# There once was a grid at ol' Carkeek

First Author[*,1], Second Author[1,2], and Third Author[2]

[1]Department of Computer Science, LaTeX University

[2]Department of Mechanical Engineering, Superfabulous University

March 4, 2016

## 1 Keywords

2 Stuff, things, neat, cool, wow, instafun, tags4likes, etc

## 3 Abstract

4 This is the text of the abstract.

## 5 Introduction

6 Biodiversity surveillance is being revolutionized by DNA-based detection of organisms from en-

7 vironmental samples. ?(specifically speed and scope of ecological studies). Many researchers are

8 justifiably cautious about the ?(adoption) of this new form of data. Their apprehension is rooted

9 in the premise that traditional survey approaches are more accurate because the chain of infer-

10 ence between observation and ecological data is usually short: A researcher sees two swans in Lake

11 Hopatcong and infers the lake is occupied by at least 2 swans. DNA based surveys, on the other

12 hand, consist of a longer chain of inference ?(including several links at which the links are complex

13 or unknown)?: Dna sequences are observed from a sequencer, the sequencer sequences products of

---

[*]first.author@funstuff.com

14 a polymerase chain reaction (PCR), PCR amplified sequences from a purified DNA sample, DNA is

15 purified (extracted) from an environmental sample, environmental samples contain DNA from or-

16 ganisms present, the organisms present are representative of the biological community about which

17 we wish to make inference. (Could tie this to the swans of Lake Hopatcong). Clearly, this process is

18 more complex than visual surveys. ?(and includes several steps at which the relationship between

19 stages is complex or unknown)

## 20 Methods

21 We use the general framework outlined by Shelton et al (CITE). That study outlined the structure

22 for estimation of the proportional biomass of a taxon ($B_i$) given the proportional counts of sequences

23 recovered from a parallel sequencing run ($Z_i$).

24     We modeled the counts of DNA sequences ($Z$) from each of a given taxon $i$, in each replicate

25 PCR $j$, from each replicate of a given location $k$ (hence, $Z_{ijk}$), as though they are ?(proportional

26 to/drawn from)? a Poisson distribution. A Poisson distribution is described by one and only one

27 parameter, $\lambda$, which is equal to both the mean and variance. Because in this case our modeled

28 values are discrete counts, we use the natural exponent, $e^{\lambda}$. Thus,

$$Z_{ijk} \sim Poisson(e^{\lambda_{ijk}}) \tag{1}$$

29     In turn, we further assume this parameter $\lambda$ is linearly proportional to a suite of taxon-, pcr-,

30 and site- specific parameters describing the variance associated with each sub-process linking the

31 amount of DNA ($Y$) of a given taxon $i$ at a given location $k$ in a DNA extract (hence $Y_{ik}$):

$$\lambda_{ijk} = \beta_0 + \beta_i + \eta_{ijk} + \epsilon_{ijk} \tag{2}$$

32     Where $\beta_0$ is a general intercept across all taxa, $\beta_i$ is a fixed effect accounting for the variance

33 associated with taxon $i$, and $\eta_{ijk}$ and $\epsilon_{ijk}$ are random effects of variance resulting from the processes

34 associated with PCR and spatial location, respectively.

## Results

We found that if you have two apples, and someone gives you another two apples, you have four apples.

## Discussion

Boy those results sure are neat. Now, the pressing question becomes: How do you like them apples?

## Acknowledgements

We wish to thank all of the little people.

## Funding

This study was funded by our super-rich uncle.

## Author Contributions

Conceived and designed the experiments: James L. O'Donnell, Ryan P. Kelly, A. Ole Shelton. Collected the data: James L. O'Donnell, Greg Williams, Natalie C. Lowell, Ryan P. Kelly, A. Ole Shelton, Jameal F. Samhouri. Conducted the analyses: . Wrote the first draft: . Edited the manuscript: .

## Data Availablity

The data and code used to generate our results can be found at the following url:

## Figures