



Herzyk, Tymon Alexander (2025) *Towards a general theory within wastewater treatment: computational and experimental examples utilising fundamental laws to describe microbial community structure in wastewater treatment systems*. PhD thesis.

<https://theses.gla.ac.uk/85429/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>
research-enlighten@glasgow.ac.uk

Towards a general theory within wastewater treatment: computational and experimental examples utilising fundamental laws to describe microbial community structure in wastewater treatment systems

Tymon Alexander Herzyk

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy in Civil Engineering

School of Engineering
College of Science and Engineering
University of Glasgow



University
of Glasgow

August 2025

Abstract

Wastewater engineering needs novel, sustainable systems designed to cope with the pressures of climate change, urbanisation and increasing water scarcity. Modern molecular methods in microbiology are allowing us to interrogate the complex microbial communities underpinning biological wastewater treatment technologies, which affords the opportunity to design and control them. The application of theory is a fundamental tool to achieve this, however, current theories, whilst numerous, are partial and idiosyncratic. A concerted research effort is required to generate generic, widely accepted theory for wastewater treatment design. This thesis approaches the problem of defining fundamental rules on the structure and dynamics of wastewater microbial communities from two different perspectives.

The first asks the question if there was a generic theory on community assembly for a wastewater microbial community, is it possible to identify its parameters using the type of data that is routinely collected? To answer this an empirical approach is used where a model is assumed, in this case the simple neutral model. It is applied to generate a range of synthetic relative abundance time series with a range of different features, such as different sampling frequencies. The model is then calibrated using these data under various scenarios to assess parameter identifiability. The process of efficiently simulating, sampling and calibrating was achieved by developing an open-source user-friendly and adaptable computational framework. The results demonstrate fundamental difficulties in identifying neutral model parameters, even when using idealised time series. This is due to misrepresentation of the expected variance and correlated parameters. As such, there is a need to independently determine some of these parameters *a priori*. Reduced sampling periods and frequencies are also shown to impact model calibration, leading to systematic errors in the estimates of certain parameters. The research highlights the need for increasing information within relative abundance time series and demonstrates that this can be achieved by inducing a perturbation on a system. Thus, producing time series during periods of non-equilibrium is shown to be beneficial. Finally, it is demonstrated that for real data sets neutral model parameters must be considered as "effective" parameters as the estimates obtained will likely reflect a myriad of complex phenomena.

The second perspective is mechanistic. Growth kinetics, known to drive inter-species competition, are explored through energy and thermodynamics. This is achieved experimentally by culturing two species of methanotrophs, *Methylomonas methanica* S1 and *Methylosinus trichosporium* OB3b, within an isothermal calorimeter. Comparisons are made between the heat dissipated by the species and how they partition the avail-

able carbon between energy-yielding and biosynthetic reactions. The methodology developed for measuring heat dissipation is novel in its application to methanotrophic bacteria. These measurements identify that a significant amount of heat is dissipated during the growth of methanotrophic bacteria. Differences in growth are also observed with *Methylosinus trichosporium* producing more CO_2 at the cost of a reduced biomass yield when compared to *Methylomonas methanica*. This difference in carbon partitioning is shown to be linearly related to the heat dissipated per unit of biomass, however, the impact this has on providing a competitive advantage is unclear. It is speculated that additional biochemical and physical phenomena need to be quantified in order to resolve the relative importance of thermodynamics on the different kinetics.

Simple models with parameters and variables that are easily understood and manipulated by engineers have a long history of being successfully deployed in engineering design. In this thesis, two approaches have been explored to describe the assembly and kinetics of microbial communities. It is shown that identifying model parameters for even the simplest descriptions of community assembly, neutral models, is fraught with conceptual and practical difficulties. Notwithstanding this, their ability to capture phenomena is apparent, and the challenge for engineering design is mapping the calibrated parameters onto characteristics of engineered systems. A more mechanistic approach attempts to use energy and thermodynamics to explain kinetics using methanotroph species as an example. Energy-based methods are the basis of many models used in engineering design because they can capture physical phenomena while eschewing much of the underlying complexities in physical or chemical processes. The thesis falls short of delivering a generic rule on the relationship between energy dissipated and kinetics. However, the empirical data generated are already being used in the design of technology for heat recovery from wastewater treatment. In general, the thesis demonstrates the merits of pursuing simplicity in models for generic application in environmental biotechnologies.

Contents

Acknowledgements	xxii
Declaration	xxiii
1 Introduction	1
1.1 Background	1
1.2 Aims and objectives	3
1.3 Thesis structure	6
2 Literature review	7
2.1 Wastewater engineering	7
2.1.1 A brief history	7
2.1.2 Current wastewater treatment technologies	9
2.1.3 The microbial community	12
2.1.4 Challenges	13
2.2 Theoretical community ecology	15
2.2.1 The role of mathematical models	15
2.2.2 Early models	17
2.2.3 Neutral theory	19
2.2.4 Application of neutral theory for wastewater engineering	22
2.3 Thermodynamics within microbial biology	24
2.3.1 A thermodynamic description of life	24
2.3.2 Thermodynamics and microbial growth	26
2.3.3 The measurement of heat dissipation	30
2.4 Methanotrophs	31
2.4.1 Methane	31
2.4.2 Aerobic methanotrophic bacteria	33
2.4.3 Application within biotechnologies	36
2.5 Summary	37

3	Identifying neutral models on the basis of species abundance time series.	38
3.1	Abstract	38
3.2	Introduction	39
3.3	Methods	40
3.3.1	Computational framework	40
3.3.2	Discrete simulation of microbial communities using NCMs	43
3.3.3	Sampling of simulated single-species relative abundance time series	45
3.3.4	Calibration of NCMs for estimation of individual parameters	45
3.4	Results and discussion	48
3.4.1	Calibration of neutral models to idealised single species relative abundance time series	48
3.4.1.1	Calibration of Hubbell's purely neutral model	49
3.4.1.2	Calibration of Sloan et al.'s near-neutral model	54
3.4.1.3	Truncated models	57
3.4.2	Calibration of neutral models to non-idealised single species relative abundance time series	58
3.4.2.1	The effects of duration and frequency of sampling on neutral model identification	58
3.4.2.2	Time series captured from periods of non-equilibrium	67
3.4.2.3	The role of effective parameters	73
3.5	Conclusion and future perspectives	76
4	Quantifying the heating potential of methanotrophs for application within decentralised wastewater treatment technologies	80
4.1	Abstract	80
4.2	Introduction	81
4.3	Methods	83
4.3.1	Materials	83
4.3.1.1	Methanotrophs	83
4.3.1.2	Media	83
4.3.1.3	Methane	84
4.3.1.4	Culture bottles	84
4.3.2	Experimental methods	84
4.3.2.1	Maintenance of stock cultures	84
4.3.2.2	Experimental conditions	85
4.3.2.3	Heat dissipation experiments	86
4.3.2.4	External growth experiments	87
4.3.2.5	Sampling	87
4.3.3	Analytical methods	88
4.3.3.1	Headspace pressure	88

4.3.3.2	Optical density	88
4.3.3.3	Cell dry weight	89
4.3.3.4	pH	89
4.3.3.5	Dissolved oxygen	89
4.3.3.6	Non-purgeable organic carbon	89
4.3.3.7	Headspace composition	90
4.4	Results and Discussion	90
4.4.1	Growth dynamics and yields	90
4.4.2	Carbon balance and partitioning	96
4.4.3	Thermal output	99
4.4.4	Relationship between heat dissipation and growth	105
4.5	Conclusion and future perspectives	108
5	Conclusion and future perspectives	109
A	European census data displaying the percentage of population connected to independent wastewater treatment	117
B	Calculation of expected mean from continuous stochastic differential equations for Hubbell's neutral model and Sloan et al.'s near-neutral model	119
C	Derivation of continuous stochastic differential equations for Hubbell's neutral model and Sloan et al.'s near-neutral model	121
D	Results from truncated version of Hubbell's neutral model when fitted to single species relative abundance time series	128
E	Results from truncated version of Sloan et al.'s near-neutral model when fitted to single species relative abundance time series	131
F	Design and volumes of bespoke culture bottles for the Calmetrix I-CAL 4000 HPC Isothermal Calorimeter	134
G	Calculation of the theoretical pressure loss within culture bottles	137
H	Post-processing of thermal power measurements from the Calmetrix I-CAL 4000 HPC Isothermal Calorimeter	142
I	Calibration curves for OD_{600} vs CDW for <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b	147
J	Measurements of dissolved oxygen for batch cultures of <i>M. trichosporium</i> OB3b and <i>M. methanica</i> S1	148

K	Calculation of total number of moles of CH_4 and CO_2 within culture bottles	149
L	Calculation of catabolic heat	155

List of Tables

1.1	Thesis and chapter-specific aims and objectives	5
3.1	Starting values, upper limits and lower limits set within the optimisation for estimating parameters N_T , η , m , p and α	47
3.2	Summary table of mean and standard deviation in statistics calculated across all 100 single-species relative abundance time series portrayed in Figure 3.5.	50
3.3	Summary table of mean and standard deviation in statistics calculated across all 100 single-species relative abundance time series portrayed in Figure 3.9. Expected values for these statistics based on the parameters used within the simulation of these time series are also provided. .	56
3.4	Summary table of mean and standard deviation in statistics across different sets of 100 relative abundance time series sampled over different sampling periods. Sampling frequency is held constant at a value of 0.0004 hours across all time series. Values are shown for sets of time series simulated according to Hubbell's and Sloan et al.'s models.	59
3.5	Summary table of mean and standard deviation in statistics across different sets of 100 relative abundance time series sampled using different sampling frequencies. Sampling period is held constant at a value of 4000 hours across all time series. Values are shown for sets of time series simulated according to Hubbell's and Sloan et al.'s models.	60
3.6	Summary of parameters fixed during the calibration of parameters within Hubbell's neutral model during the analysis of sampling regimes.	61
3.7	Summary of parameters fixed during the calibration of parameters within Sloan et al.'s near-neutral model during the analysis of sampling regimes.	61
4.1	Comparison of characteristics between <i>Methylobacterium methanica</i> S1 and <i>Methylobacterium trichosporium</i> OB3b (Bowman, 2006; Hanson & Hanson, 1996; Koh et al., 1993; Murrell & Dalton, 1983; Whittenbury et al., 1970).	83
4.2	Injection volumes of CH_4 and air used during inoculation to provide target CH_4 concentrations of 8%, 4%, and 2%v/v in air. Volumes were determined through trial and error.	87
4.3	Final values measured for the headspace concentration of CO_2 , OD_{600} , $NPOC$, headspace pressure and pH for pure batch cultures of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b from external and calorimeter experiments. Means are calculated from biological triplicates. The standard deviation in these values is also expressed.	93

4.4	Final yields of cell dry weight (Y_{CDW/CH_4}), non-purgeable organic carbon (Y_{NPOC/CH_4}) and carbon dioxide (Y_{CO_2/CH_4}) calculated for pure batch cultures of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b. Carbon conversion efficiencies calculated from cell dry weight (CCE_{CDW}) and non-purgeable organic carbon (CCE_{NPOC}) are also provided. Calculations are undertaken using the measurements provided in Figures 4.3 and 4.4 respectively. Means are calculated from biological triplicates. The standard deviation in these values is also expressed.	94
4.5	Final percentages of carbon stored in the initial volume of CH_4 converted to biomass and CO_2 , measured for calorimeter and external batch cultures of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b. The ratio of the carbon converted to CO_2 and that stored in biomass is provided. All values are calculated as means from biological triplicates, with standard deviations also shown.	98
4.6	Thermal efficiencies (η_{max} and η_{cat}) for batch cultures of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b for different regimes of initial CH_4 concentration. The maximum available heat (Q_{max}), the maximum available heat from catabolism (Q_{cat}) and the total heat dissipated (Q_{dis}) are also provided.	101
4.7	Yields of heat dissipated per cell dry weight ($Y_{Q/CDW,dis}$) and heat dissipated per mole of CH_4 consumed ($Y_{Q/CH_4,dis}$) for batch cultures of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b for different regimes of initial CH_4 concentration. For calorimeter cultures, yields are calculated using the difference in quantities over the entire growth period, as expressed by Equation 4.4. For external cultures, yields are calculated by plotting quantities at similar time points against each other and calculating the gradient of a line fitted to this plot. Figures 4.9 and 4.10 demonstrate these plots. Maximum and catabolic yields of heat dissipated per mole of CH_4 are also provided, where $Y_{Q/CH_4,max}$ refers to the standard enthalpy the oxidation of CH_4 and $Y_{Q/CH_4,cat}$ is calculated by dividing the maximum heat generated through catabolism by the number of moles of CH_4 consumed. This is done using final samples from cultures grown in the calorimeter and externally.	102
F.1	Calculation of individual bottles volume. Volumes were determined by measuring the empty weight and the filled weight of each bottle and finding the difference between these values. . .	136
L.1	Enthalpy of formation of molecules associated with the biological oxidation of methane. Values are obtained from The US National Institute of Standards and Technology (NIST).	155

List of Figures

2.1	Schematics of typical septic tanks with two compartments, for populations of up to 30. Reproduced from BSI (1983). a) combined, b) separate.	10
2.2	Diagram defining the cell boundary and processes driving the entropy balance for a growing cell. Reproduced from von Stockar et al. (2006)	25
2.3	A black box description of microbial growth, highlighting generalised reactants and products. Reproduced from Heijnen and Van Dijken (1992)	27
2.4	A energy convertor description of microbial growth, highlighting the coupling of catabolic and anabolic reactions. Reproduced from Heijnen and Van Dijken (1992)	28
2.5	Relationships within a linear energy converter. Reproduced from Stockar and Wielen (2013b). a) Efficiency against force ratio, b) output power against force ratio.	29
2.6	A diagram representing Lavoisier and Laplace’s early ice calorimeter. Reproduced from von Stockar and Marison (2005).	31
2.7	Carbon assimilation pathways of type-I and type-II methanotrophs. Reproduced from Smith et al., 2010. a) carbon assimilation pathway of type I methanotrophs utilising the ribulose monophosphate (RuMP) cycle, b) carbon assimilation pathway of type II methanotrophs utilising the serine cycle,	34
3.1	Basic overview of the computational framework developed for simulating single-species relative abundance time series, sampling such time series and calibrating neutral model parameters to these time series. In total three packages were developed that can be operated individually or in series. Packages operate according to user defined variables stored in associated spreadsheets. All files, in-depth descriptions and guides are available at: https://github.com/tymonherzyk . Created in https://BioRender.com	41

3.2	Comprehensive overview of the computational framework developed for simulating single-species relative abundance time series, sampling such time series and calibrating neutral model parameters to these time series. Processes belonging to each package are demarcated within dashed boxes. Files marked with blue tags refer to MATLAB files, and files marked with green tags refer to <i>xlsx</i> files. Each package is executed through the primary MATLAB script highlighted at the centre of each box. The order of operations carried out by this script are provided by respective numbering. Key variables are marked in clear tags, and key decisions are depicted by grey tags. Solid arrows depict the passage of information. Any processes outside a dashed box require user interaction. All files, in-depth descriptions and guides are available at: https://github.com/tymonherzyk . Created in https://BioRender.com	42
3.3	Example experimental equivalents to the simulated communities. Created in https://BioRender.com a) for Hubbell's model, b) for Sloan et al's model.	44
3.4	Example experimental equivalents to the simulation regime employed. Created in https://BioRender.com a) 100 bioreactors operating in parallel, b) a single bioreactor repeated 100 times	48
3.5	Relative abundance time series of a single species obtained through 100 realisations of a simulated community. Community dynamics were simulated from the discrete transitional probabilities described by Hubbell (2001) for a community governed by ecological drift alone (Equation 3.1). Parameter values used were: $N_T = 10,000$, $\eta = 4.00 \times 10^{-4}$, $m = 0.10$ and $p = 0.20$. Sampling was conducted after every individual replacement event (every 0.0004 hours) over a period of 8,760 hours. a) Time series plot, b) frequency plot.	49
3.6	Box plots representing the estimation of distinct parameters from Hubbell's purely neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.5. a) Estimation of N_T , b) estimation of η , c) estimation of m . d) estimation of p	51
3.7	Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of Hubbell's purely neutral parameters shown in Figure 3.6. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs. . .	52
3.8	a) Negative log-likelihood (NLL) as a function of N_T and η . The value of N_T is given on the y-axis with the value of η on the x-axis. Corresponding values for NLL are displayed as a heat map. All NLL values are calculated by fitting Hubbell's model to a single time series obtained from a simulated community governed by drift alone. During acquisition of these values all other parameters within the model were fixed at their expected values. All NLL values shown are within a range of approximately 0.55% of the minimum value. b) Estimated parameter pairs of N_T and η from naive calibration plotted on top of the centre line of minimum NLL.	53

- 3.9 Relative abundance time series of a single species obtained through 100 realisations of a simulated community. Community dynamics were simulated from the discrete transitional probabilities described by Sloan et al. (2006) for a community governed by ecological drift and selection (Equation 3.2). Parameter values used were: $N_T = 10,000$, $\eta = 4.00 \times 10^{-4}$, $m = 0.10$, $p = 0.20$ and $\alpha = 0.01$. Sampling was conducted after every individual replacement event (every 0.0004 hours) over a period of 8,760 hours. a) Time series plot, b) frequency plot. . . . 54
- 3.10 Box plots representing the estimation of distinct parameters from Sloan et al.'s near-neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.9. a) Estimation of N_T , b) estimation of η , c) estimation of m . d) estimation of p , e) estimation of α 55
- 3.11 Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of Sloan's near-neutral parameters shown in Figure 3.10. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs. . . 56
- 3.12 Heatmaps demonstrating the mean percent error (MPE) for each set of 100 estimations of Hubbell's purely neutral parameters from relative abundance time series sampled under different regimes. The sampling frequency used is represented along the the y-axis. The sampling period used is represented along the x-axis. a) N_T , b) η , c) m , d) p 62
- 3.13 Heatmaps demonstrating the coefficient of variation (CV) for each set of 100 estimations of Hubbell's purely neutral parameters from relative abundance time series sampled under different regimes. The sampling frequency used is represented along the the y-axis. The sampling period used is represented along the x-axis. a) N_T , b) η , c) m , d) p 63
- 3.14 Heatmaps demonstrating the mean percent error (MPE) for each set of 100 estimations of Sloan et al.'s near-neutral parameters from relative abundance time series sampled under different regimes. The sampling frequency used is represented along the the y-axis. The sampling period used is represented along the x-axis. a) N_T , b) η , c) m , d) p , e) α 65
- 3.15 Heatmaps demonstrating the coefficient of variation (CV) for each set of 100 estimations of Sloan et al.'s near-neutral parameters from relative abundance time series sampled under different regimes. The sampling frequency used is represented along the the y-axis. The sampling period used is represented along the x-axis. a) N_T , b) η , c) m , d) p , e) α 66

- 3.16 Relative abundance time series of a single species obtained through 100 realisations of two simulated communities. Time series represent a period of non-equilibrium. This was achieved by setting the starting relative abundance to 0 for each simulation. Sampling was conducted after every individual replacement event (0.0004h) for 200 hours. Parameter values were equal to $N_T = 10,000$, $\eta = 4.00 \times 10^{-4}$, $m = 0.10$ and $p = 0.20$. For realisations including selection $\alpha = 0.01$. a) Community dynamics were simulated from the discrete transitional probabilities described by Hubbell (2001) for a community governed by ecological drift alone (Equation 3.1), b) community dynamics were simulated from the discrete transitional probabilities described by Sloan et al. (2006) for a community governed by ecological drift and selection (Equation 3.2). 68
- 3.17 Box plots representing the estimation of distinct parameters from Hubbell's purely neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.16a. a) Estimation of N_T , b) estimation of η , c) estimation of m . d) estimation of p 69
- 3.18 Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of Hubbell's purely neutral parameters shown in Figure 3.17. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs. . . 70
- 3.19 Box plots representing the estimation of distinct parameters from Sloan et al.'s near-neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.16b. a) Estimation of N_T , b) estimation of η , c) estimation of m . d) estimation of p , e) estimation of α 71
- 3.20 Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of Sloan's near-neutral parameters shown in Figure 3.19. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs. . . 72
- 3.21 Box plots representing the estimation of distinct parameters from Hubbell's purely neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.9. a) Estimation of N_T , b) estimation of η , c) estimation of m . d) estimation of p 74

3.22	Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of Hubbell's purely neutral parameters shown in Figure 3.21. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs.	75
3.23	Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of Hubbell's effective parameters to all 100 distinct time series depicted in Figure 3.9. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs.	76
4.1	Photos of the bottom of stock culture bottles containing either <i>Methylomonas methanica</i> S1 or <i>Methylosinus trichosporium</i> OB3b grown at 30°C with no shaking. a) <i>Methylomonas methanica</i> S1, b) <i>Methylosinus trichosporium</i> OB3b.	85
4.2	Overview of the experimental layout, demonstrating both "External Growth Experiment" and "Heat Dissipation Experiment" setups. Green bottles represent active cultures while blue bottles represent negative controls. Both experiments are undertaken in unison with the same materials, growth conditions and inoculum used. All cultures were grown as pure batch cultures. Experiments were repeated for both <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b over three distinct initial concentrations of CH_4 (8%, 4% and 2%v/v in air). Created in https://BioRender.com	86
4.3	Time series of growth parameters for external and calorimeter cultures of <i>M. methanica</i> S1, measured under three distinct concentration of initial CH_4 over a period of 7 days. a) Headspace CH_4 concentration, b) Headspace CO_2 concentration, c) Optical density at 600nm (OD_{600}), d) Non-purgeable organic carbon (NPOC), e) System pressure, f) pH.	91
4.4	Time series of growth parameters for external and calorimeter cultures of <i>M. trichosporium</i> OB3b, measured under three distinct concentration of initial CH_4 over a period of 7 days. a) Headspace CH_4 concentration, b) Headspace CO_2 concentration, c) Optical density at 600nm (OD_{600}), d) Non-purgeable organic carbon (NPOC), e) System pressure, f) pH.	92
4.5	Final yields of cell dry weight (Y_{CDW/CH_4}), and carbon dioxide (Y_{CO_2/CH_4}) plotted against initial headspace concentration of CH_4 for calorimeter and external batch cultures of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b. Linear fits are made using all points plotted for <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b respectively. Equations of these lines and the R^2 values are provided in figure legends. a) Y_{CDW/CH_4} , b) Y_{CO_2/CH_4}	95
4.6	Carbon composition of CH_4 , CO_2 and biomass expressed as a percent of the initial carbon within the system. Measurements of CH_4 , CO_2 and biomass are obtained from batch cultures of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b over three different initial concentration of CH_4 , as shown in Figures 4.3 and 4.4 respectively. a) <i>M. methanica</i> S1 8%v/v CH_4 in air b) <i>M. trichosporium</i> OB3b 8%v/v CH_4 in air, c) <i>M. methanica</i> S1 4%v/v CH_4 in air, d) <i>M. trichosporium</i> OB3b 4%v/v CH_4 in air, e) <i>M. methanica</i> S1 2%v/v CH_4 in air. f) <i>M. trichosporium</i> OB3b 2%v/v CH_4 in air.	97

4.7	Carbon partitioning ratio (<i>CPR</i>) plotted against the initial concentration of CH_4 for <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b.	99
4.8	Thermal power and thermal energy dissipated during the batch growth of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b under three initial concentration of CH_4 (8%, 4% and 2%v/v in air). Thermal power dissipation was measured directly using a Calmetrix I-CAL 4000 HPC Isothermal Calorimeter. Total thermal energy was determine through integrating between measurements of thermal power and summing the results. a) Thermal power dissipation from <i>M. methanica</i> S1, b) thermal power dissipation from <i>M. trichosporium</i> OB3b, c) total thermal energy dissipation from <i>M. methanica</i> S1, d) total thermal energy dissipation from <i>M. trichosporium</i> OB3b.	100
4.9	Change in thermal energy dissipated against the change in cell dry weight (<i>CDW</i>) for batch cultures of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b under different initial concentrations of CH_4 . Heat dissipation is measured from cultures within the Calmetrix I-Cal 4000 HPC Isothermal Calorimeter. Values of <i>CDW</i> are produced from OD_{600} measurements from parallel external cultures. Values plotted against each other refer to the same time point from the experiment start time. Linear fits are produced using all points plotted for a single species. a) 8%v/v CH_4 in air, b) 4%v/v CH_4 in air, c) 2%v/v CH_4 in air.	103
4.10	Change in thermal energy dissipated against the change in moles of CH_4 for batch cultures of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b under different initial concentrations of CH_4 . Heat dissipation is measured from cultures within the Calmetrix I-Cal 4000 HPC Isothermal Calorimeter. The number of moles of CH_4 are produced from analysis of headspace composition of parallel external cultures. Values plotted against each other refer to the same time point from the experiment start time. Linear fits are produced using all points plotted for a single species. Lines corresponding to $Y_{Q/CDW,cat}$ calculated from external cultures, and $Y_{Q/CDW,max}$ are also plotted for both <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b to provide comparison. a) 8%v/v CH_4 in air, b) 4%v/v CH_4 in air, c) 2%v/v CH_4 in air.	104
4.11	Yield of heat per cell dry weight ($Y_{Q/CDW,dis}$) plotted against the carbon partitioning ratio (<i>CPR</i>) for batch cultures of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b grown under different concentration of CH_4 in air. A liner relationship is observed across all points with the line $y = 36.01x$ providing a fit with $R^2 = 0.94$	106
5.1	Diagram representing the link between neutral model parameters and their attributed mechanistic meaning when using 'idealised' and 'real' relative abundance time series. Here we define 'idealised' time series as those produced from communities only governed by the processes described by the model calibrated, whereas 'real' time series refer to those produced from communities where the structuring dynamics differ to those described by the model calibrated. a) idealised time series, b) real time series.	111

5.2	Approaches for determining the phenomena that drive effective neutral model parameters. One is an empirical approach where time series from differing communities are obtained and model parameters are calibrated to observe differences in the values obtained. The second is a mechanistic approach that looks to derive parameters from first principles.	112
A.1	European census data displaying the percentage of population connected to independent wastewater treatment. Data available at: https://ec.europa.eu/eurostat	118
D.1	Box plots representing the estimation of distinct parameters from the truncated version of Hubbell's purely neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.5. a) Estimation of N_T , b) estimation of η , c) estimation of m . d) estimation of p	129
D.2	Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of the truncated version of Hubbell's purely neutral parameters shown in Figure D.1. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs.	130
D.3	Estimated values of N_T and η obtained through naive calibration of the truncated version of Hubbell's neutral model to the 100 distinct time series depicted in Figure 3.5 plotted against each other and on top of the midline of negative log-likelihood (NLL) from Figure 3.8.	130
E.1	Box plots representing the estimation of distinct parameters from the truncated version of Sloan et al.'s near-neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.9. a) Estimation of N_T , b) estimation of η , c) estimation of m . d) estimation of p , e) estimation of α	132
E.2	Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of the truncated version of Sloan et al.'s near-neutral parameters shown in Figure E.1. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs.	133
F.1	Initial engineering drawings created for bespoke Calmetrix I-CAL 4000 HPC Isothermal Calorimeter bottles. During final production wall thickness was decreases to 2.5mm, increasing the internal volume of the bottles. The internal neck diameter was also decreased during production from 13mm to 12mm.	135

H.1	Raw thermal power measurements made using a Calmetrix I-CAL 4000 HPC Isothermal Calorimeter from batch cultures of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b under three initial concentration of CH_4 (8%, 4% and 2%v/v in air). a) <i>M. methanica</i> S1 8%v/v CH_4 , b) <i>M. trichosporium</i> OB3b 8%v/v CH_4 , c) <i>M. methanica</i> S1 4%v/v CH_4 , d) <i>M. trichosporium</i> OB3b 4%v/v CH_4 , e) <i>M. methanica</i> S1 2%v/v CH_4 , f) <i>M. trichosporium</i> OB3b 2%v/v CH_4	143
H.2	Thermal power measurements provided in figure H.1 with baseline reduction. a) <i>M. methanica</i> S1 8%v/v CH_4 , b) <i>M. trichosporium</i> OB3b 8%v/v CH_4 , c) <i>M. methanica</i> S1 4%v/v CH_4 , d) <i>M. trichosporium</i> OB3b 4%v/v CH_4 , e) <i>M. methanica</i> S1 2%v/v CH_4 , f) <i>M. trichosporium</i> OB3b 2%v/v CH_4	144
H.3	Thermal power measurements provided in figure H.1 with baseline and settling period reduction. a) <i>M. methanica</i> S1 8%v/v CH_4 , b) <i>M. trichosporium</i> OB3b 8%v/v CH_4 , c) <i>M. methanica</i> S1 4%v/v CH_4 , d) <i>M. trichosporium</i> OB3b 4%v/v CH_4 , e) <i>M. methanica</i> S1 2%v/v CH_4 , f) <i>M. trichosporium</i> OB3b 2%v/v CH_4	145
H.4	Thermal energy measurements calculated from the thermal power measurements provided in figure H.3. a) <i>M. methanica</i> S1 8%v/v CH_4 , b) <i>M. trichosporium</i> OB3b 8%v/v CH_4 , c) <i>M. methanica</i> S1 4%v/v CH_4 , d) <i>M. trichosporium</i> OB3b 4%v/v CH_4 , e) <i>M. methanica</i> S1 2%v/v CH_4 , f) <i>M. trichosporium</i> OB3b 2%v/v CH_4	146
I.1	Calibration curve for determining cell dry weight from OD_{600}	147
J.1	Time series of dissolved oxygen measurements for external and calorimeter cultures of <i>M. methanica</i> S1 and <i>M. trichosporium</i> OB3b, measured under three distinct concentration of initial CH_4 over a period of 7 days. a) <i>M. methanica</i> S1, b) <i>M. trichosporium</i> OB3b	148

List of Symbols

Basic growth equation

ΔN Change in population size

Δt Change in time

B Births

D Deaths

E Emigration

I Immigration

Verhulst's logistic growth equation

$\frac{dN}{dt}$ Rate of change of population size with respect to time

K Carrying capacity

N Population size

r Intrinsic growth rate

Lotka-Volterra model

$\frac{dx}{dt}$ Rate of change of mass of prey with respect to time

$\frac{dy}{dt}$ Rate of change of mass of predators with respect to time

a Growth rate of prey in the absence of predators

b Predation rate coefficient

c Natural death rate of predators in the absence of prey

d Reproduction rate coefficient

x Total mass of prey

y Total mass of predators

Genetic drift

i Abundance of allele A

N Population size

N_e Effective population size

Neutral theory

α Local selection regarding the monitored species

α_e Effective local selection

α_i Local selection regarding species i

η Average time between replacement events

η_e Effective average time between replacement events

$\mu_{hubbell}$ Mean (Hubbell's model)

μ_{sloan} Mean (Sloan et al.'s model)

$\sigma_{hubbell}$ Variance (Hubbell's model)

σ_{sloan} Variance (Sloan et al.'s model)

$d\tau$ Infinitesimal change in number of replacement events

dt Infinitesimal change in time

$dx_{hubbell}$ Infinitesimal change in relative abundance of the monitored species (Hubbell's model)

$dx_{i,hubbell}$ Infinitesimal change in relative abundance of species i (Hubbell's model)

$dx_{i,sloan}$ Infinitesimal change in relative abundance of species i (Sloan et al.'s model)

dx_{sloan} Infinitesimal change in relative abundance of the monitored species (Sloan et al.'s model)

m Immigration rate

m_e Effective immigration rate

N_i Number of individuals in species i

N_T Community size

N_{Te}	Effective community size
$NLL_{hubbell}$	Negative log-likelihood (Hubbell's model)
NLL_{sloan}	Negative log-likelihood (Sloan et al.'s model)
p	Relative abundance of the monitored species within the immigrating community
p_e	Effective relative abundance within the immigrating community
p_i	Relative abundance of species i within the immigrating community
W	Standard Weiner process
x	Relative abundance of the monitored species
x_i	Relative abundance of species i

Thermodynamics

$\bar{\mu}_i$	Chemical potential of i^{th} metabolite
$\bar{\mu}_x$	Chemical potential of biomass
\bar{h}_i	Partial molar enthalpy of the i^{th} metabolite
\bar{h}_i	Partial molar enthalpy of the produced biomass
\bar{s}	Partial molar entropy
\bar{s}_i	Partial molar entropy of the i^{th} metabolite
\bar{s}_x	Partial molar entropy of produced biomass
ΔG	Change in standard Gibbs free energy
ΔG_{ana}	Change in standard Gibbs free energy of the anabolic reaction
ΔG_a	Change in standard Gibbs free energy of reaction a
ΔG_b	Change in standard Gibbs free energy of reaction b
ΔG_{cat}	Change in standard Gibbs free energy of the catabolic reaction
$\Delta_r G$	Change in standard Gibbs free energy of a coupled reaction
\dot{n}_i	Molar transfer rate of the i^{th} metabolite
\dot{n}_x	Molar transfer rate of the produced biomass

\dot{Q}	Rate of heat transfer
\dot{S}_{gen}	Rate of entropy generated
\dot{S}_{in}	Rate of entropy in
\dot{S}_{out}	Rate of entropy out
\dot{S}_{prod}	Rate of entropy production
\dot{W}	Rate of work done
η	Efficiency
η^{BB}	Black box Efficiency
$\eta_{thermal}$	Thermal efficiency
$\frac{dG}{dt}$	Rate of change of Gibbs free energy with respect to time
$\frac{dH}{dt}$	Rate of change of enthalpy with respect to time
$\frac{dS}{dt}$	Rate of change of entropy with respect to time
G_{cell}	Gibbs free energy of the cell
G_{sys}	Gibbs free energy of the system
H_{cell}	Enthalpy of the cell
H_{sys}	Enthalpy of the system
J	Flux
L	Flux/force coefficient
Q	Heat transfer
r_a	Rate of reaction a
r_b	Rate of reaction b
S_{cell}	Entropy of the cell
S_{sys}	Entropy of the system
T	Absolute temperature
t	time

Z Force

Bacterial growth

CCE Carbon conversion efficiency

CCE_{CDW} Carbon conversion efficiency of cell dry weight

CCE_{NPOC} Carbon conversion efficiency of non-purgeable organic carbon

Q Heat dissipated

S Substrate

X Product

Y_{CDW/CH_4} Yield of cell dry weight per CH_4

Y_{CO_2/CH_4} Yield of CO_2 per CH_4 .

Y_{NPOC/CH_4} Yield of non-purgeable organic carbon per CH_4

$Y_{Q/CDW}$ Yield of heat dissipated per cell dry weight

Y_{Q/CH_4} Yield of heat dissipated per CH_4

$Y_{Q/X}$ Yield of heat dissipated per product

$Y_{X/S}$ Yield of biomass produced per substrate

Acknowledgements

Firstly, I would like to acknowledge and express my deep appreciation for the support both my supervisors, Professor William T. Sloan and Dr Rebeca Gonzalez-Cabaleiro, have provided throughout my PhD. I would like to especially thank my primary supervisor Professor William T. Sloan for the continued interest and motivation that drove me to complete my PhD. I look forward to expanding on the research undertaken here and hope to continue working together in the future. I would also like to take this moment to offer special thanks to Dr Tania Gomez-Borraz, who I could always rely on for help and expertise, especially in my work regarding methanotrophic cultures. Tania also taught me a significant amount regarding wastewater engineering and I thank her for her company throughout my PhD, especially during the many hours driving to and from Gauldry.

I wish to extend my gratitude to the entire water and environment group at the University of Glasgow, who provided a healthy and supportive atmosphere for conducting research. I would like to thank in particular Dr Ayo Ogundero for support and training in regard to flow cytometry, Dr Anca Elena Amariei for providing training in gas chromatography and Dr Baptiste Poursat for training me in analytical carbon measurements. I would also like to acknowledge the vital work undertaken by the entire Biological Services technical staff, especially, Julie Russell, Alysha Hunter, Megan Blaney, Charlotte Burgess, Eirinn-Rose McWilliams and Katarzyna Klimczak, whom without, this research would not have been possible. I also express my gratitude to Professor Fin Stuart, Orestis Gazetas and Alex McDonald for facilitating my gas chromatography analysis at SUERC.

A heartfelt thanks to my partner Katie Brydie. Her lasting love, compassion, and resilience provided the motivation required to endeavour over the past four years. I am inspired by the daily support, understanding and kindness you have shown me throughout this entire process, without which, it is unlikely that I would have persisted. Finally, I cannot go without thanking my family, Anna Amtmann, Pawel Herzyk and Dominik Herzyk for their unwavering love and support. The emotional and financial support they provided throughout my life is why I am in a position to achieve my PhD. It is to them that I dedicate this thesis.

Declaration

With the exception of Chapters 1 and 2, which contain introductory material, all work in this thesis was carried out by the author unless otherwise explicitly stated.

Chapter 1

Introduction

1.1 Background

The commencement of life as we understand it would not have been possible without water (Vieira et al., 2020; Westall & Brack, 2018). Indeed, our dependency on water reflects the fundamental role this resource plays in facilitating carbon-based life. In ancient cultures, the importance of water was reflected in the deities worshipped (Angelakis et al., 2012). Currently, it is expressed within the goals set by intergovernmental agencies. For example, goal six of the UN's "Sustainable Development Goals" is solely concerned with ensuring clean water and sanitation for all (United Nations, 2015). Current anthropogenic and environmental pressures endanger water availability, which is presenting new engineering challenges. (T. C. Brown et al., 2019; Gosling & Arnell, 2016; Nations, 2024; Schlosser et al., 2014; Vörösmarty et al., 2000).

Trends at a global level show that water use, across all sectors, has increased consistently since 1950 (Flörke et al., 2013). Factors such as an increasing population, urbanisation, expansion of agriculture and economic growth will act to continue this trend, driving an increase in the global water demand over the coming decades (Dawadi & Ahmad, 2013; Ercin & Hoekstra, 2014; Flörke et al., 2018; Hanasaki et al., 2013; C. He et al., 2021). In parallel, pollution, drought and ice-losses are expected to further decrease remaining freshwater supplies in many regions (Pierrat et al., 2023; Rodell et al., 2018; M. Wang et al., 2024). The cumulative result of these pressures is an increase in the number of people facing water scarcity. Current estimates show that approximately 4 billion people face severe water scarcity for at least one month of the year (Mekonnen & Hoekstra, 2016; WWAP, 2018), projections demonstrate that this could rise to over 6 billion people by 2050 (Boretti & Rosa, 2019; WWAP, 2018). This increase in water scarcity is expected to drive economic uncertainty (Dolan et al., 2021), health risks (Motoshita et al., 2011; WHO, 2019), and food insecurity (Mancosu et al., 2015). There is an additional concern in regions where trans-boundary agreements are lacking, as securing water from shared basins could exacerbate social and political tensions (De Stefano et al., 2017). Water supply shortages make it into the top five national risks of seven countries, while the combined supply shortage of food and water is identified as the fourth most severe global risk of the coming decade (World Economic Forum, 2024).

The technologies we use to dispose of and treat wastewater play a key role in water security. Wastewater irrigation, reuse and resource recovery technologies are being heralded as a necessity in tackling water scarcity (United Nations Environment Programme, 2023), however, these technologies can also be counterproductive if not managed correctly. Wastewater treatment technologies are routinely observed to fail, for mechanical (Taheriyoun & Moradinejad, 2014; K. Wang et al., 2017), biological (Heine et al., 2002) and economical reasons (Murray & Drechsel, 2011). Even treated wastewater has been shown to be discharged at poor quality (Charles et al., 2005). As a result, treated and untreated wastewater has the potential to contaminate surface and groundwater, deteriorating drinking water supplies further (United Nations, 2022b). In low-income and emerging economies, untreated wastewater is the leading factor in water quality decline (du Plessis, 2022). Wastewater treatment is also a leading contributor to greenhouse gas emissions, with centralised WWTPs estimated to consume up to 3% of the electricity produced globally (Gong et al., 2024). Although the quantity of greenhouse gas emissions are harder to estimate for smaller decentralised systems, such as septic tanks, they are known to produce direct emissions (Risch et al., 2021).

These failings in wastewater treatment reflect a design approach trapped in time. Indeed, there is a growing consensus that we should move away from design principles adopted within the 20th century, to ones that promote the hybridisation and integration of centralised and decentralised technologies (Eggimann et al., 2015; Libralato et al., 2012). There is also a push to look towards the management of these systems through microbial, rather than operational, parameters (Carballa et al., 2015; Connelly et al., 2019; Vanwonterghem et al., 2014). Progress toward this paradigm shift, is however hindered by the lack of generalised theory. Indeed, unlike other engineering disciplines, such as structural, mechanical or aeronautical, the theoretical framework informing wastewater engineering is rudimentary in comparison, and thus the development of robust theory within this field has been emphasised (T. P. Curtis et al., 2003; Prosser et al., 2007; Rittmann et al., 2006; Sloan & Gómez-Borraz, 2023; Widder et al., 2016). As such, an empirical approach is deployed, which means change is incremental and slow and the economic and social costs of water infrastructure failure are too high to justify the risk of radically new designs that are not grounded in theory. Therefore, it is incumbent on researchers to develop novel theory, that encapsulates and extends the empirical evidence, and can accelerate innovation in wastewater technologies. This thesis explores nascent theories in microbial ecology that might be deployed in biotechnologies.

1.2 Aims and objectives

The weaning from a reliance on empirical data and thus the subsequent quest towards a universal theory within wastewater treatment defines the *raison d'être* of this thesis. The assumption that such a theory must be inherently tied to the individuals that carry out the main body of work, the microbes, is at the fore of the research undertaken. As encouraged by T. Curtis and Fowler (2023), a multifaceted approach is adopted, one that delves into computational, experimental and theoretical exploration. Empirical and mechanistic methodologies are undertaken in parallel, in an attempt to tie parameters, obtained through model calibration, to the fundamental laws of nature. Indeed, phenomenological parameters, such as the Reynolds Number, Gauckler–Manning coefficient and Young's modulus, are routinely used within engineering to inform system design, however, their usefulness is often tied to our understanding of the processes that drive them. Our attempt at fulfilling this aim is expressed through two distinct research chapters, entitled:

1. Identifying neutral models on the basis of species abundance time series.
2. Quantifying the heating potential of methanotrophs for application within decentralised wastewater treatment technologies.

The first of these is positioned to adopt the hypothesis that community ecology, and by extension microbial ecology, is too complex to derive general laws using a reductionist approach (Lawton, 1999; Prosser et al., 2007; Sloan & Gómez-Borraz, 2023). Here, the aim is to tackle the lack of ecological tools for detecting and characterising underlying drivers from the trends and patterns exhibited by microbial communities; as highlighted by Prosser et al. (2007) and Prosser (2020). By applying tools of ecological theory, namely neutral models, to single-species relative abundance time series, it is assessed by purely empirical means whether community structuring processes, such as community size, replacement rate, immigration and local selection, can be calibrated accurately in specific scenarios.

To achieve this, calibrations of two common neutral models, Hubbell's purely neutral model (Hubbell, 2001) and Sloan et al.'s near-neutral model (Sloan et al., 2006), are undertaken using simulated relative abundance time series. Using synthetic rather than experimentally obtained datasets allows both models to be examined under an idealised scenario, in which these data sets are produced over a large period using the maximum temporal resolution theoretical possible. These conditions are then relaxed to quantify what effect diverging towards more realistic scenarios, that might reasonably reflect a real experimental campaign, has on the accuracy of parameter estimates. The first objective is thus to develop a computational framework to allow the simulation of such datasets and therefore facilitate this research. Once datasets have been generated estimates of distinct model parameters are produced through maximum likelihood estimation and compared to those used in generating the synthetic time series. Through this comparison, the accuracy of the estimated parameters can be determined and thus the ability to accurately identify them for the given scenario can be assessed. Based on the results this provides, the final objective is to lay out a series of recommendations that promote the production of accurate parameter estimates from neutral model calibration.

The second research chapter concedes that, even though the microbial community assembly process may be largely stochastic, the microbial dynamics must be a function of the interactions between microbes and their physical and chemical environment (Lawton, 1999; Vellend, 2010; Vellend & Orrock, 2009). Thus, if some simple, tractable laws can be derived to characterise these interactions, they could provide context for parameters determined through purely empirical means or even allow parameters to be estimated theoretically. Indeed, if we assume that accurate calibration of the neutral parameters explored within the first chapter is achievable, then knowledge of the fundamental forces that drive these parameters will increase our ability to engineer the systems they are derived from. The aim of the second research chapter is therefore to explore the formation of community structure from a mechanistic perspective, one which explores the causation tied to the second law of thermodynamics. An attempt is made to demonstrate experimentally, whether growth parameters, known to drive inter-species competition, can be linked to the heat dissipation exhibited during the growth of bacteria.

The heat dissipated from pure batch cultures of *Methylobacterium methanica* S1 and *Methylobacterium trichosporium* OB3b, two species of methanotrophs, is measured alongside culture growth in an attempt to examine whether trade-offs between these processes exist. The reason methanotrophic bacteria are used is three-fold. Firstly, the available energy provided by the catabolism of these bacteria, the oxidation of methane, is significant (Pittam & Pilcher, 1972). Thus, a wider reservoir of thermal energy is offered, potentially increasing the amount of heat dissipated and therefore our ability to measure it. Secondly, the carbon and energy source utilised by these bacteria is the same (Hanson & Hanson, 1996). Coupling between energy-yielding and biosynthetic (energy consuming) reactions is therefore reliant on a single substrate. Thirdly, methanotrophic bacteria have proved useful in emerging methane mitigation technologies for decentralised wastewater treatment systems (Gómez-Borraz et al., 2022). Analysing the heat dissipated from these bacteria therefore carries intrinsic engineering importance. The first objective for this study is to develop a methodology for measuring heat dissipation and growth parameters *in situ*. Further objectives are to observe whether the two species differ in how carbon is partitioned and the amount of heat dissipated, and whether correlations can be made between these processes and observed growth kinetics.

While the aims of both chapters may seem independent of one another, they are linked by the quest toward a general theory. For example, to arrive at a suitable generalised theory, it is crucial not only to assess whether the underlying mathematical model makes testable predictions while accurately describing the data under the scenario applied, but also to ensure it provides fundamental insights (Wieland et al., 2021). In this regard, by assessing the identifiability of model parameters using data and under scenarios relevant to wastewater treatment, the aims of the first research chapter largely tackle the first two of the aforementioned criteria, to establish whether neutral models are indeed appropriate models to base such a theory. The second research chapter aims to provide a context for the usefulness of these models, or the ability to provide fundamental insight, by attempting to derive the more generalised parameters these model rely on, such as a selection coefficient (α), from a point of fundamental physical theory. Indeed, only by achieving both sets of these aims can a convincing and useful theory be embarked towards.

A summary of the aims and objectives outlined for this thesis are laid out in Table 1.1. The aims may seem ambitious, however, any novel movement towards them can be viewed as success. Even failure, if documented, can provide a useful resource for future scientific endeavours (T. Curtis & Fowler, 2023).

Table 1.1: Thesis and chapter-specific aims and objectives

Thesis	Towards a general theory within wastewater treatment: computational and experimental examples utilising fundamental laws to describe microbial community structure in wastewater treatment systems
Aim	To develop a general theory describing microbial community structuring for use within wastewater engineering.
Research Chapter 1	Identifying neutral models on the basis of species abundance time series
Aim 1.0	To determine whether processes driving community structure can be accurately derived through the calibration of neutral models using single-species relative abundance time series.
Objective 1.1	To develop a computational framework for simulating and sampling single-species relative abundance time series, which also facilitates the calibration of Hubbell's and Sloan et al.'s neutral models to the produced time series.
Objective 1.2	To assess using idealised time series whether model parameters are structurally identifiable.
Objective 1.3	To assess using sampled time series whether model parameters are practically identifiable.
Objective 1.4	To provide a series of recommendations for ensuring accurate identification of parameters from the calibration of Hubbell's and Sloan et al.'s neutral models.
Research Chapter 2	Quantifying the heating potential of methanotrophs for application within decentralised wastewater treatment technologies
Aim 2.0	To determine experimentally whether the second law of thermodynamics can be used to explain differences in growth kinetics between species of methanotrophs.
Objective 2.1	To develop an experimental methodology for monitoring heat dissipation and growth kinetics <i>in situ</i> from pure cultures of <i>Methylomonas methanica</i> S1 and <i>Methylosinus trichosporium</i> OB3b.
Objective 2.2	To highlight any differences in growth dynamics between <i>Methylomonas methanica</i> S1 and <i>Methylosinus trichosporium</i> OB3b.
Objective 2.3	To highlight any differences in the heat dissipated between <i>Methylomonas methanica</i> S1 and <i>Methylosinus trichosporium</i> OB3b.
Objective 2.4	To explore correlations between heat dissipation and growth dynamics.

1.3 Thesis structure

- **Chapter 2 - Literature review:** Firstly, historical and current approaches to treating wastewater are laid out. The challenges facing wastewater treatment technologies are summarised and the need for novel approaches and more robust theory is highlighted. Attention is then given to theoretical community ecology. The role of mathematical models within this context is provided and important examples highlighting development within the field are discussed. An emphasis is then placed on neutral theory and the application of this theory within wastewater engineering. The literature review then pivots to describe how microbial growth is described from a thermodynamic viewpoint. This section demonstrates theoretical relationships between growth parameters and thermodynamics concepts such as entropy, enthalpy, Gibbs free energy and growth yields. The importance of heat dissipation and experimental methods used for measuring this quantity are reviewed. Finally, a short description of methanotrophic bacteria is provided.
- **Chapter 3 - Identifying neutral models on the basis of species abundance time series:** This chapter offers a comprehensive investigation into how sampling single-species relative abundance time series influences the ability to calibrate neutral model parameters, and thus the identifiability of these parameters under differing scenarios and datasets. Within this study two common neutral models are assessed; Hubbell's purely neutral model and Sloan et al.'s near neutral model. Structural identifiability is firstly examined by calibrating model parameters using idealised datasets, where the time series calibrated against are produced over a vast period and with maximum temporal resolution. Relaxation of fitting conditions and sampling regimes is then made to assess the practical identifiability of these parameters. Processes that offer relative improvement to model calibration are highlighted. Discussion primarily revolves around the role of effective parameters and how they can be used within the neutral model framework. The case is made for employing neutral models that include effective parameters as a generalised framework for characterising single-species relative abundance time series.
- **Chapter 4 - Quantifying the heating potential of methanotrophs for application within decentralised wastewater treatment technologies:** In this chapter a thermodynamic analysis of the batch growth of two species of methanotrophs, *Methylomonas methanica* S1 and *Methylosinus trichosporium* OB3b, is achieved by isothermal calorimetry. The quantity and rate of heat dissipation is measured for three different initial concentrations of methane (8%, 4% and 2%v/v in air). The headspace concentrations of methane and carbon dioxide, system pressure, pH, dissolved oxygen, optical density and non-purgeable organic carbon are measured through external growth experiments under the same growth conditions. By coupling the heat dissipated with how each species partitions carbon, a discussion is provided on whether the respective heat dissipation is a driver of growth kinetics and thus community structure.
- **Chapter 5 - Conclusion and future perspectives:** The final chapter provides a discussion summarising the primary conclusions of both research chapters. While more specific conclusions are presented within the respective research chapters themselves, here the broader importance these outcomes have regarding progress towards a general theory for use within wastewater treatment is discussed. Areas of improvement regarding the research undertaken within this thesis are identified and possible critiques are addressed pre-emptively. Avenues for continued research are identified alongside their possible outcomes, highlighting the future impact of the research presented within this thesis.

Chapter 2

Literature review

2.1 Wastewater engineering

2.1.1 A brief history

The development of engineering solutions to wastewater management is commonly traced back to examples from the Mesopotamian Empire which existed over 5,000 years ago (3500 – 2500 B.C.) (P. Cooper, 2005; De Feo et al., 2014; Lofrano & Brown, 2010). Although humans have existed for a far greater period, the lack of records make evaluating wastewater practices before this time difficult to achieve. However, it is commonly accepted that before permanent settlements were established disposal of waste was largely left to natural cycles. (Angelakis et al., 2023; Lofrano & Brown, 2010). Therefore transition to early settlements was likely plagued by the lack of waste management and a new relationship between waste and society needed to be adopted (Vuorinen et al., 2007).

In ancient urban settlements, examples of wastewater management can be found, for example, in the village of Skara Brae in Scotland (3100 - 2500 B.C.) (Childe, 1931; I. A. Simpson et al., 2006) and throughout the Minoan civilisation on Crete (3200 – 1100 B.C.) (Angelakis et al., 2013). The importance placed on developing sewage and drainage networks demonstrates the newfound veneration for water management. Indeed, in many ancient civilisations, deities of water were widely revered and worshipped (Angelakis et al., 2012). Perhaps the most striking example of wastewater infrastructure during this period is the drainage system found in Mohenjo-Daro; a city of Chalcolithic culture situated near the Indus River, which existed between 3300 – 2700 B.C. (Webster, 1962). Almost all houses in that city were equipped with a bathroom and flushable latrine and, in turn, were connected to a centralised sewage network. The most remarkable aspect of the sewer network, is the use of brick-lined sumps or cesspools, some which were found to be as large as 5ft wide, 7ft long and 6ft deep. The large size and expected low-flowrate would have allowed for solid sedimentation to occur. Therefore, it is thought that these cesspools may have been used as an early form of primary treatment (Jansen, 1989; Webster, 1962).

The grandest architectural achievements in water management were made during the Roman period. Large canal and aqueduct systems were a prominent feature of Roman hydraulic infrastructure (Rogers, 2018). The longest of these systems was built in the 4th and 5th century to ensure water supply into the city of Constantinople. The entire length of this system has been calculated to be at least 426km long (Sürmelihindi et al., 2021; Ward et al., 2017). While most of the systems implemented by Romans for managing wastewater were not new, they demonstrated increasing mastery of these

technologies, developing grand drainage systems such as the Cloaca Maxima in Rome (De Feo et al., 2014). It is also important to note that the Roman period marked advances within the governance of water resources. Romans adopted a systematic legislative approach to water. Private and public water rights were written into law and continuities within policy were observed across the empire (Bannon, 2017).

The collapse of the Roman Empire ushered in the “sanitary dark age”; a period that lasted until the 19th century (P. Cooper, 2005; Lofrano & Brown, 2010). As Roman and ancient infrastructure degraded, so did the use of these technologies, and regression in wastewater management was observed in Europe throughout the Early Medieval Age. Waste was largely disposed of by simply throwing it out the window or dumping it in the street, open defecation was a normality of life (Angelakis et al., 2023; P. Cooper, 2005). Rivers flowing through large cities were effectively open sewers and outbreaks of plague and disease, directly or indirectly driven by poor sanitation, were catastrophic for the population (Robb et al., 2021). Moving into the High and Late Middle Ages some effort was made to curb bad sanitation practices. In European cities, cesspools and latrines were ordered to be built and regulation combatting the improper dumping of waste was introduced (Ciecieznski, 2013). Improving sanitary conditions became a community endeavour (Jørgensen, 2008). Perhaps the most forward-thinking solution employed within this period was the formation of “sewage farms”, which combined waste disposal with agriculture. This strategy was first implemented in Bunzlau in 1531 and then in Edinburgh in 1650, and became a favourable solution for disposing of wastewater from cesspools within cities (Angelakis et al., 2023). Modern equivalents of this process still exist today (Mishra et al., 2023).

In the wake of the Industrial Revolution, increased urbanisation, rising mortality rates and outbreaks of cholera as well as other diseases, Europe’s major cities were in desperate need of sanitation reform. The need to transform outdated sewage systems was pushed by sanitarians such as Sir Edwin Chadwick, culminating in legislation such as The Public Health Act of 1848 (Aiello et al., 2008; Gill, 2000). As a result, large infrastructure projects were undertaken. In 1842 fire devastated the city of Hamburg, upon rebuilding the city, an extensive centralised sewer network was constructed. This marks the first of such networks to be introduced into modern European cities and signifies a pivotal shift toward active design of these systems using empirically driven calculations (Burian & Edwards, 2002; Seeder, 1999). Major improvements to Paris’s sewer network were made starting in 1854 and construction of the Bazalgette London sewer system was started in 1858 (Lofrano & Brown, 2010). Most of these sewer networks operated on the rational that *“the solution to pollution is dilution”* (P. Cooper, 2005). It was not until later in the century that more scientific treatment technologies were developed.

Towards the end of the 19th century a vast range of sewage treatment methods were developed. Between 1865 and 1875 over 400 patents regarding wastewater treatment technologies were registered (Wardle, 1893). A review of the most popular of these methods at the time was written by Sir Thomas Wardle in his book entitled *“On sewage treatment and disposal: for cities, towns, villages, private dwellings, and public institutions”*, published in 1893 (Wardle, 1893). Wardle highlights that most processes developed were various iterations of lime treatment, designed to enhance the precipitation of solids in sewage through the addition of calcium compounds. From a biological viewpoint, the most interesting method discussed is the Amines process. This process identifies microorganisms as the primary pollutant and aimed to sterilise sewage using trimethylamine. Developments during this period were not solely tied to treatment technologies. Advancements in the analysis of network design were also made (Ramalingam et al., 2012). This development of wastewater treatment methods, combined with a better understanding of sewage composition and the associated risks, prompted the

construction of wastewater treatment plants (WWTPs). The necessity for these plants was compounded by the increasing scarcity of farmland, a consequence of expanding urbanisation. As a result, many "sewage farms" were replaced by WWTPs (Angelakis et al., 2023; Seeder, 1999). The first septic tanks were also designed during this period (Cotteral & Norris, 1969).

Throughout the 20th century the advancement of environmental science and analytical chemistry prompted a new outlook on waste practices, one primarily driven by environmental considerations (Lofrano & Brown, 2010). The amount of environmental legislation introduced reflects this shift in societal concern (Shifrin, 2005). Some of the most universally relevant developments made during the beginning of the 20th century are summarised in the final report issued by the Royal Commission on Sewage Disposal (Royal Commission on Sewage Disposal, 1915). The conception of standard methods based on total suspended solids (TSS) and biochemical oxygen demand (BOD) is often credited to the eighth report produced by this commission (A., 1913). These conceptual breakthroughs paved the way for subsequent technological advancements, indeed, many of the primary and secondary treatments used today were developed within the latter half of the 20th century (Lofrano & Brown, 2010). The formation of the Activated Sludge Process as well as the biological removal of nitrogen and phosphorus through anoxic and aerobic zones are an examples of this (Barnard, 1973, 1974, 1975; Sawyer, 1965). Methods for optimising and analysing water distribution networks were also advanced, largely catalysed by the birth of the computer (Ramalingam et al., 2012). Many of the centralised WWTPs built within the 20th century are still in use today (Makisha & Gulshin, 2024).

2.1.2 Current wastewater treatment technologies

The role of current wastewater treatment infrastructure is best summarised by the targets set within the UN Sustainable Development Goal 6 (United Nations, 2015). The recent recognition of *"The human right to a clean, healthy and sustainable environment"* (United Nations, 2022a) sets the context in which these targets should be achieved. Wastewater treatment technologies are thus tasked with removing a variety of constituents, such as, suspended solids, biodegradable organics, pathogens, nutrients, priority pollutants, refractory organics, heavy metals, dissolved inorganics, pharmaceuticals, microplastics and any emerging contaminants (Khasawneh & Palaniandy, 2021; W. Liu et al., 2021; Metcalf & Eddy, 2003; Sauvé & Desrosiers, 2014; J. Wang & Wang, 2016). The accumulation of these pollutants within the environment are known to cause toxicity, antimicrobial resistance, facilitate the spread of disease, deplete natural oxygen and increase nutrient loading (Corcoran et al., 2010; Johnson et al., 2020; J. Li et al., 2018; Metcalf & Eddy, 2003; Mutuku et al., 2022; Schwarzenbach et al., 2006). Assessment of treatment effectiveness is primarily achieved through the observation of quantities such as biochemical oxygen demand (BOD), chemical oxygen demand (COD), total suspended solids (TSS), total nitrogen (TN) and total phosphorus (TP) (Metcalf & Eddy, 2003). Current practices for treating wastewater employ two forms of infrastructure, centralised and decentralised. Centralised wastewater treatment can be defined as a network of pipes collecting wastewater from multiple sources and transporting it to a large treatment facility some distance away. Decentralised systems are of smaller-scale and treat wastewater much closer to the source (Capodaglio, 2017; Wilderer & Schreff, 2000).

Septic tanks are the most commonly installed form of decentralised wastewater treatment (P. J. Withers et al., 2014). The first septic tanks were established in Europe around 1870 and introduced in the United States in 1884 (Cotteral & Norris, 1969). By the 1980s nearly one third of all housing units in the US were equipped with a septic tank system. This

equates to approximately 17 million septic tanks with a combined annual discharge of 800 billion gallons of wastewater (Canter & Knox, 1985). Since this time, the design of septic tank systems has largely remained unchanged. Indeed, current building standards still cite design practices from 1983 (Building Standards Division Directorate for Housing, 2024). According to these practices, septic tanks should utilise two separate compartments, either by partitioning a single tank or arranging two tanks in series, where the overall capacity should be split two thirds to one third, as shown in Figure 2.1 (BSI, 1983). For most septic tank systems, an accompanying soil absorption field or further treatment technology is required (BSI, 1983; Canter & Knox, 1985), however, it is not uncommon to find septic tanks discharging directly to surface water (Richards et al., 2016; P. J. A. Withers et al., 2011, 2012). More recently, designs incorporating a vertical flow direction (Moussavi et al., 2010), microalgae (Sorenson et al., 2023), solar-powered heating (Connelly et al., 2019), or a combination of these (Koottatep et al., 2020) have been proposed.

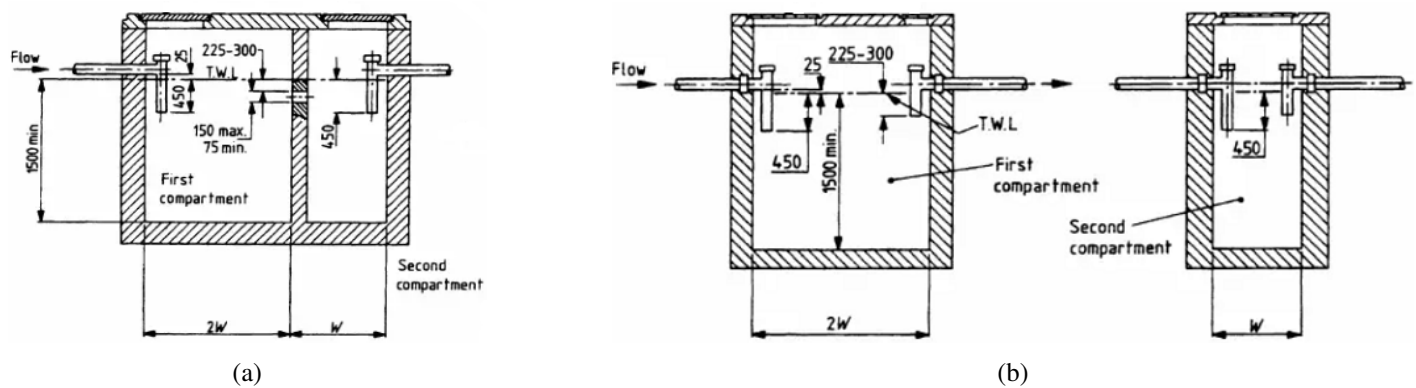


Figure 2.1: Schematics of typical septic tanks with two compartments, for populations of up to 30. Reproduced from BSI (1983). a) combined, b) separate.

The physical treatment process employed by septic tanks is the settlement of wastewater, allowing for the separation of heavier solids (sludge) and lighter particles (scum) to take place. Effective use of baffles and positioning of effluent pipes allows solids and scum to be retained within the tank rather than passed within the effluent (Canter & Knox, 1985). Biological treatment is also undergone within these systems, in the form of anaerobic digestion. This involves hydrolysis, acidogenesis, acetogenesis, and methanogenesis undertaken by microbes within the tank (Meegoda et al., 2018). Removal efficiencies of BOD, COD and TSS by septic tanks have been shown to be as high as 50%, however, nitrogen and phosphorus removal efficiencies are considerably lower. Removal efficiencies are shown to increase across all these quantities when assessing effluent from septic tanks that utilise soil absorption systems, demonstrating the importance these systems have in facilitating the removal of contaminants (Angelakis et al., 2023; Beal et al., 2005; Canter & Knox, 1985; Costa et al., 2002; Dasgupta et al., 2021; Moonkawin et al., 2023; Philippi et al., 1999).

Despite their relatively simple design, recent data shows that decentralised systems are still used extensively throughout the world. According to European census data (Appendix A), the percentage of population connected to independent wastewater treatment is high in many countries. For example, in Croatia, Ireland, Slovenia, Bulgaria and France, these percentages are reported to be 39, 31, 29, 25 and 18% respectively. In the US it was shown that in 2018 approximately a third of new housing relied on on-site wastewater treatment (NESC, 2020), and that currently around 20% of the population is not connected to a public sewer network (Yu et al., 2024). In emerging economies, septic tank systems provide a cost effective and economically viable option for improving sanitation (Hutton et al., 2013; Van Minh & Hung, 2011).

Centralised systems are more complex than their decentralised counterparts. These are commonly deployed within urban areas and have been considered the preferred form of wastewater infrastructure until recently (Pasciucco et al., 2022). The WWTPs associated with these systems employ a variety of treatment processes that can be categorised as preliminary, primary, secondary, tertiary or advanced (Fernandes et al., 2024; Metcalf & Eddy, 2003). Preliminary treatment is tasked with removing larger objects that may damage or cause problems in further treatment processes (Metcalf & Eddy, 2003). The first technology employed in this regard is screens. Typically coarse screens with a spacing of 50mm are used, as fine screens can plug due to a build up of material on the mesh causing head loss. Head loss within the screening process is typically ensured to remain below 15cm; as excessive head loss will lead to poor retention of screened objects. This is achieved through monitoring, cleaning and adjustment of flow rate (Verma, 2023). Following this, grit chambers can be employed. These systems are designed to remove large particles either through settlement, aeration or mechanical means (Fernandes et al., 2024). Although the aim of these systems is solely to remove larger objects, a small reduction in COD, suspended solids, nitrogen and phosphorus has been observed (L. He et al., 2022).

Primary treatment in centralised WWTPs acts to remove constituents by allowing particles to separate based on their density. This is usually achieved in a clarification tank where oils and grease can be skimmed off the top while heavier solids settle at the bottom (Ahammad et al., 2020). Gravity alone can facilitate the removal of 40 to 60% of the BOD and suspended solids respectively. Implementing coagulation or flocculation in aid, can bring the removal of suspended solids towards 90% (Fernandes et al., 2024; Sukmana et al., 2021; Verma, 2023). Despite these high removal efficiencies primary effluent can still be assumed to contain the majority of its organic constituents, these remain in the form of colloidal or dissolved particles and are thus not easily removed through purely physical processes. Primary clarification is therefore largely used as a precursor to the biological solutions adopted within secondary treatment technologies (Verma, 2023).

The most common form of secondary treatment is known as The Activated Sludge Process (ASP). This process of aerating wastewater leading to the propagation of suspended flocs, can be traced back to the beginning of the 20th century (Arden & Lockett, 1914). The primary mechanism underpinning this process is the promotion of aerobic bacterial growth, a biological process that was not understood until more than a decade after the technology's implementation (Buswell & Long, 1923). The growth of aerobic bacteria not only acts to degrade organic solids but also supports the formation of aggregates (flocs), which accelerates sedimentation. Indeed, it is common practice to install a secondary clarifier after the aeration chamber to allow for this to take place. Early milestones demonstrate that various configurations of ASP technologies can be implemented (Sawyer, 1965). Alongside the aerobic chamber is common to provide an anoxic compartment as these conditions facilitate nitrification and denitrification respectively (Barnard, 1973). Another important process is the returning of activated sludge to the initial aeration chamber, as this acts to maintain a high volume of active bacteria within the treatment process (Fernandes et al., 2024). To remove the need for physically returning sludge, sequencing batch reactors (SBRs) were developed. These reactors condense the ASP process within a single tank. The cycling of pumping, aeration and idling phases allows for these reactors to implement the key steps of the ASP process (A. Singh et al., 2022).

Other biological secondary treatment systems such as membrane bioreactors and attached growth systems can also be utilised within centralised WWTPs (Fernandes et al., 2024). Membrane bioreactors (MBRs) operate similarly to ASP and SBR technologies, as in they represent a suspended growth system. Where they differ is the introduction of physical

barriers in the form of microfiltration (MF) and ultrafiltration (UF) membranes (Asante-Sackey et al., 2022). These systems have been shown to remove pharmaceuticals and personal care products (PPCPs) from wastewater with an efficiency of between 41 and 95% (Y. Wang et al., 2018) and a removal rate of over 90% for BOD and COD (Hoinkis et al., 2012). Attached growth systems utilise the formation of biofilms to treat wastewater. The oldest example of these is the trickling filter in which wastewater is sprayed onto a filter bed to promote the formation of biofilm and thus biological treatment (Ahammad et al., 2020; Murshid et al., 2023; Verma, 2023). More complex iterations of this technology exist and for most of these systems removal efficiencies are shown to be greater than 80% (Murshid et al., 2023).

Tertiary and advanced treatment technologies are largely concerned with disinfection of wastewater and is thus implemented in response to the end-use of effluent (Metcalf & Eddy, 2003). Processes included are chlorination and ozonation, ultraviolet disinfection, ultrasound treatment, the Fenton Process, Electrochemical Advanced Oxidation Processes (EAOPs) and absorption (Ahammad et al., 2020; Fernandes et al., 2024; Verma, 2023). These can often be hybridised or integrated with other treatment technologies to increase application and compensate for individual process limitations (Asheghmoalla & Mehrvar, 2024; Fernandes et al., 2024; Su et al., 2016; Sudhakaran et al., 2013). Further treatment of sludge is also undertaken within WWTPs (Joe Cyril & Pawar, 2024). Anaerobic digestion is the primary form of sludge treatment and represents the sole energy-positive process utilised within common WWTPs (Jenicek et al., 2012). This is a consequence of the production of biogas, which can be used to offset energy consumption (Shen et al., 2015).

2.1.3 The microbial community

Microbial communities have been known to drive contemporary wastewater treatment processes for just over a century (Buswell & Long, 1923; Sawyer, 1965), however, quantification and characterisation of these organisms *in situ*, only became possible with the more recent advent of polymerase chain reaction (PCR) techniques and next-generation sequencing (Sanz & Köchling, 2007; Wagner et al., 2002). Research utilising quantitative polymerase chain reaction (qPCR) on 16S rRNA gene segments has been revolutionary in building a picture of the microbial communities at heart of wastewater treatment technologies (V. Singh & Dey, 2024). Indeed, by applying the aforementioned methods to samples taken from WWTPs it is shown that the process of ammonia oxidation, a key step of nitrification, within ASP technologies is dominated by chemolithotrophic ammonia-oxidising bacteria (AOB); specifically *Nitrosomonas* (H. D. Park et al., 2006; Purkhold et al., 2000; V. Singh & Dey, 2024; Zhang et al., 2011). Regarding the second step of nitrification, *Nitrosipira* have been shown to be the most diverse and ubiquitous nitrite-oxidising bacteria (NOB) (Lücker et al., 2010). Additionally, denitrifiers related to *Azoarcus*, *Zoogloea* and *Thauera* have been identified within activated sludge (H. Chen et al., 2020; Thomsen et al., 2007). Phosphorus-accumulating organisms (PAOs) from *Tetrasphaera* and *Accumulibacter*, as well as, filamentous bacteria such as *Microthrix* and *Gordonia* are also known to be present, along with many other genera (H. Chen et al., 2020; Mielczarek et al., 2012; V. Singh & Dey, 2024).

Within the anaerobic digestion process it has been shown that a "core group" of *Bacteroidetes*, *Chloroflexi*, *Proteobacteria* and *Synergistetes* exist across different anaerobic digesters (Rivière et al., 2009). Members of these phylum have been found within septic tank sludge and effluent samples, suggesting that commonality regarding the structuring of microbial communities may exist across different treatment scales (Connelly et al., 2019; Tomaras et al., 2009). On top of this, functional groups such as hydrolysers, acidogens, acetogens, and methanogens have been linked to system function-

ality and effectiveness (L. N. Nguyen et al., 2019). Indeed, it has been suggested that improved septic tank performance, due to increased temperature, may be linked to a more metabolically balanced methanogenic consortia (Connelly et al., 2019). Many other such relationships between functionality and microbial community composition within anaerobic digesters have been demonstrated (Kundu et al., 2014; Palarari et al., 2021; Regueiro et al., 2012).

This fingerprinting of microbial communities demonstrates a dramatic increase in our understanding of the biological processes underpinning current wastewater treatment technologies. As such, it has paved the way for studies investigating how operational parameters influence the overall microbial community or specific genera of interest (Y. M. Kim et al., 2011; Seuntjens et al., 2018). This, in turn, has produced practical solutions, such as the implementation of bioaugmentation strategies (Muter, 2023; Sanz & Köchling, 2007). A consensus is developing that wastewater treatment systems, previously monitored through process parameters and empirical data, should now be maintained through management of microbial communities (Carballa et al., 2015; Connelly et al., 2019; T. P. Curtis et al., 2003; Vanwonterghem et al., 2014).

2.1.4 Challenges

It is estimated that currently between 37 and 48% of all wastewater is discharged untreated (Jones et al., 2021). This discharging of untreated wastewater is one of the primary drivers of water degradation within surface water and groundwater resources (Bojanowski et al., 2022; Canter & Knox, 1985; du Plessis, 2022; Naidoo & Olaniran, 2014; United Nations, 2022b; van Wezel et al., 2018). The anticipated rise in water scarcity (Boretti & Rosa, 2019; WWAP, 2018), underscores the global importance of increasing the proportion of treated wastewater, in addition to the need to prevent, mitigate, and anticipate potential system failures. Indeed, as water resources become more valuable, contamination as a result of untreated wastewater discharge is of increasing concern (United Nations, 2022b). Thus, the challenges facing wastewater treatment must be identified and incorporated into long-term infrastructure planning strategies (United Nations Environment Programme, 2023).

Climate change is one of the greatest challenges of our time, and just as it is expected to challenge multiple, if not all, aspects of society, direct concerns regarding the effect physical climate change impacts, such as, increased air temperature, changing precipitation patterns, extreme weather events and a rising sea level, will have on wastewater treatment infrastructure have been expressed (Major et al., 2011; Vicuña et al., 2018). The consequence of these phenomena vary. For example, treatment efficiencies within decentralised wastewater treatment technologies have been shown to increase with higher temperatures (Connelly et al., 2019; Pussayanavin et al., 2015), however, lower levels of BOD may limit heterotrophic denitrification within associated soil absorption systems, and a reduction in O_2 availability, as a result of increased soil temperature, is expected to diminish phosphorus removal capacity (J. A. Cooper et al., 2016). In centralised WWTPs there are concerns that global warming could lead to a breakdown within enhanced biological phosphorus removal (EBPR) as a result of a change in microbial community composition (Lopez-Vazquez et al., 2009; Qiu et al., 2022). The increase in frequency and severity of extreme weather events, such as high temperatures and heavy precipitation, as a result of climate change is also a cause for concern (Goodess, 2013). Flooding as a result of extreme storms and precipitation is known to facilitate the discharge of untreated wastewater (J. Li et al., 2023; Nazari et al., 2022). On the opposite side of the spectrum, high temperatures leading to drought will lower dilution capacities of lotic systems, increasing the relative pollution due to wastewater effluent (Benotti et al., 2010; Marino et al., 2024). Extreme weather events are also

known to damage wastewater treatment infrastructure and thus climate change will invoke structural implications on these systems (J. Hughes et al., 2021; Hyde-Smith et al., 2022; Valdez et al., 2010).

Another fundamental challenge facing wastewater infrastructure is the expected increase in global population, driving increases in wastewater production (United Nations Environment Programme, 2023). It has been projected that under all Shared Socioeconomic Pathways (SSPs), the global population will continue to increase until 2050, with the worst-case scenario demonstrating a global population of approximately 10 billion at this point (Kc & Lutz, 2017). As a result it is expected that by 2050 wastewater production will have grown by 51% from current levels, resulting in global production of over half a trillion m^3 /year (Qadir et al., 2020). This value can be broken down by looking at per capita production. The global average per capita production rate has been estimated at $49m^3$ /year, however, for the UK this value is significantly higher at $92m^3$ /year. Indeed, per capita production rates are shown to vary drastically between regions, with Sub-Saharan Africa demonstrating the lowest value at just $11m^3$ /year, while North America has the highest per capita production rate of $210m^3$ /year (Jones et al., 2021). While regions with higher economic classification express a higher production of wastewater, the proportion of wastewater sent to treatment is also greater (Jones et al., 2021; Su et al., 2016). It has however been observed that during economic development within emerging countries a lag in infrastructure development has led to an increase in the discharge of untreated wastewater (Flörke et al., 2013). Thus, wastewater infrastructure should be at the fore of economic development. To elevate the strain on treatment capacity as a result of increased wastewater production, strategies for decreasing per capita rates should be explored. In China a reduction in wastewater production has been achieved through the implementation of water-saving policies (Xu et al., 2020).

The demand increase associated with a growing population will place greater pressure on effective management and development of urban water infrastructure (United Nations Human Settlements Programme (UN-Habitat), 2023). Indeed, urbanisation is expected to intensify the challenges facing wastewater treatment. Depending on the SSP assessed, it is predicted that by 2100 the proportion of the world's population residing in an urban setting could be between 60 and 92%, demonstrating an increase from the current level of 50% (Jiang & O'Neill, 2017). This in turn is projected to increase the global urban area by over 100 and 200% respectively (X. Li et al., 2019). Wastewater treatment systems will therefore be forced to advance and expand if further pollution is to be avoided (Strokal et al., 2021; United Nations Human Settlements Programme (UN-Habitat), 2023). Indeed, in cities in India, current wastewater generation is already exceeding treatment capacity by 60 to 70% (Amerasinghe et al., 2013), and in Africa, urbanisation without the development of suitable wastewater infrastructure, has lead to a deterioration of human and environmental health (Boadi et al., 2005).

The best method to employ in tackling these challenges is not obviously clear. There is a growing consensus that rather than blindly following design principles developed within the 20th century (S. Graham & Marvin, 2001), the degree of wastewater centralisation should be tailored based on criteria such as, treatment efficiency, type of contaminants expected, capital and O&M costs, energy demand, social implications, robustness to environmental impact, settlement dispersion and topography (Eggimann et al., 2015; Libralato et al., 2012). The conversation regarding the respective roles of centralised and decentralised technologies is still ongoing (Duque et al., 2024; Pasciucco et al., 2022; Torre et al., 2021) and although integrated and hybrid solutions are being presented as attractive alternatives, the implementation of these strategies is still rare (Keller, 2023). There is also a shift to view wastewater treatment infrastructure as a solution rather than a problem (United Nations Environment Programme, 2023). Within a circular economy, WWTPs can be categorised as resource production hubs primarily due to their capacity for producing nutrients and energy (Ali et al.,

2022; Kretschmer et al., 2015; Qadir et al., 2020; Shen et al., 2015; van Loosdrecht & Brdjanovic, 2014; Verstraete et al., 2009). Additional products can be generated from the microbes themselves, such as, cellulose, extracellular polymeric substances (EPS), volatile fatty acids (VFAs) and single-cell protein (SCP) (Kehrein et al., 2020; Lin et al., 2010; Puyol et al., 2017). Effluent reuse is also being heralded as a vital strategy for addressing water scarcity (Angelakis & Snyder, 2015; Hristov et al., 2021; Mannina et al., 2022; Salgot & Folch, 2018).

Increasing reliance on wastewater infrastructure offers both opportunities and risks. While these technologies can provide valuable resources for promoting a circular economy, failures in treatment systems could result in more severe consequences, especially if treated effluent is to be recycled (Toze, 2006). Indeed, despite technological and theoretical advancements in wastewater treatment, failure of systems is commonly observed (Cossio et al., 2018; Liang & Yue, 2021; Murray & Drechsel, 2011; Preisner, 2020). Even treated effluent has been shown to be released at a rate and quality outside of accepted guidelines (Büttner et al., 2022; Charles et al., 2005). These failures can occur due to mechanical (Taheriyoun & Moradinejad, 2014; K. Wang et al., 2017), biological (Heine et al., 2002) and economic reasons (Murray & Drechsel, 2011), and although attempts have been made at developing models for predicting effluent quality and process failures, these are primarily data-driven reflecting an empirical approach (Afan et al., 2024; Duarte et al., 2024; Weirich et al., 2015). The lack of theoretical laws governing the structuring of microbial communities at heart of treatment processes make developing solutions from a fundamental perspective a significant challenge. Indeed, there have been calls to improve our theoretical basis if we are to better engineer wastewater treatment systems (T. P. Curtis et al., 2003; Prosser et al., 2007; Sloan & Gómez-Borraz, 2023). The research presented within this thesis is a direct attempt at addressing this challenge.

2.2 Theoretical community ecology

2.2.1 The role of mathematical models

At the heart of all biological wastewater treatment technologies is a mixed community of bacteria and so developing a mathematical model of the bacterial populations would appear to be central to predicting the function of wastewater systems. In the most elementary case, the size of a population results from a trade-off between births, deaths, immigration and emigration (Nisbet & Gurney, 1982). To formulate a model, one can portray such processes mathematically. For example, the previous statement can be represented by:

$$\Delta N = (B - D + I - E)\Delta t, \quad (2.1)$$

where ΔN represents the change in the number of individuals within a population over a change in time Δt . B , D , I and E respectively refer to the rate of births, deaths, immigration and emigration. Early deterministic models largely expanded on birth and death processes (Lotka, 1920a; Verhulst, 1838). Migration and dispersal were later brought to the fore, most famously by MacArthur and Wilson in their *"Theory of Island Biogeography"* (MacArthur & Wilson, 1967). The success and popularity of such models were dependent on their ability to describe observed patterns (Pearl & Reed, 1920). As a result, models formulated initially through the description of fundamental processes have been subsequently adapted for representing specific patterns, leading to a field rich in mathematical theory yet devoid of any overarching organisation (Vellend, 2010).

The abundance of mathematical models within theoretical ecology arises from the problem of scale. In the textbook *"Modelling Fluctuating Populations"* by Nisbet and Gurney, two categories of models are defined, "tactical" and "strategic" (Nisbet & Gurney, 1982, p. 1). Tactical models refer to those aimed at accurately portraying ecological processes at a low level. They usually account for complex interactions requiring precise knowledge of the system. Strategic models are mathematically simpler, using sweeping assumptions to describe trends more broadly. These definitions highlight that the scale at which a model is applied often determines its relative complexity. Indeed, there is a trade-off between complexity and scale such that formulating methods for transcribing between models of different scales has occupied ecologists for years (Jarvis & McNaughton, 1986; Levin, 1992; Martinez-Rabert et al., 2023; Parry & Bithell, 2012).

Taken to its conclusion, some scales must exist where the complexity of processes described must be curtailed. This is especially true for exceptionally large populations or communities such as those of microorganisms (Woodcock et al., 2007). It is therefore not always the case that the most realistic model is best. Levin (1992) states this explicitly, declaring *"A good model does not attempt to reproduce every detail of the biological system; the system itself suffices for that. Rather, the objective of a model should be to ask how much detail can be ignored without producing results that contradict specific sets of observations, on particular scales of interest"*. Under such a description, mathematical models can only be expected to provide circumstantial evidence of the processes that drive the observed pattern, to state assuredly would succumb to the *"same behaviour implies same mechanism fallacy"* (Woodcock et al., 2007). What Levin highlights is a pragmatic approach, where the aim of mathematical models is to simply explain the observed pattern rather than describe all of its underlying processes and drivers (Levin, 1992).

As a consequence of describing distinct patterns, there has been a failure to define a set of universal laws governing community ecology (Lawton, 1999; Vellend, 2010). Lawton expresses that only an indication of fundamental rules can be arrived upon when considering patterns from the majority of systems. Most systems are considered too complex, and thus universal dynamics are obscured by the swathe of local interactions. Despite this observation, he concedes that all ecological patterns must emerge from the truly universal laws of nature (Lawton, 1999). Vellend argues for the need of mathematical models to be stripped back towards these fundamental laws, despite their relative importance to the case at hand (Vellend, 2010; Vellend & Orrock, 2009). He demonstrates this point by highlighting that layering complexity within mathematical models, in an attempt to describe specific observed phenomena, quickly leads to the formulation of a vast number of models, all differing in their theoretical importance and with no organisational framework (Vellend, 2010).

Unlike Lawton, Vellend does not promote the physical laws of nature as pillars to coalesce theoretical ecology around, rather he draws his set of fundamental processes from population genetics (Vellend, 2010; Vellend & Orrock, 2009), a field developed with a relative lack of empirical data, yet, shares striking similarities to community ecology such that theoretical insights and mathematical formula are often transcribed between fields (Caswell, 1976; Hedrick, 2009; Hu et al., 2006; Millstein, 2013; Sloan et al., 2021). A characterisation of the early development of population genetics is given by Lewontin (1974, p. 189) *"For many years population genetics was an immensely rich and powerful theory with virtually no suitable facts on which to operate. It was like a complex and exquisite machine, designed to process a raw material that no one had succeeded in mining"*. Vellend argues that this lack of "raw material" enforced a philosophy of process over pattern, enabling the field to arrive at theoretical descriptions based on fundamental processes. Vellend adopts the same set of fundamental processes, albeit with slightly different definitions, for use within community ecology. These are drift, selection, speciation and dispersal (Vellend, 2010).

The viewpoints articulated by Levin and Vellend present a dichotomy in regard to the role of mathematical models, where one favours the description of observable patterns, while the other emphasises expression of universal processes. However, these perspectives are not mutually exclusive. To demonstrate this the analogy described by Sloan and Gómez-Borraz (2023) is drawn upon. What they conceptualise is the success of Bernoulli's equation in informing the design of water distribution systems, despite the availability of the more comprehensive Navier-Stokes model. Indeed, what Bernoulli achieved was a refinement of universal processes to a point complex enough to inform engineering design, yet simple enough to negate the inability to solve it. This outlook is adopted within this thesis. The view is taken that the formulation of an effective generalised theory for describing microbial communities must be able to describe observed patterns produced by these communities, whilst also accounting for universal processes, even if these processes must be simplified to their most parsimonious form.

2.2.2 Early models

The field of theoretical ecology can be traced back to Thomas Malthus' publication in 1798 entitled *"Essay on the Principle of Population"*, in which obstacles to the exponential growth, mathematically described by Euler in 1748, are recognised for living populations (Malthus, 1798). These ideas were further developed in subsequent editions and largely influenced the theory of evolution developed by Charles Darwin and Alfred Russell Wallace (Bacaër, 2011d). Despite the success of these ideas, no explicit mathematical descriptions of the limitations proposed by Malthus were given. This was achieved by Pierre-Francois Verhulst in 1838 (Verhulst, 1838).

The mathematical model developed by Verhulst was coined the logistic growth equation; a common manifestation of this is given by the form:

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K} \right), \quad (2.2)$$

where N is the number of individuals within a population at time t , r is the intrinsic growth rate and K is a carrying capacity. The logistic equation marks a pivotal moment in theoretical ecology as mathematical simplicity is achieved by characterising external pressures through "vital" coefficients (Scudo & Ziegler, 1978). Indeed, the expansion of these coefficients to explicitly represent certain phenomena was undertaken in later developments of the model, as reviewed by Scudo and Ziegler (1978). Application of the logistic equation was undertaken in the years after Verhulst's formulation, yet these were done in ignorance of each other leading to "rediscoveries" of the same model (Lloyd, 1967). Indeed, it wasn't until the "golden age" of the 1920s that Verhulst's work was truly appreciated (Scudo & Ziegler, 1978).

In 1920 Raymond Pearl and Lowell J. Reed published an article *"On the Rate of Growth of the Population of the United States since 1790 and Its Mathematical Representation"* (Pearl & Reed, 1920), prompting renewed interest in the use of the logistic equation for describing population growth (Cramer, 2004). Although the formulation of the logistic equation by Pearl and Reed was achieved independently, Pearl does draw focus to Verhulst's work in subsequent publications, stating explicitly that *"The equation to the curve which has been found by experiment and observation to be descriptive of population growth in a wide variety of organisms was first discovered by the Belgian mathematician, Verhulst in 1838"* (Pearl & Slobodkin, 1976).

Over the same period, Pearl drew inspiration from a publication by Alfred James Lotka on biological oscillations, and subsequently helped him secure a scholarship at Jon Hopkins University's School of Hygiene and Public Health (Bacaër, 2011c). Lotka's primary interest was in mathematically describing predator-prey interactions. Lotka applied his previous knowledge of chemistry to these systems. To his surprise, this resulted in a series of rhythmic relations stretching out through infinite oscillations (Lotka, 1920a). Lotka first published these findings in 1920 and during his time at Jon Hopkins University developed his ideas into a book entitled "*Elements of Physical Biology*", published in 1925 (Lotka, 1925).

Lotka's model describes the abundance of both predator and prey populations through a pair of differential equations. These can be represented by:

$$\frac{dx}{dt} = ax - bxy, \quad (2.3a)$$

$$\frac{dy}{dt} = -cy + dxy, \quad (2.3b)$$

where x is the total mass of prey at time t , y is the total mass of predators at time t , a is the growth rate of prey when there are no predators and c is the rate of decrease of predators when no prey is present. b and d describe the respective impacts that each species has on one another, where b impacts the death rate and d impacts the growth rate respectively. Values for a , b , c and d are all positive. Under these conditions, the oscillatory nature of the interaction between both species can be mathematically proven. This process is explicitly shown by Lotka in another publication from 1920 (Lotka, 1920b). It is important to note that over the same period, Vito Volterra independently utilised the same pair of differential equations for describing the interactions between predator and prey in fisheries (Volterra, 1926a, 1926b). As such, the predator-prey model, defined in Equation 2.3, is commonly referred to as the Lotka-Volterra model.

The Lotka-Volterra model is significant as it provided a mathematical framework for analysing populations of multiple species. Indeed, a general case for n species was developed alongside mathematical explanations of food webs generated between numerous species (Scudo, 1971). The Lotka-Volterra model was also a necessary step in the formulation of the competitive exclusion principle defined by Gause in 1934 (Gause, 1934). Gause draws reference to the work undertaken by Lotka and Volterra throughout his book, and an early example of the exclusion principle was mathematically demonstrate by Volterra in 1927 (Volterra, 1927). A description of this is provided by Scudo (1971).

Next, focus is given to Sewall Wright's description of random genetic drift, an antecedent of the neutral models examined throughout much of this thesis. Sewall Wright first published this description in 1931 (Wright, 1931). This mathematical formulation followed from the large body of mathematical development within population genetics throughout the early 20th century; most relevant are works by Hardy (1908), Weinberg (1908) and Fisher (1923, 1930). In its simplest form it describes the change in frequency of an allele within a Mendelian population with no migration, selection or speciation, where generations are non-overlapping such that births and deaths occur simultaneously and the population is of a finite size. Under such a regime the allele type chosen for a subsequent generation is undertaken through random binomial sampling of the current pool of allele types (Bacaër, 2011e). An allele, short for allelomorph (allelo = of each other, morph = form), refers to a specific version of a gene that influences the expression of certain traits. Here it is most helpful to think of this as a fundamental unit of hereditary variation.

Examining the case where there are only two types of alleles (a and A) for each individual, and the abundance of allele A is given as i . The probability that allele A is chosen at random is defined as $\frac{i}{2N}$, where N is the number of individuals within the population. The probability that allele a is chosen is therefore $\frac{2N-i}{2N}$ which simplifies to $1 - \frac{i}{2N}$. In the case where a future generation contains a single increase in the abundance of allele A , the probability that allele A is chosen in a further birth is increased by $\frac{1}{2N}$. This demonstrates that the magnitude of probability increase is directly linked to the size of a population N . Indeed, in the case where a population is indefinitely large such that N approaches infinity then $\frac{1}{2N} \rightarrow 0$. In such a population it can therefore be assumed that the frequency of alleles remains constant through time, this is known as the Hardy-Weinberg Equilibrium (Bacaër, 2011b). Where N is of a fixed size, the increase in selection probability, attained through the random selection of the allele, will inevitably lead to fixation when $i = 2N$. This occurs as the selection probability is tied to allele abundance and thus an increase in one will lead to an increase in the other. Similarly the opposite trend is true and thus fixation can also occur at $i = 0$, where the allele is lost entirely. This tendency of the frequency of alleles to meander based solely on random mating is what is known as "genetic drift" (Bacaër, 2011e). Methods for assessing when the loss of genetic diversity may occur based on drift and the role of the population size N , more commonly referred to as the effective population size N_e , are still central in the field today (Charlesworth, 2009; J. Wang et al., 2016).

2.2.3 Neutral theory

Neutral theory was established towards the end of the 1960s in opposition to the Darwinian description of genetic evolution. A summary of the Darwinian approach at this time is given by G. G. Simpson (1964). The argument follows that if there is control over the molecular level, as stated by Weiss (1962), then if selection pressures exist at the individual or community level it must also exist at the molecular level. Mayr echoes this viewpoint, stating that the evolution of proteins is subject to constant selective pressures (Bryson & Vogel, 1965). Despite the success of Darwinian thinking, Kimura demonstrates mathematically that at the genetic level selection would induce a rate of evolution so great as to be unreasonable. The only remedy is to assume that most mutations are neutral (Kimura, 1968). King and Jukes (1969) support Kimura's rationale and provide a fuller account of the position in their definition of non-Darwinian evolution. Indeed, these early descriptions of the role that neutral mutations play in genetic evolution mark the inception of a defined neutral theory. The theory was subsequently galvanised, with Kimura publishing two books on the topic entitled "*Theoretical Aspects of Population Genetics*" and "*The Neutral Theory of Molecular Evolution*" (Kimura, 1971, 1985).

The migration of neutral theory to community ecology occurred later. In 1976, Caswell applied the neutral model, originally developed in population genetics, to community ecology. He used this model to establish a reference scale and demonstrated its capacity to produce relative-abundance patterns and species-area curves that aligned with both theoretical and empirical expectations (Caswell, 1976). This contributed to the ongoing debate over the relative importance of deterministic versus stochastic processes within community structure at the time (Connell, 1978, 1980; Connor & Simberloff, 1979; Diamond, 1975; Sale, 1977; Simberloff, 1978). Further neutral models were shown to be successful in describing observed ecological patterns (Hubbell, 1979; R. Hughes, 1986; Innis & Haefner, 1980), yet, there was reluctance to concede the theoretical merits of the theory (Bell, 2000). The idea of neutral processes governing community structure persisted to seem aberrant to ecologists subscribed to the neo-Darwinian niche-assembly view of community structure (Chesson, 1991; Gilpin & Diamond, 1984; Roughgarden, 1983). This debate was expected to continue given the legitimacy of both viewpoints (Hubbell, 1997).

The most notable development of neutral theory within community ecology came at the turn of the millennium with Hubbell's landmark book entitled *"The Unified Neutral Theory of Biodiversity and Biogeography"* (Hubbell, 2001). In this book Hubbell expanded on his previous work (Hubbell, 1979, 1995, 1997; Hubbell & Foster, 1986) and aimed to unify theories of biogeography, such as MacArthur and Wilson's *"Theory of Island Biogeography"* (MacArthur & Wilson, 1967), with those of species abundance. A review of species abundance theory at the time is given by Tokeshi (1993, 1997). The result is an in-depth formulation of a mathematical framework describing the change in a species' abundance based on neutral processes.

Neutrality in Hubbell's theory is applied through functional equivalence. In its simplest form, the theory uses a zero-sum ecological drift model for an isolated community, in which only birth and death impact a community's structure. The transitional probabilities describing the outcomes of a single discrete replacement event, in which birth occurs subsequently to death, on a species' abundance are as follows:

$$Pr\{N_i - 1 | N_i\} = \left(\frac{N_i}{N_T} \right) \left(\frac{N_T - N_i}{N_T - 1} \right), \quad (2.4a)$$

$$Pr\{N_i | N_i\} = 1 - Pr\{N_i - 1 | N_i\} - Pr\{N_i + 1 | N_i\}, \quad (2.4b)$$

$$Pr\{N_i + 1 | N_i\} = \left(\frac{N_T - N_i}{N_T} \right) \left(\frac{N_i}{N_T - 1} \right), \quad (2.4c)$$

where N_i is the number of individuals in species i and N_T is total number of individuals in the community. The probabilities for species-specific birth and death events can be separated as:

$$Pr\{Death | N_i\} = \left(\frac{N_i}{N_T} \right), \quad (2.5a)$$

$$Pr\{Birth | N_i\} = \left(\frac{N_i}{N_T - 1} \right), \quad (2.5b)$$

This produces similar dynamics to the mathematics developed by Sewall Wright in his formation of genetic drift (Wright, 1931). Indeed, the probability of an allele, or in this case and individual, being chosen is dependent on the abundance of the quantity itself and the size of pool it exists within. If one were to simulate a community based on Equations 2.4a, 2.4b, and 2.4c over an infinite number of replacement events, fixation of a single species would occur at $N_i = N_T$, while extinction of all other species would occur at $N_i = 0$. This reflects the same behaviour as exhibited by Wright's model, and thus the random changes in a species relative abundance exhibited by this process can be referred to as "ecological drift".

To incorporate the role of migration on community structure, Hubbell expands on the previous mathematical description by including terms m and p_i , where m is the probability, or rate, of immigration and p_i is the relative abundance of the i^{th} species within the source community of migrants (Hubbell, 2001). In this case we can also think of the previously

defined death of an individual as either a local death or emigration from the community. Incorporating these parameters within the discrete transitional probabilities provided in Equations 2.4a, 2.4b and 2.4c we arrive at:

$$Pr\{N_i - 1|N_i\} = \left(\frac{N_i}{N_T}\right) \left[m(1 - p_i) + (1 - m) \left(\frac{N_T - N_i}{N_T - 1}\right) \right], \quad (2.6a)$$

$$Pr\{N_i|N_i\} = 1 - Pr\{N_i - 1|N_i\} - Pr\{N_i + 1|N_i\}, \quad (2.6b)$$

$$Pr\{N_i + 1|N_i\} = \left(\frac{N_T - N_i}{N_T}\right) \left[mp_i + (1 - m) \left(\frac{N_i}{N_T - 1}\right) \right], \quad (2.6c)$$

Following the publication of Hubbell's theory, an uptake in the application of neutral models transpired. Neutral theory was applied to species abundance patterns observed from various macro-organisms with varying degrees of success (Chisholm & Pacala, 2010; Condit et al., 2002; Dornelas et al., 2006; B. Gilbert et al., 2006; Latimer et al., 2005; McGill, 2003; Muneeppeerakul et al., 2008; Ulrich & Zalewski, 2007; Volkov et al., 2007). The theory's growing popularity unsurprisingly prompted rigorous critique, (Clark, 2009, 2012; Nee, 2005; Purves & Turnbull, 2010; Ricklefs, 2006), rekindling the debate observed in the previous millennium. Explicit examples of debate can be found between Bell (2001) and Enquist et al. (2002), Condit et al. (2002) and Duivenvoorden et al. (2002), Hubbell (2003) and Ricklefs (2003) and Clark and Volkov et al. (2004)

Most criticisms of neutral theory can be expressed through four primary concerns: the underlying assumptions, the attribution of processes to patterns, the fitting methods and parameter interpretation, and community size and associated timescales (Matthews & Whittaker, 2014). In addressing the first two concerns, Hubbell concedes that disregarding competition and niche differences is fanciful, and that multiple explanations can exist for a given observed pattern (Hubbell, 2005). What he draws focus to is that the success of a model, void of any competitive interactions or selection pressures, in describing observed species abundance patterns is inherently significant; a sentiment echoed by Volkov et al. (2007). Fitting methods and parameter interpretation are largely expanded on through further publications, such as those by Etienne and others which developed a sampling theory and fitting framework based on likelihood (Etienne, 2005, 2009; Etienne & Alonso, 2007; Etienne & Olff, 2004; Etienne et al., 2006). Finally, regarding community size and timescales, Ricklefs (2003) firstly demonstrates that speciation in Hubbell's model may lead to species lifespans at either end of the acceptable spectrum. Hubbell admits that these can be viewed as theoretical extremes and offers a third method of speciation in direct response (Hubbell, 2003). Despite further advancements in how speciation is handled within neutral models, this still presents an issue within the theory (Chisholm & O'Dwyer, 2014). Ricklefs (2006) and Nee (2005) also demonstrate a lack of importance of ecological drift in large communities, illustrated by the excessively long timescales needed to observe any meaningful change. This has been recently challenged through the suggestion of an "effective community size" (Sloan et al., 2021), an approach employed successfully within population genetics for many years (Charlesworth, 2009). Despite these criticisms, there is a consensus that Hubbell's theory has merit (Alonso et al., 2006; Rosindell et al., 2012), and as time passed ecological drift was identified as a fundamental driver of community assembly (Vellend, 2010). As such, efforts have been made to bridge niche and neutral perspectives (Adler et al., 2007; Leibold & McPeck, 2006; Stokes & Archer, 2010; Tilman, 2004).

Further development of Hubbell's model has occurred since its inception. Developments include the expansion of speciation modes (de Aguiar et al., 2009; Etienne & Haegeman, 2011; Etienne et al., 2007; Haegeman & Etienne, 2009, 2010; Hubbell, 2003; Kopp, 2010; Rosindell et al., 2010) and the traversal from a spatially implicit model to spatially explicit ones (Chave & Leigh, 2002; Etienne & Rosindell, 2011; O'Dwyer & Green, 2010; Rosindell & Cornell, 2007). The zero-sum assumption has also been relaxed (Haegeman & Etienne, 2008). These developments are reviewed in an assessment of the decade succeeding the initial publication of *"The Unified Neutral Theory of Biodiversity and Biogeography"* (Rosindell et al., 2011). More recent developments include, the incorporation of habitat and environmental considerations (Abeyasinghe & Punchi-Manage, 2020; Kalyuzhny et al., 2015), new analytical expressions for species ages (Chisholm & O'Dwyer, 2014) and an adaptation for assessing transient dynamics (Fung & Chisholm, 2023). Hubbell's model has also been modified for specific use with microbial communities (Ofiteru et al., 2010; Sloan et al., 2006, 2021; Woodcock et al., 2007). Numerous examples demonstrating its application in this context can be found in recent literature (Cira et al., 2018; Grilli, 2020; Heys et al., 2020; T. Kim et al., 2021; Z. Liu et al., 2019; Sun et al., 2021). Perhaps this enduring relevance and ongoing development presents the strongest testament to the success of neutral theory.

2.2.4 Application of neutral theory for wastewater engineering

Ecological theories, such as neutral theory, are applied to wastewater engineering through the context of microbial ecology. Advances in analytical techniques and molecular tools have prompted the vast accumulation of molecular data, driving a revolution in this field (Prosser, 2020; Prosser et al., 2007). The overall size, fast turnover and ease of sampling microbial communities, means high quality time series can be obtained with relative ease; a problem that has stifled the development of ecological theory when observing larger organisms (Magurran, 2007). However, despite the eagerness to observe these communities in newfound detail, theoretical analysis of datasets at the community level has been lacking and the importance of analysing microbial communities across all scales has been recently re-emphasised (J. A. Gilbert & Lynch, 2019; Lemke & DeSalle, 2023; Stubbendieck et al., 2016). Hampering this process is the nature of these communities, which makes useful application of neutral models to these communities a non-trivial endeavour. Meeting these challenges will not only enable the useful application of neutral models within wastewater engineering but also facilitate theoretical development that can be shared between macro and micro ecology (Barberán et al., 2014).

The most obvious challenge is that microbial communities are extremely large. The vast abundance and diversity of microbes is difficult to comprehend. The total number of prokaryotes globally has been estimated to be $4 - 6 \times 10^{30}$ (Whitman et al., 1998), while 4×10^6 different taxa could exist in a ton of soil (T. P. Curtis et al., 2002). For wastewater treatment systems, it has been estimated that the global activated sludge community holds up to 6×10^{23} individuals, made up from $1 - 2 \times 10^9$ different species (Wu et al., 2019). Even at small-scale one can expect a large amount of diversity, with 70 species expected per ml of wastewater (T. P. Curtis et al., 2002). This presents a problem in regard to the application of neutral theory, firstly, due to the arguments expressed by Nee (2005) and Ricklefs (2006), in which they demonstrate the decrease in importance of ecological drift as the community size increases. Indeed, the Hardy-Weinberg equilibrium, developed for population genetics, demonstrates that for an infinite population, genetic drift is diminished to a point where allele frequency remains constant (Bacaër, 2011a, 2011b).

The effect of undersampling is also a cause for concern when applying neutral theory to such large communities. Due to the extreme diversity, within samples it is likely that the few abundant taxa will dominate, thus concealing the much greater number of rare taxa (T. P. Curtis et al., 2002). Even if a full sample census can be obtained, the discrepancy between the sample size and community size is dramatically larger for microorganisms than what one might expect for macro-organisms, resulting in a scattered and under-represented analysis of diversity (Sloan et al., 2007; Woodcock et al., 2007). One avenue for avoiding this is to base analysis on a single detectable taxa or species. Indeed, Ofiteru et al. (2010) demonstrates that through the use of temporal data, neutral model parameters can be calibrated in absence of the full distribution of taxa. However, the exact impact temporal sampling regimes have on the calibration accuracy of neutral model parameters is still unknown.

Despite these challenges, the number of studies assessing the role that stochastic processes, such as ecological drift, have in the structuring of microbial communities has been increasing (Zhou & Ning, 2017). In many cases, this is achieved by calibrating neutral models to stationary or temporal abundance data. (Fodelianakis et al., 2021; Heys et al., 2020; Konopka et al., 2015; Logares et al., 2018; Ofiteru et al., 2010; Sloan et al., 2006, 2021; Venkataraman et al., 2015; Woodcock et al., 2007). This methodology has been directly applied to microbial communities from within wastewater treatment systems, most commonly to quantify the importance of stochastic versus deterministic mechanisms in driving community assembly from stationary species abundance distributions (Ayarza & Erijman, 2011; W. Chen et al., 2022; Cheng et al., 2019, 2024). While this approach can be useful in helping uncover parameters that may be used to control these communities, Ofiteru et al. (2010) demonstrates that through the application to temporal datasets we may be able to predict the extinction of key taxa or when their abundance might drop below a critical threshold. The study highlights the ability of neutral models to explain the change in relative abundance of the most abundant ammonia oxidising bacteria and heterotrophic bacteria obtained from aeration basins at the Palo Alto Regional Water Quality Control Plant. It is demonstrated that from this, one can empirically derive neutral model parameters that hold ecological meaning and are "biologically plausible".

The ability to derive parameters that deliver biological insights is a key step to move from a good model to a useful one (Wieland et al., 2021). Indeed, a recent study aiming to develop the theoretical framework behind the use of neutral models within microbial ecology has specifically focused on the calibration of model parameters using temporal datasets (Sloan et al., 2021). The outcome reached is that while derived parameters may be biologically plausible they are often at odds with what is measured independently and present a higher importance of ecological drift than what one might expect for large communities; this is made apparent through a significant underestimation of the size of the community. This mismatch in the theoretical and experimental importance of ecological drift within large communities is not an isolated case. In population genetics it well established that genetic drift is a key driver of evolution despite the great abundance of alleles within large populations (Nei et al., 2010). The solution was to adopt a parameter known as the "effective population size", initially conceptualised by Sewall Wright in 1931 (Wright, 1931). The effect this parameter has on the importance of genetic drift within large pools of alleles is highlighted by the differences in conclusions reached by Fisher (1930) and Kimura (1968). In layman's terms this parameter effectively reduces the population size thereby increasing the impact of stochastic drift. This is exactly what is being observed within microbial ecology (Sloan et al., 2021).

The error within empirically derived neutral parameters from microbial communities look to reflect a similar phenomena observed within population genetics, however, the exact mechanisms driving this are, as of yet, unproven. While

it is hypothesised (Sloan et al., 2021) that the answer may lie in theoretical arguments with a richness as vast as that developed within population genetics (see J. Wang et al., 2016), the answer may lie in the more prosaic problem of model identification. Within microbial ecology, reduced sampling is often "excused" as analysis of samples is costly in both time and expense, thus, a pragmatic approach can be adopted (Lemke & DeSalle, 2023). The effect these "excused" sampling regimes have on the identification of neutral model parameters has not been comprehensively determined. As sampling and analysis becomes increasingly affordable, a more scientific basis for choosing sampling protocols should ensue. Indeed, if we are to use neutral model parameters, calibrated from single-species relative abundance time series, to express the fundamental processes that drive microbial community structure, then it is crucial that the errors associated with the sampling protocols employed are quantified. This question is directly addressed within the research presented in Chapter 3.

2.3 Thermodynamics within microbial biology

2.3.1 A thermodynamic description of life

Schrödinger famously wrote "*an organism feeds upon negative entropy*" (Schrödinger, 1992, p. 71). What this brief statement encapsulates is the analysis of growth, and therefore life, through the lens of the second law of thermodynamics. In the words of Clausius, the second law of thermodynamics can be summarised as "*The entropy of the universe tends to a maximum*" (Clausius, 1867, p. 365). We can rephrase this as the total entropy in an isolated system must either increase or remain constant. Growth, therefore presents a contradiction to this viewpoint. If the universal tendency is toward decay (high entropy), how can highly organised organisms (low entropy) exist, let alone grow? Schrödinger's assessment is that this is achieved through metabolism and that "*the essential thing in metabolism is that the organism succeeds in freeing itself from all the entropy it cannot help produce while alive*" (Schrödinger, 1992, p. 71). This realisation underpins his concept of "negative entropy".

To analyse the change in entropy across a system we must first decide on suitable boundaries. Indeed, the lack of system boundaries expressed in Schrödinger's formulation of negative entropy was a source of criticism (Gnaiger et al., 1994). In a microbial context we can quite easily set the system boundary as the boundary around a cell, as shown in Figure 2.2. Given that a cell is an open system (von Bertalanffy, 1950), the rate of change of entropy can be expressed as:

$$\frac{dS_{sys}}{dt} = \dot{S}_{gen} + \dot{S}_{in} - \dot{S}_{out}, \quad (2.7)$$

where, dS_{sys}/dt is the rate of change of system entropy, \dot{S}_{gen} is the rate of entropy generated, \dot{S}_{in} and \dot{S}_{out} are the rate of entropy into and out of the system. For a cell, the primary entropic processes include the production of entropy within the cell by irreversible processes, heat transfer, the import of substrates and the export of catabolic products and synthesised biomass (von Stockar et al., 2006). Based on these entropy fluxes we can recast Equation 2.7 as:

$$\frac{dS_{cell}}{dt} = \dot{S}_{prod} + \frac{\dot{Q}}{T} + \sum_i (\bar{s}_i \cdot \dot{n}_i) - \bar{s}_x \cdot \dot{n}_x, \quad (2.8)$$

where, dS_{cell}/dt is the rate of change of cell entropy, \dot{S}_{prod} is the entropy production within the cell due to irreversible processes, \dot{Q} is the rate of heat import, T is absolute temperature, \bar{s}_i and \bar{s}_x are the partial molar entropy of the i^{th} metabo-

lite and the produced biomass respectively and \dot{n}_i and \dot{n}_x are the molar rates of import and export of the i^{th} metabolite and the produced biomass respectively. Given that \dot{S}_{prod} acts to increase entropy within the cell, such that with no other processes $dS_{cell}/dt > 0$, the eventual death or disorganisation of the cell is inevitable. Thus, to facilitate current existence, a steady state must be imposed, where $dS_{cell}/dt = 0$. The solution is to ensure that the other three processes amass a value sufficiently negative to overcome \dot{S}_{prod} . Given that metabolites are imported and exported from the cell, such that \bar{n}_i can either be positive or negative, to ensure $\sum_i(\bar{s}_i \cdot \dot{n}_i)$ is negative, the metabolites exported must have a higher partial molar entropy than those imported. Indeed, this represents the primary mechanism of exporting entropy, as although some reduction in entropy can be expected as a result of the export of newly formed biomass ($\bar{s}_x \cdot \dot{n}_x$), this biomass will be of a highly organised nature, and thus have low partial molar entropy (von Stockar et al., 2006). This entropic flux, described by $\sum_i(\bar{s}_i \cdot \dot{n}_i)$, is the exact mechanism Schrödinger was referring to in his characterisation of feeding on negative entropy.

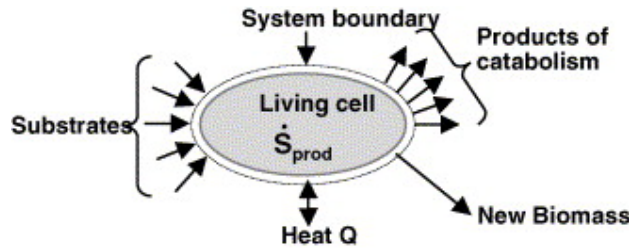


Figure 2.2: Diagram defining the cell boundary and processes driving the entropy balance for a growing cell. Reproduced from von Stockar et al. (2006)

Despite the merit and novelty of Schrödinger's outlook at the time, debate regarding the interpretation and significance of "negative entropy" began immediately after its conceptualisation (Gnaiger et al., 1994). A proportion of early controversy, pertained to the linguistic nature of the theory. As a result, Schrödinger conceded that the term "free energy" would have been a more technically sound notation to adopt (Schrödinger, 1992, p. 74). Indeed, an assessment of the importance of negative entropy within microbial growth, demonstrates that through the evaluation of Gibbs free energy, growth can be enthalpy-driven as well as entropy-driven (von Stockar & Liu, 1999). Within these explanations focus is given to the importance of heat dissipation, a process insufficiently accounted for in Schrödinger's formulation. Through calorimetric studies investigating the heat evolved during the growth of *Methanobacterium thermoautotrophicum* and *Methanosarcina barkeri*, autotrophic and acetotrophic methanogens respectively, direct evidence of entropy-retarded and enthalpy-retarded growth is presented (J.-S. Liu et al., 2001; Schill & von Stockar, 1995). Thus, it is not always the case that growth feeds solely on negative entropy (von Stockar & Liu, 1999).

The concept of "available energy" or "free energy" was developed by J. Willard Gibbs throughout the 1870s, with a summary of this work provided in 1878 (Gibbs, 1878). What this concept of available energy describes is the portion of energy that can be used to perform work for a system at constant pressure and temperature. It is now most commonly referred to as the Gibbs free energy, and the change in this quantity can be expressed as:

$$\Delta G_{sys} = \Delta H_{sys} - T \Delta S_{sys}, \quad (2.9)$$

where ΔG_{sys} is the change in Gibbs free energy of the system, ΔH_{sys} is the change in enthalpy, T is the absolute temperature and ΔS_{sys} is the change in entropy of the system. Using the cell once again as a model organism, we can defined the rate

of change in Gibbs free energy as:

$$\frac{dG_{cell}}{dt} = \frac{dH_{cell}}{dt} - T \frac{dS_{cell}}{dt}, \quad (2.10)$$

where, $\frac{dG_{cell}}{dt}$, $\frac{dH_{cell}}{dt}$ and $\frac{dS_{cell}}{dt}$ represent the the rate of change of Gibbs free energy, enthalpy and entropy with respect to time respectively. T carries the same definition as before. To arrive at a more detailed mathematical identity we can express the rate of change of enthalpy as:

$$\frac{dH_{cell}}{dt} = \dot{W} + \dot{Q} + \sum_i (\bar{h}_i \cdot \dot{n}_i) - \bar{h}_x \cdot \dot{n}_x, \quad (2.11)$$

where \dot{W} is the rate of work done on the cell, \bar{h}_i and \bar{h}_x are the partial molar enthalpy of the i^{th} metabolite and the produced biomass respectively and $\frac{dH_{cell}}{dt}$, \dot{Q} , \dot{n}_i and \dot{n}_x carry the same definitions as before (von Stockar et al., 2006). Substituting Equations 2.8 and 2.11 into Equation 2.10 and simplifying, we get:

$$\frac{dG_{cell}}{dt} = \dot{W} + \sum_i (\bar{\mu}_i \cdot \dot{n}_i) - \bar{\mu}_x \cdot \dot{n}_x - T \dot{S}_{prod}, \quad (2.12)$$

where $\bar{\mu}_i$ and $\bar{\mu}_x$ are the chemical potentials of the i^{th} metabolite and the produced biomass respectively and all other identities hold their previous definitions. For the cell to survive it must balance the Gibbs free energy such that $dG_{cell}/dt = 0$ despite processes such as the production of entropy ($T \dot{S}_{prod}$) and the exporting of biomass ($\bar{\mu}_x \cdot \dot{n}_x$), that act as sources of energy loss. Within most bacteria this is achieved through catabolism in which high energy substrate is consumed and low energy products are excreted. This process acts to drive a positive value of $\sum_i (\bar{\mu}_i \cdot \dot{n}_i)$, compensating for the effects of $\bar{\mu}_x \cdot \dot{n}_x$ and $T \dot{S}_{prod}$. In some specialised bacteria, such as phototrophs, a positive \dot{W} can be used as a compensation mechanism (von Stockar et al., 2006).

As highlighted by the previous mathematical definitions, for a living organism to exist at a far-from-equilibrium state, a constant and unrelenting decrease in Gibbs free energy must be compensated for. This is achieved most commonly through a catabolic redistribution of energy from high-energy substrate to low-energy products. Although this "feeding" on free energy as expressed by Schrödinger (1992) may explain how life is able to thrive, Schneider and Kay (1994) take this theory further and use it to express why life exists. Their theory defines living systems as gradient dissipators and suggest that *"life exists on earth as another means of dissipating the solar induced gradient and, as such, is a manifestation of the restated second law"*. Indeed, the inception of theories tackling such fundamental questions demonstrates the theoretical power that assessing living systems through the second law of thermodynamics has.

2.3.2 Thermodynamics and microbial growth

So far we have supplied examples of thermodynamic interpretations of how and perhaps why life exists. Although this carries significant scientific and philosophical importance, from an engineering point of view we are arguably more concerned with the effects these processes have on phenomena we can use within engineering contexts, such as those that drive metabolic processes exploited by biotechnologies or contribute to competitive difference between species. Indeed, attempts at linking thermodynamic concepts to microbial functions has spawned entire fields of study. Reviewing such a volume of literature here would be futile, indeed, much more qualified reviews exist (Demirel & Sandler, 2002; Harris,

2009; Nicholls, 2013). Focus here is given to literature specifically looking at how thermodynamics influence higher-level growth parameters such as growth yields and rates.

The simplest thermodynamic assessment of microbial growth views the process as a spontaneous reaction (von Stockar et al., 2008). Under this description we can view microbial growth as a black box, one which requires reactants such as, a carbon source, nitrogen source, electron donor and electron acceptor, and produces a variety of products including, biomass, CO_2 and H_2O as well as specific catabolic products and any excreted metabolites (Figure 2.3) (Heijnen & Van Dijken, 1992). Given that all reactants and products are accurately measured, a corresponding macrochemical growth equation can be obtained (Battley, 1960a). To facilitate spontaneity the Gibbs energy change (ΔG) of the overall growth reaction must be negative. By comparing the sum of the Gibbs energy of all the inputs with the sum of the Gibbs energy of all the outputs, an energy efficiency can be expressed. This efficiency was proposed by Roels and coined the black box Gibbs energy efficiency (η^{BB}) (Roels, 1983). Early approaches to correlate Gibbs energy with growth descriptors, such as biomass yield, were expressed using this parameter (Heijnen & Van Dijken, 1992; Roels, 1983). Primarily the relationship demonstrates that at $\eta^{BB} = 1$ the theoretical maximum biomass yield is observed. The black box Gibbs energy efficiency is, however, extremely sensitive to the reference frame chosen for calculation of Gibbs energy. It has been demonstrated that reversals in the relationship between η^{BB} and biomass yield can occur based on the thermodynamic reference point chosen (Heijnen & Van Dijken, 1992).

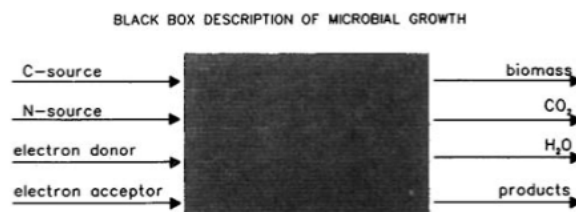


Figure 2.3: A black box description of microbial growth, highlighting generalised reactants and products. Reproduced from Heijnen and Van Dijken (1992)

Rather than assessing the total black box reaction, an energy convertor description of microbial growth was developed, assessing microbial growth through coupling catabolic and anabolic reactions, as shown in Figure 2.4 (Battley, 1960b; Heijnen & Van Dijken, 1992). At first glance, this may seem counter-intuitive, as the synthesis of complex, high-energy biomass from simpler precursors, through anabolism, surely promotes a positive value for ΔG (von Stockar et al., 2008). However, by coupling this reaction to the energy-yielding catabolic reaction we define a source of energy for the anabolic reaction to consume. Indeed, the coupling of reactions is used to explain how "unnatural" processes can be driven (Kedem & Caplan, 1965). On a cellular level this is made possible because microbial organisms transport energy between catabolic and anabolic reactions using energy carriers such as adenosine triphosphate (ATP) (Kleerebezem & van Loosdrecht, 2010).

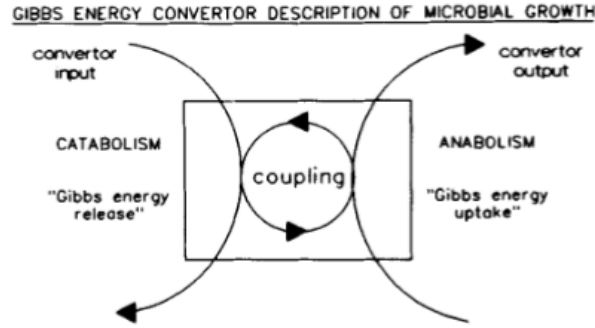


Figure 2.4: A energy convertor description of microbial growth, highlighting the coupling of catabolic and anabolic reactions. Reproduced from Heijnen and Van Dijken (1992)

Catabolic reactions are exergonic, with a negative ΔG , while anabolic reactions are endergonic, with a positive ΔG . Given that the Gibbs energy change across the overall coupled reaction ($\Delta_r G$) must remain negative, the change in Gibbs free energy associated with the anabolic reaction must be balanced by that of the catabolic reaction. The energy requirement of the anabolic reaction can therefore be thought of as a "payload" that the catabolic reaction must be able to carry (von Stockar et al., 2008). We can once again determine a thermodynamic efficiency (η) that defines the fraction of recovered energy as biomass from the energy released during the catabolic reaction. This is known as a energy transducer efficiency (Stockar & Wielen, 2013a) and can be expressed as:

$$\eta = -\frac{Y_{X/S}\Delta G_{ana}}{\Delta G_{cat}}, \quad (2.13)$$

where η is the efficiency, $Y_{X/S}$ is the yield of gram of biomass per gram of substrate, ΔG_{ana} is the change in Gibbs free energy of the anabolic reaction and ΔG_{cat} is the change in Gibbs free energy of the catabolic reaction. The negative is included to transform the efficiency into a positive value as we assume ΔG_{cat} to be a negative quantity. If ΔG_{ana} and ΔG_{cat} are fixed, Equation 2.13 demonstrates that the efficiency is directly linked to $Y_{X/S}$, where a more energy efficient organism will have a higher yield of biomass.

To express the relationship between $Y_{X/S}$ and the overall change in Gibbs energy across the coupled reaction ($\Delta_r G$) we can express the relationship:

$$\Delta_r G = \Delta G_{cat}(1 - \eta), \quad (2.14)$$

where quantities hold their previous meanings. Substituting Equation 2.13 into 2.14 we get:

$$\Delta_r G = \Delta G_{cat} + Y_{X/S}\Delta G_{ana}, \quad (2.15)$$

where again all quantities hold previous definitions. For an efficient organism, where $\eta \rightarrow 1$, $Y_{X/S}$ will tend to a maximum at $Y_{X/S}\Delta G_{ana} = -\Delta G_{cat}$. According to Equations 2.15 this will drive $\Delta_r G$, also known as the "driving force" of the reaction, to zero, representing thermodynamic equilibrium (von Stockar & Liu, 1999). A trade-off between biomass yield and this "driving force" is apparent.

Given this relationship between yield and efficiency it would seem logic to expect evolution to favour efficient growth, where η tends to 1 and thus $Y_{X/S}$ tends to its maximum, however, this applies a limitation to the rate of the coupled reactions. If we assume the case where the catabolic and anabolic reactions are perfectly and inversely coupled, such that every unit of energy from the catabolic process is used to drive the anabolic reaction in the direction opposite to its spontaneous tendency, the coupling coefficient q is equal to -1 . By favouring efficient growth under this scenario, the force ratio between the catabolic and anabolic reaction would be driven to 1, meaning the magnitude of the driving (catabolic) and driven (anabolic) forces are equal in magnitude (Stockar & Wielen, 2013b). This is demonstrated by figure 2.5a. This presents a problem, as under such conditions all the Gibbs free energy provided by the catabolic reaction would be consumed by the anabolic reaction, forcing the power output, or $\Delta_r G$, to zero and thus thermodynamic equilibrium, as highlighted by figure 2.5b. At this point the reaction would be infinitely slow and thus no growth would occur. This condition is named "level flow" (Kedem & Caplan, 1965). Based on this description, we may assume that evolution would act to drive η to a value less than 1. Indeed, it has been suggested that evolution could favour the maximisation of power output, which, for perfectly and inversely coupled reactions ($q = -1$), would occur at a force ratio of 0.5. This is achieved by reducing efficiency. In most cases it is improper to assume that the biomass production is perfectly coupled to the catabolic reaction and thus q is more likely to be > -1 . The greater q is from -1 the lower the optimum force ratio will be further driving a decrease in the 'optimal' efficiency and thus biomass yield (Stockar & Wielen, 2013b; Stucki, 1980).

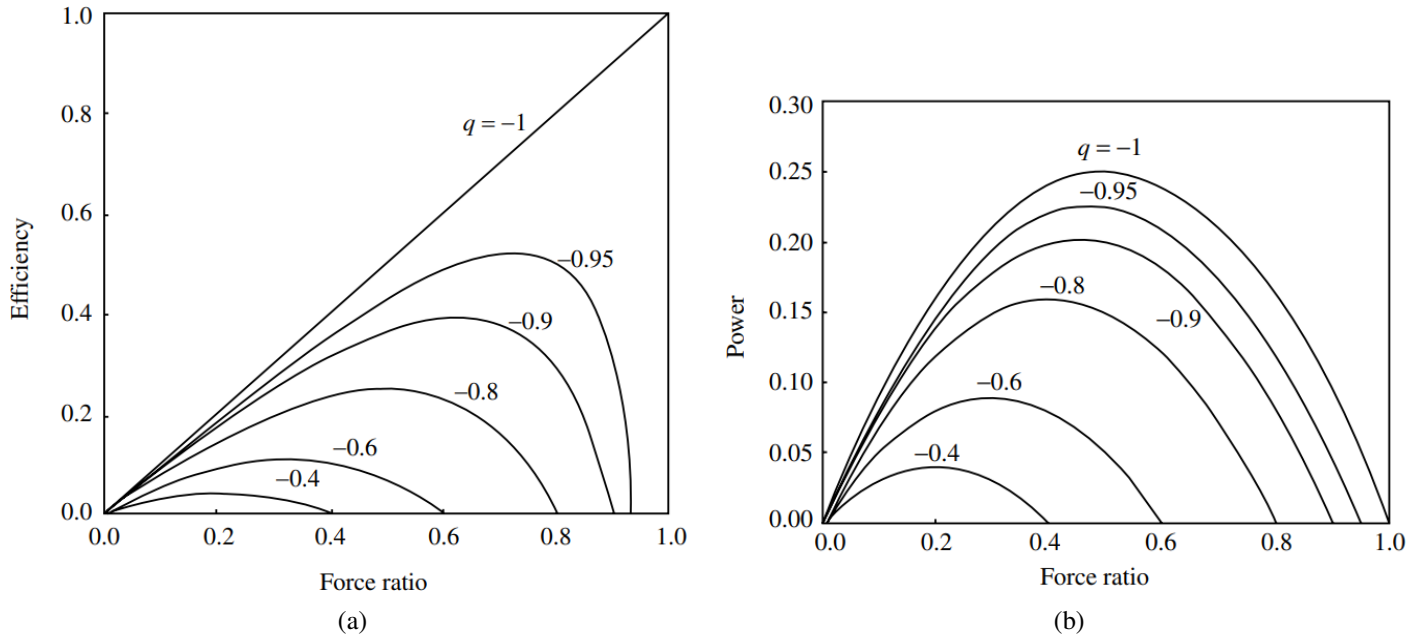


Figure 2.5: Relationships within a linear energy converter. Reproduced from Stockar and Wielen (2013b). a) Efficiency against force ratio, b) output power against force ratio.

This description presents a trade-off between $\Delta_r G$ and $Y_{X/S}$ driven by thermodynamic constraints. This trade-off is expanded to growth kinetics as well as yield, where the thermodynamic driving force ($\Delta_r G$) is compared directly to rate of catabolic and anabolic reactions (Stockar & Wielen, 2013a; von Stockar & Liu, 1999). This can be thought of similarly to an increase in steepness of physical gradient driven by an increase in potential energy and thus an increase in the velocity of the object travelling down this gradient (Stockar & Wielen, 2013a; von Stockar et al., 2008). Although there is some evidence of this relationship (Birou et al., 1987; J.-S. Liu et al., 1999; von Stockar & Liu, 1999) it is important to note

that biochemical processes, unaccounted for within this description, are known to drive, or place limits on, metabolic flux (Chubukov et al., 2014; Gerosa & Sauer, 2011; Locasale & Cantley, 2011). Revisiting the previous analogy, we can compare these to physical processes, such as the friction between the object and the slope or the terminal velocity imposed by aerodynamics (Harris, 2009). Further experimental analysis of the role that $\Delta_r G$ has on growth kinetics will provide a more robust foundation for developing theoretical relationships between these quantities.

2.3.3 The measurement of heat dissipation

Heat export plays a crucial role in maintaining cellular energy balance and dissipating excess energy. Indeed, heat dissipation is a universal process exhibited by all living organisms. When assessing heat dissipation at a microbial level, during small-scale experimental studies heat is largely lost to the environment, however, in larger systems, where significant microbial activity is expected, considerable increases in temperature have been observed (Fjelsted et al., 2020; Huete et al., 2018). For biotechnologies, the quantification of heat dissipation from microbial processes provides useful information for informing design (Gómez-Borraz et al., 2022). Observed temperature increases can also be used to directly assess microbial activity where the quantity of heat dissipation from the organisms present is known or can be estimated (Warren & Bekins, 2015).

Quantifying heat dissipation can also provide information at a more fundamental level. The previous thermodynamic description of microbial growth demonstrates that we can expect heat to be dissipated as a result of the excess quantity of $\Delta_r G$ needed to drive the coupled anabolic reaction. It has therefore been proposed that by measuring heat dissipation, an indirect observation of $\Delta_r G$ can be made (Stockar & Wielen, 2013a). These measurements can then be linked to observed growth kinetics, as demonstrated by Birou et al. (1987), allowing for further development of theoretical arguments or for developing on-line monitoring technologies.

The primary mechanism for quantifying the heat dissipation from a process is calorimetry. Calorimeters function in a variety of ways, however, all aim to measure the thermal energy dissipated from a sample added to the measurement chamber. The earliest interpretation of this technology is provided in Figure 2.6 and represents the methodology designed by Lavoisier and Laplace. This involved measuring the volume of water produced by the heat dissipated from a guinea pig housed in an enclosure and surrounded by ice and snow (Notley et al., 2023). Since this point technology has developed significantly, with recent develops including megacalorimetry, chip-calorimetry, high-throughput calorimetry, ultra-sensitive calorimetry and photocalorimetry (Maskow et al., 2010). Most calorimeters are either adiabatic or isothermal in nature. Adiabatic calorimeters allow for a temperature change to arise within the sample chamber and as a result are largely unsuited for measuring heat dissipation during microbial growth, as changes in temperature can effect growth kinetics (von Stockar & Marison, 2005). Isothermal (heat conduction) calorimeters measure the rate of heat production while maintaining the internal chamber at a fixed temperature using high-precision temperature control systems. In systems that use electrical compensation, the voltage required to maintain this constant temperature is proportional to the heat exported or imported by the sample, providing a measurement of this quantity (Wadsö & Li, 2008). In most cases a reference chamber is also used to account for any external disturbances (von Stockar & Marison, 2005).

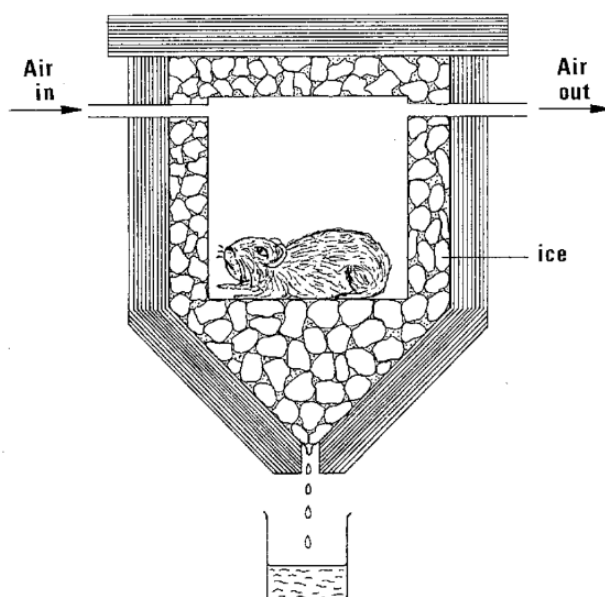


Figure 2.6: A diagram representing Lavoisier and Laplace's early ice calorimeter. Reproduced from von Stockar and Marison (2005).

Calorimetry, or more accurately microcalorimetry, is arguably underused within microbial studies and bioengineering. This is primarily due to the necessity of maintaining a fixed temperature and therefore a tightly controlled and closed environment. This makes processes common to microbial culturing such as, sampling, culture agitation and inclusion of additional probes and sensors difficult to achieve (Maskow et al., 2010). Despite these challenges, past studies utilising microcalorimetry have provided clear results and been used to forward the thermodynamic description of microbial growth (Birou et al., 1987; J.-S. Liu et al., 2001; Schill & von Stockar, 1995). Within wastewater engineering, and as a result of the previously stated challenges facing current technologies, there is a greater need to engineer carefully within energy bounds. Quantifying the amount of chemical energy available to these systems has proved fruitful for developing new technologies able to harvest it (Heidrich et al., 2011, 2013). However, to gain an insight into the bioenergetics taking place at the heart of these systems the redeployment of calorimetry is needed.

2.4 Methanotrophs

2.4.1 Methane

Methane is the second most abundant greenhouse gas in the atmosphere. Although it has a relatively short lifespan of 11.8 ± 1.8 years, it holds a global warming potential of 29.8 ± 11 and 27.0 ± 11 times that of carbon dioxide over a 100-year period, when emitted from fossil fuels or non-fossil sources respectively. (Intergovernmental Panel on Climate Change (IPCC), 2023b). From the late 1970s, when the first measurements of atmospheric methane were recorded (Blake et al., 1982), until the turn of the millennium, the global methane concentration was shown to increase, despite reductions within its growth rate (Kirschke et al., 2013). Between 2000 and 2006, when the growth rate of atmospheric methane was at its lowest, a short quasi-equilibrium phase within the global atmospheric methane concentration was observed (Dlugokencky et al., 2003; Kirschke et al., 2013). However, from 2006 onwards, the concentration of methane within the

atmosphere has been steadily increasing (Prather & Holmes, 2017). Ice core records paint an even more damning picture, with the global methane concentration almost tripling in the last 1,000 years (Etheridge et al., 1998). Measurements taken in 2019 show that the concentration of global methane has reached its highest in over 800,000 years (Intergovernmental Panel on Climate Change (IPCC), 2023a).

The abundance of methane within the atmosphere is defined by the trade-off between sources and sinks. Methane sources are abundant and diverse, yet can be categorised as thermogenic, pyrogenic, or biogenic. Thermogenic methane is formed through the decomposition of organic material due to the heat and pressure deep within the earth. This methane accounts for natural gas reserves and is released during the processing of fossil fuels. Pyrogenic methane is formed through the incomplete combustion of biomass, processes such as biofuel usage, peatland fires, and forest fires facilitate the formation of this methane. Finally, biogenic methane refers to methane produced by methanogenic organisms. These species exist in anaerobic environments and are responsible for the methane emitted from wetlands, human and animal guts, landfills and wastewater treatment systems (Saunois et al., 2020). Each process has a unique range of isotopic ratios, and as such, measuring the isotopic composition of the atmosphere can be used to define the relative importance of these sources (Monteil et al., 2011). In addition to defining methane sources by their emission process, sources are often categorised as either natural or anthropogenic. Between 2008 and 2017 anthropogenic sources were estimated to make up between 62% (top-down estimation) and 49% (bottom-up estimation) of all methane sources, with the majority of these emissions coming from agriculture and waste (Saunois et al., 2020). Natural sources, such as wetlands, freshwater, permafrost, and others, also play a significant role in the global methane budget. Yet, large uncertainties within the estimation of these sources make analysing their relative contributions difficult (O'Connor et al., 2010; Saunois et al., 2020).

Methane sinks exist in soils and the atmosphere, where atmospheric sinks account for the vast majority of methane losses. This is a result of oxidation by hydroxyl radicals, chlorine atoms, and excited oxygen. Photolysis is also known to occur within the upper atmosphere, however, due to the lack of methane within this region, this process is usually ignored within global methane budgets (Brasseur & Solomon, 2005). Out of all atmospheric oxidisers, the hydroxyl radicals are by far the most prominent, being estimated to account for 84% of all methane sinks during the 2000-2009 period. In perspective, oxidation by atmospheric chlorine was estimated to account for only 4%, while the effects of excited oxygen, assumed to be 25% of stratospheric losses, is estimated to make up 1% (Saunois et al., 2020). Indeed, given the importance of hydroxyl radicals in removing methane, it has been suggested that the observed trend within the global methane concentration since 1995 is directly linked to changes in the concentration of these molecules (Rigby et al., 2017). Apart from atmospheric losses, the only other sink to be considered within global methane budgets is soil uptake (Saunois et al., 2020). The loss of methane within soils is a direct result of methanotrophic bacteria. These bacteria undertake the oxidation of methane to form carbon dioxide through the metabolic pathway they exploit. They account for the only biological methane sink (Guerrero-Cruz et al., 2021).

2.4.2 Aerobic methanotrophic bacteria

The first species of aerobic methanotrophs was isolated in 1906 by Söhngen (Sohngen, 1906) named *Bacillus methanicus*, now known as *Methanomonas methanica* after being renamed by Orla-Jensen in 1909 (Orla-Jensen, 1909). Despite this success, early attempts to isolate methanotrophs were often unsuccessful. This difficulty in isolating methanotrophs acted to hamper early research within the field (Leadbetter & Foster, 1958). Isolation of new species began to occur in the 1950s and 60s with the isolation of *Pseudomonas methanica* (Dworkin & Foster, 1956; Leadbetter & Foster, 1958), *Methanomonas methanooxidans* (L. R. Brown et al., 1964) and *Methylococcus capsulatus* (Foster & Davis, 1966). However, it wasn't until Whittenbury, Phillips, and Wilkinson published the isolation of over a hundred aerobic methane utilising bacteria that the foundation of current taxonomic classifications were established (Whittenbury et al., 1970).

Methanotrophs fall under two phylum *Proteobacteria*, split into classes of *Alphaproteobacteria* and *Gammaproteobacteria*, and *Verrucomicrobia* (Hanson & Hanson, 1996; Op den Camp et al., 2009), where methanotrophic *Verrucomicrobia* were only isolated more recently (Dunfield et al., 2007; Islam et al., 2008; Pol et al., 2007). Rather than defining methanotrophs at taxonomic levels it is more common to categorise them under simplified functional types, type-I, type-II, and type-X, where type-X are a sub-group of type-I methanotrophs (Bowman, 2006). These types classify methanotrophs based on phylogeny, biochemical differences, internal ultrastructure and carbon assimilation pathways (Bowman, 2006; Hanson & Hanson, 1996). Type-I methanotrophs are all part of the *Methylococcaceae* family, including genera, *Methylomonas*, *Methylobacter*, *Methylococcus*, *Methylosphaera*, *Methylocaldum*, *Methylothermus*, *Methylohalobius*, *Methylochromium*, *Methylosarcina*, and *Methylosoma* (Smith et al., 2010). All type-I methanotrophs are gram-negative, aerobic, obligate methylotrophs that assimilate carbon using the ribulose monophosphate RuMP cycle (Hanson & Hanson, 1996). *Methylosinus*, *Methylocystis*, *Methylocella*, and *Methylocapsa* are examples of genera that belong to the *Methylocystaceae* family and are thus categorised as type-II methanotrophs (Bowman, 2006; Smith et al., 2010). Similarly to type-I methanotrophs they are also gram-negative, aerobic, obligate methylotrophs. Where they differ is in the use of the serine cycle for carbon assimilation (Hanson & Hanson, 1996).

Figures 2.7a and 2.7b depict the carbon assimilation pathways for type-I and type-II methanotrophs respectively. Both pathways begin from the catabolic intermediate formaldehyde. In the RuMP cycle, employed by type-I methanotrophs, formaldehyde is initial fixed using ribulose-5-phosphate and proceeds through hexulose-6-phosphate to fructose-6-phosphate, after which variation in the metabolic pathway can occur. (Smith et al., 2010). For type-II methanotrophs, serine is synthesised through the reaction of formaldehyde and glycine. This amino acid is eventually converted into glyoxylate, which leads to the regeneration of glycine, allowing the cycle to continue (Smith et al., 2010). A consequential step within the serine cycle is the conversion of phosphoenol pyruvate to oxaloacetate as this requires CO_2 , as highlighted in Figure 2.7b. This need for CO_2 increases interactions between the catabolic and anabolic pathways, as CO_2 is produced following the full oxidation of methane. Indeed, an increased lag phase is observed during the growth of type-II methanotrophs in systems where the initial concentration of CO_2 is low and thus the catabolic oxidation of methane is the sole source of CO_2 (S. Park et al., 1991). Both assimilation pathways eventually tie into more standard cycles, such as the tricarboxylic acid (TCA) cycle.

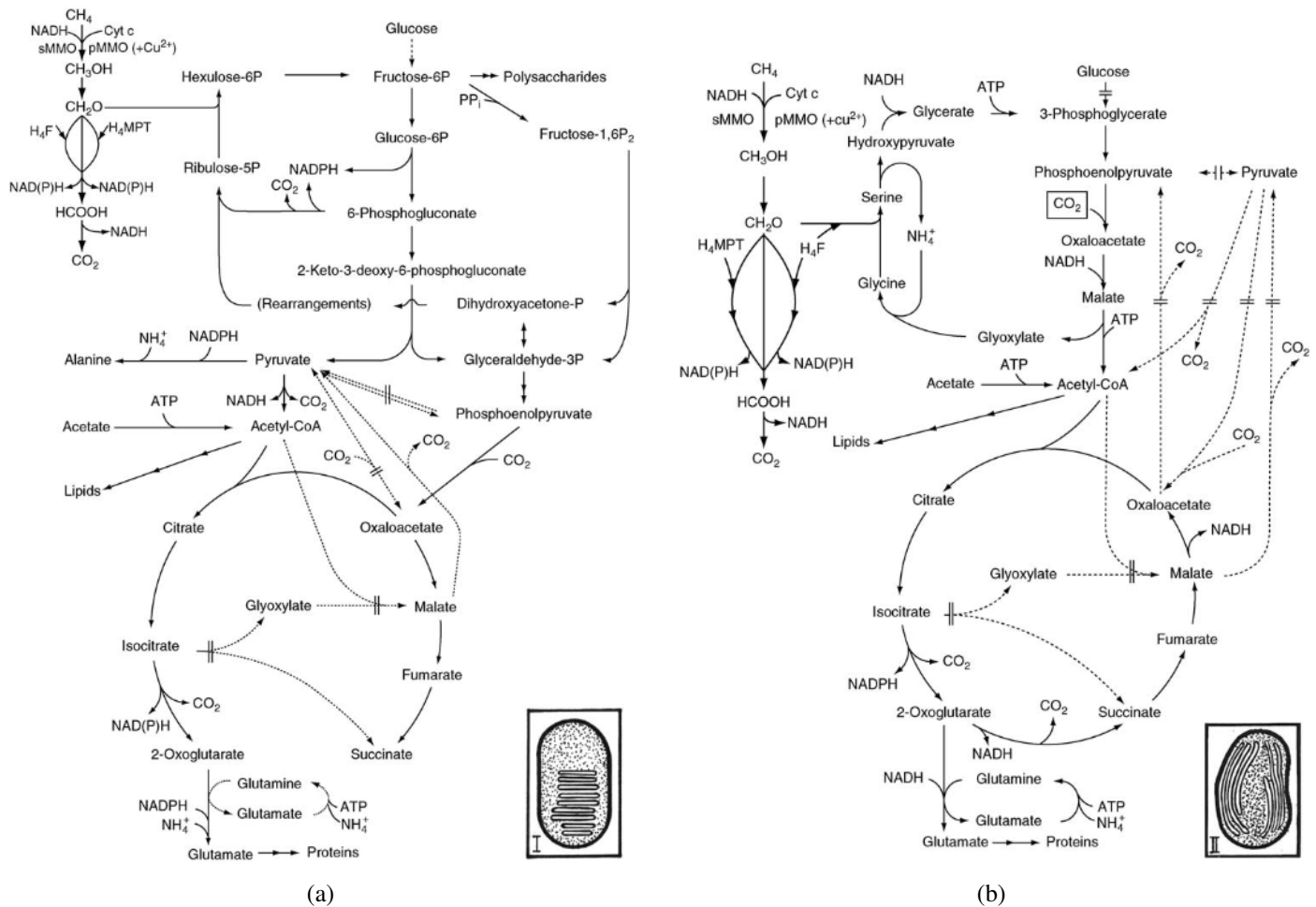


Figure 2.7: Carbon assimilation pathways of type-I and type-II methanotrophs. Reproduced from Smith et al., 2010. a) carbon assimilation pathway of type I methanotrophs utilising the ribulose monophosphate (RuMP) cycle, b) carbon assimilation pathway of type II methanotrophs utilising the serine cycle,

All aerobic methanotrophs fundamentally share the same catabolic pathway, and generate energy through the oxidation of methane to carbon dioxide via intermediates, methanol, formaldehyde and formate. As methanotrophs utilise the same compound as their energy and carbon source, not all methane is converted to carbon dioxide through this catabolic pathway. Thus the stoichiometric of the overall reaction can be expressed as:



where x is the stoichiometric coefficient, typically defined as 0.5 (De Visscher & Van Cleemput, 2003; Gómez-Borraz et al., 2022; Molins et al., 2008; Ng et al., 2015). Differences in the amount of carbon partitioned between biomass and CO_2 has been measured between type-I and type-II methanotrophs (Templeton et al., 2006). Indeed, although the general catabolism of methane is considered uniform across methanotrophs (Hanson & Hanson, 1996; Khmelenina et al., 2018), differences in the exact stoichiometry and biochemistry that facilitate this reaction have been identified.

During the oxidation of methane to methanol, methane monooxygenases (MMOs) are used to activate the methane. These enzymes have been shown to exist in two forms, a copper-dependent particulate form (pMMO) that is found within the cytoplasmic membrane, and a soluble form (sMMO) that utilises a diiron active site and is present within the cytoplasm. A recent review of the architecture of both these enzymes is given by Koo and Rosenzweig (2021). pMMO can be expressed by almost all methanotrophs either solely or in addition to the expression of sMMO. Only a handful of species belonging to the *Methylocella* and *Methyloferula* genera have been shown to only express sMMO (Dunfield et al., 2003; Vorobev et al., 2011). For species able to express both forms of methane monooxygenase, expression is regulated through a copper switch, where pMMO is the dominant enzyme within copper-rich environments and sMMO is only expressed under copper limitation (Choi et al., 2003; Stanley et al., 1983). Similar to the oxidation of methane, the oxidation of methanol is accepted to be achieved by the expression of two enzymes, a calcium-dependent type of methanol dehydrogenase MxaFI, and a lanthanide-containing XoxF type (Chu & Lidstrom, 2016; Khmelenina et al., 2018; L. Wang et al., 2020). The expression of different enzymes within oxidation steps can harbour direct consequences regarding the competition between methanotrophs. While the expression of either pMMO or sMMO is not type specific, it is generally accepted that type-I methanotrophs primarily express pMMO while type-II methanotrophs are often more versatile and can express both. As a result it has been shown that copper limitation can be used to promote the selection of type-II methanotrophs as a direct consequence of sMMO expression (D. W. Graham et al., 1993).

Other factors are also known to impact competition between type-I and type-II methanotrophs (D. W. Graham et al., 1993; Hanson & Hanson, 1996). Studies assessing microbial community composition at different soil depths, demonstrate vertical distribution of methanotrophs, where type-I methanotrophs are show to be abundant closer to the surface while type-II are more abundant deeper within the soil (Bull et al., 2000; Conrad & Rothfuss, 1991; Henckel et al., 2000). This spatial separation of methanotrophs is explained through observed differences in the oxygen and methane concentrations at differing soil depths (Frenzel et al., 1992). Experimental studies that facilitate differing oxygen and methane gradients support this hypothesis, demonstrating that type-I methanotrophs dominate in high-oxygen low-methane environments whereas type-II methanotrophs outcompete type-I in environments of low oxygen and high methane (J. Amaral et al., 1995; J. A. Amaral & Knowles, 1995). The competitive differences observed between type-I and type-II methanotrophs demonstrates occupation of different ecological niches (Henckel et al., 2000).

2.4.3 Application within biotechnologies

From an engineering viewpoint, research into methanotrophs tends to explore how the biological oxidation of methane can be utilised for industrial purposes. One example of this is the use of methanotrophs in biofuel processing. Anthropogenic sources of methane, such as landfill, wastewater treatment, and anaerobic digestion, facilitate the capture of biogas that can be used as a renewable alternative to fossil fuels. The main challenges associated with biofuel occur in the cleaning and upgrading processes required before customer use (Rafiee et al., 2021). Although cleaned biofuel can be burned to generate energy, research suggests that the conversion of biofuel to liquid fuels and desired chemicals is of greater benefit (Haynes & Gonzalez, 2014). The conversion of methane to liquid fuels is carried out predominantly by Fischer–Tropsch technologies, which are expensive in both the energy and capital costs required (Riaz et al., 2013; Wood et al., 2012). Methanotrophs are centred within emerging technologies as a result of their ability to oxidise methane and produce methanol under atmospheric conditions (Ge et al., 2014; Guerrero-Cruz et al., 2021; Haynes & Gonzalez, 2014; Priyadarsini et al., 2023; Sheets et al., 2016).

More relevant applications of the biological oxidation of methane to the research presented in this thesis, are those concerned with the direct reduction of greenhouse gas emissions from waste management systems. The use of biofilters, containing methanotrophs within the associated microbial consortium, is being pushed as a strategy to reduce anthropogenic methane emissions (Brandt et al., 2016; Gómez-Borraz et al., 2017; La et al., 2018). This technology is especially suited to small-scale systems, such as decentralised wastewater treatment systems, in which biogas may not be produced at quantities great enough for cost-effective capture. Even in systems where biogas capture is feasible, dissolved methane is being directly emitted through the effluent, where once discharged can desorb to the atmosphere (Gómez-Borraz et al., 2017). Studies suggest that up to 50% of the CH_4 produced via anaerobic digestion can be emitted to the environment via this route (Noyola et al., 1988; Souza et al., 2011). Laboratory studies have further demonstrated that by co-culturing aerobic methanotrophs with microalgae, the CO_2 produced during methane oxidation can also be removed (Ruiz-Ruiz et al., 2020).

Methanotrophs can also generate a variety of products useful to society, such as, single cell protein (SCP), internal storage polymers (PHB), lipids, ectoine and extracellular polysaccharides (EPS) (Strong et al., 2015). Genetic modification of methanotrophs has also been proposed for the biosynthesis of additional products (Lee et al., 2019; A. D. Nguyen et al., 2020; Tan et al., 2024). In most cases further research is however required if biotechnologies based solely on biosynthesis are to make it to market (Strong et al., 2015). Coupling these biosynthetic processes with the aforementioned ability of methanotrophs to reduce greenhouse gas emissions from anthropogenic sources of methane provides a promising strategy within a circular economy (AlSayed et al., 2018; García & Galán, 2022; Govindaraju et al., 2022; Guerrero Cruz & Pijuan, 2022). The goal of research within this field is to therefore develop technologies and methodologies to reduce process costs, or to identify additional products that can be used to increase the end value (Strong et al., 2015). Within this thesis we explore heat dissipation as a product of methanotrophic growth. Indeed, although temperature rises as a result of biological methane oxidation have been observed (Fjelsted et al., 2020; Huete et al., 2018; Warren & Bekins, 2015), and attempts at modelling this have been undertaken (Gómez-Borraz et al., 2022), the exact quantity of heat produced from these bacteria has not yet been quantified experimentally.

2.5 Summary

To meet current challenges facing water security there is a shift within wastewater engineering from a reliance on energy-intensive systems to processes that enable resource recovery and water recycling. To facilitate this shift, precise control of the microbial communities underpinning these technologies is vital. Within this thesis, advancement to this aim is pursued through development towards a generalised theory describing these communities, one that forgoes defining complex micro-scale interactions in an attempt to prioritise simplicity and universality.

Neutral models are adopted as the starting point for such a theory given that the number of parameters versus the number of successful predictions is large (Harte, 2004). Indeed, these models have been successfully fitted using datasets obtained from many microbial communities, including those found within wastewater treatment plants (Fodelianakis et al., 2021; Heys et al., 2020; Konopka et al., 2015; Logares et al., 2018; Ofițeru et al., 2010; Sloan et al., 2006, 2021; Venkataraman et al., 2015; Woodcock et al., 2007). However, assessing only the fit does not confirm these models to be "good" (Wieland et al., 2021). In many studies distinct model parameters are left unidentified, and in cases where they are identified there is little discussion on whether this has been done correctly. While studies accessing parameter identifiability have been undertaken when using stationary species abundance distributions (Etienne et al., 2006), this analysis has not yet been expanded to include single species relative abundance time series; a data type commonly obtained from microbial communities and assumed to increase the ability to identify distinct model parameters. As such it is unknown whether the observed mismatch between calibrated neutral model parameters and independent measurements of these parameters (Sloan et al., 2021) provides ecological insight or is simply a consequence of incorrect identification. We therefore arrive at the question, can distinct neutral model parameters be successfully identified when using single species relative abundance time series. This question is addressed by the research presented in Chapter 3 entitled *"Identifying neutral models on the basis of species abundance time series"*.

Whilst a generalised model may forgo describing micro-scale interactions and still be effective, knowledge of the mechanisms that drive the model parameters can prove beneficial. Indeed, a "good" model is only "useful" if correctly identified parameters provide biological insight (Wieland et al., 2021). Perhaps the most fundamental description of growth is provided by the second law of thermodynamics. Here we explore the difference in growth of two species of methanotrophic bacteria through the lens of thermodynamics. This is achieved through quantification of the heat dissipated by these bacteria through use of calorimetry. The quantity of heat dissipated is compared against growth parameters in an attempt to uncover if a clear relationship between these quantities exist. Due to the potential for methanotrophs to be incorporated within decentralised wastewater technologies, as a result to their ability to oxidise methane under standard conditions, the quantification of heat dissipated by these bacteria alone is an inherently valuable finding. This research is presented in Chapter 4 entitled *"Quantifying the heating potential of methanotrophs for application within decentralised wastewater treatment technologies"*.

Chapter 3

Identifying neutral models on the basis of species abundance time series.

3.1 Abstract

Guidance on the sampling regimes that allow for robust calibration of neutral models using relative abundance time series is sorely lacking. Here, we generate realisations of neutral dynamics at high temporal resolution using both neutral and near-neutral models and show that, even when sampling from these idealised species abundance time series, certain parameters are structurally unidentifiable. This is a consequence of correlated parameters and from the generic difficulties in integrating noisy stochastic data. As the times series duration and the temporal resolution at which it is sampled decrease to values that are realistic for a pragmatic experimental program, parameters previously shown to be identifiable now appear to be practically unidentifiable and are thus obscure to robust model fitting methods. The research aims to quantify the sampling regime-dependent accuracy in parameter estimates and the correlations between parameter estimates as a guide to the design of experimental campaigns.

Assessment of the errors within parameter estimates indicate that optimal values for these quantities exist and are dependent on the number and type of parameters estimated. Achieving accurate estimation of parameters, especially within the more complex near-neutral model, is shown to require further *a priori* independent knowledge of the processes governing community structure. The trade-offs in parameters are too extensive to estimate their values based on the dynamics within relative abundance displayed in the time series. Thus, even with high resolution data, to be confident that a model is true to its underlying process assumptions, independent experimental determination of model parameters is required. Alternatively, if relying on calibration with no prior knowledge, it is prudent to use compound parameters or, at least, be able to constrain the values a parameter can take and accept a level of error within the estimated parameters. Under scenarios where models under-represent the processes governing a community's structure, model parameters will most likely reflect complex dependencies rather than distinct biological mechanisms. The argument is made that by utilising effective parameters within highly simplified neutral models and calibrating these models under conditions that facilitate accurate parameter estimates, exploring processes that drive the magnitudes of these parameters can be used to inform strategies for influencing a community's structure.

3.2 Introduction

The role of demographic stochasticity was brought to the fore by researchers exploring the diversity of biological communities of birds, mammals and trees (Caswell, 1976; Hubbell, 2001; MacArthur & Wilson, 1967). Their primary sources of data were extensive surveys of diversity from multiple locations; the practical difficulties of sampling communities over multiple generations meant that time series data were rare. Therefore, model calibration aimed to match theoretical stationary species abundance distributions to observed species abundance distributions.

This was successfully employed initially to macro-organism (Duivenvoorden et al., 2002; Hubbell, 2001; Muneepeerakul et al., 2008; Ulrich & Zalewski, 2007; Volkov et al., 2007) and then, more recently, to microorganisms, for example, waterborne communities (Woodcock et al., 2007; Zhou et al., 2013), in the human lung (Venkataraman et al., 2015), the guts of insects (Adair et al., 2018) and fish (Burns et al., 2016; Heys et al., 2020), and on micro-plastics (Sun et al., 2021). Whilst these do lend weight to the view that drift plays a significant role they can only really be described as circumstantial evidence for three main reasons. Firstly, whilst the stationary distributions seem to fit the models, for large populations one might expect drift to be so slow as to be swamped by environmental fluctuations (Chisholm & O'Dwyer, 2014; Nee, 2005; Ricklefs, 2006). Secondly, it has been shown that calibration using stationary species abundance distributions is fraught with uncertainty and large regions of the parameter space can yield equally good fits (Etienne et al., 2006; Rosindell et al., 2011). Thirdly, other models have been shown to produce similar fits to species abundance distributions (Chisholm & Pacala, 2010; McGill, 2003). Indeed, the consensus is that stationary abundance distributions alone lack the information needed to explain micro-scale interactions. It has therefore been suggested that we should move away from the stationary species abundance distributions and utilise spatial or temporal data sets (Etienne et al., 2006; Jabot & Chave, 2009; Rosindell et al., 2011).

The relative ease of collecting abundance time series from microbial communities that span multiple generations, means that recent attempts to pin down the influence of drift using temporal data has been within the field of microbial ecology (Ofiteru et al., 2010; Sloan et al., 2021). An emerging finding is the apparent mismatch between calibrated neutral parameters and their independently measured counterparts, as illustrated by Sloan et al. (2021). There could be two reasons for this. It could be that the models are able to encapsulate processes that are not explicitly represented in their formulation. Thus, whilst the models may fit experimental data, the parameters have lost the direct link to the underlying phenomena that they are supposed to represent and, instead, capture the effects of more complex behaviours. Or it may be the more prosaic problem of model identification from noisy time series data.

Here, the initial focus is in the latter prosaic problem of accurately estimating model parameters from time series. It is only by demonstrating that it is possible to accurately calibrate model parameters using time series data and by guiding experimentalist towards effective sampling strategies that we can hope to answer questions like: are time series really better than spatially distributed data for model estimation; or, do the calibrated model parameters make sense with prior knowledge of the biological processes; and, consequently, is the model a good fit for the wrong reasons? The sampling regimes for most experimental studies, are usually determined heuristically. There is a trade off between obtaining as high a resolution as possible and pragmatism. Indeed, there is a certain "excused" reduction of sampling accepted due to the cost and complexity of analysis (Lemke & DeSalle, 2023).

With no ground truth, the accuracy of parameter estimates cannot be anticipated, nor can the perceived parameter mismatch be distinguished between the consequence of ecosystem functions or calibration errors. Not only this, but it is unclear whether the dynamics are unnecessarily over-sampled or so under-sampled as to negate any inference, or whether in all situations the extra effort required to collect time series data actually pays dividends. Here an attempt is made to remedy this issue, by answering if it is possible to accurately derive distinct parameters using species abundance time series as well as determining the effects sampling has on the estimation of model parameters.

Typically, quantification of the number of individuals in a particular species requires targeted molecular methods such as qPCR, which are laborious and, therefore, limit the number of species that can be tracked. Even if a full analysis of the microbial community can be undertaken the vast diversity within microbial communities make capturing and categorising every species by sampling extremely difficult (T. P. Curtis et al., 2002; Sloan et al., 2007; Woodcock et al., 2007). For this reason the analysis is limited to relative abundance time series obtained from a single detectable species. Indeed, Ofiteru et al. (2010) demonstrate that when using relative abundance time series, neutral model parameters can be calibrated for a single taxa.

3.3 Methods

3.3.1 Computational framework

To facilitate the research carried out within this study, a custom computational framework was developed. This was tasked with simulating single-species relative abundance time series under different sampling protocols using two distinct neutral models, while also facilitating the calibration of distinct neutral model parameters using these data sets. Figure 3.1 depicts a basic overview of the developed framework. In total three distinct packages were developed, a simulation package, a sampling package and a fitting package. The compartmentalisation of code allows packages to be operated individually or in series as highlighted in Figure 3.1. Outputs from each package are in the correct format for direct use by subsequent packages. Each package operates based on user provided operational parameters stored in associated spreadsheets, these are highlighted as the green files in Figures 3.1 and 3.2. A more comprehensive graphic of the framework is provided in Figure 3.2. An in-depth explanation of Figure 3.2, the overall architecture and individual scripts and functions are provided at <https://github.com/tymonherzyk>.

All scripts and functions within these packages were written using MATLAB version R2020b, in accordance with the academic license provided by the University of Glasgow. Storage of user variables for the operation of these packages is achieved through accessible xlsx files. Output and log files are produced as MATLAB files. All code is available at <https://github.com/tymonherzyk> alongside descriptions. Operational and installation guides are also provided. All code is shared under the Creative Commons (CC-BY 4.0) license. This framework constitutes the first user friendly, robust framework for simulating and fitting neutral dynamics. It is anticipated that the framework will find wide application in both macro and microbial ecology. The modular approach taken promotes future development further widening its possible application.

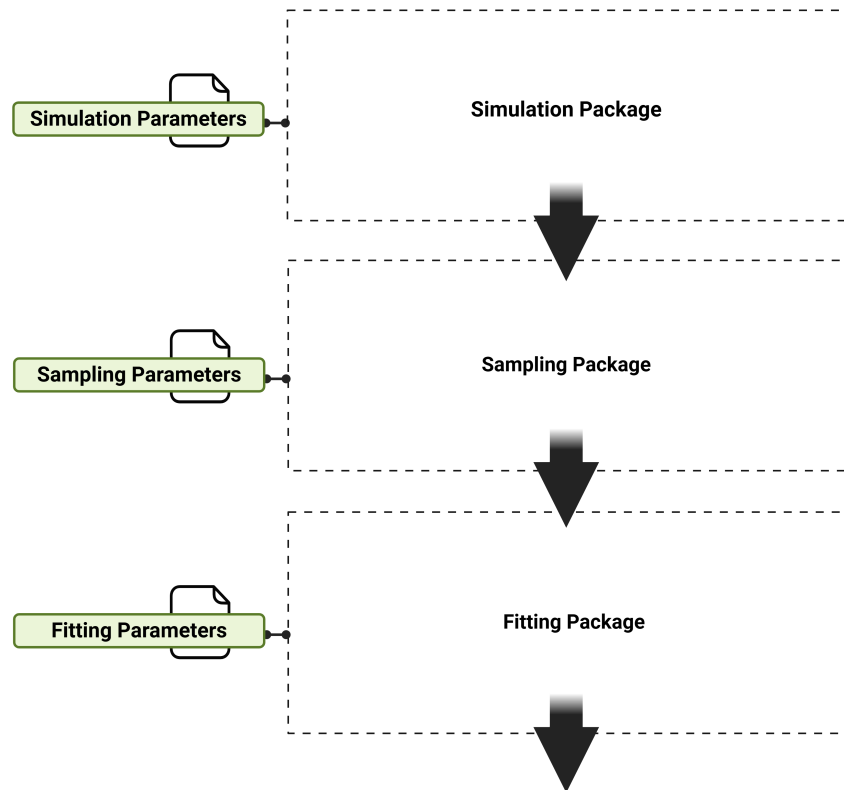
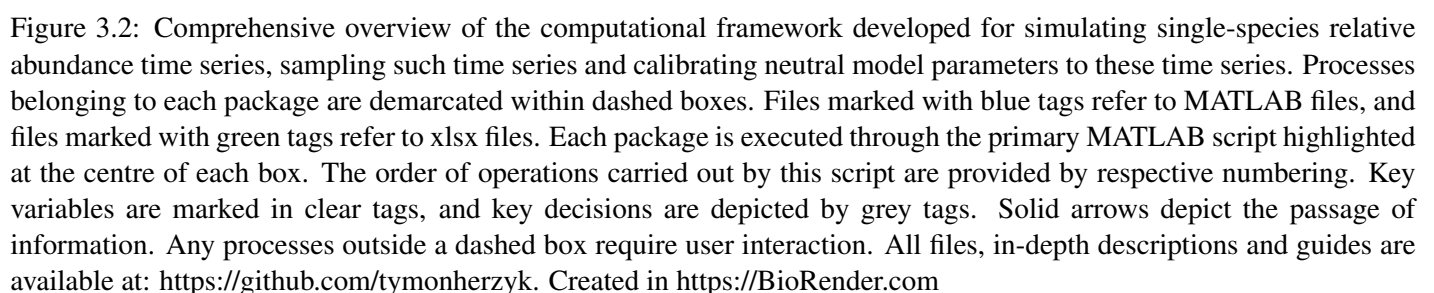


Figure 3.1: Basic overview of the computational framework developed for simulating single-species relative abundance time series, sampling such time series and calibrating neutral model parameters to these time series. In total three packages were developed that can be operated individually or in series. Packages operate according to user defined variables stored in associated spreadsheets. All files, in-depth descriptions and guides are available at: <https://github.com/tymonherzyk>. Created in <https://BioRender.com>



3.3.2 Discrete simulation of microbial communities using NCMs

To explore the calibration of neutral community models (NCMs), relative abundance time series were simulated for communities that adhere perfectly to the axioms of the models. The simulated time series were generated at exceptional high temporal resolution and over large periods; far higher resolution and longer duration than it would ever be feasible to collect for real communities. These idealised time series were generated for two commonly used NCMs: Hubbell's purely neutral model (Hubbell, 2001) and Sloan et al.'s near-neutral model (Sloan et al., 2006).

Firstly, Hubbell's neutral model represents a zero-sum game where the total number of organisms in the community stays constant and the community structure evolves through discrete random replacement events: births, deaths and external immigration. The mathematical description of this is given by transition probabilities in a Markov chain:

$$Pr\{N_i + 1 | N_i\} = \left(\frac{N_T - N_i}{N_T} \right) \left[mp_i + (1 - m) \left(\frac{N_i}{N_T - 1} \right) \right], \quad (3.1a)$$

$$Pr\{N_i - 1 | N_i\} = \left(\frac{N_i}{N_T} \right) \left[m(1 - p_i) + (1 - m) \left(\frac{N_T - N_i}{N_T - 1} \right) \right], \quad (3.1b)$$

$$Pr\{N_i | N_i\} = 1 - Pr\{N_i + 1 | N_i\} - Pr\{N_i - 1 | N_i\}, \quad (3.1c)$$

where, N_i represents the total number of individuals within the monitored i^{th} species, N_T is the total number of individuals in the community, which is assumed to be fixed, m is the probability that when an individual dies or leaves the community it is replaced by an immigrant from a source community, which is assumed to be species independent, and p_i is the relative abundance of the i^{th} species within the source community. So, for example, Equation 3.1a gives the probability that, when an individual in the community is replaced at random, the number in the i^{th} species increases from N_i to $N_i + 1$. For this to occur, there needs to be a death/emigration in some other species, with probability $\frac{N_T - N_i}{N_T}$, and either an immigration from the i^{th} species, with probability mp_i , or a birth from the i^{th} species, with probability $(1 - m) \left(\frac{N_i}{N_T - 1} \right)$.

The second, a near-neutral model developed by Sloan et al. (2006), includes local selection in addition to processes defined in Hubbell's model. The transition probabilities are therefore adapted to include a selection parameter α_i , resulting in equations:

$$Pr\{N_i + 1 | N_i\} = \left(\frac{N_T - N_i}{N_T} \right) \left[mp_i + (1 + \alpha_i)(1 - m) \left(\frac{N_i}{N_T - 1} \right) \right], \quad (3.2a)$$

$$Pr\{N_i - 1 | N_i\} = \left(\frac{N_i}{N_T} \right) \left[m(1 - p_i) + (1 - \alpha_i)(1 - m) \left(\frac{N_T - N_i}{N_T - 1} \right) \right], \quad (3.2b)$$

$$Pr\{N_i | N_i\} = 1 - Pr\{N_i + 1 | N_i\} - Pr\{N_i - 1 | N_i\}, \quad (3.2c)$$

where, parameters that are shared with Equations 3.1a, 3.1b and 3.1b keep their original meaning, while α_i represents a constant advantage assigned to the i^{th} species. Equations 3.2a and 3.2b are almost identical to their purely neutral

counterparts (Equations 3.1a and 3.1b) except for a selection advantage and disadvantage assigned to local birth given by $(1 + \alpha_i)$ and $(1 - \alpha_i)$ respectively. The differing of a single parameter between both models means that the influence of a single community structuring process, in this case local selection, can be explored by means of comparison between both simulated data sets. Indeed, if the influence of selection is deemed to be negligible, such that $\alpha_i = 0$, then the transition probabilities revert back to their purely neutral forms.

The simulated time series were produced at the highest possible temporal resolution, where every individual replacement event was represented. Therefore, simulations were conducted directly from the transitional probability equations defined for each model. Since the simulations were for single detectable species within the community, the parameters N_i , p_i and α_i can be referred to as N , p and α , where these are assumed to be specific to the monitored species.

Model parameters were kept constant throughout the simulation periods. Simulations were carried out for a community composed of 10,000 individuals ($N_T = 10,000$), with an average time between replacement events of 1.44 seconds ($\eta = 0.0004h$). This represents a community growing, for example, in a chemostat with a hydraulic retention time of four hours, as depicted in Figure 3.3. The value of N_T may seem small for a microbial community, however, Sloan et al. (2021) demonstrate that the dynamics in real populations of microbes can often be characterised by an effective community size, which is orders of magnitude smaller than the census community size. The monitored species was assumed to make up 20% of the source community ($p = 0.2$), and the immigration probability was set at 0.1 ($m = 0.1$). In a typically diverse microbial community 0.2 would reflect a relatively high abundance, or common species, however, within wastewater treatment samples, taxa with relative abundances of up to 0.4 have been observed (Ofiteru et al., 2010). An immigration probability of 0.1 is also high and so the local community is well-connected to the source. A high immigration probability was selected to ensure this process was detectable within the simulated time series. For the purely neutral community no selection advantage is given ($\alpha = 0$), while in the community where selection is active, a positive selection of 0.01 is attributed to the species ($\alpha = 0.01$).

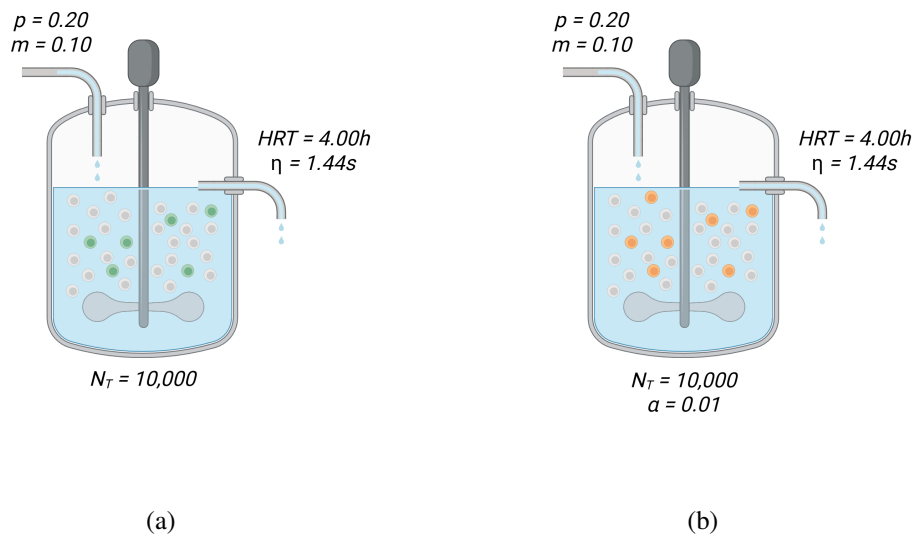


Figure 3.3: Example experimental equivalents to the simulated communities. Created in <https://BioRender.com> a) for Hubbell's model, b) for Sloan et al's model.

A stochastic process is said to be stationary if the probability distribution (in this case of species abundance) is constant in time, and therefore statistics like the mean and variance remain constant. In ecology, a stationary community where all the joint probability distributions of all species abundances stays constant is said to be in equilibrium; also sometimes referred to as dynamic equilibrium because of the random fluctuations. Historically, calibration of neutral models have assumed that the community is in equilibrium and hence the observed distribution of abundances is stationary. Here we simulate time series for stationary communities for two primary reasons, firstly to offer as a comparison to stationary species abundance distributions, and secondly because studies utilising relative abundance time series have done so for stationary communities (Ofiteru et al., 2010). We also simulate communities where the statistics change through time displaying transient (non-equilibrium) dynamics. This might occur when the community is perturbed in some way, for example, naturally through environmental events or synthetically through bioaugmentation. For simulations representing a community in equilibrium, the starting relative abundance values were set at the theoretical mean abundance values. Calculation of these values is provided in Appendix B. To ensure that these starting values did not bias the parameter estimates, especially within time series with short sampling periods, data points from the first 7 days of simulation were ignored. This process allows the starting abundance for each time series to differ and be intrinsic to each simulation. For simulations undertaken over a period of non-equilibrium all data points were considered.

3.3.3 Sampling of simulated single-species relative abundance time series

Rather than re-simulating time series for different periods and temporal resolutions, the choice was made to sample from a single initial set of simulated high-resolution time series for each set of community structuring processes simulated; as defined by Hubbell's and Sloan et al.'s models respectively. This reduces the computational time required to produce multiple time series under differing sampling protocols. Sampling was undertaken uniformly from the start point of the initial time series to a designated end point for a given frequency. In total 28 different sampling protocols were investigated, ranging in periods from 0.004h to 4000h and frequencies between 0.0004h to 400h.

3.3.4 Calibration of NCMs for estimation of individual parameters

Given that the data are re-sampled from the high-resolutions simulated time series, such that there are potentially numerous births and deaths occurring between observation points, calibration is undertaken using the non-truncated continuous equivalent of both models. These take the form of the stochastic differential equations (SDEs):

$$dx_{i,hubbell} = \frac{m(p_i - x_i)}{N_T} d\tau + \sqrt{\frac{2x_i(1 - x_i) + m(p_i - x_i)(1 - 2x_i)}{N_T^2}} dW_\tau \quad (3.3)$$

$$dx_{i,sloan} = \frac{m(p_i - x_i) + 2\alpha(1 - m)x_i(1 - x_i)}{N_T} d\tau + \sqrt{\frac{2x_i(1 - x_i) + m(p_i - x_i)(1 - 2x_i)}{N_T^2}} dW_\tau, \quad (3.4)$$

where x_i is the relative abundance of the monitored i^{th} species, and dx_i is the change in this value over an arbitrary change in number of replacement events (death or emigration followed by birth of immigration) $d\tau$. Parameters N_T , m , p_i and α_i all carry the same definition as before. A full derivation of the stochastic differential equations for each model is given in Appendix C. As we are only considering relative abundance data from a single species we can once again refer to parameters as their species-specific equivalents, such that $dx_i = dx$, $x_i = x$, $p_i = p$ and $\alpha_i = \alpha$.

To estimate individual parameters, maximum likelihood estimation is used. Here this is converted into a minimisation problem by looking for the minimum negative log-likelihood (NLL). The negative log-likelihood is calculated by making negative the sum of the log of the probability density function calculated for each value of dx . Given that W is a standard Weiner process, such that $dW_\tau \approx \sqrt{d\tau}N(0, 1)$, it can be seen that dx is distributed normally, with the mean and variance defined directly from the respective drift and diffusion terms for each of the previously given SDEs. A conversion is also made to express these quantities in real-time by assuming $\tau = t/\eta$, where t is the time between observations and η is the average time between replacement events. From this, the mean and variance are defined as:

$$\mu_{hubbelt} = \frac{m(p-x)}{N_T} \cdot \frac{dt}{\eta} \quad (3.5a)$$

$$\sigma_{hubbelt} = \sqrt{\frac{2x(1-x) + m(p-x)(1-2x)}{N_T^2}} \cdot \sqrt{\frac{dt}{\eta}} \quad (3.5b)$$

$$\mu_{sloan} = \frac{m(p-x) + 2\alpha(1-m)x(1-x)}{N_T} \cdot \frac{dt}{\eta} \quad (3.6a)$$

$$\sigma_{sloan} = \sqrt{\frac{2x(1-x) + m(p-x)(1-2x)}{N_T^2}} \cdot \sqrt{\frac{dt}{\eta}} \quad (3.6b)$$

It is common practice to to truncate the variance terms as demonstrated by Sloan et al. (2006). Although the majority of calibrations in this study are conducted using the full variance terms we do investigate the effect truncating these terms has on calibration accuracy. For this analysis the variance expressions given previously are simplified to:

$$\sigma_{hubbelt} = \sqrt{\frac{2x(1-x)}{N_T^2}} \cdot \sqrt{\frac{dt}{\eta}} \quad (3.7)$$

$$\sigma_{sloan} = \sqrt{\frac{2x(1-x)}{N_T^2}} \cdot \sqrt{\frac{dt}{\eta}} \quad (3.8)$$

The probability density function for dx in each model can be defined by substituting the required identities for μ and σ into the standard form:

$$f(dx | \mu_{hubbelt}, \sigma_{hubbelt}) = \frac{1}{\sqrt{2\pi\sigma_{hubbelt}^2}} \exp\left(-\frac{(dx - \mu_{hubbelt})^2}{2\sigma_{hubbelt}^2}\right) \quad (3.9)$$

$$f(dx | \mu_{sloan}, \sigma_{sloan}) = \frac{1}{\sqrt{2\pi\sigma_{sloan}^2}} \exp\left(-\frac{(dx - \mu_{sloan})^2}{2\sigma_{sloan}^2}\right) \quad (3.10)$$

The negative log-likelihood is thus determined by:

$$NLL_{hubbell} = -\sum \log f(dx | \mu_{hubbell}, \sigma_{hubbell}) \quad (3.11)$$

$$NLL_{sloan} = -\sum \log f(dx | \mu_{sloan}, \sigma_{sloan}) \quad (3.12)$$

Optimisation towards the minimum value is achieved by altering the values of unknown parameters using the Nelder-Mead simplex algorithm. The choice is made to use this algorithm due to its accessibility and efficiency when optimising a small number of parameters. While other algorithms have the potential to provide a more robust optimisation the exploration of optimisation algorithms is outwith the scope of this study. The optimisation calculations are conducted computationally using the MATLAB *mle* function. By defining the probability density functions in terms of the distinct model parameters, estimations of each individual parameter can be obtained. This differs from other methods in which estimations of compound parameters are determined (Latimer et al., 2005; Ofițeru et al., 2010; Sloan et al., 2021).

All optimisations were carried out under the same conditions. Estimation limits for each parameter were set at their theoretical maximums and minimums. These are available in Table 3.1 alongside the starting values for each parameter. These were set to the values used during simulation of the time series and thus offer the optimisation the easiest path to the correct estimation. The optimisation tolerance was set at the default value of 1×10^{-6} . Maximum function evaluations and iterations were set to ensure that, in all cases, the optimisation was terminated upon meeting the tolerance criterion. This was achieved using a value of 1×10^6 .

Table 3.1: Starting values, upper limits and lower limits set within the optimisation for estimating parameters N_T , η , m , p and α .

Parameter	Description	Actual value	Starting value	Upper limit	Lower limit
N_T	Community size	10000	10000	<i>Inf</i>	0
η	Time between replacement events (hours)	4×10^{-4}	4×10^{-4}	<i>Inf</i>	0
m	Immigration rate	0.1	0.1	1	0
p	Relative abundance within immigrating community	0.2	0.2	1	0
α	Local selection advantage	0.01	0.01	1	-1

3.4 Results and discussion

3.4.1 Calibration of neutral models to idealised single species relative abundance time series

Firstly we assess the structural identifiability of model parameters from single-species relative abundance time series by calibrating both models with idealised datasets, where every discrete birth, death and immigration event is explicitly represented. Here we enforce that within the data that there are no sampling errors and that the community sampled is in equilibrium under the entire sampling period. This equilibrium is defined as when fluctuations in species abundance proceed around a stationary mean, such that the expected value of the time series does not change over time.

To create a very rich data set one hundred realisations of these idealised time series of species abundance are generated. For both Hubbell's purely neutral and Sloan et al.'s near-neutral models the model parameters N_T , η , m and p were kept constant at values of 10,000, $4.00 \times 10^{-4}\text{h}$, 0.10 and 0.20 respectively. The additional parameter in Sloan et al.'s model (α) was set to 0.01, giving the species a selective advantage over other members of the community. The duration of the time series were one-year, or 8760 hours, and since every replacement was simulated (every 0.0004h) they comprised 21.9×10^6 data points. The experimental equivalent of this simulation regime is illustrated by Figure 3.4, where either 100 bioreactors are operating in parallel, or an experiment within a single bioreactor is repeated 100 times. For the real-world such a setup would be extremely costly, even if achieved, observations with the same resolution and duration provided by the simulations would be impossible using existing experimental methods.

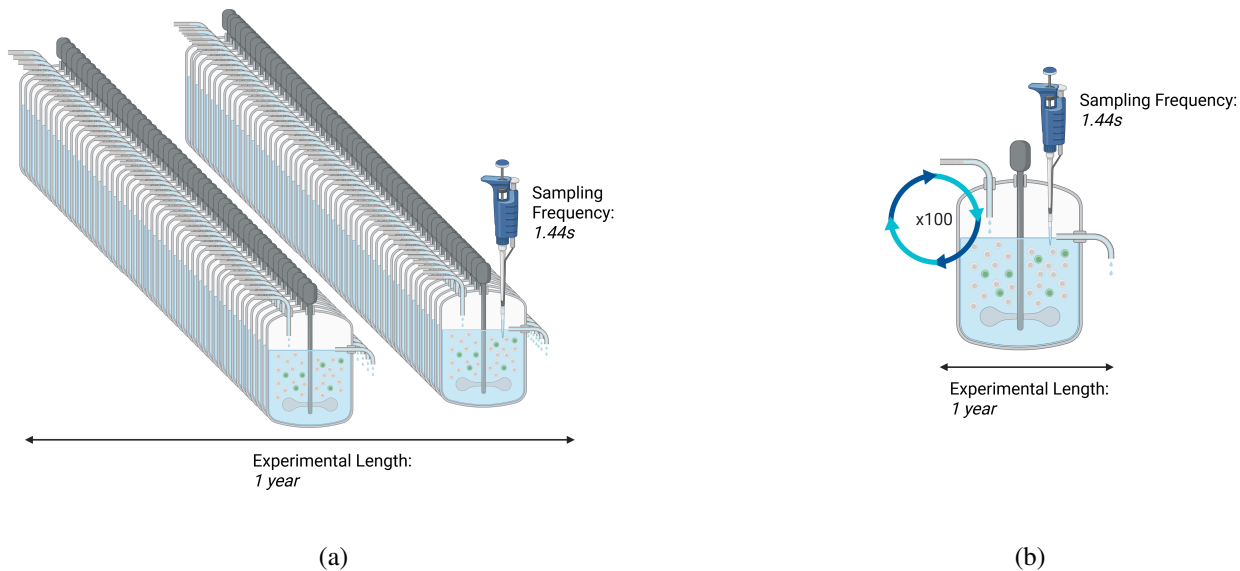


Figure 3.4: Example experimental equivalents to the simulation regime employed. Created in <https://BioRender.com> a) 100 bioreactors operating in parallel, b) a single bioreactor repeated 100 times

In this section the one hundred time series are used in a number of different ways. First, the inherent variability across the time series is characterised. This shows that, even though the time series may be perceived to be long and high resolution, they are not sufficiently so that the variance is constant from series to series. It is then shown that this has a knock-on effect on the accuracy of the parameters values estimated by calibration. Thus, even with high resolution

data there is enough variability between series to make accurately identifying parameters difficult. The analysis also goes on to highlight that using single-species relative abundance time series captured during a period of equilibrium does not resolve correlated parameters with the models. Therefore, extra effort should be made to determine certain parameters *a priori* by, for example, producing knowledge of the system from previous independent experiments. By reducing the dimensionality of the parameter space and resolving parameter correlations, identification of the remaining parameters becomes trivial when using these idealised datasets.

3.4.1.1 Calibration of Hubbell's purely neutral model

The one hundred high resolution realisations of Hubbell's model are presented in Figure 3.5a. Given that the simulations start close to the mean, p , the simulations should be a good representation of the stationary stochastic process and this can be seen in the frequency distributions plotted in Figure 3.5b. Given the resolution of the times series, one might expect the mean and variance of each of the one hundred realisations to be the same. Table 3.2 demonstrates that there is still variability in these statistics between series. Indeed, the standard deviation in variance is only one order of magnitude smaller than its average value, which suggests an even longer time series is required for the process to be truly stationary and to capture the true variance with certainty.

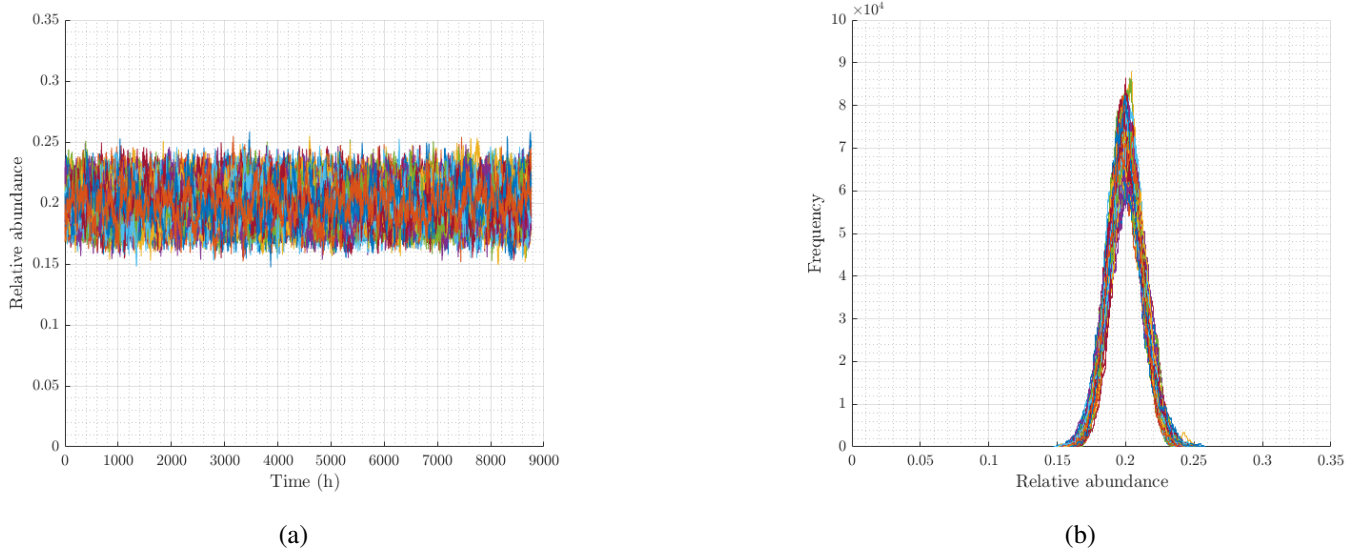


Figure 3.5: Relative abundance time series of a single species obtained through 100 realisations of a simulated community. Community dynamics were simulated from the discrete transitional probabilities described by Hubbell (2001) for a community governed by ecological drift alone (Equation 3.1). Parameter values used were: $N_T = 10,000$, $\eta = 4.00 \times 10^{-4}$, $m = 0.10$ and $p = 0.20$. Sampling was conducted after every individual replacement event (every 0.0004 hours) over a period of 8,760 hours. a) Time series plot, b) frequency plot.

Attempts are now made to retrieve the parameters used in the simulation by calibrating the continuous version of Hubbell's model for each of the realisations under differing scenarios. Under a naive scenario, an attempt is made to estimate all parameters within the model. This is akin to having an experimental set-up where none of the parameters can be fixed or known and where model calibration is the sole source of information on the parameters. This can be the case, particularly with field studies of natural systems, where there is no human intervention. In a more controlled experiment,

Table 3.2: Summary table of mean and standard deviation in statistics calculated across all 100 single-species relative abundance time series portrayed in Figure 3.5.

Statistic	Time Series	
	Mean	SD
Mean	0.200	0.001
Variance	1.585×10^{-4}	0.154×10^{-4}
Skewness	0.061	0.128

it is often possible to fix at least some of the experimental variables, such as the size of the community, the replacement rate or the immigration rate, or, indeed, one might have prior knowledge of one of these variables. Thus, only a subset of the parameters need to be estimated by calibration to define the full model. Thus, here Hubbell's model is calibrated for all combinations of fixed parameters.

The box plots in Figure 3.6 illustrate the parameter estimates obtained from calibrating Hubbell's model to all one hundred time series shown in Figure 3.5. This is done for all scenarios and thus all levels of *a priori* system knowledge, with the known/fixed parameters displayed along the x-axis. The y-axis on the left-hand side of each plot shows the estimated value obtained through calibration, while the y-axis on the right-hand side displays the equivalent percent error calculated based on the value of that parameter used within the simulation of the time series. To summarise these plots the corresponding mean percent error (MPE) and coefficient of variation (CV) for each suite of estimations are calculated. Here MPE provides a measure of bias in the estimates relative to the true value and CV provides the spread of estimates scaled by the mean. Since MPE and CV are scaled measures they facilitate direct comparison between parameters, despite differences in the magnitudes of these parameters. Plots displaying the calculated MPEs and CVs are provided in Figures 3.7a and 3.7b respectively.

Figures 3.6c and 3.7b demonstrate that there is a generic, intrinsic difficulty in estimating the immigration rate m . The results demonstrate that under every knowledge regime (combination of predetermined and calibrated parameters) there is variation between the estimates of m , even when these idealised time series are used. Although a maximum MPE of only 1% is observed (Figure 3.7a), a minimum CV of 8% holds true despite the number of fixed parameters (Figure 3.7b). Thus, under this experimental design, any m calibrated from a single realisation is likely to be erroneous.

If the variance in the simulated time series is underestimated then the fitting algorithm tends to overestimate m , the opposite is also true. This is especially the case when N_T or η is fixed. In Hubbell's model, the rate of immigration (m) drives the relative abundance of species towards that of the source community (p). This is implicit in the drift term $m(p - x)$; if the abundance of the monitored species, x , drifts above p , then the drift term is negative and pulls the next realisation in the time series back towards p and if it drops below p the drift term is positive and, similarly, draws the time series back towards p . Thus, if the simulated variance is smaller than the theoretical, then the time series does not deviate from p as much as expected and it is as if immigration is having a stronger pull towards p ; this is reflected in a higher calibrated m . Similarly, if the simulated variance is larger than the theoretical, it is as if immigration is weaker, allowing the time series to move further away from p and so the calibrated m is lower than the theoretical. Thus the estimation of m is related to how representative the variance in the realisation is.

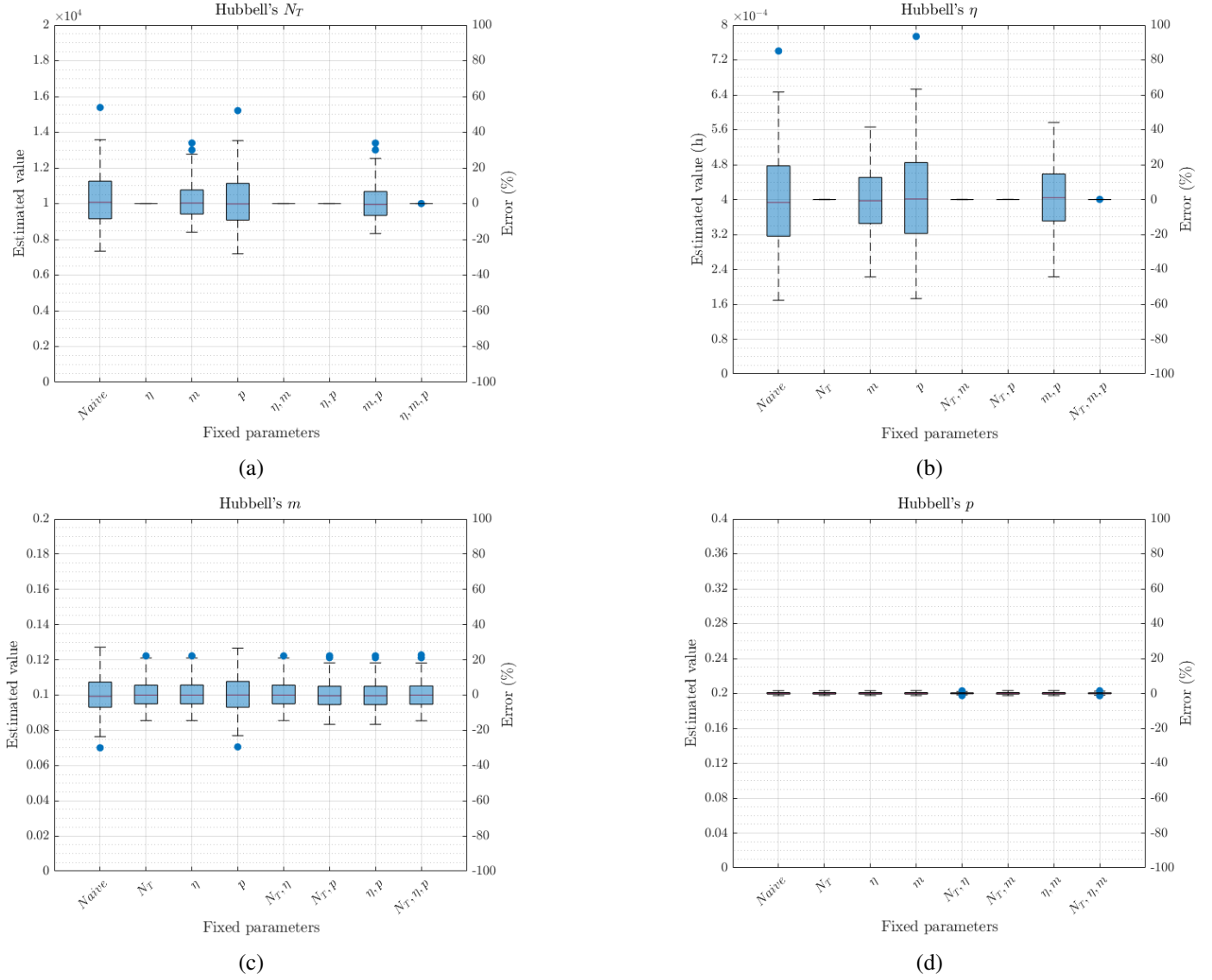


Figure 3.6: Box plots representing the estimation of distinct parameters from Hubbell's purely neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.5. a) Estimation of N_T , b) estimation of η , c) estimation of m . d) estimation of p .

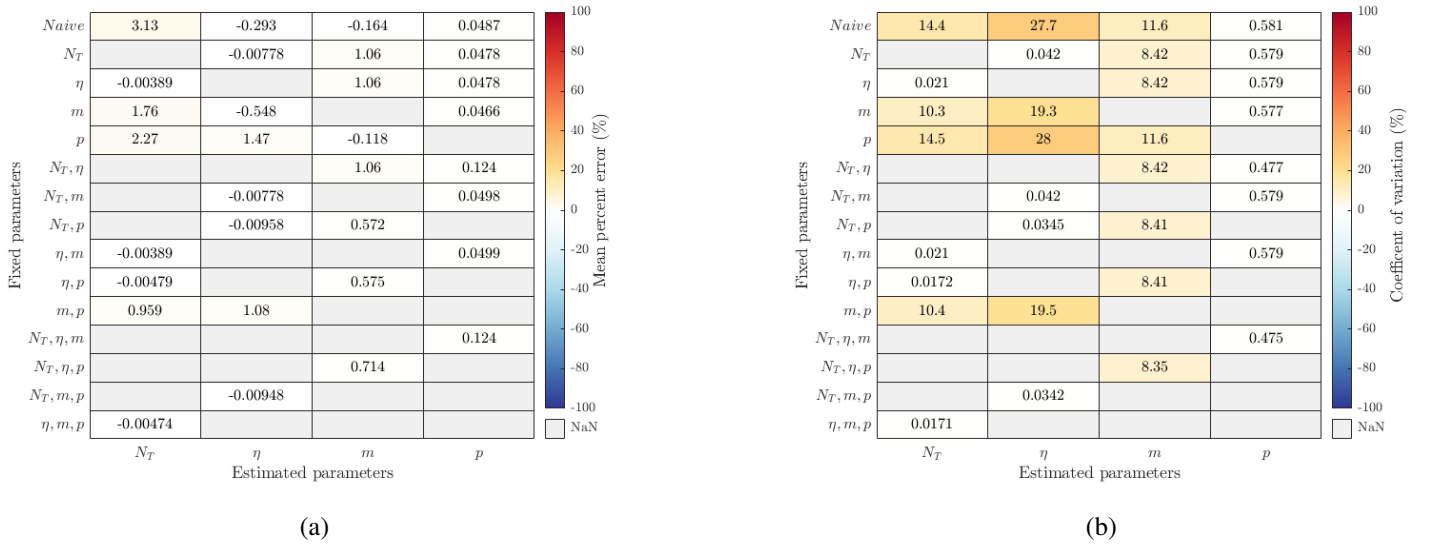


Figure 3.7: Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of Hubbell's purely neutral parameters shown in Figure 3.6. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs.

For parameters N_T and η robust calibration is observed only when either of these parameters are fixed (Figures 3.6a 3.6b and 3.7). The interdependence of N_T and η can be revealed by considering the negative log-likelihood function. Ideally, this multidimensional function would contain a well-defined global minimum, described by a unique set of optimal parameters. However, for the Hubbell (and Sloan et al.) model this is not the case. If we hold m and p constant at their theoretical values then the negative log-likelihood function is a two dimensional function of N_T and η as shown in Figure 3.8. Rather than a single definitive minimum, there is a region, in dark blue, where the function is at a perceived constant minimum. A centre line, running through this region traces out the relationship $\eta N_T^2 \approx 4 \times 10^4$.

For individual calibrations the log-likelihood will inevitably alight on a best-fit parameter pair for N_T and η . Plotting these values against each other, for all naive calibrations, we find that they all lie on the centre line, however, the exact position varies (Figure 3.8b). Thus, for each individual calibration a single minimum, albeit an extremely shallow one, must exist, driving the calibration away from the starting values used within the optimisation. This focussing of the negative log-likelihood to different points, is a result of the inclusion of m and p within the diffusion term of the model calibrated. Indeed, if the same analysis is conducted using the truncated version of Hubbell's model, which does not contain m and p within this term, all naive estimations of N_T and η lie on the exact same point of the centre line (Appendix D) demonstrating that they are structurally non-identifiable (Wieland et al., 2021). As p is shown to be estimated robustly for all calibrations undertaken (Figure 3.6d), we can conclude that the observed differences in the exact location of the minimum must be predominantly driven by m .

A further relation exists between N_T , η and m . Indeed, a change in errors is observed when m is fixed and both N_T and η are unknown (Figures 3.6a 3.6b and 3.7). Again, the errors observed can be explained through the differences within the variance of each time series. It has been established that unrealistically long simulations would be required to accurately capture the model's theoretical variance. And, therefore, the variance of the time series here often differ. Thus,

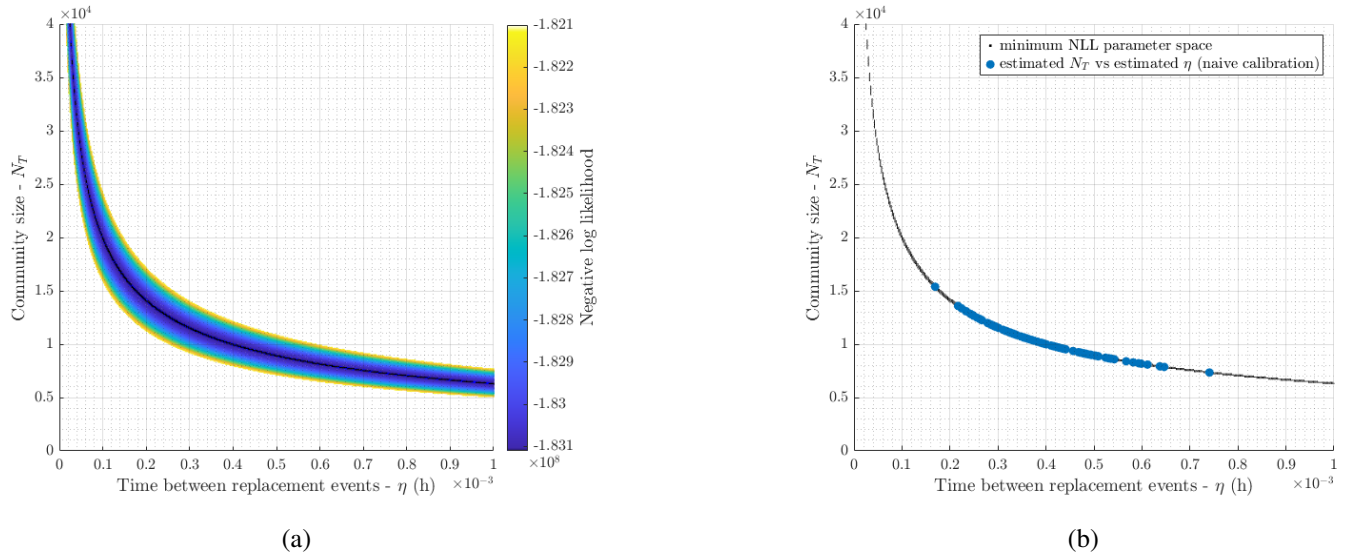


Figure 3.8: a) Negative log-likelihood (NLL) as a function of N_T and η . The value of N_T is given on the y-axis with the value of η on the x-axis. Corresponding values for NLL are displayed as a heat map. All NLL values are calculated by fitting Hubbell’s model to a single time series obtained from a simulated community governed by drift alone. During acquisition of these values all other parameters within the model were fixed at their expected values. All NLL values shown are within a range of approximately 0.55% of the minimum value. b) Estimated parameter pairs of N_T and η from naive calibration plotted on top of the centre line of minimum NLL.

if the simulated variance is incorrect and m is fixed at its theoretical value the optimisation will yield incorrect values for N_T and η in an attempt to capture the variance. This process does act, once again, to focus the negative log-likelihood to a global minimum dependent on the observed variance.

So, from an experimental design perspective, even if you could collect extremely high-resolution data on a system that you believed to be undergoing long-term neutral population dynamics, it is difficult to calibrate the correlated parameters N_T and η with any degree of confidence through optimisation alone when using single-species relative abundance time series captured under equilibrium. If one of these parameters has been independently measured, or constrained, in the experimental setup, then the other can be accurately estimated through calibration with relative ease, assuming use of sufficiently high resolution data. Note, what constitutes a sufficiently high resolution is explored later. When simultaneous calibration of N_T and η is unavoidable, the error observed will be related to the variance captured. We have shown that capturing the expected variance, even when using high-resolution data, is difficult to achieve. An option to reduce this error, when there is no alternative to calibration, is to perturb the system, as far away from equilibrium as possible, such that it is not only the diffusion term that dominates the dynamics and, so, the likelihood function for the time series of abundance is altered. This is explored later.

Calibration appears to yield accurate estimates of p , independent of the scenario, as highlighted in Figure 3.6d. Indeed, the MPEs (Figure 3.7a) and CVs (Figure 3.7b) are maintained at under 1% across all estimations of p . In Hubbell’s description of community structure, p represents the relative abundance of the monitored species within the immigrating, or source, community. For a neutral community at equilibrium, the relative abundance will fluctuate around a stationary mean, for Hubbell’s model this is defined solely by p , as demonstrated in Appendix B. Given that the simulated time

series are acquired over a long enough period such that the variation in relative abundance is Beta distributed (Figures 3.5a and 3.5b), determining this value is a straightforward endeavour.

Large errors in the estimation of p can exist, as alluded to by Etienne et al. (2006). This occurs when m tends to zero. Under such a scenario, the effect of immigration is so small that the influence of p is insignificant, making its inference increasingly difficult. This represents a unique difficulty when calibrating neutral models to populations with near-zero immigration.

3.4.1.2 Calibration of Sloan et al.'s near-neutral model

The second model analysed is Sloan et al.'s near-neutral model (Sloan et al., 2006), which includes an additional parameter α , used to quantify the magnitude of selection experienced by the monitored species relative to all other species within the community. Again, one hundred year-long time series were produced and sampled at intervals of 0.0004 hours, these are plotted in Figure 3.9. Comparison of Figure 3.9 and Table 3.3 with Figure 3.5 and Table 3.2 indicates an increase in mean and variance is induced as a result of the local advantage given to the monitored species.

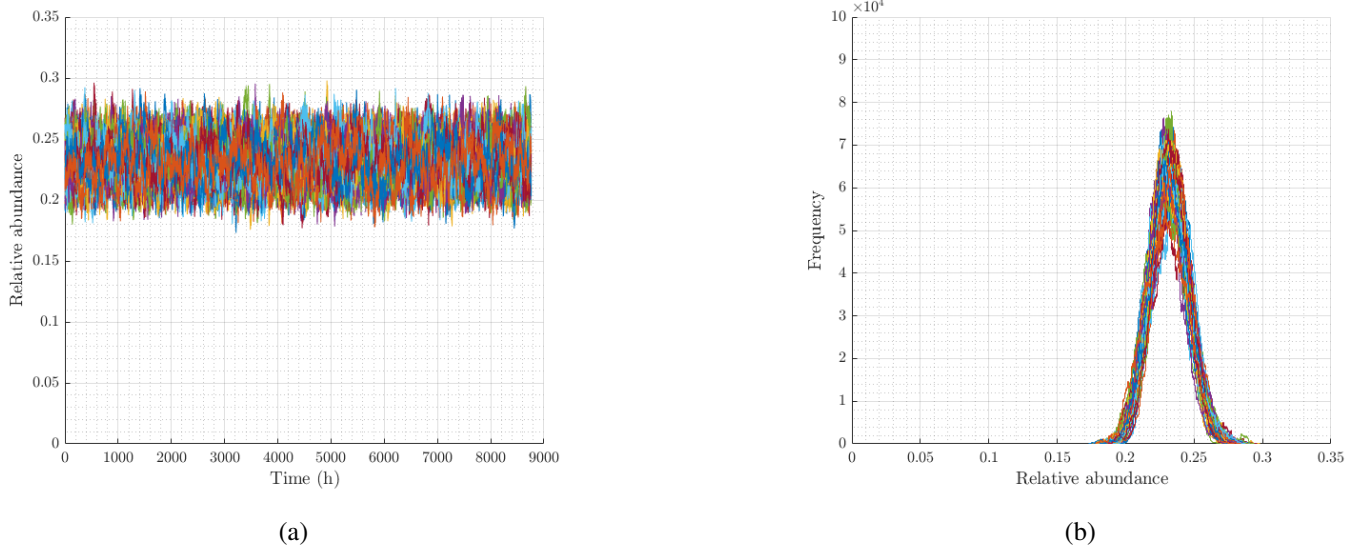
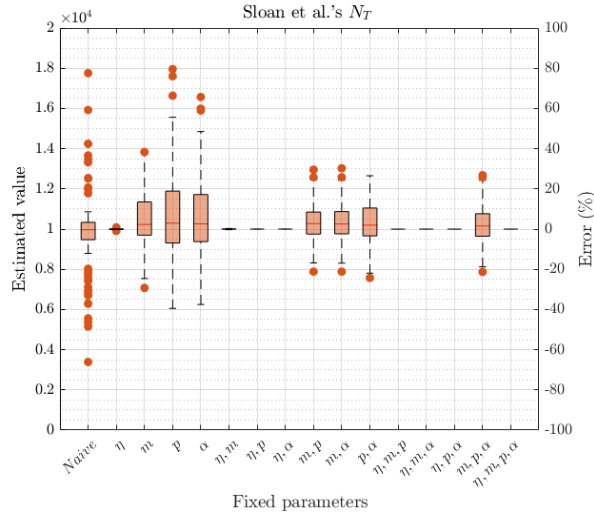
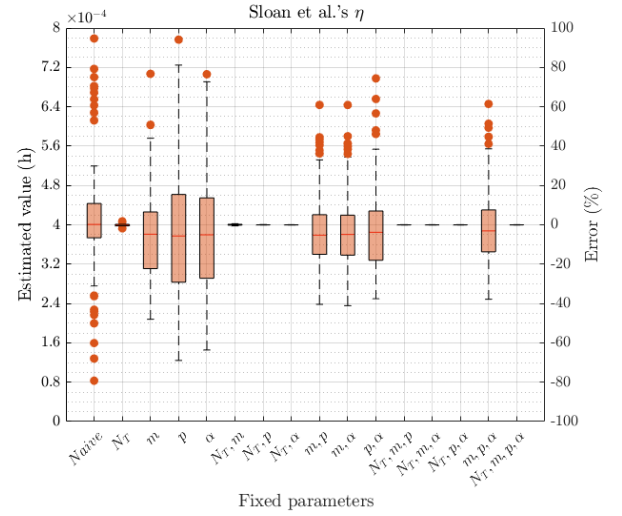


Figure 3.9: Relative abundance time series of a single species obtained through 100 realisations of a simulated community. Community dynamics were simulated from the discrete transitional probabilities described by Sloan et al. (2006) for a community governed by ecological drift and selection (Equation 3.2). Parameter values used were: $N_T = 10,000$, $\eta = 4.00 \times 10^{-4}$, $m = 0.10$, $p = 0.20$ and $\alpha = 0.01$. Sampling was conducted after every individual replacement event (every 0.0004 hours) over a period of 8,760 hours. a) Time series plot, b) frequency plot.

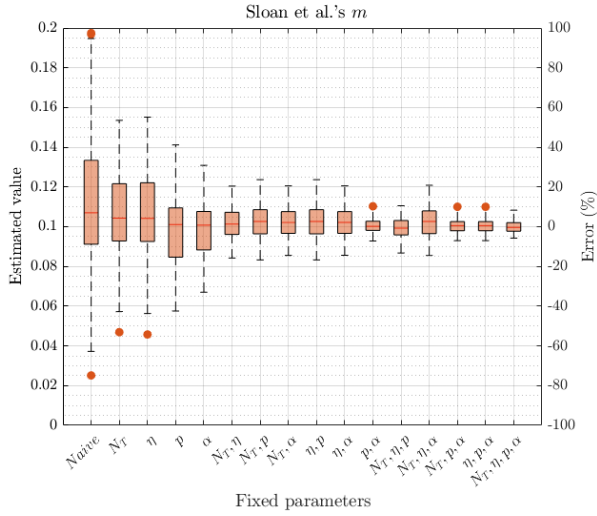
Similarly to Hubbell's model, Sloan et al.'s model is calibrated to all one hundred time series for all scenarios (combinations of fixed and free parameters), box plots summarising the estimations of each parameter are given in Figure 3.10 with the corresponding mean percent error and coefficients of variation given in Figure 3.11. In the first instance, Figures 3.10a and 3.10b indicate similar behaviour to that observed for Hubbell's model, in which accurate estimation of N_T or η is achieved when the other is fixed. As previously demonstrated this is a result of the area of similar likelihood expressed in Figure 3.8.



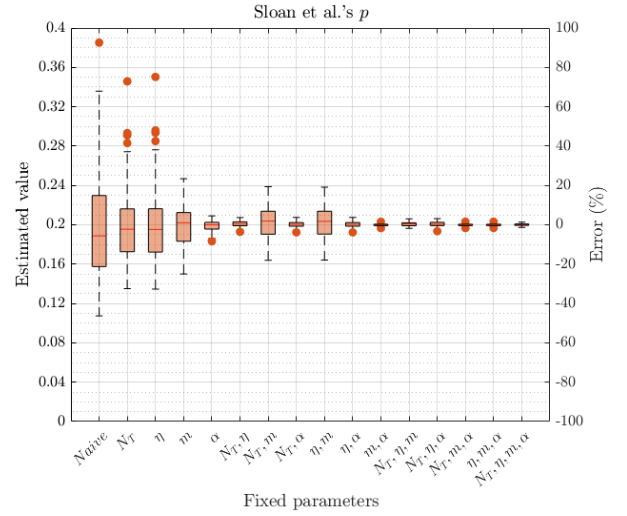
(a)



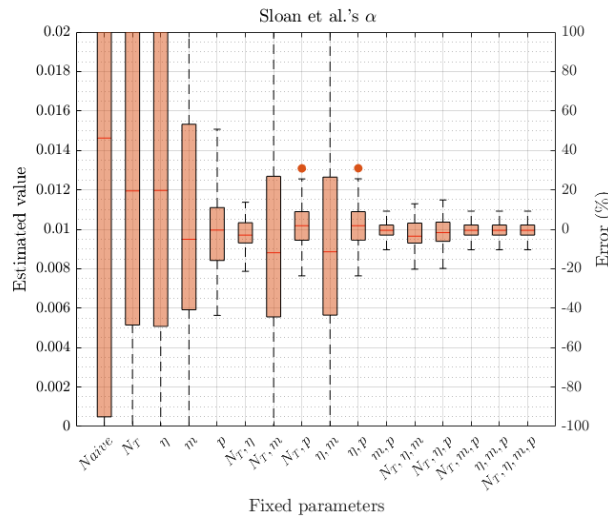
(b)



(c)



(d)

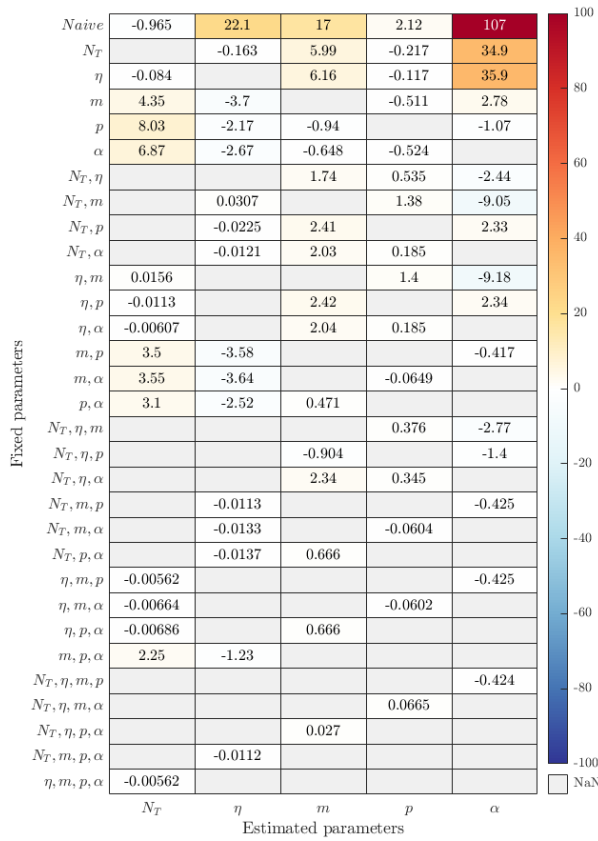


(e)

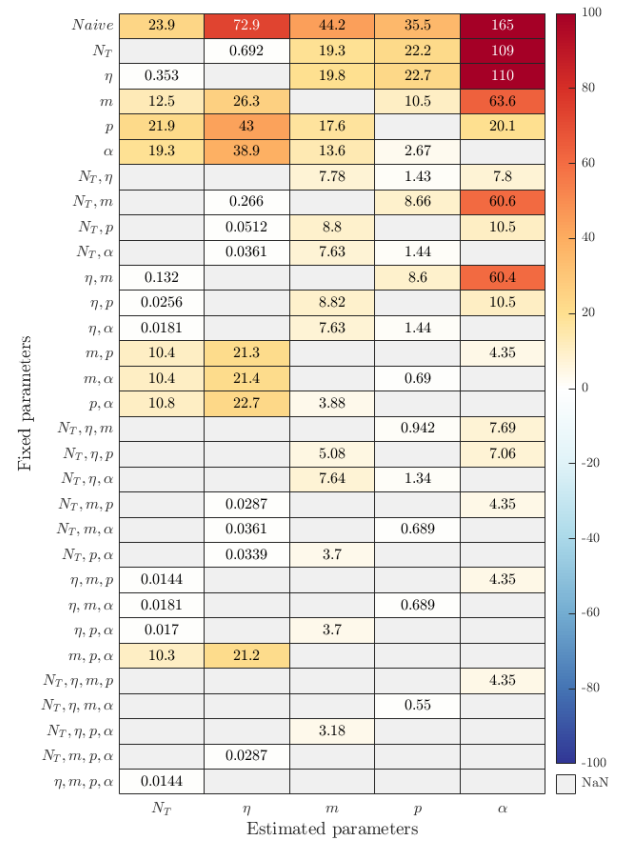
Figure 3.10: Box plots representing the estimation of distinct parameters from Sloan et al.'s near-neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.9. a) Estimation of N_T , b) estimation of η , c) estimation of m , d) estimation of p , e) estimation of α

Table 3.3: Summary table of mean and standard deviation in statistics calculated across all 100 single-species relative abundance time series portrayed in Figure 3.9. Expected values for these statistics based on the parameters used within the simulation of these time series are also provided.

Statistic	Time Series	
	Mean	SD
Mean	0.232	0.002
Variance	1.914×10^{-4}	0.203×10^{-4}
Skewness	0.073	0.143



(a)



(b)

Figure 3.11: Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of Sloan's near-neutral parameters shown in Figure 3.10. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs.

Comparing Figures 3.10c and 3.10d to their counterparts obtained from calibrating Hubbell's model (Figures 3.6c and 3.6d respectively), highlights that the addition of a single parameter between models, in this case local selection, leads to further dependencies and thus a requirement of greater *a priori* knowledge. Indeed, p can no longer be accurately estimated under the naive scenario. In the purely neutral model, the mean of abundance in the local community for a species is equal to p . In the near neutral model, if the species has an advantage, $\alpha > 0$ then the mean is a function of additional parameters, as demonstrated in Appendix B. Given that the variance in abundance is also a function of these, the possibility of compensations between parameters yielding the same mean and variances is significantly increased. Thus, there is not necessarily a unique parameter set for the time series.

3.4.1.3 Truncated models

Within the derivation of the SDEs describing the continuous version of both models, it is common to remove $m(p-x)(1-2x)$ from the diffusion term; a step not taken within our derivations. This is done because m is often so small as to negate the effect this term has on random fluctuations, and even if m is large then x tends to p making $p-x$ small, again reducing the effect of this term (Sloan et al., 2006). However, while removing m and p from the diffusion term can make solving the SDEs easier, it does impact the calibration of distinct neutral model parameters.

If we truncate both models, such that the numerator within the diffusion term is defined by $2x(1-x)$, and the same analysis as before is carried out, it is shown that, for Hubbell's model, a less accurate estimation of m is produced under scenarios where either N_T or η is fixed (Appendix D). For Sloan et al.'s model this effect is extended to p and α (Appendix E). In addition, as described previously, the area of minimum negative log-likelihood between N_T and η (Figure 3.8) indeed becomes constant and thus estimates are solely driven by the path the optimisation takes in reaching the centre line of minimum negative log-likelihood. This makes estimates of N_T and η , where both are unknown, highly dependent on the starting value used for parameter estimates.

Under a truly naive calibration in which none of the parameters can be measured and starting values of the optimisation are likely to be some distance away from the correct values, then the truncated versions of both models will perform worse when estimating each parameter distinctly. Indeed, the removal of $m(p-x)(1-2x)$ from the diffusion term is a more nuanced decision than originally observed. As such, the rationale behind the use of either the truncated or non-truncated versions of Hubbell's and Sloan et al.'s models should be provided when defining a methodology.

3.4.2 Calibration of neutral models to non-idealised single species relative abundance time series

3.4.2.1 The effects of duration and frequency of sampling on neutral model identification

Thus far, model calibration has been performed using time series recorded over a large period and produced with the highest temporal sampling resolution possible. To achieve this degree of resolution, a species' abundance must be sampled after every distinct replacement event. For most natural or engineered communities of importance, enumerating the abundance after every replacement is impossible to achieve, even using the most up-to-date molecular methods in microbiology. Even if expectationally high sampling resolution can be obtained, it would be extremely expensive and time intensive to conducted sampling at this rate over a long period. For this reason, it is necessary to quantify how the frequency and duration (or period) of a sampling regime affect the calibration of model parameters and thus analyse the practical identifiability of the model parameters.

To achieve this, the idealised time series are shortened to 10x, 100x, 1,000x, 10,000x, 100,000x, 1,000,000x and 10,000,000x the time of a single replacement, resulting in overall periods of 0.004, 0.04, 0.4, 4, 40, 400 and 4,000 hours respectively. Each set of shortened time series is then sampled using frequencies of 0.0004, 0.004, 0.04, 0.4, 4, 40 and 400 hours where possible, ensuring at least one set of sampling regimes with the same total number of samples are obtained. Employing these sampling regimes results in 28 sets of 100 time series for both the neutral and near-neutral population dynamics, this is akin to repeating the experimental setups depicted in Figure 3.4 28 times using differing sampling regimes. The broad range of time series lengths and resolutions assessed offers insight into datasets that it might be feasible to collect using different experimental setups or measurement methods. For example, a sampling regime with a short period and high sampling resolution may reflect experiments undertaken with equipment such as a microfluidic device, while regimes with longer periods and less frequent sampling are more attributed to field-scale studies.

Tables 3.4 and 3.5 demonstrate how the mean, variance and skewness across each set of one hundred time series are affected by introducing these new sampling regimes. Table 3.4 shows how these statistics are affected by time series length (or sampling period) when sampling frequency is fixed at the maximum. Although the mean value calculated from each set of one hundred time series is stable, the standard deviation in this mean increases as time series length is reduced. The largest increase is observed when decreasing the sampling period from 400 hours to 40 hours, with the standard deviation at least doubling in magnitude. This trend is also observed for the skewness calculated for each set of one hundred time series. The mean of this quantity stays relatively consistent, however, its standard deviation increases as sampling period is decreased. The largest increase is observed between the shortest periods analysed, demonstrating an exponential rise in the divergence from the previous Beta distributions (Figures 3.5b and 3.9b) as the sampling period tends to zero.

In contrast, the mean of variance calculated for each set of one hundred time series is shown to significantly reduce as the sampling period is shortened. This reduction is in orders of magnitude with the initial mean of variances across the set of time series sampled over 4,000 hours recorded at 1.581×10^{-4} and 1.928×10^{-4} , for time series produced using Hubbell's and Sloan et al.'s models respectively. For a sampling period of 4 hours these are reduced to 4.303×10^{-6} and 5.847×10^{-6} , and for a period of 0.004 hours values of 6.022×10^{-9} and 6.510×10^{-9} are recorded. Variation in the variance is also shown to reduce at a similar rate. This demonstrates that a systematic under-representation of the expected variance in relative abundance values is induced as a result of decreasing the sampling period.

Table 3.4: Summary table of mean and standard deviation in statistics across different sets of 100 relative abundance time series sampled over different sampling periods. Sampling frequency is held constant at a value of 0.0004 hours across all time series. Values are shown for sets of time series simulated according to Hubbell's and Sloan et al.'s models.

Statistic	Sampling period (h)	No. of samples	Hubbell		Sloan et al.	
			Mean	SD	Mean	SD
Mean	4000	10000001	0.200	0.002	0.232	0.002
	400	1000001	0.200	0.004	0.232	0.006
	40	100001	0.199	0.010	0.231	0.012
	4	10001	0.201	0.012	0.231	0.014
	0.4	1001	0.201	0.013	0.231	0.014
	0.04	101	0.201	0.013	0.232	0.014
	0.004	11	0.201	0.013	0.232	0.014
Variance	4000	10000001	1.581×10^{-4}	0.212×10^{-4}	1.928×10^{-4}	0.267×10^{-4}
	400	1000001	1.272×10^{-4}	0.588×10^{-4}	1.685×10^{-4}	0.803×10^{-4}
	40	100001	3.993×10^{-5}	3.164×10^{-5}	4.220×10^{-5}	3.446×10^{-5}
	4	10001	4.303×10^{-6}	2.578×10^{-6}	5.847×10^{-6}	4.372×10^{-6}
	0.4	1001	5.111×10^{-7}	4.479×10^{-7}	6.004×10^{-7}	5.763×10^{-7}
	0.04	101	4.941×10^{-8}	4.614×10^{-8}	5.477×10^{-8}	4.746×10^{-8}
	0.004	11	6.022×10^{-9}	5.002×10^{-9}	6.510×10^{-9}	5.180×10^{-9}
Skewness	4000	10000001	0.055	0.169	0.090	0.198
	400	1000001	0.114	0.363	0.124	0.387
	40	100001	0.048	0.490	0.045	0.433
	4	10001	-0.044	0.416	-0.029	0.488
	0.4	1001	0.048	0.486	0.006	0.466
	0.04	101	-0.010	0.510	0.028	0.506
	0.004	11	0.020	0.775	0.002	0.804

Table 3.5: Summary table of mean and standard deviation in statistics across different sets of 100 relative abundance time series sampled using different sampling frequencies. Sampling period is held constant at a value of 4000 hours across all time series. Values are shown for sets of time series simulated according to Hubbell's and Sloan et al.'s models.

Statistic	Sampling frequency (h)	No. of samples	Hubbell		Sloan et al.	
			Mean	SD	Mean	SD
Mean	0.0004	10000001	0.200	0.002	0.232	0.002
	0.004	1000001	0.200	0.002	0.232	0.002
	0.04	100001	0.200	0.002	0.232	0.002
	0.4	10001	0.200	0.002	0.232	0.002
	4	1001	0.200	0.002	0.232	0.002
	40	101	0.200	0.002	0.232	0.002
	400	11	0.200	0.004	0.232	0.005
Variance	0.0004	10000001	1.581×10^{-4}	0.212×10^{-4}	1.928×10^{-4}	0.267×10^{-4}
	0.004	1000001	1.581×10^{-4}	0.212×10^{-4}	1.928×10^{-4}	0.267×10^{-4}
	0.04	100001	1.581×10^{-4}	0.212×10^{-4}	1.928×10^{-4}	0.267×10^{-4}
	0.4	10001	1.581×10^{-4}	0.212×10^{-4}	1.928×10^{-4}	0.267×10^{-4}
	4	1001	1.583×10^{-4}	0.213×10^{-4}	1.929×10^{-4}	0.267×10^{-4}
	40	101	1.575×10^{-4}	0.236×10^{-4}	1.936×10^{-4}	0.296×10^{-4}
	400	11	1.662×10^{-4}	0.688×10^{-4}	1.948×10^{-4}	0.932×10^{-4}
Skewness	0.0004	10000001	0.055	0.169	0.090	0.198
	0.004	1000001	0.055	0.169	0.090	0.198
	0.04	100001	0.055	0.169	0.090	0.198
	0.4	10001	0.055	0.169	0.090	0.198
	4	1001	0.053	0.171	0.093	0.198
	40	101	0.017	0.239	0.120	0.231
	400	11	0.074	0.519	0.051	0.497

Table 3.5 demonstrates changes within the same quantities as a consequence of increasing the time between samples. The duration of the time series are fixed over a 4,000-hour period. The results within Table 3.5 demonstrate that sampling frequency has little influence on the mean, variance and skewness exhibited by each time series. Only when a sampling frequency of 40 hours is employed is there any noticeable change in the values of each quantity. Even within the lowest of resolutions analysed significant changes are only observed within the variation of each statistic. This is, therefore, likely a consequence of a reduction of the total number of samples rather than an under representation of phenomena due to the increased length between samples. The analysis does suggest that the mean, variance and skewness within the time series produced here will be accurately portrayed if the number of total samples is higher than an optimal value which lies between 100 and 1000.

To determine the impact that changes in the sampling regime have on the ability to calibrate model parameters, both models are calibrated to the sampled time series. Calibration of both models is undertaken using a scenario that facilitates structural identifiability, this ensures that errors within each parameter can be attributed in large to the sampling regime used. Tables 3.6 and 3.7 provides the scenario (fixed parameters) used to identify Hubbell's and Sloan et al.'s models. Heat maps displaying the MPEs and CVs summarising the accuracy of parameter estimates to each set of 100 sampled time series for either model, are given in Figures 3.12, 3.13, 3.14 and 3.15c. For all figures, the y-axis portrays the changes in sampling frequency and the x-axis portrays changes in sampling period. A colour is given to each box demonstrating accuracy from +100% to -100%, any errors outwith this range are considered grossly inaccurate. In many cases errors do exceed this range and thus exact values are provided within the boxes themselves.

Table 3.6: Summary of parameters fixed during the calibration of parameters within Hubbell's neutral model during the analysis of sampling regimes.

Estimated parameter	Fixed parameters
N_T	η, p
η	N_T, p
m	η, p
p	η, m

Table 3.7: Summary of parameters fixed during the calibration of parameters within Sloan et al.'s near-neutral model during the analysis of sampling regimes.

Estimated parameter	Fixed parameters
N_T	η, m, p
η	N_T, m, p
m	η, p, α
p	η, m, α
α	η, m, p

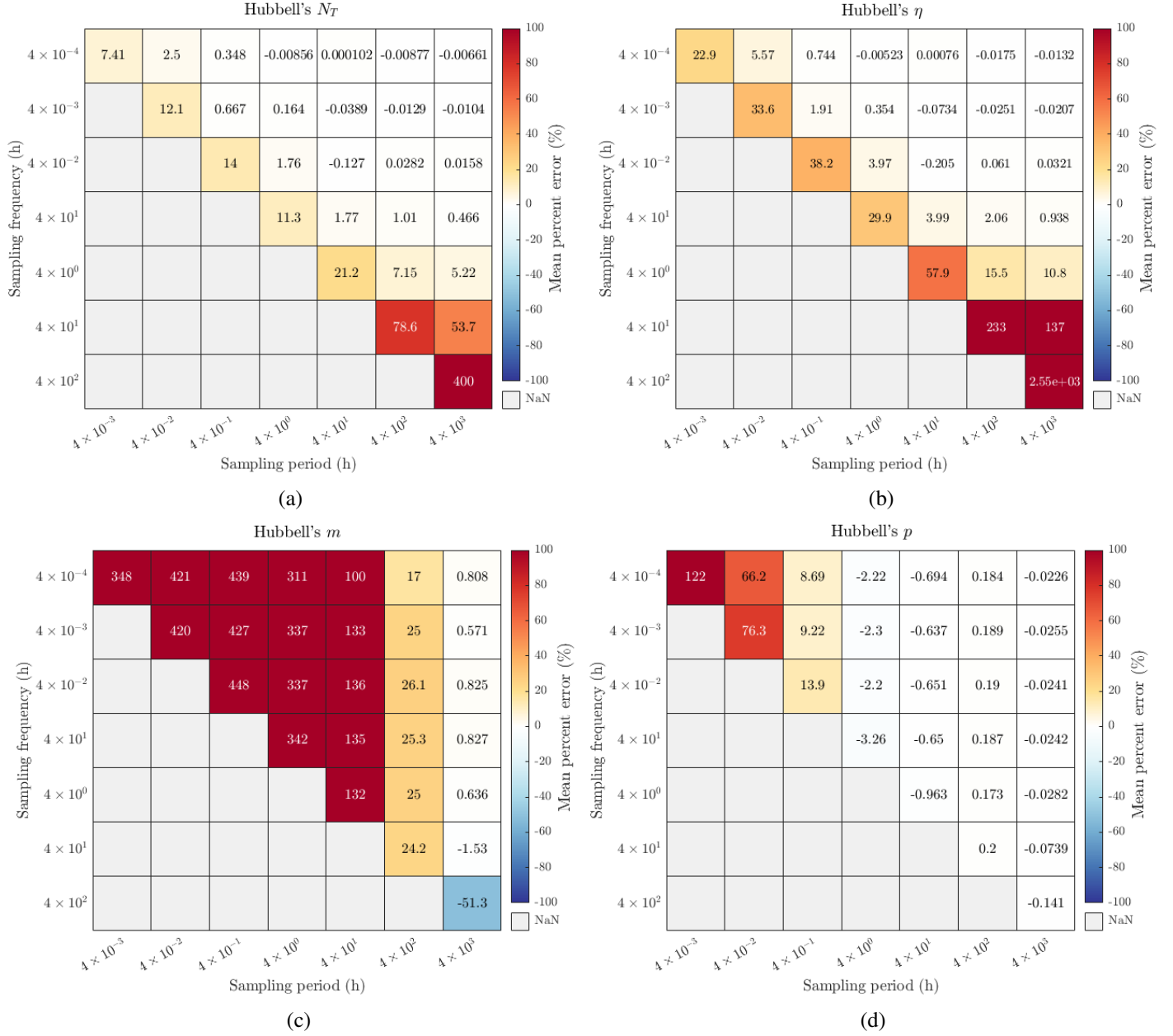


Figure 3.12: Heatmaps demonstrating the mean percent error (MPE) for each set of 100 estimations of Hubbell's purely neutral parameters from relative abundance time series sampled under different regimes. The sampling frequency used is represented along the the y-axis. The sampling period used is represented along the x-axis. a) N_T , b) η , c) m , d) p .

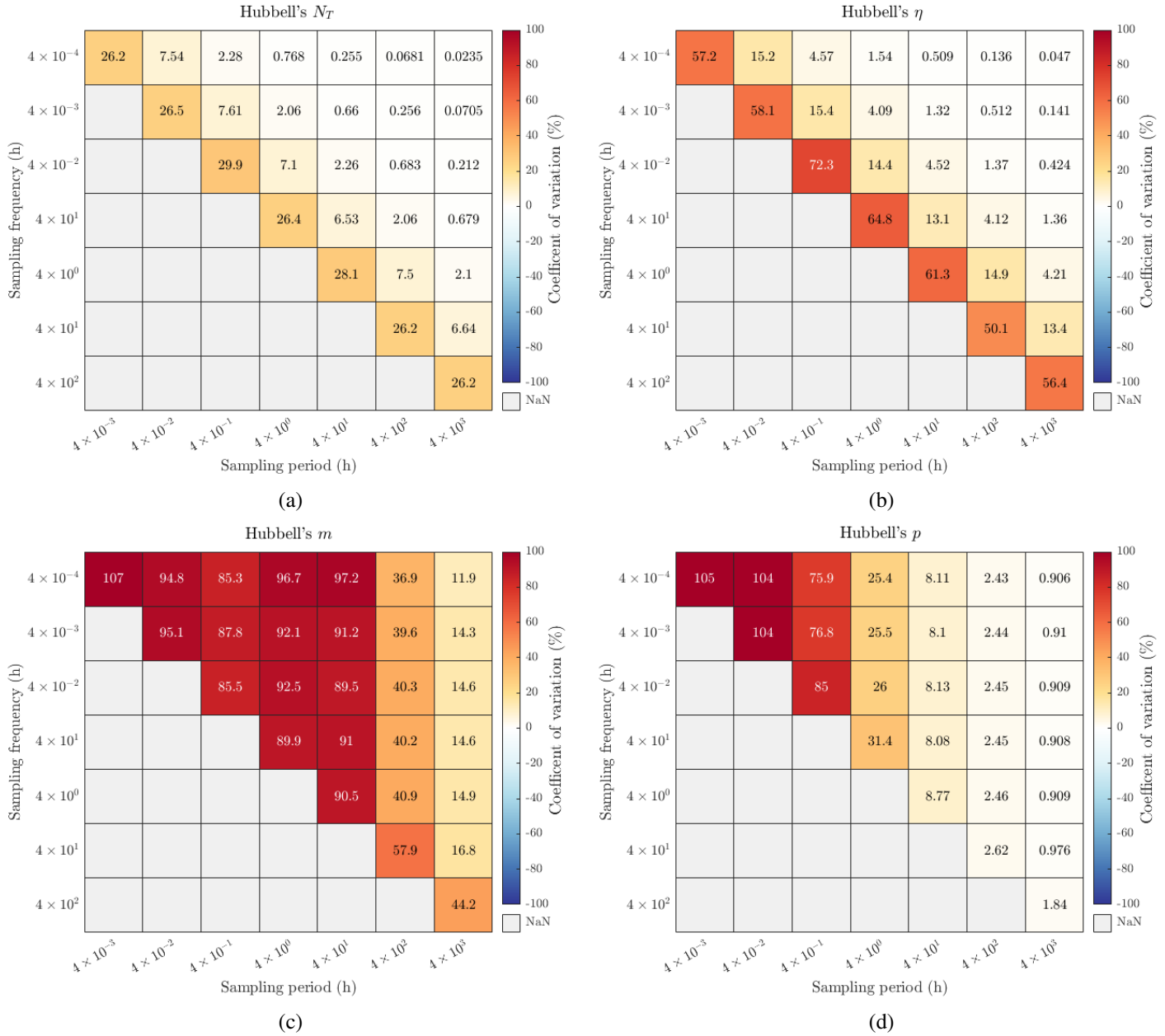


Figure 3.13: Heatmaps demonstrating the coefficient of variation (CV) for each set of 100 estimations of Hubbell's purely neutral parameters from relative abundance time series sampled under different regimes. The sampling frequency used is represented along the the y-axis. The sampling period used is represented along the x-axis. a) N_T , b) η , c) m , d) p .

For Hubbell's model, Figures 3.12c and 3.13c demonstrate that once again m is uniquely difficult to estimate and the most susceptible to changes in sampling period. The mean percent errors observed (Figure 3.12c) demonstrate significant overestimations of m as the sampling period is reduced; with the highest observed at 448% over the expected value. Based on previous explanations this would indicate under-representation of variance relative to the expected variance. Indeed, the onset of overestimation is inline with the onset of dampened variance portrayed in Table 3.4. The increase in the coefficient of variation displayed in Figure 3.13c is also in line with the increase in relative variance as the sampling period is decreased (Table 3.4). Therefore, to accurately calibrate m we need time series that best represents the expected variance. This is only achieved through long sampling periods. Our results do, however, highlight the lack of importance that sampling frequency has on the observed variance. Thus, if the sole aim of an experiment is to calibrate the immigration rate m , then sampling frequency can be reduced in order to capture the long time series needed.

The results for p display a similar trend to m . As previously shown, for time series captured at equilibrium the expected mean value is a much easier quantity to capture than the expected variance, this value will always be captured correctly given that the values of relative abundance are Beta distributed. This distribution is not represented by the small range of abundances sampled over a short duration, as suggested by Table 3.4. This leads to an increase in the variation between estimates of p as the sampling period is decreased, as shown in Figure 3.13d. Table 3.4 shows that this process happens less drastically than the damping of variance and, therefore, explains why the increase in errors is not as extreme as that observed for m . For short time series it is more likely that the sampled abundances exhibit a trend, rather than being stationary. Thus, the time series are just too short to capture the statistics of the process, and parameters, such as p , that reflect the long term mean abundance cannot be estimated.

Unlike m and p accurate estimates of N_T and η can be obtained from regimes with short sampling periods. Indeed, whilst the magnitude of the errors are not as large as for m and p , from both the MPE and CV heat maps it is apparent that the sampling frequency make the greater difference to the estimates of N_T and η . This is because the fitting algorithm is based on maximising the likelihood of dx ; successive difference in abundance. The continuous model is derived on an approximation based on small time increments. As the time between observations gets larger the poorer the theoretical expression for dx becomes.

Calibration of Sloan et al.'s parameters (Figures 3.14 and 3.15) yields similar conclusions. Like previously, errors in the estimates for N_T and η are driven by the total number of samples and sampling frequency, in particular, while errors in m , p and α are driven by the sampling period. The similarity of results between m , p and α further indicates the increased dependencies between these parameters as a result of including selection within the model processes. Figures 3.14c and 3.15c, do show that the inclusion of α within the model does stifle the increase in errors observed in m compared to what was observed for Hubbell's model when altering the sampling period (Figures 3.14c and 3.15c).

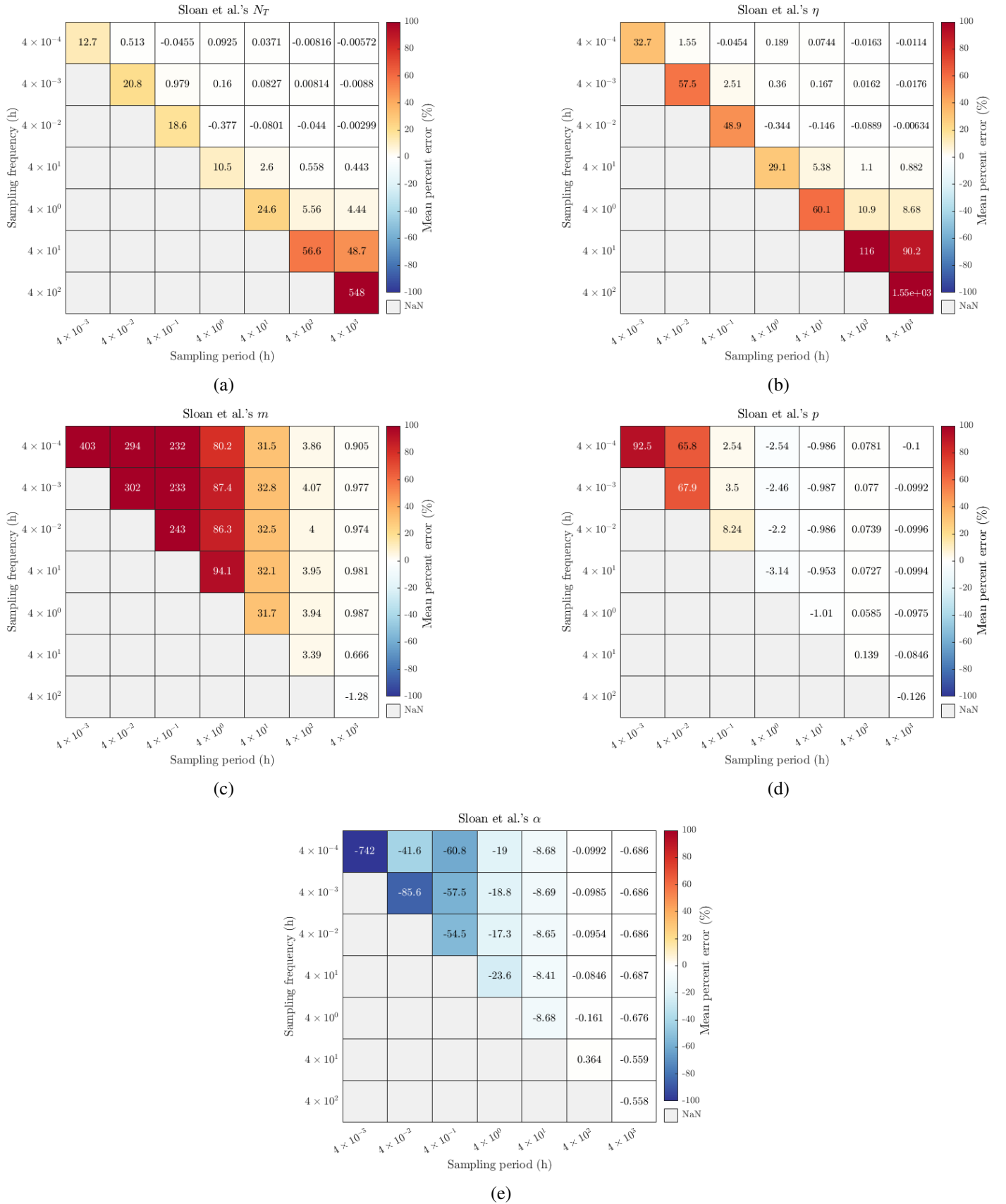


Figure 3.14: Heatmaps demonstrating the mean percent error (MPE) for each set of 100 estimations of Sloan et al.'s near-neutral parameters from relative abundance time series sampled under different regimes. The sampling frequency used is represented along the y-axis. The sampling period used is represented along the x-axis. a) N_T , b) η , c) m , d) p , e) α .

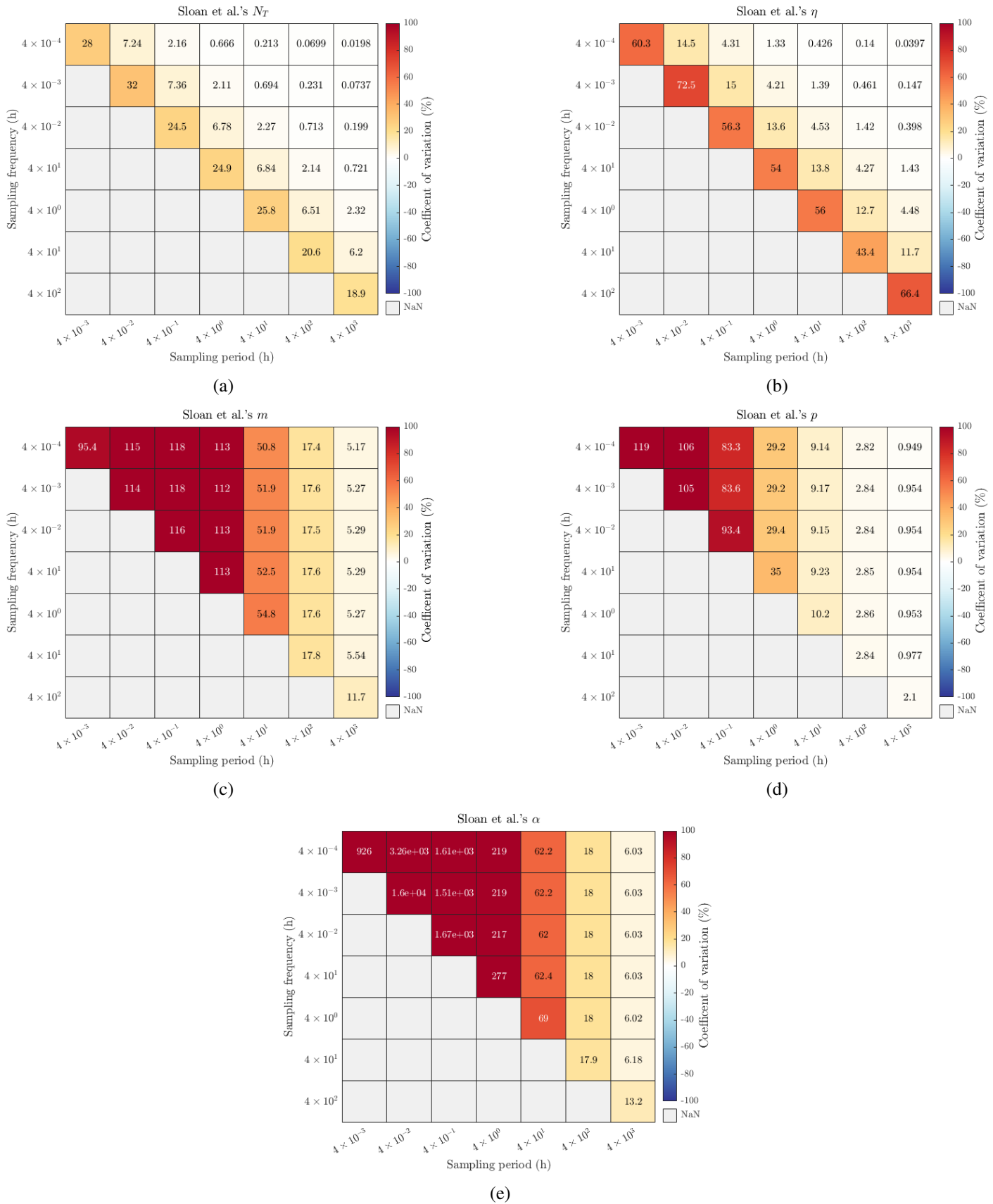


Figure 3.15: Heatmaps demonstrating the coefficient of variation (CV) for each set of 100 estimations of Sloan et al.'s near-neutral parameters from relative abundance time series sampled under different regimes. The sampling frequency used is represented along the y-axis. The sampling period used is represented along the x-axis. a) N_T , b) η , c) m , d) p , e) α .

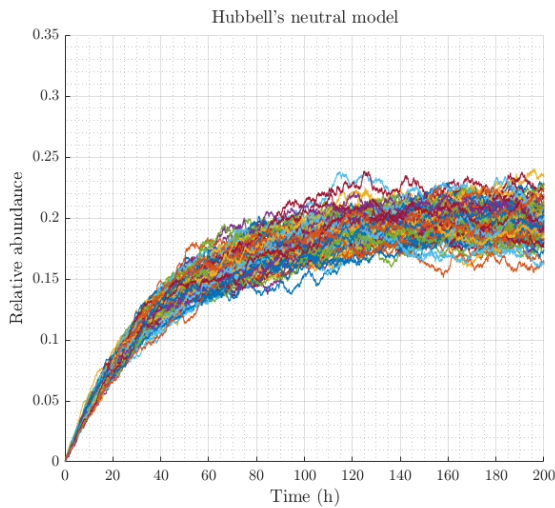
3.4.2.2 Time series captured from periods of non-equilibrium

In many natural and engineered systems the ecological processes that shape a microbial community will remain constant over time. Therefore, it is not unusual, indeed it might be the norm, that we rely on data from these communities in equilibrium to identify the processes. To achieve this in a quantitative fashion means encoding potential processes in a mathematical model and fitting the model to data. In this chapter, so far, we have explored the potential of retrieving the parameters of two very simple community assembly models, the neutral and near-neutral models, based on simulated time series data of the abundance of one species under equilibrium conditions. The results highlight some inherent difficulties in calibrating the neutral models generally and for equilibrium dynamics, in particular. The implications of trade-offs in for example N_T and η and the effects of sampling frequency and duration on model identification will be discussed later. First, however, it is worth considering whether the observations on model identification for systems in equilibrium also hold for systems that are not in equilibrium.

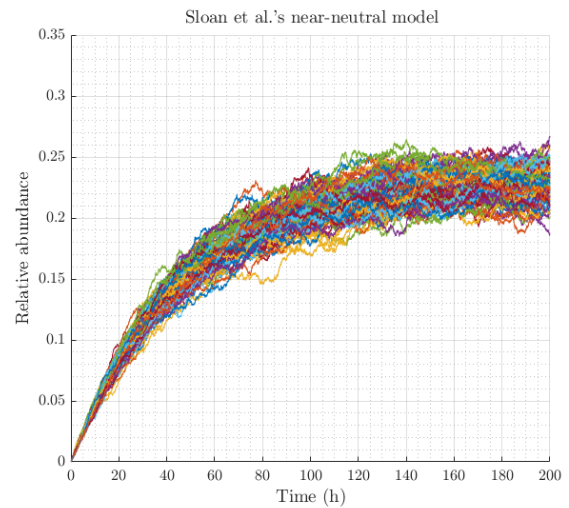
The assumption that the dynamics are in equilibrium may be compromised. For example, changes in physical conditions, such as weather or flow in fluid environments, can mean that immigration m changes over time. In engineered systems it may be desirable to perturb a system that is ostensibly in equilibrium in order to effect some positive outcome. For example, in biotechnologies, such as anaerobic digestion (AD) or probiotics in veterinary medicine, AD reactors or the guts of horses, the abundances of naturally occurring bacteria are boosted by augmentation to improve function or health. From the perspective of model identification, perhaps periods on non-equilibrium are beneficial, especially when time series can be collected. When calibrating a neutral model using stationary abundance distributions (SADs), observed at a single point in time, as is the norm, the assumption that a community is at equilibrium is necessary (Hubbell, 2001). For time series that are sufficiently long that a period of nonequilibrium can be captured in its entirety, the additional dynamics may help to separate out the parameters that trade-off when calibrated for equilibrium conditions.

To explore this, new sets of time series are simulated for both neutral systems investigated thus far. These time series are initially perturbed and then allowed to return to equilibrium dynamics. This is achieved by setting the starting relative abundance value within each simulation to zero. This is the antithesis of bioaugmentation. Rather than augmenting a species to an artificially high abundances in the local community, the species is initially removed, or is not present. Immigration then acts to grow the abundance of the species until it fluctuates around p , its relative abundance in the source community. The simulation parameters for the two models are kept consistent with what has been used thus far, such that, $N_T = 10,000$, $\eta = 0.0004\text{h}$, $m = 0.10$, $p = 0.20$ and $\alpha = 0.01$. Simulations are recorded after every replacement (0.0004 hours) for 200 hours as this allows for the period of non-equilibrium to be captured and the system to return to equilibrium dynamics. All newly simulated time series are depicted in Figures 3.16a and 3.16b.

Similarly to the first analysis undertaken both models are fitted under all scenarios and therefore all combinations of fixed and free parameters. Figures 3.17 and 3.19 display the values and accuracy in parameter estimations obtained when a period of non-equilibrium is observed. Figures 3.17 and 3.18 demonstrate that for Hubbell's model parameters N_T , η and m are estimated with better if not the same accuracy for almost all degrees of freedom when compared to their counterparts calibrated using high-resolution time series obtained under equilibrium over a period of 8,760 hours (Figure 3.7a and 3.7b). A better comparison is made against time series obtained under equilibrium over a more similar period, such as those sampled over a 400-hour period, as shown in Figures 3.12 and 3.13.



(a)



(b)

Figure 3.16: Relative abundance time series of a single species obtained through 100 realisations of two simulated communities. Time series represent a period of non-equilibrium. This was achieved by setting the starting relative abundance to 0 for each simulation. Sampling was conducted after every individual replacement event (0.0004h) for 200 hours. Parameter values were equal to $N_T = 10,000$, $\eta = 4.00 \times 10^{-4}$, $m = 0.10$ and $p = 0.20$. For realisations including selection $\alpha = 0.01$. a) Community dynamics were simulated from the discrete transitional probabilities described by Hubbell (2001) for a community governed by ecological drift alone (Equation 3.1), b) community dynamics were simulated from the discrete transitional probabilities described by Sloan et al. (2006) for a community governed by ecological drift and selection (Equation 3.2).

For m the improvement in accuracy is significant. For our time series under equilibrium sampled at a frequency of 0.0004 hours over a period of 400 hours the variation in m when η and p are fixed is 36.9%. This is reduced to 7.44% for a time series half the length that is obtained over a period of non-equilibrium. This improvement in m signifies that by monitoring periods of non-equilibrium the magnitude of tenancy towards the equilibrium value is better defined. This indicates that the drift during a period of equilibrium is too subtle to confidently identify m but that its effect on the dynamics is amplified by a perturbation to a species' relative abundance in a way that allows m to be identified. Given our previous description of the relationship between N_T , η and m this will likely also increase the ability to accurately calibrate N_T and η under calibrations where both are unknown. Even the perceived increases in variation associated with the estimation of p when comparing the results here to those from our idealised time series, do not look out of place when compared to equilibrium time series of a similar length (Figure 3.13d).

Figures 3.19 and 3.20 show that for Sloan et al.'s model monitoring a period of non-equilibrium also improves the ability to produce accurate parameter estimates. Although the variability in estimates of p and α appear higher than those for long high-resolution time series (Figures 3.10 and 3.11), these increases are in line with what is expected at equilibrium for sampling periods closer to 200 hours (Figures 3.14 and 3.14). Again the highest improvement is in estimates of m . A reduction in the CV of approximately 10% is observed when calibrating m to a time series captured during a period of non-equilibrium over a length of 200 hours (Figure 3.20b) compared to the calibration of m using time series at equilibrium sampled over a period of 400 hours (Figure 3.15c), where the same scenario is used.

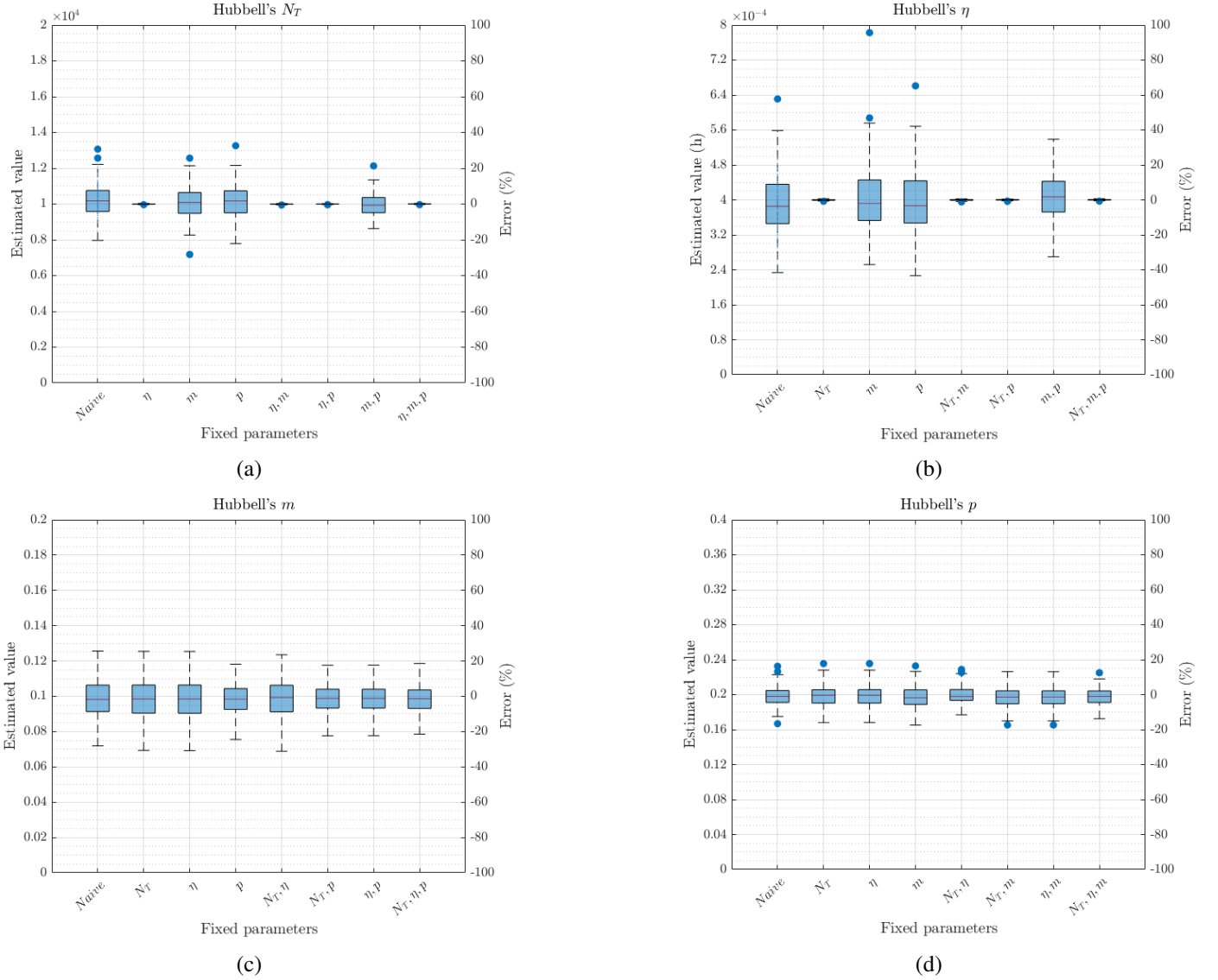
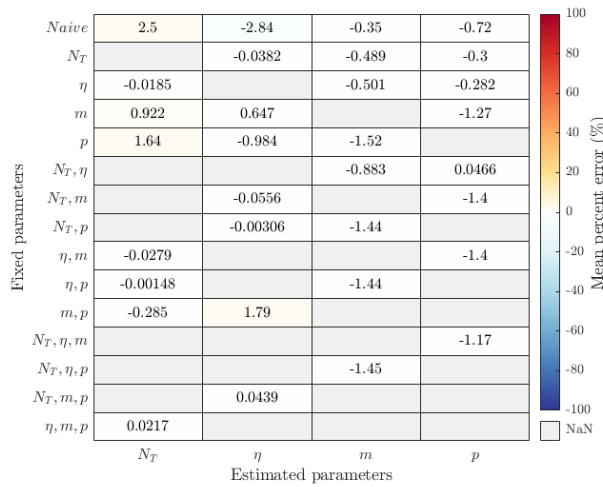
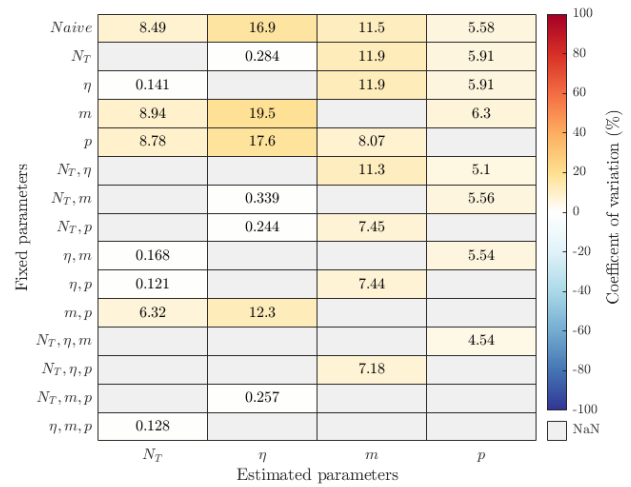


Figure 3.17: Box plots representing the estimation of distinct parameters from Hubbell's purely neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.16a. a) Estimation of N_T , b) estimation of η , c) estimation of m . d) estimation of p .



(a)



(b)

Figure 3.18: Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of Hubbell's purely neutral parameters shown in Figure 3.17. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs.

As a result the amount of prior knowledge needed to resolve the dependencies between m , p and α is reduced. Figure 3.20 shows that only small differences are observed within the estimation of m , p or α , between combinations of fixed parameters where one or two of the others are fixed. For example, there is almost no improvement in the estimation of α after p is fixed, this is contrary to what was previously observed in Figure 3.11. This is also observed across estimations for m and p . Thus, calibration of the model with higher degrees of freedom is successful.

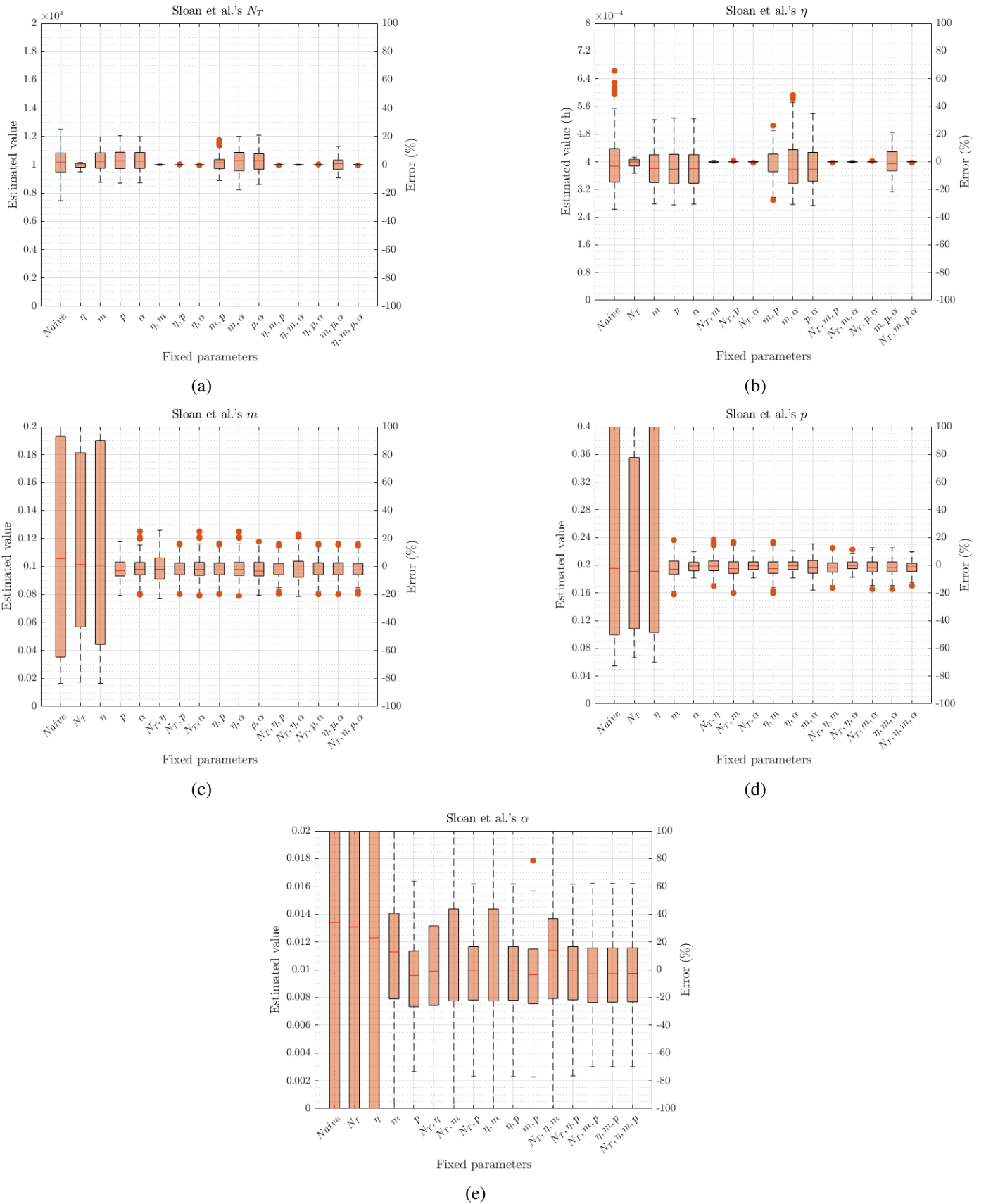
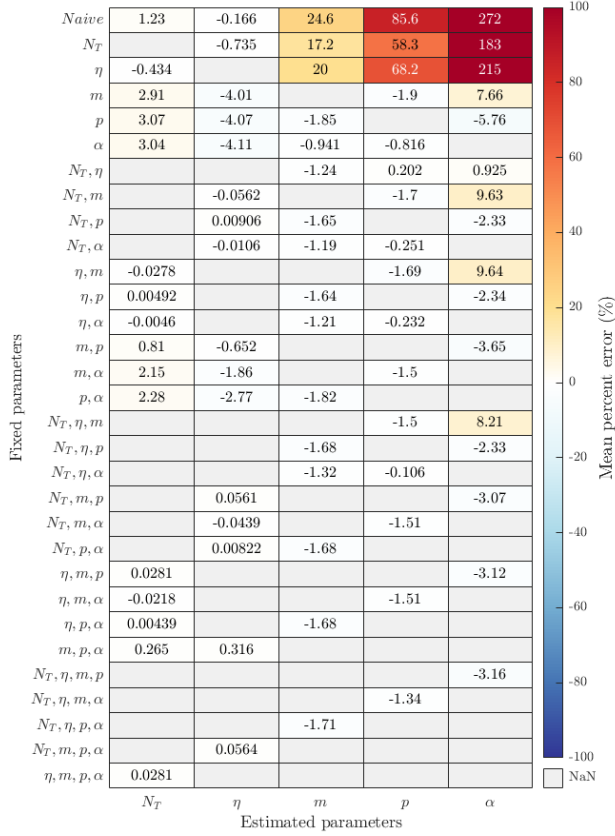
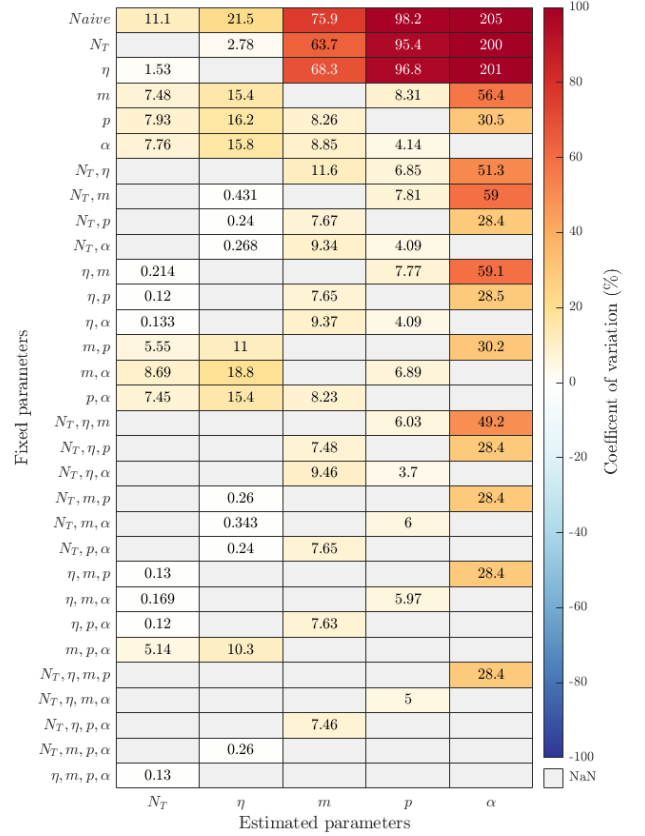


Figure 3.19: Box plots representing the estimation of distinct parameters from Sloan et al.'s near-neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.16b. a) Estimation of N_T , b) estimation of η , c) estimation of m , d) estimation of p , e) estimation of α



(a)



(b)

Figure 3.20: Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of Sloan's near-neutral parameters shown in Figure 3.19. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs.

3.4.2.3 The role of effective parameters

So far, there has been the luxury of generating synthetic time series data, so we know which model should, theoretically, fit which dataset and the challenge has been 'merely' to identify the correct parameters. In microbial ecology we do not have a fundamental, universally agreed set of laws governing community assembly. Whilst, most ecologists include drift, immigration and selection in ecological community assembly models, some, like Hubbell (2001), neglect selection and others advocate models with additional biological mechanisms, such as more explicit inter-species competition (Vellend, 2010). For a real system in nature, or in the laboratory, it is not possible to know, *a priori*, what complexity of model is warranted to capture the important ecological processes. Indeed, it is likely that when calibrating simple neutral models, as done here, additional processes will exist within the community. This begs the question whether the inclusion of such processes i.e those unaccounted for by the model calibrated, will influence the ability to accurately estimate model parameters and to what extent.

To demonstrate this in the most simple of cases, where only one additional unaccounted process exists, we repeat the initial analysis for Hubbell's model, however, we now calibrate it using the one hundred time series generated from Sloan et al.'s near-neutral model, depicted in Figure 3.9. Thus, selection is taking place within the simulated dynamics but not accounted for within the model calibrated. The results obtained from these calibrations are displayed in Figures 3.21 and 3.22.

The quantitative differences in expected mean and variance can be captured within Hubbell's parameters. Estimates of p , as shown in Figure 3.21d, are largely in line with the observed mean for the time series used (Table 3.3), and the under-estimation of m , when only N_T or/and η is fixed, represents the increase in observed variance. This however poses a problem when parameters are then independently measured, as shown when p is fixed at its true value of 0.2. The optimisation process will try to accommodate for the mismatch in the fixed and observed quantity through altering the value of other parameters. This leads to large errors in the parameter estimates.

Given that for mixed microbial communities in most natural, and even carefully engineered, systems a myriad of interactions between species and the environment could be impacting community structure. Both Hubbell's and Sloan et al.'s models will, inevitably be simplifications, and therefore parameters estimated by calibration may be artifacts of more complex interactions. The results obtained demonstrate that although these models may provide a good fit, and that the erroneous estimations of m and p are in fact in line with what is expected given the true community dynamics, this presents a problem when independently determining the value of a parameter based on its ecological meaning true to the calibrated model. This is especially worrying as we have previously demonstrated that at least some degree of *a priori* knowledge of the system is required to resolve correlated parameters. Although no differences are observed within the estimation of N_T and η , it can be assumed that the same logic applies. Indeed, parameters now describe an effective quantity and thus it may be more beneficial to name them as such.

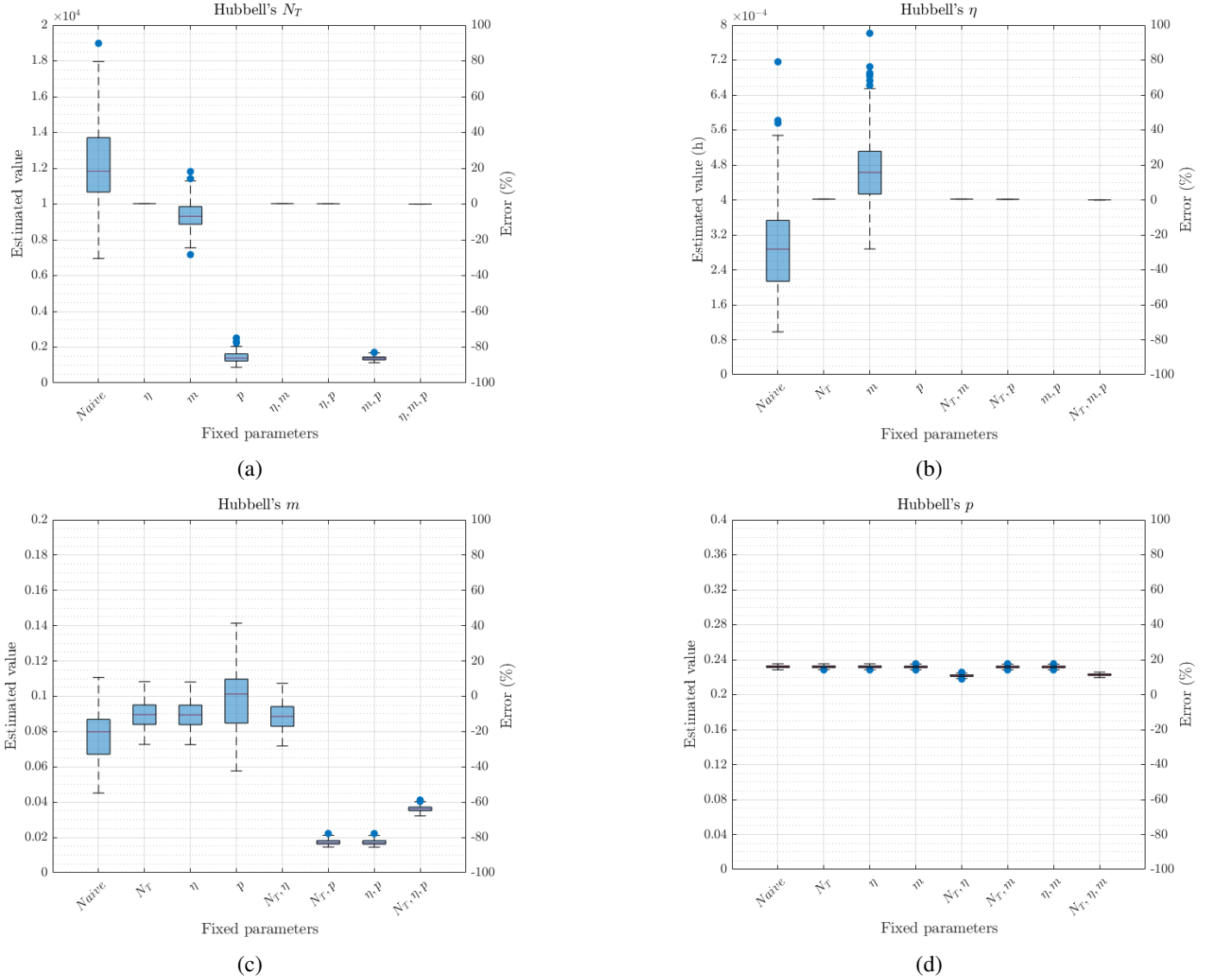


Figure 3.21: Box plots representing the estimation of distinct parameters from Hubbell's purely neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.9. a) Estimation of N_T , b) estimation of η , c) estimation of m . d) estimation of p .

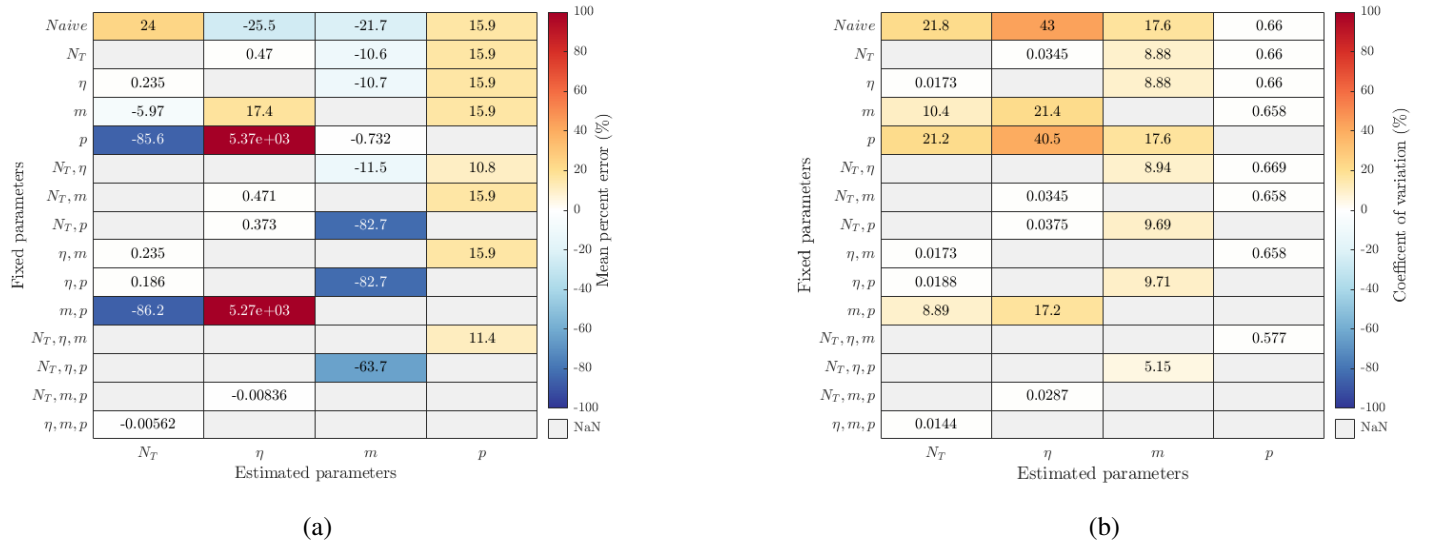


Figure 3.22: Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of Hubbell's purely neutral parameters shown in Figure 3.21. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs.

To implement this, the recommendation provided by Sloan et al. (2021) of utilising an "effective community size" is extended across all unique parameters, such that each parameter can be represented by an effective counterpart. Here, effective parameters are denoted with a small e , such that $N_T = N_{Te}$, $\eta = \eta_e$, $m = m_e$, and $p = p_e$. In the extreme case where all parameters are considered effective, the continuous form of Hubbell's purely neutral model is represented by:

$$dx_{\text{hubbell}} = \frac{m_e(p_e - x)}{N_{Te}} d\tau + \sqrt{\frac{2x(1-x) + m_e(p_e - x)(1-2x)}{N_{Te}^2}} dW_\tau, \quad (3.13)$$

where non-effective parameters hold their previous meanings but $\tau = t/\eta_e$.

To facilitate calibration of this effective model, under all combinations of fixed parameters, we must come up with expected values for the effective parameters. Here we assume $N_{Te} = N_T = 10,000$ and $\eta_e = \eta = 0.0004h$. We can calculate p_e based on the expected mean, which makes $p_e = 0.232$ (Appendix B). m_e is harder to determine mathematically and thus we simply use the mean percent error highlighted in Figure 3.22a obtained when either N_T or η is fixed. Therefore, we arrive at an approximation of m_e , where $m_e = 0.894m = 0.0864$.

Figures 3.23a and 3.23b demonstrate the errors associated with calibrating the effective version of Hubbell's model to the time series generated from Sloan et al.'s near-neutral model. While the variance in estimates is almost identical to those measured previously (Figure 3.22b), the results demonstrate that by recasting Hubbell's model in terms of effective parameters the MPEs within the calibrated parameters are significantly improved from those observed previously (Figure 3.22a). MPE values are still however not at the levels observed under our idealised conditions (Figure 3.7a). This is likely a consequence of the approximations of the effective parameters used, rather than issues with the model fitting. Defining robust methodologies for determining effective parameters independently is therefore required. For real communities in which the exact structuring process are truly unknown, this will prove challenging; as demonstrated by the difficulty in

predicting and estimating the "effective population size" in population genetics (J. Wang et al., 2016).

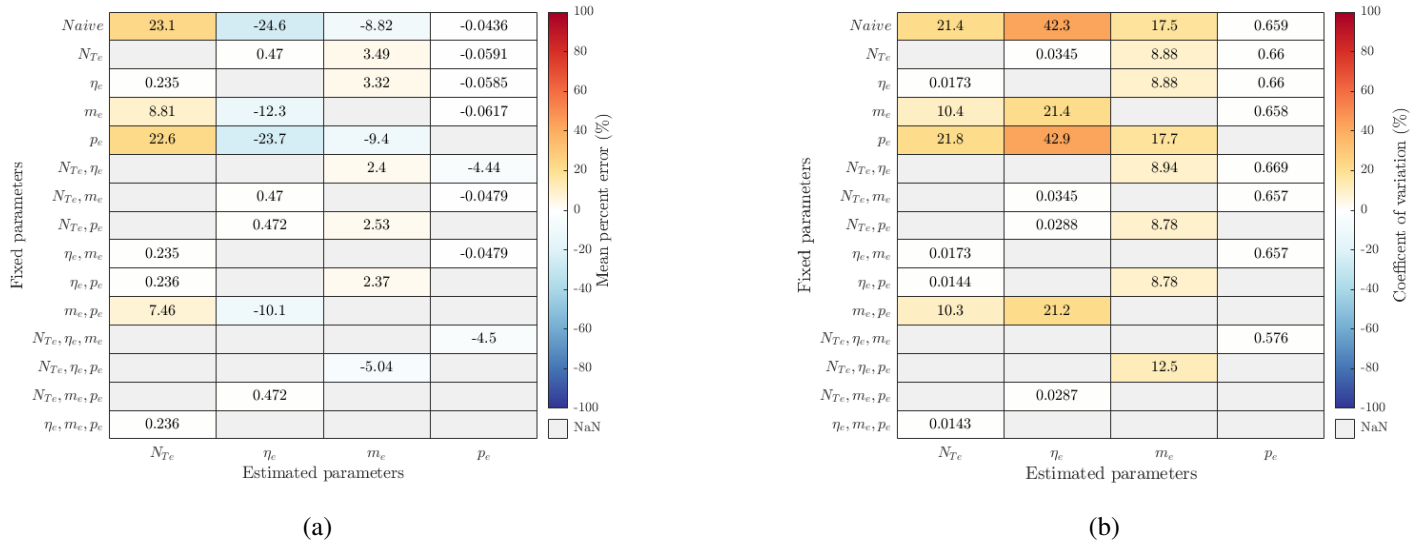


Figure 3.23: Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of Hubbell's effective parameters to all 100 distinct time series depicted in Figure 3.9. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs.

3.5 Conclusion and future perspectives

To ecologists and biologists who observe the vast range of interactions between organisms and the environment, NCMs can appear too simple to represent real ecological communities. Many critiques of neutral theory suggest its lack of complexity renders it ineffective (Clark, 2009, 2012; Purves & Turnbull, 2010). Despite this, it does encapsulate basic processes that must occur in every open microbial community (Vellend, 2010). The fact that such simplistic models fit commonly observed species abundance patterns makes the parameters within them inherently important. Understanding under what conditions estimates of these parameters is achievable is therefore important (Beeravolu et al., 2009; Etienne et al., 2006). Interpreting the calibrated parameters, and what they encapsulate, is a separate but equally interesting task.

Accurately identifying distinct parameters, even in highly simplified neutral and near-neutral models, is fundamentally difficult to achieve when using relative abundance time series data obtained from a single detectable species under the scenarios imposed. The lack of information within these time series, especially when the system is in steady-state, means that parameters remain correlated making them structurally non-identifiable. For example, when calibrating Hubbell's purely neutral model or Sloan et al.'s near-neutral model we observed a prominent correlation between N_T and η . For both these models, even though they comprise only 3 and 4 parameters respectively, the data from monitoring a single species time series of abundance for a community in equilibrium is insufficient to identify each model accurately. These simple models are, in effect, over-parametrised for this type of data, even at the very high resolutions simulated here. Thus, to identify the models based on data, there is a need to constrain or independently measure some of the parameters or, alternatively, increase the richness of information used to calibrate the model. This could involve using data from multiple species, which would potentially add to the labour of experiments, or, induce perturbations in the experiment,

such that the dynamics are not in equilibrium. The results presented also demonstrate that mismatches between expected and observed phenomena drive errors in parameter estimates and that these errors can propagate through the previously mentioned correlated parameters.

Additionally, when calibrating these models using abundance data alone and under a scenario where structuring processes are unknown, it is impossible to know if the model represents all the processes shaping the community. Where possible, in highly constrained experiments, one can eliminate some of the external forces that might add complexity, but even then internal factors, like spatial organisation may be neglected in the model. It has been shown here that building in complexity only helps if the dynamics in abundance data are sufficiently rich to identify parameters. For systems in equilibrium this is not the case. Thus neutral models, despite their simplicity, are often fitted in recognition that they are the most complex warranted given the data. The parameter values will therefore likely encapsulate the effects of processes over-and-above those that they were conceived for.

This approach is not unique to ecology. In population genetics models for the distribution of alleles in a population, when calibrated, yield values for population size that are far smaller than the census population. Population geneticists call this calibrated value the "effective population size". It is a fundamentally important number that drives the evolutions of populations and indicates, for example, the number of breeding individuals population; which was not explicitly represented in the original model (Charlesworth, 2009). A recent study from Sloan et al. (2021) pushes for the use of a similar term, coined the "effective community size", within community ecology. Utilising effective parameters thus allows for the relaxation of strict ecological definitions tied to each parameter, removing the rigidity of having to assume that the model calibrated offers an exact explanation of all processes influencing the community structure. These effective parameters thus becomes fundamental descriptors of the community that are critical to describing the dynamics. The challenge then is to identify these effective parameters.

Within this study we have provided an example of utilising such a notation and demonstrated that while this notation is required to reduce errors when fitting these models to systems more complex than the model describes, accurate independent determination of effective parameters is more nuanced. As a result, fixing these parameters within models may become difficult. Indeed, within population genetics multiple methods have been developed in order to predict effective parameters (J. Wang et al., 2016). This presents an issue, as previously mentioned producing accurate parameter estimates from single-species relative abundance time series require at least some prior knowledge of these parameters. Future research must therefore look to develop robust methodologies for independently determining these parameters.

Another approach is to only define effective parameters for those independent of strong correlations, or only in scenarios where these correlations are resolved, avoiding the need to determine them independently when estimating them through model calibration. For example, given that fixing η within Hubbell's and Sloan et al.'s models acts to facilitate accurate estimation of N_T , then under such a scenario, defining N_T as its effective counterpart (N_{Te}) and identifying it through model calibration will allow the model to capture the quantitative information provided by the time series while also solving the required trade-off between η and N_T .

Conceding the use of effective parameters such that each parameter now reflects the quantitative effect of many ecological processes, the distinction between the models developed by Hubbell and Sloan et al. is weakened. As demonstrated Hubbell's model with the inclusion of effective parameters can be calibrated to a community in which selection is active. This is arguably achieved more easily than for Sloan et al.'s model, as the lack of a selection parameter (α) untangles dependencies between m , p and α . However, the previous definitions of each parameters still provide ecological and theoretical insight. Indeed, an effective community size (N_{Te}) is a more natural concept to understand than an effective average time between replacement events (η_e), although both can be used to capture the relative step change in abundance per unit of time. It can be argued under the same logic that calibrating a single parameter that captures all processes driving the behaviour of local selection (α_e) is more theoretically compelling than capturing an effective rate of immigration (m_e) and an effective relative abundance within the source community (p_e).

Irrespective of whether model parameters hold distinct ecological meaning or are considered effective, identification of these parameters is highly dependent on the sampling protocol used. The results within this study demonstrate that sampling period and frequency effect the information portrayed within in each relative abundance time series. Reductions in sampling period are shown to lead to the under-representation of the expected variance and divergence from observed Beta distributions, while reductions in sampling period degrade the theoretical expression for dx . These trends drive errors within respective parameters, and therefore, if calibration of neutral models is undertaken to produce parameter estimates, then at least some rationale behind the sampling regime used during the acquisition of data must be provided. Processes that maintain the integrity of the time series but allow for reduced sampling should therefore be explored. Here we have demonstrated that one process which fulfils this is to induce a perturbation on a species and measure its recovery to its equilibrium. Here it has been shown that if this perturbation, or period of non-equilibrium, is captured in its entirety, errors associated with sampling period, especially in the under-representation of time series variance are largely negated. For such a system, the sampling period must therefore only cover the period of recovery. In the example provided, this allows the required length of time series to be reduced significantly. Another option, not explicitly demonstrated here, is to provide high resolution snapshots at infrequent intervals. The ideal would be to produce relative abundance measurements (x) in quick succession to get a good estimates of dx , but then spread the estimate of dx sparsely, through the larger duration need for capturing the long term dynamics.

It is important to note that the analysis conducted here is not exhaustive, the time series produced are simulated based on only a single set of parameter magnitudes: $N_T = 10,000$, $\eta = 0.0004h$, $m = 0.1$, $p = 0.2$ and $\alpha = 0.01$. Indeed, it is foreseeable that if the magnitudes of these structuring processes are altered then the exact degree of parameter estimation error will vary. Thus, the exact quantities presented within this study are unlikely to be universal. An expansion of this study, in which the magnitudes and therefore the relative importance of these processes is altered, would be beneficial for determining whether universal quantities or trends in calibration errors exists. Within this study it has been alluded to that optimal sample periods and frequencies exist for ensuring accurate calibration of specific parameters, if possible, providing a universal mathematical relationship for calculating this point would be extremely useful for informing experimental design. Investigation of other sampling processes not accounted for here, such as, sample size or the inclusion of multi-species relative abundance time series, would also be useful for experimentalists.

Future work should also assess the practical applications of the models analysed here. In many engineered systems, where communities are at the heart of the processes undertaken, it can be beneficial to maintain a certain species, genus or functional group within the community. Purposely calibrating neutral model parameters from these systems, in an attempt to monitor trends within a species' relative abundance would provide a valuable case study. In a more theoretical context, using closely controlled experiments or computational simulation, parameters produced through the calibration of neutral models could be utilised for quantifying the impact or relative importance of certain phenomena, as eluded to by Sloan et al. (2021).

The overall view taken here is that through more nuanced and thoughtful calibration of neutral models, accurate estimates of the effective parameters governing these models can be achieved. These parameters can then be used, not only to categorise the information provided by species relative-abundance time series, but also to provide an insight into how the abundance of a species may vary as time progresses. Additionally, by investigating how certain ecological, chemical or biological processes adapted the values estimated, it may be possible to develop methods for influencing the structuring of communities. The work here therefore aims to pivot the use of neutral models towards a more pragmatic and engineering related use. Indeed, much of the criticisms of neutral models is often in relation to the ecological contents of the models. If the ecological definitions are therefore relaxed, through the use of effective parameters, then the mathematical simplicity and ability of these models to fit many observed species abundance patterns can be utilised in full.

Chapter 4

Quantifying the heating potential of methanotrophs for application within decentralised wastewater treatment technologies

4.1 Abstract

In many decentralised wastewater technologies, like septic tanks or small treatment works, wastewater sludge is stored and undergoes anaerobic digestion, producing methane with significant global warming potential ($\text{GWP-100} \approx 30$). Methanotrophic bacteria oxidise methane to gain energy, even at low concentrations, and thus could potentially be deployed in mitigating the methane emissions from decentralised wastewater treatment. There is evidence that methanotrophs dissipate a significant proportion of the energy released in oxidising methane as heat. This opens up the possibility of technologies that simultaneously mitigate methane emissions and make use of the low-grade heat dissipated using, for example, heat pumps. To rationally design such technologies requires a much deeper understanding of the heat dissipated by methanotrophs. Here we develop a method of characterising and apply it to two different species of methanotrophs: *M. methanica* S1 (type-I methanotroph) and *M. trichosporium* OB3b (type-II methanotroph), known to occupy different ecological niches.

Batch cultures were grown in isothermal calorimeter chambers at 30°C with headspace concentrations of 8%, 4% and 2% v/v CH_4 in air. Across all cultures, 56 – 64% of the energy in oxidising all of the methane was dissipated as heat. Batch experiments, that identically replicated those in the calorimeter, were conducted outside the calorimeter, which allowed optical density (OD_{600}), headspace composition and pressure, dissolved oxygen, pH and non-purgeable organic carbon (NPOC) to be measured. Analysing these indicates that for *M. methanica* S1 and *M. trichosporium* OB3b approximately 43 – 59% and 49 – 66% of the carbon stored in the supplied methane is fully oxidised to carbon dioxide respectively. For *M. methanica* S1, 27 – 35% of this carbon is used to synthesis biomass, while for *M. trichosporium* OB3b, this is reduced to between 14 – 22%. We demonstrate that for species utilising the same substrate, and where this substrate is the sole source of energy and carbon, the partitioning of available carbon between energy-yielding and energy consuming reactions is linearly related to the quantity of heat dissipated per biomass. Discussion is given to the role heat dissipation has on the thermodynamics at a cellular level and how this impacts the overall growth rate.

4.2 Introduction

Waste management plays a significant role in the anthropogenic sources of methane (Saunois et al., 2020). As such, reducing methane emissions from waste and wastewater technologies has been identified as an opportunity for scaling up climate action (Intergovernmental Panel on Climate Change (IPCC), 2023c). In large-scale systems, mitigation of methane emissions is undertaken using various technologies (Pratt & Tate, 2018), the most common of which involves burning the methane in combined heat and power plants. For smaller decentralised systems, such as septic tanks or small wastewater treatment works, where sewage sludge is stored, little development since their original design has been made (P. J. A. Withers et al., 2011). The effect of emissions on the climate are often neglected and no mitigation strategies are used. Until recently, little was known about CH_4 emissions production from decentralised wastewater systems, however, recent studies are demonstrating that CH_4 is a leading constituent within greenhouse gas (GHG) emissions from these systems (Diaz-Valbuena et al., 2011; Manga & Muoghalu, 2024; Somlai et al., 2019). Not only this, but the per capita carbon footprint of rural wastewater technologies in Scotland have been shown to be 7 times that of centralised urban technologies and most of this can be attributed to methane emissions from sewage sludge (Gupta et al., 2024).

Biological solutions have shown increasing promise and cost-effectiveness for combatting GHG emissions (López et al., 2013). In this context it is often prudent to look to nature to find the most suitable solutions. In the environment, methanotrophic bacteria represent the sole biological methane sink (Guerrero-Cruz et al., 2021). Aerobic methanotrophs represent a diverse functional group of bacteria that generate energy and biomass through the oxidation of methane. Their diverse taxonomy reflects variations in preferred growth conditions, physiology, and anabolic pathways. However, while different enzymes can catalyse the catabolic stages, the overall pathway of methane oxidation to carbon dioxide via methanol, formaldehyde, and formate remains consistent across methanotrophs (Hanson & Hanson, 1996; Khmelenina et al., 2018). It is this metabolism that makes them useful in engineering applications to mitigate methane emissions.

The use of methanotrophs for GHG mitigation is an active area of research. Strategies such as co-culturing alkaliphilic methanotrophs and microalgae have shown the ability to sequester both CH_4 and CO_2 , enabling direct reduction in GHG emissions (Ruiz-Ruiz et al., 2020). Anaerobic methanotrophs have been incorporated within WWTPs to link nitrogen and carbon cycles (Guerrero-Cruz et al., 2021; Y. Wang et al., 2017). Interest extends beyond moping up methane emissions because some of the intermediates and products of methanotrophic metabolism are valuable resources. A review by Strong et al. (2015) highlights that methanotrophs can be used to generate a wide range of useful products. Perhaps the most well-studied of these is the excretion of methanol due the oxidation of methane by methane monooxygenase (MMO) in aerobic methanotrophs (D. Park & Lee, 2013; Sahoo et al., 2023; Strong et al., 2015; Xin et al., 2004). This ability is leveraged to convert biogas into liquid fuels at ambient conditions (Ge et al., 2014; Sheets et al., 2016), representing a more sustainable strategy to current energy intensive thermo-chemical processes (Riaz et al., 2013).

Our interest is in exploiting the thermodynamic properties of these bacteria. To our knowledge, literature utilising heat dissipation from methanotrophic bacteria for wastewater treatment GHG emission mitigation strategies is rare. The most likely manifestation of this would be through the use of biofilters in series with anaerobic digesters (Brandt et al., 2016; Gómez-Borraz et al., 2017; La et al., 2018). Heat produced within these filters could then be transferred back to the anaerobic reactors through use of a heat pump or utilised for other means. We hypothesise that such a system is feasible as the oxidation of methane is highly exothermic ($\Delta_c H^\circ = -890.71 \text{ kJ/mol}$) (Pittam & Pilcher, 1972), and in decentralised

anaerobic treatment technologies, such as septic tanks, raising the system temperature has been shown to enhance treatment efficiency (Connelly et al., 2019; Pussayanavin et al., 2015). Indeed, in systems where biological methane oxidation is assumed to be taking place, large temperature increases have been recorded (Fjelsted et al., 2020; Gómez-Borraz et al., 2022), and increases in subsurface temperatures have also been linked to the activity of methanotrophic bacteria (Warren & Bekins, 2015).

There is also interest in demonstrating the link between heat dissipation and kinetic parameters from a theoretical viewpoint. Indeed, perhaps the most fundamental view of bacterial growth is that offered by the second law of thermodynamics, which when applied to biomass production, governs that the biosynthesis of complex molecules from simpler substrates, requires a thermodynamic driving force (von Stockar & Liu, 1999). This is achieved, at least theoretically, by coupling the energy-yielding catabolic reaction to the energy-consuming anabolic reaction (Heijnen & Van Dijken, 1992; Stockar & Wielen, 2013b; von Stockar & Liu, 1999). It is also theorised that the magnitude of driving force dictates the rate of reactions (Stockar & Wielen, 2013b; von Stockar & Liu, 1999). Methanotrophs offer an interesting case study to experimentally test this theory. Given that methane is the most reduced form of carbon, methanotrophs have the potential to dissipate greater amounts of heat relative to other microorganisms (Birou et al., 1987). This makes the quantity of heat dissipated easier to capture. They also utilise the same substrate for anabolic and catabolic reactions (Hanson & Hanson, 1996), therefore, monitoring how methanotrophic species partition carbon should provide insight into how differences in heat dissipation between species are generated.

In lieu of the direct measurement of heat dissipation from the biological methane oxidation, theoretical approaches are used (Gómez-Borraz et al., 2022; Seto et al., 2019; Warren & Bekins, 2015). These approaches have only been tenuously linked to experimental evidence. Thus, if we are to rationally design biotechnologies to generate heat from biological methane oxidation then it is imperative that we empirically quantify the heat dissipated by methanotrophs and validate, or update, theoretical models for their energetics. Here, we have developed a robust approach to characterising the heat dissipated by different species of methanotrophs. Pure cultures of *Methylobacterium methanica* S1 and *Methylobacterium trichosporium* OB3b were cultivated separately within a Calimetrix I-CAL HPC 4000 Isothermal Calorimeter, allowing us to measure the heat dissipation *in situ*. This was done for three distinct initial concentration of CH_4 (8%, 4% and 2% v/v in air). For cultures within the calorimeter measurements of environmental and growth parameters, such as optical density (OD_{600}), headspace composition and pressure, dissolved oxygen, pH and non-purgeable organic carbon (NPOC) were taken at the start and end of each growth period. To reproduce how these values change over time, external cultures were grown in parallel to those within the calorimeter and sampled routinely over the entire growth period.

4.3 Methods

4.3.1 Materials

4.3.1.1 Methanotrophs

Two species of methanotrophs were used throughout the experiments carried out, these were:

1. *Methylomonas methanica* S1,
2. *Methylosinus trichosporium* OB3b,

where, S1 and OB3b are strain identifiers. The selection of species was made on multiple criteria. These species were selected because, whilst they grow on the same culture media and at similar temperatures, they have distinct metabolisms classified as type-I and type-II (Bowman, 2006; Hanson & Hanson, 1996; Soni et al., 1998). These types have been shown to have different growth kinetics and thus occupy different ecological niches (J. Amaral et al., 1995; J. A. Amaral & Knowles, 1995; D. W. Graham et al., 1993; Hanson & Hanson, 1996; Henckel et al., 2000). Both strains were obtained from internal freezer stocks stored at the Advanced Research Center at the University of Glasgow and were stored in 2ml cryovials at -80°C . Each vial contained 0.9ml of bacteria (OD_{600} of 0.5-1.0) with 100 μl of Dimethyl sulfoxide (DMSO). Table 4.1 provides a summary of characteristics of each strain.

Table 4.1: Comparison of characteristics between *Methylomonas methanica* S1 and *Methylosinus trichosporium* OB3b (Bowman, 2006; Hanson & Hanson, 1996; Koh et al., 1993; Murrell & Dalton, 1983; Whittenbury et al., 1970).

	Methylomonas methanica	Methylosinus trichosporium
Depositor designation	S1	OB3b
Type	I	II
Morphology	Short rods	Rod or pear shaped cells
Colour	Pink	White
Carbon assimilation	RuMP pathway	Serine pathway
pMMO	+	+
sMMO	+	+
Can fix nitrogen	-	+
Optimal temperature	25°C - 30°C	25°C - 30°C
Media	Nitrate mineral salts	Nitrate mineral salts

4.3.1.2 Media

Both *Methylomonas methanica* S1 and *Methylosinus trichosporium* OB3b were grown in nitrate mineral salts (NMS) media, as initially described by Whittenbury et al. (1970). This media contains: KNO_3 , MgSO_4 , CaCl_2 (anhydrous), FeEDTA , Na_2MoO_4 , trace elements (CuSO_4 , FeSO_4 , ZnSO_4 , H_3BO_3 , CoCl_2 , EDTA di sodium salt, MnCl_2 , NiCl_2) and phosphate buffer (Na_2HPO_4 , KH_2PO_4). Exact amounts and specific hydrates used, can be found in the recipe supplied at: https://store.ncimb.com/page/media_table.

4.3.1.3 Methane

CH_4 used within the experiments was obtained from a compressed gas cylinder stored at the Advanced Research Center at the University of Glasgow. This was decanted through a $0.2\mu m$ filter and into a 1L gasbag. All injection of CH_4 to cultures were made directly from the gasbag. The CH_4 used was of grade N2.5 (99.5%) and purchased from BOC (Part number: 270927-F).

4.3.1.4 Culture bottles

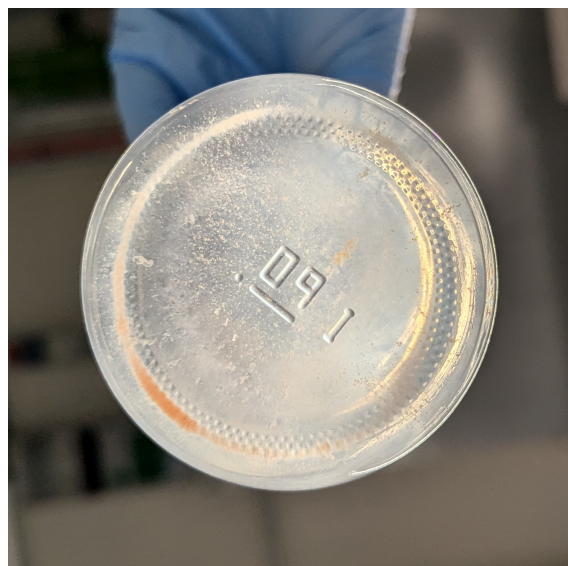
The sample containers supplied by Calmetrix for use within the I-CAL 4000 HPC Isothermal Calorimeter were unsuitable for culturing methanotrophs. As a result custom bottles were designed and manufactured. The design of these bottles was based on the dimensions of the test chambers, thus, the height and width of these bottles were designed not to exceed 60.0mm and 53.5mm respectively. Accounting for a size tolerance of approximately $\pm 2mm$, bottle height and width dimensions were set at 58mm and 52mm respectively. To ensure airtight sealing, the neck and opening were designed to accommodate 20mm crimp seals. The bottles were manufactured from 52mm (outer diameter) Simax borosilicate medium-walled tubing (2.5mm thickness) and produced with a smooth finish to ensure a uniform contact area. The manufacture of all newly designed bottles was undertaken by the glass-blowing facility at the University of Glasgow. In total 32 bottles were produced.

To ensure the experiments conducted outwith the the calorimeter replicated those within, the same bespoke culturing bottles were used for both. Between experiments bottles were left in virkon for at least 24 hours and autoclaved before use, safeguarding against the risk of contamination. Exact volumes of each bottle were determined by comparing the empty weight of each bottle against its filled weight using water. These measurements, alongside engineering drawings of the bottles, can be found in Appendix F.

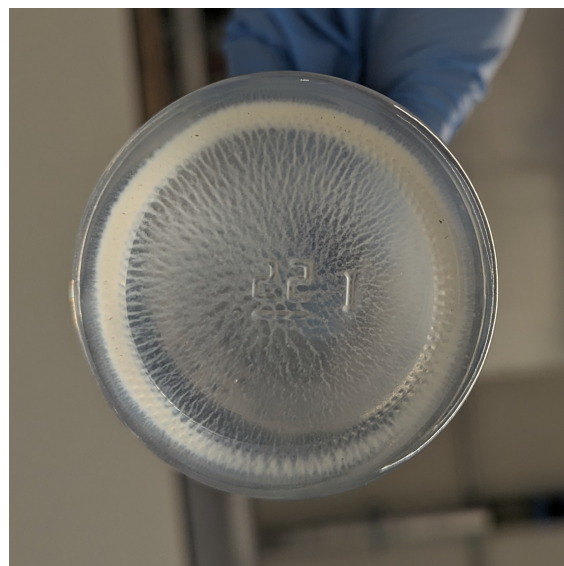
4.3.2 Experimental methods

4.3.2.1 Maintenance of stock cultures

Stock cultures of both *M. methanica* S1 or *M. trichosporium* OB3b were maintained between subsequent experiments using a semi-continuous batch method in which headspace and media were replaced on a weekly basis. Cultures were grown in 3x 250ml serum bottles with a 50:50 ratio of headspace and media. The initial CH_4 concentration was close to but lower than 8%v/v in air to avoid oxygen limitation. The bottles were pressurised through the addition of air to avoid negative pressure within the bottles; even in the event of full CH_4 consumption. Measurements of headspace pressure, pH and OD_{600} were taken on a weekly basis to assess culture health. The amount of media replaced each week was altered to ensure cultures were maintained at an OD_{600} of 0.100. Cultures were examined on a routine basis using a EVOS microscope to assess for any potential contamination. Stocks used for inoculation were in the growth phase. These cultures were grown at $30^\circ C$ with no mixing. As a result sedimentation of bacteria occurred. The physical manifestation of this differed between both species where *M. methanica* S1 favoured the formation of flocs and larger aggregates, while *M. trichosporium* OB3b developed a structured sediment. Photos highlighting this differences are provided in Figures 4.1a and 4.1b respectively.



(a)



(b)

Figure 4.1: Photos of the bottom of stock culture bottles containing either *Methylobacterium methanica* S1 or *Methylobacterium trichosporium* OB3b grown at 30°C with no shaking. a) *Methylobacterium methanica* S1, b) *Methylobacterium trichosporium* OB3b.

4.3.2.2 Experimental conditions

The aim is to capture both the growth kinetics and the thermal output, however, these cannot be measured simultaneously. This because the isothermal conditions within the calorimeter must be maintained; opening the calorimeter would release heat and affect the temperature, ruining the measurement of heat flux. Bottles within the calorimeter therefore cannot be accessed and sampling them to measure physicochemical parameters can only be undertaken prior to their addition to the calorimeter and once they are removed. Thus, the same conditions were mimicked externally using an incubator, and cultures within this incubator were grown in parallel with those within the calorimeter. These external cultures were sacrificed at daily intervals to characterise the dynamics of growth and substrate use. This experimental setup is highlighted in Figure 4.2. Green bottles indicate active cultures while blue bottles refer to negative controls. As shown, sets of bottles were separated between the acquisition of thermal output (CAL1-CAL4) and growth dynamics (C1-C7, S1-S7, D1-D7 and T1-T7). Using a batch method all cultures were grown for a period of 7 days at 30°C with no shaking. The aforementioned bespoke bottles were used for all cultures.

For the setup depicted in Figure 4.2, one litre of inoculum was used. This was made by mixing 100ml of a stock culture, at $OD_{600} = 0.100$, with 900ml of fresh NMS media. The inoculum was shaken within a MAAQ incubator at a speed of 150rpm for at least 1hr at 20°C prior to use, to ensure uniformity. Before the inoculation of active cultures, control cultures were made up by adding 30ml of NMS media to each of the control bottles. These bottles were then sealed using a rubber stopper and 20mm aluminium crimp seal. To avoid any unnecessary increase in pressure due to the sealing process, the top of each bottle was pierced with a needle to equalise the headspace pressure with the environment. To active cultures 30ml of the inoculum was added. These bottles were sealed using the same method as for the control bottles.

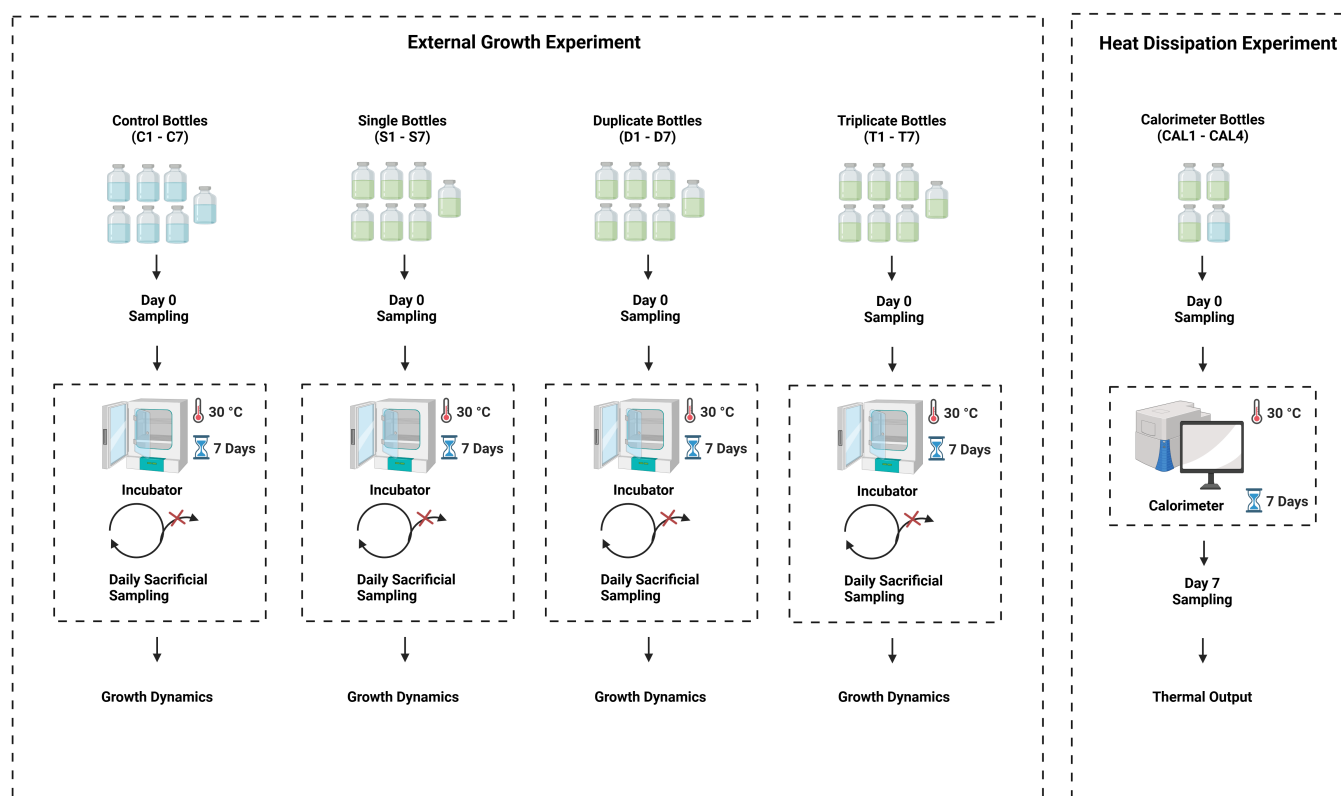


Figure 4.2: Overview of the experimental layout, demonstrating both "External Growth Experiment" and "Heat Dissipation Experiment" setups. Green bottles represent active cultures while blue bottles represent negative controls. Both experiments are undertaken in unison with the same materials, growth conditions and inoculum used. All cultures were grown as pure batch cultures. Experiments were repeated for both *M. methanica* S1 and *M. trichosporium* OB3b over three distinct initial concentrations of CH_4 (8%, 4% and 2% v/v in air). Created in <https://BioRender.com>

Gases were added to sealed bottles using a series of injections. All injections were made using a 5mL Hamilton gastight syringe model 1005 SL fitted with a sample lock and male luer RN hub (L) adapter. 25G x 5/8" needles were added to the end of this syringe to facilitate piercing the rubber stoppers. In total, four injections were made; adding a total of 19.10ml of gas to each bottle. The total amount of gas added was calculated based on the need to ensure pressurisation throughout the entire growth period; as sampling depressurised systems can invoke errors when sampling gases (Stone et al., 2019). This calculation can be found in Appendix G. A process of trial and error was used to determine the composition of the injected gas required to ensure initial concentrations of 8%, 4% and 2% v/v CH_4 in air. These compositions are given in Table 4.2.

4.3.2.3 Heat dissipation experiments

To measure the thermal power output during batch culture of *M. methanica* S1 and *M. trichosporium* OB3b, both strains were grown within a Calmetrix I-Cal 4000 HPC Isothermal Calorimeter. The Calmetrix I-Cal 4000 HPC contains four test chambers, three were used to house active cultures, representing biological triplicates, while the fourth was used as a control, as shown in Figure 4.2. Day 0 samples were taken prior to the addition of bottles to the calorimeter.

Table 4.2: Injection volumes of CH_4 and air used during inoculation to provide target CH_4 concentrations of 8%, 4%, and 2%v/v in air. Volumes were determined through trial and error.

Target CH_4 concentration (%v/v in air)	Injection 1		Injection 2		Injection 3		Injection 4		Total	
	CH_4 (ml)	Air (ml)	CH_4 (ml)	Air (ml)	CH_4 (ml)	Air (ml)	CH_4 (ml)	Air (ml)	CH_4 (ml)	Air (ml)
8	4.10	0.00	0.00	5.00	0.00	5.00	0.00	5.00	4.10	15.00
4	2.05	2.05	0.00	5.00	0.00	5.00	0.00	5.00	2.05	17.05
2	1.00	3.10	0.00	5.00	0.00	5.00	0.00	5.00	1.00	18.10

Cultures were left within the calorimeter for a period of 7 days with the internal temperature of the calorimeter set at 30°C. To ensure a stable temperature the calorimeter was set at 30°C for at least 24 hours before bottles were added. In line with manufacturer recommendations, the factory calibration was used with a reference mass of 100g. Readings of thermal power output were obtained from each chamber once per minute using the associated CalCommander software. Post-processing of results was required to remove baseline offsets and account for the settling period of the calorimeter. A full description of the post-processing steps is provided in Appendix H. On removal from the calorimeter final pressure measurements were made and headspace sample taken. Samples from the liquid culture were then taken for measuring growth parameters. In total the experiment was repeated six times, to facilitate results for both *M. methanica* S1 and *M. trichosporium* OB3b under initial CH_4 concentrations of 8%, 4% and 2%v/v in air.

4.3.2.4 External growth experiments

External growth experiments were run in parallel to the calorimeter experiments to monitor growth parameters over the entire growth period. To ensure experiments offered an accurate reflection of growth occurring within the calorimeter, materials, growth conditions and inoculum were kept identical across the two sets of experiments. Both sets of experiments were undertaken at the same time and analysis of samples was carried out in unison. External cultures were grown within an MAAQ incubator at 30°C with no shaking to match the conditions within the calorimeter.

As a large proportion of the total culture volume was required for analytical analysis, sacrificial sampling was implemented within these experiments. To conduct a daily sampling regime over the 7-day experimental period, sets of 7 identical cultures were grown. Three of these sets were required to produce biological replicates. A further 7 cultures were used as negative controls, increasing the overall total to 28. This layout of bottles is highlighted in Figure 4.2. In line with the heat dissipation experiments, the external growth experiments were conducted a total of six times, facilitating results for both *M. methanica* and *M. trichosporium* OB3b under initial CH_4 concentrations of 8%, 4% and 2%v/v in air.

4.3.2.5 Sampling

Sampling was required for the analysis of optical density (OD_{600}), headspace composition, dissolved oxygen, pH and non-purgeable organic carbon (NPOC). Headspace pressure was also measured, however, this analysis was undertaken directly on the bottles themselves and thus no sample was required. In total four separate samples were taken, 3 aqueous and 1 gaseous: a single 5ml sample for measuring OD_{600} and pH; a 3ml sample for the dissolved oxygen measurement and; a 10ml sample for the measurement of NPOC. To measure the headspace composition of CH_4 and CO_2 , a single 3ml

gaseous sample was extracted from the bottle headspace. Headspace samples were removed using 5mL Hamilton gastight syringe model 1005 SL fitted with a sample lock and male luer RN hub (L) adapter. Aqueous samples were stored in sterile 15ml centrifuge tubes, while gaseous samples were added to pre-vacuumed Exetainer® 3ml glass vials with double wadded caps.

For each bottle, sampling was conducted twice; once prior to experiments beginning (day 0), and once at the end of the desired growth period (day 7 for calorimeter bottles, day 1-7 for external cultures). Day 0 sampling was conducted on the stocks of inoculum and NMS media used. Day 0 headspace samples were taken directly from individual bottles. All final samples were taken directly from bottles, with gaseous samples taken first and aqueous samples taken thereafter. Cultures were shaken before taking aqueous samples to ensure a uniformly mixed culture. Headspace samples were stored for future analysis, which took place at the Scottish Universities Environmental Research Centre (SUERC). Samples for OD_{600} , pH and dissolved oxygen measurements were analysed within 2 hours of being taken, while samples for *NPOC* measurement were analysed within 12 hours of being taken.

4.3.3 Analytical methods

4.3.3.1 Headspace pressure

Measurements of headspace pressure were taken using a Greisinger GDH 200 Manometer by attaching a 25G x 5/8" needle and 8mm (outer diameter) PVC tubing to the positive port on the manometer. By piercing the rubber stopper of each bottle with the needle the headspace pressure could be measured. To minimise the risk of cross contamination, needles were replaced between injections and measurements were made on control bottles prior to active cultures. For each bottle, readings were obtained on three separate occasions: firstly, after the addition of gases to determine the maximum pressure; secondly, after the first headspace sample was taken to record the initial pressure and; finally, a measurement was taken after the growth period had been completed to record the final pressure. In the first two instances, the pressure was measured at room temperature ($\approx 22^{\circ}\text{C}$). For the final reading pressure was measured at 30°C . Pressure measurements were primarily used to facilitate calculations of molar mass, while also ensuring pressure was within a range that did not negatively impact methanotrophic growth (Soni et al., 1998).

4.3.3.2 Optical density

The measurement of optical density (OD_{600}) was undertaken using a HACH DR1900 spectrophotometer. Measurements of absorption were made using a single wavelength of 600nm. The HACH DR1900 was zeroed using NMS media. Samples, stored within 15ml centrifuge tubes, were mixed at 2500rpm for 5 seconds to break up biomass aggregates and ensure a uniform sample. The sample was aliquoted into 3x 1ml micro-cuvettes for measurement, with the final value determined by calculating the mean across all three measurements. In all instances, fresh cuvettes were used, removing the need for a cleaning methodology. Effort was made to remove any bubbles within a cuvette before the measurement was taken. This was done by either lightly tapping the cuvette against a surface or fishing out bubbles using a sterile needle. After measurement, samples were added back to their 15ml tubes for analysis of pH. All measurements of optical density were taken at room temperature.

4.3.3.3 Cell dry weight

A calibration curve between OD_{600} and cell dry weight (CDW) was obtained to facilitate conversion between these two quantities. To measure CDW stock cultures of *M. methanica* S1 and *M. trichosporium* OB3b were diluted to OD_{600} values of approximately 0.025, 0.050, 0.075, 0.100, 0.125 and 0.150. Samples were then vacuumed filtered using $0.2\mu m$ cellulose acetate membrane filters. Prior to use, these filters were dried at $105^{\circ}C$ for 12 hours and then left in a desiccator for 1 hour before initial measurements of weight were recorded. After filtration filters once again dried at $105^{\circ}C$ for 12 hours and then left in a desiccator for 1 hour before final measurements of weight were recorded. For samples with an OD_{600} of ≤ 0.100 100ml was filtered. For samples with an OD_{600} of > 0.100 50ml was filtered. The full calibration curves obtained can be found in Appendix I.

4.3.3.4 pH

pH was measured using an Orion Star A215 pH Meter. Three-point calibration of the probe was undertaken before each set of measurements. This was done using standards with pHs of 4, 7 and 10. pH measurements were conducted on the same samples used for the determination of optical density. The probe was submerged into 5ml of sample, stored in a 15ml centrifuge tube, and kept there until a stable reading was obtained. The probe was rinsed with Milli-Q water between each reading. Measurements of pH were taken at room temperature and only obtained after measurements of OD_{600} .

4.3.3.5 Dissolved oxygen

Dissolved oxygen measurements were made using a Vernier Go Direct Optical Dissolved Oxygen probe. The sample volume used for dissolved oxygen measurements was 3ml stored within a sterile 15ml centrifuge tube. The probe was washed with Milli-Q water and wiped dry before use. Readings were obtained via bluetooth using the Vernier Graphical Analysis app. Readings were taken over a three-minute period. To monitor the instrument variance, readings of Milli-Q water were taken on each day of use for comparison. Before readings commenced a visual check was made to ensure no bubbles were present on or around the optical probe base. If bubbles were present, the probe was removed and re-inserted. Measurements were made to primarily check that dissolved oxygen remained within the system even after the full consumption of CH_4 , ensuring growth was indeed carbon limited. These measurements can be found in Appendix J.

4.3.3.6 Non-purgeable organic carbon

Measurements of non-purgeable organic carbon ($NPOC$) were made to quantify the carbon within biomass. This is especially useful to account for the carbon partitioned for biosynthetic processes during growth. $NPOC$ was chosen instead of total organic carbon (TOC) as it acts to remove volatile organic carbon, such as CO_2 and CH_4 , from the sample before measurements of organic carbon are made. As CO_2 and CH_4 within the system is already determined through gas chromatography, the aim of $NPOC$ measurements were only to account for cellular organic carbon. Glass vials were baked within an oven at $400^{\circ}C$ overnight prior to use. Samples of 10ml were taken from each culture for measurement of $NPOC$. These were mixed and then transferred to 10ml baked glass vials. $NPOC$ was measured using a Shimadzu TOC-L. In accordance with this equipment, acidification and sparging is conducted prior to non-dispersive infrared (NDIR) detection. The Shimadzu TOC-L was calibrated up to 25mg/L.

4.3.3.7 Headspace composition

Headspace composition was determined using gas chromatography. Headspace samples were stored within 3ml DW Exetainer® Vials and transported to the Scottish Universities Environmental Research Centre (SUERC) for analysis. Measurements of CH_4 and CO_2 were made using a Hewlett Packard 5890 Gas Chromatograph with a thermal conductivity detector (TCD). Helium was used as the carrier gas and a Hayesep-N polymer packed column was used with a column head pressure of 20psi. The initial temperature was set at 50°C for 0.5 minutes followed by a ramp rate of 40°C per minute until a final temperature of 150°C. Readings were obtained from two 50µL injections of each sample, with the final reading expressed as the mean. To avoid dilution of sample due to air within the needle itself and to normalise sample pressure, 100µL was initially taken within the syringe, 50µL of this was purged immediately prior to inserting the rest of the sample for analysis. Calibration and maintenance of the equipment was facilitated by SUERC, and prior to use, the column was baked at 150°C for at least 45min or until the measurement baseline returned to a constant reading. Zeroing of the equipment was undertaken before each sample was measured.

4.4 Results and Discussion

4.4.1 Growth dynamics and yields

Measurements of the headspace concentration of CH_4 and CO_2 , OD_{600} , $NPOC$, system pressure and pH, for both *M. methanica* S1 and *M. trichosporium* OB3b are provided in Figures 4.3 and 4.4 respectively. Results are shown for three distinct initial concentrations of CH_4 (8%, 4% and 2%v/v in air). These concentrations of CH_4 were selected based on expected CH_4 concentrations within a compost biofilter, as defined by Gómez-Borraz et al. (2025), and those observed within septic tanks (to be published). The overall length of the experiments was determined by the longest period required for all CH_4 to be consumed. This was determined to be one week as demonstrated by Figure 4.4a. The sampling period was kept uniform across all experiments. Sampling was spaced equally throughout this period with sampling conducted every 24 hours. For *M. methanica* S1, when the initial concentration of CH_4 was set at 4% and 2%v/v in air, the sampling frequency was doubled to every 12 hours for the first three and two days respectively. This was done to ensure a higher resolution within the dynamic period. To achieve this, the sampling frequency of subsequent samples were reduced. This change in sampling period is displayed in Figure 4.3. As measurements from the calorimeter cultures could only be made at the start and end of the growth period, growth dynamics are primarily obtained from external cultures grown in parallel to those within the calorimeter. The final values reached for all quantities measured from both external and calorimeter cultures are expressed in Table 4.3, where the values presented demonstrate the mean across biological triplicates; the standard deviation within these values is also provided.

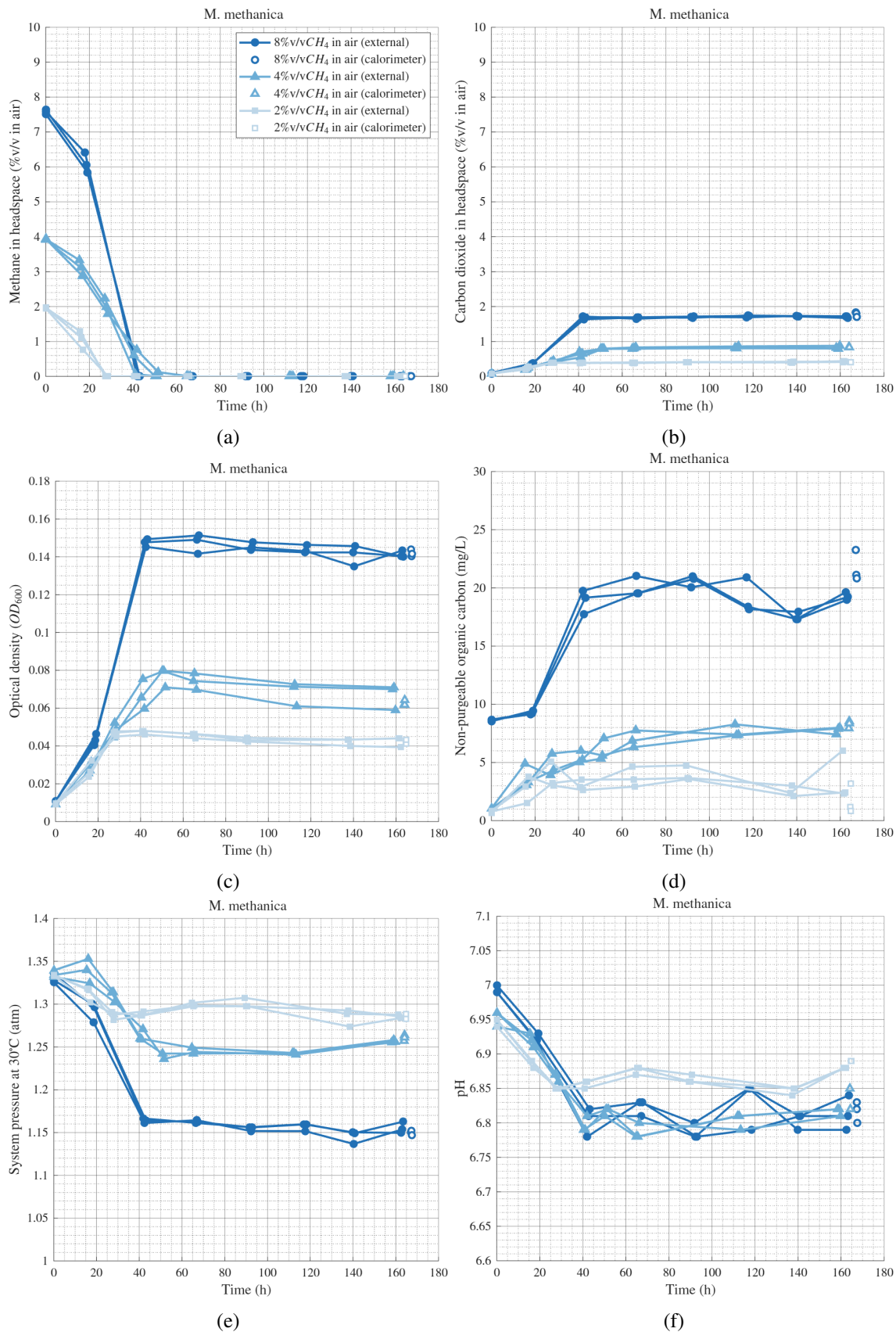


Figure 4.3: Time series of growth parameters for external and calorimeter cultures of *M. methanica* S1, measured under three distinct concentration of initial CH_4 over a period of 7 days. a) Headspace CH_4 concentration, b) Headspace CO_2 concentration, c) Optical density at 600nm (OD_{600}), d) Non-purgeable organic carbon (NPOC), e) System pressure, f) pH.

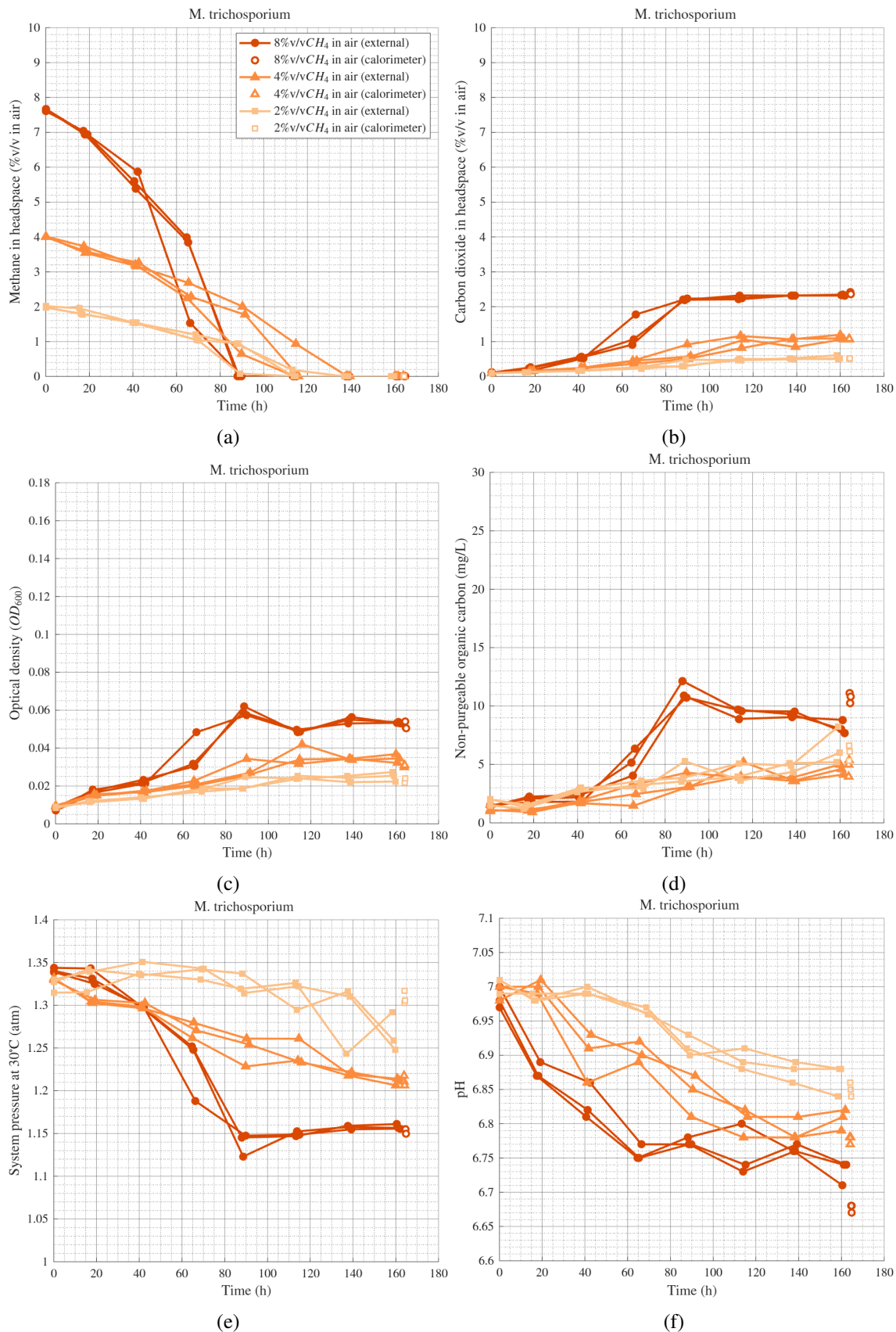


Figure 4.4: Time series of growth parameters for external and calorimeter cultures of *M. trichosporium* OB3b, measured under three distinct concentration of initial CH₄ over a period of 7 days. a) Headspace CH₄ concentration, b) Headspace CO₂ concentration, c) Optical density at 600nm (*OD*₆₀₀), d) Non-purgeable organic carbon (NPOC), e) System pressure, f) pH.

Table 4.3: Final values measured for the headspace concentration of CO_2 , OD_{600} , $NPOC$, headspace pressure and pH for pure batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b from external and calorimeter experiments. Means are calculated from biological triplicates. The standard deviation in these values is also expressed.

Species	Experiment	Initial CH_4 (%v/v in air)		CO_2 (%v/v in air)		OD_{600} (—)		$NPOC$ (mg/L)		Pressure (atm)		pH (—)	
		mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
<i>M. methanica</i>	Calorimeter	7.73	0.13	1.78	0.07	0.142	0.002	21.75	1.33	1.150	0.003	6.82	0.02
		3.91	0.06	0.83	0.00	0.063	0.002	8.30	0.31	1.261	0.004	6.84	0.02
		1.96	0.03	0.41	0.00	0.042	0.001	1.73	1.29	1.285	0.003	6.89	0.00
<i>M. methanica</i>	External	7.56	0.11	1.70	0.03	0.141	0.002	19.27	0.33	1.155	0.007	6.81	0.03
		3.89	0.06	0.83	0.04	0.067	0.007	7.77	0.32	1.256	0.002	6.82	0.01
		1.96	0.03	0.42	0.01	0.042	0.003	3.59	2.10	1.286	0.002	6.88	0.00
<i>M. trichosporium</i>	Calorimeter	7.73	0.13	2.38	0.03	0.052	0.002	10.72	0.45	1.153	0.003	6.68	0.01
		4.01	0.04	1.07	0.01	0.031	0.001	4.78	0.75	1.212	0.006	6.78	0.01
		1.98	0.02	0.51	0.00	0.023	0.001	6.00	0.63	1.308	0.009	6.85	0.01
<i>M. trichosporium</i>	External	7.67	0.03	2.33	0.01	0.053	0.000	8.19	0.57	1.158	0.003	6.73	0.02
		4.00	0.03	1.12	0.07	0.035	0.002	4.58	0.43	1.211	0.004	6.81	0.02
		2.01	0.04	0.54	0.06	0.025	0.003	6.47	1.57	1.266	0.023	6.87	0.02

Across both *M. methanica* S1 and *M. trichosporium* OB3b (Figures 4.3 and 4.4 respectively), OD_{600} and $NPOC$ is shown to increase as CH_4 is consumed, representing methanotrophic growth. An increase in the headspace concentration of CO_2 is also observed as a result of the full oxidation of CH_4 . System pressure is shown to decrease as CH_4 is consumed and pH is also shown to decrease as a consequence of CO_2 being produced. The delay in growth and the reduced growth rate experienced by *M. trichosporium* OB3b mirrors results of S. Park et al. (1991), and is thus a consequence of a lack of CO_2 initially within the system. This lack of CO_2 only affects the anabolic pathway within type-II methanotrophs by inhibiting the conversion of phosphoenolpyruvate to oxaloacetate within the serine cycle.

To summarise the differences between the growth of *M. methanica* S1 and *M. trichosporium* OB3b, yields of cell dry weight (CDW), $NPOC$ and CO_2 were calculated. This was achieved using the relationship:

$$Y_{X/S} = \frac{\Delta X}{\Delta S}, \quad (4.1)$$

where, ΔX is the change in mass of a known product over the entire growth period and ΔS is the change of mass of substrate over the same period. For calculating the yield of CO_2 , the number of moles is used instead of overall mass. To determine the total number of moles of CH_4 and CO_2 , the number of moles within the headspace, and the proportion dissolved within the media must be calculated. Calculation of the number of moles of each compound within the headspace is achieved through the ideal gas law, where the total number of moles of gas is calculated based on the measured pressure (Figures 4.4e and 4.3e), the volume of headspace in each bottle (Appendix F) and the temperature. The relative proportions of CH_4 and CO_2 are then calculated using the measured concentrations of these gases (Figures 4.4a, 4.4b, 4.3a and 4.3b). The number of moles dissolved are calculated using Henry's law, where an adjustment is made to correct Henry's coefficients based on temperature. The coefficients used and equation for adjusting them based on temperature dependence can be found in Sander (2023). A large proportion of CO_2 will be in the form of HCO_3^- based on the pH observed for both sets of experiments (Figures 4.3f and 4.4f). This quantity is calculated based on the pH observed and

the dissociation constant of H_2CO_3 adjusted to the required temperature. An example calculation determining the total number of moles of CH_4 and CO_2 can be found in Appendix K. From these values the total mass of CH_4 is subsequently determined by multiplying the total number of moles by the respective molecular weight. The mass of CDW is calculated based on the calibration curves given within Appendix I. Thus, for *M. methanica* S1 $OD_{600} = 1.0 = 358mg/L$ and for *M. trichosporium* OB3b $OD_{600} = 1.0 = 497mg/L$. $NPOC$ is directly measured as mg/L . To determine the total mass of these quantities they are multiplied by the total volume of culture media (30ml).

Carbon conversion efficiency (CCE) is calculated from yields of CDW and $NPOC$. These are determined using the equation:

$$CCE = \left(\frac{\text{carbon in biomass}}{\text{total carbon consumed}} \right) \times 100, \quad (4.2)$$

where, the carbon stored in biomass can be calculated from CDW by assuming cells contain 47% carbon (Leak & Dalton, 1986a), for $NPOC$ the total carbon in biomass is directly measured. The total carbon consumed is calculated by multiplying the initial number of moles of CH_4 previously calculated, by the molar mass of carbon.

Yields and $CCEs$ are calculated per culture, as such, the final values provided in Table 4.4 are obtained by taking the mean across biological triplicates. The standard deviation (SD) in these values is also calculated and provided alongside the mean values. Yields and $CCEs$ calculated for the external and calorimeter cultures are expressed separately.

Table 4.4: Final yields of cell dry weight (Y_{CDW/CH_4}), non-purgeable organic carbon (Y_{NPOC/CH_4}) and carbon dioxide (Y_{CO_2/CH_4}) calculated for pure batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b. Carbon conversion efficiencies calculated from cell dry weight (CCE_{CDW}) and non-purgeable organic carbon (CCE_{NPOC}) are also provided. Calculations are undertaken using the measurements provided in Figures 4.3 and 4.4 respectively. Means are calculated from biological triplicates. The standard deviation in these values is also expressed.

Species	Experiment	Initial CH_4 (%v/v in air)		Y_{CDW/CH_4} (—)		CCE_{CDW} (%)		Y_{NPOC/CH_4} (—)		CCE_{NPOC} (%)		Y_{CO_2/CH_4} (—)	
		mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
<i>M. methanica</i>	Calorimeter	7.73	0.13	0.53	0.00	33.45	0.08	0.15	0.01	20.42	1.83	0.59	0.03
		3.91	0.06	0.43	0.01	27.19	0.71	0.17	0.01	22.08	0.89	0.55	0.01
		1.96	0.03	0.56	0.01	35.04	0.68	0.05	0.06	6.32	7.51	0.43	0.04
<i>M. methanica</i>	External	7.56	0.11	0.55	0.00	34.31	0.09	0.13	0.01	16.98	0.72	0.58	0.01
		3.89	0.06	0.47	0.06	29.64	3.57	0.16	0.01	20.77	1.21	0.56	0.05
		1.96	0.03	0.55	0.04	34.32	2.35	0.13	0.09	17.38	12.54	0.48	0.03
<i>M. trichosporium</i>	Calorimeter	7.73	0.13	0.25	0.01	15.41	0.62	0.11	0.00	14.18	0.63	0.66	0.00
		4.01	0.04	0.22	0.01	13.86	0.38	0.08	0.02	10.94	2.21	0.57	0.01
		1.98	0.02	0.30	0.03	19.14	1.67	0.18	0.03	24.14	3.74	0.49	0.01
<i>M. trichosporium</i>	External	7.67	0.03	0.26	0.00	16.16	0.16	0.08	0.01	10.48	1.02	0.70	0.01
		4.00	0.03	0.27	0.03	16.94	1.61	0.08	0.01	10.07	0.79	0.64	0.02
		2.01	0.04	0.36	0.05	22.40	3.44	0.21	0.06	27.88	8.61	0.56	0.08

The yields in Table 4.4 reveal the differences between the growth of batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b under the conditions investigated. *M. methanica* S1 is shown to generate a much higher yield of CDW when compared to *M. trichosporium* OB3b, this is shown to be the case across cultures of all initial concentrations of

CH_4 (8%, 4% and 2%v/v in air). The yields of $NPOC$ do not observe the same relationship. While Y_{NPOC/CH_4} for *M. methanica* S1, for the 8% and 4%v/v CH_4 experiments, are higher than that of *M. trichosporium* OB3b, for our experiments at 2%v/v CH_4 , $NPOC$ yields are shown to be greater for *M. trichosporium* OB3b. This contradicts what is observed for Y_{CDW/CH_4} . Indeed when comparing the yields and carbon conversion efficiencies calculated from cell dry weight (CCE_{CDW}) to those calculated from measurements of non-purgeable organic carbon (CCE_{NPOC}), differences are observed between these values. The most likely explanation for this is due to the time required for equipment calibration, samples within the Shimadzu TOC-L were not analysed until at least 1 hour after they had been mixed prior to being placed within the machine. Over this period samples stood undisturbed, and as a result some sedimentation of cell mass occurred. This would happen at a greater rate for *M. methanica* S1 than *M. trichosporium* OB3b due to the prevalence of cell aggregates. Indeed some sediment was found on the bottom of these vials after $NPOC$ measurements had been made. For this reason we consider the $NPOC$ measurements to under-represent the correct carbon content and base our analysis primarily on the measurements of CDW . Regarding the yield of CO_2 (Y_{CO_2/CH_4}), the results provided in table 4.4 show that cultures of *M. trichosporium* OB3b produce more CO_2 than cultures of *M. methanica* S1 across all concentration of CH_4 investigated.

To observe if yields depend on the initial concentration of CH_4 , the final yields of cell dry weight (Y_{CDW/CH_4}) and CO_2 (Y_{CO_2/CH_4}) for each species are plotted against the initial concentration of CH_4 used (Figure 4.5). Trendlines are generated using all points for a given species. For *M. trichosporium* OB3b a weak negative correlation between Y_{CDW/CH_4} and the initial concentration of CH_4 is observed, with the R^2 value for this fit equal to 0.33 and the gradient of the linear fit calculated as -0.0120 per %v/v (Figure 4.5a). For *M. methanica* S1 no clear relationship between these two quantities is observed. Indeed, the p-value of the gradient coefficient of the fit is > 0.05 , thus, we fail to reject the null hypothesis. For both *M. methanica* S1 and *M. trichosporium* OB3b, positive linear correlations between Y_{CO_2/CH_4} and the initial concentration of CH_4 is observed (Figure 4.5b); where R^2 values are 0.61 and 0.68 respectively. The gradient of this correlation is greater for *M. trichosporium* OB3b than *M. methanica* S1, with gradients calculated as 0.027 and 0.020 per %v/v respectively.

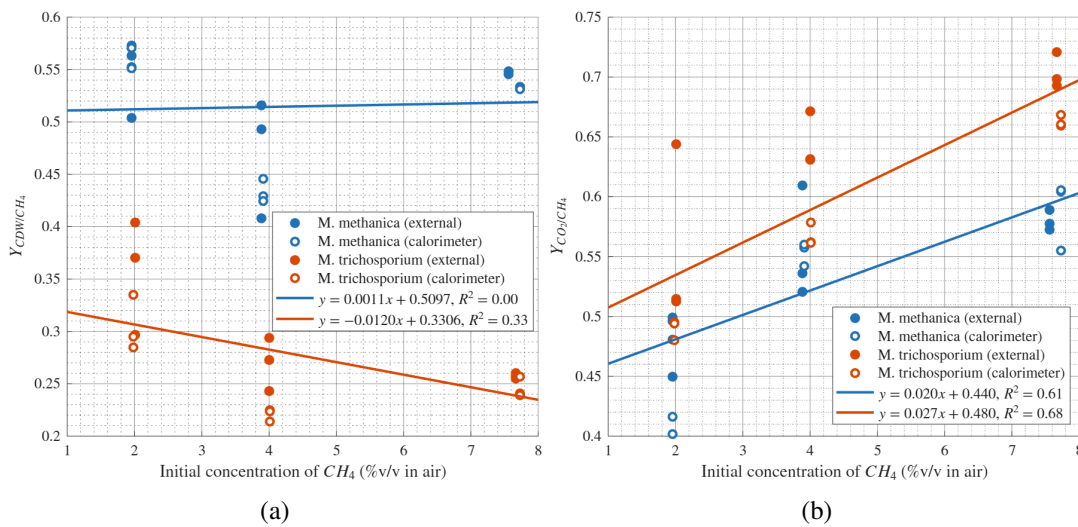


Figure 4.5: Final yields of cell dry weight (Y_{CDW/CH_4}), and carbon dioxide (Y_{CO_2/CH_4}) plotted against initial headspace concentration of CH_4 for calorimeter and external batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b. Linear fits are made using all points plotted for *M. methanica* S1 and *M. trichosporium* OB3b respectively. Equations of these lines and the R^2 values are provided in figure legends. a) Y_{CDW/CH_4} , b) Y_{CO_2/CH_4} .

4.4.2 Carbon balance and partitioning

Given that CH_4 is both the carbon and energy source for our methanotrophic cultures, we make effort to track how both strains partition carbon. First this is done by carrying out a carbon balance analysis. Here we assumed that at the start of each experiment the total amount of carbon present is made up of: the biomass within the inoculum and the CH_4 and CO_2 within the system initially. Here the amount of carbon in biomass is calculated from CDW rather than $NPOC$ for the reasons provided earlier. Again, it is assumed that CDW consist of 47% carbon (Leak & Dalton, 1986a). The final carbon contributions are calculated from the same sources. Within Figure 4.6 the proportion each source makes up of the initial carbon is displayed as a percentage. These proportions are plotted across all samples for every experiment. The first three columns of each set within Figure 4.6 represent the biological triplicates, the fourth column displays a negative control. "EX1-7" refer to samples taken from the external growth experiments over the growth period, and correspond chronologically to the "external" points displayed in Figures 4.3 and 4.4 respectively. Samples marked "CAL" refer to samples taken from the calorimeter cultures at the end of the experimental period, these correspond to the "calorimeter" sample points within Figures 4.3 and 4.4 respectively.

While the graphs displayed in Figure 4.6 clearly depict the transition from CH_4 to biomass and CO_2 . They also demonstrate that a proportion of carbon is unaccounted for. One reason for this is that during measurements of system pressure, taken after the initial headspace sample and before the final headspace sample, the total volume of the system is increased by the internal volume of the manometer. Due to the diffusion of CH_4 and CO_2 across this volume, the concentration of carbon within the headspace will decrease. By comparing the initial carbon composition to the final carbon composition across all control samples, a decrease of between 3.9% and 13.7% is observed (Figure 4.6). While this does demonstrate that indeed some carbon is lost as a result of the methodology used, for samples taken towards the end of the growth period, especially during experiments at lower initial concentration of CH_4 , this reduction does not sufficiently explain the amount of unaccounted carbon observed.

It is known that methanotrophs have the potential to secrete metabolites during growth (Kalyuzhnaya et al., 2013; Kalyuzhnaya et al., 2015; Roslev & King, 1995). Indeed, in some cases, extracellular carbon in the form of metabolites has been observed to account for 40 – 50% of the total carbon within the CH_4 consumed (Kalyuzhnaya et al., 2013). A known pathway of metabolite excretion is unique to type-II methanotrophs and arises when the equilibrium of the catabolic reaction is shifted by excess CO_2 , driving the oxidation of CH_4 in the reverse direction. Due to the inability of methane monooxygenase (MMO) to catalyse the reaction from CH_3OH to CH_4 , this energy-intensive process results in the synthesis of methanol (CH_3OH), which is then excreted to the media (Sahoo et al., 2021, 2023; Xin et al., 2007). Given that excreted metabolites are not captured within OD_{600} measurements, one would expect a reduction within system carbon as a result of their formation. Without a more detailed accounting of the carbon within the system the form of the additional unaccounted carbon can only be speculated at. Despite this, the consistency of the methodology used does facilitate the comparison of carbon partitioned between biomass and CO_2 .

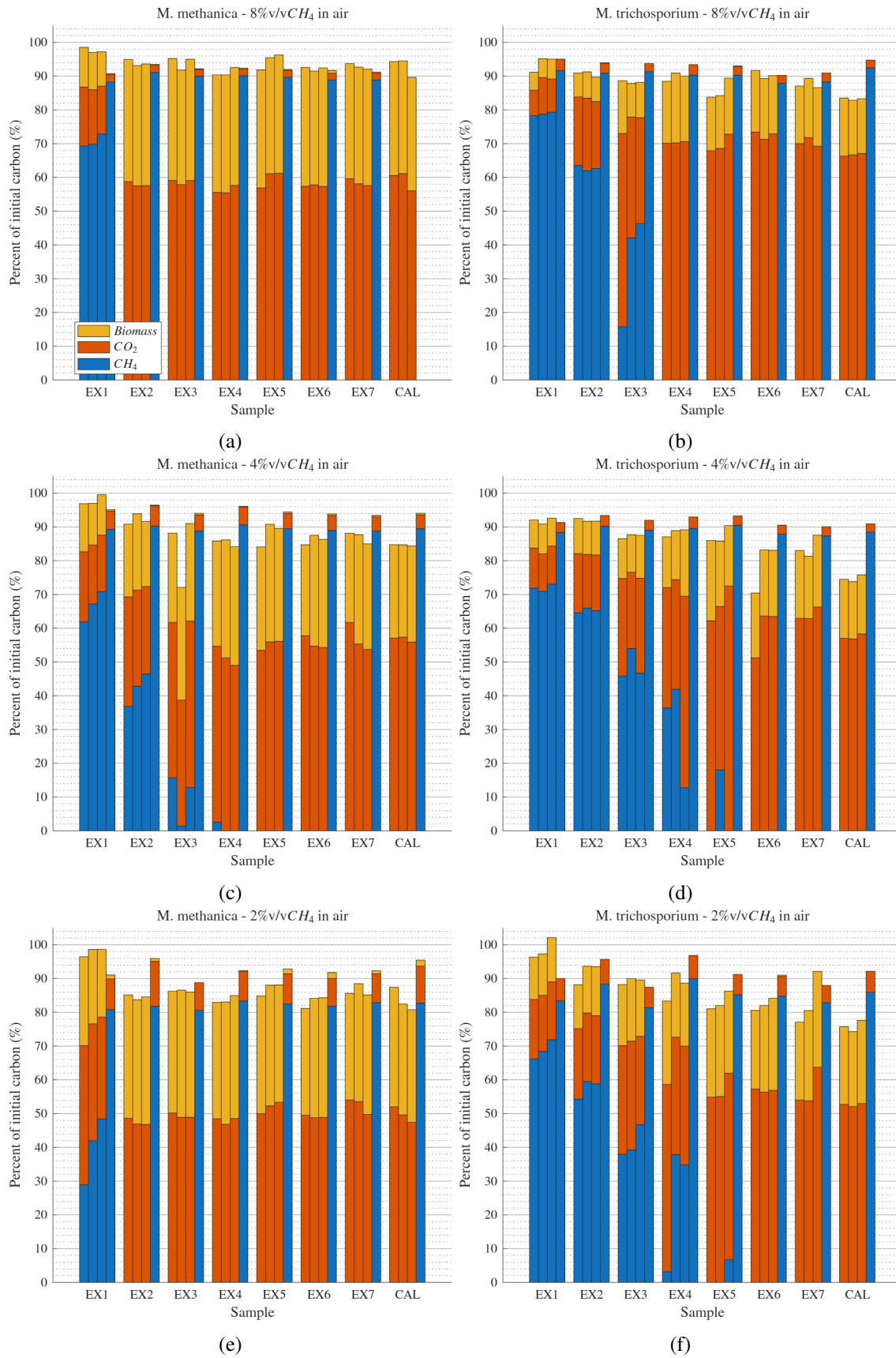


Figure 4.6: Carbon composition of CH_4 , CO_2 and biomass expressed as a percent of the initial carbon within the system. Measurements of CH_4 , CO_2 and biomass are obtained from batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b over three different initial concentration of CH_4 , as shown in Figures 4.3 and 4.4 respectively. a) *M. methanica* S1 8%v/v CH_4 in air b) *M. trichosporium* OB3b 8%v/v CH_4 in air, c) *M. methanica* S1 4%v/v CH_4 in air, d) *M. trichosporium* OB3b 4%v/v CH_4 in air, e) *M. methanica* S1 2%v/v CH_4 in air. f) *M. trichosporium* OB3b 2%v/v CH_4 in air.

To further explore the partitioning of carbon we can present the the amount of carbon in the produced biomass and CO_2 as a percent of the carbon in the CH_4 consumed, these values are supplied in Table 4.7, where only the final carbon compositions presented in Figure 4.6 are used for calculations. To summarise this trade-off, here we define a carbon partitioning ratio (CPR). This is achieved by dividing the relative percentage of carbon funnelled to CO_2 , by the relative percent of carbon converted to CDW or biomass, thus, a ratio of > 1 indicates that more of the carbon stored in the initial quantity of CH_4 is present within the CO_2 measured than the CDW , and the opposite is observed when $CPR < 1$. These ratios are calculated using the values within Table 4.7, with the ratios themselves provided in the final column.

Table 4.5: Final percentages of carbon stored in the initial volume of CH_4 converted to biomass and CO_2 , measured for calorimeter and external batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b. The ratio of the carbon converted to CO_2 and that stored in biomass is provided. All values are calculated as means from biological triplicates, with standard deviations also shown.

Strain	Experiment	Initial CH_4 (%v/v in air)		Biomass (%)		CO_2 (%)		CPR (—)	
		mean	SD	mean	SD	mean	SD	mean	SD
<i>M. methanica</i>	Calorimeter	7.73	0.13	33.45	0.08	58.84	2.88	1.76	0.08
		3.91	0.06	27.19	0.71	55.33	0.97	2.04	0.09
		1.96	0.03	35.04	0.68	43.29	4.19	1.23	0.10
<i>M. methanica</i>	External	7.56	0.11	34.31	0.09	57.96	0.84	1.69	0.03
		3.89	0.06	29.64	3.57	55.53	4.75	1.91	0.41
		1.96	0.03	34.32	2.35	48.16	2.77	1.41	0.16
<i>M. trichosporium</i>	Calorimeter	7.73	0.13	15.41	0.62	66.26	0.48	4.31	0.19
		4.01	0.04	13.86	0.38	56.74	0.96	4.10	0.10
		1.98	0.02	19.14	1.66	48.98	0.83	2.57	0.19
<i>M. trichosporium</i>	External	7.67	0.03	16.16	0.16	70.42	1.49	4.36	0.06
		4.00	0.03	16.94	1.61	64.44	2.33	3.82	0.27
		2.01	0.04	22.4	3.44	55.71	7.54	2.50	0.28

The calculated ratios displayed in Table 4.7 highlight that across all cultures there is a preference towards the full oxidation of CH_4 , as more of the available carbon is transformed to CO_2 than biomass (CDW). While this is true for both species, differences in the severity of this preference is observed between the species. *M. trichosporium* OB3b is shown to partition a much larger proportion of the available carbon to CO_2 relative to that partitioned to biomass when compared to *M. methanica* S1, especially when cultured using initial concentrations of 8% and 4% CH_4 v/v in air. When the initial concentration of CH_4 is reduced to 2% v/v in air, a decrease in the CPR is observed across cultures of *M. trichosporium* OB3b. Plotting the calculated CPR values against the initial concentration of CH_4 illustrates this (Figure 4.7). For *M. methanica* S1 no such trend is observed. Indeed, generating a linear fit to the point displayed in Figure 4.7 provides a p-value of the gradient coefficient to be > 0.05 . Interesting, this shift in carbon partitioning, observed in *M. trichosporium* OB3b at 2% v/v CH_4 in air, aligns with findings from Gómez-Borraz et al. (2025), where an anabolic phase within a methanotrophic community under the same concentration of 2% v/v CH_4 in air was observed. The effect this has on the bioenergetics is discussed subsequently.

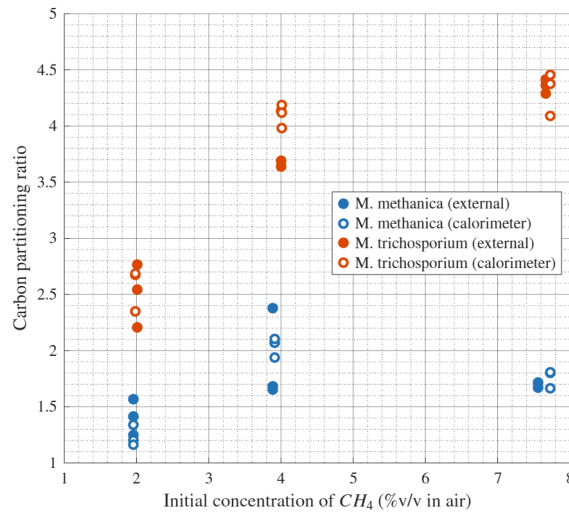


Figure 4.7: Carbon partitioning ratio (*CPR*) plotted against the initial concentration of CH_4 for *M. methanica* S1 and *M. trichosporium* OB3b.

4.4.3 Thermal output

The thermal output measured from batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b using the Calmetrix I-CAL HPC 4000 Isothermal Calorimeter are provided in Figure 4.8, for the three different concentrations of CH_4 investigated. In all cases, the power of heat dissipation increases exponentially to a peak and then drops-off sharply at a time that coincides with the total consumption of CH_4 (Figures 4.3a and 4.4a respectively). The integral of the total power of heat dissipation yields the total energy dissipated through time, shown in Figures 4.8c and 4.8d. In all cases, there is a correlation between the total energy dissipated and the initial concentration of CH_4 .

In the first instance we aim at quantifying the effectiveness of *M. methanica* S1 and *M. trichosporium* OB3b at producing heat. To do this, we firstly compare the total heat dissipated during the growth of these bacteria with the heat that could theoretically be accrued by oxidising all the CH_4 in air. Here we define this ratio as the thermal efficiency (η_{max}), which follows:

$$\eta_{max} = \frac{Q_{dis}}{Q_{max}} \times 100, \quad (4.3)$$

where, Q_{dis} refers to the total energy dissipated as heat during the growth of *M. methanica* S1 and *M. trichosporium* OB3b, and Q_{max} refers to the heat that could be produced by oxidising all the CH_4 available. Values of Q_{dis} can be taken directly from the data presented in Figure 4.8. To avoid errors associated with measurement noise, Q_{dis} is calculated by taking the mean of thermal energy values obtained after the thermal power reaches zero. The maximum available heat (Q_{max}) can be calculated from the enthalpy of combustion of CH_4 . As growth conditions are sufficiently close to standard conditions, the standard enthalpy of combustion is used (890.71 kJ/mol). Given that not all the available CH_4 is fully oxidised to CO_2 (Table 4.7), we also compare Q_{dis} to the maximum amount of heat that could theoretically be accrued from the catabolism undertaken by each culture Q_{cat} . This is different from Q_{max} as it takes into account how the available carbon is partitioned between biomass and CO_2 production. The thermal efficiency calculated from this quantity is denoted by η_{cat} . The method for calculating Q_{cat} is provided in Appendix L.

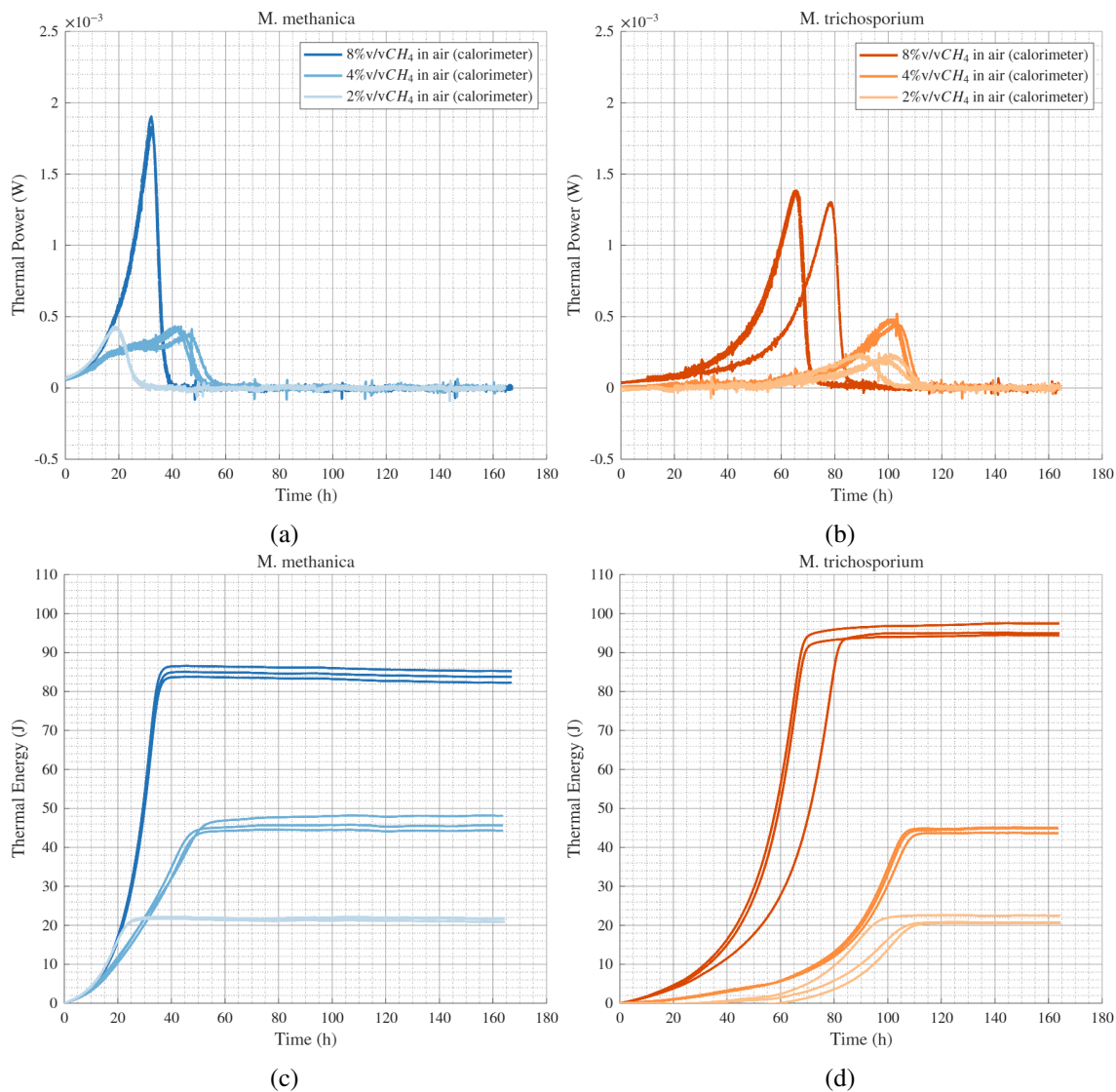


Figure 4.8: Thermal power and thermal energy dissipated during the batch growth of *M. methanica* S1 and *M. trichosporium* OB3b under three initial concentration of CH_4 (8%, 4% and 2% v/v in air). Thermal power dissipation was measured directly using a Calmetrix I-CAL 4000 HPC Isothermal Calorimeter. Total thermal energy was determine through integrating between measurments of thermal power and summing the results. a) Thermal power dissipation from *M. methanica* S1, b) thermal power dissipation from *M. trichosporium* OB3b, c) total thermal energy dissipation from *M. methanica* S1, d) total thermal energy dissipation from *M. trichosporium* OB3b.

The values of Q_{max} , Q_{cat} , Q_{dis} , η_{max} and η_{cat} for the different concentrations of CH_4 investigated are provided in Table L.1. The values displayed represent means from the biological triplicates displayed in Figure 4.8.

Table 4.6: Thermal efficiencies (η_{max} and η_{cat}) for batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b for different regimes of initial CH_4 concentration. The maximum available heat (Q_{max}), the maximum available heat from catabolism (Q_{cat}) and the total heat dissipated (Q_{dis}) are also provided.

Species	Experiment	Initial CH_4 (%v/v in air)		Q_{max} (J)		Q_{cat} (J)		Q_{dis} (J)		η_{max} (%)		η_{cat} (%)	
		mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
<i>M. methanica</i>	Calorimeter	7.73	0.14	147.98	2.01	109.32	3.06	84.43	1.43	57.06	0.20	77.19	1.91
		3.92	0.06	75.27	0.20	53.99	0.45	45.92	1.79	61.00	2.24	85.05	2.93
		1.96	0.04	37.40	1.23	23.96	0.82	21.66	0.39	57.94	1.46	90.49	4.26
<i>M. trichosporium</i>	Calorimeter	7.73	0.13	148.39	0.82	116.70	0.47	95.39	1.51	64.29	1.31	81.74	1.37
		4.01	0.04	76.66	0.20	55.67	0.61	44.46	0.70	57.99	0.84	79.86	0.98
		1.98	0.02	38.01	0.27	25.73	0.20	21.41	0.92	56.31	2.30	83.19	4.05

The results in Table L.1 demonstrate that both *M. methanica* S1 and *M. trichosporium* OB3b dissipate a significant proportion of the maximum available heat, with η_{max} calculated between 56 – 64%. When comparing the heat dissipated to that which could be accrued based on how the available carbon is partitioned (Q_{cat}), these efficiencies rise to between 77 – 90%, demonstrating that the vast majority of energy generated during the catabolic process is dissipated as heat by these bacteria. It is, however, important to note that given the observed reduction in carbon between initial and final sampling (Figure 4.6), it is likely that the CO_2 measured is an under-representation of the total amount of CH_4 fully oxidised and thus thermal efficiencies may be slightly lower than what is presented here. Additionally, it is known that in type-II methanotrophs up to 50 – 60% of cellular carbon can be obtained from CO_2 fixation via phosphoenolpyruvate (PEP) carboxylase (Hanson & Hanson, 1996; Kalyuzhanaya et al., 2013; Templeton et al., 2006), this process is not accounted for in our measurements of CO_2 .

From the measurements of heat dissipation we can also calculate the yield of heat per cell dry weight ($Y_{Q/CDW,dis}$) and the yield of heat per mole of CH_4 consumed ($Y_{Q/CH_4,dis}$). As measurements of OD_{600} and headspace composition can only be obtained from the calorimeter cultures at the end of the growth period, yields are calculated following:

$$Y_{Q/X} = \frac{\Delta Q}{\Delta X}, \quad (4.4)$$

where ΔQ is the heat dissipated over the entire growth period and ΔX is the change of a known quantity (CDW , CH_4) over the same period. These yields are calculated for each calorimeter culture individually with the final yields presented as the mean of these alongside the standard deviation. Within Table 4.7, these yields are presented alongside $Y_{Q/CH_4,cat}$ and $Y_{Q/CH_4,max}$, where $Y_{Q/CH_4,cat}$ is calculated by dividing Q_{cat} by the number of moles of CH_4 consumed, this is done solely using final samples. $Y_{Q/CH_4,max}$ is the standard enthalpy of CH_4 oxidation. $Y_{Q/CDW,dis}$ and $Y_{Q/CH_4,dis}$ can also be calculated by plotting the the change in thermal energy, measured from the calorimeter cultures, against the change in CDW or CH_4 , measured from the external cultures. This is done for both CDW and CH_4 as shown in Figures 4.9 and 4.10 respectively. As demonstrated linear fits passing through the origin can be made where the gradient refers to $Y_{Q/CDW,dis}$ and $Y_{Q/CH_4,dis}$ respectively. This reflects the same process as described by Birou et al. (1987). The values obtained by this

method are also provided in Table 4.7 where the uncertainty is based on the 95% confidence bounds of the fit. For 4.10 lines corresponding to $Y_{Q/CH_4,cat}$ and $Y_{Q/CH_4,max}$, calculated from external cultures, are also plotted to provide comparison.

Table 4.7: Yields of heat dissipated per cell dry weight ($Y_{Q/CDW,dis}$) and heat dissipated per mole of CH_4 consumed ($Y_{Q/CH_4,dis}$) for batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b for different regimes of initial CH_4 concentration. For calorimeter cultures, yields are calculated using the difference in quantities over the entire growth period, as expressed by Equation 4.4. For external cultures, yields are calculated by plotting quantities at similar time points against each other and calculating the gradient of a line fitted to this plot. Figures 4.9 and 4.10 demonstrate these plots. Maximum and catabolic yields of heat dissipated per mole of CH_4 are also provided, where $Y_{Q/CH_4,max}$ refers to the standard enthalpy the oxidation of CH_4 and $Y_{Q/CH_4,cat}$ is calculated by dividing the maximum heat generated through catabolism by the number of moles of CH_4 consumed. This is done using final samples from cultures grown in the calorimeter and externally.

Species	Experiment	Initial CH_4 (%v/v in air)		$Y_{Q/CDW,dis}$ (kJ/g)		$Y_{Q/CH_4,dis}$ (kJ/mol)		$Y_{Q/CH_4,cat}$ (kJ/mol)		$Y_{Q/CH_4,max}$ (kJ/mol)
		mean	SD	mean	SD	mean	SD	mean	SD	(-)
<i>M. methanica</i>	Calorimeter	7.73	0.14	59.46	0.21	508.21	1.75	658.63	16.25	890.71
		3.92	0.06	78.27	4.19	543.38	19.91	638.86	5.45	890.71
		1.96	0.04	57.66	1.94	516.12	12.99	570.96	23.64	890.71
<i>M. trichosporium</i>	Calorimeter	7.73	0.13	145.66	8.43	572.61	11.68	700.50	2.73	890.71
		4.01	0.04	145.89	1.94	516.56	7.48	646.82	5.42	890.71
		1.98	0.02	103.21	11.89	501.53	20.45	603.03	4.69	890.71
		mean	SD	gradient	95% bounds	gradient	95% bounds	mean	SD	(-)
<i>M. methanica</i>	External	7.73	0.14	57.96	±1.05	518.83	±7.69	653.70	4.72	890.71
		3.92	0.06	62.10	±4.66	521.48	±23.83	639.99	26.77	890.71
		1.96	0.04	55.50	±1.41	516.80	±8.93	598.43	15.62	890.71
<i>M. trichosporium</i>	External	7.73	0.13	134.46	±10.21	575.96	±35.69	723.91	8.39	890.71
		4.01	0.04	104.15	±14.29	480.99	±52.84	690.25	13.11	890.71
		1.98	0.02	77.32	±12.24	455.46	±62.13	640.98	42.52	890.71

The yields of heat dissipation provided in Table 4.7 demonstrates significant heat dissipation per cell dry weight from methanotrophic cultures of *M. methanica* S1 and *M. trichosporium* OB3b. The yields observed here are much greater than those measured for other bacteria, such as *E.coli*, which provide typical values of 11 – 15kJ/g when utilising substrates such as hexoses (Birou et al., 1987). Indeed, it has been observed that growth utilising carbon substrates that are more reduced resulted in greater yields of heat per biomass formed (Birou et al., 1987; C. Wang & Kuzyakov, 2023). Methane, being the most reduced form of carbon, will therefore facilitate much higher yields of heat per cell dry weight as demonstrated.

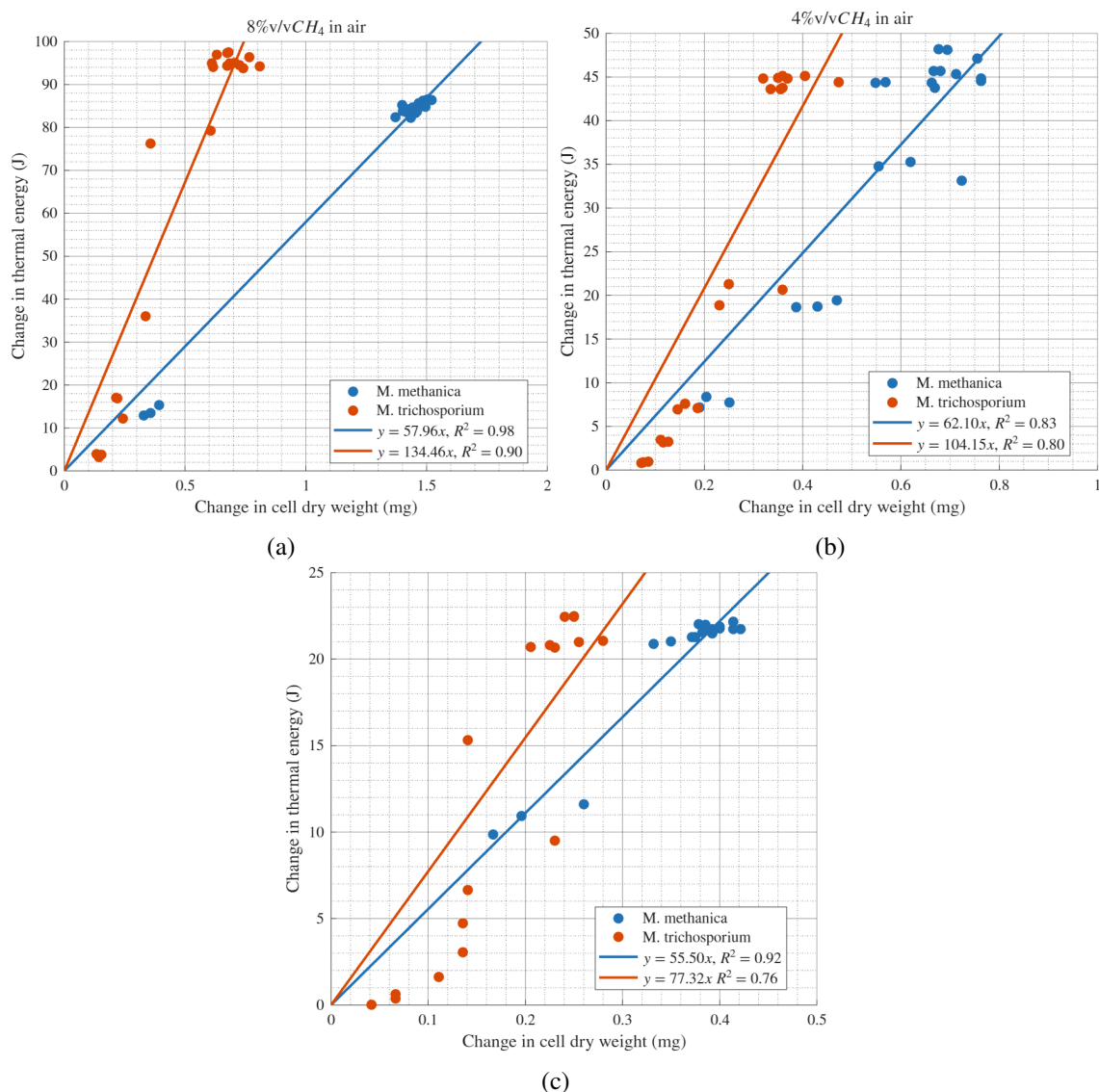


Figure 4.9: Change in thermal energy dissipated against the change in cell dry weight (CDW) for batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b under different initial concentrations of CH₄. Heat dissipation is measured from cultures within the Calmetrix I-Cal 4000 HPC Isothermal Calorimeter. Values of CDW are produced from OD₆₀₀ measurements from parallel external cultures. Values plotted against each other refer to the same time point from the experiment start time. Linear fits are produced using all points plotted for a single species. a) 8%v/vCH₄ in air, b) 4%v/vCH₄ in air, c) 2%v/vCH₄ in air.

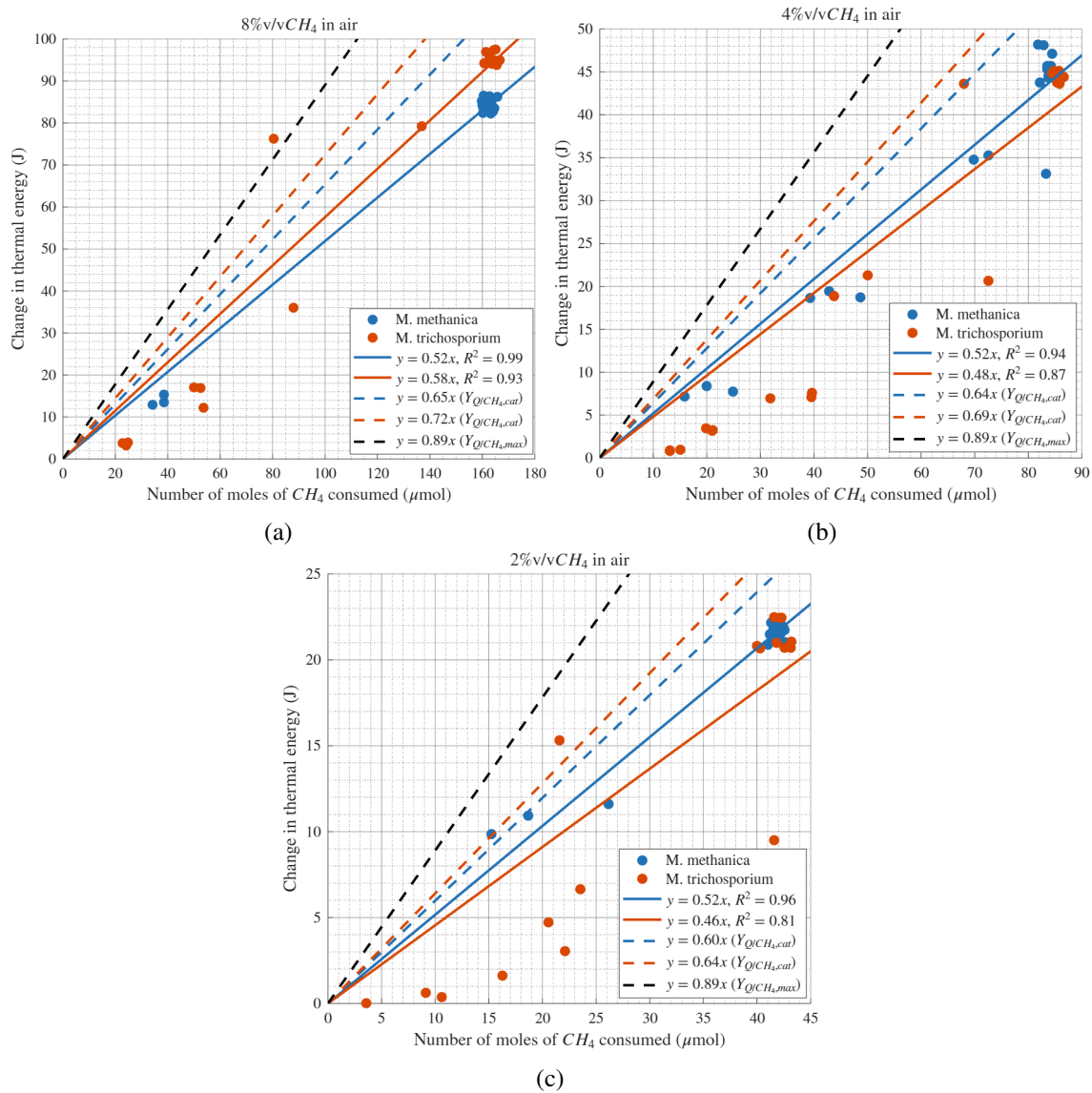


Figure 4.10: Change in thermal energy dissipated against the change in moles of CH_4 for batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b under different initial concentrations of CH_4 . Heat dissipation is measured from cultures within the Calmetrix I-Cal 4000 HPC Isothermal Calorimeter. The number of moles of CH_4 are produced from analysis of headspace composition of parallel external cultures. Values plotted against each other refer to the same time point from the experiment start time. Linear fits are produced using all points plotted for a single species. Lines corresponding to $Y_{Q/CDW,cat}$ calculated from external cultures, and $Y_{Q/CDW,max}$ are also plotted for both *M. methanica* S1 and *M. trichosporium* OB3b to provide comparison. a) 8%v/v CH_4 in air, b) 4%v/v CH_4 in air, c) 2%v/v CH_4 in air.

4.4.4 Relationship between heat dissipation and growth

For microorganisms, it has been argued that growth is driven by a thermodynamic driving force generated by the difference in the Gibbs energy change of the energy-yielding catabolic reaction (ΔG_{cat}) and the energy-consuming anabolic reaction (ΔG_{ana}) (von Stockar & Liu, 1999). This viewpoint can be summarised by the flux-force relationship developed in non-equilibrium thermodynamics, where the proportionality between flux and the conjugate force is linear:

$$J_i = L_{ii}Z_i + \sum_{j \neq i} L_{ij}Z_j, \quad (4.5)$$

where, L_{ii} is the coefficient that links the flux (J_i) to 'its' force (Z_i) and L_{ij} links the flux (J_i) to all other forces (Z_j). Assessing the coupling of only two reactions, it is shown that the rate of each individual reaction can therefore be expressed as:

$$r_a = L_{aa}(-\Delta G_a) + L_{ab}(-\Delta G_b) \quad (4.6a)$$

$$r_b = L_{bb}(-\Delta G_b) + L_{ba}(-\Delta G_a), \quad (4.6b)$$

where, r_a and r_b are the rate of reactions a and b respectively, ΔG_a and ΔG_b are the Gibbs free energy of each reaction and L_{aa} , L_{bb} , L_{ab} and L_{ba} are coefficients (Stockar & Wielen, 2013b). Given that ΔG_a and ΔG_b are both negative and assuming the cross coefficients are equal and also significantly negative, $L_{ab} = L_{ba} = -ve$, then reaction b can be driven "uphill" ($r_b < 0$) if $\|\Delta G_a\| > \|\Delta G_b\|$. This is the case in the coupling of catabolic and anabolic reactions and thus the magnitude of the rate of both of these reactions is proportional to the difference in Gibbs free energy between both reactions. This is commonly referred to as the "driving force" of the coupled reaction (Heijnen & Van Dijken, 1992; Stockar & Wielen, 2013b; von Stockar & Liu, 1999).

For two species utilising similar metabolic pathways, such that the magnitude of energy produced through the catabolic reaction and energy consumed during the anabolic pathway are consistent across species, then a trade-off is observed driven by the biomass yield. At high yield, thus efficient growth, ΔG_{ana} tends to ΔG_{cat} reducing the thermodynamic driving force and thus the rate of reactions. At low yield, the difference between ΔG_{ana} and ΔG_{cat} is maximised, driving an increased reaction rate at the cost of reduced biomass yield (Stockar & Wielen, 2013a; von Stockar et al., 2008). The growth of *M. methanica* S1 and *M. trichosporium* OB3b demonstrate this trade-off. Here we observe that batch cultures of *M. methanica* S1 produce a greater yield of biomass while producing a reduced yield of CO_2 , when compared to cultures of *M. trichosporium* OB3b grown under the same conditions. While exact yields may vary from those previously recorded, this difference in growth is in keeping with literature (Anthony, 1978; Templeton et al., 2006; Van Dijken & Harder, 1975). These differences in growth efficiency, with *M. methanica* S1 displaying more efficient growth and *M. trichosporium* displaying less efficient growth, is directly reflected in the measurements of heat dissipation made. *M. trichosporium* OB3b is observed to dissipate a larger quantity of heat per cell dry weight than that of *M. methanica* S1. Plotting the yield of heat per cell dry weight ($Y_{Q/CDW,dis}$) against the carbon partitioning ratio (CPR) results in a linear relationship, as highlighted in Figure 4.11. This mirrors the results by Birou et al. (1987), where the yield of heat per biomass is shown to be dependent on biomass yield as well as the oxidation state of the carbon substrate. Indeed the overall high values obtained for $Y_{Q/CDW,dis}$ from our methanotrophic cultures are a consequence of CH_4 occupying the

lowest possible oxidation state for carbon, while the differences between species utilising the same substrate being driven by differences in how carbon is partitioned.

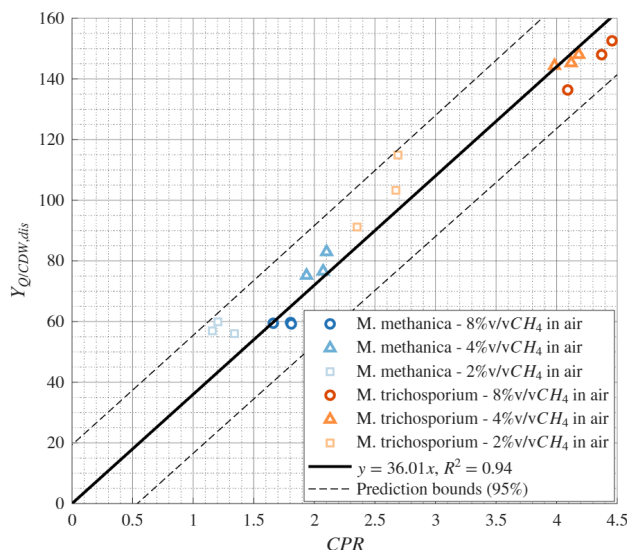


Figure 4.11: Yield of heat per cell dry weight ($Y_{Q/CDW,dis}$) plotted against the carbon partitioning ratio (CPR) for batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b grown under different concentration of CH_4 in air. A linear relationship is observed across all points with the line $y = 36.01x$ providing a fit with $R^2 = 0.94$.

Based on the relationship observed within Figure 4.11, two theoretical extremes can be characterised. At one end, where the CPR tends to 0, biomass yield is maximised providing a perfectly efficient system where all available energy is contained within the biomass produced and thus no heat is dissipated. However, under such a regime the rate of reaction would be infinitely slow (Kedem & Caplan, 1965). At the other, where CPR tends to a maximum, all energy is dissipated as heat and thus reaction rates are maximised but no biomass is produced. Under both these conditions growth is inhibited. Thus, at some point between these extremes an optimal point must exist where the growth rate is maximised. A fuller account of this argument is presented by Stockar and Wielen (2013b). According to our results we can assume that this point must lie closer to the points plotted for *M. methanica* S1 as for all experiments undertaken, batch cultures of *M. methanica* S1 express a higher growth rate than cultures of *M. trichosporium* OB3b, however, this is complicated by the lack of initial CO_2 , known to impact the growth of *M. trichosporium* (Kalyuzhanaya et al., 2013; S. Park et al., 1991).

Literature demonstrates that under high methane and high oxygen conditions growth of type-II methanotrophs is favoured over type-I (J. Amaral et al., 1995; J. A. Amaral & Knowles, 1995; Henckel et al., 2000). This is at odds with what is observed here. It is also expected and observed that the biomass yield of methanotrophs utilising the serine cycle (type-II) is greater than we recorded (Anthony, 1978; Templeton et al., 2006; Van Dijken & Harder, 1975). Accepting that this is the case, we may expect that under optimal conditions the points plotted for *M. trichosporium* OB3b within Figure 4.11, at higher concentrations of CH_4 (8% and 4% v/v), would sit much closer but still above those of *M. methanica* S1. If, as observed in literature, type-II methanotrophs enjoy a competitive advantage here then, according to the argument laid out by Stockar and Wielen (2013b), one would expect this location to better represent the optimal trade-off between the thermodynamic "driving force" and biomass yield, or in our case the CPR .

Under low CH_4 concentrations the opposite is observed and type-I methanotrophs outcompete type-II (J. Amaral et al., 1995; J. A. Amaral & Knowles, 1995; Henckel et al., 2000). A study by Graham et al. highlights that when the concentration of CH_4 in the feed gas is reduced from 3.4% v/v to 1.7% v/v, the ratio of the number of cells of *M. trichosporium* (type-II) to *M. albus* (type-I) declined (D. W. Graham et al., 1993). Within our experiments, we observe a shift in the partitioning of carbon towards greater biomass production within cultures of *M. trichosporium* OB3b when grown using an initial concentration of 2%v/v CH_4 . Carrying on from our previous hypothesis, a case could emerge that as the concentration of CH_4 is reduced the carbon partitioning ratio (*CPR*) for *M. trichosporium* OB3b is dropped to below that of *M. methanica* S1. Thus, if the optimal *CPR* lies above the values plotted for *M. methanica* S1 on Figure 4.11, then a *CPR* value below these would invoke a competitive disadvantage. This point would prove a useful indicator for confirming or disproving this hypothesis. Indeed, further experiments are needed to confirm the speculation made.

This thermodynamic trade-off between biomass yield or carbon partitioning and the "driving force" may be too simplistic to explain the complications of growth under all scenarios. In the previous hypothesis we assume that coefficients L_{ab} and L_{ba} are equal and constant across species. This is arguably not an accurate assumption. Indeed, the difficulty in predicting growth yields from energy yields has long been attributed to "uncoupled growth" or wastage of ATP (Leak & Dalton, 1986b). Under a starvation period, it is likely that the cross coefficients (L_{ab} and L_{ba}) would be affected if a physical or biochemical processes responsible for coupling catabolic and anabolic reactions is inhibited, or the energy yielded through the catabolic reaction is being used for other means and is therefore wasted in regard to biomass production. Indeed, assuming that the catabolic reaction is only coupled to biomass production is a major simplification. In actual fact a much more complex web will exist (Chubukov et al., 2014; Gerosa & Sauer, 2011; Locasale & Cantley, 2011). This uncoupling of growth will act to upset the previously argued trade-off.

It is also important to note that physical processes influence the growth rate of organisms. Under the methodology developed agitation or mixing of cultures is impossible. Cultures will therefore sediment, however, the rate and processes associated with this can differ between species. Figure 4.1 demonstrates differences in the spatial structure and sedimentation of both *M. methanica* S1 and *M. trichosporium* OB3b. This will effect the diffusion of CH_4 to the cells and thus the rate at which substrate can be consumed. These processes are completely ignored within the thermodynamic explanation.

It is most likely that a complex combination of the processes suggested as well as others unaccounted for explain the differences in growth rates observed here between *M. methanica* S1 and *M. trichosporium* OB3b. Here we only speculate on these processes, indeed a much more rigorous analysis, perhaps utilising in-depth metabolic models, is required to unlock the exact nature of the interactions between these processes. What the results here do highlight is that a trade-off between how carbon is partitioned, between energy yielding and energy consuming reactions, and the heat dissipated per biomass is apparent. However, the relative importance this trade-off has on the overall growth rate, especially under conditions that invoke non-uniform starvation between species, is difficult to uncover. Indeed, in most cases theory is developed based on model systems where we assume favourable growth conditions and spatial homogeneity. For most biotechnologies this is not the case. For example, centralised and decentralised wastewater treatment systems represents a myriad of conditions, with growth environments varying both spatially and temporally. Trying to define the limitations and importance of theory under such conditions is thus an equally important endeavour.

4.5 Conclusion and future perspectives

We have recorded that for batch cultures of both *M. methanica* S1 and *M. trichosporium* OB3b, 56 – 64% of the energy that would be released as heat during complete combustion of CH_4 is evolved as heat during the growth of these species. This accounts for 77 – 90% of the energy generated through catabolism. For *M. methanica* S1 yields of heat per gram of cell dry weight of between 56 – 78 kJ/mol are calculated and for *M. trichosporium* OB3b these are calculated between 77 – 146 kJ/mol, where the exact value is dependent on the initial concentration of CH_4 as well as the method of calculation used. When compared to species grown using less reduced forms of carbon, the results presented here demonstrate the greatest magnitude of heat dissipation from bacterial growth. Indeed, the values presented here are in excess of previously recorded maximum yields of heat per biomass of 25 kJ/g (Birou et al., 1987).

The difference in yields of heat per cell dry weight obtained within this study is directly linked with how each species partitions the carbon within the CH_4 introduced. Our cultures of *M. trichosporium* OB3b are shown to convert as little as 14% of the carbon included in the CH_4 to cell dry weight, while for *M. methanica* S1 this proportion is higher at 27 – 35%. The inverse is observed for yields of CO_2 , where within cultures of *M. trichosporium* OB3b CO_2 accounts for between 49 – 66% of the carbon stored initially as CH_4 , this percentage is reduced to between 43 – 59% for *M. methanica* S1. Plotting a ratio of how carbon is partitioned between CDW and CO_2 against yields of heat per cell dry weight demonstrates a linear relationship. Thus, it is both the yield of biomass and the oxidation state of the substrate that drive heat dissipation within bacteria.

By demonstrating the vast amount of heat dissipated by methanotrophic bacteria, we have laid a quantitative foundation for exploiting this heat as a resource, thus aiding the development of methane mitigation technologies that may recycle this heat. The design of bioreactors capable of harvesting the metabolic heat produced by methanotrophs will encounter challenges. These include: the trade-off between heat extracted and temperature shifts that will effect bacterial activity, insulation to ensure heat is not dissipated to the environment, spatial gradients in methane concentration and maintenance of optimal conditions. As a result the exact impact these challenges provide will need to be understood and solutions for overcoming them will need to be developed. The research here does however suggest that, in principal, the biological oxidation of CH_4 could be harnessed for useful heat production and rising to these research challenges would not be futile.

It is also apparent that although the link between heat dissipation and biomass yield is demonstrated here, further research is required to determine the importance this trade-off has on determining growth rates. This is especially true in scenarios, such as presented here, where cultures are grown outside of optimal conditions and thus thermodynamic explanations of growth are muddled by physical and biochemical processes unaccounted for within these descriptions. Indeed such conditions are prevalent throughout biotechnologies and therefore expanding thermodynamic theory in relation to microbial growth to account for such conditions would be of great engineering importance.

Chapter 5

Conclusion and future perspectives

Within this thesis two distinct pieces of research have been presented. The first, a comprehensive investigation of the compatibility of neutral models with single-species relative abundance time series, and the second a thermodynamic analysis of methanotrophic bacteria. Although each have merit of their own accord, here we discuss the broader implications of the research and whether it provides insight into the acquisition of general theory within wastewater treatment. To support the arguments made, motivations behind key decisions are reaffirmed, and in an attempt to address foreseen critiques, shortcomings of the research are discussed retrospectively. Avenues for further research are also made clear.

The first research chapter entitled *"Identifying neutral models on the basis of species abundance time series"* was aimed with deriving the underlying principles that govern microbial community structure from observable patterns and trends using ecological theory. Here we settled on exploring the neutral theory developed by Hubbell (2001) and an adaptation of this theory developed by Sloan et al. (2006). The selection of these theories was a deliberate decision, with the motivation behind this choice based on the logic provided by Curtis, Ofiteru, Sloan, and Woodcock, among others. Woodcock et al. (2007) break down the selection of appropriate models for developing a generalised theory into two defining criteria: validity and predictive capability. Regarding these, Harte (2004) states *"Theories are of most interest when the ratio of the number of predictions that they make to the number of assumptions and adjustable parameters in the theory is large"*. Harte directly provides Hubbell's purely-neutral model as an example of this (Harte, 2004). Ofiteru et al. (2010) states, *"the emergent parameter-rich models defy calibration for very diverse microbial communities"*, highlighting the difficulty in selecting more complex theories for calibration using data obtained from microbial communities. Indeed, our results demonstrate that even the simplest of neutral models are over-parametrised to derive all distinct parameters under a naive scenario when using single species relative abundance time series. Given the extreme diversity of microbial communities central to wastewater treatment technologies (T. P. Curtis et al., 2002), the selection of simpler models as a starting point towards a generalised theory provides a pragmatic and scientifically sound approach. Indeed, to assess systems at a broader scale, one must accept and strive for a loss of detailed intricacies (Levin, 1992). The fact that the most mathematically simple community models, Hubbell's purely neutral model and Sloan et al.'s near-neutral model, have been shown to accurately describe and predict many patterns and trends observed from complex communities, makes the parameters within these models inherently important (Hubbell, 2001; Latimer et al., 2005; Ofiteru et al., 2010; Rosindell et al., 2011; Sloan et al., 2006, 2021).

From this point the choice of using simulated time series to calibrate the models selected was made. One may argue that a more compelling analysis could have been provided using 'real' time series, obtained from highly controlled experiments in which all model parameters could be measured externally as well as derived through model calibration. However, with no indication of appropriate sampling protocols, experimental length and sampling frequencies, sampling regimes used to obtain data sets would have been difficult to determine prior to experimentation. Generating time series of similar lengths and temporal resolutions to the ones employed within our study would not have been possible experimentally using current technologies. Indeed, experimentation is a much more time intensive approach for producing large datasets than simulation. Given that time has been identified as our most valuable resource (T. Curtis & Fowler, 2023), a choice was made to adopt the strategy articulated by Martinez-Rabert et al. (2023), where computational modelling is used to develop hypotheses through a broader analysis, and where experimentation is only implemented subsequently as a tool for focused validation. It is important to make clear that this experimental validation has not yet been undertaken for the conclusions attained through the research presented in Chapter 3.

This lack of experimental validation presents an opportunity for further research. Sloan et al. (2021) demonstrate a methodology that allows for the controlled culture of microbial communities composed of two species. As a result the processes assumed to govern neutral community structuring, community size, replacement rate, immigration and selection can be determined and controlled externally to model calibration. The effects of sampling regimes on parameter estimates could be validated by producing time series of varying lengths and temporal resolutions and subsequently using these time series to calibrated Hubbell's and Sloan et al.'s neutral models. Calibration can be achieved using the computational framework developed within this thesis, which is available at <https://github.com/tymonherzyk/neutral-modelling-framework>. For validating the increase in parameter estimation accuracy associated with capturing time series over a perturbation to the species, a process for artificial inducing a perturbation on the controlled communities would need to be developed. Similar to how Sloan et al. (2021) artificially induced competitive advantages, antibiotics could be introduced to the system to reduce the relative abundance of a non-resistant species. Indeed, the methodology developed by Sloan et al. (2021) demonstrates a promising avenue for validating the results obtained within Chapter 3 through experimental means.

The results presented in Chapter 3 indicate that for all but the most controlled systems, direct quantification of the distinct processes that drive microbial community structure is unachievable through calibrating neutral models using single-species relative abundance time series. The parameter estimates obtained are shown to be influenced by additional mechanisms unaccounted for in these models and thus it must be assumed that in most cases model parameters reflect a multitude of complex dependencies and are removed from their attributed distinct mechanistic meanings. This conclusion is visualised in Figure 5.1, where direct links between model parameters and associated phenomena are maintained only when 'idealised' time series are used for calibration. Here we defined 'idealised' time series as ones produced from communities only governed by the processes described by the model calibrated. For 'real' communities, governed by more complex drivers, these links are distorted. Thus, even under suitable sampling protocols and *a priori* knowledge of the system only effective parameters can be accurately estimated. This concession may fall short of the initial aim set (Table 1.1), however, the means by which this verdict was reached does indicate fulfilment of the objectives set out.

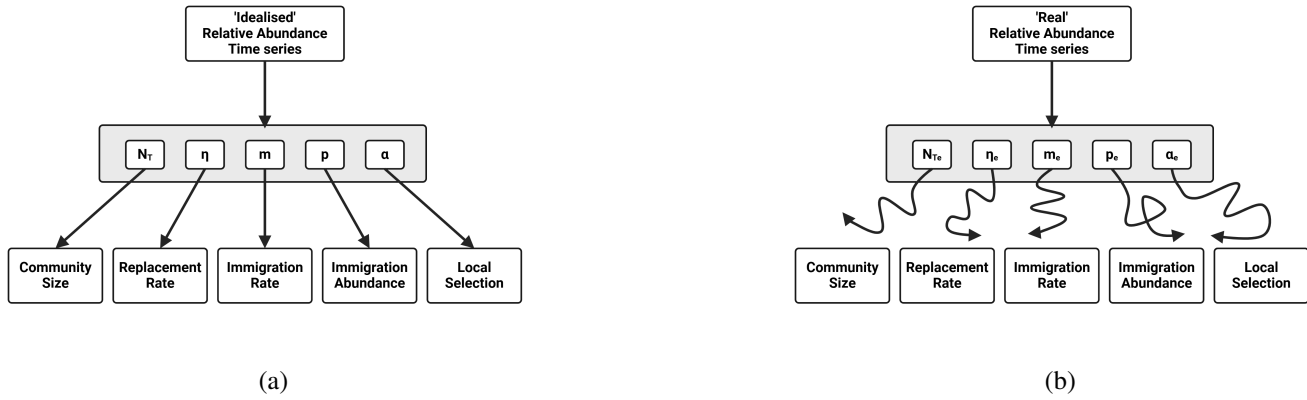


Figure 5.1: Diagram representing the link between neutral model parameters and their attributed mechanistic meaning when using 'idealised' and 'real' relative abundance time series. Here we define 'idealised' time series as those produced from communities only governed by the processes described by the model calibrated, whereas 'real' time series refer to those produced from communities where the structuring dynamics differ to those described by the model calibrated. a) idealised time series, b) real time series.

To access the progress this affords towards a general theory it is pertinent to draw focus to the role of phenomenological parameters within engineering contexts. General theory typically has the ability to offer insight into complex systems despite uncertainty (Sloan & Gómez-Borraz, 2023). To achieve this, phenomenological parameters are often used to summarise intricate interdependencies between phenomena. For example, the Reynolds Number is used in aeronautical engineering to predict flow regimes and inform aircraft design, the Gauckler–Manning coefficient is used by hydraulic engineers to portray the roughness of pipes or channels and Young's modulus is used to define the tensile and compressive stiffness of materials. Even in fields more closely related to community ecology, such as population genetics, effective parameters like the "effective population size" are harnessed to quantify genetic diversity and inform conservation strategies (J. Wang et al., 2016). Quantification of these parameters is achieved through empirical studies or simulation, as deriving them from first principles is extremely difficult.

Our results demonstrate that a similar method of categorisation can be used for microbial communities where neutral models are expected to provide a good fit. If these time series are produced through conforming to favourable sampling regimes, as those demonstrated in Chapter 3, then the parameter quantities estimated will still provide an indication of abundance dynamics. For example, if for a species of interest the effective immigration rate (m_e) is estimated at a high value, then it is likely the species will remain within the community and recover more quickly from perturbations to the system. On the other hand if this value is low, larger drift within the species abundance can be expected. Other parameters such as the effective community size N_{Te} and effective local selection α_e will provide different characteristics and can similarly be used to give indication of future dynamics. Therefore, generating high-quality relative abundance time series from biotechnologies, such as wastewater treatment plants or smaller decentralised wastewater systems, can provide useful data for predicting species abundance dynamics using the phenomenological or effective parameters expressed here. Indeed, the argument for utilising effective parameters within these models has been made previously (Sloan et al., 2021).

Where the effective parameters presented in our research differ from the engineering examples previously provided, is in a lack of a consensus regarding the underlying processes that impact them. These processes have only been postulated and not yet definitively proved (Sloan et al., 2021). For the Reynolds Number, Gauckler–Manning coefficient and Young’s modulus, knowledge of the factors that drive these values has been established and thus can be leveraged to inform engineering design. In this aspect our effective parameters are lacking. Establishing these processes can be undertaken using two approaches, as shown in Figure 5.2. The first is a purely empirical approach. Similarly to the analysis conducted in Chapter 3, time series can be produced through simulations or sampled from multiple communities in which the processes driving their structure are known and can be controlled. By calibrating models to time series from these differing communities and observing the shift in parameter estimates, the relative importance of these processes can be established. To be confident that any observed relationships between effective parameters and the tested phenomena are true, we must first be assured that the conditions under which these models are calibrated can produce reliable results. These conditions have been defined by the research presented in Chapter 3. The second approach is to try and derive parameters from first principles. This is a similar process to that undertaken in Chapter 4 and is arguably the more difficult of the two approaches. Both of these approaches have been utilised in predicting and determining drivers of the effective population size within population genetics (Charlesworth, 2009; Nunziata & Weisrock, 2018; J. Wang et al., 2016).

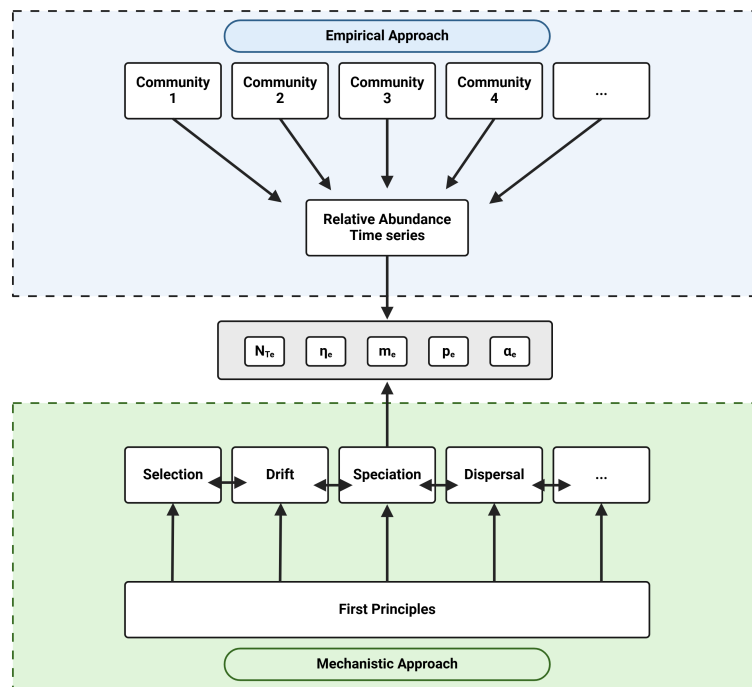


Figure 5.2: Approaches for determining the phenomena that drive effective neutral model parameters. One is an empirical approach where time series from differing communities are obtained and model parameters are calibrated to observe differences in the values obtained. The second is a mechanistic approach that looks to derive parameters from first principles.

Here we promote the exploration of both these approaches for undertaking further research. Indeed, despite the difficulty of providing relationships from first principles, developing theoretical descriptions, even if they are highly simplified, that can be used in predictive equations for determining effective parameters can be extremely useful for informing decision making and courses of action. For example, regarding the effective populations size (N_e) within population genetics, many predictive equations for this single parameter have been developed (J. Wang et al., 2016). These mathematical relationships provide a useful theoretical backbone for informing conservation practices. Developing such an abundance of mathematical theory for effective parameters within ecology, such as those presented in this thesis, is long overdue and when paired in tandem with methods for estimating these parameters from community data will provide theoretical context for the behaviour of these communities.

In the meantime, providing insight through empirical means may prove effective. For example, in calibrating Hubbell's purely neutral model from a community governed by the dynamics expressed in Sloan's near-neutral model, we were able to identify that the inclusion of a local advantage can drive systematic errors in certain parameters, leading to the conclusion that these should be considered as effective parameters. The isolation of a single structuring process, in this case selection, was made relatively easy through computational simulation of contrasting models. A similar process could be implemented to quantify the effects that other known drivers have on the estimation of neutral model parameters. To achieve this one would need to develop a system for simulating multiple relative abundance time series from different theoretical interpretations of community structure. As a result of the design principles adopted within the development of the computational framework provided in this thesis, the inclusion of additional models into the simulation package is an uncomplicated process. Given the compatibility of subsequent packages within the computational framework developed, calibrating effective versions of Hubbell's and Sloan et al.'s models to newly simulated time series would then be a straightforward endeavour. By broadly exploring how different interpretations of community structuring affect the calibration of effective neutral parameters, a picture can begin to be built into the processes that drive the estimated values of these parameters. Experimental validation is essential before drawing definitive conclusions, however, through this process of computational probing one could provide areas of focus for experimental studies.

The conclusions reached within Chapter 3, paired with the discussion presented here, indicates, that novel progress has been made towards the overall aim of the thesis based on the research undertaken regarding neutral models. Although there are bound to be further critiques, unaccounted for within the discussions provided here, and additional research could have been undertaken to strengthen the arguments made, we believe that the research provides a fundamental step necessary for deriving the effective parameters proposed using empirical data. This, in turn, will facilitate studies aimed at observing relationships between these parameters and the underlying processes that drive them. Only by establishing these relationships, through empirical observations or by developing derivations for them from first principles, can a generalised theory utilising these parameters take shape.

Now consider Chapter 4 entitled "*Quantifying the heating potential of methanotrophs for application within decentralised wastewater treatment technologies*". The purpose of this study was to offer a different approach to that provided by the aforementioned research. While the first study aimed at deriving the processes governing community structure empirically, which inevitably ended up providing parameters void of specific mechanistic meaning. The goal of the second research chapter was to build a mechanistic description of these processes from first principles. This is pursued through an attempt to provide evidence that relationships between growth kinetics and heat dissipation exist. The choice was

made to achieve this for methanotrophic bacteria. The selection of methanotrophs was made for multiple reasons. Firstly, it has been shown that heat dissipation from bacteria depends on the degree of reduction of the substrate (Birou et al., 1987). Therefore, since the carbon in methane has the lowest possible oxidation state, growth on methane should result in the highest possible yield of heat. Indeed, the values measured for the yield of heat per biomass within Chapter 4 are higher than any of those observed from less reduced substrates (Birou et al., 1987). Additionally, methanotrophs are of engineering interest due to their unique ability to oxidised methane under atmospheric conditions, facilitating reductions in green house gas emissions from biotechnologies (Ruiz-Ruiz et al., 2020; Strong et al., 2015). Thus, reporting the heat dissipated from these bacteria has great potential for applications in biotechnological and environmental engineering contexts.

To measure heat dissipation a methodology was developed utilising a Calmetrix I-CAL 4000 HPC Isothermal Calorimeter. The growth of methanotrophs presented a unique challenge with this equipment due to the need to maintain a gas-tight growth system. As a solution, custom culture bottles were designed and produced as part of the research undertaken. The design of these bottles is provided in Appendix F. While heat dissipation from bacterial growth has been previously measured *in situ* (Birou et al., 1987; J.-S. Liu et al., 2001; Schill & von Stockar, 1995; von Stockar & Birou, 1989), our methodology is novel in its application to methanotrophic bacteria. The results in Chapter 4 demonstrate the success of this methodology with the heat dissipation curves tracing out patterns that make biological and physical sense and display similar relationships to those observed elsewhere (Birou et al., 1987; J.-S. Liu et al., 1999, 2001).

The overall methodology could have been strengthened if growth and energy dissipation were measured simultaneously. As stated within Chapter 4, growth was measured in external cultures; biomass measurement within the calorimeter was limited to solely start and end times. While growth of external cultures and those within the calorimeter were shown to reach the same final states, comparing growth between these two sets of cultures must be handled with care. Another issue was in the temporal resolution used when measuring these parameters. Phases of exponential growth were too short and thus undersampled such that the clear determination of rates was difficult. This issue was exacerbated by having to grow our methanotrophs using batch cultures and thus the low volumes of the culture bottles used determined a limit to the methane, and therefore carbon available. Differences in growth kinetics between *M. methanica* S1 and *M. trichosporium* OB3b as a result of low initial CO_2 , as demonstrated by S. Park et al. (1991), also made determining a suitable yet consistent sampling regime more difficult. Indeed, repeating the experiments using a CO_2 rich headspace would provide a useful comparison for highlighting the effect of available CO_2 on growth rate and heat dissipation, especially for type-II methanotrophs. The inability to robustly observe steady-state growth rates and substrate consumption rates made assessing correlations with these rates and quantity or rate of heat dissipation difficult to achieve.

A higher resolution break down of the distribution of carbon in the system would have also been beneficial. A mismatch in the final amount of carbon calculated versus the initial carbon content is demonstrated. Here we believe the mismatch may be due to physical and biological processes unaccounted for within the methodology employed. During headspace sampling and measurements of pressure, perforations made into rubber stoppers are likely to result in the release of gas, reducing the pressure and thus the carbon content calculated. Some gas may also be lost within the internal volume of the manometer system used for measuring pressure. The methodology also fails to account for any build up of inorganic carbon or volatile organic carbon compounds that may have been excreted during growth. Indeed, excretion of carbon containing metabolites by methanotrophic bacteria has been observed via multiple pathways (Kalyuzhnaya et al.,

2013; Kalyuzhnaya et al., 2015; Sahoo et al., 2021, 2023). Although the mismatch in carbon balance may throw doubt on the exact quantities presented, it is important to note that based on the consistency of the methodology these errors would be uniform across experiments and thus conclusions based on the comparison of results remain sound. To improve on the methodology a fuller analysis should be undertaken, one that accounts for all possible forms of carbon and minimises unintentional losses.

Despite possible improvements to the methodology the results do express differences in the growth of *M. trichosporium* OB3b and *M. methanica* S1 which are in-keeping with literature, where cultures of *M. trichosporium* OB3b are expected to produce a higher yield of CO_2 than those of *M. methanica* S1 while producing a reduced yield of biomass (Anthony, 1978; Templeton et al., 2006; Van Dijken & Harder, 1975). Our results demonstrate that this trade-off in how carbon is partitioned is directly related with the heat dissipated per biomass, with this value measured much higher for cultures of *M. trichosporium* OB3b. According to thermodynamic theory, as summarised by Stockar and Wielen (2013b), one would expect the increased quantity of heat dissipation per biomass to drive the rate of catabolic and anabolic reactions, and thus, in regards to growth rate, a trade-off is observed between the heat dissipation and biomass yield, where an optimal ratio exists. As under all growth experiments we observe greater rates of growth for *M. methanica* S1 than *M. trichosporium* OB3b we would assume that *M. methanica* S1 would outcompete *M. trichosporium* OB3b under all growth conditions investigated and thus the yields for *M. methanica* S1 must lie closer to this optimal point.

Here it is however important to note that this is not observed under normal growth conditions. While under low concentrations of methane *M. methanica* S1 (type-I methanotrophs) are shown to outcompete *M. trichosporium* OB3b (type-II methanotrophs), at higher concentrations of methane they are observed to co-exist or type-II methanotrophs dominate (J. Amaral et al., 1995; J. A. Amaral & Knowles, 1995; D. W. Graham et al., 1993; Henckel et al., 2000). The primary difference between our growth conditions and those used within the previously cited literature is the initial concentration of CO_2 . Without further experiments, measuring growth yields and heat dissipation under conditions with higher initial concentrations of CO_2 it is unknown whether the ratio of yields observed for *M. methanica* S1 are indeed more optimal than those of *M. trichosporium* OB3b, or whether the results we have observed are a consequence of additional phenomena such as the uncoupling of anabolic and catabolic reactions driven by a period of starvation. An experiment of this nature would be easy to achieve with the methodology presented in Chapter 4. Indeed, the closer the methodology is to the one developed here the easier the comparison of results would be. If the methodology in Chapter 4 is followed exactly, then only the initial composition of culture headspace would need to be altered. This relatively easy study to undertake would provide a strong source of data for developing the argument between heat dissipation and growth dynamics within methanotrophic bacteria.

While the biochemical and thus thermodynamic complexity of methanotrophic growth presents a challenge for determining the relationship between heat dissipation and growth, the sheer quantity of heat dissipated, 56 – 64% of the heat dissipated from the equivalent combustion of methane under standard conditions, provides a significant result in terms of engineering application. In decentralised wastewater treatment systems, where methane is emitted at volumes and concentrations too low to support cost effective capture, this green-house gas is being emitted to the environment and the energy within this resource is lost (Gómez-Borraz et al., 2017). Using methanotrophs, for example within a biofilter, this wasted energy could be repurposed to produce heat. Research, led by Professor William Sloan and Dr Tania Gomez-Borraz, is being undertaken to quantify the amount of methane emitted from household septic tank systems and the effect

temperature has on their operation. The results from this study, coupled with the quantification of the heat dissipated from methanotrophic cultures, as reported here, will allow for an estimate of the available heat to be determined from the incorporation of such a technology. Indeed, the quantification of heat dissipated by methanotrophs presented in Chapter 4, lays out a quantitative framework for analysing the feasibility of such systems. While there is still a lot of research and analysis required to determine the implementation of such technologies, the vast amount of heat dissipated by methanotrophs provides a great enough motivation for further research in this regard. Indeed, a pilot-scale system utilising a methanotroph containing biofilter in series with a septic tank system is currently under construction. While this technology is in its infancy, if achieved it would provide a significant development of current septic tank systems facilitated by the research undertaken within this thesis. Thus, while the results in Chapter 4 have difficulty in providing an analysis of community structuring drivers from first principles, the results have already found purchase within developing wastewater treatment technologies.

Assessing the two studies in tandem demonstrates novel and tangible advancement toward achieving a generalised theory. If the effective neutral parameters, which we have shown to be accurately attainable through purely empirical means, can be linked to the mechanistic description of growth from, for example, our thermodynamic analysis, then surely a general theory will be close to hand. Although this step is arguably the most difficult, if achieved it will have reverberating consequences regarding the design of biotechnologies and thus wastewater treatment systems. Even if this ultimately cannot be achieved, the endeavour itself will undoubtable bear valuable scientific outputs. This thesis serves as a testament to this notion, with the studies presented here making novel contributions to their respective scientific fields and already providing output tied to the engineering development of wastewater treatment technologies. Therefore, the real test of success will lie, not only in the reception with which the research presented here is received, but also in the outcome of research already prompted by the results and conclusions obtained. Although within this thesis fundamentally necessary steps to move toward a generalised theory have been provided, further research is still needed to strengthen the arguments made and, ultimately, to build the critical mass necessary for driving social change and design philosophy (Centola et al., 2018).

Appendix A

European census data displaying the percentage of population connected to independent wastewater treatment



 	TIME	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
GEO											
European Union - 27 countries (from 2020)		:	:	:	:	:	:	:	:	:	:
Belgium		14.78	14.87	13.30	14.15	13.97	12.73	12.81	14.23	13.92	13.92
Bulgaria		25.30 (e)	25.12 (e)	24.50 (e)	24.31 (e)	23.97 (e)	23.81 (e)	23.50 (e)	23.75 (e)	25.21 (e)	:
Czechia		:	:	:	:	:	:	:	:	:	:
Denmark		9.10 (e)	9.00 (e)	8.00	8.30	8.20	8.10	8.00	7.80	7.80	7.58
Germany		3.23	2.80 (e)	2.52 (e)	2.17	2.12 (e)	2.00 (e)	2.04	:	:	:
Estonia		5.00 (e)	5.00 (e)	5.00	5.00	5.00	:	:	:	:	:
Ireland		32.00 (e)	31.70 (e)	31.40 (e)	31.11	31.05 (e)	30.99 (e)	30.93 (e)	30.88 (e)	30.83 (e)	30.77
Greece		:	:	:	:	:	:	:	:	:	:
Spain		1.04 (1)	1.17	1.13 (1)	1.09	0.92 (1)	0.74	0.61 (1)	0.48	0.40	:
France		18.00	18.00	18.00	18.00	18.00	18.00	18.00	18.00	18.00	18.00
Croatia		45.40	45.40	45.40	45.40	45.40	45.40	45.40	45.40	39.06 (b)	:
Italy		:	:	:	:	:	:	:	:	:	:
Cyprus		:	:	:	:	:	0.00	0.00 (1)	0.00	:	:
Latvia		0.28	0.22	0.25	0.28	0.30	0.34	0.29	0.33	0.37	0.38
Lithuania		:	:	:	:	:	:	:	:	:	:
Luxembourg		1.00	:	:	:	:	:	:	:	:	:
Hungary		:	:	:	:	:	:	:	:	:	:
Malta		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Netherlands		0.60	0.60	0.57	0.55	0.50	0.50	0.50	0.45	0.35	0.35
Austria		5.25 (1)	5.00 (e)	4.90 (1)	4.80 (e)	4.43 (1)	4.05 (e)	4.05 (e)	3.96 (e)	3.96 (e)	3.77 (e)
Poland		:	:	:	:	:	:	:	:	:	:
Portugal		:	:	:	:	:	:	:	:	:	:
Romania		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Slovenia		35.20	35.20	35.20	26.73	27.29	28.81	30.01	31.06	26.50	29.26
Slovakia		:	:	:	:	:	:	:	:	:	:
Finland		17.00 (e)	15.00	15.00 (1)	15.00 (1)	15.00 (1)	15.00	15.00	15.00	15.00	15.00
Sweden		13.00	13.00	13.00	13.00	13.00	12.00	12.00	12.00	12.00	:
Iceland		:	:	:	:	:	:	:	:	:	:
Norway		15.68	15.27	15.07	15.08	14.82	14.46	14.41	14.04	13.76	13.43
Switzerland		1.70 (e)	:	:	:	:	:	1.70 (e)	:	:	:
United Kingdom		:	:	:	:	:	:	:	:	:	:
England and Wales		:	:	:	:	:	:	:	:	:	:
Bosnia and Herzegovina		13.00 (e)	13.00 (e)	13.00 (e)	13.00 (e)	13.00 (e)	13.00 (e)	13.00 (e)	:	:	:
Albania		:	:	:	:	:	:	:	:	:	:
Serbia		:	:	:	:	:	:	:	:	:	:
Turkey		0.00 (1)	0.00 (e)	0.00 (e)	0.00 (e)	0.00 (e)	0.00 (e)	0.00 (e)	0.00 (e)	0.00 (e)	0.00 (e)

Figure A.1: European census data displaying the percentage of population connected to independent wastewater treatment. Data available at: <https://ec.europa.eu/eurostat>

Appendix B

Calculation of expected mean from continuous stochastic differential equations for Hubbell's neutral model and Sloan et al.'s near-neutral model

Hubbell's neutral model

The SDE describing the continuous version of Hubbell's neutral model is given by:

$$dx_{i,hubbell} = \frac{m(p_i - x_i)}{\eta N_T} dt + \sqrt{\frac{2x_i(1 - x_i) + m(p_i - x_i)(1 - 2x_i)}{\eta N_T^2}} dW_t \quad (\text{B.1})$$

The expected mean can be calculated by observing equation B.1 under steady-state conditions. At steady state, the expectation of the change $dx_{i,hubbell}$ is zero, and the diffusion term does not contribute to the mean because dW_t has zero expectation. Thus, the expected mean is determined by setting the drift term to zero:

$$\frac{m(p_i - \mathbb{E}[x_i])}{\eta N_T} = 0, \quad (\text{B.2})$$

where, $\mathbb{E}[x_i]$ is the expected mean. Rearranging equation B.2 we get:

$$\mathbb{E}[x_i] = p_i, \quad (\text{B.3})$$

where for our simulations $p_i = 0.2$ and thus $\mathbb{E}[x_i] = 0.2$.

Sloan et al.'s near-neutral model

The SDE describing the continuous version of Hubbell's neutral model is given by:

$$dx_{i,sloan} = \frac{m(p_i - x_i) + 2\alpha(1 - m)x_i(1 - x_i)}{\eta N_T} dt + \sqrt{\frac{2x_i(1 - x_i) + m(p_i - x_i)(1 - 2x_i)}{\eta N_T^2}} dW_t \quad (B.4)$$

The same process is employed as above such that we get:

$$\frac{m(p_i - \mathbb{E}[x_i]) + 2\alpha(1 - m)\mathbb{E}[x_i](1 - \mathbb{E}[x_i])}{\eta N_T} = 0 \quad (B.5)$$

This can be simplified by collecting like terms, resulting in the quadratic equation:

$$-2\alpha(1 - m)\mathbb{E}[x_i]^2 + (-m + 2\alpha(1 - m))\mathbb{E}[x_i] + mp_i = 0 \quad (B.6)$$

This can be solved using the quadratic formula, where our coefficients are:

$$A = -2\alpha(1 - m) \quad (B.7a)$$

$$B = -m + 2\alpha(1 - m) \quad (B.7b)$$

$$C = mp_i \quad (B.7c)$$

Substituting these into the quadratic formula we get:

$$\mathbb{E}[x_i] = \frac{-(-m + 2\alpha(1 - m)) \pm \sqrt{(-m + 2\alpha(1 - m))^2 - 4(-2\alpha(1 - m))(mp_i)}}{2(-2\alpha(1 - m))}, \quad (B.8)$$

where for our simulations $m = 0.1$, $p_i = 0.2$ and $\alpha = 0.01$. Using these values to solve the above equation we get two solutions:

$$\mathbb{E}[x_i] = -4.79 \quad (B.9a)$$

$$\mathbb{E}[x_i] = 0.232, \quad (B.9b)$$

where we ignore the negative solution, giving us $\mathbb{E}[x_i] = 0.232$ as the final answer.

Appendix C

Derivation of continuous stochastic differential equations for Hubbell's neutral model and Sloan et al.'s near-neutral model

Discrete probabilities

The derivation of the continuous stochastic differential equations (SDEs) governing Hubbell's and Sloan's models begins at the discrete transitional probability equations. These equations are commonly written for a single replacement event in which only three outcomes are possible; the abundance N of species i , will grow by one, decrease by one, or stay the same. These equations are provided below for Hubbell's neutral model and Sloan et al.'s near-neutral model respectively:

$$Pr\{N_i + 1|N_i\} = \left(\frac{N_T - N_i}{N_T}\right) \left[mp_i + (1 - m) \left(\frac{N_i}{N_T - 1}\right)\right], \quad (\text{C.1a})$$

$$Pr\{N_i - 1|N_i\} = \left(\frac{N_i}{N_T}\right) \left[m(1 - p_i) + (1 - m) \left(\frac{N_T - N_i}{N_T - 1}\right)\right], \quad (\text{C.1b})$$

$$Pr\{N_i|N_i\} = 1 - Pr\{N_i + 1|N_i\} - Pr\{N_i - 1|N_i\}, \quad (\text{C.1c})$$

$$Pr\{N_i + 1|N_i\} = \left(\frac{N_T - N_i}{N_T}\right) \left[mp_i + (1 + \alpha_i)(1 - m) \left(\frac{N_i}{N_T - 1}\right)\right] \quad (\text{C.2a})$$

$$Pr\{N_i - 1|N_i\} = \left(\frac{N_i}{N_T}\right) \left[m(1 - p_i) + (1 - \alpha_i)(1 - m) \left(\frac{N_T - N_i}{N_T - 1}\right)\right] \quad (\text{C.2b})$$

$$Pr\{N_i|N_i\} = 1 - Pr\{N_i + 1|N_i\} - Pr\{N_i - 1|N_i\} \quad (\text{C.2c})$$

where, parameters hold the same meanings as those within Chapter 3. Given that R represents a specific replacement event, the probability that the population of the species is growing, decreasing or staying the same, when the abundance of the monitored species N_i is equal to an arbitrary value X , can be expressed as $(G_X(R))$, $(D_X(R))$ and $(S_X(R))$ respectively. As such, the transitional probability equations for Hubbell's neutral model and Sloan et al.'s near neutral model can be rewritten as:

$$G_X(R) = \left(\frac{N_T - X}{N_T} \right) \left[mp_i + (1 - m) \left(\frac{X}{N_T - 1} \right) \right], \quad (\text{C.3a})$$

$$D_X(R) = \left(\frac{X}{N_T} \right) \left[m(1 - p_i) + (1 - m) \left(\frac{N_T - X}{N_T - 1} \right) \right], \quad (\text{C.3b})$$

$$S_X(R) = 1 - G_X(R) - D_X(R), \quad (\text{C.3c})$$

$$G_X(R) = \left(\frac{N_T - X}{N_T} \right) \left[mp_i + (1 + \alpha_i)(1 - m) \left(\frac{X}{N_T - 1} \right) \right], \quad (\text{C.4a})$$

$$D_X(R) = \left(\frac{X}{N_T} \right) \left[m(1 - p_i) + (1 - \alpha_i)(1 - m) \left(\frac{N_T - X}{N_T - 1} \right) \right], \quad (\text{C.4b})$$

$$S_X(R) = 1 - G_X(R) - D_X(R) \quad (\text{C.4c})$$

To find the probability that the species of interest i has an abundance of X after $R + 1$ events, the combined probability of outcomes leading to this state must be determined. As progression through time is discrete, outcomes are expressed as the result of the previous event R , where N_i can occupy one of three values, $X - 1$, X and $X + 1$. For each of these values, a single action must occur for N_i to equal X at event $R + 1$. If $N_i = X - 1$ the abundance of i must grow by one, if $N_i = X$ the abundance of i must stay the same, and if $N_i = X + 1$ the abundance of i must decrease by one. The combined probabilities of these outcomes can be expressed in a one-dimensional difference equation:

$$P_X(R + 1) = G_{X-1}(R)P_{X-1}(R) + S_X(R)P_X(R) + D_{X+1}(R)P_{X+1}(R), \quad (\text{C.5})$$

where $P_X(R)$ is equal to the probability that the monitored species has abundance X at event R . Expressing $S_X(R)$ in terms of $G_X(R)$ and $D_X(R)$ gives:

$$P_X(R + 1) = G_{X-1}(R)P_{X-1}(R) + [1 - G_X(R) - D_X(R)]P_X(R) + D_{X+1}(R)P_{X+1}(R), \quad (\text{C.6})$$

which simplifies to:

$$P_X(R + 1) = P_X(R) + [D_{X+1}(R)P_{X+1}(R) - D_X(R)P_X(R)] + [G_{X-1}(R)P_{X-1}(R) - G_X(R)P_X(R)] \quad (\text{C.7})$$

Continuous probabilities

Instead of investigating the change in abundance (N_i), it is more beneficial to examine the relative abundance, this is portrayed as a continuous variable x_i which is equal to $\frac{N_i}{N_T}$ with the domain $[0, 1]$. Given that x_i is defined as a continuous variable N_T must be limited to large values to accommodate this. This grants a further assumption that the community is large enough such that $N_T - 1 \approx N_T$. If once again N_i is equal to the arbitrary abundance value X , such that $x_i = \frac{X}{N_T}$, the discrete transitional probabilities expressed in equations C.3 and C.4, can be recast into continuous forms independent of time/events as shown below:

$$G(x_i) = (1 - x_i)(mp_i + (1 - m)x_i), \quad (\text{C.8a})$$

$$D(x_i) = x_i(m(1 - p_i) + (1 - m)(1 - x_i)), \quad (\text{C.8b})$$

$$S(x_i) = 1 - G(x_i) - D(x_i) \quad (\text{C.8c})$$

$$G(x_i) = (1 - x_i)(mp_i + (1 + \alpha_i)(1 - m)x_i), \quad (\text{C.9a})$$

$$D(x_i) = x_i(m(1 - p_i) + (1 - \alpha_i)(1 - m)(1 - x_i)), \quad (\text{C.9b})$$

$$S(x_i) = 1 - G(x_i) - D(x_i) \quad (\text{C.9c})$$

To define a continuous time variable, the discrete replacement event R is replaced by τ , representing an arbitrary event. This being the case, the one-dimensional difference equation C.7 can now be expressed in terms of $G(x_i)$ and $D(x_i)$ over any interval of replacement events $\delta\tau$. Doing this results in a probability density function (pdf) $\phi(x_i, \tau + \delta\tau)$ that describes the probability of the relative abundance of species i equalling x_i , at event $\tau + \delta\tau$:

$$\phi(x_i, \tau + \delta\tau) = \phi(x_i, \tau) + [D(x_i + \delta x_i)\phi(x_i + \delta x_i, \tau) - D(x_i)\phi(x_i, \tau)] + [G(x_i - \delta x_i)\phi(x_i - \delta x_i, \tau) - G(x_i)\phi(x_i, \tau)], \quad (\text{C.10})$$

where $\phi(x_i, \tau)$ represents the probability density function at event τ .

Solving the probability density function

To solve equation C.10 the first step is to expand and simplify terms $\phi(x_i, \tau + \delta\tau)$, $D(x_i + \delta x_i)\phi(x_i + \delta x_i, \tau)$ and $G(x_i - \delta x_i)\phi(x_i - \delta x_i, \tau)$ using Taylor series expansion, as shown below:

$$\phi(x_i, \tau + \delta\tau) \approx \phi(x_i, \tau) + \delta\tau \frac{\partial \phi}{\partial \tau}(x_i, \tau) + \dots \quad (\text{C.11})$$

$$D(x_i + \delta x_i)\phi(x_i + \delta x_i, \tau) \approx D(x_i)\phi(x_i, \tau) + (\delta x_i) \frac{\partial [D(x_i)\phi(x_i, \tau)]}{\partial (x_i)} + \frac{1}{2}(\delta x_i)^2 \frac{\partial^2 [D(x_i)\phi(x_i, \tau)]}{\partial x_i^2} + \dots \quad (\text{C.12})$$

$$G(x_i - \delta x_i)\phi(x_i - \delta x_i, \tau) \approx G(x_i)\phi(x_i, \tau) - (\delta x_i) \frac{\partial [G(x_i)\phi(x_i, \tau)]}{\partial (x_i)} + \frac{1}{2}(\delta x_i)^2 \frac{\partial^2 [G(x_i)\phi(x_i, \tau)]}{\partial x_i^2} - \dots \quad (\text{C.13})$$

Ignoring higher terms, these expansions can be substituted into equation C.10, which follows:

$$\begin{aligned} \phi(x_i, \tau) + \delta \tau \frac{\partial \phi}{\partial \tau}(x_i, \tau) &= \phi(x_i, \tau) \\ &+ \left[D(x_i)\phi(x_i, \tau) + (\delta x_i) \frac{\partial [D(x_i)\phi(x_i, \tau)]}{\partial (x_i)} + \frac{1}{2}(\delta x_i)^2 \frac{\partial^2 [D(x_i)\phi(x_i, \tau)]}{\partial x_i^2} - D(x_i)\phi(x_i, \tau) \right] \\ &+ \left[G(x_i)\phi(x_i, \tau) - (\delta x_i) \frac{\partial [G(x_i)\phi(x_i, \tau)]}{\partial (x_i)} + \frac{1}{2}(\delta x_i)^2 \frac{\partial^2 [G(x_i)\phi(x_i, \tau)]}{\partial x_i^2} - G(x_i)\phi(x_i, \tau) \right], \end{aligned} \quad (\text{C.14})$$

which simplifies to:

$$\begin{aligned} \delta \tau \frac{\partial \phi}{\partial \tau}(x_i, \tau) &= \left[(\delta x_i) \frac{\partial [D(x_i)\phi(x_i, \tau)]}{\partial (x_i)} + \frac{1}{2}(\delta x_i)^2 \frac{\partial^2 [D(x_i)\phi(x_i, \tau)]}{\partial x_i^2} \right] \\ &+ \left[-(\delta x_i) \frac{\partial [G(x_i)\phi(x_i, \tau)]}{\partial (x_i)} + \frac{1}{2}(\delta x_i)^2 \frac{\partial^2 [G(x_i)\phi(x_i, \tau)]}{\partial x_i^2} \right], \end{aligned} \quad (\text{C.15})$$

and eventually:

$$\begin{aligned} \delta \tau \frac{\partial \phi}{\partial \tau}(x_i, \tau) &= (\delta x_i) \left[\frac{\partial [D(x_i)\phi(x_i, \tau)]}{\partial (x_i)} - \frac{\partial [G(x_i)\phi(x_i, \tau)]}{\partial (x_i)} \right] \\ &+ \frac{1}{2}(\delta x_i)^2 \left[\frac{\partial^2 [D(x_i)\phi(x_i, \tau)]}{\partial x_i^2} + \frac{\partial^2 [G(x_i)\phi(x_i, \tau)]}{\partial x_i^2} \right] \end{aligned} \quad (\text{C.16})$$

The partial differential $\frac{\partial \phi}{\partial \tau}(x_i, \tau)$ can be defined by dividing through by $\delta \tau$:

$$\begin{aligned} \frac{\partial \phi}{\partial \tau}(x_i, \tau) &= \frac{(\delta x_i)}{\delta \tau} \left[\frac{\partial [D(x_i)\phi(x_i, \tau)]}{\partial (x_i)} - \frac{\partial [G(x_i)\phi(x_i, \tau)]}{\partial (x_i)} \right] \\ &+ \frac{1}{2} \frac{(\delta x_i)^2}{\delta \tau} \left[\frac{\partial^2 [D(x_i)\phi(x_i, \tau)]}{\partial x_i^2} + \frac{\partial^2 [G(x_i)\phi(x_i, \tau)]}{\partial x_i^2} \right] \end{aligned} \quad (\text{C.17})$$

Finally, this can be rearranged into the form:

$$\frac{\partial \phi}{\partial \tau}(x_i, \tau) = -\frac{\partial}{\partial x_i} \left[\frac{\delta x_i [G(x_i) - D(x_i)]}{\delta \tau} \phi(x_i, \tau) \right] + \frac{1}{2} \frac{\partial^2}{\partial x_i^2} \left[\frac{\delta x_i^2 [G(x_i) + D(x_i)]}{\delta \tau} \phi(x_i, \tau) \right] \quad (\text{C.18})$$

The Fokker-Planck equation

The Fokker-Planck equation can be defined as:

$$\frac{\partial \phi}{\partial \tau}(x_i, \tau) = -\frac{\partial}{\partial x_i}[\mu(x_i, \tau)\phi(x_i, \tau)] + \frac{1}{2} \frac{\partial^2}{\partial x_i^2}[\sigma^2(x_i, \tau)\phi(x_i, \tau)] \quad (\text{C.19})$$

where $\mu(x_i, \tau)$ is the drift coefficient, and $\sigma^2(x_i, \tau)$ is the diffusion coefficient. These identities can be picked out from equation C.18 such that:

$$\mu(x_i, \tau) = \frac{\delta x_i [G(x_i) - D(x_i)]}{\delta \tau} \quad (\text{C.20a})$$

$$\sigma^2(x_i, \tau) = \frac{\delta x_i^2 [G(x_i) + D(x_i)]}{\delta \tau} \quad (\text{C.20b})$$

Since a timestep represents a replacement event, where each event replaces only a single individual, the magnitude of the change in relative abundance per timestep is always $\frac{1}{N_T}$. In the discrete case, where $\tau = 1$, it follows that $\frac{\delta x_i}{\delta \tau} = \frac{1}{N_T}$ and $\frac{\delta x_i^2}{\delta \tau} = \frac{1}{N_T^2}$ thus:

$$\mu(x_i, \tau) = \frac{G(x_i) - D(x_i)}{N_T} \quad (\text{C.21a})$$

$$\sigma^2(x_i, \tau) = \frac{G(x_i) + D(x_i)}{N_T^2} \quad (\text{C.21b})$$

To express $\mu(x_i, \tau)$ and $\sigma^2(x_i, \tau)$ in specific functional forms as defined by Hubbell's and Sloan et al.'s models, the substitution of equations C.8 and C.9 can be made. This results in the definitions of $\mu(x_i, \tau)$ and $\sigma^2(x_i, \tau)$ for each model given below:

$$\mu(x_i, \tau)_{\text{hubbell}} = \frac{(1 - x_i)(mp_i + (1 - m)x_i) - x_i(m(1 - p_i) + (1 - m)(1 - x_i))}{N_T}, \quad (\text{C.22a})$$

$$\sigma^2(x_i, \tau)_{\text{hubbell}} = \frac{(1 - x_i)(mp_i + (1 - m)x_i) + x_i(m(1 - p_i) + (1 - m)(1 - x_i))}{N_T^2}, \quad (\text{C.22b})$$

$$\mu(x_i, \tau)_{\text{sloan}} = \frac{(1 - x_i)(mp_i + (1 + \alpha_i)(1 - m)x_i) - x_i(m(1 - p_i) + (1 - \alpha_i)(1 - m)(1 - x_i))}{N_T}, \quad (\text{C.23a})$$

$$\sigma^2(x_i, \tau)_{\text{sloan}} = \frac{(1 - x_i)(mp_i + (1 + \alpha_i)(1 - m)x_i) + x_i(m(1 - p_i) + (1 - \alpha_i)(1 - m)(1 - x_i))}{N_T^2}, \quad (\text{C.23b})$$

These simplify to:

$$\mu(x_i, \tau)_{hubbell} = \frac{m(p_i - x_i)}{N_T} \quad (C.24a)$$

$$\sigma^2(x_i, \tau)_{hubbell} = \frac{2x_i(1 - x_i) + m(p_i - x_i)(1 - 2x_i)}{N_T^2} \quad (C.24b)$$

$$\mu(x_i, \tau)_{sloan} = \frac{m(p_i - x_i) + 2\alpha(1 - m)x_i(1 - x_i)}{N_T} \quad (C.25a)$$

$$\sigma^2(x_i, \tau)_{sloan} = \frac{2x_i(1 - x_i) + m(p_i - x_i)(1 - 2x_i)}{N_T^2} \quad (C.25b)$$

The stochastic differential equations

Using equations C.24 and C.25 a stochastic differential equation can be defined for each model, given that it follows the standard form:

$$dx_i = \mu(x_i, \tau)d\tau + \sigma(x_i, \tau)dW_\tau \quad (C.26)$$

Substituting in μ and σ for each model results in the SDEs:

$$dx_{i,hubbell} = \frac{m(p_i - x_i)}{N_T}d\tau + \sqrt{\frac{2x_i(1 - x_i) + m(p_i - x_i)(1 - 2x_i)}{N_T^2}}dW_\tau \quad (C.27)$$

$$dx_{i,sloan} = \frac{m(p_i - x_i) + 2\alpha(1 - m)x_i(1 - x_i)}{N_T}d\tau + \sqrt{\frac{2x_i(1 - x_i) + m(p_i - x_i)(1 - 2x_i)}{N_T^2}}dW_\tau \quad (C.28)$$

Each SDE can be further simplified using the common approximation of the Weiner process component, such that $dW_\tau \approx \sqrt{d\tau}$. Making this simplification results in the final form of both SDEs in which the change in relative abundance dx_i can be estimated across any number of replacement events $d\tau$:

$$dx_{i,hubbell} = \frac{m(p_i - x_i)}{N_T}d\tau + \sqrt{\frac{2x_i(1 - x_i) + m(p_i - x_i)(1 - 2x_i)}{N_T^2}}\sqrt{d\tau} \quad (C.29)$$

$$dx_{i,sloan} = \frac{m(p_i - x_i) + 2\alpha(1 - m)x_i(1 - x_i)}{N_T}d\tau + \sqrt{\frac{2x_i(1 - x_i) + m(p_i - x_i)(1 - 2x_i)}{N_T^2}}\sqrt{d\tau} \quad (C.30)$$

Conversion to real time

So far the derivation of equations C.29 and C.30 has been conducted using discrete replacement events R and the continuous counterpart τ . This reflects the number of replacement events that transpired rather than the quantity of real-time. To convert between the two, H is introduced, which represents the magnitude of real-time that transpires between two

discrete replacement events. Using this identity, T can be calculated following:

$$T = R \cdot H \quad (\text{C.31})$$

Using continuous variables, this equation can be represented as:

$$t = \tau \cdot \eta \quad (\text{C.32})$$

as such the change in time can be defined as:

$$dt = d\tau \cdot \eta \quad (\text{C.33})$$

where η is equal to the average time between replacement events. By rearranging these, descriptions of $d\tau$ and $\sqrt{d\tau}$ can be made.

$$d\tau = \frac{dt}{\eta} \quad (\text{C.34})$$

$$\sqrt{d\tau} = \sqrt{\frac{dt}{\eta}} \quad (\text{C.35})$$

Finally, these can be substituted into equations C.29 and C.30, resulting in:

$$dx_{i,hubbell} = \frac{m(p_i - x_i)}{N_T} \cdot \frac{dt}{\eta} + \sqrt{\frac{2x_i(1 - x_i) + m(p_i - x_i)(1 - 2x_i)}{N_T^2}} \cdot \sqrt{\frac{dt}{\eta}} \quad (\text{C.36})$$

$$dx_{i,sloan} = \frac{m(p_i - x_i) + 2\alpha(1 - m)x_i(1 - x_i)}{N_T} \cdot \frac{dt}{\eta} + \sqrt{\frac{2x_i(1 - x_i) + m(p_i - x_i)(1 - 2x_i)}{N_T^2}} \cdot \sqrt{\frac{dt}{\eta}} \quad (\text{C.37})$$

Appendix D

**Results from truncated version of Hubbell's
neutral model when fitted to single species
relative abundance time series**

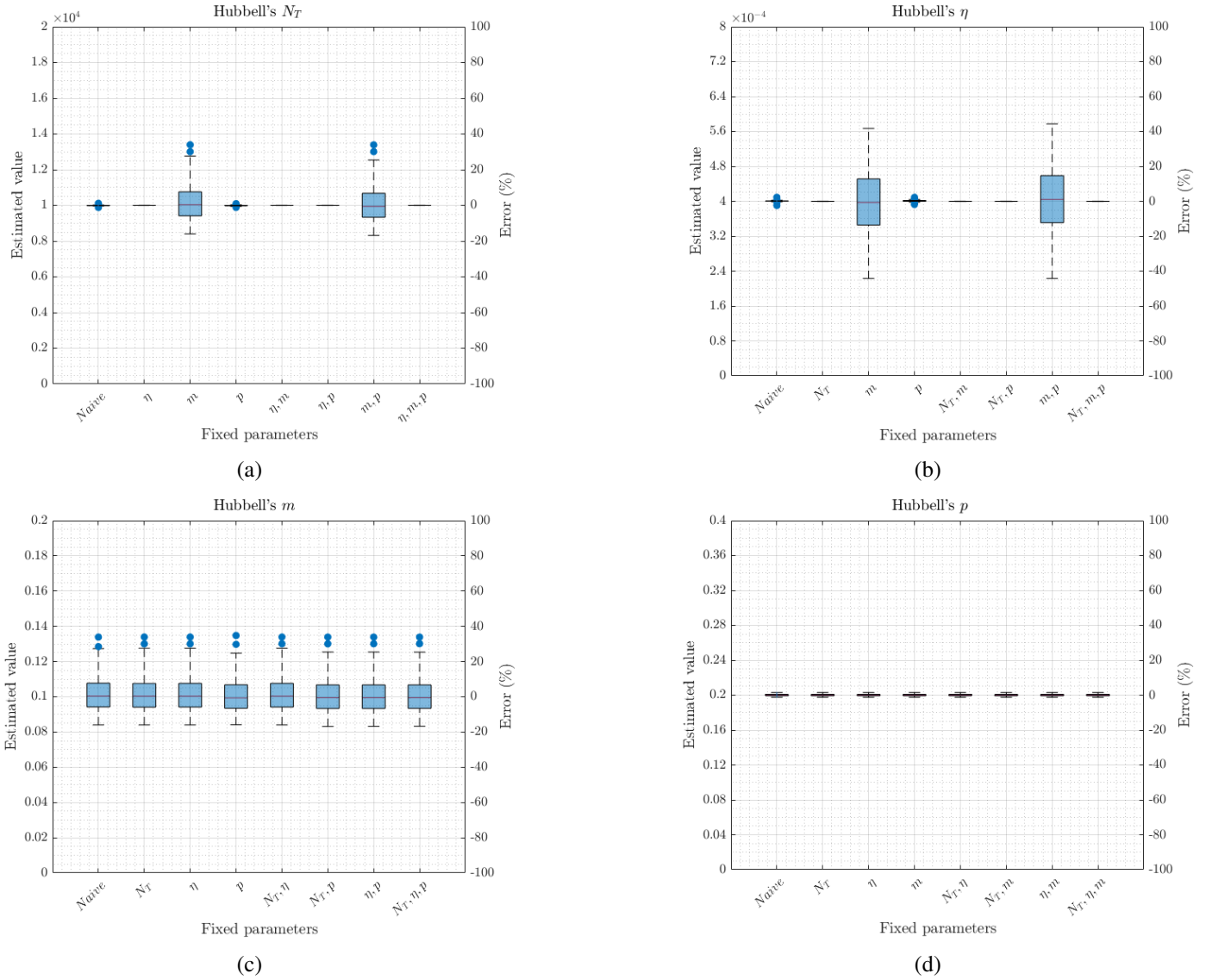
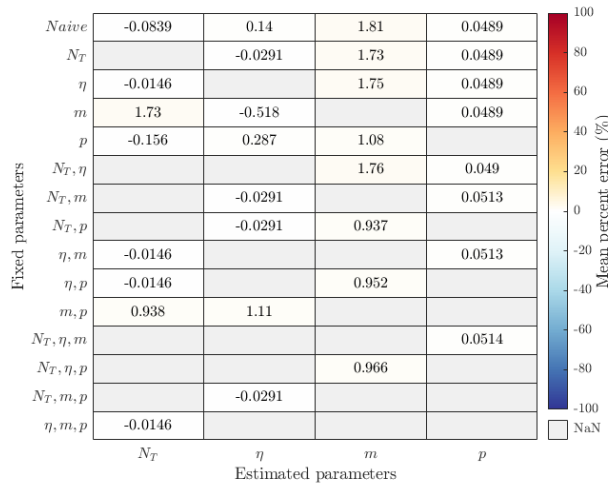
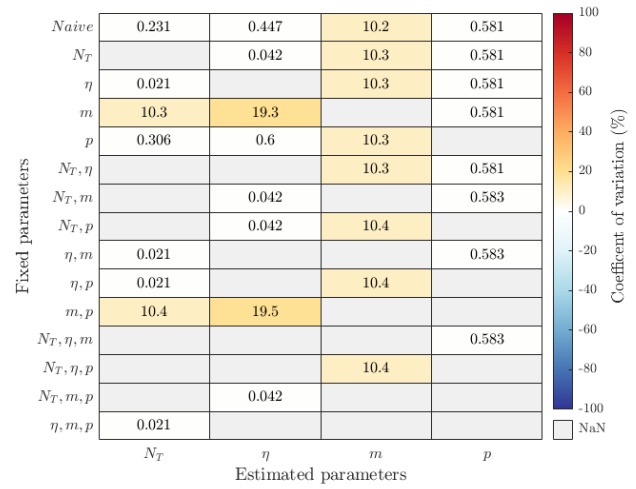


Figure D.1: Box plots representing the estimation of distinct parameters from the truncated version of Hubbell's purely neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.5. a) Estimation of N_T , b) estimation of η , c) estimation of m . d) estimation of p .



(a)



(b)

Figure D.2: Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of the truncated version of Hubbell's purely neutral parameters shown in Figure D.1. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs.

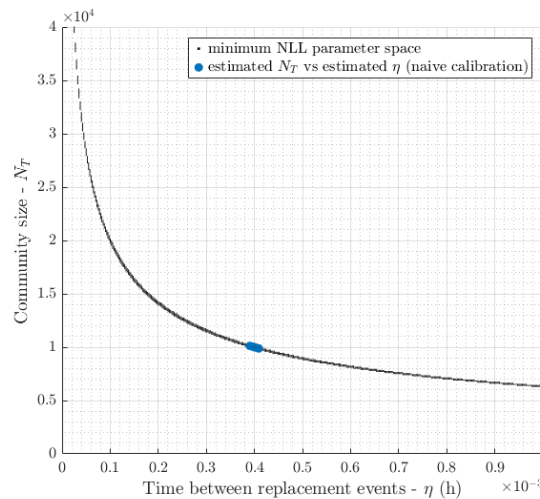
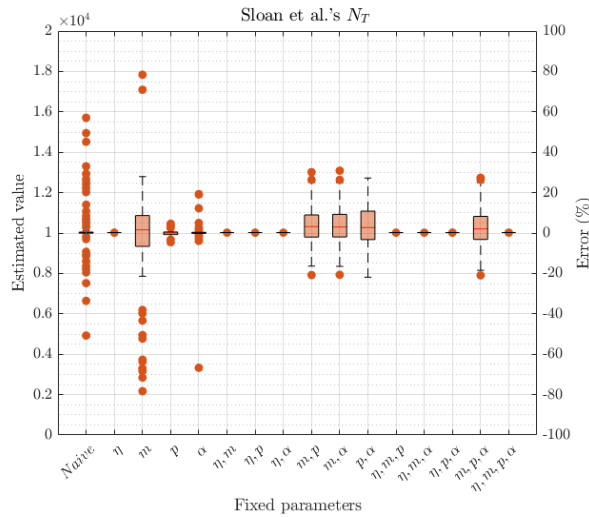


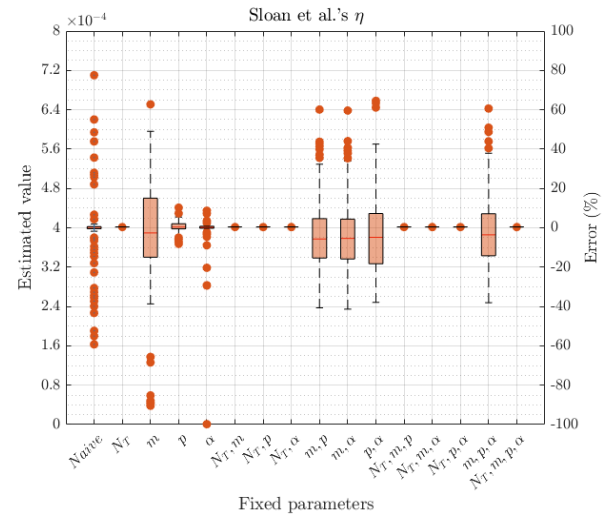
Figure D.3: Estimated values of N_T and η obtained through naive calibration of the truncated version of Hubbell's neutral model to the 100 distinct time series depicted in Figure 3.5 plotted against each other and on top of the midline of negative log-likelihood (NLL) from Figure 3.8.

Appendix E

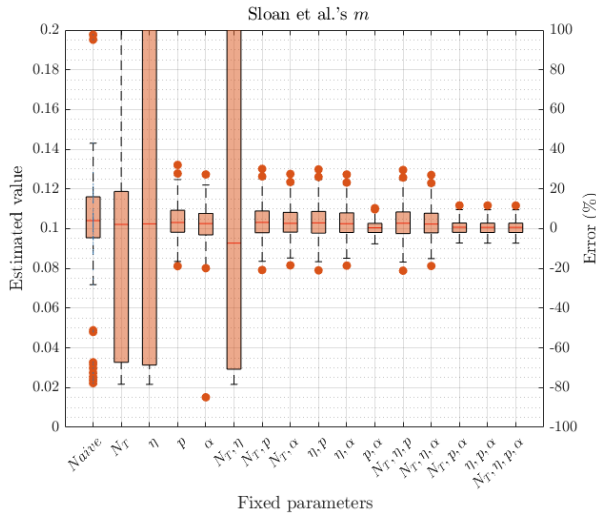
**Results from truncated version of Sloan et al.'s
near-neutral model when fitted to single species
relative abundance time series**



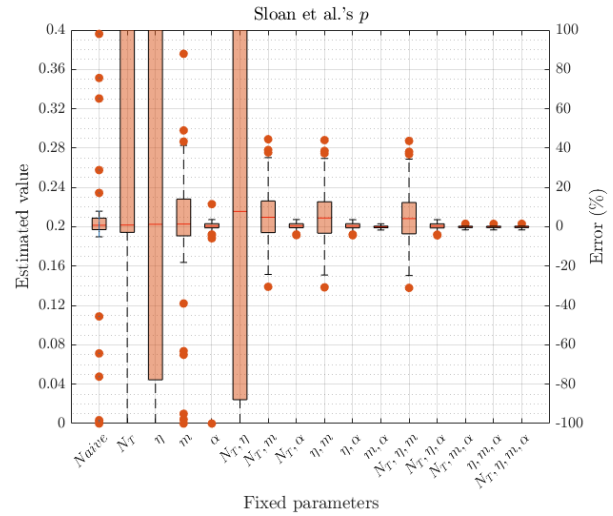
(a)



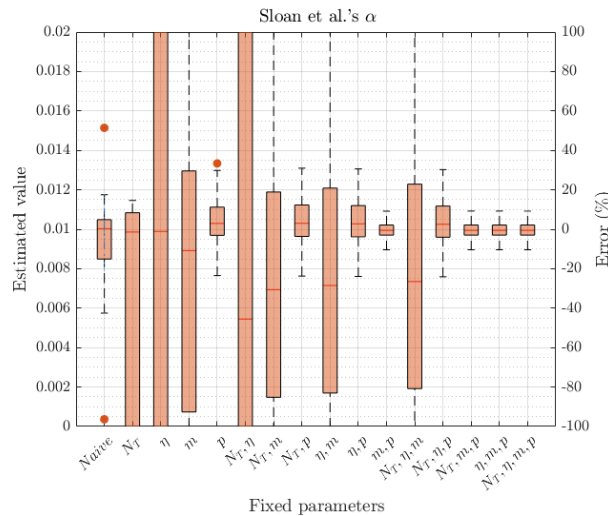
(b)



(c)

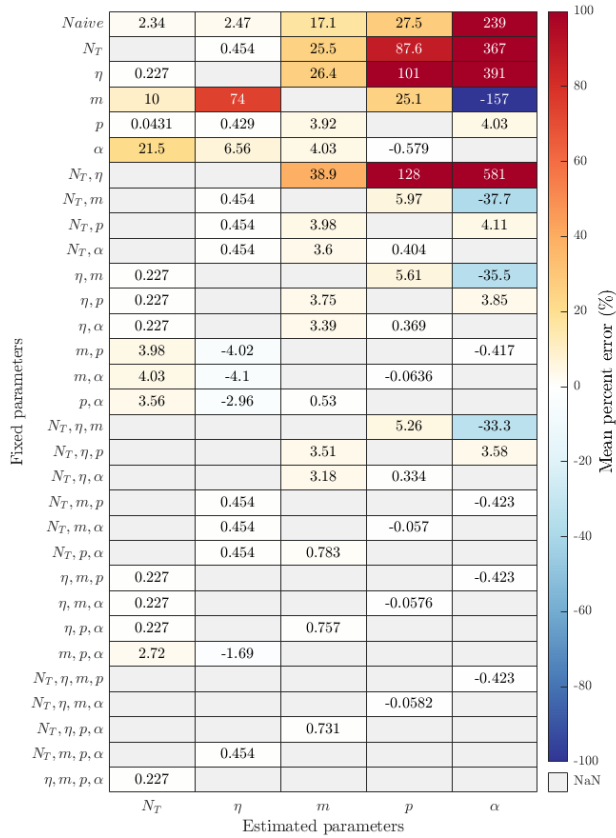


(d)

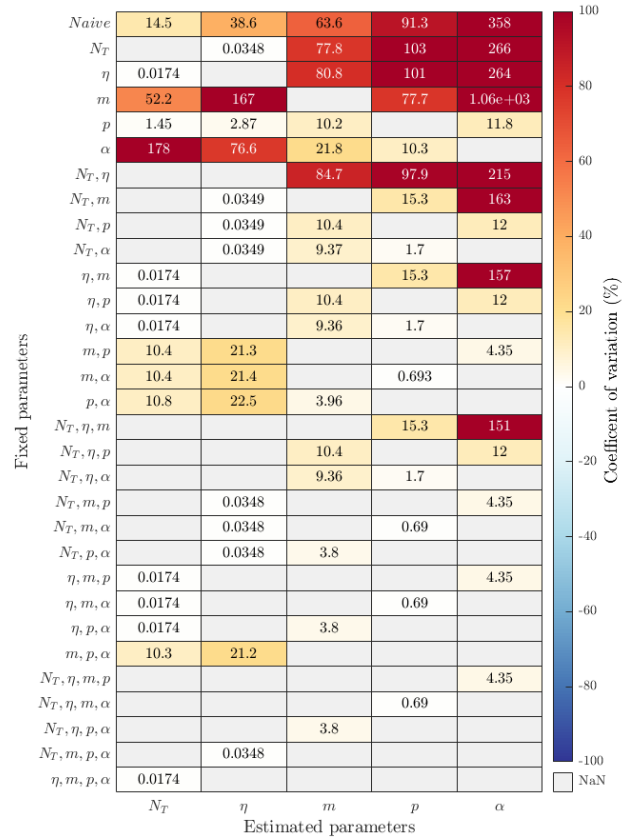


(e)

Figure E.1: Box plots representing the estimation of distinct parameters from the truncated version of Sloan et al.'s near-neutral model under all regimes of system knowledge. The degree of system knowledge is represented by the parameters fixed on the x-axis. The estimated value of the parameter is given on the left y-axis, with the respective percent error displayed on the right y-axis. Each box plot represents calibration to all 100 distinct time series depicted in Figure 3.9. a) Estimation of N_T , b) estimation of η , c) estimation of m , d) estimation of p , e) estimation of α



(a)



(b)

Figure E.2: Heatmaps demonstrating the mean percent error (MPE) and the coefficient of variance (CV) for each set of 100 estimations of the truncated version of Sloan et al.'s near-neutral parameters shown in Figure E.1. The degree of system knowledge is represented by the parameters fixed on the y-axis. The x-axis shows the estimated parameter. a) Heatmap displaying MPEs. b) Heatmap displaying CVs.

Appendix F

Design and volumes of bespoke culture bottles for the Calmetrix I-CAL 4000 HPC Isothermal Calorimeter

Engineering Drawing

The engineering drawings developed for the first prototype of the bespoke bottles are displayed in Figure I.1. During the final production of bottles a medium walled tubing was used rather than thick wall tubing. As such the wall thickness was reduced from 3.5mm to 2.5mm, increasing the internal volume of the bottles. The internal neck diameter was decreased from 13mm to 12mm in the final design

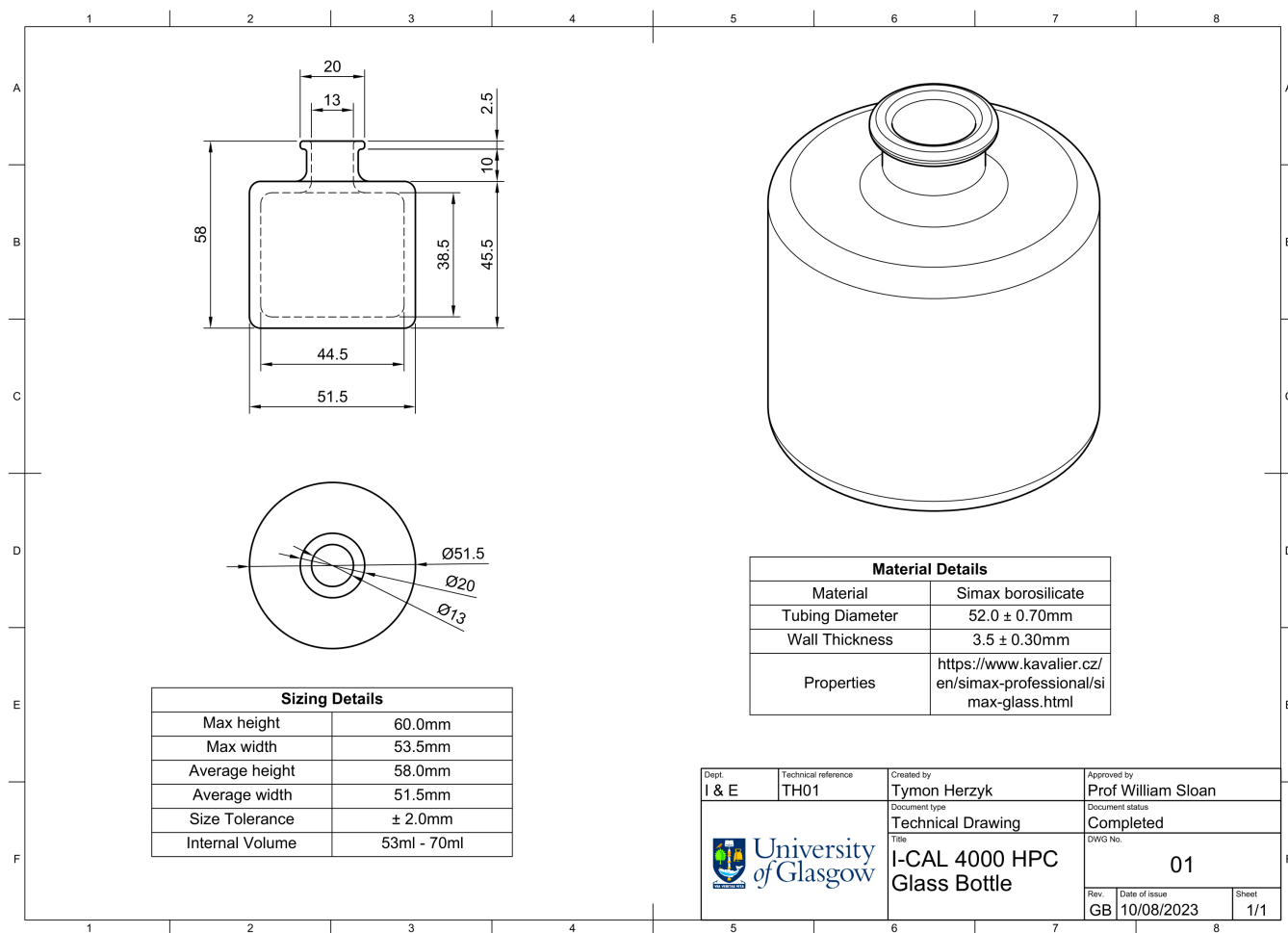


Figure F.1: Initial engineering drawings created for bespoke Calmetrix I-CAL 4000 HPC Isothermal Calorimeter bottles. During final production wall thickness was decreases to 2.5mm, increasing the internal volume of the bottles. The internal neck diameter was also decreased during production from 13mm to 12mm.

Volume measurements

name	number	empty weight (g)	filled weight (g)	volume (ml)
CON 1	6	60.021	126.758	66.737
CON 2	25	52.636	126.717	74.081
CON 3	17	55.921	122.909	66.988
CON 4	27	52.405	126.249	73.844
CON 5	31	54.929	126.548	71.619
CON 6	18	57.020	122.305	65.285
CON 7	19	54.805	118.394	63.589
SNG 1	3	53.804	122.107	68.303
SNG 2	4	61.786	129.087	67.301
SNG 3	5	61.128	129.839	68.711
SNG 4	24	56.558	127.634	71.076
SNG 5	7	59.307	128.237	68.930
SNG 6	8	58.021	127.761	69.740
SNG 7	10	60.118	128.922	68.804
DUP 1	11	59.766	127.313	67.547
DUP 2	26	51.552	121.18	69.628
DUP 3	13	58.951	127.629	68.678
DUP 4	14	53.335	124.173	70.838
DUP 5	15	55.217	123.137	67.920
DUP 6	16	54.462	121.978	67.516
DUP 7	12	58.602	127.416	68.814
TRI 1	20	57.101	123.81	66.709
TRI 2	21	49.924	118.476	68.552
TRI 3	22	50.737	121.47	70.733
TRI 4	23	50.413	119.954	69.541
TRI 5	28	50.407	118.457	68.050
TRI 6	29	56.058	124.909	68.851
TRI 7	32	56.604	125.37	68.766
CAL 1	1	58.419	127.076	68.657
CAL 2	2	57.808	126.212	68.404
CAL 3	9	55.866	125.724	69.858
CAL 4	30	58.029	126.548	68.519

Table F.1: Calculation of individual bottles volume. Volumes were determined by measuring the empty weight and the filled weight of each bottle and finding the difference between these values.

Appendix G

Calculation of the theoretical pressure loss within culture bottles

As stated earlier, the choice was made to ensure that culture bottles were pressurised to negate errors observed from sampling depressurised systems (Soni et al., 1998). There are multiple processes that will enact a loss of pressure on each system within our methodology. In order, these are:

1. Measurement of maximum pressure
2. Starting headspace sample
3. Measurement of starting pressure
4. Conversion of CH_4 to biomass during microbial growth
5. Measurement of final pressure
6. Final Headspace sample

To ensure that the pressure of the final headspace sample is $\geq 1\text{atm}$, we can work backwards to calculate the starting pressure required within each system. Here we use an average bottle volume of 68ml .

Final Headspace sample

Final headspace samples are withdrawn from each bottle on removal from the incubator. Given that the number of moles is constant, the headspace volume is 38ml at a temperature of 30°C , the headspace sample withdrawn has a volume of 3.1ml and the syringe is at room temperature (22°C). The conditions of this calculation can be given as:

- $P_1 = ?$
- $V_1 = 38\text{ml}$
- $T_1 = 303.15\text{K}$
- $P_2 = 1\text{atm}$

- $V_2 = 41.1ml$
- $T_2 = 295.15K$

From the ideal gas law, where the number of moles is constant we get the relationship:

$$\frac{PV}{T} = constant, \quad (G.1)$$

thus:

$$\frac{P_1 V_1}{T_1} = \frac{P_2 V_2}{T_2}, \quad (G.2)$$

and rearranging for P_1 we get:

$$P_1 = \frac{P_2 V_2 T_1}{V_1 T_2} \quad (G.3)$$

Substituting in the identities previously defined we get:

$$P_1 = \frac{(1)(41.1)(303.15)}{(38)(295.15)} = 1.111atm \quad (G.4)$$

Measurement of final pressure

During the measurement of pressure using a manometer some pressure within the system will be lost as the pressure will equalise across the system and the additional volume introduced. The volume of the system used to measure pressure must be defined. An estimate of this volume is calculated to be 2.391ml, based on the dimensions of the tubing used and the size of the needle. The gas within this volume will be at atmospheric pressure and as pressure measurement is taken within the incubator we can consider the temperature to be constant. From this description, we can calculate the pressure in the bottle before the pressure reading by the conservation of moles. From the ideal gas law, the total number of moles within the bottle and within the manometer can be expressed separately by:

$$n_1 = \frac{P_1 V_1}{RT_1}, \quad (G.5a)$$

$$n_2 = \frac{P_2 V_2}{RT_2}, \quad (G.5b)$$

where n_1 is the total number of moles within the bottle and n_2 is the total number of moles within the manometer. P , V and T are the respective pressure, volume and temperature of each system and R is the universal gas constant. Given that the total number of moles is conserved when these two system are added together we get:

$$n_1 + n_2 = n_3, \quad (G.6a)$$

$$\frac{P_1 V_1}{RT_1} + \frac{P_2 V_2}{RT_2} = \frac{P_3 (V_1 + V_2)}{RT_3}, \quad (G.6b)$$

since RT is constant across all three systems, this equation can be simplified to:

$$P_1 V_1 + P_2 V_2 = P_3 (V_1 + V_2). \quad (G.7)$$

From the previous calculation, we know what the final pressure within the combined system must be equal to at least $1.111atm$ (P_3). We also know both volumes (V_1 and V_2) and assume that the pressure within the manometer (P_2) is at the air pressure within the incubator. Thus these values are defined as follows:

- $P_1 = ?$
- $V_1 = 38ml$
- $P_2 = 1.027atm$
- $V_2 = 2.391ml$
- $P_3 = 1.111atm$

Rearranging equation G.7 to find P_1 gives:

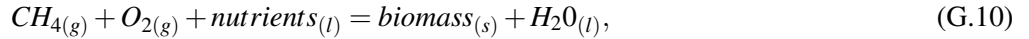
$$P_1 = \frac{P_3(V_1 + V_2) - P_2V_2}{V_1}, \quad (G.8)$$

substituting the values provided for these identities, we get:

$$P_1 = \frac{(1.111)[(38) + (2.391)] - (1.027)(2.391)}{38} = 1.116atm \quad (G.9)$$

Conversion of methane to biomass during methanotrophic growth

Under the most extreme scenario for pressure loss from the growth of methanotrophs, we assume that during the oxidation of methane, all methane is converted to biomass such that the stoichiometric reaction can be expressed as,



thus for every mole of CH_4 oxidised, one mole of O_2 is also consumed. As no gas phase products are produced, a negative net change in the moles of gas within the system is observed. If we assume enough oxygen is present to facilitate the oxidation of all methane within the system then the total number of moles consumed can be described by:

$$\xi = 2\xi_{(CH_4)}, \quad (G.11)$$

where ξ is the total number of moles consumed and $\xi_{(CH_4)}$ is the number of moles of methane consumed. From the ideal gas law, and assuming temperature and volume are constant, the relationship between the number of moles and pressure can be expressed as:

$$\frac{P_1}{n_1} = \frac{P_2}{n_2}, \quad (G.12)$$

where P refers to the pressure of the system and n is the total number of moles of gas in the system. If we assess the number of moles of gas before and after the consumption of methane then we can express n_2 as:

$$n_2 = n_1 - 2\xi_{(CH_4)}, \quad (G.13)$$

This definition can be substituted into equation G.13 resulting in:

$$\frac{P_1}{n_1} = \frac{P_2}{n_1 - 2\xi_{(CH_4)}}, \quad (G.14)$$

rearranging we get:

$$P_1 = P_2 \left(\frac{n_1}{n_1 - 2\xi_{(CH_4)}} \right), \quad (G.15)$$

Under the experimental methodology developed, the maximum amount of methane expected is $\approx 8\%$ of the headspace, and thus $\xi_{(CH_4)} = 0.08n_1$. Substituting this identity in equation G.15 we arrive at:

$$P_1 = P_2 \left(\frac{n_1}{n_1 - 2(0.08n_1)} \right) = \frac{P_2}{0.84} \quad (G.16)$$

Given that from the previous calculation $P_2 = 1.116atm$ we get:

$$P_1 = \frac{1.116}{0.84} = 1.329atm \quad (G.17)$$

Measurement of starting pressure

Before bottles are added to incubators to facilitate the growth of methanotrophs a measurement of pressure is taken. This measurement is taken at room temperature and as the system is the same as that used during the final pressure reading we can use a similar calculation process, where the conditions are:

- $P_1 = ?$
- $V_1 = 38ml$
- $P_2 = 1atm$
- $V_2 = 2.391ml$
- $P_3 = 1.294atm$

Here P_3 is lower than what was calculated under the previous step as it has been adjusted for the drop in temperature using:

$$\frac{P_1}{T_1} = \frac{P_2}{T_2}, \quad (G.18)$$

where P_1 is the pressure in the bottle before it is added to the incubator and thus $T_1 = 295.15K$ (room temperature), P_2 is the pressure in the bottle within the incubator as calculated previously to be $1.329atm$ and thus $T_2 = 303.15K$ (incubator temperature). Substituting all the required values into equation G.8 we get:

$$P_1 = \frac{(1.294)[(38) + (2.391)] - (1)(2.391)}{38} = 1.312atm \quad (G.19)$$

Starting Headspace sampling

We now have to calculate the drop in pressure associated with the starting headspace sample. Here we can use the relationship derived earlier as shown in equation G.3. Here the temperature is constant as both the bottle and syringe are at room temperature. Thus equation G.3 can be simplified to:

$$P_1 = \frac{P_2 V_2}{V_1} \quad (\text{G.20})$$

where:

- $P_1 = ?$
- $V_1 = 38ml$
- $P_2 = 1.312atm$
- $V_2 = 41.1ml$

substituting these in, we get:

$$P_1 = \frac{(1.312)(41.1)}{38} = 1.419atm \quad (\text{G.21})$$

Measurement of maximum pressure

The final step is to calculate the pressure loss as a result of measuring the maximum pressure. This is measurement is taken immediately after gases have been added to each bottle. Once again we can use equation G.8 to calculate this as the temperature is constant. Here the value for each identity will be as follows

- $P_1 = ?$
- $V_1 = 38ml$
- $P_2 = 1atm$
- $V_2 = 2.391ml$
- $P_3 = 1.419atm$

Substituting all the required values into equation G.8 we get:

$$P_1 = \frac{(1.419)[(38) + (2.391)] - (1)(2.391)}{38} = 1.445atm \quad (\text{G.22})$$

To ensure the pressure within the syringe during the final headspace sample is at ambient pressure, the initial pressure within each bottle must be $\approx 1.45atm$ dependent on the exact volume of each bottle.

Appendix H

Post-processing of thermal power measurements from the Calmetrix I-CAL 4000 HPC Isothermal Calorimeter

To obtain the measurements presented in Figure 4.8, the raw values obtained from the Calmetrix I-CAL 4000 HPC Isothermal Calorimeter must go through several steps of post processing. These are:

- Removal of baseline offset
- Removal of settling period
- Integration of thermal power measurements

Figure H.1 depicts the raw measurements of thermal power from the Calmetrix I-CAL 4000 HPC Isothermal Calorimeter for each experiment. As depicted, each chamber within the calorimeter, relating to a single curve on each plot, exhibits an offset. To remove this offset the last 12 hours of measurements for each curve are average and this value is subtracted from all prior measurements. This results in the plots provided in figure H.2.

Within figures H.1 and H.2 a settling period is observed, which impacts measurements made within the first 24 hours of an experiment. Thus rather than removing this period effort is made to firstly dampen this effect and then simulate data in its place. To dampen the effect of the settling period the first 24 hours from respective control samples (the flat line on each plot) is subtracted from the other curves. While this does dampen the settling period, due to slight differences within the time of the settling periods, removing the control measurements presents high errors within the first 10 hours. Thus the first 10 hours of each curve is removed. An exponential curve is then fit using data from the subsequent 24 hours to provide data for the first 10 hours. This new data is then appended onto the measurement curves. The final result is a full removal of the settling period as demonstrated in figure H.3.

The final processing step is the conversion of thermal power to thermal energy. Here this is undertaken within Matlab using the function "trapz" which facilitates trapezoidal integration. A unit spacing of 60 seconds is used as measurements of thermal power are taken every minute. The values obtained are summed at every time point to produce the plots of thermal energy shown in figure H.4.

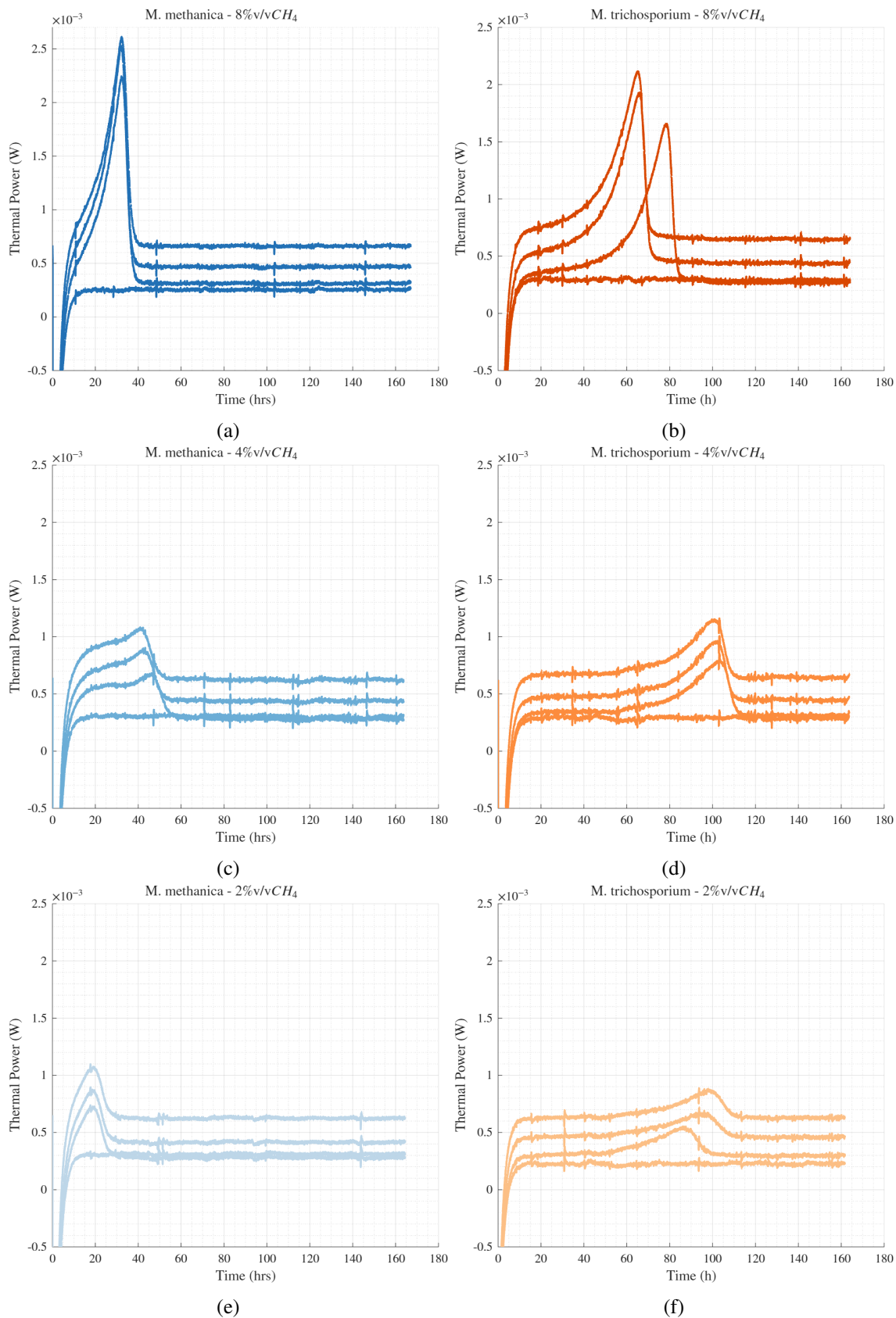


Figure H.1: Raw thermal power measurements made using a Calmetrix I-CAL 4000 HPC Isothermal Calorimeter from batch cultures of *M. methanica* S1 and *M. trichosporium* OB3b under three initial concentration of CH₄ (8%, 4% and 2%v/v in air). a) *M. methanica* S1 8%v/vCH₄, b) *M. trichosporium* OB3b 8%v/vCH₄, c) *M. methanica* S1 4%v/vCH₄, d) *M. trichosporium* OB3b 4%v/vCH₄, e) *M. methanica* S1 2%v/vCH₄, f) *M. trichosporium* OB3b 2%v/vCH₄.

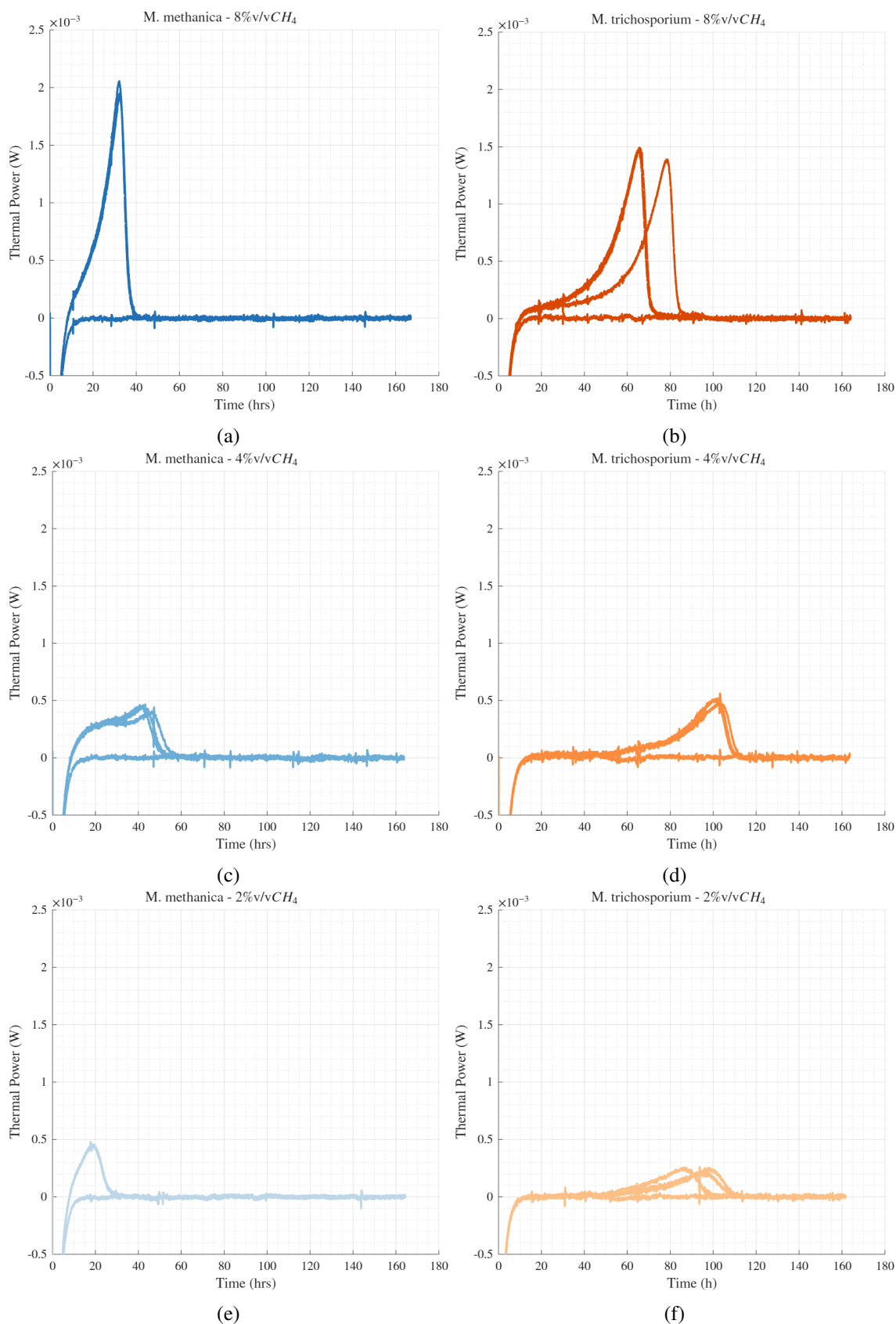


Figure H.2: Thermal power measurements provided in figure H.1 with baseline reduction. a) *M. methanica* S1 8%v/v CH_4 , b) *M. trichosporium* OB3b 8%v/v CH_4 , c) *M. methanica* S1 4%v/v CH_4 , d) *M. trichosporium* OB3b 4%v/v CH_4 , e) *M. methanica* S1 2%v/v CH_4 , f) *M. trichosporium* OB3b 2%v/v CH_4 .

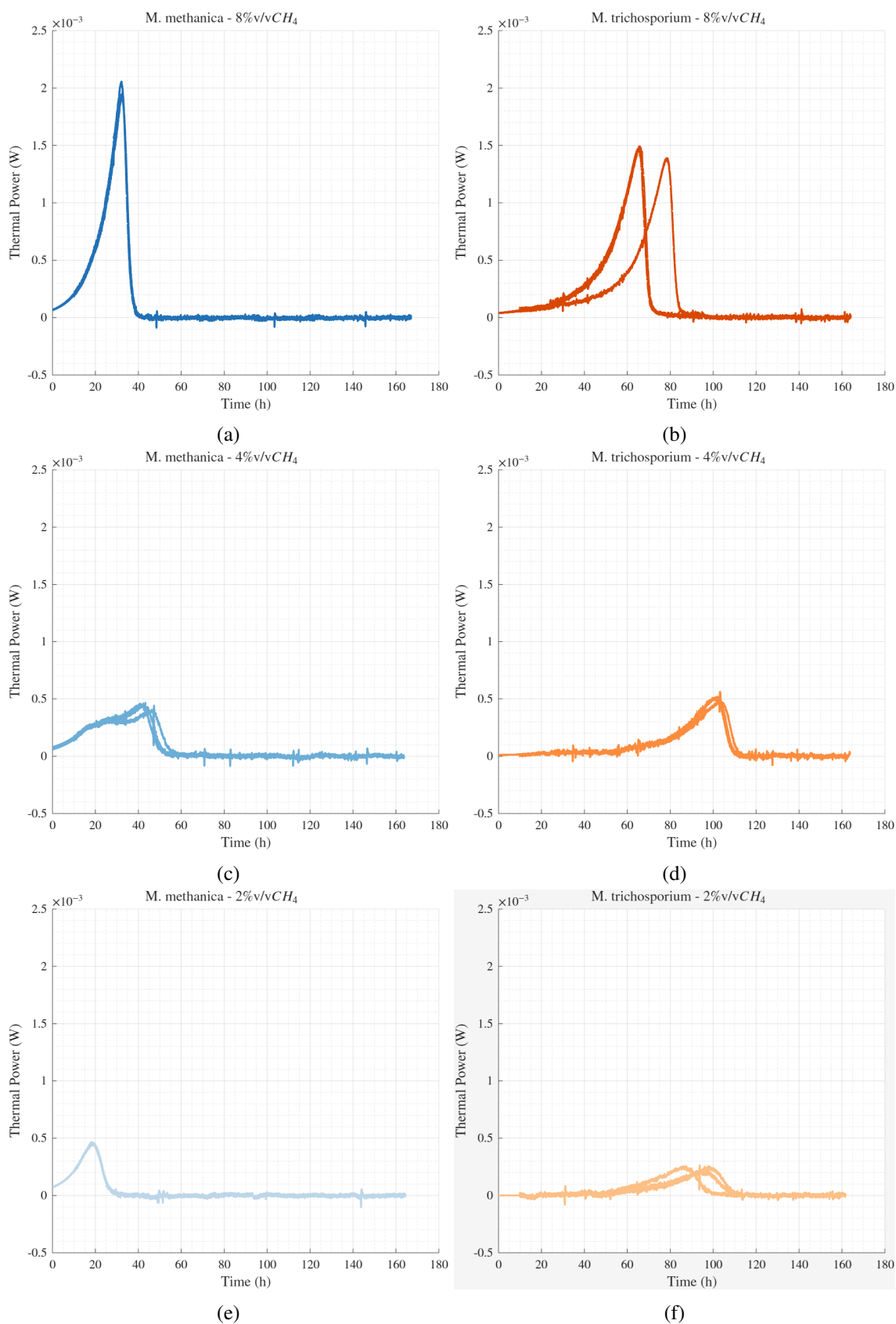


Figure H.3: Thermal power measurements provided in figure H.1 with baseline and settling period reduction. a) *M. methanica* S1 8%v/vCH₄, b) *M. trichosporium* OB3b 8%v/vCH₄, c) *M. methanica* S1 4%v/vCH₄, d) *M. trichosporium* OB3b 4%v/vCH₄, e) *M. methanica* S1 2%v/vCH₄, f) *M. trichosporium* OB3b 2%v/vCH₄.

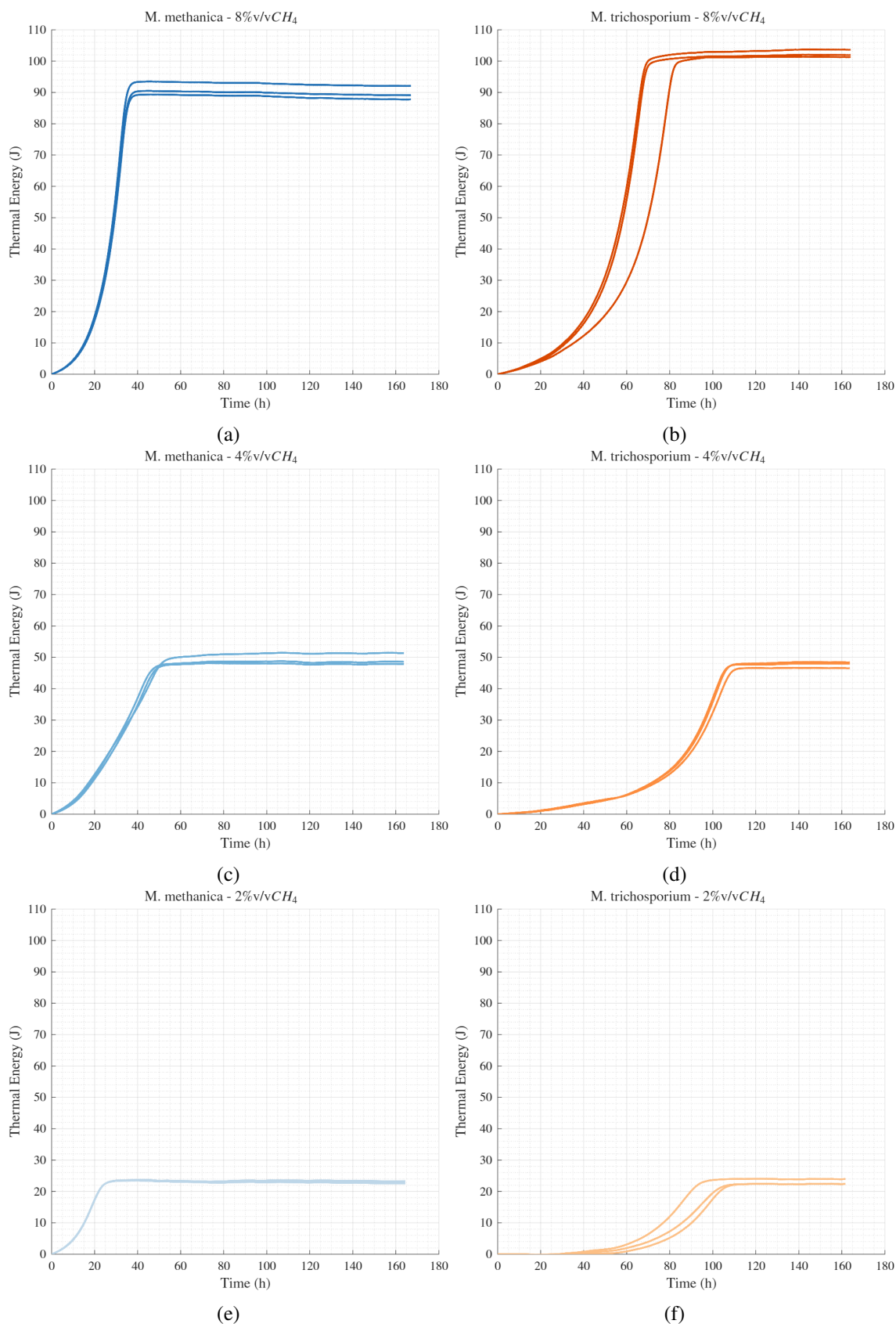


Figure H.4: Thermal energy measurements calculated from the thermal power measurements provided in figure H.3. a) *M. methanica* S1 8%v/vCH₄, b) *M. trichosporium* OB3b 8%v/vCH₄, c) *M. methanica* S1 4%v/vCH₄, d) *M. trichosporium* OB3b 4%v/vCH₄, e) *M. methanica* S1 2%v/vCH₄, f) *M. trichosporium* OB3b 2%v/vCH₄.

Appendix I

Calibration curves for OD_{600} vs CDW for *M. methanica* S1 and *M. trichosporium* OB3b

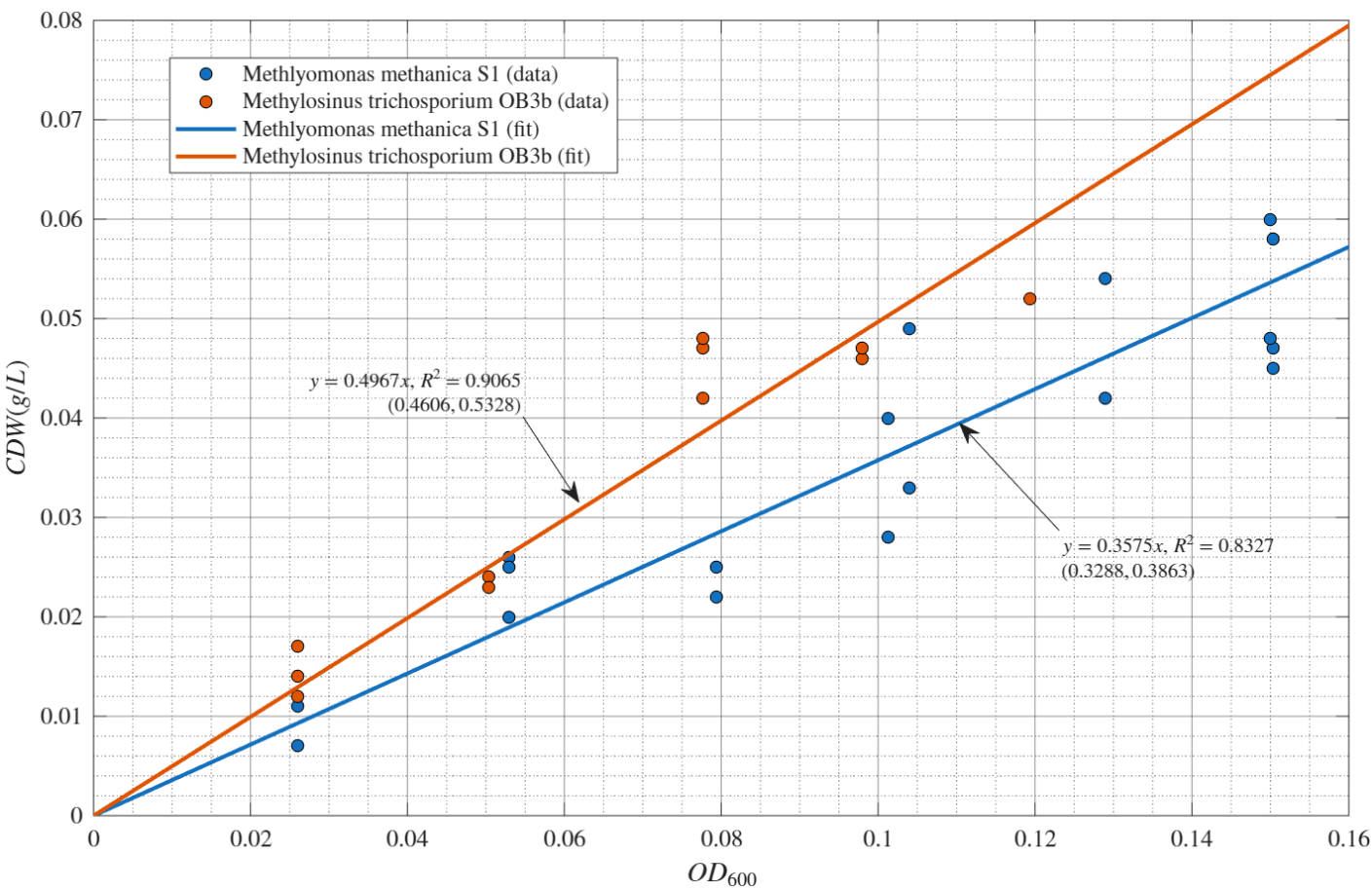


Figure I.1: Calibration curve for determining cell dry weight from OD_{600} .

Appendix J

Measurements of dissolved oxygen for batch cultures of *M. trichosporium* OB3b and *M. methanica* S1

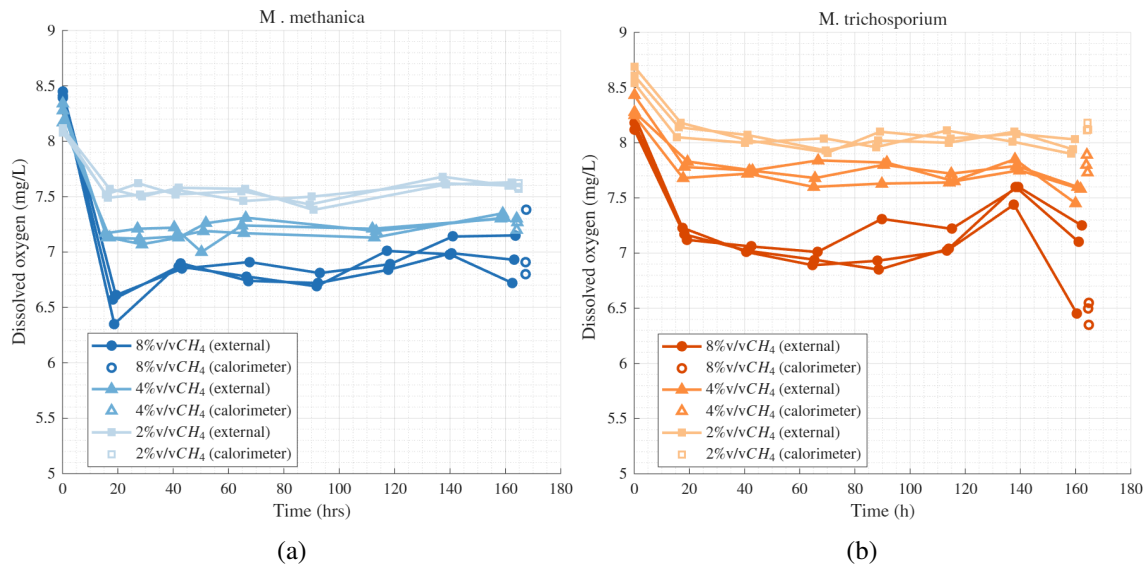


Figure J.1: Time series of dissolved oxygen measurements for external and calorimeter cultures of *M. methanica* S1 and *M. trichosporium* OB3b, measured under three distinct concentration of initial CH_4 over a period of 7 days. a) *M. methanica* S1, b) *M. trichosporium* OB3b

Appendix K

Calculation of total number of moles of CH_4 and CO_2 within culture bottles

This is an example calculation of the total number of moles and the mass of CH_4 and CO_2 in the system. Values chosen are from sample CAL1.

Known quantities

Measurements (initial)

- $\%CH_4 = 7.88\% \text{ v/v}$
- $\%CO_2 = 0.08\% \text{ v/v}$
- $p = 1.30 \text{ atm}$
- $T = 22^\circ C (295.15 K)$

Measurements (final)

- $\%CH_{4,final} = 0.00\%$
- $\%CO_{2,final} = 1.83\%$
- $p_{final} = 1.15 \text{ atm}$
- $T_{final} = 30^\circ C (303.15 K)$

Constants

- $V_{bottle} = 68.7 \text{ ml}$
- $V_{media} = 30.0 \text{ ml}$

- $V_{headspace} = 38.7ml$
- $R = 8.314J/molK$
- $R = 0.0821Latm/molK$
- $H_{s,CH_4}^{\circ} = 1.4 \times 10^{-5}mol/m^3Pa$
- $H_{s,CO_2}^{\circ} = 3.4 \times 10^{-4}mol/m^3Pa$
- $\Delta H_{H_2CO_3}^{\circ} = 9151J$
- $\frac{d \ln H_{s,CH_4}}{d(1/T)} = 1600$
- $\frac{d \ln H_{s,CO_2}}{d(1/T)} = 2300$

Number of moles in headspace

To calculate the total number of moles in the headspace the ideal gas law is used:

$$pV = nRT, \quad (K.1a)$$

$$n = \frac{pV}{RT} \quad (K.1b)$$

From initial measurements:

$$n_{total,headspace} = \frac{(1.30)(0.0387)}{(0.0821)(295.15)} = 207.62 \times 10^{-5}mol \quad (K.2)$$

From final measurements:

$$n_{total,headspace} = \frac{(1.15)(0.0387)}{(0.0821)(303.15)} = 178.82 \times 10^{-5}mol \quad (K.3)$$

The number of moles of CH_4 and CO_2 can be found using the % v/v concentrations of CH_4 and CO_2 measured in the headspace respectively.

From initial measurements:

$$n_{CH_4,headspace} = \frac{207.62 \times 10^{-5}}{100}(7.88) = 16.36 \times 10^{-5}mol, \quad (K.4)$$

$$n_{CO_2,headspace} = \frac{207.62 \times 10^{-5}}{100}(0.08) = 0.17 \times 10^{-5}mol \quad (K.5)$$

From final measurements:

$$n_{CH_4,headspace} = \frac{178.82 \times 10^{-5}}{100}(0.00) = 0.00mol, \quad (K.6)$$

$$n_{CO_2,headspace} = \frac{178.82 \times 10^{-5}}{100}(1.83) = 3.27 \times 10^{-5}mol \quad (K.7)$$

Number of moles dissolved

To calculate the number of moles dissolved we use Henry's law:

$$C = H_s p_{gas}, \quad (K.8)$$

where C is the solubility at a fixed temperature, H_s is Henry's coefficient and p_{gas} is the partial pressure of the specific gas. From literature (Sander, 2023) we can find the Henry's coefficient for both CH_4 and CO_2 at $25^\circ C$. This can then be adjusted to the temperature of the system using the Van't Hoff equation:

$$H_s = H_s^\circ \times \exp \left(B \left(\frac{1}{T} - \frac{1}{T^\circ} \right) \right), \quad (K.9)$$

where $B = \frac{-\Delta_{sol}H}{R} = \frac{d \ln H_s}{d(1/T)}$. Values for $\frac{d \ln H_s}{d(1/T)}$ can also be taken directly from literature (Sander, 2023). Thus to calculate H_{s,CH_4} follows:

From initial measurements:

$$H_{s,CH_4} = H_{s,CH_4}^\circ \times \exp \left(\frac{d \ln H_{s,CH_4}}{d(1/T)} \left(\frac{1}{T} - \frac{1}{T^\circ} \right) \right), \quad (K.10a)$$

$$H_{s,CH_4} = 1.4 \times 10^{-5} \times \exp \left(1600 \left(\frac{1}{295.15} - \frac{1}{298.15} \right) \right), \quad (K.10b)$$

$$H_{s,CH_4} = 1.4785 \times 10^{-5} \text{ mol/m}^3 \text{ Pa}, \quad (K.10c)$$

$$H_{s,CO_2} = H_{s,CO_2}^\circ \times \exp \left(\frac{d \ln H_{s,CO_2}}{d(1/T)} \left(\frac{1}{T} - \frac{1}{T^\circ} \right) \right), \quad (K.11a)$$

$$H_{s,CO_2} = 3.4 \times 10^{-4} \times \exp \left(2300 \left(\frac{1}{295.15} - \frac{1}{298.15} \right) \right), \quad (K.11b)$$

$$H_{s,CO_2} = 3.6773 \times 10^{-4} \text{ mol/m}^3 \text{ Pa} \quad (K.11c)$$

From final measurements:

$$H_{s,CH_4} = H_{s,CH_4}^\circ \times \exp \left(\frac{d \ln H_{s,CH_4}}{d(1/T)} \left(\frac{1}{T} - \frac{1}{T^\circ} \right) \right), \quad (K.12a)$$

$$H_{s,CH_4} = 1.4 \times 10^{-5} \times \exp \left(1600 \left(\frac{1}{303.15} - \frac{1}{298.15} \right) \right), \quad (K.12b)$$

$$H_{s,CH_4} = 1.2814 \times 10^{-5} \text{ mol/m}^3 \text{ Pa}, \quad (K.12c)$$

$$H_{s,CO_2} = H_{s,CO_2}^{\circ} \times \exp\left(\frac{d\ln H_{s,CO_2}}{d(1/T)}\left(\frac{1}{T} - \frac{1}{T^{\circ}}\right)\right), \quad (K.13a)$$

$$H_{s,CO_2} = 3.4 \times 10^{-4} \times \exp\left(2300\left(\frac{1}{303.15} - \frac{1}{298.15}\right)\right), \quad (K.13b)$$

$$H_{s,CO_2} = 2.9938 \times 10^{-4} \text{ mol/m}^3 \text{ Pa} \quad (K.13c)$$

As we have the respective concentrations of CH_4 and CO_2 we can calculate the partial pressures.

From initial measurements:

$$p_{CH_4} = \frac{1.30}{100}(7.88) = 0.1024 \text{ atm}, \quad (K.14)$$

$$p_{CO_2} = \frac{1.30}{100}(0.08) = 0.0010 \text{ atm} \quad (K.15)$$

From final measurements:

$$p_{CH_4} = \frac{1.15}{100}(0.00) = 0.0000 \text{ atm}, \quad (K.16)$$

$$p_{CO_2} = \frac{1.15}{100}(1.83) = 0.0210 \text{ atm} \quad (K.17)$$

Using the previously calculated identities and converting pressure from atmospheres (atm) to pascals (Pa) (multiply by 101325), the solubility of CH_4 and CO_2 can be calculated.

From initial measurements:

$$C_{CH_4} = H_{s,CH_4} p_{CH_4} = (1.4785 \times 10^{-5})(0.1024)(101325) = 0.1534 \text{ mol/m}^3 \quad (K.18)$$

$$C_{CO_2} = H_{s,CO_2} p_{CO_2} = (3.6773 \times 10^{-4})(0.0010)(101325) = 0.0373 \text{ mol/m}^3, \quad (K.19)$$

From final measurements:

$$C_{CH_4} = H_{s,CH_4} p_{CH_4} = (1.2814 \times 10^{-5})(0.0000)(101325) = 0.0000 \text{ mol/m}^3 \quad (K.20)$$

$$C_{CO_2} = H_{s,CO_2} p_{CO_2} = (2.9938 \times 10^{-4})(0.0210)(101325) = 0.6370 \text{ mol/m}^3, \quad (K.21)$$

The number of moles dissolved can be found by multiplying the solubility by the volume of media, where $1 \text{ ml} = 1 \times 10^{-6} \text{ m}^3$.

From initial measurements:

$$n_{CH_4,dissolved} = (0.1534)(30.0 \times 10^{-6}) = 0.4602 \times 10^{-5} \text{ mol}, \quad (K.22)$$

$$n_{CO_2,dissolved} = (0.0373)(30.0 \times 10^{-6}) = 0.1119 \times 10^{-5} \text{ mol}, \quad (K.23)$$

From final measurements:

$$n_{CH_4,dissolved} = (0.0000)(30.0 \times 10^{-6}) = 0.0000 \text{ mol}, \quad (K.24)$$

$$n_{CO_2,dissolved} = (0.6370)(30.0 \times 10^{-6}) = 1.9110 \times 10^{-5} \text{ mol}, \quad (K.25)$$

Accounting for CO_2 speciation

CO_2 exists in multiple forms when dissolved. Firstly, hydration of the dissolved CO_2 occurs forming carbonic acid ($CO_{2(aq)} + H_2O \rightleftharpoons H_2CO_3$), while this step is necessary, H_2CO_3 is short-lived and present in very low concentrations such that it is generally ignored in carbon calculations. Following this step, two dissociation reactions occur. The first is the dissociation from carbonic acid to bicarbonate ($H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$). The second is the dissociation from bicarbonate to carbonate ($HCO_3^- \rightleftharpoons H^+ + CO_3^{2-}$). The amount of CO_2 dissociated into these forms is dependent on the pH of the liquid. At very low pHs no CO_2 dissociates and thus these reaction can be ignored. At a pH < 7.5 only the first reaction is considered. For our samples pH is observed between 6.8 - 7.0 and thus the majority of dissolved CO_2 will be in the form of HCO_3^- and therefore must be accounted for. The proportion quantity of HCO_3^- can be calculated using the acid dissociation constant. This is given for HCO_3^- at $25^\circ C$ as $pK_{a1} = 6.35$ or $K_{a1} = 10^{-6.35}$ where $pK_a = -\log_{10}(K_a)$. To adjust this constant based on temperature the Van't Hoff equation can then be used.

From initial measurements:

$$K_{a1,T2} = K_{a1,T1} \times \exp\left(-\frac{\Delta H^\circ}{R} \left(\frac{1}{T_2} - \frac{1}{T_1}\right)\right), \quad (K.26a)$$

$$K_{a1,T2} = 10^{-6.35} \times \exp\left(-\frac{9151}{8.314} \left(\frac{1}{295.15} - \frac{1}{298.15}\right)\right), \quad (K.26b)$$

$$K_{a1,T2} = 4.3023 \times 10^{-7} \text{ mol/L}, \quad (K.26c)$$

From final measurements:

$$K_{a1,T2} = K_{a1,T1} \times \exp\left(-\frac{\Delta H^\circ}{R} \left(\frac{1}{T_2} - \frac{1}{T_1}\right)\right), \quad (K.27a)$$

$$K_{a1,T2} = 10^{-6.35} \times \exp\left(-\frac{9151}{8.314} \left(\frac{1}{303.15} - \frac{1}{298.15}\right)\right), \quad (K.27b)$$

$$K_{a1,T2} = 4.7473 \times 10^{-7} \text{ mol/L}, \quad (K.27c)$$

To find the number of moles of HCO_3^- we now use the relationship:

$$K_{a1} = \frac{[H^+][HCO_3^-]}{[H_2CO_3]}, \quad (K.28)$$

where $H^+ = 10^{-pH}$ and $H_2CO_3 \approx CO_{2(aq)}$.

From initial measurements:

$$n_{HCO_3^-,dissolved} = \frac{(4.3023 \times 10^{-7})(0.1119 \times 10^{-5})}{10^{-7.00}} = 0.4813 \times 10^{-5} mol \quad (K.29)$$

From final measurements:

$$n_{HCO_3^-,dissolved} = \frac{(4.7473 \times 10^{-7})(1.9110 \times 10^{-5})}{10^{-6.82}} = 5.9939 \times 10^{-5} mol \quad (K.30)$$

Total number of moles

To calculate the total number of moles we add the moles in the headspace to the moles dissolved. It is important to note that for the initial measurements headspace samples are taken directly after the addition of CH_4 and, as such, the assumption is made that not enough time has passed for CH_4 to dissolve.

From initial measurements:

$$n_{CH_4,total} = n_{CH_4,headspace} = 16.36 \times 10^{-5} mol, \quad (K.31)$$

$$n_{CO_2,total} = n_{CO_2,headspace} + n_{CO_2,dissolved} + n_{HCO_3^-,dissolved} = 0.7703 \times 10^{-5} mol \quad (K.32)$$

From final measurements:

$$n_{CH_4,total} = n_{CH_4,headspace} + n_{CH_4,dissolved} = 0.0000 mol, \quad (K.33)$$

$$n_{CO_2,total} = n_{CO_2,headspace} + n_{CO_2,dissolved} + n_{HCO_3^-,dissolved} = 11.1749 \times 10^{-5} mol \quad (K.34)$$

Appendix L

Calculation of catabolic heat

Enthalpy of reactions

The enthalpy of reaction is calculated for both the full oxidation of methane to carbon dioxide and the oxidation of methane to formaldehyde. These are calculated from the enthalpy of formation of products and reactants as provided below. The stoichiometric equations related to the biological oxidation of methane are as follows:

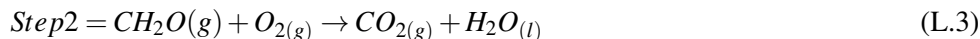
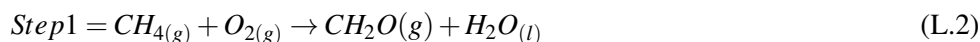
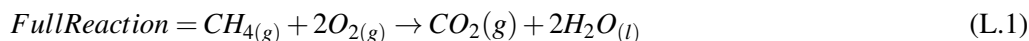


Table L.1: Enthalpy of formation of molecules associated with the biological oxidation of methane. Values are obtained from The US National Institute of Standards and Technology (NIST).

Quantity	Molecule	Value (kJ/mol)
$\Delta_f H_{gas}^\circ$	CH_4	-74.9
$\Delta_f H_{gas}^\circ$	O_2	0.0
$\Delta_f H_{gas}^\circ$	CH_2O	-115.9
$\Delta_f H_{gas}^\circ$	CO_2	-393.5
$\Delta_f H_{liquid}^\circ$	H_2O	-285.8

For the first step in the oxidation of CH_4 , the enthalpy of reaction is calculated to be -326.9 kJ/mol . For the second step, given by Equation 3, the enthalpy of reaction is calculated as -563.4 kJ/mol . The full oxidation of CH_4 , given by Equation 1, is therefore calculated as -890.3 kJ/mol . This is inline with the standard enthalpy of combustion of methane given as -890.71 kJ/mol (NIST).

Maximum amount of heat available from catabolism

Here the maximum amount of heat available (Q_{cat}) is calculated by multiplying the enthalpy of combustion of methane by the number of moles of CO_2 produced within each bottle. The remaining moles of CH_4 are assumed to be converted to CH_2O , where the change in enthalpy of this reaction is previously calculated as $-326.9 kJ/mol$. So for CAL1 (*M. methanica* S1, 8%v/v CH_4 in air) we measure that in total 1.6765×10^{-4} moles of CH_4 are consumed while only 1.1053×10^{-4} moles of CO_2 are produced. Firstly we calculate the heat theoretically generated by production of CO_2 which follows:

$$Q_{CO_2} = 1.1053 \times 10^{-4} \times 890710 = 98.45 J. \quad (L.4)$$

The next step is to calculate the heat generated from the oxidation of CH_4 to CH_2O for the remaining moles of CH_4 consumed. This follows:

$$Q_{CH_2O} = (1.6765 \times 10^{-4} - 1.1053 \times 10^{-4}) \times 326900 = 18.67 J. \quad (L.5)$$

So the final amount of maximum heat generated through the catabolic pathway is:

$$Q_{cat} = 98.45 + 18.67 = 117.12 J. \quad (L.6)$$

Bibliography

- A., E. (1913). Standards and Tests for Sewage and Sewage Effluents. *Nature*, 91(2264), 61–62. <https://doi.org/10.1038/091061a0>
- Abeyasinghe, W., & Punchi-Manage, R. (2020). Beyond neutrality: Adding habitat filtering to neutral models. *Ecological Processes*, 9(1), 27. <https://doi.org/10.1186/s13717-020-00228-7>
- Adair, K. L., Wilson, M., Bost, A., & Douglas, A. E. (2018). Microbial community assembly in wild populations of the fruit fly *Drosophila melanogaster*. *The ISME Journal*, 12(4), 959–972. <https://doi.org/10.1038/s41396-017-0020-x>
- Adler, P. B., HilleRisLambers, J., & Levine, J. M. (2007). A niche for neutrality. *Ecology Letters*, 10(2), 95–104. <https://doi.org/10.1111/j.1461-0248.2006.00996.x>
- Afan, H. A., Mohtar, W. H. M. W., Khaleel, F., Kamel, A. H., Mansoor, S. S., Alsultani, R., Ahmed, A. N., Sherif, M., & El-Shafie, A. (2024). Data-driven water quality prediction for wastewater treatment plants. *Heliyon*, 10(18). <https://doi.org/10.1016/j.heliyon.2024.e36940>
- Ahammad, S. Z., Graham, D. W., & Dolfing, J. (2020). Wastewater Treatment: Biological. In *Managing Water Resources and Hydrological Systems* (2nd ed.). CRC Press.
- Aiello, A. E., Larson, E. L., & Sedlak, R. (2008). Hidden heroes of the health revolution Sanitation and personal hygiene. *American Journal of Infection Control*, 36(10), S128–S151. <https://doi.org/10.1016/j.ajic.2008.09.008>
- Ali, M., Hong, P.-Y., Mishra, H., Vrouwenvelder, J., & Saikaly, P. E. (2022). Adopting the circular model: Opportunities and challenges of transforming wastewater treatment plants into resource recovery factories in Saudi Arabia. *Water Reuse*, 12(3), 346–365. <https://doi.org/10.2166/wrd.2022.038>
- Alonso, D., Etienne, R. S., & McKane, A. J. (2006). The merits of neutral theory. *Trends in Ecology & Evolution*, 21(8), 451–457. <https://doi.org/10.1016/j.tree.2006.03.019>
- AlSayed, A., Fergala, A., & Eldyasti, A. (2018). Sustainable biogas mitigation and value-added resources recovery using methanotrophs intergrated into wastewater treatment plants. *Reviews in Environmental Science and Bio/Technology*, 17(2), 351–393. <https://doi.org/10.1007/s11157-018-9464-3>
- Amaral, J., Archambault, C., Richards, S., & Knowles, R. (1995). Denitrification associated with Groups I and II methanotrophs in a gradient enrichment system. *FEMS Microbiology Ecology*, 18(4), 289–298. <https://doi.org/10.1111/j.1574-6941.1995.tb00185.x>
- Amaral, J. A., & Knowles, R. (1995). Growth of methanotrophs in methane and oxygen counter gradients. *FEMS Microbiology Letters*, 126(3), 215–220. [https://doi.org/10.1016/0378-1097\(95\)00012-T](https://doi.org/10.1016/0378-1097(95)00012-T)

- Amerasinghe, P. H., Bhardwaj, R. M., Scott, C., Jella, K., & Marshall, F. (2013). *Urban wastewater and agricultural reuse challenges in India*. International Water Management Institute. Retrieved January 9, 2025, from <https://hdl.handle.net/10568/39932>
- Angelakis, A. N., Capodaglio, A. G., & Dialynas, E. G. (2023). Wastewater Management: From Ancient Greece to Modern Times and Future. *Water*, 15(1), 43. <https://doi.org/10.3390/w15010043>
- Angelakis, A. N., De Feo, G., Laureano, P., & Zourou, A. (2013). Minoan and Etruscan Hydro-Technologies. *Water*, 5(3), 972–987. <https://doi.org/10.3390/w5030972>
- Angelakis, A. N., Mays, L. W., Koutsoyiannis, D., & Mamassis, N. (2012). *Evolution of Water Supply Through the Millennia* (Vol. 11). IWA Publishing. <https://iwaponline.com/ebooks/book/472/Evolution-of-Water-Supply-Through-the-Millennia>
- Angelakis, A. N., & Snyder, S. A. (2015). Wastewater Treatment and Reuse: Past, Present, and Future. *Water*, 7(9), 4887–4895. <https://doi.org/10.3390/w7094887>
- Anthony, C. (1978). The Prediction of Growth Yields in Methylophs. *Microbiology*, 104(1), 91–104. <https://doi.org/10.1099/00221287-104-1-91>
- Ardern, E., & Lockett, W. T. (1914). Experiments on the oxidation of sewage without the aid of filters. *Journal of the Society of Chemical Industry*, 33(10), 523–539. <https://doi.org/10.1002/jctb.5000331005>
- Asante-Sackey, D., Rathilal, S., Tetteh, E. K., & Armah, E. K. (2022). Membrane Bioreactors for Produced Water Treatment: A Mini-Review. *Membranes*, 12(3), 275. <https://doi.org/10.3390/membranes12030275>
- Asheghmoalla, M., & Mehrvar, M. (2024). Integrated and Hybrid Processes for the Treatment of Actual Wastewaters Containing Micropollutants: A Review on Recent Advances. *Processes*, 12(2), 339. <https://doi.org/10.3390/pr12020339>
- Ayarza, J. M., & Erijman, L. (2011). Balance of Neutral and Deterministic Components in the Dynamics of Activated Sludge Floc Assembly. *Microbial Ecology*, 61(3), 486–495. <https://doi.org/10.1007/s00248-010-9762-y>
- Bacaër, N. (2011a). Fisher and natural selection (1922). In N. Bacaër (Ed.), *A Short History of Mathematical Population Dynamics* (pp. 77–80). Springer. https://doi.org/10.1007/978-0-85729-115-8_14
- Bacaër, N. (2011b). The Hardy–Weinberg law (1908). In N. Bacaër (Ed.), *A Short History of Mathematical Population Dynamics* (pp. 59–63). Springer. https://doi.org/10.1007/978-0-85729-115-8_11
- Bacaër, N. (2011c). Lotka and stable population theory (1907–1911). In N. Bacaër (Ed.), *A Short History of Mathematical Population Dynamics* (pp. 55–58). Springer. https://doi.org/10.1007/978-0-85729-115-8_10
- Bacaër, N. (2011d). Malthus and the obstacles to geometric growth (1798). In N. Bacaër (Ed.), *A Short History of Mathematical Population Dynamics* (pp. 31–33). Springer. https://doi.org/10.1007/978-0-85729-115-8_5
- Bacaër, N. (2011e). Wright and random genetic drift (1931). In N. Bacaër (Ed.), *A Short History of Mathematical Population Dynamics* (pp. 105–109). Springer. https://doi.org/10.1007/978-0-85729-115-8_19
- Bannon, C. (2017). Fresh Water in Roman Law: Rights and Policy. *The Journal of Roman Studies*, 107, 60–89. <https://doi.org/10.1017/S007543581700079X>

- Barberán, A., Casamayor, E. O., & Fierer, N. (2014). The microbial contribution to macroecology. *Frontiers in Microbiology*, 5. <https://doi.org/10.3389/fmicb.2014.00203>
- Barnard, J. (1973). Biological Denitrification. *Water Pollution Control*, 72, 705–720.
- Barnard, J. (1974). Cut P and N Without Chemicals. *WATER WASTES ENGNG*, 11.
- Barnard, J. (1975). Biological nutrient removal without addition of chemicals. *Water Research - WATER RES*, 9, 485–490. [https://doi.org/10.1016/0043-1354\(75\)90072-X](https://doi.org/10.1016/0043-1354(75)90072-X)
- Battley, E. H. (1960a). Growth-Reaction Equations for *Saccharomyces cerevisiae*. *Physiologia Plantarum*, 13(2), 192–203. <https://doi.org/10.1111/j.1399-3054.1960.tb08023.x>
- Battley, E. H. (1960b). A Theoretical Approach to the Study of the Thermodynamics of Growth of *Saccharomyces cerevisiae* (Hansen). *Physiologia Plantarum*, 13(4), 674–686. <https://doi.org/10.1111/j.1399-3054.1960.tb08090.x>
- Beal, C. D., Gardner, E. A., & Menzies, N. W. (2005). Process, performance, and pollution potential: A review of septic tank–soil absorption systems. *Soil Research*, 43(7), 781–802. <https://doi.org/10.1071/SR05018>
- Beeravolu, C., Coutron, P., Péliissier, R., & Munoz, F. (2009). Studying ecological communities from a neutral standpoint: A review of models' structure and parameter estimation. *Ecological Modelling*, 2603–2610. <https://doi.org/10.1016/j.ecolmodel.2009.06.041>
- Bell, G. (2000). The Distribution of Abundance in Neutral Communities. *The American Naturalist*, 155(5), 606–617. <https://doi.org/10.1086/303345>
- Bell, G. (2001). Neutral Macroecology. *Science*, 293(5539), 2413–2418. <https://doi.org/10.1126/science.293.5539.2413>
- Benotti, M. J., Stanford, B. D., & Snyder, S. A. (2010). Impact of Drought on Wastewater Contaminants in an Urban Water Supply. *Journal of Environmental Quality*, 39(4), 1196–1200. <https://doi.org/10.2134/jeq2009.0072>
- Birou, B., Marison, I. W., & Stockar, U. V. (1987). Calorimetric investigation of aerobic fermentations. *Biotechnology and Bioengineering*, 30(5), 650–660. <https://doi.org/10.1002/bit.260300509>
- Blake, D. R., Mayer, E. W., Tyler, S. C., Makide, Y., Montague, D. C., & Rowland, F. S. (1982). Global increase in atmospheric methane concentrations between 1978 and 1980. *Geophysical Research Letters*, 9(4), 477–480. <https://doi.org/10.1029/GL009i004p00477>
- Boadi, K., Kuitunen, M., Raheem, K., & Hanninen, K. (2005). Urbanisation Without Development: Environmental and Health Implications in African Cities. *Environment, Development and Sustainability*, 7(4), 465–500. <https://doi.org/10.1007/s10668-004-5410-3>
- Bojanowski, D., Orlńska-Woźniak, P., Wilk, P., & Szalińska, E. (2022). Estimation of nutrient loads with the use of mass-balance and modelling approaches on the Wełna River catchment example (central Poland). *Scientific Reports*, 12(1), 13052. <https://doi.org/10.1038/s41598-022-17270-4>
- Boretti, A., & Rosa, L. (2019). Reassessing the projections of the World Water Development Report. *npj Clean Water*, 2(1), 1–6. <https://doi.org/10.1038/s41545-019-0039-9>
- Bowman, J. (2006). The Methanotrophs — The Families Methylococcaceae and Methylocystaceae. In M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, & E. Stackebrandt (Eds.), *The Prokaryotes: Volume*

- 5: *Proteobacteria: Alpha and Beta Subclasses* (pp. 266–289). Springer. https://doi.org/10.1007/0-387-30745-1_15
- Brandt, E. M. F., Duarte, F. V., Vieira, J. P. R., Melo, V. M., Souza, C. L., Araújo, J. C., & Chernicharo, C. A. L. (2016). The use of novel packing material for improving methane oxidation in biofilters. *Journal of Environmental Management*, 182, 412–420. <https://doi.org/10.1016/j.jenvman.2016.07.075>
- Brasseur, G. P., & Solomon, S. (2005). Composition and Chemistry. In *Aeronomy of the Middle Atmosphere: Chemistry and Physics of the Stratosphere and Mesosphere* (pp. 265–442). Springer Netherlands. https://doi.org/10.1007/1-4020-3824-0_5
- Brown, L. R., Strawinski, R. J., & McCleskey, C. S. (1964). The isolation and characterization of methanomonas methanooxidans brown and strawinski. *Canadian Journal of Microbiology*, 10(5), 791–799. <https://doi.org/10.1139/m64-100>
- Brown, T. C., Mahat, V., & Ramirez, J. A. (2019). Adaptation to Future Water Shortages in the United States Caused by Population Growth and Climate Change. *Earth's Future*, 7(3), 219–234. <https://doi.org/10.1029/2018EF001091>
- Bryson, V., & Vogel, H. J. (1965). Evolving Genes and Proteins. *Science*, 147(3653), 68–71. <https://doi.org/10.1126/science.147.3653.68>
- BSI. (1983, April). British Standard Code of practice for Design and installation of small sewage treatment works and cesspools.
- Building Standards Division Directorate for Housing. (2024, November). Building Standards Division Domestic Technical Handbook January 2025. <https://www.gov.scot/publications/building-standards-technical-handbook-january-2025-domestic/>
- Bull, I. D., Parekh, N. R., Hall, G. H., Ineson, P., & Evershed, R. P. (2000). Detection and classification of atmospheric methane oxidizing bacteria in soil. *Nature*, 405(6783), 175–178. <https://doi.org/10.1038/35012061>
- Burian, S., & Edwards, F. (2002). Historical Perspectives of Urban Drainage. *Global Solutions for Urban Drainage*. [https://doi.org/10.1061/40644\(2002\)284](https://doi.org/10.1061/40644(2002)284)
- Burns, A. R., Stephens, W. Z., Stagaman, K., Wong, S., Rawls, J. F., Guillemin, K., & Bohannan, B. J. (2016). Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *The ISME Journal*, 10(3), 655–664. <https://doi.org/10.1038/ismej.2015.142>
- Buswell, A. M., & Long, H. L. (1923). Microbiology and Theory of Activated Sludge. *Journal (American Water Works Association)*, 10(2), 309–321. Retrieved January 6, 2025, from <https://www.jstor.org/stable/41224985>
- Büttner, O., Jawitz, J. W., Birk, S., & Borchardt, D. (2022). Why wastewater treatment fails to protect stream ecosystems in Europe. *Water Research*, 217, 118382. <https://doi.org/10.1016/j.watres.2022.118382>
- Canter, L. W., & Knox, R. C. (1985). *Septic Tank System Effects on Ground Water Quality*. Routledge. <https://doi.org/10.1201/9780203739877>
- Capodaglio, A. G. (2017). Integrated, Decentralized Wastewater Management for Resource Recovery in Rural and Peri-Urban Areas. *Resources*, 6(2), 22. <https://doi.org/10.3390/resources6020022>

- Carballa, M., Regueiro, L., & Lema, J. M. (2015). Microbial management of anaerobic digestion: Exploiting the microbiome-functionality nexus. *Current Opinion in Biotechnology*, 33, 103–111. <https://doi.org/10.1016/j.copbio.2015.01.008>
- Caswell, H. (1976). Community Structure: A Neutral Model Analysis. *Ecological Monographs*, 46(3), 327–354. <https://doi.org/10.2307/1942257>
- Centola, D., Becker, J., Brackbill, D., & Baronchelli, A. (2018). Experimental evidence for tipping points in social convention. *Science*, 360(6393), 1116–1119. <https://doi.org/10.1126/science.aas8827>
- Charles, K., Ashbolt, N., Roser, D., McGuinness, R., & Deere, D. (2005). Effluent quality from 200 on-site sewage systems: Design values for guidelines. *Water Science and Technology*, 51(10), 163–169. <https://doi.org/10.2166/wst.2005.0363>
- Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, 10(3), 195–205. <https://doi.org/10.1038/nrg2526>
- Chave, J., & Leigh, E. G. (2002). A spatially explicit neutral model of beta-diversity in tropical forests. *Theoretical Population Biology*, 62(2), 153–168. <https://doi.org/10.1006/tpbi.2002.1597>
- Chen, H., Wang, M., & Chang, S. (2020). Disentangling Community Structure of Ecological System in Activated Sludge: Core Communities, Functionality, and Functional Redundancy. *Microbial Ecology*, 80(2), 296–308. <https://doi.org/10.1007/s00248-020-01492-y>
- Chen, W., Wei, J., Su, Z., Wu, L., Liu, M., Huang, X., Yao, P., & Wen, D. (2022). Deterministic mechanisms drive bacterial communities assembly in industrial wastewater treatment system. *Environment International*, 168, 107486. <https://doi.org/10.1016/j.envint.2022.107486>
- Cheng, H., Cheng, D., Mao, J., Lu, T., & Hong, P.-Y. (2019). Identification and characterization of core sludge and biofilm microbiota in anaerobic membrane bioreactors. *Environment International*, 133, 105165. <https://doi.org/10.1016/j.envint.2019.105165>
- Cheng, H., Monjed, M. K., Myshkevych, Y., Wang, T., & Hong, P.-Y. (2024). Accounting for the microbial assembly of each process in wastewater treatment plants (WWTPs): Study of four WWTPs receiving similar influent streams. *Applied and Environmental Microbiology*, 90(4), e02253–23. <https://doi.org/10.1128/aem.02253-23>
- Chesson, P. (1991). A need for niches? *Trends in Ecology & Evolution*, 6(1), 26–28. [https://doi.org/10.1016/0169-5347\(91\)90144-M](https://doi.org/10.1016/0169-5347(91)90144-M)
- Childe, V. G. (1931). Skara Brae: A ‘Stone Age’ Village in Orkney. *Antiquity*, 5(17), 47–59. <https://doi.org/10.1017/S0003598X00005597>
- Chisholm, R. A., & O’Dwyer, J. P. (2014). Species ages in neutral biodiversity models. *Theoretical Population Biology*, 93, 85–94. <https://doi.org/10.1016/j.tpb.2014.02.002>
- Chisholm, R. A., & Pacala, S. W. (2010). Niche and neutral models predict asymptotically equivalent species abundance distributions in high-diversity ecological communities. *Proceedings of the National Academy of Sciences*, 107(36), 15821–15825. <https://doi.org/10.1073/pnas.1009387107>
- Choi, D.-W., Kunz, R. C., Boyd, E. S., Semrau, J. D., Antholine, W. E., Han, J.-I., Zahn, J. A., Boyd, J. M., de la Mora, A. M., & DiSpirito, A. A. (2003). The Membrane-Associated Methane Monooxygenase

- (pMMO) and pMMO-NADH:Quinone Oxidoreductase Complex from *Methylococcus capsulatus* Bath. *Journal of Bacteriology*, 185(19), 5755–5764. <https://doi.org/10.1128/jb.185.19.5755-5764.2003>
- Chu, F., & Lidstrom, M. E. (2016). XoxF Acts as the Predominant Methanol Dehydrogenase in the Type I Methanotroph *Methylobacterium buryatense*. *Journal of Bacteriology*, 198(8), 1317–1325. <https://doi.org/10.1128/jb.00959-15>
- Chubukov, V., Gerosa, L., Kochanowski, K., & Sauer, U. (2014). Coordination of microbial metabolism. *Nature Reviews Microbiology*, 12(5), 327–340. <https://doi.org/10.1038/nrmicro3238>
- Ciecielnski, N. J. (2013). The Stench of Disease: Public Health and the Environment in Late-Medieval English towns and cities. *Health, Culture and Society*, 4(1), 91–104. <https://doi.org/10.5195/hcs.2013.114>
- Cira, N. J., Pearce, M. T., & Quake, S. R. (2018). Neutral and selective dynamics in a synthetic microbial community. *Proceedings of the National Academy of Sciences*, 115(42), E9842–E9848. <https://doi.org/10.1073/pnas.1808118115>
- Clark, J. S. (2009). Beyond neutral science. *Trends in Ecology & Evolution*, 24(1), 8–15. <https://doi.org/10.1016/j.tree.2008.09.004>
- Clark, J. S. (2012). The coherence problem with the Unified Neutral Theory of Biodiversity. *Trends in Ecology & Evolution*, 27(4), 198–202. <https://doi.org/10.1016/j.tree.2012.02.001>
- Clausius, R. (1867). *The Mechanical Theory of Heat: With Its Applications to the Steam-engine and to the Physical Properties of Bodies*. J. Van Voorst. Retrieved January 16, 2025, from <http://archive.org/details/mechanicaltheor04claugoog>
- Condit, R., Pitman, N., Leigh, E. G., Chave, J., Terborgh, J., Foster, R. B., Núñez, P., Aguilar, S., Valencia, R., Villa, G., Muller-Landau, H. C., Losos, E., & Hubbell, S. P. (2002). Beta-Diversity in Tropical Forest Trees. *Science*, 295(5555), 666–669. <https://doi.org/10.1126/science.1066854>
- Connell, J. H. (1978). Diversity in Tropical Rain Forests and Coral Reefs. *Science*, 199(4335), 1302–1310. Retrieved September 30, 2024, from <https://www.jstor.org/stable/1745369>
- Connell, J. H. (1980). Diversity and the Coevolution of Competitors, or the Ghost of Competition Past. *Oikos*, 35(2), 131–138. <https://doi.org/10.2307/3544421>
- Connelly, S., Pussayanavin, T., J. Randle-Boggis, R., Wicheansan, A., Jampathong, S., Keating, C., Z. Ijaz, U., T. Sloan, W., & Koottatep, T. (2019). Solar Septic Tank: Next Generation Sequencing Reveals Effluent Microbial Community Composition as a Useful Index of System Performance. *Water*, 11(12), 2660. <https://doi.org/10.3390/w11122660>
- Connor, E. F., & Simberloff, D. (1979). The Assembly of Species Communities: Chance or Competition? *Ecology*, 60(6), 1132–1140. <https://doi.org/10.2307/1936961>
- Conrad, R., & Rothfuss, F. (1991). Methane oxidation in the soil surface layer of a flooded rice field and the effect of ammonium. *Biology and Fertility of Soils*, 12(1), 28–32. <https://doi.org/10.1007/BF00369384>
- Cooper, J. A., Loomis, G. W., & Amador, J. A. (2016). Hell and High Water: Diminished Septic System Performance in Coastal Regions Due to Climate Change. *PLOS ONE*, 11(9), e0162104. <https://doi.org/10.1371/journal.pone.0162104>

- Cooper, P. (2005, February). Historical aspects of wastewater treatment. In *Decentralised Sanitation and Reuse: Concepts, Systems and Implementation*. IWA Publishing. <https://doi.org/10.2166/9781780402949>
- Corcoran, E., Nellemann, C., Baker, E., Bos, R., Osborn, D., & Savelli, H. (2010). *Sick Water? The central role of wastewater management in sustainable development. A Rapid Response Assessment*. United Nations Environment Programme. <https://wedocs.unep.org/20.500.11822/9156>
- Cossio, C., McConville, J., Rauch, S., Wilén, B.-M., Dalahmeh, S., Mercado, A., & Romero, A. M. (2018). Wastewater management in small towns – understanding the failure of small treatment plants in Bolivia. *Environmental Technology*, 39(11), 1393–1403. <https://doi.org/10.1080/09593330.2017.1330364>
- Costa, J. E., Heufelder, G., Foss, S., Milham, N. P., & Howes, B. (2002). Nitrogen removal efficiencies of three alternative septic technologies and a conventional septic system. *Environment Cape Cod*, 5(1), 15–24.
- Cotteral, J. A., & Norris, D. (1969). Septic Tank Systems. *ASCE Journal of the Sanitary Engineering Division Proceedings*, 95(SA4), 715–746.
- Cramer, J. S. (2004). The early origins of the logit model. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 35(4), 613–626. <https://doi.org/10.1016/j.shpsc.2004.09.003>
- Curtis, T., & Fowler, J. (2023). Microbial Ecology for Engineering Biology. *Interface Focus*, 13(4), 20230027. <https://doi.org/10.1098/rsfs.2023.0027>
- Curtis, T. P., Head, I. M., & Graham, D. W. (2003). Theoretical Ecology for Engineering Biology. *Environmental Science & Technology*, 37(3), 64A–70A. <https://doi.org/10.1021/es0323493>
- Curtis, T. P., Sloan, W. T., & Scannell, J. W. (2002). Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences*, 99(16), 10494–10499. <https://doi.org/10.1073/pnas.142680199>
- Dasgupta, S., Agarwal, N., & Mukherjee, A. (2021). Moving up the On-Site Sanitation ladder in urban India through better systems and standards. *Journal of Environmental Management*, 280, 111656. <https://doi.org/10.1016/j.jenvman.2020.111656>
- Dawadi, S., & Ahmad, S. (2013). Evaluating the impact of demand-side management on water resources under changing climatic conditions and increasing population. *Journal of Environmental Management*, 114, 261–275. <https://doi.org/10.1016/j.jenvman.2012.10.015>
- de Aguiar, M. a. M., Baranger, M., Baptestini, E. M., Kaufman, L., & Bar-Yam, Y. (2009). Global patterns of speciation and diversity. *Nature*, 460(7253), 384–387. <https://doi.org/10.1038/nature08168>
- De Feo, G., Antoniou, G., Fardin, H. F., El-Gohary, F., Zheng, X. Y., Reklaityte, I., Butler, D., Yannopoulos, S., & Angelakis, A. N. (2014). The Historical Development of Sewers Worldwide. *Sustainability*, 6(6), 3936–3974. <https://doi.org/10.3390/su6063936>
- De Stefano, L., Petersen-Perlman, J. D., Sproles, E. A., Eynard, J., & Wolf, A. T. (2017). Assessment of transboundary river basins for potential hydro-political tensions. *Global Environmental Change*, 45, 35–46. <https://doi.org/10.1016/j.gloenvcha.2017.04.008>
- De Visscher, A., & Van Cleemput, O. (2003). Simulation model for gas diffusion and methane oxidation in landfill cover soils. *Waste Management*, 23(7), 581–591. [https://doi.org/10.1016/S0956-053X\(03\)00096-5](https://doi.org/10.1016/S0956-053X(03)00096-5)

- Demirel, Y., & Sandler, S. I. (2002). Thermodynamics and bioenergetics. *Biophysical Chemistry*, 97(2), 87–111. [https://doi.org/10.1016/S0301-4622\(02\)00069-8](https://doi.org/10.1016/S0301-4622(02)00069-8)
- Diamond, J. M. (1975). Assembly of Species Communities. In *Ecology and Evolution of Communities* (pp. 342–444). Harvard University Press, Cambridge, Massachusetts. <https://archive.org/details/diamond-1975>
- Diaz-Valbuena, L. R., Leverenz, H. L., Cappa, C. D., Tchobanoglous, G., Horwath, W. R., & Darby, J. L. (2011). Methane, Carbon Dioxide, and Nitrous Oxide Emissions from Septic Tank Systems. *Environmental Science & Technology*, 45(7), 2741–2747. <https://doi.org/10.1021/es1036095>
- Dlugokencky, E. J., Houweling, S., Bruhwiler, L., Masarie, K. A., Lang, P. M., Miller, J. B., & Tans, P. P. (2003). Atmospheric methane levels off: Temporary pause or a new steady-state? *Geophysical Research Letters*, 30(19). <https://doi.org/10.1029/2003GL018126>
- Dolan, F., Lamontagne, J., Link, R., Hejazi, M., Reed, P., & Edmonds, J. (2021). Evaluating the economic impact of water scarcity in a changing world. *Nature Communications*, 12(1), 1915. <https://doi.org/10.1038/s41467-021-22194-0>
- Dornelas, M., Connolly, S. R., & Hughes, T. P. (2006). Coral reef diversity refutes the neutral theory of biodiversity. *Nature*, 440(7080), 80–82. <https://doi.org/10.1038/nature04534>
- du Plessis, A. (2022). Persistent degradation: Global water quality challenges and required actions. *One Earth*, 5(2), 129–131. <https://doi.org/10.1016/j.oneear.2022.01.005>
- Duarte, M. S., Martins, G., Oliveira, P., Fernandes, B., Ferreira, E. C., Alves, M. M., Lopes, F., Pereira, M. A., & Novais, P. (2024). A Review of Computational Modeling in Wastewater Treatment Processes. *ACS ES&T Water*, 4(3), 784–804. <https://doi.org/10.1021/acsestwater.3c00117>
- Duivenvoorden, J. F., Svenning, J.-C., & Wright, S. J. (2002). Beta Diversity in Tropical Forests. *Science*, 295(5555), 636–637. <https://doi.org/10.1126/science.295.5555.636>
- Dunfield, P. F., Khmelenina, V. N., Suzina, N. E., Trotsenko, Y. A., & Dedys, S. N. (2003). *Methylocella silvestris* sp. nov., a novel methanotroph isolated from an acidic forest cambisol. *International Journal of Systematic and Evolutionary Microbiology*, 53(5), 1231–1239. <https://doi.org/10.1099/ijs.0.02481-0>
- Dunfield, P. F., Yuryev, A., Senin, P., Smirnova, A. V., Stott, M. B., Hou, S., Ly, B., Saw, J. H., Zhou, Z., Ren, Y., Wang, J., Mountain, B. W., Crowe, M. A., Weatherby, T. M., Bodelier, P. L. E., Liesack, W., Feng, L., Wang, L., & Alam, M. (2007). Methane oxidation by an extremely acidophilic bacterium of the phylum Verrucomicrobia. *Nature*, 450(7171), 879–882. <https://doi.org/10.1038/nature06411>
- Duque, N., Scholten, L., & Maurer, M. (2024). When does infrastructure hybridisation outperform centralised infrastructure paradigms? – Exploring economic and hydraulic impacts of decentralised urban wastewater system expansion. *Water Research*, 254, 121327. <https://doi.org/10.1016/j.watres.2024.121327>
- Dworkin, M., & Foster, J. W. (1956). Studies on pseudomonas methanica (söhnngen) nov. comb. *Journal of Bacteriology*, 72(5), 646–659. <https://doi.org/10.1128/jb.72.5.646-659.1956>
- Eggimann, S., Truffer, B., & Maurer, M. (2015). To connect or not to connect? Modelling the optimal degree of centralisation for wastewater infrastructures. *Water Research*, 84, 218–231. <https://doi.org/10.1016/j.watres.2015.07.004>

- Enquist, B. J., Sanderson, J., & Weiser, M. D. (2002). Modeling Macroscopic Patterns in Ecology. *Science*, 295(5561), 1835–1837. <https://doi.org/10.1126/science.295.5561.1835c>
- Ercin, A. E., & Hoekstra, A. Y. (2014). Water footprint scenarios for 2050: A global analysis. *Environment International*, 64, 71–82. <https://doi.org/10.1016/j.envint.2013.11.019>
- Etheridge, D. M., Steele, L. P., Francey, R. J., & Langenfelds, R. L. (1998). Atmospheric methane between 1000 A.D. and present: Evidence of anthropogenic emissions and climatic variability. *Journal of Geophysical Research: Atmospheres*, 103(D13), 15979–15993. <https://doi.org/10.1029/98JD00923>
- Etienne, R. S. (2005). A new sampling formula for neutral biodiversity. *Ecology Letters*, 8(3), 253–260. <https://doi.org/10.1111/j.1461-0248.2004.00717.x>
- Etienne, R. S. (2009). Maximum likelihood estimation of neutral model parameters for multiple samples with different degrees of dispersal limitation. *Journal of Theoretical Biology*, 257(3), 510–514. <https://doi.org/10.1016/j.jtbi.2008.12.016>
- Etienne, R. S., & Alonso, D. (2007). Neutral Community Theory: How Stochasticity and Dispersal-Limitation Can Explain Species Coexistence. *Journal of Statistical Physics*, 128(1), 485–510. <https://doi.org/10.1007/s10955-006-9163-2>
- Etienne, R. S., Apol, M. E. F., Olff, H., & Weissing, F. J. (2007). Modes of speciation and the neutral theory of biodiversity. *Oikos*, 116(2), 241–258. <https://doi.org/10.1111/j.0030-1299.2007.15438.x>
- Etienne, R. S., & Haegeman, B. (2011). The neutral theory of biodiversity with random fission speciation. *Theoretical Ecology*, 4(1), 87–109. <https://doi.org/10.1007/s12080-010-0076-y>
- Etienne, R. S., Latimer, A. M., Silander, J. A., & Cowling, R. M. (2006). Comment on "Neutral Ecological Theory Reveals Isolation and Rapid Speciation in a Biodiversity Hot Spot". *Science*, 311(5761), 610–610. <https://doi.org/10.1126/science.1121914>
- Etienne, R. S., & Olff, H. (2004). A novel genealogical approach to neutral biodiversity theory. *Ecology Letters*, 7(3), 170–175. <https://doi.org/10.1111/j.1461-0248.2004.00572.x>
- Etienne, R. S., & Rosindell, J. (2011). The spatial limitations of current neutral models of biodiversity. *PloS One*, 6(3), e14717. <https://doi.org/10.1371/journal.pone.0014717>
- Fernandes, J., Ramísio, P. J., & Puga, H. (2024). A Comprehensive Review on Various Phases of Wastewater Technologies: Trends and Future Perspectives. *Eng*, 5(4), 2633–2661. <https://doi.org/10.3390/eng5040138>
- Fisher, R. A. (1923). XXI.—On the Dominance Ratio. *Proceedings of the Royal Society of Edinburgh*, 42, 321–341. <https://doi.org/10.1017/S0370164600023993>
- Fisher, R. A. (1930). *The Genetical Theory Of Natural Selection*. At The Clarendon Press. Retrieved January 20, 2025, from <http://archive.org/details/geneticaltheoryo031631mbp>
- Fjelsted, L., Scheutz, C., Christensen, A. G., Larsen, J. E., & Kjeldsen, P. (2020). Biofiltration of diluted landfill gas in an active loaded open-bed compost filter. *Waste Management*, 103, 1–11. <https://doi.org/10.1016/j.wasman.2019.12.005>

- Flörke, M., Kynast, E., Bärlund, I., Eisner, S., Wimmer, F., & Alcamo, J. (2013). Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development: A global simulation study. *Global Environmental Change*, 23(1), 144–156. <https://doi.org/10.1016/j.gloenvcha.2012.10.018>
- Flörke, M., Schneider, C., & McDonald, R. I. (2018). Water competition between cities and agriculture driven by climate change and urban growth. *Nature Sustainability*, 1(1), 51–58. <https://doi.org/10.1038/s41893-017-0006-8>
- Fodelianakis, S., Valenzuela-Cuevas, A., Barozzi, A., & Daffonchio, D. (2021). Direct quantification of ecological drift at the population level in synthetic bacterial communities. *The ISME Journal*, 15(1), 55–66. <https://doi.org/10.1038/s41396-020-00754-4>
- Foster, J. W., & Davis, R. H. (1966). A Methane-Dependent Coccus, with Notes on Classification and Nomenclature of Obligate, Methane-Utilizing Bacteria. *Journal of Bacteriology*, 91(5), 1924–1931. <https://doi.org/10.1128/jb.91.5.1924-1931.1966>
- Frenzel, P., Rothfuss, F., & Conrad, R. (1992). Oxygen profiles and methane turnover in a flooded rice microcosm. *Biology and Fertility of Soils*, 14(2), 84–89. <https://doi.org/10.1007/BF00336255>
- Fung, T., & Chisholm, R. A. (2023). Improving the realism of neutral ecological models by incorporating transient dynamics with temporal changes in community size. *Theoretical Population Biology*, 149, 12–26. <https://doi.org/10.1016/j.tpb.2022.12.001>
- García, J. L., & Galán, B. (2022). Integrating greenhouse gas capture and C1 biotechnology: A key challenge for circular economy. *Microbial Biotechnology*, 15(1), 228–239. <https://doi.org/10.1111/1751-7915.13991>
- Gause, G. F. (1934). *The struggle for existence*. Baltimore, The Williams & Wilkins company. <http://archive.org/details/struggleforexist00gauz>
- Ge, X., Yang, L., Sheets, J. P., Yu, Z., & Li, Y. (2014). Biological conversion of methane to liquid fuels: Status and opportunities. *Biotechnology Advances*, 32(8), 1460–1475. <https://doi.org/10.1016/j.biotechadv.2014.09.004>
- Gerosa, L., & Sauer, U. (2011). Regulation and control of metabolic fluxes in microbes. *Current Opinion in Biotechnology*, 22(4), 566–575. <https://doi.org/10.1016/j.copbio.2011.04.016>
- Gibbs, J. W. (1878). On the equilibrium of heterogeneous substances. *American Journal of Science*, s3-16(96), 441–458. <https://doi.org/10.2475/ajs.s3-16.96.441>
- Gilbert, B., Laurance, W. F., Leigh Jr., E. G., & Nascimento, H. E. M. (2006). Can Neutral Theory Predict the Responses of Amazonian Tree Communities to Forest Fragmentation? *The American Naturalist*, 168(3), 304–317. <https://doi.org/10.1086/506969>
- Gilbert, J. A., & Lynch, S. V. (2019). Community ecology as a framework for human microbiome research. *Nature Medicine*, 25(6), 884–889. <https://doi.org/10.1038/s41591-019-0464-9>
- Gill, G. (2000). Cholera and the Fight for Public Health Reform in Mid-Victorian England. *The Historical Association*. Retrieved December 4, 2024, from <https://www.history.org.uk/publications/resource/531/cholera-and-the-fight-for-public-health-reform-in>

- Gilpin, M. E., & Diamond, J. M. (1984). Are Species Co-occurrences on Islands Non-random, and Are Null Hypotheses Useful in Community Ecology? In D. R. Strong, D. Simberloff, L. G. Abele, & A. B. Thistle (Eds.), *Ecological Communities: Conceptual Issues and the Evidence* (pp. 297–315). Princeton University Press. <https://doi.org/10.1515/9781400857081.297>
- Gnaiger, E., Gellerich, F. N., & Wyss, M. (1994). *What is Controlling Life?: 50 Years After Erwin Schrödinger's What is Life?* Innsbruck University Press.
- Gómez-Borraz, T. L., González-Sánchez, A., Cabello, J., & Noyola, A. (2022). Model assessment on the non-isothermal methane biofiltration at ambient conditions. *Process Safety and Environmental Protection*, 163, 283–297. <https://doi.org/10.1016/j.psep.2022.05.042>
- Gómez-Borraz, T. L., Torres-Arévalo, Y. V., Cuetero-Martínez, Y., González-Sánchez, A., & Noyola, A. (2025). Assessment of temperature dynamics during methane oxidation in a pilot scale compost biofilter. *Biore-source Technology*, 419, 132097. <https://doi.org/10.1016/j.biortech.2025.132097>
- Gómez-Borraz, T. L., González-Sánchez, A., Bonilla-Blancas, W., Revah, S., & Noyola, A. (2017). Characterization of the biofiltration of methane emissions from municipal anaerobic effluents. *Process Biochemistry*, 63, 204–213. <https://doi.org/10.1016/j.procbio.2017.08.011>
- Gong, A., Wang, G., Qi, X., He, Y., Yang, X., Huang, X., & Liang, P. (2024). Energy recovery and saving in municipal wastewater treatment engineering practices. *Nature Sustainability*, 1–8. <https://doi.org/10.1038/s41893-024-01478-5>
- Goodess, C. M. (2013). How is the frequency, location and severity of extreme events likely to change up to 2060? *Environmental Science & Policy*, 27, S4–S14. <https://doi.org/10.1016/j.envsci.2012.04.001>
- Gosling, S. N., & Arnell, N. W. (2016). A global assessment of the impact of climate change on water scarcity. *Climatic Change*, 134(3), 371–385. <https://doi.org/10.1007/s10584-013-0853-x>
- Govindaraju, A., Good, N. M., Zytneck, A. M., & Martinez-Gomez, N. C. (2022). Employing methylotrophs for a green economy: One-carbon to fuel them all and through metabolism redesign them. *Current Opinion in Microbiology*, 67, 102145. <https://doi.org/10.1016/j.mib.2022.102145>
- Graham, D. W., Chaudhary, J. A., Hanson, R. S., & Arnold, R. G. (1993). Factors affecting competition between type I and type II methanotrophs in two-organism, continuous-flow reactors. *Microbial Ecology*, 25(1), 1–17. <https://doi.org/10.1007/BF00182126>
- Graham, S., & Marvin, S. (2001). CONSTRUCTING THE MODERN NETWORKED CITY, 1850–1960. In *Splintering Urbanism*. Routledge.
- Grilli, J. (2020). Macroecological laws describe variation and diversity in microbial communities. *Nature Communications*, 11(1), 4743. <https://doi.org/10.1038/s41467-020-18529-y>
- Guerrero Cruz, S., & Pijuan, M. (2022, January). Chapter 10 - Methanotrophic bacterial biorefineries: Resource recovery and GHG mitigation through the production of bacterial biopolymers. In A. An, V. Tyagi, M. Kumar, & Z. Cetecioglu (Eds.), *Clean Energy and Resource Recovery* (pp. 155–178). Elsevier. <https://doi.org/10.1016/B978-0-323-90178-9.00003-2>

- Guerrero-Cruz, S., Vaksmaa, A., Horn, M. A., Niemann, H., Pijuan, M., & Ho, A. (2021). Methanotrophs: Discoveries, Environmental Relevance, and a Perspective on Current and Future Applications. *Frontiers in Microbiology*, 12. <https://doi.org/10.3389/fmicb.2021.678057>
- Gupta, R., Lee, S., Lui, J., Sloan, W. T., & You, S. (2024). Carbon footprint assessment of water and wastewater treatment works in Scottish islands. *Journal of Cleaner Production*, 450, 141650. <https://doi.org/10.1016/j.jclepro.2024.141650>
- Haegeman, B., & Etienne, R. S. (2008). Relaxing the zero-sum assumption in neutral biodiversity theory. *Journal of Theoretical Biology*, 252(2), 288–294. <https://doi.org/10.1016/j.jtbi.2008.01.023>
- Haegeman, B., & Etienne, R. S. (2009). Neutral Models with Generalised Speciation. *Bulletin of Mathematical Biology*, 71(6), 1507–1519. <https://doi.org/10.1007/s11538-009-9411-0>
- Haegeman, B., & Etienne, R. S. (2010). Self-consistent approach for neutral community models with speciation. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, 81(3 Pt 1), 031911. <https://doi.org/10.1103/PhysRevE.81.031911>
- Hanasaki, N., Fujimori, S., Yamamoto, T., Yoshikawa, S., Masaki, Y., Hijioka, Y., Kainuma, M., Kanamori, Y., Masui, T., Takahashi, K., & Kanae, S. (2013). A global water scarcity assessment under Shared Socio-economic Pathways – Part 2: Water availability and scarcity. *Hydrology and Earth System Sciences*, 17(7), 2393–2413. <https://doi.org/10.5194/hess-17-2393-2013>
- Hanson, R. S., & Hanson, T. E. (1996). Methanotrophic bacteria. *Microbiological Reviews*, 60(2), 439–471. <https://doi.org/10.1128/mr.60.2.439-471.1996>
- Hardy, G. H. (1908). Mendelian Proportions in a Mixed Population. *Science*, 28(706), 49–50. <https://doi.org/10.1126/science.28.706.49>
- Harris, D. A. (2009, June). *Bioenergetics at a Glance: An Illustrated Introduction*. John Wiley & Sons.
- Harte, J. (2004). The Value of Null Theories in Ecology. *Ecology*, 85(7), 1792–1794. Retrieved February 2, 2025, from <https://www.jstor.org/stable/3450343>
- Haynes, C. A., & Gonzalez, R. (2014). Rethinking biological activation of methane and conversion to liquid fuels. *Nature Chemical Biology*, 10(5), 331–339. <https://doi.org/10.1038/nchembio.1509>
- He, C., Liu, Z., Wu, J., Pan, X., Fang, Z., Li, J., & Bryan, B. A. (2021). Future global urban water scarcity and potential solutions. *Nature Communications*, 12(1), 4667. <https://doi.org/10.1038/s41467-021-25026-3>
- He, L., Zhang, Y., Song, D., Ou, Z., Xie, Z., Yang, S., Guan, W., Dong, C., & Zhang, Y. (2022). Influence of Pretreatment System on Inorganic Suspended Solids for Influent in Wastewater Treatment Plant. *Journal of Environmental and Public Health*, 2022(1), 2768883. <https://doi.org/10.1155/2022/2768883>
- Hedrick, P. (2009, July). I.15 Population Genetics and Ecology. In S. A. Levin, S. R. Carpenter, H. C. J. Godfray, A. P. Kinzig, M. Loreau, J. B. Losos, B. Walker, & D. S. Wilcove (Eds.), *The Princeton Guide to Ecology* (pp. 109–116). Princeton University Press. <https://doi.org/10.1515/9781400833023.109>
- Heidrich, E. S., Curtis, T. P., & Dolfing, J. (2011). Determination of the Internal Chemical Energy of Wastewater. *Environmental Science & Technology*, 45(2), 827–832. <https://doi.org/10.1021/es103058w>

- Heidrich, E. S., Dolfing, J., Scott, K., Edwards, S. R., Jones, C., & Curtis, T. P. (2013). Production of hydrogen from domestic wastewater in a pilot-scale microbial electrolysis cell. *Applied Microbiology and Biotechnology*, 97(15), 6979–6989. <https://doi.org/10.1007/s00253-012-4456-7>
- Heijnen, J. J., & Van Dijken, J. P. (1992). In search of a thermodynamic description of biomass yields for the chemotrophic growth of microorganisms. *Biotechnology and Bioengineering*, 39(8), 833–858. <https://doi.org/10.1002/bit.260390806>
- Heine, W., Sekoulov, I., Burkhardt, H., Bergen, L., & Behrendt, J. (2002). Early warning-system for operation-failures in biological stages of WWTPs by on-line image analysis. *Water Science and Technology*, 46(4-5), 117–124. <https://doi.org/10.2166/wst.2002.0566>
- Henckel, T., Roslev, P., & Conrad, R. (2000). Effects of O₂ and CH₄ on presence and activity of the indigenous methanotrophic community in rice field soil. *Environmental Microbiology*, 2(6), 666–679. <https://doi.org/10.1046/j.1462-2920.2000.00149.x>
- Heys, C., Cheaib, B., Busetti, A., Kazlauskaitė, R., Maier, L., Sloan, W. T., Ijaz, U. Z., Kaufmann, J., McGinnity, P., & Llewellyn, M. S. (2020). Neutral Processes Dominate Microbial Community Assembly in Atlantic Salmon, *Salmo salar*. *Applied and Environmental Microbiology*, 86(8), e02283–19. <https://doi.org/10.1128/AEM.02283-19>
- Hoinkis, J., Deowan, S. A., Panten, V., Figoli, A., Huang, R. R., & Drioli, E. (2012). Membrane Bioreactor (MBR) Technology – a Promising Approach for Industrial Water Reuse. *Procedia Engineering*, 33, 234–241. <https://doi.org/10.1016/j.proeng.2012.01.1199>
- Hristov, J., Barreiro-Hurle, J., Salputra, G., Blanco, M., & Witzke, P. (2021). Reuse of treated water in European agriculture: Potential to address water scarcity under climate change. *Agricultural Water Management*, 251, 106872. <https://doi.org/10.1016/j.agwat.2021.106872>
- Hu, X.-S., He, F., & Hubbell, S. P. (2006). Neutral theory in macroecology and population genetics. *Oikos*, 113(3), 548–556. <https://doi.org/10.1111/j.2006.0030-1299.14837.x>
- Hubbell, S. P. (1979). Tree Dispersion, Abundance, and Diversity in a Tropical Dry Forest. *Science*, 203(4387), 1299–1309. <https://doi.org/10.1126/science.203.4387.1299>
- Hubbell, S. P. (1995). Towards a theory of biodiversity and biogeography on continuous landscapes. In *Preparing for Global Change: A Midwestern Perspective* (pp. 173–201). SPB Academic Pub.
- Hubbell, S. P. (1997). A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs*, 16(1), S9–S21. <https://doi.org/10.1007/s003380050237>
- Hubbell, S. P. (2001, January). *The Unified Neutral Theory of Biodiversity and Biogeography* (Vol. 32).
- Hubbell, S. P. (2003). Modes of speciation and the lifespans of species under neutrality: A response to the comment of Robert E. Ricklefs. *Oikos*, 100(1), 193–199. <https://doi.org/10.1034/j.1600-0706.2003.12450.x>
- Hubbell, S. P. (2005). Neutral Theory in Community Ecology and the Hypothesis of Functional Equivalence. *Functional Ecology*, 19(1), 166–172. Retrieved October 2, 2024, from <https://www.jstor.org/stable/3599285>

- Hubbell, S. P., & Foster, R. B. (1986, January). Biology, chance and history and the structure of tropical rain forest tree communities. In *Community Ecology* (pp. 314–329). Harper; Row.
- Huete, A., de los Cobos-Vasconcelos, D., Gómez-Borraz, T., Morgan-Sagastume, J. M., & Noyola, A. (2018). Control of dissolved CH₄ in a municipal UASB reactor effluent by means of a desorption – Biofiltration arrangement. *Journal of Environmental Management*, 216, 383–391. <https://doi.org/10.1016/j.jenvman.2017.06.061>
- Hughes, J., Cowper-Heays, K., Olesson, E., Bell, R., & Stroombergen, A. (2021). Impacts and implications of climate change on wastewater systems: A New Zealand perspective. *Climate Risk Management*, 31, 100262. <https://doi.org/10.1016/j.crm.2020.100262>
- Hughes, R. (1986). Theories and Models of Species Abundance. *American Naturalist - AMER NATURALIST*, 128. <https://doi.org/10.1086/284611>
- Hutton, G., Rodriguez, U.-P., Winara, A., Anh, N. V., Phyrum, K., Chuan, L., Blackett, I., & Weitz, A. (2013). Economic efficiency of sanitation interventions in Southeast Asia. *Journal of Water, Sanitation and Hygiene for Development*, 4(1), 23–36. <https://doi.org/10.2166/washdev.2013.158>
- Hyde-Smith, L., Zhan, Z., Roelich, K., Mdee, A., & Evans, B. (2022). Climate Change Impacts on Urban Sanitation: A Systematic Review and Failure Mode Analysis. *Environmental Science & Technology*, 56(9), 5306–5321. <https://doi.org/10.1021/acs.est.1c07424>
- Innis, G. S., & Haefner, J. W. (1980). A neutral model of community organization. *Journal of Theoretical Biology*, 87(3), 529–558. [https://doi.org/10.1016/0022-5193\(80\)90234-9](https://doi.org/10.1016/0022-5193(80)90234-9)
- Intergovernmental Panel on Climate Change (IPCC). (2023a). Changing State of the Climate System. In *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 287–422). Cambridge University Press. <https://doi.org/10.1017/9781009157896.004>
- Intergovernmental Panel on Climate Change (IPCC). (2023b). The Earth's Energy Budget, Climate Feedbacks and Climate Sensitivity. In *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 923–1054). Cambridge University Press. <https://doi.org/10.1017/9781009157896.009>
- Intergovernmental Panel on Climate Change (IPCC). (2023c). Summary for Policymakers. In *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 1–34). Retrieved December 9, 2024, from <https://www.ipcc.ch/report/ar6/syr/>
- Islam, T., Jensen, S., Reigstad, L. J., Larsen, Ø., & Birkeland, N.-K. (2008). Methane oxidation at 55°C and pH 2 by a thermoacidophilic bacterium belonging to the Verrucomicrobia phylum. *Proceedings of the National Academy of Sciences*, 105(1), 300–304. <https://doi.org/10.1073/pnas.0704162105>
- Jabot, F., & Chave, J. (2009). Inferring the parameters of the neutral theory of biodiversity using phylogenetic information and implications for tropical forests. *Ecology Letters*, 12(3), 239–248. <https://doi.org/10.1111/j.1461-0248.2008.01280.x>

- Jansen, M. (1989). Water Supply and Sewage Disposal at Mohenjo-Daro. *World Archaeology*, 21(2), 177–192. Retrieved December 4, 2024, from <https://www.jstor.org/stable/124907>
- Jarvis, P. G., & McNaughton, K. G. (1986). Stomatal Control of Transpiration: Scaling Up from Leaf to Region (A. MacFadyen & E. D. Ford, Eds.). *Advances in Ecological Research*, 15, 1–49. [https://doi.org/10.1016/S0065-2504\(08\)60119-1](https://doi.org/10.1016/S0065-2504(08)60119-1)
- Jenicek, P., Bartacek, J., Kutil, J., Zabranska, J., & Dohanyos, M. (2012). Potentials and limits of anaerobic digestion of sewage sludge: Energy self-sufficient municipal wastewater treatment plant? *Water Science and Technology*, 66(6), 1277–1281. <https://doi.org/10.2166/wst.2012.317>
- Jiang, L., & O'Neill, B. C. (2017). Global urbanization projections for the Shared Socioeconomic Pathways. *Global Environmental Change*, 42, 193–199. <https://doi.org/10.1016/j.gloenvcha.2015.03.008>
- Joe Cyril, H. A., & Pawar, H. S. (2024). Sludge Conditioning, Activation, and Engineering. In *Application of Sewage Sludge in Industrial Wastewater Treatment* (pp. 1–20). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119857396.ch1>
- Johnson, A. C., Jin, X., Nakada, N., & Sumpter, J. P. (2020). Learning from the past and considering the future of chemicals in the environment. *Science*, 367(6476), 384–387. <https://doi.org/10.1126/science.aay6637>
- Jones, E. R., van Vliet, M. T. H., Qadir, M., & Bierkens, M. F. P. (2021). Country-level and gridded estimates of wastewater production, collection, treatment and reuse. *Earth System Science Data*, 13(2), 237–254. <https://doi.org/10.5194/essd-13-237-2021>
- Jørgensen, D. (2008). Cooperative Sanitation: Managing Streets and Gutters in Late Medieval England and Scandinavia. *Technology and Culture*, 49(3), 547–567. Retrieved December 4, 2024, from <https://www.jstor.org/stable/40061427>
- Kalyuzhanaya, M. G., Yang, S., Matsen, J. B., Konopka, M., Green-Saxena, A., Clubb, J., Sadilek, M., Orphan, V. J., & Beck, D. (2013). Global Molecular Analyses of Methane Metabolism in Methanotrophic Alphaproteobacterium, *Methylosinus trichosporium* OB3b. Part II. Metabolomics and ¹³C-Labeling Study. *Frontiers in Microbiology*, 4. <https://doi.org/10.3389/fmicb.2013.00070>
- Kalyuzhnaya, M. G., Yang, S., Rozova, O. N., Smalley, N. E., Clubb, J., Lamb, A., Gowda, G. A. N., Raftery, D., Fu, Y., Bringel, F., Vuilleumier, S., Beck, D. a. C., Trotsenko, Y. A., Khmelenina, V. N., & Lidstrom, M. E. (2013). Highly efficient methane biocatalysis revealed in a methanotrophic bacterium. *Nature Communications*, 4(1), 2785. <https://doi.org/10.1038/ncomms3785>
- Kalyuzhnaya, M. G., Puri, A. W., & Lidstrom, M. E. (2015). Metabolic engineering in methanotrophic bacteria. *Metabolic Engineering*, 29, 142–152. <https://doi.org/10.1016/j.ymben.2015.03.010>
- Kalyuzhny, M., Kadmon, R., & Shnerb, N. M. (2015). A neutral theory with environmental stochasticity explains static and dynamic properties of ecological communities. *Ecology Letters*, 18(6), 572–580. <https://doi.org/10.1111/ele.12439>
- Kc, S., & Lutz, W. (2017). The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100. *Global Environmental Change*, 42, 181–192. <https://doi.org/10.1016/j.gloenvcha.2014.06.004>

- Kedem, O., & Caplan, S. R. (1965). Degree of coupling and its relation to efficiency of energy conversion. *Transactions of the Faraday Society*, 61(0), 1897–1911. <https://doi.org/10.1039/TF9656101897>
- Kehrein, P., Loosdrecht, M. v., Osseweijer, P., Garfí, M., Dewulf, J., & Posada, J. (2020). A critical review of resource recovery from municipal wastewater treatment plants – market supply potentials, technologies and bottlenecks. *Environmental Science: Water Research & Technology*, 6(4), 877–910. <https://doi.org/10.1039/C9EW00905A>
- Keller, J. (2023). Why are decentralised urban water solutions still rare given all the claimed benefits, and how could that be changed? *Water Research X*, 19, 100180. <https://doi.org/10.1016/j.wroa.2023.100180>
- Khasawneh, O. F. S., & Palaniandy, P. (2021). Occurrence and removal of pharmaceuticals in wastewater treatment plants. *Process Safety and Environmental Protection*, 150, 532–556. <https://doi.org/10.1016/j.psep.2021.04.045>
- Khmelenina, V. N., Colin Murrell, J., Smith, T. J., & Trotsenko, Y. A. (2018). Physiology and Biochemistry of the Aerobic Methanotrophs. In F. Rojo (Ed.), *Aerobic Utilization of Hydrocarbons, Oils and Lipids* (pp. 1–25). Springer International Publishing. https://doi.org/10.1007/978-3-319-39782-5_4-1
- Kim, T., Behrens, S., & LaPara, T. M. (2021). Direct Evidence for Deterministic Assembly of Bacterial Communities in Full-Scale Municipal Wastewater Treatment Facilities. *Applied and Environmental Microbiology*, 87(20), e01086–21. <https://doi.org/10.1128/AEM.01086-21>
- Kim, Y. M., Cho, H. U., Lee, D. S., Park, D., & Park, J. M. (2011). Influence of operational parameters on nitrogen removal efficiency and microbial communities in a full-scale activated sludge process. *Water Research*, 45(17), 5785–5795. <https://doi.org/10.1016/j.watres.2011.08.063>
- Kimura, M. (1968). Evolutionary Rate at the Molecular Level. *Nature*, 217(5129), 624–626. <https://doi.org/10.1038/217624a0>
- Kimura, M. (1971). *Theoretical aspects of population genetics*. Princeton, N.J. : Princeton University Press. Retrieved September 30, 2024, from <http://archive.org/details/theoreticalaspec0000kimu>
- Kimura, M. (1985, January). *The Neutral Theory of Molecular Evolution*. Cambridge University Press.
- King, J. L., & Jukes, T. H. (1969). Non-Darwinian Evolution. *Science*, 164(3881), 788–798. <https://doi.org/10.1126/science.164.3881.788>
- Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J. G., Dlugokencky, E. J., Bergamaschi, P., Bergmann, D., Blake, D. R., Bruhwiler, L., Cameron-Smith, P., Castaldi, S., Chevallier, F., Feng, L., Fraser, A., Heimann, M., Hodson, E. L., Houweling, S., Josse, B., ... Zeng, G. (2013). Three decades of global methane sources and sinks. *Nature Geoscience*, 6(10), 813–823. <https://doi.org/10.1038/ngeo1955>
- Kleerebezem, R., & van Loosdrecht, M. C. M. (2010). A Generalized Method for Thermodynamic State Analysis of Environmental Systems. *Critical Reviews in Environmental Science and Technology*, 40(1), 1–54. <https://doi.org/10.1080/10643380802000974>
- Koh, S.-C., Bowman, J. P., & Sayler, G. S. (1993). Soluble Methane Monooxygenase Production and Trichloroethylene Degradation by a Type I Methanotroph, *Methylobomonas methanica* 68-1. *Applied and Environmental Microbiology*, 59(4), 960–967. <https://doi.org/10.1128/aem.59.4.960-967.1993>

- Konopka, A., Lindemann, S., & Fredrickson, J. (2015). Dynamics in microbial communities: Unraveling mechanisms to identify principles. *The ISME Journal*, 9(7), 1488–1495. <https://doi.org/10.1038/ismej.2014.251>
- Koo, C. W., & Rosenzweig, A. C. (2021). Biochemistry of aerobic biological methane oxidation. *Chemical Society Reviews*, 50(5), 3424–3436. <https://doi.org/10.1039/D0CS01291B>
- Koottatep, T., Pussayanavin, T., & Polprasert, C. (2020). Nouveau design solar septic tank: Reinvented toilet technology for sanitation 4.0. *Environmental Technology & Innovation*, 19, 100933. <https://doi.org/10.1016/j.eti.2020.100933>
- Kopp, M. (2010). Speciation and the neutral theory of biodiversity. *BioEssays*, 32(7), 564–570. <https://doi.org/10.1002/bies.201000023>
- Kretschmer, F., Neugebauer, G., Kollmann, R., Eder, M., Zach, F., Zottl, A., Narodoslawsky, M., Stoeglehner, G., & Ertl, T. (2015). Resource recovery from wastewater in Austria: Wastewater treatment plants as regional energy cells. *Journal of Water Reuse and Desalination*, 6(3), 421–429. <https://doi.org/10.2166/wrd.2015.119>
- Kundu, K., Bergmann, I., Klocke, M., Sharma, S., & Sreekrishnan, T. R. (2014). Impact of abrupt temperature increase on the performance of an anaerobic hybrid bioreactor and its intrinsic microbial community. *Bioresource Technology*, 168, 72–79. <https://doi.org/10.1016/j.biortech.2014.01.093>
- La, H., Hettiaratchi, J. P. A., Achari, G., & Dunfield, P. F. (2018). Biofiltration of methane. *Bioresource Technology*, 268, 759–772. <https://doi.org/10.1016/j.biortech.2018.07.043>
- Latimer, A. M., Silander, J. A., & Cowling, R. M. (2005). Neutral Ecological Theory Reveals Isolation and Rapid Speciation in a Biodiversity Hot Spot. *Science*, 309(5741), 1722–1725. <https://doi.org/10.1126/science.1115576>
- Lawton, J. H. (1999). Are There General Laws in Ecology? *Oikos*, 84(2), 177–192. <https://doi.org/10.2307/3546712>
- Leadbetter, E. R., & Foster, J. W. (1958). Studies on some methane-utilizing bacteria. *Archiv für Mikrobiologie*, 30(1), 91–118. <https://doi.org/10.1007/BF00509229>
- Leak, D. J., & Dalton, H. (1986a). Growth yields of methanotrophs 1. Effect of copper on the energetics of methane oxidation. *Applied Microbiology and Biotechnology*, 23(6), 470–476. <https://doi.org/10.1007/BF02346062>
- Leak, D. J., & Dalton, H. (1986b). Growth yields of methanotrophs 2. A theoretical analysis. *Applied Microbiology and Biotechnology*, 23(6), 477–481. <https://doi.org/10.1007/BF02346063>
- Lee, O. K., Nguyen, D. T. N., & Lee, E. Y. (2019). Metabolic Engineering of Methanotrophs for the Production of Chemicals and Fuels. In E. Y. Lee (Ed.), *Methanotrophs: Microbiology Fundamentals and Biotechnological Applications* (pp. 163–203). Springer International Publishing. https://doi.org/10.1007/978-3-030-23261-0_6
- Leibold, M. A., & McPeck, M. A. (2006). Coexistence of the Niche and Neutral Perspectives in Community Ecology. *Ecology*, 87(6), 1399–1410. Retrieved October 3, 2024, from <https://www.jstor.org/stable/20069089>

- Lemke, M., & DeSalle, R. (2023). The Next Generation of Microbial Ecology and Its Importance in Environmental Sustainability. *Microbial Ecology*, 85(3), 781–795. <https://doi.org/10.1007/s00248-023-02185-y>
- Levin, S. A. (1992). The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. *Ecology*, 73(6), 1943–1967. <https://doi.org/10.2307/1941447>
- Lewontin, R. C. (1974). *The genetic basis of evolutionary change*. New York, Columbia University Press. Retrieved September 25, 2024, from <http://archive.org/details/geneticbasisofev00lewo>
- Li, J., Li, X., Liu, H., Gao, L., Wang, W., Wang, Z., Zhou, T., & Wang, Q. (2023). Climate change impacts on wastewater infrastructure: A systematic review and typological adaptation strategy. *Water Research*, 242, 120282. <https://doi.org/10.1016/j.watres.2023.120282>
- Li, J., Liu, H., & Paul Chen, J. (2018). Microplastics in freshwater systems: A review on occurrence, environmental effects, and methods for microplastics detection. *Water Research*, 137, 362–374. <https://doi.org/10.1016/j.watres.2017.12.056>
- Li, X., Zhou, Y., Eom, J., Yu, S., & Asrar, G. R. (2019). Projecting Global Urban Area Growth Through 2100 Based on Historical Time Series Data and Future Shared Socioeconomic Pathways. *Earth's Future*, 7(4), 351–362. <https://doi.org/10.1029/2019EF001152>
- Liang, X., & Yue, X. (2021). Challenges facing the management of wastewater treatment systems in Chinese rural areas. *Water Science and Technology*, 84(6), 1518–1526. <https://doi.org/10.2166/wst.2021.332>
- Libralato, G., Volpi Ghirardini, A., & Avezzi, F. (2012). To centralise or to decentralise: An overview of the most recent trends in wastewater treatment management. *Journal of Environmental Management*, 94(1), 61–68. <https://doi.org/10.1016/j.jenvman.2011.07.010>
- Lin, Y., de Kreuk, M., van Loosdrecht, M. C. M., & Adin, A. (2010). Characterization of alginate-like exopolysaccharides isolated from aerobic granular sludge in pilot-plant. *Water Research*, 44(11), 3355–3364. <https://doi.org/10.1016/j.watres.2010.03.019>
- Liu, J.-S., Marison, I. W., & von Stockar, U. (1999). Anaerobic Calorimetry of the Growth of *Lactobacillus Helveticus* Using a Highly Sensitive Bio-RC1. *Journal of Thermal Analysis and Calorimetry*, 56(3), 1191–1195. <https://doi.org/10.1023/A:1010121531950>
- Liu, J.-S., Marison, I. W., & von Stockar, U. (2001). Microbial growth by a net heat up-take: A calorimetric and thermodynamic study on acetotrophic methanogenesis by *Methanosarcina barkeri*. *Biotechnology and Bioengineering*, 75(2), 170–180. <https://doi.org/10.1002/bit.1176>
- Liu, W., Zhang, J., Liu, H., Guo, X., Zhang, X., Yao, X., Cao, Z., & Zhang, T. (2021). A review of the removal of microplastics in global wastewater treatment plants: Characteristics and mechanisms. *Environment International*, 146, 106277. <https://doi.org/10.1016/j.envint.2020.106277>
- Liu, Z., Cichocki, N., Hübschmann, T., Süring, C., Ofiteru, I. D., Sloan, W. T., Grimm, V., & Müller, S. (2019). Neutral mechanisms and niche differentiation in steady-state insular microbial communities revealed by single cell analysis. *Environmental Microbiology*, 21(1), 164–181. <https://doi.org/10.1111/1462-2920.14437>

- Lloyd, P. J. (1967). American, German and British Antecedents to Pearl and Reed's Logistic Curve. *Population Studies*, 21(2), 99–108. <https://doi.org/10.2307/2172714>
- Locasale, J. W., & Cantley, L. C. (2011). Metabolic Flux and the Regulation of Mammalian Cell Growth. *Cell Metabolism*, 14(4), 443–451. <https://doi.org/10.1016/j.cmet.2011.07.014>
- Lofrano, G., & Brown, J. (2010). Wastewater management through the ages: A history of mankind. *Science of The Total Environment*, 408(22), 5254–5264. <https://doi.org/10.1016/j.scitotenv.2010.07.062>
- Logares, R., Tesson, S. V., Canbäck, B., Pontarp, M., Hedlund, K., & Rengefors, K. (2018). Contrasting prevalence of selection and drift in the community structuring of bacteria and microbial eukaryotes. *Environmental Microbiology*, 20(6), 2231–2240. <https://doi.org/10.1111/1462-2920.14265>
- López, J. C., Quijano, G., Souza, T. S. O., Estrada, J. M., Lebrero, R., & Muñoz, R. (2013). Biotechnologies for greenhouse gases (CH₄, N₂O, and CO₂) abatement: State of the art and challenges. *Applied Microbiology and Biotechnology*, 97(6), 2277–2303. <https://doi.org/10.1007/s00253-013-4734-z>
- Lopez-Vazquez, C. M., Oehmen, A., Hooijmans, C. M., Brdjanovic, D., Gijzen, H. J., Yuan, Z., & van Loosdrecht, M. C. M. (2009). Modeling the PAO–GAO competition: Effects of carbon source, pH and temperature. *Water Research*, 43(2), 450–462. <https://doi.org/10.1016/j.watres.2008.10.032>
- Lotka, A. J. (1920a). Analytical Note on Certain Rhythmic Relations in Organic Systems. *Proceedings of the National Academy of Sciences*, 6(7), 410–415. <https://doi.org/10.1073/pnas.6.7.410>
- Lotka, A. J. (1920b). Undamped oscillations derived from the law of mass action. *Journal of the American Chemical Society*, 42(8), 1595–1599. <https://doi.org/10.1021/ja01453a010>
- Lotka, A. J. (1925). *Elements of physical biology*. Williams & Wilkins company. Retrieved September 17, 2024, from <http://archive.org/details/elementsofphysic0000alfr>
- Lücker, S., Wagner, M., Maixner, F., Pelletier, E., Koch, H., Vacherie, B., Rattei, T., Damsté, J. S. S., Spieck, E., Le Paslier, D., & Daims, H. (2010). A Nitrospira metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria. *Proceedings of the National Academy of Sciences*, 107(30), 13479–13484. <https://doi.org/10.1073/pnas.1003860107>
- MacArthur, R. H., & Wilson, E. O. (1967). *The Theory of Island Biogeography* (REV - Revised). Princeton University Press. Retrieved September 23, 2024, from <https://www.jstor.org/stable/j.ctt19cc1t2>
- Magurran, A. E. (2007). Species abundance distributions over time. *Ecology Letters*, 10(5), 347–354. <https://doi.org/10.1111/j.1461-0248.2007.01024.x>
- Major, D. C., Omojola, A., Dettinger, M., Hanson, R. T., & Sanchez-Rodriguez, R. (2011). Climate change, water, and wastewater in cities. In C. Rosenzweig, S. Mehrotra, S. A. Hammer, & W. D. Solecki (Eds.), *Climate Change and Cities: First Assessment Report of the Urban Climate Change Research Network* (pp. 113–144). Cambridge University Press. <https://doi.org/10.1017/CBO9780511783142.011>
- Makisha, N., & Gulshin, I. (2024). Sustainable Modernization of Wastewater Treatment Plants. *Sustainability*, 16(20), 8757. <https://doi.org/10.3390/su16208757>
- Malthus, T. (1798). *An Essay on the Principle of Population* (1st).
- Mancosu, N., Snyder, R. L., Kyriakakis, G., & Spano, D. (2015). Water Scarcity and Future Challenges for Food Production. *Water*, 7(3), 975–992. <https://doi.org/10.3390/w7030975>

- Manga, M., & Muoghalu, C. C. (2024). Greenhouse gas emissions from on-site sanitation systems: A systematic review and meta-analysis of emission rates, formation pathways and influencing factors. *Journal of Environmental Management*, 357, 120736. <https://doi.org/10.1016/j.jenvman.2024.120736>
- Mannina, G., Gulhan, H., & Ni, B.-J. (2022). Water reuse from wastewater treatment: The transition towards circular economy in the water sector. *Bioresource Technology*, 363, 127951. <https://doi.org/10.1016/j.biortech.2022.127951>
- Marino, A., Bertolotti, S., Macrì, M., Bona, F., Bonetta, S., Falasco, E., Minella, M., & Fenoglio, S. (2024). Impact of wastewater treatment and drought in an Alpine region: A multidisciplinary case study. *Heliyon*, 10(15). <https://doi.org/10.1016/j.heliyon.2024.e35290>
- Martinez-Rabert, E., Sloan, W. T., & Gonzalez-Cabaleiro, R. (2023). Multiscale models driving hypothesis and theory-based research in microbial ecology. *Interface Focus*, 13(4), 20230008. <https://doi.org/10.1098/rsfs.2023.0008>
- Maskow, T., Kemp, R., Buchholz, F., Schubert, T., Kiesel, B., & Harms, H. (2010). What heat is telling us about microbial conversions in nature and technology: From chip- to megacalorimetry. *Microbial Biotechnology*, 3(3), 269–284. <https://doi.org/10.1111/j.1751-7915.2009.00121.x>
- Matthews, T. J., & Whittaker, R. J. (2014). Neutral theory and the species abundance distribution: Recent developments and prospects for unifying niche and neutral perspectives. *Ecology and Evolution*, 4(11), 2263–2277. <https://doi.org/10.1002/ece3.1092>
- McGill, B. J. (2003). A test of the unified neutral theory of biodiversity. *Nature*, 422(6934), 881–885. <https://doi.org/10.1038/nature01583>
- Meegoda, J. N., Li, B., Patel, K., & Wang, L. B. (2018). A Review of the Processes, Parameters, and Optimization of Anaerobic Digestion. *International Journal of Environmental Research and Public Health*, 15(10), 2224. <https://doi.org/10.3390/ijerph15102224>
- Mekonnen, M. M., & Hoekstra, A. Y. (2016). Four billion people facing severe water scarcity. *Science Advances*, 2(2). <https://doi.org/DOI:10.1126/sciadv.1500323>
- Metcalf & Eddy, I. (2003). *Wastewater engineering : Treatment and reuse*. McGraw-Hill. Retrieved January 5, 2025, from <http://archive.org/details/wastewaterengine0000unse>
- Mielczarek, A. T., Kragelund, C., Eriksen, P. S., & Nielsen, P. H. (2012). Population dynamics of filamentous bacteria in Danish wastewater treatment plants with nutrient removal. *Water Research*, 46(12), 3781–3795. <https://doi.org/10.1016/j.watres.2012.04.009>
- Millstein, R. L. (2013). Exploring the Status of Population Genetics: The Role of Ecology. *Biological Theory*, 7(4), 346–357. <https://doi.org/10.1007/s13752-012-0056-0>
- Mishra, S., Kumar, R., & Kumar, M. (2023). Use of treated sewage or wastewater as an irrigation water for agricultural purposes- Environmental, health, and economic impacts. *Total Environment Research Themes*, 6, 100051. <https://doi.org/10.1016/j.totert.2023.100051>
- Molins, S., Mayer, K. U., Scheutz, C., & Kjeldsen, P. (2008). Transport and Reaction Processes Affecting the Attenuation of Landfill Gas in Cover Soils. *Journal of Environmental Quality*, 37(2), 459–468. <https://doi.org/10.2134/jeq2007.0250>

- Monteil, G., Houweling, S., Dlugockenky, E. J., Maenhout, G., Vaughn, B. H., White, J. W. C., & Rockmann, T. (2011). Interpreting methane variations in the past two decades using measurements of CH₄ mixing ratio and isotopic composition. *Atmospheric Chemistry and Physics*, 11(17), 9141–9153. <https://doi.org/10.5194/acp-11-9141-2011>
- Moonkawin, J., Huynh, L. T., Schneider, M. Y., Fujii, S., Echigo, S., Nguyen, L. P. H., Hoang, T.-H. T., Huynh, H. T., & Harada, H. (2023). Challenges to Accurate Estimation of Methane Emission from Septic Tanks with Long Emptying Intervals. *Environmental Science & Technology*, 57(43), 16575–16584. <https://doi.org/10.1021/acs.est.3c05724>
- Motoshita, M., Itsubo, N., & Inaba, A. (2011). Development of impact factors on damage to health by infectious diseases caused by domestic water scarcity. *The International Journal of Life Cycle Assessment*, 16(1), 65–73. <https://doi.org/10.1007/s11367-010-0236-8>
- Moussavi, G., Kazembeigi, F., & Farzadkia, M. (2010). Performance of a pilot scale up-flow septic tank for on-site decentralized treatment of residential wastewater. *Process Safety and Environmental Protection*, 88(1), 47–52. <https://doi.org/10.1016/j.psep.2009.10.001>
- Muneepeerakul, R., Bertuzzo, E., Lynch, H. J., Fagan, W. F., Rinaldo, A., & Rodriguez-Iturbe, I. (2008). Neutral metacommunity models predict fish diversity patterns in Mississippi–Missouri basin. *Nature*, 453(7192), 220–222. <https://doi.org/10.1038/nature06813>
- Murray, A., & Drechsel, P. (2011). Why do some wastewater treatment facilities work when the majority fail? Case study from the sanitation sector in Ghana. *Waterlines*, 30(2), 135–149. Retrieved January 7, 2025, from <https://www.jstor.org/stable/24686701>
- Murrell, J. C., & Dalton, H. (1983). Nitrogen Fixation in Obligate Methanotrophs. *Microbiology*, 129(11), 3481–3486. <https://doi.org/10.1099/00221287-129-11-3481>
- Murshid, S., Antonysamy, A., Dhakshinamoorthy, G., Jayaseelan, A., & Pugazhendhi, A. (2023). A review on biofilm-based reactors for wastewater treatment: Recent advancements in biofilm carriers, kinetics, reactors, economics, and future perspectives. *Science of The Total Environment*, 892, 164796. <https://doi.org/10.1016/j.scitotenv.2023.164796>
- Muter, O. (2023). Current Trends in Bioaugmentation Tools for Bioremediation: A Critical Review of Advances and Knowledge Gaps. *Microorganisms*, 11(3), 710. <https://doi.org/10.3390/microorganisms11030710>
- Mutuku, C., Gazdag, Z., & Melegh, S. (2022). Occurrence of antibiotics and bacterial resistance genes in wastewater: Resistance mechanisms and antimicrobial resistance control approaches. *World Journal of Microbiology and Biotechnology*, 38(9), 152. <https://doi.org/10.1007/s11274-022-03334-0>
- Naidoo, S., & Olaniran, A. O. (2014). Treated Wastewater Effluent as a Source of Microbial Pollution of Surface Water Resources. *International Journal of Environmental Research and Public Health*, 11(1), 249–270. <https://doi.org/10.3390/ijerph110100249>
- Nations, U. (2024). *The United Nations World Water Development Report 2024: Water for prosperity and peace*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000388948>
- Nazari, R., Vasiliadis, H., Karimi, M., Fahad, M. G. R., Simon, S., Zhang, T., Sun, Q., & Peters, R. (2022). Hydrodynamic Study of the Impact of Extreme Flooding Events on Wastewater Treatment Plants Con-

- sidering Total Water Level. *Natural Hazards Review*, 23(1), 04021056. [https://doi.org/10.1061/\(ASCE\)NH.1527-6996.0000531](https://doi.org/10.1061/(ASCE)NH.1527-6996.0000531)
- Nee, S. (2005). The Neutral Theory of Biodiversity: Do the Numbers Add up? *Functional Ecology*, 19(1), 173–176. Retrieved October 2, 2024, from <https://www.jstor.org/stable/3599286>
- Nei, M., Suzuki, Y., & Nozawa, M. (2010). The Neutral Theory of Molecular Evolution in the Genomic Era. *Annual Review of Genomics and Human Genetics*, 11(Volume 11, 2010), 265–289. <https://doi.org/10.1146/annurev-genom-082908-150129>
- NESC. (2020). *Onsite Wastewater Installation Assessment: Phase 2 (2015 – 2018) Report* (tech. rep.). West Virginia University. Morgantown. Retrieved December 9, 2024, from <https://actat.wvu.edu/topics-of-interest/assessment-of-u-s-onsite-system-installations-2015-through-2018>
- Ng, C. W. W., Feng, S., & Liu, H. W. (2015). A fully coupled model for water–gas–heat reactive transport with methane oxidation in landfill covers. *Science of The Total Environment*, 508, 307–319. <https://doi.org/10.1016/j.scitotenv.2014.11.037>
- Nguyen, A. D., Kim, D., & Lee, E. Y. (2020). Unlocking the biosynthesis of sesquiterpenoids from methane via the methylerythritol phosphate pathway in methanotrophic bacteria, using -humulene as a model compound. *Metabolic Engineering*, 61, 69–78. <https://doi.org/10.1016/j.ymben.2020.04.011>
- Nguyen, L. N., Nguyen, A. Q., & Nghiem, L. D. (2019). Microbial Community in Anaerobic Digestion System: Progression in Microbial Ecology. In X.-T. Bui, C. Chiemchaisri, T. Fujioka, & S. Varjani (Eds.), *Water and Wastewater Treatment Technologies* (pp. 331–355). Springer. https://doi.org/10.1007/978-981-13-3259-3_15
- Nicholls, D. G. (2013, May). *Bioenergetics*. Academic Press.
- Nisbet, R., & Gurney, W. (1982). *Modelling Fluctuating Populations*.
- Notley, S. R., Mitchell, D., & Taylor, N. A. S. (2023). A century of exercise physiology: Concepts that ignited the study of human thermoregulation. Part 2: Physiological measurements. *European Journal of Applied Physiology*, 123(12), 2587–2685. <https://doi.org/10.1007/s00421-023-05284-3>
- Noyola, A., Capdeville, B., & Roques, H. (1988). Anaerobic treatment of domestic sewage with a rotating stationary fixed-film reactor. *Water Research*, 22(12), 1585–1592. [https://doi.org/10.1016/0043-1354\(88\)90172-8](https://doi.org/10.1016/0043-1354(88)90172-8)
- Nunziata, S. O., & Weisrock, D. W. (2018). Estimation of contemporary effective population size and population declines using RAD sequence data. *Heredity*, 120(3), 196–207. <https://doi.org/10.1038/s41437-017-0037-y>
- O'Connor, F. M., Boucher, O., Gedney, N., Jones, C. D., Folberth, G. A., Coppel, R., Friedlingstein, P., Collins, W. J., Chappellaz, J., Ridley, J., & Johnson, C. E. (2010). Possible role of wetlands, permafrost, and methane hydrates in the methane cycle under future climate change: A review. *Reviews of Geophysics*, 48(4). <https://doi.org/10.1029/2010RG000326>
- O'Dwyer, J. P., & Green, J. L. (2010). Field theory for biogeography: A spatially explicit model for predicting patterns of biodiversity. *Ecology Letters*, 13(1), 87–95. <https://doi.org/10.1111/j.1461-0248.2009.01404.x>

- Ofițeru, I. D., Lunn, M., Curtis, T. P., Wells, G. F., Criddle, C. S., Francis, C. A., & Sloan, W. T. (2010). Combined niche and neutral effects in a microbial wastewater treatment community. *Proceedings of the National Academy of Sciences*, 107(35), 15345–15350. <https://doi.org/10.1073/pnas.1000604107>
- Op den Camp, H. J. M., Islam, T., Stott, M. B., Harhangi, H. R., Hynes, A., Schouten, S., Jetten, M. S. M., Birkeland, N.-K., Pol, A., & Dunfield, P. F. (2009). Environmental, genomic and taxonomic perspectives on methanotrophic Verrucomicrobia. *Environmental Microbiology Reports*, 1(5), 293–306. <https://doi.org/10.1111/j.1758-2229.2009.00022.x>
- Orla-Jensen, S. (1909). Die Hauptlinien des natrlichen Bakterien-systems. Zentralbl. Bakteriöl. Parasitenkd. *Zentrabl Bakteriöl Parasitenk Infektionskr*, 22, 305–346.
- Park, D., & Lee, J. (2013). Biological conversion of methane to methanol. *Korean Journal of Chemical Engineering*, 30(5), 977–987. <https://doi.org/10.1007/s11814-013-0060-5>
- Park, H. D., Wells, G. F., Bae, H., Criddle, C. S., & Francis, C. A. (2006). Occurrence of Ammonia-Oxidizing Archaea in Wastewater Treatment Plant Bioreactors. *Applied and Environmental Microbiology*, 72(8), 5643–5647. <https://doi.org/10.1128/AEM.00402-06>
- Park, S., Hanna, L., Taylor, R. T., & Droege, M. W. (1991). Batch cultivation of *Methylosinus trichosporium* OB3b. I: Production of soluble methane monooxygenase. *Biotechnology and Bioengineering*, 38(4), 423–433. <https://doi.org/10.1002/bit.260380412>
- Parry, H. R., & Bithell, M. (2012). Large Scale Agent-Based Modelling: A Review and Guidelines for Model Scaling. In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.), *Agent-Based Models of Geographical Systems* (pp. 271–308). Springer Netherlands. https://doi.org/10.1007/978-90-481-8927-4_14
- Pasalari, H., Gholami, M., Rezaee, A., Esrafil, A., & Farzadkia, M. (2021). Perspectives on microbial community in anaerobic digestion with emphasis on environmental parameters: A systematic review. *Chemosphere*, 270, 128618. <https://doi.org/10.1016/j.chemosphere.2020.128618>
- Pasciucco, F., Pecorini, I., & Iannelli, R. (2022). Planning the centralization level in wastewater collection and treatment: A review of assessment methods. *Journal of Cleaner Production*, 375, 134092. <https://doi.org/10.1016/j.jclepro.2022.134092>
- Pearl, R., & Reed, L. J. (1920). On the Rate of Growth of the Population of the United States since 1790 and Its Mathematical Representation. *Proceedings of the National Academy of Sciences of the United States of America*, 6(6), 275–288. <https://doi.org/10.1073/pnas.6.6.275>
- Pearl, R., & Slobodkin, L. (1976). The Growth of Populations. *The Quarterly Review of Biology*, 51, 6–24. Retrieved September 16, 2024, from <https://www.jstor.org/stable/2823040>
- Philippi, L. S., da Costa, R. H. R., & Sezerino, P. H. (1999). Domestic effluent treatment through integrated system of septic tank and root zone. *Water Science and Technology*, 40(3), 125–131. [https://doi.org/10.1016/S0273-1223\(99\)00455-2](https://doi.org/10.1016/S0273-1223(99)00455-2)
- Pierrat, É., Laurent, A., Dorber, M., Rygaard, M., Verones, F., & Hauschild, M. (2023). Advancing water footprint assessments: Combining the impacts of water pollution and scarcity. *Science of The Total Environment*, 870, 161910. <https://doi.org/10.1016/j.scitotenv.2023.161910>

- Pittam, D. A., & Pilcher, G. (1972). Measurements of heats of combustion by flame calorimetry. Part 8.—Methane, ethane, propane, n-butane and 2-methylpropane. *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases*, 68(0), 2224–2229. <https://doi.org/10.1039/F19726802224>
- Pol, A., Heijmans, K., Harhangi, H. R., Tedesco, D., Jetten, M. S. M., & Op den Camp, H. J. M. (2007). Methanotrophy below pH 1 by a new Verrucomicrobia species. *Nature*, 450(7171), 874–878. <https://doi.org/10.1038/nature06222>
- Prather, M. J., & Holmes, C. D. (2017). Overexplaining or underexplaining methane's role in climate change. *Proceedings of the National Academy of Sciences*, 114(21), 5324–5326. <https://doi.org/10.1073/pnas.1704884114>
- Pratt, C., & Tate, K. (2018). Mitigating Methane: Emerging Technologies To Combat Climate Change's Second Leading Contributor. *Environmental Science & Technology*, 52(11), 6084–6097. <https://doi.org/10.1021/acs.est.7b04711>
- Preisner, M. (2020). Surface Water Pollution by Untreated Municipal Wastewater Discharge Due to a Sewer Failure. *Environmental Processes*, 7(3), 767–780. <https://doi.org/10.1007/s40710-020-00452-5>
- Priyadarsini, A., Singh, R., Barbora, L., Maitra, S. S., & Moholkar, V. S. (2023). Methanotroph detection and bioconversion of methane to methanol by enriched microbial consortium from rice field soil. *Biore-source Technology Reports*, 22, 101410. <https://doi.org/10.1016/j.biteb.2023.101410>
- Prosser, J. I. (2020). Putting science back into microbial ecology: A question of approach. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1798), 20190240. <https://doi.org/10.1098/rstb.2019.0240>
- Prosser, J. I., Bohannan, B. J. M., Curtis, T. P., Ellis, R. J., Firestone, M. K., Freckleton, R. P., Green, J. L., Green, L. E., Killham, K., Lennon, J. J., Osborn, A. M., Solan, M., van der Gast, C. J., & Young, J. P. W. (2007). The role of ecological theory in microbial ecology. *Nature Reviews Microbiology*, 5(5), 384–392. <https://doi.org/10.1038/nrmicro1643>
- Purkhold, U., Pommerening-Röser, A., Juretschko, S., Schmid, M. C., Koops, H.-P., & Wagner, M. (2000). Phylogeny of All Recognized Species of Ammonia Oxidizers Based on Comparative 16S rRNA and amoA Sequence Analysis: Implications for Molecular Diversity Surveys. *Applied and Environmental Microbiology*, 66(12), 5368–5382. <https://doi.org/10.1128/AEM.66.12.5368-5382.2000>
- Purves, D. W., & Turnbull, L. A. (2010). Different but equal: The implausible assumption at the heart of neutral theory. *Journal of Animal Ecology*, 79(6), 1215–1225. <https://doi.org/10.1111/j.1365-2656.2010.01738.x>
- Pussayanavin, T., Koottatep, T., Eamrat, R., & Polprasert, C. (2015). Enhanced sludge reduction in septic tanks by increasing temperature. *Journal of Environmental Science and Health. Part A, Toxic/Hazardous Substances & Environmental Engineering*, 50(1), 81–89. <https://doi.org/10.1080/10934529.2015.964633>
- Puyol, D., Batstone, D. J., Hülsen, T., Astals, S., Peces, M., & Krömer, J. O. (2017). Resource Recovery from Wastewater by Biological Technologies: Opportunities, Challenges, and Prospects. *Frontiers in Microbiology*, 7. <https://doi.org/10.3389/fmicb.2016.02106>

- Qadir, M., Drechsel, P., Jiménez Cisneros, B., Kim, Y., Pramanik, A., Mehta, P., & Olaniyan, O. (2020). Global and regional potential of wastewater as a water, nutrient and energy source. *Natural Resources Forum*, 44(1), 40–51. <https://doi.org/10.1111/1477-8947.12187>
- Qiu, G., Law, Y., Zuniga-Montanez, R., Deng, X., Lu, Y., Roy, S., Thi, S. S., Hoon, H. Y., Nguyen, T. Q. N., Eganathan, K., Liu, X., Nielsen, P. H., Williams, R. B. H., & Wuertz, S. (2022). Global warming readiness: Feasibility of enhanced biological phosphorus removal at 35 °C. *Water Research*, 216, 118301. <https://doi.org/10.1016/j.watres.2022.118301>
- Rafiee, A., Khalilpour, K. R., Prest, J., & Skryabin, I. (2021). Biogas as an energy vector. *Biomass and Bioenergy*, 144, 105935. <https://doi.org/10.1016/j.biombioe.2020.105935>
- Ramalingam, D., Lingireddy, S., & Ormsbee, L. E. (2012). History of Water Distribution Network Analysis: Over 100 Years of Progress, 55–67. [https://doi.org/10.1061/40650\(2003\)6](https://doi.org/10.1061/40650(2003)6)
- Regueiro, L., Veiga, P., Figueroa, M., Alonso-Gutierrez, J., Stams, A. J. M., Lema, J. M., & Carballa, M. (2012). Relationship between microbial activity and microbial community structure in six full-scale anaerobic digesters. *Microbiological Research*, 167(10), 581–589. <https://doi.org/10.1016/j.micres.2012.06.002>
- Riaz, A., Zahedi, G., & Klemeš, J. J. (2013). A review of cleaner production methods for the manufacture of methanol. *Journal of Cleaner Production*, 57, 19–37. <https://doi.org/10.1016/j.jclepro.2013.06.017>
- Richards, S., Paterson, E., Withers, P. J. A., & Stutter, M. (2016). Septic tank discharges as multi-pollutant hotspots in catchments. *Science of The Total Environment*, 542, 854–863. <https://doi.org/10.1016/j.scitotenv.2015.10.160>
- Ricklefs, R. E. (2003). A Comment on Hubbell's Zero-Sum Ecological Drift Model. *Oikos*, 100(1), 185–192. Retrieved October 2, 2024, from <https://www.jstor.org/stable/3548273>
- Ricklefs, R. E. (2006). The Unified Neutral Theory of Biodiversity: Do the Numbers Add Up? *Ecology*, 87(6), 1424–1431. [https://doi.org/10.1890/0012-9658\(2006\)87\[1424:TUNTOB\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[1424:TUNTOB]2.0.CO;2)
- Rigby, M., Montzka, S. A., Prinn, R. G., White, J. W. C., Young, D., O'Doherty, S., Lunt, M. F., Ganesan, A. L., Manning, A. J., Simmonds, P. G., Salameh, P. K., Harth, C. M., Mühle, J., Weiss, R. F., Fraser, P. J., Steele, L. P., Krummel, P. B., McCulloch, A., & Park, S. (2017). Role of atmospheric oxidation in recent methane growth. *Proceedings of the National Academy of Sciences*, 114(21), 5373–5377. <https://doi.org/10.1073/pnas.1616426114>
- Risch, E., Boutin, C., & Roux, P. (2021). Applying life cycle assessment to assess the environmental performance of decentralised versus centralised wastewater systems. *Water Research*, 196, 116991. <https://doi.org/10.1016/j.watres.2021.116991>
- Rittmann, B. E., Hausner, M., Löffler, F., Love, N. G., Muyzer, G., Okabe, S., Oerther, D. B., Peccia, J., Raskin, L., & Wagner, M. (2006). A Vista for Microbial Ecology and Environmental Biotechnology. *Environmental Science & Technology*, 40(4), 1096–1103. <https://doi.org/10.1021/es062631k>
- Rivière, D., Desvignes, V., Pelletier, E., Chaussonnerie, S., Guermazi, S., Weissenbach, J., Li, T., Camacho, P., & Sghir, A. (2009). Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. *The ISME Journal*, 3(6), 700–714. <https://doi.org/10.1038/ismej.2009.2>

- Robb, J., Cessford, C., Dittmar, J., Inskip, S. A., & Mitchell, P. D. (2021). The greatest health problem of the Middle Ages? Estimating the burden of disease in medieval England. *International Journal of Paleopathology*, 34, 101–112. <https://doi.org/10.1016/j.ijpp.2021.06.011>
- Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoin, H. K., Landerer, F. W., & Lo, M.-H. (2018). Emerging trends in global freshwater availability. *Nature*, 557(7707), 651–659. <https://doi.org/10.1038/s41586-018-0123-1>
- Roels, J. A. (1983). *Energetics and Kinetics in Biotechnology*. Elsevier Biomedical Press.
- Rogers, D. K. (2018). Water Culture in Roman Society. *Brill Research Perspectives in Ancient History*, 1(1), 1–118. <https://doi.org/10.1163/25425374-12340001>
- Rosindell, J., & Cornell, S. J. (2007). Species–area relationships from a spatially explicit neutral model in an infinite landscape. *Ecology Letters*, 10(7), 586–595. <https://doi.org/10.1111/j.1461-0248.2007.01050.x>
- Rosindell, J., Cornell, S. J., Hubbell, S. P., & Etienne, R. S. (2010). Protracted speciation revitalizes the neutral theory of biodiversity. *Ecology Letters*, 13(6), 716–727. <https://doi.org/10.1111/j.1461-0248.2010.01463.x>
- Rosindell, J., Hubbell, S. P., & Etienne, R. S. (2011). *The Unified Neutral Theory of Biodiversity and Biogeography* at Age Ten. *Trends in Ecology & Evolution*, 26(7), 340–348. <https://doi.org/10.1016/j.tree.2011.03.024>
- Rosindell, J., Hubbell, S. P., He, F., Harmon, L. J., & Etienne, R. S. (2012). The case for ecological neutral theory. *Trends in Ecology & Evolution*, 27(4), 203–208. <https://doi.org/10.1016/j.tree.2012.01.004>
- Roslev, P., & King, G. M. (1995). Aerobic and anaerobic starvation metabolism in methanotrophic bacteria. *Applied and Environmental Microbiology*, 61(4), 1563–1570. <https://doi.org/10.1128/aem.61.4.1563-1570.1995>
- Roughgarden, J. (1983). Competition and Theory in Community Ecology. *The American Naturalist*, 122(5), 583–601. <https://doi.org/10.1086/284160>
- Royal Commission on Sewage Disposal. (1915). *Final report of the commissioners appointed to inquire and report what methods of treating and disposing of sewage (including any liquid from any factory or manufacturing process) may properly be adopted. General summary of conclusions and recommendations*. H. M. Stationery Office. Retrieved January 4, 2025, from <http://archive.org/details/cu31924003641929>
- Ruiz-Ruiz, P., Gómez-Borraz, T. L., Revah, S., & Morales, M. (2020). Methanotroph-microalgae co-culture for greenhouse gas mitigation: Effect of initial biomass ratio and methane concentration. *Chemosphere*, 259, 127418. <https://doi.org/10.1016/j.chemosphere.2020.127418>
- Sahoo, K. K., Goswami, G., & Das, D. (2021). Biotransformation of Methane and Carbon Dioxide Into High-Value Products by Methanotrophs: Current State of Art and Future Prospects. *Frontiers in Microbiology*, 12. <https://doi.org/10.3389/fmicb.2021.636486>
- Sahoo, K. K., Katari, J. K., & Das, D. (2023). Recent advances in methanol production from methanotrophs. *World Journal of Microbiology and Biotechnology*, 39(12), 360. <https://doi.org/10.1007/s11274-023-03813-y>

- Sale, P. F. (1977). Maintenance of High Diversity in Coral Reef Fish Communities. *The American Naturalist*, 111(978), 337–359. Retrieved September 30, 2024, from <https://www.jstor.org/stable/2460067>
- Salgot, M., & Folch, M. (2018). Wastewater treatment and water reuse. *Current Opinion in Environmental Science & Health*, 2, 64–74. <https://doi.org/10.1016/j.coesh.2018.03.005>
- Sander, R. (2023). Compilation of Henry's law constants (version 5.0.0) for water as solvent. *Atmospheric Chemistry and Physics*, 23(19), 10901–12440. <https://doi.org/10.5194/acp-23-10901-2023>
- Sanz, J. L., & Köchling, T. (2007). Molecular biology techniques used in wastewater treatment: An overview. *Process Biochemistry*, 42(2), 119–133. <https://doi.org/10.1016/j.procbio.2006.10.003>
- Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., Patra, P. K., Ciais, P., Arora, V. K., Bastviken, D., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., ... Zhuang, Q. (2020). The Global Methane Budget 2000–2017. *Earth System Science Data*, 12(3), 1561–1623. <https://doi.org/10.5194/essd-12-1561-2020>
- Sauvé, S., & Desrosiers, M. (2014). A review of what is an emerging contaminant. *Chemistry Central Journal*, 8(1), 15. <https://doi.org/10.1186/1752-153X-8-15>
- Sawyer, C. N. (1965). Milestones in the Development of the Activated Sludge Process. *Journal (Water Pollution Control Federation)*, 37(2), 151–162. Retrieved January 6, 2025, from <https://www.jstor.org/stable/25035231>
- Schill, N., & von Stockar, U. (1995). Thermodynamic analysis of *Methanobacterium thermoautotrophicum*. *Thermochimica Acta*, 251, 71–77. [https://doi.org/10.1016/0040-6031\(94\)02093-4](https://doi.org/10.1016/0040-6031(94)02093-4)
- Schlosser, C. A., Strzepek, K., Gao, X., Fant, C., Blanc, É., Paltsev, S., Jacoby, H., Reilly, J., & Gueneau, A. (2014). The future of global water stress: An integrated assessment. *Earth's Future*, 2(8), 341–361. <https://doi.org/10.1002/2014EF000238>
- Schneider, E. D., & Kay, J. J. (1994). Life as a manifestation of the second law of thermodynamics. *Mathematical and Computer Modelling*, 19(6), 25–48. [https://doi.org/10.1016/0895-7177\(94\)90188-0](https://doi.org/10.1016/0895-7177(94)90188-0)
- Schrödinger, E. (1992). *What is Life?: With Mind and Matter and Autobiographical Sketches*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139644129>
- Schwarzenbach, R. P., Escher, B. I., Fenner, K., Hofstetter, T. B., Johnson, C. A., von Gunten, U., & Wehrli, B. (2006). The Challenge of Micropollutants in Aquatic Systems. *Science*, 313(5790), 1072–1077. <https://doi.org/10.1126/science.1127291>
- Scudo, F. M. (1971). Vito Volterra and theoretical ecology. *Theoretical Population Biology*, 2(1), 1–23. [https://doi.org/10.1016/0040-5809\(71\)90002-5](https://doi.org/10.1016/0040-5809(71)90002-5)
- Scudo, F. M., & Ziegler, J. R. (1978). *The Golden Age of Theoretical Ecology: 1923–1940* (S. Levin, Ed.; Vol. 22). Springer. <https://doi.org/10.1007/978-3-642-50151-7>
- Seeder, H. (1999). The history of German waste water treatment. 2(5), 51–56. Retrieved December 4, 2024, from <https://silo.tips/download/the-history-of-german-waste-water-treatment>
- Seto, M., Noguchi, K., & Cappellen, P. V. (2019). Potential for Aerobic Methanotrophic Metabolism on Mars. *Astrobiology*, 19(10), 1187–1195. <https://doi.org/10.1089/ast.2018.1943>

- Seuntjens, D., Han, M., Kerckhof, F.-M., Boon, N., Al-Omari, A., Takacs, I., Meerburg, F., De Mulder, C., Wett, B., Bott, C., Murthy, S., Carvajal Arroyo, J. M., De Clippeleir, H., & Vlaeminck, S. E. (2018). Pinpointing wastewater and process parameters controlling the AOB to NOB activity ratio in sewage treatment plants. *Water Research*, 138, 37–46. <https://doi.org/10.1016/j.watres.2017.11.044>
- Sheets, J. P., Ge, X., Li, Y.-F., Yu, Z., & Li, Y. (2016). Biological conversion of biogas to methanol using methanotrophs isolated from solid-state anaerobic digestate. *Bioresource Technology*, 201, 50–57. <https://doi.org/10.1016/j.biortech.2015.11.035>
- Shen, Y., Linville, J. L., Urgun-Demirtas, M., Mintz, M. M., & Snyder, S. W. (2015). An overview of biogas production and utilization at full-scale wastewater treatment plants (WWTPs) in the United States: Challenges and opportunities towards energy-neutral WWTPs. *Renewable and Sustainable Energy Reviews*, 50, 346–362. <https://doi.org/10.1016/j.rser.2015.04.129>
- Shifrin, N. S. (2005). Pollution Management in the Twentieth Century. *Journal of Environmental Engineering*, 131(5), 676–691. [https://doi.org/10.1061/\(ASCE\)0733-9372\(2005\)131:5\(676\)](https://doi.org/10.1061/(ASCE)0733-9372(2005)131:5(676))
- Simberloff, D. (1978). Using Island Biogeographic Distributions to Determine if Colonization is Stochastic. *The American Naturalist*, 112(986), 713–726. <https://doi.org/10.1086/283313>
- Simpson, G. G. (1964). Organisms and Molecules in Evolution. *Science*, 146(3651), 1535–1538. <https://doi.org/10.1126/science.146.3651.1535>
- Simpson, I. A., Guttman, E. B., Cluett, J., & Shepherd, A. (2006). Characterizing anthropic sediments in north European Neolithic settlements: An assessment from Skara Brae, Orkney. *Geoarchaeology*, 21(3), 221–235. <https://doi.org/10.1002/gea.20101>
- Singh, A., Srivastava, A., Saidulu, D., & Gupta, A. K. (2022). Advancements of sequencing batch reactor for industrial wastewater treatment: Major focus on modifications, critical operational parameters, and future perspectives. *Journal of Environmental Management*, 317, 115305. <https://doi.org/10.1016/j.jenvman.2022.115305>
- Singh, V., & Dey, S. (2024). Biological and Microbiological Characteristics of Activated Sewage Sludge. In *Application of Sewage Sludge in Industrial Wastewater Treatment* (pp. 87–106). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119857396.ch5>
- Sloan, W. T., & Gómez-Borraz, T. L. (2023). Engineering biology in the face of uncertainty. *Interface Focus*, 13(4), 20230001. <https://doi.org/10.1098/rsfs.2023.0001>
- Sloan, W. T., Lunn, M., Woodcock, S., Head, I. M., Nee, S., & Curtis, T. P. (2006). Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental Microbiology*, 8(4), 732–740. <https://doi.org/10.1111/j.1462-2920.2005.00956.x>
- Sloan, W. T., Nnaji, C. F., Lunn, M., Curtis, T. P., Colloms, S. D., Couto, J. M., Pinto, A. J., Connelly, S., & Rosser, S. J. (2021). Drift dynamics in microbial communities and the effective community size. *Environmental Microbiology*, 23(5), 2473–2483. <https://doi.org/10.1111/1462-2920.15453>
- Sloan, W. T., Woodcock, S., Lunn, M., Head, I. M., & Curtis, T. P. (2007). Modeling Taxa-Abundance Distributions in Microbial Communities using Environmental Sequence Data. *Microbial Ecology*, 53(3), 443–455. <https://doi.org/10.1007/s00248-006-9141-x>

- Smith, T. J., Trotsenko, Y. A., & Murrell, J. C. (2010). Physiology and Biochemistry of the Aerobic Methane Oxidizing Bacteria. In K. N. Timmis (Ed.), *Handbook of Hydrocarbon and Lipid Microbiology* (pp. 765–779). Springer. https://doi.org/10.1007/978-3-540-77587-4_58
- Sohnen, N. (1906). Über bakterien, welche methan als kohlenstoffnahrung und energiequelle gebrauchen. *Zentralbl Bakteriell Parasitenk Infektionskr*, 15, 513–517.
- Somlai, C., Knappe, J., & Gill, L. (2019). Spatial and temporal variation of CO₂ and CH₄ emissions from a septic tank soakaway. *Science of The Total Environment*, 679, 185–195. <https://doi.org/10.1016/j.scitotenv.2019.04.449>
- Soni, B. K., Conrad, J., Kelley, R. L., & Srivastava, V. J. (1998). Effect of temperature and pressure on growth and methane utilization by several methanotrophic cultures. *Applied Biochemistry and Biotechnology*, 70(1), 729–738. <https://doi.org/10.1007/BF02920184>
- Sorenson, C., Ding, L., Zamalloa, C., Arsenault, A., Debeni Devi, N., & Hu, B. (2023). Illuminated septic tank – A microalgae assisted onsite wastewater treatment system. *Journal of Environmental Chemical Engineering*, 11(5), 110768. <https://doi.org/10.1016/j.jece.2023.110768>
- Souza, C. L., Chernicharo, C. A. L., & Aquino, S. F. (2011). Quantification of dissolved methane in UASB reactors treating domestic wastewater under different operating conditions. *Water Science and Technology*, 64(11), 2259–2264. <https://doi.org/10.2166/wst.2011.695>
- Stanley, S. H., Prior, S. D., Leak, D. J., & Dalton, H. (1983). Copper stress underlies the fundamental change in intracellular location of methane mono-oxygenase in methane-oxidizing organisms: Studies in batch and continuous cultures. *Biotechnology Letters*, 5(7), 487–492. <https://doi.org/10.1007/BF00132233>
- Stockar, U. v., & Wielen, L. A. M. v. d. (2013a). Biothermodynamics Of Live Cells: Energy Dissipation And Heat Generation In Cellular Cultures. In *Biothermodynamics*. EPFL Press.
- Stockar, U. v., & Wielen, L. A. M. v. d. (2013b). Live Cells As Open Non-Equilibrium Systems. In *Biothermodynamics*. EPFL Press.
- Stokes, C. J., & Archer, S. R. (2010). Niche differentiation and neutral theory: An integrated perspective on shrub assemblages in a parkland savanna. *Ecology*, 91(4), 1152–1162. <https://doi.org/10.1890/08-1105.1>
- Stone, K. A., He, Q. P., & Wang, J. (2019). Two Experimental Protocols for Accurate Measurement of Gas Component Uptake and Production Rates in Bioconversion Processes. *Scientific Reports*, 9(1), 5899. <https://doi.org/10.1038/s41598-019-42469-3>
- Stokal, M., Bai, Z., Franssen, W., Hofstra, N., Koelmans, A. A., Ludwig, F., Ma, L., van Puijenbroek, P., Spanier, J. E., Vermeulen, L. C., van Vliet, M. T. H., van Wijnen, J., & Kroeze, C. (2021). Urbanization: An increasing source of multiple pollutants to rivers in the 21st century. *npj Urban Sustainability*, 1(1), 1–13. <https://doi.org/10.1038/s42949-021-00026-w>
- Strong, P. J., Xie, S., & Clarke, W. P. (2015). Methane as a Resource: Can the Methanotrophs Add Value? *Environmental Science & Technology*, 49(7), 4001–4018. <https://doi.org/10.1021/es504242n>
- Stubbendieck, R. M., Vargas-Bautista, C., & Straight, P. D. (2016). Bacterial Communities: Interactions to Scale. *Frontiers in Microbiology*, 7. <https://doi.org/10.3389/fmicb.2016.01234>

- Stucki, J. W. (1980). The Optimal Efficiency and the Economic Degrees of Coupling of Oxidative Phosphorylation. *European Journal of Biochemistry*, 109(1), 269–283. <https://doi.org/10.1111/j.1432-1033.1980.tb04792.x>
- Su, C. X.-H., Low, L. W., Teng, T. T., & Wong, Y. S. (2016). Combination and hybridisation of treatments in dye wastewater treatment: A review. *Journal of Environmental Chemical Engineering*, 4(3), 3618–3631. <https://doi.org/10.1016/j.jece.2016.07.026>
- Sudhakaran, S., Maeng, S. K., & Amy, G. (2013). Hybridization of natural systems with advanced treatment processes for organic micropollutant removals: New concepts in multi-barrier treatment. *Chemosphere*, 92(6), 731–737. <https://doi.org/10.1016/j.chemosphere.2013.04.021>
- Sukmana, H., Bellahsen, N., Pantoja, F., & Hodur, C. (2021). Adsorption and coagulation in wastewater treatment – Review. *Progress in Agricultural Engineering Sciences*, 17(1), 49–68. <https://doi.org/10.1556/446.2021.00029>
- Sun, Y., Zhang, M., Duan, C., Cao, N., Jia, W., Zhao, Z., Ding, C., Huang, Y., & Wang, J. (2021). Contribution of stochastic processes to the microbial community assembly on field-collected microplastics. *Environmental Microbiology*, 23(11), 6707–6720. <https://doi.org/10.1111/1462-2920.15713>
- Sürmelihiindi, G., Passchier, C., Crow, J., Spötl, C., & Mertz-Kraus, R. (2021). Carbonates from the ancient world's longest aqueduct: A testament of Byzantine water management. *Geoarchaeology*, 36(4), 643–659. <https://doi.org/10.1002/gea.21853>
- Taheriyoun, M., & Moradinejad, S. (2014). Reliability analysis of a wastewater treatment plant using fault tree analysis and Monte Carlo simulation. *Environmental Monitoring and Assessment*, 187(1), 4186. <https://doi.org/10.1007/s10661-014-4186-7>
- Tan, J. N., Ratra, K., Singer, S. W., Simmons, B. A., Goswami, S., & Awasthi, D. (2024). Methane to bioproducts: Unraveling the potential of methanotrophs for biomanufacturing. *Current Opinion in Biotechnology*, 90, 103210. <https://doi.org/10.1016/j.copbio.2024.103210>
- Templeton, A. S., Chu, K.-H., Alvarez-Cohen, L., & Conrad, M. E. (2006). Variable carbon isotope fractionation expressed by aerobic CH₄-oxidizing bacteria. *Geochimica et Cosmochimica Acta*, 70(7), 1739–1752. <https://doi.org/10.1016/j.gca.2005.12.002>
- Thomsen, T. R., Kong, Y., & Nielsen, P. H. (2007). Ecophysiology of abundant denitrifying bacteria in activated sludge. *FEMS Microbiology Ecology*, 60(3), 370–382. <https://doi.org/10.1111/j.1574-6941.2007.00309.x>
- Tilman, D. (2004). Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences*, 101(30), 10854–10861. <https://doi.org/10.1073/pnas.0403458101>
- Tokeshi, M. (1993, January). Species Abundance Patterns and Community Structure. In M. Begon & A. H. Fitter (Eds.), *Advances in Ecological Research* (pp. 111–186, Vol. 24). Academic Press. [https://doi.org/10.1016/S0065-2504\(08\)60042-2](https://doi.org/10.1016/S0065-2504(08)60042-2)

- Tokeshi, M. (1997). Species Coexistence and Abundance: Patterns and Processes. In T. Abe, S. A. Levin, & M. Higashi (Eds.), *Biodiversity: An Ecological Perspective* (pp. 35–55). Springer. https://doi.org/10.1007/978-1-4612-1906-4_4
- Tomaras, J., Sahl, J. W., Siegrist, R. L., & Spear, J. R. (2009). Microbial Diversity of Septic Tank Effluent and a Soil Biomat. *Applied and Environmental Microbiology*, 75(10), 3348–3351. <https://doi.org/10.1128/AEM.00560-08>
- Torre, A., Vázquez-Rowe, I., Parodi, E., & Kahhat, R. (2021). Wastewater treatment decentralization: Is this the right direction for megacities in the Global South? *Science of The Total Environment*, 778, 146227. <https://doi.org/10.1016/j.scitotenv.2021.146227>
- Toze, S. (2006). Reuse of effluent water—benefits and risks. *Agricultural Water Management*, 80(1), 147–159. <https://doi.org/10.1016/j.agwat.2005.07.010>
- Ulrich, W., & Zalewski, M. (2007). Are ground beetles neutral? *Basic and Applied Ecology*, 8(5), 411–420. <https://doi.org/10.1016/j.baae.2006.08.002>
- United Nations. (2015). *Transforming our World: The 2030 Agenda for Sustainable Development* | Department of Economic and Social Affairs (tech. rep.). UN. <https://sdgs.un.org/publications/transforming-our-world-2030-agenda-sustainable-development-17981>
- United Nations. (2022a). The human right to a clean, healthy and sustainable environment resolution adopted by the General Assembly. Retrieved January 5, 2025, from <https://digitallibrary.un.org/record/3982659>
- United Nations. (2022b). *The United Nations World Water Development Report 2022: Groundwater: Making the invisible visible*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000380721>
- United Nations Environment Programme. (2023). *Wastewater - Turning Problem to Solution. A UNEP Rapid Response Assessment*.
- United Nations Human Settlements Programme (UN-Habitat). (2023). *The Global Report on Sanitation and Wastewater Management in Cities and Human Settlements*.
- Valdez, B., Schorr, M., Quintero, M., García, R., & Rosas, N. (2010). Effect of climate change on durability of engineering materials in hydraulic infrastructure: An overview. *Corrosion Engineering, Science and Technology*, 45(1), 34–41. <https://doi.org/10.1179/147842209X12559428167526>
- van Loosdrecht, M. C. M., & Brdjanovic, D. (2014). Anticipating the next century of wastewater treatment. *Science*, 344(6191), 1452–1453. <https://doi.org/10.1126/science.1255183>
- van Wezel, A. P., van den Hurk, F., Sjerps, R. M. A., Meijers, E. M., Roex, E. W. M., & ter Laak, T. L. (2018). Impact of industrial waste water treatment plants on Dutch surface waters and drinking water sources. *Science of The Total Environment*, 640–641, 1489–1499. <https://doi.org/10.1016/j.scitotenv.2018.05.325>
- Van Dijken, J. P., & Harder, W. (1975). Growth yields of microorganisms on methanol and methane. A theoretical study. *Biotechnology and Bioengineering*, 17(1), 15–30. <https://doi.org/10.1002/bit.260170103>
- Van Minh, H., & Hung, N. V. (2011). Economic Aspects of Sanitation in Developing Countries. *Environmental Health Insights*, 5, EHI.S8199. <https://doi.org/10.4137/EHI.S8199>

- Vanwonterghem, I., Jensen, P. D., Ho, D. P., Batstone, D. J., & Tyson, G. W. (2014). Linking microbial community structure, interactions and function in anaerobic digesters using new molecular techniques. *Current Opinion in Biotechnology*, 27, 55–64. <https://doi.org/10.1016/j.copbio.2013.11.004>
- Vellend, M. (2010). Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology*, 85(2), 183–206. <https://doi.org/10.1086/652373>
- Vellend, M., & Orrock, L. (2009, October). Ecological and Genetic Models of Diversity: Lessons across Disciplines. In J. B. Losos & R. E. Ricklefs (Eds.), *The Theory of Island Biogeography Revisited* (pp. 439–462). Princeton University Press. <https://doi.org/10.1515/9781400831920.439>
- Venkataraman, A., Bassis, C. M., Beck, J. M., Young, V. B., Curtis, J. L., Huffnagle, G. B., & Schmidt, T. M. (2015). Application of a Neutral Community Model To Assess Structuring of the Human Lung Microbiome. *mBio*, 6(1), 10.1128/mbio.02284–14. <https://doi.org/10.1128/mbio.02284-14>
- Verhulst, P. F. (1838). Notice sur la loi que la population poursuit dans son accroissement. *correspondence mathématique et physique*, 10, 113–121.
- Verma, S. (2023, July). *Water and Wastewater Engineering Technology*. CRC Press. <https://doi.org/10.1201/9781003347941>
- Verstraete, W., Van de Caveye, P., & Diamantis, V. (2009). Maximum use of resources present in domestic “used water”. *Bioresource Technology*, 100(23), 5537–5545. <https://doi.org/10.1016/j.biortech.2009.05.047>
- Vicuña, S., Redwood, M., Dettinger, M., Noyola, A., Ferguson, D. B., Guereca, L. P., Clark, C. M., Lulham, N., Jamwal, P., Wejs, A., Lall, U., Raschid, L., Omojola, A., & Major, D. C. (2018). Urban Water Systems. In C. Rosenzweig, P. Romero-Lankao, S. Mehrotra, S. Dhakal, S. Ali Ibrahim, & W. D. Solecki (Eds.), *Climate Change and Cities: Second Assessment Report of the Urban Climate Change Research Network* (pp. 519–552). Cambridge University Press. <https://doi.org/10.1017/9781316563878.021>
- Vieira, A. d. N., Kleineremanns, K., Martin, W. F., & Preiner, M. (2020). The ambivalent role of water at the origins of life. *FEBS Letters*, 594(17), 2717–2733. <https://doi.org/10.1002/1873-3468.13815>
- Volkov, I., Banavar, J. R., Hubbell, S. P., & Maritan, A. (2007). Patterns of relative species abundance in rainforests and coral reefs. *Nature*, 450(7166), 45–49. <https://doi.org/10.1038/nature06197>
- Volkov, I., Banavar, J. R., Maritan, A., & Hubbell, S. P. (2004). The stability of forest biodiversity. *Nature*, 427(6976), 696–696. <https://doi.org/10.1038/427696a>
- Volterra, V. (1926a). Fluctuations in the Abundance of a Species considered Mathematically. *Nature*, 118(2972), 558–560. <https://doi.org/10.1038/118558a0>
- Volterra, V. (1926b). *Variazioni e fluttuazioni del numero d'individui in specie animali conviventi*. Società anonima tipografica "Leonardo da Vinci".
- Volterra, V. (1927). *Variazioni e fluttuazioni del numero d'individui in specie animali conviventi*. Memorie del R. Comitato Talassografico Italiano.
- von Bertalanffy, L. (1950). The Theory of Open Systems in Physics and Biology. *Science*, 111(2872), 23–29. <https://doi.org/10.1126/science.111.2872.23>

- von Stockar, U., & Birou, B. (1989). The heat generated by yeast cultures with a mixed metabolism in the transition between respiration and fermentation. *Biotechnology and Bioengineering*, 34(1), 86–101. <https://doi.org/10.1002/bit.260340112>
- von Stockar, U., & Liu, J. -. (1999). Does microbial life always feed on negative entropy? Thermodynamic analysis of microbial growth. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1412(3), 191–211. [https://doi.org/10.1016/S0005-2728\(99\)00065-1](https://doi.org/10.1016/S0005-2728(99)00065-1)
- von Stockar, U., & Marison, I. W. (2005). The use of calorimetry in biotechnology. *Bioprocesses and Engineering*, 93–136. <https://doi.org/10.1007/BFb0009829>
- von Stockar, U., Maskow, T., Liu, J., Marison, I. W., & Patiño, R. (2006). Thermodynamics of microbial growth and metabolism: An analysis of the current situation. *Journal of Biotechnology*, 121(4), 517–533. <https://doi.org/10.1016/j.jbiotec.2005.08.012>
- von Stockar, U., Vojinović, V., Maskow, T., & Liu, J. (2008). Can microbial growth yield be estimated using simple thermodynamic analogies to technical processes? *Chemical Engineering and Processing: Process Intensification*, 47(6), 980–990. <https://doi.org/10.1016/j.cep.2007.02.016>
- Vorobev, A. V., Baani, M., Doronina, N. V., Brady, A. L., Liesack, W., Dunfield, P. F., & Dedysch, S. N. (2011). *Methyloferula stellata* gen. nov., sp. nov., an acidophilic, obligately methanotrophic bacterium that possesses only a soluble methane monooxygenase. *International Journal of Systematic and Evolutionary Microbiology*, 61(10), 2456–2463. <https://doi.org/10.1099/ijs.0.028118-0>
- Vörösmarty, C. J., Green, P., Salisbury, J., & Lammers, R. B. (2000). Global Water Resources: Vulnerability from Climate Change and Population Growth. *Science*, 289(5477), 284–288. <https://doi.org/10.1126/science.289.5477.284>
- Vuorinen, H., Juuti, P., & Katko, T. (2007). History of water and health from ancient civilizations to modern times. *Water Supply*, 7(1), 49–57. <https://doi.org/10.2166/ws.2007.006>
- Wadsö, L., & Li, X. (2008). A Simple Rate Law Experiment Using a Custom-Built Isothermal Heat Conduction Calorimeter. *Journal of Chemical Education*, 85(1), 112. <https://doi.org/10.1021/ed085p112>
- Wagner, M., Loy, A., Nogueira, R., Purkhold, U., Lee, N., & Daims, H. (2002). Microbial community composition and function in wastewater treatment plants. *Antonie van Leeuwenhoek*, 81(1), 665–680. <https://doi.org/10.1023/A:1020586312170>
- Wang, C., & Kuzyakov, Y. (2023). Energy use efficiency of soil microorganisms: Driven by carbon recycling and reduction. *Global Change Biology*, 29(22), 6170–6187. <https://doi.org/10.1111/gcb.16925>
- Wang, J., Santiago, E., & Caballero, A. (2016). Prediction and estimation of effective population size. *Heredity*, 117(4), 193–206. <https://doi.org/10.1038/hdy.2016.43>
- Wang, J., & Wang, S. (2016). Removal of pharmaceuticals and personal care products (PPCPs) from wastewater: A review. *Journal of Environmental Management*, 182, 620–640. <https://doi.org/10.1016/j.jenvman.2016.07.049>
- Wang, K., Abdalla, A. A., Khaleel, M. A., Hilal, N., & Khraisheh, M. K. (2017). Mechanical properties of water desalination and wastewater treatment membranes. *Desalination*, 401, 190–205. <https://doi.org/10.1016/j.desal.2016.06.032>

- Wang, L., Hibino, A., Suganuma, S., Ebihara, A., Iwamoto, S., Mitsui, R., Tani, A., Shimada, M., Hayakawa, T., & Nakagawa, T. (2020). Preference for particular lanthanide species and thermal stability of XoxFs in *Methylobacterium extorquens* strain AM1. *Enzyme and Microbial Technology*, 136, 109518. <https://doi.org/10.1016/j.enzmictec.2020.109518>
- Wang, M., Bodirsky, B. L., Rijneveld, R., Beier, F., Bak, M. P., Batool, M., Droppers, B., Popp, A., van Vliet, M. T. H., & Strokal, M. (2024). A triple increase in global river basins with water scarcity due to future pollution. *Nature Communications*, 15(1), 880. <https://doi.org/10.1038/s41467-024-44947-3>
- Wang, Y., Wang, D., Yang, Q., Zeng, G., & Li, X. (2017). Wastewater Opportunities for Denitrifying Anaerobic Methane Oxidation. *Trends in Biotechnology*, 35(9), 799–802. <https://doi.org/10.1016/j.tibtech.2017.02.010>
- Wang, Y., Wang, X., Li, M., Dong, J., Sun, C., & Chen, G. (2018). Removal of Pharmaceutical and Personal Care Products (PPCPs) from Municipal Waste Water with Integrated Membrane Systems, MBR-RO/NF. *International Journal of Environmental Research and Public Health*, 15(2), 269. <https://doi.org/10.3390/ijerph15020269>
- Ward, K., Crow, J., & Crapper, M. (2017). Water-supply infrastructure of Byzantine Constantinople. *Journal of Roman Archaeology*, 30, 175–195. <https://doi.org/10.1017/S1047759400074079>
- Wardle, T. (1893). *On sewage treatment and disposal : For cities, towns, villages, private dwellings, and public institutions*. Manchester : J. Heywood. Retrieved December 4, 2024, from <http://archive.org/details/b21952449>
- Warren, E., & Bekins, B. A. (2015). Relating subsurface temperature changes to microbial activity at a crude oil-contaminated site. *Journal of Contaminant Hydrology*, 182, 183–193. <https://doi.org/10.1016/j.jconhyd.2015.09.007>
- Webster, C. (1962). The Sewers of Mohenjo-Daro. *Journal (Water Pollution Control Federation)*, 34(2), 116–123. Retrieved December 4, 2024, from <https://www.jstor.org/stable/25034575>
- Weinberg, W. (1908). *Über den Nachweis der Vererbung beim Menschen*. Retrieved January 20, 2025, from <http://archive.org/details/b30613000>
- Weirich, S. R., Silverstein, J., & Rajagopalan, B. (2015). Resilience of Secondary Wastewater Treatment Plants: Prior Performance Is Predictive of Future Process Failure and Recovery Time. *Environmental Engineering Science*, 32(3), 222–231. <https://doi.org/10.1089/ees.2014.0406>
- Weiss, P. (1962). From Cell to Molecule. In P. Weiss (Ed.), *The Molecular Control of Cellular Activity* (pp. 1–73). McGraw-Hill.
- Westall, F., & Brack, A. (2018). The Importance of Water for Life. *Space Science Reviews*, 214(2), 50. <https://doi.org/10.1007/s11214-018-0476-7>
- Whitman, W. B., Coleman, D. C., & Wiebe, W. J. (1998). Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences*, 95(12), 6578–6583. <https://doi.org/10.1073/pnas.95.12.6578>
- Whittenbury, R., Phillips, K. C., & Wilkinson, J. F. (1970). Enrichment, Isolation and Some Properties of Methane-utilizing Bacteria. *Microbiology*, 61(2), 205–218. <https://doi.org/10.1099/00221287-61-2-205>

- WHO. (2019). *Safer water, better health. 2019 update*. World Health Organization. <https://www.who.int/publications-detail-redirect/9789241516891>
- Widder, S., Allen, R. J., Pfeiffer, T., Curtis, T. P., Wiuf, C., Sloan, W. T., Cordero, O. X., Brown, S. P., Momeni, B., Shou, W., Kettle, H., Flint, H. J., Haas, A. F., Laroche, B., Kreft, J.-U., Rainey, P. B., Freilich, S., Schuster, S., Milferstedt, K., ... Soyer, O. S. (2016). Challenges in microbial ecology: Building predictive understanding of community function and dynamics. *The ISME Journal*, 10(11), 2557–2568. <https://doi.org/10.1038/ismej.2016.45>
- Wieland, F.-G., Hauber, A. L., Rosenblatt, M., Tönsing, C., & Timmer, J. (2021). On structural and practical identifiability. *Current Opinion in Systems Biology*, 25, 60–69. <https://doi.org/10.1016/j.coisb.2021.03.005>
- Wilderer, P., & Schreff, D. (2000). Decentralized and centralized wastewater management: A challenge for technology developers. *Water Science and Technology*, 41(1), 1–8. <https://doi.org/10.2166/wst.2000.0001>
- Withers, P. J. A., Jarvie, H. P., & Stoate, C. (2011). Quantifying the impact of septic tank systems on eutrophication risk in rural headwaters. *Environment International*, 37(3), 644–653. <https://doi.org/10.1016/j.envint.2011.01.002>
- Withers, P. J. A., May, L., Jarvie, H. P., Jordan, P., Doody, D., Foy, R. H., Bechmann, M., Cooksley, S., Dils, R., & Deal, N. (2012). Nutrient emissions to water from septic tank systems in rural catchments: Uncertainties and implications for policy. *Environmental Science & Policy*, 24, 71–82. <https://doi.org/10.1016/j.envsci.2012.07.023>
- Withers, P. J., Jordan, P., May, L., Jarvie, H. P., & Deal, N. E. (2014). Do septic tank systems pose a hidden threat to water quality? *Frontiers in Ecology and the Environment*, 12(2), 123–130. <https://doi.org/10.1890/130131>
- Wood, D. A., Nwaoha, C., & Towler, B. F. (2012). Gas-to-liquids (GTL): A review of an industry offering several routes for monetizing natural gas. *Journal of Natural Gas Science and Engineering*, 9, 196–208. <https://doi.org/10.1016/j.jngse.2012.07.001>
- Woodcock, S., van der Gast, C. J., Bell, T., Lunn, M., Curtis, T. P., Head, I. M., & Sloan, W. T. (2007). Neutral assembly of bacterial communities. *FEMS Microbiology Ecology*, 62(2), 171–180. <https://doi.org/10.1111/j.1574-6941.2007.00379.x>
- World Economic Forum. (2024). *Global Risks Report 2024* (tech. rep.). World Economic Forum. Retrieved April 11, 2024, from <https://www.weforum.org/publications/global-risks-report-2024/in-full/>
- Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, 16(2), 97–159. <https://doi.org/10.1093/genetics/16.2.97>
- Wu, L., Ning, D., Zhang, B., Li, Y., Zhang, P., Shan, X., Zhang, Q., Brown, M. R., Li, Z., Van Nostrand, J. D., Ling, F., Xiao, N., Zhang, Y., Vierheilig, J., Wells, G. F., Yang, Y., Deng, Y., Tu, Q., Wang, A., ... Zhou, J. (2019). Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Nature Microbiology*, 4(7), 1183–1195. <https://doi.org/10.1038/s41564-019-0426-5>

- WWAP. (2018). *The United Nations world water development report 2018: Nature-based solutions for water*. <https://unesdoc.unesco.org/ark:/48223/pf0000261424>
- Xin, J.-y., Cui, J.-r., Niu, J.-z., Hua, S.-f., Xia, C.-g., Li, S.-b., & Zhu, L.-m. (2004). Production of methanol from methane by methanotrophic bacteria. *Biocatalysis and Biotransformation*, 22(3), 225–229. <https://doi.org/10.1080/10242420412331283305>
- Xin, J.-y., Zhang, Y.-x., Zhang, S., Xia, C.-g., & Li, S.-b. (2007). Methanol production from CO₂ by resting cells of the methanotrophic bacterium *Methylosinus trichosporium* IMV 3011. *Journal of Basic Microbiology*, 47(5), 426–435. <https://doi.org/10.1002/jobm.200710313>
- Xu, A., Wu, Y.-H., Chen, Z., Wu, G., Wu, Q., Ling, F., Huang, W. E., & Hu, H.-Y. (2020). Towards the new era of wastewater treatment of China: Development history, current status, and future directions. *Water Cycle*, 1, 80–87. <https://doi.org/10.1016/j.watcyc.2020.06.004>
- Yu, Q., Olesen, S. W., Duvallet, C., & Grad, Y. H. (2024). Assessment of sewer connectivity in the United States and its implications for equity in wastewater-based epidemiology. *PLOS Global Public Health*, 4(4), e0003039. <https://doi.org/10.1371/journal.pgph.0003039>
- Zhang, T., Ye, L., Tong, A. H. Y., Shao, M.-F., & Lok, S. (2011). Ammonia-oxidizing archaea and ammonia-oxidizing bacteria in six full-scale wastewater treatment bioreactors. *Applied Microbiology and Biotechnology*, 91(4), 1215–1225. <https://doi.org/10.1007/s00253-011-3408-y>
- Zhou, J., Liu, W., Deng, Y., Jiang, Y.-H., Xue, K., He, Z., Van Nostrand, J. D., Wu, L., Yang, Y., & Wang, A. (2013). Stochastic Assembly Leads to Alternative Communities with Distinct Functions in a Bioreactor Microbial Community. *mBio*, 4(2), 10.1128/mbio.00584–12. <https://doi.org/10.1128/mbio.00584-12>
- Zhou, J., & Ning, D. (2017). Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiology and Molecular Biology Reviews*, 81(4), 10.1128/mmbr.00002–17. <https://doi.org/10.1128/mmbr.00002-17>