An Agent Based Model of the Diel Vertical Migration Patterns of *Mysis* diluviana

Nick Strayer[1], Brian O'Malley[2], Sture Hansson[3], Jason Stockwell[2]

[1]*College of Engineering and Mathematical Sciences, University of Vermont*

[2]*Rubenstein School for the Environment and Natural Resources, University of Vermont*

[1]*Stockholm University*

Abstract

Recent work indicates that the macro-invertebrate *Mysis diluviana* exhibits partial diel vertical migration (DVM), whereby one part of the population remains on the lake bottom at night while the other migrates up the water column. The drivers underlying the decision to migrate remain unknown. We developed an agent-based model that can simulate thousands of individual mysids decision-making processes at an hourly time step throughout a year. The model takes into account a daily and seasonally changing environment, including light, temperature, food availability across habitats and body condition. We found that the simulated *Mysis* population is highly sensitive to changes in the energy cost of performing migration. We have also developed a graphical user interface to help disseminate the results and testing of hypotheses without the need for the researcher to edit code. In addition to testing hypotheses about migration drivers, the model, once parameters have been calibrated with real data, will help facilitate more efficient field sampling and prediction of resource availability for mysivorous fishes by evaluating the potential for seasonality in *Mysis* migration patterns.

Keywords: *Mysis*, Modeling, Ecology, Agent-based, Simulation

1. Introduction

Mysis diluviana (*Mysis*) is a macroinvertebrate crustacean that lives in deep glacial lakes. They exhibit a behavior known as diel vertical migration in which they migrate from the waters at the bottom of the lake up into the

water column on a daily time scale and in doing so transfer nutrients between the two environments. Previous research of *Mysis* DVM has focused on particular aspects of their migrations. We access the sensitivity of migration to changes in different environmental parameters using Monte-Carlo-style simulation of many individuals at an hourly time step over a year.

1.1. Why Modeling?

More traditional methods of measuring the effects of environmental changes are costly and time consuming, requiring many hours of work and producing data with the inevitable noise associated with sampling in aquatic ecosystems. By modeling we are able to investigate effects at a much lower cost and gain a picture of the trends of effects.

1.2. Modeling is not perfect

It should be noted that the results and their specific units and/or amounts may not be exact, but what is important are the trends. Our objective is that the results from this modeling-based exploration into DVM will help us explore the potential for multiple stable migration strategies and also drive future real-world sampling efforts in a more efficient and impactful direction.

2. Methods

2.1. Programming Language

The programming language used for the model is R (R Core Team [2]). R was chosen due to its high adoption rate in the ecology community. This will allow for easier dissemination of the workings of the model and expansion by future researchers.

2.2. Agent-Based

An agent-based approach was chosen for the model as it allows for the interrogation of effects on the population and also the detection of multiple stable strategies while at the same time accounting for the inherent stochasticity of variables such as weather.

2.3. Expandability

The environmental factors fed into the agent-based main model are generated using their own sub-models. This design was chosen because it allows the perturbation of environmental factors by manipulating their generating model, but also for the ease in which real data can be substituted for the generated data. As data are gathered in the field they can be gradually integrated into the model to allow for more realistic measurements of effects and potentially in the future, predictions of *Mysis* locations.

2.4. Main Model

The model runs an individual through a year on an hourly time step. At each hour the individual draws random numbers from both uniform and normal probability distributions that then decide if the individual migrates and the amount of energy the gain in feeding.

2.4.1. Migration Decision

Every hour the individual checks to see if it is either one hour before sunset or the hour of sunrise. If it is an hour before sunset, a draw from a random uniform distribution from 0 to 1 is made. If the result of this draw is below the current food availability score ($FA(h)$) then another draw is made from the same distribution. If this second draw is below the condition curve for the individuals energy reserve ($PM(E)$) then the individual enters a state of "migrating". By doing this a food availability of 0.8 results in an 80% chance of a successful draw. If either of the two draws are *above* their respective curve the individual does not migrate.

If the current hour is sunrise than the individual enters a non-migrating state (if they didn't migrate prior to sunset they could already be in this state).

2.4.2. Feeding

The amount of energy an individual gains from feeding is drawn from a normal distribution with varying means and standard deviations depending on if the individual is currently migrating or in the benthic environment.

If the individual is migrating, they draw from a normal distribution with a mean of the migration reward (M_r) scaled by the seasonable food availability ($FA(h)$) with a standard deviation of the seasonal variance in food availability ($FV(h)$).

If the individual is not migrating then their energy reward is 20% of their reward if they had migrated. This value has a floor of 0.2 units to simulate the stability in food availability in the benthic environment.

2.4.3. Main Model

(a) Model Structure
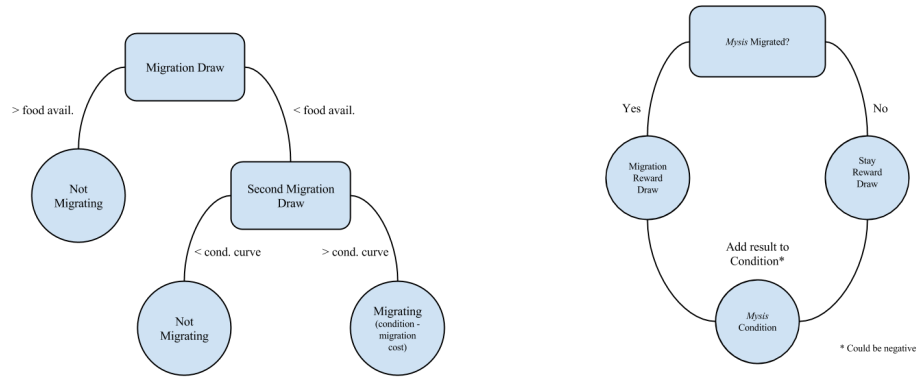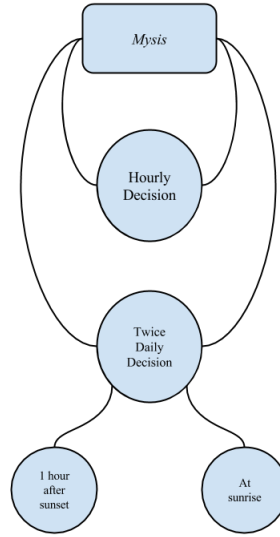
(b) Migration Decisionl   (c) Feeding

Figure 1: Graphical representation of the model structure.

2.5. Sub-Models

Sub-models were developed to generate and feed data on the environment into the main model. R scripts for the generation of all data are available at rpubs.com/nstrayer.

2.5.1. Thermocline Model:

We define the thermocline depth in the water column as the depth of $10\,°C$. This temperature was chosen as a *Mysis* threshold based upon the work of Jensen et al. [3]. To model thermocline depth $(T(h))$ at hour h over the year two sigmoidal curves with parameters differed to match seasonal trends in the environment (figure 2) were joined together.

$$\alpha = \begin{cases} k_{t1}(h - m_{t1}), & \text{if } h < 4380 \\ k_{t2}(h - m_{t2}), & \text{otherwise} \end{cases}$$

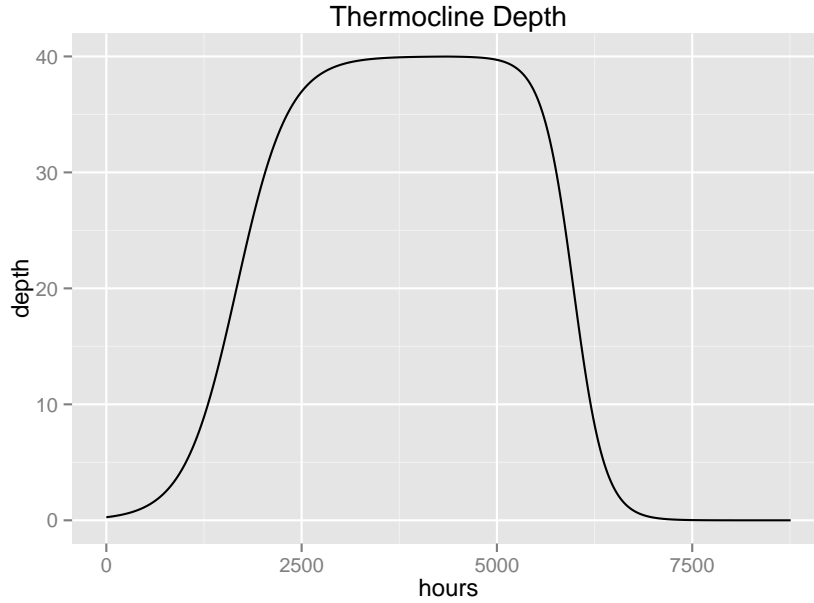$$\text{Th}(h) = \frac{M}{1 + e^{-\alpha}}$$

Figure 2: Modeled thermocline depth $(10\,°C)$ as a function of current hour.

2.5.2. Isocline Depth

Mysis are repelled by light. Boscarino et al. [4] found that *Mysis* did not exceed a threshold of 1×10^{-3} lux (1 lm/m^2). We modeled where in the water column this light level threshold was reached over the course of the entire year. To do this, data of light intensity levels in Burlington were obtained from the National Renewable Energy Laboratory (NREL [5]) and fed into the following models.

2.5.3. Moon Cycle

The raw data from the National Renewable Energy Laboratory were not sensitive enough to pick up the lunar cycles. To overcome this, nighttime light intensity levels were filled into with a model of the lunar cycle(Palmer and Johnsen [6]). The greatest light intensity value between the moon cycle model and real data was chosen for the final dataset.

$$\text{MC}(h_c) = 0.5 \Big[\cos \Big(\frac{1}{c} \cdot 2\pi \cdot h_c \Big) \Big] + 0.5$$

2.5.4. Beer's Law

Once a complete dataset was assembled for the year, the light intensity was run through an equation $(\mathrm{BL}(I_o))$ derived from Beer's Law (Hutchinson [7]). This equation takes in light intensity levels and returns the depth at which the *Mysis* light threshold is reached accounting for the dispersion of light through a particular medium. The extinction coefficient (k) used was one from Lake Superior (Jensen et al. [3]) as data were not available for Lake Champlain.

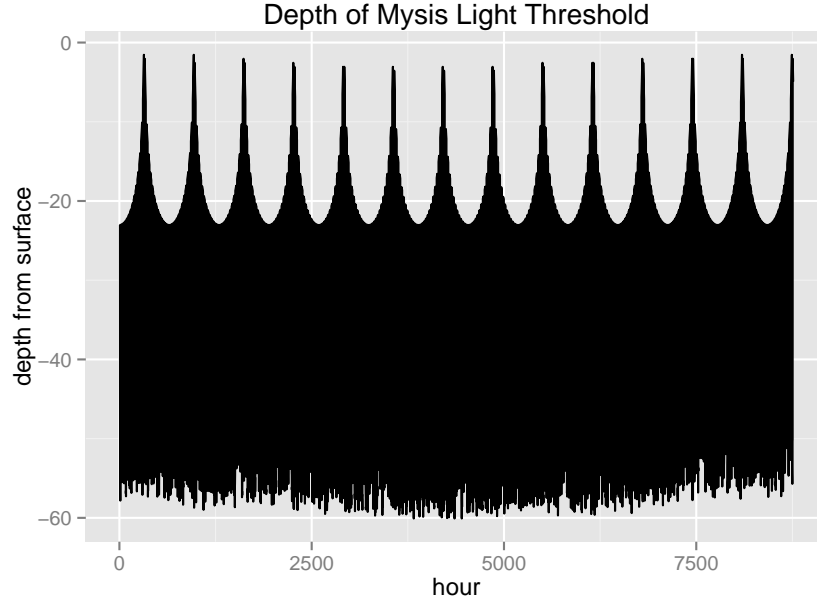$$\mathrm{BL}(I_o) = \frac{1}{k}\Big[\ln(I_o) - \ln(I_x)\Big]$$

Figure 3: Depth of *Mysis* light threshold 1×10^{-3} lux

7

2.5.5. Food Availability

The food availability ratio is a normalized measure of food quality and quantity in the pelagic (water column) to the benthic (bottom) environment. For example, a food availability score of 0.8 implies 80% of the food is available in the pelagic waters. We modeled pelagic food availability to mimic the increase in pelagic production during spring, peak in summer, and decrease in fall (figure 3). We used the following model for food availability (FA) for each hour (h) of the year:

$$\text{FA}(h) = s \cdot \cos\left[(1/8750) \cdot 2\pi \cdot (h - H) + \rho\right]$$
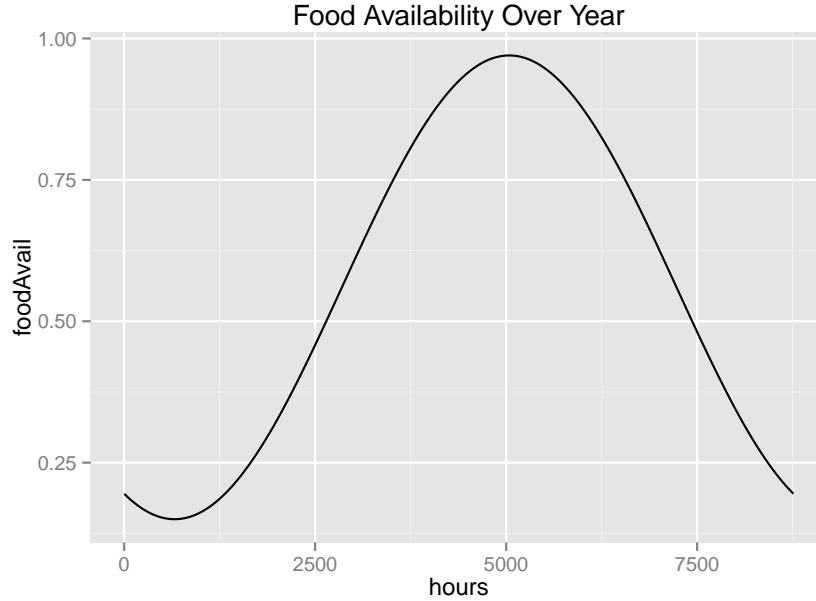

Figure 4: Food availability ratio over the course of a year.

2.5.6. Food Variability

To simulate the range of possible feeding conditions throughout the year we modeled variability in food availability as high in spring and fall and low summer and winter (figure 4). We doubled the frequencies of the food availability curve (figure 3) to capture this behavior. This value is used in the model to control the spread of the distribution upon which feeding rewards are drawn.

$$\text{FV}(h) = A \cdot \cos\left[(1/8750) \cdot 4\pi \cdot (h - H) + a\right]$$
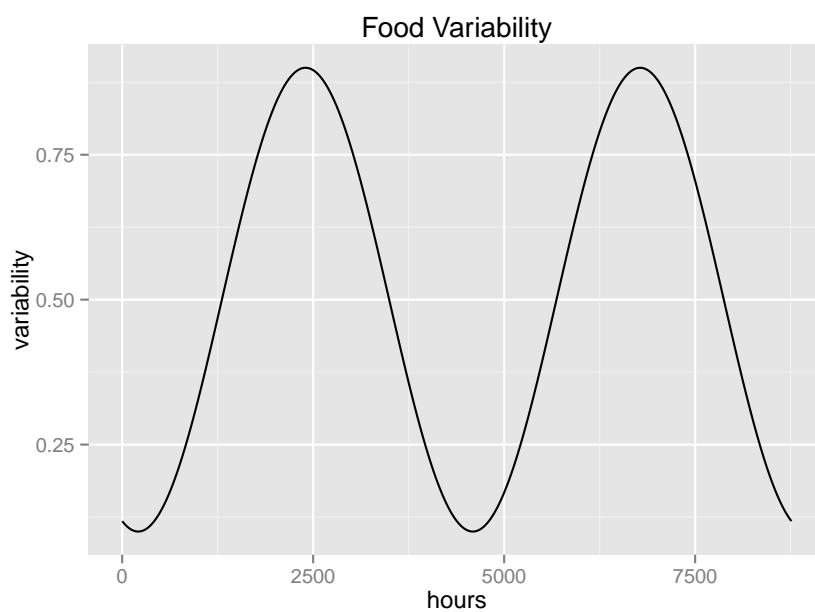

Figure 5: Variability in food availability over the course of a year.

2.5.7. Probability to Migrate

To simulate an individual's probability to migrate based upon body condition (E) after a successful draw to migrate based upon food availability, we constructed a model (PM(E)) that describes the relationship as sigmoidal. With this model a *Mysis* with a high body condition has less pressure to migrate than a *Mysis* with a low body condition (figure 6). The midpoint of the sigmoid sitting at what we decided was a 'normal' body condition, or 120 energy units.

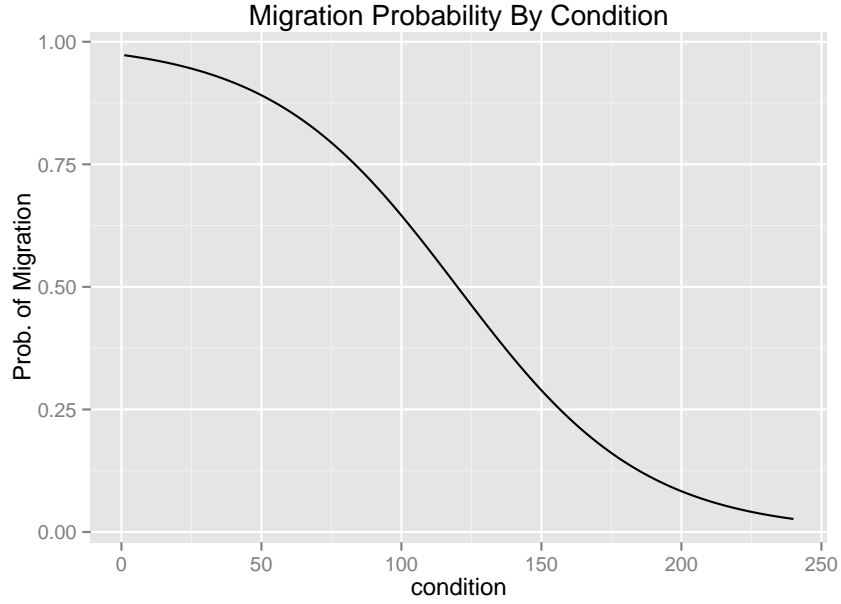$$\mathrm{PM}(E) = \frac{1}{1 + e^{-k_p(E - m_p)}}$$

Figure 6: Probability of migration by body condition E.

2.5.8. Feeding Reward

A draw from a normal distribution ($R(h)$) determines the number of energy units that an individual adds to its condition on a given hour (FR(h)). The distribution has a mean of r_h scaled by food availability at the hour (FA(h_i)) with a standard deviation of the food variability at that hour

(FV(h)). Note that this value can be negative to account for unproductive feeding efforts. If the individual is in a non-migrated state then they either take 20% of $R(h)$ or a 0.2 units, whichever is larger.

$$R(h) = N\Big(r_h \cdot [1 + FA(h)], \text{FV}(h)\Big)$$

$$\text{FR}(h) = \begin{cases} R(h), & \text{if } \textit{Mysis} \text{ migrating} \\ \max\Big[0.2 \cdot R(h), 0.2\Big], & \text{otherwise} \end{cases}$$

2.5.9. Testing of Agent-Based Model

To test the model's sensitivity to changes in variables we ran it through an entire year, simulating 250 mysids while perturbing the average hourly energy reward and migration cost. At the end of each simulation we recorded the proportion of surviving individuals with predation risk set to zero as to isolate the effects of the perturbing the variables.

2.6. Parameters

The following are the variables used throughout the model.

Variable	Value	Description	Units
s	1 for winter/spring, -1 for summer/fall	Seasonality	
M	40	Maximum Depth of $10\,°C$ in water column	Meters
K_{t1}	.003	Thermocline curve steepness (first half of year)	
K_{t2}	.005	Thermocline curve steepness (second half of year)	
K_p	.03	Probability to migrate curve steepness	
m_t	1667	Midpoint of thermocline curve	Hours
m_p	120	Midpoint of probability to migrate curve	Energy
h	1,2,...,8750	Hour of year	Hours
FA_{min}	0.15	Min value of food availability ratio	
FA_{max}	0.95	Max value of food availability ratio	
A	$(FA_{min} - FA_{max})/(2)$	Food availability curve scaler	
ρ	$(FA_{min} + FA_{max})/(2)$	Food availability curve scaler	
t	$\frac{1}{365 \cdot 24}$	Sinusoidal scaler	
H	5040	Time of peak food availability	Hours
c	$27.8 \cdot 24$	Moon cycle length	Hours
h_c	h modulous c	Position in time in the moon cycle	Hours
k	0.3	Extinction coefficient of lake (Jensen et al. [3])	m^{-1}
I_x	1×10^{-3}	*Mysis* Light threshold (Boscarino et al. [4])	Lux
I_o	input	Light level at lake surface	Lux
n	input	# of mysids	
r_h	input	Average hourly reward	Energy
m_c	input	Daily migration cost	Energy

2.7. Interactive Model Server

To help speed up analysis of variable perturbation an interactive R shiny server (Chang et al. [8]) was developed. This server allows the user through a graphical user interface to manipulate parameters of the model and in real time view the effects of those manipulations over the course of a year on the simulated individuals. In addition to serving as a quick way of testing effects, it also allows for the model to be investigated by users not comfortable enough with the code to manually change the variable values in the R script.
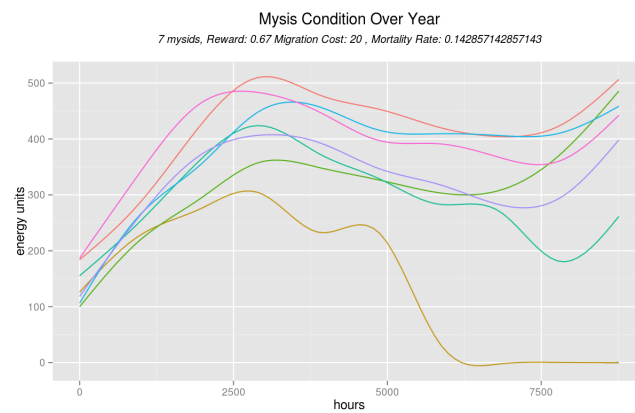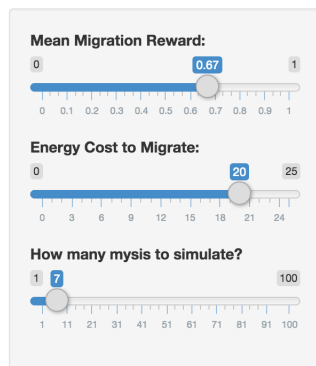
Figure 7: Screenshot of Shiny app.

3. Results

3.1. Mysocline

By merging the theoretical migration limit provided by where ten degrees celsius is in the water column with the location of the light threshold I_x we are able to get a macro-view of migration extent for mysids in Lake Champlain over the course of the year.
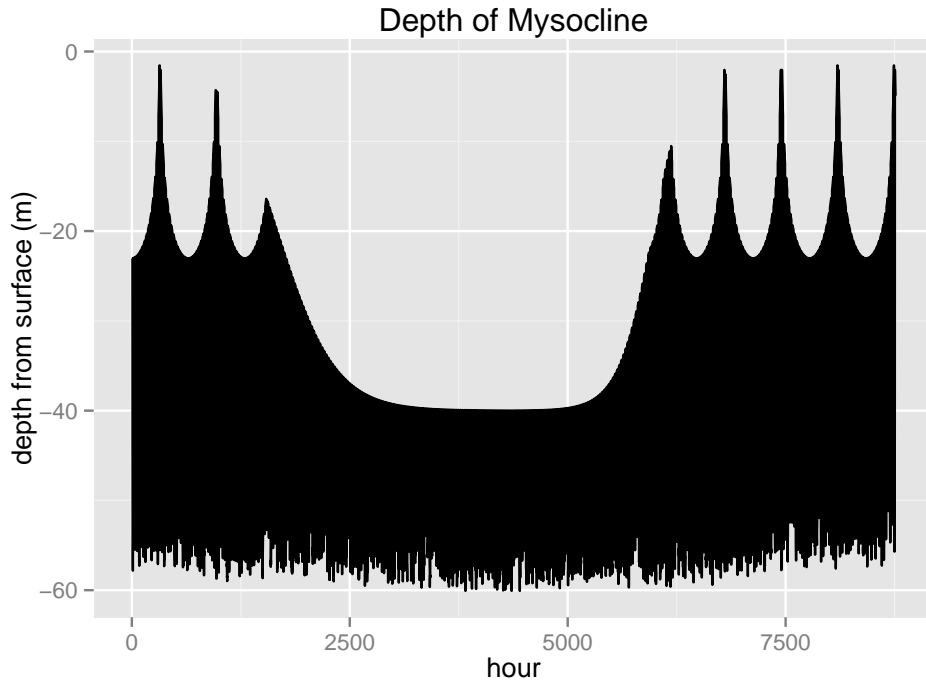

Figure 8: Mysis migration extent over the year taking into account light and temperature thresholds.

Results from the model suggest that during the winter the factor limiting *Mysis* migration is the light intensity levels, whereas in the summer and early fall it is water temperature.

3.2. Sensitivity to Rewards and Cost

Using the shiny app we located a range of average behavior and mortality rates for the parameters of average energy reward (r_h) and daily migration cost (m_c). We then set the model to run on a range of these values of 0.64

14

to 0.69 for r_h and 18 to 21.5 for m_c. At each combination of the variables we recorded mortality rates.

Condition Over Year

500 mysids, Reward: 0.66 , Migration Cost: 20 , Mortality Rate: 0.182
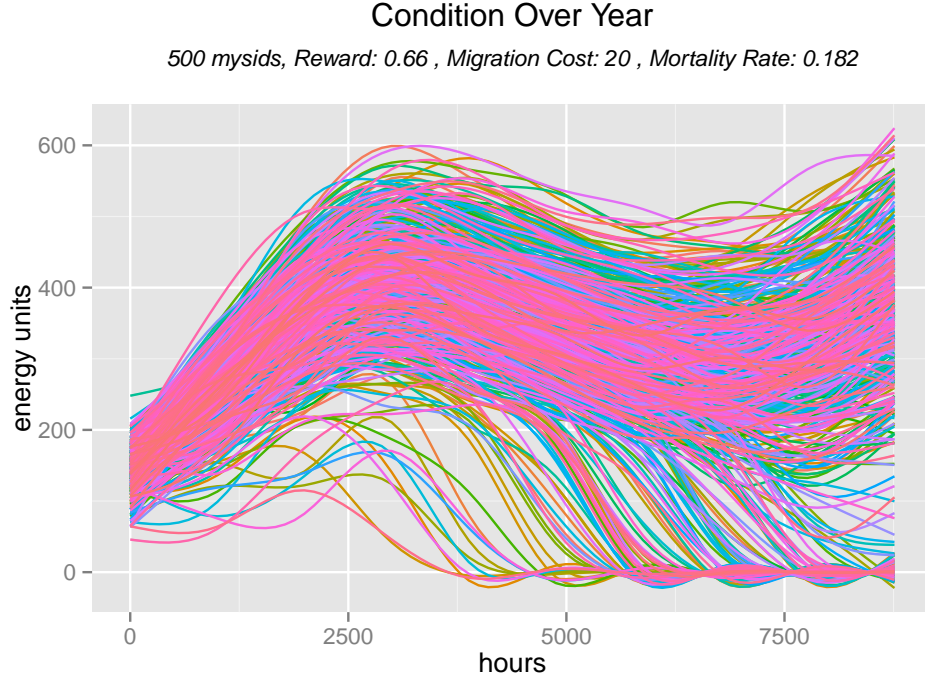

Figure 9: Condition of 500 Mysids over one year.

From this run over a year with $n = 500$ mysids we can see that most individuals fair very well in the first few months of the year, and then the whole population's condition starts to decline around late spring, continuing to decline into fall and then resuming a climb in the last couple months of the year. In the run seen in figure 9, with r_h of 0.66, and m_c of 20 has a mortality rate of 18.2%.

By repeating this simulation over the aforementioned ranges of r_h and m_c we can see the model's sensitivity to their changes.

From figure 10 we can see that if we fix the reward units and perturb the migration cost that significant mortality rates occur at every tested migration cost value, but not if we hold migration cost steady and perturb reward units. This result implies that further investigation into the energy cost of migration should be performed.
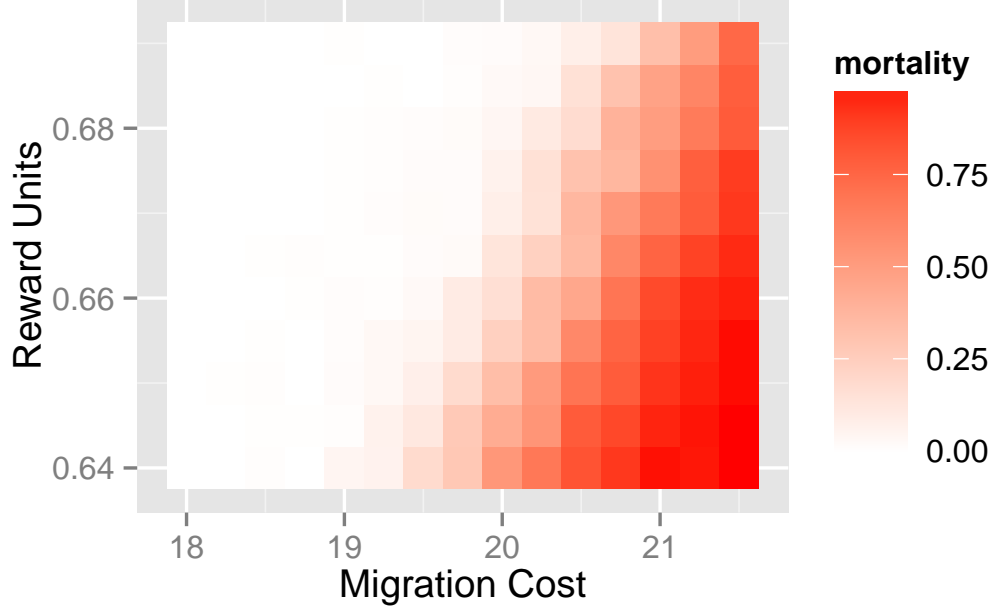
15

Figure 10: Heat map of mortality rates by changes in mean reward units and migration cost.

3.3. Multiple Migration Strategies

For each run the number of days that a given surviving individual chose to migrate was recorded and plotted. For a large majority of the runs the distribution of days migrated was normal, however, for some runs (figure 11) there appeared to be a slight bi-modal distribution. However, none of the results were significant enough to make any definitive statements about the presence of multiple stable strategies for migration behavior.

4. Conclusions and Future Directions

4.1. Environmental role

The model's response to variable perturbations hints at a high sensitivity to environmental changes. Further investigation by perturbing environmental parameters such as food availability and temperature are necessary to investigate this.
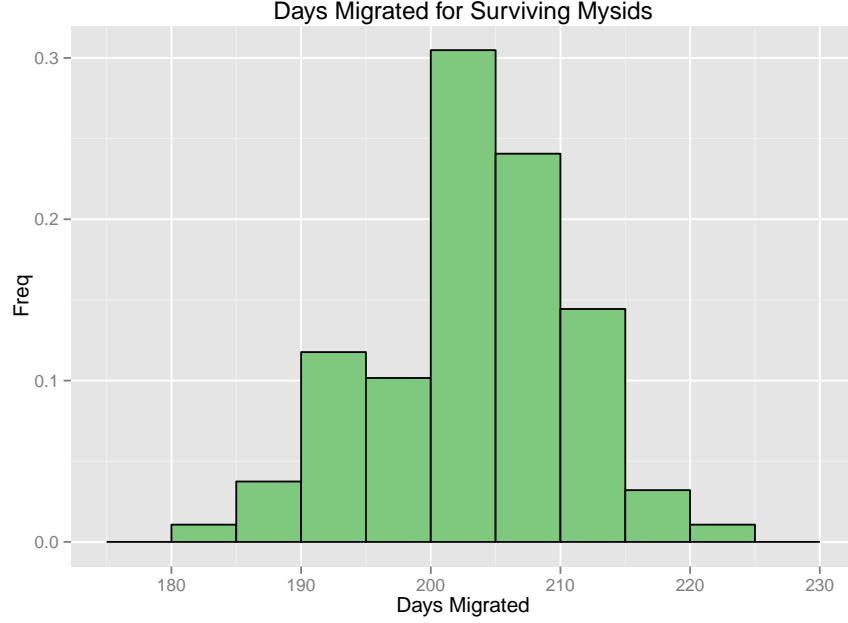
Figure 11: Number of days migrated by surviving individuals during a simulation.

4.1.1. Fall Condition Climb

For example, population level conditions were highly tied to the seasons with a large portion of the population getting close to starvation in the mid-summer months. If the recovery of body condition brought by fall cae later in the year there is a strong potential for a high impact on the survival rate of individuals.

4.1.2. Variable Extinction Coefficient

The link between algae blooms and environmental conditions have been well established (Isenstein et al. [9]). For the current model the clarity of the water is set at a constant k. Algae blooms obviously have a large impact on the dispersion of light through the water column. A higher k value would raise the *Mysis'* migration extent during the beginning and end of the year and increase the cost of migration, thus having large impacts on population levels. One way to investigate this would be to put a variable extinction coefficient into the model to simulate algae blooms and tie the migration cost to extent migrated to study the effects on the *Mysis* population.

4.1.3. Multiple Migration Strategies

In the results we saw hints at a bimodal distribution of migration strategies by looking at the number of days migrated for surviving individuals. The model in its current form does not have the ability to fully explore these patterns. Future efforts could give individuals "personalities" such as "risk prone" or "conservative" and explore the effects each personality has on the population's fitness to investigate the viability of differing strategies.

4.2. Role in the Lab

As previously mentioned the cost of sampling at a sufficient frequency to gain a clear picture of migration patterns is prohibitively expensive and time consuming. That being said, modeling, especially with so many assumptions is not an exact science. This model will never replace field sampling *Mysis* but it can have its own valuable role in the research process. As the model is tested and calibrated with real data its capacity to influence new sampling efforts and explain observed behaviors will expand. Simulation and real world sampling can help each other become more accurate and efficient and hopefully drive a new workflow in *Mysis* DVM research.

5. Acknowledgements

References

[1] J. M. Watkins, L. G. Rudstam, M. J. Connerton, T. Schaner, K. L. Rudstam, Per G. Bowen, Abundance and spatial distribution of mysis diluviana in lake ontario in 2008 estimated with 120 khz hydroacoustic surveys and net tows, AQUATIC ECOSYSTEM HEALTH & MANAGEMENT 18 (2015) 63–75.

[2] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2014.

[3] O. P. Jensen, L. Hurwicz, H. Uzawa, Diel vertical migration in the lake superior pelagic community. ii. modeling trade-offs at an intermediate trophic level, Naval Research Logistics Quarterly 8 (1961) 175–191.

[4] B. T. Boscarino, L. G. Rudstam, M. A. Minson, E. E. Freund, Laboratory-derived light and temperature preferences of juvenile mysid shrimp, Mysis diluviana, JOURNAL OF GREAT LAKES RESEARCH 36 (2010) 699–706.

[5] NREL, 2015.

[6] G. Palmer, S. Johnsen, Downwelling spectral irradiance during evening twilight as a function of the lunar phase, APPLIED OPTICS 54 (2015) B85–B92.

[7] G. Hutchinson, A treatise on limnology, volume 1, Geology, Wiley, 1957.

[8] W. Chang, J. Cheng, J. Allaire, Y. Xie, J. McPherson, shiny: Web Application Framework for R, 2015. R package version 0.11.1.

[9] E. M. Isenstein, A. Trescott, M.-H. Park, Multispectral Remote Sensing of Harmful Algal Blooms in Lake Champlain, USA, WATER ENVIRONMENT RESEARCH 86 (2014) 2271–2278.