Biol 419

**Introduction:**

Pseudo-replication, wherein an experimenter takes multiple measurements from a single experimental unit, is one of many pitfalls endemic to experimental design and can lead to an experimenter reaching improper conclusions and incorrect inference due to the inability to distinguish the difference between treatment and individual unit variability. Pseudo-replication is of special interest in ecological experiments, as Hurlbert identified that of 176 ecological studies conducted in a 17-year period that applied inferential statistics, 48% of them had pseudo-replication occurring in the design, producing potentially incorrect results (Hurlbert, 1984). Pseudo-replication has multiple causes, including limitations caused by experimental material, budgetary or time constraints, and other limiting factors.

One potential solution to the pseudo-replication is the use of a blocking structure in the experimental design. Blocking is the creation of subsets within an experimental unit, inside of which variability is expected to be roughly homogenous, but variability is expected to be heterogeneous between individual blocks. This heterogeneity between blocks allows the attribution of experimental variance to a block effect, reducing background variability and improving the precision of treatment comparisons (Welham et al, 2015). The decision for using a blocking structure requires information upon the type and magnitude of heterogeneity within proposed blocks.

Our data analysis project focuses on the design stage of a laboratory disease challenge study, which is attempting to identify species based sensitivity in abalone to the bacterial disease *Candidatus Xenohaliotis californiensis*, or Abalone Withering Syndrome. This disease is nearly always fatal upon infection, though is endemic to a large portion of the west coast of the United States, having been observed as far north as Crescent City, California; and continues to march northward, aided by increasing ocean temperatures due to global climate change and other factors (Crosson et al., 2014).

We have chosen four species, all of which are native to the endemic range of the disease, and all of which are currently the focus of aquaculture efforts in the range of the disease and are the Red (*Haliotis rufescens),* Pink (*H. corrugata),* Pinto (*H. kamtschatkana kamtschatkana),* andBlack (*H. cracherodii)* abalones*.* We chose these species based upon two different main considerations, economic and ecological. Red and Pink abalone have current commercial aquaculture importance, with sales over 3 million dollars annually in California alone from the six operating commercial aquaculture facilities (McBride and Conte, 2000). Black and Pinto are not currently commercially exploited; instead the aquaculture is focused on species restoration, as the Endangered Species Act (ESA) lists both species, with the Black abalone listed as endangered and the Pinto abalone listed as threatened.

Given that aquaculture facilities are hypothesized to serve as incubators, and potentially amplifiers of disease as well as the additional stresses placed upon animals in aquaculture conditions potentially weakening the immunes systems of animals under culture, information regarding relative susceptibility of species to the disease in question could help to inform growers, both commercial and restoration focused, in which species to focus their limited space and budgets on, maximizing their effects (Crosson, *private communication).* In aid of that goal, we have designed an experiment built around a small scale recirculating sea water system, with multiple sub-aquaria, in which we can challenge animals with the disease and measure any physiological changes associated with the disease, as well as time to mortality. Our experiment, which is severely limited in both space available, as it must be conducted in a WDFW permitted biohazard space, as well as in number of animals available, especially the Black and Pinto abalones, given their status as listed on the ESA and having to source them from growers who are highly reluctant to part with animals for experimentation, for obvious reasons. Given these limitations, we have designed an experiment built around a blocking structure to maximize the chance of seeing a likely small species effect on time to mortality when infected. In pursuit of this goal, we have designed a suite of remote sensors that can measure variability over time in temperature and water flow, between individual sub-aquaria, which will assist in the justification of both our design and subsequent analysis, if it is found that variability is heterogeneous between sub-aquaria based blocks. Upon sensor suite development, it was found that the water flow sensing capability was technically unfeasible due to limitations of microcomputer processing capabilities so our line of inquiry was scaled back to include temperature alone.

**Materials and Methods:**

Experimental Apparatus:



Figure 1: Recirculating sea water system, with individual sub-aquaria.

We designed a simple recirculating sea water system from commonly available containers and PVC plumbing piping to serve as the basis for our disease exposure experiment. Water begins in the large green sump, is pumped via a 300 gph marine pump (Supreme Classic model 3), through a UV sterilizer (Coralife 6x Turbo-Twist, pn: ENE15601), then through a 25 micron physical filtration system (Red Sea Ocean Clear Canister Filter, pn: 340), then a biological filtration system (Red Sea Ocean Clear Biofilter, pn: 325), and finally into an aquarium chiller for temperature regulation (JBJ Arctica, pn: CD-19155) before circulation into a head tank, in which known infected animals will be contained. Water is distributed from the head tank into the individual sub-aquaria via a distribution manifold

Sensor System:

Our sensor system is based upon a Raspberry Pi 2, rev. B microcomputer. Temperature sensing capability is provided via a one-wire thermometer (Maxim Integrated pn: DS18B20) and water flow is measured via a Hall effect water flow sensor (Adafruit pn: 828). Calibration of the thermometers was conducted by placing them in an incubator set at three different temperatures, set as the minimum, maximum, and mean temperature expected to be encountered during the duration of the experiment. We ran linear regression on each group and found the intercept difference from actual temperature and used that as a calibration offset for each probe. Python code for operating the thermometer is heavily adapted from Adafruit’s Raspberry Pi Lesson 11 (Monk, 2015) as it both polls the temperature sensor for instantaneous temperature, as well as provides a framework to interact with an Internet of Things (IoT) API and website for aggregation and curation. We elected to use [www.thingspeak.com](http://www.thingspeak.com) for our IoT API due to a relatively robust and well-documented API, with reasonable frequency of API call limitations and data storage durations. The thingspeak API limits us to API calls every 15 seconds, so we elected to save temperature data every 16 seconds, to allow for any inconsistencies in upload timings.

Analysis:

Once our data was collected, we analyzed it via two methods. To determine raw difference in mean we initially chose a one-way ANOVA, with sub-aquaria as factors and initially, mean temperature and water flow as response variables, but this was scaled back to temperature alone due to technical issues. One-way ANOVA identifies differences between three or more groups based upon an assumed F-distribution. For time-series analysis we elected a smoothing via moving average. A smoothing such as this allows us to identify long-term trends by creating a series of averages from different subsets of the data, with the subset being created by specifying some “lag”, which represents the number of previous points incorporated into the individual average. This process results in the loss of the initial data points equal to your lag, but produces a smoothed product, with the intent of minimizing local distortions and allowing for the identification of longer-term trends in the time series. We chose to use the SMA() function from the TTR package in R, but upon analysis of the source code found it to simply be a wrapper for the runMean() function, found in the base R package (Ulrich, 2015; R Core Team, 2015). We compared results from both functions, and found them to be equivalent for all use cases.

Results:

Actual ANOVA

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source of Variation | Degrees of Freedom | Sum of Squares | Mean Square | F- Ratio | p-value |
| Aquarium | 3 | 371 | 123.75 | 956.8 | <2e-16 |
| Residual | 37388 | 4836 | 0.13 |  |  |

Expected ANOVA

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source of Variation | Degrees of Freedom | Sum of Squares | Mean Square | F- Ratio | p-value |
| Aquarium | 3 | 368 | 122.65 | 1038 | <2e-16 |
| Residual | 37388 | 4371 | 0.12 |  |  |

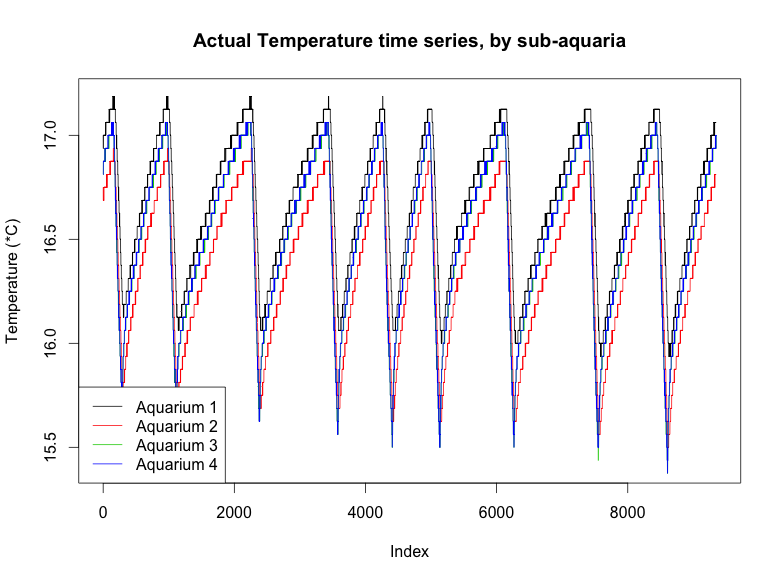


Figure : Actual temperatures recorded

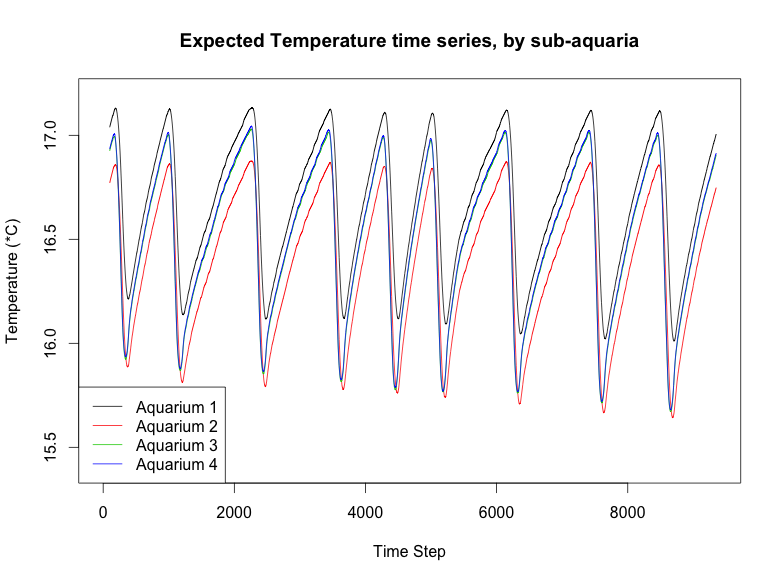


Figure : Expected temperature recorded

**Discussion:**

**-**Talk about how the processor priority of the Raspberry Pi made it impossible to process water flow reliably

- Talk about how temperature probe claims to read to 1/1000th of a degree, but only reads in 1/64th degree increments leading to the uncharacteristic steps found in the actual data

- Talk about how I probably figured out the root cause of the difference in the temperature for the hot tank at least. It’s the one closest to the chiller, so heat thrown off from the chiller may be artificially warming the tank

- Talk about how we plan a second data collection run, with sensor interaction offloaded to an arduino based microcontroller for functioning water flow

- Results interpretation: If we accept that the chiller induced heat is the cause for the warm tank, we likely have three tanks that are the same, one tank that is different (low). Low temperature one is still a mystery. As it stands, based on temperature alone we don’t have sufficient evidence to suggest that there is heterogeneity in variance enough to call our sub-aquaria blocks. Water flow may make this different.

- Further inquiry: Can we induce structural variability and not break statistical rules? I don’t honestly know the answer to that (IE, can we set one at a temp of 15, one at a temp of 16, … and then block based on that? I haven’t encountered that, so I think it’s a little sketchy). We may also look at characterizing suspended organic matter from the head tank down (Disease is fecal orally transferred, so we’re spreading poop, if different levels of poop comes out of each pipe, that’s reason for blocking. This would require multiple qPCRs for bacterial detection, or collection of fixed volume of liquid for each aquaria, spinning down of suspended material, drying, and weighing).

- Hardware programming/debugging is hard as balls.

- I generated the expected by doing a ridiculous smoothing, like a 27 minute moving average. Smooths out though and looks pretty!

References:

Crosson, L.M., Wight, N., VanBlaricom, G.R., Kiryu, I., Moore, J.D., and Friedman, C.S. 2014. Abalone withering syndrome: distribution, impacts, current diagnostic methods and new findings. Dis. Aquat. Organ. **108**(3): 261–270. doi: Doi 10.3354/Dao02713.

Hurlbert, S.H. 1984. Pseudoreplication and the Design of Ecological Field Experiments. Ecological Monographs 54(2): 187–211. [accessed 11 March 2016].

McBride, S., and Conte, F.S. 2000. California abalone aquaculture. Calif. Aquac.: 1–5. doi: ASAQ-A10.

Monk, S. 2015. Adafruit’s Raspberry Pi Lesson 11. DS18B20 Temperature Sensing. Available from https://learn.adafruit.com/adafruits-raspberry-pi-lesson-11-ds18b20-temperature-sensing/overview.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Welham, SJ., Gezan, SA., Clark, SA., and Mead, A. 2015. Statistical methods in biology. *In* CRC Press. doi: 10.1002/0471732877.emd318.

Ulrich, Joshua. (2015). TTR: Technical Trading Rules. R package version 0.23-0. https://CRAN.R-project.org/package=TTR