Advice on writing modeling papers for mixed theoretical/empirical audiences

Marissa L. Baskett

This set of guidelines originated from a set of suggestions I found myself frequently giving graduate students and postdocs who were embarking on writing their first modeling manuscript. Much of the advice below particularly applies to writing papers for ecological and evolutionary journals that publish both empirical and modeling papers. Mathematical journals have their own conventions, for which some of the guidance below does not necessarily apply.

1. **Read "The science of scientific writing"** (Gopen and Swan, 1990): the lessons about clear scientific writing apply just as well to modeling papers as they do to empirical ones.

2. The Introduction: **motivate the study biologically** – i.e., what's the biological question/gap to fill. For example, saying two things have been modeled separately but never together isn't a biological motivation. Talking first about why, biologically, we might expect those two things to affect each other, perhaps drawing from what's known about modeling each separately, provides a better biological motivation. Even better is if you can set up contrasting logical expectations, e.g. with those two things together you might expect a variety of biological outcomes depending on which dominates, because then you've set up a logical conundrum for which a model provides the resolution, using mathematics as "a way of thinking clearly" (May, 2004). For example, compare:

   - Models of marine reserves have incorporated single-species population dynamics, and some fisheries models without marine reserves have included predator-prey dynamics. Here we include predator-prey dynamics into a marine reserve model.
   - Fisheries models without marine reserves indicate that predator-prey dynamics can affect management decisions. With marine reserves, predator-prey dynamics are likely to affect population responses to reserve establishment, where prey species might decrease after reserve establishment due to an increase in harvested predators. Here we quantify the potential for such a decrease by incorporating predator-prey dynamics into a marine reserve model.
   - Fisheries models without marine reserves indicate that predator-prey dynamics can affect management decisions. With marine reserves, prey species might decrease after reserve establishment due to an increase in harvested predators. However, prey species might increase after reserve establishment if they are harvested as well. Here we determine whether, and under what conditions, prey populations will decline after reserve establishment given dynamical predator-prey interactions.

   Each successive example better provides a biological hook to intrigue the reader and sets up the need for a model to explore the array of possible outcomes.

3. The Methods:

   (a) **Build up to the model**: define all variables and parameters before they appear in an equation, and do so in the context of describing the dynamics intuitively. The goal of this approach is for a math-savvy reader to be able to picture the model structure before s/he sees it and for a non-math-savvy reader to understand where the expression comes from when s/he sees it. For example, compare:

- The population dynamics are

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right), \tag{1}$$

where r is the population growth rate, K is the carrying capacity, and N is the population size at time t.

- Given growth rate r and carrying capacity K, the change in population size N over time t is

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right). \tag{2}$$

- The population size N grows exponentially at rate r at low densities, but that growth rate slows as it approaches carrying capacity K and becomes negative if the population size exceeds K. The population change over time t is then

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right). \tag{3}$$

Each successive example gives the reader a better intuition for the model before seeing the math (note that here I use the simple, commonly-used logistic model for illustration; if I were building up to a more complicated model that included these dynamics, I might just say, "The population N exhibits logistic growth with population growth rate r and carrying capacity K...").

After having defined something, if it's been a while (a couple of paragraphs) since readers have seen a given parameter or variable symbol, re-state its definition when bringing it up again.

(b) **Include a cartoon figure of the model** that provides an intuitive illustration of the dynamics. A table of parameters and their values is also very helpful when the analysis is primarily numerical.

(c) **A brief model overview** at the outset of the methods section gives the empirical reader a math-free place to build their intuition and the modeling reader a sense of where the paper is going. Not only should this describe the model in intuitive, biological terms, it also can strengthen the paper if it indicates how the model analysis will answer the central question(s) set up in the Introduction (or this can go in a separate model analysis subsection of the methods). Similarly, in describing the model, it adds to the intuition to say why you chose the structure you did in terms of how these choices capture the dynamics that are essential to answering the question(s) posed in the Introduction (or, if you don't have a model overview section, then look to do this as you build your model).

(d) If your paper is focused on numerical (computer) simulations, still write down the steps of those simulations mathematically, as math provides the universal language for repeatability of your study.

(e) Punctuate equations, even offset/display/numbered equations, as if they are part of the sentence (e.g., having a period at the end if they're the last thing in a sentence, like in the examples of logistic growth above).

4. The Results:

(a) If you are constructing a "strategic" model where the goal is qualitative understanding (as opposed to a "tactical" model where the goal is prediction; May, 2001), then do not report numbers in the Results section, even if all of your analysis was numerical. Instead, focus on the comparative outcomes and qualitative dynamics that address the original question(s) posed.

(b) As with any paper, be careful to construct and label your figures such that they can be easily understood without having to read extra text, and make sure your figures clearly convey your take-home messages that directly address the question(s) posed in the introduction. For example, if your main take-home is that the key driver of the qualitative outcome is some factor represented in the model by a continuous parameter, then that factor should be the x-axis in your first results figure. As another example, if your key take-home is that some functional form changes the qualitative dynamics, that comparison should be overlaid or side-by-side in your first results figure. To illustrate building a figure around answering the central question:

- For the marine reserves/predator-prey question posed in the Introduction example above, a first results figure might be to have a y-axis of change in prey after reserve establishment (the response/outcome that is in question), an x-axis of the harvest rate on prey (the factor hypothesized to affect this response), and different lines for including versus excluding a harvested predator (the functional difference that represents the new dynamic proposed to change the outcome in comparison to the baseline without that dynamic).

5. The Discussion:

(a) As with any Discussion, a modeling paper Discussion typically begins with the key take-home messages, where saying why those outcomes occurred helps to distinguish the Discussion material from the Results material. Also common to any Discussion, including sub-headers throughout helps to better guide the reader, and referring to figures where appropriate helps to better clarify which points come of the results of this study.

(b) **Be sure to include both the theoretical and empirical context** (i.e., place your findings in the context of relevant papers from both the modeling and empirical literatures). How these come in will vary greatly from manuscript to manuscript, and there is no one "right" way. That said, one common way that the theoretical context comes in is by providing intuition for what might happen if a given model assumption is relaxed, and one common way that the empirical context comes in is by indicating what assumptions might be reasonable vs. unrealistic. Note the implication that discussing assumptions is integral to any Discussion section of a modeling paper.

(c) For a modeling study, the answer to the question posed in the Introduction will most likely be "it depends", but the interesting thing is *what* it depends on. For both the Discussion and Abstract, be as precise and specific as possible in describing the dependencies. For example, compare:

- Whether we expect prey species to increase or decrease after reserve establishment depends on both predator and prey harvest rates.
- A prey species is more likely to decrease after reserve establishment with lower predator harvest, and therefore lower increase in predation after establishment, as well as greater prey harvest, and therefore a greater reduction in direct mortality after establishment.
- Prey species decline after reserve establishment requires that predator harvest, and therefore the increase in predation after establishment, exceed [some relationship involving prey harvest].

Each successive example is increasingly specific about the insight provided by the model.

(d) Remember, models say what *can* happen, data say what *does* happen.

6. Apply all of the above lessons to the Abstract: motivate the study biologically, use your model overview to think about how to describe the methods, focus on qualitative results (unless your goal is predictive forecasting), and be precise and specific about dependencies, giving an intuition for why your outcomes occurred where feasible.

7. Consider **learning LaTeX:** writing and referencing equations is smooth and straightforward after the initial time investment to learn the basics (note that you can now use LaTeX commands in Word Equation Editor as well). If you're collaborating with non-LaTeX-using co-authors, you can use latex2rtf or another converter to get it in a Word-friendly format for tracked changes (Adobe Acrobat also does tracked editing for LaTeX-produced pdfs, and someone with Acrobat can enable tracked editing on a given document for those with Adobe Reader).

References

G. Gopen and J. Swan. The science of scientific writing. *American Scientist*, 78(6):550–558, 1990.

R. M. May. Uses and abuses of mathematics in biology. *Science*, 303(5659):790–793, 2004.

R.M. May. *Stability and complexity in model ecosystems.* Princeton University Press, 2001.