ISSN 1946-7664. MCFNS 2024
Vol. 16, Issue 2, pp. 27–32
Mathematical and Computational Forestry & Natural-Resource Sciences

AVAILABLE ONLINE AT HTTP://MCFNS.COM

Submitted: Jun. 15, 2024
Accepted: Jul. 8, 2024
Published: Oct. 30, 2024
Last Correction: Jul. 16, 2024

MODELING WITH STOCHASTIC DIFFERENTIAL EQUATIONS

Oscar García
Dasometrics, Concón, Chile
ORCID: 0000-0002-8995-1341

Abstract. Stochastic differential equations (SDEs) are increasingly prevalent in a variety of fields. They have become routine in areas like pharmacokinetics and finance. Forestry applications, on the other hand, remain uncommon. This is an accessible introduction to the basic concepts and practical use of SDEs, with an emphasis on forest growth modeling. I begin by briefly discussing dynamical system ideas, describing rates of change in a state space instead of using functions of time directly. Rates of change can be specified by finite differences, but a formulation in continuous time with differential equations is often more convenient. Rational parameter estimation necessitates a stochastic representation of the error structure. Besides observation errors, process noise or environmental variability can be important. Both sources of variability can be taken into account with stochastic differential equations. Simple examples are demonstrated using R software.

Keywords: Dynamical systems, state space, growth, estimation, *R* language.

1 Introduction

The use of stochastic differential equations (SDEs) is growing in many areas. They are used for describing error structures and to estimate parameters in dynamical systems, systems that develop over time subject to management and environmental interventions. Unfortunately, the mathematics and computation of general SDE models can be complex and highly technical, and that is reflected in the available literature. For modeling purposes, however, technical details are either not relevant or an intuitive understanding is sufficient. This introduction should be accessible to practitioners without a specialist mathematical background. The ideas are illustrated with examples from forest stand development, but they are much more general.

The text is structured into two main sections: *Dynamical Systems*, and *Statistics*. These are followed by examples of using the `resde` *R* package in modeling height growth.

The *Dynamical Systems* section begins discussing the limitations of modeling growth directly as functions of time, as in traditional yield tables. A more flexible alternative is to model instead rates of change, which are then accumulated to compute trajectories between any two system conditions. For that to work properly, it is necessary to describe the current condition (or state) by a sufficient number of state variables, with a rate equation for each of them. The ideas are explained initially with annual or periodic rates of change, determining the trajectory at only a discrete set of points in time. It is often more accurate and convenient to consider vanishing small time intervals, expressing the rates of change as derivatives, resulting in differential equations and continuous-time trajectories.

Once we have a (deterministic) model, we need to estimate parameters. Rational parameter estimation requires some model of the variability or uncertainty in the data, the subject of the *Statistics* section. It is common to assume that the deterministic model is exact, with only observation errors. Otherwise, if the rates of change are also uncertain or subject to perturbations, the continuous-time formulation leads to SDEs.

The `resde` package deals with one-dimensional SDEs that can be reduced to linear by a change of variables (*reducible*). The development of top height is well described by a dynamical system with a single state variable, since it is largely independent of density or other stand variables. It happens that practically all published growth models can be written as a linear differential equation on some height transformation.

It is shown first how to fit a model for a single tree or stand. In a site index model, the growth of each stand depends on a site-dependent parameter specific to the stand. The second example estimates site index models considering the parameters as either fixed unknown constants or mixed effects.

The article ends with a concise *Summary*. For additional details and literature references, one might see García (2013) on dynamical systems, and García (2023) on SDEs.

2 Dynamical Systems

2.1 Functions of time, e.g., yield tables

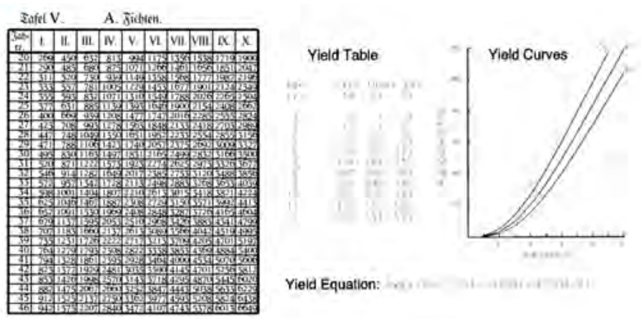

Figure 1: Yield tables (or curves, or equations).

The simplest representation of systems that develop over time is directly as functions of time, as in traditional forestry yield tables. The tables show the course of volume per hectare over age (Fig. 1). They may be parametrized by quantities specific to each stand, such as site quality or initial density:

$$V = F(t, q) \ ,$$

where t is time or age and q is a parameter [1]. In addition to volume V, other variables may be included in the table, e.g., top height H, mean diameter D, trees per hectare N.

2.2 Rates of change

For many applications the direct use of time functions is sufficient. Yield tables are commonly used in forest planning systems to predict growth of unmanaged stands, or under a small number of silvicultural prescriptions. Their simplicity can be crucial in those uses. However, in more demanding situations the direct approach has important limitations.

Assume that at age 40 there is a thinning that reduces the standing volume (Fig. 2). The yield table does not provide a forecast starting at the new point. Even without interventions, a fluctuating environment causes

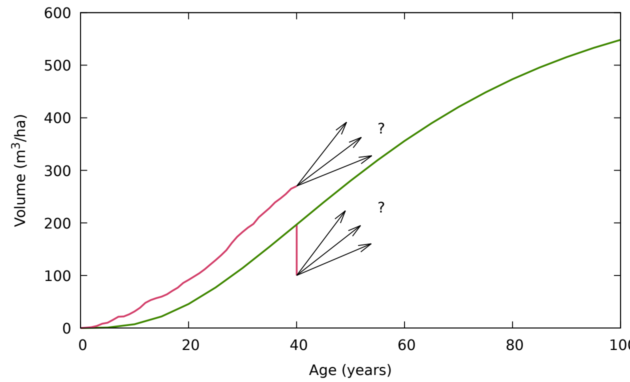

Figure 2: $V = F(t)$, effect of interventions and perturbations. Nominal yield curve, and thinning at age 40 (bottom), or deviation in realized volume (top). Future development is not specified.

deviations from the predicted trajectory. Once off the curve, you are on your own.

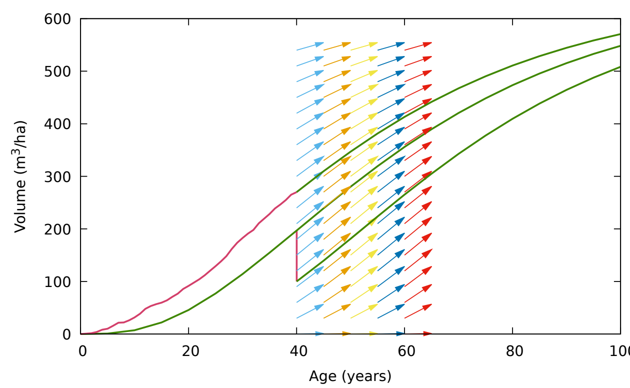

Figure 3: Rates of change (arrows). Predictions follow the arrows.

A more flexible alternative is to model rates of change. Instead of $V = F(t)$, the model predicts an annual or periodic increment

$$\Delta V = f(V)$$

for any current V, and a time increment Δt of 1, 5, or 10 years, for instance[2].

Given the rate of change at any point, trajectories are generated by iteration, accumulating increments (Fig. 3):

$$\Delta V = f(V) \quad \rightarrow \quad V = F(V_0, n\Delta t) \quad \text{(numerically)} \ ,$$

starting from any initial $V = V_0$ and for any multiple of Δt.

[1] A parameter is a quantity that, depending on circumstances, may be considered either as a constant or as a variable. For instance, variable while fitting a model and constant when applying it.

[2] Other common notations include a rate per unit time $\frac{\Delta V}{\Delta t} = f(V)$, or the iteration $V(t + \Delta t) = f[V(t)]$ (with different f's).

2.3 State

$\Delta V = f(V)$ can only be a rough approximation, because volume increment does not depend only on current volume, but varies also with stand density, and possibly with height. A better description of the *state* of the system is a list of state variables (state vector) $\boldsymbol{x} = (H, N, V)$. There is one rate equation for each variable:

$$\begin{cases} \Delta H = f_1(H, N, V) \\ \Delta N = f_2(H, N, V) \\ \Delta V = f_3(H, N, V) \end{cases}$$

Or, in vector shorthand,

$$\Delta \boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}) \ .$$

If you are not comfortable with vectors, think of $\boldsymbol{x}$ as simply the top height H, and ignore the rest of this subsection. The later examples use top height growth, which is assumed to be independent of stand density or other stand characteristics. A single equation $\Delta H = f(H)$ is then sufficient.

Some other transformation of the state could be used as well. For instance, the state vector (B, S, H) contains essentially the same information, with basal area B and average spacing S. Other variables of interest can be obtained from the state through *output functions*, such as a simple volume relationship $V = 0.4BH$, or more complex estimates of ecosystem services. The rate equations might also include *input functions*, such as representations of a fluctuating environment through time-dependent climatic variables.

2.4 Discrete time

Finite intervals Δt are usually dictated by the remeasurement periods present in the data, e.g., 1, 2, 5, or 10 years, depending on the species growth rate. Annual or periodic rate of change equations are easy to understand and use. Projections are generated by simple arithmetic, $\boldsymbol{x} = \boldsymbol{F}(\boldsymbol{x}_0, n\Delta t)$ is obtained by iteratively adding $\Delta \boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x})$ n times.

A disadvantage is that irregular measurement intervals waste information or require questionable approximations. And projections are limited to a discrete set of times, multiples of Δt. This may be inconvenient in some applications, including sub-annual estimation in fast-growing species.

2.5 Continuous time, ODEs

An alternative is to model in continuous time. Let us use the rate of change per unit time

$$\frac{\Delta \boldsymbol{x}}{\Delta t} = \boldsymbol{f}(\boldsymbol{x})$$

and think of a *very* small Δt. Then, the "instantaneous" rate of change is modeled by an (ordinary) differential equation, ODE,

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{f}(\boldsymbol{x}) \ .$$

This equation can be integrated analytically or numerically to project any state $\boldsymbol{x}_0$ at time t_0 to the state at any other time t:

$$\boldsymbol{x} = \boldsymbol{F}(\boldsymbol{x}_0, t - t_0) \ .$$

This is sometimes called a trajectory or a transition function[3]. Fig. 4 shows observed trajectories in a (B, S, H) state space.

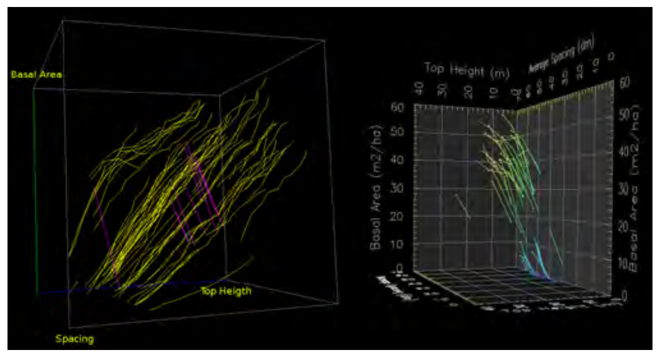

Figure 4: Observed trajectories of forest plantations in a three-dimensional state space. Left: Radiata pine in New Zealand. Right: Interior spruce in British Columbia. Thinnings are jumps from one trajectory to another.

Cohort or individual-tree models contain more state variables (possibly hundreds), but the concepts are the same.

The trick of using rates of change instead of functions of time can be attributed to Isaac Newton and is taken for granted in physics and engineering. Physics applications have some peculiarities specific to them. The fundamental principles were abstracted and generalized by System Theory in the mid-20th Century.

3 Statistics

3.1 Parameter estimation

Once the (deterministic) model is formulated, it is necessary to estimate parameter values to fit the data. Sometimes, ad-hoc methods are used, prioritizing convenience or computational feasibility. A more reasoned approach requires a model for the variability or uncertainty. A

[3]V. I. Arnold ("Ordinary Differential Equations", The MIT Press, 1973) reverses this point of view. He considers the set of all the trajectories, called a "flow", as the more fundamental concept, and derives the theory of ODEs from that.

probabilistic (stochastic, statistical) component on top of the deterministic part.

Given the statistical model, one typically finds the probability of generating the observed data for given parameter values. That is the *likelihood function*, used for obtaining maximum-likelihood (ML) or Bayesian parameter estimates.

Experience suggests that frequently the realism of the stochastic component makes little difference. An over-elaborate statistical model can be counterproductive, wasting information in estimating the stochastic part at the expense of precision in the predictions.

3.2 Observation error

It is still common to assume that the trajectory predicted by the dynamic model is exact and that there are only errors in the observations. The model, e.g., an ODE, predicts values $x(t_i)$ at a set of observation times $t_1, t_2, \ldots, t_n$. But one observes

$$y_i = x(t_i) + \varepsilon_i \ ,$$

where the ε_i are usually assumed to be identically and independently distributed, often Normal, with mean 0.[4].

3.3 Process noise, SDEs

More recently, in some applications the process is viewed as noisy, with the trajectories $x(t)$ subject to perturbations or uncertainty (Fig. 5). This leads to stochastic differential equations (SDEs).

An SDE is an ODE perturbed by noise, such as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x) + g(x)\dot{w}(t) \ ,$$

where $\dot{w}(t)$ is "white noise" and $g(x)$ is a scale factor. The scale factor may or may not depend on the state x. White noise is characterized by having the same distribution with mean 0 for every t, and being uncorrelated for any two different t's.

This notation, a *Langevin equation*, is commonly used by physicists and is perhaps easiest to understand. But completely uncorrelated noise has some weird properties, like being non-differentiable and having an infinite variance. Thus, in mathematics and statistics it is preferred to work with the cumulative noise $W(t)$, called a Brownian motion or Wiener process, and to write the SDE in terms of differentials:

$$\mathrm{d}X(t) = f[X(t)]\,\mathrm{d}t + g[X(t)]\,\mathrm{d}W(t) \ .$$

The capital letters emphasize the fact that $X(t)$ and $W(t)$ are random variables.

Still, there are technicalities related to the definition of the integral of the second term (Ito and Stratonovich integrals), unless g does not depend on X where the definitions coincide. Langevin and others formulated SDEs to simplify the approach of Einstein in his famous paper of 1905 on Brownian motion. Hotelling used SDEs for logistic growth in 1927[5]. Not until the middle of the Century it was found that a rigorous mathematical treatment was surprisingly subtle and complex. However, for modeling applications, an intuitive understanding should suffice.

Again, all this can be generalized to vectors. Observation errors can be included as discussed before.

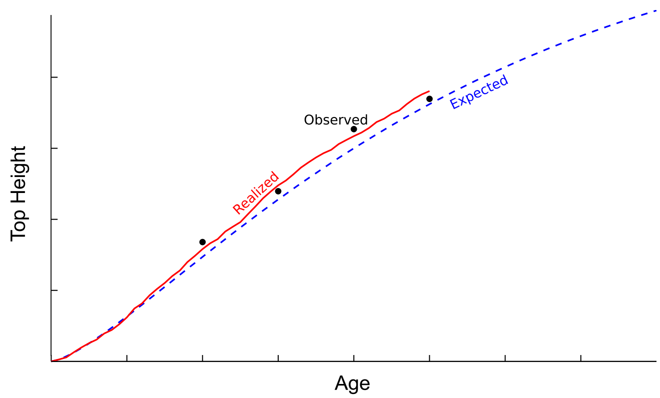

Figure 5: Sources of error in a height trajectory. The dashed blue curve is the ODE prediction. Disturbances (process noise) lead to the solid red curve. The (disturbed) process is observed at a finite set of times, possibly with observation errors.

4 Example

4.1 Height growth. Reducible SDEs

Linear SDEs are some of the few easily tractable mathematically. Fortunately, practically all growth equations can be reduced to a linear ODE/SDE through some transformation of the state variable. We shall use the *R* package `resde`, which computes ML estimates for reducible univariate SDEs. The package is available in the official *R* repository and converts the ML problem into

[4]For more than one dimension, think of these variables as vectors. In that case, only some of the components of x might be observed. In general, one can write

$$y_i = h[x(t_i)] + \varepsilon_i$$

for some observation function $h()$ such that the dimension of y may be less than the dimension of x.

[5]Harold Hotelling, "Differential Equations Subject to Error, and Population Estimates", *Journal of the American Statistical Association 22*, 283–314, 1927.

the minimization of a sum of squares. This is more efficient and reliable than using general-purpose optimization routines.

The general model form in `resde` is a linear SDE

$$\mathrm{d}Y = (\beta_0 + \beta_1 Y)\,\mathrm{d}t + \sigma_p\,\mathrm{d}W\ ,$$

where Y is some transformation $Y = \varphi(X)$ of the variable of interest X. The observations x_i can have errors according to

$$\varphi(x_i) = Y(t_i) + \sigma_m \varepsilon_i$$

(subscript p stands for *process* and m for *measurement*).

The examples use a Richards ODE, which is linear in a power transformation of H:

$$\frac{\mathrm{d}H^c}{\mathrm{d}t} = b(a^c - H^c)\ .$$

Expanding the derivative on the left-hand side and rearranging gives the more familiar form

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \tfrac{a^c b}{c} H^{1-c} - \tfrac{b}{c} H\ .$$

It is shown first how to fit the model for a single tree or stand. Then, we estimate a site index model, where a growth model is fitted simultaneously to many stands, with a site-dependent parameter that varies among the individual stands.

4.2 Single tree or stand

Install and load `resde` from the official CRAN repository. The package has two main functions, `sdemodel()` specifies the model form, and `sdefit()` performs the ML estimation.

Establish the correspondence between `resde`'s notation

$$\mathrm{d}Y = (\beta_0 + \beta_1 Y)\,\mathrm{d}t + \sigma_p\,\mathrm{d}W$$

and our

$$\mathrm{d}H^c = b(a^c - H^c)\,\mathrm{d}t + \sigma_p\,\mathrm{d}W\ :$$

```
mymodel <- sdemodel(phi = ~x^c, beta0 =
  ~b*a^c, beta1 = ~-b)
```

These arguments are *R* formulas, indicated by ~. Keep the defaults for the rest.

Extract tree number 301 from the `Loblolly` data set that comes standard with *R*:

```
lob301 <- Loblolly[Loblolly$Seed == 301, ]
```

`Seed` is the name of the column with the tree identifiers, or stand identifiers if we assume that the data corresponds to stand heights.

Now run the estimation procedure, indicating the model, data columns, and initial parameter values:

```
fit <- sdefit(mymodel, x="height", t="age",
  data=lob301, start=c(a=60, b=0.1, c=1))
fit
```

The list `fit` contains fit statistics, and ML estimates for a, b, c, σ_p, and σ_m.

4.3 A site index model

We use the same model form, already stored in `mymodel`. But now we fit the 14 trees (or stands) in the `Loblolly` data set simultaneously.

For this example, assume that the asymptote a is a *local* site-dependent parameter, different for each stand. The other parameters are *global*, common to all the stands. This generates so-called anamorphic site index curves, where the stand curves are proportional along the H-axis.

We consider two ways of handling local parameters:

(a) Fixed effects. The local is a fixed unknown value for each stand.

```
alocalF <- sdefit(mymodel, x="height",
  t="age", data=Loblolly, unit="Seed",
  local=c(a=70), global=c(b=0.1, c=0.5))
```

`unit` is the column with the stand identifiers. The parameters and starting values are separated into locals and globals.

(b) Mixed effects. The local is thought of as "random", in some sense, normally distributed across stands. The fitting command is the same as before, except that `method` is specified as "nlme" instead of the "nls" default:

```
alocalM <- sdefit(mymodel, x="height",
  t="age", data=Loblolly, unit="Seed",
  local=c(a=70), global=c(b=0.1, c=0.5),
  method="nlme")
```

Estimation methods and parametrizations can be compared with the help of the likelihood values and AIC and BIC statistics returned by `sdefit`.

For more details and examples, see the `resde` documentation.

5 Summary

- It is advantageous to model rates of change instead of direct functions of time.

- Continuous time (ODEs) may be less obvious than discrete time (finite differences), but it is often more convenient.

- Rational parameter estimation requires a (simple) stochastic model.
- Process variability or uncertainty in dynamical systems is represented by SDEs.
- The mathematical theory of SDEs is subtle and complex but rarely needed for applications.
- Software like `resde` makes modeling with SDEs accessible.

Acknowledgements

This article is based on a presentation at the Symposium on Systems Analysis in Forest Resources, Hondarribia, Spain, May 2024. I am grateful to the organizers who made possible my participation.

Further Reading

García, O. (2013). Forest Stands as Dynamical Systems: An Introduction. *Modern Applied Science 7*: 32–38. DOI: 10.5539/mas.v7n5p32 (Spanish translation at `https://www.researchgate.net/publication/352536288`)

García, O. (2023). resde: Estimation in Reducible Stochastic Differential Equations (Vignette). CRAN, The Comprehensive R Archive Network. `https://cran.r-project.org/package=resde`