burrus M., Chadwick D., Crystal-Ornelas R., Elbashandy H., Eloy Alves R., Ely K., Goldman A., Haberman T., Hendrix V., Kakalia Z., Kemner K. M., Kersting A., Merino N., O'Brien F., Perzan Z., Robles E., Sorensen P., Stegen J., Walls R., Weisenhorn P., Zavarin M. and Agarwal D. (2021) Sample identifiers and metadata to support data management and reuse in multidisciplinary ecosystem sciences. *Data Science Journal,* **20**, 1-19; doi: 10.5334/dsj-2021-011
3. Danczak R. E., Goldman A. E., Chu R. K., Toyoda J. G., Garayburu-Caruso V. A., Tolic N., Graham E. B., Morad J. W., Renteria L., Wells J. R., Herzog S. P., Ward A. S. and Stegen J. C. (2021) Ecological theory applied to environmental metabolomes reveals compositional divergence despite conserved molecular properties. *Science of the Total Environment,* **788**, 147409; doi: 10.1016/j.scitotenv.2021.147409.
4. Dwivedi D., Godsey S. E. and Scheibe T. D. (2021) Editorial: Linking hydrological and biogeochemical processes in riparian corridors. *Frontiers in Water,* **3**; doi: 10.3389/frwa.2021.693763.
5. Fudyma J. D., Chu R., Grachet N. G., Stegen J. and Tfaily M. (2021) Coupled biotic-abiotic processes control biogeochemical cycling of dissolved organic matter in the Columbia River hyporheic zone. *Frontiers in Water,* **2**, 574692; doi: 10.3389/frwa.2020.574692.
6. Graham E. B., Averill C., Bond-Lamberty B., Knelman J. E., Krause S., Peralta A. L., Shade A., Smith A. P., Cheng S. J., Fanin N., Freund C., Garcia P. E., Gibbons S. M., Van Goethem M. W., Guebila M. B., Kemppinen J., Nowicki R. J., Pausas J. G., Reed S. P., Rocca J., Sengupta A., Sihi D., Simonin M., Słowiński M., Spawn S. A., Sutherland I., Tonkin J. D., Wisnoski N. I., Zipper S. C., C. C., Staal A., Arora B., Oldfield C., Dwivedi D., Larson E., Santillan E., Aaron Hogan J., Atkins J., Zheng J., Lembrechts J., Patel K., Copes-Gerbitz K., Winker K., Mudge L., Wong M., Nuñez M., Luoto M. and Barnes R. (2021) Toward a generalizable framework of disturbance ecology through crowdsourced science. *Frontiers in Ecology and Evolution,* **9**, 588940; doi: 10.3389/fevo.2021.588940.
7. Kaufman M. H., Warden J. G., Cardenas M. B., Stegen J. C., Graham E. B. and Brown J. (2021) Evaluating a laboratory flume microbiome as a window into natural riverbed biogeochemistry. *Frontiers in Water,* **3**, 596260; doi: 10.3389/frwa.2021.596260.
8. Missik J. E. C., Liu H., Gao Z., Huang M., Chen X., Arntzen E., McFarland D. P. and Verbeke B. (2021) Groundwater regulates interannual variations in evapotranspiration in a riparian semiarid ecosystem. *Journal of Geophysical Research: Atmospheres,* **126**, e2020JD033078; doi: 10.1029/2020jd033078.
9. Mueller B. M., Schulz H., Danczak R. E., Putschew A. and Lewandowski J. (2021) Simultaneous attenuation of trace organics and change in organic matter composition in the hyporheic zone of urban streams. *Scientific Reports,* **11**, 4179; doi: 10.1038/s41598-021-83750-8.
10. Nayfach S., Roux S., Seshadri R., Udwary D., Varghese N., Schulz F., Wu D. Y., Paez-Espino D., Chen I. M., Huntemann M., Palaniappan K., Ladau J., Mukherjee S., Reddy T. B. K., Nielsen T., Kirton E., Faria J. P., Edirisinghe J. N., Henry C. S., Jungbluth S. P., Chivian D., Dehal P., Wood-Charlson E. M., Arkin A. P., Tringe S. G., Visel A., Woyke T., Mouncey N. J., Ivanova N. N., Kyrpides N. C., Eloe-Fadrosh E. A. and Consortium I. M. D. (2021) A genomic catalog of Earth's

microbiomes (Nov, 10.1038/s41587-020-0718-6, 2020). *Nature Biotechnology,* **39**, 521-521; doi: 10.1038/s41587-021-00898-4.
11. Ren H., Song X., Fang Y., Hou Z. J. and Scheibe T. D. (2021) Machine learning analysis of hydrologic exchange flows and transit time distributions in a large regulated river. *Frontiers in Artificial Intelligence,* **4**; doi: 10.3389/frai.2021.648071.
12. Song H.-S., Stegen J. C., Graham E. B. and Scheibe T. D. (2021) Historical contingency in microbial resilience to hydrologic perturbations. *Frontiers in Water,* **3**; doi: 10.3389/frwa.2021.590378.
13. Song X., Fang Y., Bao J., Ren H., Duan Z., Perkins W., Zhou H., Hou Z., Chen Y. and Scheibe T. (2021) Scale-dependent spatial variabilities of hydrological exchange flows and transit time in a large regulated river. *Journal of Hydrology,* **598**, 126283; doi: 10.1016/j.jhydrol.2021.126283.
14. Wu L., Gomez-Velez J. D., Krause S., Wörman A., Singh T., Nützmann G. and Lewandowski J. (2021) How daily groundwater table drawdown affects the diel rhythm of hyporheic exchange. *Hydrology and Earth System Sciences,* **25**, 1905-1921; doi: 10.5194/hess-25-1905-2021.
15. Wu R., Chen X., Hammond G., Bisht G., Song X., Huang M., Niu G.-Y. and Ferre T. (2021) Coupling surface flow with high-performance subsurface reactive flow and transport code PFLOTRAN. *Environmental Modelling & Software,* **137**, 104959; doi: 10.1016/j.envsoft.2021.104959.
16. Xiao Y., Sloan J., Hepworth C., Osborne C. P., Fleming A. J., Chen X. and Zhu X. G. (2021) Estimating uncertainty: A Bayesian approach to modelling photosynthesis in C3 leaves. *Plant Cell & Environment,* **44**, 1436-1450; doi: 10.1111/pce.13995.
17. Zhang Y., Zhou D., Wei W., Frame J. M., Sun H., Sun A. Y. and Chen X. (2021) Hierarchical Fractional Advection-Dispersion Equation (FADE) to quantify anomalous transport in river corridor over a broad spectrum of scales: Theory and applications. *Mathematics,* **9**, 790; doi: 10.3390/math9070790.

Accepted:

1. Arora B., Briggs M., Gomez-Velez J., Stegen J. C. and Zarnetske J. (2021) Hot spots and hot moments in the critical zone: Identification of and incorporation into reactive transport models. In *Biogeochemistry of the Critical Zone*, S. Banwart (Ed.), Springer-Nature., Accepted.
2. Chen X., Lee R. M., Dwivedi D., Son K., Fang Y., Zhang X., Graham E., Stegen J., Fisher J. B., Moulton D. and Scheibe T. D. (2020) Integrating field observations and process-based modeling to predict watershed water quality under environmental perturbations. *Journal of Hydrology*, Accepted; 125762; doi: 10.1016/j.jhydrol.2020.125762.
3. Dove N. C., Veach A. M., Muchero W., Wahl T., Stegen J., Schadt C. W. and Cregger M. (2021) Assembly of the Populus microbiome is temporally dynamic and determined by selective and stochastic factors. *mSphere*, Accepted, mSphere01316-01320.

Submitted:

1. Bao J., Chen Y., Fang Y., Song X., Perkins W., Duan Z., Ren H., Hou Z., Richmond M. C., He X. and Scheibe T. D. (2021) Use of numerical modeling to evaluate the impact of hydrodynamic pressure on hydrologic exchange fluxes and resident time for a large-scale river section over a long-term period. *Hydrological Processes*, In review; HYP-20-0169.
2. Bao Y., Feng Y., Stegen J., Wu M., Chen R., Liu W., Zhang J., Li Z. and Lin X. (2021) Integrating assembly processes, community patterns, and ecosystem function in the context of organic matter degradation. *Nature Ecology & Evolution*, Submitted; NATECOLEVOL-19067196.
3. Chen Y., Bao J., Fang Y., Song X., Perkins W., Duan Z., Ren H., Hou Z., Richmond M., He X. and Scheibe T. D. (2021) Modeling framework for evaluating the impacts of hydrodynamic pressure on

hydrologic exchange fluxes and residence time for a large-scale river section over a long-term period. *Environmental Modelling and Software*, Submitted.
4. Dwivedi D., Steefel C., Arora B., Banfield J. F., Bargar J., Boyanov M. I., Brooks S. C., Chen X., Hubbard S. S., Kaplan D., Kemner K. M., O'Loughlin E., Pierce E., Painter S., Scheibe T. D., Wainwright H. M., Williams K. and Zavarin M. (2021) From legacy contamination to watershed systems science: A review of scientific insights and technologies developed through doe-supported research in water and energy security. *Environmental Research Letters*, Submitted; 3.30.2021.
5. Fegel T., Boot C. M., Covino T., Elder K., Hall E. K., Starr B., Stegen J. C. and Rhoades C. (2021) Land cover change affects the amount and reactivity of DOM exported from old growth and regenerating forests in headwater ecosystems. *Hydrological Processes*.
6. Gomez-Velez J. D., Cardenas M. B., Chen X. and Scheibe T. D. (2021) Riverbed respiration produces a globally significant amount of CO_2. *Nature*, Submitted; 2020-2009-17406.
7. Graham E., Song H., Grieger S., Garayburu-Caruso V., Stegen J. C., Bladon K. D. and Myers-Pigg A. (2021) Inferred bioavailability of pyrogenic organic matter compared to natural organic matter from global sediments and surface waters. *Environmental Science & Technology Letters*, Submitted.
8. Graham E. B. and Smith A. P. (2021) Crowdsourcing global perspectives in ecology using social media. *EcoEvoRxiv*, ECY20-0500; doi: 10.32942/osf.io/pv3wk.
9. Hou Z., Ray J., Huang M., Bao J., Ren H. and Swiler L. (2021) Machine learning-based feature selection and surrogate-based MCMC-bayesian inversion for improving runoff simulations in the community land model. *Water Resources Research*, In revision.
10. Hou Z., Ren H., Murray C., Song X., Fang Y., Arntzen E., Chen X., Stegen J., Huang M., Gomez-Velez J., Duan Z., Perkins W., Richmond M. C. and Scheibe T. D. (2021) A novel construct for scaling groundwater-river interactions based on machine-guided hydromorphic classification. *Environmental Research Letters*, Revised manuscript submitted 6.24.21; ERL-111881.
11. Johnson T., Thomle J., Strickland C., Goldman A. and Stegen J. (2021) Riverbed temperature and 4D ERT monitoring reveals heterogenous horizontal and vertical groundwater-surface water exchange flows under dynamic stage conditions. *Water Resources Research*, Submitted 2.1.20; 2020WR027255.
12. Ju J., Dai H., Wu C., Hu B. X., Ye M., Chen X., Gui D., Liu H. and Zhang J. (2021) Quantifying the uncertainty of future climate change: an advanced hierarchical sensitivity analysis for the hydrological modeling under climate change. *Journal of Hydrometeorology*, Submitted; JHM-D-20-0003.
13. Kaufman M. H., Ghosh R., Grate J., Shooltz D., Freeman J., Ball T., Loloee R., McIntire C., Wells J., Strickland C., Vermeul V. R., Rod K. A., Mackley R. D., Lin X., Ren H. and Stegen J. (2021) Dissolved oxygen dynamics reveal biogeochemical tipping points driven by river corridor hydrology. *Groundwater*, Rejected 5/17/21.
14. Konrad C., Schmadel N., Harvey J., Schwarz G., Gomez-Velez J., Boyer E. and Scott D. (2021) Streamflow variability reduces in-channel travel time. *Water Resources Research*, In review; 2019WR025984.
15. Little C. J., Rizzuto M., Luhring T. M., Monk J. D., Nowicki R., Paseka R. E., Stegen J. C., Symons C. C., Taub F. B. and Yen J. D. (2021) Movement with meaning: Integrating information into meta-ecology. *Ecology Letters*, Rejected, in revision.
16. Liu H., Dai H., Niu J., Hu B. X., Qiu H., Gui D., Ye M., Chen X., Wu C., Zhang J. and Riley W. (2021) Hierarchical sensitivity analysis for large scale process-based hydrological modeling with application in an amazonian watershed. *Hydrology and Earth System Sciences*, In review, hess-2019-2246.
17. Moghaddam M., Ferre T., Chen X. and Chen K. (2021) Optimizing machine learning methods to infer surface water – ground water exchange with high temporal resolution temperature measurements. *Water*, Submitted.

18. Moghaddam M., Ferre T., Chen X., Chen K., Song X. and Hammond G. (2021) Applying simple machine learning tools to infer streambed flux from subsurface pressure and temperature observations. *Water Resources Research*, Submitted 2.19.20; 2020WR027375.
19. Ren H., Cromwell E., Kravitz B. and Chen X. (2021) Using long short-term memory models to fill data gaps in hydrological monitoring networks. *Hydrology and Earth System Sciences*, In review 1.29.20; doi: 10.5194/hess-2019-196.
20. Rubinstein R., Borton M. A., Zhou H., Shaffer M., Hoyt D. W., Stegen J. C., Henry C. S., Wrighton K. C. and Versteeg R. (2021) ORT: A workflow linking genome-scale meta-bolic models with reactive transport codes. *Bioinformatics*, Submitted; BIOINF-2021-0630.
21. Shuai P., Chen X., Hammond G., Song X., Chen K. and Zachara J. (2021) Hydrologic exchange flows alter river corridor thermal regimes in a dam-regulated river. *Water Resources Research*, Submitted 10/19; 2019WR026459.
22. Singh T., Gomez-Velez J., Wu L., Worman A., Hannah D. and Krause S. (2021) Impacts of antecedent peak flow events on hyporheic exchange and residence times. *Water Resources Research*, In review; 2020WR027113.
23. Ward A., Packman A. I., Bernal S., Brekenfeld N., Drummond J., Graham E., Hannah D., Klaar M., Krause S., Kurz M. J., Li A., Lupon A., Mao F., Marti Roca E., Ouellet V., Royer T., Stegen J. C., Wondzell S. and Zarnetske J. (2021) Advancing river corridor science beyond disciplinary boundaries with an inductive approach to hypothesis generation. *AGU Advances*, Submitted.
24. Yang J., Ye M., Chen X., Dai H. and Walker A. (2021) A new total-effect process sensitivity index to identify influential system processes under process model and parametric uncertainty. *Water Resources Research*, Submitted.
25. Zhu Y., Johnson T., Thomle J., Strickland C., Song X. and Stegen J. (2021) Joint hydrogeophysical inversion of vertical pressure, temperature, fluid conductivity, and bulk conductivity time-series to monitor dynamic mass flux at the groundwater/surface water interface. *Journal of Hydrology*, Submitted; HYDROL34482.

2020

1. Ahmmed B., Mudunuru M. K., Karra S., James S. C., Viswanathan H. and Dunbar J. A. (2020) PFLOTRAN-SIP: A PFLOTRAN module for simulating spectral-induced polarization of electrical impedance data. *Energies*, **13**, 6552; doi: 10.3390/en13246552.
2. Bao Y., Feng Y., Stegen J. C., Wu M., Chen R., Liu W., Zhang J., Li Z. and Lin X. (2020) Straw chemistry links the assembly of bacterial communities to decomposition in paddy soils. *Soil Biology and Biochemistry*, **148**, 107866; doi: 10.1016/j.soilbio.2020.107866.
3. Chen X., Zachara J. M., Vermuel V. R., Hammond G., Freshley M. and Fang Y. (2020) Understanding contaminant migration within a dynamic river corridor through field experiments and reactive transport modeling. *Frontiers in Water*, **2**; doi: 10.3389/frwa.2020.533796.
4. Danczak R. E., Chu R. K., Fansler S. J., Goldman A. E., Graham E. B., Tfaily M. M., Toyoda J. and Stegen J. C. (2020) Using metacommunity ecology to understand environmental metabolomes. Nature Communications, 11, 6369; doi: 10.1038/s41467-020-19989-y.
5. Danczak R. E., Daly R. A., Borton M. A., Stegen J. C., Roux S., Wrighton K. C. and Wilkins M. J. (2020) Ecological assembly processes are coordinated between bacterial and viral communities in fractured shale ecosystems. *Msystems*, **5**, e00098-00020; doi: 10.1128/mSystems.00098-20.
6. Fang Y. L., Chen X. Y., Velez J. G., Zhang X. S., Duan Z. R., Hammond G. E., Goldman A. E., Garayburu-Caruso V. A. and Graham E. B. (2020) A multirate mass transfer model to represent the interaction of multicomponent biogeochemical processes between surface water and hyporheic zones (SWAT-MRMT-R 1.0). *Geoscientific Model Development*, **13**, 3553-3569; doi: 10.5194/gmd-13-3553-2020.

7. Fang Y., Song X., Ren H., Perkins W. A., Shuai P., Richmond M. C., Hou Z., Bao J., Chen X. and Scheibe T. D. (2020) High-performance simulation of dynamic hydrologic exchange and implications for surrogate flow and reactive transport modeling in a large river corridor. *Frontiers in Water*, **2**; doi: 10.3389/frwa.2020.564211.
8. Gao Z. M., Liu H. P., Arntzen E., Mcfarland D. P., Chen X. Y. and Huang M. Y. (2020) Uncertainties in turbulent statistics and fluxes of CO2 associated with density effect corrections. *Geophysical Research Letters*, **47**; doi: 10.1029/2020GL088859.
9. Gao Z., Liu H., Chen X., Huang M., Missik J. E. C., Yao J., Arntzen E. and McFarland D. P. (2020) Enlarged nonclosure of surface energy balance with increasing atmospheric instabilities linked to changes in coherent structures. *Journal of Geophysical Research: Atmospheres*, **125**, e2020JD032889; doi: 10.1029/2020jd032889.
10. Garayburu-Caruso V. A., Danczak R. E., Stegen J. C., Renteria L., Mccall M., Goldman A. E., Chu R. K., Toyoda J., Resch C. T., Torgeson J. M., Wells J., Fansler S., Kumar S. and Graham E. B. (2020) Using community science to reveal the global chemogeography of river metabolomes. *Metabolites*, **10**, 518; doi: 10.3390/metabo10120518.
11. Garayburu-Caruso V. A., Stegen J. C., Song H.-S., Renteria L., Wells J., Garcia W., Resch C. T., Goldman A. E., Chu R. K., Toyoda J. and Graham E. B. (2020) Carbon limitation leads to thermodynamic regulation of aerobic metabolism. *Environmental Science & Technology Letters*, **7**, 517-524; doi: 10.1021/acs.estlett.0c00258.
12. Graham E. B. and Krause S. (2020) Social media sows consensus in disturbance ecology. *Nature,* **577**, 170; doi: 10.1038/d41586-020-00006-7.
13. Grant S. B., Gomez-Velez J. D., Ghisalberti M., Guymer I., Boano F., Roche K. and Harvey J. (2020) A one-dimensional model for turbulent mixing in the benthic biolayer of stream and coastal sediments. *Water Resources Research,* **56**, e2019WR026822; doi: 10.1029/2019wr026822.
14. Ji M., Kong W., Stegen J., Yue L., Wang F., Dong X., Cowan D. A. and Ferrari B. C. (2020) Distinct assembly mechanisms underlie similar biogeographical patterns of rare and abundant bacteria in Tibetan Plateau grassland soils. *Environmental Microbiology*, **22**, 2261-2272; doi: 10.1111/1462-2920.14993.
15. Kruegler J., Gomez-Velez J., Lautz L. K. and Endreny T. A. (2020) Dynamic evapotranspiration alters hyporheic flow and residence times in the intrameander zone. *Water*, 12, 424; doi: 10.3390/w12020424.
16. Li B., Liu X., Kaufman M. H., Turetcaia A., Chen X. and Cardenas M. B. (2020) Flexible and modular simultaneous modeling of flow and reactive transport in rivers and hyporheic zones. *Water Resources Research*, 56; doi: 10.1029/2019wr026528.
17. Lin X., Ren H., Goldman A., Stegen J. C. and Scheibe T. D. (2020) WHONDRS-GUI: A web application for global survey of surface water metabolites. *PeerJ*, **8**, e9277; doi: 10.7717/peerj.9277.
18. Liu W., Graham E. B., Zhong L., Zhang J., Li W., Li Z., Lin X., Feng Y. and Wang J. (2020) Dynamic microbial assembly processes correspond to soil fertility in sustainable paddy agroecosystems. *Functional Ecology*, 34, 1244-1256; doi: 10.1111/1365-2435.13550.
19. Liu W., Graham E. B., Zhong L., Zhang J., Li S., Lin X. and Feng Y. (2020) Long-term stochasticity combines with short-term variability in assembly processes to underlie rice paddy sustainability. *Frontiers in Microbiology*, 11, 873; doi: 10.3389/fmicb.2020.00873.
20. Nelson W. C., Graham E. B., Crump A. R., Fansler S. J., Arntzen E. V., Kennedy D. W. and Stegen J. C. (2020) Distinct temporal diversity profiles for nitrogen cycling genes in a hyporheic microbiome. *PLOS ONE,* **15**, e0228165; doi: 10.1371/journal.pone.0228165.
21. Ren H., Hou Z., Duan Z., Song X., Perkins W. A., Richmond M. C., Arntzen E. V. and Scheibe T. D. (2020) Spatial mapping of riverbed grain-size distribution using machine learning. *Frontiers in Water*, **2**; doi: 10.3389/frwa.2020.551627.
22. Rizzo C. B., Song X., de Barros F. P. J. and Chen X. (2020) Temporal flow variations interact with spatial physical heterogeneity to impact solute transport in managed river corridors. *Journal of Contaminant Hydrology*, 103713; doi: https://doi.org/10.1016/j.jconhyd.2020.103713.

23. Song H.-S., Stegen J. C., Graham E. B., Lee J.-Y., Garayburu-Caruso V. A., Nelson W. C., Chen X., Moulton J. D. and Scheibe T. D. (2020) Representing organic matter thermodynamics in biogeochemical reactions via substrate-explicit modeling. *Frontiers in Microbiology*, **11**; doi: 10.3389/fmicb.2020.531756.
24. Song X. H., Chen X. Y., Zachara J. M., Gomez-Velez J. D., Shuai P., Ren H. Y. and Hammond G. E. (2020) River dynamics control transit time distributions and biogeochemical reactions in a dam-regulated river corridor. *Water Resources Research*, **56**; doi: 10.1029/2019WR026470.
25. Thomle J., Strickland C., Johnson T. C., Zhu Y. and Stegen J. (2020) A flux detection probe to quantify dynamic groundwater-surface water exchange in the hyporheic zone. *Ground Water*, **58**, 892-900; doi: 10.1111/gwat.13001.
26. Tso C. M., Johnson T. C., Song X., Chen X., Kuras O., Wilkinson P., Uhlemann S., Chambers J. and Binley A. (2020) Integrated hydrogeophysical modelling and data assimilation for geoelectrical leak detection. *Journal of Contaminant Hydrology*, **234**, 103679; doi: 10.1016/j.jconhyd.2020.103679.
27. Villa J. A., Smith G. J., Ju Y., Renteria L., Angle J. C., Arntzen E., Harding S. F., Ren H., Chen X., Sawyer A. H., Graham E. B., Stegen J. C., Wrighton K. C. and Bohrer G. (2020) Methane and nitrous oxide porewater concentrations and surface fluxes of a regulated river. *Science of The Total Environment*, **715**, 136920; doi: 10.1016/j.scitotenv.2020.136920.
28. Wang J. J., Legendre P., Soininen J., Yeh C. F., Graham E., Stegen J. C., Casamayor E. O., Zhou J. Z., Shen J. and Pan F. Y. (2020) Temperature drives local contributions to beta diversity in mountain streams: Stochastic and deterministic processes. *Global Ecology and Biogeography*, 29, 420-432; doi: 10.1111/geb.13035.
29. Wu L., Gomez-Velez J. D., Krause S., Singh T., Wörman A. and Lewandowski J. (2020) Impact of flow alteration and temperature variability on hyporheic exchange. *Water Resources Research,* **56**; doi: 10.1029/2019wr026225.
30. Yao J., Liu H., Huang J., Gao Z., Wang G., Li D., Yu H. and Chen X. (2020) Accelerated dryland expansion regulates future variability in dryland gross primary production. Nature Communications, 11, 1665; doi: 10.1038/s41467-020-15515-2.
31. Zachara J. M., Chen X., Song X., Shuai P., Murray C. and Resch C. T. (2020) Kilometer-scale hydrologic exchange flows in a gravel bed river corridor and their implications to solute migration. *Water Resources Research,* **56**, e2019WR02525; doi: 10.1029/2019wr025258.
32. Zhu B., Huang M., Cheng Y., Xie X., Liu Y., Zhang X., Bisht G., Chen X., Missik J. and Liu H. (2020) Effects of irrigation on water, carbon, and nitrogen budgets in a semiarid watershed in the Pacific Northwest: A modeling study. *Journal of Advances in Modeling Earth Systems*, **12**, e2019MS001953; doi: 10.1029/2019ms001953.

APPENDIX B: LISTING OF COLLABORATIVE PROJECTS

Direct-Funded Collaborations – The SFA has directly funded the following external collaborations during FY20-21 to date:

- (2017-Current) Heping Liu, Washington State University – Install, maintain, and process data from eddy flux towers at multiple locations (RC-1).
- (2017-Current) Jesus Gomez-Velez, Vanderbilt University – Develop and apply the NEXSS model and other analytical methods to evaluate HEFs and link to other modeling efforts (RC-1, RC-2, MM)
- (2019-Current) Kelly Wrighton, Colorado State University – Collaborate on microbial community analysis including metagenomics studies for WHONDRS (RC-4)
- (2019-Current) Jeff Wiles, US Geological Survey – Install and maintain monitoring system for an irrigation return channel in the Hanford Reach and incorporate data into USGS water data system; see https://waterdata.usgs.gov/nwis/uv?site_no=12473503.
- (2020-Current) Hyun-Seob Song, University of Nebraska Lincoln – Develop methods for integrating multi-omics data into reactive transport models, and link to other modeling efforts.
- (2019-Current) John Selker, Oregon State University – Technology development for flux tool

ESS-Funded University Collaborations – The SFA has collaborated closely with several university-led projects funded by the ESS program during FY20-21:

Co-Funded Collaborators (PNNL receives limited supplemental funds under these projects):

- (FY19-21) Bayani Cardenas (University of Texas Austin): Respiration in Hyporheic Zones: Advancing the Understanding of Coupled Transport and Microbial Biogeochemistry and their Representation in Open-source Mechanistic Models
- (FY17-20) Michael Gooseff (University of Colorado) Quantifying Distributed Exchanges of Groundwater with River Corridors
- (FY20-22) Matt Ginder-Vogel (University of Wisconsin) Particulate Organic Matter (POM) Transport and Transformation at the Terrestrial-Aquatic Interface

WHONDRS – Collaborative relationships within WHONDRS are too numerous to list. However, here we note some of the key locations, investigators, and their institutions participating in the WHONDRS-Local effort through RC-4:

- St. Lawrence River – Francois Guillemette and Elizabeth Grater (Université du Québec à Trois-Rivières) – Longitudinal transect of the St Lawrence from Lake Ontario to the coast, encompassing the convergence of three rivers; Collaboration with EXCHANGE component of COMPASS project.
- Missouri Flat Creek – Aline Ortega Pieck (University of Idaho) and Sarah Roley (Washington State University) – Seasonal sampling in an agricultural stream to evaluate temporal impacts on N cycling, hydrologic connectivity, and organic matter.
- Svalbard – Jacob Yde and Gabby Kleber (Western Norway University of Applied Sciences) - Samples through space and time to evaluate metabolite variation across Arctic features, including permafrost groundwater springs, glacial outflow, mixing zones, and cryoconite holes.
- Nile River -Rehab Abdallah and Amged Ouf (American University in Cairo) - Sparse spatial transect
- Hyperspectral – Shuo Chen and Yuehan Lu (University of Alabama) - Samples through space and time paired with remotely sensed hyperspectral data to evaluate statistical relationships.
- Josh Sharp and Laura Leonard (Colorado School of Mines) – Samples through space and time used for pre- and post-laboratory chlorination to evaluate disinfection byproduct formation.

SBIR-Funded Industrial Collaborations – The SFA works with a number of small businesses to apply and test new field, laboratory and modeling technologies. For example, we recently teamed with Michael Wilde of Parallel Works, Inc. on a successful Phase II proposal entitled "A Platform for Scientific Data Management, Modeling and Analysis with Machine Learning."

APPENDIX C: REFERENCES CITED

Bird, M.I., Wynn, J.G., Saiz, G., Wurster, C.M. and McBeath, A. (2015) The Pyrogenic Carbon Cycle. Annu Rev Earth Pl Sc 43, 273-298.

Chen, Y., Bao, J., Fang, Y., Scheibe, T.D., Richmond, M.C., Perkins, W.A., Ren, H., Song, X., Duan, Z., Hou, Z. and X, H. (2020) Modeling the effects of topography heterogeneity and discharge variations on riverbed hydrodynamics in a 30-kilometer-long reach over a nine-year period using OpenFOAM. Earth and Space Science Open Archive.

Damerow, J., Varadharajan, C., Boye, K., Brodie, E.L., Burrus, M., Chadwick, D., Crystal-Ornelas, R., Elbashandy, H., Eloy Alves, R., Ely, K., Goldman, A., Haberman, T., Hendrix, V., Kakalia, Z.K., K. M., Kersting, A., Merino, N., O'Brien, F., Perzan, Z., Robles, E., Sorensen, P., Stegen, J., Walls, R., Weisenhorn, P., Zavarin, M. and Agarwal, D. (2021) Sample identifiers and metadata to support data management and reuse in multidisciplinary ecosystem sciences. Data Science Journal 20, 1-19.

Fang, Y.L., Chen, X.Y., Velez, J.G., Zhang, X.S., Duan, Z.R., Hammond, G.E., Goldman, A.E., Garayburu-Caruso, V.A. and Graham, E.B. (2020) A multirate mass transfer model to represent the interaction of multicomponent biogeochemical processes between surface water and hyporheic zones (SWAT-MRMT-R 1.0). Geosci Model Dev 13, 3553-3569.

Goldman, A.E., Emani, S.R., Pérez-Angel, L.C., Rodríguez-Ramos, J.A., Stegen, J.C. and Fox, P. (2021) Special collection on open collaboration across geosciences. Eos 102.

Graham, E.B., Averill, C., Bond-Lamberty, B., Knelman, J.E., Krause, S., Peralta, A.L., Shade, A., Smith, A.P., Cheng, S.J., Fanin, N., Freund, C., Garcia, P.E., Gibbons, S.M., Van Goethem, M.W., Ben Guebila, M., Kemppinen, J., Nowicki, R.J., Pausas, J.G., Reed, S.P., Rocca, J., Sengupta, A., Sihi, D., Simonin, M., Slowinski, M., Spawn, S.A., Sutherland, I., Tonkin, J.D., Wisnoski, N.I., Zipper, S.C. and Consortium, C. (2021) Toward a Generalizable Framework of Disturbance Ecology Through Crowdsourced Science. Front Ecol Evol 9.

Moreland, K.C., Dove, N.C., Hart, S.C., Hou, T., Filley, T., Barnes, M.E., Yan, Q., Lohse, K.A., Kumar, P. and Berhe, A.A. (2021) Organic matter dynamics in the critical zone., in: White, T. (Ed.), Critical Zone and Ecosystem Dynamics. Springer, New York.

Myers-Pigg, A.N., Louchouarn, P., Amon, R.M.W., Prokushkin, A., Pierce, K. and Rubtsov, A. (2015) Labile pyrogenic dissolved organic carbon in major Siberian Arctic rivers: Implications for wildfire-stream metabolic linkages. Geophys Res Lett 42, 377-385.

Ren, H., Song, X., Fang, Y., Hou, Z. and Scheibe, T. (2021) Machine learning analysis of hydrologic exchange flows and transit time distributions in a large regulated river. Frontiers in Artificial Intelligence 4.

Roebuck, J., Myers-Pigg, A. and E, G. (2021) Spatial and temporal responses in stream dissolved organic matter following a wildfire. JGR Biogeosciences in preparation.

Santin, C., Doerr, S.H., Merino, A., Bucheli, T.D., Bryant, R., Ascough, P., Gao, X.D. and Masiello, C.A. (2017) Carbon sequestration potential and physicochemical properties differ between wildfire charcoals and slow-pyrolysis biochars. Sci Rep-Uk 7.

Santin, C., Doerr, S.H., Preston, C.M. and Gonzalez-Rodriguez, G. (2015) Pyrogenic organic matter production from wildfires: a missing sink in the global carbon cycle. Global Change Biol 21, 1621-1633.

Song, H.S., Stegen, J.C., Graham, E.B., Lee, J.Y., Garayburu-Caruso, V.A., Nelson, W.C., Chen, X.Y., Moulton, J.D. and Scheibe, T.D. (2020) Representing Organic Matter Thermodynamics in Biogeochemical Reactions via Substrate-Explicit Modeling. Front Microbiol 11.

Song, X., Fang, Y., Bao, J., Ren, H., Duan, Z., Perkins, W., Zhou, H., Hou, Z., Chen, Y. and Scheibe, T. (2021) Scale-dependent spatial variabilities of hydrological exchange flows and transit time in a large regulated river. Journal of Hydrology 598, 126283.

Vega, J.A., Fonturbel, T., Merino, A., Fernandez, C., Ferreiro, A. and Jimenez, E. (2013) Testing the ability of visual indicators of soil burn severity to reflect changes in soil chemical and microbial properties in pine forests and shrubland. Plant Soil 369, 73-91.

Wagner, S., Brandes, J., Goranov, A.I., Drake, T.W., Spencer, R.G.M. and Stubbins, A. (2017a) Online quantification and compound-specific stable isotopic analysis of black carbon in environmental matrices via liquid chromatography-isotope ratio mass spectrometry. Limnol Oceanogr-Meth 15, 995-1006.

Wagner, S., Ding, Y. and Jaffe, R. (2017b) A New Perspective on the Apparent Solubility of Dissolved Black Carbon. Front Earth Sc-Switz 5.