FITTING THE GROWTH AND REPRODUCTION TRAJECTORIES OF UNINFECTED ANIMALS

For this fitting, I need to use parameter values that make more sense for *Daphnia dentifera*. Hall et al. 2009 report that the mass of the host is $W = \alpha L^3$, where $\alpha = 1.8 \times 10^{-3}$ mgC mm^{-3}, and they assume (Appendix A) that half of the measured dry mass is carbon. L is the observed length, rather than the "structural" length of DEB theory; Spencer's formulation of the standard DEB model didn't involve the structural length, so this parameter wasn't needed. In other words, α was estimated using observed dry weights W, observed lengths L, and an assumption that $W/2$ is the carbon mass of the host.

This is slightly different than what I have been doing, but is not incompatible with it. I had assumed that observed mass $W_{obs} = W + E$, where W is the "structural" mass and E is the reserves, and that observed length was $L_{obs} = \xi W_{obs}^q$, where ξ and q are based on a regression of observed weight on observed length. I then assumed that "structural length" L was equal to $W^{1/3}$. Basically, all I need to do is assume that $\xi = \alpha$ and $q = 3$ and I will have recovered more or less the same parameterization.

I need better values from Spencer et al. for some of the parameters, so for my first attempt at fitting Cat's data for uninfected animals, I fixed $E_R = 1.51 \times 10^{-3}$, the initial biomass as $W + E = 0.00225$, and the biomass at maturity as 0.005. I also fixed $v = 10$ for the fitting.

The results are not particularly encouraging, frankly (Fig. 1). The parameters are fairly tightly estimated at the highest likelihood parameter sets; in particular, $\rho = 0.15$, $\kappa = 0.077$, and $L_{obs} = 0.088$. However, that κ value is much too low to make biological sense. Neither f_h nor k_m are well-identified in the very highest likelihood parameter sets: for parameters sets within 1 log-likelihood unit of the maximum, f_h varies between practically 0 and 10000 (although all but one estimate is less than 40, and two are essentially 0), while k_m varies between essentially 0 and 0.008 (which is likely much too small). There is also a very strong positive correlation between the estimates of κ and k_m that I haven't noticed before in the simulation-recovery fits.

The fact that the data seem to want $f_h = 0$, implying food-independent feeding, is interesting/strange. I cannot really explain why that should be favored.

I think the best thing to do is to try fixing f_h at different values, and see how that affects the other parameter estimates. Although I doubt if there is a parameter set with a higher likelihood that was overlooked by the algorithm, my sense is that it may be that the likelihood surface is quite flat, so f_h can run off towards 0, dragging all of the other parameter estimates with it. The profile likelihood will be very informative in this regard.

Suprisingly, the profile likelihood suggests that, indeed, the likelihood does increase in the direction of smaller f_h values (Fig. 2). You can see some interesting patterns in Fig. 2. This figure shows all of the parameter estimates within two log-likelihood units of the maximum for each value of f_h. You can see that there are "lines" showing particular relationships between, e.g., ρ and f_h and the log-likelihood and f_h suggesting multiple fitness peaks with different scaling relationships between the variables.

These multiple likelihood peaks can be seen in pairwise scatterplot of the parameter estimates, for example, for $f_h = 9000$ (Fig. 3). Note, in particular, in the density plot that there are multiple likelihood peaks and multiple density peaks for each parameter.
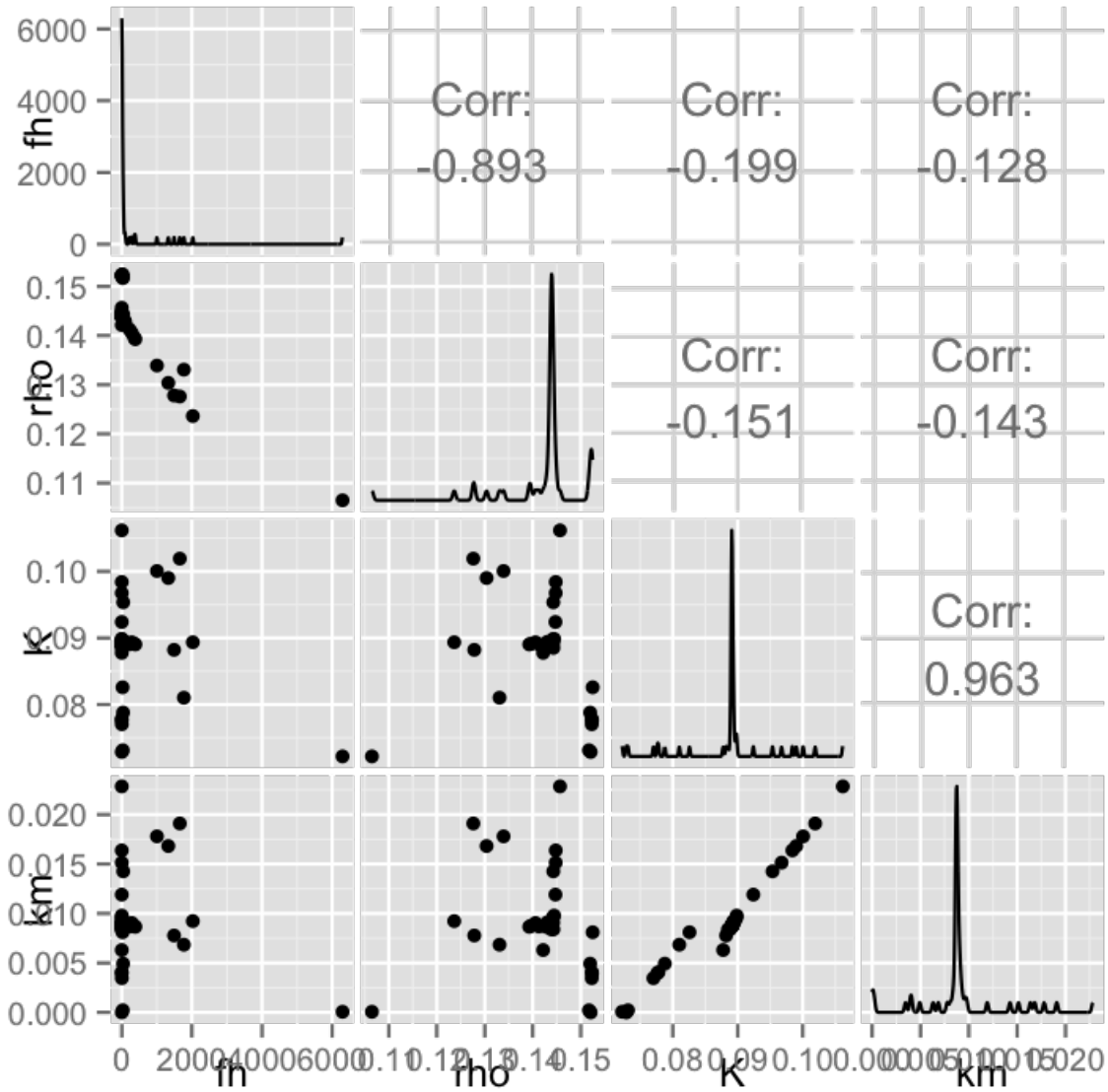
FIGURE 1. Pairwise plot of parameter estimates from fitting Cat's growth and reproduction dataset.

The larger problem, however, is that these parameter values still seem highly biologically unrealistic and there is a very strong correlation between the estimates of κ and k_m, suggesting that these might be difficult to independently estimate. One concern that I have is that some of the fixed parameter estimates might be very far from the truth. For example, the estimate of κ being less than 0.1 indicates that the vast majority of energy is being shunted into reproduction. This may be because the cost of reproduction E_R is too high. Recall that I was assuming $E_R = 1.51 \times 10^{-3}$. If this cost is equal to the mass of a neonate, this would suggest a neonate length of $(0.00151/0.0018)^{1/3} = 0.94$mm, which is much too large. I refit the model, assuming
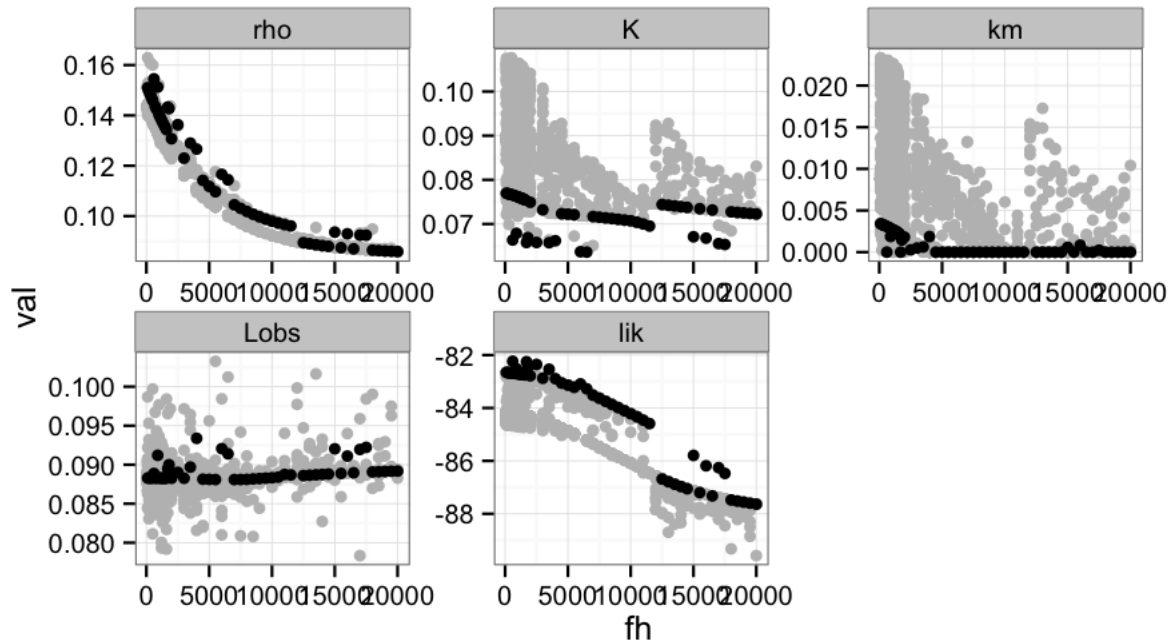
FIGURE 2. Profile likelihood for f_h showing the best-fitting parameter estimates for Cat's uninfected *Daphnia* as f_h is varied.
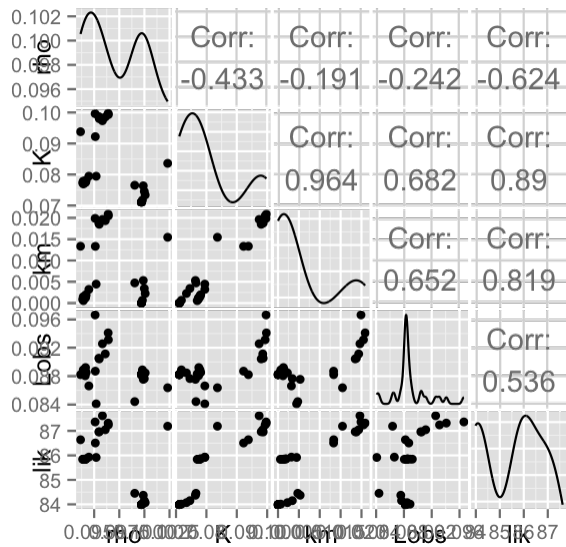

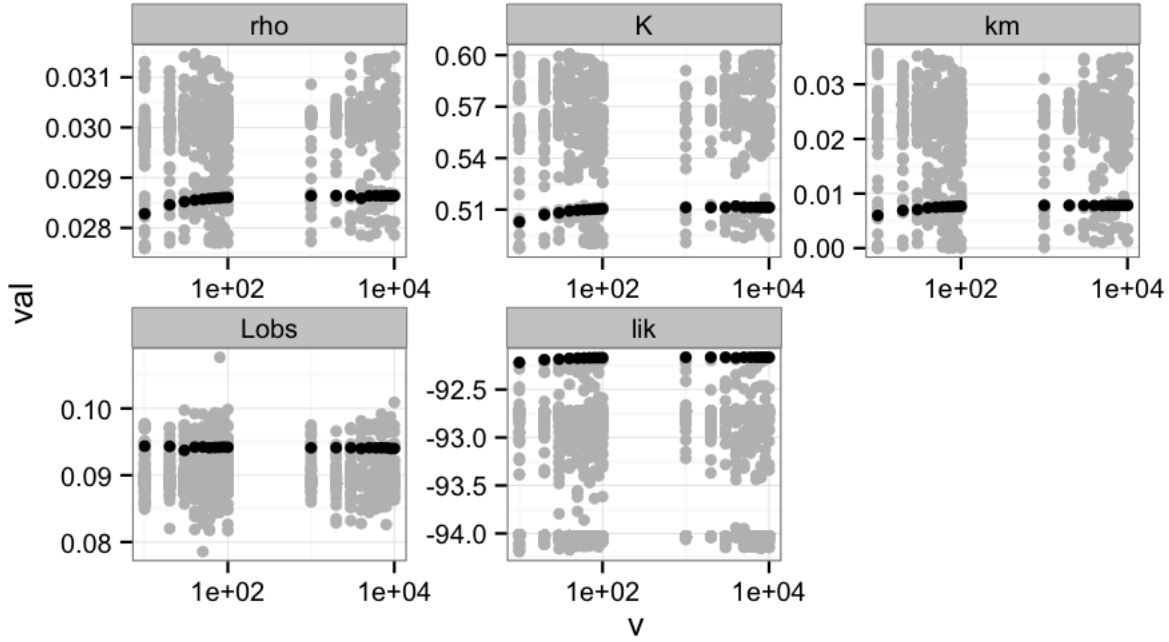FIGURE 3. Scatterplot of parameter estimates for $f_h = 9000$.

FIGURE 4. Parameter estimates as v is varied, when f_h is fixed. The black point is the maximum likelihood parameter estimate.

a cost of reproduction equal to 2.25×10^{-4}, which is equivalent to the mass of a 0.5mm neonate. I also reduced the mass at maturity from 0.005 to 0.002, based on the estimate from Hall et al. 2009. Finally, I fixed the value of f_h at 2250 cells/ml, which is equivalent to 0.1mgC/L, the estimate reported in Hall et al. 2009. Which those values fixed, I esimated only ρ, κ, k_m, and the observation error. I constructed a profile likelihood over v, just to see if there was still no evidence for an effect of v on the fits. Fig. 4 shows the results. It is clear that the parameter estimates are not much affected by v. In fact, even taking v as large as 10000 has no effect on the other parameter estimates, suggesting that the data really do support a reserve-less model. The estimates of κ and k_m, at any rate, seem more reasonable as well. However, the estimate of ρ is *very* small, suggesting that the model is perhaps inadequate in some way.

I want to point out that the seeming variation in the parameter estimates is a bit illusory. Most of the parameter estimates are essentially lying directly under the maximum likelihood estimate, as can be seen in Fig. 5, which shows the distribution of parameter estimates for $v = 10$.

However, this distribution also reveals the *very* strong correlations between the estimates of ρ, κ, and k_m.
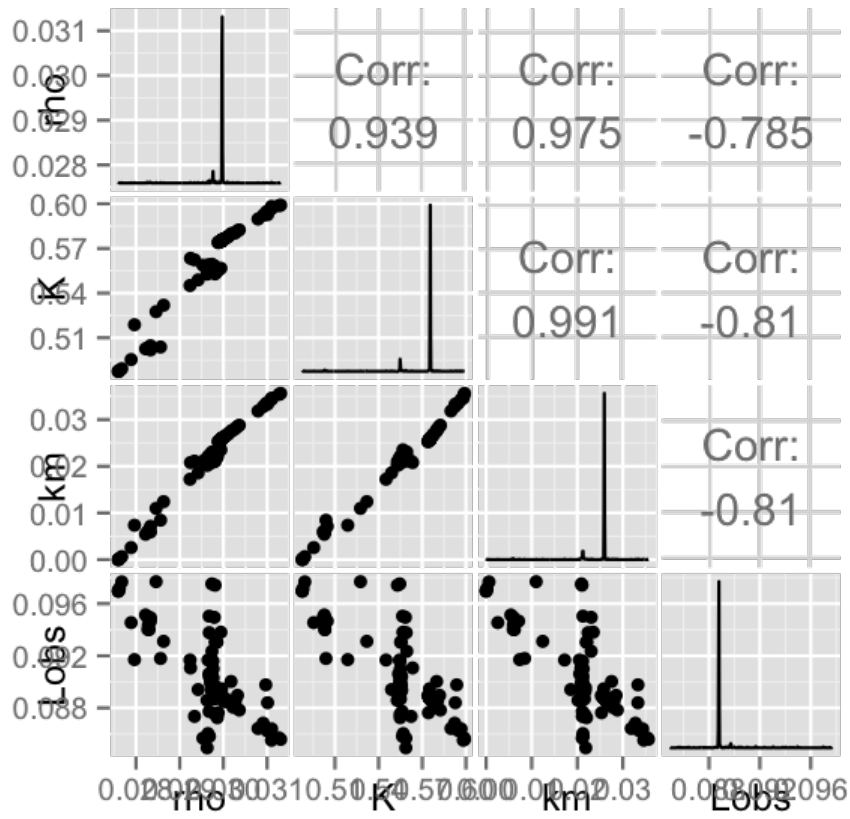
FIGURE 5. Distribution of parameter estimates when $v = 10$ and f_h is fixed.

I have a hypothesis both about why the estimate of ρ is so low and why the estimates of ρ, κ, and k_m are so strongly correlated.

As to the former question, recall the model that I am actually fitting to these data:

$$\frac{dF}{dt} = -I_{max} \frac{F}{f_h + F} L^g_{obs} \tag{1}$$

$$\frac{dE}{dt} = \rho \epsilon V I_{max} \frac{F}{f_h + F} L^g_{obs} - p_C, \tag{2}$$

$$\frac{dW}{dt} = \kappa \, p_C - k_M \, W, \tag{3}$$

$$\frac{dR}{dt} = \frac{(1 - \kappa) \, p_C}{E_R}, \tag{4}$$

$$p_C = E \left(\frac{\frac{v}{L} + k_m}{1 + \frac{\kappa E}{W}} \right). \tag{5}$$

This model is missing a term for the cost of growth. I am assuming that x mgC of reserves allocated towards growth becomes x mgC of new structural biomass. In the standard DEB model, there is a conversion cost, such that growth in weight is actually:

$$\frac{dW}{dt} = \frac{\kappa \, p_C - k_m \, E_G \, W}{E_G}, \tag{6}$$

and the mobilization rate is

$$p_C = E \left(\frac{\frac{v}{L} + k_m}{1 + \frac{\kappa E}{E_G W}} \right). \tag{7}$$

With this cost added in, the animals have to be more efficient at assimilation in order to meet all of the costs.

However, it turns out that adding this parameter doesn't make the estimation any easier or more reasonable, as you might expect based on the previous findings regarding the trade-offs between ρ and E_R (Fig. 6). I fit the standard DEB model defined above to the data, assuming the same fixed parameter values but also attempting to estimate E_G. Fig. 6 shows the distribution of parameter estimates that were within 2 log-likelihood units of the maximum. In addition to ρ still being estimated as very small, most of the κ estimates are near 1. There are also clear and strong trade-offs between ρ and E_G, as expected, as well as between ρ and κ, k_m and E_G, and κ and E_G, and ρ and k_m.

These results reinforce the question, raised earlier, about the correlation between parameter estimates. This is almost certainly because the model is overparameterized. The most efficient way of dealing with that problem is likely to nondimensionalize the model. That will reveal parameter combinations that are important to the dynamics.

To that end, let's take a look at the standard DEB model again. Here, I have replaced W with V to be totally consistent with the standard DEB model. The volume of the container is
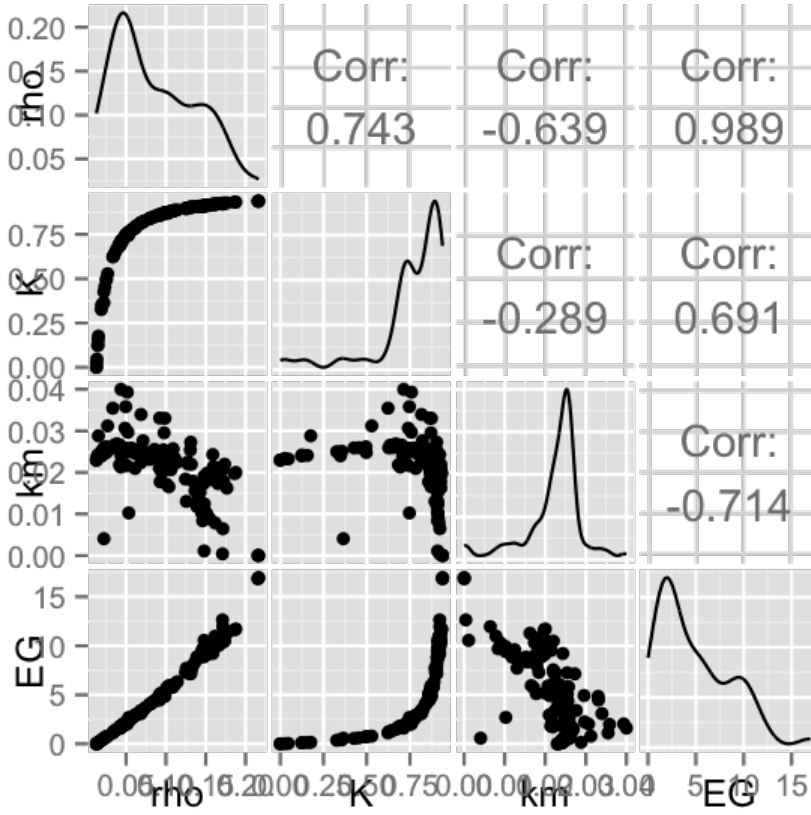
FIGURE 6. Correlations between the parameter estimates when E_G is also estimated.

given by $\mathcal{V}$.

$$\frac{dF}{dt} = -I_{max}\frac{F}{f_h + F}L_{obs}^g \tag{8}$$

$$\frac{dE}{dt} = \rho\epsilon\mathcal{V}I_{max}\frac{F}{f_h + F}L_{obs}^g - p_C, \tag{9}$$

$$\frac{dV}{dt} = \frac{\kappa\, p_C - k_m\, E_G\, V}{E_G}, \tag{10}$$

$$\frac{dR}{dt} = \frac{(1-\kappa)\, p_C}{E_R}, \tag{11}$$

$$p_C = E\left(\frac{\frac{v}{L} + k_m}{1 + \frac{\kappa E}{E_G V}}\right). \tag{12}$$

The presence of the E/V term in the denominator is suggestive: rather than working with a model parameterized in terms of E and V, I am going to work instead on a model parameterized in terms of E/V and L. It turns out that this will be much more convenient for me in the long run, and standard DEB theory often works with the variable E/V, which is

defined as the "reserve density." I will define the new variable $W = E/V$. Applying the chain rule, the dynamics of $W = E/V$ are

$$\frac{dW}{dt} = \frac{d(E/V)}{dt} = \frac{1}{V(t)}\frac{dE}{dt} - \frac{E(t)}{V(t)^2}\frac{dV}{dt} \tag{13}$$

If you plug in the expressions for dE/dt and dV/dt, you end up with the dynamics of the variable W as

$$\frac{dW}{dt} = \rho\epsilon\mathcal{V}I_{max}\frac{F}{f_h + F}\frac{L^g}{V} - W\frac{v}{L}, \tag{14}$$

which is a very simple expression that actually eliminates a lot of the parameters already. Note that I have also made the simplifying assumption that "structural length" L and the observed length L_{obs} are actually the same.

To find the equation for the dynamics of length, we need a way to convert length (which is in mm) to structural volume V. A length-dry weight regression should be almost exactly converting length to structural volume, since the drying process would probably remove most of the weight that was tied up in reserve. As such, we have that $V = \xi L^q$, and

$$\frac{dL}{dt} = \frac{1}{q\xi L^{q-1}}\frac{dV}{dt}. \tag{15}$$

Simplifying, I have that

$$\frac{dL}{dt} = \frac{L}{q}\left(\frac{\frac{\kappa}{E_G}W\frac{v}{L} - k_m}{1 + \frac{\kappa}{E_G}W}\right). \tag{16}$$

We can use this length-weight regression to eliminate V from the equation for dW/dt as well:

$$\frac{dW}{dt} = \rho\epsilon\mathcal{V}I_{max}\frac{F}{f_h + F}\frac{L^g}{\xi L^q} - W\frac{v}{L}. \tag{17}$$

Note that equations 16 and 17 make sense together. The reserve density W increases through ingestion and is depleted through the mobilization of reserve to fuel growth and reproduction by an amount $W\frac{v}{L}$. This shows up in the numerator of dL/dt, weighted by the fractional allocation to growth κ and the cost of growth E_G. There is subtraction due to maintenance k_m and than a normalization $1 + \kappa W/E_G$ that comes from the reserve mobilization equation.

A simple examination of these two equations reveals right away that neither κ nor E_G by itself influences the dynamics of the system, only their combination κ/E_G. Since κ is a fraction between 0 and 1 and E_G is the conversion cost converting mgC of reserve into mgC of structure, this is a dimensionless quantity. The purpose of nondimensionalization is to find other dimensionless quantities that affect the system dynamics.

Essentially, I want to reparameterize the system, focusing on the dynamics of the dimensionless variables f, w, ℓ, r, and τ, rather than the dimensional variables F, W, L, R, and t. Thus, define $F = \hat{f}f$, $W = \hat{w}w$, $L = \hat{\ell}\ell$, $R = \hat{r}r$, and $t = \hat{\tau}\tau$, where the "hatted" variables are combinations of parameters that scale the dimensionless stage variables, allowing you to exactly recover the dimensional dynamics.

Let's start with the food dynamics, and work from there. The dimensional dynamics are

$$\frac{dF}{dt} = -I_{max}\frac{F}{f_h + F}L^g. \tag{18}$$

The nondimensional dynamics are

$$\frac{df}{d\tau} = \frac{\hat{\tau}}{\hat{f}}\frac{dF}{dt} = \frac{\hat{\tau}}{\hat{f}}\left(-I_{max}\frac{\hat{f}f}{f_h + \hat{f}f}\hat{\ell}^g\ell^g\right). \tag{19}$$

If I define $\hat{f} = f_h$, then I can simplify this equation to

$$\frac{df}{d\tau} = -\frac{\hat{\tau}I_{max}\hat{\ell}^g}{f_h}\frac{f}{1+f}\ell^g. \tag{20}$$

Given that the units of I_{max} are cells ml^{-1} mm^{-g} d^{-1} and the units of f_h are cells ml^{-1}, and that the units of $\hat{\ell}$ must be mm, if I define

$$\hat{\tau} = \frac{f_h}{I_{max}\hat{\ell}^g},$$

then $\hat{\tau}$ will have the proper units. This scaling means that we are measuring time in terms of ingestion, which makes sense given that ingestion sets all other processes. Furthermore, since f_h, I_{max} and g are all known based on the feeding data, specifying the scaling of length automatically specifies the scaling of time. This is important because, in order to fit the dimensionless model to data, I will have to estimate these scaling parameters. With $\hat{\tau}$ specified, the dynamics of dimensionless food become

$$\frac{df}{d\tau} = -\frac{f}{1+f}\ell^g. \tag{21}$$

Now let's consider the dynamics of dimensionless length:

$$\frac{d\ell}{d\tau} = \frac{\hat{\tau}}{\hat{\ell}}\left(\frac{\hat{\ell}\ell}{q}\left(\frac{\frac{\kappa}{E_G}\hat{w}w\frac{v}{\hat{\ell}\ell} - k_m}{1 + \frac{\kappa}{E_G}\hat{w}w}\right)\right)$$

If I define $\hat{w} = E_G/\kappa$ and $\hat{\ell} = v/k_m$ then this equation simplifies to

$$\frac{d\ell}{d\tau} = k_m\hat{\tau}\frac{\ell}{q}\left(\frac{\frac{w}{\ell} - 1}{1 + w}\right) \tag{22}$$

Note that the combination $k_m\hat{\tau}$ will be dimensionless, since k_m has units of 1/time and $\hat{\tau}$ has units of time. I will define the dimensionless parameter $\alpha = k_m\hat{\tau}$.

The dynamics of dimensionless reserve density are

$$\frac{dw}{d\tau} = \frac{\hat{\tau}}{\hat{w}}\rho\epsilon\mathcal{V}I_{max}\frac{\hat{f}f}{f_h + \hat{f}f}\frac{\hat{\ell}^g\ell^g}{\xi\hat{\ell}^q\ell^q} - \frac{\hat{\tau}}{\hat{w}}\hat{w}w\frac{v}{\hat{\ell}\ell} \tag{23}$$

Simplifying this equation, we have

$$\frac{dw}{d\tau} = \frac{\rho\kappa}{E_G}\frac{f_h\epsilon\mathcal{V}}{\xi\hat{\ell}^q}\frac{f}{1+f}\ell^{g-q} - k_m\hat{\tau}\frac{w}{\ell} \tag{24}$$

Note that the quantities $\phi = \frac{f_h \epsilon \mathcal{V}}{\xi \hat{\ell}^q}$ and $\sigma = \frac{\rho \kappa}{E_G}$ are both dimensionless. The dimensionless quantity $\alpha = k_m \hat{\tau}$ also appears in this equation.

Finally, we have the equation for egg production. Note that R is already dimensionless, as it is just a number. I will also replace the E in the mobilization equation p_C with $E = W\, V = W\, \xi\, L^q$, to focus on the variables we are actually interested in:

$$\frac{dR}{d\tau} = \hat{\tau}\frac{1-\kappa}{E_R}\hat{w}w\xi\hat{\ell}^q\ell^q \left(\frac{\frac{v}{\hat{\ell}\ell} + k_m}{1 + \frac{\kappa}{E_G}\hat{w}w} \right)$$

This can also be simplified to

$$\frac{dR}{d\tau} = \frac{E_G}{E_R}\frac{1-\kappa}{\kappa}\xi\hat{\ell}^q k_m \hat{\tau}\ell^q \left(\frac{\frac{1}{\ell}+1}{1+w} \right). \tag{25}$$

I will define the new parameter

$$\beta = \frac{E_G}{E_R}\frac{1-\kappa}{\kappa}\xi\hat{\ell}^q \tag{26}$$

which is dimensionless (because E_G is dimensionless, E_R has units of mgC, κ is dimensionless, ξ has units of mgC mm^{-q}, and $\hat{\ell}^q$ has units of mm^q).

Putting all of these equations together, I have the dimensionless system:

$$\frac{df}{d\tau} = -\frac{f}{1+f}\ell^g, \tag{27}$$

$$\frac{dw}{d\tau} = \sigma\phi\frac{f}{1+f}\ell^{g-q} - \alpha\frac{w}{\ell}, \tag{28}$$

$$\frac{d\ell}{d\tau} = \alpha\frac{\ell}{q}\left(\frac{\frac{w}{\ell}-1}{1+w} \right), \tag{29}$$

$$\frac{dR}{d\tau} = \beta\alpha\ell^q\left(\frac{\frac{w}{\ell}+w}{1+w} \right), \tag{30}$$

where the state variables are nondimensionalized using

$$\hat{\tau} = \frac{f_h}{I_{max}\hat{\ell}^g}, \tag{31}$$

$$\hat{w} = \frac{E_G}{\kappa}, \text{ and} \tag{32}$$

$$\hat{\ell} = \frac{v}{k_m}, \tag{33}$$

$$\tag{34}$$

and the dimensionless parameters are defined as:

$$\alpha = k_m \hat{\tau}, \tag{35}$$

$$\sigma = \frac{\rho \kappa}{E_G}, \tag{36}$$

$$\phi = \frac{f_h \epsilon \mathcal{V}}{\xi \hat{\ell}^q}, \tag{37}$$

$$\beta = \frac{E_G}{E_R} \frac{1 - \kappa}{\kappa} \xi \hat{\ell}^q. \tag{38}$$

If I were to fit this model to data, I would need to estimate $\hat{w}$, $\hat{\ell}$, α, σ, and β. $\hat{\tau}$ and ϕ are fixed by the estimate of $\hat{\ell}$ and the feeding model and the length-dry weight regression equation, since the parameters f_h, ϵ, $\mathcal{V}$, ξ, q, and g are all known. Moreover, if I assume that v is fixed, then the estimate of $\hat{\ell}$ specifies α so that it no longer needs to be estimated.

Whether this model would prove any more amenable to fitting is unclear, though I point out that many of the estimated parameters now involve combinations of parameters that are likely to trade off against one another, such as in $\hat{w}$, σ, and β.

Another option to address this challenge is to embrace what the data-fitting is saying. In particular, the data-fitting says that v has no influence on dynamics - I can set it to any value I want without affecting the other parameter estimates, almost at all. Given that, it makes sense to consider a reserve-less model, which essentially assumes that v is infinite and ingested food is immediately allocated for growth and reproduction. In this case, the dynamics of the system are simpler:

$$\frac{dF}{dt} = -I_{max} \frac{F}{f_h + F} L^g, \tag{39}$$

$$\frac{dV}{dt} = \left(\kappa \rho \epsilon \mathcal{V} I_{max} \frac{F}{f_h + F} L^g - k_m E_G V \right) / E_G, \tag{40}$$

$$\frac{dV}{dt} = \left((1 - \kappa) \rho \epsilon \mathcal{V} I_{max} \frac{F}{f_h + F} L^g \right) / E_R. \tag{41}$$

I can also do the same nondimensionalization trick to get a slightly simpler system (again moving from volume to length as a state variable):

$$\frac{df}{d\tau} = -\frac{f}{1 + f} \ell^g, \tag{42}$$

$$\frac{d\ell}{d\tau} = \frac{1}{q} \left(\sigma \phi \frac{f}{1 + f} \ell^{g-q+1} - \mu \ell^1 \right), \tag{43}$$

$$\frac{dr}{d\tau} = \beta \frac{f}{11 + f} \ell^g, \tag{44}$$

where

$$\hat{f} = f_h, \tag{45}$$

$\hat{\ell}$ is unspecified, and (46)

$$\hat{\tau} = \frac{f_h}{I_{max}\hat{\ell}^g}. \tag{47}$$

Note that the fact that $\hat{\ell}$ is unspecified is a little bit strange, but occurs because there aren't any parameters with units of length; there are only parameters with units of length to some power, such as I_{max} and ξ. The dimensionless parameters are

$$\sigma = \frac{\rho\kappa}{E_G}, \tag{48}$$

$$\phi = \frac{f_h\epsilon\mathcal{V}}{\xi\hat{\ell}^q}, \tag{49}$$

$$\mu = -k_m\hat{\tau}, \tag{50}$$

$$\beta = \frac{\rho(1-\kappa)f_h\epsilon\mathcal{V}}{E_R}. \tag{51}$$

Here, the things that need to be estimated by the fitting are $\hat{f}$, $\hat{\ell}$, σ, μ, and β. Both τ and ϕ are specified, once $\hat{\ell}$ is specified, as the feeding and length-dry weight regression parameters are all known.

Again, it is not totally clear to me whether this will be an easier estimation problem.