A systematic review of the use of statistics in studies of restoration ecology of arid areas.

Taylor J. Noble
Department of Biology
York University
Toronto, Ontario, Canada

Abstract

Restoration ecology is the study of restoration or restoration practices in degraded areas. It is of particular importance in arid environments due to the heavy impact humans have had in these areas. Some studies of restoration may require different statistics due to the unique challenges faced when examining degraded areas. A systematic review was conducted to assess the use of statistics in the field. It was determined that the field and influence of restoration ecology had increased dramatically since its development. Statistics are widely used in the study of restoration of arid areas. Major tests are similar to those found in other ecological studies such as ANOVAs and linear regressions. There were a few less common tests used in some of the studies. These include tests such as the Mantel test which may be useful to restoration ecology and should be explored further. Finally it was determined that the description of how statistics were used in the study was particularly important. The description should be detailed to help other researchers understand the findings of the paper. This will help to advance the field and the restoration of arid environments.
Introduction

Research involving degraded ecosystems have been around for some time, however restoration ecology, research involving the restoration of degraded areas, only emerged as a separate field in the late 1980s (Cairns et al. 1996, Dobson et al. 1997). Since then it has expanded as the need for restoring degraded ecosystems as well as conserving non-disturbed environments has been recognized. Restoration ecology is not meant as a replacement for preserving natural areas, but as a counter-point. The field has come to encompass everything from restoring heavily disturbed landscapes to helping develop management technique for areas with minor disturbance (Hobbs et al. 1996, Dobson et al. 1997). Since them it has expanded as the need for restoring degraded ecosystems as well as conserving non-disturbed environments has been recognized.

The field has come to encompass everything from restoring heavily disturbed landscapes to helping develop management technique for areas with minor disturbance (Hobbs et al. 1996). Projects often focus on restoring and area to a previously documented state whether recently or further in the past (Swetnam et al. 1999). Often restoration is encourage or promoted by government agencies or non-profit organizations in order to regain the benefits of intact ecosystems. This has led to projects involving restoration becoming among the most widespread and well-funded conservation projects worldwide (Holl et al. 2003). This is especially true of arid ecosystems where a combination of factors, including having a low value associated with them historically, have led to these environments being frequently disturbed and degraded (Belnap et al. 2002). Restoration issues commonly examined include invasive species and revegetation of disturbed areas (Fleishman et al. 2003, Zhao et al. 2007).

However despite this widespread use, many restoration studies are limited in focus to one site or restricted area. There has been a push to explore the implications of these studies on a larger scale and developing comprehensive tools for restoration ((Hobbs et al. 1996, Suding et al. 2004, Piekarska-Stachowiak et al. 2014). It has also been recognized that restoration ecology studies differ than more traditional ecology studies. For examine there may be non-uniform treatment across the site or sites. Replication may also be difficult. These difference mean that the experiment design may also differ from non-restoration studies (Michener et al. 1997).

The experimental design of a study and its use of statistical analysis influence each other greatly, guided by the questions asked by the study (Michener et al. 1997, Lortie 2014). Since restoration studies can differ in design from more traditional study ecologies, the statistical test chosen to examine the data may differ as well (Michener et al. 1997). This can lead to less common statistics being used in these types of studies or to the use of more common statistical tests in different ways (Michener et al. 1997). Restoration studies may also occasionally be purely observational. These studies may not use statistics at all (Dudley et al. 2004). With the increasing number of restoration studies there is a need to assess the use of statistics in the field to determine where the field currently stands in regards to statistical analysis and to see what could be improved or changed in the future. This will help determine what relationship statistics have to other factors of the paper, for example impact factor of journal, sample size, and restoration factor examined. Not only will this help us to understand how different restoration issues are often studied, but it will help ecologists chose the correct statistic for their particular study.

In this study, a systematic review of the use of statistics in restoration ecology was conducted. Restoration studies in arid areas were focused on due to the extensive impact humans have had on arid areas. This heavy impact is due to a number of factors, including the low value historically placed on these ecosystems and the effects that current climate change is having on these ecosystems (Belnap et al. 2002). The idea was to review papers from one biome in order to eliminate variability of
experimental design across different ecosystems. The primary purpose of this paper is to assess how statistics are used in restorations studies. This was determined by examining the following objectives:

1. What statistical tests are used in restoration studies of arid areas? Do these tests differ from the statistical test that are expected in ecology studies?
2. Are there any trends or patterns of statistics use in this field?
3. Are there gaps in statistical analysis of the field? Are statistics being used correctly or are there areas where improvements could be made? Do the statistics used make sense and are clearly explained?

It was predicted that most studies would use statistics that are most familiar to ecologists (such as ANOVAs or linear regressions) although a wider range of test was expected due to the different challenges that restoration studies present (Michener et al. 1997). It was also predicted that there would be a correlation between statistical test used and the citation rate of the paper. Statistical tests that are best for the experimental design and the question examined will likely result in a more informative paper that is more useful to other scientists and therefore cited more. A better use and understanding of statistics will improve the field of restoration ecology and lead to more meaningful and useful papers. This will in turn improve our current restoration techniques.

Methods

Database

To get an overview of the use of statistics in restoration ecology, the literature of the field was systematically reviewed. Thomson Reuters Web of Science was used to search for relevant peer-reviewed papers. The search terms restor* and desert were used. Since the focus of the review was statistics used in the restoration of deserts and arid ecosystems the Boolean search operator AND was used to eliminate papers that dealt with restoration or desert ecosystems only. The wildcard symbol was used to ensure that papers using all forms of the word restoration (restore, restored, restoration, etc.) were included. This search resulted in 899 peer-reviewed papers. To further refine the number of papers, results were then restricted to the Web of Science category of ecology. This reduced the number of papers to 369. Though year of publication was not used as part of the search criteria, the resulting list of papers were published between 1989 and 2014. These papers were then reviewed for their suitability for use in the systematic review. This lead to a rejection of 255 papers. These papers were rejected because they did not address some restoration issue, did not take place in a desert or arid ecosystem, or could not be applied to restoration in some way. The remaining 114 papers were examined and included in the systematic review.

Review and Analysis

The papers included in the review were examined and the following data was extracted: the year of publication, the number of citations, whether or not statistical analysis was used in the paper, the ecological or restorative factors examined by the study, the number of statistical tests used, the name and type of test used. The sample size, country of lead author, and location of study were also extracted. An assessment of the statistical description was also made. This was done by examining how much detail the authors put into their description (none = no description, brief = very little detail, but statistics explained, detailed = a detailed description of the statistics). Journal tier was determined using
the SCImago Journal Rank Indicator (SJR) (Scimago lab, http://www.scimagojr.com). These factors were examined to look for trends in the use of statistics in restoration ecology, as well as any gaps in how this topic is analyzed using statistics.

**Results**

The 114 examined studies of arid area used 31 different statistical tests in a total of 179 analyses. 21 different restoration factors were examined, with revegetation, invasive species, cryptobiotic soil, fire and herbivory being studied the most (Noble 2014). Statistics were used in 93% of the studies (fig. 6). The remaining 7% did not use statistics or were observational studies. Of the studies using statistics, 75% explained the tests used in detail (fig. 7). 18% had a brief description, while 7% did not explain their statistics (fig. 7).

The most common tests used were ANOVAs, t-tests, linear regression analyses, Kruskal-Wallis test, Pearson’s correlation analyses, and Mantel tests. Over time the diversity of tests have increased from just ANOVAs and t-tests in the earlier tests (fig. 1). A Chi-squared test was performed and found that the distribution of tests was significantly different than expected (p<0.0005). ANOVA tests were used much more often than any other test (n=71). Citation rate also varied with test. The 6 most common tests were cited the most. However studies that used the Mantel test were cited an average of 43.5 times, followed by studies using ANOVAs (average of 23.12 citations) and linear regressions (average of 24.23 citations) (see fig. 2). Impact factor was compared by year. While the average impact factor has remained the same over the years, the range of impact factors has become wider more recently. Some papers examining restoration ecology are being published in journals with a higher impact factor than in the past (fig. 8). Post hoc tests were given for 75 of the studies. Least Significant Difference tests were used more than half the time (52%, fig. 9). Tukey’s test was another common post hoc test used, with Duncan’s multiple range test and least squared means correlation also being used (fig. 9).
Discussion

Restoration ecology is widely studied in arid areas due to the heavy impact humans have had on them ecosystems (Belnap et al. 2002). Recent concern for these ecosystems has increased the number of studies on how to reverse these impacts. Many studies dealt with immediate concerns for humans such as the prevention of erosion (Li et al. 2009). The loss of soil can be a large concern in agricultural areas. Movement of sand in deserts may also be a concern to human industry such as railroads in the Horqin Sandy Land, China (Zhang et al. 2006, Zhao et al. 2007). Many others focused on restoring biodiversity or ecosystem function to areas, often through the restoration of soil (Zou et al. 2008, Doblas-Miranda et al. 2009) or plants (El-Bana et al. 2003, Gasque et al. 2004). Removal of invasives and
the use of fire are often also helpful in promoting biodiversity, which is likely why these were also among the most commonly studied topics (Schwinning et al. 2003, Parisien et al. 2009).

Statistical test of some kind or another were used in a majority of the studies (fig. 6). This is despite the fairly frequent use of observational studies in restoration (Zhang et al. 2012). This is likely due to the usefulness of statistics when assessing change. Without some kind of comparison it is harder to quantify change and determine how effective the restoration technique at hand is. Statistical analysis also allows for more comparisons to be made over more sites or repeats such as Peters et al. with 4800 repeats (2008). Studies with statistics tended to have larger sample sizes than those without (mean sample size 4.4 repeats for sites without statistics versus 489.3 repeats for studies with) (Noble 2014). It is not feasible to compare more than a few sites by observation alone. Statistics allow a wider area to be surveyed and increase the scope of the study.

Most studies described their statistical test in good detail. This is important as an understanding of the statistics is vital to researchers who wish to replicate the study. However a few studies did not provide much detail about their statistics. This is a poor practice as it does not allow for a full understanding of the results. Some studies that take this approach do not have many stats to describe or use common stats such as ANOVAs (Rafferty et al. 2002, Li et al. 2004). In these cases it is assumed that a researcher would have a basic understanding of familiar statistical test. In other cases with more complicated statistics or multiple levels of tests, an explanation would be beneficial.

Since the emergence of the field the number of papers published per year has increased. Just looking at the articles surveyed for this review, the number of papers written in the last ten years was more than twice the number written in the previous ten years (fig. 1). This demonstrates how the field has grown, which is likely due to the importance restoration has gained among ecologists and land managers (Cairns et al. 1996). The impact factor of journals published in compared over time demonstrates the growth as well. The earliest studies in this review were published in journals with a SJR (see methods) of between 1.00 and 3.00. Over time there were still many papers published in journals of this impact level however beginning in the early 2000s some papers began being published in journals with higher impact factors (fig. 8). The baseline remained but over time there was a slight upward trend with an increase of the upper end of the range. This shows that although some papers published about restoration ecology have the same impact as papers published earlier in the field’s development, some papers today are more influential. The field itself has become more influential since the late 1990s (fig. 8). The stats used also demonstrate the growth of the field. In earlier years ANOVAs, t-tests, and linear regressions make up the diversity of tests used. More recently a wider range of tests have been used including Spearman correlations, nonmetric multidimensional scaling, and generalized linear mixed models (fig. 1).

As stated by Michener et al., restoration studies often have a less than typical experimental design, which can drive the need for different statistical tests (1997). Most studies can be analyzed by ANOVAs or linear regressions, the two major types of statistics found in ecology (Cairns et al. 1996, Michener et al. 1997). Arid area restoration studies seemed to be similar. The most common statistical tests of the studies support this, with ANOVA, t-tests, and linear regressions making up a large portion of the total number of tests (fig. 1). In addition to being appropriate for common study designs, many ecologists are also familiar with these tests, and may feel most comfortable using them. Researchers may never have studied nonmetric multidimensional scaling, but most science undergraduate students
get a basic training to these common tests in an introductory statistics class by (Michener et al. 1997). The use of common statistics can be beneficial, as a wide range of individuals will be able to interpret the results easily. This is provided of course that the statistics chosen are appropriate for the study design. There are example where more powerful or more appropriate tests should be used (Michener et al. 1997). The results of this study show that there are many other kinds of tests that can be used in restoration ecology. To bring ecology terms to the statistics there is good diversity to the tests but not much evenness. The more widespread use of some of these “uncommon” tests might benefit the field more than the dominance of a few tests if they are used correctly.

Citation rate varied for the examined papers. Most of this was due to year of publication. The more influential papers tended to be at least a few years old. Many of the top cited papers were from the early 2000s such as Schwinning et al. 2002, or Fleishman et al. 2003. Statistical test used also seemed to have an effect on citation rate. The most common tests all had citation rates of between 15-25 citations on average per paper. ANOVAs and linear regressions were on the upper end of this range with average citation rates of 24.23 (linear regression) and 23.12 (ANOVA) citations per paper (fig. 2). In addition the most highly cited paper (Schwinning et al. 2002) used an ANOVA. However one test had a much higher citation rate when compared with these tests. Studies that used Mantel tests were cited an average of 43.5 times per paper. Though only four studies used Mantel tests, they were all among the most cited papers. Mantel tests used a generalized regression to relate two matrices. Often one is a measure of spatial distance and the other a measure of temporal distance (Dutilleul et al. 2000). This comparison of space and time is utilized in many restoration studies, which may explain why the citation rate for papers that use the Mantel test are disproportionate when compared to their total number. Studies such as Fleishman et al. 2003 and Bowker et al. 2006 may be the next advancement in restoration ecology. With its comparison of temporal and spacial distance measures, the Mantel test has the potential to be very useful in looking at how ecosystems change or recover over time (Dutilleul et al. 2000). If appropriate for the individual study designs, the field would benefit from more widespread use of this test. From the citation data, it appears to present useful information to other ecologists.

For the majority of studies, the primary statistic used could have been followed by a Post Hoc test. Post Hoc tests can increase the statistical power of the analysis overall when used correctly (Conagin et al. 2008). The most used test among the review studies was some form of ANOVA, which can usually be followed by some Post Hoc test. Despite this a minority of studies used or reported using this kind of test. For those that did report Tukey’s test and Least Significant Difference test were the most common Post Hoc used. While not appropriate in all situations, Post Hocs would be beneficial in many of the studies reviewed. For papers where this is the case but one is not given, it is unclear whether the tests were not used or merely that the specific test was not reported. Either way this should be a concern for anyone publishing a study of restoration ecology. Post Hocs can increase the power of an analysis and often can test us more clearly what is going on in an experiment (Conagin et al. 2008). Furthermore a detailed description of how the statistical tests were carried out will make the results of the study more useful to other researchers.

While still relatively new in comparison with other subdivisions of ecology, it is clear that restoration ecology has matured much as a field since it developed in the 1980s. Results from these studies have helped to improve land management practices of both conserved and degraded areas. Arid areas are a common topic of restoration studies because of the heavy impacts humans have had on these areas. However improvements could still be made. A wider diversity of statistical tests, where
appropriate to study design might help to further develop the field by giving researchers more tools to look at restoration questions with. The popularity of ANOVAs will likely never be challenge, but other tests such as the Mantel test could prove useful to ecologists (Dutilleul et al. 2000). Post Hoc tests should also be considered where appropriate (Conagin et al. 2008). More detail in explaining the choice of test and how it was preformed would also benefit many studies. Further development of the statistics used in restoration ecology will help to improve research practices and ultimately help to protect and restore the arid ecosystems of concern.

Conclusions:

- The field and influence of restoration ecology had increased dramatically since its development.

- Statistics are widely used in the study of restoration of arid areas. Major tests are similar to those found in other ecological studies such as ANOVAs and linear regressions.

- Some studies may require different statistics due to the unique challenges faced when examining restoration (including non-uniform treatment across the sites and difficulty of replication). A more widespread use of different tests across the field might be beneficial because of this.

Figure 2: Citation rate averaged per paper by test. The citation rates of all the papers using a particular test were averaged and compared. The top six most cited tests are presented.
- Less common statistical tests such as the Mantel test may be useful to restoration ecology and should be explored further.

- The description of how statistics were used in the study should be detailed to help other researchers understand the findings of the paper. This includes the description of any Post Hoc test used.

References


Figure 3: PRISMA Report of the search and reviewing process for studies used in this review. 369 studies were found using the search terms restor* and desert. This was reduced to 114 papers including in the
Figure 4: Geographic location of the studies included in the review.

Figure 5: The top 12 restoration factors examined in the studies and the number of articles that look at them.
Figure 6: Graph presenting the percentage of studies that did and did not use statistical test.

Figure 7: A graph presenting the amount of detail studies including in their description of statistics (only studies that used statistics were included). A detailed description was any description that provide substantial detail on the statistics used, brief was when limited detail on the statistics was provided, and none was any study that did not have a description of the statistics.
Figure 8: SJR rating (ScImago Journal Rank Indicator (SJR), Scimago lab, http://www.scimagojr.com), comparable to impact factor, of journals where papers published by year published.

Figure 9: Graph showing Post Hoc test used, when given.