An Analysis of Long-term Data Consistency and a Proposal to Standardize Flower Survey Methods for the EISI Pollinator Project

A. Sanders[1], M. Childs[2], E. Traub[3], J. Jones[4]

[1]Arizona State University, School of Life Sciences, Tempe, AZ

[2]Whitman College, Walla Walla, WA

[3]Humboldt State University, College of Biological Sciences, Arcata, CA

[4]Oregon State University, College of Earth, Oceanic, and Atmospheric Science, Corvallis, OR

Collections of long-term ecological data can be used to answer important and unique questions about natural systems due to their temporal breadth. However, to reach meaningful conclusions, the integrity and robustness of these data must be maintained from the start. It is necessary to perform routine maintenance and assess the uncertainties in methods that arise from observer error and changing methods over time. In this analysis, we analyze variation in flower survey data taken from 2011 to 2014 at the H.J. Andrews Forest as a part of a long-term project mapping plant-pollinator interactions in montane meadows. By separating the data to see how observers counted flowers per stalk and stalks per plot, we found that total abundance tends to stay relatively stable between observers even though observers may be counting flowers and stalks differently. Also, experienced observers accompanying new observers can help maintain consistency, but does not show enough consistency to correct for uncertainty from observer turnover. Distributions across species are very different, indicating the variety of flower types and their varying degrees of difficulty in counting. Adjusting protocols after evaluations of existing data can sometimes compromise consistency, but will be better for the collection in the long run. In an attempt to minimize uncertainties, the current study also contains a standardized protocol and a flower identification key unique to this particular system.

Introduction

Long-term research at the H.J. Andrews Experimental Forest, like that at other sites in the Long-Term Ecological Research (LTER) network, requires reliable data that is used to understand processes relevant to forest management and ecosystem function (Franklin et al., 1990). Data that span large temporal scales demand careful entry and consistent protocols to minimize the effects of observer turnover and changing methods over time (Beard et al., 1999) (Magurran et al., 2010). A common problem with long-term data is its compatibility with other studies of the same scope, which makes it all the more important to design protocols that are thorough yet still temporally robust (Magurran et al., 2010).

The Eco-Informatics Summer Institute (EISI) hosted by Oregon State University has recorded five years of interactions between pollinators and flowering plants in 18 montane meadows throughout the H.J. Andrews Forest. Projects derived from this program aim to fit statistical models of pollinator preference using flower survey and interaction data. To preserve the strength of the data, it is imperative to routinely assess the accuracy of the data already compiled as well as standardize the protocol for future accuracy. The present study aims to pinpoint sources of observer error among other sources of variation and provide a detailed methods protocol for future data consistency.

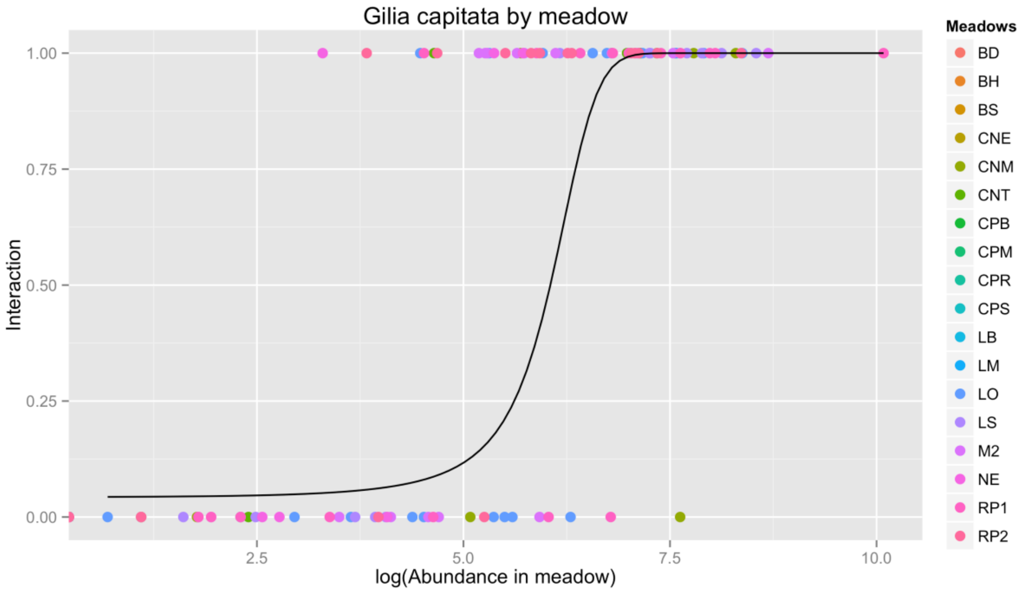


Figure 1. Childs 2015 visualized the positive effects of flower abundance on number of interactions with flower species *Gilia capitata*. There are, however, interactions that happen even when flower abundances are very low (far left) as well as no interactions when abundances are high (far right). The horizontal axis is also logarithmically-transformed, which has compressed abundance values and in turn, minimized drastically large or small values. Variable abundances suggest the data might contain unnecessary error.

It is evident that flower abundance may be related to interaction connectivity (Figure 1). However, the abundance data are very noisy and it would be helpful to reduce uncertainty in these data if possible. Figure 1 is an impetus for the questions we had about data consistency. Over-counting *Gilia capitata* abundance, which is a common mistake under the current protocol, can result in points like the green one in the far bottom right that illustrate high flower abundance but no interactions.

Some of the uncertainty may be due to avoidable observer error. Observers may have used different rules to identify and count flowers and stalks. In addition, there has been little to no communication between observers throughout the years and very few observers have participated in data collection for more than one year. Specific protocols for counting flowers have not been formalized or presented to each cohort of students, nor was there strict standardization within our own 2015 cohort before data collection started. Because of this lack in standardization, it is possible that records of number of flowers per stalk and number of stalks per plot may vary by observer, affecting total flower abundance, which is determined by multiplying flowers and stalks. We hypothesize that avoidable variation at the observer level in what is perceived to be a stalk and what is perceived to be a flower may not necessarily affect the total abundance, but could limit future studies using these data.

Estimation of total flower abundance involves several steps (Figure 2). The first bifurcation is determined by the observer's ability to correctly identify species. Once a flower is identified, its abundance then relies upon the observer's ability to determine the number of flowers per stalk and the number of stalks per plot using the same protocol as other observers. This is also influenced by natural variability, which is ultimately the data we try to isolate. These analyses aim not only to identify how observers have been counting, but how much of this variation is due to natural variability of the plants themselves.

Because communication is a key element in data consistency, we also hypothesize that the retention of someone who has collected data for more than one year will have a positive influence on counting consistency and increase the ability to see natural population fluctuations (Vos et al., 2000).

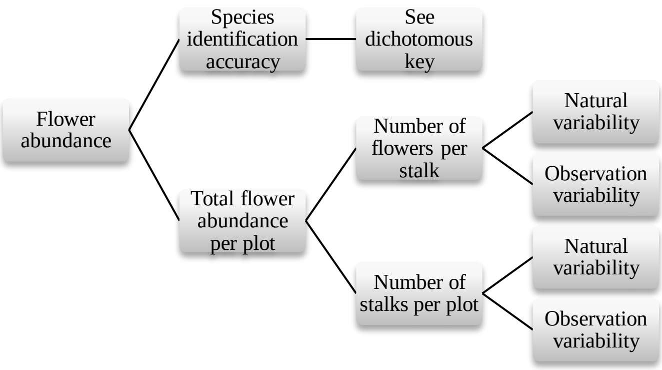


Figure 2. Flower abundance depends on a host of different variables, some desirable such as natural variability, and others undesirable, such as observation variability.

Methods

Flower survey data was recorded from 2011 to 2014 during five weeks in the summer of each year. There are data for 18 meadows, but the range has since been limited to only 12 across three complexes, which are distinct areas across the H.J. Andrews Forest that contain meadows. Each meadow has 10 3x3 meter plots. Flower surveys preceded pollinator-interaction data collection and were conducted by counting and identifying all flowers whose reproductive parts were available to pollinators. Then individual flowers per stalk and stalks per plot were recorded.

We examined flower abundance data by observer from 2011 to 2014 for selected plant species. Although 130 flower species were observed in this time frame, only seven species were isolated for this study. The species selected were *Rumex acetosella*, *Ligusticum grayi*, *Gilia capitata*, *Eriophyllum lanatum*, *Eriogonum compositum*, *Achillea millefolium*, and *Eriogonum umbellatum*. These species each have different structures and reveal varying degrees of ambiguity in counting flower abundance (Figure 3). Plots display median and range by observer in box-and-whisker plots using R software. Three types of data were plotted for each species: number of flowers per stalk, number of stalks per plot, and total flower abundance per plot. Observers are only present on a figure if they collected data for that species, which is why the quantity and selection of observers changes by species. Each observer was assigned a number to maintain anonymity.

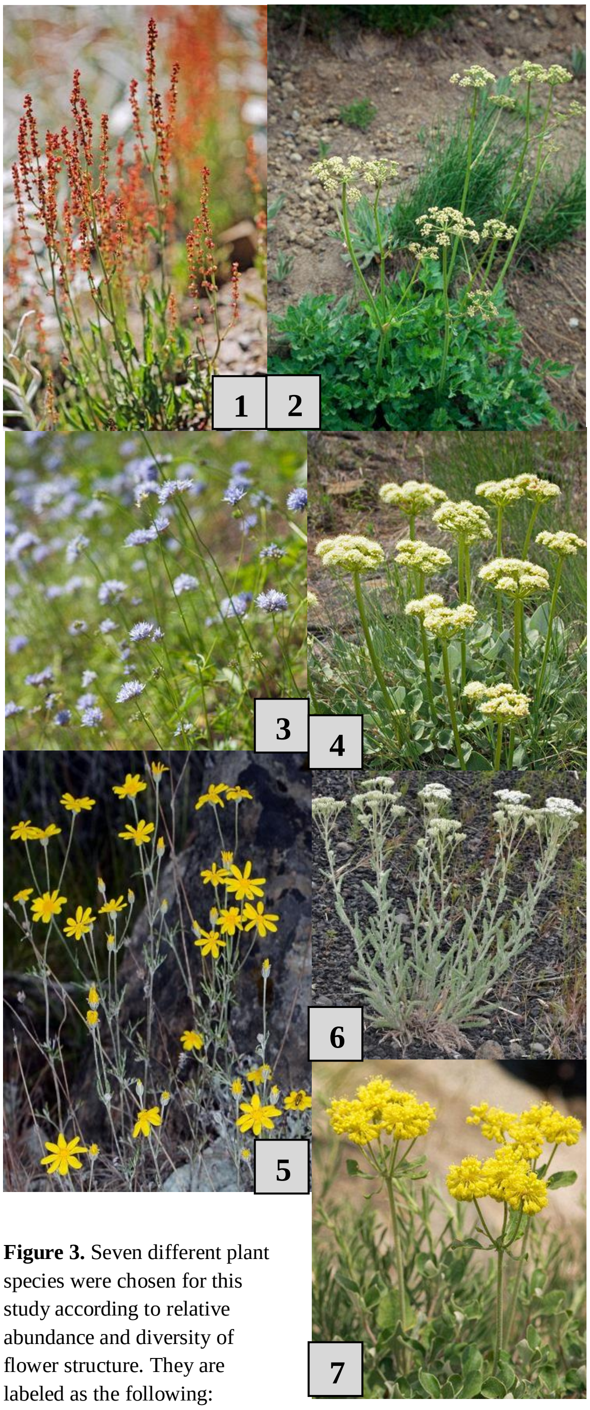


Figure 3. Seven different plant species were chosen for this study according to relative abundance and diversity of flower structure. They are labeled as the following:
1. *Rumex acetosella,*
2. *Ligusticum grayi*
3. *Gilia capitata*
4. *Eriogonum compositum*
5. *Eriophyllum lanatum*
6. *Achillea millefolium*
7. *Eriogonum umbellatum*

We also examined variation across year holding observers constant by comparing results from individual observers (Observers 7 and 49) who collected data over more than one year.

Results

The first observer of *Ligusticum grayi* in 2011 counted a median of nearly 5 stalks per plot and a median of around 300 flowers per stalk, but the third observer in 2013 counted a median of nearly 80 stalks per plot and a median of around 5 flowers per stalk, producing similar total flower abundance of nearly 1300 (Figure 5). Each year a few observers counted flowers per stalk differently than the rest of the group (Figure 5). Overall, in 2014, observers of *Ligusticum grayi* recorded higher medians in flowers per stalk and total flowers per plot (Figure 5).

Counts of flowers per stalk and stalks per plots did not vary much among observers or year for *Gilia capitata*. There is little variation within each plot, with most variability in the number of stalks per plot (Figure 6). Observers in 2014 recorded larger medians for flowers per stalk than other years (Figure 6).

Observers of *Rumex acetosella* recorded declining median flowers per stalk from 2011 to 2013, but stalks per plots and total flowers per plot did not change considerably (Figure 4). In the survey distributions of *Eriophyllum lanatum*, the data contain very high median values and outliers for flowers per stalk in 2013 and stalks per plot in 2011 (Figure 7). Medians of total flowers per plot stay fairly consistent, except the lower values across all observers in 2012.

In the *Eriogonum compositum* distributions, flowers per stalk vary widely each year and between observers. Median number of stalks per plot hikes in 2014, but total flowers per plot values level out, except for the lower median values in 2013 (Figure 8). Counts for *Achillea millefolium* stay consistent in all distributions, with intra-year variation lower than inter-year variation in flowers per stalk (Figure 10a). Outliers for this species are particularly high across all observers (Figure 10c).

Similarly, observers counted *Eriogonum umbellatum* consistently across years, however in each year there was usually one observer who counted differently. These distributions are also characterized by the abnormally high outliers.

A closer look at the breakdown of how total flower abundance is calculated reveals that even though different observers use different methods for counting flowers per stalk and stalks per plot, the total abundance does stay somewhat consistent. We often see two different observers who count inversely to one other (one counting more flowers, the other counting more stalks) but end up with the same total abundance of flowers per plot, evidenced by Observers 2 and 27 in the *Ligusticum grayi* figures (5). This phenomenon even happens across years as well. With *Ligusticum grayi*, it is apparent between 2011, 2013, and 2014 (Figure 5) and for *Eriophyllum lanatum*, it is apparent between 2011 and 2013 (Figure 7). The total flower abundance, therefore, is somewhat stabilized by this phenomenon because it is the product of the two inverse values. This discovery also suggests that the two metrics flowers per stalk and stalks per plot are not accurate measurements in themselves and may even be subject to additional variability by observer.

Observer 49 collected data in 2011 and 2012. In 2011, Observer 49 counted much higher flowers per stalk for *Rumex acetosella* than in 2012, and most of the other observers appeared to change their counting method between 2011 and 2012, resulting in reduced estimates of flower abundance from 2011 to 2012 for this species (Figure 4c). A similar change occurred for *Eriogonum compositum* (Figure 8). Observer 49 counted the other 5 species consistently between 2011 and 2012 (Figures 5, 6, 7, 9, and 10). Observer 49 also observed *Achillea millefolium* with the rest of the cohort in 2013 and most other observers recorded similar numbers in that year (Figure 9).

Observer 7 was present in most species distributions and was recorded as assisting several other observers. A joint effort was recorded as a separate observer. Observer 7 stayed very consistent counting flowers per stalk for every species they counted two years in a row (Figures 5-10). Observer 7's counting was never wildly aberrant among other observers in the same year as well (Figures 4-10). Observers 7 and 49 counted flowers per stalk similarly across all species except *Eriogonum compositum* (Figure 8) and compared to Observer 7, Observer 49 recorded more outliers (Figures 5, 7, and 9).

Discussion

Among the seven species examined, consistency of flower counting was highest for *Gilia capitata* an abundant, compound flower whose flower and stalks were easily identifiable and countable. Because there are several species that are also straight-forward to count, observer error will not be as evident, which suggests there are at least some reliable abundance data. On the other hand, the species with the greatest inconsistency was *Eriogonum compositum* because its flowers are small, abundant, and in an oddly-shaped arrangement. Species that are more difficult to count are commonly where error occurs in the data.

A source of inconsistency in the data is differentiating the observer error from the actual flower population fluctuations. We would expect that the presence of an experienced observer would result in more consistent median flowers per stalk values among observers, because that person would guide the methodology. Therefore, we would expect any variation in their median flowers per stalk values to indicate actual flower population fluctuations.

With both *Ligusticum grayi* and *Rumex acetosella*, Observers 7 and 49 did not record data in 2013 and 2014 and the median flowers per stalk values increased/decreased respectively. This change suggests some "drift" in the rules used to count these species, which may be due to the lack of an experienced observer to train novice observers for these difficult-to-count species.

However, the presence of an experienced observer did not necessarily guarantee consistency. Even though Observer 49 recorded data for *Eriophyllum lanatum* in 2013, median values became inconsistent with large outliers in flowers per stalk in this year (Figure 7a). Nevertheless, observer continuity from one year to another appears to have a positive impact on counting consistency at least for the first two years of data collection.

Following observers that collected data for more than one year assumes that they are using the same counting method each year, which may not be the case. For *Eriogonum compositum* and *Rumex acetosella*, Observer 49's median flowers per stalk values decreased from 2011 to 2012, and appeared to produce an overall reduction in the flowers per stalk of all observers, indicating that this person may have been an experienced observer guiding the novice observers (Figures 4 and 8). If we assume that the experienced observers were using the same counting methods each year for *R. acetosella*, that would mean that there was a median value of 60 flowers per stalk in 2011, then a median value of 15 flowers per stalk the next year (Figure 4a). *R. acetosella* plant structure does not naturally fluctuate that widely, which suggests that a change in protocol was made without corrections to prior data. This shift makes total abundances decrease dramatically when this might not actually be happening in reality (Figure 4c). A similar trend is seen for *E. compositum*, which was recorded to contain around 200 flowers per stalk in 2011 and 25 flowers per stalk in 2013 (Figure 8a). In addition, Observer 7's consistency did not change in these years like that of Observer 49. Technically Observer 49's trend would indicate natural fluctuations because he/she was present for these years, but change in counting technique could be a better explanation.

Furthermore, evidence of different counting methods each year exists in the distribution for *Achillea millefolium*, which is usually a difficult species to count. Even though Observer 49 stayed for 2013, there is still a substantial difference in fluctuations of flowers per stalk (Figure 9a). In both 2011 and 2012, median values stay consistent, but in 2013, ranges elongated and outliers became far more scattered (Figure 9a). One observer counted 350 flowers per stalk, which is an anatomical possibility, but it does not seem to agree with the counting methods of other observers in that year and other years (Figure 9a).

Fluctuations like this are widespread throughout the data. There is always at least one observer in each year who recorded an aberrant median value for either flowers per stalk or stalks per plot for *Eriogonum umbellatum* (Figures 10a and 10b). Observer 37 recorded a median value for stalks per plot that is almost 6 times larger than everyone else's (Figure 10b) and Observer 21 recorded almost 1200 flowers per stalk in 2013 (Figure 10a). These sorts of outliers are questionable and will never be corrected for because the counting methods were never explained for each year. We can fairly associate some of the noise in the total flower abundance data to particularly abnormal values like these.

The dilemma of extricating the natural fluctuations from avoidable and unavoidable observer error is certainly not new (Vos et al., 2000). Biodiversity monitoring in particular can be misleading because of imperfect detectability that makes population levels seem more dramatic than they actually are (Kéry and Schmid, 2004). We operate under several assumptions about sampling consistency and randomness when, for example, one observer in one year who over-counts could make the abundance of one species incorrect, yet still a data point that must be used. In the context of pollination interactions, it may make preference scores for this species lower because according to the data, there were "more" flowers of this species to choose from. Following data collected by one observer over several years is supposed to provide somewhat of a baseline, but in this study, there are too few of these observers and too much variation in the observers' consistency to establish a baseline.

The variation between observers in the same year is a difficult metric because some observers only conducted a few watches over the season and others observed all meadows throughout the season. To diagnose any variability, we need to assume that all observers in a year randomly conducted surveys in plots in all meadows, which may not be an accurate assumption. In addition, some observers may have only data from the beginning of the season when certain flowers were more successful in one plot while others have data from the whole season at every plot. Very high outliers may be the result of an unusual plot or an unskilled observer. These values may average out in some cases, but they create noise in analyses like Figure 1 and should be minimized.

In summary, total flower abundance data from 2011 to 2014 include obvious inconsistencies among observers and among years, limiting their utility for analysis of population abundances. This creates challenge for interpreting how flower abundance affects pollination (e.g., Figure 1). A more detailed protocol could reduce noise and create a sharper relationship.

Conclusions, Suggested Protocol and Dichotomous Key for Flowering Vegetation

By providing evidence that our flower survey data could be inaccurate, we are taking necessary steps in data housekeeping, which is a common practice when working with long-term data. Vos et al. (2000) detail steps that ecological monitoring projects can follow in order to ensure the livelihood of their work. The analysis mainly focuses on projects with policy objectives in mind, but a few main suggestions in the outline are particularly relevant to the EISI Pollinator Project, though no policy is involved yet.

First, at a monitored site, metrics should initially encompass most of the system until hypotheses are developed and metrics are adjusted according to usefulness and work efficiency (Vos et al., 2000). The flower surveys started out very broad but are starting to dial in on what we need according to a working hypothesis; flowers will now be counted as statistical units to cater to probability models instead of botanically accurate abundances (See Sample Protocol). We do, however, still understand that our study needs to accommodate more current and future hypotheses, which is why the surveys and interaction data are still exhaustive in each plot. Changing the protocol is not inherently bad if it makes the data collection easier and more accurate/useful in what they refer to as the "highest sustainable quality" (Vos et al., 2000).

Second, data are subject to observer turnover and should be as least subjective as possible (i.e. using counts rather than estimates) (Vos et al., 2000). Because we now understand some of the sources of avoidable variation in the flower survey data, a detailed protocol for future data collection should help remedy some of the uncertainty. Attached is a protocol proposal to be read prior to data collection in the future. In addition to this written document, it would be helpful to have a representative from the year's past to accompany the team into the field and host practice surveys (Vos et al., 2000). Field workers should each practice on the same plot then compare results until a mutual understanding is reached. For example, during the orientation week, an experienced observer can take everyone who will be collecting data to a meadow, equipped with all necessary tools for a normal watch. The experienced observer will introduce counting methods and relevant species names, as well as perform his/her own flower survey audibly on one plot. Next, everyone will conduct a trial survey on a new plot, leaving each observer to count and record independently on a data sheet. Once everyone has finished, the group can share results and justifications for any discrepancies. More trials with this technique can be repeated until all observers are consistent with each other. An open line of communication is also important to maintain when observers find new species throughout the season. New methods for certain difficult species should be added to the protocol for future reference. These techniques serve to narrow observer error and make the observations less subjective.

Along with the protocol is a flower key for all the observed species in the data. It is a user-friendly guide and is open to additions if new species are encountered. These documents will not only provide reference to future field workers for data consistency, but can be made available for other scientists who are interested in working with the data we collect.

Acknowledgements

Thank you to all of my team members, Kate Jones, Julia Jones, Tom Dietterich, Alix Gitelman, Rebecca Hutchinson, and the H.J. Andrews Forest.

Works Cited

Beard, G.R., W.A. Scott, and J.K. Adamson. 1999. The value of consistent methodology in long-term environmental monitoring. **Environmental Monitoring and Assessment** 54: 239-258.

Childs, M. 2015. Predicting Plant-pollinator Interactions. Unpublished draft. EISI Pollinator Project. Available upon request.

Franklin, J.F., C.S. Bledsoe, and J.T. Callahan. 1990. Contributions of the Long-Term Ecological Research Program. **Bioscience** 40: 509-523.

Kéry, M. and H. Schmid. 2004. Monitoring programs need to take into account imperfect species detectability. **Basic and Applied Ecology** 5: 65-73.

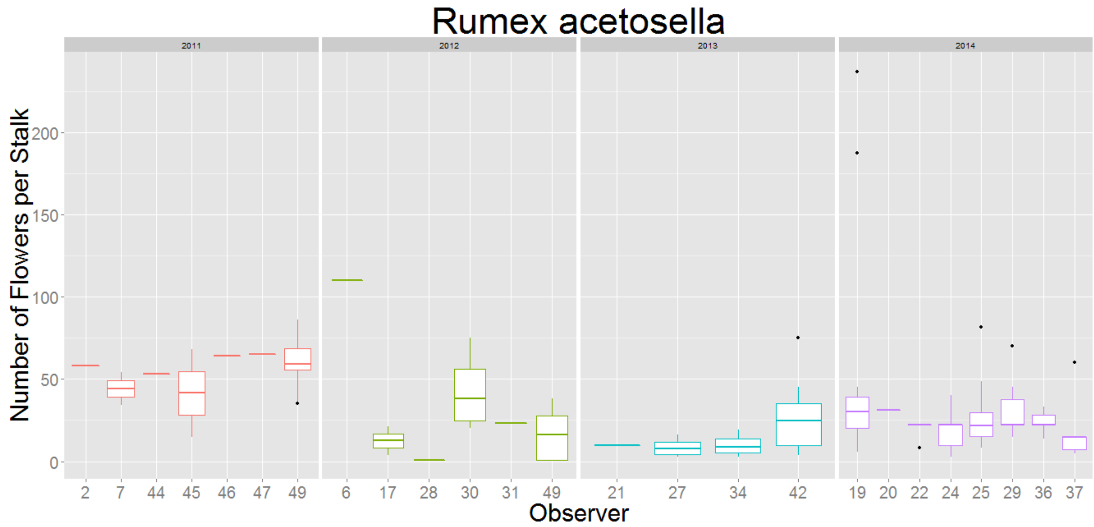
Magurran, A.E., S.R. Baillie, S.T. Buckland, J. McP. Dick, D.A. Elston, E.M. Scott, R.I. Smith, P.J. Somerfield, and A.D. Watt. 2010. Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. **Trends in Ecology and Evolution** 25: 574-582.

Vos, P., E. Meelis, and W.J. Ter Keurs. 2000. A framework for the design of ecological monitoring programs as a tool for environmental and nature management. **Environmental Monitoring** 61: 317–344.

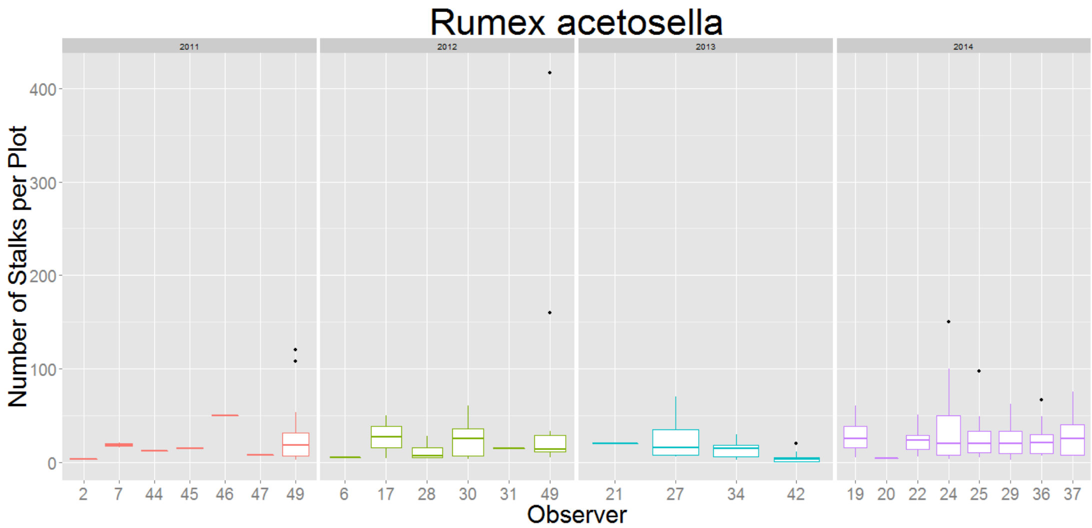
Images cited

Rumex acetosella: http://www.pnwflowers.com/flower/rumex-acetosella
Eriophyllum lanatum: http://www.botany.hawaii.edu/faculty/carr/ofp/eri_lan.htm
Ligusticum grayi: http://www.pnwflowers.com/flower/ligusticum-grayi
Achillea millefolium: http://www.botany.hawaii.edu/faculty/carr/ofp/images2/ach_mil_RLC2_6903b.jpg
Eriogonum umbellatum: http://www.pnwflowers.com/flower/eriogonum-umbellatum-var-polyanthum
Eriogonum compositum: http://www.pnwflowers.com/flower/eriogonum-compositum
Gilia capitata: http://www.pnwflowers.com/flower/gilia-capitata

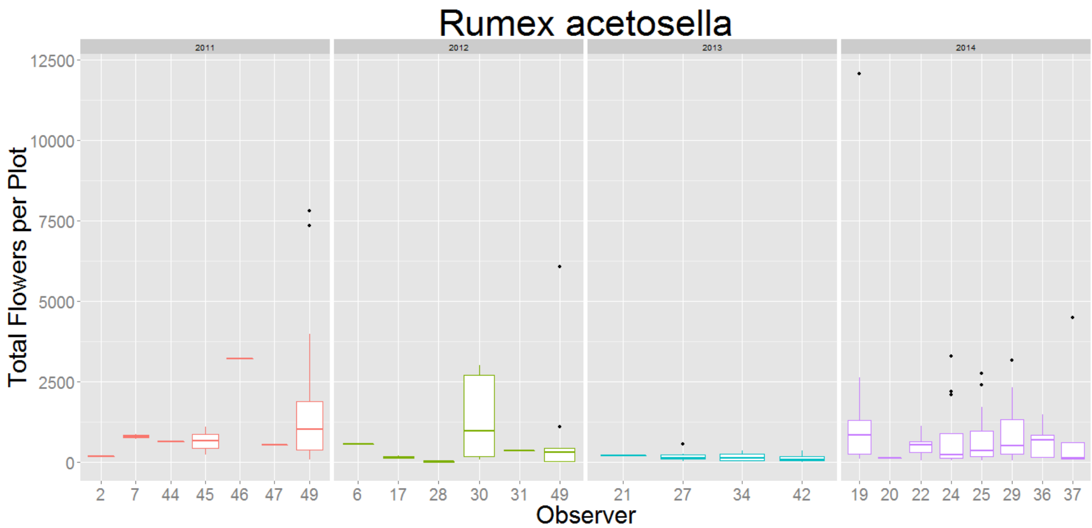
Supplementary materials



A.

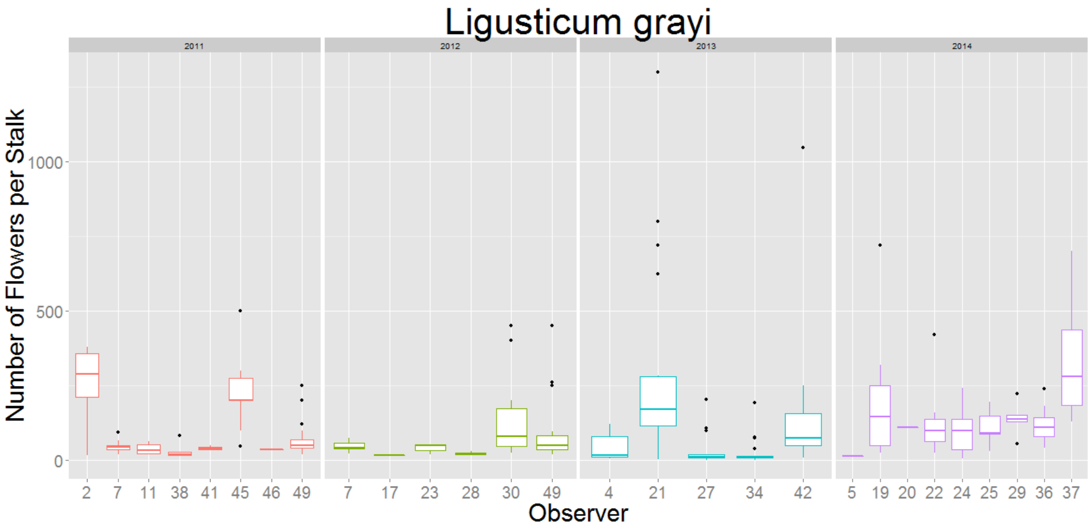


B.

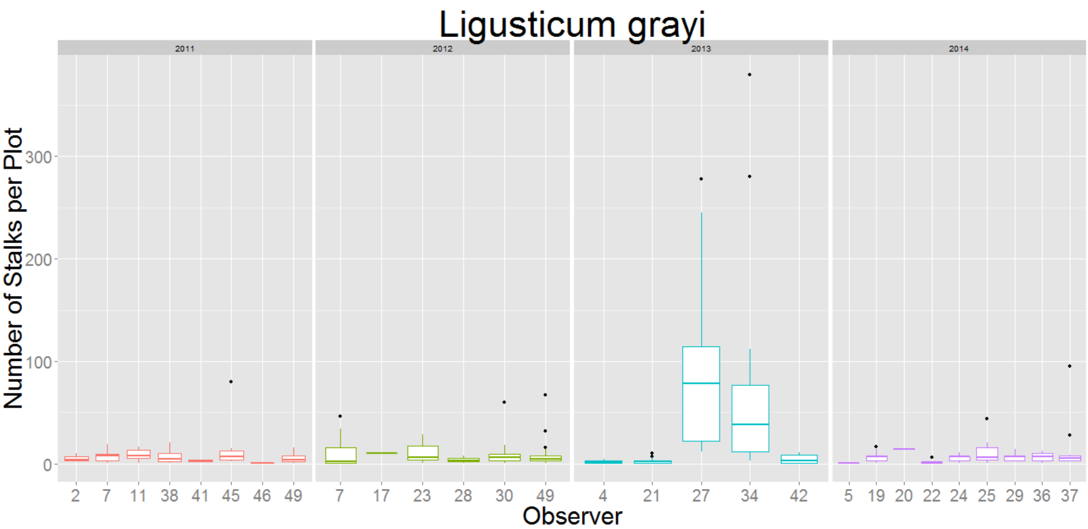


C.

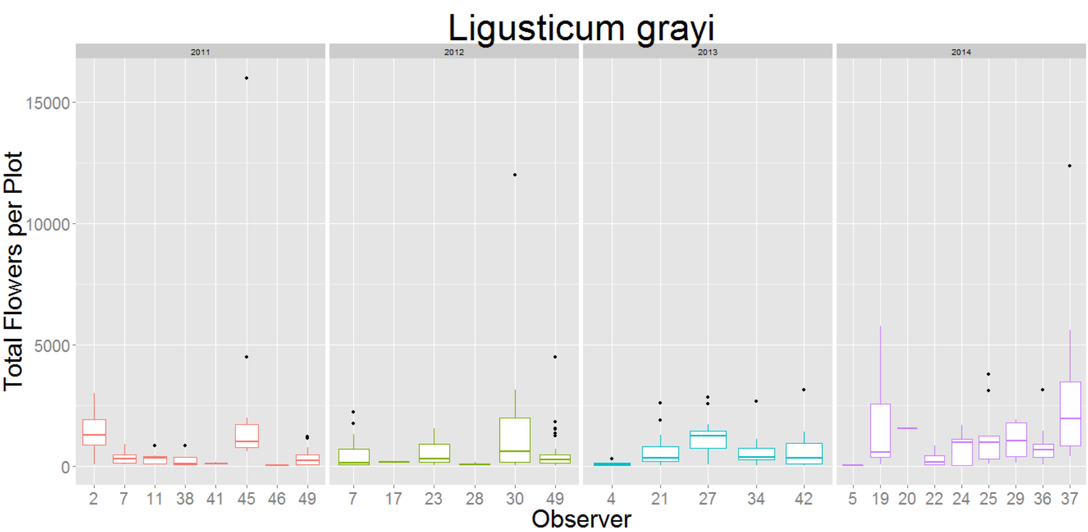
Figure 4. *Rumex acetosella* flower survey distributions across observer (1-56) and year (2011-2014) for number of flowers per stalk (A), number of stalks per plot (B), and total flowers per plot (C).



A.

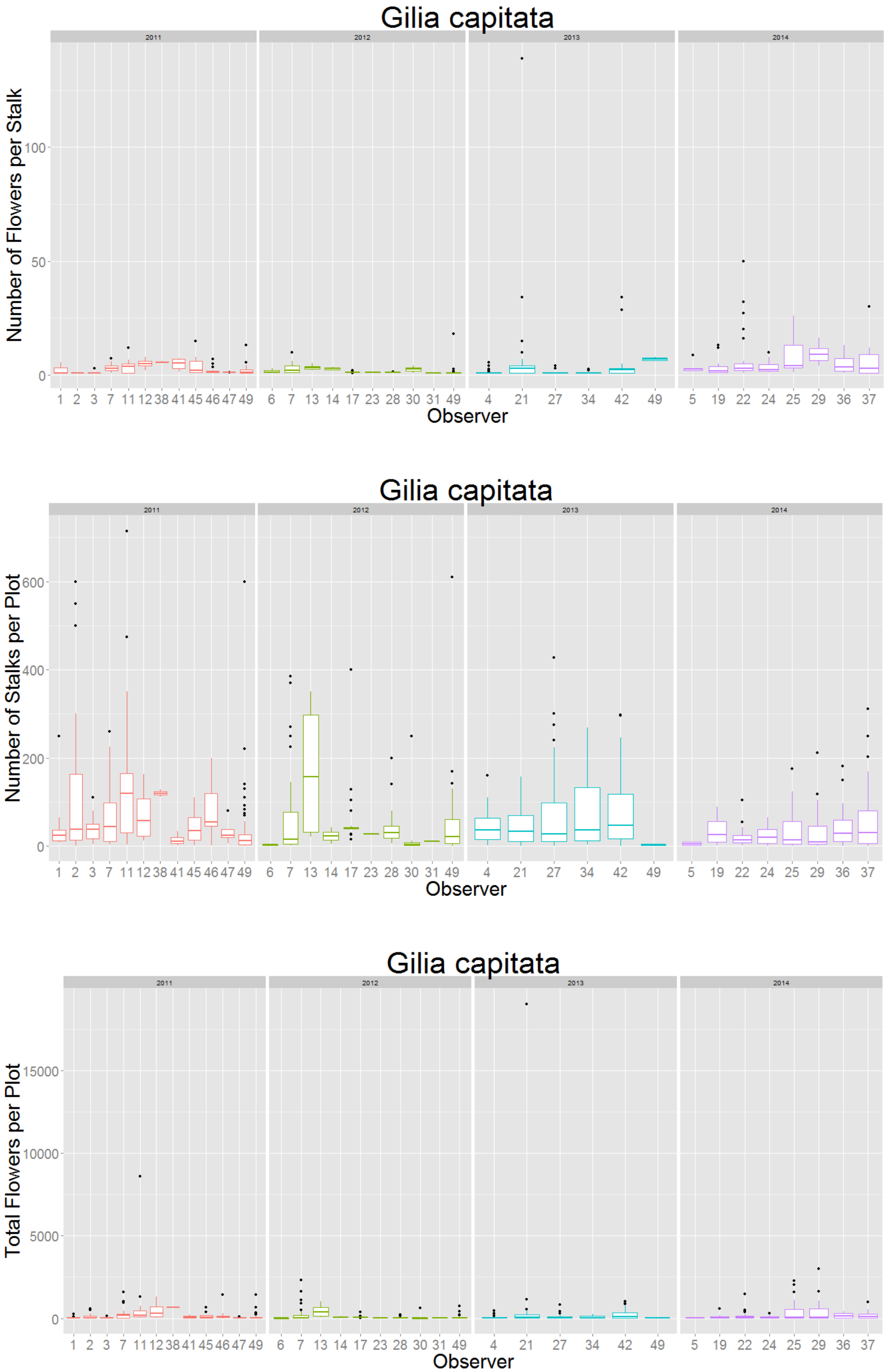


B.



C.

Figure 5. *Ligusticum grayi* flower survey distributions across observer (1-56) and year (2011-2014) for number of flowers per stalk (A), number of stalks per plot (B), and total flowers per plot (C).

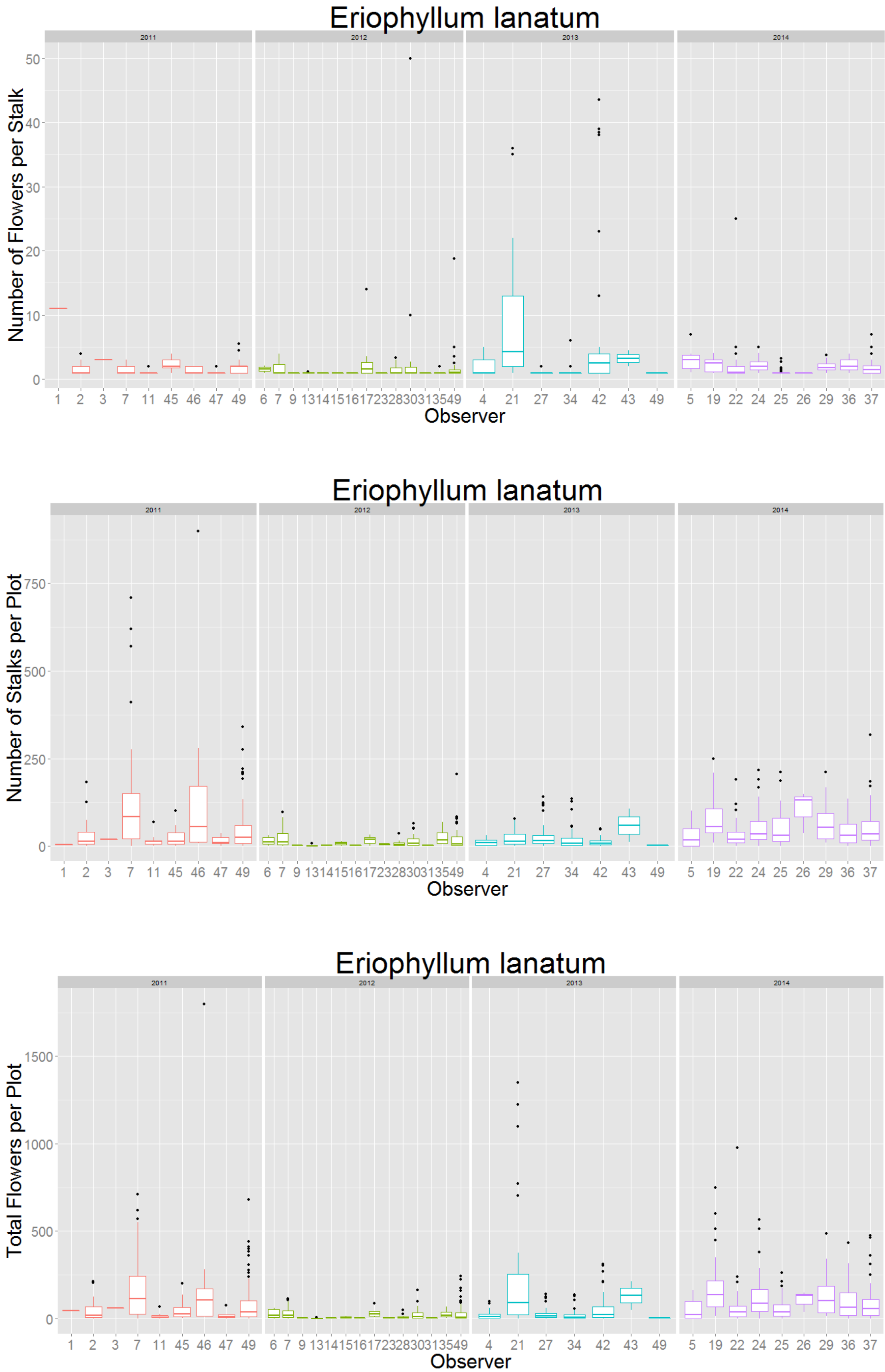


A.

B.

C.

Figure 6. *Gilia capitata* flower survey distributions across observer (1-56) and year (2011-2014) for number of flowers per stalk (A), number of stalks per plot (B), and total flowers per plot (C).

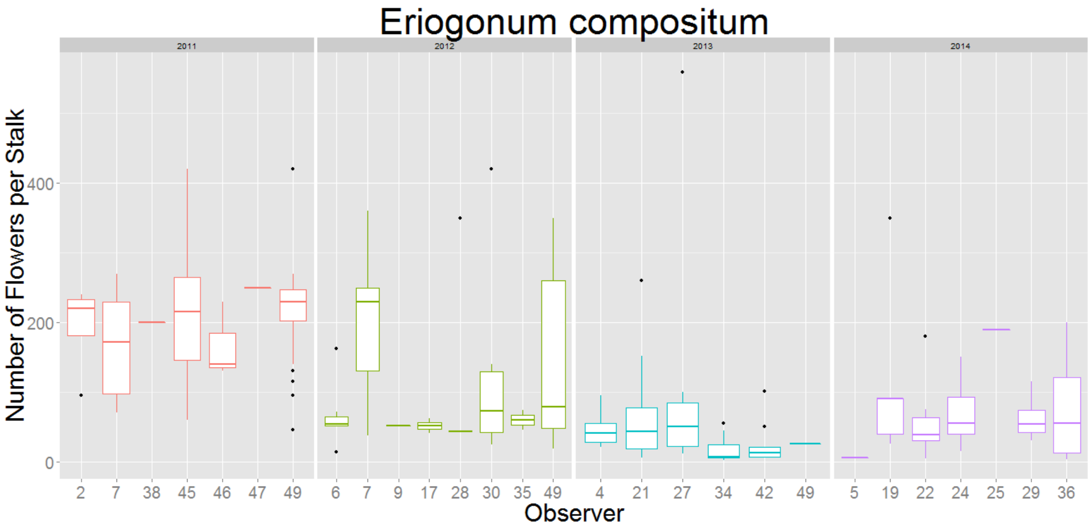


A.

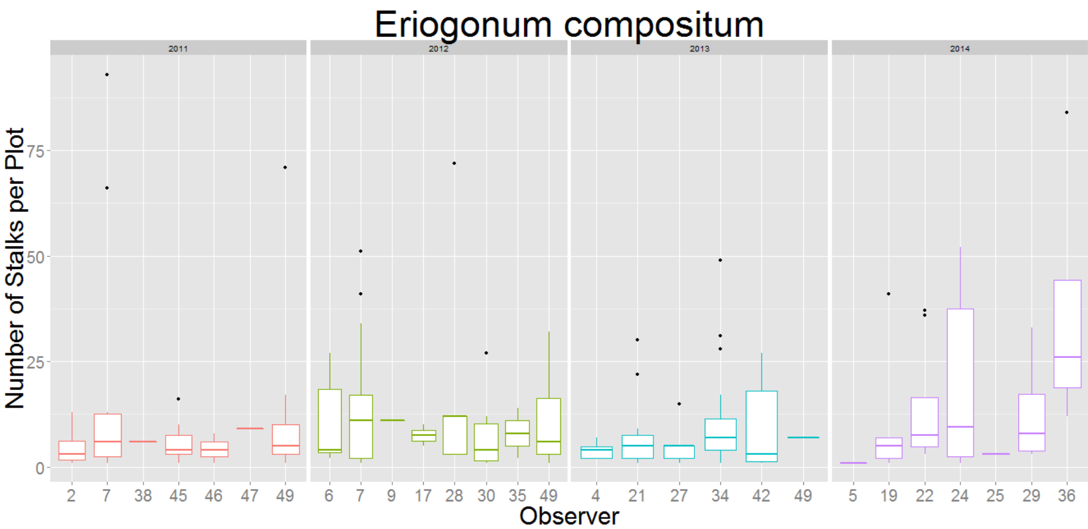
B.

C.

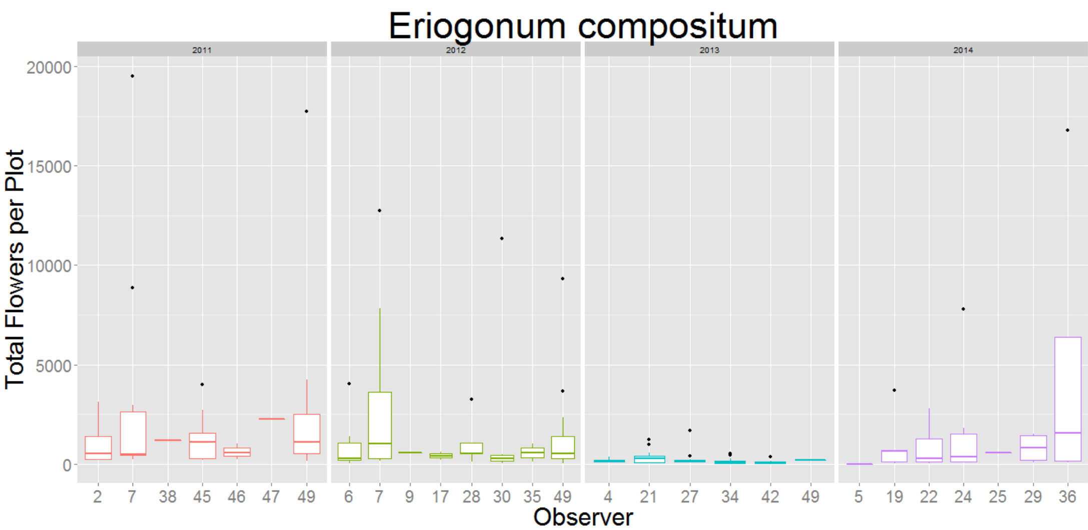
Figure 7. *Eriophyllum lanatum* flower survey distributions across observer (1-56) and year (2011-2014) for number of flowers per stalk (A), number of stalks per plot (B), and total flowers per plot (C).



A.

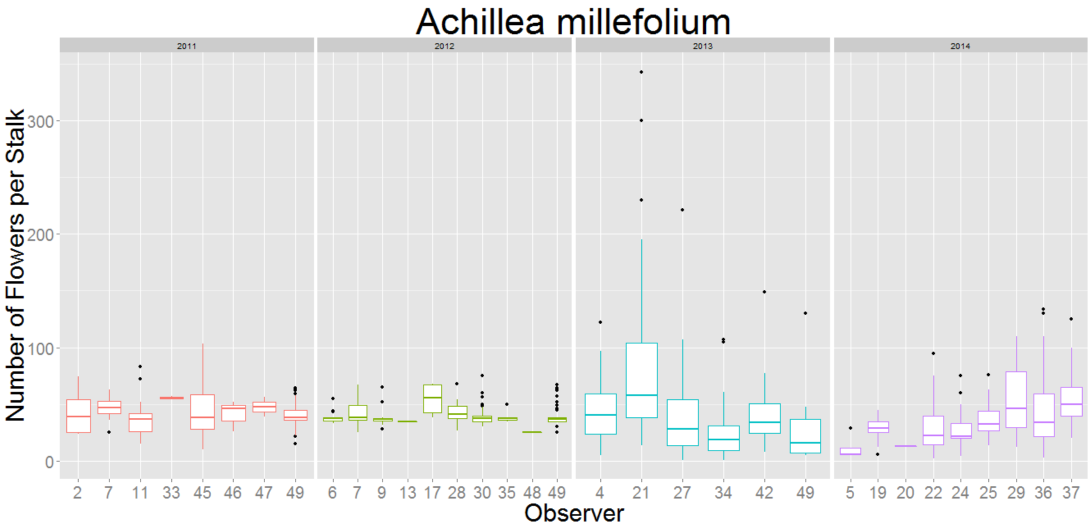


B.

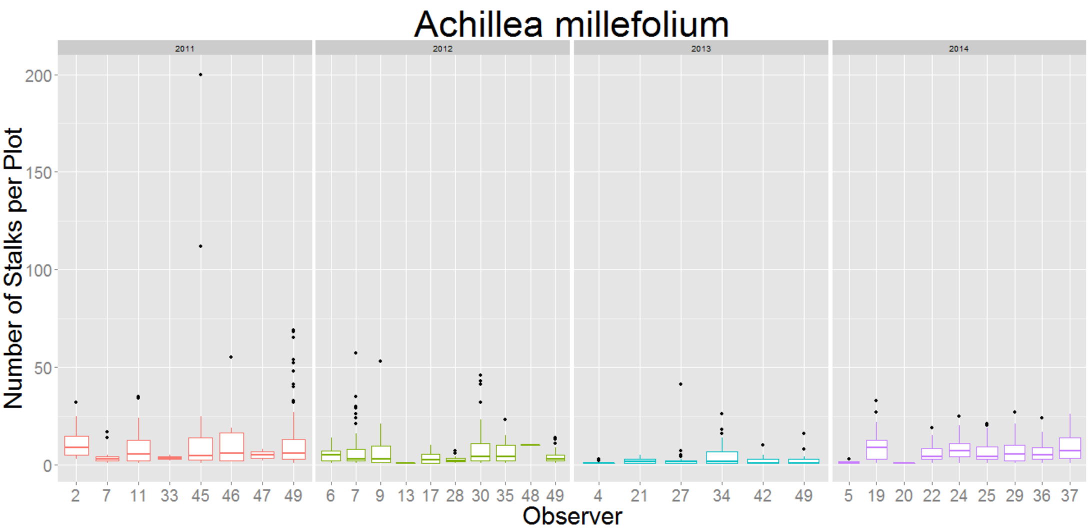


C.

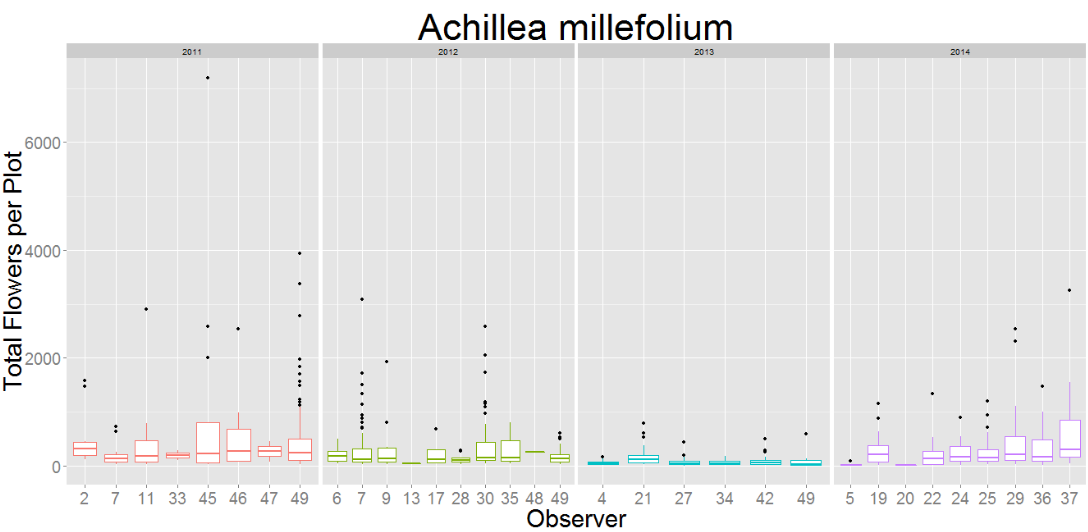
Figure 8. *Eriogonum compositum* flower survey distributions across observer (1-56) and year (2011-2014) for number of flowers per stalk (A), number of stalks per plot (B), and total flowers per plot (C).



A.

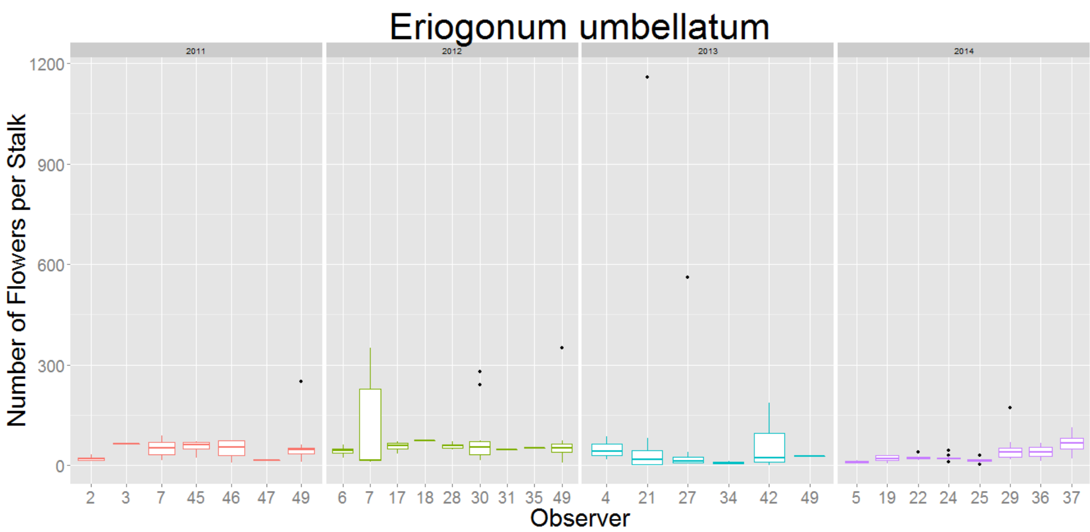


B.

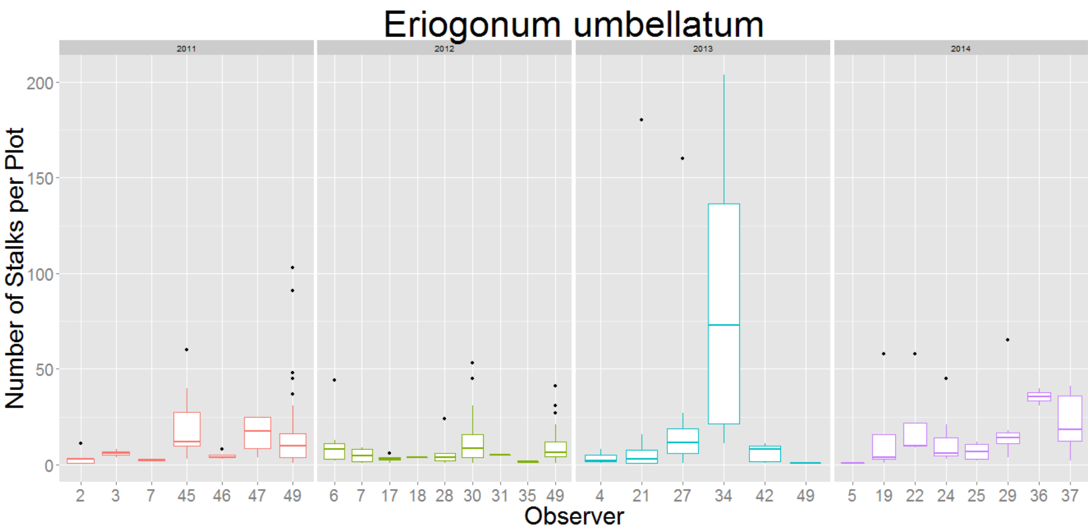


C.

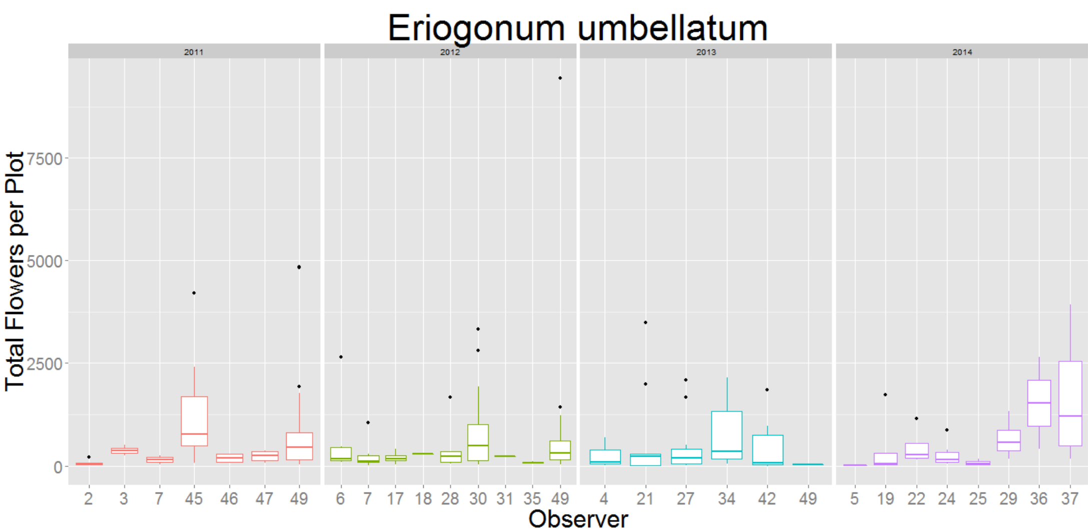
Figure 9. *Achillea millefolium* flower survey distributions across observer (1-56) and year (2011-2014) for number of flowers per stalk (A), number of stalks per plot (B), and total flowers per plot (C).



A.



B.



C.

Figure 10. *Eriogonum umbellatum* flower survey distributions across observer (1-56) and year (2011-2014) for number of flowers per stalk (A), number of stalks per plot (B), and total flowers per plot (C).