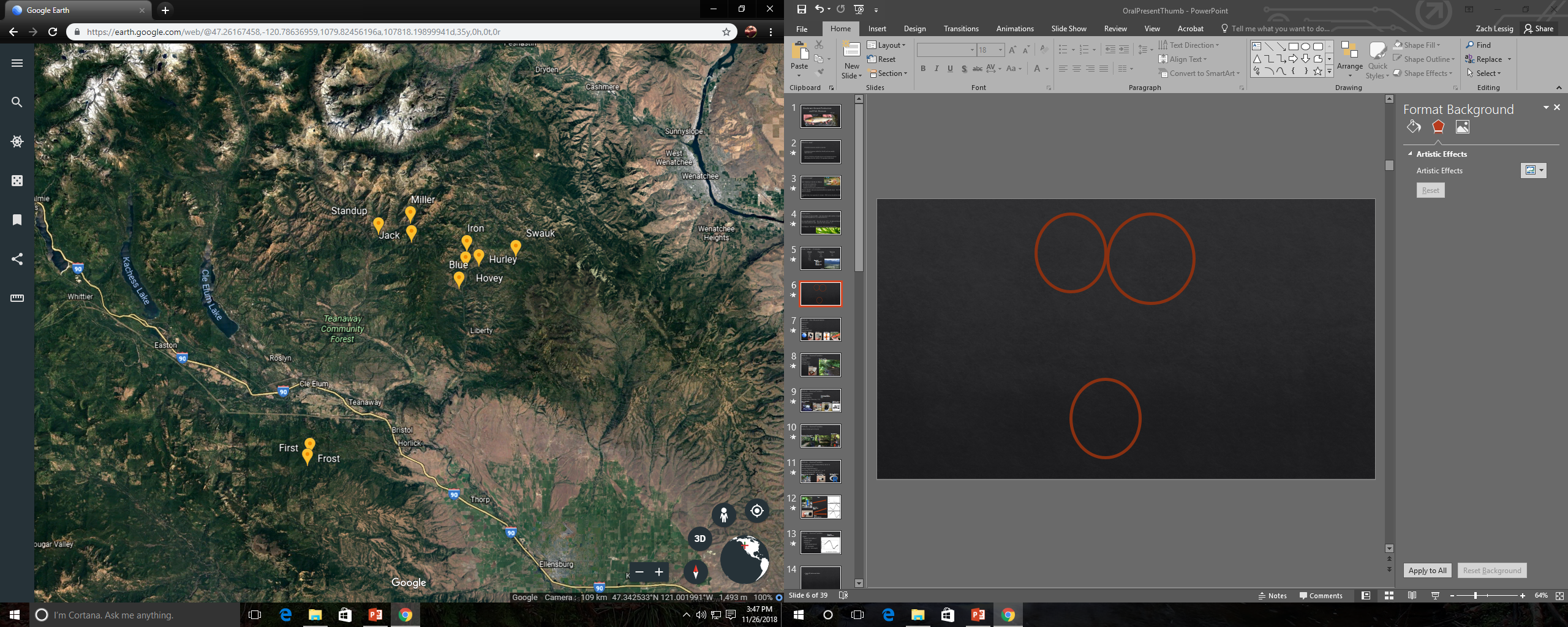
(Weigel, Peterson, & Spruell, 2002)**Methods**

Study Design

My ten study sites were located on first and second order headwater streams in the Swauk (n=5), Teanaway (n=3), and Taneum (n=2) watersheds in Kittitas County, WA. These sites, on the east slope of the Cascade Mountains in the Yakima River Basin, have a hydrograph mainly driven by snowmelt with peak runoff in May and baseflow at the end of July to beginning of October (US Bureau of Reclamation, 2018). The 5 sites in Swauk were on Blue, Hovey, Hurley, Iron, and Swauk creeks. The 2 sites in Taneum were on First and Frost creeks, and the 3 sites in Teanaway were on Jack, Miller and Standup creeks (Figure 1). 

Taneum

Swauk

Teanaway

Figure 1. Map showing Teanaway, Swauk, and Taneum watersheds with respective study sites.

I sampled these sites 3 times between 2017 and 2018 to capture seasonal variation in stream conditions. The first sampling period was in the summer of 2017 from July 19 to August 15, the second sampling period was in the fall of 2017 from November 5 to November 16, and the third and final sampling was in the summer of 2018 from Jun 26 to July 15.

At each site on these low order headwater streams I collected site description data one time; GPS coordinates (MotionX-GPS version 24.1, Fullpower Technologies on Apple iPhone 5), and elevation (Google Earth) with sites ranging from 869 to 1071m. Stream aspect (Lensatic compass, Engineer) which was converted to a value of degrees south facing (e.g. 180º is directly south and 0° is north). Stream slopes (Suunto PM-5 Clinometer) were from 2 to 10% and bank full widths ranged from 3.3 to 12.7m. I conducted a Wolman Pebble Count (Wolman, 1954) with 50 pebbles measured per stream where median pebble widths ranged from 35.5 to 92mm.

For each sampling period I took data on the following variables; riparian overstory density (canopy openness) with a densitometer (Spherical Crown Densiometer, Convex Model A, Forestry Suppliers) where values ranged from 4.9 % open for one of the most narrow streams (Frost) in the Summer to 78.1% for the widest stream (Standup) during the fall. I measured stream discharge with a portable flow meter (Flo-Mate 2000, Marsh-McBirney) according to Rantz (1982) with discharges ranging from 0.3 to 65.5 L/s. For each sampling period I also sampled stream water for nutrient analysis, conducted fish population/biomass estimates and estimated stream metabolism according to the methods below.

Stream Nutrients

I collected stream water in acid washed HDPE bottles using 1 µm glass fiber syringe filters (Type A/E Glass Fiber Filter, Pall Corporation). I acidified one of these samples intended for dissolved organic carbon (DOC) analysis with 100 µL of 0.5N HCl to ensure pH was at or below 2. These samples were transported in a cooler out of the field and stored in a freezer until analysis.

I analyzed the samples for ammonium (NH4+) using EPA-103-B Rev. 1 method (2012) with the exception that 0.025 mg/L ammonium was added to the sample to ensure concentrations were within the detection limit and later subtracted. I analyzed nitrate + nitrite (NO3- + NO2-; hereafter referred to as NO3-) using EPA-127-B Rev. 1 (2012) and added the ammonium and nitrate concentrations together to obtain a concentration of total dissolved inorganic nitrogen (DIN). I analyzed the sample for phosphate (PO43-) using EPA-155-B Rev. 0 (2016) referred to here as soluble reactive phosphorus (SRP). These samples were run on an AQ1 Discrete Analyzer (Seal Analytical). The acidified DOC sample was analyzed on a Shimadzu TOC-L (TOC-L Total Organic Carbon Analyzer, Shimadzu) with techniques outlined in the administrators manual.

Fish

I conducted a population estimate of stream salmonids (Family Salmonidae) immediately upstream of each site where DO probes were deployed and water samples were taken. The fish included in the estimate were almost all native westslope cutthroat trout (*Oncorhynchus clarkii lewisi*) with some fish displaying signs of hybridization with the native redband rainbow trout (*Oncorhynchus mykiss gairdneri*) (Weigel, Peterson, & Spruell, 2002). A few non-native eastern brook trout (*Salvelinus fontinalis*) were included in the Jack Cr. 2018 population and biomass estimate as well. Some young-of-the-year (YOY) salmonids and sculpin (*Cottus spp.*) were also encountered but not included in the estimates.

I used a backpack electrofisher (LR-20B Electrofisher, Smith Root) to collect fish from a 25 m length of stream and an assistant who caught the salmonids 50 mm or more in length with a dip net and placed them in a 5 gallon bucket. I used the two-pass depletion method (Lockwood & Schneider, 2000) and did not include block-nets. Block-nets to prevent migration were not used because these streams were relatively small and the time elapsed between the first and second pass was only a few minutes. The assumptions are met for this estimate as long as migration is negligible. I anesthetized the fish to measure and weigh them using Tricaine Methanesulfonate according to Central Washington University Institutional Animal Care and Use Committee (IACUC protocol #A041710). I calculated the fish population with the two-pass depletion method from Lockwood & Schneider (2000) as follows:

Where, C1 is the number of fish removed in the first pass, C2 is the number of fish removed in the second pass, N is the population estimate in numbers of fish and SE is the standard error of N. This population estimate was then divided by the length of stream sampled to provide a measure of fish density in fish m-1.

I estimated fish biomass by multiplying the population estimate by the average mass of the fish and then dividing by the stream length sampled. The average fish mass came from the combination of the fish caught in both passes.

Stream Metabolism

At each site and for each sampling period I deployed a dissolved oxygen (DO) probe (miniDOT Submersible Water Logger, Precision Measurement Engineering) in the stream to measure DO in mg L-1 and temperature. I also deployed a photosynthetically active radiation (PAR) logger (Odyssey Photosynthetic Active Radiation Logger, Dataflow Systems) on the stream bank within 2 meters of the DO probe to measure light as PAR in µmol photons m-2 s-1. These two instruments were synchronized to collect data every 10 minutes (first sampling period only) or every 5 minutes for at least 36 hours per stream from ~4:00 p.m. on day one to ~9:00 a.m. on day three. After this deployment, I immediately moved the loggers to a different site until all sites had been sampled.

I used the diel DO and PAR curves to estimate stream metabolism with the statistical program R Version 3.4.3 (R Core Team, 2013) by the single station open-channel method with inverse modeling from *Methods in Stream Ecology Volume2: Ecosystem Function* (Lamberti & Hauer, 2017) *Chapter 34* (Hall & Hotchkiss) using the supplemental R script from Supplemental File 34.3. Additional data needed to complete the calculation included barometric pressure calculated from elevation using the same R script, stream depth obtained from flow measurements and the air-water general gas exchange rate (K600) explained below.

Included in the R script is the option to estimate metabolism e.g. gross primary production (GPP), ecosystem respiration (ER) and K600 directly from the oxygen, temperature and light data which works well for low gradient high GPP streams (Hall Jr. & Madinger, 2018) where K600 is considered a free parameter. Another option is to supply a K600 value and use the model to estimate only GPP and ER. It is recommended that in headwater streams this method is used where the K600 is measured using tracer gas additions (Lamberti & Hauer, 2017). I did not have the tracer gas method available to me so I investigated alternative methods of deriving a K600.

One method I investigated for deriving K600 involved running the model with the option to estimate both metabolism and K600 for all samples. From this data I ran a linear model in R of the diel oxygen data vs the modeled oxygen data and obtained an R2 value for each sample. From these I selected the K600’s from a subset of the models that had +K600, +GPP, –ER and R2>0.95 to run regressions with. Lamberti & Hauer (2017) assert that the model output is erroneous if the GPP is negative or if the ER is positive and Hall Jr., et al. (2015) indicate a negative K600 can not be trusted. I ran regressions with these K600 values against all of the relevant data I had collected and found that average stream velocity had the highest R2. I used the stream velocity vs K600 to derive an equation that was then used to estimate the K600 values for the models that were thrown out:

N=14, R2=0.27, P=0.07

Where *K600* is the general gas exchange rate in units of 1/d and *velocity* is the average stream velocity in m s-1. These models were re-run with the derived K600 values and metabolism was estimated again. The metabolism estimates from all of these models were then kept if they had +GPP and –ER which resulted in 21 models.

The other method I investigated was to derive K600 values from literature data. Hall Jr. & Madinger (2018) suggest there is a strong relationship between stream slope and gas exchange and include slope data along with K600 values as determined by Argon gas injections to the stream. I used their data to derive an equation:

N=8, R2=0.68, P=0.01

Where *K600* is the general gas exchange rate in units of 1/d and *slope* is the stream slope in %. This equation was used to derive K600 values for all of the models which ultimately produced 26 models displaying +GPP and –ER after the inverse modeling was run.

I chose to continue analysis with the model output produced by the literature derived K600 values. I did this because the inverse modeling that estimates K600 as a free parameter is intended for streams that generally have a lower gradient and Hall Jr. & Madinger (2018) suggest that K600 values are unexpectedly high when measured directly in high gradient streams. Although this technique used an equation based on a relationship with a lower sample size (n=8 vs n=14) it had a larger R2, smaller P-value and 5 more usable model output values (26 vs 21).

Statistical Analysis

I developed a generalized linear model for each of the response variables GPP, ER, and trout biomass. The predictor variables were the above mentioned site, hydrologic, and nutrient data (Table 1). I conducted this analysis with R and the ‘lme4’ package (Bates, Maechler, Bolker, & Walker, 2015).

Table 1. List of response and predictor variables shown as random or fixed effects

|  |  |
| --- | --- |
| **Responses** | **Random Effects** |
| GPP (g O2 m-2 d-1) | Watershed (total of 3) |
| ER (g O2 m-2 d-1) | Site (total of 10) |
| Trout Biomass (g m-2) | Sampling Period (total of 3) |
| **Fixed Effects** | |
| Elevation (m) | Discharge (L s-1) |
| Aspect (°South Facing) | Canopy Openness (%) |
| Slope (%) | Daily PAR (mol photons m-2 d -1) |
| Bank Full (m) | Stream Temp. (°C) |
| Pebble Median (mm) | Carbon (DOC mg L-1) |
| Wetted Width (m) | Nitrogen (DIN mg L-1) |
| Depth (m) | Phosphorus (SRP mg L-1) |
| Stream Velocity (m s-1) |  |

I started model selection with a pairwise scatterplot of the response and all predictor variables. Where variables shared a collinearity value of 0.6 and greater, I kept the variable that had the best relationship with the response and removed the other variable from further analysis. I then chose a general linear model (GLM) with several predictors and no interactions and used R’s “drop 1” and “step” functions which returned AIC values associated with each predictor variable. Variables that performed poorly were removed and other unused variables were added in and the process was repeated. After I had worked through the list of variables I had a small subset remaining with which I constructed several different GLM’s. I used R’s “anova” function to compare these GLM’s with one another which returned p-values associated with the comparison. From the best of these models, I then constructed a Q-Q plot, a residual plot and ran an Anderson-Darling test for normality on the residuals. If these results showed evidence of heteroscedasticity or non-normal residuals I moved to a generalized linear model (GZLM). A different GZLM was constructed with the variables in question for each of the random effects listed in Table 1. These were then analyzed with residual plots and the anova function and based on the weight of evidence, the best of these was used in a GZLM that allowed for alternate variance structures. This process of residual analysis and comparison was then repeated for models with variance that was constant, varied by power, exponential, and constant power. If the best of these models (based on p-values and residual analysis) did not appear to meet the model assumptions, the response variable was then transformed with a mild transformation and the process of model selection was started again. I proceeded with model selection in this way working iteratively with stronger response transformations until a model was produced that best met assumptions. I then went back to the non-collinear variables that were not included in the current model and included them as an interaction term one by one and compared these to each other while analyzing the residuals. The best of these was then considered the final model.

**Results**

GPP

I estimated the mean Gross Primary Production (GPP) across all sites and sampling periods to be 0.196 g O2 m-2 d-1 (Figure 2).

N:\Thesis\Rplot.gpp.lit1.tiff

Figure 2. Graph showing mean GPP values for all sites at consecutive sampling periods, error bars represent standard error and letters above bars represent Tukey’s honest significant difference categories at p= 0.020 while the model p< 0.0001 indicates the significance of sampling period for the GPP model.

The model I determined for GPP was a linear mixed effects model with a square root of a log transformation with site as a random effect. The main effects were determined to be sampling period (Figure 2) and depth (Figure 3) both with a model p-value<<0.05. GPP did not appear to be related to daily amounts of light (PAR) or concentrations of nutrients (DOC, DIN, SRP) or was weakly negatively associated.

N:\Thesis\Rplot.depth.t.gpp1.tiff

Figure 3. Graph showing regression of transformed GPP and depth with an adjusted R2 of 0.13 and p<0.0001 from the GPP model.

ER

I estimated the mean ER across all sites and sampling periods to be -10.287

g O2 m-2 d-1 (Figure 4). Although ecosystem respiration is noted as a negative number because it is thought of as a subtraction of oxygen from the environment, it will be discussed here in terms of its absolute value (positive) to facilitate modeling and conceptualization.

N:\Thesis\Rplot.er.lit1.tiff

Figure 4. Graph showing the absolute value (magnitude) of mean ER values for all sites at consecutive sampling periods, error bars represent standard error and letters above bars represent Tukey’s honest significant difference categories at p= 0.052.

The model I determined for ER was a linear mixed effects model with a log transformation of the absolute value and site as a random effect. The main effects were determined to be depth (Figure 5) and slope (Figure 6). ER magnitude did not appear to relate positively to nutrients (DOC, DIN, SRP) however ER did relate to GPP (Figure 7).

N:\Thesis\Rplot.depth.t.er1.tiff

Figure 5. Graph showing regression of transformed ER and depth (m) with an adjusted R2 of 0.36 and p< 0.0001 from the ER model.

N:\Thesis\Rplot.slope.t.er1.tiff

Figure 6. Graph showing regression of transformed ER and Slope (%) with an adjusted R2 of 0.57 and p< 0.0001 from the ER model.

N:\Thesis\Rplot.gpp.er1.tiff

Figure 7. Graph showing regression of ER and GPP with an R2 of 0.41 and a p-value of 0.00026.

Fish

I sampled a total of 230 westslope cutthroat trout (*Oncorhynchus clarkii lewisi*) and 4 eastern brook trout (*Salvelinus fontinalis*) with a minimum fish length of 50 mm, median 79 mm and a maximum length of 215 mm (8.5 inches). I estimated the trout density in fish per meter of stream length of all sites and both fish sampling periods to range from 0 in First Cr. (Taneum Watershed) 2018 to 1.33 fish m-1 in Standup Cr. (Teanaway Watershed) 2018 (Figure 8). The mean trout mass per individual fish ranged from 3.58 g in Frost Cr. (Taneum Watershed) 2017 to 31.23 g in Jack Cr. (Teanaway Watershed) 2017 (Figure 9). I estimated trout biomass in g m-2 to range from 0 in First Cr. (Taneum Watershed) 2018 to 8.38 g m-2 in Hurly Cr. (Swauk Watershed) 2017 (Figure 10).

N:\Thesis\Rplot.trout.density1.tiff

Figure 8. Graph showing trout density (fish m-1) for each stream and year of sampling with streams arranged by increasing wetted width and grouped by watershed. Error bars represent standard error.

N:\Thesis\Rplot.trout.mass1.tiff

Figure 9. Graph showing the mean mass of individual fish per stream and year of sampling with streams arranged by increasing wetted width and grouped by watershed. Error bars represent standard error.

N:\Thesis\Rplot.biomasssquared1.tiff

Figure 10. Graph showing trout biomass in g m-2 per stream and year of sampling with streams arranged by increasing wetted width and grouped by watershed. These values were arrived at by multiplying the fish density per meter of stream by the mean weight of the individual fish, the result was then divided by the stream wetted width. Error bars represent the standard error of fish density multiplied by the average fish weight.

The model I determined for trout biomass in g m-2 of stream surface area was a generalized least squares with a log transformation. The variance was allowed to change based on an exponential function and the main effects were watershed (Figure 11) and minimum daily temperature with an interaction on canopy openness (Figure 12). Trout biomass did not appear to relate to nutrients (DOC, DIN, SRP), light (PAR), or ecosystem metabolism (ER, GPP).

N:\Thesis\Rplot.t.cut.mass.m.basin1.tiff

Figure 11. Graph showing transformed trout biomass by watershed with the trout model p-value of 0.0007. Error bars represent the minimum and maximum estimates within the watershed.

N:\Thesis\Rplot.interaction1.tiff

Figure 12. Graph showing transformed trout biomass by water temperature category and canopy openness category. There appears to be lower trout biomass at higher minimum temperatures and higher biomass for more open canopy streams at medium temperatures (minimum stream temperature from 8.1 to 9.5°C), error bars represent the minimum and maximum estimates.