Trees, Forests and People 2 (2020) 100010

Contents lists available at ScienceDirect

Trees, Forests and People

journal homepage: www.elsevier.com/locate/tfp

Editorial: Four tips for communicating clearly with readers: Designs, interpretations, and statistics

Scientists use many tools in creating ideas, testing those ideas, and communicating the new insights to peers and the public. Communication tools are fundamental: without clear, understandable communication, brilliant insights remain weak and ineffective. We offer four tips here for improving the communication of your projects in *Trees, Forests and People*, based on experiences of authors, reviewers and editors.

1. Science grows from the development of creative ideas about how the world works, coupled with experiments that have a strong chance of finding flaws in the appealing ideas. When challenging measurements or experiments fail to find flaws, the ideas can be used to understand the world with some confidence. This general way-of-knowing applies across all scales of space and time in trees and forests, but sometimes confusion about these scales can lead to misinterpretations. For example, we may have an idea that the biomass of Scots pine (*Pinus sylvestris*) trees can be predicted with precision by knowing the height and diameter of the tree. This idea could be challenged by cutting down 30 trees in a stand and determining mass by weighing, and then using statistical methods to calculate an allometric curve. The statistics may tell us that mass can be predicted with very high confidence, based on height and diameter. But could this equation be used to estimate the mass of stems of birch (*Betula pendula*) in the same stand? Probably not. Would the equation be reliable for use in a stand that was 20 years old on a similar site, or the same age stand on a higher productivity site? Maybe, but because the other sites were not in the population of inference for the original experiment, this sort of extrapolation would not warrant high confidence without further testing. Table 1 presents this same idea for two ends of an experimental spectrum, from creative ideas that might be tested across a region, or within a single tree. Both of these examples could form the basis of a strong scientific paper in *Trees, Forests, and People*, but readers benefit when authors are very clear about the statistical population of inference, the sampling scheme, and what extrapolations might be reasonable.

2. The introduction of a science article is the place to hook the curiosity of readers so they will be motivated to continue reading. Many approaches can make the opening of an introduction intriguing, but there's one form that always seems to be a weak choice:

A. This broad, general topic is very important in forests everywhere.
B. Not much is known about this smaller detail, so we measured it.

Just because something is not well known doesn't mean a reader will want to invest time to learn about it. Indeed, there's an infinite amount of details that are not well known, so we hope there was a more exciting reason why a project was developed. Why might a reader want to invest the time in reading this paper? Line "A" above can be a fine first sentence, but different approaches to "B" could be more engaging for readers' curiosity:

1. The broad, general topic shows great variation across sites, and we tested whether our idea could account for when values are extremely high or extremely low…
2. We might expect the broad, general topic would be like this for the forests we're most interested in, but surprisingly our forest was very different. We hypothesized that the divergence could be explained by…
3. The trend for the broad, general topic is generally thought to result from this important driver, but we tested whether a different factor was actually responsible…
4. The broad, general topic leads to some expectations that specific sites might be especially vulnerable, and we examined the vulnerability of three such sites to see if that seemed to be true…

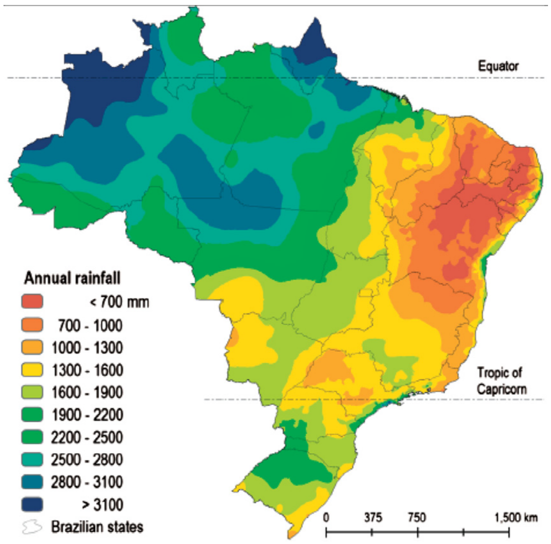
The key point is that the reason a reader should continue with a paper is that something useful or important will be learned, not just that one more bit of knowledge will be added. A good exercise would be to revisit papers that authors found to be intriguing and careful evaluate how the authors structured the first few sentences of the introduction. Which papers had the most effective "hook" that pulled readers into paragraphs that followed?

3. Interpretation is the task of distilling a story out of raw data. All projects and experiments produce large amounts of data, and scientists and authors have an obligation to evaluate the full range and power of the data. When it comes time for readers to understand insights that came from the project or experiment, authors need to provide the data summaries and insights that clearly relate to ideas, to the stories. Data cannot be relied upon to tell stories all by themselves. Fig. 1 illustrates a few points about how authors may consider distilling data to provide clear insights. These data come from forest inventories across a large area of *Eucalyptus* plantations in southeastern Brazil. The first graph (2A) shows the average rate of wood production for each year across several decades. The dots represent the averages across the area, and the jagged line simply connects the dots. Is this the best way to understand the "story" of how plantation growth changed over decades? A close look would show that the production rate fell from 37.2 in 1992 down to 33.3 in 1993, an 11% decline. However, production was back up to 37.2 in 1994, so the drop in a single year might be attributable to "noise" rather than to the long-term trend over the years. Production may rise above or fall below the long-term average as a result of factors such

https://doi.org/10.1016/j.tfp.2020.100010

Available online 28 June 2020

Table 1
The design for large experiments and for case studies need to clearly address the same core questions: what is the population of inference, how was that population sampled, and can the results be expected to apply beyond the statistical population of inference? A question tested across a large set of sites across a region has a large population of inference, whereas sampling nested within a single location has a small population of inference (even if replicate samples are taken within that site). Extrapolation of results beyond the statistical population of inference may be described based on a fundamental process or other line of reasoning, but extrapolation needs to be expressed as the authors' speculation or best guess, not as a confident outcome from the experiment's data (Brazil map provided by C.A. Alvares, see Binkley et al., 2017).



Issue	Broad experiment	Case study
Key question	How does the productivity of *Eucalyptus* plantations relate to rainfall?	How many insect species occur on this tree?
Population of inference	*Eucalyptus* plantations across Brazil	This *Eucalyptus* tree
Range for extrapolation	Perhaps *Eucalyptus* plantations in other regions	Perhaps nearby trees of the same species
Sampling design	Representative samples across a rainfall gradient	Branches and leaves from various heights in this tree
Replicated?	Yes, but only within Brazil	Yes, but only within this tree

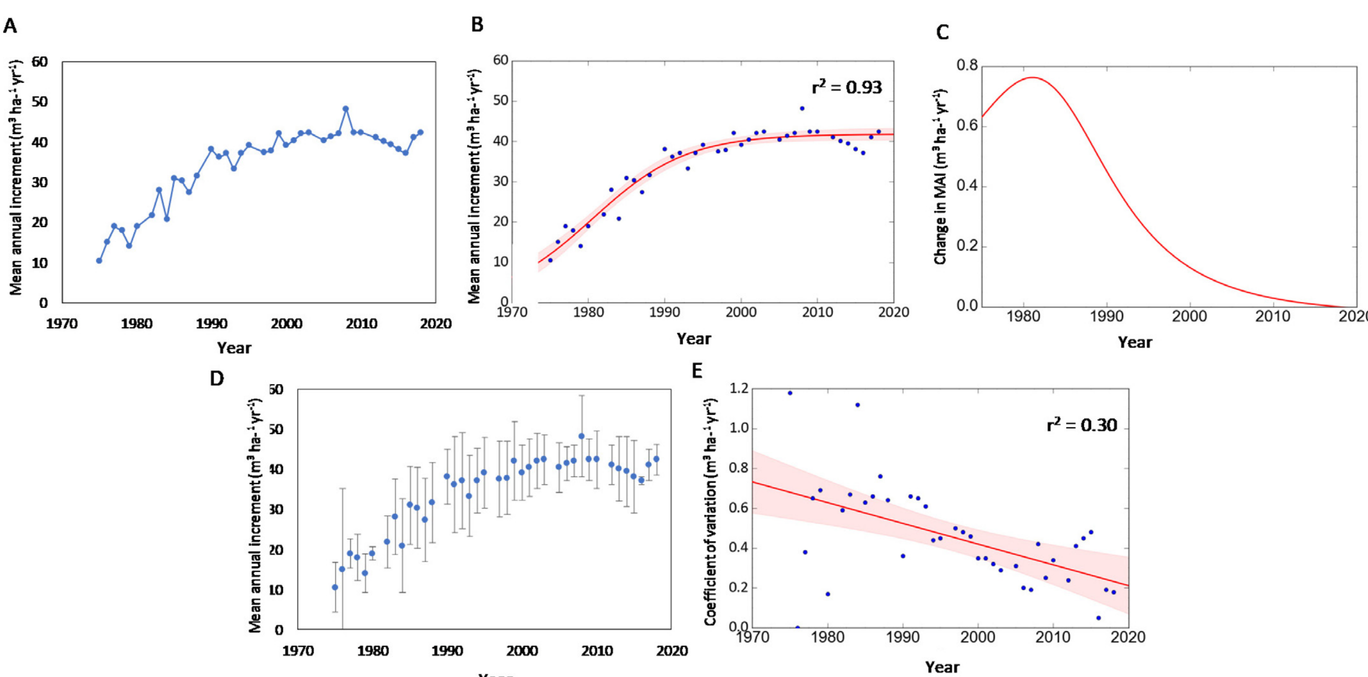


Fig. 1. The raw data from a project needs to be investigated and synthesized to inform readers of the key patterns, and the outcomes of testing hypotheses. In a large area managed for *Eucalyptus* plantations, wood production increased over a period of several decades. The choice of the which of these graphs to use in a paper depends on the insights the authors would like to convey (see text; data shared by R.E. Hakamada).

as weather fluctuations, insect or disease problems, or inconsistency in weed control. The jagged line in 2A essentially traces both the long-term "signal" in the data, and also all the "noise" from factors that may not relate to the main trend. If authors wanted to convey the core story of how growth changed in the long term, Fig. 1B aims to display the central tendency (and even has a band plotted along the line to show the domain of the 95% confidence interval around the line). The original data points may or may not be useful to keep in Fig. 1B; that decision would be based on the key points the author would like to help the readers understand. Both 2A and 2B make it clear that production increased, and that the gains in production were larger in early decades than in recent decades. Is this a key point of the story that the authors would like to help the readers understand? If so, then Fig. 1C might be worth adding to the paper to show explicitly how the rate of increased changed over the years.

The original inventory data had many plots behind the reported average for each year, and sometimes authors would like readers to understand how consistent or variable the replicate observations happened to be. Fig. 1D illustrates the coefficient of variation, or how large the standard deviations were among the replicate plots relative to the sizes of the means. Two points might be noticed about the coefficients of variation. Some years had high coefficients (such as 1.1 for 1984), and others had lower values (such as 0.6 for 1985). There may be an interesting story lurking in these year-to-year fluctuations in the coefficients of variation, but these variations might be largely "noise" if the focus is on the general trend across the years. Astute readers may notice that the coefficients of variation seem to be getting narrower over time, and indeed a direct plotting in Fig. 1E shows the coefficients declined by more than two-thirds over the decades. Clearly the management of the plantations improved the consistency of production across sites in later years. If the authors expect that readers would be interested in understanding this aspect of the long-term changes, then a figure such as 1E conveys the story much more clearly than Fig. 1D. Authors have great opportunities for understanding the key patterns that lie within their large data sets, and to develop the clearest graphics to convey those key insights to readers. Readers should not have to be data miners; authors are responsible for serving clear and accurate graphics to support the key points of the analyses.

4. Science was invaded by statistical viruses in the late 20th Century. Widespread access to computers in the 1960s and especially in the 1970s led to an explosion of statistical analyses in scientific papers. When calculations were done laboriously by hand, statistics focused on the main hypotheses of experiments. Now it's easy to find papers that are peppered with P values, and sometimes F values as well. Did the proliferation of statistics result from clear insights about what makes a good hypothesis, and how it might be challenged with data? Or did papers get covered with P because the idea (meme) caught on and was replicated, even if it didn't provide much value? Many interesting stories about the history of science explore these questions. A particularly good summary from quantitatively literate ecologists is Anderson et al. (2000; and more detail in Burnham and Anderson 2002, Hobbs and Ogle 2011). Two strong and desperate pleas from statisticians who hope to stem the tide of P are Wasserstein and Lazar (2016), and Wasserstein et al., 2019. If you would like to test whether you know what a P value is (and is not), you might enjoy: https://www.youtube.com/watch?v=OcJImS16jR4.

For papers in *Trees, Forests, and People*, perhaps a few simple guidelines would be useful. The first is that when the effects of a treatment are reported, the most interesting thing for readers to understand is whether the effect was impressively large, or too small to be interesting. The text of paragraphs (including abstracts) should emphasize whether the outcomes were large or small. Once readers have an image of how interesting the size of an effect was, they're ready to know how much confidence they might want to vest in the idea.

If authors state in a methods section that a *P* value of 0.05 was used in statistical tests, it's not useful to repeat lots of actual *P* values throughout the paper (and especially not inserted parenthetically into sentences, constipating the flow of the prose). Some situations might bring a *P* value into the forefront, such as an observed treatment effect that appeared very large, and was probably not random, but may have barely missed the binary cutoff level for the chose P value.

Many ecological ideas may be examined using information/theoretic approaches of testing how well models (simple equations or complex simulations) account for variation observed in real situations. A bottom line for scientists to consider is whether they are using the best sort of quantitative tools, rather than diving into minute details of a classical tool that might not be all that insightful. We should expect the quantitative tools that we use to challenge ideas, now and in the future, are not limited to the ones that characterized statistics courses in the 20th century.

Dan Binkley, Editor-in-Chief

E-mail addresses: dan.binkley@alumni.ubc.ca,
dan.binkley@colostate.edu

References

Anderson, D.R., Burnham, K.P., Thompson, W.L., 2000. Null hypothesis testing: problems, prevalence, and an alternative. J. Wildlife Manage. 64, 912–923.

Binkley, D., Campoe, O.C., Alvares, C., Carneiro, R.L., Cegatta, I., Stape, J.L., 2017. The interactions of climate, spacing and genetics on clonal Eucalyptus plantations across Brazil and Uruguay. For. Ecol. Manage. 405, 271–283.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference. Springer-Verlag, New York.

Hobbs, N.T., Ogle, K., 2011. Introducing data–model assimilation to students of ecology. Ecol. Appl. 21, 1537–1545.

Wasserstein, R.L., Lazar, N.A., 2016. The ASA's statement on *p*-Values: context, process, and purpose. Am Stat 70, 129–133.

Wasserstein, R.L., Schirm, A.L., Lazar, N.A., 2019. Moving to a world beyond "p < 0.05". Am Stat 73, 1–19.