Data Carpentry: Final Project

Michelle Kelly

25 July, 2018

Introduction

The conversion of solar energy, water, and carbon dioxide into metabolizable carbon products forms the foundation of the trophic ladder: supporting both producers and consumers. The conversion rate of solar energy to organic energy within an ecosystem is termed gross primary productivity (GPP), and can be measured as oxygen produced per day. The consumption of this organic energy, by both autotrophs and heterotrophs, is termed ecosystem respiration (ER). The balance between GPP and ER is referred to as net ecosystem production (NEP), and can ultimately be used as an indicator of whether an ecosystem is retaining or losing carbon (Odum 1956; Woodwell and Whittaker 1968).

In terrestrial systems, NEP often follows a predictable annual cycle. GPP and NEP tend to peak during warm, wet summer months, when conditions are most favorable for photosynthetic growth. Primary production in lakes usually syncs up with the terrestrial cycle; greater light availability, higher water temperatures, and available dissolved nutrients create the ideal cocktail for algae, moss, and macrophyte growth (Woodwell and Whittaker 1968).

However, primary production in rivers and streams doesn't often correlate with the terrestrial growing season. In small streams, light availability decreases in summer, as canopy leaf-out prevents light from reaching surface water. Variations in stream flow, which can be caused by rain events, snow melt, or drought, scour or desiccate stream beds, reducing the biomass of primary producers. Stream hydrology can follow a yearly pattern, but vary widely between biomes [Figure 1]. Furthermore, reaches may also receive a significant carbon input from non-aquatic sources, such as an influx of leaf litter during Autumn, or the flushing of soil-bound organic matter during rain events (Vannote et al. 1980). Carbon input from terrestrial sources can equal or exceed yearly GPP, blurring the seasonal pattern of GPP and ER (Bernhardt et al. 2018).
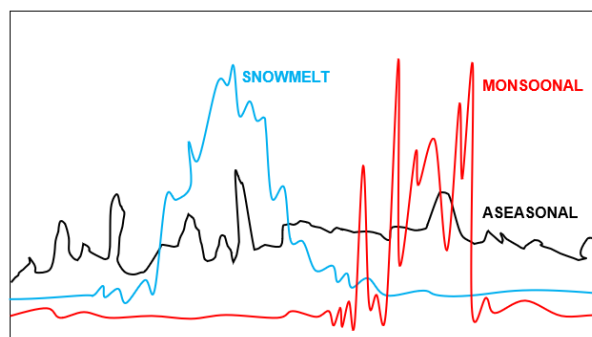

Figure 1: Conceptual yearly hydrographs for rivers experiencing aseasonal rain dynamics, a mid-year monsoon season, and a spring snowmelt (adapted from Bernhardt et al. 2018).

As stream size increases, annual ecosystem productivity is less effected by reach morphology, and is more strongly correlated with the terrestrial growing season. In wide, open rivers, canopy cover ceases to be a limiting factor [Figure 2]. Water velocity decreases with increasing channel width, reducing scouring, even under the same hydrologic regimes (Leopold and Maddock 1953). In large rivers, terrestrial organic matter tends to make up a smaller fraction of the system's total dissolved organic matter (Vannote et al. 1980).
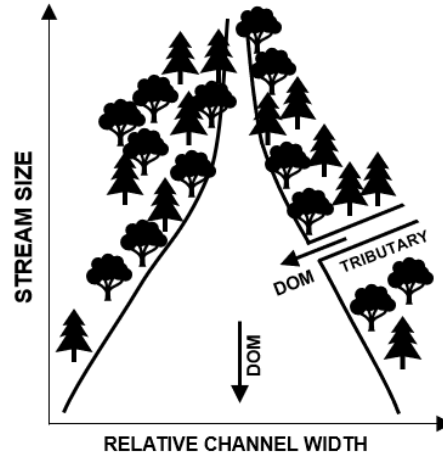
Figure 2: The relationship between stream size and stream attributes (adapted from Vannote et al. 1980).

Recently, Bernhardt et. al proposed a conceptual framework of stream "metabolic regimes", suggesting dominant annual patterns of GPP and ER dynamics for river ecosystems subjected to similar physical and biochemical conditions (2018). We propose to "ground-truth" the metabolism regime concept by asking (1) do we see these proposed regimes in real stream data, and (2) can use an annual metabolic profile to "diagnose" local stream conditions, such as annual precipitation patterns, stream size, or surrounding land use?

Methods

Site details and raw data

We selected the Mississippi River (Clinton, IA), the North Branch of the Park River (Hartford, CT), and the Sacramento River (Freeport, CA), for this analysis. These sites are located at similar latitudes to account for overall differences in light intensity, but have differences in river size, land use, and climate regime. Data was provided by the United States Geological Survey (USGS) National Water Information System database (waterdata.usgs.gov/nwis), and spanned from January 1, 2017 to November 1, 2017. Stream discharge, water temperature, and dissolved oxygen were measured at 15 minute intervals from January 1, 2017 to December 31, 2017 at the Sacramento and Park River sites, and thru November 1, 2017 at the Mississippi River site due to ice cover.

Table 1: Site characteristics of Park, Sacramento, and Mississippi River stations for 2017. Precipitation regime based on 2017 NOAA data. Land use based on USDA land cover maps. N data was not available for Park (= NA). Q = discharge, T = temperature, N = nitrate plus nitrite, min = minimum, max = maximum.

Characteristic	Park	Sacramento	Mississippi
Site location (latitude)	41°47'04.0"N	38°27'22.0"N	41°46'50.0"N
Site location (longitude)	72°42'29.0"W	121°30'01.0"W	90°15'07.0"W
USGS gage number	01191000	11447650	05420500
Watershed drainage area (km^2)	69.4	68,600	221,000

Characteristic	Park	Sacramento	Mississippi
Precipitation regime	Aseasonal	Winter (Jan-Feb)	Spring/Summer (Apr-Jul)
Dominant land use	Forested, urban	Urban, agriculture	Agriculture
Mean Q (m^3 s^{-1}) (min-max)	0.436 (0.056-25.6)	1070 (289-2730)	2340 (728-4730)
Mean T (°C) (min-max)	11 (0.0-27)	15 (7.9-23)	13 (0.9-28)
Mean turbidity (FNU) (min-max)	7.48 (0.9-584)	19.5 (3.40-129)	22.2 (2.4-256)
Mean N (mg-N L^{-1}) (min-max)	NA	0.15 (0.08-0.51)	2.4 (0.44-3.7)

Photosynthetically active radiation (PAR) was estimated using time and latitude. A more common (and accurate) approach would be to use PAR data collected at a local solar irradiance monitoring station, or collected by a portable in-field sensor. For the purposes of this preliminary proposal, we chose an estimation technique for a quick proof of concept. Depth data was back-calculated from instantaneous discharge measurements using $d = c\,Q^f$, where depth (d) is equal to the theoretical coefficient of depth at unit discharge (c, 0.409 m) multiplied by discharge (Q, m^3 s^{-1}) to the power of the theoretical exponent of depth at unit discharge (f, 0.294) (see Leopold and Maddock 1953; Raymond et al. 2012).

Missing data points were filled via interpolation, using `na.seasplit` function from the `imputeTS` R package (Moritz and Bartz-Beielstein 2017; R Core Team 2013). This function seasonally decomposes time series data, and performs imputation seperately for each resulting seasonal chunk. This interpolation method is the same as that used by the `streamPULSE` team, a national stream metabolism research group (Vlah and Berdanier 2018).

Metabolism modeling

We used the `streamMetabolizer` R package to fit a Markov chain Monte Carlo (MCMC) model to the data, using 500 samples for a warmup and 500 samples to fit the model (1000 total samples) (Appling et al. 2017). For this preliminary report, the metabolism model was fit to the Mississippi River site only. Standard methods for the direct calculation of GPP, ER, and NEP (see Odum 1956) can't allow for changing air-water gas exchange dynamics with changing discharge, which introduces considerable error into multi-day metabolism estimates (Appling et al. 2018).

Results and discussion

Values of GPP and ER both temporally varied, but GPP estimates were most dependent upon season. It is likely that GPP may be correlated with T, PAR, and Q, but this analysis has not yet been completed. In the Mississippi, GPP was lowest during colder months (January-March and October-November), and with the exception of a decrease in GPP in early July, closely follows the terrestrial growing season. Confidence intervals on ER estimates were much wider than those on GPP estimates, especially during colder months [Figure **??**]. A similar analysis of the lower Mississippi River (Baton Rouge, LA) shows a negative correlation between PAR and GPP, and suggests that PAR may be indirectly related to disturbance effects, as maximum PAR and highest Q co-occured during late summer (Dodds et al. 2013).
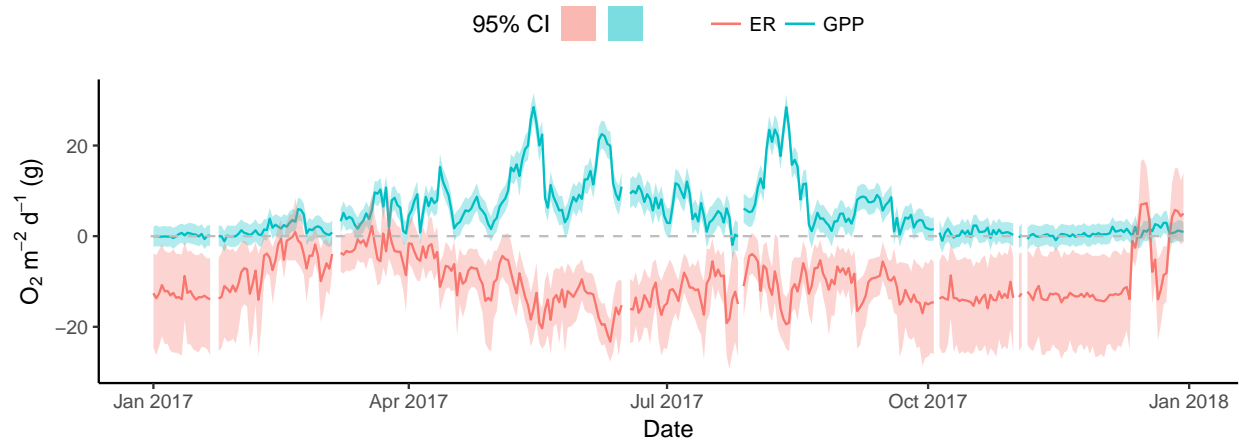
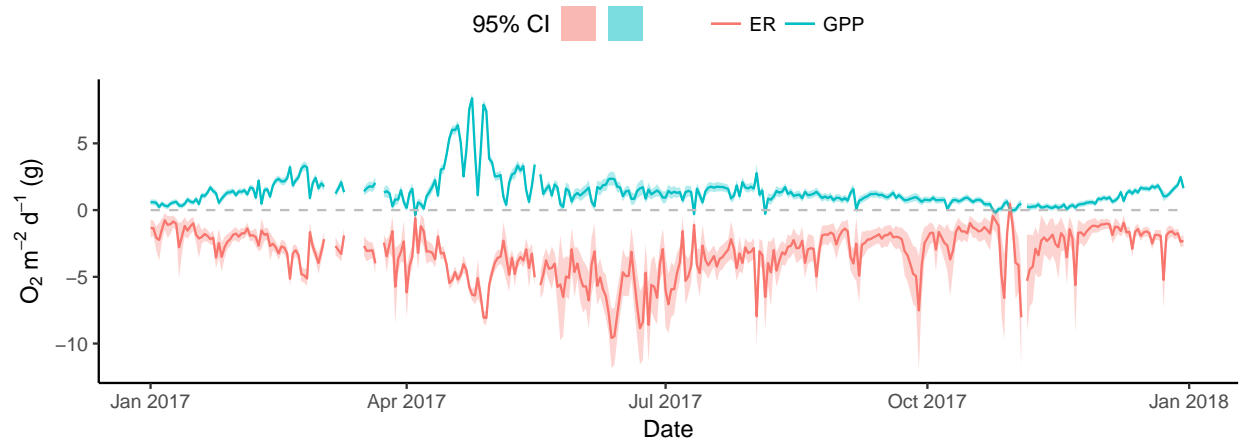Figure 3: An annual estimate of GPP and ER in the Mississippi River, 2017.

Figure 4: An annual estimate of GPP and ER in the North Branch of the Park River, 2017.
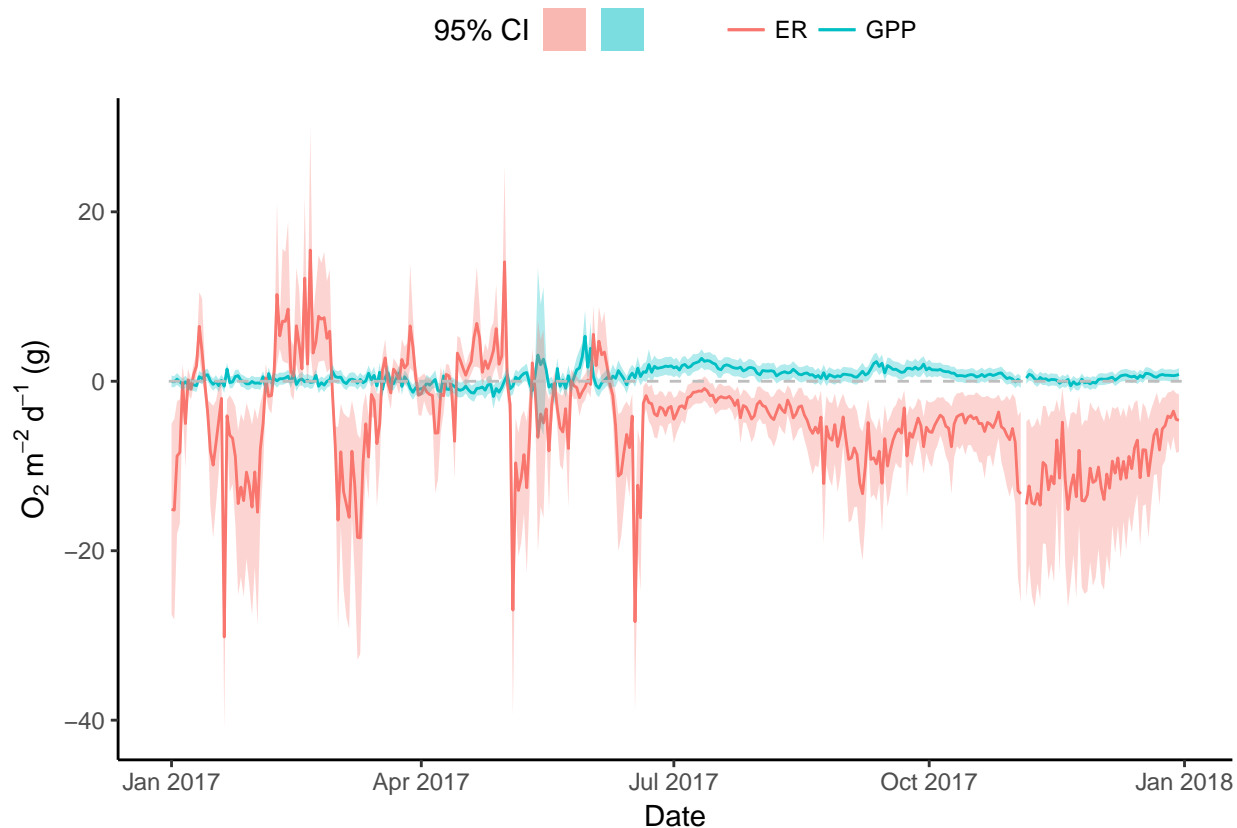
```
metabolismPlot(model_Sacramento,
               filename = "./FinalProject/figures/SacramentoR_MetabolismModel.png")
```

```
## Warning: Removed 2 rows containing missing values (geom_path).

## Warning: Removed 2 rows containing missing values (geom_path).

## Warning: Removed 2 rows containing missing values (geom_path).

## Warning: Removed 2 rows containing missing values (geom_path).
```

For the final report, we will compare and contrast the metabolism profiles of these three sites, and identify possible predictors for a resultant "regime types". We will also evaluate how ecosystem NEP varies throughout the year.

Modeling metabolism for a full year and across different rivers will allow us to explore the effects of both seasonal (T, PAR, Q) and environmental (land use, climate) constraints on ecosystem productivity. In the past, metabolism modeling over long time series was rare, due to the energy required to collect and process data. Now, with the advent of cheap in-situ sensors and modeling packages, these analyses are becoming more common. Greater data availability will allow us to have a better understanding of what drives the ecological differences between stream systems. Ultimately, this may allow us to improve our estimates of carbon retention and loss from these systems, and determine the drivers behind both processes.

References

Appling, A. P., R. O. Hall, M. Arroita, and C. B. Yackulic. 2017. *StreamMetabolizer: Models for Estimating Aquatic Photosynthesis and Respiration.* https://github.com/USGS-R/streamMetabolizer.

Appling, A. P., R. O. Hall, C. B. Yackulic, and M. Arroita. 2018. "Overcoming Equifinality: Leveraging Long Time Series for Stream Metabolism Estimation." *Journal of Geophysical Research: Biogeosciences* 123 (2): 624–45. doi:10.1002/2017JG004140.

Bernhardt, E. S., J. B. Heffernan, N. B. Grimm, E. H. Stanley, J. W. Harvey, M. Arroita, A. P. Appling, et al. 2018. "The metabolic regimes of flowing waters." *Limnology and Oceanography* 63: S99–S118. doi:10.1002/lno.10726.

Dodds, Walter K., Allison M. Veach, Claire M. Ruffing, Danelle M. Larson, Jason L. Fischer, and Katie H. Costigan. 2013. "Abiotic controls and temporal variability of river metabolism: multiyear analyses of

Mississippi and Chattahoochee River data." *Freshwater Science* 32 (4): 1073–87. doi:10.1899/13-018.1.

Leopold, Luna Bergere, and Thomas Maddock. 1953. "The Hydraulic Geometry of Stream Channels and Some Physiographic Implications." *Geological Survey Professional Paper* 252: 1–53.

Moritz, S., and T. Bartz-Beielstein. 2017. "imputeTS: Time Series Missing Value Imputation in R." *The R Journal* 9 (1): 207–18. https://journal.r-project.org/archive/2017/RJ-2017-009/index.html.

Odum, Howard T. 1956. "Primary Production in Flowing Waters." *Limnology and Oceanography* 1 (2): 102–17. doi:10.4319/lo.1956.1.2.0102.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Raymond, Peter A., Christopher J. Zappa, David Butman, Thomas L. Bott, Jody Potter, Patrick Mulholland, Andrew E. Laursen, William H. McDowell, and Denis Newbold. 2012. "Scaling the gas transfer velocity and hydraulic geometry in streams and small rivers." *Limnology and Oceanography: Fluids and Environments* 2 (1): 41–53. doi:10.1215/21573689-1597669.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell, and C.E. Cushing. 1980. "The River Continuum Concept." *Canadian Journal of Fisheries and Aquatic Sciences* 37 (1): 130–37. https://doi.org/10.1139/f80-017.

Vlah, M., and A. Berdanier. 2018. *StreamPULSE: Run Stream Metabolism Models on Streampulse Data.* https://github.com/streampulse/StreamPULSE.

Woodwell, George M, and Robert H Whittaker. 1968. "Primary Production in Terrestrial Ecosystems." *Integrative and Comparative Biology* 8 (1): 19–30. doi:https://doi.org/10.1093/icb/8.1.19.