**BOT5555**

**February 21, 2017**

**FUTURE DELIVERABLES (STREAM TEAM)**

**Daniel Greenwood, Abigail Hoffman, Atefeh Hosseini, Nathaniel Kitchel, Katherine Lawry, Nicholas Oakley**

**PURPOSE:** Ecosystem metabolism is the process that controls organic carbon cycling in stream and river networks (Hall et. al 2015). It is of great importance conceptually, as it quantitatively describes the total energetics of all autotrophic and heterotrophic organisms in the system over time (Hall et. al 2016). It encompasses all reduction of CO2 to organic carbon (C) via photosynthesis by all autotrophs in the system (Gross Primary Production, or GPP) and all oxidation from (C) back to CO2 via respiration by all heterotrophs and autotrophs in the system (Ecosystem Respiration, or ER). Its practical implications range from small, reach-scale uses (e.g. fisheries managers estimating changes in dissolved oxygen availability for stream fishes and other aquatic organisms) to much larger, broader uses (e.g. estimating CO2 fluxes from large rivers for climate models).

While estimates of ecosystem metabolism in small streams are ubiquitous within the stream ecology literature, to date few attempts have been made to estimate ecosystem metabolism in large river systems. Furthermore, the methods that have been used for estimating ecosystem metabolism on large rivers have either been shown to yield biased results in variable flow rate situations (e.g. Hall et. al 2015) or they are in a format that is not useable by the broader research community (e.g. Payn et. al (in prep)). We seek to improve upon current methods of estimating river metabolism in large rivers by developing a new model that both incorporates the uncertainty posed by variable flow conditions and is easily useable by the broader research community.

**PROBLEM:** Estimating components of ecosystem metabolism, namely gross Gross Primary Production (GPP) and Ecosystem Respiration (ER), for a river reach is a non-trivial process. Traditionally, stream ecologists have estimated GPP and ER from dissolved oxygen measurements recorded by a single oxygen sensor placed at a single point along the stream reach. The sensor is deployed for a 24 hour period, and the difference between the daytime dissolved oxygen ‘spike’ (caused by photosynthesis) and the nighttime dissolved oxygen ‘lull’ (caused by respiration) is then calculated as GPP-ER. This “single station model,” assumes homogeneity of all factors that may influence dissolved oxygen (e.g. constant channel geomorphology, shading, etc.) along the entire stream reach.

While simpler from an analysis standpoint, the single station approach is not suitable for estimating GPP and ER along river reaches where dissolved oxygen is influenced by outside factors such as aeration from waterfalls or rapids, inputs of nutrients from a wastewater treatment facilities, or irregular discharge of water from a dam (Hall 2016).  In such situations, it is necessary to install both an oxygen sensor upstream of any channel irregularity and an oxygen sensor downstream of said channel irregularity and then measure the diel change in dissolved oxygen at both points along the reach to provide a robust estimate of GPP and ER. This is referred to as the “two station model.”

Analytically, a two station setup complicates estimation of GPP and ER, as one must be able to “sync” the measurements taken by the downstream oxygen sensor with those taken earlier by the upstream oxygen sensor to ensure that both measurements are matching up with the same “packet” of water. Given constant flow rate, the solution to this problem is relatively straightforward: one has only to measure the stream velocity, and the distance travelled by each “packet” of water, and then one can calculate the time step difference between the two oxygen probes. When the assumption of constant flow is violated, however, the problem of syncing oxygen measurements becomes problematic as the time it takes each “packet” of water to travel the distance between the two oxygen sensors likewise becomes variable. Assuming a constant flow rate under flow conditions that are far from constant not surprisingly produces biased results. This is referred to as the “pernicious problem.”

          Hall and colleagues (2015) outline a Bayesian method to estimate GPP and ER with a two station model. This model was implemented using packages in the R software environment (R Core Team 2016) that have now been superseded by more capable software that is easier to understand and use, specifically JAGS.  Additionally, Hall’s two station method assumes a constant daily flow rate for a given river reach. At stream sites where the flow rate is highly variable (such as below hydroelectric dams), this model has been shown to produce biased results (Payn et. al (in prep)). While Payn and colleagues (in prep) present a solution to the “pernicious problem” of variable flow rates (Hall 2016), this model is run in Java 7 with later versions of Java presenting back compatibility problems. This model is also complex and difficult for researchers other than the original author of the model to interpret or use. As such, this model is not broadly useful to other members of this research community. Our team will incorporate variable flow rate into the two-station model developed by Hall and colleagues (2015) and do so in a way that is reproducible and widely available to researchers interested in calculating stream metabolism.

**OBJECTIVES:**

1. Modify the original R code for Hall’s constant flow rate two station model so that it can be implemented with JAGS
2. Overcome the “pernicious” problem ourselves via development of a new model that utilizes a Bayesian approach to relax the assumption of constant flow rate

\*Use the Stream Metabolizer package as a basic template and build this new model in R

1. Test the performance of the the abovementioned models (Hall’s and ours) by applying each model to simulated data.

\*Use stream metabolizer to generate a set of simulated data for us, and see if we can recover those data with Hall’s code and with our new variable flow model

\*Determine the point at which Hall’s constant flow rate model becomes significantly less accurate than our variable flow rate model

\*Quantify the level of discrepancy between the two models and determine the point at which that discrepancy increases to a level that is unacceptable

\*Make a figure with the amount of discrepancy between the two models on y axis and flow variability on x axis.

1. Confront each model with real data (from the Colorado River, etc).
2. Annotate all code extensively to ensure transparency and reproducibility

\*Via the use of RNotebooks

**WORKFLOW SCHEDULE (\*tentative and guaranteed to change\*):**

|  |  |  |
| --- | --- | --- |
| **WHAT NEEDS TO BE DONE?** | **WHO'S GOING TO DO IT?** | **WHEN MUST IT BE DONE BY?** |
|
| **future deliverables document** | **EVERYONE** | **21-Feb-17** |
| email Bob about missing data | Atefeh |
| upload new future deliverables document so Buerkle can see it | Nathan |
| Finish translating Bob's code into Jags | Abby, Atefeh, Nathan, Katie | **28-Feb-17** |
| figure out StreamMetabolizer simulation stuff | Daniel, Nicholas |
| figure out GitHub/ "issues" to do list thing | EVERYONE |
| Determine appropriate distributions for priors, parameters, hyperparameters, etc. | Katie |
| Be able to write out likelihood function | Katie |
| Upload "how do do shit in git" doc | Daniel |
| view metadata on wyocourses and figure out how to incorporate into JAGS | Abby |
| transfer this list to the github wiki | Daniel? |
| code new model that incorporates variable flow rate | Abby, Katie, Nathan | **7-Mar-17** |
| Test Bob's constant flow model with simulated data from Stream Metabolizer | Daniel, Nick, Atefeh |
| Summarize results of testing Bob's constant flow model | Daniel, Nick, Atefeh |
| **SPRING BREAK! (catch up on anything we are running behind on)** | EVERYONE | **14-Mar-17** |
| Test new variable flow model with simulated data from Stream Metabolizer | Abby, Katie, Nathan | **21-Mar-17** |
| Summarize findings of simulated data test on new variable flow model | Abby, Katie, Nathan |
| Start outlining interim report (incorporate findings of test on Bob's model) | Daniel, Nick, Atefeh |
| Quantify discrepancy between the two models | EVERYONE | **28-Mar-17** |
| Make figure showing discrepancy between the two models |
| rough draft of interim report |
| Email Bob with preliminary results and request real data for "model confrontation" |
| **Deliery of interim report on analysis** | **EVERYONE** | **4-Apr-17** |
| Confront Bob's constant flow model with real data and summarize findings | Daniel, Nick, Atefeh | **11-Apr-17** |
| Confront new variable flow model with real data and summarize findings | Abby, Katie, Nathan |
| Quantify discrepancy between the two models with REAL DATA | EVERYONE | **18-Apr-17** |
| Make figure showing discrepancy between the two models WITH REAL DATA |
| work on final report |
| Rough draft of final report "due" | EVERYONE | **25-Apr-17** |
| **Delivery of final report on analysis** | **EVERYONE** | **2-May-17** |