Received: 20 April 2018 | Accepted: 4 October 2018

DOI: 10.1111/2041-210X.13105

RESEARCH ARTICLE

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

Some guidance on using mathematical notation in ecology

Andrew M. Edwards[1,2] | **Marie Auger-Méthé**[3,4]

[1]Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, British Columbia, Canada

[2]Department of Biology, University of Victoria, Victoria, British Columbia, Canada

[3]Department of Statistics, University of British Columbia, Vancouver, British Columbia, Canada

[4]Institute for the Oceans & Fisheries, University of British Columbia, Vancouver, British Columbia, Canada

Correspondence
Andrew M. Edwards
Email: andrew.edwards@dfo-mpo.gc.ca

Handling Editor: Sean McMahon

Abstract

1. Mathematical modelling is playing an increasing role in studies of ecological systems. This requires the communication of the details of a mathematical model, including the use of mathematical notation to represent ecological variables, parameters and processes.
2. In our experience, the clarity of mathematical notation varies between papers and can often be inconsistent with general conventions. Poor notation can impede communication and understanding of ideas, and make models appear more complicated than necessary.
3. Here, we present some guidelines, including: (a) define every term in an equation, (b) use fonts appropriately (italicise mathematical symbols, use bold lower case for vectors and bold upper case for matrices), (c) use subscripts appropriately (to index quantities, for example, by year), (d) use superscripts appropriately (to indicate a power, the transpose of a matrix or the steady-state value of a quantity), (e) avoid multiletter variable names, and (f) revisit notation early on in a project to see if it should be refined.
4. Although we focus mainly on ecology, our guidelines should be of interest to researchers applying models in evolutionary biology and broader areas of biology.

KEYWORDS
ecological models, equations, mathematics, parameters, reproducibility, variables

1 | INTRODUCTION

Six decades ago, Lowry (1959) pleaded for the "increased use of mathematical notation where appropriate in ecological literature." This has certainly been realised. For example, equations appeared in only 14% of the 37 contributions in the issue of *Ecology* containing the Lowry (1959) paper, but in 38% of the 26 contributions comprising the March 2018 issue. Not surprisingly, in *Methods in Ecology and Evolution*, a relatively new journal focussed on methodological developments, there is an even higher percentage (56% of the 34 contributions in the March 2018 issue). Advantages given by Lowry (1959) include precise communication of logical thought (beyond that afforded by the written word), and the reproducibility of methods and results. He noted that advantages "accrue in proportion to the care used by the author in the use of notation."

There are many books concerning the various skills needed to produce a scientific paper that uses mathematical modelling in ecology. For example, there are books covering introductory ecology (Begon, Harper, & Townsend, 1986), general mathematics (Kreyszig, 1993), ecological modelling (Hilborn & Mangel, 1997), implementing ecological models in the programming language R (Bolker, 2008) and, the final step, writing and publishing a paper (Day, 1994). Modern tools such as Git and GitHub are recommended to streamline workflows, particularly for large collaborative projects (Lowndes et al., 2017), and recommendations exist for making computer code available and reproducible (Barnes, 2010; Mislan, Heer, & White, 2016; Wickham,

2015). However, missing from this advice are tips on developing the mathematical notation to use in a model.

Here we recommend some guidelines based on our joint experience of over 25 years of publishing papers in the area of ecological modelling. Our target audience is anyone using mathematical models in ecology, including biologists who are starting graduate training and anticipate using mathematical models, and established field biologists conducting quantitative analyses (such as population models, fisheries stock assessments and movement models). Our guidelines are appropriate for all technical documents, such as journal papers, theses, government reports, consultant reports and documentation for software.

We have sometimes found that mathematical models are described in a more complicated way than necessary—a principal cause is the lack of careful consideration of the notation used. Similarly, we have seen mathematical models that are lacking important information. We ourselves have papers with notation that could, in hindsight, be improved. Unfortunately, poor mathematical notation may mean that a reviewer of a manuscript is unable to comprehend the model used and therefore is unable to properly evaluate a manuscript. This is analogous to a manuscript that has an incomplete description of the methods of a laboratory experiment. Clarity of model descriptions is obviously important. For example, Individual Based Models were previously criticised as generally being so poorly documented that they could not be evaluated or reproduced, which motivated development of standardised protocols that has led to a more rigorous formulation of models and enhanced understanding (Grimm et al., 2010).

By "notation" we are specifically referring to the letters used to represent quantities in equations, as well as the use of subscripts, superscripts and related concepts. Letters are usually taken from the Roman (*a*, *b*, *c*, …) or Greek (α, β, γ, …) alphabets, and can be lower or upper case.

We reviewed the Author Guidelines for a sample of 14 journals: *Bulletin of Mathematical Biology*, *Canadian Journal of Fisheries and Aquatic Sciences*, *Ecology*, *Ecology Letters*, *Evolution*, *Functional Ecology*, *Journal of Animal Ecology*, *Journal of the Royal Society Interface*, *Marine Ecology Progress Series*, *Methods in Ecology and Evolution*, *Molecular Biology and Evolution*, *Nature*, *PLOS ONE* and *Science*. There is minimal guidance regarding mathematical notation. Only two (*Evolution* and *Journal of the Royal Society Interface*) have no mention of equations, but the mathematical-related guidelines for the other journals are almost exclusively restricted to typesetting aspects (such as bold for vectors) rather than the broader considerations that we present here.

Our guidance goes beyond typesetting, and is aimed to help authors think about the choice of notation to improve clarity and understanding. The onus for good notation should be on authors rather than journals (though ideally journals might adopt our guidelines), because: (a) notation should be decided early on in a study (possibly before deciding on a particular journal), (b) work may first appear in a technical document such as a thesis before being submitted to a journal, (c) the Supporting Information of a paper often contains the full mathematical details of models and is not typeset or edited by publishers, such that responsibility for clarity rests with the authors.

Our guidelines are based on those traditionally used in mathematics and are in a somewhat logical order. Our aim is for them to be useful and adopted, though we anticipate exceptions for which they are purposefully overlooked. Where appropriate, we use examples from common ecological models, our own fields of research and from evolutionary biology.

2 | GUIDELINES

2.1 | Define all terms

The number one guideline is to define every term that is used in an equation. For example, although readers will recognise the equation

$$E = mc^2, \tag{1}$$

it does not convey any information as written (since the letters are not defined). Similarly, many ecologists may be familiar with the equation

$$S = cA^z, \tag{2}$$

but it requires the definition of the terms to be understandable to all readers. For this reason, it is imperative to give the equation *and the definitions* of its terms. For example,

$$S = cA^z, \tag{3}$$

where S represents the number of species (of a particular taxonomic category) in area A, the constant c is the number of species that would be in one square unit, and the dimensionless exponent z quantifies the change in species number with area (May, Crawley & Sugihara, 2007). This indicates that (3) represents an increase (because z is typically around 0.2–0.3) in species richness with area. Thus, immediately after an equation (as part of the same sentence) any previously undefined symbols should be defined using the phrase "where …".

There is no need to define a term for a second time in the text, although a reminder may be warranted if the term was introduced much earlier in a lengthy article (and a table of notation can be helpful for complicated models). However, in line with the author guidelines of many journals, it can be desirable for tables and figures to be understandable on their own and so notation should be additionally defined in their captions.

2.2 | Use italics, boldface and capitalisation appropriately

By convention, mathematical symbols should be italicised. This distinguishes text from mathematical notation—for example, "a large value of *a*" is more comprehensible than "a large value of a." However, vectors and matrices should be Roman type (not italicised) and bold, with vectors being lower case and matrices being upper case. So $\mathbf{a}$ would be a vector and $\mathbf{A}$ would be a matrix. Element i of $\mathbf{a}$ is often denoted a_i and the element in row i and column j of $\mathbf{A}$ is hence A_{ij} (or a_{ij}, e.g. Caswell, 2001; Kreyszig, 1993). A vector $\mathbf{a}$ will usually be a column vector unless

otherwise specified. Generic random variables are usually denoted by upper-case letters (like X), with possible numeric values represented by the corresponding lower-case letter (x).

Use Roman type for standard mathematical functions such as sin, cos, log, ln and e^x, and when using other words, such as describing a statistical distribution like

$$X \sim \text{Normal}(\mu, \sigma^2) \tag{4}$$

for variable X coming from a normal distribution with mean μ and standard deviation σ. Also use a Roman d for derivative:

$$\frac{dX}{dt} \tag{5}$$

for the derivative of variable X with respect to time t. Similarly for integration, e.g. $\int f(x)dx$. Italics and Roman type may be more easily distinguishable in a serif font (see Supporting Information).

Units should always be given and also be in Roman type, e.g. the speed of the polar bear was 1 km/hr. This distinguishes units from mathematical variables, and allows use of phrases such as: let the speed be x km/hr.

2.3 | Use subscripts appropriately

Subscripts are used to represent different values of a quantity. For example, define B_t as the biomass of a population in year t, where $t = 1, 2, 3, \ldots, T$, and T represents the maximum year. Note that you cannot then use, say, B_s to represent the biomass in spatial area s. This is because setting $t = 1$ gives B_1 as the biomass in year 1. But setting $s = 1$ also gives B_1, creating obvious confusion.

In practice, there may be interest in modelling the biomass in area s in year t (for various combinations of s and t), in which case two subscripts are needed: B_{st}. Extending this idea, Fung, Farnsworth, Reid, and Rossberg (2012) analysed data from fish trawl surveys and defined B_{ijkmn} as the biomass caught per hour (g/hr) of taxonomic group i, length class j and haul k by vessel m in year n. This notation succinctly describes the detailed structure of the data and makes Fung *et al.*'s subsequent calculations clear and unambiguous. For brevity, there may be no need for a comma ($B_{s,t}$) until numbers are inserted and there could be ambiguity ($B_{3,17}$ rather than B_{317}), although this can be a matter of personal preference, and we ourselves have differing inclinations.

An abbreviation can be used as a (nonitalicised) subscript to indicate related definitions. For example, Olajos et al. (2018) used P_{fp} and P_{fn} to represent the respective probabilities of false positives and false negatives when modelling DNA records from sediments to study processes such as evolutionary divergence. Similarly, B_{MSY} is often used in fisheries science to represent the biomass of a stock at the maximum sustainable yield (MSY). So MSY is being used not as an index (like in B_t above), but as a nonitalicised acronym to distinguish B_{MSY} from the related B_t. Just avoid trying to use *MSY* alone as a variable (see below)—if Y_t represents the yield in year t, then Y_{MSY} would be notationally consistent with B_{MSY}.

Usually, a subscript that is used as an index should appear on both sides of an equation, such as in

$$y_t = 2x_t + 3, \tag{6}$$

for the relationship between two variables x_t and y_t at each time t; there is a value of y_t corresponding to each setting of t. There should generally not be a mixture of subscripts, such as

$$y_t = 2x_i + 3. \tag{7}$$

This equation implies that y_t depends on the value of x_i, and therefore on i, and so y_t should really be denoted y_{it}. Otherwise, for example, we may have $y_t = 9$ for $x_1 = 3$ (with $i = 1$), but $y_t = 13$ for $x_2 = 5$ (with $i = 2$)—but the notation y_t does not distinguish between the different values for $i = 1$ and $i = 2$.

An expression such as

$$y_{ij} = 2x_i + 3 \tag{8}$$

is valid—it just means that the ij value of y is the same for all values of the index j (since nothing on the right-hand side depends on j).

One source of confusion is having i and j as indices, but then using these again in a summation to sum N values:

$$y_{ij} = 2x_{ij} + \sum_{i=1}^{N} x_{ij}. \tag{9}$$

The problem is that i is used here as an index, such that the equation is valid for all values of i, but then is also used in the summation where it is just a dummy index. A better formulation would be

$$y_{ij} = 2x_{ij} + \sum_{k=1}^{N} x_{kj} \tag{10}$$

where k is a dummy index. Another way to understand this is to realise that the following are all equivalent:

$$\sum_{k=1}^{N} x_{kj} = \sum_{l=1}^{N} x_{lj} = \sum_{m=1}^{N} x_{mj}, \tag{11}$$

because k, l and m are just dummy indices used to indicate the terms to be added together by the summation—switching between them in Equation (10) will not affect any related equations, whereas changing i and j likely will.

For spatio-temporal situations, an alternative to the aforementioned B_{st} is using $Y_t(s)$ to represent the value of a random variable Y at time t and spatial location s (Cressie & Wikle, 2011). For times t_1, t_2, and t_3, and spatial locations s_1, s_2, and s_3, such notation enabled Cressie and Wikle (2011) to succinctly represent the spatial process at the fixed time t_1 as

$$\mathbf{Y}_{t_1} = \left(Y_{t_1}(s_1), Y_{t_1}(s_2), Y_{t_1}(s_3)\right)', \tag{12}$$

where $'$ represents the transpose, and the temporal process at the fixed spatial location s_1 as

$$\mathbf{Y}(s_1) = \left(Y_{t_1}(s_1), Y_{t_2}(s_1), Y_{t_3}(s_1)\right)'. \tag{13}$$

Note that $\mathbf{Y}_{t_1}$ and $\mathbf{Y}(s_1)$ are vectors of random variables but cannot be simultaneously lower-case bold (as vectors should be) and upper-case italics (as random variables should be), further emphasising the need for clear definitions.

2.4 | Be careful with superscripts

Superscripts can be used to distinguish two related variables, e.g. X and X'. Although be aware that, as in Equation (12), single quotes are sometimes used to designate the transpose of vectors or matrices (e.g. $\mathbf{X}'$, though $\mathbf{X}^t$ or $\mathbf{X}^T$ are also common), or to represent the derivatives of functions (e.g. $f'(x)$). Asterisks (e.g. X^*) traditionally represent the steady state of a dynamic variable.

A number or a letter (as an index) should not be used as a superscript. For example, we have seen B^t used as the biomass in year t, but this looks like B raised to the power t. When we explicitly set $t = 2$ we get B^2, which is interpreted as B squared, not the desired biomass in year 2. Another example is using u_{at}^{sg} to represent the exploitation rate of fish that are of sex s and age a being caught by fishing gear g in year t. To avoid the superscripts it is fine to use u_{atsg} like in the earlier example of multiple subscripts (B_{ijkmn}).

Subscripts and superscripts should always come after the variable, to avoid confusing notation such as ${}_jv_t$. Otherwise, if θ multiplies ${}_jv_t$ to give $\theta_j v_t$ then there is ambiguity as to whether the first component of this term should be interpreted as θ_j or just θ.

2.5 | It can be helpful to distinguish variables from parameters

It is essential to understand the differences between variables, parameters and constants in a model (Platt & Sathyendranath, 1993). To help emphasise the distinction between variables and parameters, upper-case letters are often used for variables, whereas lower-case (and Greek) letters designate parameters and constants. The dependent variable is usually on the left-hand side of an equation and depends on everything on the right-hand side. An example is the following formulation of the Ricker model (Bjorkstedt, 2012):

$$R = \alpha S e^{-\beta S}, \tag{14}$$

where R is the dependent variable (the recruitment of new fish) arising from the independent variable S (the spawning stock biomass), with parameters α (the maximum number of recruits produced by each unit of spawning stock biomass) and β (which scales the intensity of density dependence). If there is a choice (such as when developing a new model) then it is useful to use upper and lower case to distinguish variables from parameter and constants. Though it can be best to stick with established convention when this is not followed, for example in Equation (1), or for simple equations involving variables denoted as x and y with no letters used for parameters, as in Equation (6)—we used x and y there since they are often the first choice to represent unknown variables.

2.6 | Avoid multiletter variable names

Use only one letter (rather than two or more) to represent a quantity. For example, in fisheries science it is common to see the abbreviation SSB for spawning stock biomass. This is fine as an acronym in a sentence, but can become problematic when SSB is used as a mathematical quantity in an equation. A subscript t might then be added to represent time: SSB_t. But if there is a quantity S defined as, say, selectivity, and B_t is defined as the total (spawning plus nonspawning) biomass in year t, then an equation such as

$$\frac{SB_t}{SSB_t} \tag{15}$$

is very ambiguous. Can the S's be cancelled? What about the B_t's? Does the denominator represent S^2 multiplied by B_t? A solution would be to use B_t and T_t for, respectively, the spawning and total biomasses at time t.

Occasionally it may be okay to use an acronym or word as a variable name. For example, Zuur, Hilbe, and Ieno (2013) often use words as variables in their statistical models, resulting in terms such as $e^{\beta_1 + \beta_2 \times MeanDepth_i}$ which intuitively represents an exponential effect of a linear function of mean depth (β_1 and β_2 are parameters). In such statistical models there is generally no further subsequent mathematical manipulation which may avoid the problems outlined by Equation (15). The use of words or acronyms can make models more understandable to, say, stakeholders who are not quantitatively trained but have insights into the system being modelled—this also may be particularly appropriate for a presentation or a poster, especially if not all details of a model are going to be given. And using words or acronyms does not require people to remember notation. However, it should be ensured that there is no potential for confusion (which can be hard to guarantee when first defining notation) or for equations to become cumbersome and hard to understand. One approach can be to simultaneously give equations in word and notation form (e.g. Edwards & Brindley, 1996).

2.7 | Fully define probability distributions

A discrete random variable takes discrete values (e.g. 1, 2, 3, ...), whereas a continuous random variable can take any value within a specified range (e.g. between 0 and 10). The probability mass function of a discrete variable X is written as $f(x)$, and is just the probability that X takes each possible value of x, i.e. $f(x) = P(X = x)$, where $P(\cdot)$ stands for the probability of occurrence of the event in parentheses. For example, the Poisson distribution can be used for count data and is represented by

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!},\ x = 0, 1, 2, \ldots \tag{16}$$

where x are the possible values and there is just one parameter $\lambda > 0$ (e.g. Bolker, 2008). The corresponding *cumulative* distribution function is given by the upper case $F(x) = P(X \le x) = \sum_{i=0}^{x} f(i)$. Note that sometimes $F(x)$ is simply called the probability distribution function (Grimmett & Stirzaker, 1990), though this term can be ambiguous (e.g. Cressie & Wikle, 2011 used it for the probability density function) and so it may be best avoided. Explicitly stating that $F(x) = P(X \le x)$ can avoid any confusion.

For a continuous variable X, we have the continuous probability density function $f(x)$ and the cumulative probability distribution

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)du. \tag{17}$$

For example, the lognormal distribution is often used in population dynamics and has

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma x}e^{-(\log x-\mu)^2/2\sigma^2},\ x>0 \tag{18}$$

for parameters μ and σ (e.g. Bolker, 2008); log here is the natural logarithm (this is worth specifying and using $\log_{10}$ for base-10 logarithm, though ln is also used for the natural logarithm). Note the definition of the domain ($x > 0$). Other distributions may take only positive integer values (e.g. Poisson distribution), any value between two bounds (e.g. uniform distribution), or any value between two positive values (e.g. bounded power-law distribution)—see Bolker (2008) and Edwards, Robinson, Plank, Baum, and Blanchard (2017). Defining the domain also helps the author confirm that the distribution is appropriate for the question at hand. For the normal distribution, since X can take any positive or negative value there is generally no need to specify that the domain is given by $-\infty < x < \infty$. Defining the domain of indices (and parameters) is similarly necessary. For example, in state space models of animal movements, observations may start at time $t = 1$, but an unobserved initial location at $t = 0$ needs to be modelled and explicitly explained (e.g. Auger-Méthé et al., 2016).

If a second random variable, Y, is being considered, then its density function is often expressed as $f_Y(y)$, and that for X would become $f_X(x)$. Alternatively, $g(y)$ could be used. Note that $f(y)$ does not work as it is not distinguishable from $f(x)$; the f in Equation (18) represents the lognormal distribution, not the x. The analogous cumulative distribution function for Y would be $F_Y(y)$ or $G(y)$.

Some distributions have a conventional shorthand. For example, $X \sim N(\mu, \sigma^2)$ for a variable X that is normally distributed with mean μ and standard deviation σ; though when using N(0, 2) specify whether the 2 is the standard deviation or the variance. While the shorthand for the Gamma distribution is generally $X \sim \text{Gamma}(a, b)$, it is important to define the parameters as the distribution can be parameterised using shape and either rate, scale or mean.

2.8 | Give equations of a model rather than just computer code

One reason that acronyms or words get used to identify variables may be that this is how they are written in computer code. Using words in code can be helpful because it can make the code easier to read and avoid typographical errors. One solution is to write the equations first and then have a comment in the corresponding code that links the words used in the code to the corresponding mathematical notation. A modern simpler alternative is the `R` package `knitr` (Xie, 2018), that interweaves the text and computer code in a single file, easily enabling the same succinct notation to be used throughout. Although writing the equations first may seem a necessary prerequisite to writing code, some people, especially with the popularity of `R` (R Core Team, 2018), are proficient programmers but are sometimes unable to translate the code into mathematical notation. For example, for a numeric vector `x` in `R`, the command

```
y <- cumsum(x)
```

is defined as creating a vector `y` where each element is the cumulative sum of the elements of `x`. That can seem more intuitive than having to express the same idea using an equation: for a vector $\mathbf{x}$ of length n,

$$y_i=\sum_{j=1}^{i}x_j,\ i=1,2,3,\ldots,n. \tag{19}$$

However, the verbal description is ambiguous, unlike Equation (19).

Similarly, it may not be obvious how to translate `for` loops into equations. But it can be cumbersome to describe in words what the `for` loop is doing, whereas the equation can be described more succinctly. For example, consider estimating the parameter of the simple one-dimensional random walk model

$$X_t=X_{t-1}+\varepsilon_t,\ t=2,3,4,\ldots,T \tag{20}$$

where X_t is location at time t (with $X_1 = 0$) and ε_t is a random independent movement component distributed normally with mean 0 and unknown standard deviation σ. Here, σ is the only parameter to estimate and we could code the appropriate log-likelihood in `R` as:

```
loglik <- 0
for(t in 2:T){
  p <- dnorm(x[t], x[t-1], sigma)
  loglik <- loglik + log(p)
}
```

where `x` is a vector of known data (with `x[1]=0`), and for each `t` in the loop we first calculate the probability of observing the value `x[t]` based on a normal distribution with mean `x[t-1]` and standard deviation `sigma`, and then sum the log of these probabilities for each `t` from 2 to `T` to get the overall log-likelihood. This can be more succinctly described in equation form as

$$\log\left[\mathcal{L}\left(\sigma|\mathbf{x}\right)\right]=\log\left(\sum_{t=2}^{T}\frac{1}{\sqrt{2\pi}\sigma}e^{-(x_t-x_{t-1})^2/2\sigma^2}\right), \tag{21}$$

where $\mathcal{L}(\sigma|\mathbf{x})$, or just $L(\sigma|\mathbf{x})$, is the likelihood of the standard deviation σ given the data vector $\mathbf{x}$ with T elements x_t (with $x_1 = 0$), and the summation term comes from the assumption of normally distributed movement. Often $\ell(\sigma|\mathbf{x})$ represents the log-likelihood; as does $l(\sigma|\mathbf{x})$, although l can be hard to distinguish from 1 or I (capital i).

A command such as `cumsum()` may just be part of the book-keeping involved when writing code to implement a model and may not require an equation. However, a log-likelihood function is usually an essential part of the model and should be written explicitly.

So although a piece of `R` code may be more intuitive than the corresponding equations, a drawback of just supplying code is that it relies on the reader to have knowledge of `R`. Programming languages evolve and fall out of favour, and `R` may not be around

in 20 years, whereas properly documented equations will stand the test of time. In addition, while code can be added in supplementary materials, it is rarely included in the main text and properly verified by reviewers. The equations provide a clear description of what has been done and can often be incorporated in the main text.

2.9 | Abide by conventions (but still define everything)

Table 1 gives some common mathematical uses of certain letters that it is worth being aware of when creating new notation. For example, compare the following descriptions:

1. The population size is given by

$$t_\varepsilon = g t_{\varepsilon-1}(1-t_{\varepsilon-1}) + f_\varepsilon \tag{22}$$

where t_ε is the population in year ε (ε = 1, 2, 3, ..., Γ), g is the intrinsic growth rate at low population size and f_ε is normally distributed random noise with mean σ and variance μ^2.

2. The population size is given by

$$X_t = rX_{t-1}(1-X_{t-1}) + \varepsilon_t \tag{23}$$

where X_t is the population in year t (t = 1, 2, 3, ..., T), r is the intrinsic growth rate at low population size and ε_t is normally distributed random noise with mean μ and variance σ^2.

While the two equations convey the same meaning, the choice of notation makes (22) less understandable than (23). Reading (22) is quite jarring and requires extra effort for the reader (acknowledging that we purposefully chose the notation to make such a point).

Within some fields, certain notation may be fairly standard. However, notation should still be clearly defined, in particular because of the multidisciplinary nature of ecology. It may be best to try to retain the established convention of a particular field, though this may depend on how well thought-out the conventional notation was. If you decide to use nonconventional notation then maybe briefly clarify why (which may help convince others and establish a new convention). With multidisciplinary work it may be hard to retain all conventions (and appease everyone), further emphasising the need to define all notation upon first occurrence.

When learning a new subject area the equations can seem daunting at first and require careful examination to understand. But after reading a number of papers the equations (if conventions are established) often become familiar and require less effort to understand, as indeed does the subject area in general.

2.10 | Use parentheses and brackets only as necessary

Parentheses, (), are used around the arguments of a function, e.g. $f(x)$ in Equation (16), and to denote which calculations in an equation need to be done first, as in Equation (23). Square brackets, [], and braces, {}, are also used if necessary to avoid having too many slightly-different sized parentheses. But parentheses should not be included if not necessary. For example, there is no ambiguity in Equation (16), but writing it as

TABLE 1 Common mathematical usage of particular letters and symbols

Letter/symbol	Common usage
e	usually avoided to prevent confusion with nonitalicised e (=2.718...)
f, g	function, e.g. $f(x) = x^3 + 7$
i, j, k	index, e.g. the ith element of vector $\mathbf{x}$ is x_i
n, N	sample size
o, O	usually avoided to prevent confusion with number 0
t	time
u, v, w	speeds
x, y, z	variables, or co-ordinates in space
$\mathrm{P}(\cdot)$	probability of occurrence of the event in parentheses
X, Y, Z	variables
$\alpha, \beta, \gamma, \theta$	parameters
δ, Δ	difference or change in a variable, ΔX, or a parameter
ε	a small value, or random noise term
μ	mean
π	the value 3.141...
$\prod$	product of the proceeding values
σ	standard deviation
$\sum$	summation of the proceeding values
X^*	steady-state value of X
$\dot{X}$	derivative of X
$f'(x)$	derivative of $f(x)$ with respect to x
$\partial f/\partial x$	partial derivative of $f(x, y)$ with respect to x
$\hat{\theta}$	an estimate of θ

$$f(x) = \frac{(\lambda^x)\left(\mathrm{e}^{-\lambda}\right)}{x!}, \quad x = 0, 1, 2, \ldots \tag{24}$$

introduces extra unnecessary notation and makes the equation appear more complicated than it is. We have seen this happen in practice, and the extra clutter in the equation can impede comprehension. There is almost always no need to use a symbol such as × or · to convey multiplication unless using words as variable names (or to break up long equations for readability).

Parentheses are also used in the form $x \in (0, 1)$ to represent $0 < x < 1$, and similarly $x \in [0, 1]$ means $0 \le x \le 1$; a combination such as $x \in [0, 1)$ consequently means $0 \le x < 1$.

2.11 | Equations should be part of sentences

The equations above are all part of sentences. Some of those in the middle of a sentence may require a comma, as in Equation (6), or not, as in Equation (4), and those completing a sentence are

followed by a period, such as for Equation (19). It is generally worth numbering all equations, even ones that are not explicitly referred to again in the text, to make it easy for others (including your future self) to refer to an explicit equation. Single terms in equations (or very simple equations) can appear within text and not on their own line. Fractions within such lines should be written as a/b not $\frac{a}{b}$.

2.12 | Revise notation early on if necessary

As well as thinking about notation before defining a model, it can be useful to revisit the notation early on in a project once some of the details have become more fleshed out. Once a project has proceeded far enough—for example, two papers already published with thousands of lines of computer code shared with others—it can be very hard to then change the notation to make it clearer. Thus, time spent revisiting notation early on may be time well spent. This is similar to functionalising computer code—it can be hard to take a pause and rewrite code in a more user-friendly way, but such efforts tend to pay off in the future. A related point is the need for careful proofreading of equations, as equations may be reformatted during the publishing process and publishers may have their own minor typesetting rules that authors are not aware of and that differ between journals—these may or may not impact comprehension. Also, letters should have a unique definition in a single piece of work (though we violate that here because we are giving independent examples).

2.13 | An example of confusing notation

One example we have seen that highlights several of the problems outlined above concerns fish growth and is

$$\sigma_a^{s^2} = \left(\frac{sd_a^S}{L_a^S} \right)^2, \tag{25}$$

which relates the standard deviation of the length of a fish of age a and sex s (where $s = 1$ for females and $s = 2$ for males), sd_a^S, to the standard deviation of the distribution of log(length) at age a for sex s, σ_a^s, where L_a^S is the length at age a for sex s. The use of a superscript s to index sex requires another higher-level superscript in the term $\sigma_a^{s^2}$ to denote that the standard deviation term σ_a^s is being squared. The sd on the right-hand side stands for standard deviation (one letter would suffice, especially with s being used elsewhere), and the superscript S is capitalised on the right-hand side but not on the left. Equation (25) is simply scaling a standard deviation by a length value, but looks much more complicated than that due to the choice of notation.

3 | DISCUSSION

We hope these guidelines will be helpful when writing your own equations and will improve the future comprehension and reproducibility of ecological models. We re-iterate that these are guidelines but not rules, and should be overlooked when appropriate.

ACKNOWLEDGEMENTS

We thank our colleagues Jaclyn Cleary, Brooke Davis, Wayne Hajas, James Robinson, Catarina Wor and Brianna Wright for encouragement and constructive feedback. We thank three anonymous reviewers, the Associate Editor and Robert O'Hara for their insightful comments and suggestions that also improved this work. MAM acknowledges an NSERC discovery grant.

AUTHORS' CONTRIBUTIONS

A.E. conceived the idea for this work and then A.E. and M.A.M. wrote it together.

DATA ACCESSIBILITY

This work uses no data.

ORCID

Andrew M. Edwards http://orcid.org/0000-0003-2749-8198

Marie Auger-Méthé http://orcid.org/0000-0003-3550-4930

REFERENCES

Auger-Méthé, M., Field, C., Albertsen, C. M., Derocher, A. E., Lewis, M. A., Jonsen, I. D., & Flemming, J. M. (2016). State-space models' dirty little secrets: Even simple linear Gaussian models can have estimation problems. *Scientific Reports*, *6*, 26677. https://doi.org/10.1038/srep26677

Barnes, N. (2010). Publish your computer code: It is good enough. *Nature*, *467*, 753. https://doi.org/10.1038/467753a

Begon, M., Harper, J. L., & Townsend, C. R. (1986). *Ecology: Individuals, populations and communities*. Oxford, UK: Blackwell Scientific Publications Inc.

Bjorkstedt, E. P. (2012). Ricker model. In A. Hastings, & L. J. Gross (Eds.), *Encyclopedia of theoretical ecology* (pp. 632–636). Berkeley and Los Angeles, CA: University of California Press.

Bolker, B. (2008). *Ecological models and data in R*. Princeton and Oxford: Princeton University Press.

Caswell, H. (2001). *Matrix population models: Construction, analysis and interpretation*. Sunderland, MA: Sinauer Associates.

Cressie, N., & Wikle, C. K. (2011). *Statistics for spatio-temporal data*. Wiley Series in Probability and Statistics, Hoboken, NJ: John Wiley & Sons, Inc.

Day, R. A. (1994). *How to write and publish a scientific paper*. Cambridge: Cambridge University Press.

Edwards, A. M., & Brindley, J. (1996). Oscillatory behaviour in a three-component plankton population model. *Dynamics and Stability of Systems*, *11*, 347–370. https://doi.org/10.1080/02681119608806231

Edwards, A. M., Robinson, J. P. W., Plank, M. J., Baum, J. K., & Blanchard, J. L. (2017). Testing and recommending methods for fitting size spectra to data. *Methods in Ecology and Evolution*, *8*, 57–67. https://doi.org/10.1111/2041-210x.12641

Fung, T., Farnsworth, K. D., Reid, D. G., & Rossberg, A. G. (2012). Recent data suggest no further recovery in North Sea Large Fish Indicator. *ICES Journal of Marine Science*, *69*, 235–239. https://doi.org/10.1093/icesjms/fsr206

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological Modelling*, *221*, 2760–2768. https://doi.org/10.1016/j.ecolmodel.2010.08.019

Grimmett, G. R., & Stirzaker, D. R. (1990). *Probability and random processes*. Oxford, UK: Oxford University Press.

Hilborn, R., & Mangel, M. (1997). *The ecological detective: Confronting models with data*. Vol. 28, Monographs in Population Biology, Princeton, NJ: Princeton University Press.

Kreyszig, E. (1993). *Advanced engineering mathematics* (7th ed.). New York, NY: John Wiley & Sons, Inc.

Lowndes, J. S. S., Best, B. D., Scarborough, C., Afflerbach, J. C., Frazier, M. R., O'Hara, C. C., … Halpern, B. S. (2017). Our path to better science in less time using open data science tools. *Nature Ecology and Evolution*, *1*, 0160. https://doi.org/10.1038/s41559-017-0160

Lowry, W. P. (1959). On the use of mathematical notation in ecological literature. *Ecology*, *40*, 492. https://doi.org/10.2307/1929773

May, R. M., Crawley, M. J., & Sugihara, G. (2007). Communities: Patterns. In R. M. May, & A. R. McLean (Eds.), *Theoretical ecology: principles and applications* (3rd ed., pp. 111–131). New York, NY: Oxford University Press.

Mislan, K. A. S., Heer, J. M., & White, E. P. (2016). Elevating the status of code in ecology. *Trends in Ecology and Evolution*, *31*, 4–7. https://doi.org/10.1016/j.tree.2015.11.006

Olajos, F., Bokma, F., Bartels, P., Myrstener, E., Rydberg, J., Öhlund, G., … Englund, G. (2018). Estimating species colonization dates using DNA in lake sediment. *Methods in Ecology and Evolution*, *9*, 535–543. https://doi.org/10.1111/2041-210x.12890

Platt, T., & Sathyendranath, S. (1993). Estimators of primary production for interpretation of remotely sensed data on ocean color. *Journal of Geophysical Research*, *98*, 14561–14576. https://doi.org/10.1029/93jc01001

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Wickham, H. (2015). *R packages: Organize, test, document, and share your code*. Sebastopol, CA: O'Reilly Media Inc.

Xie, Y. (2018). knitr: A general-purpose package for dynamic report generation in R. R package version 1.20.

Zuur, A. F., Hilbe, J. M., & Ieno, E. N. (2013). *A beginner's guide to GLM and GLMM with R: A frequentist and Bayesian perspective for ecologists*. Newburgh, UK: Highland Statistics Ltd.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Edwards AM, Auger-Méthé M. Some guidance on using mathematical notation in ecology. *Methods Ecol Evol*. 2019;10:92–99. https://doi.org/10.1111/2041-210X.13105