CRITTERS: Climate, Resource, and Image Tracking in Tiny, Ecologically Representative Systems

Eleanor R. Stern[1], Saoirse Kelleher[1], Peter Vesk[1], Michael McCarthy[1]

School of Agriculture, Food, and Ecosystem Sciences, University of Melbourne, Parkville, Australia

Corresponding Author: ES (estern@student.unimelb.edu.au)

Abstract

1. Conservation and ecology research and practice is most effective when theory and models that underly species management are well explored and understood in experimental systems. Microcosm studies can provide experimental evidence to support theory, test model performance in different conditions, and suggest generality. However these benefits have been limited to primarily extinction and population studies, while habitat models have largely remained unexplored in newer systems.
2. There is presently lacking an experimental system that uses benchtop temperature and food manipulation combined with automated, in situ non-destructive sampling – necessary for conducting habitat and spatial structure studies in microcosms.
3. We developed Climate, Resource, and Image Tracking in Tiny, Ecologically Representative Systems (CRITTERS), an experimental microcosm system that manipulates habitat through abiotic temperature and biotic components (resources) in a benchtop design, and automates data collection through in situ observations. We propose it as a system in which to test a variety of models and theories in conservation and ecology.
4. We present the counting accuracy of CRITTERS and population responses of *Paramecium caudatum* to the temperature and resource availability. Using a modified version of the software *Wellcounter*, the correlation between automated and manual counts was r = 0.98. In logistic models, *P. caudatum* growth rate *r* responded to temperature quadratically, and carrying capacity *K* responded to resources (concentration of carbon in media) log-linearly, allowing independent control of *r* and *K*.
5. This system can be used to test a variety of habitat models in which climate and/or resources require manipulation when in situ sampling and high replication is important. We describe potential uses for testing adaptive management strategies, occupancy and N-mixture models, and Species Distribution Models.

Introduction

Microcosms are miniature experimental systems that are a powerful tool in ecology and conservation, with potential to test and improve a wide variety of models.

Microcosms have been used for over 100 years to provide experimental evidence for ecological theories, and the use of ciliate microcosms in particular as surrogates for larger, complex ecosystems is well established (Woodruff, 1908, 1911; Bonsall & Hassell, 2005; Holyoak & Lawler, 2005; Altermatt et al., 2015). In conservation biology, microcosm experiments are versatile, have management implications, and may be used for both ecosystem- or assemblage-specific hypotheses as well as broadly applied theories and models (Stern et al 2025). Furthermore, microcosms allow researchers to explore questions in critical topics such as biodiversity loss, extinction, climate change, and anthropogenic disturbance in a relatively low-risk and ethical manner compared to field experiments, which typically conduct experiments directly on focal ecosystems which may include rare or endangered species (Benton et al., 2007; Stern et al., 2025). Microcosm experiments can provide experimental backing for models, demonstrate their generalisability to many ecosystems or species, and provide insights into model strengths and limitations (Stern et al., 2025).

Microcosms are perhaps most usefully employed to complement both field and simulation studies rather than as substitutes. Microcosms provide answers from simplified systems absent of ecosystem-specific noise, allowing findings to potentially be widely used or applied, but they are limited in how complex and realistic they can be (Drake et al., 1996; Fraser & Keddy, 1997; Holyoak & Lawler, 2005; Benton et al., 2007; Altermatt et al., 2015; Stern et al., 2025). Mesocosms and field experiments, which are larger and more realistic can capture ecosystem-specific traits and dynamics in a controlled environment, but their findings may be difficult to generalise outside a given study system (Beyers & Odum, 1993; Diamond, 1983; Sagarin et al., 2016; Stewart et al., 2013). Finally, simulated studies are highly customisable and controllable and can test things impossible to explore in real systems, but they lack true natural responses and are inherently more deterministic (Peck, 2004; Petrovskii & Petrovskaya, 2012). These experiment types provide distinct, independent insights while complementing each other, and ideally should all be used to explore theories and test models. . However, the advantages of experimental manipulability, high replication and the natural variability of biological systems make microcosms very beneficial for model testing (Stern et al. 2025).

In ecology and conservation, statistical models which use data on population responses and habitat drivers are frequently used to support conservation efforts, plan protected areas, manage harvesting, and protect endangered species (Hayes et al., 1996; Guisan & Zimmermann, 2000; McCarthy & Possingham, 2007; McCarthy et al., 2011; Kéry et al., 2013; Guisan, 2017; Tucker & Runge, 2021; Ovaskainen, 2002). However, when models are inappropriate for the situation, poorly constructed, or ill-informed, conservation outcomes can be poor (Addison et al., 2013), emphasising the need to understand models and their best applications thoroughly. Many models (e.g. species distribution, occupancy models, etc.) have been tested in simulated or observational studies (Elith et al., 2006; Elith & Leathwick, 2009; Valavi et al., 2020) but lack experimental exploration, validation, and testing (Drew, 1994; J. A. Miller, 2014). Calls for increased experimental testing of conservation theory dates back at least to the 1990s (Drew, 1994).

Ecological and conservation microcosm experiments have been limited in scope and replicability by labour and the logistics of habitat manipulations. Past ciliate microcosm studies (e.g. Naeem & Li, 1997; Donahue et al., 2003; Fox et al., 2017) relied on manual counting where populations are grown in incubators (which characterises temperature habitat) and destructively sampled to estimate population densities rather than absolute population counts. The lack of benchtop habitat manipulation has prohibited experiments from easily exploring different temperature levels at the same time during the same experiment, instead repeating experiments in multiple parallel experiment systems. Some recent progress has included automation (Altermatt et al., 2015; Besson et al., 2022a; Stelzer & Groffman, 2025) and benchtop observational sampling, in which microcosms are kept on laboratory benches and populations counts are achieved by non-destructive, in-situ sampling through imaging (Besson et al., 2022a; Stelzer & Groffman, 2025), but differential benchtop temperature control has not been achieved. As such, many of these improvements have been restricted to population studies that focus on extinctions driven solely by food availability (Cerini et al., 2023) and ecotoxicology studies (Stelzer & Groffman, 2025), while distribution and habitat studies have remained limited (Table 1).

Benchtop habitat control would allow one to test more complex models and theories which rely on more than one facet of habitat (e.g. combining abiotic and biotic drivers). Benchtop temperature control can allow this variable to be integrated into a single experimental platform (unlike incubators which may require a split plot experimental design), increase habitat realism, reduce labour, and allow connectivity and spatial complexity to be more easily explored.Here, we present the new microcosm system Climate, Resource, and Image Tracking in Tiny Ecologically Representative Systems (CRITTERS), designed to test ecology and conservation models and theories through benchtop, automated direct observation of experimental microcosms that can simultaneously control temperature and food to impose habitat quality differences on microorganisms.

Table 1. Comparison of automated microcosm systems and their current uses.

Automated System, Location, and Citation	**Automated Sampling**	**Automated Counting**	**Benchtop Sampling**	**Benchtop Temperature Manipulation**	**Current Uses**
Unnamed System, University of Bristol, (Besson et al., 2022a)	✓	✓	✓		Extinction models, Population models
Wellcounter, Innsbruck	✓	✓	✓		Ecotoxicology

University, (Stelzer & Groffman, 2025)					
Unnamed System, Swiss Federal Institute of Aquatic Science and Technology, (Fronhofer et al., 2015)	✓	✓			Population models, Density dependence, range expansions (food gradient)
CRITTERS, University of Melbourne, *preprint link*	✓	✓	✓	✓	Habitat models

Methods

Overview and Rationale

This experimental system is designed to replicate closed habitat patches that are large relative to species body size, where both temperature and resources (abiotic and biotic habitat, respectively) are controlled. We factorially cross temperature and food levels during experiments to produce a variety of habitat qualities. Microcosm wells (hereafter “wells”) are constructed from acrylic and held in a frame on top of heat mats (Figure 1). *Paramecium caudatum* are grown in liquid media made up of Chalkey’s solution (filtered water with a standardized amount of salts added, the salts being necessary for osmoregulation of *P. caudutum)*, ground alfalfa, and bacteria.

Our system uses top-down imaging of microcosms, conducted using a video camera atop a stereomicroscope that is mounted on a robot moving on a 2-dimensional frame (hereafter referred to as the “gantry”). Videos are then processed in a modified version of Wellcounter (Stelzer & Groffman, 2025) to count *P. caudatum* in each video.

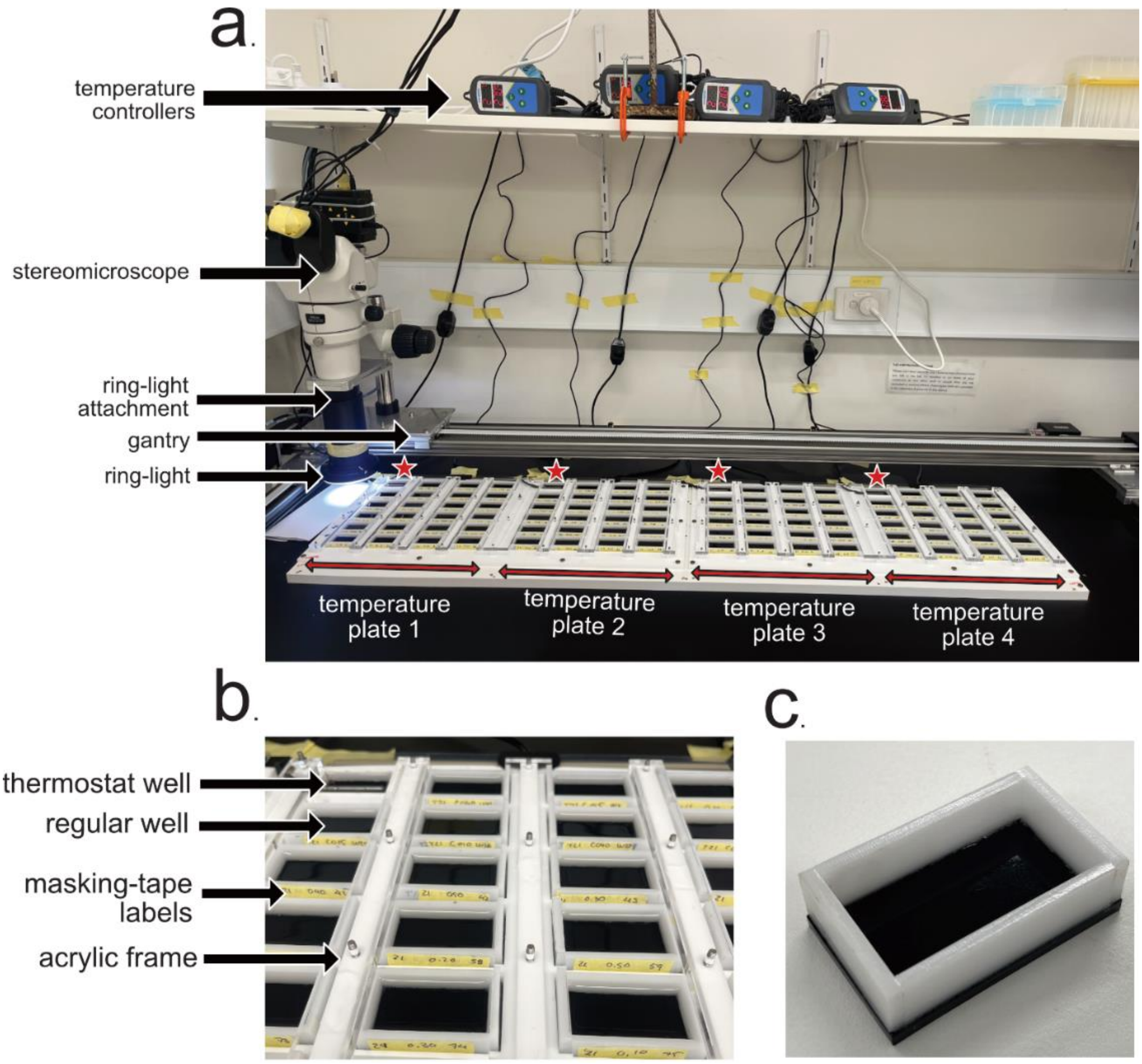

Figure 1. A) overview of the experimental system depicting the microcosms, temperature plates (red arrows), temperature controllers, thermostat wells (red stars), robotic gantry, ring light configuration, and stereomicroscope; B) microcosms in acrylic frame with media, *Paramecium* stock, and mineral oil inside, as in an experiment, C) top view of a microcosm;

Study Species

Although other similar organisms could also be used, the experimental system as described here uses *Paramecium caudatum,* a globally occurring freshwater ciliate protozoan, as its model species. *P. caudatum* measure between 170-290 µm, and are elongated and ovoid in shape, and typically reproduce asexually with a generation time of around 1 day (Wichterman, 1986). *P. caudatum* are covered in hairlike cilia that allow them to be highly active and motile (Wichterman, 1986). *Paramecium* have a long history of use in ecological microcosms and are well established as proxies for larger species (Woodruff, 1908, 1909, 1911; Woodruff & Baitsell, 1911; Gause, 1933, 1934; Gause et al., 1936; Gill, 1972; Luckinbill, 1973; Maly, 1978). *Paramecium* are a bacterial grazer (Wichterman, 1986), which allows them to take the functional role of a non-predatory animal in microcosms. Their population growth primarily depends on environmental conditions such as food, temperature, pH, waste products, oxygen and aeration, and light, and the relationship between these factors and population growth is relatively well described and understood (Wichterman, 1986).

Microcosm Configuration and Construction

During experiments, *P. caudatum* populations inhabit custom-built microcosms designed to be waterproof while optimising detection in top-down imaging. Microcosms are rectangular open boxes made of 3 mm acrylic sheet, with sides laser cut and joined with acrylic adhesive. Bases are shiny black acrylic and walls are translucent white acrylic. Note that translucent white is necessary to reduce the glare issues that occur in transparent acrylic or plastic when imaged from above. For complete construction instructions, see Appendix. Our microcosms have an internal floor area of 26 x 44 mm, 24 wells are on each temperature plate, and the acrylic frame holds 4 temperature plates (Figure 1A, 1B).

Climate manipulation through temperature

The climate is controlled through the temperature of media within which *P. caudatum* are grown during experiments. Temperature is known to drive population growth rate of *P. caudatum* (Krenek et al., 2011). *P. caudatum* tend to have a thermal preference of around 24 C , a lower critical limit of around 3-5 C and an upper critical limit of around 32 - 35 C, depending on the strain of *P. caudatum* (Krenek et al., 2011).

We control temperature using flat heat mats (Lerway reptile heat mats, 20 W) placed underneath the acrylic frame. This allows manipulation of the temperature of liquid medium in wells. Each heat mat has separate temperature control (Inkbird ITC-306T-WIFI) with a thermostat probe placed in a dedicated well on each mat (top leftmost well on each mat). Heating of the mats, and thus the microcosms, is regulated by this thermostat well.

To describe the variability of temperature in the CRITTERS system, we recorded in-well variability of the thermostat well (recordings automatically taken every 5 minutes by the probe) and additionally recorded the oil surface temperature of each well multiple times across 4 experiments using a hand-held infrared thermometer.

Resource manipulation through food source

Resources are controlled through the bacterial food source for *P. caudatum*. While bacterial concentrations are difficult to directly manipulate, they may be indirectly controlled through organic matter in media, which drives bacterial populations. We control this through the concentration of alfalfa brewed in media. Our experiments have test food levels between 0.1 and 0.5 g alfalfa / L Chalkey's solution, which presents a range of food concentrations from very poor food resources to abundant food resources. For detailed instructions on preparing media, see Appendix.

Experiment Duration and Protocol

During experiments, *P. caudatum* are grown in alfalfa broth (described above) inside the microcosms. In our system, a combined total of 5 mL media made up of alfalfa broth and *Paramecium* inoculation stock is added to microcosms. Then 3 mL of mineral oil is gently dispersed on top of the medium to prevent evaporation during the experiments, while not impairing visibility during imaging (Figure 1B). The ratio of media and stock may be determined by desired inoculation density; we generally inoculate wells with around 20 individuals which is approximately 0.2 – 0.4 mL of stock into 4.6 – 4.8 mL of media. Experimental durations were restricted to less than 3 weeks (~ 30 generations for *P. caudatum*) because mineral oil restricts oxygen diffusing into the medium, which will eventually lead to extinction in the microcosms. We typically conduct experiments for up to 14 days (~ 14 generations for *Paramecium*). We have not found that the mineral oil poses a complication for *Paramecium* populations up to this duration. The population density of *Paramecium* in wells covered by mineral oil is about 20% lower than when grown in bottles in incubators. Otherwise, populations in wells under the oil follow expected trajectories of growth and then stability at carrying capacity. Note, due to heating of the aquatic medium, it is not possible to cover the wells with lids because condensation would impair imaging.

In-Situ Sampling using Automated Video Recording

Our sampling process is similar to other automated high-throughput microcosm designs developed recently (Besson et al., 2022a; Stelzer & Groffman, 2025). We also perform non-destructive, in-situ counting of the abundance of *Paramecium* in microcosms using an automated recording system (Figure 1A).

We use a stereomicroscope mounted on the gantry, set to 1x magnification, atop a c-mount set to 0.55x (a sufficient magnification to clearly see *P. caudatum* when projected to a computer monitor), with video pixel resolution of 5440 x 3060. This is approximately 1 million pixels per square centimetre, or about 20 pixels over the length of an individual *P. caudatum*. This microscope has a video camera attached, with a ring light attached to the microscope lens. A custom, 3D printed, attachment holds the ring light approximately 0.5 cm above the surface level of the microcosms using a 3D printed extension piece (Figure 1A). Reducing the distance between the microcosm surface and the ring light minimises glare effects on the mineral oil. For details on the gantry, its construction, and its operation, see Appendix.

Sampling is done by recording 15 second videos at 4-hour intervals (6 recordings per day, roughly equivalent to 6 observations per generation).

Automated Abundance Estimates

To process videos recorded by the CRITTERS system we adapted the *'Wellcounter'* software for automated detection of microinvertebrates developed by Stelzer & Groffman (2025) (Stelzer & Groffman, 2025). The original open-source software is written in Python and uses a deterministic image-processing approach, which identifies organisms by isolating areas of movement in videos and applying geometric thresholds to determine if a shape is comparable to a microorganism.

Our CRITTERS system differs from the system in which the *Wellcounter* software was developed, in that we use rectangular rather than circular wells and are more prone to contamination by dust particles and bacterial clumps due to the open tops (rather than closed petri dishes as in Wellcounter), resulting in a high false-positive detection rate from *Wellcounter* even after optimisation. To mitigate these challenges, we made substantial modifications throughout the *Wellcounter* analysis pipeline, resulting in our *Wellcounter-Modified* software summarised in Table 2. For expanded explanations of Wellcounter-Modified, see Appendix. Our changes significantly improve detection accuracy in open-top wells.

Table 2: *Wellcounter-Modified* workflow compared against the original *Wellcounter.*

Step	***Wellcounter***	***Wellcounter-Modified***	**Reason for change**
Convert to greyscale (all frames)	Convert to grayscale	Same as *Wellcounter.*	No difference.
Detect and mask non-well areas (all frames)	Apply a circular mask based on known well dimensions to isolate the regular well area.	Detect the largest contiguous area of high-brightness pixels (the well walls) and apply as a mask to isolate the irregular well area.	*Wellcounter* used circular wells, CRITTERS custom wells are rectangular but may have rough edges.
Identify movement channels (all frames)	Not included.	For each frame of the video, calculate difference between pixel brightness and minimum value to identify areas of potential movement, then detect contours of these areas and filter to large contiguous areas that correspond with 'tracks' of moving *Paramecium. (Figure 2A)*	Isolating areas of continuous movements helps to distinguish slow-moving dust particles from active *Paramecia*.
Detect and mask bubbles	Not included.	Detect bright circular areas in the image and mask as	Air bubbles were a frequent source of

(all frames)		likely air-bubbles.(Figure 2A)	false positive detections.
Apply background subtraction (repeat for 3 frames)	For each of three frames in the video, find the difference in pixel values from the other two to identify areas of potential movement.	For each of three frames in the video, find the difference in pixel values from the median value to identify areas of potential movement. These areas are then masked by the movement channel mask.	Applying the movement channel mask reduces false positives, and using the median value rather than values from the other frames reduces stochasticity.
Convert greyscale to binary image (repeat for 3 frames)	Apply a fixed brightness threshold to the background-subtracted image.	Apply a dynamic threshold to the background-subtracted image, where the threshold is automatically determined by Otsu's binarization,	Different lighting conditions make a fixed threshold inconsistent.
Detect *Paramecium* (repeat for 3 frames)	Used automated contour detection to identify contiguous areas of positive pixels as possible *Paramecium*. (Figure 2B)	Same as *Wellcounter.* (Figure 2B)	No change.
Filter contours (repeat for 3 frames)	Filter contours of possible *Paramecium* by an area threshold.	Filter contours of possible *Paramecium* by area, eccentricity, solidity, aspect ratio, perimeter, ferret radius, and axis ratio.	Additional shape parameters help reduce false positives.
Compare counts	The three counts, the maximum count, and the mean count are returned.	Same as *Wellcounter*	No change.

The primary output from both *Wellcounter* and *Wellcounter-Modified* has three estimates of *Paramecium* abundance from each video. We take the mean of these estimates as the population estimate for a video. This process is automated and performed in batches over large quantities of videos. Because our experiments record over 6000 videos per experiment, we use the High Performance Computing system Spartan at the University of Melbourne to run these modified Wellcounter scripts.

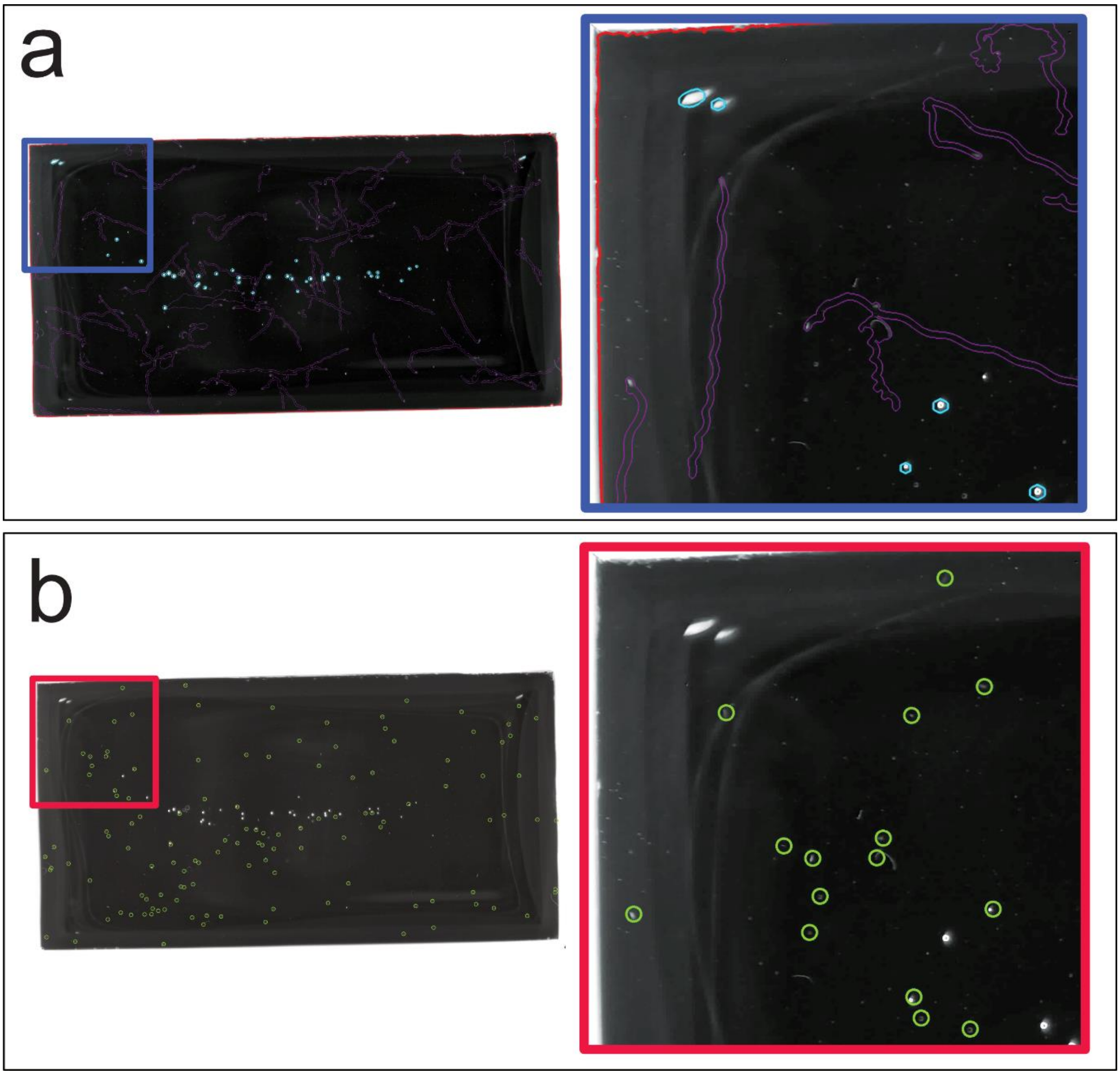

Figure 2. A) Additional masks added which identify *P. caudatum* movement channels (purple polygons), bubbles or lighting aberrations (blue circles), and microcosm walls (red polygons). All of these masks are unique to *Wellcounter-Modified*, with *Paramecium* detection limited to the purple polygons. B) *P. caudatum* identified in the same well (green circles, also present in the original Wellcounter code) within movement channels, which are unique to *Wellcounter-Modified*.

Validation

We optimised *Wellcounter-Modified* detection settings (e.g. particle size, obliqueness, blur, etc.) using a sample set of videos from trial-runs of experiments. After optimisation, accuracy of *Wellcounter-Modified* was validated on 5 independent sets of data (Experiments 0-4) to assess performance:

i. 16 videos taken of 16 microcosm wells (1 video per well) in the laboratory for the purpose of validation only, run for less than 1 day (Experiment 0),

ii. 24 videos from two experiments with 12 wells in each run for 14 days: Experiment 1 & Experiment 2 (one video was removed from Experiment 2 because it was recorded during a software malfunction). Videos from Experiment 1 & Experiment 2 were selected randomly, stratified by time in experiment (0-4 days, 5-9 days, 10-14 days) and the predicted count (< 10, 10 – 50, 50 - 200, and >200), sampling 1 video from of these crossed groups.

iii. 24 videos from two experiments with 12 wells in each run for 14 days: Experiment 3 & Experiment 4. This validation aimed to assess if error was biased in time. 4 sequential videos starting at Day 10, when populations were at equilibrium, were chosen from 3 wells sampled randomly for each experiment. The 3 wells were sampled from groups stratified by predicted count (< 10, 10 – 50, 50 – 200).

Each of these 63 videos was validated manually by a single observer. Videos had grids imposed and observed *Paramecium* were counted in each grid cell (112 cells per video, which will fit a maximum of about 10 *Paramecium* in a very dense well). For detailed counting protocol, see Appendix. Counts were summed to produce a manual population count per video. This was compared with the *Wellcounter-Modified* estimate for the same video. The 16 videos in Experiment 0 were independently validated by two observers to determine consistency between observers ($R^2 = 0.9996$, mean difference = -0.44 paramecia per well), indicating that one observer was sufficient for the remaining videos in Experiments 1-4.

Population modelling

We used all available data in Experiment 1 (87 total wells for 14 days) to model the population response to the temperature and resource manipulations. Analyses were conducted in R version 4.4.3 (R Core Team, 2022) and visualised using the package ggplot (Wickham, 2016). Population responses in individual wells were modelled using a logistic growth model:

$$N(t) = \frac{K}{1 + \left(\frac{K - N0}{N0}\right) e^{-rt}}$$

Where $N(t)$ is the population at a given timestep t in the experiment (in our experiments, a timestep = 4 hour intervals), K is carrying capacity (number of individuals), $N0$ is the starting population (number of individuals when $t = 0$, i.e. the inoculation population), and r is the per capita growth rate.

Parameters r and K were estimated for each well in experiment SDM1 (n = 87 microcosms) by fitting the population growth function for $N(t)$ versus time using Bayesian non-linear regressions in the BRMS package in R (Bürkner, 2017). The prior distribution for r was uniform between 0 and 10, the prior distribution for K was uniform between 0 and 1000, and the prior distribution for $N0$ was normal with a mean of 20 and standard distribution of 1 (wells were inoculated with an estimated 20 individuals). Some populations did not grow, especially those in unfavourable conditions. In these cases, estimates of *r* were very uncertain because the low initial population sizes were very close to the carrying capacity *K* (*r* controls that rate of population growth toward *K*). When the posterior distribution of *r* had a standard deviation greater than 0.3, the data point was excluded from further analysis. For the remaining parameter estimates, weighted linear regressions were then performed to determine how temperature and alfalfa concentration influenced the estimated values of *r* and *K*. The weights were equal to the inverse of the variance. Thus, regressions for *r* were weighted by:

$$\frac{1}{sd^2}$$

where *sd* is the standard deviation of the posterior distribution of *r.*

The regression model for *K* was based on ln(*K*), so to estimate the the appropriate weight, we assumed that uncertainty in *K* had a log-normal distribution, in which case ln(K) would have a normal distribution with variance equal to ln(cv^2+1), where *cv* is the coefficient of variation of the posterior distribution of *K*. Thus, regressions for ln(*K*) were weighted by:

$$\frac{1}{\ln{(cv^2+1)}}$$

We explored models to explain patterns in *r* and *K* as a function of concentration and/or temperature, including additive and interactive models, as well as linear and quadratic relationships. The full set of tested models to explain variation in *r & K,* separately includes:

1. *r* OR ln(*K)* ~ concentration
2. *r* OR ln(*K)* ~ temperature + temperature^2
3. *r* OR ln(*K)* ~ temperature + temperature^2 + concentration
4. *r* OR ln(*K)* ~ (temperature + temperature^2) * concentration

Results

Validation

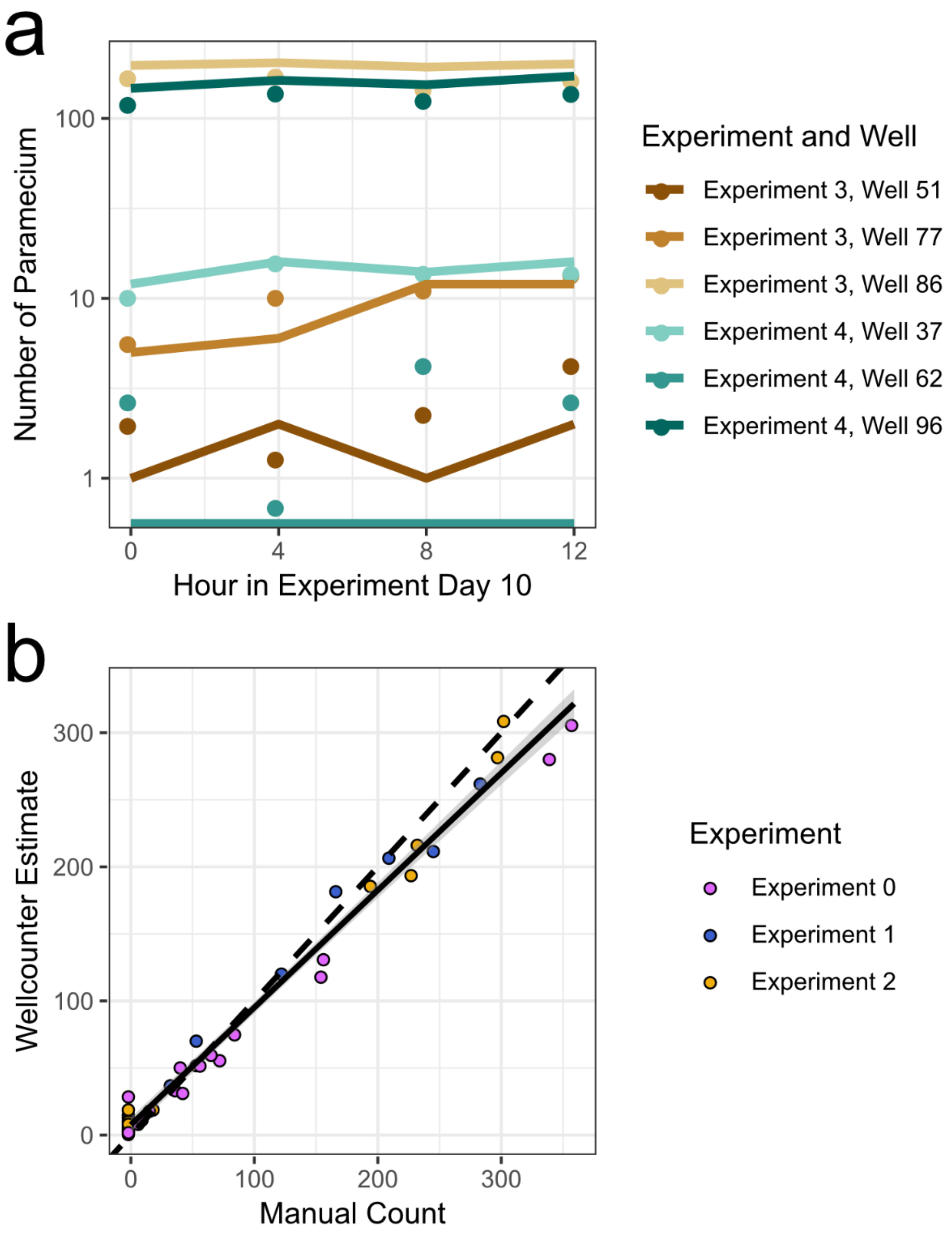

Figure 3. a) Wellcounter-Modified estimates (points) compared to manual counts (lines) over 4 sequential recording intervals for 6 wells (3 each from Experiment 3 & Experiment 4, stratified by *Paramecium* abundance) to assess autocorrelation of errors ($\log_e$ scaled counts); b) comparison of manual counts and Wellcounter-Modified estimate across all validation sets (points) with linear model (solid black line) and a 1:1 line (dashed black line) for comparison, where a 1:1 relationship would indicate perfect correspondence between manual counts and Wellcounter-Modified estimates.

Paramecium were highly detectable using the CRITTERS workflow. Wellcounter-Modified estimates were strongly correlated with manual counts (adjusted $R^2 = 0.98$, $n = 63$ across 5 experiments, linear regression slope = 0.87, intercept = 7.6) (Figure 3B). Wellcounter-Modified occasionally predicted a small number of *Paramecium* when none were present, and it tended to slightly undercount when *Paramecium* were abundant (Figure 3A).

Temperature variation

Temperatures within wells are very consistent (±1°C). Across a single heat mat, temperature may deviate by up to 5°C from the thermostat well, especially on mats kept at warmer temperatures (e.g. 24°C and above) (Figure 4). Colder mats closer to the temperature of the room or walk-in incubator (eg. 18 C and 21C) had less variation across mats. When there was significant variation within a mat, heat generally decreased in a radial pattern from the centre of the mat. Experiments including connectivity could use this variation to generate a habitat gradient within a single mat, allowing organisms to disperse or distribute according to habitat preference. For studies not considering spatial complexity, because well temperatures are autocorrelated based on location, other treatments (e.g. broth concentration) should be distributed randomly using Latin squares or similar randomisation method.

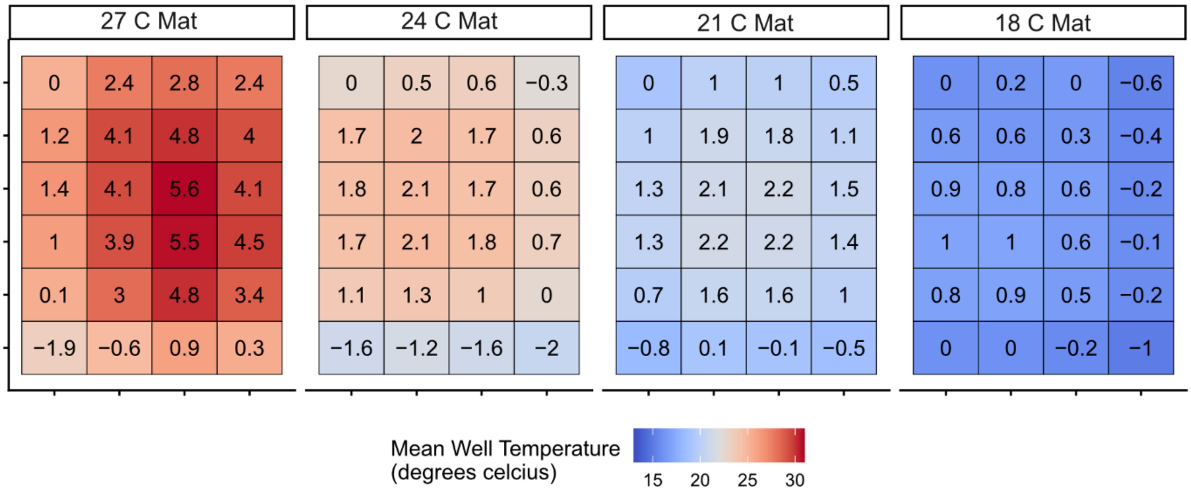

Fig. 4. Mean temperature of median inside wells (colour of tiles) and deviation from the temperature of a dedicated thermostat well (black labels) over the duration of 4 experiments.

Population responses to temperature and resources

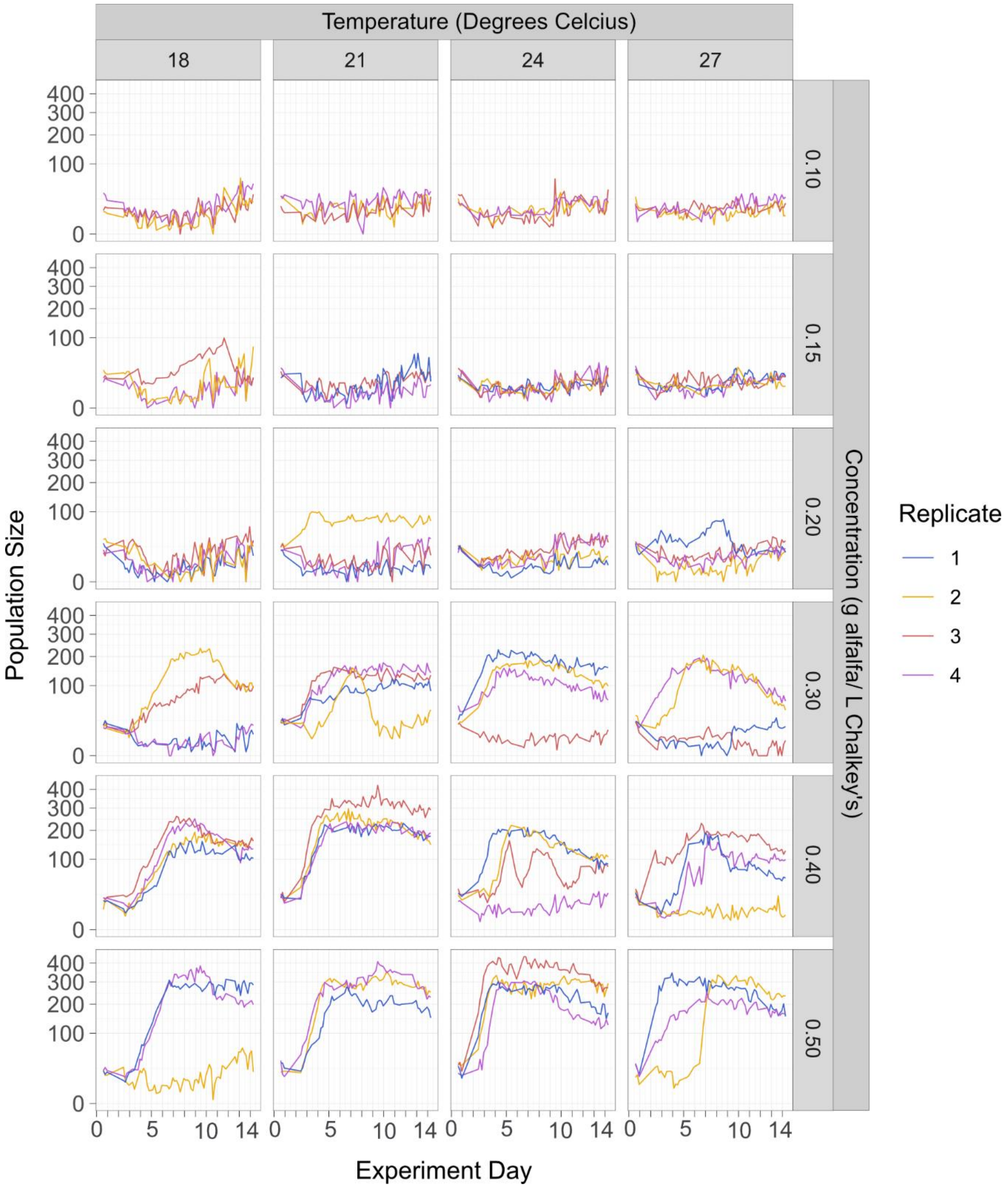

Figure 5. Population trajectories for each well in Experiment 1, grouped by temperature plate (columns) and broth concentration (rows)

Temperature and food concentration influenced the population trajectories. There was some variation among replicates and stochastically through time, with differences in trajectories being occasionally quite large (Figure 5).

Resources (broth concentration) primarily drove the carrying capacity *K* in microcosms (Figure 6A). It had little effect on the growth rate *r* in microcosms. K had a positive linear relationship with the concentration of media (Table 3). The relationship between K and broth concentration can be described by the equation produced from a GLM (Table 1, Figure 6A):

$$Ln(K) = 2.23 + 7.2C$$

where K is carrying capacity and C is the concentration of alfalfa broth in g alfalfa per L Chalkey's solution.

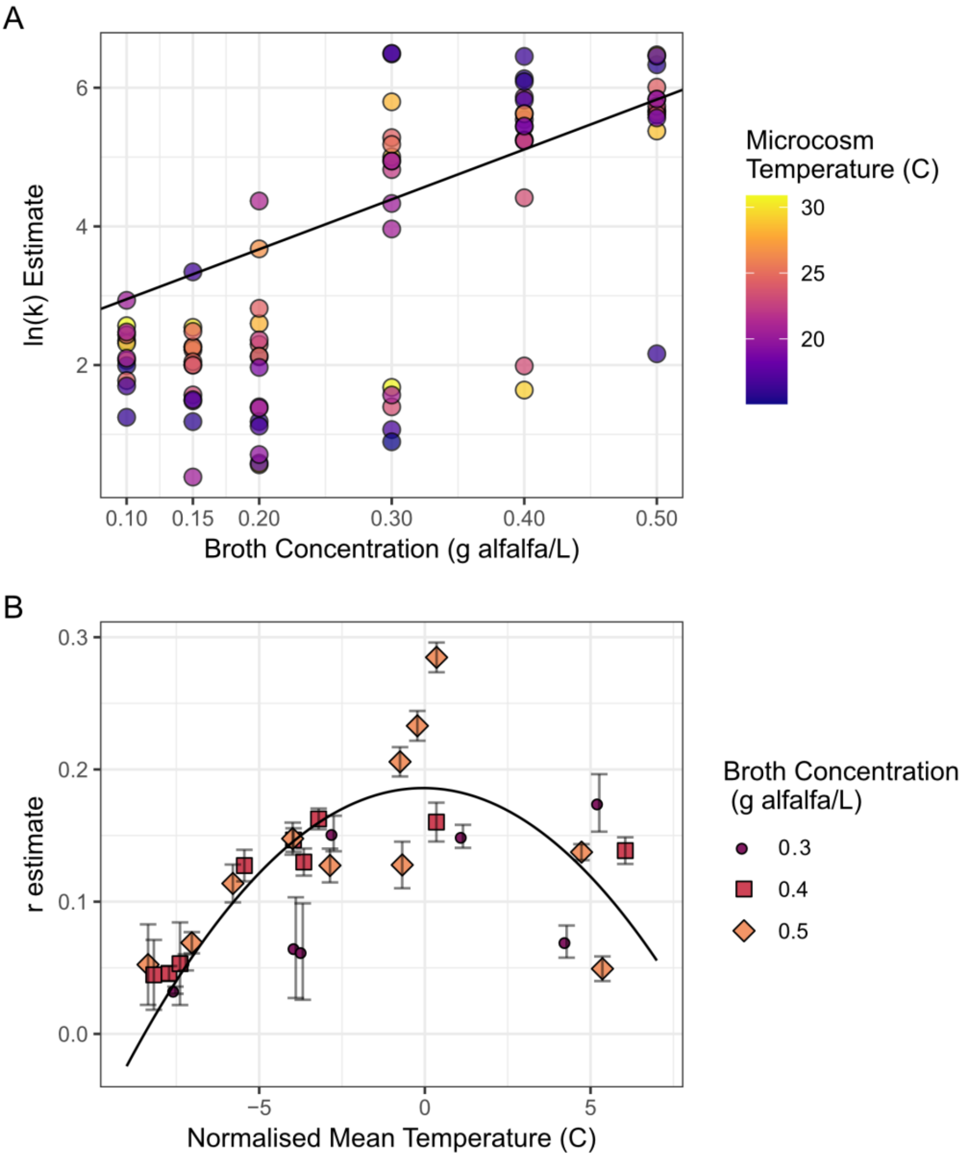

Figure 6. A) Estimate of the natural logarithm of carrying capacity ln(*K*) as a function of broth concentration B) estimate of growth rate *r* as a function of normalised mean temperature of a microcosm.

The mean temperature in a microcosm over the course of experiments drove the growth rate *r* (Figure 5B), while having little effect on the carrying capacity *K* of microcosms. Growth rate *r* had a quadratic response to temperature. The relationship between r and microcosm temperature can be described by the equation produced from the GLM (Table 3, Figure 5B):

$$r = 0186 + 2.54 * 10^{-4}(T - O) - 2.63 * 10^{-3}(T - O)^2$$

where *r* is growth rate and *T* is the average temperature of a microcosm, standardised by subtracting *O*, the approximate optimal temperature (24C for our populations).

Table 3. Response of growth rate *r (*n = 30 microcosms*)* to temperature (where temperature is normalised by subtracting the approximate optimum) and carrying capacity *K* (n = 87 microcosms) to broth concentration (resources) in *Paramecium caudatum* microcosms, derived from weighted linear regressions.

Model	**Term**	**Estimate**	**Std. Error**	**p value**
ln(K) ~ broth concentration	Intercept	2.23	0.24	<0.001
	Broth Concentration	7.21	0.57	<0.001
r ~ (temp – O) + (temp – O)^2	Intercept	0.186	1.11 x 10^{-2}	<0.001
	Temperature	-2.54x 10^{-4}	1.66 x 10^{-3}	0.879
	Temperature^2	-2.63 x 10^{-3}	43.49 x 10^{-4}	<0.001

Discussion

The CRITTERS system presented here represents an advance in replicable, realistic, and efficient benchtop experiments for developing and testing ecological models. By leveraging existing robotics and computer vision tools while introducing benchtop temperature manipulation CRITTERS allows direct and confident counts of populations to be achieved in ecologically realistic microcosm systems with relative ease. A key improvement of CRITTERS is ourbenchtop temperature and resource manipulations which improve realism relative to previous incubator-based experimental designs, permitting an expanded variety of models and theories to be tested. With the exception of experiment setup and validation, the system is entirely automated, increasing the potential for replication compared to the historically labour-intensive data collection of non-automated and/or non-benchtop systems. Similar experiments may be repeated many times, potentially in different locations by different research groups, to increase confidence in results and facilitate model testing over short time-scales at minimal cost relative to field surveys. It has been noted that more experimentation is necessary to complement the plethora of observational research in conservation (Drew, 1994), and CRITTERS is well situated to support experiments developing or testing models or theories relating to species occurrence and environmental factors.

While the simplification in microcosms means that they cannot fully replace field studies, the advantage of microcosm experiments is their simplicity and generalisability. Much as statistical models themselves simplify complex, real-world systems for the purpose of inference and prediction, “generalised microcosms” (Stern et al 2025) such as CRITTERS simplify natural systems for the same purposes. By eliminating ecosystem-specific complexities and confounders, focal phenomena can be isolated to improve mechanistic understandings of systems (Drake et al., 1996; Fraser & Keddy, 1997; Benton et al., 2007; Altermatt et al., 2015). CRITTERS can test models in simplified, generalised systems, while improving on realism from some other microcosm experiments through benchtop temperature control. Environmental drivers can act as a continuous variable (Figure 4) rather than discrete - as in homogenous incubators -and more complex manipulations are also possible such as connectivity across temperature gradients, cycling, stochastic disturbances or changes in environmental conditions (e.g. temperature) over time. Our populations showed some variation in responses that would not be present in deterministic simulations and would be difficult to foreshadow and parameterise in stochastic simulations (Figure 4A). *Paramecium* populations have natural stochasticity that provides randomness that mimics larger organisms – some populations fail even in good conditions, and populations in poor conditions may persist.

While significant work has gone towards counting micro-organisms automatically using computer vision (Stelzer & Groffman, 2025) or machine learning (Besson et al., 2022a), previous deterministic computer-vision based methods operate on largely sterile microcosms with no competing particles (e.g. Wellcounter (Stelzer & Groffman, 2025)). Other alternatives for open-top wells require computationally-taxing machine learning and classification to eliminate false positives (e.g. CommTrack (Besson et al., 2022a)), requiring additional training to adapt to new conditions and datasets. Our modifications to Wellcounter allow for automated counting of micro-organisms in non-sterile, open-top microcosms, necessary in lidded heated systems due to condensation and an inability to image from below if heating the wells. These modifications use deterministic image manipulation techniques rather than machine-learning based classification methods, avoiding an onerous increase in required computing power. Automated detection of *P. caudatum* with *Wellcounter-Modified* proved functionally equivalent to the accuracy of manual counting (Figure 3B).

Potential Uses of CRITTERS

The advances of CRITTERS create opportunities to test a wide array of models in ecology and conservation with microcosm experiments. Here, we discuss how adaptive management strategies, N-mixture and occupancy models, and species distribution models (SDMs) may be studied using CRITTERS. Many of these models are difficult to evaluate with field studies or field experiments, because comprehensive field evaluation is precluded by the required levels of experimental control, replication, and monitoring accuracy. .

A compelling example of these opportunities is adaptive management (Holling, 1978), which is specifically designed for cases where experimental manipulation is limited and the ability to manipulate habitat drivers independently may be necessary. Adaptive management aims to “learn by doing” while explicitly acknowledging limited options for in-field experimentation, meaning that full knowledge of the managed system is unattainable. Attempts to optimise adaptive management account for this uncertainty (e.g. McCarthy & Possingham, 2007)) but the limitation on experimentation in adaptive management systems precludes experimental evaluation of such optimisation with field studies (Holden & Ellner, 2016). Our pilot experiments with CRITTERS found that the carrying capacity of the microcosms (K) is largely driven by food availability, while the rate of population change (r) is largely driven by temperature (Fig. 6). These fundamental properties can be independently manipulated in CRITTERS to permit evaluation of adaptive management – with separate control of both r and K, high replication, and accurate monitoring of abundance, optimisation of adaptive management in biological systems can be feasibly tested beyond simulation studies. Wells may act as discrete units in which nominally optimal management actions can be tested, with population outcomes easily and accurately assessed. Testable ‘management’ interventions in CRITTERS could include patch size (i.e., volume of media), inoculation density relative to carrying capacity, habitat quality via temperature and food, culling of individuals from the population, management of r versus K, survey effort, genetic strain of inoculated populations, etc. Management scenarios that can be evaluated include re-introduction and captive management, eradication and population management more generally, optimal monitoring, and reserve design.

Interest in models for species occurrence or abundance that account for imperfect detection has increased recently. Examples of such hierarchical models like occupancy-detection or N-mixture models are difficult to test in natural experimental systems because the true underlying state (abundance or occupancy) is generally not directly observed. As the modifications to *Wellcounter* enable high-accuracy counts of organisms in open-top wells, CRITTERS can accurately observe these ‘latent’ states to provide a source of truth and enable evaluation. Because the locations of individuals within wells is recorded, subsampling of the populations within the wells is straight-forward, facilitating the incorporation of imperfect (and potentially biased) detection, which is fundamental to occupancy and N-mixture models. While occupancy models have been tested through observational or field data (D. A. W. Miller et al., 2015; Briscoe et al., 2021; Valente et al., 2024), these tests have limited precision of detection and monitoring, levels of replication, and experimental control. As such, it has been difficult to experimentally assess performance across a range of conditions or with violations of particular assumptions. Given the increasingly popularity of these models and a proliferation of extensions (Kelleher et al., 2024), the rapid and efficient testing of these assumptions with CRITTERS can improve confidence in novel methods.

A third example of CRITTERS' potential relates to testing Species Distribution Models (SDMs). This class of models is widely used in ecology and conservation to both predict species abundances and infer habitat relationships to support management and conservation efforts (Elith & Leathwick, 2009; Guisan, 2017). While these models are well tested in simulated studies and have been assessed observationally (Briscoe et al., 2021, 2021; Elith et al., 2006), the performance of SDMs have not yet been tested in natural experimental systems such as microcosms. To test the predictive performance of SDMs experimentally, one needs accurate population response data to multiple environmental drivers that may be controlled independently to produce a variety of habitat conditions. Prior to benchtop temperature manipulation, testing these was prohibitively difficult to achieve in incubators. CRITTERS makes it easy to explore habitat gradients through temperature and food in benchtop systems. Potential experiments to address knowledge gaps may include testing effects of imperfect detection, and assessing predictive performance in real-world experimental systems based on data quantity and quality, whether populations are at equilibrium or are establishing during data collection, and how model performance changes when conditions are not at equilibrium (e.g. climate change).

Conclusions

Stern et al. (2025) distinguished between specialised microcosms that aimed to replicate specific communities or ecosystems, and generalised microcosms that were designed to be simplified analogues of many different systems. CRITTERS, a generalised microcosm system, is best used on models with broad applicability, where fundamental validity or performance needs to be assessed. Importantly, CRITTERS in its simplicity has potential to be widely deployed to test models in separate experimental systems at different researching institutions in collaborative efforts. This opens up potential for "between study replication", which is uncommon in ecology and conservation (Filazzola & Cahill, 2021). For researchers interested in testing models or theories in CRITTERS or setting up their own version, instructions on lab setup and protocol notes are available through the Appendix. We also encourage reaching out to the CRITTERS team directly.

Acknowledgements

This research was funded by ARC DP230102663 to MM and a Research Training Program scholarship to ES. This research was additionally supported by The University of Melbourne's Research Computing Services and the Petascale Campus Initiative. We would like to additionally thank Erik S. Alonso Brust for his work maintaining the lab and assisting with experiments, Max Pearse contributions to microcosm well construction design, David P. Wilkinson for help and writing code to establish the workflow in Spartan, and Ainslie E. Macdonald and Richard A. Stern for comments and providing perspective on the draft.

Appendix

1) response data

2) experiment protocol notes

3) construction and setup information

References

Addison, P. F. E., Rumpff, L., Bau, S. S., Carey, J. M., Yung En Chee, Jarrad, F. C., McBride, M. F., & Burgman, Mark A. (2013). Practical solutions for making models indispensable in conservation decision-making. *Diversity and Distributions*, *19*, 490–502.

Altermatt, F., Fronhofer, E. A., Garnier, A., Giometto, A., Hammes, F., Klecka, J., Legrand, D., Mächler, E., Massie, T. M., Pennekamp, F., Plebani, M., Pontarp, M., Schtickzelle, N., Thuillier, V., & Petchey, O. L. (2015). Big answers from small worlds: A user's guide for protist microcosms as a model system in ecology and evolution. *Methods in Ecology and Evolution*, *6*(2), 218–231. https://doi.org/10.1111/2041-210X.12312

Benton, T. G., Solan, M., Travis, J. M. J., & Sait, S. M. (2007). Microcosm experiments can inform global ecological problems. *Trends in Ecology & Evolution*, *22*(10), Article 10. https://doi.org/10.1016/j.tree.2007.08.003

Besson, M., Alison, J., Bjerge, K., Gorochowski, T. E., Høye, T. T., Jucker, T., Mann, H. M. R., & Clements, C. F. (2022a). Towards the fully automated monitoring of ecological communities. *Ecology Letters*, *25*(12), 2753–2775. https://doi.org/10.1111/ele.14123

Besson, M., Alison, J., Bjerge, K., Gorochowski, T. E., Høye, T. T., Jucker, T., Mann, H. M. R., & Clements, C. F. (2022b). Towards the fully automated monitoring of ecological communities. *Ecology Letters*, *25*(12), 2753–2775. https://doi.org/10.1111/ele.14123

Beyers, R., J., & Odum, H., T. (1993). *Ecological Microcosms*. Springer-Verlag.

Bonsall, M. B., & Hassell, M. P. (2005). Understanding Ecological Concepts: The Role of Laboratory Systems. In *Advances in Ecological Research* (Vol. 37, pp. 1–36). Elsevier. https://doi.org/10.1016/S0065-2504(04)37001-7

Briscoe, N. J., Zurell, D., Elith, J., König, C., Fandos, G., Malchow, A., Kéry, M., Schmid, H., & Guillera-Arroita, G. (2021). Can dynamic occupancy models improve predictions of species' range dynamics? A test using Swiss birds. *Global Change Biology*, *27*(18), 4269–4282. https://doi.org/10.1111/gcb.15723

Bürkner, P. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80(1)*, 1–28.

Cerini, F., Jackson, J., Childs, D. Z., & Clements, C. F. (2023). *Multivariate signals of population collapse in a high-throughput ecological experiment.* https://doi.org/10.22541/au.170052311.18245610/v1

Diamond, J. M. (1983). Ecology: Laboratory, field and natural experiments. *Nature*, *304*(5927), 586–587. https://doi.org/10.1038/304586a0

Donahue, M. J., Holyoak, M., & Feng, C. (2003). Patterns of Dispersal and Dynamics among Habitat Patches Varying in Quality. *The American Naturalist*, *162*(3), 302–317. https://doi.org/10.1086/377185

Drake, J. A., Huxel, G. R., & Hewitt, C. L. (1996). Microcosms as Models for Generating and Testing Community Theory. *Ecology*, *77*(3), 670–677. https://doi.org/10.2307/2265489

Drew, G. S. (1994). The Scientific Method Revisited. *Conservation Biology*, *8*(2), 596–597.

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., … E. Zimmermann, N. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Filazzola, A., & Cahill, J. F. (2021). Replication in field ecology: Identifying challenges and proposing solutions. *Methods in Ecology and Evolution*, *12*(10), 1780–1792. https://doi.org/10.1111/2041-210X.13657

Fox, J. W., Vasseur, D., Cotroneo, M., Guan, L., & Simon, F. (2017). Population extinctions can increase metapopulation persistence. *Nature Ecology & Evolution*, *1*(9), 1271–1278. https://doi.org/10.1038/s41559-017-0271-y

Fraser, L. H., & Keddy, P. (1997). The role of experimental microcosms in ecological research. *Trends in Ecology & Evolution*, *12*(12), 478–481.

Fronhofer, E. A., Kropf, T., & Altermatt, F. (2015). Density-dependent movement and the consequences of the Allee effect in the model organism Tetrahymen. *Journal of Animal Ecology*, *84*, 712–722. https://doi.org/10.1111/1365-2656.12315

Gause, G. F. (1933). *The Struggle for Existence*. Dover Publications.

Gause, G. F. (1934). Experimental Analysis of Vito Volterra's Mathematical Theory of the Struggle for Existence. *Science, New Series*, *79*(2036), 16–17.

Gause, G. F., Smaragdova, N. P., & Witt, A. A. (1936). Further Studies of Interaction between Predators and Prey. *The Journal of Animal Ecology*, *5*(1), 1. https://doi.org/10.2307/1087

Gill, D. E. (1972). Density Dependence and Population Regulation in Laboratory Cultures of Paramecium. *Ecology*, *53*(4), 701–708. https://doi.org/10.2307/1934786

Guisan, A. (2017). *Habitat Suitability and Distribution Models With Applications in R*. Cambridge University Press.

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2–3), 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Hauser, C. E., Giljohann, K. M., McCarthy, M. A., Garrard, G. E., Robinson, A. P., Williams, N. S. G., & Moore, J. L. (2022). A field experiment characterizing variable detection rates during plant surveys. *Conservation Biology*, *36*(3), e13888. https://doi.org/10.1111/cobi.13888

Hayes, D. B., Ferreri, C. P., & Taylor, W. W. (1996). *Linking fish habitat to their population dynamics*. *53*.

Holden, M. H., & Ellner, S. P. (2016). Human judgment vs. Quantitative models for the management of ecological resources. *Ecological Applications*, *26*(5), 1553–1565. https://doi.org/10.1890/15-1295

Holling, C. S. (1978). *Adaptive Environmental Assessment and Management*.

Holyoak, M., & Lawler, S. P. (2005). The Contribution of Laboratory Experiments on Protists to Understanding Population and Metapopulation Dynamics. In *Advances in Ecological Research* (Vol. 37, pp. 245–271). Elsevier. https://doi.org/10.1016/S0065-2504(04)37008-X

Kelleher, S., Guillera-Arroita, G., Elith, J., & Briscoe, N. (2024). *Twenty years of dynamic occupancy models: A review of applications and look to the future*. Biodiversity. https://doi.org/10.32942/X28S5X

Kéry, M., Guillera-Arroita, G., & Lahoz-Monfort, J. J. (2013). Analysing and mapping species range dynamics using occupancy models. *Journal of Biogeography*, *40*(8), 1463–1474. https://doi.org/10.1111/jbi.12087

Krenek, S., Berendonk, T. U., & Petzoldt, T. (2011). Thermal performance curves of Paramecium caudatum: A model selection approach. *European Journal of Protistology*, *47*(2), 124–137. https://doi.org/10.1016/j.ejop.2010.12.001

Luckinbill, L. S. (1973). Coexistence in Laboratory Populations of Paramecium Aurelia and Its Predator Didinium Nasutum. *Ecology*, *54*(6), 1320–1327. https://doi.org/10.2307/1934194

Maly, E. J. (1978). Stability of the Interaction Between Didinium and Paramecium: Effects of Dispersal and Predator Time Lag. *Ecology*, *59*(4), 733–741. https://doi.org/10.2307/1938777

McCarthy, M. A., & Possingham, H. P. (2007). Active Adaptive Management for Conservation. *Conservation Biology*, *21*(4), 956–963. https://doi.org/10.1111/j.1523-1739.2007.00677.x

McCarthy, M. A., Thompson, C. J., Moore, A. L., & Possingham, H. P. (2011). Designing nature reserves in the face of uncertainty: Reserve design and uncertainty. *Ecology Letters*, *14*(5), 470–475. https://doi.org/10.1111/j.1461-0248.2011.01608.x

Miller, D. A. W., Bailey, L. L., Grant, E. H. C., McClintock, B. T., Weir, L. A., & Simons, T. R. (2015). Performance of species occurrence estimators when basic assumptions are not met: A test using field data where true occupancy status is known. *Methods in Ecology and Evolution*, *6*(5), 557–565. https://doi.org/10.1111/2041-210X.12342

Miller, J. A. (2014). Virtual species distribution models: Using simulated data to evaluate aspects of model performance. *Progress in Physical Geography: Earth and Environment*, *38*(1), 117–128. https://doi.org/10.1177/0309133314521448

Naeem, S., & Li, S. (1997). Biodiversity enhances ecosystem reliability. *Nature*, *390*(6659), 507–509. https://doi.org/10.1038/37348

Ovaskainen, O. (2002). Long-Term Persistence of Species and the SLOSS Problem. *Journal of Theoretical Biology*, *218*, 491–433. https://doi.org/10.1006/yjtbi.3089

Peck, S. L. (2004). Simulation as experiment: A philosophical reassessment for biological modeling. *Trends in Ecology & Evolution*, *19*(10), 530–534. https://doi.org/10.1016/j.tree.2004.07.019

Petrovskii, S., & Petrovskaya, N. (2012). Computational ecology as an emerging science. *Interface Focus*, *2*(2), 241–254. https://doi.org/10.1098/rsfs.2011.0083

R Core Team. (2022). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Sagarin, R. D., Adams, J., Blanchette, C. A., Brusca, R. C., Chorover, J., Cole, J. E., Micheli, F., Munguia-Vega, A., Rochman, C. M., Bonine, K., Van Haren, J., & Troch, P. A. (2016). Between control and complexity: Opportunities and challenges for marine mesocosms. *Frontiers in Ecology and the Environment*, *14*(7), 389–396. https://doi.org/10.1002/fee.1313

Stelzer, C.-P., & Groffman, D. (2025). Wellcounter: Automated high-throughput phenotyping for aquatic microinvertebrates. *Methods in Ecology and Evolution*, *16*, 958–971.

Stern, E. R., Vesk, P. A., Clements, C. C., & McCarthy, M. A. (2025). Contribution of microcosm experiments to conservation science. *Conservation Biology*, e70129. https://doi.org/10.1111/cobi.70129

Stewart, R. I. A., Dossena, M., Bohan, D. A., Jeppesen, E., Kordas, R. L., Ledger, M. E., Meerhoff, M., Moss, B., Mulder, C., Shurin, J. B., Suttle, B., Thompson, R., Trimmer, M., & Woodward, G. (2013). Mesocosm Experiments as a Tool for Ecological Climate-Change Research. In *Advances in Ecological Research* (Vol. 48, pp. 71–181). Elsevier. https://doi.org/10.1016/B978-0-12-417199-2.00002-1

Tucker, A. M., & Runge, M. C. (2021). Optimal Strategies for Managing Wildlife Harvest Under System Change. *The Journal of Wildlife Management*, *85*(5), 847–854. https://doi.org/10.1002/jwmg.22047

Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2020). *Modelling species presence-only data with random forests* (p. 2020.11.16.384164). bioRxiv. https://doi.org/10.1101/2020.11.16.384164

Valente, J. J., Jirinec, V., & Leu, M. (2024). Thinking beyond the closure assumption: Designing surveys for estimating biological truth with occupancy models. *Methods in Ecology and Evolution*, *15*(12), 2289–2300. https://doi.org/10.1111/2041-210X.14439

Wichterman, R. (1986). *The Biology of Paramecium*. Springer US. https://doi.org/10.1007/978-1-4757-0372-6

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* [Computer software]. Springer-Verlag New York.

Woodruff, L. L. (1908). The Life Cycle of Paramecium when Subjected to a Varied Environment. *The American Naturalist*, *42*(500), 520–526. https://doi.org/10.1086/278963

Woodruff, L. L. (1909). Studies on the life cycle of Paramecium. *Experimental Biology and Medicine*, *6*(5), 117–118. https://doi.org/10.3181/00379727-6-62

Woodruff, L. L. (1911). The effect of excretion products of Paramaecium on its rate of reproduction. *Journal of Experimental Zoology*, *10*(4), 557–581. https://doi.org/10.1002/jez.1400100407

Woodruff, L. L., & Baitsell, G. A. (1911). THE TEMPERATURE COEFFICIENT OF THE RATE OF REPRODUCTION OF PARAMÆCIUM AURELIA. *American Journal* *of Physiology-Legacy Content*, *29*(2), 147–155. https://doi.org/10.1152/ajplegacy.1911.29.2.147